

Article

Pharmacoinformatics Analysis Reveals Flavonoids and Diterpenoids from *Andrographis paniculata* and *Thespesia populnea* to Target Hepatocellular Carcinoma Induced by Hepatitis B Virus

Vishal S. Patil ^{1,2} , Darasaguppe R. Harish ^{1,*} , Umashankar Vetrivel ^{1,3} , Sanjay H. Deshpande ^{1,4} , Pukar Khanal ⁵ , Harsha V. Hegde ¹, Subarna Roy ^{1,*}  and Sunil S. Jalalpure ^{2,*}

¹ ICMR-National Institute of Traditional Medicine, Nehru Nagar, Belagavi 590010, India

² KLE College of Pharmacy, Belagavi, KLE Academy of Higher Education and Research, Nehru Nagar, Belagavi 590010, India

³ ICMR-National Institute for Research in Tuberculosis, Chetpet, Chennai 600031, India

⁴ Regional Centre for Biotechnology, NCR-Biotech Science Cluster, Faridabad 121001, India

⁵ Department of Pharmacology, NGSM Institute of Pharmaceutical Sciences, NITTE Deemed-to-be University, Mangalore 575018, India

* Correspondence: harish.dr@icmr.gov.in (D.R.H.); roys@icmr.gov.in (S.R.); jalalpuresunil@rediffmail.com (S.S.J.)



Citation: Patil, V.S.; Harish, D.R.; Vetrivel, U.; Deshpande, S.H.; Khanal, P.; Hegde, H.V.; Roy, S.; Jalalpure, S.S. Pharmacoinformatics Analysis Reveals Flavonoids and Diterpenoids from *Andrographis paniculata* and *Thespesia populnea* to Target Hepatocellular Carcinoma Induced by Hepatitis B Virus. *Appl. Sci.* **2022**, *12*, 10691. <https://doi.org/10.3390/app122110691>

Academic Editors: Snezana Agatonovic-Kustrin, Antonija Tomić and Dejan Agić

Received: 19 September 2022

Accepted: 19 October 2022

Published: 22 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Herbs are widely utilized in the Western Ghats region of India to treat liver diseases and viral-like infections. However, such practices lack scientific evidence at the molecular level and may often pose adverse drug reactions. This study intends to identify phytochemicals with druggability and non-toxic profiles with potential activity against hepatitis B virus-induced hepatocellular carcinoma. The details of phytochemicals in traditionally utilized herbs in the Western Ghats region were collated from chemical databases and publications. The druggability and toxicity of these compounds were predicted using MolSoft and ADVERpred, respectively. The probable targets of these phytochemicals were predicted using BindingDB. Moreover, compound-gene set pathways, cellular processes, and functional enrichment analyses were also performed using STRING and KEGG pathway databases. Subsequently, herb-compound-target-disease pathway networks were constructed using Cytoscape. The potential hub protein was virtually screened against the ligand dataset using the POAP pipeline. Finally, molecular dynamics (MD) simulations of the most potential protein-ligand complexes were performed in triplicate using Schrödinger Desmond. Amongst 274 compounds from 16 herbs studied, 36 showed drug-likeness with nontoxic properties and were also predicted to modulate 16 potential targets involved in the pathogenesis of HBV-induced HCC. Among all the molecules screened, flavonoids and diterpenoids from *Andrographis paniculata* and *Thespesia populnea* scored the highest edge count via modulating multiple targets and pathways. Moreover, molecular docking and MD simulation (100ns) also inferred that the top-ranking Andrographin and Gossypetin exhibit stable intermolecular interactions with EGFR protein, which was identified as a highly connected hub protein in the constructed network. All these findings are suggestive of identified moieties as potential therapeutics for targeting HBV-associated HCC sans adverse drug reactions.

Keywords: *Andrographis paniculata*; hepatitis B; hepatocellular carcinoma; molecular docking; molecular dynamics; network pharmacology; *Thespesia populnea*

1. Introduction

Hepatitis B, caused by the hepatitis B virus (HBV), is a life-threatening chronic liver infection and is one of the major risk factors for the development of hepatocellular carcinoma (HCC), which accounts for 70–90% of all primary liver cancers. HCC is the sixth

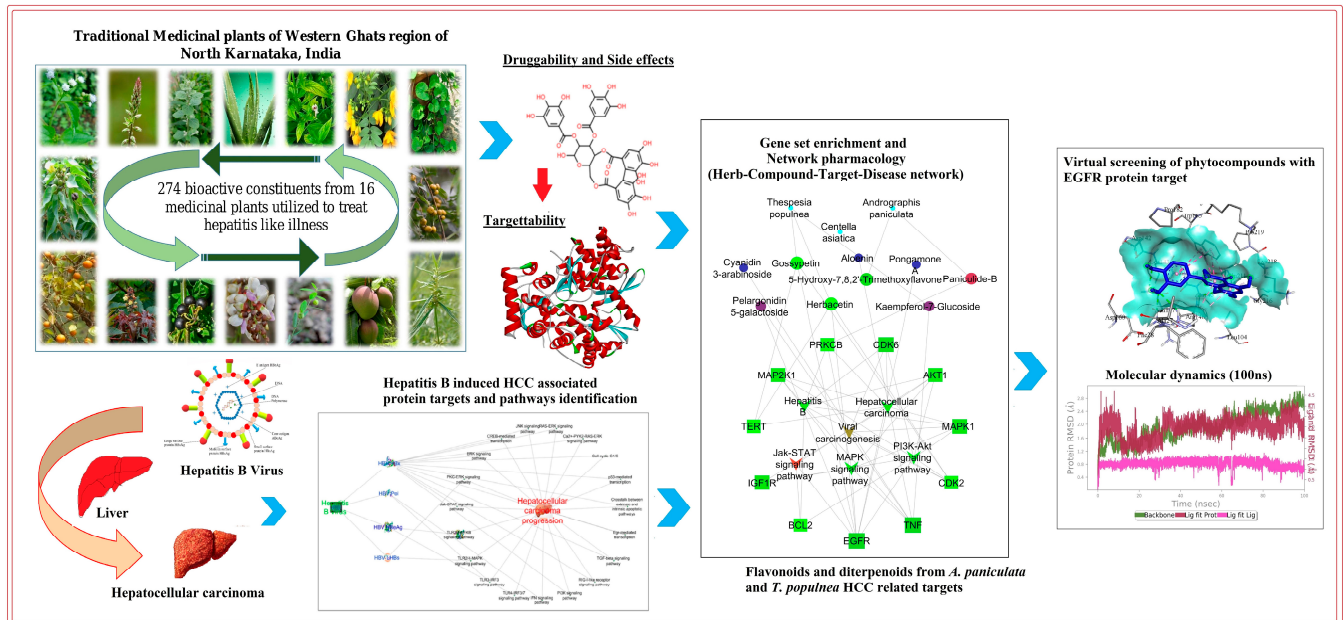
most common cancer worldwide with a significant mortality rate and is projected to increase by 40% by 2030 [1,2]. Currently, more than 248 million individuals worldwide are living with chronic HBV infection. Africa has the most elevated endemicity with 8.83% of HBsAg prevalence. About 95 million people in China have chronic HBV infection [3]. India is meso-endemic for HBV and is the second leading country after China to have about 50 million individuals with chronic HBV infection with an overall prevalence rate of 2.4% (2.2–2.7%) and a much higher rate of 15.9% (11.4–20.4%) among the tribal population [4].

Multiple studies, such as the accumulation of genetic damage due to immune-mediated hepatic inflammation and chronic hepatic oxidative stress, established a robust correlation between chronic HBV infection and the progression to HCC [5–7]. Moreover, HCC mediated by HBV is a polygenic disease that involves multi-protein, multi-pathway pathogenesis. HBV viral proteins alter the expression of endogenous genes within the host tissue. In the hepatocytes, HBV-mediated cytokine production by pro-inflammatory reactions and T-cell dysfunction as well as inflammation-mediated up- and downregulation of signaling pathways enhance the development of viral carcinogenesis and ultimately lead to HCC. Further, Signal Transducer and Activator of Transcription 3 (STAT3) and nuclear factor (NF)- κ B play key roles in the regulation of communication between cancer cells and inflammatory cells, which further aids in the regulation of tumor angiogenesis. However, HBV-mediated activation of these two pathways promotes immune escapes, production of interleukin IL-6, and IL-22, and ultimately progresses to the development of HCC [8,9]. The HBV X Interacting Protein (HBXIP) is a transcriptional transactivator that controls the transcription of viral nucleic acid and cellular promotion, i.e., it mediates cell cycle regulatory proteins including p16, p21, p27, cyclin A, D1, and B1, promoting CDK2 activity [10–15]. HBx regulates HBV replication and plays a crucial role in the development of HCC. The correlation between HBV-mediated HCC is well established. However, the molecular mechanism is not well understood, because post HBV infection, multiple cascades are involved in HCC development. In addition, numerous scientific reports suggest that the major cause of HCC is due to continuous hepatocyte destruction and regeneration [16].

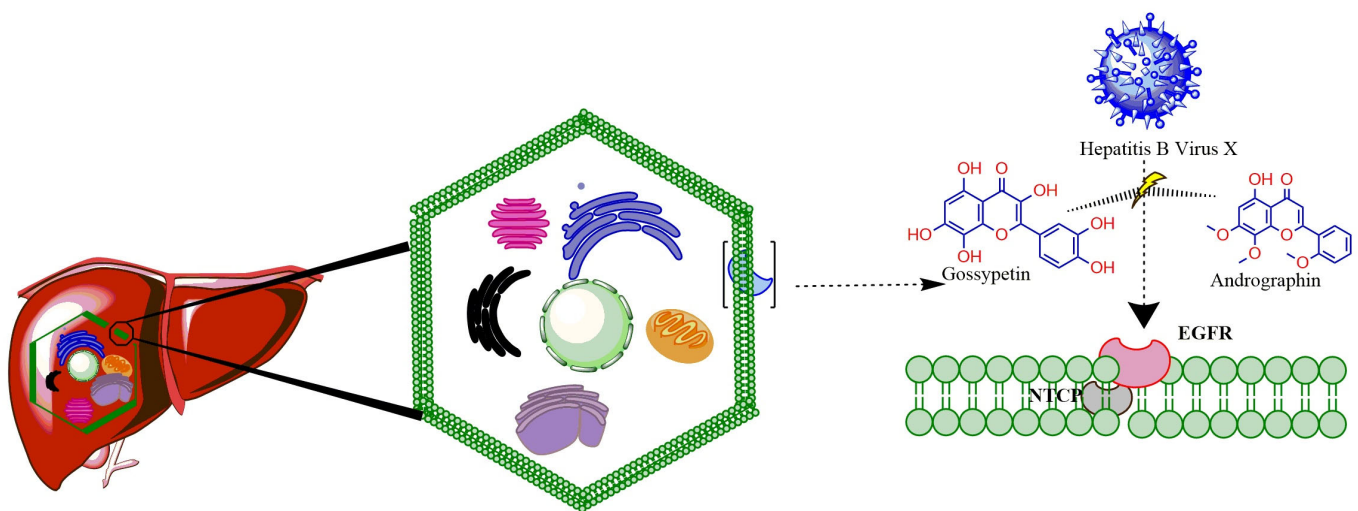
Since ancient times, herbs and herbal products have been widely utilized as therapeutics for various infectious diseases. Herbs are known to prevent viral infections, delay or arrest the progression or extend the latency period by acting as an agonist, antagonistic, or synergistic via modulating multi-proteins and multi-pathways associated with the disease pathogenesis [17]. Although the use of herbs against specific etiologies such as HBV infection and its associated HCC is not adequately documented, attempts are being continuously made to evaluate if traditionally used herbs could be used against these specific etiologies by a screening of its use against signs and symptoms resembling the disease. In the current study, we identified 16 herbs from the Western Ghats region that are traditionally utilized by the local traditional practitioners against conditions that represent viral infection, fever, inflammation, and jaundice, which are documented in earlier studies by the ICMR-National Institute of Traditional Medicine, Belagavi, India. These herbs include *Achyranthes aspera* L., *Aerva lanata* (L.) Juss. ex Schult., *Aloe vera* (L.) Burm. f., *Andrographis paniculata* (Burm.f.) Nees, *Cassia auriculata* (L.) Roxb., *Centella asiatica* (L.) Urban, *Diospyros Montana* Roxb., *Leucas aspera* (Willd.) Link, *Mangifera indica* L., *Phyllanthus virgatus* G. Forst., *Pongamia pinnata* (L.) Panigrahi, *Ricinus communis* L., *Solanum nigrum* L., *Strychnus nux-vomica* L., *Thespesia populnea* (L.) Sol. ex Corrêa, and *Vernonia cinerea* (Carl Linnaeus) H. Rob.

In the present study, we utilized the knowledge of polypharmacology and chemoinformatics approaches such as compound-gene set enrichment analysis, network pharmacology, molecular docking, and dynamics studies to identify potential classes (e.g., group of flavonoids, alkaloids, terpenoids, tannins, etc.) of bioactives from herbs with druggable characteristics, non-toxic properties, and potent beneficial effect in the management of HCC induced by HBV. Polypharmacology is the study based on the effect of a single compound on two or multiple protein targets and pathways, which involves a network-based regulation of numerous physiological responses [18]. Understanding the effect of

multi-component drug therapy using traditional knowledge provides new insight for the identification of potent therapeutics against complex diseases such as HCC. The workflow of polypharmacology and chemoinformatics approaches implemented in this study is illustrated in Figure 1A, and the findings are illustrated in Figure 1B.



(A)



(B)

Figure 1. (A). Workflow implemented in the present study. (B). Results highlight of the present study.

2. Materials and Methods

2.1. Selection of Plants and Retrieval of Phytocompounds

Sixteen medicinal plants utilized in the region of Western Ghats and North Karnataka for the treatment of fever, jaundice, inflammation, and viral infection were selected based on an online resource documented by ICMR-National Institute of Traditional Medicine (<http://nitmmmedplantsdb.in/>) (accessed on 6 March 2021) [19]. The selected medicinal plants with Ayurveda/common names and a list of references for their above-claimed uses are provided in Table S1. Reported phytocompounds from each plant were retrieved from Dr. Duke's DB (<https://phytochem.nal.usda.gov/phytochem/search>; accessed on 7 March 2021) [20], Phytochemical Interaction DB (<https://www.genome.jp/db/picdb/>)

(accessed on 7 March 2021), ChEBI (<https://www.ebi.ac.uk/chebi/>) (accessed on 7 March 2021) [21], Database of Ethnomedicinal Plants of Western Ghats (<http://nitmmedplantsdb.in/>) (accessed on 7 March 2021) [19] and other public repositories.

2.2. Druggability and Toxicity Profile

Molecular formula (MF), molecular weight (MW), hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), logP value, and canonical SMILES of each molecule were retrieved from the PubChem chemical database (<https://pubchem.ncbi.nlm.nih.gov/>) (accessed on 18 March 2021). The canonical smile data of individual compounds was queried in MolSoft online server (<https://molsoft.com/mprop/>; accessed on 15 March 2021) [22] to predict the druggability of each phytocompound based on the Lipinski rule of five. The server provides structural information, viz., MF, MW, HBA, HBD, logP, and predicts probable drug-likeness score (DLS). Further, phytocompounds with positive DLS were chosen, and their toxicity was predicted using the ADVERpred server [23]. We opted for this server, as it predicts hepatotoxicity, nephrotoxicity, arrhythmia, and myocardial infarction of unknown molecules based on the probable activity (Pa) and probable inactivity (Pi) and predicts probable side effects. Phytocompounds with a Pa value of ≤ 0.5 proceeded for further studies.

