



STRUCTURE-BASED PRIORITIZATION OF HETEROCYCLIC CANDIDATES TARGETING HDAC8 THROUGH MOLECULAR DOCKING, MOLECULAR DYNAMICS, AND ADMET PROFILING

Shireen Nadhem Saleh *

Department of Pharmaceutical Chemistry, College of Pharmacy, Tikrit University, Salah Al-Din, Iraq.

* Corresponding Author: Shireen Nadhem Saleh

ORCID Details

Author name: <https://orcid.org/0009-0003-6623-4467>

Open Access Research Journal of Chemistry and Pharmacy, 2026, 09(01), 072–089

Publication history: Received on 17 August 2026; revised on 24 September 2026; accepted on 25 September 2026

Article DOI: <https://doi.org/10.53022/oarjcp.2026.9.1.0027>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

The epigenetic target of growing interest in the discovery of anticancer drugs is the histone deacetylase 8 (HDAC8). The paper tested 14 heterocyclic (L1-L14) designs with an integrated computational workflow that included molecular docking, molecular dynamics (MD), ADME profiling and toxicity prediction. The docking of human HDAC8 (PDB ID: 1T69) resulted in L2 being the highest ranked compound by score (-5.954 kcal/mol), with L9 close behind (-5.881 kcal/mol) along with a predicted interaction with the catalytic Zn²⁺ center (2.31 Å). According to this binding geometry, L9 was prioritized for a 100-ns MD simulation. Following initial structural relaxation, the pathway facilitated further association of L9 to HDAC8 pocket by limited ligand-pocket separation, long-term hydrophobic interactions, and repeated yet discontinuous Zn378-O contact geometry. The geometric definition of the metal-contact analysis was based on the non-bonded Zn²⁺ representation used in the simulation and thus was perceived as recurrent metal-proximal interaction instead of a permanent zinc-chelation state. ADME analysis revealed that there were no Lipinski violations throughout the series, but solubility, polarity, and predicted gastrointestinal absorption among the candidates varied. L9 showed medium predicted solubility and low lipophilicity, as well as high topological polar surface area and low predicted gastrointestinal absorption. Various endpoint-specific alerts were produced in toxicity modeling of L9. Altogether, the combined findings support L9 as a high-ranking, mechanistically informative HDAC8-targeting candidate to be optimized through the designed series using structure-guided optimization.

Keywords: HDAC8; Molecular docking; molecular dynamics; ADMET; Heterocyclic compounds; Zinc-proximal interaction; In silico drug discovery.

1 INTRODUCTION

Epigenetic regulation is integral to the control of gene expression, chromatin organization, and cellular identity, and its disruption is closely associated with malignant transformation, tumor progression, and therapeutic resistance. Histone deacetylases (HDACs) are major regulators of this process, catalyzing the removal of acetyl groups from lysine residues on histone and non-histone proteins and thereby influencing transcriptional programs, cell-cycle control, apoptosis, DNA-damage responses, and other cancer-relevant pathways. These functions have established HDACs as important targets in anticancer drug discovery [1].

Among the zinc-dependent HDACs, histone deacetylase 8 (HDAC8) is a class I isoform with structural and biological features that distinguish it from other family members. Besides histones, HDAC8 deacetylates non-histone targets, such as chromosome cohesion, transcriptional regulation, stress responses and cell survival (SMC3, p53, and estrogen-related receptor alpha). Imbalanced HDAC8 expression or activity has been implicated in a variety of malignancies, such as hematological and solid tumors, which are reasons to keep studying this isoform as a therapeutically relevant molecular target [2,4]. The main issue with HDAC medicinal chemistry is whether isoselectivity can be of any significance. Even though HDAC inhibition with broad-spectrum can result in pharmacological activity, cross-inhibition of up to multiple HDAC isoforms can complicate identification of target-specific effects and raise the risk of off-target liabilities. As a result, recent HDAC8-based discovery programs are more aimed at taking advantage of isoform-based variations in active-site geometry, channel structure, surface recognition, and zinc-proximal interactions to achieve more selective molecular recognition [3,4]. Structurally, HDAC8 has a catalytic Zn^{2+} ion at the bottom of a small mostly hydrophobic channel linking the catalytic site to the protein surface. Productive inhibitor binding is regulated by zinc binding and also channel geometry of the linker region, occupancy of the hydrophobic channel and complementary interactions at the rim of the channel. An experimentally resolved reference of this binding environment, as well as a convenient structural model to compare with the ligand of interest, is the crystal structure of human HDAC8 complexed with suberoylanilide hydroxamic acid (SAHA; vorinostat), deposited as PDB ID 1T69 [4,5]. Classical HDAC inhibitors tend to be defined by zinc-binding group-linker-cap pharmacophore, but modern HDAC8 discovery has gone beyond conventional hydroxamate-based designs to other zinc-binding motifs and chemotypically varied structures. With heteroatoms in place, ring electronics, substitution patterns, and three-dimensional geometry, heterocyclic scaffolds are especially advantageous in this regard, as they can be adjusted to modify metal-proximal interactions, hydrogen bonding, aromatic contacts as well as physicochemical properties. The heterocyclic library of the N/S-containing structurally diversified library studied in this case was thus deemed to be appropriate to comparative structure-based analysis with HDAC8 [4,6]. Early-stage prioritization: Before experimental testing, computational methods can assist in prioritizing the evidence by combining complementary layers of evidence.

Molecular docking can generate plausible binding orientations and identify interaction patterns within a target site, but docking scores remain scoring-function estimates rather than direct measurements of affinity or inhibition [7]. Molecular dynamics (MD) simulations provide an additional dynamic perspective by examining ligand retention, protein flexibility, conformational adaptation, and persistence of intermolecular contacts over time [8]. Physicochemical, pharmacokinetic, drug-likeness, and toxicity predictions can further identify developability constraints that may not be apparent from binding-site analysis alone [9,10]. Rather than assigning priority from docking-score rank alone, this workflow was structured to integrate metal-proximal binding geometry, dynamic pocket retention, and developability predictions so that closely scoring candidates could be differentiated on mechanistic grounds.

Accordingly, the present study applied an integrated in silico workflow to fourteen designed heterocyclic derivatives (L1-L14) targeting human HDAC8. Molecular docking against the SAHA-bound HDAC8 crystal structure (PDB ID 1T69) was combined with detailed binding-interaction analysis, followed by 100-ns MD simulation of the prioritized complex. SwissADME and ProTox 3.0 were used to complement the structure-based analysis with physicochemical, pharmacokinetic, drug-likeness, and predicted toxicity profiling. The overall objective was to establish a coherent structure-based prioritization of the designed series and identify candidates suitable for subsequent medicinal-chemistry optimization.

1.1 Integrated Computational Workflow

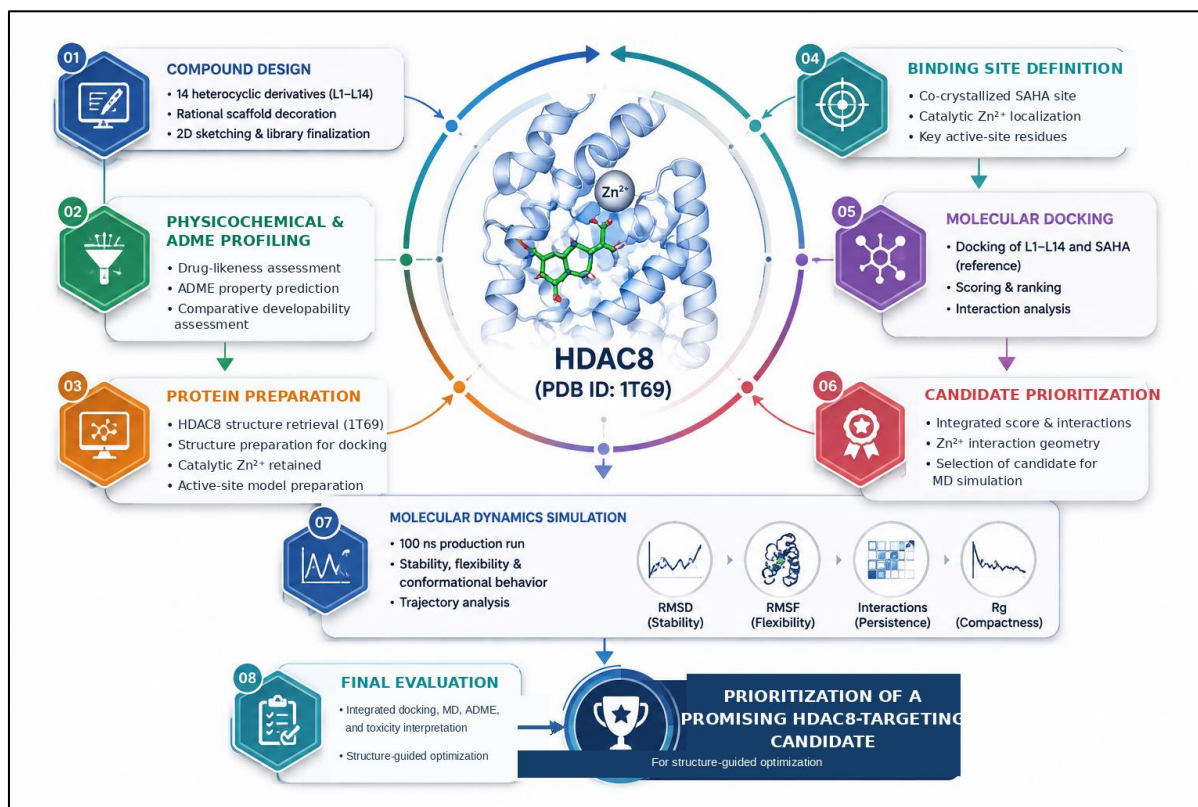
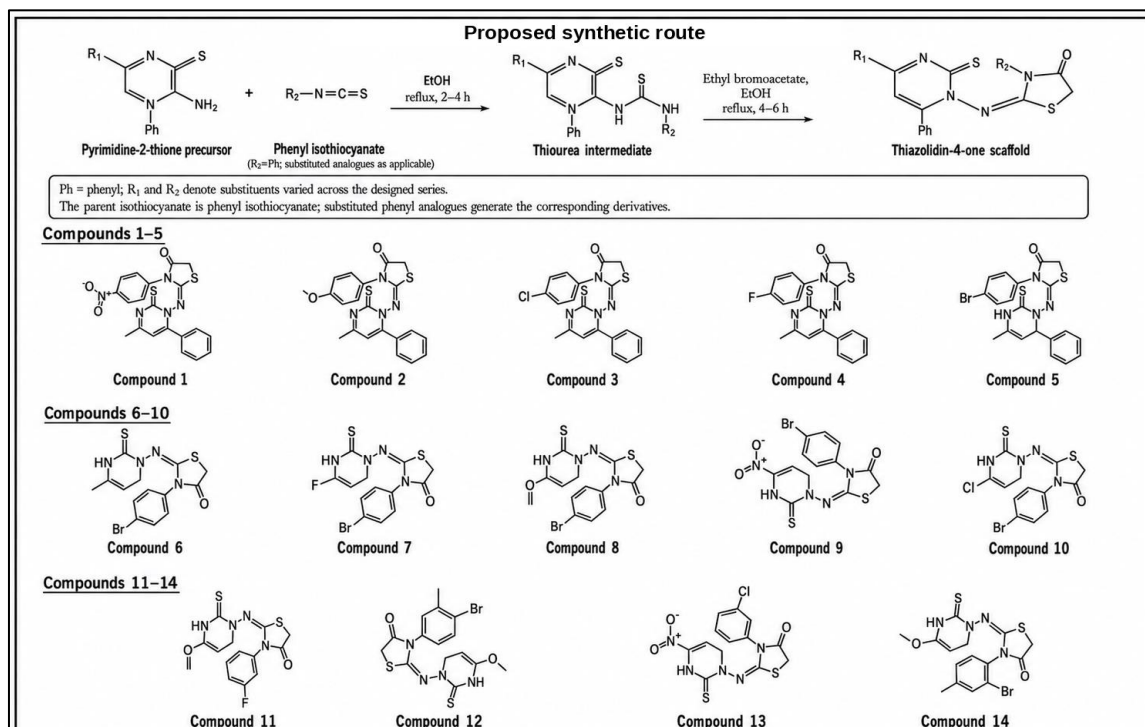


