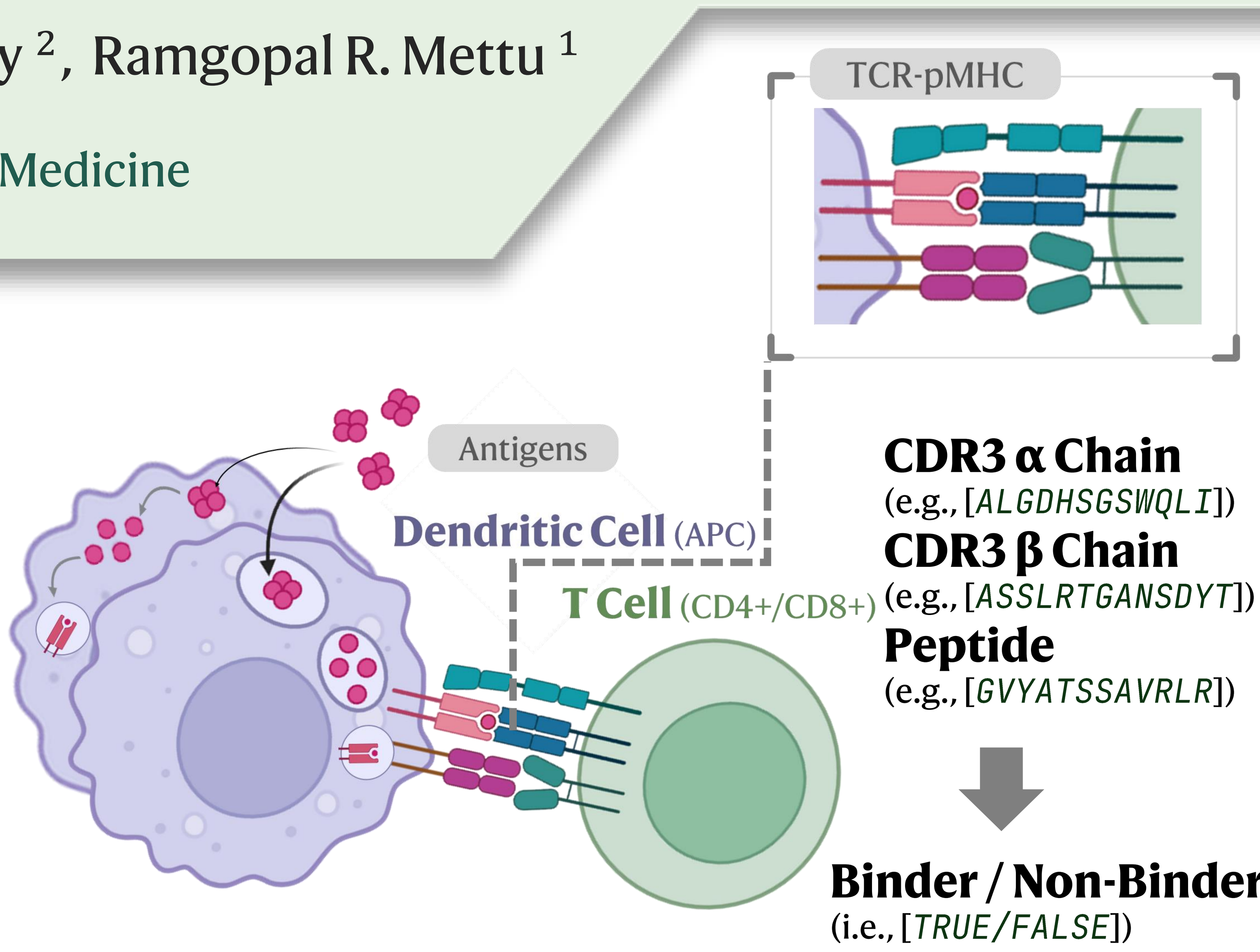
TCR-EML: Explainable Model Layers for TCR-pMHC Prediction

Jiarui Li [1], Zixiang Yin [1], Zhengming Ding [1], Samuel J Landry [2], Ramgopal R. Mettu [1]
[1] Department of Computer Science, Tulane University
[2] Biochemistry and Molecular Biology, Tulane University School of Medicine

