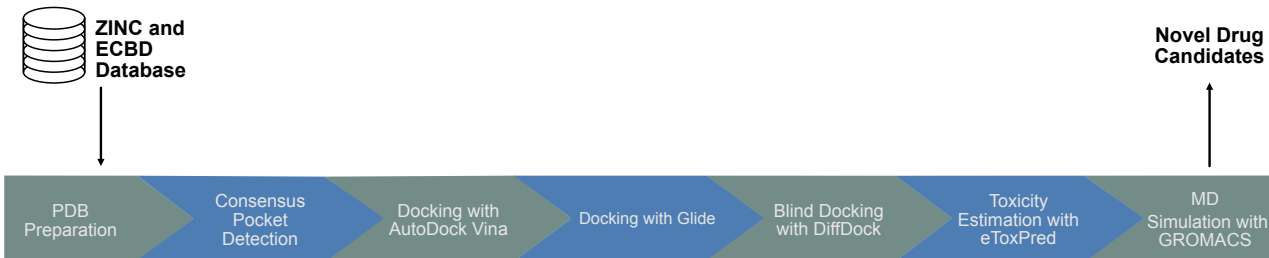MeetEU Project - Team Heidelberg - Team 1
In silico Screening for NSP13 Helicase Inhibitors as Potential Drug
Candidates in anti-SARS-CoV-2 Therapy

Linda Blaier, Paul Brunner, Selina Ernst, Valerie Segatz, and Chloé Weiler

February 2024

1   Workflow

Workflow for the discovery of NSP13 helicase inhibitors.

2   Material and Methods

1616 FDA-approved drugs from the ZINC database (Irwin et al. 2020) and 5016 files from the ECBD pilot library were downloaded. Three protein structures of NSP13 (PDB codes: 6ZSL, 5RME, and 5RM2) were aquired from RCSB PDB.

For identification of the consensus pocket, the protein structures were prepared using *PDBFixer* (Eastman et al. 2017). *Fpocket* (Le Guilloux et al. 2009), *PrankWeb 3* (Krivák and Hoksza 2018; Jendele et al. 2019; Jakubec et al. 2022), and *FTMap* (Brenke et al. 2009) were run on the structures to determine the consensus binding pocket. For further experiments, the docking box was set to be a cube with an edge length of 30 Å. Molecular docking was performed on the A chain of 6ZSL whose catalytic activity had been confirmed in previous publications (Berta et al. 2021). Molecular docking of the ligands was performed using *AutoDock Vina* (Trott and Olson 2010) and *Glide* (Friesner et al. 2004). The docking results were then validated through blind docking with *DiffDock* (Corso et al. 2022). The toxicity and synthetic accessibility of the Top 100 scorers from *AutoDock Vina* were predicted using *eToxPred* (Pu et al. 2019). In order to validate the binding of the top scorer, a molecular dynamics simulation was performed using *Gromacs* (Abraham et al. 2015). Lastly, for a deeper understanding of the potential inhibitors of NSP13, our 10 top scoring compounds from *Glide* were clustered against the top 10 best-docking ligands of the teams "Sorebonne5" and "Warsaw1" based on structural similarity as well as their physico-chemical properties.

3   Results

The location of our consensus pocket was found within the ATP binding site, between the RecA-like domains A1 and A2. The molecular docking with *AutoDock Vina* indentified 100 top scorers, which were ranked by their binding affinity, ranging from -10.3 to -8.8 $\frac{kcal}{mol}$. Passing these top 100 compounds through *Glide* returned the following top three compounds: ZINC000096077632 (popular name: Angiotensin 1-7), ZINC000008101127 and ZINC000035880991. While they exhibit good binding affinities, they still do not bind as well as ATP. *DiffDock* did not predict our Top 1 Glide scorer within our binding box, however the molecular dynamics simulation suggested a stable bond between the ligand and the helicase in our consensus pocket, as the ligand stayed inside of the pocket through the whole simulation. Multiple hydrogen bonds can be observed at the end of the 100 ns simulation, which had almost all been previously found by *Glide*. The estimation of toxicity and synthetic accessibility using *eToxPred* showed that only one out of out top 100 compounds is predicted to be toxic. Furthermore, most of the compounds seem to be accessible in their synthesis.

4 Discussion and Outlook

Molecular Docking was performed using the FDA approved ligands from ZINC and the pilot library from ECBD. The Top 100 Vina results were further analysed by Glide and their toxicity and synthetic accessibility estimated via eToxPred. Validation of our Top 10 scorer was achieved by blind docking with Diffdock and MD-Simulation using GROMACS. In conclusion, our findings suggest Angiotensin 1-7 as lead compound. This result is further supported by Mendoza-Torres et al. (2023), who proposed Angiotensin 1-7 as a new therapy to support the recovery from Covid-19. Moreover, because of its role in the Renin-Angiotensin-Aldosteron System, it seemingly could also lower the viral damage to the body by inhibiting NSP13.

In comparison, Sorbonne team 5 used a pre selection of the ECBD pilot library and subjected these to Vina. For their first ranked ligand our results predict an affinity score of -11.1 and was ranked 63 in our Vina results. The remaining 7 were mapped to lower with ranks between 123 and 2626. These results however aren't surprising, as Sorbonne 5 used a different binding pocket, compared to us.

