

Article

Mutant-Selective Binding of *Phyllanthus niruri* Phytochemicals to EGFR T790M: A Quantum-Classical Mechanistic Study

William D. Lituma-González ¹, Diksha Dinesh Kumar ¹, Amogh V. Arunprasad ¹, Tanishque Verma ¹, Shrutika Pillai ², Fabian R. Jimenez ¹, Sasikumar J. Mahalingam ³ and Shanmugamurthy Lakshmanan ^{1,*}

- ¹ Department of Integrative and Transdisciplinary Pharmacognosy, Siddha Vetha University, 211 Warren st, Newark, NJ 07103, USA; william.lituma@yachaytech.edu.ec (W.D.L.-G.); dikshadk2020@gmail.com (D.D.K.); amoghvarunprasad@gmail.com (A.V.A.); tanishque.verma@siddhavethammultiversity.org (T.V.); fabian.jimenez@nfinnovations.com (F.R.J.)
² Department of Neuroscience, University of Virginia, P.O. Box 400224, Charlottesville, VA 22904, USA; sruthika.pillai123@gmail.com
³ Department of Microbiology, Karpagam Academy of Higher Education, Coimbatore 641021, India; sasikumar.jmahalingam@kahedu.edu.in
* Correspondence: s.lakshmanan@siddhavethammultiversity.org; Tel.: +1-(201)-406-3358

Abstract

Epidermal growth factor receptor (EGFR) mutations drive hepatocellular carcinoma (HCC) progression, and the T790M gatekeeper substitution is the predominant mechanism of acquired resistance to EGFR-targeted therapies. Herein, we present multiscale quantum-classical *in silico* predictions of *Phyllanthus niruri* phytochemicals as non-covalent EGFR T790M binders, employing molecular docking, 100 ns molecular dynamics, MM-PBSA/MM-GBSA, per-residue decomposition, PCA/LDA, DFT at B3LYP-D3(BJ)/def2-TZVP, and comparative wild-type EGFR simulations. Five phytochemicals exhibited computationally predicted binding affinities against EGFR T790M exceeding the non-covalent binding component of osimertinib (-25.74 kcal/mol): corilagin (-53.71 ± 5.05 kcal/mol), eriodictyol-7-rhamnopyranoside (-44.35 ± 4.51 kcal/mol), isoquercetin (-44.23 ± 2.92 kcal/mol), rutin (-42.15 ± 4.50 kcal/mol), and kaempferol-4-rhamnoside (-41.68 ± 3.69 kcal/mol). Wild-type EGFR simulations (PDB 1M17) yielded a selectivity index (IS) of 2.76 for corilagin ($\Delta\Delta G_{\text{bind}} = +34.22$ kcal/mol), indicating T790M-preferential binding. Osimertinib reproduced its clinically established T790M selectivity under identical conditions (IS = 1.43; $\Delta\Delta G_{\text{bind}} = +7.74$ kcal/mol), providing internal methodological validation. DFT at B3LYP-D3(BJ)/def2-TZVP established the quantum-mechanical basis for corilagin's electrostatic affinity: its molecular electrostatic potential (MEP) surface minimum ($V_{s,\text{min}} = -46.92$ kcal/mol) directly predicts the largest MM-PBSA electrostatic term ($\Delta E_{\text{ele}} = -52.48$ kcal/mol), establishing quantum-classical coherence. Supervised PCA/LDA of 7416 MM-PBSA trajectory frames identified solvation energy (ΔG_{SOLV}) as the primary pharmacological class discriminant, with the first discriminant function (LD1) capturing 93.1% of inter-class binding variance. Collectively, corilagin (hydrolyzable tannin), eriodictyol-7-rhamnopyranoside (flavonoid glycoside), and phylltetralin (lignan) constitute diverse computational leads from *P. niruri* warranting experimental validation as T790M-directed agents in HCC.

Keywords: hepatocellular carcinoma; molecular dynamics simulation; MM-PBSA; per-residue energy decomposition; conceptual DFT; free energy profiling; gatekeeper resistance; kinase inhibitor resistance

Academic Editors: Giovanni Tarantino and Rolf Teschke

Received: 14 February 2026

Revised: 11 July 2026

Accepted: 12 July 2026

Published: 23 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Hepatocellular carcinoma (HCC) ranks as the sixth most frequently diagnosed malignancy worldwide, with an estimated 905,677 new cases in 2022, and constitutes the third leading cause of cancer-related mortality globally [1]. The systemic treatment landscape remains narrow: sorafenib yields objective response rates of only 2–5% [2], while lenvatinib—the current first-line standard of care [3]—achieves a 24% objective response rate complicated by acquired resistance in approximately 30% of patients within six to eight months of treatment initiation [4–6]. This clinical trajectory leaves a substantial proportion of advanced HCC patients, particularly those with underlying cirrhosis, without effective systemic options at the time of disease progression.

Epidermal growth factor receptor (EGFR) signaling occupies a central position in hepatocarcinogenesis, with EGFR protein overexpression documented in 40–80% of human HCC tissues [7,8]. EGFR functions as a convergence hub integrating mitogenic signals through the PI3K/AKT/mTOR, RAS/MAPK, and STAT3 pathways, driving hepatocyte proliferation, epithelial-to-mesenchymal transition, and apoptotic resistance in HCC models and primary tumor specimens [8]. Despite this biological rationale, clinical trials of erlotinib and gefitinib monotherapy in HCC yielded suboptimal outcomes, attributed to primary resistance and the absence of validated predictive biomarkers [7]. These observations identify sustained EGFR signaling as a therapeutically tractable pathway of escape, particularly relevant in the post-lenvatinib resistance context.

Tantiwetrueangdet et al. identified 13 missense EGFR mutations in human HCC tissues spanning exons 18–23, establishing that EGFR mutational heterogeneity in liver cancer extends well beyond the canonical lung-cancer alteration landscape [9]. HCC-derived EGFR mutants are functioning, EGF-dependent, and erlotinib-resistant [9]. The mechanistic basis of this resistance is well characterized: the T790M gatekeeper substitution increases ATP-binding affinity approximately 10-fold, competitively displacing first-generation EGFR inhibitors from the kinase pocket [10]. These observations underscore the need for inhibitory strategies engaging EGFR T790M through mechanisms distinct from ATP-competitive displacement [11,12].

EGFR reactivation constitutes an established driver of acquired lenvatinib resistance in HCC. Jin et al. demonstrated that EGFR activation limits the lenvatinib response and that pharmacological EGFR inhibition restores drug sensitivity in preclinical HCC models [13]. Downstream resistance is further sustained through an EGFR-STAT3-ABCB1 signaling axis that promotes drug efflux in lenvatinib-resistant cells [14]. Genomic profiling of lenvatinib-resistant HCC cohorts identified EGFR amplification as a recurrent molecular alteration [15], and a prospective exploratory trial of lenvatinib plus gefitinib in EGFR-overexpressing lenvatinib-resistant HCC subsequently reported an objective response rate of 30.0% by mRECIST criteria [16]. These data position EGFR as a genomically validated and therapeutically actionable target in the lenvatinib-refractory HCC setting.

Clinical evidence further corroborates the therapeutic relevance of EGFR inhibition in HCC. A phase II trial of bevacizumab combined with erlotinib reported disease control in a subset of EGFR-expressing advanced HCC patients [17], and a case report of CTC-guided EGFR-targeted therapy in a patient with metastatic HCC refractory to standard systemic agents documented a confirmed partial response maintained for 14 months [18]. These data delineate a specific therapeutic gap: the absence of non-covalent binders capable of overcoming EGFR T790M-mediated resistance with a computationally predicted hepatoprotective pharmacological profile compatible with the cirrhotic background prevalent among HCC patients.

Phyllanthus niruri L. (*Phyllanthaceae*), documented in traditional medicine systems of Asia and Latin America for hepatic conditions and viral hepatitis [19–21], harbors more than 65 bioactive constituents including hydrolyzable tannins (corilagin, MW = 634.45

Da), flavonoid glycosides (rutin, eriodictyol-7-rhamnopyranoside), and lignans (phyltetralin, niranthin) [22–24]. Three independent lines of evidence support its evaluation as a source of computationally predicted leads against EGFR T790M in HCC: pharmacological EGFR inhibition reverses lenvatinib resistance in HCC preclinical models [13]; *P. niruri* dry extract induces apoptosis in human HCC cell lines [25]; and *Phyllanthus* species exhibit antitumoral activity in HCC-relevant in vitro systems [26]. These convergences identify *P. niruri* as a source of in silico predictions grounded in mechanistically coherent biological precedent rather than arbitrary phytochemical selection [27–32].

The crystal structure PDB 7A6K captures the EGFR T790M kinase domain in an inactive α C-helix-out conformation [33]. The co-present V948R substitution was deliberately introduced to prevent asymmetric receptor dimerization during crystallization, enabling high-resolution structural characterization of the monomeric inactive state required for inhibitor design [34]. Critically, V948R resides in the distal C-lobe of the kinase domain and does not interact with the ATP-binding cleft, the T790M gatekeeper position, or the pharmacologically relevant catalytic residues (Lys745, Met793, Cys797) that govern inhibitor binding [34]. This crystallographic construct has served as the structural basis for multiple peer-reviewed EGFR T790M inhibitor design studies from the Rauh laboratory [35,36], validating its suitability for computational investigation of T790M-directed non-covalent agents.

To address these scientific questions systematically, the present study employs an integrated multiscale computational framework structured as a hierarchy of analytical levels. Molecular docking against the EGFR T790M kinase domain (PDB 7A6K) established initial binding pose candidates from the *P. niruri* phytochemical library, providing entry-level affinity ranking [37]. Candidate compounds were subjected to 100-nanosecond molecular dynamics simulations, from which MM-PBSA binding free energies were derived to provide thermodynamically converged affinity estimates [38], and per-residue MM-GBSA energy decomposition at single-residue resolution identified the hot spot architecture governing selectivity. Supervised principal component and linear discriminant analyses (PCA/LDA) of the resulting MM-PBSA trajectory data then characterized the mechanistic basis of differential binding across chemical classes. Quantum-chemical analysis at the B3LYP-D3(BJ)/def2-TZVP level of density functional theory (DFT), encompassing molecular electrostatic potential (MEP) mapping, conceptual DFT descriptors, and non-covalent interaction (NCI) analysis, established the electronic-structure foundation for these in silico predictions [39]. Comparative 100-nanosecond molecular dynamics simulations against wild-type EGFR (PDB 1M17) provided quantitative mutant selectivity indices under identical simulation conditions [40].

The present study aims to identify, from the phytochemical repertoire of *Phyllanthus niruri* L., computationally predicted T790M-selective EGFR binders characterized through an integrated multiscale computational framework encompassing quantum mechanics, classical molecular dynamics, and multivariate chemometrics, with mutant selectivity rigorously quantified against wild-type EGFR under identical simulation conditions, thereby generating ranked computational leads that warrant experimental validation as non-covalent T790M-directed agents in hepatocellular carcinoma.

2. Results

2.1. Virtual Screening of *Phyllanthus niruri* Phytochemicals Against EGFR T790M

Thirty-seven *P. niruri* phytochemicals and two clinical standards were docked against EGFR T790M/V948R (PDB 7A6K) using a validated SwissDock protocol (TAK-788 redocking RMSD = 0.80 Å; DUD-E ROC-AUC = 0.91). Fourteen phytochemicals met the selection threshold of $\Delta G \leq -7.91$ kcal/mol (Supplementary Figure S1). The mean ΔG of selected compounds (-8.93 ± 0.28 kcal/mol) differed significantly from that of non-selected

compounds (-6.43 ± 0.71 kcal/mol; Mann–Whitney U, $p = 1.2 \times 10^{-5}$; Supplementary Figure S2 and Table S1).

Among the 14 selected phytochemicals, rutin showed the most negative docking score ($\Delta G = -9.27$ kcal/mol). Osimertinib and mobocertinib scored -9.74 and -9.31 kcal/mol, respectively. Five additional phytochemicals scored within 1.0 kcal/mol of mobocertinib: nirphyllin (-9.00), 2,3-desmethoxy seco-isolintetralin (-8.98), nirurin D (-8.97), isoquercitrin (-8.52), and hypophyllanthin (-8.35 kcal/mol; Supplementary Figure S2, Supplementary Table S1).

Interaction profiling identified three residue-contact patterns (Figure 1). Flavonoid glycosides (rutin, isoquercitrin, astragalín) showed hydrogen-bonding with Met793 and Asp800, with conservation of water molecule HOH-611 (occupancy $\geq 75\%$), a pattern also observed for osimertinib (Supplementary Figure S3). Lignans (nirphyllin, hypophyllanthin, phyllanthin) showed predominant van der Waals contacts with Leu718, Val726, and Met790, with additional van der Waals contacts to Arg948; however, Arg948 arises from the V948R crystallographic stabilization mutation specific to PDB 7A6K and is not present in native EGFR T790M—these contacts are therefore construct-specific and should not be interpreted as pharmacologically relevant binding interactions (see Section 3.7). Corilagin showed a π -sulfur interaction with Cys797 and hydrogen-bonding contacts with Leu788, Arg841, and Asn842, including a double hydrogen bond with ASP855 (Figure 1). Contact frequencies are reported in Supplementary Table S2.

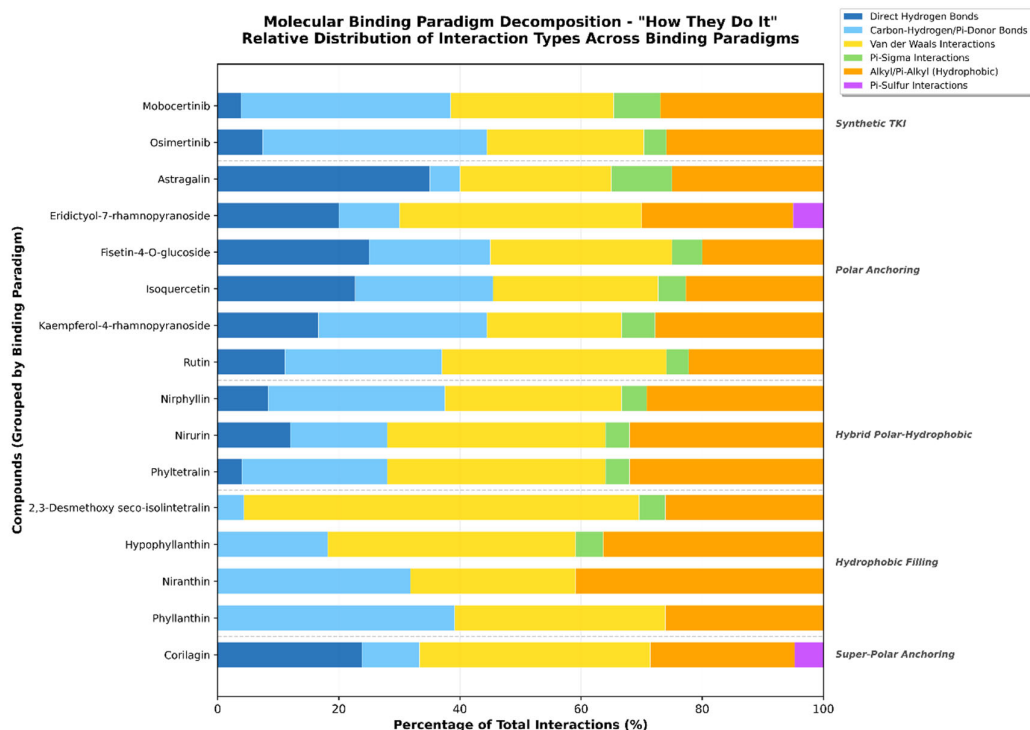


Figure 1. Residue-contact interaction profiles for 14 selected *P. niruri* phytochemicals and clinical reference compounds against EGFR T790M (PDB 7A6K). Stacked bar chart showing the relative frequency of hydrogen bonds, van der Waals, π -alkyl, π -sulfur, and cation- π interactions per compound (triplicate docking). Three interaction profiles are shown: (1) Corilagin: Hydrogen bond frequency $>40\%$, with additional van der Waals contacts. (2) Flavonoid glycosides (rutin, isoquercitrin, astragalín): Hydrogen bond frequency 25–35% with π -alkyl contacts; (3) Lignans (nirphyllin, hypophyllanthin, phyllanthin): Van der Waals contacts $>50\%$ with π -alkyl interactions 20–30%. Osimertinib and mobocertinib profiles are included for reference. Bar segments scaled to 100% per compound.

2.2. Multi-Platform Computational ADMET Profiling of 14 *P. niruri* Phytochemicals

Predicted ADMET profiles for the 14 selected phytochemicals and two clinical controls (osimertinib, mobocertinib) were assessed across three platforms (SwissADME, pkCSM, Deep-PK; Table 1, Supplementary Figures S4–S6). All values represent *in silico* predictions; experimental pharmacokinetic validation was not performed.

Eight of 14 phytochemicals (57%) showed predicted CYP3A4 inhibition confidence <0.10, including corilagin, nirurin, rutin, isoquercitrin, astragalín, kaempferol-4'-rhamnopyranoside, eriodictyol-7-rhamnopyranoside, and fisetin-4-O-glucoside. Osimertinib and mobocertinib showed CYP3A4 inhibition confidence >0.63 (Table 1). Lignans showed intermediate CYP3A4 inhibition (symbol $\pm$). Glycosides and polyphenols showed predicted UGT-centered Phase II metabolism (UGT1A1, UGT1A9, UGT1A10), whereas lignans showed predicted CYP-distributed metabolism (CYP2C8, CYP2C9, CYP2C19; Table 1, Supplementary Table S2).

Table 1. Predicted ADMET properties for 14 *P. niruri* phytochemicals and clinical reference compounds, derived from three independent computational platforms (SwissADME, pkCSM, Deep-PK). Risk symbols: + = predicted low risk (platform consensus Pa < 0.30); $\pm$ = predicted intermediate or conflicting risk (Pa 0.30–0.70); – = predicted high risk (consensus Pa > 0.70). All values are *in silico* predictions; no experimental pharmacokinetic data were obtained in this study.

Compound	Solubility (Log S) ^a	Permeability (Caco-2) ^b	P-gp Sub- strate	Primary Metabolic Pathway ^c	CYP3A 4 DDI Risk ^d	Hepatotox- icity Risk ^e	Cardiotox- icity Risk ^f
Clinical Controls							
Osimertinib	Good (–3.30)	High (1.14)	$\pm$	CYP3A4-Dependent	–	$\pm$	–
Mobocertinib	Poor (–6.48)	Moderate (0.34)	$\pm$	CYP3A4-Dependent	–	+	$\pm$
Lignans							
2,3-Desmethoxy seco-isolintetra- lin	Moderate (–4.73)	High (1.14)	$\pm$	CYP-Diversified (2C8/2C9)	$\pm$	$\pm$	+
Hy- pophyllanthin	Moderate (–4.80)	High (1.10)	+	CYP-Diversified (2C9/2C19)	$\pm$	+	+
Niranthin	Moderate (–4.66)	High (1.09)	$\pm$	CYP-Diversified (2C9/2C19)	$\pm$	$\pm$	+
Nirphyllin	Moderate (–4.69)	High (1.10)	+	CYP-Diversified (2C9/2C8)	$\pm$	$\pm$	+
Phyllanthin	Moderate (–4.54)	High (1.04)	$\pm$	CYP-Diversified (2C8/2B6)	–	+	–
Phyltetralin	Moderate (–4.57)	High (0.98)	$\pm$	CYP-Diversified (2C8/2B6)	$\pm$	+	$\pm$
Glycosides & Polyphenols							
Astragalín	Good (–3.18)	Moderate (0.31)	+	UGT-Centered (1A1/1A10)	+	$\pm$	+
Corilagin	Good (–3.92)	Very Low (–1.34)	+	UGT-Centered (1A1/1A9)	+	–	+
Eriodictyol-7- rhamnopyra- noside	Good (–3.51)	Low (–0.46)	+	UGT-Centered (1A1/1A10)	+	$\pm$	–
Fisetin-4-O-glu- coside	Good (–2.83)	Low (0.27)	+	UGT-Centered (1A1/1A10)	+	–	+
Isoquercitrin	Good (–3.35)	Low (0.24)	+	UGT-Centered (1A1/1A10)	+	$\pm$	+
Kaempferol-4'- rhamnopyra- noside	Good (–3.60)	Low (0.20)	+	UGT-Centered (1A1/1A10)	+	$\pm$	+

Nirurin	Good (−3.60)	Low (0.18)	+	UGT-Centered (1A1/1A9)	+	+	+
Rutin	Good (−3.30)	Very Low (−0.95)	+	UGT-Centered (1A1/1A10)	+	±	+

Footnotes: ^a Solubility (Log S): Consensus from SwissADME (ESOL). Good > −4; Moderate −4 to −6; Poor < −6. ^b Permeability (Caco-2): Predicted by pkCSM (log P_{app}, cm/s). High > 0.9; Moderate 0.3–0.9; Low 0–0.3; Very Low < 0. ^c Metabolic Pathway: Inferred from Way2Drug SOMP/SMP. ^d CYP3A4-Dependent: >80% metabolism via CYP3A4 (high DDI risk). CYP-Diversified: metabolism distributed across multiple CYP isoforms. UGT-Centered: predominant Phase II glucuronidation (low DDI risk). ^e Hepatotoxicity: Integrated assessment from pkCSM, Deep-PK (DILI I/II), and AD-VERPred. ^f Cardiotoxicity: Integrated assessment of hERG inhibition (pkCSM/Deep-PK) and arrhythmia/cardiac failure risk (ADVERPred).

hERG blocker confidence was >0.99 for both clinical controls. Eight phytochemicals (astragalin, corilagin, hypophyllanthin, isoquercitrin, kaempferol-4'-rhamnopyranoside, niranthin, nirphyllin, nirurin) showed hERG confidence <0.10. Phyllanthin and eriodictyol-7-rhamnopyranoside showed values in the high predicted-risk category (symbol −; Table 1, Supplementary Figures S4–S6).

DILI II confidence for osimertinib was 0.58–0.67. Corilagin and fisetin-4-O-glucoside showed high predicted DILI II confidence (symbol −); all other glycosides and polyphenols showed low predicted risk (+). Among lignans, hypophyllanthin, nirphyllin, and phyllanthin showed low predicted risk (+); niranthin and 2,3-desmethoxy seco-isolintetralin showed mixed predictions (±; Table 1).

Six lignans showed predicted high Caco-2 permeability (log P_{app} 0.98–1.14 cm/s). Glycosides and polyphenols showed predicted low-to-very-low permeability (log P_{app} −1.34 to 0.31). Phyltetralin and hypophyllanthin showed LogBB >0.2; osimertinib showed LogBB = −0.045 (Table 1).

Osimertinib and mobocertinib showed P-glycoprotein (P-gp) substrate confidence >0.75. Among phytochemicals, glycosides, polyphenols, hypophyllanthin, and nirphyllin showed P-gp substrate classification (+); remaining lignans showed mixed predictions (Table 1).

Based on predicted Caco-2 permeability and primary metabolic pathway, the phytochemicals separated into two groups: (i) lignans with high permeability and CYP-distributed metabolism; (ii) glycosides and polyphenols with low permeability, UGT-centered metabolism, and low predicted CYP3A4 inhibition (Supplementary Table S2, Supplementary Figures S4–S6). The pharmacokinetic implications are discussed in Section 3.