2.3. Target Identification

Canonical SMILES of the shortlisted phytocompounds were queried into BindingDB [24] for probable target protein identification ($p \geq 0.7$). The Gene IDs of the predicted protein targets were retrieved manually from the UniProt web server (<https://www.uniprot.org/>; accessed on 9 April 2021) [25]. Further, the reported targets involved in the pathogenesis of HBV-induced HCC were retrieved from the KEGG database (<https://www.genome.jp/kegg/>) (accessed on 9 April 2021) and public repositories.

2.4. Gene Ontology and Pathway Enrichment Analysis

To understand the molecular functions, processes of probable protein targets modulated by the phytocompounds, and to retrieve the protein–protein interactions systematically, a set of retrieved gene IDs was provided as input to STRING 11.0 (Search Tool for the Retrieval of Interacting Genes/Proteins (<https://string--db.org>) accessed on 9 April 2021) [26]. Gene ontology (GO)-based functional and process enrichment analyses were carried out using the STRING GO tool to retrieve the molecular functions and processes of protein targets. The KEGG pathway (<https://www.genome.jp/kegg/pathway.html>; accessed on 9 April 2021) [27] was utilized to identify the enriched pathways involved in HBV-induced HCC. In this study, to obtain significant functions, processes, and pathway enrichment, the default false discovery rate (FDR) p value cut-off of <0.05 was considered. The overall pathways involved in HBV-mediated HCC are shown in Figure 2.

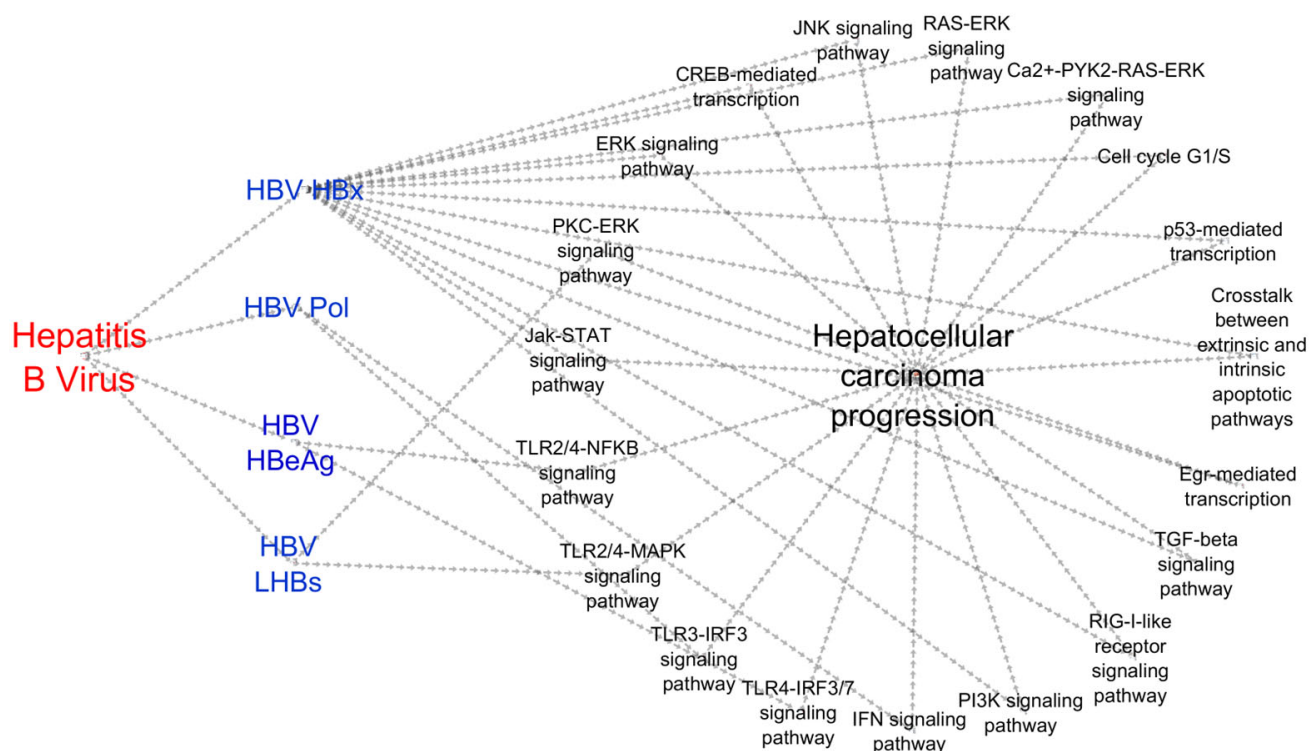


Figure 2. Network representation of pathways involved in HBV infection-induced HCC.

2.5. Network Construction

The compound–protein pathway network was constructed using Cytoscape v3.6.1 software (Washington, USA), an open-source software utilized to visualize biological pathways and complex molecular interaction networks [28]. During the construction of the network, the “Network Analyzer” command was used and treated as directed. To express the degree of connections among the compound–protein pathways, edge count calculation was applied. Further, “low values to small sizes” for node size and “low values to bright colors” for node color were applied to design the network [29].

2.6. Structure Refinement, Homology Modeling, and Assessment of Active Site Residues

Among the total targets predicted by bioactive phytocompounds, EGFR was found to score the highest edge count within the network (found to be involved in MAPK, PI3K–Akt, Ras, and Jak–STAT signaling pathways). Hence, EGFR, a highly connected target within the network and a known therapeutic target for hepatocellular carcinoma was chosen to virtually screen the bioactive intermolecular phytocompounds to infer significant intermolecular interactions at active site residues. The wild-type EGFR kinase domain in the complex with dacomitinib (PDB id: 4I23 (2.8 Å resolution) expressed in *Spodoptera frugiperda* expression system) was selected from the Protein Data Bank (PDB) (<https://www.rcsb.org/structure/4i23>; accessed on 30 April 2021) [30]. Discovery Studio Visualizer version 2019 (DSV 2019) was used to visualize the missing amino acid residues. ClustalW (<https://www.genome.jp/tools-bin/clustalw>; (accessed on 30 April 2021) [31] multi-sequence alignment online server was used to align the sequence of crystal structure (PDB 4I23) with the complete sequence of EGFR from Uniprot (UniProt accession no. P00533). This identified twenty amino acid residues data to be missing in the PDB id: 4I23 crystal structure. Further, these missing amino acid residues were fixed using Modeller9.10v software (accessed on 30 April 2021) [32]. While modeling the ligand unbound form, the crystal structure (PDB id: 4I23) was used as a template, and the sequence of UniProt accession no. P00533 entry was used as a query. By using Modeller 9.10v, 100 model structures were generated, and the model with the least discrete optimized protein energy

(DOPE) score was chosen and checked for structural homology with the template using superimpose RMSD option in PyMol. The stereochemical check for the modeled structure was assessed using the PROCHECK web server. Further, EGFR active site residues were mapped based on the crystal structure of EGFR in a complex with dacomitinib.

2.7. Least Potential Energy Conformation by Molecular Dynamics (MD) Simulation

MD simulation has become an important method in computational biology for understanding the dynamic behavior of molecular complexes and intermolecular interactions in biologically simulated conditions [33]. To obtain the lowest potential energy (PE in kcal/mol) conformation of the modeled structure of EGFR, MD simulation was performed for 50 ns using the OPLS force field by Schrodinger Desmond v6.5 software [34]. The protein structure was kept at the center of the cuboidal box and solvated using the SPC water model using periodic boundary conditions in all directions. The distance of the cubic box was set to 10 Å far from all protein edges. The system was neutralized by adding 6 Na⁺ counterions. Further, the SHAKE algorithm was applied to restrain the geometry of water molecules, bond lengths, and bond angles of heavy atoms. The Particle Mesh Ewald method was applied to estimate the long-range interactions that exist between the molecules. The Lennard–Jones interactions cut-off was set to 10 Å. The system was minimized for a 100.0 ps production run using the OPLS force field. Finally, the NPT ensemble was utilized, implementing the Thermostat “Nose-Hoover chain” method with 1.0 ps relaxation time, and the Barostat “Martyna–Tobias–Klein” method (with 2.0 ps relaxation time) was used to maintain the pressure of 1.01325 bar and 300 K temperature. The Coulombic short-range cut-off radius was set to 9.0 Å. The trajectory sampling for the entire run was recorded at an interval of 10.0 ps.

2.8. Molecular Docking

2.8.1. Preparation of Ligand and Protein Target

Based on the network analysis and polypharmacology approach, 11 phytochemicals from *Andrographis paniculata* and 3 from *Thespesia populnea* were also found to interact with EGFR (a highly connected target within the network). Thus, the 3D structures of 14 phytochemicals were downloaded from the PubChem database and virtually screened against EGFR protein. Each ligand was energy minimized by mmff94 force field using the steepest descent algorithm at 5000 energy minimization steps by executing the “prepare_ligand4.py” script [35]. The energy minimization step eliminates the clashes within ligand atoms and produces a reasonable starting pose. Further, the convergence criteria were set to default (1×10^{-6}) with van der Waals (VDW) cut-off distance to 6.0 Å. Finally, the pose with the lowest energy conformation was chosen and converted into an AutoDock vina file (.pdbqt) using the AutoDock vina POAP pipeline [35]. The EGFR structure with the lowest potential energy conformation from the MD simulation was chosen for subsequent virtual screening studies.

2.8.2. Ligand–Protein Molecular Docking

Molecular docking was carried out using AutoDock vina by POAP pipeline [35]. The grid box was created to center $x = 4.00245$, $y = 10.9442$, $z = -4.3899$; and size $x, y, z = 26$ Å; spacing 1 Å based on the active site residues inferred from an earlier ligand co-crystallized complex (PDB: 4I23) of EGFR protein. The exhaustiveness was set to 100 in vina. As a result, it generated nine docked conformations and the protein–ligand complex with the lowest RMSD value was considered as the best pose and was visualized for key intermolecular interactions using Discovery Studio Visualizer 2019. The compound with the lowest BE and the maximal number of interactions with the active site residues was chosen and examined for complex stability by MD simulation (100 ns).

2.9. Stability of Protein–Ligand Complex

After the molecular docking step, protein–ligand complexes MD simulation was performed in triplicate to infer the stability of complex formation during 100 ns of the production run, to ensure plausibility. The above-mentioned “MD simulation” procedure was followed, and the trajectories generated were analyzed using in-build Desmond. The parameters describing structural stability and their residue-wise interaction fluctuations were analyzed. RMSD of backbone, complex, ligand heavy atoms was estimated considering the first frame as a reference, a C- α RMSF, radius of gyration (rGyr) of ligand and protein, and residue-wise interactions were considered during analysis.

2.10. Statistical Analysis

False Discovery Rate < 0.05 was considered significant during gene set enrichment analysis. The network was analyzed by the “Edge count” topological parameter. Docking score was represented in kcal/mol. RMSD, RMSF, and rGyr were used to interpret the stability and fluctuations of ligand–protein interactions using MD simulation.

3. Results

3.1. Selection of Plants and Mining of Phytocompounds

Sixteen plants that are utilized by traditional practitioners/local healers in the region of Western Ghats of India for the treatment of fever, jaundice, inflammation, and viral infection were selected. Initially, 629 phytocompounds from the 16 plants were retrieved from various herbal databases and documented literature. Further, 274 phytocompounds were shortlisted by removing the duplicate and aliphatic chain compounds.

3.2. Druglikeness and Toxicity Profile

Among 274 compounds, 145 compounds showed positive drug-likeness properties. Among these 145 compounds, 36 compounds showed $Pa \leq 0.5$ and nontoxic properties. Table 1 represents the 36 compounds chemical information, druggability, and toxicity profile with their figures (downloaded from PubChem chemical database). The drug-likeness property and toxicity profile of 274 compounds are shown in Tables S2 and S3, respectively.

3.3. Target Identification

Potential protein targets of 36 phytocompounds were predicted using BindingDB. All these phytocompounds were predicted to modulate 206 protein targets (Table S4). Among these, 16 targets were potentially involved in the pathogenesis of the HBV infection-associated HCC (Table 2).

Table 1. Phytocompounds with positive for drug-like properties and their toxicity profile.

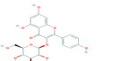
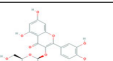
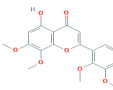
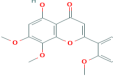
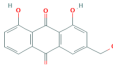
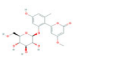
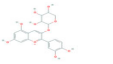
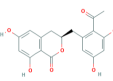
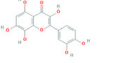
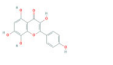
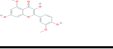
Plant Name	Phytocompounds	PubChem CID	MF	MW (g/mol)	HBA	HBD	LogP	DLS	Pa (≤ 0.5)	Pi	Side Effect	Structure
<i>Aloe vera</i>	2-Methyl-2-Phytyl-6-Chromanol	9865117	C ₂₆ H ₄₄ O ₂	388.33	2	1	10.24	0.83	0.421	0.244	Hepatotoxicity	
<i>Centella asiatica</i>	3-Glucosylkaempferol	5282102	C ₂₁ H ₂₀ O ₁₁	488.1	11	7	−0.26	0.8	0.35	0.3	Hepatotoxicity	
									0.332	0.127	Nephrotoxicity	
<i>Centella asiatica</i>	3-Glucosylquercetin	5280804	C ₂₁ H ₂₀ O ₁₂	464.1	12	8	−0.64	0.89	0.369	0.282	Hepatotoxicity	
									0.302	0.155	Nephrotoxicity	
<i>Andrographis paniculata</i>	5-Hydroxy-7,8,2',3'-Tetramethoxyflavone	5319878	C ₁₉ H ₁₈ O ₇	358.11	7	1	3.46	0.5	Nil			
<i>Andrographis paniculata</i>	5-Hydroxy-7,8,2'-Trimethoxyflavone	5318506	C ₁₈ H ₁₆ O ₆	328.09	6	1	3.49	0.61	Nil			
<i>Aloe vera</i>	Aloe-Emodin	10207	C ₁₅ H ₁₀ O ₅	270.05	5	3	1.91	0.25	0.5	0.05	Cardiac failure	
									0.426	0.24	Hepatotoxicity	
									0.263	0.198	Nephrotoxicity	
<i>Aloe vera</i>	Aloenin	162305	C ₁₉ H ₂₂ O ₁₀	410.12	10	5	−1.36	0.73	0.36	0.108	Nephrotoxicity	
<i>Mangifera indica</i>	Cyanidin 3-arabinoside	74079809	C ₂₀ H ₁₉ O ₁₀	419.1	10	7	−0.07	0.38	Nil			
<i>Aloe vera</i>	Feralolide	5317333	C ₁₈ H ₁₆ O ₇	344.09	7	4	1.58	0.82	0.281	0.178	Nephrotoxicity	
									0.242	0.229	Cardiac failure	
<i>Thespesia populnea</i>	Gossypetin	5280647	C ₁₅ H ₁₀ O ₈	318	8	6	1.61	0.62	0.392	0.265	Hepatotoxicity	
<i>Thespesia populnea</i>	Herbacetin	5280544	C ₁₅ H ₁₀ O ₇	302.04	7	5	1.99	0.63	0.411	0.25	Hepatotoxicity	
<i>Aerva lanata</i>	Isorhamnetin	5281654	C ₁₆ H ₁₂ O ₇	316.06	7	4	2.46	0.67	0.245	0.231	Nephrotoxicity	

Table 1. Cont.

