GNN for the Prediction of Neutralizing Paratopes of SARS-CoV2 Antibodies

Shamima Rashid[a], Ng Yue Hao Shaun[a,b], Ng Teng Ann[a], Kwoh Chee Keong[a]

Graph Convolutional Network with Self-Attention Pooling for the Prediction of Neutralizing Paratope Sequences of SARS-CoV2 Antibodies

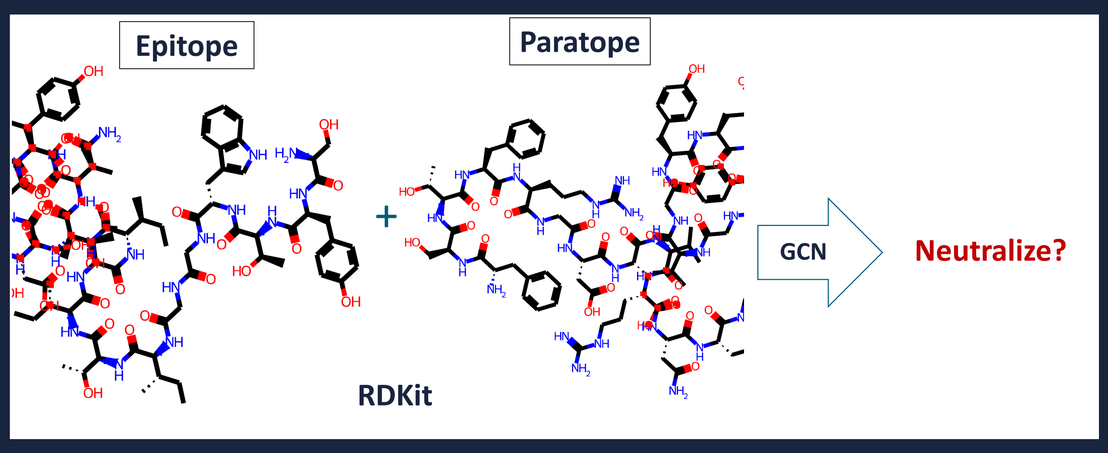
Background & Significance

- COVID-19 pandemic caused by SARS-CoV2 pathogen has caused millions of deaths & greatly disrupted society & economy
- Although the severity of the disease outbreak has been overcome therapeutics to treat COVID-19 still remain necessary. Many in the population continue to get re-infected with circulating variants
- Beneficial to have a repertoire of suitable antibody or paratope sequences, rapidly designed to meet therapeutic needs, based on emergent strains.

Introduction

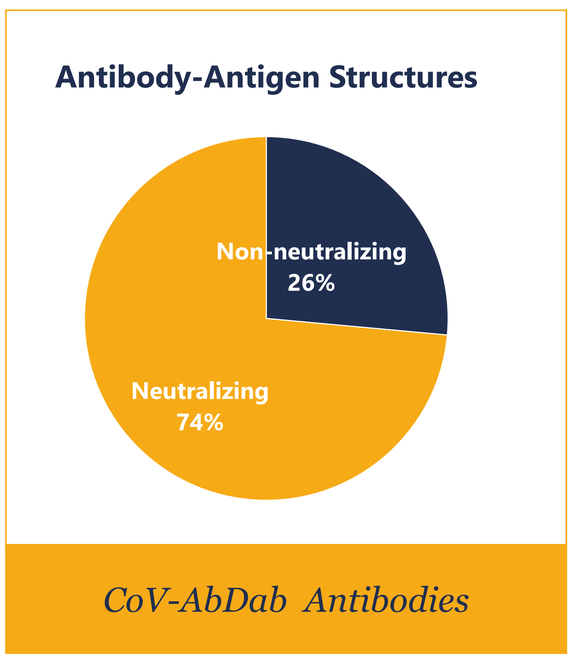
- In-silico models provided by deep graph networks are an avenue for high-throughput discoveries of neutralizing antibody sequences
- GNN have been used as models of adaptive graph relations for antibody prediction [1] & in drug-target interactions [2]
- Here, a deep graph neural network employing graph convolution with self-attention pooling was trained to detect pairs of neutralizing paratopes and epitopes from sequence data alone

