In Silico Study of Licorice (*Glycyrrhiza inflata*) Compounds on EGFR Inhibition as an Anti-Lung Cancer Agent

Nataszna Ribielle Meiru[1], Muhammad Fikri Alvaro[1], Meika Fatia Kayrani[1], Rania Salsabila[1], Adeline Margaret[2], Risna Maulida Setiawan[2], Diah Lia Aulifa[2*]

[1]Bachelor Program in Pharmacy, Faculty of Pharmacy, Universitas Padjadjaran, Sumedang 45363, Indonesia
[2]Department of Pharmaceutical Analysis and Medicinal Chemistry, Faculty of Pharmacy, Universitas Padjadjaran, Sumedang 45363, Indonesia

Abstract

Non-small cell lung cancer (NSCLC) remains one of the leading causes of cancer-related mortality worldwide, with wild-type epidermal growth factor receptor (wild-type EGFR) being an important therapeutic target. This study aimed to identify potential wild-type EGFR inhibitors from *Glycyrrhiza inflata* through an *in silico* approach. A total of 184 phytochemicals were sequentially evaluated using Lipinski's Rule of Five, ADMET prediction, pharmacophore-based virtual screening, and molecular docking against the wild-type EGFR kinase domain (PDB ID: 3POZ). The docking protocol was validated by redocking the native ligand (TAK-285), yielding a root mean square deviation (RMSD) of 1.27 Å, indicating reliable docking performance. Drug-likeness and ADMET screening identified 61 compounds with favorable physicochemical, pharmacokinetic, and safety profiles for subsequent analyses. Compared with the reference inhibitor erlotinib, Parvisoflavone A exhibited a lower binding affinity (−8.91 kcal/mol; Ki=292.70 nM vs. −9.66 kcal/mol; Ki=83.43 nM), but demonstrated the highest pharmacophore fit score (42.63%), favorable drug-likeness and ADMET properties, and interactions with the key catalytic residues Lys745, Asp855, and Thr790 within the ATP-binding pocket of wild-type EGFR. These findings identify Parvisoflavone A as a promising hit compound for wild-type EGFR inhibition and provide a basis for further molecular dynamics simulations and experimental validation for NSCLC.

Keywords: *Glycyrrhiza inflata, molecular docking, non-small cell lung cancer, pharmacophore modeling, wild-type EGFR.*

INTRODUCTION

Lung cancer remains a major global health burden and is the leading cause of cancer-related mortality worldwide. The latest Global Cancer Observatory (GLOBOCAN) 2024 estimates reported approximately 2.64 million new lung cancer cases and 1.86 million deaths globally, making lung cancer the most frequently diagnosed cancer and the leading cause of cancer-related death in 2024 (Sung, *et al*., 2026). In Indonesia, lung

Submitted: July 05, 2026
Revised: August 25, 2026
Accepted: September 04, 2026

*Corresponding author: diah.lia@unpad.ac.id

cancer is the second most frequently diagnosed cancer, with 58,446 new cases (12.3% of all cancer cases), and remains the leading cause of cancer-related mortality (Sung, *et al*., 2026). Histologically, lung cancer is broadly classified into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), with NSCLC accounting for approximately 80–85% of lung cancer cases (Oh, *et al.*, 2019; Padinharayil, *et al*., 2023). Based on the 2024 global lung cancer estimate, this proportion corresponds to approximately 2.11–2.24 million NSCLC cases worldwide. Despite advances in diagnosis and targeted therapy, advanced NSCLC continues to have a poor prognosis, highlighting the need for improved therapeutic strategies (Wang, *et al*., 2022; Friedlaender, *et al*., 2024).

The epidermal growth factor receptor (EGFR) is an important molecular target in NSCLC because its activation regulates cellular proliferation, survival, differentiation, migration, and angiogenesis (Shaban, *et al*., 2024). Although activating EGFR mutations have received considerable attention, aberrant activation and overexpression of wild-type EGFR can also contribute to NSCLC progression through persistent activation of downstream signaling pathways, including PI3K/AKT and RAS/MAPK (Wang, *et al*., 2022; Shaban, *et al*., 2024). Accordingly, EGFR tyrosine kinase inhibitors (EGFR-TKIs), including first-generation inhibitors such as erlotinib, have been developed to inhibit ATP binding within the EGFR kinase domain and suppress downstream oncogenic signaling. EGFR-TKIs have provided important therapeutic benefits in NSCLC; however, their clinical effectiveness may be limited by primary or acquired resistance and treatment-related adverse effects (Sattler, *et al*., 2023; Shaban, *et al*., 2024).

Resistance to EGFR-targeted therapy represents a major challenge in the long-term management of NSCLC because tumor cells can regain proliferative signaling despite continued EGFR inhibition. Acquired resistance has been reported following treatment with different generations of EGFR-TKIs and may arise through multiple molecular mechanisms. One well-established on-target mechanism is the EGFR T790M mutation, which can reduce the efficacy of first- and second-generation EGFR-TKIs by altering the kinase domain and increasing the relative preference for ATP. Resistance to third-generation inhibitors, including osimertinib, may subsequently occur through additional EGFR alterations such as C797S, which prevents the covalent interaction required for effective inhibition. In addition to on-target alterations, bypass signaling represents another important mechanism of resistance; for example, MET amplification can reactivate downstream signaling despite EGFR inhibition. Resistance may also involve alterations in downstream signaling pathways and histological transformation, further contributing to treatment failure (Sattler, *et al*., 2023). The diversity of these resistance mechanisms can limit the durability of EGFR-targeted therapy and supports the investigation of complementary therapeutic strategies that may modulate EGFR-associated signaling and potentially complement conventional EGFR-TKI treatment.

