ROYAL SOCIETY OF CHEMISTRY

Check for updates

Database mining of ZINC15 natural compounds reveals potential thyroid receptor β agonists for NAFLD management: an *in silico* study

Ahmet Buğra Ortaakarsu, [iD] *[a] Michel Hosny,[b] Mansour Sobeh[c] and Mohamed A. O. Abdelfattah[d]

Non-alcoholic fatty liver disease (NAFLD) is a prevalent metabolic disorder with limited therapeutic options. Thyroid receptor β (THR-β) agonists have been showing promise for controlling NAFLD *via* improving hepatic lipid metabolism. This study utilized different *in silico* tools to screen 47 199 natural compounds from the ZINC15 database to identify potential THR-β agonists. Molecular docking, molecular dynamics simulations, and advanced analyses such as PCA, TICA-FES, and MSM revealed that 4-*O*-caffeoylquinic acid (compounds **2**) and dihydroxydehydrodiconiferyl alcohol (compound **18**) are the most promising hits. Both demonstrated high binding affinity and stable agonist interactions with key THR-β residues such as Arg316 and Arg320 that stabilize the ligand binding pocket and support the agonist potential, comparable to the reference agonists resmetirom and {3,5-dichloro-4-[4-hydroxy-3-(propan-2-yl) phenoxy]phenyl}acetic acid. Long-term MD simulations confirmed their stability, and MM/GBSA calculations supported robust thermodynamic profiles. Moreover, the two hits displayed superior selectivity for THR-β over THR-α and favorable pharmacokinetic profiles with minimal toxicity alerts. These findings support compounds **2** and **18** as strong candidates for NAFLD therapy, warranting further experimental validation.

1. Introduction

Nonalcoholic fatty liver disease (NAFLD) is the most prevalent chronic hepatic disorder in the western world and a leading cause of liver diseases worldwide. As per recent estimates, 32% of individuals, globally, are diagnosed with NAFLD with higher prevalence in men (40%) than in women (26%).[1] Currently, the detrimental effects of NAFLD are becoming a greater concern for public health owing to the rising rates of diabetes and obesity across the globe.[2]

Primarily, NAFLD results from a calorie-rich diet, high intake of fructose and animal protein, lack of physical activity, and a sedentary life style in general. Obesity and type 2 diabetes mellitus are major risk factors for the development of steatosis, which is the first stage in the NAFLD spectrum. Steatosis is characterized by accumulation of fat in 5% or more of hepatocytes, when there is no obvious etiology such as drug or alcohol intake. It can then progress to a more serious condition known as nonalcoholic steatohepatitis (NASH) that is associated with lobular inflammation, hepatocyte damage, along with variable degrees of fibrosis and cirrhosis in liver. Although NASH is generally a nonprogressive disease, it might develop into liver failure and even hepatocellular carcinoma in some patients.[3,4]

Several research studies have demonstrated that disorders of significant signaling pathways play a crucial role in the pathogenesis of fatty liver disease such as sirtuin 1 (SIRT1), adenosine 5′-monophosphate-activated protein kinase (AMPK), toll-like receptor 4 (TLR4), peroxisome proliferator-activated receptor (PPAR), nuclear factor kappa-B (NF-κB), and nuclear factor erythroid 2-related factor 2 (Nrf2).[5–8]

The beneficial effects of thyroid hormones, which are key regulators of metabolism in humans, are mediated by thyroid receptors with two isoforms: THR-α and THR-β. The former is prevailing in the brain, heart, and bones, whereas the latter is predominant in the liver, brain, kidneys, and pituitary gland.[8,9] While the THR-α isoform plays a crucial role in neuronal development, TRβ is proposed to have a role in the remyelination processes. Moreover, THR-β is responsible for the thyroid hormones' beneficial effects on the lipid metabolism in the liver.[10,11] The structural differences between THR-α and THR-β are shown in Fig. 1.

