 doi:10.20944/preprints202209.0218.v1

Article

Potential Inhibitors of EGFR Tyrosine Kinase from *Scutellaria baicalensis*: A Cheminformatics and Molecular Docking Studies

Amina J. Yusuf[1,*], Abayomi E. Adegboyega [2,3,*], Abdulbasit H. Yakubu [4], Grace I. Johnson[3,5], Rita O. Asomadu[6], Mary N. Adeduro[7], Ifeoma F. Chukwuma[8], Chinenye J. Ugwah-Oguejiofor[9], Olayinka S. Okoh[10,11], Afrah F. Alkhuriji[12], Wafa Abdullah I. Al-Megrin[13], Michel De Waard[14,15,16], Gaber El-Saber Batiha[17] and Titilayo O. Johnson[3,18*]

1 Department of Pharmaceutical and Medicinal Chemistry, Faculty of Pharmaceutical Sciences, Usmanu Danfodiyo University, P.M.B. 2346, Sokoto, Nigeria; amina.yusuf@udusok.edu.ng
2 Department of Biochemistry, Faculty of Basic Medical Sciences, College of Health Sciences, University of Jos, Plateau State, Nigeria; abayomiadegboyega5@gmail.com
3 Jaris Computational Biology Center, Jos, Plateau State, Nigeria
4 Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of University, P.M.B 1069, Maiduguri, Nigeria; pharmahy071@gmail.com
5 Faculty of Clinical Sciences, College of Health Sciences, University of Jos, Plateau State, Nigeria; johnsongrace003@gmail.com
6 Department of Biochemistry, Faculty of Biological Sciences, University of Nigeria, Nsukka; rita.asomadu@unn.edu.ng
7 Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Lagos, Nigeria; madeduro@unilag.edu.ng
8 Department of Biochemistry, Faculty of Biological Sciences, University of Nigeria, Nsukka, Nigeria; chukwuma.ifeoma@unn.edu.ng
9 Department of Pharmacology and Toxicology, Faculty of Pharmaceutical Sciences, Usmanu Danfodiyo University, Sokoto, Nigeria; nenye789@yahoo.com
10 Department of Chemical Sciences, Anchor University, Lagos, Nigeria; ookoh@aul.edu.ng
11 Center of Global Health, Anchor University, Lagos, Nigeria
12 Department of Zoology, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia; aalkhuriji@ksu.edu.sa
13 Department of Biology, College of Science, Princess Nourah bint Abdulrahman, University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; waalmegrin@pnu.edu.sa
14 Smartox Biotechnology, 6 rue des Platanes, 38120 Saint-Egrève, France; michel.dewaard@univ-nantes.fr
15 L'institut du thorax, INSERM, CNRS, UNIV NANTES, F-44007 Nantes, France
16 Université de Nice Sophia-Antipolis, LabEx «Ion Channels, Science & Therapeutics», F-06560 Valbonne, France
17 Department of Pharmacology and Therapeutic, Faculty of Veterinary Medicine Damanhour University Damanhour, 22511, AlBeheira, Egypt; gaberbatiha@gmail.com
18 Department of Biochemistry, Faculty of Basic Medical Sciences, College of Health Sciences, University of Jos, Plateau State, Nigeria; titijohnson2004@gmail.com
* Correspondence: amina.yusuf@udusok.edu.ng; abayomiadegboyega5@gmail.com; titijohnson2004@gmail.com; Tel.: +2348036386793

Abstract: Epidermal Growth Factor Receptor (EGFR) tyrosine kinase is a cell surface receptor whose overexpression has been associated with different types of cancers including brain cancer (glioblastoma multiforme). The ability of the extract of *Scutellaria baicalensis* to inhibit the proliferation of malignant glioma cells have been reported. Thus, in this study we report the identification of 307 bioactive constituents responsible for the anti-glioblastoma multiforme effect from *S. baicalensis* using in silico studies such as molecular docking, binding free energy calculations, pharmacophore modelling, induced-fit docking, gene enrichment analysis, molecular dynamic simulations and ADMET predictions. A total of 307 chemical constituents of *S. baicalensis* were screened and the top 10 scoring compounds indicated different binding affinities ranging from -9.010 to -6.427 kcal/mol towards the EGFR tyrosine kinase; Ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, and tenaxin I possess higher binding affinities (-9.010 to -8.649 kcal/mol) compared to the standard ligand, erlotinib having -8.539 kcal/mol. The

 doi:10.20944/preprints202209.0218.v1

compounds interacted with amino acids of clinical importance such as MET 769, GLU 738, THR 766 via H-bond. The structural features involved in the interaction with the target were mostly two aromatic rings, H-bond donors and acceptors and some hydrophobic interactions which varies between the ligands. Better docking scores in the induced-fit docking further validates the inhibitory potential of the compounds against the flexible protein. All the top-scoring ligands from *S. baicalensis* had zero Lipinski violation and also obeyed other drug-likeness rules by Ghose, Veber, Egan, and Muegge with the exception of breviscapine. Interestingly, all the compounds are not likely to be hepatotoxic, mutagenic, or cytogenic, making them potentially safe as anti-cancer agents.

Keywords: *Scutellaria baicalensis*; phytochemicals; glioblastoma multiforme; in silico

1. Introduction

Brain cancers (glioblastoma multiforme, GBM) are tumours that develop in the brain or spinal cord [1–2] which account for not less than 3 % of all cancers in the world [3]. It has been estimated that about 18,000 cases of the disease are diagnosed annually with an approximately 13,000 mortality in the USA alone with median prognosis of 15 months despite aggressive regimens such as surgery, radiation and chemotherapy [4–5]. The incidence of GBM is higher in developed countries compared to the developing and underdeveloped countries with male being the most victims [6]. Brain cancers are associated with certain risk factors which include family history, exposure to radiation (including cell phone radiation), alcohol intake, food allergies, hormonal factors, and smoking amongst others [7– 8].

The World Health Organization (WHO) classified brain tumor into various types based on their nature, cellular origin, rate of growth, histologic appearance and progression stage [8–9]; the primary tumor which occur due to DNA mutations invading the organs directly as a result of environmental, lifestyle, and eating habits and the secondary tumor (metastasis) which originates through blood vessels [10]. Symptoms of brain cancer are heterogeneous and include frequent headaches, mood swings, speech inadequacies, concentration problems, vision loss, paralysis, seizures and memory loss [1,8]. In various studies, there are correlations existing between degree of mutation in genes and type of brain tumor [1]. Mutations in certain types of genes determines the cancer type. Epidermal growth factor receptor (EGFR) is a trans-membrane receptor in the receptor tyrosine kinase (RTK) family. Mutation occurring in EGFR will lead to increased cell cycle proliferation and increased tumor cell survival. EGFR mutation is commonly connected with primary glioblastomas (brain tumor); more than 38 % of the mutations that caused them are found inside it [11].

The burden of brain cancer disease is high; its indolence and invasiveness added to absence of promising treatments renders it highly disabling, exorbitant, and generally burdensome for caregivers and health systems [5]. Intricacies and heterogeneity of the brain cancer-forming pathways makes conventional treatment nonresponsive. Even though treatment of brain cancer with conventional approach is quite promising, development of agents with reduced side effects and high target specificity has been a challenge. Traditional medicines are suggested as the safest source of drugs with several mode of actions and little or no side effects [12].

Scutellaria baicalensis Georgi (*Lamiaceae*) is a perennial herb indigenous to numerous East Asian countries and cultivated in many European countries [13]. It has been used in traditional medicine to treat various ailments such as respiratory infections, diarrhea, dysentery, inflammation, hypertension and insomnia [14]. The anti-GBM effect of *S. baicalensis* have been validated scientifically [15], but the chemical constituents responsible for the observed effect were not identified. Thus, the aim of this study was to employ different computational tools to identify multi-target ligands from characterized compounds of the plant against EGFR tyrosine kinase, a receptor connected to brain tumor.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

2. Results

2.1. Molecular docking analysis

Molecular docking studies of the chemical constituents of S. baicalensis (Figure 1) indicated different binding affinities ranging from -9.010 to -6.427 kcal/mol towards the EGFR tyrosine kinase. The top 10 constituents include ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, tenaxin I, chrysosplenetin, scutellarin, 5,2',5'-trihydroxy-7,8-dimethoxyflavone, 5,7,2',6'-tetrahydroxyflavone, viscidulin II and (2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one. Thus, ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, and tenaxin I possess higher binding affinities compared to the standard ligand, erlotinib having -8.539 kcal/mol (Table 1). Tables S1 and S2 contain the list of all the compounds of S. baicalensis and their docking scores (in kcal/mol) against the target protein (EGFR tyrosine kinase).

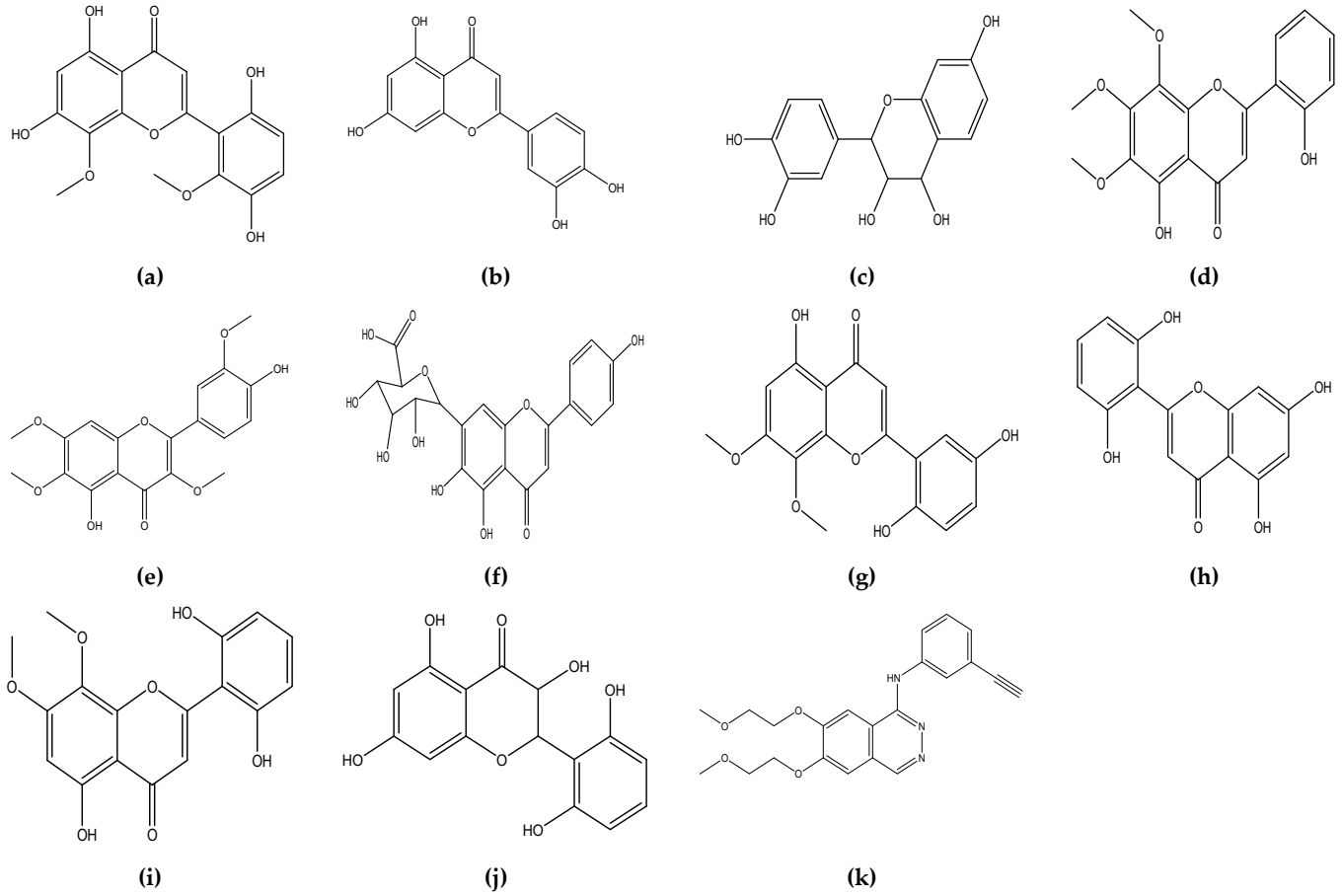


Figure 1. Chemical structures of top 10 chemical constituents of *S. baicalensis* and the standard ligand Erlotinib; (a) Ganhuangenin, (b) 5,7,2',5'-Tetrahydroxyflavone, (c) (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, (d) Tenaxin I, (e) Chrysosplenetin, (f) Scutellarin, (g) 5,2',5'-trihydroxy-7,8-dimethoxyflavone, (h) 5,7,2',6'-Tetrahydroxyflavone, (i) Viscidulin II, (j) (2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one, (k) Erlotinib (Standard ligand).

 doi:10.20944/preprints202209.0218.v1

Table 1. Top 10 screened compounds of *S. baicalensis* against EGFR tyrosine kinase.

Compound	Docking score
Ganhuangenin	-9.010
5,7,2',5'-Tetrahydroxyflavone	-8.911
(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol	-8.778
Tenaxin I	-8.649
Chrysosplenetin	-8.482
Scutellarin	-8.478
5,2',5'-trihydroxy-7,8-dimethoxyflavone	-8.462
5,7,2',6'-Tetrahydroxyflavone	-8.420
Viscidulin II	-8.397
(2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one	-8.365
Erlotinib (standard ligand)	-8.539

2.2. Biological interactions

Protein-ligand interactions of the EGFR tyrosine kinase and the compounds of S. baicalensis indicated that the compounds were able to interact with amino acids of clinical importance by forming hydrogen bonds with amino acid residues such as MET 769, GLU 738, THR 766 among others (Table 2 and Figures 2 – 3). Figures S1 and S2 shows the 2D and 3D views of the molecular interactions of the top ten ligands of S. baicalensis against EGFR tyrosine kinase.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

Table 2. Interactions of the top 10 compounds against the target protein of EGFR tyrosine kinase.

