

NANO-D

Algorithms for Modeling and Simulating Nanosystems



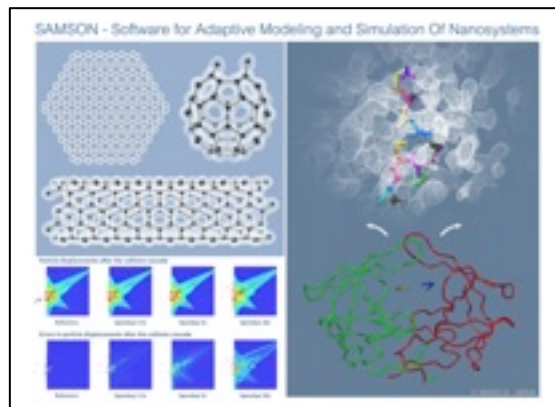
PhD Position Opening

Title : Computational developments for integrative structural biology : small-angle scattering using polynomial expansions

Application deadline : September 1st, 2018.

About research environment

The computational part of the project will be supervised by **Sergei Grudinin** from the NANO-D team at Inria Grenoble. The team develops novel multiscale, adaptive modeling and simulation methods, which automatically focus computational resources on the most relevant parts of the nanosystems under study. See more at <https://team.inria.fr/nano-d/>.



The experimental aspects of the project will be supervised by an alliance of experimentalists working at international research facilities (ILL/ESRF/EMBL). These are bioSAS experts **Anne Martel** and **Sylvain Prevost**, and **Trevor Forsyth**, and **Estelle Mossou** are the experts in 2D fiber diffraction. All of them are instrument scientists at ILL. **Martha Brennich** (EMBL) is an instrument scientist at BioSAXS and a specialist in small-angle diffraction of biological molecules.

PhD topic description : Introduction

While crystallography has been providing atomic-resolution structures of biomolecules for over half a century, the real challenge of today's biophysicists is to correlate molecules' structure and dynamics in solution with their function. Small-angle scattering (SAS) is the fundamental techniques for structural studies of biological systems in solution. Thanks to advances in instrumentation and data analysis software, Small-angle X-ray scattering (SAXS), complemented by other methods, is becoming very popular in structural biology. Over the years, a number of computational tools have been developed for the analysis of SAXS curves, calculation of theoretical profiles and low-resolution reconstruction of model shapes. Many efforts have been spent to reduce the running time of these tools without degrading the quality of their approximations, most prominently the ATSAS package developed at EMBL Hamburg. Particularly, the Crysol program calculates a model SAXS profile to test a structural hypothesis of a SAXS experiment. The number of Bio-SAXS publications exploded as a result of this effort. Comparatively, the lack of user-friendly analysis tools has hindered the development of Small Angle Neutron Scattering (SANS), more complex but providing more information.

Very recently, we developed a computational tool called Pepsi-SAXS that outperforms all the competitors for atomistic modelling in both speed (5 – 50 times faster on average) and accuracy. However, Pepsi-SAXS developments were very much restricted by the lack of human resources and, more importantly, by the lack of support and expertise from experimental teams. Also, SANS experiments are somehow more complex and bring different information than SAXS experiments because they offer the possibility to tune the scattering of the solvent to match the scattering of some compounds included in a particle to determine the structure of the compounds that remain visible. The current project is devoted to the novel computational developments for SAXS and SANS with applications to *integrative structural biology*.

PhD topic description : Challenges and tasks

The overall research topic of the current project is to extend the state-of-the-art computational methods for small-angle (both SAXS and SANS) scattering experiments and to provide the algorithmic foundation to the upcoming single-molecule experimental techniques. In particular, the practical aim of the project is the development of software tools with intuitive user-interaction feedback.

Mathematically, the project will rely on the very efficient representation of the scattering profiles based on polynomial expansions of the scattering amplitudes using **spherical harmonics**. Structural optimization will be performed using the **fast Fourier transform**-accelerated techniques and the **polynomial translation theorems** and large collective structural motions using **Normal Mode Analysis**.

Three main axes of the proposed project are : (i) Molecular flexibility and the ways of modelling it; (ii) Algorithmic developments for new experimental setups, particularly combining SAXS and SANS data, and 2D scattering; (iii) Practical software development with intuitive user interfaces. Below we detail the main research axes of the project.

More precisely, the planned developments include the following:

- Modelling thermal fluctuations (see review by Hub, 2018) using an analytical model of the scattering profile resulting from thermal molecular vibrations. This will be possible thanks to the (i) expression of normal modes in spherical coordinates expanded in the Spherical Harmonics basis, and (ii) computing molecular fluctuations in this basis analytically using the algebra for triple and quadruple overlap integrals of spherical harmonics.
- Flexible fitting of atomistic structures into small-angle scattering profiles. This includes combination the Pepsi-SAXS / Pepsi-SANS engines with the NOLB NMA method for the large-scale flexible structure optimization along the lowest-frequency normal modes. A working prototype exists, however, a proper analytical optimization technique with respect to the goodness of fit should be added. Also, the method needs to be extended to SANS data. Finally, the method needs to be made more robust and user-friendly, possibly, with a GUI inside the SAMSON modelling platform.
- Application of motion planning to adapt the sampling techniques for flexible unstructured regions of the molecules.
- In some cases it is impossible to represent the molecular structural heterogeneity continuously. This situation will require fitting the experimental SAS curves with an ensemble of structures. Therefore, we plan to extend the Pepsi-SAXS / Pepsi-SANS methods for the cases of multiple models, whose contributions will be adjusted with stochastic optimization techniques.
- We will consider the case when the molecular system can be represented with a set of rigid domains. Here we can compute scattering profiles for supra-molecular assemblies based on the precomputed scattering amplitudes of individual rigid bodies with subsequent rigid-body transforms applied to the amplitudes in the Fourier space.
- Fitting of atomic-resolution structures into multiple profiles obtained in different experimental configurations from the same sample. Indeed, the current modelling pipelines merge different experimental measurements into a single scattering profile. This approach has a significant drawback since the scaling factors for individual scattering components are generally unknown and can be erroneously guessed upon merging. A much more general, and also more computationally elegant approach would be to adjust the scaling factors of individual scattering profiles upon fitting atomistic models into multiple scattering curves. This approach would also have an obvious advantage of mixing different experimental information, e.g. SAXS and SANS experiments.
- Novel developments for 2D scattering based on the spherical harmonics and cylindrical harmonics representation. Many elongated or polymeric molecules can be aligned in space along the magnetic field. Thanks to this spatial alignment, their scattering profiles wouldn't be axially symmetric any longer and will contain much more information compared to profiles from spherically averaged particles. We plan to extend the current method for analytical cylindrical averaging of spherical harmonics with multiple advantages over the standard approaches.

Requirements and the role of the candidate

This is a very challenging project whose aim is to **change the modelling paradigm** in the small-angle community. It has many ambitious goals both on the algorithmic side and also considering software applications. The goals of the PhD student will be the following:

- Combination of the existing software tools that model molecular flexibility (deformations along the slowest normal modes, rigid-body perturbations of rigid domains, modeller of flexible loops using robotics techniques) into a unified library and link it with the Pepsi-SAXS library. Then, multiple tests should be performed, and probably novel benchmarks should be created.
- Adaptation of the developed methods for SAXS scattering and flexible fitting of molecular models into scattering profiles for SANS experiments. Computationally, it is an easy and straightforward task, however, numerous tests should be performed and very convenient software released.
- Formulation of the local refinement problem in the basis of Spherical Harmonics (SH). This is a challenging task, since no scattering methods that use analytical gradients in the SH basis have been developed.
- Developments for scattering of aligned objects and single-particle experiments
- Creation of practical interactive GUI interfaces with force feedback computed from the experimental scattering profiles. This is a challenging task, however, we have already done attempts for the integration of the NOLB NMA and Pepsi-SAXS methods inside the SAMSON modular simulation platform (see <https://www.samson-connect.net>). Again, the most difficult part will be to create convenient and interactive software tools.

The PhD candidate must be proficient in **computer science**, **computational physics** and **applied math**. He/she will have to communicate with the software developers at the side of the Inria partner, instrument scientists at the side of EMBL-ILL partners, and also with numerous users of the small-angle instruments at ILL and ESRF facilities.

The results of the PhD thesis will be novel algorithmic and software developments. We expect each tool (the algorithmic and the application part of the work) to be published in specialized high-impact international journals, such as Journal of Applied Crystallography, Journal of Structural Biology, or Nature Methods.

We are looking for creative, passionate and hard-working individuals with exceptional talent for computer science and mathematics. Excellent oral, written and interpersonal communication skills are essential (working language will be English – knowledge of French is a plus). Excellent knowledge of C++ and linear algebra are strictly required.

Salary and duration

This is a 3-year project. Gross salary is 1768,55€ / month.

About Grenoble

Grenoble is the capital city of the French Alps. Combining the urban life-style of southern France with a unique mountain setting, it is ideally situated for outdoor activities. The Grenoble area is today an important centre of industry and science (second largest in France). Dedicated to an ambitious policy in the arts, the city is host to numerous cultural institutions. With 60,000 students (including 6,000 foreign students), Grenoble is the third largest student area in France.



References:

- Spilotros, A. & Svergun, D. I. (2014). Encyclopedia of Analytical Chemistry, pp. 1–34.
- Graewert, M. A. & Svergun, D. I. (2013). Current opinion in structural biology, 23(5), 748–754.
- Petoukhov, et al. (2012). Journal of applied crystallography, 45(2), 342–350.
- Svergun, D., Barberato, C. & Koch, M. (1995). Journal of applied crystallography, 28(6), 768–773.
- S. Grudinin, M. Garkavenko & A Kazennov, Acta Cryst., 2017, D73, pp.449 – 464.
- A. Hoffmann and S. Grudinin, J. Chem. Theory Comput., vol. 13, no. 5, 2017.
- J. S. Hub, Current Opinion in Structural Biology, vol. 49. pp. 18–26, 2018.

Keywords:

structural bioinformatics; integrative structural biology; small-angle scattering; Fourier analysis; polynomial expansions; spherical harmonics