



OPEN

# Investigation of the therapeutic role of native plant compounds against colorectal cancer based on system biology and virtual screening

Abbas Alibakhshi<sup>1,4</sup>, Rahim Malekzadeh<sup>2,4</sup>, Sayedeh Azimeh Hosseini<sup>2</sup> & Hajar Yaghoobi<sup>3</sup>✉

This study investigated the anticancer effects of compounds extracted from native plants on colon cancer following drug–target–network analysis and molecular docking. Based on the ChEBI database, compounds were identified in medicinal plants and weeds in the Chaharmahal and Bakhtiari provinces of Iran. A drug–target network was constructed based on candidate colon cancer protein targets and selective compounds. Network pharmacology analysis was conducted against the identified compounds and subjected to molecular docking studies. Based on molecular dynamics simulations, the most efficient compounds were evaluated for their anticancer effects. Our study suggests that TREM1, MAPK1, MAPK8, CTSB, MIF, and DPP4 proteins may be targeted by compounds in medicinal plants for their anti-cancer effects. Multiorthoquinone, Liquiritin, Isoliquiritin, Hispaglabridin A, Gibberellin A98, Cyclomulberrin, Cyclomorusin A, and Cudraflavone B are effective anticancer compounds found in targeted medicinal plants and play an important role in the regulation of important pathways in colon cancer. Compounds that inhibit MIF, CTSB, and MAPK8-16 appear to be more effective. Additional in vitro and in vivo experiments will be helpful in validating and optimizing the findings of this study.

Colorectal cancer (CRC) is a common malignant tumor of the gastrointestinal tract with vague early and late symptoms of anemia, weight loss, and other systemic symptoms<sup>1</sup>. In addition to conventional surgical intervention, other treatments are used, including the development of chemotherapy agents, finding new drug carriers, and combination therapy. Due to the side effects of some of these treatments, many studies have been conducted using natural compounds in cancer therapy<sup>2,3</sup>. The identification of natural compounds in herbal medicines that can be used as anti-cancer drugs has significantly reduced the mortality of cancer patients. In addition, natural compounds derived from medicinal plants have excellent anticancer effects as well as unique advantages such as low toxicity, fewer side effects, and lower cost. Several studies have focused on the therapeutic effects of herbal medicines on metastatic cancers. For example, cancer treatment using berberine, curcumin, resveratrol, and other bioactive compounds extracted from medicinal plants has been shown to target different factors that regulate tumor growth and metastasis. In contrast, the malignancy of cancer cells is typically caused by a combination of cellular and molecular factors<sup>4,5</sup>. Therefore, it can be claimed that choosing the proper intracellular factor or pathway as a therapeutic target is the first step in the therapy of such disorders. Identifying the active compounds of medicinal plants that influence various targets involved in the development of cancer cells can be considered an important step in cancer treatment.

Although routine studies based on natural substances, effective substances and their targets have already been selected based on experiences or impressions from previous studies, in such studies, the possibility of interaction between the effective substance and the target based on their physical structures and pharmacokinetic properties is not considered, and the possibility of finding several new potential substances and introducing unknown or

<sup>1</sup>Molecular Medicine Research Center, Hamadan University of Medical Sciences, Hamadan, Iran. <sup>2</sup>Department of Medical Biotechnology, School of Advanced Technology, Shahrekord University of Medical Sciences, Shahrekord, Iran. <sup>3</sup>Clinical Biochemistry Research Center, Basic Health Sciences Institute, Shahrekord University of Medical Sciences, Shahrekord, Iran. <sup>4</sup>These authors contributed equally: Abbas Alibakhshi and Rahim Malekzadeh. ✉email: hajar.yaghoobi64@gmail.com

less investigated targets is lost. Therefore, one way can be the rigid analysis of previous studies using interpretation of biological data and the examination of molecules in terms of their physicochemical structures for strong interaction with their target, thereby inhibiting its activity, so that it can finally introduce several molecules and lead to a large reduction in the probability of trial and error. Investigating biological interaction networks, including the drug–target interaction network, using data mining and analyzing and simulating their interactions with the help of computer and bioinformatic methods can be very helpful in introducing potential molecules with hidden roles and finally using them in laboratory methods. High-throughput data such as genomic and proteomic data, have contributed significantly to mechanism-based drug discovery. In addition to accelerating drug target identification and screening, bioinformatics analysis also facilitates the identification of side effects and drug resistance. Laboratory investigations of molecular structures, as well as the development of molecular modeling and simulations alongside large molecular databases, have paved the way for more realistic investigations of molecular connections and more informative virtual screening<sup>6</sup>.

The purpose of this study was to use compounds identified from native plants to investigate their possible therapeutic effects on colon cancer using one or more molecular targets. This interaction has been investigated based on pharmacological approaches by modeling drug–target (DT) and drug–target–disease (DTD) networks and analyzing them using molecular docking and MD simulation approaches. Estimating this possible correlation based on the molecular targets involved in different pathways of colon cancer could introduce anticancer compounds with low toxicity and side effects that simultaneously act against different pathways of carcinogenesis, including metastasis, proliferation, and angiogenesis.

## Methods

**Preparation of compounds.** All known medicinal plants and weeds in the Chaharmahal and Bakhtiari provinces of Iran were extracted from the flora book of this province and literature review. The names of the plants were searched in the Chemical Entities of Biological Interest (ChEBI) database for potential compounds<sup>7</sup>. ChEBI is a freely available database and the ontology of molecular entities focuses on small chemical compounds. Among all the compounds in this database, the plant compounds identified in the study were downloaded in the SDF format (structural data file). After removing minor and duplicate compounds, the final compounds were selected.

**Preparation of targets.** The PharmMapper website (<http://www.lilab-ecust.cn/pharmmapper/>) was used to identify the potential target candidates for these compounds<sup>8</sup>. This web server finds the best mapping poses of the uploaded molecules against all targets in several web-accessible databases. The target set option was set to human protein targets, and the maximum number of reserved matched targets was set to 300. After eliminating duplicates, desired targets were selected for further evaluation. Colorectal cancer-related targets were identified using the NCBI database (<https://www.ncbi.nlm.nih.gov/gene>) and matched to the potential targets of the selected compounds. Common targets of the two groups were selected for further investigation.

A drug–target (DT) network was constructed using Cytoscape v3.6.0 software based on candidate protein targets and selective compounds. To provide a network diagram, the targets are represented by circular nodes and compounds by triangle-shaped nodes. The NetworkAnalyzer plugin<sup>9</sup> was used to analyze the quantitative properties of the undirected network, in which parameters such as degree and betweenness centrality were estimated, and the nodes for each target and compound with the highest degree and betweenness centrality score were selected<sup>10</sup>.

