Molecular Diversity
https://doi.org/10.1007/s11030-026-11504-7

RESEARCH

Repurposing DrugBank compounds as NAD-dependent deacetylase sirtuin 2 inhibitors via QSAR modelling with gradient boosting algorithms and all-atom molecular simulations

Yassir Boulaamane[1] · Asmae Saih[2,3] · Abdelkrim Guendouzi[4] · Amal Maurady[1,5]

Received: 30 August 2025 / Accepted: 21 February 2026

Abstract
Sirtuin 2 (SIRT2), a NAD^+-dependent histone deacetylase implicated in α-synuclein aggregation, is an emerging target for disease-modifying therapies in Parkinson's disease (PD). Here, we employed an integrated computational drug-repurposing strategy to identify potent SIRT2 inhibitors from the DrugBank database. A curated set of 949 inhibitors was used to construct quantitative structure–activity relationship (QSAR) models with four gradient-boosting algorithms, yielding CatBoost as the optimal predictor ($R^2_{val}=0.74$, $Q^2_{tenfold}=0.72$). The model screened 4947 drug-like compounds, from which 97 candidates with predicted $pIC_{50}\geq 6$ were prioritized. Molecular docking against the SIRT2 crystal structure (PDB: 4RMG) revealed high-affinity binding modes for multiple hits, notably DB14822, DB03571, and DB06506, engaging conserved residues (Phe119, Tyr139, Phe190, Ile232) through hydrophobic and π-stacking interactions. ADMET profiling indicated favorable drug-likeness and acceptable pharmacokinetic/toxicity properties for most candidates. All-atom molecular dynamics simulations (250 ns) demonstrated that top ligands maintained compact, stable complexes with low RMSD, restricted radius of gyration, and minimal solvent exposure. Principal component and free energy landscape analyses confirmed constrained global motions, while MM/GBSA calculations yielded favorable binding free energies (−32.6 to −35.7 kcal/mol) for lead compounds. Given SIRT2's established role in α-synuclein aggregation and neurodegeneration, these compounds represent potential therapeutic starting points for Parkinson's disease and merit experimental validation.

Keywords SIRT2 inhibition · Parkinson's disease · Drug repurposing · QSAR · Molecular docking · Molecular dynamics

✉ Yassir Boulaamane
boulaamane.yassir@etu.uae.ac.ma

Asmae Saih
asmae.saihbio@etu.univh2c.ma

Abdelkrim Guendouzi
guendouzzi@yahoo.fr

Amal Maurady
amaurady@uae.ac.ma

[1] Laboratory of Innovative Technologies, National School of Applied Sciences of Tangier, Abdelmalek Essaadi University, Tétouan, Morocco

[2] Virology Unit, Immunovirology Laboratory, Institut Pasteur du Maroc, 20360 Casablanca, Morocco

[3] Laboratory of Biology and Health, URAC 34, Faculty of Sciences, Ben M'Sik Hassan II University of Casablanca, Casablanca, Morocco

[4] Laboratory of Chemistry: Synthesis, Properties and Applications. (LCSPA), Faculty of Sciences, University of Saida - Dr Moulay Tahar, Saida, Algeria

[5] Faculty of Sciences and Techniques, Abdelmalek Essaadi University, Tétouan, Morocco

Published online: 12 March 2026

Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder, with its global prevalence more than doubling over the past 25 years [1]. Despite affecting over 8.5 million people and causing significant disability and death, current treatments are limited to symptom relief and do not halt disease progression [2]. Long-term therapies often lose effectiveness and cause side effects, highlighting the urgent need for disease-modifying strategies that target the molecular mechanisms driving PD [3].

PD involves two main pathological features: loss of dopamine-producing neurons in the substantia nigra and accumulation of misfolded alpha-synuclein (α-syn) protein [4]. Normally involved in synaptic function, α-syn aggregates into toxic fibrils forming Lewy bodies, which disrupt cellular processes and spread between neurons [5, 6]. This aggregation drives disease progression, making it a key therapeutic target. However, direct targeting of α-syn has proven difficult, highlighting the importance of understanding upstream processes like post-translational modifications that influence its aggregation [7, 8].

Sirtuin 2 (SIRT2) has emerged as an important enzymatic regulator in the context of PD and α-synuclein pathology (Fig. 1A) [9, 10]. SIRT2 is a member of the sirtuin family (SIRT1–7 in humans), which are class III histone deacetylases characterized by a dependence on NAD^+ as a cofactor [11]. Biochemically, SIRT2's catalytic core (~275 amino acids) consists of a large Rossmann-fold domain that binds NAD^+ and a smaller zinc-binding domain; the Zn^{2+} ion, coordinated by cysteine residues, is required for structural stability but not directly for catalysis (Fig. 1B) [12]. This conserved architecture enables SIRT2 to remove acetyl groups from lysine residues on both histone and non-histone substrates [12]. Unlike the ubiquitously nuclear SIRT1, SIRT2 is primarily cytosolic and highly expressed in the brain, especially in neurons and oligodendrocytes of the striatum, cortex, and other regions. SIRT2 regulates a broad array of cellular processes like cell cycle progression, mitochondrial function, microtubule dynamics, autophagy, oxidative stress response, and aging, through its deacetylase activity. Importantly, SIRT2 levels increase with aging and have been found to be elevated in PD models, linking this enzyme to age-related neurodegenerative mechanisms [13].

Given the need for new SIRT2 inhibitors as potential disease-modifying agents in PD, computational drug repurposing offers a fast and cost-effective strategy [14, 15]. Computational methods such as quantitative structure–activity relationship (QSAR) modeling, molecular docking, and molecular dynamics (MD) simulations allow for efficient identification of promising compounds from existing drug libraries [16–18]. QSAR models trained on known inhibitors can predict bioactivity across large chemical libraries, narrowing down candidates with high likelihood of target activity. Subsequent molecular docking evaluates the binding orientation and affinity of these compounds within the target protein, while MD simulations provide insights into the stability and dynamics of the ligand–protein complex over time. Together, these approaches offer a robust and scalable platform for prioritizing potential therapeutic candidates before experimental testing.

This work uniquely integrates a high-performance CatBoost QSAR model with multi-tiered structure-based screening, including molecular docking, all-atom MD simulations, and MM/GBSA calculations. This combined strategy provides a robust and efficient pipeline for repurposing DrugBank compounds as potential SIRT2 inhibitors. The best model was then used to virtually screen the

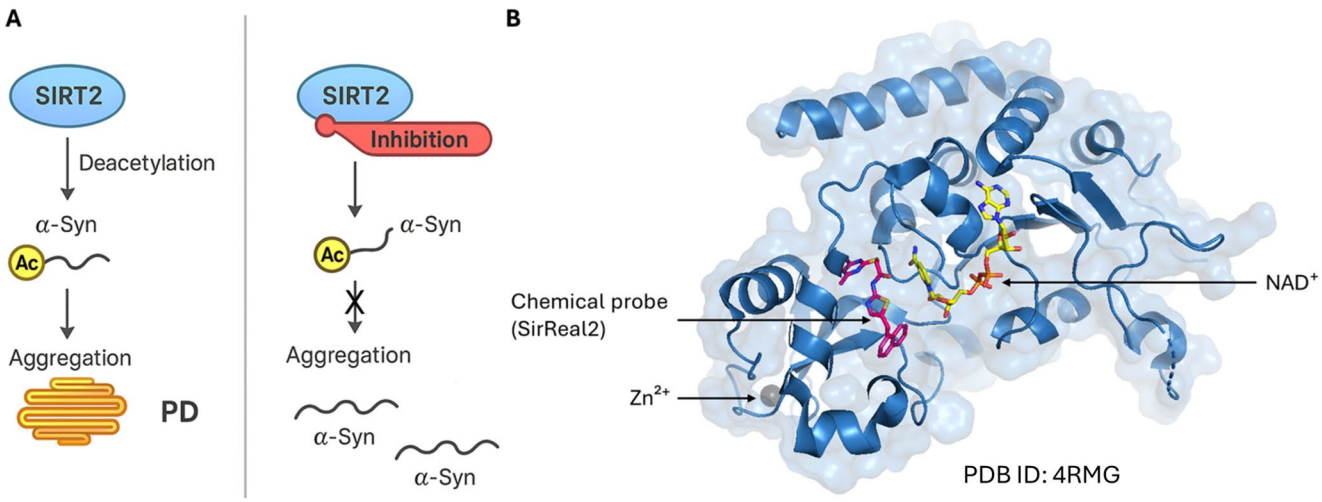

Fig. 1 SIRT2's role in α-synuclein aggregation and inhibition. **A** SIRT2 promotes α-synuclein deacetylation and aggregation in PD; inhibition prevents this process. **B** Structure of SIRT2 (PDB: 4RMG) bound to NAD^+, Zn^{2+}, and the inhibitor SirReal2

DrugBank database. High-scoring compounds were subjected to molecular docking to analyze their binding interactions with SIRT2. Prior to molecular dynamics (MD) simulations, ADMET profiling was performed to evaluate pharmacokinetic properties and toxicity risks. Selected top candidates were then evaluated using all-atom MD simulations to assess complex stability and key interactions over time. Finally, Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) binding free energy calculations were conducted to estimate the strength and stability of ligand–SIRT2 binding. This integrated workflow was designed to identify potent, drug-like SIRT2 inhibitors suitable for further in vitro and in vivo validation in the context of Parkinson's disease.

Materials and methods

Computational tools and software

All computational analyses were performed using Python 3.10 with the following packages: RDKit (v2023.09.6) [19] for molecular fingerprint generation and cheminformatics operations, Pandas (v2.0.3) [20] for data processing, Scikit-learn for model evaluation, CatBoost, XGBoost, LightGBM, and GradientBoostingRegressor for QSAR modeling, and Optuna for hyperparameter optimization [21–23]. Molecular docking was conducted using AutoDock Vina with ligand preparation performed in Meeko (v0.6.1) and Open Babel (v3.1.0) [24] for energy minimization (MMFF94 force field) [25], while PDBFixer was used for protein structure preparation [26]. Molecular dynamics simulations were performed in GROMACS (v2025.1) [27] using the CHARMM36 force field for protein and CGenFF for ligand parameterization, with MM/GBSA binding free energy calculations conducted via gmx_MMPBSA [28, 29]. Pharmacokinetic and toxicity predictions were obtained from the DeepPK web server (https://biosig.lab.umaryland.edu/deeppk/). All analyses used fixed random seeds where applicable to ensure reproducibility.

QSAR model development

Bioactivity data processing

The bioactivity data of SIRT2 inhibitors was obtained from the ChEMBL_35 database (ID: CHEMBL4462), comprising 1,806 compounds with experimentally determined IC_{50} values (Fig. 2) [30]. Most of the ChEMBL IC_{50} values used to train the QSAR model originate from in-vitro, cell-free biochemical deacetylase assays performed on recombinant human SIRT2, with fluorescence-based formats being the predominant assay type. All IC_{50} values were consistently reported in nanomolar concentrations; a small number of records with inconsistent or ambiguous units were excluded to maintain data uniformity and reliability. Data preprocessing was carried out using Pandas (v2.0.3) [31]. Compounds that did not report an exact activity value (left or right censored) were excluded to ensure consistency, and duplicate entries were aggregated by computing the mean IC_{50} to reduce variability arising from different experimental conditions. All IC_{50} values were converted to molar units and

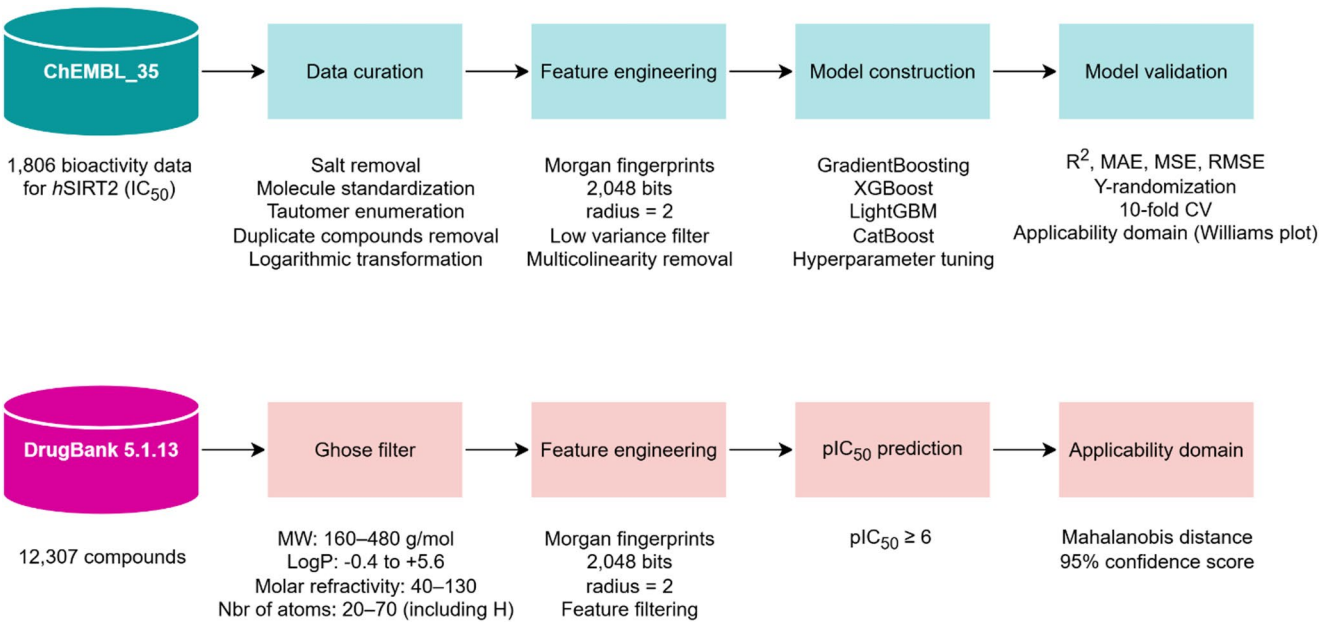

Fig. 2 Workflow for QSAR model development and virtual screening of hSIRT2 inhibitors

subsequently transformed into pIC_{50} to ensure a uniform activity scale. After curation, the final dataset included 949 compounds. Circular Extended-Connectivity Fingerprints (ECFP4) were generated using the Morgan algorithm with a radius of 2 and a length of 2,048 bits, as implemented in the RDKit cheminformatics toolkit (v2023.09.6) [19]. To ensure a balanced distribution of activity values, the pIC_{50} values were binned into five intervals: [2.35, 4.45, 4.98, 5.43, 6.24, 10.15]. The dataset was split into 80% for model development and 20% as an external test set. The 80% subset was further divided into training and validation sets. To address high dimensionality and reduce sparsity, feature selection was performed using a low variance threshold (0.1) and a pairwise correlation filter (threshold = 0.9).

