GMR
The Original

IN-SILICO IDENTIFICATION OF SHARED MOLECULAR BIOMARKERS AND CANDIDATE THERAPEUTICS LINKING OBESITY AND CARDIOVASCULAR DISEASE USING DIFFERENTIAL GENE EXPRESSION AND PATHWAY ANALYSIS

Ali A. Zaeri

Department of Laboratory Medicine, Faculty of Applied Medical Science, Al-Baha University, 65779, Saudi Arabia. Email: Azaeri@bu.edu.sa, ORCID: 0000-0003-4377-1457

ABSTRACT

Obesity and cardiovascular disease (CVD) are closely interconnected disorders driven by chronic inflammation, oxidative stress, endothelial dysfunction, and immune-metabolic dysregulation. However, the shared molecular biomarkers and therapeutically relevant targets linking these conditions remain insufficiently characterized. Publicly available GEO datasets were analyzed to detect differentially expressed genes (DEGs), followed by hub gene prioritization using ReactomeFI, cytoHubba, and MCODE. Functional interpretation was performed through Gene Ontology and KEGG pathway enrichment analyses. Candidate ligands with anti-inflammatory and antioxidant activity were further screened through ADMET profiling, molecular docking, molecular dynamics simulation, MM/GBSA energy estimation, and HOMO–LUMO analysis. A total of 846 downregulated and 1002 upregulated genes were identified in the CVD datasets, whereas the obesity datasets yielded 865 downregulated and 1564 upregulated genes. Integration of network-based analyses identified five shared candidate biomarkers between the two conditions: the downregulated genes JUN and EGFR, and the upregulated genes EP300, PIK3CG, and CCR2. Enrichment analysis indicated that these genes were associated with inflammation-related and signaling pathways relevant to cardiometabolic dysfunction. Among the screened phytochemicals, Ginkgolide A and Broussoflavan A showed favorable ADMET properties and emerged as the leading compounds. Docking analysis demonstrated target-dependent affinity, with Ginkgolide A showing the strongest interaction with PIK3CG (−9.3 kcal/mol) and Broussoflavan A showing the strongest interaction with CCR2 (−9.5 kcal/mol). Molecular dynamics simulation further supported these preferences, revealing greater structural stability of the PIK3CG–Ginkgolide A and CCR2–Broussoflavan A complexes over 200 ns. MM/GBSA analysis confirmed these findings, with binding free energies of −27.66 ± 1.52 kcal·mol^{-1} for PIK3CG–Ginkgolide A and −28.72 ± 3.79 kcal·mol^{-1} for CCR2–Broussoflavan A. Overall, this study identifies a concise shared biomarker signature linking obesity and CVD and proposes PIK3CG and CCR2 as promising therapeutic targets. The findings further suggest that Ginkgolide A and Broussoflavan A may represent potential lead compounds for the development of phytochemical-based interventions in obesity-associated cardiovascular disease.

.

KEYWORDS: Obesity; cardiovascular disease; shared biomarkers; transcriptomics; differential gene expression; PIK3CG; CCR2; Ginkgolide A; Broussoflavan A; molecular docking; molecular dynamics simulation; phytochemical therapeutics.

INTRODUCTION

Obesity is a chronic, multifactorial metabolic disorder characterized by excessive adipose tissue accumulation that impairs physiological homeostasis and increases the risk of systemic complications [1]. In adults, obesity is commonly defined by a body mass index (BMI) of at least 30 kg/m² [2]. Cardiovascular disease (CVD) is an umbrella term for disorders of the heart and blood vessels, including coronary artery disease, cerebrovascular disease, heart failure, and peripheral vascular disease [3]. Although obesity has long been recognized as a major modifiable risk factor for CVD, the molecular mechanisms that biologically connect these conditions remain incompletely understood [4]. Increasing evidence indicates that obesity promotes cardiovascular injury through chronic low-grade inflammation, oxidative stress, endothelial dysfunction, adipokine imbalance, and immune-metabolic dysregulation, all of which contribute to vascular remodeling and atherogenesis [5]. Among the most frequently reported shared biomarkers linking obesity and cardiovascular pathology are inflammatory cytokines and endothelial activation markers, including interleukin-6 (IL-6), tumor necrosis factor-alpha (TNF-α), classical inflammatory mediators such as *CCR2, ICAM-1, and VCAM-1, PIK3CG* deserves special attention because it integrates inflammatory and metabolic signaling through the *PI3K/AKT* axis and may contribute to leukocyte recruitment, endothelial dysfunction, and progression of obesity-related cardiovascular injury [6]. These molecules do not merely reflect disease burden; they actively participate in leukocyte recruitment,

vascular inflammation, extracellular matrix remodeling, nitric oxide dysregulation, and plaque instability [7]. Current literature increasingly suggests that future biomarker discovery should move beyond single-analyte approaches toward integrative multi-omics frameworks that combine transcriptomics, proteomics, metabolomics, and network analysis to identify robust blood-based biomarkers with higher diagnostic and therapeutic relevance [8].

From a therapeutic perspective, the most biologically compelling upregulated targets are proteins that occupy central positions in shared inflammatory and endothelial pathways [9]. In particular, IL-6, TNF-α, CCR2, ICAM-1, VCAM-1, and PIK3CG are attractive candidates because they lie at the intersection of adipose inflammation, endothelial activation, monocyte trafficking, and vascular lesion progression [10]. Upregulation of these proteins has been associated with impaired nitric oxide bioavailability, activation of NF-κB and JAK/STAT signaling, enhanced oxidative stress, and progressive atherosclerotic remodeling [11].

Phytochemicals have emerged as particularly interesting therapeutic candidates because they are often multi-target agents capable of modulating several of these dysregulated pathways simultaneously [12]. Polyphenols and plant-derived bioactives such as curcumin, quercetin, resveratrol, and berberine have demonstrated anti-inflammatory, antioxidant, endothelial-protective, and anti-atherogenic effects in experimental and clinical studies [13].

Previous biomarker studies in obesity, atherosclerosis, diabetes, metabolic syndrome, and other inflammatory disorders have already demonstrated the value of integrated transcriptomic and network-based approaches for identifying disease-associated molecular signatures [14]. Recent bioinformatic studies have specifically examined shared genetic mechanisms between obesity and atherosclerosis, supporting the idea that these conditions converge on common immune-inflammatory and endothelial pathways [15]. However, important gaps remain. Many studies focus on a single disease, a single tissue type, or isolated biomarkers rather than systematically identifying overlapping differentially expressed genes, shared pathways, clinically relevant hub proteins, and candidate therapeutics within a unified framework [16]. Furthermore, relatively few studies integrate biomarker discovery with therapeutic prioritization, particularly the repositioning of phytochemical or drug-like compounds against shared molecular targets [17].

In this context, an in-silico comparative transcriptomic approach offers a rational and cost-effective strategy to uncover common differentially expressed genes and signaling pathways linking obesity and cardiovascular disease. By integrating differential gene expression analysis, functional enrichment, protein-protein interaction mapping, and therapeutic target prioritization, it becomes possible to identify biomarker candidates with both mechanistic and translational relevance. Such an approach may reveal not only shared molecular determinants of disease progression but also novel opportunities for intervention through targeted inhibition of overexpressed proteins and pathway modulation by phytochemical agents. Therefore, this study is expected to contribute to the growing field of precision cardiometabolic medicine by clarifying the molecular overlap between obesity and CVD and by nominating candidate biomarkers and therapeutics for future validation

METHOD & MATERIALS

GEO dataset search

This study was conducted as an in silico transcriptomic analysis to identify shared biomarkers and candidate therapeutic targets between obesity and cardiovascular disease (CVD). Human gene expression datasets were obtained from the NCBI Gene Expression Omnibus (GEO) database based on disease relevance, case-control design, and suitability for differential expression analysis[18]. For CVD, the datasets GSE179789, GSE66360, and GSE34822 were selected. For obesity-related analysis, GSE18897, GSE128021, and GSE185011 were included (Table 1). Each dataset was analyzed independently using standard differential expression methods. Significant differentially expressed genes (DEGs) were identified using an adjusted p value < 0.05 and $|\log2 \text{ fold change}| \geq 1.0$. Shared DEGs between obesity-related and CVD datasets were then identified by overlap analysis. These common genes were further subjected to functional enrichment, pathway analysis, and protein-protein interaction network analysis to prioritize candidate biomarkers and therapeutic targets.

Table 1 summarizes the GEO datasets included in the present study.

GSE number	Title	Samples	Reference
GSE179789	Expression data of 8 healthy people and 8 patients with coronary heart disease before and after 8 weeks treatment with Danlou tablets	8 healthy controls; 8 CHD patients	[19]
GSE66360	A Whole Blood Molecular Signature for the Identification of Acute Myocardial Infarction Without Relying Upon Myonecrosis	50 healthy controls; 49 AMI patients	[20]
GSE34822	Transcriptome Analysis in Patients with Progressive Coronary Artery Disease	16 progressive CAD; 16 stable CAD	[21]
GSE18897	Whole blood expression profiling of obese diet-sensitive,	20 lean subjects; 20 obese	[22]

GSE number	Title	Samples	Reference
	obese diet-resistant and lean human subjects	subjects with repeated time-point sampling	
GSE128021	Gene expression differences in omental mesothelial cells from lean and obese human donors	3 replicate pooled lean samples; 3 replicate pooled obese samples	N/A
GSE185011	The expression of mRNA and lncRNA in peripheral blood mononuclear cells (PBMCs) of diabetes mellitus, diabetic retinopathy, diabetic peripheral neuropathy and diabetic nephropathy patients	Healthy control, T2DM, DR, DPN and DN PBMC samples	[23]

Analysis of significant DEGs for biomarker identification

Significant differentially expressed genes (DEGs) were identified in R from the TSV-formatted dataset according to their upregulated and downregulated expression patterns. The resulting DEG dataset contained both gene symbols and gene names for all significant transcripts. Gene interaction analysis was then performed using ReactomeFI in Cytoscape version 3.10.0[24,25]. To identify key hub genes, the cytoHubba plug-in was applied using 12 topological ranking methods. In addition, the MCODE algorithm in Cytoscape was used to detect highly interconnected gene clusters. Hub genes supported by at least two cytoHubba ranking methods, together with genes identified within significant MCODE clusters, were considered potential biomarkers, thereby improving the robustness and reliability of the analysis.

GO and pathway enrichment analysis

Gene Ontology (GO) and pathway enrichment analyses of the candidate biomarker genes were performed using the web-based server Weishengxin Bioinformatics (https://bioinformatics.com.cn/)[26]. GO enrichment was conducted under the three standard categories: biological process, cellular component, and molecular function. Pathway enrichment analysis was carried out using the KEGG library available through the Enricher server. Statistical significance was defined as an adjusted p value < 0.05, and correction for multiple testing was performed using the Benjamini-Hochberg method.

Significant biomarker identification and protein preparation

The most significant candidate biomarkers were selected from both the upregulated and downregulated gene sets based on the previously described screening approach. Proteins corresponding to the prioritized upregulated biomarker genes were retrieved from the UniProt database using their gene symbols and gene names. The three-dimensional protein structures were then obtained from the RCSB Protein Data Bank (RCSB PDB) in PDB format[27,28]. Prior to downstream analysis, all bound ligands, water molecules, and other heteroatoms were removed using PyMOL, and the cleaned protein structures were saved again in PDB format.

