Targeting TrmD methyltransferase Activity: A Docking Simulation and Reinforcement Learning Approach

Ladislav Buček, Zuzana Nováková, Věra Štěpánková

Faculty of Science
Charles University

INTRODUCTION

TrmD methyltransferase is a knotted bacterial enzyme. It catalyzes the methylation of guanine at the N1 position of **tRNA**, a post-transcriptional modification required for proper tRNA folding, stability, and accurate decoding during protein synthesis. Previous research suggests an important role of TrmD in bacterial growth, making it a possible target for antimicrobial drug development. The primary goal of this project was to identify potential TrmD methyltransferase **inhibitors**. Such inhibitors could hold potential for further research as novel antibiotic candidates. [1]

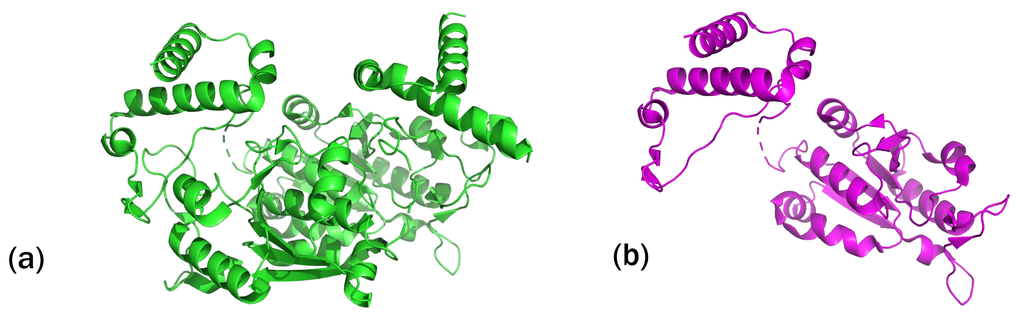


Fig. 1: Structural comparison of TrmD dimer (a) and TrmD monomer (b). (4YVG – Visualized using PyMOL.)

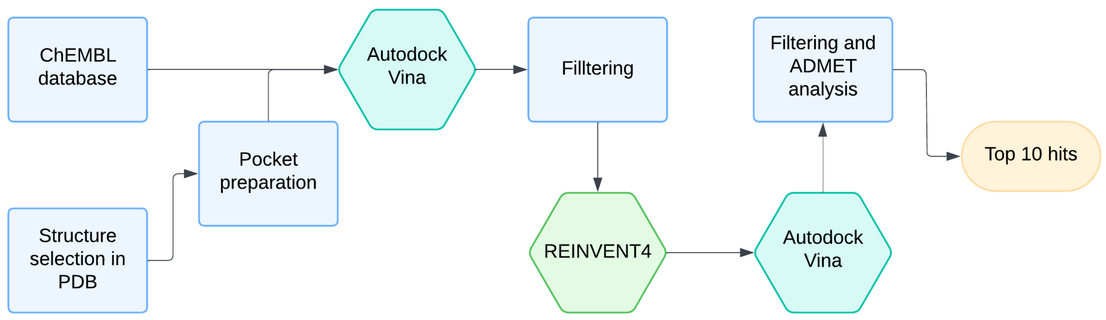


Fig. 2: Pipeline diagram

TARGET PREPARATION

The experimental structure **4YVG** from H. influenzae [1] was selected as the docking target due to being the **highest-resolution TrmD monomer** available (1.55 Å). For specific docking site **AdoMet binding pocket** was identified using machine learning based tool **P2Rank.** [2]