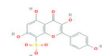
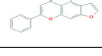
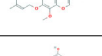
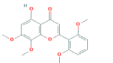
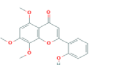
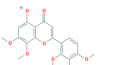
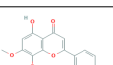
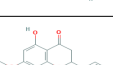
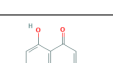
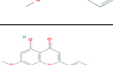
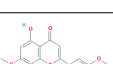
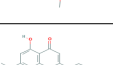
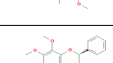
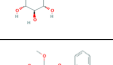
Plant Name	Phytocompounds	PubChem CID	MF	MW (g/mol)	HBA	HBD	LogP	DLS	Pa (≤ 0.5)	Pi	Side Effect	Structure
<i>Phyllanthus virgatus</i>	Kaempferol 8-C-sulfate	10500456	C ₁₅ H ₁₀ O ₉ S	366	9	5	0.41	0.27	Nil			
<i>Thespesia populnea</i>	Kaempferol-7-Glucoside	10095180	C ₂₁ H ₂₀ O ₁₁	448.1	11	7	−0.26	0.81	0.363 0.329	0.288 0.129	Hepatotoxicity Nephrotoxicity	
<i>Vernonia cinerea</i>	Luteolin-7-Mono-Beta-D-Glucoside	5280637	C ₂₁ H ₂₀ O ₁₁	448.1	11	7	−0.07	0.86	0.395 0.373	0.262 0.1	Hepatotoxicity Nephrotoxicity	
<i>Mangifera indica</i>	Myrtillin	165558	C ₂₁ H ₂₁ O ₁₂ .Cl	500.07	12	9	−0.76	0.29	Nil			
<i>Andrographis paniculata</i>	Neoandrographolide	9848024	C ₂₆ H ₄₀ O ₈	480.27	8	4	2.17	0.17	0.424	0.075	Nephrotoxicity	
<i>Andrographis paniculata</i>	Paniculide-B	101289823	C ₁₅ H ₂₀ O ₅	280.13	5	2	0.72	0.09	Nil			
<i>Cassia auriculata</i>	Pelargonidin 5-galactoside	44256698	C ₂₁ H ₂₁ O ₁₀	433.11	10	7	−0.26	0.32	Nil			
<i>Mangifera indica</i>	Peonidin galactoside	91810512	C ₂₂ H ₂₃ O ₁₁ .Cl	498.09	11	7	−0.04	0.08	Nil			
<i>Mangifera indica</i>	Petunidin 3-glucoside	176449	C ₂₂ H ₂₃ O ₁₂ .Cl	514.09	12	8	−0.41	0.42	Nil			
<i>Pongamia pinnata</i>	Pinnatin	5320607	C ₁₈ H ₁₂ O ₄	292.07	4	0	4.42	0.24	Nil			
<i>Pongamia pinnata</i>	Pongamone A	11588678	C ₂₂ H ₂₂ O ₅	366.15	5	0	5.32	0.47	0.397	0.106	Cardiac failure	
<i>Aerva lanata</i>	Tiliroside	5320686	C ₃₀ H ₂₆ O ₁₃	594.14	13	7	2.49	0.85	0.263	0.198	Nephrotoxicity	
<i>Andrographis paniculata</i>	Andrographiside	44593583	C ₂₆ H ₄₀ O ₁₀	512.26	10	6	0.55	0.08	0.333	0.126	Nephrotoxicity	

Table 1. Cont.

Plant Name	Phytocompounds	PubChem CID	MF	MW (g/mol)	HBA	HBD	LogP	DLS	Pa (≤ 0.5)	Pi	Side Effect	Structure
<i>Andrographis paniculata</i>	Altisin	15100719	C ₁₉ H ₁₈ O ₇	358.11	7	1	2.81	0.03	0.306	0.296	Arrhythmia	
<i>Andrographis paniculata</i>	2'-hydroxy-5,7,8-trimethoxyflavone	21668878	C ₁₈ H ₁₆ O ₆	328.09	6	1	3.9	0.2	Nil			
<i>Andrographis paniculata</i>	5-hydroxy-7,8,2',3',4'-pentamethoxyflavone	12315479	C ₂₀ H ₂₀ O ₈	388.12	8	1	3.4	0.36	Nil			
<i>Andrographis paniculata</i>	Wightin	12444943	C ₁₈ H ₁₆ O ₇	344.09	7	2	3.08	0.19	0.342	0.308	Hepatotoxicity	
<i>Andrographis paniculata</i>	5-hydroxy-7,8-dimethoxyflavanone	13963770	C ₁₇ H ₁₆ O ₅	300.1	5	1	3.02	0.48	0.322	0.269	Arrhythmia	
<i>Andrographis paniculata</i>	5-hydroxy-7,8-dimethoxyflavone	188316	C ₁₇ H ₁₄ O ₅	298.08	5	1	3.74	0.17	0.309	0.291	Arrhythmia	
<i>Andrographis paniculata</i>	Skullcapflavone I	5320399	C ₁₇ H ₁₄ O ₆	314.08	6	2	3.71	0.33	Nil			
<i>Andrographis paniculata</i>	5-hydroxy-7,8,2',5'-tetramethoxyflavone	10948318	C ₁₉ H ₁₈ O ₇	358.11	7	1	3.41	0.11	Nil			
<i>Andrographis paniculata</i>	5,4'-dihydroxy-7,8,2',3'-tetramethoxyflavone	13963777	C ₁₉ H ₁₈ O ₈	374.1	8	2	3.02	0.31	Nil			
<i>Andrographis paniculata</i>	Andrographidine A	13963762	C ₂₃ H ₂₆ O ₁₀	462.15	10	4	0.91	0.38	0.43	0.072	Nephrotoxicity	
<i>Andrographis paniculata</i>	Andrographidine C	5318484	C ₂₃ H ₂₄ O ₁₀	460.14	1	4	1.67	0.24	0.348	0.116	Nephrotoxicity	

MF, molecular formula; MW, molecular weight; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; LogP, partition co-efficient; DLS druglikeness score; Nil, non-toxic; Pa, probable activity to cause toxicity; Pi, probable inactivity for non-toxicity.

Table 2. Phytocompounds and their probable predicted protein targets involved in hepatitis B infection pathways and associated hepatocellular carcinoma.

Plant Name	Phytocompounds	Plant Part	Compound Type	Gene Count	Protein Targets Involved in Hepatitis B Infection and HCC Pathways
<i>Aerva lanata</i>	Isorhamnetin	Aerial	Flavonoid	7	CDK1, CDK2, CDK6, EGFR, HGF, IGF1R, TERT
	Tiliroside	Aerial	Flavonoid	1	TNF
<i>Aloe vera</i>	2-Methyl-2-Phytyl-6-Chromanol	Whole plant	Phylloquinone	1	AKT1
	Aloe-Emodin	Leaf, Juice	Anthraquinone	1	BCL2
	Aloenin	Whole plant	Glucoside	1	TNF
	Feralolide	Leaf	Coumarin	2	MAP2K1, MAPK1
<i>Andrographis paniculata</i>	5-Hydroxy-7,8,2',3'-Tetramethoxyflavone	Whole plant, Root	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	5-Hydroxy-7,8,2'-Trimethoxyflavone	Whole plant, Root	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	Neoandrographolide	Whole plant, Root	Diterpenoid	1	STAT3
	Paniculide-B	Whole plant, Stem	Diterpenoid	3	PRKCA, PRKCB, PRKCG
	Andrographiside	Whole plant, Stem	Flavonoid	4	PRKCB, PRKCA, PRKCG, STAT3
	Altisin	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	2'-hydroxy-5,7,8-trimethoxyflavone	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	5-hydroxy-7,8,2',3',4'-pentamethoxyflavone	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	Wightin	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	5-hydroxy-7,8-dimethoxyflavanone	Whole plant, Stem	Flavonoid	3	AKT1, HGF, TERT
	5-hydroxy-7,8-dimethoxyflavone	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	Skullcapflavone I	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	5-hydroxy-7,8,2',5'-tetramethoxyflavone	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	5,4'-dihydroxy-7,8,2',3'-tetramethoxyflavone	Whole plant, Stem	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	Andrographidine C	Whole plant, Stem	Flavonoid	2	EGFR, TNF
<i>Cassia auriculata</i>	Pelargonidin 5-galactoside	Leaves	Anthocyanidin	2	EGFR, TNF
<i>Centella asiatica</i>	3-Glucosylkaempferol	Leaves	Flavonoid	2	EGFR, TNF
	3-Glucosylquercetin	Leaves	Flavonoid	2	EGFR, TNF
<i>Mangifera indica</i>	Cyanidin 3-arabinoside	Peel	Anthocyanin	1	TNF
	Myrtillin	-	Phenolic compound	1	TNF
	Peonidin galactoside	Peel	Anthocyanin	1	TNF
	Petunidin 3-glucoside	Peel	Anthocyanin	1	TNF
<i>Phyllanthus virgatus</i>	Kaempferol 8-C-sulfate	Whole plant	Flavonoid	1	EGFR
<i>Pongamia pinnata</i>	Pinnatin	Seed	Flavonoid	6	CDK1, CDK6, EGFR, HGF, IGF1R, TERT
	Pongamone A	Stem	Flavonoid	1	EGFR
<i>Thespesia populnea</i>	Gossypetin	Flower, Stem, Root	Flavonoid	7	CDK1, CDK2, CDK6, EGFR, HGF, IGF1R, TERT
	Herbacetin	Flower	Flavonoid	7	CDK1, CDK2, CDK6, EGFR, HGF, IGF1R, TERT
	Kaempferol-7-Glucoside	Flowers	Flavonoid	2	EGFR, TNF
<i>Vernonia cinerea</i>	Luteolin-7-Mono-Beta-D-Glucoside	Seed	Flavonoid	2	EGFR, TNF

3.4. Gene Set Enrichment Analysis and Network Construction

Gene IDs of the 16 targets were fetched and provided as input to the STRING 11.0 search. The GO functional and process enrichment showed featured 46 molecular functions and 599 molecular processes to be potentially modulated by these 16 targets

(Tables S5 and S6 respectively). Further, the KEGG pathway enrichment analysis inferred that 139 pathways were modulated by 36 compounds. Among these 139 pathways, 14 were found to be involved in the pathogenesis of HBV-induced HCC. The 16 targets involved in 139 pathways were predicted to be potentially modulated by the phytochemicals, as shown in Table S7. Pathways that are solely involved in the HBV-induced HCC modulated by the phytochemicals are shown in Table 3.

Table 3. Pathways involved in HBV infection and associated hepatocellular carcinoma modulated by the phytochemicals.

KEGG ID	Pathway Name	Gene Count	False Discovery Rate	Modulated Gene within a Pathway
hsa05161	Hepatitis B	11	1.07E-18	AKT1, BCL2, CDK2, CDK6, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, STAT3, TNF
hsa05225	Hepatocellular carcinoma	11	3.43E-18	AKT1, CDK6, EGFR, HGF, IGF1R, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, TERT
hsa01521	EGFR tyrosine kinase inhibitor resistance	11	6.37E-21	AKT1, BCL2, EGFR, HGF, IGF1R, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, STAT3
hsa05206	MicroRNAs in cancer	9	2.07E-14	BCL2, CDK6, EGFR, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, STAT3
hsa04010	MAPK signaling pathway	10	7.51E-14	AKT1, EGFR, HGF, IGF1R, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, TNF
hsa04151	PI3K-Akt signaling pathway	10	3.22E-13	AKT1, BCL2, CDK2, CDK6, EGFR, HGF, IGF1R, MAP2K1, MAPK1, PRKCA
hsa04014	Ras signaling pathway	9	5.60E-13	AKT1, EGFR, HGF, IGF1R, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG
hsa04150	mTOR signaling pathway	8	1.52E-12	AKT1, IGF1R, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, TNF
hsa04210	Apoptosis	5	2.18E-07	AKT1, BCL2, MAP2K1, MAPK1, TNF
hsa05203	Viral carcinogenesis	5	8.12E-07	CDK1, CDK2, CDK6, MAPK1, STAT3
hsa04668	TNF signaling pathway	4	3.85E-06	AKT1, MAP2K1, MAPK1, TNF
hsa04630	Jak-STAT signaling pathway	4	1.57E-05	AKT1, BCL2, EGFR, STAT3
hsa04115	p53 signaling pathway	3	4.58E-05	CDK1, CDK2, CDK6
hsa04110	Cell cycle	3	0.00021	CDK1, CDK2, CDK6

3.5. Network Analysis

Network analysis identified Skullcapflavone I, Wightin, 2'-hydroxy-5,7,8-trimethoxyflavone, Altisin, Andrographiside, 5-Hydroxy-7,8,2'-Trimethoxyflavone, 5-hydroxy-7,8,2',3',4'-pentamethoxyflavone, 5-hydroxy-7,8,2',5'-tetramethoxyflavone, 5,4'-dihydroxy-7,8,2',3'-tetramethoxyflavone, Andrographidine C, Wightin, 5-hydroxy-7,8-dimethoxyflavanone, 5-hydroxy-7,8-dimethoxyflavone, 5-Hydroxy-7,8,2',3'-Tetramethoxyflavone and Paniculide-B from *Andrographis paniculata* (whole plant); Herbacetin, Gossypetin, Kaempferol-7-Glucoside from *Thespesia populnea* (Flowers); Isorhamnetin from *Aerva lanata* (Aerial part); and Pinnatin from *Pongamia pinnata* (Seed) to have highest edge count. Further, among 14 pathways, EGFR tyrosine kinase inhibitor resistance, hepatitis B, and hepatocellular carcinoma pathways were identified as enriched pathways via modulating EGFR, AKT1, BCL2, HGF, IGF1R, MAP2K1, MAPK1, PRKCA, PRKCB, PRKCG, STAT3, CDK2, CDK6, STAT3, TNF, and TERT as common target proteins (Tables 2 and 3). However, the above-selected compounds were considered based on polypharmacology therapy due to their probable affinity toward multiple targets and pathways. The herb-compound-protein pathway network is shown in Figure 3. For further docking studies, based on the polypharmacology analysis, phytochemicals that are predicted to modulate multiple protein targets and pathways were selected; therefore, 11 phytochemicals from *Andrographis paniculata* were predicted to target 13 protein molecules and 3 phytochemicals from *Thespesia populnea* were predicted to target 8 protein molecules involved in HBV-induced HCC and are repre-

sented in a Venn diagram (Figure S1). Phytocompound-affected checkpoints within the Hepatitis B infection pathway (hsa05161) were traced from the KEGG pathway and are represented in Figure 4.

