



Identification of 5-nitroindazole as a multitargeted inhibitor for CDK and transferase kinase in lung cancer: a multisampling algorithm-based structural study

Shaban Ahmad¹ · Khalid Raza¹

Received: 12 March 2023 / Accepted: 5 April 2023 / Published online: 14 April 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract

Lung cancer is the second most common cancer, which is the leading cause of cancer death worldwide. The FDA has approved almost 100 drugs against lung cancer, but it is still not curable as most drugs target a single protein and block a single pathway. In this study, we screened the Drug Bank library against three major proteins- ribosomal protein S6 kinase alpha-6 (6G77), cyclic-dependent protein kinase 2 (1AQ1), and insulin-like growth factor 1 (1K3A) of lung cancer and identified the compound 5-nitroindazole (DB04534) as a multitargeted inhibitor that potentially can treat lung cancer. For the screening, we deployed multisampling algorithms such as HTVS, SP and XP, followed by the MM\GSA calculation, and the study was extended to molecular fingerprinting analysis, pharmacokinetics prediction, and Molecular Dynamics simulation to understand the complex's stability. The docking scores against the proteins 6G77, 1AQ1, and 1K3A were -6.884 kcal/mol, -7.515 kcal/mol, and -6.754 kcal/mol, respectively. Also, the compound has shown all the values satisfying the ADMET criteria, and the fingerprint analysis has shown wide similarities and the water WaterMap analysis that helped justify the compound's suitability. The molecular dynamics of each complex have shown a cumulative deviation of less than 2 Å, which is considered best for the biomolecules, especially for the protein–ligand complexes. The best feature of the identified drug candidate is that it targets multiple proteins that control cell division and growth hormone mediates simultaneously, reducing the burden of the pharmaceutical industry by reducing the resistance chance.

Keywords Multitargeted docking · 5-Nitroindazole · Molecular fingerprints · WaterMap · Molecular dynamics simulation

Introduction

Lung carcinoma (LC) is the leading cause of illness and death globally among lung diseases, and it begins as a primary metastatic tumour in the lung and subsequently spreads to other body regions [1, 2]. Small-cell lung cancer, also known as SCLC and non-small-cell lung cancer that is NSCLC, is the most diagnosed type of LC. Most lung cancer statistics include SCLC and NSCLC, and the SCLC accounts for around 10% to 15% of all lung malignancies [3,

4]. The hallmark signs of LC include weight loss, breathing difficulty, bloody coughing, and chest discomfort. LC is now the fourth most significant cause of hospitalisation in people with respiratory diseases, while NSCLC is the primary cause of LC-associated death (up to 85 per cent of LC) (NSCLC). Genetic and epigenetic alterations in the cellular DNA led LC to form [5, 6]. The extensive molecular dissection of NSCLC has recognised the groundwork for developing new small therapeutic compounds that target mutations in the EGFR, K-Ras, ALK, c-MET, B-Raf, LKB1, and NKX2-1 genes, all of which are important in disease progression [7, 8]. Among these LC genes, activating mutations in the EGFR gene are seen in 10–40% of NSCLC patients [7, 9]. The EGFR gene encodes a transmembrane epidermal growth factor receptor protein that, when activated (by ligand interaction), transmits signals necessary for migration, cellular proliferation, differentiation and survival [10]. Tobacco smoking, hereditary factors, food and obesity, and environmental variables such as air pollution have all been

✉ Khalid Raza
kraza@jmi.ac.in

Shaban Ahmad
Shaban184343@st.jmi.ac.in

¹ Computational Intelligence and Bioinformatics Laboratory,
Department of Computer Science, Jamia Millia Islamia,
New Delhi 110025, India

associated with LC's aetiology [11, 12]. Lung cancers are widely classed as SCLC and NSCLC; the latter accounts for 85 per cent of all cases, with adenocarcinoma accounting for 40 per cent of NSCLCs. The most effective strategy to treat NSCLC adenocarcinoma is to target the epidermal growth factor receptor (EGFR) [13, 14]. The most prevalent EGFR-targeting medicines are Erlotinib, Gefitinib, and Afatinib. Among the challenges medicinal chemists addressed was identifying polymorphism-related kinase inhibitors, one of the critical targets for EGFR tyrosine kinase inhibitors [15]. The use of EGFR therapy to address NSCLC with the mutational resistance of T790M is a critical treatment requirement. Hyperexpression of the EGFR tyrosine kinase has been identified as the most prevalent reason behind the NSCLCs, which mainly afflict cigarette smokers and are gender-specific to females. Osimertinib and Afatinib are second- and third-generation NSCLC treatment drugs [16–18]. The first-generation reversible NSCLC treatment drugs were developed to handle EGFR L858R mutations. Second-generation irreversible NSCLC treatment drugs targeted EGFR T790M mutations [19, 20]. Third-generation irreversible NSCLC therapeutic drugs were also developed to treat EGFR T790M/L790M double mutations.

In lung cancer, there are huge lists of reported proteins and genes, and a few biomarkers have been used extensively, and this is true in various cases to design the drug and the reason behind the resistance development. The PDB ID 6G77 entitled RSK4 N-terminal kinase domain has been used for various drug repurposing as it is a promoter for drug resistance and metastasis that can be an excellent option to treat as the target [21]. PDB ID 1AQ1 is Human Cyclin-Dependent Kinase 2 that participates in the DNA replication processes and cell division in all eukaryotes, including humans. CDK2 forms complexes with cyclin E in the G and G/S phases of the cell cycle [22]. In comparison, the PDB ID 1K3A is an insulin-like growth factor 1 receptor kinase that participates in the therapeutic interventions managed by autophosphorylation

within three sites of the kinase activation loop [23]. On the other hand, the fast growth of computational biology or bioinformatics has created a fantastic chance for designing novel molecules with features to remove the resistance to EGFR-specified sensitivity. Computational techniques for predicting drug candidates for the resistance to developed mutational and creating resistance-evading medications have proved highly reliable. The molecular modelling approach is molecular docking, which is used to investigate the interaction between the 3-D structures of a ligand and a receptor and how the ligand binds firmly in the active region of the receptor. It also helps with the virtual screening of a chemical library during the pre-clinical stage of drug development. When there are numerous compounds to examine and access to physical samples is limited, the evaluation of absorption, distribution, metabolism, and excretion (ADME) in drug development is included early in the discovery phase [24–29].

In this study, we have identified three important protein targets of lung cancer, screened the drug bank against each, and identified the potential drug candidate 5-nitroindazole as a multitargeted inhibitor against lung cancer proteins. Further, we extended our analysis for the ADMET and fingerprinting, and after getting satisfactory results, we simulated the system in water for 100 ns in the NVT ensemble and analysed various constraints for all three complexes.

Methods

We have downloaded the drug library from Drug Bank, proteins from PDB, and docked. The study extended to multiple directions to explore the drug candidate's suitability. The same has been plotted in Fig. 1 to understand the methods easily. Further, the detailed methods are as follows-

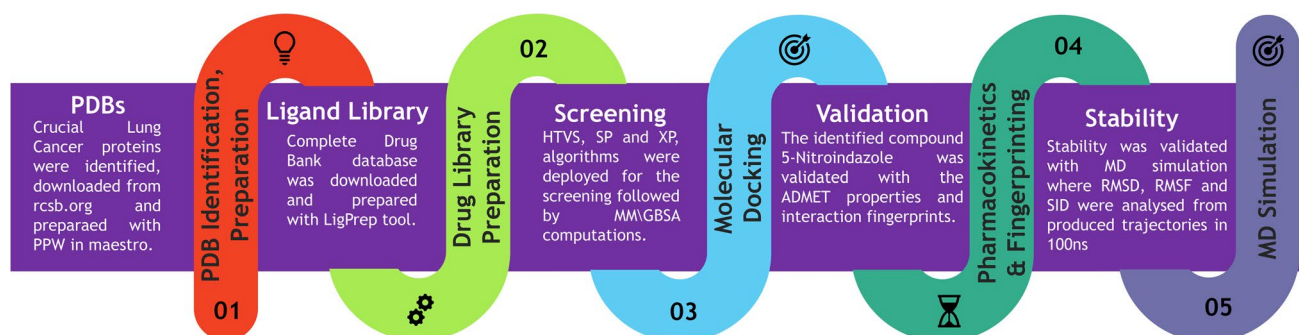


Fig. 1 The workflow of the complete study; graphical abstract showing the methods to identify the 5-nitroindazole as a multitargeted inhibitor against lung cancer

Protein preparation

We mined various through various literature and identified the protein targets responsible for an essential role in the case of lung cancer. The identified proteins were RSK4 N-terminal kinase domain, human cyclin-dependent Kinase 2, and insulin-like growth factor 1 receptor kinase that further was downloaded from the <https://www.rcsb.org/> database and their PDBID were- 6G77, 1AQ1, and 1K3A, and imported into the workspace of Schrödinger Maestro for preparation using the Protein Preparation Wizard [21–23, 30, 31]. PDBID: 6G77 has two chains, A and B, with solvent and other metals/ ions, and after preparation, we kept only clean chain A, and removed chain B, solvent and other metals/ions. In PDBID: 1AQ1 and 1K3A have only chain A and solvents. We deleted solvent from the protein during preparation and kept only chain A for the subsequent study. We maintained the same parameters for all proteins throughout the preparation process, and we chose the preprocess workspace structure tab to assign bonds using the CCD tool. We added H-atoms, created zero-order and disulfide bonds, converted selenomethionines to methionines, filled in missing side chains, and loops using Prime, and generated het states using Epik at pH 7.0 to ± 2.0 . [30, 32, 33]. We also refined our structure using crystal symmetry and optimised and removed water energy. Further, each PDB structure was minimised with the OPLS4 force field [34].