2.3. PASS and KinScreen Predictions Suggest Distinct Polypharmacological Profiles and High Hepatoprotective Probability Scores Across *P. niruri* Phytochemicals

PASS Online (ChEMBL v.21) and KinScreen (442 human kinases) predictions are probabilistic estimates (Pa = probability of activity) and do not constitute experimental confirmation. All values are in silico predictions.

Hepatoprotective Pa ≥ 0.940 was predicted for nirurin (0.974), rutin (0.968), isoquercitrin (0.961), and astragalin (0.940). Antineoplastic Pa for these compounds ranged from 0.832 to 0.849. Osimertinib and mobocertinib showed antineoplastic Pa of 0.797 and 0.682, respectively, and hepatoprotective Pa < 0.30 (Supplementary Figure S7, Supplementary Table S3). Twelve of 14 phytochemicals showed predicted carbonic anhydrase (CA-III, CA-IX, CA-XII) Pa ≥ 0.40 (Supplementary Table S4); clinical controls showed Pa < 0.30 for all three CA isoforms.

The mean number of predicted high-confidence kinase interactions (Pa ≥ 0.40) was 3.5 ± 2.1 for phytochemicals vs. 28.0 ± 12.0 for clinical controls. Flavonoid glycosides

showed predicted NEK6 interactions ($P_a = 0.71$ – 0.86); lignans showed predicted DYRK2 (0.40 – 0.51) and CLK4 (0.37 – 0.44) interactions; phyltetralin and hypophyllanthin showed predicted MRCK-A interactions ($P_a > 0.55$). Osimertinib showed CLK4 $P_a = 0.19$ and PKN1 $P_a = 0.15$ (Figure 2, Supplementary Table S4).

Collectively, the phytochemicals showed computationally predicted interactions spanning EGFR T790M docking (Section 2.1), kinase targets, P-glycoprotein modulation (P_a 0.60 – 0.87), CYP3A4 non-inhibition for 8/14 compounds (Section 2.2), and CA-IX/CA-XII targeting for 12/14 compounds.

2.4. Molecular Dynamics Simulations Reveal Class-Dependent Structural Stability, Interaction Networks, and MM-PBSA Binding Free Energies for Selected *P. niruri* Phytochemicals Against EGFR T790M

2.4.1. Conformational Stability over 100 ns Trajectories

Backbone C α RMSD and ligand heavy-atom RMSD were monitored over 100 ns NPT trajectories for 14 phytochemicals and two clinical controls against EGFR T790M/V948R (Figure 3, Supplementary Figures S10 and S11, Supplementary Tables S5 and S6). Osimertinib and mobocertinib showed mean backbone RMSD of 3.09 ± 0.41 Å ($\tau = 2.82$ ns) and 2.66 ± 0.34 Å ($\tau = 9.96$ ns), respectively; ligand RMSD was 5.37 ± 0.64 Å (1.3% stable frames) and 5.45 ± 0.55 Å (0.1% stable frames; Figure 3A, Supplementary Figure S10A).

Among glycosides, eriodictyol-7-rhamnopyranoside and kaempferol-4'-rhamnoside showed the lowest backbone RMSD ($\mu = 2.16 \pm 0.25$ and 2.21 ± 0.24 Å; $\tau = 3.54$ and 4.32 ns) and ligand RMSD of 2.31 ± 0.20 Å (100% stable) and 2.47 ± 0.29 Å (96.6% stable), respectively. Isoquercitrin showed the lowest ligand RMSD in the dataset (1.78 ± 0.30 Å; 100% stable). Secondary glycosides (astragalin, rutin, nirurin) showed intermediate backbone RMSD values (2.67 – 2.95 Å; Figure 3B, Supplementary Table S5). Corilagin showed backbone RMSD $\mu = 2.56 \pm 0.45$ Å with $\tau = 13.77$ ns (Figure 3A).

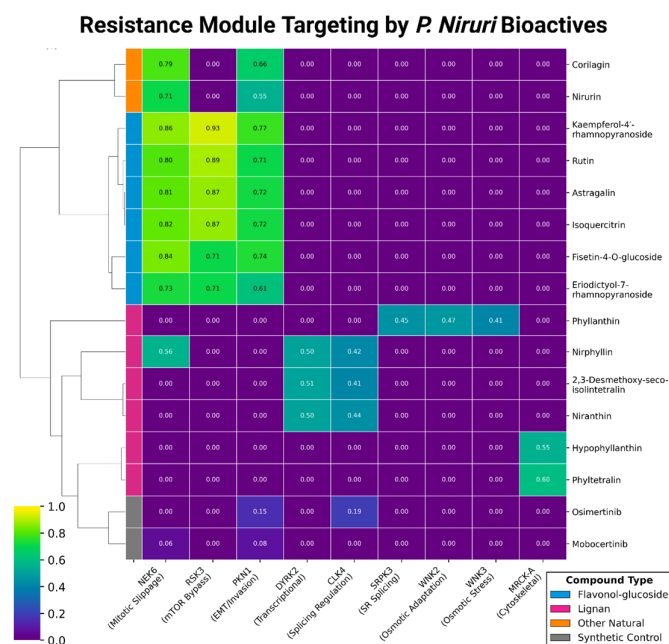


Figure 2. KinScreen-predicted kinase interaction profiles (P_a) for 14 *P. niruri* phytochemicals and clinical reference compounds across 442 human kinases. Unsupervised hierarchical clustering of P_a values ($P_a \geq 0.40$ threshold). Rows: compounds; columns: kinase targets with $P_a \geq 0.40$ in at least one compound. Color scale: P_a from 0.0 (purple) to 1.0 (yellow). Three compound clusters are observed: flavonoid glycosides (NEK6 P_a 0.71 – 0.86 ; RSK3, PKN1), lignans (DYRK2 P_a 0.40 – 0.51 ; CLK4

0.37–0.44), and a phyltetralin/hypophyllanthin sub-cluster (MRCK-A $P_a > 0.55$). Clinical controls showed $P_a < 0.30$ for NEK6, DYRK2, CLK4, and MRCK-A (osimertinib CLK4 $P_a = 0.19$; PKN1 $P_a = 0.15$). All predictions are *in silico*; experimental kinase assays were not performed.

Among lignans, four of six compounds showed backbone RMSD > 3.0 Å: phyllanthin (3.10 ± 0.50 Å), phyltetralin (3.14 ± 0.56 Å, $\tau = 11.26$ ns), nirphyllin (3.35 ± 0.47 Å), and 2,3-desmethoxy seco-isolintetralin (3.57 ± 0.53 Å; Figure 3B, Supplementary Table S5). Ligand RMSD for niranthin and phyllanthin was 7.48 ± 1.32 Å and 5.67 ± 1.33 Å, respectively (Supplementary Figure S10B, Supplementary Table S6). Phyltetralin showed the lowest lignan ligand RMSD (2.18 ± 0.47 Å; 92.6% stable). Fisetin-4-O-glucoside underwent a conformational transition at ~ 25 ns (RMSD ≈ 2.5 Å $\rightarrow >6.5$ Å; Supplementary Figure S10A).

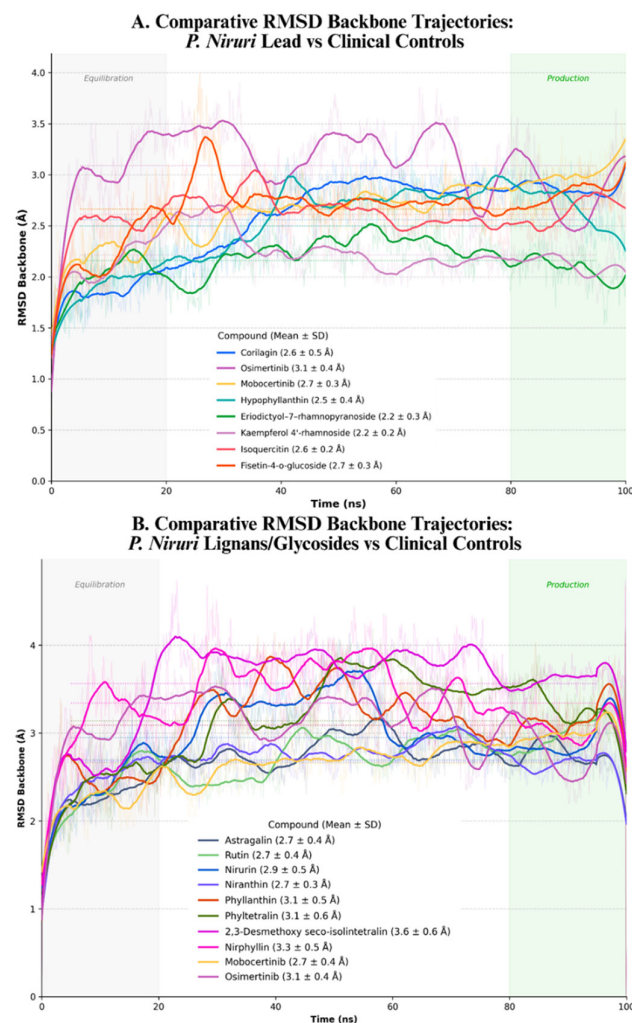


Figure 3. Backbone C α RMSD trajectories for EGFR T790M/V948R complexes over 100 ns MD simulations. Solid lines: Savitzky–Golay smoothed; shaded bands: ± 1 SD. Gray shading: equilibration phase (0–20 ns); green shading: production phase (80–100 ns). **(A)** Lead compounds and controls: eriodictyol-7-rhamnopyranoside ($\mu = 2.16 \pm 0.25$ Å, $\tau = 3.54$ ns), kaempferol-4'-rhamnoside ($\mu = 2.21 \pm 0.24$ Å, $\tau = 4.32$ ns), corilagin ($\mu = 2.56 \pm 0.45$ Å, $\tau = 13.77$ ns), isoquercitrin ($\mu = 2.61 \pm 0.25$ Å), mobocertinib ($\mu = 2.66 \pm 0.34$ Å, $\tau = 9.96$ ns), osimertinib ($\mu = 3.09 \pm 0.41$ Å). **(B)** Secondary compounds: astragalin ($\mu = 2.67 \pm 0.33$ Å), rutin ($\mu = 2.70 \pm 0.37$ Å), nirurin ($\mu = 2.95 \pm 0.47$ Å), phyllanthin ($\mu = 3.10 \pm 0.50$ Å), phyltetralin ($\mu = 3.14 \pm 0.56$ Å, $\tau = 11.26$ ns), nirphyllin ($\mu = 3.35 \pm 0.47$ Å), 2,3-desmethoxy seco-isolintetralin ($\mu = 3.57 \pm 0.53$ Å), niranthin ($\mu = 2.69 \pm 0.30$ Å). Complete metrics in Supplementary Table S5.

Mean radius of gyration (R_g) for glycoside-bound complexes ranged from 1.976 to 1.987 nm ($\sigma = 0.010$ – 0.022 nm); clinical controls showed mean $R_g = 1.996$ – 1.999 nm (Supplementary Figure S11A, Supplementary Table S7). Corilagin-bound complex showed mean $R_g = 1.994$ nm with low conformational dispersion. Nirphyllin-bound complex showed the highest mean R_g (2.042 nm, Range = 0.141 nm; Supplementary Figure S11B, Supplementary Table S7). Mean solvent-accessible surface area (SASA) for kaempferol-4'-rhamnoside-bound complex was $15,151 \pm 266 \text{ \AA}^2$; isoquercitrin $15,237 \pm 356 \text{ \AA}^2$; rutin $15,240 \pm 435 \text{ \AA}^2$ (Supplementary Figure S12A, Supplementary Table S8). Osimertinib-bound complex showed mean SASA = $15,599 \pm 370 \text{ \AA}^2$. Nirphyllin-bound complex showed mean SASA = $15,950 \pm 341 \text{ \AA}^2$ (Range = $2578 \text{ \AA}^2$; Supplementary Figure S12B, Supplementary Table S8).

2.4.2. Intermolecular Hydrogen-Bond Occupancy and Per-Residue RMSF

Intermolecular hydrogen-bond counts were computed over the 100 ns production trajectory (Figure 4, Supplementary Table S9). Corilagin showed mean H-bonds = 5.83 ± 1.16 (maximum 9; residence time 33.30 ns). Isoquercitrin, eriodictyol-7-rhamnopyranoside, and kaempferol-4'-rhamnoside showed mean H-bonds of 5.18 ± 1.40 , 4.95 ± 1.52 , and 4.84 ± 1.10 , respectively (maxima 8–9; residence times > 100 ns; Figure 4A). Osimertinib showed mean H-bonds = 1.49 ± 0.61 (maximum 3; residence time 3.60 ns); mobocertinib = 0.98 ± 0.51 bonds (maximum 3; residence time 0.86 ns; Figure 4A,B). Phylltetralin and hypophyllanthin showed mean H-bonds of 0.98 ± 0.39 and 1.05 ± 0.46 , respectively. Phyllanthin showed 0.16 ± 0.42 bonds (residence time 0.27 ns); nirphyllin = 1.05 ± 0.73 bonds (residence time 0.51 ns; Figure 4B, Supplementary Table S9).

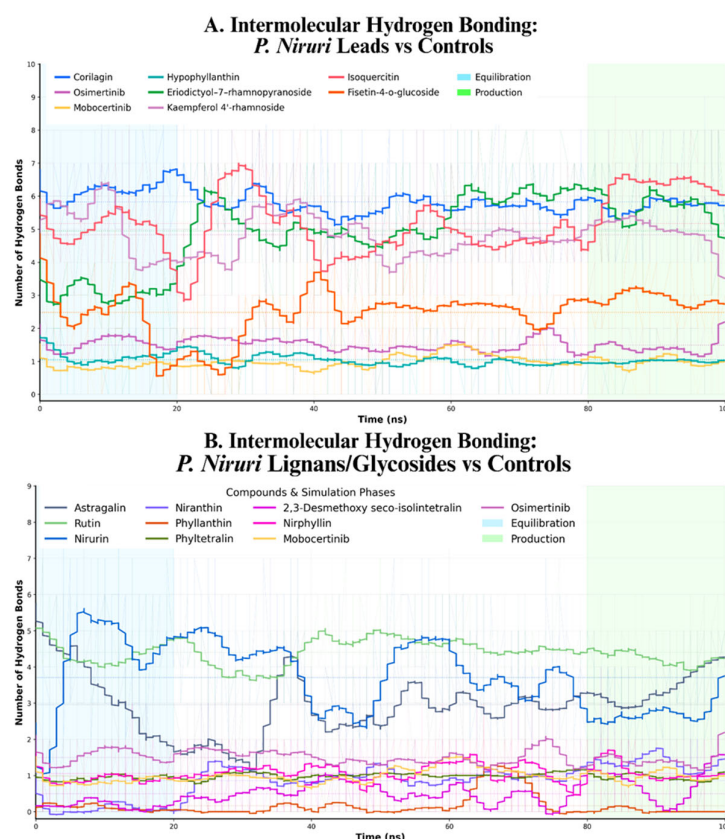


Figure 4. Intermolecular hydrogen-bond counts between each ligand and EGFR T790M/V948R over 100 ns simulations. Shaded bands: ± 1 SD. (A) Lead compounds and controls: corilagin (mean = 5.83 ± 1.16 ; maximum 9; residence time 33.30 ns), isoquercitrin (5.18 ± 1.40 ; maximum 9; residence

time > 100 ns), eriodictyol-7-rhamnopyranoside (4.95 ± 1.52 ; maximum 8; residence time > 100 ns), kaempferol-4'-rhamnoside (4.84 ± 1.10 ; maximum 9; residence time > 100 ns), fisetin-4-O-glucoside (intermediate occupancy), osimertinib (1.49 ± 0.61 ; maximum 3; residence time 3.60 ns), mobocertinib (0.98 ± 0.51). (B) Secondary compounds: phylltetralin (0.98 ± 0.39 ; maximum 2), hypophyllanthin (1.05 ± 0.46 ; maximum 3), phyllanthin (0.16 ± 0.42 ; residence time 0.27 ns), nirphyllin (1.05 ± 0.73 ; residence time 0.51 ns), 2,3-desmethoxy seco-isolintetralin (0.49 ± 0.57). Complete metrics in Supplementary Table S9.

Per-residue $\Delta\text{RMSF} = \text{RMSF}_{\text{compound}} - \text{RMSF}_{\text{osimertinib}}$ was computed over the 100 ns trajectory (Figure 5, Supplementary Figure S13, Supplementary Table S10). Corilagin showed global $\Delta\text{RMSF} = -0.212 \text{ \AA}$, with $\Delta\text{RMSF} < -0.05 \text{ \AA}$ at P-loop residues 750–757, and $\Delta\text{RMSF} < -0.10 \text{ \AA}$ at hinge/ATP-binding residues 874–876 and 887–891. Eriodictyol-7-rhamnopyranoside showed $\Delta\text{RMSF} = -0.117 \text{ \AA}$ with reductions at residues 834–837 and 861–877. Isoquercitrin showed $\Delta\text{RMSF} = -0.030 \text{ \AA}$; kaempferol-4'-rhamnoside showed reductions at residues 750–757 and 862–876. Nirphyllin showed mean $\text{RMSF} = 1.865 \text{ \AA}$ ($\Delta\text{RMSF} = +0.429 \text{ \AA}$), 2,3-desmethoxy seco-isolintetralin $\Delta\text{RMSF} = +0.347 \text{ \AA}$, and phylltetralin $\Delta\text{RMSF} = +0.328 \text{ \AA}$ (Figure 5, Supplementary Table S10). The ΔRMSF heatmap (Figure 5) shows negative ΔRMSF concentrated at P-Loop, Hinge/T790M, and A-Loop for corilagin and eriodictyol-7-rhamnopyranoside, and positive ΔRMSF distributed across all kinase domains for nirphyllin, 2,3-desmethoxy seco-isolintetralin, and phylltetralin.

Differential Structural Stability Profile: *P. Niruri* Leads Baseline Reference: Osimertinib ($\Delta = \text{RMSF}_{\text{ligand}} - \text{RMSF}_{\text{osi}}$)

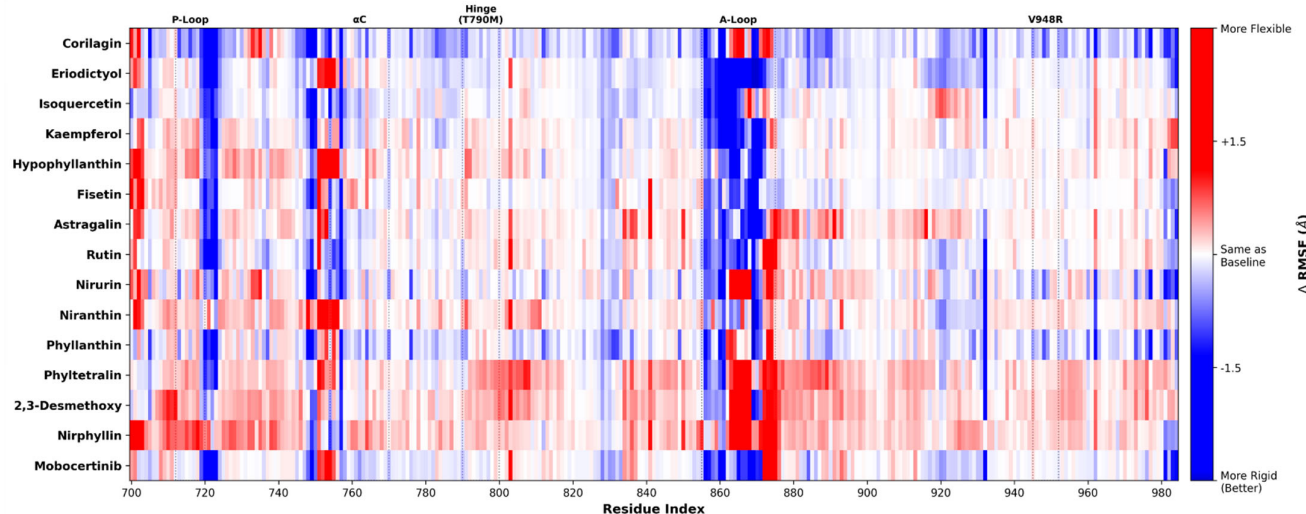


Figure 5. Per-residue differential RMSF heatmap ($\Delta\text{RMSF} = \text{RMSF}_{\text{compound}} - \text{RMSF}_{\text{osimertinib}}$) for 16 compounds against EGFR T790M/V948R over 100 ns trajectories. Color scale: $-1.5 \text{ \AA}$ (blue, reduced flexibility) to $+1.5 \text{ \AA}$ (red, increased flexibility). Compounds ordered by decreasing MM-PBSA affinity (Supplementary Table S11). Dashed vertical lines demarcate P-Loop, G-Loop, Hinge/DFG motif, and Activation Loop. Corilagin: global $\Delta\text{RMSF} = -0.212 \text{ \AA}$; $\Delta\text{RMSF} < -0.10 \text{ \AA}$ at P-Loop (750–757), Hinge/T790M (874–876), and A-Loop (855–875). Eriodictyol-7-rhamnopyranoside: $\Delta\text{RMSF} = -0.117 \text{ \AA}$ (residues 834–837, 861–877). Isoquercitrin: $\Delta\text{RMSF} = -0.030 \text{ \AA}$; kaempferol-4'-rhamnoside: reductions at 750–757 and 862–876. Nirphyllin: mean $\text{RMSF} = 1.865 \text{ \AA}$, $\Delta\text{RMSF} = +0.429 \text{ \AA}$; 2,3-desmethoxy seco-isolintetralin: $\Delta\text{RMSF} = +0.347 \text{ \AA}$; phylltetralin: $\Delta\text{RMSF} = +0.328 \text{ \AA}$. Complete per-residue values in Supplementary Table S10.

2.4.3. MM-PBSA Binding Free Energies for EGFR T790M

MM-PBSA binding free energies were computed over the 100 ns production trajectory at $T = 310$ K, $P = 1$ atm, NPT ensemble (Figure 6, Supplementary Table S11). The non-covalent binding component of osimertinib ($\Delta G_{\text{bind}} = -25.74 \pm 3.37$ kcal/mol) was used as the internal reference value for all pairwise comparisons; this value reflects exclusively the non-covalent contribution, as the irreversible Michael addition to Cys797 that governs osimertinib's clinical potency is not captured by the standard MM-PBSA formalism (see Section 3.6).

Five phytochemicals showed predicted binding affinities more negative than the osimertinib non-covalent reference. Corilagin showed the most favorable $\Delta G_{\text{bind}} = -53.71 \pm 5.05$ kcal/mol ($\Delta\Delta G$ relative to osimertinib: -27.97 kcal/mol; complete pairwise statistics in Supplementary Table S11). Eriodictyol-7-rhamnopyranoside showed $\Delta G_{\text{bind}} = -44.35 \pm 4.51$ kcal/mol ($\Delta\Delta G = -18.61$ kcal/mol). Isoquercetin showed $\Delta G_{\text{bind}} = -44.23 \pm 2.92$ kcal/mol ($\Delta\Delta G = -18.49$ kcal/mol). Rutin showed $\Delta G_{\text{bind}} = -42.15 \pm 4.50$ kcal/mol ($\Delta\Delta G = -16.41$ kcal/mol). Kaempferol-4'-rhamnoside showed $\Delta G_{\text{bind}} = -41.68 \pm 3.69$ kcal/mol ($\Delta\Delta G = -15.94$ kcal/mol).

