European **J**ournal of **Chem**istry

Check for updates

QSAR study of benzofuran and indole derivatives to predict new compounds as histone lysine methyl transferase inhibitors

Kaushik Sarkar [ID], Sraboni Ghosh [ID] and Rajesh Kumar Das [ID] *

Department of Chemistry, University of North Bengal, Darjeeling, 734013, India

* Corresponding author at: Department of Chemistry, University of North Bengal, Darjeeling, 734013, India.
 e-mail: rajeshnbu@gmail.com (R.K. Das).

RESEARCH ARTICLE

ABSTRACT

Initiation and progression of several diseases by post-translational histone modifications are considered a worldwide problem. Enhancer of Zeste Homologue 2 (EZH2), which belongs to the histone-lysine *N*-methyl transferase (HKMT) family, has been emphasised as a promising target for cancer therapy. It is a major challenge for the scientific community to find novel approaches to treating this disease. In this study, a series of 51 derivatives of the benzofuran and indole families, previously experimentally evaluated against HKMT, was used to develop the best model with promising anticancer activity. The multiple linear regression (MLR) method, implemented in QSARINS software, was used with a genetic algorithm for variable selection. According to QSARINS, the model with two descriptors (minHBint4 and Wlambdal.unity) was found to be the best and its parameters fit well, and its validation was well established. The applicability domain was also validated for this model. Furthermore, its robustness (R^2 = 0.9328), stability (Q^2_{LOO} = 0.9212, Q^2_{LMO} = 0.9187), and good predictive power (R^2_{ext} = 0.929) were also verified. Hence, this model was assumed to have predictive HKMT anticancer activity for designing active compounds. Molecular docking was also performed to identify binding interactions, and new molecules with better predicted biological activity (pIC_{50}) were designed. The binding energy of the three designed compounds demonstrated higher binding activity at the target receptor, followed by complex stability, determined by a 100 ns molecular dynamics simulation and binding free energy calculation. Density functional theory (DFT) and pharmacokinetic analyses also confirmed their drug-like properties. Finally, it can be declared that the proposed tools allow rapid and economical identification of potential anti-HKMT drugs (anticancer drugs) for further development.

1. Introduction

Several studies recommend that abnormal post-translational histone modifications have a significant character in the initiation and progression of numerous diseases [1]. Post-translational modifications (PTMs) mention chemical adjustments of a protein after translation, which controls gene expression by varying the chromatin conformation and transcription factor activity [2,3]. Small basic proteins such as histones are found in the nucleus of eukaryotic cells that are tightly bound to the DNA of eukaryotic chromosomes. The positively charged N-terminal tail of their subunits is vulnerable to covalent modifications [4]. Covalent histone modifications, such as phosphorylation, ubiquitination, acetylation, and methylation, are important for controlling chromatin dynamics and function [5]. Methylation of specific lysine residues in the N-terminal tail of histones is essential to regulate gene expression and transduction of cellular signalling [6-8]. By transferring the methyl group of S-adenosyl methionine (SAM) to equivalent specific lysine, histone lysine methyl transferases (HKMTs) behave as the 'drivers' of histone methylation, which regulate biological processes [9,10]. Lysine methyl transferases (KMT) regulate the monomethylation, dimethylation and trimethylation of N-

lysine in histone and non-histone substrates [11]. Enhancer of Zeste Homologue 2 (EZH2) belongs to the HKMTs family, which is responsible for the methylation of lysine-9 and lysine-27 of histone H3, via methyl group transfer from SAM 5 cofactor, which leads transcriptional repression of the affected genes. Trimethylation of lysine-27 suppresses many specific genes, including tumour suppressor genes [12]. Thus, these proteins have been associated with several diseases such as cancer and neurodegeneration [11]. In numerous types of cancer, it has been revealed that high EZH2 expression is associated with a poor prognosis, a high grade, and a high stage [13]. In tumours, increased EZH2 activity has been associated with suppression of differentiation. Overexpressed of EZH2 such as breast, bladder, endometrial, liver, ovarian, prostate, small cell lung cancer (SCLC), melanoma, glioblastoma, and pediatric glioma, as well as lymphomas, are correlated with progression and poor prognosis. Increasing EZH2 activity in SCLC, associated with rapid inactivation of the retinoblastoma (RB1) tumour suppression gene, regulates tumour growth and standard of care chemotherapy resistance. Somatic activating mutations in EZH2 have been recognised in diffuse large B cell lymphoma (DLBCL) and follicular lymphoma (FL), resulting in higher H3K27me3. It implies that EZH2 overexpression silences target
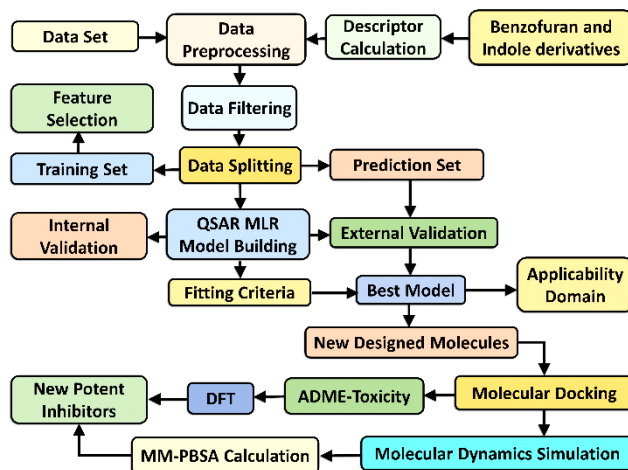
Figure 1. Flowchart of this work.

genes important in tumour growth and survival due to the dysregulation of H3K27me3 [1]. This has led to the identi-fication of EZH2 as a possible target for cancer therapy. Thus, many pharmaceutical companies and academic institutes have been involved in small molecule EZH2 inhibitors in the treatment of cancer, and several EZH2 inhibitors have been reported [14,15]. Unfortunately, there are only a few scaffolds among the known EZH2 inhibitors.

Therefore, we made the decision to find novel classes of EZH2 inhibitors that have the same scaffolds as existing EZH2 inhibitors. In this study, a quantitative structure-activity relationship (QSAR) was established on a set of benzofuran and indole derivatives. Benzofuran and indole moieties have been widely used in many drugs due to their extensive anticancer activity [16]. Nowadays, drug resistance is a universal problem of morbidity and mortality in cancer, causing great concern in cancer therapy [17,18]. Hence, novel anticancer agents are urgently needed to overcome intrinsic or developed resistance [19]. Benzofuran and indole are considered outstanding heterocyclic compounds due to their diverse biological activities [20]. The various pharmacological properties of indole include antibacterial, anticonvulsant, antifungal, anti-inflammatory, antimalarial, antitubercular, antiviral, and anti-cancer activities. The anticancer activity of these derivatives includes mechanisms such as apoptosis (myeloid cell leukemia-1 inhibitors), signal transduction (proviral insertion site in Moloney murine leukemia virus inhibitors), replication and transcription (DNA topoisomerase), epigenetic modifications (histone deacetylase inhibitors, histone acetyltransferase inhibitors, and silent mating type information regulation 2 homolog inhibitors), and cell mitosis (tubulin inhibitors) [19]. However, benzofuran derivatives have antiviral, immunosupp-ressive, antioxidant, antifungal, anti-inflammatory, analgesic, antimicrobial, and antitumor activities [20]. After our *in silico* approach was successful, we found three designed molecules: H53av, H53bb, and H53y, which bind strongly to the active site of the receptor without any significant change in the protein during the entire MD simulation. Therefore, these molecules could be exploited as novel, potent, and selective inhibitors for EZH2 and further subjected to *in vitro* and *in vivo* studies. The flow of the work of this study is shown in Figure 1.

2. Materials and methods

2.1. Database of inhibitor compounds

The database used consists of 51 molecules derived from benzofuran and indole, extracted from the bindingdb database

(https://www.bindingdb.org/bind/index.jsp), which were used to develop a model with inhibitory activity of HKMT EZH2. The chemical structures and experimentally reported biological activity data of the used molecules are shown in Tables 1-3. According to the database, the activity data was taken using the values of pIC_{50} ($9\text{-}\log_{10}IC_{50}$) as a dependent variable. The IC_{50} denotes the molar concentration of the drug required to reach 50% protection against the HKMT target.

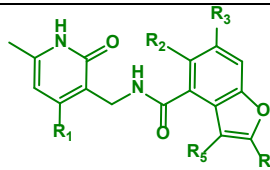### 2.2. Molecular descriptors

The PaDEL 2.18 software [21] was used to calculate the molecular descriptors. This software is an application for calculating molecular descriptors, originally developed by Chun Wei Yap, the Laboratory of Pharmaceutical Data Exploration of the National University of Singapore, and provides more than 1875 molecular descriptors that are divided into logical groups. All 1D-3D descriptor families were included in this study by removing constant or near-constant variables. To properly use 3D descriptors, it was essential to optimise all of the molecules previously, to minimise the energy of the conformation. In this sense, we use the Merck molecular force field (MMFF94) force field in the Avogadro suite [22].

