

Cannabis sativa phytoligands as potential EGFR inhibitors for lung cancer treatment

Afolabi Owoloye

afolabiowoloye@yahoo.com

Mountain Top University

Precious A. Akinnusi

University of Cincinnati

Samuel O. Olubode

University of Cincinnati

Evans A. Adu

Global Health and Infectious Disease Research Group

Natalia Shakela

Namibia University of Science and Technology

David O. Oyerinde

Federal University Oye Ekiti

Olamide L. Onawola

Federal University Oye Ekiti

Kolapo M. Oyebola

Mountain Top University

Article

Keywords: Lung cancer, EGFR, Cannabis sativa, Grossamide, Quercetin

Posted Date: October 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11024152/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

***Cannabis sativa* phytoligands as potential EGFR inhibitors for lung cancer treatment**

*Afolabi Owoloye^{1,2}, Precious A. Akinnusi³, Samuel O. Olubode³, Evans A. Adu^{4,5}, Natalia Shakela⁶, David O. Oyerinde⁷, Oluwatobi I. Akinbami⁷, Kolapo M. Oyebola^{1,2}

¹Centre for Genomic Research in Biomedicine (CeGRIB), Mountain Top University, Ibafo, Nigeria

²Nigerian Institute of Medical Research, Lagos, Nigeria

⁴Department of Chemistry, University of Cincinnati, USA

⁴Kumasi Centre for Collaborative Research in Tropical Medicine, Global Health and Infectious Disease Research Group, Kumasi, Ghana

⁵Kwame Nkrumah University of Science and Technology, School of Medicine and Dentistry, Department of Molecular Medicine, Kumasi, Ghana

⁶Namibia University of Science and Technology, Department of Biology, Chemistry and Physics

⁷Department of Biochemistry, Federal University, Oye-Ekiti, Nigeria

***Corresponding author: Afolabi Owoloye**

Abstract

Lung cancer remains a major global health burden, and epidermal growth factor receptor (EGFR) overexpression promotes tumour-cell proliferation, making EGFR an important therapeutic target. This study used an *in-silico* approach to evaluate *Cannabis sativa* compounds as inhibitors of the EGFR kinase domain. Among the compounds investigated, Luteolin-7-*O*-glucoside, grossamide, and quercetin were the leading ATP-competitive inhibitors, with binding affinities of -10.878, -10.091, and -10.000 kcal/mol, respectively. A four-feature pharmacophore model was generated from the lead compounds, which also showed favourable predicted pharmacokinetic profiles. During 100-ns molecular dynamics simulations (MDS), the EGFR-quercetin complex exhibited the lowest root mean square deviation of 0.9-1.5 Å, while fluctuation profiles revealed a relatively stable ligand-bound complex compared with the apo-EGFR. The MDS trajectory Molecular Mechanics/Generalized Born Surface Area analysis yielded binding free energies of -75.59 kcal/mol for grossamide, -56.90 kcal/mol for isovitexin, -49.11 kcal/mol for luteolin-7-*O*-glucoside, and -43.27 kcal/mol for quercetin, compared with -47.41 kcal/mol for gefitinib. Electronic profiling revealed favourable orbital overlap and softness for quercetin and flavone glycosides within EGFR, while isovitexin showed superior kinetic stability and balanced solvation, supporting strong binding and active-site entry. These findings identify luteolin-7-*O*-glucoside, grossamide, and quercetin as promising EGFR inhibitors and provide a baseline hypothesis for experimental validation.

Keywords: Lung cancer, EGFR, *Cannabis sativa*, Grossamide, Quercetin

1 Introduction

2 Lung cancer is the leading cause of cancer deaths worldwide¹. About 85% of cases are attributable to
3 smoking, which remains the most common cause of the disease². Both direct and indirect smoking have
4 been linked to equal exposure among men and women³. Histologically, lung cancers are classified into two
5 main subtypes: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). More than 80% of
6 lung cancers are NSCLC, of which 40% are adenocarcinomas, 25% to 30% are squamous cell carcinomas
7 (SCC), and large cell carcinomas (LCC) account for 10% to 15%⁴.

8 Epidermal growth factor receptor (EGFR) is a member of the receptor tyrosine kinase family and a crucial
9 oncogenic driver in the development, progression, and metastasis of lung cancer⁵. Activated by extracellular
10 growth factors, EGFR initiates complex signal transduction pathways that convey vital cellular information
11 to lung epithelial cells, thereby promoting cell proliferation and survival⁶. Consequently, blocking EGFR
12 with specific tyrosine kinase inhibitors (TKIs) or monoclonal antibodies has become a primary therapeutic
13 strategy.

14 Current treatment for advanced NSCLC is limited to small-molecule drugs such as erlotinib and gefitinib,
15 as well as monoclonal antibodies such as cetuximab. These drugs are generally expensive, and their
16 effectiveness is limited by primary or secondary drug resistance, which emerges after prolonged treatment⁷.
17 Although these therapies initially yield high tumour response rates, their long-term clinical utility is
18 severely compromised. These commercial drugs are costly⁸ and patients inevitably develop primary or
19 acquired drug resistance after prolonged treatment⁹. Therefore, there is an urgent need to identify novel,
20 cost-effective compounds that can modulate the EGFR signalling pathway and overcome resistance
21 mechanisms.

22 Cannabis broadly refers to products derived from *Cannabis sativa* L., including cannabinoids, marijuana,
23 and hemp, although the terms cannabis and marijuana are often used interchangeably¹⁰. Cannabis use
24 disorder represents a growing global health burden, contributing to over 2 million disability-adjusted life
25 years (DALYs), with a disproportionate impact on young adults and populations in high-income countries¹¹.
26 Notably, the inhalation of cannabis smoke exposes users to a spectrum of toxic and carcinogenic compounds
27 also found in tobacco smoke, including phenols, nitrosamines, polycyclic aromatic hydrocarbons,
28 acetaldehyde, formaldehyde, ammonia, and acrolein¹².

29 Despite these risks, *C. sativa* remains a rich source of structurally diverse bioactive compounds. Extensive
30 phytochemical studies have characterized both phyto-cannabinoids and non-cannabinoid metabolites with
31 potential pharmacological relevance¹³⁻¹⁷. Among these, Δ^9 -tetrahydrocannabinol (THC) and cannabidiol
32 (CBD) are the principal phytocannabinoids used to define cannabis chemotypes, with THC being
33 psychoactive and CBD non-intoxicating¹⁸. In particular, CBD has attracted considerable attention for its
34 anticancer potential, mediated in part through modulation of the endocannabinoid system (ECS), which
35 influences tumour cell proliferation, apoptosis, pain signalling, and immune regulation¹⁸.

36 Cannabidiol modulates cannabinoid receptors, which have promising antitumor effects on cell growth,
37 programmed cell death, and angiogenesis. It acts as an inverse agonist for GPR3, GPR6, and GPR12. This
38 adds to evidence that CBDs could affect EGFR, RAS/RAF isoforms, MEK1/2, and ERK1/2¹⁹. Here, we

hypothesize that the anticancer activity of *C. sativa*-derived phytoligands may be mediated, at least in part, through inverse agonism or functional suppression of EGFR signalling.

Building on this pharmacological potential, we hypothesize that selected *C. Sativa*-derived phytoligands may function as inverse agonists or negative modulators of EGFR. This study therefore investigated the inhibitory activity of *C. Sativa*-derived phytochemicals against the EGFR. A comprehensive computational framework was employed to characterize ligand-receptor interactions within the EGFR kinase domain and to identify candidate compounds with high inhibitory potential. Specifically, molecular docking, MM/GBSA binding free energy estimation, pharmacophore modelling, molecular dynamics simulations, and structure-based pharmacokinetic profiling were integrated to provide a mechanistic and dynamic understanding of ligand binding and stability.

Results

EGFR-ligand binding affinity

The simulated biological interactions revealed several compounds with potential inhibitory properties against EGFR tyrosine kinase (Table 1). The top 10 compounds are reported herein after being filtered using SP and XP algorithms, and the top 5 candidates were selected for further analyses. Luteolin 7-*O*-glucoside had the highest affinity to the target with a binding score of -10.88 kcal/mol. Grossamide ranked next to Luteolin 7-*O*-glucoside with a binding affinity of -10.09 kcal/mol. Quercetin showed robust binding with an affinity of -10.00 kcal/mol. Quercetin is a natural flavonoid present extensively in vegetables and fruits. In addition, isovitexin and orientin, in complex with EGFR kinase, had binding affinities of -9.78 kcal/mol and -9.64 kcal/mol, respectively. The interaction diagrams of the complexes were analyzed to elucidate the molecular recognition pattern and identify specific interactions between the top-scoring ligands and the kinase domain of EGFR (Figures 1 and 2).

EGFR-ligand interactions

Grossamide interacted with the kinase by forming an H-bond with Thr854. The binding of the pyrrolopyrimidine compound AEE788 to the target kinase was found to extend into the hydrophobic pocket defined by Thr790, Leu788, Lys745, and Met766 via its 4-phenylethylamine moiety, as well as Leu718 and Gly796 below via the 6-phenyl substituent. In addition, the ethylpiperazine group extended towards Asp800 and Glu804 at the edge of the active site. In that order, Luteolin 7-*O*-glucoside, isovitexin, and quercetin established hydrophobic contacts with Lys745, while grossamide made hydrophobic interactions with Met766. A detailed target-ligand interaction is presented in Additional File 1.

Structure-based pharmacokinetics

Modern drug discovery and development depend heavily on drug metabolism and pharmacokinetics, often known as Absorption, Distribution, Metabolism, Elimination, and Toxicity (ADMET). A high-quality drug candidate should exhibit appropriate ADMET characteristics at the therapeutic dose, in addition to sufficient activity against the therapeutic target. Early drug development uses *in silico* ADMET servers to model a range of pharmacokinetic and physicochemical endpoints. Additional File 1, Tables S1-3, presents the predicted physicochemical properties from the SwissADME server. The pharmacokinetic profile of the selected compounds is summarized in Table 4.