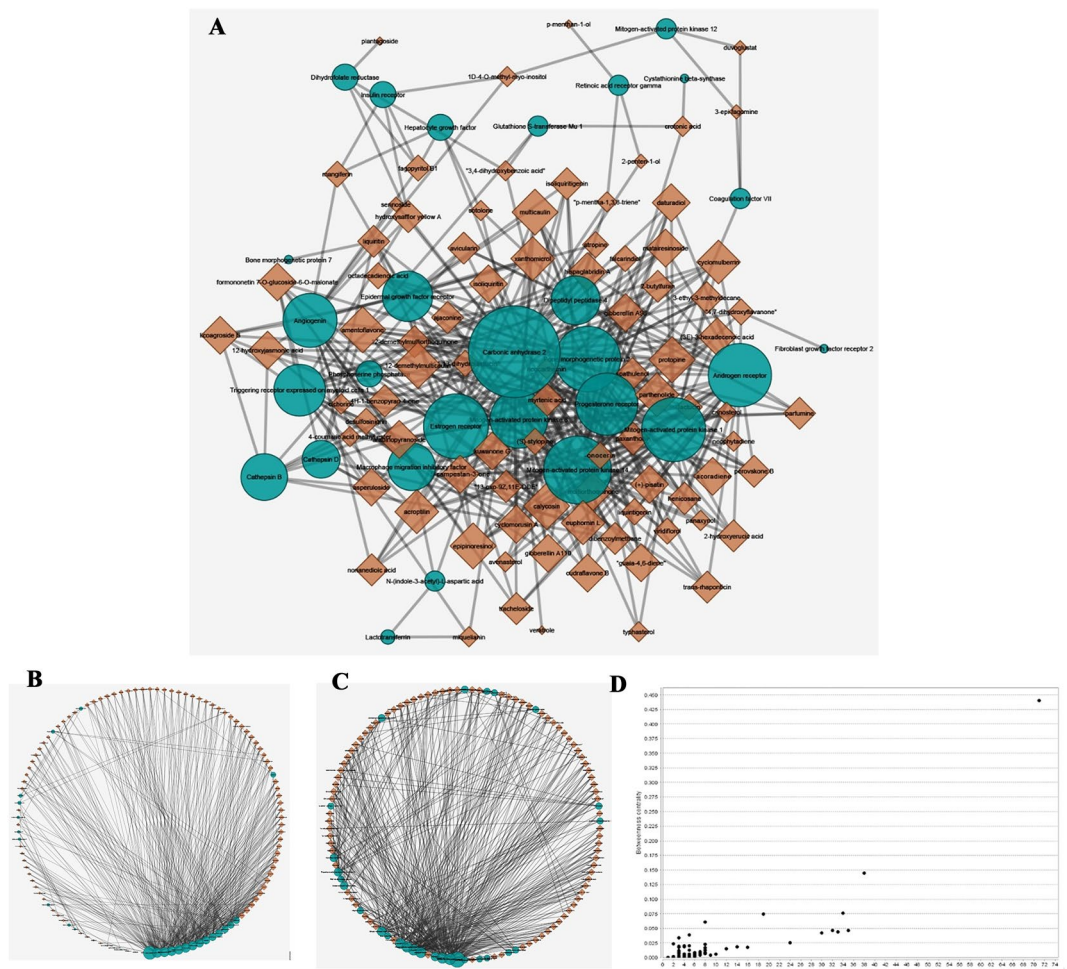
**Protein and ligand preparation and molecular docking study.** Targeted proteins were selected from the network analysis, and their structures were retrieved in PDB format from the Protein Data Bank (<http://www.rcsb.org/PDB>) and converted to pdbqt using AutoDock software<sup>11</sup>. In addition, the selected compound structures in sdf format were converted into mol2 and pdbqt formats using OpenBabel and Raccoon software, respectively<sup>12</sup>. Vina software hosted in PyRx, a virtual screening software<sup>13</sup>, was used for docking, and PyMol and DiscoveryStudio were used to analyze docked complexes.

**Pharmacokinetic and ADME study.** Compounds that showed interaction energy above  $-7.5$  kcal/mol against their targets were checked by the SwissADME website<sup>14</sup> in terms of physicochemical descriptors, as well as to predict ADME parameters (absorption, distribution, metabolism, and excretion), pharmacokinetic properties, and drug-like nature using the Lipinski rule.

**Molecular dynamic simulation.** United atom molecular dynamics (MD) simulations were performed using the GROMACS 2019.6 software package and the G43a1 force field in combination with the SPC water model for all protein and ligand–protein complexes. The system neutralization and physiological ionic strength of 0.14 M were achieved by adding suitable amounts of Na<sup>+</sup> and Cl<sup>-</sup> ions<sup>15</sup>.

First, energy minimization for the relaxation of internal constraints was performed using the steepest descent method until the system converged. Equilibration in the NVT and NPT ensembles was then performed under positional restraints for 500 and 1000 ps, respectively. Finally, an MD production run was performed for 100 ns with a 2 fs time step for all proteins and ligand–protein complexes.

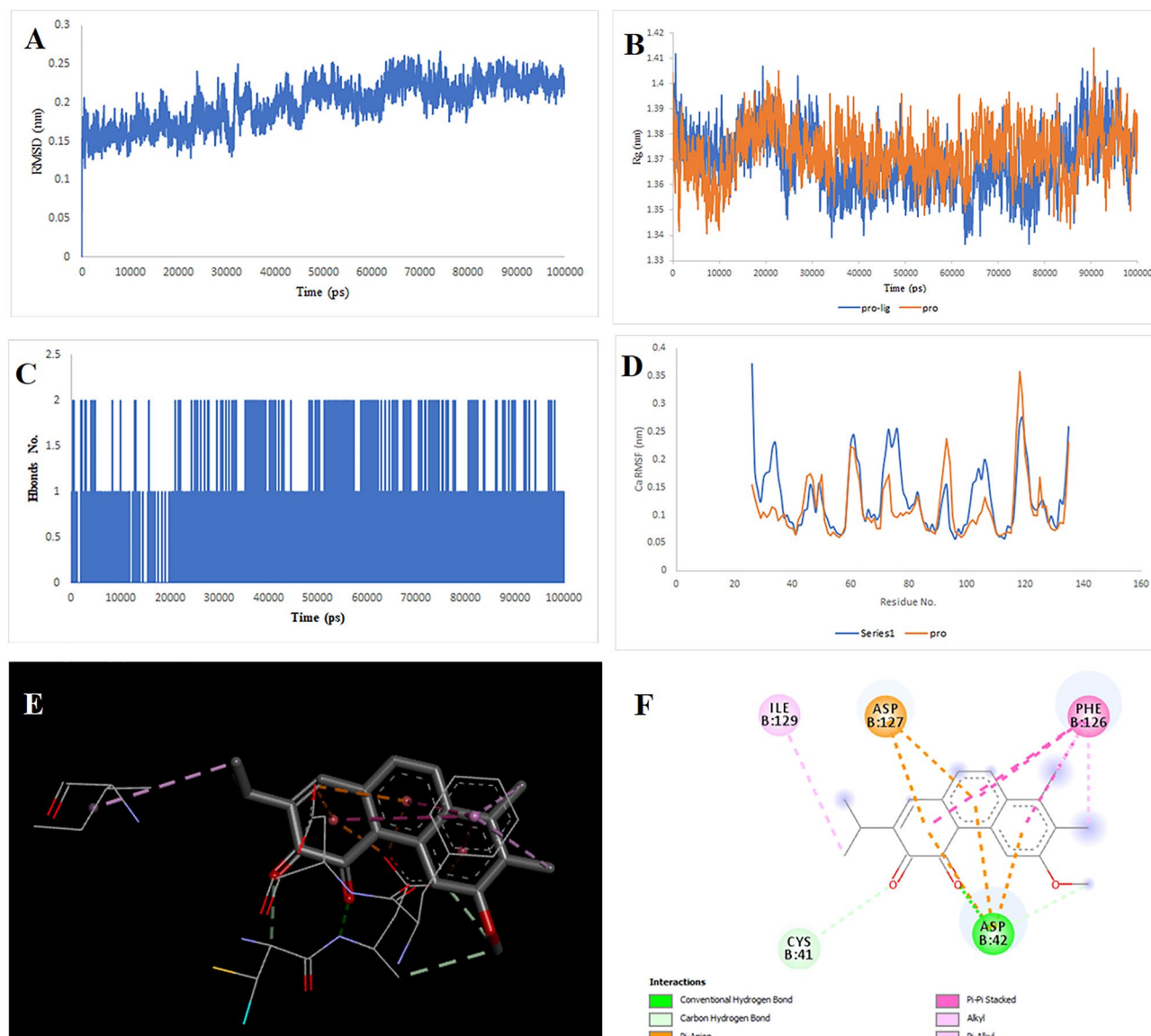
The Ewald particle mesh<sup>16</sup> was used for long-range electrostatic forces, and the lengths of all bonds were constrained using the LINCS algorithm<sup>17</sup>. The temperature was set to 310 K using a V-rescale thermostat<sup>18</sup> and the pressure was controlled at 1.0 atm using a Parrinello–Rahman barostat<sup>19</sup>.



**Figure 1.** (A) The bipartite DT network. (B) The circle layout of DT network based on degree. (C) The circle layout of DT network based on betweenness centrality. (D) The betweenness centrality of nodes. The green circular nodes represent the targets, and the brown triangular nodes represent the chemical compounds. The larger the size of the node, the greater the number of nodes and associations.

Compound name	Degree	Compound name	Degree
12-Demethylmulticaulin	8	Formononetin 7-O-glucoside-6-O-malonate	6
Amentoflavone	8	Gibberellin A110	6
Calycosin	8	Hispaglabridin A	6
Epipinoresinol	8	Isoliquiritin	6
Multicaulin	8	Licoagroside B	6
Protopine	8	Matairesinoside	6
Acroptilin	7	Cyclomorusin A	6
Cyclomulberrin	7	Daturadiol	6
Parthenolide	7	Gibberellin A98	6
Xanthomicrol	7	“Guaia-4,6-diene”	6
Euphornin L	7	Multiorthoquinone	6
Arabinopyranoside	6	β-Acoradiene	6
(+)-Pisatin	6	12-Hydroxyjasmonic acid	6
(3E)-3-Hexadecenoic acid	6	Asperuloside	6
5α±-Campestan-3-one	6	Kuwanone G	6
Cudraflavone B	6		

**Table 1.** The compounds that were selected for molecular docking based on the number of degrees.



**Figure 2.** Two-dimensional representations of the Multiorthoquinone against TREM1. (A) Root-mean square deviation of the complexes (RMSD). (B) Radius of gyration (Rg). (C) Hydrogen bond analysis from the simulation system. (D) Root-mean-square fluctuation (RMSF). (E) The binding conformation of 3D view. (F) Binding site interactions of 2D view.