Figure 1: Integrated computational workflow for structure-based assessment and prioritization of the investigated HDAC8-targeting candidates.

1.2 Designed Compound Library and Structural Diversification



Scheme 1: Proposed synthetic route and designed structures of the target thiazolidin-4-one derivatives (compounds 1-14).

2 COMPUTATIONAL METHODS

2.1 Ligand Library

A library of fourteen designed heterocyclic candidates, designated L1-L14, was organized into three structurally related series according to the substitution patterns summarized in Scheme 1. Compound numbering in Scheme 1 corresponds directly to the ligand identifiers used throughout the computational workflow (Compound 1 = L1 through Compound 14 = L14). The molecular structures were encoded using their corresponding SMILES representations, and identical compound identifiers were retained throughout the docking, ADME, toxicity, and molecular dynamics analyses. The systematic variation of peripheral substituents and heteroatom-containing motifs within related heterocyclic frameworks enabled comparative assessment of how structural diversification influences predicted HDAC8 recognition and developability.

2.2 Target Selection and Structural Model

Human histone deacetylase 8 (HDAC8) was selected as the molecular target, and the experimentally determined human HDAC8 structure deposited as PDB ID 1T69, in complex with suberoylanilide hydroxamic acid (SAHA; vorinostat), was used for the structure-based investigation [5]. The co-crystallized SAHA binding environment served as the structural reference to define the catalytic channel as well as interpret the predicted poses of L1-L14. Special focus was placed on the conformational adaptation of ligand in the channel, as well as its interactions with major active site residues and geometry with respect to catalytic Zn^{2+} center, which is a hallmark mechanistic element of zinc-dependent HDAC recognition [4,5].

2.3 Molecular Docking

Repair of the crystallographic HDAC8 structure was done with the Protein Repair and Analysis Server (PRAS) [21]. PDB2PQR was used to assign protonation states at pH 7.4 in combination with PROPKA [22,23]. Water molecules farther than 4.5 Å from the co-crystallized ligand were eliminated and the receptor was then subjected to the MOE QuickPrep protocol. In the energy minimization, atoms that were above 8 Å in distance with the ligand were fixed, and the hydrogen atoms within the binding-site area were free to move. The process of minimization was repeated until the RMS gradient had dropped to $0.1 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ with the QuickPrep restraints in place. The ligands L1-L14 were made by the addition of hydrogen atoms, assigning them corresponding protonation and tautomeric states, and energy minimization was followed by docking. Molecular docking was carried out using Molecular Operating Environment (MOE) 2019 [26]. The Triangle Matcher algorithm was used to place the ligands, and the process was then initialised using London dG. Each ligand had 20 poses created. The poses generated were optimized with the Rigid Receptor protocol and were refined and re-scored with the GBVI/WSA dG scoring function, then the two highest-ranked poses were kept and examined in more detail. Vorinostat (SAHA) was included as the reference ligand. Pose evaluation considered the docking score, orientation within the catalytic channel, hydrogen-bonding interactions, aromatic and hydrophobic contacts, electrostatic interactions, and geometry relative to the catalytic Zn^{2+} center. The docking protocol was validated by redocking the co-crystallized reference ligand; the selected redocked pose reproduced the crystallographic binding mode with an RMSD of approximately 1.4 Å.

2.4 Candidate Prioritization

Candidate prioritization was based on an integrated interpretation of docking score and predicted binding geometry rather than numerical rank alone. L2 produced the most favorable docking score (-5.954 kcal/mol), whereas L9 combined a closely comparable score (-5.881 kcal/mol) with a short predicted metal-acceptor contact to Zn378 (2.31 Å). Because geometry relative to the catalytic Zn^{2+} center is mechanistically relevant to HDAC8 ligand recognition [4,5], L9 was selected for subsequent molecular dynamics analysis. This criterion defines the candidate selected for dynamic follow-up within the present computational workflow.

2.5 Physicochemical, Pharmacokinetic, and Drug-Likeness Assessment

The fourteen compounds were evaluated computationally for key physicochemical and pharmacokinetic descriptors relevant to early-stage drug discovery. The assessed parameters included molecular weight, hydrogen-bond acceptors and donors, topological polar surface area (TPSA), lipophilicity, predicted aqueous solubility, gastrointestinal absorption, blood-brain barrier permeability, P-glycoprotein substrate status, and drug-likeness criteria. These predictions were subsequently integrated with the docking results to provide a broader assessment of the developability of the investigated series. SwissADME was used for the physicochemical, pharmacokinetic, and drug-likeness evaluation [9].

2.6 Toxicity Prediction

Predicted toxicity profiles were evaluated for compounds L1-L14 using ProTox 3.0 [10]. The assessment included predicted median lethal dose (LD_{50}), toxicity class, hepatotoxicity, neurotoxicity, nephrotoxicity, respiratory toxicity, cardiotoxicity, carcinogenicity, mutagenicity, cytotoxicity, and additional mechanistic toxicity endpoints. The predictions of toxicity were seen as computational prioritization instruments as opposed to experimental toxicological analysis.

2.7 Molecular Dynamics Simulation

The docking-derived L9-HDAC8 complex was used as the starting point of all-atom molecular dynamics simulations with GROMACS 2025 [24]. The CHARMM36 force field was used to describe the protein [25]. The catalytic Zn^{2+} ion was held as non-bonded divalent ionic species with the standard CHARMM36 ion parameters; no explicit Zn topology or parameter file, bonded Zn model or cationic dummy-atom model was added. The formal charge of the Zn^{2+} was +2; its Lennard-Jones parameters were directly copied over to CHARMM36.

L9 was parameterized separately with SwissParam [28], generating the ligand topology and parameter files LIG.itp and LIG.par, which were included in the final GROMACS topology (topol.top). The complex was placed in a dodecahedral simulation box with a 10 Å separation between the solute and the box boundary and solvated with explicit TIP3P water molecules [27]. Sodium and chloride ions were added to neutralize the system and establish a final ionic concentration of 0.15 M NaCl. The solvated and ionized system was subjected to energy minimization, followed by 100 ps of NVT equilibration at 300 K and 100 ps of NPT equilibration at 300 K and 1 atm. A 100-ns production simulation was then conducted at 300 K and 1 atm.

