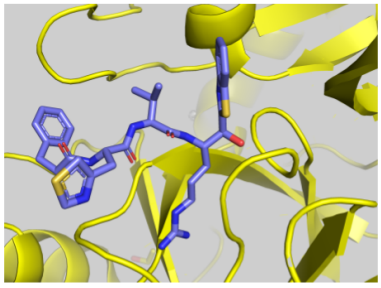
TMPRSS2 small molecule inhibitors as candidates for novel Covid-19 antivirals

Sybren van den Bedem
Wakatsuki Lab, Stanford University
August 2022
syb.vdbedem@gmail.com code: https://github.com/syb-vdbedem

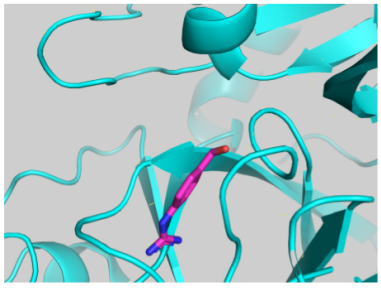
Introduction

Viruses reproduce by inserting their genetic material into the host cell and hijacking its protein synthesis pathways. Viruses mutate rapidly, reducing the effectiveness of antiviral treatments. However, viruses often rely on host proteins to enter the cell, which don't mutate as rapidly. COVID-19 is caused by the SARS-CoV-2 virus. SARS-CoV-2 uses Transmembrane Protein Serine S2 (TMPRSS2) to enter the cell. By blocking the function of TMPRSS2, we can effectively treat COVID-19. In this study, we perform a small virtual High Throughput Screen to identify small molecules that possibly inhibit TMPRSS2.

Background



TMPRSS6 in complex with benzothiazole (slate)

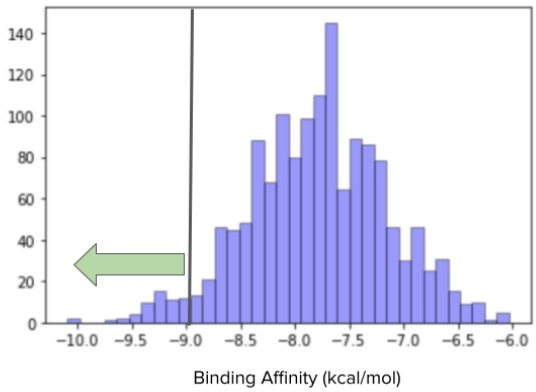


TMPRSS2 in complex with Nafamostat (magenta)

I found a recent study that identifies the peptidomimetic ligand N-0385 as a powerful inhibitor of TMPRSS2 [1]. However, N-0385 is not well-suited as a starting point for therapeutics, as peptidomimetics are often recognized as 'intruders' by the human body, and could be destroyed before reaching TMPRSS2. The study does not provide a binding mode, but recent crystal structures of TMPRSS2 in complex with Nafamostat [2] and the close homolog TMPRSS6 in complex with benzothiazole [3] yield information on the binding site. I hypothesized that there exist small molecule alternatives to N-0385 that could bind and potentially inhibit TMPRSS2.

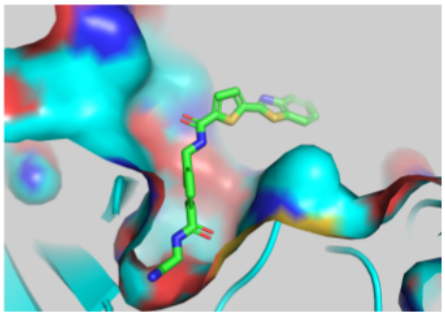
Results

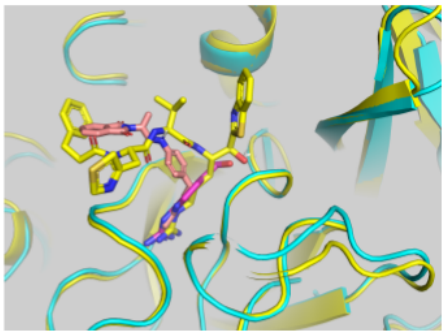
To test my hypothesis, I first searched the Pubchem ligand database for small molecules structurally similar to N-0385. Next, I docked the small molecules to TMPRSS2. I then used a Convolutional Neural Network to score how likely the ligands are to inhibit TMPRSS2. Finally, I verified the binding modes by superimposing the binding sites of N-0385, Nafamostat, and the ligands.



To search the Pubchem database, I downloaded the SMILES string representation of N-0385 and searched for molecules with Tanimoto similarity >=80% similarity, receiving 1,456 ligands. Using Smina [4], I docked the results at the Nafamostat binding site, which took ca. 10 hours on my Macbook Air. I wrote a Python script that automatically extracted the binding affinity from the Smina output files. No analogs exceeded a binding affinity of -10.09 kcal/mol (Figure). The scores approximately follow a Gaussian distribution; 36 compounds score in the top 2.5% of the data.

However, Smina considers only binding interactions, which poorly predict whether a molecule actually binds [5]. The deep learning software Gnina [6] learned additional features to predict whether a molecule binds the protein using a CNN (Convolutional Neural Network). The CNN was trained on a dataset of crystal structures of bound complexes to recognize features of 'good' protein-ligand binding. I ran Jupyter Notebooks on Google Colab to dock the Smina results with a predicted affinity better than -9 kcal/mol, corresponding to the top 3.5% of data, using Gnina and reranked them based on their Gnina scores. I superimposed the top-ranked Gnina ligand onto the TMPRSS2 and TMPRSS6 complexes to visually evaluate the binding mode.





Discussion and Conclusions

The top scoring ligand (top figure, green) reaches deep into the TMPRSS2 binding cavity (surface representation). I also observed that the top scoring ligand (bottom figure, salmon) also closely follows the binding modes of both Nafamostat (magenta) and the peptidomimetic (yellow). Thus, the top scoring Gnina ligand is a promising candidate for experimental validation.

References

[1] Shapira, T., Monreal, I.A., Dion, S.P. et al. A TMPRSS2 inhibitor acts as a pan-SARS-CoV-2 prophylactic and therapeutic. *Nature* 605, 340–348 (2022)

[2] Fraser, B.J., Beldar, S., Seitova, A. et al. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat Chem Biol* 18, 963–971 (2022)

[3] Béliveau, François, et al. Discovery and development of TMPRSS6 inhibitors modulating hepcidin levels in human hepatocytes. *Cell Chemical Biology* 26, 1559-1572 (2019)

[4] Koes, D.R. Baumgartner, M.P., and Camacho, C.J., Lessons Learned in Empirical Scoring with Smina from the CSAR 2011 Benchmarking Exercise *J. Chem. Inf. Model*. 53 1893–1904(2013)

[5] Wójcikowski, M., Ballester, P. & Siedlecki, P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci Rep* 7, 46710 (2017)

[6] McNutt, A.T., Francoeur, P., Aggarwal, R. et al. GNINA 1.0: molecular docking with deep learning. *J Cheminform* 13, 43 (2021).