ABlooper: Fast accurate antibody CDR loop structure prediction with accuracy estimation

Brennan Abanades[1], Guy Georges[2], Alexander Bujotzek[2], and Charlotte M. Deane[1]

[1]Department of Statistics, University of Oxford, Oxford, UK.
[2]Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Penzberg, Germany.

July 2021

Abstract

Antibodies are a key component of the immune system and have been extensively used as biotherapeutics. Accurate knowledge of their structure is central to understanding their function. The key area for antigen binding and the main area of structural variation in antibodies are concentrated in their six complementarity determining regions (CDRs), with the most variable being the CDR-H3 loop. The sequence and structure variability of CDR-H3 make it particularly challenging to model. Recently, deep learning methods have offered a step change in our ability to predict protein structures. In this work we present ABlooper, an end-to-end equivariant deep-learning based CDR loop structure prediction tool. ABlooper predicts the structure of CDR loops with high accuracy and provides a confidence estimate for each of its predictions. On the models of the Rosetta Antibody Benchmark, ABlooper makes predictions with an average H3 RMSD of 2.45Å, which drops to 2.02Å when considering only its 76% most confident predictions.

1 Introduction

1.1 Antibody Structure

Antibodies are a class of protein produced by B-cells during an immune response. Their ability to bind with high affinity and specificity to almost any antigen, makes them attractive for use as therapeutics [1].

Knowledge of the structure of antibodies is becoming increasingly important in therapeutic development [2]. However, experimental structure determination is time-consuming and expensive so it is not always practical or even possible to use routinely during therapeutic development. Computational modelling

tools have allowed researchers to bridge this gap by predicting large numbers of antibody structures to a high level of accuracy [3, 4]. For example, models of antibody structures have been used for virtual screening [5] and to identify coronavirus-binding antibodies that bind the same epitope with very different sequences [6].

The overall structure of all antibodies is similar and therefore can be accurately predicted using current methods [3]. The area of antibodies that it is hardest to model are the sequence variable regions that provide the structural diversity necessary to bind a wide range of antigens. This diversity is largely focused in six loops known as the complementarity determining regions (CDRs). The most diverse of these CDRs and therefore the hardest to model is the third CDR loop of the heavy chain (H3) [7].

1.2 Deep Learning for Protein Structure Prediction

At CASP14 [8], DeepMind showcased AlphaFold2 [9], a neural network capable of accurately predicting many protein structures. The method relies on the use of equivariant neural networks and an attention mechanism [10]. More recently, RoseTTAFold, a novel neural network based on equivariance and attention was shown to obtain results comparable to those of AlphaFold2 [11].

For a network to be equivariant with respect to a group, it must be able to commute with the group action. For rotations, this means that rotating the input before feeding it into the network will have the same result as rotating the output. In the case of proteins, using a network equivariant to both translations and rotations in 3D space allows us to learn directly from atom coordinates. This is in contrast to previous methods like TrRosetta [12] or the original version of AlphaFold [13] that predict invariant features, such as inter-residue distances and orientations which are then used to reconstruct the protein.

A number of approaches for developing equivariant networks have been developed recently. These can be classified into three distinct groups: those based on Lie-Algebra [14], those based on irreducible representations [15, 16] and those based on invariant features such as E(n)-Equivariant Graph Neural Networks (E(n)-EGNN) [17]. In this work, we choose to use E(n)-EGNNs due to their speed and simplicity.

1.3 Deep Learning for Antibody Structure Prediction

Deep learning based approaches have also been shown to improve structure prediction in antibodies, for example DeepH3 [18], an antibody specific version of TrRosetta, outperformed TrRosetta when predicting CDR-H3 loops. Recently DeepAb [4], was shown to outperform all currently available antibody structure prediction methods. For both DeepAb and DeepH3, deep learning is used to obtain inter-residue geometries that are then fed into energy minimisation methods to produce the final structure.

In this work we present ABlooper, a fast and accurate tool for antibody CDR loop structure prediction. By leveraging E(n)-EGNNs, ABlooper is able

to directly predict the structure of CDR loops, removing the need for computationally expensive energy minimisation methods. Additionally, by simultaneously predicting multiple structures for each loop and comparing them amongst themselves, ABlooper is capable of estimating a confidence measure for each predicted loop.

