Biomacromol. J., Vol.12, No. 1, 92-106, January 2026.

Biomacromolecular Journal
www.bmmj.org

Molecular docking, molecular dynamics studies, and MM/GBSA calculation on some of the tyrosine kinase inhibitors

Zohre Ghiasi, and Davood Ajloo*

School of Chemistry, *Damghan University, Damghan, Iran.*

(Received 10 June 2025, Accepted 10 September 2025, Published 30 January 2026)

ABSTRACT

Cancer is a life-threatening ailment characterized by the uncontrolled proliferation of cells. Because cancer is not just a disease, it is unlikely that there will ever be a single cure for it. Therefore, the need for drugs to combat this disease has increased. Worldwide efforts from scientists are underway to determine the causes of cancer and reduce mortality rates. Tyrosine kinase inhibitors (TKIs) are widely used in tumor treatment. The screened compounds were followed for SP docking, XP docking, MM-GBSA analysis, induced-fit (IFD) docking, and MD simulation. The absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of all compounds were analyzed, and a final selection was made based on the Lipinski rule of five. The ligand Imatinib revealed the highest docking score of -12.560 kcal/mol. To further validate the interactions of top-scored receptors and ligands, a molecular dynamics study of 10 ns was carried out. This indicated that the protein-ligand complex was stable throughout the simulation period, and minimal backbone fluctuations ensued in the system. Post-MM-GBSA analysis of molecular dynamics data showed a free binding energy of -68.392 kcal/mol. The MD simulation studies confirmed the stability of the protein-ligand complex. In this study, it was concluded that the imatinib drug is known as the best cancer drug among other drugs and can be used for further studies and more calculations. This molecule may emerge as a promising ligand against cancer and thus needs further detailed investigations.

Keywords: Tyrosine kinase inhibitors, molecular docking, molecular dynamic, MM-GBSA

INTRODUCTION

Cancer is one of the most common and deadly diseases around the world, and it has numerous different types. Cancer begins when cells in a part of the body grow out of control. Several other parts of the body may be affected by cancer. Cancer incidence increased to 23.6 million cases worldwide, which led to more than 10.0 million deaths in 2019 [1]. Due to advances in cancer early detection and cancer treatments, yearly cancer mortality has been decreasing since 1995 [2].

Consequently, the imperative for early detection and prognosis has emerged as a means to enhance long-term survival rates and mitigate mortality. The development of small molecules to target kinases has significantly increased due to the success of imatinib, the first tyrosine kinase inhibitor. TKIs are the first-line small-

***Corresponding author. E-mail: ajloo@du.ac.ir**

molecule drugs in many cancer treatments [3], effectively suppressing the aberrant cell growth and proliferation caused by hyperactive tyrosine kinases through inhibiting cell signal transduction pathways.

Innovations in the field of cancer therapies have led to the effective management of different cancers that were previously considered to be incurable. Although these treatments have dramatically changed the natural course of many cancers, they may result in cardiac and extracardiac complications, which can be manifested either during therapy or after the completion of treatment. Undoubtedly, TKIs have dramatically improved long-term survival outcomes and the quality of life for patients [4]. Notable examples among the widely employed TKIs include dasatinib, ibrutinib, ponatinib, sunitinib, erlotinib, lapatinib, nilotinib, sorafenib, and so on [5]. TKIs are agents that inhibit the enzyme tyrosine kinase. This enzyme transfers phosphate groups on adenosine triphosphate (ATP) to the tyrosine residues of protein by phosphorylation, sending signals to regulate cell growth and differentiation [6]. There are many receptors at the cell membrane with tyrosine kinase activity. TKIs display substantial differences concerning the spectrum of kinase inhibition, clinical application, pharmacokinetics, pharmacodynamics, and the spectrum of possible drug–drug interactions (DDIs) [7]. TKIs are a group of pharmacological agents that disrupt the signal transduction pathways of protein kinases through several modes of inhibition. The advent of imatinib (the most representative medicine for TKIs) has presented a new era of targeted therapy for tumors [8]. Furthermore, information on exposure, therapeutic efficacy, and side effects of many oral molecular-targeted anticancer drugs needs to be collected, and studies on pharmacokinetics/pharmacodynamics analysis are underway [9]. The physicochemical properties and therapeutic concentration range of each analyte greatly differ.

Traditional research and the development of new drugs are a long-cycle, high-risk, and large investment process. On average, the development of a new drug takes about 10–15 years, and the investment cost is about USD 800 million [10-12]. However, the emergence of computer-aided drug design (CADD) technology has changed this landscape. The application of CADD technology can guide the rational development of new drugs, reduce blindness and contingency, speed up the drug development process, and, ultimately, save human, material, and financial resources. With the development of computing power and efficient algorithms, CADD techniques have been widely and successfully applied in the field of drug discovery, rapidly advancing the development of related drugs [13-15]. As powerful auxiliary means for the rapid development of new drugs, some CADD techniques, including structure-based drug design (such as molecular docking and molecular dynamics simulations) and ligand-based drug design methods (such as 3D-QSAR and pharmacophore model screening), have also been applied in the research. This study aims to compare a group of tyrosine kinase drugs and obtain the best drug in terms of molecular docking and molecular dynamics calculations for further research.

In this study, we are trying to prepare 33 compounds from the tinib and fenib family, perform molecular docking and molecular dynamics calculations, and report the results. This study is the first to combine MM/GBSA with induced fit docking for the optimization of virtual screening of TKIs. This study presents a novel paradigm for screening and optimization of tyrosine kinase inhibitors (TKIs) by combining state-of-the-art structure-based computational methods using standard molecular docking (SP), high-precision docking

(XP), and induced fit docking (IFD). Induced fit docking (IFD) was used to account for protein flexibility and improve ligand interaction prediction. This combined approach reduced false positives and increased interaction prediction accuracy. MM-GBSA was used to calculate the binding free energy more accurately. This method is more accurate than conventional docking score and better evaluates electrostatic, hydrophobic, and hydrogen bonding interactions. Combining IFD and MM/GBSA for virtual screening optimization increases the accuracy of interaction prediction and allows for faster identification of drug candidates. Molecular dynamics (MD) simulations were used to investigate the stability of the complex.

MATERIALS AND METHODS

The structures of more than 33 compounds were collected and compiled into a database. Standard Precision docking (SP) docking was used to identify compounds with a higher binding affinity, resulting in the identification of 67 molecules. Following the SP docking results, a pharmacophore analysis was performed to identify and examine the common features of the compounds' docking conformations. Furthermore, these 65 compounds, utilizing XP docking and MM-GBSA Analysis, were further used to screen 33 molecules with good binding affinity. Induced Fit Docking (IFD) was performed to identify the interaction between these 33 molecules and with protein. Based on the ligand-receptor interactions and structure type (Table 1), the better ligand-receptor complexes were subjected to molecular dynamics (MD) simulation to assess their stability, and the ADME/T properties of these 18 molecules were evaluated. After careful evaluation, one compound with good MD stability and reasonable ADME/T predictions was selected.

