Fighting Future Pandemics:

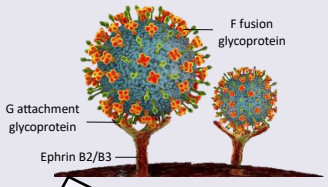
A Novel Approach to In Silico Drug Discovery for Global Health Priority Pathogen Nipah Henipavirus (NiV)

Lim Wan Qin

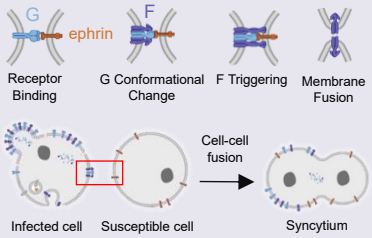
INTRODUCTION

1) A Deadly Paramyxovirus

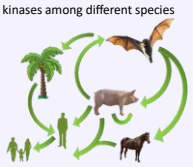
World Health Organization — Blueprint Priority Disease

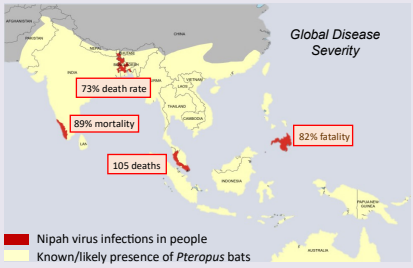


Infection Mechanism



Highly conserved receptor tyrosine kinases among different species





- Fatal encephalitis
- Acute respiratory illness
- Ataxia

2) The Long, Costly Drug Discovery Process

For a single drug: 12-15 years | >USD 2.5 billion | 90% failure rate

3) Previous Approach to Virtual Screening for NiV: Molecular Dynamics

- Require tedious preparation processes of chemical environment and proteins for accurate simulations

4) Current Drug-Target-Affinity (DTI) Models

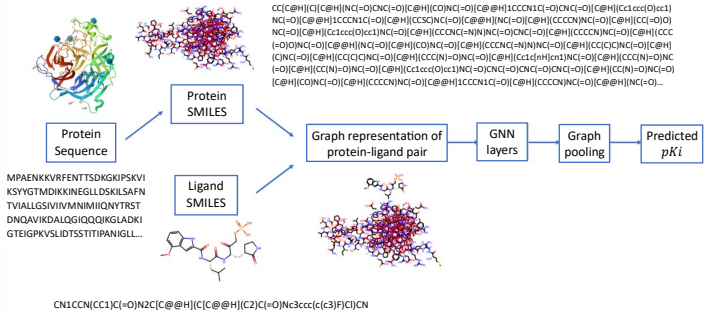
- Encode protein residues
  - Neglect atomic information

METHODOLOGY

✓ **Curate and train a novel deep learning Graph Neural Network (GNN) model that:**

- Incorporates protein atomic information
- Predicts inhibitory potential of a drug within seconds
  - Much faster than molecular simulations

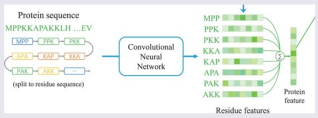
DeepGraphDTI



Trained on 410,286 binding measurements (BindingDB)

To predict on 1,040 candidate drugs (MMV Pandemic Response Box, Global Health Priority Box & Pathogen Box)

VS existing models:



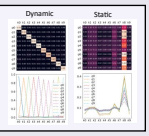
Graph Features

Node features	Atomic number, degree, formal charge, hybridization, aromaticity, radical electrons, number of Hs, in ring, chirality
Edge features	Bond type (Aromatic, DativeL, DativeR, Dative, DativeOne, Hydrogen, Ionic, ThreeCenter, 0, 1.0, 1.5, 2.0, 2.5 3.0, 3.5 4.0, 4.5, 5.0, 5.5, 6.0), in ring
Edge Index	Graph connectivity in coordinate form (adjacency matrix)
Label	$pKi = -\log\frac{Ki}{10^9}$ where Ki is the inhibition constant (concentration of drug required to produce half maximum inhibition)

Best Model Parameters

Parameter	Value/Type
Model Layers	4
GNN Type(s)	Layer 1: GATv2Conv Layers 2-4: TransformerConv
Attention heads	3
Optimiser	Stochastic Gradient Descent
Learning Rate	0.001
Activation Function	Scaled Exponential Linear Unit (SELU)
Graph Pooling Type(s)	Concatenation of Global Max Pooling and Global Mean Pooling
Loss Function	Mean Squared Error
Batch size	16
Dense Neurons	256

Computes dynamic attention: Unique ranking of attention coefficients for each node
Instead of
Monotonic static attention: Global shared ranking



Considers both extreme and average properties of the molecule

DeepGraphDTI Final Algorithm

Forward Propagation:

Input: node features x; edge features e; edge in dex ε; batch index β; trainable parameters $\mathbf{W}$; attention coefficient α; hidden size of each attention head d; global representation σ; linear transformation L; batch normalisation BN; number of layers l
Output: final representation x

