Efficient and accurate prediction of protein structures and interactions using RoseTTAFold

Minkyung Baek[1]

[1]University of Washington

minkbaek@gmail.com

An accurate protein structure prediction from its amino acid sequence is a longstanding challenge in computational biology. Considerable progress has recently been made by leveraging genetic information through deep learning-based methods. In this talk, I'll present a three-track attention-based neural network named RoseTTAFold. In this model, information at the 1D sequence level, the 2D distance map level, and the 3D coordinate level are successively transformed and integrated to generate accurate protein structures. RoseTTAFold also enables rapid generation of accurate protein-protein complex models from sequence information alone, reducing the computational cost of traditional approaches which require modeling of individual subunits followed by docking. We further use a combination of RoseTTAFold and AlphaFold2 to screen protein pairwise interactions of yeast proteins, and build structure models for 106 previously unidentified assemblies and 806 that have not been structurally characterized. These complexes play roles in almost all key processes in eukaryotic cells and provide broad insights into biological functions which can be utilized to design drugs.