

## DeepMainmast: Protein Structure Modeling for Cryo-EM Using Deep Learning

Genki Terashi<sup>1</sup>, Xiao Wang<sup>2</sup>, Devashish Prasad<sup>2</sup>, and Daisuke Kihara<sup>1,2</sup>

1. Purdue University, Biological Sciences

2. Purdue University, Computer Sciences

In recent years, there has been a significant increase in the number of protein and nucleotide structures modeled using cryo-electron microscopy (cryo-EM) maps. Despite the steady improvement in EM map resolution over time, modelers often encounter challenges when modeling biomolecular structures. Specifically, protein structure modeling from a cryo-EM map becomes particularly difficult when the resolution is lower than about 3.5 Å. To tackle this issue, we have developed an integrated protein structure modeling protocol, named DeepMainMast.

This protocol employs a new de novo protein main-chain tracing method that uses deep learning to identify the positions of  $C\alpha$  atoms and the types of amino acids. The core process of DeepMainmast employs an effective main-chain tracing approach known as the Vehicle Routing Problem Solver and Constraint Problem Solver. Additionally, the protocol can accurately assign chain identity to the structure models of homo-multimers. To further enhance the performance of the protocol, we also incorporate AlphaFold2 models when applicable. These models provide valuable information that contributes to enhancing the accuracy of the resulting protein structures. Overall, DeepMainmast is a powerful tool that can help researchers overcome the challenges of protein structure modeling from cryo-EM maps. Our benchmarking results demonstrate that DeepMainmast substantially outperforms existing methods on the benchmark dataset. Compared to AlphaFold2, DeepMainmast achieves higher accuracy on a larger number of maps within the dataset, which consists of 178 high-resolution maps. The code is available at <https://github.com/kiharalab/DeepMainMast>.