The application of AI-driven Drug Discovery technology for molecular optimization of nuclear receptor ligands.

Rafał A. Bachorz
Simulations Plus, Inc., 42505 10th Street West, Lancaster, CA 93534, USA
rafal.bachorz@simulations-plus.com

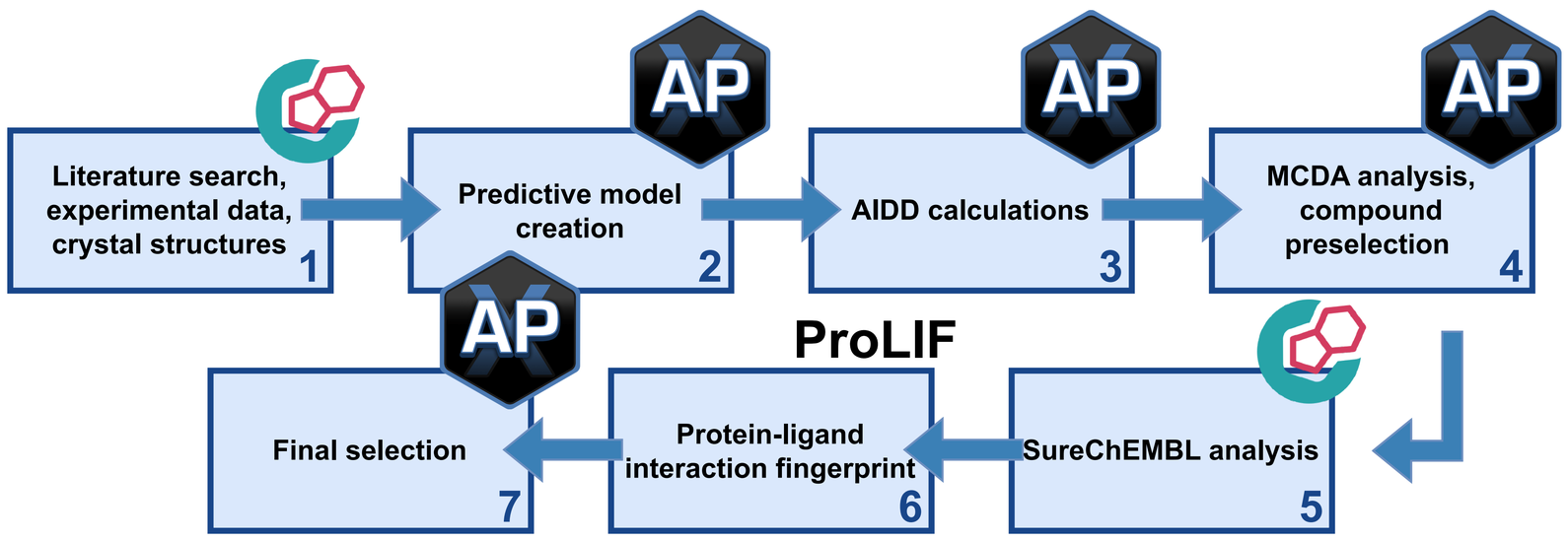
SimulationsPlus

Introduction

Nuclear receptors (NRs) are a superfamily of transcription factors whose activity is regulated upon the binding of a specific ligand. In this work, we utilized the new AI-driven Drug Design (AIDD) module within the ADMET Predictor® suite to derive novel and promising RORγT agonists with suitable predicted ADMET and PK properties. The AIDD module uses chemically intelligent SMIRKS transformations to generate new molecules based on seed compounds. The generation process is channeled towards molecules of desired properties within a multicriteria optimization loop. The objectives, including potency and selectivity at the chosen target, synthetic feasibility, and ADMET/pharmacokinetic (PK) properties are considered simultaneously. Thus, the properties of optimized molecules are on a Pareto front and are excellent candidates for experimental verification.

