# Screening for Sars-Cov-2 helicase inhibitors

Anna Audit, Hortense Beaussant, Célia Messaoudi

January 31, 2024

# 1 Introduction

One approach to prevent the spread of COVID-19 is to stop the virus from replicating. In particular, inhibiting the Sars-Cov-2 helicase : Nsp13. Indeed, helicases are necessary for the virus to replicate. We built a pipeline to select potential inhibitors, generate multiple conformations for those candidates, docked them and ranked them. Finally, we performed molecular dynamic simulation (MD simulation) on our best performing candidates in order to refine our ranking. We used the European Chemical Biology Database (ECBD ).

# 2 Materials and Methods

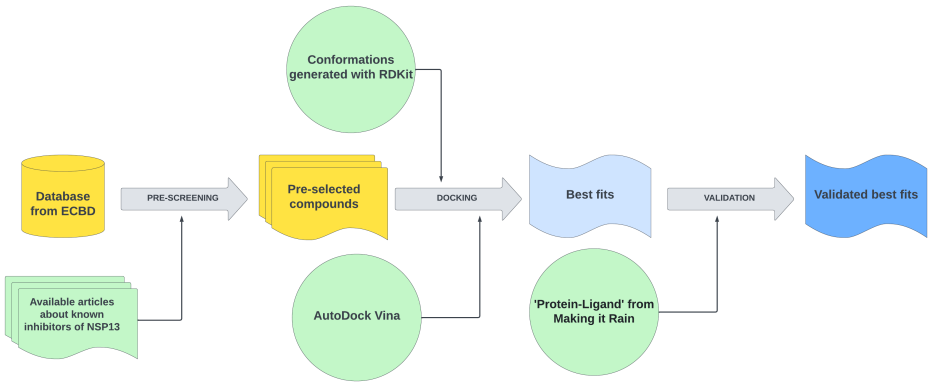


Figure 1: **Inhibitor screening pipeline**. It consists of four main steps. The pre-filtering of our database based on information found in the literature. Then, the generation of multiple conformations for our selected compounds with RDKit, the docking of those multiple conformations with the helicase using Autodock Vina and the subsequent ranking of the conformer based on their affinity score. Finally, MD simulation was performed for the top performing inhibitors.

## 2.1 Database of compounds

The database used for this project is the pilot library from ECBD. It contains 5015 compounds. We did a pre-selection and only kept molecules that met the following criteria [2]:

- availability in ZINC database,
- logP $< 4$ and
- $430 \leq$ molecular weight $\leq 470$ kDa

## 2.2 Generating 3D conformations from SMILES

We used RDKit to generate 3D conformations for out candidates from SMILES formula. We generated 3 conformers for each candidate. With our 276 candidates, we got 828 conformations in total.

## 2.3 Molecular docking

We used AutoDock Vina as our docking software [6]. The default Vina affinity score is used here.

Before docking, we prepared the ligands using MGLTools from AutoDock. Docking of all 828 conformers was done and the 10 best scoring candidates were kept.

## 2.4 Molecular Dynamics simulation

Molecular Dynamics (MD) simulations were run for the top 5 candidates found after docking to get a dynamic representation of the system over time. This step allowed us to refine our previous ranking.

We used Making-it-Rain 'Protein-Ligand' notebook [3]. We simulated 10 ns of interaction for each compound. The analysis functions provided comprised RMSD , MM-GBSA and MM-PBSA scoring [5] among others.

# 3 Results

## 3.1 Molecular docking

After the computation, two docking sites appear. Since the 7nio conformation of Nsp13 is a homodimer, the two sites may be symmetrical i.e. they are identical but located on one half of the dimer each.

## 3.2 Best scoring compounds after docking

Our best scoring ligand is ZINC000049006633. Table 1 was obtained by ranking and keeping only the best scoring conformation for each ligand and taking top 10. The scores are very close so we can't conclude on some candidates being way better than others.

| Formula | ZINC ID | Affinity score (kcal/mole) | Place |
|---|---|---|---|
| C24H17N7O3 confo 0 | ZINC000049006633 | -11.1 | 1 |
| C21H22FN3O5S confo 2 | ZINC000064889585 | -10.9 | 2 |
| C21H22ClN3O5S confo 2 | ZINC000064889599 | -10.9 | 2 |
| C24H33FO6 confo 1 | ZINC000064889599 | -10.8 | 4 |
| C24H31FO6 confo 0 | ZINC000004097305 | -10.8 | 4 |
| C24H32ClFO5 confo 0 | ZINC000004213474 | -10.7 | 6 |
| C25H25N5O4 confo 0 | ZINC000011677837 | -10.7 | 6 |
| C25H30N6O3 confo 1 | ZINC000059258964 | -10.7 | 6 |
| C27H27N5O3 confo 2 | ZINC000033173150 | -10.6 | 9 |
| C22H20F5N3O3 confo 2 | ZINC000043208642 | -10.6 | 9 |