1: **procedure** forward
2: **for** $l = 1$ **do**
3: $x_i \leftarrow a_{i,i} W_1 x_i + \sum_{j \in N(i)} a_{i,j} W_2 x_j$
where $\alpha i, j = \frac{exp\,(aTLeakyReLU(W_1 x_i + W_2 x_j + W_3 e_{i,j})}{\sum_{k \in N(i) \cup \{i\}} exp\,(aTLeakyReLU(W_1 x_i + W_2 x_k + W_3 e_{ik}))}$
4: **for** $l = 2$ to 4 **do**
5: $x_i \leftarrow W_{4,l}\, x_i + \sum_{j \in N(i)} a_{i,j} (W_{5,l} x_j + W_{8,l} e_{i,j})$
where $\alpha i, j = \text{softmax} \frac{(W_{6,l} x_i)^T (W_{7,l} x_j + W_{8,l} e_{i,j})}{\sqrt{d}}$
6: $x_i \leftarrow SELU\,(L_{1,l}(x_i))$
7: $x_i \leftarrow BN_l(x_i)$
8: $\sigma l \leftarrow$ Append $(\sigma_{i-1}$, CONCAT (GMP(xi, β), GAP(xi, β)))
9: **end for**
10: $x_i \leftarrow \sum_{k=1}^{4} \sigma_i$
11: $x_i \leftarrow SELU\,(L_2(x_i))$
12: $x_i \leftarrow SELU\,(L_3(x_i))$
13: $x_i \leftarrow L_4(x_i)$
14: **return** x_i

Performance Results

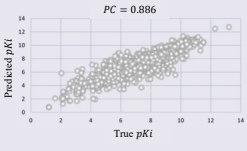
Metrics

t: ground-truth value; p: predicted value

Root-Mean-Squared-Error: $RMSE(t,p) = \sqrt{\frac{\sum_{i=1}^{n}(t_i - p_i)^2}{n}}$

Pearson Correlation: $PC(t,p) = \frac{n\Sigma tp - \Sigma t \Sigma p}{\sqrt{[n\Sigma t^2 - (\Sigma t)^2]\cdot[n\Sigma p^2 - (\Sigma p)^2]}}$

Method	RMSE↓	PC↑
DeepAffinity	0.840	0.840
DeepDTA	0.824	0.832
DeepCDA	0.808	0.844
BACPI	0.800	0.860
SSM-DTA	0.792	0.863
DeepGraphDTI	**0.742**	**0.886**



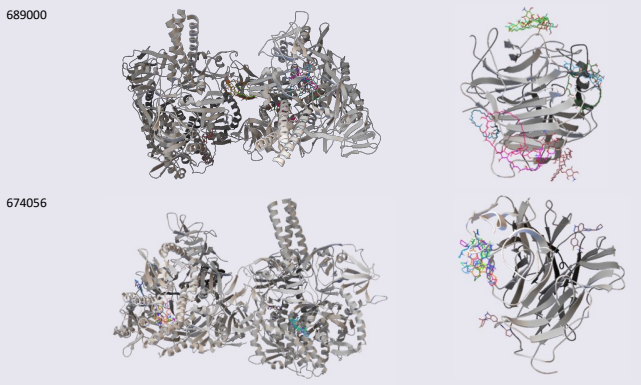
✓ **Molecular docking on 15 drug ligands giving highest pKi with NiV-F and G predicted by DeepGraphDTI ⇒ 7 potential inhibitors identified**

RESULTS

Binding Affinities of 7 Predicted Drugs with NiV-G and F Proteins

Ligand MMV ID	689000	674056	673982	1782412	687800	676461	1542798
NiV-G	-8.4	-8.7	-8.0	-7.4	-8.2	-7.9	-7.8
NiV-F	-10.6	-10.1	-9.7	-9.2	-9.1	-8.4	-8.2

Docked Poses of top 2 multi-target inhibitors



40 x 40 grid with highest weight values in NiV proteins' complexes with respective top 3 inhibitors

- Located at same selection of atoms for both NiV-F and G with all 3 ligands
- Highly similar weight values

CONCLUSION

- This study suggests the inhibitory potential of 7 drugs against NiV
- DeepGraphDTI, a novel DTA model which incorporates protein atomic information, outperformed competing DTI models, suggesting the value of incorporating protein atomic information in determining ligand-protein reaction outcomes, especially when focusing on localised interactions as in DTA tasks

FUTURE WORKS

- To confirm the pharmacological activity of the 7 potential inhibitors in vitro and vivo
- To powerfully harness DeepGraphDTI for future drug discovery tasks

References
[1] Tit-Oon, P., Tharakaraman, K., Artpradit, C., Godavarthi, A., Sungkeeree, P., Sasisekharan, V., Kerdwong, J., Miller, N. B., Mahajan, B., Khongmanee, A., Ruchirawat, M., Sasisekharan, R., & Fuangthong, M. (2020). Prediction of the binding interface between monoclonal antibody m102.4 and Nipah attachment glycoprotein using structure-guided alanine scanning and computational docking. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-75056-y [2] Looi, L.M., Chua, K.B. (2007). Lessons from the Nipah virus outbreak in Malaysia. The Malaysian Journal of Pathology. Department of Pathology, University of Malaya and National Public Health Laboratory of the Ministry of Health, Malaysia. 29 (2): 63–67. PMID 19108397 [3] Xiong, J., Xiong, Z., Chen, K., Jiang, H., & Zheng, M. (2021). Graph neural networks for automated de novo drug design. Drug Discovery Today, 26(6), 1382–1393. https://doi.org/10.1016/j.drudis.2021.02.011 [4] Bossart K.N., Tachedjian M., McEachern J.A., Crameri G., Zhu Z., Dimitrov D.S., Broder C.C., Wang L.F. Functional studies of host-specific ephrin-B ligands as Henipavirus receptors. Virology. 2008;372:357–371. doi: 10.1016/j.virol.2007.11.011 [5] Negrete O.A., Levroney E.L., Aguilar H.C., Bertolotti-Ciarlet A., Nazarian R., Tajyar S., Lee B. EphrinB2 is the entry receptor for Nipah virus, an emergent deadly paramyxovirus. Nature, 2005;436:401–405. doi: 10.1038/nature03838