Computational and Structural Biotechnology Journal 27 (2025) 1286–1295

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

Leveraging large language models to predict antibody biological activity against influenza A hemagglutinin

Ella Barkan [a], Ibrahim Siddiqui [b], Kevin J. Cheng [c], Alex Golts [a], Yoel Shoshan [a], Jeffrey K. Weber [c], Yailin Campos Mota [d], Michal Ozery-Flato [a,*,1], Giuseppe A. Sautto [d,*,1]

[a] IBM Research-Israel, Haifa, Israel
[b] Case Western Reserve University, Cleveland, OH, USA
[c] IBM TJ Watson Research Center, Yorktown Heights, NY, USA
[d] Florida Research and Innovation Center, Cleveland Clinic, Port St. Lucie, FL, USA

ARTICLE INFO

Keywords:
Monoclonal antibodies (mAbs)
Influenza virus
Hemagglutinin (HA)
Hemagglutination inhibition (HAI) assay
Binding assay
Artificial Intelligence (AI)
Machine learning (ML)
Large language models (LLM)

ABSTRACT

Monoclonal antibodies (mAbs) represent one of the most prevalent FDA-approved treatments for autoimmune, infectious, and cancer diseases. However, their discovery and development remains a time-consuming and costly process. Recent advancements in machine learning (ML) and artificial intelligence (AI) have shown significant promise in revolutionizing antibody discovery field. Models that predict antibody biological activity enable in silico evaluation of binding and functional properties; such models can prioritize antibodies with the highest likelihood of success in laboratory testing procedures. We explore an AI model for predicting the binding and receptor blocking activity of antibodies against influenza A hemagglutinin (HA) antigens. Our model is developed with the Molecular Aligned Multi-Modal Architecture and Language (MAMMAL) framework for biologics discovery to predict antibody-antigen interactions using only sequence information. To evaluate the model's performance, we tested it under various data split conditions to mimic real-world scenarios. Our model achieved an area under the receiver operating characteristic (AUROC) score of ≥ 0.91 for predicting the activity of existing antibodies against seen HAs and an AUROC score of 0.9 for unseen HAs. For novel antibody activity prediction, the AUROC was 0.73, which further declined to 0.63–0.66 under stringent constraints on similarity to existing antibodies. These results demonstrate the potential of AI foundation models to transform antibody design by reducing dependence on extensive laboratory testing and enabling more efficient prioritization of antibody candidates. Moreover, our findings emphasize the critical importance of diverse and comprehensive antibody datasets to improve the generalization of prediction models, particularly for novel antibody development.

1. Introduction

The influenza A virus is responsible for one of the most prevalent infectious diseases worldwide, with associated influenza infections impacting an estimated 20–40 million people annually in the United States alone [1,2]. Despite its wide prevalence, the influenza A virus remains a significant global health concern due to its rapid mutation rate and potential spillovers from animal reservoirs that enable persistence and continued evolution [3]. In resource-limited countries, seasonal outbreaks of influenza A can result in severe illness and even death [4]. A key factor in the virus's ability to infect host cells is hemagglutinin (HA), the main protein on the viral surface that facilitates binding to host cell receptors. Importantly, HA also serves as a primary target for the immune system: when an antibody binds to HA, that antibody can neutralize the virus and prevent infection [5]. In fact, HA is the primary antigen contained in all the current standards of care for influenza vaccine formulations, serving as the main target for developing immunity to influenza A virus.

Monoclonal antibodies are critical for the development of diagnostic, immunotherapeutic, and immunoprophylactic tools for treating/preventing diseases spanning cancer, autoimmune disorders, and infectious diseases. In particular, mAbs have shown great promise in the prevention and treatment of infections caused by respiratory pathogens like the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the

* Corresponding authors.
E-mail addresses: ozery@il.ibm.com (M. Ozery-Flato), sauttog@ccf.org (G.A. Sautto).
[1] These authors jointly supervised this work.

https://doi.org/10.1016/j.csbj.2025.03.038
Received 31 January 2025; Received in revised form 19 March 2025; Accepted 21 March 2025
Available online 24 March 2025

respiratory syncytial virus (RSV) [6].

In this context, over the past two decades, we and other groups have characterized a plethora of monoclonal antibodies (mAbs) directed against the influenza HA that are endowed with different breadth of recognition, neutralization, and protection profiles [7–19].

Though mAbs now constitute almost a third of all newly FDA-approved treatments, therapeutic antibody discovery remains a lengthy and costly process [20]. Researchers' ability to test new antibody formulations in silico represents a critical bottleneck; predicting antibody binding to influenza A virus with AI technology offers a groundbreaking opportunity to model immunology and deliver tangible benefits for global health.

In recent years, AI models have opened new pathways for developing therapeutic molecules. AlphaFold architectures [21–23] achieve impressive single-domain protein structure prediction accuracy, but encounter significant challenges with predicting antibody–antigen complex structures and, independent of application, require computationally intensive large-scale sampling [23–25]. Biomedical language models, leveraging natural language processing frameworks, have also been successfully applied to modeling the "languages" of proteins [26–28] and antibodies [29–32], enhancing computational methods for antibody design and optimization [32–35]. Rapid identification and characterization of antibodies that bind viral pathogen antigens remains a critical challenge in antibody design.

In this study, we investigate the use of the MAMMAL methodology [19] for developing and fine-tuning pretrained biomedical language models, applying it to predict antibody binding and receptor blocking activity to influenza A virus hemagglutinin (HA). We fine-tune the pretrained model ibm/biomed.omics.bl.sm.ma-ted-458m on two standardized antibody-HA datasets: the first corresponding to HA binding assays, and the second corresponding to hemagglutination inhibition (HAI) assays. Using different data splits, we address three key applications: (1) imputing missing labels for pairs of known antibodies and HAs, (2) predicting activities of known antibodies against new HA sequences, and (3) predicting activities of new and arbitrary antibody sequences against HA sequences. An overview of our study pipeline, from data generation and preprocessing to model development and evaluation, is illustrated in Fig. 1.

Our models demonstrated high performance when imputing missing labels or predicting antibody activity against new HA sequences, achieving AUROC scores of ≥ 0.9 in both scenarios. When predicting anti-HA activity for novel antibodies, our models showed moderate classification power (AUROC of 0.73) that dropped to 0.63–0.66 under stringent antibody-similarity constraints. We validated the robustness of our results through cross-validation and extensive subgroup analyses. Furthermore, we assessed the models' abilities to identify broadly protective antibodies.

Our findings highlight the potential of leveraging language models pretrained on large-scale protein and antibody sequence data to develop sequence-based AI models for predicting antibody-target activity, even with a limited labeled dataset. This approach could streamline antibody research by reducing the need for extensive laboratory testing and helping prioritize more expensive in silico screening methods. Importantly, streamlining this process may represent a paradigm shift in the context of a pandemic response and preparedness, when reducing developing time for vaccines and antivirals is crucial to limit the spread and burden of emerging viruses [6].

2. Related work

Predicting the binding affinity or functional activity of an antibody candidate against a given target antigen represents a key step in

