# Methods for interpreting small-angle scattering scattering data from membrane proteins

Lise Arleth, Structural Biophysics, Niels Bohr institute, University of Copenhagen, Denmark

Membrane proteins are an important target for structural investigations by a broad range of experimental and computational methods. During the previous decade, several carrier systems for the reconstitution of membrane proteins have been developed and refined. Also, the experimental facilities and accesories for SAXS and SANS studies of such systems have improved tremendously. The interpretation of small-angle scattering data from membrane protein samples however remains a challenge that still to a large extent requires custom-fitted solutions on a case to case basis. This is in contrast to the more general methods that have been developed for the SAS analysis of soluble proteins and which has enabled a large user community to access these. In my talk, I will start by giving an overview of some of the initial work that has been done with respect to analysing SAS data from simple cases of membrane proteins by my own group and others. I will use this as a basis for discussing how the analysis of small-angle scattering data from membrane proteins and other samples could benefit from more systematic integration of data from complementary experimental or computational sources and give an example of how we are presently developing a combined SAXS/SANS/NMR and MD approach to extract more detailed information about nanodisc samples. The perspectives for generalising this approach to membrane protein samples will be discussed.

# Disentangling the Structural Complexity of Disordered Biomolecular Assemblies

# Pau Bernadó Centre de Biochimie Structurale, Montpellier, France

In the last decade Intrinsically Disordered Proteins (IDPs) have emerged as key actors in the vast majority of cellular events related with signalling, regulation and homeostasis. The inherent plasticity of IDPs, which lack of permanent secondary and tertiary structure, facilitates specific partner recognition. Importantly, interaction events are normally performed by short regions that rigidify upon binding and induce low affinity interactions. The structural characterization of these complexes is highly challenging due to the inherent flexibility of the non-interacting segments of the IDP and the thermodynamic equilibrium between bound and unbound species present in solution. The synergistic combination of SAS, Nuclear Magnetic Resonance (NMR) and advanced computational tools is the most adapted strategy to disentangle this complexity.

Examples of intrinsically disordered complexes involved in DNA repair and regulation of gene transcription will be presented. With these examples, new tools and strategies developed in the group to model disordered proteins and to study transient complexes will be described.

# SAS of membrane proteins: the data we have, what modelling would we need?

Cécile Breyton<sup>1</sup>, Anne Martel<sup>2</sup>, Christine Ebel<sup>1</sup>

<sup>1</sup> Institut de Biologie Structurale, Grenoble, France

<sup>2</sup> Institut Laue-Langevin, Grenoble, France

Proteins are polymers that fold in a well-defined 3D structure, which can be determined by X-ray diffraction to atomic resolution, given well-diffracting crystals are available. To allow their biological function however, a certain flexibility can be allowed: in the majority of cases, the protein oscillates between two extreme conformations.

I will present two examples of membrane protein systems that we study by SANS. Indeed, membrane proteins need to be handled in detergent to avoid precipitation. We have determined conditions in which bound detergent and free micelles can be completely matched, the SAS data being thus exclusively produced by the studied protein.

The first example is a bacterial protein, SpNox, composed of a unique polypeptide chain, folded in two independent domains, linked by a linker. The atomic structure of a homologous protein is available, but the SANS data shows a larger object. We suspect flexibility between the two domains.

The second example is another bacterial protein, BmrA. This protein is formed by the combination of two polypeptide chains, and the structure of homologues have been determined in different conformations depending on the type of ligand it is incubated with. Are those structures crystallisation artefacts or do they exist in solution? Are they artefacts due to protein manipulation in detergent, or do they also exist when the protein is embedded in a membrane? Also, what is the structure of the protein when it is not incubated with any ligands? Probably an equilibrium between the different conformations...

#### SAXS data-driven modeling of RNA 3D structures

Pritha Ghosh<sup>1</sup>, Michał J. Boniecki<sup>1</sup>, Chandran Nithin<sup>1</sup>, Gay Pauline Padilla-Meier<sup>2</sup>, Grzegorz Chojnowski<sup>3</sup>, Sean A. McKenna<sup>2</sup>, Trushar R. Patel<sup>4</sup>, Janusz M. Bujnicki<sup>1,5,\*</sup>

<sup>1</sup> Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, ul. Ks. Trojdena 4, PL-02-109 Warsaw, Poland

<sup>2</sup> Department of Chemistry, University of Manitoba, Winnipeg, MB, Canada

<sup>3</sup> European Molecular Biology Laboratory (EMBL) Hamburg Outstation, c/o DESY, Notkestrasse 85, Hamburg 22607, Germany

<sup>4</sup> Alberta RNA Research and Training Institute, Department of Chemistry & Biochemistry, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4, Canada

<sup>5</sup> Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, ul. Umultowska 89, PL-61-614 Poznan, Poland

\* presenting author, <u>iamb@genesilico.pl</u>

The majority of known RNAs exert their cellular functions in complexes with other molecules (proteins and/or small molecules). Many RNAs are known to or have been predicted to interact with small molecule ligands, which emphasizes their importance as potential therapeutic targets. However, in order to discern the molecular functions of such interactions, it is important to understand the three-dimensional (3D) structure of the RNA molecules, in the presence and the absence of ligands. Due to the difficulties associated with the experimental determination of high-resolution RNA structures, experimental data-aided computational modeling has become an important approach in generating high-quality theoretical models<sup>1</sup>. The inherent flexibility of RNA molecules allows it to sample a large conformational space. This hints at the fact that its 3D structure is best represented by an ensemble of atomistic structures rather than a single structural model.

The small angle X-ray scattering technique describes the distribution of electron density in a molecule and hence can be used to interpret the low-resolution envelope of a biomolecule. My group has developed a computational workflow for SAXS data-driven modeling of RNA 3D structures and their ensembles. The workflow involves exploration of the conformational space with the SimRNA program (using a coarsegrained representation and a statistical potential to generate physically realistic structures)<sup>2</sup> to generate large sets of plausible conformations of the target RNA structure. The large sets of generated decoys are scored against experimental SAXS data using CRYSOL<sup>3,4</sup>, followed by clustering, ensemble optimization<sup>5,</sup> and refinement with QRNAS. The structural ensemble collectively explains the scattering pattern of the RNA molecule. The workflow also allows the use of data from other sources, such as information about RNA secondary structure from computational predictions or experimental probing<sup>6</sup>.

Our method is capable of modeling 3D structures of RNA molecules, as well as those of RNA-protein and RNA-ligand complexes. In case of complexes, this method in conjunction with experimental restraints from other techniques, also allows us to identify a predominant binding mode of the RNA and its partner. We have applied our methods to case studies involving RNA molecules and complexes of different size.

