Geometric Epitope and Paratope Prediction

Marco Pegoraro*
Sapienza, University of Rome
pegoraro@di.uniroma1.it

Clémentine Dominé *
Gatsby Computational Neuroscience Unit, University College London
clementine.domine.20@ucl.ac.uk

Emanuele Rodolà
Sapienza, University of Rome

Petar Veličković
Google DeepMind

Andreea Deac
Mila, Université de Montréal

Abstract

Antibody-antigen interactions play a crucial role in identifying harmful foreign molecules. In this paper, we investigate the optimal representation for predicting the binding sites in the two molecules and emphasize the importance of geometric information. Specifically, we compare different geometric deep learning methods applied to proteins' inner (I-GEP) and outer (O-GEP) structures. We incorporate 3D coordinates and spectral geometric descriptors as input features to fully leverage the geometric information. Our research suggests that surface-based models are more efficient than other methods, and our O-GEP experiments have achieved state-of-the-art results with significant performance improvements.

1 Introduction

Identifying the binding sites of antibodies is essential for developing vaccines and synthetic antibodies. These binding sites, called paratopes, can bind to antigens, wherein the corresponding binding site is known as the epitope, thus neutralizing harmful foreign molecules in the body. Experimental methods for determining the residues that belong to the paratope and epitope are time-consuming and expensive, highlighting the need for computational tools to facilitate the rapid development of therapeutics. The recent COVID-19 epidemic highlighted this need further, as mutations in the antigen were shown to impact the binding mechanism, potentially reducing the efficacy of existing treatments [1]. Predicting the binding sites of an antibody-antigen interaction requires considering the entire antigen for epitope prediction and a localized region of the antibody, known as the Complementarity-Determining Region (CDR), for paratope prediction.

The shape and structure of molecules play a crucial role in determining their interactions with other molecules, as complementary geometric shapes are required for successful binding [2]. The use of geometrical information is further justified by the emergence of technology predicting the single-protein structure, such as AlphaFold 2 [3], which has comparable accuracy to experimental methods. The integration of geometric and structural information in protein-to-protein interaction studies has led to significant progress [4, 5]. While several methods have concentrated on the 3D graph representation, few methods [5, 6] have investigated the 3D surface representation. We aim to assess the impact of utilizing the geometric representation of the antigen and antibody in the task of epitope-paratope prediction. Our approach, GEP (Geometric Epitope-Paratope) Prediction, proposes different geometric representations of the molecules to create accurate predictors for predicting antibody-antigen binding sites. In particular, we recognize the importance of the outer surface of a molecule in molecular interactions.

Our paper introduces several contributions, including the analysis of the importance of geometric information within graph learning using equivariant layers for improved predictions. Moreover, we fully leverage molecular geometric information by representing molecules as surfaces and employing spectral geometry techniques, leading to state-of-the-art performance. Additionally, we will provide

*Equal contribution.

M. Pegoraro et al., Geometric Epitope and Paratope Prediction (Extended Abstract). Presented at the Second Learning on Graphs Conference (LoG 2023), Virtual Event, November 27–30, 2023.

a dataset generation pipeline for PDB molecules, offering molecular representations in both graph and surface formats, facilitating comprehensive cross-method comparisons.

2 Related work

The structure of proteins provides crucial information about the location and orientation of the binding sites. Various approaches have been taken in the literature to address the task of epitope and paratope prediction, including sequential [7, 8] and structural [9, 10] methods. Furthermore, Geometric deep learning has emerged as a powerful tool for predicting protein-protein interactions [11], with graph-based representations being one of the most common approaches [4, 12]. These methods leverage the geometric information of the molecules to learn complex relationships between epitopes and paratopes. For instance, some approaches [10, 13] use the graph structure to compute features based on neighbouring residues, which are then aggregated to highlight the most probable region of interaction.

An alternative approach is to represent proteins as surfaces. MaSIF [14] focuses on the more general problem of protein interaction region prediction and uses a surface representation learned through convolutions defined on the surface. PiNet [5] represents the protein surface as a point cloud and employs PointNet [15] to classify points as interacting or not. On the contrary, Zhang et al. [6] model the surface of a molecule as a graph and apply an equivariant graph neural network (EGNN, [16]) for binding site prediction.

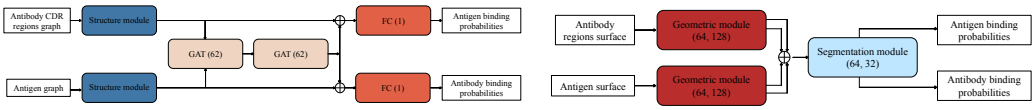
Integrating structural and geometric information has proven to be a promising approach for improving protein interaction prediction. Still, few studies have focused on the specific case of epitope and paratope prediction [17]. Our work supports this view by showing that considering the problem as a geometric one can effectively improve performance.

3 Data

Comparing methods across different molecular representations is crucial for advancing research in molecular modelling. We developed a reusable pipeline that generates a dataset to evaluate methods using inner and outer structure representations.We collected a dataset of 133 protein complexes from Epipred [9], with 103 for training and 30 for testing. The training and test sets have been selected to share no more than 90% pairwise sequence identity. The PDB files were obtained from the Sabdab database [18]. In the test set, 7.8% of antigen residues were labelled as positive. Additionally, we used a separate set of 27 protein complexes from PECAN derived from a subset of the Docking Benchmark v5 [19] to validate our results.

4 Method

We considered two scenarios: a protein represented through its inner structure (*I-GEP*) and outer structure (*O-GEP*). In both cases, we leverage the geometric information to improve the performance of epitope and paratope prediction methods. Details on the methods, including how we construct the different representations for each model and the models architectures are reported in Appendix A.



