Molecular Dynamics IN PROTO-NOOS: Comparative Analysis of EcDHFR-NADPH-Ligand Comple? Dynamics for Selected Trimethoprim Analogues

1.Abstract

Escherichia coli dihydrofolate reductase (EcDHFR) is a well-established antibacterial target, and trimethoprim serves as the reference compound for inhibitor design. Ligand recognition by EcDHFR cannot, however, be reliably described by docking alone, as it depends on protein dynamics, NADPH occupancy, ligand-residue contacts, and conformational changes within the M20 loop.

In this work, we present a comparative analysis of EcDHFR–ligand complexes for trimethoprim analogues selected as EcDHFR inhibitors by the PROTO-NOOS system. PROTO-NOOS enables the identification of promising compounds but does not encompass in-depth dynamic or energetic analysis; the proposed approach fills this gap. The EcDHFR–TMP–NADPH system is treated as the reference for candidate evaluation. PyMOL is used for preparation and quality control of input structures for molecular dynamics.

Molecular dynamics simulations in GROMACS are used to assess binding mode stability, protein and ligand RMSD, M20 loop RMSF, persistence of intermolecular contacts, hydrogen bonds, and pocket solvent accessibility. This work extends PROTO-NOOS with a chemical, dynamic, and energetic layer, supporting more rigorous selection of antibacterial compounds.

2. Introduction

The primary challenge in designing ecDHFR inhibitors is moving from docking to a lasting biological effect. While PROTO-NOOS excels at proposing candidate molecules, molecular dynamics determines their actual therapeutic potential. Our motivation is to answer critical questions: Is the binding stable? We use GROMACS to verify if docking poses survive thermal fluctuations and protein vibrations. Are there steric clashes? We investigate whether the hydrophobic "tails" of novel compounds create structural tensions that would hinder efficacy. How does the M20 loop respond? We analyze whether the candidates can maintain the M20 loop in a closed conformation, which is essential for effective enzyme inhibition.

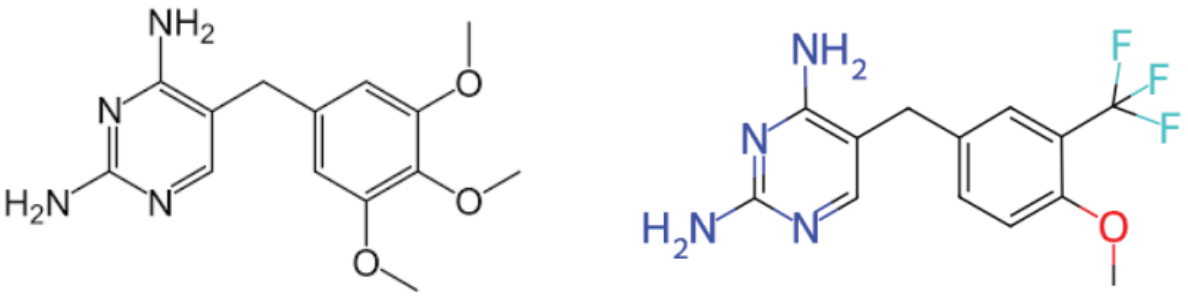

Fig. 1 From left: trimethoprim, CHEMBL277391

CHEMBL277391 is a trimethoprim-like candidate containing a 2,4-diaminopyrimidine pharmacophore. This polar ring was analyzed as the key recognition motif for Asp27 in the EcDHFR active site The substituted aromatic tail, including the CF3 group, was evaluated as a hydrophobic/steric element that may influence pocket fitting, ligand stability, and M20-loop response during MD simulation.

