Distance-AF: Flexible protein structure prediction with distance constraints using AlphaFold2

Yuanyuan Zhang[1], Zicong Zhang[1], Yuki Kagaya[2] and Daisuke Kihara[1,2,3*]

[1] Department of Computer Science, Purdue University, West Lafayette, Indiana, 47907, USA

[2] Department of Biological Sciences, Purdue University, West Lafayette, Indiana, 47907, USA

[3] Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN, 47907, USA

The three-dimensional structure of a protein plays a fundamental role in determining its function and has an essential impact on understanding biological processes. Despite significant progress in protein structure prediction, such as the AlphaFold2, challenges remain on those hard targets that Alphafold2 does not often perform well due to the complex folding of protein and a large number of possible conformations. Here we present a modified version of the AlphaFold2, called Distance-AF, which aims to improve the performance of AlphaFold2 by including distance constraints as input information. Distance-AF uses AlphaFold2's predicted structure as a starting point and incorporates distance constraints between amino acids to adjust folding of the protein structure until it meets the constraints. Distance-AF can correct the domain orientation on challenging targets, leading to much more accurate structures with much lower Root-Mean-Standard-Deviation (RMSD). The ability of Distance-AF to reposition parts of a protein's structure allows for a more accurate prediction of protein structure, particularly in cases where AlphaFold2 predicts less accurately on fitting protein structures into Cryo-electron microscopy maps. This could have significant implications for drug discovery and other areas of biological research that rely on accurate protein structure predictions.