ANTIBODY DESIGN USING PREFERENCE OPTIMIZATION AND STRUCTURAL INFERENCE

Archit Vasan, Gautham Dharuman, Ozan Gokdemir, Matthew Sinclair, Xinran Lian, Moeen Meigooni, Heng Ma, Carla Mann & Arvind Ramanathan
Argonne National Laboratory
Lemont, IL 60439, USA

ABSTRACT

Antibodies offer several advantages in therapeutic design, including high specificity to targets, reduced off-target effects, immune system engagement, and the ability to bind traditionally undruggable proteins. To harness these benefits, we propose an antibody design method that integrates large language models (LLMs), preference optimization, diffusion modeling, and molecular dynamics simulations. Our approach begins by fine-tuning an LLM on complementarity-determining region (CDR) sequences, generating new CDR sequences, and folding antibodies with antigen scaffolds. We then apply diffusion models to refine CDR backbones, followed by inverse folding to generate new amino acid sequences. These redesigned antibodies undergo molecular dynamics simulations to evaluate binding affinity, and preference data is used to iteratively improve the LLM through direct preference optimization. This method has been applied to lysozyme, where it produced antibodies with greater predicted binding affinity than native counterparts. Future directions include extending this approach to antigens that adopt multiple conformations and experimentally validating the designed antibodies. Ultimately, this framework leverages artificial intelligence and high-performance computing to accelerate the discovery of clinically relevant antibody candidates.

1 INTRODUCTION

Antibodies are effective therapeutics that offer high specificity and low toxicity, making them ideal for treating cancer (Kumar et al., 2024; Sliwkowski & Mellman, 2013) and autoimmune diseases (Du et al., 2017; Lu et al., 2020). Furthermore, they have a longer half-life compared to small molecules, enabling extended circulation times and reducing the need for frequent dosing. Moreover, antibodies possess versatile mechanisms of antigen neutralization, including direct binding, blocking receptor-ligand interactions, and inducing receptor internalization. Finally, they can effectively bind to otherwise undruggable extracellular receptors on large and complex protein structures. In light of this non-exhaustive list of advantages, antibodies represent a promising direction for successful therapeutic design campaigns. Traditionally, antibody design has relied on experimental methods such as hybridoma technology (KÖHLER & MILSTEIN, 1975), phage display (Smith, 1985; Winter et al., 1994), yeast display (Gai & Wittrup, 2007), and directed evolution (Arnold, 1996; Wang et al., 2021; Amon et al., 2020). Informed by experimental validation, these methods while reliable can be time-consuming, expensive and error-prone, limiting the screening capacity.

The integration of machine learning (ML) relies heavily on high-throughput in-silico validations to rapidly iterate and refine antibody candidates. Earlier efforts have utilized ML to predict various molecular and physicochemical properties, serving as faster surrogates to traditional experimental pipelines. More recently, generative ML techniques such as Generative Adversarial Networks (GANs) (Amimeur et al., 2020), Variational Autoencoders (VAEs) (Eguchi et al., 2022), Large Language Models (LLMs) (Melnyk et al., 2023; Barton et al., 2024; Zvyagin et al., 2022), and Denoising Probabilistic Diffusion Models (He et al., 2024) have been employed to generate innovative antibody sequences and structures. While these methods are scalable and cost-effective, integrating them with sparse experimental data remains a significant challenge. This often results in a bottle-

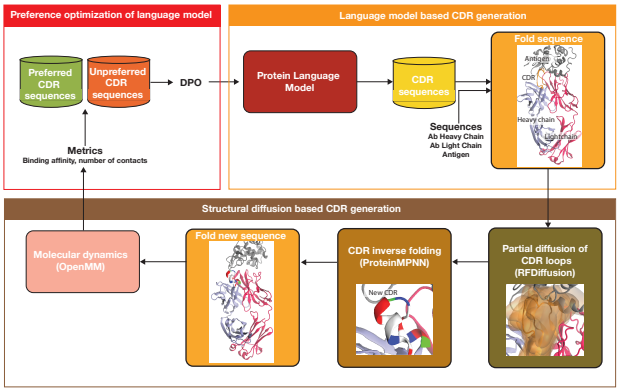

Figure 1: **Antibody generation workflow** Our workflow generates antibodies via a combination of LLM inference, protein folding, partial diffusion, inverse folding, direct preference optimization and molecular dynamics simulation.

neck due to the limited resources available for experimentally validating synthetic designs and the inherent limitations of the models themselves.