2 Methods

2.1 Data

The data used to train, test and validate ABlooper was extracted from SAbDab [19], a database of all antibody structures contained in the PDB [20]. Structures with a resolution better than 3.0Å and no missing backbone atoms within any of the CDRs were selected. The CDRs were defined using the IMGT numbering scheme [21].

For easy comparison with different pipelines, we used the 49 antibodies from the Rosetta Antibody Benchmark as our test set. For validation, 100 structures were selected at random. It was ensured that there were no structures with the same CDR sequences in the training, testing and validation sets. Sequence redundancy was allowed within the training set to expose the network to the existence of antibodies with identical sequence but different structural conformations. This resulted in a total of 3438 training structures.

Additionally, we use a secondary test set composed of 114 antibodies (SAbDab Latest Structures) with a resolution of under 2.3 Å and a maximum CDR-H3 loop length of 20, which were added to SAbDab after the initial test, train and validation sets were extracted (8^{th}November 2020 − 24^{th}May 2021). A list containing the PDB IDs of all the structures used in the train, test, and validation sets is given in the SI. ABodyBuilder was used to build models of all the structures (see SI for more details) and all tests were carried out on these models

2.2 Network

ABlooper is composed of five E(n)-EGNNs, each one with four layers, all trained in parallel. The model is trained on the position of the C_α-N-C-C_β backbone atoms for all six CDR loops plus two anchor residues at either side. An input geometry is generated by flattening each of the CDR loops while leaving the anchor residues in place. The model is given four different types of features per residue resulting in a 41-dimensional vector. These include a one-hot encoded vector describing the amino acid type, the atom type and which loop the node belongs to. Additionally, sinusoidal positional embeddings are given to each residue describing how close it is to the anchors.

Each of the five networks is trained on antibody crystal structures to independently predict the loop conformation for all of the CDRs. The RMSD between the prediction and true structure is used as the loss function. The

output from the five networks is then averaged to obtain a final prediction. To ensure the final prediction is physically plausible, an L1 loss between all intra-residue atom distances is used. More details on the implementation and training of ABlooper can be found in the SI.

3 Results

3.1 Using ABlooper to predict CDR loops on modelled antibody structures

We used ABlooper to predict the CDRs on ABodyBuilder models of the Rosetta Antibody Benchmark (RAB) and SAbDab Latest Structures (SLS) sets. The RMSD between the C_α-N-C-C_β atoms in the backbone of the crystal structure and the predicted CDRs for both test sets data sets is shown in Table 1.

Method	CDR-H1		CDR-H2		CDR-H3		CDR-L1		CDR-L2		CDR-L3	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Rosetta Antibody Benchmark												
ABodyBuilder	1.08	0.67	0.99	0.71	2.77	2.66	0.69	0.48	0.50	0.46	1.12	0.84
ABlooper	0.90	0.61	1.03	0.81	2.45	1.99	0.61	0.44	0.51	0.46	0.93	0.63
SAbDab Latest Structures												
ABodyBuilder	1.24	0.85	1.07	0.83	3.25	2.94	0.88	0.60	0.57	0.46	1.03	0.72
ABlooper	1.14	0.91	0.99	0.85	2.66	2.37	0.73	0.58	0.54	0.44	1.01	0.74

Table 1: Performance comparison between ABodyBuilder and ABlooper for two test sets. The mean and median RMSD to the crystal structure across each test set for the six CDRs is shown. RMSDs for each CDR are calculated after superimposing their corresponding chain to the crystal structure. RMSDs are given in Angstroms (Å).

ABlooper achieves lower mean and median RMSDs than AbodyBuilder for most of the CDRs (Table 1). By far the largest improvement observed is for the CDR-H3 loop, where due to the large structural diversity, homology modelling struggles the most [3]. As ABlooper is predicting on model frameworks generated by ABodyBuilder, we examined its resilience to the small deviations seen in these models and found little to no correlation between framework error and CDR prediction error (see SI).