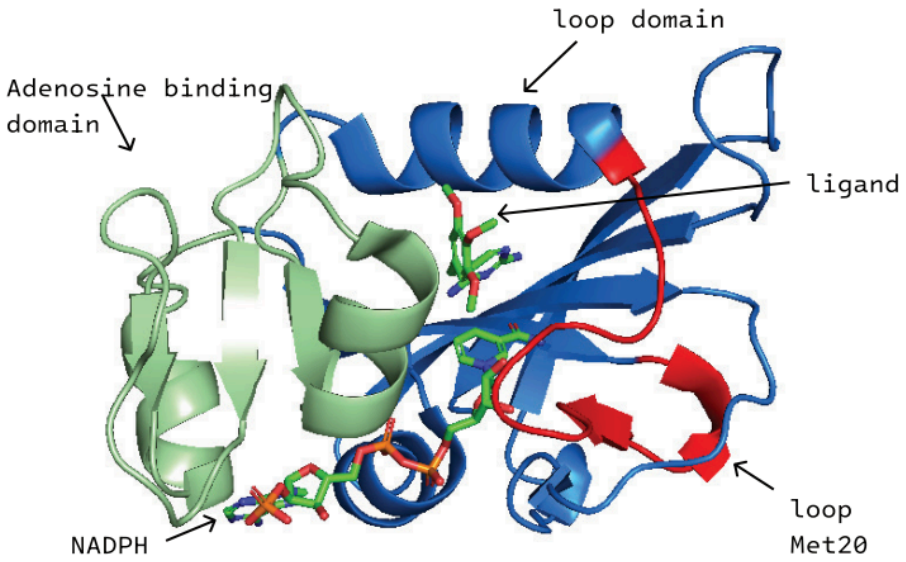

Fig. 2 Structural overview of the EcDHFR–ligand–NADPH complex.

The figure presents Escherichia coli dihydrofolate reductase in cartoon representation, with the ligand and NADPH cofactor shown as sticks. The labeled regions indicate the spatial organization of the complex: the ligand occupies the active-site pocket close to NADPH, while the M20 loop is positioned near the catalytic region and may influence active-site closure and ligand recognition. The loop domain and adenosine-binding domain form the main structural framework surrounding the binding site and contribute to cofactor positioning and inhibitor stabilization.

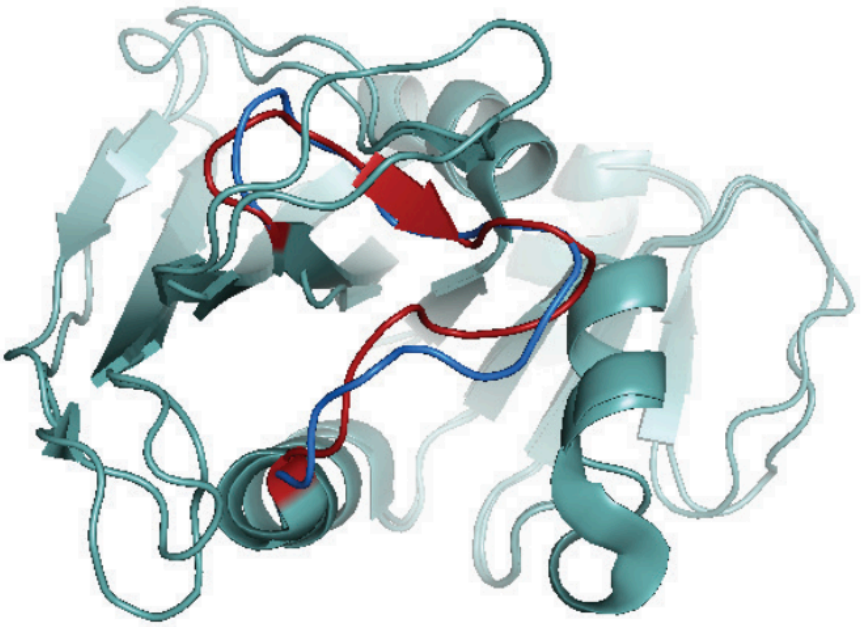

Fig. 3

M20 loop dynamics in EcDHFR are highly relevant for ligand recognition and binding-site organization. Static docking captures only a single structural pose and does not account for loop rearrangements, thermal fluctuations, or cofactor-dependent active-site plasticity. Using GROMACS molecular dynamics, we assess whether candidate ligands maintain a closed-like M20-loop geometry or whether they promote increased loop mobility and reduced binding-site.