Additional phytochemicals showed ΔG_{bind} values distributed across the following range: nirurin -35.30 ± 5.42 kcal/mol, phylltetralin -34.54 ± 4.01 kcal/mol, hypophyllanthin -30.51 ± 2.54 kcal/mol, nirphyllin -29.06 ± 2.52 kcal/mol, fisetin -27.90 ± 3.18 kcal/mol, astragalin -25.33 ± 2.71 kcal/mol, 2,3-desmethoxy seco-isolintetralin -24.83 ± 2.34 kcal/mol, niranthin -25.03 ± 2.89 kcal/mol, and phyllanthin -22.15 ± 1.94 kcal/mol (Figure 6, Supplementary Table S11). Phyllanthin showed a ΔG_{bind} value less negative than the osimertinib non-covalent reference ($\Delta\Delta G = +3.59$ kcal/mol; Supplementary Table S11). Complete pairwise Kruskal–Wallis tests with Benjamini–Hochberg correction and Cohen's d effect sizes for all inter-compound comparisons are reported in Supplementary Table S11.

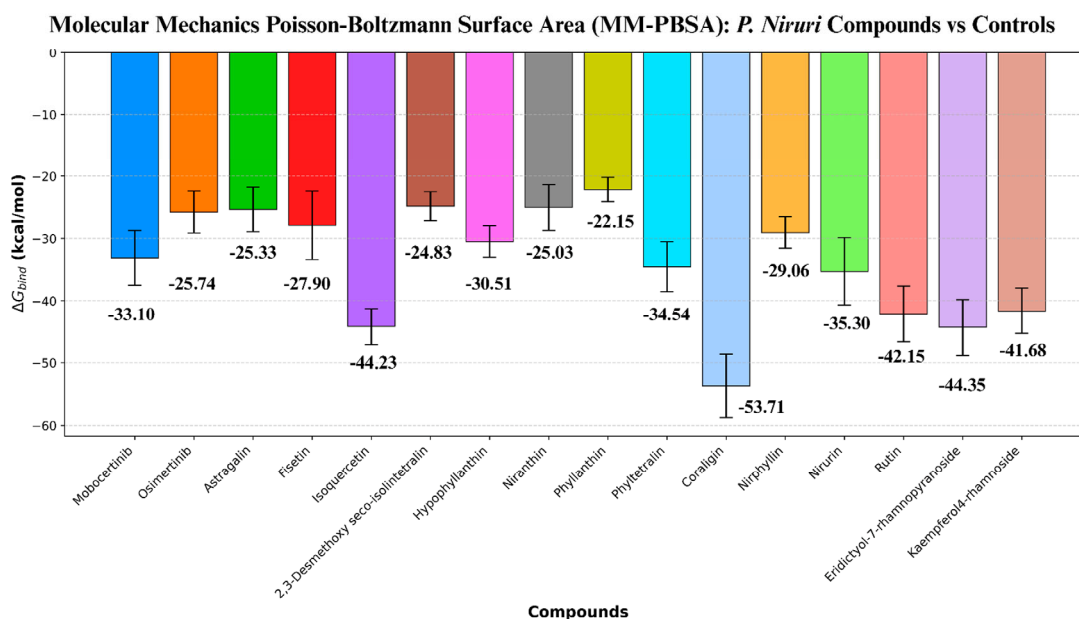


Figure 6. MM-PBSA binding free energies (ΔG_{bind} , kcal/mol) for 16 *Phyllanthus niruri* phytochemicals and two covalent reference inhibitors against EGFR T790M/V948R (PDB 7A6K; $T = 310$ K, $P = 1$ atm, NPT ensemble, 201 production frames per system, 50–100 ns window). Error bars represent ± 1 standard deviation from MM-PBSA ensemble averaging. The vertical dashed reference line indicates the non-covalent binding component of osimertinib ($\Delta G_{\text{bind}} = -25.74 \pm 3.37$ kcal/mol); compounds to the left of this line show less favorable predicted binding relative to this reference. Five phytochemicals exceeded the osimertinib non-covalent reference: corilagin (-53.71 ± 5.05 kcal/mol),

eriodictyol-7-rhamnopyranoside (-44.35 ± 4.51 kcal/mol), rutin (-42.15 ± 4.50 kcal/mol), isoquercetin (-44.23 ± 2.92 kcal/mol), and kaempferol-4-rhamnoside (-41.68 ± 3.69 kcal/mol). Osimertinib and mobocertinib form irreversible covalent bonds with Cys797 via Michael addition; the ΔG_{bind} values reported here (-25.74 ± 3.37 and -33.10 ± 4.41 kcal/mol, respectively) reflect the non-covalent binding component only and are not directly comparable with the phytochemical affinities. Complete per-compound statistics, convergence diagnostics, and pairwise Kruskal–Wallis comparisons are provided in Supplementary Table S11.

Mobocertinib yielded $\Delta G_{\text{bind}} = -33.10 \pm 4.41$ kcal/mol, reflecting solely its non-covalent binding component. As with osimertinib, mobocertinib forms an irreversible covalent bond with Cys797 via Michael addition; its clinical potency is governed by this covalent mechanism, which is not captured by standard MM-PBSA non-covalent calculations. The reported value should not be interpreted as its total binding affinity, and direct comparison with the phytochemical ΔG_{bind} values is methodologically inappropriate (see Section 3.6).

2.5. Comparative Docking and Molecular Dynamics of T790M Versus Wild-Type EGFR Quantify Predicted Mutant Selectivity

2.5.1. Redocking Protocol Validation and Comparative Docking Scores

Redocking of co-crystallized ligands validated the SwissDock 2024 protocol independently for both receptor forms. Erlotinib redocked into wild-type EGFR (PDB: 1M17) with an RMSD of $1.8315 \text{ \AA}$ relative to its crystallographic pose; TAK-788 (mobocertinib) redocked into the T790M/V948R mutant (PDB: 7A6K) with an RMSD of $0.80 \text{ \AA}$ (Supplementary Table S12). Both values fell below the standard $2.0 \text{ \AA}$ acceptance threshold, confirming protocol integrity across the two conformational states of the kinase domain (C-helix-out inactive in T790M; C-helix-in active in WT). The T790M/V948R model harbors an additional V948R substitution in the C-lobe; this residue is distal to the ATP-binding site and does not contact the binding pocket, and therefore is not expected to systematically influence docking results.

Comparative docking scores across the T790M and WT models are summarized in Table 2. Osimertinib scored -9.737 kcal/mol (T790M) and -8.0 kcal/mol (WT), a difference ($\Delta\Delta G$) of $+1.74$ kcal/mol in favor of the mutant form. Among the phytochemicals, corilagin scored -8.677 kcal/mol (T790M) and -8.4 kcal/mol (WT) ($\Delta\Delta G = +0.28$ kcal/mol). Phyltetralin scored -8.507 kcal/mol (T790M) and -8.0 kcal/mol (WT) ($\Delta\Delta G = +0.51$ kcal/mol). Eriodictyol-7-rhamnopyranoside scored -8.561 kcal/mol (T790M) and -9.7 kcal/mol (WT) ($\Delta\Delta G = -1.14$ kcal/mol). This apparent WT preference at the static docking level is inverted by the dynamic MM-PBSA analysis, which captures the polar desolvation penalty imposed on this glycosylated flavonoid in explicit solvent (Section 2.5.3).

Table 2. Comparative docking estimated binding affinities (ΔG , kcal/mol) for phytochemical candidates and controls against T790M EGFR (PDB: 7A6K) and wild-type EGFR (PDB: 1M17). Contacts = number of interacting residues. $\Delta\Delta G > 0$ indicates computationally predicted mutant preference.

Compound	ΔG T790M (kcal/mol)	ΔG WT (kcal/mol)	$\Delta\Delta G$ (kcal/mol)	Contacts T790M	Contacts WT
Corilagin	-8.677	-8.4	$+0.28$	18	36
Eriodictyol-7-rhamnopyranoside	-8.561	-9.7	-1.14	17	37
Phyltetralin	-8.507	-8.0	$+0.51$	20	28
Osimertinib	-9.737	-8.0	$+1.74$	19	38

Footnote: $\Delta\Delta G = \Delta G(\text{T790M}) - \Delta G(\text{WT})$; positive values indicate computationally predicted mutant preference at the static docking level.

2.5.2. MD Stability Classification, Catalytic Residue Rigidification, and Hydrogen-Bond Occupancy in Wild-Type EGFR Simulations

Extended 100 ns MD simulations of each ligand–WT EGFR complex were conducted to assess dynamic stability. Backbone C α RMSD values below 4.0 Å are considered indicative of structural equilibrium; ligand RMSD values below 3.0 Å indicate confinement within the binding site (Supplementary Table S13). Backbone RMSD trajectories for the WT complexes (Figure 7A) demonstrated stabilization during the final 25 ns, with mean values ranging from 9.506 ± 1.993 Å (eriodictyol-7-O-rhamnoside) to 12.197 ± 1.824 Å (erlotinib).

Ligand RMSD profiles (Figure 7B; Supplementary Table S13) classified WT-binding behavior into three stability categories. Class I (mean RMSD < 5 Å) comprised only eriodictyol-7-O-rhamnoside (3.075 ± 0.580 Å, maximum 5.332 Å), indicative of sustained binding-site occupancy throughout the 100 ns trajectory. Class II (mean RMSD 5–15 Å) encompassed osimertinib (10.324 ± 3.119 Å), corilagin (10.673 ± 4.069 Å), and phyltetralin (6.485 ± 2.835 Å), consistent with partial or fluctuating binding engagement. Class III (mean RMSD > 15 Å) comprised erlotinib alone (41.574 ± 22.278 Å, maximum 84.542 Å), which underwent complete dissociation from the WT binding site during the simulation.

Per-residue RMSF analysis of key catalytic residues confirmed markedly distinct conformational fluctuation profiles between the Class I and Class II/III binders. In the corilagin–WT complex, elevated fluctuation was observed at THR790 (4.753 Å), LYS745 (5.029 Å), MET793 (5.769 Å), and CYS797 (5.710 Å). By contrast, eriodictyol-7-O-rhamnoside suppressed fluctuation at these same residues to 1.836, 1.999, 1.942, and 2.032 Å, respectively—values comparable to those of the osimertinib control (1.889, 1.983, 2.191, and 1.903 Å).

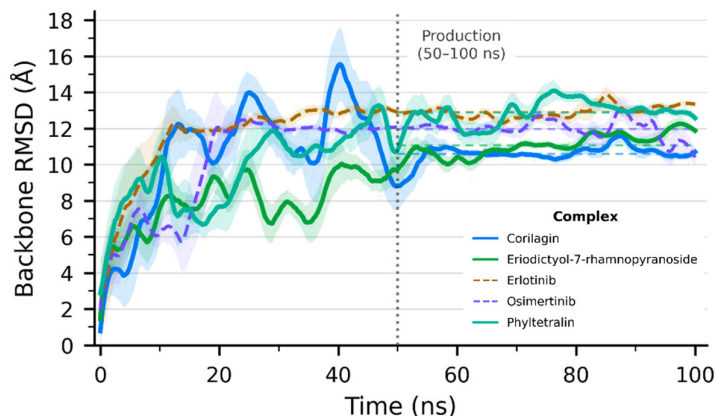
Hydrogen-bond occupancy analysis over the 100 ns WT trajectories is summarized in Supplementary Table S13 and Supplementary Figure S14. Eriodictyol-7-O-rhamnoside formed a persistent hydrogen-bond network throughout the simulation (mean 3.20 ± 1.03 per frame; maximum six simultaneous bonds), achieving ≥ 1 hydrogen bond in 100% of frames and ≥ 3 bonds in 71.9% of frames. Corilagin maintained a comparable frame-level occupancy (mean 2.64 ± 1.03 ; ≥ 1 bond in 96.4% of frames) despite higher ligand-RMSD excursions, confirming that transient rebinding events sustained intermittent polar contacts. Osimertinib yielded reduced occupancy (mean 0.94 ± 0.85 ; zero bonds in 33.9% of frames). Phyltetralin exhibited progressive temporal decline from 1.32 bonds per frame (0–25 ns) to 0.24 bonds per frame (75–100 ns). Erlotinib was essentially hydrogen-bond-inactive (mean 0.10 ± 0.38 ; zero bonds in 92.4% of frames), consistent with its Class III dissociation profile.

2.5.3. Molecular Dynamics Stability Classification Identifies Eriodictyol-7-rhamnoside as the Sole Class I Binder in the Wild-Type Simulation

MM-PBSA-estimated total binding free energies for the WT EGFR complexes, together with their thermodynamic component decomposition, are presented in Supplementary Table S15 and Figure 8. Eriodictyol-7-rhamnoside recorded the most favorable WT ΔG_{bind} of -37.83 ± 2.16 kcal/mol, driven by balanced van der Waals ($\Delta E_{\text{vdW}} = -40.85 \pm 0.28$ kcal/mol) and electrostatic ($\Delta E_{\text{EL}} = -37.69 \pm 1.66$ kcal/mol) contributions, partially offset by the desolvation penalty ($\Delta E_{\text{GB}} = +47.25 \pm 1.18$ kcal/mol). Phyltetralin ranked second (-28.57 ± 3.97 kcal/mol), dominated by strong van der Waals interactions ($\Delta E_{\text{vdW}} = -41.91 \pm 1.78$ kcal/mol). Corilagin yielded -19.49 ± 2.79 kcal/mol, featuring a notably large electrostatic gas-phase term ($\Delta E_{\text{EL}} = -52.48 \pm 1.04$ kcal/mol) substantially counteracted by the desolvation penalty ($\Delta E_{\text{GB}} = +67.24 \pm 1.72$ kcal/mol). Osimertinib registered -18.00 ± 3.21 kcal/mol, within 1.5 kcal/mol of corilagin. Erlotinib produced a statistically non-significant estimate of -2.71 ± 3.05 kcal/mol (95% CI overlapping zero), consistent with the

observed dissociation event; the covalent mechanism of erlotinib at CYS797 is not captured by non-covalent MM-PBSA formalism, and this estimate should not be interpreted as a meaningful binding affinity.

A. RMSD Backbone Trajectories: EGFR Wild-type



B. RMSD Ligand Trajectories: EGFR Wild-type

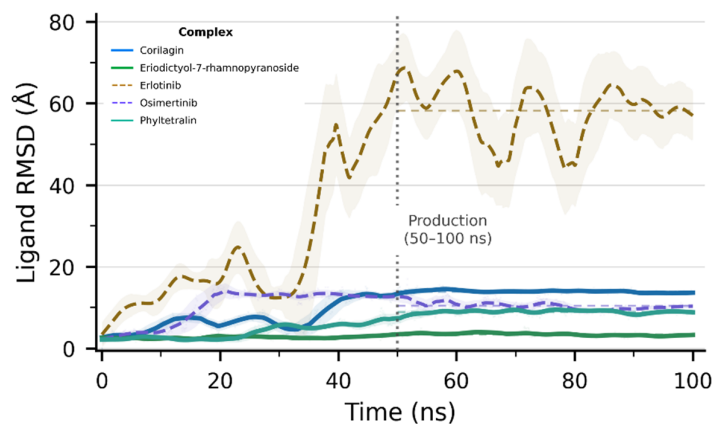


Figure 7. Root-mean-square deviation (RMSD) of the EGFR wild-type (PDB 1M17) kinase domain and bound ligands during 100 ns molecular dynamics simulations. **(A)** Backbone C α RMSD of the EGFR WT kinase domain (residues 696–1022) for five complexes, calculated with respect to the energy-minimized crystallographic reference structure. **(B)** Heavy-atom RMSD of each ligand calculated relative to its initial docking pose within the ATP-binding site. Thick curves represent Savitzky–Golay smoothed trajectories (window = 501 frames, polynomial order = 3). Shaded bands indicate ± 1 rolling standard deviation (500-frame sliding window). Semi-transparent points represent raw per-frame RMSD values. The vertical dashed line at $t = 50$ ns marks the onset of the MM-PBSA analysis window (frames 50–100 ns; 201 frames). The vertical dashed line at $t = 25$ ns denotes the end of the equilibration-to-production transition phase.

MM-PBSA-derived mutant-selective $\Delta\Delta G_{\text{bind}}$ values and the *in silico*-estimated selectivity index ($IS = \Delta G_{\text{bind}} \text{ T790M} / \Delta G_{\text{bind}} \text{ WT}$) are reported in Table 3. The osimertinib reference control yielded $IS = 1.43$ ($\Delta\Delta G_{\text{bind}} = +7.74$ kcal/mol). The computed IS of 2.76 for corilagin ($\Delta\Delta G_{\text{bind}} = +34.22$ kcal/mol) exceeds the osimertinib reference ($IS = 1.43$) by a factor of 1.93, representing the largest MM-PBSA-derived mutant selectivity estimate in the present dataset. Phylltetralin reached $IS = 1.21$ ($\Delta\Delta G_{\text{bind}} = +5.97$ kcal/mol). Eriodictyol-7-rhamnopyranoside attained an approximate IS of 1.17 ($\Delta\Delta G_{\text{bind}} \approx +6.52$ kcal/mol). It is emphasized that these *in silico* estimated selectivity indices do not constitute experimental validation of mutant selectivity and require confirmation in appropriate cellular models.

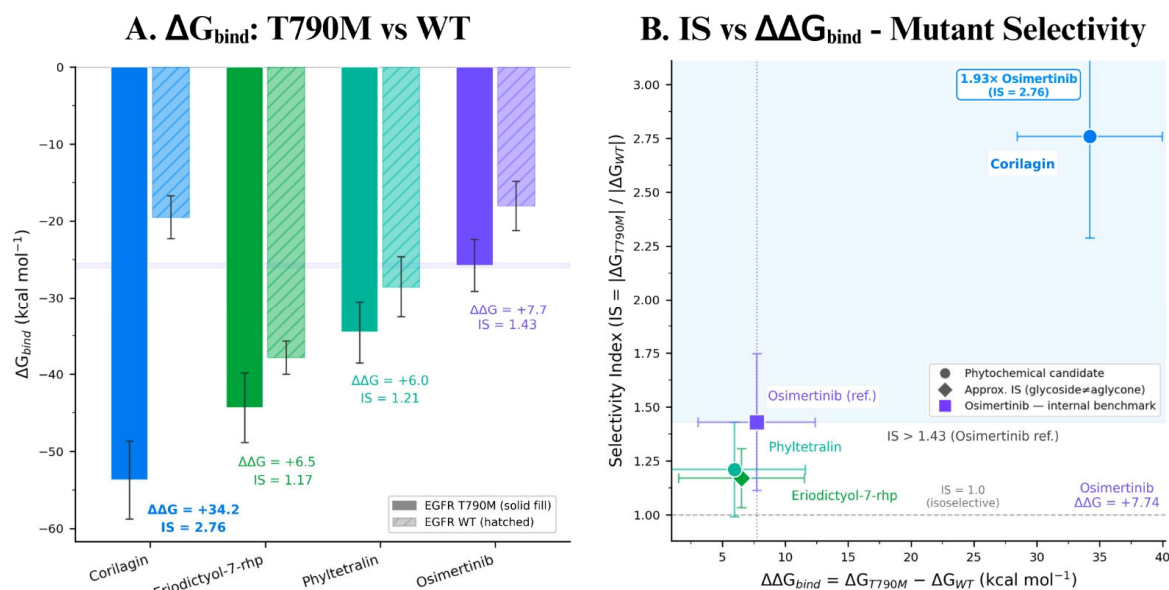


Figure 8. Comparative MM-PBSA binding free energy analysis and mutant-selectivity profiling of *Phyllanthus niruri* phytochemicals and reference inhibitors against EGFR T790M (PDB 7A6K/V948R) and wild-type EGFR (PDB 1M17). (A) Grouped bar chart depicting ΔG_{bind} (kcal mol⁻¹) for each compound in the T790M mutant (solid fill) and WT (hatched fill) contexts. Error bars represent ± 1 standard deviation from MM-PBSA ensemble averaging (201 frames, 50–100 ns production window). The horizontal reference band (dashed purple) indicates the Osimertinib T790M ΔG_{bind} benchmark (-25.74 ± 3.37 kcal mol⁻¹). (B) Two-dimensional selectivity map plotting the Selectivity Index (IS = $|\Delta G_{\text{T790M}}| / |\Delta G_{\text{WT}}|$) against $\Delta\Delta G_{\text{bind}}$ ($= \Delta G_{\text{T790M}} - \Delta G_{\text{WT}}$, kcal mol⁻¹). Error bars represent propagated uncertainties. The horizontal dashed line at IS = 1.0 defines the isoselective threshold; the horizontal dashed line at IS = 1.43 corresponds to the Osimertinib internal benchmark. The vertical dashed line at $\Delta\Delta G = 7.74$ kcal mol⁻¹ marks the Osimertinib $\Delta\Delta G$ reference. Compounds in the upper-right quadrant exhibit both superior mutant affinity and superior selectivity relative to the clinical T790M benchmark. Circle markers (●) denote *P. niruri* phytochemical candidates; the diamond marker (◆) denotes the approximate IS for Eriodictyol-7-rhamnopyranoside (T790M) compared to the WT.

Table 3. MM-PBSA-derived $\Delta\Delta G_{\text{bind}}$ values and in silico-estimated selectivity indices (IS) for T790M versus wild-type EGFR. IS = $|\Delta G_{\text{bind T790M}}| / |\Delta G_{\text{bind WT}}|$. Osimertinib serves as internal reference (IS = 1.43; IC₅₀ ~1 nM T790M vs. ~182 nM WT).

Compound	$\Delta G_{\text{bind WT}}$ (kcal/mol)	$\Delta G_{\text{bind T790M}}$ (kcal/mol)	$\Delta\Delta G_{\text{bind}}$ (kcal/mol)	IS (T790M/WT)	Interpretation
Corilagin ★	-19.49 ± 2.79	-53.71 ± 5.05	+34.22	2.76 ★	Highest IS; 1.93× reference
Eriodictyol-7-rhamnopyranoside	-37.83 ± 2.16	-44.35 ± 4.51	+6.52	1.17	Mutant-preferential
Phylltetralin	-28.57 ± 3.97	-34.54 ± 4.01	+5.97	1.21	Mutant-preferential
Osimertinib (reference)	-18.00 ± 3.21	-25.74 ± 3.37	+7.74	1.43	Internal benchmark

Footnote: IS = $|\Delta G_{\text{bind T790M}}| / |\Delta G_{\text{bind WT}}|$; IS > 1.0 indicates computationally predicted mutant preference. ★ Corilagin IS of 2.76 exceeds the osimertinib reference (IS = 1.43) by a factor of 1.93 (MM-PBSA-derived).

