Pharmaceutical Sciences and Drug Design
ISSN: 3062-4428
2023, Volume 3, Page No: 312-328

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Uncovering Anticancer Bioactives from Potentilla nepalensis: An Integrated Network Pharmacology, Molecular Docking, Molecular Dynamics, and MMGBSA-Based Approach to Therapeutic Target Identification

Sanne de Vries[1], Mark van Dijk[2]*, Petra Jansen[1]

[1]Department of Pharmaceutical Sciences, Faculty of Science, Utrecht University, Utrecht, Netherlands.
[2]Department of Drug Design, Faculty of Pharmacy, University of Copenhagen, Copenhagen, Denmark.

***E-mail** ✉ mark.vandijk@gmail.com

Received: 29 May 2023, **Revised:** 04 November 2023, **Accepted:** 09 November 2023

ABSTRACT

Potentilla nepalensis, a member of the Rosaceae family, exhibits a wide range of medicinal properties as an effective herbal remedy. Previous investigations have characterized forty phytochemical compounds (PCs) isolated from root and stem fractions using n-hexane (NR and NS) as well as methanolic (MR and MS) solvents. Nonetheless, the specific influence of these PCs on human genetic elements and their physiological functions remain unexplored until now. This research utilized approaches including network pharmacology, molecular docking, molecular dynamics simulations (MDSs), and MMGBSA techniques. SMILES structures of the PCs retrieved from PubChem served as inputs for DIGEP-Pred, resulting in the detection of 764 target genes. Enrichment analyses of these genes provided insights into their ontological classifications, associated biological pathways, linked disorders, and relevant pharmaceuticals. Protein-protein interaction (PPI) networks generated via String DB, along with topological analysis conducted in Cytoscape version 3.10, highlighted three key targets: TP53 associated with genes induced by MS, NR, and NS extracts, and HSPCB as well as Nf-kB1 linked to MR-induced genes. Among the 40 PCs evaluated, compounds 1b (from MR) and 2a (from MS) displayed superior docking energies (kcal/mol) against the p53 protein at −8.6 and −8.0, respectively, whereas compounds 3a (NR), 4a, and 4c (NS) showed strong affinities with the HSP protein at −9.6, −8.7, and −8.2. Analyses through MDS and MMGBSA confirmed the stability of these ligand-protein complexes, characterized by minimal structural fluctuations and favorable binding free energies. The targets uncovered in this work play significant roles in various malignancies. Therefore, additional experimental validation via in vitro and in vivo assays is recommended to elucidate the precise molecular actions and underlying mechanisms of these targets across multiple cancer models, incorporating the PCs derived from P. nepalensis.

Keywords: TP53, HSPCB, Nf-kB1, Potentilla nepalensis, Network pharmacology, Computational studies

How to Cite This Article: de Vries S, van Dijk M, Jansen P. Uncovering Anticancer Bioactives from Potentilla nepalensis: An Integrated Network Pharmacology, Molecular Docking, Molecular Dynamics, and MMGBSA-Based Approach to Therapeutic Target Identification. Pharm Sci Drug Des. 2023,3:312-28. https://doi.org/10.51847/TAR5F7QBDt

Introduction

Herbal remedies have played a pivotal role in the evolution of contemporary medicine, serving as foundational sources for numerous effective treatments against various illnesses and underscoring the value of botanical compounds in pharmaceutical innovation [1]. Additionally, phytochemicals obtained from botanical origins provide benefits such as enhanced availability and reduced adverse effects relative to chemically synthesized medications. Incorporating plant-based therapeutic agents into modern drug design offers potential solutions for combating issues like microbial resistance to antibiotics [2].

Known commonly as Nepal cinquefoil, Potentilla nepalensis is a species within the Rosaceae family, originating from the Himalayan areas of Nepal and Tibet, where it thrives in high-altitude alpine and subalpine zones.

Although frequently grown for its attractive blooms as a decorative species, P. nepalensis is highly regarded for its health-promoting attributes. These benefits are largely attributed to its rich profile of secondary metabolites, encompassing elevated levels of phenolic compounds, flavonoids, and terpenoids, which underpin its pharmacological effects [3].

Traditionally, P. nepalensis has been applied in diverse healing practices. It has been used for promoting wound recovery and managing dermatological issues, alongside supporting digestive health and intestinal function. Its anticarcinogenic potential has also been noted [4], complemented by antioxidant capabilities that help protect cellular structures from damage caused by oxidation. Furthermore, the species demonstrates anti-inflammatory and pain-relieving effects, making it useful for reducing discomfort and inflammatory responses [3]. In Tibetan medicinal traditions, certain Potentilla species, including P. nepalensis, have been employed to address conditions like asthma, migraines, dysentery, and respiratory infections. Extracts from the roots of P. nepalensis have shown encouraging anticancer [5] and antibacterial properties [6].

The current investigation focuses on n-hexane and methanolic fractions derived from the root and stem segments of P. nepalensis, as analyzed previously via Gas Chromatography–Mass Spectrometry (GC-MS) [7]. The objective is to clarify how these fractions interact with human genetic targets by applying a combined strategy involving network pharmacology, molecular docking, and simulation-based dynamics.

This research involved a series of methodical procedures. Initially, relevant data on the phytochemical compounds (PCs) were obtained from the PubChem repository. These compounds were then screened to pinpoint target genes with notable pharmacological probabilities exceeding 0.8. Next, protein–protein interaction networks were built for each fraction using String DB, with emphasis on determining the central genes in those networks. Molecular docking was subsequently carried out with AutoDock Vina. Essential proteins from the networks were sourced from the Protein Data Bank (PDB). Finally, molecular dynamics simulations (MDSs) and MMGBSA binding free energy computations were performed on complexes demonstrating high docking affinities. A schematic overview of the methods used in this study is illustrated in **Figure 1.**

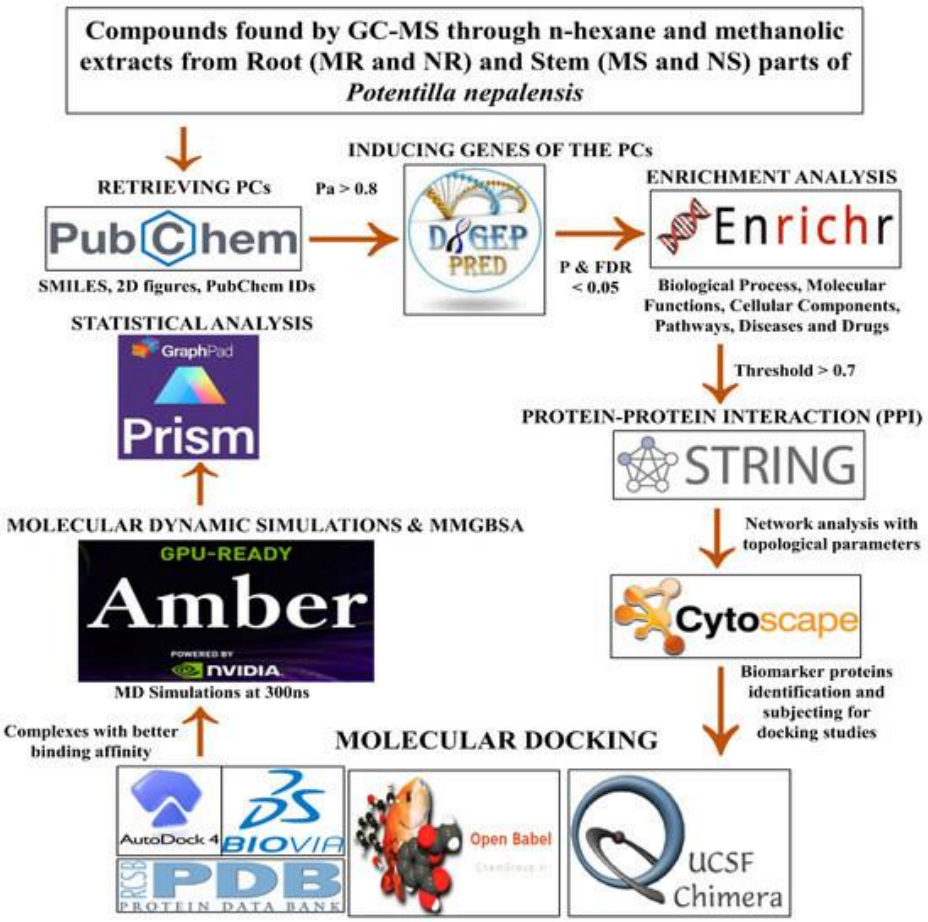


Figure 1. Flowchart illustrating the sequential computational strategies employed in this research.

