

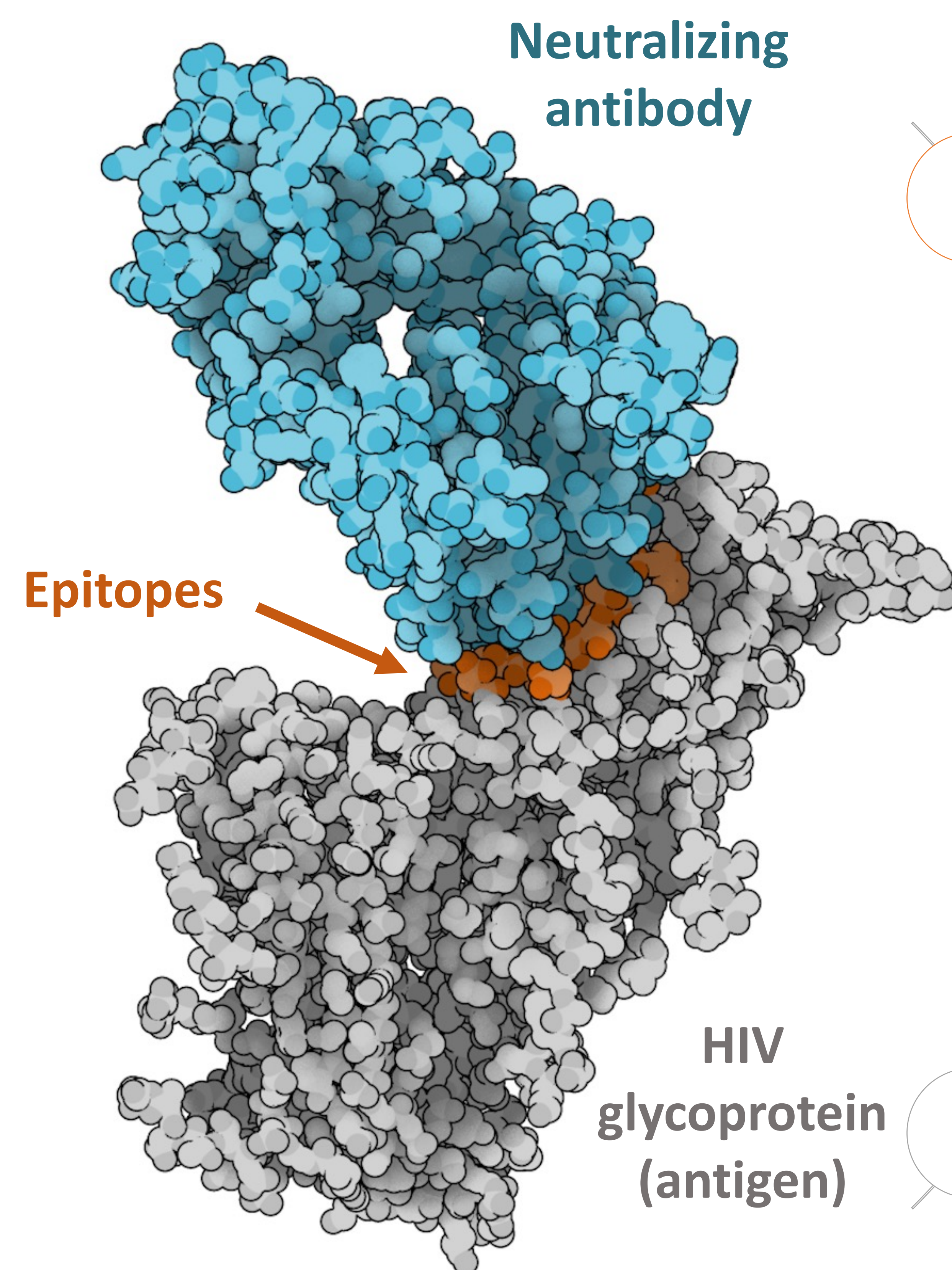


Prediction of epitopes using AlphaFold2 structures and graph neural networks