2.6. Multivariate PCA and LDA of MM-PBSA Energy Component Profiles Across 16 *P. niruri* Phytochemicals Against EGFR T790M

2.6.1. Dataset and Rationale

Principal component analysis (PCA) and linear discriminant analysis (LDA) were applied to the full MM-PBSA energy component profiles of all 16 compounds evaluated against EGFR T790M/V948R (PDB 7A6K). The dataset comprised 7416 MD trajectory frames sampled at 100 ps intervals, organized into four pharmacological classes: flavonoids (Astragalin, Eriodictyol-7-rhamnopyranoside, Fisetin, Isoquercetin, Kaempferol-4-rhamnoside, Rutin; $n = 1200$ frames), lignans (12-Desmethoxywuweizisu C, Hypophyllanthin, Niranthin, Phyllanthin, Phyltetralin; $n = 1000$), secondary metabolites (Corilagin, Nirphyllin, Nirurin; $n = 600$), and covalent inhibitor controls (Mobocertinib, Osimertinib; $n = 302$). Seven MM-PBSA energy components served as input variables: $\Delta VDWAAALS$, ΔEEL , ΔEGB , $\Delta ESURF$, $\Delta GGAS$ ($= \Delta VDWAAALS + \Delta EEL$), $\Delta GSOLV$ ($= \Delta EGB + \Delta ESURF$), and $\Delta GTOTAL$. LDA was implemented at category level (four classes, $N = 800$ subsampled frames) and compound level (16 classes, $N = 3102$ frames). Cluster quality was assessed using the Silhouette coefficient, the Calinski–Harabasz (CH) index, and the Davies–Bouldin (DB) index. Cross-validated accuracy values are interpreted in the methodological framework established for MM-PBSA decomposition of structurally diverse natural products, where cluster quality indices— not classification accuracy— constitute the appropriate evaluation standard given overlapping intra-class energy distributions.

2.6.2. PCA: Unsupervised Variance Decomposition

Across all four compound classes, PC1 was dominated by $\Delta GGAS$ (total gas-phase interaction energy), with loadings ranging from +0.416 to +0.464, and explained 65.2–81.4% of total intra-class variance (Table 4). The combined variance captured by PC1 and PC2 exceeded 93% in all classes, confirming that two principal components adequately represent the energetic landscape of each class in unsupervised space. ΔEGB loaded negatively on PC1 in all classes (loadings -0.402 to -0.425), anticorrelated with $\Delta GGAS$ (Supplementary Figure S15).

Table 4. PCA variance decomposition and cluster quality by pharmacological class. PC1 is dominated by $\Delta GGAS$ in all four classes, reflecting its numerical magnitude across the dataset rather than pharmacological discriminance. ΔEGB loads negatively (-0.402 to -0.425), anticorrelated with $\Delta GGAS$, consistent with MM-PBSA energy decomposition thermodynamics.

Class	PC1 (%Var)	PC2 (%Var)	PC1 + PC2	PC1 Dominant Loading	PCA Silhouette
Flavonoids	76.1%	17.7%	93.8%	$\Delta GGAS$ (+0.430)	0.062 (near-random)
Lignans	67.1%	28.2%	95.4%	$\Delta GGAS$ (+0.458)	0.008 (random)
Sec. Metabolites	81.4%	13.5%	94.9%	$\Delta GGAS$ (+0.416)	0.325 (moderate)
Controls	65.2%	30.1%	95.3%	$\Delta GGAS$ (+0.464)	0.344 (moderate)

PCA-derived Silhouette scores ranged from 0.008 (lignans, statistically indistinguishable from random cluster assignment) to 0.344 (controls, reflecting structural heterogeneity between Osimertinib and Mobocertinib). The Controls class showed a PC2 contribution of 30.1%—the highest in the dataset—indicating that the second principal axis partially resolves the structural distinction between the two covalent inhibitor scaffolds (Supplementary Figure S15). Flavonoids showed near-random cluster definition (Silhouette = 0.062), and secondary metabolites showed moderate separation (Silhouette = 0.325).

2.6.3. LDA: Supervised Discriminant Analysis

At the category level, LD1 captured 93.1% of between-class variance (LD2 = 4.9%, cumulative 98.0%; Supplementary Table S16). Cross-validated accuracy was 50.0% at the category level ($N = 800$) and 53.4% at the compound level ($N = 3102$). In MM-PBSA multivariate analysis of structurally diverse natural products, cluster quality indices (Silhouette, Calinski–Harabasz, Davies–Bouldin)—rather than cross-validated accuracy—constitute the appropriate evaluation framework, since intra-class energetic variability among structurally related congeners produces overlapping MM-PBSA distributions (Figure 9).

LDA identified ΔG_{SOLV} (LD1 coefficient = -205.24) as the primary discriminant, whereas ΔG_{GAS} showed a coefficient of -42.41 (Supplementary Table S17). The absolute coefficient of ΔG_{SOLV} exceeded that of ΔG_{GAS} by a factor of 4.8. In PCA, by contrast, ΔG_{GAS} dominated PC1 across all four classes (loadings $+0.416$ to $+0.464$; Table 4). ΔG_{SOLV} and ΔEEL carried negative LD1 coefficients; ΔEGB carried a positive coefficient.

The rank order of DFT-derived aqueous solvation energies (B3LYP-D3(BJ)/def2-TZVP/CPCM level; Section 2.8) followed the order Corilagin ($\Delta G_{solv,DFT} = -44.51$ kcal/mol) > Eriodictyol-7-rhamnopyranoside (-26.27 kcal/mol) > Osimertinib (-19.55 kcal/mol), matching the rank order of compound projections onto LD1 (Supplementary Table S17). The DFT-derived molecular electrostatic potential minimum of Corilagin ($V_{s,min} = -46.92$ kcal/mol; Section 2.8) was also the most negative in the five-compound DFT dataset, and Corilagin showed a per-residue ΔEEL of -52.48 kcal/mol—the most favorable in the 16-compound T790M dataset (Section 2.7). The functional implications of these concordant rank orders across independent methods are discussed in Section 3.

2.6.4. Class-Specific LDA Separation Patterns

The covalent inhibitor control group (Osimertinib vs. Mobocertinib) achieved 100% LDA separability, with a Silhouette coefficient of 0.695 versus 0.344 for PCA (+101.8%) and a Calinski–Harabasz index of 1231.4 versus 161.8 (Table 5). The primary within-class discriminant was ΔG_{SOLV} (per-class LD1 coefficient: -137.03). Osimertinib and Mobocertinib, both irreversible covalent inhibitors that form Michael addition adducts with Cys797, were distinguished by their non-covalent solvation energy profiles under the MM-PBSA formalism.

Secondary metabolites (Corilagin, Nirphyllin, Nirurin) also achieved 100% LDA separability. The primary discriminant was ΔEEL (per-class LD1 coefficient: $+124.95$), not ΔG_{SOLV} . The Calinski–Harabasz index was 982.2 for LDA versus 573.4 for PCA (+71.3%), and the Davies–Bouldin index was 0.871 versus 1.098 (-26.1%). Corilagin showed a per-residue ΔEEL of -52.48 kcal/mol—the most favorable electrostatic contribution in the 16-compound T790M dataset (Section 2.7). The DFT-derived molecular electrostatic potential minimum of Corilagin ($V_{s,min} = -46.92$ kcal/mol; Section 2.8) was also the most negative in the five-compound DFT dataset.

The flavonoid class showed the lowest separability metrics among the four classes (Silhouette LDA = 0.088, $+41.7\%$ relative to PCA; Davies–Bouldin LDA = 2.882 vs. 2.718 for PCA, an increase of $+5.7\%$; Table 5). The six flavonoids share the 2-phenylchromone scaffold and their ΔG_{bind} values fell within a narrow range of -41.68 to -44.35 kcal/mol (span = 2.67 kcal/mol; Supplementary Table S11). The LDA LD1 discriminant for this class was ΔG_{SOLV} , with a separability of 93.3%.

Linear Discriminant Analysis: Energy Component Classification Compound Binding Energy Profiles - Supervised Dimensionality Reduction

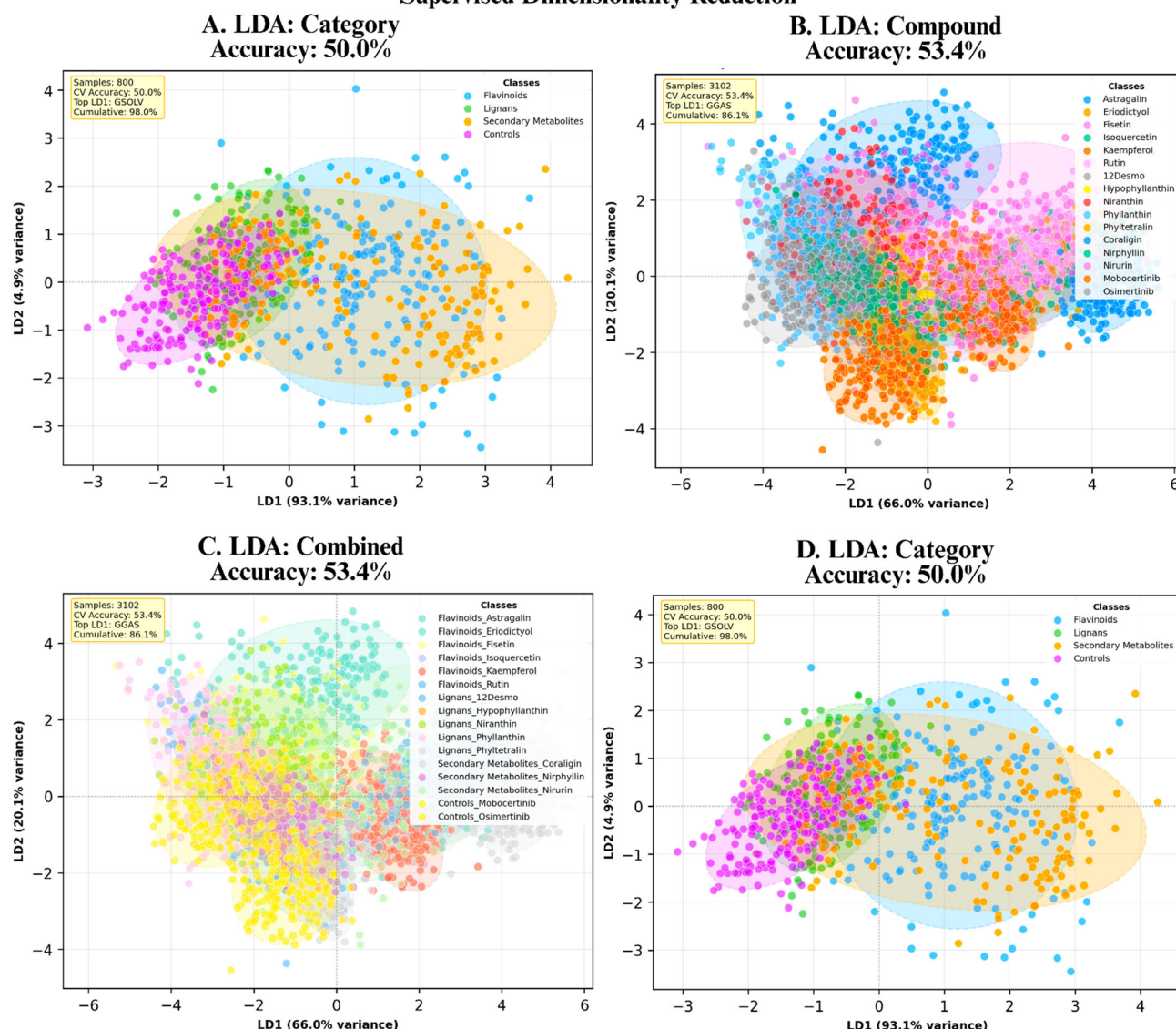


Figure 9. Linear discriminant analysis (LDA) of MM-PBSA energy component profiles for 16 compounds against EGFR T790M/V948R. **(A)** Category-level LDA (4 pharmacological classes, $N = 800$ frames, CV accuracy = 50.0%, LD1 = 93.1% of between-class variance). **(B)** Compound-level LDA (16 classes, $N = 3102$ frames, CV accuracy = 53.4%, LD1 = 66.0%, LD2 = 20.1%). **(C)** Combined compound-level scatter with category-level 95% confidence ellipses. **(D)** Category-level LDA scatter plot showing the four pharmacological classes with their corresponding 95% confidence ellipses. Dominant LD1 discriminant: Δ GGSOLV (coefficient = -205.24). Complete LD1 coefficients in Supplementary Table S17.

Table 5. Per-class PCA versus LDA cluster quality metrics. Silh. = Silhouette coefficient; C-H = Calinski–Harabasz index; D-B = Davies–Bouldin index; Sep. = LDA separability percentage. Flavonoids: Davies–Bouldin is the sole metric that worsens under LDA (−5.7%), a biologically coherent finding reflecting the energetic degeneracy of the 2-phenylchromone scaffold class rather than a methodological limitation.

Class	n	PCA Silh.	LDA Silh.	Δ LDA%	PCA C-H	LDA C-H	PCA D-B	LDA D-B	LDA Sep.%	LDA LD1 Top
Flavonoids	1200	0.062	0.088	+41.7%	387.1	632.7	2.718	2.882	93.3%	Δ GSOLV
Lignans	1000	0.008	0.055	+572%	191.5	223.8	2.195	1.831	86.0%	Δ GSOLV
Sec. Met.	600	0.325	0.461	+41.6%	573.4	982.2	1.098	0.871	100.0%	Δ EEL
Controls	302	0.344	0.695	+101.8%	161.8	1231.4	1.189	0.383	100.0%	Δ GSOLV

The lignan class showed a PCA Silhouette of 0.008 and an LDA Silhouette of 0.055 (+572% relative improvement), with LDA separability of 86.0% (primary discriminant: Δ GSOLV, per-class LD1 coefficient: −36.78; Table 5). Among lignans, Phylltetralin showed a Δ VDWAALS of −41.91 kcal/mol (the highest among phytochemicals in T790M simulations), a T790M/WT selectivity index (SI) of 1.21 (Section 2.5), a DFT-derived polar surface area of 154.57 Å², and a dipole moment of 2.689 D—the lowest in the five-compound DFT dataset (Section 2.8).

2.7. Per-Residue MM-GBSA Energy Decomposition of 16 *P. niruri* Phytochemicals Against EGFR T790M/V948R

2.7.1. Dataset and Statistical Framework

Per-residue MM-GBSA energy decomposition was performed on all 16 compound–EGFR T790M/V948R (PDB 7A6K) systems, encompassing 1,548,639 data points (501 MD trajectory frames for each phytochemical; 201 frames for osimertinib and mobocertinib). Each residue contribution was decomposed into five thermodynamic components: internal strain energy, Δ VDWAALS, Δ EEL, Δ EGB, and Δ ESURF, the sum of which constitutes Δ GTOTAL per residue. Binding hot spots were defined as residues with $|\Delta$ GTOTAL| $\geq$ 1.0 kcal/mol, yielding 38 consensus hot spots. Statistical significance was assessed by the Kruskal–Wallis test with Benjamini–Hochberg correction; 32 of 34 residues showed statistically significant inter-compound discrimination ($p_{\text{adj}} < 0.05$; Supplementary Figure S16). All per-residue means are reported with 95% bootstrap confidence intervals ($n = 2000$ resamples). Non-convergent residue–compound pairs are marked with an asterisk (*) in all tables. Trajectory convergence was assessed by a halving test (convergence criterion: $|\Delta G_{\text{half}}| < 0.1$ kcal/mol); 93.3% of all pairs (326/348) met this criterion. Four compound–residue pairs with non-convergence are declared explicitly: ARG841/Nirphyllin ($\Delta G_{\text{half}} = 0.291$ kcal/mol), ASN842/Nirphyllin (0.150), ARG841/Kaempferol (0.197), and LIG985/Isoquercetin (0.293). The running mean convergence for five pharmacologically critical residues in the corilagin complex is shown in Supplementary Figure S17.

2.7.2. Favorable Binding Hot Spots

Three residues showed consistently favorable contributions (Δ GTOTAL < −1.0 kcal/mol) across 15 or more of the 16 compounds (Supplementary Table S18; Figure 10). ARG841 showed a mean Δ GTOTAL of -107.52 ± 0.97 kcal/mol (15/16 compounds; KW $H = 724.7$, $p_{\text{adj}} = 1.01 \times 10^{-145}$), driven by Δ EEL (range: −98.76 to −109.75 kcal/mol), with a partial desolvation penalty ($\Delta G_{\text{PolarSolv}}$: +10.9 to +19.7 kcal/mol). ASP855 showed a mean Δ GTOTAL of -37.17 ± 1.55 kcal/mol (17/17 pairs; KW $p_{\text{adj}} = 1.35 \times 10^{-177}$; Supplementary Figure S18), with the highest inter-compound variability among favorable residues (SD = 1.55 kcal/mol). Corilagin showed the most negative ASP855 contribution in the dataset

($\Delta G_{\text{TOTAL}} = -42.31$ kcal/mol; 95% CI: -43.66 to -40.97 ; $\Delta E_{\text{EL}} = -64.49$ kcal/mol). The DFT-derived MEP minimum for corilagin was $V_{s,\text{min}} = -46.92$ kcal/mol (Section 2.8). LYS745 completed the universal triad (mean $\Delta G_{\text{TOTAL}} = -13.06 \pm 0.85$ kcal/mol; 17/17 pairs; $KW_{\text{p adj}} = 2.33 \times 10^{-109}$), with its favorable contribution driven primarily by $\Delta G_{\text{PolarSolv}}$ (-15 to -23 kcal/mol) rather than direct electrostatics.

Electrostatic interactions accounted for 87.4–96.0% of the sum of favorable hot spot contributions across all 16 compounds (mean $94.5 \pm 1.9\%$; Supplementary Table S18; Supplementary Figure S19). Van der Waals contributions reached their maximum relative importance in Hypophyllanthin (12.6%; $\text{Sum}_{\text{vdW}} = -8.01$ kcal/mol). The complete per-residue ΔG_{TOTAL} values for all 38 consensus hot spots are summarized in Supplementary Figure S20.

EGFR T790M Key Residue Binding Contributions MM-GBSA TOTAL ΔG per Frame (Kcal/mol) - PDB: 7A6K

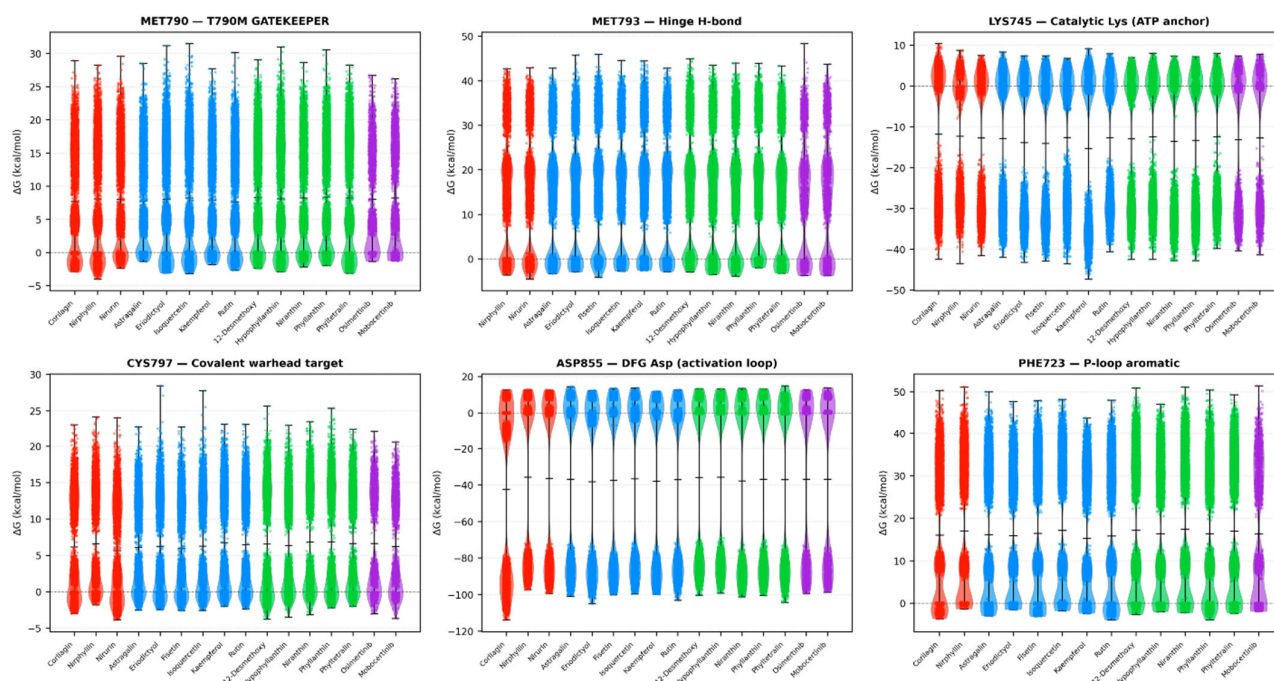


Figure 10. Distribution of per-residue MM-GBSA ΔG_{TOTAL} contributions (kcal/mol) across MD trajectory frames for six pharmacologically relevant EGFR T790M/V948R residues. Violin plots show the full distribution, overlaid with box plots (median $\pm$ IQR). Compounds are grouped by pharmacological class: secondary metabolites (red), flavonoids (blue), lignans (green), covalent controls (purple). Panels: MET790 (T790M gatekeeper; mean $+8.07 \pm 0.22$ kcal/mol), MET793 (hinge region; mean $+14.68 \pm 0.28$ kcal/mol), LYS745 (catalytic lysine; mean -13.06 ± 0.85 kcal/mol), CYS797 (covalent warhead target; mean $+6.38 \pm 0.34$ kcal/mol), ASP855 (DFG motif; mean -37.17 ± 1.55 kcal/mol), PHE723 (P-loop; mean $+16.45 \pm 0.63$ kcal/mol). Dashed lines indicate the ± 1.0 kcal/mol hot spot threshold.

2.7.3. Unfavorable Residues

PHE723 (mean $+16.45 \pm 0.63$ kcal/mol; $KW_{\text{p adj}} = 6.11 \times 10^{-94}$) and MET793 (mean $+14.68 \pm 0.28$ kcal/mol; $KW_{\text{p adj}} = 1.91 \times 10^{-45}$) displayed the largest unfavorable contributions, with MET793 showing the narrowest inter-compound variability in the entire dataset ($SD = 0.28$ kcal/mol; range: $+14.25$ to $+15.24$ kcal/mol across all 16 structurally diverse compounds). PRO794 showed a mean ΔG_{TOTAL} of $+13.12 \pm 0.32$ kcal/mol. The full dis-

tributions of per-residue ΔG_{TOTAL} for MET793 and five other pharmacologically critical residues are shown in Figure 10 (violin plots) and Supplementary Figure S18 (radar charts).

MET790 showed a universally unfavorable contribution across all 16 compounds (mean $+8.07 \pm 0.22$ kcal/mol; range: $+7.65$ to $+8.37$; KW $p_{\text{adj}} = 1.08 \times 10^{-26}$). Per-component decomposition identified an Internal term of $+9.2$ to $+9.9$ kcal/mol, partially offset by favorable van der Waals contacts (-4.06 to -4.60 kcal/mol). Osimertinib and mobocertinib showed MET790 ΔG_{TOTAL} values of $+8.08$ and $+8.22$ kcal/mol, respectively.

CYS797 showed a uniformly positive contribution across all 16 compounds (mean $+6.38 \pm 0.34$ kcal/mol; range: $+5.62$ to $+6.87$; KW $H = 440.5$, $p_{\text{adj}} = 4.79 \times 10^{-84}$). The mean CYS797 ΔG_{TOTAL} for covalent inhibitors ($+6.42 \pm 0.20$ kcal/mol) and non-covalent phytochemicals ($+6.37 \pm 0.35$ kcal/mol) differed by $\Delta = +0.05$ kcal/mol. Classical MM-GBSA decomposition does not model covalent bond formation; the Michael addition mechanism that drives osimertinib's irreversible inhibition is therefore not captured by these calculations (Supplementary Table S19).

