Learning Context-aware Structural Representations to Predict Antigen and Antibody Binding Interfaces

Srivamshi Pittala[1] and Chris Bailey-Kellogg[1]

[1]Department of Computer Science, Dartmouth College, Hanover, USA

Abstract

Understanding how antibodies specifically interact with their antigens can enable better drug and vaccine design, as well as provide insights into natural immunity. Experimental structural characterization can detail the "ground truth" of antibody-antigen interactions, but computational methods are required to efficiently scale to large-scale studies. In order to increase prediction accuracy as well as to provide a means to gain new biological insights into these interactions, we have developed a unified deep learning-based framework to predict binding interfaces on both antibodies and antigens. The framework leverages three key aspects of antibody-antigen interactions in order to learn predictive structural representations: (1) since interfaces are formed from multiple residues in spatial proximity, we employ graph convolutions to aggregate properties across local regions in a protein; (2) since interactions are specific between antibody-antigen pairs, we employ an attention layer to explicitly encode the context of the partner; (3) since more data is available for general protein-protein interactions, we employ transfer learning to leverage this data as a prior for the specific case of antibody-antigen interactions. We show that this single framework achieves state-of-the-art performance at predicting binding interfaces on both antibodies and antigens, and that each of its three aspects drives additional improvement in the performance. We further show that the attention layer not only improves performance, but also provides a biologically interpretable perspective into the mode of interaction.

1 Introduction

As one of its mechanisms to combat disease, the immune system develops B cells that secrete antibodies to specifically recognize and either neutralize or help drive functional responses against a pathogen. An antibody recognizes a particular region, called its epitope, on a particular part of the pathogen, called its antigen; the region of the antibody directly involved in the recognition is called its paratope. The interface between an epitope and paratope is crucial to the affinity and specificity of an antibody-antigen interaction, and thus the antibody's function. Characterizing antibody-antigen interactions at the epitope-paratope resolution can thus reveal mechanisms of immune recognition, and, over a set of antibodies, can even provide insights into the development of the immune response. Such characterization can also benefit the development of therapeutics and vaccines. For example, therapeutic antibodies are being used to treat many different diseases [7, 20], and early development processes typically yield large arrays of candidate antibodies from which to select. Understanding their different recognition mechanisms can aid the selection and subsequent development. Similarly, subunit vaccines are being developed to train the immune system against a pathogen by mimicking an important part but without causing actual infection [5, 11, 12]. Understanding the recognition

processes driving a beneficial response, as well as those that are not useful, can guide the development of these vaccines so as to ensure the desired immune targeting.

Experimental structure determination methods, namely x-ray crystallography, NMR spectroscopy, and cryoEM, provide the gold standard for characterizing antibody-antigen binding modes [3, 28]. Unfortunately they remain expensive and time-consuming, and cannot feasibly keep up with the exploding amount of antibody sequence data for which it is desirable to understand antigen recognition, e.g., the millions of sequences obtained from analysis of an immune repertoire [33, 50, 56]. Alternative experimental methods like H-D exchange mass spectrometry [17] and alanine scanning [53] are faster and cheaper, and of lower resolution/confidence, but still require substantial experimental effort per target. Higher-throughput methods such as multiplexed SPR can characterize many interactions simultaneously but do not provide direct localization information [38, 6]. Computational methods thus have the most promise to scale to characterization of large numbers of possible epitope-paratope interactions, but it is necessary to ensure that predictions provide sufficient grounds to support further investigations, in terms of overall accuracy as well as the underlying reasoning for a prediction.

Prediction of antibody-antigen binding interfaces can be seen as a special case of predicting protein-protein binding interfaces. However, since these interfaces have their own special characteristics [27, 15] (as do other classes of protein-protein interactions), specific methods have been developed for epitope prediction and others for paratope prediction. Many methods make predictions based on amino acid sequence alone, e.g., predicting epitopes based on neural networks [39], SVMs [14, 43], HMMs [55], and random forests [21], and paratopes using LSTMs [31, 10] and random forests [34] Though sequence-based methods can perform well on paratope prediction, most sequence-based epitope predictions are limited to the special case of a sequentially contiguous epitope [54], while in contrast most epitopes are found to be conformational (distal in sequence, but close in 3D structure) [52, 37]. Thus we and others focus on structure-based methods that leverage geometric information in making predictions. Fortunately, while the complex structure is not known, in most common scenarios the structure of the antigen by itself is available, and antibody structure prediction techniques enable confident prediction of most of the antibody's structure [44, 45]. We thus briefly review this body of most closely related work on structure-based prediction of epitopes and paratopes.

Docking: Many structure-based methods for epitope and paratope prediction rely on computational docking techniques, which estimate the most likely conformations of a complex based on complementarity (geometric, chemical, energetic) between the individual proteins in many possible poses [8, 40, 46]. The resulting docking models may be ranked using a scoring function incorporating many different geometric and physico-chemical parameters; defining a good scoring function is a challenging task that typically relies on domain expertise [35]. From the top ranked conformations, regions on one protein that are close to the partner protein can be identified as binding interfaces. Thus antibody-antigen docking can simultaneously predict epitopes and paratopes. While the recall of docking is generally fairly high if enough docking models are considered, the precision is then fairly low, prompting the development of methods that directly predict epitopes and paratopes based on properties of the proteins [4].

