Pharmaceutical Sciences and Drug Design
ISSN: 3062-4428
2023, Volume 3, Page No: 329-346

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Computational Exploration of IP6K1 Inhibition: Ligand- and Structure-Based Insights for Obesity-Related Therapeutics

Kevin O'Donnell[1], Rachel S. Murphy[2]*, Brian Nolan[1]

[1]Department of Medicinal Chemistry, School of Pharmacy, Trinity College Dublin, Dublin, Ireland.
[2]Department of Pharmaceutical Sciences, School of Pharmacy, University of Arizona, Tucson, United States.

***E-mail** ✉ rachel.murphy@outlook.com

Received: 16 August 2023, **Revised:** 29 November 2023, **Accepted:** 01 December 2023

ABSTRACT

Current investigations have revealed an encouraging method for managing the increasing worldwide challenge of obesity and its comorbid conditions. Targeting inositol hexakisphosphate kinase 1 (IP6K1) through inhibition has surfaced as a viable treatment avenue. The present work applies diverse ligand-based computational modeling approaches to examine the key structural elements needed for benzisoxazole compounds to inhibit IP6K1 effectively. Initially, we generated linear 2D Quantitative Structure–Activity Relationship (2D-QSAR) models to balance explanatory clarity with robust forecasting capability. Subsequently, pharmacophore modeling from ligands was carried out to detect the critical chemical features driving the high potency of these molecules. To elucidate the three-dimensional aspects required for greater efficacy toward the IP6K1 target, various alignment strategies were used to build 3D-QSAR models. Because no experimental X-ray structure exists for IP6K1, a dependable homology model was constructed and thoroughly verified structurally, allowing structure-based studies on the chosen compound set. In addition, molecular dynamics simulations employing the docked configurations of these molecules yielded deeper understanding. The outcomes uniformly reinforced the explanatory insights gained from ligand-based as well as structure-based methods. This research supplies practical recommendations for developing new IP6K1 inhibitors. Significantly, all analyses were conducted using only freely available, non-proprietary software, facilitating easy replication of the described models.

Keywords: IP6K1 inhibitors, QSAR, Pharmacophore mapping, Homology modeling, molecular dynamics simulations

How to Cite This Article: O'Donnell K, Murphy RS, Nolan B. Computational Exploration of IP6K1 Inhibition: Ligand- and Structure-Based Insights for Obesity-Related Therapeutics. Pharm Sci Drug Des. 2023,3:329-46. https://doi.org/10.51847/xYqaC4o5uP

Introduction

During the past forty years, obesity rates around the world have increased without pause, impacting people regardless of age, ethnicity, or sex [1, 2]. This disturbing pattern has sparked numerous related medical problems, among them type 2 diabetes mellitus (T2DM), high blood pressure, abnormal blood lipids, heart and vessel diseases, non-alcoholic fatty liver disease/non-alcoholic steatohepatitis (NAFLD/NASH), issues with reproduction, breathing difficulties, mental and neurological disorders, and particular types of cancer [3-5]. Positively, integrating drugs with changes in daily habits has proven effective against obesity [6, 7]. A small weight or fat loss of just 5–10% can markedly decrease the chances of complications linked to obesity in grown-ups [8, 9]. Nevertheless, keeping weight off over time is hard, fueling major efforts to find fresh treatment paths for obesity and its metabolic consequences [2, 10].

Enzymes called inositol hexakisphosphate kinases (IP6Ks) are essential for adding a phosphate group to inositol hexakisphosphate (InsP6), forming 5-diphosphoinositol pentakisphosphate (5-InsP7 or 5PP-IP5, commonly IP7). This reaction starts the production of inositol pyrophosphates (PP-InsPs). New data suggest that blocking the

pathway for making PP-InsPs could help treat metabolic issues, bone loss, blood clots, infections, cancer spread, and problems tied to aging [11, 12]. These inositol pyrophosphates act as powerful signaling agents in eukaryotic cells, supporting key body functions like making ATP, releasing insulin, transmitting signals inside cells, moving cells, fixing DNA, and keeping energy levels stable [13-15].

Out of the three chief IP6K types, IP6K1 and IP6K2 appear in most body tissues, but IP6K3 is mostly in heart, skeletal muscle, and brain [16]. IP6K1 works in two ways: it dampens some insulin signals yet boosts insulin release from pancreas cells. It also lowers heat production in fat cells, cutting overall energy use [17]. In mice without IP6K1, insulin works better and energy spending rises, shielding them from obesity caused by rich diets, too much insulin in blood, and resistance to insulin. Less IP6K1 also helps treat NAFLD and NASH [18]. Such results highlight IP6K1 blockers as possible drugs for obesity and connected metabolic troubles [18-21].

Compounds with ring structures containing different atoms are often used in drug research to create new starting molecules for important targets [22-25]. Recently, Zhou *et al*. described a group of benzisoxazole compounds that block IP6K1, IP6K2, and IP6K3 to different extents [26]. Tests outside living organisms on HCT116 colon cancer cells showed strong blockers of IP6K1 and IP6K2 greatly cut inositol pyrophosphate amounts while leaving other inositol phosphates mostly unchanged. One top blocker dropped those levels by 66–81% with little effect elsewhere. Inside animals, these compounds eased obesity-linked damage and cut weight without changing how much was eaten.

Using computers to help design drugs is now seen as vital in early discovery stages [27-31]. Here, we did thorough computational modeling on this key compound group to find what structure parts boost blocking of IP6K1. We noticed a solid link ($R^2 \sim 0.85$) between how well they block IP6K1 and IP6K2, meaning features for IP6K1 likely apply wider. Tools like 2D-QSAR, 3D-QSAR, and pharmacophore from ligands helped spot these parts. To check our work, we matched ligand-based computer results with dynamics simulations on an IP6K1 model built from similar proteins. This effort helps make better IP6K1 blockers and fights obesity.

Materials and Methods

Dataset collection and preparation

Structures and activity data for 36 IP6K1 inhibitors were obtained from the recent publication by Zhou *et al*. [26], where IC50 values against IP6K1 were measured using a novel enzyme-coupled assay. These IC50 values were converted to pIC50 ($-\log10(IC50/10^6)$) and used as the dependent variable for all ligand-based modeling. For uniformity, the original SMILES strings provided by Zhou *et al*. were processed into canonical SMILES via RDKit and then generated as 3D structures in .sdf format using Discovery Studio Visualizer. Additional standardization of the 3D models was achieved with the Chemaxon Standardizer tool through the following procedures: (a) addition of explicit hydrogens, (b) aromatization, (c) 2D and 3D cleaning, (d) neutralization to achieve zero net charge, and (e) removal of any counterions.

The 2D-QSAR modeling

Descriptor calculation

Molecular descriptors were computed with the alvaDesc version 2.0.4 tool, available through the freely accessible OCHEM online platform (https://www.alvascience.com/alvadesc/) (accessed on 7 September 2023) [32]. The 3D structures of the compounds were energy-minimized using the Corina module integrated in the OCHEM server [33]. The complete dataset for 2D-QSAR construction was assembled by merging these calculated descriptors with the corresponding pIC50 values of the compounds.

Dataset division and model development

The full dataset was divided into a training set (80%) and a test set (20%) employing the open-source Python program SFS-QSAR-tool_v2 (https://github.com/ncordeirfcup/SFS-QSAR-tool, accessed on 12 September 2023) [34]. This split followed an activity-sorted strategy with an initial offset of 2, whereby compounds were first ordered by descending pIC50 (starting from the second entry), and every fifth compound was assigned to the test set. Model building proceeded in two phases. First, only descriptors from eight highly interpretable categories were used: molecular properties, functional group counts, 2D atom pairs, drug-likeness indices, ring descriptors, atom-centered fragments, and constitutional descriptors. In the second phase, the entire range of alvaDesc descriptors was included.

For the 2D-QSAR methodology, a multiple linear regression (MLR) approach was chosen. Two freely available programs were utilized to construct the MLR models:
(a) SFS-QSAR-tool_v2: This provides a user-friendly interface for creating linear, interpretable 2D-QSAR models. It implements sequential forward selection (SFS), adapted from the Mlxtend library code (http://rasbt.github.io/mlxtend/, accessed on 12 September 2023). SFS is a deterministic feature selection method that supports multiple scoring functions and cross-validation options to identify the optimal descriptors. Here, four scoring functions were tested: coefficient of determination (R^2), negative mean absolute error (NMAE), negative mean Poisson deviance (NMPD), and negative mean gamma deviance (NMGD). For each, models were built without cross-validation and with 5-fold cross-validation, producing eight (=4 × 2) models per descriptor set.
(b) Genetic-Algorithm v.4.1_2 (https://dtclab.webs.com/software-tools, accessed on 14 September 2023): This tool constructs linear interpretable MLR models via a stochastic genetic algorithm (GA) procedure, details of which are reported elsewhere [35]. During preprocessing, correlation and variance thresholds were set to 0.99 and 0.0001, respectively, to retain a diverse descriptor pool while removing constant or near-identical ones.