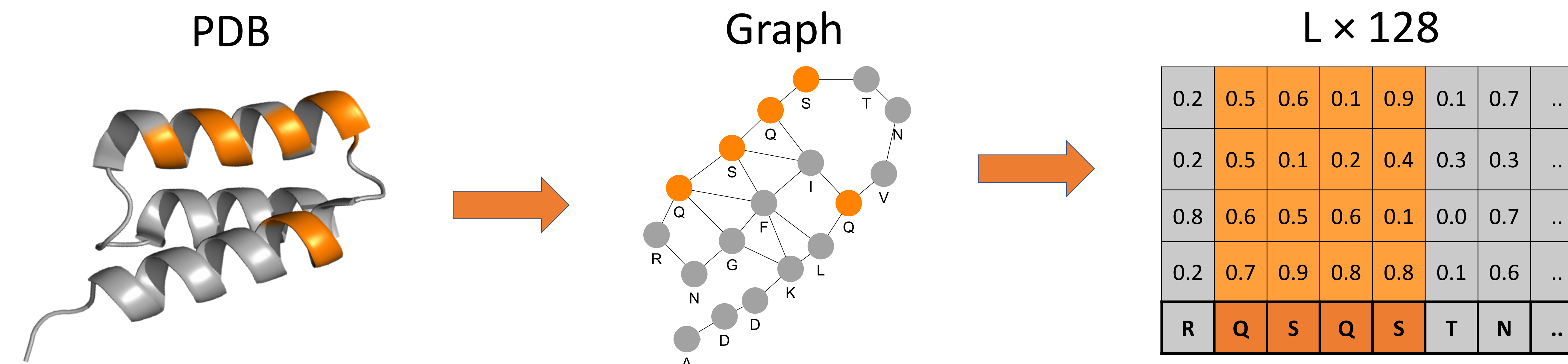
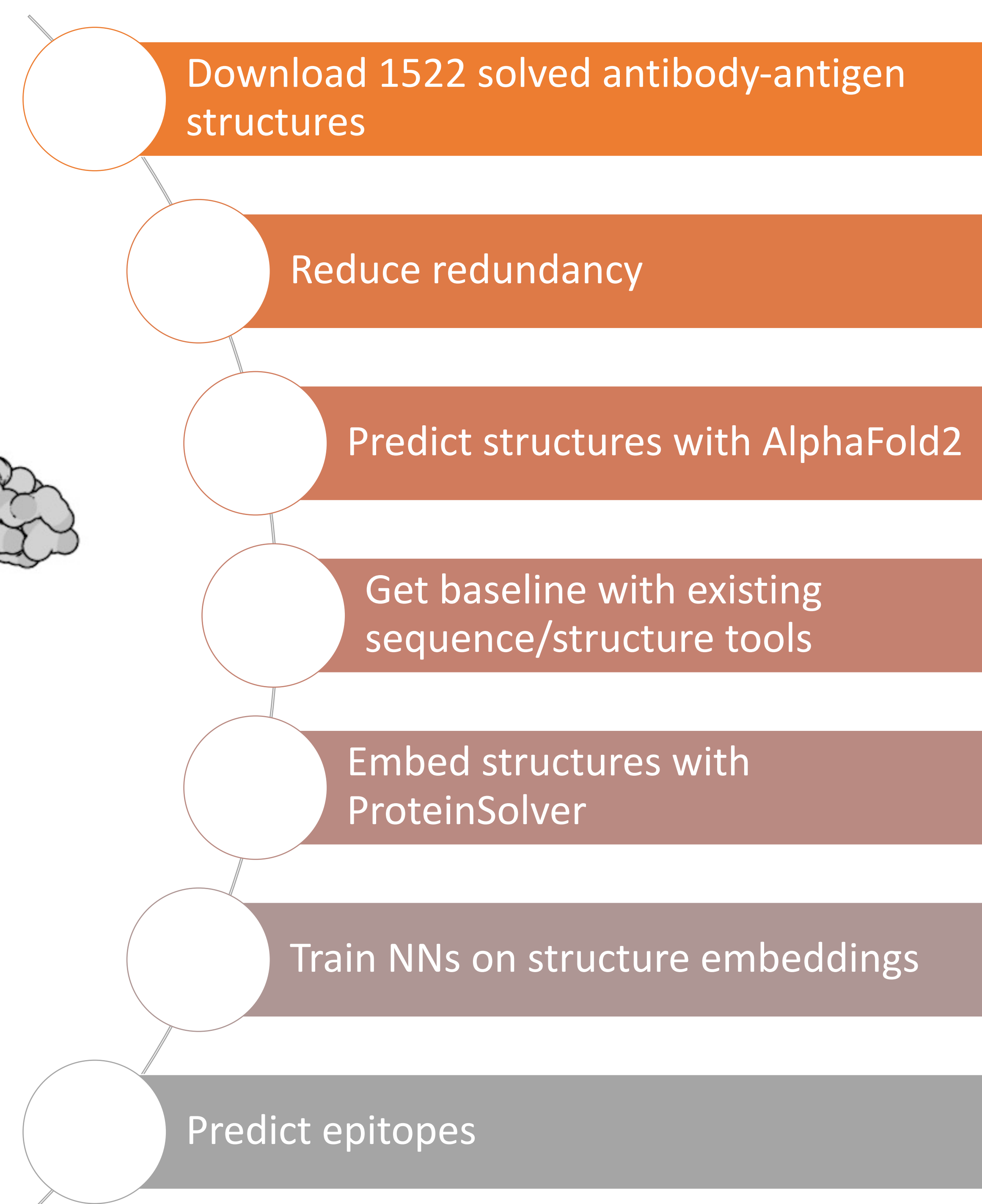
Frederik S. Gade (s184260), Magnus H. Høie (s216559), Eduardo Moreno (s180038), Ida K.S. Meitil (s153020), Supervisor: Paolo Marcatili