Molecular dynamics simulation

Root Mean Square Deviation (RMSD)

The RMSD of the protein backbone (C α) and the RMSD of the ligand heavy atoms were plotted to analyses their behaviour throughout the simulation period. In addition, the average values for each system were calculated. Fluctuations in RMSD relative to the apo-state (unbound state) indicate the stability of the ligand within the protein's binding pocket. In the initial 15 ns of the simulation, the C α of the apo system fluctuated, reaching 3.0 Å at 18 ns, and later stabilized at an average of 3.5 Å until the end of the simulation run (Figure 3A). The EGFR-quercetin complex C α increased steadily for the first 20 ns before stabilizing at an average RMSD of 2.5 Å. The EGFR-orientin and isovitexin complex C α remained the most stable, with an average RMSD of 2.3 and 2.6 Å, respectively.

The ligand RMSD plot provides insights into the impact of ligands on EGFR stability. We observed fewer structural deviations for grossamide (Figure 3B). Grossamide increased erratically for the first 40 ns (from 0 to 1.8 Å), after which the instability was between 0.9 and 1.8 Å. Gefitinib instability was observed from 5 ns, though with little fluctuation. However, in the case of the EGFR-isovitexin, EGFR-luteolin-7-*O*-glucoside and EGFR-quercetin complexes, the binding of ligands to the active pocket of EGFR demonstrated low deviation throughout the 100 ns of the simulation trajectories. The EGFR-quercetin complex displayed the lowest RMSD (between 0.2 and 0.6 Å).

Root Mean Square Fluctuation

We analyzed the root mean square fluctuation (RMSF) of the 5 complexes and the apo-state (Figure 4). The RMSF of the 5 complexes and the apo-state system exhibited similar fluctuations. However, notable distinctions were observed in the luteolin-7-*O*-glucoside peak region. It exhibited two high peaks at residues 297 and 298 (Asn996 and Phe997, respectively). Similarly, gefitinib showed its highest peak at residue 222 (Ser921).

Radius of gyration (Rg) and Solvent Accessible Surface Area (SASA)

The average Rg values for EGRF-quercetin, EGFR-orientin, EGFR-luteolin-7-*O*-glucoside and EGRF-isovitexin were observed to be 3.7 Å, 3.8 Å, 4.6 Å, and 4.8 Å, respectively (Figure 5A). Grossamide exhibited a relatively high and unstable Rg. However, the gyration was still within the range expected for a good receptor-ligand interaction. Solvent Accessible Surface Area (SASA) depicted in Figure 5B shows that grossamide, luteolin-7-*O*-glucoside, and gefitinib had high SASA values, while quercetin had lower SASA values. The differences in SASA among grossamide, luteolin-7-*O*-glucoside, isovitexin, and quercetin suggest distinct structural and functional characteristics.

Further structural and physicochemical properties of the simulated complexes, specifically the time evolution of ligand molecular surface area (MolSA), polar surface area (PSA), intramolecular hydrogen bonds (IntraHB), and aggregate surface metrics (SASA, MolSA, and PSA), are presented in Additional File 1 (Figures S2–S5). Across the 100 ns trajectory, these parameters consistently indicate stable pocket occupancy, conformational rigidity, and favourable burial within the EGFR binding cavity for the *C. sativa* phytoligands relative to the control inhibitor gefitinib.

Protein-ligand interaction

Additional File 1, Figure S6, shows the molecular dynamics trajectories of protein-ligand contacts over 100 ns. Hydrogen bonds, hydrophobic interactions, and water bridges were predominant throughout the

simulation for most residues that form contacts with the ligands. However, an ionic bond was observed in the luteolin-7-*O*-glucoside complex (Asn842 and Asn855).

MM/GBSA binding energy and validation

The thermal MM-GBSA trajectory results for the 6 evaluated ligand complexes (N=20 frames per system) are summarized and ranked below by overall binding affinity (ΔG_{bind}), with the inclusion of Coulomb energy ($\Delta G_{\text{Coulomb}}$), hydrogen-bonding energy (ΔG_{Hbond}), Lipophilic energy (ΔG_{Lipo}), and Van der Waals energy (ΔG_{vdW}), which comprise the total energy change of the complexes. The binding energy calculations showed that grossamide had the highest binding free energy among the tested compounds, with a ΔG_{bind} value of -75.59 ± 33.24 kcal/mol. This interaction was expected to be the most spontaneous. Although Luteolin 7-*O*-glucoside had the highest binding affinity to the EGFR kinase domain, the reaction between the target and grossamide was anticipated to be more spontaneous than that with Luteolin 7-*O*-glucoside due to its lower binding energy -49.11 ± 13.58 kcal/mol).

Grossamide, the top binder and a strong non-polar driver, operates primarily through extensive van der Waals (-50.05 kcal/mol) and hydrophobic interactions (-23.30 kcal/mol). In active binding frames, its affinity exceeds -100 kcal/mol. Transient unbinding/loosening events around frames 800–850 account for the elevated standard deviation. Isovitexin and Orientin, flavone C-glycosides, exhibit balanced profiles driven by cooperative Coulombic attraction (-26 to -29 kcal/mol), stable hydrogen bonding (-2.8 to -3.0 kcal/mol), and moderate van der Waals packing (-38 to -43 kcal/mol). They outperform the reference inhibitor Gefitinib by engaging polar pocket residues while preserving hydrophobic contacts.

Luteolin 7-*O*-glucoside, a flavone O-glucoside, utilizes strong directional hydrogen bonding (-3.39 kcal/mol) and electrostatic attraction (-24.00 kcal/mol), yielding a consistent binding energy (-49.11 kcal/mol) with moderate desolvation costs. Gefitinib: A typical type-I kinase inhibitor profile, predominantly driven by hydrophobic/aromatic stacking (-41.57 kcal/mol vdW, -14.59 kcal/mol Lipophilic) with minimal hydrogen-bond contribution (-0.81 kcal/mol). Quercetin, with the highest stability/lowest fluctuations, demonstrates the most stable trajectory ($SD=\pm 3.56$ kcal/mol across all frames). The planar aglycone core fits rigidly into the active cleft without conformational strain, anchored continuously by directional hydrogen bonds and low desolvation penalties.

A correlation coefficient of 0.89 (p-value = 0.03) was observed between the ChEMBL experimental data for the receptor (EGFR) and the molecular docking score for the same receptor. This demonstrates that the virtual screening scores obtained in this investigation were similar to the empirically determined inhibitory values in ChEMBL (Additional File 1, Figure S7).

DFT Analysis & Electronic Property

Frontier Molecular Orbital (FMO) Characteristics and Electronic Stability

Grossamide exhibits the least negative E_{HOMO} among all tested ligands. This high-lying HOMO indicates a comparably greater intrinsic electron-donating density from its conjugated lignanamide π -system to electrophilic residues or empty orbital sub-regions within the EGFR catalytic pocket. In contrast, the flavone C-glycosides (isovitexin and orientin) possess deeper HOMO levels, reflecting high chemical stability against oxidative degradation.

Quercetin displays the narrowest HOMO–LUMO energy band gap (ΔE), followed closely by orientin and luteolin 7-*O*-glucoside. A smaller energy gap reflects high intramolecular polarizability and facilitates facile electronic transitions upon target engagement, allowing rapid electronic adaptation to the receptor's electrostatic field.

Chemical Hardness, Softness, and Pocket Complementarity

In accordance with Pearson's Hard and Soft Acids and Bases (HSAB) principle, biological receptor clefts often favour interactions with soft, polarisable ligands capable of extensive dispersion and charge-transfer interactions.

Quercetin and the flavone glycosides exhibit relatively higher chemical softness and lower hardness compared to isovitexin and gefitinib. This increased softness promotes favourable orbital overlap, cooperative π - π stacking, and hydrogen bonding with key aromatic and polar residues in the EGFR hinge and gatekeeper regions. Isovitexin has the largest band gap and highest chemical hardness, indicating greater kinetic stability and resistance to electronic perturbation, which correlates with lower conformational deformation during dynamic simulations.

Global Electrophilicity and Solvation Tendencies

Orientin and luteolin 7-*O*-glucoside show the highest global electrophilicity indices, highlighting their tendency to act as electron acceptors when engaging nucleophilic amine and hydroxyl groups lining the ATP-binding channel.

Gefitinib exhibited the most favourable calculated aqueous solvation energy (ΔG_{solv}), followed by the polyhydroxylated C-glycosides (orientin and isovitexin). The balanced solvation profiles of the natural flavonoids reflect sufficient aqueous solubility to reach the target site, while retaining enough non-polar character to favour desolvation and entry into the hydrophobic kinase core.

***In Silico* Mutagenesis**

When Aspartate mutates to glutamate at positions 800 and 855, the local pocket remains largely unchanged. Glutamate retains the original negative charge but has a slightly longer side chain, so ligands simply adjust their angles to accommodate it. Isovitexin binds both Glu800 and Glu855, while grossamide and luteolin-7-*O*-glycoside bind Glu800. The control, gefitinib, similarly adapts to Glu855.

Mutating aspartate to asparagine removes the negative charge and introduces a neutral amide group. This amide group is a good hydrogen-bond donor and acceptor; your oxygen-rich compounds readily find new ways to anchor. Luteolin-7-*O*-glycoside and orientin adapt by binding to Asn800 and Asn855. Grossamide maintains its hold at position 800 by binding Asn800, while quercetin and gefitinib focus on Asn855.

Discussion

The epidermal growth factor receptor (EGFR), frequently regarded as a receptor tyrosine kinase (RTK)⁹. It binds EGF-family ligands and activates downstream signalling cascades that translate extracellular signals into appropriate cellular responses²⁰. The receptors play a significant role in cell proliferation, differentiation, and migration⁶. Overactivity of this receptor plays a central role in lung cancer and some other types of cancer. EGFR inhibitors are established treatments for patients with advanced NSCLC²¹. To

1 this end, compounds of *Cannabis sativa* were assessed to elucidate their potential inhibitory effects on
2 EGFR kinase.

3 The EGFR docking simulation showed good binding interactions with the *C. sativa* phyto-compounds.
4 Isovitexin and orientin are both flavones, with the former being a C-glycosyl compound containing apigenin
5 and the latter, an 8C glucoside of luteolin. Intriguingly, all five top-ranking compounds are flavonoids. This
6 finding further contributes to understanding flavonoid-specific bioactivities. Meanwhile, all five selected
7 compounds bound more strongly to the EGFR tyrosine kinase than the FDA-approved gefitinib. The
8 reported ligands were docked into the ATP-binding site; they must therefore compete with ATP for binding.
9 Owing to their impressive binding affinities, they are proposed as potential ATP-competitive EGFR kinase
10 inhibitors.

