Designing a bacterial TrmD inhibitor

Krzysztof Łukasz[1], Lidia Stadnik[1], and Michał Osiński[1]

[1]Faculty of Mathematics, Informatics, and Mechanics, University of Warsaw

Abstract

The bacterial tRNA-(N^1G37) methyltransferase (TrmD) is an essential enzyme in many pathogens, making it a promising target for novel antimicrobial agents. This study employed a comprehensive computational approach to identify and design potential TrmD inhibitors. Initially, known TrmD ligands were retrieved from the PubChem database and literature sources, and their SMILES representations were utilized to identify structurally similar compounds. Furthermore, a Variational Autoencoder (VAE) was implemented to generate novel chemical structures based on known ligands, expanding the pool of potential inhibitors. Candidate molecules from both similarity searches and VAE-generated libraries were subjected to molecular docking studies using AutoDock to evaluate their binding affinities to the TrmD active site. Docking results revealed several compounds with high binding affinities, highlighting their potential as TrmD inhibitors. These findings lay the groundwork for experimental validation and further optimization of these compounds as antimicrobial agents targeting TrmD.

1 Introduction

The global rise in antimicrobial resistance has created an urgent need for novel therapeutic strategies to combat bacterial infections. Among the promising targets for antibiotic development is the bacterial tRNA-(N^1G37) methyltransferase (TrmD), an essential enzyme involved in post-transcriptional modification of tRNA. TrmD catalyzes the methylation of guanosine at position 37 in tRNAs, a modification critical for maintaining translational fidelity and preventing frameshift errors during protein synthesis. (Hou et al. 2017) This enzymatic function is indispensable for bacterial survival and growth, as evidenced by its essentiality in numerous pathogens, including Acinetobacter baumannii, *Pseudomonas aeruginosa*, and *Mycobacterium tuberculosis.(Zhong et al. 2019)*

TrmD is structurally distinct from its eukaryotic counterpart, TRM5, which makes it an attractive target for selective inhibition with minimal off-target effects on human cells. (Christian et al. 2004; Goto-Ito, Ito, and Yokoyama 2017) Structural studies have revealed that TrmD binds S-adenosylmethionine (SAM) as a methyl donor, and its active site architecture differs significantly from TRM5, enabling the design of selective inhibitors. (Ito et al. 2015) Despite its essentiality and ligandability, translating TrmD inhibition into effective antimicrobial therapy has faced challenges. These include poor bacterial cell permeability of inhibitors and questions about the enzyme's druggability in vivo, as some studies suggest that even potent TrmD inhibitors fail to exhibit significant antibacterial activity due to high cellular enzyme abundance or insufficient target engagement. (Wilkinson et al. 2023)

Recent advances in computational drug discovery have provided efficient tools to identify potential TrmD inhibitors. Virtual screening methods, such as similarity-based compound selection and molecular docking, are increasingly used to predict binding interactions between candidate molecules and TrmD's active site. High-throughput screening (HTS) campaigns have also identified several chemical scaffolds with potent TrmD inhibitory activity. For instance, pyridine-pyrazole-piperidine derivatives have shown promising binding affinity to the SAM-binding pocket of TrmD, although their antibacterial efficacy remains limited. (Zhong et al. 2019) These findings highlight the need for further optimization of inhibitor properties, such as enhancing

bacterial permeability and minimizing efflux susceptibility. In our work, we aimed to identify novel TrmD inhibitors by employing a comprehensive computational pipeline. Initially, known ligands were retrieved from the PubChem database and literature sources to serve as reference structures for similarity-based virtual screening. Additionally, a Variational Autoencoder (VAE) was implemented to generate novel chemical structures, leveraging molecular encoding and decoding techniques to explore a broader chemical space. Candidate compounds, including both VAE-generated and similarity-screened molecules, were evaluated through molecular docking studies using AutoDock (Eberhardt et al. 2021) to assess their binding affinity and interaction profiles with the TrmD active site. By integrating these advanced computational approaches with structural insights into TrmD's mechanism, this work seeks to contribute to the development of selective and potent inhibitors that can serve as leads for novel antimicrobial agents targeting resistant bacterial pathogens.