Ligand library collection and preparation

The Drug Bank database provides an interactive platform to access information about the drug as well the structures of the compounds. That is the main reason behind taking the complete database of the Drug Bank to access its information for our studies. We downloaded the entire ligand library from <https://go.drugbank.com/>, which contains 14,940 compounds, many of which are licensed biologic medications, some of which are nutraceuticals, and some were experimental drugs updated in January 2022 with version 5.1.9 [35]. We have used the LigPrep tool in Maestro for preparation. The OPLS4 force field was used to minimise the ligands, and the size was restricted to not more than 500 atoms to filter the compounds not fitting into the good drug candidate category [34, 36]. Tautomers and stereoisomers were created with the given chiral carbon, computations were limited to not more than 32 per ligand, and the complete process produced a sum of 1,55,888 ligands that were further used for molecular docking.

Glide grid generation and multitargeted molecular docking and filtering with MM\GBSA

An integral and crucial part of the docking procedure is generating grids on the active site. The active site of the

proteins was calculated with the help of the SiteMap tool with predefined algorithms to predict the protein's active site [30]. The PDBIDs: 6G77, 1AQ1, and 1K3A were individually selected and entered into the grid box to fit on the active site and performed the gridding sets. Further, molecular docking was performed with the help of Maestro's 'Virtual Screening Workflow' (VSW) tool that offers ligand-target docking-based screening with multiple sampling algorithms at once [30, 37]. The ligand library was filtered using QikProp and Lipinski's Rule to meet the requirements for Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) characteristics [38]. Additionally, Epik state penalties for docking were generated [32]. Further, to reduce the computational cost, we performed the High-Throughput Virtual Screening (HTVS), Standard-precision (SP), and Extra precision (XP) and passed only the top 5% of the data to the next level of screening. Additionally, we have filtered out best poses with the help of Molecular mechanics with generalised born and surface area solvation (MM/GBSA). In order to further manual filtering, the calculated data were sorted to determine which compound or complex had the highest likelihood of binding to each of the chosen protein targets [33].

Pharmacokinetic properties prediction and interaction fingerprinting analysis

The molecular level of the information regarding the pharmacokinetic properties of the compounds was generated using the QikProp tool, which provides an extensive calculation and various ADMET features [30, 38]. We took the 5-nitroindazole chemical from the workspace, kept it for the calculations, and contrasted it with the standard values. Drug-likeness of 5-nitroindazole was evaluated by using various descriptors and pharmacokinetic properties like predicted octanol/ water partition coefficient (QPlogPo/w), predicted aqueous solubility (QPlogPo/w), predicted apparent Caco-2 permeability (QPPCaco), predicted brain/blood partition coefficient (QPlogBB), Lipinski's rule of 5, and human oral absorption. Further, we have also performed and analysed the interaction fingerprints of the protein ligands. The interaction fingerprints tool was used to analyse the patterns, and the protein–ligand complexes were used for this study. All the proteins had different sequences, so we aligned them and generated the fingerprints. We have selected any in the bonding types, coloured the main plot against the docking score and exported only the interacting residues to find the better pattern. Further, the data were taken into the main plot for pattern, residue interaction count, and ligand interactions for proper understanding.

WaterMap analysis

Protein–ligand complexes are stabilised in large part by water molecules. Therefore, it is critical to comprehend how water molecules behave and are distributed within a protein–ligand complex's binding site. Water map analysis is a practical method for examining how water molecules interact with the protein–ligand complex [39]. To better understand water molecules' role in the binding pocket of all three proteins, which have a significant role in ligand binding, we performed WaterMap analysis using the WaterMap module from the Schrodinger Maestro, which uses various algorithms to predict the most probable locations of water molecules in the binding site [30, 40]. The ligand was picked from the workspace for each complex individually and then analysed the waters within 5 Å. Further, in the simulation setup, we truncated the protein while using the OPLS4 forcefield and the water crystal structure was eliminated. After simulation, the selected ligands were evaluated for the possible overlapping of the water molecule's hydration site and Gibbs's energy. Insights into the hydration patterns of protein–ligand complexes can be gained by examining water maps, which can also be used to direct ligand optimization for increased binding affinity and specificity. For instance, water molecules continuously present in the binding site can be targeted for displacement by ligand changes to increase binding affinity [41].

Molecular dynamics simulation

The stability and flexibility of the atom are considerably improved in biological and material sciences by the molecular dynamic (MD) simulation, a group of algorithms. For the MD simulation analysis, we used the Desmond package developed by Shaw Research, regarded as the MD algorithm's fastest and most accurate computation [42]. In order to set up the system using the SPC water model, orthorhombic boundary conditions in a buffer system at a distance of 10 Å × 10 Å × 10 Å were implemented [30, 42]. In the complex of 5-Nitroindazole with 6G77, and 5-Nitroindazole with 1AQ1, we have added 5Cl⁻ whereas in the complex of 5-Nitroindazole with 1K3A, we added 17Na⁺ atoms to neutralise the complete system. The complete system constructed for the MD simulation was minimised after

removing the salt and ion placement inside 20 Å. Further, in the molecular dynamics simulation panel, we loaded the system builder files, and the simulation time was set for the 100 ns and recording interval of 100 ps for each trajectory that generated 100 frames. We kept the energy level of 1.2 while using the NPT ensemble class at 300 K with a pressure bar of 1.02325 and relaxed the model before simulation [43]. We chose the NPT ensemble as the complex may undergo significant volume changes and to study phase transitions and solvation processes. Our complex belongs to the biomolecules category, and the NPT ensemble allows the system's volume to fluctuate in response to changes in pressure, and the production run needs to get interactions, deviations, and fluctuations that were analysed with SID.

Results and discussion

Interaction analysis and MM\GBSA analysis

The multisampling algorithms-based screening has led us to identify the multitargeted potential of 5-nitroindazole. We have analysed its bonding configurations with the help of the ligand interaction diagram tool to get with bond and residue types. The 5-nitroindazole with ribosomal protein S6 kinase alpha-6 (6G77) showed a docking score of −6.88 kcal/mol and MM/GBSA score of −30.17 kcal/mol (Table 1) interaction with four hydrogen bonding by LYS105, THR215. Both residues individually interact with the O atom, ASP153 with the NH atom, and LEU155 with the N atom of the ligand (Fig. 2A). Interaction of 5-nitroindazole with cyclic-dependent protein kinase 2 (1AQ1) shows a docking score of −7.51 kcal/mol and MM/GBSA score of −30.34 kcal/mol while interacting with two hydrogen bonds among LEU83 with N atom, GLU81 with NH atom and pi-cation interact by PHE80 with N⁺ atom of the ligand (Fig. 2B). 5-Nitroindazole interacts with Insulin-like growth factor 1 (1K3A) and has shown a docking score of −6.75 kcal/mol and MM/GBSA score of −23.22 kcal/mol by two hydrogen bonding among GLU1050, MET1052 with NH atom and with N atom also form a salt bridge by LYS1003 with O atom of the ligand (Fig. 2C). The complete ligand interaction diagram has led us to identify the best bonding residues with different types, such as hydrogen bonds, pi-cations, and many more.

