Multi-Objective Nanobody Design via Masked Discrete Diffusion with Simplex Refinement

Anonymous Author(s)
Affiliation
Address
email

Abstract

Nanobodies are compact, stable, and highly specific binding proteins that can access epitopes inaccessible to conventional antibodies, making them ideal scaffolds for therapeutic design. We present a masked discrete denoising framework for nanobody generation (*NanoMDLM*) that learns to reconstruct CDRs on a fixed scaffold, with region-specific masking that emphasizes diversity in CDR3. For inference, we develop a platform for **N**anobody **O**ptimization for **S**elective **I**nteraction and **E**nhanced properties (**NOSIE**) via discrete simplex refinement (*DSR*), a gradient-free, black-box guidance method that samples CDR completions and reweights them using Pareto-weighted softmax over predicted binding and stability scores. At inference time, DSR steers NanoMDLM toward high-performing sequences without retraining or differentiable reward access. Across multiple antigens, including the GPCR MRGPRX2, NOSIE produces nanobodies with competitive or superior *in silico* binding, thermostability, and CDR3 quality, as assessed by NanoNet structure prediction, AlphaFold-Multimer co-folding, and feature combination-based ranking. Together, these results provide a scalable, sequence-only framework for multi-objective nanobody design, enabling numerous therapeutic applications.

1 Introduction

Nanobodies, or VHH single-domain antibodies derived from camelids, are attractive therapeutic and diagnostic agents due to their small size, stability, solubility, and ability to access buried epitopes [Jovčevska and Muyldermans, 2020, Jin et al., 2023, Salvador et al., 2019, Uchański et al., 2020]. Their binding function is localized to three complementarity-determining regions (CDRs), with CDR3 contributing most to antigen specificity [Mitchell and Colwell, 2018]. *In silico* design of nanobody CDRs offers a scalable alternative to experimental screening [Longsompurana et al., 2023, Ferraz et al., 2025], but remains challenging due to the combinatorial space of loop conformations and the need to optimize multiple properties including affinity, immunogenicity, and thermostability [Tadokoro et al., 2024].

Structure-based nanobody models, such as a fine-tuned RFdiffusion (RFantibody), have enabled scaffold-aware design of nanobodies but require accurate 3D structures [Watson et al., 2023, Sumida et al., 2024, Bennett et al., 2025]. In contrast, masked discrete diffusion language models (MDLMs) offer a scalable sequence-only alternative and have been applied to peptide and protein design [Sahoo et al., 2024, Shi et al., 2024, Wang et al., 2024, Goel et al., 2025, Tang et al., 2025, Vincoff et al., 2025a]. Recent guidance strategies for discrete diffusion, including CFG [Ho and Salimans, 2022], LaMBO-2/NOS [Gruver et al., 2023], and PepTune's MCTS-guided sampling [Tang et al., 2025], have enabled post-hoc optimization of pretrained models. However, these methods assume full-sequence mutation and do not address modular protein design settings, such as nanobodies, where

Submitted to the FM4LS Workshop at NeurIPS 2025.

CDRs are preferentially altered for re-targeting. In this work, we introduce a nanobody-specific MDLM (**NanoMDLM**) trained with CDR-only corruption and region-dependent masking rates (Figure 1), combined with a discrete simplex refinement (DSR) framework that conducts **N**anobody **O**ptimization for **S**elective **I**nteraction and **E**nhanced properties (**NOSIE**). DSR uses black-box binding and stability predictors to guide NanoMDLM generation via Pareto-weighted reweighting over sampled completions, without requiring gradients or retraining (Figure 2). Across diverse antigens, including the GPCR MRGPRX2, our NOSIE (NanoMDLM + DSR) framework generates nanobody candidates with superior *in silico* performance compared to structure-based baselines, matching or exceeding them in predicted binding, thermostability, and CDR3 quality.

Our contributions are as follows:

- We develop a CDR-constrained NanoMDLM trained with region-specific, rate-dependent masking, emphasizing CDR3 while keeping the scaffold fixed.
- We introduce discrete simplex refinement (DSR): a gradient-free, black-box inference-time guidance algorithm over categorical denoising distributions.
- We frame DSR as a constrained stochastic optimization problem over the discrete simplex, enabling modular, multi-objective generation.
- We benchmark our integrated NanoMDLM + DSR framework, NOSIE, against structure-based models for nanobody design and show superior or comparable performance in binding, stability, and structural plausibility.

2 Problem Formulation

Here we present the formulation of NOSIE. Detailed dataset and implementation methods can be found in Appendix Sections C and D.

2.1 Nanobody Sequence Representation and CDR Masking

Let $x \in \Sigma^L$ be a nanobody amino acid sequence of length L, where Σ is the alphabet of 20 canonical amino acids. The sequence is partitioned into:

$$x = (x_{\text{FR1}}, x_{\text{CDR1}}, x_{\text{FR2}}, x_{\text{CDR2}}, x_{\text{FR3}}, x_{\text{CDR3}}, x_{\text{FR4}}),$$

where the CDR regions $\mathcal{C} = \{\text{CDR1}, \text{CDR2}, \text{CDR3}\} \subset \{1, \ldots, L\}$ define the support over which generation will occur. The framework regions (FRs) are treated as fixed context, derived from a canonical humanized scaffold x_{scaffold}, and are never perturbed or predicted during training or inference.

2.2 Nanobody Masked Discrete Diffusion Model (NanoMDLM)

We define a forward corruption process $q(x^{(t)} \mid x^{(t-1)})$ over timesteps $t = 1, \ldots, T$ that progressively masks only the complementarity-determining regions (CDRs) of the nanobody sequence, while treating the scaffold residues as fixed context. Let $x^{(0)}$ denote the uncorrupted input sequence, and $x^{(T)}$ the fully masked CDRs.

We adopt a domain-specific corruption schedule, assigning each masked position $i \in \mathcal{C}$ a fixed corruption strength $\gamma(i) \in \{\gamma_{\text{weak}}, \gamma_{\text{strong}}\}$ based on its membership in the CDRs:

$$\gamma(i) = \begin{cases} \gamma_{\text{weak}} & \text{if } i \in \text{CDR1 or CDR2}, \\ \gamma_{\text{strong}} & \text{if } i \in \text{CDR3}, \end{cases} \qquad \text{with } \gamma_{\text{strong}} > \gamma_{\text{weak}}.$$

Given a masking schedule $\alpha_t \in [0, 1]$, the corruption process is defined as:

$$q(x^{(t)} \mid x^{(t-1)}) = \prod_{i \in \mathcal{C}} \text{Cat}\left(x_i^{(t)} \mid \alpha_t^{\gamma(i)} \delta_{x_i^{(t-1)}} + \left(1 - \alpha_t^{\gamma(i)}\right) u\right),$$

where u is the uniform distribution over amino acids and δ is the Kronecker delta.

Scaffold positions $j \notin \mathcal{C}$ are never corrupted and remain fixed: $x_j^{(t)} = x_j^{(0)}$ for all t.