11 In medicinal chemistry and drug discovery, it is usually of interest to identify ligand-receptor interactions,
12 the functional groups involved, and the amino acid residues involved. In a process called lead optimization,
13 the participating atoms can be altered to optimize the protein-ligand interaction between the target and the
14 drug candidate²². Luteolin 7-*O*-glucoside and quercetin interacted with the Met793 residue in the active
15 pocket of EGFR. Met793 is located in the hinge region of the kinase, which connects the N and C lobes²³.
16 Luteolin 7-*O*-glucoside and quercetin were observed to form double H-bond interactions with the hinge.
17 However, Gefitinib also interacted with Met793 via a single hydrogen bond (H-bond). This is consistent
18 with an earlier study on the binding orientation of gefitinib with the target receptor, in which gefitinib
19 formed only one hydrogen bond with Met793 at the hinge^{24,25}.

20 Studies suggest that quercetin has therapeutic potential for treating several diseases, including
21 cardiovascular and neurodegenerative diseases²⁶⁻²⁸. Its anticancer activities have also been documented,
22 particularly in inhibiting the growth, proliferation, migration, invasion, and metastasis of lung cancer²⁹.
23 Specifically, quercetin has been reported to inhibit Snail-dependent Akt activation expression pathways,
24 suppress Src and Aurora B kinase expression, and inhibit the Fn14/NF-kappaB (NF-kB) signalling
25 pathway³⁰. This study suggests that, in addition to the above, EGFR tyrosine kinase inhibition may be one
26 mechanism by which quercetin exerts its activity against lung cancer. Orientin has good binding affinity
27 for the receptor. Several studies have demonstrated the anticancer activities of orientin in bladder,
28 hepatocellular, and colorectal carcinoma³¹⁻³³.

29 A pharmacophore is the set of steric and electronic signatures that further contribute to the interactions of
30 ligands with a bioreceptor³⁴. The complex-based pharmacophore hypothesis draws advantages from both
31 ligand-based and receptor-based approaches³⁵. However, it is complex-specific and may sometimes miss
32 other potential interaction sites within the protein active site because it focuses only on the key steric and
33 electrostatic signatures of the interaction³⁴. To identify the potential pharmacophoric features essential for
34 interaction with the EGFR tyrosine kinase, the *Phase* module of the Schrödinger suite was used to screen
35 the lead complexes and predict molecular signatures that actively contribute to ligand-target interactions³⁶.
36 The Phase module in Schrodinger's suite contains six built-in pharmacophoric features, namely H-bond
37 receptor (A), H-bond donor (D), hydrophobic group (H), positively charged group (P), negatively charged
38 group (N), and aromatic ring (R)³⁷.

39 Lipophilicity was computed to understand the absorption of lead ligands in living systems. This helps
40 predict how rapidly and effectively drugs penetrate human cells. The n-octanol-to-water partition

coefficient is used to compute lipophilicity³⁸. Grossamide, a natural product found in plants, has been reported to have a high log P value, as predicted. The n-octanol-water partition coefficient of this compound is higher than that of the standard drug. This implies that grossamide is expected to have greater solubility in nonpolar solvents. The high log P value suggests that grossamide is likely to partition preferentially into non-polar environments. Although grossamide was predicted to have low bioavailability, this could be optimized using nanoparticles³⁹. A drug candidate is anticipated to exhibit ideal levels of water solubility and lipophilicity⁴⁰. A substance introduced into the body system needs to dissolve in water to aid gastrointestinal absorption. Poor lipophilicity could reduce absorption. Moreover, during distribution, drugs are circulated through a hydrophilic environment.

Blood-brain barrier (BBB) permeation is particularly relevant for drugs whose site of action is the BBB, given their high Log P values. Drugs not intended to target the central nervous system can still permeate the BBB and cause adverse drug reactions in the brain⁴¹. It was expected that grossamide would permeate the BBB. However, none of the compounds permeated these endothelial cells. The inability to cross the BBB may be due to several factors. Brain penetration involves a complex interplay of critical factors, including lipophilicity, metabolism, receptor affinity, molecular size and shape, ionization potential, and specific binding to BBB efflux⁴².

P-glycoprotein is an ATP-binding cassette (ABC) protein that expels molecules from the cell, thereby reducing drug action by preventing bioaccumulation. It is primarily found in barrier organs (such as the blood-brain barrier) and excretory organs (liver and kidneys)⁴³. Quercetin, isovitexin, orientin, and the standard drug are not P-glycoprotein substrates; therefore, these compounds are likely to accumulate at considerable concentrations within the cell. By contrast, grossamide and luteolin 7-O-glucoside are predicted to be P-glycoprotein substrates. These two compounds exhibit rapid renal clearance and limit BBB permeation.

In *in silico* drug discovery, the structural stability of a protein-ligand complex during MD simulations serves as a crucial indicator of potential in vitro binding affinity. Root Mean Square Deviation (RMSD) provides insights into the dynamic behaviour and stability of the complexes over the simulation period and measures the average atomic displacement relative to a reference frame⁴⁴. The RMSD of the C α atoms indicates complex stability, suggesting that ligand binding does not perturb the conformation during the simulation. The RMSD of the protein backbone across all systems was small and comparable. Low-intensity peaks suggest that phytoligand binding does not affect the stability of the receptor's structural region. Protein secondary structure components, such as alpha helices and beta strands, exhibit greater rigidity and hence less fluctuation than loop and unstructured regions during MD simulations.

The Root Mean Square Fluctuation (RMSF) analysis of protein secondary structure components during molecular dynamics simulations revealed consistent overall dynamics across five complexes and the apo state, suggesting that the binding events do not induce major conformational changes. Notable distinctions in the peak regions for quercetin, isovitexin, and orientin indicate ligand-specific effects on protein dynamics. The high peak values at residue 43 for quercetin and isovitexin suggest a shared interaction or stabilization site, whereas the peak at residue 155 for orientin indicates a distinct binding site. These findings pinpoint specific residues crucial for ligand binding, which may influence protein stability and function.

The radius of gyration (Rg) is a parameter that provides insights into protein tertiary structure and can be used to evaluate its stability within a biological system⁴⁵. All ligands, except grossamide, exhibited a lower radius of gyration within the receptor binding pocket, indicating a relatively stable conformation⁴⁶. Solvent Accessible Surface Area (SASA) analyses showed that ligand binding does not affect the overall folding of the protein⁴⁷. High SASA values for grossamide, Luteolin-7-*O*-glucoside, and isovitexin suggest that these molecules adopt extended conformations with larger solvent-exposed surface areas, potentially indicating greater flexibility, hydrophilicity, and accessibility for interactions. In contrast, the lower SASA values for quercetin suggest a more compact, possibly hydrophobic structure, which could affect its solvent interactions and binding affinity. This implies that quercetin could exhibit more specific binding and a better fit within a specific binding pocket of EGFR.

In protein-ligand interactions, H-bonds, hydrophobic interactions, ionic interactions, and water bridges collectively play pivotal roles in determining the stability and specificity of the binding complex. H-bonds were observed in certain amino acid residues (Ser720, Met793, Cys797, and Asp855) across all the complexes, contributing to binding specificity and strength. Hydrophobic interactions drive the burial of hydrophobic moieties within the binding pocket; these interactions were observed among hydrophobic amino acid residues (Leu718, Phe723, Val726, Ala743, Met766, Leu844, and Phe997). Ionic interactions add electrostatic contributions to the stability of the complexes, and water bridges, which were also prominent in all the complexes, mediate interactions and facilitate conformational changes^{48,49}.

Evaluating the thermodynamic components of binding free energy offers critical insight into the intermolecular forces and structural mechanisms that govern EGFR kinase inhibition by *C. sativa* secondary metabolites. The Molecular Mechanics with Generalized Born Surface Area (MM/GBSA) molecular modelling protocol was used to estimate the relative binding free energy of protein-ligand complexes⁵⁰. The binding free energy obtained from MM/GBSA calculations is based on an MD simulation trajectory and often predicts the stability of protein-ligand interactions⁵¹. Generally, more negative ΔG_{bind} values indicate greater spontaneity of the protein-ligand complex and, theoretically, more interactions.

The thermodynamics observed for grossamide highlights the pivotal roles of non-polar surface burial and steric complementarity within the ATP-binding pocket. In typical tyrosine kinase inhibitors, extensive hydrophobic contacts within deep hydrophobic sub-pockets primarily drive high-affinity binding^{52,53}. For grossamide, dominant van der Waals and lipophilic contributions indicate tight steric entrapment of its lignanamide core within the catalytic cleft.

In contrast, isovitexin, orientin, and luteolin 7-*O*-glucoside exhibit an alternative, polar-driven binding mechanism that distinguishes them from gefitinib. While gefitinib relies predominantly on hydrophobic interactions with minimal directional hydrogen bonding, the flavonoid glycosides establish extensive electrostatic networks and direct hydrogen-bonding anchors. The glucosyl moieties extend towards polar residues and the solvent-exposed pocket edge, bridging key hinge-region interactions without incurring prohibitive polar desolvation penalties⁵⁴. This balanced energetic profile suggests that glycosylation enhances target retention by stabilizing the ligand across multiple sub-regions of the kinase domain.

Quercetin exhibits a rigid-core binding mode. Its scaffold experiences negligible conformational strain and fits compactly into the binding pocket, yielding a highly consistent trajectory. The persistent hydrogen-bonding network, coupled with minimal desolvation penalties, indicates a thermodynamically efficient interaction in which favourable enthalpy readily compensates for the smaller contact surface area⁵⁵. Overall,

these thermodynamic profiles reveal distinct yet complementary binding modes across *C. sativa* phytochemicals. While grossamide relies on extensive hydrophobic packing comparable to that of potent synthetic scaffolds, the glycosylated flavonoids achieve strong affinity through cooperative polar networks. These mechanisms highlight the potential of diverse cannabis-derived chemical scaffolds as lead structures for developing targeted EGFR inhibitors for non-small cell lung cancer.

The alignment between computational predictions and empirical observations was assessed, yielding a high correlation coefficient (R^2). The empirical EGFR inhibition assessment and the computational model score showed that the ligands' binding affinity in this study was comparable to the experimentally determined inhibitory assay in the EGFR ChEMBL dataset.