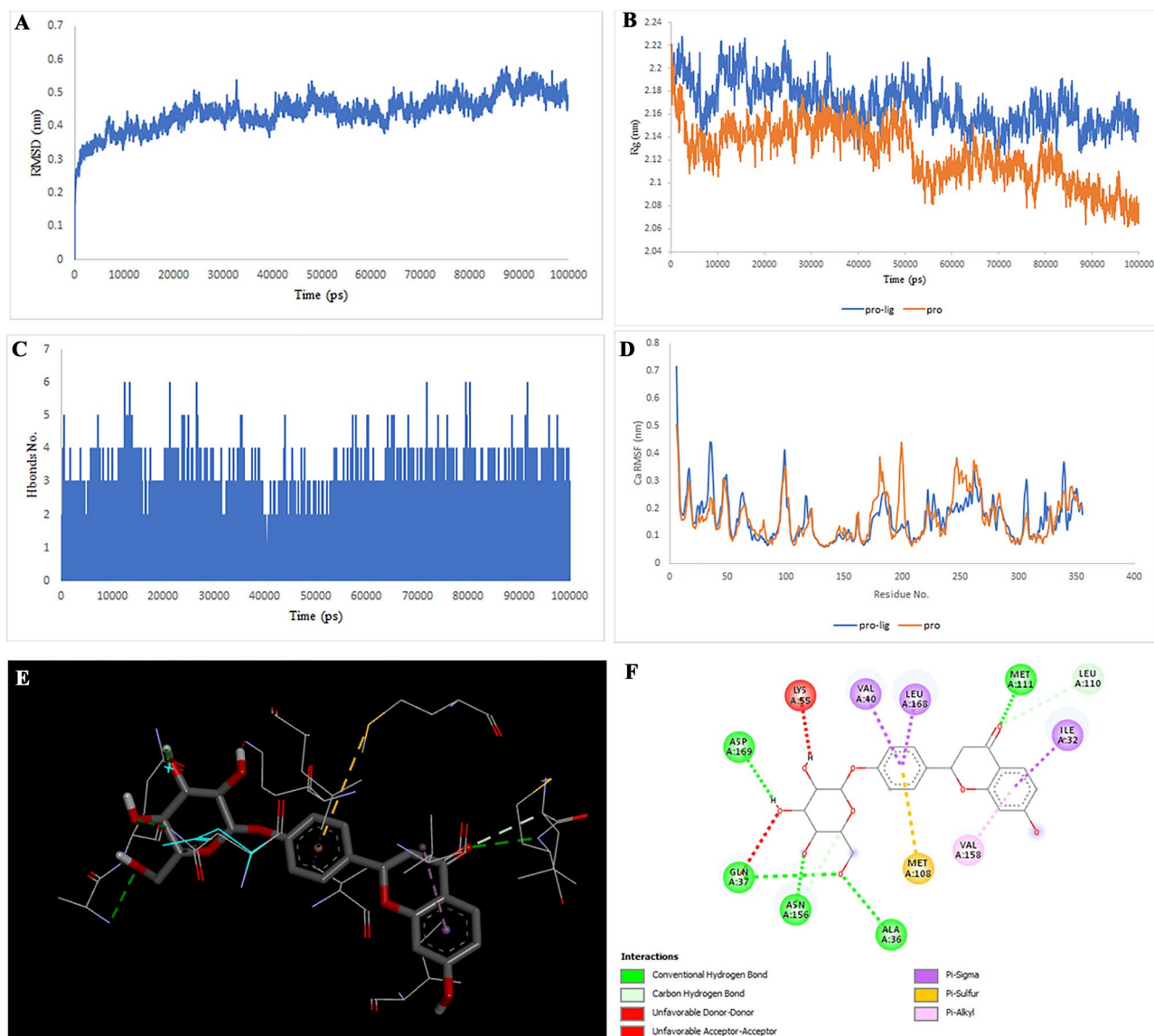
**MM/PBSA binding free energy calculation.** The molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) approach<sup>20</sup> was applied to calculate the binding free energy of protein–ligand interactions using the last 50 ns of each simulation trajectory.

The total binding energy of the system was calculated using the following equation:

$$\Delta G_{\text{Binding}} = \Delta G_{\text{Complex}} - \Delta G_{\text{Ligand}} - \Delta G_{\text{Receptor}}$$

**Principal component analysis.** We performed principal component analysis (PCA) of the  $\alpha$ -carbon atoms in the 100 ns trajectories to further explore the dynamic properties of the investigated structures. PCA was obtained as the eigenvectors of a covariance matrix consisting of  $\alpha$ -carbon displacements that reflects the overall conformational motions and flexibility of the systems during the simulations.

The first two principal components (PC1 and PC2) that dominate protein conformational fluctuations were used to analyze the proteins and protein–ligand complexes. The eigenvalue corresponding to each eigenvector represents the energy contribution of that particular part of the motion. Eigenvectors and eigenvalues were generated using the `g_covar` Gromacs utility by calculating and diagonalizing the covariance matrix. Eigenvector analysis was performed using the `g_anaeig` tool.



**Figure 3.** Two-dimensional representations of the Liquiritin against MAPK8. (A) Root-mean square deviation of the complexes (RMSD). (B) Radius of gyration (Rg). (C) Hydrogen bond analysis from the simulation system. (D) Root-mean-square fluctuation (RMSF). (E) The binding conformation of 3D view. (F) Binding site interactions of 2D view.

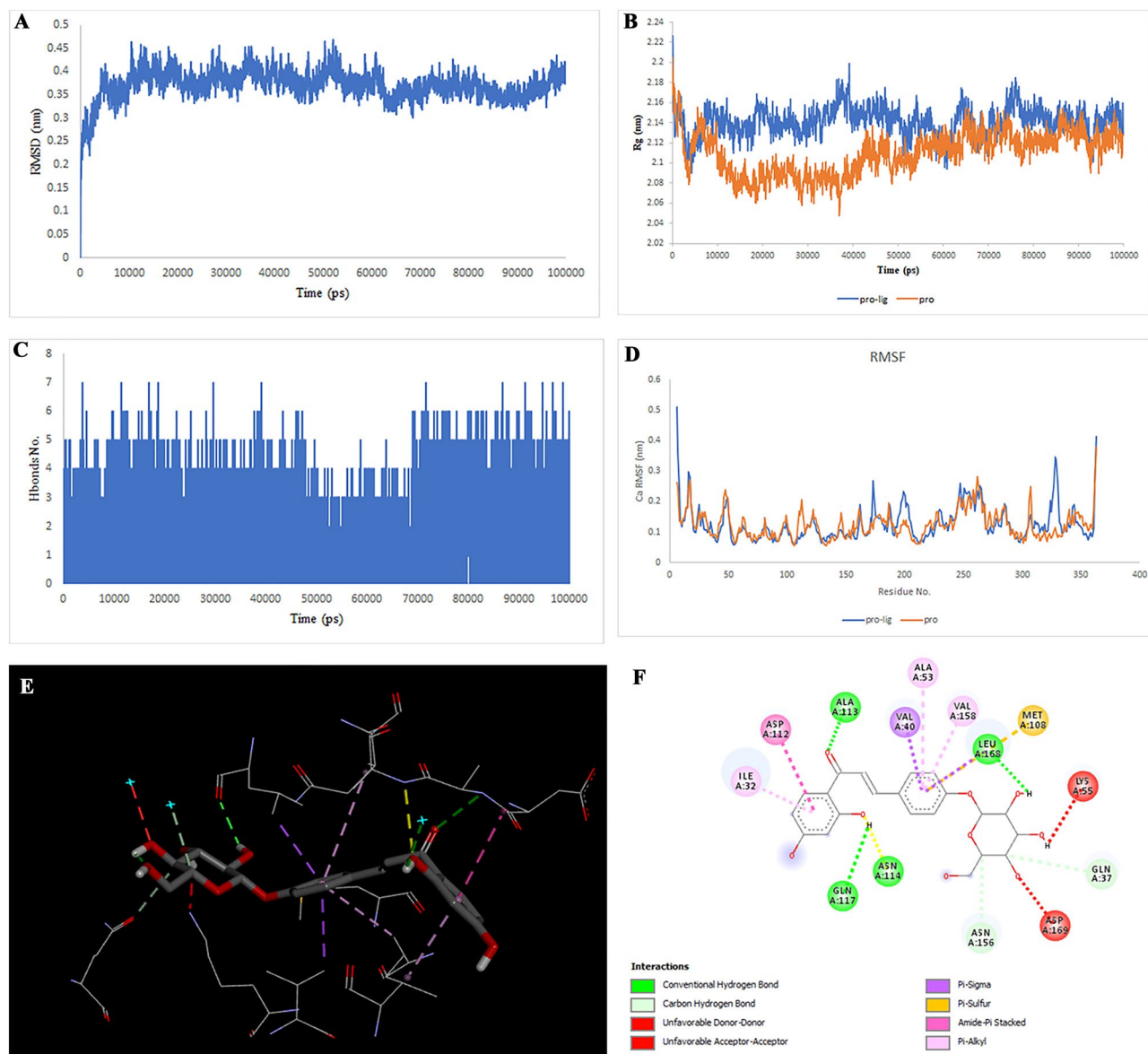
**Ethical approval.** This article does not contain any studies involving human participants or animals performed by any of the authors.

