

PROTEIN DESIGN

De novo-designed pMHC binders facilitate T cell-mediated cytotoxicity toward cancer cells

Kristoffer Haurum Johansen^{1*†}, Darian Stephan Wolff^{2†}, Beatrice Scapolo^{1,2†}, Monica L. Fernández-Quintero^{3†}, Charlotte Risager Christensen^{1,2}, Johannes R. Loeffler³, Esperanza Rivera-de-Torre², Max D. Overath², Kamilla Kjærgaard Munk¹, Oliver Morell², Marie Christine Viuff¹, Iñigo Lacunza¹, Alberte T. Damm Englund², Mathilde Due², Anant Gharpure³, Stefano Forli³, Carlos Rodriguez Pardo¹, Tripti Tamhane¹, Emma Qingjie Andersen², Kasper Haldrup Björnsson², Jordan Sylvester Fernandes², Lasse Frank Voss¹, Suthimon Thumtecho², Andrew B. Ward³, Maria Ormhøj¹, Sine Reker Hadrup^{1*†}, Timothy P. Jenkins^{2*†}

The recognition of intracellular antigens by CD8⁺ T cells through T cell receptors (TCRs) is central for adaptive immunity against infections and cancer. However, the identification of TCRs from patient material remains complex. We present a rapid de novo minibinder (miBd) design platform leveraging state-of-the-art generative models to engineer miBds targeting the cancer-associated peptide-bound major histocompatibility complex (pMHC) SLLMWITQC/HLA-A*02:01 (NY-ESO-1). Incorporating in silico cross-panning enabled computational prescreening of specificity, and molecular dynamics simulations allowed for improved predictability of in vitro success. We identified a high-affinity NY-ESO-1 binder and confirmed its structure using cryo-electron microscopy, which, when incorporated in a chimeric antigen receptor, induced killing of NY-ESO-1⁺ melanoma cells. We further designed and validated binders to a neoantigen pMHC complex, RVTDESILSY/HLA-A*01:01, with unknown structure, demonstrating the potential for precision immunotherapy.

T cells are immune cells capable of responding to intracellular antigens such as those derived from infections or cancerous mutations. CD8⁺ T cells recognize antigens through their T cell receptor (TCR), which binds peptides presented by major histocompatibility complex class I (pMHC-I) on all nucleated cells. MHC-I molecules, encoded by human leukocyte antigen (HLA) genes, adopt a globular structure composed of three extracellular domains, $\alpha 1$, $\alpha 2$, and $\alpha 3$, and a $\beta 2$ -microglobulin ($\beta 2m$) subunit. The $\alpha 1$ and $\alpha 2$ domains form a peptide-binding groove accommodating short peptides (typically nine to 11 amino acids) anchored at distinct positions, mostly second and C-terminal residues, and displays them for recognition by TCRs (1).

Cancer cells display antigens on pMHC-I that can be recognized by tumor-reactive T cells. TCRs isolated from such T cells have been tested in the clinical setting, with several clinical trials revealing durable responses in patients across several cancer indications (1). One such tumor-associated antigen, NY-ESO-1, contains a T cell epitope SLLMWITQC [NY-ESO-1⁽¹⁵⁷⁻¹⁶⁵⁾] presented in the context of HLA-A*02:01 (2). NY-ESO-1, also known as cancer/testis antigen 1B (CTAG1B), is expressed in a wide range of tumors, including melanoma, lung, ovarian, and breast cancers,

but not in normal tissues (except for the testis). It is structurally well characterized, and its broad, cancer-specific expression pattern makes it an attractive cancer immunotherapy target.

The identification of TCRs for therapeutic use is driven by the selection of relevant T cell populations from patients (3, 4) or through expansion of T cell populations from the naïve repertoire toward a given antigen (5). These approaches are laborious, technically challenging, and potentially limited or biased by the TCR repertoire in the donors used for selection. TCRs are inherently cross-reactive to confer wider protection against diseases, so most of them cross-react to several peptides (6). However, predicting the behavior and potential cross-reactivity of such TCRs to pMHCs present on healthy tissue, when applied in hosts other than the donors they are selected from, remains a challenge (7, 8).

To circumvent the challenges of TCR identification and to facilitate the clinical application of immunotherapies based on pMHC-I recognition, strong interest has emerged in developing alternative strategies to target infected or malignant cells presenting intracellular proteins on the pMHC-I complex. TCR-like molecules, which include TCR-like antibodies, TCR-bispecific engagers, and single-chain TCRs (1), are promising approaches. Such TCR-like molecules can be incorporated into chimeric antigen receptors (CARs) to induce T cell-mediated cytotoxicity toward a cell expressing a given pMHC-I target (9). TCR-like antibodies are currently found in either immune or synthetic libraries screened by display or other technologies. This process can take months to years and often yields very low success rates. Although some success has been reported, these approaches often struggle to generate antibodies that bind pMHC-I with high affinity. Hybridoma technology is limited by low throughput and labor-intensive protocols, and it skews toward the targeted HLA isoform rather than the peptide specificity (10). Phage display panning is faster and more cost-effective than hybridoma technology, but the libraries still rely on large, naïve antibody repertoires (11). Alternative methods have been explored in silico and in vitro (12, 13), but these are limited by a lack of knowledge about the specificity of individual TCRs and the pMHC-I complex to which they bind (8). Furthermore, establishing the specificity of high-affinity binders by exploration of cross-reactive binding presents a substantial challenge for current in vitro approaches (14).

Structural models such as TCRen (15) and DeepAIR (16) have predicted TCR-pMHC interaction with improved accuracy. However, advances in generative protein design provide additional opportunities for the in silico development of alternative TCR-like pMHC-I binders. Denoising diffusion models, in particular, have optimized machine learning-guided protein design because of their ability to generate protein-binder backbones that are optimized toward their target through stepwise diffusion. This has led to substantial improvements in experimental success rates.

