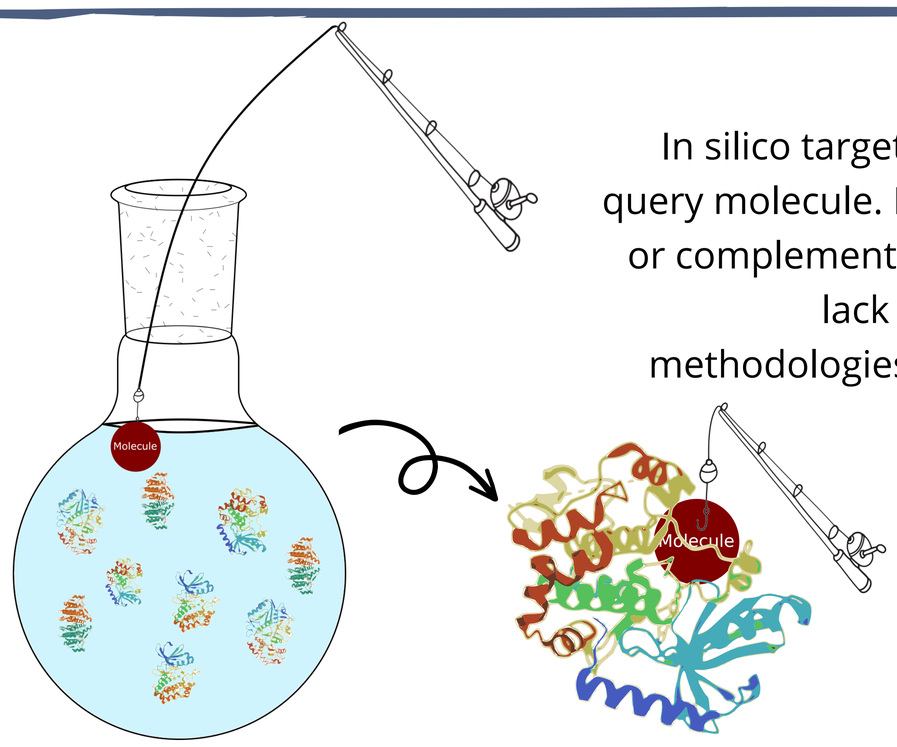
TARGET FISHING IN NEURODEGENERATIVE DISEASES: A CASE STUDY USING A NOVEL DOCKING-MACHINE LEARNING APPROACH

Fernández, Julián F.[1,2], Martinez Heredia, L. [3], Perez, C.[4] Palermo, J.A. [1,2] Rodríguez-Franco, M.I. [4], Lavecchia, M.J. [3]

[1] Departamento de Química Orgánica - FCEN - UBA, [2] UMYMFOR (UBA CONICET), Buenos Aires, Argentina, [3] CEQUINOR (UNLP CONICET), La Plata, Argentina, [4] Instituto de Química Médica, Madrid, España



Introduction

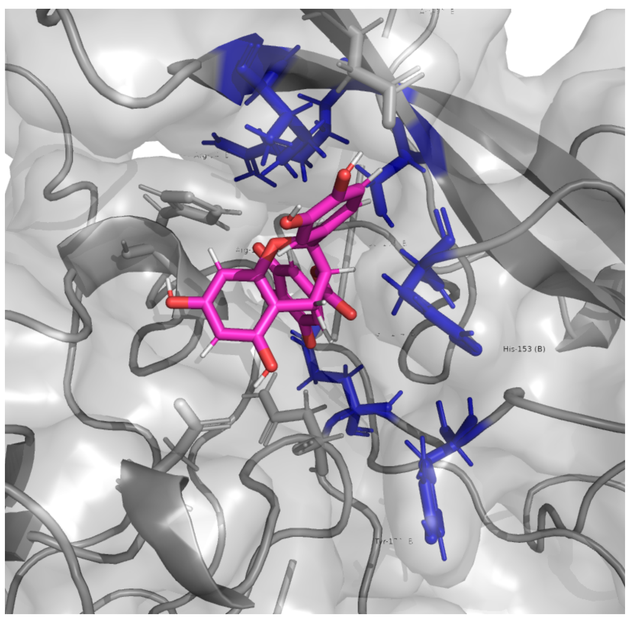
In silico target prediction or "target fishing" consists of identifying the most promising targets of a query molecule. It is a promising field in drug discovery, since it has the potential to be an alternative or complement for the time and cost consuming experimental equivalent. However, because of the lack of performance metrics and the confusing results given by the available tools, these methodologies are rarely-used in drug discovery projects. Therefore, the development of new and efficient methods remains as a challenging task.

In the present study, the first real-world application of a new, open-source, structure-based target prediction methodology named Target Fisher is presented. A library consisting of 1500 already synthesized compounds was screened looking for new inhibitors in key targets for neurodegenerative diseases.

How does it work?

