SOAPIA: Siamese-Guided Generation of Off Target-Avoiding Protein Interactions with High Target Affinity

Sophia Vincoff[*]
University of Pennsylvania

Oscar Davis[*]
University of Oxford

Ismail Ilkan Ceylan
University of Oxford

Alexander Tong
Université de Montréal
Mila – Quebec AI Institute

Avishek Joey Bose
University of Oxford
Mila – Quebec AI Institute

Pranam Chatterjee[†]
University of Pennsylvania

Abstract

Therapeutic molecules must selectively interact with a target protein while avoiding structurally or functionally similar off-targets. However, no existing generative strategy explicitly optimizes both target affinity and off-target avoidance. To address this, we introduce **SOAPIA**, a framework for the **S**iamese-guided generation of **O**ff-target-**A**voiding **P**rotein **I**nteractions with high target **A**ffinity. SOAPIA generates *de novo* peptide binders by steering the generative process of a Diffusion Protein Language Model (DPLM) using a multi-objective Monte Carlo Tree Search (MCTS). Affinity is optimized via a pre-trained predictor, while specificity is enforced using a Siamese model trained with an adaptive Log-Sum-Exp Decoy Loss. This dual-guidance scheme enables Pareto-efficient exploration of discrete sequence space without gradient access. In benchmarks across 17 fusion oncoproteins, SOAPIA consistently produces binders with strong affinity and high selectivity for over 75% of targets. For multiple clinically relevant targets, SOAPIA generated peptides that preferentially bind the fusion by engaging both its head and tail domains, while avoiding the wild-type counterparts. These results underscore SOAPIA's promise for designing safe, specific biologics for fusion-driven cancers and other rare, currently untreatable diseases.

1 Introduction

Selective modulation of pathogenic proteins is essential for drug design [Nada et al., 2024]. Off-target interactions can reduce efficacy or lead to toxicity, a challenge shared across small molecules, PROTACs, and biologics [Garon et al., 2017, Chen et al., 2023a]. Since large-scale *in vitro* screening is costly and impractical, computational methods for designing drug-target interactions (DTIs) and protein-protein interactions (PPIs) are increasingly important. Structure-based approaches offer atomistic resolution but fail on disordered and chimeric proteins and are too slow for high-throughput design [Chen et al., 2023a]. Sequence-based models for DTIs [Singh et al., 2023, McNutt et al., 2024, Gao et al., 2024], PPIs [Sledzieski et al., 2021, Singh et al., 2022], and peptide design [Bhat et al., 2025, Tang et al., 2025] address this limitation by operating directly on primary sequences.

Yet, most generative approaches optimize only for target binding, without explicitly avoiding off-target interactions. This is particularly problematic for fusion oncoproteins, which drive many pediatric

[*]Equal contribution
[†]Corresponding author: `pranam@seas.upenn.edu`

First Exploration in AI Today Workshop at ICML (EXAIT at ICML 2025).

cancers and result from chromosomal translocations, often retaining high sequence identity with their wild-type head and tail domains [Vincoff et al., 2025]. Designing binders for such targets requires a multi-objective formulation that simultaneously maximizes affinity and enforces specificity—not just against a generic background proteome, but against multiple known off-targets. Including two explicit decoys during training and generation better reflects real-world therapeutic constraints, where safe and selective binding is essential [Chen et al., 2023a].

Discrete diffusion models have become a powerful class of generative frameworks for sequence design, enabling high-quality, controllable generation at the token level without requiring 3D structures or continuous embeddings [Campbell et al., 2024, Shi et al., 2024, Sahoo et al., 2024]. These models have recently shown strong performance in protein design tasks; as examples, DPLM [Wang et al., 2024] and EvoDiff [Alamdari et al., 2023] support structure-free generation of valid, foldable protein sequences. Most recently, PepTune [Tang et al., 2025] extended this paradigm to multi-objective optimization by introducing a Monte Carlo Tree Search (MCTS) framework that guides discrete diffusion using multiple non-differentiable reward functions. Operating in therapeutic peptide SMILES space, PepTune demonstrates that MCTS can efficiently explore the denoising landscape to discover Pareto-optimal solutions, even in the absence of gradient signals [Tang et al., 2025].

We build on these insights with **SOAPIA**: a framework for the **S**iamese-guided generation of **O**ff-target-**A**voiding **P**rotein **I**nteractions with high **A**ffinity. SOAPIA combines a contrastive Siamese protein language model—trained with an adaptive Log-Sum-Exp Decoy Loss to separate binders from multiple off-targets—with a pre-trained affinity predictor. These soft-value signals define a dual-objective reward function that guides a Pareto-aware MCTS over the denoising trajectory of DPLM [Wang et al., 2024], enabling efficient sampling of short protein sequences that satisfy both constraints. We show that SOAPIA outperforms Best-of-N sampling baselines on both affinity and specificity and generates peptide-like binders that preferentially dock to fusion proteins while avoiding their head and tail domains. *In silico* docking with AlphaFold-Multimer [Evans et al., 2021] confirms SOAPIA's ability to design safe and selective binders for undruggable and isoform-sensitive targets, such as fusion oncoproteins.

2 Methods

Data curation and handling All data curation, splitting, and clustering details can be found in the Supplementary Methods.

Protein encoding We encode four protein sequences—a binder, target, and two off-targets—into a shared latent space using the 33-layer ESM-2-650M model. The final two layers are unfrozen during training. A positional multi-head attention module (*n_heads* = 10) with rotary positional embeddings (RoPE) [Su et al., 2024] captures sequence-order information. Outputs are passed through two SiLU-activated linear layers with skip connections, and attention pooling produces fixed-length embeddings (Figure A1).

Specificity loss To enforce off-target avoidance, we train a Siamese model using an Adaptive Log-Sum-Exp Decoy Loss. Let $\mathbf{b}$, $\mathbf{t}$, $\mathbf{ot}_1$, and $\mathbf{ot}_2$ represent the embeddings of the binder, target, and two off-targets. The loss is:

$$\mathcal{L} = \sum_{\mathbf{b}} \max\left(0,\, \text{dist}(\mathbf{b},\mathbf{t}) - \frac{1}{\beta}\log\left(e^{\beta\cdot\text{dist}(\mathbf{b},\mathbf{ot}_1)} + e^{\beta\cdot\text{dist}(\mathbf{b},\mathbf{ot}_2)} + \epsilon\right) + \alpha\right) \tag{1}$$

where β controls sharpness, α is a margin, and $\epsilon = 10^{-8}$ ensures numerical stability.