Machine learning models

Four gradient boosting algorithms were employed in this study: Gradient Boosting Decision Trees (GBDT), XGBoost, LightGBM, and CatBoost [21]. All are based on the principle of boosting, where an ensemble of weak learners, typically decision trees, is built sequentially to minimize prediction errors. At each iteration, a new tree is trained to approximate the negative gradient of the loss function with respect to the model's current predictions, thereby improving overall performance (Fig. S1 in Supporting Information). GBDT represents classical implementation, while XGBoost enhances this by incorporating regularization techniques and optimized handling of missing values. LightGBM introduces a histogram-based algorithm and leaf-wise tree growth strategy for improved speed and memory efficiency. CatBoost, on the other hand, addresses the challenges of categorical data and overfitting through ordered boosting and advanced encoding schemes. These implementations are highly efficient and well-suited for capturing complex, nonlinear relationships in cheminformatics datasets.

Baseline models were trained using tenfold stratified cross-validation with a fixed random seed. Model performance was evaluated using R^2, RMSE, MAE, and MSE. The best-performing model was further optimized via Optuna for hyperparameter tuning [23]. To validate model robustness, Y-randomization was performed with 100 iterations, and the average R^2 on randomized data was recorded to confirm that performance was not due to chance. Final model evaluation was based on mean tenfold CV R^2. The applicability domain (AD) was assessed using a Williams plot, combining leverage values from PCA (20 components) with standardized residuals to identify potential outliers [32]. Predictions were subsequently made on a DrugBank subset filtered using the Ghose criteria, and prediction confidence was quantified using Mahalanobis distance [33, 34]. Only predictions within the AD at a 95% confidence level were retained.

Structure-based screening

Protein and ligand dataset preparation

Several X-ray crystal structures of SIRT2 are available, many of which are complexes with inhibitors. To avoid biasing ligand placement during docking, we selected the structure with PDB ID: 4RMG, resolved at 1.88 Å, which contains the SirReal2 inhibitor and the NAD^+ cofactor [35]. OpenMM's PDBFixer (https://github.com/openmm/pdbfixer) was used to model missing residues and correct side chains. The selected DrugBank hits were energy-minimized using the MMFF94 force field in Open Babel (v3.1.0), and prepared using the Meeko (v0.6.1) package (https://github.com/forlilab/Meeko) to assign charges, atom types, and convert both receptor and ligands to PDBQT format [24].

Molecular docking protocol

To validate the docking protocol, SirReal2 was downloaded from PubChem database and prepared using the same procedure as the DrugBank ligands [36]. It was then redocked into the SIRT2 binding site to evaluate the ability of the docking tool to reproduce the native binding conformation. The grid box was centered on SirReal2 using PyMOL, with coordinates set to (x = −17.50, y = −26.54, z = 13.18) and a box size of 20 Å in each dimension. The remaining DrugBank hits were subsequently docked using the VinaScreen Python script to automate the docking workflow.

Pharmacokinetic/toxicity assessment

Recent analyses of clinical drug development indicate that about 40–50% of failures are due to lack of clinical efficacy, often linked to poor pharmacokinetics and toxicity; toxicity accounts for around 30%, and poor pharmacokinetic properties about 10–15% of failures in clinical trials, suggesting that these issues are prominent reasons for attrition even before or during clinical phases [37, 38]. To evaluate these properties for the screened hits, we used the DeepPK server (https://biosig.lab.uq.edu.au/deeppk/), which predicts key pharmacokinetic and toxicity parameters [39]. The assessed properties included: Human Intestinal Absorption, P-Glycoprotein inhibition and substrate status, Blood–Brain Barrier (CNS) penetration, inhibition and substrate classification for CYP enzymes (CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4), AMES mutagenicity, carcinogenicity, drug-induced liver injury (DILI), and hERG channel inhibition.

Molecular dynamics simulations

The top ~ 1% of compounds retained from docking were further analyzed for stability within the SIRT2 active site using molecular dynamics (MD) simulations in GROMACS over 250 ns. Each protein–ligand complex was separated and prepared individually by generating topology files: the CHARMM36 all-atom force field was used for the protein via pdb2gmx, and ligand parameters (including the NAD^+ cofactor) were obtained from the CGenFF web server (https://cgenff.umaryland.edu/), then converted to GROMACS format using the cgenff_charmm2gmx.py script. The final complexes were solvated in a TIP3P water model within a rhombic dodecahedron box of dimensions $8.99 \times 8.99 \times 6.36$ nm and neutralized with Na^+ and Cl^- ions at physiological concentration (~0.15 M). Energy minimization was carried out using the steepest descent algorithm until a maximum force (Fmax) below 1000 kJ/mol/nm was achieved. This was followed by two 1 ns equilibration steps using the NVT and NPT ensembles at 300 K and 1 bar, respectively. All simulations were performed under periodic boundary conditions using the particle mesh Ewald method for long-range electrostatics. A production MD run of 200 ns was then performed for each system. The resulting trajectories were analyzed using GROMACS tools to compute structural and dynamic descriptors, including root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), hydrogen bonds, principal component analysis (PCA), free energy landscapes (FELs), and post-MD structural alignment to visualize the extent of local rearrangements around the binding pocket after dynamic equilibration.

MM/GBSA free energy calculations

The selected SIRT2–ligand complexes were further evaluated using MM/GBSA binding free energy calculations via the gmx_MMPBSA package [29]. Snapshots were extracted every 20 ns from the MD trajectory, resulting in 10 frames for energy analysis. The goal was to estimate the total binding free energy and identify the major energy components (van der Waals, electrostatic, polar and non-polar solvation) contributing to the stability of each complex. In addition, per-residue energy decomposition was performed to pinpoint the contributions of individual residues, particularly those within the active site, to ligand binding. This analysis provides insights into key interaction hotspots that may guide future optimization of SIRT2 inhibitors.

Results

QSAR model validation and screening

The performance of four gradient boosting models, Gradient Boosting, XGBoost, LightGBM, and CatBoost, was evaluated using key regression metrics: R^2, RMSE, MAE, and MSE, across training and validation sets (Table 1). As shown in Table 1, all models demonstrated good predictive ability, with XGBoost achieving the best performance on the training set ($R^2 = 0.98$, RMSE = 0.15), indicating a near-perfect fit. However, it showed a slight drop in validation performance ($R^2 = 0.63$), suggesting possible overfitting. CatBoost exhibited strong generalization, with high R^2 scores in both training (0.96) and validation (0.71) sets and relatively low error metrics, making it the most balanced model. LightGBM and Gradient Boosting also performed reasonably well, with validation R^2 values of 0.68 and 0.64, respectively. Based on these results, CatBoost was selected as the optimal model for downstream virtual screening due to its strong balance between fitting and generalization. The full workflow for QSAR model development, hyperparameter optimization, applicability domain assessment, and bioactivity prediction has been made publicly available on GitHub for reproducibility (https://github.com/yboulaaman e/QSARBoost).

To further improve predictive performance, the CatBoost model was optimized through hyperparameter tuning using the Optuna framework (Table S1 in Supporting Information). The tuned model demonstrated robust performance with an R^2 of 0.94 on the training set, along with low error metrics (RMSE = 0.28, MAE = 0.22, MSE = 0.08), indicating a strong fit. On the validation set, the model achieved an R^2 of 0.74, confirming good generalization. Y-randomization was performed to assess robustness, yielding a mean R^2 of − 0.141 on scrambled data, indicating that the model's predictive power is not due to chance. Additionally, the

Table 1 Model performance comparison on training and validation sets using R^2, RMSE, MAE, and MSE

	Training				Validation			
	R^2	RMSE	MAE	MSE	R^2	RMSE	MAE	MSE
GradientBoosting	0.81	0.51	0.40	0.26	0.64	0.70	0.53	0.50
XGBoost	0.98	0.15	0.08	0.02	0.63	0.71	0.51	0.52
LightGBM	0.90	0.37	0.28	0.14	0.68	0.66	0.49	0.44
CatBoost	0.96	0.24	0.18	0.06	0.71	0.63	0.46	0.40

CatBoost and LightGBM show strong generalization, while XGBoost slightly overfits

model achieved a Q^2 of 0.72 in tenfold cross-validation, supporting its reliability for external prediction and virtual screening applications (Table 2).

The tuned CatBoost model was further evaluated on the untouched holdout set to assess external predictivity. The parity plot (Fig. 3A) demonstrated strong correlation between predicted and true pIC50 values, with $R^2 = 0.75$, RMSE = 0.63, and MAE = 0.46, indicating reliable performance on unseen data. The residuals histogram (Fig. 3B) showed a normal distribution centered around zero, suggesting unbiased predictions. The residuals versus predicted (Fig. 3C) and absolute error vs. predicted (Fig. 3D) plots indicated no significant pattern or heteroscedasticity, supporting model consistency across the prediction range. The Williams plot (Fig. 4E) confirmed most compounds fall within the applicability domain, with few outliers (Fig. S2 in Supporting Information) exceeding the leverage threshold ($h^* = 0.066$) or ± 3 standardized residuals, validating the model's generalization capacity and robustness for virtual screening.

For virtual screening, the validated CatBoost QSAR model was applied to the DrugBank database to predict the inhibitory potency of each compound. First, the initial set of 12,307 DrugBank molecules was filtered using the Ghose criteria to ensure drug-likeness, yielding a pre-screened subset of 4,947 structures. The optimized CatBoost model was then used to predict pIC_{50} values for every compound in this subset. Compounds achieving a predicted $pIC_{50} \geq 6$, corresponding to an estimated $IC_{50} \leq 1$ µM, were considered potential high-affinity SIRT2 inhibitors. Based on this cutoff, 97 candidates were prioritized for subsequent structure-based evaluation through molecular docking and MD simulations.

Docking results

To evaluate the binding affinity of the top-ranked DrugBank hits to the SIRT2 active site, molecular docking was performed on 97 selected compounds ($pIC_{50} \geq 6$). Prior to docking the test compounds, the docking protocol was validated by redocking the reference ligand (SirReal2) into its original binding site. The resulting pose was then superimposed onto the native crystallographic conformation to calculate the root-mean-square deviation (RMSD). An RMSD below 2.0 Å is generally considered acceptable for reliable docking. In this case, the redocking yielded a remarkably low RMSD of 0.074 Å, indicating highly accurate pose prediction, as illustrated in Fig. 4.

Table 3 summarizes the molecular docking results of the top DrugBank hits ($pIC_{50} \geq 6$) screened against the SIRT2 active site. Binding affinities were evaluated using AutoDock Vina scores, and key protein–ligand interactions were identified. Among the hits, compound DB19242 exhibited the strongest binding affinity (−12.52 kcal/mol), outperforming the reference ligand SirReal2 (−12.15 kcal/mol), and formed multiple favorable hydrophobic interactions and a hydrogen bond with His187. Similarly, DB14822, DB06506, and DB07010 showed high predicted activity ($pIC_{50} > 6.3$) and strong docking scores (−11.25 to −12.05 kcal/mol), with consistent binding to key SIRT2 residues such as Phe119, Phe131, Tyr139, Phe190, and Ile232. These residues are known to form a hydrophobic pocket and contribute to π-stacking interactions, especially with aromatic ligands. Notably, π-stacking with Phe190 and Tyr139 was frequently observed across the top compounds, reinforcing the importance of these residues in ligand stabilization. Several compounds also formed hydrogen bonds with His187 and Val233, which may enhance binding specificity. Overall, the docking results suggest that these DrugBank hits exhibit binding modes similar to the reference inhibitor, making them promising candidates for further validation.

Figure 5 illustrates the binding poses of the reference ligand SirReal2 and the top-ranked DrugBank hits within the SIRT2 active site. All compounds occupy the catalytic tunnel and establish key interactions with conserved residues such as PHE96, PHE119, TYR139, ILE169, and ASN168, which are known to play roles in ligand stabilization [40, 41]. Several candidates, including DB19242, DB14822, and DB07010, form hydrogen bonds and hydrophobic contacts similar to those observed for SirReal2, suggesting favorable binding. Notably, DB08930 and DB03571 also showed deep insertion into the binding pocket, indicating strong interaction potential. These docking results support the likelihood of these compounds acting as effective SIRT2 inhibitors, warranting further dynamic and energetic evaluations.

