



Unveiling promising phytochemicals from *Moringa oleifera* as dual inhibitors of EGFR^(T790M/C797S) and VEGFR-2 in non-small cell lung cancer through in silico screening, ADMET, dynamics simulation, and DFT analysis

Md. Masudur Rahman Munna^{a,b}, Md. Touki Tahamid Tusar^{a,1}, Saima Sajnin Shanta^{c,1}, Md. Hossain Ahmed^{a,1}, Md. Sarafat Ali^{a,*}

^a Department of Biotechnology and Genetic Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj 8100, Bangladesh

^b Dawn of Bioinformatics Limited, Dhaka 1361, Bangladesh

^c Department of Biochemistry and Molecular Biology, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj 8100, Bangladesh

ARTICLE INFO

Keywords:

EGFR
VEGFR-2
Molecular docking
Molecular dynamics simulation
Density functional theory

ABSTRACT

Non-small cell lung cancer (NSCLC) is among the main causes of mortality from cancer around the globe, affecting all genders. Current treatments mainly focus on tyrosine kinase inhibitors (TKIs) targeting the epidermal growth factor receptor (EGFR). However, resistance mechanisms, such as the emergence of T790M and C797S EGFR mutations and upregulation of VEGFR-2, often hinder the effectiveness of TKIs. Thereby, EGFR and VEGFR-2 present an intriguing opportunity for the treatment of NSCLC by developing dual-acting drugs. This research aims to evaluate prospective *Moringa oleifera* L. (MO)-originated compounds to efficiently block both of these receptors. In our research, we screened a library of 200 compounds sourced from MO, a plant known for its remarkable therapeutic potential. We identified five intriguing phytochemicals: hesperetin, gossypetin, quercetin, gallic acid, and epigallocatechin, as potential anti-cancer agents. The compounds have demonstrated notable binding affinity in virtual screening and multi-stage molecular docking analysis, surpassing the controls, Erlotinib and Bevacizumab + Rituximab. In addition, these compounds demonstrate top-notch drug-likeness and ADMET properties. The five promising drug candidates also had a strong ability to bind to receptors and stayed stable with them during the 200 ns molecular dynamics (MD) simulation and MM-GBSA calculation. Furthermore, DFT analysis indicates that hesperetin, gossypetin, and quercetin stand out as the most promising drug candidates among all others. The findings of our study suggest that these three therapeutic candidates can precisely target both EGFR and VEGFR-2 and can potentially act on both of these pathways as a single agent.

1. Introduction

Lung cancer is the most prevalent form of cancer and is responsible for the majority of cancer-related deaths in both males and females¹. It remains one of the most common and deadly cancer on a global scale and can be caused by disruptions in genetic and epigenetic processes, which can be influenced by variables such as hormone imbalances, microbial infections, smoking, exposure to second-hand smoke, and

chemical carcinogenesis^{2,3}. The incidence of lung cancer has increased over time in both industrialized and developing nations due to a variety of complex factors, such as population aging and growth, rapid socioeconomic progress, and shifts in the prevalence of related risk factors⁴. It is the most frequently diagnosed cancer and the leading cause of cancer-related deaths in 2022, with an estimated 2.4 million new cases and 1.8 million deaths in individuals of all genders and age groups⁵. Based on the Cancer Tomorrow online tool provided by GLO-

Abbreviations: EGFR, Epidermal Growth Factor Receptor; VEGFR-2, Vascular Endothelial Growth Factor Receptor- 2; MAPK, Mitogen-Activated Protein Kinase; PI3K, Phosphoinositide 3-Kinase; mTOR, Mammalian Target of Rapamycin; MM-GBSA, Molecular Mechanics With Generalised Born and Surface Area Solvation; VS, Virtual Screening; ADMET, Absorption, Distribution, Metabolism, Excretion, and Toxicity; SDF, Structured Data File; PDBQT, Protein Data Bank, Partial Charge (Q), & Atom Type (T); Kcal/mol, Kilocalorie Per Mole; SMILES, Simplified Molecular Input Line Entry System; CYP, Cytochromes P450; CID, Compound Identification Number.

* Corresponding author.

E-mail address: sarafatbiotech@bsmrstu.edu.bd (Md. Sarafat Ali).

¹ Authors are equally contributed.

<https://doi.org/10.1016/j.jgeb.2024.100406>

Received 29 May 2024; Received in revised form 27 July 2024; Accepted 29 July 2024

BOCAN, it is predicted that there will be a 64 % increase in lung cancer incidence and a 67 % increase in death cases by 2040⁶. Small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC) are the two major types of lung cancer, accounting for around 15 % and 85 % of lung malignancies in all cases⁷. NSCLC is a multifactorial disease that have multiple causes and progress through different phases. Hence, the multifactorial nature of this cancer emphasizes the necessity for a multifunctional therapeutic approach, such as the identification of a single molecule capable of regulating many disease-causing pathways.

Epidermal Growth Factor (EGF), a proangiogenic factor, can be observed in epithelial cells and binds with the EGFR for cell proliferation, differentiation and to maintain cellular development. But when mutated, EGFR plays a crucial role in NSCLC by promoting the survival of tumor cells by activating the overexpression of some anti-apoptotic pathways, like the MAPK and/or PI3K/Akt/mTOR pathways⁸. The secondary resistance mutation T790M and the tertiary resistance mutation C797S are the most prevalent EGFR mutations, occurring in nearly 60 % of NSCLC cases where individuals develop acquired resistance when taking medications⁹. EGFR also modulates the expression of VEGF-2 via transcription factor hypoxia-inducible factor HIF-1 α , as illustrated in Fig. 1. VEGFR-2 is a receptor for Vascular Endothelial Growth Factor (VEGF), a signaling molecule crucial for blood vessel development (angiogenesis) and also stimulates microvascular hyper permeability¹⁰. Both of these induce the tumor stroma formation and tumor cell spreading. There are three types of VEGFR, but VEGFR-2 is indispensable for tumor angiogenesis and associated with NSCLC progression to the advanced stage¹¹. Under hypoxic conditions, cells can directly produce VEGF or increase the synthesis of interleukin-8 (IL-8). IL-8 then binds to CXCR2 (C-X-C Motif Chemokine Receptor 2), which triggers a signaling cascade within the cell. This pathway is involved in activating the transcription factor NF- κ B (Nuclear Factor- κ B). NF- κ B can then induce the expression of VEGF¹². VEGF has five different isoforms (A, B, C, D, and E), from which VEGF-A, VEGF-B, and VEGF-C can only bind with VEGFR-2 and acti-

vate it¹³. In the downstream, VEGFR-2 also regulates the same anti-apoptotic pathways as EGFR¹⁴. In addition, VEGFR-2 may also independently can be activated by bFGFs (basic Fibroblast Growth Factors) and PDGFs (Platelet-Derived Growth Factors)¹⁵. Therefore, merely inhibiting either EGFR or VEGFR-2 may still result in the activation of PI3K/AKT and/or MAPK pathways¹⁶, hence we need a single compound that potentially limiting the activity of EGFR and VEGFR-2 and inhibiting tumor growth. The advent of computational power has revolutionized modern anti-cancer drug development, enabling the discovery of more affordable, less toxic, and potent anti-cancer agents. By targeting specific molecular targets, such as receptors, proteins, and enzymes, researchers can accelerate the drug discovery process^{17,18}. Numerous studies have leveraged computational approaches to identify potential anti-cancer agents from natural sources, as exemplified by Daoui et al.'s (2023) exploration of Cannabis sativa L. compounds as novel EGFR-TKIs¹⁹. In another study, Daoui et al. identified novel 4-phenoxy pyridine derivatives as potential c-Met inhibitors for cancer therapy¹⁸. Although these computational approaches have shown promise in drug discovery, to the best of our knowledge, no studies have investigated the simultaneous dual inhibitory potential of *Moringa oleifera* L. compounds against EGFR and VEGFR-2 in NSCLC. This study aims to fill this knowledge gap by computationally investigating MO compounds as potential dual inhibitors of these targets. The extracts of MO exhibited significant efficacy against a variety of cancerous cells as it contains a diverse range of bioactive chemicals with anticancer properties^{20,21}. Research has shown that its extracts can hinder cell proliferation, growth, and migration by influencing various signaling pathways in cells, which can be used to develop novel drugs for the treatment of different forms of cancer^{22,23}.

Drug development is a time-consuming process requiring rigorous testing, emphasizing the value of computer-aided drug design for accelerating discovery²⁴. Computational techniques have been applied in various therapeutic areas beyond anti-cancer drug development. Previous studies have explored their application in diabetes, as

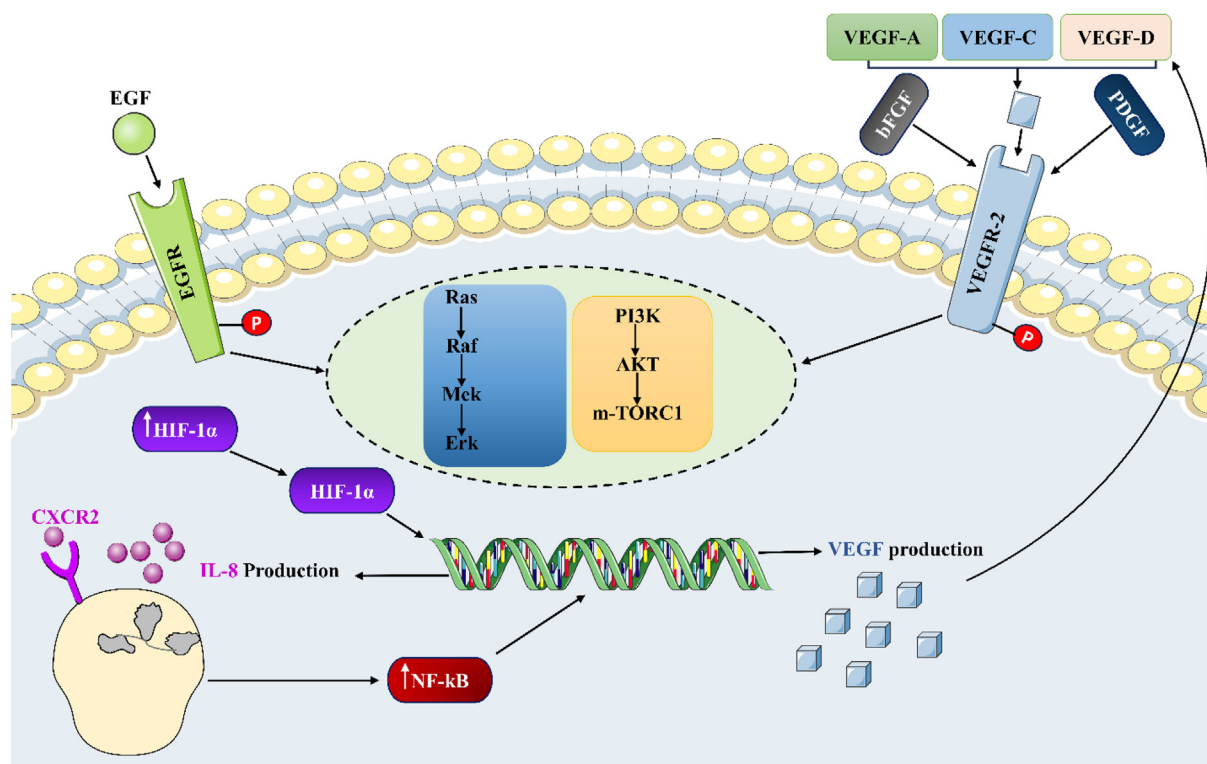


Fig. 1. The interaction between the EGFR and VEGFR-2 signaling pathways.

exemplified by the work of Abchir et al., neurodegenerative diseases, as demonstrated by Nour et al., and other therapeutic domains^{24,25}. However, though quantitative structure–activity relationship models are pivotal in this landscape, correlating molecular properties with biological activity²⁶, structure-based methods, such as molecular docking and dynamics, offer complementary insights into protein–ligand interactions²⁷. Our study aims to contribute to this field by employing a robust computational pipeline that integrates molecular docking, molecular dynamics simulations, binding free energies, and quantum mechanics calculations. In our study, we firstly identified potential phytochemical compounds from MO via virtual screening with EGFR and VEGFR-2 proteins. Then, we identified compounds that had a higher binding energy than our chosen controls and omitted the ones with unfavorable pharmacokinetic and toxicity properties by ADMET analysis. In addition, we performed molecular dynamics simulation and MM-GBSA analysis of the selected protein–ligand complexes to gain insights into the protein's stability and detect any structural changes that might occur upon the binding of compounds. Finally, we analyzed the electrical and structural features of the selected compounds through Density functional theory. Research reveals that double inhibition of the EGFR and VEGFR-2 pathways yields a more advantageous therapeutic effect than solely inhibiting these pathways. In this study, we identified the compounds hesperetin, gossypetin, and quercetagenin from MO using a rigorous computational approach. These compounds may improve the treatment approach for non-small cell lung cancer by selectively inhibiting both the EGFR^(T790M/C797S) and VEGFR-2 pathways, potentially leading to a reduction in lung cancer mortality rates.

2. Methods and materials

2.1. Preparation of proteins

The crystal 3D structures of the epidermal growth factor receptor (EGFR) and vascular endothelial growth factor receptor 2 (VEGFR-2) of humans were retrieved from the RCSB protein data bank ([rcsb.org/](https://www.rcsb.org/))²⁸, which served as the targeted receptor proteins. We selected and downloaded the crystal structure of EGFR (PDB entry code: 5ZJWJ) due to its unique mutation profile and superior resolution com-

pared to alternative structures available for this specific receptor. The structure has a resolution of 2.90 Å (R-value free: 0.272, R-value observed: 0.256, and R-value work: 0.255). Likewise, the crystallographic structure of VEGFR-2 (PDB ID-4ASD and resolution of 2.03 Å, R-Value Free: 0.230, R-Value Observed: 0.178, and R-Value Work: 0.175) was also obtained. In addition, we also analyzed the Ramachandran plot of these proteins and found that over 90 % of the residues ([supplementary file Fig.S2](#)) are in the most favored region of both proteins. The downloaded PDB structures of the proteins were processed using specific criteria, which involved removing any attached ligands and eliminating water molecules using Discovery Studio 2024²⁹. The protein structure was effectively stabilized through subsequent energy minimization, using the GROMOS96 force field of SWISS PDB Viewer (SPDV) v4.10³⁰.

2.2. Compound library preparation

Phytochemicals, a class of bioactive compounds extracted from medicinal plants, have garnered considerable attention for their pharmacological properties and potential therapeutic applications as pharmaceutical agents. Hence, the research aimed to explore the phytochemical substances that originate from *Moringa oleifera* L. (MO) and their potential inhibitory effects on targeted receptors, relying on the IMPPAT 2.0 (Indian Medicinal Plants, Phytochemistry, and Therapeutics 2.0) (imppat/) database³¹. The IMPPAT 2.0 database is a vast compilation of information, encompassing 4010 Indian medicinal plants, 17,967 phytochemicals, and 1095 therapeutic uses. This valuable resource may assist in the search for novel biologically significant molecules to support drug discovery efforts. A library of phytochemicals extracted from various plant parts of MO was compiled, followed by a search for these compounds in the NCBI PubChem ([pubchem/](https://pubchem.ncbi.nlm.nih.gov/)) database³². The PubChem database is a large repository of information regarding chemical compounds, and it offers information on their structure, formula, molecular weight, and other relevant details. The percentage of compounds identified from different parts of the MO is displayed in [Fig. 2](#). Compounds that displayed any of the four chemical safety issues, specifically irritancy, flammability, health hazards, and environmental dangers, were excluded from the study. The remaining compounds, which have no safety concerns and a molecular

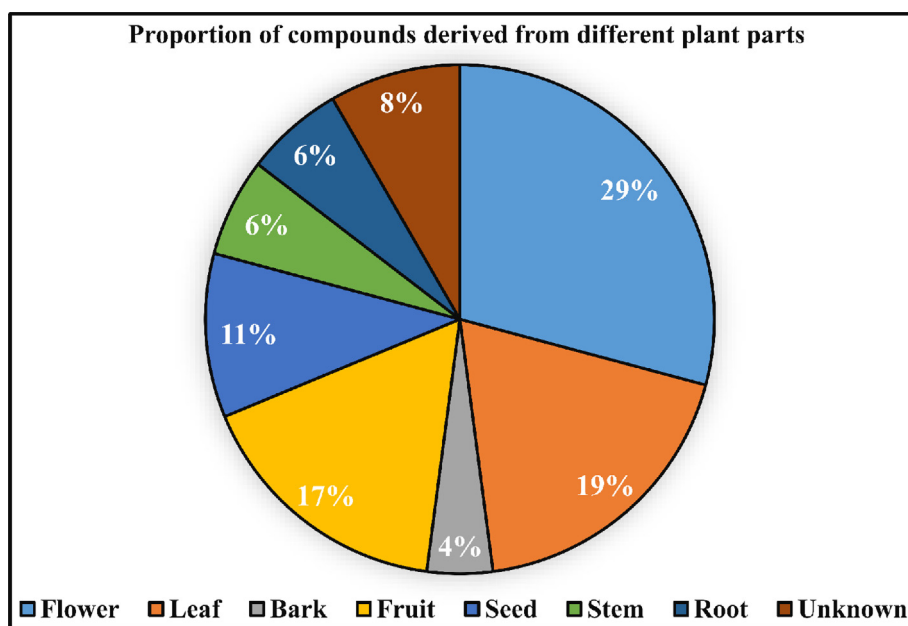


Fig. 2. Percentage of compounds retrieved from different parts of *Moringa oleifera* L.

weight below 500 g/mol, were included in the library. To prepare the compounds for virtual screening, their 3D structures were retrieved from the PubChem database and converted into a unified PDBQT file using OpenBabelGUI³³. Subsequently, Open Babel was employed to finalize the preparation of the compounds by minimizing them and adding polar hydrogens.

