Natural Product Research
Formerly Natural Product Letters

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gnpl20

In silico analysis, isolation, and cytotoxicity evaluation of the coumestans from *Cullen corylifolium* (L.) Medik

Nancy Tripathi, Himisa Shah, Nivedita Bhardwaj, Ruma Sarkar & Shreyans K. Jain

To cite this article: Nancy Tripathi, Himisa Shah, Nivedita Bhardwaj, Ruma Sarkar & Shreyans K. Jain (27 Nov 2023): *In silico* analysis, isolation, and cytotoxicity evaluation of the coumestans from *Cullen corylifolium* (L.) Medik, Natural Product Research, DOI: 10.1080/14786419.2023.2285875

To link to this article: https://doi.org/10.1080/14786419.2023.2285875

NATURAL PRODUCT RESEARCH
https://doi.org/10.1080/14786419.2023.2285875

Taylor & Francis
Taylor & Francis Group

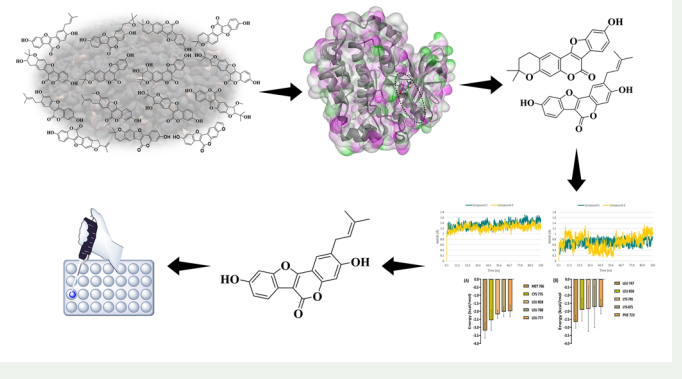
In silico analysis, isolation, and cytotoxicity evaluation of the coumestans from *Cullen corylifolium* (L.) Medik

Nancy Tripathi[a], Himisa Shah[b], Nivedita Bhardwaj[a], Ruma Sarkar[b] and Shreyans K. Jain[a]

[a]Department of Pharmaceutical Engineering & Technology, Indian Institute of Technology (Banaras Hindu University), Varanasi, India; [b]B.D. Patel Institute of Paramedical Sciences, Charotar University of Science & Technology, CHARUSAT Campus, Gujarat, India

ABSTRACT

Cullen corylifolium is well known for diverse phytoconstituents that possess multifaceted pharmacology, and one such less explored class is coumestans, which have not been well explored for their anticancer activities. One of the popular cancer targets is the Epidermal Growth Factor Receptor, a tyrosine kinase involved in various cancers, especially breast and lung cancer hence, a crucial cancer target. This work is focussed on molecular docking and molecular simulation studies on coumestans against EGFR. The rigorous docking studies resulted in two coumestans (**1** and **5**) with binding energy less than Gefitinib and Erlotinib. Compounds **1** and **5** were subjected to molecular simulation, binding free energy calculation, per-residue energy decomposition, and *in silico* ADMET prediction. The best hit, compound **1** was evaluated for its cytotoxicity against MDA-MB-231 and A549 cells *via in vitro* assay. The ligand-protein complex exhibited good stability, binding free energies, better *in silico* pharmacokinetics, low toxicity, and good cytotoxicity.



ARTICLE HISTORY
Received 3 September 2023
Accepted 12 November 2023

KEYWORDS
Cullen corylifolium; coumestans; EGFR; molecular docking; molecular simulation

CONTACT Shreyans Kumar Jain sjain.phe@iitbhu.ac.in
Supplemental data for this article can be accessed online at https://doi.org/10.1080/14786419.2023.2285875.

1. Introduction

Nature being highly evolved have remarkable ability to interact with enzymes and receptors thus, have played a pivotal role in drug discovery (Atanasov et al. 2021). One such lesser-explored class is 'Coumestans or Phytoestrogens', with polycyclic aromatic nature. They possess coumestan framework consisting of four-ring oxygen heterocycle system linked to a coumarin and a benzofuran unit through a C=C bond (Figure 1) (Tu et al. 2021). Natural coumestans are reported to exhibit oestrogenic, antimicrobial, antifungal, antioxidative, anti-osteoporotic, anti-inflammatory,