To complement the docking results, a comparative structure–activity relationship (SAR) analysis was performed between the docked DrugBank candidates and the SIRT2-selective chemical probe SirReal2. Maximum common substructure (MCS) mapping, physicochemical property

Table 2 Performance metrics for training and validation sets

	R^2	RMSE	MAE	MSE	$R^2_{scramble}$	$Q^2_{tenfold}$
Training set	0.94	0.28	0.22	0.08	−0.14	0.72
Validation set	0.74	0.64	0.49	0.41		

High R^2 and low errors indicate good model fit, while Q^2 from tenfold CV confirms robustness. Negative R^2 from Y-scrambling suggests no chance correlations

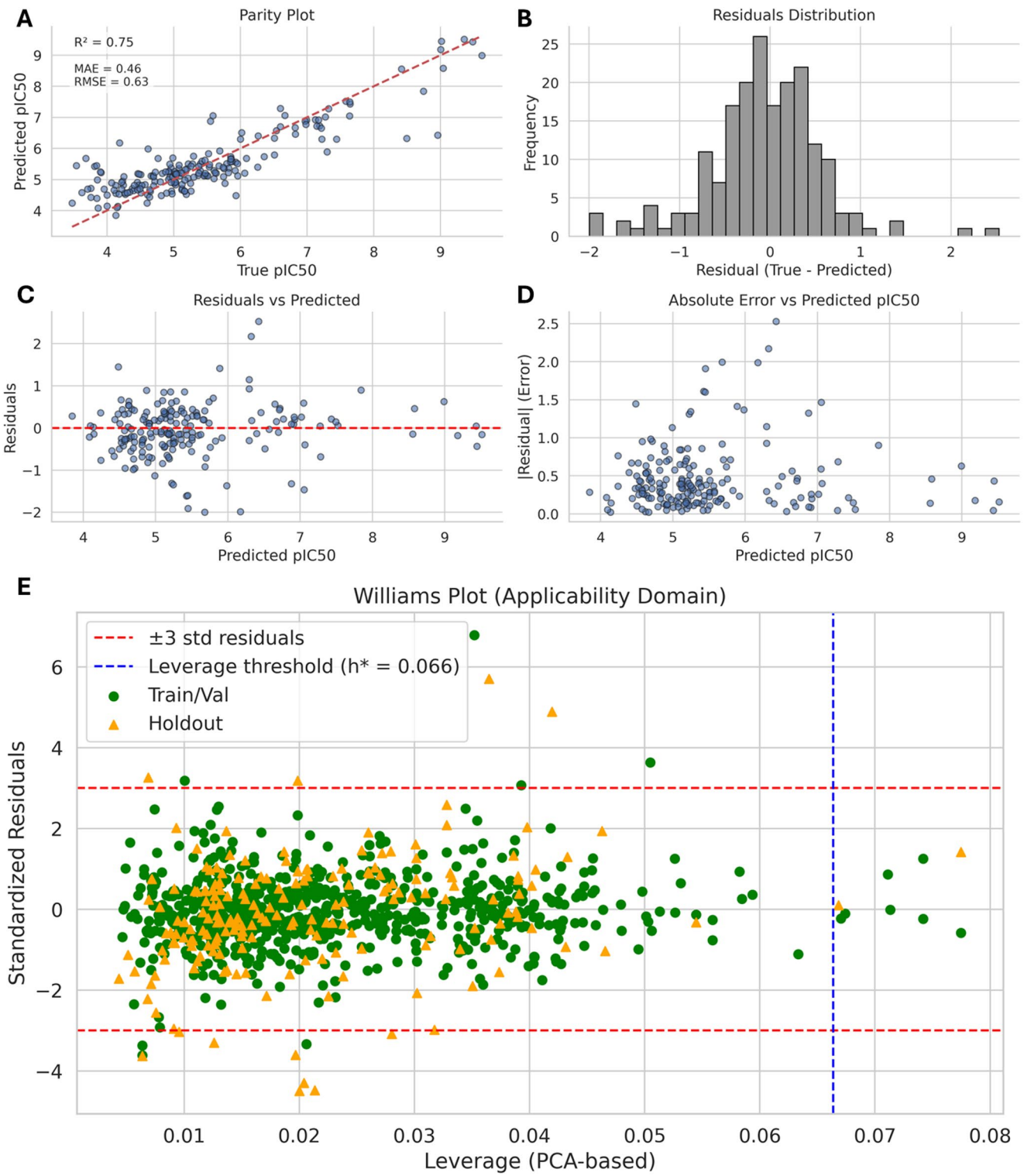

Fig. 3 Model diagnostic plots showing prediction accuracy, error distribution, and applicability domain. **A** Parity plot shows good correlation (R^2=0.75). **B** Residuals are centered around zero. **C** No clear trend in residuals versus predicted. **D** Most absolute errors are below 1.0. **E** Williams plot confirms most compounds lie within the model's applicability domain

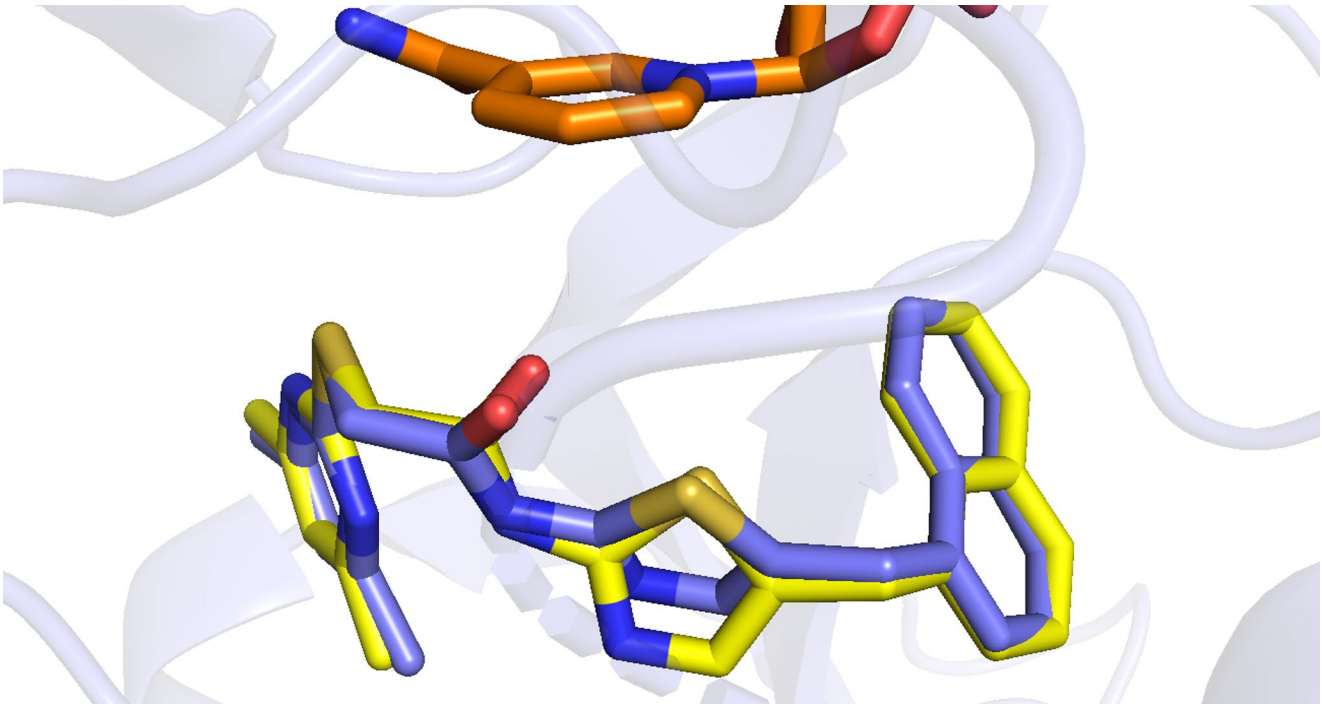

Fig. 4 Redocking validation of the SIRT2 inhibitor SirReal2. The native co-crystallized pose of SirReal2 (purple) is superimposed with the redocked pose (yellow), with the NAD^+ cofactor shown in orange. The excellent alignment, with an RMSD of 0.074 Å, confirms the reliability of the docking protocol

Z-score profiling, and molecular similarity metrics were employed to quantify structural and functional resemblance (Fig. S3 in Supporting Information). Several repurposed drugs demonstrated moderate to high scaffold and three-dimensional shape similarity to SirReal2, indicating potential to occupy the SIRT2 binding pocket in an analogous fashion. Other candidates retained key pharmacophoric features but differed in polarity or conformational flexibility, which may influence binding interactions. In the context of drug repurposing, such resemblance is advantageous as it may enhance the likelihood of target engagement while leveraging the established pharmacokinetic and safety profiles of approved therapeutics.

ADMET results

DeepPK ADMET analysis (Table 4) shows high intestinal absorption for all candidates (>80% in most, DB03571 = 70.4%) and high Caco-2 permeability (logPapp > 0.90) except DB07333 and DB07970. All are P-gp substrates, with several also acting as inhibitors, and all have low skin permeability (logKp < −2.5). Distribution predictions identify DB14822 (−1.836) and DB07010 (−1.952) as CNSp + (BBB-permeant), DB08930 (−3.597) as CNSp−, and others in the intermediate range. CYP3A4 metabolism is common, with DB19242 showing minimal CYP inhibition. Clearance varies, highest for DB06506. Toxicity predictions show most are AMES- and hERG I-negative, but many are hERG II-positive; MTD classification flags SirReal2, DB03571, DB07010, and DB07333 as high (>0.477), with hepatotoxicity predicted for most. Overall, ADMET predictions further prioritized DB14822 and DB07010 as the most favorable CNS-penetrant candidates, while DB03571 and DB06506 exhibited balanced profiles supporting further evaluation.

Molecular dynamics simulations

To evaluate the stability of the protein–ligand complexes, 250 ns molecular dynamics simulations were performed under physiological conditions. Various analyses were carried out, including RMSD, RMSF, radius of gyration, and hydrogen bond assessments to characterize their structural behaviour.

Root-mean square deviation and fluctuations

The structural stability and dynamics of SIRT2 with ten ligands, including SirReal2, were assessed over 250 ns MD simulations using backbone RMSD, RMSF, ligand RMSD, and NAD^+ RMSD (Fig. 6). All systems equilibrated within 30 ns. DB03571 and DB08930 showed the lowest backbone RMSD (~0.25–0.30 nm) and most compact residue profiles, while SirReal2 was moderately flexible

Table 3 Molecular docking results of top-ranked DrugBank compounds ($pIC_{50} \geq 6$) against the SIRT2 active site

ID	Drug group	Predicted pIC50	Vina score (Kcal/mol)	Interacting residues	Bond type	Distance (Å)
SirReal2	Chemical probe	5.68	−12.15	Phe119	Hydrophobic interaction	3.61
				Phe131	Hydrophobic interaction	3.67
				Tyr139	Hydrophobic interaction	3.47
					π-Stacking	4.88
				Phe143	Hydrophobic interaction	3.55
				Ile169	Hydrophobic interaction	3.72
				Phe190	Hydrophobic interaction	3.62
					π-Stacking	4.18
				Ile232	Hydrophobic interaction	3.93
				Phe234	π-Stacking	4.79
DB19242	Investigational	6.04	−12.52	Phe131	Hydrophobic interaction	3.43
				Leu134	Hydrophobic interaction	3.43
				Leu138	Hydrophobic interaction	3.81
				Tyr139	Hydrophobic interaction	5.05
				Phe143	Hydrophobic interaction	3.86
				Ile169	Hydrophobic interaction	3.5
				Asp170	Hydrophobic interaction	3.51
				His187	Hydrogen bond	3.05
				Phe190	Hydrophobic interaction	3.8
					π-Stacking	4.27
				Ile232	Hydrophobic interaction	3.88
DB14822	Investigational	6.35	−12.05	Phe96	Hydrophobic interaction	3.73
				Phe119	Hydrophobic interaction	3.76
				Phe131	Hydrophobic interaction	3.82
				Leu138	Hydrophobic interaction	3.76
				Ile169	Hydrophobic interaction	3.94
				His187	Hydrogen bond	2.57
				Phe190	Hydrophobic interaction	3.85
				Ile232	Hydrophobic interaction	3.68
				Phe234	Hydrophobic interaction	3.82
DB03571	Experimental	6.09	−11.81	Phe131	Hydrophobic interaction	3.34
				Leu134	Hydrophobic interaction	3.32
				Leu138	Hydrophobic interaction	3.83
				Tyr139	π-Stacking	5.15
				Ile169	Hydrophobic interaction	3.56
				Phe190	π-Stacking	4.16
				Val233	Hydrogen bond	3.17
DB07333	Experimental	6.09	−11.77	Phe96	Hydrophobic interaction	3.59
				Phe119	Hydrophobic interaction	3.67
				Leu138	Hydrophobic interaction	3.8
				Tyr139	π-Stacking	4.78
				Ile169	Hydrophobic interaction	3.77
				His187	Hydrogen bond	3.23
				Phe190	Hydrophobic interaction	3.37
					π-Stacking	4.08
				Ile232	Hydrophobic interaction	3.66
				Val233	Hydrogen bond	2.18
				Phe234	π-Stacking	5.45
DB07010	Experimental	6.68	−11.37	Phe119	π-Stacking	4.95
				Tyr139	π-Stacking	5.22
				Ile169	Hydrophobic interaction	3.55
				His187	Hydrophobic interaction	3.64
				Phe190	π-Stacking	3.99
				ILe232	Hydrophobic interaction	3.92

Table 3 (continued)