Among various de novo protein design approaches, RFdiffusion is considered one of the most promising tools due to its impressive average success rate of 10% of designed binders presenting ultrahigh target affinity (<10 nM) (17). The design pipeline consists of three independently trained models: (i) the diffusion model, which de novo generates a highly accurate scaffold, followed by (ii) ProteinMPNN (18), and finally (iii) AlphaFold2 (AF2), which can predict and evaluate a protein structure by its sequence (19). For the evaluation and down-selection of minibinder (miBd) designs, the AF2 interface predicted aligned error (iPAE) with a value <10 Å and a predicted local distance difference test (pLDDT) >90, as suggested by Bennett *et al.* (20), is widely applied (17, 21, 22). Here, we describe the power of these tools for rapidly developing de novo miBds to pMHCs that can facilitate T cell-mediated cytotoxicity toward cancer cells.

Results

Design and in silico cross-panning of NY-ESO-1 miBds

We sought to design miBds against SLLMWITQC/HLA-A*02:01 using a published crystal structure (Protein Data Bank ID 2BNR) (23) as input to RFdiffusion. To guide the binder interaction toward the

¹Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark.

²Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. ³Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA. *Corresponding author. Email: tpj@dtu.dk (T.P.J.); krisha@dtu.dk (K.H.J.); sirha@dtu.dk (S.R.H.) †These authors contributed equally to this work.

surface-accessible residues of the peptide and not just the MHC-I backbone, we specified the residues L3, M4, W5, and T7 with upward-facing side chains as contact points for the RFdiffusion de novo design. Initially, we generated amino acid backbones for the miBds through RFdiffusion-denoising trajectories and interpreted sequences using ProteinMPNN (Fig. 1A). These sequence designs were evaluated by AF2 and filtered based on an iPAE <12 Å. To diversify the panel of designs, only miBds originating from different backbones were chosen. Next, the preselected miBds were sequence diversified by resubmission to ProteinMPNN (100 sequences per candidate), resulting in a final set of 44 miBds surpassing an iPAE of 7 Å and a pLDDT of 92. We found that most of the miBds consisted of tightly packed bundles of four α helices orientated almost exclusively parallel to the peptide (fig. S1, A to G).

To drive in silico design further toward specific recognition of the peptide, we leveraged AF2 to generate binding scores (iPAE and pLDDT) for all 44 miBds against a panel of single and double mutants of the wild-type peptide SLLMWITQC (Fig. 1B). The peptide variants presented on HLA-A*02:01 were modeled using AF3 and chosen to evaluate whether a minimal change of one or two key positions with exposed residues (mainly M4, W5, or T7) could disrupt the binding interaction between the miBds and the pMHC (24). Thus, we included peptide substitutions with a subtle change of the local physical-chemical properties, e.g., M4A and W5Y, as well as substantial changes of the local physical properties, by introducing negatively charged residues in position 4 (e.g., M4E and M4D). Furthermore, a variant peptide, identified in a human proteome search and predicted to have a strong affinity to the HLA-A*02:01, was added to the panel (fig. S2A) (25, 26). Finally, applying the iPAE cutoff (<10 Å) to define a potential binding interaction between miBd and a variant pMHC, we ranked the miBds by the number of such undesired binding events and proceeded with the 44 top candidates for in vitro assessment.

Molecular dynamics identification of false positives and off-target binding likelihood

Protein design guided by machine learning can result in a high rate of false positives when predicted binders fail during experimental

