A MAMMALIAN HIGH-THROUGHPUT ASSAY TO SCREEN AI-DESIGNED PROTEIN DEGRADERS

Lin Zhao[1], **Aastha Pal**[2], **Tong Chen**[1], **Pranam Chatterjee**[1,3,4,†]
[1]Department of Biomedical Engineering, Duke University
[2]Department of Radiation Oncology, Stanford University
[3]Department of Computer Science, Duke University
[4]Department of Biostatistics and Bioinformatics, Duke University

[†]Corresponding author: pranam.chatterjee@duke.edu

ABSTRACT

The development of specific protein binders is crucial for biologics and targeted protein degradation (TPD) therapies. However, existing screening methods are low-throughput, labor-intensive, and often rely on non-human display systems such as phage, yeast, or mRNA display, limiting their translational relevance. To address this, we developed a high-throughput, human cell-based binder screening platform that enables the functional evaluation of artificial intelligence (AI)-designed peptide binders in a mammalian context. Our approach utilizes genetically-encodable, doxycycline-inducible ubiquibodies (uAbs), where a library of computationally designed "guide" peptides is fused to an E3 ubiquitin ligase domain, enabling modular, CRISPR-like TPD. By monitoring degradation through an mCherry-fused target protein, we screen and validate AI-generated binders in a physiologically relevant setting. We successfully apply this platform to identify functional binders for EWS::FLI1 and β-catenin, two highly challenging oncogenic targets. Overall, our approach facilitates the discovery of AI-designed binders for diverse therapeutic applications.

1 INTRODUCTION

A significant portion of disease-driving proteins, including transcription factors and fusion oncoproteins, are considered intractable by small molecules due to inaccessible binding pockets and largely disordered domains (Chen et al., 2023b). Biologics-based modalities, such as antibodies, nanobodies, and peptides, offer an attractive alternative for targeting these proteins, as they do not rely on active sites for binding. Recent advances in generative artificial intelligence have further accelerated the development of such protein binders. For instance, the RFDiffusion model has successfully leveraged denoising diffusion probabilistic modeling to design mini-protein binders for structured target proteins (Watson et al., 2023).

A novel targeted protein degradation (TPD) strategy has emerged that genetically fuses target-specific short "guide" peptides—designed using both structure- and sequence-based algorithms—to the ubiquitin conjugation domain of the human CHIP E3 ubiquitin ligase, creating CRISPR-like "ubiquibodies" (uAbs) for TPD (Chatterjee et al., 2020; Brixi et al., 2023; Bhat et al., 2025; Chen et al., 2023a). Recent advances in protein language models, including PepMLM, PepPrCLIP, SaLT&PepPr, and moPPIt, now enable peptide binder design purely from target sequences (Brixi et al., 2023; Bhat et al., 2025; Chen et al., 2023a; 2024). Evaluating these models to design and down-select specific binders from their expressive latent spaces is of strong interest.

High-throughput CRISPR screening (CRISPR-HTS) is a powerful genome-editing approach that enables large-scale functional studies of gene perturbations Bock et al. (2022). These screens utilize programmable nucleases, such as Cas9 for gene knockout or catalytically inactive Cas9 for transcriptional modulation, to systematically assess gene fuciton (Gilbert et al., 2014; Konermann et al., 2015). By introducing a large library of single-guide RNAs (sgRNAs) into a population of cells,

followed by selection and deep sequencing, CRISPR-HTS identifies sgRNAs that are enriched or depleted in response to functional perturbations (Shalem et al., 2015; Yang et al., 2023).

