

Research Article

Quantum-Mechanically Refined, Dynamics-Coupled, and AI-Augmented Elucidation of Epigenetic Inhibition: An In Silico Paradigm Targeting HDAC8 of *Schistosoma mansoni*

Mohd Imran ,¹ Talha Jawaid ,² Hayaa M. Alhuthali ,³ Amani A. Alrehaili ,³ Abdullah R. Alzahrani ,⁴ Zia Ur Rehman ,^{5,6} Elliot Mbunge ,⁷ and Tafadzwa Dzinamarira ,^{8,9}

¹Center for Health Research, Northern Border University, Arar 73213, Saudi Arabia

²Department of Pharmacology, College of Medicine, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

³Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁴Department of Pharmacology and Toxicology, Faculty of Medicine, Umm Al-Qura University, P.O. Box 13578, Al-Abidiyah, Makkah 21955, Saudi Arabia

⁵Health Research Centre, Jazan University, P.O. Box 114, Jazan 45142, Saudi Arabia

⁶Department of Pharmaceutical Chemistry and Pharmacognosy, Faculty of Pharmacy, Jazan University, P.O. Box 114, Jazan 45142, Saudi Arabia

⁷Department of Applied Information Systems, University of Johannesburg, Johannesburg, South Africa

⁸School of Health Systems and Public Health, University of Pretoria, Pretoria, South Africa

⁹ICAP at Columbia University, Lusaka, Zambia

Correspondence should be addressed to Tafadzwa Dzinamarira; u19395419@up.ac.za

Received 13 September 2025; Revised 6 November 2025; Accepted 12 November 2025

Academic Editor: Wei Wang

Copyright © 2025 Mohd Imran et al. Journal of Tropical Medicine published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The increasing burden of schistosomiasis, compounded by the restriction imposed by monotherapeutic regimens, highlights the pressing need for new molecules that target specific molecular pathways. *Schistosoma mansoni* histone deacetylase 8 (SmHDAC8), a zinc-dependent epigenetic regulator, has emerged as a nonredundant and druggable enzyme, critical for parasite survival, fertility, and chromatin homeostasis. In this study, we outline multiple-mode computational analysis involving structure-based virtual screening against a chemically diverse ligand library, frontier molecular orbital analysis through DFT, large-scale molecular dynamics (MD) simulations (500 ns), and molecular mechanics/gas-phase/Generalized Born (MM/GBSA) energy component analysis, complemented with machine learning-guided pIC₅₀ model building and prediction. Our screening cascade comprising docking, MM, and MD identified the lead candidate, **24374890**, with the best docking score (−9.5 kcal/mol) and desirable electronic configuration (HOMO–LUMO gap: 4.143 eV) for its optimal reactivity–stability balance. MD simulations confirmed its stability in the short term, as well as its conformational preservation in the SmHDAC8 catalytic pocket, as evidenced through low RMSD values, stable free energy basins, and sustained intermolecular interactions. Hydrogen bond analysis proved that compounds 24374890 and 24280440 kept 1–4 stable hydrogen bonds for the entire 500 ns simulation, supporting their strong and stable binding in SmHDAC8's active site. Thermodynamic calculations through MM/GBSA indicated **24374890** has the best energetics for binding ($\Delta G_{\text{total}} = -65.11$ kcal/mol), comprising largely van der Waals and nonpolar solvation energies. Finally,

the pIC₅₀ value for **24374890** was predicted, through supervised machine learning, as 8.1, better than the reference molecule. These convergent findings from quantum mechanical, molecular mechanical, and AI-based computations validate **24374890** as an SmHDAC8 inhibitor that is structurally and dynamically sound. These calculations need to be supported with in vitro enzyme inhibition experiments against recombinant SmHDAC8 and cytotoxicity profiling in schistosome cultures. Moreover, X-ray crystallography or cryo-EM analysis of the SmHDAC8–**24374890** complex would reveal detailed binding conformations.

Keywords: DFT; HDAC8; machine learning; *Schistosoma mansoni*; schistosomiasis

1. Introduction

Schistosomiasis is still one of the most debilitating of the neglected tropical diseases, infecting more than 240 million individuals globally and putting an estimated 800 million at risk, mostly in sub-Saharan Africa, certain regions of South America, the Middle East, and parts of Southeast Asia [1]. *Schistosoma mansoni* is among the significant causative agents of intestinal schistosomiasis and continues as a significant public health issue [2]. Global analyses describe its epidemiology, pathology, and the limitations of current control strategies, highlighting the urgent need for improved interventions [2]. Although praziquantel remains the mainstay of treatment, detailed studies on its metabolism, mode of action, and clinical performance have revealed reduced efficacy in juvenile worms and early evidence of drug resistance in certain regions [3, 4].

Recent WHO guidelines recommend a more integrated approach that combines preventive chemotherapy with snail control, improved water, sanitation, and hygiene (WASH), and enhanced surveillance to strengthen long-term impact [5]. Mass drug administration (MDA) remains a cornerstone of these programs, successfully reducing prevalence in many endemic areas, yet alone it is insufficient to achieve elimination without complementary public health measures [6]. Evidence from multiple endemic settings underscores the urgency of addressing the limitations of praziquantel through diversification of therapeutic options [4]. National programs, such as the comprehensive elimination campaign in Rwanda, demonstrate the effectiveness of coordinated mapping, targeted treatment, and sustained monitoring in progressing toward elimination goals [7].

In parallel, adopting a One Health framework that considers human, animal, and environmental health has been emphasized as essential for sustainable schistosomiasis control. This approach addresses zoonotic transmission cycles and environmental reservoirs alongside human treatment, providing a broader platform for achieving WHO elimination targets [8]. Children carry the highest burden in endemic areas, suffering from long-term sequelae such as anemia, growth delay, and cognitive impairments [9]. The disease also presents economic issues, decreasing working capacity and putting an overload on health care resources.

Epigenetic regulation in schistosomes has been identified as a key driver of parasite development, reproduction, and survival, positioning it as a promising therapeutic target [10]. Disruption of these mechanisms can impair egg formation and reduce transmission potential [11]. In the snail intermediate host, epigenetic plasticity influences

susceptibility to infection, highlighting new avenues for control [12]. Reviews of current knowledge reveal substantial progress but also critical gaps requiring further investigation [13]. Specific epigenetic markers have been linked to schistosomiasis, offering potential tools for diagnosis and targeted therapy [14]. Chromatin regulation and histone-modifying enzymes, in particular, show strong potential as novel drug targets [15].

Among the different isoforms of HDAC, one of them, namely, *Schistosoma mansoni* histone deacetylase 8 (SmHDAC8), has been of particular interest owing to its structural specificity, substrate choice, and drugability [16–18]. SmHDAC8, unlike its counterparts, contains specific amino acid residues in the catalytic pocket, which may be an entry point to selectively target it without cross-reacting with host HDACs [16, 17, 19]. The biological significance of SmHDAC8 is also reinforced by studies at the genetic and biochemical levels [20]. Silencing or pharmacological inhibition of SmHDAC8 perturbs parasite development, alters gene expression associated with survival and reproduction, and causes phenotypic defects, thus establishing it as a target for drug development [21, 22]. SmHDAC8 is significant in stem cell viability and tegument integrity, two processes that are paramount to the survivability of the parasite in the bloodstream of the host [23]. Lack of upregulation of other HDACs in SmHDAC8-suppressed lines further attests to its nonredundant role [15, 24]. Additionally, advancing structural information, such as crystal structures of SmHDAC8 in bound form to inhibitors, has facilitated structure-guided drug development, presenting an accessible paradigm for designing selective therapeutics [16, 25].

In spite of advances in this area, treatment of schistosomiasis is still bound by therapeutic restraints. Praziquantel, which remains the only recommended treatment by WHO, is limited in potency against immature larval stages and fails to safeguard against reinfection [3, 26]. Resistance is not yet prevalent but is increasingly of concern, especially where there is extensive use of MDA programs [6, 27]. Various small molecules and epigenetic inhibitors have also been investigated for antischistosomal activities. For instance, compounds like J1075 and PCI-34051, which were initially designed for HDAC8 in humans, proved to be potent antischistosomal agents, but their selectivity and in vivo safety remain debatable [16, 25, 28, 29]. Other scaffolds of hydroxamic acids and benzamides have also been repurposed or structurally optimized to target SmHDAC8, though most of these prospects remain at or below in vitro advances [18, 30]. Hence, identification of

potent, safe, and selective SmHDAC8 inhibitors remains an area of high-priority interest in drug development against schistosomiasis.

Due to the constraints imposed by existing schistosomiasis therapy as well as growing praziquantel resistance, a severe necessity exists for new therapeutic targets as well as drug leads. SmHDAC8 represents a nonredundant, structurally distinct enzyme vital for parasite survival that can be positioned as a promising target for selective inhibitor molecules. The hypothesis here is that the structural distinctiveness of SmHDAC8 compared to human HDACs can be exploited for selective drug development. To investigate this further, we undertook a multifaceted computer-aided approach involving structure-centered virtual screening, quantum mechanic computations, and molecular dynamic simulation, as well as machine-learned guided prediction of bioactivity for finding and validating new selective SmHDAC8 inhibitors.

Considering increasing interest in epigenetic targets and an immediate need for novel therapeutics, our research aimed to find effective inhibitors of *Schistosoma mansoni* HDAC8 using a structure-based, rational drug design strategy. The aim was to utilize both data-driven predictive modeling and principles of quantum chemistry to identify novel chemotypes of high affinity and biologic relevance against SmHDAC8. With an eye on the structural novelty of SmHDAC8 and the use of chemically diverse scaffolds, this research hopes to provide novelties amenable to further development and optimization. A detailed workflow illustrating the overall computational approach is presented in Figure 1.

2. Results

The virtual screening against SmHDAC8 identified 1500 lead candidates from the Diverse-Lib chemical library with docking scores from -9.5 to -7.0 kcal/mol, indicative of favorable protein–ligand interactions in the target enzyme catalytic domain as shown in the supporting file as Table S1. The workflow for docking subjected the ligands to stringent filtering parameters, ensuring only those ligands with desirable physicochemical properties and good docking scores were filtered for advanced analysis. Of the screened dataset, three compounds—**24374890**, **137276022**, and **24280440**—were top-rank hits according to their high docking scores and chemical diversity. The ligands performed better than most in the dataset and were shortlisted for further analysis. As reference controls for comparing the predictive accuracy in the docking assays, known HDAC8-targeting compounds were included in parallel.

This result demonstrates the chemical diversity and bindability of the Diverse-Lib database and lends credence to the belief that previously undiscovered scaffolds might be able to yield selective affinity toward the parasitic HDAC8 targets. The top three candidates were chosen based not only on energetic aspects but also on structural differentiation, paving the way for quantum mechanical calculations and dynamic simulations in the following stages of the work.

2.1. Density Functional Theory (DFT). Frontier molecular orbital (FMO) analysis of the chosen ligands was achieved through the application of the DFT in order to determine their electronic distribution and tendencies toward reactivity. Highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) isosurfaces of **24374890**, **137276022**, **24280440**, and the reference molecule appear in Figures 2(a), 2(b), 2(c), 2(d), 2(e), 2(f), 2(g), 2(h), whereas their orbital energies and gaps in energy appear in the provided dataset.

24374890 exhibited HOMO energy and LUMO energy values as -0.217776 Hartree and -0.065524 Hartree, respectively, with the HOMO–LUMO gap being 4.143 eV (Figures 2(a), 2(b)). HOMO density was extensively delocalized throughout the sulfur-containing moiety and the extended π -system, implying ready electrophile attack points, and LUMO was localized over the terminal aromatics, implying probable nucleophile attack positions.

137276022 possessed an intrinsically narrower energy gap of 3.790 eV, having HOMO and LUMO energies at the level of -0.186038 and -0.046775 Hartree, respectively (Figures 2(c), 2(d)). The relatively narrower gap indicates higher chemical softness and, probably, greater biologic reactivity. Charge delocalization patterns reveal high conjugated segment overlap, signifying high electron transport properties.

24280440, with the lowest gap value of 3.503 eV among the test compounds, indicates greater potential for charge transfer interactions (Figures 2(e), 2(f)). The HOMO energy was at -0.237295 Hartree, while the LUMO was at -0.108578 Hartree. Spatial distribution shows an extensive π -cloud through the center framework, facilitating its interaction with electrophilic residues in the pocket.

Whereas, the reference molecule showed an appreciably large HOMO–LUMO gap of 6.396 eV, having a positive LUMO energy ($+0.030079$ Hartree), reflecting less reactivity and higher electronic stability (Figures 2(g), 2(h)). The visible gap energy difference between the reference and the test compounds reflects the higher chemical flexibility of the filtered candidates.

2.2. Molecular Redocking and Interaction Analysis. The selected ligands underwent redocking at exhaustiveness = 8 after DFT optimization to enhance their binding pose and ensure spatial compatibility at the active site. The renewed docking scores confirmed the high affinity of the ligands, with **24374890**, **137276022**, and **24280440** having docking scores of -9.5 , -9.0 , and -8.5 kcal/mol, respectively, better than the reference compound, whose docking score was -7.4 kcal/mol. Also, to test sampling robustness, all docking runs were repeated with exhaustiveness = 100. The rank order, binding poses, and docking scores were identical to those obtained at exhaustiveness = 8 (**24374890**: -9.5 ; **137276022**: -9.0 ; **24280440**: -8.5 ; reference: -7.4 kcal/mol), confirming the robustness of the docking outcomes. Post-optimal docking indicates an increase in geometric fidelity and electronic consistency of the ligands with the target pocket shown in Figures 3(a), 3(b), 3(c), 3(d), where 3D

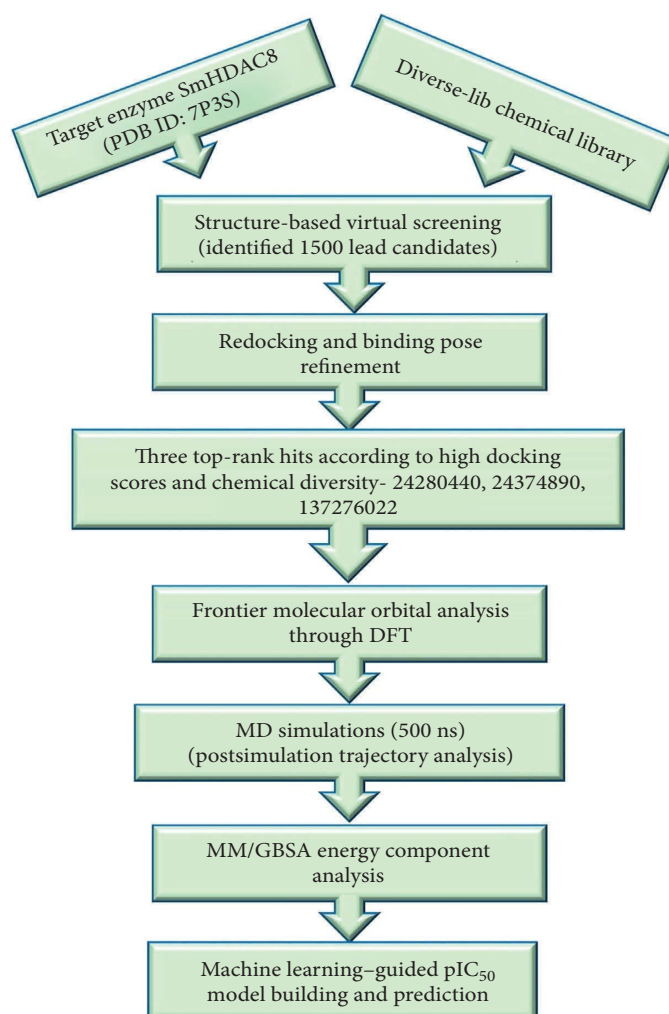


FIGURE 1: Workflow summarizing the computational methodology used in this study.

structures were generated via PyMOL and 2D through BIOVIA, and their interactions are given in Table 1 [31, 32]. To validate the docking protocol, the co-crystallized inhibitor from PDB ID: 7P3S was redocked into the SmHDAC8 active site, and the predicted pose was superimposed with the experimental structure using UCSF Chimera software. The resulting RMSD of 0.7 Å confirmed accurate pose reproduction and validated the docking methodology (Supporting Figure S4).