Materials and Methods

Compound retrieval

PubChem identifiers (IDs), two-dimensional structures, and Simplified Molecular Input Line Entry System (SMILES) notations for the PCs identified in P. nepalensis were obtained from the PubChem database https://pubchem.ncbi.nlm.nih.gov/ (accessed 5 September 2023) [8]. These compounds were previously characterized from root and stem methanolic and n-hexane extracts using GC-MS [7]. PubChem serves as a comprehensive, open-access repository of chemical compounds with detailed annotations.

Prediction of PC-induced genes

SMILES strings of the PCs were submitted to the DIGEP-Pred web server http://www.way2drug.com/GE (accessed 10 September 2023) [9] to predict induced human genes, applying a pharmacological activity probability cutoff of Pa > 0.8. This server leverages the Prediction of Activity Spectra for Substances (PASS) algorithm, employing leave-one-out cross-validation and integrating both mRNA and protein expression datasets.

Functional enrichment analysis

Enrichment profiling of PC-induced human genes was executed via the Enrichr platform https://maayanlab.cloud/Enrichr/ (accessed 20 September 2023), using thresholds of false discovery rate (FDR) < 0.05 and p-value < 0.05 [10]. Duplicate genes were eliminated to curate a refined input list, enabling identification of associated biological processes, molecular functions, cellular components, pathways, diseases, and approved therapeutics.

Protein–protein interaction networks

Interconnections among proteins encoded by the predicted induced genes were mapped using the STRING database https://string-db.org/ (accessed 28 September 2023) at a high-confidence interaction score threshold of 0.7. Networks were assembled from diverse evidence channels, encompassing text mining, experimental validations, curated databases, co-expression profiles, genomic neighborhood, gene fusion, and phylogenetic co-occurrence [11]. The resultant networks were imported into Cytoscape V3.10.0 [12] and refined using the Analyzer plugin to quantify network topology for the input gene sets. Key metrics evaluated included degree centrality, average shortest path length, clustering coefficient, closeness centrality, and betweenness centrality, providing quantitative measures of gene interdependencies and centrality within the interaction landscapes.

Molecular docking

Binding affinities for the protein–ligand complexes were evaluated using established molecular docking protocols previously described [13]. The three-dimensional structures of the TP53-binding domain (PDB ID: 6MXY) https://www.rcsb.org/structure/6MXY (accessed 5 October 2023), complexed with the native ligand K6M (N-[3-(tert-butylamino)propyl]-3-(trifluoromethyl)benzamide), nuclear factor NF-kappa-B p50 (PDB ID: 3GUT) [14], and heat shock protein 90 (PDB ID: 1UYM) bound to PU3 (9-butyl-8-(3,4,5-trimethoxybenzyl)-9H-purin-6-amine) [15] were downloaded from the RCSB Protein Data Bank https://www.rcsb.org/ (accessed 5 October 2023). These structures were resolved by X-ray crystallography with resolutions of 1.62 Å, 3.59 Å, and 2.45 Å, respectively.

The phytochemical compounds (PCs) in SMILES format from PubChem and the retrieved protein structures required preprocessing prior to docking. Ligands were optimized by adding hydrogens at physiological pH (7.4), generating three-dimensional conformations, and minimizing energy with the MMFF94 force field using Open Babel [16], which supports conversion across multiple chemical file formats. For the proteins, co-crystallized ligands and solvent molecules were deleted using Discovery Studio. Missing atoms were repaired, polar hydrogens were added, and Gasteiger partial charges were computed with AutoDock Tools [17].

Grid box parameters were defined to encompass the active sites accurately, thereby enabling reliable prediction of binding poses and affinities. Centers were positioned based on the coordinates of native ligands or the overall structure: for TP53 with K6M (X = −10.80, Y = 26.77, Z = −3.52), for HSP90 with PU3 (X = 3.60, Y = 11.13, Z = 24.75), and for NFKB (X = 28.80, Y = −23.60, Z = 58.23). Docking calculations were performed using AutoDock Vina [18], ensuring precise evaluation of protein–ligand interactions and binding scores.

Molecular Dynamics Simulations and MMGBSA

Molecular dynamics (MD) simulations were conducted on an Ubuntu 22.04 LTS workstation powered by an Intel Core i7-13700k CPU and an NVIDIA RTX 4080 GPU. All simulations and subsequent trajectory processing were executed with the Amber 22 suite and AmberTools 23 [19]. Docked complexes involving p53 and HSPCB were separated into individual ligand and protein components, with missing hydrogens supplemented via UCSF ChimeraX [20]. Ligand topologies and parameters were generated using the General Amber Force Field (GAFF) [19] through Antechamber, employing AM1-BCC charges. The protein was parameterized with the FF19SB force field [21], solvated in a TIP3P water box, and neutralized with Na^+ and Cl^- ions.

System equilibration involved 30,000 steps of energy minimization, followed by gradual heating to 300 K and density equilibration at 1 atm for 200 ps. Production runs spanned 300 ns using the pmemd.cuda module for GPU acceleration. Binding free energies were computed via the Molecular Mechanics Generalized Born Surface Area (MMGBSA) method [22] implemented in MMPBSA.py. Trajectory visualizations and graphical representations were generated using xmgrace 5.1.25 [23] and R version 4.2.3 [24].

Results and Discussion

Collection of phytochemical compounds

The n-hexane and methanolic extracts derived from the roots and stems of P. nepalensis were previously characterized using Gas Chromatography–Mass Spectrometry (GC-MS). That analysis identified a total of forty distinct phytochemicals, with ten unique compounds detected in each of the four extracts corresponding to different plant parts and solvents. In this work, detailed information for these compounds—including SMILES notations, PubChem identifiers, and two-dimensional structures—was obtained from the PubChem database. The compounds were grouped into four sets according to extraction conditions: 1a–1j from the n-hexane root extract (NR), 2a–2j from the methanolic stem extract (MS), 3a–3j from the n-hexane root extract (NR), and 4a–4j from the n-hexane stem extract (NS).

Genes targeted by the phytoconstituents

A thorough screening for gene interactions identified a total of 764 genes influenced by the phytochemicals. The methanolic root extract (MR) was linked to 149 target genes, while the methanolic stem extract (MS) influenced 217 genes. The n-hexane root extract (NR) affected 277 genes, and the n-hexane stem extract (NS) targeted 121 genes. These data offer a detailed profile of the human genes modulated by the phytoconstituents of P. nepalensis.

Enrichment analysis

To maintain focus and significance, only the ten most predictive terms were chosen from the gene ontology (GO) enrichment results, prioritizing those with the strongest functional relevance. These terms were subsequently summarized to highlight the key characteristics of the most highly enriched genes.

Biological processes (BP) refer to the larger functional roles performed by gene products [25]. Among all extracts, the majority of genes were associated primarily with two biological processes: positive regulation of nucleic acid-templated transcription and regulation of DNA-templated transcription, as detailed in **Table 1**. The methanolic stem extract (MS) showed the greatest degree of involvement in these processes.

Table 1. Gene Ontology and Pathway Enrichment Analysis of Methanolic and n-Hexane Extracts from Root and Stem Tissues of *P. nepalensis*

Category	Methanolic Stem (MS)	Methanolic Root (MR)	n-Hexane Stem (NS)	n-Hexane Root (NR)
Cellular Component (CC)	Nucleus, azurophil granule lumen	Intracellular membrane-bounded organelle, nucleus, secretory granule lumen	Intracellular organelle lumen, endoplasmic reticulum lumen, secretory granule lumen	Intracellular membrane-bounded organelle, nucleus
Biological Process (BP)	Positive regulation of nucleic-acid-templated transcription	Positive regulation of nucleic-acid-templated transcription, regulation of DNA-templated transcription	Positive regulation of nucleic-acid-templated transcription	Positive regulation of nucleic-acid-templated transcription