(a) I-GEP **(b)** O-GEP

Figure 1: Our model architecture is represented with arrows indicating data flow between modules, using color-coded blocks to represent layers or modules, with text inside each block specifying the layer type. The model takes antibody-antigen pairs as input, featuring node-level features for IGEP and surface point-level features for OGEP, and produces binding probabilities for each input node or point.

4.1 I-GEP

Our I-GEP model is a method for predicting epitopes and paratopes using a graph-based approach that captures the inner structure of a protein. The I-GEP model has two main components: a structural module that computes an embedding for each residue using the graph structure and a graph attention network (GAT) that combines information from both the antigen and antibody residues. The network then predicts both epitope and paratope residues simultaneously using a fully connected layer, as shown in Fig. 1b.

To improve the accuracy of our predictions, we integrate geometric information into the I-GEP model using two different approaches. In the first approach, EPMP$_{xyz}$, we use graph convolutional network layers in the structural module as in EPMP [10], but we include the centred 3D coordinates of residues in the input features. The second approach, $E(n)$-*EPMP*, uses the $E(n)$ invariant layer encoder from EGNN [16] instead of graph convolutional networks. This approach considers only the distances between residues, making it invariant to translations, rotations, and reflections on the residue positions in each molecule.

4.2 O-GEP

Our O-GEP model operates on the protein's surface and includes a geometric module that uses the surface's geometry to spread information across it. This process generates features that are then combined and shared between the antibody and antigen through fully connected layers (segmentation module), resulting in an interaction probability for each point on the surface, as shown in Fig. 1b.

We explore two different models for the geometric module. As a baseline, we use PointNet [15] to recreate the architecture proposed in PiNet [5]. The second model employs diffusion layers from DiffNet [20] to propagate features on the surface. This makes our model robust against surface perturbations and suitable for handling meshes and point clouds with fewer points.

We further examine the impact of using the Heat Kernel Signature (HKS) as an extra geometric descriptor input. The HKS [21] is a concise point-wise spectral signature which summarizes local and global information about the intrinsic geometry of a shape by capturing the properties of the heat diffusion process on the surface. One of the key benefits of using HKS is that it remains stable even under minor surface perturbations, thus enabling it to withstand even conformational rearrangements of the proteins. To utilize the HKS descriptor, we concatenate it with the input features at each point on the surface and then pass the concatenated data through the geometric module.

To transfer the binding probabilities from the protein's surface to the residues, we utilized the average of all the points on the surface that correspond to the same residues. This method ensures that the binding probabilities are accurately represented in the residue space, enabling us to make reliable predictions about epitope and paratope locations.

4.3 Training and evaluation

The networks were trained using the class-weighted binary cross-entropy loss and the Adam SGD optimizer to handle imbalanced binary classification tasks. We report training details in Appendix B. Given the significant disparity in class sizes, we utilize Matthew's correlation coefficient (MCC) between the residues' classification as our main benchmarking metric for model evaluation. We also report the area under the receiver operating characteristic curve (AUC ROC) and the area under the precision-recall curve (AUC PR) as used in [5, 10]. All reported values are aggregated across five random seeds to ensure the robustness of our findings.

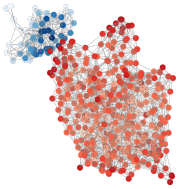
5 Results

In this section, we report the results of our experiments and demonstrate the contribution of geometric information on the task of epitope-paratope prediction.

I-GEP results We conducted experiments to evaluate the effectiveness of incorporating geometric information by comparing our proposed models from Section 4.1 with the EPMP model proposed in [10]. Our results, presented in Table 2a, demonstrate that the inclusion of geometric information leads to a meaningful increase in performance. Specifically, the use of the $E(n)$ invariant layer ($E(n)$-EPMP) resulted in an improvement in all metrics for both antibody and antigen.

	Antigen			Antibody		
	MCC	AUC ROC	AUC PR	MCC	AUC ROC	AUC PR
EPMP	0.09 ± 0.01	0.61 ± 0.01	0.12 ± 0.00	0.39 ± 0.02	0.79 ± 0.01	0.53 ± 0.01
EPMP$_{xyz}$	0.10 ± 0.01	0.63 ± 0.01	0.15 ± 0.01	0.38 ± 0.02	0.79 ± 0.01	0.53 ± 0.01
$E(n)$-EPMP	$\mathbf{0.14 \pm 0.01}$	$\mathbf{0.68 \pm 0.02}$	$\mathbf{0.16 \pm 0.01}$	$\mathbf{0.44 \pm 0.11}$	$\mathbf{0.82 \pm 0.07}$	$\mathbf{0.60 \pm 0.10}$

(a) Quantitative results as mean and standard deviation ($\pm$)

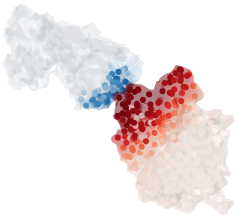


(b) Qualitative example

(c) Results from I-GEP models.