2 Materials and methods

2.1 Selection of TrmD structures

Antimicrobial resistance poses a severe global crisis with profound societal and healthcare consequences. Each year, over 2 million cases of multidrug-resistant (MDR) bacterial infections are reported, including MDR Pseudomonas aeruginosa, a species listed in the World Health Organization's 2024 Bacterial Priority Pathogens List (BPPL) as a high-priority pathogen. Also included on the BPPL is *Acinetobacter baumannii*, classified as a critical priority pathogen due to its resistance to last-resort antibiotics. The Mycobacterium abscessus complex, while not explicitly listed in BPPL, is responsible for 2.6% to 13.0% of all non-tuberculous mycobacterial pulmonary infections. These infections are notoriously difficult to treat due to the need for complex regimens, widespread drug resistance, and significant adverse effects. (Dedrick et al. 2022)

In light of these challenges, TrmD structures from these three species were selected for further analysis *Mycobacterium abscessus* (PDB: 6QRB), *Acinetobacter baumannii* (PDB: 7MYS) and *Pseudomonas aeruginosa* (PDB: 5WYQ).

Structural superimposition was performed using Chimera to validate the similarity of the TrmD active sites across these species. Pairwise root-mean-square deviation (RMSD) values of the active sites were calculated to quantify similarity:

- 5WYQ vs. 6QRB: 0.735
- 5WYQ vs. 7MYS: 0.309
- 6QRB vs. 7MYS: 0.774

The high degree of similarity in the active sites confirmed the feasibility of utilizing ligands from all three species for inhibitor design. This similarity also enabled the rationalization of performing molecular docking primarily on one structure, ensuring time efficiency while maintaining the potential for cross-species TrmD inhibition.

2.2 Selection of ligands

Two sources were used to curate ligands for study:

1. PDB Ligands: A total of 59 ligands co-crystallized with TrmD from the selected bacterial species were extracted from the PDB.
2. Literature-Derived Ligands: Ligands were also collected from three key publications, yielding an additional 16 ligands: 7 from (Wilkinson et al. 2023), 4 from (Zhong et al. 2019) and 5 from (Vlasov et al. 2023).

Ligands obtained from the literature were converted to SMILES format using the OPSIN parser and standardized to canonical SMILES using RDKit. These ligands were pooled with the PDB-derived ligands for subsequent analysis.

2.3 Compound Filtering and Similarity Search

A dataset of 2,512,731 compounds was retrieved from the PubChem database and filtered based on drug-likeness and physicochemical properties. The filtering criteria were based on Lipinski's Rules, which define acceptable ranges for molecular properties to improve oral bioavailability, and additional physicochemical constraints as outlined by (Kralj, Jukič, and Bren 2023):

- Lipinski's Rules: molecular weight $\leq$ 500 Da, hydrogen bond donors ≤ 5, hydrogen bond acceptors ≤ 10, logP (partition coefficient) ≤ 5
- Charge: Net charge between -2 and +2.
- Polar Surface Area (TPSA): ≤ 140 Å^2.

The filtering process reduced the dataset to 1,187,012 compounds. Structurally similar compounds were identified using the LINGO similarity metric, a method designed to compare SMILES strings by analyzing three-element segments (lingos) within the strings. This approach, inspired by methodologies described by (Öztürk, Ozkirimli, and Özgür 2016), operates as follows:

1. Canonical SMILES strings of compounds are used to ensure consistency in representation.
2. SMILES strings are broken into overlapping lingo segments of three characters each.
3. For each compound pair, the differences in lingo segment counts are computed.
4. The differences are aggregated into a single similarity score.