Molecular Function (MF)	Protein serine/threonine phosphatase activity	Protein homodimerization activity	DNA-binding transcription activator activity, protein serine/threonine phosphatase activity, oxidoreductase activity	DNA binding, protein homodimerization activity
Associated Diseases	Neoplasm metastasis, breast carcinoma, prostate malignant neoplasm	Neoplasm metastasis, liver carcinoma, mammary neoplasms, melanoma	Breast carcinoma, malignant breast neoplasm, neoplasm metastasis	Neoplasm metastasis, liver carcinoma
Enriched Pathways	Nuclear receptors meta-pathway, osteoblast differentiation, vitamin D receptor pathway	VEGFA-VEGFR2 signaling pathway, leptin signaling pathway, microRNAs in cardiomyocyte hypertrophy, B-cell receptor signaling pathway	Common pathways in drug addiction, melanoma, pyrimidine metabolism	Nuclear receptors meta-pathway, vitamin D receptor pathway
Associated Drugs	Trifluoperazine, Pitolisant, Cyproheptadine, Pimozide, Brompheniramine, Buprenorphine, Lidoflazine, Chlorambucil	Aprindine, Domperidone	Stearic acid, Epalrestat, Dodecanoic acid, Gamolenic acid, Vemurafenib, Bezafibrate, Gemfibrozil, Linolenic acid, Aprindine, Eicosapentaenoic acid	Mefenamic acid, Diclofenac, Flufenamic acid, Quercetin, Mezlocillin, Hydrochlorothiazide, Hydroxycarbamide, Bendroflumethiazide, Benzthiazide, Chlorambucil, Bezafibrate, Rosiglitazone, Stearic acid, Dodecanoic acid, Gamolenic acid, Aprindine, Caffeine, Eicosapentaenoic acid, Linolenic acid

Molecular function (MF) describes the specific biochemical activity or role performed by a gene product [25]. A shared molecular function signature was observed among the genes induced by the methanolic root (MR) and n-hexane root (NR) extracts. In particular, protein homodimerization activity and protein serine/threonine phosphatase activity emerged as prominent and overlapping functional categories across these two extracts. Similarly, the methanolic stem (MS) and n-hexane stem (NS) extracts displayed consistent enrichment patterns in their respective target gene sets. Furthermore, when considering the combined gene lists induced by MR+NR and MS+NS extracts, additional MF terms were significantly enriched. Specifically, DNA binding was enriched in the NR-induced gene set, while DNA-binding transcription activator activity, protein serine/threonine phosphatase activity, and oxidoreductase activity were notably enriched in the NS-induced genes **(Table 1)**. These shared and extract-specific functional enrichments highlight promising directions for further mechanistic investigation.

Cellular component (CC)

Cellular component (CC) terms indicate the subcellular localization or compartment where gene products are primarily active [25]. Genes induced by both MR and NR extracts showed substantial overlap, with predominant localization in the intracellular membrane-bounded organelle and nucleus. The MR extract additionally enriched genes associated with the secretory granule lumen. In contrast, genes targeted by the MS extract were mainly localized to the nucleus and azurophil granule lumen. The NS extract induced genes with broader and more diverse localization, including the intracellular organelle lumen, endoplasmic reticulum lumen, and secretory granule lumen **(Table 1)**.

Pathway enrichment

The genes modulated by the MR extract were significantly associated with several biologically relevant signaling cascades, including the VEGFA–VEGFR2 signaling pathway, leptin signaling pathway, microRNA regulation in cardiomyocyte hypertrophy, and B-cell receptor signaling pathway. Both the MS and NR extracts commonly targeted genes involved in the nuclear receptors meta-pathway and the vitamin D receptor pathway, with the MS extract further enriching the osteoblast differentiation pathway. In contrast, genes induced by the NS extract were linked to pathways underlying common drug addiction, melanoma, and pyrimidine metabolism **(Table 1)**.

Disease association

Analysis of disease enrichment revealed a strong and consistent association of the target genes from all four extracts with cancer-related pathologies, particularly neoplasm metastasis. This recurrent theme strongly suggests that the phytoconstituents present in the extracts may influence key processes involved in cancer progression and metastatic spread **(Table 1)**.

Protein–protein interaction (PPI) network analysis

To explore functional relationships among the induced genes, protein–protein interaction networks were constructed using the STRING database. These networks comprise nodes (representing proteins) and edges (representing interactions), with evidence drawn from multiple sources including experimental data, text mining, gene fusion, co-expression, genomic neighborhood, and curated databases. To quantitatively characterize the topological properties of the resulting PPI networks, several network parameters were calculated for each protein node, including degree centrality, average shortest path length, clustering coefficient, closeness centrality, and betweenness centrality **(Table 2)**.

Table 2. Topological Parameters of Key Hub Proteins in the Protein–Protein Interaction Networks Derived from Induced Genes

Extract	Protein	Average Shortest Path Length	Degree Centrality (DC)	Closeness Centrality (C. Cen)	Clustering Coefficient (CC)	Betweenness Centrality (BC)
MR	HSPCB	3.02	15	0.33	0.18	0.18
MR	NFKB1	2.85	15	0.34	0.25	0.33
MS	TP53	1.87	25	0.53	0.15	0.56
NR	TP53	2.32	20	0.42	0.16	0.58
NS	TP53	1.75	25	0.57	0.17	0.59

The topological parameters of the PPI networks provide insight into the structural and functional importance of individual proteins within each extract-induced network. Degree centrality (DC) reflects the number of direct interactions a protein has with others, serving as a measure of its local connectivity. Average shortest path length indicates the typical minimum number of steps required to connect a given protein to any other node in the network. The clustering coefficient (CC) quantifies the tendency of a protein's neighbors to interconnect with one another, often forming tightly knit functional modules. Betweenness centrality (BC) identifies proteins that lie on many shortest paths between other nodes, positioning them as critical bridges or bottlenecks for information flow across the network. Finally, closeness centrality (C. Cen) measures how quickly a protein can reach all other nodes, with higher values indicating greater proximity to the rest of the network.

Based on these topological descriptors, several proteins emerged as highly influential hubs. In the network derived from genes induced by the methanolic root (MR) extract, HSPCB and NFKB1 exhibited prominent connectivity and centrality measures **(Figure 2)**. Across the networks from the methanolic stem (MS), n-hexane root (NR), and n-hexane stem (NS) extracts, TP53 consistently stood out as the dominant hub protein, demonstrating the highest degree centrality, superior closeness and betweenness values, and relatively short average path lengths **(Figures 3-5)**.

In the visualized networks, these key hub proteins are prominently displayed in centrally positioned rectangular nodes with a yellow background and blue font for clear identification. The quantitative topological parameters for these proteins are summarized in **Table 2**, reinforcing their central roles. The elevated centrality scores of HSPCB,

NFKB1, and particularly TP53 underscore their potential as key mediators of the biological effects elicited by the phytochemicals from P. nepalensis. These findings highlight critical nodes within the interaction networks and strengthen the rationale for considering these proteins as promising therapeutic targets modulated by the plant's bioactive constituents. This topological evaluation thus deepens our understanding of the molecular connectivity and functional organization driven by the different extracts at the systems level.