Trajectory analysis used complementary global, local, and binding-site descriptors: protein C α RMSD, ligand RMSD in the protein-fitted frame, RMSD-based structural clustering using a 2.0 Å cutoff, radius of gyration (Rg), full-protein and ligand RMSF, ligand self-fit RMSD, pocket-focused RMSF, ligand-pocket center-of-mass (COM) distance, protein-ligand hydrogen-bond counts, and interaction-type persistence. For the pocket-focused RMSF assessment, the trajectory-based pocket analysis defined the ligand-proximal subset as residues within 5.0 Å of L9 in at least 90% of trajectory frames. This any-contact persistence criterion is distinct from the percentages reported in the interaction map, which quantify occupancies of individual interaction types; the two quantities were therefore not treated as numerically interchangeable. For the zinc-specific analysis, one oxygen atom of the L9 nitro group was monitored relative to Zn378. A frame was classified as Zn-O contact-positive when the Zn^{2+} -oxygen distance was ≤ 2.2 Å, and occupancy was calculated as (contact-positive frames $\div$ total analyzed frames) $\times$ 100. Because Zn^{2+} was represented as a non-bonded ion, this distance-based occupancy was interpreted as recurrent Zn-O proximity rather than as a measurement of coordination-bond strength, chelation thermodynamics, or electronic stabilization. The MD descriptors were interpreted collectively to distinguish early conformational relaxation from sustained late-time pocket association rather than relying on any single trajectory metric.

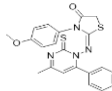
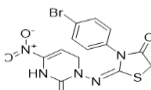
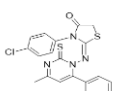
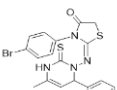
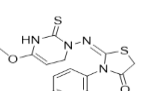
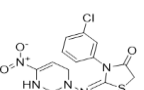
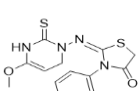
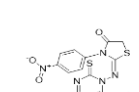
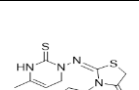
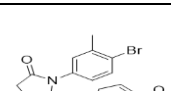
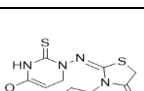
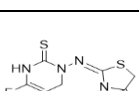
For direct comparison with the docking analysis, trajectory-derived residue labels were sequence-mapped to the crystallographic author numbering of PDB 1T69. The first modeled protein residue in 1T69 is Leu14; sequence-compatible active-site labels in the trajectory output therefore map with a +13 offset: Phe139→Phe152, Phe194→Phe207, Phe195→Phe208, Met261→Met274, Gly138→Gly151, His167→His180, Asp254→Asp267, Tyr293→Tyr306, and Zn365→Zn378. These sequence-compatible mappings were used for residue-specific cross-method interpretation, and all MD residue identifiers reported below use PDB 1T69 numbering [5].

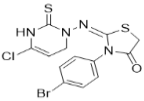
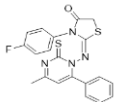
3 RESULTS AND DISCUSSION

3.1 Docking Validation, Ranking, and Binding Interaction Analysis

The docking protocol was first evaluated by redocking the co-crystallized reference ligand vorinostat (SAHA) into the HDAC8 active site of PDB ID: 1T69. The redocked pose reproduced the crystallographic binding orientation with an RMSD of approximately 1.4 Å, supporting the ability of the adopted protocol to recover the experimentally observed binding mode. This validation was used to support comparative pose interpretation; however, docking scores were not treated as experimental measures of binding affinity [7]. Across the fourteen investigated compounds, final docking scores ranged from -5.150 to -5.954 kcal/mol. L2 showed the most favorable numerical score (-5.954 kcal/mol), whereas L9 produced a closely comparable score (-5.881 kcal/mol) together with a short predicted metal-acceptor contact to Zn378 (2.31 Å). The reference ligand SAHA scored -5.230 kcal/mol and retained the expected short Zn378 interaction at approximately 1.95-1.97 Å. The minimal numerical difference between L2 and L9 (0.073 kcal/mol) was thus not overemphasised and the active-site geometry and docking score were combined and used to prioritize compounds based on their mechanistic relevance.

Table 1: Molecular docking scores and predicted favorable interactions of the studied compounds with HDAC8 (PDB ID: 1T69).

No.	Structure	Cmpd.	Docking score (kcal/mol)	Predicted favorable interactions
1		L2	-5.954	Gly151 (C-H bond); Phe152 (π -sulfur, π - π stacked); Met274 (π -sulfur); Lys33, His180, His143 (alkyl / π -alkyl)
2		L9	-5.881	Zn378 (metal-acceptor, 2.31 Å); Gly151 (H-bond, 2.92 Å); Phe208 (carbon H-bond); His143, Asp178, Asp267, His180 (attractive charge / π -cation); Phe152, Lys33, Tyr100 (alkyl)
3		L3	-5.863	Tyr100 (π -sulfur); Tyr306, Phe152 (π - π stacked); His180 (π -anion); Lys33, Met274, His143 (alkyl / π -alkyl)
4		L5	-5.857	Phe207 (π - π T-shaped); Met274 (π -sulfur); His180, Phe152, His143, Tyr100 (alkyl / π - π T-shaped)
5		L14	-5.779	His143 (carbon H-bond); His180, Phe208 (π -sulfur); Tyr100, Met274 (alkyl)
6		L13	-5.775	Tyr306 (H-bond, 2.22 Å); Asp101 (carbon H-bond); His180, Asp267 (π -cation); Zn378, Phe152 (attractive charge, 4.0-4.5 Å); Phe208 (π -sulfur); Met274 (alkyl)
7		L8	-5.766	His142 (carbon H-bond); Phe152 (π - π T-shaped); His180, Phe208 (π -sulfur); Lys33, Tyr100 (alkyl)
8		L1	-5.573	Asp101 (attractive charge / π -anion); Tyr100, Met274 (π -sulfur); Phe208 (π -alkyl); Tyr306, Phe152 (π - π stacked / T-shaped)
9		L6	-5.551	His180 (conventional H-bond); Asp101 (carbon H-bond); Phe152 (π - π stacked); His143, Phe208, Tyr100 (alkyl / π - π T-shaped)
10		L12	-5.434	His142 (carbon H-bond); His180, Phe152 (π - π T-shaped); Lys33, Tyr100 (alkyl)
11		L11	-5.399	His142, His143 (carbon H-bond); His180, Phe152, Phe208 (π -sulfur)
12		L7	-5.358	Gly206 (conventional H-bond); His142, Phe207, Met274 (alkyl); His180 (π -anion); Phe152 (π - π stacked); Phe208 (π -sulfur)

13		L10	-5.280	Zn378 (annotated metal-acceptor, 5.33 Å; too distant for short-range coordination); His180 (π -anion); Phe152 (π - π stacked); His142/143, Phe207 (alkyl)
14		L4	-5.150	Lys33 (conventional H-bond); Gly151 (carbon H-bond + halogen-F); His180 (π -anion); Phe152 (π - π stacked); Pro273, Met274 (alkyl)
Ref.	Reference ligand	Vorinostat (SAHA)	-5.230	Zn378 (metal-acceptor, 1.95-1.97 Å; reference chelation); Tyr306 (conventional H-bond, 2.60 Å); Phe208, Phe152 (alkyl / π -alkyl)

Interaction pattern through the series demonstrated repeated occupancy of the HDAC8 channel with Phe152/Phe208 and channel-lining residues Tyr100, Met274, His142/His143 and His180 repeatedly providing hydrophobic, aromatic or electrostatic stabilization. L2 scored the highest score due to positive channel contacts but did not exhibit a short direct interaction between Zn378 metal and acceptor in the pose analyzed. Conversely, L9 exhibited both competitive occupancy of channels and a 2.31 Å Zn378 contact, which is the most interesting zinc-oriented geometry of the designed compounds. An annotation of a Zn378 metal-acceptor interaction to L10 was also found, however the further distance of 5.33 Å was thought to be of a less supportive geometry to effective short-range coordination; L13 only exhibited a further Zn-associated electrostatic contact (around 4.0–4.5 Å). Based on this, L9 was selected for the 100-ns molecular dynamics analysis not based on score rank but on an integrated structural criterion. This prioritization is in line with the core importance of catalytic-metal interaction and channel identification in the design of HDAC8 inhibitors [4,5].

3.2 Detailed Binding-Mode Analysis and Docking Figures

Representative docking poses were studied to establish whether the score-based ranking was validated by a consistent predicted binding geometry in the HDAC8 catalytic channel. The interpretation was based on the orientation of the ligands, repetitive aromatic and hydrophobic interactions, hydrogen-bonding, and geometry of any annotated interaction between Zn378. Due to the predictive nature of docking, these interactions were viewed in comparison, instead of as experimental measures of affinity to bind or inhibit enzyme activity [4,5,7].

Vorinostat (SAHA) was used as the structural reference. Its docked pose showed short metal-acceptor contacts with Zn378 at approximately 1.95–1.97 Å, a conventional hydrogen bond with Tyr306 at 2.60 Å, and hydrophobic/ π -associated contacts involving Phe152 and Phe208. This reference pattern provided the benchmark for assessing whether the investigated compounds combined channel recognition with short-range zinc-oriented geometry.

L2, the highest score-ranked compound (–5.954 kcal/mol), was supported by predicted contacts involving Gly151, Phe152, Met274, Lys33, His180, and His143. No short direct Zn378 metal-acceptor contact was annotated for L2. It was therefore retained as the numerical docking leader, but its pose was not interpreted as evidence of zinc engagement.

L9 produced a closely comparable docking score (–5.881 kcal/mol) and, unlike L2, showed an annotated Zn378 metal-acceptor contact at 2.31 Å. Additional predicted interactions included a conventional hydrogen bond with Gly151 at 2.92 Å, a carbon hydrogen bond with Phe208, and electrostatic or hydrophobic contacts involving His143, Asp178, Asp267, His180, Phe152, Lys33, and Tyr100. The combination of a competitive score, channel contacts, and short zinc-oriented geometry formed the basis for selecting L9 for molecular-dynamics analysis.