Intermolecular interaction analysis indicated all investigated complexes had multiple stabilization interactions within the protein's active pocket, yet of different intensities and residue involvement. Complex 24374890 showed strong stabilization by hydrogen bonds to Lys20 and Tyr341, assisted by a tight network of hydrophobic residues (Pro19, Asp100, Gly150, Pro291, His292, Asn342, and Arg380) and π - π stacking by His142, His188, Phe216, Tyr341, and Phe343, indicating rigid anchoring around the zinc catalytic Zn²⁺ ion (Figure 3(a)). Complex 137276022 exhibited several hydrogen bonds by Lys20, Asp186, and His292, accompanied by widespread hydrophobic contacts (Phe21, Asp100, Gly139, Phe151, Phe215, Gln281, Asp285, Gly338,

and Gly339) and aromatic π - π stacking by Trp140, His141, Phe216, and Tyr341, indicating improved ligand accommodation and strong aromatic stabilization, hence an indication of large stabilization enthalpy and high affinity for proteins by the mechanism of increased accessible surface contact and aromatic stabilization assistance, respectively (Figure 3(b)). Complex 24280440 also possessed a strong interaction profile, having hydrogen bonds by His188, His292, and Tyr341 and establishing hydrophobic and π - π contacts by key residues (Lys20, Asp100, Cys152, Gly214, Ile294, Gly339, His142, His141, and Phe216) around ZN501 (Figure 3(c)). The control complex, despite retaining crucial interactions like hydrogen bonds by His14, Gly150, and Tyr341, exhibited comparatively weak hydrophobic coverage, hence less stability (Figure 3(d)). Generally, ligands 24374890, 137276022, and 24280440 possessed stronger binding complementarity and interaction diversity; hence, their increased ability to stabilize and specific protein inhibitory capability.

Overall, these findings emphasize the enhanced binding properties of the screened ligands in comparison with the reference and indicate that the high docking scores for them

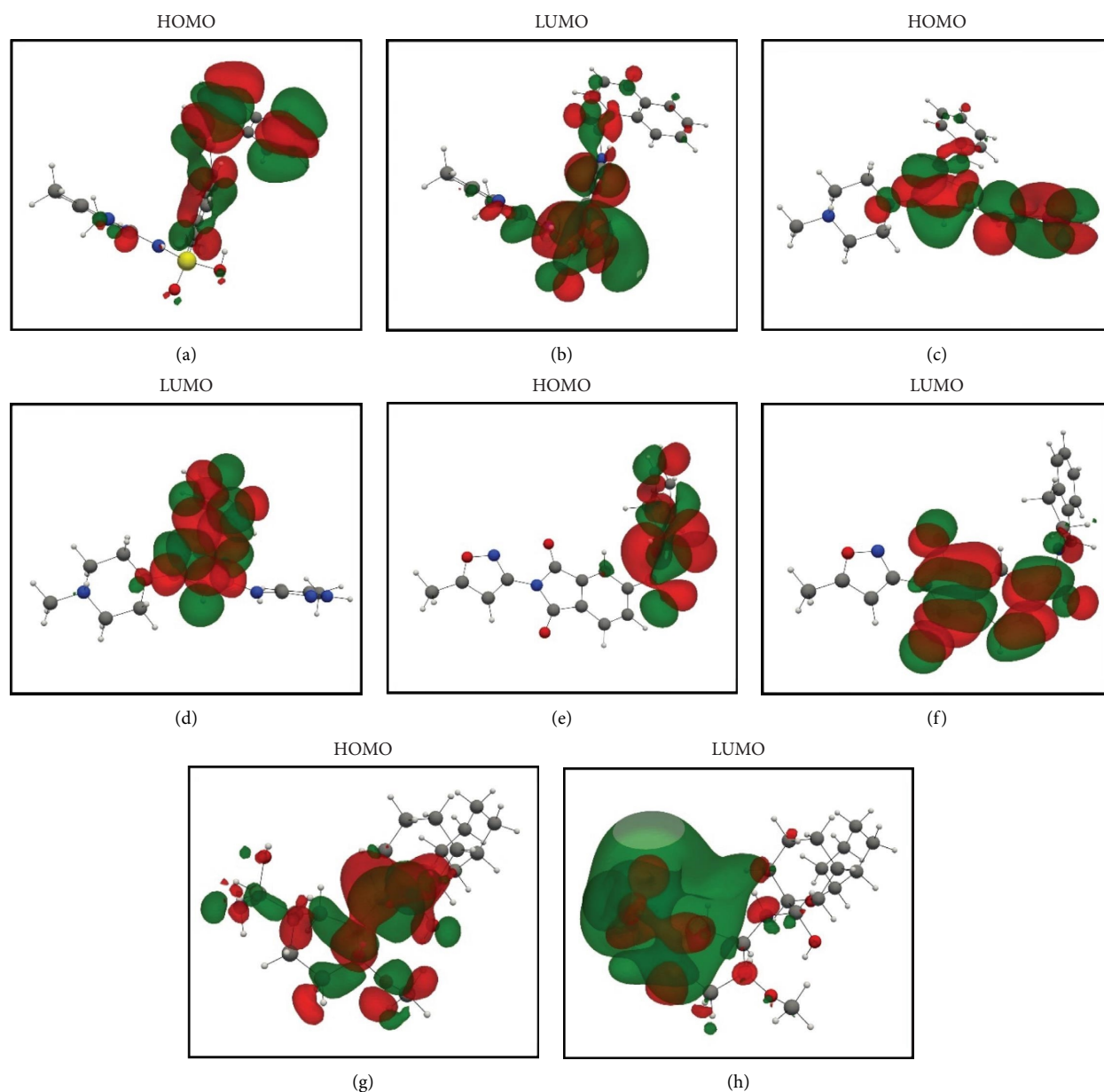


FIGURE 2: Frontier molecular orbitals of top screened ligands. Panels (a–h) display HOMO and LUMO surfaces for (a–b) 24374890, (c–d) 137276022, (e–f) 24280440, and (g–h) the reference compound.

are the result of synergistic hydrogen bonding, broad hydrophobic packing, and aromatic interactions with functionally important residues.

2.3. Stability Analysis Through Molecular Dynamics (MD) Simulation. For the analysis of the dynamic stability and conformational flexibility of the redocked complexes, 500 ns MD simulations in an explicit solvent were performed for each ligand-bound HDAC8 structure.

2.3.1. Root Mean Square Deviation (RMSD) Analysis. RMSD trajectories were created in order to monitor the backbone variations of the protein and the ligand internal

flexibility throughout the simulation, as shown in Figure 4. **24374890** demonstrated a distinctive biphasic profile throughout the simulation (Figure 4(a)). The protein backbone RMSD plateaued gradually near 3.2 Å, consistent with the presence of an overall stable fold. The ligand RMSD, in contrast, varied wildly, peaking above 14 Å in the early 300 ns. There was an evident shift in conformation near 350 ns, after which point the ligand seemed to settle in a second, more constrained pose. This indicates an intensive conformational sampling before the attainment of an enduring, semi-stable bound state toward the end of the simulation.

137276022 showed superior equilibrium already from early simulation steps. Both protein and ligand RMSDs

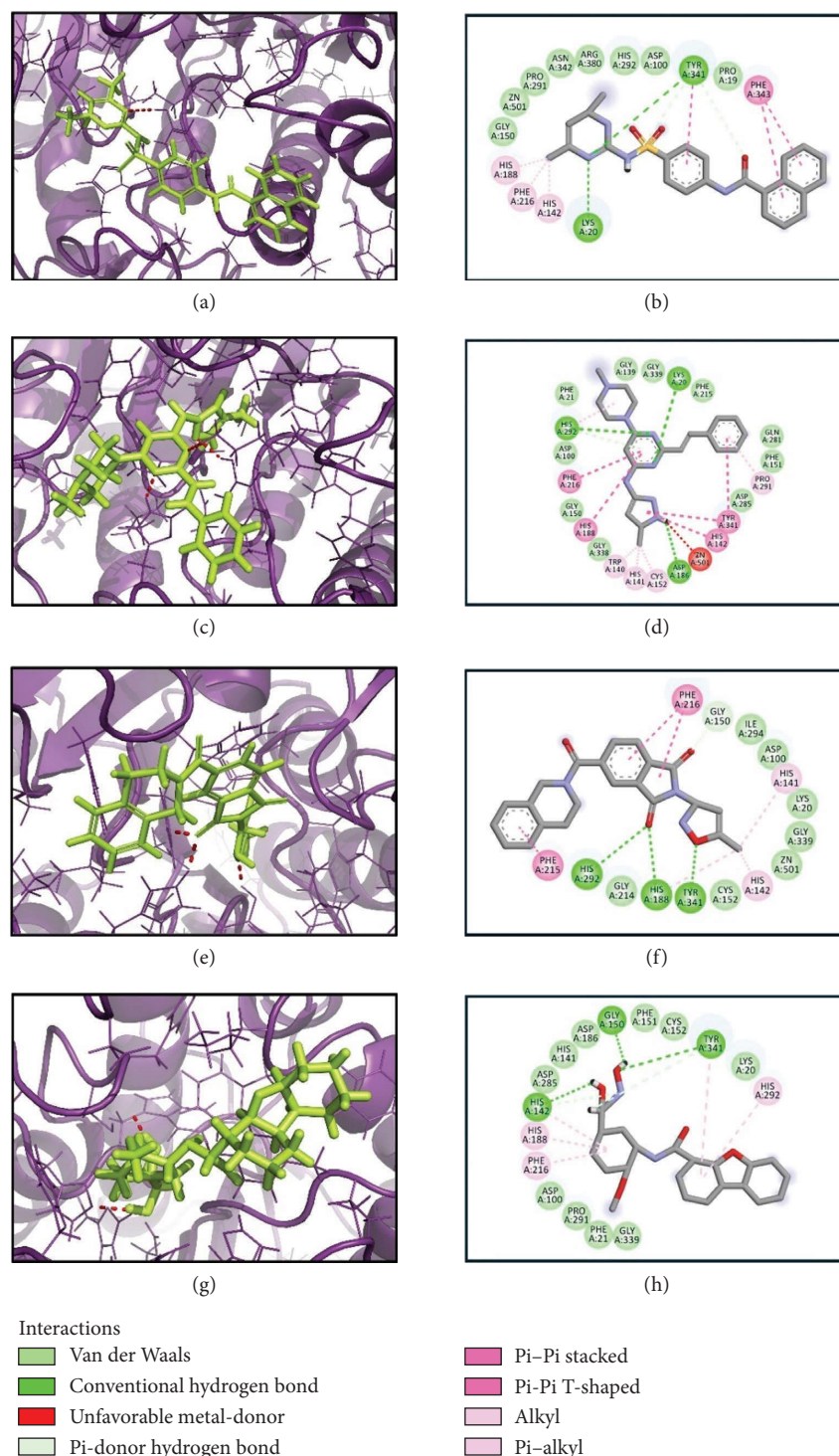


FIGURE 3: 3D and 2D interaction analysis of selected compounds and control within the active site of the target protein. Panels (a–h) represent the binding interactions of compounds (a–b) 24374890, (c–d) 137276022, (e–f) 24280440, and (g–h) the control.

varied in relatively small limits, predominantly below 4 Å (Figure 4(b)). The almost superimposed protein and ligand trajectories suggest an extremely cooperative process of stabilization, with very little drift from the initial bound conformation. The stability, evidenced from the start, reflects an optimized binding geometry and low internal strain in the complex. **24280440** exhibited an increasing trend in

ligand RMSD beyond 200 ns, indicative of progressive displacement and restructuring in the binding site (Figure 4(c)). The protein was intact with RMSD values about 3.5 Å, while the ligand showed periodic spikes in the 5–7 Å region. Such deviations are indicative of interconversion among multiple metastable conformations, perhaps controlled through partial desolvation and

TABLE 1: Illustrating the key binding interactions for selected ligand–protein complexes.

S. no.	Complex	H-bond	Hydrophobic	π - π stacking/ π - π cation
1	24374890	Lys ²⁰ , Tyr ³⁴¹	Pro ¹⁹ , Asp ¹⁰⁰ , Gly ¹⁵⁰ , Pro ²⁹¹ , His ²⁹² , Asn ³⁴² , Arg ³⁸⁰ , ZN ⁵⁰¹	His ¹⁴² , His ¹⁸⁸ , Phe ²¹⁶ , Tyr ³⁴¹ , (Phe ³⁴³), ²
2	137276022	Lys ²⁰ , Asp ¹⁸⁶ , His ²⁹²	Phe ²¹ , Asp ¹⁰⁰ , Gly ¹³⁹ , Gly ¹⁵⁰ , Phe ¹⁵¹ , Phe ²¹⁵ , Gln ²⁸¹ , Asp ²⁸⁵ , Gly ³³⁸ , Gly ³³⁹	Trp ¹⁴⁰ , His ¹⁴¹ , His ¹⁴² , Pro ²⁹¹ , Cys ¹⁵² , His ¹⁸⁸ , Phe ²¹⁶ , His ²⁹² , (Tyr ³⁴¹) ²
3	24280440	His ¹⁸⁸ , His ²⁹² , Tyr ³⁴¹	Lys ²⁰ , Asp ¹⁰⁰ , Cys ¹⁵² , Gly ²¹⁴ , Ile ²⁹⁴ , Gly ³³⁹ , ZN ⁵⁰¹	His ¹⁴² , His ¹⁴¹ , His ¹⁸⁸ , Phe ²¹⁶ , (Phe ²¹⁶) ²
5	Control	His ¹⁴ , Gly ¹⁵⁰ , Tyr ³⁴¹	Lys ²⁰ , Phe ²¹ , Asp ¹⁰⁰ , His ¹⁴¹ , Phe ¹⁵¹ , Cys ¹⁵² , Asp ¹⁸⁶ , Asp ²⁸⁵ , Pro ²⁹¹ , Gly ³³⁹	His ¹⁴² , His ¹⁸⁸ , Phe ²¹⁶ , His ²⁹² , Tyr ³⁴¹

Note: Interactions are listed for compounds (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

