Ethnomedicine claim directed in-silico prediction of epidermal growth factor receptor kinase antagonist: an untapped reservoir of prospective anticancer agents

Tope Abraham Ibisanmi [1, *] Ayodele Ifeoluwa Faleti [2], Peace Ifeoma Odjegba [1], Olamide Joshua Babatunde [1], Titiladunayo Samuel Oluwatoyin [1], Olabanji Bukola Olayinka [1], Alaba Adebola Abosede [1] and Akinlolu Oluwatoyin Felicia [1]

[1] Department of Microbiology, School of Life Sciences, Federal University of Technology Akure, P.M.B. 704, Akure Nigeria.
[2] Department of Chemistry, School of Life Sciences, Federal University of Technology Akure, P.M.B. 704, Akure Nigeria.

Abstract

The epidermal growth factor receptor (EGFR) is a tyrosine kinase (TK) that belongs to the ErbB family and governs important cellular functions like reproduction, survival, motility, and differentiation. Overexpression, intensification, and alteration of EGFR occur in a wide range of human malignancies and are associated with tumor progression and decreased anticancer drug sensitivity. As a result, EGFR has been identified as one of the primary anticancer targets. As cancer is more likely to be poorly understood in traditional medical practices, the extrapolation of an ethnomedicine-led strategy to identifying and prioritizing anticancer medicinal plants has been questioned. Nonetheless, given the challenges of developing innovative anticancer drugs that are effective, safe, inexpensive, and widely available, ethnomedicinal studies play critical roles in identifying relevant medicinal plants that can be further investigated. This study employed pharmacophore modeling, molecular docking, and molecular dynamics simulation to develop an effective agent as an inhibitor for EGFR. The final findings revealed that the selected bioactive compound stabilized the EGFR protein. The optimum orientations of the various inhibitors was Friedelin and it was chosen and subjected, along with the FDA-approved drug, to molecular dynamics modeling to determine the molecular interaction of the medication with various mutational sites in order to deduce the suitable orientation for the inhibitors. The study also attest to the ethnomedicinal claims that ethnomedicinal plants played a huge role in anticancer drug discovery and that their exploration can change the bleak picture cancer paints in our societies today.

Keywords: Epidermal growth factor receptor; Ethnomedicinal plants; Cancer; Bioactive compound

1. Introduction

Structure-function studies have solidified our understanding of the function of the epidermal growth factor receptor (EGFR), a transmembrane protein that serves as a receptor for extracellular protein ligands from the epidermal growth factor family (EGF family) [1]. Viktor Hamburger, one of the fathers of developmental neuroscience, conducted early investigations in the 1930s at St. Louis University that suggested nerve cell proliferation was dependent on an inductive agent originating from their destination. Rita Levi-Montalcini, who performed the trials in Turin (Italy) and eventually won the Nobel Prize for discovering the Nerve Growth Factor (NGF), reported comparable findings. Stanley Cohen, a biochemist who collaborated with Levi-Montalcini and eventually shared the Nobel Prize with her, later isolated NGF after finding both the second growth factor, epidermal growth factor (EGF) [2,3], and the appropriate cell surface receptor (EGFR).

---

* Corresponding author: Tope Abraham Ibisanmi

The initial association to cancer was made by Michael Waterfield and colleagues. They collected partial human EGFR amino acid sequences and determined that this receptor closely matches a portion of the v-erb-B transforming protein of the avian erythroblastosis virus (AEV. Weihua and Thomas EGFR Kinase-Independent Functions in cancer. The EGFR tyrosine kinase function has acted as a beacon leading the design of EGFR targeted cancer treatments. Although subsequent studies discovered that EGFR and some ligand growth factors are overexpressed in human tumors and that the EGFR gene is amplified or rearranged in brain tumors, frequent oncogenic point mutants of EGFR were first identified in 2004, in patients with non-small cell lung cancer (NSCLC) who were sensitive to EGFR-specific tyrosine kinase inhibitors (TKIs; e.g., gefitinib and erlotinib) [1].

However, potent EGFR inhibitors, small molecules of tyrosine kinase inhibitors (TKI) competing with ATP for kinase activation and monoclonal antibody inhibitors (mAb) preventing EGFR from being activated by its ligands, have shown limited efficacy and have been challenged in the clinic by innate and acquired resistance. Over the last decade, the discovery that EGFR has pro-survival effects independent of its kinase activity has opened a new window for a better understanding of EGFR's role in cancer and presented a unique way to targeting this strong oncogene for cancer therapy.

Currently, because of a lack of access to contemporary cancer care, as well as other circumstances, many cancer patients in the developing countries rely on traditional therapies. Given the obstacles to developing innovative anticancer drugs that are effective, safe, inexpensive, and accessible, ethnomedicinal and ethnopharmacological studies play critical roles in identifying relevant medicinal plants that can be explored further [4]. Medicinal herbs, which formed the foundation of traditional medicinal practices, have long been used by mankind for preventive and therapeutic purposes [5]. The extrapolation into western medicine has not always been direct, as in the case of vincristine, where therapeutic efficacy was only distantly tied to indigenous use. Aside from the ethnomedicinal approach's overall success, its values and performance varied across diseases. Its scientific significance has also been questioned in diseases such as cancer, which are thought to be inadequately described in traditional medicinal methods [6]. Despite ethnomedicinal study-based claims of anticancer activity made by many, the National Cancer Institute (NCI) contends that cancer, particularly internal organ neoplasms such as lung, colon, ovarian, and prostate cancers, is unlikely to be established as a single disease entity in traditional medicine.

This problem with the ethnomedicine-based approach has been exacerbated by the lack of theory-based practice in many traditional medicine practices, making the information on diseases and plants utilized less accurate. As a result, the ethnomedicinal data acquired will be in general insufficient to identify promising therapeutic plants for more costly biological experiments. Despite the general challenge of establishing cancer as a specific disease in ethno medicine, ethnomedicinal reports can provide trustworthy information on external tumors such as skin cancer, external growths and swellings, tumors of the oral cavity, and genital cancers. To correlate findings from existing scientific studies and ethnomedicinal surveys, virtual screening methods that are cost-effective, rapid, and trustworthy could be used. In retrospect, such methodological combination' could also aid understanding of the indigenously defined disease in terms of its ties to cancer [7]. To that purpose, we merged an ethnomedicine study with an in silico ligand-based technique. The study aims to provide a scientific foundation for existing conventional medical practice in a complex disease like cancer, as well as to leverage its potential for medication discovery and development. As a result, we investigated the anticancer activity and therapeutic potential of phytochemicals found in plants identified by an ethnomedicinal study.