Ligand Preparation

The ligands used as input for the binding study were plotted using Hyperchem software, and then the ligands were used for calculations. The LigPrep module, packaged by Schrödinger, was used to prepare ligands by applying certain filters and generating a customized ligand library. The preparation of the ligands is carried out by adjusting the torsion of the ligand and then assigning the protonation states [16]. Subsequent refinement of 3D structures by the addition of hydrogen atoms, generation of stereoisomers, and identification of ionization states led to ligand refinement. Furthermore, the OPLS3 force field incorporated in PHASE was used for energy minimization and to optimize the low-energy 3D structure of the ligand [17]. This minimization helps assign bond orders, add hydrogens to the ligands, and convert the 2D to 3D structure for the docking studies. The generated output file (Best conformations of the ligands) was further used for docking studies. Figure. 1 shows some of the ligands used in the Schrödinger calculations.

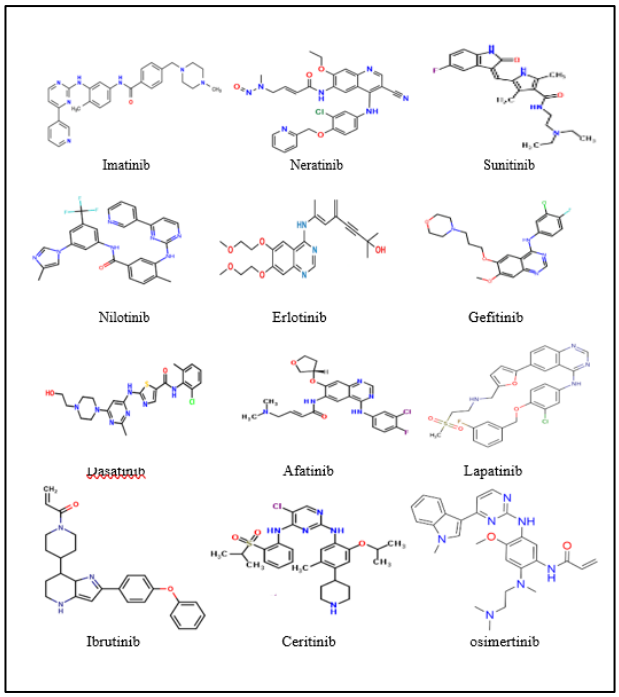

Figure 1. The chemical structures of 33 ligands used in this study showed the best docking scores.

Protein preparation

The crystal structure of the protein (PDB ID: 2HYY) was retrieved from the Protein Data Bank (PDB). 2HYY has a good resolution of 2.40 A˚. A human polypeptide enzyme, along with ABL, is a member of the tyrosine kinase family, usually in a complex with a ligand or drug. This structure shows how the function of ABL is regulated by its interaction with a specific substance. The protein structure was prepared using a protein preparation wizard (PPrep) module, implemented in the Maestro module of Schrödinger software [18]. This study obtained the protein structure from the PDB database (https://www.rcsb.org, ID: 2HYY).

This ensures high-confidence structural accuracy and converts the protein structure from crude to refined, sufficiently prepared for molecular docking and molecular dynamics studies. This process involved determining modified bond arrangements, removing water molecules and other non-specific chemical components from the crystal structure, and addition of hydrogen atoms to the protein structure for remodeling the tautomeric and ionization states of amino acid residues [19].

A protein preparation wizard was utilized to pre-process the protein structure complex by removing the water molecules from the complex, positioned within 5 Å of the heteroatom groups. Charges were assigned, and Het states using Epik at pH 7.0 ± 2.0. Pre-process the protein and refine, and modify the protein by analyzing the workspace water molecules and others. The subsequent addition of missing hydrogen atoms and hydrogen bond optimization [18-19] was followed by restrained energy minimization of protein structure by employing the Optimized Potentials for Liquid Simulations (OPLS)3 force field to achieve high accuracy. It is a widely used force field in molecular dynamics (MD) simulations, especially for protein and biomolecular systems. Although improvement of the OPLS protein force field has received less attention than for CHARMM and AMBER over the past decade, the major revision of the protein force field in OPLS3 reported here appears to display performance that is competitive with the state of the art. [20-21].

After the final step of docking with the co-crystal ligand in extra precision (XP) mode, the root mean square deviation (RMSD) was checked to validate the protein. Finally, the shape and property of the protein were presented by computing a grid box of (20 × 20 × 20 Å) using the receptor grid generation option available in the Glide module of Schrödinger software [22].

ADME/T and Lipinski parameters

The drug-likeness of all compounds was evaluated by QIKPROP to determine their Absorption, distribution, metabolism, excretion, and toxicity (ADME/T) properties. ADME/T is a very important evaluation standard in contemporary drug design and drug screening. Any compound that exhibits drug-likeness must have moderate ADME/T properties. QikProp predicts all physically relevant and pharmaceutically essential properties of a compound by comparing the properties with 95% of already known drug molecules [23].

To filter these compounds according to the Lipinski rule [24], the ADME/T properties analysis was conducted. According to the rule, molecules with Molecular Weight ≤ 500, hydrogen bond donors ≤ 5, and acceptors ≤ 10, calculated octanol-water partition coefficient, and log P ≤ 5 possess good membrane permeability [25]. The Physicochemical parameters of the obtained hits after the docking studies were in silico predicted using the Qikprop module of Schrodinger. Properties

assessed involved molecular weight (MW), brain/blood partition coefficient, lipophilicity (log Po/w), hydrophobicity, and hydrophilicity (QPPCaco, QPlogBB, and QPPMDCK), human oral absorption (HOA%). total solvent accessible surface area (SASA), number of hydrogen bond donors (HBD), number of hydrogen bond acceptors (HBA), octanol/water partition coefficient (log P), predicted apparent Caco-2 cell permeability in nm/sec (P Caco), and number of rotatable bonds (Rot).

PSA (Van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms), QPlogS (aqueous solubility, log S. S in mol dm-3 is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid), QPlogPo/w (octanol/water partition coefficient), CNS (central nervous system activity), #metab (number of likely metabolic reactions), HumanOralAbsorption, QPlogBB (brain/blood partition coefficient), QPPMDCK (apparent MDCK cell permeability in nm/sec), QPPCaco (apparent Caco-2 cell permeability in nm/sec) and QPlogHERG (IC50 value for blockage of HEGR K+ channels). Based on the collective data, we infer that the ADME/T properties of compounds are within the prescribed limits for a potential candidate drug compound. Parameters play an important role in the overall transport binding of the drug at the target site. Prediction of these properties also helps to understand metabolic stability and the fate of the molecules in the body. All these values suggested that the identified hits from the databases fall under the acceptable range of ADME parameter values, and their chemical structures and features can be further utilized for the design of potent inhibitors. Results of in silico pharmacokinetic study on these natural products showed good pharmacokinetic properties. CNS activity of these natural products was evaluated.