Thus, considerable efforts in the last two decades were devoted to developing selective agonists towards THR-β in order to be effective in managing metabolic and/or neuronal

[a]Department of Chemistry, Faculty of Science, Gazi University, Turkey. E-mail: bugra@ortaakarsu.com

[b]American University of the Middle East, Kuwait

[c]AgroBioSciences Program, College of Agriculture and Environmental Sciences, University Mohammed VI Polytechnic, Ben-Guerir 43150, Morocco

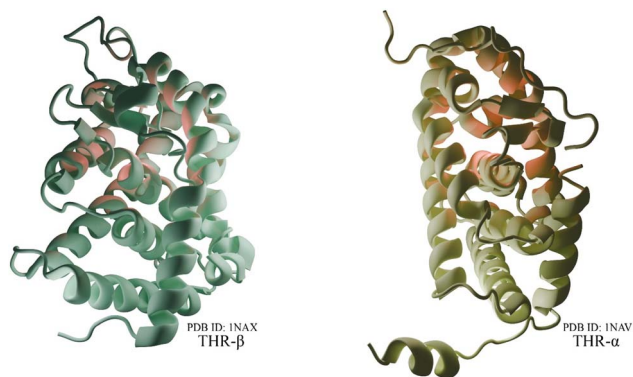*[d]College of Engineering and Technology, American University of the Middle East, Kuwait*

Fig. 1 Structural differences between THR-α and THR-β. The red illuminated area is the centre of action of the active zone centre.

disorders without exerting serious adverse effects on the bone and heart. For many years, the two molecules sobetirome and eprotirome were extensively investigated. They showed great promise for treating lipid metabolism diseases but failed in the clinical trials because of their unwanted adverse effects.[12] Resmetirom (MGL-3196) was the first FDA approved oral THR-β agonist in March 2024 to treat noncirrhotic NASH with moderate or advanced hepatic fibrosis. The drug targets the primary underlying causes of the metabolic disorder associated with NASH, and it should be administered concomitant with diet management and exercise.[13]

On the other hand, various studies revealed the beneficial effects of some plant-based phytochemicals and bioactive components in reducing the severe burden in patients with NAFLD. Among these, curcumin, cryptochlorogenic acid, epigallocatechin-3-gallate, caffeic acid, trigonelline, caffeine, 4-*p*-coumaroylquinic acid, salidroside, allicin, resveratrol, and silybin are extensively investigated and reported.[14] The potential underlying mechanisms related to the mode of action of these compounds included inhibition of oxidative stress and inflammation in liver, downregulation of fatty acid synthase expression, increasing the abundance of *Lactobacillus* in gut, hindering apoptosis and autophagy, reducing TNF-α production, and upregulating some pathways such as Nrf2, PPAR-α, AMPK, and fatty acid β-oxidation pathways.[15–17] To date, there are no comprehensive studies investigating the agonist potential of plant-based phytochemicals towards the THR-β receptor.

In the drug discovery process nowadays, where in situations where efficacy is prioritised on efficacy, diversity, and speed, various computer-aided approaches have enabled the virtual screening of large libraries of compounds, reduced experimental costs, and provided faster and more accurate calculations for parameters such as the estimated binding affinity of a drug to its target and various physico-chemical properties of newly designed drug candidates. These advances have all contributed to a rapid and more efficient drug design process.[18]

In this work, we aimed to evaluate the binding affinity and agonist potential of natural compounds in the ZINC15 database towards the THR-β receptor. For this purpose, four rounds of molecular docking were adopted, followed by short- and long-term molecular dynamics studies. The top performing compounds were subjected to thorough analysis of protein dynamics, including principle component analysis (PCA), Time-Lagged Independent Component Analysis-Free Energy Surface (TICA-FES), and Markov State Model (MSM) analyses, and their drug likeness, ADME, and toxicity profiles were evaluated. Furthermore, the selectivity of the top performing compounds towards the beta isoform of the receptor was investigated through molecular docking and molecular dynamics studies done on the alpha isoform.

2. Computational methods

In order to examine the effect of phytochemicals in the ZINC15 database on THR-β and to identify potential agonist compounds, a series of *in silico* methods were employed, including molecular docking,[19] molecular dynamics (MD),[20] assessment of drug likeness, pharmacokinetics (ADME) and toxicity profiles.[21] In addition, PCA,[22] TICA-FES with reduction free energy surface,[23] and MSM[24] were used to analyse the long-term effects of the top performing compounds on the protein structure. Unless otherwise stated, the Schrödinger Software Company's Maestro[25] software package and its modules were used throughout the study.