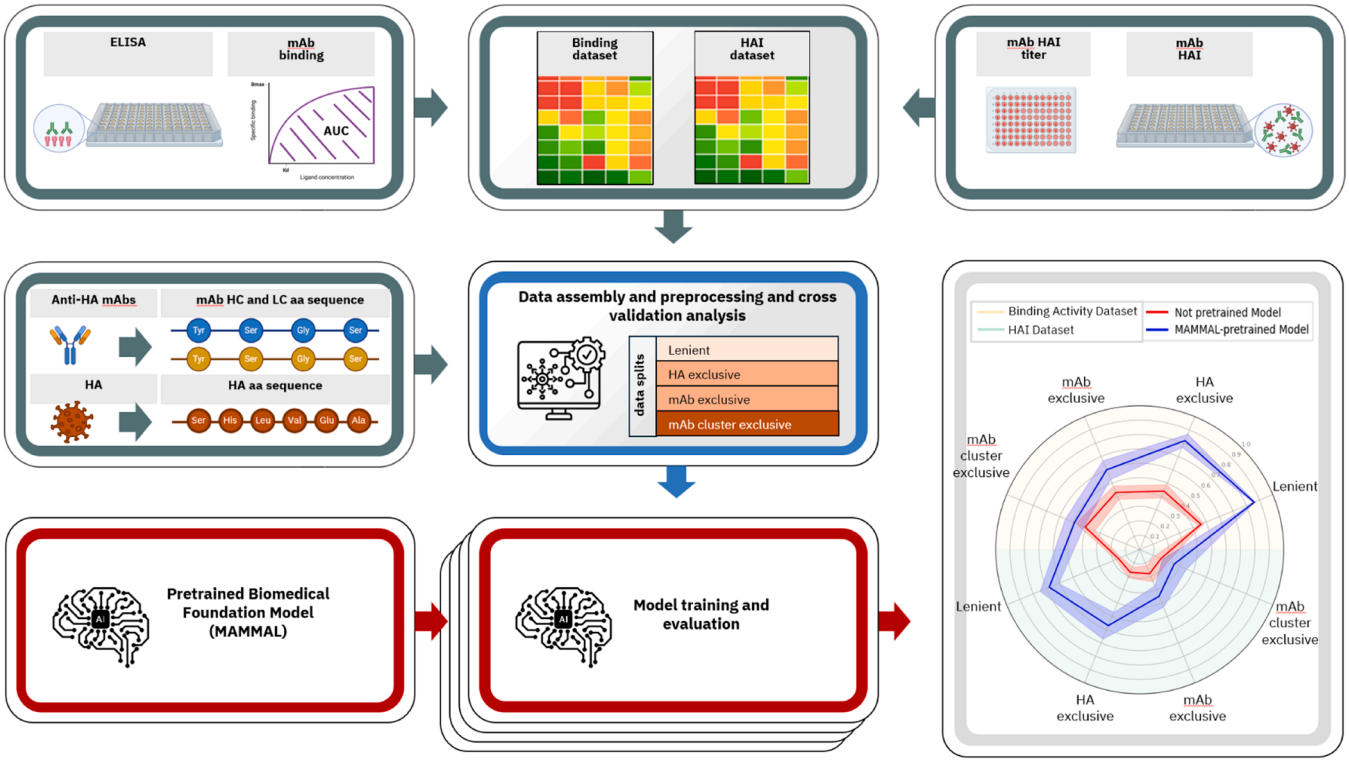

Fig. 1. Study pipeline for developing and evaluating AI models to predict antibody binding and hemagglutination inhibition (HAI) for influenza A hemagglutinin (HA). The dataset included monoclonal antibody (mAb)-HA amino acid sequences and corresponding enzyme-linked immunosorbent assay (ELISA) and HAI assay outcomes. To reflect real-world model applications, different data splits were employed, and a cross-validation approach was used to assess model robustness. Predictive models were developed by fine-tuning the MAMMAL (Molecular Aligned Multi-Modal Architecture and Language) biomedical foundation model. Performance was evaluated using the area under the receiver operating characteristic curve (AUROC) and the area under the precision-recall curve (AUPRC), with results compared to those from randomly initialized models. Performance metrics are visualized here in a spider diagram for comprehensive comparison.

therapeutic and vaccine design. Simulation methods like molecular dynamics (MD) and free energy perturbation (FEP) [36–38] can accurately capture the physical states of antibody-antigen complexes and offer one avenue toward making affinity predictions. However, these atomistic simulation approaches are computationally expensive and typically rely on accurate static structural information as a starting point. Molecular docking methods offer higher-throughput but less accurate estimates of intermolecular binding modes; docking scores exhibit notoriously weak correlations with affinity data [39,40]. Other knowledge-based scoring functions [41,42] offer alternative means of estimating binding affinities, but with limited accuracy and validity.

Machine learning (ML) methods (like those based on random forests) have set the standard for affinity prediction for small molecule-target complexes for decades. Similar approaches have since gained prevalence in the antibody discovery space. With the growing availability of large-scale molecular datasets, deep learning (DL) techniques have emerged as favored methods for both small molecule-protein [43–45] and antibody-antigen binding prediction [46,47]. The recent trend of using very large protein datasets to pretrain biomolecular large language models (LLMs) [26,27,48–50], with some models even being specific to antibodies [29–31,51], has redefined antibody-antigen binding affinity prediction as a downstream task. Most approaches in this domain rely on making predictions based on pretrained LLM embeddings or fine-tuning LLMs on labeled data [52].

Experimental three-dimensional structures of antibody-antigen complexes are scarce, expensive, and difficult to solve relative to simple protein sequencing. Breakthroughs in deep learning for protein structural modeling, particularly AlphaFold [21–23] and RoseTTAFold [53], have transformed efforts toward protein structure prediction. Despite significant advances in state-of-the-art structure prediction methods for antibody-antigen complex modeling, further refinements are needed to enhance antibody design efforts [23,54].

In the context of predicting antibody-HA interactions, previous work has relied on various computational approaches. Several studies have employed AlphaFold [55,56] for structural predictions of antibody-HA complexes, aiming to integrate empirical structural biology insights into binding predictions [56]. Other studies [34] have demonstrated the value of protein language model scores in predicting enhanced antibody binding, focusing primarily on antibody sequence analysis without explicit consideration of the target protein.

Several studies have utilized language models for predicting properties and de novo design/optimize antigen-specific antibodies, prominently SARS-CoV-2 spike-specific antibodies [57–59]. However, fewer studies leveraged these models to predict and design/optimize HA-directed antibodies. For example, Wang et al. [60] leveraged a memory B-cell language model to predict novel antibodies targeting the conserved HA stem region. Such applications underscore the potential of fine-tuning of pretrained models for improving antibody-antigen binding discrimination, as demonstrated for both SARS-CoV-2 and influenza antigens [61]. However, most language models lack the ability to directly quantify binding affinity, a task that is critical for accurately predicting and optimizing antibody efficacy.

Prior efforts in predicting antibody HAI titers have largely relied on ML models trained on virus-antiserum pairs contained within HAI titer datasets from the Worldwide Influenza Centre (WIC) [62]. Predictive strategies based on AdaBoost [63], random forest [64], and Bayesian approaches [65] have achieved notable success in this area. Additional bioinformatics methods [66,67] have provided insights into antigenic drift. Despite these advances, existing methods predominantly focus on virus-antiserum pairs and are not tailored to predict antibody-HA interactions based on sequence data alone.

3. Methods

3.1. Monoclonal antibodies, recombinant HA, and influenza viruses

The 188 human and mouse anti-HA mAbs used to generate the binding and HAI dataset featured in this study were previously described by our group or derived from the Biodefense and Emerging Infections Research Resources (BEI) and International Reagent Resources (IRR) repositories [8–19,68–70]. Recombinant HAs and influenza viruses corresponding to different H1N1 and H3N2 strains were produced and sequence confirmed as previously described [71,72]. Additionally, recombinant HA proteins and influenza viruses were validated in previously performed vaccination and serology studies, using previously characterized antisera and monoclonal antibodies [11,19,72–75].

3.2. Data description

This study's dataset comprises pairs of mAbs and influenza A HA antigens, evaluated using two assays: binding activity and HAI activity. These assays characterize 1) the breadth of binding, indicating the number of strains the mAb can bind, and 2) the breadth of receptor binding inhibition (i.e., HAI), which reflects the antibody's ability to prevent viral binding to the sialic acid receptor, thereby blocking virus entry and neutralizing the virus.

Binding activity assays were conducted as previously described [10,12,14,18] using ELISA, a method to detect and quantify antibody-antigen interactions by measuring a signal, such as color change, proportional to binding strength [5]. Binding strength is quantified as the area under the binding curve (AUC-ELISA), derived from three-fold serial dilutions of mAbs (20–0.009 μg/ML). AUC-ELISA values range from 0.5 (negative) to 20, with higher values indicating stronger binding. We considered AUC-ELISA values greater than 1 as positive binding.

HAI represents the ability of an antibody to block the interaction between HA and the sialic acid cell receptor, measured as the minimum antibody concentration required to inhibit red blood cell agglutination by the virus. In particular, the HAI assay involves mixing serially diluted mAbs with a fixed amount of virus and adding red blood cells to detect hemagglutination. Ultimately, the ability of an antibody to inhibit this agglutination is indicative of its neutralizing potency [19]. HAI assays with mAbs were performed as previously described [10–13,18]. HAI values range from 0.005 to 20 μg/ML, with lower values indicating higher potency. We classified HAI values below 10 μg/ML as positive outcomes. The mAbs in the dataset are represented by amino acid sequences of the variable regions of the heavy (HC) and light (LC) chains, excluding the constant regions. These variable regions include the framework regions (FRs) and the complementarity-determining regions (CDRs), which are essential for antigen recognition and binding.