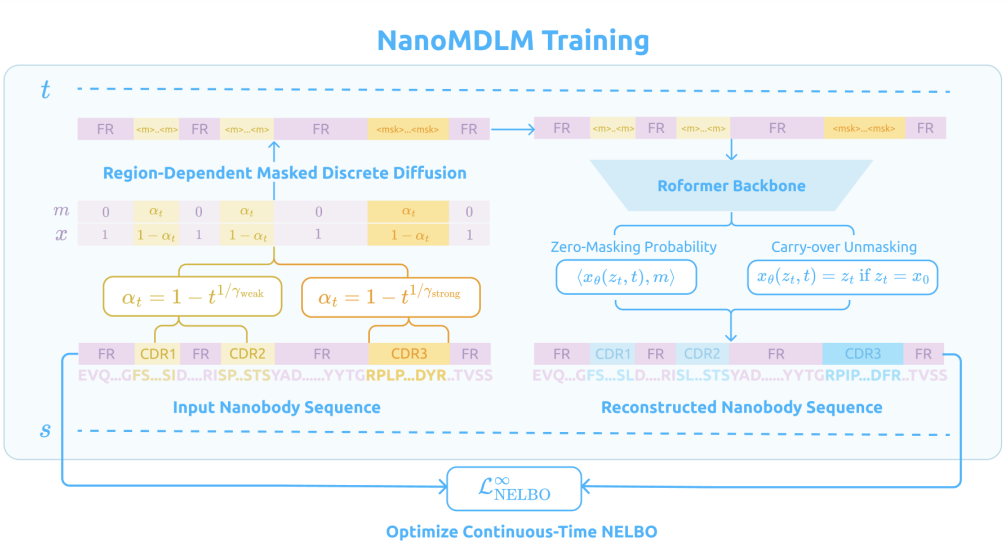

Figure 1: **NanoMDLM**. Continuous-time training of discrete masked diffusion model NanoMDLM under Region-Specific masking schedules, with SUBS parametrization and RoFormer as the denoising backbone.

The denoising model $p_\theta(x^{(t-1)} \mid x^{(t)})$ learns to invert this process by reconstructing masked CDRs given the corrupted sequence and fixed scaffold. The training objective is:

$$\mathcal{L}_{\text{MDLM}} = \mathbb{E}_{t,x^{(0)},x^{(t)}\sim q}\left[-\sum_{i\in\mathcal{C}} \log p_\theta\left(x_i^{(0)} \mid x^{(t)}\right)\right].$$

2.3 Inference-Time Optimization via Discrete Simplex Refinement (NOSIE)

Let $x^{(t)}$ denote the corrupted CDR sequence at time t, and let $p_\theta(x^{(t-1)} \mid x^{(t)})$ be the model's categorical output distribution over each masked CDR position. Denote this distribution over tokens at position i as $\pi_i^{(t)} \in \Delta^{|\Sigma|}$, where $\Delta^{|\Sigma|}$ is the probability simplex.

Our goal is to guide the generation process toward satisfying black-box reward functions:

$$f_{\text{bind}} : \Sigma^L \to \mathbb{R}, \quad f_{\text{stab}} : \Sigma^L \to \mathbb{R},$$

corresponding to predicted binding affinity and thermostability, respectively. These functions are not differentiable and cannot be backpropagated through.

At each denoising step, we sample K candidate completions from the model:

$$\{x_k^{(t-1)}\}_{k=1}^K \sim p_\theta(x^{(t-1)} \mid x^{(t)}),$$

evaluate their rewards, and define a scalarized objective:

$$R_k^{(\lambda)} = \lambda_{\text{bind}} f_{\text{bind}}(x_k^{(t-1)}) + \lambda_{\text{stab}} f_{\text{stab}}(x_k^{(t-1)}),$$

for scalarization weights $\lambda = (\lambda_{\text{bind}}, \lambda_{\text{stab}}) \in \Delta^2$.

We then compute Pareto-weighted sampling probabilities using softmax:

$$w_k = \frac{\exp(\beta R_k^{(\lambda)})}{\sum_{j=1}^K \exp(\beta R_j^{(\lambda)})},$$

and update the token probabilities via empirical reweighting:

$$\tilde{\pi}_i^{(t)}(a) = \sum_{k=1}^K w_k \cdot \mathbb{1}\{x_{k,i}^{(t-1)} = a\}, \quad \forall i \in \mathcal{C},\ a \in \Sigma.$$

This yields a refined distribution $\tilde{\pi}^{(t)}$ that remains on the categorical simplex and concentrates probability mass on reward-optimal trajectories.

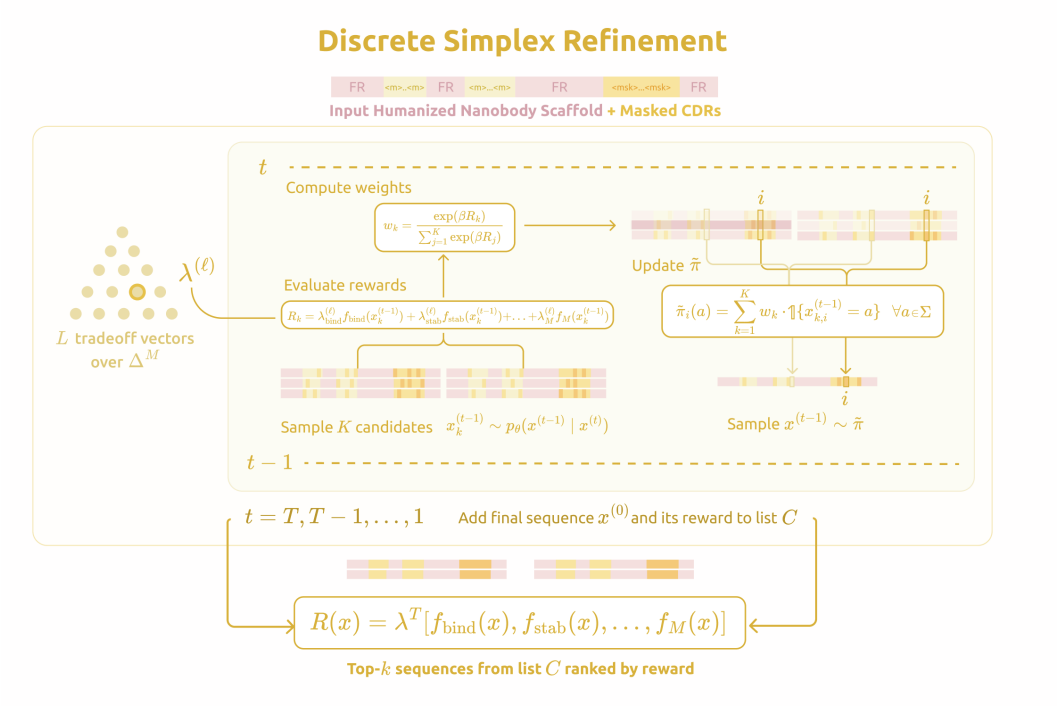

Figure 2: **Discrete Simplex Refinement**. Sampling with Discrete Simplex Refinement (DSR) where humanized scaffolds with masked CDRs are iteratively refined by candidate sampling, reward evaluation, and tradeoff vector weighting, producing ranked nanobody sequences with optimized binding, stability, and multi-objective performance.

2.4 Theoretical Framing of NOSIE

We frame this process as a constrained stochastic optimization problem over a product of categorical simplices. The denoising process is viewed as a policy over discrete action sequences $\{x^{(0)}, \ldots, x^{(T)}\}$, and the guidance procedure acts as a black-box policy improvement step at each time t, satisfying the update:

$$\tilde{\pi}^{(t)} = \arg \min_{\pi \in \Delta^{|\Sigma|^{|\mathcal{C}|}}} \mathrm{KL}(\pi \| \pi^{(t)}) - \eta \sum_{k=1}^{K} R_k^{(\lambda)} \log \pi(x_k),$$

where $\pi^{(t)}$ is the base model distribution and $\tilde{\pi}^{(t)}$ is the refined posterior, as in mirror descent with KL divergence.

This update guarantees an ascent in the expected scalarized reward under the current sampling distribution (See Appendix Section B for proof). Moreover, by varying λ over a Das-Dennis scalarization lattice on Δ^2, we approximate the Pareto frontier of reward-optimal CDR sequences across tradeoffs between binding and stability.