2.3. QSAR modelling

Based on the experience of the Insubria QSAR Research Unit, a QSAR multiple linear regression model (MLR) was established using the QSARINS software [23]. This software permits the development of QSAR-MLR models by the ordinary least squares (OLS) method according to Equation (1),

$$Y_i = b_0 \sum_{j=1}^{n} b_j X_{ij} + e_i \qquad (1)$$

In this mathematical equation, a linear relationship is shown between the studied response (Y_i) and the selected values of the descriptors (X_{ij}), where e_i is the random error. Therefore, the intersection (b_0) and the coefficients (b_j) must be calculated. The chemometric approach was used to validate the obtained models. In order to choose the best model, they were analysed by the QSARINS according to various parameters: fit (higher R^2), robustness (higher Q^2_{LOO}), stability (lower R^2-Q^2_{LOO}), correlation of descriptors (low K_{XX}), correlation with response (high dK), and the root of the mean error squared over (a) training calculation ($RMSE_{tr}$), (b) training prediction by leaving one out ($RMSE_{CV}$), and (c) external prediction set ($RMSE_{ext}$).
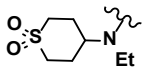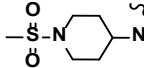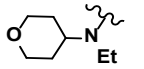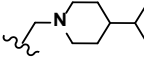
Table 1. List of chemical compounds with their IC$_{50}$ values.
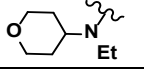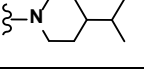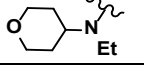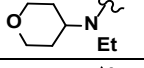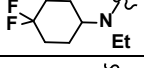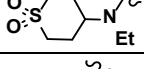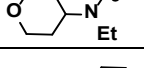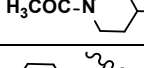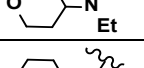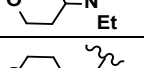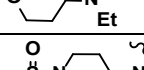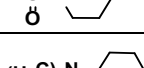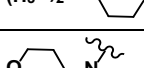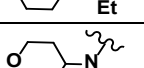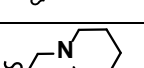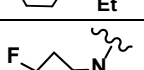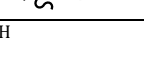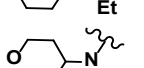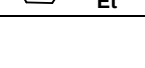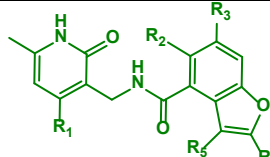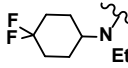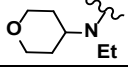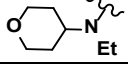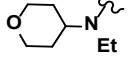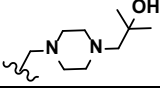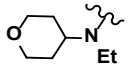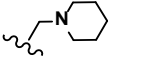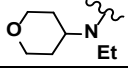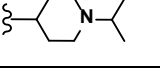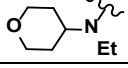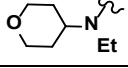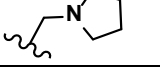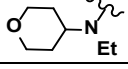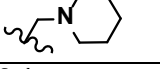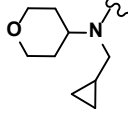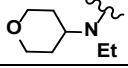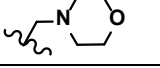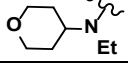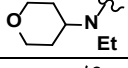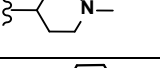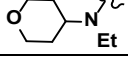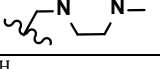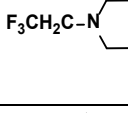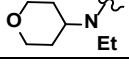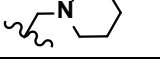
Compound	R$_1$	R$_2$	R$_3$	R$_4$	R$_5$	IC$_{50}$ (nM)
H2	Methoxy	Ethyl		H	H	0.8
H4	Methoxy	Methyl		H	H	0.9
H5	Methoxy	Ethyl			H	1.2
H6	Methoxy	Ethyl			H	1.7
H7	Methoxy	Methyl		H	H	2.3
H8	Methoxy	Methyl		Trifluoromethyl	H	2.7
H9	Methoxy	Ethyl		H	H	2.7
H10	Methoxy	Ethyl		H	H	2.7
H11	Methoxy	Methyl		Methyl	H	2.8
H12	Methoxy	Methyl		H	H	2.8
H13	Methoxy	Methyl		H	H	2.9
H14	Methoxy	Ethyl		-CH$_2$N(CH$_3$)$_2$	H	2.9
H15	Methoxy	Methyl		Methyl	H	2.9
H16	Methoxy	Methyl		H	H	3.2
H17	Methoxy	Methyl		H	H	3.4
H18	Methoxy	Ethyl		F	H	3.5
H19	Methoxy	Ethyl			H	3.5
H20	Methyl	Ethyl			H	4
H21	Methoxy	Ethyl		H	H	4
H22	Methyl	Ethyl			H	4
H23	Methoxy	Methyl		H	H	4

Table 1. Continued.
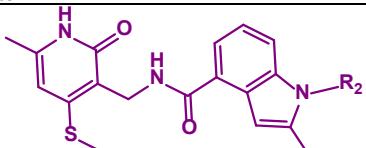
Compound	R_1	R_2	R_3	R_4	R_5	IC_{50} (nM)
H24	Methoxy	Ethyl	(difluorocyclohexyl N-Et)	H	H	4.5
H25	Methoxy	Methyl	(oxan-4-yl N-Et)	H	H	4.8
H26	Methoxy	Methyl	(oxan-4-yl N-Et)	H	-CH₃	5.6
H27	Methyl	Ethyl	(oxan-4-yl N-Et)	(piperazine with hydroxy)	H	5.6
H28	Methyl	Ethyl	(oxan-4-yl N-Et)	(piperidine)	H	5.7
H29	Methyl	Ethyl	(oxan-4-yl N-Et)	(isopropyl piperidine)	H	5.7
H30	Methoxy	Methyl	(oxan-4-yl N-Et)	H	-CH₃	5.9
H31	Methoxy	Ethyl	(oxan-4-yl N-Et)	(pyrrolidine)	H	6.2
H32	Methoxy	Ethyl	(oxan-4-yl N-Et)	(piperidine)	H	6.4
H34	Methoxy	Ethyl	(oxan-4-yl N with cyclopropylmethyl)	Cyclopropane	H	6.8
H35	Methoxy	Ethyl	(oxan-4-yl N-Et)	(morpholine)	H	7.8
H36	Methyl	Ethyl	(oxan-4-yl N-Et)	Cyclopentane	H	9.6
H37	Methoxy	Ethyl	(oxan-4-yl N-Et)	(N-methyl piperidine)	H	9.9
H38	Methoxy	Ethyl	(oxan-4-yl N-Et)	(N-methyl piperazine)	H	10
H39	Methoxy	Ethyl	(F₃CH₂C-N piperidine, N with cyclopropylmethyl)	H	H	14
H40	Methoxy	Methyl	(oxan-4-yl N-Et)	(piperidine)	H	14

Therefore, the selected model must be the most stable, predictive, and generalisable because of the least difference between the adjustment, cross-validation, and external validation. Consequently, it must have RMSE values that are as close as possible [24].

2.4. Model validation by OECD principles

It should be remembered that QSAR modelling is not a trivial approach, and the outcomes of correlation are not as easy as 'pressing a button'. In recent years, the QSAR community has given specific consideration to the validation of these models, with an emphasis on the applicability domain (AD) and predictive power to increase assurance in the consistency of the data predicted by them. According to existing regulatory principles, five fundamental characteristics (OECD principles) must be satisfied for the suggested model to be predictive [24].

Table 2. List of chemical compounds with their IC$_{50}$ values.
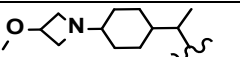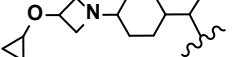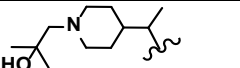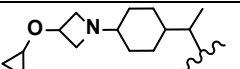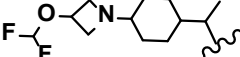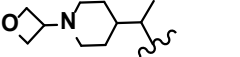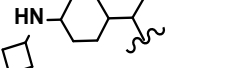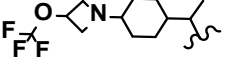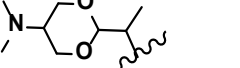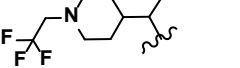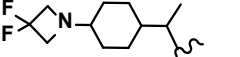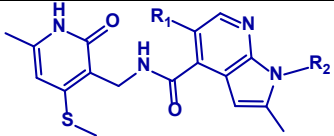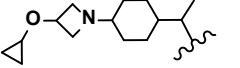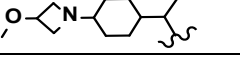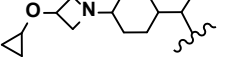
Compound	R$_2$	IC$_{50}$ (nM)
H45		0.059
H47		0.063
H48		0.067
H49		0.072
H51		0.100
H52		0.120
H53		0.130
H54		0.140
H55		0.160
H56		0.180
H57		0.180

Table 3. List of chemical compounds with their IC$_{50}$ values.