L3 (–5.863 kcal/mol) and L5 (–5.857 kcal/mol) were retained as high-ranking comparators. L10 (–5.280 kcal/mol) was included specifically because the docking output annotated a Zn378 metal-acceptor interaction at 5.33 Å. Because this distance is substantially longer than the 2.31 Å contact observed for L9, L10 was used as a geometric contrast rather than as evidence of short-range zinc engagement. Figures 2 and 3 therefore summarize the representative binding modes required to explain the ranking and candidate-selection rationale without reproducing all fourteen docking maps.

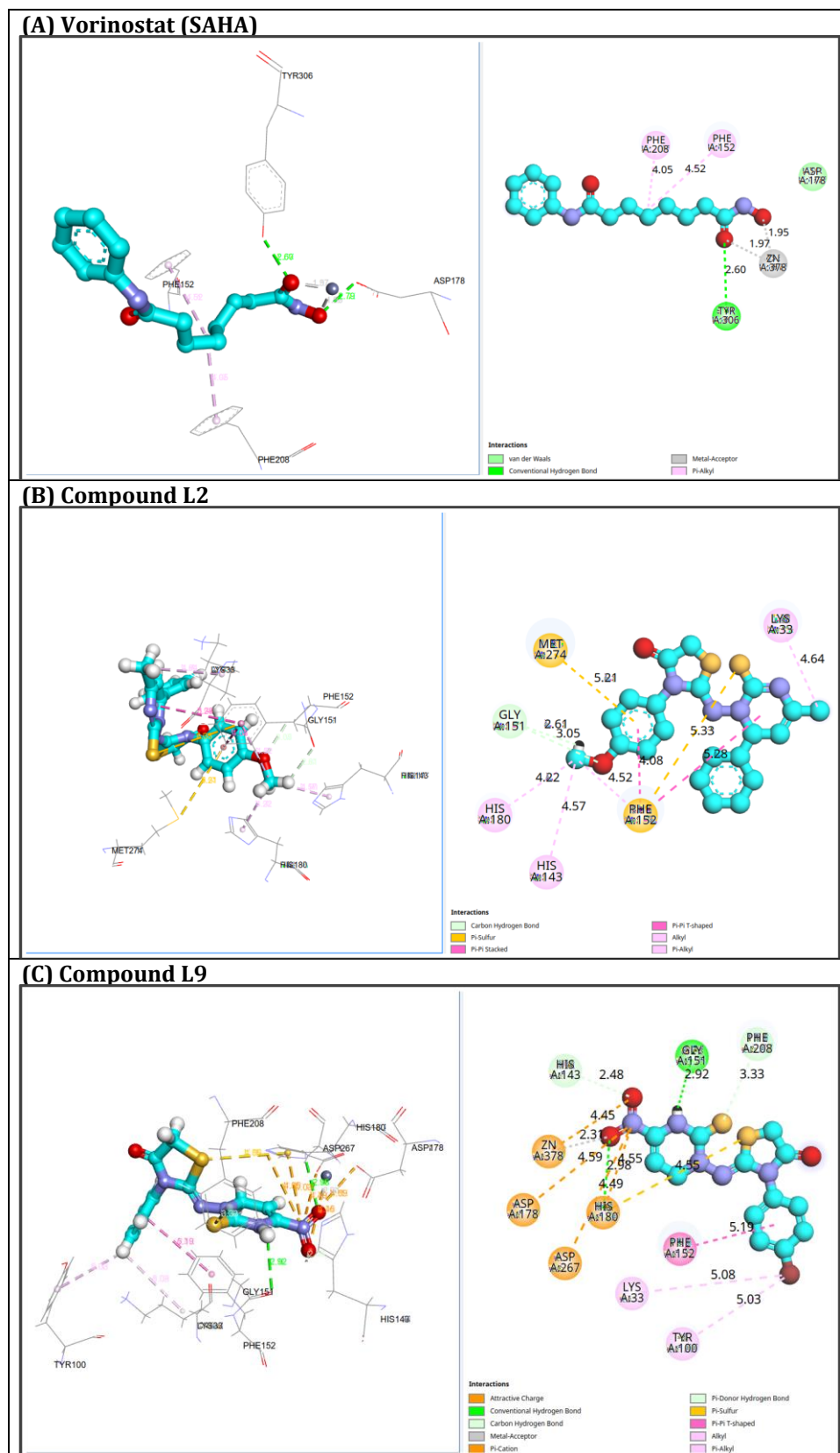


Figure 2: Representative docking poses and 2D interaction maps within the HDAC8 catalytic site (PDB ID: 1T69): (A) vorinostat (SAHA), (B) L2, and (C) L9.

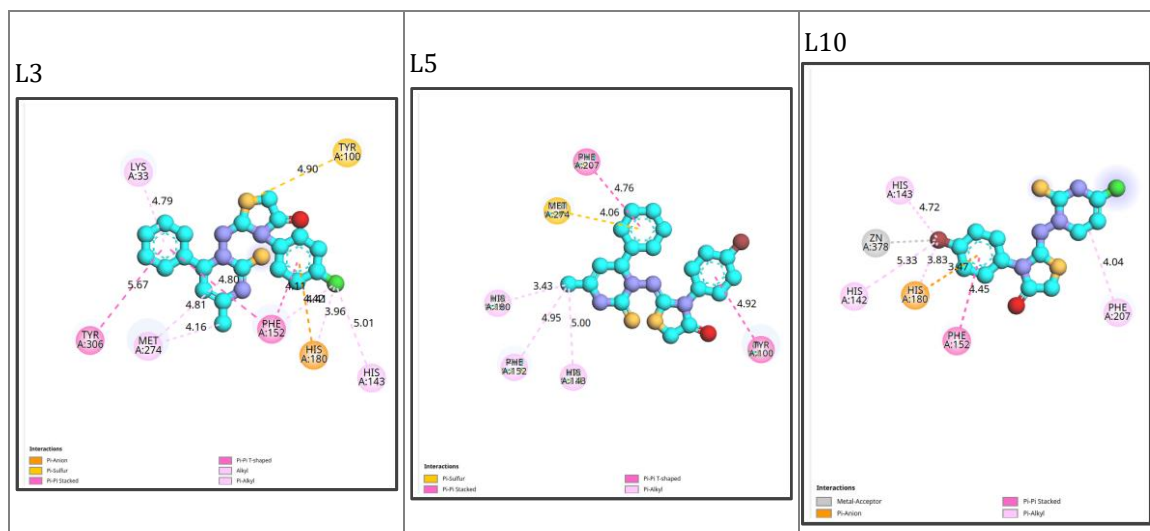


Figure 3: Representative 2D docking interaction maps for L3, L5, and L10. L3 and L5 are shown as high-ranking comparators; L10 is included to contrast its annotated Zn378 metal-acceptor distance (5.33 Å) with the shorter L9-Zn378 distance (2.31 Å).

3.3 ADME, Pharmacokinetic, and Drug-Likeness Profiling

SwissADME analysis showed that formal rule-based drug-likeness and predicted developability were not fully concordant across the series. All of the fourteen compounds met the Lipinski rule of five and had a bioavailability of 0.55. Nevertheless, significant variations in molecular size, lipophilicity, topological polar surface area (TPSA), aqueous solubility and predicted gastrointestinal absorption revealed that adherence to one rule set could not be equated with desirable exposure and oral performance. The results of the ADME, therefore, were viewed as a multidimensional prioritization layer and not a binary measure of drug-likeness [9,11-19]. There was an evident trend of solubility throughout the structural series. Compounds L1-L5 showed the most unfavorable aqueous-solubility profiles with ESOL LogS of -6.74 to -6.03 with all five compounds having a poor solubility profile. In comparison, the majority of the L6-L14 group were estimated to have a moderate solubility, whereas L11 had the best estimate (ESOL LogS = -3.59; soluble). This enhancement was generally paralleled by reduced consensus LogP of various of the subsequent analogues, which suggests that peripheral structural modification adjusted the equilibrium between hydrophobicity and aqueous behaviour. Compound developability-wise, this is a distinction of significance since a high lipophilicity and poor solubility may restrict compound handling and systemic exposure despite a desirable target-site docking [11,14]. The top candidate L9 had an informative ADME profile that is mixed. It had a molecular weight of 428.28 g/mol, a single consensus LogP of 1.65, ESOL LogS of -4.33 (moderately soluble), and a TPSA of 151.15 Å². The lower lipophilicity and better predicted solubility of L9 are good attributes relative to the more lipophilic members of the series.

Nevertheless, its high polar surface area coincided with low predicted gastrointestinal absorption and failure to satisfy the Veber, Egan, and Muegge filters. These findings identify permeability and exposure, rather than lipophilicity alone, as the principal predicted developability limitations of L9. Thus, L9 is best regarded as a mechanistically prioritized HDAC8-targeting candidate that may require physicochemical optimization before any expectation of favorable oral exposure can be justified [12,13,17,18].

None of the fourteen compounds was predicted to permeate the blood-brain barrier, and all were classified as non-substrates of P-glycoprotein. Predicted gastrointestinal absorption was high for L2-L8, L10-L12, and L14, but low for L1, L9, and L13, the three analogues with the highest TPSA values in the series.