	Antigen			Antibody		
	MCC	AUC ROC	AUC PR	MCC	AUC ROC	AUC PR
PiNet (xyz)	0.39 ± 0.05	0.89 ± 0.01	0.44 ± 0.02	0.26 ± 0.12	0.77 ± 0.03	0.52 ± 0.08
PiNet (xyz+hks)	0.30 ± 0.04	0.87 ± 0.02	0.37 ± 0.06	0.22 ± 0.05	0.74 ± 0.00	0.47 ± 0.02
DiffNet$_{pc}$ (xyz)	0.41 ± 0.06	$\mathbf{0.90 \pm 0.01}$	0.49 ± 0.02	0.30 ± 0.06	0.79 ± 0.01	0.56 ± 0.03
DiffNet$_{pc}$ (hks)	0.07 ± 0.05	0.66 ± 0.02	0.14 ± 0.01	0.44 ± 0.03	$\mathbf{0.85 \pm 0.00}$	0.68 ± 0.01
DiffNet$_{pc}$ (xyz+hks)	$\mathbf{0.44 \pm 0.03}$	$\mathbf{0.90 \pm 0.01}$	$\mathbf{0.50 \pm 0.02}$	0.23 ± 0.06	0.77 ± 0.04	0.51 ± 0.05
DiffNet$_{m}$ (xyz)	0.42 ± 0.03	$\mathbf{0.90 \pm 0.01}$	0.48 ± 0.05	0.24 ± 0.08	0.78 ± 0.02	0.52 ± 0.03
DiffNet$_{m}$ (hks)	0.09 ± 0.02	0.64 ± 0.02	0.14 ± 0.01	$\mathbf{0.49 \pm 0.01}$	$\mathbf{0.85 \pm 0.00}$	$\mathbf{0.69 \pm 0.01}$
DiffNet$_{m}$ (xyz+hks)	0.42 ± 0.06	$\mathbf{0.90 \pm 0.01}$	0.46 ± 0.07	0.28 ± 0.06	0.77 ± 0.02	0.52 ± 0.04

(d) Quantitative results as mean and standard deviation ($\pm$)



(e) Qualitative example

(f) Results from O-GEP models.

Figure 2: Left: Quantitative results evaluated on the residues. We report the MCC, the AUR ROC and the AUC PR. We write in bold the best results. **Right**: Representation of binding prediciton on the antibody-antigen complex number '4jr9'. The continuous binding predictions are represented as a color gradient in blue and red for the antigen and antibody, respectively.

O-GEP results To test the performance of O-GEP models, we consider the methods proposed in Section 4.2 with different combinations of input features. In addition to the physicochemical features, we test different combinations of geometric information: 3d coordinates (XYZ) and Heat Kernel Signature (HKS). For the DIFFNET models, we consider both the point cloud ($_{pc}$) and the mesh ($_m$) of the surface. The results are summarized in Table 2d. Incorporating diffusion layers (DIFFNET) along with 3D coordinates and Heat Kernel Signature as additional features consistently outperformed the baseline method PINET. The use of these techniques led to an MCC score three time as high as that obtained by the I-GEP models. However, unlike epitope prediction, the paratope prediction did not show the same level of improvement with O-GEP models. In this case, the best results were achieved by considering only the HKS features and diffusion layers. In Appendix C, we also show the metrics computed only on residues with a representing point on the surface.

Qualitative results We plot the binding probability on the residuals computed by the models as increasing intensity colours: blue for the antibody and red for the antigen. Figure 2b shows the results of the $E(n)$-*EPMP* on the residual graph. The epitope prediction focuses on sparse regions of the antigene, such as the spiky edges. In contrast, paratope prediction concentrates on the residues closest to the antigen. In Figure 2e, the predictions of DIFFNET$_{pc}$ (XYZ+HKS) are shown on both the surface and residues of the molecules. The predictions are highly localized on the region nearest to the binding molecule. It's worth noticing that the 3d coordinates given as input to the models are centred and randomly rotated, providing no prior knowledge of the binding region.

6 Conclusions

We investigated the effectiveness of geometric deep learning techniques in predicting antibody-antigen interactions. Our results indicate that incorporating geometric information is crucial for accurately predicting epitope and paratope regions. Specifically, the use of an invariant representation in I-GEP models outperformed previous models, and O-GEP models with diffusion layers and additional geometric features achieved state-of-the-art performance. Our study highlights the potential of geometric deep learning in computational biology. Future research could explore using spectral shape analysis to address the more complex problem of conformational rearrangement in antigen-antibody binding [22].

Acknowledgements

This work is supported by the ERC grant no.802554 (SPECGEO), PRIN 2020 project no.2020TA3K9N (LEGO.AI), and PNRR MUR project PE0000013-FAIR. C.D. was supported by the Gatsby Charitable Foundation (GAT3755) and G-research grants for PhD students.

References

[1] Emma C Thomson, Laura E Rosen, James G Shepherd, Roberto Spreafico, Ana da Silva Filipe, Jason A Wojcechowskyj, Chris Davis, Luca Piccoli, David J Pascall, Josh Dillen, et al. Circulating sars-cov-2 spike n439k variants maintain fitness while evading antibody-mediated immunity. *Cell*, 184(5):1171–1187, 2021. 1

[2] Emil Fischer. Einfluss der configuration auf die wirkung der enzyme. *Berichte der deutschen chemischen Gesellschaft*, 27(3):2985–2993, 1894. 1

[3] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021. 1

[4] Hannes Stärk, Octavian Ganea, Lagnajit Pattanaik, Regina Barzilay, and Tommi Jaakkola. Equibind: Geometric deep learning for drug binding structure prediction. In *International Conference on Machine Learning*, pages 20503–20521. PMLR, 2022. 1, 2

[5] Bowen Dai and Chris Bailey-Kellogg. Protein interaction interface region prediction by geometric deep learning. *Bioinformatics*, 37(17):2580–2588, 2021. 1, 2, 3, 8

[6] Yang Zhang, Wenbing Huang, Zhewei Wei, Ye Yuan, and Zhaohan Ding. Equipocket: an e (3)-equivariant geometric graph neural network for ligand binding site prediction. *arXiv preprint arXiv:2302.12177*, 2023. 1, 2

[7] Edgar Liberis, Petar Veličković, Pietro Sormanni, Michele Vendruscolo, and Pietro Liò. Parapred: antibody paratope prediction using convolutional and recurrent neural networks. *Bioinformatics*, 34(17):2944–2950, 2018. 2

[8] Andreea Deac, Petar Veličković, and Pietro Sormanni. Attentive cross-modal paratope prediction. *Journal of Computational Biology*, 26(6):536–545, 2019. 2

[9] 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, 2014. 2

[10] Alice Del Vecchio, Andreea Deac, Pietro Liò, and Petar Veličković. Neural message passing for joint paratope-epitope prediction. *arXiv preprint arXiv:2106.00757*, 2021. 2, 3, 10