3 Results

3.1 Guided Generation Scores

We assessed NOSIE's ability to generate nanobodies with high predicted binding affinity and thermostability across a panel of diverse antigens. For each target, we compared unconditional sampling,

MCTS-guided optimization introduced in PepTune [Tang et al., 2025], and our DSR method using DeepNano [Deng et al., 2024] and NanoMelt [Ramon et al., 2025] predictors as evaluation metrics for nanobody-antigen interaction and thermostability respectively. Figure 3 shows results for antigen PD-L1. While the unconditional model produces plausible but suboptimal sequences, both MCTS and NOSIE improve over this baseline. Notably, NOSIE consistently yields nanobody candidates with stronger predicted binding and stable thermostability, outperforming MCTS across the top-ranked sequences. These trends are further supported across additional antigens in Figure A6, with NOSIE recovering sequences that frequently match or exceed the experimental nanobodies in affinity while maintaining favorable developability.

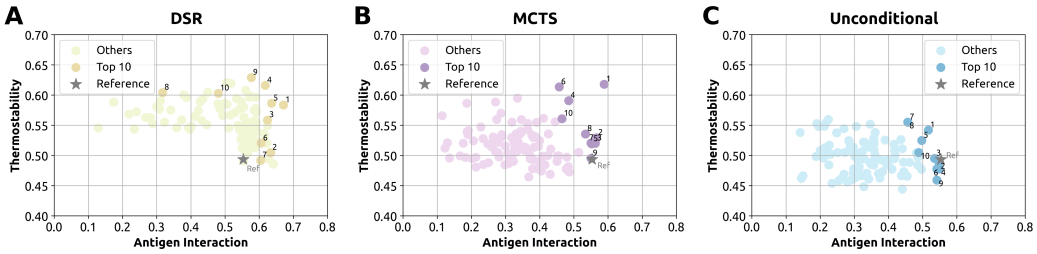

Figure 3: Comparison of generated nanobody scores with antigen PD-L1, topK=100. Reference point represents the scores of the known binding VHH (PDB: 7CZD_2). All three sampling methods were run with 32 sampling timesteps. The x-axis shows antigen interaction scores predicted by DeepNano, while y-axis shows thermostability scores regularized from NanoMelt-predicted melting temperatures. For unconditional and MCTS methods, the top 100 sequences were ranked by the average of the antigen interaction and thermostability scores, whereas NOSIE sequences were ranked according to scalar rewards defined by the algorithm.

Even in more challenging cases such as the high-affinity antigen CTLA-4 (PDB:7DV4_1) [Gan et al., 2022](Figure A6F), where reference binders are already near optimal, DSR produces competitive alternatives with improved stability. This suggests that simplex-guided denoising effectively identifies viable CDR variants even when the scope for affinity improvement is narrow, and excels in discovering high-quality solutions in moderate-to-difficult design settings.

3.2 Validation on MRGPRX2

MRGPRX2 is an orphan G protein–coupled receptor expressed on mast cells that mediates IgE-independent hypersensitivity and inflammatory responses, making it an attractive therapeutic target for drug-induced pseudo-allergic reactions and related diseases [Al Hamwi et al., 2025].

We selected MRGPRX2 as the benchmark antigen for two main reasons: (1) biologically, its shallow binding pocket, promiscuous ligand recognition, and clinical relevance make it a strong candidate for *in silico* nanobody design [Al Hamwi et al., 2022], and (2) computationally, recent work shows that AlphaFold-Multimer (AFM) [Evans et al., 2021] can reliably distinguish positive from negative binding pairs for MRGPRX2 (a discriminative capability not established for most other antigens) making it a robust test-bed for benchmarking nanobody design methods [Harvey et al., 2025].

Table 1: Comparison of average AFM metrics (mean with SEM in parentheses) across 20 pilot nanobody sequences targeting MRGPRX2, generated with RFantibody (RFa) and NOSIE, evaluated using the best-ranked AFM model for each complex. The input to RFa-AF is an AF2 predicted structure of MRGPRX2, while the input to RFa-PDB is the PDB:7VV6 structure of MRGPRX2.

Method	**LCF** ↑	**pTM** ↑	**ipTM** ↑	**pAE** ↓	**pDockQ** ↑
RFa-AF	0.039 (0.006)	0.679 (0.011)	0.402 (0.031)	15.637 (0.830)	0.290 (0.025)
RFa-PDB	0.045 (0.009)	**0.690 (0.013)**	0.428 (0.051)	14.401 (1.335)	**0.317 (0.042)**
NOSIE	**0.045 (0.005)**	0.681 (0.011)	**0.450 (0.031)**	**13.289 (0.637)**	0.288 (0.018)

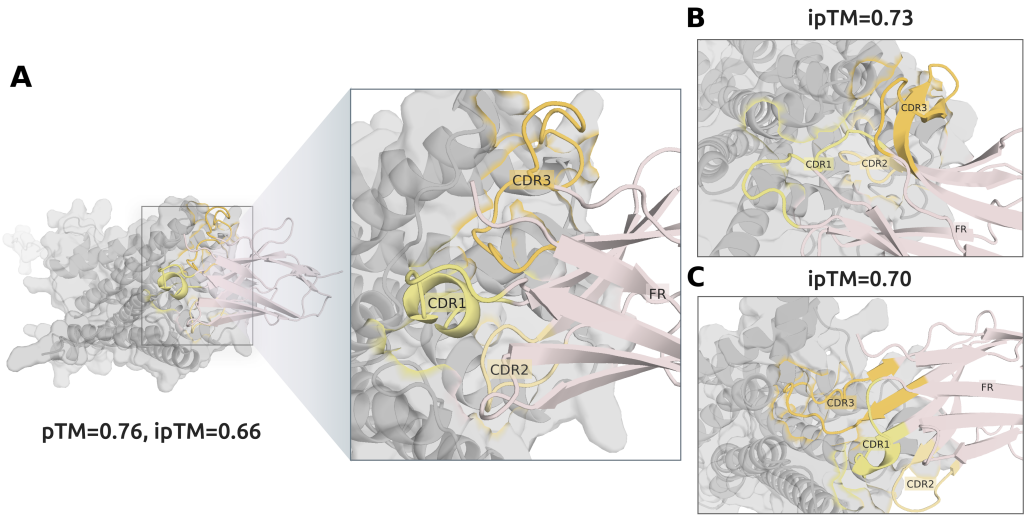

Figure 4: AlphaFold2-Multimer-predicted structures of the three top NOSIE nanobody binders to MRGPRX2

We benchmarked three nanobody libraries using AFM confidence metrics, including the Linear Combination Feature (LCF) [Harvey et al., 2025]. A pilot set of 20 RFantibody Bennett et al. [2025] sequences was compared to 20 NOSIE designs. Despite the small sample size, AFM consistently showed NOSIE performing on par or better: its mean LCF matched RFantibody (0.045 vs. 0.039/0.045) but with lower variance, and it outperformed on pAE (13.289 vs. 15.637/14.401) and ipTM (0.450 vs. 0.402/0.428) (Table 1). Predicted complexes show confident binding interfaces and plausible CDR3 structures (Figure 4), highlighting NOSIE's ability to generate high-quality nanobodies without structural input. Unlike RFantibody, which benefits from access to experimental antigen structures (as seen in the performance gap between RFa-AF and RFa-PDB) Bennett et al. [2025], NOSIE operates with only the antigen sequence, yet is still able to produce a meaningful fraction of quality binders, suggesting its strong potential for generalizable applications.

4 Discussion

Our results demonstrate that CDR-constrained masked discrete denoising, combined with black-box multi-objective simplex refinement, enables potent nanobody design – all without requiring structural input or differentiable objectives. By restricting corruption to CDRs and emphasizing CDR3 through region-dependent masking, our NanoMDLM captures functional diversity while preserving scaffold integrity. DSR further improves sequence quality through gradient-free inference-time guidance, projecting Pareto-weighted reward estimates back onto the categorical simplex. Benchmarking against RFantibody Bennett et al. [2025], our method achieves strong performance in predicted binding, buried surface area, and stability, with structurally plausible CDR3 loops validated via NanoNet prediction Cohen et al. [2022] and AFM-based co-folding Evans et al. [2021].