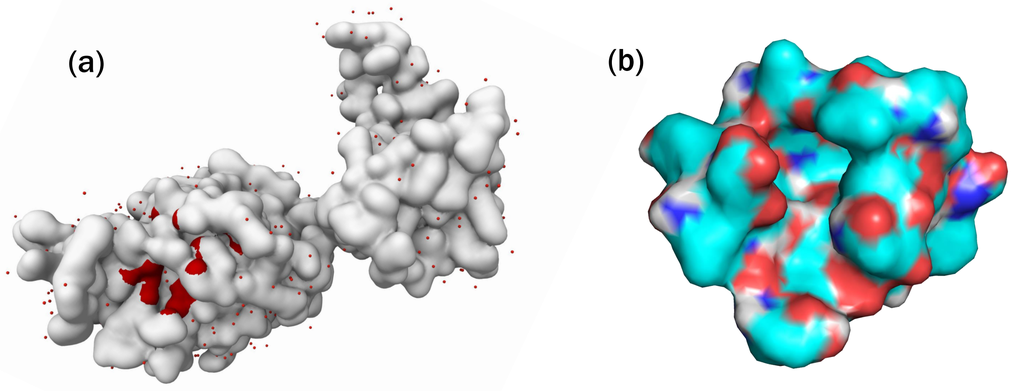
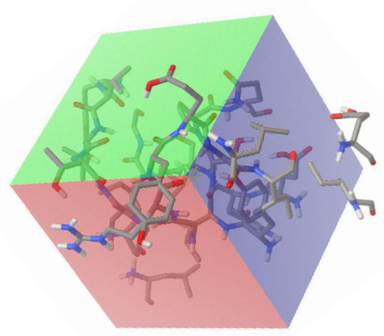


Fig. 3: AdoMet binding pocket visualization: The binding pocket of the TrmD enzyme visualized with (a) P2Rank and(b) PyMOL

INITIAL DOCKING SIMULATION

Docking simulations were performed on 69 **TrmD-associated** molecules retrieved from the **ChEMBL** database using **AutoDock Vina** [3]. Molecules scoring lower then **-8 kcal/mol** were chosen for reinforcement learning input.



center_x	center_y	center_z
45.650	1.391	11.014

Fig. 4: Configuration file preparation for docking using **AutoDock Tools**: grid size and grid center selection.

REINFORCEMENT LEARNING

A two-stage pipeline using the **Mol2Mol** mode of **REINVENT4** [4] was implemented to optimize the best TrmD inhibitors from initial screening. Stage 1 focused on exploring the chemical space, ensuring scaffold diversity, and eliminating undesirable features. Stage 2 refined these candidates to prioritize drug-like properties while maintaining structural novelty. Multinomial sampling, diversity filters, and scaffold-based priors guided the generation of **9,180** molecules.

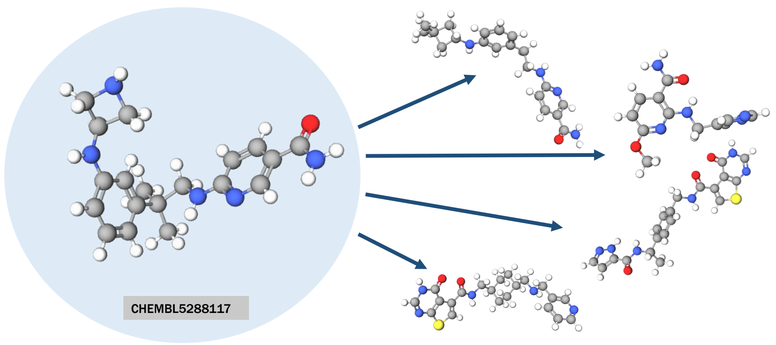


Fig.5: Reinforcment learning digram, visualised using **MolView**.

DOCKING OF GENERATED MOLECULES

REINVENT-generated molecules were further evaluated for their binding affinity to the TrmD active site through a second docking simulation. Redocking was performed on the top **500 best hits** to ensure high-quality results.