Epitope prediction: Some approaches, e.g., PEPITO [49], ElliPro [36], EPSVR [30], and DiscoTope [25], apply machine learning methods to structural features of the antigen's residues. These methods can be considered antibody-agnostic as they do not use information from the partner antibody, and thus just reveal parts of the antigen generally amenable to antibody binding [42]. For prediction of an epitope targeted by a particular antibody, the context of which residues are likely to be involved in the interaction can improve the prediction performance, as well as distinguish specificity differences among different antibodies. This aspect of antibody-antigen interactions

was leveraged by the antibody-specific prediction method EpiPred [24] to achieve state-of-the-art performance. EpiPred first performs geometric matching of patches (e.g., based on docking models) and then scores residues on the antigen with a customized binding potential specific for antibodies and antigens.

Paratope prediction: Many paratope prediction predictors focus on special regions on antibodies called complementarity determining regions (CDRs), as they are well-defined from sequence and constitute the majority of the paratope and the majority of the differences among antibodies driving antigen-specific recognition. In the non-parametric method Paratome, the query antibody's structure and sequence are compared against a non-redundant dataset of antibodies, and paratopes are predicted based on resemblance to those on the closest matching antibody. [26]. Antibody i-Patch [23] uses a scoring function derived from an analysis of antibody-antigen interactions in a non-redundant training set. Recently, Daberdaku et al. [9] achieved state-of-the-art performance with a method that applies SVMs to classify patches extracted from the surface of the antibody, based on roto-translationally invariant shape descriptors and other physico-chemical properties representing the patches.

A common drawback of current structure-based methods for epitope and paratope prediction is the use of fixed representations, which can be limited by the extent of available domain knowledge. Furthermore, epitope and paratope prediction are treated as two separate tasks, leading to the use of different representations and prediction methods for antigens and antibodies. Sequence-based methods Parapred [31] and AG-Parapred [10] demonstrate the utility of learning representations for better paratope prediction. However, there are currently no methods to learn structural representations for either epitope or paratope prediction tasks. Recently, a spatial graph convolution network was proposed to learn structural representations of proteins for interface prediction in general protein-protein interactions [16]. While graph convolution networks can encode structural representations of residues with information from their spatial neighborhood, they do not encode the context of the target protein. As shown by current methods, embedding the correct context of the target protein can improve the prediction performance [24, 23]. Therefore, there is a need to develop methods for learning context-aware structural representations for epitope and paratope prediction.

In this work, we present a unified deep learning-based framework for learning context-aware structural representations of antigens and antibodies to predict their binding interfaces. Our framework consists of a novel combination of graph convolution networks, attention, and transfer learning to capture several desired aspects of antibody-antigen interactions. We show that the models trained on our framework can overcome the limitations of current computational methods and achieve state-of-the-art performance on both epitope and paratope prediction tasks. Using the attention layer, we demonstrate the ability of our framework to reveal the mode of interaction between antigens and antibodies, enabling a deeper study of the biological factors driving their interactions. Therefore, our framework improves prediction accuracy and provides interpretable results to expedite the process of large-scale antibody-antigen characterization.

2 Learning context-aware structural representations

We propose a novel deep learning framework (Figure 1) to learn structural representations of antigens and antibodies in order to predict their binding interfaces (i.e., antigen epitopes and antibody paratopes). Our framework comprises three components to leverage biological insights: (1) graph convolutions to capture the spatial relationships of the interfaces, (2) an attention layer to enable each protein's interface predictions to account for the potential binding context provided by its partner,

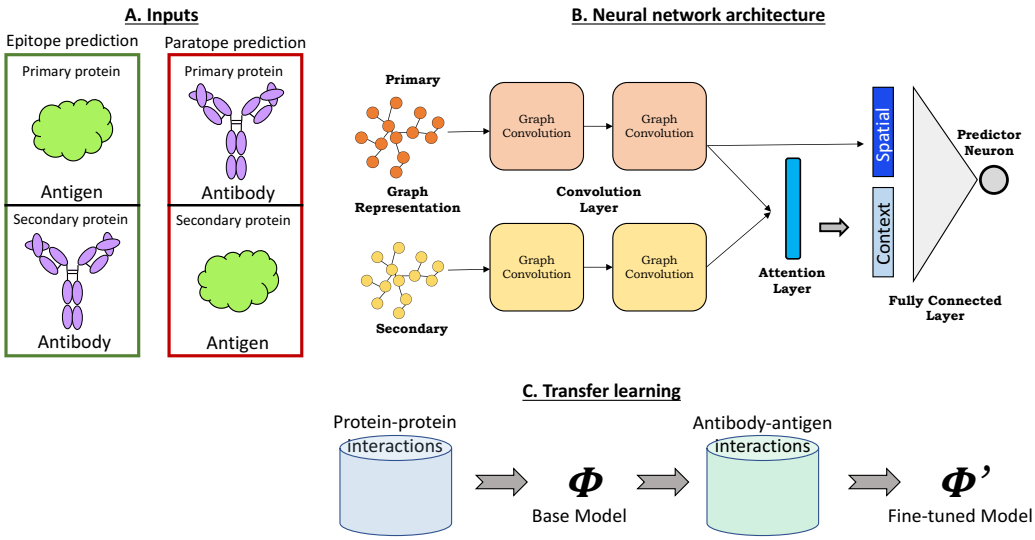

Figure 1: Schematic overview of the proposed framework

and (3) transfer learning to leverage the larger set of data available for general protein-protein interactions to provide a baseline model to be fine-tuned with antibody-antigen data.

