

## Supplementary Materials for PiTE

Table 1. Summarized representation size of different sequence encoder and trainable parameters of TCR-epitope binding affinity prediction models. We show number of total trainable parameters in the prediction model and trainable parameters in encoder layers in parentheses.

Sequence Encoder Structure	Representation Size	Trainable Parameters (in encoders)
Average Pooling (Baseline)	$1 \times 1024$	3,149,825 (0)
Transformer	$1 \times 1024$	20,082,753 (16,932,928)
BiLSTMs	$1 \times 256$	1,364,993 (574,464)
CNNs	$1 \times 1024$	11,430,657 (8,280,832)

Table 2. We conduct an ablation study for the number of heads in the multi-head attention layers and the size of binary classification layers. We report average AUC scores under different parameter settings for three runs in the below table. The bold row is the results showed in our main manuscript.

Nums of heads	Nums of neurons	AUC in TCR split	AUC in epitope split
1	1,024	97.64	89.68
<b>2</b>	<b>1,024</b>	<b>97.48</b>	<b>89.82</b>
4	1,024	97.40	89.24
8	1,024	97.34	88.04
16	1,024	97.58	89.43
2	512	97.85	89.49
2	2,048	97.50	89.69