TMPRSS2 small molecule inhibitors

Sybren van den Bedem
Wakatsuki Lab, Stanford University
August 2022

syb.vdbedem@gmail.com
code: https://github.com/syb-vdbedem

Motivation and Background

COVID-19 is caused by a virus called SARS-CoV-2

Viruses can mutate rapidly, reducing the effectiveness of treatments

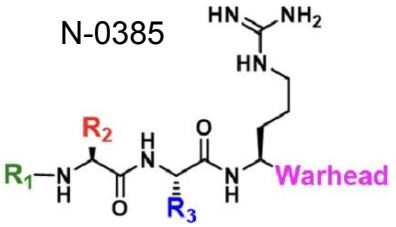
Viruses use proteins of the host

- Viruses often need host proteins to enter the cell
- Host proteins are less likely to mutate rapidly
- Transmembrane Protein Serine S2 (TMPRSS2) is a host protein that helps the SARS-CoV-2 virus enter the cell*

Can we block the function of TMPRSS2 and prevent SARS-CoV-2 from infecting host cells?

*Fraser *et al.* Nat Chem Biol **18**, 963-971 (2022)

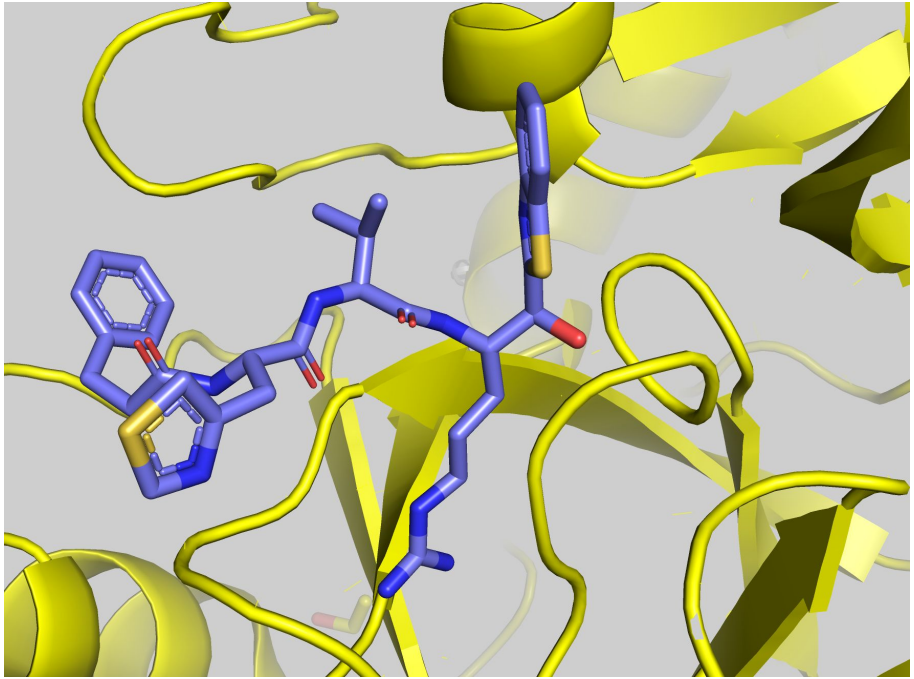
Background



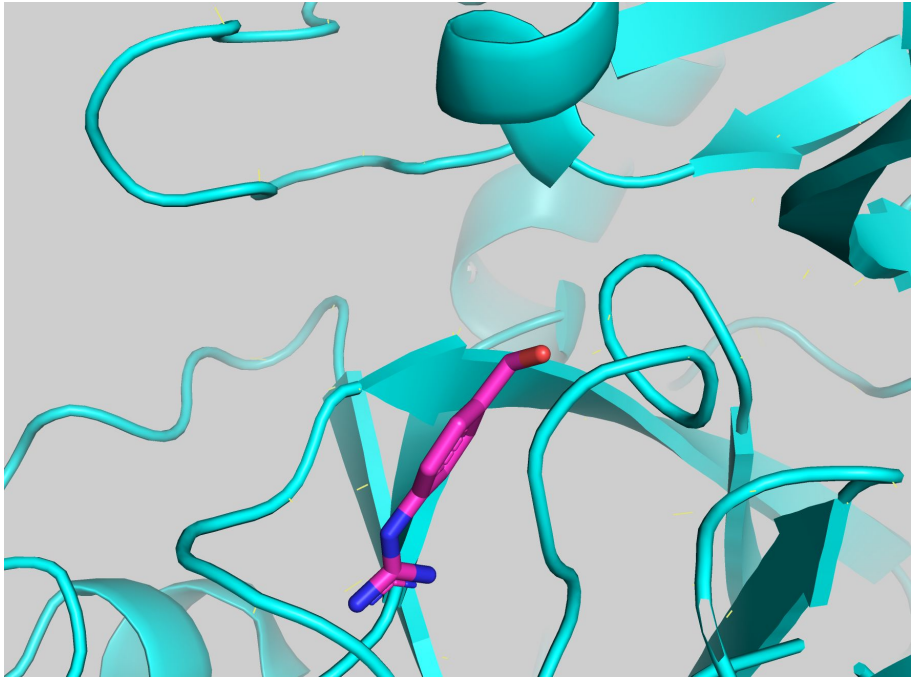
1. The study* ‘A TMPRSS2 inhibitor acts as a pan-SARS-CoV-2 prophylactic and therapeutic’ identifies the peptidomimetic ligand N-0385 as a powerful inhibitor of TMPRSS2. However, the authors do not have a crystal structure, and do not provide a molecular structure of the peptidomimetic ligand N-0385
2. Reference 32** in the paper points to a crystal structure of TMPRSS6 in complex with benzothiazole, a peptidomimetic similar to N-0385. (PDB ID 6N4T)
3. The study*** ‘Structure and Activity of Human TMPRSS2 Protease Implicated in SARS-CoV-2 Activation’ explains how TMPRSS2 inhibition prevents SARS-CoV-2 from multiplying, **and** provides a crystal structure (PDB ID 7MEQ), but does not have N-0385 in the study. Instead, it has a small peptide Nafamostat bound to TMPRSS2.

