

***In-silico* QSAR Modeling for Identification of Novel Anticancer Agent as Anaphylactic Lymphoma Kinase Inhibitor**

R.D. Jawarkar^{a,*}, P. Sharma^a, S.K. Jain^a and S. Mali^b

^aFaculty of Pharmacy, Oriental University, Indore 453555, Madhya Pradesh, India

^bDepartment of Pharmaceutical Sciences and Technology, Birla Institute of Technology, Ranchi

(Received 2 October 2023, Accepted 15 November 2023)

The ALK tyrosine kinase receptor is a promising target in lung cancer. To estimate ALK-TK inhibitory activity, we used QSAR modeling on heterocyclic compounds with varied structures and a large dataset of 1329 chemicals experimentally reported for anticancer activity against ALK-TK. The developed QSAR model meets various validation criteria, such as $R^2 = 0.79$, $Q^2_{LOO} = 0.78$, $Q^2_{LMO} = 0.78$, $R^2_{ex} = 0.77$, and $CCC_{ex} = 0.87$. In addition, we have used QSAR-based virtual screening to find 12 FDA compounds as in-silico hits, some of which could be used in clinical settings as ALK-TK inhibitors with a docking score ranging from -7.10 to -10.57 kcal mol⁻¹. For both wild-type and mutant ALK-TK, QSAR-based virtual screening predicted a pIC₅₀ of 9.18 M for the new compound ZINC000150338819 with a docking score of -10.57 kcal mol⁻¹ (RMSD 1.54 Å). MD simulation and MMGBSA investigations confirm that the ZINC000150338819-ALK TK complex is stable for 200 ns for both wild-type and mutant ALK TK. To confirm the in-silico findings, the MTT assay reveals that Ledipasvir showed more inhibition as compared to ceritinib. This study suggests that the hit compound ZINC000150338819 may be a repurposed ALK TK inhibitor in drug discovery.

Keywords: OECD, GA-MLR, MD simulation, MMGBSA, X-Ray

INTRODUCTION

In 1991, researchers discovered a membrane-bound tyrosine kinase receptor called anaplastic lymphoma kinase (ALK). Abnormal forms of ALK, such as fusion proteins, point mutations that activate ALK, and gene amplification, are found in cancer. Abnormal ALK expression is linked to the development of many types of cancer [1]. In 1997, many notable studies were conducted on ALK in its wild forms. Since then, the receptor tyrosine kinase has been the primary focus of research [2,3].

The human ALK gene is located in the chromosomal region 2p23.2p23.1. This 26-exon gene encodes the ALK protein, which is 1620 amino acids long. The full-length ALK protein has a transmembrane region, a ligand-binding region, and an intracellular tyrosine kinase region consisting

of 561 amino acids [2,3]. The 3-tyrosine motif (Tyr1278, Tyr1282, and Tyr1283), where autophosphorylation for kinase activity occurs, is also found in other members of the same family of kinases (See Fig. 1).

The tyrosine kinase enzyme ALK (ATP) converts a tyrosine residue on a substrate protein to adenosine triphosphate. It catalyzes the phosphorylation of protein tyrosine residues. Several enzymes, including kinases and phosphatases, carry out the crucial phosphorylation and dephosphorylation of proteins in various biological processes. The amino-terminal part of nucleophosmin (NPM) and an ALK kinase domain are both found in the NPMALK/p80 proteins. The ALK kinase domain can form functional dimers more efficiently. The NPMALK fusion protein is the first known ALK fusion protein that causes cancer. The full-length ALK receptor tyrosine kinase (RTK) was discovered several years ago. This protein consists of transmembrane regions, an extracellular ligand-binding

*Corresponding author. E-mail: rahuljawarkar@gmail.com

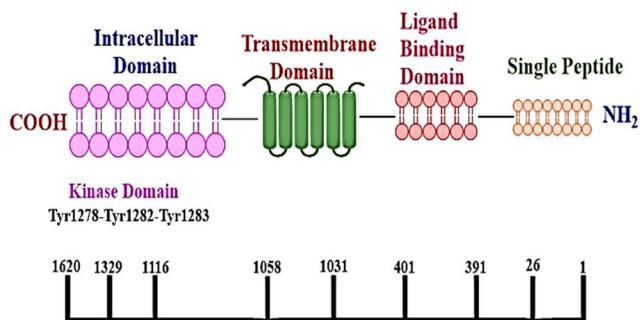


Fig. 1. Structure of ALK, showing a polypeptide of 1620 amino acids. The full-length form of ALK resembles other receptor tyrosine kinases. Kinase activity is regulated by a 3-tyrosine motif (Tyr1278, Tyr1282, and Tyr1283), which is located in an intracellular tyrosine kinase domain.

region, and an intracellular tyrosine kinase domain [2,3,4].

In mice and humans, the ALK locus encodes a receptor tyrosine kinase (RTK) with ligand-binding, transmembrane, and intracellular domains. ALK and LTK are similar enough to form a subfamily of the insulin receptor superfamily. Unlike other RTKs, ALK and LTK include glycine-rich extracellular domains. ALK comprises LDLa and MAM domains [5]. Mammals also activate ALK using Pleiotrophin (PTN) and midkine (MK) [6,7,8,9,10]. Inhibiting the ALK kinase of interest is necessary to target ALK in various

cancers. This is because oncogenic stimulation of ALK kinase activity is required for the formation of ALK fusion proteins and the acquisition of ALK's distinctive factor mutations [11,12]. Several research studies have revealed that TKIs' ability to inhibit ALK kinase activity has a substantial anticancer effect [13-20]. Moreover, numerous potent and selective ALK-TKIs have been designed to inhibit fusion proteins and activate ALK variants [21-31]. Many ALK TKIs are available for cancer therapy in different regions. The FDA has approved several ALK inhibitors for cancer therapy, including ceritinib, brigatinib, crizotinib, alectinib, TPX-0131, GSK1838705, CEP28122, AP26113, and X-396 (See Fig. 2).

Targeted therapy using the epidermal growth factor receptor (EGFR) has been discovered as a treatment for NSCLC patients with ALK activation mutations. However, most patients who receive targeted treatment relapse due to genetic changes that confer resistance [21,32]. PF02341066 (crizotinib) is an aminopyridine ALK inhibitor that is now being tested in phase III clinical studies [33].

Most QSAR studies on ALK-TK inhibitors have limited datasets, poor prediction, no mechanistic explanation, or a combination of these flaws, which limits their utility. In this study, QSAR analysis was performed on a large and diverse experimentally reported dataset of ALK-TK inhibitors in compliance with OECD regulations. To locate compounds

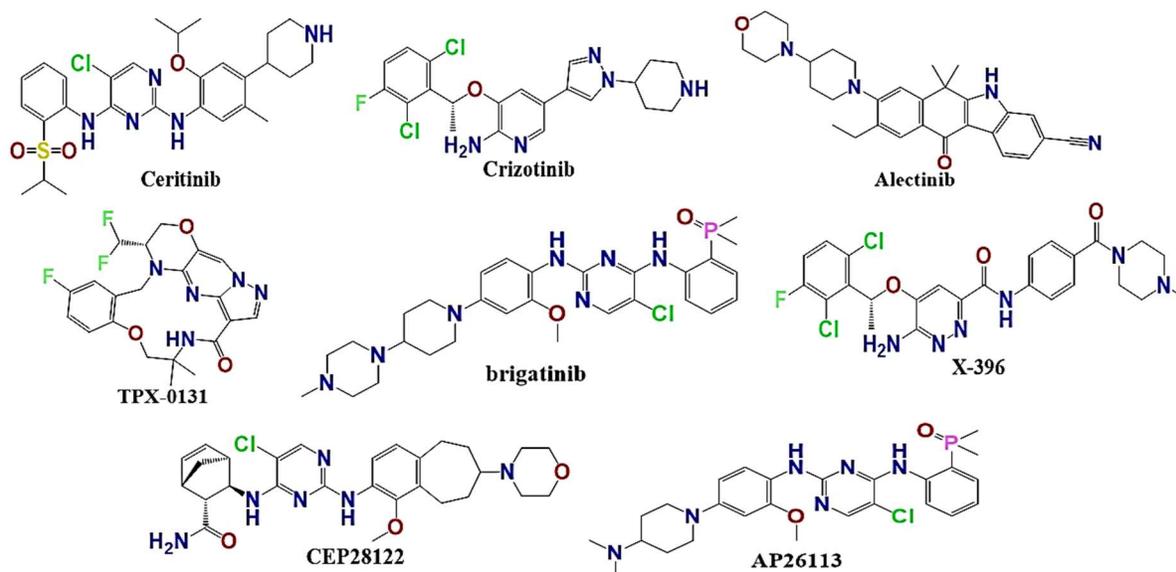


Fig. 2. Some of the ALK tyrosine kinase discovered to date and available in clinical practice.

with required characteristics in huge libraries, QSAR-based virtual screening (VS) was studied. The compounds were then experimentally validated. VS, like other computational methods, speeds up the procedure, narrows the candidate pool, and illuminates the selection process [34].

Repurposing "old" drugs for rare and multiple diseases may minimize research costs and development times [35]. In the present work, we used a QSAR-based virtual screening approach combined with drug repositioning to find a unique lead drug candidate that might effectively inhibit ALK tyrosine kinase receptors. The interactions between the newly discovered FDA drugs were investigated using molecular docking, MD modeling, and MMGBSA studies. To corroborate the results obtained virtually, we conducted an *in vitro* MTT experiment using the A549 lung cancer cell line.

EXPERIMENTAL

QSAR Methodology

The current study adheres to the conventional technique advised by the OECD and other scholars for doing QSAR analysis [36-38]. All tools were used with their default values; however, certain settings were altered, and these are detailed in the documentation, in order to generate a robust QSAR model with an equilibrium of predictive capacity and mechanistic understanding.

Step 1: Data collection and curation. To start, the Binding DB (<https://www.bindingdb.org/bind/index.jsp>; last accessed: 12/24/2021) was used to get a large dataset with 1806 IC₅₀ values for experimentally proven ALK-TK inhibitors. Data quality, and suitable curation before further processing have a significant impact on QSAR analysis [37,39-41]. After that, we filtered the data [42], which included eliminating duplicates, organometallic compounds, salts, molecules with ambiguous IC₅₀ values, and so on. This resulted in a reduction of the dataset's molecules from 1807 to 1328. In spite of the reduced size of the dataset, it still included molecules with experimental IC₅₀ (nM) values between 0.3 and 83,000 nM and the occurrence of different scaffolds such as heterocyclic rings, positional isomers, stereoisomers, *etc.*, all of which widened the chemical space and increased the model's applicability. Table 1 (See Table 1 in the Supplementary Materials) contains the SMILES (Simplified Molecular Input Line Entry System)

nomenclature for all of the compounds used in this investigation, together with their experimental IC₅₀ and pIC₅₀ (=log₁₀IC₅₀). To further illustrate the structural variety of the molecules in the collection, some illustrative examples have been provided in Fig. 3.

Table 1 shows a few representative values for IC₅₀ (nM) and pIC₅₀ (M) in the SMILES format, along with some examples of the most and least active compounds.

Step 2. Second, we utilized the default settings for OpenBabel 2.4 [43] and MOPAC 2012 (openmopac.net, obtained on March 5, 2022) to generate SMILES notations for the optimum 3D structures of the compounds (semi-empirical PM3 technique).

Step 3. If enough molecular descriptors are generated and subsequently pruned to limit the chance of overfitting from redundant noisy descriptors, then a QSAR model may strike a satisfactory balance between mechanistic interpretation and predictive ability [44]. Then, molecular descriptors were generated from 1D to 3D for every molecule. For this objective, we used PyDescriptor [45], which can calculate over 40,000 molecular descriptors for a given molecule. To reduce the number of molecular descriptors in the descriptor pool, we used QSARINS 2.2.4 to eliminate duplicates and variables with strong correlation ($|R| > 0.95$ or $> 98\%$) [46]. This resulted in a reduction from 40,000 to 2,376 but still covered a broad spectrum of molecular descriptors.

Subjective Feature Selection (SFS) Involves Separating the Dataset into a Training Set and an External Set

Splitting the dataset into a training set and an external set (also known as a prediction set or test set) is essential for developing and validating a reliable QSAR model [37,39-41]. In order to exclude any possible bias, we divided the dataset into a training set of 1062 molecules (80%) and an external collection of 266 molecules (20%) for this study. The only purpose served by the external set was model validation (predictive QSAR), whereas the molecular descriptors were selected from the training set to determine how many parameters were desired. QSARINS 2.2.4 was used to develop the model through multi-linear regression (MLR) and the Genetic Algorithm (GA). With a fitness function of Q²_{LOO} and 10,000 iterations. Determining the

right number of molecular descriptors to use in a model's creation is an important step in quantitative structure-activity relationship (QSAR) modeling. Until the value of Q^2_{LOO} rose beyond a certain threshold, the heuristic search had to

construct a large number of models, beginning with a univariate model and progressing to a multivariate model as additional molecular descriptors were included.

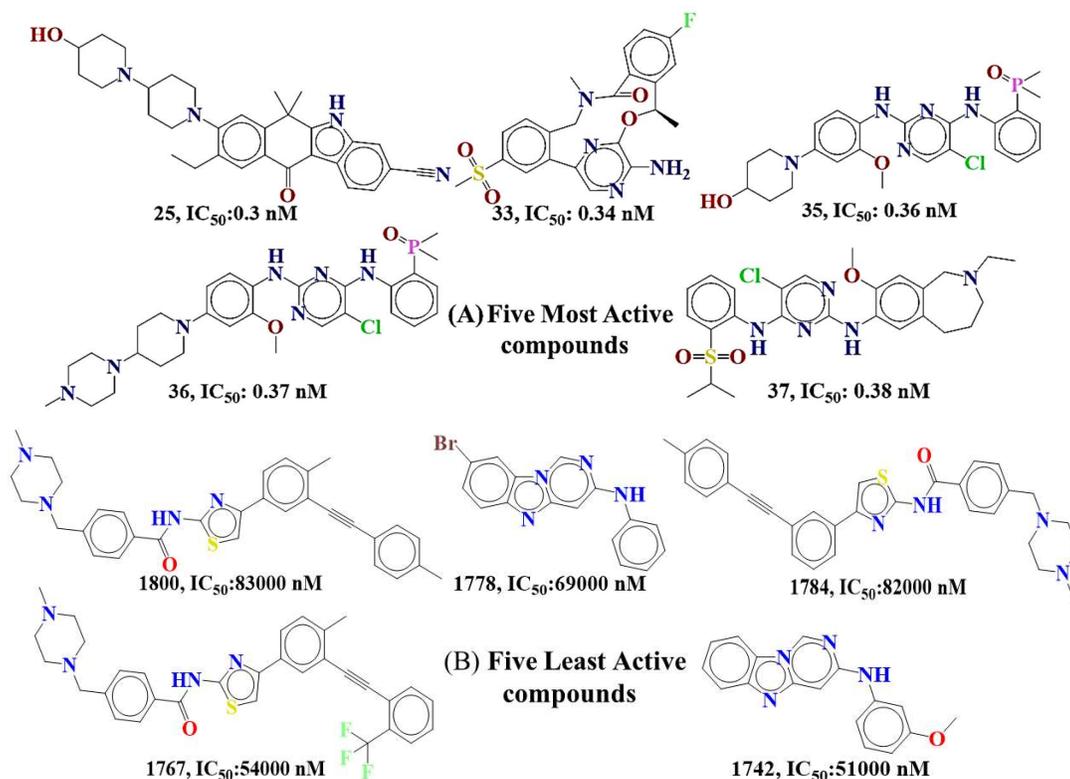


Fig. 3. The top five most active (A) and least active (B) five molecules in the provided dataset in terms of activity.