[11] Clemens Isert, Kenneth Atz, and Gisbert Schneider. Structure-based drug design with geometric deep learning. *Current Opinion in Structural Biology*, 79:102548, 2023. 2

[12] Jérôme Tubiana, Dina Schneidman-Duhovny, and Haim J Wolfson. Scannet: an interpretable geometric deep learning model for structure-based protein binding site prediction. *Nature Methods*, 19(6):730–739, 2022. 2

[13] Bruna Moreira da Silva, YooChan Myung, David B Ascher, and Douglas EV Pires. epitope3d: a machine learning method for conformational b-cell epitope prediction. *Briefings in Bioinformatics*, 23(1):bbab423, 2022. 2

[14] Pablo Gainza, Freyr Sverrisson, Frederico Monti, Emanuele Rodola, D Boscaini, MM Bronstein, and BE Correia. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods*, 17(2):184–192, 2020. 2

[15] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017. 2, 3

[16] Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pages 9323–9332. PMLR, 2021. 2, 3

[17] Gabriel Cia, Fabrizio Pucci, and Marianne Rooman. Critical review of conformational b-cell epitope prediction methods. *Briefings in Bioinformatics*, 24(1):bbac567, 2023. 2

[18] James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, Jiye Shi, and Charlotte M Deane. Sabdab: the structural antibody database. *Nucleic acids research*, 42(D1):D1140–D1146, 2014. 2

[19] 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, et al. 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. 2

[20] Nicholas Sharp, Souhaib Attaiki, Keenan Crane, and Maks Ovsjanikov. Diffusionnet: Discretization agnostic learning on surfaces. *ACM Transactions on Graphics (TOG)*, 41(3):1–16, 2022. 3, 9

[21] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer graphics forum*, volume 28, pages 1383–1392. Wiley Online Library, 2009. 3

[22] Robyn L Stanfield, Midori Takimoto-Kamimura, James M Rini, Albert T Profy, and Ian A Wilson. Major antigen-induced domain rearrangements in an antibody. *Structure*, 1(2):83–93, 1994. 4

[23] Jens Meiler, Michael Müller, Anita Zeidler, and Felix Schmäschke. Generation and evaluation of dimension-reduced amino acid parameter representations by artificial neural networks. *Molecular modeling annual*, 7(9):360–369, 2001. 7

A Methods

A.1 Data representation

For each protein, we construct a residue graph (Figure 2b), representing residues as nodes and establishing edges between the 15 nearest neighboring residues within a 10 Å radius. Each residue is characterized by a 28-dimensional physicochemical feature vector. This vector encompasses a one-hot encoding of the amino acid, encompassing 20 possible types along with one for an unclassified type. Additionally, seven other features are included that portray the physical, chemical, and structural attributes of the amino acid type. These supplementary features can be viewed as a consistent embedding, as outlined in [23].

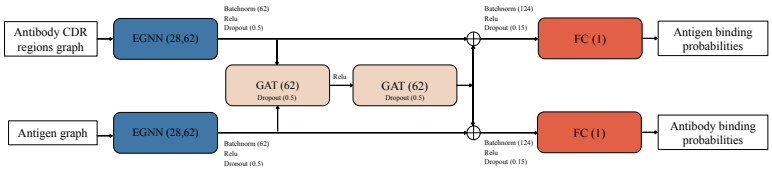
For each protein, we generated a surface mesh (Figure 2e) using the PyMOL API with a 1.4 Å water probe radius. We associated each point on the protein's surface with a residue by finding the closest atom to that point. This association was then used to transfer the feature of each residue to the points on the surface.

A.2 Model

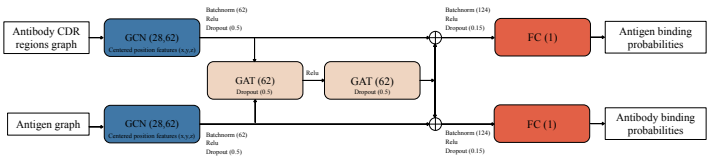
In this Section, we provide a detailed explanation of the models' architecture presented in the paper.

A.2.1 I-GEP

The baseline I-GEP architecture operates as follows: for a given antibody-antigen pair, we independently compute distinct sets of features for each protein in the pair using a Graph Convolutional Network (GCN). It's crucial to note that our model processes these protein pairs without any prior knowledge of their interactions. We then incorporate these distinct features from both the antibody and antigen and fuse them through a two-layer Graph Attention Network (GAT). Specifically, the output of GCN is concatenated after passing through batch normalization, ReLU activation, and a dropout layer. This combined output is subsequently fed into the Graph Attention Layers, responsible for learning weights across the residues of both the antigen and antibody. Finally, the outputs of both the GCN and the two Graph Attention Layers are concatenated to generate two separate predictions, one for the antigen and another for the antibody. To improve the accuracy of our predictions, we integrate geometric information into the I-GEP model using two different approaches discribed in the main text. We visually represent this pipeline in Figure 3b and 3a.



(a) EGNN I-GEP



(b) GCN O-GEP

Figure 3: Models architecture: The layers or modules are depicted using color-coded blocks, with the text inside indicating the respective layer type. In parentheses, we provide the dimensions for each layer: GCN (including inner dimensions and the number of features related to the CDR/antigen), GAT (with inner dimensions specified twice), and FC (1). The arrows indicate the data flow from one module to the next. Additional details about the transformation performed on the input are written.

A.2.2 O-GEP

The O-GEP model is an extension of the architectural foundation established in Pinet, where models store a set of features at each point on the protein surface and generate a binding probability for each point. The input features are computed as explained in A.1 and transferred to the surface representation, allowing close comparison with the two geometric representations.