The antigens in the dataset are represented as amino acid sequences of the HA protein. The amino acid sequences for 79 % of antigens were obtained from publicly available databases including NCBI GenBank and GISAID. The remaining 21 % of antigens were derived from experimental results conducted against a computationally optimized broadly reactive antigen (COBRA) HA. COBRA antigens are specifically designed to trigger broader reactivity than wild-type antigens by addressing the variability among different seasonal and/or pandemic HA strains in a single antigen [76]. In particular, the COBRA method uses multiple rounds of consensus building based upon not only the phylogenetic sequence of each isolate, but also the outbreak and specific time that each isolate was collected, thereby eliminating the bias in the number of sequences uploaded to online databases.

The following COBRA HA were included: H2_COBRA.Z7 [77], H3_NG5 [72], H3_NG7 [74], H5_COBRA2 [78], H1N1 COBRA P1 [79], H1N1 COBRA X3 [75], H1N1 COBRA X6 [80], H3N2 COBRA NG2 [81], H1N1 COBRA Y2 [82], H3N2 COBRA T10 [83], and H3N2 COBRA T11 [84]. Additional metadata available for antibodies includes the isotype

of the light and heavy chains, as well as epitope information.

3.3. Data splits

We assessed model robustness through five-fold cross validation across four distinct data splitting strategies: lenient, HA-exclusive, mAb-exclusive, and mAb-cluster exclusive. In this approach, the dataset is divided into five equal-sized subsets (folds), and the model is trained on four folds and tested on the remaining fold. This process is repeated five times, ensuring that each fold serves as the test set exactly once. The final model performance is reported as the average across all five test results. In the lenient split, mAb-HA pairs were randomly distributed between folds. For the HA-exclusive split, we ensured that all pairs containing the same HA sequence were assigned to the same fold, preventing the training and test sets from sharing any HA sequence. The mAb-exclusive split followed a similar principle, keeping all pairs involving the same mAb within a single fold. In the mAb-cluster exclusive split, we grouped mAb-HA pairs based on antibody clusters, which were determined using MMseqs2 [85] with a minimum sequence identity threshold of 50 %. We generated these four splitting strategies separately for each label type (binding activity and HAI), maintaining consistent proportions of positive antibody-HA pair labels across all folds.

3.4. Binding and HAI classification models

We developed predictive models using MAMMAL [50], a recently published biomolecular foundation model framework. MAMMAL is part of IBM's Biomedical Foundation Modeling (BMFM) technology suite.[2] Specifically, we used the model ibm/biomed.omics.bl.sm.ma-ted-458 m, which was trained on extensive multi-domain data, including proteins (UniProt [86]), antibodies(OAS [87]), and protein-protein interactions (STRING [88]) via various self-supervised tasks. For brevity, we refer to this foundation model as MAMMAL. The model code and pretrained weights are publicly available at https://github.com/BiomedSciAI/biomed-multi-alignment and https://huggingface.co/ibm/biomed.omics.bl.sm.ma-ted-458m, respectively. The developed model takes a pair of antibody and antigen sequences as input, as represented by their amino acid sequences and using the prompt syntax of the AbAg Bind task described in past work [50]. Model training and evaluation were conducted using the FuseMedML framework [89].

For each task and training fold, we fine-tuned MAMMAL with the default hyperparameters outlined in [50]: The AdamW optimizer with $\beta1 = 0.9$ and $\beta2 = 0.999$, weight decay of 0.01, and gradient clipping with a norm of 1.0. We employed 2 K warm-up steps to reach the maximum learning rate, followed by a cosine decay scheduler that reduces the learning rate to 10 % of the maximum by the end of training. The maximum input length was set to 900, ensuring input sequences are not truncated. For all models, training involved 10 K iterations on a V100–32G GPU using batch sizes of 8.

3.5. Model evaluation

We evaluated the models' classification performance using the Area Under the Receiver Operating Characteristic Curve (AUROC) and the Area Under the Precision-Recall Curve (AUPRC). Both metrics have values ranging between 0 and 1. For a random classifier, the expected AUROC is 0.5; the expected AUPRC for a random classifier corresponds to the rate of positivity in the dataset [90]. For each five-fold cross-validation experiment, we report average performance metrics along with their standard deviations across the five folds.

[2] https://research.ibm.com/projects/biomedical-foundation-models

4. Results

4.1. Data statistics

The classification dataset consists of 188 mAbs and 79 HAs. The data include 4922 unique mAb-HA pairs from 176 mAbs and 59 HAs in the binding activity assays and 5035 pairs from 186 mAbs and 59 HAs in the HAI assays. Among these, 3188 mAb-HA pairs are shared between the binding activity and HAI assays, involving 174 mAbs and 39 HAs. Positive pairs account for 35 % of the binding dataset and 11 % of the HAI dataset. On average, each mAb is involved in 28 ± 17 (mean +/- standard deviation) binding assays, while each antigen is included in 83 ± 45 binding assays. For HAI assays, each mAb is involved in 27 ± 17 assays, and each antigen is included in 85 ± 41 assays.

Tables 1 and 2 summarize the data characteristics and statistical associations with positive outcomes for the binding and HAI datasets, respectively. The majority of assays in the datasets involve antibodies (82–85 %) that originated from human hosts. H1N1 and H3N2 subtypes dominate, representing 88–95 % of the data. The mean lengths of the variable regions in the LC and HC sequences of the mAbs, determined using AbNumber [29], are 122 ± 8 and 108 ± 4, respectively. The mean length of antigen sequences is 559 ± 20.

Clustering the 176 mAbs in the binding dataset and the 186 mAbs in the HAI dataset yielded 76 clusters in total for each dataset, with cluster sizes ranging from 1 to 9. The mean cluster size was 2.2 ± 2.0 in the binding dataset and 2.4 ± 2.1 in the HAI dataset. In the binding activity dataset, the proportion of positive pairs ranged from 28 % to 41 % across all folds in the four splits (20 folds in total). For the HAI dataset, this proportion ranged from 7 % to 17 %.

4.2. Prediction of binding activity and hemagglutination inhibition

We conducted eight experiments using five-fold cross-validation, corresponding to two classification outcomes—binding and HAI—and four data splits (lenient, HA-exclusive, mAb-exclusive, and mAb-cluster-exclusive). Each experiment was repeated twice: once with random weight initialization (the "random-initialization" model) and once with initialization using MAMMAL weights (the "MAMMAL-finetuned" model). The evaluation metrics for these experiments are presented in Fig. 2 and Table 4 in the appendix. Both the random-initialization and MAMMAL-finetuned models exhibited performance metrics significantly higher than random, where a random classifier is expected to produce an AUROC of 0.5 for both datasets and AUPRCs of 0.35 and 0.11 for the binding and HAI datasets, respectively.

Table 1
Binding activity dataset characteristics and their association with positive binding classification.

		Overall	Positive, N (%)	P-Value
Total		4922	1740	
mAb host, n (%)	**Human**	4181 (84.9)	1357 (78.0)	< 0.0001
	Mouse	741 (15.1)	383 (22.0)	
HA year, n (%)	**2000–2010**	1336 (27.1)	458 (26.3)	0.02
	< 1950	903 (18.3)	291 (16.7)	
	> 2010	2190 (44.5)	795 (45.7)	
	Other/unknown	493 (10.0)	196 (11.3)	
HA subtype, n (%)	**H1**	2366 (48.1)	920 (52.9)	< 0.0001
	H2	53 (1.1)	14 (0.8)	
	H3	2210 (44.9)	740 (42.5)	
	H5	278 (5.6)	64 (3.7)	
	H7	15 (0.3)	2 (0.1)	
mAb LC ISO, n (%)	**kappa**	2806 (57.0)	1198 (68.9)	< 0.0001
	lambda	2116 (43.0)	542 (31.1)	
mAb HC ISO, n (%)	**IgA**	490 (10.0)	213 (12.2)	< 0.0001
	IgG	4432 (90.0)	1527 (87.8)	
mAb Epitope, n (%)	**Conformational**	2493 (50.7)	844 (48.5)	0.03
	Other/unknown	2429 (49.3)	896 (51.5)	

Table 2
HAI activity dataset characteristics and their association with positive HAI classification.