We use this general framework to train two separate networks for the two prediction tasks: (1) an epitope prediction network in which the antigen is the primary protein on which we want to predict the interface (epitope) and the antibody is the secondary protein providing the context for a suitable interface; (2) a paratope prediction network with the antibody as primary for interface prediction (paratope) and the antigen secondary providing the context. We note that in both tasks, the interface labels of the secondary protein are hidden during training and prediction phases, forcing the attention layer to learn the correct context of the secondary protein in an unsupervised fashion.

Problem statement: The objective is to assign a label, either $+1$ (interface) or -1 (non-interface), to each residue of the primary protein.

A. Input representation: Each protein structure is represented as a graph, with nodes for the amino acid residues and edges between residues with Cβ-Cβ distance less than 10Å. Associated with each node is a 62-dimension feature vector encoding important sequence and structural properties as used in [16]: (a) a one-hot encoding of the amino acid type (d=20); (b) a conservation profile for that position across a set of homologous proteins returned by PSI-BLAST [2] (d=20); (c) the absolute and relative solvent accessible surface area of the residue as computed by STRIDE [19] (d=2); (d) a local amino acid profile indicating the number of times each amino acid type appears within 8Å of the residue (d=20). The structure-based features (c and d) were calculated for each protein in isolation from its partner.

Since antibody CDRs drive their antigen-specific recognition and the rest of the antibody framework is quite similar across all antibodies, nodes in the antibody graph are limited to "CDR clouds" as follows: (1) identify the six CDRs using the IMGT annotation tool [29]; (2) for each CDR, consider two sequentially adjacent residues; (3) further extend these sets to include all residues within 6Å in the structure (the maximum of the minimum Cβ-Cβ distance between any two CDR residues in the training sets).

B. Neural network: The neural network consists of graph convolution, attention, and fully

connected layers. Given two input graphs, primary $P = \{p_i\}_{i=1}^{N}$ and secondary $S = \{s_i\}_{i=1}^{M}$, the network assigns to each node $p_i \in P$ a probability of belonging to the positive class (i.e., binding interface).

B.1 Graph convolutions: Graph convolution [16] enables order-independent aggregation of properties over a neighborhood of residues that together contribute to the formation of a binding interface. For a node x_i and its receptive field consisting of K spatial neighbors $G_i = \{g_j\}_{j=1}^{K}$ from the input graph, the convolution operation results in a vector $\hat{z}_i \in \mathbb{R}^v$, where v is a specified number of filters for the layer (Eqn. 1). The parameters of this operation represent the aggregation weight matrix $\mathbf{W^c}$ for the center node, the aggregation weight matrix $\mathbf{W^g}$ for the neighboring nodes, and the bias vector $\hat{b}$. Thus, the convolution operation for a node x_i results in a spatial vector representation $\hat{x}'_i$ in the latent space $\mathbb{R}^v$.

$$\hat{z}_i = \text{ReLU}\left(\mathbf{W^c}\hat{x}_i + \frac{1}{|G_i|}\sum_{j=1}^{K}\mathbf{W^g}\hat{g}_j + \hat{b}\right) \tag{1}$$

Multiple layers can be stacked to produce high-level representations for each node. Each convolution layer has two weight-shared graph convolution modules, one for the primary graph and one for the secondary graph.

B.2 Attention: An attention layer encodes the context of secondary graph in the residue-level representations of the primary graph, providing information for each primary residue about secondary residues that are likely to interact with it. An attention score a_{ij} is computed between all node pairs $p_i \in P$ and $s_j \in S$ after projecting them into a latent space via an attention weight matrix $\mathbf{W^a}$ (Eqn. 2). The dimensions of $\mathbf{W^a}$ are determined by the number of neurons in the final convolution layer and the desired dimension of the latent space. This dot product style of computing attention scores was used to directly estimate complementarity between hidden representations as in [32, 10]. The context vector $\hat{c}_i \in \mathbb{R}^v$ for node p_i is then computed by aggregating the node-level representations of S using normalized attention scores (Eqn. 3). The normalization was performed to calibrate the score between each pair with respect to scores across all possible pairs. The normalized score α_{ij} can therefore be interpreted as a pair-wise interaction potential between p_i and s_j.

$$\begin{aligned} a_{ij} &= \text{ReLU}(h_i^T h_j), \\ \hat{h}_i = \mathbf{W^a}\hat{p}'_i, \quad \hat{h}_j &= \mathbf{W^a}\hat{s}'_j. \end{aligned} \tag{2}$$

$$\hat{c}_i = \sum_{j=1}^{M}\alpha_{ij}\hat{s}'_j, \quad \text{where } \alpha_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i,j=1}^{i=N,j=M} a_{ij}^2}} \tag{3}$$

B.3 Node classification: A final fully-connected layer performs classification for each primary node p_i based on its spatial vector $\hat{p}'_i$ and context vector $\hat{c}_i$ (Eqn. 4). A logit function transforms each node's output y_i to indicate the probability of belonging to the positive class.

$$y_i = \hat{W}^T\left(c_i||p'_i\right) + b, \qquad \text{where } || \text{ indicates concatenation} \tag{4}$$

C. Transfer learning: A base network ϕ for interface prediction is learned for a relatively larger set of general protein-protein interactions. The learned weights from the base model are then used to initialize weights for training the two task-specific networks, essentially fine-tuning the general base network for epitope and paratope prediction using antibody-antigen data.