This metric effectively captures structural nuances in SMILES strings, enabling robust identification of compounds resembling known ligands. The integration of this method facilitated the prioritization of candidates for subsequent analysis, optimizing the identification of potential inhibitors from the filtered dataset.

2.4 Compound generation

To generate novel chemical structures based on known ligands and similar compounds, a Variational Autoencoder (VAE) was implemented using a framework based on Keras' example for molecule generation (`https://keras.io/examples/generative/molecule_generation/`). (Gómez-Bombarelli et al. 2018; Cao and Kipf 2022) The methodology involved encoding and decoding molecular structures into a latent space using a matrix-based format. A connection matrix was used, representing bond types as triangular matrices (1 = single bond, 2 = double bond, etc.). Feature vectors captured atom-specific properties, such as indices in an atom dictionary. Molecules were encoded into a compressed latent representation. New molecular structures were generated by decoding from the latent space.

This approach, originally adapted from image generation techniques, facilitated the creation of novel molecular structures by capturing key patterns in molecular data. Initial results demonstrated the successful generation of simple molecules with up to seven atoms and single branches. However, challenges arose when attempting to extend this approach to more complex molecules.

2.4.1 Challenges and Potential Solutions

Several challenges were encountered. The complexity of molecular structures led to high loss values during training. Also, generated molecules predominantly featured carbon chains with occasional nitrogen or oxygen atoms. Furthermore, training on longer molecules resulted in large input vectors (e.g., 281 kB for 70-atom molecules).

To address these issues, the following strategies were proposed:

- Pre-training the model on small molecules with simple chains, followed by fine-tuning for target molecules.
- Reducing input size by encoding common substructures as tokens to simplify the representation of larger molecules.

2.5 Molecular docking

AutoDock Vina is one of the fastest and most widely used open-source docking engines with over 33,000 citations reported in Google Scholar. It is a turnkey computational docking program that is based on a simple scoring function and rapid gradient-optimization conformational search. (Eberhardt et al. 2021)

2.5.1 Receptor Preparation

The receptor protein structure was prepared using a custom Python script that automates the process of receptor preparation for docking. The script utilizes the ProDy library (Bakan, Meireles, and Bahar 2011) to parse the Protein Data Bank (PDB) file of the target protein and select relevant receptor atoms based on user-defined criteria. Ligand atoms were also selected to define the docking box center. The receptor preparation involved the following steps:

- The PDB file of TrmD from Pseudomonas aeruginosa (under accesion ID 5WYQ) was downloaded from the RCSB Protein Data Bank, receptor and ligand atoms were selected and their coordinates were saved in separate PDB files.
- The docking box center was calculated as the geometric center of the ligand atoms, while the box size was defined to be cubical with size equal to 30Å.
- The receptor PDB file was processed with Reduce2 (Grosse-Kunstleve et al. 2002) to add missing hydrogens and optimize protonation states.

2.5.2 Ligand Preparation

Ligand molecules were prepared using a pipeline that converts SMILES strings into docking-ready PDBQT files. For this purpose we have chosen top 5000 ligands sorting by similarity measure as returned by LINGOSIM. The ligands generated by VAE were not docked doe to their relativley small size. A script was employed for this purpose:

- SMILES strings of known ligands and structurally similar compounds were processed using RDKit to generate 3D molecular structures.
- Ligands were optimized for pH 7.35 with skipping of tautomerization and acid-base state adjustment toggled on.

2.5.3 Docking Procedure

Molecular docking was performed using AutoDock Vina to predict the binding affinity of ligands to TrmD. Each of the selected ligands was docked into the receptor's binding site using AutoDock Vina with an exhaustiveness parameter set to 32 for thorough sampling. The docking results were saved in PDBQT format and for each ligand top docking poses were reported by AutoDock with corresponding affinities.