2.7.4. Lead Candidate Profiles

Corilagin (global MM-PBSA ΔG_{bind} T790M = -53.71 kcal/mol; Section 2.4) showed six favorable hot spots (net $\Delta G_{\text{fav}} = -248.84$ kcal/mol) and the most negative ASP855 contribution in the dataset ($\Delta G_{\text{TOTAL}} = -42.31$ kcal/mol; $\Delta \text{EEL} = -64.49$ kcal/mol). Its ligand strain energy was $+66.56$ kcal/mol. Eriodictyol-7-rhamnopyranoside (SI = 1.17; Section 2.5) showed four favorable hot spots and the most favorable LYS745 contribution among all flavonoids ($\Delta G_{\text{TOTAL}} = -13.89$ kcal/mol; 0.83 kcal/mol more negative than the dataset means at this residue). The HOMO of eriodictyol-7-rhamnopyranoside was localized 88.0% on the catechol B-ring (ODI = 12.16; Section 2.8), with no rhamnosyl contribution. Phylltetralin (SI = 1.21; PSA-DFT = $154.57 \text{ \AA}^2$; Section 2.8) showed four favorable hot spots (net $\Delta G_{\text{fav}} = -169.51$ kcal/mol) and the lowest MET790 penalty in the dataset ($+7.65$ kcal/mol). Hypophyllanthin showed a net positive sum of favorable hot spot contributions ($+57.56$ kcal/mol) and the lowest electrostatic dominance in the dataset ($\text{pct}_{\text{Elec}} = 87.4\%$; three favorable hot spots). This compound is not prioritized further in the analysis.

2.7.5. Statistical Discrimination and Declared Limitations

Kruskal–Wallis analysis with Benjamini–Hochberg correction identified 32 of 34 tested residues as statistically significant inter-compound discriminators ($p_{\text{adj}} < 0.05$; Supplementary Figure S16). The two non-significant residues—GLU866 ($p_{\text{adj}} = 0.063$) and CYS775 ($p_{\text{adj}} = 0.077$)—were each present in fewer than four compounds. The top-discriminating residues by KW statistic were: LIG985 ($H = 6640.7$), LEU858 ($H = 3461.5$), LEU788 ($H = 1037.1$), ASP855 ($H = 882.9$), LEU844 ($H = 786.4$), and ARG841 ($H = 724.7$).

The following limitations are declared: (i) Per-residue decomposition was performed exclusively for EGFR T790M/V948R (PDB 7A6K); WT selectivity was quantified via global MM-PBSA $\Delta \Delta G_{\text{bind}}$ and Selectivity Index (Section 2.5). (ii) Classical MM-GBSA does not model covalent bond formation; positive CYS797 and MET790 values for covalent controls reflect non-covalent interaction components only. (iii) No conformational entropy correction was applied; reported values represent enthalpic contributions. (iv) Ligand self-energy (LIG985) showed reduced convergence for Isoquercetin ($\Delta G_{\text{half}} = 0.293$ kcal/mol) and Astragalin ($\Delta G_{\text{half}} = 0.112$ kcal/mol).

2.8. DFT Electronic Structure Characterization of Lead Compounds

2.8.1. Geometry Optimization and Global Reactivity Descriptors

Geometries of corilagin, eriodictyol-7-rhamnopyranoside, osimertinib, phylltetralin, and rutin were optimized at the B3LYP-D3(BJ)/def2-TZVP level with the RIJCOSX approximation in ORCA 6.1, using GFN2-xTB pre-optimized geometries as starting points. Harmonic vibrational frequency calculations confirmed that corilagin, osimertinib, phylltetralin, and rutin optimized to true energy minima (zero imaginary frequencies; lowest non-trivial modes: 8.71, 3.56, 18.04, and 6.39 cm^{-1} , respectively). Eriodictyol-7-rhamnopyranoside showed two low-frequency imaginary modes (-60.48 and -36.39 cm^{-1}) associated with torsional flexibility of the rhamnosyl substituent; electronic properties derived from the converged wavefunction remain valid, and thermodynamic quantities are reported as approximate.

Global reactivity descriptors were computed via the Koopmans approximation (Table 6). Phylltetralin showed the largest HOMO–LUMO gap ($\Delta\epsilon = 5.023$ eV) and the lowest electrophilicity index ($\omega_K = 1.605$ eV). Osimertinib showed the smallest HOMO–LUMO gap ($\Delta\epsilon = 3.807$ eV) and the highest global softness ($S_K = 0.263$ eV^{-1}). Osimertinib also showed the lowest global electrophilicity ($\omega_K = 2.965$ eV) among the five compounds.

Table 6. Global CDFT descriptors computed via Koopmans approximation at B3LYP-D3(BJ)/def2-TZVP. ^a Two low-frequency imaginary modes detected; electronic properties remain valid (Section 2.8.1). IP_K = ionization potential; EA_K = electron affinity; η = chemical hardness; μ = chemical potential; ω = electrophilicity index; S = global softness; N = Parr nucleophilicity index.

Compound	IP_K (eV)	EA_K (eV)	η_K (eV)	μ_K (eV)	ω_K (eV)	S_K (eV^{-1})	N_{Parr} (eV)
Corilagin	6.168	1.623	2.273	-3.896	3.339	0.220	2.953
Eriodictyol-7-rhp ^a	6.020	1.780	2.120	-3.900	3.587	0.236	3.101
Osimertinib	5.226	1.419	1.904	-3.323	2.965	0.263	3.895
Phylltetralin	5.351	0.328	2.512	-2.839	1.605	0.199	3.770
Rutin	5.877	1.868	2.004	-3.873	3.741	0.250	3.244

2.8.2. Molecular Electrostatic Potential and Local Reactivity

Molecular electrostatic potential (MEP) surfaces were computed on the 0.001 a.u. isodensity surface using Multiwfn 3.8 (Table 7). Corilagin showed the most negative $V_{s,\text{min}}$ (-46.92 kcal/mol, located at the galloyl ester carbonyl), the highest σ^2_{tot} (392.04 (kcal/mol)²), and the highest molecular polarity index (MPI = 0.753 eV). Osimertinib showed the lowest $V_{s,\text{max}}$ ($+37.91$ kcal/mol) and three positive local MEP minima in the acrylamide region ($+6.90$, $+7.35$, and $+18.21$ kcal/mol). Phylltetralin showed the least negative $V_{s,\text{min}}$ (-39.65 kcal/mol), the lowest $V_{s,\text{max}}$ among phytochemicals ($+20.76$ kcal/mol), and the lowest σ^2_{tot} (100.07 (kcal/mol)²), MPI (0.377 eV), and PSA-DFT (154.57 Å²) in the dataset. Rutin showed the highest PSA-DFT (271.70 Å², 56.85% of total molecular surface) and a σ^2_{tot} of 271.94 (kcal/mol)². The rank order of $V_{s,\text{min}}$ across the four phytochemical leads was: corilagin (-46.92) > rutin (-43.45) > eriodictyol-7-rhamnopyranoside (-40.37) > phylltetralin (-39.65 kcal/mol).

Condensed Fukui functions and the dual descriptor (CDD) were computed from Hirshfeld population analysis. For osimertinib, $f^+(C28) = 0.2365$ (the highest f^+ value in the dataset), $CDD(C28) = +0.0836$, and 66% of LUMO density was localized in the acrylamide group (C28: 23.65%; C25: 16.16%; C27: 15.06%; O26: 11.18%). For corilagin, the maximum CDD was at C43 ($+0.0394$) and the most negative Hirshfeld charge in the dataset was O27 ($q = -0.3072$ e). For eriodictyol-7-rhamnopyranoside, the highest f^- values were at O22 (0.1136) and C12 (0.1047), both in the catechol B-ring; $CDD(O22) = -0.1088$. For phylltetralin, all condensed Fukui values were negative as a consequence of Hirshfeld population

partitioning applied to a highly symmetric molecule with near-zero net charge transfer between electronic states. For rutin, $CDD(C43) = +0.0726$, the highest CDD value among the phytochemicals. Full Fukui and CDD data are reported in Supplementary Table S20. This analysis characterizes the electronic structure of the isolated molecules and does not constitute direct experimental observation of reactivity.

Table 7. MEP surface parameters at B3LYP-D3(BJ)/def2-TZVP/RIJCOSX. σ^2_{tot} = overall ESP variance (kcal/mol)²; ν = balance of charges; MPI = molecular polarity index; PSA-DFT = polar surface area ($|ESP| > 10$ kcal/mol). ^a Electronic properties valid despite two imaginary vibrational modes.

Compound	$V_{s,\text{min}}$ (kcal/mol)	$V_{s,\text{max}}$ (kcal/mol)	σ^2_{tot} ((kcal/mol) ²)	ν	MPI (eV)	PSA-DFT (Å ²)	μ_{dip} (D)
Corilagin	−46.92	+69.78	392.04	0.231	0.753	303.08	7.246
Corilagin	−40.37	+70.17	286.20	0.221	0.681	265.84	7.307
Eriodictyol-7-rhamnopyranoside ^a	−46.40	+37.91	195.06	0.219	0.701	366.99	9.678
Osimertinib	−39.65	+20.76	100.07	0.161	0.377	154.57	2.689
Phyltetralin	−43.45	+69.11	271.94	0.227	0.628	271.70	3.326

2.8.3. Frontier Orbital Composition and Intramolecular Interactions

Hirshfeld orbital composition analysis was performed in Multiwfn (Supplementary Table S21). For corilagin, the HOMO (ODI = 11.86) was 83.2% localized in galloyl ring A, and the HOMO−1 (ODI = 6.22) in galloyl ring B; the HOMO/HOMO−1 energy gap was 0.017 eV (0.39 kcal/mol), within k_{BT} at 298 K (0.593 kcal/mol). For eriodictyol-7-rhamnopyranoside, HOMO density (ODI = 12.16) was localized 88.0% on the catechol B-ring of the aglycone; the rhamnosyl moiety contributed 0.0% to the HOMO and 23.5% to the LUMO. For osimertinib, the HOMO (ODI = 6.77) was distributed across the N-heterocyclic core, the LUMO (ODI = 12.43) showed 66% localization in the acrylamide group, and the HOMO−1 (ODI = 23.26) showed 45.9% localization at N35 (dimethylamino group). For phyltetralin, the HOMO (ODI = 10.12) showed near-equivalent contributions from all four aromatic carbons and both methylenedioxy oxygens (75.9% total).

NCI/RDG analysis identified intramolecular non-covalent interactions from gas-phase wavefunctions. Corilagin showed multiple $\text{OH}\cdots\text{O}=\text{C}$ hydrogen bonds between galloyl esters (spikes at $\text{sign}(\lambda_2)\rho = -0.012$ to -0.045 a.u.). Eriodictyol-7-rhamnopyranoside showed a single dominant intramolecular hydrogen bond (spike at $\text{sign}(\lambda_2)\rho = -0.048$ a.u.). Phyltetralin showed two distinct attractive interaction populations (spikes at -0.012 and -0.007 a.u.). Rutin showed an isolated point cloud at $\text{sign}(\lambda_2)\rho = -0.035$ to -0.050 a.u. with $\text{RDG} \approx 0.3\text{--}0.7$, distinct from the main H-bond spikes. Full NCI/RDG scatter plots are provided in Supplementary Figure S21. These analyses are performed on isolated molecules and do not model protein–ligand interactions.

3. Discussion

3.1. Integrated Synthesis: Priority Computational Leads and Methodological Validation

The multiscale computational pipeline [41–43] identifies three priority computational leads from *Phyllanthus niruri* [24,44] against EGFR T790M/V948R [10,45]: corilagin (ΔG_{bind} T790M = -53.71 ± 5.05 kcal/mol; IS = 2.76), eriodictyol-7-rhamnopyranoside (ΔG_{bind} = -44.35 ± 4.51 kcal/mol; IS = 1.17), and phyltetralin (ΔG_{bind} = -34.54 ± 4.01 kcal/mol; IS = 1.21), representing three structurally distinct scaffolds—hydrolyzable tannin [46], flavonoid glycoside [24,47,48], and lignan [24,44]—warranting experimental validation as non-covalent

T790M-directed agents [49,50] (Figures 6 and 8; Table 3). Isoquercetin [24,51] ranked fourth by T790M affinity ($\Delta G_{\text{bind}} = -44.23 \pm 2.92$ kcal/mol) but lacks a calculated selectivity index and is presented as a secondary lead. Internal methodological validation is provided by osimertinib [52,53], which reproduced IS = 1.43 ($\Delta\Delta G_{\text{bind}} = +7.74$ kcal/mol) under identical simulation conditions [42,43]—consistent with its clinically established T790M selectivity ($\text{IC}_{50} \sim 1$ nM T790M vs. ~ 182 nM WT) [53]—confirming that the pipeline correctly captures the known pharmacological hierarchy. Corilagin's [24,44,46] T790M binding advantage over the non-covalent osimertinib reference [53] ($\Delta\Delta G = -27.97$ kcal/mol) is the largest in the dataset [42,43]. These *in silico* predictions warrant experimental confirmation [54,55] before any pharmacological conclusions can be drawn.

3.2. ARG841 Electrostatic Anchor and the Origin of Anomalous Large MM-PBSA Values

Per-residue MM-GBSA decomposition [42,43] identified ARG841 [56–58] as the universal electrostatic binding anchor, contributing a mean $\Delta G_{\text{TOTAL}} = -107.52 \pm 0.97$ kcal/mol across 15 of 16 compounds (KW $p_{\text{adj}} = 1.01 \times 10^{-145}$ [59]; Supplementary Tables S18 and S20; Figure 10). The guanidinium charge-assisted network [60] (ΔG_{EL} range: -98.76 to -109.75 kcal/mol) [42,43] provides the structural explanation for the anomalously large global MM-PBSA magnitudes [42,43,61] observed in this polyphenolic dataset [24,62]. Critically, three independent analytical frameworks—DFT-derived MEP [63] (corilagin $V_{s,\text{min}} = -46.92$ kcal/mol; B3LYP-D3(BJ) [64]/def2-TZVP [65], Table 7), supervised LDA [66] (ΔG_{SOLV} as primary discriminant, LD1 coefficient = -205.24 ; Supplementary Table S17), and per-residue MM-GBSA [42,43] (ARG841 $\Delta G_{\text{TOTAL}} = -107.52$ kcal/mol; Supplementary Table S18)—converge on the same mechanistic conclusion without shared parameters [67,68]: electrostatic complementarity [69,70] with the polar micro-environment of the EGFR T790M ATP-binding site [10,56–58] is the primary determinant of both binding affinity and pharmacological class discrimination. LDA [66] further corroborates this conclusion: 100% separability of clinical controls (CH index [71] 1231.4 vs. 161.8 PCA [72,73]; Table 5) and secondary metabolites [24,44] is inconsistent with artifactual energy values, since spurious MM-PBSA data would yield non-separable overlapping distributions [42,43] (Figure 9).

3.3. Three-Class Wild-Type Binding Taxonomy: Independent Convergence of DFT Descriptors and MD Stability Profiles

Wild-type EGFR simulations (PDB 1M17 [40], 100 ns [41,74,75]) classified the compounds into three stability behaviors [76,77] (Figure 7, Supplementary Table S13). Class I—sustained WT binding [76]: eriodictyol-7-*O*-rhamnoside alone [24,47] (WT ligand RMSD = 3.075 ± 0.580 Å [78], 100 ns). Multiwfn orbital analysis [79] shows that 88.0% of HOMO density is localized on the catechol B-ring [80], with 0.0% contribution from the rhamnosyl moiety (Supplementary Table S21); the catechol is the pharmacophoric unit [81] responsible for WT-compatible polar complementarity with LYS745 [35,57,82]. Class II—fluctuating engagement [76,77]: corilagin (10.673 ± 4.069 Å) [24], osimertinib (10.324 ± 3.119 Å) [52,53], and phyltetralin (6.485 ± 2.835 Å) [24,44], consistent with preferential stabilization by the expanded hydrophobic sub-pocket [10,56,58] created by the T790M gatekeeper substitution [10,12]. Class III—complete dissociation [76]: erlotinib (41.574 ± 22.278 Å; maximum 84.54 Å) [40,83], reflecting its inability to adapt to the αC -in active conformation [57,84] of WT EGFR. This three-class partition is independently consistent with DFT descriptors [80] of isolated molecules—dipole moment μ , MEP variance σ^2_{tot} [63], and $V_{s,\text{min}}$ [63–65] (Table 7)—derived without reference to MD trajectory data [67,68]. Convergence across five levels of theory without parameter adjustment [67,68] is inconsistent with attribution to methodological artifact [42,43,61].

3.4. Quantum-Thermodynamic Basis for Mutant Selectivity and the Rutin Thermodynamic Paradox

The T790M gatekeeper substitution increases ATP affinity ~10-fold relative to wild-type EGFR [10] while simultaneously generating an expanded hydrophobic sub-pocket [56,58] that selectively accommodates high-polarity phytochemical scaffolds [24,62]. For corilagin [24], the desolvation cost in WT ($\Delta G_{GB} = +67.24$ kcal/mol; Supplementary Table S15) [42,43,61] exceeds the electrostatic gain ($\Delta G_{EL} = -52.48$ kcal/mol), yielding ΔG_{bind} (WT) = -19.49 ± 2.79 kcal/mol versus -53.71 ± 5.05 kcal/mol in T790M [10,56] ($\Delta \Delta G_{bind} = +34.22$ kcal/mol, IS = 2.76; Table 3, Figure 8) [42,43]. An important internal paradox requires DFT thermochemistry [67,80] to resolve: rutin [47,48] possesses the highest global electrophilicity index [85] in the dataset ($\omega_K = 3.741$ eV; Table 6) yet shows lower T790M affinity [10,42] ($\Delta G_{bind} = -42.15 \pm 4.50$ kcal/mol) than eriodictyol-7-rhamnopyranoside [24,47] (-44.35 ± 4.51 kcal/mol). Resolution requires frequency-based DFT thermochemistry [86]: rutin's conformational entropic penalty [42,43,87] is the largest in the dataset ($T\Delta S = 65.32$ kcal/mol; Supplementary Table S20), and its quantum-derived polar surface area [63,88] (PSA-DFT = $271.70 \text{ \AA}^2$, 56.85% polar surface; Table 7) imposes a proportionally large desolvation cost [42,43,61] upon binding. Global electrophilicity [85] does not predict binding affinity [42,43] when conformational entropy is the dominant thermodynamic contribution [86,87].

3.5. $\Delta G_{GAS} \rightarrow \Delta G_{SOLV}$ Paradigm Shift: Implications for Inhibitor Design

Unsupervised PCA [72,73] identified ΔG_{GAS} [42,43] as the dominant source of variance across all four compound classes (PC1 loadings +0.416 to +0.464; 65.2–81.4% of variance; Table 4, Supplementary Figure S15), reflecting numerical magnitude [89] rather than pharmacological discriminance [66]. Supervised LDA [66] inverted this hierarchy [89]: ΔG_{SOLV} [42,43,61] emerged as the primary discriminant (LD1 coefficient = -205.24 vs. $\Delta G_{GAS} = -42.41$; 4.8-fold ratio), capturing 93.1% of between-class variance [66] (Figure 9; Supplementary Table S17). The 100% LDA [66] separability for clinical controls and secondary metabolites [24,44] (Calinski–Harabasz index [71] 1231.4 and 982.2 vs. PCA [72,73] 161.8 and 573.4; Table 5) validates the computational pipeline [42,43,67,68] rather than revealing its limitation. The design implication is mechanistically grounded: future computational leads targeting EGFR T790M [15,49,50] should optimize electrostatic complementarity [69,70] with the ARG841-anchored polar micro-environment [56–58]—maximizing ΔG_{SOLV} [42,43,61]—rather than prioritizing van der Waals contact area [68,90], a conclusion consistent with the per-residue hot spot architecture [91] (Figure 10).

3.6. Covalent/Non-Covalent Asymmetry: DFT Justification of the Reference Comparison

Osimertinib [52,53] (ΔG_{bind} T790M = -25.74 ± 3.37 kcal/mol) and mobocertinib [92–94] (-33.10 ± 4.41 kcal/mol) form irreversible Michael addition adducts with CYS797 [95,96]; their clinical potency depends on this covalent mechanism [96,97], which is not modeled by standard MM-PBSA [42,43]. The values reported [42,43] reflect the non-covalent binding component exclusively, and direct affinity comparison with non-covalent phytochemicals [24,55] is therefore methodologically inappropriate [42,43,98] (Figure 1; Table 3). DFT analysis [67,80] provides the complementary mechanistic information that MM-PBSA [42,43] cannot supply: $f(C28) = 0.2365$ [99] (highest in the dataset), $CDD(C28) = +0.0836$ [79], and 66% LUMO density localized on the acrylamide group [79] (Table 6; Supplementary Table S20) unambiguously identify C28 as the electrophilic Michael acceptor site [95,96], characterizing the warhead [96] that governs irreversible inhibition [71,97]. Per-residue MM-GBSA [42,43] corroborates this interpretation: CYS797 [95,96] showed uniformly positive contributions (+5.62 to +6.87 kcal/mol for all 16 compounds, including osimertinib [53] +6.62 kcal/mol; Figure 10), the physically expected result for a non-covalent

model [42,43] in which the thiol is a weak H-bond partner [70]. QM/MM simulations [68,87] at the ONIOM B3LYP/6-31G*/AMBER level [100,101] are proposed to quantify the covalent ΔG component [67,68] and rigorously contextualize the MM-PBSA comparison [42,43].

3.7. PDB 7A6K/V948R as Field Standard: Informing Rather than Limiting the Structural Analysis

The structural model PDB 7A6K [33] incorporates a V948R crystallographic stabilization mutation [45,102], introduced to prevent asymmetric EGFR dimerization during crystallization [102]. This residue is located in the distal C-lobe of the kinase domain [33,102] and does not contact the ATP-binding pocket residues (Lys745 [35,82], Met793 [57,102], Cys797 [95,96]) that govern selective inhibitor binding [35,57,102]. The T790M/V948R construct [33,45] has been used to resolve multiple co-crystal structures with diverse EGFR T790M inhibitors [35,49,57,102], establishing it as the field standard for structure-based drug design [35,82,102] targeting this resistance mutation [10,12]. The inclusion of comparative MD simulations [41,74,75] with wild-type EGFR (PDB 1M17 [40]) and the calculated $\Delta\Delta G_{\text{bind}}$ values [42,43] provide independent verification [67,76] that the observed binding selectivity is attributable to T790M [10,56] rather than V948R [33,45] (Table 3). Per-residue decomposition [42,43] (Figure 10) confirms that Val948/Arg948 contacts are exclusively van der Waals in nature [69,90], inconsistent with a driving role in the electrostatic selectivity pattern [69,70] documented here. Accordingly, Arg948 contacts reported in the docking interaction profiles (Section 2.1) are specific to the V948R-containing crystallographic construct and are not representative of interactions in native EGFR T790M.