2.1. Molecular docking

2.1.1. Preparation of ligands. The ligands employed in the study comprised 47 199 compounds, within the natural compound class in the ZINC15 (ref. 26) database. The codes of the ligands utilized in the study are provided in the SI 1. Furthermore, resmetirom[27] and the cognate ligand bound to THR-β (PDB ID: 1NAX), namely {3,5-dichloro-4-[4-hydroxy-3-(propan-2-yl)phenoxy]phenyl}acetic acid, were employed in the *in silico* studies as reference agonists. The ionisation states of all compounds whose 3D structures were downloaded were estimated for pH 7.4 using the Epik[28] embedded software in the LigPrep[29] module. The ten most stable conformations were generated by considering tautomerisation states and subjected to the OPLS4 (ref. 30) force field.

2.1.2. Preparation of the protein structure. The X-ray crystal structure of THR-β receptor in complex with {3,5-dichloro-4-[4-hydroxy-3-(propan-2-yl)phenoxy]phenyl}acetic acid (cognate ligand) was downloaded from the protein data bank (PDB ID: 1NAX, 2.70 Å resolution).[31] The protein was appropriately coated with water molecules using HydraProt,[32] which was trained with deep learning models to accurately detect the positions of water molecules in 4000 crystal protein structures in the PDB. This enables the complete and accurate modelling of the solvent effect. The structure was then transferred to the Protein Preparation Wizard[33] and subjected to preprocessing, which included the addition of hydrogens, the formation of disulfide bonds and the estimation of the ionisation state of the protein structure at pH 7.4. Subsequently, the hydrogen bonds were optimised for pH 7.4 and minimised with the OPLS4 force field.

2.1.3. Molecular docking protocols. The binding affinity of all the natural ligands towards the cognate ligand site in the

THR-β receptor was evaluated using the Glide[34] module. Docking the ligands involved four rounds, including high-throughput virtual screening (HTVS), standard sensitivity (SP), extra sensitivity (XP), and induced fit docking (IFD). A total of 20 poses were retained for each ligand throughout the whole docking process. Elimination of irrelevant compounds following each docking round, utilizing the settings with the lowest root mean square deviation (RMSD) values, is outlined as follows: after the first docking round conducted with HTVS, a total of 7974 ligands were successfully placed within the agonist binding site of the target receptor and ranked based on their binding affinity to the receptor. The best 50% of these ligands (3984 compounds) were subjected to the second docking round with SP to obtain 3951 top ranking ligands. Consequently, the best 20% of these ligands (791 of them) underwent docking with XP. Among these, 78 ligands ($\approx 10\%$ of the ligands retained from the previous round) were able to reveal appreciable binding affinity and showed similar binding interactions to those obtained by the reference agonists. Thus, these 78 ligands were finally employed in IFD (flexible docking) and ranked according to their docking scores.[35] To proceed with molecular dynamics simulations, we selected the top 22 ligands that represent approximately 20% of the ligands retained after the induced fit docking. A redocking study was performed for protocol validation and presented as Fig. S1. The ratios per round were determined statistically in order to prevent the compounds that may be effective under physical conditions from being missed during the study. Assuming HTVS scores follow a normal distribution, dropping the lowest-scoring 50% removes ligands about one standard deviation below the mean, while moving 50% into SP keeps false negatives low, thanks to the strong HTVS–SP agreement. Cutting to the top 20% in XP and then the top 10% in IFD, at a Benjamini–Hochberg FDR $\approx$ 0.1, channels resources toward quartile–decile leaders and pares false positives. Finally, forwarding the best 20% to MD and confirming flexible agonist-like binding motifs averts conformational escape and safeguards physiologically active hits. The force field used in the molecular docking calculation was OPLS4, and it was chosen because the interactions of the force field with the active site residues were calculated similarly to the interactions of the protein crystal structure. In particular, it was found that the carboxylic acid end rotating away from the arginine residues, which is one of the key interactions of the OPLS2005 force field, is more similar to the crystal data in OPLS4. The data showing this are given in Fig. S3 and Table S3.

2.2. Molecular dynamics studies

The obtained top 22 ligands, following IFD, were subject to short-term (10 ns) molecular dynamics (MD) simulations to examine the dynamic behavior of the protein–ligand complexes in the cellular environment within a limited time frame. This helps ascertain the continuity of the interactions, the stability of the complexes and the effects of the ligands on the protein structure.[36] Noteworthily, the ligands' data obtained from the short-term MD study were compared with those of the reference agonists (resmetirom and the cognate ligand), which were

employed in the same study. Six ligands that demonstrated the most similar behavior to the reference agonists in the short-term MD, as well as the two reference agonists, were then subject to a long-term MD study (500 ns).