*Shapira, T. *et al.* Nature **605**, 340-348 (2022).
Beliveau et al, Cell Chem Biol **26, 1559-1572 (2019)
***Fraser *et al.* Nat Chem Biol **18**, 963-971 (2022)

Crystal structures



TMPRSS6 in complex with benzothiazole (slate)



TMPRSS2 in complex with Nafamostat (magenta)

Why not peptidomimetic N-0385?

Peptidomimetics are molecules that look like small proteins

Peptidomimetics are often recognized as ‘intruders’ by the human body, and could be destroyed before reaching TMPRSS2

A small molecule alternative could be optimized towards a drug

Hypothesis:

There exist small molecule alternatives for N-0385 that bind and potentially inhibit TMPRSS2

Approach

Computationally combine the findings of these previous studies to find a small molecule alternative to N-0385

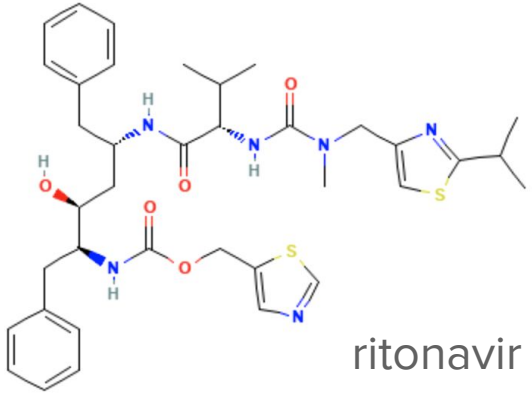
- Search ligand databases for small molecules similar to peptidomimetic N-0385
- Dock the small molecules to TMPRSS2 (PDB 7MEQ)
- Score how likely they are to inhibit TMPRSS2 with a CNN
- Verify the binding modes by superimposing PDBs 7MEQ and 6N4T, i.e., the verified inhibitor

Results

PubChem results

PubChem is a database of chemical matter and activities against biological assays

1. I downloaded the SMILES string representation of the peptidomimetic from the crystal structure with PDB ID 6N4T
2. I searched for similar compounds (‘analogs’) in the PubChem database using the SMILES string. I limited the search to results with >=80% similarity to the query, and received 1,456 results
3. The top hit was ritonavir, a protease inhibitor used to treat HIV/AIDS and COVID-19 (and an active ingredient of Paxlovid!)