In response to existing challenges, we propose a novel, experiment-in-the-loop pipeline that iteratively refines a protein language model to enhance antibody design capabilities. Our pipeline optimizes antibodies by leveraging direct preference optimization (DPO), a replacement to reinforcement learning. DPO increases the probability of sequences similar to preferred sequences instead of directly optimizing metrics, improving multi-objective optimization. Recently, DPO has been applied to miniprotein binder and small molecule optimization; however, these approaches have not incorporated structural diffusion derived sequences into the preference data (Mistani & Mysore, 2024; Dharuman et al., 2024; Park et al., 2023).

Our pipeline first generates multiple CDR sequences from a protein language model finetuned on naturally occuring complementarity-determining region (CDR) sequences and grafts these CDRs into an antibody scaffold of interest. We then predict the antibody-antigen complex 3D structure and use the structure to initialize partial diffusion and inverse folding of CDR loops. The diffusion modeling increases sampling of novel CDR loops that may have been missed by the PLM. This is followed by molecular dynamics simulations of each complex to assess binding affinities. These simulations serve as preference signals in our direct preference optimization (DPO) loop in which the PLM is optimized to generate effective designs. This method not only accelerates the antibody discovery process but also expands the scope of designable antibodies, enabling access to novel sequences and structures that are difficult to achieve through experimental methods alone. By optimizing the generation of preferred CDRs, our framework embodies a closed-loop discovery process that significantly enhances the speed and breadth of antibody development.

2 METHODS

This methodology intends to transform an arbitrary antibody to recognize a particular antigen of interest. Our workflow for generating novel antibodies involves six general steps (Fig 1): (1) LLM inference to generate antibody CDR loop sequences, (2) folding model inference to predict a structural model from antibody heavy/light and antigen sequences, (3) diffusion model inference to generate new backbone structures, (4) inverse folding to predict new sequences based on backbone structures, (5) molecular dynamics simulations to evaluate energetics of antibody-antigen interactions, (6) direct preference optimization to increase likelihood of LLM generating favorable antibody CDR loop sequences.

2.1 LLM INFERENCE TO GENERATE ANTIBODY CDR LOOP SEQUENCES

To generate biologically relevant CDR loop sequences, the GPT-NEOX model that serves as the basis for GenSLM Zvyagin et al. (2022) is finetuned on a dataset of 670 CDRH3 loop sequences Vita et al. (2015). We divide the dataset into training, validation, and test sets with an 8:1:1 split. Then, the model is finetuned on the training set with 10 epochs. Following this, the finetuned GPT-NEOX

model can be used to autoregressively generate CDR loop sequences. A max length of 26 and minimum length of 8 is employed during the generation to approximate the size of a CDRH3 loop. Each generated CDR loop is then grafted onto a reference antibody FAB sequence.

2.2 FOLDING SEQUENCES

Then, we take the antibody sequence with the newly grafted CDR loop and input this into the folding module. We used the state-of-the-art model, CHAI-1 for folding the heavy and light chains of the antibody in complex with the antigen. CHAI-1 has previously shown success on folding antibodies (Escarra-Senmarti et al., 2025); on a subset of 268 interfaces across 129 structures, CHAI-1 outperformed AF2.3 significantly (DockQ < 0.23). In our method, we ran CHAI-1 inference directly using embeddings from the ESM model without incorporating full MSAs. It has previously been shown on antibody benchmarks that when omitting MSAs, similar performance was achieved Discovery et al. (2024).