Graph Network with **Self-Attention Pooling (SAG)** to predict neutralizing **paratopes** for **SARS-CoV2** Spike **epitopes**



Methods

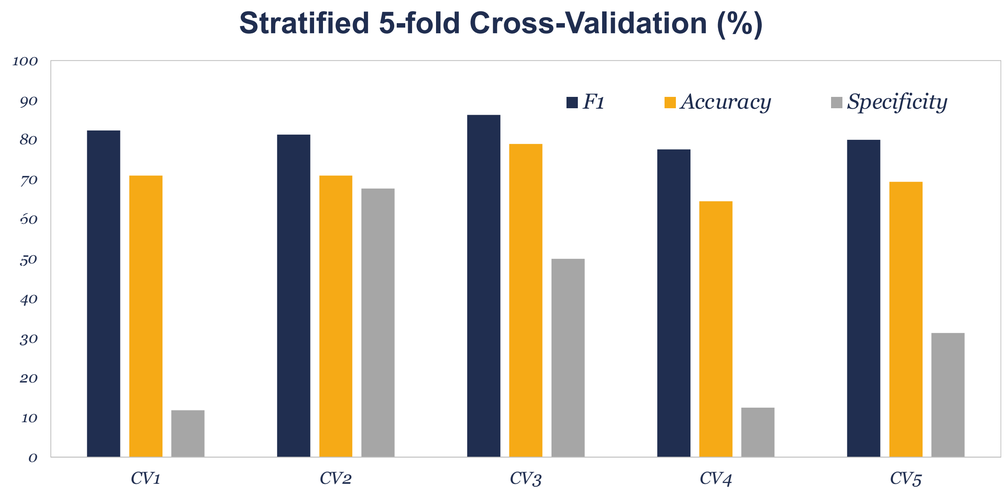
- Dataset contains 310 pairs of SARS-CoV2 epitope-paratope pairs, extracted from the Coronavirus Antibodies Database[3] (CoV-AbDab)
- Graph-based input features generated with RDKit[4] library
- Network architecture has 2 stacked [GCN+SAG] layers & 2 linear output layers (+softmax)
- Each GCN layer has 128 hidden units
- Predictions (Neutralizing=1 Non-neutralizing=0)



Results

- Performance was evaluated with **two** data-splitting strategies
- **Train-Test Split**: 2/3 train, 1/3 test

Sensitivity = 91.4%
F1 score = 81.5%
Accuracy = 71.0%
ROC-AUC = 57.4%
Specificity = 23.3%



Conclusions

- Good performance for neutralizing paratopes (True Positives) was achieved
- Further work is needed to detect non-neutralizing paratopes (True Negatives)
- Graph convolutional network with SAG pooling can capture paired representations using only the corresponding epitope and paratope sub-sequences
- Knowledge of the VH or VL regions (sequences or structures) may not be required to make the prediction
- More epitope-paratope pairs & advanced graph featurization and representation approaches can be used to rapidly detect neutralizing epitope-paratope pairs

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

References

1. "Predicting unseen antibodies' neutralizability via adaptive graph neural networks," doi: 10.1038/s42256-022-00553-w

2. "Drug-Target Interaction Prediction with Graph Attention networks." http://arxiv.org/abs/2107.06099

3. CoV-AdDab by Oxford Protein Informatics Group http://opig.stats.ox.ac.uk/webapps/covabdab/

4. RDKit: Open-source cheminformatics. https://www.rdkit.org

Dr. Shamima Rashid
School of Computer Science & Engineering [a]
sham0012@e.ntu.edu.sg
School of Chemical & Biomolecular Engineering [b]

Scan the QR code to download the **abstract**

SCAN ME