Deep Learning Methods for Structure Prediction and De Novo Design of T and B Cell Receptors Targeting Peptide-MHC Neoantigens

Amir Motmaen

A dissertation

submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

David Baker, Chair

Jorge A Marchand

Cole A. DeForest

Program Authorized to Offer Degree:

Bioengineering

University of Washington

Abstract

Deep Learning Methods for Structure Prediction and De Novo Design of T and B Cell Receptors Targeting Peptide-MHC Neoantigens

Amir Motmaen

Chair of the Supervisory Committee:

David Baker

Biochemistry

Adaptive immune recognition depends on the specific interaction between peptides presented by major histocompatibility complexes (MHCs) and the receptors that survey them. Advances in protein structure prediction and design now allow us to computationally model these interactions with unprecedented fidelity and engineer them with remarkably higher success rates. In this dissertation, we develop two complementary structure-based deep learning approaches: one for predicting peptide–MHC specificity by fine-tuning structure prediction networks directly on binding data, and another for designing de novo T cell receptors and TCR-mimic antibodies that recognize peptide–MHC targets with high accuracy. Together, these methods illustrate how integrating structurally-informed deep learning protein design and structure prediction frameworks enable both robust generalization and precise molecular engineering. These tools lay the foundation for programmable, therapeutic recognition of diseased cells and expand our understanding of adaptive immune specificity.

Acknowledgements:

I am deeply grateful to my advisors and mentors for their guidance throughout this work. I thank Dr. David Baker for his vision, support, and insistence on ambitious ideas, and Dr. Philip Bradley for his insight, patience, and dedicated mentorship. I like to express my gratitude to members of my supervisory and reading committees who generously provided me with guidance and advice throughout my PhD. I am equally indebted to the many collaborators who contributed intellectually and experimentally to this dissertation—including colleagues who helped develop the computational and the experimental validation assays in my thesis.

I want to acknowledge the members of the Institute for Protein Design and the UW Bioengineering community for providing an inspiring environment filled with creativity, rigor, and camaraderie. Special thanks to friends and labmates who made even the most difficult moments enjoyable, and to those who offered thoughtful discussions that shaped both the science and its interpretation.

Finally, I thank my family for their continuous love and encouragement, and for supporting my pursuit of scientific curiosity across countries, institutions, and years. This work would not have been possible without them.

Table of Contents

Chapter 1

Peptide binding specificity prediction using fine-tuned protein structure prediction networks

As originally published in Motmaen, A., Dauparas, J., Baek, M., Abedi, M.H., Baker, D. and Bradley, P., 2023. Peptide-binding specificity prediction using fine-tuned protein structure prediction networks. *Proceedings of the National Academy of Sciences*, *120*(9), p.e2216697120.

Abstract

Peptide binding proteins play key roles in biology, and predicting their binding specificity is a long-standing challenge. While considerable protein structural information is available, the most successful current methods use sequence information alone, in part because it has been a challenge to model the subtle structural changes accompanying sequence substitutions. Protein structure prediction networks such as AlphaFold model sequence-structure relationships very accurately, and we reasoned that if it were possible to specifically train such networks on binding data, more generalizable models could be created. We show that placing a classifier on top of the AlphaFold network and fine-tuning the combined network parameters for both classification and structure prediction accuracy leads to a model with strong generalizable performance on a wide range of Class I and Class II peptide-MHC interactions that approaches the overall performance of the state-of-the-art NetMHCpan sequence-based method. The peptide-MHC optimized model shows excellent performance in distinguishing binding and non-binding peptides to SH3 and PDZ domains. This ability to generalize well beyond the training set far exceeds that of sequence only models, and should be particularly powerful for systems where less experimental data is available.

Introduction

Sequence-based methods utilize large sets of experimentally validated binding and non-binding peptides to assemble position specific weight matrices or more sophisticated neural networks with several layers to discriminate binder from non-binder peptides (1–7). Methods such as NetMHCpan are the current state-of-the-art to address key biological challenges like MHC peptide binding specificity which is central to the adaptive immune system (T-cell surveillance, differentiation, etc), since they can readily optimize parameters over large sets of binding and non-binding peptides. However, sequence-based methods are limited by their inability to incorporate detailed structural information, and as a result, they have reduced generalizability, particularly in cases where there is less training data. While structure-based methods have shown promise to fill this gap, they have been limited by their inability to accurately predict protein and peptide backbone changes which can affect both affinity and specificity, and more importantly, they lack a way to optimize many model parameters on the large amounts of peptide binding data that are often available (8).

Results

We reasoned that the recent advances in protein structure prediction could help overcome both limitations: that of structure-based methods in utilizing large amounts of known binding data, and sequence-based methods, in using structural information. AlphaFold (9) and RoseTTAFold (10) predict highly accurate structures (11) and structure quality confidence metrics that have been used to distinguish pairs of proteins which bind from those that don't with some success (12, 13). However, while these methods can be readily trained with structural data, in their current form it is not straightforward to train on binding data.

We set out to extend these networks to enable simultaneous training on structure and binding data. Because of the importance of the peptide-MHC interaction to adaptive immunity, and the very large datasets available for training and testing, we began by focusing on this binding interaction. We first explored how to provide available sequence and structure data as inputs to AlphaFold to obtain the most accurate peptide-MHC structure models (see Methods), and obtained best performance using single sequence (rather than MSA information) and peptide-MHC structure templates as inputs with the query peptide positionally aligned to the template peptides. This approach modeled peptide-MHC structures with a median peptide RMSD of 0.8 Å on backbone and 1.8 Å on all atoms for Class I (**Fig. S1A, B**) and 0.8 Å on backbone and 1.2 Å on all atoms for Class II

(**Fig. S1C, D**) complexes. These levels of accuracy are consistent with other recent work predicting peptide-MHC structures (8, 14–17).

We found previously in studies of designed proteins bound to targets (18) that both the AlphaFold residue-residue accuracy estimate, PAE, and the per-residue accuracy estimate pLDDT, are able to partially discriminate binders from non-binders. We carried out predictions for sets of binding and non-binding peptides from (1, 5, 19), and found that both the PAE, calculated between MHC and peptide, and the per-residue accuracy estimate pLDDT averaged over the peptide residues, provided some discrimination of binders from non-binders (**Fig. S2**). However, the discrimination between binders and non-binders was considerably poorer than NetMHCpan, and AlphaFold tended to dock non-binding peptides in the MHC peptide-binding groove (**Fig. 2A**).

We reasoned that the relatively poor discrimination compared to NetMHCpan reflected the lack of training of AlphaFold on negative binding examples generally, and on peptide-MHC binding data in particular. To overcome this limitation, we set out to implement a hybrid structurally-informed classification approach for fine-tuning AlphaFold that incorporates both binder and non-binder peptide-MHC pairs. To enable direct training on binding data, we added a simple logistic regression layer on top of AlphaFold which converts the peptide-MHC PAE values into a normalized binder/non-binder score. We experimented with different training regimens for optimizing the combined model parameters, and obtained the best results by first fitting the logistic regression parameters, and then subsequently fine-tuning all of the AlphaFold parameters (**Fig. 1** and **Fig. S3**). To maintain structure prediction accuracy, we included peptide-MHC structure prediction examples during training. For cases where the structure was not known, we followed a "distillation" procedure (9) in which AlphaFold models of the complex were used as the ground truth conformation, but with the loss calculated only over the MHC to allow the peptide conformation to adjust during training. Leveraging the wealth of data available on peptide-MHC interactions, we assembled a diverse training set consisting of 10340 peptide-MHC examples, 203 structurally characterized and 5102 modeled peptide-MHC binder examples, and 5035 non-binder examples, distributed across 68 Class I alleles and 39 Class II allele pairs (see Methods). Training the combined structure prediction-classification model on this set produced consistent drops in the combined loss function (**Fig. S4**), suggesting that the model might be learning to better discriminate binder from non-binder peptides on the basis of structural information. The fine-tuned model

developed a tendency to exclude non-binder peptides completely out of the MHC's peptide-binding groove (**Fig. 2A**).

We assessed prediction performance on an independent test set containing 2402 binder and 2717 non-binder peptide-MHC pairs distributed across 32 Class I and 26 Class II alleles (**Fig. 2**). This set contained no peptides with fewer than two mismatches to any peptide in the training set (**Fig. S5**). During training, the Class I examples were restricted to 9mer peptides, but in the test set we included Class I peptide-MHC pairs containing peptides of lengths 8, 9, and 10 residues to better assess generalizability of our fine-tuning approach. The combined structure prediction-classification model performs significantly better on this test set than AlphaFold with default parameters (**Fig. 2B, C**). A direct comparison to the state-of-the-art sequence-based predictor NetMHCpan on classical HLA alleles is challenging since the vast majority of available binding data is already contained in its training set. Despite our test examples being contained within NetMHCpan's training data, our fine-tuned model performed competitively with NetMHCpan, particularly on Class II, achieving AUROC of 0.97 for Class I (**Fig. 2B**) and 0.93 for Class II (**Fig. 2C**) indicating excellent classification. Beyond classification, our model also has reasonable performance on peptide-MHC binding affinity, with a Pearson correlation coefficient of 0.79 for 9-residue peptides binding to HLA-A*02:01 (compared to 0.57 for AlphaFold with default parameters and 0.78 for NetMHCpan; **Fig. S6**). Our combined prediction-classification approach appears to learn the biochemical space of peptide-MHC binding efficiently rather than relying solely on sequence patterns observed during training and retains good peptide structure prediction performance (**Figs. S7-8**).

Divergence from NetMHCpan on 10mer peptides

Our structure-based approach to peptide-MHC binding prediction is an orthogonal and hence potentially complementary method to purely sequence-based tools such as NetMHCpan. As an illustration, we noticed a striking divergence between structural predictions and NetMHCpan predictions in a scan of the Hepatitis C virus (HCV) genome polyprotein for 10mer peptides predicted to bind to the HLA-A*02:01 allele. Surprisingly, the two top-ranked NetMHCpan peptides (AKLMPQLPGI and ADLMGYIPLV) both had charged amino acids at the first anchor position (peptide position 2, a position which strongly prefers hydrophobic amino acids in known HLA-A*02:01 binders; **Fig. 3B**). By contrast, these two peptides are not predicted to bind to HLA-A*02:01 using our joint

structure and binding prediction model (**Fig. 3A**). When predicting affinities for class I peptides of length greater than 9, NetMHCpan aligns target peptides to a 9mer binding model, allowing consecutive positions to "gap out" to match the 9mer model length. For both of these peptides, peptide position 2 was not aligned to the 9mer model, making position 3 the anchor position for the purposes of scoring. This likely explains the highly favorable scores assigned to these peptides, as position 3 is a leucine (a strongly preferred anchor amino acid) in both. To assess more broadly the relationship between the NetMHCpan gap position and the consistency with our predictions, we assigned a rank error equal to the absolute difference in ranks assigned by NetMHCpan and our structural approach to each 10mer peptide window in the target and grouped these rank error values based on the NetMHCpan gap position. As shown in Figure 3C, the positions with highest mean rank error are the two canonical HLA-A*02:01 anchor positions (2 and 10, for a 10mer peptide). These results suggest that there may be differences in the reliability of NetMHCpan predictions for peptides of length greater than 9, depending on the location of the gap position(s), and more generally, that our combined structure prediction-classification model may be more robust than purely sequence-based models to the inevitable biological variations in peptide length and other properties.

Combined structure prediction-classification model extends to other protein-peptide systems

To evaluate the generalizability of our combined structure prediction-classification model beyond the peptide-MHC binding data it was trained on, we assessed performance on two additional systems: peptide recognition by PDZ domains, which bind C-terminal peptides, and peptide recognition by SH3 domains, which recognize proline-rich peptides. The PDZ dataset was taken from Ref. (20) and consists of 17 PDZ domains with an experimentally determined co-complex structure and peptide binders from phage-display selection experiments (21) (Table S1). The SH3 dataset consists of 19 SH3 domains with an experimentally determined co-complex structure and peptide binders from phage-display available in the PRM-DB database (http://www.prm-db.org/) (22) (**Table S2**). For comparison with experiments, we performed an "in silico selection" experiment in which 20,000 random peptides were modeled and ranked by our structure prediction-classification model (**Fig. 4A**). Position weight matrices (PWMs) were constructed from the top-ranked binding sequences, and these PWMs were compared with PWMs constructed from the experimentally selected binders using two PWM divergence measures proposed by Smith and Kortemme (20). The fraction of top-ranked sequences used to build the PWM

(fPWM) is a free parameter in this approach; we varied this parameter from 10-4 (i.e, only 2 sequences used to build the PWM) to 1 (all sequences used) and evaluated the accuracy of the resulting PWMs. The prediction errors, plotted as a function of fPWM, are shown in Figure 4B for our fine-tuned parameters (purple lines) and for AlphaFold with default parameters (green lines). From these error curves we can see that there appears to be an optimal value for fPWM somewhere around 0.01: using too few top-ranked sequences likely introduces noise into the PWMs, while using too many of the randomly sampled sequences reduces the degree of enrichment of preferred amino acids. Notably, fine-tuning AlphaFold's parameters on peptide-MHC binding data does not degrade performance on these other, structurally distinct, classes of protein-peptide interactions; the fine-tuned parameters perform better than the default parameters (see **Fig. S11** for a comparison to the sequence-based machine learning method HSM (23) on these two systems).

Discussion

The outstanding performance of AlphaFold on protein structure prediction and NetMHCpan on peptide MHC binding specificity prediction illustrates the power of networks trained on very large datasets; the PDB in the first case, experimental peptide-MHC binding data in the second. Here, we show how to train a single network on both structural and binding data simultaneously, and demonstrate that this enables considerable generalization of binding specificity prediction beyond the available training sets. While sequence-only networks such as NetMHCpan can achieve outstanding performance within their target domains, combined sequence and structure models can be more robust to biological variation and limited training data, and they are able to generalize to new biological systems. Moving forward, there is much to explore in network architectures for combined structure and specificity prediction; our combination of a simple classification module (logistic regression) on top of the AlphaFold network provides a starting benchmark for comparison. One notable limitation of our framework is the dichotomous treatment of binding affinity: peptides are either 'binders' or 'non-binders', which neglects variation in binding affinity that can have important biological consequences. Although the network trained on binder/non-binder classification does have some ability to predict binding strength (**Fig. S6**), it may be possible to incorporate binding affinity data into the training process and thereby produce a more accurate model.

Materials and Methods

Peptide-MHC dataset

Structures of peptide-MHC complexes were downloaded from the RCSB Protein Databank (24) with an inclusion cutoff of 2021-08-05. Peptide-MHC binding data was extracted from the training sets generously provided by the NetMHCpan developers on their website (https://services.healthtech.dtu.dk/service.php?NetMHCpan-4.1). The training set consisted of 10340 peptide-MHC examples, 203 structurally characterized and 5102 modeled peptide-MHC binder examples, and 5035 non-binder examples, distributed across 68 Class I alleles and 39 Class II allele pairs. Twenty-three mouse peptide-MHC structures were included in the training set; the remainder of the training and testing data was restricted to human alleles. We assessed prediction performance on an independent test set containing 2402 binder and 2717 non-binder peptide-MHC pairs distributed across 32 Class I and 26 Class II alleles. This set contained no peptides with fewer than two mismatches to any peptide with the same binder/non-binder label in the training set (**Fig. S5**). For Class I, only 9 residue peptides were used for training; for testing, we included 21 8-residue and 475 10-residue peptides. Class II peptides were modeled as the core 9 residue window together with one amino acid of context on either side (11 residues total). During inference, longer Class II peptides were broken into overlapping 11-residue windows for modeling and the best scoring window was used to make the binding prediction. During training, Class II peptides were represented by their best scoring 11-residue window; thus each peptide, regardless of length, mapped to a single training example. All examples used for training and testing are provided in the Zenodo repository associated with this manuscript.

Peptide-MHC template search and query-to-template mapping

The amino acid sequence of the α1 and α2 domains for MHC Class I and the α1 and β1 domains for MHC Class II were considered as the MHC query sequence and were pairwise aligned to their corresponding domains across all entries in the PDB template set. The pairwise alignment was done with the Needleman-Wunsch local alignment algorithm using the BLOSUM62 substitution score matrix while setting gap open penalty to -11 and gap extend penalty to -1. The Biopython package (25) was used to run the alignment algorithm. The alignments were used to generate query-to-template mapping strings that map each

query residue to the corresponding residue in the template crystal structure. To generate query-to-template mapping for the peptide, sequence identity was ignored and mappings were generated based on position along the peptide sequence. For Class I molecules, the query peptide was mapped to the template one to one if the length of the two peptides were equal, and if not, the first three and the last three residues of the query were mapped to the corresponding residues of the template. For Class II queries, the core 9mer residues of the peptide and their immediate neighboring residues (11mer residues in total) were mapped to the corresponding residues in the templates. If the 9mer core was not determined in the query, each possible consecutive 11-residue window along the peptide was used as the query peptide to model the complex with all possible positions for the 9mer core and the best model was selected by picking the one with the lowest inter-chain PAE among the windows. Finally, the MHC and peptide mapping strings were combined by putting the peptide either in the N-terminus or the C-terminus of the MHC mapping string.

AlphaFold modeling of peptide-MHC pairs

Query peptide sequence was appended to the N-terminus or C-terminus of the query MHC sequence. At chain break junctions, the residue indices were shifted by 200. Single sequence and templates were used as inputs and AlphaFold's model_1, model_1_ptm, model_2, or model_2_ptm weights were used for inference. The crystal structures of the top 4 hits from the template search based on the MHC alignment scores were used to construct the template features in AlphaFold. Both PDB outputs of the predictions and predicted AlphaFold metrics were saved for analysis purposes.

Hybrid AlphaFold Fine-Tuning on Structure and Binding Data

Combined structure prediction-classification approach for fine-tuning AlphaFold was performed in two stages (**Fig. S3**). First, the binder model parameters were fitted using logistic regression while the AlphaFold parameters were kept fixed at their starting values. The AlphaFold MHC-peptide inter-PAE scores for the training set, together with their binder/non-binder labels, were provided to the LogisticRegression class from the Scikit-learn (26) package linear_model. This produced slope and intercept values that defined a linear mapping between inter-PAE values and logit values. We noted that these fitted parameters differed between MHC class I and MHC class II: a model fitted on class I peptides had a higher midpoint PAE (8.03) than a model fitted on class II peptides (4.34).

This difference likely reflects structural differences in the peptide binding mode, with class I peptides bulging out of the MHC pocket, anchored at the ends, while class II peptides follow a direct and fully extended path with termini protruding at the ends of the pocket. To account for these differences, and allow for future applications in which even more disparate binding data might be combined, we allowed the logistic regression midpoint parameter to vary as a function of a predefined class value (here I or II) for each training example.

The binder model parameters were then fixed at their optimal values while the AlphaFold parameters were fine-tuned in the context of a hybrid structure and binder loss function. The loss for a single training example was equal to the softmax cross entropy of the binder/non-binder prediction for that example plus the AlphaFold structure loss, which was multiplied by a weight of 1.0, if the ground truth structure for that example was an experimentally determined structure, or 0.25, if the ground truth structure was an AlphaFold predicted structure. Predicted structural models were used as the ground truth conformations for the non-binder peptides and for binder peptides without solved structures; for these examples, the AlphaFold structure loss was computed only over the MHC to allow the peptide conformation to adjust during fine-tuning. For the parameters evaluated here, we chose a fairly conservative stopping point after two epochs of training (when each training example had been seen twice) to avoid excessive drift in the AlphaFold model. A Python script that performs parameter fine tuning, together with command line parameters and example inputs, is provided in the github repository associated with this manuscript. Full inputs for fine-tuning, including the predicted AlphaFold models used as ground truth structures, are provided in the Zenodo repository associated with this manuscript.

Evaluation of classification performance on peptide-MHC

Pairs of peptide-MHC from the dataset under evaluation were modeled with AlphaFold or Fine-Tuned AlphaFold parameters. Inter-chain PAE terms corresponding to the peptide-MHC interactions were averaged. Mean inter-chain PAE values and binder non-binder class labels were used to construct Receiver Operating Characteristic (ROC) curves and the area under the curve values were used to compare classification performance of different models. To compare our prediction results to NetMHCpan, we used NetMHCpan-4.1 and

NetMHCIIpan-4.1 (1, 2). These software tools were installed on our server and executed through the command line.

AlphaFold modeling of PDZ and SH3 domains

The PDZ domain dataset (**Table S1**) was taken from Ref. (20) and consisted of 17 human PDZ domains with experimentally determined structures. Binder peptides for the 17 PDZ domains were downloaded from supplemental data for Ref. (21) (https://baderlab.org/Data/PDZ). Experimental amino acid frequency matrices (PWMs) were constructed from the PDZ binder peptides, with clone frequency weighting. For the AlphaFold simulations, 20,000 random peptide sequences of length equal to the peptide in the experimental structure were generated using NNK codon frequencies to match the amino acid bias in the phage display libraries. The experimental structure listed in the 'template' column of **Table S1** was used as the sole template, with the random peptide sequences aligned to the template peptide and single-sequence MSA information.

The SH3 domain dataset (**Table S2**) consisted of 19 SH3 domains with experimentally determined structures extracted from the Database of Peptide Recognition Modules (http://prm-db.org/) (22). Experimental PWMs were downloaded from the PRM-DB. SH3 domains can bind peptides in two orientations, denoted class I and class II, which have opposite chain orientations: class I peptides often match a +XXPXXP sequence motif, where '+' denotes a positively charged amino acid and X is any amino acid; class II peptides often match a PXXPX+ motif. SH3 peptide PWMs from PRM-DB were annotated as class I or class II by choosing the class whose sequence motif had the highest PWM frequency (averaged over the three motif positions). Five of the SH3 domains had multiple PWMs in the PRM-DB, one of which was assigned as class I and one as class II; these domains were modeled twice, once in each orientation. For AlphaFold modeling, the native PDB structure listed in **Table S2** ('SH3 template' column) was used as the template for the SH3 domain. Four peptide-SH3 structures with peptides in the desired orientation (i.e., class I or class II) were chosen as peptide templates based on SH3 domain sequence identity (**Table S2**, 'Peptide templates' column). The peptides in these structures were transformed into the reference frame of the SH3 domain template by structural superimposition to create hybrid template models for AlphaFold. Multiple structural alignment was used to identify the core motif positions (+XXPXXP or PXXPX+) in each template peptide. The peptide sequence

modeled in the AlphaFold runs consisted of the core motif together with one residue on either side (9 residues for class I and 8 residues for class II).

For comparison with experimental PWMs, predicted PWMs were constructed from the top ranked peptide sequences. Peptides were ranked by protein-peptide inter-PAE: the sum of the residue-residue PAE scores for all (protein,peptide) and (peptide,protein) residue pairs, where PAE is AlphaFold's 'predicted aligned error' accuracy measure. The experimental PWMs were derived from phage display experiments with random peptide libraries of size 109 and greater, whereas the predicted PWMs were based on 20,000 modeled peptides. To account for this differential and better match the entropy of the amino acid frequency distributions, we squared the predicted amino acid frequencies and renormalized them to sum to 1. This had the effect of increasing the information content of the predicted PWMs without changing the order of amino acid preference. The exponent of 2 can be partly rationalized by the approximate two-fold differential in log search-space size between predictions and experiments. Following Ref. (20), predicted and experimental PWMs for PDZ domains were compared over the last 5 C-terminal peptide positions. Predicted and experimental PWMs for SH3 domains were compared over the core 7 (for class I) or 6 (for class II) positions of the SH3 motif together with the immediately adjacent positions, if those positions were present in the experimental PWM. Two measures of PWM column divergence were used to assess predictions: average absolute deviation (AAD) and the Frobenius metric. The AAD for a single PWM position equals the sum of the absolute frequency deviations for all amino acids, divided by 20; AAD ranges from 0.0 (perfect agreement) to 0.1 (maximal divergence). The Frobenius metric for a single PWM position equals the square root of the sum of the squared frequency deviations; it ranges from 0.0 (perfect agreement) to the square root of 2 (maximal divergence). The AAD and Frobenius values in Figure 4 were averaged over all PWM columns.

HCV experiment

We scanned the sequence of the Hepatitis C Virus (HCV) genome polyprotein (3011 amino acids; Uniprot ID P27958) to find potential 10mer peptide epitopes presented by HLA-A*02:01. We ran NetMHCpan with FASTA format input and default parameters. NetMHCpan compares 10mer peptides to its internal 9mer binding model by dropping one "gap" position from the peptide:model alignment. We parsed the location of the gap position from the output columns ('Of', 'Gp', and 'Gl') and confirmed by matching to the sequence in

the 'Core' column. To assess divergence between our structure-based predictions and NetMHCpan, raw scores from each method were sorted and converted to rank scores. Peptides were grouped by the NetMHCpan-assigned gap position, and the mean absolute difference in rank score was computed for each gap position.

Data and code availability

Fine tuning the AlphaFold model parameters required small changes to the AlphaFold software. The modified version of the AlphaFold package used here as well as additional python scripts for training and prediction will be made available prior to publication at https://github.com/phbradley/alphafold_finetune. Full datasets for training and testing the structure prediction-classification model, including predicted AlphaFold structures for binder and non-binder examples, will be made available prior to publication through the manuscript github repository (https://github.com/phbradley/alphafold_finetune).

References

1. B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 48, W449–W454 (2020).

2. B. Reynisson, et al., Improved Prediction of MHC II Antigen Presentation through Integration and Motif Deconvolution of Mass Spectrometry MHC Eluted Ligand Data. J. Proteome Res. 19, 2304–2315 (2020).

3. J. Jin, et al., Deep learning pan-specific model for interpretable MHC-I peptide binding prediction with improved attention mechanism. Proteins 89, 866–883 (2021).

4. G. Venkatesh, A. Grover, G. Srinivasaraghavan, S. Rao, MHCAttnNet: predicting MHC-peptide bindings for MHC alleles classes I and II using an attention-based deep neural model. Bioinformatics 36, i399–i406 (2020).

5. S. Sarkizova, et al., A large peptidome dataset improves HLA class I epitope prediction across most of the human population. Nat. Biotechnol. 38, 199–209 (2020).

6. H. Zeng, D. K. Gifford, DeepLigand: accurate prediction of MHC class I ligands using peptide embedding. Bioinformatics 35, i278–i283 (2019).

7. D. Gfeller, et al., Predictions of immunogenicity reveal potent SARS-CoV-2 CD8+ T-cell epitopes. bioRxiv, 2022.05.23.492800 (2022).

8. M. A. S. Perez, M. A. Cuendet, U. F. Röhrig, O. Michielin, V. Zoete, Structural Prediction of Peptide-MHC Binding Modes. Methods Mol. Biol. 2405, 245–282 (2022).

9. J. Jumper, et al., Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021).

10. M. Baek, et al., Accurate prediction of protein structures and interactions using a three-track neural network. Science 373, 871–876 (2021).

11. J. Jumper, et al., Applying and improving AlphaFold at CASP14. Proteins 89, 1711–1721 (2021).

12. I. R. Humphreys, et al., Computed structures of core eukaryotic protein complexes. Science 374, eabm4805 (2021).

13. L. Chang, A. Perez, AlphaFold encodes the principles to identify high affinity peptide binders. bioRxiv, 2022.03.18.484931 (2022).

14. M.-S. Park, S. Y. Park, K. R. Miller, E. J. Collins, H. Y. Lee, Accurate structure prediction of peptide-MHC complexes for identifying highly immunogenic antigens. Mol. Immunol. 56, 81–90 (2013).

15. J. M. Khan, S. Ranganathan, pDOCK: a new technique for rapid and accurate docking of peptide ligands to Major Histocompatibility Complexes. Immunome Res. 6 Suppl 1, S2 (2010).

16. T. Liu, et al., Subangstrom accuracy in pHLA-I modeling by Rosetta FlexPepDock refinement protocol. J. Chem. Inf. Model. 54, 2233–2242 (2014).

17. D. A. Antunes, D. Devaurs, M. Moll, G. Lizée, L. E. Kavraki, General Prediction of Peptide-MHC Binding Modes Using Incremental Docking: A Proof of Concept. Sci. Rep. 8, 4327 (2018).

18. N. Bennett, et al., Improving de novo Protein Binder Design with Deep Learning. bioRxiv, 2022.06.15.495993 (2022).

19. J. G. Abelin, et al., Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. Immunity 46, 315–326 (2017).

20. C. A. Smith, T. Kortemme, Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. J. Mol. Biol. 402, 460–474 (2010).

21. R. Tonikian, et al., A specificity map for the PDZ domain family. PLoS Biol. 6, e239 (2008).

22. J. Teyra, et al., Large-scale survey and database of high affinity ligands for peptide recognition modules. Mol. Syst. Biol. 16, e9310 (2020).

23. J. M. Cunningham, G. Koytiger, P. K. Sorger, M. AlQuraishi, Biophysical prediction of protein-peptide interactions and signaling networks using machine learning. Nat. Methods 17, 175–183 (2020).