Table 1: Summary of datasets used for training, validation, and testing.

Epitope prediction dataset

Data split	# Complexes	% Epitopes
Training	103	8.9%
Validation	38	9.7%
Test	30	7.8%

Paratope prediction dataset

Data split	# Complexes	% Paratopes
Training	205	8.8%
Validation	103	8.9%
Test	152	9.4%

3 Experiments

We evaluate our approach in head-to-head benchmark comparisons against state-of-the-art epitope and paratope predictors, showing that our unified framework outperforms approaches specifically targeted to each. Furthermore, we elaborate general precision and recall trends in the full architecture, as well as versions enabling characterization of the contributions of convolution, attention, and transfer learning. Finally, we explore the ability of the attention layer to provide insights into the basis for the model's predictions.

3.1 Datasets

Epitope prediction: The dataset from EpiPred [24] consists of 148 antibody-antigen complexes, 118 for training and 30 for testing. Since a separate validation set was not used, we constructed one from the antibody-antigen complexes in the Docking Benchmarking Dataset (DBD) v5 [51], ensuring that there were no overlapping complexes (by PDB id) among the three sets. This yielded resulted in 103 complexes for training, 38 for validation, and 30 for testing.

Paratope prediction: The dataset from Daberdaku et al. [9] consists of 471 antibody-antigen complexes, with 213 complexes for training, 106 for validation, and 152 for testing. Since our framework accepts only proteins, we discarded complexes with non-protein antigens (e.g., DNA), resulting in 205 complexes for training, 103 for validation, and 152 for testing.

Transfer learning: To facilitate unbiased transfer learning, the DBD v5 dataset [51] was processed to discard complexes that were categorized as antibody-antigen, resulting in a dataset of 191 protein-protein complexes.

Following previous studies [24, 9], residues were labeled as part of the interface if they had any non-Hydrogen atoms within 4.5Å of any non-Hydrogen atoms of residues on the other protein.

Table1 summarizes these dataset characteristics.

3.2 Implementation Details

The framework was implemented in TensorFlow [1]. The validation sets were used to find the optimal set of network training parameters for final evaluation. A grid search was performed over the following parameters: (a) Optimizer: Stochastic gradient descent, Momentum [48], Adagrad [13], or Adam [22]]; (b) learning rates: 0.0001, 0.001, 0.005, 0.01, 0.05, or 0.1; (c) batch size: 32, 64, pr 128; (d) dropout: 0.5 or 0.8. For each combination, networks were trained until the performance on validation set stopped improving or for a maximum of 250 epochs. For both epitope and paratope prediction, the best validation set performance was achieved when training till 120 epochs using the Momentum optimizer with Nesterov accelerated gradients [47] at a learning rate of 0.001, with batch size of 32 and 50% dropout rate. Training was carried out by minimizing the weighted cross-entropy

loss function as in [16]. The same network settings were used for training on general protein-protein complexes, but the fine-tuning was carried out on antibody-antigen complexes for half the original time (i.e., 60 epochs). The graph convolution layers were set to have 32 filters and the latent space dimension for attention was also set to 32. All weight matrices were initialized as in [18] and biases set to zero. For graph convolution, the receptive field (spatial neighborhood) for each node was set to include the 15 nearest nodes in the graph.

3.3 Evaluation

Epitope and paratope prediction networks were trained using the validation-set optimized hyper-parameters from above. Per-protein prediction performance was measured on the test sets by comparing predicted scores against ground truth labels. While our networks output a probability for each residue, in order to enable a direct comparison to other epitope prediction methods we computed precision and recall by predicting as interface residues those with logits above 0.5. To further elaborate precision-recall trade-offs, we considered all such classification thresholds and computed the area under the precision recall curve (AUC-PR). Though some previous methods have used the AUC-ROC metric, AUC-PR is more suitable here since the emphasis is on predicting binding interfaces (positive class) and the negative class constitutes roughly 90% of the samples. To summarize the performance, AUC-PR was averaged over all proteins in the test set. To provide robust estimates of performance, the training and testing procedures were repeated five times, and the mean and standard error reported. Our evaluations included two learning schema: task-specific learning (i.e., just using antibody-antigen data) and transfer learning (i.e., fine-tuning from a model trained with general protein-protein data). For each schema, five networks were evaluated: one network with a single fully-connected layer (No convolution), one with a single graph convolution layer (Conv1-layer) and one with two (Conv2-layer), and likewise one network with the attention layer following a single graph convolution layer (Conv1-layer+Attn) and one following two convolution layers (Conv2-layer+Attn).

3.4 Results

Epitope prediction

Table 2 summarizes the epitope test set prediction performance for our different neural network implementations along with the state-of-the-art Epipred [24] and DiscoTope [25]. Our networks all perform better than Epipred and DiscoTope in terms of precision and recall at the 0.5 cutoff. Elaborating performance for a range of cut-offs via AUC-PR enables further comparison among our architecture implementations (these numbers are not available for the other methods). The network with an attention layer after two convolution layers achieves the best performance, confirming the utility of embedding the context of the target antibody into the representation of antigen's residues in addition to information from their spatial neighbors. Furthermore, the improvements in performance of all models after transfer learning illustrates the benefits of leveraging data from general protein-protein interactions to establish a base model that can be fine-tuned with antibody-antigen data.