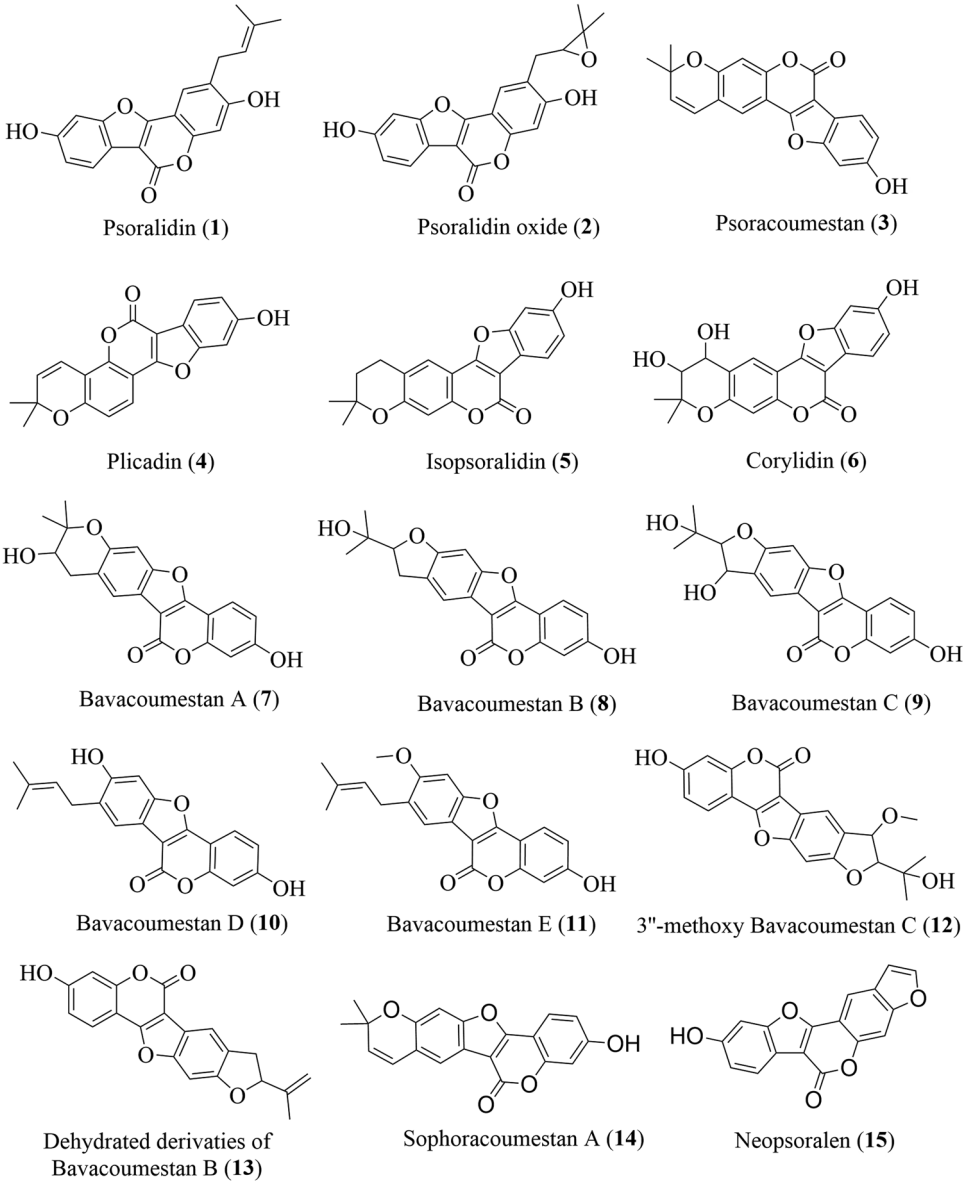


Figure 1. Chemical structures of coumestans present in *C. corylifolium*.

antihemorrhagic, antimyotoxic, hepatoprotective, antifibrotic, antiproteolytic, anti-diabetic, anticancer, neuroprotective, and immunomodulatory properties (Dutta et al. 2021; Tu et al. 2021).