Compound	R$_1$	R$_2$	IC$_{50}$ (nM)
H44	H		0.057
H46	F		0.060
H50	Cl		0.092

Principle I is related to the definition of the measurement point in which it measures the physicochemical, biological, or pharmacological properties of the developed model. The goal of this approach is to provide transparency at the measurement point anticipated by a particular model [24]. In this study, the measurement point is the molar concentration of the inhibitor required to achieve 50% protection against the cancer effect of HKMT EZH2 (IC$_{50}$), expressed as pIC$_{50}$.

In Principle II, the QSAR model is expressed in the form of definite algorithms, taking into consideration how the descriptors of the chemical structure are related to biological activity [24]. Here, the algorithm used in this QSAR modelling is a mathematical model of multiple linear regression (MLR).

In accordance with Principle III, the applicability domain (AD) of a QSAR model is mentioned. It is difficult to assume accurate predictions of the modelled property for the whole universe of chemical compounds, even from a strong, significant, and validated QSAR model. In fact, extrapolations of the models are not considered valid; only predictions for compounds that fall inside the application domain are considered [24].
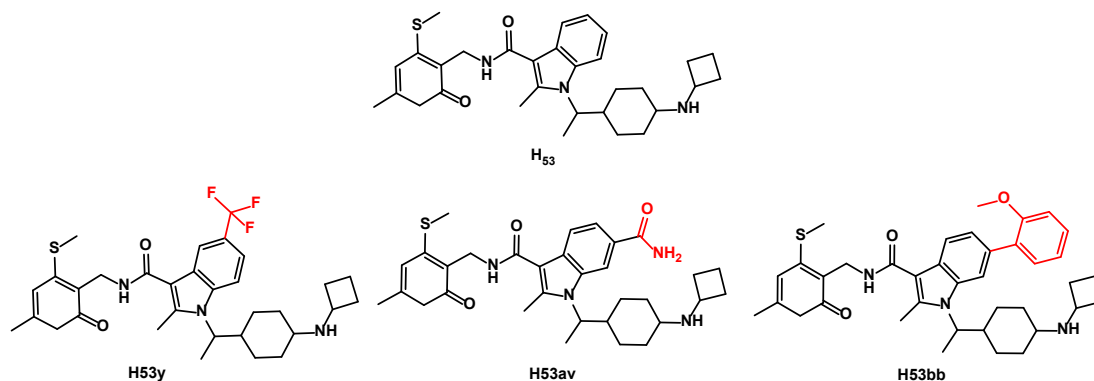
Figure 2. Compounds used for virtual screening derived from compound H53 of the database.

In this study, two approaches provided by QSARINS were used to explain the AD of the model. First, the Williams plot was used to ensure that compounds outside the range can be considered outliers, while the accurate prediction of the model has leverage values below the critical leverage of ±3 standard deviations. Second, the Insubria plot was used to estimate the reliability of the predictions of compounds that exhibit no experimental response and to compare them with the predictions from the database.

On the other hand, appropriate measurement of goodness-of-fit, robustness, and predictivity of the model are discussed in Principle IV [24]. In this regard, the coefficient of determination R^2 is used to assess the fit, and the effectiveness of the model is evaluated using the modified form of R^2_{adj}.

In this study, the cross-validation strategies implemented in QSARINS were also used to check the robustness of the model. LOO (leaving one out), which is known as a widely used technique in QSAR, was also discussed here, in which a compound is repeatedly excluded from the data set and the model with the remaining compounds is calculated. However, a single compound disturbance is too small to establish the real robustness of the model in large databases. On the other hand, LMO (leaving many out), implemented in QSARINS, is also used to study the behaviour of the model when a large number of compounds are excluded [24]. The Y-randomization procedure is applied to demonstrate that the model is not the result of a casual correlation.

When the model is internally validated and the probability of casual correlation is excluded, an external validation can be executed, which predicts new compounds. The procedure is performed by applying the model equation, obtained from the training set compounds, to a prediction set of compounds, which are excluded and will never be used in the calculation of the model. Performance can be measured using a variety of factors, including $RMSE_{ext}$, Q^2_{F1}, Q^2_{F2}, Q^2_{F3}, and CCC. Although Principle V is optional, we will come to an understanding based on the relationship established between the inhibitory activity of benzofuran and indole derivatives against HKMT EZH2, and the molecular descriptors included in the model.

2.5. Molecular docking study

The molecular docking procedure is broadly used for calculating the binding affinities for a number of ligand molecules. Transitional steps, such as PDBQT files for protein and ligand preparation and grid box creation, were accomplished using the graphical user interface of Autodock Tools (ADT) [25]. It allocates polar hydrogens, aggregate atom Kollman charges, solvation parameters, and fragmental volume to the protein structure. The three-dimensional X-ray crystal structure of the human enhancer of zeste homolog 2 (EZH2)

protein, deposited in the RCSB Protein Data Bank (https://www.rcsb.org/), with PDB ID: 4MI0, was downloaded. The cofactors and water molecules were excluded from the protein moiety. Autodock saved the organized file in PDBQT format. Auto Grid was used for preparing the grid map using a grid box. The size of the grid box was set to 60×60×60 xyz points with a spacing of 0.375 Å and the centre of the grid box was selected at dimensions ($x, y,$ and z): 13.862, 49.626 and 20.597. In order to minimize the computation time, a scoring grid is calculated from the ligand structure. Autodock Vina [26] was used to perform the docking using protein and ligand information along with grid box properties. During docking, both the protein and ligands were considered rigid. The pose with the highest negative energy of binding affinity was selected and aligned with the protein structure for further analysis.

2.6. Virtual screening

The validation movements designated for the proposed model have proven their effectiveness for the application and are therefore used in the identification or selection of new compounds as HKMT EZH2 inhibitors. However, it is not suitable to screen structurally different compounds based on the characteristics of the database (all compounds have a common scaffold). Because it is very likely that they were outside the domain of application of the model, they would not be considered reliable predictions. In response, the three best-designed compounds derived from H53 in this database were screened using the model formed by the described techniques. The chemical structure of these compounds is shown in Figure 2.

2.7. Molecular dynamics (MD) simulation

MD simulation was performed on the H53y-associated complex from the molecular docking trajectory with GROMACS 5.1 software. The force fields utilised to simulate the protein complex were the CHARMM36 force field [27] and the TIP3P water model [28]. Hydrogens were ignored during the production of the protein topology file because they will be added later. Periodic boundary conditions were applied to the simulation at 310 K and 1 atm pressure. The Avogadro software [22] was used to add hydrogen atoms to the docked H53y molecule. All topology information for the ligand was done with the CHARMM General Force Field (CGenFF) server [29,30].

Now, the equilibration of the system was carried out in two phases: NVT ensemble (constant number of particles, volume, and temperature), followed by the NPT ensemble (constant number of particles, pressure, and temperature). Here, 1000 ps NVT equilibration was accompanied by positional restraint to

the protein and ligand. The algorithm is based on a leap-frog integrator with a modified Berendsen thermostat temperature coupling with two FS time steps. Furthermore, the NPT equilibration was executed using the same protocol. Finally, a 100 ns MD simulation run was executed for data collection and analysis.

2.8. MM-PBSA (molecular mechanic-Poisson Boltzmann surface area) calculation

The binding free energy for the protein-H53y complex was calculated from simulated MD trajectories using the g_mmpbsa script of the MM-PBSA method [31,32]. The free binding energy ($\Delta G_{binding}$) of the protein-ligand complex is expressed by Equation (2):

$$\Delta G_{binding} = G_{complex} - (G_{protein} + G_{ligand}) \tag{2}$$

where $G_{complex}$ is the total free energy of the protein-ligand complex, and $G_{protein}$ and G_{ligand} are the total free energies of the protein and ligand in the solvent, respectively. The combined effects of solvent-accessible surface area (SASA), polar solvation, van der Waals, and electrostatic energies represent the final binding free energy of the complex.

2.9. In silico drug-likeness and toxicity prediction

In this study, the designed compounds were assessed using the PreADMET web server (https://preadmet.webservice.bmdrc.org/) for testing drug-likeness and bioactivity scores. Furthermore, the toxicological evaluation was performed using the OSIRIS [33] property explorer, which was checked for maximum toxicity.