2.3 STRUCTURAL INFERENCE OF NEW CDR LOOPS

Then, we use the CDR fold as a basis for RFDiffusion to generate new CDR backbones Watson et al. (2023). RFDiffusion is an addendum to generative LLM which incorporates structural context into the pipeline. Diffusion is performed only on the CDR loop of interest while keeping the antibody scaffold and antigen fixed. To condition structural generation, the folded CDR from the previous step is only partially noised for 10 steps before sampling new backbones.

RFDiffusion is an example of a denoising diffusion probabilistic model which aim to approximate a distribution by reversing a discrete diffusion process Watson et al. (2023). In forward diffusion, x^0 is sampled out of the target data distribution P^0 and noised in T steps towards a final $x^T \in P^T$ that is not dependent on P^0 Ho et al. (2020). Then, the real distribution P is approximated by the distribution $\hat{P}$ which is parameterized by a transition kernel $\hat{p}(x^{t-1}|x^t)$ at each timestep, t. RFDiffusion uses a trained neural network to parameterize each transition kernel. The denoising process works by (1) sampling x^T from the reference distribution P^T, and (2) at each timestep t sampling new x^{t-1} from $\hat{p}(x^{t-1}|x^t)$ until reaching $x^0 \sim P^0$ Ho et al. (2020).

Once new backbones are sampled, we apply inverse folding on the CDR loops using the ProteinMPNN model Dauparas et al. (2022). This model contains two components: (1) an encoder layer that embeds backbone coordinates via a graph neural network (GNN), and (2) a decoder layer that conditions generation of sequences on the GNN embeddings. The new CDR loop sequences from inverse folding are then grafted onto the antibody scaffold sequence and the sequence is folded using the CHAI-1 model.

2.4 MOLECULAR DYNAMICS SIMULATIONS

To verify stability for the antibody/antigen complex structures, we build and simulate the complexes in implicit solvent using OpenMM. Each complex is built in the GBn2 solvation model in the AMBER forcefield, with solute and solvent dielectrics of 1.0 and 80.0, respectively with a Debye-Huckel screening parameter of 1.0/nm. Then each complex is energy minimized and simulated in an NVT ensemble using Langevin dynamics with a temperature of 300K, friction coefficient of 1/ps, and a timestep of 4 fs. Hydrogen bonds are constrained during the simulations.

2.5 DIRECT PREFERENCE OPTIMIZATION OF CDR LOOPS

Then, we finetune the LLM generating CDR loops using direct preference optimization (DPO) to improve the binding energetics of generated antibody structures.

To apply DPO, we input a list of preferred and unpreferred CDR sequences according to a metric. Our metric is the energy weighted contacts (W_i) between the antibody and antigen. A contact is defined as an antibody residue C_α that comes within 8 Åof an antigen residue C_α. We apply a Boltzmann weight towards lower energy interactions that typically define the stability of a complex. The interaction energy is calculated between the antibody and antigen within the same implicit solvent conditions as during simulation using OpenMM. We determine the energy of the two components

together (E_{AB}), the energy of each component individually (E_A, E_B), and subtract these terms to get the interaction energy ($E_{int} = E_{AB} - (E_A + E_B)$

With this metric, we divide the CDR loop datasets into preferred and unpreferred according to the following criteria:

$$\mathcal{D}^+ = \{s_i | W_i(s_i) > W_{cut}\}_{s_i \in S}$$
$$\mathcal{D}^- = \{s_i | W_i(s_i) < W_{cut}\}_{s_i \in S},$$

where $\mathcal{D}^+$,$\mathcal{D}^- \subset \mathcal{D}$, and W_{cut} is the cutoff weighted contact number. Then, we use $\mathcal{D}^+$ and $\mathcal{D}^-$ in our DPO loop to align the LLM.

DPO simplifies model alignment compared to Proximal Policy Optimization (PPO) methods by not requiring a trained reward model Rafailov et al. (2024). Instead, DPO directly updates the preference policy directly using preference information. During alignment, the preference policy is optimized to fit the preference data using a binary cross entropy objective Rafailov et al. (2024). Although DPO removes the reward model directly from LLM alignment, an implicit reward function is still used to define the preference data. Thus, it is essential to choose a sensible metric to divide the dataset.