3.2 Prediction diversity as a measure of prediction quality

ABlooper predicts five structures for each loop. We found that the average RMSD between predictions can be used as a measure of certainty of the final averaged prediction. If all five models agree on the same conformation, then it is more likely that it will be the correct conformation, if they do not, then the final prediction is likely to be less accurate (Figure 1B,1C). This allows ABlooper to

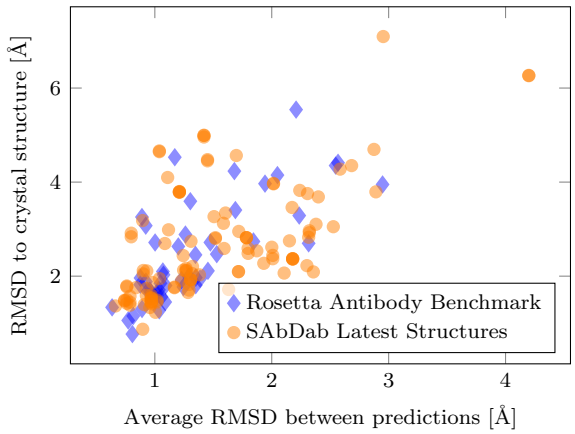
give a confidence score for each predicted loop. As shown in Figure 1D, this score can be used as a filter, removing structures which are expected to be incorrectly modelled by ABlooper. For example, by setting a 1.5Å inter-prediction RMSD cut-off on structures from the Rosetta Antibody Benchmark, the average CDR-H3 RMSD for the set can be reduced from 2.45Å to 2.02Å while keeping 76% of predictions. Examining the same set of CDR-H3s selected out by ABlooper, there is also a reduction in RMSD for ABodyBuilder predictions from 2.77Å to 2.31Å. As expected, accuracy filtering has a tendency to remove longer CDR-H3 predictions but it is not exclusively correlated to length (see SI).

4 Discussion

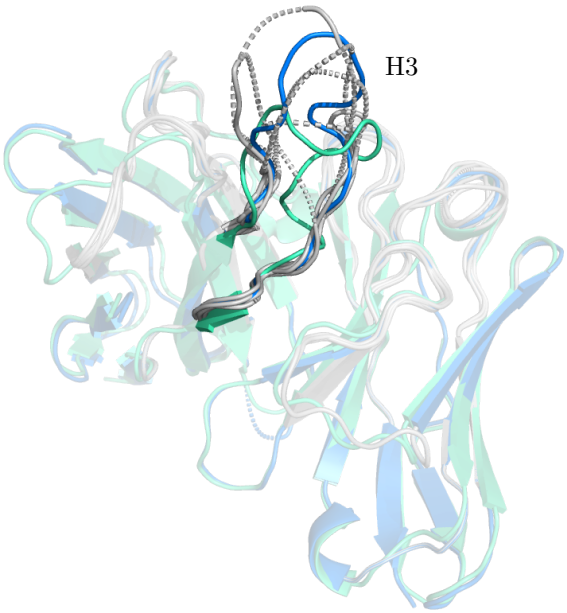
We present ABlooper, a fast and accurate tool for predicting the structure of the CDR loops in antibodies. It builds on recent advances in equivariant graph neural networks to improve CDR loop structure prediction. Furthermore, by predicting each loop multiple times, it is capable of producing an accuracy estimate for each generated loop structure.

ABlooper outperforms ABodyBuilder, one of the current state of the art methods for antibody structure prediction. On an NVIDIA Tesla V100 GPU, ABlooper can predict the CDRs for one hundred structures in under five seconds. Preliminary results suggest that ABlooper produces antibody models of similar accuracy to both AlphaFold2 and DeepAb, but on a far faster timescale. The model used for ABlooper is available at: https://github.com/oxpig/ABlooper.

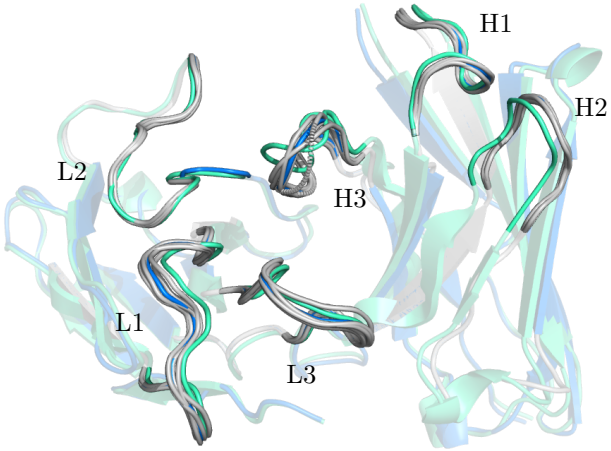
A



B



C



D

Threshold [Å]	RAB [Å]	SLS [Å]
–	2.45 (100%)	2.66 (100%)
2.00	2.18 (86%)	2.39 (75%)
1.50	2.02 (76%)	2.29 (59%)
1.25	1.90 (57%)	2.11 (43%)
1.00	1.74 (27%)	1.71 (27%)