Grid generation

Receptor grid generation of the Maestro Suite was used to create the active binding site of the protein structure around the residues ASP 381, GLU 286, THR 315, and MET 318. The site map module of maestro Schrödinger suite was applied to validate the grid box, which was generated at (X: 14.04, Y: 14.61, Z: 17.05) and diameter midpoint box of (X, Y, Z: 10 Å), site maps with site score > 1.088 were more valid than the others. The ligand docking was performed using the Glide of the Maestro suite in two precision steps. A large number of ligands were screened quickly employing high-throughput virtual screening (HTVS), followed by the extra precision (XP) method used to dock the best 10% of poses with excess precision. Flexible ligand sampling was applied to both protocols, and the output result was expressed as a docking score. Grid generation is essential in molecular docking studies because we must specify the binding pocket in proteins before docking (Figure. 2). The best way is to pick up the active sites of the protein from the crystal structure or the published research papers reported earlier.

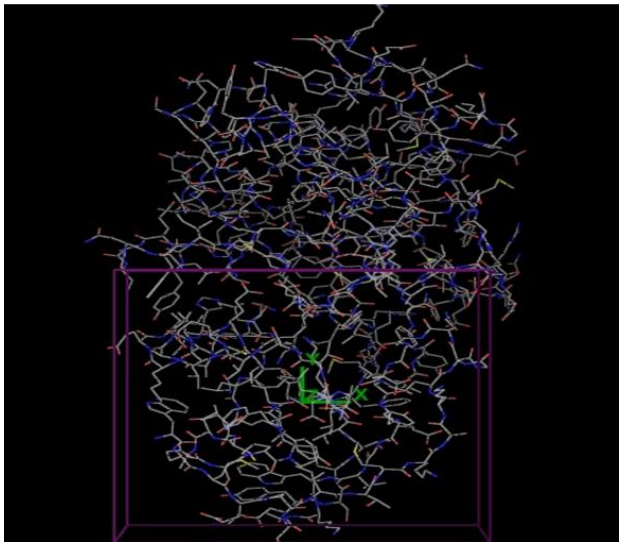

Figure 2. Grid box resulting from active site identification.

Calculation of binding free energy using MM-GBSA

To evaluate the free binding energy between the protein and docked ligand, the MMGBSA of the Prime module was employed [26]. The prime module of Schrodinger software was used to calculate the optimal binding energy of the selected complexes, whose docking score was the lowest among all. The prime module of Schrödinger was used to estimate the theoretical binding free energies of the hits obtained from XP docking, based on the glide scores, which were utilized for MM-GBSA analysis [27]. Ligand-protein complex binding free energy was calculated using two equations: molecular mechanics generalized Born surface area (MM-GBSA) and molecular mechanics Poisson–Boltzmann surface area (MM-PBSA). The Molecular Mechanics/Generalized Born Surface Area dG (MMGBSA dG) for the optimized Receptor-Ligand Complex was calculated to determine the ligand binding affinities. During the calculation of binding free energy, the VSGB solvation model and OPLS3 force field were applied. For the analysis, the VSGB (a novel model for high-resolution protein modeling) model was exploited, having an OPLS-3 force field inclusive of an implicit solvent model in addition to physics-based modifications for p-p interactions, hydrophobic interactions, and hydrogen bonding self-contact interactions. The binding energy was calculated based on the following equation:

$$\Delta G_{binding} = E_{complex(minimized)} - E_{ligand(minimized)} - E_{receptor(minimized)}$$

The MM-GBSA calculations entail maintaining the rigidity of all protein atoms while relaxing the compound's atoms. Furthermore, the ranking of protein compound complexes is carried out using binding free energy calculations. During MM-GBSA energy calculation, many jobs run like optimization of receptor/ligand/complex, receptor from optimized complex, and ligand from optimized complex. MM-GBSA requires a pose-viewer file in pv. maegz format that can be obtained from the XP mode of docking, and the distance from the receptor was set to 5.0 Å [28].

Molecular docking studies

Docking studies of the database compounds were performed using the Glide module of Schrodinger software, which was installed. The selected target protein structure was retrieved from the RCSB protein data bank. The minimized molecule was docked with Ligand Docking in Glide. All molecules bind to the protein's active site by SP/XP docking and flexible docking. Post-docking minimization was used to optimize the ligand-protein complexes. The docking results are analyzed and displayed in Maestro. The docking affinity energy of the compound with the protein is calculated by the following equation: where R is the Boltzmann gas constant (R = 1.987 cal/mol. K), T is the default temperature of simulated docking (T = 298 K), and Ka is the binding free affinity of docking.

The docking studies revealed the presumed binding modes in the active site of the selected targets and exhibited the maximum docking scores. The co-crystal ligands of the selected targets revealed lower docking scores when compared with the top-hit natural product database compounds. The compound with the highest binding score against the target (PDB 2HYY) was selected for comprehensive analysis.

High-throughput virtual screening

The compounds with the best fitness scores were subsequently subjected to high-throughput virtual screening (HTVS). HTVS is quick and more tolerant to suboptimal fits than standard precision (SP) and extra precision (XP) docking studies. The retrieved compounds and 20 × 20 ×

20 Å receptor grid were subjected to a virtual screening workflow in the Glide module [29]. The retrieved ligands were subsequently processed for SP and XP docking. The scaling factor for van der Waals radii was set to 0.80 Å, with a partial charge cut-off of 0.15 Å. In each step of docking, the top 10% of ligands were retrieved with the highest docking score. Finally, 15 compounds were retrieved based on their binding free energies and docking poses at the catalytic site. Structure-based high-throughput screening was employed to minimize the number of false positives retrieved from the database. This workflow employed fast preliminary screening by HTVS, followed by accurate docking studies using SP or XP mode.

Standard precision docking

To evaluate the affinity of the ligands to the protein in the active site, molecular docking with standard precision (SP docking) was performed on all 33 molecules. The molecular structures of the compounds were drawn in 2D Sketcher of Maestro and subsequently prepared for docking through the use of LigPrep. Before docking, the protein underwent preparation and energy optimization with the aid of the Wizard in Glide. The minimized molecules were then docked to the grid box, utilizing standard precision docking.