This trend aligns with the physicochemical nature of the BOILED-Egg model where lipophilicity and polarity are key predictors of the predicted passive gastrointestinal absorption and brain penetration. These classifications are thus helpful in prioritization by comparison but must not be taken as alternatives to experimental measurements of permeability or pharmacokinetics [15]. Further structural scrutiny was offered by the use of medicinal-chemistry filters. None of the compounds raised a PAINS alarm, which alleviates the worry about the particular substructures that filter set depicts. Brenk alerts were however found throughout the series, mainly in conjunction with imine, thiocarbonyl, nitro or N-C-halo motifs. These alerts are rightly considered as warnings to take a closer look at chemical and assay results, not to assume they represent nonspecific activity or toxicity. The synthetic-accessibility scores were relatively small (3.48-4.29) indicating that there is no significant expected difference in the overall complexity of synthesis among the fourteen candidates [20].

Table 2A: Physicochemical and predicted aqueous-solubility properties of compounds L1-L14 obtained using SwissADME.

Cmpd	MW	HBA	HBD	TPSA	Cons. LogP	ESOL LogS	ESOL class
L1	437.49	5	0	153.70	2.99	-6.20	Poor
L2	422.52	4	0	117.11	3.75	-6.22	Poor
L3	426.94	3	0	107.88	4.43	-6.74	Poor
L4	410.49	4	0	107.88	4.19	-6.30	Poor
L5	473.41	2	1	105.33	4.10	-6.03	Poor
L6	397.31	2	1	105.33	2.76	-4.44	Moderate
L7	401.28	3	1	105.33	2.77	-4.42	Moderate
L8	413.31	3	1	114.56	2.39	-4.34	Moderate
L9	428.28	4	1	151.15	1.65	-4.33	Moderate
L10	417.73	2	1	105.33	2.95	-4.86	Moderate
L11	352.41	4	1	114.56	2.06	-3.59	Soluble
L12	427.34	3	1	114.56	2.74	-4.65	Moderate
L13	383.83	4	1	151.15	1.57	-4.02	Moderate
L14	427.34	3	1	114.56	2.74	-4.65	Moderate

MW: molecular weight (g/mol); HBA/HBD: hydrogen-bond acceptors/donors; TPSA: topological polar surface area ($\text{\AA}^2$); Cons. LogP: SwissADME consensus octanol/water partition coefficient; ESOL LogS: estimated aqueous solubility.

Table 2B. Predicted pharmacokinetic and drug-likeness characteristics of compounds L1-L14 obtained using SwissADME.

Cmpd	GI abs.	BBB	P-gp	Lipinski viol.	Veber	Egan	Muegge	BA score
L1	Low	No	No	0	No	No	No	0.55
L2	High	No	No	0	Yes	Yes	No	0.55
L3	High	No	No	0	Yes	Yes	No	0.55
L4	High	No	No	0	Yes	Yes	No	0.55
L5	High	No	No	0	Yes	Yes	Yes	0.55
L6	High	No	No	0	Yes	Yes	Yes	0.55
L7	High	No	No	0	Yes	Yes	Yes	0.55
L8	High	No	No	0	Yes	Yes	Yes	0.55
L9	Low	No	No	0	No	No	No	0.55
L10	High	No	No	0	Yes	Yes	Yes	0.55
L11	High	No	No	0	Yes	Yes	Yes	0.55
L12	High	No	No	0	Yes	Yes	Yes	0.55
L13	Low	No	No	0	No	No	No	0.55
L14	High	No	No	0	Yes	Yes	Yes	0.55

GI abs.: predicted gastrointestinal absorption; BBB: predicted blood-brain barrier permeation; P-gp: P-glycoprotein substrate status; BA score: bioavailability score. Lipinski violations are reported as counts; Veber, Egan, and Muegge entries indicate pass/fail status.

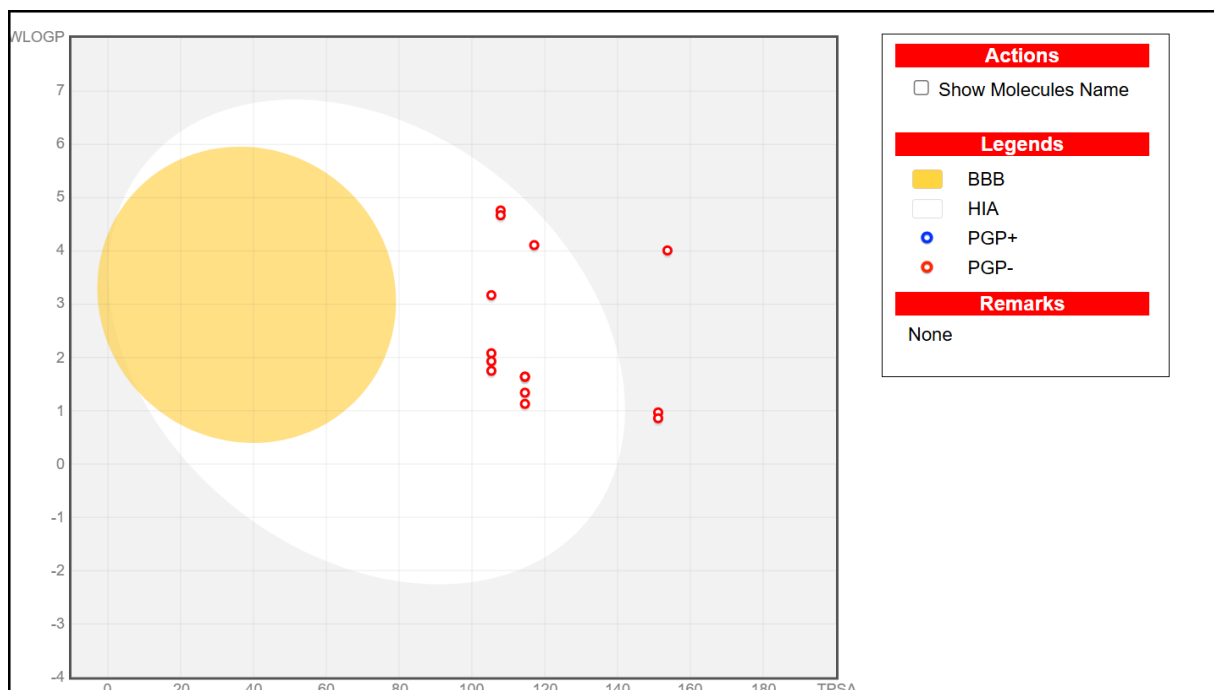


Figure 4: BOILED-Egg representation of compounds L1-L14 based on SwissADME predictions. The white region represents physicochemical space associated with a high probability of gastrointestinal absorption, whereas the yellow yolk denotes the region associated with blood-brain barrier permeation; red markers indicate compounds predicted to be P-glycoprotein non-substrates.

The BOILED-Egg representation was consistent with the tabulated pharmacokinetic predictions and highlighted the permeability trade-offs within the series. None of the compounds occupied the BBB-permeant region, whereas most analogues were positioned within or close to the region associated with gastrointestinal absorption. In contrast, L1, L9, and L13, which exhibited the highest TPSA values, were predicted to have low gastrointestinal absorption. All compounds were classified as P-glycoprotein non-substrates. Collectively, these findings support one of the key distinctions, namely that positive rule-based drug-likeness can be found alongside significant permeability restrictions, especially with highly polar analogues like L9 [15].

3.4 Predicted Toxicity and Safety Profiling

ProTox 3.0 was used as an adjunctive in silico safety-screening layer to determine the possible acute and endpoint-specific liabilities of the fourteen compounds [10]. The oral LD₅₀ values were predicted to be between 300 and 1125 mg/kg, with toxicity class 4 being predicted to contain 13 compounds, with L5 being predicted to have the lowest oral LD₅₀ (300 mg/kg). In comparison, the highest predicted LD₅₀ values in the series were found in L6-L9 (1125 mg/kg). These acute-toxicity estimates give a convenient comparative frame of reference, yet do not alone reflect the complete toxicity profile of a candidate and were thus combined with the respective organ- and mechanism-specific toxicity endpoints. A more differentiated pattern was observed in endpoint-level predictions. All fourteen compounds were predicted to be inactive in nephrotoxicity, cardiotoxicity, and cytotoxicity but respiratory toxicity was predicted to be active across the series. Most analogues were predicted to be neurotoxic but L1, L9, and L13 were not. Only L9 and L13 were predicted active for hepatotoxicity in the reported output. L1, L2, L6, L7, L9 and L13 were classified as carcinogenic, whereas L1, L9 and L13 were predicted to be mutagenic. The probabilities associated were mostly moderate and not conclusive, supporting the purpose of these outputs as prioritization indicators and not as evidence of experimentally determined toxicity.

The safety profile of L9 illustrates why acute-toxicity ranking and endpoint-specific risk should be considered separately. L9 showed one of the highest predicted oral LD₅₀ values (1125 mg/kg; class 4), and neurotoxicity, nephrotoxicity, cardiotoxicity, and cytotoxicity were predicted inactive. However, hepatotoxicity (0.52), respiratory toxicity (0.78), carcinogenicity (0.59), and mutagenicity (0.76) were predicted active. These alerts do not alter the structural rationale for selecting L9 for molecular-dynamics analysis, because that selection was based on binding geometry. Within the integrated computational profile, they instead identify toxicity-oriented structural features that can be considered during subsequent medicinal-chemistry optimization.

The predictive confidence indicators associated with the ProTox output were considered during interpretation. Structural similarity to reference compounds ranged from 36.02% to 39.49%, and the reported oral-toxicity accuracy

field was 23% across the series. Accordingly, the active/inactive labels in Tables 3A and 3B were used as comparative screening alerts within the integrated computational profile, particularly when probabilities were close to the classification threshold [10].

Table 3A. ProTox 3.0-predicted oral acute-toxicity and organ-toxicity endpoints for compounds L1-L14.