2.3. Virtual screening

Virtual screening (VS) is now a commonly used technique in drug development for finding potential hits by screening enormous collections of compounds against a specific protein target. VS utilizes efficient algorithms to efficiently screen compounds and effectively identify prospective therapeutic candidates³⁴. We employed the PyRx virtual screening tool, utilizing its AutoDock Vina wizard to identify optimal binding poses for the phytocompounds within the targeted receptor proteins. In our study, the phytocompounds were subjected to VS analysis against two receptor proteins by blind docking. Blind molecular docking simplifies the identification of optimal complexes based on energy, regardless of the binding site's location. This method is valuable for investigating protein–ligand interactions and predicting inhibitors for therapeutic targets³⁵. During VS, we also include Erlotinib (Pubchem CID: 176870) as a control for EGFR and Bevacizumab + Rituximab (Pubchem CID: 24801581) for VEGFR-2. Erlotinib is an FDA-approved EGFR inhibitor used to treat NSCLC with EGFR mutations³⁶. On the other hand, Bevacizumab + Rituximab are drug combinations used for the treatment of certain kinds of cancer, particularly those linked to dysfunctions in VEGFR-2³⁷. The grid box parameter for EGFR was set to $x = 5.3371$, $y = 29.8879$, and $z = 3.9183$, and the dimensions were set to $x = 41.077$ Å, $y = 58.774$ Å, and $z = 62.285$ Å to accurately represent the bioactive conformation of the protein and ligand. The VEGFR-2 grid box was defined with a parameter of $x = -21.9259$, $y = -1.1298$, and $z = -3.7084$, while the dimensions were set to $x = 52.462$ Å, $y = 51.565$ Å, and $z = 57.045$ Å. The exhaustiveness was set to 32, and the VS was run using the default setup parameters of the PyRx virtual screening tool, and the drug candidates that had a higher negative binding energy than the controls with an RMSD value of 0 were selected. As the main goal of this study is to identify dual inhibitors of EGFR and VEGFR-2, we then identified the compounds that have a better binding affinity with both receptors and performed drug-likeness, ADME, and toxicity analyses of these compounds. Following ADMET analysis, the compounds that showed the highest drug-likeness and excellent ADME properties with no toxicity had their docking scores revalidated using the single molecular docking method of the AutoDock Vina³⁸ and BINDSURF³⁹.

2.4. Prediction of drug-like properties

In assessing new drug candidates, it is essential to comprehend their drug-like properties. These properties are essential for predicting how effectively a drug will behave in the body. In this study, following the VS process, an evaluation of the physicochemical features, lipophilicity, and drug-likeness properties of the selected compounds was conducted utilizing the SwissADME ([swissadme/](http://www.swissadme.ch)) web server⁴⁰. Identifying compounds suitable for drug use involves adhering to specific rules such as Lipinski's⁴¹, Veber's⁴², and Egan's⁴³. Compounds that demonstrate multiple violations of the Lipinski, Veber, and Egan rules may potentially encounter pitfalls related to their pharmacokinetic properties pertaining to ADMET. Only a small percentage of drugs (<10 %) that advance to the clinical trial phase fail to meet these criteria set by Lipinski, Veber, and Egan⁴⁴. In addition, we also evaluated two more factors including the topological polar surface area (TPSA) and number of rotatable bonds (n-ROTB). These criteria can be used to determine if the drug molecule exhibits a flexible or inflexible interaction with the receptor⁴⁵. Although certain com-

pounds may not follow these guidelines and are currently being used in clinical settings, these provides a strong basis for any in-silico screening process.

2.5. Prediction of ADME properties

Poor pharmacokinetic and safety features of potential therapeutic agents make drug development harder, leading to higher failure rates in advancing to clinical trials. Computational methods possess the capability to tackle and mitigate these problems. The report and research have emphasized the significant benefits of utilizing ADME to predict the pharmacokinetics of drug molecules before clinical or pre-clinical trials⁴⁶. The web-based server pkCSM ([pkcsm/prediction](http://pkcsm.biopparadis-lab.fr))⁴⁷ is a convenient and freely available resource for predicting the pharmacokinetic properties of drugs, including their absorption, distribution, metabolism, and excretion (ADME) characteristics. We used this machine learning method to analyze the ADME parameters of our screened compounds. We evaluated several ADME properties of the compounds, including human intestinal absorption, water solubility (Log S), blood–brain barrier (BBB) permeability, VDss (human), CYP450 substrates, CYP450 inhibitors, and total clearance (ml/min/kg).

2.6. Prediction of toxicological properties

Considering the potential risks associated with medication, it was important to assess the toxicological properties of the screened natural compounds. The ProTox-3.0 server ([protox3/](http://protox3.0.fraunhofer-efp.de))⁴⁸ and pkCSM online server⁴⁷, well-known for their effectiveness in drug discovery, were also utilized to forecast the toxicological analysis of selected compounds. Parameters such as AMES Toxicity, maximum tolerated dose, human Ether-a-go-go-Related gene, oral rat acute toxicity, skin sensitization, hepatotoxicity, neurotoxicity, immunotoxicity and cytotoxicity were utilized to filter out compounds. The compounds obtained from PubChem were used as the input to the web servers, utilizing their canonical smiles. Oral medicines must exhibit drug-like qualities to be considered pharmacologically consistent with their bioactivity score. We utilized the Molinspiration Cheminformatics software ([cgi/properties](http://www.molinspiration.com/cgi/properties))⁴⁹ to analyze the bioactivity score of our screened compounds encompassing enzyme inhibitors, kinase inhibitors nuclear receptor ligands, protease inhibitors, G protein-coupled receptor (GPCR) ligands, and ion channel modulators.

2.7. Molecular docking analysis

Molecular docking is a reliable approach in structural biology, particularly in computer-aided drug design (CADD) processes. This technique allows for predicting the optimal attachment pose of small molecules to target macromolecules⁵⁰. Based on our drug-likeness, ADME, and toxicity studies, we selected the five compounds gallic acid (CID: 65084), epigallocatechin (CID: 72277), hesperetin (CID: 72281), gossypetin (CID: 5280647), and quercetagenin (CID: 5281680) for molecular docking analysis as they have a great potential to be successful drugs. Erlotinib and Bevacizumab + Rituximab were also used as controls during molecular docking analyses for EGFR and VEGFR-2, respectively. Initially, the co-crystallized ligands EAI045 and sorafenib were redocked to the EGFR and VEGFR-2 receptors, respectively, to verify the docking approach. If the empirical binding pattern of the co-crystallized ligand effectively replicates in the redocking stage with a low RMSD value, then it suggests the modified docking procedure is appropriate for the present docking investigation. We used AutoDock Vina³⁸ to perform molecular docking between selected ligands and the EGFR and VEGFR-2 receptor proteins. Before the molecular docking process we cleaned the receptors using Biovia Discovery studio 2024²⁹ by removing water molecules and other attached ligand molecules and energy was minimized using

SWISS PDB Viewer (SPDV) 4.10³⁰. The receptors were then prepared using AutoDock v4.2.6⁵¹, which involved adding polar hydrogens to the protein atoms and then assigning Kollman charges. The prepared and cleaned crystal structure of the receptors was then converted into PDBQT format, which is necessary for docking with AutoDock Vina. Similar to the receptor preparation, the ligand structures were prepared for docking using AutoDock v4.2.6. This involved adding polar hydrogens and assigning Kollman charges. Then we set the center and dimensions of the grid box for the two receptors as they were during virtual screening; thus, this docking also served as a validation of virtual screening. The exhaustiveness was set to 32, and three distinct molecular docking simulations for each complex were then performed. Nine binding poses with binding affinity score in kcal/mol were identified following the docking process. The docking pose with a zero RMSD value and the lowest binding energy was deemed the most precise prediction. We also utilized BINDSURF (bindsurf/)³⁹ for docking analysis. BINDSURF provided an additional layer of validation of the docking scores for the docked complexes derived from Vina. Discovery Studio Visualizer 2024 was used to analyze the molecular interactions between the selected compounds and proteins.

2.8. Molecular dynamics simulation

The use of molecular dynamics (MD) simulation in the contemporary drug development process has expanded. MD simulation analyzes the movements of atoms and molecules, revealing details about the structural integrity and conformational changes within protein–ligand complexes. The objective of MD simulation is to ascertain information regarding the kinetic, thermodynamic, and structural characteristics of a protein–ligand complex. Since docking offers a static viewpoint on a compound's binding to a protein's active site, molecular dynamics simulation is used to track the atomic movement across time using Newton's classical equation of motion, allowing for a dynamic perspective⁵². The trajectory file will provide results with details on the kinetic and thermodynamic properties of the system obtained from the simulation study conducted within a specific time frame (ns). The top five pharmacological candidates, gallo catechin (CID: 65084), epigallocatechin (CID: 72277), hesperetin (CID: 72281), gossypetin (CID: 5280647), and quercetagenin (CID: 5281680), that satisfied ADMET criteria and had a higher docking score were selected for MD simulation to assess their stability when interacting with EGFR and VEGFR-2 receptors in the body's physiological environment. We used Desmond software version 7.4⁵³ from Schrödinger for performing MD simulations for a time-frame of 200 ns in a Linux environment. To compare the stability of the receptors with the selected compounds and controls, we also include both apo-proteins (only proteins, without any ligands) and the control's complex during simulation. We then individually imported ten complexes of the top five selected ligands and two controls with two specific receptors generated by AutoDockVina and the apo-proteins into Schrödinger's Maestro interface v13.6⁵⁴ and run the MD simulation. Before running the MD, the systems were relaxed using the Desmond default protocol of relaxation. The Maestro Protein Preparation Wizard⁵⁵ from Schrödinger Suite 2023–2 was used to preprocess the protein–ligand complex, involving advanced optimization and minimization steps. All necessary steps were accomplished, including filling in missing amino acid side chains, assigning bond orders, adding hydrogens, loops with hydrogen-bond assignment optimization, and sampling water orientations at pH 7.0. The OPLS_2005 force field⁵⁶ was used to subject the protein–ligand complex to the system builder. The protein–ligand complex was placed in an orthorhombic box by maintaining a distance of 10 Å between the protein and the box edge. The box was filled with water molecules using the TIP3P model to solvate the system for optimization. Both Na⁺ and Cl[−] ions were added with a salt concentration of 0.15 M to neutralize the entire system and simulate a physiological buffer. The MD simulation was performed in the NPT ensemble⁵⁷ with a Nose-Hoover ther-

mostat at 300 K and Martyna-Tobias-Klein barostats at 1.01 bar pressure. The simulation maintained constant volume using the Smooth Particle Mesh Ewald (SPME) method. The Simulation Interaction Diagram (SID) of the Schrödinger package was used to analyze MD simulation outcomes. The trajectory data were recorded at intervals of 200.0 ps, with 1000 frames and the stability of the protein–ligand complexes was assessed by calculating root mean square deviation (RMSD), root mean square fluctuation (RMSF), and other ligand properties like ligand RMSD, radius of gyration (Rg), MolSa, and PSA.

2.9. Binding free energy (MM-GBSA) and DFT calculation

The estimation of binding free energies offers an important predictive criterion for ranking compounds based on their binding affinities⁵⁸. The binding free energies of each of the ligands with the EGFR and VEGFR-2 were computed using the Molecular Mechanics Generalized Born Surface Area (MM-GBSA) approach⁵⁹. The MM-GBSA method looked at the stability of ligand–protein complexes following the interaction between the proteins and ligands to minimize the occurrence of inaccurate positive outcomes in molecular docking analysis. The MM-GBSA computes the binding complex's free energies by using solvation models and molecular mechanics computations⁶⁰. We used thermal_mmgbsa.py script of the Prime module⁶¹ in the Schrödinger suite 2023–2 to process the trajectory file by partitioning it into several snapshots. Then performs calculations on each 20-nanosecond trajectory to determine the binding energy across the entire 200-nanosecond time span. This module calculates the binding free energy by using the following specific formula:

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

Where, the term “ ΔG_{bind} ” refers to the total binding free energy, the symbol for the complex's free energy is G_{complex} , the symbol G_{protein} represents the target protein's free energy, and the ligand's free energy is denoted as G_{ligand} . The calculations of binding free energy (expressed as kcal/mol) between the ligands and proteins were allowed by the calculation of several energy modules, including van der Waals (vdW), covalent, Generalized Born electrostatic solvation, coulomb, hydrogen bonding (Hbond), and lipophilic binding energies (Lipo). Furthermore, we used Density Functional Theory (DFT) to analyze a quantum mechanical calculation, evaluating the electrical and structural features of the five most promising compounds. It is an important method for studying biomolecular systems in order to determine a reaction's transition state and reaction coordinates^{62,63}. Additionally, it serves as a critical and useful tool for analyzing the correlation between the geometric and electrical characteristics of bioactive compounds. The calculations were performed using the B3LYP/6-31G⁶⁴ (d, p) basic set of the Gaussian 09 W software⁶⁵ and calculated the energies of the frontier molecular orbitals (HOMO, LUMO, and energy gap) and additional reactivity parameters based on HOMO and LUMO energy, including ionizing potential, electron affinity, electronegativity, chemical hardness, softness, and electrophilicity index.

3. Results and discussion

3.1. Virtual screening

Structure-based drug design heavily relies on virtual screening techniques to identify potential lead compounds for target proteins from a vast array of natural compounds⁶⁶. Therefore, by using the crystal structure of human EGFR (PDB entry code: 5ZJW) and VEGFR-2 (PDB entry code: 4ASD), a virtual screening-based docking study was conducted. For the EGFR receptor, all phytochemicals exhibited binding energy scores ranging from −9.4 to −4.8 kcal/mol, while control Erlotinib showed −7.5 kcal/mol. The compounds, on

the other hand, bind to the VEGFR-2 receptor with an affinity range of -9.5 to -4.3 kcal/mol, while the control compound, Bevacizumab + Rituximab, had an affinity of -6.9 kcal/mol. To analyze drug likeness, ADMET, and toxicological properties, we selected the 17 compounds that bound to EGFR and VEGFR-2 proteins with an affinity greater than control and have better interactions with both receptors. Fig. 3 depicted the binding affinity score of all of these compounds. The higher binding affinity of selected compounds to the tar-

get receptors than controls suggests these can have better efficacy in treating the disease.

3.2. Analysis of drug-like properties

The evaluation results of the drug-like properties of the selected compounds are depicted in Table 1. According to Lipinski's rule of thumb, compounds that fail to meet multiple guidelines have inad-

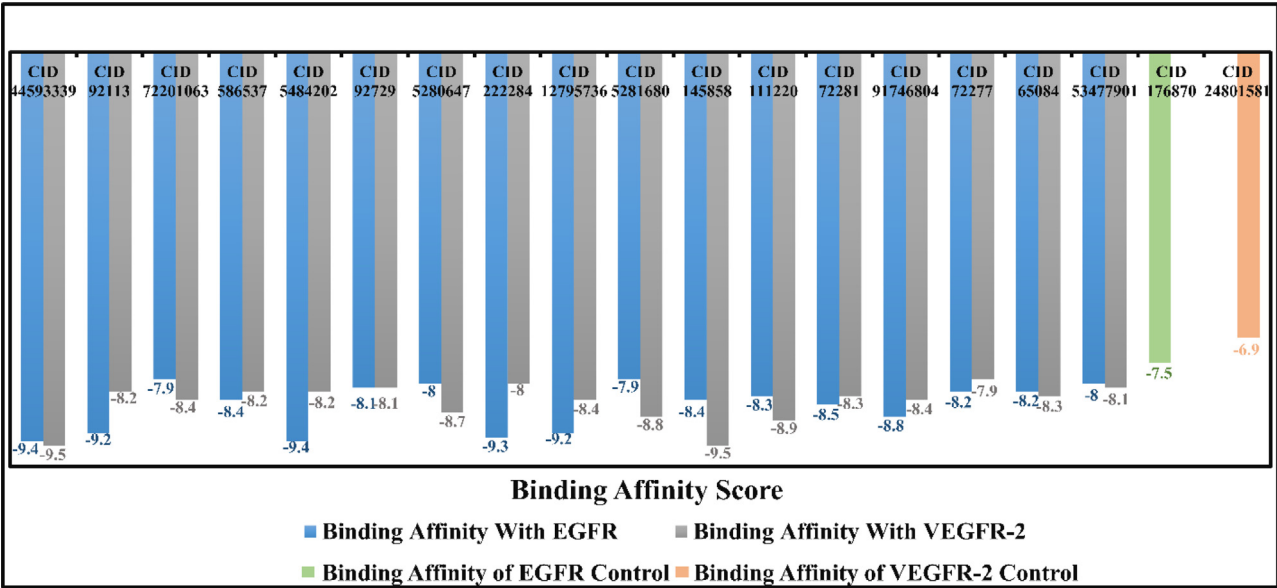


Fig. 3. Binding affinity score of all the compounds with a higher binding score than controls with both EGFR and VEGFR-2 proteins after virtual screening in the PyRx tool.

Table 1
Assessment of the physiochemical properties, lipophilicity and drug-like characteristics of the selected compounds according to SwissADME.

Compound CID	Phytochemical Name	Properties										
		Physiochemical Properties						Lipophilicity		Drug-likeness		
		Rules										
		< 500	< 5	< 10	≥ 40 – ≤ 130	≤ 140	< 10	≤ 5		≤ 1	≤ 1	≤ 1
MW (Da)	HBD	HBA	MR	TPSA (Å ²)	NRB	iLOGP		Lipinski's violation	Veber Violation	Egan Violation	BS	
72281	Hesperetin	302.28	3	6	78.06	96.22	2	2.24	0	0	0	0.55
5280647	Gossypetin	318.24	6	8	80.06	151.59	1	1.33	1	1	1	0.55
12795736	delta7-Avenasterol	412.69	1	1	132.75	20.23	5	5.07	1	0	1	0.55
5281680	Quercetagenin	318.24	6	8	80.06	151.59	1	1.73	1	1	1	0.55
92113	24-Methylenecholesterol	398.66	1	1	127.95	20.23	5	4.78	1	0	1	0.55
145858	Flavylum	207.25	0	1	66.06	13.14	1	−0.76	0	0	0	0.55
72277	Epigallocatechin	306.27	6	7	76.36	130.61	1	0.98	1	0	0	0.55
111220	Bauerenol	426.72	1	1	135.14	20.23	0	4.71	1	0	1	0.55
222284	Beta-Sitosterol	414.71	1	1	133.23	20.23	6	4.79	1	0	1	0.55
72201063	Pterygospermin	406.52	0	2	123.32	89.12	4	3.45	0	0	0	0.55
586537	6-Chromanol, 2,8-dimethyl-2-(4,8,12-trimethyltridecyl)	402.65	1	2	129.34	29.46	12	5.33	1	1	1	0.55
65084	Galocatechin	306.27	6	7	76.36	130.61	1	0.98	1	0	0	0.55
92729	gamma-Tocopherol	416.68	1	2	134.31	29.46	12	5.76	1	1	1	0.55
91746804	28-Isoavenasterol acetate	468.75	0	2	147.04	26.30	7	5.73	1	0	1	0.55
44593339	Roridin E	514.61	1	8	1	103.82	135.56	3.62	1	0	0	0.55
5484202	Stigmast-4-en-3-one	412.69	0	1	6	17.07	132.27	4.69	1	0	1	0.55
53477901	(3S,9S,10R,13R,14S,17R)-10,13-dimethyl-17-[(2S)-6-methyl-5-methylideneheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-3-ol	398.66	1	1	127.95	20.23	5	4.78	1	0	1	0.55

quate oral bioavailability⁶⁷. In our analysis, three out of 17 candidate drugs meet all Lipinski rules, which include criteria like $MW \leq 500$, $HBA \leq 10$, $HBD \leq 5$, $\text{Log } P \text{ (octanol/water)} \leq 5$, and $40 \leq MR \leq 130$. The remaining 14 compounds adhere to the Lipinski rule with only 1 violation, which is also acceptable. Furthermore, these compounds were reviewed based on the Veber and Egan rules. Out of all the compounds, only four had one acceptable violation of the Veber rule, while the rest had no violations. When it comes to the Egan rule, six compounds satisfied all the criteria, while 11 compounds fell short of meeting one rule. All the compounds show excellent bioavailability scores, with each one reaching around 55 % bioavailability.