Table 1 Showing the docking score (Kcal/mol) and MM\GBSA (kcal/mol) score and other important calculations against each protein and 5-nitroindazole compound

S. no.	PDB ID	Compound name	Docking score	MM\GBSA	Prime H-bond	Prime vdW
1	6G77	5-Nitroindazole	−6.884	−30.17	−153	−1376.02
2	1AQ1	5-Nitroindazole	−7.515	−30.34	−143.6	−1319.21
3	1K3A	5-Nitroindazole	−6.754	−23.22	−165.96	−1319.54

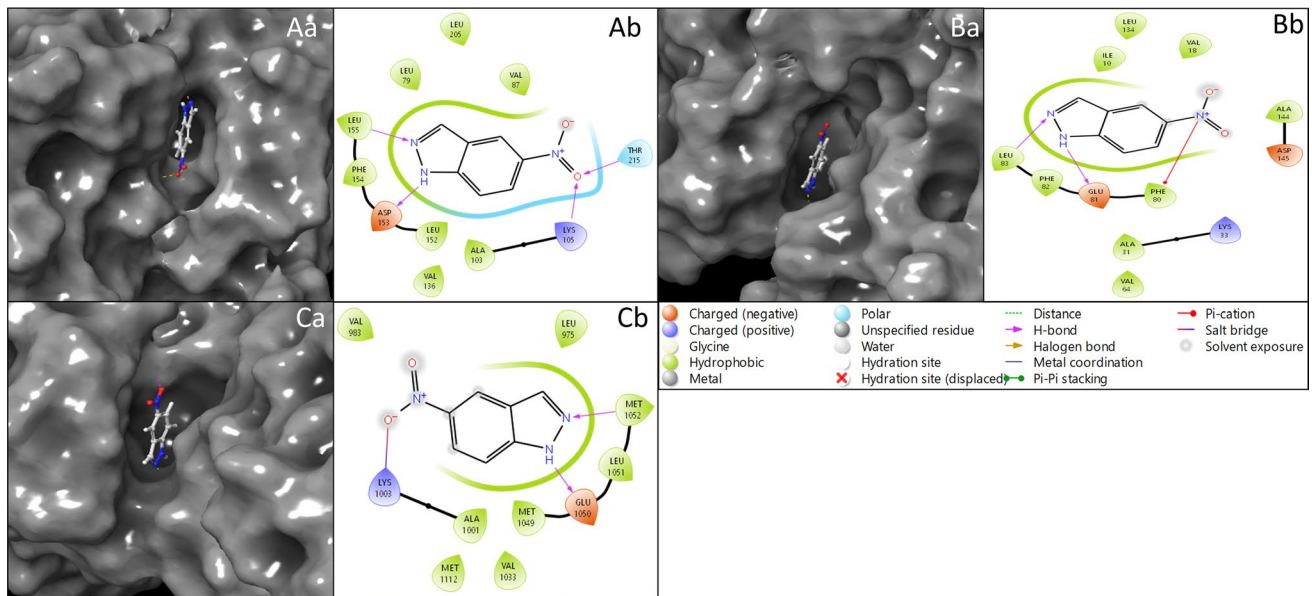


Fig. 2 Showing the 3-D and 2-D protein–ligand interactions. The 3-D representation of **Aa** 6G77, **Ba** 1Aq1, **Ca** 1K3A are shown while the 2-D representation of **Ab** 6G77, **Bb** 1Aq1, **Cb** 1K3A, and the legend is shown for proper interpretations of the bonds and residues

Also, we have identified the coverage in the main pocket of the proteins. The identified compound is compact in the protein's pocket and has shown broad interaction types such as hydrogen bonding, pi-cation, and many more that make the structure much stronger and stable during the treatment or even at the validation level. This interaction has also predicted that the molecular dynamics simulation might have fewer deviations and fluctuations, and it also has provided insight to take the study for the WaterMap computations.

Pharmacokinetic and interaction pattern Identification

The ADMET analysis has revealed that the identified compound 5-nitroindazole can be a better multitargeted therapeutic, and its higher doses can be administered to make it more efficient for cancer cells. The compound is also inactive for the CNS, which boosts the level of understanding of how this compound cannot harm other brain and nervous systems. The compound has a molecular weight of 163.135 g/mol, considered among the optimised ones, and the compound has no amines, amidines, acids, or amides. Also, the compound's SASA is 336.065, FOSA is 0, FISA is 165.954, PISA is 170.11, and WPSA is 0 (Table 2). There is also one donor and two acceptor hydrogen bond capacity to stabilise the compound with the protein targets. The complete ADMET analysis has revealed that 5-nitroindazole can be one of the prominent compounds to treat lung cancer, and its multitargeted approach can also accept and will show a boosted performance to cure lung cancer [26,

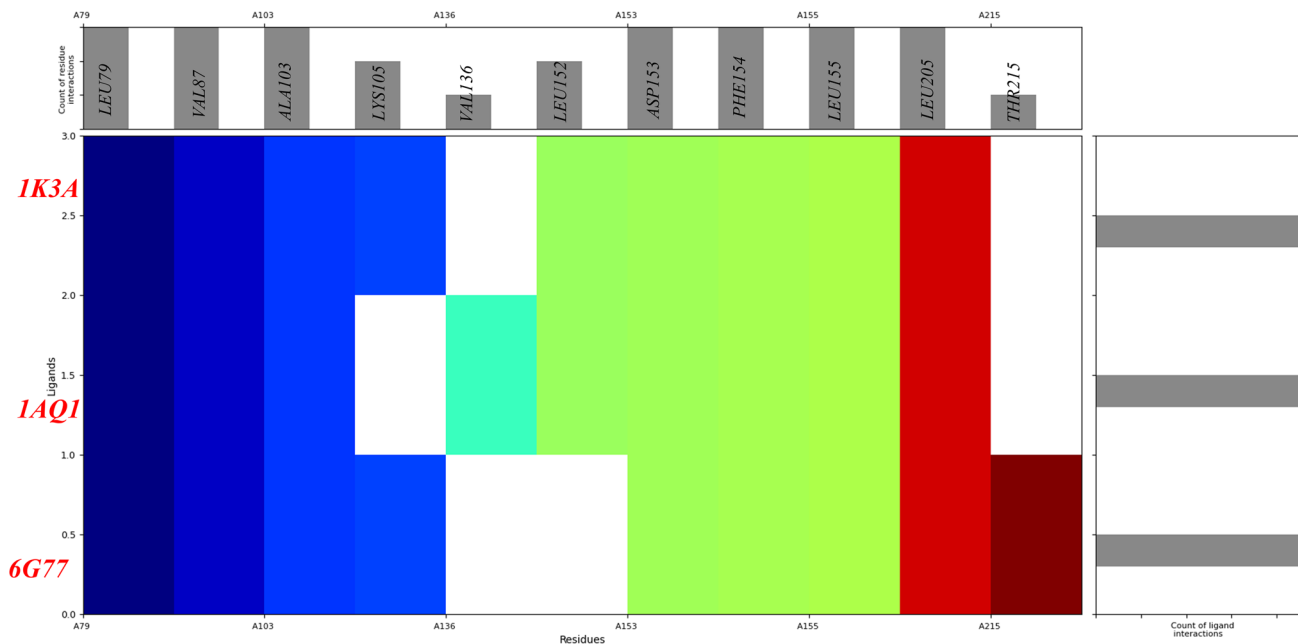
38, 44–47]. The interaction pattern (Fig. 3) analysis has revealed that the compound is widely interacting with an identified and expected pattern. LEU79, VAL87, ALA103, LYS105, VAL136, LEU152, LEY152, ASP153, PHE154, PHE154, LEU155, LEU205, and THR215 are the interacting residues with 5-nitroindazole. The coloured plot shows the interaction patterns against the position of the amino acids of 6G77. 1Aq1 and 1K3A proteins, while the left plot shows the ligand interaction counts. Most of the interactions were found in the initial sequences of the proteins and the residues after 152 positions. The pattern analysis has shown that the compound 5-nitroindazole has enough potential to bind multiple targets, and its higher doses might block multiple targets together and can lead to the shrinking of the lung cancer cells. The compound's H-bond acceptor and donor capacity make it unique and boost its stability with multiple protein targets.

WaterMap analysis hydration site prediction

A computational technique called the WaterMap analysis investigates how a ligand interacts with a protein when water molecules are present. We have used Desmond for the backend computations and the WaterMap tool from Schrodinger's front end to make the computations using the GPU. A protein–ligand complex's hydration patterns can be understood using this technique, and the water molecules essential to the complex's stabilisation can be found. It often entails protein–ligand–water molecular dynamics simulations, which offer a thorough understanding of the

Table 2 Showing the ADMET properties of 5-nitroindazole calculated with the QikProp tool against the standard values