Density Functional Theory (DFT) calculations provide crucial quantum-mechanical insights into the electronic structure, charge-transfer capacity, and chemical reactivity that govern interactions between *Cannabis sativa* secondary metabolites and the EGFR kinase domain.

The frontier molecular orbital characteristics highlight distinct electronic mechanisms across the lignanamide and flavonoid classes. Grossamide features a notably high-lying highest occupied molecular orbital (HOMO), reflecting a strong propensity to donate electron density from its conjugated π -electron network to electrophilic amino acid side chains and orbital sub-regions within the catalytic cleft⁵⁶. This elevated electron-donating capacity correlates with extensive non-polar and dispersion-driven interactions observed in dynamic simulations. Conversely, the flavone C-glycosides, isovitexin and orientin, exhibit deeper HOMO levels, which confer chemical stability against oxidative degradation while preserving specific, localised charge-transfer capabilities⁵⁷.

The energy band gap (ΔE) and global chemical hardness (η) are direct indicators of molecular polarisability and kinetic reactivity. Quercetin, followed closely by orientin and luteolin 7-*O*-glucoside, exhibits a narrow energy band gap and higher chemical softness (σ). According to the Hard and Soft Acids and Bases (HSAB) principle, softer, more polarizable molecules undergo facile electronic reorganisation in an external electrostatic field, facilitating favourable orbital overlap with polar and aromatic residues lining the kinase hinge region⁵⁸. In contrast, isovitexin displays greater chemical hardness, reflecting high kinetic stability and resistance to electronic deformation.

Furthermore, the high electrophilicity indices (ω) observed for orientin and luteolin 7-*O*-glucoside suggest their capacity to stabilize electron density acquired during hydrogen-bonding interactions with nucleophilic residues within the ATP-binding channel. Alongside their favourable aqueous solvation energies, these electronic properties indicate that *C. sativa*-derived flavonoids and lignanamides have well-balanced electronic and polarisability profiles, making them promising molecular templates for targeted EGFR inhibition in non-small cell lung cancer.

The hit compounds' ability to maintain hydrogen bonds with mutated EGFR residues supports the findings reported in "Cannabis Phytoligands as Potential EGFR Inhibitors for Lung Cancer". This persistence indicates that the ligands have flexible binding modes that adapt to active-site changes, a key trait for overcoming drug resistance. This flexibility means that if a tumour develops these specific mutations to evade treatment, the compounds could remain active. They avoid relying on a rigid connection to the original aspartate, allowing them to adapt and maintain a stable grip on the protein.

Conclusions

This study demonstrated the potential of *C. sativa* compounds, specifically luteolin 7-*O*-glucoside, grossamide, and quercetin, as ATP-competitive inhibitors of the EGFR kinase domain for lung cancer treatment. *In silico* analyses showed strong binding affinities and notable binding free energies for these compounds, surpassing those of the FDA-approved gefitinib. However, further validation through in vitro investigations is warranted. This research opens avenues for novel therapeutic strategies that explore *C. sativa* compounds to combat lung cancer by targeting EGFR overactivity.

Methods

Protein targets and ligand library

The 2D structures of the phytoligands and the 3D structure of the target receptor were obtained from compound and protein structure repositories, respectively. Specifically, the 3D structure of EGFR (PDB ID: 3POZ) was retrieved from the Protein Data Bank (PDB; <https://www.rcsb.org/>). Similarly, all available *C. sativa* ligands were retrieved in structure data file (SDF) format from the PubChem repository⁵⁹.

Protein preparation

The 3D structure of EGFR, obtained from the PDB, was imported into the Schrödinger Maestro suite (version 12.9) for preparation. The Protein Preparation Wizard module was used to perform this procedure⁶⁰. Initially, the receptor was preprocessed to identify structural deficiencies in the polypeptide chain. Subsequently, water and all other bound non-standard ligands were removed. Overall, the preparation involved correcting charges, adding missing hydrogen bonds, and assigning bond orders. Missing loops were filled, and missing side chains were added. Thereafter, the structure was processed using ProtAssign to automatically optimize hydroxyl, Asn, Gln, and His states. Finally, a backbone-restrained minimization was performed on the optimized structure. A final visual and quantitative check and correction for residual clashes and rotamers of the binding pocket was performed in Coot

Ligand Preparation

The retrieved SDFs of the ligands were imported into the Maestro Schrödinger suite (version 12.9) for preparation. Using Epik, the LigPrep tool of the Maestro Schrödinger suite, we prepared the compounds at physiological pH. The ligands were subsequently converted into their corresponding 3D structures.

Grid generation

The grid generation module in the Maestro Schrödinger suite (version 12.9) was used to map the active site of the receptor's kinase domain. A Glide docking grid box (dimensions of 25.9 each for x, y and z) was automatically generated based on the position and binding of the co-crystallized ligand in 3POZ. A grid box essentially encompasses the region where ligands interact with bioreceptors. This was done to implement a site-directed molecular docking protocol to guide the automated placement of ligands within the active region of the EGFR tyrosine kinase.

Structure-based virtual screening

A molecular docking approach was employed to screen the prepared phytoligands of *C. sativa* for their binding interactions with EGFR tyrosine kinase. The compounds were docked into the prepared target grid to estimate their binding affinities. The molecular docking protocol was executed at two levels of precision⁶¹. Initially, the Standard Precision (SP) algorithm was used to screen the prepared compounds. The SP protocol allows for lower torsional refinement and greater flexibility in ligand sampling.

Subsequently, the top 50% of the ranking returned by the SP docking protocol was selected and rescored using the Extra Precision (XP) algorithm. The XP protocol uses a more rigorous approach and includes more extensive torsional refinement. The XP results were exported, and the specific intermolecular interactions contributing to ligand binding to the target were investigated.

Pharmacophore hypothesis

A complex-based pharmacophore hypothesis was designed from complexes formed by interactions between the lead compounds and the EGFR kinase domain, using the Phase module in the Maestro Schrödinger suite. This was used to map out the combination of molecular signatures that contributed to the binding of the top-scoring ligands to EGFR tyrosine kinase. The hypothesis set a maximum feature-feature distance of 2.0 Å, and a receptor-based excluded-volume shell was created with radii equivalent to the Van der Waals radii of receptor atoms.

Structure-based pharmacokinetic and pharmacodynamic profiling

Structure-based pharmacokinetic and pharmacodynamic profiling was carried out on the lead compounds to understand their pharmacokinetic and pharmacodynamic behaviour using SwissADME and the ProTox-II server, respectively^{38,62}.

Molecular dynamics simulation

A comprehensive analysis was conducted using a 100-nanosecond molecular dynamics simulation of pivotal protein-ligand complexes with the Desmond tool in the Maestro Schrödinger Suite. This simulation aimed to replicate the dynamic intricacies of biological environments, including the influence of water molecules and lipid membranes. Protein-ligand system preparation used the System Builder tool in the Maestro Schrödinger Suite (version 12.9) in a Linux environment. Utilizing the OPLS4 force field, the transferable intermolecular interaction potential three-point (TIP3P) solvent model was incorporated, adopting an orthorhombic box configuration. The system was kept neutral, counterions were added, and NaCl (0.15 M) was included to replicate physiological conditions.

The isothermal-isobaric (NpT) ensemble was chosen for the simulation, implementing the Martyna–Tuckerman–Klein Barostat at a temperature of 300° Kelvin and atmospheric pressure⁶³. Before the simulation began, the model was relaxed, and trajectories were saved at 100 ps intervals for subsequent analysis. The stability assessment involved examining the Root Mean Square Deviation (RMSD) of both the ligand and the protein over time. Following this, the Root Mean Square Fluctuation (RMSF) was analyzed. The protein-ligand interactions were also examined. This provides valuable insights into the dynamic behaviour and stability of the systems.

Binding Free Energy Calculations (Thermal MM-GBSA)

Molecular Mechanics Generalised Born Surface Area (MM-GBSA) calculations were performed to assess the thermodynamic stability and binding free energy of the ligand–protein complexes using the thermal_mmgbsa.py script within the Schrödinger Suite’s Prime module⁵⁰. Calculations were carried out on the post-equilibration Molecular Dynamics (MD) simulation trajectories. A total of 20 equidistant frames were sampled at 50-frame intervals (from frame 0 to frame 1000) using the OPLS4 force field and the

VSGB 2.0 implicit solvation model. The total binding free energy (ΔG_{bind}) was determined using the following equation:

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

The individual thermodynamic components were decomposed into Coulombic/electrostatic interactions ($\Delta G_{\text{Coulomb}}$), van der Waals contacts (ΔG_{vdW}), lipophilic/hydrophobic contributions (ΔG_{Lipo}), hydrogen-bonding interactions (ΔG_{Hbond}), and **Non-Strained Binding Affinity** ($\Delta G_{\text{bind(NS)}}$). The resulting values were reported as the mean $\pm$ standard deviation across the sampled trajectory frames.

Van der Waals Interaction Energy (ΔG_{vdW}) was calculated using the standard Lennard-Jones 12-6 potential function representing dispersion attractions and steric repulsions:

$$\Delta G_{\text{vdW}} = E_{\text{vdW}}^{\text{complex}} - (E_{\text{vdW}}^{\text{receptor}} + E_{\text{vdW}}^{\text{ligand}})$$

$$E_{\text{vdW}} = \sum_{i < j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)$$

Where r_{ij} is the distance between non-bonded atoms i and j , and A_{ij} and B_{ij} are force field Lennard-Jones parameters.

Electrostatic / Coulombic Energy ($\Delta G_{\text{Coulomb}}$) was calculated using Coulomb's law for all interatomic partial charge interactions:

$$\Delta G_{\text{Coulomb}} = E_{\text{Coulomb}}^{\text{complex}} - (E_{\text{Coulomb}}^{\text{receptor}} + E_{\text{Coulomb}}^{\text{ligand}})$$

$$E_{\text{Coulomb}} = \sum_{i < j} \frac{q_i q_j}{4\pi \epsilon_0 \epsilon_r r_{ij}}$$

Where q_i, q_j are atomic partial charges, ϵ_0 is the vacuum permittivity, and ϵ_r is the relative dielectric constant (typically set to 1 in molecular mechanics gas-phase calculations).