2.10. DFT study

Quantum chemical calculations of the designed molecules with better inhibitory activity and binding affinity were performed using Gaussian 16W [34] and GaussView 6.0 [35], a molecular visualisation programme. The chemical structure of designed compounds was optimized by density functional theory (DFT) using restricted Becke's three parameter hybrid functional (B3LYP) with the latest 6-311++G(d,p) basis set [36]. The DFT studies were completed to identify the energies of the highest occupied molecular orbital (E_{HOMO}), lowest unoccupied molecular orbital (E_{LUMO}), and the band gap (ΔE_{HL}), which correlate with biological activity. Using E_{HOMO} and E_{LUMO}, other important global reactivity descriptor parameters such as the ionization potential (I), electron affinity (A), chemical potential (μ), global hardness (η), global softness (S), electronegativity (σ) and electrophilicity index (ω) were calculated:

$$I = -E_{HOMO} \tag{3}$$

$$A = -E_{LUMO} \tag{4}$$

$$\mu = \frac{1}{2}(E_{LUMO} + E_{HOMO}) \tag{5}$$

$$\eta = \frac{1}{2}(E_{LUMO} - E_{HOMO}) \tag{6}$$

$$S = \frac{1}{\eta} \tag{7}$$

$$\sigma = -\mu \tag{8}$$

$$\omega = \frac{\mu^2}{2\eta} \tag{9}$$

3. Results and discussion

3.1. QSAR-MLR model

After selection of the descriptors, thousands of different models were calculated during a typical QSAR session. Models with inadequate quality were excluded to avoid a useless list of final outputs. In this study, a model with lower correlation (Table S1) between the descriptors was selected, in which the model has low multicolinearity and a good correlation with the modelled response.

The best MLR model obtained with its statistical parameters is

$$pIC_{50} = 10.2512 - 1.2283(minHBint4) - 0.3777(Wlambdal.unity) \tag{10}$$

Fitting criteria: R^2: 0.9328, R^2_{adj}: 0.930, R^2-R^2_{adj}: 0.0034, LOF: 0.0467, Kxx: 0.4597, Delta K: 0.0903, RMSE$_{tr}$: 0.196, MAE$_{tr}$: 0.1557, RSS$_{tr}$: 1.6521, CCC$_{tr}$: 0.9652, s: 0.2032, F: 277.734

Internal validation criteria: Q^2_{LOO}: 0.9212, R^2-Q^2_{LOO}: 0.0117, RMSE$_{cv}$: 0.2124, MAE$_{cv}$: 0.1683, PRESS$_{cv}$: 1.939, CCC$_{cv}$: 0.9592, Q^2_{LMO}: 0.9187, R^2_{Yscr}: 0.0471, Q^2_{Yscr}: -0.1027

External validation criteria: RMSE$_{ext}$: 0.2191, MAE$_{ex}$: 0.1978, PRESS$_{ext}$: 0.384, R^2_{ext}: 0.929, Q^2_{F1}: 0.9302, Q^2_{F2}: 0.9249, Q^2_{F3}: 0.9161, CCC$_{ext}$: 0.9612.

where R^2, R^2_{adj} and LOF are the coefficient of determination, adjusted coefficient of determination, and Friedman's lack of adjustment, respectively. Furthermore, Kxx is the global correlation between the descriptors and Delta K is the difference in the correlation between the descriptors (K$_X$) and the response (K$_{XY}$). RMSE$_{tr}$ is the root of the mean error squared over the training calculation, MAE$_{tr}$ is the absolute error (average) in the adjustment calculated in the training series, RSS$_{tr}$ is the sum of the squared residuals in the adjustment, s is the standard estimation error, CCC$_{tr}$ is the correlation coefficient of concordance in the training and F is the Fisher value.

The best model has an R^2 value of 0.9328; therefore, it exhibits a good fit to the model with the inhibition of the HKMT EZH2 target. Additionally, it has an R^2_{adj} value of 0.930, which indicates the ease of adding new descriptors to the model. As a result, with a low LOF parameter of 0.0467, it can be concluded that the model is not overfit because it has a good fit with a minimum of two descriptors.

Since the K$_{XX}$ is low (0.4597) and the correlation between the model's descriptors is small, we may assume that the descriptors we have chosen do not include any redundant information. Furthermore, according to the Delta K parameter (0.0903), with a small error on training calculations and parameter estimation (RMSE$_{tr}$ = 0.196; MAE$_{tr}$ = 0.1557; s = 0.2032), the correlation between the descriptors and the modelled response is satisfactory. The plot of predicted *versus* experimental responses (Figure 3) enables easy detection of systematic trends or grouping of data with probable outliers.

This graph compares the experimental values of pIC50 with the values predicted by the model equation for the training and prediction series. It is seen that most of the points are close to the line, and no unusual behaviour of the other compounds in the database has been found. The experimental pIC$_{50}$ for HKMT EZH2 and the results predicted by the MLR model for the training and prediction series are shown in Table S2.

3.2. Analysis of the model with OECD regulatory principles

Although the first stage in a QSAR analysis is to calculate the model, which is not sufficient to guarantee its validity.
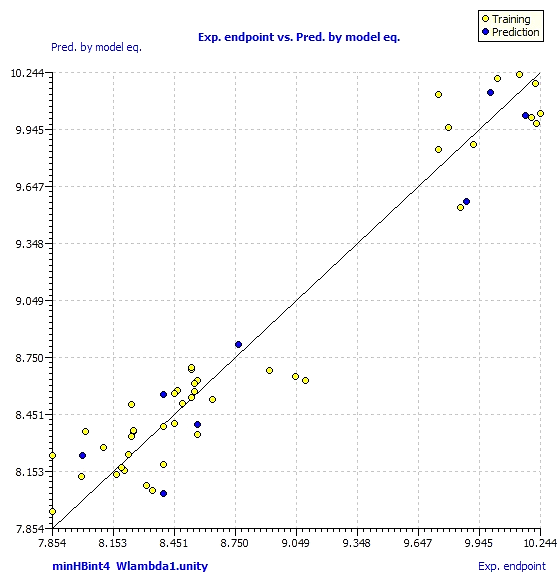
Figure 3. Values predicted by the model against observed values. The light circles represent the training set compounds and the dark circles represent the prediction set.
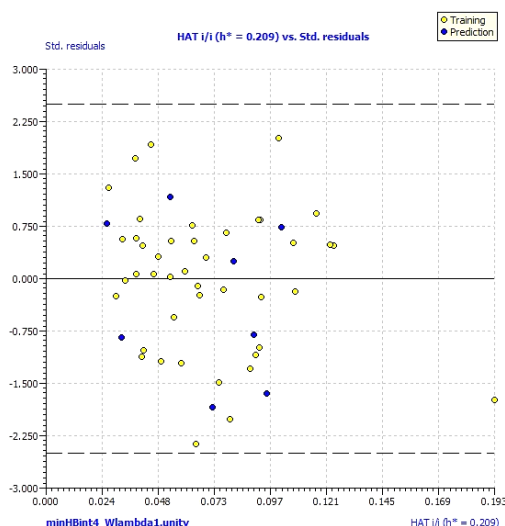
Figure 4. Williams graph for the applicability domain of the model.

The model must be thoroughly validated using the chemometric technique, paying close attention to conformity with the regulatory guidelines outlined by the OECD. In this regard, the section titled "Model validation by OECD principles" clarifies Principles I and II [24].

In this work, the leverage (h) and standardised residuals techniques were used to test AD (Principle III). The Williams graph of the model for the training and prediction series is shown in Figure 4. All compounds are included in the AD of the model with critical leverage, as seen in this figure ($h^* = 0.209$).

Furthermore, QSARINS software offers a new method for determining AD based on leverage and model predictions. On the other hand, the Insubria graph is very valuable to assess the position of a molecule.

The Insubria graph (Figure S1) shows that the results resemble the Williams plot, so it must be remembered that these compounds might be considered influential compounds.

Several criteria were taken into account to evaluate Principle IV. In the previous section, the fit of the model was discussed. The results obtained from the internal validation show the robustness and stability of the model. According to the variance explained in the prediction by LOO ($Q^2_{LOO} = 0.9212$), it confirms good internal predictions. The Q^2_{LOO} parameter (0.9212) has a high value similar to $R^2 = 0.9328$, and therefore the model is considered internally stable or robust, with small errors in the predictions of $RMSE_{cv}$ (0.2124) and MAE_{cv} (0.1683).

Figure S2 shows the predicted values by LOO *versus* the experimental values of pIC_{50} for the training and prediction sets. Because the compounds are close to the straight line, as predicted by the model, this provides the hypothesis that the model is stable and robust.