2. Material and methods

2.1. Ligands selection and preparation

A systematic literature search of recent years across multiple databases yielded 30 bioactive compounds from ethnomedicinal plants (Table 1) in an effort to uncover a viable therapeutic bioactive chemical and repurpose it against Epidermal Growth Factor Receptor Kinase. The PubChem online library (https://pubchem.ncbi.nlm.nih.gov/) was utilized to retrieve the previously selected naturally derived Product for docking against epidermal growth factor receptor kinase [8].

The FDA-approved medication library (Erlotinib zinc15 ID: ZINC000001546066 for) was retrieved from the ZINC15 database. This is an online public repository that contains information about chemicals and their biological activity. The 3D structures were obtained from the database in 3D SDF format and converted separately to PDBQT files using OpenBabel-3.1.0.
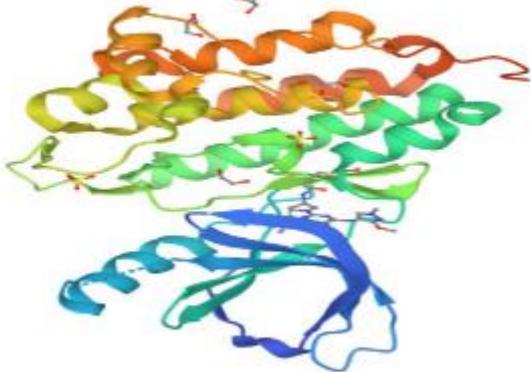
2.2. Protein preparation for docking

The crystallographic structure of epidermal growth factor receptor kinase (SUG9) at a minimal resolution was obtained from Protein Data Bank (www.rscb.org), and the water molecules of these proteins were removed using BIOVIA Discovery Studio version 21.1. Chimera 1.16 was used to prepare these proteins for docking, while Discovery Studio version 21.1 and CASTp23 online tools were utilized to estimate the active site of each protein, which was then cross-checked against the target literature (Table 2).

Table 1 Ethnomedicinal plants and their bioactive compounds

Plant Name	Bioactive compounds	Ref
Tithonia diversifolia	Daturametelin C, 5-O-Methylvisamminol, Trigonelline, Kaempferol, Quercimeritrin, and 4-O-Caffeoylquinic acid	[9]
Vernonia Cinerascens	Vernodalol, hydroxyvernolide, Vernodalin, Vernolide,and 3'-Methylquercetin	[10]
Barleria eranthemoides	Barlerin, Isoverbascoside, Syringic acid, Melilotic acid , Vanillic acid, Scutellarin,and Shanzhiside	[11]
Lannea Schweinfurthi	Quinic acid, 3-O Caffeoylquinic acid, Feruloylquinic acid, Ligustroside Lupenone, Sitosterol and Rutin	[12,13]
Terminalia mollis	Friedelin	[14]
Acacia tortilis		
Ficus cycamorus	Kaempferol	[15,16,17,18]
Indigofera Colutea	Hetranthin B	[19]
Stachytarpheta jamaicaensis	α-spinasterol and hispidulin	[20]

Table 2 Information about the target protein (EGFR kinase domain)

Title	EGFR Kinase Domain
Structure	

Structure validation	
RCSBPDB ID	5UG9
Experimental method	X-RAY DIFFRACTION
Resolution	1.33 Å
Source organism	Homo sapiens
Expression system	Spodoptera aff. frugiperda 2 RZ-2014

2.3. Organ toxicity and toxicity endpoints analysis

Toxicity refers to a drug's ability to poison the body. The toxicity of a medicine can have an effect on a whole organism (animal, plant, bacteria). The PKCSM web server (http://biosig.unimelb.edu.au/pkcsm/prediction) is an online database that may simply assess the drug molecule by entering its canonical smiles to examine its toxicity. PKCSM has all toxicity parameters, including molecular weight, AMES toxicity, oral rat acute toxicity (LD50), hepatotoxicity, minnow toxicity, HERG-I inhibitor, maximum tolerated dose (human), and so forth.

2.4. ADMET profiling

To develop the ADMET profile of all 30 bioactive chemical ligands and FDA-approved Erlotinib, Protox II online server was used to predict their absorption, distribution, metabolism, excretion, and toxicity. SWISSADME prediction was used to calculate the drug-like features of the ligands, and drug-like characteristics were estimated using Lipinski's rule of five.

2.5. Molecular target

Target prediction Swiss Target Prediction is a web-based tool for predicting the macromolecular target of a small bioactive molecule. It is critical to identify the phenotypic side effects and potential cross-reactivity of small biomolecules. It is based on the similarity principle, which states that two molecules with similar structures are likely to have comparable properties. In order to forecast the target of our hit compound, the canonical smile is entered and analyzed in the search box.

2.6. Docking strategy

Only 15 of 30 bioactive compounds passed Lipinski's rule of five after ADMET screening. To discover a hit medication that could potentially inhibit the function of a protein to target epidermal growth factor receptor kinase, specific protein-ligand docking was performed using Autodock Vina (21) and PyRx 8.0. The target proteins were centered on the grid map for docking calculations [22]. The software BIOVIA Discovery Studio version 21.1 was used to simulate non-bonded polar and hydrophobic interactions in the inhibitor site of 5UG9. Standard docking protocols were used for subsequent docking procedures.

The hit compounds, which exhibited the predicted interactions with the essential amino acids present in the protein's active region, could be potent antagonists of 5UG9. Prior to pose rating for fitness within the binding pocket of each protein for continuum and discrete bond interactions, the molecular docking was evaluated based on the free binding energy of the ligands with the various proteins. pymol, ligplus, and BIOVIA Discovery Studio version 21.1 were used for post-docking analysis, 3D, and 2D image production. The docking studies were validated by analyzing the ligand

4

interactions with 5UG9. The most energy-minimized conformation (highest binding free energy) for each of the ligands was observed to interact with higher numbers of bonds and the nature of bonds.

2.7. Molecular dynamics simulation

In this study, we performed molecular dynamics simulations of a system using the massively parallel simulation package OpenMM version 7.5 [23] on a GPU-enabled platform. The system was prepared using CHARMM-GUI solution builder [24] and the input files were exported in the Gromacs format. We used a Gromacs.gro file to initialize the positions of the particles in the system and a topology file to define the interactions between the particles. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method [23], and a cutoff distance of 1.2 nm was applied for Van der Waals interactions.

The system was equilibrated in the isothermal-isobaric (NPT) ensemble, and the final simulation was run in the canonical (NVT) ensemble at a temperature of 300 K for a total of 50 ns. The simulation was run using a Langevin integrator, with a timestep of 1 fs. The simulation was run on a GPU-enabled platform, using a device with an Intel(R) Core(TM) i9-10920X CPU @ 3.50GHz and a NVIDIA GeForce RTX 3090 graphics card. We utilized the Mdanalysis package [25] to analyze the computed trajectories and calculate important parameters such as the root mean square fluctuation (RMSF), the radius of gyration (RG), and root mean square deviation (RMSD) of the protein and protein-ligand complexes. These parameters allowed us to examine the deviations of the protein and complexes and assess their stability over the course of the simulation.