ID	Drug group	Predicted pIC50	Vina score (Kcal/mol)	Interacting residues	Bond type	Distance (Å)
DB07970	Experimental	6.09	−11.32	Ile93	Hydrophobic interaction	3.21
				Phe96	Hydrophobic interaction	3.16
				Ile118	Hydrophobic interaction	3.93
				Phe119	Hydrophobic interaction	3.27
					Halogen bonds	2.93
				Phe131	π-Stacking	4.96
				Leu134	Hydrophobic interaction	3.8
				Leu138	Hydrogen bond	2.31
				Tyr139	Hydrophobic interaction	3.88
				Ile169	Hydrophobic interaction	3.41
				Asp170	Hydrogen bond	2.79
				Phe190	π-Stacking	3.69
				Ile232	Hydrophobic interaction	3.54
DB06506	Investigational	6.06	−11.25	Phe96	Hydrophobic interaction	3.25
				Phe119	π-Stacking	5.05
				Phe131	Hydrophobic interaction	3.51
				Leu138	Hydrophobic interaction	3.89
				Tyr139	Hydrophobic interaction	3.53
					π-Stacking	4.64
				Phe143	Hydrophobic interaction	3.88
				Ile169	Hydrophobic interaction	3.01
				Asp170	Hydrophobic interaction	3.92
				Phe190	Hydrophobic interaction	3.6
					π-Stacking	4.03
				Ile232	Hydrophobic interaction	3.83
				Phe234	π-Stacking	4.8
DB08349	Experimental	6.17	−11.09	Ile93	Hydrophobic interaction	3.58
				Phe96	Hydrophobic interaction	3.88
				Phe119	π-Stacking	4.93
				Leu134	Hydrophobic interaction	3.55
				Leu138	Hydrophobic interaction	3.65
				Pro140	Hydrophobic interaction	3.38
				Asp170	Hydrophobic interaction	3.35
				His187	Hydrophobic interaction	3.38
DB08930	Approved	6.5	−11.06	Phe119	Hydrophobic interaction	3.32
				Tyr139	π-Stacking	4.98
				169 Ile	Hydrophobic interaction	3.8
				Asp170	Hydrophobic interaction	3.93
				Phe190	π-Stacking	4.02

(~0.30–0.40 nm). DB14822, DB07333, DB07010, and DB07970 exceeded 0.40 nm, indicating higher conformational variability; RMSF confirmed increased loop/terminal mobility, with DB14822 and DB07333 peaking > 1.0 nm. Ligand RMSD was lowest (< 0.20 nm) for DB03571, DB08930, and SirReal2, moderate (~0.20–0.25 nm) for DB19242, DB06506, DB08349, and DB07970, and highest (> 0.30 nm) for DB14822, DB07333, and DB07010. NAD^+ remained most stable (< 0.20 nm) with DB03571, DB08930, and SirReal2, moderate (~0.20–0.30 nm) with DB19242, DB06506, DB08349, and DB07970, and most perturbed (> 0.35 nm) by DB14822 and DB07010. Overall, DB03571 and DB08930 best preserved SIRT2 structure and cofactor binding, SirReal2 maintained strong binding with moderate flexibility, while DB14822, DB07333, and DB07010 induced the greatest destabilization.

Radius of gyration

Radius of gyration (Rg) analysis (Fig. 7) showed values between 2.050 and 2.200 nm for all complexes, indicating no major unfolding. DB03571 and DB08930 had the lowest, most stable Rg (2.050–2.100 nm), preserving high compactness. SirReal2 remained compact (2.080–2.120 nm). DB14822, DB07333, and DB07010 often exceeded 2.150 nm, suggesting transient expansion. DB07970 and

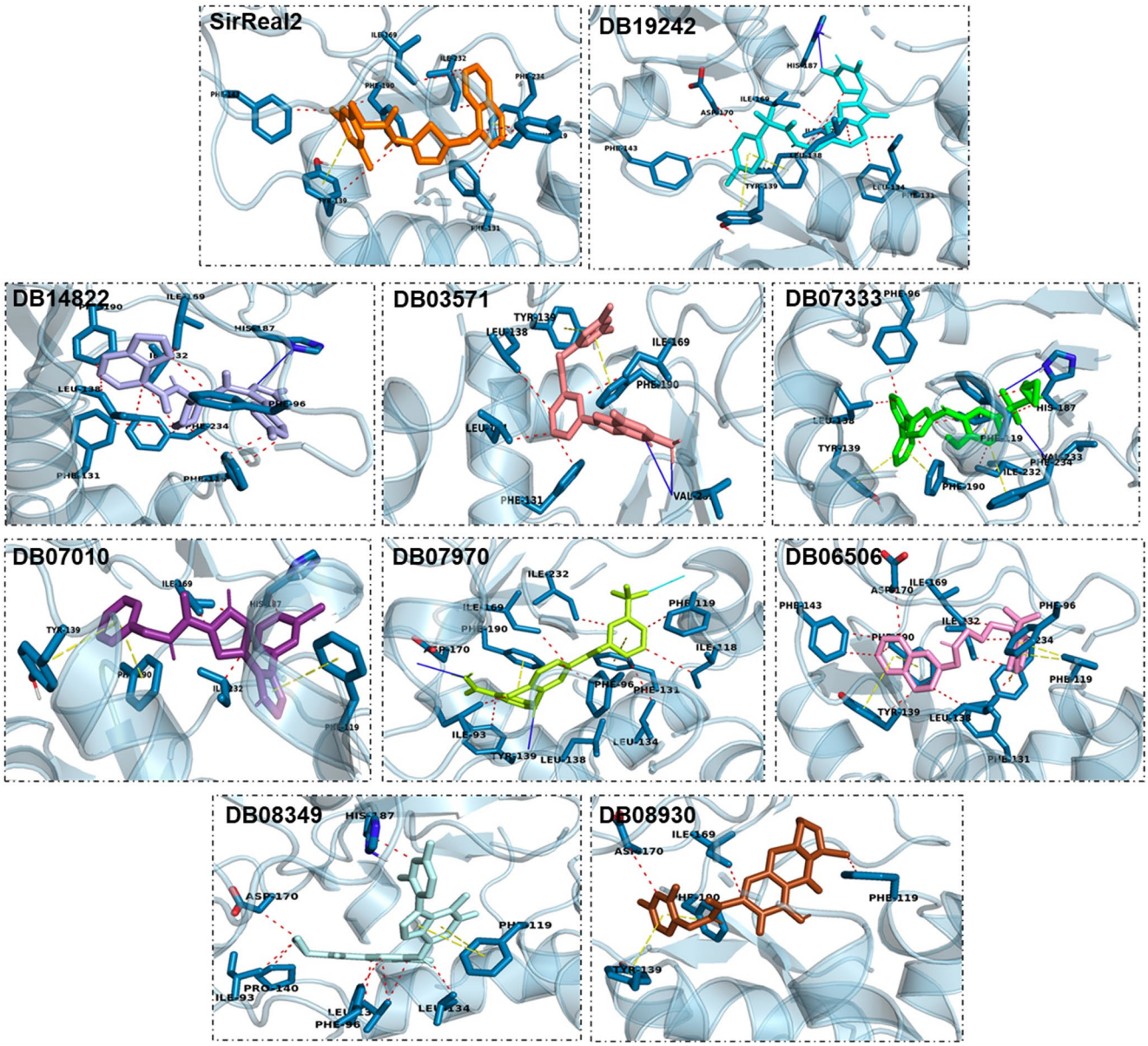

Fig. 5 Binding poses of top-ranked DrugBank hits in the SIRT2 active site

DB06506 showed moderately higher values (~2.120–2.170 nm) with more oscillations, while DB19242 and DB08349 were intermediate and stable. Overall, DB03571 and DB08930 best maintained compactness, whereas DB14822 and DB07333 increased conformational variability, consistent with RMSD/RMSF trends.

Solvent-accessible surface

SASA analysis (Fig. 8) showed values of 145–175 nm^2 across all complexes, indicating no major unfolding. DB03571 and DB08930 had the lowest SASA (145–155 nm^2), maintaining compact structures with minimal solvent exposure. SirReal2 remained slightly higher (150–160 nm^2) but stable. DB14822 had the highest (~165–175 nm^2), suggesting increased flexibility, while DB07333, DB07010, and DB07970 also exceeded 160 nm^2. DB19242, DB08349, and DB06506 were moderate (155–165 nm^2). These results align with Rg and RMSD trends, with DB03571 and DB08930 best preserving compactness and DB14822 and DB07333 promoting greater solvent exposure.

Principal component analysis

The essential dynamics of the SIRT2–ligand systems were investigated using principal component analysis (PCA), which captures large-scale collective motions during the 250 ns simulation. Figure 9 presents the 2D projection

Table 4 Pharmacokinetic and toxicity profiles of top-ranked DrugBank hits compared to SirReal2, predicted using the DeepPK server

Pharmacokinetic properties		Selected ligands									
Properties	Model name	SirReal2	DB19242	DB14822	DB03571	DB07333	DB07010	DB07970	DB06506	DB08349	DB08930
Absorption	Water solubility	−5.471	−4.166	−4.886	−3.183	−4.079	−3.588	−4.077	−3.735	−3.587	−3.331
	$CaCo_2$ permeability	1.023	0.628	1.068	1.046	−0.035	1.101	0.21	1.007	1.187	0.993
	Intestinal absorption (human)	92.829	80.579	90.746	70.421	100	91.965	86.214	94.3	95.913	81.948
	Skin permeability	−2.702	−3.02	−2.759	−2.735	−2.736	−2.735	−2.737	−2.704	−2.739	−2.823
	P-glycoprotein substrate	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	P-glycoprotein I inhibitor	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
	P-glycoprotein II inhibitor	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Distribution	VDss (human)	0.171	−0.496	0.276	0.452	0.099	−0.046	0.181	0.737	0.395	0.153
	Fraction unbound (human)	0.136	0.027	0	0	0.096	0	0	0.123	0.052	0.205
	BBB permeability	−0.799	−0.97	−0.427	−1.437	−0.397	−0.943	−1.232	−0.61	−0.336	−1.03
	CNS permeability	−2.102	−2.667	−1.836	−2.799	−2.43	−1.952	−2.013	−2.599	−2.372	−3.597
Metabolism	CYP2D6 substrate	No	No	No	No	No	Yes	No	No	No	No
	CYP3A4 substrate	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
	CYP1A2 inhibitor	Yes	No	No	No	No	Yes	Yes	No	No	No
	CYP2C19 inhibitor	Yes	No	Yes	No	Yes	Yes	Yes	No	Yes	No
	CYP2C9 inhibitor	Yes	No	Yes	No	Yes	Yes	Yes	No	Yes	No
	CYP2D6 inhibitor	No	No	No	No	No	No	No	No	No	No
	CYP3A4 inhibitor	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No
Excretion	Total clearance	−0.022	−0.412	0.178	−0.418	0.54	0.193	−0.168	0.723	−0.002	−0.06
	Renal OCT2 substrate	No	No	Yes	No	No	No	No	Yes	No	No

Table 4 (continued)

Pharmacokinetic properties		Selected ligands									
Properties	Model name	SirReal2	DB19242	DB14822	DB03571	DB07333	DB07010	DB07970	DB06506	DB08349	DB08930
Toxicity	AMES toxicity	Yes	No	No	No	No	No	No	No	No	No
	Max. tolerated dose (human)	0.63	−0.449	0.037	0.797	0.358	0.619	0.241	−0.211	0.298	0.035
	hERG I inhibitor	No	No	No	No	No	No	No	No	No	No
	hERG II inhibitor	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No
	Oral Rat Acute Toxicity (LD50)	2.732	2.16	2.866	2.342	3.031	3.145	2.539	2.765	3.227	2.289
	Oral rat chronic toxicity (LD50)	0.326	1.527	0.69	2.978	1.617	1.762	1.119	1.976	1.489	1.943
	Hepatotoxicity	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Skin sensitization	No	No	No	No	No	No	No	No	No	No
	T. Pyriformis toxicity	0.316	0.412	0.31	0.285	0.286	0.29	0.401	0.477	0.288	0.313
	Minnow toxicity	1.038	1.482	−0.994	0.513	−0.285	3.704	1.17	1.095	0.035	3.107

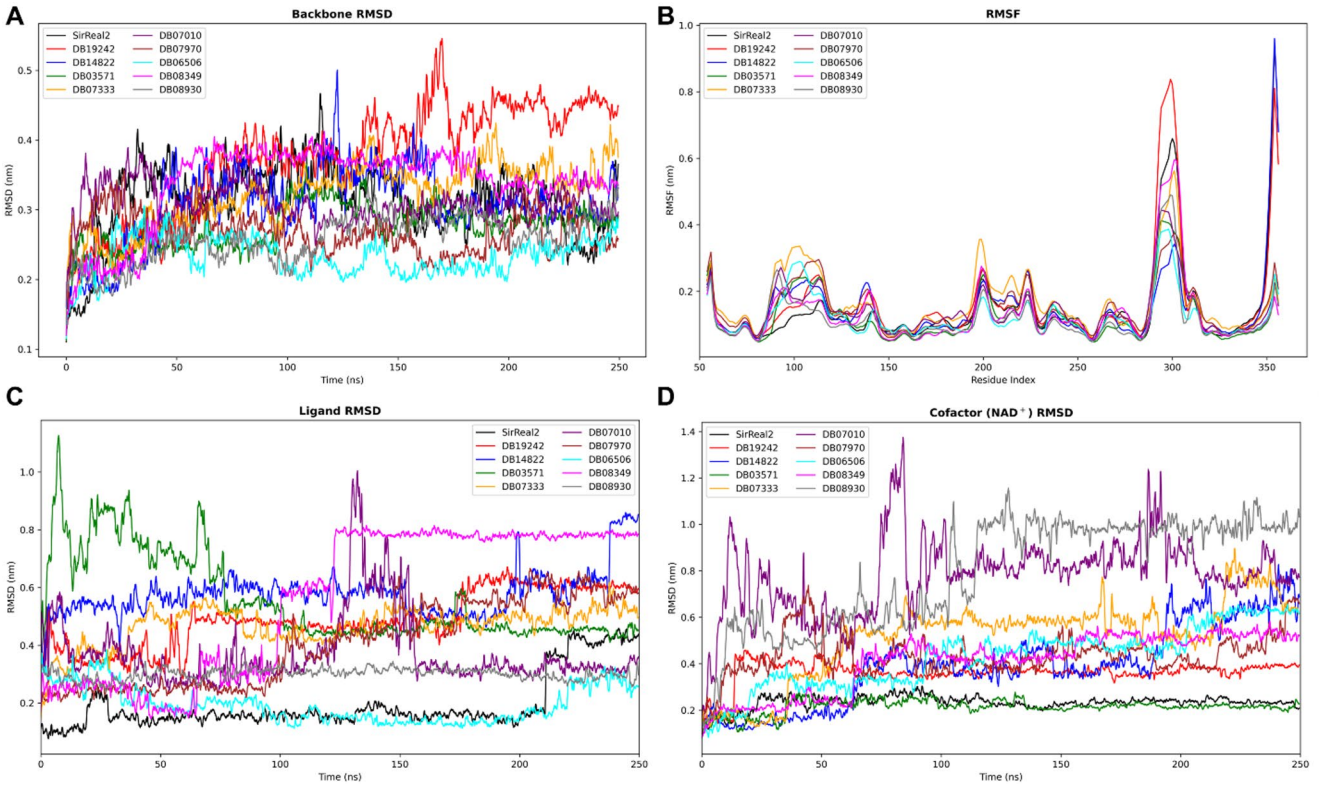

Fig. 6 **A** Backbone RMSD profiles of SIRT2-ligand complexes over 250 ns of MD simulation. **B** Residue-wise RMSF of SIRT2, highlighting flexible regions across the protein sequence. **C** Ligand RMSD trajectories showing positional stability within the binding pocket. **D** RMSD of the NAD^+ cofactor across all complexes. The plots illustrate the structural integrity of SIRT2, the localized flexibility of specific residues, and the differential stability of ligands and cofactor during the simulation