Legumes are notably rich in coumestans and many coumestans has been successfully extracted from Leguminosae/Fabaceae family (Tu et al. 2021). Within this family, *Cullen corylifolium* (L.) Medik. (syn. *Psoralea corylifolia* L., Fabaceae) is a rich source of coumestans (Chopra et al. 2013) that are concentrated primarily in seeds. Their oestrogenic properties support the potential role of *C. corylifolium* in osteoporosis and cardiovascular diseases (Liu et al. 2014; Qi et al. 2023). The coumestan of *C. corylifolium* includes psoralidin, psoralidin-2',3'-oxide, psoracoumestan, plicadin, etc. (Figure 1) and they still remain largely unexplored for their pharmacological potential (Zhang et al. 2016).

Epidermal Growth Factor Receptor (EGFR) is ubiquitously expressed multifunctional signal transducer. It orchestrates various cellular processes including migration, proliferation, cell fate determination, and apoptosis (Sigismund et al. 2018). It binds to various growth factor ligands like EGF, betacellulin, amphiregulin, epiregulin, epigen etc. and trigger responses (Uribe et al. 2021). Upon ligand binding, EGFR undergoes dimerisation leading to autophosphorylation and triggers resultant signalling pathways like PI3 kinase, Ras-Raf-MAPK, JNK, and PLCγ (Masuda et al. 2012). ErbB1, the EGFR family representative, is linked to cancer development through mutations or overexpression. EGFR overexpression are common in various cancers, notably breast and lung cancers, making it a focal point for several cancer treatments (Sigismund et al. 2018).

EGFR overexpression is a key therapeutic target in breast cancer, causing poor tumour differentiation, larger tumour sizes, and disruption of EGFR pathways (Masuda et al. 2012). Across all breast cancer subtypes, its overexpression is more prevalent in triple-negative breast cancer (TNBC) and hence, its negative impact is more apparent in TNBC. Hence, EGFR a promising target for TNBC treatment (Masuda et al. 2012).

EGFR growth factors play a crucial role in lung cancer initiation, progression, and metastasis affecting both non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). Mutations in EGFR, damage signalling pathways culminating in aggressive and metastatic lung adenocarcinoma. EGF upregulation is observed in lung cancer, while amphiregulin upregulation is linked to poor prognosis and reduced survival rates in NSCLC (Liu et al. 2017).

Molecular docking is a vital method to explore the interaction of ligands with protein in a computational setting (Meng et al. 2011) and molecular dynamics (MD) helps in gaining deeper insights into ligand-protein interactions within a simulated biological environment. In this study, we employed *in silico* approaches to assess the anticancer capabilities of coumestans of *C. corylifolium* against EGFR. It involved rigorous molecular docking and MD simulations, followed by *in silico* estimations of pharmacokinetics. The results were validated by focussed isolation of identified lead from *C. corylifolium* and assessing its cytotoxic effects on TNBC and NSCLC cell lines. This study offers an opportunity to delve into the extensive anticancer potential of coumestans from *C. corylifolium*, potentially leading to the discovery of novel EGFR inhibitors derived from this plant source.

2. Result and discussion

2.1. Ligand preparation, protein-energy minimisation and protein preparation

The energy minimisation of ligands resulted in the optimisation of the bond length and bond angle before docking study. The GAFF in Open Babel 3.1.1 employs simple harmonic form for optimising angles and bonds of the molecules, and it covers most of the organic chemical space. Protein-energy minimisation is crucial for global minimisation in docking and simulation studies and it was done using Amber 20 module. This step resulted in the potential energy minimisation of protein from −137060 kcal/mol to −204320 kcal/mol (Supplementary material Figure S1). Further, the protein was processed to add polar hydrogens, merge non-polar hydrogens, assign atom types, and Gastieger charge in AutoDock Tools 1.5.6 and converted to PDBQT format.

2.2. Molecular docking studies

The molecular docking approach has remained an important tool in the drug discovery process during the initial screening of potential candidates (Meng et al. 2011). The co-crystallized ligand displayed various interactions with Phe 723, Lys 745, Glu 749, Arg 776, Leu 788, Met 790, Asp 855, and Leu 858 (Supplementary material Figure S2). These residues were used to ascertain grid box size (Morris et al. 2009). The grid validation between docked ligand and co-crystalised ligand showed RMSD of 0.3409 Å (Supplementary material Figure S3) that shows less deviation in the selected grid and thus, can be used for docking studies.