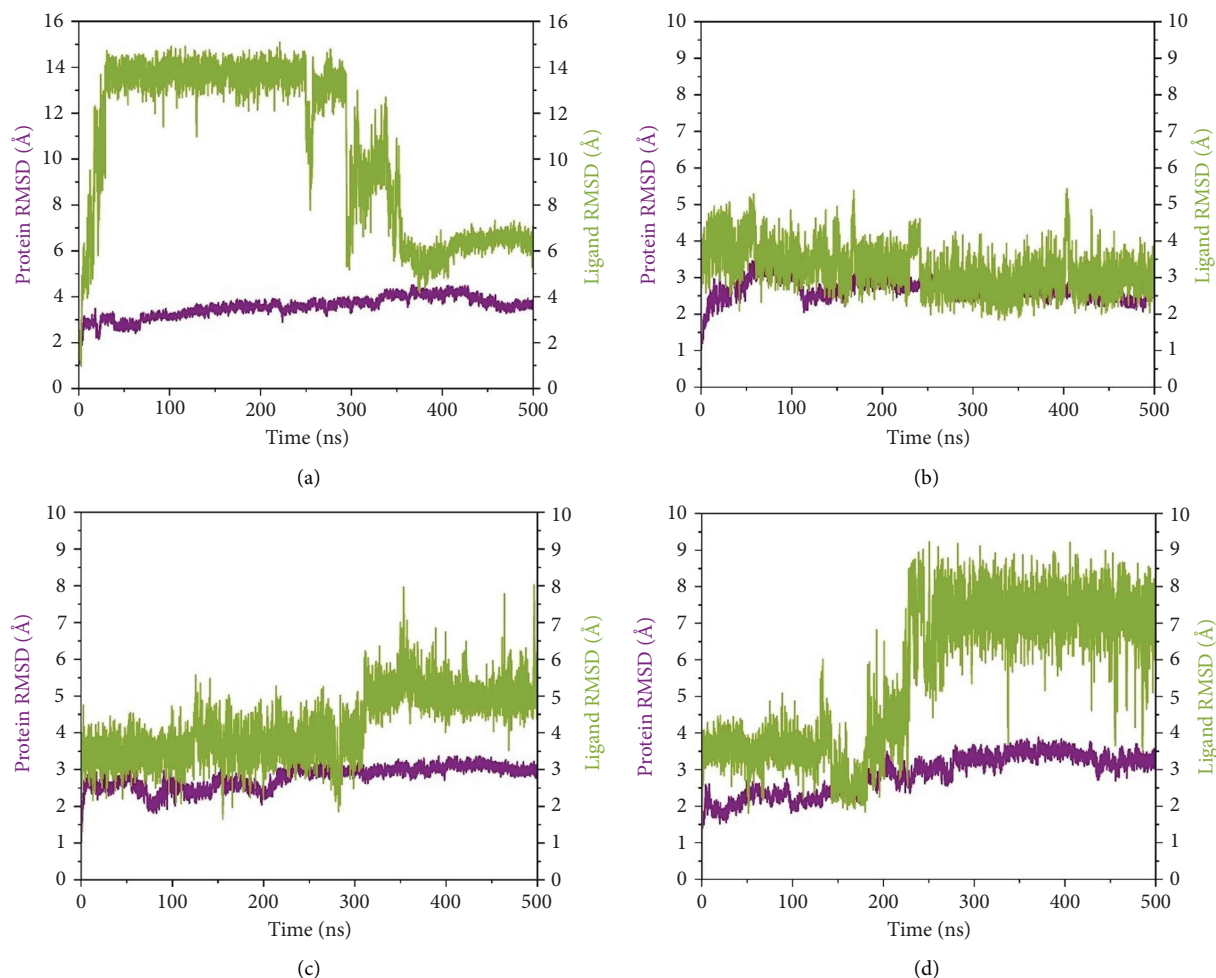


FIGURE 4: Root mean square deviation (RMSD) plots of selected complexes over 500 ns. Panels (a–d) show backbone and ligand RMSD plots for compounds (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

sidechain reorganizations in the region of the binding site. In the reference complex (Figure 4(d)), the protein stayed dynamically stable (RMSD \sim 3.0 Å), while the ligand RMSD rapidly rose after 180 ns and fluctuated persistently above 6 Å. The trend indicates the deterioration in the stability of the bindings with the lapse in time, perhaps because of poor intermolecular contacts or failure in the anchoring contacts. The path indicates movement toward the less bound state or the surface-exposed state.

To establish a comparative baseline for receptor flexibility, the apo form of SmHDAC8 was also simulated for 500 ns under identical conditions. The RMSD trajectory (Supporting Figure S5) showed a rapid equilibration within the first 50 ns and remained stable thereafter, fluctuating around 3.2 Å, indicating overall structural convergence and maintenance of the native fold throughout the simulation.

2.3.2. Root Mean Square Fluctuation (RMSF) Analysis. For residue-level characterization of the protein flexibility in response to ligand association, we calculated RMSF along the 500 ns trajectories for all the complexes. The plots of RMSF capture the trend in atomic displacement along the

amino acid residues of *Schistosoma mansoni* HDAC8 and represent the quantitative fingerprint for local structural mobility, as shown in Figure 5.

Results in the **24374890**-bound complex showed substantial fluctuations in multiple loop areas, specifically near the locations 150–160, and 220–230, and there was a sharp spike past residue 290 (Figure 5(a)). The amplitude reached as high as 6 Å, indicating that **24374890** induces dynamic motions in flexible loops, possibly being responsible for adaptive conformational breathing in the binding cleft. The **137276022** complexes showed a more limited fluctuation profile (Figure 5(b)). Most of the residues oscillated in the region below 2.5 Å, except for an unusually high peak at residue \sim 310, reaching just beyond 6 Å. The single high peak would seem likely to represent an accessible loop or helix near the active site, influencing accessibility or ligand accommodation. Overall, the plug-like nature of this complex is in agreement with the RMSD data, reflecting firm containment in the structure. The RMSF curve for **24280440** presented several clear fluctuation crests throughout the structure (Figure 5(c)). Peaks were visible at residues 160, 260, and 300, with values ranging from 5 to 7 Å. The high

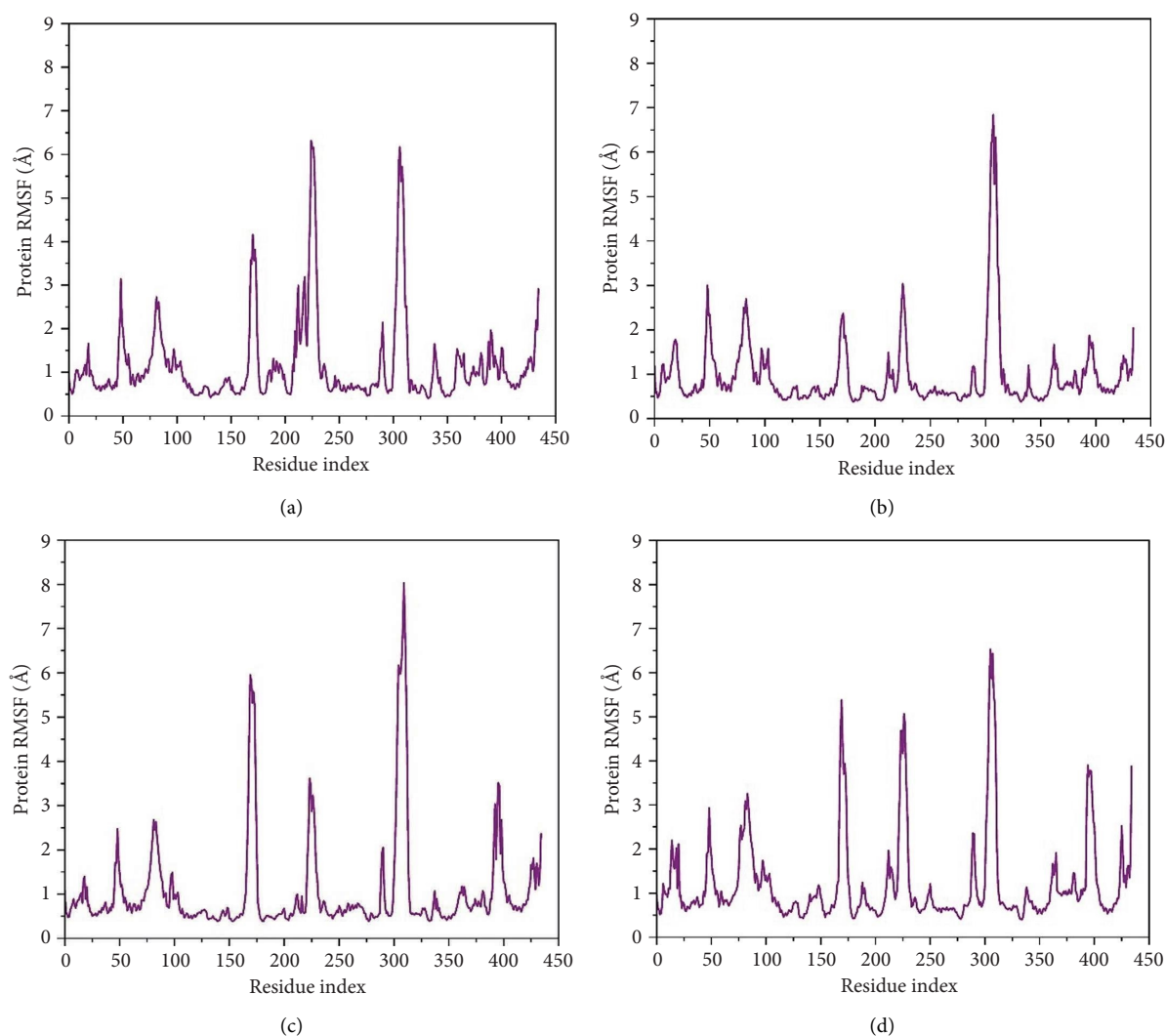


FIGURE 5: Root mean square fluctuation (RMSF) profiles of protein residues in complex with selected ligands. (a–d) Present RMSF values for (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

flexibility in these areas may indicate a compensation movement mechanism for ligand adjustment and adaptive fit, particularly for the bulkier moieties in **24280440**.

In contrast, the reference complex exhibited greater baseline fluctuations for several areas (Figure 5(d)). Abrupt departures near residues 90–100, 160–170, and 290–310 resembled a structurally less restrained system, perhaps indicative of a less tightly bound conformation. Of particular interest, the RMSF maximum at ~ 310 Å approached 7 Å, further suggesting the dynamic nature of the nearby loop in the absence of optimal ligand anchorage.

The RMSF profile (Supporting Figure S6) of the apo-SmHDAC8 protein showed higher flexibility in specific loop regions, particularly between residues 200–250 and 300–310. These areas correspond to surface loops known to contribute to conformational changes near the active site. When compared with the ligand-bound systems, the fluctuations in these regions were noticeably lower, indicating that ligand binding helps stabilize the enzyme structure and reduces local motion around the catalytic pocket.

2.3.3. Comparison Between Final Pose Interactions and Initial Docking. The comparative analysis between the original docking poses and the final conformations after MD simulation (Table 2) provides crucial information for the presence, changes, or disruption of the major contacts that determine ligand stability in the SmHDAC8 protein binding site.

The analysis of interaction showed positive and stabilizing interactions among all assayed complexes, indicating a good binding atmosphere for the ligands. Complex 24374890 showed a strong hydrogen bond interaction with Asp285 and was enveloped by a number of hydrophobic residues, including His141, Gly150, Phe151, Asp186, Leu187, His188, Gly283, Ala284, Gly338–Gly340, Tyr341, and Phe343. Moreover, π -cation interaction with Lys20 also increased its stability, indicating profound accommodation into the pocket. Complex 137276022 cassette arranged two significant hydrogen contacts with Asp186 and His188 and a large hydrophobic scaffold including Phe21, Gly139, Trp140, Ser149, Gly150, Phe151, Phe216, Pro217, Gln281,

TABLE 2: Interaction profile of final MD simulation poses for selected ligand–protein complexes: (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

S. no.	Complex	H-bond	Hydrophobic	π - π stacking/ π - π cation
1	24374890	Asp ²⁸⁵	His ¹⁴¹ , Gly ¹⁵⁰ , Phe ¹⁵¹ , Asp ¹⁸⁶ , Leu ¹⁸⁷ , His ¹⁸⁸ , Gly ²⁸³ , Ala ²⁸⁴ , Gly ³³⁸ , Gly ³³⁹ , Gly ³⁴⁰ , Try ³⁴¹ , Phe ³⁴³	Lys ²⁰
2	137276022	Asp ¹⁸⁶ , His ¹⁸⁸	Phe ²¹ , Gly ¹³⁹ , Trp ¹⁴⁰ , Ser ¹⁴⁹ , Gly ¹⁵⁰ , Phe ¹⁵¹ , Phe ²¹⁶ , Pro ²¹⁷ , Gln ²⁸¹ , Asp ²⁸⁵ , Gly ³³⁸ , Gly ³³⁹ , Gly ³⁴⁰ , Phe ³⁴³	His ¹⁴¹ , (His ¹⁴²) ² , Cys ¹⁵² , His ¹⁸⁸ , Tyr ³⁴¹
3	24280440	—	Arg ²⁴ , Gly ¹³⁹ , His ¹⁴² , Gly ¹⁵⁰ , Phe ¹⁵¹ , His ¹⁸⁸ , Gly ²¹⁴ , Gly ³³⁸	Trp ¹⁴⁰ , His ¹⁴¹ , Cys ¹⁵² , Phe ²¹⁵ , Phe ²¹⁶
4	Control	Asp ¹⁸⁶	His ¹⁴¹ , Lys ¹⁴⁴ , Ser ¹⁴⁹ , Gly ¹⁵⁰ , Asp ²⁸⁵ , Gly ³³⁹ , Tyr ³⁴¹	His ¹⁴² , (Phe ²¹⁶) ²

Asp285, Gly338–Gly340, and Phe343. π – π and π –cation contacts of considerable strength between His141, His142, Cys152, His188, and Tyr341 improved its aromatic stabilization, reflecting the best fitting and high binding efficiency. Complex 24280440, despite lacking hydrogen contacts, possessed a predominating role of hydrophobic and π – π contacts, including Arg24, Gly139, His142, Gly150, Phe151, His188, Gly214, Gly338, Trp140, His141, Cys152, Phe215, and Phe216, indicating a π –electron-centered binding of strong force. The control complex exhibited a hydrogen bond of Asp186 along with a poor hydrophobic contact (His141, Lys144, Ser149, Gly150, Asp285, Gly339, and Tyr341) and a few π – π contacts, including His142 and Phe216. Overall, the ligands, especially 137276022 and 24374890, exhibited a rich network of interactions and strong complementation with the binding cavity, indicating their potency as a more stable and effective inhibitor compared to a control complex.

2.3.4. Hydrogen Bond Analysis. Figures 6(a), 6(b), 6(c), 6(d) shows the hydrogen bond forming profiles from the MD simulations for compounds (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control complex. The y -axis is the number of formed hydrogen bonds, and the x -axis is the simulation time in picoseconds. Up to two stable hydrogen bonds were formed by compound 24374890 along much of the trajectory, reflecting a strong, consistent protein interaction Figure 6(a)). Compound 137276022 showed one to four hydrogen bonds, demonstrating dynamic but persistent binding stability Figure 6(b)). By comparison, compound 24280440 established just one occasional hydrogen bond, reflecting weaker interplay and limited stability within the binding site Figure 6(c)). The control complex established one to three hydrogen bonds throughout the simulation, which is what is expected for a given structure with stable binding activity Figure 6(d)). These findings reflect that compounds 24374890 and 137276022 show more persistent, stoichiometric hydrogen bonding than 24280440, which is in accord with the compounds exhibiting a higher binding affinity and structure stability within the active site.