Introduction

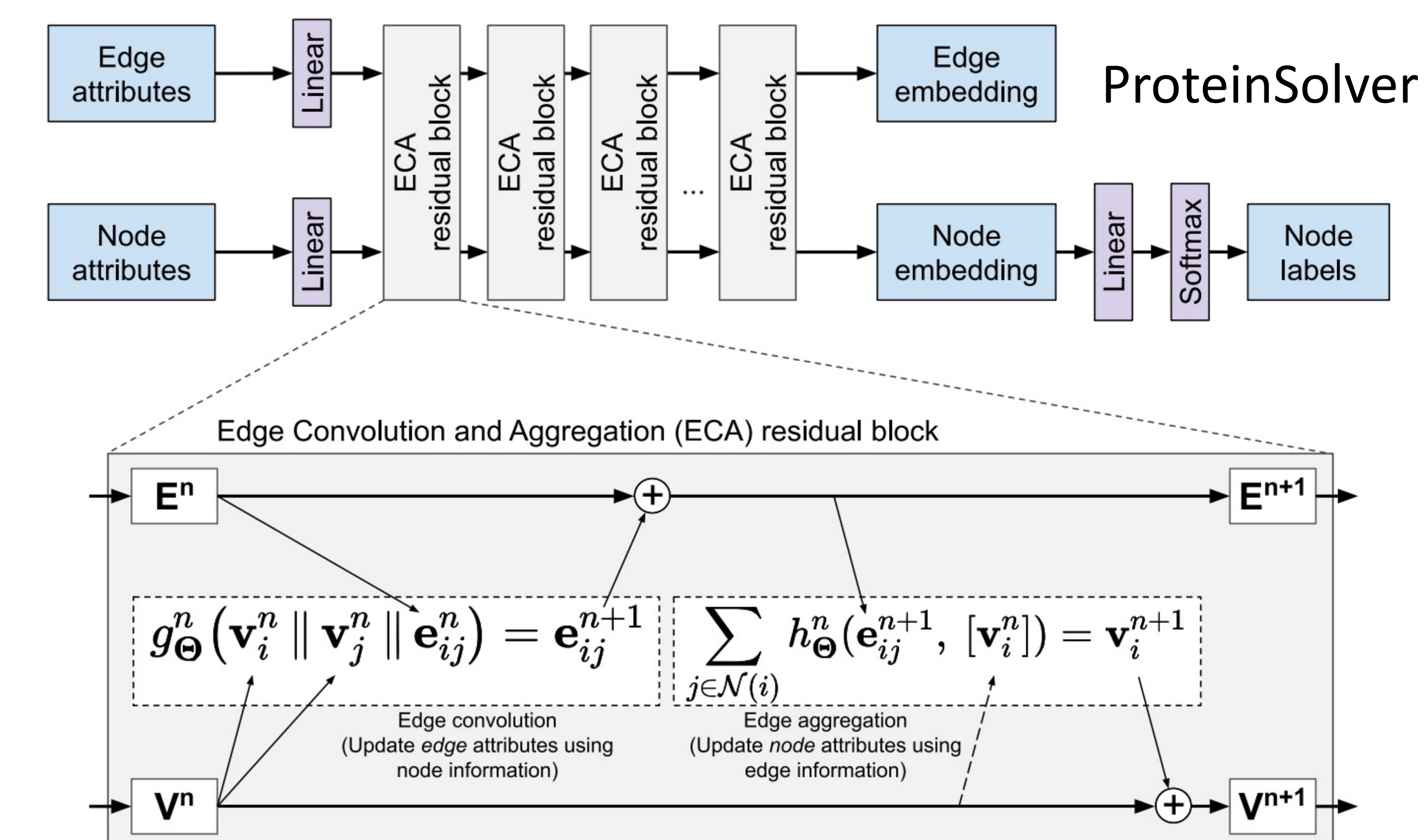
Epitopes are surface-exposed regions of a pathogenic molecule or antigen, which are targeted by the adaptive immune system by e.g. B-cell receptors. Binding is largely determined by the surface features of the target molecule. AlphaFold2 is a deep-learning protein folding model achieving near experimental quality prediction for many proteins. Furthermore, graph-based neural networks such as ProteinSolver allow structural representation of proteins suitable for tasks such as epitope prediction. We investigate improved epitope prediction using Alphafold2 modelled structures over sequence-only models.