Compound ID	Compound name	Interactions	
		H-Bond	Hydrophobic and Others
5271991	Ganhuangenin	THR830, GLU738, THR766, MET769	CYS773, GLY772, LEU694, PRO770, LEU768, LEU820, GLN764, LEU764, ALA719, LYS721, MET742, PHE832, ASP831, VAL702, PHE699
5487756	5,7,2',5'-Tetrahydroxyflavone	MET769, THR766, GLU738	VAL702, PHE832, ASP831, THR830, MET742, LYS721, ALA719, LEU764, LEU820, GLN767, LEU768, PRO770, LEU694, GLY772
124708188	(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol	ASP831, THR766, MET769	VAL702, THR830, MET742, GLU738, LYS721, ALA719, LEU764, LEU820, GLN767, LEU768, PRO770, GLY772, LEU694
159029	Tenaxin I	MET769, THR766	LEU694, GLY695, PHE699, VAL702, THR830, ASP831, GLU738, MET742, CYS773, GLY772, PRO770, LEU768, GLN767, THR766, LEU764, LYS721, ALA719, LEU820
5281608	Chrysosplenetin	MET769, GLU738	GLY772, PRO770, LEU768, GLN767, THR766, ILE765, LEU764, ALA719, ILE720, LYS721, CYS751, LEU820, MET742, ASP831, THR830, VAL702, LEU694, GLY695
185617	Scutellarin	ASP776, ASP831, LYS721, MET769	GLY695, LEU694, THR830, ALA719, VAL702, LEU820, LEU768, PRO770, GLY772, CYS773
21122623	5,2',5'-trihydroxy-7,8-dimethoxyflavone	MET769, THR766, GLU738	VAL702, THR830, ASP831, PHE832, MET742, LYS721, ALA719, LEU764, GLN767, LEU820, LEU768, PRO770, GLY772, CYS773
5321865	5,7,2',6'-Tetrahydroxyflavone	ASP931, THR766, MET769	LEU694, VAL702, THR830, GLU738, MET742, LYS721, ALA719, LEU764, GLN767, LEU820, LEU768, PRO770, GLY772, CYS773
5322059	Viscidulin II	ASP831, THR766, MET769	VAL702, THR830, GLU738, MET742, LYS721, ALA719, LEU764, GLN767, LEU820, LEU768, PRO770, GLY772, CYS773, LEU694

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

86311089	(2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydro-chromen-4-one	MET769, THR766, THR830, ASP831	VAL702, GLU738, MET742, LYS721, ALA719, LEU764, GLN767, CYS751, LEU820, LEU768, PRO770, LEU694, GLY772, CYS773
176870	Erlotinib	CYS773, MET769	LEU694, VAL702, THR830, ASP831, GLU738, MET742, LEU753, LEU764, ILE765, THR766, GLN767, LEU768, LYS721, LEU820, ALA719, PRO776, PHE771, GLY772, ASP776

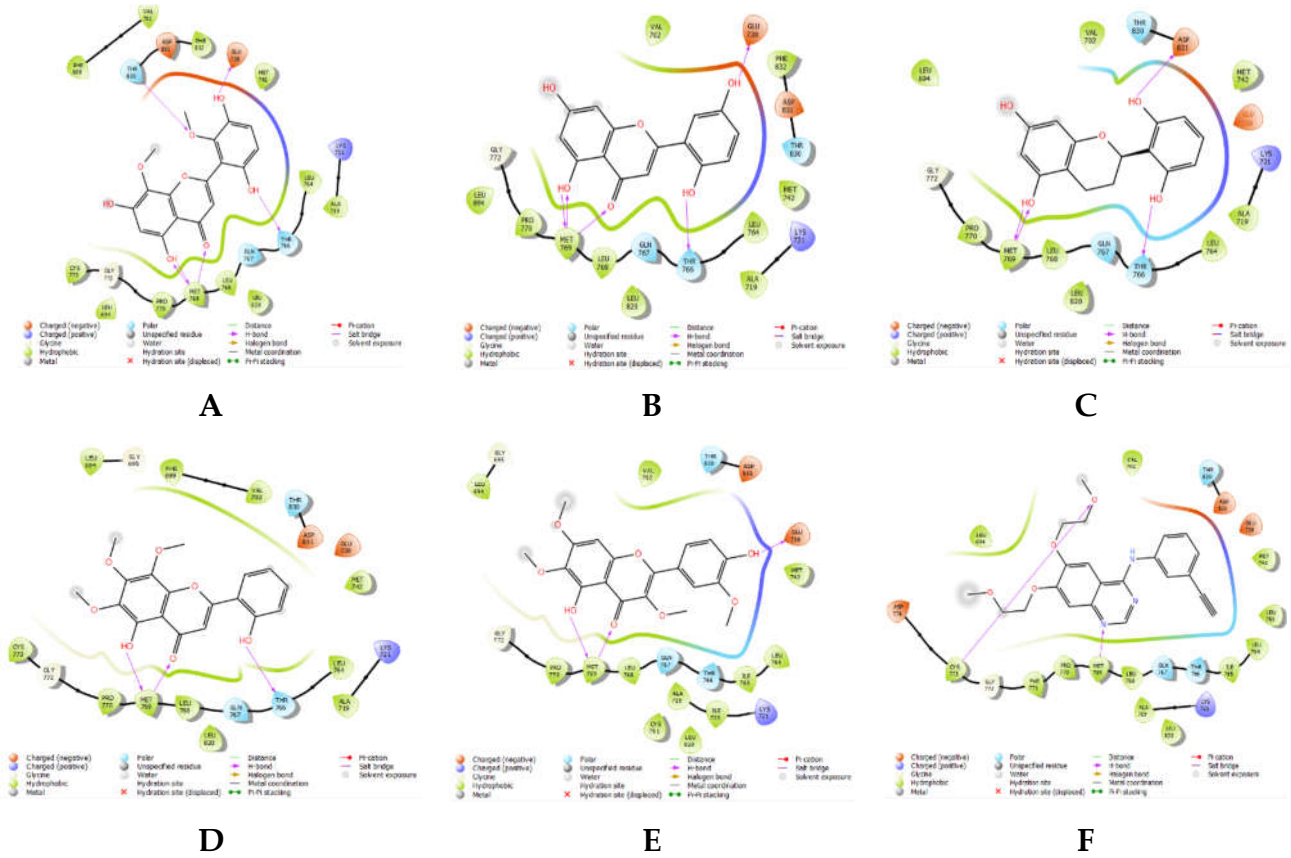


Figure 2. 2D view of the molecular interactions of A=Ganhuangenin, B=5,7,2',5'-Tetrahydroxyflavone, C=(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, D=Tenaxin I, E=Chrysosplenetin, F=Erlotinib with EGFR tyrosine kinase.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

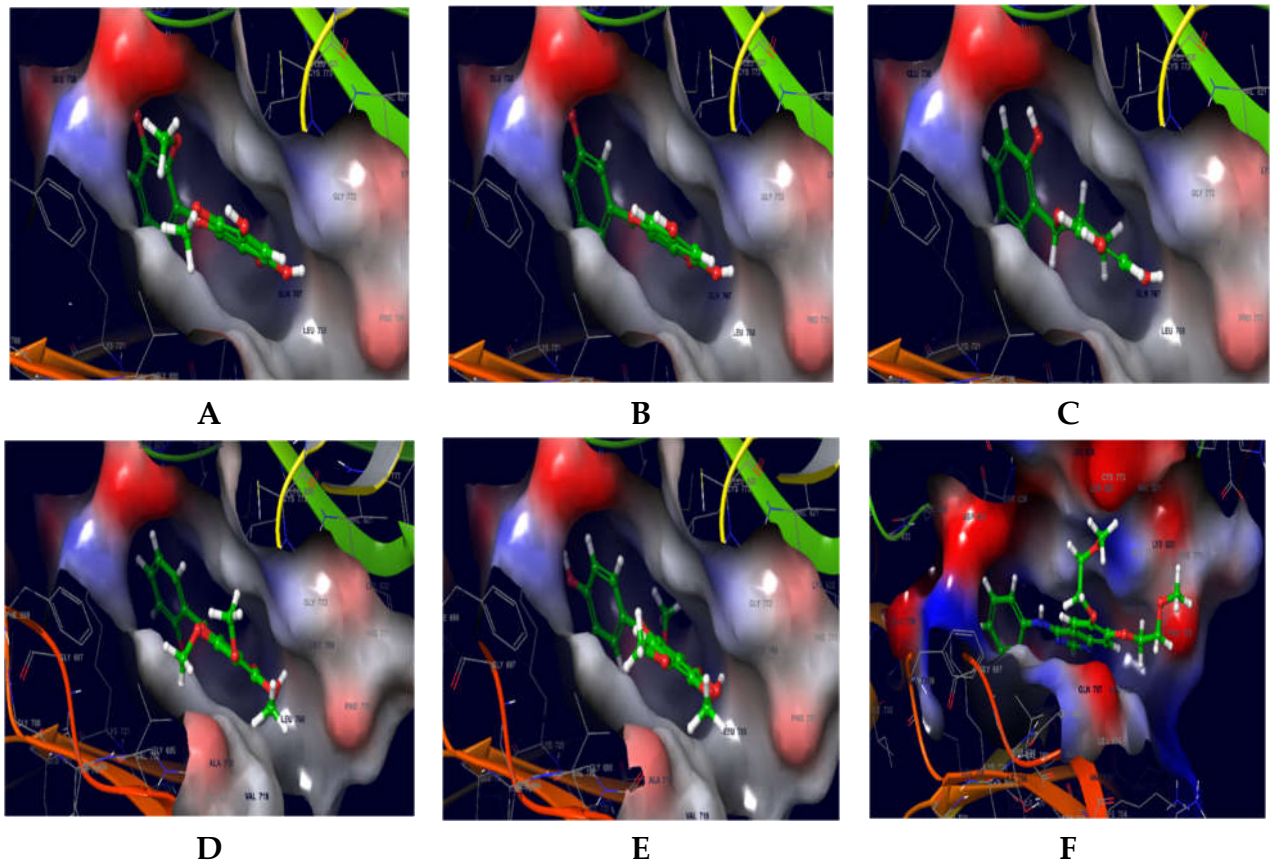


Figure 3. 3D view of the molecular interactions of A=Ganhuangenin, B=5,7,2',5'-Tetrahydroxyflavone, C=(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, D=Tenaxin I, E=Chrysosplenetin, F=Erlotinib with EGFR tyrosine kinase.

2.3. Binding free energy

The free binding energy computational analysis of the phytoconstituents obtained from *S. baicalensis* is shown in Table S3 while Figure 4 shows the scatter plot of the binding free energy; MMGBSA dG binding (ΔG Bind) versus docking score (kcal/mol) of the compounds against EGFR tyrosine kinase. The results obtained indicated that all the compounds have strong and good stability with the target protein; highest stability was obtained from (2R,3R,4S,5R,6S)-6-(5,6-dihydroxy-4-oxo-2-phenylchromen-7-yl)oxy-3,4,5-trihydroxyoxane-2-carboxylic acid (-56.02 kcal/mol), followed by chrysin-6-C-glucoside-8-C-arabinoside (-54.52 kcal/mol) and 2-(2,6-dihydroxyphenyl)-5,6,7,8-tetrahydroxychromen-4-one (-53.10 kcal/mol) respectively; while 2'-(β-D-glucopyranosyloxy)-5,6',7-trihydroxyflavone has the least binding energy (-33.89 kcal/mol). A positive correlation was obtained from the scatterplot in Figure 4 between the docking score and that of the free binding energy, presented by the regression line.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

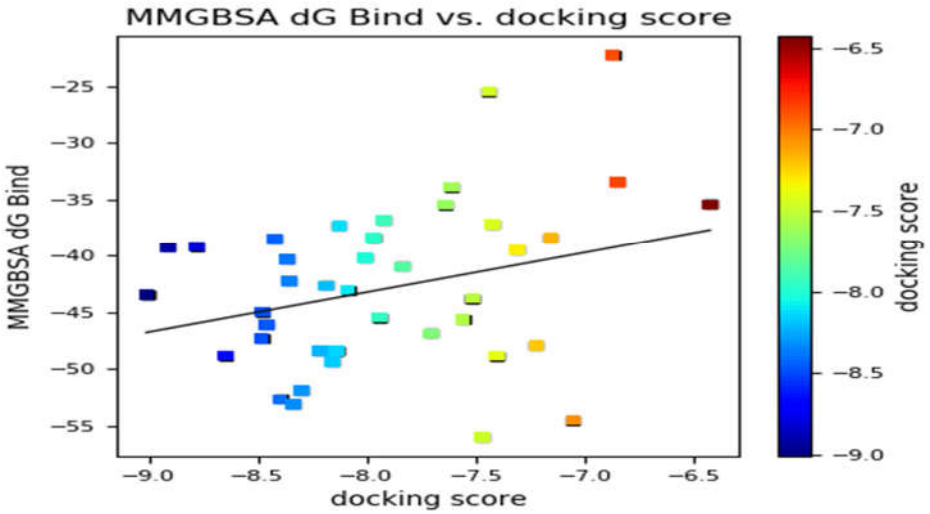


Figure 4. The binding free energy MMGBSA dG Bind (ΔG_{bind}) versus the docking score (kcal/mol) of the phytochemical constituents of *S. baicalensis* against EGFR tyrosine kinase target.

