Unravelling the complexity of proteins

1 December 2014

Structures determined as part of the SPINE project, using high-throughput technologies developed and implemented in SPINE laboratories. Credit: SPINE project team

Knowledge of the three-dimensional structures of proteins is essential for understanding biological processes.

Structures help to explain molecular and biochemical functions, visualize details of macromolecular interactions, facilitate understanding of underlying biochemical mechanisms and define biological concepts.

The human genome and follow-up sequencing projects have revolutionized biology and medicine; structural genomic programmes have developed and applied structure-determination pipelines to a wide range of protein targets, facilitating the visualization of macromolecular interactions and the understanding of their molecular and biochemical functions.

A paper recently published by Mizianty et al. (2014). *Acta Cryst*. D70, 2781-2793; DOI: 10.1107/S1399004714019427 seeks to address the fundamental question of whether the three-dimensional structures of all proteins and all functional annotations can be determined using X-ray crystallography.

The researchers set out to perform the first large scale analysis of its kind covering all known complete proteomes (the sets of proteins thought to be expressed by an organism whose genome has been completely sequenced, as defined by the UniProt Consortium in 2012) and all functional and localization annotations available in the Gene Ontology for the corresponding proteins.

The Canadian and US team show that current X-ray crystallographic knowhow combined with homology modeling can provide structures for 25% of modelling families (protein clusters for which structural models can be obtained through homology modelling), with at least one structural model produced for each Gene Ontology functional annotation. The coverage varies between super-kingdoms, with 19% for eukaryotes, 35% for bacteria and 49% for archaea, and with those of viruses following the coverage values of their hosts. It is shown that the crystallization propensities of proteomes from the taxonomic super kingdoms are distinct. The use of knowledge-based target selection is shown to substantially increase the ability to produce X-ray structures.

Talking to the IUCr Mizianty commented "We believe our method has helped to advance our understanding of the coverage by X-ray structures of proteins and complete proteomes on a global scale".

Provided by International Union of Crystallography