Machine Learning Models Assisting the Development of Antibody Therapeutics and Vaccines - an Emerging Trend

Felipe Leno da Silva, Mikel Landajuela, Edwin A. Saada, Piyush Karande, Sudeep Sarma, Igor D'Angelo, Simone Conti, Daniel Faissol
Lawrence Livermore National Laboratory
7000 East Ave
Livermore, CA 94550 USA
{leno, landajuelala1, saada1, karande1, sarma3, dangelo5, conti3, faissol1}@llnl.gov

Abstract

The development of novel effective medical treatments is one of the most important and expected beneficial effects of the AI revolution. This decade is witnessing the rise of AI models able to predict complex properties for protein-protein interactions that hold great promise in assisting in the development of antibody therapeutics and vaccines, including for diseases that long eluded us in the pursuit of an effective treatment. This paper introduces this area of research in a language accessible to an AI researcher, exploring the biological problems that can be solved by AI models, as well as the general context to make solutions feasible in practical scenarios. We survey the main current trends and works in this research area and point towards current still unsolved challenges and trade offs. We expect this paper will be extremely helpful for AI researchers trying to join the field, as well as for researchers already working in one of the subtopics that wish to have a better understanding of the general context around it.

Context

Humans rely on their immune system as the primary line of defense when infected with a pathogen as illustrated in Figure 1. Following exposure, our immune system will execute a process (affinity maturation) similar to a genetic algorithm to fight the infection. Germinal centers within our lymph nodes will proliferate B cells (which are slightly mutated throughout proliferation), and those B cells will try to bind (attach) to samples of pathogen cells. As more successful the B cells are in binding the pathogen, more likely they are to be replicated and further differentiated, and eventually the particularly successful B cells develop into plasma cells, which will then release antibodies into the blood to fight the infection in the rest of the body. Some B cells that produced successful antibodies are turned into memory cells, which are recalled in the event of a future re-infection (hence for many diseases we are infected just once in our lifetime and acquire immunity). This process is called adaptive immune response.

Therefore, the ability to produce strong antibodies protects us from a variety of diseases caused by microorganisms. However, there are a myriad of organisms that cause acute infections faster than our adaptive immune response can produce effective antibodies, causing serious illness that result in high mortality and complication rates.

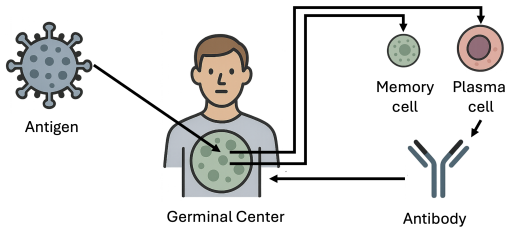

Figure 1: Illustration of the affinity maturation process. When infected with an antigen, our germinal centers proliferate B cells in search of an antibody that fights the infection. Eventually, successful B-cells are differentiated into memory and plasma cells, and the latter will release antibodies to fight the infection.

Modern medicine can follow two different paths leveraging this knowledge of how our immune system works to produce treatments:

1. **Antibody Therapeutics** - which consists of manufacturing the antibodies that target a specific pathogen in a laboratory and administering it as a medicine, providing an immediate, passive immunity;
2. **Vaccines** - they focus on leveraging our immune systems to build up protection against the disease. We manufacture an *antigen* at a lab, which after administered will guide the recipient's immune system towards developing immunization, inducing active immunity and long-term memory.

Both approaches have drawbacks and advantages (e.g. vaccines do not work for immunocompromised people and antibody therapeutics do not provide long-term protection). However, a good part of the challenge for developing those treatments is identifying an antibody that works.

The traditional approach is illustrated in Figure 2 (a). It consists of examining the surviving population among those infected by the disease and trying to extract from their immune cells (which are floating in blood) the antibody that was successful in neutralizing the pathogen. Additionally, we can also try to generate repertoires from animal hosts