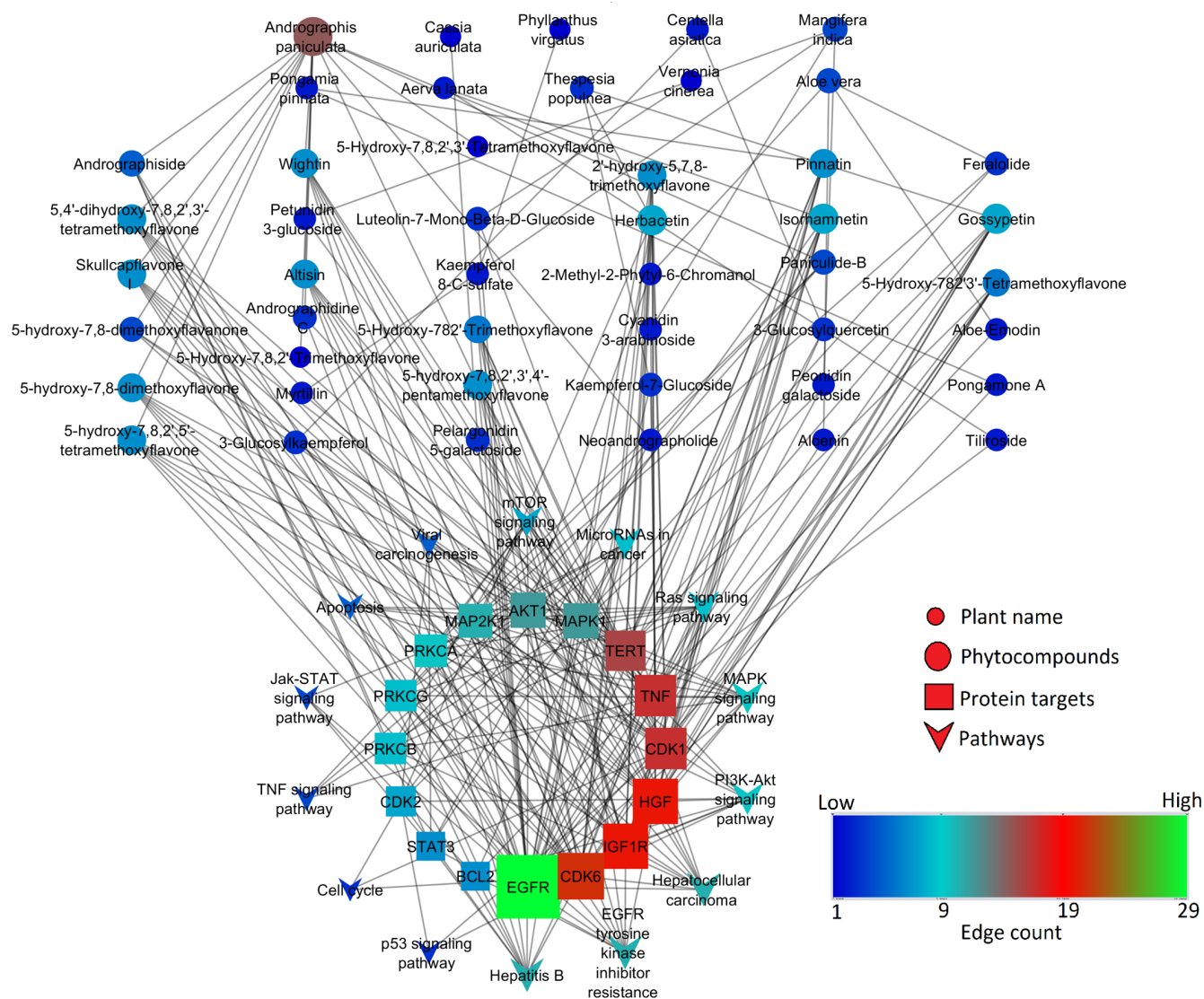


Figure 3. Network representation of herbs–phytocompounds–proteins pathways of HBV infection and HCC. The network node size “Low to High” representation is based on the degree of connection among the nodes and was analyzed by a network topological parameter “edge count”. It is highly evident that the EGFR node scored the highest edge count of 29, compared to all the other targets studied.

Further, the study was extended to molecular docking and dynamics simulations. As discussed earlier, EGFR was chosen as the most potential target, as it was found to be the highly connected protein target within the network and is also a well-known therapeutic target in many types of cancer, viz., lung cancer [36], glioblastoma [37], breast carcinoma [38], HCC [39,40], and so on. In this study, EGFR was found to be involved in hepatocellular carcinoma, MAPK, PI3K-Akt, Jak-STAT, and Ras signaling pathways, and the EGFR tyrosine kinase inhibitor resistance pathway. Based on the docking study, 11 phytocompounds from *Andrographis paniculata* and 3 from *Thespesia populnea* were identified to potentially target EGFR protein.

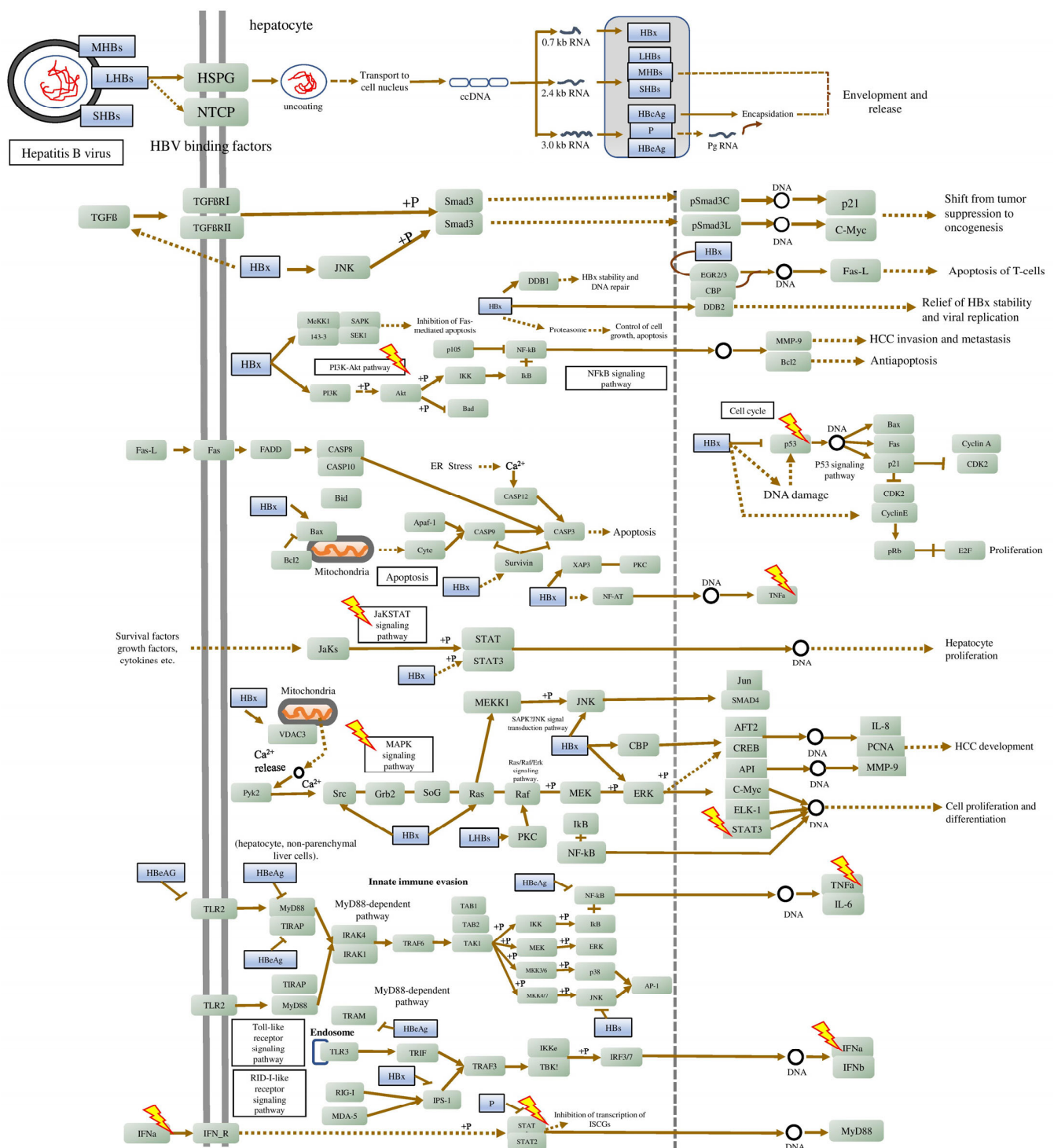


Figure 4. Major checkpoints to be affected by the phytochemicals. ⚡ Affected checkpoints.

3.6. Structure Refinement, Homology Modeling, and Active Sites Assessment

As the WT-EGFR crystal structure (PDB id: 4I23) had missing amino acids, the entire structure was rebuilt by homology modeling. To fill the missing residues and to obtain a plausible structure, 100 homology structural models of WT-EGFR were generated using Modeller 9.10v software. Among these, Model number 30 was chosen as the best model, as it featured the lowest DOPE score, and it had an RMSD of 0.292 Å w.r.t. template PDB id: 4I23. The Ramachandran plot also revealed about 89.9% and 8.7% residues in the most favored and additionally allowed region (Figure S2). The active site residues of WT-EGFR

reported in the co-crystallized structure are Leu23, Ala48, Lys50, Glu67, Met71, Leu93, Thr95, Gln96, Met98, Pro99, and Gly101 [30].

3.7. Model Stability and Lowest Potential Energy Conformation Sampling by MD Simulation

The stability of the WT-EGFR protein modeled structure was validated by an MD run of 50 ns, and it was found to be highly stable with few perturbations. The lowest PE conformation for the WT-EGFR protein was found to be $-129,809.51$ kcal/mol at 45,252.0th ps (45.2 ns) (Figure S3B). The lowest PE conformations were extracted and used further for docking studies.

3.8. Molecular Docking Study

Molecular docking of 11 phytocompounds from *Andrographis paniculata* and 3 from *Thespesia populnea* were performed with WT-EGFR, and the corresponding binding energies (BE) and intermolecular interactions are summarized in Table 4. First, all 11 compounds from *A. paniculata* were predicted to form interactions with the active site residues. Among these, Andrographin scored the lowest BE of -8.3 kcal/mol via forming two hydrogen bonds, i.e., Met98...=O, Met98...OH, and forming ten hydrophobic interactions, i.e., Thr159, Lys50, Val31 (2), Ala48 (2), Leu149 (2), Leu23 (2) residues. Among these, seven interactions were found to span the active site residues. The intermolecular interactions of Andrographin with WT-EGFR are shown in Figure 5A. Second, all three compounds from *T. populnea* were predicted to form stable interactions with active site residues. Among these, Gossypetin scored the lowest BE of -8.5 kcal/mol via forming three hydrogen bond interactions, i.e., Met98...OH, Thr95...OH (2) and eight hydrophobic bonds, i.e., Thr159, Ala48, Val31 (2), Leu149 (2), Leu23 (2). Among these, six interactions were found to span with active site residues. The interactions of Gossypetin with WT-EGFR are shown in Figure 5B.

3.9. Stability of Protein–Ligand Complexes

Based on the lowest BE and maximum intermolecular interactions with active site residues, two compounds were prioritized for MD simulation, i.e., Andrographin from *A. paniculata* and Gossypetin from *T. populnea*. The stability of complex formation of the WT-EGFR-Andrographin complex and WT-EGFR-Gossypetin complex were validated by MD simulation for 100 ns. The simulation for both complexes was carried out in triplicate, to ensure plausibility.

3.9.1. Andrographin and WT-EGFR Complex

The Andrographin with WT-EGFR simulation of the system was neutralized by adding 8 Na^+ ions (12.777 mM concentration) and was subjected to a 100 ns run of production run with a 10 ps recording interval. The backbone RMSD fluctuation was within 0.7 to 3.6 Å , and the average RMSD was 2.28 Å (Figure 6A). The complex RMSD was found within the range of 0.3 to 3.5 Å from 0 to 100 ns in all three replicates. The average RMSD of the complex was 1.24 Å (Figure 6B). Similarly, the initial average RMSD of ligand w.r.t. ligand for all three replicas was 0.356 Å , and the final was 0.45 Å . The total average RMSD of ligand w.r.t. ligand for all three replicas was 0.44 Å (Figure 6C). The RMSF analysis was performed to infer residue-wise fluctuations observed during the production run. The average RMSF value of residues was $\sim 0.7\text{ Å}$. The c-terminal loop region shows maximum fluctuations due to the flexibility of terminal residues' maximum fluctuation ($\sim 2.5\text{ Å}$), whereas the RMSF of active residues involved in the ligand interactions, viz., Gly24, Val31, Met98, and Leu149 was $\sim 1.0\text{ Å}$ (Figure 7A). Additionally, the radius of gyration (rGyr) trajectory over the 100 ns production run was analyzed. The rGyr deviation of Andrographin with EGFR was found within 3.60 to 3.68 Å , which indicates the higher compactness of the Andrographin-EGFR complex (Figure 7B). Similarly, in all three replicas, the protein rGyr was also found to be uniform (within ~ 20 to 20.8 Å) throughout the simulation period (Figure 7C).

Table 4. Binding energy and molecular interactions of phytocompounds with WT-EGFR protein target.