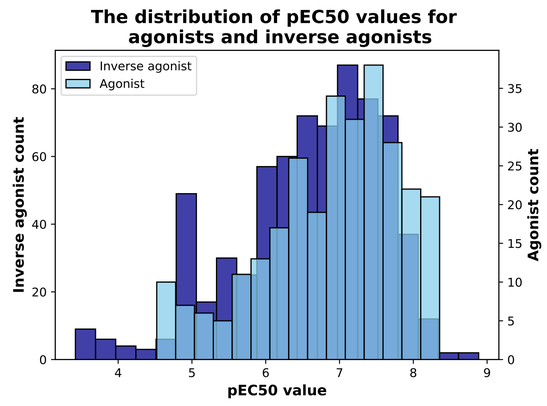
Proposed workflow

The entire project has been carried out in a systematic way according to the flowchart presented below. The crucial elements of the workflow are supported by the ADMET Predictor® program.



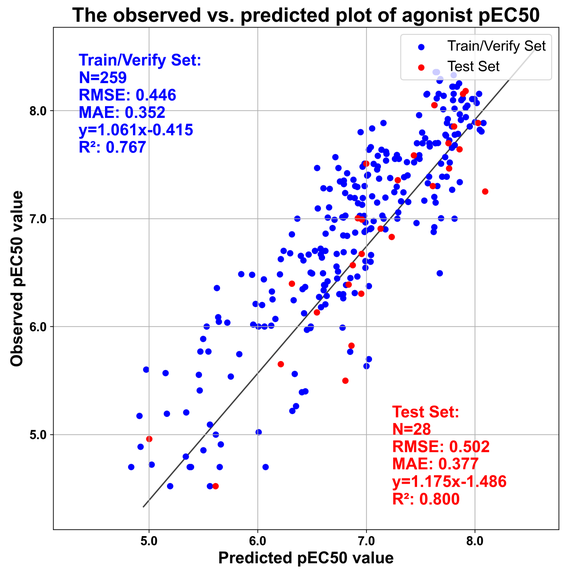
1. Experimental data

Biological data for 3527 number of molecules was extracted from the ChEMBL database, including both IC50 and EC50 data. Expert knowledge was used to categorize molecules as agonists, inverse agonists, and antagonists.



2. Predictive models

Following data curation, we used ADMET Modeler™ , part of the ADMET Predictor® suite, to develop classification and regression models to predict RORγT agonist activity.



3. Generative chemistry

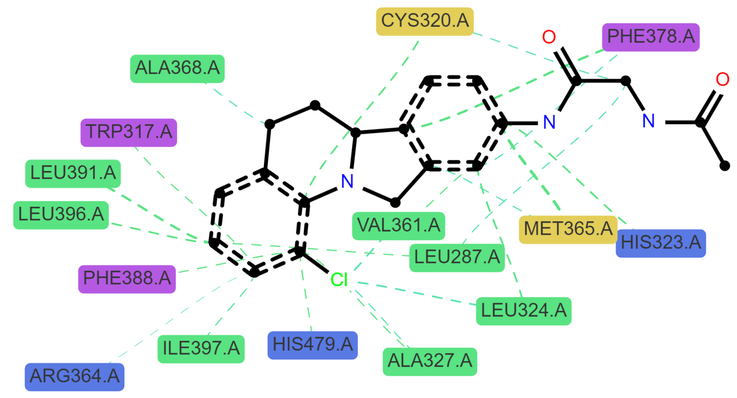
The *de novo* drug design was carried out with the AIDD™ module integrated in ADMET Predictor® program. New compounds, created within the evolutionary loop, are subject to quality estimation with certain objective functions. Within this work we chose 4 goal functions: predicted agonistic pEC50, 3D pharmacophoric similarity, absorption risk and synthetic difficulty.

5. SureChEMBL analysis

An important factor related to the *de novo* drug design is the novelty and patentability of generated compounds. To address this issue, each AIDD-created molecule has been confronted with the SureChEMBL data base containing comprehensive collection of chemical structures extracted from patent literature.

6. ProLIF analysis

To enhance the compound selection process, we applied the ProLIF[1] Protein-Ligand fingerprint analysis. The ProLIF framework analysis is capable of quantifying this interaction based on the complex geometry. The comparison with the reference interaction fingerprint of the known agonist crystal structure is an additional presumption in the selection process.