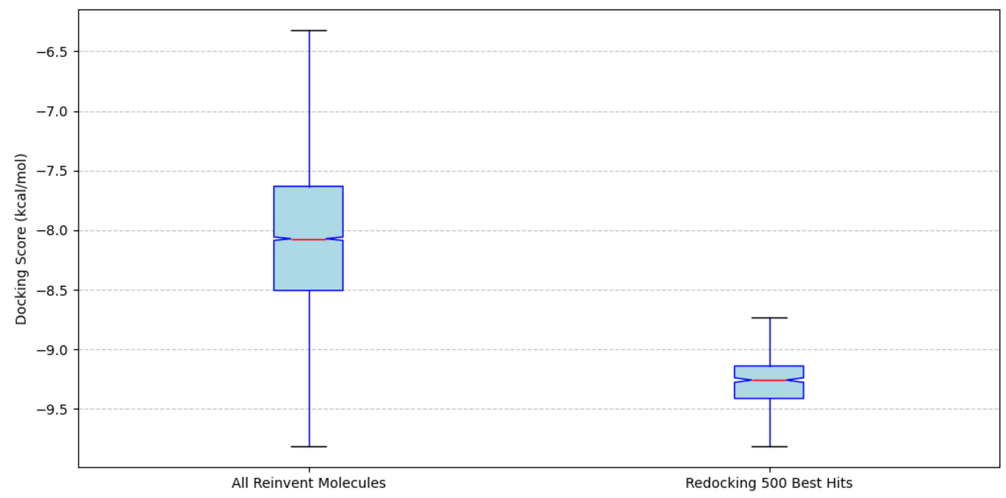


Fig. 6: Box plot of **affinity scores** for **Reinvent-generated** molecules: The plot compares the distribution of binding affinity scores for all 9180 molecules generated by Reinvent and the top 500 ligands with the highest scores.

ADMET SCREENING

Molecules with the highest predicted affinity were assessed for their **drug-like properties** and **toxicity** using computational tools such as **RDKit** [5], **SwissADME** [6], and **ADMET-AI** [7]. After filtering, approximately 96 % of our top 500 ligands successfully adhered to **Lipinski's Rule of Five.** [8]

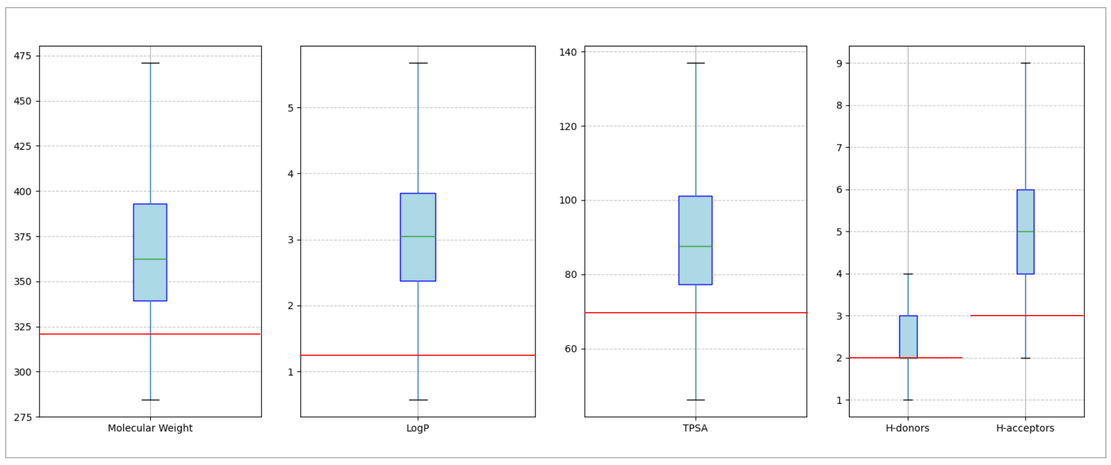
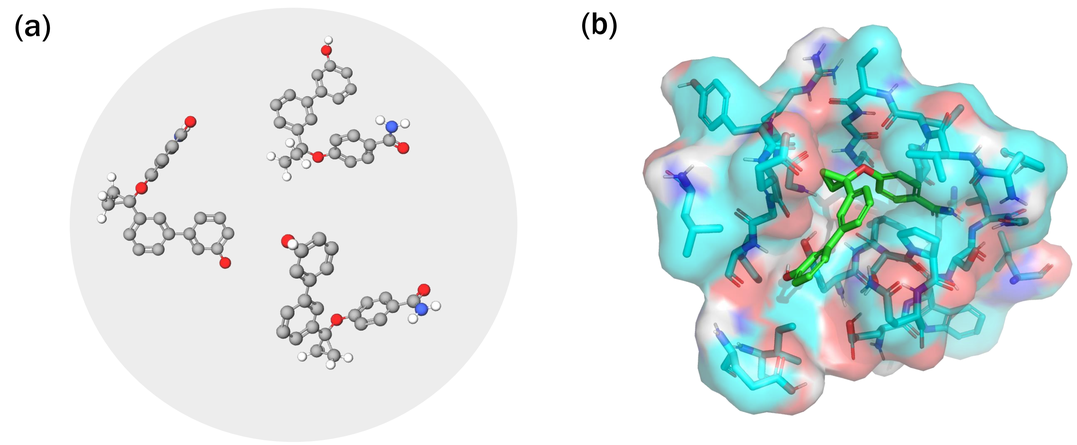