LMO, where 30% of the compounds of the training set are removed, is a more powerful technique used in QSARINS. The model is considered robust and stable because the values of R^2 (0.9328), Q^2_{LOO} (0.9212), and Q^2_{LMO} (0.9187) are comparable. Figure 5 shows Q^2_{LMO} *versus* K_{XY} (correlation between descriptors and inhibition of HKMT EZH2). It is observed from the figure that the Q^2_{LMO} values (red circles) are very similar to each other and comparable with the K_{XY} values, which agrees that the model is stable and robust.
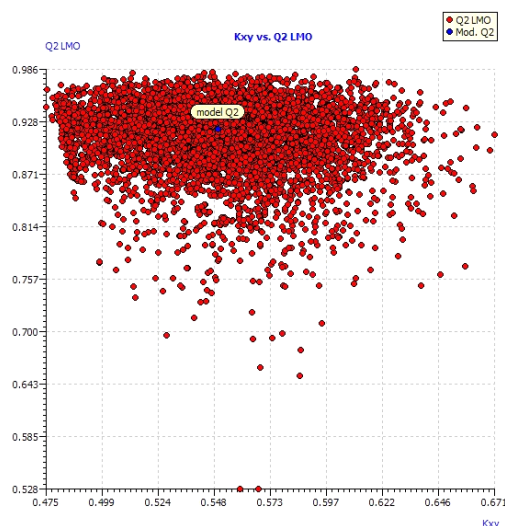
Figure 5. LMO values in cross-validation based on the correlation of the descriptors with the modelled response.
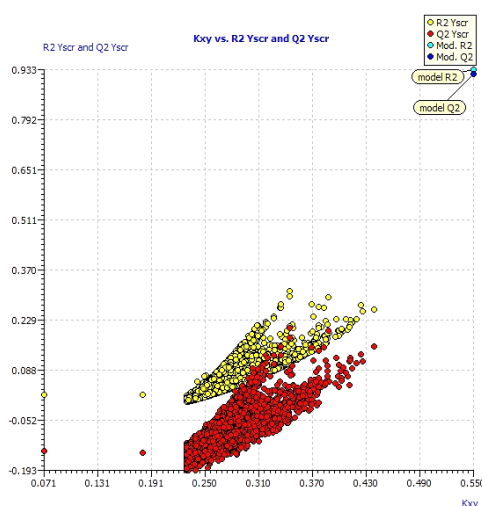
Figure 6. Y-scrambling graph in internal validation.

Additionally, the Y-scrambling graph states that the model is not obtained by casual correlation because the performance of the model decreases significantly with the placement of random answers (R^2_{Yscr} = 0.0471 and Q^2_{Yscr} = −0.1027). Figure 6 displays the values of R^2_{Yscr} and Q^2_{Yscr} *vs* R^2 and Q^2 of the model. Here, the last two parameters are very far from the values of the Yscr, which specifies that the model is not obtained by casual correlation.

However, external validation parameters validate their capacity to predict new compounds. The obtained parameters, such as R^2_{ext} = 0.929, provide accurate results for the R^2 value of the model, and RMSE$_{ext}$ = 0.2191, MAE$_{ext}$ = 0.1978, PRESS$_{ext}$ = 0.384, Q^2_{F1} = 0.9302, Q^2_{F2} = 0.9249, Q^2_{F3} = 0.9161, CCC$_{ext}$ = 0.9612, r2m aver = 0.8952, and r2m delta = 0.0558 are also good. On the basis of these consequences, the model has good predictive power and is suitable for the prediction of novel compounds that have not undergone experimental evaluation. Table S2 displays the predictions of the MLR model for the prediction set.

According to the descriptors present in the model, Principle V is taken into consideration. With just two predictive variables, the developed model can explain the variance of the experimental data to inhibit HKMT EZH2. In the model equation, these variables negatively influence the value of the

pIC$_{50}$ response, so the value of pIC$_{50}$ should increase with a decrease in the value of these molecular descriptors. The first descriptor consists of the two-dimensional descriptor minHBint4 [37], which is an atom-type electrotopological state that mainly encodes the minimum E-state descriptor of strength for potential hydrogen bonds of path length 4 in the molecule. However, the second descriptor Wlambdal.unity is a three-dimensional Weighted Holistic Invariant Molecular (WHIM) that encodes the atomic environment, weighted by unit weights.

Additionally, molecular docking studies were performed with these 51 inhibitors to validate the virtual screening structure-activity relationship. Here, the high values of the biological property (pIC$_{50}$) would correspond to a decrease in the values of the variables that negatively contribute to the model, which are minHBint4 and Wlambdal.unity. From the original dataset, few molecules were found to have better predicted pIC$_{50}$ values (>10 nM). A representative molecule (H53) from the database has the highest negative binding energy of −8.8 kcal/mol with a considerable predicted pIC$_{50}$ (>10 nM). Therefore, new molecules with an extremely predicted pIC$_{50}$ (>10 nM) were designed by various substitutions at the H53 molecule, followed by a better binding affinity (Table S3).
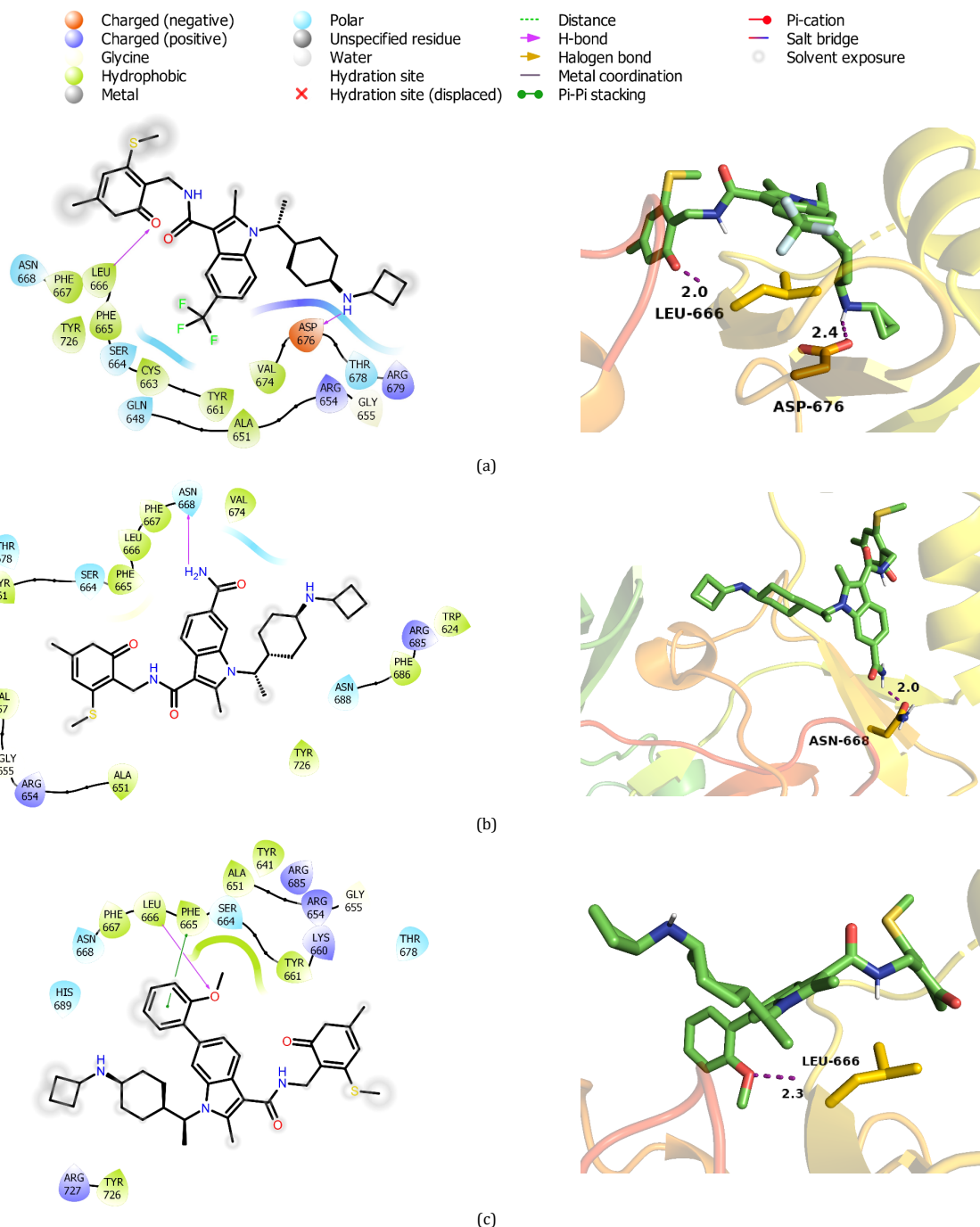
Figure 7. Three-dimensional and two-dimensional molecular docking interactions of three designed compounds (a) H53y, (b) H53av, and (c) H53bb with the target receptor of human EZH2 (PDB ID: 4MI0).

3.3. Molecular docking studies

The molecular docking was performed in order to study possible interactions between the receptor and the ligand molecules [38]. In the present manuscript, we attempt to study the interactions of new benzofuran and indole hybrids with EZH2. In order to study the binding efficacy of all the designed compounds, molecular docking studies were performed in the binding pockets of human EZH2 [PDB ID: 4MI0]. Compound H53 from the data set is taken as a reference standard due to its highest binding energy (−8.8 kcal/mol). On the other hand, all the designed compounds with considerable biological activity

(pIC$_{50}$) exhibited better binding scores, ranging from −6.9 to −9.3 kcal/mol.