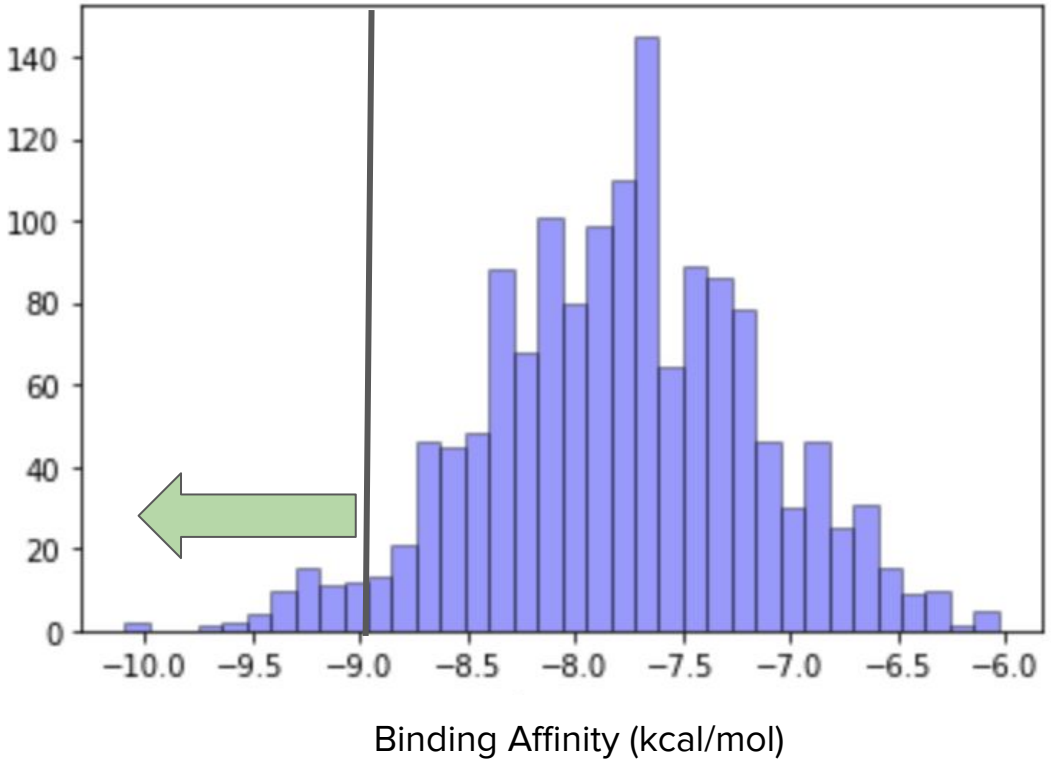
ritonavir

Dock the N-0385 analogs to TMPRSS2

1. I used the program Smina — an open source version of Autodock Vina — to dock the 1,456 analogs to TMPRSS2
2. I used the command “autodock” to dock the analogs to where the peptide Nafamostat would normally bind in the active site for TMPRSS2 (pdb 7MEQ)
3. Docking 1,456 compounds took ~10 hours on my Macbook Air
4. I wrote a python script in a Jupyter Notebook that automatically searched the Smina output for the small molecules with the greatest affinity

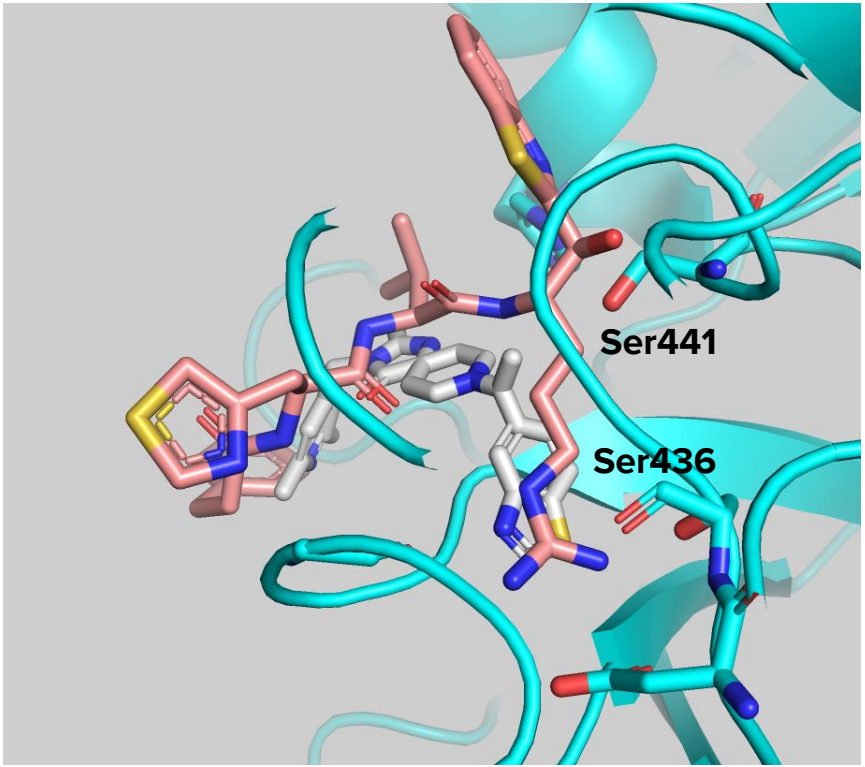
Distribution of smina scores

1. Smina scores have unit kcal/mol
2. No analogs exceeded a binding affinity of -10.09 kcal/mol.
3. kcal/mol = total energy to form complex
4. 53 peptides scored better than -9 kcal/mol.