3 Results

3.1 Docking Results

A total of 5000 compounds was selected for docking with AutoDock Vina. A mean value of binding affinity was -7.15 kcal/mol with standard deviation 0.80 and ranging from -4.24 to -9.98 kcal/mol. The top identified molecules and their binding affinity are summarised in table 1 and their 2D structures in figure 1.

	PubChem ID	Affinity (kcal/mol)
0	155792348	-9.98
1	8171885	-9.89
2	54844255	-9.87
3	8842407	-9.76
4	155792345	-9.64
5	54811489	-9.64
6	54844164	-9.64
7	45861272	-9.59
8	54841777	-9.54
9	56295120	-9.52

Table 1: Identified potetntial inhibitors and their binding affinities.

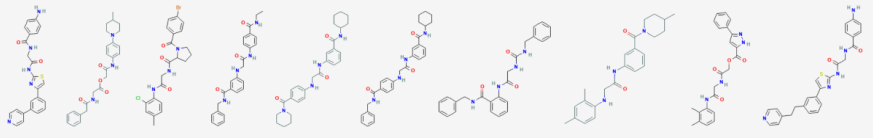

Figure 1: 2D structures of compounds with lowest binging energy (ordered).

Molecules with the lowest binding affinity tend to share common structural features: relatively long chains with either aromatic or non-aromatic rings at both ends, nitrogen atoms within the chain, and oxygen atoms as side groups. These characteristics likely facilitate binding to the target protein through hydrophobic interactions and hydrogen bonding.

None of the top-selected compounds have associated BioAssay data available in PubChem, suggesting they are not established inhibitors of TrmD or any other target. Among these, we identified molecule 155792348 as our lead compound. It was included in our screening due to its structural similarity to molecule 71724899, a confirmed lignad of bacterial TrmD, whose crystal complex with *Mycobacterium tuberculosis* TrmD is available under PDB accession number 6JOF. For comparison, molecule 71724899 yielded a binding affinity of -9.21 kcal/mol in our docking study (using PDB 5WYQ as the target). Additionally, natural TrmD ligand, S-adenosylmethionine (SAM), showed a binding affinity of -11.30 kcal/mol under the same docking conditions. Here, we present the predicted complex of compound 155792348 with TrmD in figure 2, alongside an overlay of the same complex with SAM positioned as observed in its crystal structure.

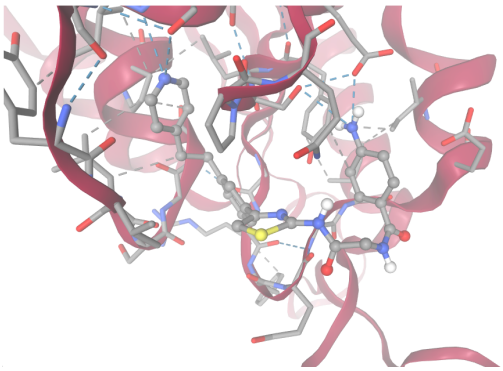

Figure 2: Predicted docking pose of our lead. Displayed are its interactions with TrmD (blue dashed line - hydrogen bonds, grey line - hydrophobic interactions).

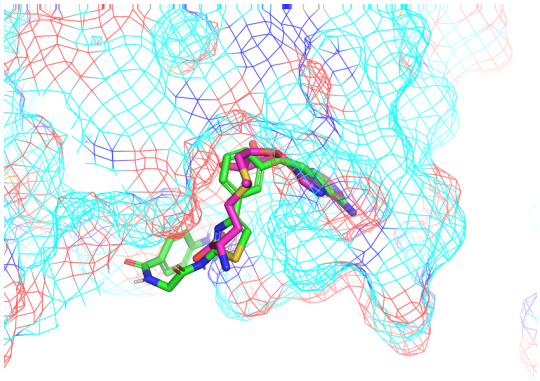

Figure 3: Our hit (green) and SAM (pink) positioned based on the crystallographically determined binding site.