Natural products represent an important source of structurally diverse compounds for anticancer drug discovery and may provide potential candidates for complementary therapeutic strategies. *Glycyrrhiza inflata* is particularly rich in bioactive constituents, including flavonoids, chalcones, coumarins, and triterpenoid saponins, several of which have demonstrated pharmacological and anticancer activities (Sharifi-Rad, *et al*., 2021; Wang, *et al*., 2020; Wu, *et al*., 2025). Flavonoid and chalcone derivatives from *Glycyrrhiza* species have also been reported to modulate oncogenic signaling pathways associated with EGFR. For example, licochalcone B has been reported to inhibit the growth of NSCLC cells through dual targeting of EGFR and MET (Oh, *et al*., 2019), while licochalcone derivatives have demonstrated diverse anticancer mechanisms involving multiple molecular targets (Deng, *et al*.,

2023). These findings suggest that constituents of *G. inflata* may have potential as complementary candidates for further investigation in lung cancer therapy. However, despite the reported biological activities of *G. inflata* constituents, their potential interactions with wild-type EGFR in the context of NSCLC have not been systematically characterized. *Glycyrrhiza inflata* was selected in the present study because of its diverse chemical composition and the reported pharmacological activities of its constituents, particularly flavonoids and chalcones, which have been associated with anticancer and oncogenic signaling-modulating effects. This chemical diversity provides a relevant compound library for identifying structurally distinct candidates that may interact with cancer-related molecular targets such as EGFR. In particular, an integrated evaluation combining drug-likeness, predicted pharmacokinetic and toxicity profiles, pharmacophore compatibility, and molecular docking against the wild-type EGFR kinase domain remains limited.

We hypothesized that bioactive constituents of *G. inflata* possess favorable physicochemical and predicted pharmacokinetic properties and may exhibit structural compatibility with the ATP-binding site of wild-type EGFR. To address this gap, the present study aimed to investigate the potential of *G. inflata* constituents as candidates for complementary or adjuvant strategies in lung cancer by evaluating their potential interactions with wild-type EGFR using an integrated computer-aided drug discovery (CADD) workflow. The study combined drug-likeness and ADMET prediction, pharmacophore-based virtual screening, and molecular docking against the wild-type EGFR kinase domain (PDB ID: 3POZ). This approach was used to prioritize *G. inflata* compounds with favorable physicochemical and predicted pharmacokinetic properties and to characterize their potential interactions within the EGFR ATP-binding site. The resulting computational evidence is intended to provide a preliminary basis for further investigation of selected *G. inflata* compounds as potential complementary candidates in lung cancer therapy.

MATERIALS AND METHODS

Materials

This study was conducted on Lenovo ThinkPad X1 Yoga 4th Gen running Windows 11 Pro 64-bit 25H2, equipped with an Intel(R) Core(TM) i7-8665U CPU @ 1.90GHz (2.11 GHz), and Intel (R) UHD Graphics 620 128 MB, and 16 GB LPDDR3 RAM @ 2133 MHz. The software environment integrated LigandScout (pharmacophore modeling), ChemDraw Ultra 12.0 (2D structure sketching), Chem3D Pro 12.0 (geometry optimization using the MM2 molecular mechanics force field), AutoDockTools 1.5.7/AutoDock v.4.2.6 (molecular docking), and BIOVIA Discovery Studio Visualizer v25.1.0.24284 (interaction analysis). Web-based utilities included KNApSAcK (identification of compounds reported in *Glycyrrhiza* inflata), PubChem (3D molecular structure retrieval), RCSB Protein Data Bank (EGFR crystal structure and co-crystallized ligand retrieval), SwissADME (ADME and drug-likeness prediction), PreADMET (ADME and toxicity prediction), DUD-E (active and decoy compounds retrieval for pharmacophore model validation).

Lipinski's Rule of Five Analysis, Pharmacokinetics and Toxicity

Lipinski's Rule of Five (RO5) parameters were predicted using SwissADME (http://www.swissadme.ch/) based on canonical SMILES obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) for both reference and test compounds. Compounds were considered to comply with RO5 when they had a molecular weight ≤500 Da, Log *P* ≤ 5, hydrogen bond donors (HBD) ≤5, and hydrogen bond acceptors (HBA) ≤10. The pharmacokinetic and toxicity properties of the reference and test compounds were predicted using the PreADMET

web server (https://preadmet.webservice.bmdrc.org/). The .mol files were submitted to the ADME Prediction and Toxicity Prediction modules, and the predicted outputs included human intestinal absorption (HIA), Caco-2 permeability, plasma protein binding (PPB), blood-brain barrier (BBB) penetration, Ames mutagenicity, and carcinogenicity. Compounds predicted to be non-mutagenic and non-carcinogenic were retained for subsequent analyses.

Pharmacophore Validation and Screening

Pharmacophore models were constructed using LigandScout from active and decoy compound data from the DUD-E database (https://dude.docking.org/targets) to identify the common structural features responsible for EGFR inhibition. The pharmacophore features considered included hydrogen bond donors, hydrogen bond acceptors, hydrophobic features, aromatic rings, and ionizable groups. Models with an area under the curve (AUC) >0.70 were considered to have acceptable discriminatory performance and selected as primary screening filters (Hamza, *et al*., 2012). The optimized pharmacophore model was then applied to the *Glycyrrhiza inflata* compound library to identify candidate compounds sharing the essential pharmacophoric features required for EGFR inhibition.

Preparation of Ligand and Receptor

The three-dimensional crystal structure of the wild-type EGFR receptor (PDB ID: 3POZ) was retrieved from the RCSB Protein Data Bank (https://www.rcsb.org/). The receptor structure was processed using BIOVIA Discovery Studio Visualizer v25.1.0.24284 to remove the native ligand and was saved in .pdb format. The ligand structures were subjected to geometry optimization using Chem3D Pro 12.0 with the MM2 molecular mechanics force field before ligand preparation. Receptor and ligand preparation was then performed using AutoDockTools 1.5.7. For receptor preparation, polar hydrogen atoms and Kollman charges were added, while all hydrogen atoms and Gasteiger charges were added to the ligands, followed by torsion tree definition. The prepared receptor and ligand structures were then exported in pdbqt format for subsequent molecular docking analysis.

Validation of Molecular Docking

Validation of the molecular docking protocol was performed by re-docking the native ligand (TAK-285) into the binding pocket of the wild-type EGFR receptor using AutoDock 4.2.6. The resulting docked pose was superimposed onto the native ligand conformation, and the root mean square deviation (RMSD) was calculated to evaluate docking accuracy. An RMSD value of ≤2.0 Å was used as the acceptance criterion, confirming the reliability of the docking parameters for subsequent molecular docking analyses.