After completion of docking studies, the interactions of all ligands were visualised in Discovery Studio Visualiser 2021. Total 15 coumestans of *C. corylifolium* along with Gefitinib and Erlotinib were docked to EGFR, and all possessed binding energy less than −8 kcal/mol. The binding energy cut-off was set to −10 kcal/mol and two ligands with energy less than this cut-off were selected for further studies. The binding energy of ligand **1** and ligand **5** was better than Gefitinib and Erlotinib (Supplementary material Table S1). The 2D and 3D interaction diagrams of ligands **1** and ligand **5** with EGFR have been shown in Supplementary Material Figure S4. The binding energies, ligand efficiency, and interaction diagram of rest of the ligands have been provided in Supplementary material Table S1 and Figure S5.

2.3. In silico ADME and toxicity prediction

ADMET profiling constitutes an important part of a compound's clinical efficacy. The PreADMET server provides a user-friendly interface to predict ADMET parameters. Since bioavailability and drug response depend on absorption, a drug must have Human intestinal absorption (HIA) above 70% for good absorption. The other important criteria is Blood-Brain Barrier (BBB) permeability where a range of 0.1 to 2.0 for predicted BBB is an indicator of moderate penetration through BBB (Ma et al. 2005). Another crucial factor to be considered is carcinogenicity and it must not possess carcinogenicity. One overlooked factor is human ether-a-go-go-related gene (hERG) inhibition that leads to cardiac dysfunctions (Danker and Möller 2014). One desirable

trait of a drug candidate is that it should not inhibit liver microenzymes. Compounds **1** and **5** possessed good *in silico* ADMET properties summarised in Supplementary material Table S2. The ADMET properties of the rest of the compounds are also provided in Supplementary Material Table S2.

2.4. Molecular dynamics simulation study

MD simulation is an advanced technique simulating systems with biological relevance and involves studying the trajectories and motion of the molecules in presence of other complexes over time, thus aids in studying the conformational changes occurring in the molecules (Hospital et al. 2015). It also gives information regarding the structural features of protein within the system and ligand-protein interactions. Here two ligand-protein complexes-Compound **1**-EGFR and Compound **5**-EGFR were studied through MD simulations using Pmemd.cuda module of the Amber 20 software package (Case et al. 2020).

The information about structural conformations of protein and stability of ligand with reference to protein during the simulation run is obtained from protein RMSD and ligand RMSD, respectively. The RMSD plot of protein-ligand complexes is shown in Supplementary material Figure S6. The ligand showed protein backbone stability during the simulation run with an average protein RMSD of 1.34 and 1.23 Å for Compound **1**-EGFR and Compound **5**-EGFR, respectively. The stability of the ligand was estimated by calculating the RMSD values of heavy atoms present in the ligands, and the values demonstrated that the ligand stands stable throughout the simulation time (Supplementary material Figure S6). The average ligand RMSD for Compound **1**-EGFR and Compound **5**-EGFR was 0.68 and 0.71 Å, respectively.

The RMSF calculation is useful for determining fluctuations in the amino acid residues along the protein chain. The peaks in the RMSF plot are residues that fluctuate most throughout the simulation. In the study, the RMSF analysis was done for C-α atoms of residues, and the RMSF plot (Supplementary material Figure S7) obtained showed stability in the secondary conformation of the protein at the time of simulation. The average RMSF values of 0.92 and 0.91 Å were obtained for Compound **1**-EGFR and Compound **5**-EGFR, respectively. The high fluctuations at the C-terminal region and N-terminal region are inevitable and hence, could be seen in the current plot also. The active site residues showed apparently lower values of RMSF and so possessed good stability during the run. The RMSF analysis showed both complexes were stable throughout the MD simulation.

The apparent protein folding or unfolding at the time of simulation run is determined *via* radius of gyration (RoG) analysis. The average values of RoG were found to be 19.73 ± 0.06 Å for Compound **1**-EGFR and 19.62 ± 0.07 Å for Compound **5**-EGFR (Supplementary material Figure S8).

Throughout the simulation, protein interactions with the ligand were observed. Understanding the stability of the projected protein-ligand complex requires an understanding of the H-bond interaction analysis. The fitting of ligand into the binding site depends largely on H-bonding. Lys 745, Cys 775, and Leu 777 displayed H-bond interaction with Ligand **1**. Lys 745, Thr 854, and Asp 855 displayed H-bond interaction with ligand **5** (Supplementary material Figure S9).