The Ramachandran Plot of the protein was generated using online server of Molecular Biology Institute of University of California, Los Angeles.

2.8. Phylogenetic analysis

The genetic relatedness of the target protein sequences (5UG9) under study was compared with sequences of their nearest neighbors. 10 EGFR kinase protein sequences (5UGA, 5UGC, 5UG8, 4I1Z, 5UWD, 2ITT, 2ITU, 2ITV, 2EB3 and 2ITZ) were aligned in Alignment Explorer of MEGA4 software. Trimming and verification of the sequence alignment were carried out using the MUSCLE (UPGMA) algorithm. The Maximum Composite Likelihood and Neighbor-Joining methods were used to compute the evolutionary distances and history respectively. The robustness of the tree was assessed by bootstrap analysis with 1000 replication.

3. Results and discussion

3.1. Drug Likeness Prediction

Drug discovery is a challenging, time consuming and costly process. In small-molecule drug development, the initial emphasis is largely on efficacy which makes addressing the issue of drugability secondary. This practice often leads to high failure rates and developmental costs. Hence, we also assessed the general and antineoplastic drugability of relevant phytochemicals in addition to a similarity based prediction of anticancer activity. The drug-likeness parameters of all the selected phytoconstituents were assessed using swissadme web server. Among all the selected Rutin and Isoverbascoside failed 3 rules, Quercimeritrin, Scutellarin, Ipolamiide, Shanzhiside, and Ligustroside failed 2 lipinski's rules. phytocompounds Friedelin, Vernodalin, Hispidulin, 5-O-Methylvisamminol, 3'-Methylquercetin, Kaempferol, Vernolide, Feruloylquinic Acid, Hydroxyvernolide, Melilotic Acid, Vernodalol, Quinic Acid, Syringic Acid, Vanillic Acid and Trigonelline passed the Lipinski's rule of five: not more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms), not more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms, a molecular weight under 500 g/mol and a partition coefficient log P less than 5. Hence, they can be used as future oral drugs (Table 3). It means there is a high tendency for molecules that passed might be pharmacologically active. In a null shell these molecules are said to have good absorption, low toxicity level, orally bioavailable, and permeable.

3.2. Toxicity Prediction

The toxicity of the selected fifteen phytoconstituents was analyzed by using the Protox II online tool. Toxicity data showed that all the selected compounds that passed Lipinski's rule of five do not cause hepatotoxicity, hERG I inhibitor and hERG I inhibitor. 5-O-Methylvisamminol, Vernodalol, and Vernodalin do cause AMES toxicity. The docking investigation meets the toxicity endpoint safety requirements (table 4).

3.3. Phylogenetic analysis

Evolutionary relationship of Epidermal Growth Factor Receptor Kinase, 10 different PDB of target protein was used for this analysis. The phylogenetic relationship between the protein target sequences and the corresponding selected homologs for EGFR infers that a likely comparable result from the docking bond affinity can be generated with all the 10 protein and closely related with 5UG8 AND 5UGA with 52% bootstrap confidence. The phylogenetic tree generated helps to understand the evolutionary relationship of EGFR among different proteins. The protein target sequences phylogenetic relationship shows similarity to all the protein sequence implying the likelihood of reoccurring results of the Insilico profiling (figure 1).

Figure 1 Phylogenetic relationship between the test EGFR kinase and other homologous sequences

Table 3 Drug Likeness Prediction of selected bioactive compound from ethinomedicinal plants

Molecule	Formula	MW	#Heavy atoms	#Aromatic heavy atoms	Fraction Csp3	#Rotatable bonds	#H-bond acceptors	#H-bond donors	MR	TPSA	Lipinski #violations	log+AL12+AM1:A0+AM1:A031	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor
Daturametelin I	C34H48O10	616.74	44	0	0.76	6	10	5	160.39	162.98	1	-8.44	No	No	No	No	No
Trigonelline	C7H7NO2	137.14	10	6	0.14	1	2	0	35.05	44.01	0	-6.77	No	No	No	No	No
Kaempferol	C15H10O6	286.24	21	16	0	1	6	4	76.01	111.13	0	-6.7	Yes	No	No	Yes	Yes
5-O-Methylvisamminol	C16H18O5	290.31	21	10	0.44	2	5	1	78.91	68.9	0	-6.73	Yes	Yes	No	Yes	No
Quercimeritrin	C21H20O12	464.38	33	16	0.29	4	12	8	110.16	210.51	2	-8.88	No	No	No	No	No
4-O-Caffeoylquinic acid	C16H18O9	354.31	25	6	0.38	5	9	6	83.5	164.75	1	-8.76	No	No	No	No	No
Vernodalol	C20H24O8	392.4	28	0	0.45	8	8	2	98.05	119.36	0	-7.86	No	No	No	No	No
Vernodalin	C19H20O7	360.36	26	0	0.42	5	7	1	89.97	99.13	0	-8.28	No	No	No	No	No
Vernolide	C19H22O7	362.37	26	0	0.58	3	7	1	89.51	94.59	0	-7.85	No	No	No	No	No
Hydroxyvernolide	C19H22O8	378.37	27	0	0.58	4	8	2	90.67	114.82	0	-8.84	No	No	No	No	No
3'-Methylquercetin	C16H12O7	316.26	23	16	0.06	2	7	4	82.5	120.36	0	-6.9	Yes	No	No	Yes	Yes
Barlerin	C19H28O12	448.42	31	0	0.79	7	12	5	98.3	181.44	1	-10.48	No	No	No	No	No
Isoverbascoside	C29H36O15	624.59	44	12	0.48	11	15	9	148.42	245.29	3	-10.46	No	No	No	No	No
Syringic acid	C9H10O5	198.17	14	6	0.22	3	5	2	48.41	75.99	0	-6.77	No	No	No	No	No
Melilotic acid	C9H10O3	166.17	12	6	0.22	3	3	2	44.82	57.53	0	-6.3	No	No	No	No	No
Vanillic acid	C8H8O4	168.15	12	6	0.12	2	4	2	41.92	66.76	0	-6.31	No	No	No	No	No
Scutellarin	C21H18O12	462.36	33	16	0.24	4	12	7	108.74	207.35	2	-8.59	No	No	No	No	No
Ipolamiide	C17H26O11	406.38	28	0	0.82	5	11	6	88.6	175.37	2	-10.32	No	No	No	No	No
Shanzhiside	C16H24O11	392.36	27	0	0.81	4	11	7	84.25	186.37	2	-10.78	No	No	No	No	No