Looking ahead, the NOSIE framework is extensible to additional black-box objectives such as immunogenicity, pH sensitivity, or expression propensity, and can be integrated with high-throughput wet-lab selection Xia et al. [2025], which are currently ongoing. Moreover, by conditioning reward models on target isoforms Vincoff et al. [2025b,a] or post-translational states Peng et al. [2025], future work can expand this strategy toward highly specific nanobody generation that discriminates between closely related proteoforms, enabling a new class of programmable and precision-targeted biologics.

References

Ivana Jovčevska and Serge Muyldermans. The therapeutic potential of nanobodies. *BioDrugs*, 34(1): 11–26, 2020.

Bo-Kyung Jin, Steven Odongo, Magdalena Radwanska, and Stefan Magez. Nanobodies®: a review of diagnostic and therapeutic applications. *International journal of molecular sciences*, 24(6): 5994–5994, 2023.

J-Pablo Salvador, Lluïsa Vilaplana, and M-Pilar Marco. Nanobody: outstanding features for diagnostic and therapeutic applications. *Analytical and bioanalytical chemistry*, 411(9):1703–1713, 2019.

Tomasz Uchański, Els Pardon, and Jan Steyaert. Nanobodies to study protein conformational states. *Current opinion in structural biology*, 60:117–123, 2020.

Laura S Mitchell and Lucy J Colwell. Comparative analysis of nanobody sequence and structure data. *Proteins: Structure, Function, and Bioinformatics*, 86(7):697–706, 2018.

Phoomintara Longsompurana, Thanyada Rungrotmongkol, Nongluk Plongthongkum, Kittikhun Wangkanont, Peter Wolschann, and Rungtiva P Poo-Arporn. Computational design of novel nanobodies targeting the receptor binding domain of variants of concern of sars-cov-2. *Plos one*, 18(10):e0293263, 2023.

Matheus VF Ferraz, W Camilla S Adan, Tayná E Lima, Adriele JC Santos, Sérgio O de Paula, Rafael Dhalia, Gabriel L Wallau, Rebecca C Wade, Isabelle FT Viana, and Roberto D Lins. Design of nanobody targeting sars-cov-2 spike glycoprotein using cdr-grafting assisted by molecular simulation and machine learning. *PLOS Computational Biology*, 21(4):e1012921, 2025.

Takashi Tadokoro, Harumi Tsuboi, Kota Nakamura, Tetsushi Hayakawa, Reo Ohmura, Izumi Kato, Masaki Inoue, Shin-ichi Tsunoda, Sayaka Niizuma, Yukari Okada, et al. Thermostability and binding properties of single-chained fv fragments derived from therapeutic antibodies. *Protein Science*, 33(7):e5084, 2024.

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.

Kiera H Sumida, Reyes Núñez-Franco, Indrek Kalvet, Samuel J Pellock, Basile IM Wicky, Lukas F Milles, Justas Dauparas, Jue Wang, Yakov Kipnis, Noel Jameson, et al. Improving protein expression, stability, and function with proteinmpnn. *Journal of the American Chemical Society*, 146(3):2054–2061, 2024.

Nathaniel R Bennett, Joseph L Watson, Robert J Ragotte, Andrew J Borst, DéJenaé L See, Connor Weidle, Riti Biswas, Yutong Yu, Ellen L Shrock, Russell Ault, et al. Atomically accurate de novo design of antibodies with rfdiffusion. *bioRxiv*, pages 2024–03, 2025.

Subham Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin Chiu, Alexander Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. *Advances in Neural Information Processing Systems*, 37:130136–130184, 2024.

Jiaxin Shi, Kehang Han, Zhe Wang, Arnaud Doucet, and Michalis Titsias. Simplified and generalized masked diffusion for discrete data. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL `https://openreview.net/forum?id=xcqSOfHt4g`.

Xinyou Wang, Zaixiang Zheng, Fei YE, Dongyu Xue, Shujian Huang, and Quanquan Gu. Diffusion language models are versatile protein learners. In *Forty-first International Conference on Machine Learning*, 2024. URL `https://openreview.net/forum?id=NUAbSFqyqb`.

Shrey Goel, Vishrut Thoutam, Edgar Mariano Marroquin, Aaron Gokaslan, Arash Firouzbakht, Sophia Vincoff, Volodymyr Kuleshov, Huong T Kratochvil, and Pranam Chatterjee. Memdlm: De novo membrane protein design with property-guided discrete diffusion. In *Learning Meaningful Representations of Life (LMRL) Workshop at ICLR 2025*, 2025.

Sophia Tang, Yinuo Zhang, and Pranam Chatterjee. Peptune: De novo generation of therapeutic peptides with multi-objective-guided discrete diffusion. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=FQoy1Y1Hd8.

Sophia Vincoff, Oscar Davis, Ismail Ilkan Ceylan, Alexander Tong, Joey Bose, and Pranam Chatterjee. SOAPIA: Siamese-guided generation of off target-avoiding protein interactions with high target affinity. In *ICML 2025 Workshop on Scaling Up Intervention Models*, 2025a. URL https://openreview.net/forum?id=j0OpIG7leX.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

Nate Gruver, Samuel Stanton, Nathan Frey, Tim GJ Rudner, Isidro Hotzel, Julien Lafrance-Vanasse, Arvind Rajpal, Kyunghyun Cho, and Andrew G Wilson. Protein design with guided discrete diffusion. *Advances in neural information processing systems*, 36:12489–12517, 2023.

Juntao Deng, Miao Gu, Pengyan Zhang, Mingyu Dong, Tao Liu, Yabin Zhang, and Min Liu. Nanobody–antigen interaction prediction with ensemble deep learning and prompt-based protein language models. *Nature Machine Intelligence*, 6(12):1594–1604, 2024.

Aubin Ramon, Mingyang Ni, Olga Predeina, Rebecca Gaffey, Patrick Kunz, Shimobi Onuoha, and Pietro Sormanni. Prediction of protein biophysical traits from limited data: a case study on nanobody thermostability through nanomelt. *mAbs*, 17(1), January 2025. ISSN 1942-0870. doi: 10.1080/19420862.2024.2442750. URL http://dx.doi.org/10.1080/19420862.2024.2442750.

Xin Gan, Qianqian Shan, He Li, Rick Janssens, Yuqiang Shen, Yun He, Fei Chen, Rien Van Haperen, Dubravka Drabek, Jin Li, et al. An anti-ctla-4 heavy chain–only antibody with enhanced treg depletion shows excellent preclinical efficacy and safety profile. *Proceedings of the National Academy of Sciences*, 119(32):e2200879119, 2022.

Ghazl Al Hamwi, Mohamad Wessam Alnouri, Sven Verdonck, Piotr Leonczak, Shaswati Chaki, Stefan Frischbutter, Pavel Kolkhir, Michaela Matthey, Constantin Kopp, Marek Bednarski, et al. Subnanomolar mas-related g protein-coupled receptor-x2/b2 antagonists with efficacy in human mast cells and disease models. *Signal transduction and targeted therapy*, 10(1):128, 2025.

Ghazl Al Hamwi, Yvonne K Riedel, Sophie Clemens, Vigneshwaran Namasivayam, Dominik Thimm, and Christa E Müller. Mas-related g protein-coupled receptors x (mrgprx): Orphan gpcrs with potential as targets for future drugs. *Pharmacology & Therapeutics*, 238:108259, 2022.

Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, et al. Protein complex prediction with alphafold-multimer. *biorxiv*, pages 2021–10, 2021.

Edward P Harvey, Jeffrey S Smith, Joseph D Hurley, Alyana Granados, Ernst W Schmid, Jason G Liang-Lin, Steffanie Paul, Emily M Meara, Matthew P Ferguson, Victor G Calvillo-Miranda, et al. In silico discovery of nanobody binders to a g-protein coupled receptor using alphafold-multimer. *bioRxiv*, pages 2025–03, 2025.