Compound Name	BE (kcal/mol)	Hydrogen Bond Interactions (No. of Interactions)	Van der Waals, Pi-Alkyl, CH, Pi-Cation, Pi-Sigma, Pi-Pi Stacked, Pi-Pi T-Shaped Interactions (No. of Interactions)	Active Sites Residues within Interactions	No. of Interactions with Active Site Residues
Kaempferol-7-Glucoside	−8.8	Asp105 ... OH, Thr95 ... OH, Gln96 ... OH	Val31, Leu149, Leu23 (3), Cys102	Thr95, Gln96	5
Herbacetin	−8	Thr95 ... OH, Met98 ... OH (2), Gly101 ... OH	Leu23 (2), Leu149 (2), Val31 (2), Ala48, Cys80, Thr159	Thr95, Met98, Gly101, Leu23, Ala48	7
Gossypetin	−8.5	Met98...OH, Thr95 ... OH (2)	Thr159, Ala48, Val31 (2), Leu149 (2), Leu23 (2)	Thr95, Met98, Leu23, Ala48	6
Andrographidine C	−8.3	Thr95 ... OH (2), Gln96 ... OH, Gly26	Arg146, Val31, Leu149, Leu23 (2), Cys102 (2), Asp105	Thr95, Gln96, Leu23, Cys102	7
5-Hydroxy-7,8,2'-Trimethoxyflavone (Andrographin)	−8.3	Met98 ... =O, Met98 ... OH	Thr159, Lys50, Val31 (2), Ala48 (2), Leu149 (2), Leu23 (2)	Met98, Lys50, Ala48, Leu23	7
Skullcapflavone I	−8.1	Met98 ... =O, Met98 ... OH	Thr159, Val31 (2), Lys50, Ala48 (2), Leu23 (2), Leu149, Asp105	Met98, Lys50, Ala48, Leu23	7
2'-hydroxy-5,7,8-trimethoxyflavone	−8.1	Nil	Gly101, Leu23 (2), Leu149 (2), Val31 (2), Gly24	Leu23	2
5,4'-dihydroxy-7,8,2',3'-tetramethoxyflavone	−7.7	Nil	Gln96, Ala48, Thr159, Lys50, Val31 (2), Leu149, Leu23 (2), Cys102 (2), Asp105	Gln96, Ala48, Lys50, Leu23, Cys102	7
5-hydroxy-7,8-dimethoxyflavone	−8.2	Met98 ... =O	Leu23 (2), Thr159, Val31 (2), Lys50, Ala48 (2), Leu149 (2), Asp105	Met98, Lys50, Ala48, Leu23	5
5-hydroxy-7,8,2',5'-tetramethoxyflavone	−8.4	Met98 ... =O	Val31 (2), Lys50, Thr159, Ala48, Leu23 (2), Leu149 (2), Asp105	Met98, Lys50, Ala48, Leu23	5
5-hydroxy-7,8,2',3',4'-pentamethoxyflavone	−8	Nil	Lys50, Thr159 (2), Val31 (2), Gln96, Ala48, Leu149, Leu23 (2), Cys102 (2)	Lys50, Gln96, Ala48, Leu23	5
Wightin	−8.3	Met98 ... =O, Met98 ... OH	Val31 (2), Lys50, Thr159, Ala48 (2), Leu23 (2), Leu149 (2), Asp105	Met98, Lys50, Leu23, Ala48,	7
5-Hydroxy-7,8,2',3'-Tetramethoxyflavone	−8.6	Met98 ... =O	Asp105, Val31 (2), Thr159, Lys50, Leu23 (2), Leu149 (2), Ala48, Gln96	Met98, Lys50, Leu23, Ala48, Gln96	6
Altisin	−7.8	Met98 ... O-	Val31 (2), Ala48, Leu149 (2), Leu23 (2)	Met98, Ala48, Leu23	4

BE, binding energy.

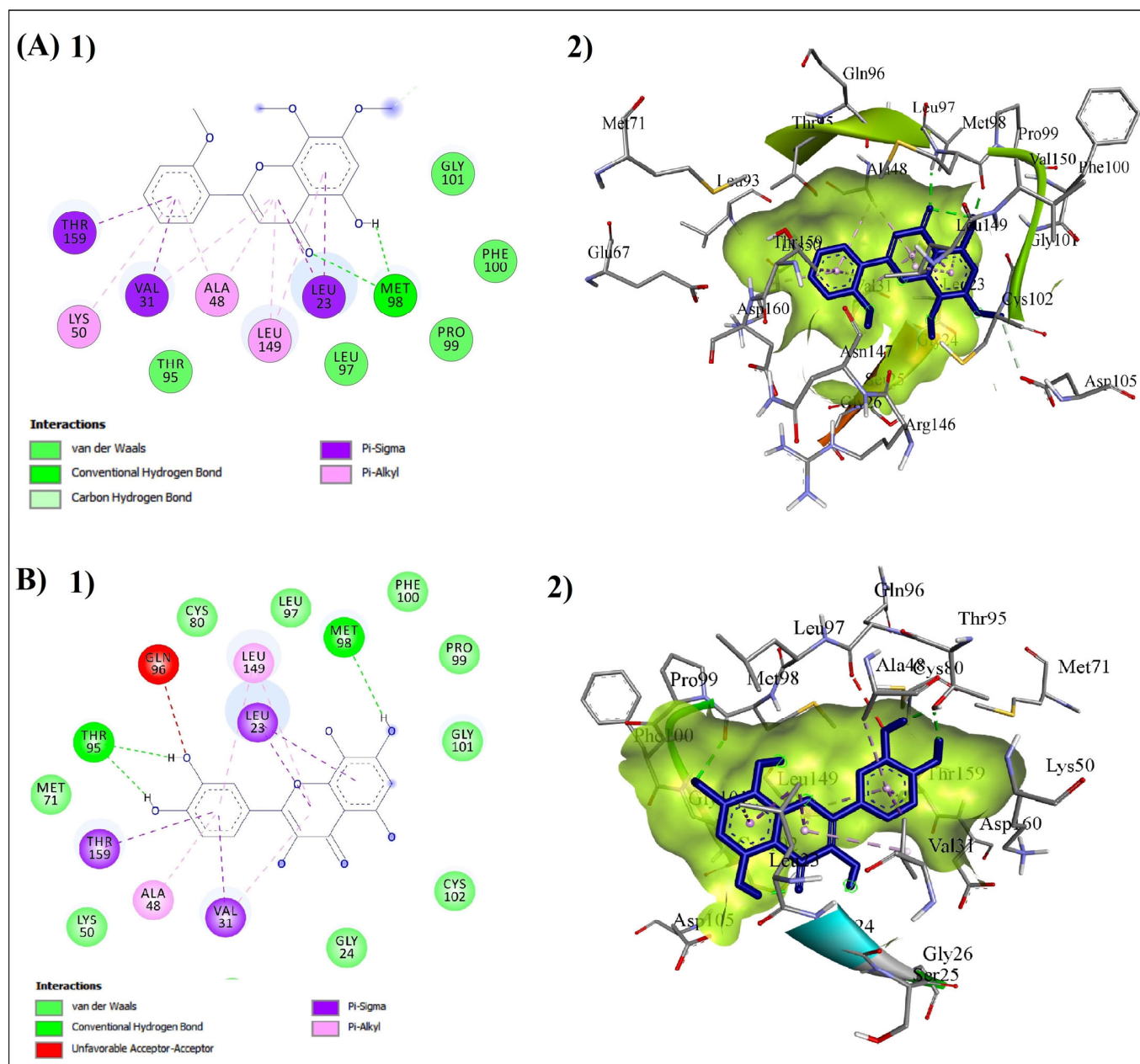


Figure 5. Intermolecular interactions of (A) Andrographin and (B) Gossypetin with WT-EGFR; (1) 2D representation and (2) 3D representation of compounds within the binding pocket.

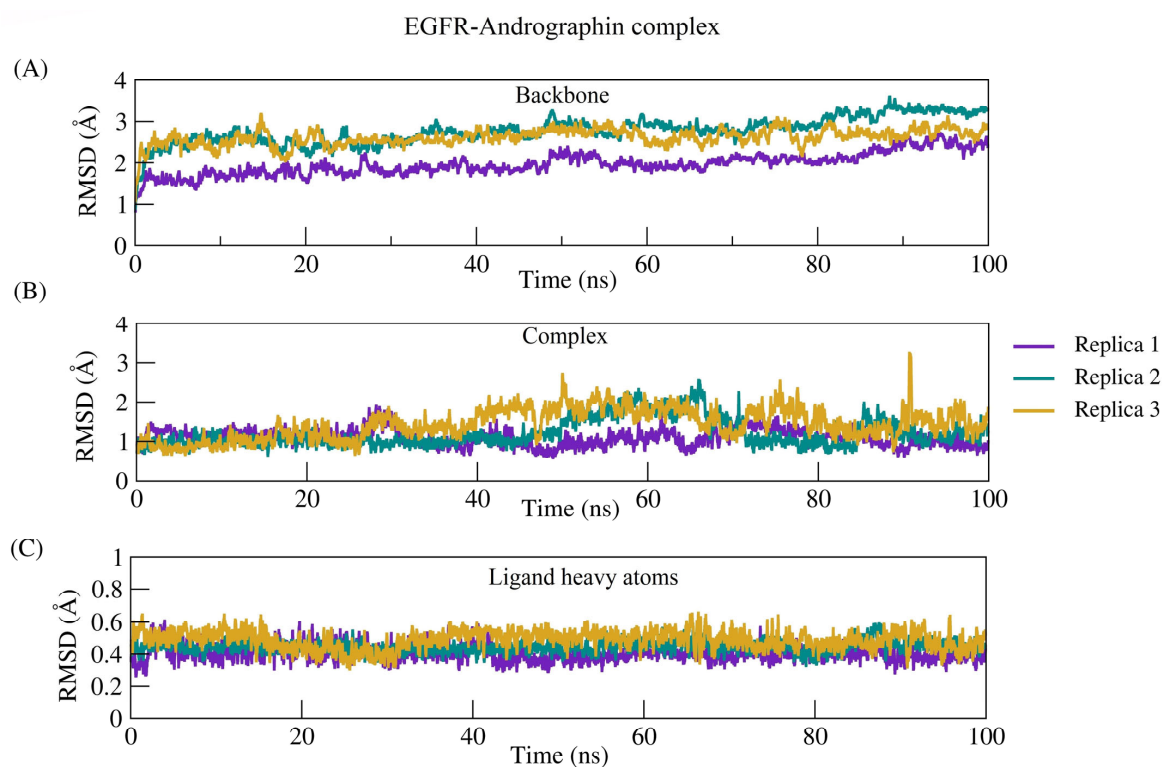


Figure 6. Andrographin with WT-EGFR 100 ns MD production run. (A) RMSD plot depicting the “Protein Backbone”. (B) Plot depicting the docked complex stability “ligand–protein complex RMSD”. (C) Plot depicting the stability of “ligand w.r.t. ligand”.

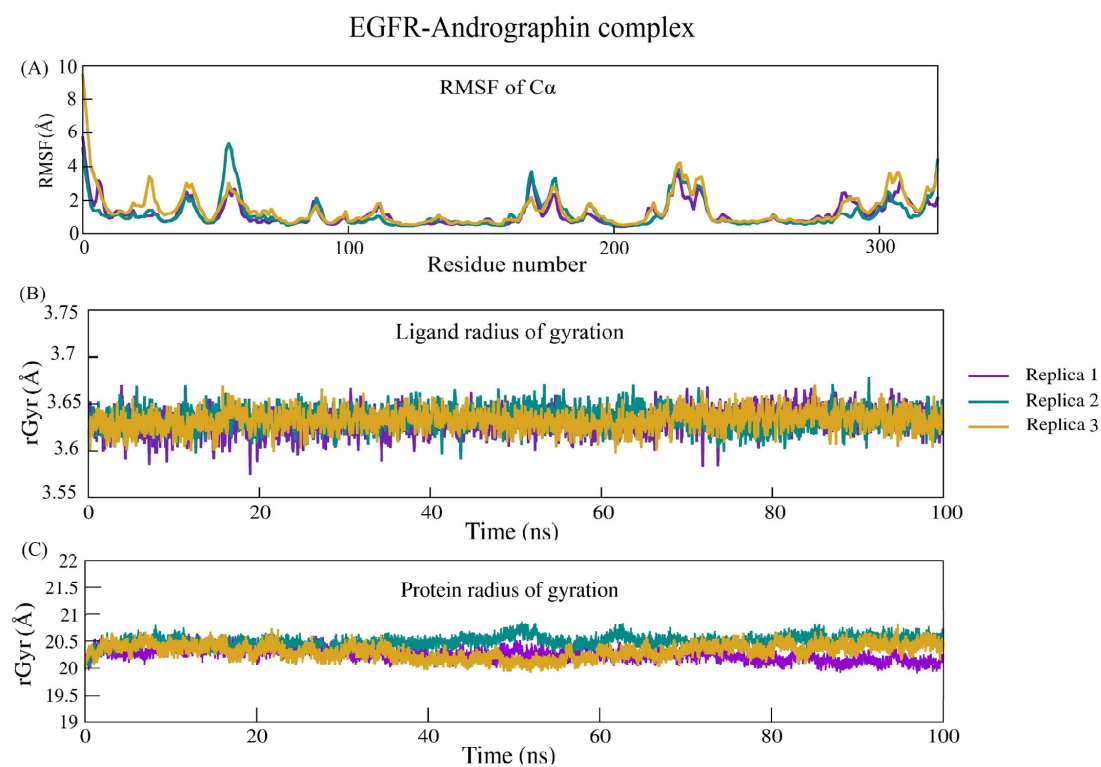


Figure 7. EGFR–Andrographin complex 100 ns MD production run. (A) Protein RMSF plot. (B) A gyration plot depicting the extendedness of a ligand. (C) A gyration plot of protein during a 100 ns production run.

Further, Andrographin–EGFR contact analysis concluded that Met98 formed stable hydrogen-bonded interactions for around 98–99% and 73–77% of the duration with =O and OH group of Andrographin, respectively, in all three replicates. Pro99 and Cys102 and Asp105 were found to be involved in water-bridge interactions and showed around 10–15%, 12–20%, and 19–32% of the time with interaction fractions in all three replicates, respectively. Figure 6, Figure 7, and Figure 8 represent the RMSD, RMSF, rGyr, and contacts of the Andrographin WT–EGFR complex.

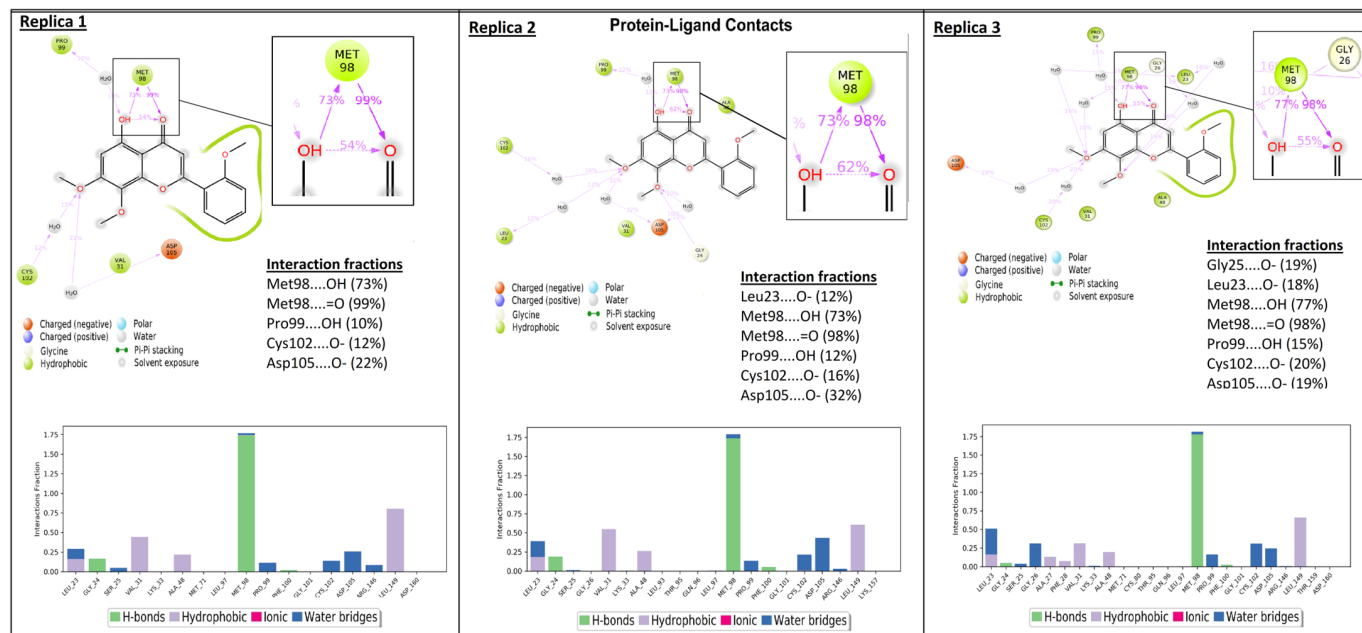


Figure 8. EGFR–Andrographin complex sustained intermolecular interactions and percentage interaction fraction throughout 100 ns production run.