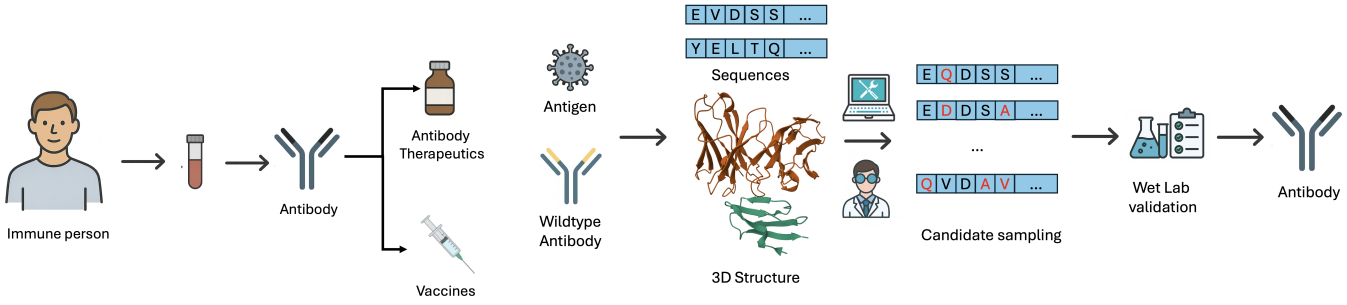

Figure 2: The traditional approach consists of analyzing the blood of people that either survived or are immune to the disease (a); while rational design tries to predict relevant biophysical properties from the 3D structures and sequences (either computationally or manually) with the goal to directly finding an effective antibody (b). After finding the antibody that neutralizes the pathogen, we can start the strenuous process of developing antibody therapeutics and vaccines.

(which has its own set of issues such as lack of immunogenic response and/or B cell bias). While this experimental approach is still somewhat prevailing, it is not only very hard and expensive, it is also slow since you have to wait for enough people to be contaminated that rare immunity arises.

With increase in availability of computational (and other lab-related) resources, *rational design* became more prevalent, where treatments are developed from leveraging information of protein structure and expert knowledge to engineer improved desired properties (Lutz 2010) with minimal wet lab assay, as illustrated in Figure 2 (b).

Early computational approaches focused on molecular dynamics simulations by directly modeling force fields derived from the atoms that compose antigens and antibodies. Although AI techniques were not mature, early results from tools such as Rosetta (Barlow et al. 2018) and FoldX (Schymkowitz et al. 2005) resulted in an excitement that created a new subfield in structural biology and motivated institutions to invest heavily in extremely optimized versions of those molecular dynamics simulations (Shaw et al. 2010).

Although a correlation has been observed (Yamashita 2018), those tools still have a course enough prediction of the binding strength that significant work and/or experimental budget were required to get to a promising antibody or antigen design.

Then, in what is today considered the "ImageNet Moment" for AI in Protein Design, AlphaFold2 (Jumper et al. 2021) achieved, at the time, unprecedented results at the CASP (Critical Assessment of Structure Prediction) competition (Kryshtafovych et al. 2021).

CASP consists of predicting the 3D structure of proteins given their amino acid sequence. While not directly solving the antibody design problem, having a model able to make this prediction indicates implicit knowledge of how the atoms interact with each other, and consequentially makes feasible AI-based rational design.

This watershed moment led to numerous research directions developing and applying ML models to problems related to antibody therapeutics and vaccine design.

In the rest of this paper, we will explore some key challenges that have been tackled by AI models. Nowadays, Protein AI models are an integral part of all major AI conferences, and are considered to be so important for the development of future treatments that the Nobel Prize in Chemistry 2024 was shared between the AlphaFold team leads and David Baker for their impact in computational protein design (Baker, Hassabis, and Jumper 2024).

(Simplified) AI Problems Description

For the purpose of this paper, we consider that both Antibodies and Antigens are proteins[1] - which are in turn composed of amino acids. There are in total 20 canonical amino acids, and there is tendency for a sequence of amino acids to fold to a particular 3D structure in the same environment (Anfinsen 1973). Therefore any protein can be conveniently described either by its 3D structure (usually saved in .pdb files (Jiménez-García et al. 2021)) or a "string" encoding the sequence of amino acids (A for Alanine, R for Arginine, etc.), as illustrated in Figure 2(b), which naturally makes it convenient to adapt NLP tools to handle protein data. The most important challenge to be solved for Antibody Therapeutics is, for a target antigen $Ag^{\oplus}$, find the antibody $Ab^{\star}$ that binds to $Ag^{\oplus}$ as strongly as possible:

$$Ab^{\star} = \arg max_{Ab\ in A} \boldsymbol{b}(Ag^{\oplus}, Ab), \tag{1}$$

where $\boldsymbol{b}(x, y)$ is the binding strength between antibody y and antigen x and A is the space of all possible antibodies. Conversely, once $Ab^{\star}$ is found, designing a vaccine requires finding an antigen $Ag^{\star}$ that excites our immune system to learn how to produce $Ab^{\star}$:

$$Ag^{\star} = \arg \max_{\substack{Ag \in D \\ \text{s.t.} risk(Ag) \leq t}} \boldsymbol{P_i}(Ag, Ab^{\star}) \tag{2}$$

where $\boldsymbol{P_i}(Ag, Ab)$ is the probability of a healthy person developing antibodies Ab after being vaccinated with Ag, $risk(Ag)$ represents the risk of infection or serious adverse

[1]While there are types of non-protein antigens, they would need models and techniques other than the explored in this paper.

reactions when vaccinated with Ag, and t represents a tolerable risk.