Paratope prediction

Table 3 summarizes the paratope test set prediction performance of our different neural networks and state-of-the-art structure-based methods Daberdaku et al. [9] and Antibody i-Patch [23]. To

Table 2: Epitope prediction performance summary. The measures for EpiPred and DiscoTope were taken from [24]

	Task-specific learning			Transfer learning		
Method	AUC-PR	Precision	Recall	AUC-PR	Precision	Recall
No Convolution	0.177±0.035	0.153	0.601	0.191±0.021	0.162	0.500
Conv1-layer	0.193±0.005	0.152	0.725	0.191±0.003	0.160	0.582
Conv1-layer+Attn	0.187±0.004	0.147	0.683	0.202±0.002	0.160	0.580
Conv2-layer	0.192±0.005	0.147	0.737	0.226±0.004	0.157	0.678
Conv2-layer+Attn	**0.212±0.007**	0.154	0.691	**0.242±0.006**	0.165	0.579
Epipred [24]	NA	0.136	0.436	-	-	-
DiscoTope [25]	NA	0.214	0.110	-	-	-

Table 3: Paratope prediction performance summary. The performance for Daberdaku et al.[9] and Antibody-iPatch[23] were taken from [9]

	Task-specific learning		Transfer learning	
Method	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC
No Convolution	0.652±0.012	0.938±0.004	0.658±0.003	0.937±0.001
Conv1-layer	**0.700±0.002**	0.956±0.000	0.696±0.001	0.957±0.000
Conv1-layer+Attn	**0.700±0.001**	**0.957±0.000**	**0.703±0.000**	**0.958±0.001**
Conv2-layer	0.691±0.004	0.957±0.000	0.689±0.002	0.957±0.000
Conv2-layer+Attn	0.692±0.004	0.957±0.001	0.697±0.001	0.958±0.000
Daberdaku et al.[9]	0.658	0.950	-	-
Antibody i-Patch[23]	0.376	0.840	-	-

enable a direct comparison to previous studies, we predict for the entire structure of the antibody Fv region instead of just the CDR clouds as described in our methods. Our networks perform better than the other methods on both AUC-PR and AUC-ROC, establishing the superior performance of learned features over pre-defined features as used by Daberdaku et al. The network with a single layer of convolution and attention achieves the best performance, but the attention layer provides only a small performance improvement over convolution. We hypothesize that since paratopes are mostly localized to regions around the CDRs, the context of the antigen may not provide much more information regarding exact paratope location than the structural properties already captured by convolution. Nonetheless, as we show in the next section, the attention layer offers the benefit of making the network interpretable, which can be a difficult task for convolution layers alone.

Assessing the contributions of attention

The attention layer provides the opportunity to study the mode of interaction by revealing the learned context of the target protein without requiring additional inference techniques. The attention score between every pair of residues can be visualized as a matrix, as Figure 2A illustrates for the complex on which our epitope prediction network performed best. In this heatmap, epitopes have a substantially distinct attention profile compared to other residues on the antigen, which results