Table 1. SMILES Notation, IC_{50} (nM), and pIC_{50} (M) Values for the Five most and Least Active Compounds in the Selected Data Set

Id	SMILES	IC_{50} in nM	pIC_{50}
25	<chem>CCc1cc2C(=O)c3c([nH]c4cc(ccc34)C#N)C(C)(C)c2cc1N1CCC(CC1)N1CCC(O)CC1</chem>	0.3	9.523
33	<chem>C[C@H]1Oc2nc(enc2N)-c2cc(ccc2CN(C)C(=O)c2ccc(F)cc12)S(C)(=O)=O</chem>	0.34	9.469
35	<chem>COc1cc(ccc1Nc1cc(Cl)c(Nc2ccccc2P(C)(C)=O)n1)N1CCC(O)CC1</chem>	0.36	9.444
36	<chem>COc1cc(ccc1Nc1cc(Cl)c(Nc2ccccc2P(C)(C)=O)n1)N1CCC(CC1)N1CCN(C)CC1</chem>	0.37	9.432
37	<chem>CCN1CCCc2cc(Nc3ncc(Cl)c(Nc4ccccc4S(=O)(=O)C(C)C)n3)c(OC)cc2C1</chem>	0.38	9.42
1791	<chem>CN1CCN(CC1)c1ccc(Nc2ncc3ccc(-c4ccccc4C(N)=O)n3n2)cc1</chem>	726	6.139
1793	<chem>COc1ccccc1C#Cc1ccnc2[nH]c3ccc(cc3c12)-c1ccc(cc1)N1CCN(C)CC1</chem>	750	6.125
1800	<chem>CN1CCN(Cc2ccc(cc2)C(=O)Nc2nc(cs2)-c2ccc(C)c(c2)C#Cc2ccc(C)cc2)CC1</chem>	83000	4.081
1803	<chem>CC1(C)c2[nH]c3cc(ccc3c2C(=O)c2ccc(OC[C@H](O)CO)cc12)-c1ccn[nH]1</chem>	770	6.114
1805	<chem>CC(C)S(=O)(=O)c1ccccc1Nc1nc(Nc2nc3CCN(C)CCc3s2)ncc1Cl</chem>	773	6.112

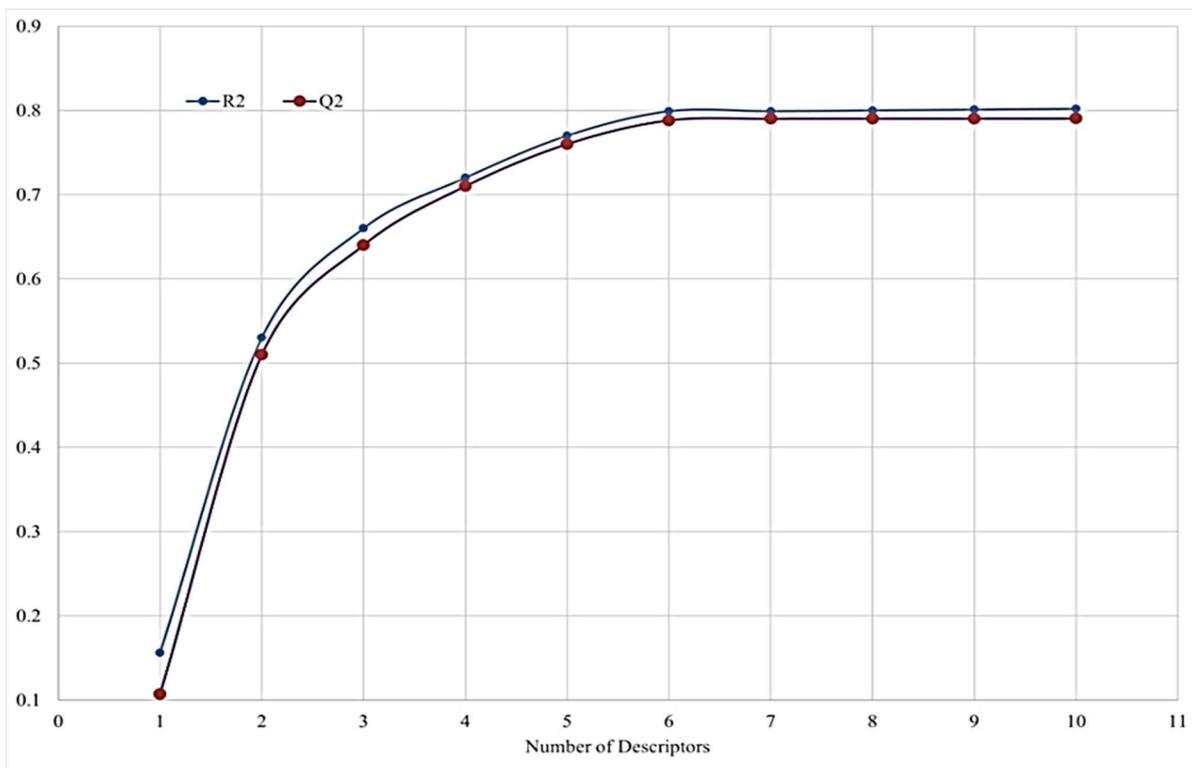


Fig. 4. Graph for the number of descriptors vs. leave-one-out Coefficient. Q^2_{LOO} for optimal descriptor count.

Figure 27 is a two-dimensional graph showing the correlation between the Q^2_{LOO} values and the number of chemical descriptors employed by the models. Adding more molecular descriptors did not increase the model's statistical performance; hence, this threshold was used to determine the ideal number of variables to include in the model. The investigation established critical values for six independent variables [47,48] (See Fig. 4). This meant that the QSAR models with more than six descriptors had to be rejected.

Constructing a Valid Regression Model and its Validation

Several validation methods, such as cross-inter validation, external validation, Y-randomization analysis, and the applicability domain (William's plot), can be used to test how accurate and reliable a QSAR model is. A well-validated QSAR model is very helpful for virtual screening, lead/hit optimization, decision-making, and other activities. Several types of validation-cross-inter, external, Y-randomization, and applicability domain (William's plot)-can

be used to approximate the robustness and wide application of a QSAR model [37,49-52].

The following are the typical criteria for assessing QSAR models, together with their corresponding threshold values for validation parameters [47,48]. There have been several successful model generation efforts using GA-MLR. The best possible model was selected using the following stringent parameters and criteria: $R^2_{tr} \geq 0.6$, $Q^2_{loo} \geq 0.5$, $Q^2_{LMO} \geq 0.6$, $R^2 > Q^2$, $R^2_{ex} \geq 0.6$, $RMSE_{tr} < RMSE_{cv}$, $K \geq 0.05$, $CCC \geq 0.80$, $Q^2_{Fn} \geq 0.60$, $r^2_m \geq 0.5$, $(1-r^2/ro^2) < 0.1$, $0.9 \leq k \leq 1.1$, or $(1-r^2/r'o^2) < 0.1$, $0.9 \leq k' \leq 1.1$, $|ro2-r'o2| < 0.3$, $RMSE_{ex}$, MAE_{ex} , R^2_{ex} , Q^2_{F1} , Q^2_{F2} , Q^2_{F3} , and low R^2_{Yscr} , $RMSE$ and MAE .

Validation of quantitative structure-activity relationship (QSAR) models entails establishing the model's applicability. We used a Williams plot to determine the extent to which the QSAR model was reliable. (Table 2 in the supplemental materials lists the computed descriptors used to create the QSAR model.) Table 3 of the supplementary materials provides the formulas for determining these

statistical features.

QSAR-Based Virtual Screening for Drug Repositioning (Repurposing)

Due to the rising need for novel anticancer treatments, drug repurposing has fascinated the cancer research community. Even though there are many ways to treat cancer, such as chemotherapy and targeted therapies, cancer is defined by its inability to respond to current medicines and drugs. Because of this, studying potential novel cancer treatments is a thriving field of study [35,53,54]. Nowadays, drug repositioning is an emerging area of study; therefore, we utilized this information to conduct a QSAR-based virtual screening using the zinc database and its 1615 FDA compounds. Since then, QSAR-based VS has used 1615 FDA substances. Before doing molecular descriptor calculations, the 3D structures of molecules were constructed in the same way as a modeling set. The ALK-TK inhibitory action of 1615 FDA-approved drugs was predicted using a completely validated six-parameter quantitative structure-activity relationship (QSAR) model, which was derived using estimated chemical descriptors. The chemical details and predicted IC₅₀ values for the zinc FDA 1615 compounds are included in Table 4 of the supplemental materials.

Molecular Docking Analysis

The ALK TK wild type (pdb-4cmu) and mutant (pdb-4clj) Protein Data Bank (pdb) data were obtained from the protein data bank [55]. The pdb:4cmu and 4clj were chosen on the basis of X-ray resolution and completion of amino acid sequences. Ramachandran's plot was used to determine the protein's health before running docking simulations. The protein, after optimization, passes muster for docking studies. Both PDB files had their native ligands removed before docking analysis could begin. To facilitate comparison between the wild-type (4cmj) and mutant (4clj) strains, all 12 hit molecules from the QSAR-based virtual screening were docked into the active sites of both. For convenience, we've included the docking position of Ledipasvir, the most active molecule.

Molecular docking analysis was performed in NRGSuite software package [56]. This is an open-access tool that is available at no cost as a plugin for the PyMOL software (www.pymol.org as of March 9, 2022). FlexAID can help

you find cavities on protein surfaces to ensure that it can be utilized in docking simulation targets [57]. Covalent docking, conformational search using a genetic algorithm, and the mobility of ligands and side chains are all modeled. For optimal performance in this study, NRGsuite was run with the following flexible-rigid docking with default parameters: Input Method for Boundary Sites: HET groups contain water molecules; have a van der Waals permeability of 0.1; have gone through 1000 generations; use the share fitness metric; reproduce using the population explosion model; have five TOP complexes; have a cylindrical form (diameter: 19); have a three-dimensional grid spacing of 0.367; have no side-chain mobility; be ligand-adaptable; not have a ligand posture for comparison; have no constraints; and so on. For both the wild-type and mutant strains, the accuracy of molecular docking was tested using two molecules: PF-06463922 and (10R) -7-amino-12-fluoro-1, 3, 10, 16-tetramethyl-16, 17-dihydro-1H-8, 4-(metheno) pyrazole (4, 3-) (2, 5, 11). ((10R) -7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16-tetrahydro-2H-8) 4-(metheno) pyrazole (4, 3-h) (2,5,11) is a well-characterized inhibitor of ALK TK that was used to verify the docking method.

Molecular Dynamics (MD) Simulations

Molecular dynamics and simulation (MDS) methods were used to look into how stable and convergent the interaction between ledipasvir and ALK TK was. This study examined both wild (pdb-4cmu) and mutant (pdb-4clj) strains for their stability. The system builder was used to construct intricate systems for the strains of Ledipasvir-wild, Ledipasvir-mutant, Ceritinib-wild, and Ceritinib-mutant. This action was undertaken in order to facilitate the execution of simulations. The system used the OPLS-2005 force field and included an explicit solvent model using SPC water molecules [58,59]. The baseline parameters for the explicit SPC water model's orthorhombic box measuring 7.0 x 7.0 x 7.0 meters were established using Desmond 2018-4 [60]. The neutralization of both wild-type and mutant ALK-TK complexes was achieved by introducing NaCl salt at a concentration of 0.15 M Na⁺ ions. In the Desmond system builder panel, the neutralise option has been selected to introduce a predetermined quantity of counterions. In the molecular dynamics simulation approach, 18 sodium (NA) ions and 15 chloride (Cl) ions were included. After using the

ASL module to choose certain residues of the ligand and protein molecules, the Desmond default relaxation technique was employed to improve the performance of the resulting systems. To learn more about each complex, we conducted molecular dynamics (MD) simulations. In the previous production run, we kept the temperature and pressure (NPT) constant and ran a molecular dynamics simulation (MDS) for 200 ns. The Nosé-Hoover chain coupling method was used to create the NPT ensemble, and the final simulation was run at 300 K with a relaxation period of 1 ps throughout the whole dynamics [60,61]. With a relaxation time of just 2 picoseconds [62], the pressure was controlled using a barostat based on the Martyna Tuckerman-Klein chain coupling system. The Desmond simulation used the isotropic Martyna-Tobias-Klein barostat and the Nose-Hoover thermostat to regulate the pressure at 1 atmosphere and the temperature at 300 Kelvin. The NPT ensemble was used in all runs, with a temperature of 300 K and a pressure of 1 bar. We successfully estimated the bonding interactions by utilizing a time step of 2 femtoseconds and the RESPA integrator. Using the particle mesh Ewald method, and keeping the radius for Coulomb interactions at 9 [63], we were able to calculate the long-range electrostatic interactions between the particles. This investigation describes the remaining possible setups. After finishing the last simulation run, the simulated trajectories of the wild-type and mutant Ledipasvir strains were analysed. Root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and hydrogen-bond formation were the primary areas of study in this examination. Binding energies for the complexes were estimated using the MM-GBSA technique, which was applied to 200 individual 1 ns trajectories. Standard deviations and mean binding energies were calculated from the obtained data.

Molecular Mechanics: Generalised Borne Surface Area

Docked complexes of ledipasvir and ceritinib were analyzed to determine their binding free energy (G_{bind}) with the help of the MM-GBSA module. During molecular dynamics (MD) simulations, the ALK complex was attached to both natural (4cmu) and mutant (4clj) strains, allowing for this estimate to be made. The New York-based Schrodinger Suite, LLC, version 2023-24, was used to run the simulations. Binding free energy was determined using a rotamer search

and calculated with the OPLS 2005 force field and the VSGB solvent model [64]. After an MD experiment was completed, a time window of 10 ns was used to choose the frames of the trajectories. By using Eq. (1), the comprehensive free energy of binding was successfully determined.

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

Where,

ΔG_{bind} = binding free energy,
 G_{complex} = free energy of the complex,
 G_{protein} = free energy of the target protein,
and G_{ligand} = free energy of the ligand.

The trajectories of the MMGBSA results were analyzed to learn more about the structural changes that occurred after the dynamics were applied.

***In-Vitro* Evaluation of Anticancer Activity by MTT Assay**

The A549 lung cancer cell line, at passage 68, was purchased from NCCS, Pune, India. F-12K medium, antibiotic-antimycotic solution, HEPES solution, and 10% fetal bovine serum were used to cultivate the cells after they were frozen. The experiment used the 3(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) assay to assess mitochondrial function. This assay relies on the reducing properties of MTT, which lead to the formation of insoluble formazan crystals specifically inside viable mitochondria. In summary, 1×10^4 A549 cells were seeded into each well of a 96-well plate, and the plate was then incubated for 24 h at 37 °C in 5% CO₂. Following removal of the medium, the cells were treated with Ceritinib (5 M), Posaconazole (10 M), Ledipasvir (20 M), and Ledipasvir (40 M) for 24 h in triplicate. The chemical ceritinib was used as a standard. Each well was given a media volume of 300 l. After 24 h of treatment, 25 l of MTT solution (5 mg ml⁻¹) was added to each well, and the cells were incubated for 4 h at 37 °C in a 5% CO₂ atmosphere. Following the dissolution of the formazan crystals in a volume of 100 l of dimethyl sulfoxide (DMSO), the absorbance was then determined at a wavelength of 570 nanometers using an Epoch Microplate Spectrophotometer manufactured by Biotek Instrument. The

IC₅₀ values were calculated using GraphPad Prism (version 7) software. Examining a nonlinear plot of the percentage of cell inhibition against the logarithm of concentration enabled this. The calculation of cell growth inhibition percentage was performed using the below formula:

$$\%Cell\ Viability = (AT/AU) \times 100$$

Where; AT = Absorbance of Treated Cells (Drug)

AU = Absorbance of Untreated Cells, %Cell Inhibition = 100 - %Cell Viability

RESULTS AND DISCUSSION

In this study, we employed molecular docking and quantitative structure-activity relationship analysis to determine structural elements essential for ALK-TK inhibition. The QSAR paradigm links structural features to simple chemical descriptors. The six-parametric GA-MLR model's unambiguous molecular descriptors and structural interpretation make it an excellent external predictor. Although it is possible to explain the effect of a specific descriptor by looking at the IC₅₀ values of the molecules in the dataset, it is important to note that the combined or inverse effect of unknown factors or other molecular descriptors could have a significant impact on a molecule's IC₅₀ value.

QSAR Model

$$pIC_{50} = 5.577 (\pm 0.087) + 1.21 (\pm 0.102) * aroC_sumpc + 0.086 (\pm 0.006) * ringC_plaN_6B + 0.221 (\pm 0.029) * fnotringNsp3C4B + -0.269 (\pm 0.027) * faroNC8B + 1.059 (\pm 0.103) * fdonnotringN5B + -0.595 (\pm 0.061) * fnotringNringN4B+$$

R²:0.7909, R²adj:0.7897, R²-R²adj:0.0012, LOF:0.2544, K_{XX}: 0.3058, Delta K:0.0661, RMSE_{tr}: 0.4987, MAE_{tr}: 0.4215, RS_{Str}: 264.3315, CCC_{tr}: 0.8832, s: 0.5003, F: 665.6663, Q²_{loo}: 0.7882, R²-Q²_{loo}: 0.0027, RMSE_{cv}: 0.5019, MAE_{cv}: 0.4242, PRESS_{cv}: 267.7734, CCC_{cv}: 0.8817, Q²_{LMO}: 0.7885, R²_{Y_{scr}}:0.0056, Q²_{Y_{scr}}:-0.0076, RMSEAV_{Y_{scr}}:1.0874, RMSE_{ext}:0.5285, MAE_{ext}:0.4460, PRESS_{ext}:74.0260, R²_{ext}:0.7710, Q²_{-F1}:0.7708, Q²_{-F2}:0.7690, Q²_{-F3}:0.7651, CCC_{ext}:0.8739, r²_{m aver.}:0.6744, r²_{m delta}:0.1472,

R²: 0.7882, R²_o: 0.7334, k': 0.9956, Clos': 0.0695, r²_m: 0.6038.

Pred(x) vs. Exp(y): R²: 0.7882, R²_o: 0.7882, k: 1.0000, Clos: 0.0000, r²_m: 0.7869, Exp(x) vs. Pred(y): R²: 0.7710, R²_o: 0.7223, k': 0.9998, Clos': 0.0632, r²_m: 0.6008, R²: 0.7710, R²_o: 0.7701, k: 0.9951, Clos: 0.0012, r²_m: 0.7480.

Figure 5 depicts the many plots used to evaluate the model's applicability, including the Williams plot (see Fig 5b), the Insubria plot (see Fig. 5c), and a comparison of the experimental and anticipated pIC₅₀. For several statistical characteristics, the generated QSAR model meets the threshold values that are suggested (formulas and statistical parameter definitions are in the supplemental material). A molecule's IC₅₀ value may be affected by unknown variables or other chemical characteristics.

R²_{tr} (Coefficient of determination) values are nearing 1, indicating that the model incorporates a sufficient number of variables without resorting to over-fitting, R²_{adj}. (Adjusted coefficient of determination), LOF (Lack of fit), and R²_{cv} (Q²_{loo}) (Cross-validated coefficient of determination for leave-one-out). Acceptable model internal validation is shown by a high value of Q²_{LMO} (the cross-validated coefficient of determination for leave-many-out). R²_{tr}, R²_{adj}, LOF, and R²_{cv} (Q²_{loo}) (Cross-validated coefficient of determination for leave-one-out) are all close, demonstrating that the model contains the appropriate number of variables and does not over-fit. A high Q²_{LMO} (the cross-validated coefficient of determination for leave-many-out) indicates that the model has satisfactory internal validation.

In order to avoid overfitting, the current QSAR model uses an inter-correlation coefficient cutoff of 0.95. Table 2 also displays the feature-feature correlation matrix. As can be observed in Table 1, the existing QSAR model does not exhibit any link between the various descriptors.