3.9.2. Gossypetin and WT–EGFR Complex

The Gossypetin with WT–EGFR simulation of the system was neutralized by adding eight Na^+ ions (12.778 mM concentration) and being subjected to a 100 ns run of production with a 10 ps recording interval. The backbone RMSD fluctuation was to be within 1.2 to 3.0 Å, and the average RMSD was found to be ~ 2.6 Å for all three replicas (Figure 9A). The complex RMSD fluctuation was found to be within 0.699 to 3.53 Å, and the average RMSD was found to be ~ 2.6 Å. In all three replicates, the complex RMSD was gradually increased with slight fluctuation from 0 to 50 ns (~ 0.7 to ~ 2.2 Å) and again attained stability from 50 to 100 ns (Figure 9B). Similarly, the initial average RMSD of ligand w.r.t. ligand for all three replicas was 0.294 Å, and the final was 0.313 Å. The total average RMSD of ligand w.r.t. ligand for all three replicas was 0.3 Å. However, replica 3 of Gossypetin showed slight fluctuation from 35 to 50 ns (0.3 to 0.7 Å) and again attained stability from 50 ns at 0.7 Å. Further, a decrease in RMSD was seen at 75 ns (0.7 to 0.3 Å) and was found to be stable throughout 100 ns (Figure 9C). The RMSF analysis was performed to infer residue-wise fluctuations observed during the production run. The average RMSF value of residues was ~ 1.4 Å. The c-terminal loop region showed maximum fluctuations due to the flexibility of terminal residues' maximum fluctuation (~ 2.5 Å), whereas the RMSF of active residues involved in the ligand interactions, viz., Ser25, Lys50, Cys80, Thr95, Gln96, Met98, Pro99, Leu149, and Thr159 was < 1.3 Å (Figure 10A). The radius of gyration (rGyr) trajectory over the 100 ns production run was also analyzed. The rGyr deviation of Gossypetin with EGFR was found within 3.7 to 3.8 Å, which indicates the higher compactness of the Gossypetin–EGFR complex (Figure 10B). Similarly, in all three replicas, the protein rGyr was also found to be uniform (within ~ 20 to 20.8 Å) throughout the simulation period (Figure 10C).

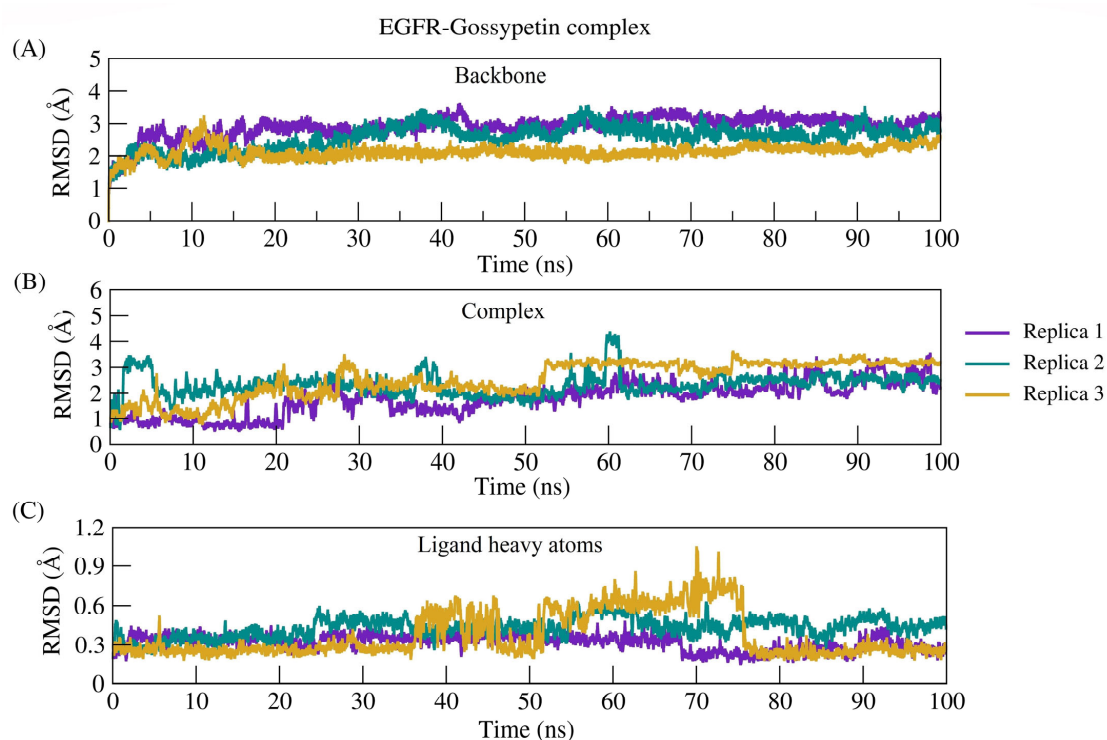


Figure 9. Gossypetin with WT-EGFR 100ns MD production run. (A) RMSD plot depicting the “protein backbone”. (B) Plot depicting the docked complex stability “ligand–protein complex RMSD”. (C) Plot depicting the stability of “ligand w.r.t. ligand”.

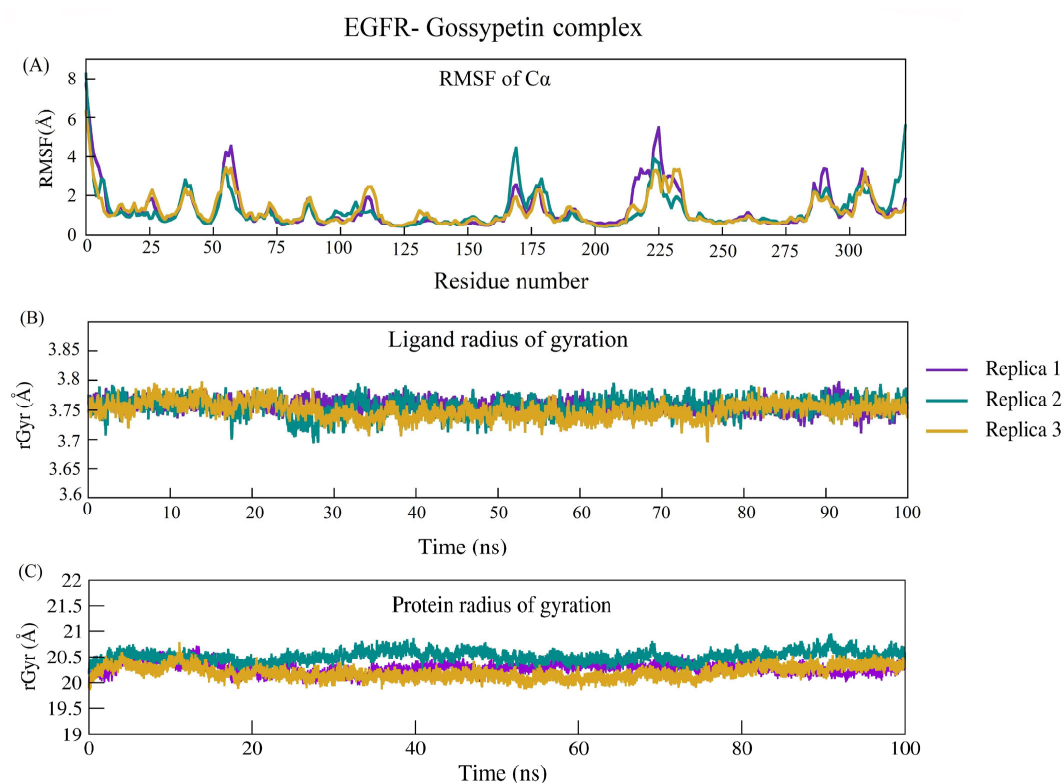


Figure 10. EGFR–Gossypetin complex 100 ns MD production run. (A) Protein RMSF plot. (B) Gyration plot depicting the extendedness of a ligand. (C) Gyration plot of protein during a 100 ns production run.

Further, Gossypetin–EGFR contact analysis concluded that Gln96 formed stable hydrogen-bonded interactions for around ~93% and 21% of the duration with the OH group of Gossypetin in replicas 1 and 3 and around 53% and 27% of the water–bridge in replicas 2 and 3. Met98, Thr95, and Asp105 residues were found to be common interactive residues and to be involved in hydrogen bonds, hydrophobic bonds, and water–bridge interactions in all three replicates. Figure 9, Figure 10, and Figure 11 represent the RMSD, RMSF, rGyr, and contacts of the Gossypetin–EGFR complex.

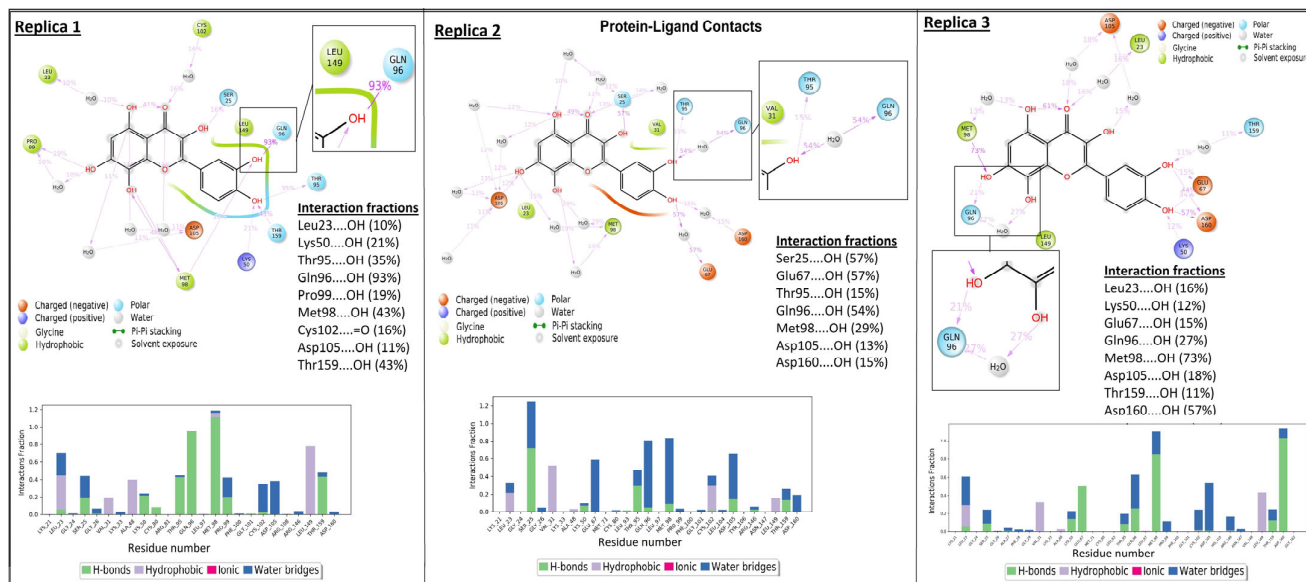


Figure 11. EGFR–Gossypetin complex sustained intermolecular interactions and percentage interaction fraction throughout 100 ns production run.

4. Discussion

Hepatocyte damage due to HBV infection and pathogenesis of liver disease are not directly associated with HBV; however, they are caused by immune-mediated host HBV interactions [41]. The complex mechanisms that cause HBV-mediated liver disease are not fully understood due to the involvement of complex protein–protein interactions with multiple host factors [42]. However, the utilization of already existing single drug molecules against a cluster of HBV-mediated liver disease pathways may cause adverse drug effects by disrupting the host tissue mechanisms. The utilization of chemotherapeutic agents during HCC may alter hepatic metabolism, reduce efficacy, or may cause liver toxicity by the disposition of the drug in liver tissue [43]. Further, physicians utilize palliative treatments such as systemic chemotherapy, interferon, transarterial chemo-embolization, and hormone therapy against HCC when there are no curative treatment options. However, the utilization of these therapies for HCC patients is challenging because HCC is highly resistant to systemic therapies [44]. In the current study, we elucidated the molecular mechanisms and probable safe use of traditionally used herbs of Western Ghats, Karnataka against jaundice, fever, inflammation, and viral infections via cluster and enrichment analysis of potential HBV disease protein targets that participate in the progression of immune processes and inflammatory responses. The results obtained from the current study predict that flavonoids and diterpenoids from *A. paniculata* and flavonoids from *T. populnea* could potentially prevent disease progression and could be effective therapeutics against chronic HBV-induced HCC.

Identification of potent candidates from herbal medicines is not only the key focus of new drug discovery against HBV-induced HCC, but drug safety is also a key concern. This means that ensuring that the drugs have a positive benefit–risk balance is the goal [45]. Hepatotoxicity induced by herbs is one of the most important factors limiting treatment. The issue has often been responsible for post-marketing warnings and withdrawal [46].

Many researchers have well documented adverse hepatic reactions and idiosyncratic liver damage caused by herbs. The major hepatotoxic herbs reported thus far are *Teucrium chamaedrys*, *Larrea tridentate*, *Cimicifuga racemosa*, *Scutellaria lateriflora*, and *Scutellaria baicalensis* [47]. In the current study, we screened 274 phytocompounds from 16 herbs for their druggable characteristics using the Lipinski rule of five, and we selected only compounds with positive drug-like properties. Further, these compounds were screened for major adverse drug reactions and side effects, i.e., hepatotoxicity, cardiac toxicity, nephrotoxicity, and potential to cause myocardial infarction using the ADVERpred online prediction. To minimize the toxicity, compounds with a probable toxic activity (Pa) with a p value ≤ 0.5 were considered. Further, compounds were also predicted for their probable affinity toward potential drug targets using BindingDB (p value ≥ 0.7), and they segregated the target associated with HBV infection and HCC with the aid of published literature, KEGG pathway, and Therapeutic Target Database (TTD).

STRING database was also utilized to identify the molecular functions, biological processes, and pathways regulated by the predicted targets, wherein the targets were meticulously chosen solely based on potential molecular functions, processes, and pathways with higher significance ($p \leq 0.05$) and those that are only associated with HBV-induced HCC. Based on the results obtained from STRING, we constructed a compound–protein pathway network building to understand the relations between compounds, target proteins, and pathways. One pathway containing many protein molecules modulated by one compound is more essential than the function of one protein that involves many pathways modulated by one drug molecule. The impact of one target in numerous pathways may be less, and the impact of one pathway with multiple protein targets modulated by one compound could be large [48]. Based on this theory, we chose 11 phytocompounds, viz., Andrographidine C, 5-Hydroxy-7,8,2'-Trimethoxyflavone (Andrographin), Skullcapflavone I, 2'-hydroxy-5,7,8-trimethoxyflavone, 5,4'-dihydroxy-7,8,2',3'-tetramethoxyflavone, 5-hydroxy-7,8-dimethoxyflavone, 5-hydroxy-7,8,2',5'-tetramethoxyflavone, 5-hydroxy-7,8,2',3',4'-pentamethoxyflavone, Wightin, 5-Hydroxy-7,8,2',3'-Tetramethoxyflavone, and Altisin from *A. paniculata* (are of Andrographolide derivatives) contained in *A. paniculata* whole plant and Gossypetin, Herbacetin, and Kaempferol-7-Glucoside contained in *T. populnea* flowers [49]. These phytocompounds were identified as flavonoids and diterpenoids and were found to modulate thirteen protein molecules involved in the progression of HBV infection to HCC, i.e., CDK1, CDK2, CDK6, EGFR, HGF, IGF1R, TERT, STAT3, PRKCA, PRKCB, PRKCG, AKT1, and TNF. Further, these protein targets were also found to modulate 14 major molecular pathways, i.e., Hepatitis B pathway, hepatocellular carcinoma, MicroRNAs in cancer, MAPK signaling pathway, PI3K-Akt signaling pathway, Ras signaling pathway, mTOR signaling pathway, apoptosis, viral carcinogenesis, TNF signaling pathway, Jak-STAT signaling pathway, p53 signaling pathway, and cell cycle. The roles of each protein target associated with the identified pathways involved in the pathogenesis of HBV infection and progression of HCC are shown in Table S8.