Implementation details All models were implemented using PyTorch Lightning [Falcon, 2019] and trained on 4xA100 NVIDIA GPUs with an effective batch size of 128. Learning rate was initialized at 1×10^{-4} and decayed using cosine annealing with 200 warmup steps. Training was stopped after loss plateaued for three epochs (10 total). See Table A2 for full hyperparameters.

Embedding separation visualization After each epoch, we computed Euclidean distances between binder-target, binder-off-target 1, and binder-off-target 2 embeddings across the training set, and visualized them using `matplotlib v3.8.2` (Figure A2).

Binder recovery screen To test specificity, 50,000 random amino acid sequences (lengths 56–856) were generated per target and scored using the Siamese model. The true binder was ranked by:

$$D = \text{dist}(\mathbf{b}, \mathbf{t}) - \frac{1}{\beta} \log \left(e^{\beta \cdot \text{dist}(\mathbf{b}, \mathbf{ot}_1)} + e^{\beta \cdot \text{dist}(\mathbf{b}, \mathbf{ot}_2)} + \epsilon \right) + \alpha \tag{2}$$

Lower D implies better specificity. Rankings were evaluated across multiple thresholds (Figure A3).

Masked discrete diffusion Binder generation is based on a masked discrete diffusion model (MDM) [Sahoo et al., 2024, Shi et al., 2024, Campbell et al., 2024]. The forward process corrupts a clean sequence $\mathbf{x}_0$ with:

$$p_t(\mathbf{x}_t|\mathbf{x}_0) = \prod_{i=1}^{n} \text{Cat}(x_t^i; \alpha_t \delta(x_0^i) + (1 - \alpha_t)\delta(m)), \tag{3}$$

and the reverse process denoises via:

$$p_t(x_{t-1}^i | x_t^i, x_0^i) \propto \alpha_{t-1} \delta(x_0^i) + (1 - \alpha_{t-1})\delta(m). \tag{4}$$

Training minimizes weighted cross-entropy:

$$\mathcal{L} = -\mathbb{E}_{\mathbf{x}_t \sim p_t} \sum_{i=1}^{n} x_0^i \log \mu_\theta(x_t^i, t). \tag{5}$$

Sampling and generation strategies (SOAPIA) SOAPIA generates peptide binders by applying MCTS over the denoising trajectory of DPLM [Wang et al., 2024], guided by two soft-value reward functions which are maximized for the optimal binder: one reward from the Siamese model trained for off-target avoidance as described here ($r(x) = -D$) (2), and one from a pre-trained peptide-protein affinity predictor Chen et al. [2025]. At each expansion step, candidate sequences are sampled from the DPLM transition distribution $p_\theta(x_{t-1} \mid x_t, t)$ and completed via ancestral decoding. Each fully unmasked sequence is scored along both objectives and compared to a dynamically updated Pareto frontier. Soft reward vectors—computed from overlap with frontier members—encourage exploration of diverse trade-offs, while heavily dominated sequences are pruned.

We evaluate multiple guidance mechanisms to direct sequence refinement, including SVDD [Li et al., 2024], which selects the best of m candidates at each step; simple guidance [Schiff et al., 2025], which applies reward-informed local updates; and NOS [Gruver et al., 2023], which performs iterative local search through the reward landscape. These samplers and guidance methods are combined with higher-level generation strategies. The *Basic* strategy generates N samples without filtering; *Best-of-N* selects the top N from a larger pool; *Scalarized Best-of-N* applies a weighted reward $R(x) = \lambda r_1(x) + (1 - \lambda) r_2(x)$; and *Pareto Best-of-N* returns non-dominated sequences. The *MCTS* strategies used by SOAPIA integrate sampling and guidance at each expansion and rollout step, updating the Pareto frontier online as sequences are generated. *MCTS* uses specificity scores only, while *MultiMCTS* performs multi-objective guidance with both specificty and affinity scores. See Algorithm 1 for details.

3 Results

Siamese model results First, we trained our Siamese model (Figure A1) to obtain the specificity predictor used for SOAPIA's dual-guided sampling. The model was trained to embed binders closer to their target than to either of two off-targets. Over the course of training, we observed progressive separation in the embedding space: binder–target distances decreased while binder–off-target distances increased, as shown in Figure A2. To assess its ranking performance, we evaluated the model against 50,000 randomly generated binders per test case. The true binder was ranked in the top-1 for 19.2% of cases and in the top-25% for 56.8% (Figure A3). Finally, protein role distributions across train, validation, and test splits are shown in Figure A4, confirming dataset diversity and generalization.

Generation of specific, high-affinity peptides SOAPIA successfully generates novel, short binding proteins with a variety of sampling strategies (Figure 1). When different methodologies were evaluated on 20 examples from the Siamese specificity model's test set, *MultiMCTS* achieved the optimal results. Binders "passed" the specificity task if their reward scores were positive, indicating that the binder is embedded closer to the target than the two off-targets in the model's latent space. For the affinity task, a passing threshold of 6.5 was selected, as indicated in MOG-DFM [Chen et al., 2025]. All three guidance methods - NOS, simple, and SVDD - performed well in the expansion and rollout steps of *MultiMCTS*. Unguided generation was also competitive. NOS achieved the best results, with a dual-objective pass rate of 48.12%, and only 5.31% of generated samples failing to meet the criteria for either objective (Figure 1).

