[BIOINFO 2025]

Predicting binding affinity between transcription factors and ligands using a graph-based model

Haechan Sung[1], Mun Su Kwon[1], Taeyoung Lee[2], and Hyun Uk Kim[1,2*]

[1] *Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141, Republic of Korea*

[2] *Graduate School of Engineering Biology, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141, Republic of Korea*

** Corresponding author: ehukim@kaist.ac.kr*

Accurate prediction of protein–ligand binding affinity is crucial for metabolic engineering and the construction of synthetic genetic circuits. Transcription factors, as key regulators of gene expression, require precise affinity modeling to enable the integration of novel sensors responsive to new inducers. Here, we present an AI-based framework that combines protein sequence embeddings with ligand graph representations. This design enables the model to exploit sequence-derived biological information alongside structural and chemical features of ligands. Using PDBbind complexes, we generated embeddings through protein language models and ligand graphs from molecular pretraining. A graph neural network captures intramolecular features within proteins and ligands, while cross-attention highlights intermolecular interactions between them. The combined representations are then processed by a multilayer perceptron to predict binding affinity. This method provides a versatile tool to accelerate transcription factor design and integration in synthetic biology.