Computational and Structural Biotechnology Journal 27 (2025) 171–183

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research Article

Structure-based identification of HNF4α agonists: Rosmarinic acid as a promising candidate for NAFLD treatment

Xi Chen [a,1], Xinqi Zhu [a,1], Gang Wu [b], Xiaobo Wang [b], Yu Zhang [a,*], Nan Jiang [a,*]

[a] National Vaccine Innovation Platform, Scholl of Pharmacy, Nanjing Medical University, Nanjing 211166, China
[b] NanjingMinova Pharmaceutical Co., Ltd.Jiangsu Biotech Innovation Park, Nanjing 211166, China

ARTICLE INFO

Keywords:
NAFLD
HNF4α
MD simulation
Rosmarinic acid
Drug discovery

ABSTRACT

The prevention and treatment of metabolic disorders, such as non-alcoholic fatty liver disease (NAFLD), have emerged as critical global health challenges. Current lipid-lowering pharmacotherapies are associated with side effects, including hepatotoxicity, rhabdomyolysis, and decreased erythrocyte counts, underscoring the urgent need for safer therapeutic alternatives. Hepatocyte nuclear factor 4α (HNF4α) has been identified as a pivotal regulator of lipid metabolism, making it an attractive target for drug development. In this study, we investigated the structural characteristics and binding interactions of four HNF4α agonists: Alverine, Benfluorex, N-trans caffeoyltyramine (NCT), and N-trans feruloyltyramine (NFT). Our results indicate that the conjugated structure formed by the amide bond and the aromatic ring in NCT and NFT enhances electron density, potentially contributing to their increased specificity for HNF4α relative to Alverine and Benfluorex. Additionally, electrostatic interactions between the aromatic moieties of the compounds and HNF4α residues were found to play a crucial role in ligand binding. Leveraging these insights, we performed a high-throughput virtual screening of 2131 natural compounds, using the binding modes of NCT and NFT as reference templates. Rosmarinic acid emerged as a promising HNF4α agonist, exhibiting a high consensus score and favorable binding affinity. Subsequent biological assays demonstrated that rosmarinic acid significantly inhibited HepG2 cell proliferation which related to the enhancement of autophagy. After the knockdown of P2 isoform of HNF4α, HepG2 was more sensitive to the administration of NCT and rosmarinic acid. Furthermore, the proliferation of DLD-1 cell, which only expresses the P2 isoform of HNF4α, was not significantly inhibited by the administration of NCT and rosmarinic acid. Collectively, these findings suggest that rosmarinic acid is a promising HNF4α agonist which is more effective to activate the P1 isoform of HNF4α and holds potential as an effective treatment for NAFLD, providing a foundation for the development of novel lipid-lowering drugs with enhanced efficacy and reduced side effect.

Data Availability: Data will be made available on request.

1. Introduction

In recent decades, non-alcoholic fatty liver disease (NAFLD) has emerged as one of the most prevalent chronic liver disorders, affecting approximately 25 % of the global population [1]. Evidence indicates that obesity and lipid metabolism dysregulation associated with NAFLD are key drivers of metabolic diseases, including diabetes and hyperlipidemia, and are strongly linked to an increased risk of cardiovascular and cerebrovascular diseases, as well as certain cancers [2–5]. Consequently, the prevention and management of NAFLD and related metabolic disorders have become pressing global public health priorities.

Currently, first-line pharmacotherapies for obesity and associated lipid metabolism disorders primarily include statins, fibrates, and niacin, which function by inhibiting cholesterol synthesis or facilitating the catabolism of excess lipids. However, prolonged use of these agents is associated with side effects, such as hepatotoxicity, rhabdomyolysis, and reduced erythrocyte count, thereby limiting their applicability in the treatment of NAFLD. Consequently, elucidating the molecular mechanisms underlying obesity and lipid dysregulation, identifying

* Corresponding authors.
E-mail addresses: zhangyu@njmu.edu.cn (Y. Zhang), jiangnan@njmu.edu.cn (N. Jiang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2024.12.014
Received 17 August 2024; Received in revised form 6 December 2024; Accepted 18 December 2024
Available online 27 December 2024

novel therapeutic targets to mitigate obesity and reduce lipid accumulation, and developing more effective weight-loss and lipid-lowering agents to overcome the limitations of current therapies represent pressing challenges in both fundamental research and clinical practice

Hepatocyte nuclear factor 4α (HNF4α) is a member of the orphan steroid hormone nuclear receptor superfamily [6]. It is a nuclear receptor transcription factor characterized by highly conserved DNA-binding and ligand-binding domains. HNF4α is predominantly expressed in the β-cells of visceral endothelial cells, as well as in the liver, intestine, kidney, and pancreas, where it plays critical roles in cellular development, differentiation, glucose metabolism, the urea cycle, and lipid metabolism [7–10]. For instance, in models of NAFLD, HNF4α overexpression has been shown to upregulate carboxylesterase 2 (CES2), thereby enhancing fatty acid oxidation [11]. Research by Iannis Talianidis and colleagues demonstrated that HNF4α significantly promotes fatty acid oxidation during fasting by mediating the expression of carnitine palmitoyltransferase 1 (CPT1), thrombomodulin receptor (CD36), and fibroblast growth factor 21 (FGF21) [10]. Moreover, Lee et al. reported that HNF4α activity is closely associated with lipid storage, where its activation markedly enhances lipid catabolism, underscoring HNF4α's potential as a key target in the treatment of obesity and related lipid metabolism disorders. Currently identified HNF4α agonists include Alverine, Benfluorex, N-trans caffeoyltyramine (NCT), and N-trans feruloyltyramine (NFT) (Fig. 1), with NCT and NFT demonstrating superior activation of HNF4α [12]. In vitro studies have shown that NCT and NFT promote lipid clearance in palmitic acid-loaded cells. Furthermore, in diet-induced obesity (DIO) mouse models, NCT was found to restore hepatic HNF4α expression and alleviate fatty liver [12]. These findings suggest that NCT and NFT hold promise as potential therapeutic agents for the treatment of lipid metabolism disorders.

Structural analysis of Alverine, Benfluorex, NCT, and NFT revealed significant structural similarities, particularly between NCT and NFT, with a similarity coefficient of 0.76. All four compounds contain aromatic moieties, which are likely instrumental in their pharmacological activity. Examination of the molecular interactions and binding free energy decomposition indicated that hydrophobic interactions between the aromatic rings of these compounds and amino acid residues of HNF4α are crucial forces in stabilizing ligand binding. Despite these similarities, notable structural differences exist. Compared to Alverine and Benfluorex, which feature amine groups linked to flexible alkyl chains, NCT and NFT possess simpler structures characterized by an amide bond conjugated to an aromatic ring derived from caffeic acid and ferulic acid, respectively. We hypothesize that the conjugated structure formed by the amide bond and aromatic ring in NCT and NFT enhances electrondensity, thereby contributing to their greater specificity for HNF4α relative to Alverine and Benfluorex. Building on these structural insights, the present study aims to further elucidate the structure-activity relationships and binding site characteristics of these four compounds with HNF4α, develop and validate computational models, and ultimately identify novel, more potent HNF4α agonists.

Utilizing molecular docking, molecular dynamics (MD) simulations, and binding free energy calculations, we determined that NCT and NFT exhibit superior binding efficacy compared to Alverine and Benfluorex, corroborating the experimental findings of Lee et al. Subsequently, leveraging the binding conformations of NCT and NFT with HNF4α, we screened a library of 2131 natural compounds, selecting candidates with high docking scores and similar structural features for further validation through microscale thermophoresis (MST), Cell Counting Kit-8 (CCK-8), and Western Blot (WB) assays. The knockdown of the P2 isoform of HNF4α enhanced the inhibitory effect of rosmarinic acid on HepG2 cells. Conversely, rosmarinic acid exhibited limited inhibitory activity on DLD-1 cells, which predominantly express the P2 isoform of HNF4α. These findings suggest that the beneficial effects of rosmarinic acid are more dependent on the activation of the P1 isoform of HNF4α. We hope this work provide novel insights and a theoretical basis for the development of innovative therapeutic agents targeting lipid metabolism disorders associated with NAFLD.

2. Methods and computational details

2.1. Preparation of the ligands and protein

The optimization processes of proteins and ligands were completed using Discovery Studio (DS) 4.0 software [13]. The structures of Alverine (CAS: 150–59–4) and Benfluorex (CAS: 23642–66–2) were obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov), and the structures of NCT and NFT were obtained from the literature by Lee [14] The ligprep module of DS was used to prepare small molecule compounds, using the CHARMm force field [15], with the maximum number of iterations set to 2000 and the root mean square gradient set to 0.01.

The highest resolution crystal structure of HNF4α was selected from the RCSB protein data band (PDB) with PDB code of 3FS1 [16,17], as the three-dimensional structure of HNF4α. The protein structure was prepared and optimized using the protein prepare module of DS, with the force field also chosen as CHARMm.