Evaluation of the models

The performance of the 2D-QSAR models was evaluated using standard validation metrics, primarily Q^2LOO (leave-one-out cross-validated R^2) and R^2Pred/Q^2F1 (external predictive R^2) [36, 37]. The former measures internal robustness on the training set, while the latter gauges predictive power on the external test set. Given the generation of multiple models per descriptor set through both stochastic and deterministic selection, these metrics were key for identifying the best-performing model.
Further statistical indicators included R^2, adjusted R^2 (R^2adj), mean absolute error (MAE), rm^2LOO with Δrm^2LOO (training set), rm^2test with Δrm^2test (test set), Q^2F2, and root mean square error of prediction (RMSEP) [38]. Descriptor inter-correlation was checked via the cross-correlation matrix, and multicollinearity in final models was quantified by variance inflation factor (VIF) [39]. Model robustness against chance correlation was verified through Y-randomization, using the cRp^2 parameter [40]. The applicability domain (AD) of selected models was defined via Williams plots, relating standardized residuals (for response outliers) to leverage values (for structural outliers) [38, 41].

Ligand-based pharmacophore modeling

Structure-based pharmacophore models were generated using the open-source Quantitative Pharmacophore Activity Relationship (QPHAR) tool recently made available [42]. For every compound, 50 conformers were produced separately via genetic algorithm and Confab methods, powered by Open Babel software.
QPHAR models were trained after partitioning the dataset into 26 training and 10 test compounds with the splitData.py script included in QPHAR. The underlying QPHAR methodology is thoroughly explained by Kohlbacher *et al.* [42] and in our earlier work [43]. Model training relied exclusively on the training set with the train.py script, applying random forest (RF) regression and these settings: fuzzy: True, weight type: distance, threshold: 1.5, number of estimators: 10, maximum depth: 3, metric: R^2. Activity predictions for the test set were obtained using the predict.py tool. Internal performance was judged by R^2, root mean square error (RMSE), standard error (SE), and median error (ME), whereas external performance was evaluated solely through R^2Pred. The derived pharmacophore models also served for compound alignment. The profile3DActivity.py tool in QPHAR was applied to produce pharmacophore-superimposed structures, which were later used for 3D-QSAR development [44].

The 3D-QSAR modeling

For 3D-QSAR analysis, the Open3DQSAR software was utilized, incorporating two distinct variable selection methods: (a) Fractional Factorial Design-based variable SELection (FFD-SEL) and (b) Uninformative Variable Elimination-based Partial Least Square (UVE-PLS) [45, 46].
Model construction involved testing two alignment approaches. The first was an unsupervised rigid body alignment procedure. Input .sdf files were first geometry-optimized via steepest descent minimization under the MMFF94 force field. Post-optimization, 500 conformers per ligand were produced using rdMolAlign.GetCrippenO3A in RDKit before alignment—a higher number than the typical 100, as it delivered

superior models here. A custom Python script named "alignment.py" handled the atom-based superimposition and is hosted at the GitHub repository: https://github.com/ncordeirfcup/InsilicoModeling_RdRp (accessed on 1 October 2023).

Model performance was evaluated through metrics including R^2, F-test values, leave-one-out (Q^2LOO), leave-two-out (Q^2LTO), and leave-many-out (Q^2LMO with 5 groups and 20 iterations) cross-validation, plus external R^2Pred. Contour maps were visualized with isocontour levels at PLS coefficients of +0.002 (green) and −0.002 (yellow) for steric fields, and +0.001 (blue) and −0.001 (red) for electrostatic fields. A comprehensive overview of the Open3DQSAR workflow is available elsewhere [47].

Homology modeling

The IP6K1 homology model was constructed via the SWISS-MODEL server using UniProt entry Q92551 (https://www.uniprot.org/, accessed on 8 October 2023) [48]. After template identification and multi-template modeling, structural validation was performed with the MolProbity server (http://molprobity.biochem.duke.edu/index.php, accessed on 10 October 2023) [49, 50], integrated within SWISS-MODEL. This evaluation pinpointed the AlphaFold structure (https://alphafold.ebi.ac.uk/, accessed on 10 October 2023) as yielding the highest-quality initial model [51, 52]. To enhance this AlphaFold-derived structure further, molecular dynamics (MD) refinement was conducted in Amber 20 [53], adhering to the protocol outlined by Nurisso *et al*. [54]. The refinement steps comprised: (i) two-phase minimization of the explicitly solvated system—first solvent/ions only, then the full system, (ii) gradual heating under NVT conditions followed by 2 ns NPT equilibration, (iii) a 50 ns production MD run in explicit solvent, and (iv) final minimization (5000 cycles of conjugate gradient) of the protein alone under Generalized Born implicit solvation.

These procedures ensured a high-quality IP6K1 homology model, with post-refinement validation again performed using MolProbity.

Molecular docking analysis

Docking was performed with the advanced CB-Dock2 pipeline, an improved iteration of CB-Dock developed by Yang Cao and colleagues [55, 56]. Accessible at http://cao.labshare.cn/clab/index.html (accessed on 12 October 2023), CB-Dock2 excels in binding-site prediction through its CurPocket algorithm, which detects cavities based on protein surface curvature. This proved particularly valuable for our homology model and for locating unknown ligand pockets. The three largest CurPocket-identified cavities (by volume) were chosen for docking, which was executed via the embedded AutoDock Vina engine [57].

Molecular dynamics simulations

The resulting docked complexes were subjected to 50 ns MD simulations following established protocols detailed previously [58, 59]. Ligand topologies were generated in Leap (Amber 14) using the general AMBER force field (GAFF) via Antechamber. Simulations employed the ff99SB force field with explicit TIP3P water in a cubic box maintaining an 8 Å buffer around the complex. Pressure and temperature were controlled with the Berendsen barostat and Langevin thermostat, respectively. Protein protonation states at pH 7.0 were assigned using the PDB2PQR server (https://server.poissonboltzmann.org/pdb2pqr, accessed on 14 October 2023) [60]. Trajectories were processed with PTRAJ and CPPTRAJ [61], and results visualized/plotted in QtGrace (https://sourceforge.net/projects/qtgrace/, accessed on 25 October 2023). Hydrogen-bond occupancy between ligands and receptor residues was quantified from trajectories. Enthalpic binding free energies were computed via MM-GBSA using MMPBSA.py in AMBER [62]. Entropy terms ($T\Delta S$) were estimated through normal-mode analysis on 100 frames sampled from the final 10 ns, employing a quasi-harmonic approach based on the covariance matrix of atomic fluctuations. Per-residue energy decomposition was also performed with the Amber MM-GBSA module to dissect contributions from binding-site residues [58, 62]. All energy terms—van der Waals, electrostatic, polar solvation, and non-polar solvation—were derived from 200 snapshots taken from the last 10 ns. These analyses yielded essential details on ligand–receptor contacts and complex stability.

Finally, the ADMET (adsorption, distribution, metabolism, excretion, and toxicity) properties of the three top-performing compounds (namely, 21, 15, and 20) were assessed via the admetSAR-2.0 online tool (http://lmmd.ecust.edu.cn/admetsar2, accessed on 06 January 2026) [63]. Owing to their close structural resemblance, the three molecules showed similar ADMET characteristics. Each exhibited low acute oral toxicity (category III: LD50 > 500 mg/kg but < 5000 mg/kg), moderate-to-low aqueous solubility, and positive profiles for human oral bioavailability and intestinal absorption. Additionally, all demonstrated blood–brain barrier (BBB) penetration, as evidenced by the BOILED-Egg diagram, computed with the SwissADME platform (http://www.swissadme.ch/, accessed on 06 January 2026) [64]. Predictions indicated potential hepatotoxicity, reproductive toxicity, myopathy (as OATP1B1 inhibitors), and respiratory toxicity for these molecules. In contrast, no evidence emerged for carcinogenicity, mutagenicity (Ames test), cardiac toxicity (hERG inhibition), nephrotoxicity, or skin sensitization.

Results and Discussion

The 2D-QSAR modeling

As per the plan in Materials and Methods, we started by finding top straight-line models connecting blocking strength to alvaDesc calculations, using sequential forward selection (SFS) and genetic algorithm (GA) to pick features. At first, multiple linear regression (MLR) 2D-QSAR setups used easy-to-understand alvaDesc items. Then, every descriptor was added to see if forecasting got better. All key findings appear in **Table 1**, covering models (M01–M09) from clear descriptors and (M10–M18) from the full set.

Table 1. Summary of the statistical results obtained for the MLR models based on different types of descriptors a.

Model ID	Evaluation Metric	CV Setting	Interpretable Descriptors			Model ID	All Descriptors		
			Q^2LOO	R^2Pred	Mean Value		Q^2LOO	R^2Pred	Mean Value
M01	R^2	None	0.733	0.427	0.580	M10	0.812	0.688	0.750
M02	NMAE	None	0.654	0.799	0.727	M11	0.820	0.868	0.844
M03	NMPD	None	0.733	0.427	0.580	M12	0.812	0.688	0.750
M04	NMGD	None	0.427	0.427	0.427	M13	0.829	0.405	0.617
M05	R^2	5	0.671	0.680	0.676	M14	0.704	0.526	0.615
M06	NMAE	5	0.662	0.253	0.458	M15	0.839	0.870	0.855
M07	NMPD	5	−0.898	0.605	−0.147	M16	0.823	0.849	0.836
M08	NMGD	5	−0.898	0.605	−0.147	M17	0.823	0.849	0.836
M09	GA-LDA	NA	0.800	0.785	0.793	M18	0.840	0.801	0.821

a The best models found are depicted in bold. b Scoring function used in SFS-based feature selection. c Cross-validation used in the SFS-based feature selection. d Descriptors belonging to the eight specific categories, as outlined in the Materials and Methods section.