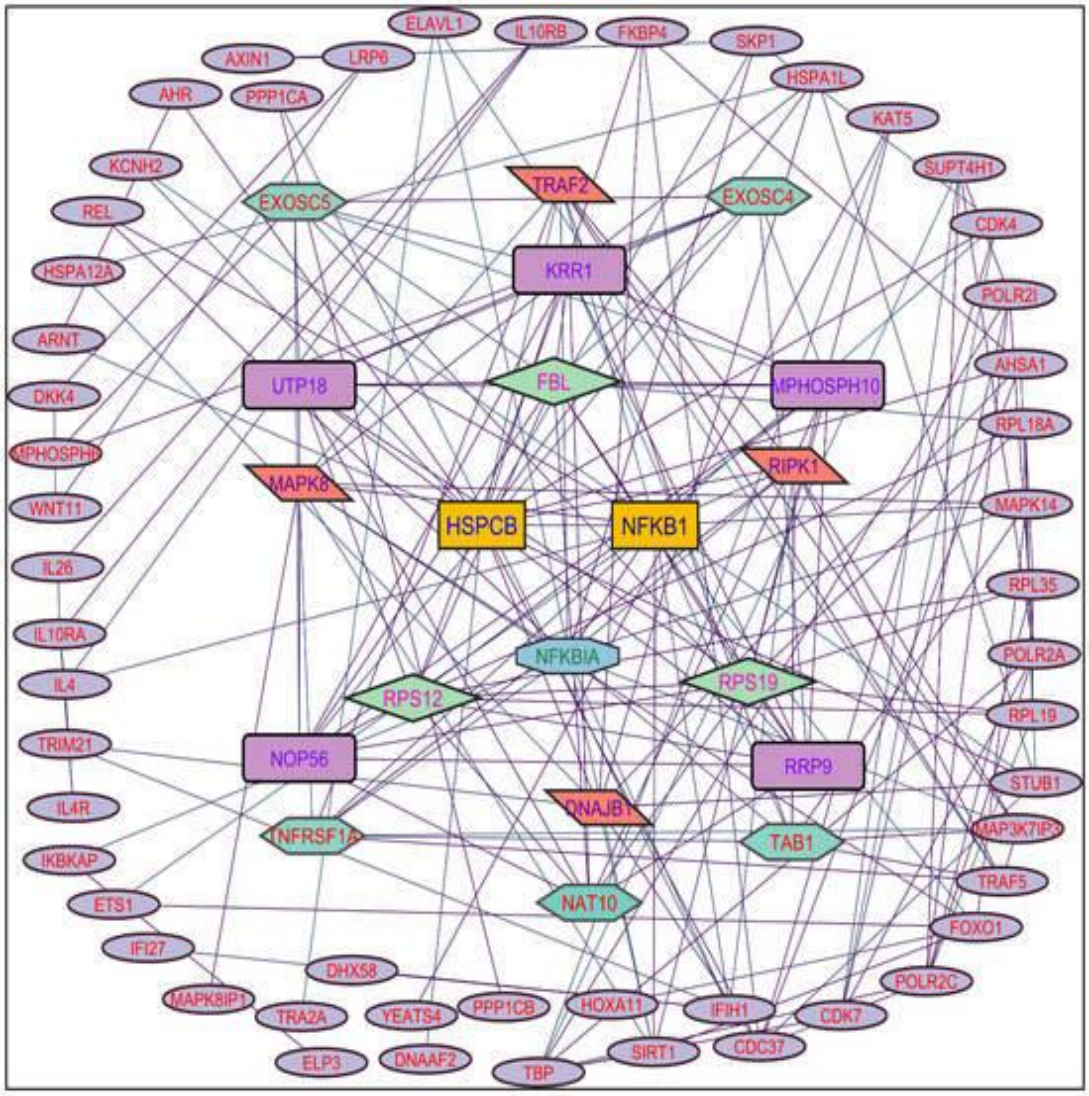


Figure 2. PPI Network Constructed from Genes Induced by the Methanolic Root Extract

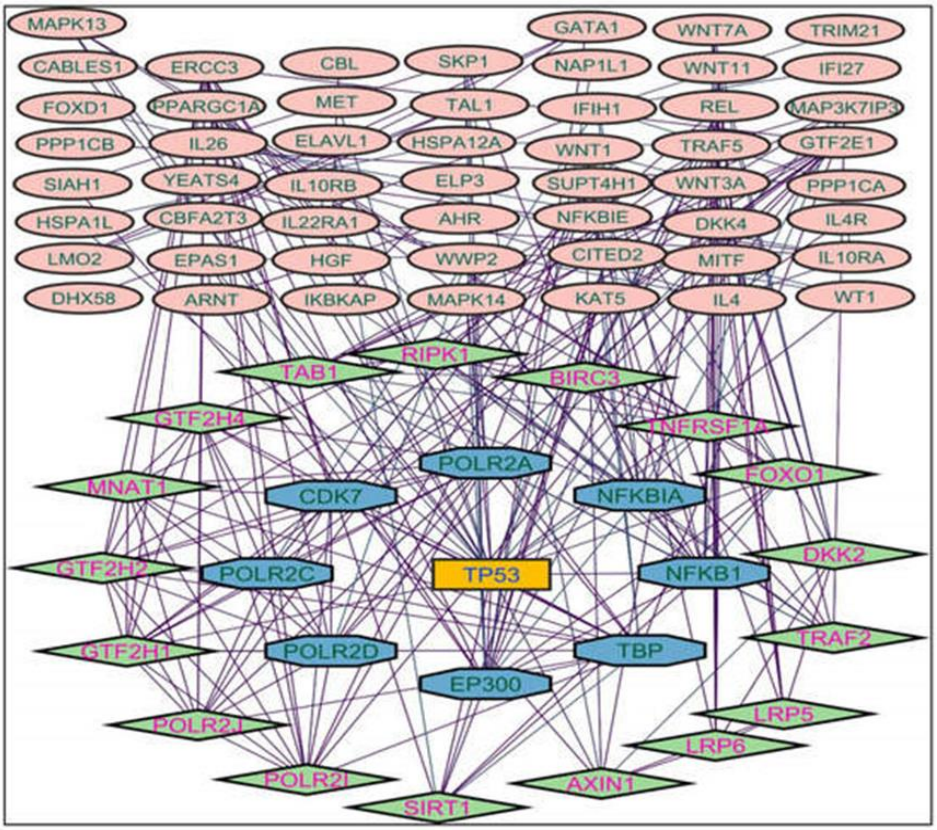


Figure 3. PPI network illustrating interactions between proteins encoded by genes activated by the methanolic extract of shoots.

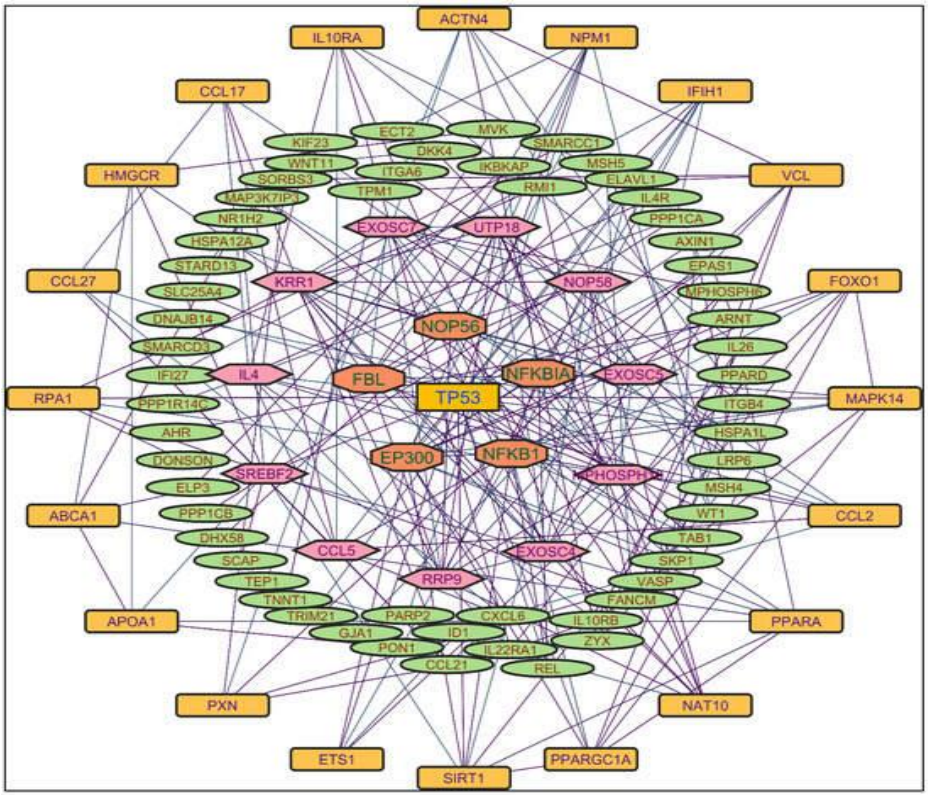


Figure 4. PPI network showing interactions between proteins encoded by genes activated by the n-hexane root extract.