QSAR Mechanistic Interpretation

AroC_sumpc shows how important the sum of partially charged aromatic fragment carbon atoms is as one of the variables that are positively correlated in the QSAR model that was built. These descriptions emphasize the total and partial charges associated with aromatic and fragment carbon atoms. The positive coefficient for this descriptor indicates that a higher activity profile correlates with a higher aroC_sumpc value. Because the total of the partial charges on

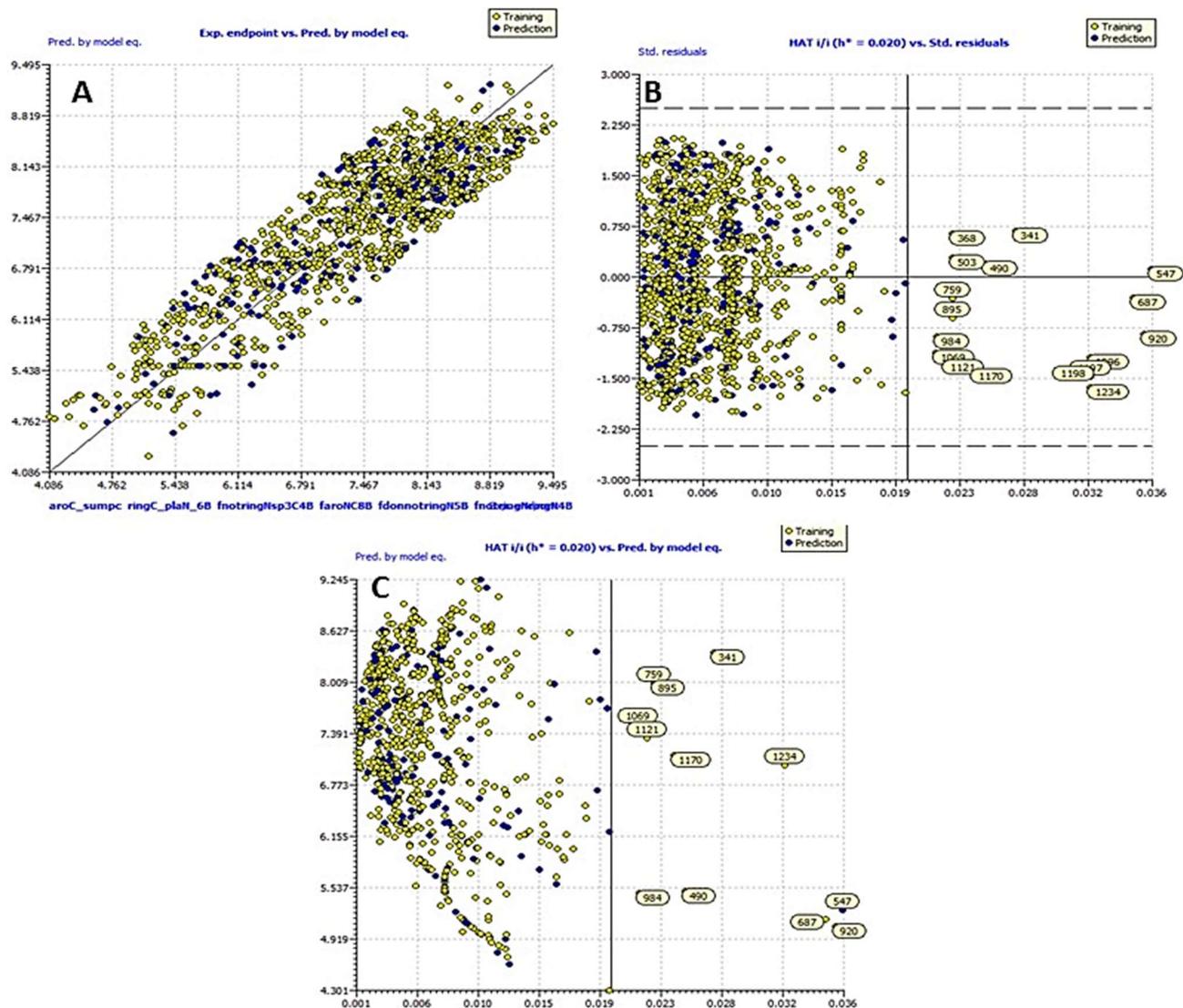


Fig. 5. QSAR model development and accompanying graphs (a) A scatter plot contrasting the expected and experimental pIC_{50} values (b) a portrayal of a Williams plot to assess the model's applicability domain; and (c) a presentation of an Insubria plot.

Table 2. Display of the Descriptor Correlation Matrix Used in the Construction of the QSAR Model

	aroC sumpc	ringC plaN 6B	fnotringNsp3C4B	faroNC8B	fdonnotringN5B	fnotringNringN4B
aroC sumpc	1					
ringC plaN 6B	0.2106	1				
fnotringNsp3C4B	0.6601	0.2798	1			
faroNC8B	0.3668	0.4161	0.3341	1		
fdonnotringN5B	-0.3209	0.1311	-0.3352	-0.3236	1	
fnotringNringN4B	0.1114	0.0567	0.1066	0.027	-0.1131	1

aromatic carbon atoms is usually positive, this property raises the pIC_{50} value. This may be determined by evaluating the subsequent pair of molecules: 502 ($IC_{50} = 4.67$ nM, $aroC_sumpc = 0.65$, $ringCplus_sumpc = 0.97$) and 1030 ($IC_{50} = 19.05$ nM, $aroC_sumpc = 0.49$, $ringCplus_sumpc = 0.85$) (See Fig. 6). Further evidence from a study comparing the following two molecules: 1060 ($IC_{50} = 54.95$ nM, $aroC_sumpc = 0.902$, $ringCplus_sumpc = 1.215$) with 1312 ($pIC_{50} = 295.1$ nM, $aroC_sumpc = 0.524$, $ringCplus_sumpc = 0.850$), As a corollary, it is interesting to note that 54 of the most active molecules (IC_{50} in the range of 1 to 0.372 nM) have higher positive values for $aroC_sumpc$, with the exception of molecules 37, 76, 86, and 90, which have $aroC_sumpc$ values in the negative range; conversely, more than 200 of the least active molecules (pIC_{50} in the range of 4 to 6) have $aroC_sumpc$ values in the negative range with low pIC_{50} . This data supports the observation that increasing the amount of partial positive charges on the aromatic or positively charged ring carbon increases lipophilicity, which promotes hydrophobic interactions, and thus improves ALK-TK inhibitory efficacy. In addition, molecule 502 is composed of a phenyl ring as a substituent at the 5th position of the pyridine ring, but in molecule 1030, it has been replaced by another pyridine ring. This result suggests that the difference in ALK-TK inhibitory activity between molecule 502 ($ringCplus_sumpc = 0.97$) and 1030 ($ringCplus_sumpc = 0.85$) may be explained by the absence of a phenyl ring in molecule 502.

Moreover, the phenyl ring contains more positively charged carbon atoms than the pyridine ring. Therefore, it can be concluded that aromatic or ring carbon atoms are essential in determining ALK-TK inhibitory activity. The molecular descriptor $aroC_sumpc$ also highlights this. Also, by replacing $aroC_sumpc$ with $ringCplus_sumpc$ (the sum of

partial charges on positively charged ring carbon atoms), a QSAR model with better statistical overall performance was made. The molecular descriptor $ringCplus_sumpc$ has a more substantial relationship ($R: 0.82$) with the pIC_{50} . As a result, keeping a positively charged carbon atom improves ALK and TK inhibitory activity. According to prior research, positively charged ring carbon atoms have a higher inhibitory impact on ALK-TK than aromatic carbon atoms. As a result, including positively charged ring carbon atoms that promote hydrophobic interaction with ALK TK is one of the strategies to modify lead compounds in the future.

$RingC_plan_6B$ showed that ring carbon and planar nitrogen are important for TK ALK to work as an inhibitor. It indicates the presence of ring carbons in the six bonds of planar nitrogen atoms. When the value of the positive factor for this descriptor in the QSAR model goes up, the inhibitory effect of TK-ALK gets much better. One way to show this is by contrasting the molecule 1065 ($IC_{50} = 21.38$ nM, $ringC_plan_6B = 20$, $C_plan_5B = 26$, $aroC_plan_6B = 16$) with the molecule 1778 ($IC_{50} = 69183$ nM, $ringC_plan_6B = 16$, $C_plan_5B = 16$, $aroC_plan_6B = 12$) (See Fig. 7). As a result, the value of the $ringC_plan_6B$ was found in the range of 16 to 25 in the 459 most active (IC_{50} in the range of 0.398 to 10 nM) molecules, while the least active molecules had a value of zero for the $ringC_plan_6B$ and the remaining least active molecules had a value in the range of 10 to 15 for the $ringC_plan_6B$. This data emphasizes the hypothesis that increasing the number of carbon and nitrogen in a combination like this boosts the inhibitory activity of ALK TK.

Interestingly, Weisheng Huang and his colleagues also reported making brigatinib (AP26113), a powerful inhibitor of anaplastic lymphoma kinase that works when taken by mouth. Their research detailed how the same basic (amine) side chain can alter ALK activity, particularly cellular potency. The molecular descriptor $ringC_plan_6B$ in the QSAR model further highlighted the relevance of the basic amine nitrogen side chain. In addition to proving the major side chain's significance, the C_plan_6B descriptor ring also pinpointed the basic (planer) nitrogen that is necessary to boost cellular potency and ALK-TK inhibitory action⁶⁵. After that, the molecular descriptor $ringC_plan_6B$ is changed to $aroC_plan_6B$, which means that there is an aromatic carbon atom inside the 6-bond ring of the planar nitrogen atom. This

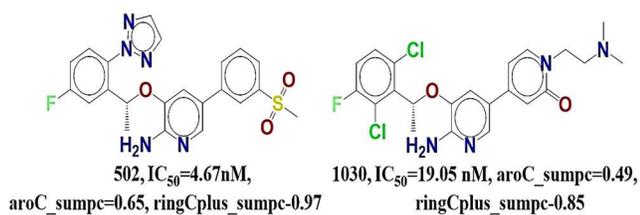


Fig. 6. The molecular descriptor $aroC_sumpc$, is shown exclusively for molecules 502 and 1030.

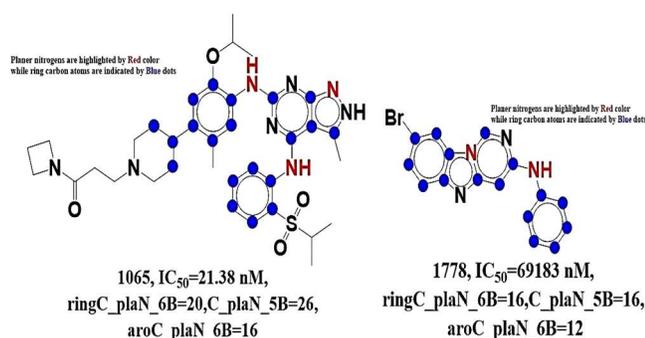


Fig. 7. For molecules 1065 and 1778, the molecular descriptor ringC_plaN_6B is shown.

slightly improves the statistical performance of the existing QSAR model ($R = 0.81$).

Conversely, swapping out the chemical descriptor ring C_plaN_6B for C_plaN_5B (the presence of a carbon atom within the 5-bond ring of the planar nitrogen atom) significantly boosts the statistical power ($R = 0.94$) of the initially created QSAR model.

This observation shows that aromatic or ring carbons are essential for ALK-TK inhibition. Still, adding a single carbon atom at or near the five bonds of the planar nitrogen atom increased the TK ALK inhibitory activity, as shown by a statistical correlation with the descriptor C_plaN_5B. This finding also shows that there should be five bonds between the carbon and the planar nitrogen atom for the best TK-ALK inhibitory activity. This observation is supported by taking a peek at the crystal structure of human anaphylactic lymphoma kinase in complex with 3-(R)-1-(5-fluoro-2(2-H1,2,3-triazol-2-yl) ethoxy) and 5-methyl-1-H-pyrazol(4yl) pyridine-2-amine (See Fig. 8).

The planar nitrogen atoms responsible for ring C_plaN_6B have been marked in red, while the carbon atom at or within 6 bonds is shown with blue dotted circles. Figure 3 shows that all of the ring's A, B, C, and D in the pdb (4ccb) ligand had hydrophobic interactions (pi-sigma, alkyl, pi-alkyl) with the following residues: Val1130, Ala1148, Leu1122, Leu1256, Leu1198, *etc.*, while the terminal NH₂ acts as a hydrogen bond donor and forms a normal hydrogen bond with the Glu1197 residue. So, this aromatic combination of carbon and nitrogen is good for improving polar and hydrophobic interactions with the ALK-TK

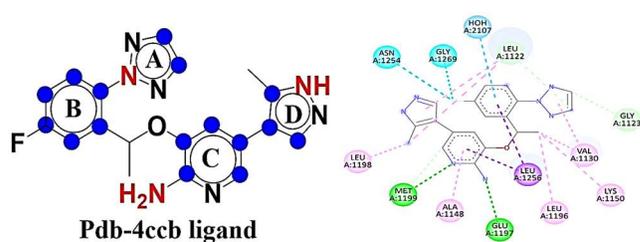


Fig. 8. The ringC_plaN_6B molecular descriptor representation in the co-crystallized ligand (pdb id = 4ccb) is shown. (Planer nitrogens are highlighted by red color while ring carbon atoms by blue bold dots).

receptor. Thus, our QSAR investigation revealed essential structural features that are visible in the X-ray-resolved crystal structure of the same target ALK-TK inhibitor

FnotringNsp3C4B

The descriptor fnotringNsp3C4B is used to show how often Sp³ hybridized carbon atoms are found within 4 bonds of acyclic or non-ring nitrogen atoms. The value of the molecular descriptor fnotringNsp3C4B increases with the activity profile since it has a positive coefficient. The fnotringNsp3C4B is not calculated if the identical sp³ hybrid carbon atom is present in 3 or 5 bonds from the carbon atom. When comparing molecule 135 ($IC_{50} = 1$ nM, fnotringNsp3C4B = 2, fnotringNnotringC4B = 2, fplaNsp3C4B = 2) to molecule 1670 ($IC_{50} = 371.5$ nM, fnotringNsp3C4B = 1, fnotringNnotringC4B = 1, fplaNsp3C4B = 1), this can be seen (See Fig. 9). A few more pairs of molecules: 25 ($IC_{50} = 0.302$ nM, fnotringNsp3C4B = 0, fnotringNnotringC4B = 0, fplaNsp3C4B = 1), 1800 ($IC_{50} = 83176.3$ nM, fnotringNsp3C4B = 0, fnotringNnotringC4B = 0, fplaNsp3C4B = 0), 35 ($IC_{50} = 0.372$ nM, fnotringNsp3C4B = 3, fnotringC4B = 3, fplaNsp3C4B = 3), & 1631 ($IC_{50} = 15488.1$ nM, fnotringNsp3C4B = 0). Both molecules 135 and 1670 include a pyrazole ring; however, although molecule 1670's pyrazole ring is completely unsubstituted, the pyrazole ring in molecule 135 is substituted with a methyl group at the N1 position and a carbonitrile group at the 5 locations. This could be a possible explanation for the variation in inhibitory efficacy between the two compounds.

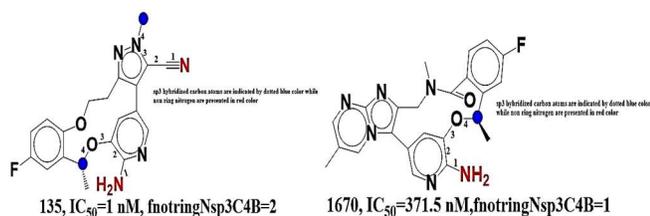


Fig. 9. For molecules 135 and 1670, the $f_{\text{notringNsp3C4B}}$ molecular descriptor is shown.

Additionally, Ted W. Johnson et al. reported the synthesis of the macrocyclic anaplastic lymphoma kinase (ALK) inhibitor PF-06463922. He underlined the relevance of the pyrazole ring in increasing the inhibitory impact against ALK-TK. He said that it was a good way to get selectivity because the cyano part of pyrazole is only one heavier than the methyl part, which is not selective. Furthermore, the benzylic chiral carbon that has been sp^3 hybridized aids in identifying the biological target (ALK TK). The descriptor $f_{\text{notringNsp3C4B}}$ visualizes these features as well. As a result, QSAR analysis not only revealed the reported pharmacophoric properties but also determined the sort of chiral carbon atom that aids in biological target identification [65]. This finding demonstrates that the total number of sp^3 hybridised carbons is necessary for ALK TK inhibitory activity; however, the model's performance is greatly enhanced ($R = 0.85$) by replacing the molecular descriptor $f_{\text{notringNsp3C4B}}$ with $f_{\text{notringNnotringC4B}}$ (frequency of the sp^3 hybridised carbon atom occurring precisely four bonds away from the noncyclic or non-ring nitrogen atom). Nonetheless, replacing $f_{\text{notringNsp3C4B}}$ with $f_{\text{donsp3C4B}}$ (the frequent occurrence of the sp^3 hybridised carbon atom with exactly 4 donor atom bonds) increases the predictive performance ($R = 0.88$) of the established QSAR model. In addition, the model's statistical performance ($R = 0.76$) is somewhat diminished when $f_{\text{plaNsp3C4B}}$ (the frequency of occurrence of sp^3 -hybridized carbons precisely at 4 bonds from the planar nitrogen atoms) is substituted for $f_{\text{notringNsp3C4B}}$. Recent research has shown that sp^3 -hybridized carbons play a critical role and that interacting with the donor atom at the optimal distance of 4 bonds greatly enhances the inhibitory impact on ALK TK. As a means of enhancing the ALK-TK inhibitory effect, it has been suggested that drugs in the future make use of

donor nitrogen or any donor atom at 4 bonds.

In most compounds, the sp^3 -hybridized carbon acts as a linkage or substituent between two rings, improving structural flexibility and allowing the molecule to assume a bioactive conformation or lipophilic characteristic. According to this theory, the chiral features of the sp^3 hybridized carbons reported in molecules 135 and 1670 confer lipophilicity, promote enantioselectivity for binding sites, and are required for the flexible alignment of these molecules inside the ALK-TK receptor's active site.

Crizotinib, when x-ray crystallized, binds to ALK TK through the same motif ($f_{\text{notringNsp3C4B}}$; pdb id 2xp2). The non-ring nitrogen atom anchored Glu 1197 of the binding site through a conventional hydrogen bond, while the sp^3 -hybridized carbon atom gave the overall flexibility to acquire the bioactive conformation at the binding site for maximum interaction with ALK TK. The X-ray crystal structure of the putative enzyme inhibitor crizotinib (See Fig. 10) corroborated the QSAR findings.