As evident, the genetic algorithm (GA) produced the most predictive MLR model among those using interpretable descriptors, whereas the sequential forward selection (SFS) algorithm, employing the NMAE scoring function with five-fold cross-validation, generated the most robust MLR model when incorporating all descriptors. A comparison of models M09 (based on interpretable descriptors) and M15 (based on all descriptors) demonstrated a marked enhancement in predictive performance with the inclusion of the full descriptor set. Nevertheless, model M09 exhibited acceptable predictive capability, particularly given the small number of descriptors involved. Furthermore, models M10–M18 were predominantly composed of 2D and 3D topological descriptors. Thus, both M09 and M15 warrant consideration: the former facilitates interpretation of structural requirements, while the latter offers superior predictivity and identifies descriptors that more precisely capture those requirements. Detailed statistical parameters for M09 and M15 are provided in **Table 2.**

Table 2. Statistical results for the best 2D-QSAR models found, M09 and M15.

Equation	Statistical Results
Model M09 (Interpretable descriptors) **pIC_{50} = +0.169(±0.028) F04[C-C] −0.447(±0.139) CMC-50 −0.378(±0.131) nRCONHR −0.275(±0.029) H-047 +0.1(±0.031) CATS2D_01_LL + 5.125(±0.485)**	$N_{training}$ = 29, R^2 = 0.857, R^2_{adj} = 0.826, Q^2_{LOO} = 0.800, MAE = 0.201, $r_m^2{}_{LOO}$ = 0.724, $\Delta r_m^2{}_{LOO}$ = 0.088 N_{test} = 7, R^2_{Pred}/Q^2_{F1} = 0.785, Q^2_{F2} = 0.765, RMSEP = 0.309, $r_m^2{}_{test}$ = 0.706, $\Delta r_m^2{}_{test}$ = 0.125
Model M15 (All descriptors) **pIC_{50} = +0.223(±0.083) VE3sign_B(s) +3.079(±0.574) MATS4m +10.797(±1.593) SpMax2_Bh(v) −6.694(±1.071) G3i +10.984(±3.07) R5e+ −33.064(±6.179)**	$N_{training}$ = 29, R^2 = 0.890, R^2_{adj} = 0.866, Q^2_{LOO} = 0.839, MAE = 0.181, $r_m^2{}_{LOO}$ = 0.772, $\Delta r_m^2{}_{LOO}$ = 0.117 N_{test} = 7, R^2_{Pred}/Q^2_{F1} = 0.870, Q^2_{F2} = 0.858, RMSEP = 0.240, $r_m^2{}_{test}$ = 0.740, $\Delta r_m^2{}_{test}$ = 0.120

Notably, model development initially incorporated five descriptors, given the training set of 29 compounds, thereby adhering to the recommended 1:5 ratio of independent variables to training samples. To evaluate whether fewer descriptors could yield a statistically sound model, a 5% improvement threshold was applied via the SFS-QSAR-tool_v2, where a new descriptor was added only if it increased Q²LOO by at least 5%. For both M09 and M15, exactly five descriptors were retained, confirming the necessity of this number.

Plots of observed versus predicted activity for M09 and M15 are presented in **Figure 1**. In addition to Q²LOO and R²Pred, other internal and external validation metrics were satisfactory (e.g., MAE: 0.201 for M09, 0.181 for M15). Key assumptions for linear regression were also fulfilled, including low inter-descriptor correlation (maximum R: 0.469 and 0.404). Variance inflation factor (VIF) values were all below 2.0, ruling out multicollinearity. Moreover, cRp² values of 0.761 for M09 and 0.794 for M15 affirmed that both models were robust and not derived by chance.

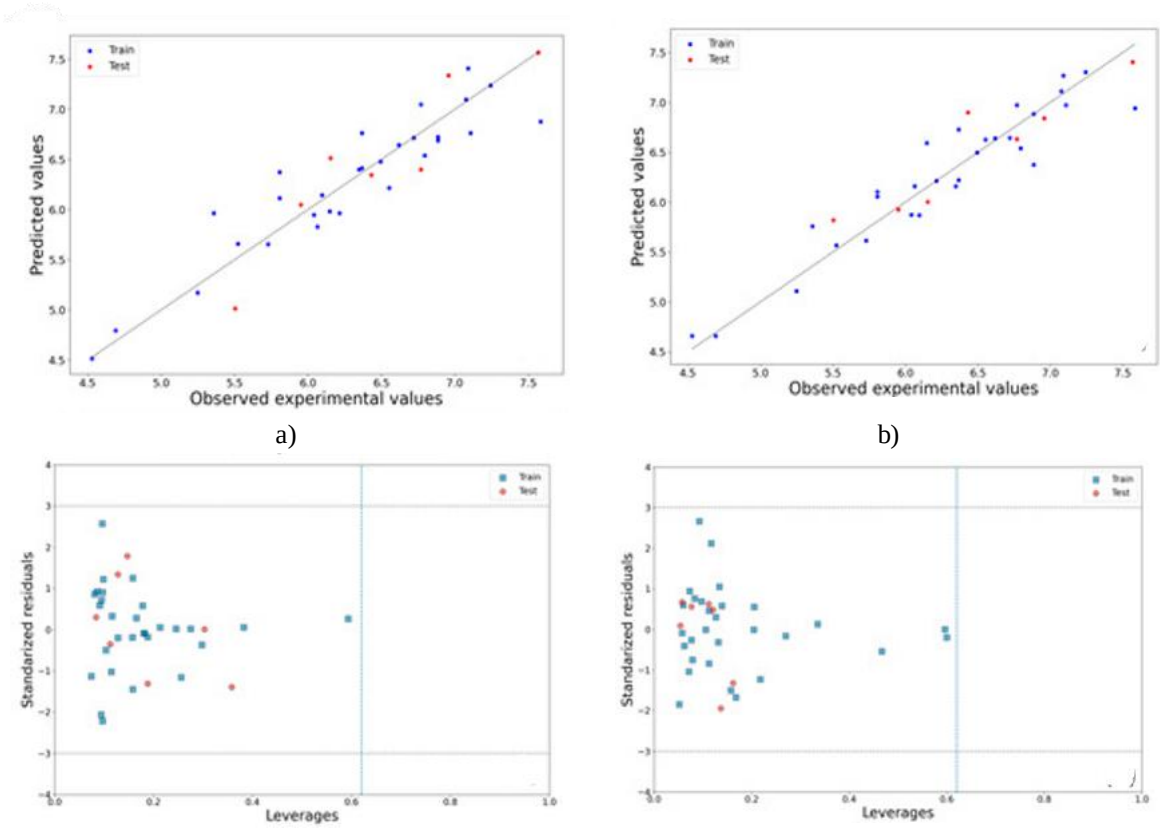


Figure 1. Observed vs. predicted activity plots for models M09 (a) and M15 (b), along with their corresponding Williams plots: M09 (c) and M15 (d).

The Williams plots in **Figure 1** further illustrate the applicability domain of these optimal 2D-QSAR models, revealing no structural or response outliers in either case. The relative importance of descriptors, based on standardized coefficients, is depicted in **Figure 2.**

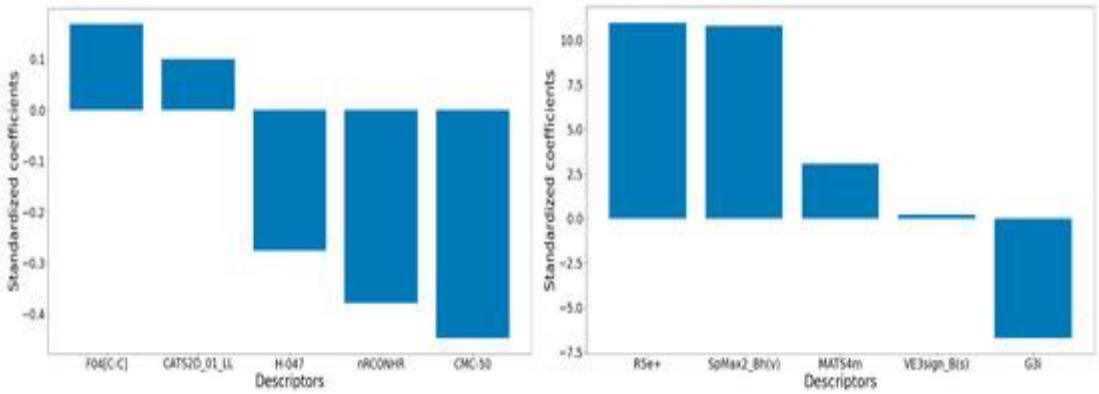


Figure 2. Relative significance of the descriptors of models M09 (left) and M15 (right).

Attention is first directed to the descriptors comprising model M09. The dominant descriptor is CMC-50, a drug-likeness index corresponding to the Ghose–Viswanadhan–Wendoloski Comprehensive Medicinal Chemistry (CMC) drug-like score at 50%. Its negative coefficient indicates that lower CMC-50 values favor greater biological activity. CMC-50 is a binary descriptor determined by ALOGP (Ghose–Crippen octanol–water partition coefficient), AMR (Ghose–Crippen molar refractivity), MW (molecular weight), and nAT (number of atoms). Values of 1 are assigned when ALOGP falls between 1.3 and 4.1, AMR between 70 and 110, MW between 230 and 390, and nAT between 30 and 55, otherwise, it is 0 **(Figure 3).** Closer examination shows that most high-activity compounds have low CMC-50 values. In particular, elevated lipophilicity (higher ALOGP) appears to be the primary driver, implying that hydrophobic interactions substantially influence the inhibitory potency of these compounds.