2.5. Binding free energy and per-residue decomposition studies

The binding free energy calculation is done by MM-GBSA and MM-PBSA in which the molecular mechanics (MM) energies and solvation energies are taken into consideration. The implicit solvent model (PB or GB) is used for calculating polar solvation energy while the SASA is used for non-polar solvation energy calculation (Miller et al. 2012). The energy contribution of non-bonded interaction energies, i.e. Vader Waals (ΔE_{vdw}) and electrostatic energy (ΔE_{ele}) for the two complexes was determined by MM-GBSA and MM-PBSA methods. The complexes were found to exhibit high stability as the net binding free energies calculated by both methods were low. In contrast, the gas phase energy contribution (ΔE_{vdw} and ΔE_{ele}) was high which indicates that the stability of the complex was primarily because of ligand conformation with reference to the receptor. The ligand **1** and ligand **5** displayed binding energy of -50.92 ± 1.83 and -42.94 ± 1.92, respectively in the GB solvation model. Both coumestans exhibited good binding free energy in the PB solvation model as well. The contribution of all the energies in the MM-GBSA assay and MM-PBSA assay has been shown in Supplementary material Figure S10.

Per-residue binding energy decomposition analysis revealed the contribution of various amino acids towards total binding energy. From the decomposition analysis, the contributions of the consistently interacting amino acids were extracted. In case of compound **1**, the decomposition energy for residues Met 766, Cys 775, Leu 858, Leu 788, and Leu 777 were −3.19, −2.54, −2.18, −2.02, and −1.98, respectively. While for compound **5**, the decomposition energy for Leu 747, Leu 858, Lys 745, Lys 875, and Phe 723 were −2.65, −1.90, −1.84, −1.72, and −1.71, respectively (Supplementary material Figure S11).

2.6. Isolation and purification of Compound 1 (psoralidin)

The seed extract of *C. corylifolium* was subjected to silica-gel column chromatography for the isolation of particularly psoralidin. The fraction containing psoralidin was subjected to repeated column chromatography to obtain a pure compound. The structure of the isolated compound was confirmed by NMR spectroscopy, mass spectrometry, and compared with that of the reported literature (Supplementary material Figure S12-S15). The isolated compound **1** was evaluated for cytotoxicity against MDA-MB-231 and A549 cell lines.

2.7. Cytotoxicity screening

The best hit, i.e. Compound **1** (Psoralidin) was evaluated against two cancer cell lines viz MDA-MB-231 and A549 for cytotoxicity screening using MTT assay for 48 h for validation of *in silico* results. The findings revealed significant cytotoxicity of compound **1** on the cancer cells with IC_{50} values of 22.21 ± 1.65 µM and 23.64 ± 0.39 against MDA-MB-231 cells and A549 cells, respectively. The good cytotoxicity shown by compound **1** in *in vitro* settings validated the protocol of *in silico* analysis and proved the robustness of *in silico* results.

3. Conclusion

The study conducted molecular docking and molecular modelling on 15 coumestans from *C. corylifolium* against EGFR, revealing two best hits: psoralidin (compound **1**) and isopsoralidin (compound **5**), which showed good inhibitory activity against EGFR. MD simulations showed stability in complexes with EGFR in simulated biological environments. The *in vitro* cytotoxicity performed on the best hit and isolated psoralidin against MDA-MB-231 and A549 cells also validated the credibility of the *in silico* method and protocol. The psoralidin displayed good cytotoxic potential against MDA-MB-231 and A549 cells with IC_{50} values of 22.21 and 23.64 µM, respectively. Since both cell lines overexpress EGFR, it could be inferred that the possible mechanism of action of Psoralidin (compound **1**) is inhibition of EGFR. The study's findings demonstrate the efficacy and potential of coumestans in inhibiting EGFR receptors, providing prospective leads for researchers in drug discovery for cancer therapy.

Acknowledgement

The authors extend their gratitude to Professor David A. Case, Department of Chemistry & Chemical Biology, Rutgers University, New Jersey, USA, for granting Amber 20 license. NT and NB are thankful to IIT (BHU) and the Ministry of Education (MoE) for providing teaching assistantship.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by SERB-CRG Grant (Grant No. CRG/2022/001637).

ORCID

Nancy Tripathi http://orcid.org/0000-0003-4193-6748
Nivedita Bhardwaj http://orcid.org/0000-0002-7857-9441
Shreyans K. Jain http://orcid.org/0000-0001-6160-8755

References

Atanasov AG, Zotchev SB, Dirsch VM, Orhan IE, Banach M, Rollinger JM, Barreca D, Weckwerth W, Bauer R, Bayer EA, et al. 2021. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov. 20(3):200–216. doi:10.1038/s41573-020-00114-z.