24. H. M. Berman, et al., The Protein Data Bank. Nucleic Acids Res. 28, 235–242 (2000).

25. P. J. A. Cock, et al., Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423 (2009).

26. Pedregosa, Varoquaux, Gramfort, Scikit-learn: Machine learning in Python. of machine Learning ... (2011).

Figures

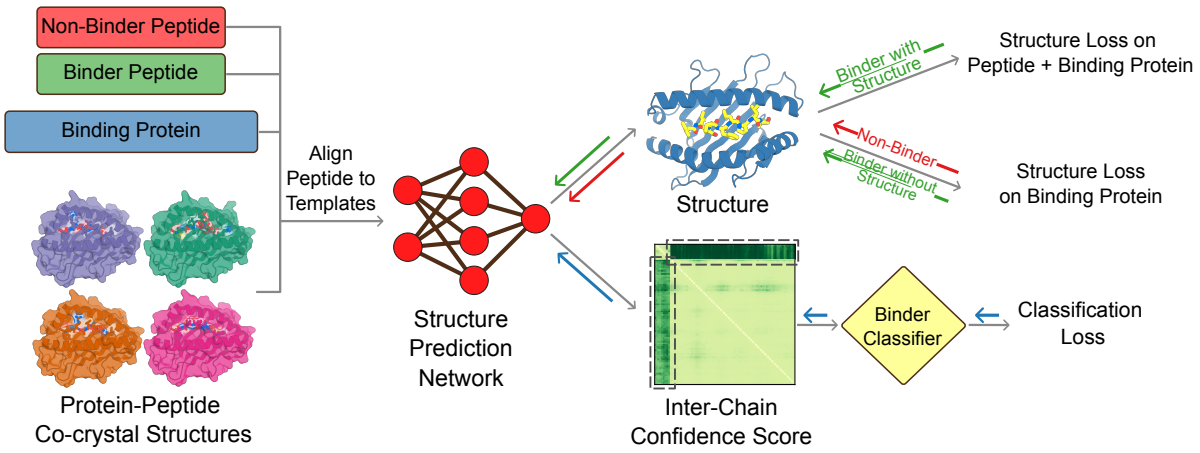


Figure 1: Extending structure prediction networks for binder classification. For fine-tuning the combined model to distinguish binding from non-binding peptides, inputs (left) are sets of known binding (green) and non-binding (red) sequences, the sequence of the target protein(s), and the available co-crystal structures which provide templates for modeling the complex. The peptide sequence is positionally aligned to the peptides in the co-crystal structures, and the structure of the complex is predicted with AlphaFold. A logistic regression classifier converts the output PAE values into a normalized binder/non-binder score and prediction. During training, the combined model parameters were optimized to minimize the loss terms in the final column, and model weights were updated through gradient backpropagation as indicated by the solid arrows. The classification loss (cross entropy) is computed on all training examples; the structure loss is computed over the entire complex for binding peptides with co-crystal structures, and over the binding protein alone for binding peptides without co-crystal structures and for non-binding peptides.

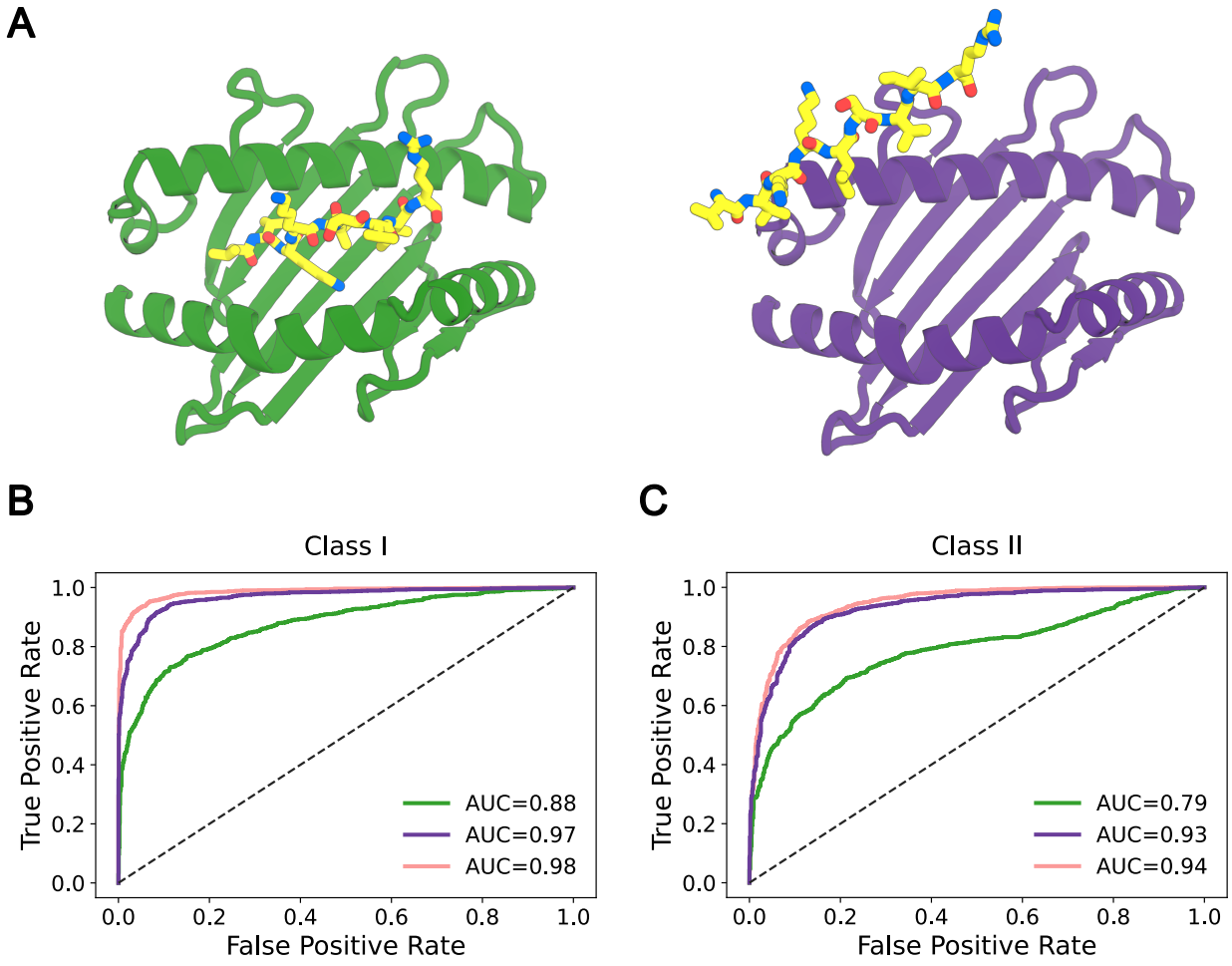


Figure 2: Fine-tuning improves pMHC-peptide binder classification to state-of-the-art levels. (A) Structure prediction for a non-interacting peptide-MHC pair using AlphaFold before (left) and after (right) fine-tuning on joint structure prediction and binder classification. The peptide is incorrectly predicted to bind the MHC in the first case but not the second. (B, C) Binder/non-binder peptide classification ROC curves on test sets (see Methods): **(B)** MHC Class I, **(C)** MHC Class II. AlphaFold (green), combined structure prediction-classification model (purple), and NetMHCpan (pink).

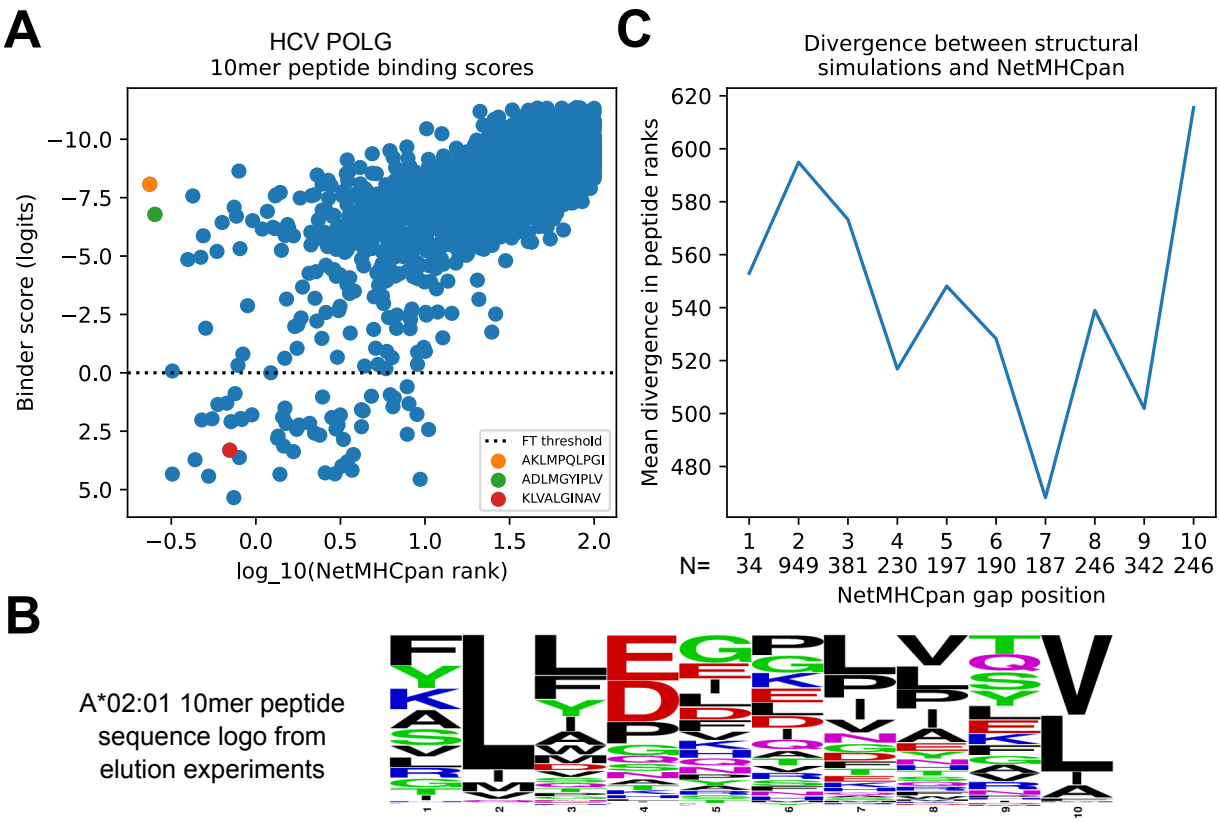


Figure 3. Prediction of 10mer peptide binding highlights differences between sequence- and structure-based approaches. (A) Scatterplot of NetMHCpan rank score (x-axis, $\log_{10}$-transformed) versus combined structure prediction-classification model binder score (y-axis, reversed so that predicted strong binders for both methods are in the lower left) reveals strong overall correlation (R^2=0.49) as well as peptides with divergent predictions (upper left and lower right corners). The top two NetMHCpan predicted peptides (orange and green markers) have charged amino acids at the first anchor position (peptide position 2). A known immunodominant T cell epitope (red marker) scores well by both methods. **(B)** Sequence logo built from 10mer peptides eluted from A*02:01 (5) shows strong preference for hydrophobic amino acids at the anchor positions. **(C)** Average disagreement in peptide rank between NetMHCpan and the structural model (y-axis) is highest when the NetMHCpan gap position is located at one of the two anchor positions (2 and 10).

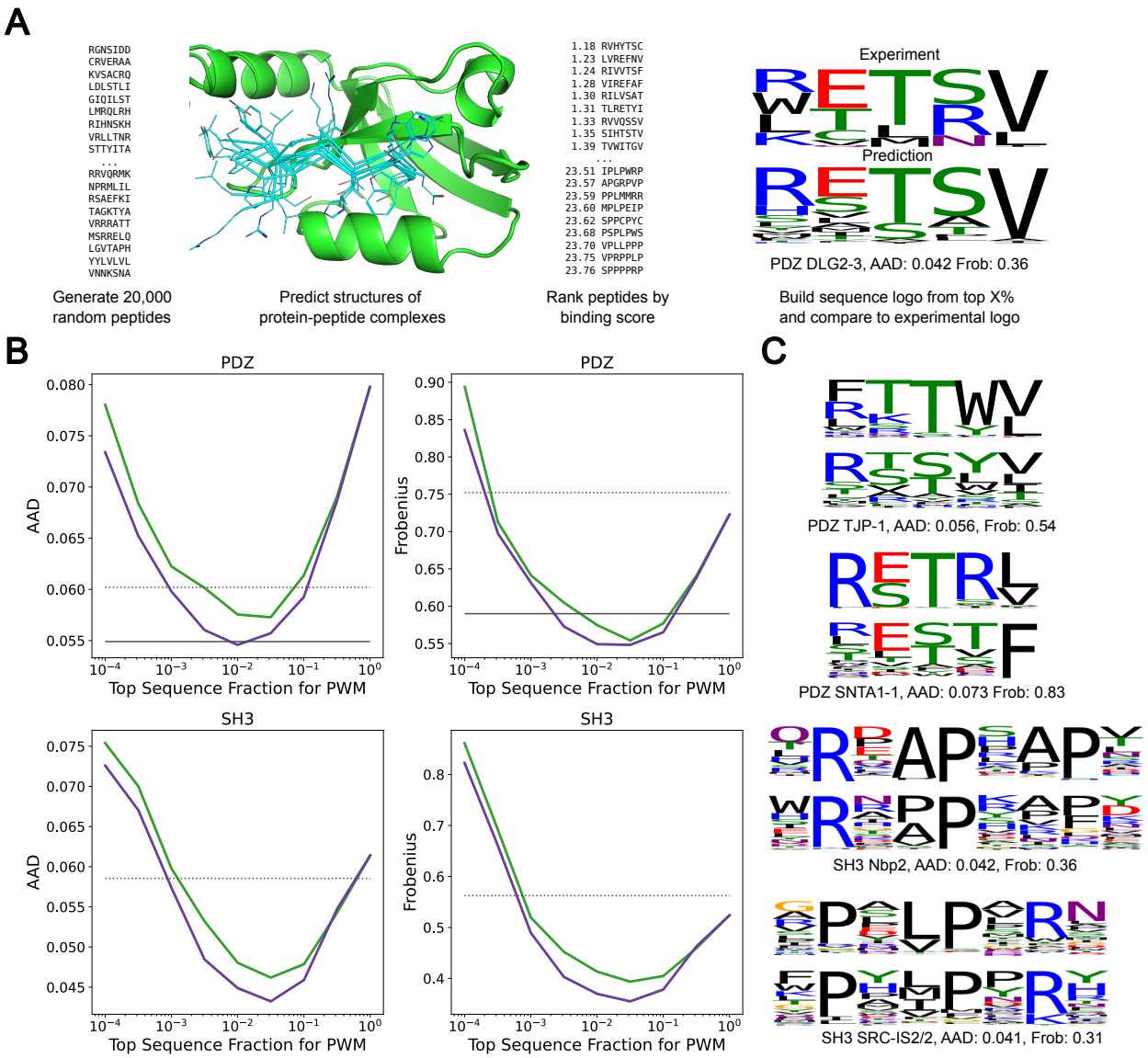


Figure 4: Accurate prediction of PDZ-peptide and SH3-peptide interactions using the combined structure prediction-classification network. (A) Prediction of position-weight matrices by random sequence generation followed by network evaluation. **(B)** Combined network fine-tuned on peptide-MHC interactions outperforms AlphaFold with default parameters in recapitulating experimental PWMs. AAD: average absolute difference in amino acid frequency; Frobenius: average Frobenius distance between PWM columns (lower is better for both; see Methods). AlphaFold (green), and fine-tuned AlphaFold (purple). The solid black lines in the PDZ panels indicate the performance of the structure-based method from Ref. (20), which was optimized on PDZ domains. Dashed black lines indicate the average AAD and Frobenius scores of PWMs constructed from the sequences of the template PDB peptides used for modeling. **(C)** Examples of predicted and experimental sequence logos (See **Figs. S9-10** for a complete set of sequence logos).

Acknowledgements

Peptide binding specificity prediction using fine-tuned protein structure prediction networks

Amir Motmaen[1,2,3], Justas Dauparas[1,2], Minkyung Baek[1,2], Mohamad H. Abedi[1,2,4], David Baker[1,2,4*], Philip Bradley[2,5*]

[1] Department of Biochemistry, University of Washington, Seattle, WA, USA.

[2] Institute for Protein Design, University of Washington, Seattle, WA, USA.

[3] Bioengineering Graduate Graduate Program, University of Washington, Seattle, WA, USA.

[4] Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA.

[5] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

* To whom correspondence should be addressed: David Baker, Philip Bradley

Email: dabaker@uw.edu, pbradley@fredhutch.org

Author Contributions: A.M., J.D., M.B., M.A., and P.B. performed research. A.M., J.D., M.B., and P.B. contributed new reagents/analytic tools. A.M., J.D., P.B. analyzed data. A.M., P.B., and D.B. wrote the paper. P.B. and D.B. designed research.

The authors wish to thank Frank DiMaio, Sergey Ovchinnikov, Ivan Anishchanka, Amijai Saragovi, Aditya Krishnakumar, and Ian Humphreys for helpful discussions, Morten Nielsen for assistance with NetMHCpan, and Luki Goldschmidt for technical IT support. We are grateful to the developers of NetMHCpan for making their training and testing datasets publicly available.

We thank AWS and Microsoft for providing computing resources. This work was supported with funds provided by a Microsoft gift (A.M., J.D., M.B., D.B.); the Audacious Project at the Institute for Protein Design (D.B.); the Howard Hughes Medical Institute (M.A., D.B.); and NIH R35 GM141457 (P.B.). The Jane Coffin Childs Memorial Fund for Medical Research (M.A).

Competing Interest Statement: Authors declare that they have no competing interests.

Supplementary Material for Peptide binding specificity prediction using fine-tuned protein structure prediction networks

Supplementary discussion on model training

The combined structure prediction-classification model adds two additional "logistic regression" parameters to the internal AlphaFold model parameters. These parameters, a slope and a switch-point, transform the mean MHC:peptide residue-residue confidence measure (the 'PAE' score) into a binder/non-binder probability that ranges from 0 to 1. A central hypothesis of this study was that fine-tuning the internal AlphaFold parameters in the context of binder classification would lead to significant improvement in performance beyond what could be achieved with the default parameters. A key question was how best to optimize both the AlphaFold internal parameters and the two logistic regression parameters. One alternative would be to simultaneously fit both the AlphaFold parameters and the logistic regression parameters. Since AlphaFold has 100s of millions of parameters that collectively summarize information on the protein sequence-structure mapping, and we are fine-tuning on a very limited domain (peptide-MHC interactions), we were concerned about moving too far from the starting parameters and thereby losing general structural knowledge that could be relevant for new protein-peptide systems not seen during training. This desire to be conservative in the fine-tuning guided many of our decisions, from stopping the AlphaFold fine-tuning after just two epochs (i.e., each example seen twice), to this question of how to train the binder model. For example, if we started with randomly initialized logistic regression parameters, early training steps might perturb the internal AlphaFold parameters in unphysical directions. Thus, we decided to first fit the two logistic regression parameters in the context of the starting AlphaFold parameters, and then keep them fixed during the fine-tuning of the AlphaFold parameters. This also avoids 'mixing' of different parameter types during training (i.e., logistic regression parameters and AlphaFold internal parameters). To explore this question further, we re-ran model training several times with random initial logistic regression parameters and saw that AUROC values of the final models were comparable to or lower than the model described in the main text. We also tried initializing the logistic regression parameters to their fitted values and letting them vary during training, and we found that they remained essentially unchanged during fine-tuning of the AlphaFold parameters, and the final model performed equivalently to the one described above.

Supplementary Figures

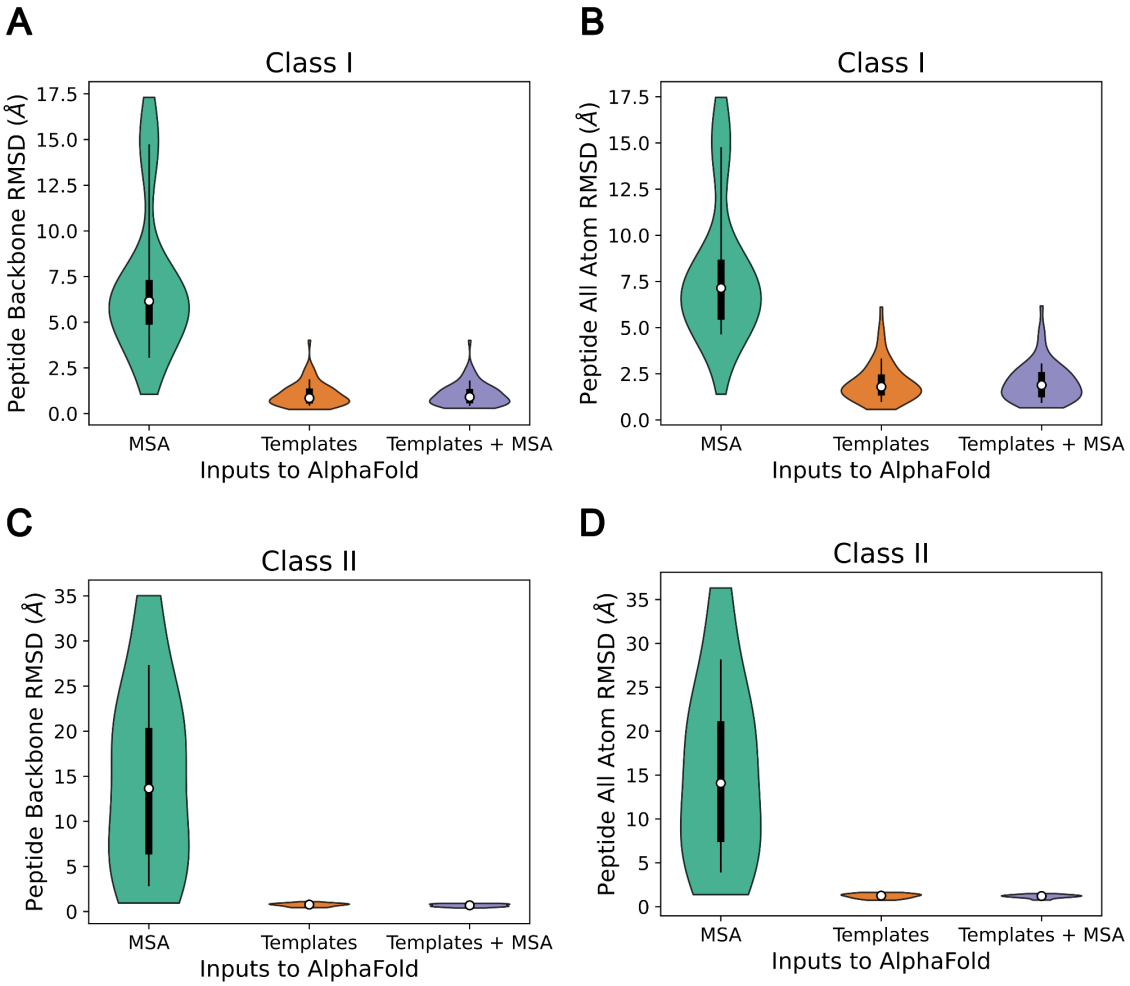


Fig. S1. Comparison of peptide-MHC structure modeling quality of AlphaFold with different inputs. Violin plots of peptide RMSD distribution of peptide-MHC models produced by AlphaFold with MSA, Templates, or MSA and Templates as inputs to the network. **(A)** peptide backbone RMSD for Class I, **(B)** peptide all atom RMSD for Class I **(C)** peptide backbone RMSD for Class II and (D) peptide all atom RMSD for Class II complexes. White circles represent the median, thick lines the interquartile range, and thin lines the range between 10th and 90th percentiles.

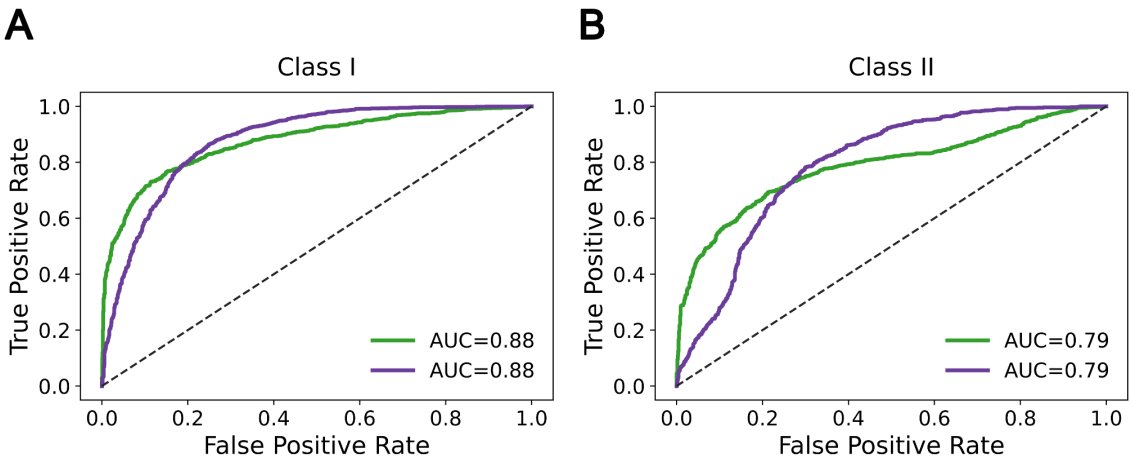


Fig. S2. Peptide-MHC classification with AlphaFold structure accuracy metrics. ROC curves comparing AlphaFold's confidence metrics in binder/non-binder peptide classification for **(A)** MHC Class I and **(B)** MHC Class II. Mean inter-chain PAE (green) and mean peptide pLDDT (purple).

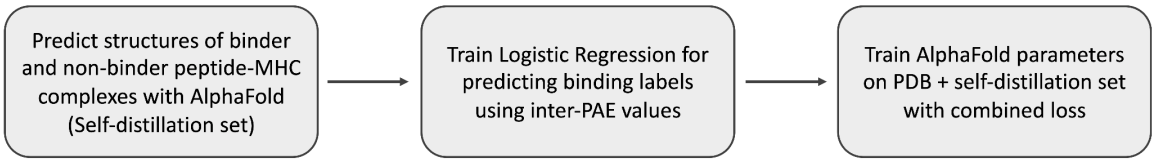


Fig. S3. Flowchart of the combined structure prediction-classification fine-tuning on peptide-MHC dataset. Overview of the steps regarding the combined fine-tuning procedure explained in the Methods section.

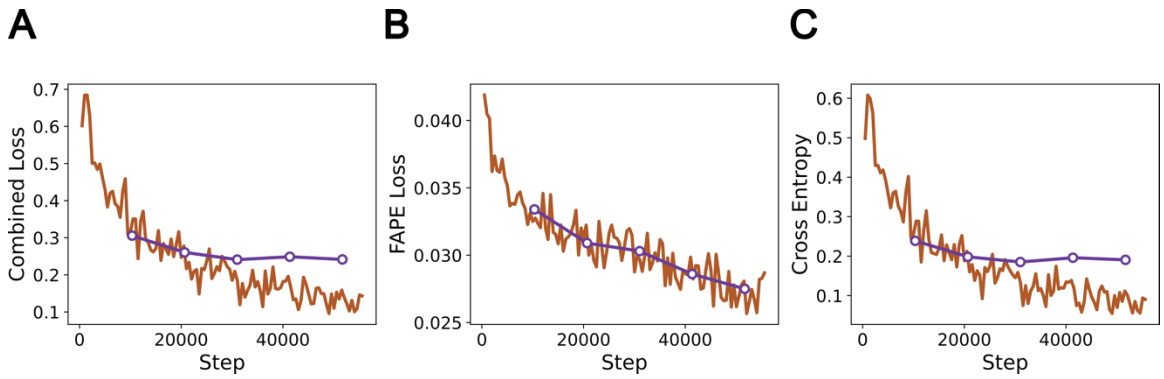


Fig. S4. Loss progression curve during fine-tuning of AlphaFold on peptide-MHC structures. A-C. Plots of **(A)** combined fine-tuning loss, **(B)** FAPE loss on structures (crystal structures and predicted self-distillation structures), and **(C)** cross entropy of peptide-MHC binding classification against training steps in the combined structural and classification fine-tuning of AlphaFold parameters on peptide-MHC data. Training loss (maroon) and validation loss (purple).

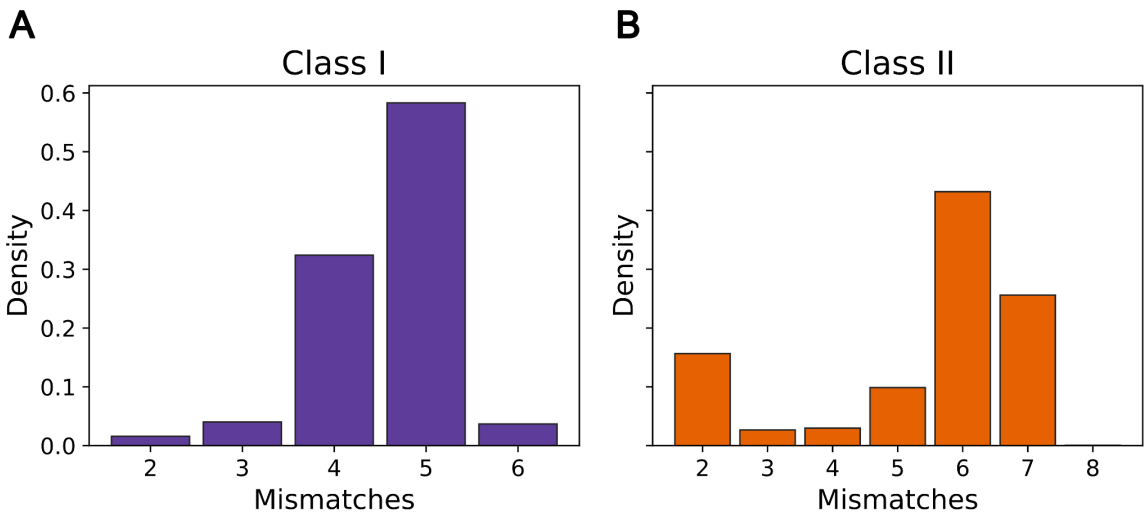


Fig. S5. Histogram of peptide distance in validation set compared to train set. Minimum number of mismatches of peptides in the validation set compared to any peptide in the training set for **(A)** MHC Class I and **(B)** MHC Class II.