Descriptors	Standard Values	5 – Nitroindazole	Descriptors	Standard Values	5 – Nitroindazole
#Stars	0 – 5	0	QPlogS	– 6.5 to 0.5	– 1.73
#Amine	0 – 1	0	CIQPlogS	– 6.5 to 0.5	– 2.244
#Amidine	0	0	QPlogHERG	concern below – 5	– 3.457
#Acid	0 – 1	0	QPPCaco	< 25 poor, > 500 great	264.335
#Amide	0 – 1	0	QPlogBB	– 3.0 to 1.2	– 0.773
#Rotor	0 – 15	1	QPPMDCK	< 25 poor, > 500 great	117.427
#rtvFG	0 – 2	0	QPlogKp	– 8.0 to – 1.0	– 3.883
CNS	– 2 (inactive), + 2 (active)	– 1	IP(eV)	7.9 – 10.5	9.544
Mol MW	130.0 – 725.0	163.135	EA(eV)	– 0.9 to 1.7	1.223
Dipole	1.0 – 12.5	6.586	#metab	1 – 8	1
SASA	300.0 – 1000.0	336.065	QPlogKhsa	– 1.5 to 1.5	– 0.375
FOSA	0.0 – 750.0	0	HumanOralAbsorption	N/A	3
FISA	7.0 – 330.0	165.954	PercentHumanOralAbsorption	> 80% is high, < 25% is poor	77.139
PISA	0.0 – 450.0	170.11	SAfluorine	0.0 – 100.0	0
WPSA	0.0 – 175.0	0	SAamideO	0.0 – 35.0	0
Volume	500.0 – 2000.0	520.507	PSA	7.0 – 200.0	75.765
DonorHB	0.0 – 6.0	1	#NandO	2 – 15	5
AccptHB	2.0 – 20.0	2	RuleOfFive	Maximum is 4	0
dip^2/V	0.0 – 0.13	0.083336	RuleOfThree	Maximum is 3	0
ACxDN^5/SA	0.0 – 0.05	0.0059512	#ringatoms	N/A	9
glob	0.75 – 0.95	0.931165	#in34	N/A	0
QPpolrz	13.0 – 70.0	15.659	#in56	N/A	9
QPlogPC16	4.0 – 18.0	5.814	#noncon	N/A	0
QPlogPoct	8.0 – 35.0	8.667	#nonHatm	N/A	12
QPlogPw	4.0 – 45.0	6.028	Jm	N/A	0.577
QPlogPo/w	– 2.0 to 6.5	1.168			

**Fig. 3** The interaction patterns of 5-nitroindazole against the considered proteins 6G77, 1AQ1, and 1K3A and coloured with blue to red to understand the position of the residue from N to C terminal

system's dynamic activity throughout time. Each picture of the system created by the simulations represents a different conformation of the complex. The WaterMap algorithm then determines the portions of the protein with a high affinity for water by examining the chemical interactions and solvation energies of the water molecules around the protein and ligand in each image. The tool then groups the water molecules into stable water clusters that contribute to the protein–ligand complex's hydration pattern based on their positions and interaction energies. It is a helpful tool in the drug discovery process and can assist researchers in comprehending the mode of action of already available medications and designing new medications with increased efficacy. This approach can help design novel ligands with enhanced binding affinity, and more potent medications can be suggested. Analysing the apoprotein area that is 5 Å of the binding pocket of the protein revealed the presence of 21 water molecules around the ligand molecular 5-nitroindazole, which was interestingly found to be shared for each computation. The drug docking pocket contains water molecules with the hydration energy (ΔG) ranging from -2.82 to 6.58 kcal/mol with the occupancy ranging from 0.98 to 0.28, indicating the wide range of water molecules around the pocket, as shown in Table 3 to make it more apparent along with other values. Further, in Fig. 4, we have also shown the 3-D and 2-D representation of how the water molecules can be involved to interfere with 5-nitroindazole.

Molecular dynamics simulation analysis

In order to analyse all complex deviations, fluctuations, and interactions in natural-like systems, we performed the MD simulation study for 100 ns to understand the stability of the complexes. We built a water system as solute and in neutralised conditions to simulate each complex individually. The detailed analysis is as follows-

Root-mean-square deviation

The root mean square deviation (RMSD) is a value that provides the deviation of the protein or ligand at a given time frame, and it also reflects how stable is the protein–ligand complex. The deviation analysis has enough potential to predict how the experimental costs will perform during the validations. The ribosomal protein S6 kinase alpha-6 (6G77) in complex with 5-nitroindazole initially deviated at 0.10 ns, showing protein at 1.93 Å and for the ligand at 1.72 Å. During the simulation period, it shows stable performance. At 100 ns, the protein deviated at 2.82 Å and the ligand at 5.19 Å. After neglecting the initial phase of the simulation, the RMSD result shows acceptable deviations, and the protein deviation can be termed as -0.89 Å (Fig. 5A). The cyclic-dependent protein kinase 2 (1AQ1)

with 5-nitroindazole started deviating and then showed a stabilised performance in the case of protein deviating at 1.91 Å and ligand 1.26 Å at 0.10 ns, further analysed at 100 ns, the protein deviated at 3.18 Å, and the ligand deviated at 1.41 Å. After ignoring the initial deviation, the noted deviation can be trusted (Fig. 5B). The Insulin-like growth factor 1 (1K3A) in complex with 5-nitroindazole has shown steady, after a few deviating at the initial level at 0.10 ns, for protein noted at 1.25 Å and ligand 1.13 Å. At 100 ns, it shows protein deviated of 3.78 Å and ligand of 0.86 Å, during simulation. After ignoring the first 1 ns, the overall complex deviation is acceptable (Fig. 5C). The complete deviation analysis has revealed that the protein and ligands are stabilised after certain nanoseconds, and the initial deviation was the result of a change in solute systems, the addition of ions, and, most important, the initial heat shocks. There was also a pattern of deviation among the amino acids that tended to deviate. However, their role in interaction was not much, so they were ignored.

Root-mean-square fluctuation

The root mean square fluctuation (RMSF), used to examine the protein's fluctuation and its interactions with the ligand 5-nitroindazole, provides protein and ligand fluctuation values at residual levels shown in Fig. 6. The residues of ribosomal protein S6 kinase alpha-6 (6G77) have performed stable during simulation. Only a few residues fluctuate beyond 2 Å are VAL51, GLY82, SER83, ARG117, VAL118, ARG119, THR120, LYS121, MET122, GLN227, LYS229, ALA230, TYR231, SER232, CYS234, GLY235, ALA288, and while during the simulation, 5-nitroindazole interacted 19 times with amino acids with LEU79, GLN81, PHE84, VAL87, ALA103, LYS105, VAL136, LEU152, ASP153, PHE154, LEU155, ARG156, GLY157, GLY158, GLU202, ILE204, LEU205, LEU214, and ASP216 (Fig. 6A). The cyclic-dependent protein kinase 2 (1AQ1) also performed so steadily while the simulation period, with some residues slightly more than 2 Å, those are GLY13, LEU25, THR41, GLU42, ASP92, ALA93, SER94, ALA95, LEU96, THR97, ALA151, PHE152, GLY153, VAL154, PRO155, VAL156, ARG157, TYR159, HIS161, VAL293, PRO294, HIS295, LEU296, ARG297, LEU298, and NMA298, and ligand atoms interact with protein 21 times at ILE10, GLY11, GLU12, GLY16, VAL18, VAL30, LYS33, VAL64, PHE80, GLU81, PHE82, LEU83, HIS84, ASP86, LYS129, PRO130, ASN132, LEU134, ALA144, ASP145, and PHE146 (Fig. 6B). The ligand 5-nitroindazole interacts with protein Insulin-like growth factor 1 (1K3A) 14 times at LEU975, GLN977, SER979, VAL983, ALA1001, LYS1003, VAL1033, MET1049, MET1052, ASP1056,

Table 3 Predicted hydration site, occupancy, and thermodynamic properties of the available hydration sites within five angstroms 5-nitroindazole in each protein–ligand complex