## Results

**Libraries preparation.** A total of 265 medicinal plants were selected and introduced into the ChEBI database to extract the effective compounds. After removing duplicates, 87 potential compounds were obtained from the 265 native plant species. The PharmMapper website has introduced 423 potential target candidates for these compounds. In addition, the NCBI database has introduced 3312 colorectal cancer-related targets. Finally, 34 human protein molecules were selected, which could be targets of these compounds and play a role in colorectal cancer.

Of these compounds, 7 did not have common targets in their top 30 targets. The names and 2D structures of both the groups of compounds used in the screening and their targets are listed in Supplementary Tables S1–S3.

**DT network construction and analysis.** A DT network with 87 compounds and 34 targets was constructed consisting of 114 nodes and 416 degrees. The NetworkAnalyser results calculate parameters such as the number of degrees, betweenness, and edge betweenness values, which are used to select the best molecular targets and potential chemical compounds (Fig. 1). Figure 1A shows a bipartite network configured with node sizes based on the number of degrees; the larger the node size, the greater is the number of associa-

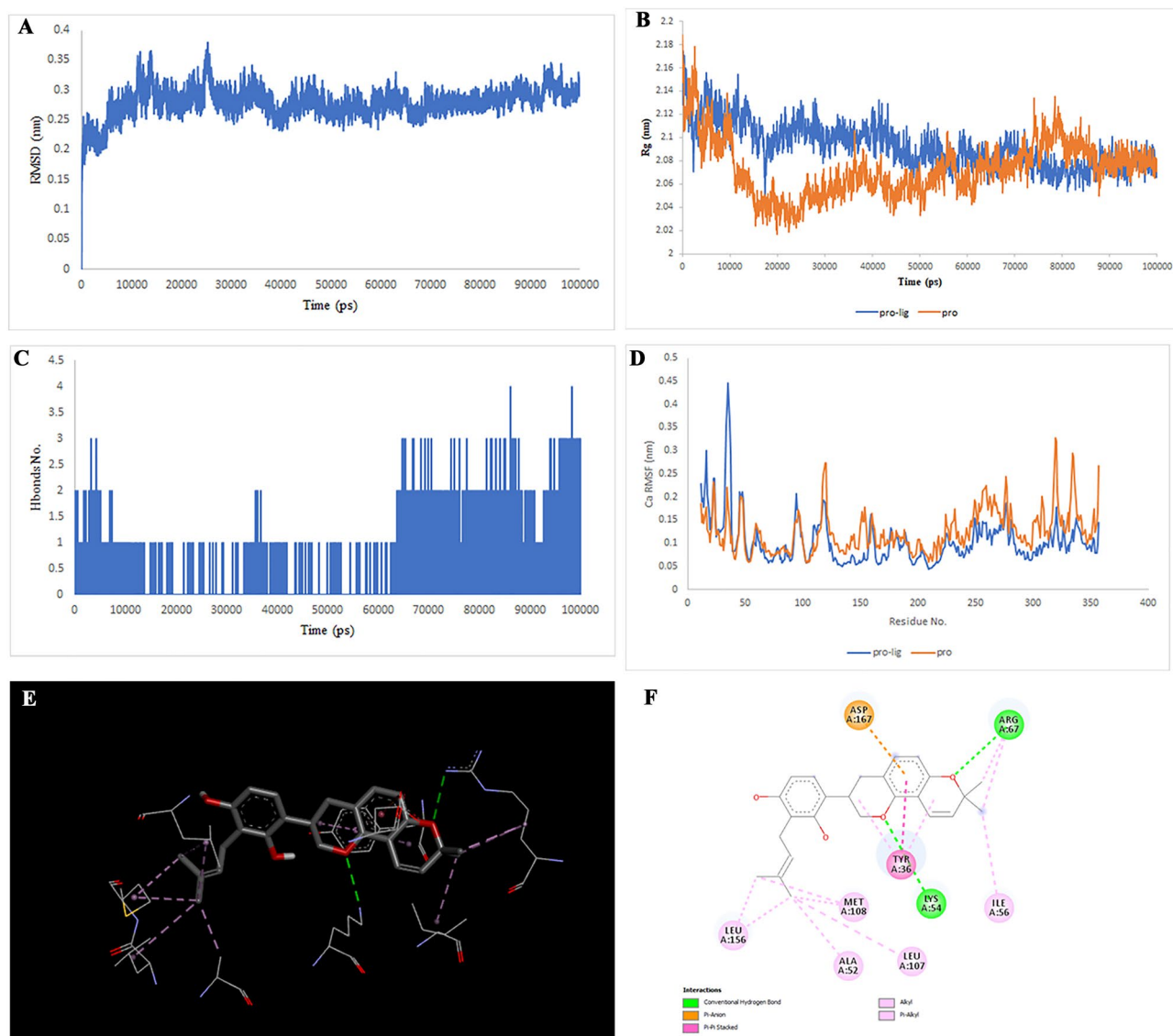


**Figure 4.** Two-dimensional representations of the Isoliquiritin against MAPK8. (A) Root-mean square deviation of the complexes (RMSD). (B) Radius of gyration (Rg). (C) Hydrogen bond analysis from the simulation system. (D) Root-mean-square fluctuation (RMSF). (E) The binding conformation of 3D view. (F) Binding site interactions of 2D view.

tions. Based on the network analysis, the compounds showed degrees ranging from 1 to 8. To screen the best compounds as ligands for molecular docking analysis against targets, compounds with six or more nodes were selected. Therefore, 31 compounds listed in Table 1 were selected for further study. Carbonic anhydrase 2, with a score of 0.44, showed the highest centrality among the targets, followed by mitogen-activated protein kinase 14 (0.14), estrogen receptor (0.07), and angiogenin (0.07). The remaining molecules had a score of less than 0.05. Moreover, Carbonic anhydrase 2 (71), mitogen-activated protein kinase 14 (38), bone morphogenetic protein 2 (35), and estrogen receptor (34) showed the most connections, based on the number of degrees associated with the nodes.

Finally, the compounds showed a number of degrees ranging from 1 to 8. To screen the best compounds as ligands for the molecular docking step against the targets, all compounds with six or more nodes were selected, in which there were 31 compounds with six, seven, or eight nodes.