		All,N(%)	Positive HAI, N (%)	P-Value
Total		5035	572	
mAb host, n (%)	**Human**	4112 (81.7)	439 (76.7)	0.002
	Mouse	923 (18.3)	133 (23.3)	
HA year, n (%)	**2000–2010**	1093 (21.7)	140 (24.5)	< 0.0001
	< 1950	1851 (36.8)	72 (12.6)	
	> 2010	1761 (35.0)	322 (56.3)	
	Other/unknown	330 (6.6)	38 (6.6)	
HA subtype, n (%)	**H1**	3048 (60.5)	233 (40.7)	< 0.0001
	H3	1832 (36.4)	339 (59.3)	
	H5	155 (3.1)		
mAb LC ISO, n (%)	**kappa**	2955 (58.7)	263 (46.0)	< 0.0001
	lambda	2080 (41.3)	309 (54.0)	
mAb HC ISO, n (%)	**IgA**	462 (9.2)	53 (9.3)	0.998
	IgG	4573 (90.8)	519 (90.7)	
mAb Epitope, n (%)	**Conformational**	2517 (50.0)	257 (44.9)	0.01
	Other/unknown	2518 (50.0)	315 (55.1)	

Comparing the two models, the MAMMAL-finetuned model consistently outperformed the random-initialization model, except for the AUROC metric on the mAb-cluster-exclusive split for the HAI classification task. However, this difference fell within the range of the estimated standard errors (SEs), calculated as the standard deviation divided by the square root of 5, and was therefore not statistically significant ($P = 0.43$, one-sided *t*-test). Given that the HAI dataset is imbalanced (with only an 11 % positivity rate), the AUPRC metric may be more relevant [90]. A significant difference was observed in AUPRC, with the MAMMAL-finetuned model showing superior performance. For the MAMMAL-finetuned model in isolation, performance was higher in the lenient and HA-exclusive splits but lower in the mAb-exclusive and mAb-cluster-exclusive splits. These findings suggest that the model generalizes more effectively to unseen HA sequences than to unseen mAb sequences.

We assessed the robustness of the model's AUROC measures across various HA subgroups. As shown in Fig. 3, the trend of higher performance in the lenient and HA-exclusive splits, coupled with reduced performance in the mAb-exclusive splits, was consistent across all subgroups. Performance within each subgroup remained largely stable, with the exception of poorer results on (1) binding prediction for mouse antibodies and (2) HAI prediction for HAs from earlier years in the mAb-exclusive splits.

4.3. *Prediction of antibody breadth of protection*

We assessed the models' ability to predict antibody breadth for binding activity and HAI across the HA subtypes H1 and H3. The analysis was conducted separately for each HA subtype (H1 and H3) and assay type (binding activity and HAI). Antibodies featured in fewer than five assays were filtered out. For each antibody, we measured its breadth of protection by aggregating assay results and calculating the proportion of positive assays. We then computed a prediction score for broad protection by averaging the predictions from validation folds across all assays in which the antibody appears. Only antibody-exclusive splits were considered, as the analysis aimed to predict the broad protection of unseen antibodies.

To evaluate the predictive power of the aggregated model scores, we calculated the Pearson correlation between these scores and the antibodies' rates of positive assay results. Additionally, we calculated the AUROC for scores associated with at least 30 % positive assays. Table 3 presents antibody statistics for each assay type (binding activity or HAI) and HA subtype (H1 and H3), along with the Pearson correlation and AUROC metrics.

5. Discussion

In this study, we developed and evaluated AI models to predict antibody binding and HAI activity on influenza A HA using antibody HC and LC variable region sequences as input. Our approach leveraged MAMMAL, a language model pre-trained on extensive biomedical datasets, which we fine-tuned using laboratory-derived antibody-HA binding and HAI assay data. We evaluated model performance through comprehensive cross-validation analyses, employing multiple split paradigms to simulate diverse real-world scenarios and assess the robustness of our models.

To address the computational challenges posed by high-dimensional sequence data, we employed transformer architectures [91], which

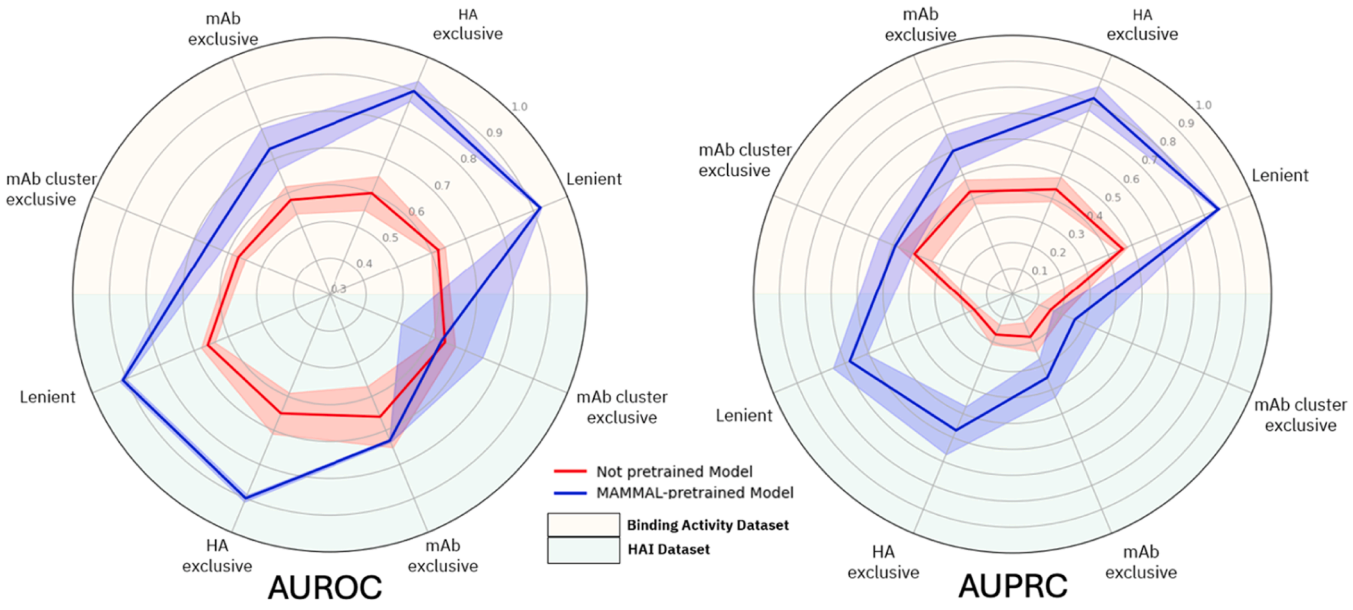

Fig. 2. Evaluation of binding activity and hemagglutination inhibition (HAI) classification models under various experimental conditions using the AUROC and AUPRC metrics. Performance of fine-tuned MAMMAL models (blue) is compared to randomly initialized models (red). Solid lines represent the mean, while shaded areas indicate the standard deviation across the five-fold cross-validation.