First, DPO defines a fixed reference model,π_{ref} and a policy model, π_θ that is aligned during DPO. Both π_{ref} and π_θ output tokens with probability ($\hat{P}$) given by:

$$\hat{P}(a_t | a_{<t}; \epsilon) = \pi_\epsilon(a_{<t}),$$

where ϵ is the parameters of π, a_t is the token at position t, and $a_{<t}$ are all tokens before t.

Then π_{ref} and π_θ are given the preferred and unpreferred samples to align the π_θ model. A loss function is defined incorporating the log-probabilities of the preferred and unpreferred samples after being input to π_{ref} and π_θ:

$$\mathcal{L}_{DPO}(\pi_\theta; \pi_{ref}) = -\mathbb{E}_{(r,s^+,s^-)\sim\mathcal{D}}[\log \sigma(\beta \log \frac{\pi_\theta(s^+|r)}{\pi_{ref}(s^+|r)} - \beta \log \frac{\pi_\theta(s^-|r)}{\pi_{ref}(s^-|r)})]. \quad (1)$$

Here, s^+ and s^- define the positive and negative CDR sequences, r is the prompt sequence inputted into the model (e.g. a single amino acid), and β is a scaling factor. For the antibody design problem, model alignment should promote strong antibody-antigen binding.

3 RESULTS

We apply our workflow to engineer a strong antibody-antigen interface using an antibody-antigen pair not known to natively bind by only altering the sequence of its CDR loops. The operative antibody in our experiment is the Broadly Neutralizing Fab PGT122 targeting HIV Padlan et al. (1989), which we engineer to target lysozyme, a common enzyme commonly found in secretions. We benchmark antibody-antigen energetics for our designs to the antibody-lysozyme interface found in the stable crystal structure (3HFM) Padlan et al. (1989). PGT122 is not known to natively target lysozyme.

We generate 914 CDR sequences using RFDiffusion and 684 sequences from our aligned language model. We find the CDRs generated by RFDiffusion to be significantly enhanced compared to via the initial LLM inference (Fig 2A). LLM inference is biased by the data used to train and finetune the model; thus, the inferred sequences would carry that inherent bias. Compared to LLMs, the diffusion/inverse folding models tend to produce structurally and sequentially more diverse outputs, as they directly sample from a continuous structural space and map back to sequence space as opposed to using autoregressoin. To understand the impact on optimizing antibody-antigen binding, we simulate each complex in implicit solvent and calculate the antibody-antigen interaction energy.

We find most generated structures have improved Lennard Jones energy than the 3HFM antibody-lysozyme complex. The heavy chain CDR3 loop in PGT122 is significantly more extended than the native lysozyme antibody, increasing accessible surface area and promoting favorable nonpolar interactions. In contrast, few generated structures improve electrostatic interaction energy between the antibody and antigen as opposed to the native complex.

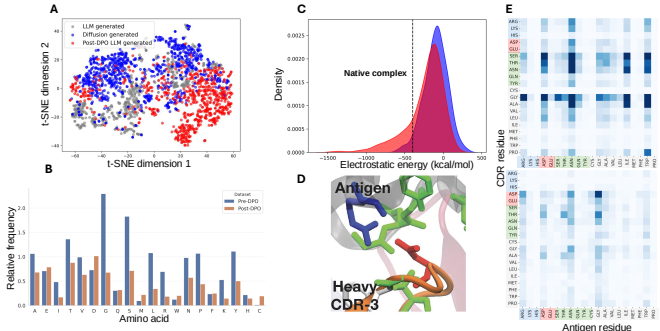

Figure 2: **DPO optimizes antibody-antigen binding** A) ESM embeddings for LLM, diffusion, and post-DPO CDR sequences, projected onto 2 t-SNE components. B) Relative frequency of each amino acid in diffusion-generated (blue) and post-DPO generated CDR sequences (orange). C) Distribution of electrostatic interaction energies for diffusion generated (blue) and post-DPO (red) antibody/antigen complexes. Interaction energies calculated in implicit solvent (dielectric constant = 1.0 and 80.0 for solute and solvent, respectively). Dashed line indicates the energy for the native lysozyme-antibody complex (PDB ID: 3HFM). D) Snapshot of strong antibody/antigen binding. Red, blue, and green ball and stick representations indicate acidic, basic, and polar residues, respectively. E) contact map between CDR and antigen residues for diffusion generated (top) and post-DPO (bottom) simulations (contact: 8 Å C-α to C-α distance)