PDB IDs	Site	Occu pancy	Overlap	ΔH	$-T\Delta S$ (kcal/mol)	ΔG (Kcal/mol)	PDB IDs	Site	Occu pancy	Overlap	ΔH	$-T\Delta S$ (kcal/mol)	ΔG (Kcal/mol)
1K3A	1	0.95	0	-1.42	4.35	2.93	1AQ1	9	0.51	0	1.99	1.45	3.44
1K3A	2	0.91	1	1.33	3.25	4.58	1AQ1	10	0.44	0	-0.82	1.27	0.45
1K3A	3	0.89	0	-6.28	3.46	-2.82	1AQ1	11	0.44	0	1.88	1.35	3.23
1K3A	4	0.88	0	-3.16	3.39	0.23	1AQ1	12	0.41	1	2.9	1.14	4.04
1K3A	5	0.85	1	1.15	2.86	4.01	1AQ1	13	0.4	1	-0.32	1.12	0.8
1K3A	6	0.66	0	-2.48	2.44	-0.04	1AQ1	14	0.4	0	0.39	1.46	1.85
1K3A	7	0.63	0	0.75	1.99	2.74	1AQ1	15	0.39	0	2.11	1.05	3.16
1K3A	8	0.53	1	1.62	1.56	3.18	1AQ1	16	0.38	0	0.26	1.05	1.31
1K3A	9	0.53	0	-0.13	1.59	1.46	1AQ1	17	0.38	0.89	0.6	1.01	1.61
1K3A	10	0.51	1	2.11	1.44	3.55	1AQ1	18	0.37	0	0.86	1.06	1.92
1K3A	11	0.49	0	2.26	2.3	4.56	1AQ1	19	0.31	1	0.91	0.99	1.9
1K3A	12	0.47	0	-0.07	1.33	1.26	1AQ1	20	0.3	0	1.43	0.79	2.22
1K3A	13	0.47	0.21	-0.82	1.23	0.41	1AQ1	21	0.28	0	0.14	0.75	0.89
1K3A	14	0.45	0	2.03	1.31	3.34	6G77	1	0.98	0	-3.44	4.2	0.76
1K3A	15	0.44	0	0.16	1.65	1.81	6G77	2	0.92	1	3.3	3.28	6.58
1K3A	16	0.42	0	1.77	1.39	3.16	6G77	3	0.86	0	0.83	2.75	3.58
1K3A	17	0.37	0	-2.15	1.18	-0.97	6G77	4	0.84	1	3.73	2.69	6.42
1K3A	18	0.35	0	-0.11	0.97	0.86	6G77	5	0.74	0	-1.67	2.85	1.18
1K3A	19	0.34	0	-0.5	0.9	0.4	6G77	6	0.7	1	1.23	2.42	3.65
1K3A	20	0.34	0	0.31	0.96	1.27	6G77	7	0.66	0	-0.43	2.72	2.29
1K3A	21	0.33	1	2.37	0.94	3.31	6G77	8	0.65	1	-0.76	2.22	1.46
1K3A	22	0.32	0	0.3	0.94	1.24	6G77	9	0.62	0	-0.61	1.9	1.29
1K3A	23	0.32	0.08	0.89	0.85	1.74	6G77	10	0.61	0.04	-0.23	1.9	1.67
1K3A	24	0.31	0	-1.03	0.94	-0.09	6G77	11	0.55	0	-0.95	1.55	0.6
1K3A	25	0.29	0	1.17	0.76	1.93	6G77	12	0.54	0	0.94	1.75	2.69
1K3A	26	0.29	0	4.31	0.87	5.18	6G77	13	0.49	0.3	-1.51	1.49	-0.02
1AQ1	1	0.96	1	2.08	3.94	6.02	6G77	14	0.46	0.97	2.17	1.29	3.46
1AQ1	2	0.94	0	-1.14	3.9	2.76	6G77	15	0.43	0	-0.99	1.28	0.29
1AQ1	3	0.92	1	1.79	3.02	4.81	6G77	16	0.36	0	-3.46	1.26	-2.2
1AQ1	4	0.91	0.33	-2.02	3.14	1.12	6G77	17	0.34	1	1.42	1	2.42
1AQ1	5	0.73	0	0.63	2.48	3.11	6G77	18	0.34	0	0.98	0.92	1.9
1AQ1	6	0.71	0	1.04	2.21	3.25	6G77	19	0.33	0	0.41	0.97	1.38
1AQ1	7	0.59	0	-1.89	1.86	-0.03	6G77	20	0.3	0	1.55	0.84	2.39
1AQ1	8	0.53	1	1.67	1.59	3.26	6G77	21	0.29	0	-1.48	0.87	-0.61

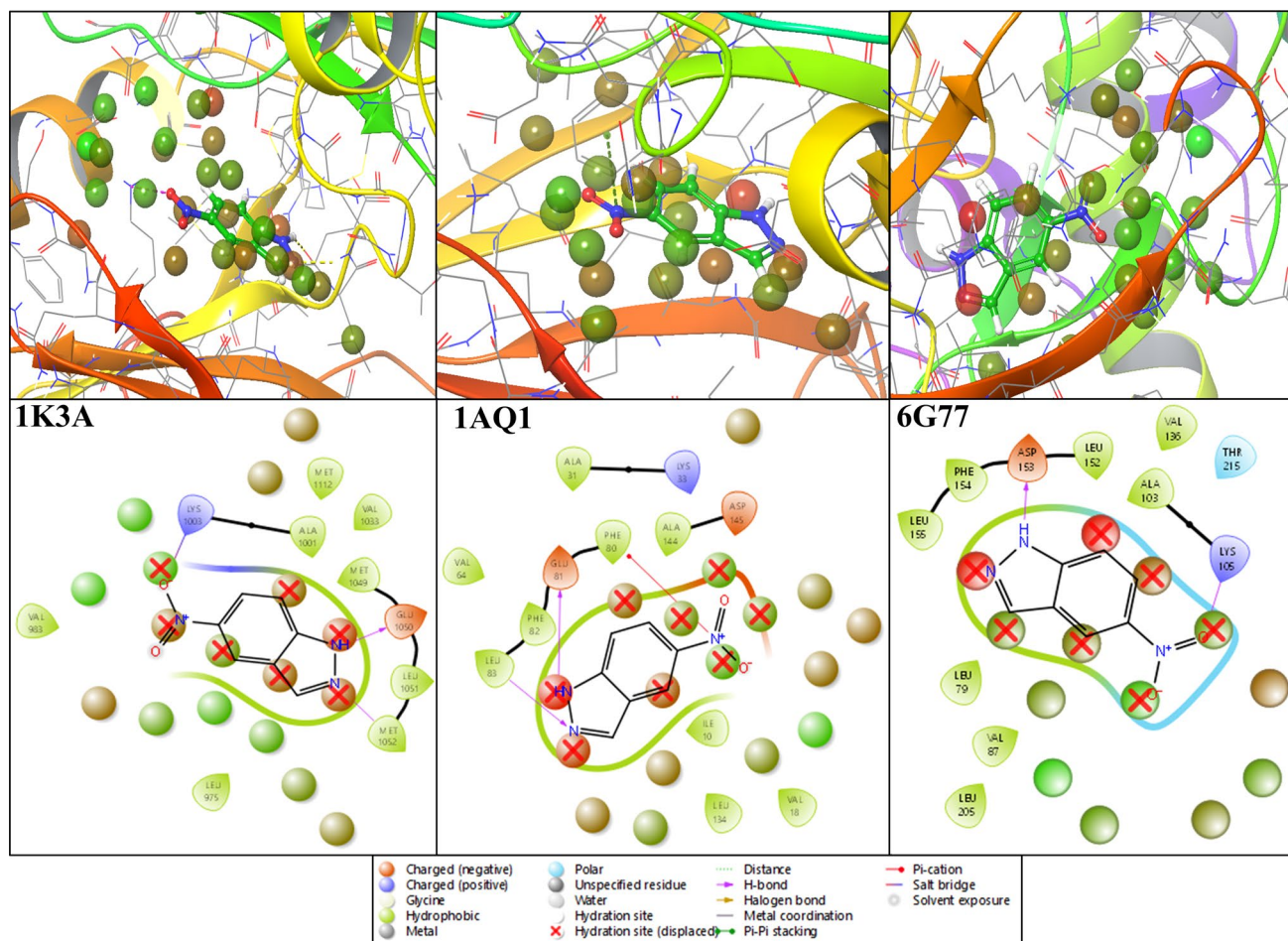


Fig. 4 The 3-D and 2-D poses of 5-nitroindazole with each protein used for simulation showing the hydrogen bond in yellow, Pi–Pi stacking in blue, Pi–cation in green, and salt bridge in lavender also shown the legend to correlate the amino acid types

ARG1109, MET1112, GLY1122, and ASP1123, and only a few residues fluctuated are VAL958, PRO959,

ASP960, GLU961, GLN977, GLY978, SER979, PHE980, GLY990, VAL991, VAL992, LYS993, ASP994, GLU995, PRO996, ASN1006, GLU1007, ALA1008, ALA1009, SER1010, MET1011, ARG1012, GLU1013, ARG1014, ILE1015, GLU1016, ASN1019, GLU1020, VAL1023, PRO1066, GLU1067, MET1068, GLU1069, ASN1070, ASN1071, PRO1072, VAL1073, LEU1074, ALA1075, GLY1139, GLY1140, LYS1141, GLY1142, LEU1143, LEU1144, PRO1145, and NMA1256 (Fig. 6C). The fluctuation study concludes that there is enormous interaction with amino acids that are least fluctuated while a few have shown the slightest fluctuation. However, the calculated fluctuation is acceptable in biological systems as they are not a natural part of the interaction with the ligand. The fluctuating residues are the one that provides an insight that there were a higher number of GLY that

tended to show the fluctuations, and the importance in interaction was negligible.

