PROTEIN FOLDING AND PEPTIDE BINDING PREDICTION WITH ALPHAFOLD

Liwei Chang and Alberto Perez

Department of Chemistry, University of Florida, Gainesville, FL, USA

Machine learning approaches have recently revolutionized structural biology by predicting 3D structures of proteins at atomistic level from their sequence with high confidence and accuracy. The community is pushing the limits of interpretability and application of these algorithms beyond their original objective. Here, we present our studies on what deep neural network models for protein structure prediction such as AlphaFold can shed light on two fundamental biological processes: (1) protein folding and (2) protein-peptide binding. We first demonstrate our studies on deciphering the folding mechanism of protein G, L and their mutants, two pair of proteins with low sequence similarity that fold into the same topology. By using large-scale molecular dynamics simulations with advanced enhanced sampling methods and Markov state modeling analysis, we observed how the sequence difference impacts folding pathways and rate in unprecedented detail. Then, we show a novel fragment decomposition approach using AlphaFold to identify preferences for secondary structure element combinations that follow the order of events observed in the folding pathways. Standing upon the highly sensitive mapping of AlphaFold between the sequence and structure, we discovered how to directly use the model to identify high affinity peptide binders using a competitive binding assay. For systems in which the individual structures of the peptides are well predicted, the assay captures the higher affinity binder in the bound state, and the other peptide in the unbound form. The speed and robustness of the method will make it readily applicable in screening libraries of peptide sequences to prioritize for detailed experimental characterization.