Lipophilic / Hydrophobic Interaction Energy (ΔG_{Lipo}): In Prime (VSGB model), non-polar interactions and surface hydrophobic burial are modeled as an empirical lipophilic free energy term based on atomic surface area exposure:

$$\Delta G_{\text{Lipo}} = G_{\text{Lipo}}^{\text{complex}} - (G_{\text{Lipo}}^{\text{receptor}} + G_{\text{Lipo}}^{\text{ligand}})$$

$$G_{\text{Lipo}} = \sum_i \sigma_i \cdot \text{SASA}_i$$

Where σ_i is the lipophilic/hydrophobic surface tension coefficient for atom i , and SASA_i is its Solvent Accessible Surface Area.

Hydrogen Bonding Energy (ΔG_{Hbond}), an empirical directional potential evaluating the geometry and distance of donor–acceptor hydrogen bonds:

$$\Delta G_{\text{Hbond}} = E_{\text{Hbond}}^{\text{complex}} - (E_{\text{Hbond}}^{\text{receptor}} + E_{\text{Hbond}}^{\text{ligand}})$$

$$E_{\text{Hbond}} = \sum_{\text{H-bonds}} f(r_{\text{DA}}) \cdot g(\theta_{\text{DHA}})$$

Where r_{DA} is donor-acceptor distance, θ_{DHA} is donor-hydrogen-acceptor angle, and $f(r)$, $g(\theta)$ are scaling geometric functions.

Non-Strained Binding Affinity ($\Delta G_{\text{bind(NS)}}$), the binding free energy computed excluding the energetic cost of ligand conformational strain:

$$\Delta G_{\text{bind(NS)}} = \Delta G_{\text{bind}} - \Delta E_{\text{strain}}$$

ChEMBL Validation

Validation of bioactivity data for the target compound involved extracting canonical SMILES representations of 4,097 compounds from the ChEMBL database (ChEMBLID: ChEMBL1824). We processed, cleaned, and converted the SMILES data into a comma-separated values (CSV) file using Microsoft Excel 2016. Next, the CSV file containing the SMILES was converted to a 2D structure data file (.sdf) using DataWarrior (version 5.5.0). Further refinement and optimization of the 2D structures were performed using the LigPrep module within Maestro Schrödinger version 12.9. The resulting ligands were docked into the target protein using the Glide tools, employing methodologies consistent with previous docking analyses. We assessed docking-score predictive accuracy using correlation analysis. Docking scores for 1,300 compounds, randomly selected from the screening against the target, were plotted against their respective inhibitory values retrieved from the ChEMBL database. This comparative analysis aimed to identify any relationship between computational docking scores and experimentally determined bioactivity. Subsequently, the correlation coefficient was calculated to quantify the alignment between computational predictions and empirical observations.

Density Functional Theory (DFT) Calculations

Quantum-mechanical calculations were performed to evaluate the electronic properties and gas-phase conformation of the isolated ligand. The initial coordinates of the ligand were extracted directly from the docking pocket of the docked protein-ligand complex in Schrödinger Maestro version 2021-2. All DFT calculations were performed using the Jaguar module (v11.2) of the Schrödinger suite. Structural refinement and geometry optimizations were carried out within the Density Functional Theory (DFT) framework. We used the hybrid functional *B3LYP-D3 with the 6-31G*** basis set.

The model environmental effects mimic physiological conditions during relaxation; implicit solvation was treated with the Poisson-Boltzmann Finite Element (PBF) solvent model, using water as the solvent.

The optimized geometry represented a true local minimum on the potential energy surface rather than a transition-state saddle point. A vibrational frequency calculation was performed at the same level of theory.

The absence of imaginary (negative) frequencies confirmed that the optimized structure was a stable thermodynamic minimum.

Electronic properties were evaluated and mapped onto the optimized structure. These included the spatial distributions of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). The electrostatic potential (ESP) was computed on the molecular surface, and partial atomic charges were simultaneously calculated using Mulliken population analysis, ESP fitting on atomic centres, and Charge Model 5 (CM5) partitioning schemes to comprehensively evaluate the electronic distribution and local reactivity of the ligand structure.

***In Silico* Mutagenesis and Structural Reassessment**

The structural and electrostatic dependencies of *Cannabis sativa* phytoligands within the active site were investigated, and *in silico* site-directed mutagenesis was performed on the human Epidermal Growth Factor Receptor (EGFR). Point mutations were selectively introduced at codons 800 and 855 using the PyMOL Mutagenesis Wizard (PyMOL, v3.14). The optimal rotamer conformation was chosen by screening the rotamer library integrated into PyMOL, prioritizing conformations that minimized immediate steric clashes with neighbouring residues. The targeted amino acids, Asp800 and Asp855, are key, highly conserved residues located within the principal catalytic binding pocket of the EGFR kinase domain, where they mediate critical hydrogen-bonding networks and electrostatic interactions with incoming ligands.

The mutation relieved local geometric strain and eliminated close van der Waals contacts; the local structural environment within a 5.0 Å radius of the mutated site was optimized. Local minimization was achieved using PyMOL's sculpting and cleanup algorithms, allowing minor conformational adjustments to adjacent side chains. The overall molecular topology was subsequently refreshed using the rebuild command to resolve any graphical or structural artefacts. The evaluation of how structural, spatial, and electronic alterations at these key pocket positions modulate the resilience of phytoligand binding was carried out. Two distinct mutant receptor variants were engineered.

Aspartate-to-Glutamate Substitution (D800E/D855E): This variant is a conservative, charge-preserving mutation. Both aspartic acid and glutamic acid carry a net negative charge at physiological pH; however, glutamic acid has a longer side chain due to an additional methylene group (-CH₂). This substitution preserved the native negative electrostatic profile of the active site but introduced local steric hindrance and altered the spatial positioning of the carboxylate hydrogen-bond acceptors. This allowed testing whether the binding pocket could accommodate bulkier side chains without disrupting phytoligand affinity.

Aspartate-to-Asparagine Substitution (D800N/D855N): This variation served as an isosteric, charge-neutralizing mutation. Asparagine is structurally and spatially nearly identical to aspartic acid, but it replaces the negatively charged carboxylate group (-COO⁻) with a neutral, polar carboxamide functional group (-CONH₂). This mutation preserved the physical geometry of the binding-pocket residues while eliminating their ionic charge, allowing us to determine whether phytoligand binding depends strictly on a negative electrostatic field or on spatial stabilization.

Following the *in silico* amino acid substitutions, local energy minimization protocols were applied directly to the mutated regions. This minimization step was crucial for relieving localized thermodynamic strain, optimizing bond lengths, and resolving severe atom-atom steric clashes introduced by the new side-chain rotamers before initiating the docking phase.

1 The structural integrity, backbone topology, and local side-chain orientations of the relaxed wild-type and
2 mutated EGFR proteins were subsequently validated using the MolProbity server. Stereochemical
3 parameters, including steric atomic overlaps (clashscores), Ramachandran dihedral angles (ϕ and ψ), and
4 side-chain rotamer outliers, were assessed.

References

- 1 Leiter, A., Veluswamy, R. R. & Wisnivesky, J. P. The global burden of lung cancer: current status and future trends. *Nature reviews. Clinical oncology* **20**, 624-639, (2023).
- 2 Samarasekera, U. WHO launches 10th report on the global tobacco epidemic. *The Lancet. Oncology*, (2025).
- 3 Sharma, P. *et al.* Emerging trends in the novel drug delivery approaches for the treatment of lung cancer. *Chemico-biological interactions* **309**, 108720, (2019).
- 4 Alduais, Y., Zhang, H., Fan, F., Chen, J. & Chen, B. Non-small cell lung cancer (NSCLC): A review of risk factors, diagnosis, and treatment. *Medicine* **102**, e32899, (2023).
- 5 Friedlaender, A., Perol, M., Banna, G. L., Parikh, K. & Addeo, A. Oncogenic alterations in advanced NSCLC: a molecular super-highway. *Biomark Res* **12**, 24, (2024).
- 6 Tito, C., Masciarelli, S., Colotti, G. & Fazi, F. EGF receptor in organ development, tissue homeostasis and regeneration. *J Biomed Sci* **32**, 24, (2025).
- 7 Zubair, T. & Bandyopadhyay, D. Small molecule EGFR inhibitors as anti-cancer agents: discovery, mechanisms of action, and opportunities. *Int J Mol Sci* **24**, 2651, (2023).
- 8 Arrieta, O. *et al.* Cost-effectiveness analysis of first and second-generation EGFR tyrosine kinase inhibitors as first line of treatment for patients with NSCLC harboring EGFR mutations. *BMC cancer* **20**, 829, (2020).
- 9 Shyamsunder, S., Lu, Z., Takiar, V. & Waltz, S. E. Challenges and resistance mechanisms to EGFR targeted therapies in head and neck cancers and breast cancer: Insights into RTK dependent and independent mechanisms. *Oncotarget* **16**, 508-530, (2025).
- 10 Riboulet-Zemouli, K. 'Cannabis' ontologies I: Conceptual issues with Cannabis and cannabinoids terminology. *Drug Science, Policy and Law* **6**, 2050324520945797, (2020).
- 11 Zhu, D. T. *et al.* Global burden of disease due to opioid, amphetamine, cocaine, and cannabis use disorders, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *PloS one* **20**, e0328276, (2025).
- 12 Jett, J., Stone, E., Warren, G. & Cummings, K. M. Cannabis Use, Lung Cancer, and Related Issues. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* **13**, 480-487, (2018).
- 13 Akinloye, D. I. *et al.* Cannabis sativa demonstrates anti-hepatocellular carcinoma potentials in animal model: in silico and in vivo studies of the involvement of Akt. *Journal of Cannabis Research* **5**, 1-13, (2023).
- 14 Chalannavar, R. K., Hosamani, P. A., Gunagi, V. U., Divakar, M. S. & Malabadi, R. B. Cannabis sativa: Phytochemistry-An updated Review. (2026).
- 15 El-Mernissi, R. *et al.* A comparative study on phytochemical analysis and biological properties of three varieties of cannabis sativa L. seeds. *Open Life Sci* **21**, 20251211, (2026).
- 16 Schanknecht, E., Bachari, A., Nassar, N., Piva, T. & Mantri, N. Phytochemical Constituents and Derivatives of Cannabis sativa; Bridging the Gap in Melanoma Treatment. *Int J Mol Sci* **24**, 859, (2023).
- 17 Montle, R. M., Prinsloo, G. & Mtunzi, F. Cannabis sativa Phytochemicals in Cancer Therapy: Molecular Mechanisms and Therapeutic Potential. *Frontiers in Pharmacology* **17**, 1768210, (2026).
- 18 Heider, C. G., Itenberg, S. A., Rao, J., Ma, H. & Wu, X. Mechanisms of Cannabidiol (CBD) in Cancer Treatment: A Review. *Biology 2022, Vol. 11, Page 817* **11**, 817, (2022).