Since uAbs are genetically encodable, we aim to develop a CRISPR-like HTS approach to identify target-specific degraders by leveraging language model-derived binder peptides for proteome editing. To achieve this, we first establish HEK293T cell lines expressing an intended target fused to an mCherry reporter. We then genetically engineer these target cells to express doxycycline-inducible uAbs, where a library of AI-designed peptides are fused to an E3 ubiquitin ligase domain, CHIPΔTPR Brixi et al. (2023); Chatterjee et al. (2020); Bhat et al. (2025). Functional binding peptides therefore lead to uAb-mediated degradation, resulting in decreased mCherry fluorescence and enabling the selection of mCherry-decreased cells for sequencing. Our results demonstrate that this high-throughput assay facilitates the down-selection of AI-designed peptide binders for challenging targets such as oncofusion protein, EWS::FLI1 May et al. (1993) and transcription factor, β-catenin Yu et al. (2021). These results highlight the potential of our platform to evaluate new binder design models and develop CRISPR-like HTS for diverse therapeutic targets.

2 METHODS

2.1 BINDER DESIGN

Novel binding peptides designed in this study were generated using either the target sequence-conditioned PepMLM algorithm (Chen et al., 2023a) (`https://huggingface.co/ChatterjeeLab/PepMLM-650M`) or the motif-specific peptide generator moPPIt algorithm (Chen et al., 2024) (`https://huggingface.co/ChatterjeeLab/moPPIt`).

2.2 LENTIVIRAL PRODUCTION

For target-reporter packaging and pooled peptide packaging, HEK293T cells were seeded in a 6-well plate and transfected at approximately 50% confluency. For each well, 0.5 μg pMD2.G (Addgene #12259), 1.5 μg psPAX2 (Addgene #12260), and 0.5 μg of the target-mCherry reporter transfer vector were transfected with Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocol. The medium was exchanged 8 hours post-transfection, and the viral supernatant was harvested at 48 and and 72 hours post-transfection. The viral supernatant was concentrated to 100x in 1x DPBS using Lenti-X Concentrator (Clontech, 631232) according to the manufacturer's instruction, and stored at -80 °C for further use.

2.3 TARGET-MCHERRY REPORTER MONOCLONAL CELL LINE GENERATION

For the target-reporter cell line generation, 1x10^5 HEK293T cells were mixed with 20 uL of the concentrated virus in a 6-well plate. The medium was changed 24 hours after transduction. Antibiotic selection was started 36 hours post transduction by adding 2 ug/mL puromycin (Sigma, P8833). Cells were harvested for sorting 5 days after antibiotic selection and a single mCherry-positive cell was seeded in a 96 well plate. Genomic PCR was performed after cell growth to validate the genotype of the monoclonal cell line.

2.4 GENERATION OF PLASMIDS

The uAb plasmid was generated from the standard pcDNA3 vector, which harbors a cytomegalovirus (CMV) promoter, a C-terminal P2A-eGFP cassette, and a human phosphoglycerate kinase 1 (hPGK) promoter followed by a blasticidin selection cassette. An Esp3I restriction site was introduced immediately upstream of the CHIPΔTPR CDS along with a flexible GSGSG linker via the KLD Enzyme Mix (NEB) following PCR amplification with mutagenic primers (Genewiz).

2.5 PEPTIDE LIBRARY DESIGN AND CLONING

For each target, the library included a set of polyG and polyA peptides as negative controls. A total of 999 peptides were generated using the PepMLM algorithm (Chen et al., 2023a) and 999 peptides were generated from the moPPIt algorithm (Chen et al., 2024). A total of 2000 peptides

were synthesized from Twist as single-strand DNA and were PCR amplified and cloned into the uAb architecture by using Gibson assembly.

2.6 PEPTIDE POOLED LENTIVIRAL TITRATION

The titer of the lentiviral peptide library pools for the single peptide libraries was determined by transducing 6×10^4 cells with serial dilutions of lentivirus and measuring the percent GFP expression four days after transduction with an Accuri C6 flow cytometer (BD). All lentiviral titrations were performed in the Target-mCherry cell line.