Tomer Cohen, Matan Halfon, and Dina Schneidman-Duhovny. Nanonet: Rapid and accurate end-to-end nanobody modeling by deep learning. *Frontiers in Immunology*, 13, August 2022. ISSN 1664-3224. doi: 10.3389/fimmu.2022.958584. URL http://dx.doi.org/10.3389/fimmu.2022.958584.

Baolong Xia, Ah-Ram Kim, Feimei Liu, Myeonghoon Han, Emily Stoneburner, Stephanie Makdissi, Francesca Di Cara, Stephanie E Mohr, Aaron M Ring, and Norbert Perrimon. Phage-displayed synthetic library and screening platform for nanobody discovery. *eLife*, June 2025. doi: 10.7554/elife.105887.2. URL http://dx.doi.org/10.7554/eLife.105887.2.

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 2025b. ISSN 2041-1723. doi: 10.1038/s41467-025-56745-6. URL http://dx.doi.org/10.1038/s41467-025-56745-6.

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.

Jian Ma, Fandi Wu, Tingyang Xu, Shaoyong Xu, Wei Liu, Divin Yan, Qifeng Bai, and Jianhua Yao. An adaptive autoregressive diffusion approach to design active humanized antibody and nanobody. *bioRxiv*, pages 2024–10, 2024.

Constantin Schneider, Matthew IJ Raybould, and Charlotte M Deane. Sabdab in the age of biotherapeutics: updates including sabdab-nano, the nanobody structure tracker. *Nucleic acids research*, 50(D1):D1368–D1372, 2022.

Cecile Vincke, Remy Loris, Dirk Saerens, Sergio Martinez-Rodriguez, Serge Muyldermans, and Katja Conrath. General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *Journal of Biological Chemistry*, 284(5):3273–3284, 2009.

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

A Algorithm

We present the full pseudocode for nanobody generation via masked discrete denoising and inference-time multi-objective guidance. The framework comprises two stages: training a CDR-constrained MDLM, and guided generation using discrete simplex refinement.

Algorithm 1 Training CDR-Constrained MDLM for Nanobody Generation

Input: Nanobody sequence dataset $\mathcal{D}$, scaffold mask $\mathcal{S}$, CDR mask $\mathcal{C}$, corruption schedule $\{\alpha_t\}_{t=1}^{T}$, CDR-specific exponents $\gamma_{\text{weak}}, \gamma_{\text{strong}}$
for each training step **do**
  Sample sequence $x^{(0)} \sim \mathcal{D}$
  Initialize $x^{(t)} \leftarrow x^{(0)}$
  for each position $i \in \mathcal{C}$ **do**
    Determine $\gamma(i) \leftarrow \begin{cases} \gamma_{\text{weak}}, & i \in \text{CDR1 or CDR2} \\ \gamma_{\text{strong}}, & i \in \text{CDR3} \end{cases}$
    Sample masking with probability $1 - \alpha_t^{\gamma(i)}$
  end for
  Construct corrupted input $x^{(t)}$ with sampled masks; leave $x_j^{(t)} = x_j^{(0)}$ for $j \in \mathcal{S}$
  Compute model logits $p_\theta(x_i^{(0)} \mid x^{(t)})$ for $i \in \mathcal{C}$
  Compute cross-entropy loss:

$$\mathcal{L}_{\text{MDLM}} = -\sum_{i \in \mathcal{C}} \log p_\theta(x_i^{(0)} \mid x^{(t)})$$

  Update model parameters θ using gradient $\nabla_\theta \mathcal{L}_{\text{MDLM}}$
end for

Algorithm 2 Multi-Objective Discrete Simplex Refinement (DSR) for Nanobody Generation

Input: Trained MDLM $p_\theta(x^{(t-1)} \mid x^{(t)})$, reward models f_{bind}, f_{stab}, temperature β, number of tradeoffs L, number of seeds S, candidates per seed K, top-k output size
Generate tradeoff vectors $\{\lambda^{(\ell)}\}_{\ell=1}^{L}$ over Δ^2 (e.g., via Das-Dennis lattice)
Initialize empty list $\mathcal{C} \leftarrow \emptyset$
for $\ell = 1$ to L **do** ▷ Loop over tradeoff weights
  for $s = 1$ to S **do** ▷ Loop over independent seeds
    Initialize $x^{(T)} \sim$ MASKED_CDR_NOISE()
    for $t = T$ to 1 **do**
      Sample $\{x_k^{(t-1)}\}_{k=1}^{K} \sim p_\theta(x^{(t-1)} \mid x^{(t)})$
      for $k = 1$ to K **do**
        Evaluate rewards:

$$R_k = \lambda_{\text{bind}}^{(\ell)} f_{\text{bind}}(x_k^{(t-1)}) + \lambda_{\text{stab}}^{(\ell)} f_{\text{stab}}(x_k^{(t-1)})$$

      end for
      Compute softmax weights:

$$w_k = \frac{\exp(\beta R_k)}{\sum_{j=1}^{K} \exp(\beta R_j)}$$

      for each CDR position $i \in \mathcal{C}$ **do**
        for each token $a \in \Sigma$ **do**
          Update:

$$\tilde{\pi}_i(a) = \sum_{k=1}^{K} w_k \cdot \mathbb{1}\{x_{k,i}^{(t-1)} = a\}$$

        end for
      end for
      Sample $x^{(t-1)}$ from $\tilde{\pi}$ at each $i \in \mathcal{C}$
    end for
    Add final sequence $x^{(0)}$ and its reward to list $\mathcal{C}$
  end for
end for
Sort $\mathcal{C}$ by scalarized reward: $R(x) = \lambda^T [f_{\text{bind}}(x), f_{\text{stab}}(x)]$
Return: Top-k nanobody sequences from $\mathcal{C}$

B Discrete Simplex Refinement Proof

We provide a proof that Discrete Simplex Refinement (DSR) updates at each denoising step guarantee a monotonic improvement in expected scalarized reward.

Theorem B.1 (DSR Monotonic Improvement)**.** *Let $\pi^{(t)}$ be the uniform distribution over K samples $\{x_k\}_{k=1}^K$ drawn from the base model at timestep t, i.e., $\pi^{(t)}(x_k) = \frac{1}{K}$ for all k. Let $R_k := R^{(\lambda)}(x_k) \in \mathbb{R}$ be the corresponding scalarized rewards. Define the DSR-updated distribution as:*

$$\tilde{\pi}^{(t)}(x_k) := \frac{\exp(\beta R_k)}{\sum_{j=1}^K \exp(\beta R_j)} =: w_k,$$

where $\beta > 0$ is the temperature parameter. Then:

$$\mathbb{E}_{x\sim\tilde{\pi}^{(t)}}[R(x)] \geq \mathbb{E}_{x\sim\pi^{(t)}}[R(x)],$$

with strict inequality whenever the rewards $\{R_k\}$ are not all equal.

Proof. The expected reward under the original uniform distribution is:

$$\mathbb{E}_{x\sim\pi^{(t)}}[R(x)] = \sum_{k=1}^K \frac{1}{K} R_k = \frac{1}{K}\sum_{k=1}^K R_k.$$

The expected reward under the DSR-updated distribution is:

$$\mathbb{E}_{x\sim\tilde{\pi}^{(t)}}[R(x)] = \sum_{k=1}^K w_k R_k = \sum_{k=1}^K \frac{\exp(\beta R_k)}{\sum_{j=1}^K \exp(\beta R_j)} R_k.$$

To establish monotonic improvement, we show that:

$$\sum_{k=1}^K w_k R_k \geq \frac{1}{K}\sum_{k=1}^K R_k.$$

Case 1: If all rewards are equal, i.e., $R_k = c$ for some constant c and all k, then:

$$w_k = \frac{\exp(\beta c)}{\sum_{j=1}^K \exp(\beta c)} = \frac{\exp(\beta c)}{K\exp(\beta c)} = \frac{1}{K},$$

so $\tilde{\pi}^{(t)} = \pi^{(t)}$ and equality holds.