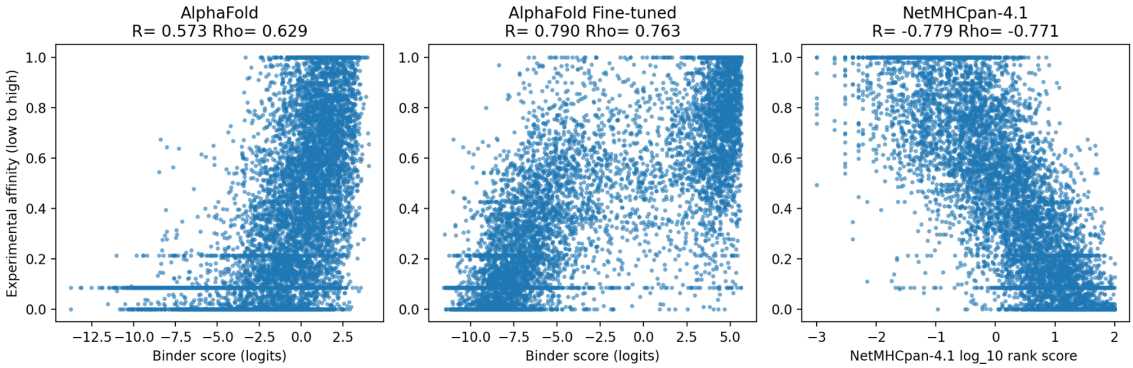


Fig. S6. Correlation between predicted and experimental binding affinities for 9 residue peptides binding to the HLA allele A*02:01. Predictions for AlphaFold are shown on the left, for the fine-tuned binder model in the middle, and for NetMHCpan-4.1 on the right. Experimental binding affinities were taken from the NetMHCpan binding affinity training set (1). Affinities in that set have been log-transformed and scaled so that a value of 0 corresponds to weak binding (>= 50 mM) and a value of 1 corresponds to strong binding (<= 1 nM). Only peptides that did not overlap with the binder fine-tuning training set for any HLA-A*02 allele were included in the analysis. Pearson linear correlation coefficients (R values) and Spearman rank correlations (rho values) are shown in the panel titles. The NetMHCpan-4.1 rank score was $\log_{10}$-transformed (the untransformed rank score showed a weaker Pearson correlation coefficient of R=-0.537).

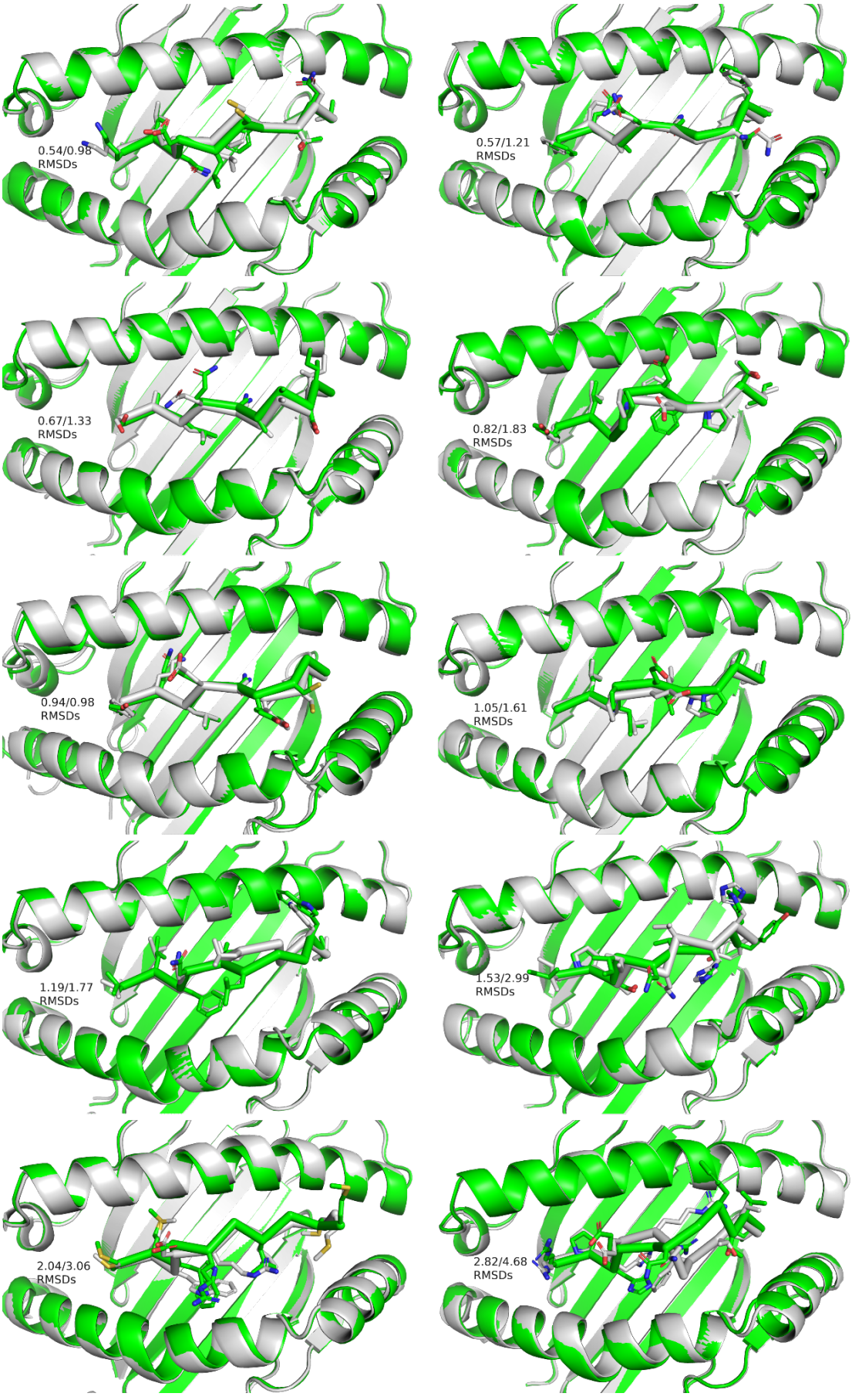

0.54/0.98
RMSDs
0.57/1.21
RMSDs
0.67/1.33
RMSDs
0.82/1.83
RMSDs
0.94/0.98
RMSDs
1.05/1.61
RMSDs
1.19/1.77
RMSDs
1.53/2.99
RMSDs
2.04/3.06
RMSDs
2.82/4.68
RMSDs

Fig. S7. Representative class I peptide-MHC structural models generated by the fine-tuned binder model. Target peptides have at least 2 mismatches to any binder peptide in the training set. Modeled structure is shown in green and native in gray. Backbone/all-atom RMSD values are shown to the left of the peptide, which is oriented with the N-terminus on the left. Representative structures were chosen uniformly along the backbone RMSD distribution (median 0.98 Å backbone and 1.83 Å all-atom RMSD).

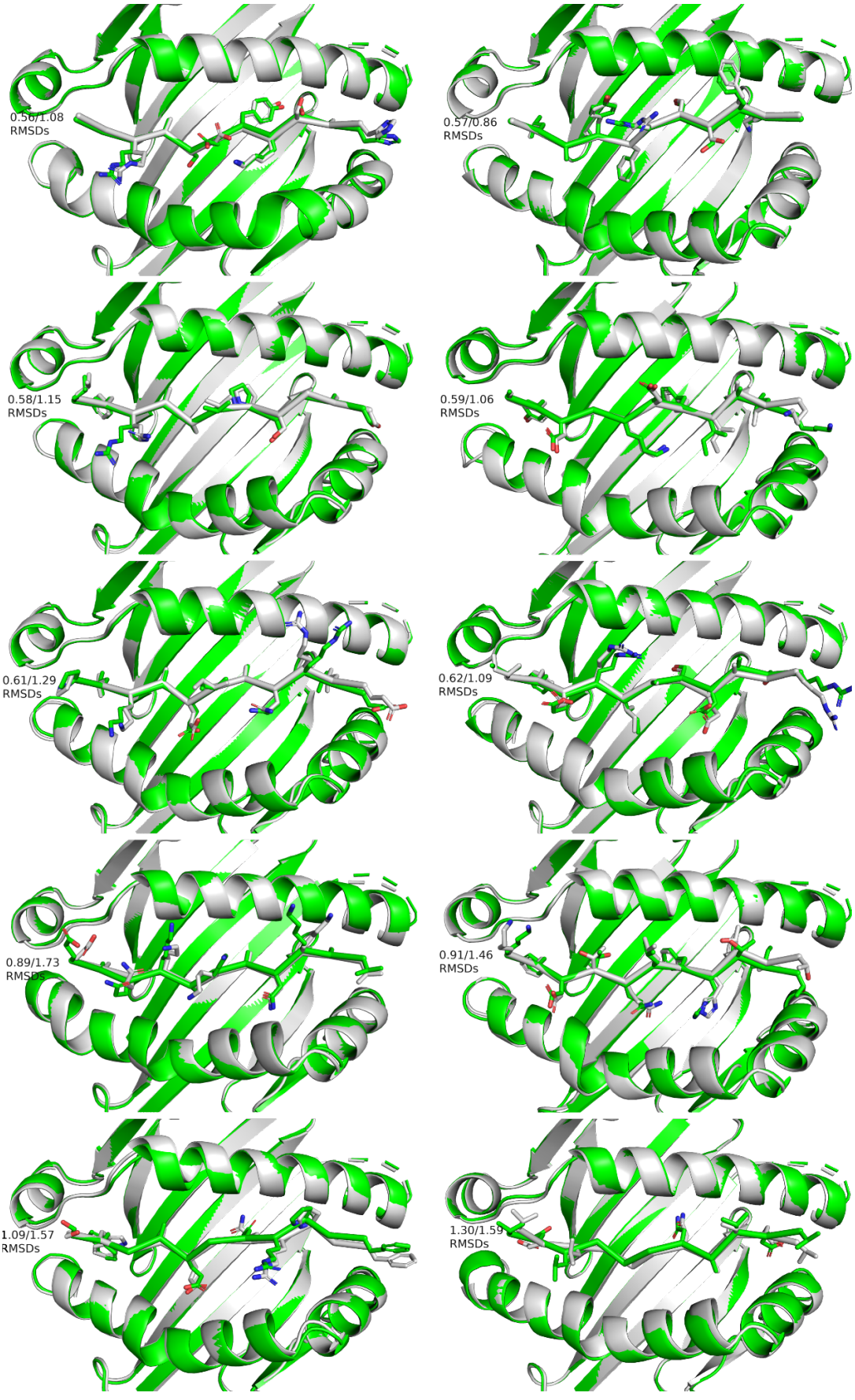

0.56/1.08
RMSDs
0.57/0.86
RMSDs
0.58/1.15
RMSDs
0.59/1.06
RMSDs
0.61/1.29
RMSDs
0.62/1.09
RMSDs
0.89/1.73
RMSDs
0.91/1.46
RMSDs
1.09/1.57
RMSDs
1.30/1.59
RMSDs

Fig. S8. Representative class II peptide-MHC structural models generated by the fine-tuned binder model. Target peptides have at least 2 mismatches to any binder peptide in the training set. Modeled structure is shown in green and native in gray. Backbone/all-atom RMSD values are shown to the left of the peptide, which is oriented with the N-terminus on the left. Representatives were chosen uniformly along the backbone RMSD distribution (median 0.61 Å backbone and 1.15 Å all-atom RMSD).

Experimental Logo, PDLIM4-1, PDBID: 4q2o

Predicted Logo, AAD: 0.033, Frobenius: 0.36

Experimental Logo, DLG1-3, PDBID: 2i0i

Predicted Logo, AAD: 0.038, Frobenius: 0.40

Experimental Logo, DLG2-3, PDBID: 2he2

Predicted Logo, AAD: 0.042, Frobenius: 0.36

Experimental Logo, PTPN13-2, PDBID: 1d5g

Predicted Logo, AAD: 0.044, Frobenius: 0.38

Experimental Logo, DLG1-2, PDBID: 2i0l

Predicted Logo, AAD: 0.046, Frobenius: 0.45

Experimental Logo, DLG4-3, PDBID: 1tp5

Predicted Logo, AAD: 0.050, Frobenius: 0.50

Experimental Logo, MPDZ-13, PDBID: 2fne

Predicted Logo, AAD: 0.053, Frobenius: 0.51

Experimental Logo, MLLT4-1-hi, PDBID: 2ain

Predicted Logo, AAD: 0.054, Frobenius: 0.44

Experimental Logo, SLC9A3R2-2, PDBID: 2he4

Predicted Logo, AAD: 0.055, Frobenius: 0.58

Experimental Logo, MPDZ-10, PDBID: 2opg

Predicted Logo, AAD: 0.055, Frobenius: 0.53

Experimental Logo, ERBB2IP-1-hi, PDBID: 1n7t

Predicted Logo, AAD: 0.055, Frobenius: 0.60

Experimental Logo, TJP1-1, PDBID: 2h2b

Predicted Logo, AAD: 0.056, Frobenius: 0.54

Experimental Logo, MPDZ-7, PDBID: 2iwq

Predicted Logo, AAD: 0.058, Frobenius: 0.56

Experimental Logo, MPDZ-12, PDBID: 2iwp

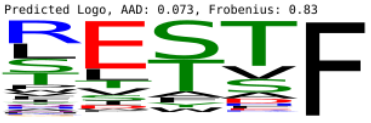
Predicted Logo, AAD: 0.065, Frobenius: 0.70

Experimental Logo, DVL2-1, PDBID: 1l6o

Predicted Logo, AAD: 0.073, Frobenius: 0.75

Experimental Logo, SNTA1-1, PDBID: 1qav

Predicted Logo, AAD: 0.073, Frobenius: 0.83



Experimental Logo, CASK-1, PDBID: 1kwa

Predicted Logo, AAD: 0.078, Frobenius: 0.86

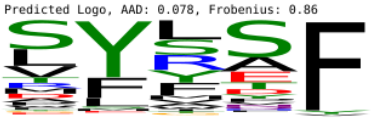


Fig. S9. Predicted and experimental sequence logos for the 17 PDZ domains. Predicted logos were built from the top 1% of random peptides ranked by the combined structure prediction-classification model. Domains are ordered by increasing AAD between predicted and experimental logos.

Fig. S10. Predicted and experimental sequence logos for the 24 SH3 domains. Predicted logos were built from the top 1% of random peptides ranked by the combined structure prediction-classification model. Domains are grouped by SH3 class and ordered within each class by increasing AAD between predicted and experimental logos.

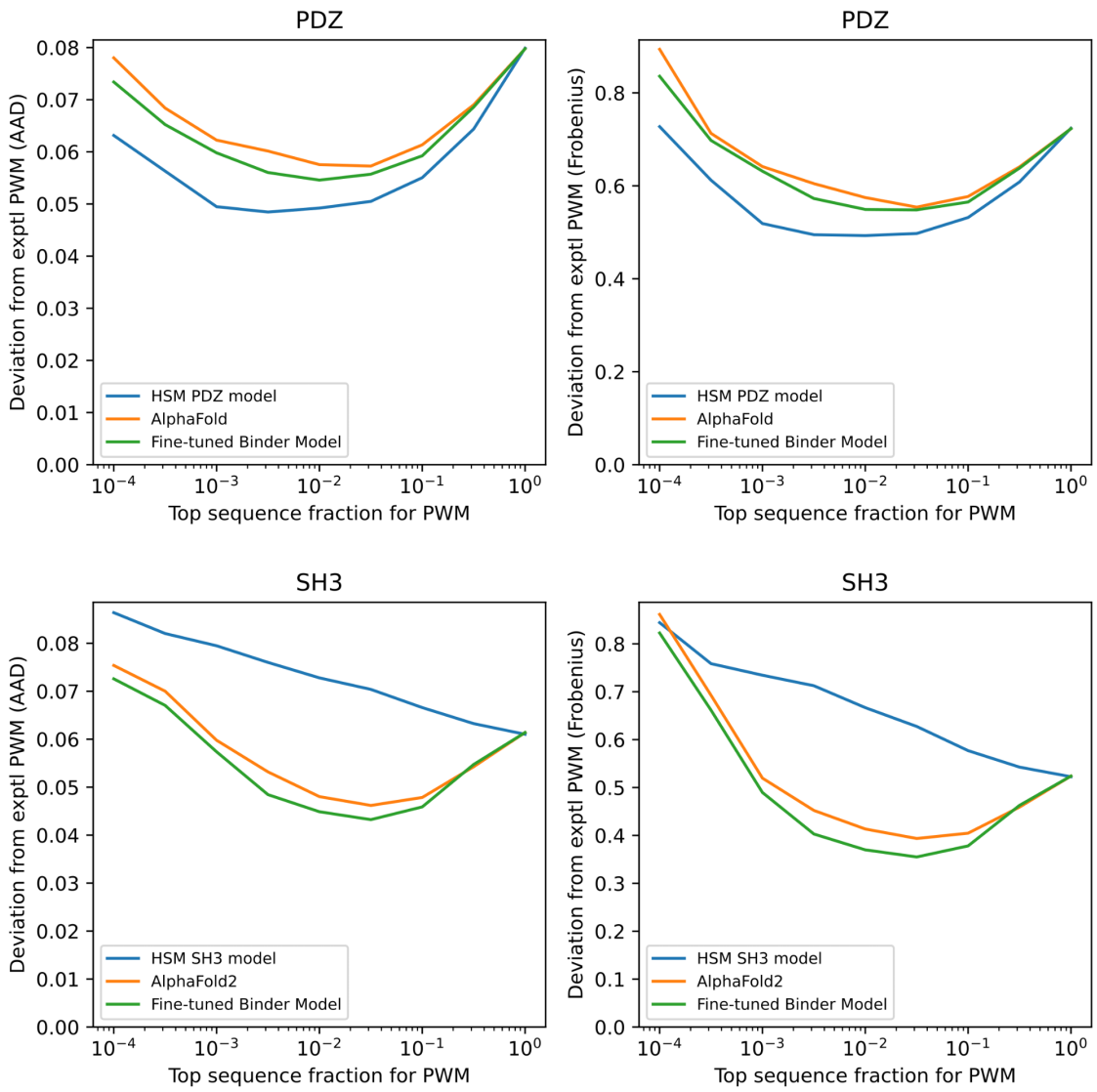


Fig. S11. PWM prediction accuracy for structure-based models and for the HSM machine learning approach. As described in the main text, twenty thousand random peptides were ranked with the structure-based approaches (standard AlphaFold with inter-PAE ranking and the fine-tuned binder model) and with the sequence-based machine learning approach hierarchical statistical mechanical modeling (HSM) (2). PWMs were generated from the top fraction of the ranked sequences as indicated on the x-axis, and these PWMs were compared with experimental PWMs using the AAD (left) and Frobenius (right) measures (see Methods). PDZ PWMs were aligned based on the position of the C-terminus; SH3 PWMs were aligned based on the PxxP motif (structure-based approaches) or by trying all registers and taking the one with minimal deviation (HSM). Pre-trained HSM PDZ and SH3 models and prediction scripts were downloaded from https://github.com/aqlaboratory/hsm. The HSM approach outperforms the structure-based models for PDZ domains, suggesting that further performance gains could be achieved by including PDZ binding data in the training set used for fine-tuning.

Supplementary Tables

Table S1. PDZ domain dataset

PDZ domain	Template[a]	Template peptide
CASK-1	1kwa_AB	VPSYREF
DLG1-2	2i0l_AC	RRETQV
DLG1-3	2i0i_AD	RRETQV
DLG2-3	2he2_AB	KIHETSV
DLG4-3	1tp5_AB	KKETWV
DVL2-1	1l6o_AD	LKLMTTV
ERBB2IP-1-hi	1n7t_AB	TGWETWV
MLLT4-1-hi	2ain_AB	LFSTEV
MPDZ-7	2iwq_AA	SIISTRL
MPDZ-10	2opg_AB	PYKSTRL
MPDZ-12	2iwp_BA	DVSETSV
MPDZ-13	2fne_BA	SSDETSV
PDLIM4-1	4q2o_AB	VESPWL
PTPN13-2	1d5g_AB	NEQVSAV
SLC9A3R2-2	2he4_AA	VGPSTRL
SNTA1-1	1qav_AB	THLETTF
TJP1-1	2h2b_AA	WRRTTYL

[a]PDB ID followed by PDZ and peptide chain identifiers

Table S2. SH3 domain dataset

PRM-DB identifier[a]	SH3 class	Protein name	Uniprot identifier	SH3 template[b]	Peptide templates[c]
PRM_0152	1	Sho1	P40073:287-367	2vkn_A	2vkn_AC,3ua7_AE,4rtz_AB,5sxp_BF
PRM_0157	1	GRB2	P62993:1-58	1gbq_A	5sxp_BF,1io6_AB,2bz8_AC,3ua7_AE
PRM_0161	2	Pex13	P80667:295-386	1n5z_A	1n5z_AP,1gbq_AB,1gbq_AB,5ul6_AM
PRM_0162	1	Nbp2	Q12163:103-180	2lcs_A	2lcs_AB,2bz8_AC,1io6_AB,4rtz_AB
PRM_0166	2	Lsb3	P43603:385-451	1ssh_A	1ssh_AB,3ua7_DF,2d1x_DQ,5ul6_AM
PRM_0190	2	Bbc1	P47068:1-78	1zuk_B	2rqu_AB,4ln2_AB,2d1x_DQ,2rpn_AB
PRM_0193	2	Abp1	P15891:535-592	2rpn_A	2rpn_AB,2d1x_DQ,4wci_EF,2rqu_AB
PRM_0210	1	GRB2	P62993:1-57	1gbq_A	5sxp_BF,1io6_AB,2bz8_AC,3ua7_AE
PRM_0210	2	GRB2	P62993:1-57	1gbq_A	1gbq_AB,1gbq_AB,2d1x_DQ,5xhz_AC
PRM_0215	1	SRC-IS2/2	P12931:87-144	4rtz_A	4rtz_AB,3ua7_AE,5sxp_BF,1io6_AB
PRM_0215	2	SRC-IS2/2	P12931:87-144	4rty_A	4rty_AB,3ua7_DF,5ul6_AM,4f14_AB
PRM_0231	2	ASAP1	Q9ULH1:1073-1131	2rqu_A	2rqu_AB,5xhz_AC,2rpn_AB,2d1x_DQ
PRM_0238	2	ARHGEF7	Q14155:187-242	5sxp_B	5xhz_AC,3u23_AB,2d1x_DQ,3ua7_DF
PRM_0266	1	FYN	P06241:85-142	3ua7_A	3ua7_AE,4rtz_AB,5sxp_BF,1io6_AB
PRM_0266	2	FYN	P06241:85-142	3ua7_D	3ua7_DF,4rty_AB,5ul6_AM,1gbq_AB
PRM_0270	2	CRK	P46108:135-191	5ul6_A	5ul6_AM,4rty_AB,1gbq_AB,1gbq_AB
PRM_0272	2	CD2AP	Q9Y5K6:111-166	3u23_A	3u23_AB,5xhz_AC,4wci_EF,2d1x_DQ
PRM_0273	2	CD2AP	Q9Y5K6:2-58	4wci_E	4wci_EF,5xhz_AC,3u23_AB,2rpn_AB
PRM_0282	2	CTTN	Q14247:459-512	2d1x_D	2d1x_DQ,4f14_AB,2rpn_AB,1gbq_AB
PRM_0285	1	SH3KBP1	Q96B97:2-57	2bz8_A	2bz8_AC,5sxp_BF,1io6_AB,3ua7_AE
PRM_0285	2	SH3KBP1	Q96B97:2-57	2bz8_A	4wci_EF,2d1x_DQ,5xhz_AC,3u23_AB
PRM_0291	1	BIN1	O00499:523-593	5i22_A	5sxp_BF,2bz8_AC,1io6_AB,3ua7_AE
PRM_0298	1	SORBS1	Q9BX66:870-927	4ln2_A	5sxp_BF,2bz8_AC,1io6_AB,3ua7_AE
PRM_0298	2	SORBS1	Q9BX66:870-927	4ln2_A	4ln2_AB,1gbq_AB,1gbq_AB,2rqu_AB

[a]Internal identifier used in the PRM-DB database (http://prm-db.org/)

[b]PDB and chain identifier of SH3 domain modeling template

[c]PDB identifier followed by SH3 and peptide chain identifiers of peptide modeling templates

SI References

1. B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 48, W449–W454 (2020).

2. J. M. Cunningham, G. Koytiger, P. K. Sorger, M. AlQuraishi, Biophysical prediction of protein-peptide interactions and signaling networks using machine learning. Nat. Methods 17, 175–183 (2020).

Chapter 2

Targeting peptide-MHC complexes with designed T cell receptors and antibodies

As originally published in Motmaen, A., Jude, K.M., Wang, N., Minervina, A., Feldman, D., Lichtenstein, M.A., Ebenezer, A., Correnti, C., Thomas, P.G., Garcia, K.C. and Baker, D., 2025. Targeting peptide-MHC complexes with designed T cell receptors and antibodies. bioRxiv, pp.2025-11.

Abstract

Class I major histocompatibility complexes (MHCs) present peptides derived from intracellular proteins for surveillance by T cells. The precise recognition of foreign or mutated peptide–MHC (pMHC) complexes by T cell receptors (TCRs) is central to immune defense against pathogens and tumors. Here, we introduce and validate a structure-based deep learning strategy for the creation of TCRs and TCR-mimic antibodies by designing dozens of binders to specific pMHCs of interest. Experimental structures of designed antibodies bound to their pMHC targets demonstrate atomic-level accuracy at the recognition interface. Computationally designed TCRs and antibodies targeting pMHC complexes could enable a broad range of therapeutic applications, from cancer immunotherapy to autoimmune disease treatment, while also providing insight into the specificity and activity of natural TCRs.

Introduction

The peptide–MHC complexes (pMHCs) on the surface of a cell provide a window onto its internally expressed proteome, enabling the detection of viral infection or oncogenic transformation from external cues. Thus disease-associated pMHCs are an attractive recognition target for delivery of cytotoxic therapies, and this is the basis of the cellular arm of the adaptive immune system, in which T cell receptors (TCRs) mediate recognition (1, 2). Highly variable TCRs are produced by a genome rearrangement process—V(D)J

recombination—that generates sequence diversity by random joining of germline-encoded gene segments together with nucleotide insertion and deletion at the segment junctions (3). TCR sequences are filtered during T cell development in the thymus by positive and negative selection processes that guarantee a baseline but not excessively strong level of binding to self-pMHCs (4). Native TCRs derived from patients are being explored as therapeutics, but their reactivity toward self or near-self targets may be constrained by thymic selection. Computational protein design represents an alternative path to therapeutic TCRs, but de novo generation of the loop-mediated binding modes employed by TCRs remains challenging, and past work on TCR design has been limited to the reengineering of existing binders (5–9). The composite nature of the pMHC interface—consisting of a few variable peptide residues surrounded by many conserved MHC positions—represents an additional barrier to peptide-specific recognition.

Deep learning protein design methods based on generative diffusion models (10) or gradient backpropagation through structure prediction networks (11) have been widely used to generate binders against many targets. Though successful, these methods can be compute- and memory-intensive for large systems like TCR:pMHC complexes, and network backpropagation may lead to inflated model confidence and exploration of adversarial sequences and structures in some cases. We reasoned that the highly conserved binding modes of TCRs for their pMHC targets would be well-learned by structure prediction networks trained on the PDB and fine-tuned on TCRs, enabling the use of simpler inference-based methods directly for conformational sampling. We set out to explore the possibility of combining such an approach for conformational sampling with ProteinMPNN-based sequence design (12) to generate pMHC-specific TCRs and 'TCR-mimic' antibodies.

Results

Design of pMHC-targeting TCRs and antibodies

We developed a structure-based deep learning strategy, ADAPT (Antigen-receptor Design Against Peptide-MHC Targets), for the design of pMHC-targeting TCRs and antibodies. The ADAPT design pipeline (**Fig. 1**) combines fine-tuned Alphafold2 (13) (AF2) conformational sampling with ProteinMPNN (12) fixed-backbone sequence design. Diversity is introduced into the pipeline through the use of multiple TCR/antibody template structures (the choice

of template determines the sequence and structure of the receptor outside the Complementarity-Determining Region (CDR) loops) and by initializing each design run with random CDR3 sequences taken from a large library of paired human TCRs. A single design run consists of four steps (**Fig. 1, top**): initial docking, sequence design of the 6 CDR loops, redocking, and design model evaluation. TCR-like binding modes are enforced by fine-tuning AF2 on TCR:pMHC structures and binding data (14, 15) and by providing AF2 with generic TCR-like docking templates as input features (see Methods). The re-docking step allows the binding mode to adjust to the designed CDR sequences and provides AF2 quality metrics for design ranking. To obtain a second view of design quality, a version of the RoseTTAFold2 (RF2) network (16, 17) fine-tuned on antibodies and TCRs is used to model the final design sequence, providing a confidence estimate as well as a predicted structure that can be compared with the original design. This process is repeated many thousands of times with different template framework structures and different randomly selected CDR3 pairs, resulting in a large ensemble of candidate designs that are ranked by a quality score that combines the AF2 and RF2 confidence scores as well as the similarity between their structural models (see Methods). The top-ranked 100-500 designs for each pMHC target are then taken as starting points for a genetic algorithm (GA) refinement procedure, conducted in parallel across multiple concurrent processes operating on a shared design pool (**Fig. 1, bottom**). Each process repeatedly selects a random design from the current pool, mutates two interface positions, redocks and redesigns the interface, and then replaces an existing design in the current pool if the refined design has superior metrics (with limits to prevent a single lineage from taking over the pool, see **Methods and fig. S1**). ADAPT refinement can be viewed as iterative alternation between ProteinMPNN sequence design and AF2 redocking, with interface mutations introducing diversity and the GA framework enabling intensified exploration of promising regions of design space.