3.3. ADME properties analysis

Drug research relies heavily on pharmacokinetic properties (PKs), which help understand and predict biological consequences, like the favorable or negative influence of a chemical on a particular system. The pharmacokinetic properties evaluate whether the drug compound has been efficiently absorbed from the gastrointestinal system, accurately transported to its target site, metabolized adequately, and removed from the body without causing any negative consequences. Properties related to absorption, distribution, metabolism, and excretion (ADME) play a crucial role in determining the pharmacokinetics of a molecule, which are pivotal steps for the success of a drug. The results of our ADME properties analysis are shown in Table 2. Exploring the ADME properties of the mentioned drugs, including water solubility, intestinal absorption in humans (HIA), VDss in humans, blood–brain barrier permeability, CYP2D6/CYP3A4 substrates, and CYP1A2/CYP2C19/CYP2C9/CYP2D6/ CYP3A4 inhibitors, as well as drug total clearance. Inspection of these characteristics is essential to guaranteeing a drug's efficacy by assessing its biological activity and precise targeting of the intended organ at a specific concentration. Based on the research, each compound has a unique water solubility range. The identified compounds exhibit a solubility range of -2.90 to -7.812 , suggesting that most of them are moderately to highly soluble in physiological environments. Substances with values from -10

to -6 exhibit poor solubility; those greater than -6 and less than -4 indicate moderate solubility; compounds between -4 and -2 are considered soluble; and compounds with values between -2 and 0 are highly soluble⁶⁸. Gallocatechin (CID: 65084), Epigallocatechin (CID: 72277), Hesperetin (CID: 72281), Gossypetin (CID: 5280647), and Quercetageitin (CID: 5281680) have been identified as soluble compounds with a range of -2.9 to -3.047 . Low absorption is indicated when the percentage of absorption by the human intestines is less than 30 %. All the compounds showed a range of human intestinal absorption (HIA) from 54.128 % to 98.556 %, indicating moderately high to very high absorption rates. The compound's distribution was evaluated by determining their volume of distribution in humans (VDss) and their ability to cross the blood–brain barrier (BBB). A log VDss value less than -0.15 suggests a lower distribution, while a value greater than 0.45 indicates a higher distribution of a drug. With a low volume of dispersion Stigmast-4-en-3-one is classified as undrug-gable. The remaining substances showed a moderate to high volume of distribution. The blood–brain barrier protects the central nervous system (CNS) by preventing specific substances, mainly harmful ones, from accessing the brain tissues. The BBB regulates the entry of external compounds to maintain the stability of the CNS⁶⁹. Based on the pkCSM guidelines, 35 % (6/17) of the compounds are deemed potential drug candidates due to their low likelihood of penetrating the BBB, while the rest of the compounds are capable of crossing the BBB. Detoxifying foreign compounds and metabolizing drugs involves heme-containing enzymes called cytochrome P450s (CYPs). Drugs can either act as inhibitors or inducers of cytochrome P450 enzymes, causing potential drug interactions that could lead to negative effects or reduced effectiveness of treatments⁷⁰. Evaluation of compounds for their effects on cytochrome P450 (CYP) is critical since CYP inhibitors may impact drug metabolism. The CYP types (1A2, 2C9, 2C19, 2D6, and 3A4) are essential in the biotransformation of more than 90 % of medications in the initial metabolism phase. Two different forms, 2D6 and 3A4, have profound effects on drug metabolism⁴⁴. In metabolism, only CID:72201063 was identified as a CYP2D6 substrate, while none of the compounds acted as CYP2D6 inhibitors. Among

Table 2
The results of in-silico ADME profiling of the potential hit molecules.

Compound CID	Phytochemical Name	Absorption		Distribution		Metabolism							Excretion
		Water solubility	Intestinal absorption (human)	VDss (human)	BBB permeability	Substrate CYP		Inhibitor					Total Clearance
		Numeric (Log mol/L)	Numeric (%) Absorbed)	(Log L/kg)	Numeric (Log BB)	2D6	3A4	1A2	2C19	2C9	2D6	3A4	Numeric (Log ml/min/kg)
72281	Hesperetin	− 3.047	70.277	0.746	− 0.719	No	No	No	No	No	No	No	0.044
5280647	Gossypetin	− 2.9	68.009	1.552	− 1.472	No	No	Yes	No	No	No	No	0.304
12795736	delta7-Avenasterol	− 6.715	94.642	0.179	0.764	No	Yes	No	No	No	No	No	0.613
5281680	Quercetageitin	− 2.904	62.773	1.424	− 1.664	No	No	Yes	No	No	No	No	0.307
92113	24-Methylencholesterol	− 7.041	94.691	0.426	0.777	No	Yes	No	No	No	No	No	0.604
145858	Flavylium	− 4.852	96.182	0.24	0.454	No	Yes	Yes	Yes	No	No	No	0.716
72277	Epigallocatechin	− 2.969	54.128	1.301	− 1.377	No	No	No	No	No	No	No	0.328
111220	Bauerenol	− 6.36	94.567	0.231	0.668	No	Yes	No	No	No	No	No	0.116
222284	Beta-Sitosterol	− 6.773	94.464	0.193	0.781	No	Yes	No	No	No	No	No	0.628
72201063	Pterygospermin	− 4.779	90.428	0.555	0.581	Yes	Yes	No	Yes	No	No	Yes	− 0.196
586537	6-Chromanol, 2,8-dimethyl-2-(4,8,12-trimethyltridecyl)	− 7.812	89.6	0.902	0.698	No	Yes	No	No	No	No	No	0.824
65084	Gallocatechin	− 2.969	54.128	1.301	− 1.377	No	No	No	No	No	No	No	0.328
92729	gamma-Tocopherol	− 7.602	90.043	0.732	0.739	No	Yes	No	No	No	No	No	0.821
91746804	28-Isoavenasterol acetate	− 6.996	96.606	0.071	0.72	No	Yes	No	No	No	No	No	0.48
44593339	Roridin E	− 5.387	96.706	0.118	− 0.802	No	Yes	No	No	No	No	No	0.715
5484202	Stigmast-4-en-3-one	− 6.212	98.556	− 0.141	0.847	No	No	No	No	No	No	No	0.56
53477901	(3S,9S,10R,13R,14S,17R)-10,13-dimethyl-17-[(2S)-6-methyl-5-methylideneheptan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-3-ol	− 7.041	94.691	0.426	0.777	No	Yes	No	No	No	No	No	0.604

17 compounds tested, 65 % (11/17) were identified as CYP3A4 substrates, while only compound CID:72201063 found to be a potent CYP3A4 inhibitor. Moreover, CID:5280647, CID:5281680, and CID:145858 were found to be strong inhibitors of CYP1A2, while CID:72201063, and CID:145858 were identified as inhibitors of CYP2C19. None of the substances exhibited inhibitory activity against CYP2C9. The bioavailability and half-life of xenobiotic compounds are affected by total clearance (TC). The frequency and dosage of a particular drug are greatly affected by this factor. Hence, the prediction establishes a basis for the initial dose in in vivo studies and aids in evaluating the prospect of therapeutic dosing ⁷¹. A lower clearance index value suggests a longer persistence of drugs in the body. In this study, we scrutinized the excretion properties to understand the stability of the compounds as drugs in the body prior to their elimination. The predictive values indicate that the total clearance index of certain compounds falls within a range from -0.196 to 0.361, suggesting very low excretion for these specific compounds, which make up 41 % (7/17) of the total. In contrast, the remaining compounds are expected to have a moderate to high excretion rate.

We have identified five compounds with favorable ADMET characteristics from a pool of 17 compounds for additional toxicity evaluation. After evaluating the in silico ADMET properties of compounds, it was discovered that five particular compounds: Gallicocatechin, Epigallocatechin, Hesperetin, Gossypetin and Quercetagenin met all the assessed pharmacokinetic criteria. Each of the five compounds examined showed a favorable absorption rate for intestinal absorption in humans and was found to be water-soluble. All the selected compounds met the BBB criteria. Compounds are likely distributed throughout the body based on the VDss parameter. Because none of the five compounds inhibit any of the several types of CYP isoenzymes, they are all readily excreted from the body after a therapeutic response. The low score in the total clearance analysis suggests that these lead compounds are expected to have minimal drug excretion, leading to a prolonged presence in the body. The comprehensive pharmacokinetic and ADME analyses indicated that all five selected compounds can be converted into drugs and are appropriate for toxicological evaluation and further testing.

3.4. Toxicity analysis

Drug toxicity is a significant concern and poses a serious health issue in the medical field, as harmful side effects are often cited as the main reasons for failures in the development of new drugs in later stages ⁴¹. Toxicity is the measure of the potential harm that a chemical substance might do to an organism or its various components, including cells and organs. Hence, the timely identification of toxicity would be immensely advantageous ⁷². Table 3 highlights the outcome regarding the potential toxicity of all molecules being examined. Based on the outcomes, it is anticipated that none of the compounds will exhibit hepatotoxicity, hERG I and II (human Ether-a-go-go-Related gene) inhibition, skin sensitization, or AMES toxicity. The highest dose that most

patients can tolerate is referred to as the maximally tolerated dose (MTD). The compounds with an MTD value exceeding 0.477 logs (mg/kg/day) may be suitable for higher dosage administration. In this inquiry, the highest dose that humans could tolerate was 0.506, 0.473, 0.486, 0.25, and 0.506 Log mg/kg/day for epigallocatechin, gossypetin, quercetagenin, hesperetin, and gallicocatechin, pointing to a moderate level of dosage adhering to the established protocols. The Oral Rat Acute Toxicity (LD50) values for epigallocatechin, gossypetin, quercetagenin, hesperetin, and gallicocatechin were 2.492, 2.527, 2.537, 2.042, and 2.492 mol/kg, respectively, revealing an adequate safety profile. Based on the ProTox-3.0 server, the compounds (epigallocatechin, gossypetin, quercetagenin, hesperetin, and gallicocatechin) are reported to be non-hepatotoxic, non-neurotoxic, non-cytotoxic, and non-immunotoxic. All the toxicity parameters evaluated were equally reassuring, confirming a high level of safety and indicating that these compounds are strong contenders for further research.

3.5. Evaluation of the bioactivity

Inspection of the potential biological activity of compounds is of the utmost importance and can be accomplished by evaluating their bioactivity score. To properly evaluate the bioactivity score of prospective drugs, it is vital to scrutinize the impact on four key factors: the control of ion channels, the interaction between ligands and G protein-coupled receptors, the inhibition of protease, kinase, and other enzymes, and the impact on nuclear receptor ligands. The molecular inspiration cheminformatics program was employed to compute the pharmacological activity of the five selected compounds, and the outcomes are listed in Table 4. In cases where the compound's score exceeds 0.00, it is more likely to show significantly higher physiological activity. Predictions illustrate a moderate level of activity when the score is within the range of -0.50 to 0. From a biological standpoint, molecules with scores below -0.50 are deemed inactive ⁷³. Based on the observations, it appears that multiple pathways can influence the physiological effects of the compounds. This could also be a result of interactions involving GPCR ligands, nuclear receptor ligands, and the inhibition of protease and other enzymes. The bioactivity evaluations of the compound revealed a strong correlation with all the drug targets in the current study, with all selected compounds demonstrating highly promising bioactivity scores.

3.6. Molecular docking analysis

Molecular docking analysis facilitates the comprehension of interactions between ligands and proteins, enabling the prediction of a most likely binding pose for each compound in the protein's active region. Additionally, the outcomes of docking offers an effective means to create new potential inhibitors. Table 5 displays the 2D structures of the ligands with binding affinity scores from both AutoDock Vina and BINDSURF tools. These tools give us the same results in terms of binding affinity and interaction residues. All of our selected com-

Table 3

Evaluation of the toxicity profile of the five potent compounds using pkCSM and Protox-3.0 Online server.

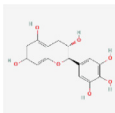
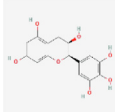
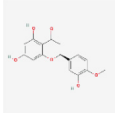
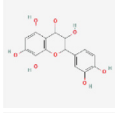
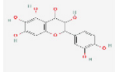
Compound CID	Phytochemical Name	pkCSM						ProTox			
		AT	MTD	hERG I inhibitor	hERG II inhibitor	ORAT	SS	Hepatotoxicity	Neurotoxicity	Immunotoxicity	Cytotoxicity
65084	Gallicocatechin	No	No	No	No	2.492	No	Inactive	Inactive	Inactive	Inactive
72277	Epigallocatechin	No	0.506	No	No	2.492	No	Inactive	Inactive	Inactive	Inactive
72281	Hesperetin	No	0.25	No	No	2.042	No	Inactive	Inactive	Active	Inactive
5280647	Gossypetin	No	0.473	No	No	2.527	No	Inactive	Inactive	Inactive	Inactive
5281680	Quercetagenin	No	0.486	No	No	2.537	No	Inactive	Inactive	Inactive	Inactive

AT: AMES Toxicity; MTD: Maximum Tolerated Dose; hERG: human Ether-a-go-go-Related gene; ORAT: Oral Rat Acute Toxicity; HT: Hepatotoxicity; SS: Skin Sensitization.

Table 4
Evaluation of the potent compound’s bioactivity using Molinspiration Cheminformatics Software.

Compound CID	Phytochemical Name	GPCR ligand	Ion channel modulator	Kinase inhibitor	Nuclear receptor ligand	Protease inhibitor	Enzyme inhibitor
65084	Gallocatechin	0.40	0.14	0.14	0.57	0.29	0.49
72277	Epigallocatechin	0.40	0.14	0.14	0.57	0.29	0.49
72281	Hesperetin	0.04	−0.26	−0.20	0.38	−0.13	0.16
5280647	Gossypetin	−0.09	−0.18	0.30	0.30	−0.23	0.30
5281680	Quercetagetin	−0.11	−0.28	0.26	0.22	−0.30	0.26

Table 5
Molecular docking results of five selected compounds in AutoDock Vina and BIND SURF with their 2D structures.

Compound CID	Phytochemical Name	2D Structure	EGFR		VEGFR-2	
			AutoDockVina	BINDSURF	AutoDockVina	BINDSURF
65084	Gallocatechin		− 8.2	− 8.4	− 8.1	− 8.0
72277	Epigallocatechin		− 8.0	− 7.9	− 7.8	− 7.8
72281	Hesperetin		− 8.3	− 8.1	− 8.1	− 8.2
5280647	Gossypetin		− 7.9	− 7.9	− 8.6	− 8.6
5281680	Quercetagetin		− 8.0	− 8.0	− 8.7	− 8.6

pounds have a docking score −7.8 to −8.7 kcal/mol with EGFR and VEGFR-2 receptors. The control Erlotinib showed a docking score of −7.5 kcal/mol, while Bevacizumab + Rituximab exhibited −6.7 kcal/mol. However, as our selected compounds have a higher negative value than controls, they can be more effective inhibitors than these two well-known drugs. Fig. 4 to Fig. 8 displays the interactive residues following docking in AutoDock Vina, and interacting residues of controls with the receptors are presented in supplementary file (Fig.S1). In addition, the supplementary file (Figure S3 and S4) contains the interactive residues following BINDSURF docking. All the binding residues and their specific interaction types with the compounds are presented in Table 6 and Table 7. Moreover, in our re-docking analysis, the co-ligands were accurately replicated into EGFR and VEGFR2 receptors, with corresponding RMSD values of 0.87 Å and 1.23 Å compared to the co-crystal poses. It is reported that if the RMSD difference is less than 2.0 Å in re-docking, the scoring function employed was effective and validated the docking approach ^{74,75}.

The human EGFR protein contains a sequence of 1210 amino acids (AA) and has a molecular weight of around 134 kDa. The tyrosine kinase domain (TKD) of this protein are divides into two lobes, encompassing residues 669 AA – 979 AA. The N-lobe consists of five β-sheet strands (β1–5) and a αC-helix that stretches from position 729 AA to 744 AA. The C-lobe is larger than the N-lobe and is made up of five α helices (αI, αH, αG, αF, and αE) ^{76,77}. The C-lobe plays an important role in EGFR activation through residues from Asp831-Val852, which has a unique Asp-Phe-Gly (DFG) motif at its core. In the active state, the DFG motif assumes a distinct shape that facilitates the binding of

ATP and the transfer of a phosphate group to tyrosine residues ⁷⁸. Phosphorylation of the TKD is crucial for the signaling pathway of EGFR. This domain is responsible for the catalytic function of EGFR. The kinase domain also has a highly conserved ATP-binding region ⁷⁹. According to our structural analysis study of the EGFR tyrosine kinase domain, certain residues in the TK domain's α-helix, β-sheets, and hinge regions are located in the ATP binding site, specifically residues Leu_844, Met_793, Ala_743, Val_726, and Leu_718. These residues are located at β1-3, β-6, and the hinge region and have the most frequent contact with the ligands in all crystal structures. This region is the most important part for the design of inhibitors due to its direct role in the aberrant function of the protein associated with cancer ⁸⁰. Our molecular docking analysis showed, the compound gallocatechin (CID: 65084) forms conventional hydrogen bonds with EGFR residues LEU_788, MET_793, and THR_854, and the compound epigallocatechin (CID: 72277) forms conventional hydrogen bonds with residues LEU_788 and MET_790. Both compounds interact with residues LEU_718, VAL_726, ALA_743, and LYS_745 by the pi-alkyl/alkyl bond. The EGFR residue MET_790 makes a pi-alkyl and pi-sulfur bond, and residue LEU_844 makes a pi-sigma bond with gallocatechin. In contrast, epigallocatechin interacts with residue MET_790 by also forming a pi-sulfur bond, and residue LEU_844 by forming both pi-sigma and pi-alkyl bonds and these outcomes are presented in Fig. 4 and Fig. 5.

The compound hesperetin (CID: 72281) has the most advantageous interactions with EGFR as an inhibitor because it binds to both the mutant residues MET_790 and SER_797 inside the protein.