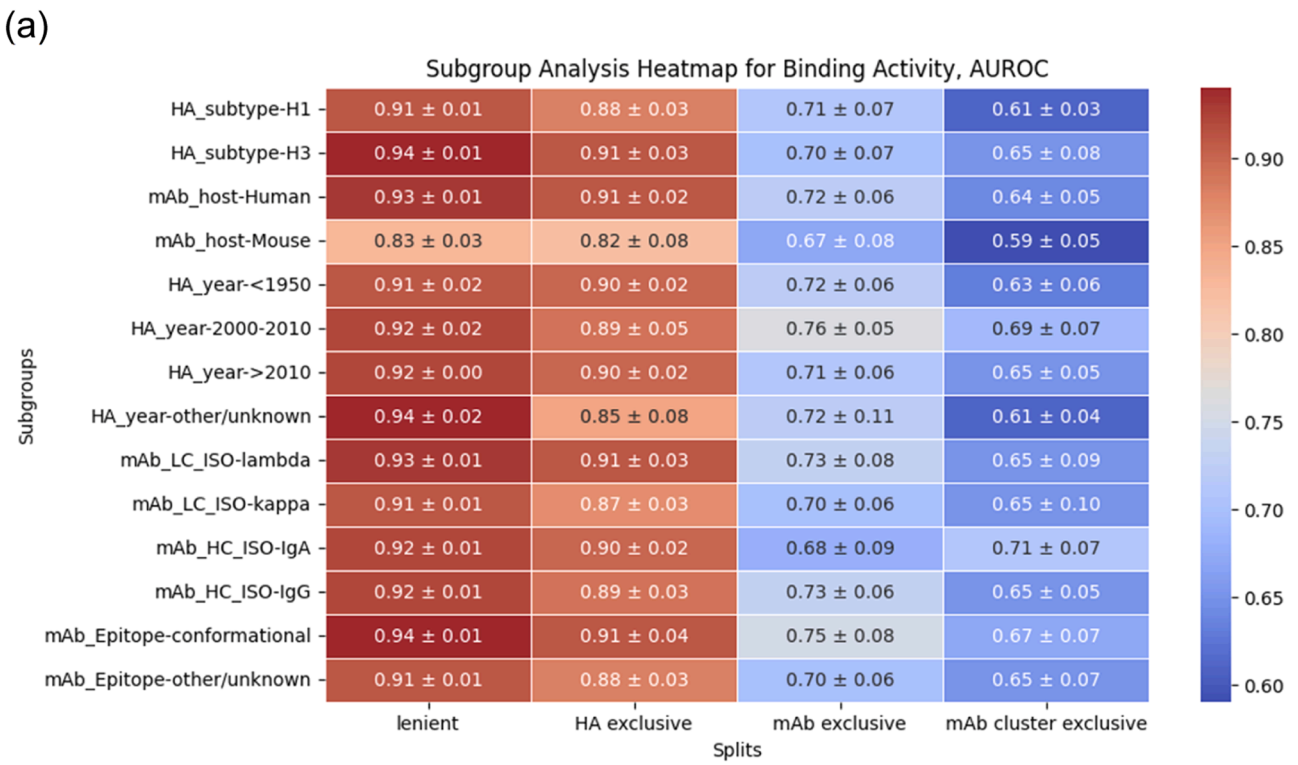

Fig. 3. Subgroup analysis for binding activity (a) and HAI (b). The AUROC metric was calculated as an average over five folds of predictions for each of the subgroups. This metric is applied to each of the four splits evaluated in this study. Subgroup names combine a data characteristic name (e.g., *HA_subtype*) with its category value (e.g., *H1*).

excel at capturing complex patterns in protein sequences data [27,28,92]. The superior performance of models initialized with MAMMAL weights compared to those with random initialization underscores the value of transfer learning from large-scale pre-trained biomedical models, particularly when working with limited task-specific data. These results underscore the transformative potential of pretraining on large-scale protein and antibody datasets to boost model performance on small, specialized datasets.

The high predictive performance (AUROC $>$ 0.9) of our fine-tuned MAMMAL models under lenient split conditions highlights those models' ability to generalize across randomly held-out antibody-HA pairs. This result implies potential utility in reducing experimental

Table 3
Analysis of antibody breadth of protection: Statistics and prediction metrics.

Task	HA subtype	mAb split	#mAbs	Pearson	p-value	Positive assays rate > 0.3	AUROC
Binding activity prediction	H1	Exclusive	145	0.45	2.E−08	59 %	0.73
		Cluster exclusive	145	0.25	0.003	64 %	0.69
	H3	Exclusive	101	0.32	0.001	54 %	0.68
		Cluster exclusive	101	0.25	0.01	40 %	0.62
HAI prediction	H1	Exclusive	142	0.26	0.001	5 %	0.73
		Cluster exclusive	142	0.24	0.005	5 %	0.64
	H3	Exclusive	110	0.49	6.E−08	28 %	0.72
		Cluster exclusive	110	0.14	0.1	34 %	0.53

workloads by predicting outcomes for untested combinations within known sequence spaces. Similarly, the strong performance (AUROC = 0.9) in the HA-exclusive split scenario demonstrates reliable prediction capabilities for novel HA sequences against previously analyzed antibodies, offering support for strain surveillance and antibody evaluation against emerging variants.

The more challenging antibody-exclusive split scenario, relevant for novel antibody design, showed moderate performance (AUROC = 0.73) for both binding activity and HAI prediction. This decrease in performance compared to other scenarios suggests current limitations in generalizing predictions to entirely new antibody sequences. The further reduction in performance under the HA-cluster-exclusive split (AUROC = 0.66 for binding, 0.63 for HAI) highlights particular challenges in extrapolating to more divergent HA sequences. Despite these limitations, the models showed promising results toward identifying broadly protective antibodies, especially against H3 subtypes (AUROC 0.64–0.73). Performance limitations likely arise because of the scarcity of characterized mAbs in datasets used in this study.

A key factor to improving model performance regards the expansion of training datasets through automated collection and curation of influenza neutralization assays from public repositories. Training on larger datasets with more diverse antibody sequence data is likely needed to enhance model generalization, particularly for novel antibody sequences. In this work, considering the extensive panel of antigens (HA) and influenza viruses used to test the mAbs, binding and HAI activity were the only assays that could provide a broad dataset for each mAb and that could be leveraged to build this model. Future work will focus on incorporating a larger number of mAb sequences and datasets, including those from publicly available antibody databases (e.g., SAbDab [93], PLAbDab [94], and IEDB [95]). Additionally, other biological datasets will be included such as affinity measurements (Kd values), *in vitro* and *in vivo* potency of neutralization and protection, as well as antibody-antigen structural data. While these data may be limited and include a restricted number of antibody-antigen combinations, they may be important for refining and enhancing the overall predictive accuracy of the model.

However, incorporating data from other groups may pose challenges due to variations in methodologies used to determine mAb biological activities (e.g., binding and HAI activity) and inconsistencies in units of measurement. Laboratory validation of in silico predictions is essential to advance AI models' performance on novel antibody-HA pairs. Such validation can also provide feedback to inform the training process and support an iterative approach to developing more robust and accurate predictive frameworks. Assessing model accuracy on new antibodies from antibody repertoire sequences and in silico design of novel antibodies will be the focus of future studies.

6. Conclusion

Our findings demonstrate the potential of fine-tuned language models for predicting antibody-HA interactions across various practical scenarios. While performance on novel antibodies, particularly those divergent from training data, remains an area for improvement, these models hold significant promise for accelerating influenza research and antibody design. Integration with computational antibody design pipelines, including AI-based systems, could enable rapid in silico assessment of candidate antibodies and significantly expedite the antibody design process.

Data and code Availability

The antibody-HA datasets are provided in the supplemental material, which also includes contact points for obtaining certain antibody and antigen sequences. The pretrained model weights and tokenizer are available on the Hugging Face model hub at https://huggingface.co/ibm/biomed.omics.bl.sm.ma-ted-458m. The model architecture and fine-tuning framework can be accessed at https://github.com/BiomedSciAI/biomed-multi-alignment, which provides detailed instructions and examples for fine-tuning and inference. The models will be released upon obtaining legal approval.

CRediT authorship contribution statement

Siddiqui Ibrahim: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Cheng Kevin J.:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology. **Golts Alex:** Writing – review & editing. **Shoshan Yoel:** Writing – review & editing, Software, Methodology, Investigation. **Barkan Ella:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Weber Jeffrey K.:** Writing – review & editing, Supervision. **Campos Mota Yailin:** Formal analysis, Data curation. **Ozery-Flato Michal:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Conceptualization. **Sautto Giuseppe Andrea:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declarations of Competing Interest

We declare no conflict of interest.

Acknowledgement

This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID), a component of the U.S. National Institutes of Health (NIH), Department of Health and Human Services, under contract 75N93019C00052, by the Cleveland Clinic Foundation and by the Discovery Accelerator Cleveland Clinic and IBM partnership.