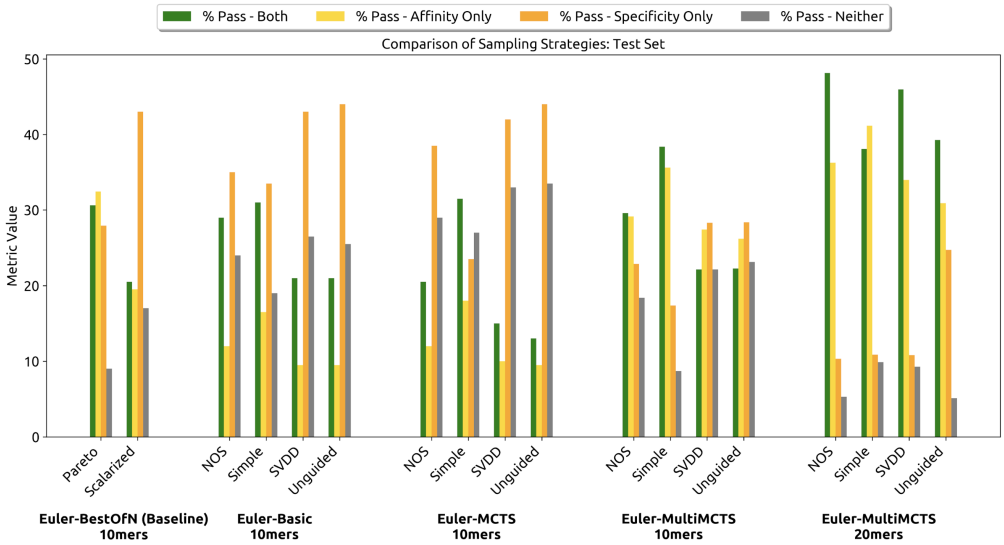

Figure 1: **Performance of different sampling strategies on the Siamese model test set**. The passing thresholds for specificity and affinity rewards are 0 and 6.5, respectively. For each input (target, off-target 1, and off-target 2), 10 novel binding peptides of length 10 or 20 were generated and scored.

Targeting fusion oncoproteins Next, we sought to determine whether SOAPIA can be applied to target proteins which have no known specific binders. We evaluated *MultiMCTS* on 17 fusion oncoproteins, whose off-targets are their corresponding wild-type heads and tails (Figure 2). For all guidance strategies tested, the maximum specificity and affinity values were greater than 12 and 8 respectively, indicating very strong selectivity and affinity for the best samples. Mean affinity scores were greater than 7 for all methods, indicating that on average, SOAPI-designed fusion oncoprotein binders pass the affinity criterion. The four strategies were slightly more differentiated in average specificity, with simple guidance achieving the highest value of 2.84 (Figure 2A). Accordingly, simple guidance also produced the highest simultaneous dual-objective hit rate of 68.27% (Figure 2B, Table A4).

Prediction of target-binder complexes To validate that SOAPIA's rewards are meaningful, we co-folded all peptides in the "green zone" (passing both criteria) using AlphaFold-Multimer. All four guidance strategies produced dozens of hits, where target ipTM exceeded both off-target ipTMs. Unguided and NOS produced the highest and lowest hit rates, respectively (36.1%, 27.3%) (Table A5). For 46 (8.1%) green zone binders across all four strategies (Table A6), target ipTMs were above 0.7 in addition to exceeding off-target ipTMs, indicating both high specificity and high affinity. We visualize four top-performing complexes for clinically relevant targets, each associated with a specific cancer: APTX::ARL5B with lung squamous cell carcinoma, ETHE1::PLAUR with esophageal carcinoma, DHRSX::RPS4Y1 with pancreatic adenocarcinoma, and AMACR::NDUFAF2 with prostate adenocarcinoma Vincoff et al. [2025] (Figure 3, Table A3). The predicted structures of the binder-target complexes imply selective binding modes, where the peptide engages both the head and tail portions of the fusion oncoprotein. In total, these findings demonstrate SOAPIA's potential to

design peptides that selectively recognize fusion oncoproteins while minimizing interactions with their wild-type counterparts.

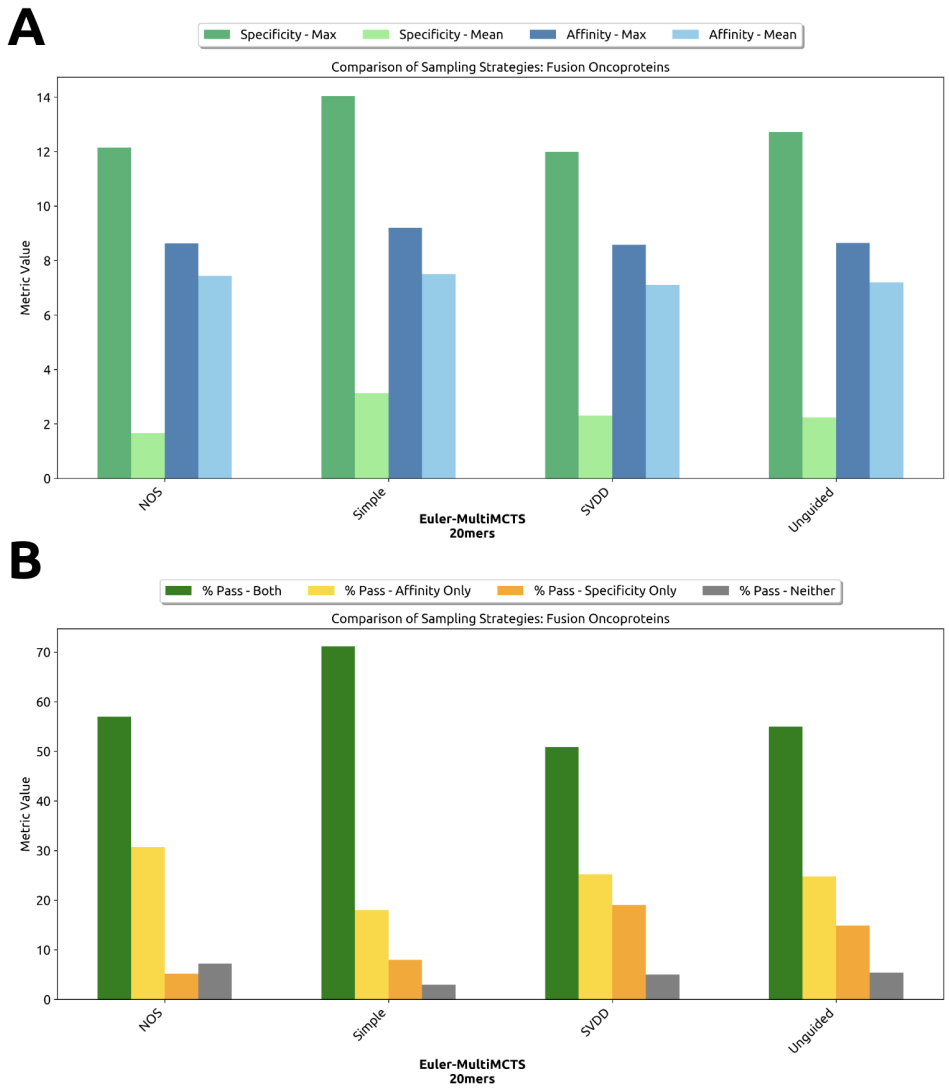

Figure 2: **Performance of different guidance methods with *MultiMCTS* on a set of fusion oncoproteins**. **A** Maximum and mean rewards across 10 samples per target, for 18 targets. **B** Pass rates for each objective across all generated samples.