2.2.1. System setup. The selected protein–ligand complexes were immersed in an orthorhombic solvent box of dimensions 10 Å × 10 Å × 10 Å. The solvent box consisted of water molecules of TIP3P[37] type. System neutralization was done by the addition of 0.15 M NaCl, which is the equivalent of the intracellular physiological medium, utilizing the Monte Carlo method. The system setup was completed by applying the OPLS4 (ref. 30) force field. The OPLS4 force field was used because of its success in the calculation of protein–ligand interactions, especially non-covalent interactions. In addition, the hybrid machine-learning & physics-based fitting strategy of the OPLS force field was found to be advantageous compared to other force fields because it produces realistic data under physical conditions.

2.2.2. Molecular dynamics system protocols. As MD simulations necessitate comprehensive control over the system, the simulation was conducted with a constant number of particles at constant temperature and pressure. This was achieved through the introduction of high energy barriers around the solvent box. The simulation was conducted at a constant temperature of 310 K, utilising a Nose–Hoover thermostat.[38,39] A constant pressure of 1 bar was maintained using a Martyna–Tobias–Klein barostat.[40] The Desmond module was employed to perform the simulations, which comprised two distinct stages. The initial stage entailed a 2 ps relaxation phase, while the subsequent stage encompassed 10 ns and 500 ns simulation generation phase for short-term and long-term simulations, respectively. Each simulation was performed as three independent simulations initiated with 310 K random Maxwell–Boltzmann velocity seeds modified from the configuration file. The portion of each iteration up to the point labelled as equilibrium was discarded. In this particular context, given that the Desmond module's algorithms were employed during the equilibration process, the consequences of atoms released without evaluating force field effects were meticulously examined in the ps regime. Subsequently, community averages were evaluated on non-overlapping blocks at the nanosecond level. The non-overlapping blocks were represented by RMSD values sliced at 25 ns for the three different simulations. Each iteration of the simulation was executed by analysing the mean values of atomic positions that attained equilibrium positions.

2.3. Analyses of protein dynamics

The top performing two compounds identified based on the long-term MD simulation results were subjected to detailed protein structure dynamics analyses along with the two reference agonists for comparison.

The detailed analyses included PCA, in which the multidimensional and complex datasets comprising the MD data were employed. By reducing the system into two dimensions and constructing energy surfaces, this approach elucidates the conformational shifts within the protein structure and the fundamental

configurations restricted by the energy barriers.[41] In addition, TICA-FES analysis was conducted with lag values of 10 (5 ns) and 100 (50 ns) to detect the short-term and minor conformational alterations induced by ligands in protein structures. Conversely, a lag value of 500 (250 ns) was employed to identify events such as protein folding, which are likely to occur within the protein structure in the long-term dynamics.[42,43] Finally, MSM analysis was employed to determine the time-dependent transition probabilities of the changes in protein conformation and to detect the changes in the energy landscape. Frame intervals containing related conformations, with the same energy profiles, were clustered using the K-means algorithm. The energies of other ensembles were calculated based on the energies of the frame series that had the lowest energy.[44–46] The Python scripts used to perform these analyses are available on Zenodo (https://doi.org/10.5281/zenodo.17194142) and GitHub (https://github.com/cannabinoid13/MDScripts). Version updates were made to all scripts under a single version number, and the first version was used in this article. The libraries used in the scripts included NumPy[47] v2.0, MDAnalysis[48] v2.7.0, PyEMMA[49] v2.5.7, Matplotlib[50] v3.9.0, and deeptime[51] v0.4.4.

2.4. Molecular mechanics with generalized born surface area (MM/GBSA)

The stability of the protein–ligand complexes formed by the two identified hits and the two reference agonists was evaluated by MM/GBSA at 200 frame intervals, within the MD simulations, using the Prime[52] module. The calculation was performed with the VSGB[53] solvent model and the OPLS4 force field.

2.5. Evaluating the selectivity towards the THR-β receptor

The same previously adopted molecular docking and molecular dynamics protocols were employed to estimate the binding affinity of the two identified hits towards the second isoform of the thyroid receptors, THR-α (PDB ID: 1NAV).