Extra precision docking

The extra precision docking (XP docking) mode is more advanced and filters out false positives, and provides a better association between docking score and good poses [30]. Therefore, XP docking was employed to re-dock the top 65 molecules, including imatinib, with protein and rank compounds by their XP binding energy.

Induced-fit docking (IFD)

Induced-fit docking (IFD) protocol was used for accurately redocking selected compounds in a relaxed residue binding pocket. Standard and OPLS-3 were employed as protocols and force fields, respectively. Eventually, the IFD score was explained in terms of kcal/mol.

The standard docking protocol sometimes misleads the docking results, as many proteins undergo side-chain or backbone movements on ligand binding. This computational limitation has been overcome by IFD. The IFD study was performed for the compound after SP docking. IFD offers more accurate receptor-ligand interactions, as it adjusts receptor flexibility according to the ligand binding modes. The IFD follows three consecutive steps: ligand docking into a rigid receptor; a maximum of 33 ligand poses were considered for protein structural refinements; subsequently, a grid box was centered on the active site [31]. Furthermore, the prime module was used to generate induced-fit protein-ligand complexes. 65 ligand poses from SP docking were subjected to protein side chain and backbone refinements. For protein structure refinement and conformational search, residues within 5 Å were considered, while residues outside the range were fixed. Finally, the ligands were re-docked with receptor structures with lower energy in the Glide XP module with the default setting. The IFD score was used to calculate the protein-ligand interaction energy as well as to rank IFD poses. The potential hits with the highest IFD scores exhibit favorable binding interactions and are considered for the next screening steps. The protein-ligand interactions were visualized using the Discovery Visualizer and Ligand Interaction Diagram module available in Maestro. Flexible docking techniques have been employed to determine the binding interactions of designed compounds. The ligand and receptor residues are imparted with flexibility, which improves the binding site prediction of the molecules.

Regarding ligand-receptor interaction, the conformation of the receptor, especially around the binding site, was also induced to be altered, better matching the shape and binding pattern of

the ligand molecule. In this study, IFD is applied for further virtual screening to obtain more realistic models of protein-molecule interactions. The 65 compounds were chosen to be re-docked into the protein by applying Induced Fit Docking (IFD) as a control and their Induced Fit Docking binding energy. Multiple docking results are produced during IFD docking due to the various poses of the compounds, and we considered the results with the highest docking energy of each compound. IFD docking facilitated stronger binding and more protein interactions with filtered molecules than SP or XP docking, further highlighting the affinity differences between these compounds and proteins.

Molecular dynamics simulation

Virtual screening and molecular docking have shown the binding ability of different compounds to the protein. In contrast, molecular dynamics simulations can provide insights into the dynamic stability of receptor-ligand complexes under physiological conditions. Thus, based on the binding affinity, ligand-receptor complex interaction, and structural diversity of compounds, one compound was selected to further carry out molecular dynamics simulations to investigate its binding stability. Throughout the molecular dynamics simulation, one frame of trajectory was generated every 10 ps, resulting in 1000 frames after 10 ns of operation. A thorough analysis of the resulting data was conducted, including Root Mean Square Deviation (RMSD) and Protein-Ligand Contacts analysis. The conformation of the ligand-protein complex at t = 0 ns was used as a reference for the calculation of RMSD values for all 1000 frames. Fluctuations in RMSD values are usually associated with changes in ligand-protein interaction. As for another ligand-protein complex with good dynamic stability, they showed stable RMSD values over long periods.

The stability of the protein-ligand complexes and various binding modes of the ligand were determined using an MD simulation study of 10 ns for the complex. The highest-scoring compound in the complex was submitted to MD simulations. The MD simulation was carried out using the Macro Model program from the Schrödinger software package. Moreover, the root mean square deviation (RMSD) was calculated to evaluate the structure's optimum energy. The conformation state was considered with a maximum half angstrom of superimposition error in RMSD. Ligand energy minimization was carried out by an OPLS3 force field in an aqueous solvent. Using Polak-Ribier Conjugate Gradient (PRCG) methods, the MD of the protein is analyzed using stochastic dynamics mechanisms. Dynamic simulation conditions were provided at 300 K, the time step of 1.5 femtoseconds, the equilibration time of 1 picosecond, and the simulation time of 10000 picoseconds. Finally, potential calculations were performed using the OPLS3 force field with solvent water [32]. The RMSD analysis offers insights into the extent of structural deviation and conformational stability observed throughout the MD simulation period [33].

One of the most critical variables in a dynamic system is the temperature, which allows the computational analysis to push through the relative potential and get closer to the global minimum. The criteria to achieve equilibrium and convergence in the dynamic system have been described in the terminal cycles to reproduce comparable energies. Using quantum dynamics can lead to increased accuracy and calculations, but quantum calculations are very expensive and require powerful and advanced computer systems for processing. One way to use Schrödinger software is that it is not quantum dynamics, and we use Newtonian dynamics.

RESULTS AND DISCUSSIONS

ADMET

The design of new kinase-selective inhibitors and analysis of the compounds using ADMET filters and molecular docking studies may be used to summarize the current investigation. The analysis results of 19 selected compounds are shown in Table 1. The first column is related to the molecular weight gr/mol of the compounds. The lower molecular weight of the compound is related to the higher percentage of oral absorption of the compound. Molecular weights less than 480 daltons reduce the percentage of absorption. The best Molecular weight for this group of compounds is 480 daltons.

The second column is the solubility of the compounds in water. Factors such as very high molecular weight and volume will reduce the percentage of oral absorption of the compounds. Recommended values are -6.5 to -0.5.

QPlogPo/w and QPlogS are important parameters that give an idea of the hydrophobic properties and aqueous solubility of the molecule. All compounds possess at least one hydrogen bond donor and four hydrogen bond acceptors. QPPCaco is a prediction of Caco-2 nm/s cell permeability in nm/s, which is poor for values <25 and excellent for values >500. The results of this study suggested that all the retrieved hits were predicted with QPlogPo/w values in the range of 2.908 to 5.806, QPlogS values in the range of −7.912 to 670.985 mol/dm3, and QPPCaco cm/s values in the range of 39.855 – 4844.87. According to this study, the partition coefficient (QPlog Po/w) was within the permissible range for the selected natural products. The calculated PSA Å2 was within the range of 45.42-122.9 Å. It has been suggested that the log P value must be less than five, and high log P results in poor absorption.

The human oral absorption percentage of compounds was in the appropriate range of 63 to 100%. Sorafenib, regorafenib, quizartinib, motesanib, ceritinib, cabozantinib, imatinib, axitinib, and afatinib showed 100% oral absorption. Also, all the hit compounds showed over 60% oral absorption. It can be observed in Table 1 that more compounds of Lipinski's rule (polar surface area, molecular weight, number of hydrogen donors, and acceptors) were found for hit compounds, and all mentioned properties were within the allowed range, thereby indicating their potential as drug-like molecules.