4 Discussion

SOAPIA is a new framework for multi-objective binder design that jointly optimizes specificity and affinity using a Pareto-guided MCTS over a DPLM prior. Inspired by PepTune [Tang et al., 2025], SOAPIA combines soft-value signals from a Siamese contrastive model and a trained affinity predictor Chen et al. [2025] to guide tree-based exploration during denoising, without requiring gradient access to either objective. This dual-guidance mechanism allows SOAPIA to perform competitively against Best-of-N sampling baselines across both objectives, generating peptide sequences that exhibit strong binding while avoiding homologous off-targets. Notably, SOAPIA is capable of generating binders to fusion oncoproteins that show preferential docking to the full fusion but not to the individual head or tail proteins. This level of selectivity is critical for therapeutic applications, particularly in pediatric fusion-driven cancers, where targeting the aberrant fusion protein while sparing wild-type counterparts is essential for minimizing toxicity Vincoff et al. [2025].

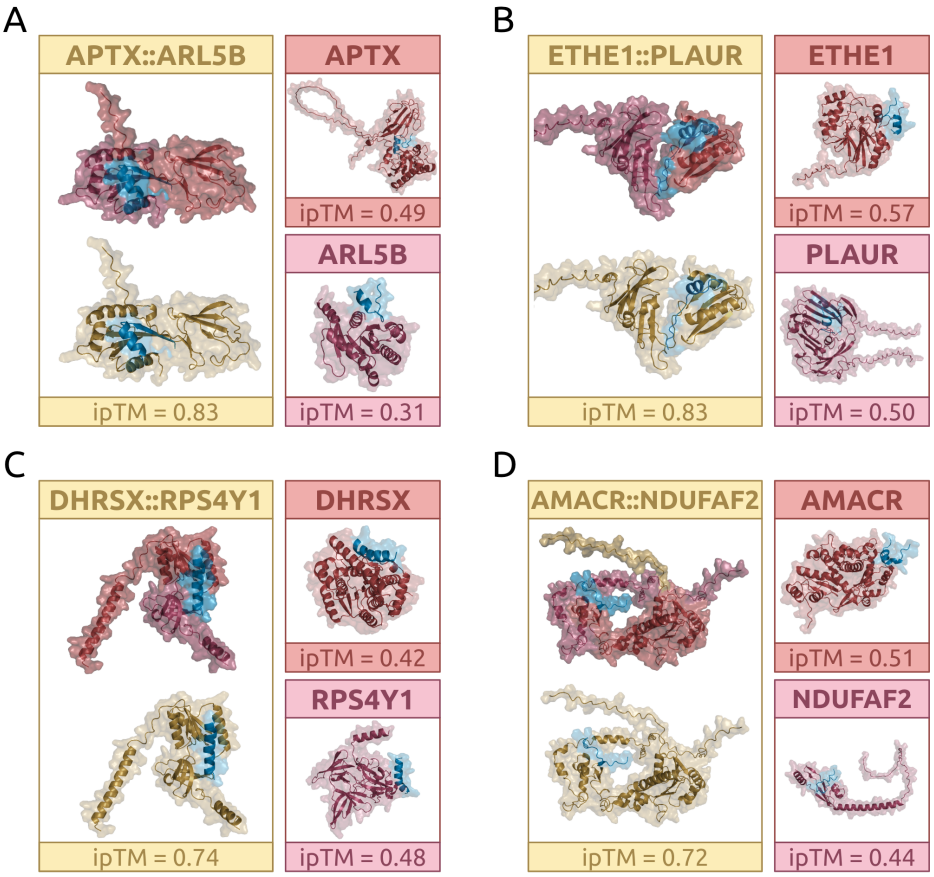

Figure 3: **AlphaFold-Multimer-predicted structures of top-performing binders.** (**A**-**D**) For each fusion oncoprotein, the binder is displayed in complex with the target, off-target 1 (wild-type head protein), and off-target 2 (wild-type tail protein)

Looking ahead, we will refine SOAPIA's guidance weights and rollout policies to further improve sample efficiency, and conduct experimental validation in cellular systems focused on disordered fusion oncoproteins. For this, we plan to integrate FusOn-pLM embeddings [Vincoff et al., 2025] to better capture breakpoint-localized context, and incorporate PTM-Mamba embeddings [Peng et al., 2025] to enable peptide design sensitive to post-translational modification states. In addition to MCTS, recent work on multi-objective guided flow matching (MOG-DFM [Chen et al., 2025]) offers a potential alternative decoding framework for joint optimization in discrete sequence space. Finally, when paired with experimental platforms such as ubiquibodies (uAbs) and deubiquibodies (duAbs) for targeted protein degradation [Brixi et al., 2023, Bhat et al., 2025, Chen et al., 2023b] and stabilization [Hong et al., 2025], SOAPIA provides a generalizable, programmable approach to modulating previously undruggable proteins, with potential impact across oncology, rare disease, and immunotherapy.

References

Hossam Nada, Yongseok Choi, Sungdo Kim, Kwon Su Jeong, Nicholas A. Meanwell, and Kyeong Lee. New insights into protein–protein interaction modulators in drug discovery and therapeutic advance. *Signal Transduction and Targeted Therapy*, 9(1), December 2024. ISSN 2059-3635. doi: 10.1038/s41392-024-02036-3. URL http://dx.doi.org/10.1038/s41392-024-02036-3.

Sarah L. Garon, Rebecca K. Pavlos, Katie D. White, Nancy J. Brown, Cosby A. Stone, and Elizabeth J. Phillips. Pharmacogenomics of off-target adverse drug reactions. *British Journal of Clinical Pharmacology*, 83(9):1896–1911, April 2017. ISSN 1365-2125. doi: 10.1111/bcp.13294. URL http://dx.doi.org/10.1111/bcp.13294.

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*, 28:100496, December 2023a. ISSN 2468-4511. doi: 10.1016/j.cobme.2023.100496. URL http://dx.doi.org/10.1016/j.cobme.2023.100496.

Rohit Singh, Samuel Sledzieski, Bryan Bryson, Lenore Cowen, and Bonnie Berger. Contrastive learning in protein language space predicts interactions between drugs and protein targets. *Proceedings of the National Academy of Sciences*, 120(24):e2220778120, 2023.