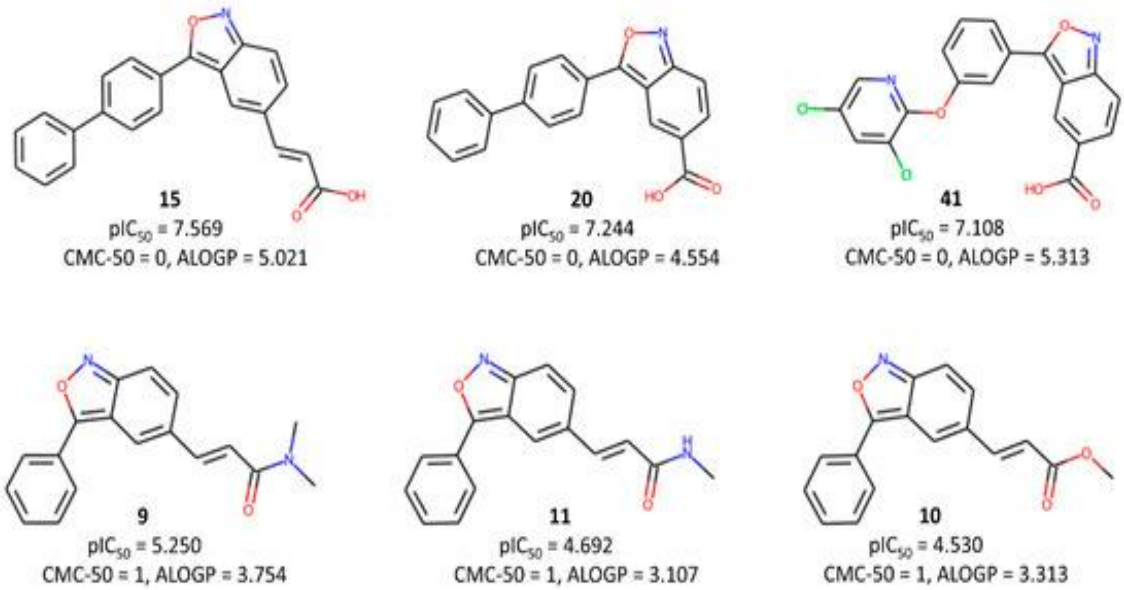


Figure 3. Typical examples highlighting the importance of the CMC-50 descriptor to the inhibitory activity.

The second most important descriptor in M09 is nRCONHR, the count of secondary amides (aliphatic). Its negative coefficient suggests that a higher count correlates with reduced activity. The third key descriptor, H-047 (hydrogens attached to $C^1(sp^3)/C^0(sp^2)$, where C^1 denotes sp^3 carbon without heteroatom attachment and C^0 denotes sp^2 carbon without heteroatom attachment), is elevated in lower-activity compounds **(Figure 4).** This pattern arises mainly from amide or ester side chains and unsubstituted phenyl rings in less active molecules.

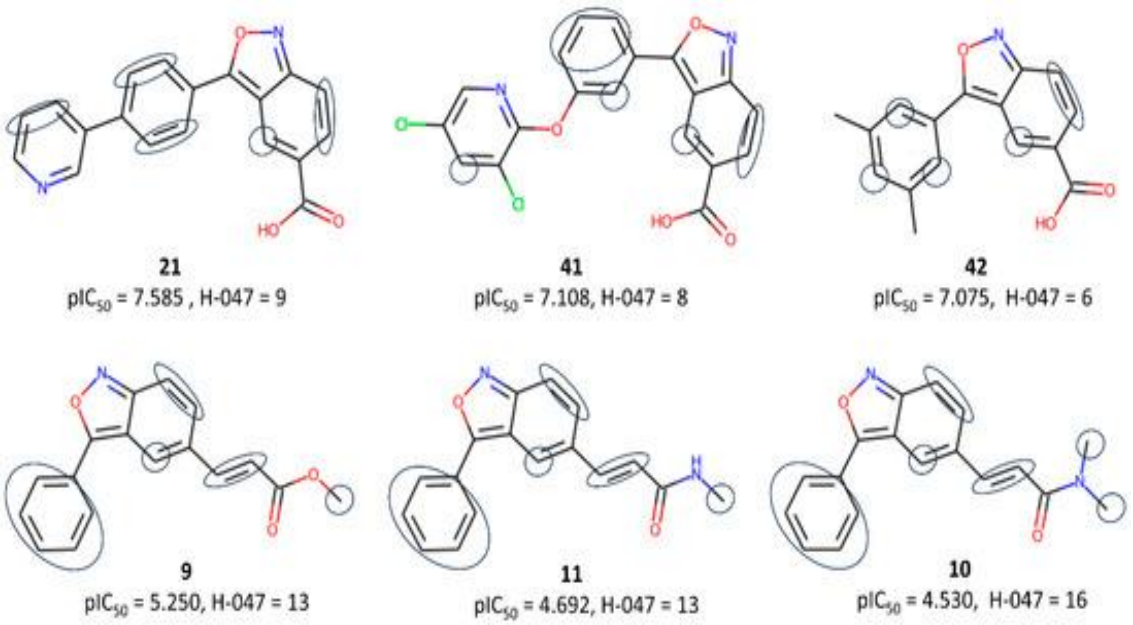


Figure 4. Typical examples highlighting the importance of the H-047 descriptor to the inhibitory activity.

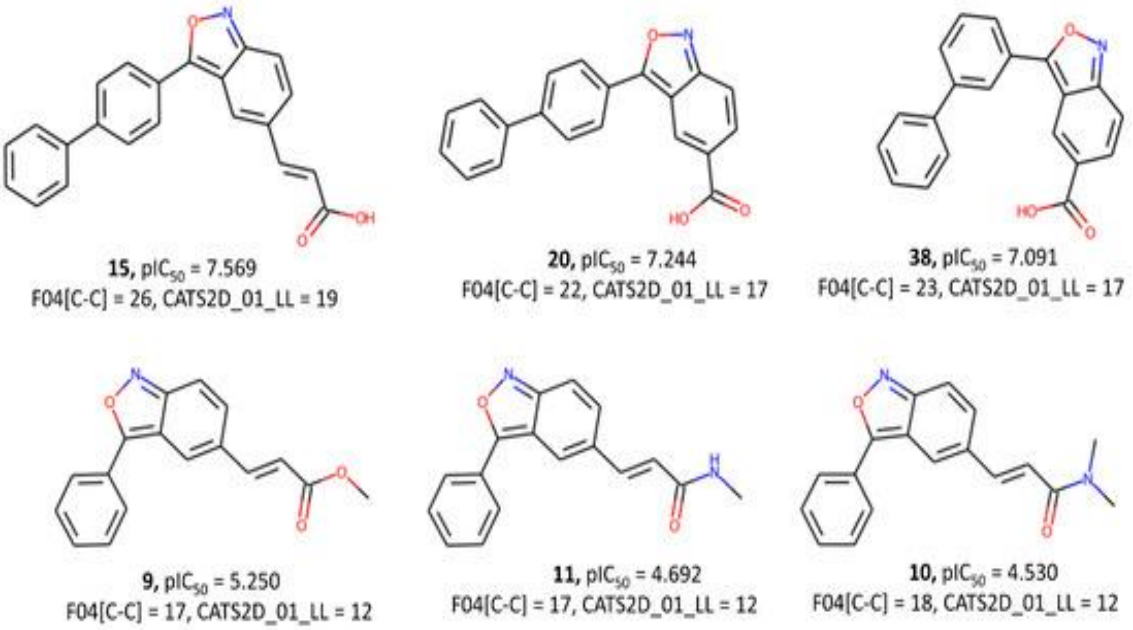


Figure 5. Typical examples illustrating the role of both F04[C-C] and CATS2D_01_LL descriptors in inhibitory activity.

Shifting focus to the descriptors within model M15, **Table 3** provides explanations for these terms. Notably, every descriptor in this model consists of intricate graph-derived topological indices. The descriptors R5e+ and G3i fall under 3D descriptors, with their magnitudes influenced by the precise 3D geometry of the molecules. The others qualify as 2D descriptors. The exceptional statistical predictivity of this model underscores that the exact topological arrangement of the molecules critically governs their biological potency. Nonetheless, the top two descriptors, R5e+ and SpMax2_Bh(v), exhibit nearly equivalent relative importance. These are respectively weighted by electronegativity and van der Waals volume. Hence, this points to the involvement of the molecules' 3D shape, alongside electrostatic and hydrophobic forces, in elevating biological activity. The third leading descriptor links to ionization potential, tied to molecular polarity, while MATS4m relates to atomic mass. In

summary, M09 primarily emphasizes hydrophobicity's role in boosting activity, whereas M15's descriptors reveal the need for equilibrium between hydrophobic/steric and electrostatic forces for optimal activity.

Table 3. The five descriptors present in the 2D-QSAR model M15.