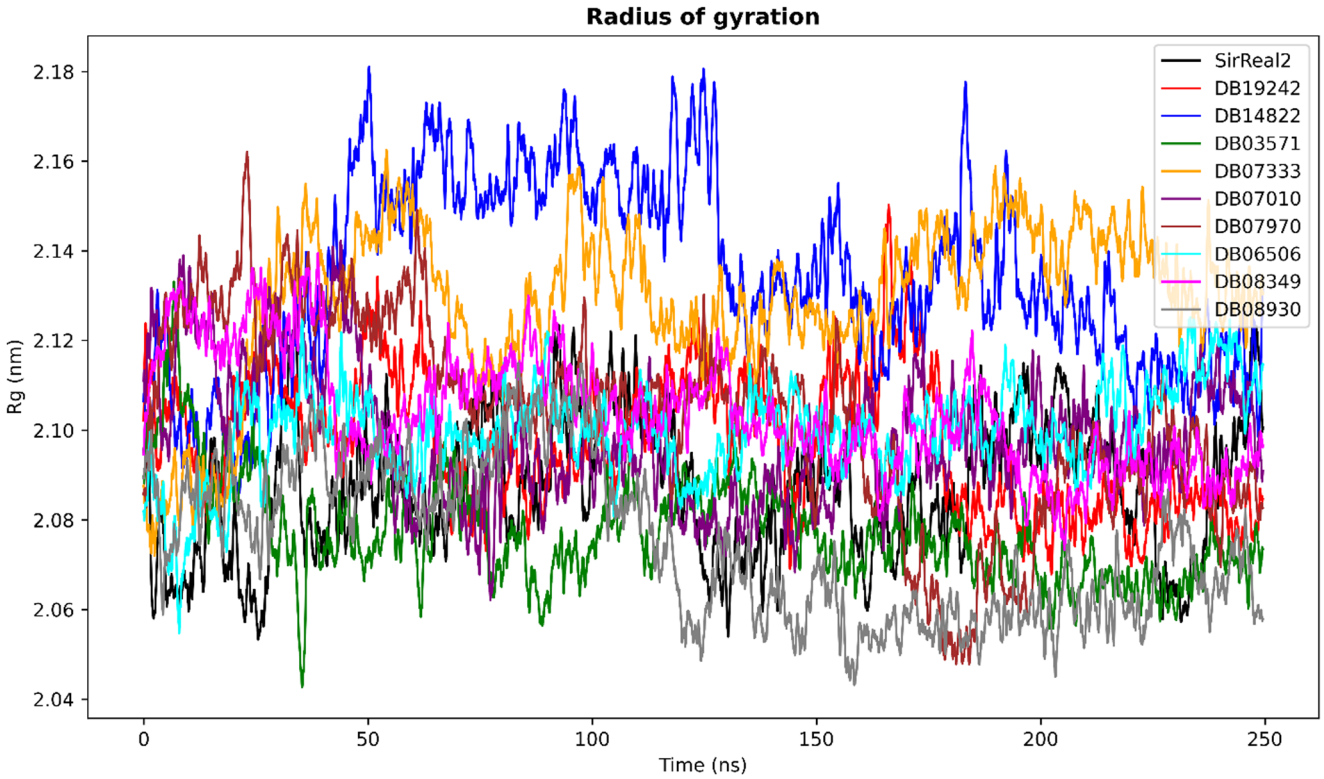

Fig. 7 Radius of gyration (Rg) of the SIRT2-ligand complexes over 250 ns MD simulations for DB03571 (green), SiReal2 (blue), DB14822 (orange), DB07333 (red), DB06506 (purple), and DB08349 (brown). Rg was calculated to monitor the compactness and structural stability of the protein throughout the simulation

along PC1 and PC2, highlighting differences in conformational sampling across complexes. DB03571, DB08930, and SirReal2 formed tight, centralized clusters, indicative of restricted motion and well-confined dynamics. In contrast, DB19242, DB14822, and DB07333 generated broader, more dispersed trajectories, reflecting enhanced conformational flexibility and induced plasticity. These patterns suggest that ligand binding modulates both local flexibility and global dynamics of SIRT2. Dense clustering for DB03571 and DB08930 reflects conserved structural behavior, whereas the widespread projections of DB14822 and DB19242 are consistent with their ability to promote broader fluctuations. Overall, the PCA results align with RMSD, Rg, and SASA analyses, reinforcing that DB03571 and DB08930 constrain SIRT2 dynamics, while DB14822 and DB19242 facilitate larger-scale motions.

Hydrogen bonding analysis

Hydrogen bond analysis highlighted clear differences in ligand interactions with SIRT2 (Fig. 10). DB03571, DB07010, and DB08349 established the most persistent protein–ligand H-bonds, often maintaining 2–4 contacts across the simulation, indicative of strong anchoring within the pocket. DB08930 and DB06506 also formed steady bonds, though at slightly lower levels. By contrast, DB19242, DB14822, and DB07333 engaged in fewer and more transient contacts, reflecting weaker stabilization, while SirReal2 showed minimal H-bonding. Cofactor–ligand H-bonds were generally infrequent; intermittent contacts were observed for DB19242, DB07010, and DB06506, whereas DB03571 and DB08930 showed negligible interaction with NAD^+. These results suggest that stable protein–ligand H-bonds, rather than cofactor engagement, underpin the strong binding observed for DB03571, DB07010, and DB08349.

Free energy landscape analysis

The free energy landscape (FEL) analysis provides insight into the conformational stability and dynamics of the SIRT2–ligand complexes. FEL plots, constructed from the first two principal components (PC1 and PC2), capture dominant motions in each system (Fig. 11). Narrow, deep basins indicate stable conformations, while broader or multi-basin profiles reflect greater flexibility and sampling of multiple substates. SirReal2, DB07333, and DB07010 showed well-defined single minima with deep energy wells, suggesting high stability and restricted dynamics. DB03571, DB06506, and DB08349 displayed more complex landscapes with

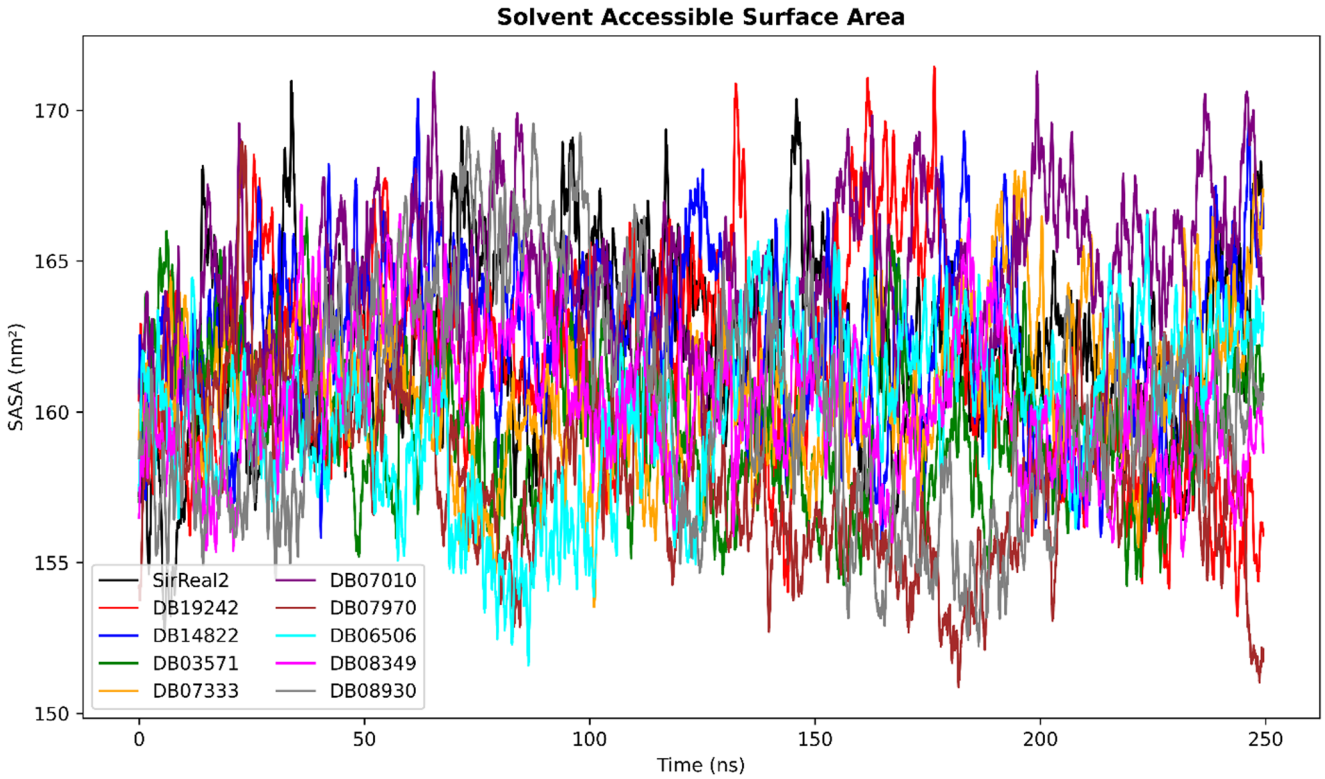

Fig. 8 Solvent accessible surface area (SASA) of the SIRT2-ligand complexes over 250 ns molecular dynamics simulations for DB03571 (green), DB08930 (gray), SirReal2 (black), DB19242 (red), DB08349 (fuchsia), DB06506 (cyan), DB07970 (brown), DB07010 (violet), DB07333 (orange), and DB14822 (blue)

multiple minima, reflecting increased flexibility and conformational transitions. DB14822, DB19242, and DB08930 had broader, less defined distributions, consistent with higher fluctuations and reduced compactness. DB07970 showed an intermediate pattern with a moderately deep well and limited scattering. These trends align with RMSD, SASA, and PCA analyses, indicating that ligands forming narrow FEL basins stabilize SIRT2, whereas broader basins promote structural variability.

MM/GBSA free energy analysis

The MM/GBSA free energy analysis was performed over the entire 250 ns trajectory, extracting energies every 250 frames from a total of 25,000 frames, yielding 101 representative snapshots. This sampling was sufficient to estimate the binding affinities of the ten ligands to SIRT2 by decomposing the total binding energy into van der Waals ($\Delta E_{VDWAALS}$), electrostatic (ΔE_{EL}), polar solvation (ΔE_{GB}), and non-polar solvation (ΔE_{SURF}) components. The resulting energy profiles revealed clear differences among the complexes (Table 5).

DB14822 showed the most favorable binding free energy (−35.72 kcal/mol), followed closely by DB03571 and SirReal2 (−34.28 kcal/mol) and DB06506 (−32.63 kcal/mol). These values, consistent with their low RMSD and Rg profiles, indicate tight, stable association with the protein, supported by strong van der Waals and gas-phase contributions. DB07333 (−27.35 kcal/mol) and DB19242 (−24.93 kcal/mol) had moderately favorable energies, matching their intermediate dynamics from RMSD and SASA. DB07970 (−25.05 kcal/mol) and DB08930 (−23.02 kcal/mol) also bound favorably but with larger polar solvation penalties offsetting strong gas-phase interactions. DB08349 (−11.25 kcal/mol) and DB07010 (−16.6 kcal/mol) showed the weakest affinities, linked to less favorable electrostatics, higher solvation penalties, and dynamic instability with higher SASA values. Per-residue decomposition (Fig. 12) confirmed van der Waals and electrostatic forces as the main contributors.