3.8. ADMET Limitations and Formulation Strategies

Predicted ADMET profiles [103–106] reveal a pharmacokinetic bifurcation (Table 1; Supplementary Figures S5 and S6). Group 1—polar high-affinity leads (corilagin [24], rutin [47,48], isoquercetin [51]): Very low predicted Caco-2 permeability [103,105] (corilagin $\log P_{\text{app}} = -1.34$) and quantum-derived PSA substantially exceeding the Ertl threshold [88,107] for poor intestinal absorption (corilagin PSA-DFT = 303.08 Å²; rutin = 271.70 Å²; Table 7). Group 2—lignans (phyltetralin [24,44], hypophyllanthin [24,44]): High predicted permeability [103,105] ($\log P_{\text{app}}$ 0.98–1.14), phyltetralin PSA-DFT = 154.57 Å² [88], and IS = 1.21 [42,43]. Mitigation strategies include lipid nanoparticulate carriers (NLC/SLN) [108] with demonstrated oral bioavailability enhancement for analogous polyphenolic scaffolds [108]; parenteral administration in oncological contexts [109–111]; and the UGT-centered metabolic pathway predicted for glycosides [112] (Supplementary Table S2), which is predicted to reduce CYP3A4-mediated DDI risk [112] in cirrhotic HCC patients with compromised CYP activity [113–115].

3.9. PASS and KinScreen Predictions as Hypothesis-Generating Tools

PASS bioactivity predictions [116–118] and KinScreen kinome-wide profiling [119,120] are employed in this study as hypothesis-generating computational tools [121,122]. PASS predictions [116,118] are probabilistic and exploratory in nature—Pa values [116] represent statistical likelihood of activity based on structural training data [117,118], not confirmed biological activity. Similarly, KinScreen results [119,120] indicate potential binding propensity across the kinome [120,123] based on structural descriptors [121,124], not experimentally validated interactions [123]. These *in silico* predictions [116,118–120] are presented as preliminary hypotheses [121,122] to guide experimental prioritization [55,125] (Figure 2; Supplementary Figures S7–S9), not as evidence of confirmed polypharmacological effects [124,126]. The 8-fold difference in predicted kinase

interaction count between phytochemicals [24,122,125] and clinical controls [119,120,123] reflects predicted structural diversity of target engagement [124,127], not confirmed kinase activity [123,126]. Any inference about a computational polypharmacology hypothesis [121,122,124] requires biochemical kinome profiling [120,123] before it can be considered pharmacologically meaningful [126,127].

3.10. Limitations and Prioritized Future Experiments

The present study has several important limitations that must be acknowledged [54,55]. First and foremost, all results reported are based exclusively on computational predictions [54,55,128] and have not been validated experimentally. Binding free energy calculations (MM-PBSA [42], MM-GBSA [43]), while widely applied in computational drug discovery [42,43], represent approximations subject to force field uncertainties [75,87], conformational sampling limitations [41,76], and the absence of explicit entropic contributions [42,43,87]. PASS predictions [116,118] and KinScreen analyses [119,120] are probabilistic [116,117] and cannot substitute for biochemical validation [123,126]. Second, the structural model (PDB 7A6K [33], T790M/V948R) contains an engineered mutation (V948R) [45,102] that, while standard in EGFR structural biology [33,57,102], may introduce minor conformational biases absent in native T790M EGFR [10,56]. Third, corilagin [24], rutin [47,48], and isoquercetin [51] exhibit predicted low membrane permeability [103,105] that would limit oral bioavailability [88,107] and require formulation strategies [108–111] for clinical translation (Table 1; Supplementary Figure S5). These limitations do not diminish the hypothesis-generating value of the predictions [121,122,125] but mandate experimental validation [54,55,123] before pharmacological conclusions can be drawn.

Five prioritized future experiments with defined methodology are proposed to address these limitations:

1. Enzymatic IC_{50} assay [120,123]: Inhibition of recombinant EGFR T790M [10] and wild-type kinase (Carna Biosciences or ProQinase) with corilagin [24], eriodictyol-7-rhamnopyranoside [24,47], phylltetralin [24,44], and isoquercetin [51] (0.01–100 μ M); determination of experimental IS (IS_{exp}) [120,123] to validate $IS_{predicted} = 2.76$ for corilagin and $IS = 1.21$ for phylltetralin [42,43].
2. Phospho-EGFR Western blot (pY1068) [129]: HCC cell lines [13,14,130] expressing endogenous or stably transfected EGFR T790M [10] treated with 1–50 μ M of the three priority leads; downstream signaling (pAKT, pERK1/2) [14] as functional confirmation of target engagement [35,49].
3. Microsomal metabolic stability [131]: Incubation with pooled human liver microsomes for corilagin [24] and phylltetralin [24,44]; determination of CL_{int} and in vitro $t_{1/2}$ [131]; confirmation of UGT-dominant vs. CYP-dependent pathways [113–115] predicted by SOMP analysis [132] (Supplementary Table S2).
4. NLC/SLN nanoformulation for corilagin [24,108]: Preparation of lipid nanoparticle carriers [108] and Caco-2 evaluation of $\log P_{app}$ to determine whether nanoencapsulation overcomes the predicted permeability barrier ($\log P_{app} = -1.34$).
5. QM/MM simulations [68,87]: ONIOM-QM/MM at B3LYP/6-31G*/AMBER level [100,101] to characterize the covalent ΔG component of osimertinib [53,96] and mobocertinib [92]; comparison of $\Delta G_{covalent} + \Delta G_{non-cov}$ with ΔG_{bind} (corilagin) to rigorously contextualize MM-PBSA magnitudes [42,43] (Section 3.6).

4. Methods

4.1. Ligand Preparation

Thirty-seven authenticated *Phyllanthus niruri* phytoconstituents were retrieved from PubChem (National Center for Biotechnology Information, Bethesda, MD, USA) [133] in

SMILES format and prepared using Avogadro 1.98.0 (Avogadro Chemistry, New York, NY, USA, Open-source) [134] with the MMFF94 force field [135]: explicit hydrogen addition with stereochemistry preservation, Gasteiger charge assignment [136], energy minimization to $0.01 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-1}$ convergence, and torsional sampling for compounds exceeding five rotatable bonds. Protocol validation against five Cambridge Structural Database (Cambridge Crystallographic Data Centre, Cambridge, UK) [137] reference compounds achieved geometric fidelity of $0.8 \pm 0.3 \text{ \AA}$ RMSD versus experimental conformations.

4.2. Protein Preparation

The EGFR T790M/V948R crystal structure (PDB 7A6K, $2.0 \text{ \AA}$ resolution; co-crystallized with mobocertinib/TAK-788) [138] was refined using UCSF ChimeraX 1.8 (Resource for Biocomputing, Visualization, and Informatics, University of California, San Francisco, CA, USA) [139] and WinCoot 0.9.8.95 (MRC Laboratory of Molecular Biology, Cambridge, UK) [140]. Structural refinement achieved MolProbity score 0.84 (95th percentile) [141], 98.2% favored rotamers, and complete elimination of bond-length and angle outliers. TAK-788 coordinates were preserved for centroid validation during docking assessment, and protonation states were assigned at physiological pH 7.4. For wild-type EGFR selectivity comparisons, the PDB 1M17 structure ($2.6 \text{ \AA}$ resolution; erlotinib co-crystal) [40] was prepared identically after removal of the co-crystallized erlotinib.

4.3. Binding-Site Characterization

Cavity detection via DoGSiteScorer (Center for Bioinformatics, University of Hamburg, Hamburg, Germany) [142,143] identified the ATP-binding pocket as the sole druggable cavity (volume $1200 \text{ \AA}^3$; druggability score 0.89), validated by 92% co-occupancy with Reactome pathway (Ontario Institute for Cancer Research, Toronto, ON, Canada) [144], BindingDB (University of California San Diego, La Jolla, CA, USA) [145] EGFR inhibitors, and UniProt (EMBL-EBI, Hinxton, UK) [146] domain mapping to the tyrosine kinase catalytic region (residues 696–1022). Grid dimensions ($26 \times 19 \times 35 \text{ \AA}$; center coordinates: $x = 8, y = 8, z = 36$) encompassed all critical binding residues (MET790, MET793, LEU718, VAL726, LEU844, CYS797) within $8 \text{ \AA}$ of the TAK-788 crystallographic centroid, achieving $0.32 \text{ \AA}$ centroid accuracy confirmed by CASTp 3.0 analysis (University of Illinois at Chicago, Chicago, IL, USA) [147].

4.4. Molecular Docking

Docking was performed using SwissDock 2024 (Swiss Institute of Bioinformatics, Lausanne, Switzerland) [148] with Attracting Cavities 2.0 (Swiss Institute of Bioinformatics, Lausanne, Switzerland) [149] in Accurate mode with FACTS implicit solvation (Swiss Institute of Bioinformatics, Lausanne, Switzerland) [150] and medium exhaustivity settings, generating approximately 50 poses per ligand. Protocol validation against the crystallographic complex demonstrated: (i) TAK-788 redocking RMSD = $0.80 \pm 0.10 \text{ \AA}$ in triplicate (PDB 7A6K [138]); (ii) erlotinib redocking RMSD = $1.83 \text{ \AA}$ (PDB 1M17 [40]), both within the standard $2.0 \text{ \AA}$ acceptance threshold; (iii) DUD-E (Shoichet Laboratory, University of California, San Francisco, CA, USA) [151] decoy discrimination with 50 compounds yielding ROC-AUC = 0.91 and enrichment factor = 12.3; and (iv) correct hinge-region interactions for covalent reference inhibitors mobocertinib [138] and osimertinib. Compound selection for molecular dynamics was based on composite scoring ($0.4 \times \text{AC energy}$ [148,149] + $0.3 \times \text{pharmacophore}$ [81] + $0.2 \times \text{cluster weight}$ + $0.1 \times \text{interaction fingerprint}$), retention of Met793 hinge hydrogen bonding, $\geq 60\%$ ATP pocket volumetric overlap, and absence of predicted critical ADMET liabilities (see Section 4.5). For wild-type EGFR selectivity analysis, the top four phytochemical candidates (corilagin, eriodic-

tyol-7-rhamnopyranoside, and phylltetralin, plus osimertinib as clinical reference) and erlotinib (co-crystallized ligand of PDB 1M17 [40]) were docked into PDB 1M17 [40] using the identical protocol.

4.5. ADMET and Biological Activity Prediction

Biological activity was predicted using PASS Online 2024 (Way2Drug, Institute of Biomedical Chemistry, Moscow, Russia) [116,118] with dual-criteria thresholds: probability of activity $P_a \geq 0.70$ and confidence index $(P_a - P_i) \geq 0.30$. ADMET properties were evaluated via tri-platform integration: using pkCSM (University of Melbourne, Melbourne, VIC, Australia) [103], Deep-PK (University of Queensland, Brisbane, QLD, Australia) [104], and SwissADME (Swiss Institute of Bioinformatics, Lausanne, Switzerland) [105], achieving >88% inter-platform concordance across major endpoints. Absorption parameters included human intestinal absorption ($HIA \geq 80\%$), Caco-2 permeability, and P-glycoprotein profiles [103–105]. Metabolism profiling evaluated CYP450 inhibition across isoforms 1A2, 2C19, 2C9, 2D6, and 3A4 [103–105]. Toxicity assessment included hERG inhibition, AMES mutagenicity, and hepatotoxicity [103–105]. Drug-likeness was evaluated against Lipinski, Ghose, Veber, Egan, and Muegge filters [105]. Off-target kinome profiling was conducted using KinScreen (Way2Drug, Institute of Biomedical Chemistry, Moscow, Russia) [152] across 442 human kinases and PASS Targets (Way2Drug, Institute of Biomedical Chemistry, Moscow, Russia) [117,153] across 930 human proteins, applying $P_a \geq 0.70$ and confidence ≥ 0.30 thresholds.

4.6. Molecular Dynamics Simulations

All molecular dynamics simulations were performed using GROMACS 2023.4 (KTH Royal Institute of Technology, Stockholm, Sweden) [75] for 100 ns per system under NPT ensemble conditions at physiological temperature (310 K) and pressure (1 bar). Systems comprised: (i) fourteen *P. niruri* phytochemicals plus osimertinib and mobocertinib against EGFR T790M/V948R (PDB 7A6K [138]); and (ii) corilagin, eriodictyol-7-rhamnopyranoside, phylltetralin, osimertinib, and erlotinib against wild-type EGFR (PDB 1M17 [40]). All systems were constructed via CHARMM-GUI v2024 (Lehigh University, Bethlehem, PA, USA) [154–156] using the CHARMM36m force field [157,158] for protein and CGenFF 4.7 [159,160] for ligands; all ligand penalty scores were below the 50-unit curation threshold. Systems were solvated in cubic boxes (1.2 nm minimum solvent buffer) with TIP3P water [161] and 0.15 M NaCl to replicate physiological ionic conditions. The V948R crystallographic stabilization mutation present in PDB 7A6K [138] was retained in its entirety; this substitution resides in the distal C-lobe and does not contact ATP-binding pocket residues.

Energy minimization employed 2500 steps steepest descent ($1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ backbone restraints) followed by 2500 steps L-BFGS optimization ($400/40 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ backbone/side-chain restraints), converging to maximum force $<500 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$ [75]. NVT equilibration (125 ps, 310 K) used the Nosé–Hoover thermostat [162,163] ($\tau_T = 1.0$ ps); NPT equilibration (1 ns, 1 bar) used the Parrinello–Rahman barostat [164] ($\tau_P = 5.0$ ps; compressibility $4.5 \times 10^{-5} \text{ bar}^{-1}$). Equilibration convergence was confirmed by density stabilization ($1.045 \pm 0.005 \text{ g} \cdot \text{cm}^{-3}$), pressure equilibration (± 5 bar), and backbone RMSD <0.25 nm. Production simulations employed the leap-frog integrator (2 fs timestep) [75], LINCS constraints on hydrogen-containing bonds [165], Verlet cutoff (1.2 nm), force-switch van der Waals modifiers (1.0–1.2 nm), and Particle Mesh Ewald electrostatics (1.2 nm real-space cutoff) [166]. Trajectory frames were saved every 10 ps.

4.7. Trajectory Analysis

The initial 10 ns were discarded to exclude residual equilibration; all analyses were performed on the 90 ns production phase using GROMACS utilities [75]. Structural stability was assessed via backbone C α RMSD (gmx rms, least-squares fitting; stable complex criterion <4.0 Å) [75,77] and ligand heavy-atom RMSD after protein alignment (stable binding criterion <3.0 Å) [75,77]. Per-residue flexibility was quantified via C α RMSF (gmx rmsf) [75,77]. Hydrogen bond persistence was assessed using gmx hbond [75,167] with geometric criteria: donor–acceptor distance ≤ 0.35 nm and D–H...A angle $\geq 120^\circ$ [70]; bonds present in $\geq 30\%$ of production frames were classified as persistent. Global compactness was quantified via radius of gyration (gmx gyrate) [75] and solvent-accessible surface area (gmx sasa) [75].

4.8. MM-PBSA Binding Free Energy Calculations

Binding free energies ($\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}}$) [42,43] were computed using gmx_MMPBSA v1.6.4 (University of Calgary, Calgary, AB, Canada) [168] implementing the OBC-II Generalized Born model (igb = 5, mbondi2 radii) [169] with interior dielectric constant $\epsilon_{\text{in}} = 1$, exterior dielectric $\epsilon_{\text{out}} = 80$, and SASA-LCPO nonpolar solvation [170] ($\gamma = 0.02268 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$, $\beta = 3.849 \text{ kJ}\cdot\text{mol}^{-1}$) at 0.150 M ionic strength. The conformational entropy term ($-T\Delta S$) was omitted following established consensus [42,43] for structurally diverse natural product datasets. Based on autocorrelation analysis ($\tau_{\text{int}} = 4\text{--}6$ ps) [77,168], snapshots were extracted every 10 ps from the final 50 ns of each trajectory, yielding 501 frames per phytochemical compound and 201 frames per wild-type system. Trajectory convergence was verified using a halving test: simulations were deemed converged when $|\Delta \Delta G_{\text{bind}}| < 0.1 \text{ kcal}\cdot\text{mol}^{-1}$ between the first and second halves of the analysis window [42,168]. For the four compounds simulated against both T790M and wild-type EGFR, mutant selectivity was quantified via the Selectivity Index ($\text{IS} = |\Delta G_{\text{bind}}(\text{T790M})|/|\Delta G_{\text{bind}}(\text{WT})|$) [42,43] and differential binding free energy ($\Delta \Delta G_{\text{bind}} = \Delta G_{\text{bind}}(\text{WT}) - \Delta G_{\text{bind}}(\text{T790M})$) [42,43], with uncertainty propagated from standard deviations of both ΔG_{bind} values.

4.9. Multivariate PCA/LDA Analysis of MM-PBSA Energy Profiles

Principal component analysis (PCA) [171,172] and linear discriminant analysis (LDA) [171] were applied to the full MM-PBSA energy component profiles of all 16 compounds across 7416 MD trajectory frames, using seven energy components as input variables ($\Delta \text{VDWAALS}$, ΔEEL , ΔEGB , ΔESURF , ΔGGAS , ΔGSOLV , ΔGTOTAL) grouped into four pharmacological classes (flavonoids, lignans, secondary metabolites, and covalent controls). PCA was implemented as unsupervised dimensionality reduction maximizing total variance; LDA was applied at category level (4 classes, $N = 800$ subsampled frames) and compound level (16 classes, $N = 3102$ frames) using scikit-learn 1.3.0 (Python Software Foundation, Wilmington, DE, USA) in Python 3.10. Cluster quality was assessed independently via the Silhouette [173], the Calinski–Harabasz index [71], and the Davies–Bouldin index [174]. Cross-validated accuracy (leave-one-out) [171] is reported as a secondary metric given the inherent energetic overlap among structurally homologous natural product scaffolds; cluster quality indices [71,173,174] constitute the appropriate primary evaluation framework for this class of analysis.

4.10. Per-Residue MM-GBSA Energy Decomposition

Per-residue free energy decomposition [42,43] was performed on all 16 compound–EGFR T790M/V948R complexes using gmx_MMPBSA v1.6.4 (idecomp = 2) [168], encompassing 1,548,639 data points (501 frames $\times$ 21–26 residues per phytochemical; 201 frames $\times$ 21–25 residues per covalent control). Each residue contribution was decomposed into

internal strain (ΔG_{INT}), van der Waals (ΔG_{VDW}), electrostatic (ΔG_{EL}), generalized Born polar solvation (ΔG_{GB}), and nonpolar solvation (ΔG_{NP}) components [42,43,168]. Binding hot spots were defined as residues with $|\Delta G_{TOTAL}| \geq 1.0$ kcal·mol⁻¹ [42,43]. Inter-compound differences at each residue were assessed by the Kruskal–Wallis H-test [59] with Benjamini–Hochberg false discovery rate correction for multiple comparisons [175]. All per-residue means are reported with 95% bootstrap confidence intervals ($n = 2000$ resamples) [176]. Non-convergent residue–compound pairs (halving test $\Delta|\Delta G| \geq 0.1$ kcal·mol⁻¹) [42,168] are indicated with an asterisk (*) in all decomposition tables.

4.11. DFT Electronic Structure Characterization

The five top-ranked compounds (corilagin, eriodictyol-7-rhamnopyranoside, osimertinib, phyltetralin, and rutin) were characterized by density functional theory. Initial geometries were pre-optimized using GFN2-xTB (xTB 6.7.1, Grimme Group, University of Bonn, Bonn, Germany) [87,177,178] with the GBSA implicit water model and tight convergence criteria. Geometries were subsequently refined in a two-step DFT workflow in ORCA 6.1 (Max Planck Institute for Kohlenforschung, Mülheim an der Ruhr, Germany) [179,180]: B3LYP-D3(BJ) [64,181–183]/def2-SVP [65] followed by B3LYP-D3(BJ)/def2-TZVP with the RIJCOSX approximation [184] and the def2/J auxiliary basis set [185]; Grimme's D3 dispersion correction [183] with Becke–Johnson damping [64] was applied throughout. Harmonic vibrational frequency calculations at the B3LYP-D3(BJ)/def2-TZVP level [64,65,181–183] confirmed true energy minima (zero imaginary frequencies) for four of five compounds. Eriodictyol-7-rhamnopyranoside exhibited two low-frequency imaginary modes (−60.48 and −36.39 cm⁻¹) associated with rhamnosyl torsional flexibility; electronic properties derived from the converged wavefunction remain valid, and thermodynamic quantities for this compound are reported as approximate. Single-point SMD continuum solvation calculations (water) [186] were performed on final gas-phase geometries for solvation free energy estimates. Wavefunction analysis was conducted with Multiwfn 3.8 (Beijing Kein Research Center for Natural Sciences, Beijing, China) [79] on neutral (N), radical anion (N⁻¹), and radical cation (N⁻¹) electronic states. Global conceptual DFT reactivity descriptors [85,99] were computed via the Koopmans approximation ($IP = -\epsilon_{HOMO}$; $EA = -\epsilon_{LUMO}$) [187]. MEP surfaces were mapped on the 0.001 a.u. Bader isodensity surface [79]. Condensed Fukui functions (f^+ , f^- , f^0) [99,188,189] and the dual descriptor ($CDD = f^+ - f^-$) [190] were derived from Hirshfeld population analysis [191]. Hirshfeld orbital composition analysis quantified per-atom contributions to frontier molecular orbitals [79,191]. Non-covalent interactions were visualized via the reduced density gradient (NCI/RDG) approach [192,193].

4.12. Statistical Analysis

Statistical analyses were performed using Python SciPy 1.11.1 (NumFOCUS, Austin, TX, USA) [194] and R 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria) [195]. Normality was assessed by the Kolmogorov–Smirnov test [195]; 95% confidence intervals were computed by bootstrap resampling ($n = 2000$) [176]. Group comparisons employed two-tailed independent *t*-tests with Bonferroni correction [196] (significance threshold: $p < 0.05$). Effect sizes were quantified using Cohen's *d* [197]. Trajectory convergence required effective sample size $N_{eff} > 5$ [42,168] and drift < 10% across five 18 ns block windows.

5. Conclusions

A multiscale quantum-classical computational investigation of 37 *Phyllanthus niruri* phytoconstituents against EGFR T790M/V948R (PDB 7A6K) identified three structurally diverse computational leads—corilagin (hydrolyzable tannin), eriodictyol-7-rhamnopyra-

noside (flavonoid glycoside), and phylltetralin (lignan)—exhibiting predicted T790M binding affinities of -53.71 ± 5.05 , -44.35 ± 4.51 , and -34.54 ± 4.01 kcal/mol, respectively. Wild-type EGFR comparative simulations (PDB 1M17, 100 ns) yielded Selectivity Indices of $IS = 2.76$ ($\Delta\Delta G_{\text{bind}} = +34.22$ kcal/mol), $IS = 1.17$, and $IS = 1.21$, indicating T790M-preferential non-covalent binding across three structurally distinct scaffolds. The clinical reference osimertinib reproduced $IS = 1.43$ under identical conditions, providing internal methodological validation. Per-residue MM-GBSA decomposition (1,548,639 data points; 38 consensus hot spots) identified ARG841 (mean $\Delta G_{\text{TOTAL}} = -107.52 \pm 0.97$ kcal/mol; $KW_{\text{adj}} = 1.01 \times 10^{-145}$), ASP855, and LYS745 as the universal binding triad, with electrostatic interactions accounting for 87.4–96.0% of favorable binding energy—mechanistically explaining the observed MM-PBSA magnitudes. Supervised PCA/LDA analysis of 7416 trajectory frames identified solvation energy (ΔG_{SOLV} , LD1 coefficient = -205.24) rather than gas-phase energy as the primary pharmacological class discriminant, capturing 93.1% of between-class variance. DFT characterization at B3LYP-D3(BJ)/def2-TZVP established the quantum-mechanical basis for corilagin's electrostatic primacy ($V_{\text{s,min}} = -46.92$ kcal/mol), characterized osimertinib's Michael acceptor site ($f^+(\text{C}28) = 0.2365$; $\text{CDD}(\text{C}28) = +0.0836$), and independently corroborated the solvation-driven selectivity mechanism—constituting tri-level convergence across quantum mechanics, multivariate statistics, and classical energy decomposition. Corilagin, eriodictyol-7-rhamnopyranoside, and phylltetralin are positioned as priority candidates warranting experimental investigation, including recombinant EGFR T790M kinase inhibition assays, cellular phospho-EGFR signaling studies in HCC models, and pharmacokinetic profiling. These *in silico* predictions do not constitute experimental validation and require biochemical confirmation before any pharmacological conclusions can be drawn.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ijms27156568/s1>.