In Pinet the process begins by individually processing both the antigen and antibody as a point cloud representing the protein surface. A Spatial Transformation Network is employed to ensure invariance to rigid-body transformations for each protein. Subsequently, a multi-layer perceptron (MLP) extracts local surface characteristics. These local surface features are combined into a comprehensive protein feature vector. This process is iteratively applied to generate two representations: a local representation after one iteration and a global representation after multiple iterations, where local features of each point are pooled into a single vector. Once both proteins have undergone this treatment that we refer as the geometric module in Fig 1b, the local surface features of each protein and their global protein features are combined and subjected to further segmentation through a set of 1D convolution neural network. Importantly, the trainable weights for canonical transformations, local feature extraction, and global feature extraction are shared between the two proteins, as in [5]. Finally, the models predict a binding probability for each point on the point could.

In our modification of this architecture, we replace PointNet with DiffNet. This change is advantageous because DiffNet can compute features on both point clouds and meshes independently. We also explore the inclusion of geometric features computed over the protein surface, such as the Heat Kernel Signature. These additional features enhance the model's comprehension of protein interactions, particularly at a finer, more local level. For a more detailed examination of the influence of this mapping and experimental results, please refer to Section C.

B Hyper-parameters

During training, we combined the losses from both tasks, paratope and epitope prediction. To enhance model robustness, we applied random rotations to dataset instances. Hyperparameter tuning involved a search for the optimal learning rate from the set $\{10^{-2}, 10^{-3}, 5 \times 10^{-3}, 10^{-5}\}$ and kept the model with the best performance on the validation set. . After the hyperparameter search, we found that the best learning rates were: 10^{-3} for EPMP and PiNet, 10^{-2} for $E(n)$-EPMP, 5×10^{-3} for DiffNet. All models were trained for 200 epochs to ensure validation loss saturation, and the weights yielding the best validation metrics during training were selected. We conducted training with five random seeds for each model, evaluating performance using the weights yielding the best validation set results in each run.

The surface generated by PyMOL is composed of around 14k points. To ease and fast the training procedure we subsampled the surface considering only 2k points. In the case of point clouds, we used a random subsampling during training, while for the mesh we used a simplification method base on quadric error metrics.

B.1 Layer dimensions

For the EPMP$_{xyz}$ model, we use a graph convolution layer with inner dimension 31 and two GAT layers with inner dimension 62. In contrast, for the $E(n)$-EPMP, we use one $E(n)$-invariant layer with an inner dimension of 28 and two GAT layers with inner dimension 56.

For all the O-GEP models, the geometric module comprises two layers with dimensions 64 and 128, while the segmentation module is composed of two layers with dimensions 64 and 32.

C Outer residues

The outer representation can't include the inner residues because they are too far from the protein's surface representation. As a result, the O-GEP model can't predict those residues. To see how this affects the predictions, we show the results for both I-GEP and O-GEP in Table 1, considering only the outer residues represented by the surface.

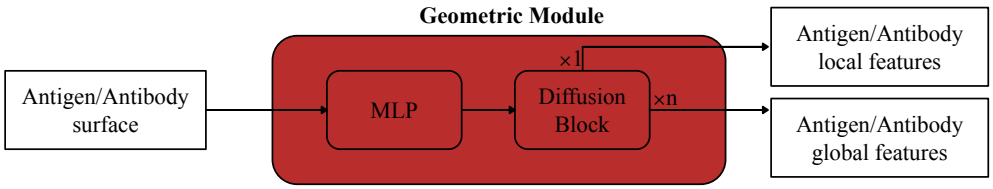


Figure 4: Geometric module: The arrows indicate the data flow from one module to the next. The protein representation, is first past through a MLP layer before entering the diffusion block as defined in [20]. The local and global features are computed by applying the diffusion block a single and n times respectively. Additional details about the transformation performed on the input are written.

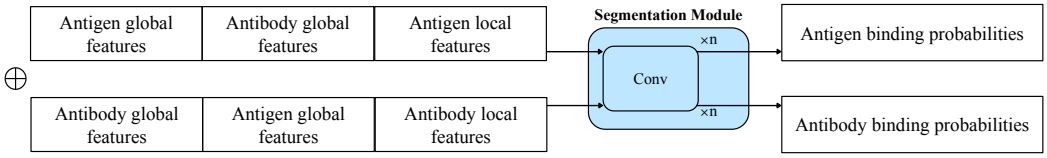


Figure 5: Segmentation Module: The output of the geometric module are concatenated in to two vectors for the antigen and the antibody respectively. These representation are then sent through the segmentation module to respectively output the binding prediction on the antigen and antibody respectivaly. The segmentation module is shared accross the two representations and consiste of a set of convolutional layers.

D Qualitative representations

It's important to recognize that the challenges posed by prediction on antibody and antigen are inherently different, leading to varying degrees of specialization among different methods and representations as observed in Figure 2f and 6j for the O-GEP with DiffNet.

In the context of antibody, it is often sufficient to rely primarily on sequence information, as these interactions typically involve localized binding with relatively consistent structural patterns across different antibody-antigen pairs. Moreover the interaction of the anitbody is limited to a small portion of the protein (CDR). This could explain why surface features like the heat kernel signature (HKS) prove effective, as they capture localized and specific structural characteristics.

Conversely, predicting antigen binding sites requires a broader and more global set of features due to the nature of antigen-protein interactions. This is evident from the fact that using 3D coordinates alone already yields reasonable predictive performance as found in the Table 2f and represented in Figure 6b,6h,6e.

On the contrary, HKS alone may not be sufficient to propagate information globally across the antigene leading to high binding predictions across all geometry of the antigen (see Figures 6g and 6d). This trend is also observed in the failure case of our model shown in Figure 7j and 8j.