Post-MD visual inspection

The MD simulations provided a detailed assessment of the conformational behavior, stability, and binding dynamics of the selected SIRT2-ligand complexes. Collectively, the RMSD, RMSF, Rg, and SASA analyses confirmed that several repurposed DrugBank compounds, particularly DB14822, DB03571, and DB06506 maintained stable and compact protein–ligand interactions over the 250 ns

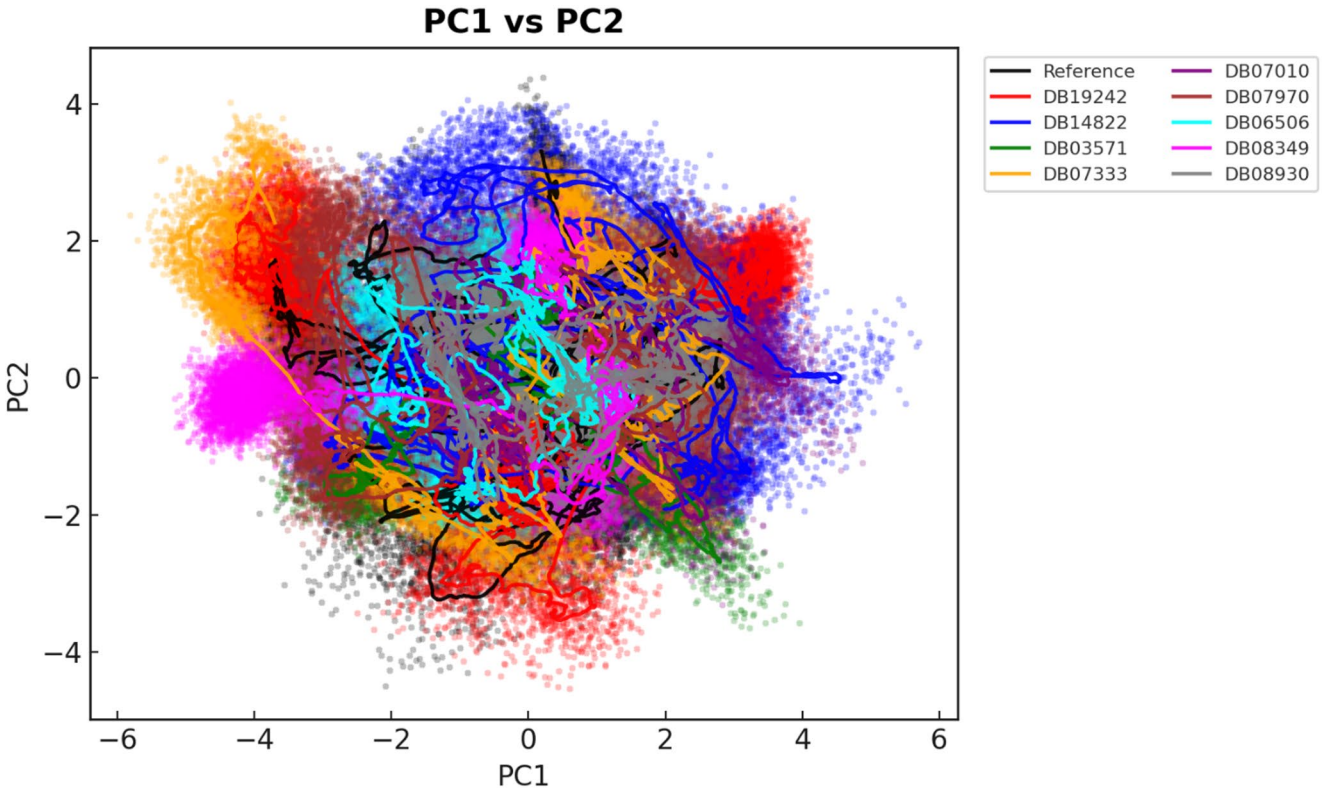

Fig. 9 Principal component analysis (PCA) of SIRT2-ligand complexes over 250 ns molecular dynamics simulations for DB03571 (green), DB08930 (gray), SirReal2 (black), DB19242 (red), DB08349 (fuchsia), DB06506 (cyan), DB07970 (brown), DB07010 (violet), DB07333 (orange), and DB14822 (blue)

simulation period. PCA and free energy landscape results further revealed that these ligands constrained SIRT2's global motions, preserving a low-energy conformational ensemble comparable to the reference inhibitor SirReal2. Hydrogen bonding and MM/GBSA energy decomposition supported these findings, showing consistent intermolecular interactions and favorable binding energetics.

RMSD analysis and PyMOL superposition after 250 ns MD simulations (Fig. 13) showed that all SIRT2–ligand complexes remained bound, with deviations ranging from 1.521 Å (DB07970) to 2.603 Å (DB07333). The reference inhibitor SirReal2 (1.643 Å) and the candidates DB03571 (1.731 Å) and DB06506 (2.113 Å) retained stable poses, consistent with minimal structural drift. Higher RMSD values, as seen for DB07333 and DB07010 (>2.4 Å), indicated greater flexibility or partial reorientation, but without ligand dissociation.

Surface views before and after MD (Fig. 14) confirmed the conserved overall fold of SIRT2 and stable NAD positioning. Ligands with lower RMSD values maintained tighter packing, whereas those with higher deviations showed minor reorientations. These structural results align with other MD stability metrics, supporting DB14822, DB03571, and DB06506 as the most promising, structurally stable SIRT2 inhibitors for further experimental evaluation in the context of PD therapeutics.

Analysis of the top three leads

To further assess the robustness and reproducibility of the MD results, triplicate 500 ns MD simulations were performed for the top three SIRT2 candidates (DB14822, DB03571, and DB06506). Figure 15 provides a comprehensive overview of their structural and dynamic behaviour, including backbone RMSD, ligand RMSD, RMSF, and radius of gyration analyses. The backbone RMSD profiles (Fig. 15B) indicate that DB14822 and DB06506 rapidly reach stable plateaus and maintain average deviations around ~ 0.30–0.40 nm, reflecting preserved global stability of the SIRT2 complexes. In contrast, DB03571 exhibits higher RMSD values, exceeding ~ 0.50 nm during later stages, suggesting increased protein flexibility. This behaviour is mainly attributed to loop regions, as confirmed by RMSF analysis (Fig. 15D), while the catalytic core remains relatively rigid across all systems. Ligand RMSD trajectories (Fig. 15C) demonstrate that DB14822 retains the most stable binding pose throughout the simulations, whereas DB03571 and DB06506 undergo moderate positional

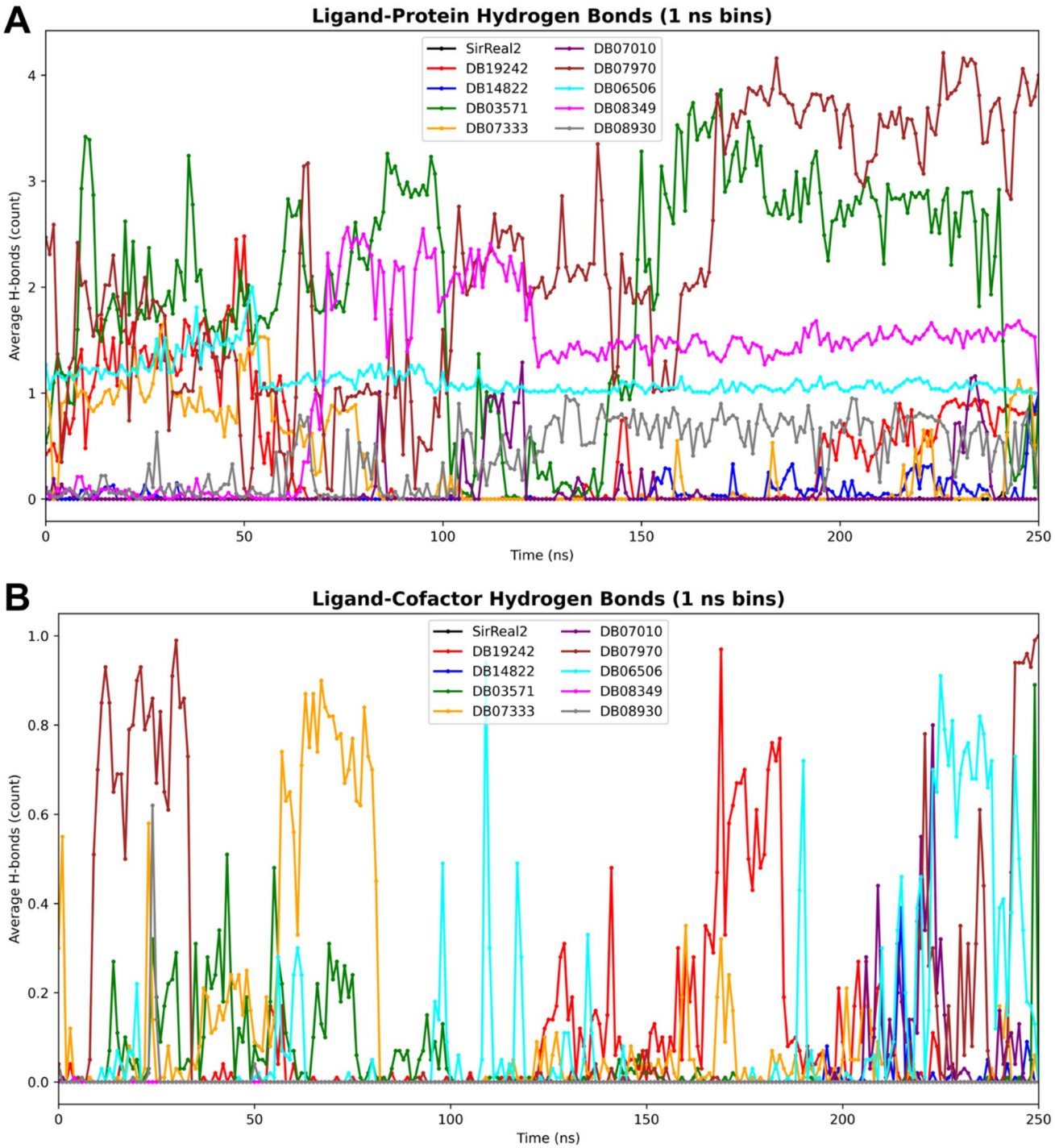

Fig. 10 Time evolution of hydrogen bonds between ligand–protein complexes (**A**) and ligand-cofactor (**B**) and their respective interaction partners over 250 ns molecular dynamics simulations. Each panel represents one ligand, showing hydrogen bonds formed between the protein and the ligand (blue) and between the NAD^+ cofactor and the ligand (green)

adjustments yet remain confined within the binding pocket, indicating sustained ligand engagement. Furthermore, the radius of gyration profiles (Fig. 15E) remain consistent across replicas, fluctuating narrowly around ~ 2.08–2.12 nm, confirming preserved global compactness and absence of large-scale unfolding. Collectively, these triplicate simulations validate the reproducibility of the observed dynamics and demonstrate that, despite ligand-dependent differences

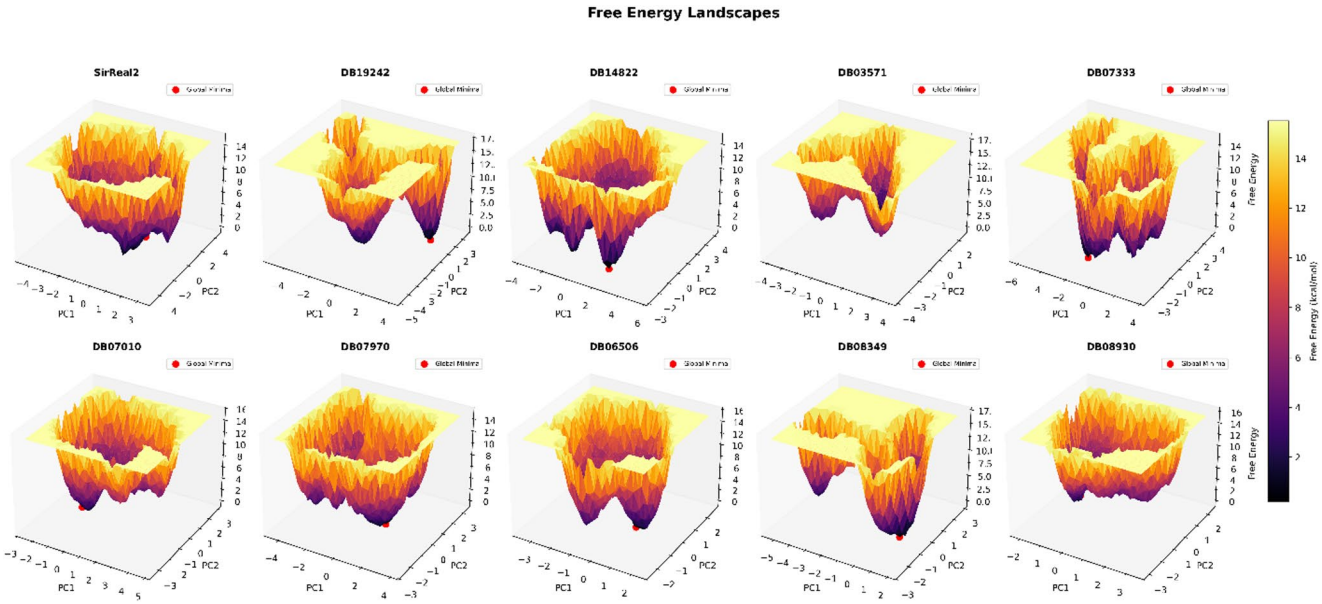

Fig. 11 3D free energy landscapes (FEL) of SIRT2-ligand complexes based on principal component analysis (PC1 vs. PC2) from 250 ns molecular dynamics simulations. The color gradient represents the relative free energy in kcal/mol**,** where darker regions indicate low-energy, stable conformations. Red spheres mark the global minimum for each system. Basin topology and depth reveal ligand-specific effects on the conformational stability and flexibility of SIRT2

Table 5 MMGBSA energy decomposition of SIRT2-ligand complexes

Complex	MM/GBSA (kcal/mol) per energy component						
	$\Delta E_{VDWAALS}$	ΔE_{EL}	ΔE_{GB}	ΔE_{SURF}	ΔE_{GAS}	ΔE_{SOLV}	ΔE_{TOTAL}
SirReal2	−42.51	−32.63	46.53	−5.67	−75.14	40.86	−34.28
DB19242	−43.2	−7.54	31.45	−5.64	−50.74	25.81	−24.93
DB14822	−51.61	−3.3	25.58	−6.39	−54.91	19.19	−35.72
DB03571	−42.51	−32.63	46.53	−5.67	−75.14	40.86	−34.28
DB07333	−53.81	−8.44	41.8	−6.9	−62.24	34.9	−27.35
DB07010	−39.11	−3.18	31.03	−5.35	−42.28	25.68	−16.6
DB07970	−39.23	−68.39	88.76	−6.2	−107.62	82.56	−25.05
DB06506	−46.85	−239.13	259.85	−6.51	−285.98	253.34	−32.63
DB08349	−36.73	−12.71	42.99	−4.8	−49.44	38.19	−11.25
DB08930	−43.12	−150.14	176.05	−5.81	−193.26	170.24	−23.02

Values represent the average contributions (in kcal/mol) of van der Waals (ΔVDWAALS), electrostatics (ΔEEL), polar solvation (ΔEGB), non-polar solvation (ΔESURF), gas-phase energy (ΔGGAS), solvation energy (ΔGSOLV), and total binding free energy (ΔTOTAL) for each ligand

$\Delta E_{VDWAALS}$, van der Waals energy; ΔE_{EL}, Electrostatic energy; ΔE_{GB}, Polar solvation energy; ΔE_{SURF}, Nonpolar solvation energy; ΔE_{GAS}, $\Delta E_{VDWAALS}+\Delta E_{EL}$; ΔE_{SOLV}, $\Delta E_{GB}+\Delta E_{SURF}$; ΔE_{TOTAL}, $\Delta E_{GAS}+\Delta E_{SOLV}$ (overall binding energy)

in local flexibility, all three compounds form stable SIRT2 complexes. Notably, DB14822 emerges as the most conformationally stable system, while DB03571 exhibits higher backbone mobility but maintains persistent binding, supporting a flexible yet energetically favorable interaction mode.