#### References

- 1 Ponce-Salvatierra, A. et al. Biosci. Rep. 39, BSR20180430 (2019)
- 2 Boniecki, M. J. et al. Nucleic Acids Res. 44, e63 (2016)

**3** Svergun, D. *et al. J. Appl. Crystallogr.* 28, 768–773 (1995)

- 4 Konarev, P. V. et al. J. Appl. Crystallogr. 39, 277–286 (2006)
- 5 Tria, G. et al. IUCrJ 2, 207–217 (2015)
- 6 Patel, T. R. et al. Methods 118-119, 146-162 (2017)

# Small angle X-ray scattering data assisting protein structure prediction with coarse-grained UNRES simulations.

<u>Cezary Czaplewski<sup>1</sup></u>, Agnieszka S. Karczyńska<sup>1</sup>, Magdalena A. Mozolewska<sup>1,2</sup>, Paweł Krupa<sup>1,3</sup>, Artur Giełdon<sup>1</sup>, Adam Liwo<sup>1</sup>

<sup>1</sup>Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, Gdańsk 80-308, Poland
<sup>2</sup>Institute of Computer Science, Polish Academy of Sciences, Jana Kazimierza 5, Warsawa 01-248, Poland

<sup>3</sup>Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, Warsawa 02-668, Poland

A new approach to small angle X-ray scattering (SAXS) assisted protein-structure prediction is presented, which is based on running multiplexed replica exchange molecular dynamics simulations with the coarse-grained UNRES force field. Distance distribution derived from SAXS data is included as a maximum-likelihood penalty term with the coarse-grained UNRES energy function. To improve the quality of the resulting models and to speed up the search, we also use restraints derived from knowledge-based models of simulated proteins.

The coarse-grained UNRES model developed in our laboratory is a highly reduced model of polypeptide chains, with two interaction sites per residue: united side chains and united peptide groups. The UNRES force field is physics-based and originates from the potential of mean force of polypeptide chains in water, in which the degrees of freedom not present in the model have been averaged out. The UNRES force field is based on Kubo's cluster-cumulant expansion and includes the multibody terms responsible for the formation of regular secondary structures. The complete pseudo-energy function used in simulations consists of the UNRES energy function, the template-restraint penalty term, and the SAXS-restraint term. We use the distance distribution derived from the SAXS data to restraint the simulations instead of the scattering intensity usually employed in SAXS penalty terms. The distance distribution is more closely related to a conformation than its Fourier transform, the scattering intensity, and, therefore, seems to be more sensitive to conformation changes. Calculating distance distribution we approximated the contribution from a given pair of residues as a Gaussian distribution with the standard deviation calculated from the Stokes radii of these residues. The Gaussian approximation to the distance distribution due to a pair of coarse-grained sites partially accounts for the solvation-shell water.

# KAP1 is an antiparallel dimer with a natively functional asymmetry

## Matteo Dal Peraro

Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, 1015 Switzerland

KAP1 (KRAB-domain associated protein 1) plays a fundamental role in regulating gene expression in mammalian cells by recruiting different transcription factors and altering the chromatin state. In doing so, KAP1 acts both as a platform for macromolecular interactions and as an E3 SUMO ligase. Using an integrative modelling approach, we sheds light on the overall organization of the full-length protein combining solution scattering diffraction data and single-molecule experiments. We show that KAP1 is an elongated antiparallel dimer with a native asymmetry at the C-terminal domain. This conformation supports our finding that the RING domain contributes to KAP1 auto-SUMOylation. Importantly, this intrinsic asymmetry has key functional implications for the KAP1 network of interactions, as the heterochromatin protein 1 (HP1) occupies only one of the two putative HP1 binding sites on the KAP1 dimer, resulting in an unexpected stoichiometry, even in the context of chromatin fibers.

#### Micelles & Monolayers: Exploring use of models to understand scattering data Karen J Edler<sup>1</sup>, Andrew McCluskey<sup>1,2</sup>, Steve C Parker<sup>1</sup>, Daniel Bowron<sup>3</sup>

1. Department of Chemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK

2. Diamond Light Source, Harwell Campus, Didcot, OX11 0DE, UK.

3. ISIS, Science and Technology Facilities Council Rutherford Appleton Laboratory, Harwell Oxford, Didcot OX11 0QX, UK

With the continuing improvements to small angle scattering instruments it is now possible to routinely collect data out to 1 Å<sup>-1</sup> or even higher in many SAS experiments. In addition, instruments such as NIMROD at ISIS are extending their low angle range into the small angle range. This extended range of data brings new opportunities for data analysis, which can potentially cover structures from molecular level, or at least coarse-grained levels through to colloidal particles. The major question is how to efficiently fit experimental data with atomistic models on feasible timescales. I will present some of our work on analysing micelle structure using Empirical Potential Structure Refinement (EPSR) and on looking at appropriate levels of coarse graining and potentials for generating models of lipid monolayers at the air-water interface.

# The atsas software suite for small angle scattering from macromolecular solutions

# Daniel Franke

European Molecular Biology Laboratory, Hamburg, Germany

The ATSAS software suite for the analysis of small angle scattering of biological macromolecules in solution is a comprehensive cross-platform package with applications for all stages of the Small Angle X-ray and Neutron Scattering workflow, from primary data processing to ab-initio and rigid body modelling (Franke, et. al., J Appl Cryst, 2017). It includes applications to integrate information from X-ray crystallography, nuclear magnetic resonance spectroscopy, electron microscopy and atomistic homology modelling to construct hybrid models based on the scattering data.

Here an overview of well-established ATSAS applications and more recent developments will be provided. This includes, but is not limited to: CRYSOL/CRSYON to calculate SAXS and SANS patterns from atomistic models and to fit them to experimental data; DAMIN/DAMMIF for ab-initio bead modelling, with the recent addition of DAMMIX, a tool to restore the shape of an unknown component in evolving systems ab initio, e.g. during fibril formation (Konarev & Svergun, IUCrJ, 2018); SASREF for rigid-body modelling; and EOM to approximate ensembles.

Further, a new application will be introduced: IMGSIM simulates realistic 2D scattering images from a 1D input file, e.g. calculated from an atomic model by CRYSOL. Based on the input data, IMGSIM generates random scattering events onto a virtual detector plane, taking into account the detector distance, dimensions and pixel size, incoming virtual photon/neutron flux and wavelength, as well as the sample concentration (for zero concentration, a flat background is generated). The resulting 2D images can be readily radially averaged by existing tools and the variation in the simulated data exhibits the expected statistical properties (normal distribution, statistically independent). With initial input on absolute scale, [cm<sup>-1</sup>]/c[mg/ml], the simulated data frames may be scaled to absolute scale by DATABSOLUTE, and the final I(0) after subtraction of the background will be proportional to the molecular weight of the input model. The angular range and spacing, as well as perceived noisiness depends on the input parameters and may be set to mimic the experimental data obtained at existing instruments. Thus, the simulated data may directly be used to validate existing data processing and modelling procedures, as well as for development of new methods.

ATSAS is freely available for academic users to download at https://www.embl-hamburg.de/biosaxs/software.html.

# Acknowledgements

The research was funded by Horizon 2020 programme of the European Union, grant iNEXT project number 653706, and Bundesministerium für Bildung und Forschung, grant TT-SAS number 05K2016.