Cmpd	LD ₅₀ mg/kg	Class	Hepato	Neuro	Nephro	Resp.	Cardio
L1	1000	4	I (0.51)	I (0.64)	I (0.66)	A (0.67)	I (0.82)
L2	1000	4	I (0.54)	A (0.66)	I (0.61)	A (0.71)	I (0.75)
L3	500	4	I (0.54)	A (0.71)	I (0.67)	A (0.73)	I (0.88)
L4	1000	4	I (0.50)	A (0.70)	I (0.68)	A (0.74)	I (0.87)
L5	300	3	I (0.50)	A (0.66)	I (0.66)	A (0.70)	I (0.84)
L6	1125	4	I (0.55)	A (0.68)	I (0.63)	A (0.77)	I (0.82)
L7	1125	4	I (0.52)	A (0.68)	I (0.64)	A (0.76)	I (0.81)
L8	1125	4	I (0.53)	A (0.63)	I (0.54)	A (0.76)	I (0.72)
L9	1125	4	A (0.52)	I (0.58)	I (0.58)	A (0.78)	I (0.75)
L10	690	4	I (0.54)	A (0.69)	I (0.62)	A (0.76)	I (0.82)
L11	1000	4	I (0.53)	A (0.63)	I (0.55)	A (0.75)	I (0.72)
L12	690	4	I (0.53)	A (0.63)	I (0.54)	A (0.74)	I (0.73)
L13	680	4	A (0.51)	I (0.55)	I (0.57)	A (0.78)	I (0.75)
L14	1000	4	I (0.53)	A (0.63)	I (0.54)	A (0.74)	I (0.73)

Table 3B. ProTox 3.0-predicted carcinogenicity, mutagenicity, cytotoxicity, and model-similarity information for compounds L1-L14.

Cmpd	Carcino	Mutagen	Cyto	Similarity %	Reported accuracy %
L1	A (0.71)	A (0.83)	I (0.80)	36.43	23
L2	A (0.52)	I (0.53)	I (0.88)	37.03	23
L3	I (0.57)	I (0.63)	I (0.84)	37.83	23
L4	I (0.53)	I (0.61)	I (0.79)	37.08	23
L5	I (0.51)	I (0.58)	I (0.70)	37.11	23
L6	A (0.51)	I (0.56)	I (0.72)	39.49	23
L7	A (0.50)	I (0.55)	I (0.71)	38.35	23
L8	I (0.52)	I (0.55)	I (0.69)	37.66	23
L9	A (0.59)	A (0.76)	I (0.67)	36.02	23
L10	I (0.51)	I (0.56)	I (0.72)	39.17	23
L11	I (0.52)	I (0.55)	I (0.73)	37.83	23
L12	I (0.55)	I (0.55)	I (0.68)	36.9	23
L13	A (0.56)	A (0.77)	I (0.72)	37.44	23
L14	I (0.55)	I (0.55)	I (0.68)	37.16	23

A = predicted active; I = predicted inactive; values in parentheses are model probabilities. Similarity and reported accuracy values are reproduced from the ProTox output. All entries are computational predictions and should not be interpreted as experimentally confirmed toxicities.

When the docking, ADME, and toxicity layers are considered together, no single compound is uniformly optimal across all computational criteria. L2 remains the strongest score-ranked ligand, whereas L9 provides the most compelling short-range zinc-oriented docking geometry among the designed compounds and was therefore selected for dynamic follow-up. Simultaneously, the forecasted permeability and toxicity liabilities of L9 suggest not to refer to it as an optimized lead. Instead, it can be considered a mechanistically prioritized HDAC8-targeting candidate the binding behavior of which deserves dynamic description and whose developability and safety profile is an obvious focus of further optimization. This was followed by a molecular-dynamics analysis to explore the hypothesis that the docking-derived association of L9-HDAC8 was dynamically stabilized in the catalytic pocket.

3.5 Molecular Dynamics Behavior of the L9-HDAC8 Complex

Docking-derived L9-HDAC8 complex was characterized by a 100-ns molecular dynamics (MD) simulation following relaxation of the structure. The process of interpretation combined complementary protein and ligand displacement, structural clustering, local flexibility, ligand-pocket separation, hydrogen bonding, and interaction persistence measures instead of relying on any single measure of stability [8]. The trajectory thus gave a dynamic evaluation of the anticipated binding mode and its stability in the HDAC8 catalytic pocket.

3.5.1 Structural Relaxation and Conformational Behavior

The RMSD traces identified the early adjustment phase and the subsequent confined regime. Protein C α RMSD increased from 2–3 Å to 5–6 Å in the initial 15–20 ns and was then irregularly fluctuating around 6–7 Å without any long-lasting late-time drift. The fact that this plateau is fairly high is interpreted to be due to conformational reorganization and not a sign of a rigid complex. Even the ligand RMSD in the protein-fitted frame also became bounded at late times; this value encompasses positional/orientational change in the pocket along with internal ligand deformation and thus not to be confused with the ligand self-fit RMSD (about 0.10–0.20 nm) below. Neither of the two RMSD measures was, individually, used to make an inference on the presence of retention of the ligand.

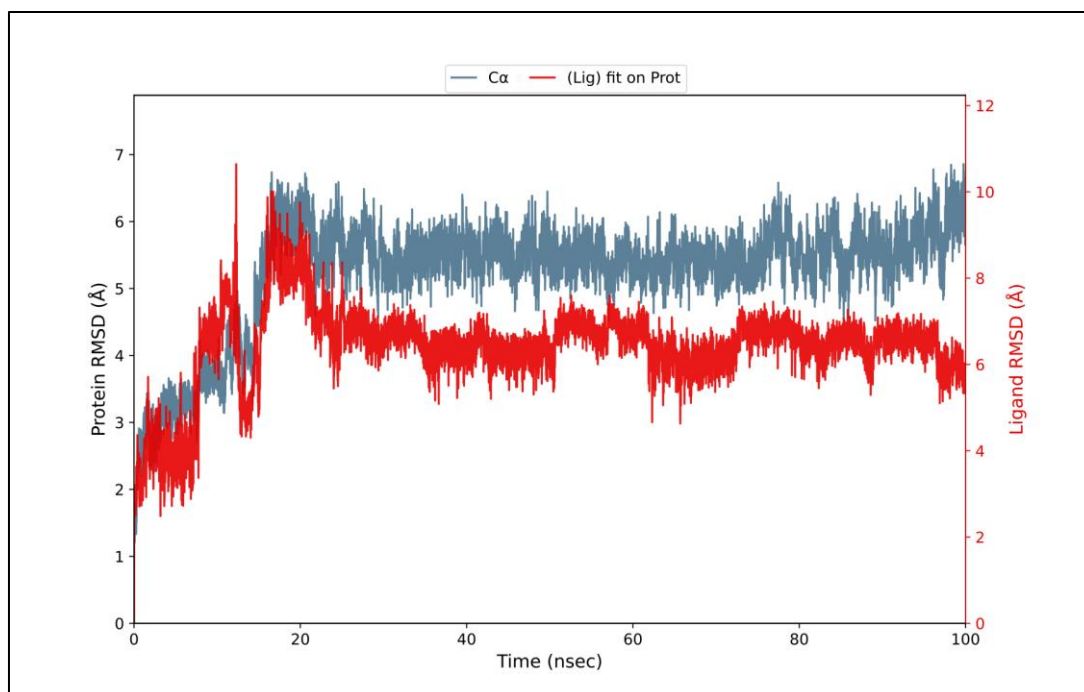


Figure 5: Profiles of protein C α and ligand RMSD throughout the L9-HDAC8 trajectory (100 ns).

Clustering with RMSD and a cutoff of 2.0 Å gave 21 clusters. The four biggest populations constituted around 13, 12, 10 and 8 percent of the trajectory showing that there was no dominant conformation in the simulation. The complex thus explored a variety of related substates, which were compared to the late-time RMSD and ligand-pocket separation instead of being treated as instability itself. In excellent agreement, as a complementary compactness trace, the radius-of-gyration (Rg) trace also evolved into a late-time fluctuating bounded regime following the first relaxation. Rg was thus employed as an auxiliary measure of overall compactness and not as an indicator of ligand binding or a rigid protein conformation.

3.5.2 Local Protein and Ligand Flexibility

Protein RMSF was typically low to moderate and most residues were within the range of 1–2 Å. The major flexible area was more or less associated with PDB residues 108–128, with RMSF values of approximately 3.5 Å. The localized flexible motion is evidenced by the limited distribution of this higher mobility, as opposed to uniformly high protein motion. Tyr306, His180, Phe208 and Met274 appeared in the ligand-proximal pocket subset with 100% presence in any frame by using the $\geq 90\%$ any-contact criterion within 5.0 Å, with Phe152 meeting the criterion at 90%. Their approximate RMSF values were 1.50 Å for Tyr306, 1.45 Å for His180, 0.62 Å for Phe152, 0.60 Å for Phe208, and 0.50 Å for Met274. Therefore, the ability to maintain pocket proximity was accompanied by residue-specific mobility, which is in accordance with an engaged but conformationally flexible binding environment. L9 had a localised internal motion. The majority of the ligand atoms varied between 0.2 and 0.9 Å; atom 17 was at around 1.4 Å, whereas atoms 22 and 23 were at around 1.1 and 1.0 Å respectively. The ligand-fit RMSD was widely spread between 0.10–0.20 nm. These data suggest that the scaffold maintained a consistent overall geometry with the peripheral positions selected experiencing more local adjustment.