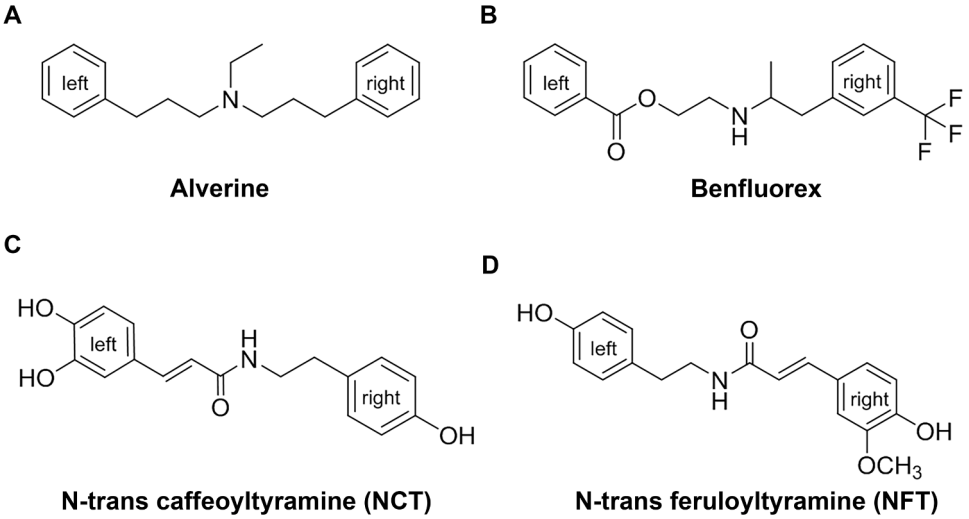

Fig. 1. Structural formulas of four HNF4α agonists. (A) Alverine, (B) Benfluorex, (C) NCT, (D) NFT.

2.2. Molecular docking

By comparing the binding modes of proteins and ligands in different crystal structures of HNF4α (e.g., PGC1α, SRC1, and EBP1), we found that HNF4α contains a conserved LXXLL motif for targeting ligands [16] (Fig. 2). Here, L represents leucine (LEU), and X represents any other amino acid residue. Therefore, we selected the LXXLL motif at the HNF4α-ligand interaction interface as the active site for generating the docking pocket. The grid center coordinates of the active site region were X: 79.715, Y: 96.805, Z: 2.423, with the grid volume and cluster radius set to 1.0nanometer (nm).

Subsequently, molecular docking was performed within CHARMm force field using the Dock Ligands (CDOCKER) module, with dynamic steps set to 1000, dynamic target temperature set to 100 K, heating steps set to 2000, and heating target temperature set to 300 K. Finally, the affinity of small molecules binding to HNF4α was evaluated based on the -CDOCKER_INTERACTION_ENERGY score [18] (Table 1). The higher the absolute value, the stronger the binding ability between the protein and ligand.

2.3. MD simulations

To thoroughly investigate the binding affinity and interaction modes between candidate ligand molecules and the receptor protein target, we used the binding conformations of the candidate compounds with HNF4α protein obtained from high-throughput screening as initial structures [19]. MD simulations were performed using Amber 18 software [20] for 50 nanoseconds (ns) based on the molecular force field.

First, the charge parameters of the ligands were calculated using Gaussian 09 [21]. The Hartree-Fock (HF) method with the 6–31 G(d,p) basis set was used to optimize the gas-phase geometry of the ligands, followed by fitting the electrostatic potential (ESP) using the B3LYP/cc-pVTZ method [22]. The integral equation formalism (IEF) of the polarizable continuum model (PCM) was used to simulate the organic solvent environment (ε = 4.0), and partial charges and restrained electrostatic potential (RESP) were assigned in the ANTECHAMBER module [23].

The force field files and coordinate files of the complexes were prepared using the tleap module in Amber, with the ligands adopting the GAFF small molecule force field and the proteins adopting the FF14SB force field [24]. First, the small molecule ligand bond parameter files, charge files, and mol2 files fitted by Gaussian 09 were imported, followed by importing the PDB files of the complexes. An octahedral water box and periodic boundary conditions were added, with water molecules using the classic TIP3P model and the minimum distance between atoms and the box edge set to 10.0 Å. Sodium ions were added to neutralize the system. Finally, the topology files and coordinate files of the complexes were saved for subsequent MD simulations.

To make the initial structure more reasonable, a two-step energy minimization was performed: (1) fixing the HNF4α protein structure with a constraint force of 500 kcal/mol, optimizing the protein using the steepest descent method for 2500 steps, and then optimizing the solvent water molecules using the conjugate gradient method for 2500 steps; (2) removing the constraints on the HNF4α protein and ligand molecules, optimizing the entire system using the steepest descent method for 1000 steps, and then optimizing using the conjugate gradient method for 9000 steps. The system was heated from 0 K to 300 K within 0.1 ns, followed by MD simulations in the NVT ensemble. The time step was 2 fs, and the simulation duration was 50 ns, with trajectory coordinates recorded every 1 ns.

Table 1
Docking analysis, including protein-ligand electrostatic (Elec, kcal/mol), Van der Waals (VdW, kcal/mol), potential energy (PE, kcal/mol) and total interaction energy (E_{tot}, kcal/mol) of agonists with HNF4α.

Complexes	Elec	PE	VdW	E_{tot}
HNF4α-Alverine	23.23	31.17	−3.82	32.13
HNF4α-Benfluorex	−2.62	4.41	0.18	28.65
HNF4α-NCT	−38.18	−28.89	0.75	32.74
HNF4α-NFT	−30.02	−19.75	1.32	27.07

2.4. Binding energy calculation

To calculate the interaction energy and solvation free energy between HNF4α protein and small molecules, we used the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method [25,26]. For each system, we extracted the final 40,000 frames of stable MD simulation trajectories, covering a time range from 11 to 50 nanoseconds. The Gibbs free energy (ΔG_bind) was defined as the difference between the Gibbs free energy of the complex (G_complex) and the sum of the Gibbs free energy of the isolated ligand (G_ligand) and protein (G_protein), calculated using the following formula:

$$\Delta G_{bind} = G_{complex} - \left(G_{ligand} + G_{protein}\right) \quad (1)$$

In equation (1), the Gibbs free energy change is calculated using the

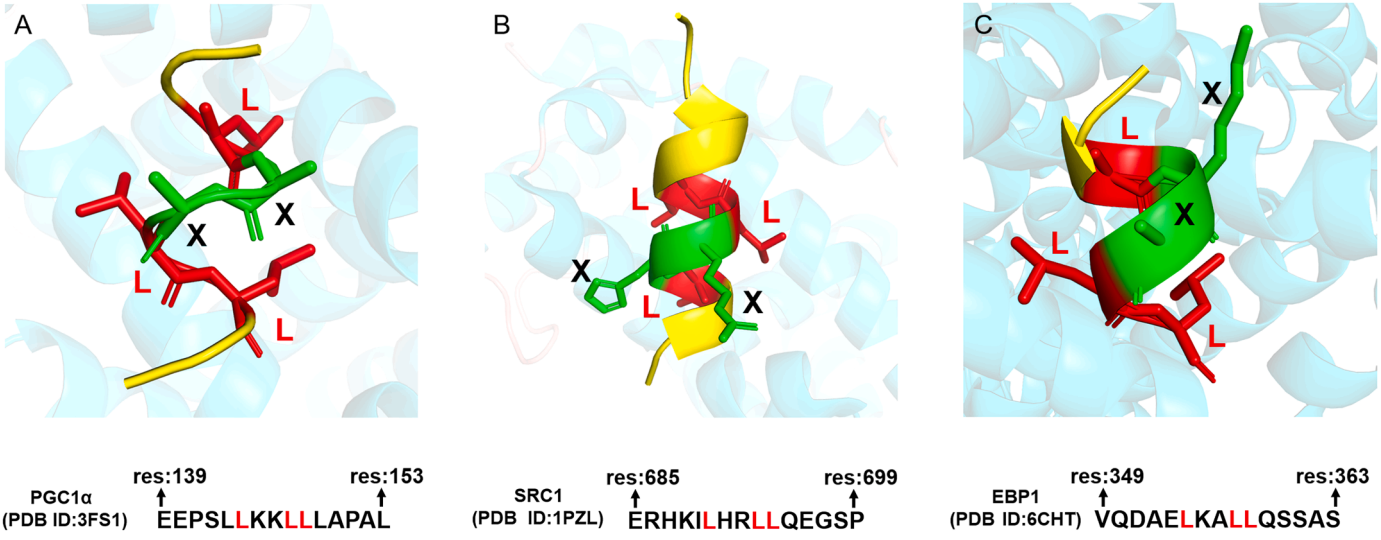

Fig. 2. Conserved LXXLL motifs (in cartoon and stick representations)in human HNF4α⊡ (A) PGC1α, (B) SRC1 and (C) EBP1 for targeting ligands. The highlighted red part L is leucine, and the green part X is any amino acid residue.

Poisson-Boltzmann (PB) equation. ΔG is calculated using the following equation (2):

$$\Delta G = E_{bond} + E_{ele} + E_{VdW} + G_{pol} + G_{np} - TS \quad (2)$$

In this formula, the first three terms are the standard MM energy terms obtained by molecular mechanics calculation, expressed as bond (bond length, bond angle, and dihedral angle), electrostatic interaction, and van der Waals interaction, respectively. G_pol and G_np denote the polar and nonpolar solvation free energies obtained by solving the PB equation, respectively. For the last term, the entropy (TS) is calculated using classical statistical thermodynamics and normal mode analysis [27].

2.5. High-throughput screening

The natural product database contains 2131 compounds. These compounds were optimized using SYBYL software [28]. First, the small molecules were converted from their 2D structures in the sdf file to 3D structures, and the Gasteiger-Hückel charge was assigned using the Tripos force field. The gradient was set to 0.005 kcal/(mol·Å) and the maximum number of iterations was set to 1000 for energy minimization. After optimization, the optimized sdf file was saved. High-throughput virtual screening of the natural compound database was performed using SYBYL software. The docking method used was Surflex-Dock Screen (SFXC) [29] in high-throughput screening mode, and the docking pocket was located in the LXXLL motif of HNF4α protein. A series of ligand molecules with potential binding activity to HNF4α were preliminarily screened, and the docking scores were compared and ranked.