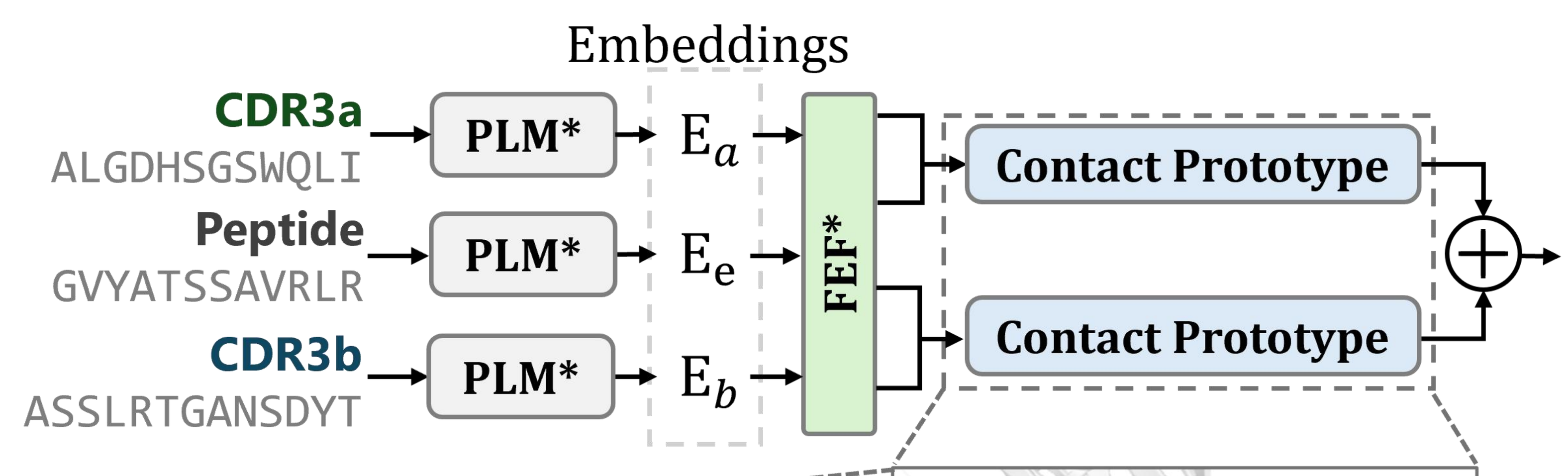
TULANE UNIVERSITY

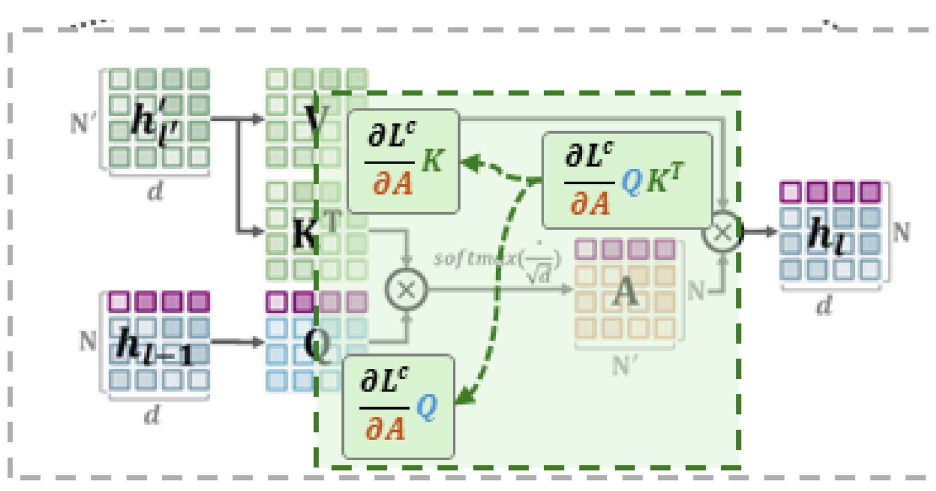


T cell receptor (TCR) recognition of peptide-MHC (pMHC) complexes is central to adaptive immunity and important for vaccine design, cancer immunotherapy, and autoimmune disease research. Although recent machine learning methods improve TCR-pMHC binding prediction, the most accurate models are black-box models. Post-hoc explanations provide limited insight and do not explicitly capture biochemical mechanisms such as binding regions. We propose **TCR-EML**, an explain-by-design model layer that uses protein language models to enable interpretable TCR-pMHC binding prediction.