Ligand selection and preparation

As the present study focused on cardiovascular disease, ligands were selected based on their reported anti-inflammatory and antioxidant biological activities. Candidate phytochemicals were identified from Dr. Duke's Phytochemical and Ethnobotanical Databases [29]. The corresponding SMILES notations and Compound IDs (CIDs) for the selected compounds were then retrieved from the PubChem database for further computational analysis [30].

ADME and toxicity prediction

The pharmacokinetic properties of the selected ligands were evaluated using the SwissADME web server [31]. Parameters including blood-brain barrier (BBB) permeation, gastrointestinal (GI) absorption, and Lipinski's rule of five were assessed to determine drug-likeness. Compounds showing high GI absorption, acceptable pharmacokinetic behavior, and zero Lipinski rule violations were considered favorable for further analysis. Toxicity prediction was subsequently performed using the pkCSM platform [32]. The selected ligands were screened for AMES toxicity, hepatotoxicity, and skin sensitization, and compounds predicted to be non-AMES toxic, non-hepatotoxic, and non-skin-sensitizing were prioritized.

Molecular docking analysis

Docking analysis was conducted in PyRx using the AutoDock Vina application (vina.exe) to assess ligand-protein interactions [33,34]. Docking was performed with an exhaustiveness value of 8. For one target protein, the grid box center was set to (35.2311, -20.2535, 26.4866) with dimensions of 106.3894 × 91.6845 × 72.9460 Å. For the other target protein, the grid center was defined as (9.1949, 20.2424, 187.3155) with dimensions of 48.6133 × 70.0411 × 97.6276 Å. The best binding conformations were selected according to the docking scores and subsequently visualized in PyMOL for further analysis.

Molecular Dynamics Simulation

All-atom molecular dynamics (MD) simulations were performed using GROMACS 2024.5 to investigate the structural stability and dynamic behavior of the protein-ligand complexes [35]. The CHARMM36 force field was applied for the protein, while ligand parameters were generated using CGenFF to maintain force-field compatibility [36]. Each complex was placed in a triclinic TIP3P water box with a minimum distance of 1.0 nm from the box edges, followed by system neutralization with Na^+/Cl^- ions and adjustment to a final salt

concentration of 0.15 M NaCl. Energy minimization was first carried out using the steepest-descent algorithm, after which the systems were equilibrated under NVT and NPT conditions for 100 ps each. A 200 ns production simulation was then conducted with a 2 fs integration timestep, and trajectory coordinates were saved every 10 ps. Trajectory analyses were performed using standard GROMACS utilities to determine root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), per-residue SASA, and protein-ligand contact patterns.

To further characterize collective structural motions, principal component analysis (PCA) and dynamic cross-correlation matrix (DCCM)/covariance analysis were carried out on the backbone atoms. In addition, binding free energies were estimated after simulation using gmx_MMPBSA from representative snapshots extracted from the production trajectories. Overall, these analyses provided an integrated assessment of complex stability, interaction energetics, and ligand-induced conformational changes.

HOMO LUMO molecular orbital profiling

Density Functional Theory (DFT) calculations were performed for the top-ranked ligands to examine their electronic characteristics related to stability and reactivity [37]. Geometry optimization and orbital calculations were carried out using the B3LYP/6-31+G(d,p) level of theory [38]. HOMO and LUMO energies were determined, and the energy gap was calculated as E_LUMO − E_HOMO. To examine the influence of binding-induced conformational changes, these parameters were evaluated for ligand structures before MD simulation and for representative conformations obtained after the 200 ns MD run. The resulting values were compared to assess changes in electronic stability and reactivity after ligand accommodation in the binding pocket[39–41].

RESULT

Identification of Biomarker genes

A total of 846 downregulated and 1002 upregulated genes were identified from the CVD datasets, whereas 865 downregulated and 1564 upregulated genes were obtained from the obesity datasets. These significant genes were subsequently subjected to hub gene screening using 12 cytoHubba topological methods in Cytoscape. In parallel, MCODE analysis was performed to detect highly interconnected clusters within the gene interaction networks. From the upregulated gene sets, MCODE identified a total of eight clusters, while ten clusters were retrieved from the downregulated gene sets. These clustered genes were considered potential biomarker candidates and were further compared with the cytoHubba-ranked genes to identify the most robust shared biomarkers between CVD and obesity.

By integrating the cytoHubba ranking outputs with MCODE cluster analysis from both the cardiovascular disease (CVD) and obesity networks, a set of shared hub biomarkers was identified. For CVD 4 down regulate genes have been found and 8 up regulated, for the obese dataset 7 down regulated and 5 up regulated genes have found (Table S1). Cross-comparison of the two disease groups showed that JUN and EGFR were common downregulated biomarkers, whereas EP300, PIK3CG, and CCR2 were common upregulated biomarkers. These overlapping genes were therefore considered the final candidate biomarkers linking obesity and CVD (Table 1).

Table 1. Shared candidate biomarkers identified from combined cytoHubba and MCODE analyses of CVD and obesity datasets.

Biomarker	Regulation in CVD	Regulation in obesity	cytoHubba support in CVD	cytoHubba support in obesity	Final status
JUN	Downregulated	Downregulated	10 methods	5 methods	Shared biomarker
EGFR	Downregulated	Downregulated	8 methods	5 methods	Shared biomarker
EP300	Upregulated	Upregulated	10 methods	10 methods	Shared biomarker
PIK3CG	Upregulated	Upregulated	4 methods	3 methods	Shared biomarker
CCR2	Upregulated	Upregulated	2 methods	10 methods	Shared biomarker

Analysis of GO and Pathway

GO enrichment analysis of the 24 prioritized hub genes demonstrated distinct but related functional patterns in CVD and obesity. A cross-comparison of the CVD and obesity hub gene sets revealed five overlapping biomarkers. Among them, **JUN** and **EGFR** were identified as the common **downregulated** biomarkers, whereas **EP300, PIK3CG, and CCR2** were identified as the common **upregulated** biomarkers. These shared genes were therefore considered the **final candidate biomarkers** linking obesity and CVD. Functionally, **JUN** is involved in transcriptional regulation and cellular stress responses, while **EGFR** is associated with cell signaling and growth regulation. In contrast, **EP300** acts as a transcriptional coactivator, **PIK3CG** participates in inflammatory and PI3K-mediated signaling, and **CCR2** is a key mediator of chemokine-driven immune cell recruitment. The enrichment patterns shown in the figure indicate that these overlapping genes are involved in biologically relevant processes associated with inflammation, signaling regulation, and cardiometabolic dysfunction (Figure 1).

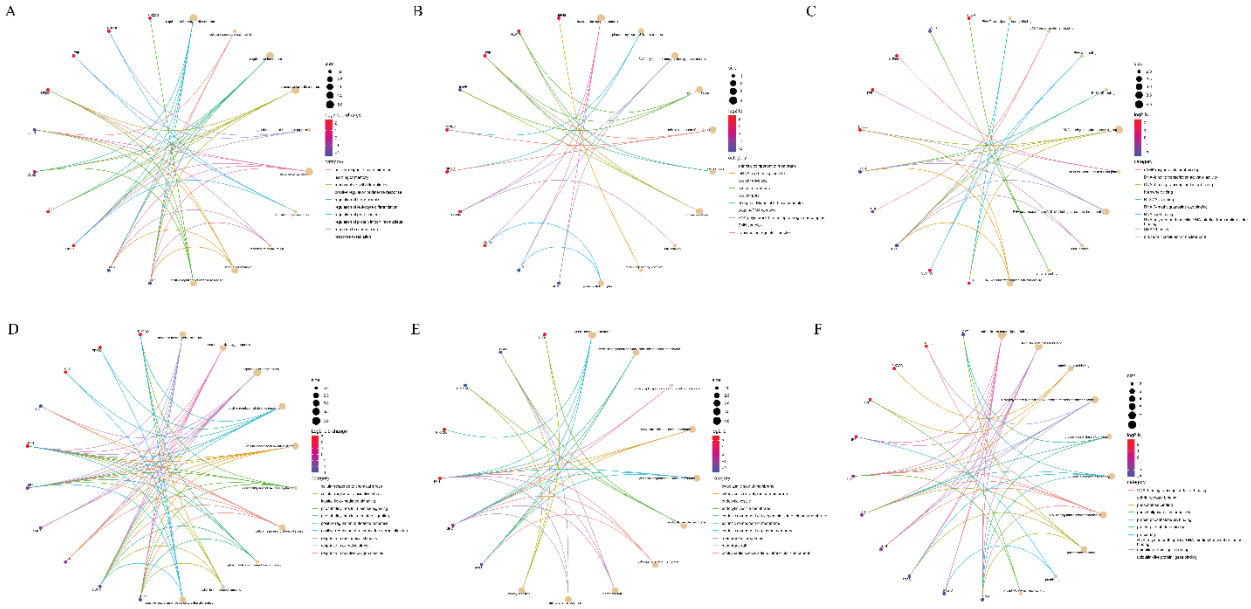


Figure 1. Gene Ontology (GO) enrichment analysis of prioritized hub genes in cardiovascular disease (CVD) and obesity.

A total of 24 prioritized hub genes were subjected to GO enrichment analysis. Panels A-C represent the GO enrichment results for the CVD-associated hub genes, including biological process (A), cellular component (B), and molecular function (C). Panels D-F represent the corresponding GO enrichment results for the obesity-associated hub genes, including biological process (D), cellular component (E), and molecular function (F). In each panel, the left-side nodes indicate the enriched genes, whereas the right-side nodes represent the significantly enriched GO terms. Node color reflects the log2 fold change, and node size indicates the relative significance or contribution of the enriched terms.

In the CVD (Figure 2A), the top enriched pathways included relaxin signaling pathway, estrogen signaling pathway, hepatitis B, colorectal cancer, endocrine resistance, choline metabolism in cancer, chemical carcinogenesis-receptor activation, prostate cancer, growth hormone synthesis, secretion and action, and breast cancer. In the obesity (Figure 2B), the most significantly enriched pathways included Kaposi sarcoma-associated herpesvirus infection, shigellosis, hepatitis B, endocrine resistance, mitophagy-animal, thyroid hormone signaling pathway, pancreatic cancer, chemokine signaling pathway, ErbB signaling pathway, and focal adhesion.