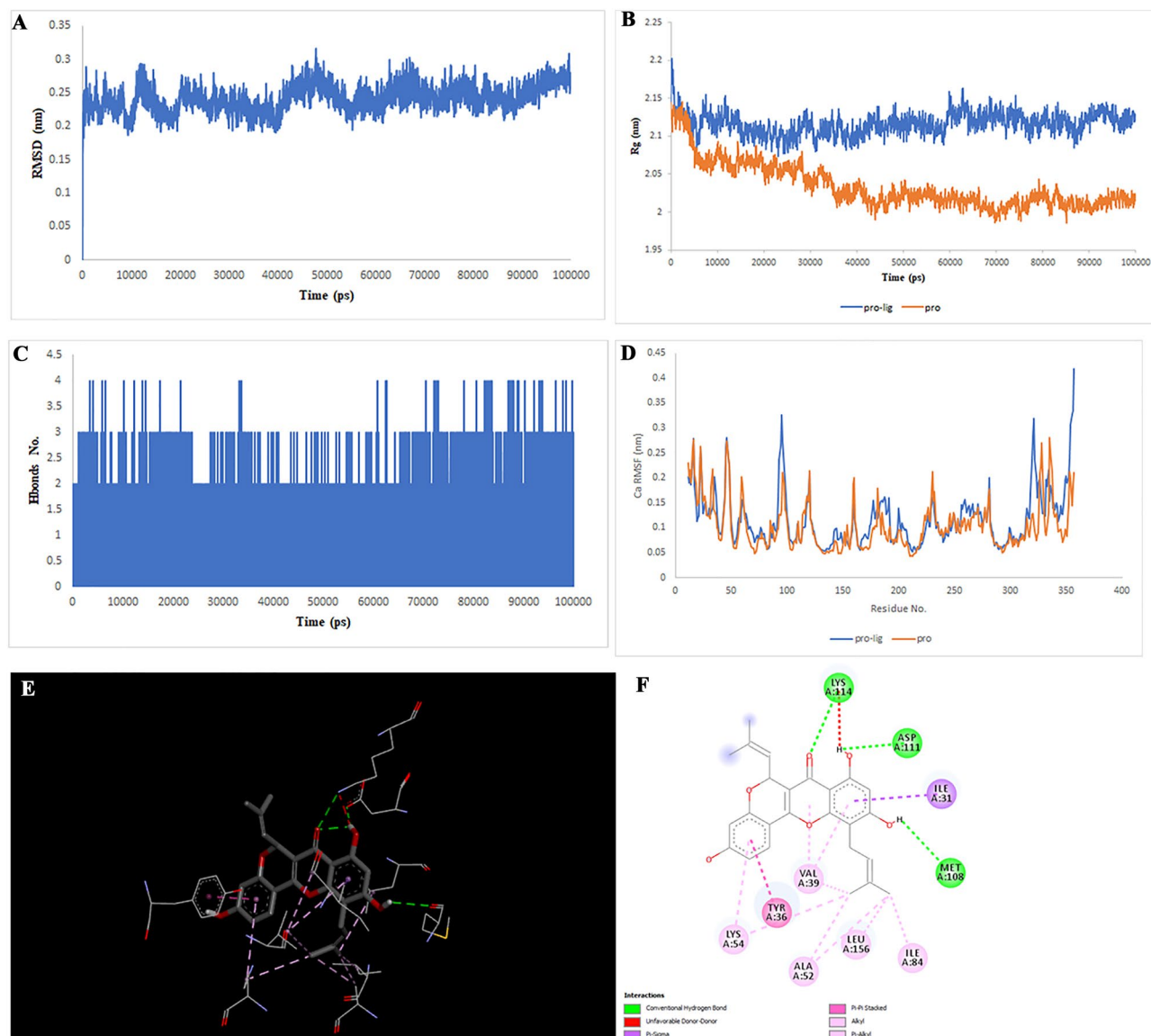
**Molecular docking and simulation.** After checking using the GEPIA database, the TOP targets that showed increased expression in colorectal cancer were selected. The target structures were downloaded from the PDB database (PDB IDs: 6AY2 for CTSB, 5T4E for DPP4, 6FVE for MIF, 7AUV for MAPK1, 4QTD for MAPK8, and 1SMO for TERM1). The Chimera 1.8.1 software was used for essential protein preparations, including removing water, ATP, ligands, and adding hydrogen charges. Critical ligand-binding sites were considered and



**Figure 5.** Two-dimensional representations of the Hispaglabridin A against MAPK1. **(A)** Root-mean square deviation of the complexes (RMSD). **(B)** Radius of gyration (Rg). **(C)** Hydrogen bond analysis from the simulation system. **(D)** Root-mean-square fluctuation (RMSF). **(E)** The binding conformation of 3D view. **(F)** Binding site interactions of 2D view.

docked to 31 compounds using the PyRx software. 31 molecules were selected after analyzing the pharmacokinetic properties and parameters of ADME. The centroid of the binding sites for the targets was calculated as the coordinates of the centroid of the ligand-binding sites using the UniProt database. The docked complexes were analyzed using PyMol and DiscoveryStudio software. Finally, molecules that presented the highest interaction energy against their target (above  $-7.5$  for TERM1,  $-8$  for CTSB and MIF, above  $-9$  for DPP4 and MAPK1 and above  $-10$  for MAPK8), respectively and had more binding with the amino acids of the binding site in the study with DiscoveryStudio. These findings led to the selection of Multiorthoquinone, Liquiritin, Hispaglabridin A, Isoliquiritin, Gibberellin A98, Cyclomulberrin, Cyclomorusin A, Cudraflavone B for simulation (MD), which produced a more stable complex with a lower energy level than TREM1, MAPK8, MAPK1, MAPK8, CTSB, MAPK1, MIF, and DPP4, respectively. The positions and amino acids involved in the binding are illustrated in Fig. 2, 3, 4, 5, 6, 7, 8 and 9, Supplementary Tables S4–S6. The RMSD score showed slight fluctuations and was approximately 0.2 nm. These findings indicate the stability of the compounds in the target complex. The RMSF values for all protein structures were computed to accurately determine how the binding of the compounds affects flexibility.

Analysis of the liquiritin- and isoliquiritin-associated RMSF plots revealed that Liquiritin and Isoliquiritin flexibility were significantly different in the five regions of 35, 183, 199, 246, and 307 and two regions in 200 and 328 in MAPK8, respectively. In addition, multiorthoquinone flexibility was extremely high in the three regions 34, 76, and 107 in TREM1, which may be attributed to a lack of interaction between the three regions and multiorthoquinone. For Cyclomulberrin and Hispaglabridin A residues 94–320, and 35, 300–350 in MAPK1, the flexibility of amino acids was lower and higher, respectively. The RMSF plot associated with DPP4 indicated high

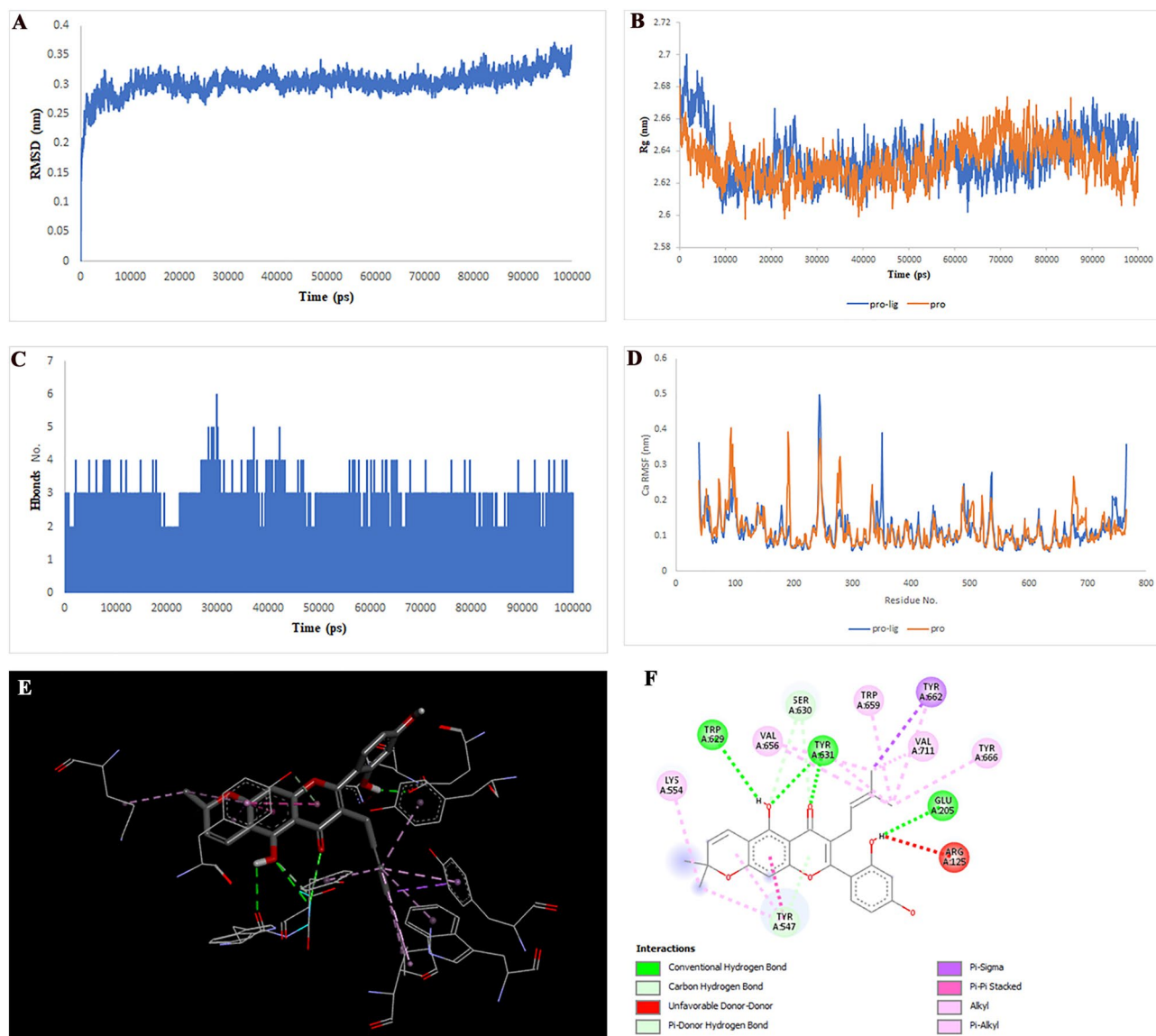


**Figure 6.** Two-dimensional representations of the Cyclomulberrin against MAPK1. (A) Root-mean square deviation of the complexes (RMSD). (B) Radius of gyration (Rg). (C) Hydrogen bond analysis from the simulation system. (D) Root-mean-square fluctuation (RMSF). (E) The binding conformation of 3D view. (F) Binding site interactions of 2D view.