- 19 du Plessis, J., Deroubaix, A., Omar, A. & Penny, C. A Bioinformatic Analysis Predicts That Cannabidiol Could Function as a Potential Inhibitor of the MAPK Pathway in Colorectal Cancer. *Curr Issues Mol Biol* **46**, 8600-8610, (2024).
- 20 Rodriguez, S. M. B. *et al.* An Overview of EGFR Mechanisms and Their Implications in Targeted Therapies for Glioblastoma. *Int J Mol Sci* **24**, (2023).
- 21 Marin-Acevedo, J. A., Pellini, B., Kimbrough, E. O., Hicks, J. K. & Chiappori, A. Treatment Strategies for Non-Small Cell Lung Cancer with Common EGFR Mutations: A Review of the History of EGFR TKIs Approval and Emerging Data. *Cancers* **15**, (2023).
- 22 Barcelos, M. P. *et al.* in *Research topics in bioactivity, environment and energy: experimental and theoretical tools* 481-500 (Springer, 2022).
- 23 Guo, Y., Du, Z. & Shi, T. Structural analysis of interactions between epidermal growth factor receptor (EGFR) mutants and their inhibitors. *Biophysica* **3**, 203-213, (2023).
- 24 Ahmadi, A., Mohammadnejadi, E. & Razzaghi-Asl, N. Gefitinib derivatives and drug-resistance: A perspective from molecular dynamics simulations. *Comput Biol Med* **163**, 107204, (2023).
- 25 Amelia, T., Kartasasmita, R. E., Ohwada, T. & Tjahjono, D. H. Structural Insight and Development of EGFR Tyrosine Kinase Inhibitors. *Molecules* **27**, 819, (2022).
- 26 Aghababaei, F. & Hadidi, M. Recent Advances in Potential Health Benefits of Quercetin. *Pharmaceuticals* **16**, 1020, (2023).
- 27 Ay, M. *et al.* in *Nutraceuticals* 749-755 (Elsevier, 2021).
- 28 Bayazid, A. B. & Lim, B. O. Quercetin Is An Active Agent in Berries against Neurodegenerative Diseases Progression through Modulation of Nrf2/HO1. *Nutrients* **14**, (2022).
- 29 Panthi, V. K., Imran, M., Chaudhary, A., Paudel, K. R. & Mohammed, Y. The significance of quercetin-loaded advanced nanoformulations for the management of diabetic wounds. *Nanomedicine* **18**, 391--411, (2023).
- 30 Zhang, Y. Q., Li, K., Guo, Q. & Li, D. A New Risk Model Based on 7 Quercetin-Related Target Genes for Predicting the Prognosis of Patients With Lung Adenocarcinoma. *Front Genet* **13**, 890079, (2022).
- 31 Ahmed, E. A. & Rajendran, P. Orientin and Cancer Suppression: Molecular Mechanisms and Synergistic Effects. *Journal of Cancer* **17**, 847-854, (2026).
- 32 Fahmy, M. I. *et al.* Orientin: a comprehensive review of a promising bioactive flavonoid. *Inflammopharmacology* **33**, 1713-1728, (2025).
- 33 Szkudlarek, J., Piwowarczyk, L. & Jelinska, A. Antioxidants Acteoside and Orientin as Emerging Agents in Synergistic Cancer Therapy: A Focus on Innovative Applications. *Antioxidants (Basel)* **14**, (2025).
- 34 Kanhed, A. M. *et al.* Identification of potential Mpro inhibitors for the treatment of COVID-19 by using systematic virtual screening approach. *Molecular diversity* **25**, 383-401, (2021).
- 35 Ghosh, R., Roy, S., Rakshit, G., Singh, N. K. & Maiti, N. J. Pharmacophore modeling in drug design. *Computational Methods for Rational Drug Design*, 167-194, (2025).
- 36 Rao, M. S. *et al.* Novel Computational Approach to Predict Off-Target Interactions for Small Molecules. *Front Big Data* **2**, 25, (2019).
- 37 Mirzaei, S., Ghodsi, R., Hadizadeh, F. & Sahebkar, A. 3D-QSAR-Based Pharmacophore Modeling, Virtual Screening, and Molecular Docking Studies for Identification of Tubulin Inhibitors with Potential Anticancer Activity. *BioMed research international* **2021**, 6480804, (2021).

- 38 Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports* 2017 7:1 7, 1--13, (2017).
- 39 Wu, K. *et al.* Overcoming Challenges in Small-Molecule Drug Bioavailability: A Review of Key Factors and Approaches. *Int J Mol Sci* **25**, (2024).
- 40 Sharapova, A., Ol'khovich, M., Blokhina, S. & Perlovich, G. L. Experimental Examination of Solubility and Lipophilicity as Pharmaceutically Relevant Points of Novel Bioactive Hybrid Compounds. *Molecules* **27**, (2022).
- 41 Wu, D. *et al.* The blood-brain barrier: structure, regulation, and drug delivery. *Signal Transduct Target Ther* **8**, 217, (2023).
- 42 Neumaier, F., Zlatopolskiy, B. D. & Neumaier, B. Drug Penetration into the Central Nervous System: Pharmacokinetic Concepts and In Vitro Model Systems. *Pharmaceutics* **13**, 1542, (2021).
- 43 Rana, H., Pandey, R. & Singh, P. in *Herb-Drug Interactions: A Medicinal Approach* 159-172 (Bentham Science Publishers, 2025).
- 44 Akash, S. *et al.* Antimalarial drug discovery against malaria parasites through haplopine modification: An advanced computational approach. *Journal of cellular and molecular medicine* **27**, 3168--3188, (2023).
- 45 Parray, Z. A., Shahid, M. & Islam, A. Insights into Fluctuations of Structure of Proteins: Significance of Intermediary States in Regulating Biological Functions. *Polymers (Basel)* **14**, (2022).
- 46 Rezaei, S., Sefidbakht, Y., Uskoković, V. J. J. o. B. S. & Dynamics. Comparative molecular dynamics study of the receptor-binding domains in SARS-CoV-2 and SARS-CoV and the effects of mutations on the binding affinity. **40**, 4662-4681, (2022).
- 47 Pal, P., Debnath, R., Jana, B. & Chakraborty, S. Solvent accessible surface area normalized protein-water hydrogen bonds define protein folded state stability and amyloid formation. *Physical chemistry chemical physics : PCCP* **28**, 4321-4332, (2026).
- 48 Chen, D. *et al.* Regulation of protein-ligand binding affinity by hydrogen bond pairing. *Sci Adv* **2**, (2016).
- 49 Hassan, A. M. *et al.* Evaluating the Binding Potential and Stability of Drug-like Compounds with the Monkeypox Virus VP39 Protein Using Molecular Dynamics Simulations and Free Energy Analysis. *Pharmaceutics (Basel)* **17**, (2024).
- 50 Hayes, J. M., Archontis, G. J. M. d.-s. o. s. & macromolecules, b. MM-GB (PB) SA calculations of protein-ligand binding free energies. 171-190, (2012).
- 51 Rao, P., Goswami, D. & Rawal, R. M. Revealing the molecular interplay of curcumin as *Culex pipiens* Acetylcholine esterase 1 (AChE1) inhibitor. *Scientific Reports* 2021 11:1 **11**, 1-18, (2021).
- 52 Salaritabar, A. *et al.* Targeting Hedgehog signaling pathway: Paving the road for cancer therapy. *Pharmacol Res* **141**, 466-480, (2019).
- 53 Zhao, L., Zhi, M. & Frenking, G. J. I. J. o. Q. C. The strength of a chemical bond. **122**, e26773, (2022).
- 54 Du, X. *et al.* Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci* **17**, (2016).
- 55 Silva-Pinto, P. A. *et al.* Phytochemical insights into flavonoids in cancer: Mechanisms, therapeutic potential, and the case of quercetin. *Heliyon* **11**, e42682, (2025).

- 56 Zhang, N. *et al.* Benzothiadiazole-Fused Cyanoindone: A Superior Building Block for Designing Ultra-Narrow Bandgap Electron Acceptor with Long-Range Ordered Stacking. *Angewandte Chemie (International ed. in English)* **64**, e202420090, (2025).
- 57 Yang, Y., Zhai, X., Jiang, Y., Wang, J. & Yao, C. Spectra-Stability Relationships in Organic Electron Acceptors: Excited-State Analysis. *Molecules* **30**, (2025).
- 58 Ayers, P. W., Mohamed, M. & Heidar-Zadeh, F. The hard/soft acid/base rule: A perspective from conceptual density-functional theory. *Conceptual Density Functional Theory: Towards a New Chemical Reactivity Theory* **1**, 263-279, (2022).
- 59 Kim, S. *et al.* PubChem Substance and Compound databases. *Nucleic acids research* **44**, D1202--D1213, (2016).
- 60 Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **27**, 221-234, (2013).
- 61 Friesner, R. A. *et al.* Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of medicinal chemistry* **49**, 6177--6196, (2006).
- 62 Banerjee, P., Eckert, A. O., Schrey, A. K. & Preissner, R. ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic acids research* **46**, W257-W263, (2018).
- 63 Owoloye, A. J. *et al.* Molecular docking, simulation and binding free energy analysis of small molecules as PfHT1 inhibitors. *PloS one* **17**, e0268269, (2022).

Acknowledgements

The authors appreciate the students and scientists at the Centre for Genomic Research in Biomedicine (CeGRIB), Mountain Top University, for the clerical support during manuscript drafting.

Funding

Kolapo Oyebola was supported by the Fogarty Emerging Global Leader Grant (NIH-K43TW011926) from the US National Institutes of Health and an APTI-18-07 Grant from the African Academy of Sciences in partnership with the Bill and Melinda Gates Foundation. The funders had no role in study design, data collection, analysis, the decision to publish, or manuscript preparation.