Appendix

Table 4
Evaluation of binding and HAI classification models. For each experiment (task + data split) and metric, we compare random weight initialization ("random-initialization") with MAMMAL weight initialization ("MAMMAL-finetuned"). Bolded values indicate the higher performance for each comparison

Task	Data Split	Random-Initialization Model		MAMMAL-finetuned Model	
		AUROC	AUPRC	AUROC	AUPRC
Binding Prediction	Lenient	0.62 ± 0.02	0.46 ± 0.02	**0.92** ± 0.004	**0.86** ± 0.007
	HA exclusive	0.60 ± 0.05	0.44 ± 0.05	**0.90** ± 0.03	**0.82** ± 0.05
	mAb exclusive	0.58 ± 0.04	0.43 ± 0.05	**0.73** ± 0.06	**0.60** ± 0.07
	mAb cluster exclusive	0.57 ± 0.02	0.41 ± 0.07	**0.66** ± 0.04	**0.49** ± 0.07
HAI Prediction	Lenient	0.66 ± 0.02	0.16 ± 0.02	**0.91** ± 0.009	**0.68** ± 0.07
	HA exclusive	0.65 ± 0.06	0.17 ± 0.04	**0.90** ± 0.01	**0.57** ± 0.10
	mAb exclusive	0.66 ± 0.09	0.18 ± 0.06	**0.73** ± 0.04	**0.35** ± 0.08
	mAb cluster exclusive	**0.64** ± 0.03	0.16 ± 0.05	0.63 ± 0.12	**0.26** ± 0.09

References

[1] Centers for Disease Control and Prevention (CDC). CDC Yellow Book 2024: Health Information for International Travel. Oxford University Press; 2023.
[2] Flu (Influenza): Causes, Symptoms, Types & Treatment. Cleveland Clinic n.d. https://my.clevelandclinic.org/health/diseases/4335-influenza-flu (accessed January 4, 2025).
[3] Low ZY, Wong KH, Wen Yip AJ, Choo WS. The convergent evolution of influenza A virus: implications, therapeutic strategies and what we need to know. Curr Res Micro Sci 2023;5:100202. https://doi.org/10.1016/j.crmicr.2023.100202.
[4] Fischer WA, Gong M, Bhagwanjee S, Sevransky J. Global burden of influenza: contributions from resource limited and low-income settings. Glob Heart 2014;9:325–36. https://doi.org/10.1016/j.gheart.2014.08.004.
[5] Wu NC, Wilson IA. Structural biology of influenza hemagglutinin: an amaranthine adventure. Viruses 2020;12:1053. https://doi.org/10.3390/v12091053.
[6] Sautto GA, Ross TM. Respiratory viruses: preventive and therapeutic approaches to diverse pathogens. CCJM 2024;91:S4–10. https://doi.org/10.3949/ccjm.91.s1.02.
[7] Sun X, Ma H, Wang X, Bao Z, Tang S, Yi C, et al. Broadly neutralizing antibodies to combat influenza virus infection. Antivir Res 2024;221:105785. https://doi.org/10.1016/j.antiviral.2023.105785.
[8] Burioni R, Canducci F, Mancini N, Clementi N, Sassi M, De Marco D, et al. Monoclonal antibodies isolated from human B cells neutralize a broad range of H1 subtype influenza A viruses including swine-origin Influenza virus (S-OIV). Virology 2010;399:144–52. https://doi.org/10.1016/j.virol.2009.12.014.
[9] Burioni R, Canducci F, Mancini N, Clementi N, Sassi M, De Marco D, et al. Molecular cloning of the first human monoclonal antibodies neutralizing with high potency swine-origin influenza A pandemic virus (S-OIV). N Microbiol 2009;32:319–24.
[10] Solforosi L., Mancini N., Canducci F., Clementi N., Sautto G.A., Diotti R.A., et al. A phage display vector optimized for the generation of human antibody combinatorial libraries and the molecular cloning of monoclonal antibody fragments n.d.
[11] Sautto GA, Kirchenbaum GA, Abreu RB, Ecker JW, Pierce SR, Kleanthous H, et al. A Computationally optimized broadly reactive antigen subtype–specific influenza vaccine strategy elicits unique potent broadly neutralizing antibodies against hemagglutinin. J Immunol 2020;204:375–85. https://doi.org/10.4049/jimmunol.1900379.
[12] Skarlupka AL, Reneer ZB, Abreu RB, Ross TM, Sautto GA. An Influenza virus hemagglutinin computationally optimized broadly reactive antigen elicits antibodies endowed with group 1 heterosubtypic breadth against swine influenza viruses. J Virol 2020;94:e02061-19. https://doi.org/10.1128/JVI.02061-19.
[13] Forgacs D, Abreu RB, Sautto GA, Kirchenbaum GA, Drabek E, Williamson KS, et al. Convergent antibody evolution and clonotype expansion following influenza virus vaccination. PLOS ONE 2021;16:e0247253. https://doi.org/10.1371/journal.pone.0247253.
[14] Sautto GA, Ecker JW, Ross TM. An H1N1 Computationally Optimized Broadly Reactive Antigen Elicits a Neutralizing Antibody Response against an Emerging Human-Infecting Eurasian Avian-Like Swine Influenza Virus. 10.1128/jvi.02421-20 J Virol 2021;95. https://doi.org/10.1128/jvi.02421-20.
[15] Gadush MV, Sautto GA, Chandrasekaran H, Bensussan A, Ross TM, Ippolito GC, et al. Template-assisted de novo sequencing of SARS-CoV-2 and influenza monoclonal antibodies by mass spectrometry. J Proteome Res 2022;21:1616–27. https://doi.org/10.1021/acs.jproteome.1c00913.
[16] Nagashima K, Dzimianski JV, Han J, Abbadi N, Gingerich AD, Royer F, et al. The pre-existing human antibody repertoire to computationally optimized influenza H1 hemagglutinin vaccines. J Immunol 2022;209:5–15. https://doi.org/10.4049/jimmunol.2101171.
[17] Dzimianski JV, Han J, Sautto GA, O'Rourke SM, Cruz JM, Pierce SR, et al. Structural insights into the broad protection against H1 influenza viruses by a computationally optimized hemagglutinin vaccine. Commun Biol 2023;6:1–13. https://doi.org/10.1038/s42003-023-04793-3.
[18] Nagashima KA, Dzimianski JV, Yang M, Abendroth J, Sautto GA, Ross TM, et al. Structural basis for the broad antigenicity of the computationally optimized influenza hemagglutinin X6. Structure 2024;32:1079–1089.e6. https://doi.org/10.1016/j.str.2024.05.001.
[19] Raju N, Kramer KJ, Cavallaro M, Diotti RA, Shiakolas AR, Campos Mota Y, et al. Multiplexed antibody sequencing and profiling of the human hemagglutinin-specific memory B cell response following Influenza Vaccination. J Immunol 2024;213:1605–19. https://doi.org/10.4049/jimmunol.2400326.
[20] Samaranayake H, Wirth T, Schenkwein D, Räty JK, Ylä-Herttuala S. Challenges in monoclonal antibody-based therapies. Ann Med 2009;41:322–31. https://doi.org/10.1080/07853890802698842.
[21] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596:583–9. https://doi.org/10.1038/s41586-021-03819-2.
[22] Evans R., O'Neill M., Pritzel A., Antropova N., Senior A., Green T., et al., Protein complex prediction with AlphaFold-Multimer 2022:2021.10.04.463034. https://doi.org/10.1101/2021.10.04.463034.
[23] Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w.
[24] Yin R, Pierce BG. Evaluation of AlphaFold antibody–antigen modeling with implications for improving predictive accuracy. Protein Sci 2024;33:e4865. https://doi.org/10.1002/pro.4865.
[25] Wallner B. AFsample: improving multimer prediction with AlphaFold using massive sampling. Bioinformatics 2023;39:btad573. https://doi.org/10.1093/bioinformatics/btad573.
[26] Rives A, Meier J, Sercu T, Goyal S, Lin Z, Liu J, et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. Proc Natl Acad Sci 2021;118:e2016239118. https://doi.org/10.1073/pnas.2016239118.
[27] Meier J., Rao R., Verkuil R., Liu J., Sercu T., Rives A. Language models enable zero-shot prediction of the effects of mutations on protein function 2021. https://doi.org/10.1101/2021.07.09.450648.
[28] Elnaggar A., Heinzinger M., Dallago C., Rihawi G., Wang Y., Jones L., et al. ProtTrans: Towards Cracking the Language of Life's Code Through Self-Supervised Deep Learning and High Performance Computing 2021. https://doi.org/10.48550/arXiv.2007.06225.
[29] Olsen TH, Moal IH, Deane CM. AbLang: an antibody language model for completing antibody sequences. Bioinforma Adv 2022;2:vbac046. https://doi.org/10.1093/bioadv/vbac046.
[30] Kuan D., Farimani A.B. AbGPT: De Novo Antibody Design via Generative Language Modeling 2024. https://doi.org/10.48550/arXiv.2409.06090.
[31] Xu X, Xu T, Zhou J, Liao X, Zhang R, Wang Y, et al. AB-Gen: antibody library design with generative pre-trained transformer and deep reinforcement learning. Genom, Proteom Bioinforma 2023;21:1043–53. https://doi.org/10.1016/j.gpb.2023.03.004.
[32] Kenlay H, Dreyer FA, Kovaltsuk A, Miketa D, Pires D, Deane CM. Large scale paired antibody language models. PLOS Comput Biol 2024;20:e1012646. https://doi.org/10.1371/journal.pcbi.1012646.
[33] Kim J, McFee M, Fang Q, Abdin O, Kim PM. Computational and artificial intelligence-based methods for antibody development. Trends Pharm Sci 2023;44:175–89. https://doi.org/10.1016/j.tips.2022.12.005.
[34] Hie BL, Shanker VR, Xu D, Bruun TUJ, Weidenbacher PA, Tang S, et al. Efficient evolution of human antibodies from general protein language models. Nat Biotechnol 2024;42:275–83. https://doi.org/10.1038/s41587-023-01763-2.
[35] Shanker VR, Bruun TUJ, Hie BL, Kim PS. Unsupervised evolution of protein and antibody complexes with a structure-informed language model. Science 2024;385:46–53. https://doi.org/10.1126/science.adk8946.
[36] Jespers W, Åqvist J, Gutiérrez-de-Terán H. Free Energy Calculations for Protein–Ligand Binding Prediction. In: Ballante F, editor. Protein-Ligand Interactions and Drug Design. New York, NY: Springer US; 2021. p. 203–26. https://doi.org/10.1007/978-1-0716-1209-5_12.