2.4. Pharmacophore modelling

The pharmacophore models of the top ten scoring ligands against EGFR tyrosine kinase are shown in Figure 5. The structural features revealed four sorts of characteristic involved in the interaction, which are A: Hydrogen acceptor, D: Hydrogen donor, H: Hydrophobic and R: Aromatic ring. Two aromatic rings, two hydrogen bond donors and acceptors contributed to the binding and molecular interactions of ganhuangenin, scutellarin and (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol to the target protein with the latter having only one hydrogen bond donor. The structural features of tenaxin I, chrysosplenetin, viscidulin II and (2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one) that contributed to the binding and molecular interactions towards the target were two aromatic rings, two hydrogen bond donors and acceptors while 5,7,2',5'-tetrahydroxyflavone and 5,7,2',6'-tetrahydroxyflavone had two aromatic rings, three hydrogen bond donors and acceptors and 5,2',5'-trihydroxy-7,8-dimethoxyflavone had only two aromatic rings and three hydrogen bond donors.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

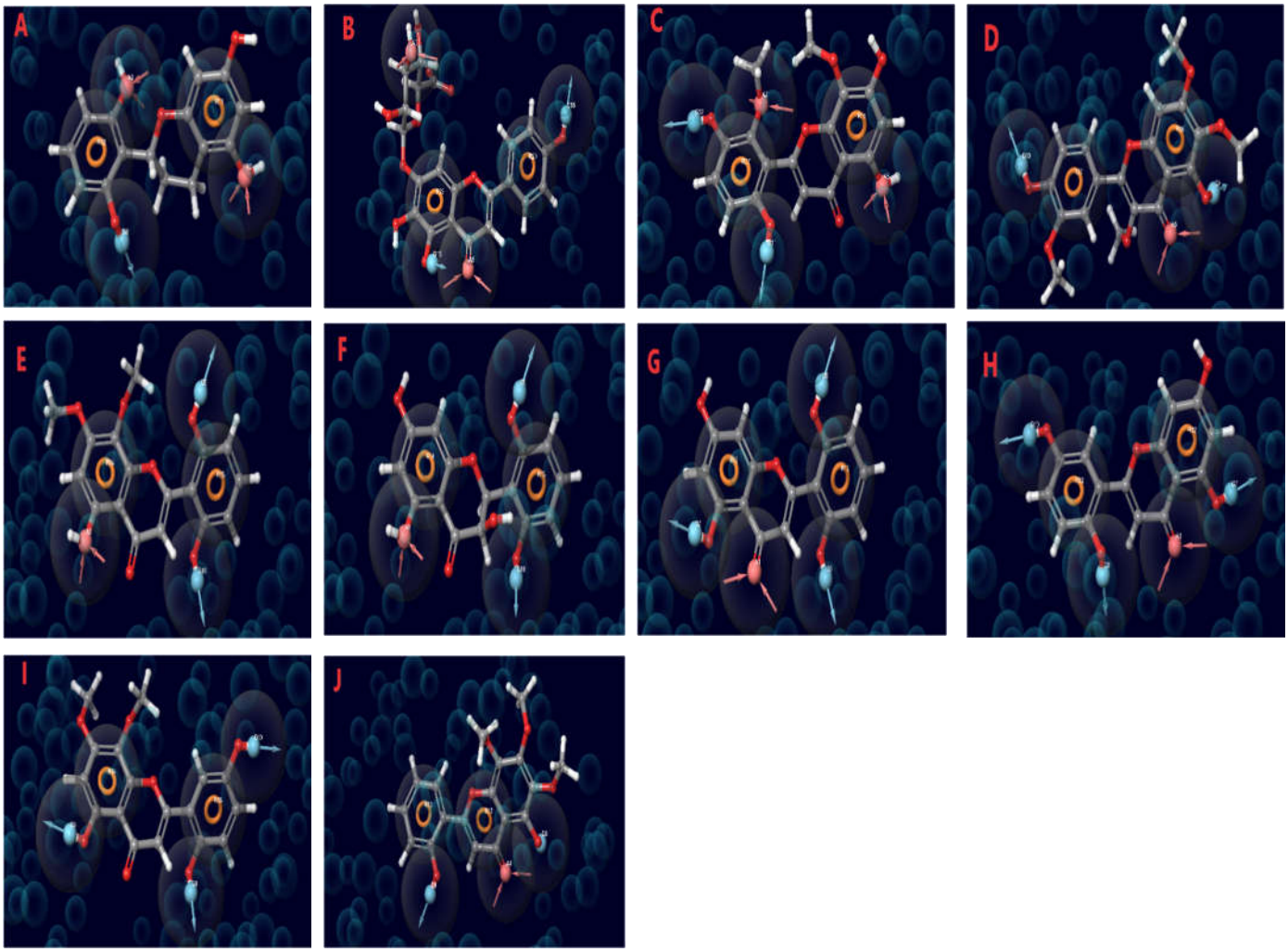
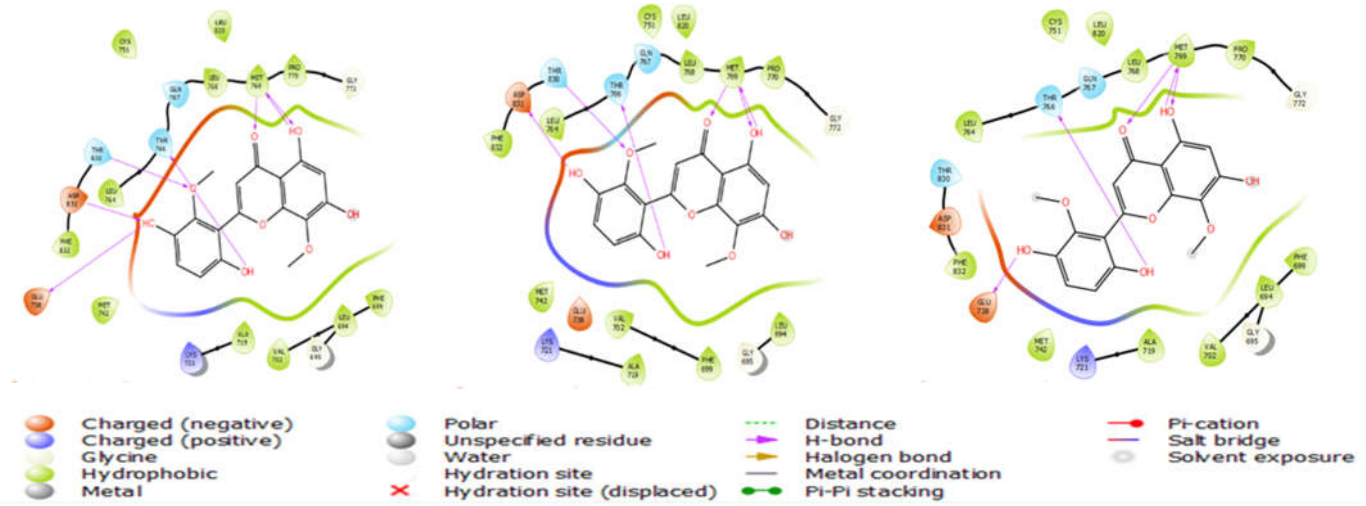


Figure 5. The receptor-ligand complex pharmacophore models of A=(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, B=Scutellarin, C=Ganhuangenin, D=Chrysosplenetin, E=Viscidulin II, F=(2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one, G=5,7,2',6'-Tetrahydroxyflavone, H=5,7,2',5'-Tetrahydroxyflavone, I=2-(2,5-Dihydroxyphenyl)-5-hydroxy-7,8-dimethoxy-4H-1-benzopyran-4-one, J=Tenaxin I on EGFR tyrosine kinase.

2.5. Induced fit docking

Induced fit docking analyses of ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, chrysosplenetin, tenaxin I and (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol was carried out. The docking scores are shown in Table S4 and the 2D representations of the conformations with the highest scores are shown in Figures 6 – 10. The first conformation of ganhuangenin with the target showed a hydrogen bond with GLU 738, ASP 831, THR 830, THR 766 and three with MET 769 as well as hydrophobic interactions with MET 742, PHE 832, LEU 764, CYS 751, LEU 768, LEU 820, PHE 699, MET 769, PRO 770, CYS 773, LEU 694, VAL 702 and ALA 719; the second conformation had similar interactions, however, there was no interaction with GLU 738 while the third conformation indicated a hydrogen bond with GLU 738, THR 766 and three with MET 769 (Figure 6).

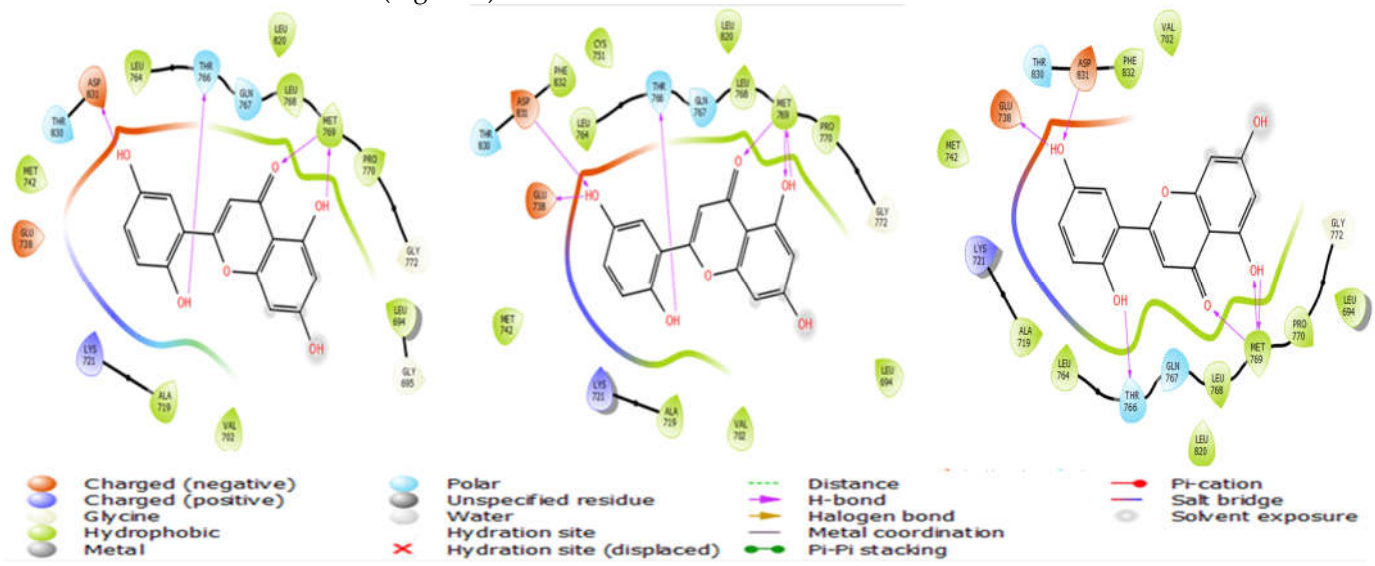
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1



(**a**) (**b**) (**c**)

Figure 6. 2D representation of Induced Fit Docking between Ganhuangenin and EGFR tyrosine kinase.

The first conformation of 5,7,2',5'-tetrahydroxyflavone formed one hydrogen bond with ASP 831, THR 766 and two with MET 769 while the hydrophobic interactions were similar to that of ganhuangenin, however, no interaction was noted with CYS 751; the second and third conformations had similar hydrogen bond interactions, but the compound formed additional H-bond with GLU 738 and hydrophobic interaction with CYS 751 (Figure 7).

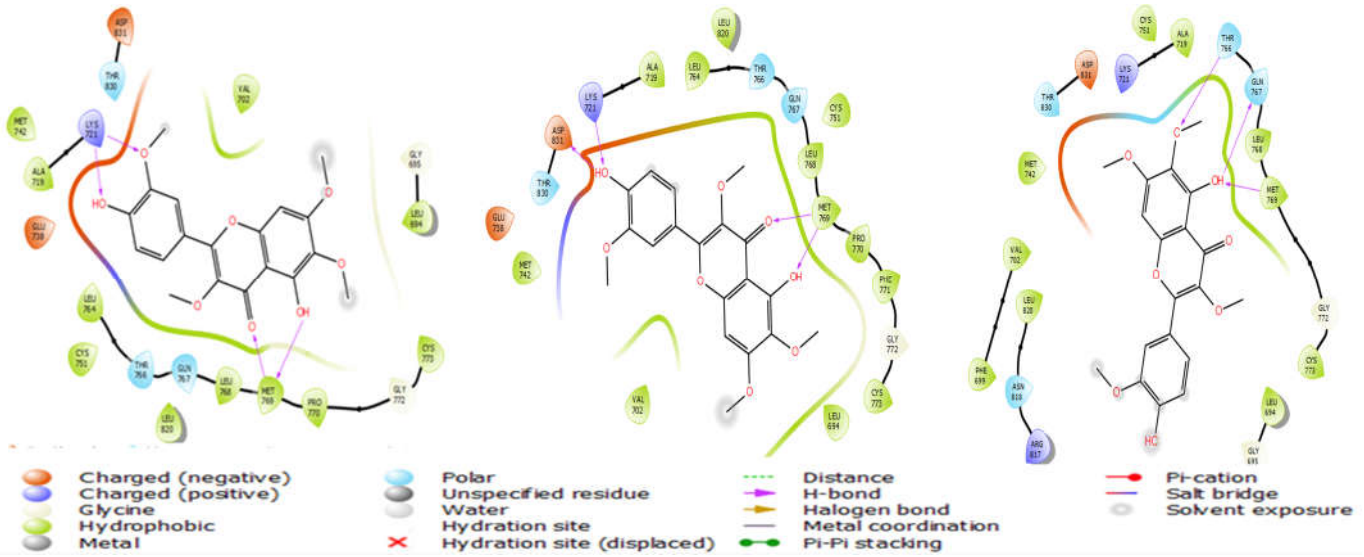


(**a**) (**b**) (**c**)

Figure 7. 2D representation of Induced Fit Docking between 5,7,2',5'-Tetrahydroxyflavone and EGFR tyrosine kinase.