Molecular Docking

Molecular docking was performed using AutoDock Tools 1.5.7 and AutoDock 4.2.6. The docking grid was set with a center at X=18.284, Y=31.913, and Z=11.626, with a grid size of 32×32×32 points and a grid spacing of 0.375 Å. Docking calculations were performed with 100 genetic algorithm runs, and the ligands were docked into the defined binding site. The resulting docking poses were evaluated based on their predicted binding affinities and ligand–receptor interactions. The binding site was defined based on the co-crystallized ligand binding site. Visualization of molecular docking results was conducted using AutoDock Tools and BIOVIA Discovery Studio Visualizer (https://www.3ds.com/products/biovia/discovery-studio). The docked complexes were displayed in 3D formats for examination of the ligand binding orientation within the receptor.

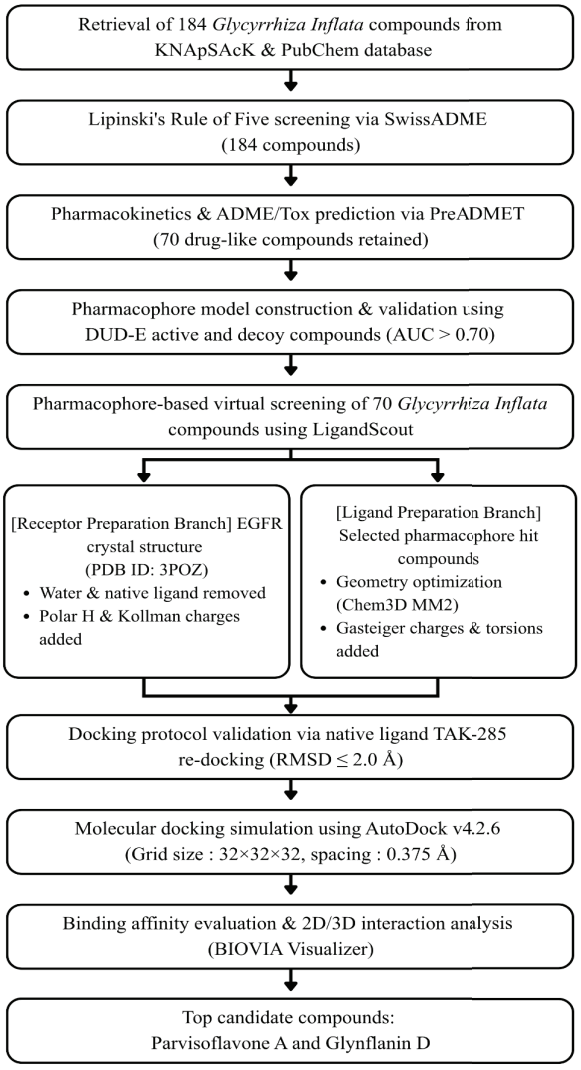


Figure 1. Virtual screening workflow for identifying potential EGFR inhibitors from *Glycyrrhiza inflata*.

RESULT

Lipinski's Rule of Five Analysis, Pharmacokinetics and Toxicity

Lipinski's Rule of Five is commonly used to predict oral bioavailability, where compounds are considered drug-like when they fulfill the criteria of molecular weight below 500 Da, Log *P* below 5, hydrogen bond donors fewer than 5, and hydrogen bond acceptors fewer than 10 (Lipinski, *et al.*, 2001). Lipinski's Rule of Five and ADMET prediction were performed independently for all 184 compounds identified from *Glycyrrhiza inflata* to evaluate their drug-likeness, pharmacokinetic characteristics, and toxicity profiles. Compounds exhibiting one or more violations of Lipinski's Rule of Five were considered not to satisfy the drug-likeness criteria. Only compounds that simultaneously complied with all Lipinski's Rule of Five criteria and passed the ADMET evaluation were retained for subsequent pharmacophore-based virtual screening.

Based on Lipinski's Rule of Five, most flavonoid and chalcone derivatives fulfilled the

established criteria, whereas several compounds, particularly glycyrrhizin analogs and uralsaponins, exhibited one or more violations related to molecular weight, lipophilicity, or hydrogen bonding parameters. ADMET prediction showed that several compounds possessed high human intestinal absorption (HIA) values, with Kanzonol E, Abyssinone III, and Glyinflanin D demonstrating HIA values exceeding 95%. In addition, a number of compounds were predicted to be non-mutagenic and non-carcinogenic. Overall, 61 compounds satisfied both the Lipinski's Rule of Five and ADMET screening criteria and were subsequently subjected to pharmacophore-based virtual screening.

Pharmacophore Validation and Screening

The pharmacophore model was validated using receiver operating characteristic (ROC) curve analysis to evaluate its ability to discriminate between active and inactive compounds. Prior to virtual screening, the performance of the generated pharmacophore model was evaluated using receiver operating characteristic (ROC) analysis to determine its ability to discriminate active compounds from inactive (decoy) compounds. For pharmacophore validation, a benchmark dataset consisting of 250 experimentally reported active compounds and 1,000 decoy compounds was obtained from the DUD-E database for EGFR. The active and decoy compounds were used to evaluate the ability of the generated pharmacophore models to discriminate known EGFR-active compounds from presumed inactive compounds. Several statistical parameters, including the area under the curve (AUC), enrichment factor (EF), goodness-of-hit (GH) score, sensitivity, specificity, and accuracy, were calculated, and the validation results are summarized in Table 1. The pharmacophore model validation produced an AUC value of 0.82, indicating good discriminatory performance. The validated model was subsequently used for the virtual screening of *Glycyrrhiza inflata* compounds against the wild-type EGFR receptor. Table 2 presents the top 10 compounds with the highest pharmacophore fit scores among the 61 compounds that passed the initial Lipinski's Rule of Five and ADMET screening.

Although the pharmacophore model produced an AUC of 0.82 and demonstrated high sensitivity (0.988), its specificity was relatively low (0.061), accompanied by a high number of false-positive predictions (939) and a low overall accuracy (0.246). The low goodness-of-hit score (0.025) and enrichment factor (1.041) further indicate that the model had limited ability to effectively enrich active compounds over decoys. These findings suggest that the model was highly sensitive to compounds sharing the defined pharmacophoric features but had limited selectivity toward true EGFR-active compounds.