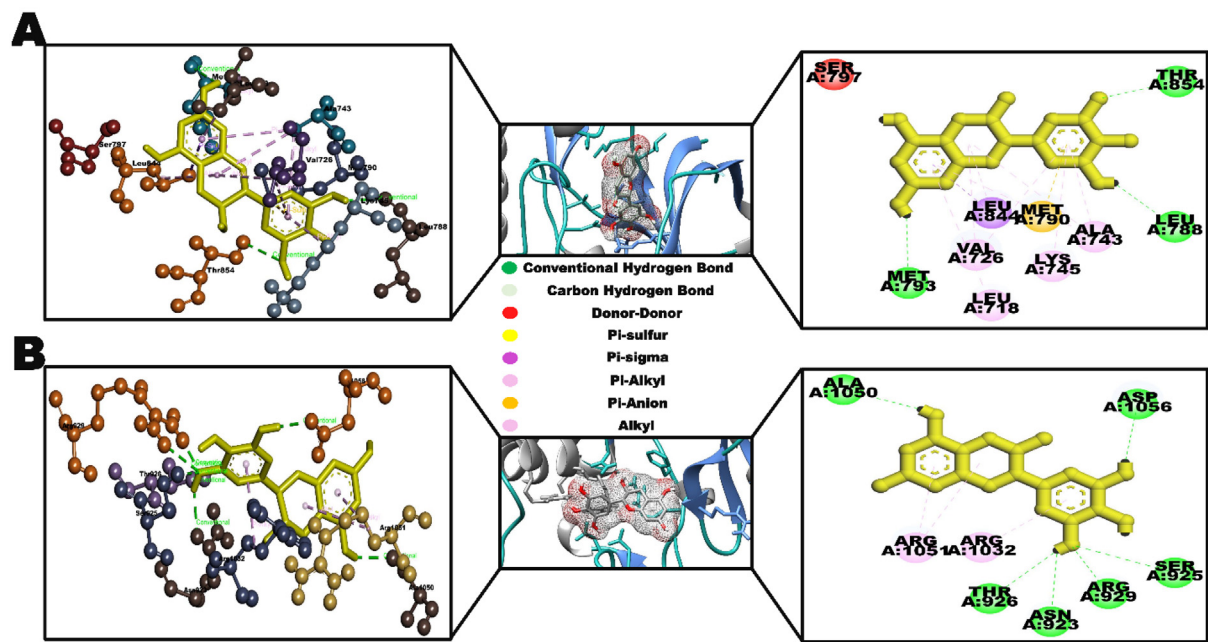


Fig. 4. Depiction of the 3D binding orientation of the gallocatechin (middle) within the binding cavity of EGFR (A) and VEGFR-2 (B) and the corresponding residues involved in the binding of the compounds, presented in 3D (left) and 2D (right).

Table 6
Residues involved in the binding of selected compounds and EGFR receptors.

Protein	Ligand	H-bond interaction	Hydrophobic interaction
EGFR (5ZWJ)	Gallocatechin	LEU788, MET793, THR854	Pi-Sigma: LEU844 Pi-Alkyl: LEU718, VAL726, ALA743, LYS745 Pi-Sulfur: MET790
	Epigallocatechin	LEU788, MET790	Pi-Sigma: LEU844
	Hesperetin	ALA743, LYS745, SER797	Pi-Alkyl: LEU718, VAL726, LYS745, ALA743
	Gossypetin	LEU788, ASP855, GLY796	Pi-Sigma: LEU844 Pi-Alkyl: LEU718, VAL726, LEU777, LEU778
	Quercetagenin	LYS745, THR854, PHE856	Pi-Sigma: LEU844 Pi-Alkyl: VAL726, ALA743, LYS745 Pi-Sulfur: MET790 Pi-Sigma: LEU788 Pi-Alkyl: LEU747, MET766, LEU777, LEU858 Pi-Anion: ASP855

Table 7
Residues involved in the binding of selected compounds and VEGFR-2 receptors.

Protein	Ligand	H-bond interaction	Hydrophobic interaction
VEGFR-2 (4ASD)	Gallocatechin	ASN923, SER925, THR926, ARG929, ALA1050, ASP1056	Pi-Alkyl: ARG1032, ARG1051
	Epigallocatechin	ASN923, SER925, THR926, ARG929, ARG1051, LYS1055, ASP1056	N/A
	Hesperetin	ASN923, SER925, THR926, ARG929, ASP1058	Pi-Sigma: ARG1051
	Gossypetin	ASN923, ARG929, ARG842, ARG1051	Pi-Alkyl: ARG1032
	Quercetagenin	ARG842, THR926, ARG1032, ALA1050, PHE1047	Pi-Alkyl: ARG1032 N/A

This compound interacts with residues ALA_743, LYS_745, and SER_797 with conventional hydrogen bonds and forms pi-alkyl/alkyl bonds with residues LEU_718, VAL_726, ALA_743, and LYS_745. In addition, residue MET_790 forms a pi-sulfur and residue LEU_844 forms a pi-sigma bond with hesperetin (Fig. 6). The compound gossypetin (CID: 5280647) interacts with residues LEU_788 and ASP_855 by forming conventional hydrogen bonds and with residue GLY_796 by a carbon hydrogen bond. The residues VAL_726, ALA_743, LYS_745, and LEU_844 interact by pi-alkyl/alkyl bonds with this compound. This compound also forms a pi-sulfur bond with residue MET_790 and a

pi-sigma bond with residue LEU_844 of the receptor (Fig. 7). Quercetagenin (CID: 5281680) interacts with LYS_745, THR_854, and PHE_856 by conventional hydrogen bonding and with LEU_747, MET_766, LEU_777, LEU_788, and LEU_858 by several pi-alkyl/alkyl bonds. In addition, this compound interacts with residue LEU_788 by Pi-Sigma and residue ASP_855 by an Anion bond (Fig. 8). Conversely, the whole VEGFR-2 consists of 1356 amino acids, which include a mature protein (20~1356 AA) and a signal peptide (1~19 AA)⁸¹. The catalytic tyrosine kinase domain (TKD) is the most conserved region within the VEGFR-2 protein. This domain includes

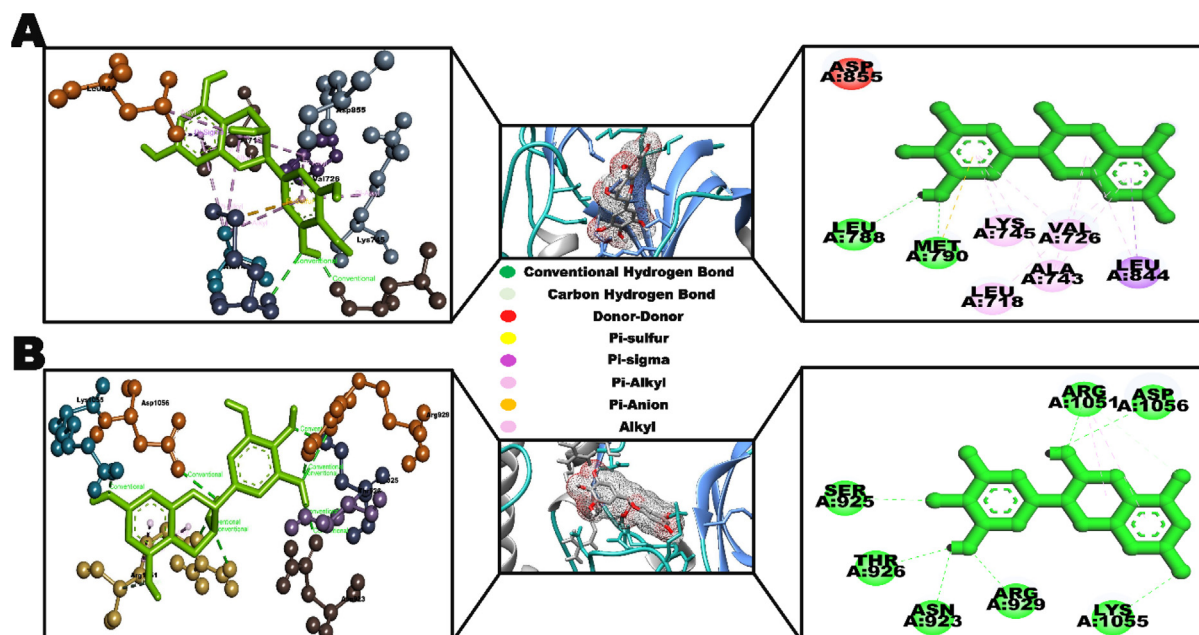


Fig. 5. Depiction of the 3D binding orientation of the epigallocatechin (middle) within the binding cavity of EGFR (A) and VEGFR-2 (B) and the corresponding residues involved in the binding of the compounds, presented in 3D (left) and 2D (right).

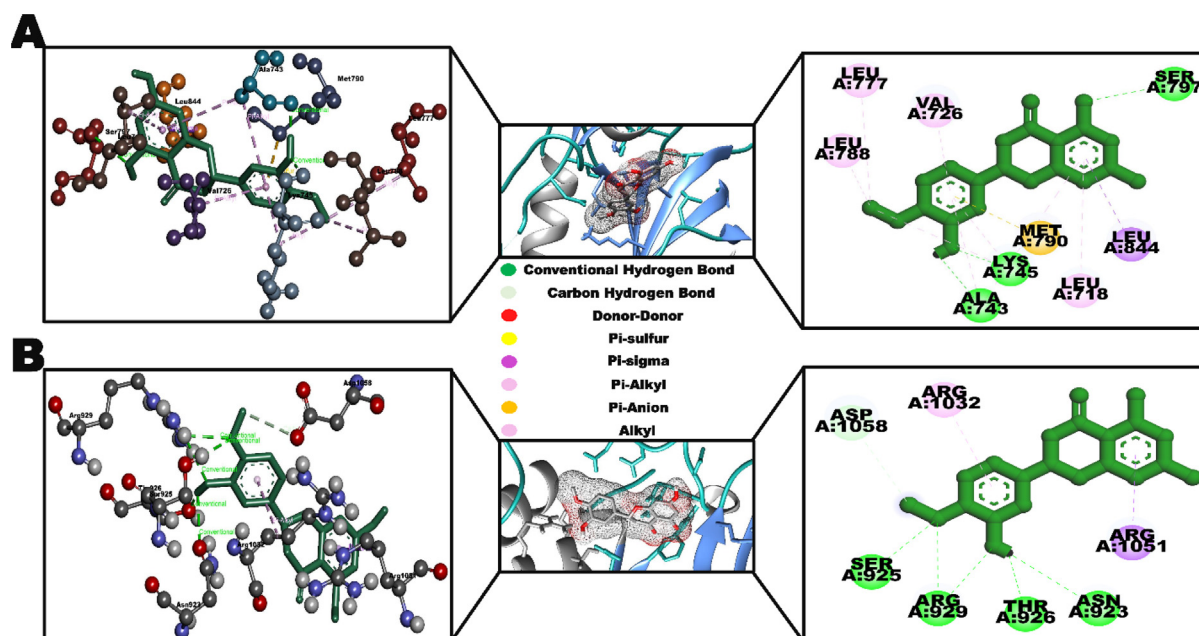


Fig. 6. Depiction of the 3D binding orientation of the hesperetin (middle) within the binding cavity of EGFR (A) and VEGFR-2 (B) and the corresponding residues involved in the binding of the compounds, presented in 3D (left) and 2D (right).

an ATP binding domain (TKD1), a phosphotransferase domain (TKD2), a flexible C-terminal domain (CTD) and a kinase insert domain (KID). The N-terminus of the cytoplasmic tyrosine kinase domain contains a hydrophobic pocket that includes a glycine-rich ATP phosphate receptor loop (GXGXXG, 841AA–846 AA) within the β -sheet structures⁸². There are several α -helical structures in TKD's C-terminal. These include an activation loop (A-loop, 1045 AA–1075 AA) and a catalytic loop (HRDLAARN, 1026 AA–1033 AA), which are very important for VEGFR-2's catalytic properties⁸³. In our study, the compound galliccatechin interacts with the VEGFR-2 protein through six conventional

hydrogen bonds. These bonds include residues ASN_923, SER_925, THR_926, ARG_929, ALA_1050, and ASP_1056. It also forms three pi-alkyl bonds with ARG residues 1032 and 1051. The receptor maintains a strong interaction with epigallocatechin via eight conventional hydrogen bonds with residues ASP_923, SER_925, THR_926, ARG_929, ARG_1051, LYS_1055, and ASP_1056, as well as two pi-alkyl bonds with ARG_1051.

Hesperetin shows contiguity with residues ASN_923, SER_925, THR_926, and ARG_929 by a conventional hydrogen bond and with ASP_1058 by a carbon hydrogen bond. The residues ARG_842,

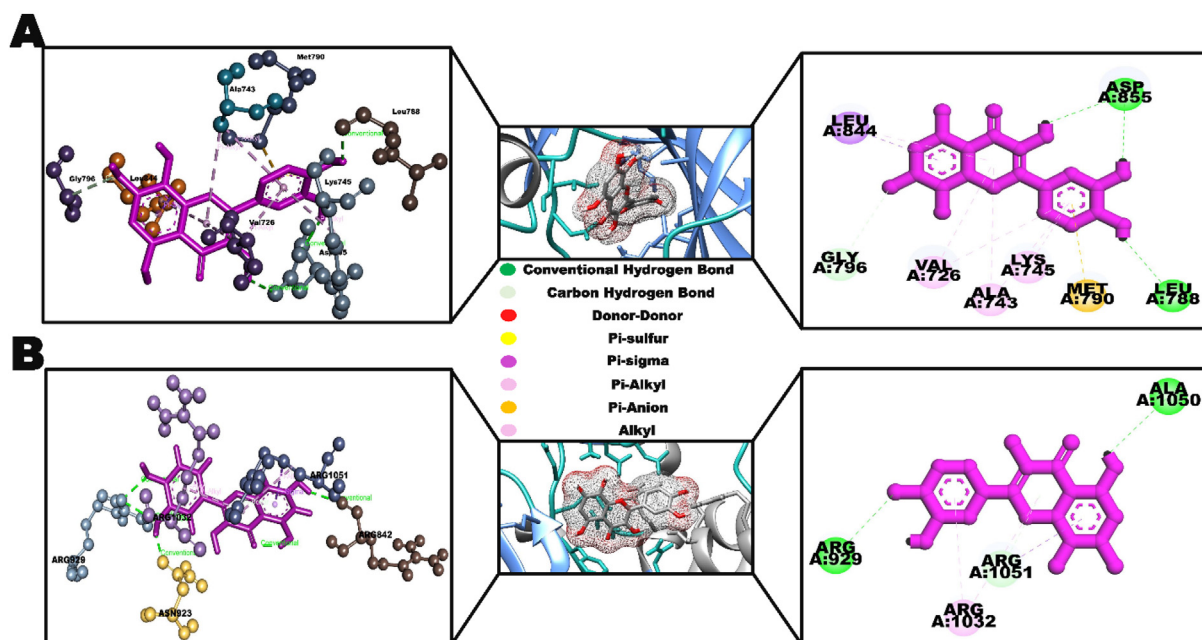


Fig. 7. Depiction of the 3D binding orientation of the quercetagenin (middle) within the binding cavity of EGFR (A) and VEGFR-2 (B) and the corresponding residues involved in the binding of the compounds, presented in 3D (left) and 2D (right).

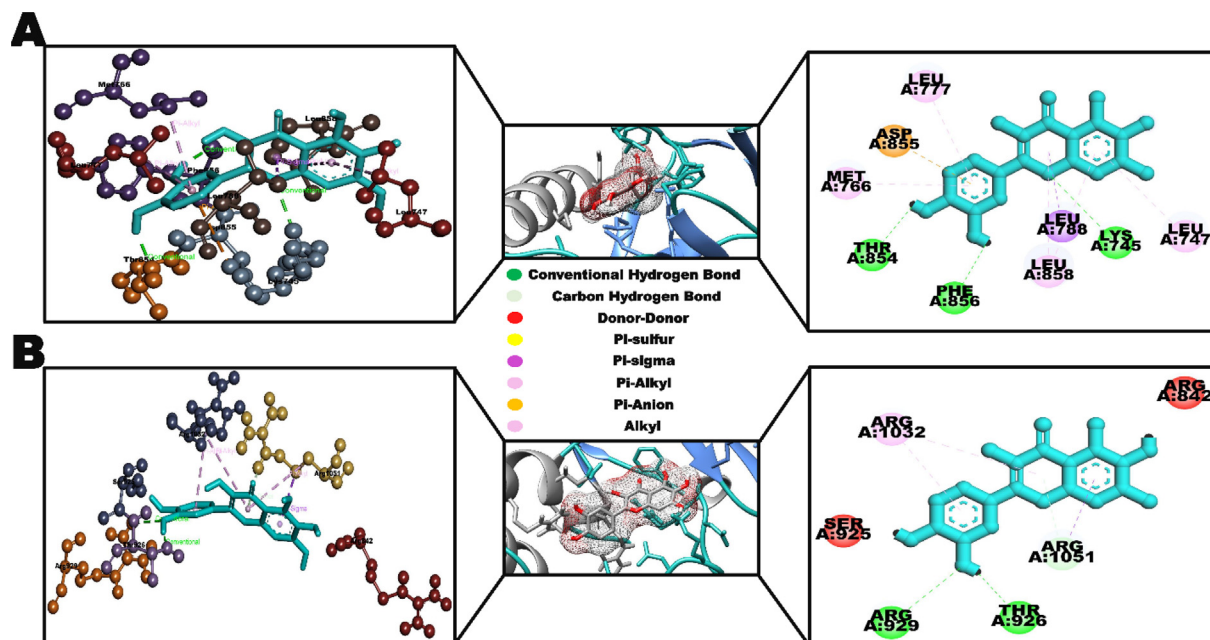


Fig. 8. Depiction of the 3D binding orientation of the gossypetin (middle) within the binding cavity of EGFR (A) and VEGFR-2 (B) and the corresponding residues involved in the binding of the compounds, presented in 3D (left) and 2D (right).

ASN_923, and ARG_926 form four conventional hydrogen bonds with the compound gossypetin. The residue ARG_1051 forms a carbon hydrogen and a pi-alkyl bond, while ARG_1032 forms two pi-alkyl/alkyl bonds with this compound. There are four conventional hydrogen bonds and a carbon hydrogen bond between quercetagenin and residues ARG_842, THR_926, ARG_1032, ALA_1050, and PHE_1047 of the receptor. In addition, residue ARG_1032 forms two pi-alkyl bonds with the compound. In order to effectively inhibit EGFR T790M/C797S mutations, next-generation NSCLC drugs must compete with ATP while also being able to target the mutant EGFR. This is a challeng-

ing task due to the strong binding of ATP to EGFR T790M. All of our compounds, except quercetagenin, interact strongly with MET_790 through multiple hydrogen and hydrophobic bonds, indicating that they can effectively compete with ATP for this specific mutant residue. The compounds also interact with the DFG motif's residues. Specific inhibitors that bind to the residues of this motif can efficiently block the attachment of phosphate groups to tyrosine residues, thus inhibiting the uncontrollable cell proliferation caused by EGFR mutations.