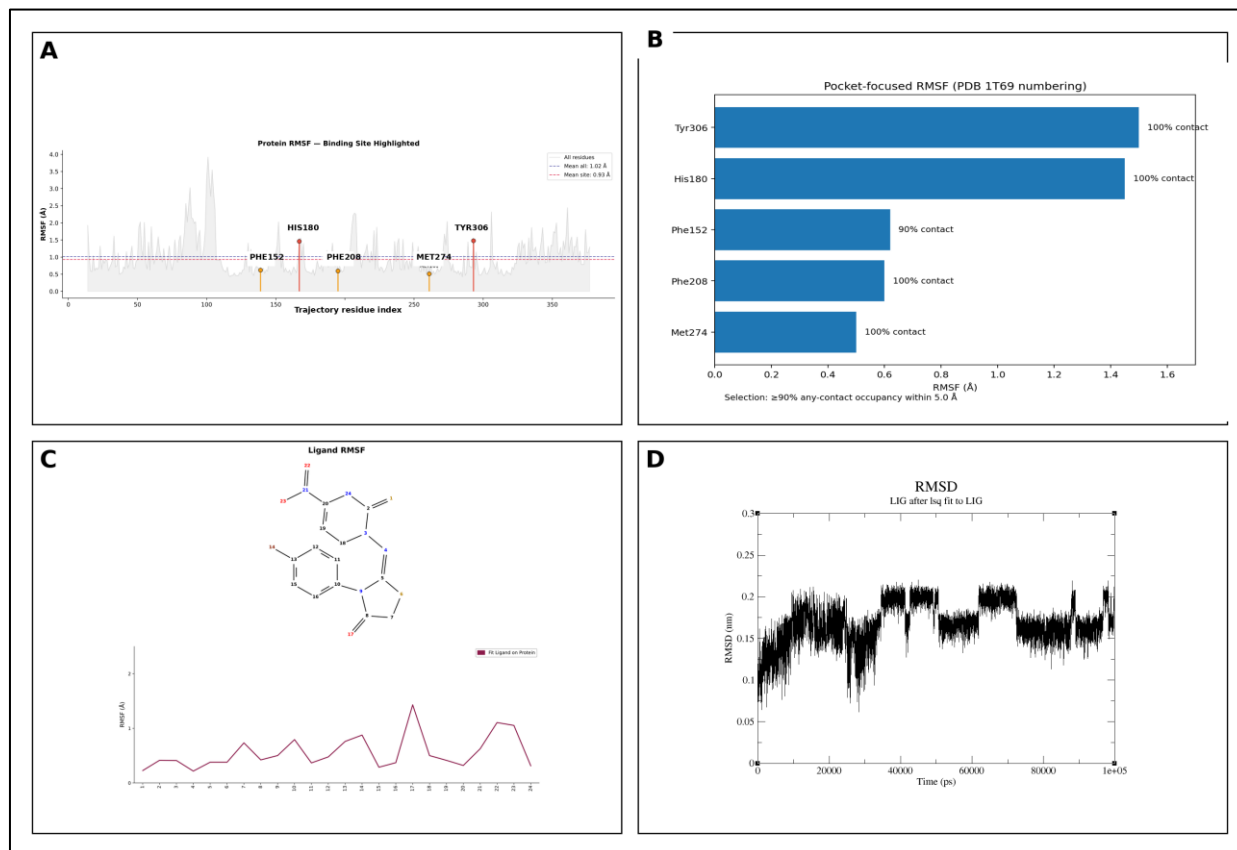


Figure 6. Flexibility analysis of the L9-HDAC8 complex: (A) protein RMSF with sequence-mapped pocket residues highlighted, (B) pocket-focused RMSF using the $\geq 90\%$ any-contact/5.0 Å subset, (C) ligand atom-wise RMSF, and (D) ligand self-fit RMSD. Residue labels are reported using PDB 1T69 numbering; the time in panel D is shown in ps (1×10^5 ps = 100 ns).

3.5.3 Binding-Site Retention and Interaction Persistence

The most direct geometrical measure of retention was the ligand-pocket center-of-mass (COM) distance. During the early trajectory, the separation was variable and generally occupied the 3.5–5 Å range. It constricted to a maximum size of 2–4 Å after about 20 ns, and had no later increase. This behavior aids in the persistence of space association of L9 with the HDAC8 pocket in the equilibrated part of the trajectory. Protein-ligand hydrogen bonding was discontinuous with two-thirds of the time involving none or one hydrogen bond and only infrequent two-hydrogen bond increases. The complex retained thus was not reliant on a single perpetual hydrogen bond, rather, the trajectory depicted an exchangeable interaction network where the hydrophobic, water-mediated, and metal-centered contacts played a role at various times.

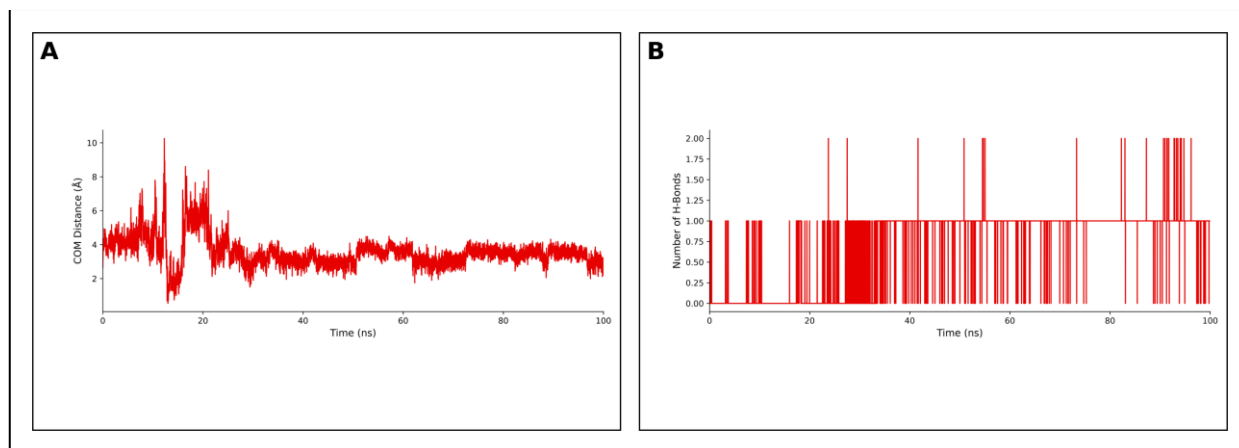


Figure 7: Binding-site retention during the 100-ns simulation: (A) L9-HDAC8 pocket center-of-mass distance and (B) protein-ligand hydrogen-bond count.

After sequence mapping to PDB 1T69 numbering, the interaction-type analysis identified hydrophobic occupancies of approximately 88% for Phe152, 34% for Phe207, 82% for Met274, and 96% for Phe208. Water-mediated or hydrogen-bond pathways involved Gly151, His180, Asp267, and Met274. The zinc-proximal analysis specifically monitored one nitro-group oxygen of L9 relative to Zn378. Using the predefined Zn-O distance criterion of ≤ 2.2 Å, contact-positive geometry was observed in approximately 32% of analyzed frames. These interaction-type occupancies are not expected to equal the $\geq 90\%$ any-contact values used for pocket-focused RMSF selection because they quantify specific interaction categories rather than simple ligand proximity. Importantly, the 32% value is a distance-defined occupancy derived from a non-bonded Zn^{2+} model; it therefore supports recurrent, non-continuous Zn378-O proximity but does not establish a persistent covalent or bonded coordination state. Together with the bounded ligand-pocket COM distance, the results support a hydrophobically anchored and dynamically exchanging pocket association with intermittent metal-proximal engagement.

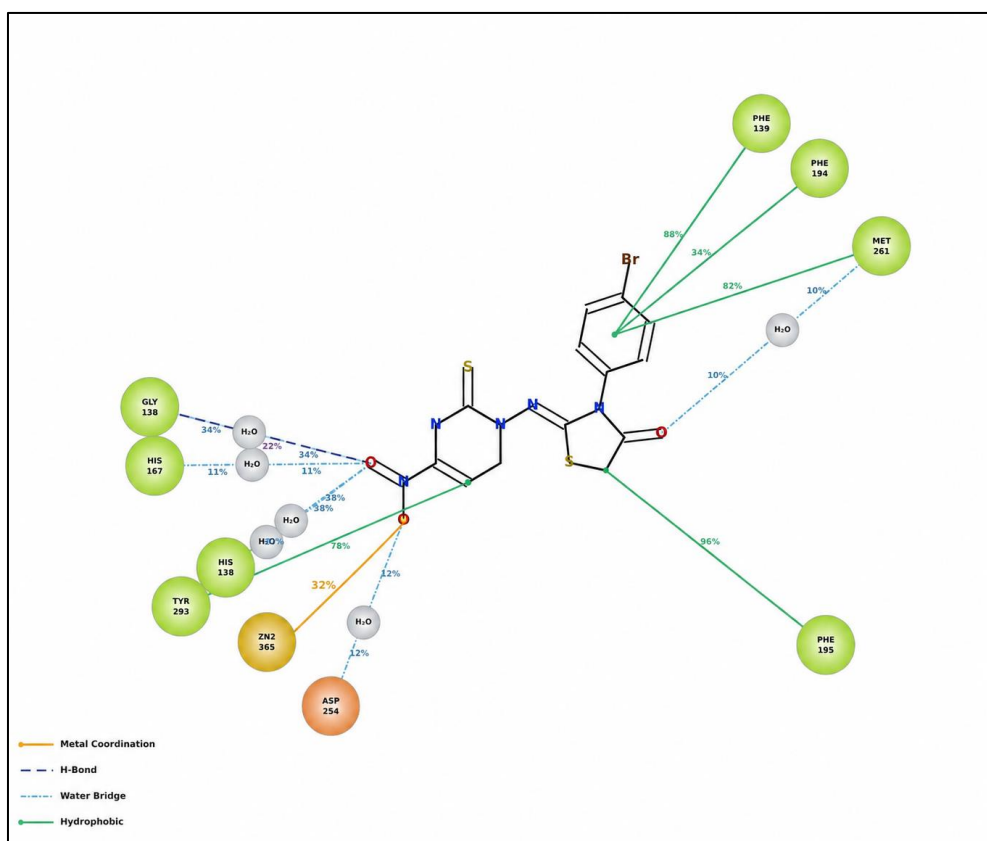


Figure 8: Protein-ligand interaction map for L9 during the 100-ns MD trajectory (trajectory numbering shown).