The first two conformations of chrysosplenetin were though the first interacted with LYS 721 and MET 769 via H-bond while the second has one conformation with LYS 721, ASP 831 and two with MET 769; the third has H-bond interactions with THR 766, GLU 767 and MET 769 and no hydrophobic interactions with PHE 832 and PRO 770 compared to the first two (Figure 8).

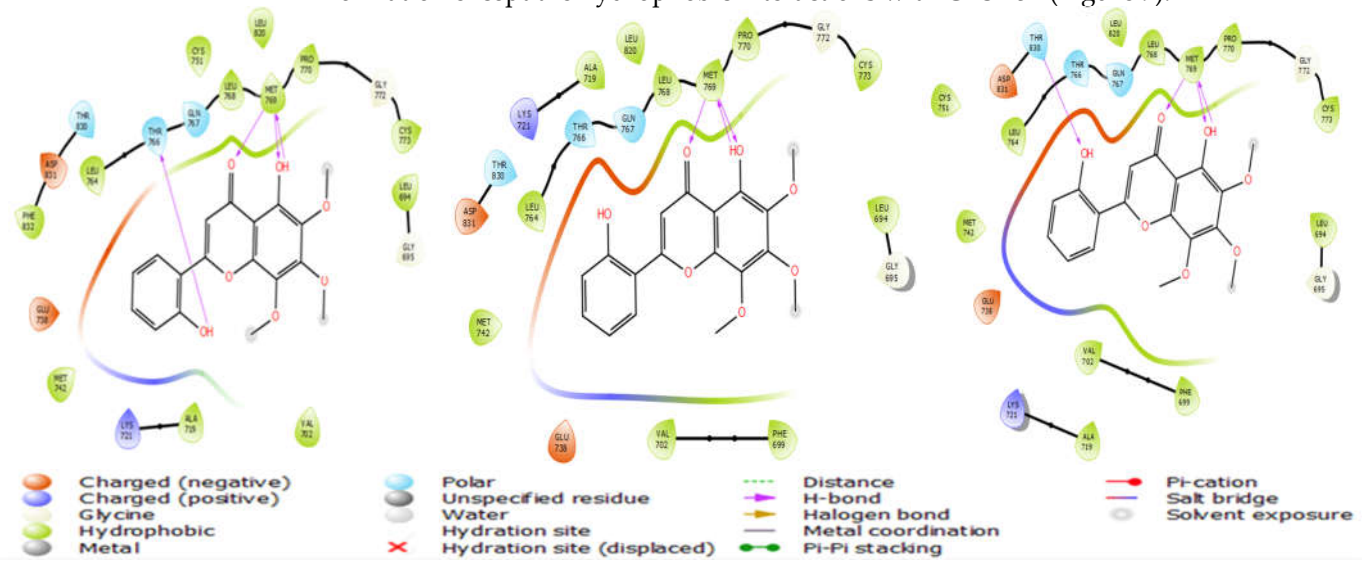
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1



(a) (b) (c)

Figure 8. 2D representation of Induced Fit Docking between Chrysosplenetin and EGFR tyrosine kinase.

The first conformation of tenaxin I with the target showed a hydrogen bond with THR 766 and three with MET 769 and hydrophobic interactions with MET 742, PHE 832, LEU 764, CYS 751, LEU 768, MET 769, PRO 770, CYS 773, LEU 694, VAL 702 and ALA 719; the second conformation had three hydrogen bonding interactions with MET 769 alongside additional hydrophobic interactions with PHE 699 and LEU 820, however, no interaction with CYS 751 and PHE 832 was observed while the third conformation interacted with THR 830 and MET 769 via H-bond and similar interactions as the second conformation except the hydrophobic interactions with CYS 751 (Figure 9).



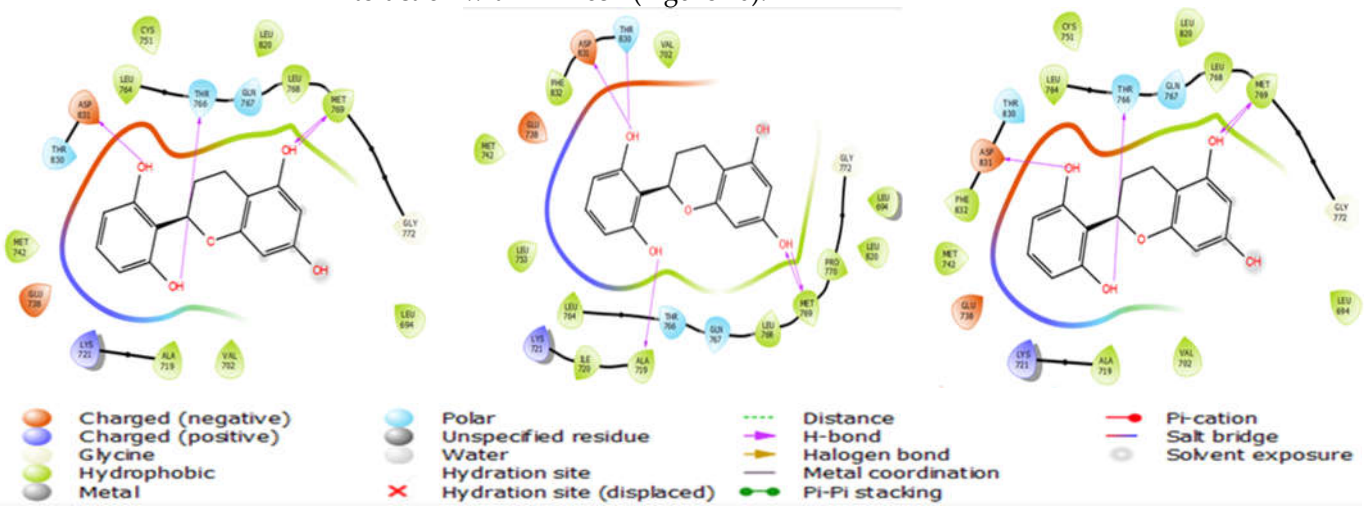
(a) (b) (c)

Figure 9. 2D representation of Induced Fit Docking between Tenaxin I and EGFR tyrosine kinase.

The first conformation of 2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol with the target revealed a hydrogen bond interactions with ASP 831, THR 766 and two with MET 769 and it indicated hydrophobic interactions with MET 742, LEU 764, CYS 751, LEU 768, MET 769, PRO 770, CYS 773, LEU 694, VAL 702 and ALA 719 while the second interacted with ASP 831, THR 830, ALA 719, MET 769 via H-bond alongside hydrophobic interactions similar to that of the first with an additional interactions with ILE

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

720 and PHE 832; the interactions of the third were also similar with the first but had an interaction with PHE 832 (Figure 10).



(**a**) (**b**) (**c**)

Figure 10. 2D representation of Induced Fit Docking between (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol and EGFR tyrosine kinase.

2.6. Gene enrichment analysis

2.6.1. Protein-protein interactions

The PPI network constructed by STRING revealed 11 nodes connected by 42 edges (Figure 11).

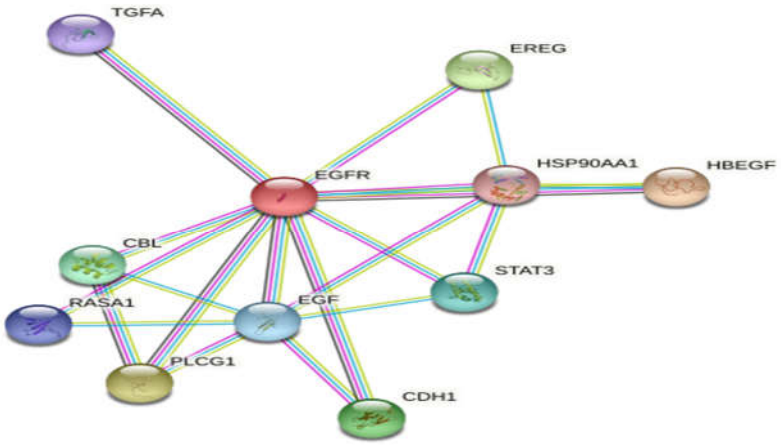


Figure 11. Protein-protein interaction (PPI) network of cancer-related targets interacting with EGFR.

2.6.2. Gene Ontology Functional Enrichment

Using the STRING database, gene ontology analysis revealed 7 molecular functions of identified genes. The most enriched functions include: protein binding, signaling receptor binding, enzyme binding, epidermal growth factor receptor binding and protein kinase binding (Table 3).

 doi:10.20944/preprints202209.0218.v1

Table 3. Molecular functions of identified genes.

Term ID	Term description	Observed gene count	False discovery rate	Matching proteins in network
GO:0005154	Epidermal growth factor receptor binding	5	3.56e-08	HBEGF, EREG, CBL, EGF, TGFA
GO:0005102	Signaling receptor binding	9	1.15e-05	HBEGF, PLCG1, EREG, CBL, STAT3, EGF, RASA1, EGFR, TGFA
GO:0008083	Growth factor activity	4	0.0012	HBEGF, EREG, EGF, TGFA
GO:0005515	Protein binding	11	0.0083	HBEGF, PLCG1, EREG, CDH1, CBL, STAT3, EGF, RASA1, EGFR, TGFA, HSP90AA1
GO:0019901	Protein kinase binding	5	0.0089	PLCG1, CBL, STAT3, EGFR, HSP90AA1
GO:1990782	Protein tyrosine kinase binding	3	0.0102	PLCG1, CBL, HSP90AA1
GO:0019899	Enzyme binding	7	0.0202	PLCG1, CBL, STAT3, EGF, RASA1, EGFR, HSP90AA1

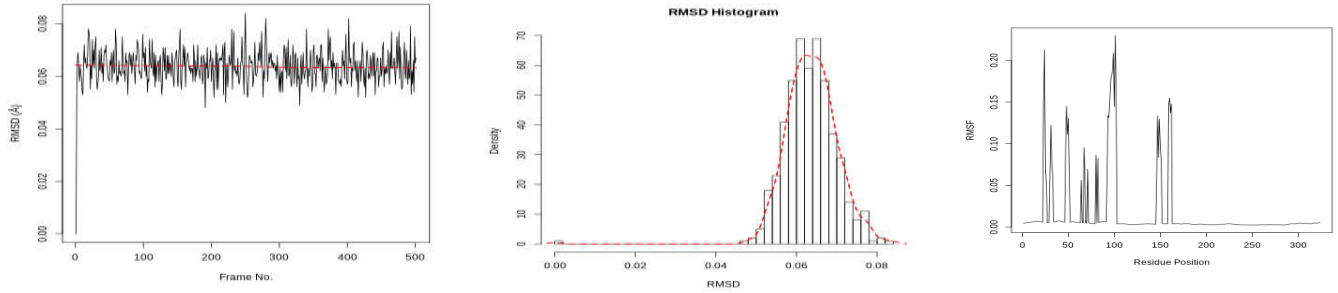
2.6.3. KEGG pathway analysis of potential target genes functions

Most enriched pathway include: Pathways in cancer, ErbB signaling pathway, non-small cell lung cancer, EGFR tyrosine kinase inhibitor resistance, proteoglycans in cancer, Ras signaling pathway, MAPK signaling pathway and PI3K-Akt signaling pathway (Table S5).