Quinic acid	C7H12O6	192.17	13	0	0.86	1	6	5	40.11	118.22	0	-9.15	No	No	No	No	No
3-O-Caffeoylquinic acid	C16H18O9	354.31	25	6	0.38	5	9	6	83.5	164.75	1	-8.76	No	No	No	No	No
Feruloylquinic acid	C17H20O9	368.34	26	6	0.41	6	9	5	87.97	153.75	0	-8.62	No	No	No	No	No
Ligustroside	C25H32O12	524.51	37	6	0.52	11	12	5	125.25	181.44	2	-9.56	No	No	No	No	No
Lupenone	C30H48O	424.7	31	0	0.9	1	1	0	134.18	17.07	1	-2.1	No	No	No	No	No
Sitosterol	C29H50O	414.71	30	0	0.93	6	1	1	133.23	20.23	1	-2.2	No	No	No	No	No
Fiedel	C15H12O6	288.25	21	12	0.13	1	6	4	73.59	107.22	0	-6.62	No	No	No	No	Yes
Rutin	C27H30O16	610.52	43	16	0.44	6	16	10	141.38	269.43	3	-10.26	No	No	No	No	No
Hetranthin B	C25H28O13	536.48	38	16	0.4	8	13	5	130.06	186.74	2	-8.64	No	No	No	No	Yes
α-spinasterol	C30H50O	426.72	31	0	0.97	0	1	0	134.39	17.07	1	-1.94	No	No	No	No	No
Hispidulin	C16H12O6	300.26	22	16	0.06	2	6	3	80.48	100.13	0	-6.01	Yes	No	No	Yes	Yes

Table 4 Toxicity analysis of the selected bioactive compound from ethinomedicinal plants

Molecule	AMES toxicity	Max. tolerated dose	hERG I inhibitor	hERG I inhibitor	Oral Rat Acute Toxicity (LD50)	Oral Rat Chronic Toxicity (LOAEL)	Hepatotoxicity	Skin Sensitisation	T.Pyriformis toxicity	Minnow toxicity
Friedel	No	0.279	No	No	2.402	1.834	No	No	0.333	1.84
Vernodalin	Yes	0.236	No	No	2.285	1.768	No	No	0.314	2.36
Hispidulin	No	0.279	No	No	2.402	1.634	No	No	0.361	1.64
5-O-Methylvisamminol	Yes	-0.051	No	No	2.402	1.443	No	No	0.844	1.379
3'-Methylquercetin	No	0.576	No	No	2.407	2.499	No	No	0.296	2.206
Kaempferol	No	0.531	No	No	2.449	2.505	No	No	0.312	2.885
Vernolide	No	-0.324	No	No	3.467	1.107	No	No	0.295	3.007
Feruloylquinic Acid	No	1.285	No	No	2.025	4.485	No	No	0.285	4.876
Hydroxyvernolide	No	0.11	No	No	3.949	2.087	No	No	0.287	3.852
Melilotic Acid	No	1.092	No	No	2.168	2.555	No	No	0.319	1.556
Vernodalol	Yes	0.501	No	No	2.388	1.971	No	No	0.29	3.646
Quinic Acid	No	1.626	No	No	No	No	No	No	0.285	4.869
Syringic Acid	No	1.374	No	No	2.157	2.415	No	No	0.2554	2.554
Vanillic Acid	No	0.719	No	No	2.454	2.032	No	No	0.265	1.926
Trigonelline	No	0.743	No	No	1.878	0.454	No	No	-0.323	2.536
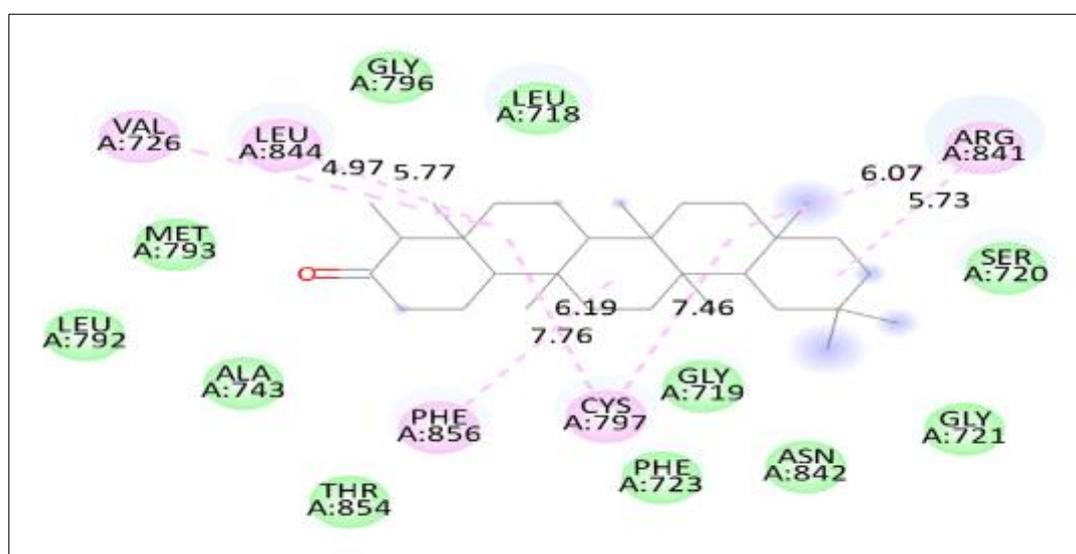
A
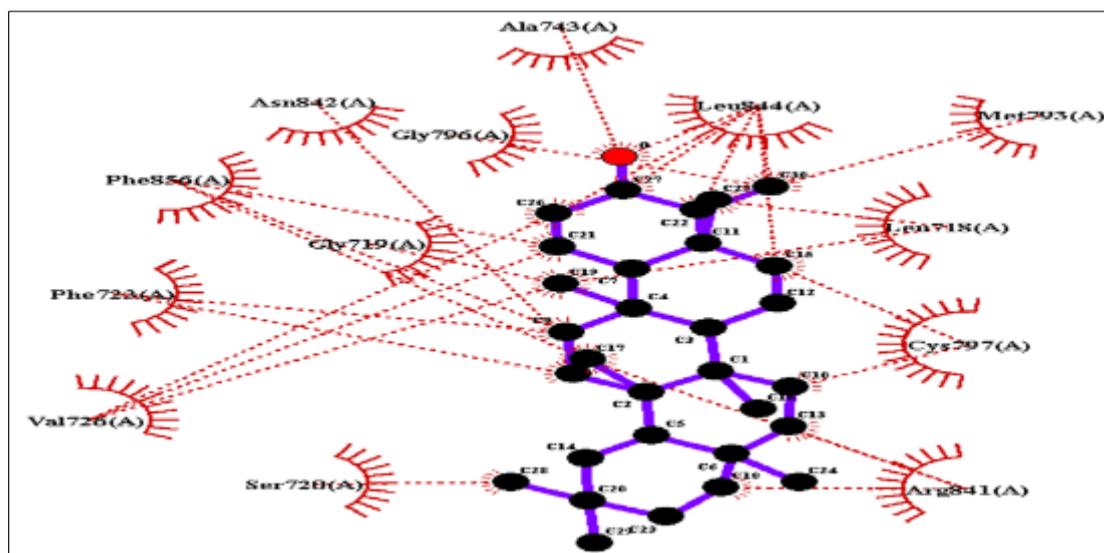
B