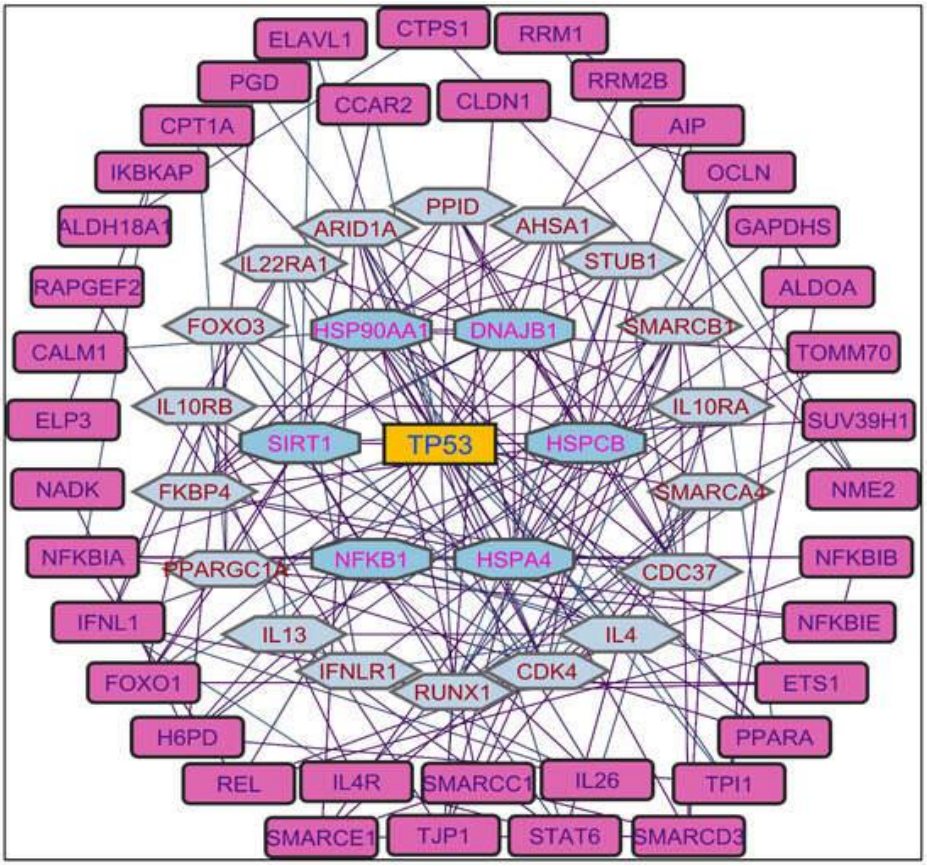


Figure 5. Protein–protein interaction network of genes induced by the n-hexane shoot extract.

Node colors and shapes indicate the degree (number of interactions) for each gene in the network:

- First description: Rectangular nodes with yellow background and blue font represent genes with 15 interactions, octagonal nodes with light blue background and green font indicate 12 interactions, diamond nodes with light green background and pink font denote 11 interactions, round rectangular nodes with purple background and pink font show 10 interactions, parallelogram nodes with orange background and blue font correspond to 9 interactions, hexagonal nodes with light green background and red font represent 8 interactions, elliptical nodes with light blue background and red font indicate fewer than 8 interactions.
- Second description: Rectangular nodes with yellow background and blue font represent genes with 25 interactions, octagonal nodes with light blue background and green font indicate 14–18 interactions, diamond nodes with light green background and pink font denote 8–13 interactions, elliptical nodes with light brown background and green font show fewer than 8 interactions.
- Third description: Rectangular nodes with yellow background and blue font represent genes with 20 interactions, octagonal nodes with brown background and green font indicate 11–15 interactions, hexagonal nodes with light pink background and pink font denote 9–10 interactions, round rectangular nodes with orange background and purple font correspond to 5–8 interactions, elliptical nodes with light green background and red font show fewer than 5 interactions.
- Fourth description: Rectangular nodes with yellow background and blue font represent genes with 25 interactions, octagonal nodes with light blue background and pink font indicate 10–16 interactions, hexagonal nodes with light green background and red font denote 7–9 interactions, rectangular nodes with light pink background and blue font show fewer than 7 interactions.

Molecular docking

Molecular docking studies were performed to investigate the binding interactions between the phytochemical compounds (PCs) and selected therapeutic target proteins using methanolic and n-hexane extracts from the roots and stems of P. nepalensis. The aim was to evaluate the binding affinities of these PCs toward three key proteins: p53 (encoded by the TP53 gene), heat shock protein (encoded by the HSPCB gene), and nuclear factor kappa light

chain (encoded by the NFKB1 gene) **(Table 2)**. Among these compounds, five (1b, 2a, 3a, 4a, and 4c) exhibited notably higher binding affinities with both p53 and heat shock proteins.

Figure 6 and Table 3 provide details on the binding affinities and the specific amino acid residues involved in the interaction types. For instance, compound 1b showed a binding affinity of −8.6 kcal/mol with the p53 protein, mediated by a single conventional hydrogen bond with the residue Aser1503. Furthermore, two alkyl bonds formed, connecting the alkyl ends of 1a with the alkyl groups of BMet1584 in the p53 protein. Seventeen π-alkyl bonds emerged, linking the π-alkyl groups of 4ATrp1495, ATyr1502, 2APhe1519, ATyr1523, 3BTrp1495, 3BTyr1502, 2BPhe1519, and BTyr1523 with the π-orbitals of 1b. PC 2a demonstrated −8.0 kcal/mol of binding affinity with the p53 protein. AMet1584 participated in an alkyl bond formation with alkyl ends. Additionally, a set of twelve π-alkyl bonds connecting the π-alkyl groups of 2ATrp1495, 2ATyr1502, 2APhe1519, 2BTrp1495, BTyr1502, 2BPhe1519, and BTyr1523 with the π-orbitals of 2a.

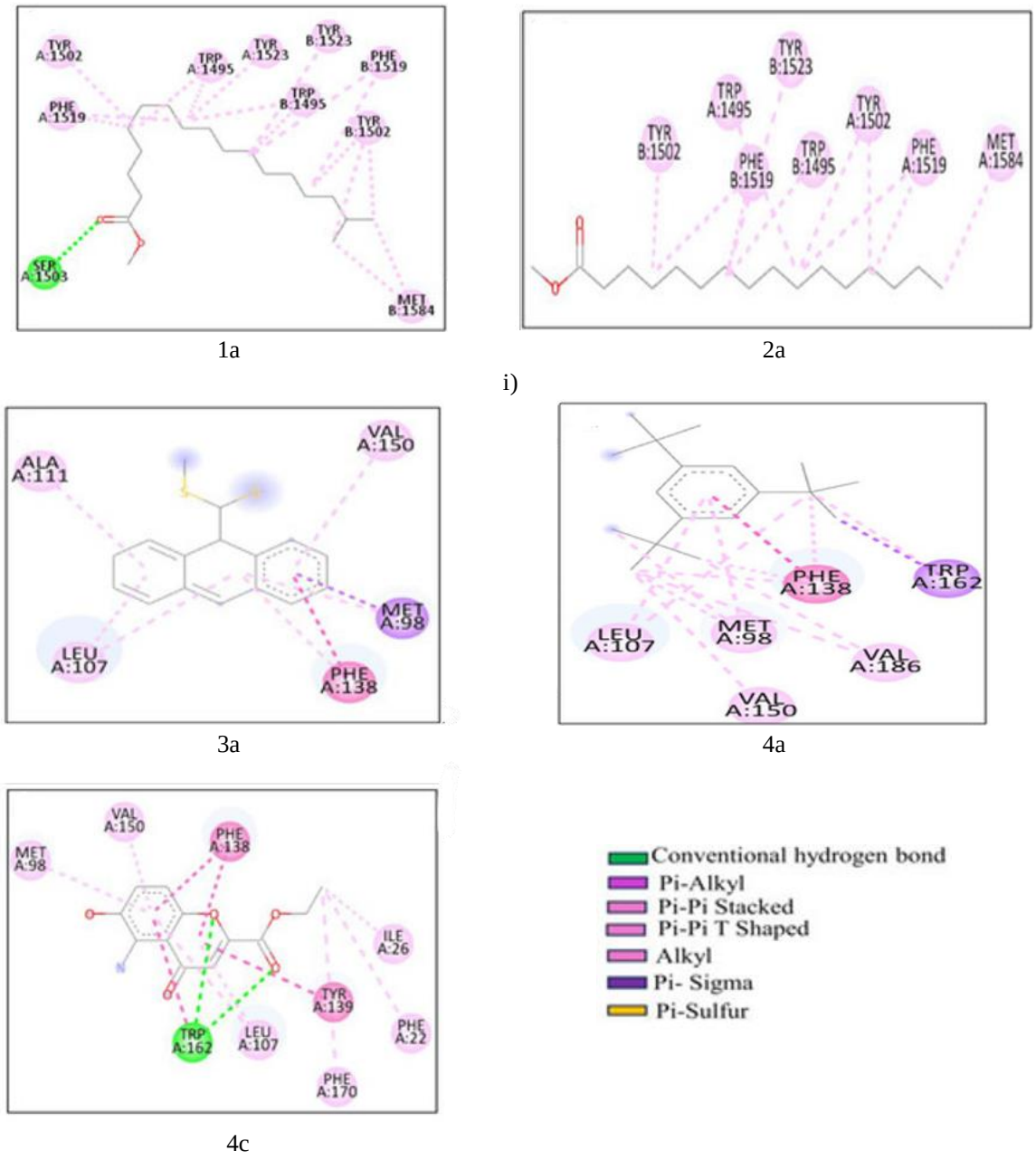


Figure 6. Molecular interactions within the docked complexes. (i) p53 protein in complex with phytochemical compounds 1a and 2a, (ii) heat shock protein (HSP) in complex with phytochemical compounds 3a, 4a, and 4c.