2.6. Drug likeness, pharmacokinetics (ADME), and toxicity profile analyses
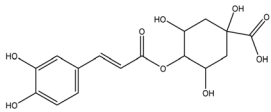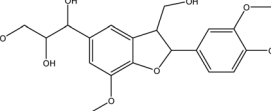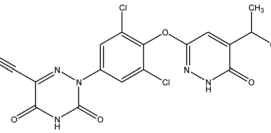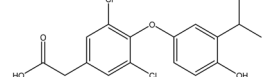
The SwissADME platform was accessed to evaluate the drug likeness criteria, pharmacokinetics, and toxicity profiles of the two identified hits and the reference agonists. This platform provides a wide predictive spectrum of various physico-chemical properties such as the logarithm of partition coefficient (log P) and water solubility (log S), total polar surface area (TPSA), number of H-bond acceptors and donors, number of rotatable bonds as well as the fraction of sp³ carbon atoms (C_{sp^3}). SwissADME also offers useful predictions about the absorption and bioavailability of drug candidates following oral administration, potential of metabolism by various cytochrome P450 enzymes, in addition to the medicinal chemistry parameters, and Brenk and pan-assay interference compounds (PAINS) alerts, which help predict the presence of unstable and/or toxic fragments within the chemical structure.

3. Results and discussion

3.1. Molecular docking

The objective of the molecular docking study was to rank the screened 47 199 natural compounds according to their binding affinities towards the THR-β receptor. This was achieved by gradually increasing the sensitivity of the molecular docking protocol over four rounds, where the number of compounds was reduced to 78. Among these, 22 top performing compounds were identified based on their docking scores and the similarity of their binding modes to those of the reference agonists,

Table 1 Docking scores (kcal mol⁻¹) afforded by the top ranking compounds and the reference agonists following induced fit docking towards the THR-β receptor

Compound number	ZINC15 ID	Compound structure	Docking score (kcal mol⁻¹)
2	100 038 257		−19.78
18	14 710 225		−16.7
23	Resmetirom		−12.95
24	Cognate ligand		−13.041

considering the afforded amino acid interactions. Table 1 lists the docking scores (kcal mol^{-1}) of the identified top ranking compounds; however, the data (docking scores and amino acid interactions) of all the 78 compounds are given in Tables S1 and S2 (the nomenclature and molecular shapes of the compound ZINC15 codes) in SI 2. Moreover, the 2D interaction maps of the docking poses retained by some of the top compounds are presented in Fig. 1. Docking results revealed that the primary interactions afforded by the reference agonists were due to the presence of some electron dense groups such as carboxylic acid and nitrile groups, which enabled the interactions with the amino acid residues of the target protein. The dominant H-bond and salt bridge interactions occurred via Arg316 and Arg320 in addition to various hydrophobic contacts within the

agonist binding site of the receptor. These two arginine residues harbour structural features responsible for the selectivity of the receptor.[54] These features provide conformational constraint in ensuring the stability of the agonist conformation in the receptor by substituting the amino acid side chain in the alcohol units present in resmetirom and the cognate ligand. THR-β is sensitive to the presence of a carbon chain, which enhances the polar interaction with arginine residues in the pocket and acts as a skeleton in which substituents lock in an active upright conformation relative to units such as the terminal phenyl ring.[55]