#### Solving the phase problem in solution scattering

Thomas D. Grant<sup>1,2</sup>

<sup>1</sup>SUNY University at Buffalo, Buffalo, NY <sup>2</sup>Hauptman-Woodward Institute, Buffalo, NY

Small angle scattering is an experimental technique used to analyze the molecular structures of a wide variety of biological and non-biological samples in solution. In contrast to X-ray crystallography and cryoelectron microscopy, where 3D electron density maps are calculated, available methods for generating 3D structural information from 1D solution scattering data rely exclusively on modeling. Many modeling algorithms rely on an implicit assumption that electron density is uniform inside the particle envelope. This assumption breaks down at resolutions better than approximately 10 - 15 Å where fluctuations in electron density contribute significantly to scattering and for particles with large scale conformational dynamics or containing mixed density species. Here I present a method<sup>1</sup> for calculating electron density maps directly from solution scattering data. Using only few simple restraints such as solvent flattening, this method avoids many of the assumptions limiting the resolution and accuracy of conventional modeling algorithms. The algorithm has been applied to publicly available experimental scattering data from twelve different biological macromolecules. In each case the electron density maps closely match known atomic models, including complex shapes with multiple density components. These results demonstrate that accurate and complex electron density maps can be reconstructed from small angle scattering data and with significantly fewer restraints than imposed by existing modeling methods.

<sup>1</sup>Grant, T. D. (2018). Ab initio electron density determination directly from solution scattering data. *Nature Methods* **15**, 191-193.



# Simulations of concentrated solutions of monoclonal antibodies and sticky cylindrical particles using open-source software

H. W. Hatch (speaker),<sup>1</sup> R. Murphy,<sup>1</sup> M. A. Blanco,<sup>1,2</sup> H. Kang,<sup>1,2</sup> N. A. Mahynski,<sup>1</sup> J. E. Curtis,<sup>1</sup> N. J. Wagner,<sup>3</sup> V. K. Shen<sup>1</sup>

1. National Institute of Standards and Technology, USA

2. Institute for Bioscience and Biotechnology Research, USA

3. University of Delaware, USA

The advantages of using simulation to analyze and interpret small angle scattering data are exemplified in recent work on the solution structure of coarse-grained models of monoclonal antibodies (mAbs) under concentrated conditions, and on the clustering of cylindrical colloids with short-range attractions. For the first example, rational strategies need to be developed to avoid physical instabilities in mAb formulations such as aggregation and increased viscosity. In the second example, cylindrical particles are promising materials for applications of fillers in nanocomposite materials and additives to control rheological properties of colloidal suspensions. In both examples, we performed flat histogram Monte Carlo simulation of these systems with specialized algorithms using the Free Energy and Advanced Sampling Simulation Toolkit (FEASST). FEASST is freely available at <a href="https://pages.nist.gov/feasst">https://pages.nist.gov/feasst</a>. We highlight the effect of intra-domain flexibility of the mAbs on the small angle scattering data. We also show how scattering data for cylindrical particles was used to collapse the attractive-driven gel boundary over a range of aspect ratios.

## **Aggregation in Surfactant-Free Ternary Mixtures**

#### Dominik Horinek

Institut für Physikalische und Theoretische Chemie, Universität Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany

In recent years it has become more and more evident that presumably simple solutions can show a variety of nanoscale aggregation structures. Ternary mixtures of three liquids, in which two show only partial mutual solubility, resemble different types of microemulsions even in the absence of classical surfactants. We present fully atomistic molecular dynamics simulations of octanol/ethanol/water mixtures, a typical representative of these "surfactant-free microemulsions". We compare MD simulations results with experimental SAXS/WAXS and neutron scattering data and present a detailed analysis of the observed structures. We recognize three types of microemulsions: i) a direct oil-in-

water microemulsion consisting of octanol rich aggregates in a water rich phase, ii) a bicontiuous sponge-like microemulsion, and iii) a reverse water-in-oil mircoemulsion, in which the polar phase comprises a network of hydrogen bondend structures.



1A 2 nm thick slice through a snapshot of an octanol/ethanl/water mixture in a region of the phase diagram where a bicontinuous microemulsion is formed

# References

1) Schöttl,S.; Lopian, T.; Prévost, S.; Touraud, D.; Grillo, I.; Diat, O.; Zemb, T.; Horinek, D. Combined MD/SAS analysis of organization on a nanometer-scale in ternary solvent solutions containing a hydrotrope. J. Coll. Inter. Sci. 2019, 540, 623-633.

2) Schöttl, S.; Horinek, D. Aggregation in Detergent-Free Ternary Mixtures with Microemulsion-Like Properties Curr. Op. Coll. Inter. Sci. 2016, 22, 8–13.

3) Lopian, T.; Schöttl, S.; Prevost, S.; Pellet-Rostaing, S.; Horinek, D.; Kunz, W.; Zemb, T. Morphologic transitions for microemulsions based on hydrotropic co-solvents near the miscibility gap of ternary solutions in the presence of additives ACS Cent. Sci. 2016, 2, 467–475.

4) Schöttl, S.; Marcus, J.; Diat, O.; Touraud, D.; Kunz, W.; Zemb, T.; Horinek, D. Emergence of Surfactant-Free Micelles from Ternary Solutions Chem. Sci. 2014, 5, 2949-2954.

# SAXS-guided structure and ensemble refinement using explicit-solvent MD simulations. Recent developments and remaining challenges

# Prof. Dr. Jochen Hub

Theoretical Biophysics Group, Department of Physics, Saarland University, Campus E2 6, Zi. 4.11, 66123 Saarbruecken, Germany

The interpretation of the SAS data by computational methods is complicated by the low information content of the data, by scattering contributions from the hydration layer and excluded solvent, and by unknown systematic errors. Explicit-solvent MD simulations may help to overcome such challenges because they add physical information to the low-information experiential data. Specifically, modern force fields may help to identify energetically reasonable conformations during structure refinement, and they provide an accurate description of the hydration layer and of the excluded solvent. After a brief introduction to MD-based SAXS data interpretation, recent developments are presented with some focus on Bayesian methods, maximum-entropy refinement, and intrinsically disordered proteins.

# Asymptotic analysis of quasielastic neutron scattering data from human Acetylcholinesterase reveals subtle dynamical changes upon ligand binding

<u>Gerald Kneller<sup>1,2</sup></u>, Melek Saouessi<sup>1,2</sup>, Judith Peters<sup>3</sup>,

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS and Université d'Orléans, Orléans, France <sup>2</sup>Synchroton SOLEIL, Gif-sur-Yvette, France <sup>3</sup>Institut Laue Langevin, 71 avenue des Martyrs, 38042 Grenoble Cedex 9, France; Université Grenoble Alpes, CNRS, Laboratoire Interdisciplinaire de Physique, F-38000 Grenoble, France

# E-mail: gerald.kneller@cnrs.fr

We report on the application of a new, "model-free" approach to analyzing quasielastic neutron scattering spectra from protein powder samples [1] to previously collected QENS data [2] from the enzyme Human Acetylcholenisterase (hACE) with and without the non-covalently bound inhibitor HuperZine A. Our QENS analysis is based on the fact that the form of the quasielastic line for small frequencies can be directly related to the relaxation of the intermediate scattering function for long times. To account for the multiscale dynamics of proteins, we describe the relaxation by a stretched Mittag-Leffler function, which displays slow power law decay for long times and a broad spectrum of relaxation rates. Using Zwanzig's model for diffusion in a rough potential [3], we translate the relaxation rate spectrum into a distribution of energy barriers of the protein energy landscape of global harmonic form [4]. Our analysis reveals that binding of Huperzine A increases the atomic motional amplitudes and slightly slows down its internal diffusive motions. This can be visualized in terms of a softening and roughening of the potential energy surface. A corresponding normal mode analysis is presented and discussed.

