A High-Throughput Human Display Screen to Identify Target-Specific Binder Proteins via Chimeric Antigen Receptors

Anonymous Authors
Abstract Under Double-Blind Peer Review

Abstract

The design of specific protein binders is a core-goal of biologics design. However, current techniques rely on medium-throughput, laborious, well-by-well screening, or standard phage, yeast, or mRNA display, which screen binders in non-human contexts. To create a human cell-based binder screening platform, here, we developed a high-throughput screening approach based on genetically engineering chimeric antigen receptors (CAR)-expressing cells with libraries of binding proteins, and co-culturing them with target-expressing cells. Using an NFAT-inducible GFP reporter to monitor CAR activation, we establish a novel platform to screen experimentally and artificial intelligence-generated binders toward diverse target substrates on the surface of human cells for downstream therapeutics development.

Introduction

A significant portion of disease-driving proteins, from transcription factors to membrane proteins, are considered intractable by small molecules due to inaccessible binding pockets and largely disordered domains.[1] Biologics-based modalities, such as antibodies, nanobodies, and peptides, are attractive alternatives for target binding, as they do not rely on the presence of active sites.[1] Moreover, generative artificial intelligence-based computational methods have accelerated the development of these protein binders for structured and disordered targets. As a recent example, the RFDiffusion model has successfully leveraged denoising diffusion probabilistic modeling to design mini protein binders to structured target proteins.[2] New language model-based methods, such as PepMLM, PepPrCLIP, and SaLT&PepPr,[3–5] now offer methods for peptide binder design conditioned purely on the target sequence. The evaluation of these models to design and down-select high-affinity and specific binders from their expressive latent spaces is of strong interest.

Chimeric antigen receptor (CAR)-T cell therapy has been revolutionary, as it has produced remarkably effective and durable clinical responses for numerous cancer indications.[6] The antigen binding domain of CARs, specifically, confers target antigen specificity to activatory signals on CAR-expressing T cells.[6,7] While the original CAR-based therapies were primarily targeted at the CD19 receptor, a protein expressed on the surface of both normal and malignant B cells, genetic engineering efforts have redirected CARs to new targets, thus making additional diseases tractable via this therapeutic modality.[8,9] In this study, we have repurposed the CAR platform to create a human display assay to screen target-binding proteins on the surface of CARs. To do this, we genetically engineered Jurkat76 cells to express lentiviral-integrated libraries of binders in CARs, and screen them against Chinese Hamster Ovary (CHO) cells expressing their intended targets. By leveraging an NFAT-responsive dsGFP reporter, we have observed green fluorescence upon target binding in Jurkat76 cells, thus motivating downstream applications to evaluate new binder design models and develop CARs for diverse targets of interest.

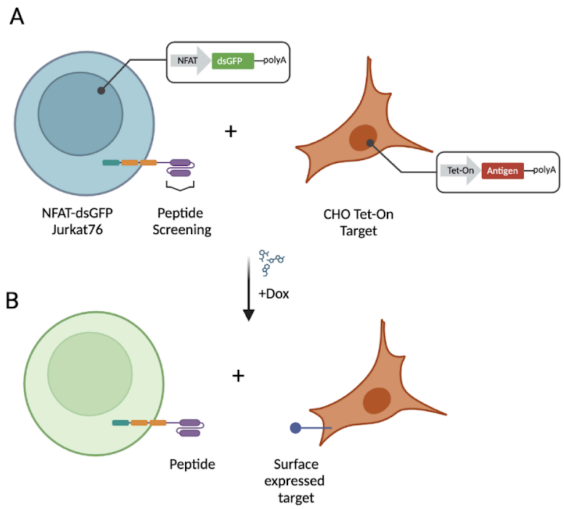


Figure 1. Concepts of modular chimeric antigen receptor derived platform. (**A**) Jurkat76-derived cell line engineered to express destabilized GFP under the control of a NFAT inducible promoter. CHO cells serve as a target cell line with inducible target expression under the control of Tet-on promoter. (**B**) Target of interest was expressed on the cell surface after adding doxycycline, the Jurkat76 cells expressed dsGFP when antigen-antigen binding domain interactions occur.