Andrew T McNutt, Abhinav K Adduri, Caleb N Ellington, Monica T Dayao, Eric P Xing, Hosein Mohimani, and David R Koes. Sprint enables interpretable and ultra-fast virtual screening against thousands of proteomes. *arXiv preprint arXiv:2411.15418*, 2024.

Bowen Gao, Bo Qiang, Haichuan Tan, Yinjun Jia, Minsi Ren, Minsi Lu, Jingjing Liu, Wei-Ying Ma, and Yanyan Lan. Drugclip: Contrasive protein-molecule representation learning for virtual screening. *Advances in Neural Information Processing Systems*, 36, 2024.

Samuel Sledzieski, Rohit Singh, Lenore Cowen, and Bonnie Berger. D-script translates genome to phenome with sequence-based, structure-aware, genome-scale predictions of protein-protein interactions. *Cell Systems*, 12(10):969–982, 2021.

Rohit Singh, Kapil Devkota, Samuel Sledzieski, Bonnie Berger, and Lenore Cowen. Topsy-turvy: integrating a global view into sequence-based ppi prediction. *Bioinformatics*, 38(Supplement_1): i264–i272, 2022.

Suhaas Bhat, Kalyan Palepu, Lauren Hong, Joey Mao, Tianzheng Ye, Rema Iyer, Lin Zhao, Tianlai Chen, Sophia Vincoff, Rio Watson, Tian Z. Wang, Divya Srijay, Venkata Srikar Kavirayuni, Kseniia Kholina, Shrey Goel, Pranay Vure, Aniruddha J. Deshpande, Scott H. Soderling, Matthew P. DeLisa, and Pranam Chatterjee. De novo design of peptide binders to conformationally diverse targets with contrastive language modeling. *Science Advances*, 11(4), January 2025. ISSN 2375-2548. doi: 10.1126/sciadv.adr8638. URL http://dx.doi.org/10.1126/sciadv.adr8638.

Sophia Tang, Yinuo Zhang, and Pranam Chatterjee. Peptune: De novo generation of therapeutic peptides with multi-objective-guided discrete diffusion. *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 2025.

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.

Andrew Campbell, Jason Yim, Regina Barzilay, Tom Rainforth, and Tommi Jaakkola. Generative flows on discrete state-spaces: enabling multimodal flows with applications to protein co-design. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org, 2024.

Jiaxin Shi, Kehang Han, Zhe Wang, Arnaud Doucet, and Michalis K. Titsias. Simplified and generalized masked diffusion for discrete data. In *Advances in Neural Information Processing Systems*, 2024.

Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T Chiu, Alexander Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. *arXiv preprint arXiv:2406.07524*, 2024.

Xinyou Wang, Zaixiang Zheng, Fei Ye, Dongyu Xue, Shujian Huang, and Quanquan Gu. Diffusion language models are versatile protein learners. *arXiv preprint arXiv:2402.18567*, 2024.

Sarah Alamdari, Nitya Thakkar, Rianne van den Berg, Neil Tenenholtz, Bob Strome, Alan Moses, Alex Xijie Lu, Nicolo Fusi, Ava Pardis Amini, and Kevin K Yang. Protein generation with evolutionary diffusion: sequence is all you need. *BioRxiv*, pages 2023–09, 2023.

Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper, and Demis Hassabis. Protein complex prediction with alphafold-multimer. October 2021. doi: 10.1101/2021.10.04.463034. URL http://dx.doi.org/10.1101/2021.10.04.463034.

Jianlin Su, Murtadha Ahmed, Yu Lu, Shengfeng Pan, Wen Bo, and Yunfeng Liu. Roformer: Enhanced transformer with rotary position embedding. *Neurocomputing*, 568:127063, 2024.

William A Falcon. Pytorch lightning. *GitHub*, 3, 2019.

Tong Chen, Yinuo Zhang, Sophia Tang, and Pranam Chatterjee. Multi-objective-guided discrete flow matching for controllable biological sequence design. *arXiv preprint arXiv:2505.07086*, 2025.

Xiner Li, Yulai Zhao, Chenyu Wang, Gabriele Scalia, Gokcen Eraslan, Surag Nair, Tommaso Biancalani, Shuiwang Ji, Aviv Regev, Sergey Levine, and Masatoshi Uehara. Derivative-free guidance in continuous and discrete diffusion models with soft value-based decoding, 2024. URL https://arxiv.org/abs/2408.08252.

Yair Schiff, Subham Sekhar Sahoo, Hao Phung, Guanghan Wang, Sam Boshar, Hugo Dalla-torre, Bernardo P. de Almeida, Alexander Rush, Thomas Pierrot, and Volodymyr Kuleshov. Simple guidance mechanisms for discrete diffusion models, 2025.

Nate Gruver, Samuel Stanton, Nathan C. Frey, Tim G. J. Rudner, Isidro Hotzel, Julien Lafrance-Vanasse, Arvind Rajpal, Kyunghyun Cho, and Andrew Gordon Wilson. Protein design with guided discrete diffusion, 2023. URL https://arxiv.org/abs/2305.20009.

Fred Zhangzhi Peng, Chentong Wang, Tong Chen, Benjamin Schussheim, Sophia Vincoff, and Pranam Chatterjee. Ptm-mamba: a ptm-aware protein language model with bidirectional gated mamba blocks. *Nature Methods*, 22(5):945–949, April 2025. ISSN 1548-7105. doi: 10.1038/s41592-025-02656-9. URL http://dx.doi.org/10.1038/s41592-025-02656-9.

Garyk Brixi, Tianzheng Ye, Lauren Hong, Tian Wang, Connor Monticello, Natalia Lopez-Barbosa, Sophia Vincoff, Vivian Yudistyra, Lin Zhao, Elena Haarer, Tianlai Chen, Sarah Pertsemlidis, Kalyan Palepu, Suhaas Bhat, Jayani Christopher, Xinning Li, Tong Liu, Sue Zhang, Lillian Petersen, Matthew P. DeLisa, and Pranam Chatterjee. Saltamp;peppr is an interface-predicting language model for designing peptide-guided protein degraders. *Communications Biology*, 6(1), October 2023. ISSN 2399-3642. doi: 10.1038/s42003-023-05464-z. URL http://dx.doi.org/10.1038/s42003-023-05464-z.

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*, 2023b. doi: 10.48550/ARXIV.2310.03842. URL https://arxiv.org/abs/2310.03842.