Strikingly, electrostatic energy of structures with post-DPO generated CDR loops have a significantly left-shifted distribution indicating favorable electrostatic interactions (Fig 2C). Moreover, the CDR loops post-DPO show a greater frequency of acidic residues than pre-DPO (Fig 2E). In particular an increased frequency of aspartate/glutamate-arginine interactions help stabilize the CDR loop-antigen interaction. Interestingly, prior to DPO, the contact matrix of CDR and antigen residues shows more high frequency interactions(Fig 2E) involving serine and glycine for the CDR loop, possibly due to inherent bias in the inverse folding model (Fig 2B). While, post-DPO CDR sequences have fewer overall antigen interactions, each interaction is more energetically favorable. It is important to note that while the left tail is larger for the post-DPO distribution vs pre-DPO, the peak is still at a higher energy than the native complex from 3HFM. In the native 3HFM heavy chain, the electrostatic interactions are predominantly formed by CDR1 and CDR2. We chose to optimize the PGT122 CDR3 since this was the longest segment; however, to improve the energetics signicantly, it would be beneficial to optimize multiple CDR loops.

4 Conclusion

Our distinct approach, integrates DPO, structural diffusion modeling, and molecular simulation to engineer antibody CDR loops. We demonstrate binding affinity optimization for the Fab PGT122 antibody for a non-native antigen, lysozyme. We find our method is capable of performing optimization in the low data regime (1000 samples) using an autoregressive LLM with few parameters (25 million). The diffusion/inference models operating in structural space allows a larger variety of generated sequences that should still bind the antigen. The usage of partial diffusion allows us to condition structure generation on the LLM-generated CDR sequences. Overall, the workflow builds on the work performed on protein and small molecule

In the future, we would like to extend this workflow beyond antibody-antigen binding optimization. This includes optimizing developability such as mitigating the risk of low thermal stability or high aggregability (Raybould et al., 2019). Additionally, the approach could be generalized to the lead optimization platforms for minibinders, small molecules and PROTACS. Ultimately, we would want to incorporate *in vitro* and *in vivo* experiments into the loop to generate validated preferred/unpreferred pairs for DPO finetuning.

References

Tileli Amimeur, Jeremy M. Shaver, Randal R. Ketchem, J. Alex Taylor, Rutilio H. Clark, Josh Smith, Danielle Van Citters, Christine C. Siska, Pauline Smidt, Megan Sprague, Bruce A. Kerwin, and

Dean Pettit. Designing feature-controlled humanoid antibody discovery libraries using generative adversarial networks. April 2020. doi: 10.1101/2020.04.12.024844. URL http://dx.doi.org/10.1101/2020.04.12.024844.

Ron Amon, Ronit Rosenfeld, Shahar Perlmutter, Oliver C. Grant, Sharon Yehuda, Aliza Borenstein-Katz, Ron Alcalay, Tal Marshanski, Hai Yu, Ron Diskin, Robert J. Woods, Xi Chen, and Vered Padler-Karavani. Directed evolution of therapeutic antibodies targeting glycosylation in cancer. *Cancers*, 12(10):2824, September 2020. ISSN 2072-6694. doi: 10.3390/cancers12102824. URL http://dx.doi.org/10.3390/cancers12102824.

Frances H. Arnold. Directed evolution: Creating biocatalysts for the future. *Chemical Engineering Science*, 51(23):5091–5102, December 1996. ISSN 0009-2509. doi: 10.1016/s0009-2509(96)00288-6. URL http://dx.doi.org/10.1016/S0009-2509(96)00288-6.