Table 3. Binding affinity (B.A.) and interactions of the heat shock protein (HSP), nuclear factor NF-kappa-BP (NFKB1), and TP53-binding protein (p53) with the extracted PCs of P. nepalensis.

Complex		B.A. (kcal /mol)	Hydrogen Bonds	Hydrophobic Bonds					Other Bonds
Proteins	PCs		CHB	π-Alkyl	Alkyl	π-π Stacked	π-π T Shaped	π-Sigma	
p53	1b	−8.6	ASer1503	^{4}ATrp1495, ATyr1502, ^{2}APhe1519, ATyr1523, ^{3}BTrp1495, ^{3}BTyr1502, ^{2}BPhe1519, BTyr1523	^{2}BMet1584	-	-	-	-
	2a	−8.0	-	^{2}ATrp1495, ^{2}ATyr1502, ^{2}APhe1519, ^{2}BTrp1495, BTyr1502, ^{2}BPhe1519, BTyr1523	AMet1584	-	-	-	-
HSP	3a	−9.6	-	APhe138, Aval150	AMet98, ^{2}ALeu107, AAla111	APhe138	-	AMet98	AMet98
	4a	−8.7	-	^{3}APhe138, ^{2}ATrp162, AMet98, ALeu107	^{2}AVal186, ^{2}AMet98, AVal150, ALeu107	APhe138	-	^{2}Trp162	-
	4c	−8.2	^{2}ATrp162	APhe22, APhe170, ^{2}ALeu107, AMet98, AVal150	AIle26	^{2}APhe138	ATyr139, ^{2}ATrp162	-	-

PCs—phytoconstituents, CHB—conventional hydrogen bond, other bond—sulfur bond.

Compounds 3a, 4a, and 4c demonstrated strong binding affinities to the HSP protein, with docking scores of −9.6, −8.7, and −8.2 kcal/mol, respectively. Compound 3a formed nine distinct interactions with key residues of the HSP protein. Notably, Met98 participated in three types of bonds: a sulfur interaction, a π-sigma bond, and an alkyl interaction with 3a. Additionally, three alkyl interactions occurred between the alkyl groups of Leu107 (two instances) and Ala111 and the ligand 3a. A π-π stacked interaction was established between the aromatic ring of Phe138 and 3a, enhancing complex stability. Two π-alkyl interactions were also observed involving Phe138 and Val150 with the aromatic system of 3a.

For compound 4a, three π-alkyl interactions and one π-π stacked bond involved the Phe138 residue. Furthermore, two π-alkyl and two π-π T-shaped interactions formed with Trp162. Additional bonds included one π-alkyl and two alkyl interactions with Met98, one alkyl bond with Val186, and both a π-alkyl and an alkyl interaction with Leu107.

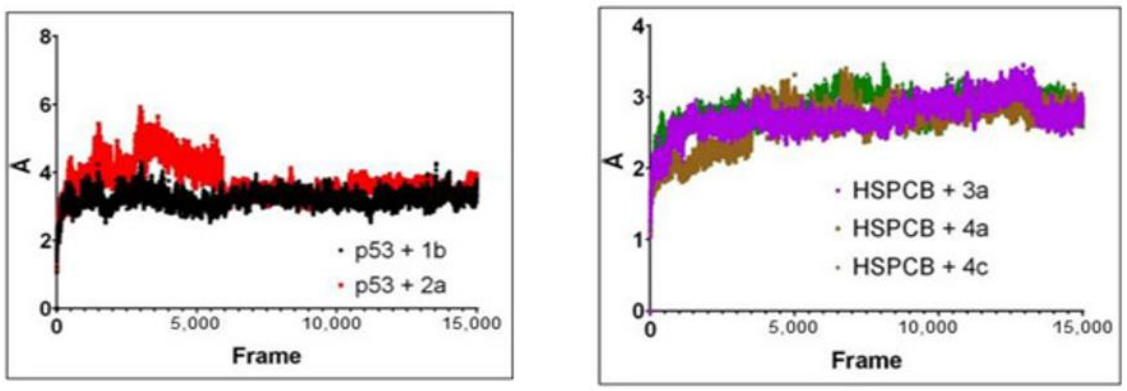
Compound 4c displayed two carbon-hydrogen bonds with Trp162. Six π-alkyl interactions linked the alkyl side chains of Phe22, Phe170, Leu107 (two instances), Met98, and Val150 to the aromatic rings of 4c. An alkyl interaction was also formed with Ile26. Additionally, two π-π stacked bonds involved Phe138 (two instances),

while three π-π T-shaped interactions occurred with Tyr139 and Trp162 (two instances). These findings elucidate the binding patterns and specific interactions of compounds 3a, 4a, and 4c with the HSP protein, suggesting their potential utility in therapeutic development.

Molecular dynamics simulations and MMGBSA analysis

Molecular dynamics (MD) simulations were conducted to evaluate the stability and dynamic behavior of the protein–ligand complexes in an aqueous environment, offering detailed atomic-level insights into their interactions. Based on the top docking poses, five complexes were selected for 300 ns MD simulations followed by MMGBSA calculations: (i) p53 bound to 1b, (ii) p53 bound to 2a, (iii) HSP complexed with 3a, (iv) HSP bound to 4a, and (v) HSP bound to 4c. Each system underwent preprocessing, including energy minimization and equilibration under NVT and NPT conditions. Trajectories were analyzed for root mean square deviation (RMSD), root mean square fluctuation (RMSF), and MMGBSA binding free energies.

The RMSD profiles of the simulated complexes are shown in **Figure 7**. In panel A **(Figure 7a),** comparing the p53 complexes with ligands 1b (p53+1b) and 2a (p53+2a), the p53+2a system initially displayed greater fluctuations than p53+1b. The p53+1b complex exhibited consistent deviations that remained below 5 Å throughout. Both systems achieved stable equilibrium around 6000 frames (approximately 130 ns), with RMSD values stabilizing between 2–3 Å. In both cases, the ligands remained firmly within the binding pocket for the entire simulation duration, indicating robust complex stability.



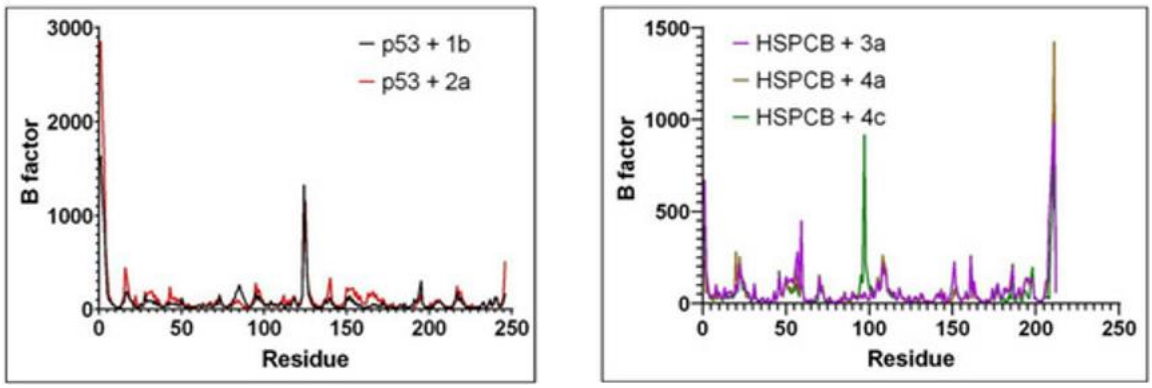
a) b)

Figure 7. Root mean square deviation (RMSD) profiles obtained from the 300 ns molecular dynamics simulations, segmented into 15,000 frames. Panel A illustrates the RMSD trajectories for the p53 protein complexes with ligands 1b (red trace) and 2a (black trace). Panel B shows the corresponding RMSD values for the HSP complexes with ligands 3a (pink), 4a (brown), and 4c (green).

The RMSD data for the HSP complexes—HSP+3a, HSP+4a, and HSP+4c—are depicted in **Figure 7b**. These trajectories demonstrate that conformational equilibrium was reached within the first 100 ns of simulation. Initial fluctuations were observed across all three systems, with deviations temporarily rising early in the runs, however, even in the HSP+4c complex, these did not surpass 3.5 Å. Throughout the production phase, all three protein–ligand systems maintained stable RMSD values in the range of 2–3 Å. This consistent behavior confirms the structural integrity of the complexes, with no evidence of ligand dissociation from the binding pockets over the entire simulation period.

