Drug Repurposing for COVID-19

Angela Kralevska[1], Ivana Vichentijevikj[2], Monika Simjanoska Misheva[1]
[1] Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje
[2] iReason LLC, Skopje

Why Drug Repurposing?

- Faster & cheaper than de novo drug discovery
- Critical in urgent settings (COVID-19)
- Computational screening enables rapid prioritization

Goal

Compare two paradigms for COVID-19 drug repurposing

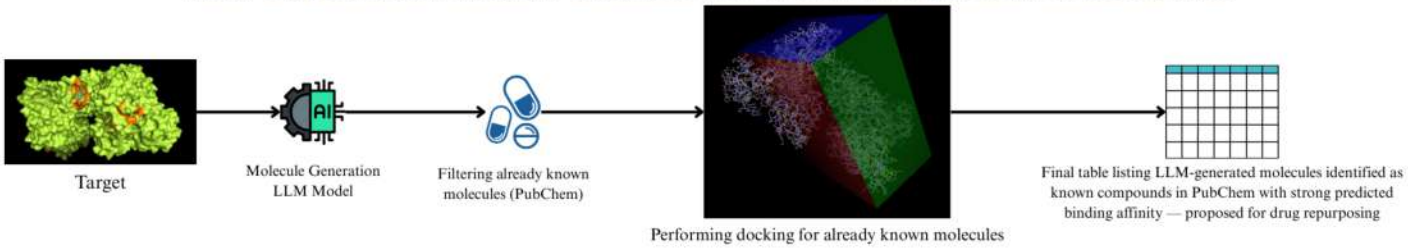
- Deep Learning (prediction)
- LLMs (generation)

Same molecular docking protocol used for validation

Study Design

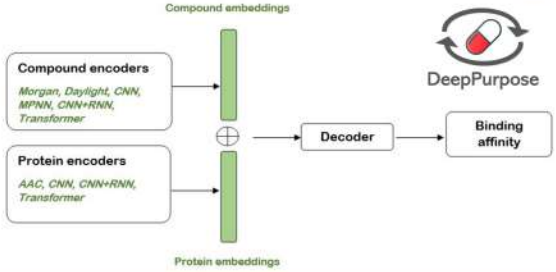
LLM Approach

DrugGen was used to generate target-conditioned SMILES from the SARS-CoV-2 3CLPro sequence. Generated molecules matched to known compounds via PubChem. Candidates were evaluated using the same AutoDock Vina.



Deep Learning Approach

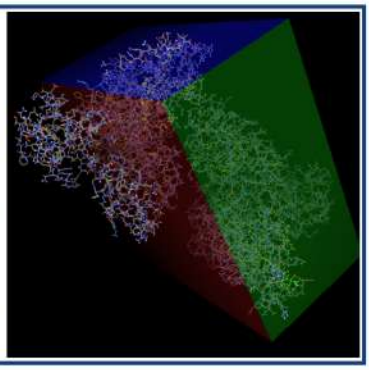
DeepPurpose models were pretrained on Davis and KIBA, fine-tuned on a 3CLPro bioassay, and applied to the CAS COVID-19 dataset for virtual screening.



Molecular Docking (Validation)

Molecular docking for selected drug candidates was performed, for validation, against the SARS-CoV-2 main protease (3CLPro, PDB ID: 6M2N).

Grid box coordinates: −48, −29, 27, size 64 × 126 × 122 Å, software: AutoDock Vina.



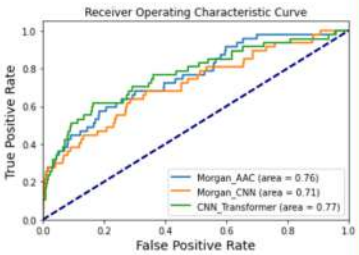
Results

Binding affinity prediction on Davis Dataset

Compound encoder	Protein encoder	MSE	Pearson Correlation	Concordance index
CNN	Transformer	0.24	0.84	0.88
MPNN	AAC	0.31	0.79	0.86
MPNN	CNN	0.27	0.82	0.87

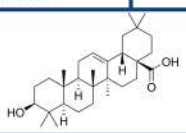
Best Model after fine-tuning on bioassay data for COVID-19 by high throughput screening assay to identify inhibitors of the 3CLPro.

Compound encoder	Protein encoder	AUROC	AUPRC	F1	Cross-entropy loss
CNN	Transformer	0.77	0.24	0.26	1.43



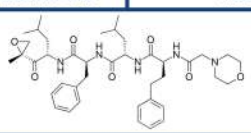
Docking Ranking of Compounds Prioritized by a Model Fine-Tuned on SARS-CoV-2 3CLPro Bioassay Data

Compound	Docking score (kcal/mol)
Oleanoic acid	-9.2
Paxlovid	-9.2
Aurintricarboxylic acid	-8.8
Sofosbuvir	-8.5



Docking Ranking of Compounds from LLM Workflow

Compound	Docking score (kcal/mol)
Carfilzomib	-9.505
Flualprazolam	-7.897
Amlexanox	-7.366
2-Chloroadenosine	-7.358



Conclusion

This work evaluates two complementary computational routes for COVID-19 drug repurposing. A deep learning–based drug–target interaction model prioritized candidate compounds, with **oleanolic acid** emerging as a top-ranked candidate after structure-based docking validation. In parallel, an LLM-driven, dataset-independent molecule generation approach produced target-conditioned candidates, identifying **Carfilzomib** among the strongest-scoring compounds. Using a shared docking protocol as a common validation step, the results highlight how predictive screening and generative modeling can support rapid candidate prioritization.

Funding

Funded by the European Union under Horizon Europe project (ChatMED Grant Agreement ID: 101159214). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Ex-ecutive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Huang, K., Fu, T., Glass, L. M., Zitnik, M., Xiao, C., and Sun, J. (2020). DeepPurpose: a deep learning library for drug-target interaction prediction. Bioinformatics, 36(22-23):5545-5547.
Organization, W. H. (2022). Who coronavirus (covid-19) dashboard. https://covid19.who.int/. Accessed: 2022-03-15.
Sheikholeslami, M., Mazrouei, N., Gheisari, Y., Fasihi, A., Irajpour, M., and Motahharynia, A. (2024). Druggen: Advancing drug discovery with large language models and reinforcement learning feedback. arXiv.
Trott, O. and Olson, A. J. (2010). Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, 31(2):455-461.