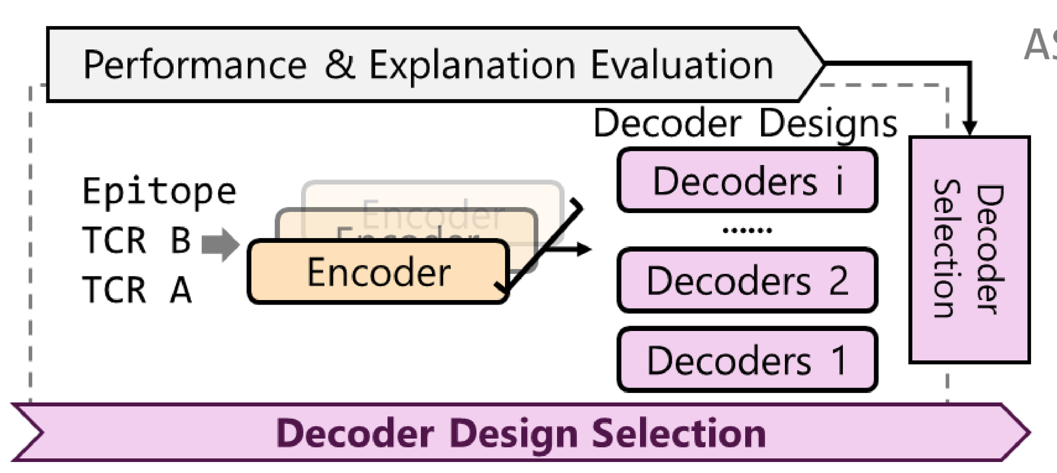
Our design consists of two components:

1. Feature Enhancement and Fusion (FEF) utilizing a rationally designed cross-attention design.
2. Contact prototype layers.