interaction between most regions of DPP4 and Cudraflavone B. Additionally, Cyclomorusin A and Gibberellin A98 flexibility were extremely different in the two regions of 50–70 and 150–170 in the CT SB, and 80–100 region in the MIF, respectively.

RG nature was constant for the individual domains during the entire simulation period associated with Gibberellin A98, Cyclomorusin A, and Cudraflavone B, indicating that the individual domains did not melt or unfold. These compounds did not affect the secondary structures of CT SB, MIF, or DPP4. However, the RG value associated with Liquiritin, Isoliquiritin and Hispaglabridin A, Cyclomulberrin throughout the MD simulation led to unfolding and activation of MAPK8 and MAPK1, respectively.

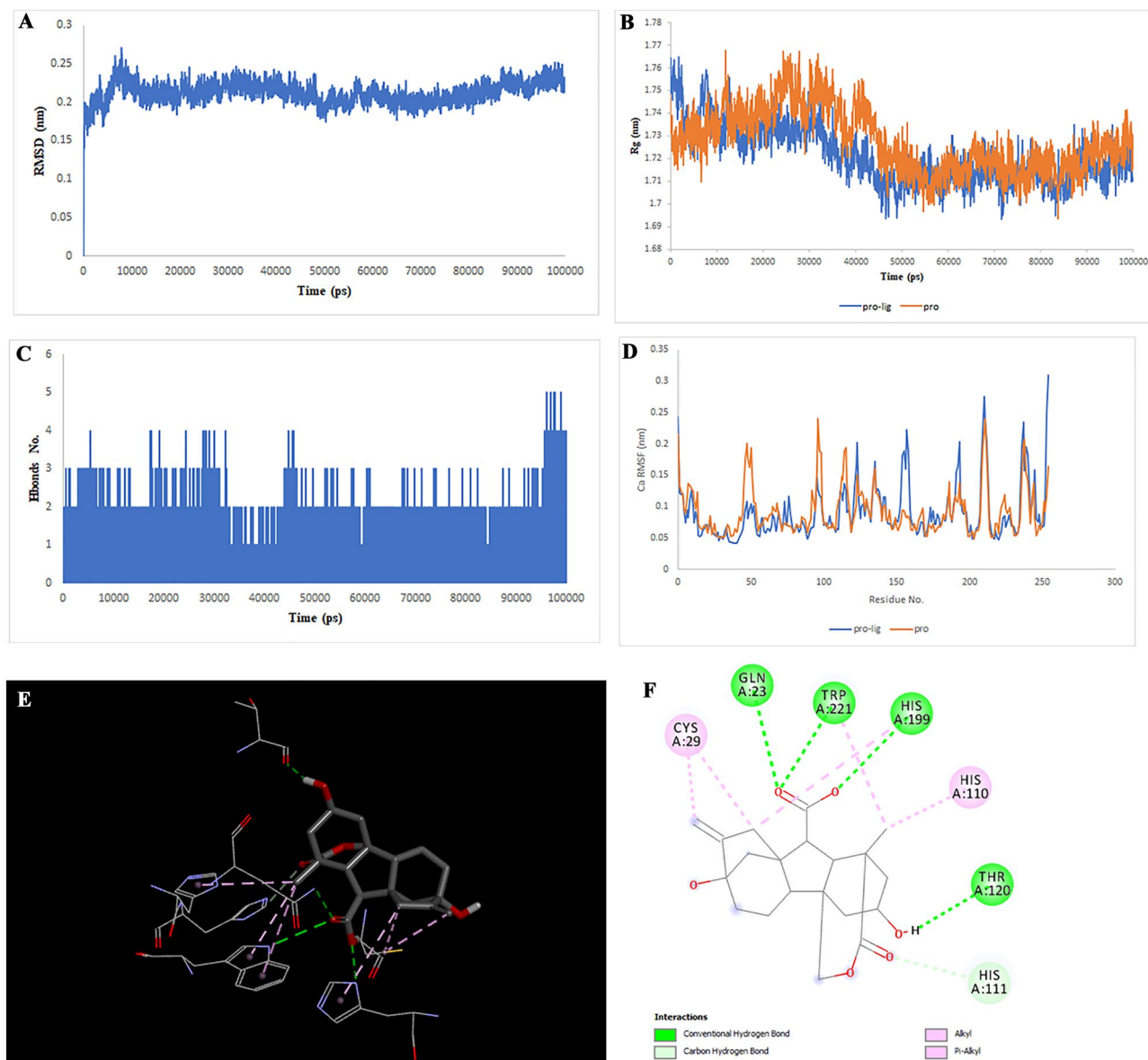




**Figure 7.** Two-dimensional representations of the Cudraflavone B against DPP4. **(A)** Root-mean square deviation of the complexes (RMSD). **(B)** Radius of gyration (Rg). **(C)** Hydrogen bond analysis from the simulation system. **(D)** Root-mean-square fluctuation (RMSF). **(E)** The binding conformation of 3D view. **(F)** Binding site interactions of 2D view.

**Principal component analysis.** Supplementary Fig. S1 shows the projection of the trajectories on special eigenvectors (vectors 1 and 2) and time-dependent motions of the components in a particular vibration mode. The overall analysis of the eigenvector plots indicated that most vibrations occurred along eigenvector 1. According to the first two PCs, TREM1 and MIF proteins have almost the same trace values of the covariance matrix for the bound and unbound states with a slight shift. This indicates that the ligand is well equilibrated and stabilized with the protein, as reflected by the least conformational changes due to reduced collective motions from unbound states. However, in other cases, 2D projection plots of the trajectories revealed that the ligand reduced the conformational diversity during the simulations, leading to a more compact cluster distribution. Sampling of different regions and showing different movement behaviors of protein–ligand complexes compared to unbound proteins points to the binding of ligand effects on the rigidity of the structural conformation, which also affects the function of proteins.

**MM/PBSA binding free energy.** The MM/PBSA binding free energy results are shown in Table 2 including the van der Waals energy (kJ/mol), electrostatic energy (kJ/mol), polar solvation energy (kJ/mol), SASA energy (kJ/mol), SAV energy (kJ/mol), WCA energy (kJ/mol), and binding energy (kJ/mol), are shown in Table 2. Compared to liquiritin, isoliquiritin had the lowest binding energy score of 183.04 kJ/mol interacting with MAPK8 and formed a stronger binding. In addition, hypoglabin A and cyclomulbrin interact with MAPK1 at almost the same binding energy (−154 kJ/mol). However, the number of hydrogen bonds in the