2.7 PEPTIDE POOLED INTEGRATED CELL LINE CULTURE

Target-fused mCherry HEK293T cell lines were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 100 units/mL penicillin, 100 mg/mL streptomycin, and 10% fetal bovine serum (FBS). For lentiviral peptide library integration selection, 8 ug/mL Blasticidin was added 72 hrs post transduction.

2.8 FUNCTIONAL PEPTIDE BINDERS SCREENS

Each target-specific binder screen was performed in duplicate with independent transductions. For each duplicate, 100 ng/uL doxycycline was added to express the uAb. Cells were harvested 5 days after the doxycycline supplement. for sorting were sorted post 5 days doxycycline supplement. $1x10^7$ target-mCherry cells were washed once with 1x PBS, dissociated using Accutase, filtered through a 30 mm CellTrics filter and resuspended in FACS buffer (0.5)% BSA (Sigma, A7906), 2 mM EDTA (Sigma, E7889) in PBS). The highest and lowest 10% of cells were sorted based on mCherry expression. Sorting was done with a SH800 FACS Cell Sorter (Sony Biotechnology). After sorting, genomic DNA was harvested with the Spin gDNA Extraction Kit (NEB T3010).

2.9 PEPTIDE LIBRARY SEQUENCING

The peptide libraries were amplified from each genomic DNA sample in 100 μL of PCR reactions using Q5 Hot Start Polymerase (NEB, M0493) with 1 μg of genomic DNA per reaction. PCR amplification was performed according to the manufacturer's instructions, using 25 cycles at an annealing temperature of 60C with the following primers:

Fwd: ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTGGCTAGCGGATCCGCCAC-CATG Rev: GACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGTTCAGCCGGCCAGAAC-CGCTGCC

The amplified libraries were purified using the QIAquick Gel extraction kit and the amplicons were sent to Genewiz for next-generation sequencing.

2.10 DATA ANALYSIS

The `MAGeCK` and `MAGeCK-VISPR` packages were used to conduct a systematic analysis of functional peptides across two experimental conditions Li et al. (2015). `MAGeCK MLE` leverages maximum-likelihood estimation (MLE) for robust identification of CRISPR-screen hits. This computational pipeline evaluates the biological significance of each gene by calculating beta scores (equivalent to $\log_2 FC$) and their associated statistical metrics. A positive beta score indicates a positive selection, and a negative beta score indicates a negative selection.

To visualize these results, a volcano plot was generated to display both the magnitude of effect (beta score) and statistical significance ($-\log_{10} p$-value) for each peptide. The grey/red dots indicate the negative/positive selection genes with the beta score threshold of 0.5 and p-value threshold < 0.05 to select for enriched peptides.

2.11 CELL FRACTIONATION AND IMMUNOBLOTTING

On the day of harvest, cells were detached by addition of 0.05% trypsin-EDTA and cell pellets were washed twice with ice-cold $1\times$ PBS. Cells were then lysed and subcellular fractions were isolated