2.4. Principal Component Analysis (PCA) of MD Trajectories. PCA was performed in order to describe the major motions in the MD trajectories for the SmHDAC8–ligand complexes. Atomic fluctuations were projected onto the two leading eigenvectors (PC1 and PC2) in order to simplify the conformational landscape, describing collective motion modes and dynamic sampling behavior in the course of time, as shown in Figure 7.

In Figure 7(a), corresponding to 24374890, the conformational space is split into two prominent clusters with clear occupancy, suggesting that the complex alternated between two pervasive macrostates. The initial frames (yellow) transition smoothly toward deeper purple, charting a time-dependent path from an ensemble that is flexible toward an ensemble that is more compact. That such segregation occurs indicates extensive structural rearrangement, probably resulting from ligand-induced modulation

of local motions (Figure 7(b)). shows 137276022, in which the PCA scatter plot indicates an extremely smooth transition with overlapping clusters. The frame-wise color gradient indicates smooth, continuous movement with no sudden jumps, suggesting that the complex was in dynamic stability and sampled along a narrow conformational corridor. This is in support of the RMSD and RMSF results, further affirming the structural integrity of the 137276022 complex along the course of the simulation (Figure 7(c)). with 24280440 has evidence for a unidirectional trajectory sweep. The frames proceed from a tightly packed initial cluster toward an extended, broad area along PC2, representing a steplike drift in conformation. The model may be representative, in turn, for slow, incremental relaxation or for partial movement of the ligand into the pocket, consistent with its rising ligand RMSD (Figure 7(d)). is the reference compound, with the same bifurcated landscape as before, though it has a very different topology. Tightly packed subspace occupancy is the early phase (yellow cluster) for an extended stretch into an extremely dissociated terminal cluster. It is an indicator of structural instability and points toward a change in conformation, possibly associated with loss of ligand retention.

2.5. Free Energy Landscape (FEL) Analysis. The FEL for each complex was reconstructed from the top two principal components (PC1 and PC2) obtained from the Cartesian coordinate PCA. The FEL plots provide a thermodynamic account of the protein–ligand complex conformational heterogeneity and stability throughout the 500 ns simulation, where energy basins correspond to the most thermodynamically favored conformational states, as shown in Figure 8.

In the case of 24374890 in (Figure 8(a)), the energy surface shows several clearly separated minima along the PC1–PC2 plane. There is an evident global minimum in the limited energy basin at about 1.5 kJ/mol, with shallow local minima surrounding it, reflecting that the complex reached many metastable conformations. The distribution indicates plasticity in the binding and the possibility of the ligand sampling other orientations for a short while before adopting the deeper energy well. The FEL for 137276022 in (Figure 8(b)) displayed a sharply bounded and deep energy well with the minimum at 1.0 kJ/mol. The steep slope encircling the well indicates a thermodynamically stable and well-localized conformation ensemble, consistent with low structural drift and high enthalpic retention of the ligand. This bounded landscape corroborates the previously noted 137276022 complex RMSD and PCA stability. In Figure 8(c), 24280440 exhibited an energy profile that was broader and shallower with global minima at 2.2 kJ/mol. The gradient and extent of the basins indicate that the complex was in an area in the conformational space that was greater, with reduced discriminatory energy among states. Such behavior is consistent with its intermediate RMSD profile and dynamic fluctuations in RMSF, pointing toward an interaction mechanism that is less rigid, yet still energy-guided.

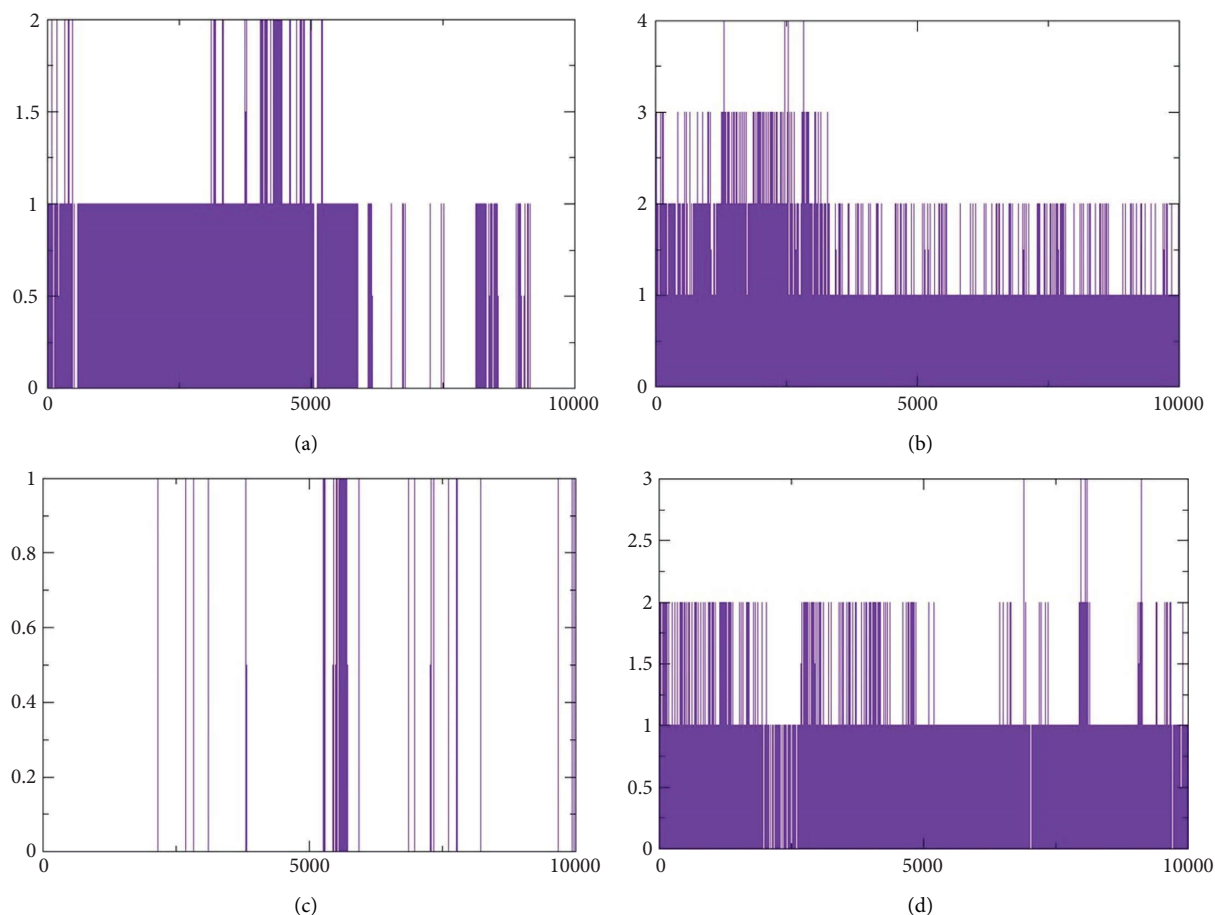


FIGURE 6: Hydrogen bond occupancy profiles during MD simulations for compounds (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

The reference compound, shown in (Figure 8(d)), displayed a rough topography with numerous shallow wells and small energetic basins. The most stable conformational state was at about 2.5 kJ/mol, but the state changes seem frequent and extensive, suggesting greater conformation flexibility and less ligand anchoring. Such an FEL profile indicates weaker thermodynamic stabilization compared with the newly discovered hits.

2.6. Minima Structure Extraction and Superimposition Analysis. To determine the protein–ligand complex conformation stability in the course of time and ensure the constancy of the docking poses, representative minima structures were sampled from the FELs and aligned onto the corresponding initial docked pose. The comparison provided a structure cross-check between predicted and dynamically achieved binding conformations, as shown in Figure 9.

The three images in the initial set (green, blue, and magenta ribbons) describe the lowest-energy conformers derived from different energy basins, identified in the analysis performed using the FEL. The second image set shows the structural superimpositions for these energy-minimized conformations with their starting docking

positions. Tight spatial correlation between the two indicates consistency in ligand placement and pocket occupancy in the course of the trajectory, as shown in Supporting Figure S1.

The superposition RMSDs between the minima-extracted and the initial docked positions of the ligands were calculated in order to determine the conformation retention. The observed values were **24374890**: 1.329 Å, **137276022**: 1.353 Å, **24280440**: 1.233 Å, and Reference: 1.231 Å. All the values remain below 1.5 Å, indicative of minimal ligand orientation shifts throughout the simulation and verifying the hypothesis of consistent binding fidelity. Of the candidates, **24280440** exhibited the nearest overlap (1.233 Å), suggesting its docked conformation was already in an almost optimal energy state and needed little correction. **24374890** and **137276022** exhibited slightly greater RMSDs, but their positioning still indicated structurally consistent transitions, confirming the stability of their original positions. Interestingly, the reference compound showed the lowest RMSD (1.231 Å), although care should be exercised in its interpretation. In spite of the geometrical superposition, dynamic and energy analysis earlier showed weaker temporal binding stabilization, inferring that geometrical rigidity might not translate into functional stability.

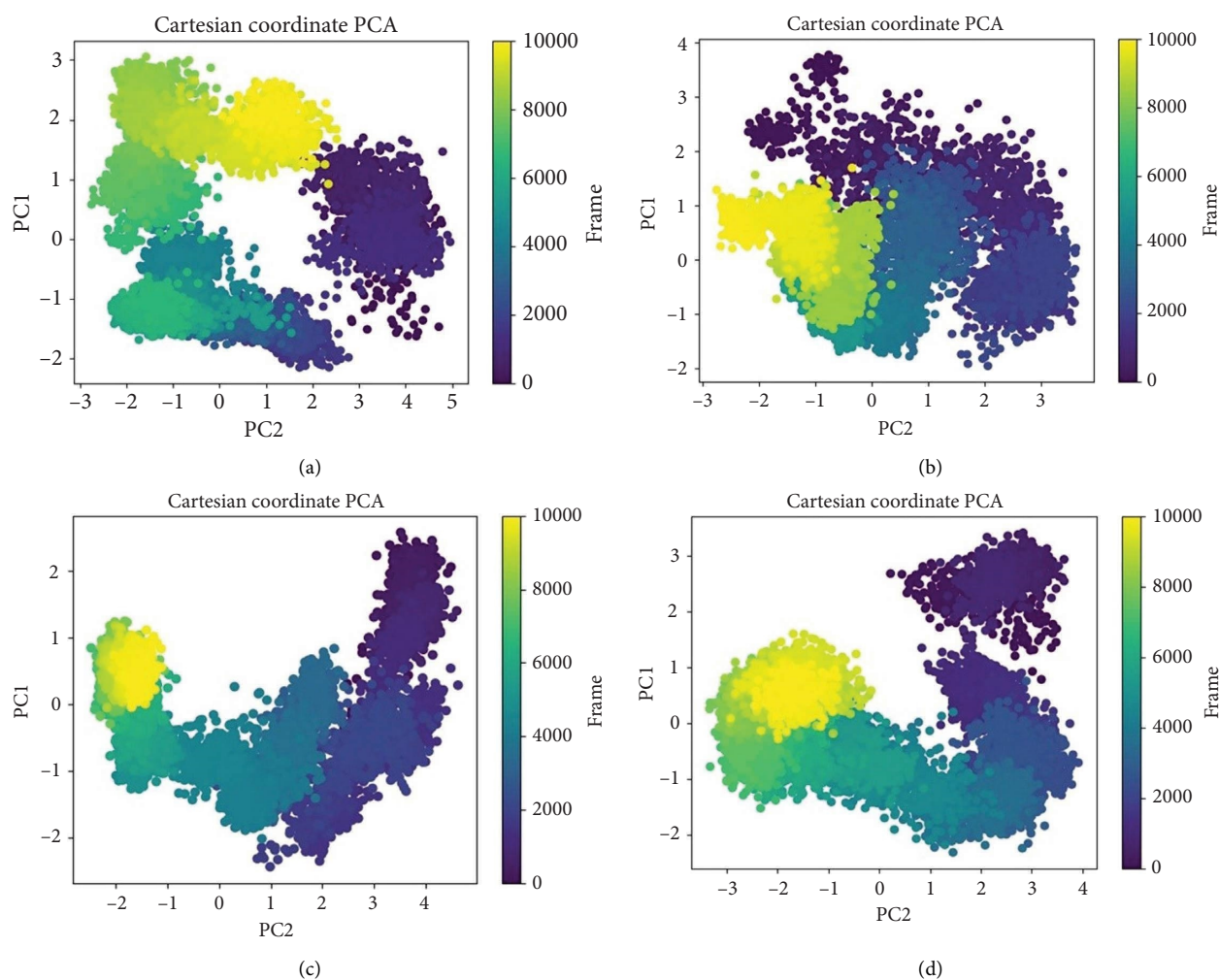


FIGURE 7: Principal component analysis (PCA) of MD trajectories based on Cartesian coordinates. Panels (a–d) show conformational space sampled: (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

2.7. Molecular Mechanics/Gas-Phase/Generalized Born (MM/GBSA) Binding Free Energy Analysis. The MM/GBSA method was used to calculate the binding affinity for each protein–ligand complex, partitioning the total binding free energy into the van der Waals, electrostatic, polar solvation, and nonpolar solvation terms. The averaged values from the 1000 snapshots from the last 50 ns MD simulation trajectories represented the thermodynamic insight into ligand retention and strength of interactions, as shown in Table 3.

24374890 was the most energetically favorable binder, having an overall free energy (ΔG_{total}) value of -65.11 ± 24.08 kcal/mol. It was largely accomplished through a dominant van der Waals term (-53.06 ± 3.16 kcal/mol) and significant electrostatic stabilization (-22.09 ± 14.37 kcal/mol), balanced against the partial cost of polar solvation ($+23.06$ kcal/mol). The nonpolar solvation contribution (-13.02 kcal/mol) further contributed to the ligand's thermodynamic profile. The low net solvation energy ($+10.04$ kcal/mol) amplified its liking for the hydrophobic core of the active site. **137276022** ranked second,

with ΔG_{total} being -54.97 ± 15.55 kcal/mol. It showed slightly weaker van der Waals interactions (-47.72 ± 3.42 kcal/mol) than **24374890**, while its better nonpolar solvation energy (-18.16 ± 1.79 kcal/mol) suggested stronger hydrophobic burial. Its electrostatic interactions (-16.29 kcal/mol) were moderate, while the polar solvation penalty ($+27.21$ kcal/mol) was similar to **24374890**. **24280440** produced a ΔG_{total} value of -41.02 ± 17.30 kcal/mol, reflecting moderate affinity. It showed the lowest in the three hits' van der Waals energy (-37.81 ± 2.85 kcal/mol), and its electrostatic component (-11.63 kcal/mol) was fairly weak as well. Surprisingly, the penalties for both polar and nonpolar solvation were balanced, implying a less hydrophobically optimized pocket orientation.