3. Materials and Methods

3.1 Ligand Generation and Screening (PROTO-NOOS)

De novo compounds in the PROTO-NOOS pipeline are generated by REINVENT4, a reinforcement learning-based molecular generative model that leverages a prior model trained on ChEMBL 25—a comprehensive database of bioactive molecules with drug-like properties. The generation process employs a sophisticated reinforcement learning scoring function that combines multiple optimization criteria: QED (Quantitative Estimate of Drug-likeness) at 30% weight, AutoDock Vina docking scores at 60% weight, and custom toxic group filters at 10% weight. To ensure structural diversity among generated compounds, the system implements a Scaffold Similarity diversity filter with configurable parameters for bucket size, minimum score, and minimum similarity thresholds. The raw REINVENT4 output is initially written to summarize CSV files and subsequently post-processed into the pipeline. After traversing the complete multi-stage pipeline (Stages 1-6), all compounds undergo final aggregation and comprehensive scoring in Stage 7, which produces the definitive final_output.csv file run_scoring.py:688-736 . This final output provides a holistic evaluation of each compound through multiple metrics: the composite GUB_score (ranging from 0-1 with higher values indicating better performance),

the overall GUB_rank position, pareto_rank for Pareto-optimal front identification, classification_6b for bactericidal/bacteriostatic/resistant categorization, detailed quality flags documenting stage-specific quality assessments and drop reasons, and an availability_mask indicating which evidence blocks—penetration, binding, molecular dynamics, and systems pharmacology—are available for each compound.

3.2 Molecular Dynamics Simulation Protocol

Molecular dynamics (MD) simulations were performed using the GROMACS software package. The protein structure was parameterized with the AMBER99SB-ILDN force field. Topologies and parameters for the cofactor (NADPH) and the investigated ligand were generated based on the General AMBER Force Field (GAFF) using the ACPYPE tool. The assembled protein-cofactor-ligand complex was placed in the center of a cubic simulation box, maintaining a minimum distance of 1.2 nm from the box edges, and subsequently solvated. To neutralize the system, an appropriate number of K+ and Cl- ions were added.

The relaxation of the system began with energy minimization utilizing the steepest descent algorithm. This was followed by a two-step equilibration process in the NVT and NPT ensembles, during which position restraints were applied to the protein atoms to allow for solvent relaxation. The temperature of the system was maintained at 310 K using the V-rescale thermostat. Throughout all simulation stages, the lengths of bonds involving hydrogen atoms were constrained using the LINCS algorithm, allowing for an integration time step of 2 fs. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method. Finally, an unrestrained production MD simulation was carried out for 1 ns to analyze the stability of the complex.