# **References:**

- [1] G. R. Kneller, "Franck–Condon picture of incoherent neutron scattering", *PNAS USA*, vol. 115, no. 38, pp. 9450–9455, 2018.
- [2] M. Trapp, et al, "Correlation of the dynamics of native human acetylcholinesterase and its inhibited huperzine A counterpart from sub-picoseconds to nanoseconds," Journal of the Royal Society Interface, vol. 11, pp. 20140372–20140372, May 2014.
- [3] R. Zwanzig, "Diffusion in a rough potential.," PNAS USA, vol. 85, pp. 2029–2030, 1988.
- [4] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, "The energy landscapes and motions of proteins", *Science*, vol. 254, no. 5038, pp. 1598–1603, 1991.
- [5] M. Saouessi, J. Peters, and G.R. Kneller, "Asymptotic analysis of quasielastic neutron scattering data from human Acetylcholinesterase reveals subtle dynamical changes upon ligand binding", manuscript submitted.

# Knotted and unknotted ring polymers under shear

Maximilian Liebetreu and Christos N. Likos

## Faculty of Physics, University of Vienna

The behavior of single knotted rings in a fluid under shear is compared to their unknotted ring polymer counterparts. We simulate flexible polymers of a fixed size in a thermostatted Multi-Particle Collision Dynamics (MPCD) solvent (with fixed control parameters) with Lees-Edwards boundary conditions. We primarily investigate the differences in shape parameters for knotted and unknotted rings in dependency of shear rate, as well as characteristics of average number of beads being part of the knot (knot size), angle between knot center of mass and first principal axis relative to the polymer's center of mass, and correlations between quantities. We compute the relaxation time of knotted rings and present results on their tumblingand tank-treading dynamics. We obtain evidence suggesting that on a knotted ring, the knot itself develops a tendency to be located near those beads closest to the orientational axis, aligned with the flow. We also show the average knot size decreasing with increasing shear. Preliminary findings indicate the knotted ring responding to lower shear rates than their unknotted counterparts, and suggest a binary-state behavior for the 31-knotted ring under strong shear, with the knot size alternating between rather stable tight and relatively unstable delocalised configurations. Special attention is paid to the correlation between alignment angle and knot size.

# A fast and accurate SAS calculation from MD simulations

Massimo Marchi<sup>1,2</sup>

1 Commissariat à l'Énergie Atomique, DRF/Joliot/SB2SM/LBMS & CNRS UMR 9198, Gif sur Yvette 91191 Cedex, France

2 Institut de Biologie Intégrative de la Cellule (I2BC), Institut Frédéric Joliot, CEA, CNRS, Univ Paris-Sud, Université Paris-Saclay, Gif-Sur-Yvette F-91198 Cedex, France

Our talk will focus on the calculation of SAXS (and SANS) profiles from molecular dynamics simulations. Applications to globular and membrane proteins along with DNA and RNA will show that SAXS calculations can be used to probe and validate force fields for bio-molecular systems in solution. The SAXS method that we have developed uses 3D fast Fourier transforms (FFTs) and particle meshing interpolation, similar to the one used in smooth particle mesh Ewald for electrostatic energies and forces, combined with a uniform solvent density FFT padding scheme to obtain a convenient SAXS spectral resolution. As a FFT's approach, our method scales favorably for systems of increasing size and is devoid of many of the artifacts of other approaches proposed so far in the literature.

# Synthesis routes and crowding effects on single-chain polymeric nanoparticles: Combining simulations and small-angle neutron and X-ray scattering

**Angel J. MORENO**<sup>1,2</sup>, Marina GONZÁLEZ-BURGOS<sup>1,3</sup>, Arantxa ARBE<sup>1</sup>, Federica LOVERSO<sup>2,4</sup>, Aurel RADULESCU<sup>5</sup>, José A. POMPOSO<sup>1,3,6</sup>, Juan COLMENERO<sup>1,2,3</sup>

<sup>1</sup>Centro de Física de Materiales (CSIC, UPV/EHU) and Materials Physics Center (MPC), San Sebastián, Spain

<sup>2</sup>Donostia International Physics Center, San Sebastián, Spain

<sup>3</sup>Universidad del País Vasco (UPV/EHU), Departamento de Física de Materiales, San Sebastián, Spain

<sup>4</sup>University of Southern Denmark, Department of Physics, Chemistry and Pharmacy, Odense, Denmark

<sup>5</sup>Forschungszentrum Jülich GmbH, Jülich Centre for Neutron Science at Heinz Maier-Leibnitz Zentrum, Garching, Germany

<sup>6</sup>IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Polymeric single-chain nanoparticles (SCNPs) are an emergent class of soft nano-objects of molecular size of 5-20 nm, resulting from the purely intramolecular cross-linking of the reactional functional groups of single polymer precursors. A growing interest is being devoted in recent years to develop a SCNP-based technology with multiple applications in catalysis, nanomedicine, or rheology, amongst others. To this end, we need good control over the size and shape of SCNPs, as well as a deeper understanding of their behaviour in complex situations as macromolecular crowding. By means of computer simulations of coarse-grained models and validation by small-angle neutron and X-ray scattering, we design and investigate different protocols leading to SCNPs with specific structures and different properties in solution.

The analysis of the conformations of SCNPs synthesized in good solvent reveals that they share basic ingredients with intrinsically disordered proteins (IDPs), as topological polydispersity, sparse conformations, and compact local domains [1]. Unlike in the case of linear macromolecules, crowding leads to collapsed conformations of SCNPs resembling those of crumpled globules [1,2], at volume fractions (about 30%) that are characteristic of crowding in cellular environments. This result is apparently universal and independent of the architecture of the polymers crowding the environment of the SCNP [2]. Our results for SCNPs - a model system free of specific interactions - propose a general scenario for the effect of steric crowding on IDPs.

[1] A.J. Moreno, F. Lo Verso, A. Arbe, J.A. Pomposo, J. Colmenero, Journal of Physical Chemistry Letters **7**, 838 (2016)

[2] M. González-Burgos, A. Arbe, A.J. Moreno, J.A. Pomposo, A. Radulescu, J. Colmenero, Macromolecules **51**, 1573 (2018)

E-mail of the corresponding author: angeljose.moreno@ehu.es

# Insights on Water Structure from the Modeling of Scattering Data

Roland R. Netz, Philip Loche, Alexander Schlaich, Emanuel Schneck

Department of Physics, Free University Berlin, Arnimallee 14, 14195 Berlin, Germany

The structural and dynamic properties of water in bulk and at interfaces are relevant for many physicochemical processes. Insight can be gained from all-atomistic simulations of interfacial water that nowadays reach the experimentally relevant length and time scales. This is demonstrated with a few examples:

i) For bulk water, simulations reproduce the main features of the experimental SAXS structure factor S(q), including the minimum at small q. Spatial correlations between local density and structural fluctuations are shown to be weak, suggesting that features in density-density correlations (such as measured by the structure factor) are not straightforwardly related to structural spatial correlations in liquid water. [1]

ii) Hydrophobic surfaces in contact with water show a pronounced depletion layer where the water density is highly reduced[2]. The combination of X-ray reflectivity (XR) and atomistic molecular dynamics (MD) simulations provide new insight into simple oil-water interfaces. [3]

iii) The bending rigidity and the structure of the water-vapor interface is obtained from an analysis of the capillary interfacial wave spectra from simulations and grazing-incidence x-ray scattering. A pronounced scale dependence of the bending rigidity points to a crossover at a length scale of about one nanometer: At smaller scales the interface reconstructs and nano-ripples proliferate, at larger scales the interface is governed by tension and bending rigidity. [4]

iv) The combination of x-ray scattering experiments and molecular dynamics simulations yields insight into the structure and the swelling mechanics of hydrated collagen. With a proper choice of simulation force fields the experimental packing distances as well as osmotic pressures for different collagen types are reproduced.