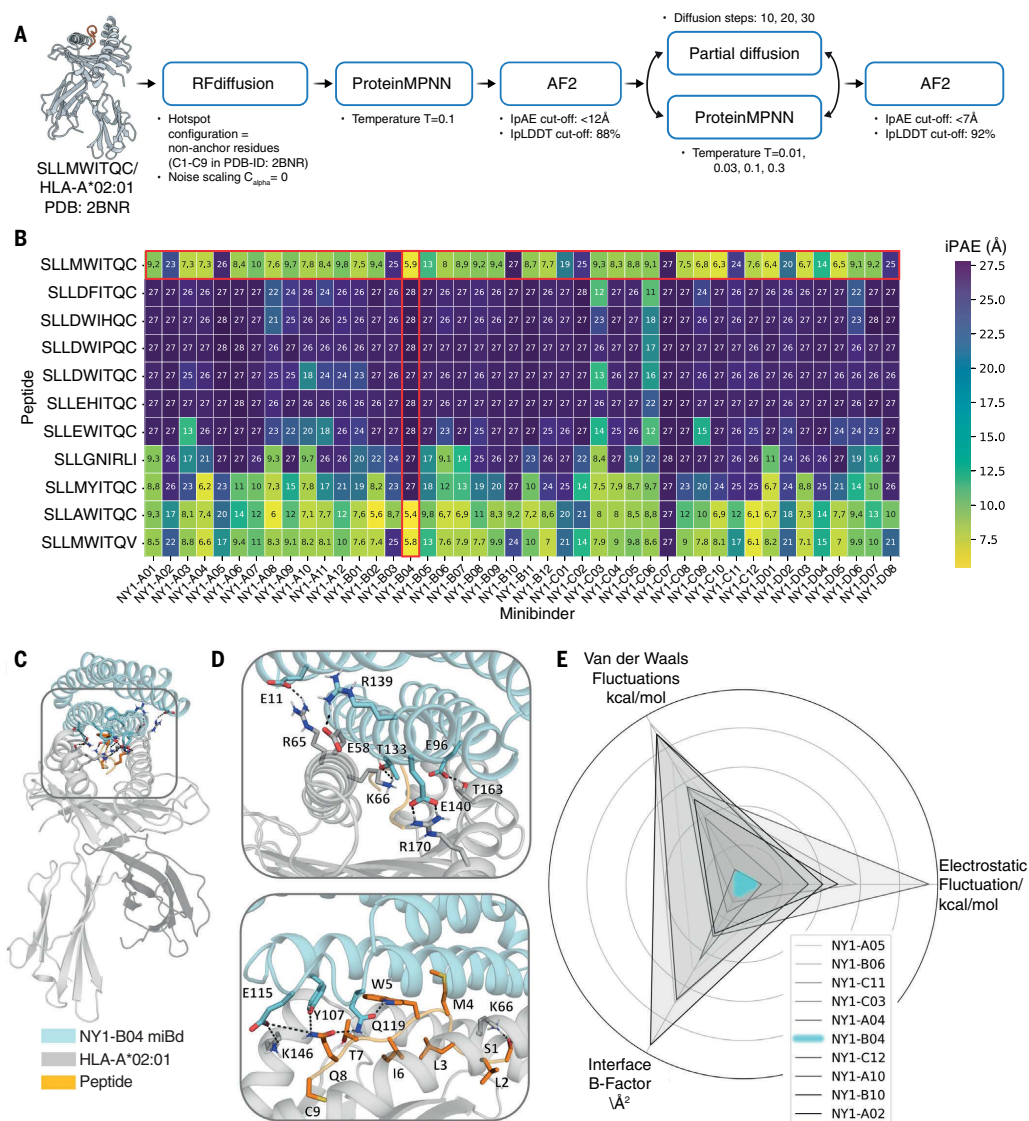


Fig. 1. Design of miBds against NY-ESO-1⁽¹⁵⁷⁻¹⁶⁵⁾ presented on pMHC and down-selection with conformational diversity estimators from MD simulations. (A) Pipeline for the identification of miBds for the SLLMWITQC/HLA-A*02:01 pMHC complex using the crystal structure (PDB accession code 2BNR) as input, with specifications indicated. We generated 5500 backbone structures of the miBd (45 s each on an Nvidia A100 GPU with 40 GB). Next, the backbone structures were subjected to a combination of sequence interpretation using ProteinMPNN (four sequences per backbone) and AF2 structure prediction based on the sequence input for each miBd design (backbone with sequence) that took 20 s and 1.5 min, respectively. (B) Heatmap showing iPAE scores in Angstrom (Å) predicted using AF2 as the parameter for potential binding interaction of the 44 miBds selected for in vitro characterization toward the original SLLMWITQC peptide and 10 selected variant peptides. Numbers in each square represent the iPAE for the given complex, and red squares mark the target SLLMWITQC peptide (horizontal) and the NY-B04 miBd (vertical). (C) Characterization of dominant salt-bridge and hydrogen-bond interactions involved in stabilizing NY1-B04 with the target pMHC SLLMWITQC/HLA-A*02:01. (D) Magnification highlighting the salt-bridge and hydrogen-bond interactions of NY1-B04 (cyan) with HLA-A*02:01 (gray) (top) and the interactions of the peptide (orange) SLLMWITQC with NY1-B04 (bottom). (E) MD simulations based on down-selection using criteria such as fluctuations in electrostatic and van der Waals interactions combined with an interface-residue-B-factor analysis as estimators for conformational flexibility in the binding interface identify miBd NY1-B04 (cyan) as the most promising miBd-pMHC complex. For each miBd, 4 × 500 ns were calculated, with an average performance on a GPU of 200 ns/day.

testing. Thus, identifying properties and features that can decrease the false-positive rate is crucial. Protein-protein binding is determined by an interplay of enthalpy and entropy, so considering both components is critical when characterizing and predicting protein-protein interfaces. To define the protein-protein interfaces for residue-wise flexibility (B-factor), electrostatic, and van der Waals interaction

energy fluctuations, we considered the salt-bridge and hydrogen-bond interactions for the miBd-pMHC complexes that occur at a rate of >10% throughout the simulations (Fig. 1C). The AF2-predicted complex highlighted salt-bridge and hydrogen-bond interactions of the NY1-B04 miBd with HLA-A*02:01 and the SLLMWITQC peptide (Fig. 1D). Residues E11, E96, T133, R139, and E140 of NY1-B04 formed long-lasting interactions with HLA-A*02:01, and residues Y107, E115, and Q119 interacted with the peptide. The radar plot based on flexibility and/or conformational entropy features showed that NY1-B04 had the lowest conformational diversity in the binding interface and the lowest interaction energy fluctuations, representing the most promising miBd-pMHC complex within the dataset (Fig. 1E). Therefore, we hypothesized that NY1-B04 was the most likely to bind the target pMHC, and we proceeded to test all 44 miBds to explore this prediction.

In vitro selection and specificity screening of SLLMWITQC/HLA-A*02:01 miBds

To validate binding and specificity toward pMHC-I experimentally, synthetic genes encoding all 44 miBd designs targeting SLLMWITQC/HLA-A*02:01 (data S1) were cloned into a lentiviral CD28-CD3 ζ CAR vector for expression on the cell surface (Fig. 2A). Using a mammalian surface display system based on miBd-CAR-transduced Jurkat cells, we evaluated the capacity of the miBd library to bind to SLLMWITQC/HLA-A*02:01 tetramers (Fig. 2B and fig. S3A). From this library, we identified one candidate, NY1-B04, that was >2 log₂FC enriched when comparing sorted cells with cells before sorting (Fig. 2, B and C, and data S2). NY1-B04 pMHC-binding characteristics aligned with the molecular dynamics (MD) predictions to deselect false-positive binders. NY1-B04 was individually cloned, and binding of SLLMWITQC/HLA-A*02:01 tetramers was confirmed in CD3 knockout (KO) Jurkat cells (Fig. 2D and fig. S3A). We further measured the binding affinity using biolayer interferometry and obtained a low nanomolar dissociation constant ($K_d = 6.92$ nM with an SD of 0.7 nM) (Fig. 2E).

To confirm the AF2-predicted binding pose of NY1-B04 to SLLMWITQC/HLA-A*02:01, we determined a cryo-electron microscopy (cryo-EM) structure of the complex (Fig. 2F and figs. S4 and S5). We obtained a 3.8-Å-resolution structure for the ~60-kDa complex (figs. S4 and S5, A to E), and found the same interacting residues in the binding interface as predicted by AF2 and MD simulations (Fig. 1, C to E, and fig. S5E). The overlay of the cryo-EM structure with the AF2 model (fig. S4) revealed an overall highly similar

binding pose and angle of approach of NY1-B04 to the pMHC (overall C-alpha root mean square deviation (RMSD): 1.8 Å, peptide Ca-RMSD: 0.92 Å). Figure 2F shows a magnified image of the binding interface, highlighting the critical interacting residues of NY1-B04 with the peptide and emphasizing the importance of residues 1, 4, 5, 7, and 8 of the peptide in stabilizing the complex.