In this most general description, this is an incredibly challenging problem for a few reasons:

1. A complete antibody sequence has over one thousand amino acids, meaning the design space is in the order of 20^{1000}, which is orders of magnitude higher than the estimated number of atoms in the observable universe.
2. Techniques such as surface plasmon resonance (Malmqvist 1993) do protein-protein interaction studies and are probably the closest we can get to having a ground truth label *in vitro* at the time this paper is being written. However, analysis costs are high and considered to be prohibitive to search over the design space
3. Since we do not understand the affinity maturation process to the level we can exactly predict which antibodies a person will develop when vaccinated, it is even more difficult to evaluate the effectiveness of vaccines, since we can only have a ground truth of the effectiveness by observing the percentage of vaccinated people to contract the disease compared to a control group, which can be very difficult to do for diseases with low infection rate such as HIV.

Hence, it is reasonable that most of the progress in the area focused on a relaxed version of the problem. The interaction relevant for binding affinity happens in an "interface", where the target-specific complementarity determining regions (CDRs) (Wu and Kabat 1970) of the antibody interact with the surface of the antigen. The *epitope* and *paratope* are respectively the parts of the antigen and antibody sequences that compose the interface. So, as long as you identify (or estimate) the paratope and epitope for $Ag^{\oplus}$ and an existing antibody Ab_w, your design task can focus on performing a few mutations on the CDR of Ab_w, drastically reducing the design space to be explored:

$$\tilde{Ab}^{\star} = Ab_w \cup \arg max_{\boldsymbol{m} \in CDR(Ab_w)} \boldsymbol{b}(Ag^{\oplus}, Ab_w \cup \boldsymbol{m}), \tag{3}$$

where we denote $Ab \cup \boldsymbol{m}$ as the operation of applying mutations $\boldsymbol{m}$ to Ab and $CDR(Ab)$ as the power set of possible mutations for the CDR of Ab (i.e., the combination of all possible 20 amino acids for each CDR position). This type of "Redesign" task has been extensively explored with great success (Desautels et al. 2024; Zhu et al. 2025).

Lines of Work

Given the general challenge described in the last section, finding $Ab^{\star}$ and/or $Ag^{\star}$ directly from $Ag^{\oplus}$ is still a very difficult problem mainly due to how expensive it is to gather wet lab data, severely limiting the amount of training data we have available for models. Thus, most of the work has focused instead on key functionalities that would equip biologists with the knowledge to perform an informed heuristic search on a reduced design space. The most relevant emerging trends for AI models include:

Folding prediction

Rational design normally requires reasoning over how Ab_w and $Ag^{\oplus}$ interact in their 3D structure, but inspecting what is the amino acid sequence or 3D structure for a desired antigen or antibody have completely different costs. The sequencing can be performed with Next-Generation Sequencing (NGS) which is not only relatively inexpensive, it also scales gracefully and can be used to sequence a batch of thousands of targets at the same time. On the other hand, finding the structure is done through techniques such as X-Ray Crystallography which requires expensive equipment, is very labor intensive, and does not scale well.

Therefore, there is significant interest in being able to predict to which structure a given sequence will "fold", since we would be able to get a structure through simple NGS, as well as (relatively) quickly estimating the corresponding structure for any new design we create in the sequence space. Thus the folding challenge can be described as learning a model f_{fold} to:

$$f_{fold} : \mathrm{seq}(P) \rightarrow \mathrm{stru}(P) \tag{4}$$

where seq and stru are respectively the amino acid sequence and 3D structure of a protein P.

AlphaFold3 (Abramson et al. 2024) can be considered the gold standard for general protein folding prediction at the time this paper is being written. SOTA results are achieved by performing both a template search on the input sequence and calculating multiple sequence alignments (MSA) to gather evolutionary information. This information is used to estimate "pairformer" embeddings through a transformer architecture, which are then fed into a diffusion model which iteratively estimates the 3D structure. Despite the desirable accuracy, calculating MSAs as well as forward passes in its large architecture makes running AlphaFold3 very time and resource consuming.

ESM3 (Hayes et al. 2025) achieves faster prediction by directly feeding the sequence into a transformer with geometric attention (avoiding the computation of MSAs). Both AlphaFold3 and ESM3 are big models concerned with folding properly all kinds of proteins, hence there is a margin to gain performance or at least significantly reducing computational cost by building antibody-specific models.