After gene ontology, enrichment, and network analysis, we further extended the study to examine the intermolecular interactions between phytocompounds and predicted protein targets using docking and MD simulation studies. As a result, we have chosen EGFR, as it earned the top-scoring node with the highest edge count of 29. Thus, this formed the sole rationale for the initial blind selection of EGFR as a potential target. On secondary literature analysis, it was found to be well documented as a therapeutic target for cancer (identified to be involved in the MAPK, PI3-Akt, Ras, and Jak-STAT signaling pathways within the network). Moreover, earlier studies have demonstrated that overexpression of EGFR is common in HCC, meaning that it can play a role in the pathogenesis and treatment of the disease. Hence, on cumulative analysis of blind network topology prediction and the documented studies, EGFR was finalized as the most potential target. It is well known that NTCP host receptors are essential for HBV entry. However, NTCP expression alone is not enough for the entry of HBV. EGFR is a host cofactor that regulates the host cell's ability to support HBV internalization and is essential for NTCP's viral receptor activity [50].

After HBV binds to NTCP on the cell surface, the HBV–NTCP complex is internalized via EGFR endocytosis machinery, which includes EGFR phosphorylation and the subsequent recruitment of adaptor molecules such as “AP2A1 and EPS15” [51–53]. The EGFR-sorting machinery (involving EGFR ubiquitination, STAM, and LAPTM4B) coordinates HBV transport in the endosomal network (early and late endosome and lysosome) [51]. In addition, researchers examined the EGF-related host factor(s) in HBV infection by siRNA-mediated knockdown of endogenous EGFR in HepG2-NTCP cells, HepaRG cells, and primary human hepatocytes and observed a significant reduction in HBV infection. The researchers reported that HBV viral attachment was not decreased in EGFR-knockdown cells, but the level of HBV DNA internalized into cells was significantly reduced by knockdown of EGFR [52]. From early inflammation and hepatocellular proliferation to fibrogenesis and neoplastic transformation, the EGFR signaling mechanisms have been recognized as a central player in all stages of the liver response to injury [39,40]. Schiffer et al. [54] reported Gefitinib, a potent EGFR tyrosine kinase inhibitor to curtail HCC development in rats exposed to diethylnitrosamine. Further, Zheng et al. [55] reported that Tropomodulin 3 levels rise in HCC conditions, especially when it spreads outside the liver. Tmod3 expression was also linked to aggressive malignancy and poor patient survival in HCC patients [56]. Tmod3 enhances matrix metalloproteinase-2, -7, and -9 transcription, which is dependent on PI3K-AKT. The interactions between Tmod3 and EGFR, which promote EGFR phosphorylation, are required for signaling activation of PI3K-Akt. Similarly, VersicanV1 was also found to enhance the EGFR–PI3K–Akt signaling pathway in HCC cells [56]. A recent study by Liu et al. [57] demonstrated that in HCC cells, the EGFR-P38 MAPK axis may upregulate PD-L1 via miR-675-5p and downregulate HLA-ABC via HK2, and this might be responsible for the immune suppression in HCC, and they suggest that EGFR signaling might be targeted for HCC immunotherapy. Hence, based on the potential role of EGFR in HCC and key network hub status in the current study, the phytochemicals were screened against EGFR. Table S9 provides information on phytochemicals targeting protein target and modulated pathways (edge count was mentioned for w.r.t. protein target, i.e., number of compounds and number of pathways connecting to respective protein targets). First, the WT-EGFR X-ray crystallographic protein structure was obtained from the PDB, and the missing residues were fixed by homology modeling. The structural stability of the best model of the EGFR protein was confirmed by molecular dynamics for a 50 ns production run. The least potential energy conformation from MD simulations was retrieved for utilization in docking studies. The molecular docking of 11 compounds from *A. paniculata* and 3 from *T. populnea* showed significant intermolecular interactions with active site residues. Further, based on the BE and highest interactions with active site residues, 5-Hydroxy-7,8,2'-Trimethoxyflavone (Andrographin) from *A. paniculata* and Gossypetin (*T. populnea*) were identified as lead hits against WT-EGFR. All shortlisted compounds were predicted to interact with active site residues. On post docking studies, the MD method was used to delineate the effects of protein structure stability on ligand binding. To explain the structural modification of the protein upon ligand binding, we ran three replicates of 100 ns MD simulations for Andrographin and Gossypetin with WT-EGFR to ensure the plausibility of complex formation. The protein–ligand interactions for RMSD were found to be stable and formed stable interactions throughout 100 ns production runs in all three replicates.

Andrographis paniculata and *Thespesia Populnea*

Andrographis paniculata Nees. is known as the “king of bitters”, a member of an Acanthaceae family, and is native to India and Sri Lanka. In Ayurveda, the roots and leaves are used for medicinal purposes [58]. *A. paniculata* is a key ingredient of many herbal formulations for the treatment of liver disease, hepatitis, diabetes, cancer, etc. [59,60]. A previous study by Tan et al. [61] demonstrated Andrographolide to downregulate EGFR expressed in epidermoid carcinoma (A-431) cells. Andrographolide and Neoandrographolide from *A. paniculata* showed a protective effect against tert-butyl hydroperoxide intoxication-induced hepatotoxicity in mice [62,63] and protective activity against galactosamine-induced hepa-

toxicity in rats [64]. The crude extract of *A. paniculata* showed potent inhibitory activity against the HBV surface antigen [65]. Dehydroandrographolide and Andrographolide showed potent inhibitory activity against HBV DNA replication [66]. Neoandrographolide is a potent anti-inflammatory compound that acts via inhibition of iNOS and COX-2 expression through inhibiting p38 MAPK activation [67]. It also suppresses NO production in LPS-activated macrophages [65].

Thespesia populnea Soland ex. Correa belongs to the family Malvaceae. It is mainly found in tropical regions and the coastal forests of India. Ayurvedic physicians use bark decoction for the treatment of skin and liver diseases. The bark and flower possess hepatoprotective, antioxidant, anti-inflammatory, memory-improving, and cholesterol-lowering activities, and have been reported for the management of Alzheimer's disease [68]. In the current study, we identified Gossypetin, Herbacetin, and Kaempferol-7-Glucoside flavonoids from *T. populnea* flower as potent hepatoprotective hits against HBV-induced HCC. This supports earlier reports of these compounds as potent antiviral flavonoids [49,69,70]. Arthanari et al. also demonstrated that methanolic flower extracts of *T. populnea* have a strong antiviral potency and possess nontoxic properties [71].

The idea of modulating multiple proteins by a single molecule has triggered researchers' interest in finding the first strike on specific targets. Phytocompounds from *A. paniculata* and *T. populnea* have also been shown to have antiviral and hepatoprotective properties in the past. As a result, the current research discovered 11 compounds from *A. paniculata* and 3 from *T. populnea* that may have effectiveness against hepatocellular carcinoma caused by the HBV virus, which has further prompted us to investigate the binding affinity and intermolecular interactions of these compounds with active site residues of EGFR, a well-known cancer therapeutic target. The identified small molecules possess druggability and a non-toxic profile, which ensures safety and efficacy.

5. Conclusions

Currently, there is no satisfactory therapy for HBV infection and HCC. The present study identified a group of druggable and nontoxic phytocompounds from herbs modulating the therapeutic targets involved in HBV infection and HCC through in silico techniques involving compound–protein pathway enrichment analysis coupled with network pharmacology, molecular docking, and dynamics simulation studies. Results of this study suggest that flavonoid and a diterpenoid fraction of *A. paniculata* and flavonoid fraction of *T. populnea* could be potent as a combined therapy for the treatment of patients with HBV infection and its associated HCC without expecting adverse drug reactions. Theoretical and predicted knowledge generated in this study needs to be further confirmed using in vitro studies including in viral cell lines and animal models.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/app122110691/s1>, Table S1: List of references for the selected medicinal plant for the treatment of hepatitis-like illness including jaundice, hepatitis, fever, inflammation and for hepatoprotective and antiviral potential; Table S2: List of plants with their phytocompounds and druglikeness properties; Table S3: Toxicity profile of phytocompounds; Table S4: Overall probable targets modulated by phytocompounds (Predicted using BindingDB); Table S5: Gene Ontology (GO) functional enrichment analysis: regulated molecular functions of 16 protein targets modulated by the phytocompounds; Table S6: GO process enrichment analysis: molecular processes of 16 protein targets modulated by phytocompounds. Table S7: Pathway enrichment analysis: 16 protein targets modulated by 139 pathways; Table S8: Role of 16 protein targets in the pathogenesis of hepatitis B viral infection and hepatocellular carcinoma (HCC); Table S9: Interaction of phytocompounds with protein targets and pathways and Edge Count w.r.t. protein target. Figure S1: Venn diagram representing the modulated protein targets involved in HBV-induced HCC by the phytocompounds from *Andrographis paniculata* and *Thespesia populnea*; Figure S2. WT- EGFR model 30. (a) Ramachandran plot for amino acid distribution (b) PyMOL superimposed (Red: Template (PDB file 4I23) and Blue: Model 30), (c) Clustal W template and model 30 alignment; Figure S3. WT-EGFR model-30 RMSD plot showing the deviation and minimum energy conformation observed at 45,252.0 thps (−129,809.51 kcal/mol).

Author Contributions: Methodology, Validation, Formal analysis, Investigation, Data curation, and Writing—original draft, V.S.P.; Conceptualization, Resources, Writing—original draft, Writing—review and editing, Supervision, Project administration, and Funding acquisition, D.R.H.; Formal analysis, Co-supervision, Writing—review and editing, U.V.; Data curation, review and editing, S.H.D. and P.K.; Resources, Writing—review and editing, and Funding acquisition, H.V.H. and S.R. Supervision and Writing—review and editing, S.S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Indian Council of Medical Research, Department of Health Research, New Delhi (ICMR grant no. ISRM/12(61)2019) and ICMR-National Institute of Traditional Medicine, Belagavi, India.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors confirm that the data supporting the findings of this study are available within the article and/or its Supplementary Materials.

Acknowledgments: The authors gratefully acknowledge the extramural funding of the Indian Council of Medical Research, Department of Health Research, New Delhi, and the Bioinformatics center of ICMR-NITM for providing resources. V.S.P. is thankful to the KLE Academy of Higher Education and Research, Belagavi, India, for supporting his studies.

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

References

1. Liu, C.; Yang, S.; Wang, K.; Bao, X.; Liu, Y.; Zhou, S. Alkaloids from Traditional Chinese Medicine against Hepatocellular Carcinoma. *Biomed. Pharmacother.* **2019**, *120*, 109543. [[CrossRef](#)] [[PubMed](#)]
2. Torresi, J.; Tran, B.M.; Christiansen, D.; Earnest-Silveira, L.; Schwab, R.H.M.; Vincan, E. HBV-Related Hepatocarcinogenesis: The Role of Signalling Pathways and Innovative Ex Vivo Research Models. *BMC Cancer* **2019**, *19*, 707. [[CrossRef](#)] [[PubMed](#)]
3. Niu, B.; Hann, H. Hepatitis B Virus-Related Hepatocellular Hepatitis Hepatocellular Carcinoma: Carcinoma: Carcinogenesis, Prevention, and Treatment. *Update Liver Cancer* **2017**, *13*, 986.
4. Reddy, D.C.S. Elimination of Viral Hepatitis: Evolution and India's Response. *Indian J. Public Health* **2019**, *63*, 275–276. [[CrossRef](#)] [[PubMed](#)]
5. Wang, Z.; Li, Z.; Ye, Y.; Xie, L.; Li, W. Oxidative Stress and Liver Cancer: Etiology and Therapeutic Targets. *Oxid. Med. Cell. Longev.* **2016**, *2016*, 7891574. [[CrossRef](#)]
6. Tarocchi, M.; Polvani, S.; Marroncini, G.; Galli, A. Molecular Mechanism of Hepatitis B Virus-Induced Hepatocarcinogenesis. *World J. Gastroenterol.* **2014**, *20*, 11630–11640. [[CrossRef](#)]
7. Ayub, A.; Ashfaq, U.A.; Haque, A. HBV Induced HCC: Major Risk Factors from Genetic to Molecular Level. *Biomed. Res. Int.* **2013**, *2013*, 810461. [[CrossRef](#)]
8. He, G.; Karin, M. NF- κ B and STAT3—Key Players in Liver Inflammation and Cancer. *Cell Res.* **2010**, *21*, 159–168. [[CrossRef](#)]
9. Fan, Y.; Mao, R.; Yang, J. NF- κ B and STAT3 Signaling Pathways Collaboratively Link Inflammation to Cancer. *Protein Cell* **2013**, *4*, 176–185. [[CrossRef](#)]
10. Ahn, J.Y.; Chung, E.Y.; Kwun, H.J.; Jang, K.L. Transcriptional Repression of P21 Waf1 Promoter by Hepatitis B Virus X Protein via a P53-Independent Pathway. *Gene* **2001**, *275*, 163–168. [[CrossRef](#)]
11. Bouchard, M.; Giannakopoulos, S.; Wang, E.H.; Tanese, N.; Schneider, R.J. Hepatitis B Virus HBx Protein Activation of Cyclin A–Cyclin-Dependent Kinase 2 Complexes and G 1 Transit via a Src Kinase Pathway. *J. Virol.* **2001**, *75*, 4247–4257. [[CrossRef](#)] [[PubMed](#)]
12. Leach, J.K.; Qiao, L.; Fang, Y.; Ly, S.; Gilfor, D.; Fisher, P.B.; Grant, S.; Hylemon, P.B.; Peterson, D.; Dent, P. Regulation of P21 and P27 Expression by the Hepatitis B Virus X Protein and the Alternate Initiation Site X Proteins, AUG2 and AUG3. *J. Gastroenterol. Hepatol.* **2003**, *18*, 376–385. [[CrossRef](#)] [[PubMed](#)]
13. Mukherji, A.; Janbandhu, V.C.; Kumar, V. HBx-Dependent Cell Cycle Deregulation Involves Interaction with Cyclin E/A–Cdk2 Complex and Destabilization of P27 Kip1. *Biochem. J.* **2007**, *256*, 247–256. [[CrossRef](#)] [[PubMed](#)]
14. Gearhart, T.L.; Bouchard, M.J. The Hepatitis B Virus HBx Protein Modulates Cell Cycle Regulatory Proteins in Cultured Primary Human Hepatocytes. *Virus Res.* **2011**, *155*, 363–367. [[CrossRef](#)]
15. Slagle, B.L.; Bouchard, M.J. Hepatitis B Virus X and Regulation of Viral Gene Expression. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, a021402. [[CrossRef](#)]
16. Alqahtani, A.; Khan, Z.; Alloghbi, A.; Said Ahmed, T.S.; Ashraf, M.; Hammouda, D.M. Hepatocellular carcinoma: Molecular mechanisms and targeted therapies. *Medicina* **2019**, *55*, 526. [[CrossRef](#)]