$f_{\text{donnotringN5B}}$ (The probability of the occurrence of a non-ring nitrogen atom is precisely 5 bonds apart from the donor atoms.) In the classic QSAR model, this descriptor has a positive coefficient; hence, an increase in its value leads to a larger degree of ALK-TK inhibition. When comparing molecule 245 ($IC_{50} = 2.13$ nM, $f_{\text{donnotringN5B}} = 1$) to molecule 1449 ($IC_{50} = 812.8$ nM, $f_{\text{donnotringN5B}} = 0$), this can be seen. This might explain the differences in the inhibitory activity of ALK TKs (See Fig. 11). This description demonstrates the significance of donor features and ring nitrogens in ALK-TK inhibition. The pIC_{50} value

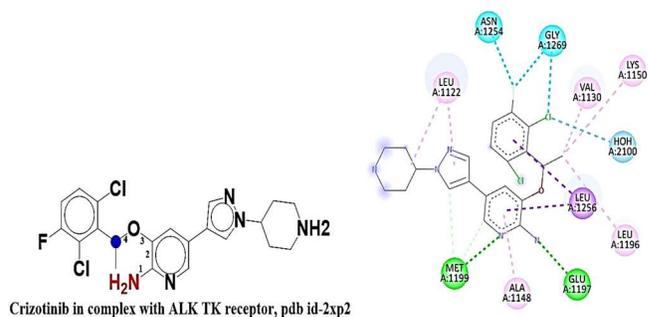


Fig. 10. Illustration of the molecular descriptor $f_{\text{notringNsp3C4B}}$ for the crizotinib in complex with ALK TK (pdb id-2xp2).

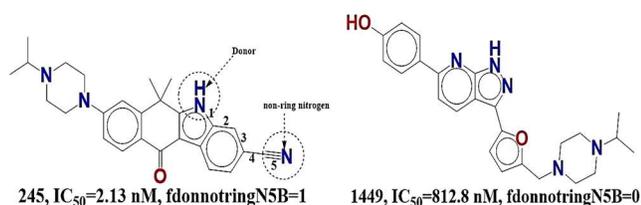


Fig. 11. For molecules 245 and 1449, the molecular descriptor $fdonnotringN5B$ is shown.

(bioactivity) of molecule 1449 rises by around 2.59 units if we modify the value of the descriptor $fdonnotringN5B$ from 0 to 1 (about a 25-fold increase in the ALK TK inhibitory potency). WeiSheng Huang et al. revealed the crystal structure of resolved brigatinib (AP 26113) in association with anaphylactic lymphoma kinase, which supports this observation [65]. The ring nitrogen of brigatinib forms hydrogen bonds with the HOH A: 1639 water molecule, whereas the non-ring nitrogen creates hydrogen bonds with the Met A: 1199 residue. The molecule is more flexible since the nitrogen is not part of a ring. In the ALK-TK binding pocket, it facilitates the transition to an active conformation.

Thus, the combination of these non-ring nitrogen and ring nitrogen atoms is required to interact with the TK-ALK receptor. So, the $fdonnotringN5B$ descriptor found in the QSAR model can also be seen in the structure of brigatinib as described (See Fig. 12). Hence, the results of the QSAR support the published data needed for developing novel ALK-TK inhibitors.

Statistical findings ($R = 0.76$) are substantially altered when the donor nitrogen ($fdonnotringN5B$) is utilized in the QSAR model instead of the molecular descriptor $fringNnotringN5B$ (the number of times the non-ring nitrogen is precisely 5 bonds distant from the ring nitrogen). Simultaneously, statistical performance ($R = 0.86$) improves when $fplaNnotringN5B$ (frequency of occurrence of non-ring nitrogen atoms exactly at 5 bonds from the planar nitrogen atom) is substituted for $fdonnotringN5B$. Therefore, the planar nitrogen with donor characteristics may be a better choice than the donor atom at the optimal bond spacing of 5 bonds to augment TK-ALK inhibitory activity.

$FaroNC8B$ denotes the frequency of the occurrence of aromatic nitrogen atoms that include carbon atoms in precisely 8 bonds. As the proposed QSAR model shows a

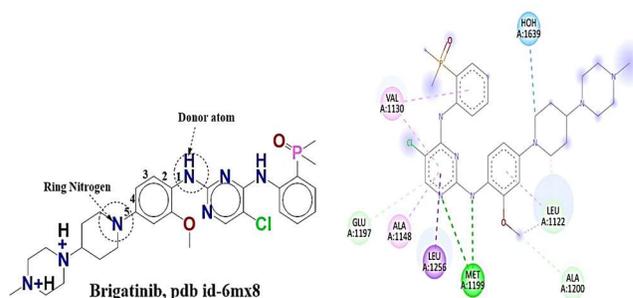


Fig. 12. Presentation of the molecular descriptor $fdonnotringN5B$ in the Brigatinib (pdb id-6mx8).

negative coefficient for this descriptor, raising its value would further reduce the inhibitory effects of ALK TK. So, further improvements to the molecule should focus on lowering the value of $faroNC8B$ to make ALK TK as effective as possible at stopping cell growth. According to this analysis, carbon and nitrogen are primarily responsible for the inhibitory activity of ALK TK. For the purposes of the $faroNC8B$ computation, a carbon atom is disregarded if it is also part of 7 or 9 other bonds. This remark can be explained by comparing molecule 1168 ($IC_{50} = 107.1$ nM, $faroNC8B = 2$) to molecule 1118 ($IC_{50} = 81.8$ nM, $faroNC8B = 1$). Decreasing the $faroNC8B$ descriptor value from 2 to 1 for molecule 1168 increases the pIC_{50} by approximately 0.12 units (roughly a 1-fold increase in ALK TK inhibitory potency for molecule 1168) (See Fig. 13). Further, switching the chemical descriptor $faroNC8B$ to $faroNsp3C8B$ (the frequency of sp^3 -hybridized carbon atoms at exactly 8 bonds in the aromatic nitrogen atom) will greatly improve the prediction accuracy of the QSAR model to $R = 0.84$. QSAR models' statistical performance ($R = 0.89$) may be greatly improved by switching back to the $faroNaroC8B$ ring from the $faroNC8B$ molecular descriptor (the frequency of occurrence of the ring carbon atom is precisely 8 bonds in the aromatic nitrogen atom).

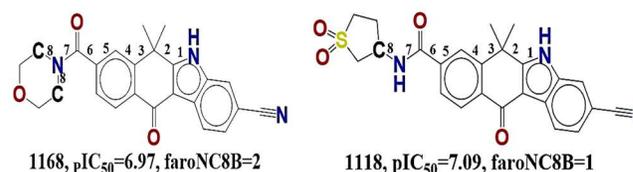


Fig. 13. Illustration of the molecular descriptor $faroNC8B$ for the molecules 1168 and 1118.

Therefore, evidence suggests that rings of carbon atoms are preferable to single carbon atoms for increasing ALK-TK's inhibitory action. For ALK TK to be most effective as an inhibitor, the ring carbon and aromatic nitrogen atoms should be separated by eight bonds. A nitrogen atom next to an aromatic ring may enhance interactions with polar regions of the receptor (ALK TK) since most nitrogen atoms function as donors or acceptors for exhibiting hydrogen bonds. In addition, the lipophilic character of aromatic rings is implied by the descriptors to play a crucial role.

$f_{\text{noNringNringN4B}}$ (frequency of occurrence of a ring nitrogen atom exactly at 4 bonds from a non-ring nitrogen atom) As the value of the descriptor goes up in the QSAR model that was made, negative coefficients of the descriptor make the inhibitor less active. This makes the ALK-TK inhibitor less active. This can be demonstrated by comparing molecule 192 ($pIC_{50} = 8.77$, $f_{\text{noNringNringN4B}} = 0$) with molecule 1343 ($pIC_{50} = 6.43$, $f_{\text{noNringNringN4B}} = 1$). The ALK-TK inhibitory activity of molecule 1343 is increased by about 2.34 units (the ALK-TK inhibitory effectiveness of molecule 1343 is increased by about 23-fold) when the value of the molecular descriptor is decreased from 1 to 0 (See Fig. 14).

Therefore, the optimum distance between ring nitrogen and non-ring nitrogen atoms should be 5 bonds, not 4 bonds. The chemical combination with hALK, as shown by X-ray crystallography (pdb id-3aox). Ethyl-N-(1-(2,4-difluorophenyl))-1(S)-N-1-(2,4-difluorophenyl)-1-butanone-3-(3-methyl-1H-Pyrazol-5-yl). The imidine (2, 1-b) pyridazin-6-amine offers more confirmation of this observation. In addition, unlike molecule 1343, molecule 192 has the identical non-ring nitrogen atom present precisely at 2, 3, 5, and 6 links, rather than 4.

The absence of a molecular descriptor in molecule 192

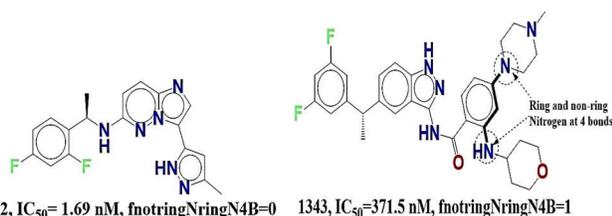


Fig. 14. The $f_{\text{noNringNringN4B}}$ molecular descriptor is presented only for molecules 192 and 1343.

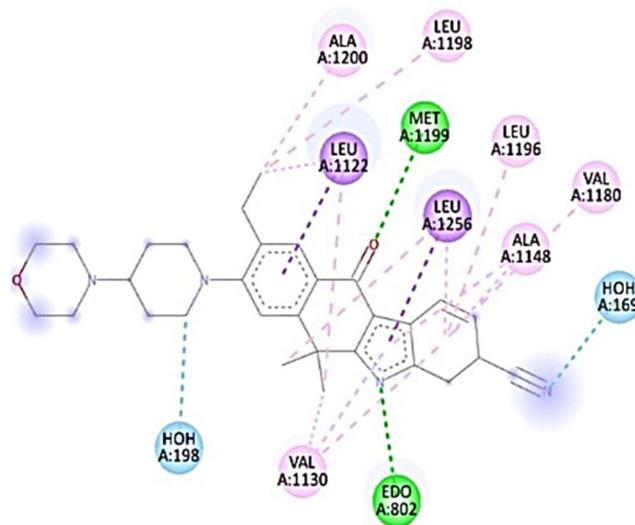


Fig. 15. Portrayal of 1(S)-N-(1-(2, 4-difluorophenyl) ethyl)-3-(3-methyl-1H-pyrazol-5-yl) imidazo [1,2-b] pyridazin-6-amine in complex with the human ALK TK (pdb id-3aox) shows the evidence for the optimal distance of 5 bonds between ring nitrogen and non-ring nitrogen atoms.

may be a feasible explanation for their different activities. (See Fig. 15) So, the new data showed that five bonds are the best distance between the nitrogen atoms in the ring and the nitrogen atoms that are not in the ring for a better ALK-TK inhibitory profile.

Drug Repositioning and QSAR-Based Virtual Screening

After making the QSAR model, we used it to predict the ALK-TK inhibitory activity of 1650 FDA molecules through a QSAR-based virtual screening. The 12 hit molecules were obtained as repurposed drug candidates against the ALK-TK receptor. Based on its pIC_{50} , the molecule ZINC000150338819 (Ledipasvir) was chosen as a major hit among the top hit molecules. The structures of the top 12 hit molecules are depicted in Fig. 16.

Applicability Domain Study of the Identified Hit Molecules

To evaluate the scope of our QSAR model's applicability, we employed a dataset consisting of 1329 molecules for training and a set of 12 hit molecules for prediction. QSAR-based virtual screening yielded 12 compounds, and these

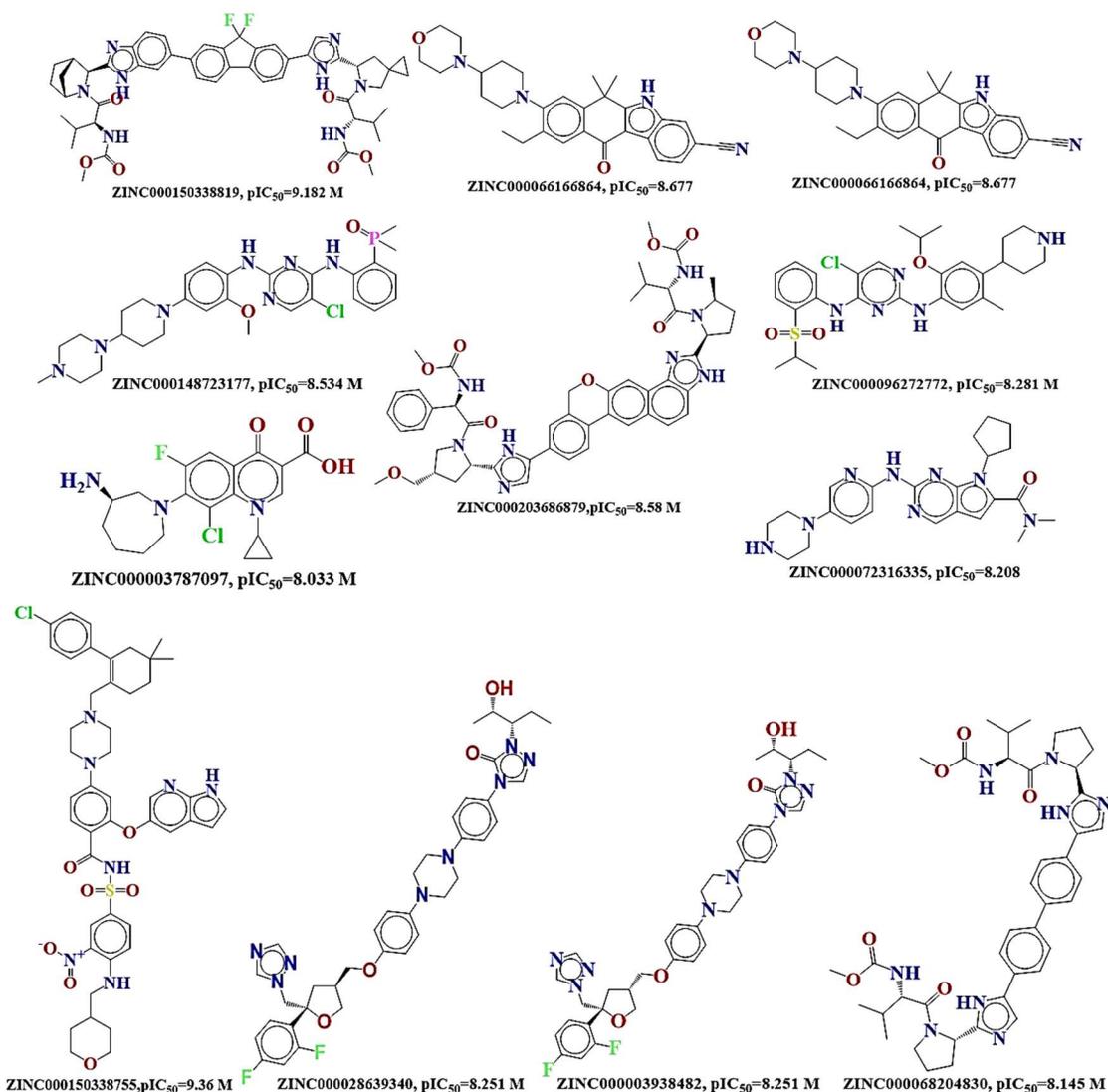


Fig. 16. Depiction of the structures of 12 hit molecules along their zinc database i.d. and predicted pIC_{50} values obtained in the QSAR Based Virtual Screening.

were characterized. HAT $i/i h^* = 0.016$ is a low leverage value, and the hit molecules ZINC000150338819, ZINC000150588351, ZINC000203686879, and ZINC000068204830 are located on the edge of the applicability domain in the Williams plot (See Fig. 17). The leverage numbers show the degree to which the structure of each chemical affects the model. The low leverage value shows that the prediction set (seven out of twelve hit molecules) is found to be within the applicability domain of the training set molecules, providing credence to the predicted data for those molecules.

Molecular Docking Analysis

Signal peptides (1-18), ligand-binding domains (19-1038), transmembrane regions (1039-1059), and tyrosine kinase domains (1060-1620) make up the human ALK protein. ALK fusions with a wide variety of partner genes, such as NPM-ALK and EML4-ALK, may be triggered by chromosomal rearrangements on the 2p23 chromosome tract. The kinase domain of ALK consists of two lobes, one tiny at the amino terminus (N-lobe) and one large at the carboxyl terminus (C-lobe), connected by a hinge region that may open or close. The two lobes are separated by a fissure that serves

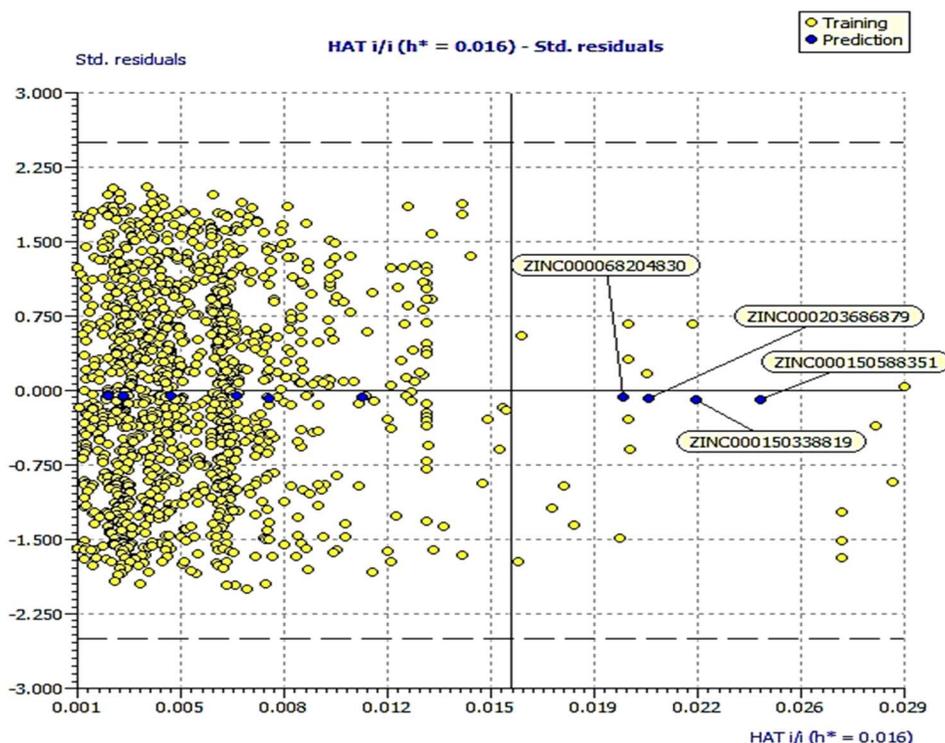


Fig. 17. The 12 hit molecules from the QSAR-based virtual screening are shown in a Williams plot for their potential applicability domain (blue dots represent the hit molecules).

as an ATP-binding pocket. The tiny N-lobe is made up of the controlling C-helix and five stranded sheets (numbered 1-5) [66]. Coordination of ATP39 and phosphates is facilitated by a glycine-rich loop (G-loop) between the N-lobe's strands 1 and 2. The big C-lobe has six conserved helices (D-I) and two short conserved strands (7-8) between the E- and F-helices⁶⁶.⁶⁷ Unphosphorylated ALK kinase domains include eight strands within the activation loop (A-loop), with an extra helix (EF) following. Hydrophobic non-contiguous motifs of ALK's N- and C-lobes spanning regulatory (R-spine) and catalytic (C-spine) regions are also maintained.