4. Results and Discussion

4.1 Protonation states

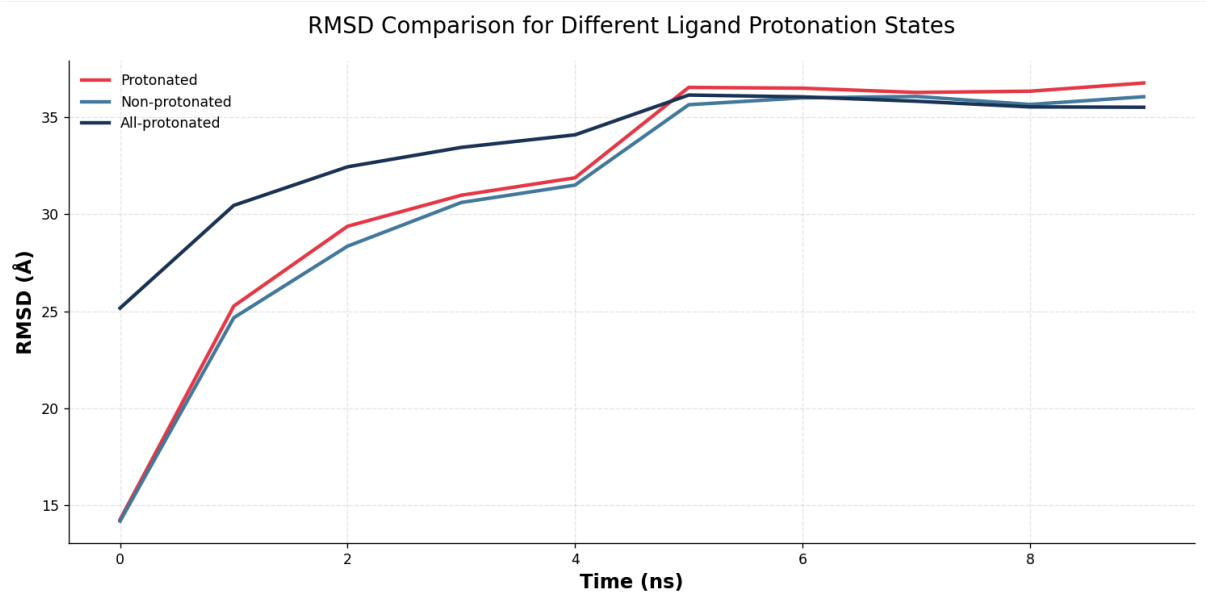

Fig. 4 Impact of Protonation States on Binding Dynamics

Root Mean Square Deviation (RMSD) was utilized as a primary indicator to evaluate the overall structural stability and convergence of the simulated systems. The structural response of the EcDHFR binding pocket exhibited an exceptional sensitivity to the ligand's electrostatic surface potential, as demonstrated by the rapid divergence of the non-protonated, protonated, and all-protonated trajectories within the first 4 ns of the production run. The simulation highlighted significant protonation-dependent conformational shifts dictating the initial relaxation behavior of the inhibitor. The all-protonated variant exhibited a distinct kinetic pattern characterized by a gradual transition and a delayed increase in RMSD. This slower binding-pose relaxation suggests that the additional positive charges on the ligand may facilitate transient, stabilizing electrostatic interactions with adjacent acidic residues, most notably the catalytic Asp27, before reaching equilibrium. While the different protonation states displayed distinct early-stage equilibration paths, all trajectories eventually converged toward a shared structural plateau after approximately 5–6 ns. Following this convergence, the RMSD values stabilized and fluctuated within a narrow, well-defined range of 4.0–5.0 Å. The achievement of this prolonged plateau phase confirms the absence of drastic macromolecular conformational changes and underscores the thermodynamic stability of the verified docking poses within the EcDHFR-NADPH active site architecture.