Fig. 7: Box plot of **drug-like properties** for top **500 best scoring ligands**: The plot showcases the distribution of key drug-like properties, including molecular weight, LogP, topological polar surface area (TPSA), hydrogen donors, and hydrogen acceptors, as calculated by **RDKit.** Red lines showing values for ligand 3664.

RESULTS AND CONCLUSIONS

After evaluating the **top 10** ligands across various **toxicity** parameters, including the **AMES** test (mutagenicity), **hERG** inhibition (cardiotoxicity), and hepatotoxicity, we selected molecule number 10 (indexed as **3664**) as the final TrmD inhibitor candidate with a binding affinity score of **-9.918** kcal/mol. This ligand demonstrated the most balanced performance across key criteria such as toxicity, solubility, and bioavailability, while maintaining a **high affinity score**, making it the most promising option for further development.



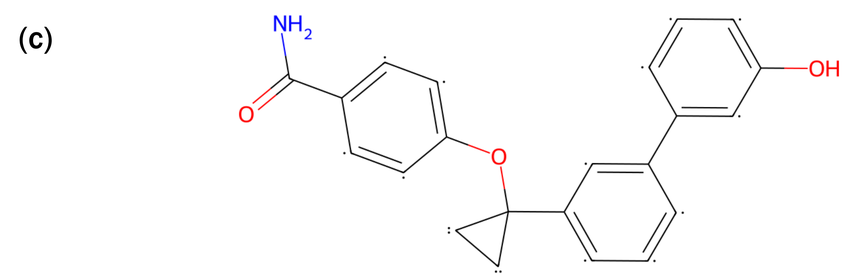


Fig. 8: Visualization of compound 3664: (a) 3D model of the molecular structure, (b) molecule bound in the binding pocket (visualized with PyMOL), (c) 2D representation of the molecular structure

We found that the reinforcement learning approach, utilizing state-of-the-art **REINVENT4** models, yielded promising results in designing molecules with basic drug-like properties and high synthetic accessibility. However, balancing binding affinity with potential **ligand toxicity** proved to be a significant challenge. Notably, all of our top 10 molecules showed a high probability of liver injury. Future research could benefit greatly from incorporating **docking simulations on reinforcement learning generated ligands** in drug design, but thorough experimental validation remains essential

References: 1. Ito, T. *et al.* Structural basis for methyl-donor–dependent and sequence-specific binding to tRNA substrates by knotted methyltransferase TrmD. *Proc. Natl. Acad. Sci.* **112**, (2015). 2. Krivák, R. & Hoksza, D. P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *J. Cheminformatics* **10**, 39 (2018). 3. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **61**, 3891–3898 (2021). 4. Loeffler, H. H. Reinvent 4: Modern AI–driven generative molecule design. (2024). 5. RDKit: Cheminformatics and Machine Learning Software. https://www.rdkit.org/ (2013). 6. Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 42717 (2017). 7. Swanson, K. *et al.* ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. *Bioinformatics* **40**, btae416 (2024). 8. Lipinski, C., Lombardo, F., Dominy, B. & Feeney, P. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* (1997) doi:10.1016/s0169-409x(00)00129-0.