The docking score of the resmetirom compound is compatible with the cognate ligand's molecular structure in terms of the presence of common groups, and the presence of the
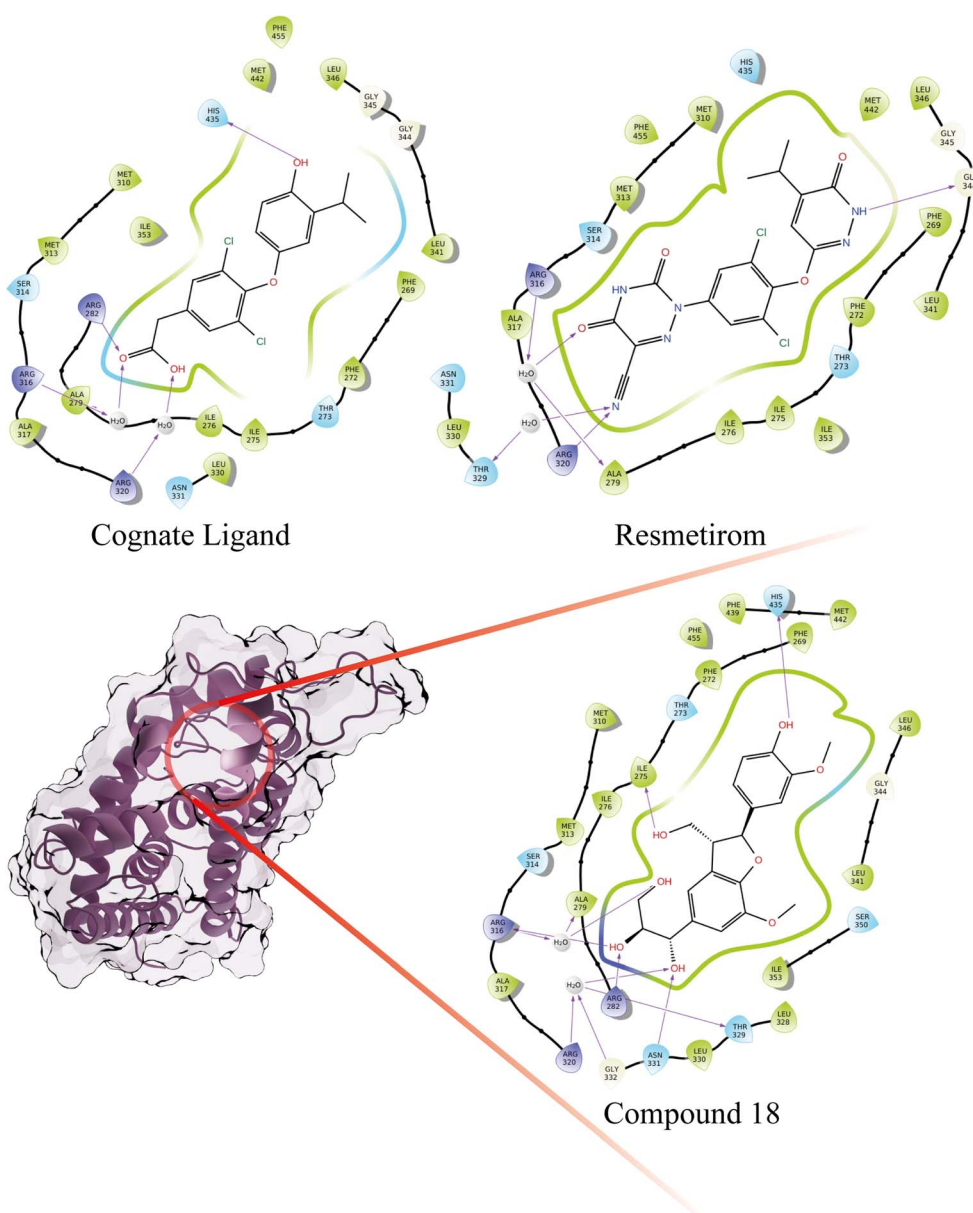

Cognate Ligand

Resmetirom

Compound 18

Fig. 2 2D binding poses of the compound and reference agonists that localised with the highest affinity following induced fit docking to the THR-β receptor.

halogen structure of resmetirom in the cognate ligand in the scores obtained clearly explains the close docking scores. The cause and effect explanation of this situation is clearly seen from the protein–ligand binding modes illustrated in Fig. 2.

The final docking round, which employed induced fit docking, revealed that the majority of the compounds that interacted with the arginine residues in the agonist binding site of the THR-β receptor were negatively charged compounds at physiological pH conditions. Compounds **1**, **7**, **11**, **14**, and **18** are of particular interest as they afforded similar binding modes to the reference agonists, which is presumably due to the presence of the negatively charged carboxylate moiety and other hydroxyl groups in the regions extending to residues Arg316 and Arg320. Nevertheless, compound **2**, which was previously reported to have a strong potential for treatment of fatty liver disease, exhibited the highest degree of similarity to the reference agonists. This compound afforded interactions with Gly332,

Thr329, Arg316, Arg320, Arg382 and Ala279 residues *via* water bridges and numerous direct hydrogen bonds, as illustrated in Fig. 1. Moreover, it was able to access the Arg316 and Arg320 residues by virtue of its carboxylate moiety and the nearby hydroxyl group while forming a hydrogen bond with Gly344 *via* the two hydroxyl groups at its other end.