IMPAC-T cell-mediated cytotoxicity toward cancer cells

To evaluate the functional potential of NY1-B04 as a cancer-targeting molecule and not just as a target-binding molecule, we investigated whether the NY1-B04 miBd could facilitate the in vitro killing of cells presenting SLLMWITQC/HLA-A*02:01 complexes on their surface when incorporated into a CAR. We generated primary human T cells expressing an NY1-B04-CD8 α hinge-CD28-CD3 ζ CAR called

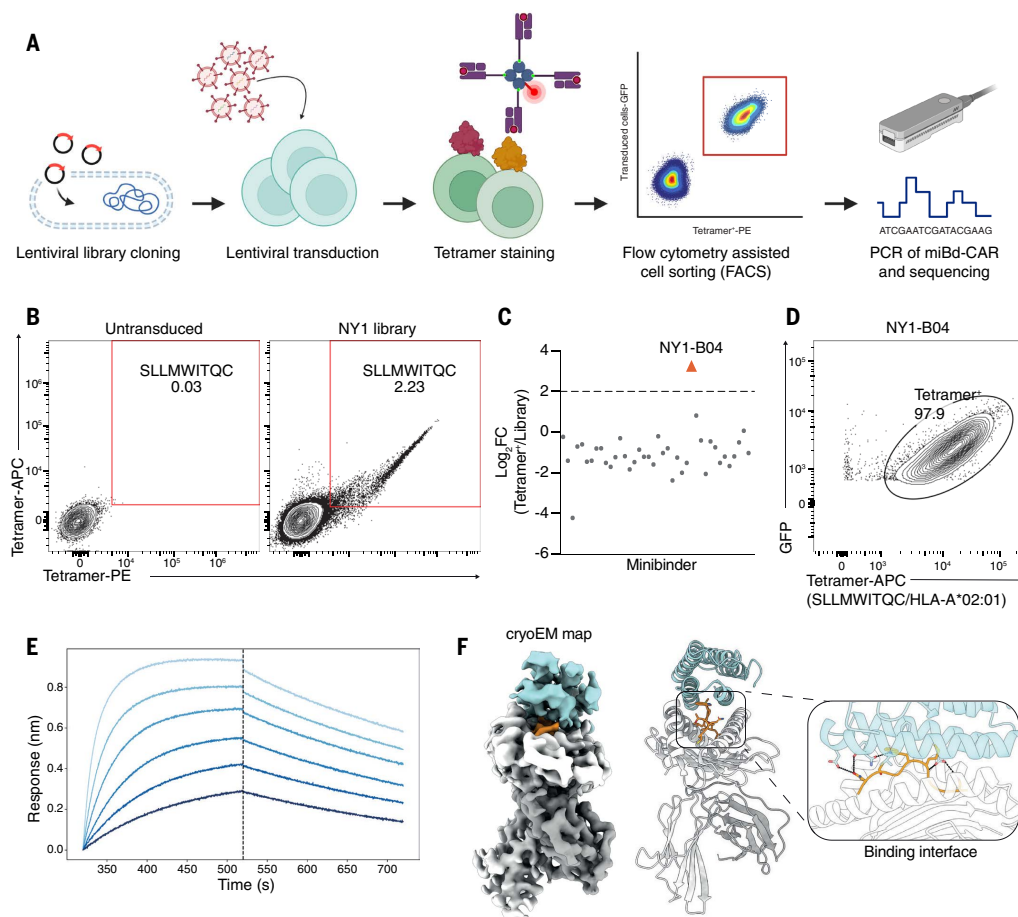


Fig. 2. Experimental selection and characterization of miBds. (A) Mammalian surface display workflow. (B and C) A library of 44 de novo-generated miBds (NY1) was lentivirally transduced into CD3 KO Jurkat cells. Cells were stained with SLLMWITQC/HLA-A*02:01 tetramers [using phycoerythrin (PE) and allophycocyanin (APC)], double-positive cells were sorted, and the integrated miBds were sequenced. Data are representative of three screens ($N = 3$). (B) Flow plots showing the sorted cells gated on singlet/live/CD3⁺/GFP⁺ cells (fig. S3A) and (C) log₂FC of miBds in the sorted tetramer⁺ cells compared with the cells before sorting for the identification of NY1-B04 miBd targeting SLLMWITQC/HLA-A*02:01 (fig. S3A). Each dot represents a miBd in the library. miBds for which it was not possible to determine the log₂FC were excluded from the plot. Log₂FC = 2 is marked with a dotted line. (D) NY1-B04 binding to SLLMWITQC/HLA-A*02:01 was confirmed by a single tetramer-APC stain gated on singlet/live/CD3⁺/GFP⁺ cells (as shown in fig. S3A). Data are representative of two stains ($N = 2$). (E) Affinity measurement of NY1-B04 with biolayer interferometry showing the association and dissociation of pMHC SLLMWITQC/HLA-A*02:01. Each shade of blue represents a concentration of the pMHC at a twofold serial dilution starting from 1 μ M (lightest blue curve). 6xHis-tagged NY1-B04 was captured on HIS1K biosensor tips. Binding data were fitted globally using 1:1 model. The dissociation constant was determined to be 6.9 nM (SD of 0.7 nM). (F) Cryo-EM map (3.8-Å resolution) and structure model of NY1-B04 in complex with SLLMWITQC/HLA-A*02:01 (PDB accession code 9NNF) confirm the predicted binding mode from AF2 (fig. S3A). The NY1-B04 is shown in cyan, the peptide is shown in orange, and the MHC molecule is shown in gray. The magnification of the binding interface highlights the interactions of NY1-B04 with the peptide.

“immune-targeting minibinder for pMHC activated cytotoxic T” (IMPAC-T) cells by lentiviral transduction (Fig. 3, A and B). Subsequently, we evaluated the capacity of these cells to kill NY-ESO-1⁺ HLA-A*02:01⁺ A375 melanoma cells modified to express mCherry. After 24 hours of exposure to NY1-B04 IMPAC-T cells, A375 cell death was higher in IMPAC-T conditions compared with the nontransduced controls (Fig. 3, C and D, and movies S1 and S2). This provides proof of principle that de novo-designed miBds have the potential to be used in cell-based therapies.