2.6. Cell viability assay

The viability of HepG2 and DLD-1 cells was assessed using the CCK-8 (Biosharp) assay. Briefly, the cells were seeded in a 96-well microplate at a density of 3 × 10^3 cells/well. After 24 h, the BI6015 (TargetMol, T21867), licochalcone B (TargetMol, T4S0350), rosmarinic acid (TargetMol, T2765), cynarin (TargetMol, T6S1529), and cichoric acid (TargetMol, TL0006) stock solutions were serially diluted to concentrations of 0.78125 μM, 1.5625 μM, 3.125 μM, 6.25 μM, 12.5 μM, 25 μM, 50 μM, 100 μM, and 200 μM in DMEM and added to the designated wells. After 48 h of incubation, the DMEM medium was replaced with 10 μL of CCK-8 solution and incubated for an additional hour at 37 °C in the dark. The absorbance of each well was measured at 450 nm using a microplate reader (BioTek).

The shRNA sequence targeting the HNF4α P2 isoform class is according to previous report [30]: 5′-CCGGCAGTGGAGAGTTCTTACG ACACTCGAGTGTCGTAAGAACTCTCCACTGTTTTTT-3′. It is purchased from Beijing Tsingke Biotech Co., Ltd. and cloned into the Plko.1-EGFP-PURO vector between *Age*I and *Eco*RI.

2.7. WB experiments

HepG2 cells were seeded in a 6-well plate at a density of 3 × 10^5 cells/well and treated with media containing different concentrations of rosmarinic acid, cichoric acid, cynarin, and licochalcone B for 48 h. Cells were then collected. Briefly, the quantified proteins were separated by SDS–polyacrylamide gel electrophoresis and then transferred onto a PVDF membrane. The PVDF membrane was then blocked with a blocking solution for 2 h and incubated with HNF4α antibody (1:1000 dilution, Affinity, AF6297/20) overnight at 4 °C. The PVDF membrane was washed three times with WB washing solution for 10 min each time. The membranes were then incubated for 1 h at room temperature with horseradish peroxidase-labeled secondary antibodies: SQSTM1/p62 (Cell Signaling Technology, #5114), LC3B Antibody (Cell Signaling Technology, #2775), Beclin-1 (Cell Signaling Technology, #3738), mTOR (Cell Signaling Technology, #2983), p-mTOR (Cell Signaling Technology, #2974), and β-actin (ProteinTech, 81115–1-RR). The membrane was then incubated in horseradish peroxidase-conjugated IgG secondary antibody (Cell Signaling Technology) for 1 h at room temperature. Finally, the protein bands were observed using high-signal ECL WB substrates.

2.8. MST measurements

HNF4α protein was labeled with MonolithTM RED-NHS protein labeling kit (NanoTemper, MO-L011), and then mixed with rosmarinic acid, cichoric acid, cynarin, and licochalcone B. The measurements were performed using the RED channel of the Monolith NT.115 MST device (NanoTemper Technologies) and the corresponding Monolith capillaries (NanoTemper, MO-K022). The data were collected using MO. Control Version 2 and evaluated using MO. Affinity Version 2.3 software (NanoTemper).

2.9. Statistical analysis

Standard deviation (SD) is used as indicator of experimental imprecision. So, all experimental data were expressed as mean ± SD. One-way ANOVA or Student's *t*-test in GraphPad Prism 8 software was used for statistical analysis. Differences between data were considered significant at $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***), $P < 0.0001$ (****).

3. Results and discussion

3.1. Analysis of agonist interaction with HNF4α

3.1.1. Molecular docking

Upon examining the structural characteristics of HNF4α agonists, we identified that Alverine, Benfluorex, NCT, and NFT all possess aromatic moieties, suggesting that these structural features may be fundamental to their pharmacological activity (Fig. 1). However, notable chemical differences are present among these compounds. Compared to Alverine and Benfluorex, NCT and NFT exhibit simpler architectures, each characterized by an amide bond conjugated to aromatic rings derived from caffeic acid and ferulic acid, respectively, whereas Alverine and Benfluorex contain amine groups linked to flexible alkyl chains. We hypothesize that the conjugated structures formed by the amide bonds and aromatic rings in NCT and NFT contribute to increased electron density, thereby enhancing their specificity for HNF4α.

To elucidate the preferred binding conformations of these agonists with HNF4α, we conducted molecular docking using DS software. The docking results indicated that NCT exhibited the highest binding affinity to HNF4α among the four agonists (Table 1). Furthermore, protein-ligand interaction diagrams provided a detailed visualization of the interaction patterns between these agonists and the active site pocket of HNF4α (Fig. 3). Specifically, Alverine's two phenyl rings formed π-π stacking interactions; one phenyl ring interacted with VAL190 through π-alkyl interactions, while the other phenyl ring participated in carbon, electrostatic, salt bridge, and conventional interactions with HNF4α (Fig. 3A). Benfluorex's phenyl ring interacted with VAL52 and LEU222 through π-π stacking and π-alkyl interactions, and further formed salt bridge and alkyl interactions with GLU363 and VAL70 (Fig. 3B). NCT established π-alkyl interactions with LEU204 and MET364, while its hydroxyl group interacted with LYS194 via electrostatic and conventional interactions. Another hydroxyl group on the aromatic ring engaged in conventional interactions with GLU363, and NCT also formed π-π T-shaped and π-sigma interactions (Fig. 3C). NFT, on the other hand, formed π-alkyl interactions with LEU187 and VAL208, and its methoxy and hydroxyl groups interacted with GLU191 and LYS194 through carbon-based interactions, while an additional hydroxyl group established conventional interactions with GLU363 (Fig. 3D). Based on our analysis of the binding modes of HNF4α agonists, we conclude that the aromatic ring plays a pivotal role in protein-ligand interactions, primarily by participating in electrostatic interactions with charged

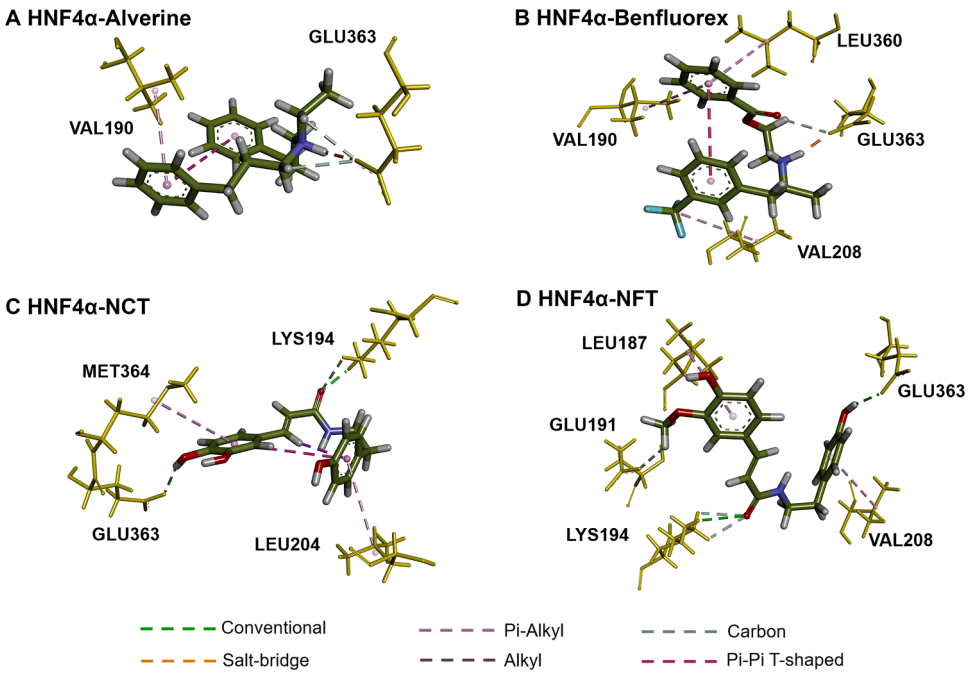

Fig. 3. Interaction pattern diagram of the docking conformation. (A) HNF4α-Alverine, (B) HNF4α-Benfluorex, (C) HNF4α-NCT, and (D) HNF4α-NFT in the LXXLL motifs of HNF4α.

amino acid residues within HNF4α, particularly GLU363. This observation underscores the significance of the aromatic ring as a key pharmacophore contributing to the pharmacological activity of these agonists.

3.1.2. *MD Simulation*

To evaluate the stability of the binding conformations of the four agonists within the LXXLL motif cavity of HNF4α, we performed 50-ns MD simulations using the optimal docking conformations. The MD trajectories provided insight into the dynamics and binding stability of the proteins and ligands over the simulation period. The conformations at the initial and final stages of simulations were compared using Gaussian 09 and PyMol (Fig. 4). The results revealed (suggested) conformational changes in the HNF4α-Alverine complex compared to the initial structure (Supplementary Fig. S1A), whereas the HNF4α-NCT and HNF4α-NFT complexes remained stable and exhibited extended conformations within the binding pocket (Fig. S1C—D). Throughout the MD simulations, HNF4α displayed limited conformational flexibility, with HNF4α-NCT notably transitioning from a folded to an extended conformation (Supplementary Fig. S1C).

