Title

Enhancing Intrinsically Disordered Region Conformational Ensembles through MSA Tuning and Energy-Based Clustering to Guide AlphaFold2 Predictions

Abstract

Intrinsically disordered regions and proteins play crucial roles in a wide range of biological functions and processes, yet their structural characterization remains challenging due to their dynamic nature. While AlphaFold2 has made significant advances in protein structure prediction, it often struggles to capture the conformational heterogeneity of IDRs and IDPs. Previous work has demonstrated the systematic identification of conditionally folded IDRs using AF2, as well as the prediction of multiple conformations through sequence clustering based AF2 approaches. Building on these efforts, this proposal aims to further enhance the modeling of conformational ensembles for IDRs and IDPs by leveraging MSA tuning and energy-based clustering techniques. By integrating Rosetta energy functions with disorder propensity metrics, we propose a novel framework that segments MSA sequences into distinct clusters, each potentially representing a different conformational state. This approach not only addresses limitations of AlphaFold2 in modeling IDRs and IDPs, such as capturing Boltzmann-weighted conformational states, but also provides a more interpretable mapping between sequence variation and structural ensembles. The proposed methodology holds promise for advancing our understanding of IDR and IDP function and for informing drug discovery efforts targeting these elusive protein regions.

Draft Figure

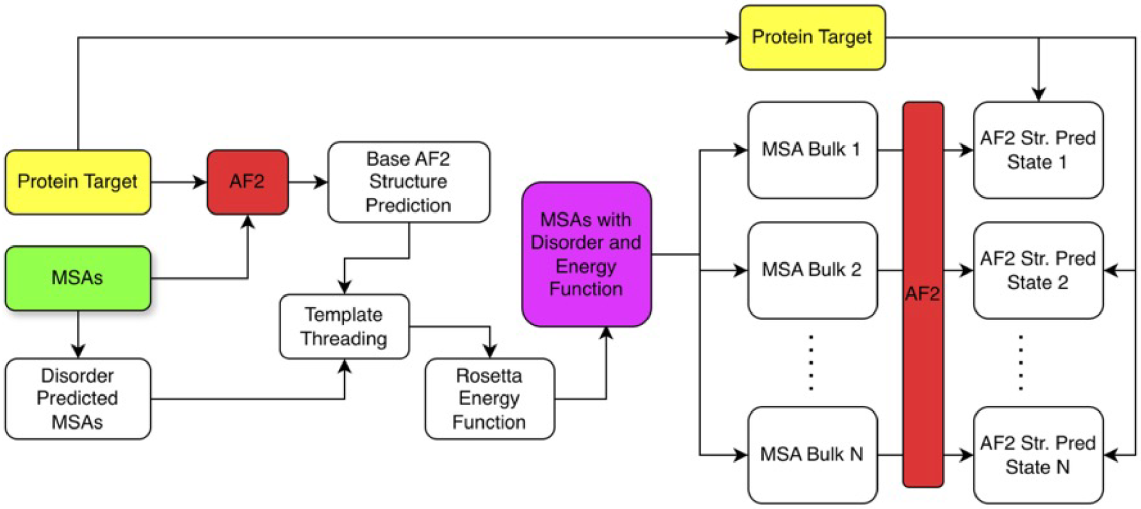

Figure 1: Proposed Method Protocol Overview