A118 · Scaling CAR-T Targeting of HLA-presented Intracellular Antigens with AI-Driven Experimentation

Presenter: Elizabeth B. Wood — JURA Bio

Authors: Xiao-Bing Cui, Kerry Dobbs, Andrei Slabodkin, Alan N. Amin, Mattia G. Gollub, Kristina Gurung, Eli N. Weinstein, Elizabeth B. Wood

Session: Day 2 · Fri Oct 2, 2026 · 4:15–5:15 PM

The majority of cancer-driving proteins are intracellular, and so can only be recognized by immunotherapies through short peptide fragments displayed on the human leukocyte antigen (HLA). We developed a lab-in-the-loop system to learn the rules of scFv-pHLA protein-protein interactions on human cells. We use generative models of proteins and of screens to design, synthesize, and test interactions between tens of millions of scFvs and 100 pHLAs in a single multiplexed experiment, producing large scale training datasets. Transformers trained on the data predict unseen interactions and exhibit reliable scaling laws, with steady model improvements against seen and unseen pHLAs as experiments continue. Overall, AI-driven experimentation enables models to systematically learn to design TCR mimicking antibodies.

Scaling CAR-T Targeting of HLA-presented Intracellular Antigens with AI-Driven Experimentation

Xiao-Bing Cui[*1], Kerry Dobbs[*1], Andrei Slabodkin[*1], Alan N. Amin[*1], Mattia G. Gollub[*1], Kristina Gurung[*1], Eli N. Weinstein[†1,2], Elizabeth B. Wood[†1]

August 14, 2026

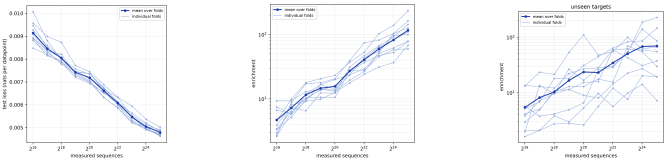
Abstract

The majority of cancer-driving proteins are intracellular, and so can only be recognized by immunotherapies through short peptide fragments displayed on the human leukocyte antigen (HLA). We developed a lab-in-the-loop system to learn the rules of scFv-pHLA protein-protein interactions on human cells. We use generative models of proteins and of screens to design, synthesize, and test interactions between tens of millions of scFvs and 100 pHLAs in a single multiplexed experiment, producing large scale training datasets. Transformers trained on the data predict unseen interactions and exhibit reliable scaling laws, with steady model improvements against seen and unseen pHLAs as experiments continue. Overall, AI-driven experimentation enables models to systematically learn to design TCR mimicking antibodies.

Most cancer driver proteins are intracellular and hence hidden from the immune system, except for short peptide fragments presented on the cell surface in complex with HLA molecules (pHLA/pMHC). TCR mimicking (TCRm) antibodies that recognize pHLAs have substantial clinical potential, but discovering and designing them poses a major challenge. Peptides are in a dynamic complex with the HLA molecule, whose peptide-binding groove presents a flat surface with limited opportunities for high-affinity antibody binding.

We developed a machine learning-guided wetlab system that can generate data to train models of interactions between pHLAs and human antibody scFvs displayed on human cells. We reasoned that by using generative models to plan and steer the experiments and data collection, large scale screens could produce large and informative datasets for training complex sequence-to-function models. We found the approach enables training of large transformer models for TCRm discovery and design, which exhibit continual and predictable learning as experimental data collection continues.

[0]* equal contribution. 1 JURA Bio, Inc. 2 Technical University of Denmark. † equal contribution.



(a) (b) (c)

Figure 1: (a) Scaling law. Sequence-activity model performance improves smoothly according to a power law relation as dataset size increases. (b,c) The model's ability to enrich for specific TCRms against seen (b) and unseen (c) pHLA targets improves smoothly, as training data increases.