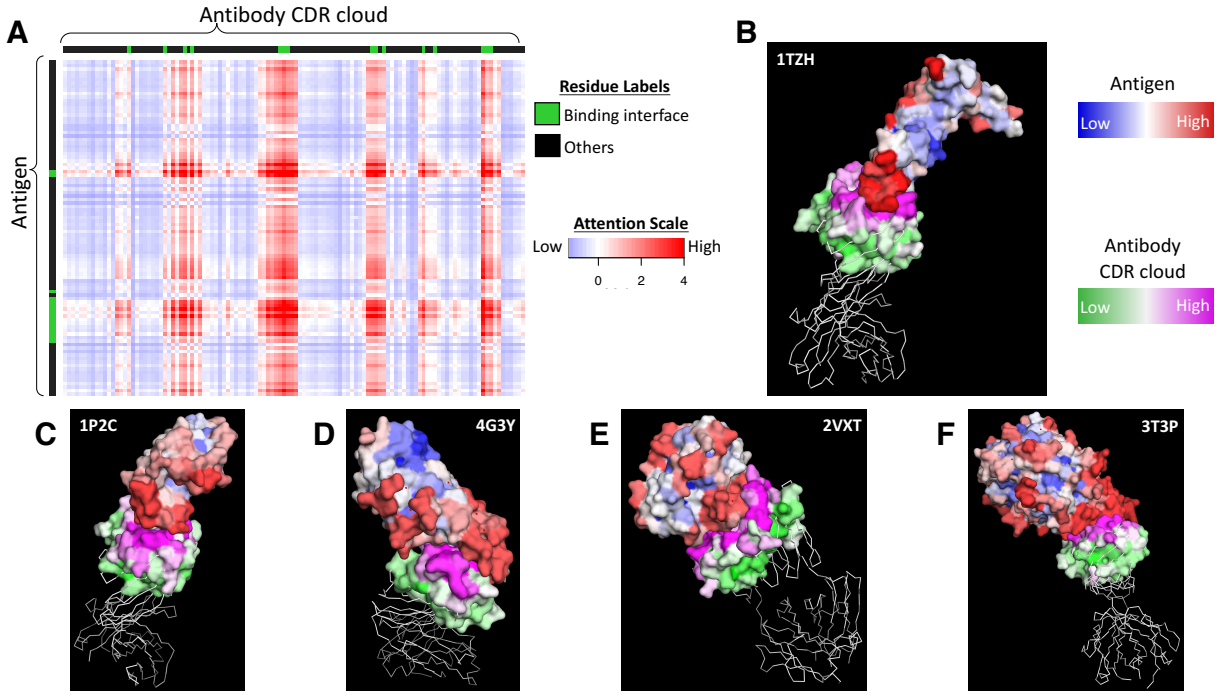

Figure 2: Attention visualization (A) Heatmap of the attention score matrix for an antibody in complex with VEGFA (vascular endothelial growth factor A) (PDB ID 1TZH). The scores were normalized to have zero mean and unit variance, and truncated to the range [-4,4]. (B-E) Projection of max-pooled attention scores onto structures of example antigen and antibody complexes in the test set. Structural visualizations made using PyMOL [41].

in improved epitope prediction (AUC-PR: 0.6) compared to convolution alone (AUC-PR: 0.45). These attention scores can further be projected onto the structures (Figure 2B) by taking for each residues the maximum of its scores with partner residues. This projection shows that attention is high between residues in and around the actual interface region, suggesting that the attention layer encodes the correct context (i.e., paratopes) of the antibody for epitopes. The same pattern of high attention scores near the interface regions was also observed for other antibody-antigen complexes (Figure. 2C-F illustrates the next 4 top performers).

Intrigued by the attention layer's ability to localize the appropriate context during epitope prediction, we hypothesized that the same ability could benefit paratope prediction. We thus performed a "cross-task evaluation", in which a network was trained to predict *epitopes* using the *antigens* in the epitope prediction training set. This epitope prediction network was then evaluated for its performance at also predicting *paratopes* for the *antibodies* in the epitope prediction test set—the reciprocal task to that for which it was trained. For reference, paratope prediction networks were trained on the antibodies from the epitope prediction training set and applied (as normal) to the antibodies from the epitope prediction test set. As expected, the paratope prediction networks perform significantly better at predicting epitopes than do the networks trained to predict epitopes. However, the results from cross-task evaluation show that even though none of the networks were trained to predict paratopes, those with an attention layer perform better than convolution-only networks. This suggests that the attention layer is indeed able to better capture the specificity of

Table 4: Performance summary from "cross-task evaluation", predicting paratopes using networks trained only for predicting epitopes.

	Cross-task evaluation		Task-specific learning	
Method	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC
Conv1-layer	0.612±0.028	0.844±0.026	0.770±0.000	0.970±0.007
Conv1-layer+Attn	0.624±0.026	0.875±0.035	0.768±0.003	0.971±0.008
Conv2-layer	0.588±0.058	0.910±0.051	0.766±0.005	0.971±0.001
Conv2-layer+Attn	**0.628±0.025**	**0.938±0.029**	0.765±0.007	0.971±0.012

antibody-antigen interactions, thereby also benefiting paratope prediction.

4 Conclusion

We have presented a unified deep learning framework for predicting binding interfaces on antibodies and antigens. Our results demonstrate that the networks learn structural representations that capture many desired aspects of antibody-antigen interactions and simultaneously achieve state-of-the-art performance on both epitope and paratope prediction tasks. We also show that the attention layer successfully encodes the context of partner proteins, improving prediction performance and providing an interpretable view of the mode of interaction. Future work includes including additional residue features while imposing sparsity constraints on the attention matrix, applying the same framework to other large protein families with specific recognition modes, and using predictions to focus docking as well as experimental evaluation.

Acknowledgments

This research was supported in part by NIH grant 2R01GM098977, with computational resources from NSF award CNS-1205521 and discovery clusters at Dartmouth. We thank Margaret Ackerman, Karl Griswold, and members of the Bailey-Kellogg lab for helpful comments, and Casey Hua for sharing antibody art in Figure 1.

Code availability

The source code will be made available upon publication of the manuscript.

References

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[2] Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402, 09 1997.

[3] X. C. Bai, G. McMullan, and S. H. Scheres. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.*, 40(1):49–57, Jan 2015.

[4] R. Brenke, D. R. Hall, G. Y. Chuang, S. R. Comeau, T. Bohnuud, D. Beglov, O. Schueler-Furman, S. Vajda, and D. Kozakov. Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics*, 28(20):2608–2614, Oct 2012.

[5] B. Briney, D. Sok, J. G. Jardine, D. W. Kulp, P. Skog, S. Menis, R. Jacak, O. Kalyuzhniy, N. de Val, F. Sesterhenn, K. M. Le, A. Ramos, M. Jones, K. L. Saye-Francisco, T. R. Blane, S. Spencer, E. Georgeson, X. Hu, G. Ozorowski, Y. Adachi, M. Kubitz, A. Sarkar, I. A. Wilson, A. B. Ward, D. Nemazee, D. R. Burton, and W. R. Schief. Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies. *Cell*, 166(6):1459–1470, Sep 2016.

[6] B. D. Brooks, A. R. Miles, and Y. N. Abdiche. High-throughput epitope binning of therapeutic monoclonal antibodies: why you need to bin the fridge. *Drug Discov. Today*, 19(8):1040–1044, Aug 2014.

[7] P. J. Carter. Potent antibody therapeutics by design. *Nat. Rev. Immunol.*, 6(5):343–357, May 2006.

[8] R. Chen, L. Li, and Z. Weng. ZDOCK: an initial-stage protein-docking algorithm. *Proteins*, 52(1):80–87, Jul 2003.