Case 2: If the rewards are not all equal, we use the following key lemma:

Lemma B.2 (Softmax Improvement Property)**.** *For any non-constant vector $(R_1, \ldots, R_K)$ and $\beta > 0$:*

$$\sum_{k=1}^K \frac{\exp(\beta R_k)}{\sum_{j=1}^K \exp(\beta R_j)} R_k > \frac{1}{K}\sum_{k=1}^K R_k.$$

Proof of Lemma. Define $f(\beta) = \sum_{k=1}^K \frac{\exp(\beta R_k)}{\sum_{j=1}^K \exp(\beta R_j)} R_k$. Note that:

- $f(0) = \frac{1}{K}\sum_{k=1}^K R_k$ (uniform weighting)
- $f'(\beta) = \mathrm{Var}_{\mathrm{softmax}_\beta}(R) \geq 0$, with strict inequality when R_k are not all equal

To see this, compute:

$$f'(\beta) = \sum_{k=1}^K w_k R_k^2 - \left(\sum_{k=1}^K w_k R_k\right)^2 = \mathbb{E}[R^2] - (\mathbb{E}[R])^2 = \mathrm{Var}(R),$$

where the expectation is taken with respect to the softmax distribution with temperature β^{-1}.

Since $\mathrm{Var}(R) > 0$ when R_k are not all equal, we have $f'(\beta) > 0$ for all $\beta > 0$. Therefore, $f(\beta) > f(0)$ for any $\beta > 0$. □

C Data Curation and Processing

C.1 Pre-training Dataset and Tokenization

The pre-training dataset comprises 3,395,594 unpaired heavy chain sequences curated by HuDiff [Ma et al., 2024], sourced from the OAS database, and aligned using the IMGT numbering scheme. All sequences were aligned to a length of 152 residues. For tokenization, we employed an amino acid tokenizer with a vocabulary of 23 tokens, comprising the 20 canonical amino acids, an 'X' token for unknown residues, a '-' token for padding, and a '<mask>' token for masked positions.

C.2 Benchmark Dataset

The experimental antigens and their paired reference nanobodies were obtained from SAbDab-nano [Schneider et al., 2022]. After filtering for *Homo Sapiens* species, a resolution cutoff of 3.0 Å, and complexes containing only two molecules, the initial pool consisted of 35 complex structures. Manual screening was then performed to remove redundant antigens with identical sequences, yielding 25 distinct antigen sequences used for unconditional sequence generation screening in Figure A1.

C.3 Humanized Scaffold for Sampling Inference

During inference, we employed a universal humanized nanobody scaffold [Vincke et al., 2009] for the framework regions, while sampling the CDRs from our model. To construct the input sequence, the h-NbBcII10$_{FGLA}$ sequence was first aligned to a fixed length of 152 residues using the IMGT numbering scheme, consistent with our training sequences. The CDR regions were then substituted with mask tokens for model inference.

D NanoMDLM Implementation Details

D.1 Region-Dependent Masking

Building on state-dependent [Shi et al., 2024] and bond-dependent masking for peptides [Tang et al., 2025], we propose a masking schedule where the probability of masking tokens in CDR1 and CDR2 increases more slowly at early timesteps t compared to CDR3 tokens. This design reflects the more conserved nature of CDR1 and CDR2, while allowing greater diversity in CDR3 sequences (as shown in Figure A2). We define the discrete-time log-linear masking schedule $\sigma(t) = -\log(1 - t^{1/\gamma_{\text{weak}}})$ for CDR1 and CDR2, and the log-polynomial schedule $\sigma(t) = -\log(1 - t^{1/\gamma_{\text{strong}}})$ for CDR3.

The masking probabilities $\alpha_t = exp(-\sigma(t))$ for the CDR tokens are thus:

$$\alpha_t^{\gamma(i)} = \begin{cases} 1 - t^{1/\gamma_{\text{weak}}} & \text{if } i \in \text{CDR1 or CDR2}, \\ 1 - t^{1/\gamma_{\text{strong}}} & \text{if } i \in \text{CDR3}, \end{cases} \qquad \text{with } \gamma_{\text{strong}} > \gamma_{\text{weak}}.$$

In our implementation, we set $\gamma_{\text{weak}} = 1/3$ and $\gamma_{\text{strong}} = 1$.

For the forward corruption process $q(x^{(t)}|x^{(t-1)})$, this approach ensures a slower masking rate in CDR1 and CDR2 at earlier stages of the diffusion process, while CDR3 are masked more quickly with a stronger exponent γ_{strong}. This enables the model to first focus

D.2 Training Strategy and Model Architecture

Pretraining of NanoMDLM for nanobodies was conducted on 4 NVIDIA DGX B200 GPUs, using the AdamW optimizer with a learning rate of 0.0003 and weight decay of 0.075. At each reverse step, token probabilities are modeled using a RoFormer network [Su et al., 2024]. The full set of model hyperparameters is listed below.

Table 2: Roformer Architecture Hyperparameters

Hyperparameter	PepTune
Input Dimension	23 (vocab size)
Hidden Dimension	768
Intermediate Dimension	3072
Number of Layers	8
Attention Heads	8
Max Positional Embeddings	256
Hidden and Attention Dropout Probability	0.1

Table 3: Training and Validation Loss of NanoMDLM. Loss values are taken after convergence at 45 epochs when training NanoMDLM on 3 million unpaired human heavy chain sequences.

Model	Train Loss (↓)	Train PPL (↓)	Val Loss (↓)	Val PPL (↓)
Standard Masking	0.661	2.025	0.676	1.988
Region-Dependent Masking	0.600	1.983	0.642	1.948

E Additional Studies on Sampling

E.1 Unconditional Generation Quality

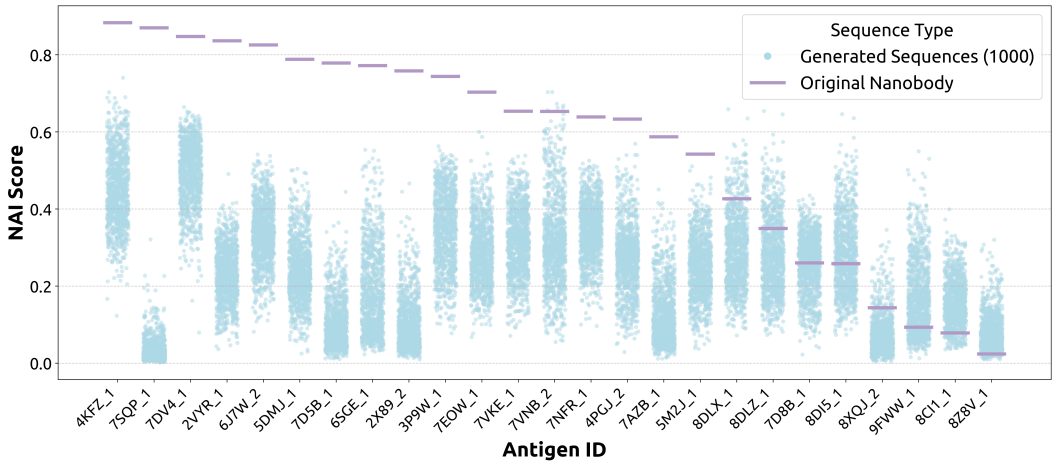

Figure A1: Interaction scores with human antigens of experimental and unconditionally generated nanobody sequences. Experimental human antigens and nanobody sequences were obtained from SAbDab-nano. After filtering for one-to-one interaction complexes and applying redundancy reduction, 25 distinct antigens were selected. Interaction scores were predicted using DeepNano. For each antigen, the score corresponding to its original experimental nanobody is shown as a purple line. Scores for 1,000 unconditionally generated nanobody sequences from the base MDLM are represented as skyblue dots.