Justin Barton, Aretas Gaspariunas, David A. Yadin, Jorge Dias, Francesca L. Nice, Danielle H. Minns, Olivia Snudden, Chelsea Povall, Sara Valle Tomas, Harry Dobson, James H. R. Farmery, Jinwoo Leem, and Jacob D. Galson. A generative foundation model for antibody sequence understanding. May 2024. doi: 10.1101/2024.05.22.594943. URL http://dx.doi.org/10.1101/2024.05.22.594943.

Justas Dauparas, Ivan Anishchenko, Nathaniel Bennett, Hua Bai, Robert J Ragotte, Lukas F Milles, Basile IM Wicky, Alexis Courbet, Rob J de Haas, Neville Bethel, et al. Robust deep learning–based protein sequence design using proteinmpnn. *Science*, 378(6615):49–56, 2022.

Gautham Dharuman, Kyle Hippe, Alexander Brace, Sam Foreman, Väinö Hatanpää, Varuni K Sastry, Huihuo Zheng, Logan Ward, Servesh Muralidharan, Archit Vasan, et al. Mprot-dpo: Breaking the exaflops barrier for multimodal protein design workflows with direct preference optimization. In *SC24: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–13. IEEE, 2024.

Chai Discovery, Jacques Boitreaud, Jack Dent, Matthew McPartlon, Joshua Meier, Vinicius Reis, Alex Rogozhnikov, and Kevin Wu. Chai-1: Decoding the molecular interactions of life. *bioRxiv*, pp. 2024–10, 2024.

Fanny Huynh Du, Elizabeth A. Mills, and Yang Mao-Draayer. Next-generation anti-cd20 monoclonal antibodies in autoimmune disease treatment. *Autoimmunity Highlights*, 8(1), November 2017. ISSN 2038-3274. doi: 10.1007/s13317-017-0100-y. URL http://dx.doi.org/10.1007/s13317-017-0100-y.

Raphael R. Eguchi, Christian A. Choe, and Po-Ssu Huang. Ig-vae: Generative modeling of protein structure by direct 3d coordinate generation. *PLOS Computational Biology*, 18(6):e1010271, June 2022. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1010271. URL http://dx.doi.org/10.1371/journal.pcbi.1010271.

Marta Escarra-Senmarti, Michael Chungyoun, Dylan Ferris, Jeffrey J Gray, and Felipe Andrade. Anti-citrullinated protein antibodies arise during affinity maturation of germline antibodies to carbamylated proteins in rheumatoid arthritis. *bioRxiv*, pp. 2025–03, 2025.

S Annie Gai and K Dane Wittrup. Yeast surface display for protein engineering and characterization. *Current Opinion in Structural Biology*, 17(4):467–473, August 2007. ISSN 0959-440X. doi: 10.1016/j.sbi.2007.08.012. URL http://dx.doi.org/10.1016/j.sbi.2007.08.012.

Xin-heng He, Jun-rui Li, James Xu, Hong Shan, Shi-yi Shen, Si-han Gao, and H. Eric Xu. Ai-driven antibody design with generative diffusion models: current insights and future directions. *Acta Pharmacologica Sinica*, September 2024. ISSN 1745-7254. doi: 10.1038/s41401-024-01380-y. URL http://dx.doi.org/10.1038/s41401-024-01380-y.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

G. KÖHLER and C. MILSTEIN. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256(5517):495–497, August 1975. ISSN 1476-4687. doi: 10.1038/256495a0. URL http://dx.doi.org/10.1038/256495a0.

Mukesh Kumar, Akansha Jalota, Sushil Kumar Sahu, and Shabirul Haque. Therapeutic antibodies for the prevention and treatment of cancer. *Journal of Biomedical Science*, 31(1), January 2024. ISSN 1423-0127. doi: 10.1186/s12929-024-00996-w. URL http://dx.doi.org/10.1186/s12929-024-00996-w.

Ruei-Min Lu, Yu-Chyi Hwang, I-Ju Liu, Chi-Chiu Lee, Han-Zen Tsai, Hsin-Jung Li, and Han-Chung Wu. Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science*, 27(1), January 2020. ISSN 1423-0127. doi: 10.1186/s12929-019-0592-z. URL http://dx.doi.org/10.1186/s12929-019-0592-z.