[9] Sebastian Daberdaku and Carlo Ferrari. Antibody interface prediction with 3D Zernike descriptors and SVM. *Bioinformatics*, 11 2018.

[10] Andreea Deac, Petar Veličković, and Pietro Sormanni. Attentive cross-modal paratope prediction. *arXiv e-prints*, page arXiv:1806.04398, Jun 2018.

[11] I. Delany, R. Rappuoli, and E. De Gregorio. Vaccines for the 21st century. *EMBO Mol Med*, 6(6):708–720, Jun 2014.

[12] N. A. Doria-Rose and M. G. Joyce. Strategies to guide the antibody affinity maturation process. *Curr Opin Virol*, 11:137–147, Apr 2015.

[13] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.*, 12:2121–2159, July 2011.

[14] Y. El-Manzalawy, D. Dobbs, and V. Honavar. Predicting linear B-cell epitopes using string kernels. *J. Mol. Recognit.*, 21(4):243–255, 2008.

[15] R. Esmaielbeiki, K. Krawczyk, B. Knapp, J. C. Nebel, and C. M. Deane. Progress and challenges in predicting protein interfaces. *Brief. Bioinformatics*, 17(1):117–131, Jan 2016.

[16] Alex Fout, Jonathon Byrd, Basir Shariat, and Asa Ben-Hur. Protein interface prediction using graph convolutional networks. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 6530–6539. Curran Associates, Inc., 2017.

[17] E. S. Gallagher and J. W. Hudgens. Mapping Protein-Ligand Interactions with Proteolytic Fragmentation, Hydrogen/Deuterium Exchange-Mass Spectrometry. *Meth. Enzymol.*, 566:357–404, 2016.

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *arXiv e-prints*, page arXiv:1502.01852, Feb 2015.

[19] M. Heinig and D. Frishman. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.*, 32(Web Server issue):W500–502, Jul 2004.

[20] P. Holliger and P. J. Hudson. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.*, 23(9):1126–1136, Sep 2005.

[21] Martin Closter Jespersen, Bjoern Peters, Morten Nielsen, and Paolo Marcatili. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research*, 45(W1):W24–W29, 05 2017.

[22] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *arXiv e-prints*, page arXiv:1412.6980, Dec 2014.

[23] Konrad Krawczyk, Terry Baker, Jiye Shi, and Charlotte M. Deane. Antibody i-Patch prediction of the antibody binding site improves rigid local antibody–antigen docking. *Protein Engineering, Design and Selection*, 26(10):621–629, 09 2013.

[24] Konrad Krawczyk, Xiaofeng Liu, Terry Baker, Jiye Shi, and Charlotte M. Deane. Improving B-cell epitope prediction and its application to global antibody-antigen docking. *Bioinformatics*, 30(16):2288–2294, 04 2014.

[25] Jens Vindahl Kringelum, Claus Lundegaard, Ole Lund, and Morten Nielsen. Reliable b cell epitope predictions: Impacts of method development and improved benchmarking. *PLOS Computational Biology*, 8(12):1–10, 12 2012.

[26] V. Kunik, S. Ashkenazi, and Y. Ofran. Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. *Nucleic Acids Res.*, 40(Web Server issue):W521–524, Jul 2012.

[27] V. Kunik and Y. Ofran. The indistinguishability of epitopes from protein surface is explained by the distinct binding preferences of each of the six antigen-binding loops. *Protein Eng. Des. Sel.*, 26(10):599–609, Oct 2013.

[28] H. Lee, S. A. Brendle, S. M. Bywaters, J. Guan, R. E. Ashley, J. D. Yoder, A. M. Makhov, J. F. Conway, N. D. Christensen, and S. Hafenstein. A cryo-electron microscopy study identifies the complete H16.V5 epitope and reveals global conformational changes initiated by binding of the neutralizing antibody fragment. *J. Virol.*, 89(2):1428–1438, Jan 2015.

[29] Marie-Paule Lefranc, Christelle Pommié, Manuel Ruiz, Véronique Giudicelli, Elodie Foulquier, Lisa Truong, Valérie Thouvenin-Contet, and Gérard Lefranc. Imgt unique numbering for immunoglobulin and t cell receptor variable domains and ig superfamily v-like domains. *Developmental & Comparative Immunology*, 27(1):55 – 77, 2003.

[30] S. Liang, D. Zheng, D. M. Standley, B. Yao, M. Zacharias, and C. Zhang. EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. *BMC Bioinformatics*, 11:381, Jul 2010.

[31] E. Liberis, P. Velickovic, P. Sormanni, M. Vendruscolo, and P. Lio. Parapred: antibody paratope prediction using convolutional and recurrent neural networks. *Bioinformatics*, 34(17):2944–2950, 09 2018.

[32] Minh-Thang Luong, Hieu Pham, and Christopher D. Manning. Effective Approaches to Attention-based Neural Machine Translation. *arXiv e-prints*, page arXiv:1508.04025, Aug 2015.

[33] E. Miho, A. Yermanos, C. R. Weber, C. T. Berger, S. T. Reddy, and V. Greiff. Computational Strategies for Dissecting the High-Dimensional Complexity of Adaptive Immune Repertoires. *Front Immunol*, 9:224, 2018.