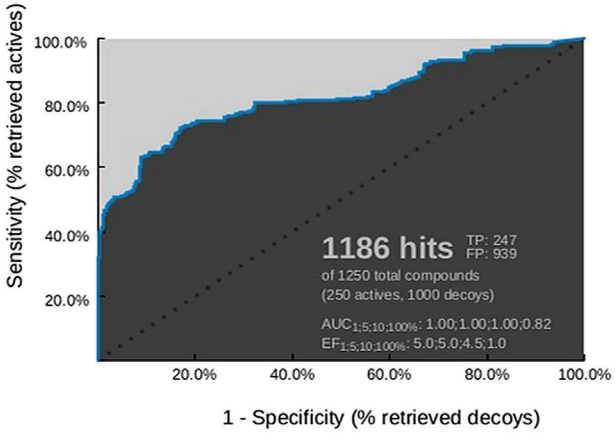


Figure 2. ROC curve of the pharmacophore model.

Table 1. Pharmacophore modeling validation parameters.

Parameter	Result
Total compounds in database (D)	1250
Active total in database (A)	250
Total hits (Ht)	1186
Active hits (Ha)/True Positive (TP)	247
False Positive (FP)	939
True Negatives (TN = D − FP)	61
False Negatives (FN = A − Ha)	3
Area Under ROC Curve (AUC)	0.82
Enrichment Factor (EF = [Ha × D]/[Ht × A])	1.041
$(GH = \left[\frac{\{Ha(3A+Ht)\}}{4HtA}\right]\left\{1-\frac{(Ht-Ha)}{(D-A)}\right\})$	0.025
Sensitivity (TPR = TP/A)	0.988
Specificity (TNR = TN/D)	0.061
Accuracy (ACC = (TP + TN)/(A + D))	0.246

Based on the pharmacophore screening results presented in Table 2, several compounds demonstrated high fit score values toward the validated pharmacophore model. Angustone B exhibited the highest fit score (42.84%), followed by Parvisoflavone A (42.63%), 5,2',4'-trihydroxy-7-methoxy-6-prenylisoflavone (42.62%), and Licoisoflavone B (42.61%). High fit score values indicate favorable alignment between ligand chemical features and the pharmacophore model, including hydrogen bond acceptors, hydrogen bond donors, hydrophobic regions, and aromatic interaction features (Yang, 2010).

Validation of Molecular Docking

The molecular docking protocol was validated through redocking of the native ligand TAK-285 into the active site of the wild-type-EGFR receptor. As shown in Table 3, the redocking process produced an RMSD value of 1.27 Å. The native ligand exhibited a binding energy of −10.06 kcal/mol with an inhibition constant (Ki) of 42.55 nM, while erlotinib as the comparative drug showed a binding energy of −9.66 kcal/mol. The overlay visualization between the native ligand and the redocked ligand (Figure 3) showed a high degree of conformational similarity within the

Table 2. Pharmacophore screening fit score results.

No	Compound	Fit Score (%)
1	Angustone B	42.84
2	Parvisoflavone A	42.63
3	5,2',4'-trihydroxy-7-methoxy-6-prenylisoflavone	42.62
4	Licoisoflavone B	42.61
5	Genistein	42.57
6	Alpinumisoflavone	42.56
7	Biochanin A	42.56
8	Wighteone	42.56
9	Licoisoflavone	42.53
10	Lupiwighteone	42.49

Table 3. Validation of molecular docking results and parameters.

Ligand	Binding Energy (kcal/mol)	Ki (nM)	RMSD (Å)	Gridbox	Gridbox size
TAK-285 (Native ligand)	−10.06	42.55	1.27	X = 18.284 Y = 31.913 Z = 11.626	32×32×32
Erlotinib (Comparative drug)	−9.66	83.43	34.80		

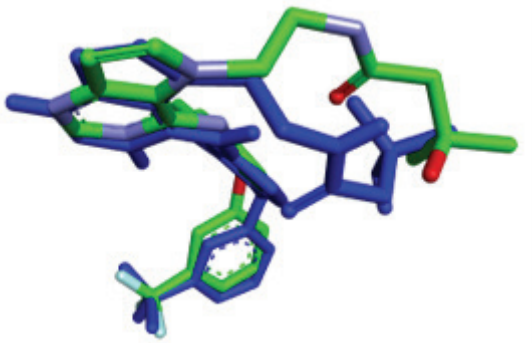


Figure 3. Overlay visualization of the native ligand (green) and redocked ligand (blue) within the wild-type-EGFR binding site.

binding pocket, further confirming the reliability of the docking protocol and supporting the obtained RMSD value of 1.27 Å.

Molecular Docking

The molecular docking results of the top 10 compounds with the most favorable binding energies among the 61 compounds subjected to molecular docking against the EGFR receptor are summarized in Table 4. Although Parvisoflavone A was not included among the top 10 compounds presented in the molecular docking table, its docking results were considered in the subsequent candidate selection. The complete molecular docking results are provided in the Supplementary Data. Binding affinity was evaluated based on binding energy and inhibition constant (Ki), where lower binding energy values indicate stronger and more stable ligand-receptor interactions (Morris & Lim-Wilby, 2008). Among the tested compounds, Glyinflanin D exhibited the lowest binding energy (−10.56 kcal/mol) and Ki value of 18.06 nM which demonstrated binding affinities comparable to the reference inhibitor erlotinib.

Based on the molecular docking and pharmacokinetic evaluation results, Glyinflanin D and Parvisoflavone A were selected as the most promising compounds for further interaction visualization analysis. Although several compounds exhibited lower binding energy values, the selection of these compounds was also supported by their favorable ADMET characteristics, non-mutagenic and non-carcinogenic toxicity profiles, acceptable Lipinski's Rule of Five parameters, and high pharmacophore fit scores. Glyinflanin D demonstrated high intestinal absorption (96.27%),

Table 4. Molecular docking results.