[1] Spatial Correlations of Density and Structural Fluctuations in Liquid Water: A Comparative Simulation Study Felix Sedlmeier, Dominik Horinek, and Roland R. Netz

J. Am. Chem. Soc. 2011, 133, 1391–1398

[2] Interfacial Water at Hydrophobic and Hydrophilic Surfaces: Depletion versus Adsorp-

tion, J. Janecek and R.R. Netz, Langmuir 23, 8417 (2007).

[3] Nanoscale Structure of the Oil-Water InterfaceM. Fukuto, B. M. Ocko, D. J. Bonthuis, R. R. Netz, H.-G. Steinrück, D. Pontoni, I. Kuzmenko, J. Haddad, and M. Deutsch

PRL 117, 256102 (2016)

[4] Nanoroughness, Intrinsic Density Profile, and Rigidity of the Air-Water Interface Felix Sedlmeier, Dominik Horinek, and Roland R. Netz

PRL 103, 136102 (2009)

# Asymmetry, Registration, Flip-Flop, and Phase Separation in Lipid Bilayers

# Peter Olmsted and John Williamson

Lipid bilayers *in vivo* are multicomponent mixtures that are capable of phase separation. Model studies of bilayers often demonstrate complex phase behaviour, which in most cases is symmetric across the bilayer, so that the two leaflets locally have the same phase. However, physiological bilayers have very different compositions in the two leaflets, and the environments on either side of the bilayers are often very different (pH, ionic strength, different molecular species, ...). Hence there are numerous reasons to consider phase separation in which the two leaflets have different local phases and/or compositions. One can even have anti-registered phase separation, in which leaflets are locally *always* in different phases! I will explore the kinetics and equilibrium behaviour of these kinds of phase separation based on different levels of modeling, and discuss signatures of asymmetry and anti-registration in experiments in the literature, as well as challenges for scattering experiments.

# **Recent Developments in Fluctuation X-ray Scattering**

Kanupriya Pande<sup>1,2</sup>, Jeffrey J. Donatelli<sup>1,3</sup>, and Petrus H. Zwart<sup>1,2</sup>

 Center for Advanced Mathematics in Energy Research Applications (CAMERA), Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA
Molecular Biophysics and Integrated Bioimaging (MBIB) Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA
Department of Mathematics, Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA

Fluctuation X-ray scattering (FXS) is a biophysical technique that overcomes the low data-to-parameter ratios encountered in traditional X-ray scattering methods used for studying non-crystalline samples [1-3]. In an FXS experiment solution scattering data are collected from particles in solution using ultrashort X-rays of pulse lengths shorter than the rotational diffusion time of the particles. The resulting data contains angularly varying information that yields structures with a greater level of detail than those obtained using tradition SAXS.

After successful application of FXS to single-particle data [4], we have recently demonstrated the experimental feasibility of this technique on data from an ensemble of PBCV-1 virus particles in solution [5, 6]. We found that by using advanced noise-filtering methods the required number of images to obtain decent correlation data is far lower than originally expected, thereby reducing the required data collection time to less than a few minutes. In this talk I will outline the data processing techniques for analysis of FXS data, and present an assessment of the effect of concentration and various sources of noise on experimental data.

This research was supported by the Advanced Scientific Computing Research and the Basic Energy Sciences programs of the Office of Science of the US Department of Energy (DOE) under, and the National Institute of General Medical Sciences of the National Institutes of Health (NIH).

- 1 Kam Z. Determination of Macromolecular Structure in Solution by Spatial Correlation of Scattering Fluctuations. Macromolecules **10** (5): 927--34 (1977).
- 2 Kam Z, Koch MH, Bordas J. Fluctuation X-ray scattering from biological particles in frozen solution by using synchrotron radiation. Proceedings of the National Academy of Sciences of the United States of America 78 (6): 3559—62 (1981).
- Malmerberg E, Kerfeld CA, Zwart PH. Operational properties of fluctuation X-ray scattering data. IUCrJ **2**(Pt 3): 309--16 (2015).
- 4 Kurta RP, *et al.* Correlations in Scattered X-Ray Laser Pulses Reveal Nanoscale Structural Features of Viruses. Phys Rev Lett. **119** (15): 158102 (2017).
- 5 Pande K, Donatelli JJ, Malmerberg E, Foucar L, Bostedt C, Schlichting I, Zwart PH. Ab initio structure determination from experimental fluctuation X-ray scattering data. Proceedings of the National Academy of Sciences of the United States of America **115** (46): 11772—11777 (2018).
- 6 Pande K, *et al.* Free-electron laser data for multiple-particle fluctuation scattering analysis. Sci Data. **5**:180201 (2018).

# Memprot & Dadimodo: programs for modeling the detergent belt in solubilized membrane protein complexes & re-orienting domains of multi-domain proteins

Maciej Baranowski<sup>1</sup>, Olga Rudenko<sup>2</sup>, Aurélien Thureau<sup>1</sup>, Javier Pérez<sup>1</sup>

1 Swing Beamline, Synchrotron Soleil, L'Orme des Merisiers, Saint-Aubin, France

2 Transverse Operations, Synchrotron Soleil, L'Orme des Merisiers, Saint-Aubin, France

# E-mail: javier.perez@synchrotron-soleil.fr

Membrane proteins participate in many critical cellular processes, including apoptosis, drug import/export and cell division, and are thus very important for biomedical research. Despite their importance, the research on membrane proteins is hampered by insufficient understanding of methods used to solubilize them, like detergent micelles, detergent bicelles or polymer/detergent nanodiscs. Moreover, the ability to model the solubilizing agent around a transmembrane protein is often necessary for further investigations of the protein itself. Therefore, to facilitate research into structure of

solubilizing agents (alone and in complex with membrane proteins), the program Memprot<sup>1</sup> was developed in SWING beamline. Memprot builds a low resolution model of the solubilizing agent belt that surrounds a protein of known atomic coordinates by trying to fit the experimental SAXS scattering curve of the complex. The initial version of Memprot could be used to model a corona of single detergent around a protein. Recent developments in new versions of Memprot include, among others:

- Models of detergent bicelles and detergent/polymer nanodiscs
- Adaptive Shape Algorithms which allow for deformation of detergent corona based on the shape of the trans-membrane region of the protein
- Parallelization with MPI and code optimizations

Based on a genetic algorithm, the program Dadimodo is meant to "refine" the 3D structure of a multi-domain protein complex against SAXS experimental data, the degrees of freedom being the backbone dihedral angles of the protein flexible fragments. A complete (all-atoms) structure representation is used with energy control for every newly generated model so as to prevent steric clashes and converge to a physically feasible structure. The program takes its roots from [2, 3]. The new version offers an automated analysis of the structure topology (fixed vs. variable parts of the structure) which extends the applicability of the software to a larger class of topologies and opens its use to a larger community of users. Moreover, the presented update allows for a better exploration of the search space at the advanced evolution stage thanks to an adaptive mechanism of decreasing the amplitude of the conformation changes. In terms of implementation, the calculation is speeded up as a result of parallel execution enabled through the DEAP package [4]. Dadimodo addresses domain re-orientation of both soluble proteins and membrane proteins with a belt modeled using Memprot. Since last year Dadimodo has been made available to the research community as a web application [5].