It's the combination of both 3D coordinates and HKS that allows our proposed model to integrate both local and global information effectively, resulting in improved performance as depicted in Figure 6i,6f. While geometric information significantly contributes to prediction accuracy, there are cases where it can lead to incorrect predictions. This is evident in the antibody-antigen complex '1n8z' (refer to Figure 8). In this particular configuration, multiple geometrically fitting regions exist on the antigen. Consequently, the majority of O-GEP models tend to make incorrect predictions for this region (as illustrated in Figures 8e, 8b, 8h, 8f, and Figure 8i). An important observation is that even when the epitope and paratope predictions are incorrect, they are still close in the 3D space. This suggests that our model effectively learns to establish meaningful communication between protein features without prior 3D knowledge of their relative position.

Regarding the choice between mesh representation and point cloud, it's important to note that when the input features are consistent, the performance difference between the two representations is

Table 1: Quantitative results evaluated on the surface residues. We report the mean and standard deviation ($\pm$) over multiple runs.

(a) I-GEP

	Antigen			Antibody		
	MCC	AUC ROC	AUC PR	MCC	AUC ROC	AUC PR
EPMP	0.08 ± 0.01	0.58 ± 0.01	0.13 ± 0.00	0.33 ± 0.03	0.74 ± 0.01	0.56 ± 0.01
EPMP$_{xyz}$	0.08 ± 0.01	0.60 ± 0.01	**0.16** ± 0.01	0.33 ± 0.02	0.74 ± 0.01	0.56 ± 0.01
$E(n)$-EPMP	**0.11** ± 0.01	**0.64** ± 0.01	**0.16** ± 0.01	**0.39** ± 0.11	**0.78** ± 0.07	**0.63** ± 0.08

(b) O-GEP

	Antigen			Antibody		
	MCC	AUC ROC	AUC PR	MCC	AUC ROC	AUC PR
PiNet (xyz)	0.38 ± 0.04	0.87 ± 0.01	0.45 ± 0.02	0.26 ± 0.12	0.77 ± 0.03	0.52 ± 0.08
PiNet (xyz+hks)	0.29 ± 0.05	0.84 ± 0.02	0.37 ± 0.04	0.13 ± 0.06	0.64 ± 0.01	0.47 ± 0.04
DiffNet$_{pc}$ (xyz)	0.40 ± 0.05	**0.88** ± 0.01	0.49 ± 0.02	0.26 ± 0.06	0.71 ± 0.02	0.56 ± 0.04
DiffNet$_{pc}$ (hks)	0.05 ± 0.04	0.58 ± 0.03	0.14 ± 0.01	0.40 ± 0.02	**0.81** ± 0.01	0.69 ± 0.01
DiffNet$_{pc}$ (xyz+hks)	**0.43** ± 0.03	**0.88** ± 0.01	**0.50** ± 0.02	0.19 ± 0.05	0.68 ± 0.06	0.51 ± 0.05
DiffNet$_{m}$ (xyz)	0.41 ± 0.03	**0.88** ± 0.01	0.49 ± 0.05	0.20 ± 0.07	0.69 ± 0.03	0.53 ± 0.03
DiffNet$_{m}$ (hks)	0.05 ± 0.01	0.56 ± 0.02	0.14 ± 0.01	**0.43** ± 0.02	0.80 ± 0.01	**0.70** ± 0.01
DiffNet$_{m}$ (xyz+hks)	0.41 ± 0.06	**0.88** ± 0.02	0.46 ± 0.07	0.23 ± 0.06	0.68 ± 0.04	0.52 ± 0.04

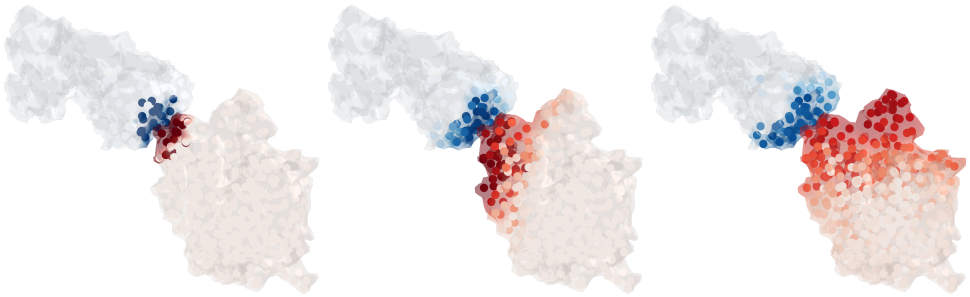
relatively small, typically within a range of 0.4 and usually overlapping in our experiments (see Table 2d). Factors contributing to this slight difference may include the computation of eigenvectors and the inherent structure of the representations themselves. Point cloud representation offers greater flexibility as it doesn't impose connectivity constraints on neighboring nodes, allowing for a more adaptable representation for proteins 6j.

In summary, the observed specialization of different methods and representations in O-GEP can be attributed to the inherent differences in the protein interaction prediction tasks, where localized interactions in the antibody contrast with the more global features required for antigen binding. Our model leverages these distinctions by tailoring its approach to each task, improving predictive performance.

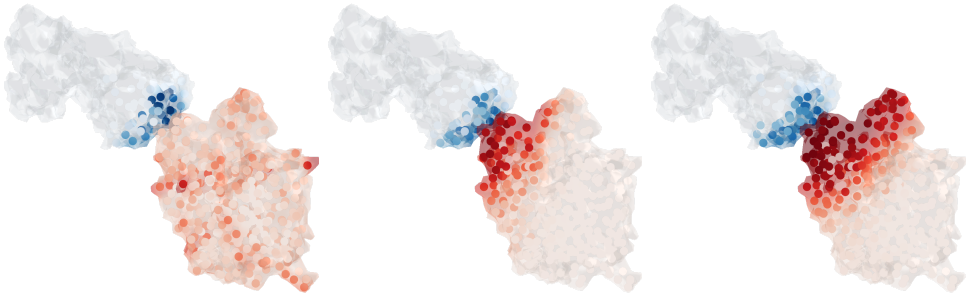
An important observation to make is that the I-GEP model does not encounter a similar issue because it utilizes the same set of positional features acrosss versions of the model, which is illustrated in Figure 6. It's worth noting that the I-GEP model does not face a similar issue since it consistently employs the same set of positional features across different model versions, as demonstrated in Figure 6.