Workflow



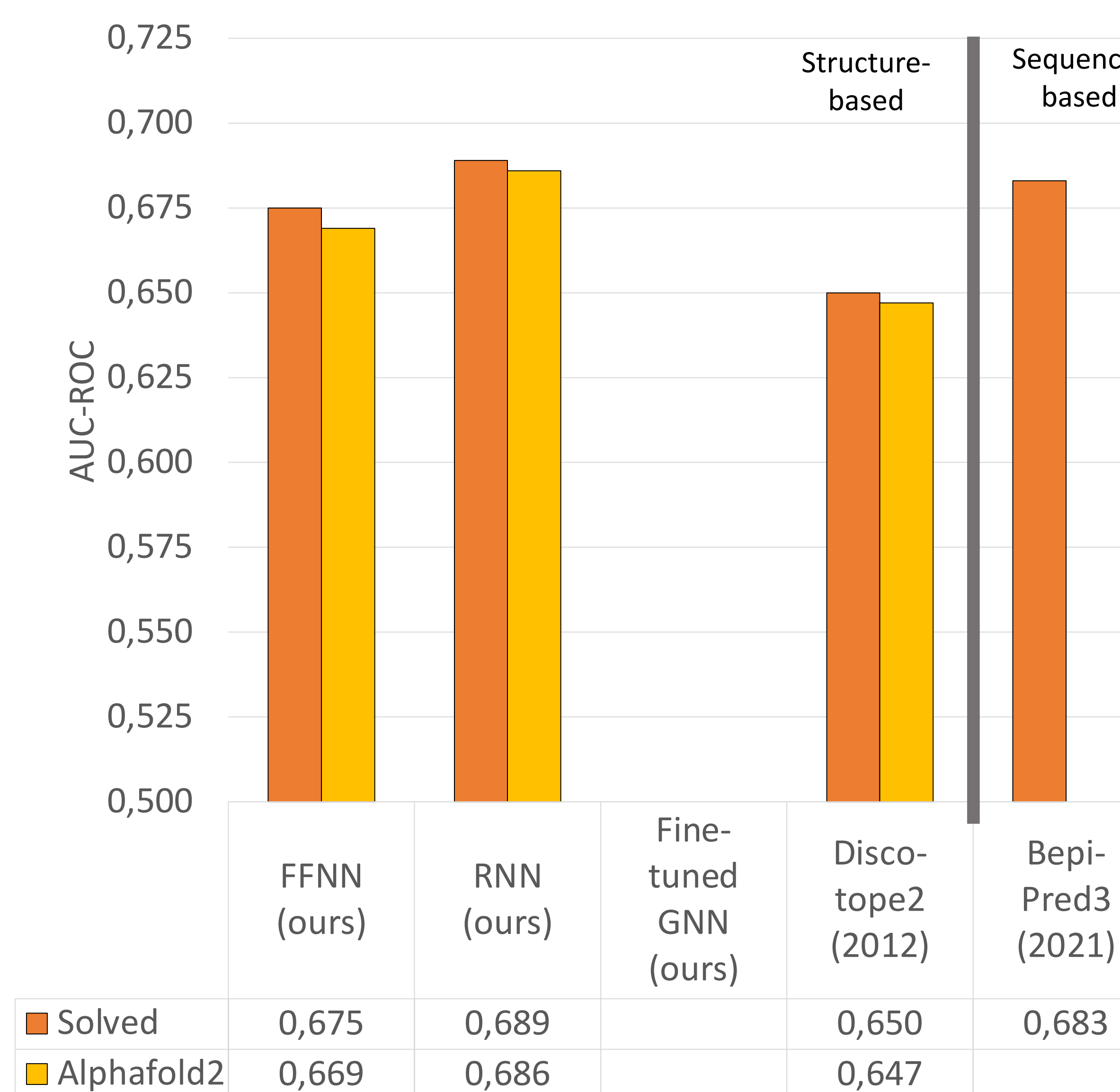
1. Model antigen structure with **AlphaFold2** OR use experimentally solved structure
2. Embed structure in pre-trained graph neural network (**ProteinSolver**)
3. Extract structure residue embeddings from nodes



Network architectures:

- FFNN – Individual residues only (128)
- RNN – Along sequence residues ($L \times 128$)
- GNN – Adapted from ProteinSolver ($L \times 128$)
- DiscoTope2 – Random forest model, surface accessibility
- BepiPred3 – Sequence language model based on ESM

Epitope prediction performance



* Test set limited to PDBs deposited after AlphaFold2 training dataset collection date 30th April 2018

References

Jespersen, M. C., Peters, B., Nielsen, M., & Marcatili, P. (2017). BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research*, 45(W1), W24–W29. <https://doi.org/10.1093/nar/gkx346>

Jumper, J et al (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, May, 1–12. <https://doi.org/10.1038/s41586-021-03819-2>

Kringelum, J. V., Lundegaard, C., Lund, O., & Nielsen, M. (2012). Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. *PLoS Computational Biology*, 8(12). <https://doi.org/10.1371/journal.pcbi.1002829>

Strokach, A., Becerra, D., Corbi-Verge, C., Perez-Riba, A., & Kim, P. M. (2020). Fast and Flexible Protein Design Using Deep Graph Neural Networks. *Cell Systems*, 11(4), 402–411.e4. <https://doi.org/10.1016/j.cels.2020.08.016>