Lauren Hong, Tianzheng Ye, Tian Z. Wang, Divya Srijay, Howard Liu, Lin Zhao, Rio Watson, Sophia Vincoff, Tianlai Chen, Kseniia Kholina, Shrey Goel, Matthew P. DeLisa, and Pranam Chatterjee. Programmable protein stabilization with language model-derived peptide guides. *Nature Communications*, 16(1), April 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-58872-6. URL http://dx.doi.org/10.1038/s41467-025-58872-6.

Rose Oughtred, Jennifer Rust, Christie Chang, Bobby-Joe Breitkreutz, Chris Stark, Andrew Willems, Lorrie Boucher, Genie Leung, Nadine Kolas, Frederick Zhang, et al. The biogrid database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Science*, 30(1):187–200, 2021.

Noemi Del Toro, Anjali Shrivastava, Eliot Ragueneau, Birgit Meldal, Colin Combe, Elisabet Barrera, Livia Perfetto, Karyn How, Prashansa Ratan, Gautam Shirodkar, et al. The intact database: efficient access to fine-grained molecular interaction data. *Nucleic acids research*, 50(D1):D648–D653, 2022.

Anton Bushuiev, Roman Bushuiev, Anatolii Filkin, Petr Kouba, Marketa Gabrielova, Michal Gabriel, Jiri Sedlar, Tomas Pluskal, Jiri Damborsky, Stanislav Mazurenko, et al. Learning to design protein-protein interactions with enhanced generalization. *arXiv preprint arXiv:2310.18515*, 2023.

Philipp Blohm, Goar Frishman, Pawel Smialowski, Florian Goebels, Benedikt Wachinger, Andreas Ruepp, and Dmitrij Frishman. Negatome 2.0: a database of non-interacting proteins derived by literature mining, manual annotation and protein structure analysis. *Nucleic acids research*, 42 (D1):D396–D400, 2014.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.

Martin Steinegger and Johannes Söding. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology*, 35(11):1026–1028, 2017.

A Algorithm

Algorithm 1 SOAPIA: Pareto-Guided MCTS with Dual Objective Scoring

Input: Denoising model $p_\theta(x_{t-1} \mid x_t, t)$, specificity model $f_{\text{spec}}(x)$, affinity model $f_{\text{aff}}(x)$, number of MCTS iterations N_{mcts}, number of children N_{child}
Output: Pareto frontier of guided binders $\mathcal{P}^*$
Initialize: $x_T \leftarrow [M]^L$ (fully masked), $\mathcal{P}^* \leftarrow \{\}$, $t \leftarrow T$
for $i = 1$ to N_{mcts} **do**
  Selection: $x_{\text{leaf}} \leftarrow \textsc{SelectLeaf}(x_T)$
  Expansion:
  for $j = 1$ to N_{child} **do**
    Sample $x_{t-1}^{(j)} \sim p_\theta(\cdot \mid x_{\text{leaf}}, t)$
    Unmask k positions to form $x_{\text{child}}^{(j)}$
    Add $x_{\text{child}}^{(j)}$ to children(x_{leaf})
  end for
  Rollout:
  for each $x_{\text{child}}^{(j)}$ **do**
    $\tilde{x}^{(j)} \leftarrow \textsc{RolloutToCompletion}(x_{\text{child}}^{(j)})$
    Compute scores: $\mathbf{s}^{(j)} \leftarrow \left[f_{\text{spec}}(\tilde{x}^{(j)}),\ f_{\text{aff}}(\tilde{x}^{(j)})\right]$
    Compute soft reward vector: $\mathbf{r}^{(j)} \leftarrow \textsc{MultiCompareParetoFront}(\tilde{x}^{(j)}, \mathbf{s}^{(j)}, \text{tokens})$
    Update $\mathcal{P}^*$ with $\tilde{x}^{(j)}$ if non-dominated
  end for
  Backpropagation: $\textsc{Backpropagate}(x_{\text{leaf}}, \{\mathbf{r}^{(j)}\})$
end for
return $\mathcal{P}^*$

MultiCompareParetoFront provides a mechanism for prioritizing items on the Pareto front to optimize performance. Over the first few MCTS iterations, binders may be added to the Pareto front even if they are dominated. This ensures that the user receives the total number of binders they have requested. In later iterations, newly generated binders may dominate multiple members of this initially sub-optimal front. MultiCompareParetoFront tracks which Pareto front members dominate each other so that in later iterations, a new superior candidate will replace the weakest current member of the front.

B Supplementary Methods

B.1 Data collection

Each sample in the training data is a protein quadruplet consisting of a binder, target, off-target 1, and off-target 2. Positive protein-protein interactions (PPIs) were collected from BioGRID (October 2022) [Oughtred et al., 2021], IntAct (October 2022) [Del Toro et al., 2022], and PPIRef (January 2025) [Bushuiev et al., 2023]. Negative interactions were collected from Negatome2.0 [Blohm et al., 2014], a manually curated database of proteins with experimental evidence indicating the absence of direct interaction. Positive PPIs were filtered by cleaning (e.g. dropping sequences with non-natural amino acids), swapping targets and binders to double the dataset size, removing duplicate homomer interactions, applying a length limit of 1022 amino acids for both target and binder sequences, and only retaining rows where one partner is included in the Negatome. The Negatome was filtered by removing any listed target and off-target pairs that are included as binder and target pairs in the PPI database. This produced a dataset of 245,587 positive PPIs and 7,198 negative interactions.

B.2 Quadruplet selection

The training, validation, and testing datasets were designed to (1) enhance generalizability by including diverse binders, targets, and off-targets, (2) prevent rigid role assignments by allowing any sequence to act in any interaction context, and (3) maximize difficulty to improve learning. Quadruplet selection was formulated as a linear programming problem using length-averaged ESM-2-650M [Lin et al., 2023] embeddings. With `PuLP v2.9.0`, quadruplets were optimized for difficulty while minimizing role repetition (Supplementary Algorithm **??**). Difficulty increased with higher

cosine similarity between target and off-target embeddings, ensuring the Siamese model learned to distinguish subtle differences. Euclidean distance was used in the loss function, but cosine similarity was preferred for selection due to its bounded range (-1 to 1). Selecting closely related targets and off-targets better reflects SOAPI's real-world applications, such as designing binders that avoid wild-type protein interactions. To ensure the model learned relationships rather than predefined roles, four constraints were imposed: (1) each binder appears at most ten times, (2) each target-off-target grouping appears only once, (3) each target appears once per binder, and (4) each off-target appears once per binder. Constraint 1 was pre-enforced by subsampling positive PPIs, and Constraint 2 required a tiebreaker term ($\lambda = 0.001$) based on Euclidean binder-target distance. A total of 1,352 quadruplets were selected, consisting of 565 unique binders, 1,054 targets, and 969 off-targets.