Inverse folding

Inverse folding (IF) deals with the inverse problem:

$$f_{IF} : \mathrm{stru}(P) \rightarrow \mathrm{seq}(P), \tag{5}$$

that is, designing an amino acid sequence that is most likely to fold into the desired structure. It might sound counterintuitive we care about this task given extracting the structure is more expensive than the sequence. However, biologists use the structure to make informed decisions regarding properties the antibody/antigen is desired to have. Thus, it is possible to design a 3D structure manually (Wang et al. 2022) then use an IF model to get its corresponding sequence.

ProteinMPNN (Dauparas et al. 2022) uses a message passing neural network to predict the sequence and can be considered a watershed model, which became extremely

popular due to being open source, having a good performance, and being a small network architecture that makes a prediction in a few seconds even on a laptop CPU.

Other models have built upon ProteinMPNN or have a similar architecture leveraging Graph Neural Networks. Notably, ESM-IF (Hsu et al. 2022) is a geometric transformer trained using not only experimentally gathered structures (like ProteinMPNN) but also adding a large number of structures predicted from sequences using AlphaFold2.

Like AlphaFold, those models are general protein models and not specifically trained to design Antibody sequences. Thus a few works finetuned those protein models with antibody-specific data to improve performance, resulting in models such as AbMPNN (Dreyer et al. 2023) and AntiFold (Høie et al. 2025).

IF models can also be very helpful for vaccines based on scaffolding (Burton 2010). Scaffolding consists of building an antigen that will have the same epitope, after it has been identified, expressed on the surface of another "body", so that a neutral and harmless molecule can mimic the pathogen for our immune system while having zero risk of infection through vaccination. In principle, one could design the desired scaffold and use an IF model to discover the sequence (although as far as we know, the current IF models have not reached sufficient performance to enable that).

Binding prediction

Binding prediction (***b*** in Eq. 1) is arguably the most important challenge to be solved, since if the designed antibody is not a binder to the pathogen, everything else is pointless. Binding is often measured in terms of $\Delta\Delta G$ (delta delta G), the change in the change in Gibbs free energy, in other words, what was the change in binding strength after a mutation was applied in an antibody/antigen. A negative $\Delta\Delta G$ means the mutation is stabilizing and thus more favorable.

There has been extensive work training models to predict antibody-antigen binding (Agarwal et al. 2025; Zhang et al. 2024; Jin et al. 2024; Yuan et al. 2023). However, to the best of our knowledge, no model has achieved performance high enough to be fully trusted in general antibody-antigen prediction problems. Good part of the problem comes from the difficulty in accessing experimental training data. Many of the proposed models have been trained and/or evaluated in simulated data. Even for models trained with wetlab data, using the open sourced wet lab datasets brings a number of issues, from those datasets being very unbalanced (targets such as SARS-Cov-2 have much more abundant data than others) to a mismatch of training and deployment conditions. Datasets composed of experiments gathered only in a single wet lab might carry artifacts of that particular environment, which will not be replicated for a different environment. Therefore, reported performance is often not replicated when trying to use the models in practical discovery work.

A correlation has been observed between the log likelihood of logits from IF models and the true $\Delta\Delta G$ of protein mutations (Riesselman, Ingraham, and Marks 2018), thus some works have directly used those models as proxy for change in binding strength. A recent work (Deng et al. 2025) went a step further and modeled the binding $\Delta\Delta G$ as a function of individual folding $\Delta\Delta G$s, giving us a principled way of estimating binding from IF models. Their finetuned model achieved promising results although no wet lab validation has been performed yet.

Automated Epitope/Paratope Identification

As mentioned before, in order to solve Equation (1) or (2) the design space is normally first pruned by constraining the sampling process only for amino acids in the interface between the antigen and the antibody (paratope and epitope).

However, identifying with precision where the correct epitope and paratope are may not be an easy task (Peri et al. 2013), especially if the epitope is buried within the protein structure (i.e., it is not easily identifiable on the surface).

Therefore, another line of research explores identifying the epitope and/or paratope for an arbitrary antigen-antibody pair. A number of approaches focus on leveraging protein embeddings from models trained for the other prediction tasks, finetuning them to be able to predict epitope-paratope pairs for known antibody-antigen pairs (Kalemati et al. 2024; Wang, Wang, and Zhang 2024).

A key limitation of the approaches in this area is that most papers are trained to mimic known paratope-epitope pairs, whereas there might be unidentified possible interfaces that would allow for even stronger binding. We expect that future works in the area will explore considering this problem is a hierarchical learning problem, where for each sampled interface there is a different optimal design to optimize binding.