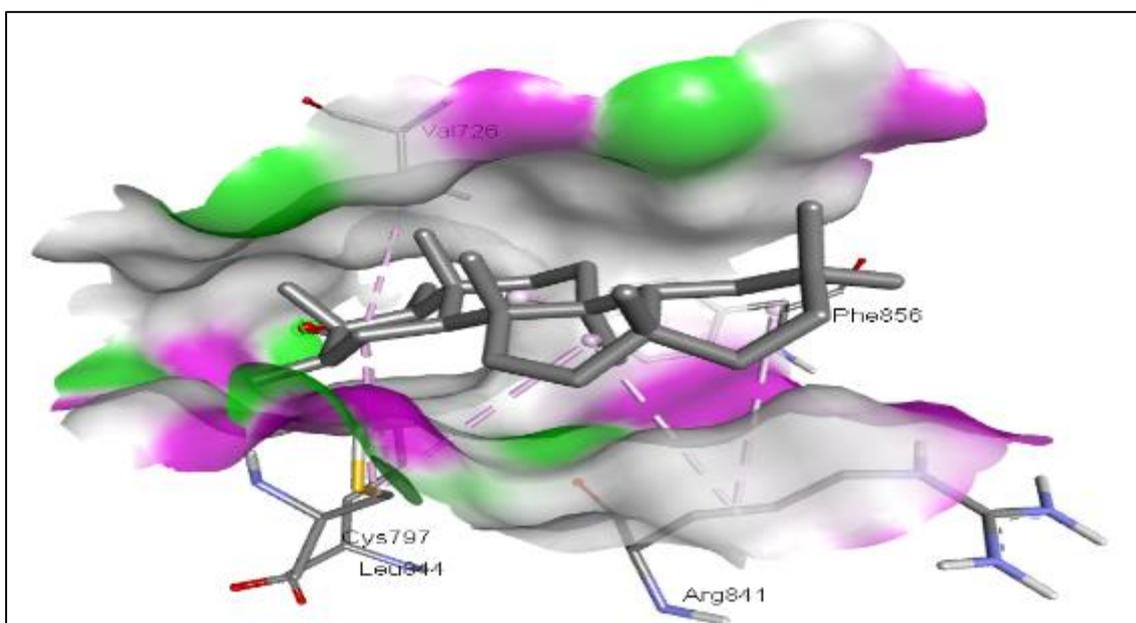**Figure 2** (A, B) 2D structure of 5UG9-Friedelin with their distance of interaction

Figure 3 3D structure of 5UG9- Friedelin binding pocket

Table 5 Molecular docking results of selected ethinomedicinal bioactive compounds against EGFR kinase domain (5UG9)

Compound name	Binding affinity (kcal/mol)
Standard compound	
Erlitinib	-6.6
Screened bioactive compounds	
Friedel	-9.6
Vernodalin	-8.5
Hispidulin	-8.4
5-O-Methylvisamminol	-8.4
3'-Methylquercetin	-8.2
Kaempferol	-8.2
Vernolide	-8
Feruloylquinic Acid	-7.8
Hydroxyvernolide	-7.6
Melilotic Acid	-6.4
Vernodalol	-6.1
Quinic Acid	-6
Syringic Acid	-5.7
Vanillic Acid	-5.7
Trigonelline	-5.5

3.4. Molecular Target Analysis

After the screening and molecular docking study, the molecular targets of our hit chemical, Friedelin and FDA-approved drug (Figure 2), are investigated further. 20% of enzymes, 20% of lyase, 20% of oxidoreductase, 13.3% of Family A G protein-coupled receptor, 13.3% of nuclear receptor, 6.7% of protease and 6.7% of cytochrome P450 were predicted for Friedelin, whereas only 93.3% kinase and 6.7% enzyme were found in the FDA approved.
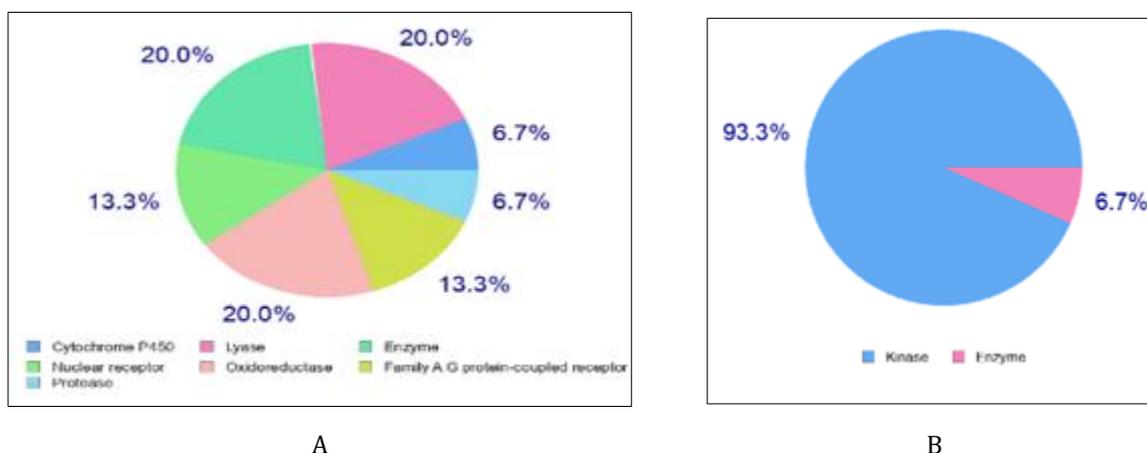

A   B

Figure 4 (A) Molecular Target Analysis result of Friedelin, (B) Molecular Target Analysis chart of FDA approved Erlotinib