B.3 Clustering and splitting

Quadruplets were clustered on binder sequence using MMSeqs2 easy clustering module [Steinegger and Söding, 2017] with a minimum sequence identity of 30% and a coverage threshold of 70%. The resulting clusters were randomly split at 80-10-10 ratio using `sklearn v1.2.0` into a training set (1,111 quadruplets, 82.2%), validation set (116 quadruplets, 8.6%), and test set (125 quadruplets, 9.2%). The distribution of roles (binder, target, off-target) played by each sequence in the full dataset and individual splits can be found in Figure A4.

C Supplementary Tables

Table A1: SOAPI loss on training data.

Split	Size	Loss
Train	1111	0.04
Validation	116	1.43
Test	125	1.57

Table A2: Siamese specificity model architecture and training hyperparameters

Hyperparameter	**Value**
Model Architecture	
ESM Model Base	ESM2_t33_650M_UR50D
Embedding Dimension	1280
ESM Unfrozen Layers	2
Linear Layers	2
Positional Attention Head: `n_heads`	10
Adaptive Log-Sum Decoy Loss	
α	5
β	0.5
ϵ	1e-8
Training	
Max Sequence Length	1022
Batch Size / Device	16
Effective Batch Size	128
Dataloader `num_workers`	30
Learning Rate (LR)	1e-4
LR Scheduler: Warmup Steps	200
LR Scheduler: Total Steps	9000
LR Scheduler: Min/Max LR Ratio	0.1
Gradient Clipping	0.5

Table A3: Peptide sequences for fusion oncoprotein visualizations.

Target	**Method**	**Specificity**	**Affinity**	**Proposed sequence**
APTX::ARL5B	NOS	3.67	6.97	AEMQIWMWGTLKDVESMKQF
ETHE1::PLAUR	Unguided	12.72	6.81	MTCAYRGLKLQDYMRLYPDL
DHRSX::RPS4Y1	Simple	11.13	7.71	CQWLWRQRCVEQLKISLSWS
AMACR::NDUFAF2	SVDD	4.18	7.11	QFLSERDRGYGIVLKVLPPN

Table A4: Performance of *MultiMCTS* guidance strategies in producing “green zone" binders (passing both objectives; specificity > 0, affinity > 6.5). The model was tasked with producing 10 binders for each of 17 fusion oncoproteins. However, *MultiMCTS* will return additional binders if they lie on the Pareto Front.

Guidance	Green Zone Binders	Green Zone Targets (/17)
NOS	143 / 251 (57.0%)	13 (76.5%)
Simple	170 / 239 (71.1%)	13 (76.5%)
SVDD	123 / 242 (50.8%)	13 (76.5%)
Unguided	133 / 252 (55.0%)	13 (76.5%)
Total	569 / 984 (57.8%)	13/17 (76.5%)

Table A5: Performance of green zone binders for each *MultiMCTS* guidance strategy in AlphaFold-Multimer. A peptide is considered a hit when its ipTM with the target is higher than its ipTM with either off-target.

Guidance	AFM Top-25 Hit Rate	AFM Hit Rate	Targets Hit
NOS	6 (24%)	39 / 143 (27.3%)	8 / 13 (61.5%)
Simple	7 (28%)	59 / 170 (34.7%)	7 / 13 (53.8%)
SVDD	4 (16%)	34 / 123 (27.6%)	12 / 13 (92.3%)
Unguided	4 (16%)	48 / 133 (36.1%)	12 / 13 (92.3%)
Total	21 (21%)	180 / 569 (31.6%)	12 / 13 (92.3%)

Table A6: AlphaFold-Multimer results indicating high specificity (target ipTM $>$ offtarget1 ipTM and offtarget2 ipTM) and high affinity (target ipTM > 0.7) on fusion oncoproteins, across all four *MultiMCTS* guidance strategies.

Guidance	AFM Dual-Objective Hit Rate
NOS	11 / 143 (7.7%)
Simple	16 / 170 (9.4%)
SVDD	5 / 123 (4.1%)
Unguided	14 / 133 (10.5%)
Total	46 / 569 (8.1%)

D Supplementary Figures

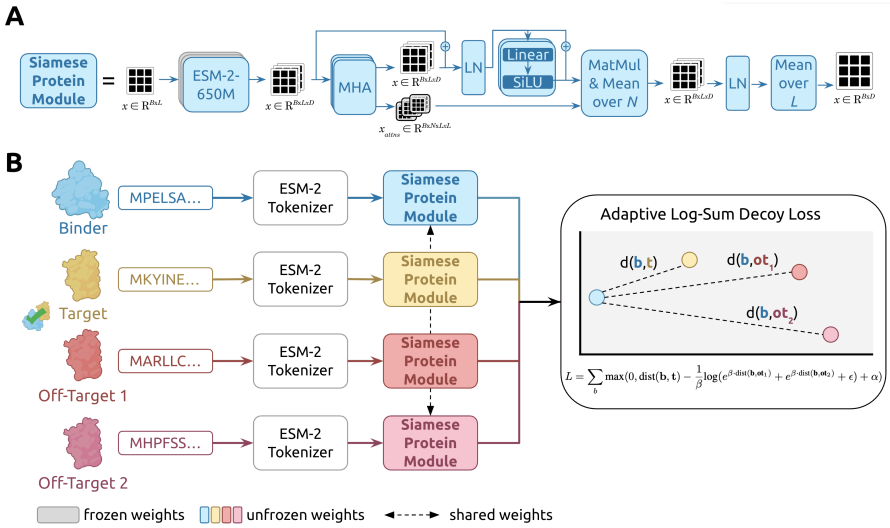

Figure A1: **Siamese specificity model architecture. A** Siamese Protein Module, the core unit of the quadruplet network. ESM-2-650M encodes sequences into [B,1280] embeddings, refined via positional attention (10 heads) with rotary embeddings, skip-connected linear layers, and attention pooling. **B** Full SOAPI pipeline. Binder, target, and off-target sequences pass through the Siamese module with shared weights. Euclidean distances between embeddings define a loss that pulls the binder toward the target while pushing it away from off-targets.