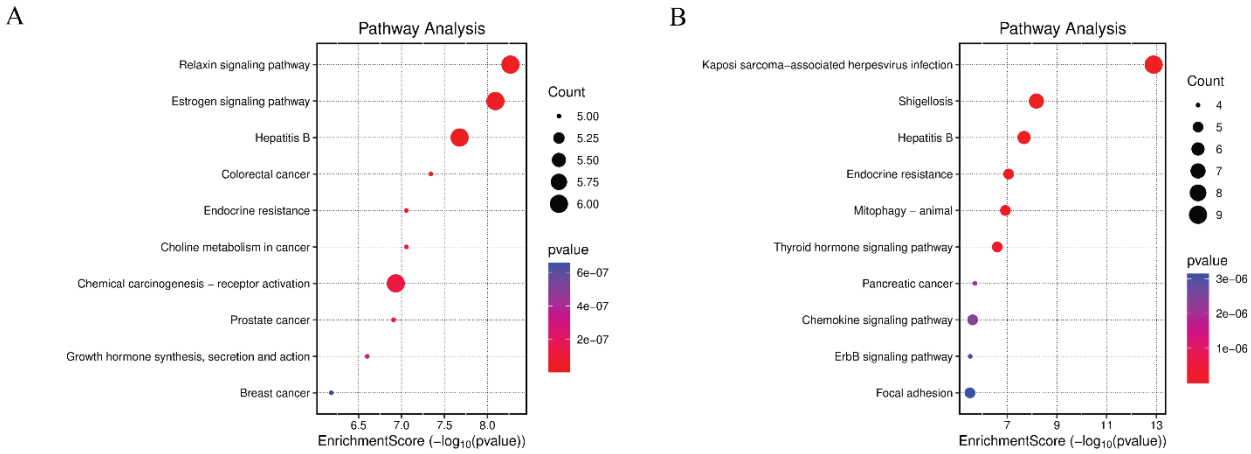


Figure 2. KEGG pathway enrichment analysis of prioritized hub genes in cardiovascular disease (CVD) and obesity.

Pathway enrichment analysis was performed using the 24 prioritized hub genes identified from the combined network-based screening. Panel A shows the significantly enriched pathways for the CVD-associated genes, whereas Panel B shows the enriched pathways for the obesity-associated genes. In both panels, bubble size represents the number of genes enriched in each pathway, and bubble color indicates the significance level (p value).

Comparative interpretation of the pathway results, together with the GO and network analyses, identified five

overlapping candidate biomarkers linking obesity and CVD. Specifically, JUN and EGFR were the common downregulated biomarkers, whereas EP300, PIK3CG, and CCR2 were the common upregulated biomarkers. These shared genes were therefore considered the final candidate biomarkers, as they are functionally associated with transcriptional regulation, inflammatory signaling, chemokine-mediated immune responses, and cardiometabolic dysfunction.

ADMET Analysis

The ADMET assessment revealed a generally favorable pharmacokinetic and toxicity profile for the screened phytochemicals (Table 2). All compounds demonstrated high gastrointestinal (GI) absorption, suggesting good oral absorption potential. None of the compounds were predicted to be blood-brain barrier (BBB) permeant, indicating a low likelihood of central nervous system exposure. In addition, all screened ligands showed zero Lipinski violations, supporting their overall drug-likeness and compliance with key physicochemical criteria for oral bioavailability.

Table 2. ADMET and drug-likeness properties of the selected phytochemicals.

Chemical	CID	Formula	GI absorption	BBB permeant	Lipinski #violations	AMES toxicity	Hepatotoxicity
Alpha-Lipoic-Acid	864	C8H14O2S2	High	No	0	No	No
Tryptophan	6305	C11H12N2O2	High	No	0	No	No
Catechin	9064	C15H14O6	High	No	0	No	No
Sinapic-Acid	10743	C11H12O5	High	No	0	No	No
(-)-Epicatechin	72276	C15H14O6	High	No	0	No	No
6,7,4'-Trihydroxyisoflavanone	125100	C15H12O5	High	No	0	No	No
Taxifolin	439533	C15H12O7	High	No	0	No	No
Artecanin	442147	C15H18O5	High	No	0	Yes	No
Cnicin	5281435	C20H26O7	High	No	0	No	No
Cimiracemate-A	5315874	C19H18O7	High	No	0	No	No
Andrographolide	5318517	C20H30O5	High	No	0	No	No
Marmin	6450230	C19H24O5	High	No	0	Yes	No
Ginkgolide-A	9909368	C20H24O9	High	No	0	No	No
Broussoflavan-A	44257178	C25H30O6	High	No	0	No	No
S-Allyl-Cysteine-Sulfoxide	129668924	C6H11NO3S2	High	No	0	No	No

Toxicity prediction further showed that the majority of the compounds possessed acceptable safety characteristics (Table 2). Among the 15 screened phytochemicals, only Artecanin and Marmin were predicted to be toxic AMES, whereas the remaining compounds were classified as non-mutagenic. Notably, none of the analyzed ligands showed predicted hepatotoxicity, indicating a favorable hepatic safety profile across the dataset. These results suggest that most of the screened compounds combine suitable absorption properties with low predicted toxicity risk.

Importantly, the two lead compounds identified from docking and simulation analyses, Ginkgolide A and Broussoflavan A, also exhibited favorable ADMET characteristics (Table 2). Both compounds showed high GI absorption, no BBB permeability, no Lipinski rule violations, no AMES toxicity, and no hepatotoxicity, indicating that they possess not only strong binding potential toward their respective targets but also promising drug-likeness and safety profiles. Overall, the ADMET findings support the further consideration of Ginkgolide A and Broussoflavan A as potential lead candidates for subsequent experimental validation.

Molecular Docking

Molecular docking against PIK3CG (PDB ID: 5JHB) and CCR2 (PDB ID: 5T1A) revealed a clear target-dependent binding preference for the two lead phytochemicals (Table 3). For PIK3CG, Ginkgolide A showed the most favorable docking score at −9.3 kcal/mol, whereas Broussoflavan A bound with a lower affinity of −8.6 kcal/mol (Table 3). This indicates a 0.7 kcal/mol advantage for Ginkgolide A over Broussoflavan A in the 5JHB pocket. In addition, Ginkgolide A outperformed the next-best docked compounds for PIK3CG, which clustered at −8.1 kcal/mol, further supporting its preferential affinity toward this target (Table 3). In contrast, the binding trend was reversed for CCR2, where Broussoflavan A achieved the best docking score of −9.5 kcal/mol, compared with −9.2 kcal/mol for Ginkgolide A (Table 3). Although both ligands exhibited strong affinity for 5T1A, Broussoflavan A retained a modest energetic advantage of 0.3 kcal/mol and also exceeded the next-best compounds, which showed docking scores of −8.4 kcal/mol (Table 3). Collectively, these docking

data indicate that Ginkgolide A is more favorable for PIK3CG, whereas Broussoflavan A shows superior affinity for CCR2.

Table 3. Molecular docking scores of the screened phytochemicals against PIK3CG (PDB ID: 5JHB) and CCR2 (PDB ID: 5T1A).

Compound	CID	Binding affinity against PIK3CG (5JHB) (kcal/mol)	Binding affinity against CCR2 (5T1A) (kcal/mol)
Alpha-lipoic acid	864	-5.0	-5.0
Tryptophan	6305	-6.5	-6.9
Catechin	9064	-7.9	-7.8
Sinapic acid	10743	-6.1	-5.6
(-)-Epicatechin	72276	-7.8	-7.8
6,7,4'-Trihydroxyisoflavanone	125100	-8.1	-7.5
Taxifolin	439533	-8.1	-7.9
Artecanin	442147	-8.1	-7.3
Cnicin	5281435	-7.4	-8.4
Cimiracemate-A	5315874	-7.4	-8.0
Andrographolide	5318517	-8.0	-8.0
Marmin	6450230	-7.8	-8.4
Ginkgolide A	9909368	-9.3	-9.2
Broussoflavan A	44257178	-8.6	-9.5
S-allyl-cysteine-sulfoxide	129668924	-4.5	-4.4

The 2D interaction analysis provided structural support for these docking preferences. Within the PIK3CG-Broussoflavan A complex, the ligand formed conventional hydrogen bonds with Arg849, Gln291, and His658, together with additional hydrophobic contacts involving Lys298 and Pro866 (Figure 3A). By comparison, the PIK3CG-Ginkgolide A complex engaged residues Arg277, Leu791, Ile828, Tyr787, Trp292, and His295, indicating a binding mode dominated by a broader combination of polar, π-mediated, and hydrophobic contacts (Figure 3B). Despite the greater number of conventional hydrogen-bonding interactions observed for Broussoflavan A in PIK3CG, Ginkgolide A still exhibited the stronger docking score, suggesting that the overall packing efficiency and nonbonded complementarity within the 5JHB binding pocket were more favorable for Ginkgolide A (Figure 3A,B; Table 3).

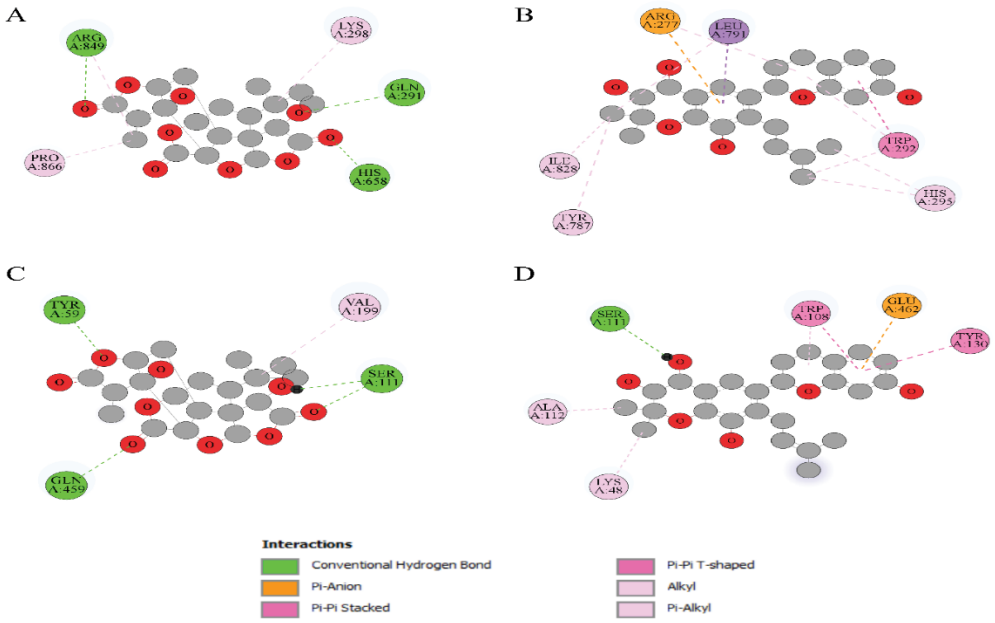


Figure 3. Two-dimensional interaction diagrams of the top-ranked ligand-protein complexes identified from molecular docking analysis against PIK3CG and CCR2. (A) Binding interaction profile of Broussoflavan A with PIK3CG (PDB ID: 5JHB). (B) Binding interaction profile of Ginkgolide A with PIK3CG (PDB ID: 5JHB). (C) Binding interaction profile of Broussoflavan A with CCR2 (PDB ID: 5T1A). (D) Binding interaction profile of Ginkgolide A with CCR2 (PDB ID: 5T1A). The diagrams illustrate the key intermolecular contacts formed within the binding pocket, including conventional hydrogen bonds, π-anion, π-π stacked, π-π T-shaped, alkyl, and π-alkyl interactions.