17. Abdel-Hamid, N.M.; Nazmy, M.H.; Mahmoud, A.W.; Fawzy, M.A. A Survey on Herbal Management of Hepatocellular Carcinoma. *World J. Hepatol.* **2011**, *3*, 175–183. [\[CrossRef\]](#)
18. Hopkins, A.L. Network Pharmacology: The next Paradigm in Drug Discovery. *Nat. Chem. Biol.* **2008**, *4*, 682–690. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Upadhyay, V.; Hegde, H.V.; Mesta, D.; Belchad, S.; Hampannavar, V.; Kholkute, S.D. Digital Database on Ethno-Medicinal Plants of Western Ghats. *Curr. Sci.* **2010**, *99*, 1658–1659.
20. Dr. Duke's Phytochemical and Ethnobotanical Databases Home Page. Available online: <https://phytochem.nal.usda.gov/phytochem/search> (accessed on 7 March 2021).
21. Degtyarenko, K.; de Matos, P.; Ennis, M.; Hastings, J.; Zbinden, M.; McNaught, A.; Alcántara, R.; Darsow, M.; Guedj, M.; Ashburner, M. ChEBI: A Database and Ontology for Chemical Entities of Biological Interest. *Nucleic Acids Res.* **2008**, *36*, D344–D350. [\[CrossRef\]](#)
22. Molsoft, L.C.C. Available online: <https://molsoft.com/mprop/> (accessed on 15 March 2021).
23. Ivanov, S.M.; Lagunin, A.A.; Rudik, A.V.; Filimonov, D.A.; Poroikov, V.V. ADVERPred-Web Service for Prediction of Adverse Effects of Drugs. *J. Chem. Inf. Model.* **2018**, *58*, 8–11. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R.N.; Gilson, M.K. BindingDB: A Web-Accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201. [\[CrossRef\]](#)
25. UniProt. Available online: <https://www.uniprot.org/> (accessed on 9 April 2021).
26. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [\[CrossRef\]](#)
28. Otasek, D.; Morris, J.H.; Bouças, J.; Pico, A.R.; Demchak, B. Cytoscape Automation: Empowering Workflow-Based Network Analysis. *Genome Biol.* **2019**, *20*, 185. [\[CrossRef\]](#)
29. Patil, V.S.; Khatib, N.A. Triterpene Saponins from *Barringtonia acutangula* (L.) Gaertn as a Potent Inhibitor of 11 β -HSD1 for Type 2 Diabetes Mellitus, Obesity, and Metabolic Syndrome. *Clin. Phytosci.* **2020**, *6*, 61. [\[CrossRef\]](#)
30. Gajiwala, K.S.; Feng, J.; Ferre, R.; Ryan, K.; Brodsky, O.; Weinrich, S.; Kath, J.C.; Stewart, A. Insights into the aberrant activity of mutant EGFR kinase domain and drug recognition. *Structure.* **2013**, *21*, 209–219. [\[CrossRef\]](#)
31. Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinform.* **2003**, *1*, 2–3. [\[CrossRef\]](#)
32. Eswar, N.; Webb, B.; Marti-Renom, M.A.; Madhusudhan, M.S.; Eramian, D.; Shen, M.; Pieper, U.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Protein Sci.* **2007**, *50*, 1–31. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Hollingsworth, S.A.; Dror, R.O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129–1143. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Bowers, K.J.; Chow, D.E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D.; et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In Proceedings of the IEEE ACM/IEEE SC 2006 Conference (SC'06), Tampa, FL, USA, 11–17 November 2006; p. 43.
35. Samdani, A.; Vetrivel, U. POAP: A GNU Parallel Based Multithreaded Pipeline of Open Babel and AutoDock Suite for Boosted High Throughput Virtual Screening. *Comput. Biol. Chem.* **2018**, *74*, 39–48. [\[CrossRef\]](#)
36. Spicer, J.F.; Rudman, S.M. EGFR Inhibitors in Non-Small Cell Lung Cancer (NSCLC): The Emerging Role of the Dual Irreversible EGFR/HER2 Inhibitor BIBW 2992. *Target. Oncol.* **2010**, *5*, 245–255. [\[CrossRef\]](#)
37. Westphal, M.; Maire, C.L.; Lamszus, K. EGFR as a Target for Glioblastoma Treatment: An Unfulfilled Promise. *CNS Drugs* **2017**, *31*, 723–735. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Maennling, A.E.; Tur, M.K.; Niebert, M.; Klockenbring, T.; Zeppernick, F.; Gattenlöhner, S.; Meinhold-Heerlein, I.; Hussain, A.F. Molecular targeting therapy against EGFR family in breast cancer: Progress and future potentials. *Cancers* **2019**, *11*, 1826. [\[CrossRef\]](#)
39. Berasain, C.; Avila, M.A. The EGFR Signalling System in the Liver: From Hepatoprotection to Hepatocarcinogenesis. *J. Gastroenterol.* **2014**, *49*, 9–23. [\[CrossRef\]](#)
40. Komposch, K.; Sibilila, M. EGFR Signaling in Liver Diseases. *Int. J. Mol. Sci.* **2015**, *17*, 30. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Lu, C.; Fu, W.; Zhou, R.; Hu, W. Network Pharmacology-Based Study on the Mechanism of Yiganling Capsule in Hepatitis B Treatment. *BMC Complement. Med. Ther.* **2020**, *20*, 37. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Dandri, M.; Petersen, J. Mechanism of Hepatitis B Virus Persistence in Hepatocytes and Its Carcinogenic Potential. *Clin. Infect. Dis.* **2016**, *62*, 281–288. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Kayaaslan, B.; Guner, R. Adverse effects of oral antiviral therapy in chronic hepatitis B. *World J. Hepatol.* **2017**, *9*, 227. [\[CrossRef\]](#)
44. Li, Y.; Martin, R.C.G., II. Herbal Medicine and Hepatocellular Carcinoma: Applications and Challenges. *Evid. Based Complementary Altern. Med.* **2011**, *2011*, 541209. [\[CrossRef\]](#)
45. Alshammari, T.M. Drug Safety: The Concept, Inception and Its Importance in Patients' Health. *Saudi Pharm. J.* **2016**, *24*, 405–412. [\[CrossRef\]](#)
46. Raschi, E.; Ponti, F. De Drug- and Herb-Induced Liver Injury: Progress, Current Challenges and Emerging Signals of Post-Marketing Risk. *World J. Hepatol.* **2015**, *7*, 1761–1771. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Stickel, F.; Patsenker, E.; Schuppan, D. Herbal Hepatotoxicity. *J. Hepatol.* **2005**, *43*, 901–910. [\[CrossRef\]](#) [\[PubMed\]](#)

48. Chen, S.; Jiang, H.; Cao, Y.; Wang, Y.; Hu, Z.; Zhu, Z.; Chai, Y. Drug Target Identification Using Network Analysis: Taking Active Components in Sini Decoction as an Example. *Sci. Rep.* **2016**, *6*, 24245. [\[CrossRef\]](#)
49. Todorov, D.; Hinkov, A.; Shishkova, K.; Shishkov, S. Antiviral Potential of Bulgarian Medicinal Plants. *Phytochem. Rev.* **2014**, *13*, 525–538. [\[CrossRef\]](#)
50. Herrscher, C.; Roingeard, P.; Blanchard, E. Hepatitis B Virus Entry into Cells. *Cells* **2020**, *9*, 1486. [\[CrossRef\]](#)
51. Iwamoto, M.; Saso, W.; Nishioka, K.; Ohashi, H.; Sugiyama, R.; Ryo, A.; Ohki, M.; Yun, J.H.; Park, S.Y.; Ohshima, T.; et al. The Machinery for Endocytosis of Epidermal Growth Factor Receptor Coordinates the Transport of Incoming Hepatitis B Virus to the Endosomal Network. *J. Biol. Chem.* **2020**, *295*, 800–807. [\[CrossRef\]](#)
52. Iwamoto, M.; Saso, W.; Sugiyama, R.; Ishii, K.; Ohki, M.; Nagamori, S.; Suzuki, R.; Aizaki, H.; Ryo, A.; Yun, J.H.; et al. Epidermal Growth Factor Receptor Is a Host-Entry Cofactor Triggering Hepatitis B Virus Internalization. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 8487–8492. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Chen, S.W.; Himeno, M.; Kouji, Y.; Sugiyama, M.; Nishitsuji, H.; Mizokami, M.; Shimotohno, K.; Miyajima, A.; Kido, T. Modulation of Hepatitis B Virus Infection by Epidermal Growth Factor Secreted from Liver Sinusoidal Endothelial Cells. *Sci. Rep.* **2020**, *10*, 14349. [\[CrossRef\]](#)
54. Schiffer, E.; Housset, C.; Cacheux, W.; Wendum, D.; Desbois-Mouthon, C.; Rey, C.; Clergue, F.; Poupon, R.; Barbu, V.; Rosmorduc, O. Gefitinib, an EGFR Inhibitor, Prevents Hepatocellular Carcinoma Development in the Rat Liver with Cirrhosis. *Hepatology* **2005**, *41*, 307–314. [\[CrossRef\]](#)
55. Zheng, H.; Yang, Y.; Hong, Y.G.; Wang, M.C.; Yuan, S.X.; Wang, Z.G.; Bi, F.R.; Hao, L.Q.; Yan, H.L.; Zhou, W.P. Tropomodulin 3 Modulates EGFR-PI3K-AKT Signaling to Drive Hepatocellular Carcinoma Metastasis. *Mol. Carcinog.* **2019**, *58*, 1897–1907. [\[CrossRef\]](#)
56. Zhangyuan, G.; Wang, F.; Zhang, H.; Jiang, R.; Tao, X.; Yu, D.; Jin, K.; Yu, W.; Liu, Y.; Yin, Y.; et al. VersicanV1 Promotes Proliferation and Metastasis of Hepatocellular Carcinoma through the Activation of EGFR-PI3K-AKT Pathway. *Oncogene* **2020**, *39*, 1213–1230. [\[CrossRef\]](#)
57. Liu, Z.; Ning, F.; Cai, Y.; Sheng, H.; Zheng, R.; Yin, X.; Lu, Z.; Su, L.; Chen, X.; Zeng, C.; et al. The EGFR-P38 MAPK Axis up-Regulates PD-L1 through MiR-675-5p and down-Regulates HLA-ABC via Hexokinase-2 in Hepatocellular Carcinoma Cells. *Cancer Commun.* **2021**, *41*, 62–78. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Ghosh, N.; Ghosh, R.; Mandal, V.; Mandal, S.C. Recent Advances in Herbal Medicine for Treatment of Liver Diseases. *Pharm. Biol.* **2011**, *49*, 970–988. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Jayakumar, T.; Hsieh, C.; Lee, J.; Sheu, J. Experimental and Clinical Pharmacology of *Andrographis paniculata* and Its Major Bioactive Phytoconstituent Andrographolide. *Evid. Based Complement. Altern. Med.* **2013**, *2013*, 846740. [\[CrossRef\]](#)
60. Jadhav, A.K.; Karuppayil, S.M. *Andrographis paniculata* (Burm. F) Wall Ex Nees: Antiviral Properties. *Phyther. Res.* **2021**, *35*, 5365–5373. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Tan, Y.; Chiow, K.; Huang, D.; Wong, S. Andrographolide Regulates Epidermal Growth Factor Receptor and Transferrin Receptor Trafficking in Epidermoid Carcinoma (A-431) Cells: Research Paper. *Br. J. Pharmacol.* **2010**, *159*, 1497–1510. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Kapil, A.; Koul, I.B.; Banerjee, S.K.; Gupta, B.D. Antihepatotoxic Effects of Major Diterpenoid Constituents of *Andrographis paniculata*. *Biochem. Pharmacol.* **1993**, *46*, 182–185. [\[CrossRef\]](#)
63. Handa, S.S.; Sharma, A. Hepatoprotective Activity of Andrographolide from *Andrographis paniculata* against Carbontetrachloride. *Indian J. Med. Res.* **1990**, *92*, 276–283.
64. Handa, S.S.; Sharma, A. Hepatoprotective Activity of Andrographolide against Galactosamine & Paracetamol Intoxication in Rats. *Indian J. Med. Res.* **1990**, *92*, 284–292. [\[PubMed\]](#)
65. Batkhuu, J.; Hattori, K.; Takano, F.; Fushiya, S.; Oshiman, K.; Fujimiya, Y. Suppression of NO Production in Activated Macrophages In Vitro and Ex Vivo by Neoandrographolide Isolated from *Andrographis paniculata*. *Biol. Pharm. Bull.* **2002**, *25*, 1169–1174. [\[CrossRef\]](#)
66. Chen, H.; Ma, Y.B.; Huang, X.Y.; Geng, C.A.; Zhao, Y.; Wang, L.J.; Guo, R.H.; Liang, W.J.; Zhang, X.M.; Chen, J.J. Synthesis, structure–activity relationships and biological evaluation of dehydroandrographolide and andrographolide derivatives as novel anti-hepatitis B virus agents. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2353–2359. [\[CrossRef\]](#)
67. Liu, J.; Wang, Y.-T.; Ji, L.-L. Inhibitory Effects of Neoandrographolide on Nitric Oxide and Prostaglandin E 2 Production in LPS-Stimulated Murine Macrophage. *Mol. Cell Biochem.* **2007**, *2*, 49–57. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Yuvaraj, P.; Subramoniam, A. Short Communication Hepatoprotective Property of *Thespesia populnea* against Carbon Tetrachloride Induced Liver Damage in Rats. *J. Basic Clin. Physiol. Pharmacol.* **2009**, *20*, 169–178. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Jo, S.; Kim, S.; Shin, D.H.; Kim, M.; Kim, M. Inhibition of SARS-CoV 3CL Protease by Flavonoids Inhibition of SARS-CoV 3CL Protease by Flavonoids. *J. Enzyme Inhib. Med. Chem.* **2020**, *35*, 145–151. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Behbahani, M.; Sayedipour, S.; Pourazar, A.; Shanehsazzadeh, M. In Vitro Anti-HIV-1 Activities of Kaempferol and Kaempferol-7-Oglucoside Isolated from *Securigera securidaca*. *Res. Pharm. Sci.* **2014**, *9*, 463–469. [\[PubMed\]](#)
71. Arthanari, S.; Renukadevi, P.; Vanitha, J. Evaluation of Antiviral and Cytotoxic Activities of Methanolic Extract of *Thespesia populnea* (Malvaceae) Flowers. *J. Herbs Spices Med. Plants* **2011**, *17*, 34–37. [\[CrossRef\]](#)