The reference compound exhibited the lowest binding energy with $\Delta G_{\text{total}} = -34.57 \pm 16.14$ kcal/mol, while its strongest electrostatic component was observed (-30.90 ± 6.45 kcal/mol). Nevertheless, its benefit was offset by the high polar solvation penalty ($+27.89$ kcal/mol) and the lowest van der Waals contribution (-21.74 ± 3.30 kcal/

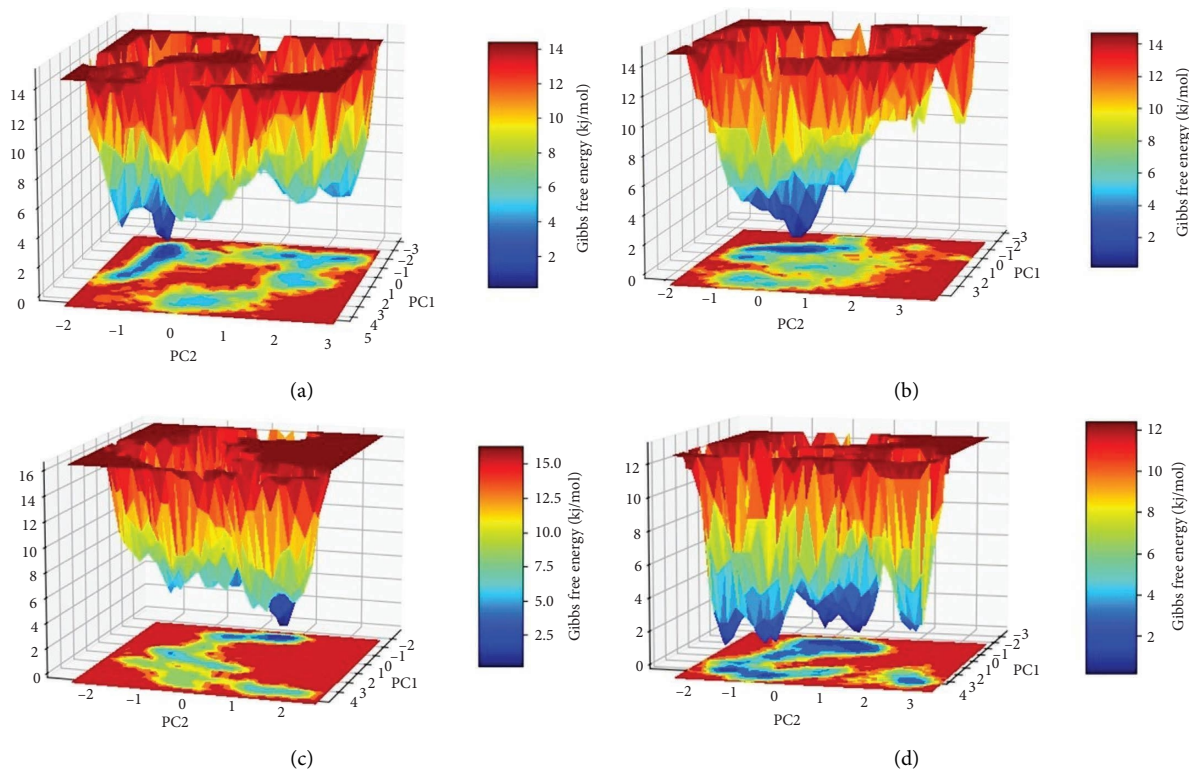


FIGURE 8: Free energy landscapes (FELs) of protein–ligand complexes projected onto the first two principal components (PC1 and PC2). Panels (a–d) represent the Gibbs free energy surfaces for compounds (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

mol), reflecting poor packing in the active site. Moreover, its net solvation energy (+18.07 kcal/mol) indicates poor aqueous compatibility and weaker complex stability.

2.8. Machine Learning-Based Bioactivity Prediction Outcomes. To identify the inhibitory activities of the chosen ligands against *Schistosoma mansoni* HDAC8, an automated ligand-based machine learning pipeline was applied. The procedure included similarity profiling, clustering, model benchmarking, and pIC₅₀ prediction. Results at each stage are described below, along with supporting graphical data.

2.8.1. Molecular Similarity and Clustering. Figure 10 gives the molecular similarity distributions between the lead compounds and the reference inhibitor, calculated using both MACCS and Morgan fingerprints with Tanimoto and Dice similarity measures. The spread witnessed on similarity values—especially wider for Morgan fingerprints—is a clear indication that the compound collection encompasses a chemically diverse space. Interestingly, MACCS-based comparisons prove to have a wider similarity score, picking up common substructural features more sensitively. This diversity is a desirable aspect of virtual screening, given increasing chances of discovering novel chemotypes with disparate binding profiles. In spite of such wide chemical variation, Supporting Figure S2 demonstrates that the structural clustering method adequately clustered

compounds with high internal similarity. The majority of clusters have their mean and median Tanimoto similarity greater than 0.8, reflecting strong internal cohesion and structural homogeneity for clusters. Clusters 5 and 7–9, for instance, have especially narrow similarity distribution, implying highly conserved chemotypes that might be good candidates for narrow-focus optimization. Clusters 0–4, on the other hand, retain high intra-cluster similarity while tolerating relatively broader structural variation, which might be effective for exploring SARs based on scaffolds. In combination, such analyses verify that the compound set obtains a good compromise between global diversity—to cover wide chemical space—and local similarity inside clusters—to allow for rational prioritization and reduction of redundancy. This validates the fingerprint-based similarity evaluation as well as the workflow of clustering as informative ingredients of the ligand selection strategy.

2.8.2. Fingerprint Enrichment Performance. To determine whether MACCS and Morgan fingerprints were able to distinguish actives from inactives early in the ranked list, cumulative enrichment curves were created. Morgan fingerprints were slightly better than MACCS in locating the actives higher in the dataset, as indicated by the faster trajectory of the curve toward the best diagonal (Figure 11). Both curves clearly performed better than the random baseline, implying that the descriptors in use caught meaningful chemical signals for predicting activity.

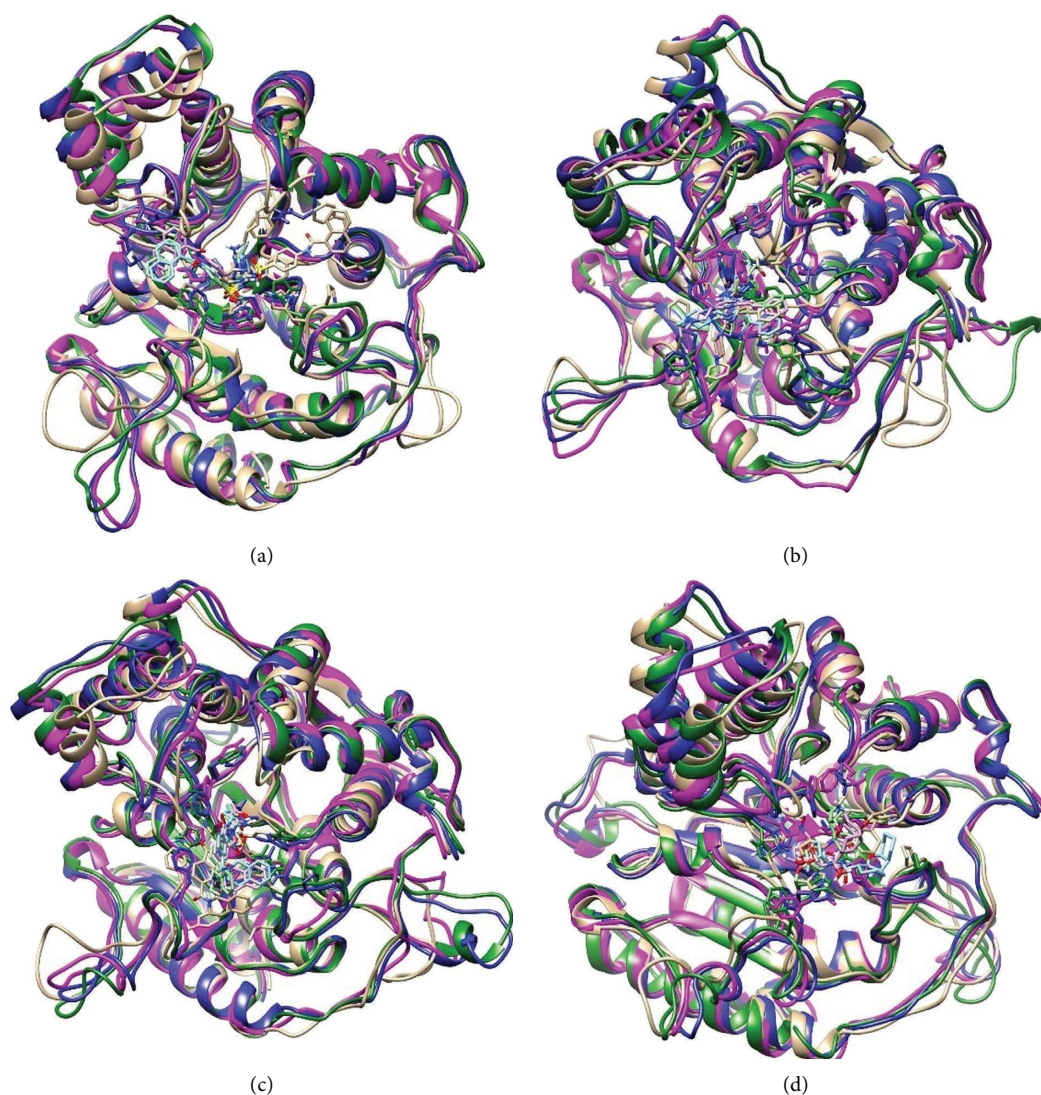


FIGURE 9: Superimposition of the three lowest-energy conformations and the initial docking pose for each protein–ligand complex. Panels (a–d) show structural alignment: (a) 24374890, (b) 137276022, (c) 24280440, and (d) the control.

TABLE 3: MM-GBSA binding free energy decomposition for ligand–protein complexes.

Energy components/complexes	24374890 (ΔG)	137276022 (ΔG)	24280440 (ΔG)	Reference (ΔG)
Van der Waals energy ($\Delta VDWAAALS$)	-53.06 ± 3.16	-47.72 ± 3.42	-37.81 ± 2.85	-21.74 ± 3.30
Electrostatic energy (ΔEEL)	-22.09 ± 14.37	-16.29 ± 2.35	-11.63 ± 5.12	-30.90 ± 6.45
Polar solvation energy (ΔEGB)	23.06 ± 22.15	27.21 ± 7.97	25.32 ± 7.78	27.89 ± 4.99
Nonpolar solvation energy ($\Delta ESURF$)	-13.02 ± 1.90	-18.16 ± 1.79	-16.89 ± 1.54	-9.81 ± 1.39
Net gas phase energy ($\Delta GGAS$)	-75.15 ± 17.53	-64.01 ± 5.77	-49.44 ± 7.97	-52.65 ± 9.76
Net solvation energy ($\Delta GSOLV$)	10.04 ± 6.54	9.04 ± 9.77	8.42 ± 9.32	18.07 ± 6.38
Total binding free energy (ΔG_{total})	-65.11 ± 24.08	-54.97 ± 15.55	-41.02 ± 17.30	34.57 ± 16.14

2.8.3. Regression Model Comparison. Twenty-two models were trained and validated with cross-validation and independent test sets. HistGradientBoostingRegressor, GradientBoostingRegressor, and LGBMRegressor performed the best, with test set R^2 values approaching 0.7, as can be seen in the R^2 bar chart (Supporting Figure S3). They performed better than LinearRegression, MLPRegressor, and DecisionTreeRegressor, often with negative or low values of

R^2 . The large gap in training and test set values for R^2 for some models, notably ARDRegressor and MLPRegressor, suggested the possibility of overfitting.

2.8.4. Mean Squared Error (MSE) Analysis. As an accompaniment to the R^2 analysis, the MSE for all the models was calculated. The MSE plot (Figure 12) corroborated the

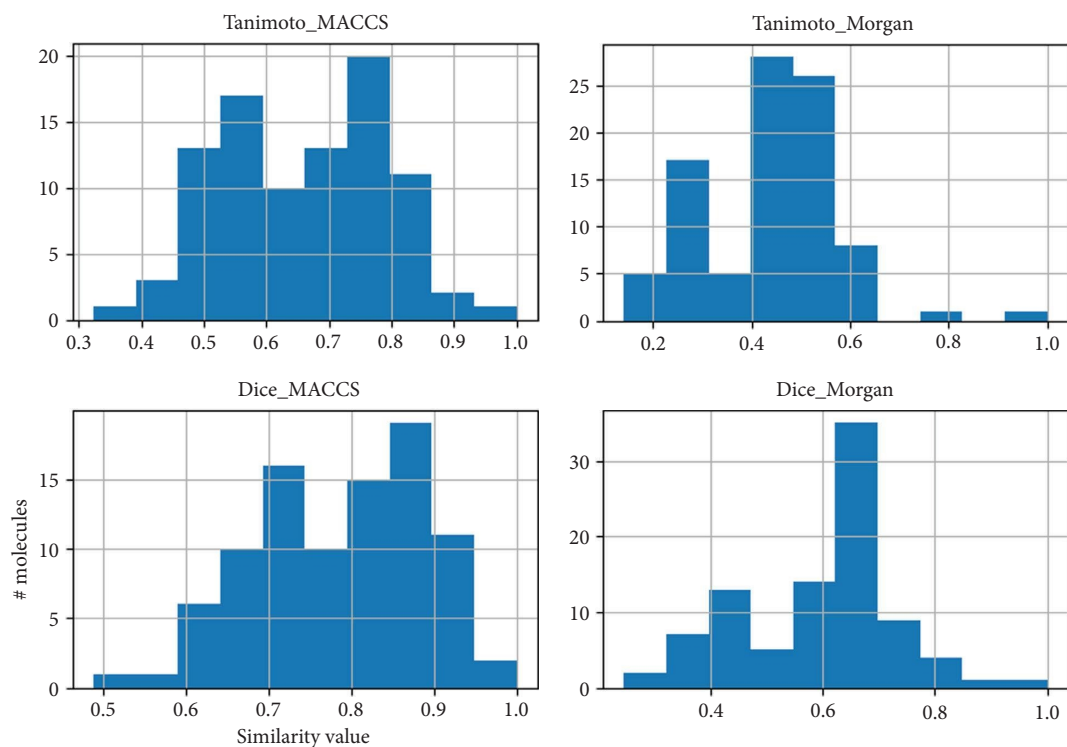


FIGURE 10: Distribution of pairwise molecular similarity scores using Tanimoto and Dice coefficients with MACCS and Morgan fingerprints. Each subplot represents the frequency of similarity values across the dataset. The x-axis shows the similarity score (0–1), and the y-axis indicates the number of molecule pairs. Tanimoto_Morgan and Dice_Morgan exhibit broader distributions, suggesting higher structural diversity captured by Morgan fingerprints.