Designed TCRs bind their pMHC targets

We applied this TCR design pipeline to a panel of 9 pMHC targets including cancer-associated epitopes, tumor neoepitopes, and viral peptides (**Table 1**). Top-ranked sequence designs were introduced into TCR-null Jurkat T cells and assessed for binding to on- and off-target pMHCs by flow cytometric analysis with multimerized pMHC staining reagents (see **Methods**). For four targets (A01-EVD, A02-LLW, A02-GLM, and A02-TLM; **Table 1**), successful binder designs were identified in small-scale clonal screens consisting of 30-60 sequences. Binders for two additional targets (A02-ALY and B44-SEI)

were found by screening larger libraries of 1000-5000 designs built by cloning chip-synthesized DNA fragments into template-specific backbones (see **Methods**). No successful designs were identified for the remaining three targets (A02-HMT, A03-ALH, and A11-VVV). Structural models and binding data are shown in **Fig. 2A** for a single design for each of the six successful targets. On- and off-target binding for 35 nonredundant designs from 17 distinct refinement lineages, along with a dendrogram showing their sequence relationships, are summarized in **Fig. 2B** and **Fig. S2**. The scatter plot in **Fig. 2C** shows binding data for a larger set of successful and failed designs; note that the off-target pMHCs span a range of similarities to the design target, from the highly similar Titin-derived off-target for A01-EVD to more distant and therefore less challenging discrimination targets (**Table S1**). The designed CDR sequences are highly mutated relative to their template structures, with 70-80% of the designable positions changing in a typical loop (**Table S2**). To visualize the diversity of binding modes in these designs, we calculated the 6 rigid-body orientational parameters relating pMHC and TCR for each design and for a non-redundant set of MHC class-I restricted native TCR structures (see **Methods**). Visualization of the docking geometries by principal components analysis (PCA) shows that the designs are well distributed across the landscape of native TCR binding modes (**Fig. 2D**) and that designs sample binding modes that are distinct from those of the experimental structures used to template their framework regions (**Fig. S3**). Retrospective modeling of the designed sequences with the default Alphafold3 (18) pipeline (without interface templates or fine-tuning) recapitulated the designed binding geometries with high confidence (fig. S4A). Finally, we performed an all-vs-all binding screen matching the six representative designs against all six target pMHCs and observed binding to the intended target along with some off-target binding, particularly to the A02-TLM pMHC (**Fig. 2E**). Alphafold3 metrics showed a substantial correlation with experimental binding results (**Fig. S4B-C**, **Fig. S5**). Taken together, these results demonstrate that the ADAPT pipeline can generate diverse, novel TCRs that bind to their intended pMHC targets.

Antibody binding modes confirmed by high-resolution structures

Like TCRs, antibodies are generated through a genetic recombination process that produces diverse CDR loop regions which mediate target binding. Because they function in part as secreted proteins, antibodies often make better soluble reagents than TCRs, which remain anchored to the T cell surface. Thus pMHC binding antibodies might have advantages over TCRs when a soluble targeting molecule is desired, for example as a component of a bispecific T cell engager (19, 20). To date, a handful of pMHC-binding

antibodies have been identified by traditional immunization and screening approaches (21–23). By replacing the TCR template structures with a set of antibody variable region structures, ADAPT can design antibodies that target pMHC complexes using TCR-like binding modes. We used ADAPT to generate TCR-mimic antibodies that specifically recognize three different pMHC targets with relatively high success rates in small-scale screens (**Fig. 3A-C**, **Fig. S6**, **Table 2**). Monovalent affinities of these designs for their targets ranged from 5nM to 700 nM, with most designs binding in the high nanomolar range (**Table S3, Fig. S7**). Cryogenic electron microscopy (cryoEM) structures were determined for two of the designed complexes—vAB-30 bound to the A01-EVD pMHC and vAB-66 bound to A02-TLM—revealing atomic accuracy at the interface (**Fig. 3D-E**). Cα-RMSD calculations over full complexes yielded values of 0.7Å and 0.6Å for vAB-30 and vAB-66 respectively; Cα-RMSDs of 0.6Å (vAB-30) and 1.2Å (vAB-66) were calculated for the CDR loop regions after aligning on the MHC structures. For the vAB-30 complex, peptide recognition is primarily mediated by the heavy chain CDR2, with a hydrogen bond between Arg54HC and the C-terminal Tyr9 main-chain oxygen. Additional stabilization arises from a Pro–aromatic CH–π interaction between Light chain CDR1 Tyr32LC to peptide residue Pro4. Contacts to the MHC are distributed, involving the α1 helix from CDR2HC, CDR3HC and α2 helix from CDR1LC, CDR2LC, and CDR3LC (**Fig. S8A**). For the vAB-66 complex, peptide recognition is driven by the heavy chain CDR1, with a hydrogen bond from Asp33HC to peptide residue Ser4, as well as van der Waals contacts from CDR3HC. Additional contacts from Phe32LC and Trp92LC extend the reach toward the C terminus of the peptide. Contacts to the MHC are more distributed, with contact to the α1 helix from CDR3HC, CDR1LC, and CDR2LC, and contact to the α2 helix from CDR1HC (**Fig. S8B**).

Existing structures of antibodies with TCR-like recognition modes have shown two broad classes of docking geometry which are related by 180° rotation of the antibody about its pseudo-symmetry axis: one in which the antibody heavy chain aligns with the TCR beta chain ('forward' docking), and one in which the heavy chain aligns with the TCR alpha chain ('reverse' docking). This is controlled in the ADAPT pipeline at the docking and redocking steps (**Fig. 1**) through the use of docking-geometry templates that are aligned to the template antibody structure in either the forward or reverse orientation. The forward or reverse orientation is randomly selected at the start of each individual design calculation in order to sample candidate designs with both orientations. As shown in **Fig. 3** and **Table 2**, the successful pMHC binders exhibit both orientations. To visualize the designed binding modes relative to native TCRs, we performed PCA analysis of their orientational parameters, now dropping the single parameter describing rotation of the antibody about its internal pseudosymmetry axis (which corresponds to the choice of forward vs reverse

dock; see **Methods**). This analysis showed that the designed binding modes fall within the space sampled by natural TCRs (**Fig. 3F**).

To explore potential clinical applications, we formulated three of the A02-ALY specific antibodies as bi-specific T cell engagers (see **Methods**) and evaluated their ability to trigger T cell activation in the presence of target cells pulsed with peptide variants. Two of the designs (vAB-246 and vAB-250) showed specificity for the target peptide and its near variants (**Fig. 3G**), with the pattern of sensitivity to peptide mutations for vAB-246 appearing to correlate with extent of antibody contact at the designed interface (**Fig. 3H**). The third design induced activation even in the absence of pulsed peptide, suggesting that it binds to off-target pMHC complexes on the target cell surface.

Designed TCRs activate T cells

For application as cytotoxic targeting reagents, TCRs must have high specificity toward their target epitope in order to avoid killing healthy cells bearing structurally similar self-pMHCs. This is challenging because the TCR recognizes a composite interface consisting of a few residues from the target peptide together with multiple MHC residues that are shared with many self-pMHCs. T cell activation provides a sensitive read-out of TCR specificity due to the signal-amplifying properties of the activation signaling cascade (24). Using transgenic T cells expressing designed TCRs co-cultured with antigen presenting cells in the presence and absence of pulsed target peptide, we found that many of the designed TCRs—despite being able to discriminate their target peptide from specific off-target competitors in binding screens—showed relatively high levels of non-specific activation in the absence of target peptide (**Fig. 4A-B**; **Fig. S9**); the more specific designs tend to have weaker on-target binding than non-specific designs (**Fig. 4C**).

Overall, there was a strong positive correlation between on-target binding and on-target activation: designs that bound above a threshold tended to support T cell activation and vice versa (**Fig. 4D**; Pearson's r = 0.87). Several designs targeting the highly polar HLA-A*02:01-restricted epitope ALYDKTKRI from terminal deoxynucleotidyl transferase (TdT) showed low levels of non-specific activation (lower right corner of **Fig. 4A**). We evaluated their fine specificity by additional activation experiments with peptides containing individual alanine mutations and similar peptides from the human proteome (**Fig. 4E**).

Several designs were able to discriminate the target peptide from some, though not all, of the nearby mutant peptides, and the pattern of mutation sensitivity could be rationalized by examination of the designed structures, with mutations at non-contacted peptide positions being better tolerated than mutations at positions with extensive TCR contacts (**Fig. 4F**).

Discussion

The specificity of T cell receptors for their pMHC epitopes underpins the exquisite precision of the adaptive immune system, providing protection from pathogens and elimination of transformed cells. Computationally designed TCRs and antibodies that specifically recognize tumor-associated pMHCs could form the basis of soluble or cell-based cancer immunotherapies; designs targeting pathogen-derived epitopes or immunogenic self-peptides could have applications treating infectious or autoimmune diseases, respectively. There has been exciting recent progress designing mini-protein binders for pMHC targets (25–27), yet there remain compelling reasons to design TCRs and antibodies for this purpose: engineered TCRs introduced into transgenic T cells benefit from fully native signaling pathways; there are well-established pipelines for development of antibody therapeutics; TCRs and antibodies may be less immunogenic than fully synthetic mini-proteins; and TCR design may yield insights into the function of natural TCRs, with implications for disease diagnosis and basic immunology. Our structure-based pipeline for the computational design of TCRs and antibodies that bind to pMHC targets iterates between conformational sampling with a fine-tuned structure prediction network and fixed-backbone sequence design. Application of this approach to nine distinct pMHC targets produced TCR binders for 6 of them, 4 from small-scale clonal screens. Two of the successful targets lacked a previously determined pMHC structure (**Table 1**), which suggests that it should be feasible to target new and less well-characterized epitopes. A simple extension to antibody design yielded binders from small screens for all three targets attempted, including a highly polar peptide (ALYDKTKRI) in complex with HLA-A*02:01. The pipeline's success in designing challenging, loop-mediated interactions at relatively high success rates demonstrates that structure prediction algorithms like AF2 can be powerful sampling engines for protein design, provided that sampling can be focused on relevant conformations (here through the use of fine-tuning and pMHC docking templates) and that sufficient diversity can be introduced (in our case through the use of diverse seed CDR3 sequences taken from natural human TCRs). Requiring only two forward passes through the AF2 network per design or refinement simulation (**Fig. 1**), the ADAPT pipeline has

modest GPU hardware requirements relative to more memory-intensive gradient-backpropagation approaches (the designs reported here were generated on commodity NVIDIA GPUs with 11 GB of VRAM). We additionally found in preliminary testing that ADAPT was better able to generate native-like TCR docking modes than a fine-tuned version of RFdiffusion (**Fig. S10**), which may suggest that structure-prediction based design approaches can be advantageous in geometrically-constrained settings with multiple, diverse examples for training.

The biophysical determinants of T cell activation are incompletely understood (28). Our large panel of ADAPT-designed TCRs, generated without the influence of positive or negative thymic selection, provides insight into the relationship between TCR binding and T cell activation. We found a significant ($p<10^{-80}$) positive correlation (**Fig. 4D**) between the levels of on-target binding, as measured by staining with multimerized pMHC reagents, and on-target activation from peptide-pulsing experiments. With some exceptions, designs that showed measurable binding to their targets also mediated T cell activation by presenter cells pulsed with target peptide. Given that all of these TCRs were designed to bind in canonical docking modes within the envelope of previously characterized TCR structures (**Fig. 2D**), our results show that binding with a canonical docking mode can support TCR activation (29, 30). Within the range of affinities sampled by these designs, we do not see strong evidence for additional biophysical or interface chemistry constraints beyond canonical binding.

While the ability of our design pipeline to generate TCRs that bind pMHCs and activate T cells is an advance for the field, there is still considerable room for improvement to generate more specific designs. The specificity of nearly all of these in silico-generated sequences is likely insufficient for targeted killing of specific cell populations, and hence further optimization by experimental or computational means will be necessary. Our pipeline could likely be improved by favoring contacts to the peptide over the MHC and by encouraging more native-like sequences in the CDR1 and CDR2 loops (which are currently being enriched for hydrophobic amino acids, **Fig. S11**). More broadly, our difficulty in designing TCRs that activate exclusively in the presence of a target peptide highlights the importance of negative thymic and peripheral selection steps in the development of natural T cells.

Materials and Methods

Computational design methods

In the first stage of the pipeline, many independent docking and design calculations are carried out, each with a different random choice of template antigen-receptor (TCR or antibody) structure and initial CDR3 loop sequences (**Fig. 1, top**). The template receptor structure determines the sequence and structure of the non-CDR framework regions and of the CDR1 and CDR2 loops. TCR template structures were taken from the RCSB protein databank (31). Antibody template structures and metadata were downloaded from the SAbDab database (32). The CDR3 loop sequences, taken from a paired human T cell receptor, are spliced into the template sequence to create a hybrid receptor sequence that is provided to the fine-tuned AF2 network along with the sequences of the target MHC and peptide. AF2 is run without MSA input but is provided with four multichain templates assembled by combining sequence-similar peptide-MHC structures together with the structure of the template receptor positioned in four different docking modes relative to the pMHC. These four 'generic' docking geometries, which are fixed and identical for every design calculation, were chosen to minimize the rigid-body distance to a set of canonical ternary TCR:pMHC structures from the RCSB protein databank (31). In other words, the four generic docks were chosen to optimally cover the space of potential docks as inferred from experimentally determined TCR:pMHC structures. The database of paired CDR3 sequences contains 1,688,863 sequences assembled from publicly available single-cell genomics studies on T cells (33). The fine-tuned AF2 model was trained on experimentally-determined TCR:pMHC structures and binding data using the "alphafold_finetune" approach described in Ref. (15) (https://github.com/phbradley/alphafold_finetune).

After AF2 docking, the sequences of the 6 CDR loops (CDRs 1-3 on both the alpha/light and beta/heavy chains; IMGT (34) loop definitions are used) are redesigned using the ProteinMPNN network (12) with the following parameters: temperature=0.1, model_name='v_48_020', num_seq_per_target=3, number of edges=48, training noise level=0.2Å. After sequence design, a new docked structure is generated using AF2 with docking templates exactly as in the initial docking step. This structure represents the final design model, and the AF2 quality metrics are recorded for selecting the top designs.

To evaluate design model quality, the mean AF2 predicted aligned error (PAE) score for all (pMHC,TCR) and (TCR,pMHC) residue pairs is computed. The designed sequence is also input to a RoseTTAFold2 (RF2) model fine-tuned on antibody and TCR complex structures (16), and the CDR3 RMSD between the design model and the RF2 prediction is calculated after aligning on the MHC chains. Scatter plots of these design quality metrics for representative design runs are shown in Figure S1. A single weighted ranking score is calculated according to the formula:

ranking_score = 2 * AF2_pMHC_TCR_PAE + 1 * RF2_pMHC_TCR_PAE + 0.5 * CDR3_RMSD

Minimizing this score selects for lower TCR-pMHC PAE values (ie, higher structural confidence) and lower RMSDs (ie, greater similarity) between the AF2 and RF2 models of the designed interface. The CDR3 RMSD value is calculated after aligning both models on the MHC chain, so it captures both the internal conformation of the CDR3 loops as well as their docked orientation relative to the MHC.

In the second stage of the pipeline, the top 100-500 designs by ranking score are selected to form an initial pool for subsequent refinement with a parallel genetic algorithm. During the refinement process, multiple independent simulations operate independently on the pool of designs, each selecting targets for refinement from the current pool at random and adding the refined models to the pool if they improve it, subject to a diversity constraint that prevents a single lineage from taking over the pool. This is necessary because refined designs do not replace their parent in the pool, which allows for intensified sampling in promising regions. The default diversity constraint caps the number of designs allowed in the pool from the same lineage at 10. Here a single lineage is defined as the set of all designs descended from the same founding member of the refinement pool.

Each independent refinement trajectory consists of a perturbation, a redesign, and a redocking calculation. Perturbation consists of randomly mutating two of the CDR positions to new amino acids and redocking with AF2. All positions in the six CDR loops are then redesigned with ProteinMPNN, and a final model with quality metrics is generated with AF2 redocking and RF2 reprediction. At this point the current refinement pool is read back into memory and any refined designs that are superior to existing pool members

replace those old designs (subject to the lineage diversity constraint). Thus a fixed pool size is maintained during the refinement process.

Docking Geometry Analysis

The TCRdock software (14) (https://github.com/phbradley/TCRdock) was used to assign six rigid-body docking geometry parameters to each design model and to a set of nonredundant native class I TCR:pMHC ternary structures. In this analysis, reference coordinate frames are first defined for the MHC and for the TCR. In both cases, the reference frame is defined by a set of residue pairs related by an approximate C2 symmetry axis. In the case of the MHC, these residues are taken from the antiparallel beta sheet that forms the floor of the peptide binding groove (**Fig. S12A**). For the TCR, they consist of a set of 13 structurally conserved framework positions that superimpose closely when aligning the TCR alpha and beta chains (or their analogous aligned positions in the antibody light and heavy chains) (**Fig. S12B-C**). In each case, the x-axis of the reference frame is defined by the axis about which a rotation optimally superimposes the residue pairs (is, the approximate symmetry axis of the residue pairs), with the frame origin and z-axis defined by the centers of mass of the residue sets being aligned (**Fig. S12D**). With these choices, the x-axes of the TCR and MHC frames are approximately aligned and antiparallel (red arrows in **Fig. S12D**). Once the frames are defined, the six orientational parameters are taken to be (1-2) the y- and z-coordinates of a unit vector from MHC to TCR represented in the MHC local coordinate system, (3-4) the y- and z-coordinates of a unit vector from TCR to MHC represented in the TCR local coordinate system, (5) the distance between the reference frame origins, and (6) the dihedral angle formed by the y-axis vectors along the rotation axis connecting the frame origins. For TCRs binding in a canonical orientation, this last parameter clusters around 180 degrees, whereas for TCR-mimic antibodies, dihedral angles of ~0 and ~180 degrees are observed, corresponding to the reverse and forward docking modes, respectively.

Lentivirus preparation and transduction

Lentiviral particles were produced in HEK293T cells (ATCC, CRL-3216) using a standard third-generation packaging system. Cells were seeded to reach ~80% confluence after 24 h

and transfected with pMD2.G (Addgene #12259), psPAX2 (Addgene #12260), and the lentiviral transfer vector at a 2:3:4 plasmid mass ratio using Lipofectamine 3000 (Thermo Fisher Scientific, L3000015) in Opti-MEM. 800 and 150 ng total DNA was transfected for 24-well and 96-well scale. After 4 h, the transfection medium was replaced with fresh DMEM containing 10% FBS and 1x penicillin–streptomycin (Thermo Fisher Scientific, 15140122) or Antibiotic-Antimycotic (Thermo Fisher Scientific, 15240062). Viral supernatants were harvested 36–48 h post-transfection, clarified through 0.45 µm low-protein binding filters, used fresh or aliquoted, and stored at -80 °C.

For transduction, target cells in log growth phase were resuspended in complete medium optionally supplemented with 8 µg/mL polybrene (Santa Cruz Biotechnology, sc-134220) and mixed with 5 to 200 uL of fresh or thawed lentivirus. Optionally, cells were spinned at 1000 × g for 2 h at 33 °C, then resuspended and incubated under standard culture conditions. Media were refreshed after 3 h, and cells were passaged or expanded 24 h post-infection. Puromycin selection (Thermo Fisher Scientific, A1113803) was initiated 48 h after transduction at 500-1000 ng/mL. Transduction efficiency was assessed by flow cytometry for mTagBFP2 expression 48 h post-transduction.

Transposon based stable cell line production

iOn plasmids (47) were co-transfected with pCAG-hyPBase at a 3:1 molar ratio using the Lonza 4D-Nucleofector system (Lonza). For each reaction, 4×10^5 cells were used per well in 20 µL nucleocuvettes with a total of 2 µg DNA and nucleofected according to the manufacturer's recommended program for the respective cell type. Following nucleofection, cells were transferred to pre-warmed complete medium and incubated under standard culture conditions. Cells were expanded the next day and Puromycin selection started at 500 ng/mL 48 hours post nucleofection. For library nucleofection, a number of 8 reactions were performed for a total of 3.2×10^6 nucleofected cells to achieve a coverage of at least 100x of the library complexity.

Variant plasmid preparation

Plasmids containing TCR and BCR variants used in clonal assays were either obtained from Genscript, IDT, or Twist, or cloned in house by cloning DNA fragments containing the

coding sequence into receiving vectors using golden gate assembly with BsaI. When cloned in-house, ZymoPure Plasmid Miniprep kit (Zymo Research) was used.

Pooled plasmid cloning

The original iOn and lenti backbone were modified to enable compatibility with BsaI, BsmBI, PaqCI BbsI, and SapI. Due to the higher cloning efficiency, the iOn plasmid was used for assembling pooled libraries. Pooled libraries were obtained as linear fragments from Twist (Multiplexed Gene Fragment product). Each DNA fragment contained four designed sequence regions (CDR1b–FW2–CDR2b, CDR3b, CDR1a–FW2a–CDR2a, and CDR3a), arranged in order and separated by three pairs of antiparallel Type IIS restriction sites (BsmBI, PaqCI, and BbsI) to enable directional Golden Gate assembly. The entire construct was flanked by outward-facing BsaI sites and subpool-specific adapter sequences for PCR amplification and library integration (Fig. S13). After receiving the double stranded library, we performed qPCR to amplify the subpools corresponding to designs made using a given framework against a specific target. Amplified subpools were either gel extracted or purified using SPRIselect magnetic beads (Beckman Coulter) and quality controlled with gel electrophoresis. We then performed golden gate assembly, PCR clean-up (DNA Clean and Concentrator-5 by Zymo Research), electroporation into Endura electrocompetent E. coli (Biosearch Technologies) and grown overnight in 37°C in shake flasks containing media with antibiotics before harvesting and plasmid prep. We used Plasmidsaurus whole plasmid sequencing to quality control each prep.

Cell based binding assay

Jurkat cells deficient in TRAC and TRBC containing a NFAT-RE_EGFP construct (Jurkat ab-) were transduced with lentivirus, undergone puromycin selection, and then stained with APC-conjugated anti-CD3 antibody (clone UCHT1 Biolegend) and PE-conjugated peptide MHC Dextramer (Immudex) or Tetramer (Fred Hutch immune monitoring core) and analyzed with an Attune NxT Flow Cytometer (Thermo Fisher).

TCR hits against B44-SEI targets were initially tested by transient transfection into HEK293T cells stably expressing CD3 ε, δ, ζ, γ, then stained and analyzed 48h post-transfection similar to Jurkat ab- cells.

T cell co-culture stimulation assay

Jurkat ab- cells transduced with TCRs and selected for puromycin expression were harvested, washed, with fresh cRPMI and co-incubated with appropriate antigen presenting cells in a ratio of 1:1 in the presence of anti-CD49d (BD Biosciences) and anti-CD28 (BD Biosciences) each at a 1:1000 dilution. Then, appropriate peptides were pulsed in at the desired concentration (with 1 uM as the most common concentration used) and incubated overnight at 37°C with 5% CO2. Then the media was replaced with DPBS and the cells were analyzed with flow cytometry to determine the fraction of EGFP+ cells among mTagBFP2+ population. In experiments containing T cell engagers (TCEs) the appropriate TCE protein was added at the desired concentration (50 nM as the most commonly used concentration).

Screening scFvs through yeast display

We performed yeast surface display for screening some of the designed antibodies with the standard protocol described in Cao et. al. (35). The scFvs were fused through the N-term to Aga2p with a VH - GS Linker - VL orientation. Staining was performed using pMHC Dextramers or Tetramers and anti-Myc or anti-IgG1 antibodies.

Mammalian display of antibodies

We created vectors for mammalian display of antibodies by fusing them to PDGFR transmembrane domain in either the iOn or Lenti backbone.IgG1-Fabs and full-length IgG1s were fused to PDGFR through the C-term of their heavy chain while the light chain was not fused to PDGFR. Staining was performed similar to staining Jurkats harboring TCRs.

Bi-specific T cell engager production

Chimeric IgG based T-cell engagers were cloned by fusing an anti-CD3 scFv to the C-terminus of the human IgG1κ gene and swapping the canonical Fc domain with a silenced Fc variant (36). These variants were first screened for binding on the HEK293T surface using a displayed protein A based capture system. Variants with correct binding were moved forward and protein preps for these constructs were obtained from Genscript through CHO-S or Genewiz using Expi293 cell expression and Protein A based purification.

Affinity measurement using Surface Plasmon Resonance (SPR)

Binding affinity measurements were performed on a Biacore 8K instrument (Cytiva) at 25 °C using HBS-EP+ buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% v/v Tween-20). For the A01-EVD system, human IgG1s were immobilized on a Protein A sensor chip and serial dilutions of biotinylated peptide–MHC complexes (pMHCs) were injected as analytes. For all other targets, biotinylated pMHCs were immobilized on a streptavidin (SA) sensor chip and monovalent IgG1-Fab analytes were injected at multiple concentrations. Association and dissociation phases were monitored under continuous buffer flow, and reference subtraction (blank and control surfaces) was applied to remove nonspecific signal. Measurements were performed using single-cycle kinetics (sequential analyte injections without surface regeneration). Data were globally fitted using a 1:1 Langmuir binding model to extract kinetic rate constants (k_{on}, k_{o}ff) and equilibrium dissociation constants (K_D). Monovalent Fabs for SPR were prepared by digestion of full-length IgG1s (acquired from Genscript or Genewiz) with LysC or Papain and removing the undigested or Fc portion with Protein A based agarose columns.

Cryo-EM sample preparation and data collection

Fab proteins of vAB30 and vAB66 for structural characterization were obtained from Genscript. The designed variable domains were fused to the human CH1 domain of IgG1 including the hinge domain, cloned in pcDNA3.4, expressed in CHO-S cells and purified with CH1 resin.

For A01-EVD, the pMHC protein used for cryo-EM was expressed in Expi293 cells (Thermo Fisher Scientific). A single-chain trimer (SCT) construct (37) was cloned into the pD649 vector, consisting of the peptide, a $(G_3S)_3$ linker, β2-microglobulin, a $(G_4S)_4$ linker, the HLA-A*01:01 α chain, and a C-terminal His tag. A total of 200 µg of SCT plasmid was transfected into 400 million Expi293 cells following the manufacturer's instructions. After four days, cells were pelleted at 500 × g for 10 min, and the supernatant was collected and diluted 1:1 with PBS. 4 mL Ni–NTA resin were added, and the mixture was incubated overnight at 4 °C with gentle rotation. The resin was collected on a gravity column, washed once with PBS (pH 7.2) containing 20 mM imidazole, and eluted with PBS (pH 7.2) containing 200 mM imidazole. The eluate was concentrated using a 30 kDa cutoff filter (Millipore) and further purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare) using an ÄKTA Purifier. Fractions containing the target protein were pooled for downstream analyses.

A02-TLM was refolded from inclusion bodies (IBs) by the as previously reported with modifications (38). HLA A*02:01 and β-2-microglobulin (β2m) were cloned into pET28a and pETDuet vectors and expressed separately in E. coli strain BL21 as IBs. The cells were lysed by sonication in 50 mM Tris pH 8.0, 1% v/v Triton X-100 (Sigma Aldrich), 100 mM NaCl, 5 mM MgCl2, 10 mM DTT, and 0.2 mM PMSF, and benzonase. Following lysis, EDTA was added to 10 mM and IBs were pelleted by centrifugation at 10,000 g for 15 minutes. IBs were washed three times in buffer containing 50 mM Tris pH 8.0, 0.5% Triton X-100, 100 mM NaCl, 1 mM Na-EDTA, 1 mM DTT, and 0.2 mM PMSF, and once in the same buffer omitting the Triton X-100. IBs were solubilized in 8 M urea, 50 mM Tris pH 8.0, 0.5 M EDTA, and 1 mM DTT and stored as frozen aliquots.

To refold, 20 mg of TLM peptide (Genscript) in DMSO was added dropwise to 400 ml refolding buffer (100 mM Tris pH 8.0, 2 mM EDTA, 5 M urea, 500 mM L-arginine HCL, 0.2 mM PMSF, 1x protease inhibitor cocktail (Sigma)) stirring at 4°C. This was followed by addition of 0.5 mM oxidized and 5 mM reduced glutathione. Aliquots containing 20 mg of HLA A*02:01 and of β2m were mixed and added dropwise. The refolding reaction was left stirring overnight, then dialyzed for two days at 4° C against 4 L of 100 mM Tris, pH 8.0, with three changes of buffer. The dialyzed protein was filtered with a 0.45 µm PMSF filter and a glass prefilter, concentrated using 10,000 MWCO centrifugal filters (Amicon), injected onto a Mono Q column (Cytiva) and eluted with a gradient to 400 mM NaCl. Protein-containing fractions were further purified on a Superdex S200 Increase column (Cytiva).

A plasmid encoding the β2m-targeting nanobody AD01 (39) was transiently transfected into expi293 cells. Supernatant was harvested after five days, passed over a Ni-IMAC resin (Thermo Scientific) and eluted with 250 mM imidazole in 150 mM NaCl, 10 mM HEPES pH 7.3 (HBS). Protein was concentrated and injected on a Superdex S75 Increase column (Cytiva) in HBS buffer.

Protein complexes for cryoEM were prepared by combining equimolar amounts of Fab, pMHC, and nanobody AD01, and purifying by size exclusion chromatography on a Superdex S200 Increase column (Cytiva). Complex formation and purity was assessed by SDS-PAGE.

To prepare cryo-EM specimens, 3.0 µL of each protein complex was applied to glow-discharged Quantifoil Au R1.2/1.3, 200-mesh (vAB-30) or 300-mesh (vAB-66) grids. Excess liquid was blotted for 1.0 s with filter paper, and grids were plunge-frozen in liquid ethane cooled by liquid nitrogen using a Vitrobot Mark IV (Thermo Fisher Scientific) operated at 8 °C and 100% humidity. Data were collected on a Titan Krios electron microscope (Thermo Fisher Scientific) operating at 300 kV and equipped with a Gatan K3 direct electron detector. Movies of the vAB-30/A01-EVD and vAB-66/A02-TLM complexes were recorded in super-resolution mode with SerialEM (40), at calibrated pixel sizes of 0.4135 Å and 0.94 Å, respectively. Patch motion correction was applied with binning to physical pixel sizes of 0.827 Å and 1.88 Å. Each movie stack contained 50 frames, with total exposures of 60 $e^{-}/Å^{2}$ (vAB-30) or 50 $e^{-}/Å^{2}$ (vAB-66), and defocus values ranged from -1.0 to -2.0 µm (vAB-30) or -0.8 to -2.0 µM (vAB-66). All movies were processed and quality-assessed in cryoSPARC Live.

