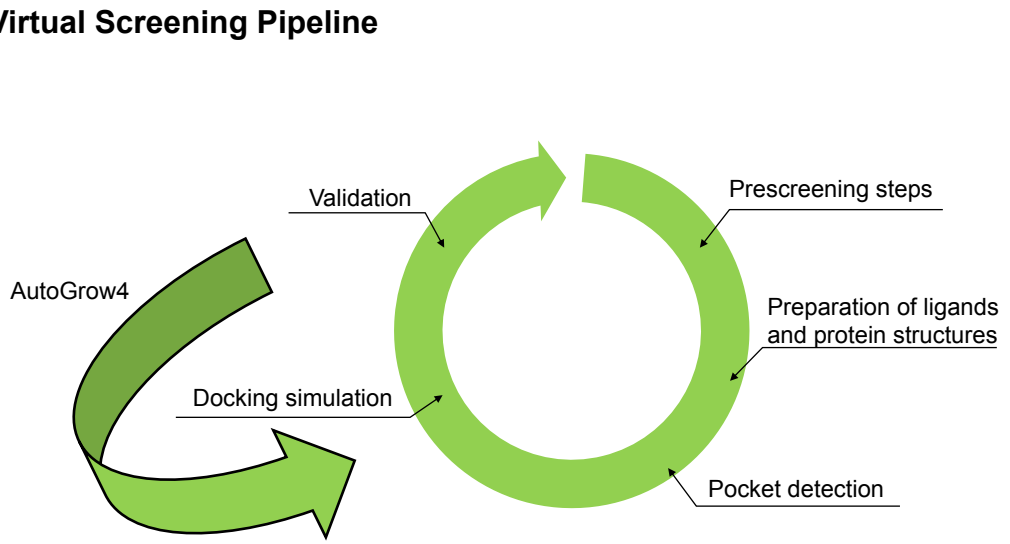
MeetEU Project - Team Heidelberg - Team 1 – Identification of novel Sars-CoV-2 NSP13 helicase inhibitors

Anastasiya Vladimirova, Darius Szablowski, Duc Thien Bui, Wangjun Hu

February 2023

1 Workflow



[2]

Figure 1: **Virtual Screening Pipeline.** It consists of the following steps: Prescreening steps, e.g. Toxicology Check, Preparation of ligands and protein structures with *PyMOL* by removing water molecules and converting to the correct file format, Pocket detection via *Fpocket*, Docking simulation with *AutoDock Vina* using the *Vinardo* scoring function, using the results from that to assimilate new ligands using *AutoGrow 4* to rerun the Docking simulation and lastly Validation using Molecular Dynamics Simulations with *GROMACS*.

2 Material and Methods

2.1 Structure and Binding Pocket Detection

The 6ZSL[1] structure from the RCSB PDB database was chosen for the detection of binding pockets. Fpocket [Guilloux et al., 2009] and Prankweb[2] [Jendele et al., 2019] were used to predict the binding pockets.

2.2 Test Dataset from ZINC15

A drug-like subset of 95,438,992 molecules was downloaded from the ZINC15-Database [Sterling and Irwin, 2015] for initial consideration.

2.3 Toxicology prediction using *eToxPred*

The machine learning based tool *eToxPred* [Pu et al., 2019] was used for toxicology prediction of the given molecules. Based on the information given by the paper the threshold was set to 0.58 for a substance to be considered toxic.

[1] `http://dx.doi.org/10.2210/pdb6zsl/pdb`
[2] `https://prankweb.cz/`

2.4 Docking

For docking 5000 molecules out of the molecules that were predicted to be non-toxic were randomly chosen for docking. Docking was performed using *AutoDock Vina* [Forli et al., 2016] using the *Vinardo* scoring function [Quiroga and Villarreal, 2016].

2.5 PDB Preparation

The crystallized protein structures of NSP13 were downloaded as previously described at RCSB [3]. Missing residues and atoms were added using the Python package *'pdbfixer'* [Eastman et al., 2013].

2.6 Molecular Dynamics Simulation using *GROMACS*

GROMACS was used to validate the docking results by evaluating the ability of the top ligands to stay in the binding pocket [Bekker et al., 1993].

2.7 Visualization and Analysis of MD simulation results

Using the program *VMD* it is possible to look at the trajectory of the MD simulations of the different molecules [Humphrey et al., 1996].

2.8 Second Docking Step

Another Docking round again using *AutoDock Vina* with the *Vinardo* scoring function was performed. As per the task of this project a subset of the European Chemical Biology Database (ECBD)[4], specifcally the subset: `https://ecbd.eu/compound/#lib{value='4'}`.

2.9 Generating New Molecules

AutoGrow 4 can generate new ligands on the basis of ligands and a specified binding pocket which are likely to have a similar or even higher affinity than the ligands used as input [Spiegel and Durrant, 2020].

3 Results

3.1 General Docking Results ZINC15 dataset subset

The resulting binding energy of our 5000 starting molecules were loaded visualized in a histogram using R (see Figure **??**). Most of the molecules have binding affinities between – 2.5 and -10 $[\frac{kJ}{mol}]$. A small fraction of molecules has binding affinities below – 10 $[\frac{kJ}{mol}]$.