A similar comparative pattern was observed for CCR2. The CCR2-Broussoflavan A complex was stabilized by conventional hydrogen bonds with Tyr59, Ser111, and Gln459, along with an additional hydrophobic interaction involving Val199 (Figure 3C). In contrast, the CCR2-Ginkgolide A complex interacted with Ser111, Trp108, Glu462, Tyr130, Ala112, and Lys48, reflecting a mixed network of hydrogen-bonding, π-associated, and hydrophobic contacts (Figure 3D). Nevertheless, the lower docking score of Broussoflavan A against CCR2 indicates that its interaction pattern was more energetically favorable within the 5T1A cavity (Figure 3C, D; Table 3). Overall, the docking results, together with the 2D interaction profiles, consistently demonstrate a protein-specific ligand preference, with Ginkgolide A favored for PIK3CG and Broussoflavan A favored for CCR2.

MD Simulation analysis

Molecular dynamics (MD) simulation was conducted to evaluate the time-dependent stability and flexibility of the investigated protein-ligand complexes under dynamic conditions. In MD analysis, the root mean square deviation (RMSD) is used to assess the overall conformational stability of a complex throughout the simulation period, whereas the root mean square fluctuation (RMSF) reflects residue-wise flexibility and helps identify structurally mobile regions. In the present study, RMSD and RMSF analyses revealed distinct ligand-dependent stability patterns for both PIK3CG and CCR2 over the 200 ns simulation period. To further visualize the docking orientation of the lead compounds within the target binding pockets, the top two docked ligands for each receptor were superimposed in three-dimensional space. In the PIK3CG binding cavity, Ginkgolide A and Broussoflavan A occupied a common ligand-binding region, but the pose of Ginkgolide A appeared more favorably accommodated, consistent with its superior docking score against PIK3CG (−9.3 kcal/mol vs. −8.6 kcal/mol). In contrast, both ligands also localized within the CCR2 binding pocket, where Broussoflavan A displayed the more favorable binding orientation, in agreement with its stronger docking score relative to Ginkgolide A (−9.5 kcal/mol vs. −9.2 kcal/mol). These structural overlays support the observed target-specific ligand preference, with Ginkgolide A favored for PIK3CG and Broussoflavan A favored for CCR2.

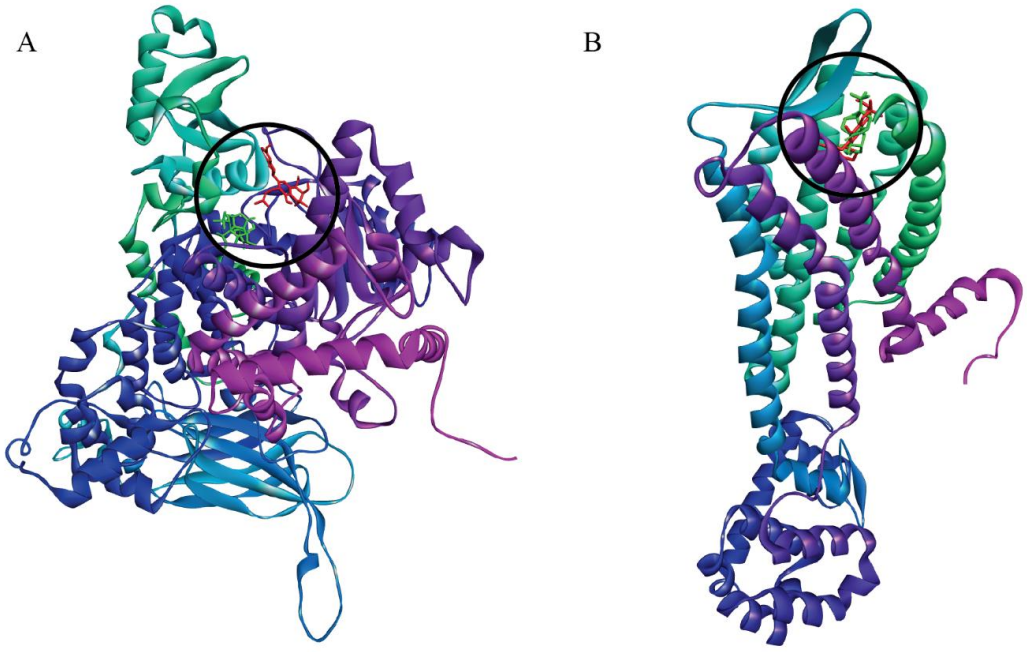


Figure 4. Three-dimensional visualization of the binding pockets of PIK3CG and CCR2 with the top two docked ligands. (A) Superimposed docked conformations of Ginkgolide A and Broussoflavan A within the binding pocket of PIK3CG. (B) Superimposed docked conformations of Ginkgolide A and Broussoflavan A within the binding pocket of CCR2.

The circled regions indicate the ligand-binding cavities of the respective proteins. The structural overlays illustrate that both ligands occupy comparable pockets in each target; however, their binding orientations differ according to receptor context. In agreement with the docking analysis, Ginkgolide A showed the most favorable affinity toward PIK3CG, whereas Broussoflavan A exhibited the strongest affinity toward CCR2, supporting a clear target-dependent binding preference. The RMSD trajectories demonstrated marked differences in the dynamic behavior of the complexes. In the PIK3CG system, the complex with Ginkgolide A maintained a highly stable trajectory, with RMSD values remaining largely within 0.2-0.5 nm throughout the simulation and without any pronounced structural drift (Figure 5A). In contrast, the PIK3CG-Broussoflavan A complex exhibited substantial deviation from its initial conformation. After an early increase to approximately 1.2-1.5 nm within the first 20 ns, the trajectory fluctuated between 1.0 and 2.5 nm up to nearly 100 ns, after which a major conformational transition occurred, resulting in a sharp rise in RMSD to approximately 5-10 nm, with

peaks approaching 12 nm between about 115 and 155 ns (Figure 5A). Although partial re-equilibration was observed during the final phase of the simulation, the RMSD remained elevated at approximately 6.4-6.6 nm, indicating markedly lower structural stability relative to the Ginkgolide A-bound complex. An opposite trend was observed for CCR2, where Broussoflavan A displayed considerably greater stability than Ginkgolide A (Figure 5C). The CCR2-Broussoflavan A complex remained compact throughout the simulation, with RMSD values mostly confined to 0.2-0.6 nm, except for a modest transient increase to nearly 0.8 nm around 115-125 ns (Figure 5C). By comparison, the CCR2-Ginkgolide A complex showed pronounced structural instability, with RMSD increasing to about 1.5-2.0 nm by 35-40 ns, followed by repeated large-amplitude deviations, including peaks of approximately 8.5-9.0 nm at 45-55 ns, sustained fluctuations of about 6.0-7.5 nm between 70 and 90 ns, and a maximum exceeding 10 nm near 95-100 ns (Figure 5C). A second unstable phase was also evident between 145 and 170 ns, where the trajectory remained predominantly within 5.5-6.8 nm, before declining to approximately 2.3-2.8 nm at the end of the simulation.

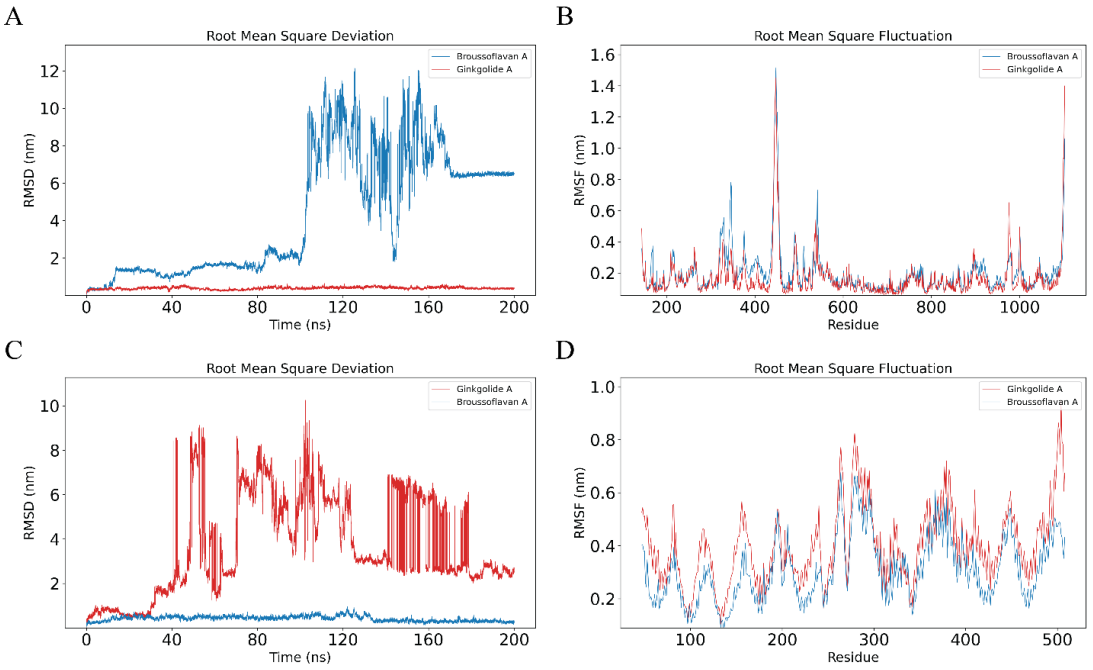


Figure 5. Molecular dynamics-based RMSD and RMSF analyses of PIK3CG and CCR2 in complex with Ginkgolide A and Broussoflavan A over a 200 ns simulation period. (A) RMSD profile of PIK3CG complexed with Ginkgolide A (red) and Broussoflavan A (blue). (B) RMSF profile of PIK3CG residues in the corresponding ligand-bound complexes. (C) RMSD profile of CCR2 complexed with Ginkgolide A (red) and Broussoflavan A (blue). (D) RMSF profile of CCR2 residues in the corresponding ligand-bound complexes. RMSD was used to evaluate the overall structural stability of the protein-ligand complexes during the simulation, whereas RMSF was used to assess residue-wise flexibility and identify locally dynamic regions.

The RMSF profiles provided further insight into residue-level flexibility within these complexes. In the PIK3CG systems, most residues in both ligand-bound forms fluctuated below 0.3 nm, indicating that the majority of the protein remained locally stable during the simulation (Figure 5B). However, several regions exhibited notable flexibility. The highest fluctuation was detected near residue 450, where RMSF reached approximately 1.5 nm in both systems (Figure 5B). Additional peaks were observed around residues 340-360 and 530-550, where the Broussoflavan A complex displayed somewhat greater fluctuations than the Ginkgolide A complex. In contrast, the terminal region near residues 1110-1120 showed a more prominent peak in the Ginkgolide A complex, reaching approximately 1.4 nm (Figure 5B). In the CCR2 systems, the difference between the two ligands was more pronounced (Figure 5D). The Broussoflavan A-bound complex exhibited generally restrained residue fluctuations, with most values lying within 0.1-0.4 nm, whereas the Ginkgolide A-bound complex showed higher overall flexibility, with most residues fluctuating between 0.2 and 0.55 nm (Figure 5D). Distinct differences were observed around residues 265-290 and 360-380, where RMSF in the Ginkgolide A complex increased to approximately 0.8 nm and 0.7 nm, respectively, while the corresponding values for Broussoflavan A remained lower (Figure 5D). The greatest divergence occurred near the C-terminal region around residues 500-510, where RMSF reached nearly 0.9-0.95 nm in the Ginkgolide A complex but remained close to 0.5 nm in the Broussoflavan A complex (Figure 5D).