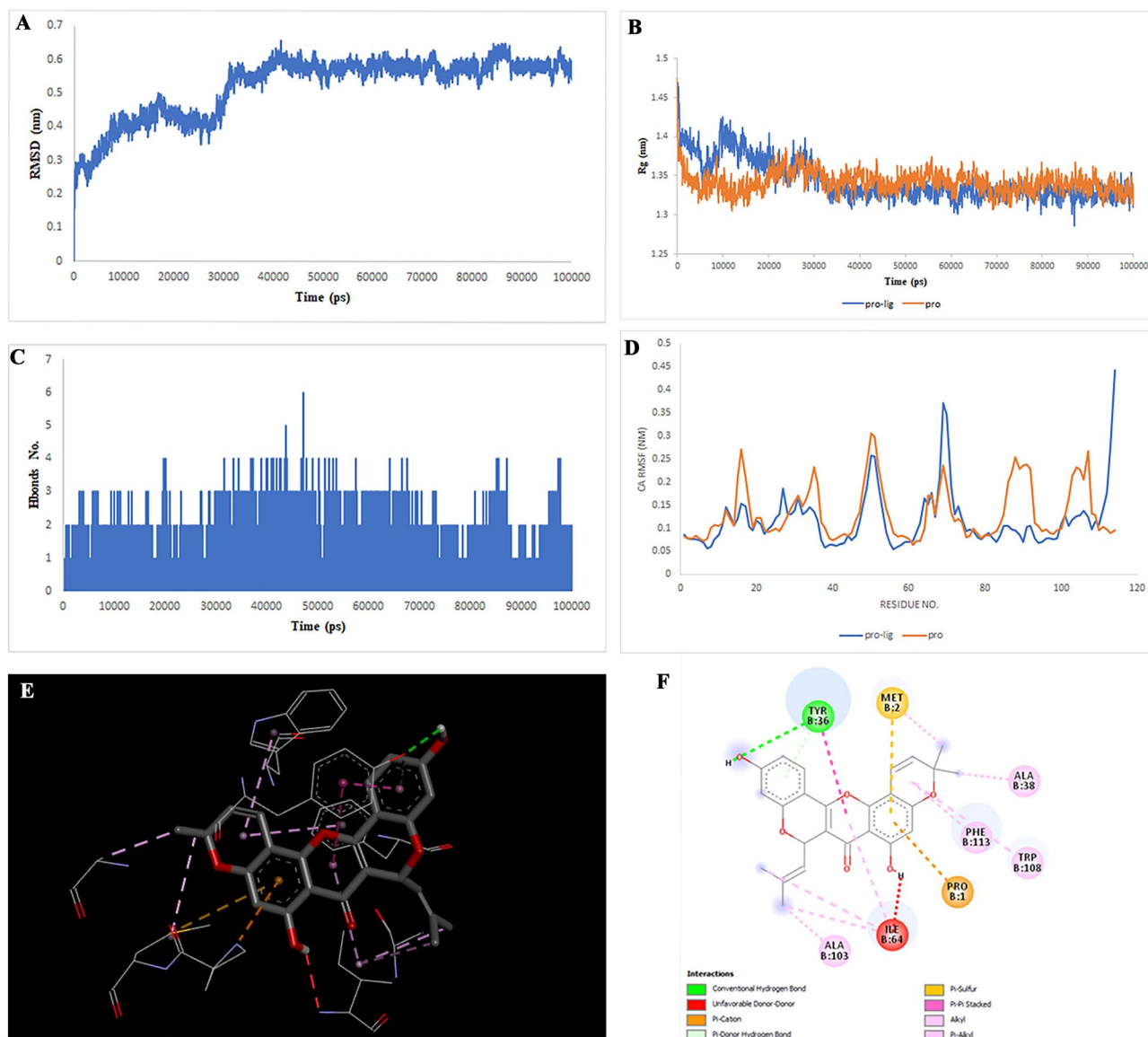
**Figure 8.** Two-dimensional representations of the Gibberellin A98 against CTSB. **(A)** Root-mean square deviation of the complexes (RMSD). **(B)** Radius of gyration (Rg). **(C)** Hydrogen bond analysis from the simulation system. **(D)** Root-mean-square fluctuation (RMSF). **(E)** The binding conformation of 3D view. **(F)** Binding site interactions of 2D view.

interaction with hypoglabin A is relatively high. MIF-Cyclomorusin A, with the most negative energy score ( $-243.768$  kJ/mol), showed the strongest interaction among all docked complexes.

## Discussion

Recently, phytoscience researchers have described a wide range of bioactive chemicals with various biological effects on plants. This issue has led to the allocation of pharmaceutical resources for the investigation and introduction of novel drugs from herbal sources. In this regard, identifying natural bioactive components and therapeutically active chemicals through the screening of plant extracts is the first step in the development of these drugs. A network pharmacology analysis, represented by the bipartite DT network, was performed to evaluate the potential interactions between the chemical compounds of medicinal plants and proteins identified in colon cancer. The results of the investigation of the target genes related to the signaling pathways of colon cancer and the names of the plant compounds are presented in Supplementary Tables S1 and S2.

Findings identified the genes MAPK1, MAPK8, TREAM1, CTSB, MIF, and DPP4 interacted most effectively with the compounds. Liquiritin and Isoliquiritin from *Glycyrrhiza glabra* effectively interact with MAPK8. In addition, Cyclomulberrin and Hispaglabridin A from *Morus* species and *Glycyrrhiza glabra* effectively inhibited MAPK1. Previously, Hispaglabridin A isoflavonoid isolated from *Glycyrrhiza glabra* was introduced as a potential agent for the development of an effective drug for colon cancer<sup>21,22</sup>. Many previous studies have provided evidence



**Figure 9.** Two-dimensional representations of the Cyclomulberrin against MIF. (A) Root-mean square deviation of the complexes (RMSD). (B) Radius of gyration (Rg). (C) Hydrogen bond analysis from the simulation system. (D) Root-mean-square fluctuation (RMSF). (E) The binding conformation of 3D view. (F) Binding site interactions of 2D view.

Energy Forms (kJ/mol)	TREM1— Multiorthoquinone	MAPK8— Liquiritin	MAPK8— Isoliquiritin	MAPK1— Hispaglabridin A	MAPK1— Cyclomulberrin	CTSB— Gibberellin A98	MIF— Cyclomorusin A	DPP4— Cudraflavone B
Van der Waal	-158.643 ± 1.833	-200.289 ± 3.237	-207.328 ± 3.231	-228.656 ± 1.11	1.591 ± 1.309	-176.028 ± 0.805	-259.193 ± 1.262	-259.028 ± 1.065
Electrostatic	-30.848 ± 1.237	-4.575 ± 0.806	-33.317 ± 1.116	-3.96 ± 0.334	1.591 ± 0.907	-22.762 ± 1.556	-38.108 ± 0.644	-25.021 ± 0.6
Polar solvation	49.704 ± 1.332	59.473 ± 1.569	76.214 ± 1.464	99.315 ± 1.622	1.591 ± 1.63	68.8 ± 0.737	72.135 ± 1.018	84.784 ± 0.694
SASA	-12.637 ± 0.113	-18.136 ± 0.313	-18.967 ± 0.307	-20.1 ± 0.1	1.591 ± 0.092	-13.791 ± 0.056	-18.568 ± 0.089	-19.461 ± 0.071
SAV	0	0	0	0	0	0	0	0
WCA	0	0	0	0	0	0	0	0
Binding	-152.141 ± 2.088	-163.657 ± 3.341	-183.044 ± 3.251	-153.478 ± 1.591	± 1.591 ± 1.678	-143.916 ± 1.308	-243.768 ± 1.184	-218.717 ± 1.12

**Table 2.** MM-PBSA results. SASA solvent accessible surface area, SAV solvent accessible volume, WCA weeks-chandler-andersen potential.

that Liquiritin and Isoliquiritin have anticarcinogenic activity in various types of cancer cells, especially colon cancer cells, by regulating different pathways, including PI3K/AKT, P53/P21, and miR-671/HOXB3 signaling. These compounds show anticancer activity, such as inhibition of tumor growth and cell proliferation and promotion of chemosensitivity and cell apoptosis<sup>23–29</sup>. MAPKs are one of the most important factors that regulate cell proliferation in cancer by regulating growth factor receptors. Disrupting their regulation could be an approach in cancer therapy. This study confirmed the inhibitory effects of these compounds on cell proliferation through the introduction of a new target.

C-cudraflavone derivatives showed anticancer activity in human colorectal, melanoma, and oral cancer by regulating the MAPK, SIRT1, NF- $\kappa$  B, and PI3K/AKT pathways<sup>30–32</sup>. The results of this study showed that the B-cudraflavone derivative also interacted with DPP4. Dipeptidyl peptidase 4 (DPP4) is a cell surface protein that acts as a tumor suppressor or activator depending on its expression. An in vitro analysis is required to investigate the function of DPP4 and the effect of cudraflavone on its function in colon cancer.

Autophagy is one of the strategies used by cancer cells to escape cell death. Increased CTSB levels are required to complete autophagy, and studies have shown that treatment with compounds that induce ERK phosphorylation reduces CTSB levels and blocks autophagy. Several studies have confirmed the oncogenic role of this protein in cell proliferation, invasion, metastasis, and drug resistance. Chemotherapy based on CTSB-depleting compounds suppresses autophagy rescue in cancer cells, ultimately leading to increased antitumor effects<sup>33–35</sup>. This study confirmed that cyclomulberrin induced CTSB folding and inhibited CTSB function. ERK-activating compounds, such as Liquiritin and Isoliquiritin, can be used in combination with cyclomulberrin to study the inhibition of drug resistance and proliferation in solid tumors, especially in colon cancer.