During insulin receptor activation, the R-spine is put together. Its disassembly controls the cycles of ALK41 activation and deactivation by changing the structure of the A-loop and C-helix. ALK's catalytic activity depends on the amino acid residues Lys 1150, Glu 1167, Asp 1249, and Asp 1270 (the K/E/D/D signature) in the ATP binding region [66, 67]. Biochemical studies have shown that auto-phosphorylation of the Y'RAS'YY motif in the A-loop can change how well ALK binds to substrates and how well it works as a catalyst [66,67]. Comparing the structures of wild

and mutant ALK TK, one can see key differences, such as an expanded opening in the former. Based on this finding, it appears that molecules with more flexible characteristics can more readily access the binding pocket of mutant ALK-TK (See Fig. 18).

The co-crystal structures of wild-type (pdb-4cmu) and mutant L1196M ALK (pdb-4clj) were analysed to determine the binding interactions between identified hit molecules and ALK TK wild-type and mutant strains. In addition to some novel clinically established antiviral agents (Ledipasvir, Elbasvir, Velpatasvir, and Daclatasvir), the QSAR-based virtual screening successfully identified some of the clinically established (FDA-approved) ALK TK inhibitors (ZINC000066166864 (Alectinib), ZINC000148723177 (Brigatinib), and ZINC000096272772 (Ceritinib)). All reported hit compounds were docked to ALK TK wild (pdb-4cum wt) and mutant strains (pdb-4clj mutant) to analyze the binding relationships. The QSAR-based virtual screening anticipated several FDA compounds' ALK-TK inhibitory activity. The docking scores and predicted activity (pIC_{50}) of the 12 hit molecules for the wild-type ALK TK and the

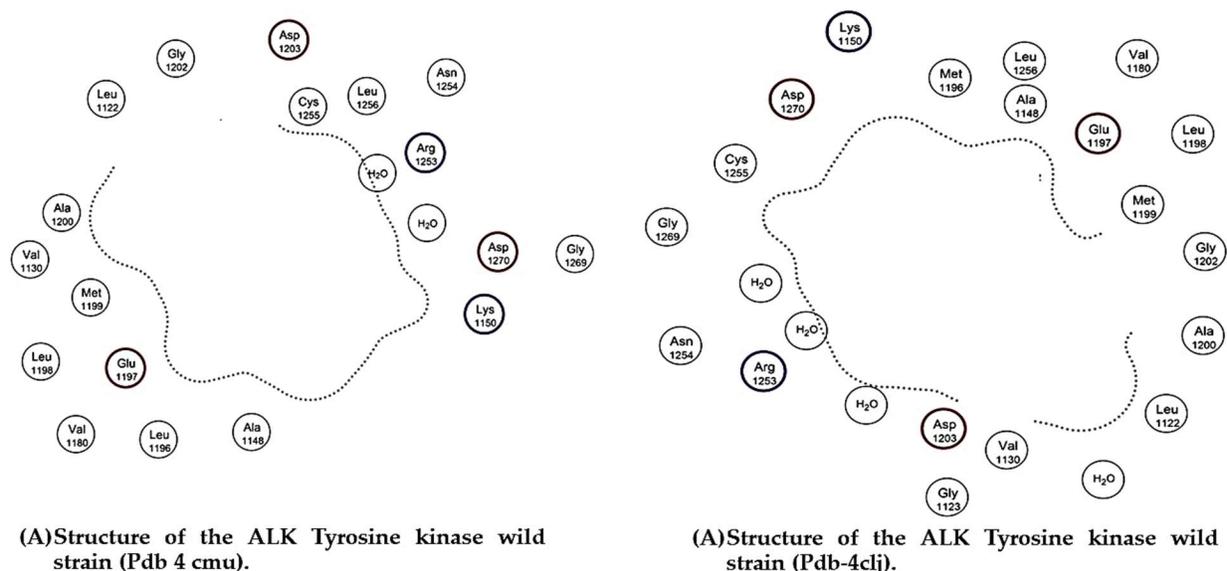


Fig. 18. Protein Structures of wild (A) and mutant (B) ALK Tyrosine kinase.

mutant L1196M TK are shown in Tables 1 and 2, respectively. The QSAR-based virtual screening showed that all antiviral drugs had a greater pIC_{50} against both wild and mutant ALK TK strains than the clinically proven ALK TK inhibitors (See Tables 3 and 4). ZINC000150338819 (Ledipasvir) had the highest docking score of $-10.57 \text{ kcal mol}^{-1}$ against the ALK TK (pdb-4cmu) wild strain and $-8.52 \text{ kcal mol}^{-1}$ against the ALK TK mutant strain (pdb-4clj, mutant) among the 12 discovered hit molecules, which was higher than ceritinib. Ledipasvir exhibited a higher docking score against the ALK TK (pdb-4cmu) wild strain than the clinically established ALK TK inhibitors: ZINC000066166864 (Alectinib, docking score of $-7.65 \text{ kcal mol}^{-1}$), ZINC000148723177 (Brigatinib, docking score of $-8.45 \text{ kcal mol}^{-1}$), and ZINC000096272772 (Ceritinib, docking score of $-8.17 \text{ kcal mol}^{-1}$) (Fig. 19 depicts the interaction of ZINC000150338819 (Ledipasvir) with the ALK TK wild and mutant strains).

Also, Ledipasvir's docking score against the ALK TK mutant strain (pdb-4clj) was $-8.52 \text{ kcal mol}^{-1}$, which was higher than Alectinib, Brigatinib, and Ceritinib. Also, the other hit molecules, ZINC000150588351 (Elbasvir), ZINC000203686879 (Velpatasvir), and ZINC000068204830 (Daclatasvir) had higher docking scores against both wild and mutant strains of the ALK TK than Alectinib, Brigatinib, and ceritinib, which are already used to stop the ALK TK

from working. Ledipasvir has better binding energy than ceritinib ($-38.64 \text{ kcal mol}^{-1}$ against wild-typed ALK TK and $-34.64 \text{ kcal mol}^{-1}$ against mutant ALK TK strain) and the other ALK TK inhibitors ($-38.64 \text{ kcal mol}^{-1}$ against wild-typed ALK TK and $-34.64 \text{ kcal mol}^{-1}$ against mutant ALK TK strain). Furthermore, the remaining antiviral drugs had increased binding affinity when compared to ceritinib and other ALK-TK inhibitors. This finding demonstrated that Ledipasvir has a greater affinity for the ALK TK in both wild-type and mutant strains.

In the ALK TK wild-type-Ledipasvir complex, we saw that the binding energy was $-3.56 \text{ kcal mol}^{-1}$ and that residue L1198 in the hinge segment was only 4.13 atoms away from Ledipasvir. This shows that residue L1198 and ledipasvir have a direct hydrophobic interaction.

The fact that the ALK TK mutant strain interacts with Ledipasvir more hydrophobically than the natural ALK TK shows that it has a high affinity. The residue Arg1253 formed conventional hydrogen bonds with Ledipasvir, while water formed hydrogen bonds with HOH2155, HOH2151, and HOH2151. In the active "DFG-in" shape of wild-type ALK TK, it takes $-0.33 \text{ kcal mol}^{-1}$ of energy for ledipasvir to interact with Met1199. In the mutant strain, it takes $-0.88 \text{ kcal mol}^{-1}$. This finding indicates the increased affinity of Ledipasvir for mutant ALK TK as compared to the wild strain, despite the fact that the contact distance between both

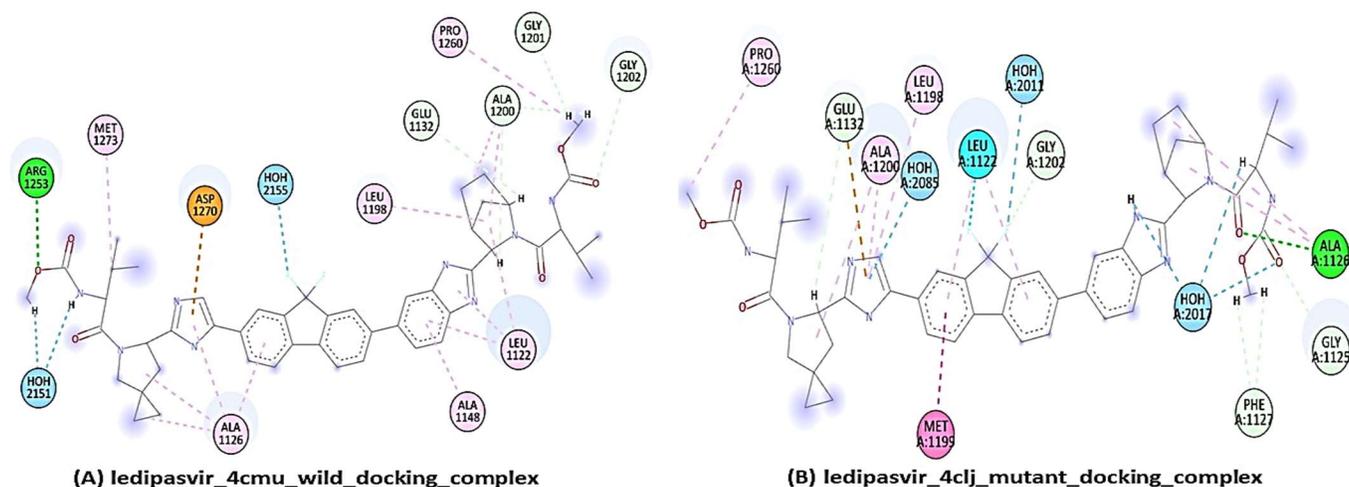


Fig. 19. Depiction of 2D interaction) of ZINC000150338819 (Ledipasvir) with ALK TK wild strain (A) and mutant strain (B).

Table 3. Depiction of the Docking Results of the 12 hits for ALK TK (pdb-4cmu, wild Strain)

SN	ZINC I.D. hit molecules	pIC_{50} by QSAR VS	Docking score (kcal mol ⁻¹)	RMSD Å	Binding free energy
1	ZINC000150338755 (Venetoclax)	9.36	-9.90	1.64	-38.63
2	ZINC000150338819 (Ledipasvir)	9.18	-10.57	1.54	-55.19
3	ZINC000150588351 (Elbasvir)	9.02	-9.93	2.37	-61.26
4	ZINC000066166864 (Alectinib)	8.67	-7.65	1.96	-34.99
5	ZINC000203686879 (Velpatasvir)	8.58	-10.08	2.10	-51.97
6	ZINC000148723177 (Brigatinib)	8.53	-8.45	1.78	-29.96
7	ZINC000096272772 (Ceritinib)	8.28	-8.17	1.58	-38.64
8	ZINC000028639340	8.25	-8.94	2.65	-50.07
9	ZINC000003938482 (Posaconazole)	8.25	-9.70	1.59	-49.49
10	ZINC000072316335 (Ribociclib)	8.20	-7.48	1.12	-34.62
11	ZINC000068204830 (Daclatasvir)	8.14	-10.10	2.15	-49.15
12	ZINC000003787097 (Besifloxacin)	8.03	-7.10	0.99	-32.95

strains was around 3.88. The terminal azaspiro ring forms a good hydrophobic interaction with the Ala1200, and the nearby methyl isobutyl carbamate showed two carbon-hydrogen bonding contacts with the Gly1201-Gly1202 of the hinge segment. This may be why Ledipasvir is more selective and effective than ceritinib, brigatinib, and alectinib. At the same time, the increased potency of Ledipasvir may be due in large part to the stronger interactions between the methyl group in isobutyl carbamate and Leu1198. Also, the mutant

ALK TK had two carbon-hydrogen bonding contacts between residue Gly1202 and both fluorine atoms as acceptors in the Ledipasvir, with a binding energy of -1.86 kcal mol⁻¹.

It was determined that the contact distance between Gly1202 and fluorine atoms was 4.14. No similar contact was discovered between Gly1202 and fluorine atoms in the wild ALK TK, which had a boat-shaped structure. Moreover, the orientation of fluorine atoms in the hydrophobic pocket

Table 4. Depiction of the Docking Results of 12 hits for ALK TK (pdb-4clj, Mutant Strain)

sn	ZINC I.D.	pIC ₅₀ by QSAR VS	Docking Score (kcal mol ⁻¹)	RMSD A0	Binding free energy
1	ZINC000150338755 (Venetoclax)	9.36	-8.87	2.28	-46.77
2	ZINC000150338819(Ledipasvir)	9.18	-8.52	2.09	-37.33
3	ZINC000150588351(Elbasvir)	9.02	-8.95	4.25	-39.68
4	ZINC000066166864 (Alectinib)	8.67	-7.70	1.47	-32.47
5	ZINC000203686879 (Velpatasvir)	8.58	-9.35	1.70	-31.60
6	ZINC000148723177(Brigatinib)	8.53	-8.33	1.70	-40.17
7	ZINC000096272772(Ceritinib)	8.28	-8.36	1.60	-34.64
8	ZINC000028639340	8.25	-7.07	2.19	-23.41
9	ZINC000003938482(Posaconazole)	8.25	-8.29	3.46	-44.94
10	ZINC000072316335(Ribociclib)	8.20	-7.93	3.70	-39.19
11	ZINC000068204830(Daclatasvir)	8.14	-8.43	4.77	-41.96
12	ZINC000003787097(Besifloxacin)	8.03	-7.27	1.89	-30.17

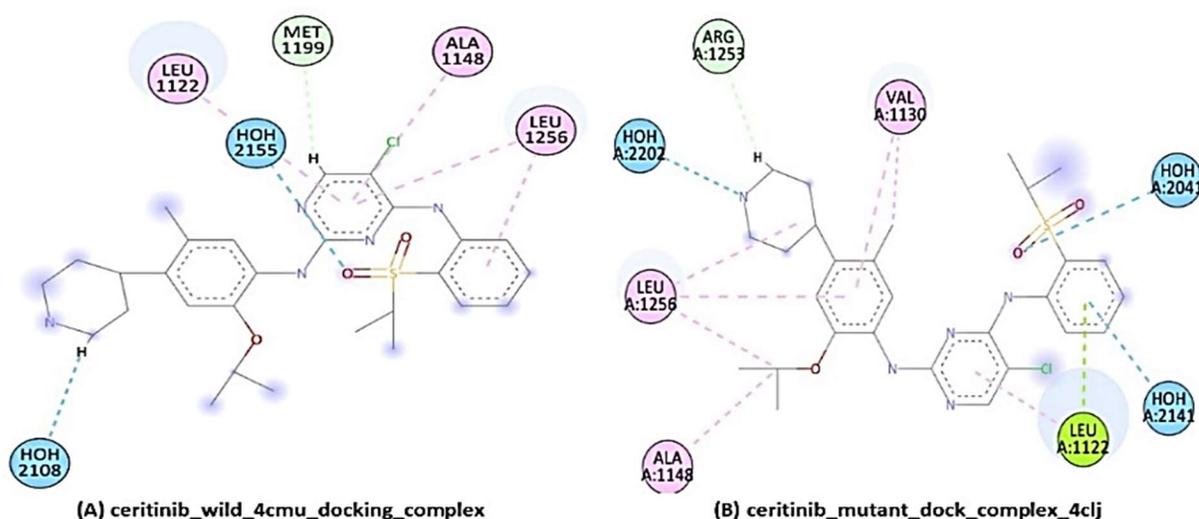


Fig. 20. Depiction of 2D interaction) of ZINC000096272772 (Ceritinib) with ALK TK wild strain (A) and mutant strain(B).

makes closer contact with the Asn1254 residue of the mutant ALK-TK, thereby enhancing intermolecular interactions and stabilizing its chair-shaped conformation. This is analogous to the orientation of fluorine atoms towards the Asn1254 residue of the pdb-4clj ligand. The differential in binding free energy between the two strains might be attributed to Ledipasvir's differing conformations towards normal and

mutant ALK TK. Furthermore, the wild ALK TK had a fluorine atom arrangement with a dihedral angle of 52.2 degrees, whereas the mutant ALK TK had an angle of 70 degrees, showing an 18-degree rotational alteration. This demonstrates that the mutant ALK TK undergoes significant conformational changes that alter its ability to bind Ledipasvir. The 2D interactions for ceritinib are revealed in Fig. 20.

Molecular Dynamics (MD) Simulations

Using MD simulations, we compared the effect of mutation on the conformational dynamics of wild-type, mutant, and apo-ALK-TK (noncomplex proteins) by measuring the root-mean-square deviation (RMSD) of the C atoms in the A-loop of the two variants. The apo-ALK tyrosine kinase (TK) exhibited overall stability over the 200-nanosecond simulation, with the exception of a notable fluctuation seen between 60 and 85 nanoseconds. Over this period, the root mean square deviation (RMSD) ranged from 2.0 to 2.2 angstroms, specifically in relation to the α -carbon atoms. Ultimately, the system reached convergence at an RMSD value of 2.27 angstroms. It was observed that the apo-ALK tyrosine kinase (TK) had greater fluctuations in both the back region, with a root-mean-square deviation (RMSD) of 2.3 Å, and the sidechain region, in comparison to the C- and backbone regions, which displayed an RMSD of 3.2 (see Fig. 1B). Further, Fig. 1 shows that after 40 ns throughout the simulation, the RMSD of the A-loop in the wild type converged to a value of 1.8. However, in the L1196 mutant, the A-loop RMSD increased to 2.4 at 40 ns and fluctuated between 40 and 50 ns. The value then reached 3.1 between 80 and 90 ns before decreasing again to 2.4 between 100 and 110 ns. Eventually, it climbed from 2.6 to 2.8 during the remainder of the MD simulation. This shows that the A-loop shape of the L1196 mutant is more flexible than that of the normal type (See Fig. 21).