In the case of an I-GEP failure ('3raj' in Figure 7), this model tends to assign a high probability to spiky edges, even when the actual binding region on the antigen is in a flat region.

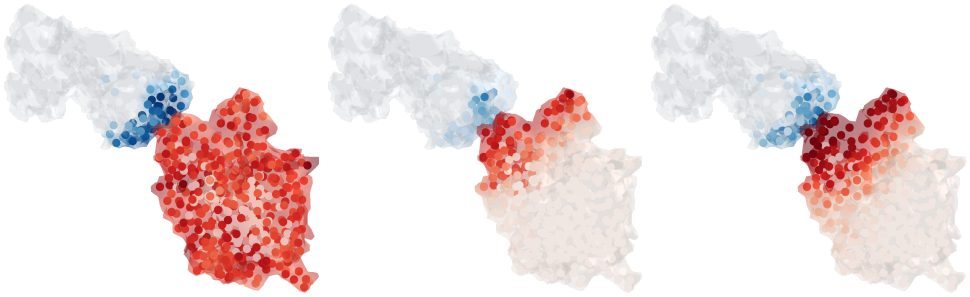
These qualitative examples highlight the significance of both I-GEP and O-GEP models on geometric information. As shown, this focus can be both advantageous and disadvantageous. We hypothesize that this behavior may also be influenced by the relatively low number of physicochemical features (28) compared to those used in similar studies, such as 63 features in [10]. We leave this analysis for future work.



(a) Ground Truth **(b)** PiNet (xyz) **(c)** PiNet (xyz+hks)

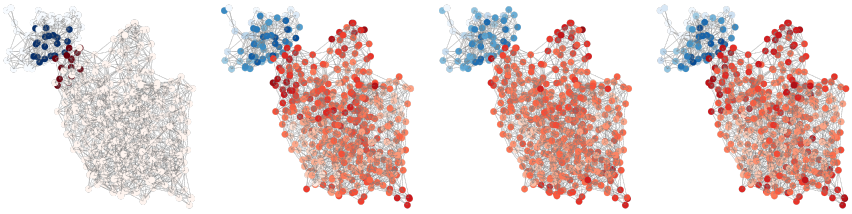


(d) DiffNet$_{pc}$ (hks) **(e)** DiffNet$_{pc}$ (xyz) **(f)** DiffNet$_{pc}$ (xyz+hks)



(g) DiffNet$_{mesh}$ (hks) **(h)** DiffNet$_{mesh}$ (xyz) **(i)** DiffNet$_{mesh}$ (xyz+hks)

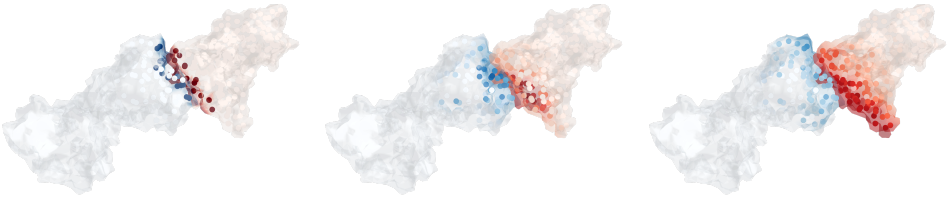
(j) O-GEP results. Plot of the surface along with the corresponding residues.



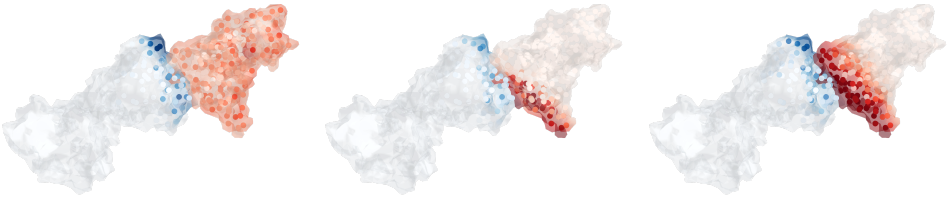
(k) Ground Truth **(l)** EPMP **(m)** EPMP$_{xyz}$ **(n)** $E(n)$-EPMP

(o) I-GEP results. Plot of the graph of residues with the continuous binding predictions

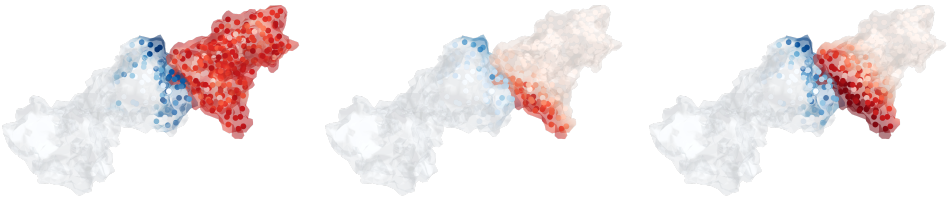
Figure 6: All O-GEP and I-GEP models results for the antibody-antigen complex '4jr9'. The continuous binding predictions are represented as a color gradient in blue and red for the antigen and antibody, respectively.