On the other hand, except for epigallocatechin, all the selected compounds have a strong interaction with the ARG_1032 residue of

VEGFR-2. It is crucial because the VEGFR-2 residue ARG_1032 is now a potential target for NSCLC cancer therapy. Numerous studies have shown that alteration in this specific residue of VEGFR-2 are associated with resistance against several cancer treatments and enhanced tumor malignancy. All of our compounds also bind to the active site residues of the catalytic kinase domain, specifically the ATP binding residues of both receptors. This makes them more effective as both EGFR and VEGFR-2 inhibitors.

3.7. Molecular dynamics simulation

Molecular dynamics (MD) simulation is utilized to discern the dynamics of a protein–ligand interaction within a defined time frame. MD simulation provides an accurate depiction of the dynamic behavior of protein–ligand interactions and enables a thorough comprehension of their demeanor in a biologically significant environment. This is an important and well-established structure-based approach to understanding the atomic-level properties of protein–ligand complexes through deviation, fluctuation, and intermolecular interaction analysis^{84,85}. Furthermore, by confirming that changes in temperature or pressure have no effect on the structure's conformation, MD simulation will validate the docking complex's stability⁸⁶. During the whole simulation period, it was apparent that all the selected ligands from our study remained attached to the protein within the binding pocket which implies that even in a physiological environment, the ligands will establish persistent interactions with the protein. In addition, some other important analyses, such as the root mean square deviation (RMSD) of atomic distances, root mean square fluctuation (RMSF), molecular surface area (MolSA), radius of gyration (Rg), protein–ligand contact map, and polar surface area (PSA), were analyzed following the MD simulation to enhance understanding of the resilience and interactions of the protein–ligand complex. All the results of these analyses are presented from Fig. 9 to Fig. 13.

3.7.1. RMSD analysis

The stability and alterations in the structural features of a docked complex can be assessed by calculating the root mean square deviation (RMSD). The intended behavior of the protein–ligand complex is reliant upon the folding and shifting of the atoms within the overall protein structure. The RMSD analysis enabled us to understand the structural alterations that transpired in the side-chain, C α , and backbone atoms of the proteins throughout the simulation period. Minimal fluctuations in the RMSD value are considered favorable indications of the docked complex's stability. If the protein–ligand complex shows consistent changes in the range of 1–3 Å than apo-protein, it implies the system has slowly reached equilibrium, and fluctuations in this range are entirely acceptable⁸⁷. But significant shifts above the aforementioned magnitude suggest that the protein undergoes a conformational change throughout the simulation. In our study, the RMSD of carbon alpha atoms in apo-proteins was used as a reference frame for the RMSD calculations. Fig. 9 (A and B) show how the RMSD values changed over time for the C-alpha atoms of the ligand-bound proteins and reference proteins. The RMSD value of EGFR apo-protein (5ZWJ) fluctuates between 1.77 Å and 3.52 Å until 55 ns, after which, with some minor fluctuations, it stabilizes with a mean RMSD of 3.18 Å until 175 ns. Then the value shows a small downward peak until the end of the simulation. The control, Erlotinib, did not show any major fluctuations throughout the simulation period and remained stable with the protein. The RMSD value of the EGFR protein with compound gallocatechin is lower than that of the reference apo-protein, which implies the ligand stabilizes protein structure by interacting with favorable residues. The highest and lowest RMSD values for the protein–ligand complex are 3.58 Å and 1.41 Å, with an average RMSD of 2.55 Å, while the reference protein shows values ranging between 4.30 Å and 1.67 Å with a mean value of 3.01 Å. The EGFR-epigallocatechin complex shows almost the same RMSD values as the reference protein until 25 ns, after which the RMSD of the complex gradually increases up to the end of the simulation. Following the

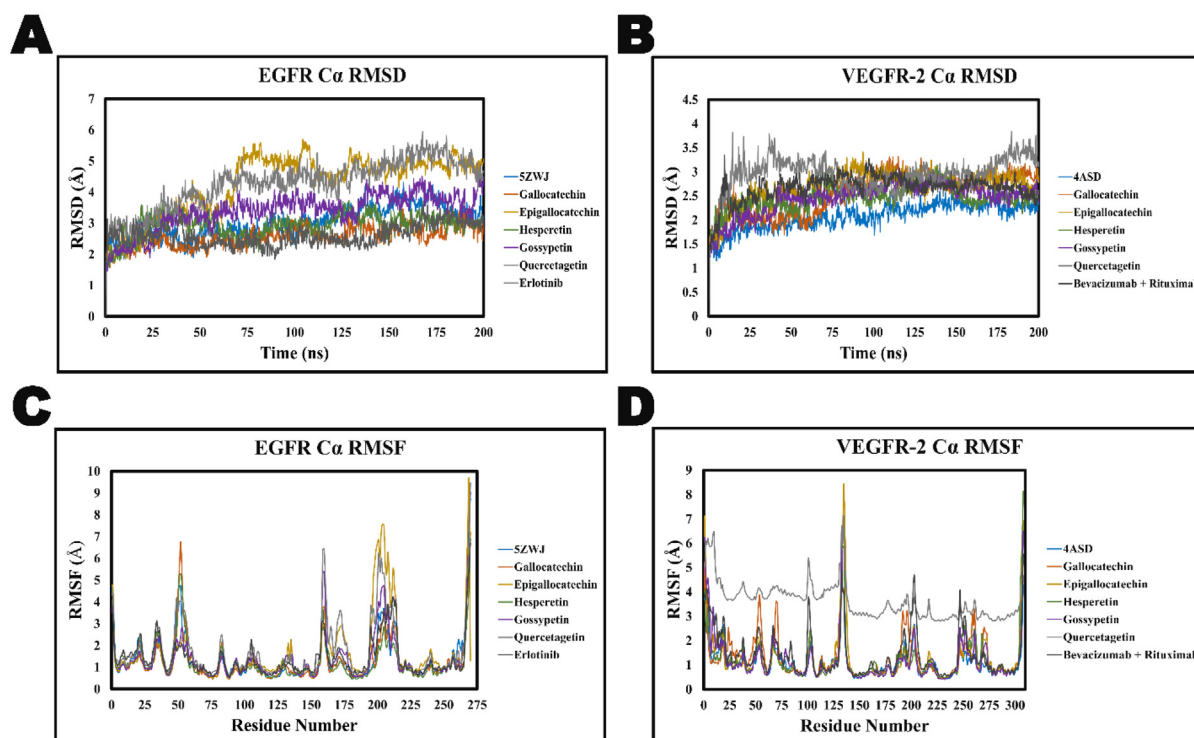


Fig. 9. RMSD (A and B) and RMSF (C and D) plots of the EGFR and VEGFR-2 carbon alpha in complex with selected five compounds for 200 ns.

peaks of 5.45 Å, 5.56 Å, and 5.59 Å at 73 ns, 79 ns, and 82 ns, the complex reached its maximum RMSD of 5.69 Å at 104 ns. The complex exhibits mild variations from 40 ns to 140 ns but then stabilizes until 200 ns with a small fluctuation at 183 ns. The apo-protein shows RMSD values of 2.94 Å, 2.84 Å, 2.46 Å, and 3.19 Å at time points 73 ns, 79 ns, 82 ns, and 104 ns, indicating that the RMSD change of the complex falls within the permissible range of 1–3 Å. The EGFR protein with compound hesperetin shows some fluctuations from the starting point of the simulation until 25 ns. The complex exhibits a slight rise in RMSD value from 26 ns to 55 ns, and then the value decreased compared to the reference protein until the end of the simulation, with a modest increase observed at points 98 ns to 100 ns. The mean RMSD for this complex is 2.92 Å, and the maximum and minimum values are 3.65 Å and 2.08 Å at 30 ns and 81 ns, respectively. The complex EGFR-gossypetin has a lower RMSD than apo-protein until 25 ns. Subsequently, the complex exhibits a consistent and enduring RMSD peak from 26 ns to 135 ns. The complex had some oscillations in the value from 136 ns to the end of the simulation. The average RMSD of this complex is 3.38 Å. At time point 1.38 ns, the complex has a higher value of 4.41 Å, while at time point 1.2 ns, it has a lower value of 1.67 Å. The RMSD values of the EGFR-quercetagen complex show a progressive increase during the simulation, though they remain within the allowed range at all times. The maximum value of RMSD for the complex is 5.92 Å at 167 ns, while the reference protein's RMSD at this point is 3.54 Å. The lowest RMSD count for the complex at time point 1 ns is 1.55 Å, while the reference protein's RMSD is 1.67 Å. The complex has a mean RMSD of 4.27 Å.

The RMSD value of this protein is less than 3 Å with all the compounds except compound quercetagen, which means the protein remains structurally stable with the compounds. There are small fluctuations in the VEGFR-2 apo-protein (4ASD) up to 20 ns, after which it reaches a stable state. The protein has a bit higher RMSD value than the adjacent time frame between 110 and 175 ns, and it also reaches its maximum RMSD value at 2.76 Å during this time frame. The control Bevacizumab + Rituximab exhibited significant fluctuations, ranging from 1.2 to 3.48, during the first 23 ns, before stabilizing at an average RMSD of 2.67 until the end of the simulation. The lowest value of the VEGFR-2 and gallicocatechin complex is 1.15 Å, while its highest value is 3.31 Å, observed at 4 ns and 115 ns, respectively. The mean RMSD of the complex is 2.55 Å. The RMSD of the VEGFR-2 and compound epigallocatechin complex slowly increases over the course of the simulation compared to the reference protein. The mean RMSD of the complex is 2.67 Å, while the mean RMSD of the VEGFR-2 apo protein is 2.07 Å, but the RMSD value of the complex never exceeds the acceptable range of the RMSD difference. During 93 ns, the highest difference was observed between the complex and apo-protein, while the value was 2.01 Å for apo and 3.41 Å for complex. The VEGFR-2 and hesperetin complex remain stable with slight variations until 110 ns, then show some upper peaks between 120 ns and 150 ns. The RMSD value of the complex is approximately equal to that of the apo-protein for the remaining period of the simulation. The average RMSD for the complex is 2.45 Å, while the maximum value is 3.15 Å at 132 ns and the minimum value is 1.3 Å at 1 ns. There is approximately an equal RMSD until 40 ns compared to the apo-protein for the complex VEGFR-2-gossypetin. Subsequently, the complex attains stability and exhibits a slightly higher RMSD value compared to the reference protein. The complex has the highest value of 3.00 Å at 110 ns and the lowest value of 1.23 Å at 1 ns, with a mean RMSD of 2.46 Å. The protein with compound quercetagen exhibits instability until 40 ns, as some noticeable peaks were observed in this time frame. Then the complex's structure remained stable until 80 ns, after which the value slightly decreased from the previous time frame until 160 ns. The graph showed a small increase again after 160 ns. The total RMSD of the complex is marginally greater than that of the reference protein. A peak of 3.83 Å is noticed for the complex at 183 ns, but the reference protein's RMSD was 2.19 Å at the same time.

A second peak of 3.81 Å is observed at 14 ns, while the reference protein's value is 1.59 Å. Accordingly, in the complex, the next higher peaks are 3.79 Å, 3.75 Å, and 3.72 Å at 36 ns, 198 ns, and 21 ns, respectively. The RMSD of the apo-protein at 36 ns, 198 ns, and 21 ns are 1.82 Å, 2.41 Å, and 1.81 Å, which means the RMSD changes never exceed the acceptable range. Overall, our RMSD analysis showed that none of the complexes has a RMSD value that is significantly larger than the reference proteins and higher than the acceptance range. This indicates that the docked structures remain stable and the ligands are effectively interacting with the protein molecules. In addition, from the RMSD values of the protein–ligand complexes, it is apparent that the ligand's binding to the proteins does not significantly change the conformational structure of the proteins.

3.7.2. RMSF analysis

The root mean square fluctuation (RMSF) analysis is an effective approach for identifying fluctuations of the specific amino acid residues in the protein chain, notably those in the active region. The RMSF represents the mean changes observed in every single amino acid residue of the protein throughout the simulation period. Fig. 9 (C and D) displays the residue-wise RMSF values of EGFR and VEGFR-2, both proteins, in their unbound and bound states with the selected compounds. Typically, the N and C-terminal regions of the protein exhibit higher fluctuations compared to other regions of the protein, and the secondary structure constituents, such as beta strands and alpha helices, are normally more firm than the loosely organized portion (loop region) of the protein. Consequently, they show fewer fluctuations compared to loop regions. Typically, large peaks in an RMSF plot represent targeted protein residues with the highest level of variation. The RMSF value of the N- and C-terminal regions of VEGFR-2 is slightly high with all the compounds, though EGFR does not show any fluctuations in these regions with any compound. The selected controls for both proteins showed about the same RMSF outcomes as our selected ligands. Throughout the simulation period, we did not observe any significant fluctuations in the RMSF value, neither with controls nor with the studied compounds. The EGFR protein, with few exceptions, shows fewer fluctuations compared to the reference protein during the whole simulation period, which indicates the ligand binding efficiently stabilizes the protein. The protein EGFR does not show any noticeable fluctuation peaks with the compounds gallicocatechin, hesperetin, and gossypetin compared to the apo-protein 5ZWJ. With compound epigallocatechin, the EGFR protein shows some higher peaks from residue LYS_913 to GLU_931, and with compound quercetagen, it shows some moderate fluctuations at residues VAL_876, PRO_877, and GLY_911 to ILE_926. However, these protein residues have not been involved in ligand interactions. More importantly, the RMSF value of the EGFR protein's mutant residues is lower than that of the apo-protein. This suggests that these compounds may be able to stop the abnormal activity and bring the protein back to its normal function, or they may make the protein more receptive to drug bindings, which could help overcome drug resistance by stabilizing the mutant residues. The VEGFR-2 protein with the compounds gallicocatechin and epigallocatechin shows some mild fluctuations at some points, including the N and C terminal regions, but the fluctuated residues do not interact with the ligands. The compounds hesperetin and gossypetin with the protein VEGFR-2 do not show any noticeable peak in the graph except an N and C terminal peak. The RMSF value of the VEGFR-2 protein with the compound quercetagen has higher peaks in all of its residues than the reference protein. But the RMSF value of most of the residues never crosses the acceptable range. Apart from the N and C terminal's elevated peaks, this complex shows slightly higher peaks above the allowed range of RMSF at residues ALA_844, LYS_868, GLU_885, LEU_901, ILE_915, SER_925, and ILE_1034. However, most of these residues are from the loop region of the protein. The other protein residues maintain a constant acceptable range of RMSF values during the whole simulation period. Our primary struc-

ture analysis of VEGFR-2 protein reveals that the residues where VEGFR-2 protein shows fluctuations in complex with the compounds, none of these fluctuated residues are involved in the active site of the protein. If there is no fluctuation observed in the active site or the fluctuations in the active region are less that indicates the protein is structurally rigid with the compounds⁸⁸.

The RMSF analysis of the EGFR and VEGFR-2 proteins, along with the compounds gallicocatechin, epigallocatechin, hesperetin, gossypetin, and quercetagetin revealed that the ligands mostly remained confined within the binding pockets and interacted with the amino acid residues of the active sites. Except at some points of the VEGFR-2-quercetagetin complex, both the proteins, along with all the compounds, never fluctuate above the acceptable range of 3 Å at any point. Additionally, as a result of the minimal movement of protein residues

with all the compounds over the simulation timeframe, the entire structure of proteins maintained a compact form in these complexes.

3.7.3. Protein–ligand interaction analysis

The amino acid and atomic-level interactions analysis between protein and ligand provides valuable insights on the conformational robustness and correlation of the protein, hence enhancing the comprehension of simulations. The amino acid interactions analysis reveals which residues are most important for the protein–ligand interaction, while the atomic-level interaction analysis is crucial for accurately predicting the mode of attachment of the protein and ligands throughout the whole simulation period⁸⁹. Therefore, the interacted amino acid residues of the protein with their Interactive Fraction Value (IFV) are presented as a stacked bar chart in Fig. 10 for EGFR

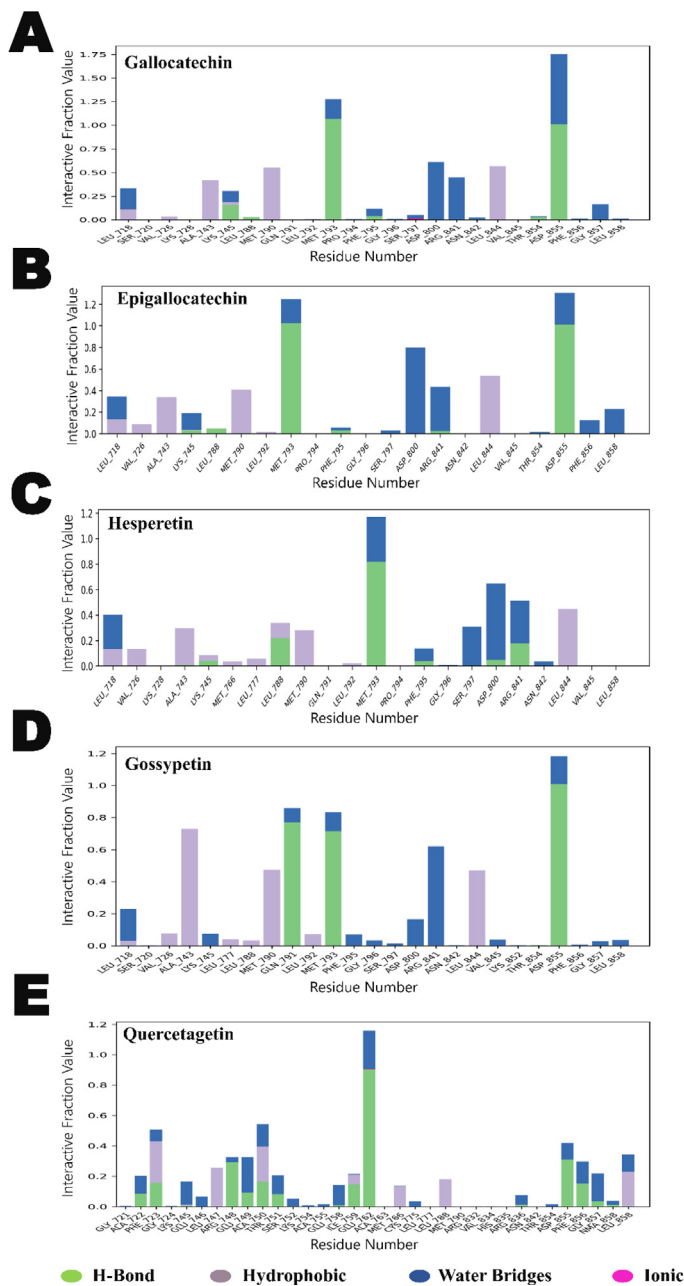


Fig. 10. Simulation interactions diagram of the selected compounds with EGFR in stacked bar chart format.