2.7. Molecular dynamics simulation

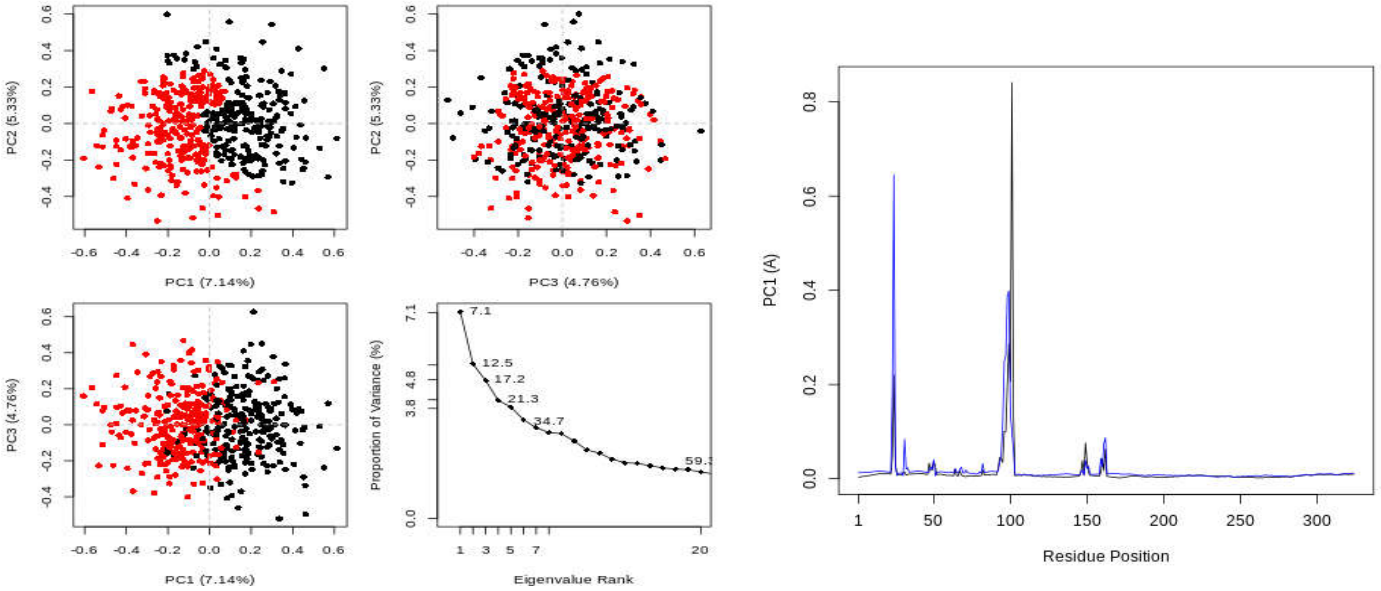
In order to assess the stability of the simulation system, the Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) of all Cα atoms in the entire Molecular dynamics (MD) trajectories were obtained as indicated in Figure 12. Figure 13 shows the results of the Principal component analysis (PCA) which include the PCA scatter plots of PC2 vs PC1, PC2 vs PC3, PC3 vs PC1 colored from red and black in order of time, and an eigenvalue rank plot for EGFR tyrosine kinase complex as well as the plot of PC1 versus Residue position. The top 20 PCs of the EGFR tyrosine kinase-ligand complex system generated 59.3% of the total contribution to the simulated trajectories. The first three PCs (PC1, PC2, and PC3) contributed 7.1, 5.4, and 4.7%, respectively, to the EGFR tyrosine kinase system.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1



(a) (b) (c)

Figure 12. Molecular dynamics simulation analysis. (a) The RMSD values of all Cα atoms for the EGFR tyrosine kinase complex over 500 ns MD simulation; (b) the RMSF fluctuations of residues for EGFR tyrosine kinase complex over 500 ns MD simulation; (c) Histogram representing the deviation points that occur in 1-nanosecond (μs) molecular dynamics (MD) simulation.



A B

Figure 13. Principal component analysis (PCA) results which include variance contribution of the principal components and PCA scatter plots graphs of PC2 vs PC1, PC2 vs PC3, PC3 vs PC1 colored from red and black in order of time, and an eigenvalue rank plot for EGFR tyrosine kinase complex (A) and PC1 versus Residue position (B).

2.8. Drug likeness and pharmacokinetics prediction output of test compounds

The drug likeness and pharmacokinetics of the phytoconstituents of *S. baicalensis* are presented in Table 4. All the top scoring compounds had molecular weight ranging from 274.27 – 462.36 g/mol, H-bond donors from 2 – 7, H-bond acceptors from 5 – 12, iLog P values from 1.66 – 3.35, Nrot from 1–5, molar refractivity values from 73.17–108.74, and TPSA values from 90.15–207.35A^2. The test compounds had zero Lipinski, Ghose, Veber, Egan, and Mugge violations; a favorable bioavailability score of 0.55; and high GI absorption except breviscapine, which violated Lipinski, Veber, Egan, and Mugge rules and had a bioavailability score of 0.11 and low GI absorption. In addition, all compounds showed no BBB permeability, but 2',5,6',7-tetrahydroxyflavane, and breviscapine were Pgp substrates, which will restrict them from keying into the site of action. All the ligands are potential non-inhibitors of CYP2CI9, whereas breviscapine and 2',3,5,6',7-pentahydroxyflavanone are inhibitors of CYP1A2 and CYP3A4. Furthermore, all the compounds

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

are expected to be permeable to the skin due to their LogKp values ranging from 6.16 – 8.59 (cm/s).

Table 4. Drug likeness and pharmacokinetics prediction output of test compounds.

Parameters	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Drug likeness										
M.W (g/mol)	346.29	286.24	274.27	344.32	374.34	462.36	330.29	286.24	330.29	304.25
# Rotatable bonds	3	1	1	4	5	4	3	1	3	1
iLog P	2.37	1.77	1.71	3	3.35	1.11	2.53	1.7	2.43	1.66
# H-bond acceptor	8	6	5	7	8	12	7	6	7	7
# H-bond donor	4	4	4	2	2	7	3	4	3	5
Molar refractivity	89.00	76.01	73.17	91.44	97.93	108.74	86.97	76.01	86.97	74.76
TPSA (A^2)	129.59	111.13	90.15	98.36	107.59	207.35	109.36	111.13	109.36	127.45
Lipinski violations	0	0	0	0	0	2	0	0	0	0
Ghose violations	0	0	0	0	0	0	0	0	0	0
Veber violations	0	0	0	0	0	1	0	0	0	0
Egan violations	0	0	0	0	0	3	0	0	0	0
Muegge violations	0	0	0	0	0	1	0	0	0	0
Bioavailability score	0.55	0.55	0.55	0.55	0.55	0.11	0.55	0.55	0.55	0.55
Pharmacokinetics										
GI absorption	High	High	High	High	High	Low	High	High	High	High
BBB permeant	No	No	No	No	No	No	No	No	No	No
P-gp substrate	No	No	Yes	No	No	Yes	No	No	No	No
CYP1A2 inhibitor	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No
CYP2C19 inhibitor	No	No	No	No	No	No	No	No	No	No
CYP2C9 inhibitor	Yes	No	No	Yes	Yes	No	Yes	No	Yes	No
CYP2D6 inhibitor	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	No
CYP3A4 inhibitor	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Log K_p (cm/s) (skin permeation)	-6.81	-6.16	-6.28	-6.31	-6.37	-8.59	-6.46	-6.16	-6.46	-7.48

Key: M1: Ganhuangenin, M2: 5,7,2',5'-Tetrahydroxyflavone, M3: 2',5,6',7-Tetrahydroxyflavane, M4: Tenaxin I, M5: Chrysosplenetin, M6: Breviscapine, M7: 5,2',5'-Trihydroxy-7,8-Dimethoxyflavone, M8: 5,7,2',6'-Tetrahydroxyflavone, M9: Viscidulin II and M10 2',3,5,6',7-Pentahydroxyflavanone

2.9. Toxicity prediction output of test compounds

Acute toxicity prediction outcome in Table 5 revealed that all the lead compounds from *S. baicalensis* had LD_{50} > 2000 mg/kg belonging to mainly class 5 with the exception of 2',3,5,6',7-pentahydroxyflavanone which belong to class 4. None of the test compounds showed a tendency for hepatotoxicity, mutagenicity, or cytotoxicity. Of all the ten compounds, breviscapine is likely to be carcinogenic while all the compounds except 2',5,6',7-tetrahydroxyflavane, Breviscapine, and 5,7,2',6'-tetrahydroxyflavone are likely to be immunotoxic.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

Table 5. Toxicity prediction output of test compounds.

Target	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Predicted LD_{50} (mg/kg)	3919	3919	2500	3919	5000	5000	3919	3919	4000	2000
Predicted Toxicity Class	5	5	5	5	5	5	5	5	5	4
Hepatotoxicity	-	-	-	-	-	-	-	-	-	-
Carcinogenicity	-	-	-	-	-	+	-	-	-	-
Immunotoxicity	+	+	-	+	+	-	+	-	+	+
Mutagenicity	-	-	-	-	-	-	-	-	-	-
Cytotoxicity	-	-	-	-	-	-	-	-	-	-

Key: +=Active; - = Inactive; M1: Ganhuangenin, M2: 5,7,2',5'-Tetrahydroxyflavone, M3: 2',5,6',7-Tetrahydroxyflavane, M4: Tenaxin I, M5: Chrysosplenetin, M6: Breviscapine, M7: 5,2',5'-Trihydroxy-7,8-Dimethoxyflavone, M8: 5,7,2',6'-Tetrahydroxyflavone, M9: Viscidulin II and M10 2',3,5,6',7-Pentahydroxyflavanone

3. Discussion

New drug discovery traditionally last between 10 – 20 years [16–17]. The presence of diverse chemical matrix in the plant extract makes the process of obtaining drugs from natural origin cumbersome [18]. Hence, the versatile computational modelling and docking studies have now greatly aided quicker drug target identification and design of potential anticancer ligands that are both sterically and chemically compatible with the binding site of target molecules [17,19]. Thus, in this study, 307 compounds of *S. baicalensis* were screened for their inhibitory effect against EGFR tyrosine kinase and 10 compounds were selected for further analysis based on their docking scores viz ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, tenaxin I, chrysosplenetin, scutellarin, 5,2',5'-trihydroxy-7,8-dimethoxyflavone, 5,7,2',6'-tetrahydroxyflavone, viscidulin II and (2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one.

EGFR play a vital role in regulating several cell phenotypes including cell migration, adhesion, proliferation and immune response in the human skin and thus, this makes it an important key target for inhibiting the growth of tumors [18]. Of the 10 compounds studied, ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol and tenaxin I had higher binding affinities against the target protein compared to the standard ligand, Erlotinib. In a study by Singh and Bast [20], the standard ligand, Erlotinib was able to interact and form hydrogen bond with MET 769, LEU 820, LEU 768, GLY 772, and LEU 694 and was involved in hydrophobic interaction with EGFR tyrosine kinase inhibitor. MET 769 and ASP 800 played a vital role in the inhibitory activity of EGFR [21]. Thus, further validating our findings.

The MM/GBSA methods are among the widely regarded and accepted end-point methods in calculating absolute free binding energy as they are more accurate and computationally less demanding and do not require the simulation of intermediate states as in the alchemical methods including thermodynamic integration (TI) and free-energy perturbations (FEP) methods, which are computationally expensive [22 –24]. This method involves the calculation of the solvation free energy contribution by solving the linearized Poisson Boltzmann or Generalized Born equation, the gas phase energy contribution by performing a molecular mechanics calculation, and the entropy contribution by performing normal mode analysis [22]. The stability of protein-ligand complexes is determined by binding free energy, which is one of the most accurate post docking tools for confirming docking score data. This means that a favorable binding free energy coincides with a reliable molecular docking score result. The higher the binding free energy, the more favorable and stable the ligand-bound protein [25]. The results of this investigation

 doi:10.20944/preprints202209.0218.v1

demonstrated that when the phytoconstituents of *S. baicalensis* bound to the drug targets, they might produce a stable complex, confirming the reliability of the docking score.

Pharmacophore modeling is one of the parameters used in evaluating potential drug candidate which gives detailed information on the molecular structures and the interactions involve in the protein-ligand interaction such as hydrogen bonding and aromatic rings [26]. In this study, the target protein of interest is EGFR tyrosine kinase, which is one of the transmembrane growth factor receptor protein [27]. The function is activated by it binding with its ligand epidermal growth factor leading to autophosphorylation and activation of signal pathway promoting proliferation [27]. The pharmacophore modeling showed the involvement of aromatic rings and hydrogen bond formation, in the interaction of the test ligands with the enzyme, which might have contributed to the higher binding affinity of the ligands. Aromatic rings are important residues for molecular interactions and they often exist in several protein-ligand and protein-protein interactions. Owing to their natural existence in amino acid residues like histidine, tryptophan, phenylalanine, and tyrosine, they are considered essential for protein stability and molecular recognition processes [28]; therefore, they play an important role in the improvement of binding affinity and specificity of drug like molecules [29] as indicated in our findings.

Induced fit docking, a computational approach for modeling the conformational changes prompted by ligand binding was carried to determine the best pose produced after the binding of the compounds with the target [17]. The findings indicated that the five compounds had better docking scores than that of the standard precision docking, signifying that they might have better inhibitory activities at such conformations. The poses formed hydrogen bond interactions with MET 769 and hydrophobic interactions with LEU 694, VAL 702, ASP831, LEU 764, ILE 765, THR 766, LEU 820, ALA 719, and ASP 776 similar to the standard ligand. More so, MET 769 is one of the most important amino acids that inhibit EGFR tyrosine kinase [21], further expressing the potentials of the selected compounds as inhibitors of the receptor.

EGFR/ErbB1 is implicated in diseases of cellular proliferation [30], which is the only enriched disease in this study. Its gene amplification and overexpression is well-recognized in cancer, especially brain cancer [31–34]. Therefore, studying the EGFR signaling network and its associated pathways are essential to enhance drug efficacy, demystify the mechanisms of drug resistance, and develop better therapeutic agents. Targeting EGFR signaling is therefore necessary for therapeutic intervention. EGFR can be activated by high-affinity ligands: pro-epidermal growth factor (EGF), protransforming growth factor alpha (TGFA), proheparin-binding EGF-like growth factor (HBEGF) and betacellulin (BTC). It can also be activated by low-affinity ligands, such as proepiregulin (EREG), epigen (EPGN) and amphiregulin (AREG) [35]. Some of these ligands are shown in the PPI network as the human EGFR interacted with 10 proteins which are involved in mediating the functions of EGFR (Figure 1). The other genes which might play important roles in mediating the oncogenic actions of EGFR include: 1-phosphatidylinositol-4,5-biphosphate phosphodiesterase gamma-1 (PLCG1), cadherin-1 (CDH1), E3-ubiquitin protein ligase CBL (CBL), signal transducer and activator of transcription 3 (STAT3), Ras GTPase-activating protein 1 (RASA1) and heat shock protein HSP 90-alpha (HSP90AA1). The functions of gene groups are effectively studied using gene enrichment and annotation [36]. Molecular functions of the genes were found to be: EGFR receptor binding, signaling receptor binding, growth factor activity, protein binding, protein kinase binding, protein tyrosine kinase binding and enzyme binding. By binding to EGFR, the active constituents of *S. baicalensis* are probably able to modulate the oncogenic actions of the receptor through the afore-mentioned molecular functions, especially protein binding which involve all the identified genes as shown in Table S5. Three domains: EGF-like domain, Src homology 2 (SH2) domain and SH2 domain superfamily were significantly enriched in this study. p85α SH2 domain mutation is involved in activating a spectrum of receptor tyrosine kinases, including EGFR [37].

KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed to recognize the signaling pathways associated with the identified target genes

 doi:10.20944/preprints202209.0218.v1

[38]. KEGG pathway analysis showed that the targets are involved in 47 pathways, with majority of the pathways implicated in cancer. EGFR acts in 40 of the identified pathways. Targeting EGFR will therefore reduce its influence on cancer-related pathways, especially pathways in cancer and ErbB signaling pathway which are the most enriched pathways with 7 and 8 genes, respectively. These pathways also have the lowest false discovery rate. Worthy of note is the role of EGFR, PLCG1, EGF and TGFA in glioma, as the role of EGFR in gliomas is well-recognized [34], and it is the most amplified receptor tyrosine kinase in glioblastoma, promoting invasion, proliferation and resistance to therapy [35]. The KEGG pathway analysis also revealed the non-small cell lung cancer, EGFR tyrosine kinase inhibitor resistance, proteoglycans in cancer, MAPK, PI3K-AKT and Ras pathways as part of the most enriched pathways with 5 genes (Table 2). It has been shown that activation of EGFR leads to the downstream activation of the RAS/RAF/MAPK pathway, phosphatidylinositol-3-kinase (PI3K)/protein kinase (AKT) pathway, src kinase and janus kinase (JAK)/signal transducers and activators of transcription (STAT) signaling pathway [33–35]. The PI3K-AKT pathway involves various processes such as protein synthesis, cell cycle progression, angiogenesis, cell proliferation, metabolism, DNA repair and cell survival [39]. Proteoglycans in the cancer pathway showed that proteoglycans are involved in the growth and metastasis of cancer cells [40–41]. MAPK pathway regulates cell differentiation, proliferation and stress responses [42]. These pathways are therefore important for cancer therapy. EGFR was found to be one of the hub genes of *S. baicalensis* in the treatment of Oral leukoplakia - which is a precancerous condition - through the PI3K/AKT signalling [43]. Ethanol root extract of *S. baicalensis* was also effective in treating EGFR tyrosine kinase inhibitors (TKIs)-resistant cells in lung cancer via inactivation of STAT3 [44–45]. This EGFR tyrosine kinase inhibitor resistance pathway was revealed to be one of the most enriched pathways in this research. Thus, *S. baicalensis* may probably play its anti-cancer effect through EGFR via the most enriched KEGG pathways.

In order to assess the stability of the simulation system, the RMSD values of all Cα atoms in the entire MD trajectories were obtained as indicated in Figures 12 - 13. The calculation of a time variable concerning the root means square deviation (RMSD) across Cα atoms from generated trajectories were executed to determine the consistency and efficiency of the simulated EGFR tyrosine kinase complex. Generally, the lower the RMSD value of the skeleton, the higher the stability of the system [46]. The association between the RMSD of the EGFR tyrosine kinase complex system and frame number (500ns) is shown in Figure 12a. It was discovered that the RMSD trajectories of the EGFR tyrosine kinase complex system stabilized at 0.06 to 0.08Å during the simulation and attained equilibrium at roughly 10 ns, plateauing at 0.08Å. As a result, the EGFR tyrosine kinase complex was stable, supporting the idea that the ligand maintained a stable conformation in the active site of the protein. The RMSD density graphs shown in Figure 12b support this. The histogram's peaks indicate that just one major conformation was present and that it was the one that was accessed during the route trajectory. Between the range count of 55 and 68 times, or approximately, the major count of the RMSD was between 0.06 and 0.08 nm. The RMSD graph and this histogram result demonstrate the stability of the EGFR tyrosine kinase complex. The RMSF of all the side chain atoms in the protein was examined in order to learn more about the flexibility of each amino acid residue in the protein system. The RMSF value for each residue in the EGFR tyrosine kinase complex system is displayed in Figure 12c. The calculation determined that the complex system's average RMSF ranged from 0.06 to 0.26 ns. As shown in the figure, flexible zones were defined as values over 0.16 nm. The flexibility of the residues in the simulation system is typically inversely correlated with the RMSF value. Conversely, the smaller the RMSF value, the lower the flexibility of the residues [46]. Higher RMSF values are most likely loop regions that confer to more conformational flexibility.

We can explore the structure and behavior of molecular systems in great detail using Classical Molecular Dynamics (MD) simulations. Numerous analysis techniques have been created in order to obtain a brief but accurate interpretation of the ever-increasing amount of simulation data. These techniques include network theory, Markov state

 doi:10.20944/preprints202209.0218.v1

models, stochastic approaches like the Langevin equation, and a variety of dimensionality reduction techniques. Principal component analysis (PCA), a straightforward and popular method, is a linear transformation that diagonalizes the covariance matrix and eliminates the instantaneous linear correlations between the coordinates [47]. In this study, PCA was performed on the simulation trajectories of EGFR tyrosine kinase protein system and EGFR tyrosine kinase-Ligand complex system to investigate the influence of ligand on the conformational state of the target protein. It showed the statistically meaningful conformations in the EGFR tyrosine kinase-ligand complex trajectory. It is possible to identify the main motions in the trajectory and the crucial motions required for conformational changes. The associated eigenvalues represent the percentage change of the fluctuation of the atomic location captured in each dimension, while the eigenvector depicts the overall motion of the Cα atom in the simulation system [46]. Figure 13a indicated that 59.3 % of the total contribution of the simulated trajectories was contributed by the top 20 PCs of the EGFR tyrosine kinase-ligand complex system. The first three PCs (PC1, PC2, and PC3) provided 7.1, 5.4 and 4.7 % to the EGFR tyrosine kinase system, respectively. The first three eigenvectors could essentially reflect the changes in the conformational state, as shown by the fact that PC1, PC2, and PC3 recorded the majority of modifications in the original distribution of the conformational space of the simulation system. Projecting along the first three PC1, PC2, and PC3 directions yielded the PCA scatter plots for these simulation systems. The simulation system alternated between the two conformational states depicted in the picture by the red and black dots, which stood in for the unstable and stable conformational states, respectively. The scattered dots were arranged into two discrete groupings along the PC1 plane, and the clustering of the red and black points was evenly distributed parallel on both sides of the diagonal, indicating a non-periodic conformational shift. The points were dispersed in an erratic and disorganized configuration, suggesting that the motions are periodic, although the groupings along the PC2 and PC3 planes do not entirely cluster apart. This suggests that PC1 is associated with an active site motion that restricts the motion to a crucial region in the target protein's amino acid residues (Figure 13b), as well as its stability as indicated by the RMSD and RMSF studies

In the drug discovery and development process, about 50 and 40 % of drug candidates failed clinical trials due to unacceptable efficacy and toxicity, respectively [48]. Recently, in silico models such as SwissADME and PROTOX-II web servers are employed in drug development to predict the druggability and safety of drug candidates with high precision thereby saving time and cost of developing drugs that will fail clinical trials [25,49]. Drug likeness prediction of the test compounds was conducted based on the rule of five (RO5) by Lipinski et al. [50] which states that an orally active drug must not violate more than one of these criteria's viz MW ≤ 500, HBD ≤ 5, HBA ≤ 10 and log P ≤ 5. All the top-scoring ligands from *S. baicalensis* had zero Lipinski violation and also obeyed other drug-likeness rules by Ghose, Veber, Egan, and Muegge with the exception of breviscapine. Potential drug candidates that obey drug-likeness rules tend to have a lower attrition rate in phases I, II, and III clinical trials [51]. This suggests that *S. baicalensis* ligands could be used as oral drug for the treatment of brain cancer. Hence, the efficacy of *S. baicalensis* as an anticancer agent for glioblastoma multiforme reported by Scheck et al. [15] could be attributed to these metabolites. Their potential as good oral drugs was also validated by their bioavailability scores of 0.55; this bioavailability score predicts that the ligands have about a 55 % probability of a minimum of 10 % oral absorptivity in rats or humans [26].

The pharmacokinetic behavior of drugs can easily be evaluated with gastrointestinal absorption (GIA) and brain access [52]. All the compounds but breviscapine showed high GIA which is an indication that they can easily be absorbed in the intestinal tract upon oral administration. However, breviscapine with low GIA could be administered by other routes such as lipoidal diffusion, para-cellular diffusion, passive, and uptake by transporter protein [53]. In the same vein, drug metabolism is vital for the optimization of lead compounds for optimal pharmacokinetics, and pharmacodynamic properties, identification of new chemical entries, minimization of toxicity effects of drugs, and prediction of appropriate human dose [54]. The first-pass metabolism or biotransformation of drugs

 doi:10.20944/preprints202209.0218.v1

and xenobioticsis done by cytochrome P450 (CYP450) isoenzymes [55]. The isoforms CYP450 2C19 and 2C9 oxidases mediate the plasma concentrations of the drugs, while CYP450 2D6 and 3A4 are responsible for the O-dealkylation reactions [56]. All the screened compounds of *S. baicalensis* were non-inhibitors of CYP2C19, whereas breviscapine and 2',3,5,6',7-pentahydroxyflavanone have the potential to inhibit all the other isoforms. Non-inhibitory action of the compounds denotes the potential of being transformed, thereby preventing toxic effects associated with high accumulation, while inhibition of the CYP450 isoenzymes denotes that the compounds can be metabolized slowly and could accumulate at high concentration in the plasma. Also, the compounds were not substrates for P-glycoprotein (P-gp) except 2',5,6',7-tetrahydroxyflavane, and breviscapine, which indicates that they can reach their site of action without being effluxed by P-gp, a transmembrane ATP-binding transporter responsible for transporting drugs out of the cell [28]**.** Skin permeability (Logkp) is a key physicochemical property used to evaluate drugs that require transdermal administration. The skin permeation results showed that all the screened compounds from *S. baicalensis* are likely to be skin permeants, having logkp values < -2. 5. According to Naspiah et al. [57], compounds with high skin permeation have Logkp < -2.5 while those with a value of > -2.5 have low skin permeability.

The toxicity prediction outcome revealed that 9 of the screened ligands have an LD_{50} ≥ 2000 mg/kg and belong to the safe toxicity class (class 5), which shows little toxic effect by ingestion. Human hepatotoxicity predicts the extent of the drug target to cause harm or damage to the hepatocyte, which could lead to organ failure or eventually death. The mutagenicity test (Ames test) evaluates the possibility of alteration of DNA or protein, while cytogenicity measures interaction with immune cells [58]. Interestingly, all the compounds are not likely to be hepatotoxic, mutagenic, or cytogenic, making them safe drug targets for brain cancer.

4. Materials and Methods

4.1. Protein preparation

The protein, (EGFR) tyrosine kinase, was prepared as previously described [59–60]. The crystal structure of the protein (1M17), which was obtained from the Protein Data Bank (PDB) repository, was prepared in Glide (Schrödinger Suite 2021-2) using the protein preparation wizard panel. During the process, hydrogen was added, bond orders were assigned, disulfide bonds were created, and the missing side chains and loops were replaced with prime. Water molecules outside 3.0 Å of the heteroatoms were eliminated, and the protein structure was minimized and optimized using OPLS4 and PROPKA respectively.

4.2. Generation of receptor grid

The receptor grid was developed to define the location and size of the active site of the protein for ligand docking. This was accomplished with Schrödinger Maestro 12.8's receptor grid generation tool. The active site of the protein's co-crystalized ligand (Erlotinib) was employed as the scoring grid [59–60].

4.3. Ligand preparation

Three hundred and seven (307) bioactive compounds of *S. baicalensis* were retrieved from Dr. Duke's Phytochemical and Ethnobotanical Databases, and they were prepared with the standard ligand (Erlotinib) using the Ligprep panel of Maestro 12.8, Schrödinger Suite 2021-2, as previously reported [28]. Low-energy 3D structures with acceptable chiralities were generated. Each ligand's ionization state was formed at a physiological pH of 7.2 ± 0.2. Stereoisomers of each ligand were computed by keeping certain chiralities constant while varying others.

 doi:10.20944/preprints202209.0218.v1

4.4. Protein-Ligand Docking

The molecular docking analysis was performed on Schrödinger Suite 2021-2 utilizing the Glide-Ligand Docking panel of Maestro 12.8. The prepared ligands and the receptor grid file were loaded into Maestro's work space, and the ligands were docked into the protein's binding site. The vdW radius scaling factor was set to 0.80 for ligand atoms, with a partial charge cut-off of 0.15, and the flexible ligand sampling option was employed [17,28].

4.5. Binding free energy calculation

The Schrödinger Suite's MM-GBSA Prime panel was used to calculate the binding free energy of the receptor-ligand complex. The ligands were first prepared with ligprep, and the proteins were prepared with the protein preparation wizard, as previously described [60]. The active sites of the proteins were identified using Sitemap, and the compounds were docked using glide extra precision (XP) docking. The OPLS3 force field was used, with the VSGB continuum solvent model. For the other choices, the default settings were used.