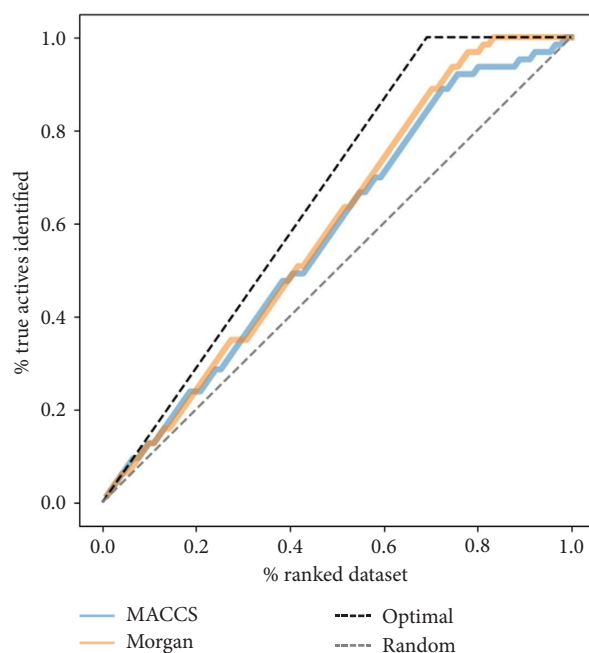


FIGURE 11: Cumulative enrichment plots for MACCS and Morgan fingerprints; both descriptors outperform random ranking, with Morgan showing a slight edge.

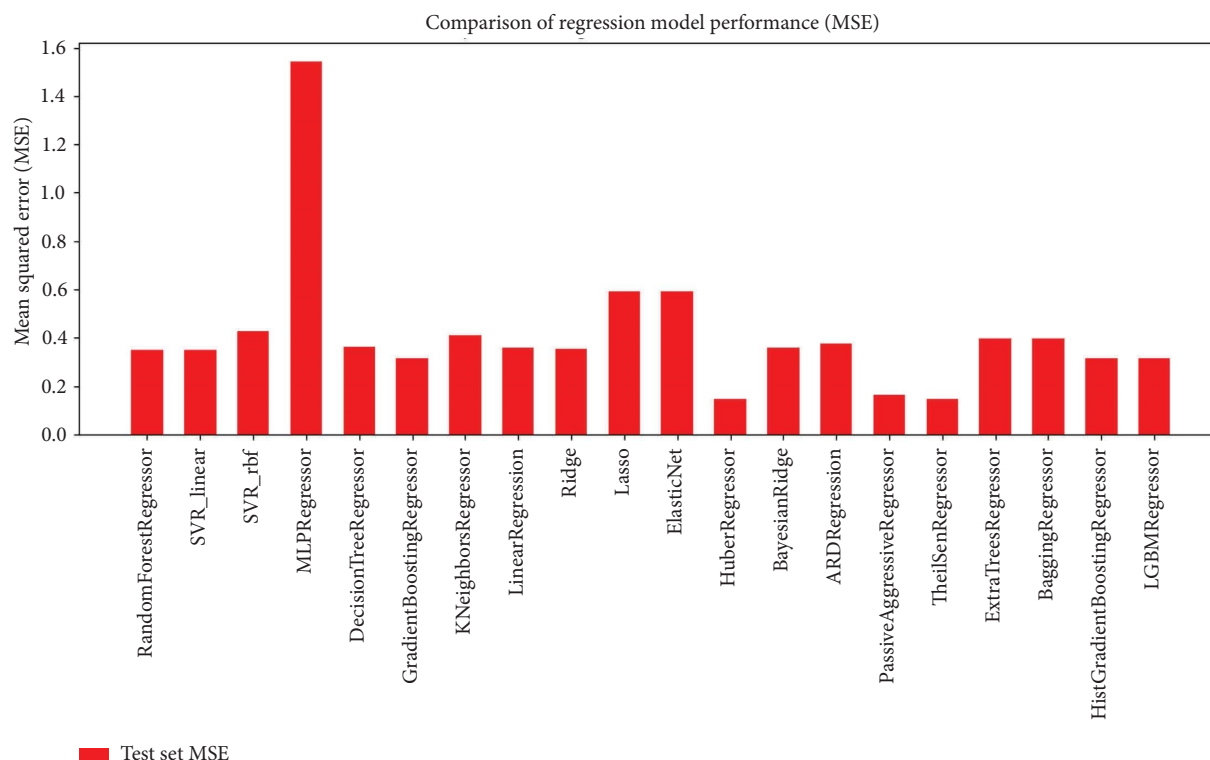


FIGURE 12: Mean squared error (MSE) for test set predictions of each regression model; ensemble models outperformed linear and tree-based variants.

above findings—HistGradientBoostingRegressor, ExtraTreesRegressor, and LGBMRegressor showed minimal prediction error (MSE < 0.2), evidencing high predictive accuracy. MLPRegressor and ElasticNet showed MSE > 1.0, signifying poor generalizability.

2.8.5. Model Selection and Predicted pIC_{50} for Selected Compounds. A total of 22 regression models were evaluated using MSE and coefficient of determination (R^2) on a handpicked SmHDAC8 bioactivity dataset downloaded from ChEMBL (UniProt ID: A5H660). As shown in Figure 12 (MSE) and Supporting Figure S3 (R^2), ensemble-based models—HistGradientBoostingRegressor, LGBMRegressor, and ExtraTreesRegressor—had better performance than linear regressors, shallow trees, and neural networks. Based on a low MSE (< 0.2), R^2 close to 0.7, and similar performance across folds, which reflects high generalizability with little overfitting, the final model chosen was HistGradientBoostingRegressor. This sort of model is also highly suitable for modeling structure–activity relationships that are nonlinear, common in cheminformatics. Conversely, such linear or overly flexible models like MLPRegressor or ElasticNet resulted in a negative R^2 outcome and a somewhat greater MSE, validating their unsuitability for this task. Model verification was carried out using a threefold cross-validation approach on the whole dataset, which enabled all compounds to act both for training and verification purposes, yet employed the available limited high-quality SmHDAC8-specific data to the maximum extent feasible.

Crucially, the dataset comprised just the experimentally verified records specific to SmHDAC8, thus ensuring species-specific relevance without any domain transfer assumptions being involved. Finally, the developed model was used for the prediction of pIC_{50} values for four lead compounds that were given a priority rank. Compound 24374890 exhibited the highest potency prediction ($pIC_{50} = 8.1$), indicating subnanomolar inhibition, whereas the second-best compound 24280440, had a $pIC_{55} = 6.8$, while compound 137276022, which had strong structural/dynamic stability, had a relatively lower pIC_{55} prediction of 5.5, which could be due to undesirable electronic or steric features detrimental to good activity. Reference compound prediction was at $pIC_{50} = 7.5$, which is according to its known inhibitory activity, yet again confirming compound 24374890's much better profile (Figure 13). The results confirm here how machine learning-oriented predictions, when complemented with a quantum mechanical analysis, along with molecular dynamic analyses, facilitate a multi-dimensional assessment of prospective SmHDAC8 inhibitors.

3. Discussion

HDAC8 has become an attractive druggable epigenetic target owing to its chromatin remodeling, parasite development, and tegument integrity roles. In our work here, we applied an integrated computational protocol—virtual screening, DFT, MD simulations, MM/GBSA estimated

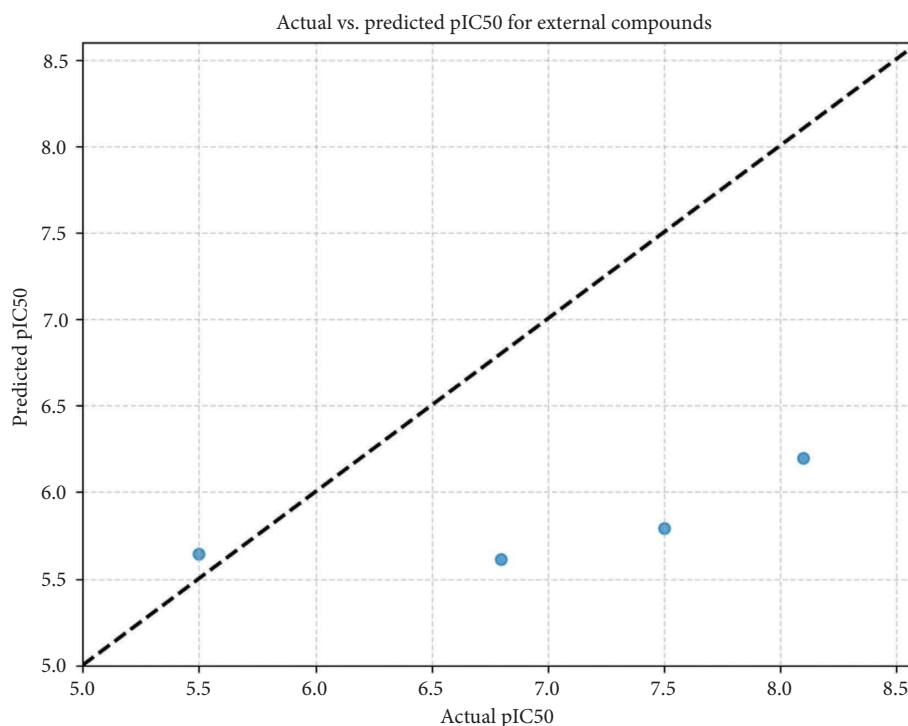


FIGURE 13: Predicted pIC₅₀ values of selected compounds using the best model (HistGradientBoostingRegressor), with Comp_1 emerging as the most potent inhibitor.

binding energies, and machine learning-based pIC₅₀ predictions—to discover and validate new HDAC8 inhibitors among a small molecule chemical diversity set. **24374890** showed superior performance in all the computational assessments, reflecting its high likelihood as an antischistosomiasis lead molecule.

Our docking result showed that **24374890** had a docking score of -9.5 kcal/mol, much greater than most known HDAC8 inhibitors in structure-based reports. For comparison, recent *in silico* analyses targeting parasitic HDACs in most cases produced docking scores ranging from about -8.0 kcal/mol, with limited focus on quantum descriptors and machine learning [33]. Unlike such a conventional strategy, we incorporated a DFT-based FMO analysis, in which we found that **24374890** exhibited a HOMO–LUMO gap of 4.143 eV, an ideal sign of chemical stability and reactivity index. Ligands known to have such gap values were indicated in the literature as being well-suited for engaging in the transfer of electron density in the charged cavities in enzymes [34].

The 500 ns molecular dynamic simulations showed sustained interactions and minimal fluctuation in the **24374890** complex, with RMSD values reaching stability after 300 ns. Unlike prior 100 ns simulations employed for the assessment of probable HDAC inhibitors [35]. Our longer simulations observed multiple energy minima and conformations, emphasizing the dynamic binding retention of the ligand in the active site. RMSF analysis showed localized flexibility, in particular in the vicinity of the loops, without sacrificing core structural stability. Hydrogen bond analysis showed that compounds **24374890** and **24280440**

had firm hydrogen bonds for the entire simulation, aiding in complex stability, while **137276022** presented weaker and less repetitive bonds, validating the high binding affinity of top candidates.

The binding free energy calculated using MM/GBSA revealed that **24374890** had the most favorable ΔG_{total} value (-65.11 ± 24.08 kcal/mol), outperforming both experimental controls and other screened candidates. This value is significantly lower than those typically observed for similar HDAC8 inhibitors, which usually range between -40 and -55 kcal/mol [36]. The van der Waals and electrostatic contributions were dominant, suggesting optimal non-covalent anchoring within the catalytic cleft. Notably, this thermodynamic profile is also consistent with ligands reported in HDAC8 inhibitor studies for other neglected diseases [37].

In order to complement the structural and energy-based validation, we performed machine learning-based prediction of pIC₅₀ using a benchmarked ensemble model (HistGradientBoostingRegressor). **24374890** showed the best predicted pIC₅₀ value (8.1), indicative of high potency compared to the reference compound (7.5). The findings support other recent AI-based reports in which pIC₅₀ predictions were incorporated into screening pipelines for parasitic and microbial targets [38]. In contrast to other machine learning models that depend exclusively on topological descriptors or simple fingerprints, our model uses molecular, dynamic, and DFT-derived characteristics, providing for a multidimensional prediction framework.

In comparison with other *in silico* work on SmHDAC8 and its homologous parasitic HDACs, the approach here is

unique in the combination of quantum chemical analysis, dynamic simulations over an extended length scale, and AI-based prediction in a unified pipeline. One recent work highlighted the synergistic benefit in model quality, as well as lead prioritization, through combined usage [38], in agreement with our findings.

In summary, the convergence among all the computational methods strongly favors **24374890** as an extremely good candidate for experimental verification. Its uniform superiority in molecular docking, DFT energies, conformational dynamics, MM/GBSA binding energy, and pIC₅₀ prediction highlights its potential as an HDAC8 inhibitor with selectivity and high potency for the treatment of schistosomiasis.

4. Methodology

4.1. Virtual Screening of Diverse-Lib Database. The investigation began with an extensive virtual screening exercise to identify possible HDAC8 inhibitors of *Schistosoma mansoni*. A structured compound library, denoted as Diverse-Lib, comprising structurally diverse, drug-like small molecules, was made ready for screening runs. The library was constructed by the MTiOpenScen web server developer to cover a wide chemical diversity while preserving preferable physicochemical properties [39–41]. These filtered molecules were screened using the MTiOpenScen web server against the SmHDAC8 crystallographic structure (PDB ID: 7P3S) to select molecules that possessed ideal interaction schemes [25]. The protein was prepared in the UCSF Chimera program [42]. The primary consideration for shortlisting high-scoring compounds was binding affinities to target molecules that showed interaction within SmHDAC8's zinc-binding domain and adjacent active site pockets.

4.2. Quantum Chemical Calculations Based on DFT. For optimizing the highest-ranked ligands and investigating their electronic properties, DFT calculations were executed with the aid of the PySCF package of quantum chemistry [43, 44]. The aim was to investigate molecular reactivity and electronic transitions of ligands through FMO analysis.

The SDF file of every chosen compound was initially read using RDKit's molecular reader to obtain atomic coordinates and connectivity [45, 46]. A robust geometry extraction was performed from the 3D conformer that is embedded to obtain an entire atomic list as well as spatial locations in correspondence therewith. The coordinates were further converted to PySCF-compatible format to construct the molecular system for quantum chemical modeling.

Every molecule was described by employing the correlation-consistent polarized valence double-zeta (cc-pVDZ) basis set, which provides an appropriate trade-off between orbital accuracy and computational efficiency. DFT was performed within the B3LYP functional, known to provide well-reproduced orbital energies and electronic band gaps for systems of biological molecules. The position of the SCF total energy was converged using an aggressive

convergence criterion, and the electron density was optimally iterated.

The number of occupied orbitals was calculated automatically from the number of electrons within the molecule, which gave rise to indices of the HOMO and LUMO. The molecular orbital coefficients of DFT output were next used to create cube files of HOMO and LUMO electron densities, which were used for visualization and were saved as homo.cube and lumo.cube.

In addition, orbital energies were expressed in terms of Hartree and were further converted to electronvolts (eV) using the Hartree-to-eV conversion factor (27.2114 eV/Hartree). The calculated HOMO–LUMO energy gap was an important descriptor of reactivity and molecular stability, where smaller gaps were associated with higher reactivity and also higher interaction energy with the target protein. These results were archived for every ligand as a reference for further choice and chemoprioritization.