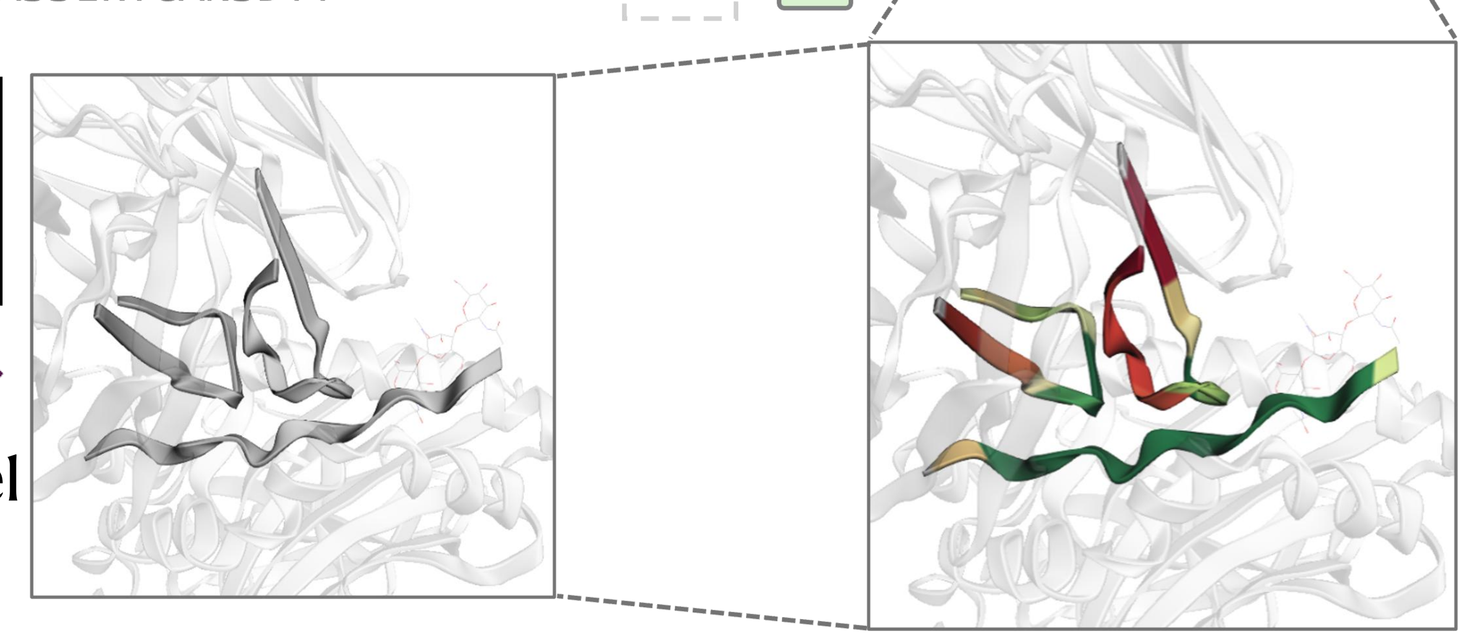


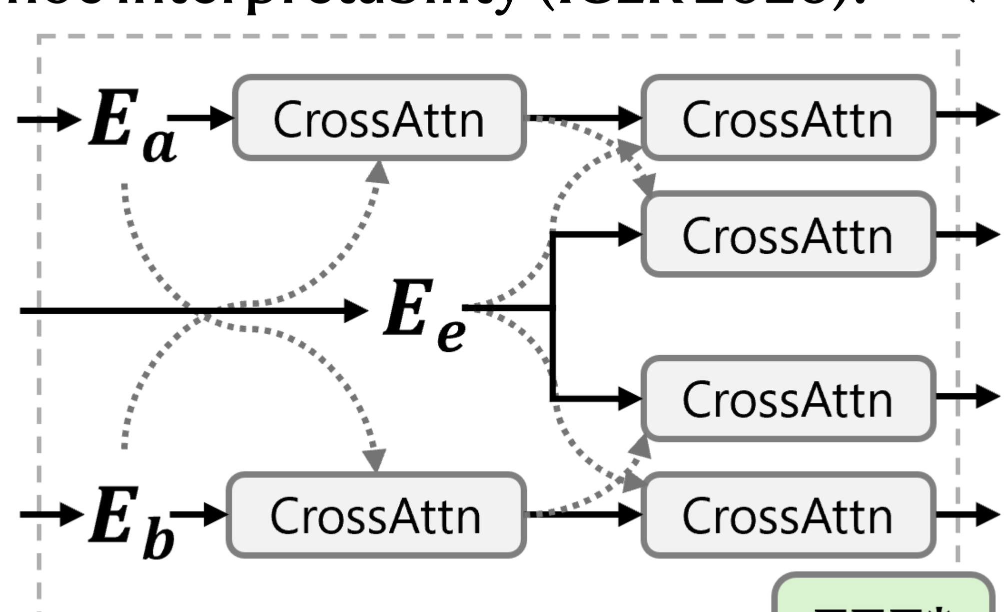


Quantifying Cross-Attention Interaction (QCAI) is used for post-hoc interpretability (ICLR 2026).