4.1 Limitations of the project

Regarding molecular dynamics (MD) simulation, the project was held back by the tight time schedule. Further simulations should have been performed in order to validate our pipeline and compare our top scorer. Moreover, since MD simulations are a very stochastic process, they should always be estimated using replicas. Multiple runs of the same simulation, using replica exchange with dynamical scaling (Raubenolt et al. 2022) could have therefore deepened our confidence in the results.

Of further interest could have been the implementation of *AutoGrow4* (Spiegel and Durrant 2020). This could have helped, alongside toxicity and synthetic accessibility predictions via *eToxPred* to improve our lead drugs and to generate novel compounds.

References

Abraham, M. J., T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl (2015). "GRO-MACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". In: *SoftwareX* 1, pp. 19–25. ISSN: 2352-7110. DOI: 10.1016/j.softx.2015.06.001.

Berta, D., M. Badaoui, S. A. Martino, P. J. Buigues, A. V. Pisliakov, N. Elghobashi-Meinhardt, G. Wells, S. A. Harris, E. Frezza, and E. Rosta (2021). "Modelling the active SARS-CoV-2 helicase complex as a basis for structure-based inhibitor design". In: *Chemical Science* 12.40, pp. 13492–13505.

Brenke, R., D. Kozakov, G. Y. Chuang, D. Beglov, D. Hall, M. R. Landon, C. Mattos, and S. Vajda (2009). "Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques". In: *Bioinformatics* 25.5, pp. 621–7. ISSN: 1367-4803 (Print) 1367-4803. DOI: 10.1093/bioinformatics/btp036.

Corso, G., H. Stärk, B. Jing, R. Barzilay, and T. Jaakkola (2022). *DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking*. DOI: 10.48550/arXiv.2210.01776.

Eastman, P., J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande (2017). "OpenMM 7: Rapid development of high performance algorithms for molecular dynamics". In: *PLOS Computational Biology* 13.7, e1005659. DOI: 10.1371/journal.pcbi.1005659. URL: https://doi.org/10.1371/journal.pcbi.1005659.

Friesner, R. A., J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin (2004). "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy". In: *Journal of Medicinal Chemistry* 47.7. doi: 10.1021/jm0306430, pp. 1739–1749. ISSN: 0022-2623. DOI: 10.1021/jm0306430. URL: https://doi.org/10.1021/jm0306430.

Irwin, J. J., K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, and R. A. Sayle (2020). "ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery". In: *Journal of chemical information and modeling* 60.12, pp. 6065–6073. DOI: 10.1021/acs.jcim.0c00675.

Jakubec, D., P. Skoda, R. Krivak, M. Novotny, and D. Hoksza (2022). "PrankWeb 3: accelerated ligand-binding site predictions for experimental and modelled protein structures". In: *Nucleic Acids Research* 50.W1, W593–W597. ISSN: 0305-1048. DOI: 10.1093/nar/gkac389. URL: https://doi.org/10.1093/nar/gkac389.

Jendele, L., R. Krivak, P. Skoda, M. Novotny, and D. Hoksza (2019). "PrankWeb: a web server for ligand binding site prediction and visualization". In: *Nucleic Acids Research* 47.W1, W345–W349. ISSN: 0305-1048. DOI: 10.1093/nar/gkz424. URL: https://doi.org/10.1093/nar/gkz424.

Krivák, R. and D. Hoksza (2018). "P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure". In: *Journal of Cheminformatics* 10.1, p. 39. ISSN: 1758-2946. DOI: 10.1186/s13321-018-0285-8. URL: https://doi.org/10.1186/s13321-018-0285-8.

Le Guilloux, V., P. Schmidtke, and P. Tuffery (2009). "Fpocket: An open source platform for ligand pocket detection". In: *BMC Bioinformatics* 10.1, p. 168. ISSN: 1471-2105. DOI: 10.1186/1471-2105-10-168. URL: https://doi.org/10.1186/1471-2105-10-168.

Mendoza-Torres, E., P. Luna, M. Fernanda Pérez, J. Castellar-Lopez, A. Chang, Y. Montoya, J. Bustamante, and W. Rosales-Rada (Jan. 2023). "Potential of Angiotensin-(1-7) in COVID-19 Treatment". In: *Current Protein & Peptide Science* 24.1, pp. 89–97. ISSN: 1389-2037. DOI: 10.2174/1389203724666221130140416.

Pu, L., M. Naderi, T. Liu, H.-C. Wu, S. Mukhopadhyay, and M. Brylinski (2019). "e toxpred: A machine learning-based approach to estimate the toxicity of drug candidates". In: *BMC Pharmacology and Toxicology* 20, pp. 1–15.

Raubenolt, B. A., N. N. Islam, C. M. Summa, and S. W. Rick (2022). "Molecular dynamics simulations of the flexibility and inhibition of SARS-CoV-2 NSP 13 helicase". In: *Journal of Molecular Graphics and Modelling* 112, p. 108122. ISSN: 1093-3263. DOI: 10.1016/j.jmgm.2022.108122.

Spiegel, J. O. and J. D. Durrant (2020). "AutoGrow4: an open-source genetic algorithm for de novo drug design and lead optimization". In: *Journal of Cheminformatics* 12.1, p. 25. ISSN: 1758-2946. DOI: 10.1186/s13321-020-00429-4.

Trott, O. and A. J. Olson (2010). "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, pp. 455–461. DOI: 10.1002/jcc.21334.