The root mean square deviation (RMSD) was employed to monitor the temporal changes in the conformation of the complexes relative to their initial structures [31]. The RMSD profiles indicated fluctuations during the initial 5 ns, with all four complexes reaching equilibrium after 10 ns. The average RMSD values for the HNF4α-Alverine, HNF4α-Benfluorex, HNF4α-NCT, and HNF4α-NFT complexes were 2.47, 2.61, 2.48, and 2.30 Å, respectively (Fig. 5A). Based on RMSD fluctuations throughout the 50-ns simulation, the HNF4α-Alverine and HNF4α-Benfluorex complexes exhibited greater variability, whereas the HNF4α-NCT and HNF4α-NFT complexes converged to stable equilibrium values during the final 10 ns, suggesting that NCT and NFT display higher stability under these simulation conditions.

The root mean square fluctuation (RMSF) was used to assess the structural mobility of individual atoms relative to a reference conformation, thereby reflecting their degrees of freedom. The RMSF analysis indicated considerable fluctuations for the HNF4α-Alverine and HNF4α-Benfluorex complexes, whereas HNF4α-NCT and HNF4α-NFT exhibited comparatively stable fluctuations, with HNF4α-NCT showing the lowest

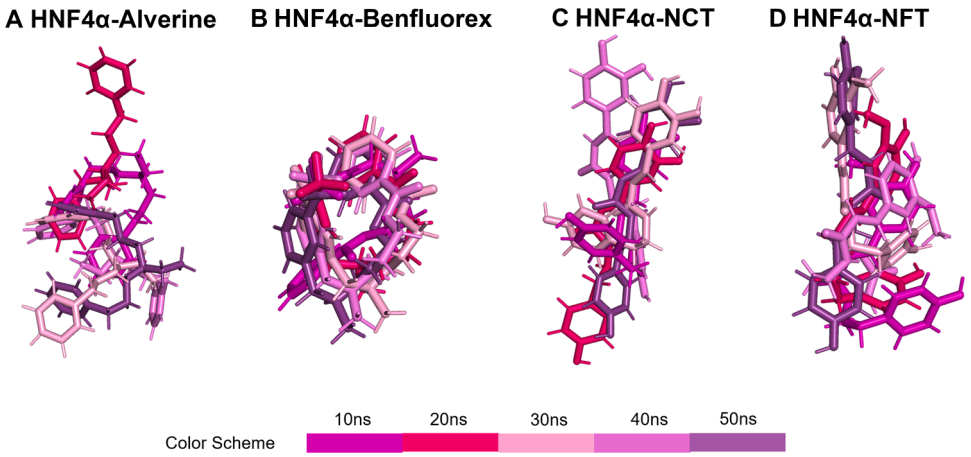

Fig. 4. The superposition of agonists in MD in LXXLL motif cavities and their docking binding states are formed at intervals of 10 ns.

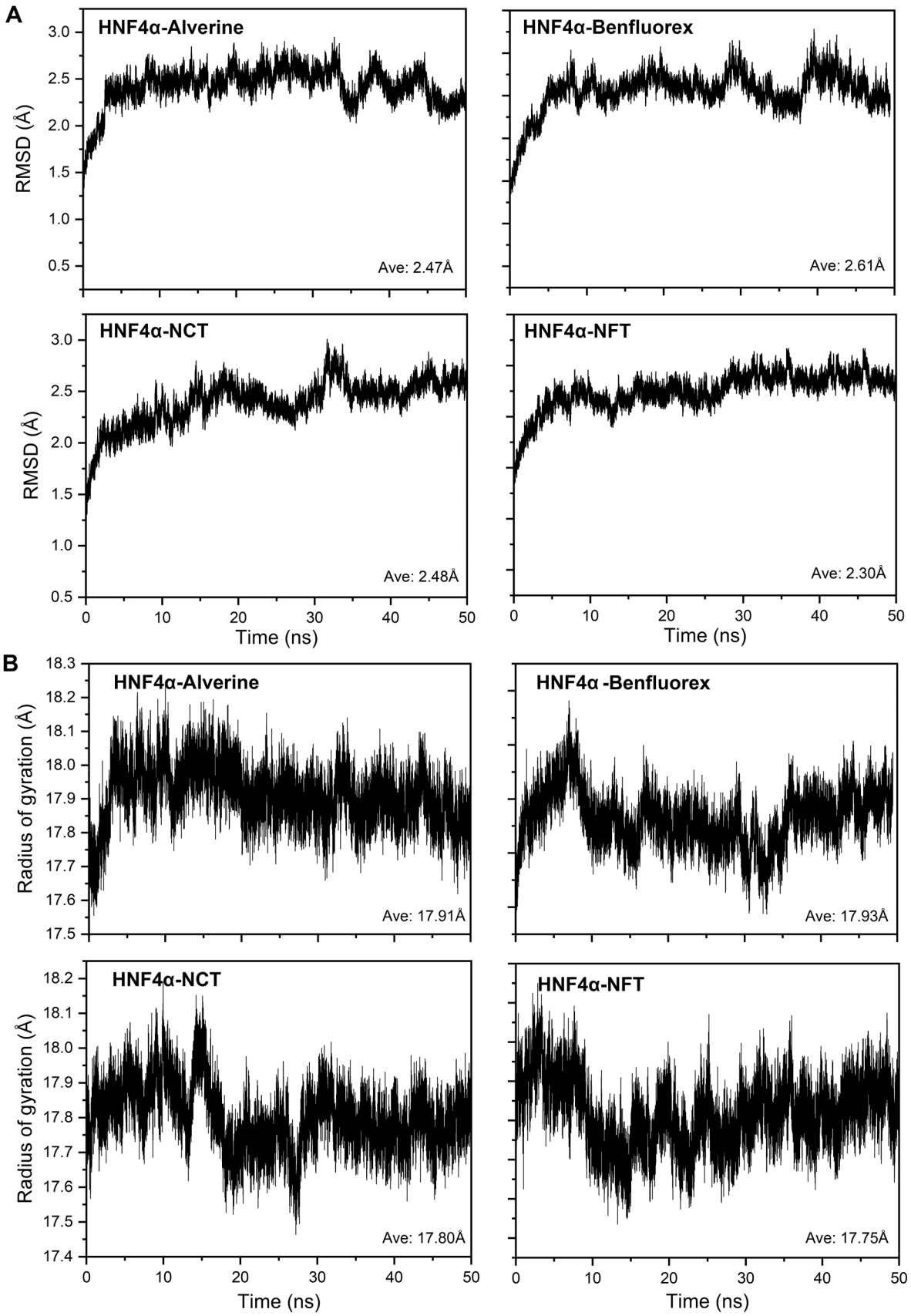

Fig. 5. Dynamics parameter changes during 50 ns MD simulation. (A) RMSD over time for complexes during 50 ns MD simulations. (B) Radius of gyration over time for complexes during 50 ns MD simulations.

RMSF values ranging from 9.75 to 18.45 Å, suggesting the most stable binding between NCT and HNF4α (Supplementary Fig. S2).

The radius of gyration (Rg) was calculated to assess the distribution of atomic mass along a specific axis, providing insights into the compactness and folding characteristics of the molecule during the simulations. The average Rg values for the HNF4α-Alverine and HNF4α-Benfluorex complexes were 17.91 and 17.93 Å, respectively. In contrast, the HNF4α-NCT and HNF4α-NFT complexes maintained lower and more stable average Rg values of 17.80 and 17.75 Å, respectively, indicating that binding with NCT and NFT may enhance the molecular compactness and stability of HNF4α (Fig. 5B).

In summary, these findings collectively indicate that the interactions of NCT and NFT with HNF4α exhibit relatively greater stability compared to those of Alverine and Benfluorex, suggesting that NCT and NFT may represent more promising candidates for further therapeutic exploration.

3.1.3. Calculation of binding energy

To further evaluate the binding affinities of the four agonists with HNF4α, we conducted 50-ns MD simulations on the protein-ligand complexes and analyzed the binding energies of the equilibrium segments of these systems [32] . Binding affinity is inversely correlated with binding energy; hence, a smaller (more negative) binding energy signifies a stronger protein ligand interaction. The results demonstrated that the HNF4α-NCT complex exhibited the most favorable (highest) binding energy, which was consistent with our docking scores and the experimental findings of Lee et al. (Fig. 6) [12] .

Furthermore, an analysis of the individual energy components contributing to the overall binding energy revealed that electrostatic interactions were the predominant contributors. This observation suggests that the distinct electronic structure of the aromatic rings present in the compounds facilitates interactions with charged residues in HNF4α, thereby significantly influencing the electrostatic component of the binding energy and resulting in a greater electrostatic contribution.

To elucidate the contribution of specific amino acid residues in HNF4α to the overall binding energy, we performed binding energy decomposition on the residues in proximity to the ligand and identified the ten residues with the highest contributions. The results indicated that GLU363 made the most significant contribution to the binding free energy in the HNF4α-NCT complex, suggesting that the interaction between NCT and the negatively charged GLU363 serves as the primary driving force for protein-ligand binding (Table 2).