Descriptor	Reworded Description	Descriptor Class
R5e+	Maximum autocorrelation at lag 5 calculated using Sanderson electronegativity weighting	GETAWAY
SpMax2_Bh(v)	Second highest eigenvalue derived from the Burden matrix, weighted by van der Waals volume	Burden eigenvalues
G3i	WHIM symmetry index (third directional component) computed with ionization potential weighting	WHIM
MATS4m	Moran index autocorrelation at lag 4, weighted according to atomic mass	2D autocorrelations
VE3sign_B(s)	Logarithmic sum of coefficients from the final Burden matrix eigenvector, weighted by intrinsic state (I-State)	2D matrix-based

Ligand-based pharmacophore mapping

Following the elucidation of structural needs via 2D-QSAR, ligand-based pharmacophore mapping was applied to create predictive models and discern pharmacophore elements linked to superior potency versus IP6K1. The QPHAR software generated these ligand-based pharmacophore models, splitting the dataset into 70% training and 30% test portions. Conformers produced by the genetic algorithm outperformed those from Confab in model quality. Statistical outcomes for the top pharmacophore model appear in **Table 4.**

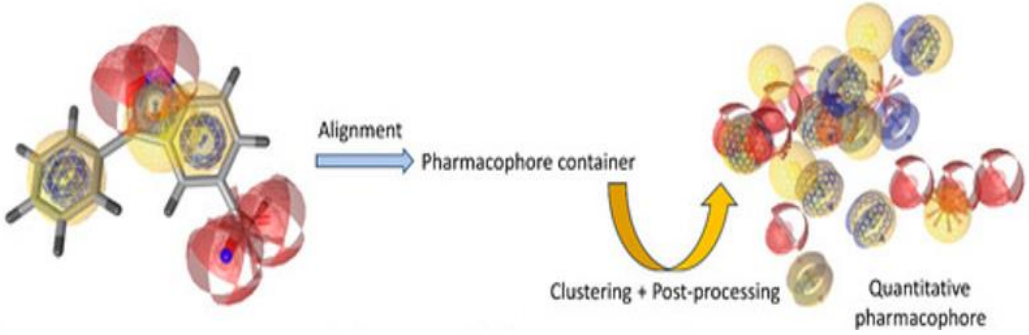
Table 4. Statistical results for the best QPHAR-based pharmacophore model found.

Parameter	Training	Test
N	26	10
R^2	0.845	
RMSE	0.309	
ME	0.248	
SE	0.183	
R^2_{Pred}		0.565
R^2_{Pred}[a]		0.716

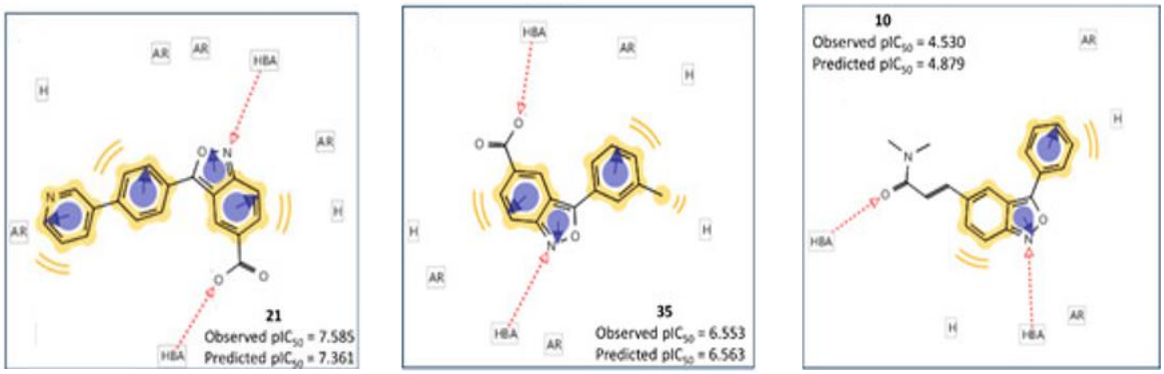
a After removal of one outlier.

This model showed solid internal predictivity, with R^2 of 0.845 and RMSE of 0.309. External validation yielded R^2Pred exceeding the 0.50 threshold, rising to 0.716 after excluding one compound. QPHAR notably picks the dataset's most rigid molecule as the template.

The template pharmacophore, aligned to compound 17, appears in **Figure 6**, including the pharmacophore container and final quantitative pharmacophore (hpmodel). Pharmacophore alignments for three compounds (21, 35, and 10) of varying activities are also shown.



a)



b) c) d)

Figure 6. (a) Pharmacophore features of the template molecule alongside the generated quantitative pharmacophore. The pharmacophore-aligned structures of compounds 21 (b), 35 (c), and 10 (d), along with the pharmacophore features, are also displayed.

Of these, compound 21 displays highest potency (pIC50 = 7.585), 10 the lowest (pIC50 = 4.530), and 35 intermediate (pIC50 = 6.553). The QPHAR model precisely forecasted their IP6K1 potencies. Hydrophobic and aromatic ring features declined progressively with falling affinity, while hydrogen bond acceptor features stayed fixed at two across all. These align with 2D-QSAR results, indicating hydrophobic contacts' key part in IP6K1 binding. Yet, the full QSAR model (M15), incorporating 2D/3D descriptors, suggested polar contacts' importance too. Compound 21 has four aromatic ring (AR) features, 35 three, and 10 two. Both 21 and 10 match two hydrophobic features, 35 matches three. Versus 10, 35 adds two key features—one AR and one hydrophobic—potentially accounting for their activity gap.

The 3D-QSAR analysis

For deeper understanding of potency requirements against IP6K1, 3D-QSAR analysis was conducted with open-source Open3DQSAR, adhering to prior alignment methods and feature selection. Rigid body alignment produced stronger, more predictive statistics, as in **Table 5**. The vital alignment process is shown visually in **Figure 6**.

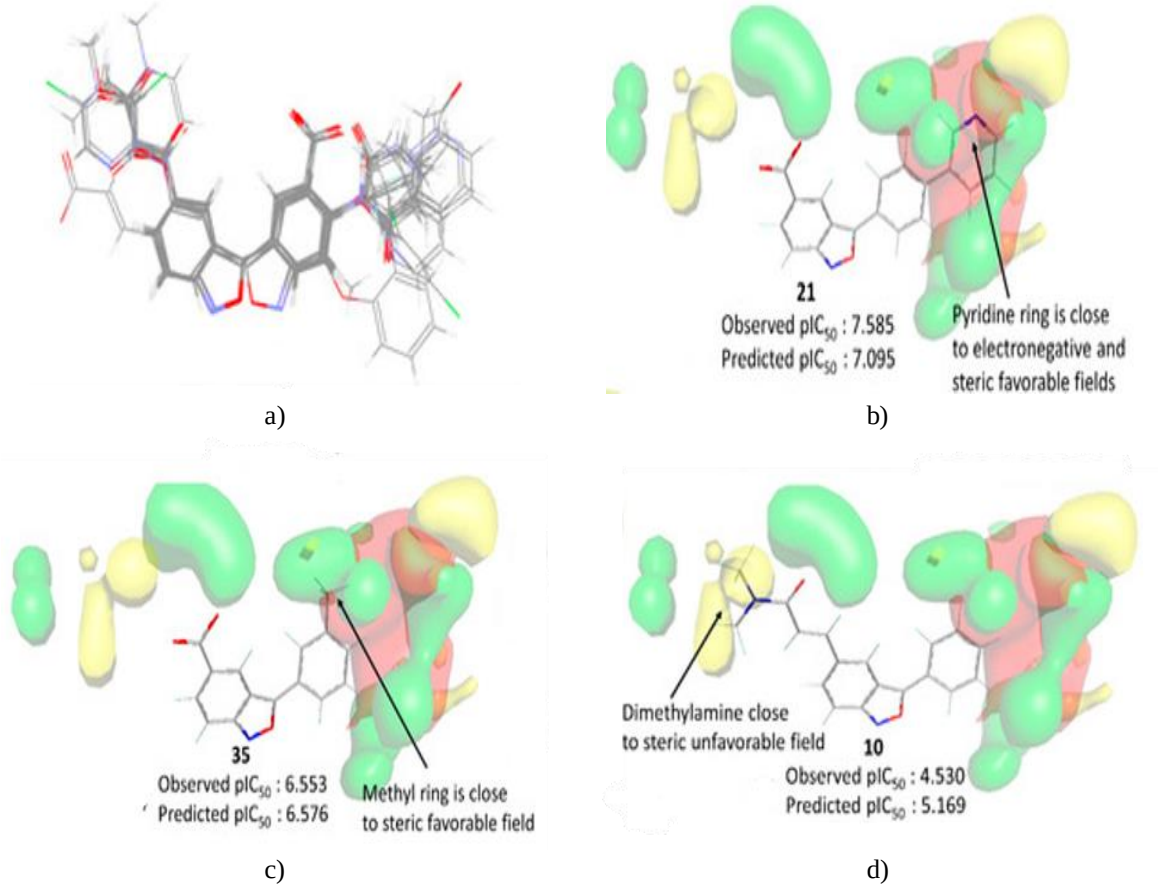
Table 5. Statistical results of 3D-QSAR models using different feature selection techniques a.

Parameter [b]	**FFD-SEL**	**UVE-PLS**
$\mathbf{N_{training}}$	29	29
NC[b]	4	3
$\mathbf{R^2}$ **(SDEC)**	0.912 (0.176)	0.856 (0.224)
F	62.157	33.847
$\mathbf{Q^2_{LOO}}$ **(SDEP)**	0.637 (0.357)	0.370 (0.471)
$\mathbf{Q^2_{LTO}}$ **(SDEP)**	0.626 (0.363)	0.361 (0.474)
$\mathbf{Q^2_{LMO}}$ **(SDEP)**	0.573 (0.387)	0.311 (0.492)
$\mathbf{N_{test}}$	7	7
$\mathbf{R^2_{Pred}}$ **(SDEP)**	0.747 (0.564)	0.668 (0.646)
$\mathbf{Q^2_s}$	0.428	---

a FFD-SEL: Fractional Factorial Design-based variable SELection, UVE-PLS: Uninformative Variable Elimination-based Partial Least Square. b NC: Number of principal components, SDEC: Standard error of calculation, SDEP: Standard error of prediction.