Evaluation of cross-reactivity of NY1-B04

To investigate the cross-reactivity profile of NY1-B04, we stained the NY1-B04 IMPAC-T cells with a variety of pMHC tetramers. We assessed the ability of NY1-B04 to bind the peptide-HLA-A*02:01 variants tested during the *in silico* cross-panning (Fig. 3, E and F, and figs. S3A and S6A). NY1-B04 interacted with the single mutants (mutants marked in bold) SLLAWITQC (M4A), SLLMYITQC (W5Y), and SLLMWITQV (C9V), of which M4A and C9V were ranked with a low iPAE (<10 Å). Furthermore, NY1-B04 also bound SLL~~EH~~ITQC to some extent but did not bind SLL~~EW~~ITQC (Fig. 4, E and F). Structurally, the negatively charged M4E mutation disrupts binding, and the histidine in W5H could partially stabilize through a hydrogen bond (fig. S7A). Overall, we identified complete abrogation of NY1-B04 binding when introducing polar, negatively charged amino acids (aspartic acid and glutamic acid) in position 4. These data indicate that the *in silico* cross-panning was capable of correctly predicting binding for eight of 10 tested variants (accuracy $N = 10 = 8/10$) (fig. S7B).

To further investigate how subtle electrostatic changes, such as in alanine and glycine, and positively charged side chains, such as in arginine, affected the binding interactions, we performed single-amino acid substitution scans both *in vitro* and *in silico*. *In vitro*, we found that alanine and glycine substitutions on three of nine positions (positions L2, M4, and T7) decreased pMHC binding, whereas substitution with arginine disrupted or decreased binding in seven of nine positions (positions 1 to 7) (Fig. 3, G and H, and fig. S6B). The *in silico* binding predictions revealed that the binary categorization based on iPAE values

led in 19 of 27 cases to the correct identification of binding and non-binding (accuracy _{$N=27$} = 19/27) (fig. S7C). Structurally, the NY1-B04 encloses position M4 and W5 in a combined hydrophobic pocket (M93, A97, A100, F123, L127, and A130) (fig. S7A). This entails the repulsion of polar residues, especially charged residues such as M4E, M4D, M4R, and W5R.