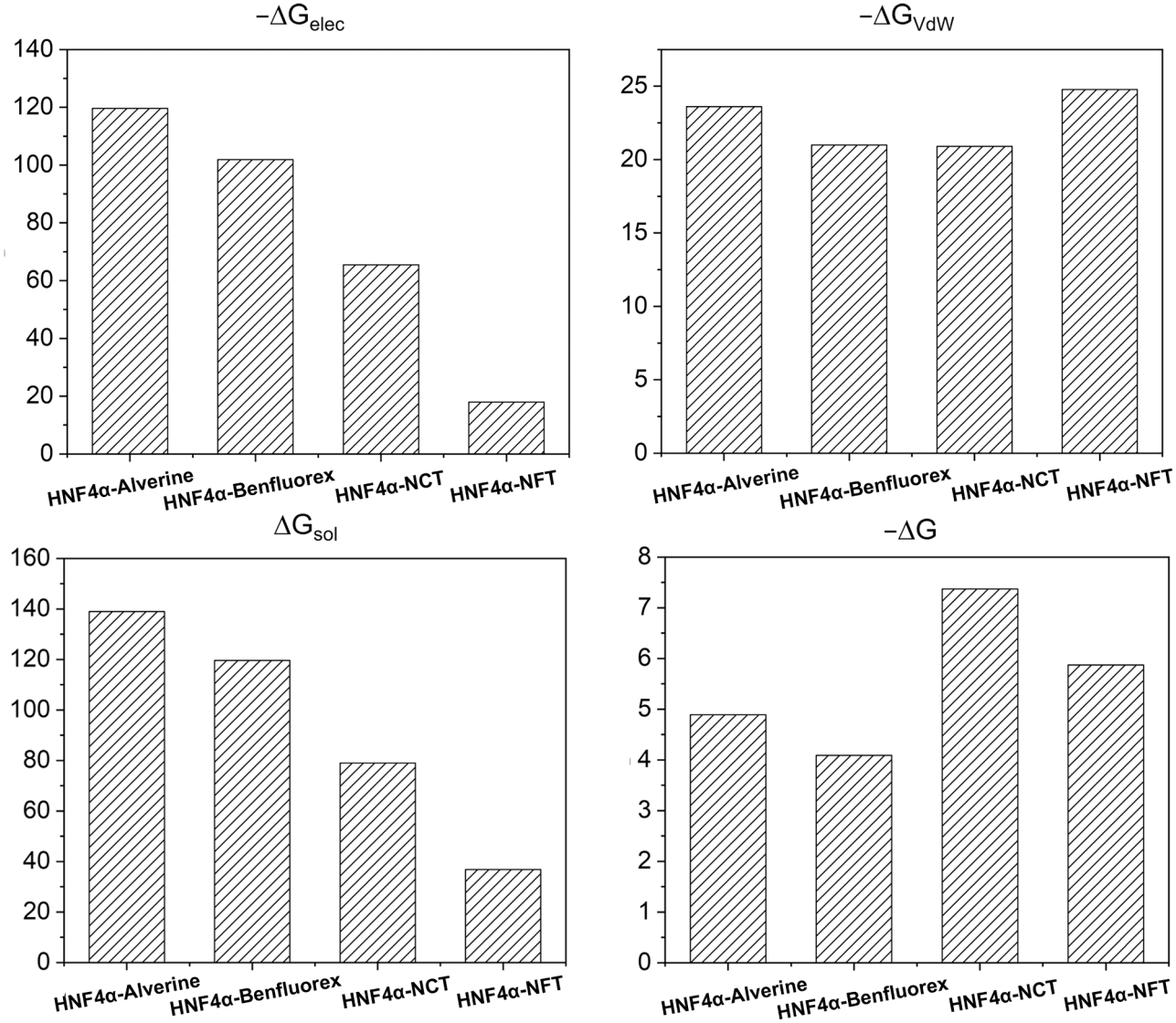

Fig. 6. Contributions of electrostaticenergy, van der Waals energy, solvation energy, and binding free energy during 50-ns MD simulations for HNF4α complexes. (A) HNF4α-Alverine, (B) HNF4α-Benfluorex, (C) HNF4α-NCT, (D) HNF4α-NFT.

Table 2
The top 10 amino acids of HNF4α that contribute the most to the binding free energy (ΔG, kcal/mol) of the four systems are listed in the table.

HNF4α-Alverine		HNF4α-Benfluorex		HNF4α-NCT		HNF4α-NFT	
Residue	ΔG	Residue	ΔG	Residue	ΔG	Residue	ΔG
VAL52	−2.17	LEU222	−1.27	GLU363	−5.81	LEU222	−2.22
VAL70	−1.36	VAL70	−1.12	LEU66	−2.08	GLU363	−1.79
GLU363	−1.30	VAL52	−1.06	VAL70	−1.73	ASN221	−1.74
PHE61	−1.02	LEU73	−0.97	LYS56	−1.66	LEU66	−1.66
LEU73	−0.99	MET363	−0.77	VAL52	−1.60	VAL52	−1.22
LEU222	−0.98	LEU66	−0.75	LEU222	−1.47	LEU73	−0.90
GLU53	−0.35	GLU225	−0.32	LEU73	−1.04	MET226	−0.84
GLN69	−0.28	ASP220	−0.25	MET226	−0.95	LYS56	−0.53
ALA55	−0.26	ASP67	−0.19	ASN221	−0.56	LEU49	−0.38
LEU49	−0.21	GLU53	−0.19	ARG74	−0.41	GLU225	−0.32

3.1.4. Hydrogen bond interaction analysis

Hydrogen bonds (H-bonds) are critical determinants of the stability of protein-ligand interactions. In a 50-ns MD simulation, the frequency and number of H-bonds formed between NCT and GLU363 were found to be the highest among the observed interactions (Fig. S3 & Table 3). This finding suggests that H-bonding with GLU363 plays a pivotal role in stabilizing the binding of NCT to HNF4α. Based on binding free energy decomposition analysis, we propose that GLU363 serves as a key residue mediating the interaction between NCT and HNF4α, thereby contributing significantly to the stability of the complex.

3.2. Compound library screening based on NCT structure

Based on the structural features of the four previously discussed agonists and the interaction pattern between NCT and HNF4α, we hypothesize that the highly conjugated electronic structure formed by the amide bonds and aromatic rings in NCT and NFT represents the key pharmacophore for the interaction with HNF4α. To identify novel lipid metabolism-regulating agents targeting HNF4α, we employed the core scaffold of NCT as a reference to screen a natural compound library for structurally similar candidates. An initial screening using DS software yielded 69 compounds with potential interactions at the HNF4α-LXXLL motif interface. Further refinement of these candidates through consensus scoring (CScore) and binding energy calculations resulted in the identification of four promising compounds: licochalcone B, rosmarinic acid, cynarin, and chicoric Acid. These compounds exhibit structural similarity to NCT and display favorable virtual interactions with the LXXLL motif (Fig. S4A—B).

Table 3
Frequency of hydrogen bonding between agonists and HNF4α in MD. The acceptor is any atom in the amino acid residue of HNF4α, and the donor is the atom in the agonists. In the MD process, the frequency of hydrogen bonds formed between the amino acid residues of the receptor protein and the donor agonist is expressed as a percentage.

Complexes	Acceptor	DonorH	Frac (%)
HNF4α-Alverine	LEU73-HB2	H25	5.27 %
	VAL70-HA	H26	4.90 %
HNF4α-Benfluorex	GLN69-OE1	H2	4.90 %
	LEU66-HB3	H3	3.82 %
HNF4α-NCT	GLU225-OE1	H16	75.73 %
	GLU225-OE1	H17	71.24 %
	GLU225-CD	H17	33.98 %
	GLU225-CD	H16	31.10 %
	GLU225-OE2	H16	23.85 %
	GLU225-OE2	H17	20.99 %
HNF4α-NFT	GLU225-OE1	H17	18.16 %
	GLU225-OE2	H17	17.26 %
	GLN69-OE1	H17	14.08 %
	GLU225-CD	H17	12.93 %

3.3. Physical interaction between rosmarinic acid and HNF4α

To validate the binding activity of these compounds with HNF4α and their potential to regulate lipid metabolism, we performed cellular experiments. HNF4α protein was labeled with a fluorescent probe and incubated with the four selected compounds, followed by microscale thermophoresis (MST) to assess protein-ligand binding. The dissociation constant (Kd) is the key parameter for evaluating the strength of molecular interactions. As shown in Fig. 7, rosmarinic acid exhibited the strongest binding affinity for HNF4α, with a Kd value of 34.9 μM. This binding affinity was superior to that of licochalcone B (78.2 μM) and cynarin (3106 μM). In contrast, no significant binding interaction was observed between chicoric Acid and HNF4α.

3.4. Rosmarinic acid inhibits cell proliferation and induces autophagy by activating the P1 promoter

To assess the effect of the virtual screening compounds on the viability of HepG2 cells, we treated HepG2 cells with different concentrations (0.78125 μM to 200 μM) of the compounds. Comparing the inhibitory activity of the four compounds on HepG2 cell viability, we found that rosmarinic acid and licochalcone B dose-dependently inhibited the proliferation of HepG2 cell (Fig. 8A). As previous reports, HNF4α can be driven by two types of promoters (P1 and P2) for expression [30]. The P1 promoter-driven isoform class is primarily expressed in differentiated, mature liver cells and regulates metabolic-related genes, with phosphorylation inhibiting its transcriptional activity. The P2 promoter-driven isoform class is mainly expressed in undifferentiated or highly proliferative cells, with phosphorylation generally enhancing its transcriptional activity and promoting proliferation. It is important to evaluate the potential different influence of candidate agonists on these isoforms. As reported, HepG2 cells have both P1 and P2 isoform expression [30]. So, we knocked down P2 isoform of HNF4α in HepG2 cells (Fig. S5) and selected DLD-1 cells, mainly expressing P2 expression, to evaluate the effect of the HNF4α agonist NCT and other candidate agonists. As shown in Fig. 8B, the knockdown of P2 isoform increased the sensitivity of HepG2 on the administration of NCT and rosmarinic acid. However, neither the HNF4α agonist NCT nor other candidate agonists inhibited the proliferation of DLD-1 cells (Fig. 9).