Overall, comparative analysis of the MD trajectories demonstrated a clear target-dependent ligand preference. Ginkgolide A showed superior dynamic stability toward PIK3CG, as evidenced by its consistently low RMSD and the absence of major conformational drift despite localized residue fluctuations (Figure 5A, B). In contrast, Broussoflavan A exhibited greater conformational compatibility with CCR2, showing both lower global deviation and reduced residue-level flexibility throughout the simulation period (Figure 5C, D). These findings

indicate that the dynamic behavior of the investigated phytochemicals is strongly influenced by the structural environment of the target protein. Molecular dynamics-derived solvent-accessible surface area (SASA) and radius of gyration (Rg) analyses were further used to evaluate changes in solvent exposure and overall structural compactness of the ligand-bound complexes during the 200 ns simulation. In general, SASA provides insight into the extent of protein surface exposed to the solvent environment, whereas Rg reflects the degree of structural compactness of the protein over time. The PIK3CG complexes showed relatively stable SASA profiles throughout the simulation, with only minor fluctuations in both ligand-bound systems (Figure 6A). The Ginkgolide A-bound complex fluctuated largely within approximately 529-541 nm², while the Broussoflavan A-bound complex remained within about 524-539 nm². Although both systems were generally stable, the Ginkgolide A complex tended to maintain a slightly higher SASA profile, suggesting a modestly greater solvent-exposed surface without evidence of major structural disruption. A comparable trend was observed for CCR2, where both complexes exhibited tightly clustered SASA trajectories with no pronounced long-term drift (Figure 6C). The CCR2-Ginkgolide A and CCR2-Broussoflavan A systems remained largely within 242-250 nm² and 241-250 nm², respectively, indicating that ligand binding did not induce marked alterations in solvent accessibility in either case.

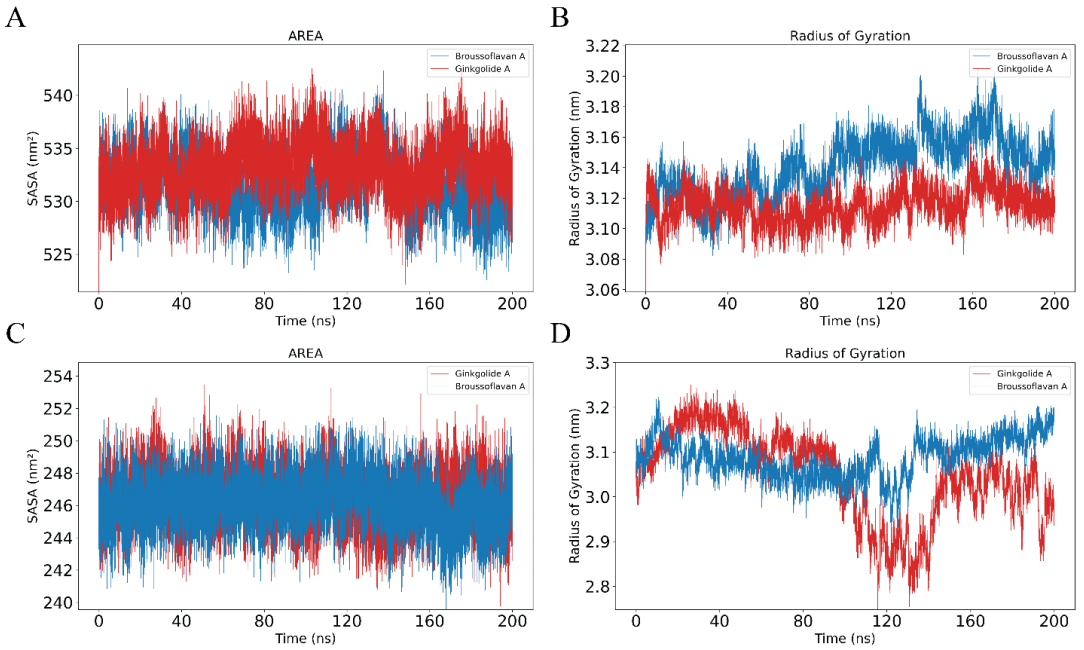


Figure 6. Molecular dynamics-based solvent-accessible surface area (SASA) and radius of gyration (Rg) analyses of PIK3CG and CCR2 in complex with Ginkgolide A and Broussoflavan A during the 200 ns simulation period. (A) SASA profile of PIK3CG in complex with Ginkgolide A (red) and Broussoflavan A (blue). (B) Rg profile of PIK3CG in complex with Ginkgolide A (red) and Broussoflavan A (blue). (C) SASA profile of CCR2 in complex with Ginkgolide A (red) and Broussoflavan A (blue). (D) Rg profile of CCR2 in complex with Ginkgolide A (red) and Broussoflavan A (blue). SASA was used to evaluate changes in the solvent-exposed surface area of the protein-ligand complexes, whereas Rg was used to assess their overall structural compactness during the simulation.

More distinct differences between the ligands were evident from the Rg trajectories. In the PIK3CG system, the Ginkgolide A-bound complex maintained a relatively narrow and stable Rg distribution, largely within 3.09-3.14 nm, with only slight fluctuations over the full simulation period (Figure 6B). By contrast, the Broussoflavan A-bound complex showed a gradual upward shift in Rg after the early stage of the simulation, increasing from approximately 3.11-3.14 nm to about 3.15-3.20 nm, particularly after nearly 70 ns, indicating reduced compactness of the protein structure relative to the Ginkgolide A-bound form (Figure 6B). In the CCR2 system, the difference was even more pronounced (Figure 6D). The Broussoflavan A-bound complex showed a comparatively more consistent Rg pattern, remaining mostly within 3.00-3.18 nm and exhibiting a gradual recovery and stabilization during the later stage of the simulation. In contrast, the Ginkgolide A-bound CCR2 complex displayed substantial fluctuation, with an early increase to approximately 3.16-3.22 nm, followed by a marked decline to nearly 2.80-2.95 nm around 110-135 ns, and continued irregular behavior thereafter (Figure 6D). This pronounced variation suggests considerable conformational rearrangement rather than stable compactness. Taken together, the SASA and Rg data indicate that although both ligands maintained broadly comparable solvent exposure profiles, Ginkgolide A preserved a more compact and structurally consistent conformation in PIK3CG, whereas Broussoflavan A exhibited superior compactness stability in CCR2, supporting the target-specific dynamic behavior observed in the preceding trajectory analyses.

Hydrogen bond analysis was performed to evaluate the persistence of intermolecular polar interactions between

the ligands and their respective target proteins during the 200 ns simulation period. Because stable hydrogen bonding often contributes to stronger ligand retention within the binding pocket, the temporal distribution of hydrogen bonds was examined to further clarify the dynamic behavior of the complexes. In the PIK3CG system, Ginkgolide A maintained a consistently higher number of hydrogen bonds than Broussoflavan A throughout most of the simulation (Figure 7A). The PIK3CG-Ginkgolide A complex generally preserved approximately 1-4 hydrogen bonds, with frequent occupancy at 2-3 hydrogen bonds and intermittent increases to 5-6 hydrogen bonds, indicating persistent and favorable polar interactions over time. In contrast, the PIK3CG-Broussoflavan A complex showed a less stable hydrogen-bonding pattern, with the number of hydrogen bonds remaining predominantly within 0-2, interrupted only occasionally by short-lived increases to 3-5. Several intervals of complete hydrogen bond loss were also observed in the Broussoflavan A-bound system, suggesting weaker and less persistent intermolecular contact formation relative to Ginkgolide A.

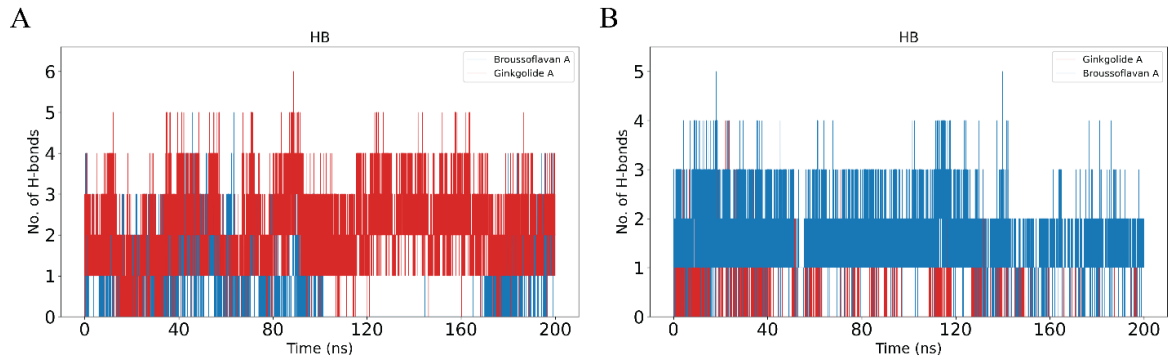


Figure 7. Hydrogen bond (HB) profiles of PIK3CG and CCR2 in complex with Ginkgolide A and Broussoflavan A during the 200 ns molecular dynamics simulation. (A) Number of hydrogen bonds formed between PIK3CG and Ginkgolide A (red) or Broussoflavan A (blue) over time. (B) Number of hydrogen bonds formed between CCR2 and Ginkgolide A (red) or Broussoflavan A (blue) over time. The hydrogen bond analysis was used to assess the persistence of intermolecular polar interactions and the stability of ligand retention within the binding pocket throughout the simulation trajectory.

A reverse trend was observed for CCR2, where Broussoflavan A exhibited more persistent hydrogen-bonding behavior than Ginkgolide A (Figure 7B). The CCR2-Broussoflavan A complex maintained mostly 1-3 hydrogen bonds over the course of the trajectory, with frequent occupancy at 2 hydrogen bonds and occasional increases to 4-5, reflecting sustained ligand-receptor interactions. By comparison, the CCR2-Ginkgolide A complex displayed a notably weaker hydrogen-bonding profile, with the number of hydrogen bonds remaining mostly within 0-1 and only sporadic short-duration increases to 2-4. Thus, comparative analysis of the hydrogen-bond trajectories indicates that Ginkgolide A formed more stable polar interactions with PIK3CG, whereas Broussoflavan A showed stronger hydrogen-bond persistence with CCR2. These findings are in good agreement with the preceding RMSD, RMSF, SASA, and Rg analyses, further supporting a clear target-specific ligand preference.