Table 1. ADMET results obtained from QikProp of Maestro Schrödinger software.

molecule	mol_MW	QPlogS	QPlogPo/w	accept HB	donor HB	CNS	Percent Human Oral Absorption	QPPCaco	Metab	Human oral absorption	QPlogBB	PSA
vatalanib	501.513	-7.458	5.806	6	0	-2	85.615	670.985	4	1	-1.011	122.90
tandutinib	386.470	-6.820	4.120	5	2	-2	94.908	281.401	1	1	-1.369	86.64
sunitinib	496.651	-6.348	3.859	8.2	0	1	90.051	183.426	2	1	-0.316	72.96
sorafenib	485.945	-7.912	4.673	9.2	2	1	100	371.230	5	1	-0.227	71.56
imatinib	519.816	-7.640	4.932	13.2	6	0	100	16.110	4	1	-0.331	91.81
regorafenib	346.818	-5.463	4.222	4	1	0	100	1783.460	4	3	-0.115	58.53
quizartinib	475.360	-5.544	4.839	6	1	2	100	1028.520	4	3	0.456	45.42
nilotinib	560.670	-6.769	3.870	9.75	2	-2	69.568	69.083	3	1	-1.219	108.80
neratinib	532.567	-4.864	3.570	9.5	1	1	63.533	39.855	6	2	-0.076	74.97
motesanib	437.518	-6.087	3.800	6.5	2	0	100	735.550	6	3	-0.662	84.00
gefitinib	375.404	-7.169	4.203	3	5	-2	83.957	342.183	1	1	-1.326	83.93
erlotinib	426.858	-4.337	2.908	6.58	1.33	2	68.633	126.396	5	3	0.575	63.56
dasatinib	595.087	-7.311	5.237	9.75	1	0	71.395	165.328	5	1	-0.802	94.08
dacomitinib	493.610	-5.118	3.617	10.5	2	1	79.433	56.128	8	2	-0.408	88.64
ceritinib	393.441	-5.169	4.150	7.4	1.5	0	100	4844.870	6	3	-0.474	62.99
cabozantinib	469.945	-6.196	4.396	7.75	2	1	100	359.656	4	3	-0.228	74.56
brivanib	450.342	-5.559	4.049	5.25	3	0	91.803	199.120	5	3	-0.145	74.74
axitinib	443.547	-6.289	4.508	6.25	2	0	100	1293.200	2	1	-0.540	66.16
afatinib	450.512	-6.248	4.895	6	1	1	100	766.047	3	1	0.161	60.60

Aqueous solubility (logS) has been among the most significant factors in the brain/blood partition coefficient and percent human oral absorption. Approximately all of the HTVS results possessed the predicted QPIogBB permissible range (-3 to 1.2), 70% of compounds had up to 80% human oral absorption, and about 99% were within the considered range of predicted QPIogPo/w (-2 to

5). Also, direct relations between QPIogPo/w, QPIogBB, and absorption percentage, and an opposite relation between QPIogPo/w and QPIogS can be seen. Caco-2 cell permeability predicted in silico for compounds vatalanib, regorafenib, quizartinib, motesanib, ceritinib, imatinib, and afatinib was greater than 500, which are excellent values and indicate that we may have a high absorption process.

MM-GBSA calculation

The results and data from free energy calculations are given in Table 2. The first column is related to the binding energy of the ligand and the receptor, taking into account all the binding energies. The next column is related to the calculation of the electrostatic binding energy of the compound with the drug.

Table 2. Results of energy calculations in molecular docking

Molecule	dG_Bind	dG_Bind_ Coulomb	dG_Bind_ Covalent	dG_Bind_ Hbond	dG_Bind_ Lipo	dG_Bind_ Packing
Linifanib	-41.935	-19.850	7.563	-3.096	-24.203	-3.335
Imatinib	-68.392	-21.225	5.560	-2.172	-18.038	-2.855
Neratinib	-68.314	-20.886	4.815	-2.177	-17.860	-2.763
gefitinib	-54.022	-14.761	15.690	-1.016	-33.606	-0.543
Pazopanib	-51.873	-9.315	6.314	-2.488	-25.087	-3.154
Motesanib	-46.399	-11.341	6.065	-1.058	-17.298	-0.624
Axitinib	-33.615	1.838	11.437	-0.392	-19.280	-1.787
Crizotinib	-19.539	6.120	15.230	-0.115	-31.937	-2.287
Ibrutinib	-7.669	-0.677	26.388	-1.244	-30.854	-0.798
Erlotinib	-46.458	-8.272	4.389	-1.122	-18.933	-0.026
Brivanib	-12.144	-1.476	11.734	-0.187	-20.878	-2.113
Crenolanib	-25.284	5.916	4.969	-0.592	-15.130	-1.571
Sunitinib	-3.302	-10.206	27.125	-1.253	-27.935	-0.200
Lenvatinib	-25.186	-7.043	19.298	-1.099	-20.063	-0.751
Vandetinib	-6.426	-0.578	19.971	-0.084	-23.217	-0.012

The next column is related to the covalent binding energy for each compound, which is a positive value because there is no possibility of forming a covalent bond with the receptor. The fourth column shows the hydrogen bond binding energy, which is one of the key indicators of the strength of the compound. The fifth column is related to the lipophilic binding energy. This energy indicates the degree of binding of the compound to the hydrophobic amino acids of the protein active site. The more negative this energy is, the stronger the bonds formed, meaning that most of the protein active site has hydrophobic amino acids, which is a particularly important property in the design of compounds. The last column is related to the foot-foot interactions; the lower the number, the stronger the bonds.

MM-GBSA score of the imatinib molecule was -68.392, which is the largest amount, whereas the MM-GBSA score of the Sunitinib molecule showed a significantly lower value of -3.302 (Table 2). With these results, it can be concluded that the binding free energy score of imatinib was better than the Sunitinib compound. The results indicated that most drugs had binding energy between -3.302 to -68.392 kcal/mol. As the number of hydrogen bonds in a molecular complex increase, the binding energy (the ability of molecules to bind or separate) increases. Each additional hydrogen bond can strengthen the interactions between molecules. Systems with a higher number of hydrogen bonds tend to have more stable structures, especially in the case of proteins and nucleic acids, where hydrogen bonds play a key role in maintaining their three-dimensional structure [34].

Molecular docking

The HTVS docking identified SP docking displayed 67, and XP docking retrieved 65 molecules with a binding energy of ≤ -3.3 kcal/mol. The best-docked molecules exhibited hydrogen bonding interactions with ASP 381 and GLU 286 residues and hydrophobic Pi–Pi interactions with PHE 382 and TYR 253.

