Developing statistical, topological and geometrical techniques for ab-initio protein structure prediction:

*Primary Supervisor: Dr Christopher Prior (Durham: Mathematical Sciences)*

*Secondary Supervisors: Dr Matteo Degiacomi (Durham: Physics), Dr Rob Rambo (Diamond light source).*

Proteins are large biopolymers folding in characteristic three dimensional shapes, and are responsible for the vast majority of processes in life. As their specific biological task is determined by their shape, a vast amount of experimental efforts (e.g. X-ray crystallography) has been dedicated to elucidating their atomic structure. Proteins, however, are not rigid. It is now clear that their function is also critically dependent on their specific dynamics. Indeed, proteins are often found to have highly flexible sections, or can completely alter their shape to suit their environment. A particular class of proteins, called intrinsically disordered proteins (IDPs), has been found to be crucial to the survival of organisms in extreme environments. Despite their importance, their flexibility makes them difficult to study, and thus very little about them is known.

 A visualisation of the protein model and its fit to SAXS data. The initial prediction (a) does not fit the data (b). The model is optimized until the final model (c) fits well (d).

The focus of this project is to develop methods to characterize protein structure and dynamics using Small angle X-ray scattering (SAXS) data. This experimental technique gives information on the protein in its native state, and is suitable to study flexible molecules. The primary supervisor has recently developed a novel discrete curve model which allows the interpretation of SAXS data, by fitting a topologically realistic folded curve model to the data. This enables making predictions at a far higher resolution than previously thought possible (see Figure). This method was used to identify a novel structure whose shape could not be determined by other experimental techniques. Currently, the method is limited to making static predictions.

In this project, we will expand the model to incorporate dynamic complexity, so that it can provide information on intrinsically disordered proteins. This will require coupling our curve model with atomistic molecular dynamics simulations of the target proteins (under the guidance of Dr Matteo Degiacomi). Our goal is to develop our data interpretation method such that it can be regularly used on SAXS data obtained at the UK's national Synchrotron facility: the Diamond light source. This aspect will be overseen by Dr Rob Rambo who is head of the Soft matter village at this facility.

The student will learn to utilize the topological techniques underpinning the description of the protein curve model, and sophisticated statistical sampling methods used to model the dynamical evolution of the structure underpinning molecular simulations. This will provide the student with a cutting edge skill-set which will prepare them for a career in protein science, whilst simultaneously advancing a method for providing significant potential insight and impact in many fields of biology.