Table 1: Top 10 candidates after docking

## 3.3 Best scoring compounds after MD simulations

After MD simulation ZINC000043208642 is the better option (top 5 to top 1) whereas ZINC000049006633 - top 1 before - is in the lower part of the ranking. Our ranking are based on the MM-PBSA values as it is more accurate that MM-GBSA

| Formula | ZINC ID | MM-GBSA (kcal/mole) | MM-PBSA (kcal/mole) | RMSD (Å) | Place |
|---|---|---|---|---|---|
| C24H17N7O3 confo 0 | ZINC000049006633 | -6.9806 | 4.6739 | 2.78 ±0.69 | 3 |
| C21H22FN3O5S confo 2 | ZINC000064889585 | -3.1628 | 4.2446 | 3.53 ±0.64 | 2 |
| C21H22ClN3O5S confo 2 | ZINC000064889599 | -8.2222 | 4.7031 | 2.99 ±0.62 | 4 |
| C24H33FO6 confo 1 | ZINC000064889599 | -8.7427 | 5.0721 | 2.64 ±0.67 | 5 |
| C24H31FO6 confo 0 | ZINC000043208642 | -9.7767 | -1.8422 | 3.94 ±0.85 | 1 |

Table 2: Top 5 candidates after MD simulations

## 3.4 Comparison with other teams

Our top ligands are different from the ones found by other teams, even though the Heidelberg team used a mix of ECBD and ZINC databases too. This can be due to the parameters used in the docking step. Also, our affinity scores are close but a little higher than theirs. Same observation for affinity scores from the Warsaw 1 team, they have higher values. The use of various databases and docking softwares (some teams used DiffDock [4]) may explain the differences in our results. Some of us also used different conformations of Nsp13 such as 7nn0 and 7nio.

# 4 Discussion

It would be interesting to use DiffDock [4] instead of AutoDock Vina for docking and compare results.

Making-it-Rain uses the cloud resources of Google Colab this allows us to run a whole MD simulation, in around 24 hours when it takes GROMACS a day and a half. However it can not be run without Colab pro. Therefore, using other tools like GROMACS [1], NAMD, AMBER and OpenMM can be a good alternative. Deep learning methods such as DeePMD-kit [7] can also be used to study biomolecular interactions.

Another area to improve would be the parameters used for docking and MD simulations, for instance : changing the solvation parameters for equilibrium before MD simulation or simulating more than 10 ns.

# References


[1] Mark James Abraham et al. "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". In: *SoftwareX* 1-2 (Sept. 2015), pp. 19–25. ISSN: 2352-7110. DOI: 10.1016/j.softx.2015.06.001. URL: https://www.sciencedirect.com/science/article/pii/S2352711015000059 (visited on 01/31/2024).

[2] Yazdi Ak et al. "Kinetic Characterization of SARS-CoV-2 nsp13 ATPase Activity and Discovery of Small-Molecule Inhibitors". en. In: *ACS infectious diseases* 8.8 (Dec. 2022). ISSN: 2373-8227. DOI: 10.1021/acsinfecdis.2c00165. URL: https://pubmed.ncbi.nlm.nih.gov/35822715/ (visited on 11/21/2023).

[3] Pablo R. Arantes et al. "Making it Rain: Cloud-Based Molecular Simulations for Everyone". In: *Journal of Chemical Information and Modeling* 61.10 (Oct. 2021), pp. 4852–4856. ISSN: 1549-9596. DOI: 10.1021/acs.jcim.1c00998. URL: https://doi.org/10.1021/acs.jcim.1c00998 (visited on 01/01/2024).

[4] Gabriele Corso et al. "DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking". en. In: Sept. 2022. URL: https://openreview.net/forum?id=kKF8_K-mBbS (visited on 01/31/2024).

[5] Samuel Genheden and Ulf Ryde. "The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities". In: *Expert Opinion on Drug Discovery* 10.5 (May 2015), pp. 449–461. ISSN: 1746-0441. DOI: 10.1517/17460441.2015.1032936. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4487606/ (visited on 01/31/2024).

[6] Oleg Trott and Arthur J. Olson. "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading". In: *Journal of computational chemistry* 31.2 (Jan. 2010), pp. 455–461. ISSN: 0192-8651. DOI: 10.1002/jcc.21334. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3041641/ (visited on 01/31/2024).

[7] Han Wang et al. "DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics". In: *Computer Physics Communications* 228 (July 2018), pp. 178–184. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2018.03.016. URL: https://www.sciencedirect.com/science/article/pii/S0010465518300882 (visited on 01/30/2024).