Simulative interaction analysis

The simulation interaction diagram (SID) shows interactions between proteins and ligands during simulations while improving ligand interaction poses. It also provides how the ligand atoms have contacted the protein residues, how and to what percentage they were contacted, and what new interaction formed to make the structure stable while mimicking the natural systems. The simulation interaction diagram usually introduces various new interactions to stabilise the structure, such as water bridges and cation and pi–pi interactions. These interactions are essential in making the system more stable and not allowing for its higher deviation and fluctuations. The ribosomal protein S6 kinase alpha-6 (6G77) in complex with 5-nitroindazole has shown a strong bonding with hydrogen bonds among THR215,

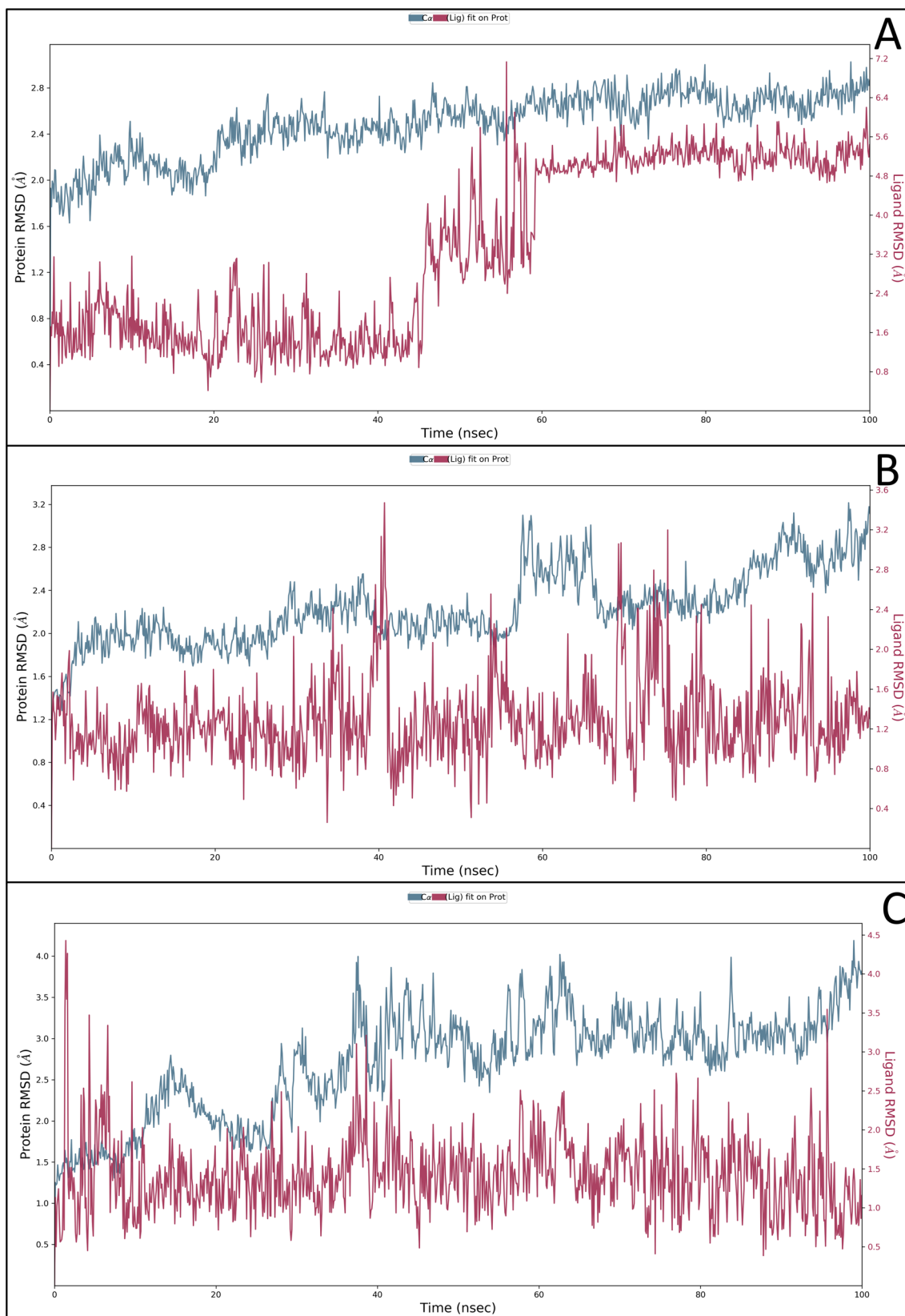


Fig. 5 Showing the root mean square deviation (RMSD) of the ligand and protein complex of 5-nitroindazole and **A** 6G77, **B** 1AQ1, **C** 1K3A. The red colour shows the deviation in the ligand, and the val-

ues are placed on the right side of the graph, while the blue shows the deviation in the proteins whose values are on the left side

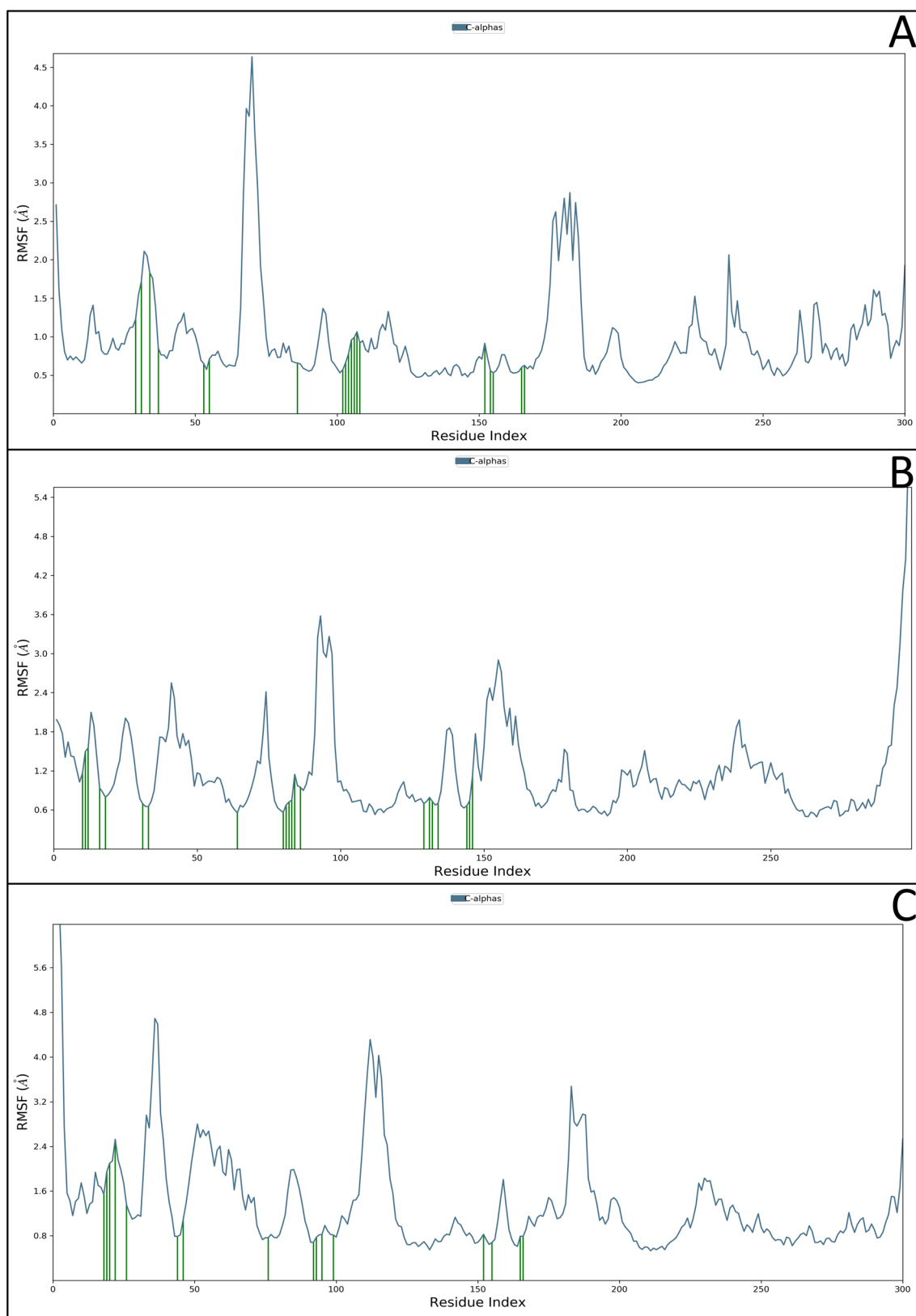


Fig. 6 Showing the root mean square fluctuation (RMSF) of the proteins with PDB ID **A** 6G77, **B** 1AQ1, **C** 1K3A. The blue colour shows the fluctuation in the protein residues, and the values are

placed on the left side of the graph, while the green bars show the ligand's contact with the amino acids at a given period

LEU155 with N atom, ASP153, LEU155 with NH atom, and also among GLU202, LYS105, and LEU155 with O atom of the ligand. Pi-cation contact by PHE84 with an N^+ atom also forms one salt bridge by GLU202 with an N^+ atom and involves nine water molecules (Fig. 7). The 5-nitroindazole complex with cyclic-dependent protein kinase 2 (1AQ1) involves four water molecules for balancing the interaction. The N atom and NH atom form hydrogen bonds with LEU83 and GLU81; the O atom also forms a hydrogen bond with ASP145 and LYS33. The pi-pi stacking and pi-cation interact with PHE80 residue with a benzene ring and N⁺ atom, forming a salt bridge by ASP145 with

the N^+ atom of the ligand (Fig. 7). Insulin-like growth factor 1 (1K3A) in complex with 5-nitroindazole forms hydrogen bonds with GLN977, SER979, ASP1123, LYS1003, LEU975, MET1052, and GLU1050 interact with O, N, and NH atom of the ligand. It also forms one salt bridge with ASP1123 by N^+ atom and eight water molecules involved (Fig. 7). Further, we have exported the count of interactions during the simulation periods and represented them into the histogram to make it more understandable which complex has shown how many interactions with types. Figure 7 also shows shown respective histogram or the simulation interaction diagram representing how stronger bonding configurations were formed to make the complex stable, and it is proof that the complex has excellent stability and can perform better experimentally.