Generative Protein Design

While most of the work in the area is concerned with predicting properties for arbitrary antibody-antigen pairs, sampling efficiently the design space is not fully addressed by the predictive models.

Approaches that explore in the sequence space normally perform an autoregressive sampling where the joint probability of the sequence is given as:

$$P(seq) = \prod_{i}^{|seq|} P(seq_i | seq_1, \ldots, seq_n) \tag{6}$$

capturing sequential and context-dependent dependencies across positions. In practice, the amino acids are sequentially sampled conditioned in the partially sampled sequence. The main family of algorithms to perform this task are *Deep Symbolic Optimization* (Silva et al. 2023a) and *GFlowNets* (Bengio et al. 2021).

Another approach is to sample directly in the structure space, giving as output an all-atom structure of the desired design. This task is normally performed by diffusion models, where RFDiffusion (Watson et al. 2023) is the most well-known models among the first ones to perform this task[2].

[2]Technically, the original RFDiffusion only outputs a backbone structure and needs an inverse folding model to give it side chains. The most recent version (Butcher et al. 2025) gets closer to an end-to-end all-atom approach

Affinity Maturation Simulation

Finding a vaccine antigen requires reasoning over the affinity maturation process (i.e., which antibodies the person will develop once vaccinated). Gathering data from affinity maturation is extremely difficult and thus there are no diverse and big datasets. On the other hand, biologists built enough knowledge of the mechanisms behind affinity maturation that Agent-based models have achieved relative success in predicting memory B cells state after a vaccination (i.e., which antibodies the person is able to produce) (Conti et al. 2022). However a limiting factor of those models is a large number of binding estimation calculations needed, which in published work resulted in limited experimentation considering a subset of the B cell repertoire that would be available in a real human. With the development of more effective and faster ML-based binding affinity prediction models, we can also expect significant improvement in the affinity maturation simulations.

Notably, BIOVAX (Faris et al. 2024b) proposes to combine an autoregressive antigen generator with agent-based modeling of affinity maturation, effectively combining solutions for many of the challenges discussed in this paper to build a pipeline for automated screening of antigen candidates. In *in silico* validation, the BIOVAX design performed better than other computationally designed antigens in providing immunization to a panel of real HIV viruses.

Therefore, one of the most important yet unexplored directions in this field is the development of models that can accurately predict the production of antibodies after vaccination to an arbitrary antigen.

Developability/Safety/Stability

As with most practical applications, finding an antibody or antigen is not solved by simply picking the design that optimizes binding. The ability to bind a given target on an antigen does not necessarily imply an inability to bind a different surface on the same antigen, or a different protein altogether. This phenomenon, known as cross-reactivity, can be especially concerning if cross-reactive antibodies recognize host proteins, as this can lead to a plethora of serious autoimmune diseases (e.g., Rheumatoid Arthritis).

While it is unfeasible to model all cross-reactivity or adverse reactions, a common practice is to train models on large repositories of known human antibodies (Olsen, Boyles, and Deane 2022), effectively creating a proxy model for "humanness" (Vashchenko et al. 2022) that could be used to flag designs that are different from any antibody ever produced by a human and thus likely to fail as a therapy.

Furthermore, designs submitted for approval to agencies like the U.S. Food and Drug Administration (FDA) have been painstakingly crafted to optimize every aspect of manufacturability and safety, thus this list can be used to train developability models. However, clinical trials follow a phased approach and it might not be trivial to know if a design was abandoned at a certain phase because it did not work or for other unrelated reason (IP issues, deprioritization from the pharmaceutical company, etc.). Moreover, the number of samples in those datasets are very low. Nevertheless, we believe that applying approaches that can methodically make use of few high-quality samples such as PU-learning (Gong et al. 2018) in a conjunction of (unlabeled) human antibodies and FDA-approved designs is a promising direction.

Despite the challenges explored in this subsection, developability and safety concerns are biophysical and biochemical properties of the designs and, in principle, could be modeled/learned. If in the future we have models that can reliably predict at least some of those properties, Equations (1) and (2) can be rewritten as a multiobjective problem, taking into account developability constraints since early on in the optimization process. In fact, a few exploratory works proposed multiobjective learning algorithms to sample antibody designs (Faris et al. 2024a).

Active Learning Lab Automation

Leveraging the recent advances in automation, *Biofoundries* are highly automated facilities where wet lab experimentation can be performed in a fully autonomous way through the use of robotic platforms (Hillson et al. 2019).