[37] Wang L, Chambers J, Abel R. Protein–Ligand Binding Free Energy Calculations with FEP+. In: Bonomi M, Camilloni C, editors. Methods and Protocols. Biomolecular Simulations. New York, NY: Springer; 2019. p. 201–32. https://doi.org/10.1007/978-1-4939-9608-7_9.

[38] Ross GA, Lu C, Scarabelli G, Albanese SK, Houang E, Abel R, et al. The maximal and current accuracy of rigorous protein-ligand binding free energy calculations. Commun Chem 2023;6:1–12. https://doi.org/10.1038/s42004-023-01019-9.

[39] Pantsar T, Poso A. Binding affinity via docking: fact and fiction. Molecules 2018; 23:1899. https://doi.org/10.3390/molecules23081899.

[40] Ballester PJ, Mitchell JBO. A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. Bioinformatics 2010;26:1169–75. https://doi.org/10.1093/bioinformatics/btq112.

[41] Xiong P, Zhang C, Zheng W, Zhang Y. BindProfX: Assessing Mutation-Induced Binding Affinity Change by Protein Interface Profiles with Pseudo-Counts. J Mol Biol 2017;429:426–34. https://doi.org/10.1016/j.jmb.2016.11.022.

[42] Meli R, Morris GM, Biggin PC. Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: a review. Front Bioinform 2022;2. https://doi.org/10.3389/fbinf.2022.885983.

[43] DeepAffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks | Bioinformatics | Oxford Academic n.d. https://academic.oup.com/bioinformatics/article/35/18/3329/5320555 (accessed January 28, 2025).

[44] DeepDTA: deep drug–target binding affinity prediction | Bioinformatics | Oxford Academic n.d. https://academic.oup.com/bioinformatics/article/34/17/i821/5093245 (accessed January 28, 2025).

[45] Lee I, Keum J, Nam H. DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences. PLOS Comput Biol 2019;15: e1007129. https://doi.org/10.1371/journal.pcbi.1007129.

[46] Huang Y, Zhang Z, Zhou Y. AbAgIntPre: a deep learning method for predicting antibody-antigen interactions based on sequence information. Front Immunol 2022;13. https://doi.org/10.3389/fimmu.2022.1053617.

[47] Zhang J, Du Y, Zhou P, Ding J, Xia S, Wang Q, et al. Predicting unseen antibodies' neutralizability via adaptive graph neural networks. Nat Mach Intell 2022;4: 964–76. https://doi.org/10.1038/s42256-022-00553-w.

[48] Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science 2023;379:1123–30. https://doi.org/10.1126/science.ade2574.

[49] Brandes N, Ofer D, Peleg Y, Rappoport N, Linial M. ProteinBERT: a universal deep-learning model of protein sequence and function. Bioinformatics 2022;38: 2102–10. https://doi.org/10.1093/bioinformatics/btac020.

[50] Shoshan Y., Raboh M., Ozery-Flato M., Ratner V., Golts A., Weber J.K., et al. MAMMAL – Molecular Aligned Multi-Modal Architecture and Language 2024. https://doi.org/10.48550/arXiv.2410.22367.

[51] Choi Y. Artificial intelligence for antibody reading comprehension: AntiBERTa. PATTER 2022;3. https://doi.org/10.1016/j.patter.2022.100535.

[52] Yuan Y, Chen Q, Mao J, Li G, Pan X. DG-Affinity: predicting antigen–antibody affinity with language models from sequences. BMC Bioinforma 2023;24:430. https://doi.org/10.1186/s12859-023-05562-z.

[53] Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science 2021;373:871–6. https://doi.org/10.1126/science.abj8754.

[54] Hitawala F.N., Gray J.J. What has AlphaFold3 learned about antibody and nanobody docking, and what remains unsolved? 2025:2024.09.21.614257. https://doi.org/10.1101/2024.09.21.614257.

[55] Gao M, Skolnick J. Improved deep learning prediction of antigen–antibody interactions. Proc Natl Acad Sci 2024;121:e2410529121. https://doi.org/10.1073/pnas.2410529121.

[56] Zeng D, Xin J, Yang K, Guo S, Wang Q, Gao Y, et al. A Hemagglutinin Stem Vaccine Designed Rationally by AlphaFold2 Confers Broad Protection against Influenza B Infection. Viruses 2022;14:1305. https://doi.org/10.3390/v14061305.

[57] He H, He B, Guan L, Zhao Y, Jiang F, Chen G, et al. De novo generation of SARS-CoV-2 antibody CDRH3 with a pre-trained generative large language model. Nat Commun 2024;15:6867. https://doi.org/10.1038/s41467-024-50903-y.

[58] Acar DD, Witkowski W, Wejda M, Wei R, Desmet T, Schepens B, et al. Integrating artificial intelligence-based epitope prediction in a SARS-CoV-2 antibody discovery pipeline: caution is warranted. EBioMedicine 2024;100:104960. https://doi.org/10.1016/j.ebiom.2023.104960.

[59] Loomis CM, Lahlali T, Van Citters D, Sprague M, Neveu G, Somody L, et al. AI-based antibody discovery platform identifies novel, diverse, and pharmacologically active therapeutic antibodies against multiple SARS-CoV-2 strains. Antib Ther 2024;7:307–23. https://doi.org/10.1093/abt/tbae025.

[60] Wang Y, Lv H, Teo QW, Lei R, Gopal AB, Ouyang WO, et al. An explainable language model for antibody specificity prediction using curated influenza hemagglutinin antibodies. Immunity 2024;57:2453–2465.e7. https://doi.org/10.1016/j.immuni.2024.07.022.

[61] Wang M, Patsenker J, Li H, Kluger Y, Kleinstein SH. Supervised fine-tuning of pre-trained antibody language models improves antigen specificity prediction. bioRxiv 2024:2024.05.13.593807. https://doi.org/10.1101/2024.05.13.593807.

[62] Annual and interim reports. Crick 2025. https://www.crick.ac.uk/research/platforms-and-facilities/worldwide-influenza-centre/annual-and-interim-reports (accessed January 20, 2025).

[63] Shah SAW, Palomar DP, Barr I, Poon LLM, Quadeer AA, McKay MR. Seasonal antigenic prediction of influenza A H3N2 using machine learning. Nat Commun 2024;15:3833. https://doi.org/10.1038/s41467-024-47862-9.