Author contributions

Afolabi Owoloye, Samuel Olubode, & Precious Akinnusi conceptualized the idea & performed the preliminary analysis; Evans Asamoah Adu & Natalia Shakela performed the molecular dynamics simulation & DFT analysis; David Oyerinde & Olamide Onawola prepared the first draft; Kolapo Oyebola wrote the review & edited; all authors reviewed the manuscript.

Availability of Data and Materials

The dataset supporting the conclusions of this article is within the manuscript.

Declarations

Competing Interests: The authors declare no competing interests.

Figure

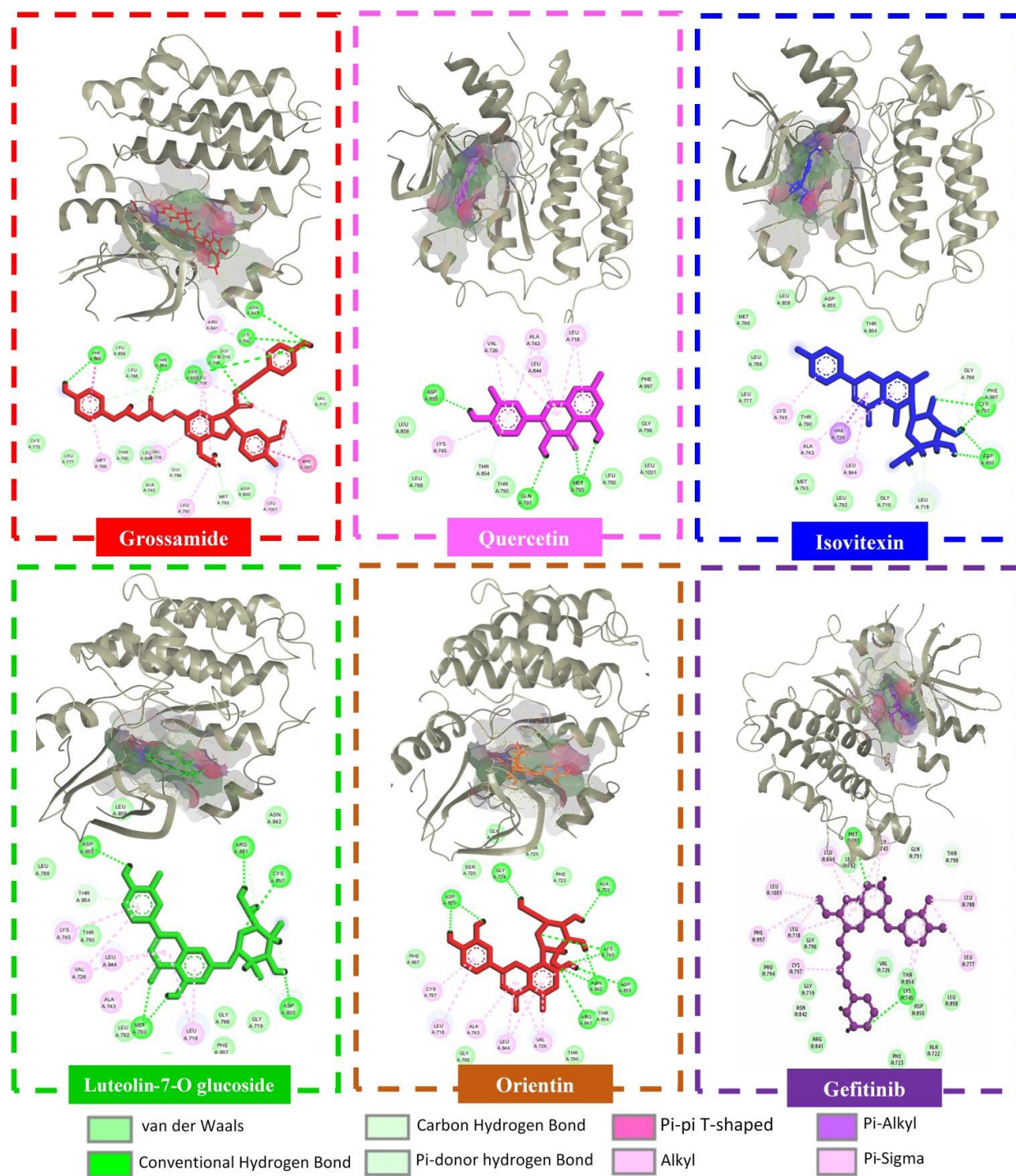


Figure 1: Intermolecular interactions between the test compounds and EGFR

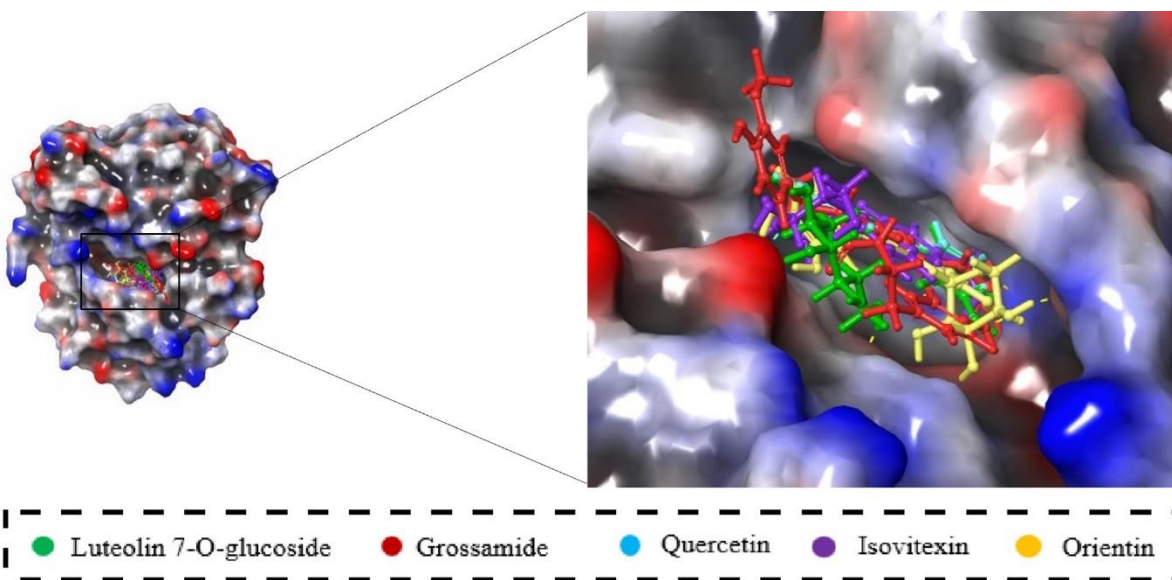


Figure 2: The top-scoring compounds in the binding pocket of EGFR tyrosine kinase

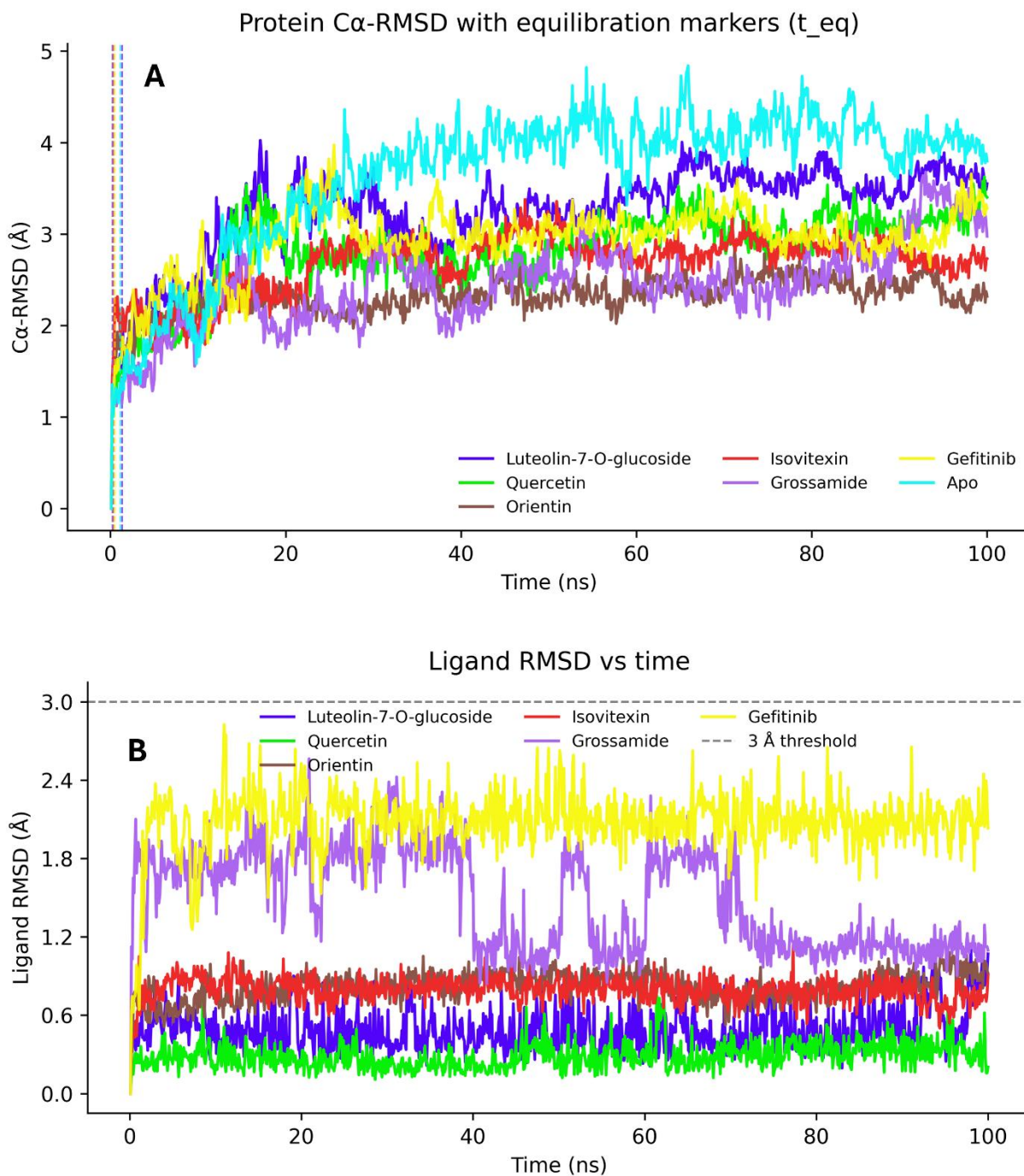


Figure 3: Line representation of the evolution of C α (A) and ligand (B) root mean square deviation (RMSD) throughout the MD simulations of the EGFR (3POZ) complex with the lead compounds.