Case D, Ben-Shalom I, Brozell S, Cerutti D, Cheatham T, III, Cruzeiro V, Darden T, Duke R, Ghoreishi D, Gilson M. 2020. San Francisco. AMBER 2020; 2020.

Chopra B, Dhingra AK, Dhar KL. 2013. *Psoralea corylifolia* L. (Buguchi)—Folklore to modern evidence: review. Fitoterapia. 90:44–56. doi:10.1016/j.fitote.2013.06.016.

Danker T, Möller C. 2014. Early identification of hERG liability in drug discovery programs by automated patch clamp. Front Pharmacol. 5:203. doi:10.3389/fphar.2014.00203.

Dutta S, Mahalanobish S, Sil PC. 2021. Chapter 8 – Phytoestrogens as novel therapeutic molecules against breast cancer. In: Brahmachari G, editor. Discovery and development of

anti-breast cancer agents from natural products. Elsevier; p. 197–229. doi:10.1016/B978-0-12-821277-6.00008-8.

Hospital A, Goñi JR, Orozco M, Gelpí JL. 2015. Molecular dynamics simulations: advances and applications. Adv Appl Bioinform Chem. 8:37–47. doi:10.2147/AABC.S70333.

Liu TC, Jin X, Wang Y, Wang K. 2017. Role of epidermal growth factor receptor in lung cancer and targeted therapies. Am J Cancer Res. 7(2):187–202. eng.

Liu X, Nam J-W, Song YS, Viswanath ANI, Pae AN, Kil Y-S, Kim H-D, Park JH, Seo E-K, Chang M. 2014. Psoralidin, a coumestan analogue, as a novel potent estrogen receptor signaling molecule isolated from Psoralea corylifolia. Bioorg Med Chem Lett. 24(5):1403–1406. doi:10.1016/j.bmcl.2014.01.029.

Ma X-l, Chen C, Yang J. 2005. Predictive model of blood-brain barrier penetration of organic compounds. Acta Pharmacol Sin. 26(4):500–512. doi:10.1111/j.1745-7254.2005.00068.x.

Masuda H, Zhang D, Bartholomeusz C, Doihara H, Hortobagyi GN, Ueno NT. 2012. Role of epidermal growth factor receptor in breast cancer. Breast Cancer Res Treat. 136(2):331–345. doi:10.1007/s10549-012-2289-9.

Meng X-Y, Zhang H-X, Mezei M, Cui M. 2011. Molecular docking: a powerful approach for structure-based drug discovery. Curr Comput Aided Drug Des. 7(2):146–157. doi:10.2174/157340911795677602.

Miller BR, III, McGee TD, Jr., Swails JM, Homeyer N, Gohlke H, Roitberg AE. 2012. MMPBSA.py: an efficient program for end-state free energy calculations. J Chem Theory Comput. 8(9):3314–3321. doi:10.1021/ct300418h.

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 30(16):2785–2791. doi:10.1002/jcc.21256.

Qi S-Z, Li B-B, Liu M-S, Zhang Z, Miao S, Gong K-K. 2023. Chemical constituents from the seeds of *Cullen corylifolium* and their inhibitory activity on diacylglycerol acyltransferase. Nat Prod Res. 37(10):1601–1607. doi:10.1080/14786419.2022.2103126.

Sigismund S, Avanzato D, Lanzetti L. 2018. Emerging functions of the EGFR in cancer. Mol Oncol. 12(1):3–20. doi:10.1002/1878-0261.12155.

Tu Y, Yang Y, Li Y, He C. 2021. Naturally occurring coumestans from plants, their biological activities and therapeutic effects on human diseases. Pharmacol Res. 169:105615. doi:10.1016/j.phrs.2021.105615.

Uribe ML, Marrocco I, Yarden Y. 2021. EGFR in cancer: signaling mechanisms, drugs, and acquired resistance. Cancers (Basel). 13(11):2748. doi:10.3390/cancers13112748.

Zhang X, Zhao W, Wang Y, Lu J, Chen X. 2016. The chemical constituents and bioactivities of *Psoralea corylifolia* Linn.: a review. Am J Chin Med. 44(1):35–60. doi:10.1142/S0192415X16500038.