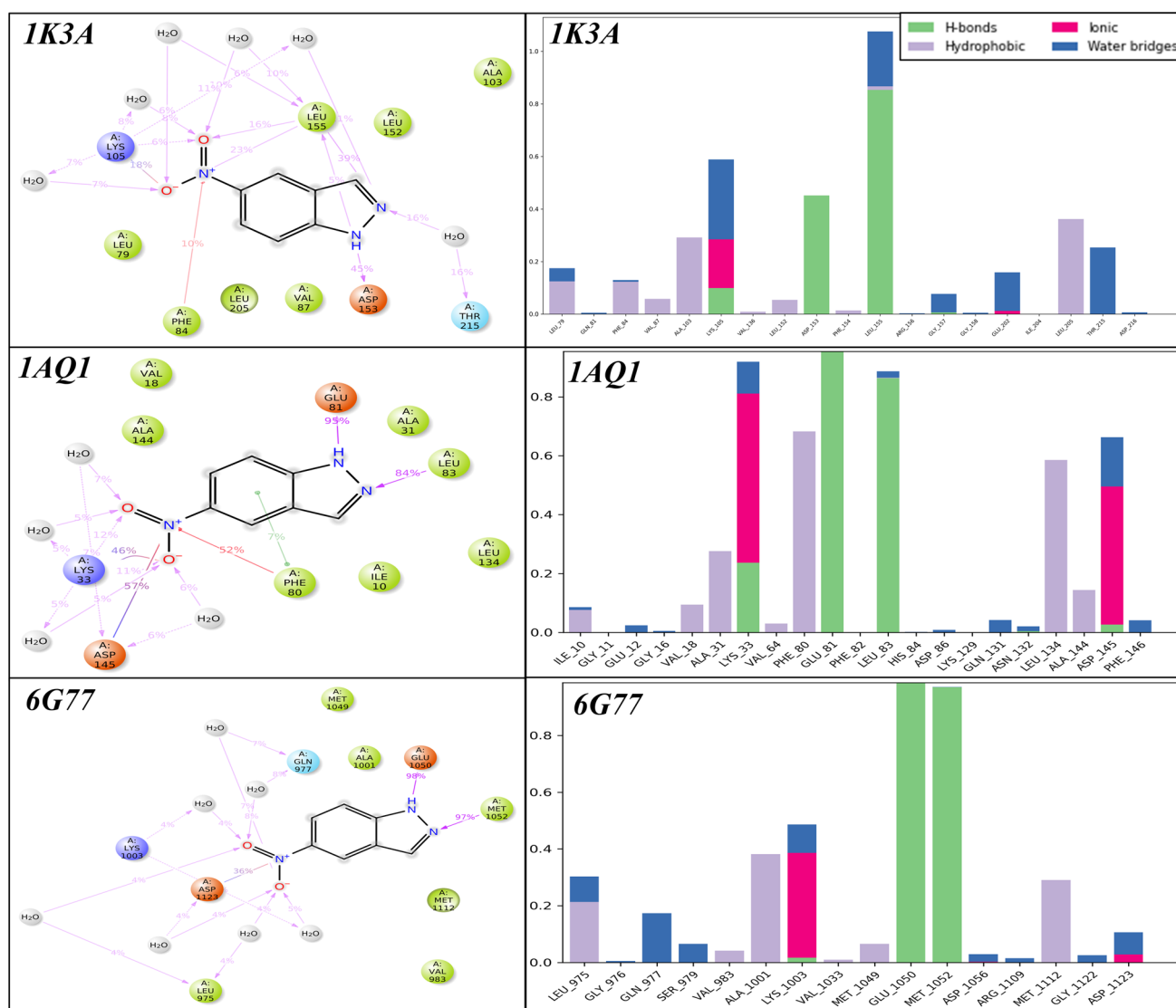


Fig. 7 Showing the Simulation Interaction Diagram and interactions count of 5-nitroindazole and proteins with PDB ID **A** 6G77, **B** 1AQ1, **C** 1K3A. The complexes have shown various interactions, and a new type of interaction, many water bridges, has been introduced

Conclusion

The food and drug administration has approved almost 100 drugs against SCLC and NSCLC, which are being used actively. However, this is unfortunate; despite so much expenditure, the world frequently faces the drug resistance problem and needs the multitargeted drug. This study includes multisampling algorithms based on screening, the pharmacokinetics of identified compound, interaction pattern, WaterMap for each P-L complex, and MD simulation for 100 ns in the SPC water medium that identified 5-nitroindazole as a multitargeted inhibitor for CDK2 and other lung cancer transferase kinase. We validated with computational methods and proven as a prominent candidate with less chance of developing resistance, or it might take longer. It can be experimentally validated and used for the welfare of humankind.

Acknowledgements The authors would like to thank Jamia Millia Islamia, New Delhi, for providing computational resources and software solutions.

Author contribution SA contributed to conceptualisation, data collection/curation, analysis, writing, and extensive editing of the first draft. KR contributed to supervision, computational resources, reviewing, and editing.

Declarations

Conflict of interest The authors declare no potential competing or conflict of interest.

Ethical approval Since this study is entirely in silico, ethical obligations are not applicable because they do not directly involve humans or other organisms.

Consent for publication Both authors agree to submit the manuscript to the journal.

References

1. Organization WH. WHO report on cancer: setting priorities, investing wisely and providing care for all. 2020. <https://www.who.int/publications/i/item/9789240001299>.
2. Organization WH. Gender in lung cancer and smoking research. 2004. <https://apps.who.int/iris/handle/10665/43086>.
3. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A et al (2021) Cancer statistics for the year 2020: an overview. *Int J Cancer* 149(4):778–789. <https://doi.org/10.1002/ijc.33588>
4. Sharma R (2022) Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of lung cancer in 2020 and 2050. *Int J Clin Oncol* 27(4):665–675. <https://doi.org/10.1007/s10147-021-02108-2>
5. Minna JD, Roth JA, Gazdar AF (2002) Focus on lung cancer. *Cancer Cell* 1(1):49–52. [https://doi.org/10.1016/S1535-6108\(02\)00027-2](https://doi.org/10.1016/S1535-6108(02)00027-2)
6. Suzuki K, Watanabe S-i, Wakabayashi M, Saji H, Aokage K, Moriya Y et al (2022) A single-arm study of sublobar resection for ground-glass opacity dominant peripheral lung cancer. *J Thorac Cardiovasc Surg* 163(1):289–301.e2. [https://doi.org/10.1016/S1535-6108\(02\)00027-2](https://doi.org/10.1016/S1535-6108(02)00027-2)
7. Shaik NA, Al-Kreathy HM, Ajabnoor GM, Verma PK, Banaganapalli B (2019) Molecular designing, virtual screening and docking study of novel curcumin analogue as mutation (S769L and K846R) selective inhibitor for EGFR. *Saudi J Biol Sci* 26(3):439–448. <https://doi.org/10.1016/j.sjbs.2018.05.026>
8. Viktorsson K, Lewensohn R, Zhivotovsky B (2014) Systems biology approaches to develop innovative strategies for lung cancer therapy. *Cell Death Dis* 5(5):e1260-e. <https://doi.org/10.1038/cddis.2014.28>
9. Gazdar A (2009) Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene* 28(1):S24–S31. <https://doi.org/10.1038/onc.2009.198>
10. Yarden Y (2001) The EGFR family and its ligands in human cancer: signalling mechanisms and therapeutic opportunities. *Eur J Cancer* 37:3–8. [https://doi.org/10.1016/S0959-8049\(01\)00230-1](https://doi.org/10.1016/S0959-8049(01)00230-1)
11. Schraufnagel DE, Balmes JR, Cowl CT, De Matteis S, Jung S-H, Mortimer K et al (2019) Air pollution and noncommunicable diseases: a review by the forum of international respiratory societies' environmental committee, part 2: Air pollution and organ systems. *Chest* 155(2):417–426. <https://doi.org/10.1016/j.chest.2018.10.041>
12. Wang J, Jiang Y, Liang H, Li P, Xiao H, Ji J et al (2012) Attributable causes of cancer in China. *Ann Oncol* 23(11):2983–2989. <https://doi.org/10.1093/annonc/mds139>
13. Liu T-C, Jin X, Wang Y, Wang K (2017) Role of epidermal growth factor receptor in lung cancer and targeted therapies. *Am J Cancer Res* 7(2):187
14. Prabhu VV, Prabhu V (2017) Epidermal growth factor receptor tyrosine kinase: a potential target in treatment of non-small-cell lung carcinoma. *J Environ Pathol, Toxicol Oncol*. <https://doi.org/10.1615/jenvironpatholtoxicoloncol.2017018341>
15. Tan L, Zhang J, Wang Y, Wang X, Wang Y, Zhang Z et al (2022) Development of dual inhibitors targeting epidermal growth factor receptor in cancer therapy. *J Med Chem* 65(7):5149–5183. <https://doi.org/10.1021/acs.jmedchem.1c01714>
16. Jiang W, Cai G, Hu PC, Wang Y (2018) Personalized medicine in non-small cell lung cancer: a review from a pharmacogenomics perspective. *Acta Pharm Sin B* 8(4):530–538. <https://doi.org/10.1016/j.apsb.2018.04.005>
17. Levitzki A, Klein S (2010) Signal transduction therapy of cancer. *Mol Aspects Med* 31(4):287–329. <https://doi.org/10.1016/j.mam.2010.04.001>
18. Facchinetti F, Rossi G, Bria E, Soria J-C, Besse B, Minari R et al (2017) Oncogene addiction in non-small cell lung cancer: focus on ROS1 inhibition. *Cancer Treat Rev* 55:83–95. <https://doi.org/10.1016/j.ctrv.2017.02.010>
19. Lu X, Yu L, Zhang Z, Ren X, Smail JB, Ding K (2018) Targeting EGFR L858R/T790M and EGFR L858R/T790M/C797S resistance mutations in NSCLC: Current developments in medicinal chemistry. *Med Res Rev* 38(5):1550–1581. <https://doi.org/10.1002/med.21488>
20. Sullivan I, Planchard D (2017) Next-generation EGFR tyrosine kinase inhibitors for treating EGFR-mutant lung cancer beyond first line. *Front Med* 3:76. <https://doi.org/10.3389/fmed.2016.00076>
21. Chrysostomou S, Roy R, Prisci F, Thamlikitkul L, Chapman KL, Mufti U et al (2021) Repurposed floxacins targeting RSK4 prevent chemoresistance and metastasis in lung and bladder cancer. *Science Translational Medicine* 13(602):eaba4627. <https://doi.org/10.1126/scitranslmed.aba4627>
22. Lawrie AM, Noble M, Tunnah P, Brown NR, Johnson LN, Endicott JA (1997) Protein kinase inhibition by staurosporine revealed