As observed, the selected molecule fits into the binding pocket, occupying the position of SAM. Its longer chain enhances stability by forming hydrogen bonds, facilitated by the nitrogen atom at its terminal end.

To further assess these findings, we conducted docking simulations with a TrmD from diffrrent species species (*Acinetobacter baumannii*, 7MYS). Notably, our lead compound consistently ranked among the top hits across both targets, achieving a favorable binding energy of -9.93 kcal/mol in the M. tuberculosis TrmD. However, a correlation coefficient of 0.78 between the docking results for the two targets underscores the importance of careful target selection and the potential limitations of relying solely on a single target for screening campaigns. A more comprehensive approach, involving broader screening efforts, may be necessary to identify versatile molecules and prevent the inadvertent exclusion of potential hits that might exhibit strong binding affinity towards specific TrmD homologs.

Nonetheless, in our case selected compound bound strongly two both homologous TrmDs. It seems likely, that it could serve as a good starting point in designing a good inhibitor for TrmD.

4 Discussion

The identification of bacterial TrmD inhibitors holds significant promise for developing novel antimicrobial agen ts, particularly in combating multidrug-resistant pathogens. In our work we employed a comprehensive computational approach, leveraging known ligands for virtual screening and molecule generation methods to explore potential candidates for TrmD inhibition. Molecular docking with AutoDock Vina revealed several compounds with high binding affinities to the TrmD active site, indicating their potential efficacy. Notably, the lead compound 155792348 exhibited favorable binding characteristics, closely resembling the established TrmD ligand71724899. However, our in silico findings would need to be thouroughly validated experimentally. It's important to remember that high binding affinity alone does not guarantee a successful inhibitor. While low binding affinities may not always translate to effective inhibition, studies have shown that lower affinities can still be sufficient for achieving clinically useful bactericidal effects, as demonstrated in (Hafeez, Zafar Paracha, and Adnan 2024).

Our work highlights the critical considerations in inhibitor design, including meticulous target protein selection, appropriate tool and screening library selection, and careful parameter optimization. Even slight adjustments can significantly impact the results.

We believe our approach has identified a promising lead compound that warrants further optimization and experimental investigation as a potential TrmD inhibitor, ultimately contributing to the development of novel antimicrobial therapies.

References

Hou, Ya-Ming, Ryuma Matsubara, Ryuichi Takase, Isao Masuda, and Joanna I. Sulkowska. 2017. "TrmD". In *RNA Modification*, 89–115. Elsevier. https://doi.org/10.1016/bs.enz.2017.03.003.

Zhong, Wenhe, Kalyan Kumar Pasunooti, Seetharamsing Balamkundu, Yee Hwa Wong, Qianhui Nah, Vinod Gadi, Shanmugavel Gnanakalai, et al. 2019. "Thienopyrimidinone Derivatives That Inhibit Bacterial TRNA (Guanine37-N1)-Methyltransferase (TrmD) by Restructuring the Active Site with a Tyrosine-Flipping Mechanism". *Journal of Medicinal Chemistry* 62 (17). https://doi.org/10.1021/acs.jmedchem.9b00582.

Christian, Thomas, Caryn Evilia, Sandra Williams, and Ya-Ming Hou. 2004. "Distinct Origins of TRNA(m1G37) Methyltransferase". *Journal of Molecular Biology* 339 (4). https://doi.org/10.1016/j.jmb.2004.04.025.

Goto-Ito, Sakurako, Takuhiro Ito, and Shigeyuki Yokoyama. 2017. "Trm5 And TrmD: Two Enzymes from Distinct Origins Catalyze the Identical TRNA Modification, m1G37". *Biomolecules* 7 (1). https://doi.org/10.3390/biom7010032.

Ito, Takuhiro, Isao Masuda, Ken-ichi Yoshida, Sakurako Goto-Ito, Shun-ichi Sekine, Se Won Suh, Ya-Ming Hou, and Shigeyuki Yokoyama. 2015. "Structural Basis for Methyl-Donor–Dependent and Sequence-Specific Binding to TRNA Substrates by Knotted Methyltransferase TrmD". *Proceedings of the National Academy of Sciences* 112 (31). https://doi.org/10.1073/pnas.1422981112.