The binding energy of the two designed compounds (H53p and H53ck) has shown the same binding energy as that of the target. Only three molecules, such as H53y, H53av, and H53bb, emerged as the best derivatives because of their higher binding energies compared to reference. Docking interactions of the best active compounds are shown in Table S4, while Figure 7 represents the binding interactions of the best-designed compounds on the human EZH2 protein. Among them, H53y has shown the highest binding energy of −9.2 kcal/mol, having two hydrogen bonds with the residues Asp676 and Leu666 of the receptor.
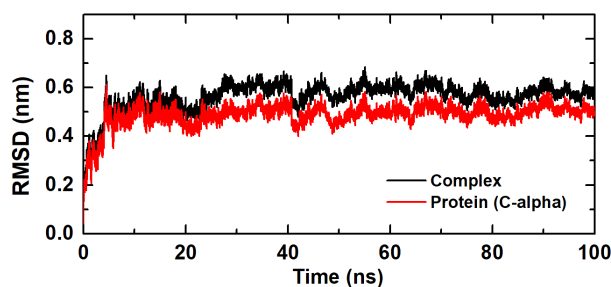
Figure 8. RMSD of protein-ligand complex during 100 ns MD simulation.
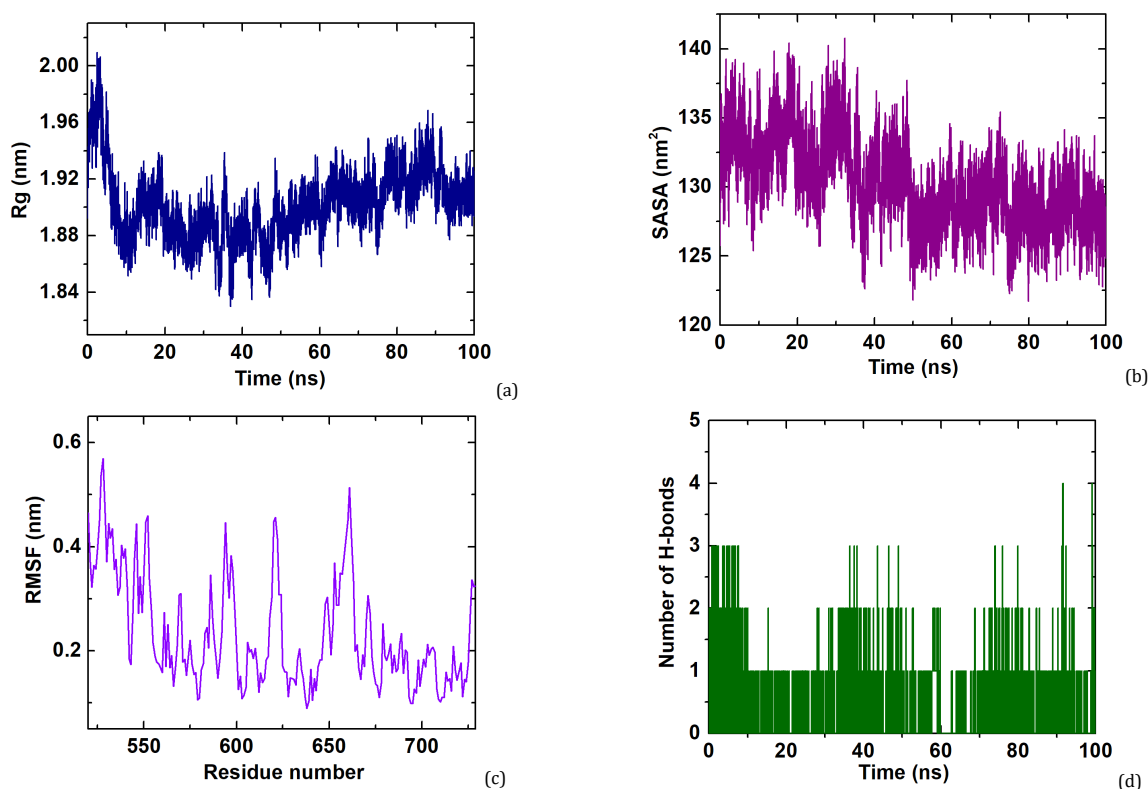
Figure 9. (a) Rg, (b) SASA, (c) RMSF, and (d) H-bond plot of protein-ligand complex during 100 ns MD simulation.

The remaining two compounds (H53av and H53bb) have also shown the same but higher binding energy (−8.9 kcal/mol) compared to the reference. In this case, one hydrogen bond is formed by both molecules.

3.4. MD simulation analysis

Molecular docking is unable to distinguish the conformational changes of the protein-ligand complex. Hence, MD simulation is carried out providing depth and detail in the structural information of the docked protein-ligand system. It gives complete information about the stability or flexibility of the ligand molecule on the protein surface [39]. Here, the best docked molecule (H53y) was subjected to MD simulation for a time period of 100 ns, and various MD simulation analyses were carried out.

The root mean square deviation (RMSD) is an important parameter that gives information on the structural stability of protein-ligand complexes [40]. It was observed that the RMSD for H53y increased by up to 5 ns (Figure 8). Subsequently, the RMSD of the simulated complex was stabilised and equilibrated throughout the simulation time period. The average RMSD value for H53y was 0.49 nm.

Radius of gyration (Rg) gives information about the compactness of the simulated complex during MD simulation [40]. Here, the Rg of the protein for the H53y-complex was calculated during the simulation, as shown in Figure 9a. The average Rg of the protein in the complex was found to be 1.90 nm, satisfying no considerable change in the Rg value and maintaining stability. Similarly, the solvent accessible surface area (SASA) provides information related to the contact area of the ligand with the receptor [40]. Here, the average SASA value of the protein with the ligand was found to be 130.31 nm^2 (Figure 9b).

Furthermore, root mean square fluctuation (RMSF) was performed to analyse the fluctuation of each amino acid residue of the protein during MD simulation. The fluctuation of amino acids in the complex was determined for a period of 100 ns. Figure 9c represents the RMSF plot of H53y. The amino acid residues at positions 656-663 showed higher fluctuations, but none of the residues (666, 668, and 676) constitute H-bond interactions at the active site of the receptor. Overall, the critical residues of the active site showed slight fluctuations during simulation.

Figure 10. Contribution energy (kJ/mol) of each amino acid residue after MD simulation.

Figure 11. HOMO-LUMO of the best-designed compounds.

In addition, the stability of a protein-ligand complex is determined by the hydrogen bond between the ligand and the receptor. Hydrogen bond increase stability and specificity due to their favourable distance and geometric 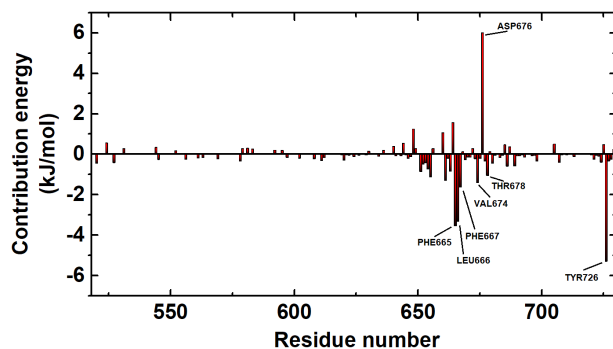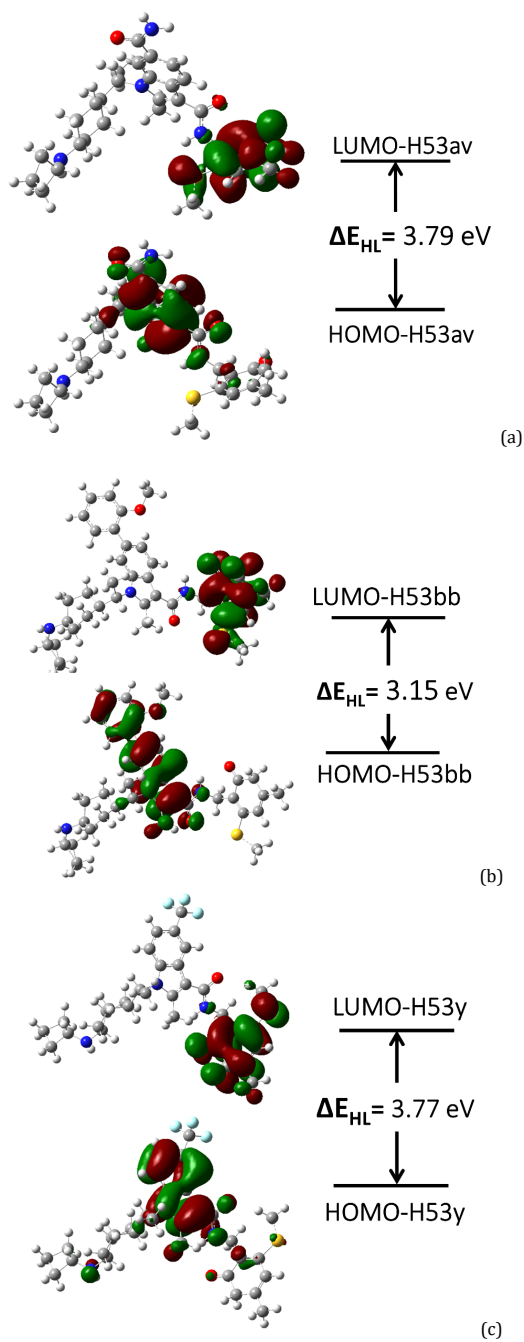restrictions. Various computational techniques are now used to access the number of hydrogen bonds in several studies of drug discovery [40]. In this study, a hydrogen bond analysis of H53y was performed to validate the result. Here, the H53y formed approximately two hydrogen bonds (Figure 9d) during the simulation, which was supported by molecular docking interactions.