3.5.4 Integrated Dynamic Interpretation

The MD results converge on one defensible interpretation: after an initial structural adjustment, L9 remained associated with the HDAC8 binding pocket while the complex sampled multiple related conformational substates. The strongest evidence for retention is the narrowed late-time ligand-pocket COM distance together with persistent hydrophobic contacts; protein/ligand RMSD, clustering, and RMSF define the accompanying conformational reorganization. Hydrogen bonds were transient, and the Zn378-specific analysis showed a distance-defined Zn-O contact in approximately 32% of frames using the ≤ 2.2 Å criterion. Because the catalytic zinc was represented as a non-bonded CHARMM36 divalent ion, this behavior is most appropriately described as recurrent metal-proximal engagement rather than continuous zinc chelation.

This dynamic behavior strengthens the rationale for prioritizing L9 after docking: the 2.31 Å docking-stage Zn378 interaction is complemented by late-time pocket retention and recurrent Zn378-O contact geometry during MD. The convergence of pocket retention, hydrophobic persistence, and intermittent distance-defined Zn-O contact supports L9 as the computationally prioritized HDAC8-targeting candidate for structure-guided optimization. Within the non-bonded CHARMM36 zinc representation, the 2.2 Å occupancy criterion was used specifically to characterize recurrent metal-proximal geometry rather than coordination-bond energetics. The MD results therefore provide a focused dynamic characterization of the prioritized L9-HDAC8 complex within the modeled system.

3.6 Integrated Computational Interpretation

The integrated workflow combined comparative molecular docking, focused 100-ns molecular dynamics follow-up of L9, physicochemical and pharmacokinetic profiling, and toxicity prediction. Docking scores were interpreted together with active-site geometry, while the MD analysis characterized conformational relaxation, ligand-pocket retention, interaction persistence, and distance-defined Zn378-O proximity under the non-bonded CHARMM36 zinc representation. SwissADME and ProTox 3.0 outputs were incorporated as complementary developability and safety-prediction layers. Together, these analyses provide a coherent basis for structure-based ranking and medicinal-chemistry prioritization across the designed series [7-10].

4 CONCLUSION

This work defines a coherent structure-based prioritization of fourteen designed heterocyclic molecules over HDAC8 instead of straightforward docking prioritization. The validated docking protocol (SAHA redocking RMSD of the order 1.4 Å) posed L2 as the first to be placed, with a score (-5.954 kcal/mol) but L9 as the more informative of the mechanistic candidates since it had a comparable score (-5.881 kcal/mol) with a short predicted Zn378 contact (2.31 Å) and complementary channel interactions. The 100-ns trajectory revealed a rapid conformational relaxation and a bounded dynamic regime with L9 spatially bound to the catalytic pocket, with recurrent, non-continuous, Zn378-O proximity among many related conformational substates. The zinc analysis employed the non-bonded CHARMM36 representation of Zn^{2+} and a distance definition of 2.2 Å, where Zn-O occupancy was approximately 32 percent; this finding thus describes intermittent metal-proximal interaction in the system being modeled. The developability profile revealed that L9 was predicted to be moderately soluble and to have relatively low lipophilicity, with a high TPSA and low predicted gastrointestinal absorption, and endpoint-specific computational toxicity alerts were revealed through ProTox 3.0. Collectively, the combined docking, MD, ADME, and toxicity data show that L9 is the most informative, mechanistically, member of the designed series and a logical target in the further optimization, which is guided by structure.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The author would like to express their gratitude to the Department of Pharmaceutical Chemistry, College of Pharmacy, Tikrit University, for providing the necessary facilities, computational resources, and support to complete this research.

Funding Source

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Disclosure of conflict of interest

The author declares no conflicts of interest regarding this manuscript.

REFERENCES

- [1] Liang T, Wang F, Elhassan RM, Cheng Y, Tang X, Chen W, Fang H, Hou X. Targeting histone deacetylases for cancer therapy: Trends and challenges. *Acta Pharm Sin B*. 2023;13(6):2425-2463. doi:10.1016/j.apsb.2023.02.007.
- [2] Kim JY, Cho H, Yoo J, Kim GW, Jeon YH, Lee SW, Kwon SH. Pathological Role of HDAC8: Cancer and Beyond. *Cells*. 2022;11(19):3161. doi:10.3390/cells11193161.
- [3] Fontana A, Cursaro I, Carullo G, Gemma S, Butini S, Campiani G. A Therapeutic Perspective of HDAC8 in Different Diseases: An Overview of Selective Inhibitors. *Int J Mol Sci*. 2022;23(17):10014. doi:10.3390/ijms231710014.
- [4] Zhao Q, Liu H, Peng J, Niu H, Liu J, Xue H, Liu W, Liu X, Hao H, Zhang X, Wu J. HDAC8 as a target in drug discovery: Function, structure and design. *Eur J Med Chem*. 2024;280:116972. doi:10.1016/j.ejmech.2024.116972.
- [5] Somoza JR, Skene RJ, Katz BA, et al. Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. *Structure*. 2004;12(7):1325-1334. doi:10.1016/j.str.2004.04.012. RCSB PDB entry: 1T69.
- [6] Mushtaq A, Wu P, Naseer MM. Recent drug design strategies and identification of key heterocyclic scaffolds for promising anticancer targets. *Pharmacol Ther*. 2024;254:108579. doi:10.1016/j.pharmthera.2023.108579.
- [7] Paggi JM, Pandit A, Dror RO. The Art and Science of Molecular Docking. *Annu Rev Biochem*. 2024;93:389-410. doi:10.1146/annurev-biochem-030222-120000.
- [8] Ahmed M, Maldonado AM, Durrant JD. From byte to bench to bedside: molecular dynamics simulations and drug discovery. *BMC Biol*. 2023;21(1):299. doi:10.1186/s12915-023-01791-z.
- [9] 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*. 2017;7:42717. doi:10.1038/srep42717.
- [10] Banerjee P, Kemmler E, Dunkel M, Preissner R. ProTox 3.0: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res*. 2024;52(W1):W513-W520. doi:10.1093/nar/gkaf303.
- [11] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001;46(1-3):3-26. doi:10.1016/S0169-409X(00)00129-0.
- [12] Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. 2002;45(12):2615-2623. doi:10.1021/jm020017n.
- [13] Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem*. 2000;43(20):3714-3717. doi:10.1021/jm000942e.
- [14] Delaney JS. ESOL: estimating aqueous solubility directly from molecular structure. *J Chem Inf Comput Sci*. 2004;44(3):1000-1005. doi:10.1021/ci034243x.
- [15] Daina A, Zoete V. A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem*. 2016;11(11):1117-1121. doi:10.1002/cmdc.201600182.
- [16] Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem*. 1999;1(1):55-68. doi:10.1021/cc9800071.
- [17] Egan WJ, Merz KM Jr, Baldwin JJ. Prediction of drug absorption using multivariate statistics. *J Med Chem*. 2000;43(21):3867-3877. doi:10.1021/jm000292e.
- [18] Muegge I, Heald SL, Brittelli D. Simple selection criteria for drug-like chemical matter. *J Med Chem*. 2001;44(12):1841-1846. doi:10.1021/jm015507e.
- [19] Martin YC. A bioavailability score. *J Med Chem*. 2005;48(9):3164-3170. doi:10.1021/jm0492002.
- [20] Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, Wyatt PG. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem*. 2008;3(3):435-444. doi:10.1002/cmdc.200700139.
- [21] Nnyigide OS, Nnyigide TO, Lee SG, Hyun K. Protein Repair and Analysis Server: A Web Server to Repair PDB Structures, Add Missing Heavy Atoms and Hydrogen Atoms, and Assign Secondary Structures by Amide Interactions. *J Chem Inf Model*. 2022;62(17):4232-4246. doi:10.1021/acs.jcim.2c00571.
- [22] Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res*. 2004;32(Web Server issue):W665-W667. doi:10.1093/nar/gkh381.

- [23] Olsson MHM, Sondergaard CR, Rostkowski M, Jensen JH. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J Chem Theory Comput.* 2011;7(2):525-537. doi:10.1021/ct100578z.
- [24] Abraham MJ, Murtola T, Schulz R, Pall S, Smith JC, Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1-2:19-25. doi:10.1016/j.softx.2015.06.001.
- [25] Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem.* 2013;34(25):2135-2145. doi:10.1002/jcc.23354.
- [26] Molecular Operating Environment (MOE), 2019. Chemical Computing Group ULC, Montreal, QC, Canada.
- [27] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79(2):926-935. doi:10.1063/1.445869.
- [28] Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: a fast force field generation tool for small organic molecules. *J Comput Chem.* 2011;32(11):2359-2368. doi:10.1002/jcc.21816.