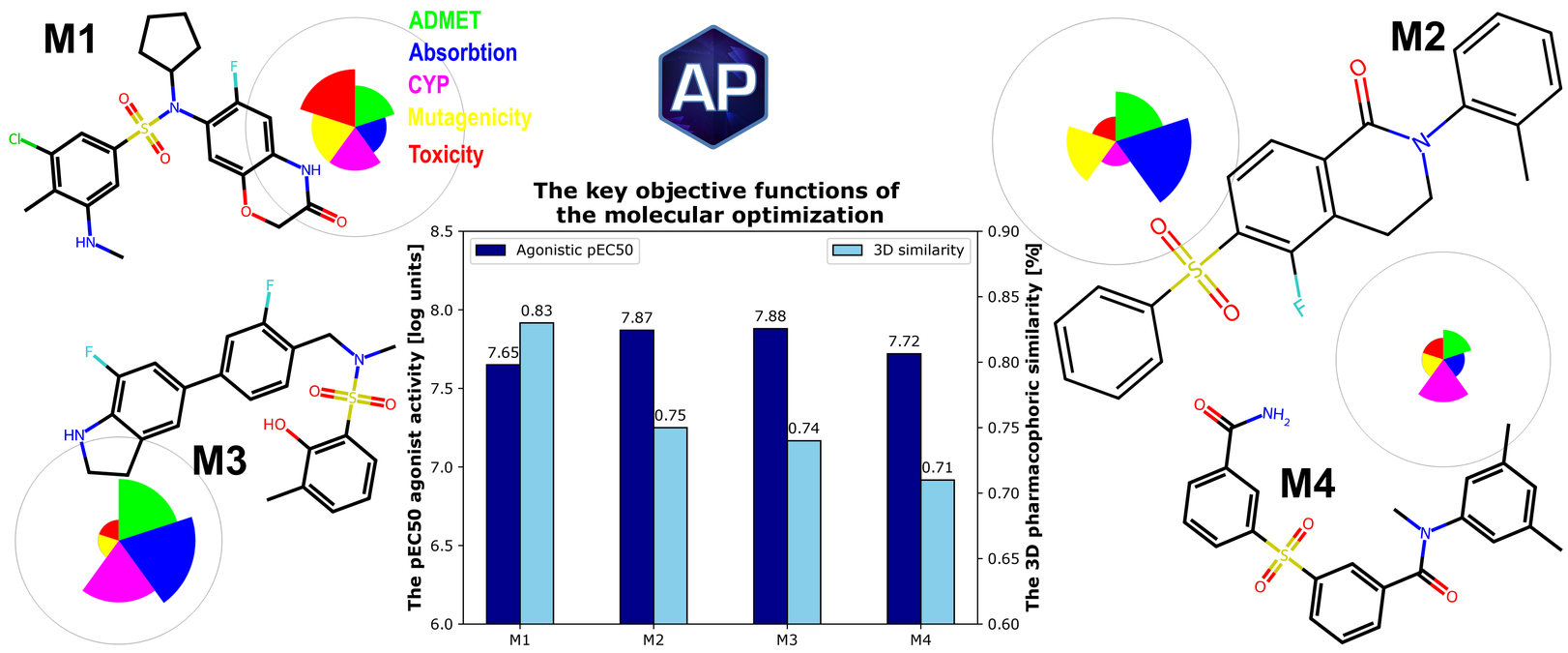


4. MCDA analysis

Created molecules need to be carefully inspected and a final selection of compounds has to be made. To support the analysis we have applied here well-known Multicriteria Decision Analysis (MCDA) technique. The Vikor method [2] takes into account the maximum group utility and the minimum group regret. From the practical perspective it accepts the user preferences in terms of weights and provides the compounds sorted according to a global utility function.

7. Final selection and summary

For the final verification we have chosen ca. 100 potential agonists created within the AIDD™ generative chemistry module.



References

[1] C. Bouysset and S. Fiorucci. Prolif: a library to encode molecular interactions as fingerprints. *Journal of Cheminformatics*, 13(1):72, Dec 2021.

[2] P. L. Yu. A class of solutions for group decision problems. *Management Science*, 19(8):936–946, 1973.