MM/GBSA Analysis

The MM/GBSA binding free energy analysis further substantiated the target-specific interaction pattern observed during the molecular dynamics simulations (Table 4). Among the PIK3CG complexes, Ginkgolide A exhibited a markedly more favorable binding free energy than Broussoflavan A, with ΔGbind values of −27.66 ± 1.52 and −9.01 ± 6.15 kcal·mol⁻¹, respectively. This stronger binding affinity of Ginkgolide A was primarily driven by more favorable van der Waals and electrostatic contributions, reflected by ΔEvdW and ΔEelec values of −29.19 ± 0.86 and −34.92 ± 0.57 kcal·mol⁻¹, respectively, compared with −16.57 ± 4.98 and −8.33 ± 3.45 kcal·mol⁻¹ for Broussoflavan A. Although the solvation penalty was higher for Ginkgolide A (+36.46 ± 0.74 kcal·mol⁻¹) than for Broussoflavan A (+15.89 ± 0.85 kcal·mol⁻¹), the substantially stronger gas-phase interaction energy of Ginkgolide A (−64.12 ± 1.34 kcal·mol⁻¹) compensated for this unfavorable contribution, resulting in a notably more stable complex with PIK3CG.

Table 4. MM/GBSA-based binding free energy estimation of PIK3CG and CCR2 complexes

Protein / Complex	ΔEvdW (kcal·mol⁻¹)	ΔEelec (kcal·mol⁻¹)	ΔGgas (kcal·mol⁻¹)	ΔGsolv (kcal·mol⁻¹)	ΔGbind (kcal·mol⁻¹)
PIK3CG-Ginkgolide A	−29.19 ± 0.86	−34.92 ± 0.57	−64.12 ± 1.34	+36.46 ± 0.74	−27.66 ± 1.52
PIK3CG-Broussoflavan A	−16.57 ± 4.98	−8.33 ± 3.45	−24.90 ± 6.09	+15.89 ± 0.85	−9.01 ± 6.15
CCR2-Ginkgolide	−10.51 ± 0.20	−2.71 ± 2.08	−13.22 ± 2.22	+5.94 ± 1.63	−7.28 ± 2.76

A					
CCR2-Broussoflavan A	−41.39 ± 0.98	−25.04 ± 3.06	−66.43 ± 3.28	+37.70 ± 1.92	−28.72 ± 3.79

A reverse trend was observed for the CCR2 complexes (Table 4). Broussoflavan A showed a significantly more favorable binding free energy than Ginkgolide A, with ΔGbind values of −28.72 ± 3.79 and −7.28 ± 2.76 kcal·mol⁻¹, respectively. The enhanced binding affinity of Broussoflavan A was associated with strongly favorable van der Waals and electrostatic interactions, with ΔEvdW and ΔEelec values of −41.39 ± 0.98 and −25.04 ± 3.06 kcal·mol⁻¹, respectively, whereas the corresponding values for Ginkgolide A were only −10.51 ± 0.20 and −2.71 ± 2.08 kcal·mol⁻¹. Although Broussoflavan A incurred a larger solvation penalty (+37.70 ± 1.92 kcal·mol⁻¹) than Ginkgolide A (+5.94 ± 1.63 kcal·mol⁻¹), its markedly favorable gas-phase energy (−66.43 ± 3.28 kcal·mol⁻¹) was sufficient to yield the most favorable overall binding among the CCR2 systems.

Overall, the MM/GBSA results demonstrated a clear protein-dependent binding preference (Table 4). Ginkgolide A was energetically more favorable toward PIK3CG, whereas Broussoflavan A showed stronger binding affinity toward CCR2. Notably, the CCR2-Broussoflavan A complex displayed the most favorable total binding free energy among all analyzed systems (−28.72 ± 3.79 kcal·mol⁻¹), closely followed by PIK3CG-Ginkgolide A (−27.66 ± 1.52 kcal·mol⁻¹). These findings are in agreement with the preceding trajectory-based analyses and further support the selective dynamic and energetic compatibility of Ginkgolide A with PIK3CG and Broussoflavan A with CCR2.

PCA and DCCM Result

Principal component analysis (PCA) and dynamic cross-correlation matrix (DCCM) analysis were carried out to further characterize the large-scale collective motions and residue-wise dynamic communication patterns of the most stable complexes identified from the MD trajectories. PCA reduces the multidimensional trajectory into a limited number of dominant motion modes, whereas DCCM describes the extent of correlated and anti-correlated atomic fluctuations during the simulation. For the PIK3CG-Ginkgolide A complex, the first three principal components accounted for 50.23%, 30.21%, and 5.94% of the total variance, respectively, yielding a cumulative contribution of 86.38% (Figure 8A). The projection of the trajectory onto the PC1-PC2, PC2-PC3, and PC1-PC3 subspaces indicated that the conformational motions were mainly governed by the first two components, with broad but continuous sampling across the essential subspace. This suggests that the dynamic behavior of the complex was dominated by a limited number of collective motions rather than randomly dispersed structural fluctuations. Consistent with this observation, the DCCM profile of the PIK3CG-Ginkgolide A system revealed multiple strongly correlated and anti-correlated residue blocks, indicating coordinated long-range intramolecular communication (Figure 8B). Prominent positive correlation clusters were observed within several residue segments, while distinct off-diagonal blue regions reflected anti-correlated motions between spatially separated parts of the protein, suggesting a dynamically organized conformational network in the ligand-bound state.

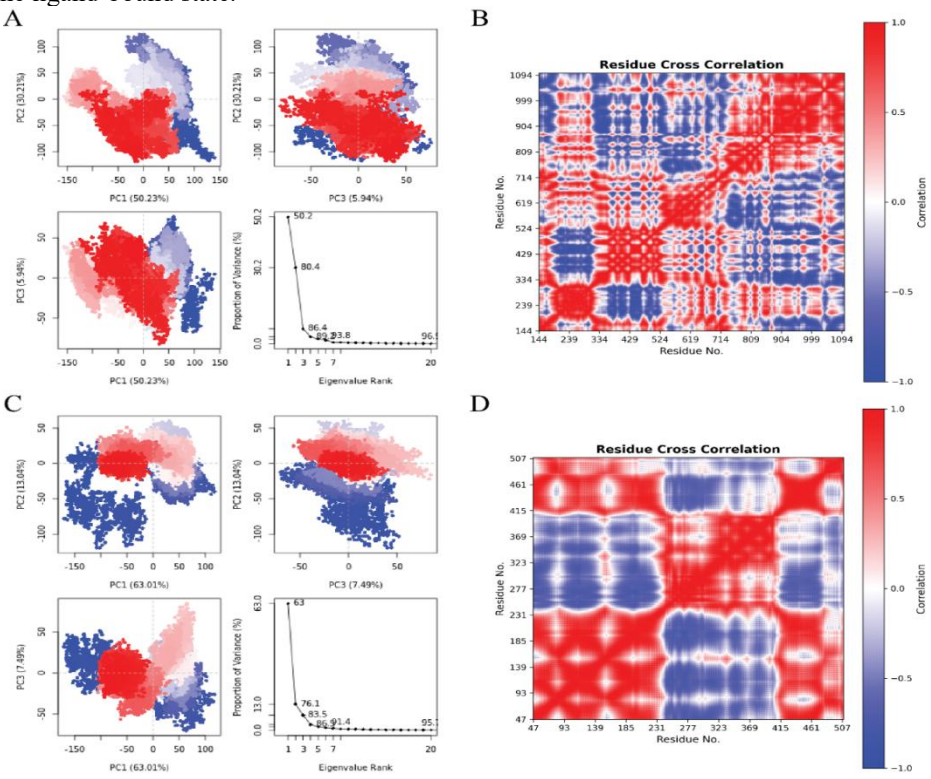


Figure 8. Principal component analysis (PCA) and dynamic cross-correlation matrix (DCCM) analysis of the PIK3CG-Ginkgolide A and CCR2-Broussoflavan A complexes obtained from the 200 ns molecular dynamics

simulation. (A) PCA projection of the PIK3CG-Ginkgolide A complex onto the PC1-PC2, PC2-PC3, and PC1-PC3 subspaces, together with the scree plot showing the proportion of variance explained by the first 20 eigenvectors. (B) DCCM heatmap of residue-wise correlated motions in the PIK3CG-Ginkgolide A complex. (C) PCA projection of the CCR2-Broussoflavan A complex onto the PC1-PC2, PC2-PC3, and PC1-PC3 subspaces, together with the scree plot showing the proportion of variance explained by the first 20 eigenvectors. (D) DCCM heatmap of residue-wise correlated motions in the CCR2-Broussoflavan A complex. Red indicates positive correlation, blue indicates negative correlation, and white indicates weak or near-zero correlation.

A similar analysis of the CCR2-Broussoflavan A complex showed that the first three principal components explained 63.01%, 13.04%, and 7.49% of the total variance, respectively, corresponding to a cumulative variance of 83.54% (Figure 8C). The higher contribution of PC1 indicates that the dominant motion of this complex was even more strongly concentrated in a single principal mode. The trajectory projections showed a relatively confined conformational distribution compared with the broader essential motion observed in the PIK3CG system, indicating that the CCR2-Broussoflavan A complex sampled a more restricted yet coherent conformational space during the simulation. The DCCM map further demonstrated well-defined correlation patterns throughout the receptor structure (Figure 8D). Strong positive correlations were evident within several local residue clusters, whereas marked anti-correlated motions were observed between the central region and the flanking structural segments, reflecting coordinated internal communication within the receptor. Taken together, the PCA and DCCM analyses indicate that both complexes possessed organized collective dynamics; however, the PIK3CG-Ginkgolide A complex displayed a broader multidimensional motion governed by the first two dominant components, whereas the CCR2-Broussoflavan A complex exhibited a more concentrated and comparatively restricted principal motion dominated by PC1.

HOMO–LUMO Molecular Orbital Analysis

To further investigate the electronic characteristics of the most stable protein–ligand systems, HOMO–LUMO molecular orbital analysis was performed for the and CCR2–Broussoflavan A complexes. The results revealed distinct differences in their frontier orbital energies and derived reactivity descriptors (Table 5). The PIK3CG–Ginkgolide A complex (5JHB) showed an E_HOMO of −0.21729 eV and an E_LUMO of −0.19965 eV, producing a very small energy gap of 0.01764 eV. In contrast, the CCR2–Broussoflavan A complex (5T1A) exhibited an E_HOMO of −0.25720 eV and an E_LUMO of −0.15726 eV, with a comparatively larger energy gap of 0.09994 eV. The markedly lower HOMO–LUMO gap of the PIK3CG–Ginkgolide A complex indicates greater electronic softness and higher chemical reactivity, whereas the larger gap of the CCR2–Broussoflavan A complex suggests comparatively lower reactivity and greater kinetic stability.

Table 5. HOMO and LUMO Energies and Energy Gap for Complexes.

Compound	EHOMO (eV)	ELUMO (eV)	Gap (eV)	IP (eV)	EA (eV)	Electronegativity χ (eV)	Chemical Potential μ (eV)	Hardness η (eV)	Electrophilicity ω (eV)
PIK3CG–Ginkgolide A	–0.21729	–0.19965	0.01764	0.21729	0.19965	0.20847	–0.20847	0.00882	9.85482
CCR2–Broussoflavan A	–0.25720	–0.15726	0.09994	0.2572	0.15726	0.20723	–0.20723	0.04997	1.7188

This interpretation was further supported by the calculated global reactivity descriptors. The hardness (η) was substantially lower for PIK3CG–Ginkgolide A (0.00882 eV) than for CCR2–Broussoflavan A (0.04997 eV), confirming the softer electronic nature of the former complex. Likewise, the electrophilicity index (ω) was much higher for PIK3CG–Ginkgolide A (9.85482 eV) than for CCR2–Broussoflavan A (1.7188 eV), indicating a stronger tendency to accept electrons and participate in charge-transfer interactions.