These results indicated that the proliferation of cells expressing P1 isoform of HNF4α might be more significantly influenced by the HNF4α agonists and the function of P1 isoform of HNF4α might be influenced by the expression of P2 isoform of HNF4α. The limited influence of HNF4α agonists on DLD-1 might be the different affinity of compounds on P1 isoform and P2 isoform, or the balance of intracellular signaling pathways might influence the response of cell to these agonists.

To explore the influence of candidate agonists on lipid metabolism, we treated HepG2 cells with 100 μM compounds for 24 h and the expression levels of key proteins in lipid metabolism were evaluated by Western-blot assay (Fig. 10 and S6). LC3I/II is a marker of autophagy,

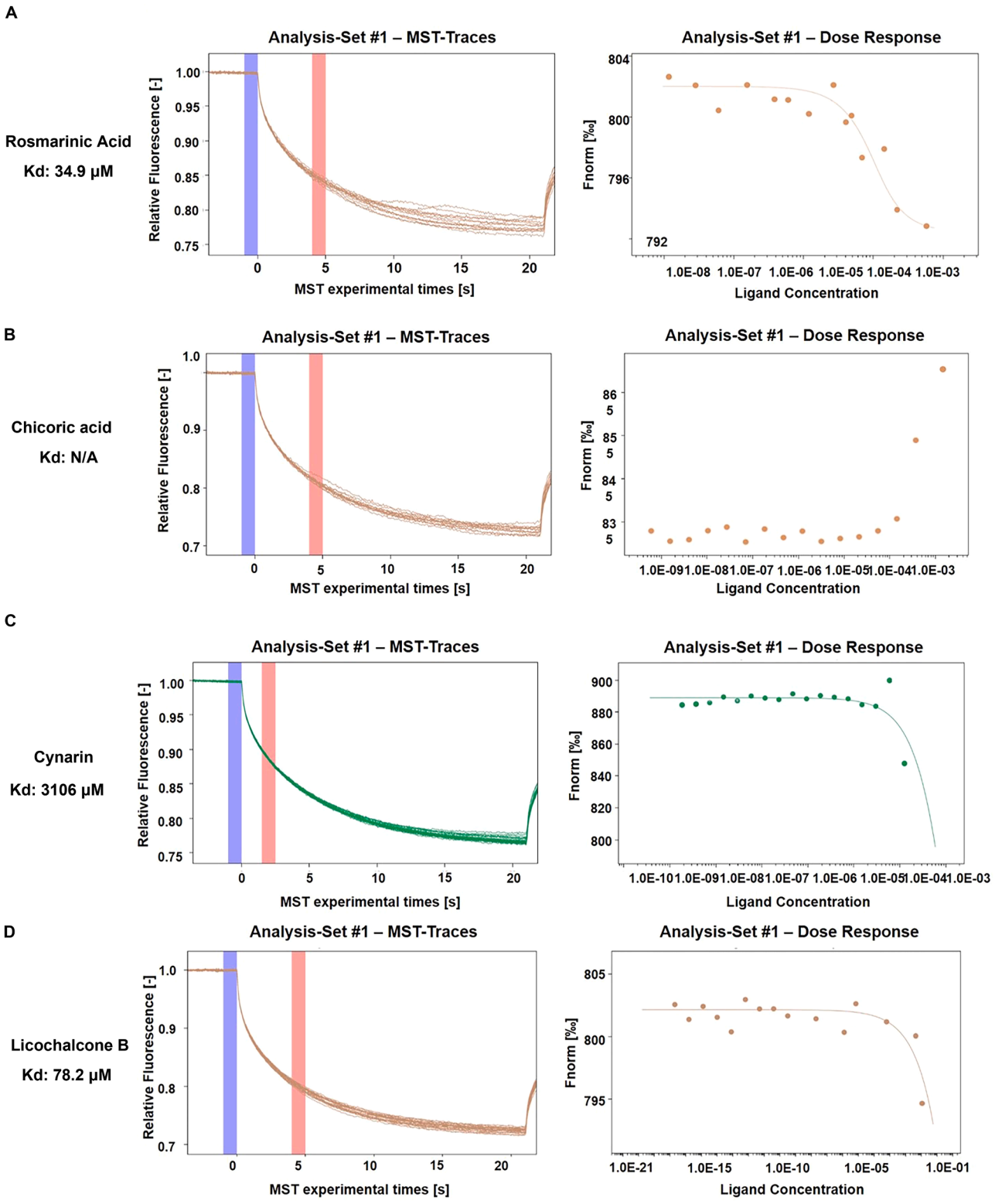

Fig. 7. MST analysis of HNF4α binding.

and changes in the LC3I/II ratio can be assessed by WB to evaluate autophagy formation. The results showed that the expression level of LC3I/II increased significantly after rosmarinic acid treatment, which is opposite to the result of the HNF4α antagonist BI6015 and is consistent with the result of the known HNF4α agonists NCT. Rosmarinic acid downregulates the mTOR signaling pathway, which plays an important role in cell growth, proliferation, and autophagy regulation. p-mTOR, the phosphorylated form of mTOR and usually upregulated during cell proliferation [33,34], is downregulated by rosmarinic acid. These results indicate its ability to inhibit the p-mTOR/mTOR signaling pathway and induce autophagy. Furthermore, the upregulation of Beclin1, an autophagy-related protein, also supports that rosmarinic acid enhances the autophagy in HepG2. Increased activity of HNF4α is reported to control liver fat storage through induction of autophagy related genes

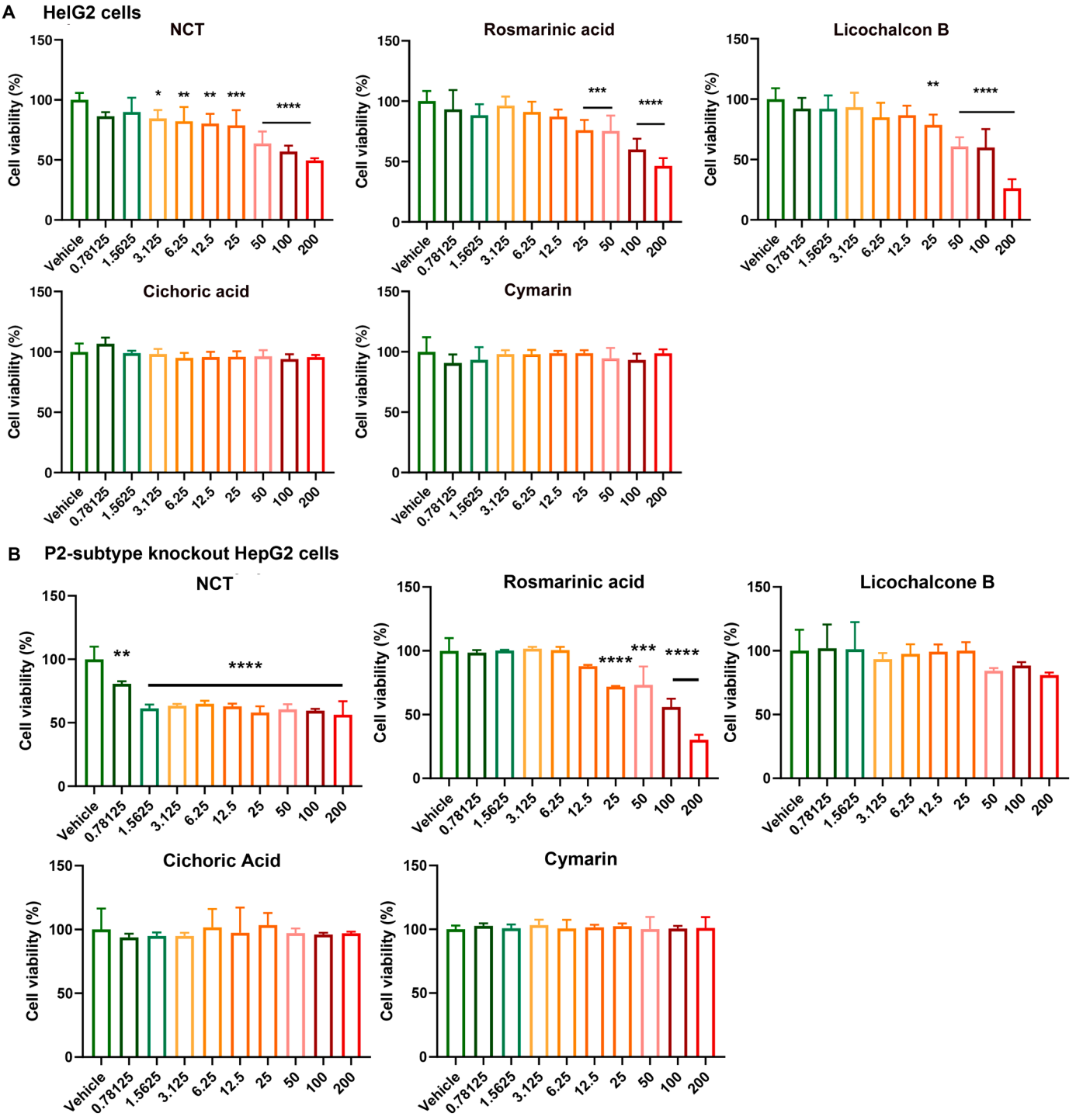

Fig. 8. Proliferation inhibition rates of HepG2 and P2-subtype knockout HepG2 cells. measured by the CCK-8 assay after 24-h treatment with various concentrations of NCT, rosmarinic Acid, licochalcone B, and chicoric Acid. (A) HepG2 cells, (B) P2-subtype knockout HepG2 cells.