3.5. Molecular docking

The binding mode of ethnomedicinal bioactive compounds that passed Lipinski's rule of five was docked against the active site of EGFR (receptor) (Table 4). From Table, The metabolites had docking scores ranging from -9.6 kcal/mol to -5.5 kcal/mol, Friedelin was identified to have the highest binding energy (-9.6 kcal/mol) whereas Trigonelline had the lowest binding affinity with a binding score of -5.5kcal/mol. Friedelin having the highest binding affinity formed sole pi-alkyl interaction with VAL A:726, LEU A:844, PHE A:856, CYS A:797 ,and ARG A:841 in the active site of EGFR (figure 2-3). Vernodalin with a binding energy of -8.5 kcal/mol and having the second highest binding affinity, formed three similar pi-alky interactions with Friedelin in the active site of EGFR with the addition of 2 conventional hydrogen bonds interacting with ARG A:841 and SER A:720 at the active site of EGFR. Looking at Trigonelline which is lowest in the ranking in terms of binding energy (-5.5 kcal/mol). It formed only 3 bonds: 3 pi-Alky Bonds interacting with VAL A:726, LEU A:844 and ALA A:743 with only one conventional hydrogen interaction (table 5).
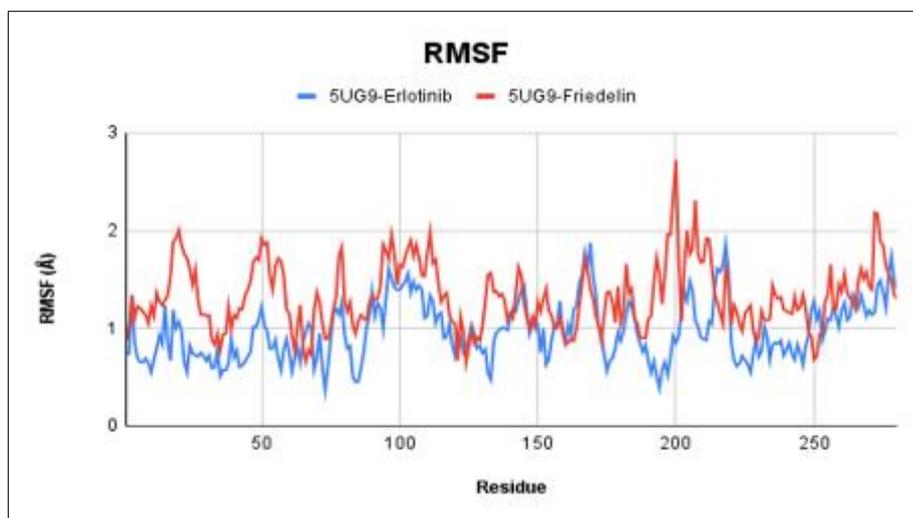

Figure 5 Analysis plots of RMSF for 5UG9-Fiedelin and 5UG8-Erlotinib complex and during 50 ns MD simulation
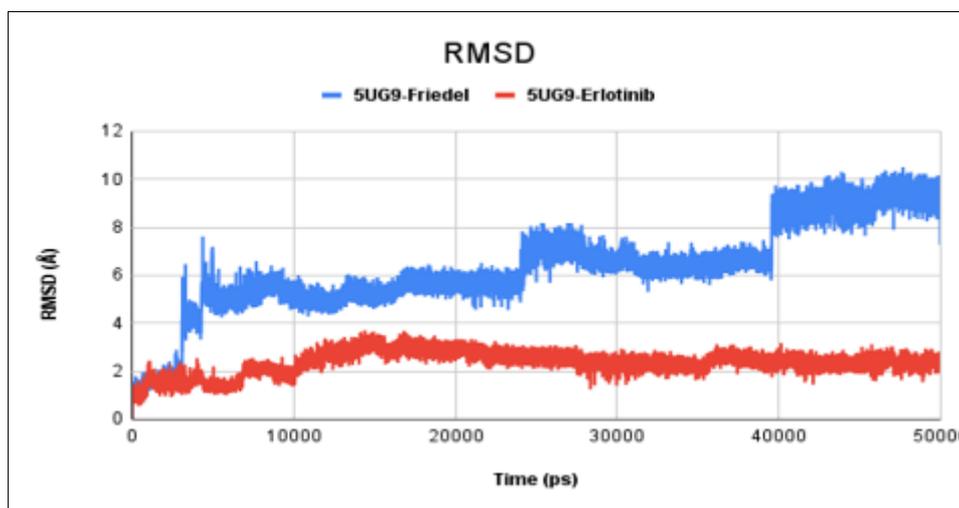

Figure 6 Analysis plots of RMSD for 5UG9-Fiedelin and 5UG8-Erlotinib complex and during 50 ns MD simulation
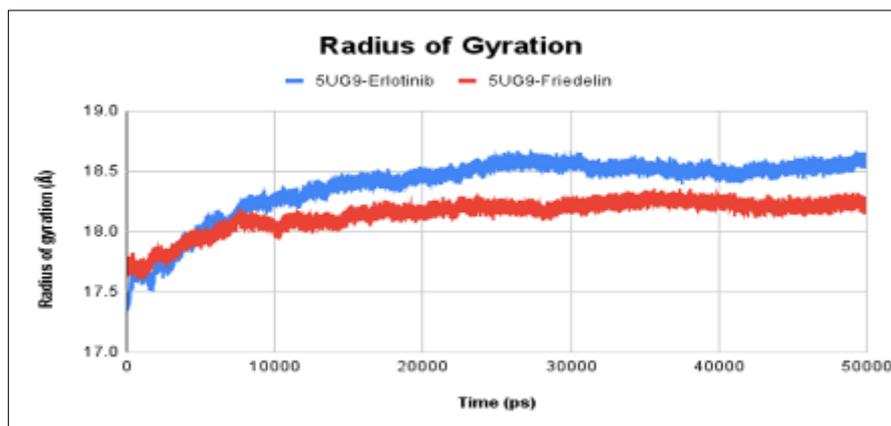
Figure 7 Analysis plots of Radius of Gyration for 5UG9-Fiedelin and 5UG8-Erlotinib complex and during 50 ns MD simulation.

3.6. Molecular dynamics simulation

Figure 3 depicts the molecular dynamics investigations of Friedelin and Erlotinib in contact with the EGFR kinase domain of the target protein (5UG9). The EGFR protein–ligand of the Friedelin and FDA-approved drug graphs are exported into Figure 5-7 in proper sequence. While the average RMSF values show that 5UG9-Friedelin RMSF is a little higher than the 5UG9-Erlotinb complex with no abrupt differences in the plot, the high flexibility values were found 250 residues of the structure. The root mean square fluctuation (RMSF) measures the movement of a subset of atoms with respect to the average structure over the entire simulation. RMSF indicates the flexibility of different regions of a protein. The root mean square fluctuation (RMSF) of the residue can be used for evaluating structural movement and flexibility. While the average RMSF values show that 5UG9-Friedelin RMSF is a little higher than the 5UG9-Erlotinb complex with no abrupt differences in the plot, the high flexibility values were found 250 residues of the structure. It is widely accepted that the binding site is mainly composed of several key residues in the active pocket, and monitoring the behavior of these residues would contribute to the study of protein–ligand interaction. After the MD experiment, RMSD curves of all the binding sites had been collected on the basis of the MD trajectory. As shown in Figure 6, After the MD experiment, RMSD curves of all the binding sites had been collected on the basis of the MD trajectory.