Results

Modular chimeric antigen receptor-derived platform

To generate a novel mammalian cell surface binding assay, a modular CAR-derived platform was designed (**Figure 1**). Briefly, antigen binding domains were cloned in lieu of conventional scFvs into a lentivirus construct expressing a 3rd generation CAR construct followed by a puromycin selection and a truncated NGFR co-expression marker. Jurkat76 T cells were chosen for proof-of-concept experiments due to their ease of use, homogeneity, and availability of reporter cell lines. Importantly, the NFAT signaling machinery is highly conserved in the Jurkat76 cell line [10]. We produced lentivirus and used it to transduce a novel Jurkat76 derived cell line engineered to express destabilized GFP under the control of an NFAT inducible promoter thereby linking GFP fluorescence intensity to CAR-Antigen binding affinity. To create target cell lines, we cloned antigens of interest into a custom Tet-On

compatible lentivirus construct and used it to transduce Tet-On CHO cells (**Figure 1A**). Doxycycline-inducible target cells and binding domain expressing CAR-NFAT dsGreen Jurkat76 cells were cocultured in different concentrations of the doxycycline to evaluate the effect of antigen density on antigen-antigen binding domain interactions (**Figure 1B**).

To test this chimeric antigen receptor-derived platform for binding interaction evaluation, we first chose CD19 as a target. The Tet-on CD19 CHO cells were co-cultured with the CD19 scFV induced Jurkat76 cell line. After adding various concentrations of doxycycline, the Jurkat76 dsGFP reporter cell line showed a distinct dosage-dependent increase in dsGFP expression level(**Figure 2**). These results indicate that the system may serve as a strong platform for further high-throughput screening.

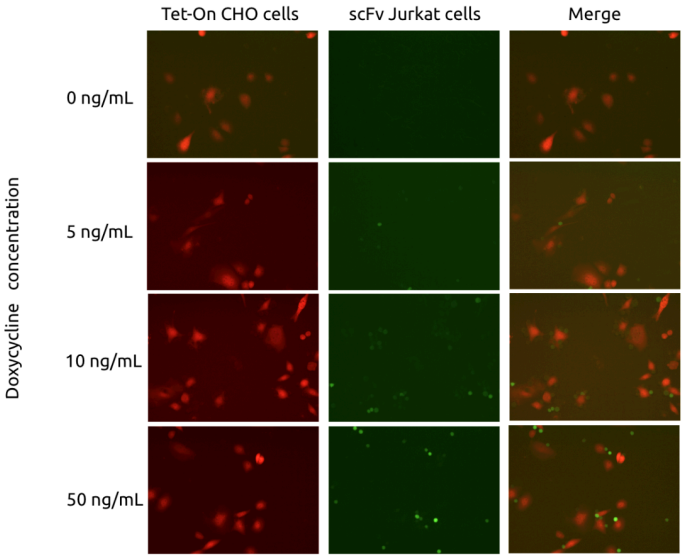


Figure 2. Binding interaction evaluation of CD19-binding scFv under various dosages of doxycycline.

High-throughput screening of language model-designed peptides

To develop a preliminary high-throughput screen with our CAR assay, we designed 1000 peptides for a critical target in glioblastoma multiforme, EGFRvIII, by leveraging both structure-based and sequence-based models that can derive peptides binders conditioned only on the target structure or sequence (**Table 1**).[2–4] We first synthesized a Twist oligo pool encoding these peptides and inserted them into a transfer vector for lentivirus packaging. The lentivirus of the peptide library was then transduced into a dsGFP reporter Jurkat76 cell line. Concurrently, we expressed EGFRvIII on the cell surface of CHO cells under the control of doxycycline supplement. As a result, target-binder interaction activated the downstream of dsGFP reporter and all dsGFP positive cells were therefore sorted using Fluorescence-activated Cell Sorting (FACS). Genomic DNA was extracted and the peptide coding sequences were further amplified for final Next Generation Sequencing (**Figure 3**). Finally, we calculated the significance vs. fold change for each peptide between the post-sorted and pre-sorted populations. Peptide enrichment was defined using a paired two-tailed DESeq2 test with Benjamini–Hochberg correction.[11] As this work is still ongoing, we anticipate that these screening results will demonstrate the binding capability and the large-scale screening provide a platform for functional validation of binders generated from generative artificial intelligence algorithms.