(a) Ground Truth **(b)** PiNet (xyz) **(c)** PiNet (xyz+hks)

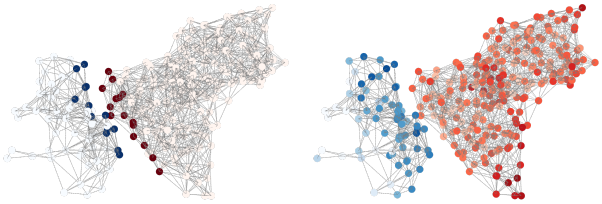


(d) DiffNet$_{pc}$ (hks) **(e)** DiffNet$_{pc}$ (xyz) **(f)** DiffNet$_{pc}$ (xyz+hks)

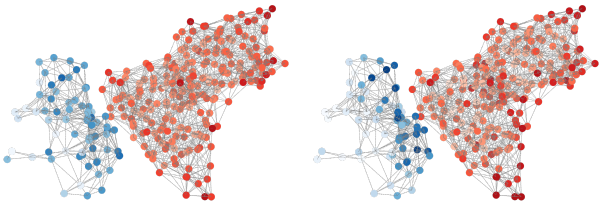


(g) DiffNet$_{mesh}$ (hks) **(h)** DiffNet$_{mesh}$ (xyz) **(i)** DiffNet$_{mesh}$ (xyz+hks)

(j) O-GEP results. Plot of the surface along with the corresponding residues.



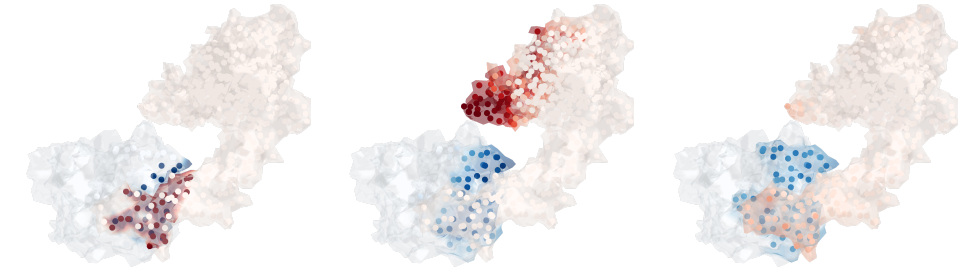
(k) Ground Truth **(l)** EPMP



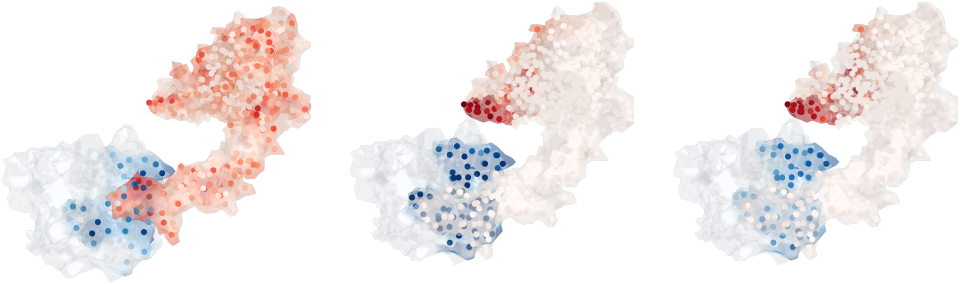
(m) EPMP$_{xyz}$ **(n)** $E(n)$-EPMP

(o) I-GEP results. Plot of the graph of residues with the continuous binding predictions

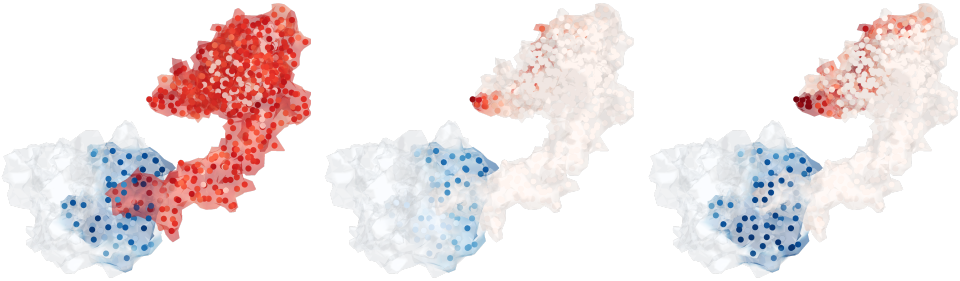
Figure 7: All O-GEP and I-GEP models results on the antibody-antigen complex with the lowest MCC on the epitope prediction performed by $E(n)$-EPMP: '3raj'.



(a) Ground Truth **(b)** PiNet (xyz) **(c)** PiNet (xyz+hks)

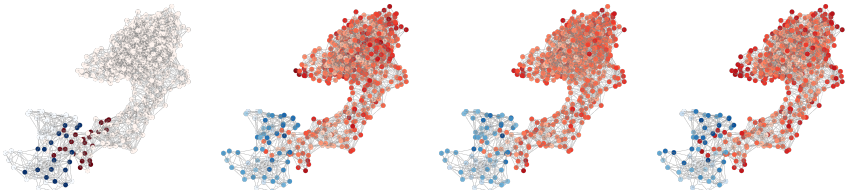


(d) DiffNet$_{pc}$ (hks) **(e)** DiffNet$_{pc}$ (xyz) **(f)** DiffNet$_{pc}$ (xyz+hks)



(g) DiffNet$_{mesh}$ (hks) **(h)** DiffNet$_{mesh}$ (xyz) **(i)** DiffNet$_{mesh}$ (xyz+hks)

(j) O-GEP results. Plot of the surface along with the corresponding residues.



(k) Ground Truth **(l)** EPMP **(m)** EPMP$_{xyz}$ **(n)** $E(n)$-EPMP

(o) I-GEP results. Plot of the graph of residues with the continuous binding predictions

Figure 8: All O-GEP and I-GEP models results on the antibody-antigen complex with the lowest MCC on the epitope prediction done by DiffNet$_{pc}$ (xyz+hks): '1n8z'. The continuous binding predictions are represented as a color gradient in blue and red for the antigen and antibody, respectively.