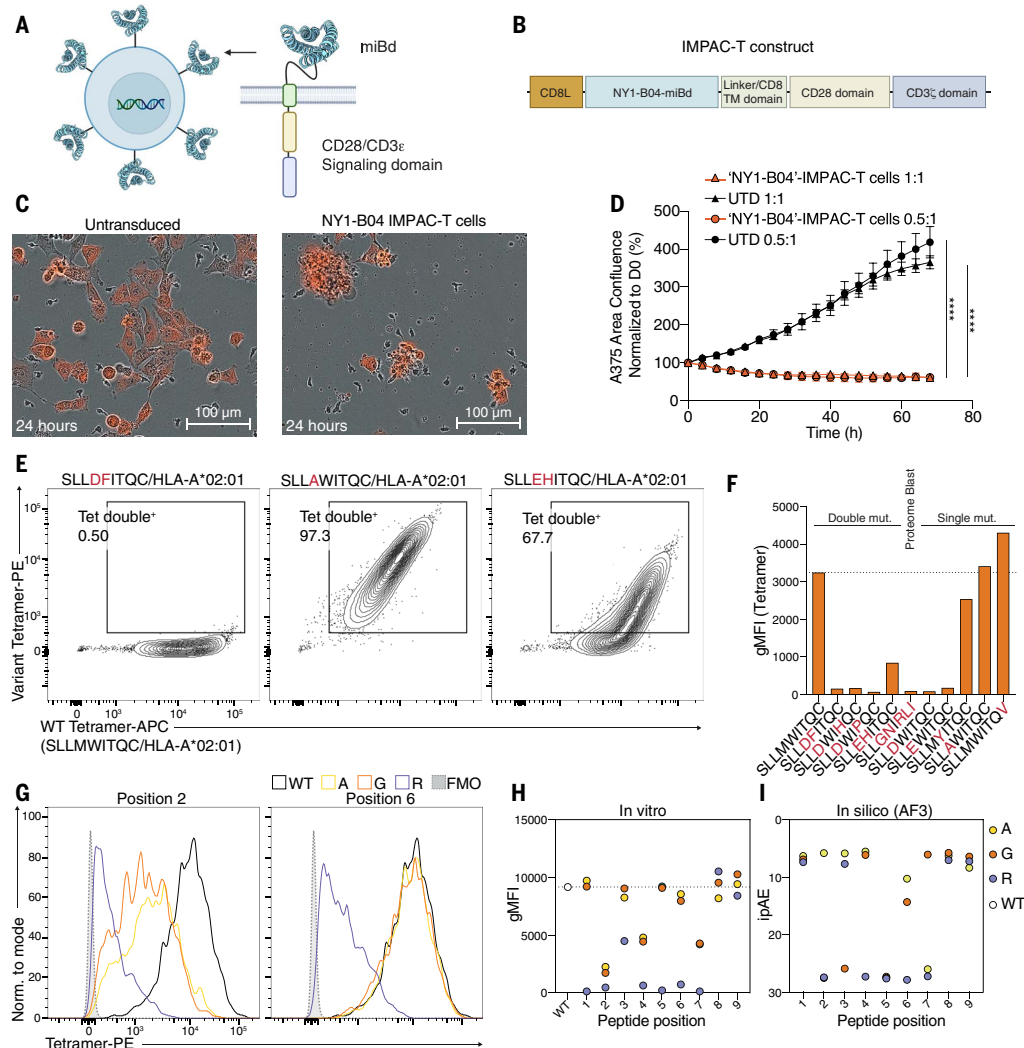


Fig. 3. IMPAC-T cell-mediated cytotoxicity against NY-ESO-1⁺ A375 cancer cells. (A) Overview of the de novo IMPAC-T cell concept with miBds fused extracellularly to intracellular signaling domains. (B) Construct design of the NY1-B04-IMPAC-T construct with a CD8 leader, miBd, CD8a hinge region, and CD28/CD3 ζ intracellular signaling domains. (C) Phase contrast images overlaid with mCherry fluorescence showing the mCherry⁺ A375 cancer cells after 24 hours of incubation of untransduced (UTD) primary T cells (left) or primary NY1-B04-IMPAC-T cells (right). Scale bar, 100 μ m. (D) Cell death of mCherry⁺ A375 cancer cells was monitored and quantified over a period of ~2 days using Incucyte live-cell imaging, with the red fluorescence area normalized to the D0 area as a measure of confluency. The assay was set up with 0.5:1 and 1:1 effector-to-target ratios. (E) Cross-reactivity screening of NY1-B04 expressed in CD3 KO Jurkat cells against variant peptides showing representative plots of three examples of nonbinding and binding peptides after double staining with SLLMWITQC/HLA-A*02:01 tetramers conjugated to APC and variant tetramers conjugated to PE. Cells were gated on singlet/live/CD3⁺/GFP⁺ cells, as shown in fig. S3. (F) Bar plot showing representative geometric mean fluorescence intensity (gMFI) of single-stained variant tetramer stains. Mutations are shown in red, and the wild-type (WT) tetramer gMFI is indicated with a dotted line. (G and H) Representative histograms (G) and gMFI plot (H) of single substitution scanning data with alanine (A), glycine (G), and arginine (R) in primary T cells transduced with NY1-B04. In (H), gMFI is shown at each amino acid position of the peptide, with the dotted line indicating the gMFI of the WT peptide. (I) iPAE values for A, G, and R single-amino acid substitutions predicted with AF2 initial guess based on AF3 structure prediction of the pMHC. The y axis is inverted, given that a low iPAE value indicates predicted binding to match the data shown in (H). Data shown in (G) and (H) are gated on singlet/live/CD3⁺/CD8⁺/GFP⁺ cells, as shown in fig. S6B. Data in (D) are means $\pm$ SD ($N = 3$ technical replicates) representative of two independent experiments ($N = 2$), and data in (E) and (F) are representative gMFIs of two experiments ($N = 2$). **** $P < 0.0001$ in (D) as evaluated by a Wilcoxon matched-pairs signed rank test comparing the area confluence of each time point between UTD and NY1-B04-IMPAC-T samples (0.5:1 and 1:1 are compared separately).

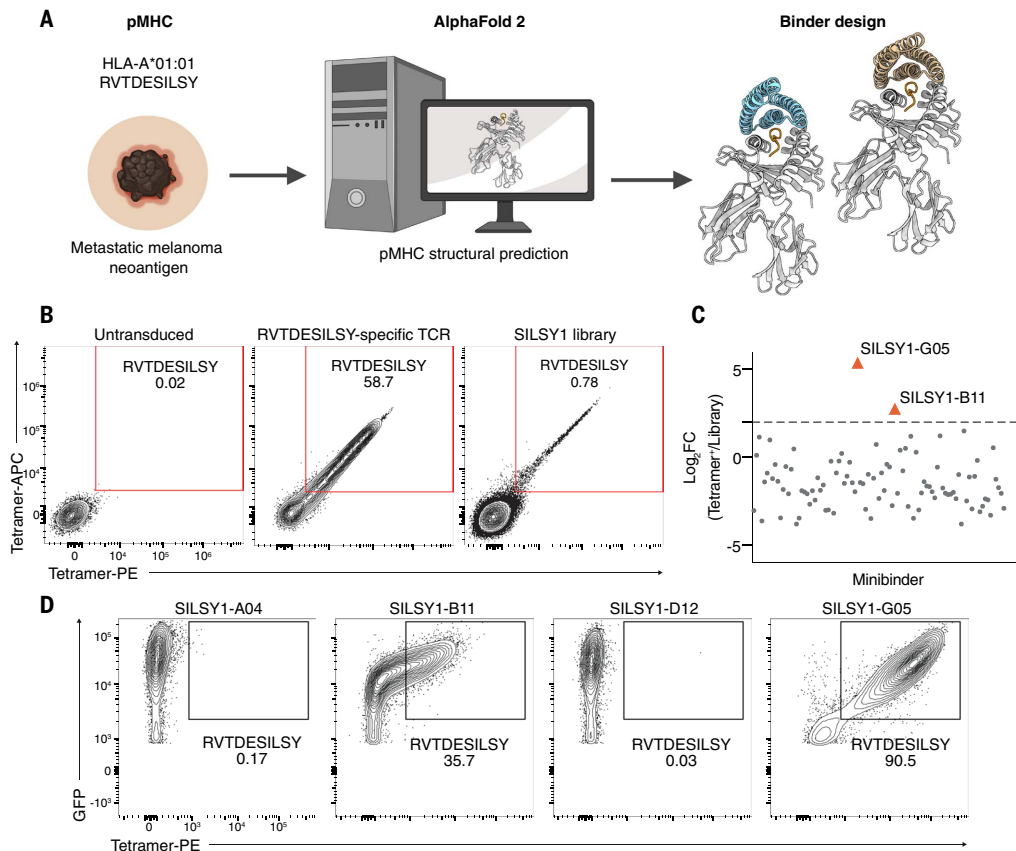


Fig. 4. Proof of concept of applying the design pipeline to neoantigens. (A) The structure of the neoantigen RVTDESILSY in complex with HLA-A*01:01 was predicted using AF2 and miBds designed following the previously described pipeline. (B) A library of 95 predicted miBds (SILSY1) was lentivirally transduced in CD3 KO Jurkat cells, and the cells were stained with RVTDESILSY/HLA-A*01:01 tetramers (PE and APC). Double-positive populations were sorted, and the integrated miBds were sequenced. (C) Log₂FC of miBds in the sorted population (tetramer⁺) compared with the cells before sorting (library) for the identification of two miBds targeting RVTDESILSY/HLA-A*01:01. Each dot represents a miBd in the library. miBds for which it was not possible to determine the log₂FC were excluded from the plot. Log₂FC = 2 is marked with a dotted line. (D) Flow plots of RVTDESILSY/HLA-A*01:01 tetramer–PE-stained CD3 KO Jurkat cells expressing each of the four highest-scoring miBds from (B) and (C). Flow plots in (B) and (D) show cells gated on singlet/live/CD3⁺/GFP⁺ cells, as shown in fig. S3A. Data in (B) and (C) are representative data from two screens (*N* = 2), and data in (D) is from an experiment performed once.