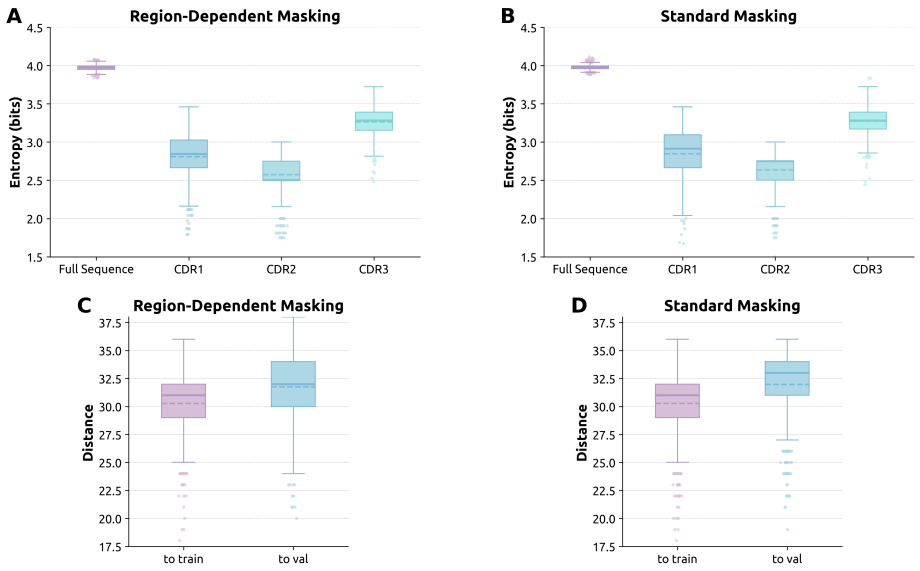

Figure A2: Unconditional-generated nanobody sequence analysis. A-B: Shannon entropy of the unconditionally generated sequences (N=1000); C-D: Minimum Levenshtein distance of the unconditionally generated sequences (N=1000)

To evaluate sequence diversity, we calculated Shannon entropy and Jaccard similarity. Shannon entropy $H(X) = -\sum p(x) \log_2 p(x)$ measures the variability of the amino acid distribution, with higher values indicating greater diversity. As shown in Figure A2 (Panels A-B), region-dependent masking results in slightly higher entropy across all CDR regions, suggesting it promotes greater sequence diversity, particularly in CDR3.

Jaccard similarity $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$ assesses the overlap with reference sequences, with higher values indicating better fidelity. Table 4 shows that standard masking yields higher Jaccard similarity to the training and validation sequences, suggesting a stronger resemblance to the existing database. In contrast, region-specific masking slightly reduces similarity but increases diversity.

Table 4: Mean pairwise Jaccard similarity for unconditionally generated nanobody sequence (N=1000)

Masking Method	**Full Sequence**	**CDR1**	**CDR2**	**CDR3**
Standard Masking	0.980503	0.529960	0.511030	0.580592
Region-Specific Masking	0.977800	0.463944	0.418981	0.528402

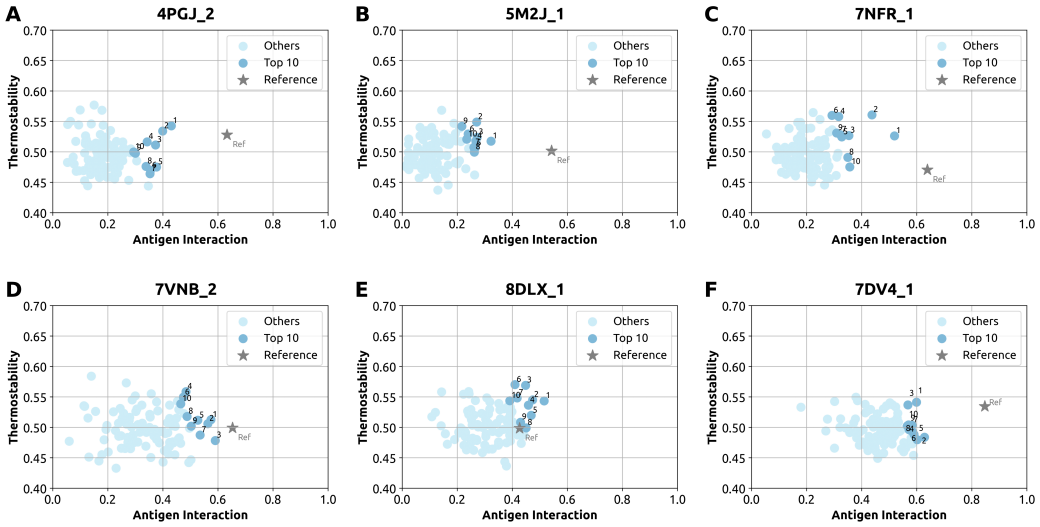

Figure A3: Scores of unconditionally generated nanobodies with selected antigens. The antigens in A-E are with experimental nanobodies that exhibit moderate interaction scores (0.4-0.65).

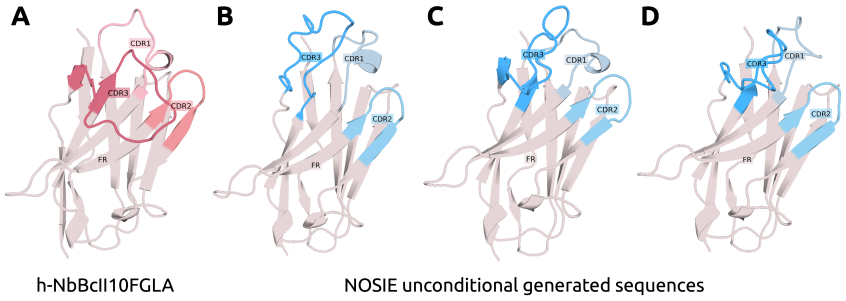

Figure A4: NanoNet-predicted structures of NbBCII10 humanized (FGLA mutant) and three NanoMDLM unconditional generated nanobodies

E.2 DSR Generation Quality

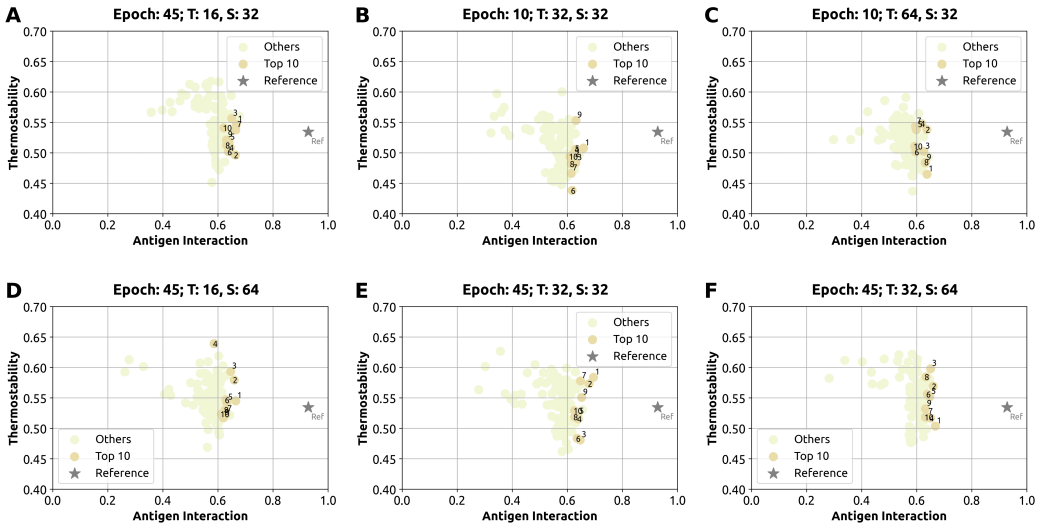

Figure A5: Hyperparameter search for DSR sampling conducted on antigen CTLA-4 (PDB:7DV4_1), topK=100.