4.6. Receptor-ligand complex pharmacophore modelling

As previously described, the PHASE module of the Schrödinger Suite 2021-2 was utilized to build an auto/e-pharmacophore model [26]. The process was configured to generate a maximum of 7 features with a minimum feature-feature distance of 2.00 and a minimum distance of 4.00 between features of the same type.

4.7. Induced fit Docking

The molecular interaction of ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, chrysosplenetin, tenaxin I and (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol with receptor EGFR tyrosine kinase was investigated using Maestro 12.8's induced fit docking panel and the approach provided in Schrödinger 2015. The process commenced with a constrained receptor minimization and then uses a softening potential to achieve the first glide docking of the ligand. Twenty sets of docked poses were sent to Prime for further refinement. After prime side-chain prediction and minimization, the best receptor structures for each ligand were returned to Glide for redocking. According to the extended sampling protocol, residues were selected for trimming and atom-specific van der Waals scaling factors were calculated based on solvent-accessible surface areas, B-factors, salt bridges, and rotamer searches. The initial docking stage produces a large number of poses, which are subsequently combined and filtered to produce up to 80 poses per ligand, which are then passed on to the prime step. The final docked poses were scored using the Glide SP.

4.8. Gene-set enrichment analysis

An effective method for connecting a disease phenotype to a collection of genes and/or proteins is gene-set enrichment analysis. STRING database version 11.5 was used to decipher the network interaction of the genes and proteins associated with the disease of interest. The target protein sequence EGFR tyrosine kinase from *homo sapiens* were queried on STRING. The network type was set to cover both functional and physical protein associations at the highest confidence level of 0.900, the maximum number of interactors to show at ≤ 10. The resulting networks of interest were extracted as a high-resolution bitmap image and tab-delimited files respectively [61].

4.9. Molecular dynamics simulation

Molecular dynamics simulations were carried out using the Simulation module of the Molecular Operating Environment MOE 2019.01 software (Chemical Computing Group Inc., 2016). The protein and protein-ligand complex were protonated, energy was minimized, and parameterized using the AMBER 10: EHT force field at intervals to get

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 September 2022 doi:10.20944/preprints202209.0218.v1

the stable conformer for EGFR tyrosine kinase-ligand complex in an R-Field implicit solvation system. The simulations were run in three stages. To begin, the molecular system was heated to 310 K (37 °C). This was followed by a 100-picosecond equilibration step at 310 K (37 °C). The trajectory of the molecular system was then constructed for 1000 picoseconds at 310 K using the Nose-Poincare-Andersen (NPA) method (the time step of each simulation was set to 0.02 picoseconds). The VMD software and the Galaxy Europe platform Bio3D platform were used for visualization and data processing. The principal components analysis (PCA) was used to model the system's essential dynamics [62]. The simulation data set was scaled down to a few essential components that define the directions to the largest variance. The key structural variants within the protein structure ensemble were captured by ranking the principal components as eigenvectors based on the variance. The fraction of variance attributed to each principal component was displayed using an eigenvalue rank plot. Following that, structural clustering based on the resulting principal components was performed, as well as residue-wise loadings to determine how much each residue contributed to the first two principal components.

4.10. ADMET Profiling

In silico predictive models were used to determine the ADMET properties of the selected test compounds. The SwissADME server was used to assess the ADME properties of the compounds, which include: lipophilicity (Log P), water solubility (ESOL Log S), drug-likeness, and pharmacokinetic properties [52]. The ProTox-II online server was used to determine the compounds' acute toxicity class, LD_{50}, and toxicological potentials [58].

5. Conclusions

In conclusion, of the 307 bioactive compounds of *S. baicalensis* screened for possible anti-GBM activity ganhuangenin, 5,7,2',5'-tetrahydroxyflavone, (2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, and tenaxin I possess higher binding affinities and they interacted with amino acids of clinical importance such as MET 769, GLU 738, THR 766. Two aromatic rings, H-bond donors and acceptors and some hydrophobic interactions were involved in the molecular interactions of the compounds of tested with the EGFR tyrosine kinase. Better docking scores in the induced-fit docking signifies that the compounds might have better inhibitory activities at different conformations. The top-scoring ligands of the plant indicated no Lipinski violation and also obeyed other drug-likeness rules by Ghose, Veber, Egan, and Muegge with the exception of breviscapine. Interestingly, all the compounds are likely not hepatotoxic, mutagenic, or cytogenic, making them safe drug targets for brain cancer. The findings of this study have shown that the plant *S. baicalensis* contain potential leads for glioblastoma multiforme inhibition and thus, should be studied further for development as effective therapeutic agents against brain cancer.

Supplementary Materials: Table S1. High throughput virtual screening of the constituents of *S. baicalensis* against EGFR. **Table S2.** Standard Precision Docking scores of the constituents of *S. baicalensis* against EGFR tyrosine kinase. **Table S3.** Binding free energy calculation of hit compounds with EGFR tyrosine kinase. **Table S4.** Induced Fit Docking scores and biological interactions of five top scoring compounds of *S. baicalensis* against EGFR tyrosine kinase. **Table S5.** KEGG pathways of identified genes. **Figure S1.** 2D view of the molecular interactions of A=Ganhuangenin, B=5,7,2',5'-Tetrahydroxyflavone, C=(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, D=Tenaxin I, E=Chrysosplenetin, F=Scutellarin, G=5,2',5'-trihydroxy-7,8-dimethoxyflavone, H=5,7,2',6'-Tetrahydroxyflavone, I=Viscidulin II, J=(2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one, K=Erlotinib with EGFR tyrosine kinase. **Figure S2.** 3D view of the molecular interactions of A=Ganhuangenin, B=5,7,2',5'-Tetrahydroxyflavone, C=(2R)-2-(2,6-dihydroxyphenyl)-3,4-dihydro-2H-chromene-5,7-diol, D=Tenaxin I, E=Chrysosplenetin, F=Scutellarin, G=5,2',5'-trihydroxy-7,8-dimethoxyflavone, H=5,7,2',6'-Tetrahydroxyflavone, I=Viscidulin II, J=(2S,3R)-2-(2,6-dihydroxyphenyl)-3,5,7-trihydroxy-2,3-dihydrochromen-4-one, K=Erlotinib with EGFR tyrosine kinase.

Author Contributions: Conceptualization, A.A.E., J.T.O., and Y.A.J.; methodology and investigation, Y.A.J., A.A.E., J.T.O., Y.A.H., J.G.I., A.R.O., A.M.N., C.I.F., U.C.J., O.O.S, A.A.F., A.W.A.I, D.Q.M., B.G.E.S.; A.A.F., A.W.A.I, D.W.M.; writing—original draft preparation, Y.A.J.; writing—review and editing, All authors; supervision, J.T.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding and The APC was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R39), and Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Researchers Supporting Project number (RSP-2021/97), King Saud

Data Availability Statement: The original contributions presented in the study are included in the article, and further inquiries can be directed to the corresponding author(s).

Acknowledgments: We acknowledged the efforts of Jaris Computational Biology Center, Jos, Nigeria for providing research softwares and facility.

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

Sample Availability: Samples of the compounds are not available from the authors.

References

1. Tandel, G.S.; Biswas, M.; Kakde, O.G.; Tiwari, A.; Suri, H.S.; Turk, M., Laird, J.R.; Asare, C.K., Ankrah, A.A.; Khanna, N.N.; Madhusudhan, B.K. A review on a deep learning perspective in brain cancer classification. *Cancers*. **2019**, *11*, 111. doi: 10.3390/cancers11010111
2. Mohammadi, E.; Ghasemi, E.; Azadnajafabad, S.; Rezaei, N.; Saeedi, M.S.; Ebrahimi, M.S.; Fattahi, N., Habibi, Z.; Karimi, Y.K.; Amirjamshidi, A.; Nejat, F. A global, regional, and national survey on burden and Quality of Care Index (QCI) of brain and other central nervous system cancers; global burden of disease systematic analysis 1990-2017. *PLoS One*. **2021**, *16*: e0247120. DOI: 10.1371/journal.pone.0247120
3. Miranda-Filho, A.; Piñeros, M.; Soerjomataram, I.; Deltour, I.; Bray, F. Cancers of the brain and CNS: global patterns and trends in incidence. Neuro-oncology. **2017**, *19*, 270-80. DOI: 10.1093/neuonc/now166.
4. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer*. **2015,** *136*, E359-86. https://doi.org/10.1002/ijc.29210.
5. Patel, A.P.; Fisher, J.L.; Nichols, E.; Abd-Allah, F.; Abdela, J.; Abdelalim, A.; Abraha, H.N.; Agius, D.; Alahdab, F.; Alam, T.; Allen, C.A. Global, regional, and national burden of brain and other CNS cancer, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. **2019**, *18*, 376-93. DOI: 10.1016/S1474-4422(18)30468-X
6. Hassanipour, S.; Namvar, G.; Fathalipour, M.; Ghorbani, M.; Abdzadeh, E.; Zafarshamspour, S.; Riahi, S.; Mohammadian-Hafshejani, A.; Salehiniya, H. The incidence of brain tumors in Iran: a systematic review and meta-analysis. *Advances Hum. Biol.* **2019,** *9*, 2-7. DOI: 10.4103/AIHB.AIHB_60_18.
7. McNeill, K.A. Epidemiology of brain tumors. *Neurol Clin*. **2016**, *34*, 981-98. DOI: 10.1016/j.ncl.2016.06.014
8. Farmanfarma, K.K.; Mohammadian, M.; Shahabinia, Z.; Hassanipour, S.; Salehiniya, H. Brain cancer in the world: an epidemiological review. *World J. Cancer Res*. **2019,** *5*, e1356. http://bsid.bums.ac.ir/dspace/handle/bums/6594
9. Louis, D.N.; Ohgaki, H.; Wiestler, O.D.; Cavenee, W.K.; Burger, P.C.; Jouvet, A.; Scheithauer, B.W.; Kleihues, P. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol*. **2007**, *114*, 97-109. doi: 10.1007/s00401-007-0243-4.
10. Sneed, P.K.; Suh, J.H.; Goetsch, S.J.; Sanghavi, S.N.; Chappell, R.; Buatti, J.M.; Regine, W.F.; Weltman, E.; King, V.J.; Breneman, J.C.; Sperduto, P.W. A multi-institutional review of radiosurgery alone vs. radiosurgery with whole brain radiotherapy as the initial management of brain metastases. *Int. J. Radiat. Oncol. Biol. Phys.*, **2002**, *53*, 519-26. DOI: 10.1016/s0360-3016(02)02770-0.

11. Hegi, M.E.; Murat, A.; Lambiv, W.L.; Stupp, R. Brain tumors: molecular biology and targeted therapies. *Ann. Oncol*. **2006**, 191-7. DOI: 10.1093/annonc/mdl259.
12. Sharma, A.; Saggu, S.K.; Mishra, R.; Kaur, G. Anti-brain cancer activity of chloroform and hexane extracts of *Tinospora cordifolia* Miers: an in vitro perspective. *Ann. Neurosci*. **2019**, *26*, 10-20. doi: 10.5214/ans.0972.7531.260104.
13. Shang, X.; He, X.; He, X.; Li, M.; Zhang, R.; Fan, P.; Zhang, Q.; Jia Z. The genus *Scutellaria* an ethnopharmacological and phytochemical review. *J Ethnopharmacol*. **2010**, *128*, 279-313. DOI: 10.1016/j.jep.2010.01.006.
14. Zhao, Q.; Chen, X.Y.; Martin, C. *Scutellaria baicalensis*, the golden herb from the garden of Chinese medicinal plants. *Sci. Bull*. **2016**, *61*, 1391-8. doi: 10.1007/s11434-016-1136-5.
15. Scheck, A.C.; Perry, K.; Hank, N.C.; Clark, W.D. Anticancer activity of extracts derived from the mature roots of *Scutellaria baicalensis* on human malignant brain tumor cells. *BMC Complement Altern. Med*. **2006**, *6*, 1-9. doi: 10.1186/1472-6882-6-27.
16. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod.*, **2020,** *83*, 770-803. https://doi.org/10.1021/acs.jnatprod.9b01285.
17. Johnson, T.O.; Adegboyega, A.E.; Adeleke, O.O.; Yusuf, A.J.; Opeyemi, I.; Ugwah-Oguejiofor, C.J.; Onyekachukwu, A.R.; Chukwuma, I.F.; Adakole, E.S. A Computational approach to elucidate the interactions of chemicals from *Artemisia annua* targeted toward SARS-CoV-2 main protease inhibition for COVID-19 treatment. *Front. Med.*, **2022,** *9*, 907583. DOI: 10.3389/fmed.2022.907583.
18. Desai, V.; Bhushan A. Natural Bioactive Compounds: Alternative Approach to the Treatment of Glioblastoma Multiforme. *BioMed Res. Int*., **2017,** ID 9363040, 1 – 10. https://doi.org/10.1155/2017/9363040.
19. Malik, A.; Manan, A.; Mirza, M.U. Molecular docking and in silico ADMET studies of silibinin and glycyrrhetic acid anti-inflammatory activity. *Trop. J. of Pharm. Res*., **2017,** *16*, 67–74. DOI: 10.4314/tjpr.v16i1.9
20. Singh, P.; Bast, F. In silico molecular docking study of natural compounds on wild and mutated epidermal growth factor receptor. *Med Chem Res* **2014,** 23, 5074–5085. http://dx.doi.org/10.1007%2Fs00044-014-1090-1
21. Bathini, R.; Sivan, S.K.; Fatima, S. *et al.* Molecular docking, MM/GBSA and 3D-QSAR studies on EGFR inhibitors. *J. Chem. Sci.* **2006,** *128*, 1163–1173. https://doi.org/10.1007/s12039-016-1103-3.
22. Schaub, A.J.; Moreno, G.O.; Zhao, S.; Truong, H.V.; Luo, R.; Tsai, S.C. Computational structural enzymology methodologies for the study and engineering of fatty acid synthases, polyketide synthases and nonribosomal peptide synthetases. *Methods Enzymol.*, **2019,** *622*, 375-409. DOI: 10.1016/bs.mie.2019.03.001.
23. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z., & Hou, T. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem. Rev.*, **2019,** *119*, 9478-9508. DOI: 10.1021/acs.chemrev.9b00055.
24. Forouzesh, N.; Mishra, N. An effective MM/GBSA protocol for absolute binding free energy calculations: A case study on SARS-CoV-2 spike protein and the human ACE2 receptor. *Molecules*, **2021,** *26*, 2383. DOI: 10.3390/molecules26082383.
25. Ojo, O.A.; Ojo, A.B.; Okolie, C.; Abdurrahman, J.; Barnabas, M.; Evbuomwan, I.O., Batiha, G. E. S. Elucidating the interactions of compounds identified from *Aframomum melegueta* seeds as promising candidates for the management of diabetes mellitus: A computational approach. *Inform. Med. Unlocked*, **2021,** *26*, 100720. https://doi.org/10.1016/j.imu.2021.100720.
26. Johnson, T.O.; Abolaji, A.O.; Omale, S.; Longdet, I.Y.; Kutshik, R.J.; Oyetayo, B.O.; Sagay, A. Benzo[a]pyrene and Benzo[a]pyrene-7,8-dihydrodiol-9,10-epoxide induced locomotor and reproductive senescence and altered biochemical parameters of oxidative damage in Canton-S *Drosophila melanogaster*. *Toxicol. Rep.*, **2021,** *8*, 571–580. DOI: 10.1016/j.toxrep.2021.03.001.
27. Van Roosbroeck, K., Calin, G.A. Cancer Hallmarks and MicroRNAs: The Therapeutic Connection. *Adv Cancer Res*. **2017**; *135*, 119-149. DOI: 10.1016/bs.acr.2017.06.002.
28. Ojo, O.A.; Ojo, A.B.; Okolie, C.; Nwakama, M.C.; Iyobhebhe, M.; Evbuomwan, I.O.; Nwonuma, C.O.; Maimako, R.F.; Adegboyega, A.E.; Taiwo, O.A.; Alsharif, K.F.; Batiha, G.E. Deciphering the Interactions of Bioactive Compounds in Selected

