MS11 Opportunities from combining structural biology and fold prediction

MS11-01

How predicted protein models help to illuminate the full protein universe

J. Pereira [1], J. Durairaj [1], L. Pantolini [1], G. Studer [1], X. Robin [1], A. Waterhouse [1], S. Bienert [1], G. Tauriello [1], T. Schwede [1]

[1]Biozentrum and SIB Swiss Institute of Bioinformatics, University of Basel - Basel ***(Switzerland)***

Abstract

Protein structural biology is going through a major revolution, with deep learning-based methods enabling the prediction of three-dimensional protein structures and complexes at an unprecedented level of accuracy. Together with the availability of large amounts of computational power, these methods facilitate the large-scale modelling of millions of proteins across the tree of life and thus complement the experimental structural information deposited in the Protein Data Bank. For example, there are currently about 1 million protein structure models in the AlphaFold database, and this number is expected to increase 150-fold by the summer of 2022 by including models for all proteins in UniRef90. This means that structural biologists will have structural models for virtually all catalogued proteins at their disposal.

But what does this mean exactly? Is experimental structural biology obsolete now, or is this a turning point and an accelerator for further developments? In this talk, I will discuss what the implications of such a structural coverage are and how large-scale protein structure prediction accelerates our understanding of natural proteins.