We used MD simulations to measure the root-mean-square deviation (RMSD) of the C atoms in the A-loop of the wild and mutant states of ALK TK. This helped us Figure out how mutation affects the conformational dynamics of wild and mutant ALK TK. Figure 19 shows that after 40 ns throughout the simulation, the RMSD of the A-loop in the wild type converged to a value of 1.8. Conversely, the L1196 mutant showed an increase in the A-loop RMSD to 2.4 at 40 ns, with values ranging from 40 to 50 ns. The value then reached 3.1 between 80 and 90 ns before decreasing again to 2.4 between 100 and 110 ns. Eventually, it climbed from 2.6 to 2.8 during the remainder of the MD simulation. This shows that the A-loop shape of the L1196 mutant is more flexible than that of the normal type.

Figure 22 shows that the Rg plot of the C-backbone of ALK TK wild-type bound to Ledipasvir has the least

compactness because fluctuations go from 20.1 to 20.7, with a mean of 20.3 over 200 ns. On the other hand, when the ALK-TK mutant was bound to Ledipasvir, the C backbone showed a consistent gyration at 8.5 with few changes. When Rg decreases, it means that the protein-ligand complex is firmly bound. Mutant ALK-TK complexes coupled with Ledipasvir were much more stable than wild-type ALK-TK complexes associated with Ledipasvir. Throughout the simulation, the RMSF plot showed that each amino acid residue was in its most stable shape. This showed that the amino acid residues in the complex of ledipasvir-bound wild-type and mutant ALK TK were the least likely to change (See Fig. 23).

Global quality analysis of RMSD and Rg shows that ledipasvir has a big effect on the stability of both wild-type and mutant ALK TK targets after they have been bound in the binding cavities. Analysis of root mean square fluctuation (RMSF) plots using a time function of 200 nanoseconds showed that both the wild-type and mutant ALK-TK proteins had significant RMSF at certain residues. In contrast, the apo-ALK TK demonstrated comparatively lower levels of fluctuation. There are a few peaks that fluctuate from residue index 60 to 90 in the 200-ns simulation run of ALK TK depicted in Fig. 21, but they eventually settle in wild-type ALK TK. Comparing the ALK-TK wild-Ledipasvir complex to the docking data, it was shown that Arg1253 was engaged in traditional hydrogen bond formation with an RMSF of 0.602 (Fig. 23). In contrast, the residues Gly1202 and Leu1198 exhibited carbon-hydrogen bonding interactions with RMSF values of 0.78 and 0.68, respectively. Hydrogen bonding contributes to the stability of the complex, as shown by the smaller RMSF fluctuations of these residues compared to others. The ALK TK mutant Ledipasvir-bound complex exhibited two typical hydrogen bond forms with residues Asp1203 and Ala1126, with RMSFs of 0.63 and 1.56, respectively. To sum up, the RMSF plot analysis revealed that the ALK wild-type and mutant strains showed different binding patterns and fluctuations during the 200 ns simulation. All of these RMSFs are within the range that would be expected to stabilize the protein-ligand combination. So, RMSF plots suggest that the ALK-TK mutant's binding to Ledipasvir was stable across many simulated protein configurations.

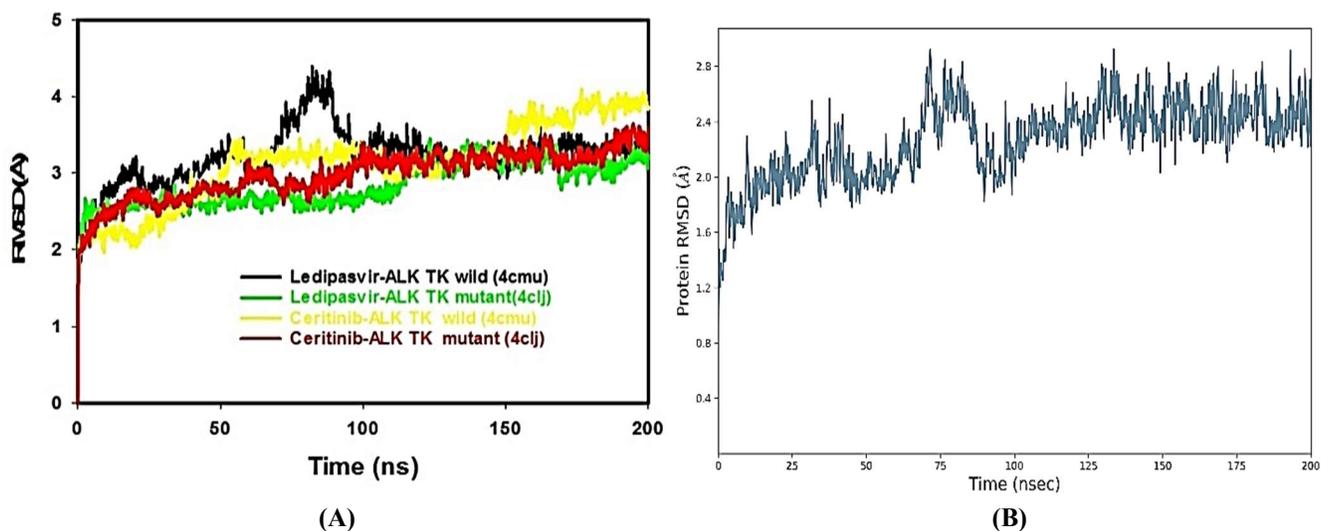


Fig. 21. (A) Root Mean Square Deviation (RMSD) analysis of MD simulation trajectories for (i) wild (4cmu) Ledipasvir-ALK TK, (ii) mutant (green, yellow) (4clj) Ledipasvir-ALK TK, (iii) wild (4cmu) Ceritinib-ALK TK, and (iv) mutant (red, 4clj) Ceritinib-ALK TK. (B) RMSD plot for the Apo-4cmu (non-complex) ALK TK protein.

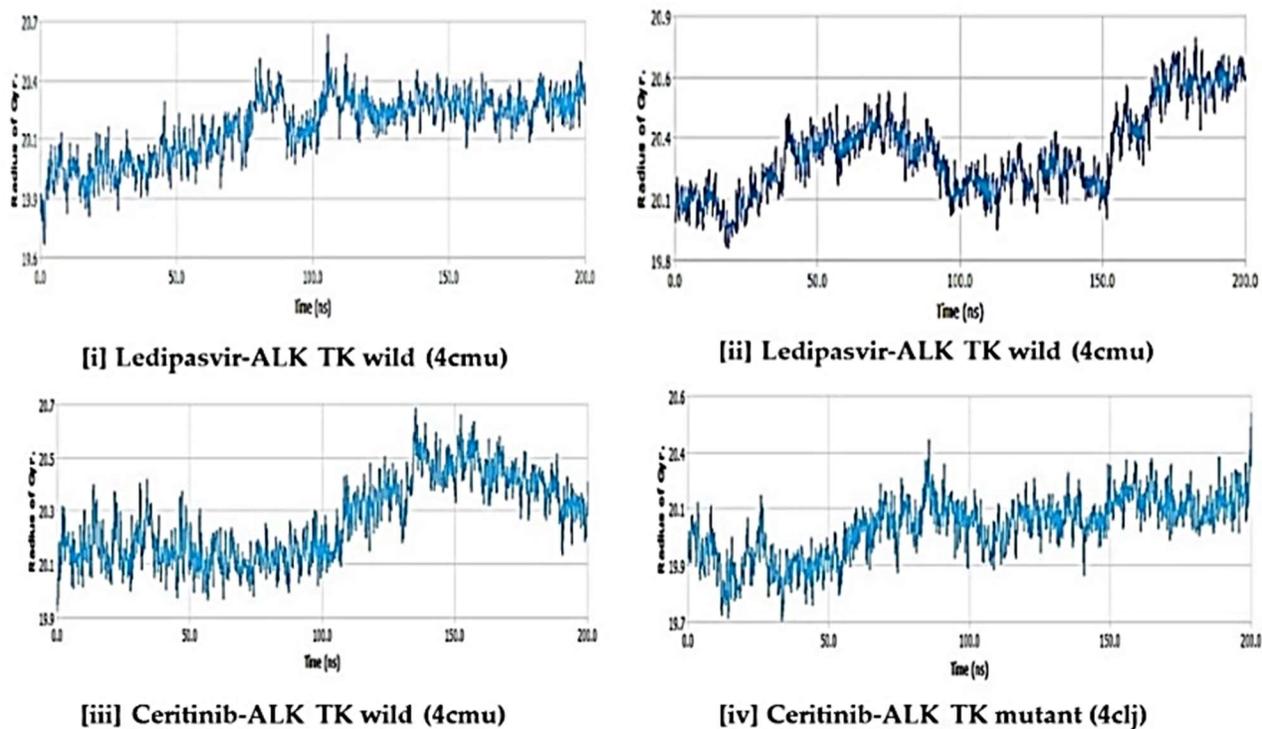


Fig. 22. Radius of gyration (Rg) trajectory study using MD simulations for [i] wild (4cmu) and [ii] mutant (4clj) Ledipasvir-ALK TK, [iii] wild (4cmu) and [iv] mutant (4clj) Ceritinib-ALK TK.

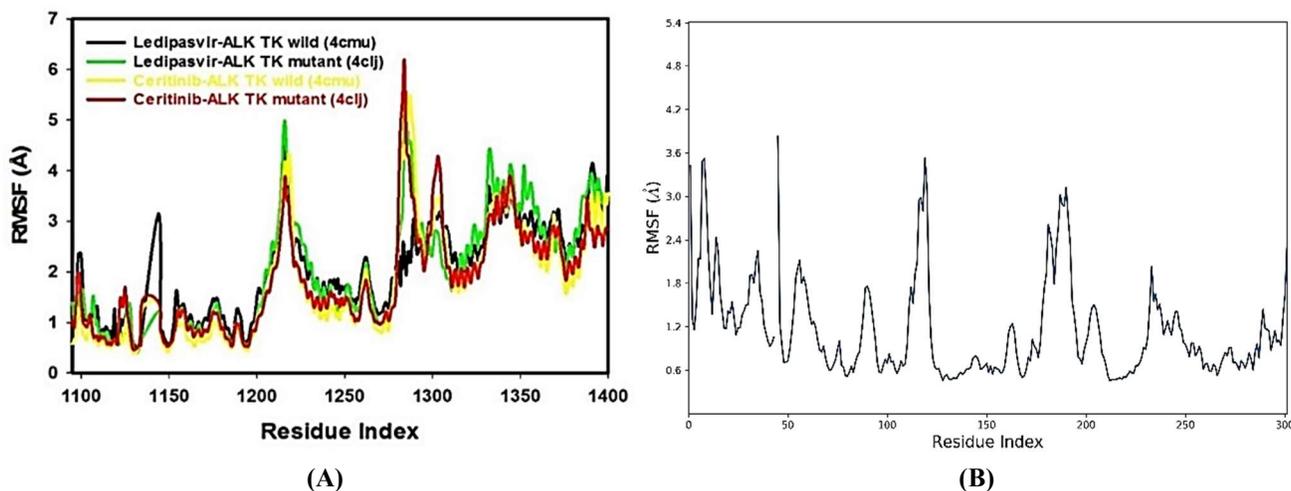


Fig. 23. (A) Root Mean Square Fluctuations (RMSF) in MD simulations of [i] wild (4cmu) and [ii] mutant (4clj) Ledipasvir-ALK TK, [iii] wild (4cmu) and [iv] mutant (4clj) Ceritinib-ALK TK. (B) Root Mean Square Fluctuations (RMSF) in MD simulations of Apo-ALK TK (non-complexed ALK TK).

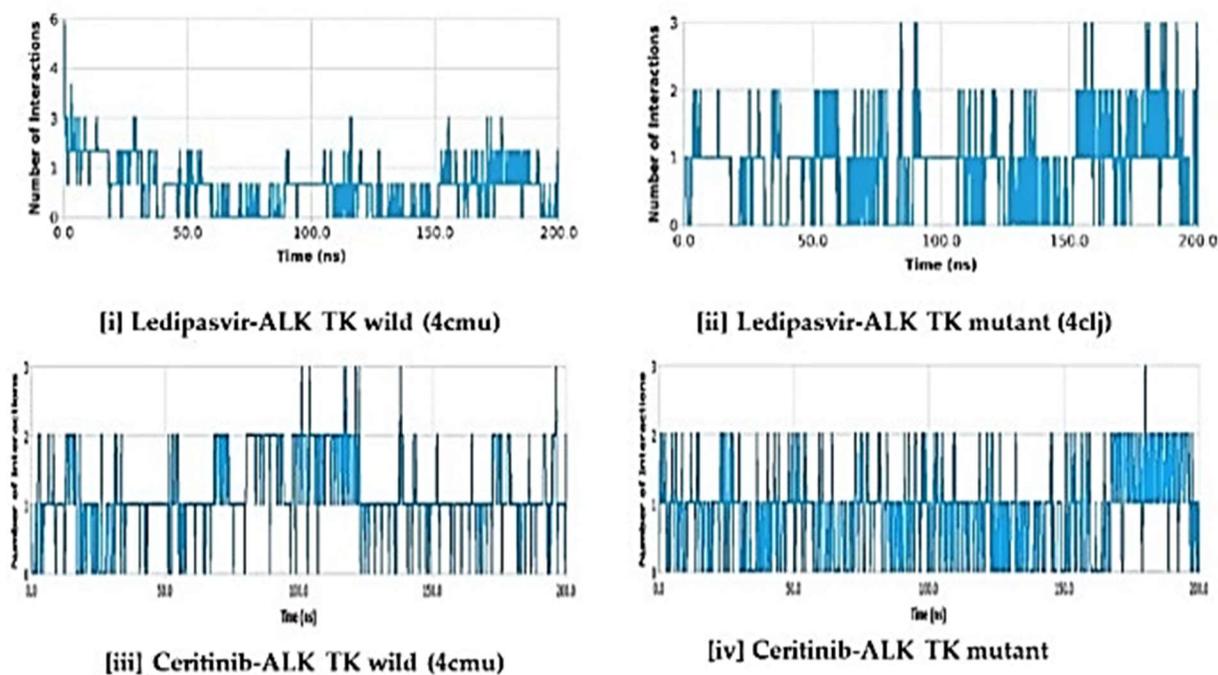


Fig. 24. Hydrogen bond (H-Bond) trajectory study using MD simulations for [i] wild (4cmu) and [ii] mutant (4clj) Ledipasvir-ALK TK, [iii] wild (4cmu) and [iv] mutant (4clj) Ceritinib-ALK TK, and [v] mutant (4clj).

During the 200 ns simulation, average hydrogen bonds between ledipasvir and the other proteins were also kept track of Fig. 24. From 0 to 200 ns, the MD simulation of Ledipasvir

showed that hydrogen bonds formed with both wild-type and mutant ALK TK. Similar numbers of hydrogen bonds were established in the wild-type Ledipasvir-ALK TK complex

(Fig. 24 iv). In addition, docking with wild-type and mutant strains of ALK and TK verified the tendency for two to three hydrogen bonds to be formed after 200 ns of molecular dynamics. Hydrogen bonding between wild-type and mutant ALK-TK strains and Ledipasvir enhanced binding and allowed the conformation to form a more stable complex during simulation.

A stepwise trajectory analysis performed at 50 ns intervals revealed that both wild-type and mutant ALK tyrosine kinases moved with respect to the 0 ns structure after treatment with ledipasvir and ceritinib (Fig. 25). Conformational stability and convergence in the ligand ledipasvir have been observed through structural angular movement in the final frame. However, mutants linked to ledipasvir exhibit behavior consistent with structural stability.

Molecular Mechanics Generalised Born Surface Area (MMGBSA)

For the most part, the MMGBSA technique is used to determine the degree of ligand-protein binding. Ledipasvir-ALK TK wild-type (4cmu) and mutant (4clj) complexes' binding free energies were determined, as were the binding free energies of ceritinib-ALK TK wild-type (4cmu) and mutant (4clj) complexes and the influence of other non-bonded interaction energies. The binding energies of the Ledipasvir ligand to the wild-type (4cmu) and mutant (4clj) ALK TK complexes were -47.77 and -61.68 kcal mol⁻¹, respectively. Table 5 displays the average binding energies of the wild-type (4cmu) and mutant (4clj) ceritinib-ALK TK complexes, which are -58.49 and -51.31 kcal mol⁻¹,

respectively. Ledipasvir showed a much higher binding affinity for mutant ALK TK than for natural ALK TK, in contrast to ceritinib. Gbind is controlled by a variety of interactions that don't include a covalent bond, including GbindCoulomb, GbindCovalent, GbindHbond, GbindLipo, GbindSolvGB, and GbindvdW. Energy contributions from GbindLipo and GbindCoulomb were less than those from GbindvdW, but all three energies played a role in the binding affinity between wild-type and mutant ALK TK and ceritinib. An essential function for GbindvdW in drug-receptor interactions has been uncovered. Contrarily, the GbindSolvGB and Gbind covalent energies had the smallest impact on the overall mean binding energies.

Both wild-type and mutant complexes of ledipasvir and ceritinib-ALK TK were found to form stable hydrogen bonds with amino acid residues. This was shown by their GbindHbond interaction values. Each molecule was shown to have a negative energy contribution by both GbindSolvGB and GbindCovalent, suggesting that they resisted binding. Inside the binding pockets of wild-type and mutant ALK-TK proteins (200 ns), the shape of ledipasvir and ceritinib has changed dramatically, going from curved to straight. Because of these changes in the conformation, there is more contact between the binding pocket and the residues, which increases stability and binding energy.

Since MM-GBSA predictions were made using MD simulation trajectories, the docking data used to figure out the binding energy agreed with them. Ledipasvir and ceritinib also shifted from their positions in the 0 ns trajectory to their positions in the 200 ns trajectory in the last frame of the MMGBSA, indicating that they are in a better position to fit into the protein's binding cavity (See Fig. 26).