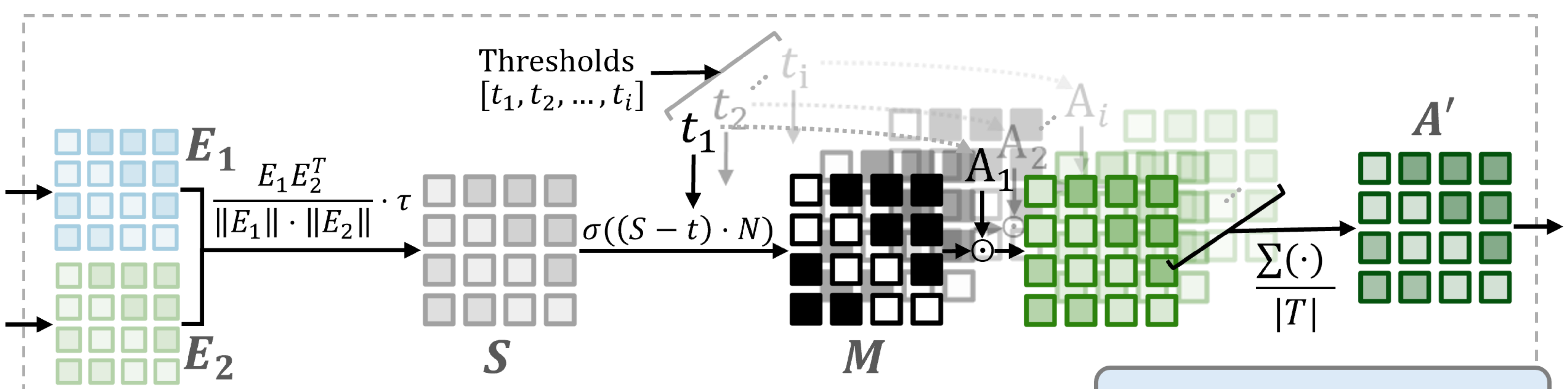


We use QCAI for cross-attention model design to design the FEF component (ACM BCB 2025).