4.3. Redocking and Binding Pose Refinement. Prior to molecular docking, the crystal structure of SmHDAC8 (PDB ID: 7P3S) was restored with UCSF Chimera 1.17. Virtual hydrogen atoms were incorporated, and protonation states for titratable residues were set at physiological pH (7.4) with Chimera's automatic protonation scheme. This provided proper orientation of ionizable side chains, correct charge distribution, and consistent presentation of the catalytic environment for subsequent docking simulations. [42, 47]. Ligands that established stable bidentate contacts with the Zn²⁺ ion and adequately occupied the hydrophobic channel were selected for further detailed analysis. The docking grid box was centered at $x = 45.08$, $y = 36.84$, and $z = 87.51$, with dimensions of $20 \times 20 \times 20 \text{ \AA}$ to fully encompass the SmHDAC8 active site and adjacent binding regions. A grid spacing of 0.375 \AA was applied to achieve optimal resolution and accurate sampling of ligand conformations within the catalytic pocket. The reproducibility of binding modes and consistency of interactions between replicates were evaluated to identify the most reliable candidates for subsequent MD simulations.

4.4. MD Simulations. MD simulations are carried out to assess the temporal stability and conformational dynamics of top-ranking protein–ligand complexes [48]. The AMBER 24 simulation package was used for the entire production run [49, 50]. The parameterization and system construction were performed using LEaP, the molecular modeling tool bundled within AMBER. The ff14SB force field for proteins and the general AMBER force field (GAFF2) for ligands were used to parameterize the protein–ligand complexes initially [51]. Ligand charges were generated from a restrained electrostatic potential (RESP) using antechamber utilities [52].

Each box was solvated using TIP3P water molecules, extending at least 12 \AA beyond the boundary of the complex. Counterions were included to neutralize the system's overall charge. The systems were prepared, and then energy minimization was conducted in two phases: in the first,

constraints were applied to the protein–ligand complex to relax the water molecules, and in the second, there was full system minimization.

The equilibrium was performed in the NVT ensemble (constant number of particles, volume, and temperature) for 500 ps and further equilibrated using the NPT ensemble (constant pressure) for 1 ns, gradually warming up the system to 300 K. Temperature regulation used Langevin dynamics, while Berendsen's barostat regulated pressure to 1 atm. The hydrogen bonds were constrained using the SHAKE algorithm, enabling an integration timestep of 2 fs [53].

Production MD simulations were run for 500 ns for each of the associated complexes. Trajectory snapshots were saved every 50 ps, and simulation stability was tracked by using RMSD and total energy profiles. Analyses of local flexibility and dynamic interactions were determined using RMSF and last pose analysis, respectively.

4.5. Hydrogen Bond Analysis. Hydrogen bond analysis was performed with visual molecular dynamics (VMD) for assessing inter- and intramolecular interactions at intervals during the MD simulations carried out with AMBER. The AMBER topology (.prmtop) and trajectory (.nc) files were opened within VMD, and hydrogen bonds were detected using VMD's built-in analysis module for HBonds. A donor–acceptor distance of 3.0 Å and an angle of 150° were used to arbitrate valid hydrogen bonds. This analysis was carried out on a frame-by-frame basis throughout the trajectory to monitor dynamic variations in hydrogen bonding moles. The output consisted of both the quantity and identity of hydrogen bonds over time, allowing for the assessment of the stability and interactional behavior of the system. This approach enabled a simple, repeatable analysis of hydrogen bonding directly from AMBER simulation output.

4.6. PCA and FEL. To obtain the fundamental motions of the protein–ligand systems within the simulation, PCA was used to process the MD trajectories in the Geo-measures PyMOL plugin [31, 54]. The covariance of atomic displacements (Ca atoms) was calculated, and eigenvectors of the top principal components were selected. The first two PCs were projected to create two-dimensional conformational space. FEL was obtained by projecting onto PC1 and PC2, and Boltzmann distribution was used to calculate free energy in the Geo-measures PyMOL plugin. Energy minima were distinguished, which were the most stable conformational states visited throughout the simulation. Conformers were taken out for comparative study.

4.7. Superimposition of Extracted Minima and Initial Binding Modes. To confirm preservation of ligand conformation within the binding site, every one of these FEL-derived minimum structures was superimposed onto its respective starting docked pose in USCF Chimera [42]. Structural drift was measured quantitatively in terms of RMSD for ligand

atoms. Consistent alignment and retention of significant contacts confirmed conformational stability, whereas significant deviation denoted ligand repositioning as an indication of dynamic sampling.

4.8. Binding Free Energy Calculation of MM/GBSA. Relative binding affinities of the ligands were calculated using the MM/GBSA method. Equilibrated snapshots from MD trajectories were used to obtain energy components, which comprised van der Waals, electrostatics, polar solvation, and nonpolar solvation energy. 1000 evenly distributed frames were taken from the last 50 ns of the production run and were used to perform MM/GBSA analysis, and the average free energy was calculated to rank ligands based on predicted affinities for binding. These calculations were carried out using the MMPBSA.py module, which is a standard utility provided within the AMBER Tools suite [50, 55].

The binding free energy (ΔG_{bind}) of the protein–ligand complex was computed using the MM/GBSA approach as implemented in the Amber suite, according to the following equation:

$$\Delta G_{\text{bind}} = \Delta G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}}). \quad (1)$$

Each free energy term (G) was estimated as

$$G = E_{\text{gas}} + G_{\text{Solv}} - TS. \quad (2)$$

Since the entropy (TS) contribution was not considered due to computational expense, the equation simplifies to:

$$\Delta G_{\text{bind}} = \Delta E_{\text{gas}} + \Delta G_{\text{Solv}}. \quad (3)$$

The gas-phase interaction energy (ΔE_{gas}) and solvation free energy (ΔG_{Solv}) were further decomposed as follows:

$$\begin{aligned} \Delta E_{\text{gas}} &= \Delta E_{\text{vdw}} + \Delta E_{\text{ele}}, \\ \Delta G_{\text{Solv}} &= \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}}. \end{aligned} \quad (4)$$

Here, ΔE_{vdw} : van der Waals interaction energy, ΔE_{ele} : electrostatic interaction energy, ΔG_{polar} : polar solvation free energy (GB model), $\Delta G_{\text{nonpolar}}$: nonpolar solvation energy, estimated from solvent accessible surface area (SASA).

4.9. Machine Learning-Based Prediction of Biological Activity. A cheminformatics-led machine learning workflow to predict the biological potency of candidate molecules against *Schistosoma mansoni* HDAC8 was established using experimentally verified bioactivity data. A combination of molecular descriptor extraction, bioactivity conversion, as well as supervised regression modeling was used to forecast compound inhibitory potency as pIC₅₀ values.

4.9.1. Data Collection and Curation. First, bioactivity information for an appropriate HDAC8 homolog (Uniprot ID: A5H660) was downloaded from the ChEMBL database with the chembl_webresource_client [56–58]. The records were

filtered to include only assay type “B” and IC_{50} values expressed in nanomolar to allow for standard measure unit consistency. Non-numeric, duplicate, or non-nanomolar records were discarded to filter down to the final dataset. The curated final dataset comprised ChEMBL molecule IDs, standard IC_{50} values, and related assay metadata.

4.9.2. Molecular Structure Retrieval and Cleaning. The standard SMILES of every molecule were retrieved from ChEMBL’s endpoint of molecule structures and cross-matched to the bioactivity dataset using distinct molecule ChEMBL IDs. The molecules that were missing or had invalid SMILES were filtered. Structures obtained were used to calculate physicochemical descriptors and create molecular fingerprints necessary for model input.

4.9.3. IC_{50} to pIC_{50} Transformation. In order to achieve logarithmic scaling of biological activity data and reduce regression learning skewness, IC_{50} values (in nM) were converted to pIC_{50} values.

4.9.4. Feature Extraction. Molecular weight, logP, hydrogen bond donors (HBD), and acceptors (HBA) molecular descriptors were calculated using RDKit’s Descriptors package [45, 46]. A compound was checked against Lipinski’s Rule of Five (Ro5) to filter for drug-likeness. A compound was counted as Ro5-compliant if it complied with three or more of these parameters: $MW < 500$ Da, $\log P < 5$, $HBD < 5$, and $HBA < 10$.

4.9.5. Data Preparation. The Ro5-compliant molecules were kept to allow predictions to center on pharmacologically potent candidates. The ending dataset used to perform machine learning included SMILES strings, calculated molecular properties, and related pIC_{50} values. Exploratory data analysis was done in a series of steps to graphically visualize the distribution of pIC_{50} and descriptor variation.

4.9.6. Model Construction. The exact model fitting section does not exist in the first 70 cells under review, yet the workflow was definitely aimed to end in supervised regression, probably using algorithms like Random Forest, Support Vector Regressor, or Gradient Boosting Regressor, which are commonly utilized in pIC_{50} prediction within cheminformatics pipelines [59–61].

4.9.7. Prediction of Query Compounds. A collection of query molecules from Diverse-Lib was transformed to molecular descriptors employing identical descriptor functions. The descriptors were fed to the trained model for developing predicted pIC_{50} values. The compounds with $pIC_{50} > 6.5$ were shortlisted as possible leads with significant inhibitory activity.

5. Conclusion

This work provides a unified computational workflow consisting of structure-based virtual screening, structure analysis using quantum chemistry, MD simulation, estimation of binding free energy, and activity prediction using machine learning for discovering *Schistosoma mansoni* HDAC8 inhibitors. Out of several screened molecules, 24374890 appeared as the best candidate with the highest binding affinity ($\Delta G_{total} = -65.11$ kcal/mol), highest electronic stability (HOMO–LUMO gap = 4.143 eV), and highest potency ($pIC_{50} = 8.1$). Supporting analyses, such as RMSD, RMSF, PCA, FEL, and hydrogen bond occupancy, verified the compound’s appropriate interaction stability and conformational stability at the catalytic pocket. These results cumulatively indicate that compound 24374890 is a good lead compound with SmHDAC8-selective and energetically favorable binding. This combination of MM, quantum descriptors, and AI prediction folds into a strong, repeatable, and multidimensional framework for discovering antiparasitic drugs. Experimental validation, such as in vitro enzyme inhibition assays and in vivo cytotoxicity profiling, remains for future experiments for validating the computed predictions. This workflow also provides a transferable template for discovering epigenetic inhibitors against other neglected tropical targets.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No funding was received for this manuscript.

Acknowledgments

The authors extend their appreciation to the Northern Border University, Saudi Arabia, for supporting this work through project number (NBU-CRP-2025-2043).

Supporting Information

Additional supporting information can be found online in the Supporting Information section. (*Supporting Information*)

Supporting File: Table S1: List of compounds obtained from virtual screening against SmHDAC8 within the binding energy range of -9.5 to -7 kcal/mol.

Figure S1: Representative low-energy conformations of ligand-bound protein complexes extracted from the free energy landscape (FEL).

Figure S2: Violin plot showing intracluster Tanimoto similarity distributions for each cluster (indexed 0–9). The width of each violin represents the density of similarity values within the cluster.

Figure S3: Comparison of R^2 scores across 22 regression models using cross-validation and independent test sets.

Figure S4: Superimposition of the experimental (brown) and redocked (cyan) poses of the cocrystallized SmHDAC8 inhibitor (PDB ID: 7P3S), showing RMSD = 0.7 Å.

Figure S5: RMSD of the apo-SmHDAC8 backbone over 500 ns.

Figure S6: RMSF of the apo-SmHDAC8 backbone over 500 ns.