No	Compound	Binding Energy (kcal/mol)	Ki (nM)	Important Amino Acid Residues	Amino Acid Residues Interaction		
					Hydrogen Bond	Van der Waals Bond	Other Bonds
1	Glyinflanin D	−10.56	18.06	Asp A855, Met A793, Cys A797, Thr 790	Gln A791 Met A793	Thr A790 Leu A792 Gly A796 Cys A797 Asn A842 Leu A718 Thr A854 Asp A855	**Pi-Sigma** Cys A775 **Alkyl/Pi-Alkyl** Leu A788 Ala A743 Arg A841 Val A726 Lys A745 Leu A844
2	Licoricidin	−9.96	49.87	Asp A855, Cys A797, Thr 790	Asp A855 **Carbon H-Bond** ASN A842	Ile A789 Ala A743 Cys A797 Ser A720 Gly A719 Thr A854 Leu A858 Arg A776 Met A766 Leu A788	ASP A855 **Carbon H-Bond** ASN A842
3	Kanzonol B	−9.86	59.25	Asp A855, Met A793	Asp A855 Phe A856	Leu A792 Ala A743 Met A793 Leu A844 Leu A718 Leu A788 Leu A777	**Pi-Sigma** Val A726 Met A766 **Alkyl/Pi-Alkyl** Leu A792 Ala A743 Met A793 Leu A844 Leu A718 Leu A788 Leu A777 **Pi-Pi T-shaped** Phe A856
4	Gancaonin H	−9.69	78.88	Asp A855, Cys A797, Thr 790	Asn A842 Lys A745 Asp A855	Val A726 Leu A844 Thr A790 Cys A775 Arg A776 Leu A788 Leu A858	**Pi-Sigma** Phe A856 **Alkyl/Pi-Alkyl** Cys A797 Arg A841 Leu A777 Met A766 Phe A856 **Pi-Lone Pair** Thr A854
5	Ononin	−9.68	80.86	Asp A855, Cys A797	Thr A854 Phe A856 Thr A854	Leu A858 Leu A788 Asp A855 Thr A780 Gly A719 Ser A720 Cys A797 Leu A844 Arg A841 Arg A776 Cys A775 Met A766	**Pi-Alkyl** Ala A743 Lys A745 Val A726

Table 4. Molecular docking results (continuous)

No	Compound	Binding Energy (kcal/mol)	Ki (nM)	Important Amino Acid Residues	Amino Acid Residues Interaction		
					Hydrogen Bond	Van der Waals Bond	Other Bonds
6	Shinpterocarpin	−9.58	94.52	Asp A855, Thr 790	Met A766 Asp A855	Val A769 Arg A776 Cys A775 Thr A790 Ala A743 Thr A854 Leu A858	**Pi-Sigma** Leu A777 Met A766 **Alkyl/Pi-Alkyl** Lys A745 Leu A844 Val A726 Leu A788 Met A766 Leu A777 **Pi-Pi T-shaped** Phe A856
7	Glycyuralin A	−9.54	101.44	Asp A855, Thr 790	Lys A745 Asn A842 **Carbon H-Bond** Leu A788	Arg A776 Arg A841 Asp A855	**Pi-Sigma** Thr A790 Thr A854 **Alkyl/Pi-Alkyl** Phe A856 Ala A743 Leu A844 Val A726 Leu A777 Met A766 Cys A775 LYS A745 **Pi-Sulfur** Cys A775
8	Euchrenone A5	−9.51	106.16	Asp A855, Met A793, Cys A797, Thr 790	Met A793	Gln A791 Leu A792 Gly A796 Cys A797 Asn A842 Thr A854 Asp A855	**Pi-Sigma** Leu A718 Thr A790 **Alkyl/Pi-Alkyl** Leu A777 Leu A788 Leu A858 Leu A844 Ala A743 Val A726 Lys A745 Met A766
9	Licochalcone D	−9.49	111.34	Asp A855, Met A793, Thr 790	Met A793 **Carbon H-Bond** Gly A796 Gln A791	Leu A1001 Leu A718 Val A726 Thr A854 Asp A855 Leu A788 Ile A789 Thr A790 Leu A792	**Pi-Sigma** Leu A844 **Alkyl/Pi-Alkyl** Leu A777 Met A766 Phe A856 Cys A775 Ala A743 Leu A844 Met A793 Lys A745

Table 4. Molecular docking results (continuous)

No	Compound	Binding Energy (kcal/mol)	Ki (nM)	Important Amino Acid Residues	Amino Acid Residues Interaction		
					Hydrogen Bond	Van der Waals Bond	Other Bonds
10	1-methoxyphaseollin	−9.49	110.67	Asp A855, Thr 790	Met A766 Phe A856 Asp A855	Arg A841 Leu A844 Thr A790 Val A769 Arg A776 Cys A775 Thr A854 Asn A842	**Pi-Sigma** Met A766 **Alkyl/Pi-Alkyl** Val A726 Leu A777 Leu A788 Leu A858 Phe A856 Met A766 **Pi-Cation** Lys A745 **Pi-Pi T-Shaped** Phe A856

favorable Caco-2 permeability (47.44 nm/sec), and negative mutagenicity and carcinogenicity predictions, indicating good drug-likeness and safety profiles. Meanwhile, Parvisoflavone A exhibited a high pharmacophore fit score (42.63%), good absorption characteristics, and favorable toxicity prediction results, suggesting strong compatibility with the wild-type-EGFR binding pocket.

The chemical structures and three-dimensional binding interactions of Glyinflanin D and Parvisoflavone A are presented in Figures 4 and 5, respectively. The structural features of these compounds may contribute to their accommodation within the ATP-binding pocket of wild-type EGFR. The EGFR ATP-binding pocket contains several functionally important residues, including Lys745, Thr790, Met793, Cys797, and Asp855, which contribute to ATP recognition, kinase activity, and

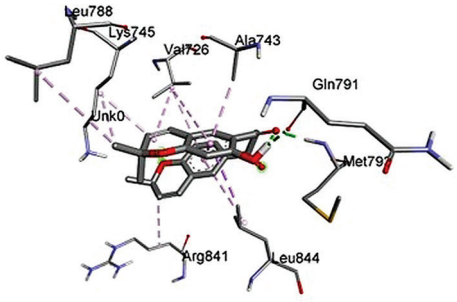


Figure 4. 3D diagram of Glyinflanin D on wild-type-EGFR active sites.

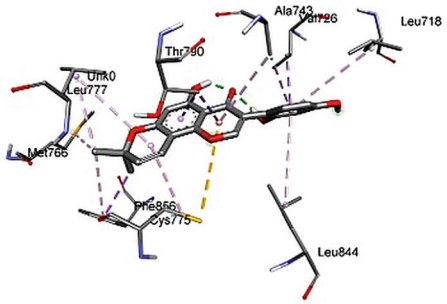


Figure 5. 3D diagram of Parvisoflavone A on wild-type-EGFR active sites.

inhibitor binding. Met793 is located in the hinge region and participates in hydrogen bonding with ATP, Thr790 functions as the gatekeeper residue, Lys745 serves as the conserved catalytic lysine, Cys797 is positioned at the edge of the ATP-binding cleft, and Asp855 forms part of the conserved DFG motif involved in kinase catalysis (Damghani, *et al*., 2023). Glyinflanin D formed interactions with key residues including Met793, Thr790, Cys797, and Asp855, whereas Parvisoflavone A interacted with Lys745, Asp855, and Thr790. These residues are located within or surrounding the ATP-binding site and are frequently involved in the binding of EGFR inhibitors (Niggenaber, 2020). The observed interactions suggest that the structural features of these compounds may contribute to their favorable positioning within the EGFR ATP-binding pocket.