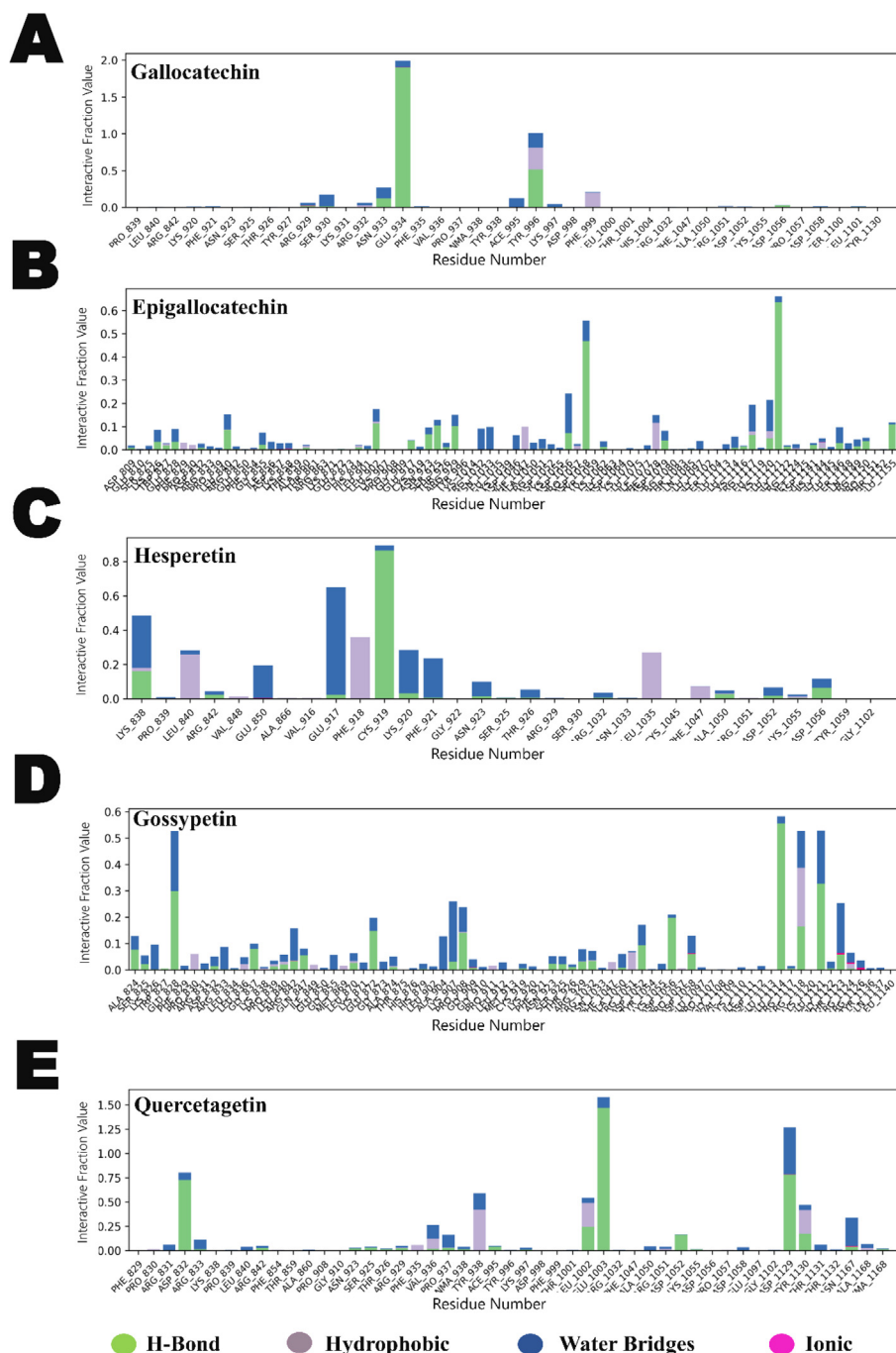


Fig. 11. Simulation interactions diagram of the selected compounds with VEGFR-2 in stacked bar chart format.

and Fig. 11 for VEGFR-2 and the IFV for both controls are presented in [supplementary file \(Fig. S5\)](#). The Y-axis of the stacked bar chart revealed interactive fraction values, while the X-axis pointed out the amino acid residues. In the stacked bar chart, IFVs 1.0 indicate the corresponding residue attached to the ligand for 100 % of the simulation time, and values over 1.0 in the chart suggest that the residue interacts multiple times (over 100 %) with the identical ligand subtype³⁵. The amino acid residues MET_793 and ASP_855 of the EGFR protein in complex with the compounds gallicocatechin and epigallocatechin maintain a stable interaction over 100 % of simulation time through hydrogen bond and water bridge interactions. The IFVs for these residues are 1.75 and 1.28 with gallicocatechin and 1.30 and 1.25 with epigallocatechin. The residues MET_790, ASP_800, and LEU_844 have an

IFV of 0.55, 0.61, and 0.57 with gallicocatechin, which indicates these residues maintain interaction over 50 % of the simulation time. In the complex, residues MET_790 and LEU_844 are attached by a hydrophobic bond, and residue ASP_800 by a water bridge. The residues ASP_800 and LEU_844 in the EGFR-epigallocatechin complex maintain interaction over 80 % and 50 % of the simulation time, respectively, with an interactive fraction value of 0.80 and 0.50 by water bridge and hydrophobic bond. The residue MET_793 of the protein has an interactive fraction value of 1.17, and residues ASP_800 and ARG_841 have a value of 0.65 and 0.51 with the compound hesperetin. It means residue MET_793 maintains interaction over 100 % and residues ASP_800 and ARG_841 maintain over 50 % of the simulation time by hydrogen bond and water-bridge. The EGFR residues

743_ALA, 791_GLN, 793_MET, 841_ARG, and 855_ASP have an interaction value of 0.73, 0.86, 0.84, 0.62, and 1.19, respectively, with the compound gossypetin. Which reveals residue 855_ASP prolonged interaction over 100 % and the rest maintain interaction over 50 % of the simulation time. The EGFR residue 762_GLU maintains interaction with an interaction value of 1.16, and the 723_PHE and 750_ALA residues have IFVs of 0.51 and 0.54 in the complex with quercetagenin. The VEGFR-2 residues GLU_934 and TYP_996 maintained the interaction over 100% of the simulation time with the compound gallicocatechin with IFVs of 1.98 and 1.01 by hydrogen, hydrophobic, and water bridge interactions. The residues ASP_1058 and GLU_1121 of the VEGFR-2 protein interact with epigallocatechin over 50 % of the simulation time by hydrogen bond and water bridge interaction with IFVs of 0.56 and 0.66. In the complex VEGFR-2-hesperetin, residues GLU_917 and CYS_919 interact 60 % and 90 % of the simulation time by hydrogen bond and water bridge, respectively, with an IFV of 0.62 and 0.90. The VEGFR-2 protein's residues GLU_828, GLU_1114, ARG_1118, and GLU_1121 with the compound gossypetin maintain interaction for 50 % of the simulation time with IFVs of 0.52, 0.58, 0.52, and 0.52, respectively. The residues GLU_1003 and ASP_1129 have an IFV of 1.56 and 1.27, which implies these residues maintain interaction over 100 % of the simulation time with the compound quercetagenin by hydrogen, ionic, and water bridges. In addition, the complex residues ASP_832, TYR_938, LEU_1002, and TYR_1130 are attached by hydrogen, hydrophobic, and water bridges with IFVs of 0.77, 0.70, 0.65, and 0.60. According to the interaction analysis, it is apparent that hydrophobic, hydrogen, ionic, and water molecules participating water bridge interactions are the four main interactions that occur between our selected ligands and proteins. Most of the interactions involve active site residues of both proteins, and these residues maintain a constant interaction with the ligands to prolong a stable attachment with them.

3.7.4. Ligand properties

We analyzed various properties of the ligands after the simulation to observe whether they are stably bound to the protein or had undergone any structural changes in their binding state. A number of criteria were analyzed to examine the structural details of the ligands, such as RMSD, molecular surface area (MolSA), radius of gyration (Rg), and polar surface area (PSA). Fig. 12 and Fig. 13 illustrate the outcomes of the ligand properties analysis. The control ligand, Erlotinib, has the highest RMSD value compared to the selected compounds. This compound showed a large fluctuation from the beginning of the simulation to 125 ns at 0.4 Å to 2.5 Å and then showed quiet stability. All the ligands except quercetagenin have an RMSD value lower than 1.0 Å with the EGFR protein, and none of them show any significant fluctuations during the whole simulation period, which is highly satisfactory. The RMSD values of the ligands are presented in Fig. 12 (A and B). From the figure, we can observe that the ligand quercetagenin has a maximum RMSD value of 1.47 Å, and it exhibits some fluctuations until 120 ns, and then it gradually stabilizes for the remainder of the simulation. On the other hand, all of these compounds except control Bevacizumab + Rituximab, when bound to the protein VEGFR-2, have an RMSD of less than 1.5 Å, which is likewise satisfactory. The control has quite a larger value of RMSD than 1.5 Å. The compound gallicocatechin with VEGFR-2 protein has shown some fluctuations until 25 ns, and then it stabilized with a mean RMSD of 1.10 Å until 135 ns. Then the RMSD shows a reduced peak with a mean value of 0.36 Å. The maximum RMSD value for this compound is 1.8 Å. The compound epigallocatechin has an average RMSD of 0.46 Å and a maximum RMSD of 1.19 Å when bound with the VEGFR-2 protein. Compound hesperetin initially exhibits a few oscillations before completely stabilizing with an average RMSD of 0.40 Å throughout the simulation period. Both the compounds gossypetin and quercetagenin show mild fluctuations throughout the simulation

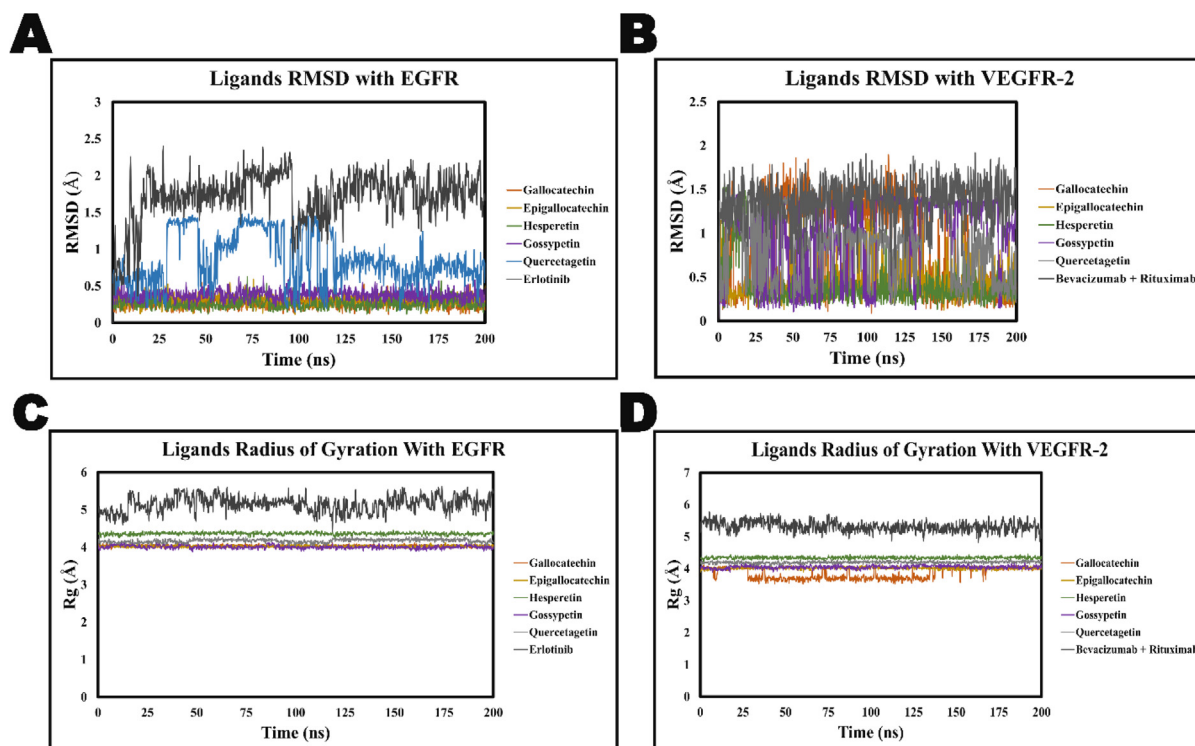


Fig. 12. Time dependent RMSD values (A) with EGFR and (B) with VEGFR-2 and Rg values (C) with EGFR and (D) with VEGFR-2 of the ligands interacting with proteins during the 200 ns.

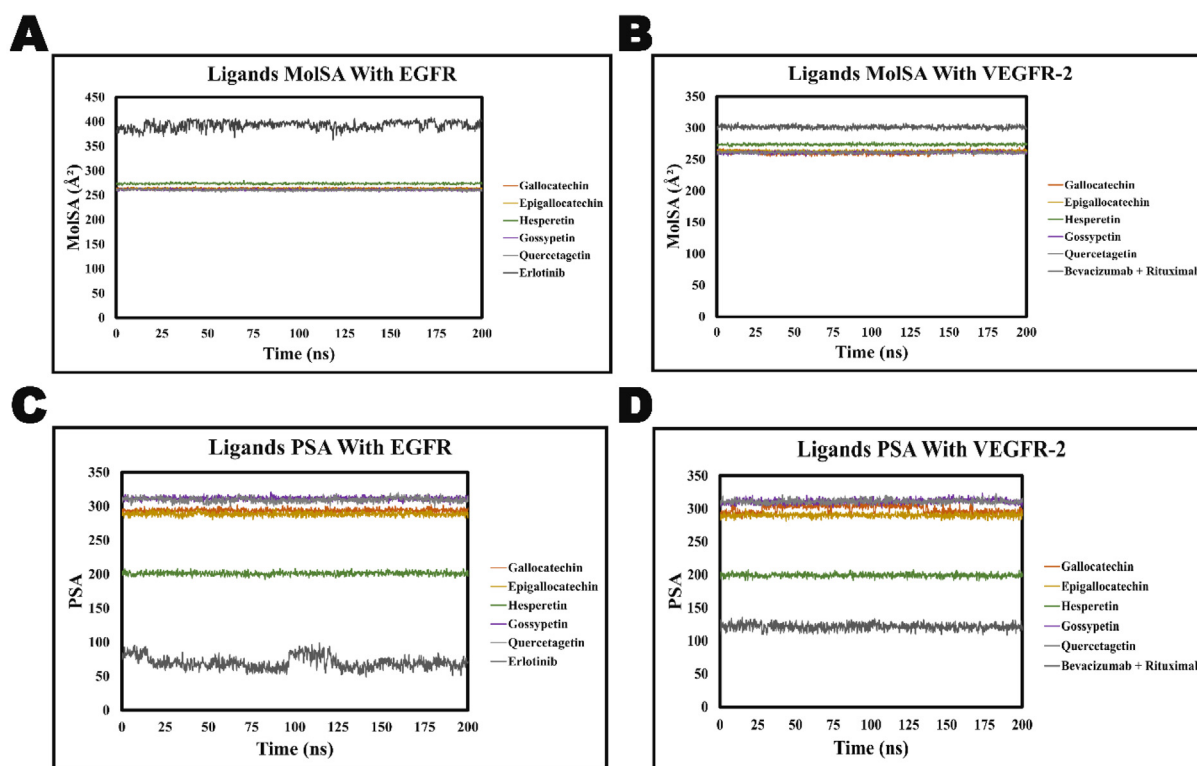


Fig. 13. MolSA (A and B) and PSA values (C and D) of the ligands in complex with proteins during the simulation time frame.

period, with an average RMSD value of 1.10 Å and 0.85 Å. The highest RMSD values for these compounds are 1.44 Å and 1.46 Å. We also examined the stiffness of the compounds with the target proteins during the whole simulation period by analyzing their radius of gyration, as shown in Fig. 12 (C and D). The structural activity of proteins can be reassessed by analyzing the radius of gyration (Rg). It quantifies the distance that a molecule expands from its center of mass during the simulation. It reflects the root-mean-square distance of the atoms from the center of mass of a molecule⁹⁰. A large difference between the lowest and highest points of Rg values suggests compounds are dissociating from the protein, while smaller differences indicate a stronger interaction between the protein and ligands⁹¹.

The lowest and highest Rg values gallicocatechin, epigallocatechin, hesperetin, gossypetin, and quercetagenin in complex with EGFR are 3.96 Å and 4.14 Å, 3.95 Å and 4.10 Å, 4.26 Å and 4.44 Å, 3.92 Å and 4.08 Å, and 4.04 Å and 4.24 Å, respectively. The average Rg values for each of these complexes are 4.02 Å, 4.02 Å, 4.36 Å, 4.00 Å, and 4.17 Å. However, the control for EGFR has an Rg value higher than around 5 Å, with significant fluctuations throughout the period. In contrast, all the compounds have a mean Rg value of 3.85 Å, 4.01 Å, 4.34 Å, 4.05 Å, and 4.20 Å with the VEGFR-2 protein. The lowest and highest Rg values with this protein are 3.6 Å and 4.0 Å for gallicocatechin, 3.92 Å and 4.08 Å for epigallocatechin, 4.24 Å and 4.40 Å for hesperetin, 3.9 Å and 4.1 Å for gossypetin, and 4.08 Å and 4.24 Å for the compound quercetagenin. The control has an Rg value from 5.40 Å to 5.90 Å with this protein. The Rg value differences for all the compounds with the EGFR protein are lower than 0.20 Å. On the other hand, except for gallicocatechin and gossypetin, the rest of the compounds have Rg value differences below 0.20 Å. The Rg value difference for gallicocatechin is 0.40 Å and 0.30 Å for gossypetin. The MolSA is equal to the van der Waals surface area, calculated with a 1.4 Å probe radius. The integrity of the protein–ligand interaction is additionally related to MolSA. If the MolSA value is high, it suggests

an unstable protein–ligand complex, while a low MolSA value indicates a more stable complex^{92,93}. The findings of the MolSA analysis in our study are depicted in Fig. 13 (A and B). All the compounds except hesperetin have MolSA values ranging from 255 Å² to 265 Å² with both the proteins EGFR and VEGFR-2. The hesperetin shows a MolSA value ranging from 270 Å² to 277 Å² with both proteins. These ranges of MolSA values are completely acceptable. In addition, the MolSA in this study remained constant without significant variations throughout the whole simulation period. The control, Erlotinib, has a MolSA value of about 365 Å² to 405 Å², and for the control, Bevacizumab + Rituximab, it varies from around 295 Å² to 305 Å². We also measured the polar surface area (PSA) because it is a descriptor that provides details on the potential polar interactions, permeability, and solubility of a molecule's. Furthermore, the significance of PSA lies in the fact that highly polar molecules have trouble passing across blood brain barrier. The selected compounds displayed satisfactory PSA averages, as illustrated in Fig. 13 (C and D). The average PSA of gallicocatechin, epigallocatechin, hesperetin, gossypetin, and quercetagenin is 293.77 Å², 288.56 Å², 201.25 Å², 311.38 Å², and 309.65 Å², respectively. These values indicate that the selected compounds will not pass through the blood brain barrier (BBB), which we also observed during ADME analysis. However, there are also no major fluctuations observed in the PSA graph throughout the simulation period. The control drug Erlotinib, on the other hand, has a PSA value of 69.82 Å², and the control drug combination Bevacizumab + Rituximab has a PSA value of 121.47 Å². This suggests that they possess a strong ability to cross the blood–brain barrier, leading to substantial damage to the nervous system through heightened toxicity and can influence possible side effects like delirium⁹⁴.