Multi-orthoquinones, Gibberellin A98, and Cyclomorulin A interact with TREM1, MAPK1, and MIF signaling pathways and have selective anti- or tumor-inducing activities. These results provided new insights into the mechanisms of action of these compounds. However, further studies are required to determine its exact effects.

## Conclusion

Network pharmacology analysis revealed the key pathways involved in the anti-cancer activities induced by natural compounds found in medicinal plants. Eight pathways were obtained from the drug–target network analysis (MAPK1, MAPK8, TREM1, CTSB, MIF, and DPP4), and molecular docking studies of these pathways revealed that the identified chemical compounds had strong binding affinities with these pathway components. The current study revealed that Multiorthoquinone, Liquiritin, Isoliquiritin, Hispaglabridin A, Gibberellin A98, Cyclomulberrin, Cyclomorulin A, and Cudraflavone B are effective anticancer compounds found in target medicinal plants and play an important role in the regulation of important pathways in colon cancer. Compounds that inhibit MIF, CTSB, and MAPK8-16 appear to be more effective. Additional in vitro and in vivo experiments will be helpful for validating and optimizing the findings of this study.

## Data availability

The datasets generated and/or analyzed during the current study are available upon request from the corresponding author.

Received: 14 January 2023; Accepted: 4 July 2023

Published online: 15 July 2023

## References

1. Yuan, J. *et al.* The effects of mesenchymal stem cell on colorectal cancer. *Stem Cells Int.* **2021**, 1–14 (2021).
2. El Bali, M., Bakkach, J. & Bennani Mechita, M. Colorectal cancer: From genetic landscape to targeted therapy. *J. Oncol.* **2021**, 1–17 (2021).
3. Vega, P., Valentín, F. & Cubiella, J. Colorectal cancer diagnosis: Pitfalls and opportunities. *World J. Gastrointest. Oncol.* **7**, 422–433 (2015).
4. Park, J., Jeong, D., Song, M. & Kim, B. Recent advances in anti-metastatic approaches of herbal medicines in 5 major cancers: From traditional medicine to modern drug discovery. *Antioxidants* **10**, 527 (2021).
5. Yin, S. Y., Wei, W. C., Jian, F. Y. & Yang, N. S. Therapeutic applications of herbal medicines for cancer patients. *Evid. Based Complement. Altern. Med.* **2013**, 1–15 (2013).
6. Xia, X. Bioinformatics and drug discovery. *Curr. Top. Med. Chem.* **17**, 1709–1726 (2017).
7. Hastings, J. *et al.* ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res.* **44**, D1214–D1219 (2016).
8. Wang, X. *et al.* PharmMapper 2017 update: A web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.* **45**, W356–W360 (2017).
9. Assenov, Y., Ramirez, F., Schelhorn, S. E. S. E., Lengauer, T. & Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **24**, 282–284 (2008).
10. Omer, A. & Singh, P. An integrated approach of network-based systems biology, molecular docking, and molecular dynamics approach to unravel the role of existing antiviral molecules against AIDS-associated cancer. *J. Biomol. Struct. Dyn.* **35**, 1547–1558 (2017).
11. Forli, S. *et al.* Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **11**, 905–919 (2016).
12. Boyle, N. M. O. *et al.* Open Babel: An open chemical toolbox. *J. Cheminform.* **3**(1), 1–14 (2011).
13. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461 (2011).
14. Ahmad, S., Abbasi, H. W., Shahid, S., Gul, S. & Abbasi, S. W. Molecular docking, simulation and MM-PBSA studies of nigella sativa compounds: A computational quest to identify potential natural antiviral for COVID-19 treatment. *J. Biomol. Struct. Dyn.* **39**, 4225–4233 (2021).
15. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

16. Allouche, A. Force field-dependant structural divergence revealed during long time simulations of Calbindin d9k. *J. Comput. Chem.* **32**, 174–182 (2012).
17. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
18. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
19. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
20. Homeyer, N. & Gohlke, H. Free energy calculations by the molecular mechanics Poisson–Boltzmann surface area method. *Mol. Inform.* **31**, 114–122 (2012).
21. Chandarakesan, A., Muruhan, S. & Sayanam, R. R. A. Morin inhibiting photocarcinogenesis by targeting ultraviolet-B-induced oxidative stress and inflammatory cytokines expression in swiss albino mice. *Int. J. Nutr. Pharmacol. Neurol. Dis.* **8**, 41–46 (2018).
22. Hasan, M. K., Ara, I., Mondal, M. S. A. & Kabir, Y. Phytochemistry, pharmacological activity, and potential health benefits of *Glycyrrhiza glabra*. *Heliyon* **7**, e07240 (2021).
23. He, S., Liu, H., Zhou, Y. & Yue, Q. Liquiritin (LT) exhibits suppressive effects against the growth of human cervical cancer cells through activating Caspase-3 in vitro and xenograft mice in vivo. *Biomed. Pharmacother.* **92**, 215–228 (2017).
24. Yang, L., Zhang, Y. & Feng, J. Liquiritin represses proliferation, migration and invasion of colorectal cancer cells through inhibition of the miR-671/HOXB3 signaling pathway, pp. 1–26 (2020).
25. Zhou, Y. & Ho, W. S. Combination of liquiritin, isoliquiritin and isoliquiritigenin induce apoptotic cell death through upregulating p53 and p21 in the A549 non-small cell lung cancer cells. *Oncol. Rep.* **31**, 298–304 (2014).
26. Wei, F., Jiang, X., Gao, H. Y. & Gao, S. H. Liquiritin induces apoptosis and autophagy in cisplatin (DDP)-resistant gastric cancer cells in vitro and xenograft nude mice in vivo. *Int. J. Oncol.* **51**, 1383–1394 (2017).
27. Amernic, J. Perspectives on the role of isoliquiritigenin in cancer. *Routledge Companion to Account. Commun.* <https://doi.org/10.4324/9780203593493-16> (2013).
28. Status, C. Liquiritin represses proliferation, migration and invasion of colorectal cancer cells through inhibition of the miR-671/HOXB3 signaling pathway, pp. 1–31 doi:<https://doi.org/10.21203/rs.3.rs-16350/v1>.
29. Meng, F. C. & Lin, J. K. Liquiritigenin inhibits colorectal cancer proliferation, invasion, and epithelial-to-mesenchymal transition by decreasing expression of runt-related transcription factor 2. *Oncol. Res.* **27**, 139–146 (2019).
30. Soo, H. C. *et al.* Cudraflavone C induces tumor-specific apoptosis in colorectal cancer cells through inhibition of the phosphoinositide 3-kinase (PI3K)-AKT pathway. *PLoS ONE* **12**, 1–20 (2017).
31. Lee, C. W. *et al.* Cudraflavone C induces apoptosis of A375.S2 melanoma cells through mitochondrial ROS production and MAPK activation. *Int. J. Mol. Sci.* **18**, 1508 (2017).
32. Lee, H. J. *et al.* Growth inhibition and apoptosis-inducing effects of cudraflavone B in human oral cancer cells via MAPK, NF- $\kappa$ B, and SIRT1 signaling pathway. *Planta Med.* **79**, 1298–1306 (2013).
33. Bian, B. *et al.* Cathepsin B promotes colorectal tumorigenesis, cell invasion, and metastasis. *Mol. Carcinog.* **55**, 671–687 (2016).
34. Peng, S. *et al.* CTSB knockdown inhibits proliferation and tumorigenesis in HL-60 cells. *Int. J. Med. Sci.* **18**, 1484–1491 (2021).
35. Chao, A. & Wang, T. H. Molecular mechanisms for synergistic effect of proteasome inhibitors with platinum-based therapy in solid tumors. *Taiwan. J. Obstet. Gynecol.* **55**, 3–8 (2016).

## Acknowledgements

We would like to express our gratitude to the Shahrekord University of Medical Sciences for financial support (Grant Number: SKUMS-3828). The founder had no role in the study design, data collection, analysis, decision to publish, or manuscript preparation. In addition, the authors acknowledge and appreciate the anonymous reviewers for their thoughtful comments, which helped improve the quality of the article.

## Author contributions

A.A., R.M.: implementation, methodology. S.A.H.: manuscript writing, reviewing and editing. H.Y.: Conceptualization, supervision the project, writing, reviewing and editing. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-38134-5>.

**Correspondence** and requests for materials should be addressed to H.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023