Evidently, FFD-SEL proved highly effective, delivering strong internal/external predictivity: Q^2LOO of 0.637 and R^2Pred of 0.747. Amid 3D-QSAR's sensitivity—particularly for small datasets—this model maintains steady moderate-to-good predictivity. Its distinction lies in inherent robustness, shown by sharp cross-validation decline after response scrambling.

Figure 7 presents contour maps from FFD-SEL, for compounds 25, 31, and 10. These depict steric/electrostatic influences, with steric at 60% and electrostatic at 40%. Contours show two main zones: right side blends steric/electrostatic, left side steric only. Right-side contours prove pivotal for potency.



a) b) c) d)

Figure 7. (a) Aligned structures for the dataset compounds. Contour maps obtained using the best 3D-QSAR model for compounds (b) 21, (c) 35, and (d) 10. The color codes used range from green (steric favorable) to red (electrostatic favorable).

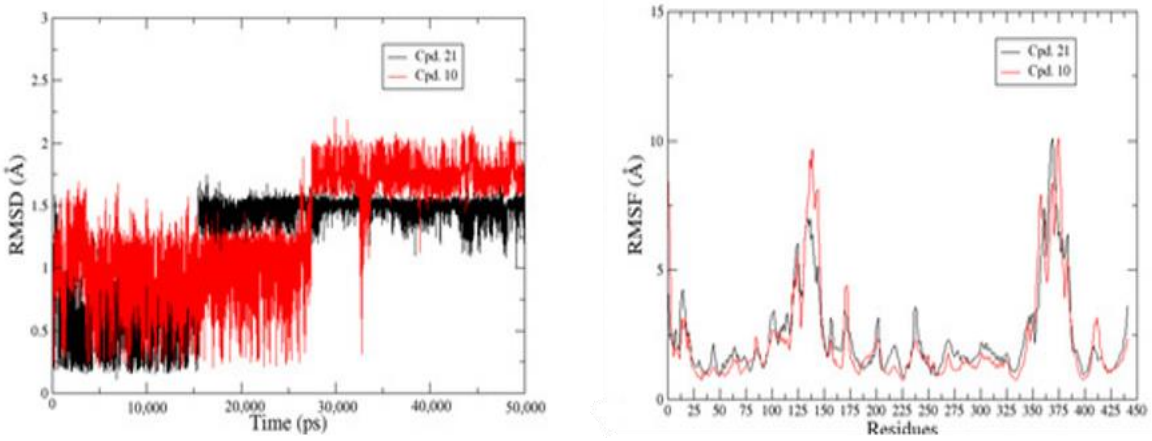
For potent compound 21, favorable steric/electronegative regions near its pyridine imply dual hydrophobic/electrostatic receptor engagement. Intermediate 35 misses some electronegative contour but gains via methyl in hydrophobic zone. Least active 10 evades both, its dimethylamine side chain hitting steric unfavorable area. Strikingly, no contours appear near benzisoxazole rings, confirming rigid body alignment success.

Homology modeling of IP6K1 and MD simulations

As of now, no experimentally determined X-ray structure exists for IP6K1. For this reason, all structure-based work had to rely on a computed homology model. In this study, the model was built using the UniProt accession Q92551 through the SWISS-MODEL server. Five high-ranking templates based on GMQE (Global Model Quality Estimation) scores were considered, and the AlphaFold DB structure—with a MolProbity score of 1.70—stood out as the best, outperforming alternatives that scored above 2.0. That said, the starting AlphaFold model still needed optimization, given its clash score of 0.58 and Ramachandran statistics of 86.10% favored and 4.10% outliers. Refinement was achieved by running molecular dynamics (MD) simulations on the initial model, producing a final version with a much improved MolProbity score of 0.79. This polished model showed 95.90% Ramachandran favored residues and zero outliers. The optimized homology model was then used for all subsequent docking and MD studies.

Despite the structural improvements, pinpointing credible ligand-binding pockets in IP6K1 proved challenging. To tackle this, binding-site detection was performed with the CB-Dock2 online tool, followed by docking of the

compounds into the three highest-scoring cavities using AutoDock Vina. Strikingly, the strongest and weakest inhibitors both preferred the same pocket (centered at X = 63 Å, Y = 77 Å, Z = 40 Å, volume 801 Å^3) and achieved the best Vina docking scores. Each complex was then subjected to 50 ns of explicit-solvent MD simulation to explore its dynamic properties. Attention was first paid to whether the ligands remained stably bound in the predicted site. Ligand RMSD trajectories **(Figure 8)** indicate that the more potent compound 21 initially moved away from its docking pose but settled into a stable position after approximately 15 ns. The weaker compound 10 similarly drifted early in the simulation and only partially stabilized around 30 ns, yet showed greater overall mobility than compound 21. Corresponding protein RMSF analysis revealed considerably larger fluctuations in the segment comprising residues 125–180 when bound to compound 10 compared to compound 21.



a) b)

Figure 8. Ligand RMSD trajectories (a) and protein RMSF profiles (b) for the potent compound 21 and the weaker compound 10.

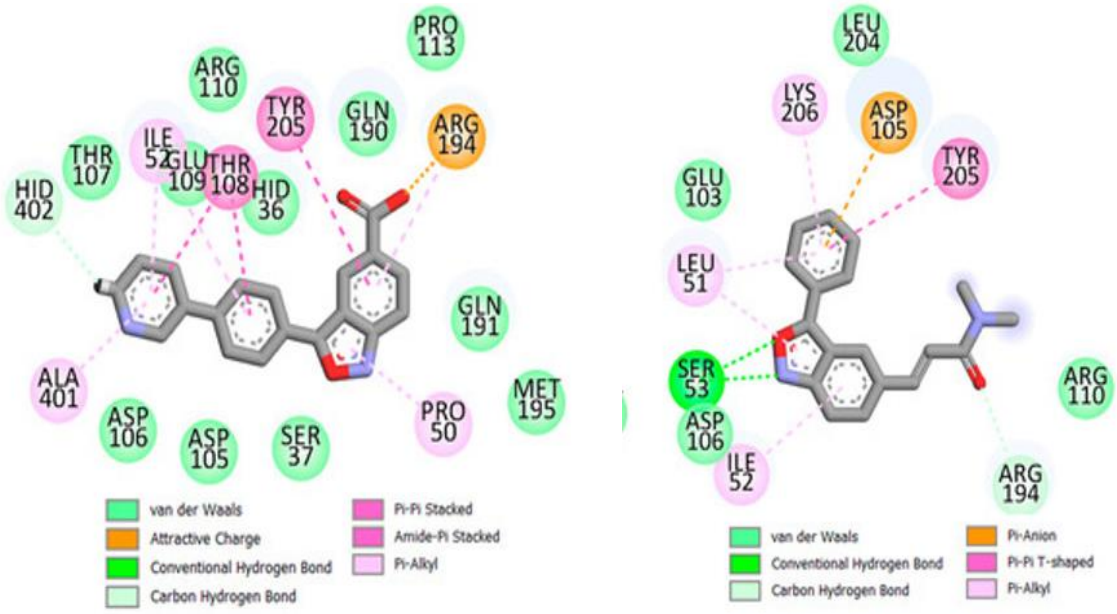
Binding free energies were subsequently estimated via MM-GBSA calculations on the MD frames to quantify enthalpic contributions. Data presented in **Table 6** clearly show that compound 21 enjoys a far more favorable predicted binding energy than compound 10, driven mainly by markedly stronger electrostatic terms. Both polar and non-polar contact energies were superior for compound 21. While compound 10 gained some advantage from a lower solvation penalty—likely owing to its compact size—its substantially weaker residue interactions ultimately led to poorer overall affinity.

Table 6. Calculated binding free energies [ΔGbind(T)] for selected IP6K1 complexes a. All the components shown are in kcal/mol.

Complexes	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{polar}	$\Delta G_{non\text{-}polar}$	$\Delta G_{solvation}$	$T\Delta S$	$\Delta G_{bind}(T)$
21	−42.45	−125.22	−167.67	+133.91	−5.78	+128.13	−21.20	−18.35
10	−38.26	−5.73	−43.99	+18.32	−4.25	+14.07	−24.26	−5.67

a ΔGbind(T): theoretical binding free energy (ΔGbind(T) = ΔEvdW + ΔEelec + ΔGpolar + ΔGnon-polar − TΔS) and its components, namely ΔEvdW: van der Waals interaction energy, ΔEelec: electrostatic interaction energy, ΔGpolar: polar solvation free energy, ΔGnon-polar: non-polar solvation free energy, TΔS: entropy.

Representative structures averaged over the last 10 ns of each trajectory are displayed in **Figure 9**. The more active compound 21 forms substantially more contacts with pocket residues than compound 10, with the extra interactions predominantly hydrophobic in character. A prominent polar contact is also visible between the carboxylate of compound 21 and Arg194. This same carboxylate engages firmly with both Gln190 and Arg194, playing a central role in anchoring the ligand securely.



a) b)

Figure 9. Final binding modes of compounds 21 (a) and 10 (b) in the IP6K1 complexes as obtained from the MD simulations.