The existence of those facilities would in principle enable modeling any of the challenges discussed in this paper as an Active Learning problem where (expensive but available) wet lab experiments could be strategically selected to give high confidence labels for certain designs to update model predictions. Although very few of such facilities are currently available, there has been some exploratory works integrating protein model learning with biofoundries (Zhang et al. 2025; Yu et al. 2023). In principle, the infrastructure for model training and the biofoundries could even be in different locations and/or institutions.

This is highly related to the exploratory works in multifidelity learning for generative models (Hernandez-Garcia et al. 2024; Silva et al. 2023b), where those models explicitly consider that there is both a cheaper and faster simulation to learn from and a more expensive but reliable source of experimentation. Although published papers in the area so far only considered different *in silico* tools, they directly generalize for biofoundries. Integrating biofoundaries with AI driven training would dramatically accelerate antibody and protein design, reducing experimental cost and improving design quality.

Conclusion

Machine Learning (ML) holds great promise to unlock the development of antibody therapeutics and vaccines for diseases that are currently untreatable. This paper described in an accessible language what are the associated challenges in which ML models have been used or are likely to be useful for in the short term. We explored the main high level problems biologists would like AI to be able to solve, as well as key models to represent the state of the art in this area. This paper can be used as an initial resource to have a high level overview of how our community can be most impactful in this area, from which the reader can choose which sub area suits their interest better for a deeper dive. By combining AI capabilities with experimental biology, the community is poised to accelerate the development of safe and effective therapeutics.

Acknowledgments

We thank Ian Humphreys for feedback on the content of the paper. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. LLNL-CONF-2010953.

References

Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; et al. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016): 493–500.

Agarwal, A. A.; Harrang, J.; Noble, D.; McGowan, K. L.; Lange, A. W.; Engelhart, E.; Lahman, M. C.; Adamo, J.; Yu, X.; Serang, O.; et al. 2025. AlphaBind, a domain-specific model to predict and optimize antibody–antigen binding affinity. In *Mabs*, volume 17, 2534626. Taylor & Francis.

Anfinsen, C. B. 1973. Principles that Govern the Folding of Protein Chains. *Science*, 181(4096): 223–230.

Baker, D.; Hassabis, D.; and Jumper, J. 2024. The Nobel Prize in Chemistry 2024. NobelPrize.org. Awarded for computational protein design (Baker) and protein structure prediction (Hassabis Jumper).

Barlow, K. A.; Ó Conchúir, S.; Thompson, S.; Suresh, P.; Lucas, J. E.; Heinonen, M.; and Kortemme, T. 2018. Flex ddG: Rosetta ensemble-based estimation of changes in protein–protein binding affinity upon mutation. *The Journal of Physical Chemistry B*, 122(21): 5389–5399.

Bengio, E.; Jain, M.; Korablyov, M.; Precup, D.; and Bengio, Y. 2021. Flow network based generative models for non-iterative diverse candidate generation. *Neural Information Processing Systems (NeurIPS)*, 34: 27381–27394.

Burton, D. R. 2010. Scaffolding to build a rational vaccine design strategy. *Proceedings of the National Academy of Sciences*, 107(42): 17859–17860.

Butcher, J. K. V.; Krishna, R.; Mitra, R.; Brent, R. I.; Li, Y.; Corley, N.; Kim, P.; Funk, J.; Mathis, S. V.; Salike, S.; et al. 2025. De novo design of all-atom biomolecular interactions with RFDiffusion3. *bioRxiv*.

Conti, S.; Ovchinnikov, V.; Faris, J. G.; Chakraborty, A. K.; Karplus, M.; and Sprenger, K. G. 2022. Multiscale affinity maturation simulations to elicit broadly neutralizing antibodies against HIV. *PLoS computational biology*, 18(4).

Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; and Baker, D. 2022. Robust deep learning–based protein sequence design using ProteinMPNN. *Science*, 378(6615): 49–56.

Deng, A.; Householder, K.; Wu, F.; Thrun, S.; Garcia, K. C.; and Trippe, B. 2025. Predicting mutational effects on protein binding from folding energy. In *International Conference on Machine Learning (ICML)*. PMLR.

Desautels, T. A.; Arrildt, K. T.; Zemla, A. T.; Lau, E. Y.; Zhu, F.; Ricci, D.; Cronin, S.; Zost, S. J.; Binshtein, E.; Scheaffer, S. M.; et al. 2024. Computationally restoring the potency of a clinical antibody against Omicron. *Nature*, 629(8013): 878.