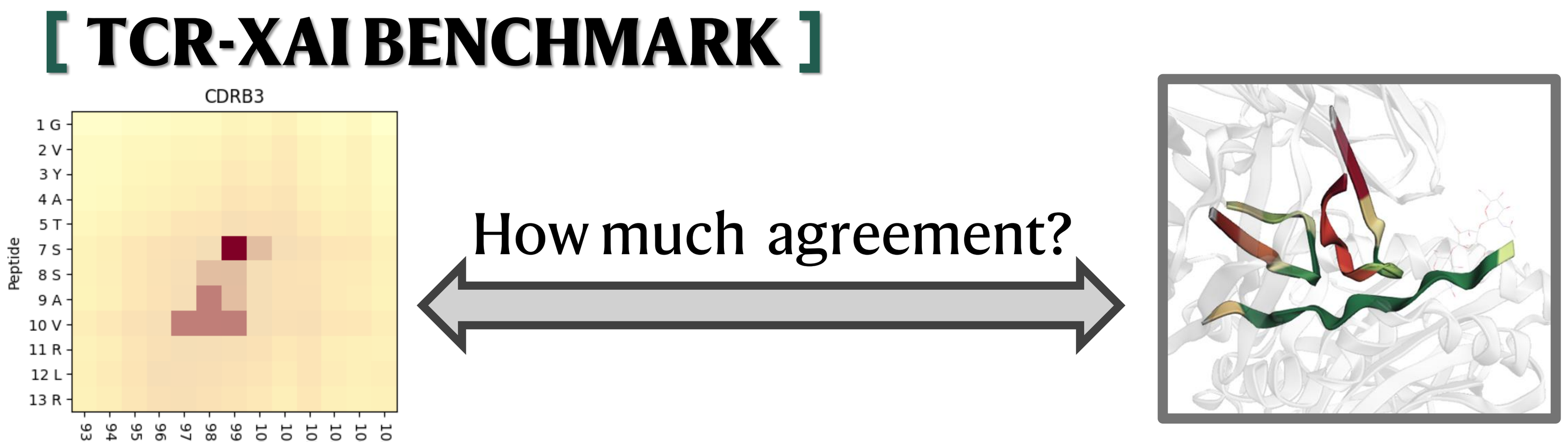


*Feature Enhancement & Fusion



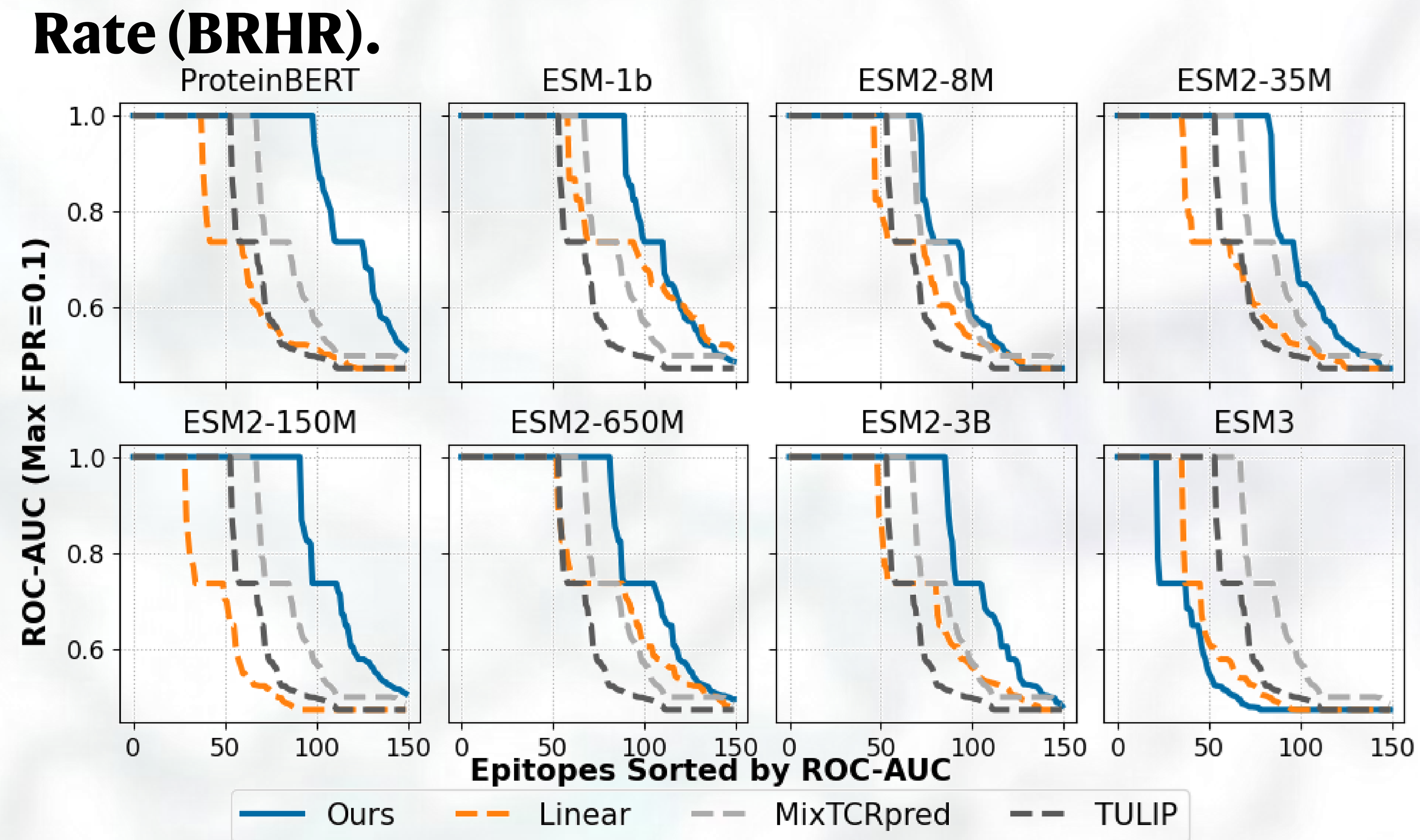
Contact Prototype

[TCR-XAI BENCHMARK]



Contact Map from TCR-EML

We have compiled a ground-truth TCR-pMHC dataset of 274 high resolution X-ray structures from STCRDab and TCR3D 2.0. We demonstrate the interpretability enabled by QCAI of TULIP predictions on this dataset using a quantitative metric we call **Binding Region Hit Rate (BRHR).**