# References

[1]. J. Pérez, A. Koutsioubas, Memprot: a program to model the detergent corona

around a membrane protein based on SEC-SAXS data, Acta Crystallographica Section D, 2015, 71, 86-93.

[2] G. Evrard, F. Mareuil, F. Bontems, C. Sizun, and J. Pérez (2011), DADIMODO: a program for refining the structure of multidomain proteins and complexes against small-angle scattering data and NMR-derived restraints. Journal of Applied Crystallography 44, 1264–1271.

[3] F. Mareuil, C. Sizun, J. Pérez, M. Schoenauer, JY. Lallemand, and F. Bontems (2007). A simple genetic algorithm for the optimization of multidomain protein homology models driven by NMR residual

dipolar coupling and small angle X-ray scattering data. Eur. Biophys J 37(1), 95–104.

[4] Félix-Antoine Fortin, François-Michel De Rainville, Marc-André Gardner, Marc Parizeau, and Christian Gagné (2012). DEAP: Evolutionary Algorithms Made Easy. J. Mach.Learn. Res. 13, 1 (2012), 2171–2175.

[5] O. Roudenko, A. Thureau and J. Pérez (2018) Dadimodo: Refining Atomic Models of Multi-Domain Proteins against SAXS Data, <u>https://dadimodo.synchrotron-soleil.fr</u>

# CCP-SAS – A COMMUNITY CONSORTIUM FOR THE ADVANCED MODELLING OF SCATTERING DATA – APPLICATIONS AND IMPLEMENTATION

#### Stephen J. Perkins and Jayesh S. Bhatt

University College London, London, WC1E 6BT, UK

The CCP-SAS [1] project (http://www.ccpsas.org/) began as a joint UK/US funded consortium to address the increasingly complex needs of scattering data analysis. A typical bench scientist synthesis, purifies and characterizes samples, collects the scattering data, and interprets the results, often using simplistic models. It is rare that the same individual also has the skills to use advanced atomistic simulation software, for example. CCP-SAS's initial goals were to build a web-based GUI front-end termed GenApp [2], coupled to a high-performance back-end and to develop advanced analysis modules and new simulation methods within a SASSIE-web workflow (https://sassie-web.chem.utk.edu/sassie2/) [3]. It allows non-computational experts to perform molecular simulations of biological and soft matter systems, calculate their theoretical SAXS or SANS profiles, and determine the best-fit atomistic structures against their experimental data.

In various applications of CCP-SAS to date, an increasing range of mostly biological macromolecules have been tackled. The first part of this presentation will discuss recently-completed applications to antibodies and the complement proteins by the University College London group in London, UK. Human IgG antibodies exist as four subclasses IgG1-IgG4, in which the two Fab and one Fc regions are connected by hinges of different lengths. The combination of molecular dynamics and Monte Carlo simulations with **both** SAXS and SANS experimental data revealed the power of the joint scattering approach in identifying the best-fit solution structures for IgG1, IgG2 and IgG4. This work showed how these best fit structures account for different receptor binding properties between the IgG subclasses [4]. For the complement proteins, which correspond to a diverse range of macromolecular structural types, we have looked at Factor H (FH) which is a major regulator of the activated form C3b of C3 in innate immunity. FH contains 20 small SCR domains. Combined with improved SAXS data sets, molecular dynamics and Monte Carlo simulations revealed two folded-back domain structures that interestingly expose either the N-terminus or the C-terminus of FH to interact with either C3b or the breakdown product of C3b respectively [5]. Other applications of molecular dynamics and Monte Carlo simulations to complement included the identification of a flexible and bendable collagen triple helix structure in solution, and notable salt-dependent conformational changes of C3b and C4b in solution that explained different functional reactivities.

More importantly for CCP-SAS, the process of community building, of connecting people and projects, and of fostering contributions to existing tools and infrastructures, needs to accelerate. With this goal in mind, we are presently installing the SASSIE-web package on a high performance computing (HPC) facility at UCL in London in order to realise its full potential in a local multi-user environment. We have found that this installation on HPC is a non-trivial task. We will highlight some of the challenges involved, and the best practice to be followed for such installations.

#### References

[1] S. J. Perkins et al., J. Appl. Cryst. 49 (2016) 1861-1875.

[2] A. Savelyev et al., Fut. Gen. Comp. Sys. (2017) doi:10.1016/j.future.2017.09.069.

[3] J. E. Curtis et al., Comp. Phys. Comm. 183 (2012) 382-389.

[4] D. W. Wright, S. J. Perkins et al., (2019) submitted/reviewed

[5] A. J. Osborne et al., J. Biol. Chem. 293 (2018), 17166-17187.

# Integrating molecular simulation with neutron scattering to study flexible biological systems

#### Loukas Petridis

# Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

Neutron scattering will address frontier challenges in biological research by providing crucial information about the three-dimensional (3D) structure and dynamics of biological systems, which underpins the molecular basis of biological processes. Small-angle neutron scattering (SANS) is an ideal technique to investigate dynamic and flexible biosystems, whose importance in cellular function is increasingly recognized. However, a critical factor limiting the impact of SANS is the interpretation of experimental data on flexible systems. The specific challenge is to determine the "conformational ensemble" i.e., the 3D structures that give rise to the experimental SANS signal. Overcoming this challenge requires the integration of SANS with high performance molecular dynamics simulation. We discuss targeted computational methods that scale on the supercomputers and permit interpretation of SANS data with realistic physical models.

# **Reduction of SAXS data into sets of structural moments**

Frédéric Poitevin<sup>1,2</sup>, Raul O. Girbal<sup>1</sup>, Rasmus Fonseca<sup>1</sup>, Jérôme Houdayer<sup>3</sup>

- 1) Levitt Lab, Dpt of Structural Biology, Stanford School of Medicine, CA, USA
- 2) CryoEM & Bioimaging Dpt, SLAC National Accelerator Laboratory, CA, USA
- 3) IPhT, CEA-Saclay, France

How much information is encoded in a small-angle scattering (SAS) curve? What does it tell us about the structure of the scattering object? Many have provided insights to answer those questions in the past, and we provide here an approach that we hypothesize will help refine shape classification of scattering objects.