References

- [1] World Health Organization, *Schistosomiasis* (2023), <https://www.who.int/news-room/fact-sheets/detail/schistosomiasis>.
- [2] D. G. Colley, A. L. Bustinduy, W. E. Secor, and C. H. King, “Human Schistosomiasis,” *The Lancet* 383, no. 9936 (2014): 2253–2264, [https://doi.org/10.1016/S0140-6736\(13\)61949-2](https://doi.org/10.1016/S0140-6736(13)61949-2).
- [3] N. Vale, M. J. Gouveia, G. Rinaldi, P. J. Brindley, F. Gärtner, and J. M. Correia Da Costa, “Praziquantel for Schistosomiasis: Single-Drug Metabolism Revisited, Mode of Action, and Resistance,” *Antimicrobial Agents and Chemotherapy* 61, no. 5 (2017): e02582, <https://doi.org/10.1128/AAC.02582-16>.
- [4] G. Eastham, *A Review of Praziquantel Resistance in Schistosomes and the Future of Schistosomiasis Control* (2025).
- [5] N. C. Lo, F. S. M. Bezerra, D. G. Colley, et al., “Review of 2022 WHO Guidelines on the Control and Elimination of Schistosomiasis,” *The Lancet Infectious Diseases* 22, no. 11 (2022): e327–e335, [https://doi.org/10.1016/S1473-3099\(22\)00221-3](https://doi.org/10.1016/S1473-3099(22)00221-3).
- [6] W. Wang and Y. Liang, “Mass Drug Administration (MDA) for Schistosomiasis,” *Journal of Infectious Diseases* 211, no. 5 (2015): 848–849, <https://doi.org/10.1093/infdis/jiu506>.
- [7] N. Ladislav, M. Aimable, O. Alison, et al., “From Mapping to Elimination: A Comprehensive Overview of Rwanda’s Schistosomiasis Program,” *International Journal of Infectious Diseases* 152 (2025): 107799, <https://doi.org/10.1016/j.ijid.2025.107799>.
- [8] A. V. Diaz, M. Walker, and J. P. Webster, “Reaching the World Health Organization Elimination Targets for Schistosomiasis: The Importance of a One Health Perspective,” *Philosophical Transactions of the Royal Society B* 378, no. 1887 (2023): 20220274, <https://doi.org/10.1098/rstb.2022.0274>.
- [9] S. H. Sokolow, C. L. Wood, I. J. Jones, et al., “Global Assessment of Schistosomiasis Control Over the Past Century Shows Targeting the Snail Intermediate Host Works Best,” *PLoS Neglected Tropical Diseases* 10, no. 7 (2016): e0004794, <https://doi.org/10.1371/journal.pntd.0004794>.
- [10] A. Cabezas-Cruz, J. Lancelot, S. Caby, G. Oliveira, and R. J. Pierce, “Epigenetic Control of Gene Function in Schistosomes: A Source of Therapeutic Targets?” *Frontiers in Genetics* 5 (2014): 317, <https://doi.org/10.3389/fgene.2014.00317>.
- [11] V. C. Carneiro, I. C. De Abreu Da Silva, E. J. L. Torres, et al., “Epigenetic Changes Modulate Schistosome Egg Formation and are a Novel Target for Reducing Transmission of Schistosomiasis,” *PLoS Pathogens* 10, no. 5 (2014): e1004116, <https://doi.org/10.1371/journal.ppat.1004116>.
- [12] M. Knight, W. Ittiprasert, H. D. Arican-Goktas, and J. M. Bridger, “Epigenetic Modulation, Stress and Plasticity in Susceptibility of the Snail Host, *Biomphalaria glabrata*, to *Schistosoma Mansoni* Infection,” *International Journal for Parasitology* 46, no. 7 (2016): 389–394, <https://doi.org/10.1016/j.ijpara.2016.03.003>.
- [13] W. Liu, “Epigenetics in Schistosomes: What We Know and What We Need Know,” *Frontiers in Cellular and Infection Microbiology* 6 (2016): 149, <https://doi.org/10.3389/fcimb.2016.00149>.
- [14] R. A. Gomes Assenço, E. Alves Mota, V. F. De Oliveira, W. De Castro Borges, and R. Guerra-Sá, “Epigenetic Markers Associated With Schistosomiasis,” *Helminthologia* 58, no. 1 (2021): 28–40, <https://doi.org/10.2478/helm-2021-0009>.
- [15] R. J. Pierce, F. Dubois-Abdesselem, S. Caby, et al., “Chromatin Regulation in Schistosomes and Histone Modifying Enzymes as Drug Targets,” *Memorias do Instituto Oswaldo Cruz* 106, no. 7 (2011): 794–801, <https://doi.org/10.1590/S0074-02762011000700003>.
- [16] M. Marek, S. Kannan, A.-T. Hauser, et al., “Structural Basis for the Inhibition of Histone Deacetylase 8 (HDAC8), A Key Epigenetic Player in the Blood Fluke *Schistosoma Mansoni*,” *PLoS Pathogens* 9, no. 9 (2013): e1003645, <https://doi.org/10.1371/journal.ppat.1003645>.
- [17] T. Bayer, A. Chakrabarti, J. Lancelot, et al., “Synthesis, Crystallization Studies, and In Vitro Characterization of Cinnamic Acid Derivatives as *Sm* HDAC8 Inhibitors for the Treatment of Schistosomiasis,” *ChemMedChem* 13, no. 15 (2018): 1517–1529, <https://doi.org/10.1002/cmdc.201800238>.
- [18] T. Heimbürg, A. Chakrabarti, J. Lancelot, et al., “Structure-Based Design and Synthesis of Novel Inhibitors Targeting HDAC8 From *Schistosoma mansoni* for the Treatment of Schistosomiasis,” *Journal of Medicinal Chemistry* 59, no. 6 (2016): 2423–2435, <https://doi.org/10.1021/acs.jmedchem.5b01478>.
- [19] J. Lancelot, S. Caby, F. Dubois-Abdesselem, et al., “Schistosoma Mansoni Sirtuins: Characterization and Potential as Chemotherapeutic Targets,” *PLoS Neglected Tropical Diseases* 7, no. 9 (2013): e2428, <https://doi.org/10.1371/journal.pntd.0002428>.
- [20] L. Pagliazzo, S. Caby, J. Lancelot, et al., “Histone Deacetylase 8 Interacts With the Gtpase SmRho1 in *Schistosoma Mansoni*,” *PLoS Neglected Tropical Diseases* 15, no. 11 (2021): e0009503, <https://doi.org/10.1371/journal.pntd.0009503>.
- [21] K. C. L. Whatley, G. Padalino, H. Whiteland, et al., “The Repositioning of Epigenetic Probes/Inhibitors Identifies New Anti-Schistosomal Lead Compounds and Chemotherapeutic Targets,” *PLoS Neglected Tropical Diseases* 13, no. 11 (2019): e0007693, <https://doi.org/10.1371/journal.pntd.0007693>.
- [22] L. Anderson, M. S. Amaral, F. Beckedorff, et al., “Schistosoma Mansoni Egg, Adult Male and Female Comparative Gene Expression Analysis and Identification of Novel Genes by RNA-Seq,” *PLoS Neglected Tropical Diseases* 9, no. 12 (2015): e0004334, <https://doi.org/10.1371/journal.pntd.0004334>.
- [23] J. J. Collins Iii, B. Wang, B. G. Lambrus, M. E. Tharp, H. Iyer, and P. A. Newmark, “Adult Somatic Stem Cells in the Human Parasite *Schistosoma Mansoni*,” *Nature* 494, no. 7438 (2013): 476–479, <https://doi.org/10.1038/nature11924>.
- [24] N. M. El-Lakkany, S. H. S. El-Din, A.-N. A.-A. Sabra, and O. A. Hammam, “Pharmacodynamics of Mefloquine and Praziquantel Combination Therapy in Mice Harboring Juvenile and Adult *Schistosoma Mansoni*,” *Memorias do Instituto Oswaldo Cruz* 106, no. 7 (2011): 814–822, <https://doi.org/10.1590/S0074-02762011000700006>.
- [25] E. Ghazy, T. Heimbürg, J. Lancelot, et al., “Synthesis, Structure-Activity Relationships, Cocrystallization and Cellular Characterization of Novel smHDAC8 Inhibitors for the

- Treatment of Schistosomiasis,” *European Journal of Medicinal Chemistry* 225 (2021): 113745, <https://doi.org/10.1016/j.ejmech.2021.113745>.
- [26] S. D. Melman, M. L. Steinauer, C. Cunningham, et al., “Reduced Susceptibility to Praziquantel Among Naturally Occurring Kenyan Isolates of *Schistosoma Mansoni*,” *PLoS Neglected Tropical Diseases* 3, no. 8 (2009): e504, <https://doi.org/10.1371/journal.pntd.0000504>.
- [27] N. C. Lo, D. G. Addiss, P. J. Hotez, et al., “A Call to Strengthen the Global Strategy Against Schistosomiasis and Soil-Transmitted Helminthiasis: The Time is now,” *The Lancet Infectious Diseases* 17, no. 2 (2017): e64–e69, [https://doi.org/10.1016/S1473-3099\(16\)30535-7](https://doi.org/10.1016/S1473-3099(16)30535-7).
- [28] J. Y. Kim, S. Y. Han, J. Yoo, et al., “HDAC8-Selective Inhibition by PCI-34051 Enhances the Anticancer Effects of ACY-241 in Ovarian Cancer Cells,” *International Journal of Molecular Sciences* 23, no. 15 (2022): 8645, <https://doi.org/10.3390/ijms23158645>.
- [29] S. Balasubramanian, J. Ramos, W. Luo, M. Sirisawad, E. Verner, and J. J. Buggy, “A Novel Histone Deacetylase 8 (HDAC8)-Specific Inhibitor PCI-34051 Induces Apoptosis in T-Cell Lymphomas,” *Leukemia* 22, no. 5 (2008): 1026–1034, <https://doi.org/10.1038/leu.2008.9>.
- [30] S. Banerjee, B. Ghosh, T. Jha, and N. Adhikari, “A Patent Review of Histone Deacetylase 8 (HDAC8) Inhibitors (2013–present),” *Expert Opinion on Therapeutic Patents* 34, no. 10 (2024): 1019–1045, <https://doi.org/10.1080/13543776.2024.2391289>.
- [31] W. L. DeLano, “Pymol: An Open-Source Molecular Graphics Tool,” *CCP4 Newsletter on Protein Crystallography* 40, no. 1 (2002): 82–92.
- [32] D. Studio, *Discovery Studio* (2008).
- [33] S. C. Ja’afaru, A. Uzairu, M. S. Sallau, G. I. Ndukwe, M. T. Ibrahim, and A. U. Danazumi, “Structure-Based Design of Potential Anti-Schistosomiasis Agent Targeting SmHDAC8: an In Silico Approach Utilizing QSAR, MD Simulation and ADMET Prediction,” *Chemistry Africa* 7, no. 2 (2024): 725–745, <https://doi.org/10.1007/s42250-023-00777-2>.
- [34] C. V. Simoben, E. Ghazy, P. Zeyen, et al., “Binding Free Energy (BFE) Calculations and Quantitative Structure–Activity Relationship (QSAR) Analysis of *Schistosoma Mansoni* Histone Deacetylase 8 (smHDAC8) Inhibitors,” *Molecules* 26, no. 9 (2021): 2584, <https://doi.org/10.3390/molecules26092584>.
- [35] C. Simoben, D. Robaa, A. Chakrabarti, et al., “A Novel Class of *Schistosoma Mansoni* Histone Deacetylase 8 (HDAC8) Inhibitors Identified by Structure-Based Virtual Screening and In Vitro Testing,” *Molecules* 23, no. 3 (2018): 566, <https://doi.org/10.3390/molecules23030566>.
- [36] O. S. Onile, O. Raji, V. Omoboyede, et al., “Structure-Based Discovery of Phytocompounds From *Azadirachta Indica* as Potential Inhibitors of Thioredoxin Glutathione Reductase in *Schistosoma mansoni*,” *Cell Biochemistry and Biophysics* 83, no. 2 (2024): 1677–1688, <https://doi.org/10.1007/s12013-024-01577-2>.
- [37] D. R. Chowdhury, D. Sen, S. J. Roy, S. Majumdar, and S. Debnath, “Discovery of Potential Selective HDAC8 Binders: Using Combination of Pharmacophore-Based–Structure-Based Virtual Screening, ADME Filtration, and Molecular Dynamics Studies,” *ChemistrySelect* 10, no. 15 (2025): e202500646, <https://doi.org/10.1002/slct.202500646>.
- [38] S. C. Ja’afaru, A. Uzairu, V. K. Mishra, M. S. Sallau, M. T. Ibrahim, and A. Dubey, “Virtual Screening and Molecular Dynamics Studies of Novel Small Molecules Targeting *Schistosoma mansoni* DHODH: Identification of Potential Inhibitors,” *In Silico Pharmacology* 12, no. 2 (2024): 113, <https://doi.org/10.1007/s40203-024-00281-6>.
- [39] C. M. Labbé, J. Rey, D. Lagorce, et al., “MTiOpenScreen: A Web Server for Structure-Based Virtual Screening,” *Nucleic Acids Research* 43, no. W1 (2015): W448–W454, <https://doi.org/10.1093/nar/gkv306>.
- [40] Y. Wang, T. Suzek, J. Zhang, et al., “PubChem BioAssay: 2014 Update,” *Nucleic Acids Research* 42, no. D1 (2014): D1075–D1082, <https://doi.org/10.1093/nar/gkt978>.
- [41] D. Lagorce, J. Maupetit, J. Baell, et al., “The FAF-Drugs2 Server: A Multistep Engine to Prepare Electronic Chemical Compound Collections,” *Bioinformatics* 27, no. 14 (2011): 2018–2020, <https://doi.org/10.1093/bioinformatics/btr333>.
- [42] E. F. Pettersen, T. D. Goddard, C. C. Huang, et al., “UCSF Chimera—A Visualization System for Exploratory Research and Analysis,” *Journal of Computational Chemistry* 25, no. 13 (2004): 1605–1612, <https://doi.org/10.1002/jcc.20084>.
- [43] Q. Sun, T. C. Berkelbach, N. S. Blunt, et al., “PySCF: The Python-Based Simulations of Chemistry Framework,” *Wiley Interdisciplinary Reviews: Computational Molecular Science* 8, no. 1 (2018): e1340, <https://doi.org/10.1002/wcms.1340>.
- [44] K. Hanasaki, Z. A. Ali, M. Choi, M. Del Ben, and B. M. Wong, “Implementation of Real-Time TDDFT for Periodic Systems in the Open-Source PySCF Software Package,” *Journal of Computational Chemistry* 44, no. 9 (2023): 980–987, <https://doi.org/10.1002/jcc.27058>.
- [45] A. P. Bento, A. Hersey, E. Félix, et al., “An Open Source Chemical Structure Curation Pipeline Using RDKit,” *Journal of Cheminformatics* 12 (2020): 1–16, <https://doi.org/10.1186/s13321-020-00456-1>.
- [46] G. Landrum, “RDKit: A Software Suite for Cheminformatics, Computational Chemistry, and Predictive Modeling,” *Greg Landrum* 8, no. 31 (2013): 5281.
- [47] O. Trott and A. J. Olson, “AutoDock Vina: Improving the Speed and Accuracy of Docking With a New Scoring Function, Efficient Optimization, and Multithreading,” *Journal of Computational Chemistry* 31, no. 2 (2010): 455–461, <https://doi.org/10.1002/jcc.21334>.
- [48] T. T. H. Do, T. U. Nguyen, T. T. H. Nguyen, et al., “Essential Oils From the Leaves, Stem, and Roots of *Blumea lanceolaria* (Roxb.) Druce in Vietnam: Determination of Chemical Composition, and In Vitro, In Vivo, and In Silico Studies on Anti-Inflammatory Activity,” *Molecules* 27, no. 22 (2022): 7839, <https://doi.org/10.3390/molecules27227839>.
- [49] D. A. Case, T. E. Cheatham III, T. Darden, et al., “The Amber Biomolecular Simulation Programs,” *Journal of Computational Chemistry* 26, no. 16 (2005): 1668–1688, <https://doi.org/10.1002/jcc.20290>.
- [50] D. A. Case, H. M. Aktulga, K. Belfon, et al., “AmberTools,” *Journal of Chemical Information and Modeling* 63, no. 20 (2023): 6183–6191, <https://doi.org/10.1021/acs.jcim.3c01153>.
- [51] G. A. Özpinar, W. Peukert, and T. Clark, “An Improved Generalized AMBER Force Field (GAFF) for Urea,” *Journal of Molecular Modeling* 16, no. 9 (2010): 1427–1440, <https://doi.org/10.1007/s00894-010-0650-7>.
- [52] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, “A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model,” *Journal of Physical Chemistry* 97, no. 40 (1993): 10269–10280, <https://doi.org/10.1021/j100142a004>.
- [53] Y. Lin, D. Pan, J. Li, L. Zhang, and X. Shao, “Application of Berendsen Barostat in Dissipative Particle Dynamics for

- Nonequilibrium Dynamic Simulation,” *The Journal of Chemical Physics* 146, no. 12 (2017): 124108, <https://doi.org/10.1063/1.4978807>.
- [54] L. P. Kagami, G. M. das Neves, L. F. S. M. Timmers, R. A. Caceres, and V. L. Eifler-Lima, “Geo-Measures: A PyMOL Plugin for Protein Structure Ensembles Analysis,” *Computational Biology and Chemistry* 87 (2020): 107322, <https://doi.org/10.1016/j.compbiolchem.2020.107322>.
- [55] B. R. Miller III, T. D. McGee Jr, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, “MMPBSA. py: An Efficient Program for End-State Free Energy Calculations,” *Journal of Chemical Theory and Computation* 8, no. 9 (2012): 3314–3321, <https://doi.org/10.1021/ct300418h>.
- [56] A. Gaulton, L. J. Bellis, A. P. Bento, et al., “ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery,” *Nucleic Acids Research* 40, no. D1 (2012): D1100–D1107, <https://doi.org/10.1093/nar/gkr777>.
- [57] B. Zdrzil, “Fifteen Years of ChEMBL and Its Role in Cheminformatics and Drug Discovery,” *Journal of Cheminformatics* 17, no. 1 (2025): 1–9, <https://doi.org/10.1186/s13321-025-00963-z>.
- [58] T. UniProt Consortium, A. Bateman, M.-J. Martin, et al., “UniProt: The Universal Protein Knowledgebase in 2025,” *Nucleic Acids Research* 53, no. D1 (2025): D609–D617, <https://doi.org/10.1093/nar/gkae1010>.
- [59] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, “Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling,” *Journal of Chemical Information and Computer Sciences* 43, no. 6 (2003): 1947–1958, <https://doi.org/10.1021/ci034160g>.
- [60] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, “Support Vector Machines,” *IEEE Intelligent Systems and Their Applications* 13, no. 4 (1998): 18–28, <https://doi.org/10.1109/5254.708428>.
- [61] I. Babajide Mustapha and F. Saeed, “Bioactive Molecule Prediction Using Extreme Gradient Boosting,” *Molecules* 21, no. 8 (2016): 983, <https://doi.org/10.3390/molecules21080983>.