Learning Antibody-Specific Epitopes
with Domain-Specialized Language Model Embeddings

Yujeong Oh[1], and Minkyung Baek[1,2*]

[1] *School of Biological Sciences, Seoul National University*
[2] *Interdisciplinary Program in Artificial Intelligence, Seoul National University*
*Corresponding author: minkbaek@snu.ac.kr

Accurate epitope prediction requires modeling antibody specificity in addition to antigen context. Existing protein language model-based approaches largely treat epitopes as antigen-only properties, limiting antibody awareness. This work introduces a graph-based, residue-level framework that conditions predictions on paired antibody–antigen inputs by fusing domain-specialized and general-purpose embeddings. Antibody sequences are encoded with Ablang2, trained on clustered antibody repertoires, while antigens are encoded with ESM-2. The resulting representations are lifted to graphs and integrated via cross-attention to capture interaction-aware features. Training data are derived from SAbDab with contact-based residue labels under leakage-controlled, antigen-cluster splits. The evaluation design comprises ablations of the antibody encoder and attention modules and comparisons to antigen-only baselines under identical graph heads, targeting robust, antibody-aware epitope scoring.