Dreyer, F. A.; Cutting, D.; Schneider, C.; Kenlay, H.; and Deane, C. M. 2023. Inverse folding for antibody sequence design using deep learning. In *Workshop on Computational Biology @ ICML*.

Faris, J. G.; Hayes, C. F.; Goncalves, A. R.; Sprenger, K. G.; Petersen, B. K.; Landajuela, M.; and Silva, F. L. 2024a. Pareto Front Training For Multi-Objective Symbolic Optimization. In *Adaptive and Learning Agents (ALA) Workshop at AAMAS*.

Faris, J. G.; Landajuela, M.; Sprenger, K. G.; Faissol, D.; and Silva, F. L. d. 2024b. Computational Antigen Optimization through Symbolic Optimization and Affinity Maturation Simulation. In *Workshop on AI for new Drug Modalities at NeurIPS (AIDrugX)*.

Gong, T.; Wang, G.; Ye, J.; Xu, Z.; and Lin, M. 2018. Margin based PU learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

Hayes, T.; Rao, R.; Akin, H.; Sofroniew, N. J.; Oktay, D.; Lin, Z.; Verkuil, R.; Tran, V. Q.; Deaton, J.; Wiggert, M.; Badkundri, R.; Shafkat, I.; Gong, J.; Derry, A.; Molina, R. S.; Thomas, N.; Khan, Y. A.; Mishra, C.; Kim, C.; Bartie, L. J.; Nemeth, M.; Hsu, P. D.; Sercu, T.; Candido, S.; and Rives, A. 2025. Simulating 500 million years of evolution with a language model. *Science*.

Hernandez-Garcia, A.; Saxena, N.; Jain, M.; Liu, C.-H.; and Bengio, Y. 2024. Multi-Fidelity Active learning with GFlowNets. *Transactions on Machine Learning Research*.

Hillson, N.; Caddick, M.; Cai, Y.; Carrasco, J. A.; Chang, M. W.; Curach, N. C.; Bell, D. J.; Le Feuvre, R.; Friedman, D. C.; Fu, X.; et al. 2019. Building a global alliance of biofoundries. *Nature communications*, 10(1): 2040.

Hsu, C.; Verkuil, R.; Liu, J.; Lin, Z.; Hie, B.; Sercu, T.; Lerer, A.; and Rives, A. 2022. Learning inverse folding from millions of predicted structures. In *International Conference on Machine Learning (ICML)*.

Høie, M. H.; Hummer, A. M.; Olsen, T. H.; Aguilar-Sanjuan, B.; Nielsen, M.; and Deane, C. M. 2025. AntiFold: improved structure-based antibody design using inverse folding. *Bioinformatics Advances*, 5(1): vbae202.

Jiménez-García, B.; Teixeira, J. M.; Trellet, M.; Rodrigues, J. P.; and Bonvin, A. M. 2021. PDB-tools Web: A User-friendly Interface for the Manipulation of PDB Files. *Proteins: Structure, Function, and Bioinformatics*, 89(3): 330–335.

Jin, R.; Ye, Q.; Wang, J.; Cao, Z.; Jiang, D.; Wang, T.; Kang, Y.; Xu, W.; Hsieh, C.-Y.; and Hou, T. 2024. AttAB-seq: an attention-based deep learning prediction method for antigen–antibody binding affinity changes based on protein sequences. *Briefings in Bioinformatics*, 25(4): bbae304.

Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873): 583–589.

Kalemati, M.; Noroozi, A.; Shahbakhsh, A.; and Koohi, S. 2024. ParaAntiProt provides paratope prediction using antibody and protein language models. *Scientific Reports*, 14(1).

Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; and Moult, J. 2021. Critical assessment of methods of protein structure prediction (CASP)—Round XIV. *Proteins: Structure, Function, and Bioinformatics*, 89(12): 1607–1617.

Lutz, S. 2010. Beyond directed evolution—semi-rational protein engineering and design. *Current Opinion in Biotechnology*, 21(6): 734–743.

Malmqvist, M. 1993. Surface plasmon resonance for detection and measurement of antibody-antigen affinity and kinetics. *Current Opinion in Immunology*, 5(2): 282–286.

Olsen, T. H.; Boyles, F.; and Deane, C. M. 2022. Observed Antibody Space: A diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Science*, 31(1): 141–146.

Peri, C.; Gagni, P.; Combi, F.; Gori, A.; Chiari, M.; Longhi, R.; Cretich, M.; and Colombo, G. 2013. Rational epitope design for protein targeting. *ACS chemical biology*, 8(2): 397–404.