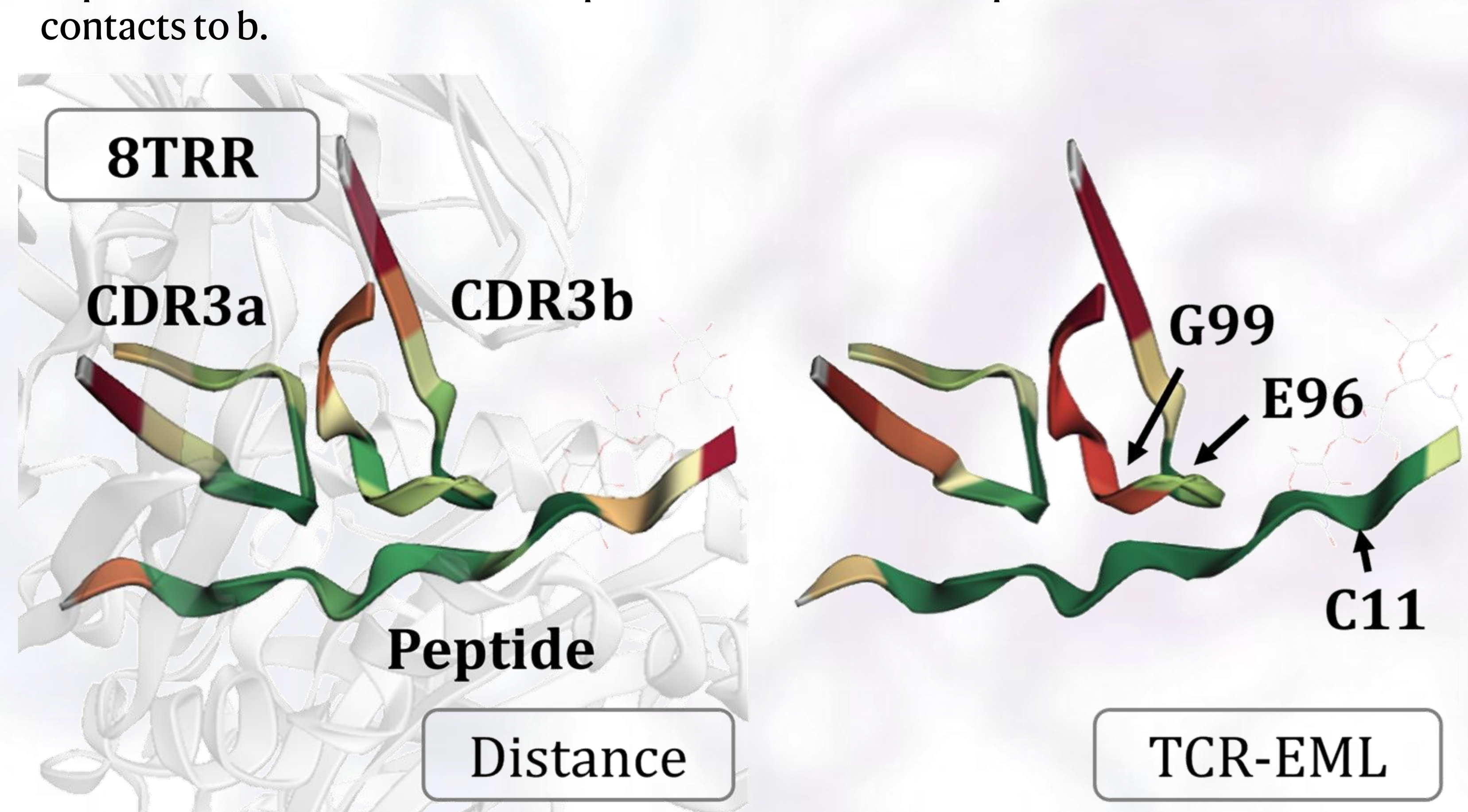


ROC-AUC Top-k	ProteinBERT Linear	ProteinBERT Ours	ESM-1b Linear	ESM-1b Ours	ESM2-150M Linear	ESM2-150M Ours	MixTCRpred	TULIP
100	0.772	**0.999**	0.900	**0.982**	0.713	**0.985**	0.906	0.821
150	0.675	**0.895**	0.795	**0.854**	0.633	**0.860**	0.773	0.706
200	0.625	**0.792**	0.716	**0.759**	0.593	**0.765**	0.698	0.648

ROC-AUC comparison at Top-100, Top-150, and Top-200 peptides. Columns report results for PLM backbones (ESM-1b, ESM-2, and ProteinBERT) with either a linear classifier or our method, with MixTCRpred and TULIP included as reference baselines. Our methods significantly outperformed all other methods with all PLM backbones.

$a \to b$	ProteinBERT	ESM-1b	ESM2-150M	MixTCRpred	TULIP
Peptide→CDR3a	0.839	0.842	0.897	0.718	0.702
Peptide→CDR3b	0.842	0.877	0.850	0.723	0.634
CDR3a→Peptide	0.736	0.812	0.773	0.795	0.798
CDR3b→Peptide	0.790	0.813	0.792	0.675	0.646

Binding Region Hit Rate (t=0.25) across different PLM backbones for Peptide → CDR3a, Peptide → CDR3b, CDR3a → Peptide, and CDR3b → Peptide, where a→b denotes a contacts to b.



A case study on HLA-DR4-bound vimentin-64cit59-71.