Root mean square fluctuation (RMSF) values, represented as B-factors, were calculated to assess per-residue flexibility, as shown in **Figure 8**. In panel A, the RMSF profiles for the p53+1b and p53+2a complexes were highly comparable, displaying no significant peak differences. This similarity indicates that neither ligand induced substantial differential perturbations at the binding site.

For the HSP complexes **(Figure 8b)**, the RMSF patterns of HSP+3a, HSP+4a, and HSP+4c were largely consistent. A notable exception occurred in the HSP+4c system, where residue 97 exhibited elevated fluctuation. Importantly, this residue is located outside the ligand-binding cavity and therefore does not impact the stability of the bound ligand.



a) b)

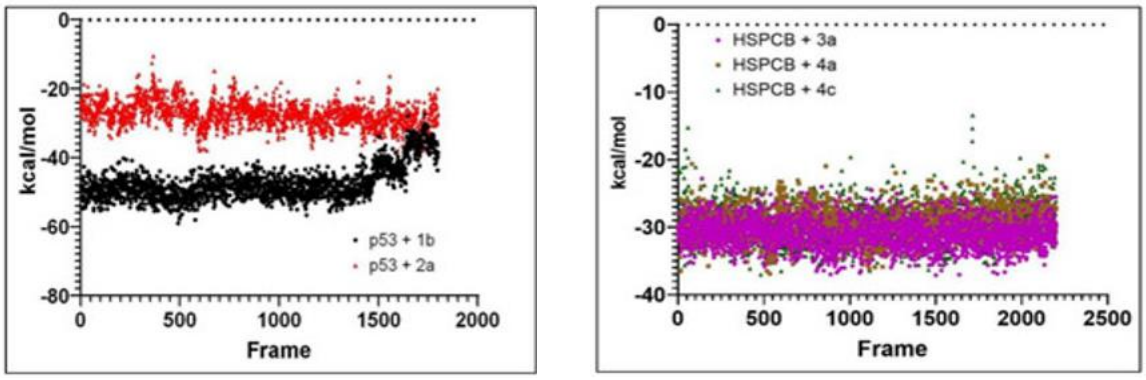
Figure 8. Root mean square fluctuation (RMSF) profiles, expressed as B-factors, for each residue across the simulated complexes. Panel A shows the RMSF data for the p53 protein in complex with ligands 1b (red trace) and 2a (black trace). Panel B displays the corresponding RMSF values for the HSP complexes with ligands 3a (pink), 4a (brown), and 4c (green).

Binding free energy estimates were obtained through MMGBSA calculations using the cpptraj module on the simulation trajectories of all five complexes. For the p53+1b and p53+2a systems, analysis began at frame 6000 with sampling every five frames. For the HSP+3a, HSP+4a, and HSP+4c complexes, processing started at frame 4000, also with five-frame intervals. Frame selection for each complex was guided by autocorrelation analysis to ensure statistically independent snapshots.

The MMGBSA-derived energy profiles revealed differences between the p53 complexes, with p53+1b initially exhibiting more favorable binding free energies compared to p53+2a. However, toward the end of the simulation, the values for both systems converged to similar levels, as illustrated in **Figure 9a**. Overall, both ligands demonstrated thermodynamically favorable interactions with the p53 receptor. Statistical comparison via t-tests confirmed significant differences in binding energies between the two complexes.

For the HSP complexes, MMGBSA results indicated sustained thermodynamically favorable binding energies throughout the simulations, supporting effective targeting of the HSP protein. One-way ANOVA revealed that the binding free energy for ligand 3a was significantly different from those of 4a and 4c, whereas no significant difference was observed between 4a and 4c **(Figure 9b)**.

A summary of the MMGBSA binding free energy values across all complexes is depicted as box-and-whisker plots in **Figure 10.**



a) b)

Figure 9. Time-series profiles of MMGBSA-derived binding free energies across the simulation trajectories. Panel A displays the energy values for the p53 protein complexes with ligands 1b (red trace) and 2a (black trace). Panel B shows the corresponding MMGBSA energy profiles for the HSP complexes with ligands 3a (pink), 4a (brown), and 4c (green).

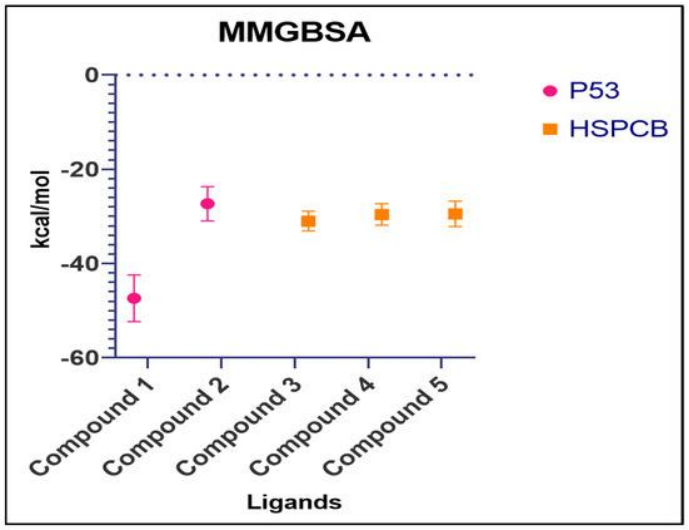


Figure 10. Box-and-whisker plots summarizing the MMGBSA binding free energy distributions for the p53 (pink) and HSPCB (yellow) complexes. The datasets correspond to Compound 1 (1b), Compound 2 (2a), Compound 3 (3a), Compound 4 (4a), and Compound 5 (4c).

The phytochemical constituents (PCs) identified in the root (NR) and stem (NS) n-hexane extracts, as well as the methanolic extracts (MR and MS) of P. nepalensis via gas chromatography-mass spectrometry (GC-MS), were systematically evaluated in this investigation. Induced genes and their encoded proteins were pinpointed as prospective therapeutic targets through gene enrichment analysis, construction of protein–protein interaction (PPI) networks, molecular docking, molecular dynamics simulations (MDS), and MMGBSA binding free energy computations.

From the 764 genes induced by the PCs (detailed in Text S1–S4), predictions were generated using DIGEP-Pred with a probability of activity (Pa) threshold exceeding 0.8. Functional enrichment analysis was subsequently applied to annotate these genes in terms of gene ontology descriptors (biological processes, molecular functions, cellular components), associated pathways, diseases, and marketed drugs. Each extract's gene list was inputted into StringDB to generate PPI networks.

PPI networks delineate the encoded proteins from these genes, their functional roles, and interconnectivities, where nodes represent proteins and edges denote interactions. The topological properties of the most highly connected genes within these networks are tabulated in **Table 2**. For the MR-extract-induced genes, HSPCB and NFKB1 exhibited degree centrality (DC) values of 15, which were lower than those of TP53 in the networks derived from the other extracts. Three proteins displayed average shortest path lengths ranging from 1.75 to 3.02, signifying closer interconnectivity among their associated nodes relative to the broader gene set. The clustering coefficient (CC) for TP53 was modestly lower than for HSPCB and NFKB1. Closeness centrality (C. cen) values for TP53 (0.53, 0.42, and 0.57) diverged slightly from those of HSPCB and NFKB1 (0.33 and 0.34). TP53 possessed the highest betweenness centrality (BC) values (0.56, 0.58, and 0.59), surpassing HSPCB and NFKB1 (0.18 and 0.33). These metrics underscore the pivotal network positions of these three genes, rendering them high-priority targets. Consequently, TP53, HSPCB, and NFKB1 were selected as promising therapeutic targets for the PCs derived from P. nepalensis.

HSPCB encodes a member of the heat shock protein 90 family, functioning as a pseudogene ortholog of heat shock protein 90 kDa protein 1 beta. It plays critical roles in signal transduction, gastric apoptosis, protein folding, and inflammatory responses. Prior research has implicated HSPCB as a viable target in cancer cell lines, including breast cancer [26] and ovarian cancer tissues [27].

Nuclear factor kappa B subunit 1 (NFKB1), a nuclear transcription regulator, translocates into the nucleus upon stimulation by cytokines or oxidative free radicals to drive gene transcription [28]. Dysregulated NFKB1 activation is associated with multiple inflammatory disorders, while its sustained suppression impairs immune cell maturation and proliferation [29]. NFKB1 serves as a key therapeutic target in diabetic cardiomyopathy [30], and its inhibition attenuates pro-inflammatory signaling in relevant pathways [31, 32]. Antagonists targeting NFKB1 can profoundly alter core gene expression profiles in leukemogenesis [33].

