COMPUTATIONAL STRUCTURAL BIOLOGY

Predicting protein structure from cryo-EM data*Proc. Natl Acad. Sci. USA* **118**, e2017525118 (2021)

Proteins play an important role in many crucial biological processes, and determining their structure is a critical step to understand their functionality: the structure of a protein dictates whether and how it can interact with other molecules. Researchers can then use this structural information, for instance, to assist in the development of new drugs and vaccines. Predicting protein structure is, however, a challenging problem and it has been an active research topic for many years.

In a recently published work, Dong Si and colleagues take advantage of cryoelectron microscopy (cryo-EM) data to predict the structure of proteins. Cryo-EM, a 2017 Nobel prize-awarded technology, has gained popularity for capturing 3D maps of macromolecules at an incredible near-atomic resolution. The authors propose a tool called **DeepTracer**, which takes as input a protein's cryo-EM map and amino acid sequence, and outputs its all-atom structure using a tailored deep learning framework. Different from other cryo-EM model determination methods, DeepTracer has the advantage of performing multichain prediction, requiring no manual processing steps, and achieving more accurate results.

The proposed method relies on a convolutional neural network that consists of four U-Nets, each of them designed to

predict a specific structural aspect: the locations of amino acids, the location of the backbone, the secondary structure elements, and the amino acid types.

A series of fully automated post-processing steps are then applied to the outputs of these U-Nets to ultimately return the predicted final structure. When compared to state-of-the-art methods (for example, Phenix, Rosetta and MAINMAST), the authors demonstrated that DeepTracer has a better accuracy: for instance, when compared to Phenix using a set of coronavirus-related data, it improved coverage (the proportion of residues that have a matching interpreted residue) in over 30%, and it decreased the root-mean-square deviation value by more than 0.40 Å. In addition, the tool was shown to be computationally efficient when running on a graphics processing unit (GPU): as an example, the tool traced a cryo-EM map containing approximately 60,000 residues within two hours. Overall, DeepTracer is an exciting new method for protein prediction that will certainly help move the field forward.

Fernando Chirigati

Published online: 22 February 2021

<https://doi.org/10.1038/s43588-021-00035-w>