Further, position L2 seems to determine the binding pose of the peptide in the MHC, which could be crucial for the interaction of E96 in NY1-B04 toward the backbone at position M4 (fig. S7A).

Extending the miBd discovery pipeline to neoantigens

A potential hurdle for broader application of machine learning–guided design of pMHC-I miBds is that crystal structures of pMHC-I rarely are available. Therefore, we sought to further test the limits of our approach by targeting the neoantigen RVTDESILSY derived from the mutation AKAP9^{P1796L} identified in metastatic melanoma, presented on HLA-A*01:01. The RVTDESILSY/HLA-A*01:01 complex was previously demonstrated as subject for T cell recognition in both blood and tumor-infiltrating lymphocyte infusion product in a patient with melanoma (27). On the basis of structure predictions, the neoantigen resulted in a targetable protruding leucine in the MHC-I groove (27). Because no experimental structure existed, we modeled the pMHC-I complex using AF2. Upon selecting the most confident structure based on AF2 (pLDDT of 95.7), we designed miBds using our pipeline (Fig. 4A).

We selected 96 miBds (data S1), introduced them for surface expression as a pool in CD3 KO Jurkat cells, and evaluated their binding to RVTDESILSY/HLA-A*01:01 tetramers in vitro. We found that the

fluorochrome intensity of the tetramer⁺ miBd library was comparable to that of the positive control primary T cells with CRISPR/Cas9 knock-in of a RVTDESILSY/HLA-A*01:01-specific TCR (Fig. 4B and fig. S3A).

Sequencing of the sorted population identified two distinct miBds (SILSY1-B11 and SILSY1-G05), constituting a success rate of ~2% (Fig. 4, C and D; fig. S3A; and data S2). This demonstrated the possibility of applying our design pipeline in the context of personalized cancer therapy for neoantigens for which experimental structures are not available.

Discussion

A challenge in cancer immunotherapy has been targeting pMHC with high specificity and efficacy (28). Traditionally, this has been achieved using TCRs identified by screening T cell responsiveness and specificity from cancer patients (1, 29–31). However, this is extremely time and resource intensive, and the identified TCRs often lack the required specificity, which may either prevent treatment or provide a risk of fatal side effects (28). Thus, there has been a need to develop approaches to overcoming these challenges and unlocking the potential of targeted T cell therapy. We leveraged leading-edge protein design approaches to generatively design SLLMWITQC/HLA-A*02:01 pMHC-I miBds de novo. We modified current design pipelines (17, 21, 22) to design pMHC-I miBds while screening for specificity in silico with the aim of reducing

the risk of off-target pMHC-I binding. We demonstrated that these rapid (<10 s per miBd–target complex) computational predictions largely correlated with in vitro specificity screening results, substantiating the potential that this approach has in de-risking T-cell therapy. Although further improvement in the current predictors and possibly a combination of several orthogonal metrics are needed, our results show the potential benefit of computational off-target binding screens. We propose that by using a mammalian surface display in a CAR construct when identifying binders, miBds that would not bind the target when presented in a CAR were deselected in the pooled in vitro screen. We further found that the miBds we designed facilitated target cell killing when using the miBd as the recognition element in a CAR format to generate IMPAC-T cells. This indicates that a key next step is the in vivo validation of CARs incorporating machine learning–guided pMHC miBds that can direct T cells toward the pMHC-I antigen of interest and activate engagement, signaling, and, consequently, cytotoxicity toward pMHC-I-expressing target cells.

Machine learning–guided design holds great promise for a rapid therapeutic discovery pipeline to target both shared cancer antigens, such as NY-ESO-1, and patient-specific pMHC targets, such as neoantigens (32). The timelines for identifying an miBd for both SLLMWITQC/

HLA-A*02:01, where the crystal structure is available, and for RVTDESILSY/HLA-A*01:01, where no prior pMHC-I structure was described, were both ~4 to 6 weeks. Given the success for both known crystallized and structurally predicted targets, we suggest that the pipeline we present here could be broadly applicable to the generation of miBds to other peptides and HLAs.

Another strength of this approach is that specific miBds were identified with a relatively high success rate (1:50), which we demonstrated could be increased by using MD approaches. We found that considering flexibility in the binding interface accelerated the identification of potential binders, and a direct comparison of the binding interfaces was facilitated by the high similarity of the interaction sites among these potential binders. Thus, considering interface flexibility and dynamics can substantially reduce the false-positive rate and further expedite the identification of the most promising miBd-pMHC-I complex. A limitation of applying de novo proteins in vivo is the risk that they could induce an immune response that could include endogenous TCR-driven immunogenicity against IMPAC-T cells and antidrug antibodies against de novo proteins. Because of their small size, miBds are unlikely to raise substantially larger immunogenicity than animal-derived antibodies, nonetheless, thorough in vitro specificity testing and well-structured in vivo studies evaluating and mitigating this risk will be needed.

Beyond cancer applications, the versatility of miBds makes them promising for targeting viral antigens presented on infected cells. This could be transformative for diseases such as HIV and hepatitis C, where broad-spectrum targeting remains challenging because of high mutation rates. Additionally, miBds could be used to block autoimmune TCR-pMHC interactions, as proposed for TCR-like antibodies (33), expanding their therapeutic scope. Further, the findings shown here suggest the use of pMHC-binding miBds as carriers of payloads or pMHC-targeting bispecific T cell engagers. Given the previous success of generatively designed miBds in vivo (21), these findings are likely to not only advance the field of cancer immunotherapy by introducing a rapid and precise method for de novo pMHC targeting but also pave the way for targeted patient-centric and personalized therapies.

