Contribution ID: **64** Type: **Oral Presentation**

Exploring the conformational space of proteins by enumeration in the frame of the Distance Geometry Problem

Wednesday, 27 March 2024 15:00 (30 minutes)

The continuous development of the methods for the protein structure prediction was taking advantage from the precious experimental information obtained by structural biology as well as by sequencing of multiple organisms. Indeed, the general developed pipeline is based on determining conformations of protein fragments, and then using multiple sequence alignments to obtain long-range distances allowing to get predicted protein conformations. The introduction of the deep learning techniques has recently permitted a important jump in the results obtained in this framework, as illustrated by the success of AlphaFold2, RoseTTAFold and ESMFold. Nevertheless, in some cases, long-range restraints cannot be obtained, as for intrinsically disordered proteins or regions (IDP/IDR), or in the case of orphan proteins for which not enough statistical signal can be obtained by sequence alignment. Here, we propose to investigate another point of view, in which local information would be mainly used to determine the protein folding, this local conformational information being extracted directly from the primary sequence. Two major obstacles arise from: (i) the variability in protein stereochemistry, inducing a drift of the protein backbone when the residues are successively added; (ii) the size of the protein conformational space described by local conformations, as it would not be reduced by the long-range proximity information. In the framework of the Distance Geometry Problem (DGP), the Branch-and-Prune (BP) approach, based on a graph description of proteins, brings an answer to the problem of the size of the conformational space by performing a systematic enumeration of all protein conformations satisfying a given set of geometric constraints. Applications of the BP approach will thus be presented for the reconstruction of folded structures of proteins, with propositions for dealing with the variability of stereochemistry. The efficiency of the BP approach will be also demonstrated on the cases of disordered or flexible proteins, in particular on the Small EDRK-rich factor 1.

Primary authors: Dr CHANG, Chi-Fon; Dr HUANG, Shu-Yu; LIN, Jung-Hsin; Dr SHIH, Orion; Dr MALLIAVIN, Therese (CNRS)

Presenter: Dr MALLIAVIN, Therese (CNRS)

Session Classification: Health & Life Science Applications

Track Classification: Track 2: Health & Life Sciences (including Pandemic Preparedness Applications)