Figure 1: (**A**) CDR-H3 loop RMSD between final averaged prediction and crystal structure compared to average RMSD between predictions for both the Rosetta Antibody Benchmark and the SAbDab Latest Structures set. (**B**) An example of a poorly predicted CDR-H3 loop. All five predictions are given in white, with the final averaged prediction in blue and the crystal structure in green. The predictions from the five networks are very different, indicating an incorrect final prediction. (**C**) Example of correctly predicted CDR loops. All five predictions are similar, indicating a high confidence prediction. (Colours are the same as in (B)) (**D**) Effect of removing structures with a high CDR-H3 inter-prediction RMSD on the averaged RMSD for the set. The number of structures remaining after each quality cut-off is shown as a percentage. Data shown for the Rosetta Antibody Benchmark (RAB) and the SAbDab Latest Structures (SLS) set.

References

[1] Paul J Carter and Greg A Lazar. Next generation antibody drugs: pursuit of the'high-hanging fruit'. *Nature Reviews Drug Discovery*, 17(3):197, 2018.

[2] Mark L Chiu, Dennis R Goulet, Alexey Teplyakov, and Gary L Gilliland. Antibody structure and function: The basis for engineering therapeutics. *Antibodies*, 8(4):55, 2019.

[3] Jinwoo Leem, James Dunbar, Guy Georges, Jiye Shi, and Charlotte M Deane. ABodyBuilder: Automated antibody structure prediction with data–driven accuracy estimation. In *MAbs*, volume 8, pages 1259–1268. Taylor & Francis, 2016.

[4] Jeffrey A Ruffolo, Jeremias Sulam, and Jeffrey J Gray. Antibody structure prediction using interpretable deep learning. *bioRxiv*, 2021.

[5] Constantin Schneider, Andrew Buchanan, Bruck Taddese, and Charlotte M Deane. DLAB-deep learning methods for structure-based virtual screening of antibodies. *bioRxiv*, 2021.

[6] Sarah A Robinson, Matthew IJ Raybould, Claire Marks, Constantin Schneider, Wing Ki Wong, and Charlotte M Deane. Epitope profiling of coronavirus-binding antibodies using computational structural modelling. *bioRxiv*, 2021.

[7] Alexey Teplyakov, Jinquan Luo, Galina Obmolova, Thomas J Malia, Raymond Sweet, Robyn L Stanfield, Sreekumar Kodangattil, Juan Carlos Almagro, and Gary L Gilliland. Antibody modeling assessment ii. structures and models. *Proteins: Structure, Function, and Bioinformatics*, 82(8):1563–1582, 2014.

[8] Andriy Kryshtafovych, Torsten Schwede, Maya Topf, Krzysztof Fidelis, and John Moult. Critical assessment of methods of protein structure prediction (CASP)—Round XIII. *Proteins: Structure, Function, and Bioinformatics*, 87(12):1011–1020, 2019.

[9] 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*, pages 1–11, 2021.

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

[11] Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch,

R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 2021.

[12] Jianyi Yang, Ivan Anishchenko, Hahnbeom Park, Zhenling Peng, Sergey Ovchinnikov, and David Baker. Improved protein structure prediction using predicted interresidue orientations. *Proceedings of the National Academy of Sciences*, 2020.

[13] Andrew W Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander WR Nelson, Alex Bridgland, et al. Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792):706–710, 2020.

[14] Marc Finzi, Max Welling, and Andrew Gordon Wilson. A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. *arXiv preprint arXiv:2104.09459*, 2021.

[15] Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3D point clouds. *arXiv preprint arXiv:1802.08219*, 2018.

[16] Fabian B Fuchs, Daniel E Worrall, Volker Fischer, and Max Welling. SE (3)-transformers: 3D roto-translation equivariant attention networks. *arXiv preprint arXiv:2006.10503*, 2020.

[17] Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. *arXiv preprint arXiv:2102.09844*, 2021.

[18] Jeffrey A Ruffolo, Carlos Guerra, Sai Pooja Mahajan, Jeremias Sulam, and Jeffrey J Gray. Geometric potentials from deep learning improve prediction of CDR H3 loop structures. *Bioinformatics*, 36(Supplement_1):i268–i275, 2020.

[19] 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.

[20] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.

[21] 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.