Discussion

In this work, a high-performance CatBoost QSAR model was employed to repurpose compounds from DrugBank as potential SIRT2 inhibitors, followed by molecular docking and molecular dynamics (MD) simulations for structural refinement. Among the screened compounds, DB14822, DB03571, and DB06506 emerged as top candidates, exhibiting MM/GBSA binding free energies in the range of −32 to −35 kcal mol^{-1} and demonstrating stable MD profiles comparable to, or exceeding, the reference inhibitor SirReal2. Consistent with prior crystallographic and computational studies of SIRT2 inhibitors, these ligands engaged key hydrophobic and π–π stacking interactions with residues such as Phe119, Tyr139, and Phe190, which are critical for anchoring within the active site [41, 42].

Comprehensive MD analyses, including RMSD, Rg, SASA, PCA, and FEL, revealed that DB03571 and DB08930

Fig. 12 Per-residue energy decomposition of the SIRT2–ligand complexes based on MM/GBSA calculations over 101 snapshots extracted from 250 ns simulations. Each plot shows the contribution of van der Waals (VDW), electrostatic (EEL), polar solvation (EGB), and non-polar solvation (ESURF) energy terms to the total binding free energy for individual residues interacting with the ligand. Key hotspot residues with strong energetic contributions are highlighted across different complexes, indicating variable interaction profiles depending on the ligand structure. Energy values are expressed in kcal/mol

formed particularly stable and compact complexes, with constrained conformational sampling and favorable energetic profiles. DB14822 and DB19242 also exhibited strong binding affinities but induced greater backbone flexibility in the later stages of the simulations, suggesting subtle conformational adjustments within the orthosteric pocket. The convergence of structural stability and energetic favorability supports these compounds as promising SIRT2 inhibition candidates. The apparent differences between dynamic stability metrics and MM/GBSA energetics reflect complementary aspects of ligand recognition. DB14822, despite its higher conformational plasticity, achieved the most favourable binding free energy, supporting a flexible but energetically optimized association with the active site. Conversely, DB03571 combined high structural stability with favourable binding energetics, representing a balanced candidate.

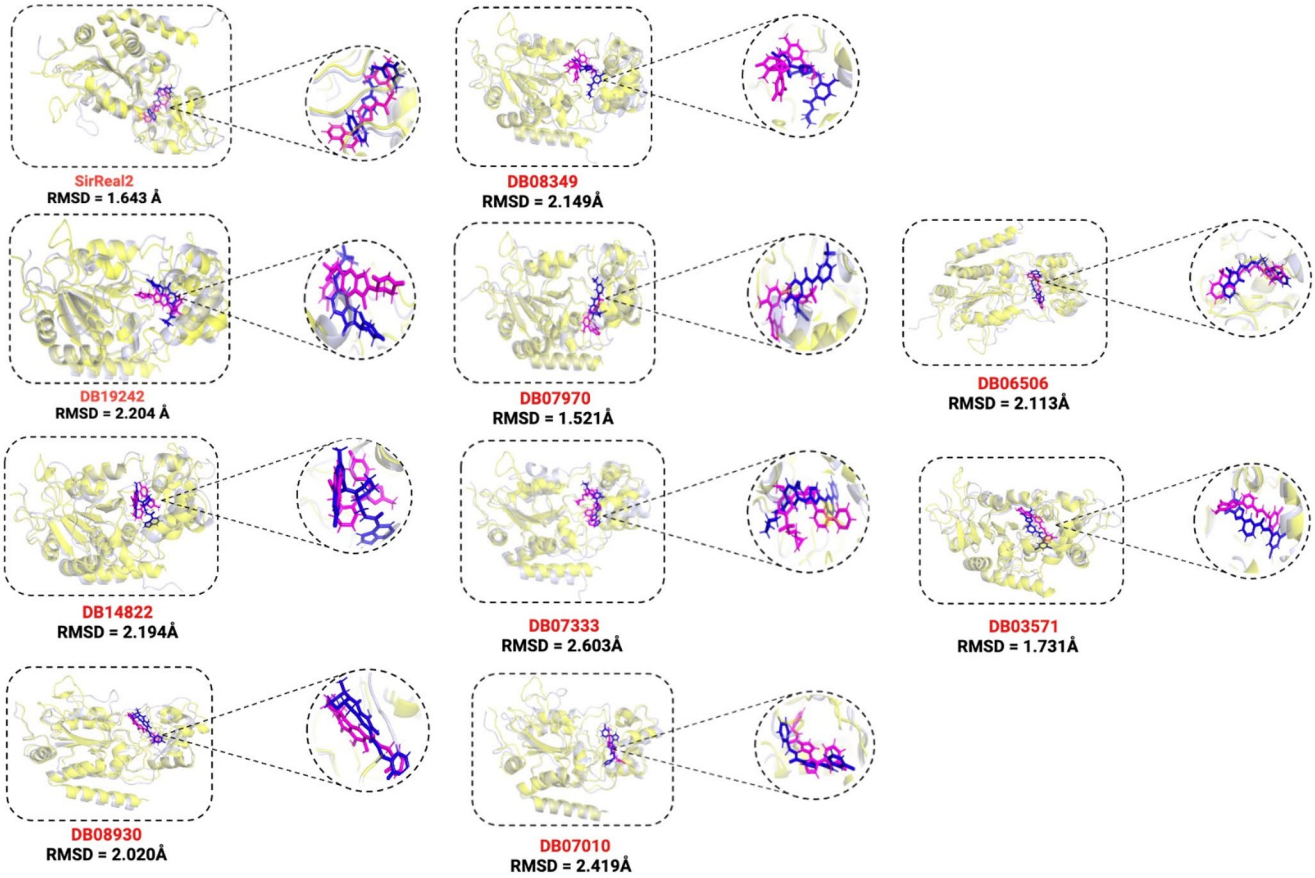

Fig. 13 Structural superposition of ligands before MD (blue) and after MD (magenta) within the SIRT2 binding site, showing conformational changes after 250 ns simulations

In contrast, DB08930's strong stability was accompanied by lower MM/GBSA affinity, likely driven by a pronounced polar solvation penalty (ΔEGB) offsetting gas-phase interactions. Together, these findings highlight the importance of integrating stability and affinity measures when prioritizing SIRT2 inhibitor candidates.

Therapeutically, SIRT2 inhibition has been associated with neuroprotection in models of Parkinson's and Huntington's diseases, partly through the reduction of α-synuclein aggregation and restoration of autophagic flux [43]. Given SIRT2's high expression in the central nervous system and its upregulation in Parkinson's disease, central nervous system (CNS) penetration is an important design consideration. DB14822 and DB07010 were classified as CNS^{+} ($logPS \geq -2$) based on in silico predictions, suggesting an enhanced likelihood of blood–brain barrier penetration. These pharmacokinetic attributes, coupled with their favourable binding characteristics, highlight these molecules as compelling candidates for further preclinical evaluation. A comparative assessment against established SIRT2 chemotypes, including SirReal2, thienopyrimidinone/thiazole-based scaffolds, and constrained nicotinamide analogs, indicates that the repurposed molecules DB14822, DB03571, and DB06506 exhibit MM/GBSA binding free energies that are comparable to, or surpass, those of benchmark inhibitors [44]. Structural analysis further demonstrates that these repurposed hits retain key interaction motifs characteristic of high-affinity SIRT2 ligands, notably π–π stacking and hydrophobic contacts with Phe119, Tyr139, and Phe190, which are similarly engaged by SirReal2 and related thienopyrimidinone derivatives in crystallographic and computational studies. Collectively, these data indicate that the identified hits are aligned with established structural determinants of SIRT2 inhibition while simultaneously introducing previously unreported chemotypes for this target, thereby expanding the chemical space available for SIRT2-focused drug discovery [41].

However, further work is needed. The 250 ns MD trajectories provide robust but finite sampling of the conformational landscape; longer simulations or enhanced sampling methods may reveal additional dynamic features. Furthermore, alchemical free energy perturbation (FEP) calculations could refine binding affinity estimates. Experimental validation, encompassing in vitro enzymatic assays, selectivity

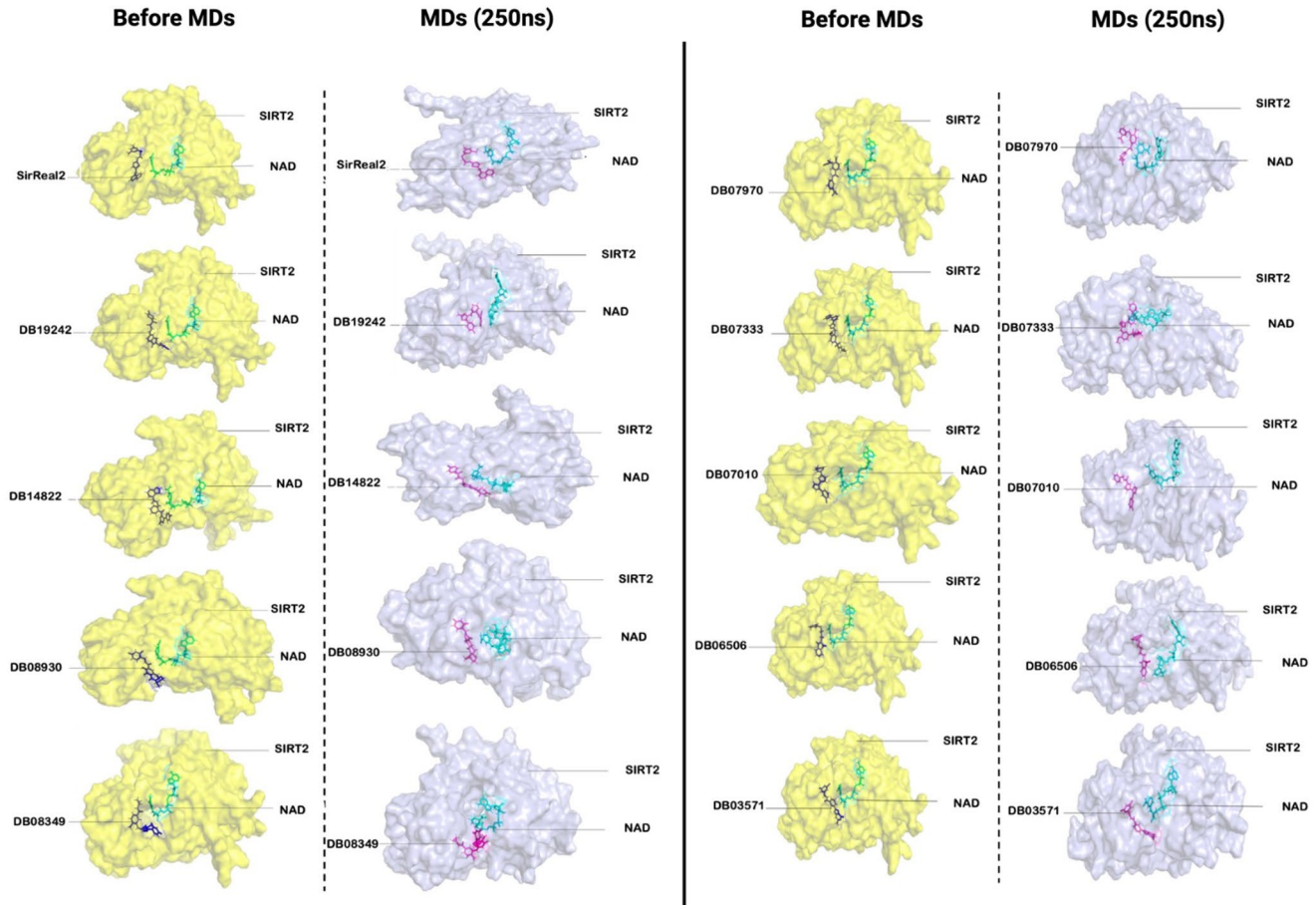

Fig. 14 Surface representation of SIRT2–ligand complexes before MD (blue) and after MD (magenta), with NAD shown in cyan, highlighting positional stability after 250 ns simulations

profiling against other sirtuin isoforms, and ADMET characterization, remains essential to confirm the computational predictions and assess safety liabilities, including potential hERG channel inhibition and CYP-mediated metabolism.