from lysates using a 1:100 dilution of protease inhibitor cocktail (Millipore Sigma, Cat # P8340) in Pierce RIPA buffer (ThermoFisher, Cat # 89900). Specifically, the protease inhibitor cocktail-RIPA buffer solution was added to the cell pellet, the mixture was placed at 4 °C for 30 min followed by centrifugation at 15,000 rpm for 10 min at 4 °C. The supernatant was collected immediately to pre-chilled PCR tubes and quantified using the Pierce BCA Protein Assay Kit (ThermoFisher, Cat # 23227). Twenty micrograms of lysed protein was mixed with 4× Bolt™ LDS Sample Buffer (ThermoFisher, Cat # NP0007) with 5% β-mercaptoethanol (Millipore Sigma, Cat # M3148) in a 3:1 ratio and subsequently incubated at 95 °C for 10 min prior to immunoblotting, which was performed according to standard protocols. Briefly, samples were loaded at equal volumes into Bolt™ Bis-Tris Plus Mini Protein Gels (ThermoFisher, Cat # NW04125BOX) and separated by electrophoresis. iBlot™ 2 Transfer Stacks (Invitrogen) were used for membrane blot transfer, and following a 1 h room-temperature incubation in 5% milk-TBST, proteins were probed with mouse anti-beta-Catenin antibody (R&D Systems, Cat # MAB13292; diluted 1:1000), and mouse anti-GAPDH (Santa Cruz Biotechnology, Cat # sc-47724; diluted 1:10,000) for overnight incubation at 4 °C. The blots were washed three times with 1× TBST for 5 min each and then probed with a secondary antibody, goat anti-mouse IgG (H + L) Poly-HRP (ThermoFisher, Cat # 32230, diluted 1:5000) for 1 h at room temperature. Following three washes with 1× TBST for 5 min each, blots were detected by chemiluminescence using a BioRad ChemiDoc™ Touch Imaging System (Bio-Rad). Densitometry analysis of protein bands in immunoblots was performed using ImageJ software as described here: `https://imagej.nih.gov/ij/docs/examples/dot-blot/`. Briefly, bands in each lane were grouped as a row or a horizontal "lane" and quantified using FIJI's gel analysis function. Intensity data for the duAb bands was first normalized to band intensity of either GAPDH in each lane then to the average band intensity for polyG uAb vector control cases across duplicates.

3 RESULTS

3.1 GENERATION OF A DUAL REPORTER CELL LINE FOR SCREENING OF FUNCTIONAL BINDER PEPTIDES

To generate a novel mammalian cell line, we designed a dual-reporter system (Figure 1A). Briefly, we cloned the coding sequences of targets of interest into vectors fused with mCherry and co-expressed them with puromycin resistance cassettes for antibiotic selection. After establishing a monoclonal cell line to ensure consistent target protein expression levels (Figure 1B), we inserted a pooled oligonucleotide library into the uAb vector, under the control of a Tet-on promoter and fused with a GFP reporter to visualize uAb construct expression. The uAb construct also includes a Blasticidin resistance cassette as a secondary selection marker to ensure successful integration of both constructs during the second round of selection. To induce uAb expression, we added doxycycline, which activated GFP fluorescence (Figure 1C). When functional peptide binders engaged their target, uAb-mediated degradation led to a reduction in red fluorescence, corresponding to decreased mCherry levels (Figure 1D).

3.2 HIGH-THROUGHPUT SCREENING IDENTIFIES TARGET-ENGAGING AND DEGRADING PEPTIDES

To develop a preliminary high-throughput screen with our platform, we designed 2,000 peptides targeting two distinct proteins: β-catenin, a core transcriptional regulator whose dysregulation often leads to cancer proliferation (Yu et al., 2021; Zhan et al., 2017), and EWS::FLI1, an oncogenic fusion protein that is pathogenic for Ewing sarcoma (May et al., 1993; Braun et al., 1995; Vincoff et al., 2025). We first synthesized a Twist oligo pool encoding these peptides and cloned them into a transfer vector for lentiviral packaging. The resulting lentiviral library was transduced into monoclonal target-mCherry cell lines.

Peptide binders that successfully engaged their target protein led to a reduction in mCherry fluorescence. The lowest 10% of cells, based on mCherry fluorescence intensity, were isolated using Fluorescence-Activated Cell Sorting (FACS), followed by genomic DNA extraction and amplification of peptide-coding sequences for Next-Generation Sequencing (NGS) (Figure 2).