In this analysis, we conducted a hyperparameter search for DSR sampling on antigen CTLA-4 [Gan et al., 2022]. Each subplot compares different combinations of training epochs, sampling timesteps (T), and sampling seeds (S), showing the antigen interaction and thermostability scores of the top 100 sequences generated in each configuration. Panels B–C correspond to the base model trained for only 10 epochs. As observed, while the antigen interaction scores exhibit moderate variation, these shorter training runs result in noticeably lower thermostability scores. This suggests that longer training allows the base model to better capture the underlying biological properties of the CDRs. In contrast, the results from the epoch-45 base model (panels A, D–F) highlight that intermediate sampling timesteps and number of seeds—particularly in panels E—achieve a favorable balance, producing high-quality candidates while avoiding excessive variance.

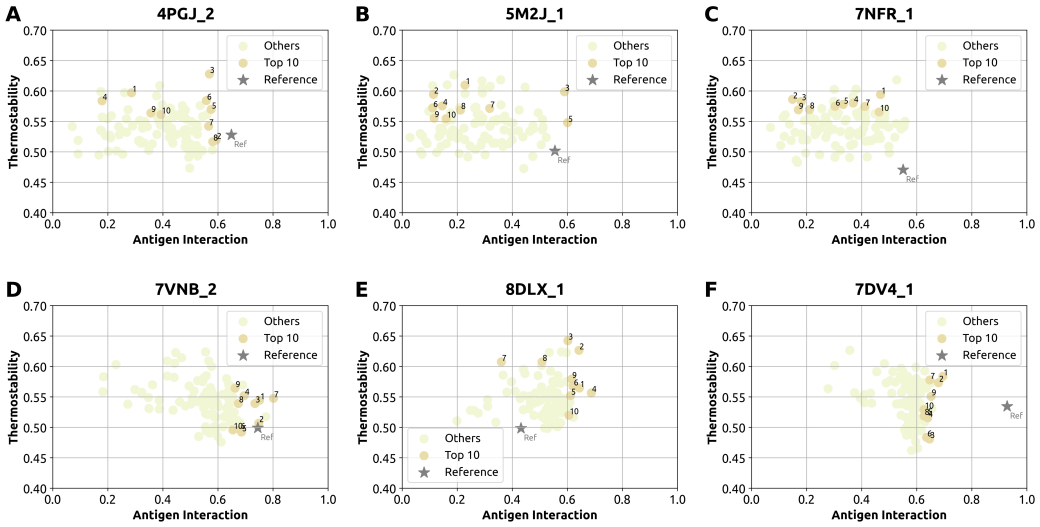

Figure A6: Scores of DSR-generated nanobodies with selected antigens. The antigens in A-E are with experimental nanobodies that exhibit moderate interaction scores (0.4-0.65).

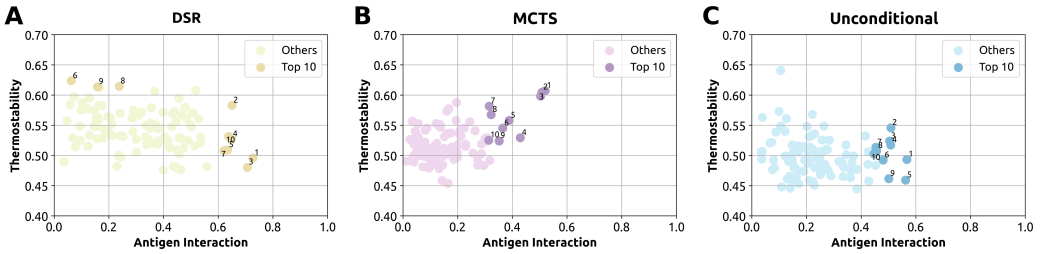

Figure A7: Comparison of generated nanobody scores with antigen MRGPRX2, topK=100. All three sampling methods were run with 32 sampling timesteps. There are no reference scores for this antigen as known nanobody-MRGPRX2 pair was not found in PDB.

E.3 MCTS Generation Quality

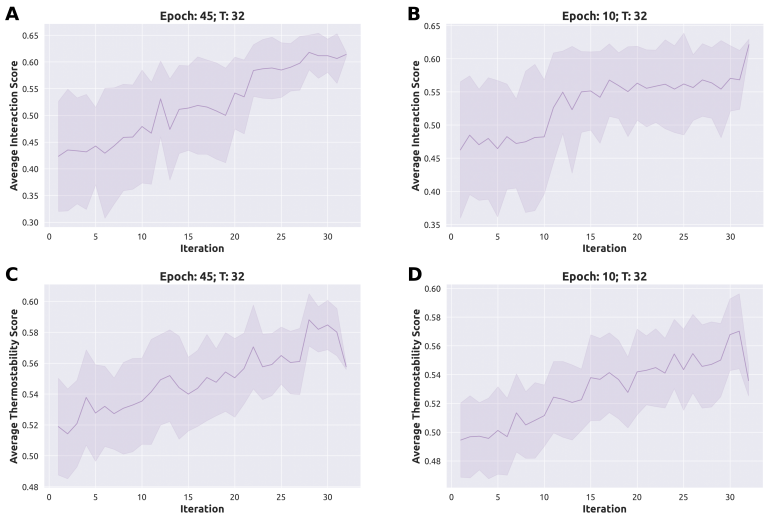

Figure A8: MCTS property scores over iteration on two epochs version of the base MDLM model

Figure A8 presents the progression of antigen interaction and thermostability scores among MCTS iterations for the two base MDLM models trained for 45 or 10 epochs. Consistent with previous observations in DSR sampling, panels B and D (10-epoch models) show that while antigen interaction scores improve with iterations, the thermostability scores remain comparatively lower than those of the 45-epoch models (panels A and C). This echoes the earlier finding that sufficient training until convergence enables the model to better capture the underlying biological properties of CDRs, leading to more stable and higher-quality sequence generation.

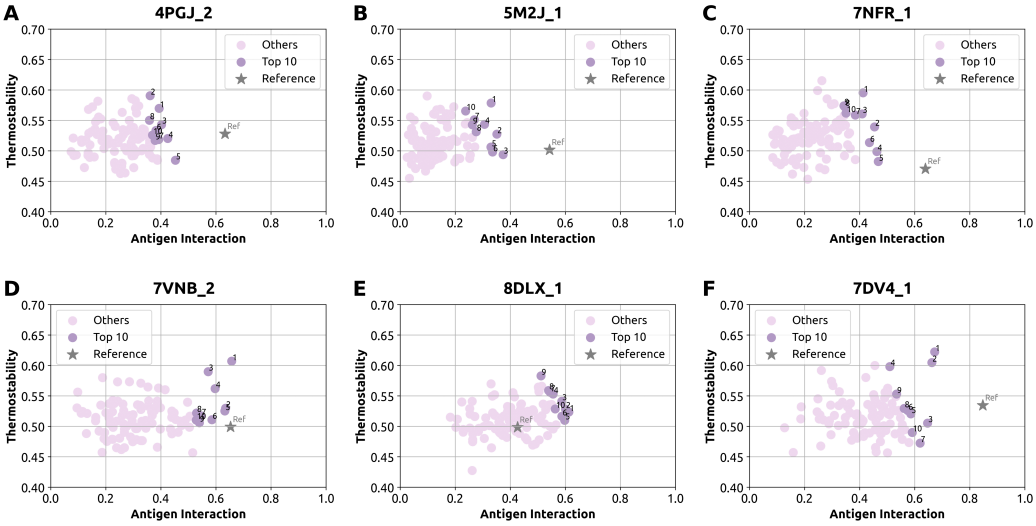

Figure A9: Scores of MCTS-generated nanobodies with antigens (topK=100). The antigens shown here are the same as in A6. The ones in A-E are with a experimental nanobodies that exhibit moderate interaction scores (0.4-0.65).