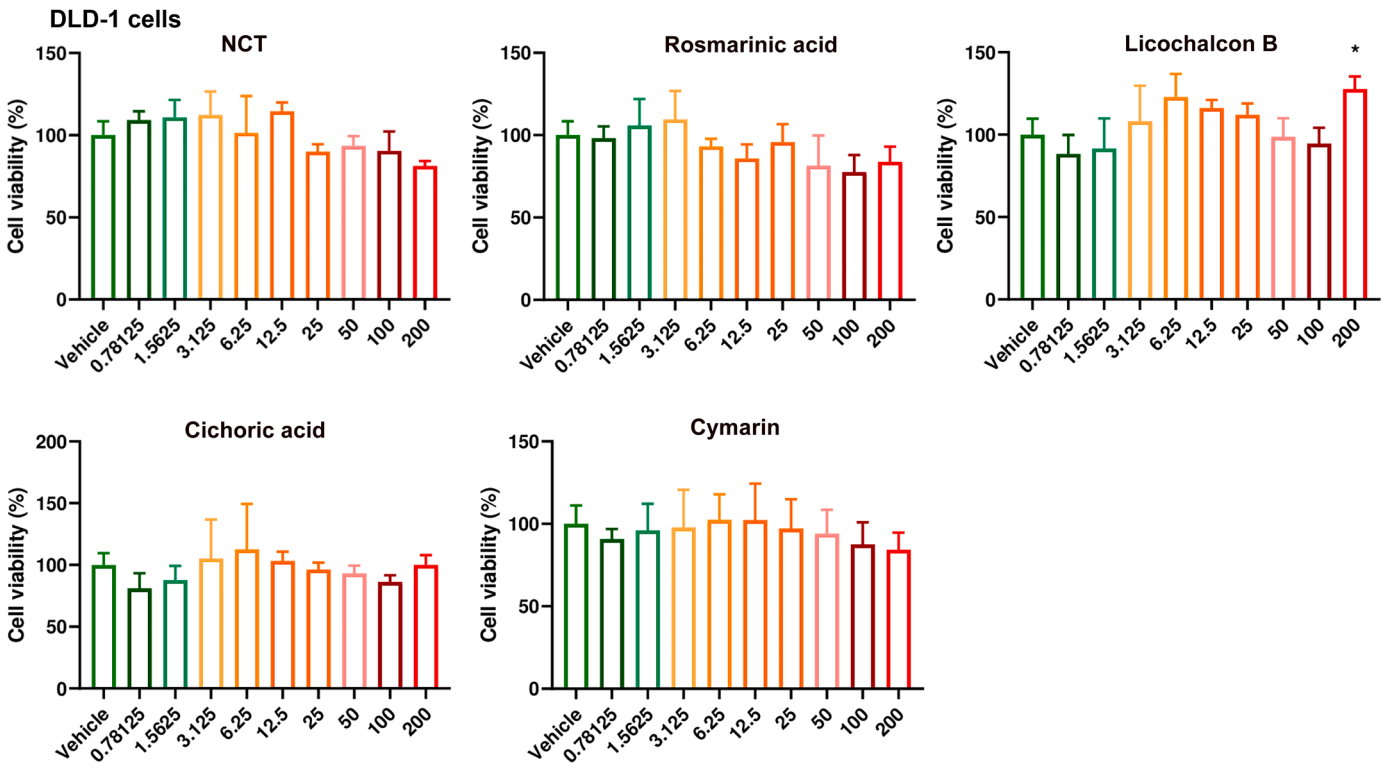

Fig. 9. Proliferation inhibition rates of DLD-1 cells measured by the CCK-8 assay after 24-h treatment with various concentrations of NCT, rosmarinic Acid, licochalcone B, and chicoric Acid.

[35] and it is similar to our results that both of NCT and rosmarinic acid increased the expression of LC3 I/II and p62. These results may explain the inhibitory effect of rosmarinic acid on HepG2 cell proliferation and its potential usage in the treatment for NAFLD.

Licochalcone B showed inhibitory effect and increased the autophagy on HepG2 cells, but its effect was not enhanced by the P2 isoform knockdown and it slightly promoted proliferation of DLD-1 in highest dose (Fig. 9 and Fig. S5). It indicated the effect of Licochalcone B is complex which may not only dependent on the HNF4α activity, such as the inhibition of NF-κB p65, the decreased production of ROS, increased apoptosis [36–38].

Therefore, rosmarinic acid primarily exert their inhibitory effects on cell proliferation and lipid metabolism without.

3.5. Discussion on the binding specicity of rosmarinic acid to LBD domain

In this study, the ligand-binding domain (LBD) fragment of HNF4α was utilized for screening, structural, and interaction analyses, owing to the pivotal role of the conserved LXXLL motif in agonist binding [16]. To evaluate the specificity of rosmarinic acid for the LXXLL binding pocket, global docking simulations were performed using the CD-DOCK2 online server [39], allowing us to explore potential interactions with other regions of HNF4α. Four potential binding sites for rosmarinic acid on HNF4α were identified (Fig. S5A). Sites 1 and 2 are located within the LXXLL motif of each chain, and consequently, they are identical in structure. Sites 3 and 4 are positioned near residues 305—308 and 316—318, respectively (Fig. S5A).

Further MD simulations revealed that rosmarinic acid, initially bound to Site 3, dissociated from this binding site around 20 ns, suggesting an unstable interaction at this location (Fig. S7B). The binding energies of rosmarinic acid at Sites 1 and 4 were found to be −16.5 and −6.7 kcal/mol, respectively (Fig. S7C), indicating a strong binding preference for the LXXLL motif. These findings highlight the binding specificity of rosmarinic acid for the LXXLL motif of HNF4α. However, it is important to note that the interactions of rosmarinic acid with other regions of HNF4α remain underexplored, and further experimental investigations are needed to fully characterize these potential interactions.

4. Conclusion

This study systematically identified potential HNF4α agonists and elucidated their binding modes through both computational and experimental approaches. The findings underscore critical interactions and key residues that contribute to binding stability. High-throughput virtual screening identified rosmarinic acid as a potent HNF4α agonist, demonstrating significant biological activity, including the inhibition of HepG2 cell proliferation and the induction of autophagy. The benefit effect of rosmarinic acid might be more related to the function of P1 isoform of HNF4α. These results offer a robust theoretical foundation for the development of novel therapeutics targeting NAFLD and related lipid metabolism dysregulation. Although the influences of rosmarinic acid on the expression levels of key proteins in autophagy signaling pathway were evaluated, the direct influences of rosmarinic acid on DNA binding and transcription ability of HNF4α are still needed to be investigated in future to explore the mechanism of rosmarinic acid on NAFLD through HNF4α.