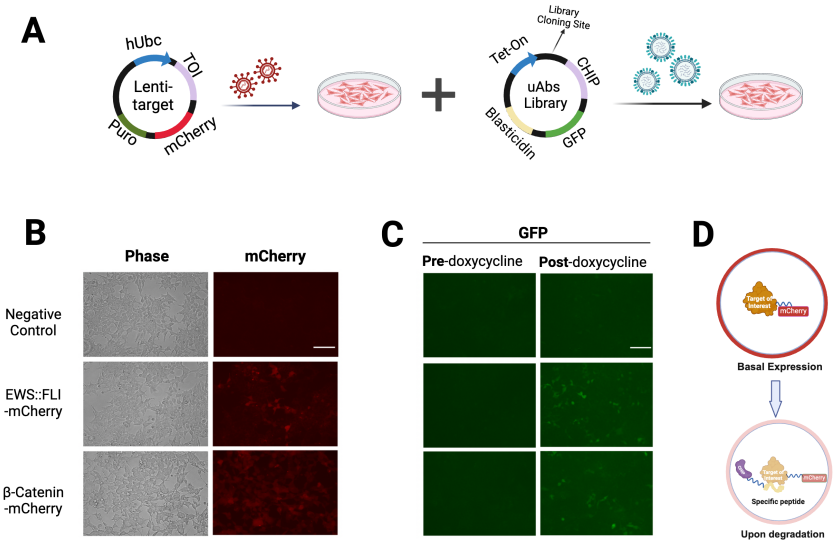

Figure 1: Reporter cell line generation for screening of functional binder peptide. A) Reporter cell line integration for both the target-mCherry and uAb-GFP constructs. B) Target expression validation by mCherry fluorescence. C)uAb expression validation by GFP fluorescence following doxycycline supplementation. D) Schemetic of functional peptide binders leading to uAb-mediated degradation in the target-mCherry reporter cell line

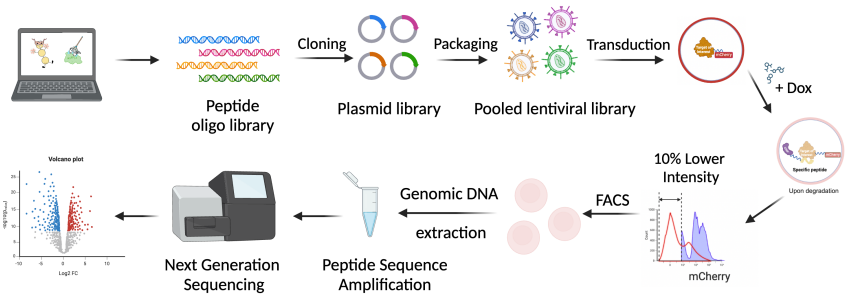

Figure 2: Schematic of high-throughput screening of language model-designed peptides to identify target-specific degraders.

To quantify enrichment, we calculated the beta score for each peptide by comparing post-sorted and pre-sorted populations. Peptides were ranked based on the differential beta score, determined by subtracting the control beta score from the treatment beta score. Red dots represent peptides with an increased beta score after treatment, while gray dots indicate peptides with a decreased beta score (Figure 3A,B). We first tested top 15 peptides-guided uAbs in β-catenin-mCherry stable monoclonal cell line. Our results demonstrated that all peptides induce statistically significant degradation of β-catenin-mCherry (Figure 3C). We further tested those 15 peptides for endogenous degradation validation, and the immunoblot analysis with an anti- β-catenin antibody revealed that a portion of the peptide-guided uAbs induced statistically significant degradation of endogenous β-catenin. We are currently cloning the top 15 candidate peptides for EWS::FLI1 target into individual uAb constructs (Table 1) for degradation validation in both mCherry reporter cell line as well as A673 Ewing sarcoma cell lines.