Author Contributions: S.L. Conceptualization, Collaboration management, Funding acquisition, Project administration, Supervision, Validation, Writing—review and editing. W.D.L.-G. Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Original draft writing, Visualization, Project administration—Docking, MD simulations, MM-PBSA/MM-GBSA, PCA/LDA, Per-residue decomposition, DFT/Multiwfn, WT EGFR selectivity analysis. T.V. Methodology, Formal analysis, Investigation, Data curation, Visualization—PCA/LDA multivariate analysis. S.J.M. Investigation, Writing—review and editing: Biological context and HCC EGFR relevance. A.V.A. Validation, Formal analysis, Writing—review and editing: WT EGFR selectivity analysis and ADMET profiling. F.R.J. Methodology, Formal analysis, Investigation, Data curation—MM-PBSA and MM-GBSA calculations. S.P. Formal analysis, Investigation, Data curation, Visualization—DFT/Multiwfn and per-residue energy decomposition. D.D.K. Conceptualization, Validation, Formal analysis, Writing—review and editing. The manuscript was written through contributions of all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data, analysis results, and input files supporting the findings of this study are publicly available in a permanent Zenodo repository at the following digital object identifier (DOI): <https://doi.org/10.5281/zenodo.18614515>. The deposited data package is organized into the following main folders, each containing a specific README.txt file for detailed explanations: Protein Structures—Initial and optimized coordinates of EGFR T790M/V948R (PDB 7A6K) and EGFR wild-type (PDB 1M17) used for docking and molecular dynamics simulations. Ligand Structures—SMILES strings, 3D conformers, and force-field parameters for the 37 *Phyllanthus niruri* phytochemicals and two clinical reference compounds (osimertinib, mobocertinib). Molecular

Docking—Raw and processed output files from SwissDock 2024 for all 37 phytochemicals against EGFR T790M, plus comparative docking against wild-type EGFR for the four lead candidates. ADMET and Bioactivity Predictions—Complete outputs from SwissADME, pkCSM, Deep-PK, PASS Online, and KinScreen platforms. Molecular Dynamics Simulations—GROMACS input files (.mdp, topology, index), compressed trajectories (.xtc), and processed analyses (RMSD, RMSF, Rg, SASA, hydrogen bonds) for all 16 T790M complexes (100 ns each) and all five wild-type EGFR complexes (100 ns each). MM-PBSA Binding Free Energies and Per-Residue Decomposition—gmx_MMPBSA input files, raw energy component data (7416 frames for PCA/LDA; 1,548,639 data points for per-residue decomposition), bootstrap confidence intervals, and Kruskal–Wallis statistical test outputs. PCA/LDA Multivariate Analysis—Python scripts (scikit-learn), input CSV matrices, and complete output tables (PCA loadings, LDA coefficients, Silhouette/Calinski–Harabasz/Davies–Bouldin metrics). DFT Electronic Structure Characterization—ORCA 6.1 input and output files for geometry optimization and frequency calculations at B3LYP-D3(BJ)/def2-TZVP level; Multiwfn 3.8 input files and complete surfanalysis outputs (MEP surfaces, Fukui functions, dual descriptor, Hirshfeld orbital composition, NCI/RDG) for the five top-ranked compounds. All software and web servers used in this study are publicly available and detailed in the Methodology section and the repository’s README files.

Acknowledgments: The authors would like to express their sincere gratitude to Siddha Vetha University for their generous funding and support, which made this research project possible.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Definition
ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicity
AMES	Ames mutagenicity test
ATP	Adenosine Triphosphate
BBB	Blood–Brain Barrier
CA	Carbonic Anhydrase
Caco-2	Human colorectal adenocarcinoma cell line (permeability model)
CDD	Dual Descriptor
CDFT	Conceptual Density Functional Theory
CYP	Cytochrome P450
DDI	Drug–Drug Interaction
DFT	Density Functional Theory
DILI	Drug-Induced Liver Injury
DUD-E	Directory of Useful Decoys Enhanced
DYRK2	Dual-specificity tyrosine phosphorylation-regulated kinase 2
EA	Electron Affinity
EGFR	Epidermal Growth Factor Receptor
HIA	Human Intestinal Absorption
hERG	human Ether-à-go-go-Related Gene
HOMO	Highest Occupied Molecular Orbital
HCC	Hepatocellular Carcinoma
IC ₅₀	Half-maximal Inhibitory Concentration
IP	Ionization Potential
IS	Selectivity Index
KW	Kruskal–Wallis
LDA	Linear Discriminant Analysis
LINCS	Linear Constraint Solver
LUMO	Lowest Unoccupied Molecular Orbital
MD	Molecular Dynamics
MEP	Molecular Electrostatic Potential

MMFF94	Merck Molecular Force Field 94
MM-GBSA	Molecular Mechanics Generalized Born Surface Area
MM-PBSA	Molecular Mechanics Poisson–Boltzmann Surface Area
MPI	Molecular Polarity Index
NCI	Non-Covalent Interactions
NEK6	NIMA-related kinase 6
NPT	Isothermal–isobaric ensemble
NVT	Canonical ensemble
ODI	Orbital Delocalization Index
Pa	Probability of Activity
PASS	Prediction of Activity Spectra for Substances
PCA	Principal Component Analysis
PDB	Protein Data Bank
P-gp	P-glycoprotein
Pi	Probability of Inactivity
PME	Particle Mesh Ewald
PSA	Polar Surface Area
QM/MM	Quantum Mechanics/Molecular Mechanics
RDG	Reduced Density Gradient
Rg	Radius of Gyration
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
ROC-AUC	Receiver Operating Characteristic Area Under the Curve
SAR	Structure–Activity Relationship
SASA	Solvent Accessible Surface Area
SI	Selectivity Index
SMD	Solvation Model based on Density
SMILES	Simplified Molecular Input Line Entry System
T790M	Threonine 790 to Methionine mutation
TAK-788	Mobocertinib (EGFR/HER2 inhibitor)
UGT	UDP-glucuronosyltransferase
V948R	Valine 948 to Arginine mutation
WT	Wild-Type

References

- Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* **2024**, *74*, 229–263. <https://doi.org/10.3322/caac.21834>.
- Llovet, J.M.; Ricci, S.; Mazzaferro, V.; Hilgard, P.; Gane, E.; Blanc, J.F.; de Oliveira, A.C.; Santoro, A.; Raoul, J.L.; Forner, A.; et al. Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* **2008**, *359*, 378–390. PMID: 18650514.
- Food and Drug Administration. FDA Approves Lenvatinib for Unresectable Hepatocellular Carcinoma. Available online: https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/206947s011lbl.pdf (accessed on 13 October 2025).
- Ermi, A.G.; Sarkar, D. Resistance to Tyrosine Kinase Inhibitors in Hepatocellular Carcinoma (HCC): Clinical Implications and Potential Strategies to Overcome the Resistance. *Cancers* **2024**, *16*, 3944.
- Wang, J.; Shi, J.; Mi, L.; Li, N.; Han, X.; Zhao, M.; Duan, X.; Han, G.; Hou, J.; Yin, F. Identification and validation of a lenvatinib resistance–related prognostic signature in HCC, in which PFKFB4 contributes to tumor progression and lenvatinib resistance. *BMC Gastroenterol.* **2025**, *25*, 287.
- Kudo, M.; Finn, R.S.; Qin, S.; Han, K.-H.; Ikeda, K.; Piscaglia, F.; Baron, A.; Park, J.-W.; Han, G.; Jassem, J.; et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: A randomised phase 3 non-inferiority trial. *Lancet.* **2018**, *391*, 1163–1173. PMID: 29433850.
- Berasain, C.; Avila, M.A. Epidermal growth factor receptor (EGFR) crosstalks in liver cancer. *Cancers.* **2011**, *3*, 2444–2461. PMID: 24213365. <https://doi.org/10.3390/cancers3022444>.
- Komposch, K.; Sibilia, M. EGFR signaling in liver diseases. *Int. J. Mol. Sci.* **2016**, *17*, 30. PMID: 26729094. <https://doi.org/10.3390/ijms17010030>.

9. Tantiwettrueangdet, A.; Panvichian, R.; Asavarut, P. HCC-derived EGFR mutants are functioning, EGF-dependent, and erlotinib-resistant. *Cell Biosci.* **2020**, *10*, 33. PMID: 32190291. <https://doi.org/10.1186/s13578-020-00407-1>.
10. Yun, C.-H.; Mengwasser, K.E.; Toms, A.V.; Woo, M.S.; Greulich, H.; Wong, K.K.; Meyerson, M.; Eck, M.J. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2070–2075.
11. Sundaresan, T.K.; Sequist, L.V.; Heymach, J.V.; Riely, G.J.; Jänne, P.A.; Koch, W.H.; Sullivan, J.P.; Fox, D.B.; Maher, R.; Muzikansky, A.; et al. Detection of T790M, the Acquired Resistance EGFR Mutation, by Tumor Biopsy versus Noninvasive Blood-Based Analyses. *Clin. Cancer Res.* **2016**, *22*, 1103–1110.
12. Bakir, A.; Aru, E.E. The Role of the T790M Mutation in Lung Cancer: A Molecular Perspective Based on Genetic Data. *Acta Haematol. Oncol. Turc.* **2024**, *57*, 101–108.
13. Jin, H.; Shi, Y.; Lv, Y.; Yuan, S.; Ramirez, C.F.A.; Lieftink, C.; Wang, L.; Wang, S.; Wang, C.; Dias, M.H.; et al. EGFR activation limits the response of liver cancer to lenvatinib. *Nature* **2021**, *595*, 730–734.
14. Hu, B.; Zou, T.; Qin, W.; Shen, X.; Su, Y.; Li, J.; Chen, Y.; Zhang, Z.; Sun, H.; Zheng, Y.; et al. Inhibition of EGFR Overcomes Acquired Lenvatinib Resistance Driven by STAT3–ABCBI Signaling in Hepatocellular Carcinoma. *Cancer Res.* **2022**, *82*, 3845–3857.
15. Lim, M.; Yang, D.; Imperial, R.J.L.; Majeed, U.; Tsai, J.; Franses, J.W.; Hsieh, D. Effect of EGFR and ERBB2 amplifications and activating alterations on efficacy of lenvatinib in hepatocellular carcinoma. *J. Clin. Oncol.* **2023**, *41*, 600.
16. Shi, Y.; Cui, D.; Xia, L.; Shi, D.; Jin, G.; Wang, S.; Lin, Y.; Tang, X.; Chi, J.; Wang, T.; et al. Efficacy and safety of lenvatinib plus gefitinib in lenvatinib-resistant hepatocellular carcinomas: A prospective, single-arm exploratory trial. *Signal Transduct. Target. Ther.* **2024**, *9*, 335. PMC: PMC11628597.
17. Thomas, M.B.; Morris, J.S.; Chadha, R.; Iwasaki, M.; Kaur, H.; Lin, E.; Kaseb, A.; Glover, K.; Davila, M.; Abbruzzese, J. Phase II trial of the combination of bevacizumab and erlotinib in patients who have advanced hepatocellular carcinoma. *J. Clin. Oncol.* **2009**, *27*, 843–850.
18. Chiu, W.; Chen, S.C.; Chao, Y.; Lee, J.M. Molecular profiling of circulating tumor cells guides effective EGFR inhibitor treatment in advanced hepatocellular carcinoma: A case report. *Front. Oncol.* **2025**, *15*, 1608604. PMC: PMC12484127.
19. Xia, Y.; Luo, H.; Liu, J.P.; Gluud, C. *Phyllanthus* species versus antiviral drugs for chronic hepatitis B virus infection. *Cochrane Database Syst. Rev.* **2013**, *4*, CD009004.
20. Freitas, A.M.; Schor, N.; Boim, M.A. The effect of *Phyllanthus niruri* on urinary inhibitors of calcium oxalate crystallization and other factors associated with renal stone formation. *BJU Int.* **2002**, *89*, 829–834.
21. Kiemer, A.K.; Hartung, T.; Huber, C.; Vollmar, A.M. *Phyllanthus amarus* has anti-inflammatory potential by inhibition of iNOS, COX-2, and cytokines via the NF- κ B pathway. *J. Hepatol.* **2003**, *38*, 289–297.
22. Mediani, A.; Abas, F.; Khatib, A.; Tan, C.P.; Ismail, I.S.; Shaari, K.; Ismail, A.; Lajis, N.H. Relationship Between Metabolites Composition and Biological Activities of *Phyllanthus niruri* Extracts Prepared by Different Drying Methods and Solvents Extraction. *Plant Foods Hum. Nutr.* **2015**, *70*, 184–192.
23. Mediani, A.; Abas, F.; Maulidiani, M.; Khatib, A.; Tan, C.P.; Ismail, I.S.; Shaari, K.; Ismail, A. Characterization of Metabolite Profile in *Phyllanthus niruri* and Correlation with Bioactivity Elucidated by Nuclear Magnetic Resonance Based Metabolomics. *Molecules* **2017**, *22*, 902.
24. Kaur, N.; Kaur, B.; Sirhindi, G. Phytochemistry and Pharmacology of *Phyllanthus niruri* L.: A Review. *Phytother. Res.* **2017**, *31*, 980–1004.
25. de Araújo Júnior, R.F.; de Souza, T.P.; Pires, J.G.L.; Soares, L.A.L.; de Araújo, A.A.; Petrovick, P.R.; Mâcedo, H.D.O.; Oliveira, A.L.C.d.S.L.; Guerra, G.C.B. A dry extract of *Phyllanthus niruri* protects normal cells and induces apoptosis in human liver carcinoma cells. *Exp. Biol. Med.* **2012**, *237*, 1281–1288.
26. Liu, L.; Wang, B.; Ma, Y.; Sun, K.; Wang, P.; Li, M.; Dong, J.; Qin, M.; Li, M.; Wei, C.; et al. A review of *Phyllanthus urinaria* L. in the treatment of liver disease: Viral hepatitis, liver fibrosis/cirrhosis and hepatocellular carcinoma. *Front. Pharmacol.* **2024**, *15*, 1443667.
27. Zhang, L.; Li, Y.; Mao, J. Research progress on natural products against hepatocellular carcinoma. *Biocell* **2024**, *48*, 905.
28. Mushtaq, S.; Abbasi, B.H.; Uzair, B.; Abbasi, R. Natural products as reservoirs of novel therapeutic agents. *EXCLI J.* **2018**, *17*, 420–451.
29. Lautié, E.; Russo, O.; Ducrot, P.; Boutin, J.A. Unraveling Plant Natural Chemical Diversity for Drug Discovery Purposes. *Front. Pharmacol.* **2020**, *11*, 397.
30. Al Saiqali, M.; Aziz, S.S.; Reddy, A.V. Current Status on Phytochemicals Classification, Structure-Activity Relationship, Stereochemistry and AI-Driven Applications: A Systematic Review. *Pharmacogn. Rev.* **2025**, *19*, 186–211.

31. Bharate, S.B.; Lindsley, C.W. Natural Products Driven Medicinal Chemistry. *J. Med. Chem.* **2024**, *67*, 20723–20730.
32. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803.
33. Lategahn, J.; Tumbirink, H.L.; Schultz-Fademrecht, C.; Heimsoeth, A.; Werr, L.; Niggenaber, J.; Keul, M.; Parmaksiz, F.; Baumann, M.; Menninger, S.; et al. Insight into Targeting Exon20 Insertion Mutations of the Epidermal Growth Factor Receptor with Wild Type-Sparing Inhibitors. *J. Med. Chem.* **2022**, *65*, 6643–6655.
34. Niggenaber, J.; Hardick, J.; Lategahn, J.; Rauh, D. Complex crystal structures of EGFR with third-generation kinase inhibitors and simultaneously targeting PROTACs. *ACS Med. Chem. Lett.* **2020**, *11*, 2484–2490. <https://doi.org/10.1021/acsmchemlett.0c00472>.
35. Beyett, T.S.; To, C.; Heppner, D.E.; Rana, J.K.; Schmoker, A.M.; Jang, J.; De Clercq, D.J.H.; Gomez, G.; Scott, D.A.; Gray, N.S.; et al. Molecular basis for cooperative binding and synergy of ATP-site and allosteric EGFR inhibitors. *Nat. Commun.* **2022**, *13*, 2530.
36. To, C.; Beyett, T.S.; Jang, J.; Feng, W.W.; Bahcall, M.; Haikala, H.M.; Shin, B.H.; Heppner, D.E.; Rana, J.K.; Leeper, B.A.; et al. An allosteric inhibitor against the therapy-resistant mutant forms of EGFR in non-small cell lung cancer. *Nat. Cancer* **2022**, *3*, 402–417.
37. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Aided Mol. Des.* **2009**, *23*, 597–608. PMID: 19340590.
38. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods: The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82. PMID: 21082818.
39. Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual density functional theory. *Chem. Rev.* **2003**, *103*, 1793–1874. PMID: 12744694.
40. Stamos, J.; Sliwkowski, M.X.; Eigenbrot, C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J. Biol. Chem.* **2002**, *277*, 46265–46272. PMID: 12196540. <https://doi.org/10.1074/jbc.M207135200>.
41. Hata, H.; Phuoc Tran, D.; Marzouk Sobeh, M.; Kitao, A. Binding free energy of protein/ligand complexes calculated using dissociation Parallel Cascade Selection Molecular Dynamics and Markov state model. *Biophys. Physicobiol.* **2021**, *18*, 305–316.
42. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z.H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478–9508.
43. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert. Opin. Drug Discov.* **2015**, *10*, 449–461.
44. Lee, N.Y.; Khoo, W.K.; Adnan, M.A.; Mahalingam, T.P.; Fernandez, A.R.; Jeevaratnam, K. The pharmacological potential of *Phyllanthus niruri*. *J. Pharm. Pharmacol.* **2016**, *68*, 953–969.
45. Zhao, P.; Yao, M.-Y.; Zhu, S.-J.; Chen, J.-Y.; Yun, C.-H. Crystal structure of EGFR T790M/C797S/V948R in complex with EAI045. *Biochem. Biophys. Res. Commun.* **2018**, *502*, 332–337.
46. Gupta, A.; Singh, A.K.; Kumar, R.; Ganguly, R.; Rana, H.K.; Pandey, P.K.; Sethi, G.; Bishayee, A.; Pandey, A.K. Corilagin in Cancer: A Critical Evaluation of Anticancer Activities and Molecular Mechanisms. *Molecules* **2019**, *24*, 3399.
47. Namdeo, A.G.; Boddu, S.H.S.; Amawi, H.; Ashby CR Jr Tukaramrao, D.B.; Trivedi, P.; Babu, R.J.; Tiwari, A.K. Flavonoids as multi-target compounds: A special emphasis on their potential as chemo-adjuvants in cancer therapy. *Curr. Pharm. Des.* **2020**, *26*, 1712–1728.
48. Xia, J.F.; Gao, J.J.; Inagaki, Y.; Kokudo, N.; Nakata, M.; Tang, W. Flavonoids as potential anti-hepatocellular carcinoma agents: Recent approaches using HepG2 cell line. *Drug Discov. Ther.* **2013**, *7*, 1–8.
49. Jia, Y.; Yun, C.H.; Park, E.; Ercan, D.; Manuia, M.; Juarez, J.; Xu, C.; Rhee, K.; Chen, T.; Zhang, H.; et al. Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. *Nature* **2016**, *534*, 129–132.
50. To, C.; Jang, J.; Chen, T.; Park, E.; Mushajiang, M.; De Clercq, D.J.; Xu, M.; Wang, S.; Cameron, M.D.; Heppner, D.E.; et al. Single and Dual Targeting of Mutant EGFR with an Allosteric Inhibitor. *Cancer Discov.* **2019**, *9*, 926–943.
51. Shui, L.; Wang, W.; Xie, M.; Ye, B.; Li, X.; Liu, Y.; Zheng, M. Isoquercitrin induces apoptosis and autophagy in hepatocellular carcinoma cells via AMPK/mTOR/p70S6K signaling pathway. *Aging (Albany NY)* **2020**, *12*, 24318–24332.
52. Remon, J.; Steuer, C.E.; Ramalingam, S.S.; Felip, E. Osimertinib and other third-generation EGFR TKI in EGFR-mutant NSCLC patients. *Ann. Oncol.* **2018**, *29*, i20–i27.
53. Cross, D.A.; Ashton, S.E.; Ghioreghiu, S.; Eberlein, C.; Nebhan, C.A.; Spitzler, P.J.; Orme, J.P.; Finlay, M.R.V.; Ward, R.A.; Mellor, M.J.; et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov.* **2014**, *4*, 1046–1061.