For the vAB-30/A01-EVD dataset, 5,682 aligned micrographs were retained for further processing in cryoSPARC (41). A total of 4,636,242 particles were initially extracted using a 320/160 binned box. After two rounds of 2D classification, 1,864,082 particles were selected and subjected to non-uniform (NU) refinement, yielding a 3.4 Å reconstruction. Refinement employed multiple references, including one accurate and two biased models (42). Following duplicate removal, 1,317,109 particles remained. Re-extraction with a 320/240 binned box was followed by global CTF refinement, local CTF refinement, and NU-refinement, resulting in a final map at 2.6 Å resolution (**Fig. S14**). CryoEM maps have been deposited in the electron microscopy databank (EMDB) with accession EMD-73460 and movie stacks at the Electron Microscopy Public Image Archive (EMPIAR).

For the vAB-66/A02-TLM, 10,869 aligned micrographs were retained for further processing in cryoSPARC. An initial set of 639,144 blob-picked particles was 2D-classified and used as templates for picking a full 6,762,438 particle set extracted in a 256 pixel box downsampled to 72 pixels. This was reduced through two rounds of 2D classification to 1,807,378 particles, then by two rounds of heterogenous 3D refinement to 880,830 particles, with reextraction to a 256/200 pixel box. Particles were finally extracted without downsampling, followed by global CTF refinement, local CTF refinement, and NU refinement, resulting in a final map at 2.5 Å resolution (**Fig. S15**). CryoEM maps have been deposited in the EMDB with accession EMD-73458 and movie stacks at EMPIAR.

Model building and refinement

For the vAB-30/A01-EVD structure, predicted atomic models of the complex were docked into the cryo-EM density maps, while for the vAB-66/A02-PAP structure model components from PDB IDs 7YV1, 3SOB, 7SRO, and 9NMV were docked into the map, using UCSF ChimeraX (43), followed by manual adjustment and rebuilding in COOT (44). Real-space refinement was carried out in PHENIX with secondary structure and geometry restraints applied (45). Model validation was performed with MolProbity, including assessment of Ramachandran statistics and overall scores (Table S4) (46). Coordinates have been deposited in the protein databank with accessions 9YTF (vAB-30/A01-EVD) and 9YTD (vAB-66/A02-PAP).

References

1. M. M. Davis, P. J. Bjorkman, T-cell antigen receptor genes and T-cell recognition. Nature 334, 395–402 (1988).

2. K. C. Garcia, M. Degano, R. L. Stanfield, A. Brunmark, M. R. Jackson, P. A. Peterson, L. Teyton, I. A. Wilson, An αβ T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex. Science 274, 209–219 (1996).

3. D. Jung, C. Giallourakis, R. Mostoslavsky, F. W. Alt, Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annu. Rev. Immunol. 24, 541–570 (2006).

4. K. M. Ashby, K. A. Hogquist, A guide to thymic selection of T cells. Nat. Rev. Immunol. 24, 103–117 (2024).

5. V. Zoete, M. Irving, M. Ferber, M. A. Cuendet, O. Michielin, Structure-Based, Rational Design of T Cell Receptors. Front. Immunol. 4, 268 (2013).

6. B. G. Pierce, L. M. Hellman, M. Hossain, N. K. Singh, C. W. Vander Kooi, Z. Weng, B. M. Baker, Computational design of the affinity and specificity of a therapeutic T cell receptor. PLoS Comput. Biol. 10, e1003478 (2014).

7. J. N. Haidar, B. Pierce, Y. Yu, W. Tong, M. Li, Z. Weng, Structure-based design of a T-cell receptor leads to nearly 100-fold improvement in binding affinity for pepMHC. Proteins 74, 948–960 (2009).

8. L. M. Hellman, K. C. Foley, N. K. Singh, J. A. Alonso, T. P. Riley, J. R. Devlin, C. M. Ayres, G. L. J. Keller, Y. Zhang, C. W. Vander Kooi, M. I. Nishimura, B. M. Baker, Improving T Cell Receptor On-Target Specificity via Structure-Guided Design. Mol. Ther. 27, 300–313 (2019).

9. D. Karthikeyan, S. N. Bennett, A. G. Reynolds, B. G. Vincent, A. Rubinsteyn, Conditional generation of real antigen-specific T cell receptor sequences. Nat. Mach. Intell. 7, 1494–1509 (2025).

10. J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, N. Hanikel, S. J. Pellock, A. Courbet, W. Sheffler, J. Wang, P. Venkatesh, I. Sappington, S. V. Torres, A. Lauko, V. De Bortoli, E. Mathieu, S. Ovchinnikov, R. Barzilay, T. S. Jaakkola, F. DiMaio, M. Baek, D. Baker, De novo design of protein structure and function with RFdiffusion. Nature 620, 1089–1100 (2023).

11. M. Pacesa, L. Nickel, C. Schellhaas, J. Schmidt, E. Pyatova, L. Kissling, P. Barendse, J. Choudhury, S. Kapoor, A. Alcaraz-Serna, Y. Cho, K. H. Ghamary, L. Vinué, B. J. Yachnin, A. M. Wollacott, S. Buckley, A. H. Westphal, S. Lindhoud, S. Georgeon, C. A. Goverde, G. N. Hatzopoulos, P. Gönczy, Y. D. Muller, G. Schwank, D. C. Swarts, A. J. Vecchio, B. L. Schneider, S. Ovchinnikov, B. E. Correia, One-shot design of functional protein binders with BindCraft. Nature 646, 483–492 (2025).

12. J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N. P. King, D. Baker, Robust deep learning–based protein sequence design using ProteinMPNN. Science 378, 49–56 (2022).

13. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021).

14. P. Bradley, Structure-based prediction of T cell receptor:peptide-MHC interactions. Elife 12 (2023).

15. A. Motmaen, J. Dauparas, M. Baek, M. H. Abedi, D. Baker, P. Bradley, Peptide-binding specificity prediction using fine-tuned protein structure prediction networks. Proc. Natl. Acad. Sci. U. S. A. 120, e2216697120 (2023).

16. N. R. Bennett, J. L. Watson, R. J. Ragotte, A. J. Borst, D. L. See, C. Weidle, R. Biswas, E. L. Shrock, P. J. Y. Leung, B. Huang, I. Goreshnik, R. Ault, K. D. Carr, B. Singer, C. Criswell, D. Vafeados, M. G. Sanchez, H. M. Kim, S. V. Torres, S. Chan, D. Baker, Atomically accurate de novo design of single-domain antibodies, bioRxiv (2024)p. 2024.03.14.585103.

17. M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. C. Garcia, N. V. Grishin, P. D. Adams, R. J. Read, D. Baker, Accurate prediction of protein structures and interactions using a three-track neural network. Science 373, 871–876 (2021).

18. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill,

D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, doi: 10.1038/s41586-024-07487-w (2024).

19. T. Arvedson, J. M. Bailis, T. Urbig, J. L. Stevens, Considerations for design, manufacture, and delivery for effective and safe T-cell engager therapies. Curr. Opin. Biotechnol. 78, 102799 (2022).

20. A. Segués, S. Huang, A. Sijts, P. Berraondo, D. M. Zaiss, Opportunities and challenges of bi-specific antibodies. Int. Rev. Cell Mol. Biol. 369, 45–70 (2022).

21. A. Y. Chang, R. S. Gejman, E. J. Brea, C. Y. Oh, M. D. Mathias, D. Pankov, E. Casey, T. Dao, D. A. Scheinberg, Opportunities and challenges for TCR mimic antibodies in cancer therapy. Expert Opin. Biol. Ther. 16, 979–987 (2016).

22. X. Yang, D. Nishimiya, S. Löchte, K. M. Jude, M. Borowska, C. S. Savvides, M. Dougan, L. Su, X. Zhao, J. Piehler, K. C. Garcia, Facile repurposing of peptide-MHC-restricted antibodies for cancer immunotherapy. Nat. Biotechnol. 41, 932–943 (2023).

23. M. Yarmarkovich, Q. F. Marshall, J. M. Warrington, R. Premaratne, A. Farrel, D. Groff, W. Li, M. di Marco, E. Runbeck, H. Truong, J. S. Toor, S. Tripathi, S. Nguyen, H. Shen, T. Noel, N. L. Church, A. Weiner, N. Kendsersky, D. Martinez, R. Weisberg, M. Christie, L. Eisenlohr, K. R. Bosse, D. S. Dimitrov, S. Stevanovic, N. G. Sgourakis, B. R. Kiefel, J. M. Maris, Targeting of intracellular oncoproteins with peptide-centric CARs. Nature 623, 820–827 (2023).

24. G. Gaud, R. Lesourne, P. E. Love, Regulatory mechanisms in T cell receptor signalling. Nat. Rev. Immunol. 18, 485–497 (2018).

25. B. Liu, N. F. Greenwood, J. E. Bonzanini, A. Motmaen, J. Meyerberg, T. Dao, X. Xiang, R. Ault, J. Sharp, C. Wang, G. M. Visani, D. K. Vafeados, N. Roullier, A. Nourmohammad, D. A. Scheinberg, K. C. Garcia, D. Baker, Design of high-specificity binders for peptide-MHC-I complexes. Science 389, 386–391 (2025).

26. K. H. Johansen, D. S. Wolff, B. Scapolo, M. L. Fernández-Quintero, C. Risager Christensen, J. R. Loeffler, E. Rivera-de-Torre, M. D. Overath, K. Kjærgaard Munk, O. Morell, M. C. Viuff, I. Lacunza, A. T. Damm Englund, M. Due, A. Gharpure, S. Forli, C. Rodriguez Pardo, T. Tamhane, E. Qingjie Andersen, K. Haldrup Björnsson, J. S. Fernandes, L. F. Voss, S. Thumtecho, A. B. Ward, M. Ormhøj, S. Reker Hadrup, T. P. Jenkins, De novo-designed pMHC

binders facilitate T cell-mediated cytotoxicity toward cancer cells. Science 389, 380–385 (2025).

27. K. D. Householder, X. Xiang, K. M. Jude, A. Deng, M. Obenaus, Y. Zhao, S. C. Wilson, X. Chen, N. Wang, K. C. Garcia, De novo design and structure of a peptide-centric TCR mimic binding module. Science 389, 375–379 (2025).

28. L. V. Sibener, R. A. Fernandes, E. M. Kolawole, C. B. Carbone, F. Liu, D. McAffee, M. E. Birnbaum, X. Yang, L. F. Su, W. Yu, S. Dong, M. H. Gee, K. M. Jude, M. M. Davis, J. T. Groves, W. A. Goddard 3rd, J. R. Heath, B. D. Evavold, R. D. Vale, K. C. Garcia, Isolation of a structural mechanism for uncoupling T cell receptor signaling from peptide-MHC binding. Cell 174, 672–687.e27 (2018).

29. J. J. Adams, S. Narayanan, B. Liu, M. E. Birnbaum, A. C. Kruse, N. A. Bowerman, W. Chen, A. M. Levin, J. M. Connolly, C. Zhu, D. M. Kranz, K. C. Garcia, T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. Immunity 35, 681–693 (2011).

30. P. Zareie, C. Szeto, C. Farenc, S. D. Gunasinghe, E. M. Kolawole, A. Nguyen, C. Blyth, X. Y. X. Sng, J. Li, C. M. Jones, A. J. Fulcher, J. R. Jacobs, Q. Wei, L. Wojciech, J. Petersen, N. R. J. Gascoigne, B. D. Evavold, K. Gaus, S. Gras, J. Rossjohn, N. L. La Gruta, Canonical T cell receptor docking on peptide-MHC is essential for T cell signaling. Science 372 (2021).

31. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The protein data bank. Nucleic Acids Res. 28, 235–242 (2000).

32. J. Dunbar, K. Krawczyk, J. Leem, T. Baker, A. Fuchs, G. Georges, J. Shi, C. M. Deane, SAbDab: the structural antibody database. Nucleic Acids Res. 42, D1140–6 (2014).

33. S. Schattgen, K. Vegesana, W. D. Hazelton, A. Minervina, S. Valkiers, K. Slowikowski, N. Smith, MGH COVID-19 Team, A.-C. Villani, P. G. Thomas, P. Bradley, Diverse modes of T cell receptor sequence convergence define unique functional and cellular phenotypes, bioRxivorg (2025)p. 2025.05.31.657155.

34. M.-P. Lefranc, V. Giudicelli, P. Duroux, J. Jabado-Michaloud, G. Folch, S. Aouinti, E. Carillon, H. Duvergey, A. Houles, T. Paysan-Lafosse, S. Hadi-Saljoqi, S. Sasorith, G. Lefranc, S. Kossida, IMGT®, the international ImMunoGeneTics information system® 25 years on. Nucleic Acids Res. 43, D413–22 (2015).

35. L. Cao, B. Coventry, I. Goreshnik, B. Huang, W. Sheffler, J. S. Park, K. M. Jude, I. Marković, R. U. Kadam, K. H. G. Verschueren, K. Verstraete, S. T. R. Walsh, N. Bennett, A. Phal, A. Yang, L. Kozodoy, M. DeWitt, L. Picton, L. Miller, E.-M. Strauch, N. D. DeBouver, A. Pires, A. K. Bera, S. Halabiya, B. Hammerson, W. Yang, S. Bernard, L. Stewart, I. A. Wilson,

H. Ruohola-Baker, J. Schlessinger, S. Lee, S. N. Savvides, K. C. Garcia, D. Baker, Design of protein-binding proteins from the target structure alone. Nature 605, 551–560 (2022).

36. I. Lin, P. B. Rupert, K. Pilat, R. O. Ruff, D. J. Friend, M. K. Chan, M. Clarke, B. G. Hoffstrom, J. Carter, S. Meshinchi, A. D. Bandaranayake, C. Mehlin, J. M. Olson, R. K. Strong, C. E. Correnti, Novel mesothelin antibodies enable crystallography of the intact mesothelin ectodomain and engineering of potent, T cell-engaging bispecific therapeutics. Front. Drug Discov. (Lausanne) 3, 1216516 (2023).

37. Y. Y. Yu, N. Netuschil, L. Lybarger, J. M. Connolly, T. H. Hansen, Cutting edge: single-chain trimers of MHC class I molecules form stable structures that potently stimulate antigen-specific T cells and B cells. The Journal of Immunology 168, 3145–3149 (2002).

38. D. N. Garboczi, D. T. Hung, D. C. Wiley, HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. Proc. Natl. Acad. Sci. U. S. A. 89, 3429–3433 (1992).

39. K. A. K. Finton, P. B. Rupert, D. J. Friend, A. Dinca, E. S. Lovelace, M. Buerger, D. V. Rusnac, U. Foote-McNabb, W. Chour, J. R. Heath, J. S. Campbell, R. H. Pierce, R. K. Strong, Effects of HLA single chain trimer design on peptide presentation and stability. Front. Immunol. 14, 1170462 (2023).

40. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. 152, 36–51 (2005).

41. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296 (2017).

42. N. Wang, X. Jiang, S. Zhang, A. Zhu, Y. Yuan, H. Xu, J. Lei, C. Yan, Structural basis of human monocarboxylate transporter 1 inhibition by anti-cancer drug candidates. Cell 184, 370–383.e13 (2021).

43. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82 (2021).

44. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 (2004).

45. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: a

comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010).

46. V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 (2010).

47. Kumamoto T, Maurinot F, Barry-Martinet R, Vaslin C, Vandormael-Pournin S, Le M, Lerat M, Niculescu D, Cohen-Tannoudji M, Rebsam A, Loulier K. Direct readout of neural stem cell transgenesis with an integration-coupled gene expression switch. Neuron. 2020 Aug 19;107(4):617-30.

Acknowledgments:

Authors: Amir Motmaen[1,2,3], Kevin M. Jude[4,5]†, Nan Wang[4,5]†, Anastasia Minervina[6], David Feldman[7], Mauriz A. Lichtenstein[1,8], Abishai Ebenezer[1,9], Colin Correnti[1], Paul G. Thomas[6], K. Christopher Garcia[4,5]*, David Baker[1,2,10]*, Philip Bradley[1,2,11]*

Affiliations:

[1]Institute for Protein Design, University of Washington, Seattle, WA, USA.

[2]Department of Biochemistry, University of Washington, Seattle, WA, USA

[3]Bioengineering Graduate Program, University of Washington, Seattle, WA, USA

[4]Departments of Molecular and Cellular Physiology and Structural Biology, Stanford University School of Medicine, Stanford, CA, USA

[5]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA.

[6]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

[7]Inceptive, Berlin, Germany

[8]Institute for Chemistry and Biochemistry, Freie Universität Berlin, Berlin, Germany

[9]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, USA

[10]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA

[11]Herbold Computational Biology Program, Fred Hutchinson Cancer Center, Seattle, WA, USA

† These authors contributed equally to this work.

* Corresponding author. Email: kcgarcia@stanford.edu (K.C.G), dabaker@uw.edu (D.B.), pbradley@fredhutch.org (P.B.)

We acknowledge K. VanWormer, L. Goldschmidt, and Fred Hutch Scientific Computing for their outstanding lab and computational support. We benefitted from discussions with and advice from B. Vogelstein, T. Nichakawade, T. Schumacher, J. Watson, N. Bennett. We thank R. Yan, J. Jung, X. and S.Yang from the Howard Hughes Medical Institute Janelia Research Campus Cryo-EM Facility for their assistance with data collection. We thank J. Sims, M. Bauer, and M. Glögl for helping with SPR. We thank the NIH Tetramer Core Facility

(NIH Contract 75N93020D00005 and RRID:SCR_026557) for providing some of the pMHC tetramers and monomers used in this manuscript.

Author contributions:

Conceptualization: PB, AM, DB

Experimental strategy: AM, DB, KCG, AMi, DF, CC, KMJ, NW

Experimental testing: AM, KMJ, NW

Software design: PB

Software testing and refinement: PB, AM, MAL

Protein design calculations: PB, AM, MAL, AE

Funding Acquisition: DB, PB, KCG, PGT

Writing – original draft: PB, AM, KMJ, NW, DB

Writing – review & editing: PB, AM, KMJ, NW, DB, AMi, KCG, PGT, MAL

Funding:

National Institutes of Health grant R35 GM141457 (P.B.).

National Institutes of Health grant R01 AI136514 (P.B., P.G.T.).

National Institutes of Health grant R01 AI103867 (K.C.G.)

Parker Institute for Cancer Immunotherapy (K.C.G.)

Cancer Research UK CGC Team MATCHMAKERS CGCATF-2023/100006 (N.W., K.J., K.C.G.)

Cancer Research UK CGC Team MATCHMAKERS CGCATF-2023/100008 (A.M., D.B.)

Howard Hughes Medical Institute (K.C.G, D.B.)

National Institutes of Health grant ORIP S10OD028685 (P.B.)

Competing interests:

A provisional patent (US Provisional Patent Application Serial No. 63/910,643) covering TCR and antibody sequences presented in this paper has been filed by the Fred Hutchinson Cancer Center. K.C.G. is a consultant for Xaira Therapeutics.

Data and materials availability:

Open-source Python scripts for running the ADAPT design pipeline are freely available under an MIT license at https://github.com/phbradley/ADAPT . Fine-tuned AF2 and RF2 parameter sets along with template and CDR3 database files can be downloaded from the Zenodo repository with DOI 10.5281/zenodo.17488258. Designed TCR and antibody sequences are provided in Data S1. Protein structure coordinates have been deposited in the RCSB Protein Databank with accessions 9YTF (vAB-30/A01-EVD) and 9YTD (vAB-66/A02-PAP).

Figures

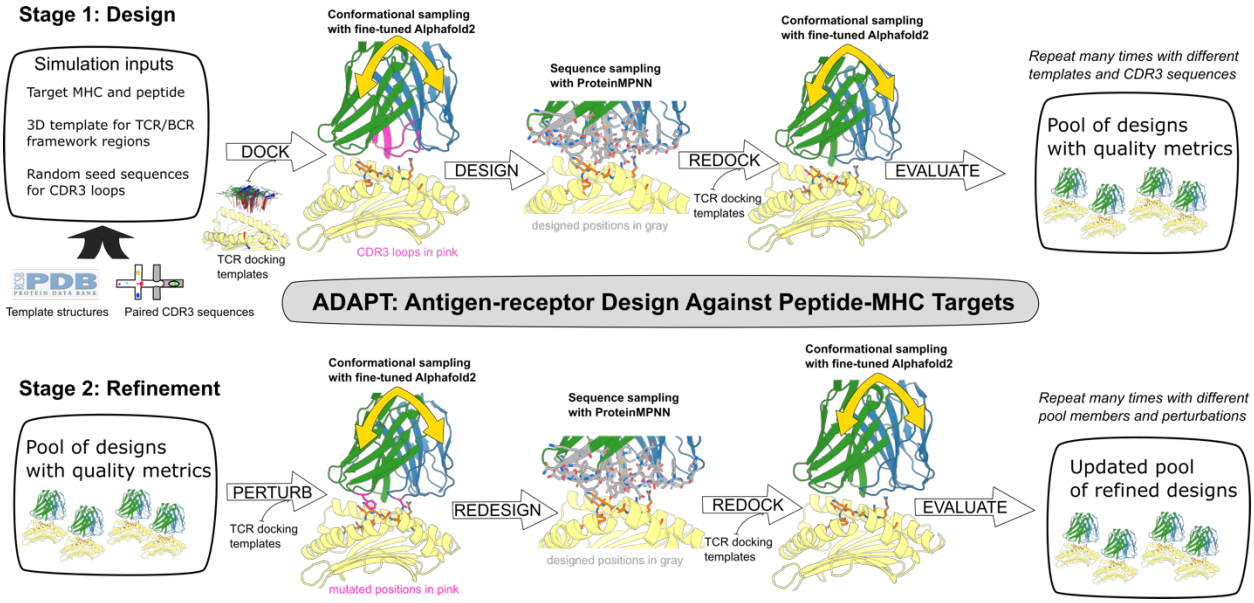


Fig. 1. The ADAPT pipeline. In the design stage (top), many independent design simulations are conducted with different template structures for the TCR or antibody framework regions and different starting CDR3 sequences. Together with the pMHC target, these inputs determine the complete sequence of the system, which is modeled with a version of Alphafold2 (AF2) that was fine-tuned on TCR:pMHC structures and binding data ("Dock" arrow). Sampling of canonical docking geometries is encouraged by providing AF2 with four generic TCR:pMHC docking orientations spanning the observed docking modes in native structures ("TCR docking templates"). After docking, the sequence of the CDR1, 2, and 3 loops is optimized with the ProteinMPNN algorithm ("Design"). The structural model is then updated ("Redock"), and design quality is assessed using AF2 and RoseTTAFold2 ("Evaluate"). In the refinement stage (bottom), the highest quality designs are iteratively refined by a genetic algorithm that includes random interface mutations followed by AF2 redocking ("Perturb") and CDR sequence redesign ("Redesign").

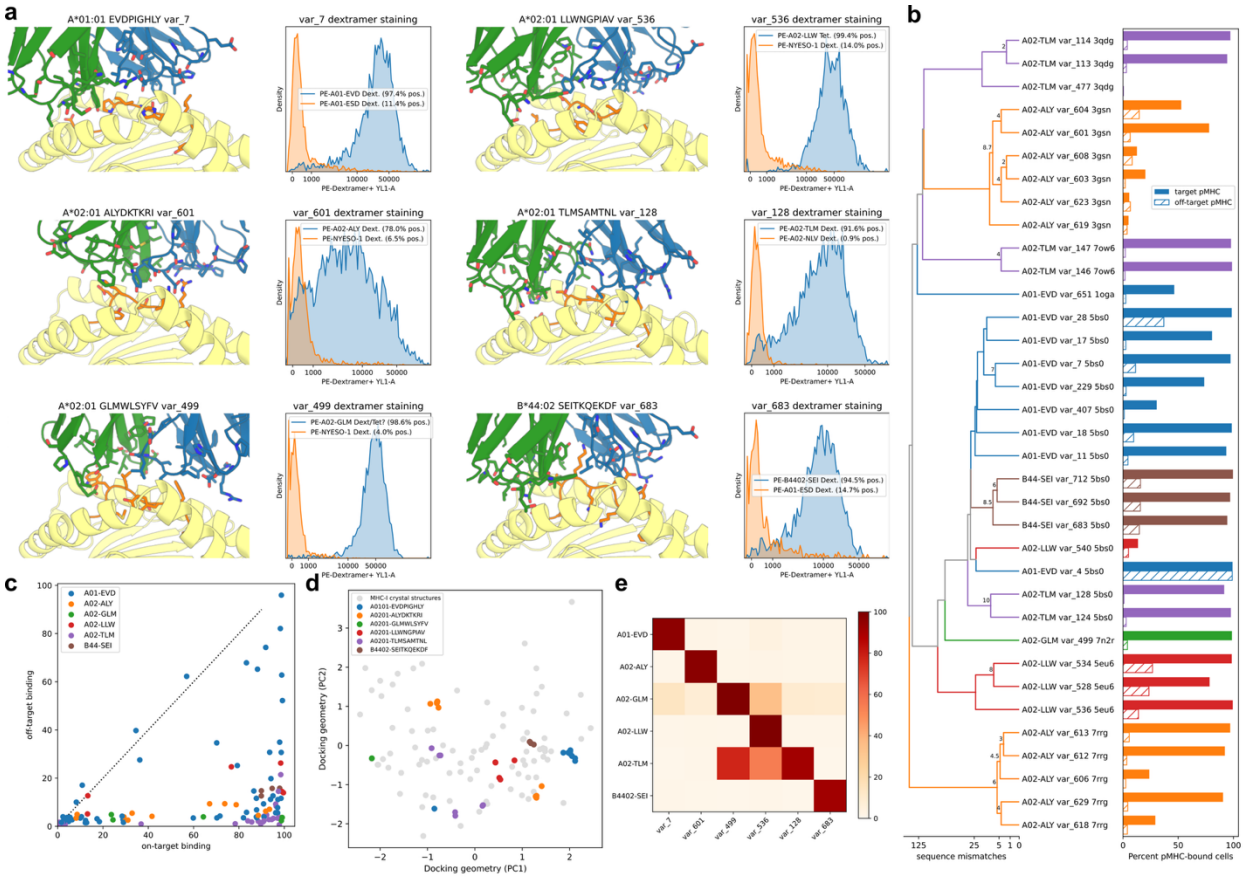


Fig. 2. Designed TCRs bind their targets. (A) For six representative designed TCRs, one from each successful pMHC target, are shown design models (MHC in yellow, peptide in orange, TCRα in green, and TCRβ in blue) and flow cytometry binding histograms for on-target (blue) and off-target (orange) pMHC multimers. **(B)** Summary bar plots of on- and off-target binding for 35 selected designs along with a dendrogram of their sequence relationships. **(C)** Scatter plot of on- vs off-target binding for a larger set of successful and failed designs, colored by pMHC target. **(D)** Docking geometry landscape of native TCR:pMHC interfaces (gray points) and designed TCR interfaces (colored by pMHC target) visualized by projecting the 6-dimensional rigid-body transform space to its top 2 principal components (PC1 and PC2). **(E)** Heatmap showing all-vs-all flow-cytometric binding data (percent cells bound by pMHC target) when the six representative designs are paired with all six pMHC targets.