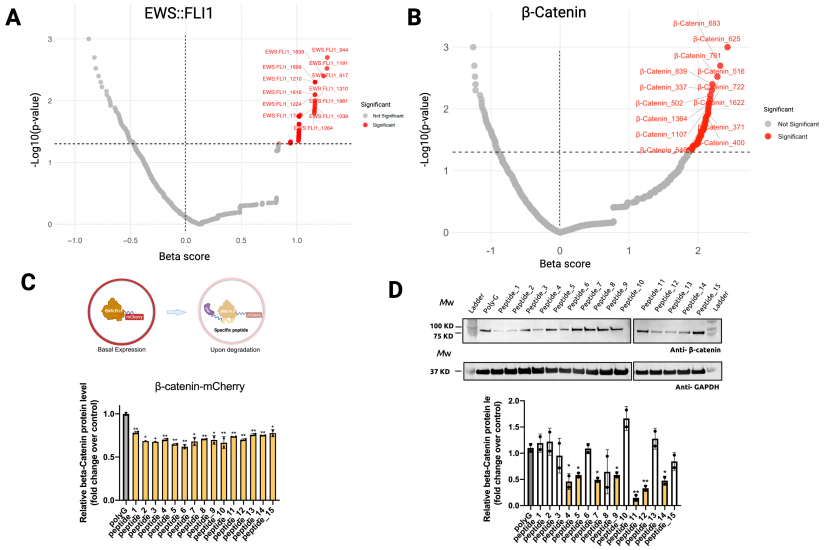

Figure 3: A) Scatter plot of high-throughput screening in EWS::FLI1-mCherry reporter cell line, B) Scatter plot of high-throughput screening in β-catenin-mCherry reporter cell line, both A and B scatter plots aiming to systematically identify functional peptide binders. Red dots represent peptides whose beta score increased after treatment. Gray dots represent genes whose beta score decreased after treatment.C) Degradation in EWS::FLI1-mCherry reporter cell line with top 15 candidate peptides. D) Degradation of endogenouse of β-catenin in proteiin extracts of DLD1 cell analyzed by immunoblotting. DLD1 cells were transfected with a pcDNA plasmid encoding top 14 cnadidate uAbs while transfection with polyG peptide-guided uAb served as a control.

4 Conclusion

The CRISPR-Cas system has revolutionized genome editing due to its modular and programmable nature. The ability to design an sgRNA for a specific genomic sequence enables precise Cas-mediated editing (Pacesa et al., 2024; Wang & Doudna, 2023). Beyond gene disruption, CRISPR has been adapted for single-base editing, gene insertion, and transcriptional modulation (Pacesa et al., 2024; Wang & Doudna, 2023). Inspired by this RNA-guided system, our peptide-guided uAb degraders are genetically encodable, allowing us to integrate them into a CRISPR-like high-throughput screening (HTS) platform to identify functional peptide binders.

In this work, we have established a mammalian high-throughput assay to screen peptide binders for targets of interest. We are actively screening binders generated from diverse target-conditioned AI models and can rapidly extend this platform to additional targets, including highly disordered proteins Bhat et al. (2025); Brixi et al. (2023); Chen et al. (2023a). We anticipate that our HTS approach will accelerate the development of therapeutic modalities dependent on selective target binders and facilitate the design of novel therapies across a range of disease indications.

References

Suhaas Bhat, Kalyan Palepu, Lauren Hong, Joey Mao, Tianzheng Ye, Rema Iyer, Lin Zhao, Tianlai Chen, Sophia Vincoff, Rio Watson, et al. De novo design of peptide binders to conformationally diverse targets with contrastive language modeling. *Science Advances*, 11(4):eadr8638, 2025.

Christoph Bock, Paul Datlinger, Florence Chardon, Matthew A. Coelho, Matthew B. Dong, Keith A. Lawson, Tian Lu, Laetitia Maroc, Thomas M. Norman, Bicna Song, Geoff Stanley, Sidi Chen, Mathew Garnett, Wei Li, Jason Moffat, Lei S. Qi, Rebecca S. Shapiro, Jay Shendure, Jonathan S. Weissman, and Xiaowei Zhuang. High-content crispr screening. *Nature Reviews Methods*

Primers, 2(1), February 2022. ISSN 2662-8449. doi: 10.1038/s43586-021-00093-4. URL http://dx.doi.org/10.1038/s43586-021-00093-4.