Table 5. Presentation of the MMGBSA Results for the Ledipasvir and Ceritinib Wild and Mutant ALK TK Strains

Energies (kcal mol ⁻¹)*	Ledipasvir-ALK-TK wild (4cmu)	Ledipasvir-ALK-TK mutant (4clj)	Ceritinib-ALK-TK wild (4cmu)	Ceritinib-ALK TK-mutant (4clj)
ΔGbind	-47.77 ± 6.95	-61.68 ± 8.16	-58.49 ± 4.42	-51.31 ± 6.29
ΔGbindLipo	-17.98 ± 1.95	-22.92 ± 2.55	-19.20 ± 1.46	-17.51 ± 2.01
ΔGbindvdW	-52.58 ± 6.71	-66.37 ± 5.71	-54.49 ± 3.75	-54.21 ± 5.77
ΔGbindCoulomb	-17.51 ± 15.17	9.10 ± 13.51	10.96 ± 10.63	11.57 ± 3.17
ΔGbindHbond	-0.85 ± 0.47	-0.93 ± 0.53	-0.95 ± 0.50	-0.60 ± 0.35
ΔGbindSolvGB	9.98 ± 13.37	14.32 ± 12.01	-1.40 ± 9.36	5.82 ± 3.12
ΔGbindCovalent	6.86 ± 3.59	5.59 ± 2.66	6.59 ± 3.12	3.64 ± 2.17

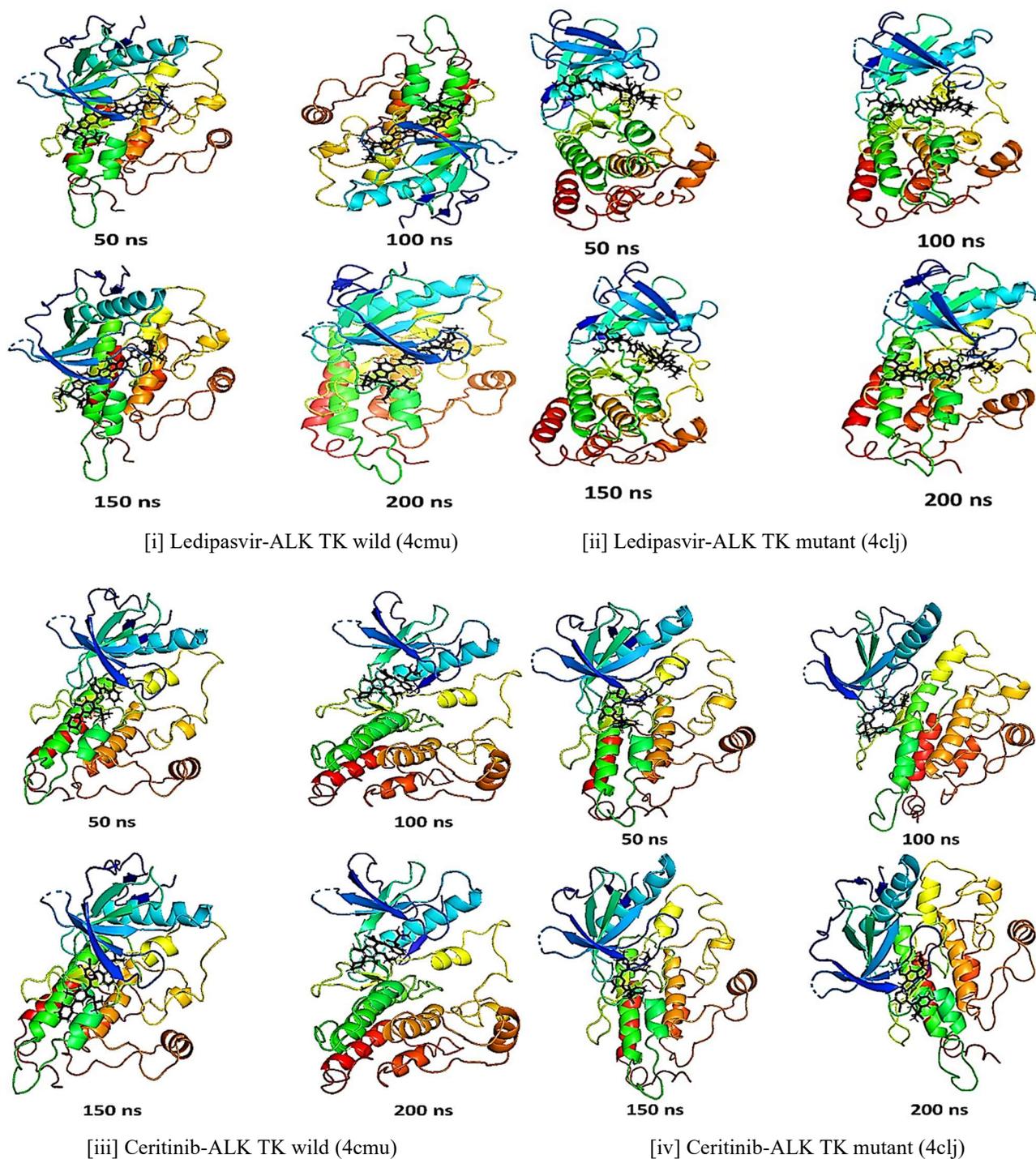


Fig. 25. After simulating the protein and ligand for 200 ns, a step-by-step trajectory analysis shows the configuration of the protein and ligand every 25 ns for [i] Ledipasvir-ALK TK wild (4cmu) and [ii] Ledipasvir-ALK TK mutant(4clj), [iii] Ceritinib-ALK TK wild (4cmu), and [iv] Ceritinib-ALK TK mutant (4clj).

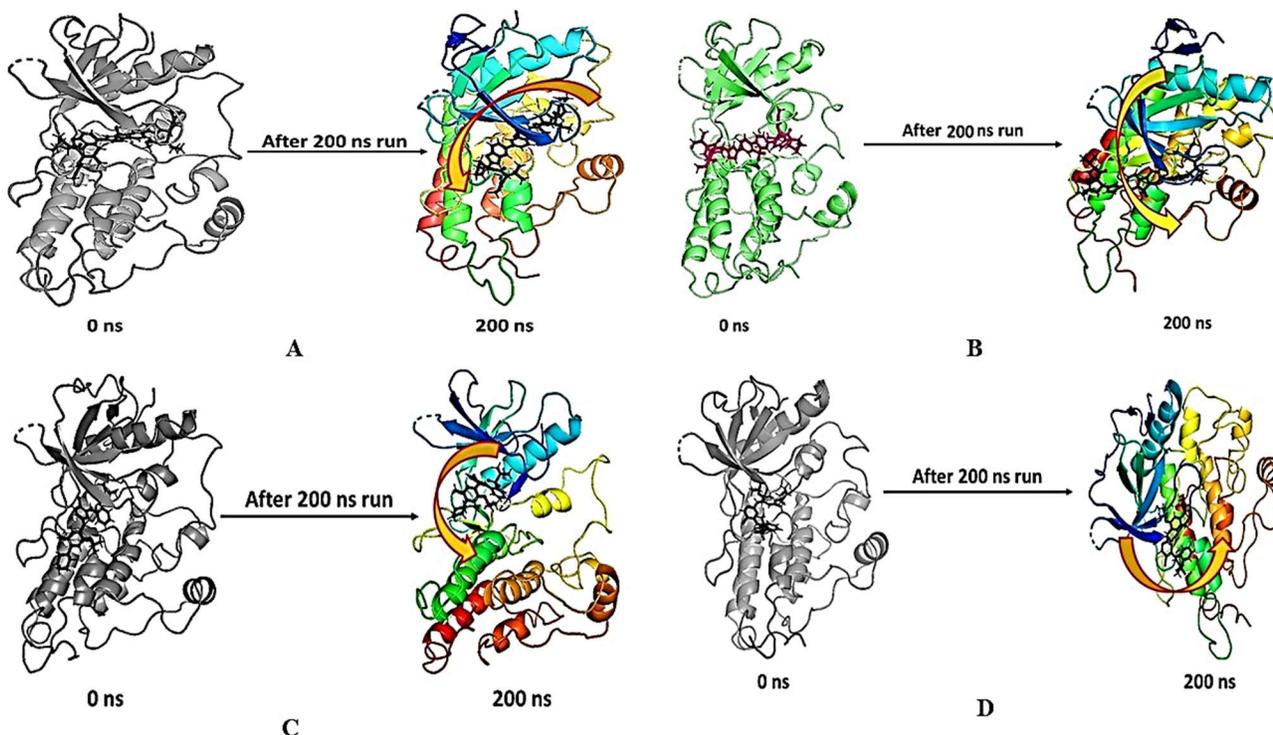


Fig. 26. A. Ledipasvir-ALK TK wild (4cmu); B. Ledipasvir-ALK TK mutant (4clj); C. Ceritinib-ALK TK wild (4cmu); D. Ceritinib-ALK TK mutant (4clj) demonstrated conformational alterations in the MMGBSA trajectory (0 ns, before simulation and 200 ns, after simulation). Ledipasvir (green) and Ceritinib (red) are shown at their respective binding sites for ALK wild type (pdb-4cmu) and ALK mutant type (pdb-4clj) strains, with their relative movement and orientation indicated by arrows.

MTT Assay (*In-Vitro* Cell Line Study)

We have used A549 cell lines for *in vitro* investigations to determine the anticancer potential of our most active hits, Ledipasvir and posaconazole. In the current investigation, ledipasvir and posaconazole were chosen for *in vitro* cell line studies because they are both located in or very close to the applicability domain in the QSAR model.

According to the results of our study (Table S5 I Supplementary Material) of *in vitro* anticancer activity on the A549 lung cancer cell line, we found that the percentage of inhibition for our acquired hit, which is Ledipasvir, was somewhat greater than that of the standard reference molecule (See Fig. 27), Ceritinib (which is a kinase inhibitor). The results are in agreement with the computational models, suggesting that the identical mechanism that suppressed the A549 lung cancer cell line may be responsible for suppressing the ALK TK target. On

the other hand, we feel that more enzyme examinations are necessary in order to get further insights.

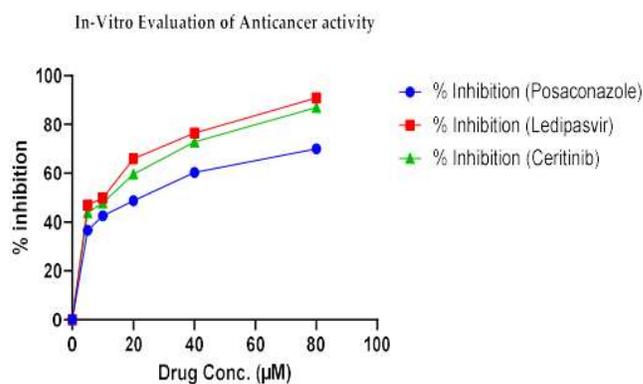


Fig. 27. Depiction of *in-vitro* evaluation of anticancer activity.

CONCLUSIONS

For the relatively large data set of 1328 compounds with TK-ALK inhibitory activity (IC_{50}), a QSAR model was made using six descriptors. The robust and predictive QSAR model fulfills all the required threshold values, such as $R^2 = 0.79$, $Q^2_{LOO} = 0.78$, $Q^2_{LMO} = 0.78$, $R^2_{ex} = 0.77$, $CCC_{ex} = 0.87$, etc. Several hidden pharmacophoric properties, such as the sum of partial charges of aromatic carbons, sp^3 -hybridized carbons within 4 bonds of non-cyclic carbons, and planer nitrogen within 6 bonds of the ring carbon atom, have been effectively detected using the created model. Based on the findings of the Quantitative Structure-Activity Relationship (QSAR) analysis, it was revealed that the presence of the aromatic or ring carbon atom, the ring nitrogen atom, or the non-ring nitrogen atom had a substantial impact on the determination of the inhibitory activity of ALK TK. X-ray-resolved structures validated the combination of disclosed and concealed structural characteristics, which was similarly seen in reported TK ALK inhibitors such as crizotinib, brigatinib, etc. In addition to this, a blend of QSAR-based virtual screening and drug repositioning recognized a ledipasvir as an FDA-approved molecule that has been tested in clinical trials and has an IC_{50} of 0.65 nM ($pIC_{50} = 9.18$ M). The molecular docking study of ledipasvir displayed a docking score of -10.57 kcal mol $^{-1}$ against the ALK TK wild strain (pdb-4cmu) and a docking score of -8.5286303 kcal/mol against the ALK TK mutant strain (pdb-4clj, mutant) which was higher than ceritinib. Based on docking analysis, ledipasvir has a higher binding energy than ceritinib (-38.64 kcal mol $^{-1}$ against wild ALK TK and -34.64 kcal/mol against the mutant ALK TK strain) and other drugs that block ALK TK. Thus, the ledipasvir exhibited a significant affinity, as the ALK-TK mutant strain made more hydrophobic contact with it than the wild strain. The MD simulation and MMGBSA analysis demonstrated that the drug-receptor complex was stable over 200 ns, and the binding energy was very high. The mutant ALK TK-ledipasvir complex was substantially more stable than the wild ALK TK complex. Thus, molecular dynamics (MD) simulation trajectories and binding energy docking data supported MM-GBSA predictions. Furthermore, the MTT assay demonstrated that the ledipasvir had slightly higher anticancer activity against the lung cancer cell line A549 as compared to the clinically

established anticancer agent; ceritinib. As a result, the findings from the in-silico study corroborate the in vitro anticancer efficacy. Therefore, the present findings might be very beneficial in developing a brand-new ALK-TK as a chemotherapeutic drug.

ACKNOWLEDGMENTS

Rahul D. Jawarkar, the author, is grateful to Dr. R. K. Jain, Dean of Oriental University, Indore, for giving extended facilities during the whole of the study endeavor. The authors express their gratitude to the Department of Pharmaceutical Sciences at Birla Institute of Technology, Mesraa, Ranchi, for providing the use of computing tools.

REFERENCES

- [1] Morris, S. W.; Kirstein, M. N.; Valentine, M. B.; Dittmer, K.; Shapiro, D. N.; Look, A. T.; Saltman, D. L., Sequence Correction. *Sci.* **1995**, *267* (5196), 316-317. DOI: 10.1126/science.267.5196.316-b.
- [2] Morris, S. W.; Naeve, C.; Mathew, P.; James, P. L.; Kirstein, M. N.; Cui, X.; Witte, D. P., ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). *Oncogene* **1997**, *14* (18), 2175-2188. DOI: 10.1038/sj.onc.1201062.
- [3] Iwahara, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Ratzkin, B.; Yamamoto, T., Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. *Oncogene* **1997**, *14* (4), 439-449. DOI: 10.1038/sj.onc.1200849.
- [4] Lorén, C. E.; Scully, A.; Grabbe, C.; Edeen, P. T.; Thomas, J.; McKeown, M.; Hunter, T.; Palmer, R. H., Identification and characterization of DAlk: a novel *Drosophila melanogaster* RTK which drives ERK activation in vivo. *Genes to Cells* **2001**, *6* (6), 531-544. DOI: 10.1046/j.1365-2443.2001.00440.x.
- [5] Palmer, Ruth H.; Vernersson, E.; Grabbe, C.; Hallberg, B., Anaplastic lymphoma kinase: signalling in development and disease. *Biochem J.* **2009**, *420* (3), 345-361. DOI: 10.1042/bj20090387.