Igor Melnyk, Vijil Chenthamarakshan, Pin-Yu Chen, Payel Das, Amit Dhurandhar, Inkit Padhi, and Devleena Das. Reprogramming pretrained language models for antibody sequence infilling, 2023. URL https://arxiv.org/abs/2210.07144.

Pouria Mistani and Venkatesh Mysore. Preference optimization of protein language models as a multi-objective binder design paradigm. *arXiv preprint arXiv:2403.04187*, 2024.

Eduardo A Padlan, Enid W Silverton, Steven Sheriff, Gerson H Cohen, Sandra J Smith-Gill, and David R Davies. Structure of an antibody-antigen complex: crystal structure of the hyhel-10 fab-lysozyme complex. *Proceedings of the National Academy of Sciences*, 86(15):5938–5942, 1989.

Ryan Park, Ryan Theisen, Navriti Sahni, Marcel Patek, Anna Cichońska, and Rayees Rahman. Preference optimization for molecular language models. *arXiv preprint arXiv:2310.12304*, 2023.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36, 2024.

Matthew IJ Raybould, Claire Marks, Konrad Krawczyk, Bruck Taddese, Jaroslaw Nowak, Alan P Lewis, Alexander Bujotzek, Jiye Shi, and Charlotte M Deane. Five computational developability guidelines for therapeutic antibody profiling. *Proceedings of the National Academy of Sciences*, 116(10):4025–4030, 2019.

Mark X. Sliwkowski and Ira Mellman. Antibody therapeutics in cancer. *Science*, 341(6151): 1192–1198, September 2013. ISSN 1095-9203. doi: 10.1126/science.1241145. URL http://dx.doi.org/10.1126/science.1241145.

George P. Smith. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science*, 228(4705):1315–1317, June 1985. ISSN 1095-9203. doi: 10.1126/science.4001944. URL http://dx.doi.org/10.1126/science.4001944.

Randi Vita, James A Overton, Jason A Greenbaum, Julia Ponomarenko, Jason D Clark, Jason R Cantrell, Daniel K Wheeler, Joseph L Gabbard, Deborah Hix, Alessandro Sette, et al. The immune epitope database (iedb) 3.0. *Nucleic acids research*, 43(D1):D405–D412, 2015.

Yajie Wang, Pu Xue, Mingfeng Cao, Tianhao Yu, Stephan T. Lane, and Huimin Zhao. Directed evolution: Methodologies and applications. *Chemical Reviews*, 121(20):12384–12444, July 2021. ISSN 1520-6890. doi: 10.1021/acs.chemrev.1c00260. URL http://dx.doi.org/10.1021/acs.chemrev.1c00260.

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.

Greg Winter, Andrew D. Griffiths, Robert E. Hawkins, and Hennie R. Hoogenboom. Making antibodies by phage display technology. *Annual Review of Immunology*, 12(1):433–455, April 1994. ISSN 1545-3278. doi: 10.1146/annurev.iy.12.040194.002245. URL http://dx.doi.org/10.1146/annurev.iy.12.040194.002245.

Maxim Zvyagin, Alexander Brace, Kyle Hippe, Yuntian Deng, Bin Zhang, Cindy Orozco Bohorquez, Austin Clyde, Bharat Kale, Danilo Perez-Rivera, Heng Ma, Carla M. Mann, Michael Irvin, J. Gregory Pauloski, Logan Ward, Valerie Hayot-Sasson, Murali Emani, Sam Foreman, Zhen Xie, Diangen Lin, Maulik Shukla, Weili Nie, Josh Romero, Christian Dallago, Arash Vahdat, Chaowei Xiao, Thomas Gibbs, Ian Foster, James J. Davis, Michael E. Papka, Thomas Brettin, Rick Stevens, Anima Anandkumar, Venkatram Vishwanath, and Arvind Ramanathan. Genslms: Genome-scale language models reveal sars-cov-2 evolutionary dynamics. October 2022. doi: 10.1101/2022.10.10.511571. URL http://dx.doi.org/10.1101/2022.10.10.511571.