The docking scores of the ligand and natural product database compound are -12.560, and the IFD score is -674.37. ASP 381 is the amino-acid residue that commonly plays its role in the active hydrogen bond formation with these hit molecules. Water plays an essential role in the formation of water-mediated interactions. All these interactions strongly suggested that these top-hit molecules may act as potential inhibitors. Upon keen observation of all

hydrogen bond distances with the surrounding amino-acid residues and the water molecules, the distances are very close to the molecules, and these connections are also tightly bonded to the molecules, which may further, in turn, contribute to its significant docking score when compared with its co-crystal ligand (Figure. 3). The calculations obtained from the docking results show and compare the highest binding scores of the four compounds after imatinib in relation to the bonds formed with the protein. As can be seen in structures A to D, the amino acids ASP 381, GLU 286 and GLU 316 have negative charges and the amino acid LYS 271 has a positive charge.

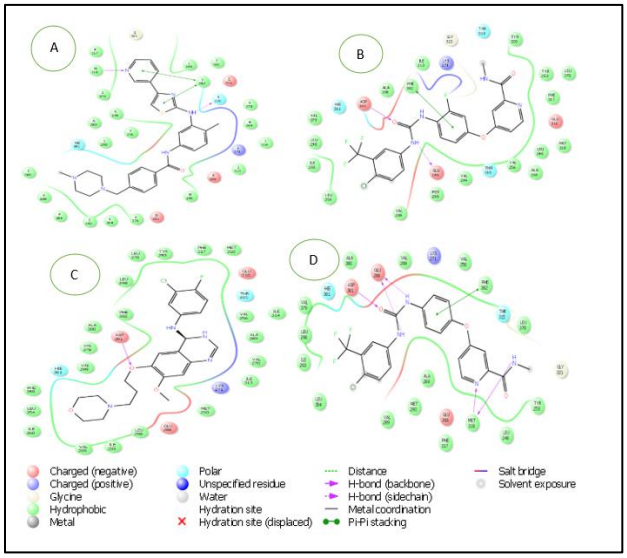

Figure 3. The 2D ligand-protein interaction of reference compound with the active site of amino acid residues: A) Masitinib, THR 315, MET 318, and TYR 353; B) Regorafenib, ASP 381, GLU 280, and PHE 382; C) Gefitinib, ASP 381; D) Sorafenib, GLU 286, ASP 381, MET 318, and PHE 382 hydrogen bond and Pi-Pi-interaction.

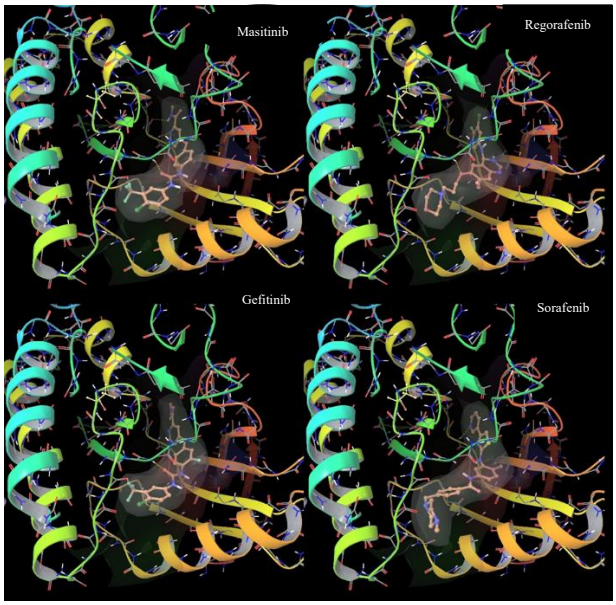

Figure 4. The 3D ligand-protein interaction of the reference compound with the active site of amino acid residues.

As can be seen from Figure. 3 and Figure. 4, part A, masitinib with a docking score of -12.154 as the best docking score after imatinib, formed two hydrogen bonds of length 2.03 with THR 315 residue and 2.40 Å with MET 318 residue, and two Pi-Pi stacking bonds of length 5.34 Å and 5.36 Å with THR 315 residue. The Part B compound, the Regorafenib ligand with a docking score of -11.784, formed two hydrogen bonds of length 2.00 Å with ASP 381 residue and length 2.67 Å with GLU 280 residue, and a Pi-Pi stacking of length 4.70 Å with PHE 382 residue. The compound in part C is gefitinib with a docking score of -11.343, which has only one hydrogen bond of length 1.97 Å with the ASP 381 residue. The last compound, which is the sorafenib ligand, with a docking score of -11.059, was one of four compounds with higher docking scores after imatinib, forming three hydrogen bonds of 2.14 Å with residue ASP 381, two bonds of 2.48 Å and 2.56 Å with residue GLU 286, two bonds of 2.50 Å and 2.74 Å with residue MET 318, and one bond of Pi-Pi stacking of length 4.82 Å with residue PHE 382.

Multiple docking results are produced during IFD docking due to the various poses of the compounds, and we recorded the results with the highest docking energy of each compound in Table S1. Table S1 includes Docking scoring functions and Induced Fit Docking calculation. Compared to SP or XP docking, IFD docking facilitated stronger binding and more protein interactions with filtered molecules, further accentuating differences in affinity between these compounds and proteins. As shown in Table S1, imatinib interacts with protein mainly through hydrogen bonds, consistent with that reported in the literature. For the screened compounds, they can engage more amino acids in protein-ligand interactions and display new interactions. show docking scores of newly designed ligands in the range of -7.799 to -12.560 and glide score in the range of -7.796 to -12.560. To illustrate, compounds Imatinib and Masitinib have docking scores of -12.560, -12.154, and glide scores of -12.560, -12.229, respectively.

Molecular Dynamics

As shown in Table 3, Imatinib interacts with protein mainly through hydrogen bonds, consistent with the literature. The screened compounds can engage more amino acids in protein-ligand interactions and display new interactions. Considering the diversity of structural features and binding interaction models, one molecule (Imatinib) is selected as a representative to illustrate the molecule-protein interactions. For the Imatinib ligand, the Protein RMSD values exhibited negligible variation, suggesting that the system was in equilibrium throughout the simulation. The results of dynamics for the processed compound (Imatinib) are represented in Table 3, Figure. 5. The OPLS3 force field RMSD index was 0.25 Å at its best ligand, confirming the precision of the measurements (The standard RMSD index of the OPLS3 force field is below 2 Å). The use of a short initial simulation series is required to minimize the impact of the system on the speed at which the system reaches a stable RMSD profile. The lower RMSD value indicates that the docking protocol could be reliable for the final docking studies of the selected compounds against the targets. The results of protein and ligand RMSD calculations at 10 nanoseconds are reported in diagram 1.