Indonesian Journal of Cancer Chemoprevention, October 2025
ISSN: 2088–0197
e-ISSN: 2355-8989

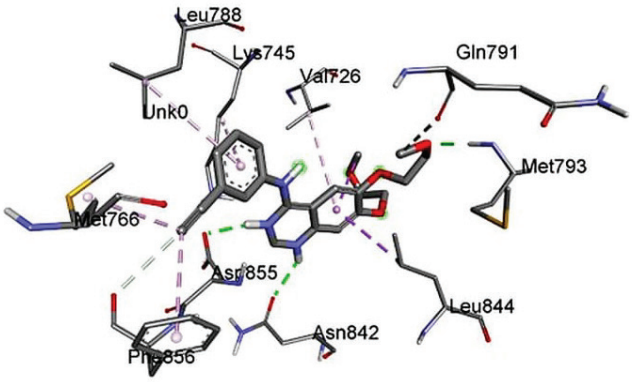


Figure 6. 3D diagram of erlotinib on wild-type-EGFR active sites.

Figure 6 presents the three-dimensional interaction visualization of Erlotinib as the comparative inhibitor within the wild-type-EGFR active site. Erlotinib exhibited a binding energy of −9.66 kcal/mol and interacted with several important amino acid residues surrounding the ATP-binding pocket, including Lys A745, Thr A790, Asp A855, and Phe A856. Comparison of the interaction profiles revealed that both Glyinflanin D and Parvisoflavone A occupied binding regions similar to erlotinib and shared several key amino acid interactions within the active site. The similarity of these interaction patterns suggests that the selected compounds may possess inhibitory mechanisms comparable to the reference EGFR inhibitor.

DISCUSSION

Lipinski screening serves as an early stage assessment of drug-likeness by estimating the likelihood of adequate absorption and membrane permeability. Compounds with molecular weights greater than 500 Da, >5 hydrogen-bond donors and >10 hydrogen-bond acceptors, or high lipophilicity are generally associated with poor oral bioavailability and were therefore considered less favorable for further development. A total of 184 compounds identified from *Glycyrrhiza inflata* were initially screened using Lipinski's Rule of Five to evaluate their suitability as orally-administered drug candidates. Most flavonoids, isoflavonoids, and chalcone derivatives passed Lipinski's criteria, whereas many glycosylated flavonoids and triterpenoid saponins exhibited multiple violations. These findings are consistent with the presence of multiple sugar moieties, which increase molecular size and polarity which may hinder passive diffusion across biological membranes.

Although Lipinski's Rule of Five provides physicochemical properties of oral drug-likeness, these criteria does not solely guarantee favorable pharmacokinetic behaviour *in vivo*. Therefore, compounds that passed the initial drug-likeness screening were further evaluated through ADMETox analysis to obtain their pharmacokinetic safety profiles. Several compounds demonstrated predicted intestinal absorption values exceeding 90%, suggesting favorable oral bioavailability. Conversely, BBB permeability was not considered a critical determinant of candidate selection because EGFR-driven NSCLC primarily affects peripheral lung tissue rather than the central nervous system. Mutagenicity and carcinogenicity predictions were also incorporated into the selection process because compounds exhibiting unfavorable safety profiles may have limited therapeutic potential.

While ADMET screening helps identify favorable pharmacokinetic behaviour and safety, it does not provide information regarding their ability to interact with the intended molecular target. As

a ligand-based drug design (LBDD) approach, pharmacophore modelling focuses on the spatial arrangement of molecular features that share key chemical characteristics with known EGFR inhibitors. Prior to virtual screening, the quality of the generated pharmacophore models were evaluated using receiver operating characteristic (ROC) curve analysis. The resulting AUC values was 0.82, indicating excellent discrimination between active and decoy compounds and a high probability that active compounds are ranked above inactive molecules. These results suggest that the essential chemical features required for EGFR inhibition from *Glycyrrhiza inflata* were successfully captured. The high fit scores obtained by Angustone B, Parvisoflavone A, and related isoflavonoids suggest that these compounds mimic the essential recognition pattern of known EGFR inhibitors. Higher fit scores indicate greater similarity between the spatial and chemical characteristics of a candidate compound.

From the ligand-based perspective, Angustone B and Parvisoflavone A emerged as the most promising compounds due to their superior pharmacophore fit scores. High fit scores indicate substantial overlap between the spatial and chemical features of a candidate compound and the validated pharmacophore model. Consequently, these compounds are predicted to possess structural characteristics similar to those of established EGFR inhibitors. Interestingly, both compounds also maintained interactions with Lys745 and Asp855 during docking analysis, indicating that the structural features responsible for their high pharmacophore scores translate into productive interactions within the receptor binding pocket. This convergence between ligand-based and structure-based approaches substantially increases confidence in their predicted activity because two independent computational methodologies identify the same compounds as promising candidates.

Parvisoflavone A deserves particular attention because it demonstrated a balanced profile across the screening workflow. In addition to exhibiting one of the highest pharmacophore fit scores, the compound maintained predicted interactions with Lys745, Asp855, and Thr790 while satisfying drug-likeness criteria and displaying favorable predicted ADMET properties. Lys745 contributes to ATP positioning and catalytic activity, while Asp855, as part of the conserved DFG motif, plays an important role in protein catalysis and kinase conformational states (Amelia, *et al*., 2022). Based on these established roles, the predicted interaction of Parvisoflavone A with Lys745 may interfere with ATP accommodation, whereas its interaction with Asp855 may contribute to ligand positioning within the kinase domain. However, these interpretations remain hypotheses based on the predicted docking pose, and their functional consequences cannot be established from molecular docking alone. The predicted interaction with Thr790 may further support accommodation within the ATP-binding cleft. Although its docking score was slightly lower than those of Angustone B and Glyinflanin D, its overall balance of receptor affinity, pharmacophore compatibility, drug-likeness, and predicted pharmacokinetic properties makes Parvisoflavone A an attractive candidate for further investigation.