Traditional Medicinal Plants against Alzheimer's Diseases via Pharmacophore Modeling, Auto-QSAR, and Molecular Docking Approaches. *Molecules*, **2021,** *26,* 1996. DOI: 10.3390/molecules26071996.

29. Lanzarotti, E.; Defelipe, L.A.; Marti, M.A.; Turjanski, A.G. Aromatic clusters in proteinprotein and protein-drug complexes. *J. Cheminf.*, **2020**, *12*, 1–9. https://doi.org/10.1186/s13321-020-00437-4.
30. Zheleznova, N.N.; Wilson, P.D.; Staruschenko, A. Epidermal growth factor-mediated proliferation and sodium transport in normal and PKD epithelial cells. *Biochim Biophys Acta Mol Basis Dis BBA-MOL BASIS DIS,* **2011**, *1812*, 1301–1313. https://doi.org/10.1016/j.bbadis.2010.10.004.
31. Hatanpaa, K.J.; Burma, S.; Zhao, D.; Habib, A.A. Epidermal Growth Factor Receptor in Glioma: Signal Transduction, Neuropathology, Imaging, and Radioresistance. *Neoplasia*, **2010,** *12*, 675–684. DOI: 10.1593/neo.10688.
32. Zahonero, C.; Sánchez-Gómez, P. EGFR-dependent mechanisms in glioblastoma: Towards a better therapeutic strategy. *Cell. Mol. Life Sci*, **2014**, *71*, 3465–3488. DOI: 10.1007/s00018-014-1608-1.
33. Liu, F.; Mischel, P. S. Targeting epidermal growth factor receptor co-dependent signaling pathways in glioblastoma. Wiley Interdiscip Rev Syst Biol Med, **2018**, *10*, e1398. DOI: 10.1002/wsbm.1398.
34. Saadeh, F. S.; Mahfouz, R.; Assi, H. I. EGFR as a clinical marker in glioblastomas and other gliomas. *Int. J. Biol. Markers*, **2018**, *33*, 22–32. DOI: 10.5301/ijbm.5000301.
35. Oprita, A.; Baloi, S.C.; Staicu, G.A.; Alexandru, O.; Tache, D.E.; Danoiu, S.; Micu, E.S.; Sevastre, A.S. Updated insights on EGFR signaling pathways in glioma. *Int. J. Mol. Sci*., **2021**, *22*, 1–21. DOI: 10.3390/ijms22020587.
36. Feng, P.; Che, Y.; Chen, D.Q. Molecular mechanism of action of Liuwei Dihuang pill for the treatment of osteoporosis based on network pharmacology and molecular docking. *Eur. J.Integr. Med.*, **2020,** *33*, 101009. https://doi.org/10.1016/j.eujim.2019.101009.
37. Li, X.; Lau, A.Y.T.; Ng, A.S.N.; Aldehaiman, A.; Zhou, Y.; Ng, P.K.S.; Arold, S.T.; Cheung, L.W.T. (2021). Cancer-associated mutations in the p85α N-terminal SH2 domain activate a spectrum of receptor tyrosine kinases. Proceedings of the National Academy of Sciences, **2021,** *118*, e2101751118. DOI: 10.1073/pnas.2101751118.
38. Ibrahim, R.S.; El-Banna, A.A. Network pharmacology-based analysis for unraveling potential cancer-related molecular targets of Egyptian propolis phytoconstituents accompanied with molecular docking and *in vitro* studies. *RSC Adv*., **2021**, *11*, 11610-11626. doi: 10.1039/d1ra01390d. DOI: 10.1039/d1ra01390d.
39. Li, N.; Liu, K.; Yu, M.; Liu, M.; Li, S.; Cai, W.; Tian, A. Systematic elucidation of the traditional Chinese medicine prescription Danxiong particles via network pharmacology and molecular docking. *Trop. J. Pharm. Res*., **2022**, *21*, 981–987. DOI: 10.4314/tjpr.v21i5.11.
40. Ahrens, T.D.; Bang-Christensen, S.R.; Jørgensen, A.M.; Løppke, C.; Spliid, C.B.; Sand, N.T.; Clausen, T.M.; Salanti, A.; Agerbæk, M.Ø. The Role of Proteoglycans in Cancer Metastasis and Circulating Tumor Cell Analysis. *Front Cell Dev Biol*., **2020**, *26*, 749. doi: 10.3389/fcell.2020.00749.
41. Yang, Z.; Lu, S.; Tang, H.; Qu, J.; Wang, B.; Wang, Y.; Pan, G.; Rao, B. Molecular Targets and Mechanisms of *Hedyotis diffusa* - *Scutellaria barbata* Herb Pair for the Treatment of Colorectal Cancer Based on Network Pharmacology and Molecular Docking. *Evid.-based Complement. Altern. Med,* **2022**, *2022*, 1–15. doi: 10.1155/2022/6186662.
42. Guo, Y.; Pan, W.; Liu, S.; Shen, Z.; Xu, Y.; Hu, L. ERK/MAPK signalling pathway and tumorigenesis (Review). *Exp. Ther. Med*., **2020,** *19*, 1997–2007. DOI: 10.3892/etm.2020.8454.
43. Hou, F.; Yu, Z.; Cheng, Y.; Liu, Y.; Liang, S.; Zhang, F. Deciphering the pharmacological mechanisms of *Scutellaria baicalensis* Georgi on oral leukoplakia by combining network pharmacology, molecular docking and experimental evaluations. *Phytomedicine*, **2022,** *103*, 154195. https://doi.org/10.1016/j.phymed.2022.154195.
44. Park, H. J.; Park, S. H.; Choi, Y. H.; Chi, G. Y. The Root extract of *Scutellaria baicalensis* Induces Apoptosis in EGFR TKI-Resistant Human Lung Cancer Cells by Inactivation of STAT3. *Int. J. Mol. Sci*., **2021**, *22*, 5181. DOI: 10.3390/ijms22105181.

45. Park, S.-H.; Park, H.-J. Root extract of *Scutellaria baicalensis* increases Gefitinib Sensitivity in H1975 Human Non-small Cell Lung Cancer Cells. *Journal of Physiology & Pathology in Korean Medicine*, **2021**, *35*, 117–123. https://doi.org/10.15188/kjopp.2021.08.35.4.117.
46. Hu, L.C.; Ding, C.H.; Li, H.Y.; Li, Z.Z.; Chen, Y.; Li, L.P.; Li, W.Z.; Liu, W. S. Identification of potential target endoribonuclease NSP15 inhibitors of SARS-COV-2 from natural products through high-throughput virtual screening and molecular dynamics simulation. *J. Food Biochem*., **2022** *46*, e14085. DOI: 10.1111/jfbc.14085.
47. Sittel, F.; Jain, A.; Stock, G. Principal component analysis of molecular dynamics: On the use of Cartesian vs. internal coordinates. *J. Chem. Phys*., **2014**, *141*, 014111. DOI: 10.1063/1.4885338.
48. Oyinloye, E.B.; Iwaloye, O.; Ajiboye, O.B. Polypharmacology of *Gongronema latifolium* leaf secondary metabolites against protein kinases implicated in Parkinson's disease and Alzheimer's disease. *Sci. African*, **2021,** *12*, e00826. https://doi.org/10.1016/j.sciaf.2021.e00826.
49. Apeh, V.O.; Njoku, O.U.; Nwodo, F.O.C. et al. In silico drug-like properties prediction and in vivo antifungal potentials of *Citrullus lanatus* seed oil against *Candida albicans*. *Arab. J. Chem.*, **2022,** *15*, 103578. https://doi.org/10.1016/j.arabjc.2021.103578.
50. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev*., **2012,** *64*, 4–17. DOI: 10.1016/s0169-409x(00)00129-0.
51. Kikiowo, B.; Ogunleye, J.A.; Iwaloye, O.; Ijatuyi, T.T. Therapeutic potential of *Chromolaena odorata* phyto-constituents against human pancreatic α-amylase. *J Biomol Struct Dyn*, **2022,** *40*, 1801 – 1812. DOI: 10.1080/07391102.2020.1833758.
52. Daina, A., Zoete, V. A BOILED-Egg to predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem*, **2016,** 1117–1121. DOI: 10.1002/cmdc.201600182.
53. Harley, B.K.; Amponsah, I.K.; Ben, I.O.; Adongo, D.W.; Mireku-Gyimah, N.A.; Baah, M.K.; Mensah, A.Y.; Fleischer, T.C. *Myrianthus libericus*: Possible mechanisms of hypoglycaemic action and in silico prediction of pharmacokinetics and toxicity profile of its bioactive metabolite, friedelan-3-one. *Biomed Pharmacother*. **2021**, *137*, 111379. DOI:10.1016/j.biopha.2021.111379.
54. Zhang, Z.; Tang, W. Drug metabolism in drug discovery and development. *Acta Pharm. Sin. B*, **2018,** *8*, 721-732. DOI: 10.1016/j.apsb.2018.04.003.
55. Patel, R.; Barker, J.; Elshaer, A. Pharmaceutical Excipients and Drug Metabolism : A Mini-Review. *Int J Mol Sci*, **2020,** *21*, 8224. DOI: 10.3390/ijms21218224.
56. Da Rocha, M.N.; Marinho, E.S.; Marinho, M.M.; Dos Santos, H.S. Virtual screening in pharmacokinetics, bioactivity, and toxicity of the *Amburana cearensis* secondary metabolites. *Biointerface Res Appl Chem*., **2022,** *12*, 8471–8491. https://doi.org/10.33263/BRIAC126.84718491.
57. Naspiah, N.; Pratama, M.R. Xanthine Oxidase Inhibition activity and ADMET properties of Terap (*Artocarpus odoratissimus* Blanco) leaves metabolites: Phytochemical Screening and in silico Studies. *Pharmacogn. J*., **2021**, *13*, 1150-1160. DOI:10.5530/pj.2021.13.148.
58. Banerjee, P.; Eckert, A.O.; Schrey, A.K.; Preissner, R. ProTox-II: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Res., **2018,** *46*, W257–W263. doi: 10.1093/nar/gky318.
59. Olsson, M.H.; Søndergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: Consistent treatment of internal and surface residues in empirical pK(a) predictions. *J. Chem. Theory Comput*., **2011,** *7*, 525-537. DOI: 10.1021/ct100578z.
60. Iwaloye, O.; Elekofehinti, O.O.; Oluwarotimi, E.A.; Kikiowo, B.I.; Fadipe, T.M. Insight into glycogen synthase kinase-3β inhibitory activity of phyto-constituents from Melissa officinalis: in silico studies. *In Silico Pharmacol*. **2020**, *8*, 2. DOI: 10.1007/s40203-020-00054-x.
61. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery

in genome-wide experimental datasets. Nucleic Acids Res. 2019 Jan 8;47(D1): D607-D613. doi: 10.1093/nar/gky1131. PMID: 30476243; PMCID: PMC6323986.

62. Bray, S.A., Senapathi, T., Barnett, C.B. *et al.* Intuitive, reproducible high-throughput molecular dynamics in Galaxy: a tutorial. *J. Cheminform.*, **2020,** *12*, 54. https://doi.org/10.1186/s13321-020-00451-6.