Benjamin S Braun, Richard Frieden, Stephen L Lessnick, William A May, and Christopher T Denny. Identification of target genes for the ewing's sarcoma ews/fli fusion protein by representational difference analysis. *Molecular and Cellular Biology*, 1995.

Garyk Brixi, Tianzheng Ye, Lauren Hong, Tian Wang, Connor Monticello, Natalia Lopez-Barbosa, Sophia Vincoff, Vivian Yudistyra, Lin Zhao, Elena Haarer, et al. Salt&peppr is an interface-predicting language model for designing peptide-guided protein degraders. *Communications Biology*, 6(1):1081, 2023.

Pranam Chatterjee, Manvitha Ponnapati, Christian Kramme, Alexandru M Plesa, George M Church, and Joseph M Jacobson. Targeted intracellular degradation of sars-cov-2 via computationally optimized peptide fusions. *Communications Biology*, 3(1):715, 2020.

Tianlai Chen, Madeleine Dumas, Rio Watson, Sophia Vincoff, Christina Peng, Lin Zhao, Lauren Hong, Sarah Pertsemlidis, Mayumi Shaepers-Cheu, Tian Zi Wang, Divya Srijay, Connor Monticello, Pranay Vure, Rishab Pulugurta, Kseniia Kholina, Shrey Goel, Matthew P. DeLisa, Ray Truant, Hector C. Aguilar, and Pranam Chatterjee. Pepmlm: Target sequence-conditioned generation of therapeutic peptide binders via span masked language modeling. *arXiv*, 2023a. doi: 10.48550/ARXIV.2310.03842. URL https://arxiv.org/abs/2310.03842.

Tianlai Chen, Lauren Hong, Vivian Yudistyra, Sophia Vincoff, and Pranam Chatterjee. Generative design of therapeutics that bind and modulate protein states. *Current Opinion in Biomedical Engineering*, pp. 100496, 2023b.

Tong Chen, Yinuo Zhang, and Pranam Chatterjee. moppit: De novo generation of motif-specific binders with protein language models. *bioRxiv*, 2024.

Luke A Gilbert, Max A Horlbeck, Britt Adamson, Jacqueline E Villalta, Yuwen Chen, Evan H Whitehead, Carla Guimaraes, Barbara Panning, Hidde L Ploegh, Michael C Bassik, et al. Genome-scale crispr-mediated control of gene repression and activation. *Cell*, 159(3):647–661, 2014.

Silvana Konermann, Mark D Brigham, Alexandro E Trevino, Julia Joung, Omar O Abudayyeh, Clea Barcena, Patrick D Hsu, Naomi Habib, Jonathan S Gootenberg, Hiroshi Nishimasu, et al. Genome-scale transcriptional activation by an engineered crispr-cas9 complex. *Nature*, 517 (7536):583–588, 2015.

Wei Li, Johannes Köster, Han Xu, Chen-Hao Chen, Tengfei Xiao, Jun S. Liu, Myles Brown, and X. Shirley Liu. Quality control, modeling, and visualization of crispr screens with mageck-vispr. *Genome Biology*, 16(1), December 2015. ISSN 1474-760X. doi: 10.1186/s13059-015-0843-6. URL http://dx.doi.org/10.1186/s13059-015-0843-6.

William A. May, Stephen L. Lessnick, Benjamin S. Braun, Micheal Klemsz, Brain C. Lewis, Lynn B. Lunsford, Robert Hromas, and Christopher T. Denny. The ewing's sarcoma ews/fli-1 fusion gene encodes a more potent transcriptional activator and is a more powerful transforming gene than fli-1. *Molecular and Cellular Biology*, 13(12):7393–7398, December 1993. ISSN 1098-5549. doi: 10.1128/mcb.13.12.7393-7398.1993. URL http://dx.doi.org/10.1128/mcb.13.12.7393-7398.1993.