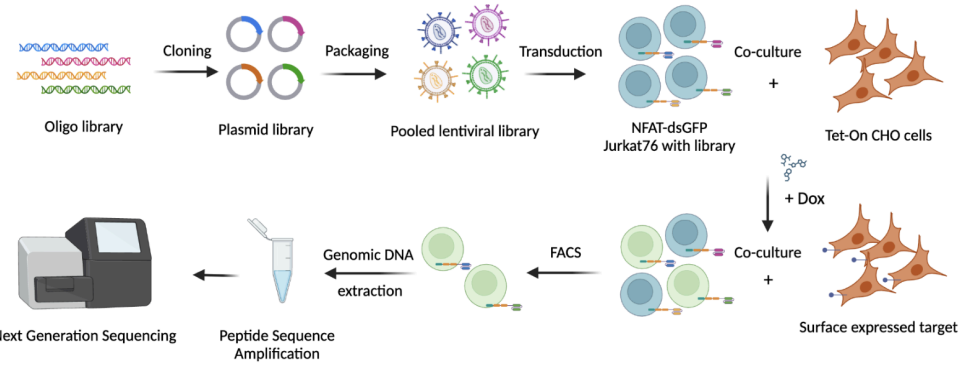
Table 1. Peptides for high-throughput CAR screening.

Algorithm	# of Peptides	Reference
PepMLM	250	Chen T, et al. PepMLM: Target Sequence-Conditioned Generation of Peptide Binders via Masked Language Modeling. *arXiv* (2023)
RFDiffusion	250	Watson, et al. *De novo* design of protein structure and function with RFdiffusion. *Nature* (2023)
PepPrCLIP	499	Bhat, et al. *De Novo* Generation and Prioritization of Target-Binding Peptide Motifs from Sequence Alone. *bioRxiv* (2023)
Positive control	1	Mansour, S. et al. Identification of a novel peptide ligand for the cancer-specific receptor mutation EGFRvIII using high-throughput sequencing of phage-selected peptides. *Sci Rep* (2022).

Conclusions

Overall, in this work, we have laid the foundation for a high-throughput, human cell-based display screen to evaluate target binders to any protein of interest. We are currently screening binders derived for diverse target-conditioned generative artificial intelligence algorithms, and can rapidly extend screening to other targets, including those that are post-translationally modified or heavily disordered. We anticipate that our CAR-based screening approach can enable the design of numerous therapeutic modalities dependent on viable target binders, as well as the development of new CAR-based therapies to diverse disease indications.

Figure 3. Schematic of high-throughput CAR screening of target binders.



References

1. Generative design of therapeutics that bind and modulate protein states. *Current Opinion in Biomedical Engineering* **28**, 100496 (2023).
2. Watson, J. L. *et al.* De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089–1100 (2023).
3. Chen, T. *et al.* PepMLM: Target Sequence-Conditioned Generation of Peptide Binders via Masked Language Modeling. *ArXiv* (2023) doi:10.48550/ARXIV.1706.03762.
4. Bhat, S. *et al.* De Novo Generation and Prioritization of Target-Binding Peptide Motifs from Sequence Alone. *bioRxiv* 2023.06.26.546591 (2023) doi:10.1101/2023.06.26.546591.
5. Brixi, G. *et al.* SaLT&PepPr is an interface-predicting language model for designing peptide-guided protein degraders. *Communications Biology* **6**, 1–10 (2023).
6. Sterner, R. C. & Sterner, R. M. CAR-T cell therapy: current limitations and potential strategies. *Blood Cancer J.* **11**, 1–11 (2021).
7. De Marco, R. C., Monzo, H. J. & Ojala, P. M. CAR T Cell Therapy: A Versatile Living Drug. *Int. J. Mol. Sci.* **24**, (2023).
8. Park, J. H., Geyer, M. B. & Brentjens, R. J. CD19-targeted CAR T-cell therapeutics for hematologic malignancies: interpreting clinical outcomes to date. *Blood* **127**, 3312–3320 (2016).
9. Marhelava, K., Krawczyk, M., Firczuk, M. & Fidyt, K. CAR-T Cells Shoot for New Targets: Novel Approaches to Boost Adoptive Cell Therapy for B Cell-Derived Malignancies. *Cells* **11**, (2022).
10. Liang, Y., Zhou, Y. & Shen, P. NF-kappaB and its regulation on the immune system. *Cell. Mol. Immunol.* **1**, 343–350 (2004).
11. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).