Target Fisher works by first docking the query molecule in each one of the selected targets using Autodock Vina. Secondly, an interaction fingerprint between the ligand and the protein is extracted using the best scoring pose in each case. These energy fingerprints are then evaluated in models trained specifically for each target using machine learning. The final prediction is made by a hybrid approach between the docking score and the result obtained from the machine learning model.

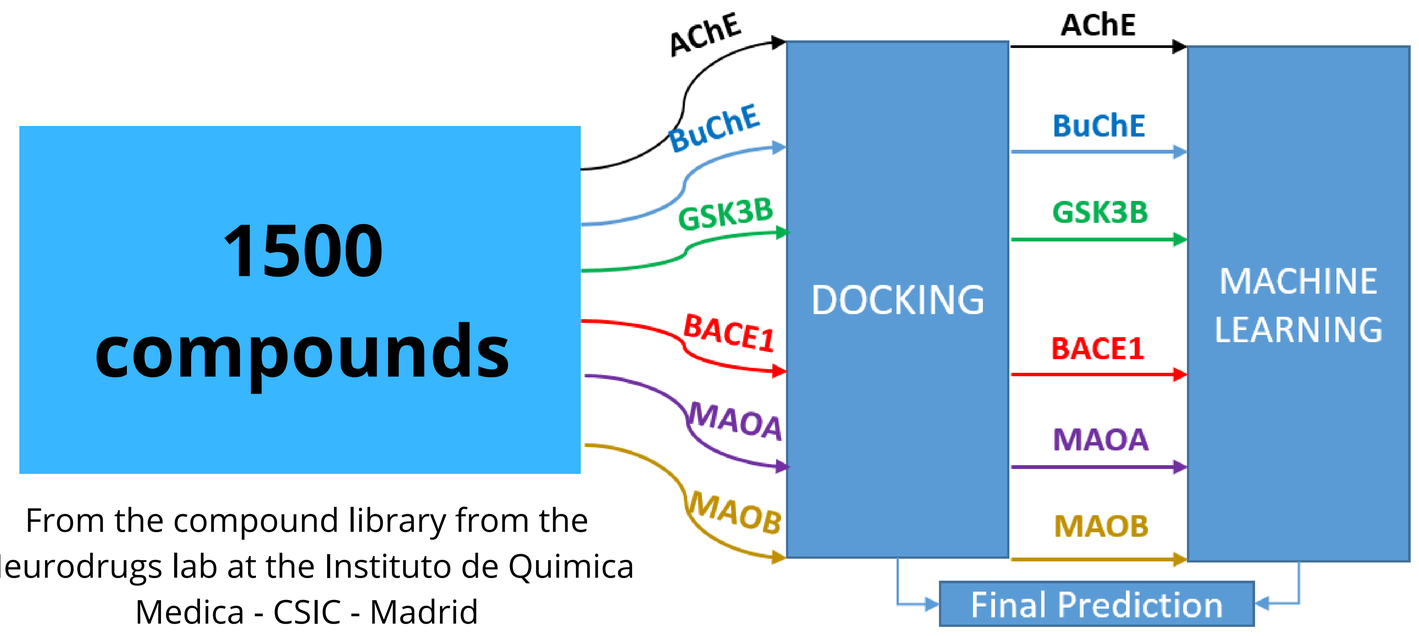
A preliminary version of Target Fisher can be found at https://gitlab.com/gqc/target-fisher with more than 30 available targets.



Tools and Software

Docking Software: Autodock Vina
Interaction Fingerprints: Home-made adaptation of OBEnergy (Open Babel) for ligand-protein complexes.
Forcefield: MMFF94
Activity data: ChEMBL Database
Machine Learning: Sci-kit learn. All models were trained using cross validation and hyperparameters tuning was made considering the best combination of multiple metrics.

Experimental Validation



From the compound library from the Neurodrugs lab at the Instituto de Quimica Medica - CSIC - Madrid

- **10 compounds selected for testing**
- **1 novel active molecule in GSK3B**
- **20 compounds from the same family tested in GSK3B**
- **2 more active compounds**
- **Clear structure-activity motif**

A library of 1500 compounds was screen against 6 key targets for neurodegenerative diseases. Those unplished and untested compounds that displayed an activity profile similar to the known active compounds both in docking and machine learning models were selected. From the tested products, one resulted in a new inhibitor of GSK3B in the uM scale (**7,91 µM**). Consequently, 20 more compounds from the same family were also tested, finding that 2 of them were also active against this kinase (**7,33 µM and 55 nM**).

Final Conclusions & Perspectives

✓ Target Fisher was applied succesfully to a 1500 compound library, enabling the discovery of a three new GSK3B inhibitors

SOON

- Synthesis of new derivatives with the active core structure
- Online platform that lets the user dock and predict activities using the available targets. A preview can be seen in: https://gqc.quimica.unlp.edu.ar/targetfisher/

Acknowledgments

- University of Buenos Aires for Julián F. Fernández doctoral's fellowship and funding his research stay in Madrid
- The financial support of the Spanish Research Agency is gratefully acknowledged (grant PID2021-122650OB-I00).

jffernandez@qo.fcen.uba.ar

SPANISH Drug Discovery Network