In this talk, I will briefly outline the foundational answers to these questions. Then I will show how one can exploit the invariant properties of the scattering intensity to expand the SAS curve to a simple sum over a basis of functions well separated in the scattering space. We call the weights of the basis functions Kn that, together with the radius R of the support of the scattering object, give a complete theoretical description of the SAS curve. Those weights also entirely describe the pair distance distribution function (PDDF) of the object, as they are related to its moments. Adding an overall constant, these parameters are easily fitted against experimental data giving a concise comprehensive description of the data.

In addition to the understanding they bring, these invariants can also be used to reliably estimate structural moments beyond the radius of gyration, thereby rigorously defining the actual set of model-free quantities one can extract from experimental SAS data. I will conclude on how these quantities can help narrow down possible structural models for the scattering object.

Houdayer & Poitevin. Acta Cryst. A73, 317-332 (2017)

# Bayesian inference of conformational ensembles from limited structural data

Wojciech Potrzebowski

ESS, Lund, Sweden

The small-angle scattering (SAS) from proteins in solution samples the ensemble average of the randomly oriented structures, and ensemble modelling for proteins with flexible regions against SAS data is increasingly popular. However, the smooth SAS profile can typically be defined by as few as 10-15 points, and the ensemble model has many more degrees of freedom. Typically, a very large ensemble (10,000 or more) is generated within some constrained set, and a population weighted sub-set of structures is identified that predicts a profile that best-fits the data. Representative structures are selected based on clustering analysis to aid in visualizing the nature of the ensemble, but their accuracy and what minimal set is justified by the data are outstanding questions.

In this study we use a model evidence to select ensembles with an optimal number of members. Model evidence (ME) is widely used in Bayesian model comparison and provides an automatic Occam's razor effect by balancing between fit to data and model complexity, thereby providing a rigorous approach to combat overfitting. However, ME is a multidimensional integral that can be very difficult to evaluate and this serves as a significant barrier to its use in ensemble selection. Our ensemble method is based on an approximate variational Bayesian inference (VBI) method for model selection. The VBI approach has two major benefits. First, it is significantly faster than complete Bayesian inference, which enables the use of large structural libraries. Second, VBI implicitly leads to maximization of ME without the need for evaluation of a multidimensional integral. However, a downside of the VBI approach is the approximation of the posterior inherent in method. Hence, after arriving at the optimal ensemble with VBI we carry out a complete Bayesian inference of weights. This enables quantification of uncertainties in the ensemble model and population weights.

A significant benefit of Bayesian methods is that multiple experimental observations can be rigorously combined in both model selection and weight inference to gain insight into the underlying ensemble. It is also possible to combine experimental data with information from simulations and force fields. This is exemplified in this study where we demonstrate how data from SAXS, NMR and structural energy values of individual conformers can be combined into one probabilistic model. The inference machinery is applied study the conformational ensembles of two-domain proteins studied by SAXS. We demonstrate the feasibility of full Bayesian inference from large structural libraries from detailed all atom simulations and show by simulations that the method is capable of accurate recovery of population weights and ensemble sizes. We also investigate how noise in the experimental data impacts the accuracy of ensemble inference and show that information encoded in energy functions can compensate for noisy SAXS data. The method is applied to two systems with experimental data from SAXS and NMR: calmodulin and cardiac myosin binding protein C. This analysis demonstrated how the combination of data from SAXS, NMR and structural energies from conformational sampling simulations can be used in synergy for improved ensemble inference.

### Macromolecular structure and dynamics based on SAXS profiles

#### Dina Schneidman

The Hebrew University of Jerusalem, Jerusalem, Israel.

Proteins generally populate multiple structural states in solution. Transitions between these states are important for function, such as allosteric signaling and enzyme catalysis. Structures solved by X-ray crystallography provide valuable, but static, atomic resolution structural information. In contrast, Small angle X-ray scattering (SAXS) profiles, while limited in resolution, contain information about conformational and compositional states of the system in solution. Moreover, SAXS profiles can be rapidly collected for a variety of experimental conditions, such as ligand-bound and unbound protein samples, different temperatures, or pH values. The challenge lies in the data interpretation since the profiles provide rotationally, conformationally, and compositionally averaged information about protein shape in solution.

We have developed a novel computational method, MultiFoXS<sup>1</sup> that simultaneously uncovers the set of structural states and their population weights for multiple input SAXS profiles. The input is a single atomic structure, a list of flexible residues, and one or more SAXS profile(s) for the protein. The method proceeds in two steps. In the first step, it samples the input structure by exploring the space of the  $\varphi$  and  $\psi$  main chain dihedral angles of the user-defined flexible residues with a Rapidly exploring Random Trees (RRTs) algorithm. In the second step a SAXS profile is calculated for each sampled conformation with FoXS, followed by a branch-and-bound enumeration of the multi-state models that are consistent with the SAXS profile. The method was benchmarked on over 30 cases with experimental SAXS profiles, including large multi-domain proteins and proteins with long disordered fragments. Moreover, comparison of conformations and their weights between the ligand-bound and unbound SAXS profiles can help in determining the allosteric mechanism. The applicability of the method extends beyond SAXS and it has been applied to datasets from cross-linking mass spectrometry, Electron Microscopy, and residual dipolar couplings.

In protein-protein docking with SAXS profile of the complex, in addition to conformational heterogeneity (multiple conformations of the same protein or complex), we also observe a compositional heterogeneity (complex and separate protein components in solution). I will present a protein-protein docking method that can account for both types of heterogeneity in SAXS data. The method relies on docking algorithms to sample possible interaction complexes<sup>2, 3</sup>, and on the SAXS scoring function<sup>4</sup> from MultiFoXS<sup>1</sup> to account for heterogeneity. The method was applied to antibody-antigen and additional complexes<sup>5</sup>.

References:

- 1. Schneidman-Duhovny, D. Hammel, M. Tainer, JA. Sali, A, Nucleic Acids Res 2016, 424-9.
- 2. Schneidman-Duhovny, D. Inbar, Y. Nussinov, R. Wolfson, HJ, Nucleic Acids Res 2005, 363-7.
- 3. Schneidman-Duhovny, D. Hammel, M. Sali, A, J Struct Biol 2011, 173, 461-71.
- 4. Schneidman-Duhovny, D. Hammel, M. Tainer, JA. Sali, A, Biophys J 2013, 105, 962-74.
- 5. Carter, L, et al., Biophys J 2015, 109, 793-805.

# Utilizing coarse-grained modelling and Monte Carlo simulations to understand and predict SAXS-data for intrinsically disordered proteins

Marie Skepö

Theoretical Chemistry, Lund University, Sweden Lund Institute of advanced Neutron and X-ray Science LINXS, Sweden

Computer simulations and modelling in combination with small angle X-ray scattering (SAXS) is a fruitful approach to achieve a molecular understanding of the underlying physics of a system. In this talk I will present how coarse-grained modelling in combination with Monte Carlo simulations can be used to understand and predict solution behavior of intrinsically disordered proteins (IDPs). The model used is based on the primitive model,<sup>1</sup> which is sometimes also referred to as the bead-necklace model. In this model each monomer corresponds to a bead with a given radius. In addition, each individual bead can be appointed a charge. The water is always treated as a dielectric continuum.