[64] Yao Y, Li X, Liao B, Huang L, He P, Wang F, et al. Predicting influenza antigenicity from Hemagglutintin sequence data based on a joint random forest method. Sci Rep 2017;7:1545. https://doi.org/10.1038/s41598-017-01699-z.

[65] Spatial exclusion leads to. tug-of-war" ecological dynamics between competing species within microchannels | PLOS Computational Biology n.d. https://journals.plos.org/ploscompbiol/article?id= 10.1371/journal.pcbi.1010868 (accessed January 20, 2025).

[66] Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses | PNAS n.d. https://www.pnas.org/doi/10.1073/pnas.1525578113 (accessed January 20, 2025).

[67] Harvey WT, Benton DJ, Gregory V, Hall JPJ, Daniels RS, Bedford T, et al. Identification of low- and high-impact hemagglutinin amino acid substitutions that drive antigenic drift of influenza A(H1N1) viruses. PLoS Pathog 2016;12: e1005526. https://doi.org/10.1371/journal.ppat.1005526.

[68] Vukovich MJ, Shiakolas AR, Lindenberger J, Richardson RA, Bass LE, Barr M, et al. Isolation and characterization of IgG3 glycan-targeting antibodies with exceptional cross-reactivity for diverse viral families. PLOS Pathog 2024;20:e1012499. https://doi.org/10.1371/journal.ppat.1012499.

[69] Ross TM, Gokanapudi N, Ge P, Shi H, Richardson RA, Pierce SR, et al. Kinetic of the Antibody Response Following AddaVax-Adjuvanted Immunization with Recombinant Influenza Antigens. Vaccines 2022;10:1315. https://doi.org/10.3390/vaccines10081315.

[70] Pilewski KA, Wall S, Richardson SI, Manamela NP, Clark K, Hermanus T, et al. Functional HIV-1/HCV cross-reactive antibodies isolated from a chronically co-infected donor. Cell Rep 2023;42. https://doi.org/10.1016/j.celrep.2023.112044.

[71] Ecker JW, Kirchenbaum GA, Pierce SR, Skarlupka AL, Abreu RB, Cooper RE, et al. High-yield expression and purification of recombinant influenza virus proteins from stably-transfected mammalian cell lines. Vaccines 2020;8:462. https://doi.org/10.3390/vaccines8030462.

[72] Carter DM, Darby CA, Lefoley BC, Crevar CJ, Alefantis T, Oomen R, et al. Design and Characterization of a Computationally Optimized Broadly Reactive Hemagglutinin Vaccine for H1N1 Influenza Viruses. J Virol 2016;90:4720–34. https://doi.org/10.1128/jvi.03152-15.

[73] Carlock MA, Allen JD, Hanley HB, Ross TM. Longitudinal assessment of human antibody binding to hemagglutinin elicited by split-inactivated influenza vaccination over six consecutive seasons. PLoS One 2024;19:e0301157. https://doi.org/10.1371/journal.pone.0301157.

[74] Wong TM, Allen JD, Bebin-Blackwell A-G, Carter DM, Alefantis T, DiNapoli J, et al. Computationally optimized broadly reactive hemagglutinin elicits hemagglutination inhibition antibodies against a panel of H3N2 influenza virus cocirculating variants. 10.1128/jvi.01581-17 J Virol 2017;91. https://doi.org/10.1128/jvi.01581-17.

[75] Sautto GA, Kirchenbaum GA, Ecker JW, Bebin-Blackwell A-G, Pierce SR, Ross TM. Elicitation of broadly protective antibodies following infection with influenza viruses expressing H1N1 computationally optimized broadly reactive hemagglutinin antigens. ImmunoHorizons 2018;2:226–37. https://doi.org/10.4049/immunohorizons.1800044.

[76] Sautto GA, Kirchenbaum GA, Ross TM. Towards a universal influenza vaccine: different approaches for one goal. Virol J 2018;15:17. https://doi.org/10.1186/s12985-017-0918-y.

[77] Nuñez IA, Ross TM. Human COBRA 2 vaccine contains two major epitopes that are responsible for eliciting neutralizing antibody responses against heterologous clades of viruses. Vaccine 2020;38:830–9. https://doi.org/10.1016/j.vaccine.2019.10.097.

[78] Allen JD, Ray S, Ross TM. Split inactivated COBRA vaccine elicits protective antibodies against H1N1 and H3N2 influenza viruses. PLOS ONE 2018;13: e0204284. https://doi.org/10.1371/journal.pone.0204284.

[79] Allen JD, Jang H, DiNapoli J, Kleanthous H, Ross TM. Elicitation oF Protective Antibodies Against 20 Years of Future H3N2 cocirculating influenza virus variants in ferrets preimmune to historical H3N2 influenza viruses. 10.1128/jvi.00946-18 J Virol 2019;93. https://doi.org/10.1128/jvi.00946-18.

[80] Reneer ZB, Jamieson PJ, Skarlupka AL, Huang Y, Ross TM. Computationally optimized broadly reactive H2 HA influenza vaccines elicited broadly cross-reactive antibodies and protected mice from viral challenges. 10.1128/jvi.01526-20 J Virol 2020;95. https://doi.org/10.1128/jvi.01526-20.

[81] Allen JD, Ross TM. Next generation methodology for updating HA vaccines against emerging human seasonal influenza A(H3N2) viruses. Sci Rep 2021;11:4554. https://doi.org/10.1038/s41598-020-79590-7.

[82] Huang Y, França MS, Allen JD, Shi H, Ross TM. Next generation of computationally optimized broadly reactive HA vaccines elicited cross-reactive immune responses and provided protection against H1N1 virus infection. Vaccines 2021;9:793. https://doi.org/10.3390/vaccines9070793.

[83] Allen JD, Ross TM. Evaluation of next-generation h3 influenza vaccines in ferrets pre-immune to historical H3N2 viruses. Front Immunol 2021;12. https://doi.org/10.3389/fimmu.2021.707339.

[84] Allen JD, Ross TM. Bivalent H1 and H3 COBRA recombinant hemagglutinin vaccines elicit seroprotective antibodies against H1N1 and H3N2 influenza viruses from 2009 to 2019. J Virol 2022;96:e01652-21. https://doi.org/10.1128/jvi.01652-21.

[85] Mirdita M, Steinegger M, Söding J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics 2019;35:2856–8. https://doi.org/10.1093/bioinformatics/bty1057.

[86] Suzek BE, Wang Y, Huang H, McGarvey PB. Wu CH, the UniProt Consortium. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. Bioinformatics 2015;31:926–32. https://doi.org/10.1093/bioinformatics/btu739.

[87] Olsen TH, Boyles F, Deane CM. Observed antibody space: a diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. Protein Sci 2022;31:141–6. https://doi.org/10.1002/pro.4205.
[88] Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 2023;51:D638–46. https://doi.org/10.1093/nar/gkac1000.
[89] Golts A, Raboh M, Shoshan Y, Polaczek S, Rabinovici-Cohen S, Hexter E. FuseMedML: a framework for accelerated discovery in machine learning based biomedicine. J Open Source Softw 2023;8:4943. https://doi.org/10.21105/joss.04943.
[90] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE 2015; 10:e0118432. https://doi.org/10.1371/journal.pone.0118432.
[91] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Atten Is All You Need 2023. https://doi.org/10.48550/arXiv.1706.03762.
[92] Xu M, Zhang Z, Lu J, Zhu Z, Zhang Y, Chang M, et al. PEER: a comprehensive and multi-task benchmark for protein sequence understanding. Adv Neural Inf Process Syst 2022;35:35156–73.
[93] Dunbar J, Krawczyk K, Leem J, Baker T, Fuchs A, Georges G, et al. SAbDab: the structural antibody database. Nucleic Acids Res 2014;42:D1140–6. https://doi.org/10.1093/nar/gkt1043.
[94] Patent and Literature Antibody Database (PLAbDab): an evolving reference set of functionally diverse, literature-annotated antibody sequences and structures | Nucleic Acids Research | Oxford Academic n.d. https://academic.oup.com/nar/article/52/D1/D545/7424429 (accessed January 30, 2025).
[95] Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res 2019;47: D339–43. https://doi.org/10.1093/nar/gky1006.