Conversely, the benzisoxazole ring in the weaker compound 10 established hydrogen bonds with Ser53 and Asp106. These contacts were also highlighted by the pharmacophore model. In line with the pharmacophore findings, all three aromatic groups in compound 21 (comprising one phenyl and one benzisoxazole) participated in π–π or π–amide stacking with Thr108 and Tyr205, whereas only a single aromatic group in compound 10 formed such a contact with Tyr205. Notably, the polar side chain of compound 10 extended toward the solvent and contributed fewer interactions, while the hydrophobic portions of compound 21 formed extensively more contacts with surrounding residues. These patterns align with the 2D-QSAR results, where descriptors such as CMC-50 and SpMax2_Bh(v) ranked highest in importance, and with the 3D-QSAR outcomes, where the side chain of compound 10 occupied a sterically unfavorable region.

The 3D-QSAR analysis especially underscored the importance of the pyridine ring in compound 21, positioned near favorable electronegative and steric contours. This ring engaged in multiple contacts, including π–amide, π–alkyl, van der Waals, and carbon–hydrogen bonds.

Given the differing polar contacts that these ligands make with various residues, evaluating hydrogen-bonding contributions is essential. Trajectory examination was therefore performed to quantify hydrogen bonds over time. On average, compound 21 maintained more hydrogen bonds than compound 10. Detailed inspection showed that the dominant contacts for 21 involved its carboxylate group with Arg194 and Gln190, likely providing critical anchoring stability within the pocket. For compound 10, the primary hydrogen bonds originated from its benzisoxazole ring with Ser53.

To thoroughly assess the individual contributions of binding-site residues to the affinity of compounds 21 and 10, per-residue energy decomposition was conducted. As illustrated in **Figure 10**, Arg194 and Gln190—which form hydrogen bonds with the carboxylate of 21—delivered the strongest energetic contributions to the binding of 21. Their impact was 2–3 times greater than that of other residues. The lack of comparable contacts in compound 10 explains the large gap in electrostatic energy (ΔEelec) between the two ligands. For better visualization of remaining interactions, contributions from these two residues were omitted in **Figure 10**. This adjusted view clearly shows that contacts with residues like His36, Ser37, Asp106, Thr107, Thr108, Glu109, and Glu191 were substantially stronger for compound 21 than for 10. Moreover, most of these enhanced interactions with 21 were non-polar in nature. This detailed breakdown offers key insight into the contrasting binding modes of compounds 21 and 10 with the receptor residues.

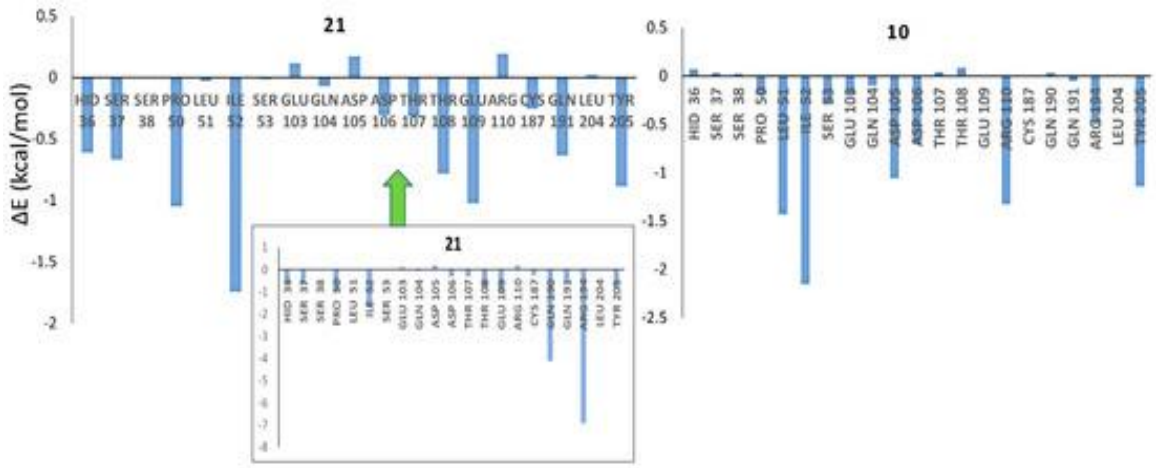


Figure 10. Total energy contributions of amino acid residues from per-residue decomposition analysis for the complexes with compounds 21 and 10.

Conclusion

In recent years, the therapeutic potential of IP6K1 has been thoroughly investigated, with growing data indicating that blockers of this kinase and its isoforms could offer effective options for managing obesity and associated metabolic conditions. Despite these promising insights, only a handful of IP6K1-targeted inhibitors have been documented to date. As a result, the present research marks the inaugural extensive ligand-based in silico modeling study dedicated to IP6K1 inhibitors. Of particular note, the employed dataset spanned over three orders of magnitude in inhibitory potency against IP6K1, even with constrained structural variation, making it especially suitable for detailed analysis.

The lack of an experimental crystal structure for human IP6K1 further highlights the value of ligand-based approaches in inhibitor development. This study's multi-layered goals commenced with the careful construction of robust, predictive ligand-based models. Among them, the 2D-QSAR model built from the complete alvaDesc descriptor set achieved the strongest statistical performance. Key structural insights derived from these models stressed the necessity of equilibrium between hydrophobic and electrostatic forces, combined with specific 3D topological/geometrical elements (e.g., CATS2D_01_LL [65], F04[C-C]), for effective receptor binding. Specifically, the optimal 2D-QSAR model, based on a small descriptor subset, revealed critical structural and topological determinants of high versus low potency. The ligand-based pharmacophore model further identified the essential roles of aromatic ring, hydrophobic, and hydrogen bond acceptor moieties in driving activity. Although yielding somewhat restrained details, the 3D-QSAR analysis clearly demonstrated the impact of electrostatic and steric factors on anti-IP6K1 activity. Complementing this, the IP6K1 homology model—coupled with docking and MD simulations—uncovered substantial contributions from polar contacts involving residues such as Glu190 and Arg194, especially with the negatively charged carboxylate functionality, to overall potency. Beyond polar interactions, hydrophobic, π–π, and π–amide contacts were found vital for stabilizing ligand occupancy in the pocket. Overall, these results offer valuable direction for future design of innovative IP6K1 inhibitors. Notably, the dataset's most potent entry (compound 21) displays approximately 26-fold selectivity for IP6K1 and IP6K2 over IP6K3. Compound 20 shows 4.2-fold preference for IP6K1 relative to IP6K3. Compound 20 has additionally proven effective in mitigating obesity-associated issues, such as enhancing glucose control, reducing liver steatosis, and limiting weight increase, independent of food consumption changes. This contrasts with prior IP6K inhibitors like SC-919 (IP6K1: IC50 < 5.6 nM, IP6K3: IC50 = 0.65 nM) and N2-(m-trifluorobenzyl)-N6-(p-nitrobenzyl)-purine (TNP) (IP6K1: IC50 = 270 nM, IP6K3: IC50 = 260 nM) [13, 26]. Such selectivity profiles suggest distinct therapeutic potentials for these new compounds, differentiating them from known agents and meriting deeper investigation for obesity therapy. The developed ligand-based models are suitable for screening novel analogs' activities, and the exclusive use of freely available software enhances the practical applicability of the findings. As a prospective extension, the docked complexes could be simulated over longer timescales (e.g., 500 ns) in future MD studies.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Twig G, Yaniv G, Levine H, Leiba A, Goldberger N, Derazne E, et al. Body-mass index in 2.3 million adolescents and cardiovascular death in adulthood. N Engl J Med. 2016;374:2430–40.
2. Müller TD, Blüher M, Tschöp MH, DiMarchi RD. Anti-obesity drug discovery: advances and challenges. Nat Rev Drug Discov. 2021;21:201–23.
3. Blüher M. Obesity: global epidemiology and pathogenesis. Nat Rev Endocrinol. 2019;15:288–98.
4. Srivastava G, Apovian C. Future pharmacotherapy for obesity: new anti-obesity drugs on the horizon. Curr Obes Rep. 2018;7:147–61.
5. Safaei M, Sundararajan EA, Driss M, Boulila W, Shapi'I A. A systematic literature review on obesity: understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. Comput Biol Med. 2021;136:104754.
6. Wadden TA, Berkowitz RI, Sarwer DB, Prus-Wisniewski R, Steinberg C. Benefits of lifestyle modification in the pharmacologic treatment of obesity. Arch Intern Med. 2001;161:218–27.
7. Kim GW, Lin JE, Blomain ES, Waldman SA. Antiobesity pharmacotherapy: new drugs and emerging targets. Clin Pharmacol Ther. 2013;95:53–66.
8. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, Torres-Gonzalez A, Gra-Oramas B, Gonzalez-Fabian L, et al. Weight loss through lifestyle modification significantly reduces features of nonalcoholic steatohepatitis. Gastroenterology. 2015;149:367–78.e5.
9. Wing RR, Lang W, Wadden TA, Safford M, Knowler WC, Bertoni AG, et al. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. Diabetes Care. 2011;34:1481–6.
10. Williams DM, Nawaz A, Evans M. Drug therapy in obesity: a review of current and emerging treatments. Diabetes Ther. 2020;11:1199–216.
11. Cridland C, Gillaspy G. Inositol pyrophosphate pathways and mechanisms: what can we learn from plants? Molecules. 2020;25:2789.
12. Li H, Datunashvili M, Reyes RC, Voglmaier SM. Inositol hexakisphosphate kinases differentially regulate trafficking of vesicular glutamate transporters 1 and 2. Front Cell Neurosci. 2022;16:926794.
13. Moritoh Y, Abe SI, Akiyama H, Kobayashi A, Koyama R, Hara R, et al. The enzymatic activity of inositol hexakisphosphate kinase controls circulating phosphate in mammals. Nat Commun. 2021;12:4847.
14. Illies C, Gromada J, Fiume R, Leibiger B, Yu J, Juhl K, et al. Requirement of inositol pyrophosphates for full exocytotic capacity in pancreatic β cells. Science. 2007;318:1299–302.
15. Wilson MSC, Livermore TM, Saiardi A. Inositol pyrophosphates: between signalling and metabolism. Biochem J. 2013;452:369–79.
16. Liao G, Ye W, Heitmann T, Ernst G, DePasquale M, Xu L, et al. Identification of small-molecule inhibitors of human inositol hexakisphosphate kinases by high-throughput screening. ACS Pharmacol Transl Sci. 2021;4:780–9.
17. Zhu Q, Ghoshal S, Tyagi R, Chakraborty A. Global IP6K1 deletion enhances temperature modulated energy expenditure which reduces carbohydrate and fat induced weight gain. Mol Metab. 2016;6:73–85.