The RMSDs of Erlotinib and Friedelin complex with 5UG9 showed little fluctuation between 0–1000 ps (which is in the acceptable range), and afterwards it remained stable up to 50000 ps (50 ns) at 2-3 Å for 5UG9-Erlotinib complex. Whereas, in the case of 5UG9-Friedelin complex, higher RMSD value was recorded compare to the FDA-approved drug, increase in the 5UG9- Friedelin was noticed reaching a peak value between 40 ns and 50ns of the simulation. RMSDs plot showed little fluctuation between 0–10ns, 20-30ns and at 40n, which remained stable between 40 to 50 ns thereafter. Figure 7 reveals that both EGFR complexes binding sites had very stable structures. The root mean square deviation (RMSD) is a frequently used measure of the differences between the structures sampled during the simulation and the reference structure. After the MD experiment, RMSD curves of all the binding sites had been collected on the basis of the MD trajectory. The RMSDs of Erlotinib and Friedelin complex with 5UG9 showed little fluctuation which is in the acceptable range, and afterward, it remained stable up to 50000 ps (50 ns). Whereas, in the case of 5UG9-Friedelin complex, a higher RMSD value was recorded compared to the FDA-approved drug, increase in the 5UG9- Friedelin was noticed reaching a peak value between 40 ns and 50ns of the simulation. This is probably related to the structure of corresponding inhibitors. However, what is interesting is that the Friedelin might contribute to the fluctuation of 5UG9 when comparing to FDA approved drug, mainly due to the fact that these two protein complexes contain different inhibitors.

According to the foregoing conjecture, EGFR-Friedelin binding site probably led to the variation of Rg during MD simulation. The radius of gyration (Rg) can be used for indicating protein stability during simulation. In other words, the Rg curve of this protein would reach a plateau during simulation when a protein had been folded well. Changes in the structure of a protein can be monitored by Rg fluctuation. The plot reveals that both EGFR complexes' binding sites had very stable structures. According to the foregoing conjecture, EGFR-Friedelin binding site probably led to the variation of Rg during MD simulation.

3.7. Ramachandran Plot

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favorable region, the results obtained after the Ramachandran plot of this protein shows 93.1% in favorable region.
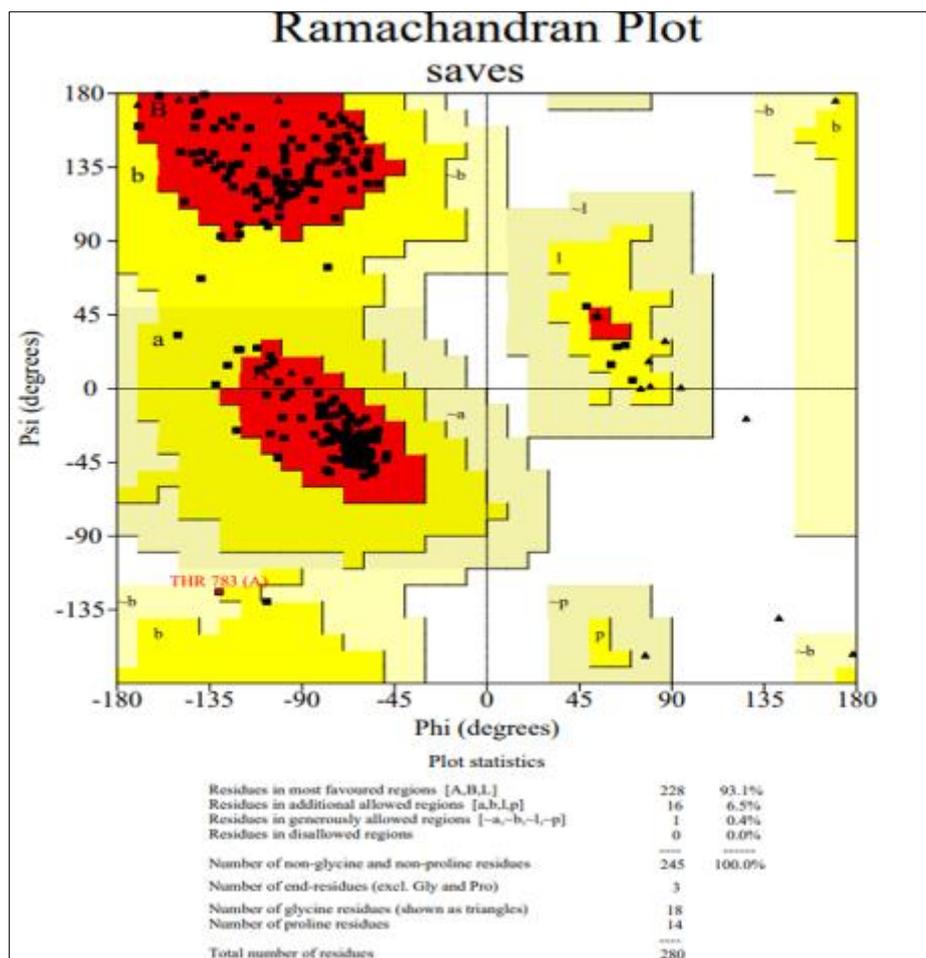

Figure 8 Ramachandran plot of the EGFR protein

4. Conclusion

Strategies for EGFR inhibition will be beneficial in patients impacted by the EGFR signaling pathway. This work employed pharmacophore modeling, molecular docking, and molecular dynamics simulation to develop an effective chemical as an inhibitor for EGFR. The final findings revealed that the selected bioactive compound stabilized the EGFR protein. The optimum orientations of the various inhibitors were chosen and subjected to molecular dynamics modeling to determine the molecular interaction of the medications with various mutational sites to deduce the suitable orientation for the inhibitors. These findings would aid in the development of a treatment strategy based on the molecular profiles of patients who respond specifically to an agent. It is concluded that EGFR inhibitors have the potential to be a powerful class of molecularly targeted anticancer medicines. The study also attests to the ethnomedicinal claims that ethnomedicinal plants played a huge role in anticancer drug discovery and their exploration can change the bleak picture cancer paint in our societies today.

Compliance with ethical standards

Disclosure of conflict of interest

Authors have declared that no conflict of interests exists.