First an introduction to coarse-grained modelling will be given with emphasis on pros and consutilizing this approach. Thereafter I will show examples of how it can be applied to study dilute as well as concentrated IDP solutions, where the effect of electrostatic interactions and oligomerization will be highlighted.<sup>2-4</sup> Throughout the talk, comparisons will be made with SAXS-experiments, and when applicable, also atomistic molecular dynamics simulations and polymer theory.<sup>5-7</sup>

1. D. A. McQuarrie, *Statistical Mechanics*, 1st ed. (University Science Books, Sausalito, Califonia, 2000).

2. C. Cragnell, D. Durand, B. Cabane and M. Skepo, Proteins-Structure Function and Bioinformatics **84** (6), 777-791 (2016).

3. C. Cragnell, E. Rieloff and M. Skepo, Journal of molecular biology (2018).

4. E. Rieloff, M. D. Tully and M. Skepo, Journal of molecular biology (2018).

5. J. Henriques, L. Arleth, K. Lindorff-Larsen and M. Skepo, Journal of Molecular Biology (430), 2521-2539 (2018).

6. J. Henriques, C. Cragnell and M. Skepö, JCTC 11 (7), 3420-3431 (2015).

7. J. Henriques and M. Skepö, JCTC 12, 3407 (2016).

# Deciphering nanometer scale heterogeneity in ordered lipid phases using molecular simulation and small angle neutron scattering

Alexander Sodt

# NIH/NICHD

Our lab is interested in lipid correlations at the nanometer scale, that is, below the scale of disordered and ordered domains. This information will provide valuable information for how bilayers respond to stresses, especially curvature. We apply molecular dynamics and neutron scattering. As we see it, molecular dynamics supplies complete information but absolutely requires validation, while neutron scattering provides accurate but heavily filtered information for that validation. Although it depends on scattering contrast, the lateral correlation signal we seek is typically small compared with that of the bilayer form. Moreover, simulations use periodic boundary conditions, introducing spurious correlations. First I will present our lab's computational software that extracts useful lateral correlations to small angle neutron scattering experiments.

# What went wrong and right in trying to improve models for CASP-SAXS targets?

Yu Yamamori, Tsukasa Nakamura, Hiroyuki Fukuda, Yuko Tsuchiya and Kentaro Tomii

Small-angle X-ray scattering (SAXS) provides information about the native states of proteins in solution. Such information is regarded as invaluable to elucidate protein function(s). Using SAXS data, we strove to improve prediction models of target proteins in the "data-assisted" category of the 13th Community-Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP13).

For this category, we combined our prediction protocol based on template-based modeling and the selection procedure of protein models using SAXS data. Our prediction protocol to construct models for a target protein comprises the following steps: 1) prediction of intrinsically disordered region (IDR) of the target protein, 2) profile construction for the target sequence except IDR, 3) profile–profile alignment and scoring against profile libraries of template sequences performed by FORTE series, that are our own profile–profile comparison methods, and 4) model construction based on the alignments using MODELLER. Subsequently, we selected models with better values in terms of two metrics, i.e., volatility of ratio and  $\chi^2$ . For this selection procedure, we measured these two values using FoXS between the calculated SAXS scattering profile from a model and the experimental SAXS scattering profile.

Results show that, by considering SAXS data and also occasionally considering models from server predictions, we improved prediction models of protein complexes for some CASP-SAXS targets. However, contrary to our expectations, we worsened models of monomers. In some cases, the SAXS scattering profiles calculated from crystal structures differ from experimentally obtained profiles in terms of the two metrics. We will present examples of these problems and discuss issues that should be addressed to avoid them.

# Time-resolved solution X-ray scattering of photoactive proteins

#### Sebastian Westenhoff<sup>1</sup>

<sup>1</sup>Department of Chemistry and Molecular Biology, University of Gothenburg, westenho@chem.gu.se

Photoactive enzymes modulate their activity according to light conditions. We have studied a series of photosensitive proteins, with the goal to identify their structural changes when they are photoswitched. I will present these results. The structural changes range from refolding and quaternary rearrangements in phytochromes, a coiled-coil transition in a light-gated histidine kinase, and redox-driven rearrangements in cryptochromes.

In order to obtain structural information in solution, we have used time-resolved solution X-ray scattering. I provide an overview over the technique and discuss how it can be used to provide specific information on structural changes in solution. In particular, I will present how model-free structural parameters can be extracted and how more detailed structural models can be generated and validated against the difference scattering data. We have also investigated some of the proteins with NMR spectroscopy, infrared spectroscopy, and crystallography. I will discuss the scope and limitations of time-resolved X-ray solution scattering against this background.

## References

- O. Berntsson, R. P. Diensthuber, M. R. Panman, A. Björling, A. J. Hughes, L. Henry, S. Niebling, G. Newby, M. Liebi, A. Menzel, R. Henning, I. Kosheleva, A. Möglich, and S. Westenhoff, "Time-Resolved X-Ray Solution Scattering Reveals the Structural Photoactivation of a Light-Oxygen-Voltage Photoreceptor," Structure, vol. 25, no. 6, p. 933–938.e3, Jun. 2017.
- O. Berntsson, R. P. Diensthuber, M. R. Panman, A. Björling, E. Gustavsson, M. Hoernke, A. J. Hughes, L. Henry, S. Niebling, H. Takala, J. A. Ihalainen, G. Newby, S. Kerruth, J. Heberle, M. Liebi, A. Menzel, R. Henning, I. Kosheleva, A. Möglich, and S. Westenhoff, "Sequential conformational transitions and α-helical supercoiling regulate a sensor histidine kinase," Nat. Comm., vol. 8, no. 1, p. 284, Aug. 2017.
- A. Björling, O. Berntsson, H. Lehtivuori, H. Takala, A. J. Hughes, M. Panman, M. Hoernke, S. Niebling, L. Henry, R. Henning, I. Kosheleva, V. Chukharev, N. V. Tkachenko, A. Menzel, G. Newby, D. Khakhulin, M. Wulff, J. A. Ihalainen, and S. Westenhoff, "Structural photoactivation of a full-length bacterial phytochrome," Sci. Adv., vol. 2, no. 8, p. e1600920, Aug. 2016.
- 4. A. Björling, S. Niebling, M. Marcellini, D. Van Der Spoel, and S. Westenhoff, "Deciphering solution scattering data with experimentally guided molecular dynamics simulations," **J. Chem. Theory Comput.**, vol. 11, no. 2, pp. 780–787, Jan. 2015.
- H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J. A. Ihalainen, and S. Westenhoff, "Signal amplification and transduction in phytochrome photosensors," Nature, vol. 509, no. 7499, pp. 245–248, 2014.
- J. A. Ihalainen, E. Gustavsson, L. Schroeder, S. Donnini, H. Lehtivuori, L. Isaksson, C. Thöing, V. Modi, O. Berntsson, B. Stucki-Buchli, A. Liukkonen, H. Häkkänen, E. Kalenius, S. Westenhoff, and T. Kottke, "Chromophore-Protein Interplay during the Phytochrome Photocycle Revealed by Step-Scan FTIR Spectroscopy," JACS, vol. 140, no. 39, pp. 12396–12404, 2018.