Transcription of the TP53 gene yields the p53 tumor suppressor protein, which governs cell cycle progression to prevent uncontrolled division [34]. p53 is extensively recognized for its tumor-suppressive functions and is a validated target across diverse malignancies due to its involvement in early oncogenic events [35]. Current inhibitors, such as piperidinone derivatives, spirooxindoles, nutlins, and isoquinolinones, target the p53-MDM2 interaction [36] Ongoing efforts focus on developing small-molecule ligands to reactivate mutant p53 and reinstate its anti-tumor efficacy [37].

Aligning with these established findings and the results of the present work, the PCs from P. nepalensis demonstrate therapeutic potential through specific binding to these validated targets. Molecular docking analyses **(Table 3)** identified PCs 1b and 2a as exhibiting superior binding affinities to p53, while 3a, 4a, and 4c preferentially bound HSPCB. The predominant interactions in both p53 and HSPCB complexes were hydrophobic in nature. Strategic incorporation of functional groups capable of forming hydrogen bonds could further enhance binding potency. MDS trajectories, as evidenced by RMSD profiles in **Figure 7** and RMSF in **Figure 8**, confirmed complex stability with deviations confined to 2–3 Å and no excessive fluctuations. MMGBSA free energy profiles in **Figure 9** affirmed persistent thermodynamic favorability for HSPCB targeting across the full simulation timeframe.

Conclusion

P. nepalensis has long been valued for its medicinal properties, largely due to bioactive phytochemical constituents (PCs) characterized in earlier work. The current study identified 764 genes modulated by these compounds. Detailed examination of protein–protein interaction networks highlighted three key therapeutic targets: TP53, influenced by compounds from MS, NR, and NS extracts, and HSPCB and NFKB1, linked to the MR extract.

Strong binding was observed for compounds 1b (from MR) and 2a (from MS) with p53 protein, achieving docking scores of −8.6 and −8.0 kcal/mol, respectively. Compounds 3a (NR), 4a, and 4c (NS) demonstrated robust affinities to HSP with scores of −9.6, −8.7, and −8.2 kcal/mol. Molecular dynamics simulations and MMGBSA calculations confirmed the structural stability of these complexes, with minimal deviations and thermodynamically favorable binding free energies.

Given the established roles of TP53, HSPCB, and NFKB1 in multiple malignancies, these findings highlight the capacity of P. nepalensis-derived PCs to modulate cancer-related genes and proteins. This work provides a solid foundation for future investigations into their therapeutic potential.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Harvey AL. Natural products in drug discovery. Drug Discov Today. 2008;13:894–901.
2. Mathur S, Hoskins C. Drug development: lessons from nature. Biomed Rep. 2017;6:612–14.
3. Kumari S, Seth A, Sharma S, Attri C. A holistic overview of different species of Potentilla, a medicinally important plant, along with their pharmaceutical significance: a review. J Herb Med. 2021;29:100460.
4. Tomczyk M, Pleszczyńska M, Wiater A. Variation in total polyphenolics contents of aerial parts of Potentilla species and their anticariogenic activity. Molecules. 2010;15:4639–51.
5. Tomczyk M, Paduch R, Wiater A, Pleszczyńska M, Kandefer-Szerszeń M, Szczodrak J. The influence of aqueous extracts of selected Potentilla species on normal human colon cells. Acta Pol Pharm. 2013;70:523–31.
6. Tomczyk M, Leszczyńska K, Jakoniuk P. Antimicrobial activity of Potentilla species. Fitoterapia. 2008;79:592–4.

7. Sharma S, Kumar V, Yaseen M, Sabouzied A, Arshad A, Bhat MA, et al. Phytochemical analysis, in vitro biological activities, and computer-aided analysis of Potentilla nepalensis Hook compounds as potential melanoma inhibitors based on molecular docking, MD simulations, and ADMET. Molecules. 2023;28:5108.
8. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2023 update. Nucleic Acids Res. 2023;51:D1373–D1380.
9. Lagunin A, Ivanov S, Rudik A, Filimonov D, Poroikov V. DIGEP-Pred: web service for in silico prediction of drug-induced gene expression profiles based on structural formula. Bioinformatics. 2013;29:2062–3.
10. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90–W97.
11. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51:D638–D646.
12. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–2504.
13. Praveen M, Morales-Bayuelo A. Drug designing against VP4, VP7 and NSP4 of rotavirus proteins—In silico studies. Moroc J Chem. 2023;11:729–41.
14. Stroud JC, Oltman A, Han A, Bates DL, Chen L. Structural basis of HIV-1 activation by NF-κB—A higher-order complex of p50:RelA bound to the HIV-1 LTR. J Mol Biol. 2009;393:98–112.
15. Wright L, Barril X, Dymock B, Sheridan L, Surgenor A, Beswick M, et al. Structure-activity relationships in purine-based inhibitor binding to HSP90 isoforms. Chem Biol. 2004;11:775–85.
16. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: an open chemical toolbox. J Cheminform. 2011;3:33.
17. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem. 2009;30:2785–91.
18. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2009;31:455–61.
19. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem. 2004;25:1157–74.
20. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 2021;30:70–82.
21. Tian C, Kasavajhala K, Belfon KA, Raguette L, Huang H, Migues AN, Bickel J, et al. ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. J Chem Theory Comput. 2020;16:528–52.
22. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov. 2015;10:449–61.
23. Turner PJ. XMGRACE, Version 5.1.19. Beaverton (OR): Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology; 2005.
24. R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2018. Available from: https://www.R-project.org/ (accessed 11 Nov 2023).
25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25:25–9.
26. Liu LL, Zhao H, Ma TF, Ge F, Chen CS, Zhang YP. Identification of valid reference genes for the normalization of RT-qPCR expression studies in human breast cancer cell lines treated with and without transient transfection. PLoS ONE. 2015;10:e0117058.
27. Fu J, Bian L, Zhao L, Dong Z, Gao X, Luan H, et al. Identification of genes for normalization of quantitative real-time PCR data in ovarian tissues. Acta Biochim Biophys Sin. 2010;42:568–74.
28. Cartwright T, Perkins ND, Wilson CL. NFKB1: a suppressor of inflammation, ageing and cancer. FEBS J. 2016;283:1812–22.
29. Li J, Lei WT, Zhang P, Rapaport F, Seeleuthner Y, Lyu B, et al. Biochemically deleterious human NFKB1 variants underlie an autosomal dominant form of common variable immunodeficiency. J Exp Med. 2021;218:e20210566.

30. Guo Q, Zhu Q, Zhang T, Qu Q, Cheang I, Liao S, et al. Integrated bioinformatic analysis reveals immune molecular markers and potential drugs for diabetic cardiomyopathy. Front Endocrinol. 2022;13:933635.
31. Weinmann D, Mueller M, Walzer SM, Hobusch GM, Lass R, Gahleitner C, et al. Brazilin blocks catabolic processes in human osteoarthritic chondrocytes via inhibition of NFKB1/p50. J Orthop Res. 2018;36:2431–8.
32. Siddique MH, Bukhari S, Khan IU, Essa A, Ali Z, Sabir U, et al. In silico, in vitro, and in vivo evaluation of caffeine-coated nanoparticles as a promising therapeutic avenue for AML through NF-kappa B and TRAIL pathways modulation. Pharmaceuticals. 2023;16:1742.
33. Reikvam H. Inhibition of NF-κB signaling alters acute myelogenous leukemia cell transcriptomics. Cells. 2020;9:1677.
34. Damineni S, Rao VR, Kumar S, Ravuri RR, Kagitha S, Dunna NR, Digumarthi R, et al. Germline mutations of TP53 gene in breast cancer. Tumor Biol. 2014;35:9219–27.
35. Hassin O, Oren M. Drugging p53 in cancer: one protein, many targets. Nat Rev Drug Discov. 2023;22:127–44.
36. Nayak SK, Khatik GL, Narang R, Monga V, Chopra HK. P53-Mdm2 interaction inhibitors as novel nongenotoxic anticancer agents. Curr Cancer Drug Targets. 2018;18:749–72.
37. Zhao D, Tahaney WM, Mazumdar A, Savage MI, Brown PH. Molecularly targeted therapies for p53-mutant cancers. Cell Mol Life Sci. 2017;74:4171–87.