Martin Pacesa, Oana Pelea, and Martin Jinek. Past, present, and future of crispr genome editing technologies. *Cell*, 187(5):1076–1100, February 2024. ISSN 0092-8674. doi: 10.1016/j.cell.2024.01.042. URL http://dx.doi.org/10.1016/j.cell.2024.01.042.

Ophir Shalem, Neville E Sanjana, and Feng Zhang. High-throughput functional genomics using crispr–cas9. *Nature Reviews Genetics*, 16(5):299–311, 2015.

Sophia Vincoff, Shrey Goel, Kseniia Kholina, Rishab Pulugurta, Pranay Vure, and Pranam Chatterjee. Fuson-plm: a fusion oncoprotein-specific language model via adjusted rate masking. *Nature Communications*, 16(1), February 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-56745-6. URL http://dx.doi.org/10.1038/s41467-025-56745-6.

Joy Y Wang and Jennifer A Doudna. Crispr technology: A decade of genome editing is only the beginning. *Science*, 379(6629):eadd8643, 2023.

Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

Xin Yang, Shoufu Duan, Zhiming Li, Zhe Wang, Ning Kon, Zhiguo Zhang, and Wei Gu. Protocol of crispr-cas9 knockout screens for identifying ferroptosis regulators. *STAR protocols*, 4(4):102762, 2023.

Fanyuan Yu, Changhao Yu, Feifei Li, Yanqin Zuo, Yitian Wang, Lin Yao, Chenzhou Wu, Chenglin Wang, and Ling Ye. Wnt/-catenin signaling in cancers and targeted therapies. *Signal Transduction and Targeted Therapy*, 6(1), August 2021. ISSN 2059-3635. doi: 10.1038/s41392-021-00701-5. URL http://dx.doi.org/10.1038/s41392-021-00701-5.

Tailan Zhan, Niklas Rindtorff, and Michael Boutros. Wnt signaling in cancer. *Oncogene*, 36(11): 1461–1473, 2017.

A SUPPLEMENTARY INFORMATION

Name	**Amino Acid Sequence**	**Beta Score**
EWS::FLI1_1	NPQPQIRPSDYIRDNLAQQS	2.66047
EWS::FLI1_2	NVPPQQRLSESYTALQQQQQ	2.53356
EWS::FLI1_3	NVQKQISLSDYYRALAQQLQ	2.42819
EWS::FLI1_4	DPLPQKSASEEYRDLLDAQY	2.33758
EWS::FLI1_5	CPPWPKPKKEYYDEHWDNDV	2.22516
EWS::FLI1_6	CKPPRKEKDSYYDWDWDNKV	2.14294
EWS::FLI1_7	PVQVLQRALDYYSANAQQQS	2.13213
EWS::FLI1_8	WPPKPKEKDEYYDKHRDNHW	2.11680
EWS::FLI1_9	NPQPQQSASDYGSDLQDAAS	2.11197
EWS::FLI1_10	DPPKQIAASDSYRALLAQAS	2.10348
β-catenin_1	FQYQEQLHAQQNQLRELRAQ	4.75735
β-catenin_2	DATSERVTAYQERIAELQAL	4.71257
β-catenin_3	DKERERMAAELLLDLVALLK	4.52595
β-catenin_4	SEEQERMSESDLDDDAVLVK	4.50363
β-catenin_5	DEVQREMSSELQDNDVVVLL	4.45841
β-catenin_6	SEVRRELSSSLDLDDLALAL	4.44399
β-catenin_7	DKEQRRMSESQQDDLVLVLK	4.44222
β-catenin_8	EPVVERMLEELQDDAVLVAK	4.39449
β-catenin_9	DEERVEMSSSQQDDLAALVL	4.38488
β-catenin_10	FTTREQLCSVEEELNALQAK	4.27063

Table 1: Enriched peptide sequences and their associated beta scores for EWS::FLI1 and β-catenin.