3.2 Reranking our top 10 molecules based on the total protein-ligand interaction based on MD simulation results

Based on the MD simulations the ranking of the ligands amongst each other changed slightly even though the overall trend of the best ligands staying the best can be observed for most of them.

3.3 Visualizing pocket stability by calculating the RMSD between protein and ligand

While the best ligands mostly stayed the best even after reevaluation it could be observed that only the RMSD values for the top 3 ligands were particularly low.

[3] `https://www.rcsb.org`
[4] `https://ecbd.eu`

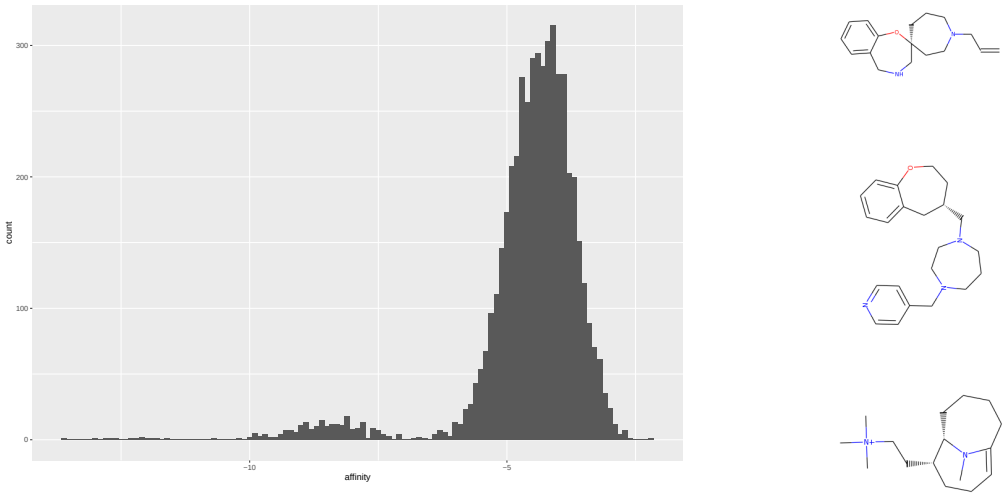


Figure 2: On the left: **A histogram of the frequency-counts of the determined binding affinities in** $[\frac{kJ}{mol}]$**.**
On the right: **Our top three candidates(ZINC000019015192, ZINC000095523345 and ZINC000101042701) in descending order.**

References

[Bekker et al., 1993] Bekker, H., Berendsen, H. J. C., Dijkstra, E., Achterop, S., Vondrumen, R., Vanderspoel, D., Sijbers, A., Keegstra, H., and Renardus, M. (1993). Gromacs - a parallel computer for molecular-dynamics simulations.

[Eastman et al., 2013] Eastman, P. K., Friedrichs, M. S., Chodera, J. D., Radmer, R. J., Bruns, C. M., Ku, J. P., Beauchamp, K. A., Lane, T. J., Wang, L., Shukla, D., Tye, T., Houston, M., Stich, T. J., Klein, C., Shirts, M. R., and Pande, V. S. (2013). Openmm 4: A reusable, extensible, hardware independent library for high performance molecular simulation. *Journal of chemical theory and computation*, 9 1:461–469.

[Forli et al., 2016] Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, 11(5):905–919.

[Guilloux et al., 2009] Guilloux, V. L., Schmidtke, P., and Tuffery, P. (2009). Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics*, 10(1).

[Humphrey et al., 1996] Humphrey, W. F., Dalke, A., and Schulten, K. (1996). Vmd: visual molecular dynamics. *Journal of molecular graphics*, 14 1:33–8, 27–8.

[Jendele et al., 2019] Jendele, L., Krivak, R., Skoda, P., Novotny, M., and Hoksza, D. (2019). PrankWeb: a web server for ligand binding site prediction and visualization. *Nucleic Acids Research*, 47(W1):W345–W349.

[Pu et al., 2019] Pu, L., Naderi, M., Liu, T., Wu, H.-C., Mukhopadhyay, S., and Brylinski, M. (2019). eToxPred: a machine learning-based approach to estimate the toxicity of drug candidates. *BMC Pharmacology and Toxicology*, 20(1).

[Quiroga and Villarreal, 2016] Quiroga, R. and Villarreal, M. A. (2016). Vinardo: A scoring function based on autodock vina improves scoring, docking, and virtual screening. *PLOS ONE*, 11(5):e0155183.

[Spiegel and Durrant, 2020] Spiegel, J. O. and Durrant, J. D. (2020). Autogrow4: an open-source genetic algorithm for de novo drug design and lead optimization. *Journal of cheminformatics*, 12(1):1–16.

[Sterling and Irwin, 2015] Sterling, T. and Irwin, J. J. (2015). ZINC 15 – ligand discovery for everyone. *Journal of Chemical Information and Modeling*, 55(11):2324–2337.