Table 3. Molecular dynamic simulation results for Imatinib ligand.

Items	Results
Total Energy	-61062.9883 kJ/mol
Stretch	708.3950 kJ/mol
Bend	2408.1514 kJ/mol
Torsion	1876.1947 kJ/mol
Improper Torsion	93.5313 kJ/mol
VDW	-2923.5383 kJ/mol
Explicit Hydrogen Bonds	0.00 kJ/mol
Cross Terms	0.00 kJ/mol
Solvation	-11867.2109 kJ/mol
T.E. for cross-checking	-61062.9883 kJ/mol
Iterations	1650 out of 2500
Conf 1 E	-61062.988 (0.123) kJ/mol

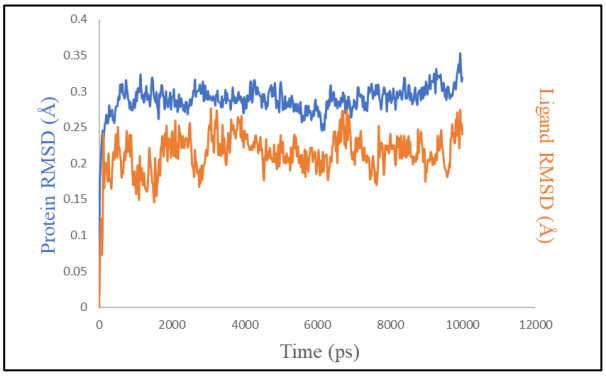

Diagram 1. Protein and ligand RMSD calculations.

The results of protein and ligand RMSD calculations are shown in Diagram 1. As shown in the figure, the best results were obtained for the RMSD of imatinib.

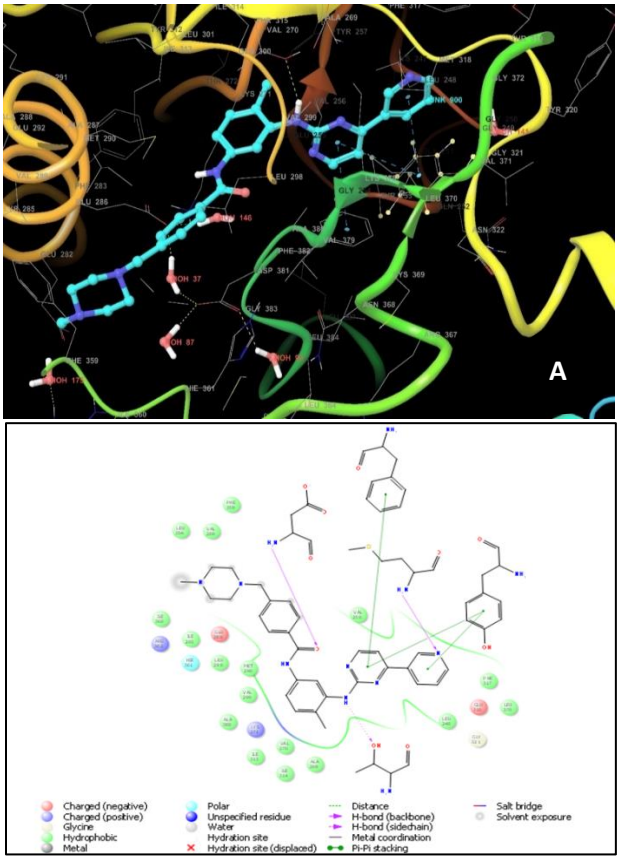

Figure 5. The 2D and 3D ligand-protein interactions of reference compounds with the active site of amino acid residues, ASP 381, THR 315, MET 318.

As can be seen in the image, part B, the purple lines represent hydrogen bonds in the complex structure, and the green lines represent the Pi-Pi stacking. In Figure. 4, part A, the three fixed hydrogen bonds were observed in amino acids ASP 381, THR 315, and MET 318 with a mean distance of 2.05, 2.38, and 2.12 Å. In the formation of hydrogen bonds at sufficient ligand-protein distances, water molecules have played a significant role. The diagram of potential energy in terms of temperature is shown in Figure. 6.

High binding energy leads to greater stability of the system. When molecules are more stable (i.e., have higher binding energy), their activity is usually reduced because they are less likely to react and change state to products. In reactions where the inhibitory activity is low, a high binding energy usually indicates that the reactants are more stable against conversion to products. Therefore, as the binding energy increases, the inhibitory activity (rate of reaction) decreases [35]. The shorter the bond distance, the stronger the hydrogen bond because the Schrödinger software itself eliminates distances of less than 5 angstroms in the calculations.

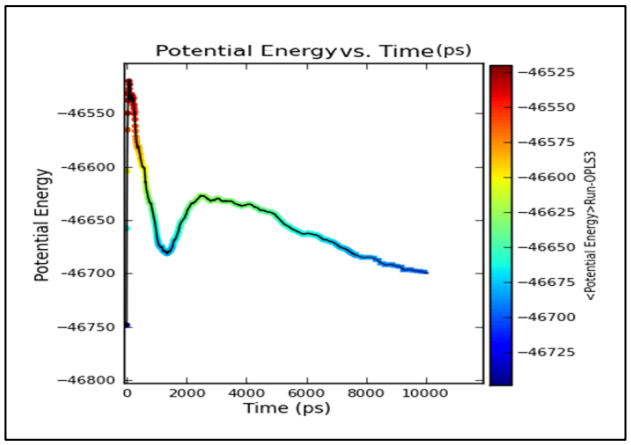

Figure 6. Achieve convergence in terms of energy in OPLS3 force fields at MD: potential energy vs. time.

Figure 6, illustrates that the mechanism has achieved a stable equilibrium. In the project, the achievement of convergence was defined in terms of energy. According to the standards, the convergence threshold was assessed as 0.05 kJ/mol. This means that if two or more calculated potential energy values differ by only 0.05 kJ/mol, the system is iteratively optimized. Moreover, if an energy value reaches -46700 kJ/mol in 8 to 10 nanoseconds, the system has reached energy stability. Furthermore, the low RMSD index (less than 1 Å) indicates that the compounds are very well placed in the active site and that the system has properly identified the active site. The standard deviation of the natural ligand is negligible. The analysis clearly shows that the amino acid ASP number 381 plays a vital role. The amino acid binds to the compound, leading the compound to the active site. Next, other hydrogen bonds cause the

persistence of the compound in the active site (Figure. 7).