18. Mukherjee S, Chakraborty M, Ulmasov B, McCommis K, Zhang J, Carpenter D, et al. Pleiotropic actions of IP6K1 mediate hepatic metabolic dysfunction to promote nonalcoholic fatty liver disease and steatohepatitis. Mol Metab. 2021;54:101364.
19. Rao F, Xu J, Fu C, Cha JY, Gadalla MM, Xu R, et al. Inositol pyrophosphates promote tumor growth and metastasis by antagonizing liver kinase B1. Proc Natl Acad Sci U S A. 2015;112:1773–8.
20. Chakraborty A. The inositol pyrophosphate pathway in health and diseases. Biol Rev. 2017;93:1203–27.
21. Moritoh Y, Oka M, Yasuhara Y, Hozumi H, Iwachidow K, Fuse H, et al. Inositol hexakisphosphate kinase 3 regulates metabolism and lifespan in mice. Sci Rep. 2016;6:32072.
22. Mousavi H. A comprehensive survey upon diverse and prolific applications of chitosan-based catalytic systems in one-pot multi-component synthesis of heterocyclic rings. Int J Biol Macromol. 2021;186:1003–1166.
23. Kumar A, Singh AK, Singh H, Vijayan V, Kumar D, Naik J, et al. Nitrogen containing heterocycles as anticancer agents: a medicinal chemistry perspective. Pharmaceuticals. 2023;16:299.
24. Hu S, Chen J, Cao JX, Zhang SS, Gu SX, Chen FE. Quinolines and isoquinolines as HIV-1 inhibitors: chemical structures, action targets, and biological activities. Bioorg Chem. 2023;136:106549.
25. De S, Aamna B, Sahu R, Parida S, Behera SK, Dan AK. Seeking heterocyclic scaffolds as antivirals against dengue virus. Eur J Med Chem. 2022;240:114576.
26. Zhou Y, Mukherjee S, Huang D, Chakraborty M, Gu C, Zong G, et al. Development of novel IP6K inhibitors for the treatment of obesity and obesity-induced metabolic dysfunctions. J Med Chem. 2022;65:6869–87.
27. Vemula D, Jayasurya P, Sushmitha V, Kumar YN, Bhandari V. CADD, AI and ML in drug discovery: a comprehensive review. Eur J Pharm Sci. 2023;181:106324.
28. Sabe VT, Ntombela T, Jhamba LA, Maguire GEM, Govender T, Naicker T, et al. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review. Eur J Med Chem. 2021;224:113705.
29. Mousavi H, Zeynizadeh B, Rimaz M. Green and efficient one-pot three-component synthesis of novel drug-like furo[2,3-d]pyrimidines as potential active site inhibitors and putative allosteric hotspots modulators of both SARS-CoV-2 MPro and PLPro. Bioorg Chem. 2023;135:106390.
30. Cox PB, Gupta R. Contemporary computational applications and tools in drug discovery. ACS Med Chem Lett. 2022;13:1016–29.
31. Bassani D, Moro S. Past, present, and future perspectives on computer-aided drug design methodologies. Molecules. 2023;28:3906.
32. Mauri A. AlvaDesc: a tool to calculate and analyze molecular descriptors and fingerprints. In: Roy K, editor. Ecotoxicological QSARs. Methods in Pharmacology and Toxicology. New York: Humana; 2020. p. 801–20.
33. Sushko I, Novotarskyi S, Körner R, Pandey AK, Rupp M, Teetz W, et al. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. J Comput Aided Mol Des. 2011;25:533–54.
34. Halder AK, Delgado AHS, Cordeiro MNDS. First multi-target QSAR model for predicting the cytotoxicity of acrylic acid-based dental monomers. Dent Mater. 2022;38:333–46.
35. Ambure P, Aher RB, Gajewicz A, Puzyn T, Roy K. "NanoBRIDGES" software: open access tools to perform QSAR and nano-QSAR modeling. Chemom Intell Lab Syst. 2015;147:1–13.
36. Tetko IV, Tanchuk VY, Villa AEP. Prediction of n-octanol/water partition coefficients from PHYSPROP database using artificial neural networks and E-state indices. J Chem Inf Comput Sci. 2001;41:1407–21.
37. Golbraikh A, Tropsha A. Beware of q2! J Mol Graph Model. 2002;20:269–76.
38. Gramatica P. Principles of QSAR models validation: internal and external. QSAR Comb Sci. 2007;26:694–701.

39. Yoo W, Mayberry R, Bae S, Singh K, He QP, Lillard JW Jr. A study of effects of multicollinearity in the multivariable analysis. Int J Appl Sci Technol. 2014;4:9–19.
40. Ojha PK, Roy K. Comparative QSARs for antimalarial endochins: importance of descriptor-thinning and noise reduction prior to feature selection. Chemom Intell Lab Syst. 2011;109:146–61.
41. Gajewicz-Skretna A, Wyrzykowska E, Gromelski M. Quantitative multi-species toxicity modeling: does a multi-species, machine learning model provide better performance than a single-species model for the evaluation of acute aquatic toxicity by organic pollutants? Sci Total Environ. 2023;861:160590.
42. Kohlbacher SM, Langer T, Seidel T. QPHAR: quantitative pharmacophore activity relationship: method and validation. J Cheminform. 2021;13:57.
43. Mitra S, Halder AK, Ghosh N, Mandal SC, Cordeiro MNDS. Multi-model in silico characterization of 3-benzamidobenzoic acid derivatives as partial agonists of Farnesoid X receptor in the management of NAFLD. Comput Biol Med. 2023;157:106789.
44. Kohlbacher SM, Schmid M, Seidel T, Langer T. Applications of the novel quantitative pharmacophore activity relationship method QPhAR in virtual screening and lead-optimisation. Pharmaceuticals. 2022;15:1122.
45. Tosco P, Balle T, Shiri F. Open3DALIGN: an open-source software aimed at unsupervised ligand alignment. J Comput Aided Mol Des. 2011;25:777–83.
46. Tosco P, Balle T. Open3DQSAR: a new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields. J Mol Model. 2010;17:201–8.
47. Sar S, Mitra S, Panda P, Mandal SC, Ghosh N, Halder AK, et al. In silico modeling and structural analysis of soluble epoxide hydrolase inhibitors for enhanced therapeutic design. Molecules. 2023;28:6379.
48. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47:D506–D515.
49. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. 2018;27:293–315.
50. Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res. 2004;32:W615–W619.
51. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46:W296–W303.
52. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 2003;31:3381–5.
53. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. WIREs Comput Mol Sci. 2012;3:198–210.
54. Nurisso A, Daina A, Walker RC. A practical introduction to molecular dynamics simulations: applications to homology modeling. In: Orry A, Abagyan R, editors. Homology Modeling. New York: Humana Press; 2012. p. 137–73.
55. Liu Y, Grimm M, Dai WT, Hou MC, Xiao ZX, Cao Y. CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. Acta Pharmacol Sin. 2020;41:138–44.
56. Liu Y, Yang X, Gan J, Chen S, Xiao ZX, Cao Y. CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. Nucleic Acids Res. 2022;50:W159–W164.
57. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455–61.
58. Ghosh A, Panda P, Halder AK, Cordeiro MNDS. In silico characterization of aryl benzoyl hydrazide derivatives as potential inhibitors of RdRp enzyme of H5N1 influenza virus. Front Pharmacol. 2022;13:1004255.
59. Halder AK, Cordeiro MNDS. Multi-target in silico prediction of inhibitors for mitogen-activated protein kinase-interacting kinases. Biomolecules. 2021;11:1670.

60. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res. 2007;35:W522–W525.
61. Roe DR, Cheatham TE III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput. 2013;9:3084–95.
62. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate—DNA helices. J Am Chem Soc. 1998;120:9401–9.
63. Yang H, Lou C, Sun L, Li J, Cai Y, Wang Z, Li W, Liu G, et al. admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. Bioinformatics. 2019;35:1067–9.
64. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717.
65. Reutlinger M, Koch CP, Reker D, Todoroff N, Schneider P, Rodrigues T, et al. Chemically advanced template search (CATS) for scaffold-hopping and prospective target prediction for 'orphan' molecules. Mol Inform. 2013;32:133–8.