REFERENCES AND NOTES

- C. A. Klebanoff, S. S. Chandran, B. M. Baker, S. A. Quezada, A. Ribas, *Nat. Rev. Drug Discov.* **22**, 996–1017 (2023).
- Y.-T. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1914–1918 (1997).
- R. Pétremand *et al.*, *Nat. Biotechnol.* **43**, 323–328 (2025).
- M. Arnaud *et al.*, *Nat. Biotechnol.* **40**, 656–660 (2022).
- E. Strønen *et al.*, *Science* **352**, 1337–1341 (2016).
- A. K. Sewell, *Nat. Rev. Immunol.* **12**, 669–677 (2012).
- F. Drost *et al.*, *Cell Genom.* **4**, 100634 (2024).
- M. Nielsen *et al.*, *Immunoinformatics (Amst.)* **16**, 100045 (2024).
- M. Yarmarkovich *et al.*, *Nature* **623**, 820–827 (2023).
- A. Porgador, J. W. Yewdell, Y. Deng, J. R. Bennink, R. N. Germain, *Immunity* **6**, 715–726 (1997).
- T. Dao *et al.*, *Sci. Transl. Med.* **5**, 176ra33 (2013).
- H. Du *et al.*, *Nat. Biotechnol.* (2024).
- G. Stewart-Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5784–5788 (2009).
- C. H. Lee *et al.*, *Front. Immunol.* **11**, 565096 (2020).
- V. K. Karnaukhov *et al.*, *Nat. Comput. Sci.* **4**, 510–521 (2024).
- Y. Zhao *et al.*, *Sci. Adv.* **9**, eabo5128 (2023).
- J. L. Watson *et al.*, *Nature* **620**, 1089–1100 (2023).
- J. Dauparas *et al.*, *Science* **378**, 49–56 (2022).
- J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
- N. R. Bennett *et al.*, *Nat. Commun.* **14**, 2625 (2023).
- S. Vázquez Torres *et al.*, *Nature* **639**, 225–231 (2025).
- S. Vázquez Torres *et al.*, *Nature* **626**, 435–442 (2024).
- J.-L. Chen *et al.*, *J. Exp. Med.* **201**, 1243–1255 (2005).
- J. Abramson *et al.*, *Nature* **630**, 493–500 (2024).
- B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, *Nucleic Acids Res.* **48**, W449–W454 (2020).
- M. Johnson *et al.*, *Nucleic Acids Res.* **36**, W5–W9 (2008).
- N. P. Kristensen *et al.*, *J. Clin. Invest.* **132**, e150535 (2022).
- G. P. Linette *et al.*, *Blood* **122**, 863–871 (2013).
- R. A. Morgan *et al.*, *Science* **314**, 126–129 (2006).
- W. Chour *et al.*, *Commun. Biol.* **6**, 528 (2023).
- E. Giannakopoulou *et al.*, *Nat. Cancer* **4**, 1474–1490 (2023).
- F. Huber *et al.*, *Nat. Biotechnol.* (2024).
- L. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2656–2661 (2014).
- M. L. Fernández-Quintero *et al.*, *MAbs* **15**, 2175319 (2023).

ACKNOWLEDGMENTS

We thank B. Liu, N. F. Greenwood, and J. Bonzanini from the Baker lab for discussions and knowledge exchange, as well as for timing submissions, and S. Sebbaha, research manager from the Flow Cytometry and Imaging Core (FLIC) at DTU Health Tech, for supporting flow cytometry experimental design and data analysis; and the EuroHPC Joint Undertaking for awarding us access to MeluXina, Luxembourg (M.L.F.Q.). BioRender was used for some of the figures. **Funding:** This work was supported by the Alliance program under a EuroTech Universities agreement (T.P.J.); Innovation Fund Denmark [Industrial PhD stipend 2052–00010B (2022) to D.S.W.]; the Lundbeck Foundation (postdoctoral grant R347-2020-2174 to K.H.J.); the Novo Nordisk Foundation (Bioscience and Biomedicine Project Grant 0087077 to S.R.H.); and the European Research Council (ERC Consolidator Grant MIMIC 101045517 to S.R.H.). **Author contributions:** T.P.J., K.H.J., and S.R.H. conceived the project, provided funding, and supervised the work. T.P.J., K.H.J., D.S.W., B.S., M.L.F.-Q., and S.R.H. wrote and reviewed the manuscript. D.S.W., C.R.C., J.S.F., and T.P.J. designed and in silico screened the pMHC miBds. D.S.W., M.D., and A.T.D.E. initiated the design of neoantigen pMHC miBds with input from C.R.C. E.Q.A. and D.S.W. analyzed the secondary structure composition and orientation of the miBds. K.H.J., B.S., and C.R.C. prepared binder libraries, performed mammalian surface display screening, and sequenced hits. M.L.F.-Q., A.G., and A.B.W. performed cryo-EM and data analysis. M.L.F.-Q., K.H.B., S.T., and J.R.L. performed affinity measurements and data analysis. M.L.F.-Q., J.R.L., S.F., and A.B.W. performed MD simulations and data analysis. K.H.J., B.S., C.R.C., M.C.V., I.L., C.R.P., and L.F.V. conducted cell assays and in vitro flow cytometric analysis of miBds. K.K.M. analyzed predicted miBd structures and assisted with sequencing and data analysis. D.S.W., M.D.O., and O.M. developed and improved in silico specificity screening protocols. M.O. designed CAR vectors. E.R.d.T. and T.T. expressed the miBds and MHC monomers. All authors reviewed and approved the final manuscript. **Competing interests:** K.H.J., D.S.W., B.S., S.R.H., and T.P.J. are inventors on a patent related to this work under filing number US 240809-P7151EP00. The remaining authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the study are available in the main text or the supplementary materials. The code for binder design using RFDiffusion has been published previously (17). The Pymol script for assessment and depiction of potential clashes of co-complexes was also published previously (34). The high-resolution structure of miBd B04 in complex with SLLMWITQC/HLA-A*02:01 was deposited in the Protein Data Bank with accession code 9NNF. **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adv0422
Materials and Methods; Figs. S1 to S7; Tables S1 to S3; References (35–50); Data S1 and S2; Movies S1 and S2

Submitted 3 December 2024; accepted 16 May 2025

10.1126/science.adv0422