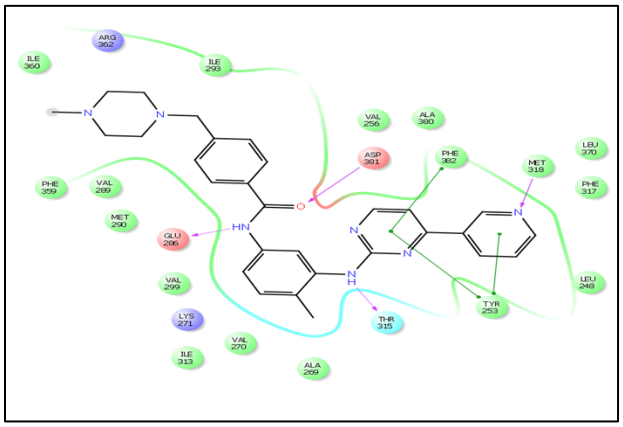

Figure 7. The 2D representation of the first contact of the ligand by the protein via the amino acid ASP 381.

The formation of hydrogen bonds between ligands and proteins, which is commonly used for molecular simulation and modeling, can lead to important results, including the following: Hydrogen bonds can be an important factor in the stability and strength of interactions between ligands and target proteins. In simulations, the study of these bonds helps to more precisely determine the key binding points between ligands and proteins and can stabilize or deform proteins, which helps to predict biological and molecular functions. It can also provide information about the optimal binding sites of ligands to proteins. In general, hydrogen bond simulations lead to a better understanding of molecular interactions and can help to design drugs better, optimize biological activities, and predict the behavior of complex biological systems. Table S2 includes MD simulation settings.

CONCLUSION

This comprehensive computational study evaluated the binding interactions and stability of 33 tyrosine kinase inhibitors (TKIs) against the ABL tyrosine kinase (PDB: 2HYY) using integrated molecular modeling approaches. The docking scores for selected compounds varied from -7.799 to -12.560 kcal/mol. The PSA, the number of hydrogen binding acceptors and donors, molecular weights, and partition coefficients were all within the allowable range for all selected compounds. In conclusion, based on ADME and the free energy binding values of the affected compounds, Imatinib was chosen as a potential inhibitor. It showed the highest liaison affinity with the protein.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Research Council of Damghan University for providing adequate facilities to do this research.

Authors contribution

Zohre Ghiasi, PhD candidate, Conceptualization, Methodology, Software, Validation, Data curation, writing the original draft

Davood Ajloo: Supervisor, Conceptualization, Editing original draft.

DISCLOSURE STATEMENT

The authors have no conflicts of interest to declare.

FUNDING

There is no funding for this research.

REFERENCES

[1] J.M. Kocarnik, K. Compton, F.E. Dean, et al., JAMA Oncol. (2022) 8(3):420–444

[2] R.L. Siegel, A.N. Giaquinto, A. Jemal, CA Cancer J. Clin. (2024) 74:12–49

[3] E. Cardoso, M. Guidi, B. Blanchet, Ther. Drug Monit. (2020) 42:33–44

[4] N. Verougstraete, V. Stove, A.G. Verstraete, Talanta (2021) 226:122140

[5] H.H. Huynh, C. Pressiat, H. Sauvageon, Ther. Drug Monit. (2017) 39:43–54

[6] J. Zhang, P.L. Yang, N.S. Gray, Nat. Rev. Cancer (2009) 9(1):28–39

[7] C. Neul, E. Schaeffeler, A. Sparreboom, et al., Trends Pharmacol. Sci. (2016) 37:904–932
[8] D.L. Longo, N. Engl. J. Med. (2017) 376:982–983
[9] R.B. Verheijen, H. Yu, J.H.M. Schellens, Clin. Pharmacol. Ther. (2017) 102:765–776
[10] S.Y. Pan, S.F. Zhou, S.H. Gao, Evid.-Based Complement Altern. Med. (2013) 2013:627375
[11] S.M. Paul, D.S. Mytelka, C.T. Dunwiddie, Nat. Rev. Drug Discov. (2010) 9:203–214
[12] M. Dickson, J.P. Gagnon, Nat. Rev. Drug Discov. (2004) 3:417–429
[13] M. Xiang, Y. Cao, W. Fan, et al., Comb. Chem. High Throughput Screen. (2012) 15:328–337
[14] M.H. Baig, K. Ahmad, G. Rabbani, et al., Comb. Chem. High Throughput Screen. (2018) 16:740–748
[15] B. Kamaraj, A.M. Al-Subaie, F. Ahmad, J. Biomol. Struct. Dyn. (2021) 39:5235–5247
[16] K. Banerjee, U. Gupta, S. Gupta, et al., Bioinformation (2011) 7(6):285–290
[17] G. Van Den Driessche, D. Fourches, J. Cheminform. (2017) 9:13
[18] F. De Falco, C. Di Giovanni, C. Cerchia, Biochem. Pharmacol. (2016) 104:83–94
[19] G.M. Sastry, M. Adzhigirey, T. Day, et al., J. Comput. Aided Mol. Des. (2013) 27(3):221–234
[20] E. Harder, W. Damm, J. Maple, et al., J. Chem. Theory Comput. (2016) 12(1):281–296
[21] M.J. Robertson, J. Tirado-Rives, W.L. Jorgensen, J. Chem. Theory Comput. (2015) 11(7):3499–3509
[22] H. Sirous, G. Campiani, V. Calderone, et al., Comput. Biol. Med. (2021) 137:104808
[23] M.P. Jacobson, D.L. Pincus, C.S. Rapp, et al., Proteins Struct. Funct. Bioinform. (2004) 55(2):351–367
[24] V. Gaddaguti, T.V. Rao, A.P. Rao, 3 Biotech (2016) 6(1):26
[25] B.C. Doak, B. Over, F. Giordanetto, et al., Chem. Biol. (2014) 21:1115–1142
[26] M. Ylilauri, O. Pentikäinen, J. Chem. Inf. Model. (2013) 53:2626–2633
[27] D. Cappel, M.L. Hall, E.B. Lenselink, et al., J. Chem. Inf. Model. (2016) 56(12):2388–2400
[28] A. Mili, S. Das, K. Nandakumar, J. Biomol. Struct. Dyn. (2022) 3:1–18
[29] T. Fawcett, Pattern Recognit. Lett. (2006) 27(8):861–874
[30] R.A. Friesner, J.L. Banks, R.B. Murphy, et al., J. Med. Chem. (2004) 47:1739–1749
[31] H. Zhong, L.M. Tran, J.L. Stang, J. Mol. Graph. Model. (2009) 28:336–346
[32] L. Duan, X. Guo, Y. Cong, et al., Front. Chem. (2019) 7:540
[33] M. Sripal, I. Celik, J. Khan, et al., J. King Saud Univ. Sci. (2022) 34(3):101826
[34] H. Meiselbach, A.H. Horn, T. Harrer, H. Sticht, J. Mol. Model. (2007) 13:297–304
[35] G.R. Baker, P.D. Carr, Biochemistry (2002) 41(20):6489–6500