In addition to ligand-based screening, a structure-based drug design (SBDD) approach was conducted through molecular docking to directly evaluate ligand interactions within the three-dimensional structure of the target protein. Molecular docking was performed using the crystal structure of wild-type epidermal growth factor receptor (wild-type-EGFR) kinase domain (PDB ID: 3POZ). The selection of wild-type-EGFR as the molecular target is highly relevant in the context of non-small cell lung cancer (NSCLC), as EGFR signaling remains one of the most important drivers of tumor proliferation and survival. EGFR belongs to the ErbB family of receptor tyrosine kinases and functions as a transmembrane signaling protein that regulates numerous cellular processes

including proliferation, differentiation, migration, angiogenesis, and apoptosis. Under physiological conditions, activation of EGFR occurs following ligand binding, which induces receptor dimerization and subsequent autophosphorylation of intracellular tyrosine residues. These phosphorylation events initiate a cascade of downstream signaling pathways, particularly the PI3K/AKT, RAS/RAF/MEK/ERK, and JAK/STAT pathways, ultimately promoting cellular growth and survival. However, in many NSCLC cases, EGFR is overexpressed or aberrantly activated, resulting in persistent signaling that drives uncontrolled cell proliferation, resistance to apoptosis, enhanced angiogenesis, and metastatic progression. Consequently, inhibition of the EGFR kinase domain remains a well-established therapeutic strategy aimed at disrupting these oncogenic signaling networks and suppressing tumor progression.

The ATP-binding pocket of wild-type-EGFR contains several highly conserved amino acid residues that are essential for kinase function and inhibitor recognition. Among these residues, Lys745 plays a crucial catalytic role by stabilizing ATP within the active site and facilitating phosphate transfer during kinase activation. Met793 forms part of the hinge region connecting the N-terminal and C-terminal lobes of the kinase domain and serves as a key anchoring residue for many ATP-competitive inhibitors. Hydrogen bonding with Met793 is frequently observed among clinically approved EGFR tyrosine kinase inhibitors and is often considered a hallmark of productive binding. Cys797 is located adjacent to the hinge region and contributes significantly to ligand stabilization within the binding cavity. Furthermore, several irreversible EGFR inhibitors exploit this residue as a covalent attachment site. Asp855, a component of the highly conserved DFG motif, plays a central role in kinase activation and conformational regulation, whereas Thr790 contributes to the architecture of the ATP-binding cleft and influences ligand accommodation within the pocket. Therefore, interactions involving Lys745, Met793, Cys797, Asp855, and Thr790 were considered important indicators of potential wild-type-EGFR inhibitory activity.

The docking protocol was first validated using Erlotinib, a first-generation EGFR tyrosine kinase inhibitor that served as the reference ligand. Erlotinib exhibited a binding energy of −9.66 kcal/mol and established interactions with Met793, Gln791, Thr790, Cys797, and Asp855. These interactions indicate stable occupancy of the ATP-binding pocket and are consistent with the established mechanism of Erlotinib as a reversible ATP-competitive inhibitor. The interaction with Met793 is particularly noteworthy because the hinge region represents the primary anchoring site responsible for maintaining inhibitor positioning within the catalytic cleft. Likewise, interactions involving Asp855 and Cys797 contribute to stabilization of the inhibitor-receptor complex and support efficient competition with endogenous ATP. Consequently, the interaction profile of Erlotinib provides an important benchmark against which the phytochemicals identified from *Glycyrrhiza inflata* can be evaluated.

Among the screened compounds, Glyinflanin D demonstrated the strongest docking affinity, surpassing even the reference ligand. The compound maintained interactions with Met793, Thr790, Cys797, and Asp855, indicating extensive engagement with critical residues throughout the ATP-binding pocket. The retention of Met793 interaction suggests effective anchoring within the hinge region, a characteristic commonly associated with potent kinase inhibitors. Simultaneously, interactions with Cys797 and Asp855 indicate favorable positioning within the catalytic cleft and may facilitate interference with ATP binding and kinase activation. From a mechanistic perspective, occupation of these residues could prevent ATP from accessing the catalytic site, thereby suppressing receptor autophosphorylation and subsequent downstream signaling. Because

autophosphorylation serves as the initiating event for activation of the PI3K/AKT and MAPK pathways, inhibition at this stage may effectively disrupt the proliferative signaling network that sustains NSCLC growth. The superior docking affinity observed for Glyinflanin D therefore suggests a strong theoretical capacity to inhibit wild-type-EGFR activity. Nevertheless, despite its favorable docking profile, consideration of its ADMET characteristics remains necessary before advancing this compound for further development. Another notable candidate was Gancaonin H, which exhibited a docking score comparable to Erlotinib while maintaining interactions with Lys745, Asp855, Thr790, Cys797, and Phe856. The interaction with Lys745 is particularly significant because this residue is directly involved in ATP stabilization and catalytic phosphate transfer. Occupation of the region surrounding Lys745 may therefore interfere with ATP positioning and reduce kinase activity. The simultaneous preservation of interactions with Asp855 and Cys797 further supports stable occupancy within the catalytic cleft. Interestingly, Gancaonin H appears to adopt a binding orientation that differs from Erlotinib while still engaging many of the same functionally relevant residues. This observation suggests that the compound may inhibit wild-type-EGFR through an alternative binding mode while maintaining sufficient interactions to disrupt kinase activity. Such alternative binding orientations may be advantageous because they expand the range of molecular interactions available for inhibitor optimization and may improve resilience against structural variations within the kinase domain.

Collectively, the integration of pharmacophore modeling and molecular docking indicates that several *Glycyrrhiza inflata* constituents possess the structural and physicochemical characteristics required for wild-type-EGFR inhibition. The ability of these compounds to occupy the ATP-binding pocket and engage catalytically important residues such as Lys745, Met793, Cys797, Asp855, and Thr790 suggests a potential mechanism involving competitive inhibition of ATP binding and suppression of receptor autophosphorylation. Such inhibition would be expected to attenuate downstream signaling through the PI3K/AKT, MAPK, and JAK/STAT pathways, thereby reducing cellular proliferation and survival in EGFR-driven NSCLC.