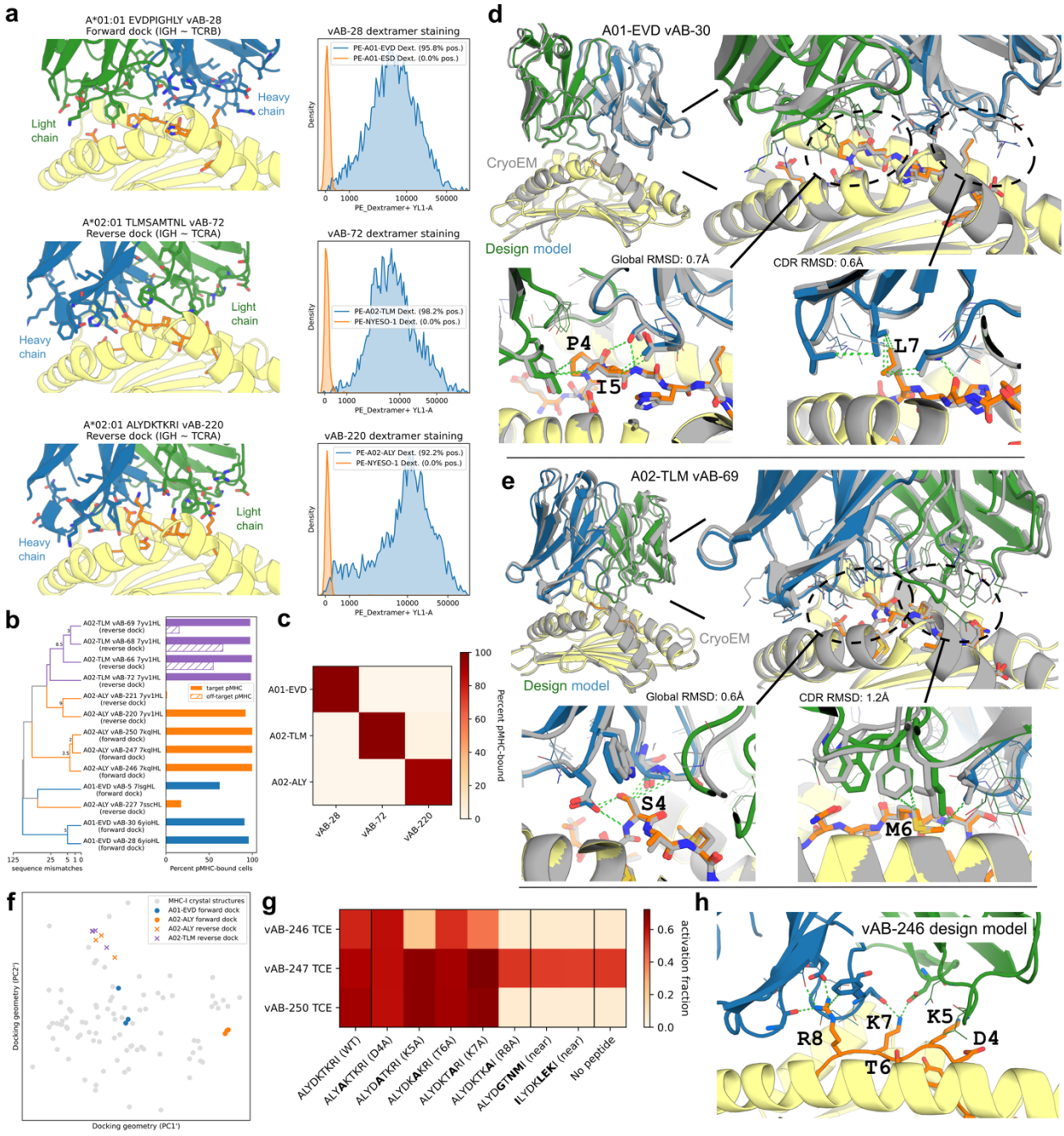

a
A*01:01 EVDPIGHLY vAB-28
Forward dock (IGH ~ TCRB)
Heavy chain
Light chain
vAB-28 dextramer staining
PE-A01-EVD Dext. (95.8% pos.)
PE-A01-ESD Dext. (0.0% pos.)
Density
PE_Dextramer+ YL1-A
A*02:01 TLMSAMTNL vAB-72
Reverse dock (IGH ~ TCRA)
Light chain
Heavy chain
vAB-72 dextramer staining
PE-A02-TLM Dext. (98.2% pos.)
PE-NYESO-1 Dext. (0.0% pos.)
A*02:01 ALYDKTKRI vAB-220
Reverse dock (IGH ~ TCRA)
Light chain
Heavy chain
vAB-220 dextramer staining
PE-A02-ALY Dext. (92.2% pos.)
PE-NYESO-1 Dext. (0.0% pos.)
b
A02-TLM vAB-69 7yv1HL (reverse dock)
A02-TLM vAB-68 7yv1HL (reverse dock)
A02-TLM vAB-66 7yv1HL (reverse dock)
A02-TLM vAB-72 7yv1HL (reverse dock)
A02-ALY vAB-221 7yv1HL (reverse dock)
A02-ALY vAB-220 7yv1HL (reverse dock)
A02-ALY vAB-250 7kqlHL (forward dock)
A02-ALY vAB-247 7kqlHL (forward dock)
A02-ALY vAB-246 7kqlHL (forward dock)
A01-EVD vAB-5 7lsgHL (forward dock)
A02-ALY vAB-227 7sscHL (reverse dock)
A01-EVD vAB-30 6yioHL (forward dock)
A01-EVD vAB-28 6yioHL (forward dock)
target pMHC
off-target pMHC
sequence mismatches
Percent pMHC-bound cells
c
A01-EVD
A02-TLM
A02-ALY
vAB-28
vAB-72
vAB-220
Percent pMHC-bound
d
A01-EVD vAB-30
CryoEM
Design model
Global RMSD: 0.7Å
CDR RMSD: 0.6Å
P4
I5
L7
e
A02-TLM vAB-69
CryoEM
Design model
Global RMSD: 0.6Å
CDR RMSD: 1.2Å
S4
M6
f
MHC-I crystal structures
A01-EVD forward dock
A02-ALY forward dock
A02-ALY reverse dock
A02-TLM reverse dock
Docking geometry (PC2')
Docking geometry (PC1')
g
vAB-246 TCE
vAB-247 TCE
vAB-250 TCE
ALYDKTKRI (WT)
ALYAKTKRI (D4A)
ALYDATKRI (K5A)
ALYDKAKRI (T6A)
ALYDKTARI (K7A)
ALYDKTKAI (R8A)
ALYDGTNMI (near)
ILYDKLEKI (near)
No peptide
activation fraction
h
vAB-246 design model
R8
K7
K5
D4
T6

Fig. 3. Designed antibodies bind their targets. (A) For three representative designed antibodies, one from each pMHC target attempted, are shown design models and flow cytometry binding histograms for on-target (blue) and off-target (orange) pMHC multimers. **(B)** Summary bar plots of on- and off-target binding for 13 selected designs along with a dendrogram of their sequence relationships. **(C)** Heatmap showing all-vs-all flow-cytometric binding data (percent cells bound by pMHC target) when the three representative antibodies are paired with all three pMHC targets. **(D)** Structural comparison of cryoEM structure (in gray) and design model (MHC in yellow, peptide in orange, heavy chain in blue, and light chain in green) for designed antibody vAB-30 bound to the HLA-A*01:01-EVDPIGHLY pMHC complex. **(E)** Structural comparison of cryoEM structure and design model (same colors as in panel d) for designed antibody vAB-66 bound to the HLA-A*02:01-TLMSAMTNL pMHC complex. **(F)** Docking geometry landscape of native TCR:pMHC interfaces (gray points) and designed antibody:pMHC interfaces (colored by pMHC target) visualized by projecting the 6-dimensional rigid-body transform space to its top 2 principal components (PC1' and PC2') after excluding rotations about the antibody pseudosymmetry axis. The excluded rigid-body dimension differentiates between forward-docking antibodies (disks) and reverse-docking antibodies (x's); canonical native TCR:pMHC complexes are forward docking). **(G)** Heatmap of induced activation for bi-specific T cell engager (BiTE) constructs incorporating three designed antibodies (rows) in the presence of presenter cells pulsed with various target peptide variants (columns). **(H)** Design model for vAB-246 antibody bound to A02-ALY pMHC shows more contacts at peptide positions whose mutation to alanine reduces T cell activation in the presence of the vAB-246 BiTE construct (R8 and K7 versus T6 and D4).

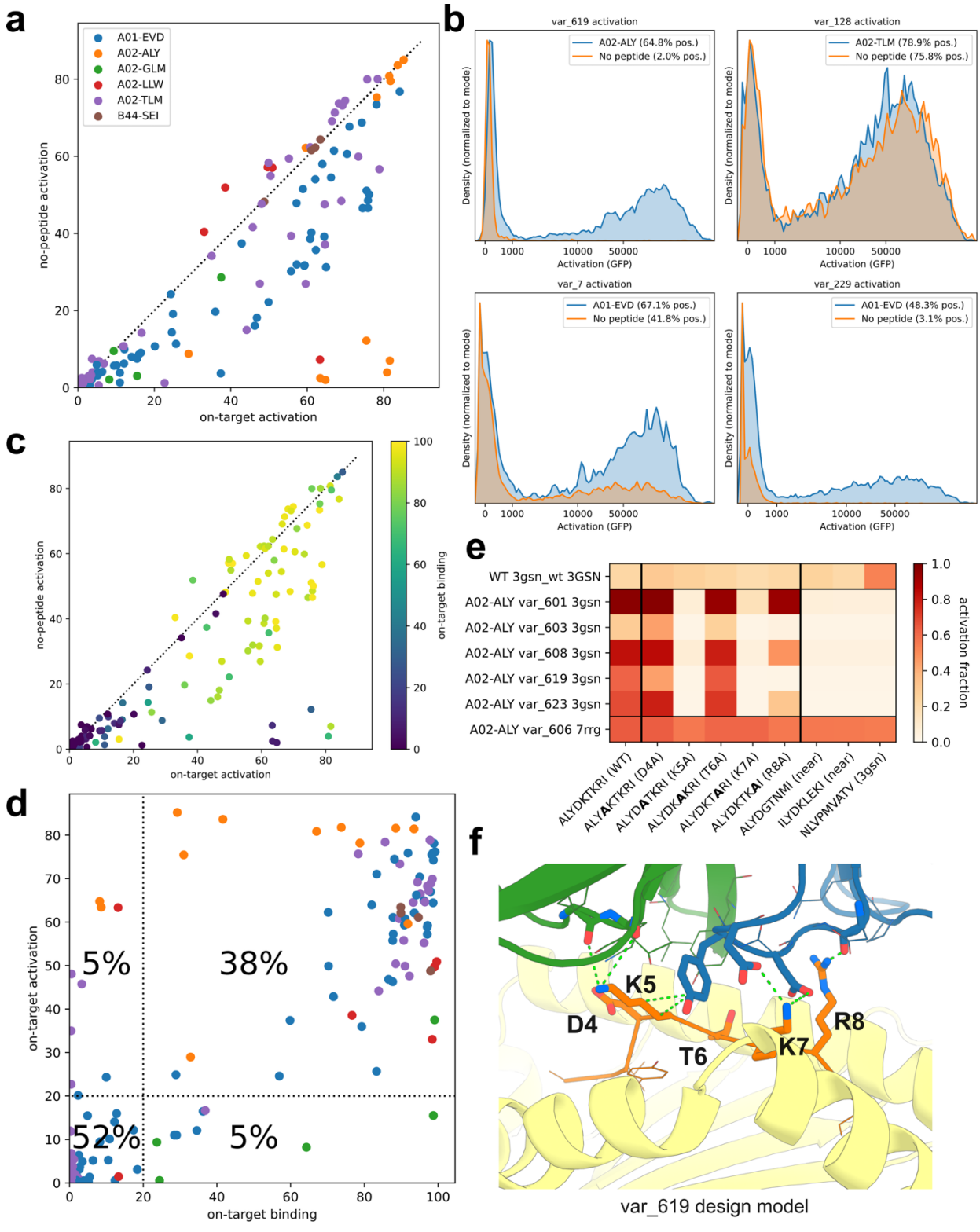

a
A01-EVD
A02-ALY
A02-GLM
A02-LLW
A02-TLM
B44-SEI
no-peptide activation
on-target activation
b
var_619 activation
A02-ALY (64.8% pos.)
No peptide (2.0% pos.)
var_128 activation
A02-TLM (78.9% pos.)
No peptide (75.8% pos.)
var_7 activation
A01-EVD (67.1% pos.)
No peptide (41.8% pos.)
var_229 activation
A01-EVD (48.3% pos.)
No peptide (3.1% pos.)
Density (normalized to mode)
Activation (GFP)
c
no-peptide activation
on-target activation
on-target binding
d
5%
38%
52%
5%
on-target activation
on-target binding
e
WT 3gsn_wt 3GSN
A02-ALY var_601 3gsn
A02-ALY var_603 3gsn
A02-ALY var_608 3gsn
A02-ALY var_619 3gsn
A02-ALY var_623 3gsn
A02-ALY var_606 7rrg
ALYDKTKRI (WT)
ALYAKTKRI (D4A)
ALYDATKRI (K5A)
ALYDKAKRI (T6A)
ALYDKTARI (K7A)
ALYDKTKAI (R8A)
ALYDGTNMI (near)
ILYDKLEKI (near)
NLVPMVATV (3gsn)
activation fraction
f
K5
D4
T6
K7
R8
var_619 design model

Fig. 4. Designed TCRs activate T cells. (A) Scatter plot (colored by pMHC target) of T cell activation (percentage of transgenic T cells expressing a GFP activation marker) measured by flow cytometry for T cell populations expressing individual TCR designs in the presence of presenter cells pulsed with target peptide (x-axis) or presenter cells alone (y-axis). **(B)** Flow cytometry cell density histograms of activation marker expression for four designs (named in panel titles) with varying levels of off-target activation in the absence of target peptide. **(C)** Scatter plot of T cell activation in the presence or absence of pulsed target peptide colored by a measure of on-target binding strength for the corresponding TCR design (percentage TCR-transgenic cells bound by multimerized pMHC target). **(D)** Scatter plot of on-target binding (percentage multimer-bound cells) versus on-target activation (percentage activation-marker expressing cells) shows positive correlation across different TCR designs (colored by pMHC target). **(E)** T cell activation percentage for different A02-ALY-targeting designed TCRs or a wildtype antiviral TCR (rows) as a function of pulsed peptide identity (columns). Pulsed peptides include (left to right) the ALY target peptide from TdT, 5 alanine variants at positions 4-8, similar peptides from the human proteome, and the viral peptide target of the TCR template used for several of the designs (PDBID 3gsn). Rows labeled with the target, design name, and TCR template on which the design was based (3gsn or 7rrg) or "3gsn_wt" for the wildtype antiviral TCR from the 3gsn PDB structure. **(F)** The design model of the var_619:pMHC interface shows that peptide positions where alanine mutations disrupt activation (K5/K7/R8) tend to have more interface contacts than positions like T6 where mutations are tolerated.

Tables

Table 1. TCR design targets and outcomes

MHC	Peptide	Success rate	Peptide source	Successful templates[3]	Template pMHC structure available
A*01:01	EVDPIGHLY	16.7% (clonal)	MAGE-A3 (CAA[1])	5BS0[4] [A*01:01] 1OGA [A*02:01]	YES
A*02:01	TLMSAMTNL	20.0% (clonal)	PAP (CAA)	3QDG [A*02:01] 5BS0 [A*01:01] 7OW6 [A*11:01]	YES
A*02:01	ALYDKTKRI	0.3% (library)	TdT (CAA)	3GSN [A*02:01] 7RRG [A*03:01]	NO
A*02:01	GLMWLSYFV	4.2% (clonal)	SARS (viral)	7N2R [B*27:05]	YES
A*02:01	HMTEVVRHC	0% (library)	p53 (NEO[2])	none	YES
A*02:01	LLWNGPIAV	20.8% (clonal)	YFV (viral)	5EU6 [A*02:01] 5BS0 [A*01:01]	YES
A*03:01	ALHGGWTTK	0% (clonal)	PIK3CA (NEO)	none	YES
A*11:01	VVVGADGVGK	0% (clonal)	KRAS (NEO)	none	YES
B*44:02	SEITKQEKDF	0.1% (library)	PIK3CA (NEO)	5BS0 [A*01:01]	NO

[1]CAA = cancer-associated antigen

[2]NEO = cancer neoepitope

[3]The native MHC restriction of each template is given in brackets

[4]The 5BS0 template is a MAGE-A3 reactive TCR, which may have increased success rates for this target/template combination.

Table 2. Antibody design targets and outcomes

MHC	peptide	Success rate	Successful templates[1]	Successful binding orientations	RMSDs to cryoEM structure[2]
A*01:01	EVDPIGHLY	6.25% (clonal)	6YIO [CD25] 7LSG [TBEV EDIII]	Forward	vAB-30: 0.7Å, 0.6Å
A*02:01	TLMSAMTNL	8.33% (clonal)	7YV1 [K-Ras]	Reverse	vAB-66: 0.6Å, 1.2Å
A*02:01	ALYDKTKRI	5.21% (clonal)	7KQL [Tim-3] 7SSC [HCMV peptide] 7YV1 [K-Ras]	Forward + Reverse	-

[1] The target of the wild-type antibody template is given in brackets. None of the wild-type targets are pMHC complexes.

[2] The global Ca-RMSD and the Ca-RMSD for the CDR loops after aligning on the MHC

Supplementary Material for Targeting peptide–MHC complexes with designed T cell receptors and antibodies

Supplementary Figures

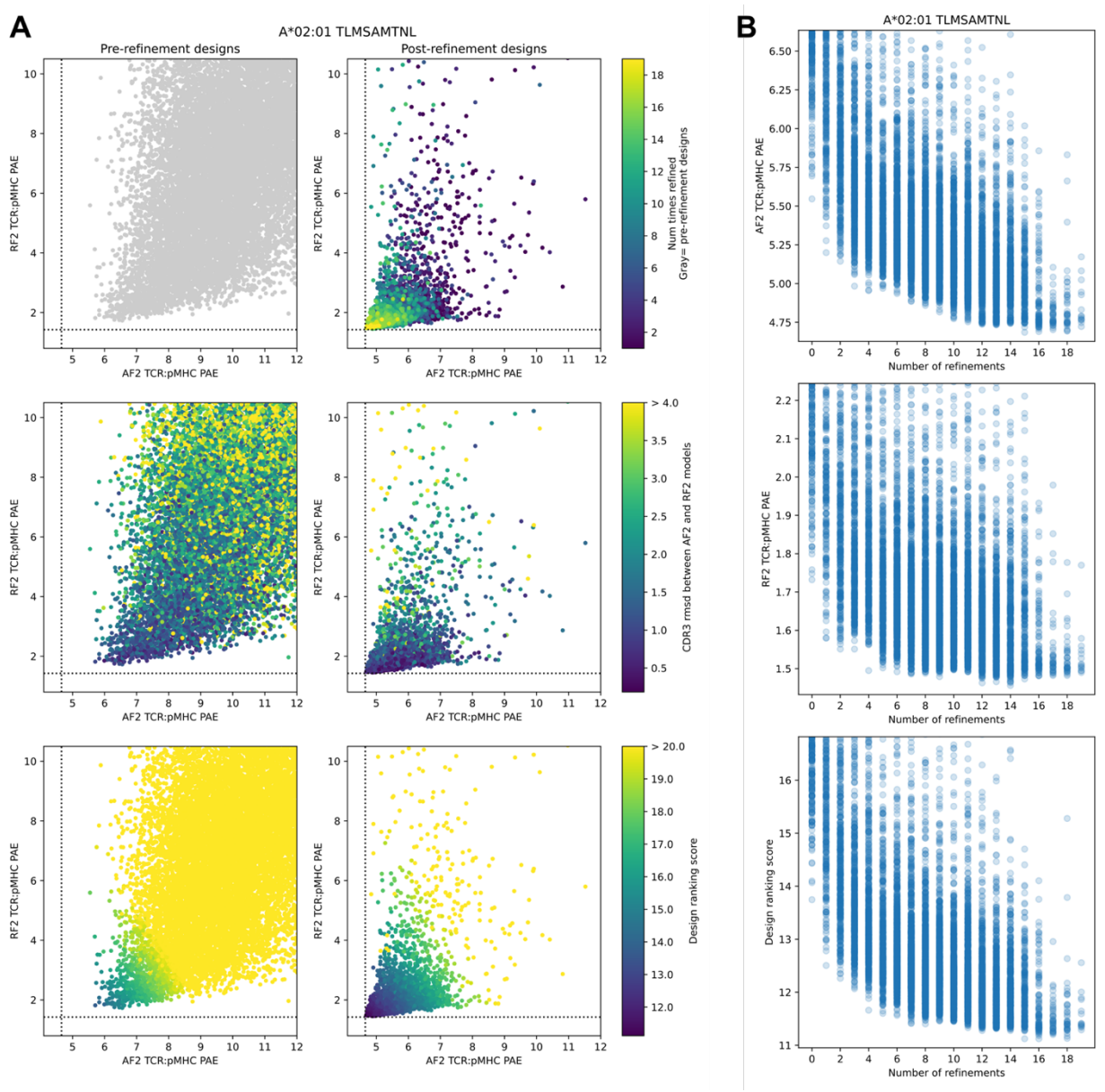


Fig. S1. Design quality metrics before and after refinement. (A) Scatter plots of AF2 predicted aligned error (PAE) between TCR and pMHC (x-axis) versus RF2 PAE between TCR and pMHC (y-axis) for pre- and post-refinement designs (left and right columns, respectively). Each dot represents an individual design. Scatter plots are colored by the number of refinement cycles (top row), the CDR3 RMSD between the AF2 and RF2 models of the design (middle row, calculated after superimposing on the MHC), and the design ranking score (bottom row), which is a weighted combination of the two inter-PAE values and the CDR3 RMSD (see Methods). **(B)** Scatter plots of design metrics (y-axis) as a function of the number of refinement cycles (x-axis); pre-refinement designs are plotted at x=0.

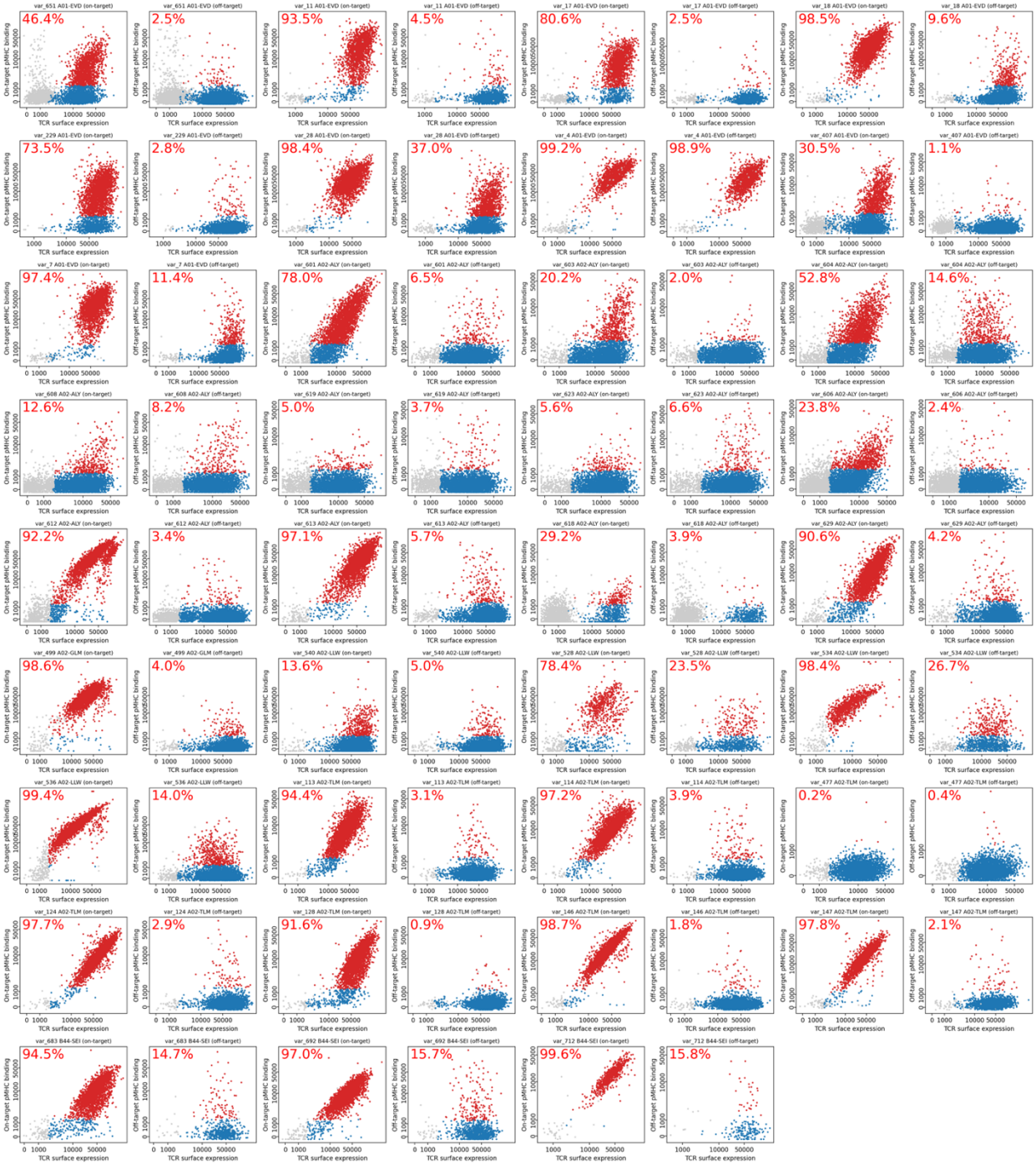
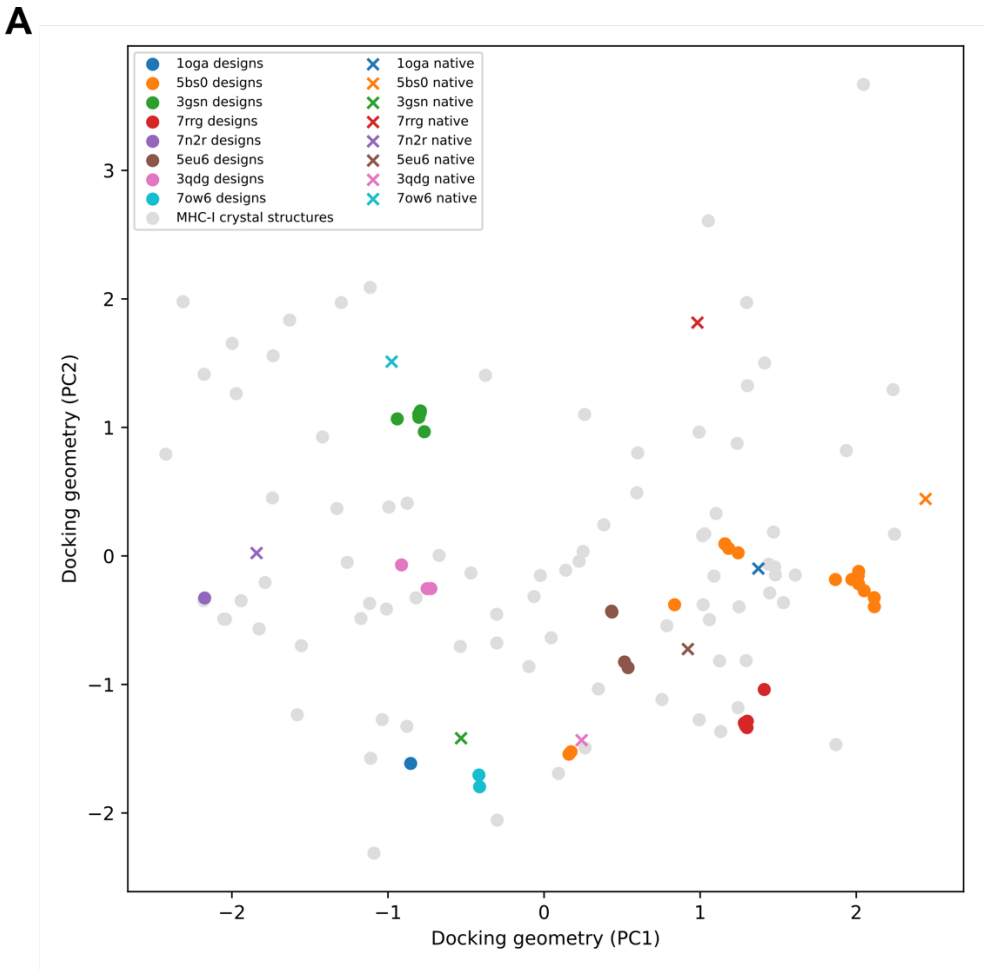
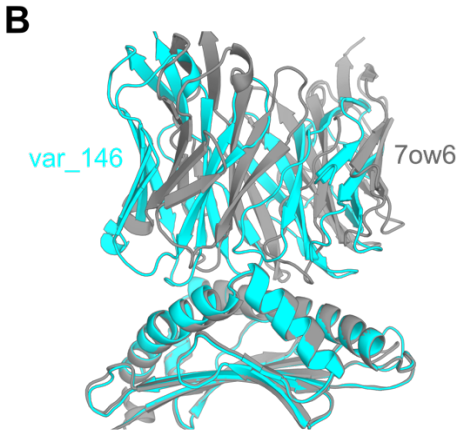

var_651 A01-EVD (on-target)
46.4%
var_651 A01-EVD (off-target)
2.5%
var_11 A01-EVD (on-target)
93.5%
var_11 A01-EVD (off-target)
4.5%
var_17 A01-EVD (on-target)
80.6%
var_17 A01-EVD (off-target)
2.5%
var_18 A01-EVD (on-target)
98.5%
var_18 A01-EVD (off-target)
9.6%
var_229 A01-EVD (on-target)
73.5%
var_229 A01-EVD (off-target)
2.8%
var_28 A01-EVD (on-target)
98.4%
var_28 A01-EVD (off-target)
37.0%
var_4 A01-EVD (on-target)
99.2%
var_4 A01-EVD (off-target)
98.9%
var_407 A01-EVD (on-target)
30.5%
var_407 A01-EVD (off-target)
1.1%
var_7 A01-EVD (on-target)
97.4%
var_7 A01-EVD (off-target)
11.4%
var_601 A02-ALY (on-target)
78.0%
var_601 A02-ALY (off-target)
6.5%
var_603 A02-ALY (on-target)
20.2%
var_603 A02-ALY (off-target)
2.0%
var_604 A02-ALY (on-target)
52.8%
var_604 A02-ALY (off-target)
14.6%
var_608 A02-ALY (on-target)
12.6%
var_608 A02-ALY (off-target)
8.2%
var_619 A02-ALY (on-target)
5.0%
var_619 A02-ALY (off-target)
3.7%
var_623 A02-ALY (on-target)
5.6%
var_623 A02-ALY (off-target)
6.6%
var_606 A02-ALY (on-target)
23.8%
var_606 A02-ALY (off-target)
2.4%
var_612 A02-ALY (on-target)
92.2%
var_612 A02-ALY (off-target)
3.4%
var_613 A02-ALY (on-target)
97.1%
var_613 A02-ALY (off-target)
5.7%
var_618 A02-ALY (on-target)
29.2%
var_618 A02-ALY (off-target)
3.9%
var_629 A02-ALY (on-target)
90.6%
var_629 A02-ALY (off-target)
4.2%
var_499 A02-GLM (on-target)
98.6%
var_499 A02-GLM (off-target)
4.0%
var_540 A02-LLW (on-target)
13.6%
var_540 A02-LLW (off-target)
5.0%
var_528 A02-LLW (on-target)
78.4%
var_528 A02-LLW (off-target)
23.5%
var_534 A02-LLW (on-target)
98.4%
var_534 A02-LLW (off-target)
26.7%
var_536 A02-LLW (on-target)
99.4%
var_536 A02-LLW (off-target)
14.0%
var_113 A02-TLM (on-target)
94.4%
var_113 A02-TLM (off-target)
3.1%
var_114 A02-TLM (on-target)
97.2%
var_114 A02-TLM (off-target)
3.9%
var_477 A02-TLM (on-target)
0.2%
var_477 A02-TLM (off-target)
0.4%
var_124 A02-TLM (on-target)
97.7%
var_124 A02-TLM (off-target)
2.9%
var_128 A02-TLM (on-target)
91.6%
var_128 A02-TLM (off-target)
0.9%
var_146 A02-TLM (on-target)
98.7%
var_146 A02-TLM (off-target)
1.8%
var_147 A02-TLM (on-target)
97.8%
var_147 A02-TLM (off-target)
2.1%
var_683 B44-SEI (on-target)
94.5%
var_683 B44-SEI (off-target)
14.7%
var_692 B44-SEI (on-target)
97.0%
var_692 B44-SEI (off-target)
15.7%
var_712 B44-SEI (on-target)
99.6%
var_712 B44-SEI (off-target)
15.8%
On-target pMHC binding
Off-target pMHC binding
TCR surface expression

Fig. S2. Representative pMHC binding data for 35 designed TCRs analyzed by flow cytometry. TCR surface expression as measured by CD3 staining is plotted on the x-axis; on- or off-target binding to multimerized pMHC is plotted on the y-axis. Cells gated as TCR-positive are shown in blue (if gated as pMHC-negative) or red (if gated as pMHC-positive). Labels in the upper left corner represent the percentage of TCR-positive cells that are pMHC positive.