Wilkinson, Andrew J., Nicola Ooi, Jonathan Finlayson, Victoria E. Lee, David Lyth, Kathryn S. Maskew, Rebecca Newman, et al. 2023. "Evaluating the Druggability of TrmD, a Potential Antibacterial Target, through Design and Microbiological Profiling of a Series of Potent TrmD Inhibitors". *Bioorganic Medicinal Chemistry Letters* 90 (June). https://doi.org/10.1016/j.bmcl.2023.129331.

Eberhardt, Jerome, Diogo Santos-Martins, Andreas F. Tillack, and Stefano Forli. 2021. "AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings". *Journal of Chemical Information and Modeling* 61 (8). https://doi.org/10.1021/acs.jcim.1c00203.

Dedrick, Rebekah M, Bailey E Smith, Madison Cristinziano, Krista G Freeman, Deborah Jacobs-Sera, Yvonne Belessis, A Whitney Brown, et al. 2022. "Phage Therapy of Mycobacterium Infections: Compassionate Use of Phages in 20 Patients With Drug-Resistant Mycobacterial Disease". *Clinical Infectious Diseases* 76 (1). https://doi.org/10.1093/cid/ciac453.

Vlasov, Sergiy, Hanna Severina, Olena Vlasova, Oleksandr Borysov, Pavlo Shynkarenko, Olga Golovchenko, Yulian Konechnyi, and Victoriya Georgiyants. 2023. "Synthesis, Docking Study and Antimicrobial Activity Evaluation of Pyridyl Amides of Thieno[2,3-d]Pyrimidine-4-Carboxylic Acid". *ScienceRise: Pharmaceutical Science*, no. 5(45) (October). https://doi.org/10.15587/2519-4852.2023.286008.

Kralj, Sebastjan, Marko Jukič, and Urban Bren. 2023. "Molecular Filters in Medicinal Chemistry". *Encyclopedia* 3 (2). https://doi.org/10.3390/encyclopedia3020035.

Öztürk, Hakime, Elif Ozkirimli, and Arzucan Özgür. 2016. "A Comparative Study of SMILES-Based Compound Similarity Functions for Drug-Target Interaction Prediction". *BMC Bioinformatics* 17 (1). https://doi.org/10.1186/s12859-016-0977-x.

Gómez-Bombarelli, Rafael, Jennifer N. Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, Ryan P. Adams, and Alán Aspuru-Guzik. 2018. "Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules". *ACS Central Science* 4 (2). https://doi.org/10.1021/acscentsci.7b00572.

Cao, Nicola De, and Thomas Kipf. 2022. "MolGAN: An Implicit Generative Model for Small Molecular Graphs". https://arxiv.org/abs/1805.11973.

Bakan, Ahmet, Lidio M. Meireles, and Ivet Bahar. 2011. "ProDy: Protein Dynamics Inferred from Theory and Experiments". *Bioinformatics* 27 (11). https://doi.org/10.1093/bioinformatics/btr168.

Grosse-Kunstleve, Ralf W., Nicholas K. Sauter, Nigel W. Moriarty, and Paul D. Adams. 2002. "TheComputational Crystallography Toolbox: Crystallographic Algorithms in a Reusable Software Framework". *Journal of Applied Crystallography* 35 (1). https://doi.org/10.1107/s0021889801017824.

Hafeez, Sidrah, Rehan Zafar Paracha, and Fazal Adnan. 2024. "Designing of Fragment Based Inhibitors with Improved Activity against E. Coli AmpC -Lactamase Compared to the Conventional Antibiotics". *Saudi Journal of Biological Sciences* 31 (1). https://doi.org/10.1016/j.sjbs.2023.103884.