3.2 Comparative analysis between TMP and protonated CHEMBL277391

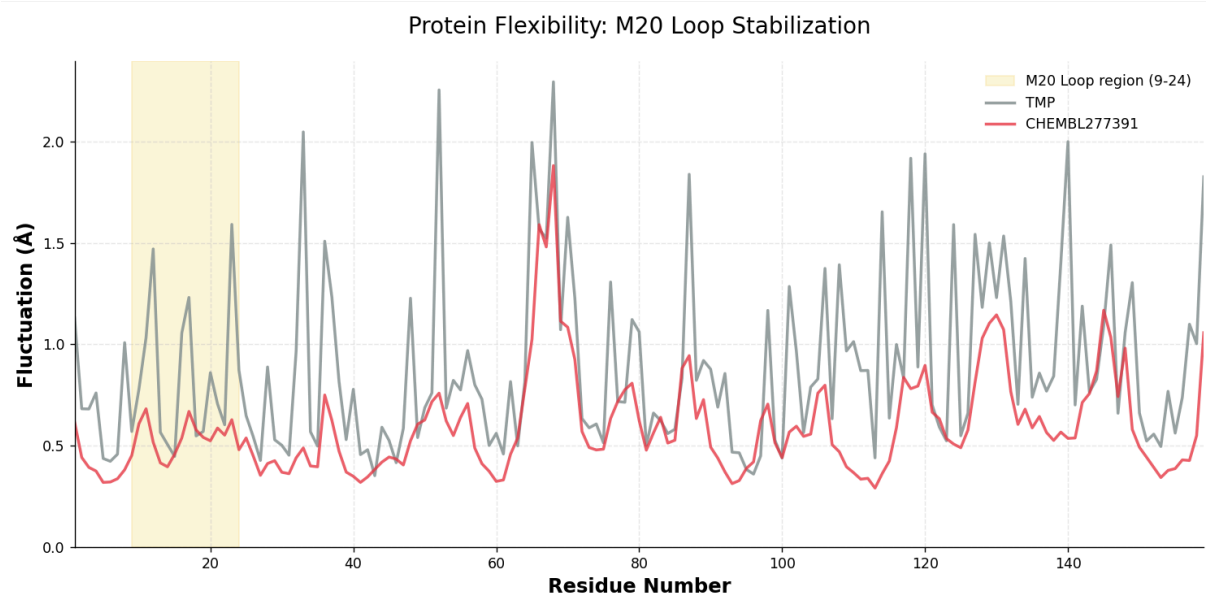

Fig. 5 RMSF Analysis & M20 Loop Dynamics

Analysis of the Root Mean Square Fluctuation (RMSF) provides insights into the dynamic behavior of the EcDHFR complex, concerning the catalytic M20 loop. As illustrated in Figure 5, the complex bound with CHEMBL277391 exhibits a distinctly lower local RMSF within the M20-loop region when compared to the reference drug, Trimethoprim (TMP). This measurable reduction in fluctuation suggests that the novel candidate significantly restricts the overall mobility of the loop within the simulated environment. The M20 loop naturally retains a degree of inherent flexibility. The observed structural variations are interpreted primarily as localized internal motions rather than large-scale conformational transitions between fully open and occluded states. This hypothesis of structural containment is further corroborated by monitoring key spatial metrics, specifically the distance between residues Asn18 and His45. Throughout the entire 10 ns molecular dynamics trajectory, this crucial inter-residue distance consistently remains within the ~7–8 Å range, a measurement strongly characteristic of a stable, closed-like conformational regime.

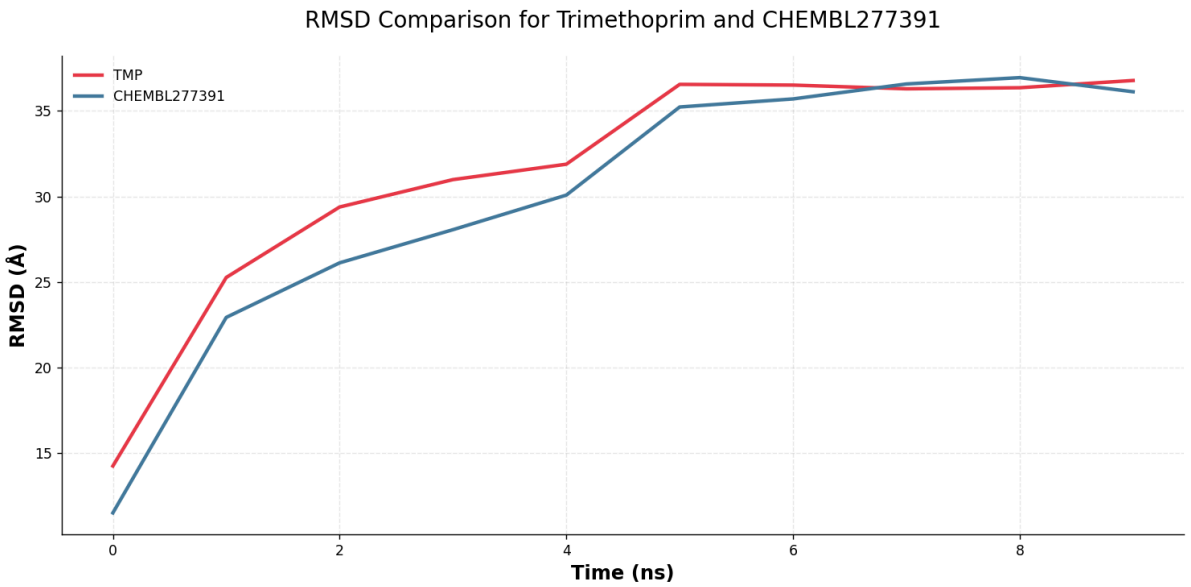

Fig. 6 Comparative RMSD analysis of CHEMBL277391 and Trimethoprim

The simulation demonstrates that CHEMBL277391 maintains a more constrained conformation during the equilibration phase (first 5 ns) compared to TMP, ultimately reaching a stable structural plateau consistent with the reference inhibitor.