A
1oga designs
5bs0 designs
3gsn designs
7rrg designs
7n2r designs
5eu6 designs
3qdg designs
7ow6 designs
MHC-I crystal structures
1oga native
5bs0 native
3gsn native
7rrg native
7n2r native
5eu6 native
3qdg native
7ow6 native
Docking geometry (PC2)
Docking geometry (PC1)
3
2
1
0
−1
−2
−2
−1
0
1
2


B
var_146
7ow6

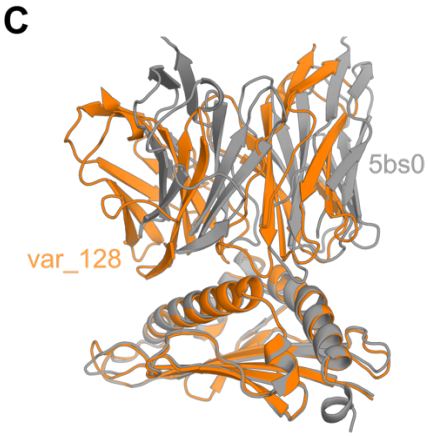

C
5bs0
var_128

Fig. S3. Comparison of docking geometries in design models and their framework-region template structures. (A) Scatter plot of the top two of 6 principal components of the TCR:pMHC rigid-body docking space for binding modes from a set of successful design models (disks colored by the native structural template for their framework region) and a set of native class I TCR:pMHC structures (gray disks and colored Xs for those template TCRs). Designed binding modes are not simply borrowed from the framework template, as expected since only the internal coordinates of the template TCR, and not the docking geometry, are supplied to AF2 during ADAPT design. **(B-C)** Example superpositions of design models (cyan and orange) and framework template structures (gray), aligned on the MHC chain.

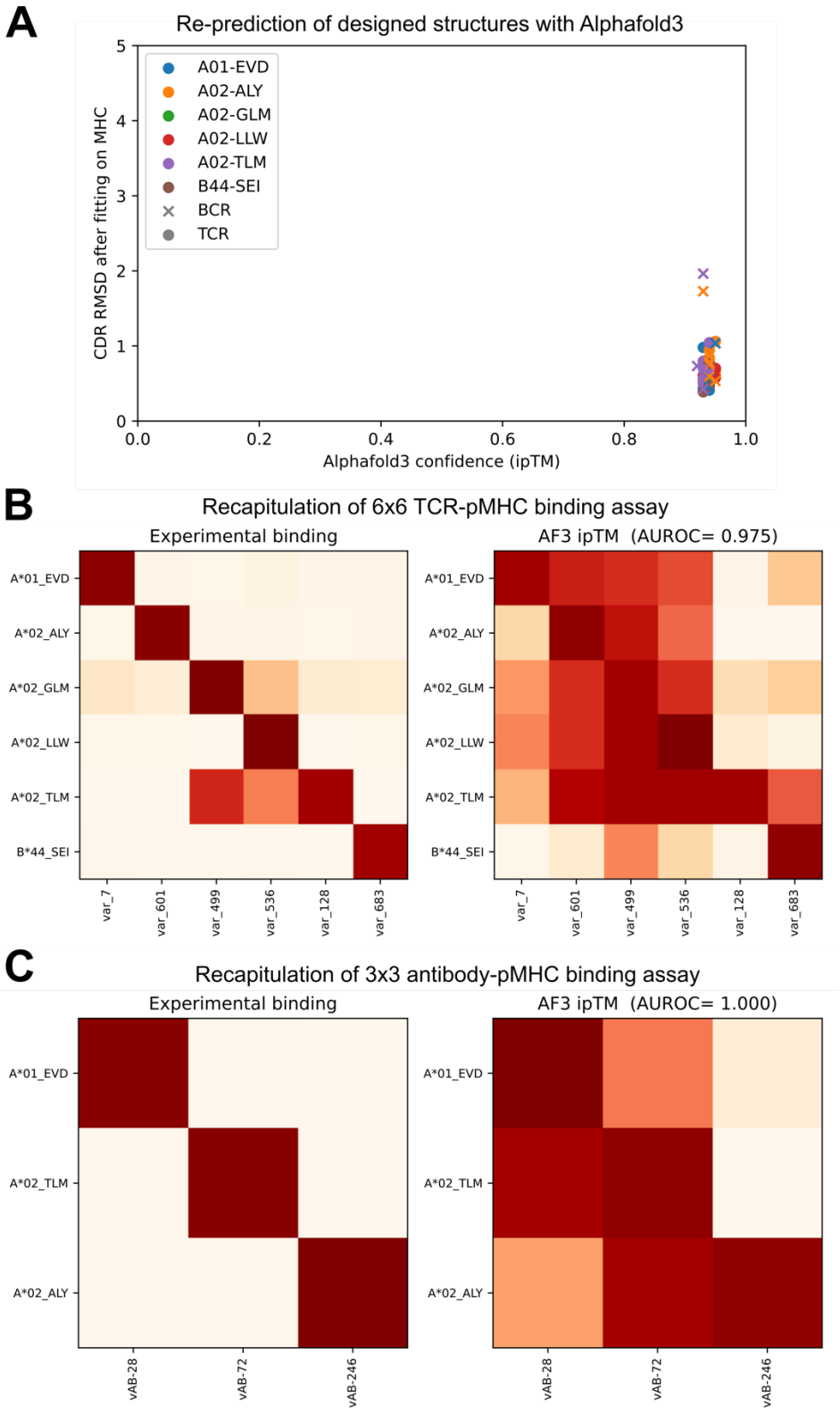

A
Re-prediction of designed structures with Alphafold3
A01-EVD
A02-ALY
A02-GLM
A02-LLW
A02-TLM
B44-SEI
BCR
TCR
CDR RMSD after fitting on MHC
Alphafold3 confidence (ipTM)
B
Recapitulation of 6x6 TCR-pMHC binding assay
Experimental binding
AF3 ipTM (AUROC= 0.975)
A*01_EVD
A*02_ALY
A*02_GLM
A*02_LLW
A*02_TLM
B*44_SEI
var_7
var_601
var_499
var_536
var_128
var_683
C
Recapitulation of 3x3 antibody-pMHC binding assay
Experimental binding
AF3 ipTM (AUROC= 1.000)
A*01_EVD
A*02_TLM
A*02_ALY
vAB-28
vAB-72
vAB-246

Fig. S4. Retrospective analysis of TCR and antibody designs with Alphafold3. (A) The default Alphafold3 pipeline was used to predict the designed interface structures using as input only the designed sequences (without interface templates or fine-tuning). Ca-RMSD values calculated over the CDR loops after aligning on MHC (y-axis) were below 1Å for nearly all the designs, and interface confidence was high as indicated by ipTM scores (x-axis) consistently above 0.9. **(B)** Heatmaps of experimental binding (left) and AF3 ipTM scores (right) for all 6x6=36 pairings of the 6 pMHC targets and a single representative TCR design for each. Ranking the possible pairings by ipTM and considering the designed interactions (diagonal entries) as true interactions gives an area under the receiver operating characteristic curve (AUROC) value of 0.975. **(C)** Experimental (left) and AF3-simulated (right) cross-binding heatmaps for three pMHC targets and a single representative antibody design for each. Ranking the possible pairings by AF3 ipTM score gives an AUROC of 1.0 for discriminating the designed pairings (i.e., the diagonal of the heatmap) from off-target interactions (off-diagonal entries in the heatmap).

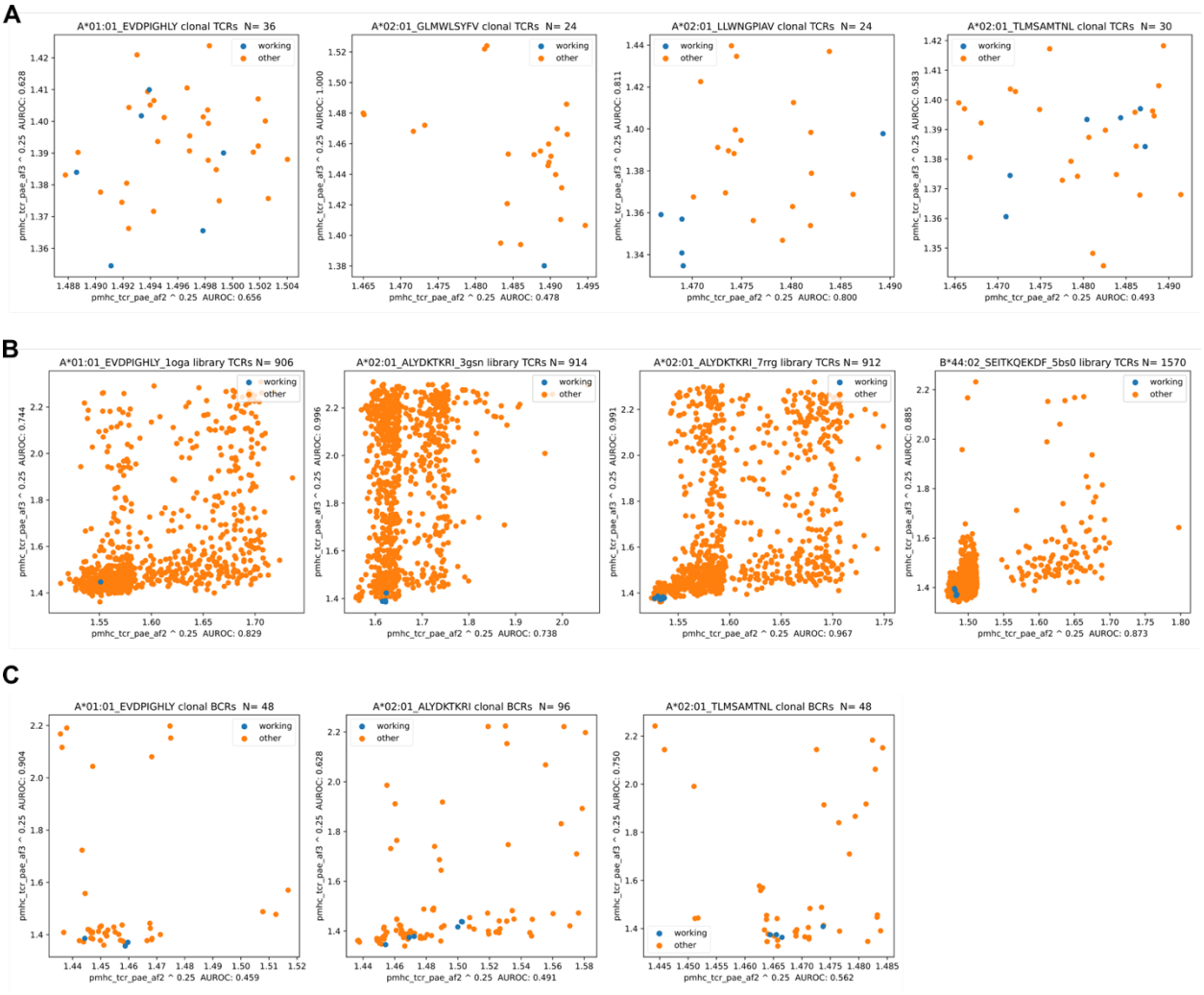


Fig. S5. Comparison of Alphafold interface metrics for working and non-working designs. Scatter plots of Alphafold2 (x-axis) versus Alphafold3 (y-axis) TCR-pMHC interface PAE scores for working and non-working designs; lower values indicate lower predicted error and hence higher confidence in the designed interface. Raw PAE values are raised to the 1/4th power to focus on the lowest scores and compress the higher scores corresponding to low-confidence designs. Area under the receiver operating characteristic curve (AUROC) values measuring discrimination of working from non-working designs are provided in the corresponding axis labels. Discrimination of working designs is modest for clonal TCR designs **(A)** but stronger for TCR libraries **(B)** and antibodies **(C)**.

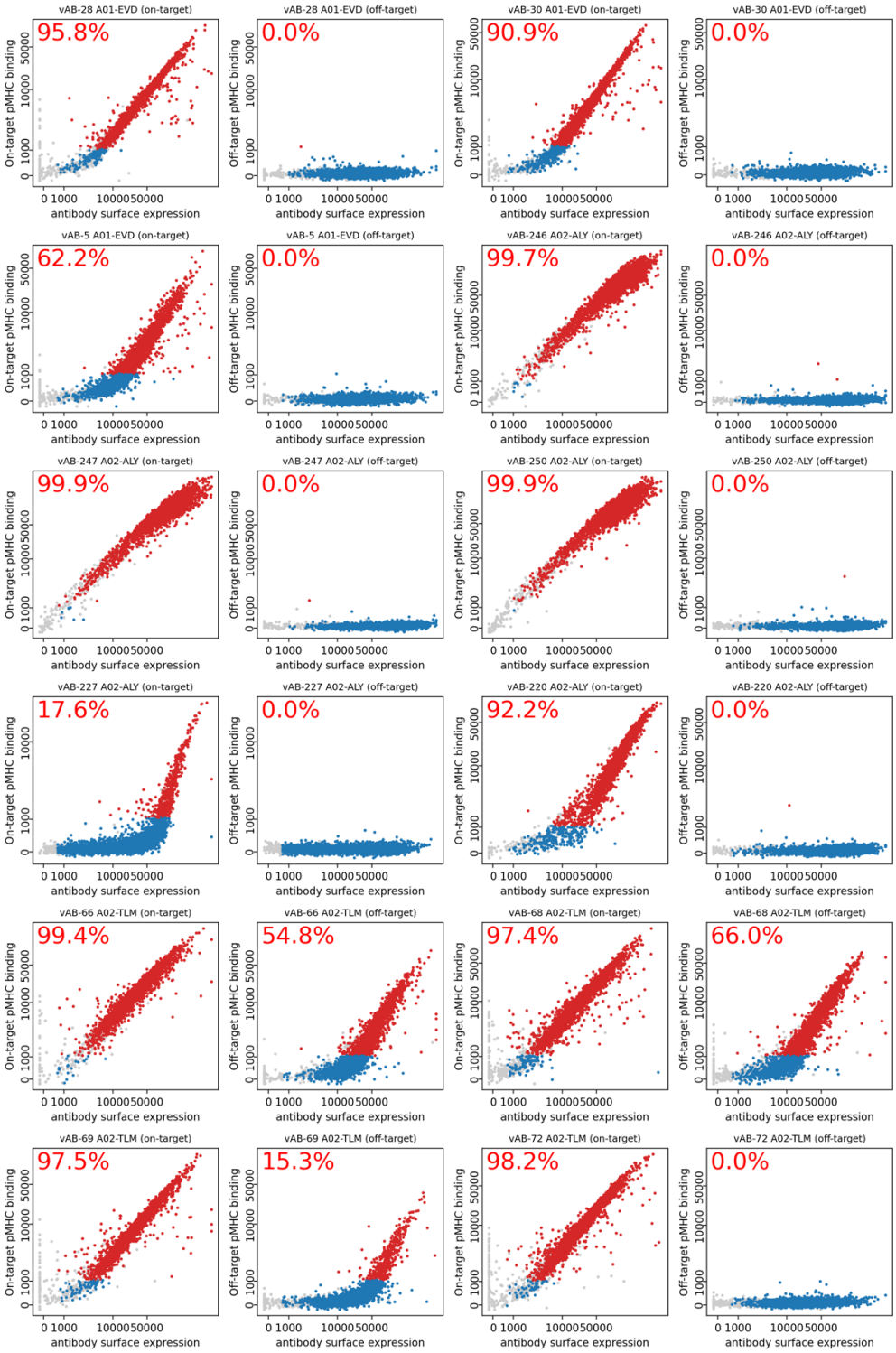

vAB-28 A01-EVD (on-target)
95.8%
vAB-28 A01-EVD (off-target)
0.0%
vAB-30 A01-EVD (on-target)
90.9%
vAB-30 A01-EVD (off-target)
0.0%
vAB-5 A01-EVD (on-target)
62.2%
vAB-5 A01-EVD (off-target)
0.0%
vAB-246 A02-ALY (on-target)
99.7%
vAB-246 A02-ALY (off-target)
0.0%
vAB-247 A02-ALY (on-target)
99.9%
vAB-247 A02-ALY (off-target)
0.0%
vAB-250 A02-ALY (on-target)
99.9%
vAB-250 A02-ALY (off-target)
0.0%
vAB-227 A02-ALY (on-target)
17.6%
vAB-227 A02-ALY (off-target)
0.0%
vAB-220 A02-ALY (on-target)
92.2%
vAB-220 A02-ALY (off-target)
0.0%
vAB-66 A02-TLM (on-target)
99.4%
vAB-66 A02-TLM (off-target)
54.8%
vAB-68 A02-TLM (on-target)
97.4%
vAB-68 A02-TLM (off-target)
66.0%
vAB-69 A02-TLM (on-target)
97.5%
vAB-69 A02-TLM (off-target)
15.3%
vAB-72 A02-TLM (on-target)
98.2%
vAB-72 A02-TLM (off-target)
0.0%
On-target pMHC binding
Off-target pMHC binding
antibody surface expression
0
1000
10000
50000

Fig. S6. Representative pMHC binding data for 12 designed antibodies analyzed by flow cytometry. IgG surface expression is plotted on the x-axis; on- or off-target binding to multimerized pMHC is plotted on the y-axis. Cells gated as IgG-positive are shown in blue (if gated as pMHC-negative) or red (if gated as pMHC-positive). Labels in the upper left corner represent the percentage of IgG-positive cells that are pMHC positive.

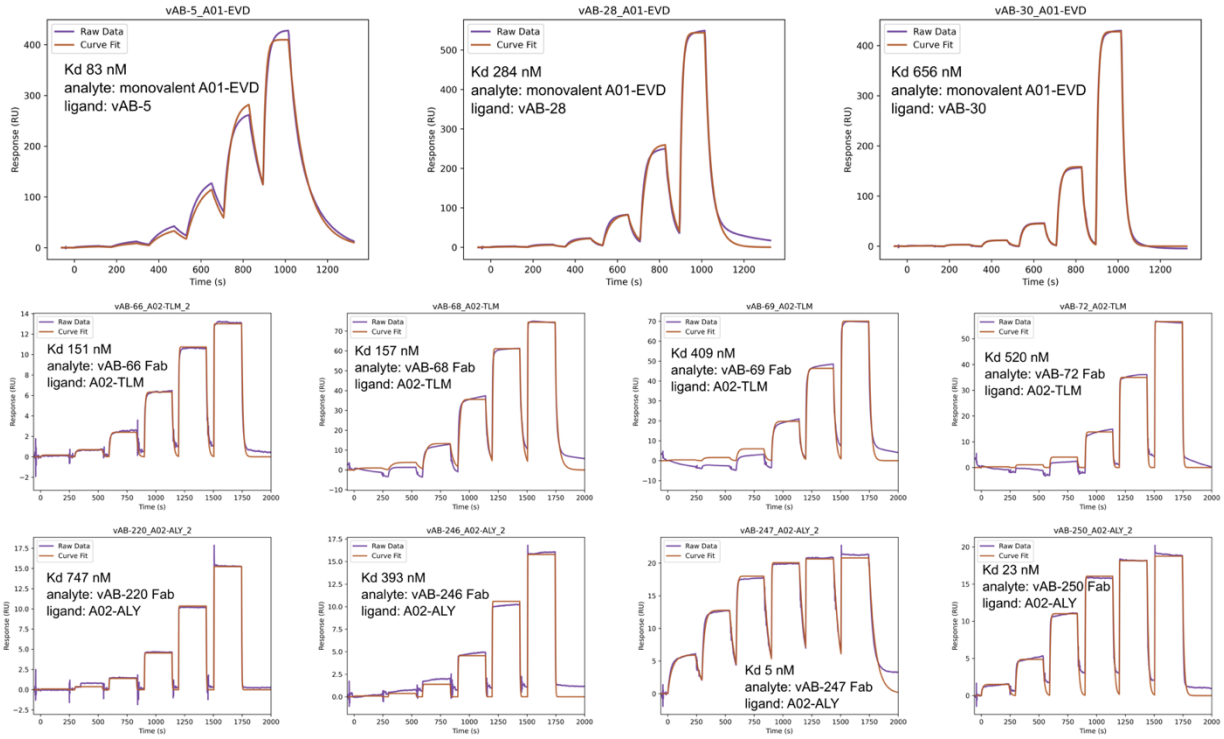


Fig. S7. Surface plasmon resonance (SPR) analysis of designed antibody–pMHC interactions. Binding responses were measured on a Biacore 8K at 25 °C in HBS-EP+ buffer. For the A01-EVD system (top row), IgG1s were captured on a Protein A chip and monovalent biotinylated A01-EVD analytes were injected. For all other targets (A02-TLM, A02-ALY), biotinylated pMHCs were immobilized on a streptavidin chip and IgG1-Fab analytes were tested using single-cycle kinetics. Raw sensorgrams (purple) and global 1:1 Langmuir fits (maroon) are shown with dissociation constants (Kd) indicated.

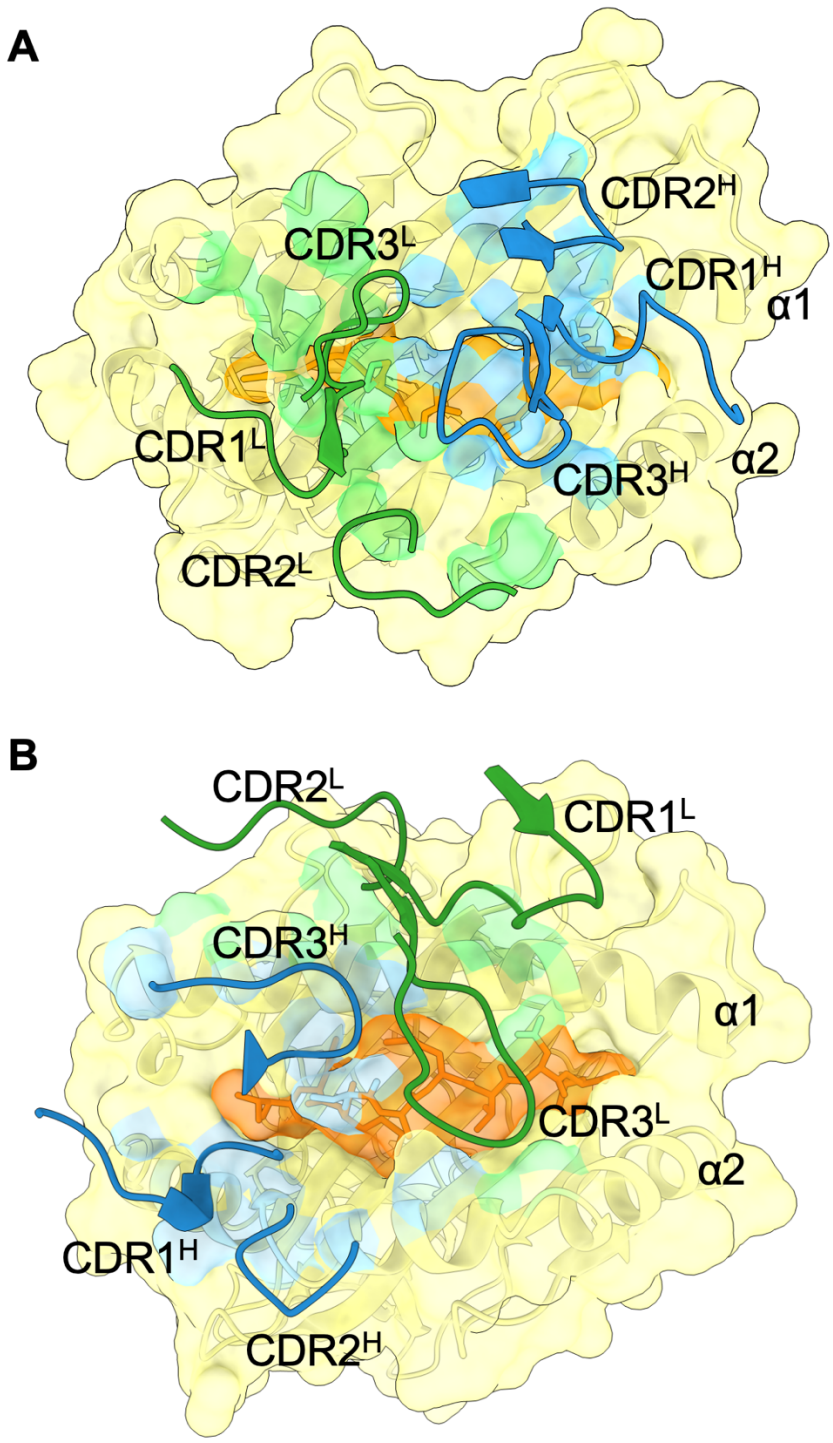


Fig. S8. CDR loop interactions with pMHC in cryoEM structures. MHC is shown in pale yellow, peptide in orange, Fab heavy chains in blue, and Fab light chains in green. Contacts to pMHC from the heavy chain are shown as light blue and from the light chain are shown as light green. **(A)** vAB-30/A01-EVD **(B)** vAB-66/A02-TLM.

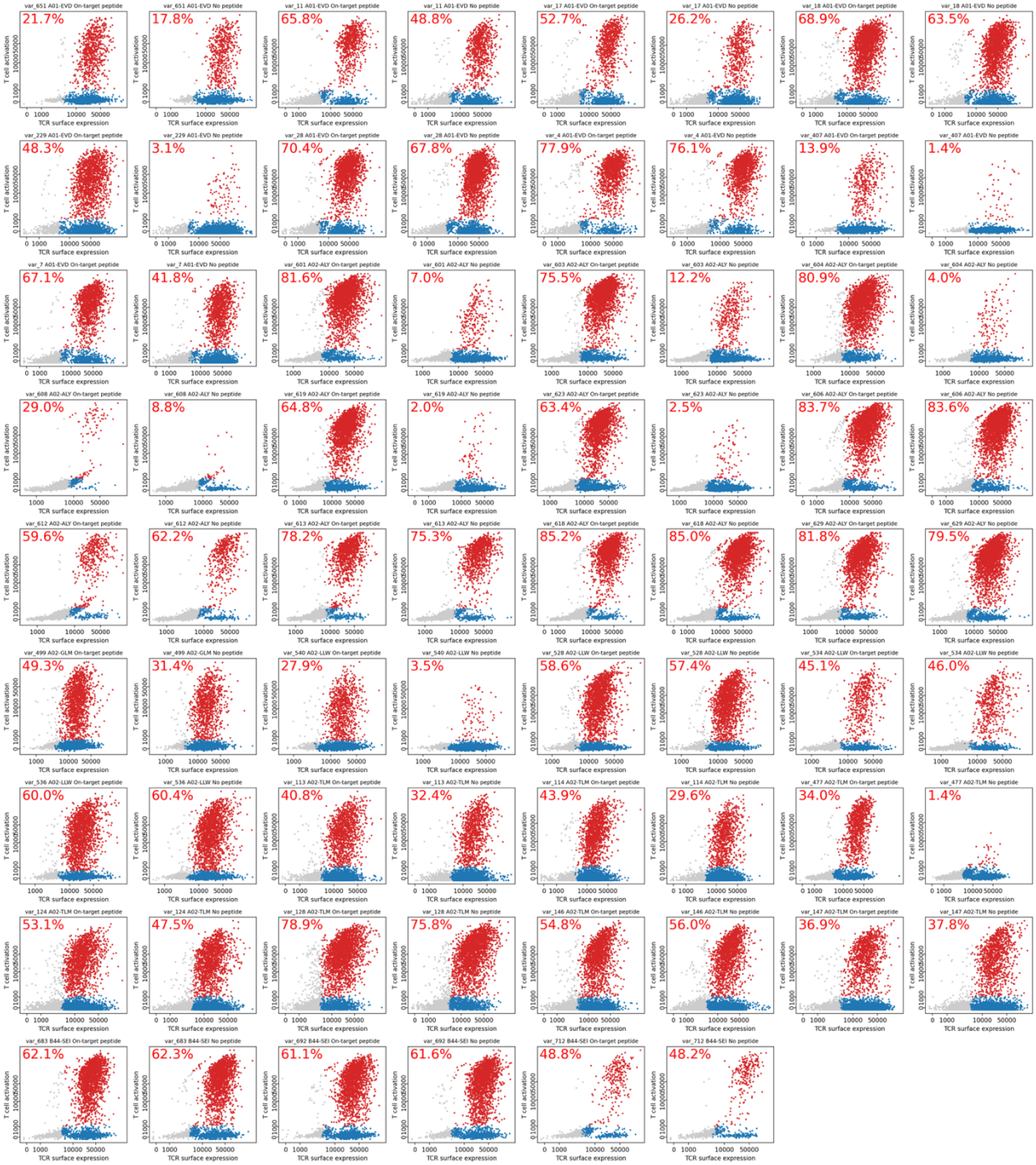

var_651 A01-EVD On-target peptide
21.7%
var_651 A01-EVD No peptide
17.8%
var_11 A01-EVD On-target peptide
65.8%
var_11 A01-EVD No peptide
48.8%
var_17 A01-EVD On-target peptide
52.7%
var_17 A01-EVD No peptide
26.2%
var_18 A01-EVD On-target peptide
68.9%
var_18 A01-EVD No peptide
63.5%
var_229 A01-EVD On-target peptide
48.3%
var_229 A01-EVD No peptide
3.1%
var_28 A01-EVD On-target peptide
70.4%
var_28 A01-EVD No peptide
67.8%
var_4 A01-EVD On-target peptide
77.9%
var_4 A01-EVD No peptide
76.1%
var_407 A01-EVD On-target peptide
13.9%
var_407 A01-EVD No peptide
1.4%
var_7 A01-EVD On-target peptide
67.1%
var_7 A01-EVD No peptide
41.8%
var_601 A02-ALY On-target peptide
81.6%
var_601 A02-ALY No peptide
7.0%
var_603 A02-ALY On-target peptide
75.5%
var_603 A02-ALY No peptide
12.2%
var_604 A02-ALY On-target peptide
80.9%
var_604 A02-ALY No peptide
4.0%
var_608 A02-ALY On-target peptide
29.0%
var_608 A02-ALY No peptide
8.8%
var_619 A02-ALY On-target peptide
64.8%
var_619 A02-ALY No peptide
2.0%
var_623 A02-ALY On-target peptide
63.4%
var_623 A02-ALY No peptide
2.5%
var_606 A02-ALY On-target peptide
83.7%
var_606 A02-ALY No peptide
83.6%
var_612 A02-ALY On-target peptide
59.6%
var_612 A02-ALY No peptide
62.2%
var_613 A02-ALY On-target peptide
78.2%
var_613 A02-ALY No peptide
75.3%
var_618 A02-ALY On-target peptide
85.2%
var_618 A02-ALY No peptide
85.0%
var_629 A02-ALY On-target peptide
81.8%
var_629 A02-ALY No peptide
79.5%
var_499 A02-GLM On-target peptide
49.3%
var_499 A02-GLM No peptide
31.4%
var_540 A02-LLW On-target peptide
27.9%
var_540 A02-LLW No peptide
3.5%
var_528 A02-LLW On-target peptide
58.6%
var_528 A02-LLW No peptide
57.4%
var_534 A02-LLW On-target peptide
45.1%
var_534 A02-LLW No peptide
46.0%
var_536 A02-LLW On-target peptide
60.0%
var_536 A02-LLW No peptide
60.4%
var_113 A02-TLM On-target peptide
40.8%
var_113 A02-TLM No peptide
32.4%
var_114 A02-TLM On-target peptide
43.9%
var_114 A02-TLM No peptide
29.6%
var_477 A02-TLM On-target peptide
34.0%
var_477 A02-TLM No peptide
1.4%
var_124 A02-TLM On-target peptide
53.1%
var_124 A02-TLM No peptide
47.5%
var_128 A02-TLM On-target peptide
78.9%
var_128 A02-TLM No peptide
75.8%
var_146 A02-TLM On-target peptide
54.8%
var_146 A02-TLM No peptide
56.0%
var_147 A02-TLM On-target peptide
36.9%
var_147 A02-TLM No peptide
37.8%
var_683 B44-SEI On-target peptide
62.1%
var_683 B44-SEI No peptide
62.3%
var_692 B44-SEI On-target peptide
61.1%
var_692 B44-SEI No peptide
61.6%
var_712 B44-SEI On-target peptide
48.8%
var_712 B44-SEI No peptide
48.2%
T cell activation
TCR surface expression

Fig. S9. Representative T cell activation data for 35 designed TCRs analyzed by flow cytometry. TCR surface expression as measured by CD3 staining is plotted on the x-axis; T cell activation (NFAT RE eGFP expression) is plotted on the y-axis. Cells gated as TCR-positive are shown in red if gated as active or blue otherwise. Labels in the upper left corner represent the percentage of TCR-positive cells that are activated. The two panels for each design represent activation in the presence or absence of pulsed target peptide.