For each complex, the highest occupied molecular orbital (HOMO) is shown on the left and the lowest unoccupied molecular orbital (LUMO) is shown on the right. The orbital distributions illustrate the electronic regions potentially involved in charge transfer and intermolecular interaction. In agreement with the calculated descriptors, the PIK3CG–Ginkgolide A complex exhibited a smaller HOMO–LUMO energy gap, indicating higher electronic reactivity, whereas the CCR2–Broussoflavan A complex showed a larger energy gap, suggesting greater kinetic stability (Figure 9).

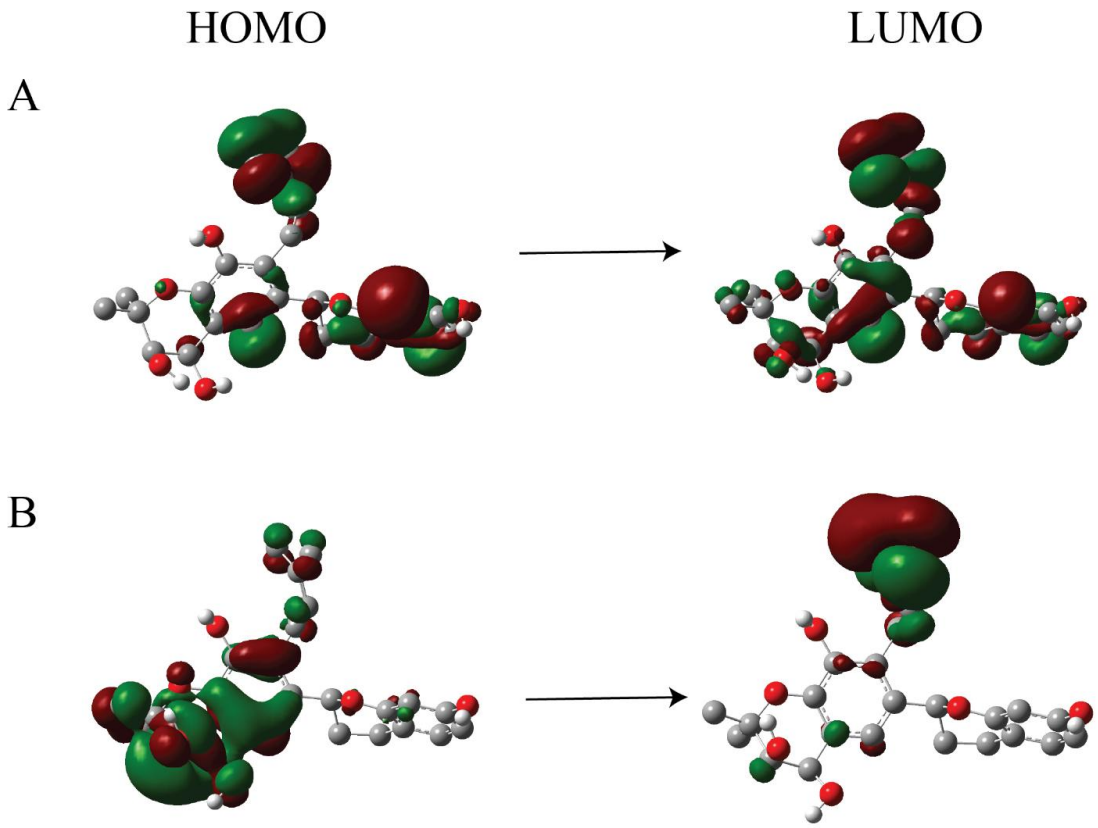


Figure 9. HOMO–LUMO molecular orbital surfaces of the top-ranked ligand-bound complexes. (A) Frontier molecular orbital distribution of the **PIK3CG–Ginkgolide A** complex. **(B)** Frontier molecular orbital distribution of the **CCR2–Broussoflavan A** complex.

Although both complexes displayed similar electronegativity values, the overall descriptor profile suggests that the PIK3CG–Ginkgolide A system possesses enhanced electronic responsiveness, whereas the CCR2–Broussoflavan A complex retains a more electronically stable configuration. These findings are consistent with the target-specific binding behavior observed in the docking, molecular dynamics, and MM/GBSA analyses.

DISCUSSION

The present study was undertaken to clarify the molecular overlap between obesity and cardiovascular disease (CVD), two pathologies that are tightly linked by chronic low-grade inflammation, endothelial dysfunction, oxidative stress, and immune-metabolic imbalance. This cardiometabolic continuum is now widely recognized as a major driver of atherosclerotic progression and vascular injury, but the shared transcriptomic determinants and tractable therapeutic targets remain incompletely defined. Methodologically, this work used an integrated in silico framework that combined differential expression analysis across public obesity- and CVD-related GEO datasets with network prioritization, GO/KEGG enrichment, molecular docking, ADMET filtering, 200 ns molecular dynamics simulation, MM/GBSA energy estimation, and HOMO–LUMO analysis. This multistep design is a strength because it moves beyond simple DEG listing and instead links transcriptomic discovery with target prioritization and ligand-level mechanistic assessment.

The core findings of the study support a shared inflammatory and signaling axis between obesity and CVD. Specifically, 846 downregulated and 1002 upregulated genes were identified in the CVD datasets, whereas the obesity datasets yielded 865 downregulated and 1564 upregulated genes. After 12 cytoHubba rankings and MCODE clustering, 24 prioritized hub genes were retained, and cross-comparison reduced these to five final shared biomarkers: the common downregulated genes JUN and EGFR, and the common upregulated genes EP300, PIK3CG, and CCR2. At the therapeutic level, Ginkgolide A showed the best docking score against PIK3CG (−9.3 kcal/mol) and a favorable MM/GBSA binding free energy (−27.66 ± 1.52 kcal·mol^{-1}), whereas Broussoflavan A showed the best docking score against CCR2 (−9.5 kcal/mol) and the most.

Lau and colleagues showed that obesity is associated with a broad circulating proteomic signature overlapping with abdominal adiposity, insulin resistance, and other cardiometabolic traits, supporting the concept that obesity-related biomarkers are deeply embedded in cardiovascular risk biology. Together, these reports align well with the present strategy of seeking overlapping molecular nodes rather than disease-specific markers in isolation [42]. Among the upregulated genes, PIK3CG and CCR2 are especially compelling because both occupy central positions in immune-cell trafficking and inflammatory signaling. Earlier experimental work showed that blockade or genetic deficiency of PI3Kγ/PIK3CG ameliorates obesity-induced inflammation, improves insulin resistance, and reduces macrophage-driven metabolic dysfunction, while broader reviews have

placed PI3K signaling at the center of atherosclerotic progression and regression [43]. The CCL2–CCR2 axis is similarly well established in both obesity and vascular disease: CCR2 deficiency modulates the inflammatory and metabolic consequences of obesity, CCR2-directed antagonism inhibits atherogenesis in ApoE-deficient mice, and recent human genetic evidence suggests that damaging CCR2 variants are associated with lower atherosclerotic disease risk [44]. In this context, the emergence of PIK3CG and CCR2 as shared upregulated biomarkers in your analysis is biologically coherent and therapeutically meaningful. The phytochemical results are also noteworthy. Ginkgolide A is not merely a docking hit; it belongs to a well-studied class of Ginkgo biloba terpenoids with recognized cardiovascular relevance [45]. For Broussoflavan A, the literature is more limited, which makes your result potentially more novel. Direct cardiovascular studies on Broussoflavan A remain scarce; however, it is a Broussonetia papyrifera-derived prenylated flavan, and compounds from this genus have repeatedly been reported to possess anti-inflammatory, antioxidant, antidiabetic, anti-obesity, antibacterial, and antiviral properties[46].

This study has several strengths, it used multiple public datasets and retained only genes with consistent cross-disease relevance, which improves biological plausibility. Second, hub selection was not based on a single network measure but on 12 cytoHubba methods plus MCODE, making the biomarker nomination more robust. The convergence between transcriptomic, docking, dynamic, and energetic results strengthens the conclusion that PIK3CG-Ginkgolide A and CCR2-Broussoflavan A are the most promising target–ligand pairs emerging from this dataset.

Several limitations should also be acknowledged. The included GEO datasets differ in tissue source, disease phenotype, and sample size, which may introduce biological and technical heterogeneity. In particular, the obesity arm combines blood-derived, PBMC-derived, and tissue-derived profiles, and one included dataset is centred on diabetic complications rather than uncomplicated obesity, which may enrich inflammatory signal but also introduce confounding. Future work should therefore proceed in three directions. First, the five shared biomarkers should be validated in independent clinical cohorts using RT-qPCR, ELISA, targeted proteomics, and ideally single-cell or spatial transcriptomics to determine their cell-type specificity in blood, adipose tissue, and vascular lesions. Second, mechanistic studies should test whether PIK3CG and CCR2 inhibition can directly modulate the inflammatory and endothelial phenotypes linking obesity and CVD, while the roles of EP300, JUN, and EGFR should be explored in epigenetic and stress-response signaling.

CONCLUSION

In conclusion, this study provides an integrated in silico framework for uncovering the shared molecular basis of obesity and cardiovascular disease. By combining transcriptomic screening, network analysis, functional enrichment, and computational therapeutic evaluation, the study identified JUN, EGFR, EP300, PIK3CG, and CCR2 as the principal shared biomarkers linking the two disorders. These genes collectively reflect a common pathological axis centered on inflammatory signaling, transcriptional regulation, and immune-cell recruitment. Among the prioritized targets, PIK3CG and CCR2 emerged as the most actionable therapeutic nodes. Computational screening of phytochemicals further revealed a clear target-specific preference, with Ginkgolide A showing superior compatibility with PIK3CG and Broussoflavan A showing superior compatibility with CCR2. Their favorable docking scores, stable molecular dynamics behavior, acceptable ADMET profiles, and strong MM/GBSA energies support their potential as promising lead compounds for future validation. Taken together, the present findings advance understanding of the molecular convergence between obesity and CVD and provide a focused set of biomarkers and candidate phytochemical therapeutics for further translational investigation. Experimental validation in cellular, animal, and clinical settings will be necessary to confirm the diagnostic and therapeutic potential of these targets and compounds.

Author Contributions.

Conflicts of Interest.

The authors declare no conflict of interest.

Funding.

We received no funding for this study.

Ethics approval and consent to participate.

Not applicable.

Consent for publication.

Not applicable.

Availability of data and materials.

No new data were generated or analyzed for this study.

Acknowledgment.

None.

