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.