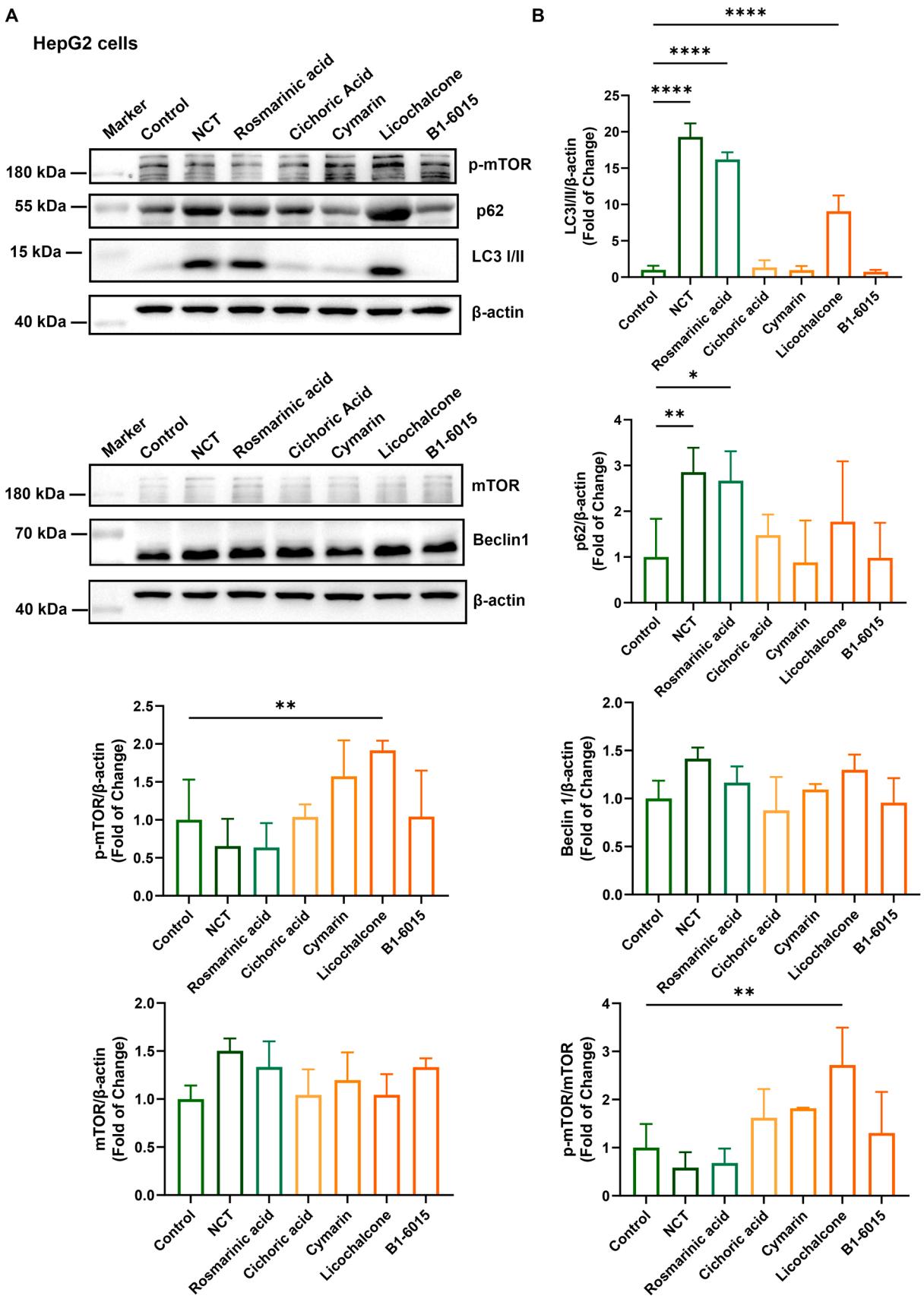

Fig. 10. Expression of p-mTOR, mTOR, p62, Beclin1, LC3I/II, and β-actin in HepG2 cells treated with 100 μM compounds for 24 h, as detected by WB experiment. (A) Immunoblot bands, (B) Quantification result of the immunoblot bands.

Funding

This work was supported by grants from the Natural Science Foundation of Jiangsu Province (BK20211254), Graduate Practice Innovation Program of Jiangsu Province (SJCX22_0677 and SJCX23_0717).

Author's statement

We, the authors, declare that the manuscript titled " Structure-Based Identification of HNF4α Agonists: Rosmarinic Acid as a Promising Candidate for NAFLD Treatment" represents original research, and that it has not been published elsewhere nor is it currently under consideration for publication by any other journal. All authors have contributed significantly to the work and have approved the final version of the manuscript.

We confirm that the study was conducted in accordance with ethical standards, and appropriate approvals were obtained for experiments involving human or animal subjects (if applicable). We also acknowledge any financial support received for this research and declare any conflicts of interest.

The data supporting the findings of this study are available from the corresponding author upon reasonable request. We are committed to ensuring that the manuscript meets the highest standards of academic integrity and scientific rigor.

CRediT authorship contribution statement

Nan Jiang: Writing – review & editing, Validation, Supervision, Resources, Funding acquisition, Conceptualization. **Xibo Wang:** Validation, Methodology. **Xi Chen:** Validation, Methodology, Data curation. **Xinqi Zhu:** Validation, Methodology, Data curation. **Gang Wu:** Validation, Methodology. **Yu Zhang:** Supervision, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2024.12.014.

References

[1] Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016;64:73–84.
[2] Kasper P, Martin A, Lang S, et al. NAFLD and cardiovascular diseases: a clinical review. Clin Res Cardiol 2021;110:921–37.
[3] Powell EE, Wong VW, Rinella M. Non-alcoholic fatty liver disease. Lancet 2021; 397:2212–24.
[4] Younossi Z, Anstee QM, Marietti M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol 2018;15:11–20.
[5] Targher G, Byrne CD, Tilg H. NAFLD and increased risk of cardiovascular disease: clinical associations, pathophysiological mechanisms and pharmacological implications. Gut 2020;69:1691–705.
[6] Dhe-Paganon S, Duda K, Iwamoto M, et al. Crystal structure of the HNF4 alpha ligand binding domain in complex with endogenous fatty acid ligand. J Biol Chem 2002;277:37973–6.
[7] Parviz F, Matullo C, Garrison WD, et al. Hepatocyte nuclear factor 4alpha controls the development of a hepatic epithelium and liver morphogenesis. Nat Genet 2003; 34:292–6.
[8] Gonzalez FJ. Regulation of hepatocyte nuclear factor 4 alpha-mediated transcription. Drug MetabPharmacokinet 2008;23:2–7.
[9] Hwang-Verslues WW, Sladek FM. HNF4α—role in drug metabolism and potential drug target? Curr OpinPharmacol 2010;10:698–705.
[10] Martinez-Jimenez CP, Kyrmizi I, Cardot P, et al. Hepatocyte nuclear factor 4alpha coordinates a transcription factor network regulating hepatic fatty acid metabolism. Mol Cell Biol 2010;30:565–77.
[11] Li Y, Zalzala M, Jadhav K, et al. Carboxylesterase 2 prevents liver steatosis by modulating lipolysis, endoplasmic reticulum stress, and lipogenesis and is regulated by hepatocyte nuclear factor 4 alpha in mice. Hepatology 2016;63: 1860–74.
[12] Lee SH, Veeriah V, Levine F. Liver fat storage is controlled by HNF4α through induction of lipophagy and is reversed by a potent HNF4α agonist. Cell Death Dis 2021;12:603.
[13] Accelrys Inc. Discovery Studio Visualizer Software, Version 4.0, 2012.
[14] Lee SH, Veeriah V, Levine F. Liver fat storage is controlled by HNF4α through induction of lipophagy and is reversed by a potent HNF4α agonist. Cell Death Dis 2021;12:603.
[15] Vanommeslaeghe K, Hatcher E, Acharya C, et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem 2010;31:671–90.
[16] Rha GB, Wu G, Shoelson SE, et al. Multiple binding modes between HNF4α and the LXXLL motifs of PGC-1α lead to full activation. J Biol Chem 2009;284:35165–76.
[17] G.B. Rha, G. Wu, S.E. Shoelson, et al. "Multiple binding modes between HNF4α and the LXXLL motifs of PGC-1α lead to full activation." (2009).
[18] Wu G, Robertson DH, Brooks 3rd CL, et al. Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm. J Comput Chem 2003;24:1549–62.
[19] Berendsen HJC, Postma JPM, van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–90.
[20] Case DA, Betz RM, Cerutti DS, et al. AMBER 18 [CP]. San Francisco, CA: University of California; 2018.
[21] Frisch MJ, Trucks GW, Schlegel HB, et al. Gaussian 09 (revision D.01) [CP]. Wallingford, CT: Gaussian Inc; 2009.
[22] Ditchfield R, Hehre WJ, Pople JA. Self-consistent molecular orbital methods. 9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. J Chem Phys 1971;54:724.
[23] Wang J, Wolf RM, Caldwell JW, et al. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–74.
[24] Maier JA, Martinez C, Kasavajhala K, et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11(8):3696–713.
[25] Wang E, Sun H, Wang J, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. Chem Rev 2019; 119:9478–508.
[26] Wang C, Nguyen PH, Pham K, et al. Calculating protein-ligand binding affinities with MMPBSA: Method and error analysis. J Comput Chem 2016;37:2436–46.
[27] Xu B, Shen H, Zhu X, et al. Fast and accurate computation schemes for evaluating vibrational entropy of proteins. J Comput Chem 2011;32:3188–93.
[28] Tripos Inc. SYBYL-X 1.3 [EB/OL].
[29] Spitzer R, Jain AN. Surflex-dock: docking benchmarks and real-world application. J Comput Aided Mol Des 2012;26:687–99.
[30] Babeu JP, Jones C, Geha S, Carrier JC, Boudreau F. P1 promoter-driven HNF4α isoforms are specifically repressed by β-catenin signaling in colorectal cancer cells. J Cell Sci 2018;131.
[31] Lohning AE, Levonis SM, Williams-Noonan B, Schweiker SS. A practical guide to molecular docking and homology modelling for medicinal chemists. Curr Top Med Chem 2017;17(18):2023–40.
[32] Berendsen HJC, Postma JPM, van Gunsteren WF, et al. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–90.
[33] Yu JS, Cui W. Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. Development 2016;143(17): 3050–60.
[34] Peng Y, Wang Y, Zhou C, Mei W, Zeng C. PI3K/Akt/mTOR Pathway and its role in cancer therapeutics: are we making headway? Front Oncol 2022;12:819128.
[35] Oh HN, Lee MH, Kim E, Yoon G, Chae JI, Shim JH. Licochalcone B inhibits growth and induces apoptosis of human non-small-cell lung cancer cells by dual targeting of EGFR and MET. Phytomedicine 2019;63:153014.
[36] Li Q, Feng H, Wang H, et al. Licochalcone B specifically inhibits the NLRP3 inflammasome by disrupting NEK7-NLRP3 interaction. EMBO Rep 2022;23: e53499.
[37] Zhang J, Si H, Lv K, et al. Licarin-B Exhibits Activity Against the *Toxoplasma gondii* RH Strain by Damaging Mitochondria and Activating Autophagy. Front Cell Dev Biol 2021;9:684393.
[38] Oh HN, Lee MH, Kim E, Yoon G, Chae JI, Shim JH. Licochalcone B inhibits growth and induces apoptosis of human non-small-cell lung cancer cells by dual targeting of EGFR and MET. Phytomedicine 2019;63:153014.
[39] Yang Liu, et al. CB-Dock2: improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. Nucleic Acids Res 2022.