REFERENCES

1. Młynarska E, Bojdo K, Bulicz A, Frankenstein H, Gąsior M, Kustosik N, et al. Obesity as a Multifactorial Chronic Disease: Molecular Mechanisms, Systemic Impact, and Emerging Digital Interventions. Curr Issues

Mol Biol. 2025;47: 787. doi:10.3390/cimb47100787
2. Lin X, Li H. Obesity: Epidemiology, Pathophysiology, and Therapeutics. Front Endocrinol. 2021;12: 706978. doi:10.3389/fendo.2021.706978
3. Frąk W, Wojtasińska A, Lisińska W, Młynarska E, Franczyk B, Rysz J. Pathophysiology of Cardiovascular Diseases: New Insights into Molecular Mechanisms of Atherosclerosis, Arterial Hypertension, and Coronary Artery Disease. Biomedicines. 2022;10: 1938. doi:10.3390/biomedicines10081938
4. Gutiérrez-Cuevas J, Sandoval-Rodriguez A, Meza-Rios A, Monroy-Ramírez HC, Galicia-Moreno M, García-Bañuelos J, et al. Molecular Mechanisms of Obesity-Linked Cardiac Dysfunction: An Up-Date on Current Knowledge. Cells. 2021;10: 629. doi:10.3390/cells10030629
5. Sanghera DK, Bejar C, Sharma S, Gupta R, Blackett PR. Obesity Genetics and Cardiometabolic Health: Potential for Risk Prediction. Diabetes Obes Metab. 2019;21: 1088–1100. doi:10.1111/dom.13641
6. Popko K, Gorska E, Stelmaszczyk-Emmel A, Plywaczewski R, Stoklosa A, Gorecka D, et al. Proinflammatory cytokines IL-6 and TNF-α and the development of inflammation in obese subjects. Eur J Med Res. 2010;15: 120–122. doi:10.1186/2047-783X-15-S2-120
7. Ramji DP, Davies TS. Cytokines in atherosclerosis: Key players in all stages of disease and promising therapeutic targets. Cytokine Growth Factor Rev. 2015;26: 673–685. doi:10.1016/j.cytogfr.2015.04.003
8. Sanches PHG, de Melo NC, Porcari AM, de Carvalho LM. Integrating Molecular Perspectives: Strategies for Comprehensive Multi-Omics Integrative Data Analysis and Machine Learning Applications in Transcriptomics, Proteomics, and Metabolomics. Biology. 2024;13: 848. doi:10.3390/biology13110848
9. Kong P, Cui Z-Y, Huang X-F, Zhang D-D, Guo R-J, Han M. Inflammation and atherosclerosis: signaling pathways and therapeutic intervention. Signal Transduct Target Ther. 2022;7: 131. doi:10.1038/s41392-022-00955-7
10. Feijóo-Bandín S, Aragón-Herrera A, Moraña-Fernández S, Anido-Varela L, Tarazón E, Roselló-Lletí E, et al. Adipokines and Inflammation: Focus on Cardiovascular Diseases. Int J Mol Sci. 2020;21: 7711. doi:10.3390/ijms21207711
11. Li W, Jin K, Luo J, Xu W, Wu Y, Zhou J, et al. NF-κB and its crosstalk with endoplasmic reticulum stress in atherosclerosis. Front Cardiovasc Med. 2022;9: 988266. doi:10.3389/fcvm.2022.988266
12. Kim HK, Kim SJ, Gil WJ, Yang C-S. Exploring the therapeutic potential of phytochemicals: challenges and strategies for clinical translation. Phytomedicine. 2025;145: 157090. doi:10.1016/j.phymed.2025.157090
13. Khurana S, Venkataraman K, Hollingsworth A, Piche M, Tai TC. Polyphenols: Benefits to the Cardiovascular System in Health and in Aging. Nutrients. 2013;5: 3779–3827. doi:10.3390/nu5103779
14. Shi L, Xu Y, Zhai H, Zhao C, Xia W, Zheng Y, et al. A novel multi-omics-machine learning pipeline reveals immune and metabolic links between type 2 diabetes and atherosclerosis. Diab Vasc Dis Res. 2025;22: 14791641251407597. doi:10.1177/14791641251407597
15. Wang Z, Nakayama T. Inflammation, a Link between Obesity and Cardiovascular Disease. Mediators Inflamm. 2010;2010: 535918. doi:10.1155/2010/535918
16. McDermott JE, Wang J, Mitchell H, Webb-Robertson B-J, Hafen R, Ramey J, et al. Challenges in Biomarker Discovery: Combining Expert Insights with Statistical Analysis of Complex Omics Data. Expert Opin Med Diagn. 2013;7: 37–51. doi:10.1517/17530059.2012.718329
17. Kalra RS, Singh S, Munshi A, Bariwal J. Editorial: Biological aspects of targeted drug discovery: Development of novel targets and/or chemotherapies, and drug repurposing. Front Oncol. 2022;12: 1106610. doi:10.3389/fonc.2022.1106610
18. Clough E, Barrett T. The Gene Expression Omnibus database. Methods Mol Biol Clifton NJ. 2016;1418: 93–110. doi:10.1007/978-1-4939-3578-9_5
19. Li Z, Cheng Q, He Y, Wang S, Xie J, Zheng Y, et al. Effect of Dan-Lou tablets on coronary heart disease revealed by microarray analysis integrated with molecular mechanism studies. Heliyon. 2023;9: e15777. doi:10.1016/j.heliyon.2023.e15777
20. Muse ED, Kramer ER, Wang H, Barrett P, Parviz F, Novotny MA, et al. A Whole Blood Molecular Signature for Acute Myocardial Infarction. Sci Rep. 2017;7: 12268. doi:10.1038/s41598-017-12166-0
21. Nührenberg TG, Langwieser N, Binder H, Kurz T, Stratz C, Kienzle R-P, et al. Transcriptome analysis in patients with progressive coronary artery disease: identification of differential gene expression in peripheral blood. J Cardiovasc Transl Res. 2013;6: 81–93. doi:10.1007/s12265-012-9420-5
22. Ghosh S, Dent R, Harper ME, Stuart J, McPherson R. Blood gene expression reveal pathway differences between diet-sensitive and resistant obese subjects prior to caloric restriction. Obesity. 2011;19: 457–463. doi:10.1038/oby.2010.209
23. Hui Z, Chen Y-M, Gong W-K, Lai J-B, Yao B-B, Zhao Z-J, et al. Shared and specific biological signalling pathways for diabetic retinopathy, peripheral neuropathy and nephropathy by high-throughput sequencing analysis. Diab Vasc Dis Res. 2022;19: 14791641221122918. doi:10.1177/14791641221122918
24. Ono K, Fong D, Gao C, Churas C, Pillich R, Lenkiewicz J, et al. Cytoscape Web: bringing network biology to the browser. Nucleic Acids Res. 2025;53: W203–W212. doi:10.1093/nar/gkaf365
25. Wu G, Dawson E, Duong A, Haw R, Stein L. ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. F1000Research. 2014;3: 146. doi:10.12688/f1000research.4431.2
26. Zhao J, He K, Du H, Wei G, Wen Y, Wang J, et al. Bioinformatics prediction and experimental verification

of key biomarkers for diabetic kidney disease based on transcriptome sequencing in mice. PeerJ. 2022;10: e13932. doi:10.7717/peerj.13932
27. Rose Y, Duarte JM, Lowe R, Segura J, Bi C, Bhikadiya C, et al. RCSB Protein Data Bank: Architectural Advances Towards Integrated Searching and Efficient Access to Macromolecular Structure Data from the PDB Archive. J Mol Biol. 2021;433: 166704. doi:10.1016/j.jmb.2020.11.003
28. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res. 2004;32: D115–D119. doi:10.1093/nar/gkh131
29. Lans C, van Asseldonk T. Dr. Duke's Phytochemical and Ethnobotanical Databases, a Cornerstone in the Validation of Ethnoveterinary Medicinal Plants, as Demonstrated by Data on Pets in British Columbia. In: Máthé Á, editor. Medicinal and Aromatic Plants of North America. Cham: Springer International Publishing; 2020. pp. 219–246. doi:10.1007/978-3-030-44930-8_10
30. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem Substance and Compound databases. Nucleic Acids Res. 2016;44: D1202–D1213. doi:10.1093/nar/gkv951
31. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7: 42717. doi:10.1038/srep42717
32. Pires DEV, Blundell TL, Ascher DB. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. J Med Chem. 2015;58: 4066–4072. doi:10.1021/acs.jmedchem.5b00104
33. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Methods Mol Biol. 2015;1263: 243–250. doi:10.1007/978-1-4939-2269-7_19
34. 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. 2010;31: 455–461. doi:10.1002/jcc.21334
35. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013;29: 845–854. doi:10.1093/bioinformatics/btt055
36. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. Nat Methods. 2017;14: 71–73. doi:10.1038/nmeth.4067
37. Orio M, Pantazis DA, Neese F. Density functional theory. Photosynth Res. 2009;102: 443–453. doi:10.1007/s11120-009-9404-8
38. Tirado-Rives J, Jorgensen WL. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. J Chem Theory Comput. 2008;4: 297–306. doi:10.1021/ct700248k
39. Alhawarri MB, Al-Thiabat MG, Dubey A, Tufail A, Fouad D, Alrimawi BH, et al. ADME profiling, molecular docking, DFT, and MEP analysis reveal cissamaline, cissamanine, and cissamdine from Cissampelos capensis L.f. as potential anti-Alzheimer's agents. RSC Adv. 2024;14: 9878–9891. doi:10.1039/D4RA01070A
40. Deghady AM, Hussein RK, Alhamzani AG, Mera A. Density Functional Theory and Molecular Docking Investigations of the Chemical and Antibacterial Activities for 1-(4-Hydroxyphenyl)-3-phenylprop-2-en-1-one. Molecules. 2021;26: 3631. doi:10.3390/molecules26123631
41. Ayalew ME. DFT Studies on Molecular Structure, Thermodynamics Parameters, HOMO-LUMO and Spectral Analysis of Pharmaceuticals Compound Quinoline (Benzo[b]Pyridine). J Biophys Chem. 2022;13: 29–42. doi:10.4236/jbpc.2022.133003
42. Lau ES, Paniagua SM, Zarbafian S, Hoffman U, Long MT, Hwang S, et al. Cardiovascular Biomarkers of Obesity and Overlap With Cardiometabolic Dysfunction. J Am Heart Assoc Cardiovasc Cerebrovasc Dis. 2021;10: e020215. doi:10.1161/JAHA.120.020215
43. Kobayashi N, Ueki K, Okazaki Y, Iwane A, Kubota N, Ohsugi M, et al. Blockade of class IB phosphoinositide-3 kinase ameliorates obesity-induced inflammation and insulin resistance. Proc Natl Acad Sci U S A. 2011;108: 5753–5758. doi:10.1073/pnas.1016430108
44. Bot I, Ortiz Zacarías NV, de Witte WEA, de Vries H, van Santbrink PJ, van der Velden D, et al. A novel CCR2 antagonist inhibits atherogenesis in apoE deficient mice by achieving high receptor occupancy. Sci Rep. 2017;7: 52. doi:10.1038/s41598-017-00104-z
45. Silva H, Martins FG. Cardiovascular Activity of Ginkgo biloba—An Insight from Healthy Subjects. Biology. 2022;12: 15. doi:10.3390/biology12010015
46. Son KH, Kwon SJ, Chang HW, Kim HP, Kang SS. Papyriflavonol A, a new prenylated flavonol from Broussonetia papyrifera. Fitoterapia. 2001;72: 456–458. doi:10.1016/s0367-326x(00)00329-4