References

[1] Thomas R and Weihua Z Rethink of EGFR in Cancer With Its Kinase Independent Function on Board. Front. Oncol. 2019, 9:800. doi: 10.3389/fonc.2019.00800

[2] Uribe, M.L.; Marrocco, I.; Yarden, Y. EGFR in Cancer: Signaling Mechanisms, Drugs, and Acquired Resistance. Cancers 2021, 13, 2748. https://doi.org/10.3390/cancers 13112748

[3] Yarden Y, Pines G. The ERBB network: at last, cancer therapy meets systems biology. Nat Rev Cancer. 2012, 12:553–63. doi: 10.1038/nrc3309

[4] Ibisanmi, T.A and Aribisala. J. O. Evaluation of antioxidant, phytochemicals, and antibacterial potential of Mormordica charantia (linn) against pathogenic bacteria isolated from ready to eat food sold in Akure metropolis, Nigeria. Bacterial empire: 2022, 5(1), e389 https://doi.org/10.36547/BE.389

[5] Girma, b. Eshetu Mulisa, ShibruTessema, Wote Amelo. Ethnomedicine Claim Directed in Silico Prediction Of Anticancer Activity. Ethiop J Health Sci.2017;28(1):83 doi:http://dx.doi.org/10.4314/ejhs.v28i1.1

[6] Sawadogo WR, Schumacher M, Teiten M-H, Dicato M, Diederich M. Traditional West African pharmacopeia, plants and derived Ethiop J Health Sci. Vol. 28, No. 1 January 2018 DOI: http://dx.doi.org/10.4314/ejhs.v28i1.10 92 compounds for cancer therapy. Biochem Pharmacol 2012; 84:1225–1240.

[7] Prada-Gracia D, Huerta-Yépez S, MorenoVargas LM. Application of computational methods for anticancer drug discovery, design, and optimization. Bol Med Hosp Infant Mex 2016; 73(6):411- 423.

[8] Akinwumi, I A.; Faleti, A. I.; Owojuyigbe, A. P.; Raji, F. M.; Alaka, O. M. In silico Studies of Bioactive Compounds Selected from Four African Plants with Inhibitory Activity against Plasmodium falciparum Dihydrofolate Reductase-Thymidylate Synthase (pfDHFR-TS). J. Adv. Pharm. Res. 2022, 6 (3), 107-122. DOI: 10.21608/aprh.2022.139794.1175

[9] Yazid, F., Salim, S. O., Rahmadika, F. D., Rosmalena, R., Artanti, N., Sundowo, A., & Prasasty, V. D.Antidiabetic effects of Tithonia diversifolia and Malus domestica leaf extracts in alloxan-induced Sprague Dawley rats. Syst Rev Pharm, 2021, 12(1), 1630-163

[10] Ahmad, I., Ashraf, M., Chaudhary, B. A., Janbaz, K. H., & Uzair, M.. Urease inhibitors and antioxidants from vernonia cinerascens. Journal of the Chemical Society of Pakistan, 2011, 33(1), 114-117.

[11] Gangaram, S., Naidoo, Y., Dewir, Y. H., & El-Hendawy, S. Phytochemicals and Biological Activities of Barleria (Acanthaceae). Plants, 2022, 11(1), 82. https://www.mdpi.com/2223-7747/11/1/82

[12] Maroyi, A.. Review of Ethnomedicinal, Phytochemical and Pharmacological Properties of Lannea schweinfurthii (Engl.) Engl. Molecules, 2019, 24(4). https://doi.org/10.3390/molecules24040732

[13] Sobeh, M., Mahmoud, M. F., Hasan, R. A., Abdelfattah, M. A. O., Sabry, O. M., Ghareeb, M. A., El-Shazly, A. M., & Wink, M.. Tannin-rich extracts from Lannea stuhlmannii and Lannea humilis (Anacardiaceae) exhibit hepatoprotective activities in vivo via enhancement of the anti-apoptotic protein Bcl-2. Sci Rep, 2018, 8(1), 9343. https://doi.org/10.1038/s41598-018-27452-8

[14] Liu, M., Katerere, D. R., Gray, A. I., & Seidel, V. Phytochemical and antifungal studies on Terminalia mollis and Terminalia brachystemma. Fitoterapia, 2009, 80(6), 369-373. https://doi.org/10.1016/j.fitote.2009.05.006

[15] Kimutai, N., Ariya, O., Mutai, C., & Jeruto, P. Ethnomedicinal study of selected medicinal plants used against bacterial infections in Nandi County, Kenya. Journal of Medicinal Plants Studies, 2019, 7, 103-108.

[16] Maregesi, S., Van Miert, S., Pannecouque, C., Haddad, M. H. F., Hermans, N., Wright, C. W., Vlietinck, A. J., Apers, S., & Pieters, L. Screening of Tanzanian medicinal plants against Plasmodium falciparum and human immunodeficiency virus. Planta medica, 2019, 76(02), 195-201.

[17] Pistelli, L., Chiellini, E. E., & Morelli, I.. Flavonoids from Ficus pumila. Biochemical Systematics and Ecology, 2000, 28(3), 287-289. https://doi.org/https://doi.org/10.1016/S0305-1978(99)00064-2

[18] Qi, Z.-Y., Zhao, J.-Y., Lin, F.-J., Zhou, W.-L., & Gan, R.-Y. . Bioactive Compounds, Therapeutic Activities, and Applications of Ficus pumila L. Agronomy, 2021, 11(1), 89. https://www.mdpi.com/2073-4395/11/1/89

[19] Gerometta, E., Grondin, I., Smadja, J., Frederich, M., & Gauvin-Bialecki, A.. A review of traditional uses, phytochemistry, and pharmacology of the genus Indigofera. Journal of Ethnopharmacology, 2020, 253, 112608. https://doi.org/https://doi.org/10.1016/j.jep.2020.112608

[20] Putri; Kurnianingsih, Nia; and Faisal, Mohammad Reza "An in Silico Study of the Cathepsin L Inhibitory Activity of Bioactive Compounds in Stachytarpheta jamaicensis as a Covid-19 Drug Therapy," Makara Journal of Science 2022

[21] Sharma J, Bhardwaj VK, Rahul S, Vidya R, Rituraj P, Sanjay K. An in-silico evaluation of different bioactive molecules of tea for their inhibition potency against non-structural protein-15 of SARS-CoV-2 Food Chemistry 2021; 346:128933 https://doi.org/10.1016%2Fj.foodchem.2020.128933

[22] Selvaraj A, Antony S, Hakdong S. Anti-methanogenic efect of rhubarb (Rheum spp.)-An in silico docking studies on methyl-coenzyme M reductase (MCR). Saudi J Biol Sci 2019; 26:1458–1462. https://doi.org/10.1016/j.sjbs.2019.06.008

[23] Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., ... & Pande, V. S. . OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS computational biology, 2017, 13(7), e1005659.

[24] Jo, S., Kim, T., Iyer, V. G., & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. Journal of computational chemistry, 2008, 29(11), 1859-1865.

[25] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., & Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. Journal of computational chemistry, 2011, 32 (10), 2319-2327.