Conclusion

The present work identifies promising repurposed candidates for SIRT2 inhibition by integrating QSAR modeling with gradient boosting algorithms and applicability domain assessment. CatBoost-based predictions enabled the prioritization of DrugBank compounds with strong predicted potency and favorable drug-like properties, underscoring their potential in Parkinson's disease therapy. The QSAR-Boost framework developed herein is target-agnostic, reproducible, and adaptable to other drug discovery projects, providing a robust platform for accelerating lead identification from large chemical libraries. While the computational workflow identifies DB14822, DB03571, and DB06506 as strong SIRT2 inhibitor candidates, experimental validation remains essential. Future work will include biochemical inhibition assays, selectivity testing across sirtuin isoforms, and detailed ADMET characterization to substantiate the therapeutic potential of these repurposed compounds in Parkinson's disease.

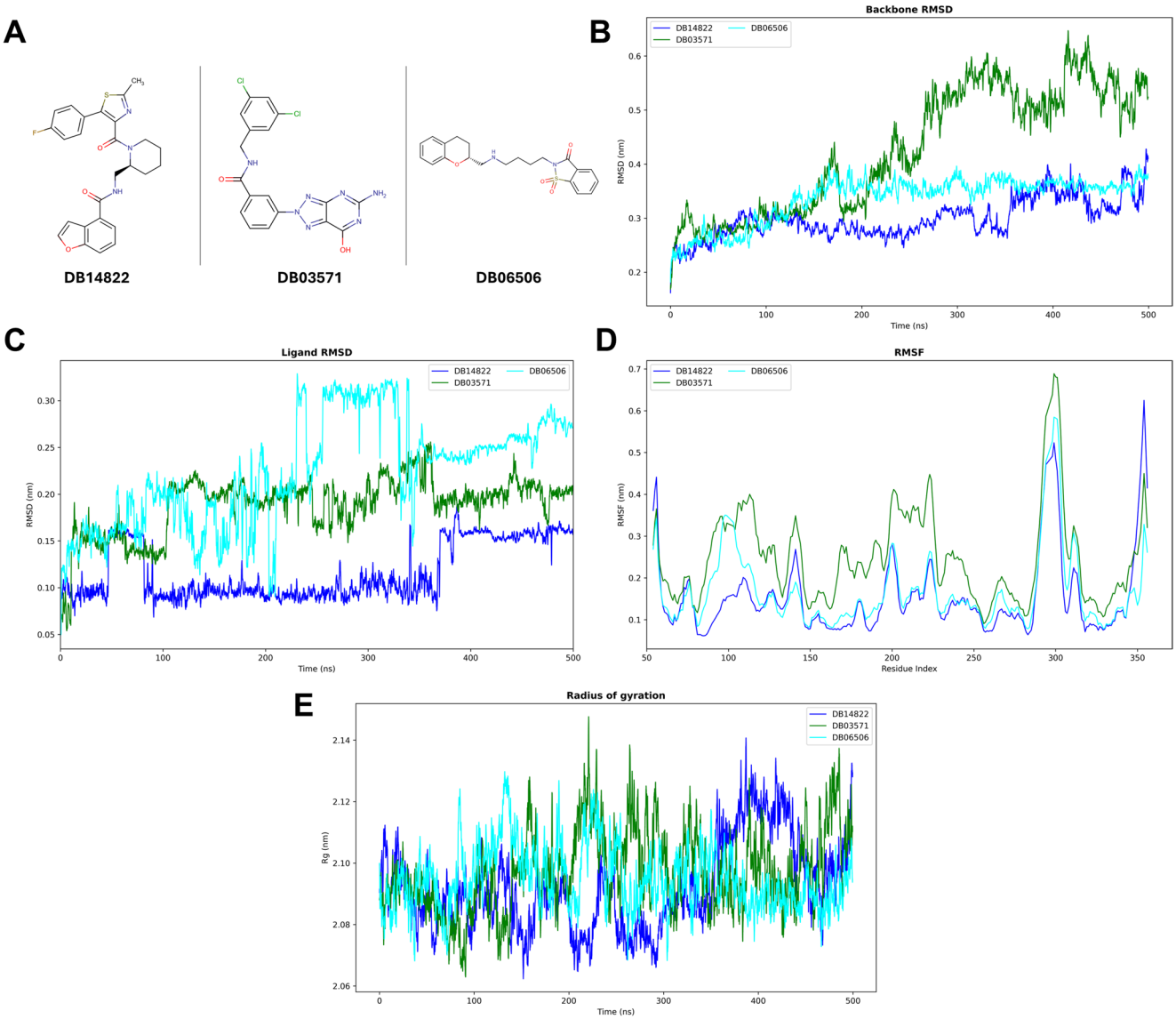

Fig. 15 **A** 2D chemical structures of the top three SIRT2 hits: DB14822, DB03571, and DB06506. **B** Backbone RMSD profiles over 500 ns. **C** Ligand RMSD plots. **D** Residue-wise RMSF analysis. **E** Radius of gyration (Rg) plots. The triplicate simulations demonstrate consistent structural stability and preserved binding behavior across independent trajectories

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11030-026-11504-7.

Author contributions Conceptualization, methodology, writing—original draft, software, project administration, writing—review and editing, investigation, resources; A.S.: Methodology, formal analysis, visualization, writing—review and editing; A.G. Formal analysis, resources, software, A.M. Formal analysis, writing—original draft, methodology, supervision, visualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding This research has received no funding.

Data availability All datasets, QSAR models, and the full virtual screening workflow used in this study are publicly available on GitHub: https://github.com/yboulaamane/QSARBoost.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Parkinson J (2002) An essay on the shaking palsy. 1817. J Neuropsychiatry Clin Neurosci. https://doi.org/10.1176/jnp.14.2.223
2. Hirsch L, Jette N, Frolkis A, Steeves T, Pringsheim T (2016) The incidence of Parkinson's disease: a systematic review and meta-analysis. Neuroepidemiology 46(4):292–300. https://doi.org/10.1159/000445751

3. Colombo D, Pnevmatikou P, Melloni E, Keywood C (2020) Therapeutic innovation in Parkinson's disease: a 2020 update on disease-modifying approaches. Expert Rev Neurother 20(10):1047–1064. https://doi.org/10.1080/14737175.2020.1800454
4. Fujiwara H et al (2002) A-Synuclein is phosphorylated in synucleinopathy lesions. Nat Cell Biol 4(2):160–164. https://doi.org/10.1038/ncb748
5. Beyer K, Domingo-Sàbat M, Ariza A (2009) Molecular pathology of Lewy Body diseases. Int J Mol Sci 10(3):724. https://doi.org/10.3390/IJMS10030724
6. F. R. Cabrero and E. H. Morrison, "Lewy Bodies," *StatPearls*, July 2023, Accessed 15 Mar 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK536956/
7. Bisi N et al (2021) α-Synuclein: an all-inclusive trip around its structure, influencing factors and applied techniques. Front Chem. https://doi.org/10.3389/fchem.2021.666585
8. Hernandez SM, Tikhonova EB, Karamyshev AL (2020) Protein-protein interactions in Alpha-Synuclein biogenesis: new potential targets in Parkinson's disease. Front Aging Neurosci 12:8. https://doi.org/10.3389/fnagi.2020.00072
9. Liu Y, Zhang Y, Zhu K, Chi S, Wang C, Xie A (2020) Emerging role of Sirtuin 2 in Parkinson's disease. Front Aging Neurosci 11:372. https://doi.org/10.3389/fnagi.2019.00372
10. "The mechanism of sirtuin 2–mediated exacerbation of alpha-synuclein toxicity in models of Parkinson disease | PLOS Biology". Accessed 12 July 2025. [Online]. Available: https://doi.org/10.1371/journal.pbio.2000374
11. Haigis MC, Sinclair DA (2010) Mammalian sirtuins: biological insights and disease relevance. Annu Rev Pathol Mech Dis 5(Volume 5):253–295. https://doi.org/10.1146/annurev.pathol.4.110807.092250
12. Sanders BD, Jackson B, Marmorstein R (2010) Structural basis for sirtuin function: what we know and what we don't. Biochim Biophys Acta 1804(8):1604–1616. https://doi.org/10.1016/j.bbapap.2009.09.009
13. Harrison IF, Smith AD, Dexter DT (2018) Pathological histone acetylation in Parkinson's disease: neuroprotection and inhibition of microglial activation through SIRT 2 inhibition. Neurosci Lett 666:48–57. https://doi.org/10.1016/j.neulet.2017.12.037
14. Elkouzi A, Vedam-Mai V, Eisinger RS, Okun MS (2019) Emerging therapies in Parkinson disease—repurposed drugs and new approaches. Nat Rev Neurol 15(4):204–223. https://doi.org/10.1038/s41582-019-0155-7
15. Pushpakom S et al (2019) Drug repurposing: progress, challenges and recommendations. Nat Rev Drug Discov 18(1):41–58. https://doi.org/10.1038/nrd.2018.168
16. Boulaamane Y et al (2023) Exploring natural products as multi-target-directed drugs for Parkinson's disease: an in-silico approach integrating QSAR, pharmacophore modeling, and molecular dynamics simulations. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2023.2260879
17. Boulaamane Y, Jangid K, Britel MR, Maurady A (2023) Probing the molecular mechanisms of α-synuclein inhibitors unveils promising natural candidates through machine-learning QSAR, pharmacophore modeling, and molecular dynamics simulations. Mol Divers. https://doi.org/10.1007/S11030-023-10691-X/METRICS
18. Boulaamane Y et al (2025) Computational screening of natural products as tryptophan 2,3-dioxygenase inhibitors: insights from CNN-based QSAR, molecular docking, ADMET, and molecular dynamics simulations. Comput Biol Med 191:110199. https://doi.org/10.1016/j.compbiomed.2025.110199
19. Landrum G (2013) RDKit: a software suite for cheminformatics, computational chemistry, and predictive modelling. Greg Landrum 8
20. McKinney W (2011) Pandas: a foundational Python library for data analysis and statistics. Python High Perform Sci Comput 14(9):1–9
21. Boldini D, Grisoni F, Kuhn D, Friedrich L, Sieber SA (2023) Practical guidelines for the use of gradient boosting for molecular property prediction. J Cheminform 15(1):73. https://doi.org/10.1186/s13321-023-00743-7
22. Chen T et al. (2015) Xgboost: extreme gradient boosting
23. Akiba T, Sano S, Yanase T, Ohta T, Koyama M (2019) Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. KDD '19. Association for Computing Machinery, New York, pp 2623–2631. https://doi.org/10.1145/3292500.3330701
24. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open Babel: an open chemical toolbox. J Cheminform 3(10):1–14. https://doi.org/10.1186/1758-2946-3-33/TABLES/2
25. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 17(5–6):5–6. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P
26. "OpenMMDL—Simplifying the complex: building, simulating, and analyzing protein–ligand systems in OpenMM. J Chem Inf Model. [Online]. Available: https://doi.org/10.1021/acs.jcim.4c02158. Accessed 14 July 2025
27. Abraham MJ et al (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1:19–25
28. Huang J, Mackerell AD (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem. https://doi.org/10.1002/JCC.23354
29. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E (2021) gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS. J Chem Theory Comput. https://doi.org/10.1021/acs.jctc.1c00645
30. Gaulton A (2012) ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 40:1100–1107
31. McKinney W (2011) Pandas: a foundational Python library for data analysis and statistics. Python High Perform Sci Comput 14(9):1–9
32. Gramatica P (2007) Principles of QSAR models validation: internal and external. QSAR Comb Sci 26(5):694–701. https://doi.org/10.1002/qsar.200610151
33. A new strategy of outlier detection for QSAR/QSPR—Cao (2010) J Comput Chem. Wiley Online Library. Accessed 22 July 2025. [Online]. Available: https://doi.org/10.1002/jcc.21351
34. Weaver S, Gleeson MP (2008) The importance of the domain of applicability in QSAR modeling. J Mol Graph Model 26(8):1315–1326. https://doi.org/10.1016/j.jmgm.2008.01.002
35. Selective Sirt2 inhibition by ligand-induced rearrangement of the active site. Nat Commun. Accessed 22 July 2025. [Online]. Available: https://www.nature.com/articles/ncomms7263
36. Kim S et al (2019) PubChem 2019 update: improved access to chemical data. Nucleic Acids Res 47(D1):D1102–D1109. https://doi.org/10.1093/NAR/GKY1033
37. Sun D, Gao W, Hu H, Zhou S (2022) Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B 12(7):3049–3062. https://doi.org/10.1016/j.apsb.2022.02.002
38. Modernizing preclinical drug development: the role of new approach methodologies. ACS Pharmacol Transl Sci. [Online]. Available: https://doi.org/10.1021/acsptsci.5c00162. Accessed 24 July 2025
39. Myung Y, de Sá AGC, Ascher DB (2024) Deep-PK: deep learning for small molecule pharmacokinetic and toxicity prediction.

Nucleic Acids Res 52(W1):W1. https://doi.org/10.1093/nar/gkae254
40. Rumpf T et al (2015) Selective Sirt2 inhibition by ligand-induced rearrangement of the active site. Nat Commun 6:6263. https://doi.org/10.1038/ncomms7263
41. Abbotto E et al (2023) Novel Thiazole-based SIRT2 inhibitors discovered via molecular modelling studies and enzymatic assays. Pharmaceuticals 16(9):9. https://doi.org/10.3390/ph16091316
42. Ai T, Wilson DJ, Chen L (2023) 5-((3-Amidobenzyl)oxy)nicotinamides as SIRT2 inhibitors: a study of constrained analogs. Molecules 28(22):7655. https://doi.org/10.3390/molecules28227655
43. de Oliveira RM et al (2017) The mechanism of sirtuin 2–mediated exacerbation of alpha-synuclein toxicity in models of Parkinson disease. PLoS Biol 15(3):e2000374. https://doi.org/10.1371/journal.pbio.2000374
44. Sundriyal S et al (2017) Thienopyrimidinone based sirtuin-2 (SIRT2)-selective inhibitors bind in the ligand induced selectivity pocket. J Med Chem 60(5):1928–1945. https://doi.org/10.1021/acs.jmedchem.6b01690