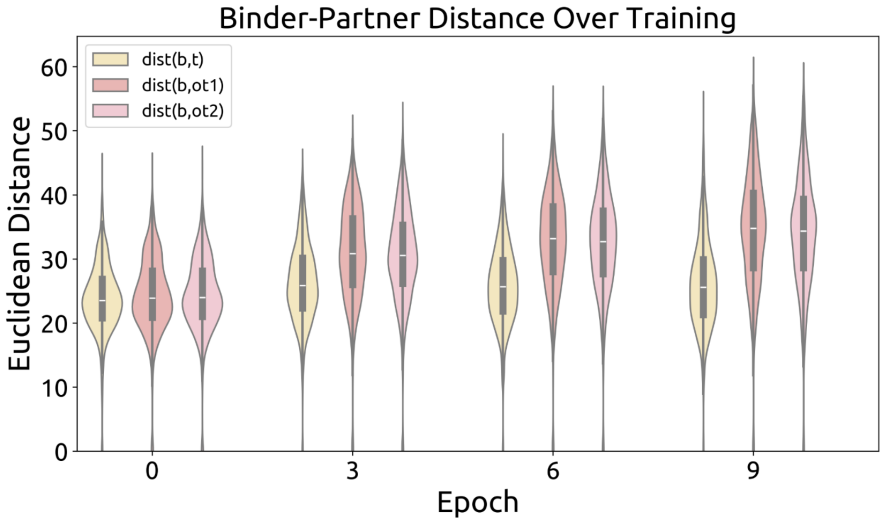

Figure A2: **Embedding separation throughout model training.** Euclidean distances between SOAPI embeddings of the proteins in each training quadruplet: dist($\mathbf{b}, \mathbf{t}$) (binder and target), dist($\mathbf{b}, \mathbf{ot}_1$) (binder and off-target 1), and dist($\mathbf{b}, \mathbf{ot}_2$) (binder and off-target 2). The inner box of each violin plot indicates the median (white) and inter-quartile range, representing the middle 50% of distances (grey box). Distances are plotted every three epochs throughout training, starting at the end of epoch 1.

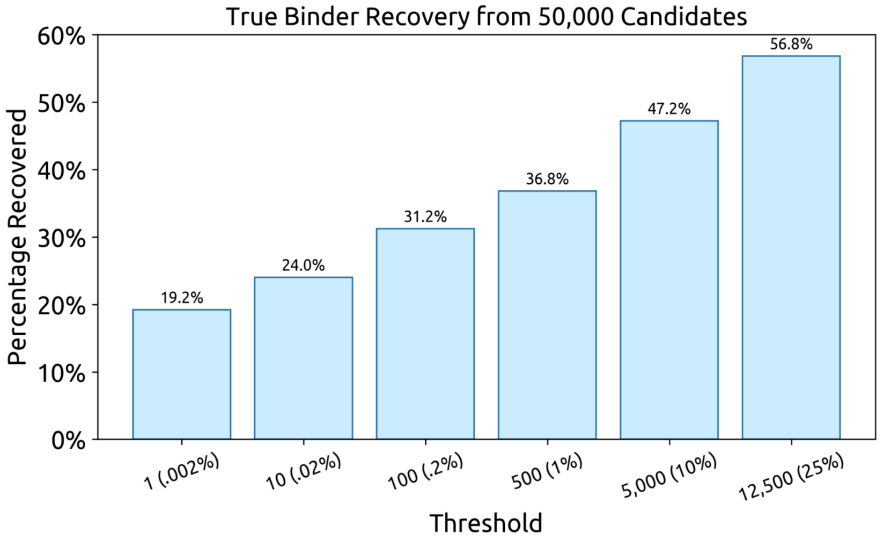

Figure A3: **Siamese model ranks true binders against 50,000 randomly generated candidates.** SOAPI produced a specificity-based ranking of 50,000 randomly generated binders and one true binder to 125 targets from the test set. Relative distance metrics D were calculated using Equation (2). Any target where SOAPI ranked the true binder among the top-N or top-$N\%$ was considered a hit. Top-1, top-10, top-100, top-1%, top-10%, and top-25% evaluations were conducted.

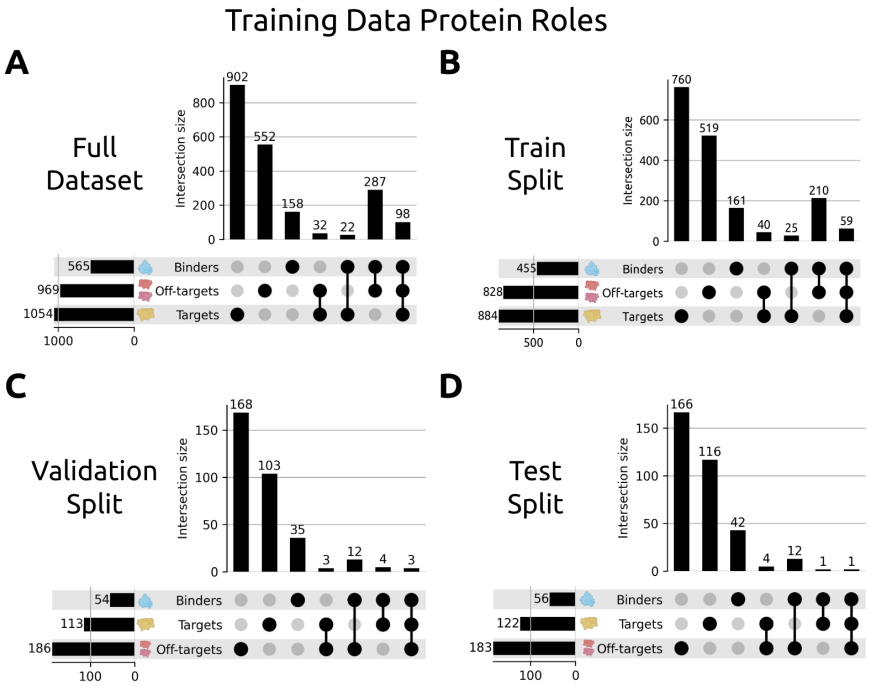

Figure A4: **Training Data Protein Roles.** Distribution of roles (binder, target, off-target) played by proteins in the quadruplets comprising the **A** full training dataset (1,352 quadruplets), **B** train split (1,111 quadruplets), **C** validation split (116 quadruplets), and **D** test spit (125 quadruplets).