54. Das, R.; Bhattarai, A.; Karn, R.; Tamang, B. Computational investigations of potential inhibitors of monkeypox virus envelope protein E8 through molecular docking and molecular dynamics simulations. *Sci. Rep.* **2024**, *14*, 19585.
55. Tejaswini, A.; Alekhya, K.P.; Manisha, P.; Nageswari, P.; Swathi, K. Screening of *Phyllanthus niruri* Plant Active Constituents for Anticancer and Antifungal Activity by Insilico Methods. *Adv. Exp. Med. Biol.* **2023**, *1423*, 123–132.
56. Doss, G.P.; Rajith, B.; Chakraborty, C.; NagaSundaram, N.; Ali, S.K.; Zhu, H. Structural signature of the G719S-T790M double mutation in the EGFR kinase domain and its response to inhibitors. *Sci. Rep.* **2014**, *4*, 5868.
57. Gajiwala, K.S.; Feng, J.; Ferre, R.A.; Ryan, K.; Brodsky, O.; Weinrich, S.; Kath, J.C.; Stewart, A. Insights into the Aberrant Activity of Mutant EGFR Kinase Domain and Drug Recognition. *Structure* **2013**, *21*, 209–219.
58. Kannan, S.; Venkatachalam, G.; Lim, H.H.; Surana, U.; Verma, C. Conformational landscape of the epidermal growth factor receptor kinase reveals a mutant specific allosteric pocket. *Chem. Sci.* **2018**, *9*, 5212–5222.
59. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621.
60. Gilli, G.; Gilli, P. Towards an unified hydrogen-bond theory. *J. Mol. Struct.* **2000**, *552*, 1–15.
61. Forouzesh, N.; Mishra, N. An Effective MM/GBSA Protocol for Absolute Binding Free Energy Calculations: A Case Study on SARS-CoV-2 Spike Protein and the Human ACE2 Receptor. *Molecules* **2021**, *26*, 2383.
62. Li, S.; Tan, H.Y.; Wang, N.; Cheung, F.; Hong, M.; Feng, Y. The Potential and Action Mechanism of Polyphenols in the Treatment of Liver Diseases. *Oxid. Med. Cell Longev.* **2018**, *2018*, 8394818.
63. Murray, J.S.; Politzer, P. The electrostatic potential: An overview. *WIREs Comput. Mol. Sci.* **2011**, *1*, 153–163.
64. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
65. Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
66. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **1936**, *7*, 179–188.
67. Ehrlich, S.; Göller, A.H.; Grimme, S. Towards full Quantum-Mechanics-based Protein-Ligand Binding Affinities. *Chemphyschem* **2017**, *18*, 898–905.
68. Molani, F.; Webb, S.; Cho, A.E. Correction to “Combining QM/MM Calculations with Classical Mining Minima to Predict Protein–Ligand Binding Free Energy”. *J. Chem. Inf. Model.* **2024**, *64*, 1105.
69. Myslinski, J.M.; DeLorbe, J.E.; Clements, J.H.; Martin, S.F. Protein-ligand interactions: Thermodynamic effects associated with increasing nonpolar surface area. *J. Am. Chem. Soc.* **2011**, *133*, 18518–18521.
70. Madushanka, A.; Moura, R.T., Jr.; Verma, N.; Kraka, E. Quantum Mechanical Assessment of Protein–Ligand Hydrogen Bond Strength Patterns: Insights from Semiempirical Tight-Binding and Local Vibrational Mode Theory. *Int. J. Mol. Sci.* **2023**, *24*, 6311.
71. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.* **1974**, *3*, 1–27.
72. Du, J.; Li, J.; Wen, J.; Liu, J.; Xiao, H.; Zhang, A.; Yang, D.; Sun, P.; Zhou, H.; Xu, J. A Systematic Hierarchical Virtual Screening Model for RhlR Inhibitors Based on PCA, Pharmacophore, Docking, and Molecular Dynamics. *Int. J. Mol. Sci.* **2024**, *25*, 8000.
73. Chen, J.; Wang, J.; Yang, W.; Zhao, L.; Zhao, J.; Hu, G. Molecular Mechanism of Phosphorylation-Mediated Impacts on the Conformation Dynamics of GTP-Bound KRAS Probed by GaMD Trajectory-Based Deep Learning. *Molecules* **2024**, *29*, 2317.
74. Karplus, M.; McCammon, J.A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652.
75. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
76. Pavan, M.; Menin, S.; Bassani, D.; Sturlese, M.; Moro, S. Qualitative Estimation of Protein-Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2022**, *62*, 5715–5728.
77. Wu, N.; Zhang, R.; Peng, X.; Fang, L.; Chen, K.; Jestilä, J. Elucidation of protein-ligand interactions by multiple trajectory analysis methods. *Phys. Chem. Chem. Phys.* **2024**, *26*, 6134–6145.
78. Lemer, C.M.-R.; Rooman, M.J.; Wodak, S.J. Protein structure prediction by threading methods: Evaluation of current techniques. *Proteins* **1995**, *23*, 337–355.
79. Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580–592.
80. Almasoudi, H.H.; Nahari, M.H. Targeting Plasmodium falciparum Schizont Egress Antigen-1 in Infected Red Blood Cells: Docking-Based Fingerprinting, Density Functional Theory, Molecular Dynamics Simulations, and Binding Free Energy Analysis. *Pharmaceuticals* **2025**, *18*, 237.
81. Wolber, G.; Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* **2005**, *45*, 160–169.

82. Wittlinger, F.; Heppner, D.E.; To, C.; Günther, M.; Shin, B.H.; Rana, J.K.; Schmoker, A.M.; Beyett, T.S.; Berger, L.M.; Berger, B.-T.; et al. Design of a “Two-in-One” Mutant-Selective Epidermal Growth Factor Receptor Inhibitor That Spans the Orthosteric and Allosteric Sites. *J. Med. Chem.* **2022**, *65*, 1370–1383.
83. Cooper, A.J.; Sequist, L.V.; Lin, J.J. Third-generation EGFR and ALK inhibitors: Mechanisms of resistance and management. *Nat. Rev. Clin. Oncol.* **2022**, *19*, 499–514.
84. Li, Y.; Li, X.; Ma, W.; Dong, Z. Conformational Transition Pathways of Epidermal Growth Factor Receptor Kinase Domain from Multiple Molecular Dynamics Simulations and Bayesian Clustering. *J. Chem. Theory Comput.* **2014**, *10*, 3503–3511.
85. Parr, R.G.; Szentpály, L.v.; Liu, S. Electrophilicity index. *J. Am. Chem. Soc.* **1999**, *121*, 1922–1924.
86. Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **2012**, *18*, 9955–9964.
87. Christensen, A.S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem. Rev.* **2016**, *116*, 5301–5337.
88. Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714–3717.
89. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150202.
90. Durham, E.; Dorr, B.; Woetzel, N.; Staritzbichler, R.; Meiler, J. Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *J. Mol. Model.* **2009**, *15*, 1093–1108.
91. Trezza, A.; Visibelli, A.; Roncaglia, B.; Barletta, R.; Iannielli, S.; Mahboob, L.; Spiga, O.; Santucci, A. Unveiling Dynamic Hotspots in Protein–Ligand Binding: Accelerating Target and Drug Discovery Approaches. *Int. J. Mol. Sci.* **2025**, *26*, 3971.
92. Huang, W.S.; Li, F.; Gong, Y.; Zhang, Y.; Youngsaye, W.; Xu, Y.; Zhu, X.; Greenfield, M.T.; Kohlmann, A.; Taslimi, P.M.; et al. Discovery of mobocertinib, a potent, oral inhibitor of EGFR exon 20 insertion mutations in non-small cell lung cancer. *Bioorg Med. Chem. Lett.* **2023**, *80*, 129084.
93. Vasconcelos, P.E.N.S.; Kobayashi, I.S.; Kobayashi, S.S.; Costa, D.B. Preclinical Characterization of Mobocertinib Highlights the Putative Therapeutic Window of This Novel EGFR Inhibitor to EGFR Exon 20 Insertion Mutations. *JTO Clin. Res. Rep.* **2021**, *2*, 100105.
94. Duke, E.S.; Stapleford, L.; Drezner, N.; Amatya, A.K.; Mishra-Kalyani, P.S.; Shen, Y.-L.; Maxfield, K.; Zirkelbach, J.F.; Bi, Y.; Liu, J.; et al. FDA Approval Summary: Mobocertinib for Metastatic Non–Small Cell Lung Cancer with EGFR Exon 20 Insertion Mutations. *Clin. Cancer Res.* **2023**, *29*, 508–512.
95. Chen, L.; Fu, W.; Zheng, L.; Liu, Z.; Liang, G. Recent Progress of Small-Molecule Epidermal Growth Factor Receptor (EGFR) Inhibitors against C797S Resistance in Non-Small-Cell Lung Cancer. *J. Med. Chem.* **2018**, *61*, 4290–4300.
96. Bradshaw, J.M.; McFarland, J.M.; Paavilainen, V.O.; Bisconte, A.; Tam, D.; Phan, V.T.; Romanov, S.; Finkle, D.; Shu, J.; Patel, V.; et al. Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat. Chem. Biol.* **2015**, *11*, 525–531.
97. Lu, H.; Tonge, P.J. Drug-target residence time: Critical information for lead optimization. *Curr. Opin. Chem. Biol.* **2010**, *14*, 467–474.
98. Hayek-Orduz, Y.; Vásquez, A.F.; Villegas-Torres, M.F.; Caicedo, P.A.; Achenie, L.E.K.; González Barrios, A.F. Novel covalent and non-covalent complex-based pharmacophore models of SARS-CoV-2 main protease (M^{Pro}) elucidated by microsecond MD simulations. *Sci. Rep.* **2022**, *12*, 14030.
99. Yang, W.; Parr, R.G. Hardness, softness, and the Fukui function in the electronic theory of metals and catalysis. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6723–6726.
100. Maseras, F.; Morokuma, K. IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.* **1995**, *16*, 1170–1179.
101. Svensson, M.; Humbel, S.; Froese, R.D.J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A multilayered integrated MO + MM method for geometry optimizations and single point energy predictions. A test for Diels–Alder reactions and Pt(P(*t*-Bu)₃)₂ + H₂ oxidative addition. *J. Phys. Chem.* **1996**, *100*, 19357–19363.
102. Heppner, D.E.; Günther, M.; Wittlinger, F.; Laufer, S.A.; Eck, M.J. Structural Basis for EGFR Mutant Inhibition by Trisubstituted Imidazole Inhibitors. *J. Med. Chem.* **2020**, *63*, 4293–4305.
103. Pires, D.E.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072.
104. Myung, Y.; de Sá, A.G.C.; Ascher, D.B. Deep-PK: Deep learning for small molecule pharmacokinetic and toxicity prediction. *Nucleic Acids Res.* **2024**, *52*, W469–W475.

105. 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*, 42717.
106. Al Azzam, K. SwissADME and pkCSM Webservers Predictors: An integrated Online Platform for Accurate and Comprehensive Predictions for In Silico ADME/T Properties of Artemisinin and its Derivatives. *Kompleksn Ispolz. Miner. Syra.* **2023**, *325*, 14–21.
107. Clark, D.E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J. Pharm. Sci.* **1999**, *88*, 807–814.
108. Jacob, S.; Rao, R.; Gorain, B.; Boddu, S.H.S.; Nair, A.B. Solid Lipid Nanoparticles and Nanostructured Lipid Carriers for Anti-cancer Phytochemical Delivery: Advances, Challenges, and Future Prospects. *Pharmaceutics* **2025**, *17*, 1079.
109. Sarkar, S.; Kiren, S.; Gmeiner, W.H. Review of Prodrug and Nanodelivery Strategies to Improve the Treatment of Colorectal Cancer with Fluoropyrimidine Drugs. *Pharmaceutics* **2024**, *16*, 734.
110. Zhang, Y.; Cui, H.; Zhang, R.; Zhang, H.; Huang, W. Nanoparticulation of Prodrug into Medicines for Cancer Therapy. *Adv Sci (Weinh)* **2021**, *8*, e2101454.
111. Chen, K.J.; Plaunt, A.J.; Leifer, F.G.; Kang, J.Y.; Cipolla, D. Recent advances in prodrug-based nanoparticle therapeutics. *Eur. J. Pharm. Biopharm.* **2021**, *165*, 219–243.
112. Xie, S.; Chen, Y.; Chen, S.; Zeng, S. Structure–metabolism relationships for the glucuronidation of flavonoids by UGT1A3 and UGT1A9. *J. Pharm. Pharmacol.* **2011**, *63*, 297–304.
113. El-Khateeb, E.; Achour, B.; Al-Majdoub, Z.M.; Barber, J.; Rostami-Hodjegan, A. Non-uniformity of Changes in Drug-Metabolizing Enzymes and Transporters in Liver Cirrhosis: Implications for Drug Dosage Adjustment. *Mol. Pharm.* **2021**, *18*, 3563–3577.
114. McConn, D.J., 2nd; Lin, Y.S.; Mathisen, T.L.; Blough, D.K.; Xu, Y.; Hashizume, T.; Taylor, S.L.; Thummel, K.E.; Shuhart, M.C. Reduced duodenal cytochrome P450 3A protein expression and catalytic activity in patients with cirrhosis. *Clin. Pharmacol. Ther.* **2009**, *85*, 387–393.
115. Albarmawi, A.; Czock, D.; Gauss, A.; Ehehalt, R.; Bermejo, J.L.; Burhenne, J.; Ganten, T.M.; Sauer, P.; Haefeli, W.E. CYP3A activity in severe liver cirrhosis correlates with Child-Pugh and model for end-stage liver disease (MELD) scores. *Br. J. Clin. Pharmacol.* **2014**, *77*, 160–169.
116. Lagunin, A.; Stepanchikova, A.; Filimonov, D.; Poroikov, V. PASS: Prediction of activity spectra for biologically active substances. *Bioinformatics* **2000**, *16*, 747–748.
117. Lagunin, A.; Filimonov, D.; Poroikov, V. Multi-targeted natural products evaluation based on biological activity prediction with PASS. *Curr. Pharm. Des.* **2010**, *16*, 1703–1717.
118. Filimonov, D.A.; Lagunin, A.A.; Gloriovova, T.A.; Rudik, A.V.; Druzhilovskii, D.S.; Pogodin, P.V.; Poroikov, V.V. Prediction of the biological activity spectra of organic compounds using the PASS online web resource. *Chem. Heterocycl. Comp.* **2014**, *50*, 444–457.
119. Kooijman, J.J.; van Riel, W.E.; Dylus, J.; Prinsen, M.B.W.; Grobben, Y.; de Bitter, T.J.J.; van Doornmalen, A.M.; Melis, J.J.T.M.; Uitdehaag, J.C.M.; Narumi, Y.; et al. Comparative kinase and cancer cell panel profiling of kinase inhibitors approved for clinical use from 2018 to 2020. *Front Oncol.* **2022**, *12*, 953013.
120. Uitdehaag, J.C.M.; de Roos, J.A.D.M.; van Doornmalen, A.M.; Prinsen, M.B.W.; de Man, J.; Tanizawa, Y.; Kawase, Y.; Yoshino, K.; Buijsman, R.C.; Zaman, G.J.R.; et al. Comparison of the cancer gene targeting and biochemical selectivities of all targeted kinase inhibitors approved for clinical use. *PLoS ONE* **2014**, *9*, e92146.
121. Chaudhari, R.; Tan, Z.; Huang, B.; Zhang, S. Computational polypharmacology: A new paradigm for drug discovery. *Expert. Opin. Drug Discov.* **2017**, *12*, 279–291.
122. Fang, J.; Liu, C.; Wang, Q.; Lin, P.; Cheng, F. In silico polypharmacology of natural products. *Brief. Bioinform.* **2018**, *19*, 1153–1171.
123. Davis, M.I.; Hunt, J.P.; Herrgard, S.; Ciceri, P.; Wodicka, L.M.; Pallares, G.; Hocker, M.; Treiber, D.K.; Zarrinkar, P.P. Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2011**, *29*, 1046–1051.
124. Feldmann, C.; Bajorath, J. Advances in computational polypharmacology. *Mol. Inform.* **2022**, *41*, e2200190.
125. Choo, M.Z.Y.; Chai, C.L.L. The polypharmacology of natural products in drug discovery and development. *Annu. Rep. Med. Chem.* **2023**, *61*, 55–100.
126. Ravikumar, B.; Aittokallio, T. Improving the efficacy-safety balance of polypharmacology in multi-target drug discovery. *Expert. Opin. Drug Discov.* **2018**, *13*, 179–192.
127. Hopkins, A.L. Network pharmacology: The next paradigm in drug discovery. *Nat. Chem. Biol.* **2008**, *4*, 682–690.
128. Ja'afaru, S.C.; Uzairu, A.; Bayil, I.; Sallau, M.S.; Ndukwe, G.I.; Ibrahim, M.T.; Moin, A.T.; Mollah, A.K.M.M.; Absar, N. Unveiling potent inhibitors for schistosomiasis through ligand-based drug design, molecular docking, molecular dynamics simulations and pharmacokinetics predictions. *PLoS ONE* **2024**, *19*, e0302390.

129. Turke, A.B.; Zejnullahu, K.; Wu, Y.L.; Song, Y.; Dias-Santagata, D.; Lifshits, E.; Toschi, L.; Rogers, A.; Mok, T.; Sequist, L.; et al. Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. *Cancer Cell* **2010**, *17*, 77–88.
130. Yang, J.D.; Hainaut, P.; Gores, G.J.; Amadou, A.; Plymoth, A.; Roberts, L.R. A global view of hepatocellular carcinoma: Trends, risk, prevention and management. *Nat. Rev. Gastroenterol. Hepatol.* **2019**, *16*, 589–604.
131. Knights, K.M.; Stresser, D.M.; Miners, J.O.; Crespi, C.L. In Vitro Drug Metabolism Using Liver Microsomes. *Curr. Protoc. Pharmacol.* **2016**, *74*, 7.8.1–7.8.24.
132. Rudik, A.; Dmitriev, A.; Lagunin, A.; Filimonov, D.; Poroikov, V. SOMP: Web server for in silico prediction of sites of metabolism for drug-like compounds. *Bioinformatics* **2015**, *31*, 2046–2048.
133. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **2020**, *49*, D1388–D1395.
134. Hanwell, M.D.; Curtis, D.E.; Lonie, D.C.; Vandermeersch, T.; Zurek, E.; Hutchison, G.R. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminf.* **2012**, *4*, 17.
135. Halgren, T. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comp. Chem.* **1996**, *17*, 490–519.
136. Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity — A rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219–3228. [https://doi.org/10.1016/0040-4020\(80\)80168-2](https://doi.org/10.1016/0040-4020(80)80168-2).
137. Groom, C.R.; Bruno, I.J.; Lightfoot, M.P.; Ward, S.C. The Cambridge Structural Database. *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72*, 171–179.
138. Gonzalez, F.; Vincent, S.; Baker, T.E.; Gould, A.E.; Li, S.; Wardwell, S.D.; Nadworny, S.; Ning, Y.; Zhang, S.; Huang, W.-S.; et al. Mobocertinib (TAK-788): A Targeted Inhibitor of EGFR Exon 20 Insertion Mutants in Non-Small Cell Lung Cancer. *Cancer Discov.* **2021**, *11*, 1672–1687.
139. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **2023**, *32*, e4792.
140. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501.
141. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **2018**, *27*, 293–315.
142. Volkamer, A.; Griewel, A.; Grombacher, T.; Rarey, M. Analyzing the topology of active sites: On the prediction of pockets and subpockets. *J. Chem. Inf. Model.* **2010**, *50*, 2041–2052.
143. Volkamer, A.; Kuhn, D.; Grombacher, T.; Rippmann, F.; Rarey, M. Combining global and local measures for structure-based druggability predictions. *J. Chem. Inf. Model.* **2012**, *52*, 360–372.
144. Milacic, M.; Beavers, D.; Conley, P.; Gong, C.; Gillespie, M.; Griss, J.; Haw, R.; Jassal, B.; Matthews, L.; May, B.; et al. The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res.* **2024**, *52*, D647–D656.
145. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R.N.; Gilson, M.K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.
146. UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51*, D523–D531.
147. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, W363–W367. <https://doi.org/10.1093/nar/gky473>.
148. Bugnon, M.; Röhrig, U.F.; Goullieux, M.; Perez, M.A.S.; Daina, A.; Michielin, O.; Zoete, V. SwissDock 2024: Major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic Acids Res.* **2024**, *52*, W324–W332.
149. Röhrig, U.F.; Goullieux, M.; Bugnon, M.; Zoete, V. Attracting Cavities 2.0: Improving the flexibility and robustness for small-molecule docking. *J. Chem. Inf. Model.* **2023**, *63*, 3925–3940.
150. Haberthür, U.; Caflisch, A. FACTS: Fast analytical continuum treatment of solvation. *J. Comput. Chem.* **2008**, *29*, 701–715. <https://doi.org/10.1002/jcc.20839>.
151. Mysinger, M.M.; Carchia, M.; Irwin, J.J.; Shoichet, B.K. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594.
152. Pogodin, P.V.; Lagunin, A.A.; Rudik, A.V.; Filimonov, D.A.; Druzhilovskiy, D.S.; Nicklaus, M.C.; Poroikov, V.V. How to Achieve Better Results Using PASS-Based Virtual Screening: Case Study for Kinase Inhibitors. *Front. Chem.* **2018**, *6*, 133. <https://doi.org/10.3389/fchem.2018.00133>.
153. Pogodin, P.V.; Lagunin, A.A.; Filimonov, D.A.; Poroikov, V.V. PASS Targets: Ligand-based multi-target computational system based on a public data and naïve Bayes approach. *SAR QSAR Environ. Res.* **2015**, *26*, 783–793.

154. Kern, N.R.; Lee, J.; Choi, Y.K.; Im, W. CHARMM-GUI Multicomponent Assembler for modeling and simulation of complex multicomponent systems. *Nat. Commun.* **2024**, *15*, 5459.
155. Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S.J.; Patel, D.S.; Beaven, A.H.; Lee, K.I.; Rui, H.; Park, S.; et al. CHARMM-GUI 10 years for biomolecular modeling and simulation. *J. Comput. Chem.* **2017**, *38*, 1114–1124.
156. Lee, J.; Cheng, X.; Swails, J.M.; Yeom, M.S.; Eastman, P.K.; Lemkul, J.A.; Wei, S.; Buckner, J.; Jeong, J.C.; Qi, Y.; et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405–413.
157. Matsubara, D.; Kasahara, K.; Dokainish, H.M.; Oshima, H.; Sugita, Y. Modified Protein-Water Interactions in CHARMM36m for Thermodynamics and Kinetics of Proteins in Dilute and Crowded Solutions. *Molecules* **2022**, *27*, 5726.
158. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B.L.; Grubmüller, H.; MacKerell, A.D., Jr. CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71–73.
159. Kim, S.; Lee, J.; Jo, S.; Brooks, C.L., 3rd; Lee, H.S.; Im, W. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *J. Comput. Chem.* **2017**, *38*, 1879–1886.
160. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
161. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
162. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511–519.
163. Hoover, W.G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.
164. Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
165. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
166. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
167. Gorelov, S.; Titov, A.; Tolicheva, O.; Konevega, A.; Shvetsov, A. Determination of Hydrogen Bonds in GROMACS: A new implementation to overcome memory limitation. *J. Chem. Inf. Model.* **2024**, *64*, 6241–6246.
168. Valdés-Tresanco, M.S.; Valdés-Tresanco, M.E.; Valiente, P.A.; Moreno, E. gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17*, 6281–6291.
169. Onufriev, A.; Bashford, D.; Case, D.A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004**, *55*, 383–394.
170. Weiser, J.; Shenkin, P.S.; Still, W.C. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J. Comput. Chem.* **1999**, *20*, 217–230.
171. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
172. Maharjan, R.; Gyawali, K.; Acharya, A.; Khanal, M.; Ghimire, M.P.; Lamichhane, T.R. Artemisinin derivatives as potential drug candidates against Mycobacterium tuberculosis: Insights from molecular docking, MD simulations, PCA, MM/GBSA and AD-MET analysis. *Mol. Simul.* **2024**, *50*, 717–728.
173. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
174. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *1*, 224–227.
175. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300.
176. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall/CRC: Boca Raton, FL, USA, 1993.
177. Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
178. Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *WIREs Comput. Mol. Sci.* **2021**, *11*, e1493.
179. Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.

180. Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108.
181. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
182. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
183. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
184. Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A ‘chain-of-spheres’ algorithm for the Hartree-Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109.
185. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
186. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
187. Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms. *Physica* **1934**, *1*, 104–113.
188. Parr, R.G.; Yang, W. Density functional approach to the frontier-electron theory of chemical reactivity. *J. Am. Chem. Soc.* **1984**, *106*, 4049–4050.
189. Yang, W.; Mortier, W.J. The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines. *J. Am. Chem. Soc.* **1986**, *108*, 5708–5711.
190. Morell, C.; Grand, A.; Toro-Labbé, A. New dual descriptor for chemical reactivity. *J. Phys. Chem. A* **2005**, *109*, 205–212.
191. Hirshfeld, F.L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta* **1977**, *44*, 129–138.
192. Johnson, E.R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A.J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.
193. Contreras-García, J.; Johnson, E.R.; Keinan, S.; Chaudret, R.; Piquemal, J.P.; Beratan, D.N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632.
194. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272.
195. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022. Available online: <https://www.R-project.org/> (accessed on 20 October 2025).
196. Dunn, O.J. Multiple comparisons among means. *J. Am. Stat. Assoc.* **1961**, *56*, 52–64.
197. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988; ISBN 978-0-8058-0283-2.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.