Our ligand properties analysis indicates the ligands are stably attached to the protein, as there are no large fluctuations observed in the ligand RMSD graph. The significant variation in the RMSD graph implies the ligand has shifted away from the protein during simula-

tion. However, in our study, the ligands exhibit uniform behavior throughout the simulation, also indicating the system's complicated stability. In addition, Rg values suggest that the binding of ligand molecules does not significantly alter the structural state of the protein's binding region. The MolSA values further supported the stability of our complexes. All the compounds also exhibited a significant PSA value in the presence of both proteins. Based on our ligand properties analysis, all the selected compounds exhibited ideal RMSD, Rg, MolSA, and PSA values, making them strong prospects for therapeutic development.

3.8. Binding free energy (MM-GBSA) and DFT calculation

MM-GBSA is broadly used and reliable method of predicting the binding free energy immediately following a simulation. It takes into

account protein flexibility, polarizability, entropy, and solvation, which are sometimes not considered in docking methods⁹⁵. The MM-GBSA calculations establish a correlation between the stability of the ligand within the binding pocket and a more negative binding free energy (ΔG_{bind}), which signifies stronger binding and a more pronounced correlation with the inhibitory ability of drugs⁹⁵. The binding energy statistics of MM-GBSA, as illustrated in Fig. 14, are represented in the form of a bar chart diagram. The binding free energy value is presented on the left Y-axis of the graph. On the X-axis, the graph displays the individual energy components.

We observed that the compounds epigallocatechin and galocatechin have the highest average binding energy (ΔG_{bind}) of -54.002 and -52.28 kcal/mol, respectively, in the case of EGFR. The findings indicated that the key energy components contributing to the stability of the simulated epigallocatechin-EGFR and galocatechin-EGFR com-

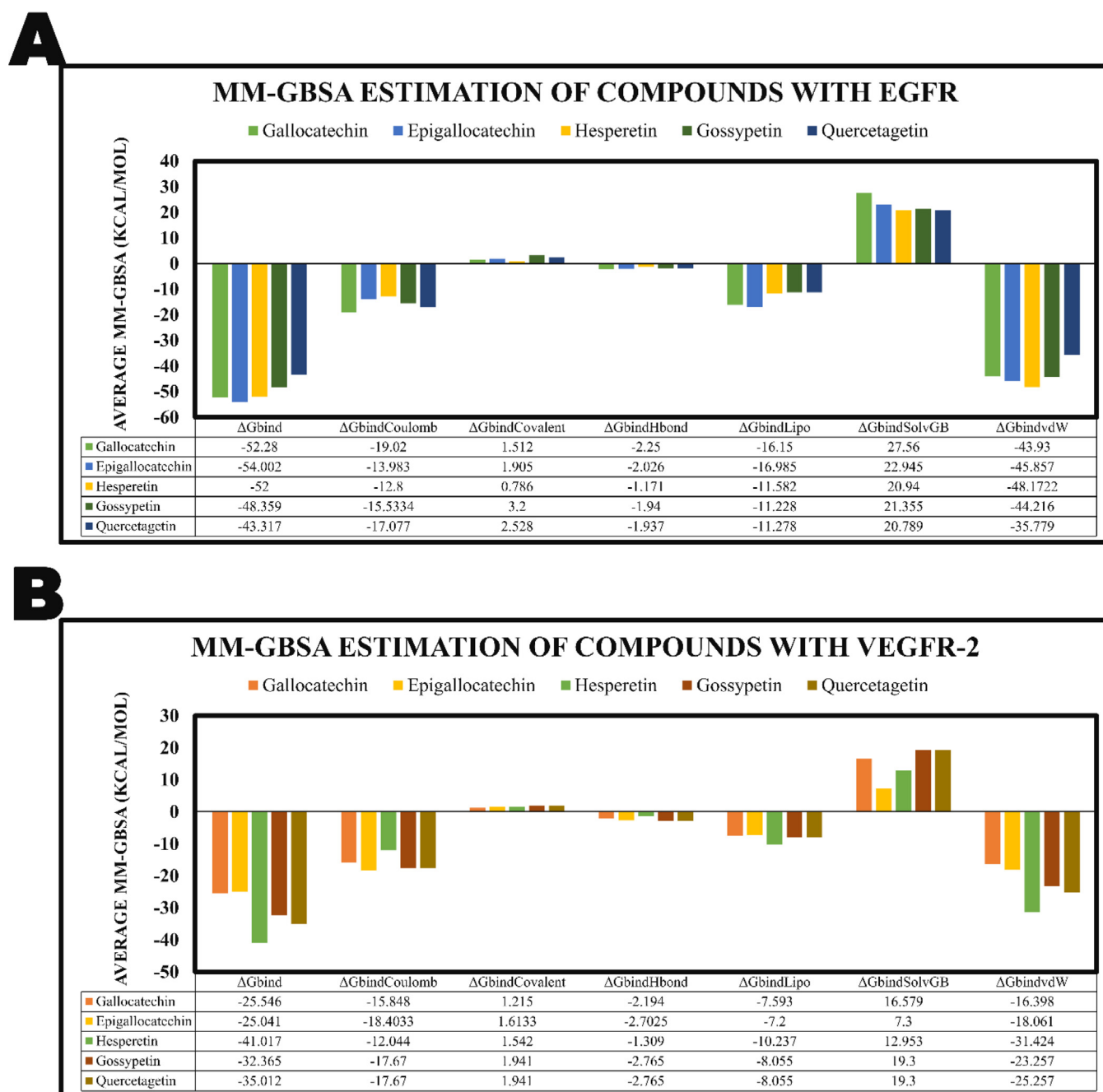


Fig. 14. The outcomes of binding free energy analysis by the MM-GBSA approach for the selected compounds in complex with (A) EGFR and (B) VEGFR-2.

plexes were $\Delta G_{\text{bindVdW}}$, $\Delta G_{\text{bindCoulomb}}$, and $\Delta G_{\text{bindLipo}}$, which had the greatest impact on ΔG_{bind} . However, in contrast, these compounds exhibit the lowest average binding energy (ΔG_{bind}) of -25.41 and -25.546 kcal/mol, respectively, for VEGFR-2. The energy components $G_{\text{bindSolvGb}}$, $G_{\text{bindCovalent}}$, and $G_{\text{bindHbond}}$ were observed to demonstrate unfavorable energy contributions for the epigallocatechin and galliccatechin in both scenarios of forming complexes with EGFR and VEGFR-2. Hesperetin's binding-free energy demonstrated strong consistency with both EGFR and VEGFR-2, as in the hesperetin-EGFR complex, the ΔG_{bind} value was -52 kcal/mol, while in the hesperetin-VEGFR-2 complex, the ΔG_{bind} score was -41.017 kcal/mol. The compound hesperetin is perhaps the most ideal among the five candidates since it has an optimum binding-free energy consistency with both EGFR and VEGFR-2. The ΔG_{bind} values of hesperetin in the hesperetin-EGFR complex and hesperetin-VEGFR-2 complex indicate that hesperetin exhibits strong stability in the binding sites of both proteins and can inhibit both proteins simultaneously. The hesperetin-EGFR complex's optimum binding energy correlated with the higher energy values reported for G_{bindVdW} (-48.17 kcal/mol), $G_{\text{bindCoulomb}}$ (-12.8 kcal/mol), and G_{bindLipo} (-11.58 kcal/mol). Similarly, these energy component's relative contributions to the binding energies in the hesperetin-VEGFR-2 complex remained fairly consistent. In this instance, the computed values of $\Delta G_{\text{bindVdW}}$, $\Delta G_{\text{bindCoulomb}}$, and $\Delta G_{\text{bindLipo}}$ are -31.42 kcal/mol, -12.04 kcal/mol, and -10.2 kcal/mol, respectively.

The calculated overall binding free energy of the gossypetin-EGFR complex is -48.36 kcal/mol, whereas the binding free energy of the gossypetin-VEGFR-2 complex is -32.36 kcal/mol, positioning it as the second most promising candidate. Quercetagenin demonstrates a binding energy of -43.31 kcal/mol, which is slightly lower but comparable to gossypetin when it interacts with EGFR to form a complex. However, it demonstrates a marginally stronger attraction to VEGFR-2 (35.01 kcal/mol) in comparison to gossypetin. Furthermore, in every instance, the van der Waals (VdW) energy seems to make the most significant contribution to the binding energy, followed by coulomb and lipophilic interactions. As per the binding free energy values, the compounds can be ranked in the following order:

Hesperetin > Gossypetin > Quercetagenin > Epigallocatechin > Galliccatechin.

The findings appear to indicate that these drugs have the potential to target the EGFR and VEGFR-2 pathways simultaneously rather than individually, indicating the possibility of a synergistic effect. In addition, we performed the DFT calculation and presented the detailed outcome in the [supplementary file](#) from pages 7 to 10. However, our DFT calculation reveals that the compounds hesperetin, gossypetin, and quercetagenin exhibit a lower band gap and hardness value, along with a higher softness and electrophilicity index, indicating their potential as more prominent inhibitors than galliccatechin and epigallocatechin. Hesperetin, galliccatechin, and epigallocatechin have also shown better outcomes in molecular docking, simulation, and the MM-GBSA approach than controls. So, we conclude that these three compounds may have more chemical potential for achieving therapeutic effects in biological systems by targeting both EGFR and VEGFR-2.

4. Conclusion

In contemporary times, the concept of “one drug, one target” has transformed, giving rise to the strategy of “one drug, multiple targets” in the domain of drug design. With the vulnerability of single-target drugs to drug resistance caused by single nucleotide polymorphisms (SNPs), there is growing interest in the research and development of dual-or multi-target inhibitors. This research performed a comprehensive in silico analysis to identify potential multitargeting compounds extracted from *M. oleifera* L. against NSCLC, specifically targeting the EGFR and VEGFR-2. The EGFR demonstrates resistance to multiple

generations of tyrosine kinase inhibitors (TKIs) due to the acquisition of certain mutations (T790M and C797S). The correlation between EGFR resistance and the overexpression of VEGFR-2 offers a promising avenue for the development of novel anticancer treatments. We identified the prospective lead compounds with the capacity to function as novel dual-EGFR and VEGFR-2 inhibitors, thereby enhancing their anti-cancer efficacy and exerting a faster inhibitory effect on both targets. The virtual screening procedure resulted in an array of 17 compounds that had a particularly excellent binding affinity with both receptor proteins than two renowned control molecules Erlotinib and Bevacizumab + Rituximab. After applying the pre-ADMET filter and ADMET prediction to these 17 compounds, a group of five drug candidates, namely hesperetin, gossypetin, quercetagenin, galliccatechin, and epigallocatechin, were found and considered as promising drug candidates. To verify the accuracy of the outcomes from multi-stage docking protocols, we re-assessed the stability and reliability of the binding poses using MD simulations and MM-GBSA binding free energy calculations. In this in silico study, the drug candidates exhibited potent anti-cancer properties, as evidenced by their reliable stability and strong binding strength. Overall, our DFT computation study suggests among these five compounds hesperetin, gossypetin and quercetagenin can work as promising inhibitors and can potentially act on both EGFR and VEGFR-2 as a single agent. However, our hypothesis suggests that co-administering these three therapeutic candidates in the body could enhance their inhibitory action and create a synergistic effect, thereby preventing NSCLC. We urge rigorous in-vivo and in-vitro testing of these compounds to develop promising novel treatments for NSCLC.

CRedit authorship contribution statement

Md. Masudur Rahman Munna: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Md. Touki Tahamid Tusar:** Writing – review & editing, Writing – original draft, Formal analysis, Visualization, Methodology. **Saima Sajnin Shanta:** Writing – review & editing, Writing – original draft. **Md. Hossain Ahmed:** Writing – review & editing, Visualization, Formal analysis. **Md. Sarafat Ali:** Writing – review & editing, Supervision, Software, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to acknowledge the Department of Biotechnology and Genetic Engineering under the Faculty of Life Sciences, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj-8100, Bangladesh, for giving us an opportunity to do computational work.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jgeb.2024.100406>.

References

- de Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. *Transl Lung Cancer Res.* 2018;7:220–233. <https://doi.org/10.21037/tlcr.2018.05.06>.
- Perlikos F, Harrington KJ, Syrigos KN. Key molecular mechanisms in lung cancer invasion and metastasis: A comprehensive review. *Crit Rev Oncol Hematol.* 2013;87:1–11. <https://doi.org/10.1016/j.critrevonc.2012.12.007>.