3.5. MM-PBSA analysis

MM-PBSA is a computational method that accesses the free binding energy of the ligand and protein in a dynamic environment. The binding free energy of the H53y-associated complex is calculated for 100 ns of MD simulated trajectories. The van der Waals energy, electrostatic energy, polar solvation energy, and SASA energy are contributed towards the overall free binding energy calculated in MM-PBSA. Here, the H53y complex exhibited a free binding energy of −63.557 ± 0.365 kJ/mol (Table S5).

Apart from the binding energy, the contribution of each amino acid residue was also analysed towards the binding free energy of the selected complex. The analysis gives information on the contribution of important amino acid residues towards the binding site of the receptor (Figure 10). Amino acid residues Phe665 (−3.542 kJ/mol), Leu666 (−3.3181 kJ/mol), Phe667 (−1.6175 kJ/mol), Val674 (−1.4084 kJ/mol), Thr678 (−1.0595 kJ/mol), and Tyr726 (−5.2858 kJ/mol) contributed more favourably in the case of the H53y-complex.

All these analyses showed that the molecule H53y performed better during the entire MD simulation. The protein-ligand complex remains stable at the binding site of the receptor and comes out as potent, as suggested by MM-PBSA analysis.

3.6. Drug-likeness and toxicity prediction

Many computational techniques have been established to identify the drug-likeness ranging from basic schemes such as Lipinski's rule of five from PreADMET (Table S6) in predicting the activity and toxicity prediction from OSIRIS property explorer (Table S7). One designed compound (H53av) has satisfied Lipinski's rule of five, as it exhibits better drug like physicochemical properties, namely, the LogPo/w values for this compound were <5. The remaining two compounds violate this criterion because of their higher molecular weight (MW) and LogPo/w. Plasma protein binding (PPB) data of these three compounds have represented moderate binding (>80%). On the other hand, human intestinal absorption (HIA) data have shown good absorptivity (>94%) for these designed compounds. Additionally, the values of blood-brain barrier (BBB) penetration data tabulated in Table S6 have shown that the two molecules (H53av and H53bb) have moderate absorptivity, while H53y has high absorptivity toward the central nervous system.

In evaluating the toxicity for the designed compounds, all molecules were devoid of important organ toxicities like mutagenicity, tumorogenicity, irritant, and reproductive effects. Topological polar surface area (TPSA) indicates drug permeability and is an important parameter for oral absorption and CNS penetration. The designed molecules were within the range of not exceeding 140 Å^2. Furthermore, their drug score values ranged from 0.16 to 0.37.

3.7. DFT analysis

The chemical reactivity and kinetic stability of the compound depend on the frontier molecular orbitals of HOMO and LUMO [41]. The energies between the molecular orbitals (HOMO and LUMO) were calculated, listed in Table S8. The HOMO and LUMO energies describe the electron-donating and accepting abilities of the molecules, respectively. The gap between them, referred to as the orbital energy band gap (ΔE_{HL}), plays a vital role in determining the electron transportation capacity of the molecule. A larger gap alludes to the fact that the molecule is chemically hard, stable, and unreactive, while a molecule with a smaller orbital energy gap is soft, unstable, and reactive [41].

It was found from the Table S8 that the ΔE_{HL} for the compounds studied ranged from 3.1 to 3.8 eV, in which H53y was the lowest (3.15 eV) (Figure 11). Hence, it was considered the softest molecule among them, being both highly polarizable as well as highly reactive. H53y was found to have the highest dipole moment value of 7.94 debye, which indicates a better tendency to participate in strong intermolecular interactions. The softness parameter, which is a characteristic of the chemical reactivity of the molecule, and the inverse values of the hardness. The chemical softness of the designed compounds was found to be between 0.53 and 0.63 eV^{-1}, while the chemical hardness was found to be between 1.58 and 1.89 eV.

4. Conclusions

In this present study, the QSAR study confirmed that the model with two descriptors (minHBint4 and Wlambdal.unity) was found to be robust and stable and successfully predicted the pIC$_{50}$ values associated with anticancer activity. The MLR approach in the QSARINS suite was used to develop the best model equation using molecular descriptors of benzofuran and indole derivatives, calculated by PaDEL software. The OECD principles were successfully followed to design and select new compounds for anticancer drugs with robustness, stability, and good predictive power of R^2 = 0.9328, Q^2_{LOO} = 0.9212, Q^2_{LMO} = 0.9187, and R^2_{ext} = 0.929, respectively. In addition, none of the compounds were found to be outliers in the Williams plot. The compound H53 from the original data set was set as a reference in the molecular docking. Finally, three designed compounds (H53av, H53bb, and H53y) with predicted pIC$_{50}$ values greater than 10 were found to be the best due to better binding activity, as well as pharmacological properties, modified by the rational drug design of the reference molecule. Therefore, these three molecules could be exploited as the most promising inhibitors in *in vitro* and *in vivo* experiments against the human EZH2 target.

Disclosure statement DS

Conflict of interest: The authors declare that they have no conflict of interest.
Ethical approval: All ethical guidelines have been adhered to.
Sample availability: Samples of the compounds are available from the author.

CRediT authorship contribution statement CR

Conceptualization: Kaushik Sarkar, Rajesh Kumar Das; Methodology: Kaushik Sarkar, Sraboni Ghosh, Rajesh Kumar Das; Software: Kaushik Sarkar, Sraboni Ghosh, Rajesh Kumar Das; Validation: Kaushik Sarkar, Sraboni Ghosh, Rajesh Kumar Das Analysis: Kaushik Sarkar, Rajesh Kumar Das; Investigation: Kaushik Sarkar, Rajesh Kumar Das; Data Curation: Kaushik Sarkar, Sraboni Ghosh, Rajesh Kumar Das; Writing - Original Draft: Kaushik Sarkar, Rajesh

Kumar Das; Writing - Review and Editing: Rajesh Kumar Das; Visualization: Kaushik Sarkar, Sraboni Ghosh, Rajesh Kumar Das; Supervision: Rajesh Kumar Das.

ORCID 🆔 and Email ✉

Kaushik Sarkar
✉ kaushikchemistry@gmail.com
🆔 https://orcid.org/0000-0002-3782-0948
Sraboni Ghosh
✉ israbonighosh@gmail.com
🆔 https://orcid.org/0000-0001-7290-9924
Rajesh Kumar Das
✉ rajeshnbu@gmail.com
🆔 https://orcid.org/0000-0001-9433-4711

References

[1]. Verma, S. K.; Tian, X.; LaFrance, L. V.; Duquenne, C.; Suarez, D. P.; Newlander, K. A.; Romeril, S. P.; Burgess, J. L.; Grant, S. W.; Brackley, J. A.; Graves, A. P.; Scherzer, D. A.; Shu, A.; Thompson, C.; Ott, H. M.; Van Aller, G. S.; Machutta, C. A.; Diaz, E.; Jiang, Y.; Johnson, N. W.; Knight, S. D.; Kruger, R. G.; McCabe, M. T.; Dhanak, D.; Tummino, P. J.; Creasy, C. L.; Miller, W. H. Identification of potent, selective, cell-active inhibitors of the histone lysine methyltransferase EZH2. *ACS Med. Chem. Lett.* **2012**, *3*, 1091–1096.

[2]. Son, M. J.; Kim, W. K.; Park, A.; Oh, K.-J.; Kim, J.-H.; Han, B. S.; Kim, I. C.; Chi, S.-W.; Park, S. G.; Lee, S. C.; Bae, K.-H. Set7/9, a methyltransferase, regulates the thermogenic program during brown adipocyte differentiation through the modulation of p53 acetylation. *Mol. Cell. Endocrinol.* **2016**, *431*, 46–53.

[3]. Batista, I. de A. A.; Helguero, L. A. Biological processes and signal transduction pathways regulated by the protein methyltransferase SETD7 and their significance in cancer. *Signal Transduct. Target. Ther.* **2018**, *3*, 19.