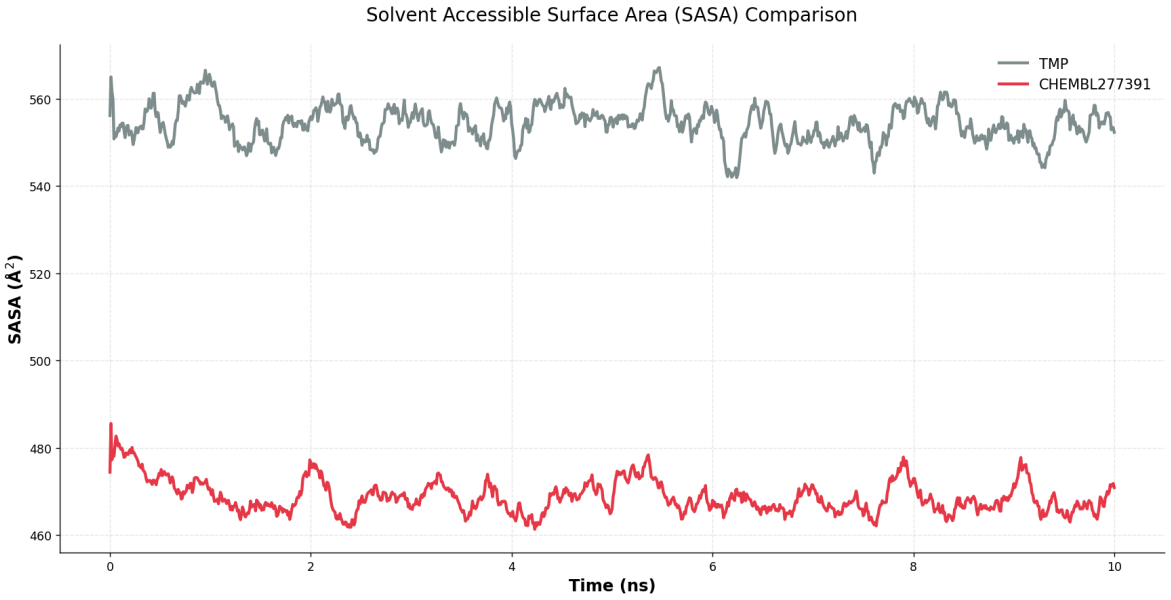

Fig. 7 Solvent Accessible Surface Area (SASA) Analysis

CHEMBL277391 consistently exhibits lower SASA values (avg. 470$Å^2$) compared to Trimethoprim (avg. 555$Å^2$), indicating a more compact protein-ligand complex The significant reduction in surface exposure (approx. 85$Å^2$) suggests that the candidate ligand buries deeper into the binding pocket, effectively displacing solvent molecules and enhancing hydrophobic stabilization Together with RMSF analysis, this supports good structural compatibility of CHEMBL27739

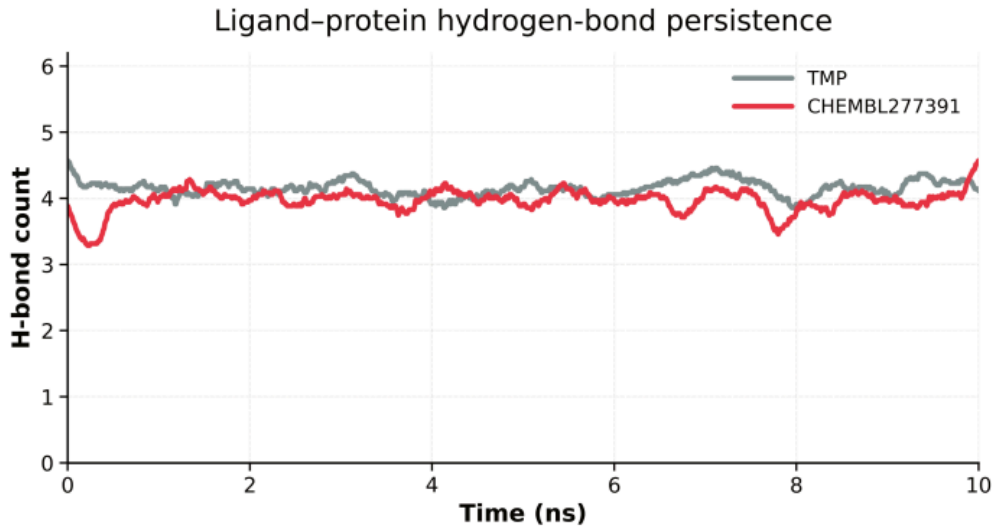

Fig. 8 Ligand–Protein Hydrogen-Bond Persistence

Both ligands establish a robust and stable hydrogen-bonding network within the ecDHFR active site, maintaining an average of ~4.0–4.2 bonds throughout the 10 ns trajectory LIGP (CHEMBL277391) demonstrates comparable binding persistence to the reference TMP, effectively mimicking its essential interaction pattern The persistence of these interactions, together with reduced local M20-loop flexibility and lower ligand solvent exposure, supports stable binding pocket retention of CHEMBL277391 during the 10 ns simulation.