- in details of the molecular interaction with CDK2. *Nat Struct Biol* 4(10):796–801. <https://doi.org/10.1038/nsb1097-796>
23. Favelyukis S, Till JH, Hubbard SR, Miller WT (2001) Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. *Nat Struct Biol* 8(12):1058–1063. <https://doi.org/10.1038/nsb721>
 24. Ahmad S, Bano N, Qazi S, Yadav MK, Ahmad N, Raza K (2022) Multitargeted molecular dynamic understanding of butoxypheneser against SARS-CoV-2: an in silico study. *Nat Prod Commun* 17(7):1-934578X221115499. <https://doi.org/10.1177/1934578X221115499>
 25. Ahmad S, Bhanu P, Kumar J, Pathak RK, Mallick D, Uttarkar A et al (2022) Molecular dynamics simulation and docking analysis of NF- κ B protein binding with sulindac acid. *Bioinformation* 18(3):170–179. <https://doi.org/10.6026/97320630018170>
 26. Ahmad S, Pasha KM, Raza K, Rafeeq MM, Habib AH, Eswaran M et al (2022) Reporting dinaciclib and theodrenaline as a multitargeted inhibitor against SARS-CoV-2: an in-silico study. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2022.2060308>
 27. Alghamdi YS, Mashraqi MM, Alzamami A, Alturki NA, Ahmad S, Alharthi AA et al (2022) Unveiling the multitargeted potential of N-(4-Aminobutanoyl)-S-(4-methoxybenzyl)-L-cysteinyglycine (NSL-CG) against SARS CoV-2: a virtual screening and molecular dynamics simulation study. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2022.2110158>
 28. Ramlal A, Ahmad S, Kumar L, Khan FN, Chongtham R (2021) From molecules to patients: the clinical applications of biological databases and electronic health records. *Translational bioinformatics in healthcare and medicine*. Academic Press, London, pp 107–25. <https://doi.org/10.1016/B978-0-323-89824-9.00009-4>
 29. Yadav MK, Ahmad S, Raza K, Kumar S, Eswaran M, Pasha KMM (2022) Predictive modeling and therapeutic repurposing of natural compounds against the receptor-binding domain of SARS-CoV-2. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2021.2021993>
 30. Schrödinger Release 2023–1: Maestro, Schrödinger, LLC, New York, NY, 2021. <https://www.schrodinger.com/citations>.
 31. Schrödinger Release 2023–1: Protein preparation wizard; Epik, Schrödinger, LLC, New York, NY, 2021; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2021. 32. Release S. 1: Epik.(2020). Schrödinger Release. 2020;1. <https://www.schrodinger.com/citations>.
 32. Schrödinger Release 2023–1: Prime, Schrödinger, LLC, New York, NY, 2021. <https://www.schrodinger.com/citations>.
 33. Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W et al (2021) OPLS4: Improving force field accuracy on challenging regimes of chemical space. *J Chem Theor Comput* 17(7):4291–4300. <https://doi.org/10.1021/acs.jctc.1c00302>
 34. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR et al (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 46(1):1074–82. <https://doi.org/10.1093/nar/gkx1037>
 35. Schrödinger Release 2023–1: LigPrep, Schrödinger, LLC, New York, NY, 2021. <https://www.schrodinger.com/citations>.
 36. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT et al (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47(7):1750–9. <https://doi.org/10.1021/jm030644s>
 37. Schrödinger Release 2023–1: QikProp, Schrödinger, LLC, New York, NY, 2021. <https://www.schrodinger.com/citations>.
 38. Targowska-Duda KM, Maj M, Drączkowski P, Budzyńska B, Boguszewska-Czubara A, Wróbel TM et al (2022) WaterMap-guided structure-based virtual screening for acetylcholinesterase inhibitors. *ChemMedChem* 17(8):e202100721. <https://doi.org/10.1002/cmdc.202100721>
 39. Schrödinger Release 2023–1: WaterMap, Schrödinger, LLC, New York, NY, 2021. <https://www.schrodinger.com/citations>.
 40. Cappel D, Sherman W, Beuming T (2017) Calculating water thermodynamics in the binding site of proteins—applications of WaterMap to drug discovery. *Curr Top Med Chem* 17(23):2586–2598. <https://doi.org/10.2174/1568026617666170414141452>
 41. Schrödinger Release 2023–1: Desmond molecular dynamics system, D. E. Shaw Research, New York, NY, 2021. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2021. <https://www.schrodinger.com/citations>.
 42. McDonald I (1972) NpT-ensemble Monte Carlo calculations for binary liquid mixtures. *Mol Phys* 23(1):41–58. <https://doi.org/10.1080/00268977200100031>
 43. Karwasra R, Ahmad S, Bano N, Qazi S, Raza K, Singh S et al (2022) Macrophage-targeted punicalagin nanoengineering to alleviate Methotrexate-Induced Neutropenia: a molecular docking, DFT, and MD simulation analysis. *Molecules* 27(18):6034. <https://doi.org/10.3390/molecules27186034>
 44. Ahmad S, Sayeed S, Bano N, Sheikh K, Raza K (2022) In-silico analysis reveals quinic acid as a multitargeted inhibitor against cervical cancer. *J Biomol Struct Dyn*. <https://doi.org/10.1080/07391102.2022.2146202>
 45. Famuyiwa SO, Ahmad S, Fakola EG, Olusola AJ, Adesida SA, Obagunle FO et al (2023) Comprehensive computational studies of naturally occurring kuguacins as antidiabetic agents by targeting visfatin. *Chem Afr*. <https://doi.org/10.1007/s42250-023-00604-8>
 46. Shah AA, Ahmad S, Yadav MK, Raza K, Kamal MA, Akhtar S (2023) Structure-based virtual screening, molecular docking, molecular dynamics simulation, and metabolic reactivity studies of quinazoline derivatives for their anti-EGFR activity against tumor angiogenesis. *Curr Med Chem*. <https://doi.org/10.2174/0929867330666230309143711>
 47. Tripathi MK, Ahmad S, Tyagi R, Dahiya V, Yadav MK (2022) Fundamentals of molecular modeling in drug design. *Computer aided drug design (CADD): From ligand-based methods to structure-based approaches*. Elsevier, Amsterdam, pp 125–55. <https://doi.org/10.1016/B978-0-323-90608-1.00001-0>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.