[4]. Pradhan, S.; Chin, H. G.; Estève, P.-O.; Jacobsen, S. E. SET7/9 mediated methylation of non-histone proteins in mammalian cells. *Epigenetics* **2009**, *4*, 383–387.

[5]. Peterson, C. L.; Laniel, M.-A. Histones and histone modifications. *Curr. Biol.* **2004**, *14*, R546-51.

[6]. Biggar, K. K.; Li, S. S.-C. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 5–17.

[7]. Hahm, J. Y.; Kim, J.-Y.; Park, J. W.; Kang, J.-Y.; Kim, K.-B.; Kim, S.-R.; Cho, H.; Seo, S.-B. Methylation of UHRF1 by SET7 is essential for DNA double-strand break repair. *Nucleic Acids Res.* **2019**, *47*, 184–196.

[8]. Song, Y.; Zhang, J.; Tian, T.; Fu, X.; Wang, W.; Li, S.; Shi, T.; Suo, A.; Ruan, Z.; Guo, H.; Yao, Y. SET7/9 inhibits oncogenic activities through regulation of Gli-1 expression in breast cancer. *Tumour Biol.* **2016**, *37*, 9311–9322.

[9]. Takemoto, Y.; Ito, A.; Niwa, H.; Okamura, M.; Fujiwara, T.; Hirano, T.; Handa, N.; Umehara, T.; Sonoda, T.; Ogawa, K.; Tariq, M.; Nishino, N.; Dan, S.; Kagechika, H.; Yamori, T.; Yokoyama, S.; Yoshida, M. Identification of cyproheptadine as an inhibitor of SET domain containing lysine methyltransferase 7/9 (Set7/9) that regulates estrogen-dependent transcription. *J. Med. Chem.* **2016**, *59*, 3650–3660.

[10]. Fujiwara, T.; Ohira, K.; Urushibara, K.; Ito, A.; Yoshida, M.; Kanai, M.; Tanatani, A.; Kagechika, H.; Hirano, T. Steric structure–activity relationship of cyproheptadine derivatives as inhibitors of histone methyltransferase Set7/9. *Bioorg. Med. Chem.* **2016**, *24*, 4318–4323.

[11]. Feoli, A.; Viviano, M.; Cipriano, A.; Milite, C.; Castellano, S.; Sbardella, G. Lysine methyltransferase inhibitors: where we are now. *RSC Chem. Biol.* **2022**, *3*, 359–406.

[12]. Chang, C.-J.; Hung, M.-C. The role of EZH2 in tumour progression. *Br. J. Cancer* **2012**, *106*, 243–247.

[13]. Woo, J.; Kim, H.-Y.; Byun, B. J.; Chae, C.-H.; Lee, J. Y.; Ryu, S. Y.; Park, W.-K.; Cho, H.; Choi, G. Biological evaluation of tanshindiols as EZH2 histone methyltransferase inhibitors. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2486–2492.

[14]. Nasveschuk, C. G.; Gagnon, A.; Garapaty-Rao, S.; Balasubramanian, S.; Campbell, R.; Lee, C.; Zhao, F.; Bergeron, L.; Cummings, R.; Trojer, P.; Audia, J. E.; Albrecht, B. K.; Harmange, J.-C. P. Discovery and optimization of tetramethylpiperidinyl benzamides as inhibitors of EZH2. *ACS Med. Chem. Lett.* **2014**, *5*, 378–383.

[15]. Bradley, W. D.; Arora, S.; Busby, J.; Balasubramanian, S.; Gehling, V. S.; Nasveschuk, C. G.; Vaswani, R. G.; Yuan, C.-C.; Hatton, C.; Zhao, F.; Williamson, K. E.; Iyer, P.; Méndez, J.; Campbell, R.; Cantone, N.; Garapaty-Rao, S.; Audia, J. E.; Cook, A. S.; Dakin, L. A.; Albrecht, B. K.; Harmange, J.-C.; Daniels, D. L.; Cummings, R. T.; Bryant, B. M.; Normant, E.; Trojer, P. EZH2 inhibitor efficacy in non-Hodgkin's lymphoma does

[16]. not require suppression of H3K27 monomethylation. *Chem. Biol.* **2014**, *21*, 1463–1475.

[16]. Siddiqui, S. K.; SahayaSheela, V. J.; Kolluru, S.; Pandian, G. N.; Santhoshkumar, T. R.; Dan, V. M.; Ramana, C. V. Discovery of 3-(benzofuran-2-ylmethyl)-1H-indole derivatives as potential autophagy inducers in cervical cancer cells. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 127431.

[17]. Al-Akra, L.; Bae, D.-H.; Leck, L. Y. W.; Richardson, D. R.; Jansson, P. J. The biochemical and molecular mechanisms involved in the role of tumor micro-environment stress in development of drug resistance. *Biochim. Biophys. Acta Gen. Subj.* **2019**, *1863*, 1390–1397.

[18]. Mu, L.-M.; Ju, R.-J.; Liu, R.; Bu, Y.-Z.; Zhang, J.-Y.; Li, X.-Q.; Zeng, F.; Lu, W.-L. Dual-functional drug liposomes in treatment of resistant cancers. *Adv. Drug Deliv. Rev.* **2017**, *115*, 46–56.

[19]. Jia, Y.; Wen, X.; Gong, Y.; Wang, X. Current scenario of indole derivatives with potential anti-drug-resistant cancer activity. *Eur. J. Med. Chem.* **2020**, *200*, 112359.

[20]. Napiórkowska, M.; Cieślak, M.; Kaźmierczak-Barańska, J.; Królewska-Golińska, K.; Nawrot, B. Synthesis of new derivatives of benzofuran as potential anticancer agents. *Molecules* **2019**, *24*, 1529.

[21]. Yap, C. W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2011**, *32*, 1466–1474.

[22]. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **2012**, *4*, 17.

[23]. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132.

[24]. Cañizares-Carmenate, Y.; Campos Delgado, L. E.; Torrens, F.; Castillo-Garit, J. A. Thorough evaluation of OECD principles in modelling of 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine derivatives using QSARINS. *SAR QSAR Environ. Res.* **2020**, *31*, 741–759.

[25]. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

[26]. Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.

[27]. Lee, S.; Tran, A.; Allsopp, M.; Lim, J. B.; Hénin, J.; Klauda, J. B. CHARMM36 united atom chain model for lipids and surfactants. *J. Phys. Chem. B* **2014**, *118*, 547–556.

[28]. Boonstra, S.; Onck, P. R.; Giessen, E. van der CHARMM TIP3P water model suppresses peptide folding by solvating the unfolded state. *J. Phys. Chem. B* **2016**, *120*, 3692–3698.

[29]. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2009**, *31*, 671–690.

[30]. Vanommeslaeghe, K.; MacKerell, A. D., Jr Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.

[31]. Kumari, R.; Kumar, R.; Open Source Drug Discovery Consortium; Lynn, A. g_mmpbsa--a GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962.

[32]. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037–10041.

[33]. Molecular properties prediction - Osiris property explorer. https://www.organic-chemistry.org/prog/peo/ (accessed February 9, 2023).

[34]. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc. , Wallingford CT, 2016.

[35]. Dennington, R.; Keith, T. A.; Millam, J. M. GaussView, Version 6, Semichem Inc.; Shawnee Mission, KS, 2016.

[36]. Tirado-Rives, J.; Jorgensen, W. L. Performance of B3LYP density functional methods for a large set of organic molecules. *J. Chem. Theory Comput.* **2008**, *4*, 297–306.

[37]. Abdulfatai, U.; Uzairu, A.; Uba, S. Molecular docking and QSAR analysis of a few Gama amino butyric acid aminotransferase inhibitors. *Egypt. J. Basic Appl. Sci.* **2018**, *5*, 41–53.

[38]. Subramani, A. K.; Sivaperuman, A.; Natarajan, R.; Bhandare, R. R.; Shaik, A. B. QSAR and molecular docking studies of pyrimidine-coumarin-triazole conjugates as prospective anti-breast cancer agents. *Molecules* **2022**, *27*, 1845.

[39]. Sarkar, K.; Das, R. K. Repurposing of existing pharmaceutical drugs against monkey-pox virus: An *in silico* study. *Anal. Chem. Lett.* **2022**, *12*, 655–670.

[40]. Sharma, B.; Bhattacherjee, D.; Zyryanov, G. V.; Purohit, R. An insight from computational approach to explore novel, high-affinity phosphodiesterase 10A inhibitors for neurological disorders. *J. Biomol. Struct. Dyn.* **2022**, 1–13.

[41]. Al-Ostoot, F. H.; Geetha, D. V.; Mohammed, Y. H. E.; Akhileshwari, P.; Sridhar, M. A.; Khanum, S. A. Design-based synthesis, molecular docking analysis of an anti-inflammatory drug, and geometrical optimization and interaction energy studies of an indole acetamide derivative. *J. Mol. Struct.* **2020**, *1202*, 127244.