[34] P. P. Olimpieri, A. Chailyan, A. Tramontano, and P. Marcatili. Prediction of site-specific interactions in antibody-antigen complexes: the proABC method and server. *Bioinformatics*, 29(18):2285–2291, Sep 2013.

[35] M. Pedotti, L. Simonelli, E. Livoti, and L. Varani. Computational docking of antibody-antigen complexes, opportunities and pitfalls illustrated by influenza hemagglutinin. *Int J Mol Sci*, 12(1):226–251, Jan 2011.

[36] J. Ponomarenko, H. H. Bui, W. Li, N. Fusseder, P. E. Bourne, A. Sette, and B. Peters. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*, 9:514, Dec 2008.

[37] Marc H.V. Van Regenmortel. Mapping epitope structure and activity: From one-dimensional prediction to four-dimensional description of antigenic specificity. *Methods*, 9(3):465 – 472, 1996.

[38] P. Safsten. Epitope mapping by surface plasmon resonance. *Methods Mol. Biol.*, 524:67–76, 2009.

[39] S. Saha and G. P. Raghava. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins*, 65(1):40–48, Oct 2006.

[40] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, and H. J. Wolfson. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.*, 33(Web Server issue):W363–367, Jul 2005.

[41] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. November 2015.

[42] I. Sela-Culang, Y. Ofran, and B. Peters. Antibody specific epitope prediction-emergence of a new paradigm. *Curr Opin Virol*, 11:98–102, Apr 2015.

[43] Harinder Singh, Hifzur Rahman Ansari, and Gajendra P. S. Raghava. Improved method for linear b-cell epitope prediction using antigen's primary sequence. *PLOS ONE*, 8(5):1–8, 05 2013.

[44] A. Sircar. Methods for the homology modeling of antibody variable regions. *Methods Mol. Biol.*, 857:301–311, 2012.

[45] A. Sircar, E. T. Kim, and J. J. Gray. RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Res.*, 37(Web Server issue):W474–479, Jul 2009.

[46] Aroop Sircar and Jeffrey J. Gray. Snugdock: Paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. *PLOS Computational Biology*, 6(1):1–13, 01 2010.

[47] Weijie Su, Stephen Boyd, Emmanuel J. C, and ès. A differential equation for modeling nesterov's accelerated gradient method: Theory and insights. *Journal of Machine Learning Research*, 17(153):1–43, 2016.

[48] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton. On the importance of initialization and momentum in deep learning. In Sanjoy Dasgupta and David McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 1139–1147, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

[49] M. J. Sweredoski and P. Baldi. PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. *Bioinformatics*, 24(12):1459–1460, Jun 2008.

[50] J. Truck, M. N. Ramasamy, J. D. Galson, R. Rance, J. Parkhill, G. Lunter, A. J. Pollard, and D. F. Kelly. Identification of antigen-specific B cell receptor sequences using public repertoire analysis. *J. Immunol.*, 194(1):252–261, Jan 2015.

[51] Thom Vreven, Iain H. Moal, Anna Vangone, Brian G. Pierce, Panagiotis L. Kastritis, Mieczyslaw Torchala, Raphael Chaleil, Brian Jiménez-García, Paul A. Bates, Juan Fernandez-Recio, Alexandre M.J.J. Bonvin, and Zhiping Weng. Updates to the integrated protein–protein interaction benchmarks: Docking benchmark version 5 and affinity benchmark version 2. *Journal of Molecular Biology*, 427(19):3031 – 3041, 2015.

[52] Gernot Walter. Production and use of antibodies against synthetic peptides. *Journal of Immunological Methods*, 88(2):149 – 161, 1986.

[53] G. A. Weiss, C. K. Watanabe, A. Zhong, A. Goddard, and S. S. Sidhu. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci. U.S.A.*, 97(16):8950–8954, Aug 2000.

[54] Bo Yao, Dandan Zheng, Shide Liang, and Chi Zhang. Conformational b-cell epitope prediction on antigen protein structures: A review of current algorithms and comparison with common binding site prediction methods. *PLOS ONE*, 8(4):1–4, 04 2013.

[55] L. Zhao, L. Wong, and J. Li. Antibody-specified B-cell epitope prediction in line with the principle of context-awareness. *IEEE/ACM Trans Comput Biol Bioinform*, 8(6):1483–1494, 2011.

[56] J. Zhu, G. Ofek, Y. Yang, B. Zhang, M. K. Louder, G. Lu, K. McKee, M. Pancera, J. Skinner, Z. Zhang, R. Parks, J. Eudailey, K. E. Lloyd, J. Blinn, S. M. Alam, B. F. Haynes, M. Simek, D. R. Burton, W. C. Koff, J. C. Mullikin, J. R. Mascola, L. Shapiro, P. D. Kwong, J. Becker, B. Benjamin, R. Blakesley, G. Bouffard, S. Brooks, H. Coleman, M. Dekhtyar, M. Gregory, X. Guan, J. Gupta, J. Han, A. Hargrove, S. L. Ho, T. Johnson, R. Legaspi, S. Lovett, Q. Maduro, C. Masiello, B. Maskeri, J. McDowell, C. Montemayor, J. Mullikin, M. Park, N. Riebow, K. Schandler, B. Schmidt, C. Sison, M. Stantripop, J. Thomas, P. Thomas, M. Vemulapalli, and A. Young. Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. *Proc. Natl. Acad. Sci. U.S.A.*, 110(16):6470–6475, Apr 2013.