3. Mishra N, Kumar Maurya A. Novel drug discovery against breast and lung cancer using pharmacophore based 2H-1-benzopyran-2-one derivative. Published online 2022. <https://doi.org/10.21203/rs.3.rs-1541154/v1>.
4. Zheng R, Zhang S, Zeng H, et al. Cancer incidence and mortality in China, 2016. *J Natl Cancer Cent.* 2022;2:1–9. <https://doi.org/10.1016/j.jncc.2022.02.002>.
5. Chhikara BS, Parang K. Global Cancer Statistics 2022: The trends projection analysis. *Chem Biol Lett.* 2023;10:1–16.
6. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394–424. <https://doi.org/10.3322/caac.21492>.
7. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer.* 2007;7:169–181. <https://doi.org/10.1038/nrc2088>.
8. Metro G, Crinò L. Advances on EGFR mutation for lung cancer. *Transl Lung Cancer Res.* 2012;1:5–13. <https://doi.org/10.3978/j.issn.2218-6751.2011.12.01>.
9. Nie W, Tang L, Zhang H, et al. Structural analysis of the EGFR TK domain and potential implications for EGFR targeted therapy. *Int J Oncol.* 2012;40:1763–1769. <https://doi.org/10.3892/ij.2012.1356>.
10. Huang L, Huang Z, Bai Z, Xie R, Sun L, Lin K. Development and strategies of VEGFR-2/KDR inhibitors. *Future Med Chem.* 2012;4:1839–1852. <https://doi.org/10.4155/fmc.12.121>.
11. Jantus-Lewintre E, Sanmartín E, Sirera R, et al. Combined VEGF-A and VEGFR-2 concentrations in plasma: Diagnostic and prognostic implications in patients with advanced NSCLC. *Lung Cancer.* 2011;74:326–331. <https://doi.org/10.1016/j.lungcan.2011.02.016>.
12. D, Inés Fernández-Ulibarri V, Vilella M, et al. Diacylglycerol Is Required for the Formation of COPI Vesicles in the Golgi-to-ER Transport Pathway. *Mol Biol Cell.* 2007;18:3250–3263. doi:10.1091/mbc.E07.
13. Patel SA, Nilsson MB, Le X, Cascone T, Jain RK, Heymach JV. Molecular mechanisms and future implications of VEGF/VEGFR in cancer therapy. *Clin Cancer Res.* 2023;29:30–39. <https://doi.org/10.1158/1078-0432.CCR-22-1366>.
14. Shibuya M. Vegf-vegfr signals in health and disease. *Biomol Ther.* 2014;22:1–9. <https://doi.org/10.4062/biomolther.2013.113>.
15. Wang Q, Zeng A, Zhu M, Song L. Dual inhibition of EGFR-VEGF: An effective approach to the treatment of advanced non-small cell lung cancer with EGFR mutation (Review). *Int J Oncol.* 2023;62:1–10. <https://doi.org/10.3892/ijo.2023.5474>.
16. Saryedine L, Zibara K, Kassem N, Badran B, El-Zein N. EGF-Induced VEGF exerts a PI3K-Dependent positive feedback on ERK and AKT through VEGFR2 in Hematological in vitro models. *PLoS One.* 2016;11:1–16. <https://doi.org/10.1371/journal.pone.0165876>.
17. El Rhabori S, El Aissouq A, Chtita S, Khalil F. 3D-QSAR, molecular docking and ADMET studies of thioquinazolinone derivatives against breast cancer. *J Indian Chem Soc.* 2022;99. <https://doi.org/10.1016/j.jics.2022.100675> 100675.
18. Daoui O, Nour H, Abchir O, Elkhattabi S, Bakhouch M, Chtita S. A computer-aided drug design approach to explore novel type II inhibitors of c-Met receptor tyrosine kinase for cancer therapy: QSAR, molecular docking, ADMET and molecular dynamics simulations. *J Biomol Struct Dyn.* 2023;41:7768–7785. <https://doi.org/10.1080/07391102.2022.2124456>.
19. Daoui O, Mali SN, Elkhattabi K, Elkhattabi S, Chtita S. Repositioning cannabinoids and terpenes as novel EGFR-TKIs candidates for targeted therapy against cancer: A virtual screening model using CADD and biophysical simulations. *Heliyon.* 2023;9. <https://doi.org/10.1016/j.heliyon.2023.e15545> e15545.
20. Wu YY, Xu YM, Lau ATY. Anti-cancer and medicinal potentials of moringa isothiocyanate. *Molecules.* 2021;26. <https://doi.org/10.3390/molecules26247512>.
21. Mumtaz MZ, Kausar F, Hassan M, Javaid S, Malik A. Anticancer activities of phenolic compounds from Moringa oleifera leaves: In vitro and in silico mechanistic study. *Beni-Suef Univ J Basic Appl Sci.* 2021;10. <https://doi.org/10.1186/s43088-021-00101-2>.
22. Marilyn Susanti Junias, I GustiNgurah Budiana, Apris A. Adu, R. Pasifikus Christa Wijaya, Indra Yohanes Kiling. Bioactive Compounds and Anticancer Activities of Moringa Oleifera of East Nusa Tenggara Origin. *Indian J Forensic Med Toxicol.* 2021;15:3554–3560. doi:10.37506/ijfnt.v15i3.15850.
23. Bhadrasha K, Thakore V, Brahmabhatt J, Upadhyay V, Jain N, Rawal R. Anticancer effect of Moringa oleifera leaves extract against lung cancer cell line via induction of apoptosis. *Adv Cancer Biol - Metastasis.* 2022;6. <https://doi.org/10.1016/j.adcanc.2022.100072> 100072.
24. Abchir O, Daoui O, Nour H, et al. Exploration of Cannabis constituents as potential candidates against diabetes mellitus disease using molecular docking, dynamics simulations and ADMET investigations. *Sci African.* 2023;21. <https://doi.org/10.1016/j.sciaf.2023.e01745>.
25. Nour H, Daoui O, Abchir O, Elkhattabi S, Belaidi S, Chtita S. Combined computational approaches for developing new anti-Alzheimer drug candidates: 3D-QSAR, molecular docking and molecular dynamics studies of liguiritigenin derivatives. *Heliyon.* 2022;8. <https://doi.org/10.1016/j.heliyon.2022.e11991> e11991.
26. Daoui O, Mazoir N, Bakhouch M, et al. 3D-QSAR, ADME-Tox, and molecular docking of semisynthetic triterpene derivatives as antibacterial and insecticide agents. *Struct Chem.* 2022;33:1063–1084. <https://doi.org/10.1007/s11224-022-01912-4>.
27. Chtita S, Larif M, Ghamali M, Bouachrine M, Lakhli T. Quantitative structure-activity relationship studies of dibenzo[a, d]cycloalkenimine derivatives for non-competitive antagonists of N-methyl-d-aspartate based on density functional theory with electronic and topological descriptors. *J Taibah Univ Sci.* 2015;9:143–154. <https://doi.org/10.1016/j.jtusc.2014.10.006>.
28. Berman HM, Westbrook J, Feng Z, et al. The protein data bank. *Nucleic Acid Res.* 2000;28:235–242.
29. Dassault Systèmes D studio modeling environment. BIOVIA. Published online 2024.
30. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis.* 2009;30(SUPPL. 1):162–173. <https://doi.org/10.1002/elps.200900140>.
31. Vivek-Ananth RP, Mohanraj K, Sahoo AK, Samal A. IMPPAT 2.0: An enhanced and expanded phytochemical atlas of Indian medicinal plants. *ACS Omega.* 2023;8:8827–8845. <https://doi.org/10.1021/acsomega.3c00156>.
32. Kim S, Chen J, Cheng T, et al. PubChem 2023 update. *Nucleic Acids Res.* 2023;51: D1373–D1380. <https://doi.org/10.1093/nar/gkac956>.
33. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel. *J Cheminform.* 2011;3:1–14. <https://doi.org/10.1186/1758-2946-3-33>.
34. Bielska E, Lucas X, Czerwoniec A, Kasprzak JM, Kaminska KH, Bujnicki JM. Virtual screening strategies in drug design - methods and applications. *Biotechnologia.* 2011;92:249–264. <https://doi.org/10.5114/bta.2011.46542>.
35. Munna MMR, Islam MA, Shanta SS, Monty MA. Structural, functional, molecular docking analysis of a hypothetical protein from Talaromyces marneffe and its molecular dynamic simulation: An in-silico approach. *J Biomol Struct Dyn.* 2024;1–20. <https://doi.org/10.1080/07391102.2024.2314264>.
36. Markóczy Z, Sárosi V, Kudaba I, et al. Erlotinib as single agent first line treatment in locally advanced or metastatic activating EGFR mutation-positive lung adenocarcinoma (CEETAC): An open-label, non-randomized, multicenter, phase IV clinical trial. *BMC Cancer.* 2018;18:1–7. <https://doi.org/10.1186/s12885-018-4283-z>.
37. Hainsworth JD, Greco FA, Raefsky EL, et al. Rituximab with or without bevacizumab for the treatment of patients with relapsed follicular lymphoma. *Clin Lymphoma Myeloma Leuk.* 2014;14:277–283. <https://doi.org/10.1016/j.clml.2014.02.010>.
38. Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. *J Chem Inf Model.* 2021;61:3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>.
39. Sánchez-Linares I, Pérez-Sánchez H, Cecilia JM, García JM. High-Throughput parallel blind Virtual Screening using BINDSURF. *BMC Bioinf.* 2012;13(SUPPL 1). <https://doi.org/10.1186/1471-2105-13-S14-S13>.
40. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep.* 2017;7:1–13. <https://doi.org/10.1038/srep42717>.
41. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2012;64(SUPPL.):4–17. <https://doi.org/10.1016/j.addr.2012.09.019>.
42. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem.* 2002;45:2615–2623. <https://doi.org/10.1021/jm020017n>.
43. Egan WJ, Merz KM, Baldwin JJ. Prediction of drug absorption using multivariate statistics. *J Med Chem.* 2000;43:3867–3877. <https://doi.org/10.1021/jm000292e>.
44. Daoui O, Elkhattabi S, Chtita S, Elkhallabi R, Zgou H, Benjelloun AT. QSAR, molecular docking and ADMET properties in silico studies of novel 4,5,6,7-tetrahydrobenzo[D]-thiazol-2-Yl derivatives derived from dimedone as potent anti-tumor agents through inhibition of C-Met receptor tyrosine kinase. *Heliyon.* 2021;7. <https://doi.org/10.1016/j.heliyon.2021.e07463> e07463.
45. Chtita S, Aoudite A, Belhassan A, et al. QSAR study of N -substituted oseltamivir derivatives as potent avian influenza virus H5N1 inhibitors using quantum chemical descriptors and statistical methods. *New J Chem.* 2020;44:1747–1760. <https://doi.org/10.1039/c9nj04909f>.
46. Akash S, Aovi FI, Azad MAK, et al. A drug design strategy based on molecular docking and molecular dynamics simulations applied to development of inhibitor against triple-negative breast cancer by Scutellarein derivatives. *PLoS One.* 2023;18:1–23. <https://doi.org/10.1371/journal.pone.0283271>.
47. Pires DEV, Blundell TL, Ascher DB. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem.* 2015;58:4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104>.
48. Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-II: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* 2018;46:W257–W263. <https://doi.org/10.1093/nar/gky318>.
49. Bai SB, Geethavani M, Ramakrishna C. Synthesis characterization and molinspiration analysis, anti-bacterial activity of novel 2,4,6-tri substituted pyrimidines. *J Young Pharm.* 2021;14:174–178. <https://doi.org/10.5530/jyp.2022.14.33>.
50. Roy AS, Savrav MSS, Hossain MS, et al. In silico identification of potential inhibitors with higher potency than bumetanide targeting NKCC1: An important ion co-transporter to treat neurological disorders. *Informatics Med Unlocked.* 2021;27. <https://doi.org/10.1016/j.imu.2021.100777> 100777.
51. Morris GM, Ruth H, Lindstrom W, et al. Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30:2785–2791. <https://doi.org/10.1002/jcc.21256>.
52. Owoloye AJ, Ligali FC, Enejoh OA, et al. Molecular docking, simulation and binding free energy analysis of small molecules as Pf HT1 inhibitors. *PLoS One.* 2022;17:1–18. <https://doi.org/10.1371/journal.pone.0268269>.
53. D. E. Shaw Research, New York, NY 2023. Desmond Molecular Dynamics System.
54. Maestro, Schrödinger, LLC, New York, NY 2024. Schrödinger Release 2023-2:

55. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des.* 2013;27:221–234. <https://doi.org/10.1007/s10822-013-9644-8>.
56. Shivakumar D, Harder E, Damm W, Friesner RA, Sherman W. Improving the prediction of absolute solvation free energies using the next generation opls force field. *J Chem Theory Comput.* 2012;8:2553–2558. <https://doi.org/10.1021/ct300203w>.
57. Kim M, Kim E, Lee S, Kim JS, Lee S. New method for constant- NPT molecular dynamics. *Chem A Eur J.* 2019;123:1689–1699. <https://doi.org/10.1021/acs.jpca.8b09082>.
58. Yazdani B, Sirous H, Enguita FJ, Brogi S, Wing PAC, Fassihi A. Discovery of novel direct small-molecule inhibitors targeting HIF-2 α using structure-based virtual screening, molecular dynamics simulation, and MM-GBSA calculations. *Mol Divers.* 2023. doi:10.1007/s11030-023-10650-6.
59. Lemkul J, Genheden S, Ryde U, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy.pdf. *J Chem Inf Model.* 2015;10:449–461 <http://www.gromacs.org/@api/deki/files/198/=gmx-tutorial.pdf>.
60. Lamb DJ, Lebedenko CG, McCallum PA, Banerjee IA. Molecular Dynamics, MMGBSA, and Docking Studies of Natural Products Conjugated to Tumor-Targeted Peptide for Targeting BRAF V600E and MERTK Receptors. Vol 27. Springer International Publishing; 2023. doi:10.1007/s11030-022-10430-8.
61. Jacobson MP, Pincus DL, Rapp CS, et al. A hierarchical approach to all-atom protein loop prediction. *Proteins Struct Funct Genet.* 2004;55:351–367. <https://doi.org/10.1002/prot.10613>.
62. Ma X, Chang D, Zhao C, et al. Geometric structures and electronic properties of the Bi 2 X 2 Y (X, y = O, S, Se, and Te) ternary compound family: A systematic DFT study. *J Mater Chem C.* 2018;6:13241–13249. <https://doi.org/10.1039/c8tc04587a>.
63. Evecen M, Tanak H. Quantum chemical studies on the molecular structure, spectroscopic and electronic properties of (6-Methoxy-2-oxo-2H-chromen-4-yl)-methyl pyrrolidine-1-carbodithioate. *Mater Sci Pol.* 2016;34:886–904. <https://doi.org/10.1515/msp-2016-0115>.
64. Bouzzine SM, Bouzakroui S, Bouachrine M, Hamidi M. Density functional theory (B3LYP/6-31G*) study of oligothiophenes in their aromatic and polaronic states. *J Mol Struct (Theochem).* 2005;726:271–276. <https://doi.org/10.1016/j.theochem.2005.04.023>.
65. Frisch, M.J., et al. (2009) Gaussian 09, Revision A.02. Gaussian, Inc., Wallingford.
66. Yadav AK, Singh TR. Novel inhibitors design through structural investigations and simulation studies for human PKMTs (SMYD2) involved in cancer. *Mol Simul.* 2021;47:1149–1158. <https://doi.org/10.1080/08927022.2021.1957882>.
67. J K, D C, M R. Molecular Docking, Drug-likeness Studies and ADMET Prediction of Quinoline Imines for Antimalarial Activity. *J Med Chem Drug Des.* 2019;2(1):1–7. doi:10.16966/2578-9589.113.
68. Yadav R, Imran M, Dhamija P, Chaurasia DK, Handu S. Virtual screening, ADMET prediction and dynamics simulation of potential compounds targeting the main protease of SARS-CoV-2. *J Biomol Struct Dyn.* 2021;39:6617–6632. <https://doi.org/10.1080/07391102.2020.1796812>.
69. Grey Arora, T., Thomas, J., Saneh, A., Tohme, P., & Abi-habib, R. I. Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information. *Psychiatry Res.* 2020;14:293.
70. Bultum LE, Tolossa GB, Kim G, Kwon O, Lee D. In silico activity and ADMET profiling of phytochemicals from Ethiopian indigenous aloes using pharmacophore models. *Sci Rep.* 2022;12:1–19. <https://doi.org/10.1038/s41598-022-26446-x>.
71. Durán-Iturbide NA, Díaz-Eufracio BI, Medina-Franco JL. In silico ADME/Tox profiling of natural products: A focus on BIOFACQUIM. *ACS Omega.* 2020;5:16076–16084. <https://doi.org/10.1021/acsomega.0c01581>.
72. Lagorce D, Douguet D, Miteva MA, Villoutreix BO. Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. *Sci Rep.* 2017;7:1–15. <https://doi.org/10.1038/srep46277>.
73. Sabarees G, Velmurugan V, Solomon VR. Molecular docking and molecular dynamics simulations discover curcumin analogs as potential wound healing agents. *Chem Phys Impact.* October 2023;2024. <https://doi.org/10.1016/j.chphi.2023.100441>.
74. Taghour MS, Elkady H, Eldehna WM, et al. Design, Synthesis, Anti-Proliferative Evaluation, Docking, and MD Simulations Studies of New Thiazolidine-2,4-Diones Targeting VEGFR-2 and Apoptosis Pathway. Vol 17; 2022. doi:10.1371/journal.pone.0272362.
75. Riadi Y, Alamri MA, Geesi MH, et al. Synthesis, characterization, biological evaluation and molecular docking of a new quinazolinone-based derivative as a potent dual inhibitor for VEGFR-2 and EGFR tyrosine kinases. *J Biomol Struct Dyn.* 2022;40:6810–6816. <https://doi.org/10.1080/07391102.2021.1890221>.
76. Roskoski R. Small molecule inhibitors targeting the EGFR/ErbB family of protein-tyrosine kinases in human cancers. *Pharmacol Res.* November 2018;2019:395–411. <https://doi.org/10.1016/j.phrs.2018.11.014>.
77. Amelia T, Kartasmita RE, Ohwada T, Tjahjono DH. Structural insight and development of EGFR tyrosine kinase inhibitors. *Molecules.* 2022;27. <https://doi.org/10.3390/molecules27030819>.
78. Martín-Fernández ML, Clarke DT, Roberts SK, Zanetti-Domingues LC, Gervasio FL. Structure and dynamics of the EGF receptor as revealed by experiments and simulations and its relevance to non-small cell lung cancer. *Cells.* 2019;8:1–34. <https://doi.org/10.3390/cells8040316>.
79. Xie L, Zhao Z, Xie L, Bourne PE. Delineation of polypharmacology across the human structural kinome using a functional site interaction fingerprint approach. *J Med Chem.* 2016;59:4326–4341. <https://doi.org/10.1021/acs.jmedchem.5b02041>.
80. Zheng Zhao1, Lei Xie2, 3 and PEB. Structural insights into characterizing binding sites in EGFR kinase mutants. *Physiol Behav.* 2018;176:139–148. doi:10.1021/acs.jcim.8b00458.Structural.
81. Wang X, Bove AM, Simone G, Ma B. Molecular bases of VEGFR-2-mediated physiological function and pathological role. *Front Cell Dev Biol.* 2020;8:1–12. <https://doi.org/10.3389/fcell.2020.599281>.
82. Leppänen VM, Protta AE, Jeltsch M, et al. Structural determinants of growth factor binding and specificity by VEGF receptor 2. *PNAS.* 2010;107:2425–2430. <https://doi.org/10.1073/pnas.0914318107>.
83. Kim DY, Park JA, Kim Y, et al. SALM4 regulates angiogenic functions in endothelial cells through VEGFR2 phosphorylation at Tyr1175. *FASEB J.* 2019;33:9842–9857. <https://doi.org/10.1096/fj.2018025168R>.
84. Agarwal SM, Nandekar P, Saini R. Computational identification of natural product inhibitors against EGFR double mutant (T790M/L858R) by integrating ADMET, machine learning, molecular docking and a dynamics approach. *RSC Adv.* 2022;12:16779–16789. <https://doi.org/10.1039/d2ra00373b>.
85. Sharma VK, Nandekar PP, Sangamwar A, Pérez-Sánchez H, Agarwal SM. Structure guided design and binding analysis of EGFR inhibiting analogues of erlotinib and AEE788 using ensemble docking, molecular dynamics and MM-GBSA. *RSC Adv.* 2016;6:65725–65735. <https://doi.org/10.1039/c6ra08517b>.
86. Srivastava N, Garg P, Srivastava P, Seth PK. A molecular dynamics simulation study of the ACE2 receptor with screened natural inhibitors to identify novel drug candidate against COVID-19. *PeerJ.* 2021;9:1–18. <https://doi.org/10.7717/peerj.11171>.
87. Kumar H, Raj U, Gupta S, Varadwaj PK. In-silico identification of inhibitors against mutated BCR-ABL protein of chronic myeloid leukemia: a virtual screening and molecular dynamics simulation study. *J Biomol Struct Dyn.* 2016;34:2171–2183. <https://doi.org/10.1080/07391102.2015.1110046>.
88. Ahmad I, Shaikh M, Surana S, Ghosh A, Patel H. p38 α MAP kinase inhibitors to overcome EGFR tertiary C797S point mutation associated with osimertinib in non-small cell lung cancer (NSCLC): Emergence of fourth-generation EGFR inhibitor. *J Biomol Struct Dyn.* 2022;40:3046–3059. <https://doi.org/10.1080/07391102.2020.1844801>.
89. Alturki NA, Mashrafi MM, Alzamami A, et al. In-silico screening and molecular dynamics simulation of drug bank experimental compounds against SARS-CoV-2. *Molecules.* 2022;27:1–12. <https://doi.org/10.3390/molecules27144391>.
90. Bhrdwaj A, Abdalla M, Pande A, et al. Structure-based virtual screening, molecular docking, molecular dynamics simulation of EGFR for the Clinical treatment of glioblastoma. *Appl Biochem Biotechnol.* 2023;195:5094–5119. <https://doi.org/10.1007/s12010-023-04430-z>.
91. Akash S, Bibi S, Biswas P, et al. Revolutionizing anti-cancer drug discovery against breast cancer and lung cancer by modification of natural genistein: An advanced computational and drug design approach. *Front Oncol.* 2023;13:1–15. <https://doi.org/10.3389/fonc.2023.1228865>.
92. Patel CN, Goswami D, Jaiswal DG, Parmar RM, Solanki HA, Pandya HA. Pinpointing the potential hits for hindering interaction of SARS-CoV-2 S-protein with ACE2 from the pool of antiviral phytochemicals utilizing molecular docking and molecular dynamics (MD) simulations. *J Mol Graph Model.* February 2020;2021. <https://doi.org/10.1016/j.jmgm.2021.107874>.
93. Beura S, Chetti P. In-silico strategies for probing chloroquine based inhibitors against SARS-CoV-2. *J Biomol Struct Dyn.* 2021;39:3747–3759. <https://doi.org/10.1080/07391102.2020.1772111>.
94. Matsuoka H, Yoshiuchi K, Koyama A, Otsuka M, Nakagawa K. Chemotherapeutic drugs that penetrate the blood – brain barrier affect the development of hyperactive delirium in cancer patients. *Palliat Support Care.* 2015;13:859–864. <https://doi.org/10.1017/S1478951514000765>.
95. Alsagaby SA, Iqbal D, Ahmad I, et al. In silico investigations identified Butyl Xanalterate to competently target CK2 α (CSNK2A1) for therapy of chronic lymphocytic leukemia. *Sci Rep.* 2022;12:1–21. <https://doi.org/10.1038/s41598-022-21546-0>.