- [6] Kadomatsu, K.; Muramatsu, T., Midkine and pleiotrophin in neural development and cancer. *Cancer Lett.* **2004**, *204* (2), 127-143. DOI: 10.1016/s0304-3835(03)00450-6.
- [7] Muramatsu, T., Midkine and Pleiotrophin: Two Related Proteins Involved in Development, Survival, Inflammation and Tumorigenesis. *J. Biochem.* **2002**, *132* (3), 359-371. DOI: 10.1093/oxfordjournals.jbchem.a003231.
- [8] Perez-Pinera, P.; Zhang, W.; Chang, Y.; Vega, J. A.; Deuel, T. F., Anaplastic Lymphoma Kinase Is Activated Through the Pleiotrophin/Receptor Protein-tyrosine Phosphatase β/ζ Signaling Pathway. *J Bio Chem.* **2007**, *282* (39), 28683-28690. DOI: 10.1074/jbc.M704505200.
- [9] Mathivet, T.; Mazot, P.; Vigny, M., In contrast to agonist monoclonal antibodies, both C-terminal truncated form and full length form of Pleiotrophin failed to activate vertebrate ALK (anaplastic lymphoma kinase)? *Cell Signal.* **2007**, *19* (12), 2434-2443. DOI: 10.1016/j.cellsig.2007.07.011.
- [10] Moog-Lutz, C.; Degoutin, J.; Gouzi, J. Y.; Frobert, Y.; Carvalho, N. B.-d.; Bureau, J.; Créminon, C.; Vigny, M., Activation and Inhibition of Anaplastic Lymphoma Kinase Receptor Tyrosine Kinase by Monoclonal Antibodies and Absence of Agonist Activity of Pleiotrophin. *J. Bio. Chem.* **2005**, *280* (28), 26039-26048. DOI: 10.1074/jbc.M501972200.
- [11] Moteji, A.; Fujimoto, J.; Kotani, M.; Sakuraba, H.; Yamamoto, T., ALK receptor tyrosine kinase promotes cell growth and neurite outgrowth. *J. Cell Sci.* **2004**, *117* (15), 3319-3329. DOI: 10.1242/jcs.01183.
- [12] Mourali, J.; Bénard, A.; Lourenço, F. C.; Monnet, C.; Greenland, C.; Moog-Lutz, C.; Racaud-Sultan, C.; Gonzalez-Dunia, D.; Vigny, M.; Mehlen, P.; *et al.* Anaplastic Lymphoma Kinase Is a Dependence Receptor Whose Proapoptotic Functions Are Activated by Caspase Cleavage. *Mol. Cell Bio.* **2023**, *26* (16), 6209-6222. DOI: 10.1128/mcb.01515-05.
- [13] Sakamoto, H.; Tsukaguchi, T.; Hiroshima, S.; Kodama, T.; Kobayashi, T.; Fukami, Takaaki A.; Oikawa, N.; Tsukuda, T.; Ishii, N.; Aoki, Y., CH5424802, a Selective ALK Inhibitor Capable of Blocking the Resistant Gatekeeper Mutant. *Cancer Cell* **2011**, *19* (5), 679-690. DOI: 10.1016/j.ccr.2011.04.004.
- [14] Guan, J.; Tucker, E. R.; Wan, H.; Chand, D.; Danielson, L. S.; Ruuth, K.; El Wakil, A.; Witek, B.; Jamin, Y.; Umapathy, G.; *et al.* The ALK inhibitor PF-06463922 is effective as a single agent in neuroblastoma driven by expression of ALK and MYCN. *Dis. Model Mech.* **2016**. DOI: 10.1242/dmm.024448.
- [15] Lu, J.; Guan, S.; Zhao, Y.; Yu, Y.; Woodfield, S. E.; Zhang, H.; Yang, K. L.; Bieerkehazhi, S.; Qi, L.; Li, X.; *et al.* The second-generation ALK inhibitor alectinib effectively induces apoptosis in human neuroblastoma cells and inhibits tumor growth in a TH-MYCN transgenic neuroblastoma mouse model. *Cancer Lett.* **2017**, *400*, 61-68. DOI: 10.1016/j.canlet.2017.04.022.
- [16] Gettinger, S. N.; Bazhenova, L. A.; Langer, C. J.; Salgia, R.; Gold, K. A.; Rosell, R.; Shaw, A. T.; Weiss, G. J.; Tugnait, M.; Narasimhan, N. I.; *et al.* Activity and safety of brigatinib in ALK-rearranged non-small-cell lung cancer and other malignancies: a single-arm, open-label, phase 1/2 trial. *Lancet Oncol.* **2016**, *17* (12), 1683-1696. DOI: 10.1016/s1470-2045(16)30392-8.
- [17] Infarinato, N. R.; Park, J. H.; Krytska, K.; Ryles, H. T.; Sano, R.; Szigety, K. M.; Li, Y.; Zou, H. Y.; Lee, N. V.; Smeal, T.; *et al.* The ALK/ROS1 Inhibitor PF-06463922 Overcomes Primary Resistance to Crizotinib in ALK-Driven Neuroblastoma. *Cancer Discov.* **2016**, *6* (1), 96-107. DOI: 10.1158/2159-8290.Cd-15-1056.
- [18] Zhang, S.; Anjum, R.; Squillace, R.; Nadworny, S.; Zhou, T.; Keats, J.; Ning, Y.; Wardwell, S. D.; Miller, D.; Song, Y.; *et al.* The Potent ALK Inhibitor Brigatinib (AP26113) Overcomes Mechanisms of Resistance to First- and Second-Generation ALK Inhibitors in Preclinical Models. *Clin. Cancer Res.* **2016**, *22* (22), 5527-5538. DOI: 10.1158/1078-0432.Ccr-16-0569.
- [19] Carneiro, B. A.; Pamarthy, S.; Shah, A. N.; Sagar, V.; Unno, K.; Han, H.; Yang, X. J.; Costa, R. B.; Nagy, R. J.; Lanman, R. B.; *et al.* Anaplastic Lymphoma Kinase Mutation (ALK F1174C) in Small Cell Carcinoma of the Prostate and Molecular Response to Alectinib. *Clin Cancer Res.* **2018**, *24* (12), 2732-2739. DOI: 10.1158/1078-0432.Ccr-18-0332.
- [20] Friboulet, L.; Li, N.; Katayama, R.; Lee, C. C.; Gainor, J. F.; Crystal, A. S.; Michellys, P. -Y.; Awad, M. M.; Yanagitani, N.; Kim, S.; *et al.* The ALK Inhibitor

- Ceritinib Overcomes Crizotinib Resistance in Non-Small Cell Lung Cancer. *Cancer Discov.* **2014**, *4* (6), 662-673. DOI: 10.1158/2159-8290.Cd-13-0846.
- [21] Choi, Y. L.; Soda, M.; Yamashita, Y.; Ueno, T.; Takashima, J.; Nakajima, T.; Yatabe, Y.; Takeuchi, K.; Hamada, T.; Haruta, H.; *et al.* EML4-ALK Mutations in Lung Cancer That Confer Resistance to ALK Inhibitors. *N Engl J Med.* **2010**, *363* (18), 1734-1739. DOI: 10.1056/NEJMoa1007478.
- [22] Heuckmann, J. M.; Hölzel, M.; Sos, M. L.; Heynck, S.; Balke-Want, H.; Koker, M.; Peifer, M.; Weiss, J.; Lovly, C. M.; Grütter, C.; *et al.* ALK Mutations Conferring Differential Resistance to Structurally Diverse ALK Inhibitors. *Clin Cancer Res.* **2011**, *17* (23), 7394-7401. DOI: 10.1158/1078-0432.Ccr-11-1648.
- [23] Doebele, R. C.; Pilling, A. B.; Aisner, D. L.; Kutateladze, T. G.; Le, A. T.; Weickhardt, A. J.; Kondo, K. L.; Linderman, D. J.; Heasley, L. E.; Franklin, W. A.; *et al.* Mechanisms of Resistance to Crizotinib in Patients with ALK Gene Rearranged Non-Small Cell Lung Cancer. *Clin Cancer Res.* **2012**, *18* (5), 1472-1482. DOI: 10.1158/1078-0432.Ccr-11-2906.
- [24] Katayama, R.; Shaw, A. T.; Khan, T. M.; Mino-Kenudson, M.; Solomon, B. J.; Halmos, B.; Jessop, N. A.; Wain, J. C.; Yeo, A. T.; Benes, C.; *et al.* Mechanisms of Acquired Crizotinib Resistance in ALK-Rearranged Lung Cancers. *Sci Transl Med.* **2012**, *4* (120). DOI: 10.1126/scitranslmed.3003316.
- [25] Shaw, A. T.; Kim, D.-W.; Mehra, R.; Tan, D. S. W.; Felip, E.; Chow, L. Q. M.; Camidge, D. R.; Vansteenkiste, J.; Sharma, S.; De Pas, T.; *et al.* Ceritinib in ALK-Rearranged Non-Small-Cell Lung Cancer. *N Engl J Med.* **2014**, *370* (13), 1189-1197. DOI: 10.1056/NEJMoa1311107.
- [26] Katayama, R.; Friboulet, L.; Koike, S.; Lockerman, E. L.; Khan, T. M.; Gainor, J. F.; Iafrate, A. J.; Takeuchi, K.; Taiji, M.; Okuno, Y.; *et al.* Two Novel ALK Mutations Mediate Acquired Resistance to the Next-Generation ALK Inhibitor Alectinib. *Clin Cancer Res.* **2014**, *20* (22), 5686-5696. DOI: 10.1158/1078-0432.Ccr-14-1511.
- [27] Ou, S.-H. I.; Ahn, J. S.; De Petris, L.; Govindan, R.; Yang, J. C.-H.; Hughes, B.; Lena, H.; Moro-Sibilot, D.; Bearz, A.; Ramirez, S. V.; *et al.* Alectinib in Crizotinib-Refractory ALK-Rearranged Non-Small-Cell Lung Cancer: A Phase II Global Study. *J Clin Oncol.* **2016**, *34* (7), 661-668. DOI: 10.1200/jco.2015.63.9443.
- [28] Rotow, J.; Bivona, T. G., Understanding and targeting resistance mechanisms in NSCLC. *Nat. Rev. Cancer.* **2017**, *17* (11), 637-658. DOI: 10.1038/nrc.2017.84.
- [29] Gainor, J. F.; Dardaei, L.; Yoda, S.; Friboulet, L.; Leshchiner, I.; Katayama, R.; Dagogo-Jack, I.; Gadgeel, S.; Schultz, K.; Singh, M.; *et al.* Molecular Mechanisms of Resistance to First- and Second-Generation ALK Inhibitors in ALK-Rearranged Lung Cancer. *Cancer Discov.* **2016**, *6* (10), 1118-1133. DOI: 10.1158/2159-8290.Cd-16-0596.
- [30] Holla, V. R.; Elamin, Y. Y.; Bailey, A. M.; Johnson, A. M.; Litztenburger, B. C.; Khotskaya, Y. B.; Sanchez, N. S.; Zeng, J.; Shufean, M. A.; Shaw, K. R.; *et al.* ALK: a tyrosine kinase target for cancer therapy. *Mol. Case Stud.* **2017**, *3* (1). DOI: 10.1101/mcs.a001115.
- [31] Rashda, S.; Gerber, D. E., A crowded, but still varied, space: brigatinib in anaplastic lymphoma kinase-rearranged non-small cell lung cancer. *Transl. Cancer Res.* **2017**, *6* (S1), S78-S82. DOI: 10.21037/ter.2017.02.12.
- [32] Engelman, J. A.; Jänne, P. A., Mechanisms of Acquired Resistance to Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. *Clin. Cancer Res.* **2008**, *14* (10), 2895-2899. DOI: 10.1158/1078-0432.Ccr-07-2248.
- [33] Kwak, E. L.; Bang, Y. -J.; Camidge, D. R.; Shaw, A. T.; Solomon, B.; Maki, R. G.; Ou, S. -H. I.; Dezube, B. J.; Jänne, P. A.; Costa, D. B.; *et al.* Anaplastic Lymphoma Kinase Inhibition in Non-Small-Cell Lung Cancer. *N Engl. J. Med.* **2010**, *363* (18), 1693-1703. DOI: 10.1056/NEJMoa1006448.
- [34] Neves, B. J.; Braga, R. C.; Melo-Filho, C. C.; Moreira-Filho, J. T.; Muratov, E. N.; Andrade, C. H., QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Front Pharmacol.* **2018**, *9*. DOI: 10.3389/fphar.2018.01275.
- [35] Pushpakom, S.; Iorio, F.; Eyers, P. A.; Escott, K. J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; *et al.* Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug*

- Discov.* **2018**, *18* (1), 41-58. DOI: 10.1038/nrd.2018.168.
- [36] Dearden, J. C.; Cronin, M. T. D.; Kaiser, K. L. E., How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). *SAR and QSAR in Environ. Res.* **2009**, *20* (3-4), 241-266. DOI: 10.1080/10629360902949567.
- [37] Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; *et al.* QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57* (12), 4977-5010. DOI: 10.1021/jm4004285.
- [38] Huang, J.; Fan, X., Why QSAR Fails: An Empirical Evaluation Using Conventional Computational Approach. *Mol. Pharmaceutics.* **2011**, *8* (2), 600-608. DOI: 10.1021/mp100423u.
- [39] Muratov, E. N.; Bajorath, J.; Sheridan, R. P.; Tetko, I. V.; Filimonov, D.; Poroikov, V.; Oprea, T. I.; Baskin, I. I.; Varnek, A.; Roitberg, A.; *et al.* QSAR without borders. *Chem. Soc. Rev.* **2020**, *49* (11), 3525-3564. DOI: 10.1039/d0cs00098a.
- [40] Golbraikh, A.; Muratov, E.; Fourches, D.; Tropsha, A., Data Set Modelability by QSAR. *J. Chem. Inf. Model.* **2014**, *54* (1), 1-4. DOI: 10.1021/ci400572x.
- [41] Martin, T. M.; Harten, P.; Young, D. M.; Muratov, E. N.; Golbraikh, A.; Zhu, H.; Tropsha, A., Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling? *J. Chem. Inf. Model.* **2012**, *52* (10), 2570-2578. DOI: 10.1021/ci300338w.
- [42] Fourches, D.; Muratov, E.; Tropsha, A., Trust, But Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. *J. Chem. Inf. Model.* **2010**, *50* (7), 1189-1204. DOI: 10.1021/ci100176x.
- [43] O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An open chemical toolbox. *J. Cheminformatics.* **2011**, *3* (1). DOI: 10.1186/1758-2946-3-33.
- [44] Tetko, I. V.; Sushko, I.; Pandey, A. K.; Zhu, H.; Tropsha, A.; Papa, E.; Öberg, T.; Todeschini, R.; Fourches, D.; Varnek, A., Critical Assessment of QSAR Models of Environmental Toxicity against *Tetrahymena pyriformis*: Focusing on Applicability Domain and Overfitting by Variable Selection. *J. Chem. Inf. Model.* **2008**, *48* (9), 1733-1746. DOI: 10.1021/ci800151m.
- [45] Masand, V. H.; Rastija, V., PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors. *Chemom. Intell. Lab. Syst.* **2017**, *169*, 12-18. DOI: 10.1016/j.chemolab.2017.08.003.
- [46] 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* (24), 2121-2132. DOI: 10.1002/jcc.23361.
- [47] Zaki, M. E. A.; Al-Hussain, S. A.; Bukhari, S. N. A.; Masand, V. H.; Rathore, M. M.; Thakur, S. D.; Patil, V. M., Exploring the Prominent and Concealed Inhibitory Features for Cytoplasmic Isoforms of Hsp90 Using QSAR Analysis. *Pharmaceutics* **2022**, *15* (3). DOI: 10.3390/ph15030303.
- [48] Zaki, M. E. A.; Al-Hussain, S. A.; Masand, V. H.; Sabnani, M. K.; Samad, A., Mechanistic and Predictive QSAR Analysis of Diverse Molecules to Capture Salient and Hidden Pharmacophores for Anti-Thrombotic Activity. *Int. J. Mol. Sci.* **2021**, *22* (15). DOI: 10.3390/ijms22158352.
- [49] Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct. Prop. Relatsh* **2020**, *5* (3), 61-97. DOI: 10.4018/IJQSPR.20200701.oa1.
- [50] Gramatica, P., On the Development and Validation of QSAR Models. In *Computational Toxicology, Methods Mol. Biol.* **2013**, pp. 499-526.
- [51] Chirico, N.; Gramatica, P., Real External Predictivity of QSAR Models: How To Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient. *J. Chem. Inf. Model.* **2011**, *51* (9), 2320-2335. DOI: 10.1021/ci200211n.
- [52] Gramatica, P., Principles of QSAR models validation: internal and external. *QSAR Comb. Sci.* **2007**, *26* (5), 694-701. DOI: 10.1002/qsar.200610151.
- [53] Sleire, L.; Førde, H. E.; Netland, I. A.; Leiss, L.; Skeie, B. S.; Enger, P. Ø., Drug repurposing in cancer. *Pharmacol. Res.* **2017**, *124*, 74-91. DOI: 10.1016/j.phrs.2017.07.013.

- [54] Eid, A. H., Drug Repurposing in Cancer: Now and Beyond. *Curr. Med. Chem.* **2021**, *28* (11), 2083-2084. DOI: 10.2174/092986732811210416084626.
- [55] Johnson, T. W.; Richardson, P. F.; Bailey, S.; Brooun, A.; Burke, B. J.; Collins, M. R.; Cui, J. J.; Deal, J. G.; Deng, Y. -L.; Dinh, D.; *et al.* Discovery of (10R)-7-Amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(metheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a Macrocyclic Inhibitor of Anaplastic Lymphoma Kinase (ALK) and c-ros Oncogene 1 (ROS1) with Preclinical Brain Exposure and Broad-Spectrum Potency against ALK-Resistant Mutations. *J. Med. Chem.* **2014**, *57* (11), 4720-4744. DOI: 10.1021/jm500261q.
- [56] Gaudreault, F.; Morency, L. -P.; Najmanovich, R. J., NRGsuite: a PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinform.* **2015**, *31* (23), 3856-3858. DOI: 10.1093/bioinformatics/btv458.
- [57] Gaudreault, F.; Najmanovich, R. J., FlexAID: Revisiting Docking on Non-Native-Complex Structures. *J. Chem. Inf. Model.* **2015**, *55* (7), 1323-1336. DOI: 10.1021/acs.jcim.5b00078.
- [58] Bowers, K. J.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; *et al.* Molecular dynamics---Scalable algorithms for molecular dynamics simulations on commodity clusters. In Proceedings of the 2006 ACM/IEEE conference on Supercomputing - SC '06, **2006**.
- [59] Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W., Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6* (5), 1509-1519. DOI: 10.1021/ct900587b.
- [60] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926-935. DOI: 10.1063/1.445869.
- [61] Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177-4189. DOI: 10.1063/1.467468.
- [62] Martyna, G. J.; Klein, M. L.; Tuckerman, M., Nosé-Hoover chains: The canonical ensemble *via* continuous dynamics. *J. Chem. Phys.* **1992**, *97* (4), 2635-2643. DOI: 10.1063/1.463940.
- [63] Di Pierro, M.; Elber, R.; Leimkuhler, B., A Stochastic Algorithm for the Isobaric-Isothermal Ensemble with Ewald Summations for All Long Range Forces. *J. Chem. Theory Comput.* **2015**, *11* (12), 5624-5637. DOI: 10.1021/acs.jctc.5b00648.
- [64] Piao, L.; Chen, Z.; Li, Q.; Liu, R.; Song, W.; Kong, R.; Chang, S., Molecular Dynamics Simulations of Wild Type and Mutants of SAPAP in Complexed with Shank3. *Int. J. Mol. Sci.* **2019**, *20* (1). DOI: 10.3390/ijms20010224.
- [65] Huang, W. -S.; Liu, S.; Zou, D.; Thomas, M.; Wang, Y.; Zhou, T.; Romero, J.; Kohlmann, A.; Li, F.; Qi, J.; *et al.* Discovery of Brigatinib (AP26113), a Phosphine Oxide-Containing, Potent, Orally Active Inhibitor of Anaplastic Lymphoma Kinase. *J. Med. Chem.* **2016**, *59* (10), 4948-4964. DOI: 10.1021/acs.jmedchem.6b00306.
- [66] Marcu, M. G.; Schulte, T. W.; Neckers, L., Novobiocin and Related Coumarins and Depletion of Heat Shock Protein 90-Dependent Signaling Proteins. *J. Natl. Cancer Inst.* **2000**, *92* (3), 242-248. DOI: 10.1093/jnci/92.3.242.
- [67] Zhang, T.; Hamza, A.; Cao, X.; Wang, B.; Yu, S.; Zhan, C. -G.; Sun, D., A novel Hsp90 inhibitor to disrupt Hsp90/Cdc37 complex against pancreatic cancer cells. *Mol. Cancer Ther.* **2008**, *7* (1), 162-170. DOI: 10.1158/1535-7163.Mct-07-0484.