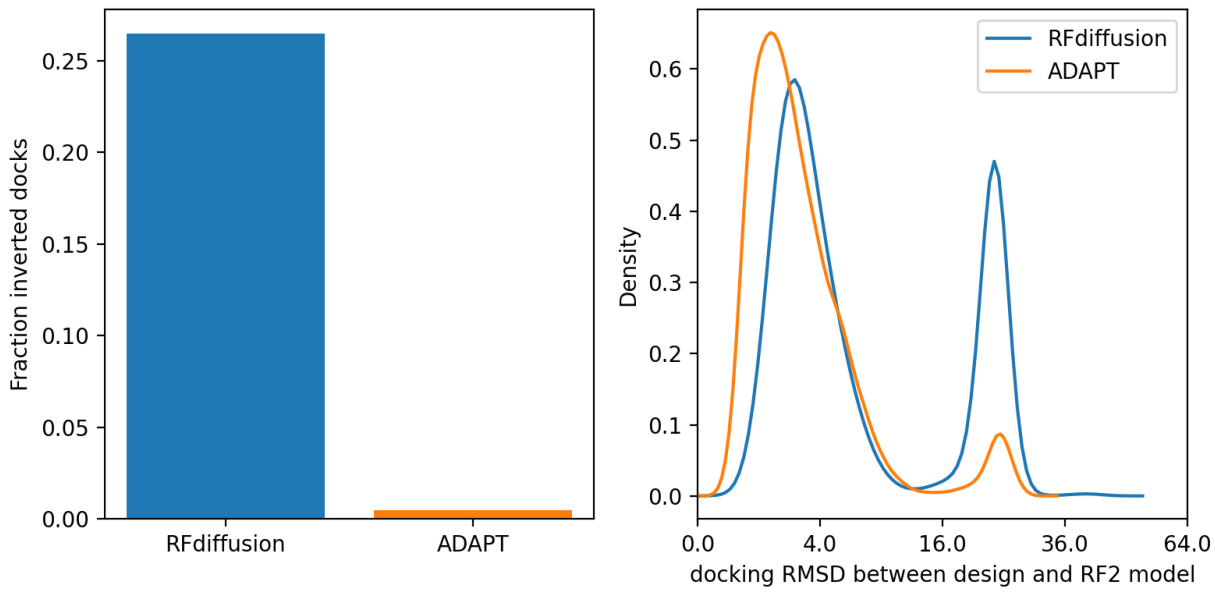
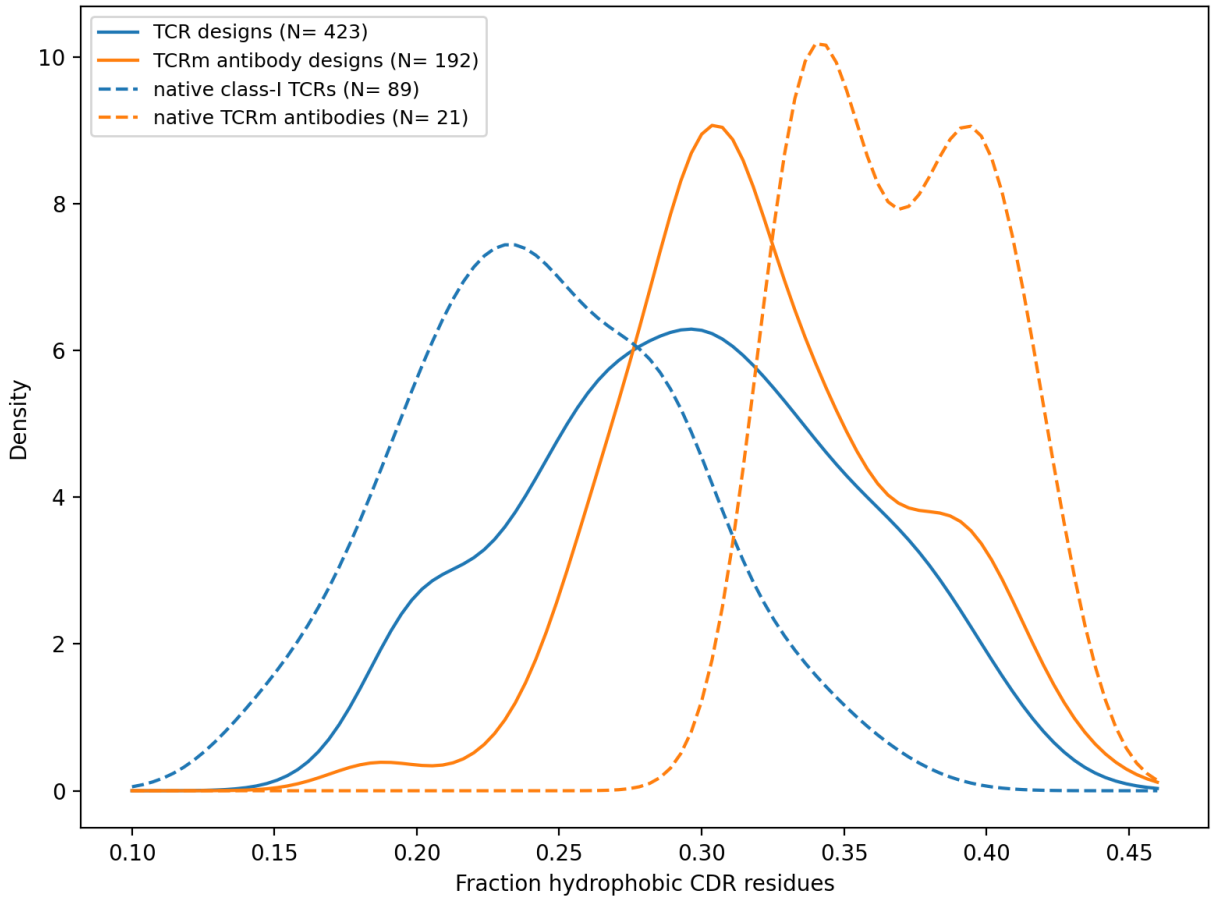


Fig. S10. Comparison between RFdiffusion- and ADAPT-generated TCR:pMHC binding modes. A version of RFdiffusion that was fine-tuned on native TCR:pMHC complexes generated reversed binding modes at higher frequencies than the ADAPT pipeline (left panel). Reprediction of designed binding modes by RosettaFold2 was less successful for RFdiffusion designs than for ADAPT-generated designs as measured by higher docking RMSD values between the design model and the RF2 prediction (right panel).



Supplementary Figure 11. The CDR loops of designed TCRs and TCR-mimic antibodies are more hydrophobic than those of native TCRs. Smoothed kernel density estimates of CDR loop hydrophobicity as measured by the fraction of CDR loop positions occupied by the amino acids {V,I,L,M,F,W,Y,C}. Interestingly, native TCR-mimic antibodies show even greater CDR loop hydrophobicity (orange dashed line).

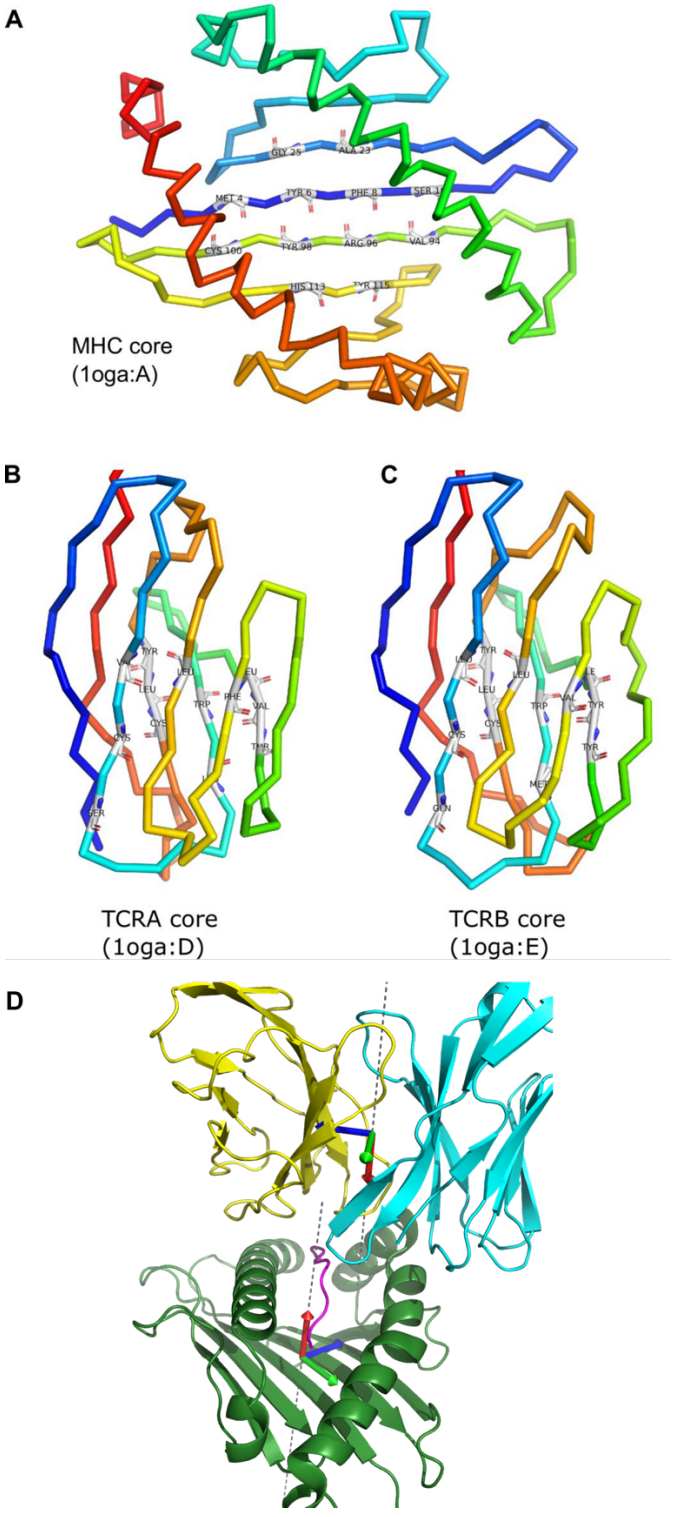

A
MHC core
(1oga:A)
B
TCRA core
(1oga:D)
C
TCRB core
(1oga:E)
D

Fig. S12. Construction of reference coordinate frames for MHC and TCR structures. As described in the Methods, the reference frames are defined by the approximate 2-fold symmetry axis relating a set of corresponding residue pairs in the MHC beta sheet **(A)** and in the TCR alpha and beta chain framework regions **(B-C)**. **(D)** The frame origin is located at the center of mass of the aligned residues. The x-axis (red arrow) is parallel to the symmetry axis, and the z-axis (blue arrow) is directed between the centers of mass of the individual residue subsets (ie, from the N-terminal to C-terminal MHC beta-sheet halves and from the TCR alpha to beta chains).

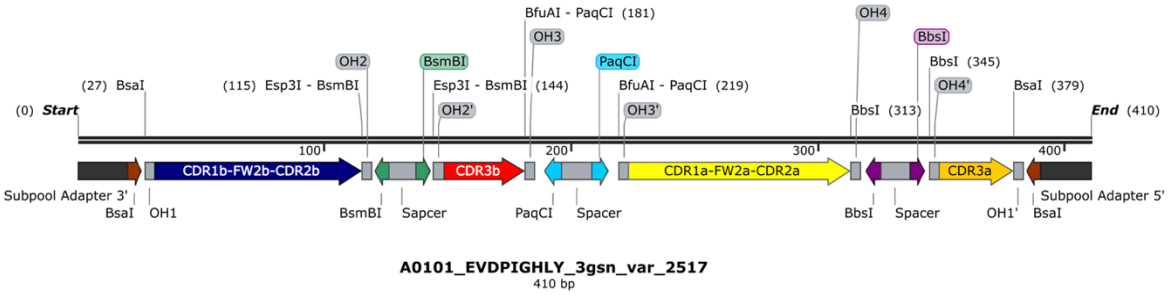


Fig. S13. Schematic layout of pooled fragment design. CDR1b–FW2–CDR2b, CDR3b, CDR1a–FW2a–CDR2a, and CDR3a sequences corresponding to the designed TCRs with a specified framework were interspersed with golden gate assembly sites to enable sequential cloning for integrating the invariant intermediating DNA fragments as described in **Methods**.

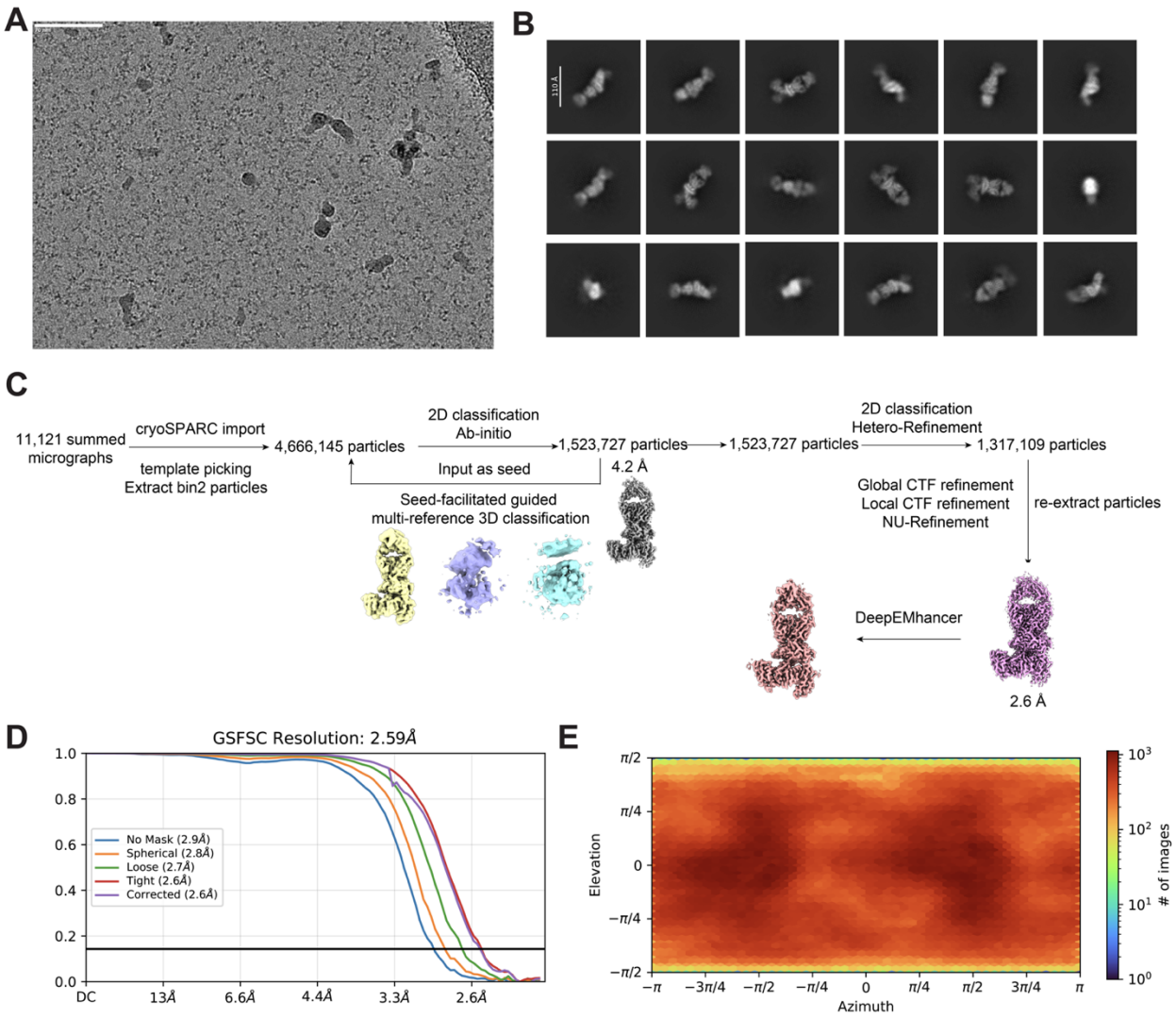


Fig. S14. Workflow for cryo-EM data processing for vAB-30/A01-EVD complex. (A) A representative micrograph of the vAB-30/A01-EVD complex. Scale bar = 70 nm. **(B)** Representative 2D class averages illustrating the varied orientations of the vAB-30/A01-EVD. **(C)** Flow chart for data processing of the vAB-30/A01-EVD complex. **(D)** Gold standard FSC curve illustrating the resolution determination for the vAB-30/A01-EVD complex, achieved through iterative 3D refinement. The resolution at the FSC 0.143 criterion is 2.6 Å. **(E)** Angular distribution of the particles used for the final reconstructions.

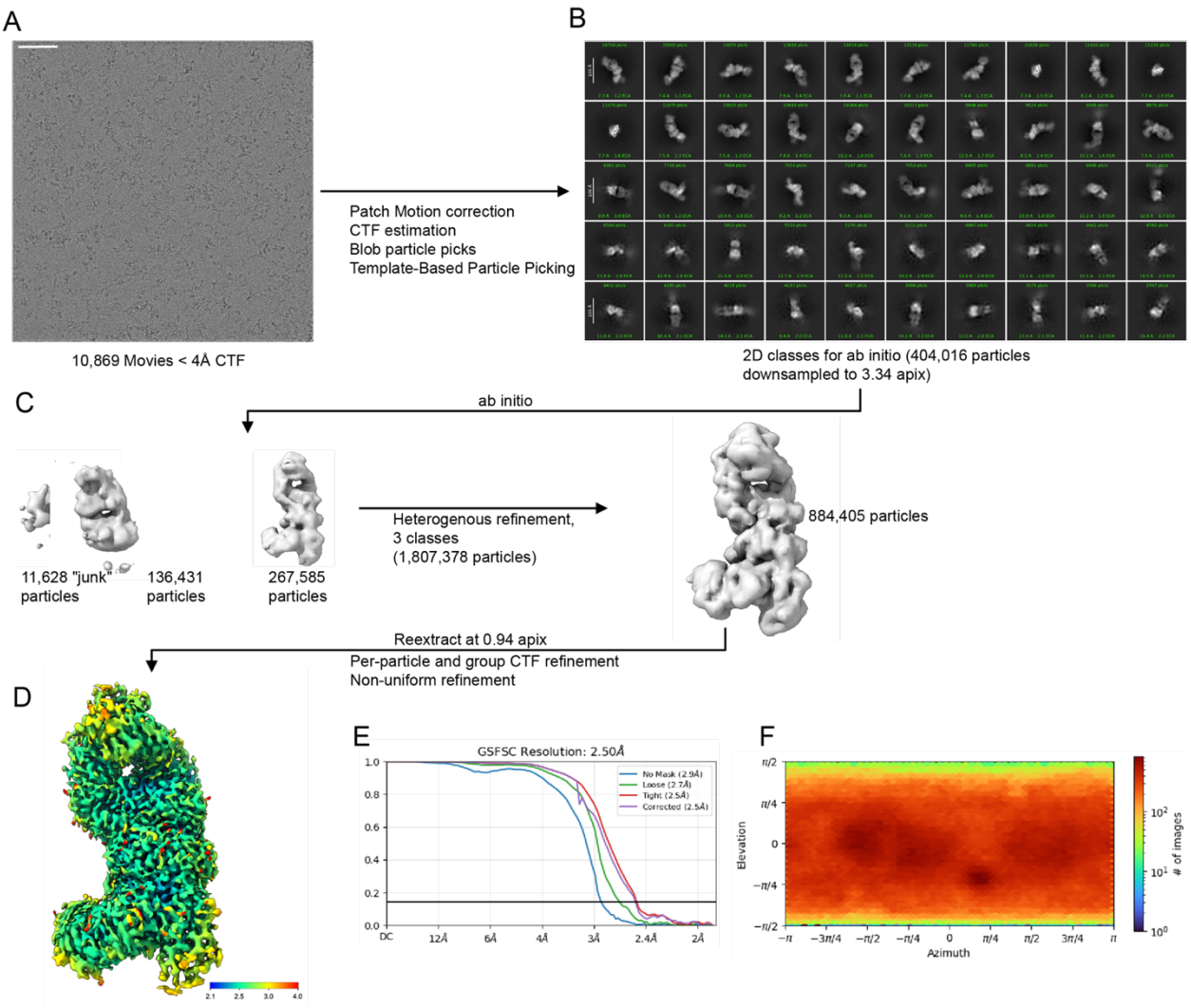


Fig. S15. Workflow for cryo-EM data processing for vAB-66/A02-TLM complex. (A) A representative micrograph of the vAB-66/A02-TLM complex. Scale bar = 50 nm. **(B)** Representative 2D class averages illustrating the varied orientations of the vAB-66/A02-TLM. **(C)** Flow chart for data processing of the vAB-66/A02-TLM complex. **(D)** Final map colored by local resolution. **(E)** Gold standard FSC curve illustrating the resolution determination for the vAB-66/A02-TLM complex, achieved through iterative 3D refinement. The resolution at the FSC 0.143 criterion is 2.5 Å. **(F)** Angular distribution of the particles used for the final reconstructions.

Supplementary Tables

Table S1. Off-target peptides for flow cytometry binding experiments.

MHC	Peptide	Off-target MHC	Off-target peptide
A*01:01	EVDPIGHLY	A*01:01	ESDPIVAQY
A*02:01	TLMSAMTNL	A*02:01	NLVPMVATV
A*02:01	ALYDKTKRI	A*02:01	SLLMWITQC
A*02:01	GLMWLSYFV	A*02:01	SLLMWITQC
A*02:01	LLWNGPIAV	A*02:01	SLLMWITQC
B*44:02	SEITKQEKDF	A*01:01	EVDPIGHLY

Table S2. Designed sequence changes in the CDR loops.

Loop	Number of mutations per loop mean (min - max)	Percent mutated positions per loop mean (min - max)
CDR1A	5.0 (2 - 7)	79.8% (33.3% - 100.0%)
CDR2A	5.1 (3 - 8)	76.5% (42.9% - 100.0%)
CDR3A[1]	8.8 (6 - 10)[2]	88.7% (80.0% - 100.0%)
CDR1B	3.5 (2 - 5)	70.0% (40.0% - 100.0%)
CDR2B	4.4 (3 - 6)	73.2% (50.0% - 100.0%)
CDR3B[1]	7.4 (4 - 10)[2]	82.8% (50.0% - 100.0%)

[1]CDR3 statistics exclude the first three ('CXX') and last two (typically 'XF') positions, which are not varied in the design process.

[2]CDR3 mutation counts include indels since the length of the designed loop does not necessarily match the length of the CDR3 in the framework template.

Table S3. Antibody binding affinities from surface plasmon resonance.

Complex	Analyte	Ligand	Kd (nM)
vAB-5: A01-EVD	monovalent pMHC	IgG	83
vAB-28: A01-EVD	monovalent pMHC	IgG	284
vAB-30: A01-EVD	monovalent pMHC	IgG	656
vAB-66: A02-TLM	monovalent IgG Fab	pMHC	151
vAB-68: A02-TLM	monovalent IgG Fab	pMHC	157
vAB-69: A02-TLM	monovalent IgG Fab	pMHC	409
vAB-72: A02-TLM	monovalent IgG Fab	pMHC	520
vAB-220: A02-ALY	monovalent IgG Fab	pMHC	747
vAB-246: A02-ALY	monovalent IgG Fab	pMHC	393
vAB-247: A02-ALY	monovalent IgG Fab	pMHC	5
vAB-250: A02-ALY	monovalent IgG Fab	pMHC	23

Table S4. Summary of cryoEM data collection and model statistics.

	vAB-30/A01-EVD	**vAB-66/A02-TLM**
EM equipment	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300
Detector	K3	Falcon 4i
Pixel size (Å)	0.827	0.94
Electron dose ($e^-/Å^2$)	60	50
Defocus range (μm)	(-1.0, -2.0)	(-0.8, -2.0)
Reconstruction		
Software	CryoSPARC	CryoSPARC
Number of used Particles	1,317,109	880,830
Map sharpening Method	deepEMhancer (visualization)	Uniform
Final Resolution (Å)	2.6	2.5
Model building and refinement		
Software	PHENIX & COOT	PHENIX & COOT
Initial models used (PDB codes)	Predicted structure	3SOB, 7YV1, 7SR0, 9NMV
Model composition		
Non-hydrogen atoms	7270	7368
Protein residues	932	933
Water molecules	0	64
B factors ($Å^2$)		
Protein	47.37	39.28
Water	——	35.36

R.m.s deviations		
Bonds length (Å)	0.004	0.003
Bonds Angle (°)	0.503	0.548
Ramachandran plot statistics (%)		
Preferred	98.04	98.70
Allowed	1.96	1.30
Outlier	0.00	0.00
Validation		
Molprobity score	1.24	1.17
Clash score	4.62	2.51
Rotamer outliers (%)	0.00	1.49

Discussion

In this work, we demonstrate that combining structural reasoning with wetlab labeled data during model training can drastically assist in protein prediction and design tasks. Specifically, we show that fine-tuning structure prediction networks on peptide–MHC binding information produces a hybrid model that retains the geometric priors of deep structural networks while incorporating new biochemical constraints from experimental data. This enables highly accurate peptide–MHC binder discrimination, robust generalization across alleles and peptide lengths, and even cross-domain transfer to systems such as PDZ and SH3 domains, illustrating how structural models enriched with labeled data can overcome limitations faced by sequence-only approaches.

And we show that these same principles extend beyond prediction into design. Using structure-guided conformational sampling and sequence diversification across templates and CDR loops, the ADAPT design pipeline produces de novo TCRs and TCR-mimic antibodies with native-like binding geometries to targeted peptide–MHC complexes. These designs demonstrate atomic-level agreement in structural validation and functional activity in T-cell assays, revealing that structure-based generative pipelines can navigate the stringent geometric and biochemical constraints of the pMHC interface and highlighting both the power and the remaining challenges of creating highly specific, potent immune receptors.

Despite these advances, several limitations remain. Fine-tuning structural models on peptide–MHC binding data improves discrimination, but the approach still depends on the availability, quality, and coverage of wet-lab measurements, which remain uneven across MHC alleles and peptide lengths. The model also inherits biases from both the structure predictor and the experimental datasets, which may affect generalization to rare alleles, low-abundance peptide registers, or unconventional binding modes. On the design side, while the ADAPT pipeline can generate receptors with accurate geometries and functional activity, achieving strict specificity without off-target activation remains challenging, and the current framework does not yet incorporate affinity tuning, developability constraints, or comprehensive off-target profiling. Additionally, experimental validation—particularly high-resolution structural determination—remains rate-limiting compared to the speed of computational generation. Addressing these limitations will require integrating more

diverse training data, expanding cross-reactivity modeling, and incorporating downstream engineering constraints directly into the generative process.

We hope that these advances enable next-generation targeted immune therapies and establish a generalizable framework for programmable molecular recognition across diverse therapeutic and synthetic-biology applications.