Visualizing binding modes of N-0385 analogs

The most potent analog (grey) formed strong bonds with TMPRSS2 (cyan). It binds similarly as the arginine substructure in benzothiazole* (brown), forming interactions with Ser436.



*Mahoney *et al* Proc Nat Acad Sci **118** e2108728118 (2021)

Re-ranking bound ligand poses with deep learning

- Smina only considers binding interactions, which poorly predicts whether a molecule actually binds.
- The deep learning software Gnina* learned additional features to predict whether a molecule binds the protein
- Gnina re-ranks smina ligand poses with 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 used a Jupyter Notebook to dock the top 53 Smina results using Gnina on Google Colab

*McNutt, A.T. *et al. J Cheminform,* **13** (2021)

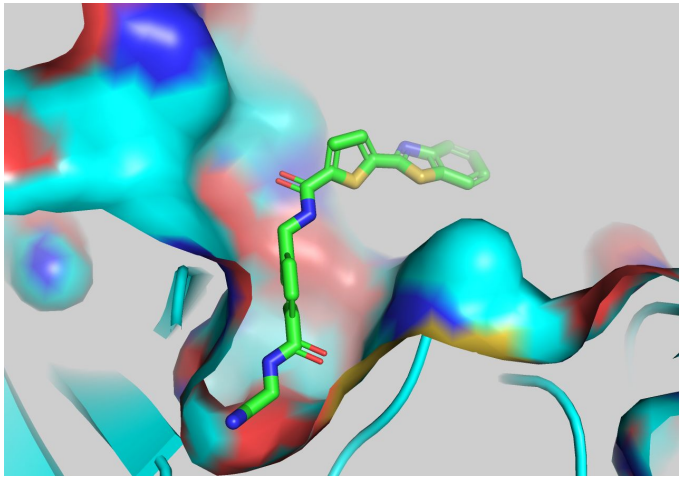
Top Gnina ligand binding modes

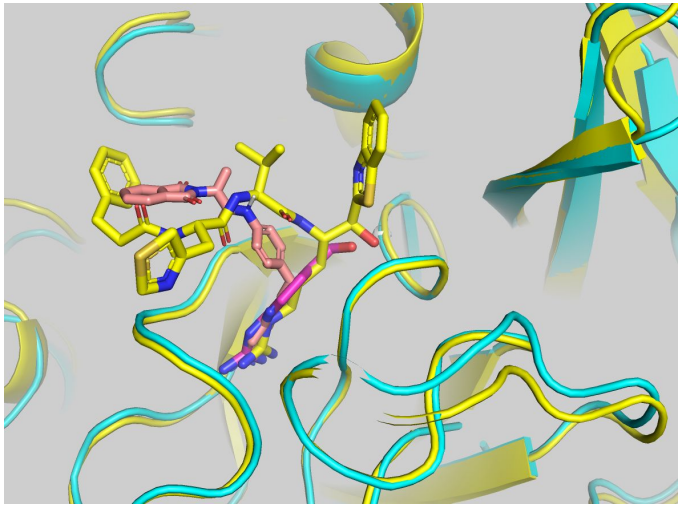
I ranked the ligands by ‘CNNaffinity’ score

Top figure: The top scoring ligand (green) reaches deep into the TMPRSS2 binding cavity

Bottom figure: The top scoring ligand (salmon) follows the binding modes of both Nafamostat (magenta) and the peptidomimetic (yellow)

CID	CNNaffinity
25480493	5.824950
45354655	5.824950
25480483	5.824950
25480488	5.824950
25480496	5.824950
50995510	5.781636
50995511	5.779104
30735249	5.730596
30735262	5.730596
30735258	5.730596
16299242	5.730596
30735254	5.730596
38284587	5.634814
52935051	5.624274
52934819	5.623214
9533674	5.308640
9533672	5.308640





Summary and Conclusions

Summary and Conclusions

- Smina and Gnina are efficient tools to screen compounds for bioactivity
- Deep Learning screening tools can consider more features than physical interactions to determine the most promising binders
- I computationally verified my hypothesis: Among only 1,456 PubChem hits there are ligands with binding modes similar to known peptidomimetics and (covalent) benzothiazole
- The top compounds I found need further optimization. Next steps:
    - Find the most common ligand-protein interactions among the top compounds
    - Search larger databases (billions) for similar compounds (ZINC database)
    - Experimentally validate the predicted ligands