Riesselman, A. J.; Ingraham, J. B.; and Marks, D. S. 2018. Deep generative models of genetic variation capture the effects of mutations. *Nature methods*, 15(10): 816–822.

Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; and Serrano, L. 2005. The FoldX web server: an online force field. *Nucleic acids research*, 33: W382–W388.

Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; and Wriggers, W. 2010. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science*, 330(6002): 341–346.

Silva, F. L. d.; Goncalves, A.; Nguyen, S.; Vashchenko, D.; Glatt, R.; Desautels, T.; Landajuela, M.; Faissol, D.; and Petersen, B. 2023a. Language model-accelerated deep symbolic optimization. *Neural Computing and Applications*, 1–17.

Silva, F. L. d.; Yang, J.; Landajuela, M.; Goncalves, A.; Ladd, A.; Faissol, D.; and Petersen, B. 2023b. Toward Multi-Fidelity Reinforcement Learning for Symbolic Optimization. In *Workshop on Adaptive and Learning Agents (ALA) at AAMAS*.

Vashchenko, D.; Nguyen, S.; Goncalves, A.; Silva, F. L. d.; Petersen, B.; Desautels, T.; and Faissol, D. 2022. AbBERT: Learning Antibody Humanness via Masked Language Modeling. In *Workshop on Healthcare AI and COVID-19 @ ICML*.

Wang, J.; Lisanza, S.; Juergens, D.; Tischer, D.; Watson, J. L.; Castro, K. M.; Ragotte, R.; Saragovi, A.; Milles, L. F.; Baek, M.; et al. 2022. Scaffolding protein functional sites using deep learning. *Science*, 377(6604): 387–394.

Wang, Z.; Wang, Y.; and Zhang, W. 2024. Improving paratope and epitope prediction by multi-modal contrastive learning and interaction informativeness estimation. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 6053–6061.

Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; et al. 2023. De novo design of protein structure and function with RFdiffusion. *Nature*, 620(7976): 1089–1100.

Wu, T. T.; and Kabat, E. A. 1970. An analysis of the sequences of the variable regions of bence jones proteins and myeloma light chains and their implications for antibody complementarity. *Journal of Experimental Medicine*, 132(2): 211–250.

Yamashita, T. 2018. Toward rational antibody design: Recent advancements in molecular dynamics simulations. *International immunology*, 30(4): 133–140.

Yu, T.; Boob, A. G.; Singh, N.; Su, Y.; and Zhao, H. 2023. In vitro continuous protein evolution empowered by machine learning and automation. *Cell Systems*, 14(8): 633–644.

Yuan, Y.; Chen, Q.; Mao, J.; Li, G.; and Pan, X. 2023. Dg-affinity: predicting antigen–antibody affinity with language models from sequences. *BMC bioinformatics*, 24(1): 430.

Zhang, Q.; Chen, W.; Qin, M.; Wang, Y.; Pu, Z.; Ding, K.; Liu, Y.; Zhang, Q.; Li, D.; Li, X.; et al. 2025. Integrating protein language models and automatic biofoundry for enhanced protein evolution. *Nature Communications*, 16(1): 1553.

Zhang, Y.; Dong, M.; Deng, J.; Wu, J.; Zhao, Q.; Gao, X.; and Xiong, D. 2024. Graph masked self-distillation learning for prediction of mutation impact on protein–protein interactions. *Communications Biology*, 7(1): 1400.

Zhu, F.; Rajan, S.; Hayes, C. F.; Kwong, K. Y.; Goncalves, A. R.; Zemla, A. T.; Lau, E. Y.; Zhang, Y.; Cai, Y.; Goforth, J. W.; Landajuela, M.; Gilchuk, P.; Kierny, M.; Dippel, A.; Amofah, B.; Kaplan, G.; Peano, V. C.; Morehouse, C.; Sparklin, B.; Gopalakrishnan, V.; Tuffy, K. M.; Nguyen, A.; Beloor, J.; Kijak, G.; Liu, C.; Dijokaite-Guraliuc, A.; Mongkolsapaya, J.; Screaton, G. R.; Petersen, B. K.; Desautels, T. A.; Bennett, D.; Conti, S.; Segelke, B. W.; Arrildt, K. T.; Kaul, S.; Grzesiak, E. A.; Silva, F. L. d.; Bates, T. W.; Earnhart, C. G.; Hopkins, S.; Sundaram, S.; Esser, M. T.; Francica, J. R.; Faissol, D. M.; and consortium, L. G. U. I. D. E. G. 2025. Preemptive optimization of a clinical antibody for broad neutralization of SARS-CoV-2 variants and robustness against viral escape. *Science Advances*, 11(13): eadu0718.