Among the evaluated compounds, Parvisoflavone A appears to exhibit the most favorable balance between pharmacophore alignment, receptor binding, drug-likeness, and predicted pharmacokinetic behavior. However, several considerations should be taken into account when interpreting these findings. All results were derived from computational approaches and have not yet been experimentally validated. Although the pharmacophore model showed an ROC-AUC of 0.82, its limited selectivity, reflected by the relatively high number of false positives and low specificity, should be considered. In addition, molecular docking provides a static representation of ligand–receptor interactions and cannot confirm binding stability, conformational changes, or functional EGFR inhibition.

Therefore, Parvisoflavone A should be considered a potential hit or early lead candidate rather than a confirmed EGFR inhibitor. Future studies involving molecular dynamics simulations, enzymatic kinase assays, and NSCLC cell-based experiments will be necessary to validate the stability and biological significance of the predicted interactions and further assess its therapeutic potential.

CONCLUSION

In conclusion, redocking of the native ligand (TAK-285) yielded a root mean square deviation (RMSD) of 1.27 Å, confirming the reliability of the docking protocol. Among the phytochemicals from *Glycyrrhiza inflata*, Parvisoflavone A exhibited a binding energy of −8.91 kcal/mol and an inhibition

constant (Ki) of 292.70 nM. Parvisoflavone A also showed the highest pharmacophore fit score (42.63%) and favorable interactions with the key catalytic residues Lys745, Asp855, and Thr790. These results identify Parvisoflavone A as the most promising hit compound for further development as a wild-type EGFR inhibitor against NSCLC. Nevertheless, these findings remain computational predictions. The activity, binding stability, and pharmacological potential of Parvisoflavone A require further validation through molecular dynamics simulations and experimental studies.

SUPPLEMENTARY

Supplementary material: https://bit.ly/4iQrhUK

ACKNOWLEDGEMENT

The authors would like to thank the Faculty of Pharmacy, Universitas Padjadjaran, for providing academic support throughout this study.

REFERENCES

Amelia, T., Kartasasmita, R.E., Ohwada, T., and Tjahjono, D.H., 2022, Structural Insight and Development of EGFR Tyrosine Kinase Inhibitors, *Molecules*, **27**(3), 819.

Damghani, T., Wittlinger, F., Beyett, T.S., Eck, M.J., Laufer, S.A., and Heppner, D.E., 2023, Structural elements that enable specificity for mutant EGFR kinase domains with next-generation small-molecule inhibitors, *Methods in enzymology*, **685**, 171–198.

Deng, N., Qiao, M., Li, Y., Liang, F., Li, J., and Liu, Y., 2023, Anticancer effects of licochalcones: A review of the mechanisms, *Frontiers in Pharmacology*, **14**, 1074506.

Friedlaender, A., Perol, M., Banna, G.L., Parikh, K., and Addeo, A., 2024, Oncogenic alterations in advanced non-small cell lung cancer: A molecular super-highway, *Biomark. Res.*, **12**(1), 24.

Hamza, A., Wei, N.N., and Zhan, C.G., 2012, Ligand-Based Virtual Screening Approach using a New Scoring Function, *Journal of Chemical Information and Modeling*, **52**(4), 963–974.

Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J., 2001, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.*, **46**(1–3), 3–26.

Morris, G.M. and Lim-Wilby, M., 2008, Molecular docking, in Kukol, A. (eds) Molecular Modeling of Proteins, *Methods Molecular Biology*, **443**, 365–382.

Niggenaber, J., Heyden, L., Grabe, T., Müller, M.P., Lategahn, J., and Rauh, D., 2020, Complex Crystal Structures of EGFR with Third-Generation Kinase Inhibitors and Simultaneously Bound Allosteric Ligands, *ACS medicinal chemistry letters*, **11**(12), 2484–2490.

Oh, H.-N., Lee, M.-H., Kim, E., Yoon, G., Chae, J.-I., and Shim, J.-H., 2019, Licochalcone B inhibits growth and induces apoptosis of human non-small-cell lung cancer cells by dual targeting of EGFR and MET, *Phytomedicine*, **63**, 153014.

Padinharayil, H., Varghese, J., John, M. C., Rajanikant, G.K., Wilson, C.M., Al-Yozbaki, M., *et al.*, 2023, Non-small cell lung carcinoma (NSCLC): Implications on molecular pathology and advances in early diagnostics and therapeutics, *Genes & Diseases*, **10**(3), 960–989.

Sattler, M., Mambetsariev, I., Fricke, J., Tan, T., Liu, S., Vaidehi, N., *et al.*, 2023, A Closer Look at EGFR Inhibitor Resistance in Non-Small Cell Lung Cancer through the Lens of Precision Medicine, *J. Clin. Med.*, **12**(5), 1936.

Shaban, N., Kamashev, D., Emelianova, A., and Buzdin, A., 2024, Targeted Inhibitors of EGFR: Structure, Biology, Biomarkers, and Clinical Applications, *Cells*, **13**(1), 47.

Sharifi-Rad, J., Quispe, C., Herrera-Bravo, J., Herrera-Belén, L., Kaur, R., Kregiel, D., *et*

al., 2021, *Glycyrrhiza* Genus: Enlightening Phytochemical Components for Pharmacological and Health-Promoting Abilities, *Oxid. Med. Cell. Longev.*, **2021,** 7571132.

Sung, H., Filho, A.M., Laversanne, M., Ferlay, J., Siegel, R.L., Soerjomataram, I., *et al.*, 2026, Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries, *CA: a cancer Journal for Clinicians*, **76**(4), e70090.

Wang, C., Chen, L., Xu, C., Shi, J., Chen, S., Tan, M., *et al.*, 2020, A Comprehensive Review for Phytochemical, Pharmacological, and Biosynthesis Studies on *Glycyrrhiza* spp, *The American journal of Chinese medicine*, **48**(1), 17–45.

Wang, Z., Xing, Y., Li, B., Li, X., Liu, B., and Wang, Y., 2022, Molecular pathways, resistance mechanisms and targeted interventions in non-small-cell lung cancer, *Mol. Biomed.*, **3**, 42.

Wu, L., Ma, T., Zang, C., Xu, Z., Sun, W., Luo, H., *et al.*, 2025, *Glycyrrhiza,* a commonly used medicinal herb: Review of species classification, pharmacology, active ingredient biosynthesis, and synthetic biology, *Journal of Advanced Research*, **75,** 249–270.

Yang, S.Y., 2010, Pharmacophore modeling and applications in drug discovery: Challenges and recent advances, *Drug Discovery Today*, **15**(11-12), 444–450.