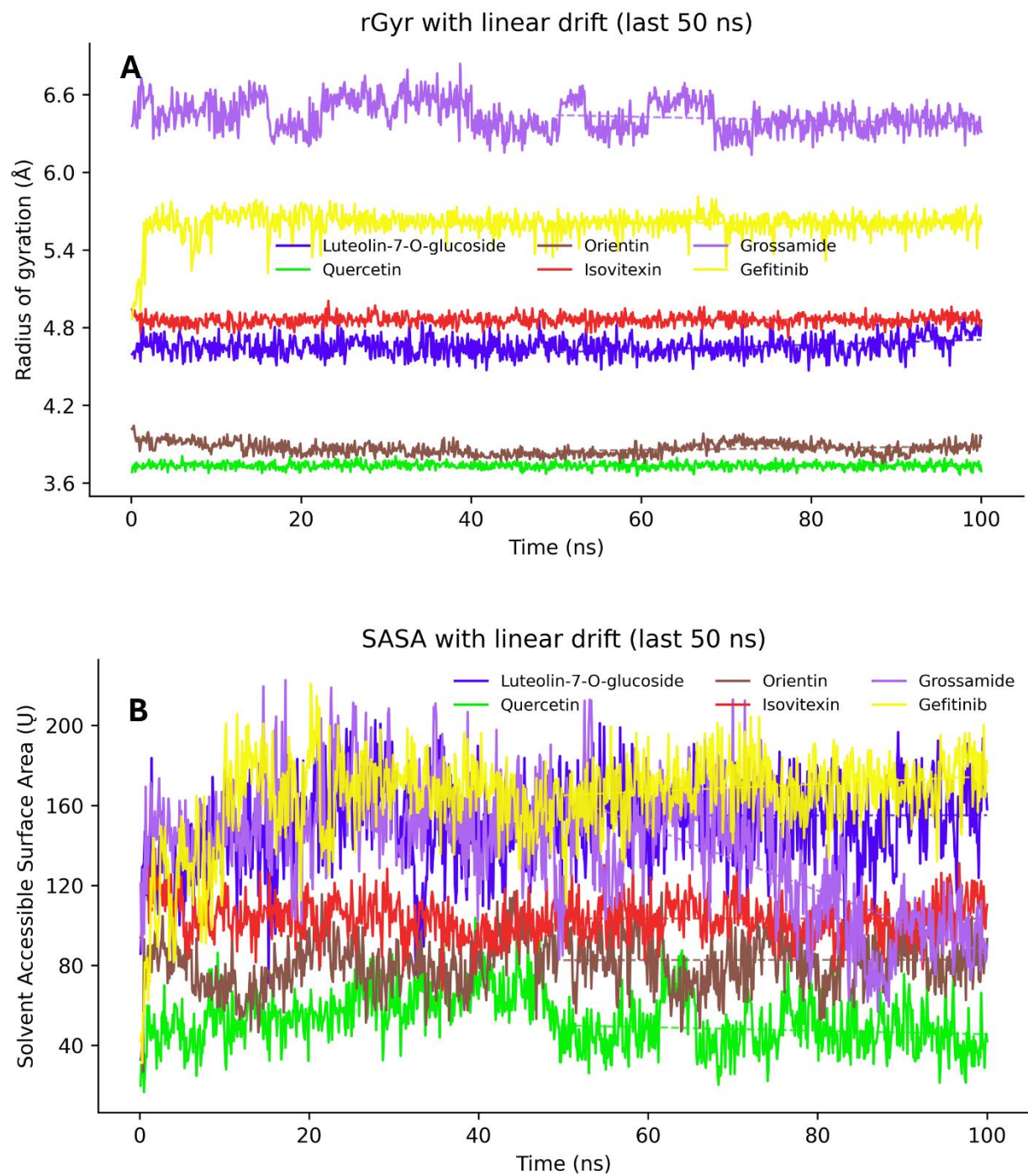


Figure 4: Root mean square fluctuation (RMSF) of the lead complexes

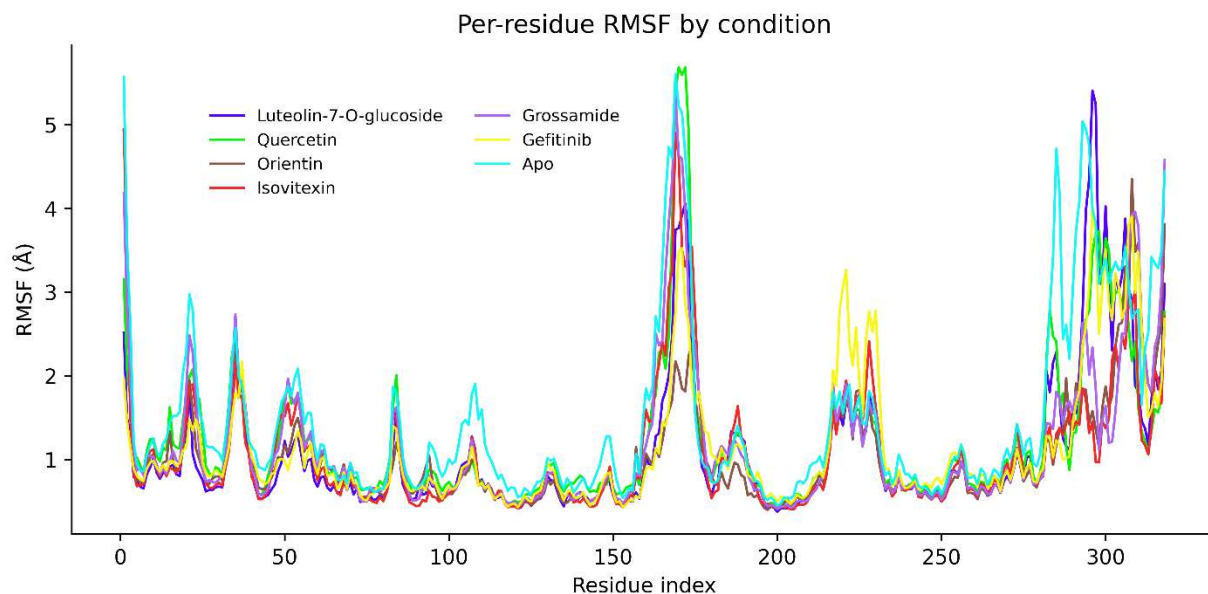


Figure 5: Radius of gyration (A) and Solvent Accessible Surface Area, SASA (B) of the selected complexes calculated during the molecular dynamics simulation evolution

Table 1: The binding affinities of the top-scoring compounds

S/No	Compounds	Binding affinity (kcal/mol)
1	Luteolin 7- <i>O</i> -glucoside	-10.88
2	Grossamide	-10.09
3	Quercetin	-10.00
4	Isovitexin	-9.78
5	Orientin	-9.64
6	Vitexin	-9.533
7	Pinoresinol	-9.483
8	Kaempferol	-9.434
9	Luteolin	-9.409
10	Secoisolariciresinol	-9.227
11	Gefitinib (Standard drug)	-9.579

Table 5: Binding free energy components matrix

Energy Term (kcal/mol)	Grossamide	Isovitexin	Orientin	Luteolin 7- <i>O</i> - glucoside	Quercetin	Gefitinib
ΔG_{bind}	-75.59 ± 33.24	-56.90 ± 16.45	-53.05 ± 18.29	-49.11 ± 13.58	-43.27 ± 3.56	-47.41 ± 28.51
ΔG_{vdW}	-50.05 ± 37.83	-43.46 ± 18.93	-38.14 ± 16.72	-40.30 ± 13.93	-37.34 ± 2.46	-41.57 ± 24.67
$\Delta G_{\text{Coulomb}}$	-21.80 ± 12.48	-28.71 ± 8.84	-26.57 ± 13.27	-24.00 ± 8.98	-18.18 ± 2.94	-24.28 ± 20.82
ΔG_{Lipo}	-23.30 ± 17.95	-11.77 ± 5.15	-11.71 ± 5.18	-7.92 ± 2.81	-6.94 ± 0.56	-14.59 ± 8.72
ΔG_{Hbond}	-1.77 ± 1.37	-2.82 ± 1.01	-3.03 ± 1.39	-3.39 ± 1.31	-3.59 ± 0.49	-0.81 ± 0.58
$\Delta G_{\text{bind(NS)}}$	-68.49 ± 51.75	-60.01 ± 26.59	-55.17 ± 24.49	-53.56 ± 19.29	-46.35 ± 3.61	-50.85 ± 30.59

Total Binding (ΔG_{bind}); Van der Waals (ΔG_{vdW}); Coulombic ($\Delta G_{\text{Coulomb}}$); Lipophilic (ΔG_{Lipo}); Hydrogen Bond (ΔG_{Hbond}); Net Free Energy ($\Delta G_{\text{bind(NS)}}$).

Table 6: Quantum Chemical Parameters and Global Reactivity Descriptors

Compound	Grossamide	Quercetin	Gefitinib	Luteolin 7- <i>O</i> - glucoside	Orientin	Isovitexin
E_{HOMO} (eV)	-5.66	-5.81	-5.86	-5.98	-6.01	-6.05
E_{LUMO} (eV)	-1.59	-1.89	-1.75	-1.94	-1.98	-1.74
Energy Gap ΔE (eV)	4.07	3.92	4.12	4.04	4.02	4.31
Hardness η (eV)	2.03	1.96	2.06	2.02	2.01	2.16
Softness σ (eV ⁻¹)	0.25	0.26	0.24	0.25	0.25	0.235
Electronegativity χ (eV)	3.62	3.85	3.8	3.96	4.00	3.90
Electrophilicity ω (eV)	3.22	3.78	3.52	3.88	3.97	3.53
ΔG_{solv} (kcal/mol)	-36.25	-26.3	-65.33	-32.48	-42.01	-39.19

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementary.docx](#)