5. Conclusion

The 10 ns molecular dynamics simulations successfully validated the inhibitory potential of CHEMBL277391, a novel trimethoprim analogue generated by the PROTO-NOOS pipeline. Our computational investigation yielded several key findings regarding the stability and conformational dynamics of the EcDHFR-ligand-NADPH complex:

1. **Active-Site Compatibility:** CHEMBL277391 remained highly compatible with the EcDHFR–NADPH active-site architecture throughout the 10 ns MD simulation, confirming its structural stability and supporting its use as a highly promising TMP-like candidate for further *in silico* and experimental validation.
2. **Pharmacophore Recognition and Pocket Fit:** The 2,4-diaminopyrimidine-like pharmacophore provides a robust and plausible motif for Asp27-directed polar recognition, maintaining a stable hydrogen-bonding network (averaging ~4.0 bonds). Concurrently, the CF3-substituted aromatic tail appears well-tolerated in the binding pocket on the simulated timescale, contributing to tighter hydrophobic shielding and a reduced Solvent Accessible Surface Area (SASA).
3. **M20 Loop Dynamics:** M20-loop analysis indicates reduced local mobility (RMSF) relative to TMP, effectively shielding the active site. Importantly, this stabilization does not result in rigid locking; the

loop retains functional local fluctuations while remaining predominantly within a closed-like conformational regime necessary for enzyme inhibition.

4. **Protonation-Dependent Stability:** The ligand's protonation state strongly affected binding-pose stability. While the singly protonated variant remained fully compatible with the active site, the all-protonated state showed partial displacement from the binding pocket, suggesting that excessive ligand charge and electrostatic repulsion may destabilize pocket retention.

Future Work: To further rigorously evaluate the binding kinetics and thermodynamic stability of CHEMBL277391 and other top-ranking PROTO-NOOS candidates, future studies will employ PLUMED-driven metadynamics. This advanced enhanced-sampling technique will enable the reconstruction of ligand dissociation free-energy profiles and the precise identification of metastable binding states.

Bibliography

1. Kordus, S. L.; Baughn, A. D. Revitalizing antifolates through understanding mechanisms that govern susceptibility and resistance. MedChemComm (2019).
2. Giladi, M.; Altman-Price, N.; Levin, I.; Levy, L.; Mevarech, M. FolM, A New Chromosomally Encoded Dihydrofolate Reductase in Escherichia coli. Journal of Bacteriology (2003).
3. Estrada, A.; Wright, D. L.; Anderson, A. C. Antibacterial Antifolates: From Development through Resistance to the Next Generation. Cold Spring Harbor Perspectives in Medicine (2016).
4. Sehrawat, R.; Rathee, P.; Khatkar, S.; Akkol, E. K.; Khayatkashani, M.; Nabavi, S. M.; Khatkar, A. Dihydrofolate Reductase (DHFR) Inhibitors: A Comprehensive Review. Current Medicinal Chemistry (2022).
5. Srinivasan, B.; Rodrigues, J. V.; Tonddast-Navaei, S.; Shakhnovich, E.; Skolnick, J. Rational Design of Novel Allosteric Dihydrofolate Reductase Inhibitors Showing Antibacterial Effects on Drug-Resistant Escherichia coli Escape Variants. ACS Omega (2017).
6. Wan, Q.; Bennett, B. C.; Wilson, M. A.; Kovalevsky, A.; Langan, P.; Howell, E. E.; Dealwis, C. Toward resolving the catalytic mechanism of dihydrofolate reductase using neutron and ultrahigh-resolution X-ray crystallography. Proceedings of the National Academy of Sciences (PNAS) (2014).
7. Babu, C. S.; Lim, C. Sensitivity of Functional Loop Conformations on Long-Range Electrostatics: Implications for M20 Loop Dynamics in E. coli Dihydrofolate Reductase. Journal of Chemical Theory and Computation (2020).