Noteworthily, the chiral compounds **1** and **18** showed "*R*" configuration of the chiral centers having the hydroxyl groups that extended to the arginine residues, which represents a distinguishing feature in the emergence of strong interactions.

3.2. Molecular dynamics

3.2.1. Short-term molecular dynamics simulations. The top ranking 22 compounds that were identified according to the docking study were subjected to a short-term (10 ns) molecular dynamics study. The RMSD and RMSF plots of the protein–ligand complexes (Fig. 3 and 4) were analyzed to identify the
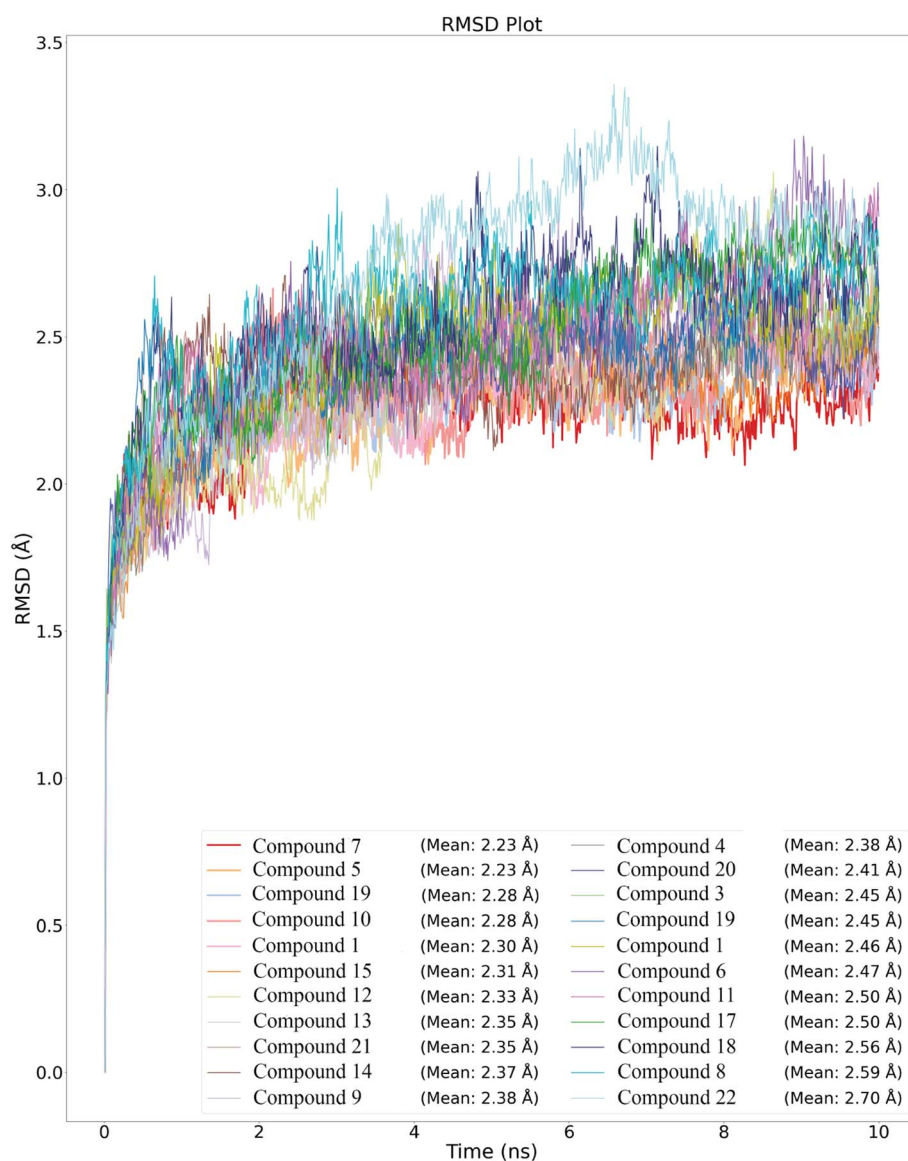

Fig. 3 RMSD data of the top ranking compounds following the short-term (10 ns) MD simulations.
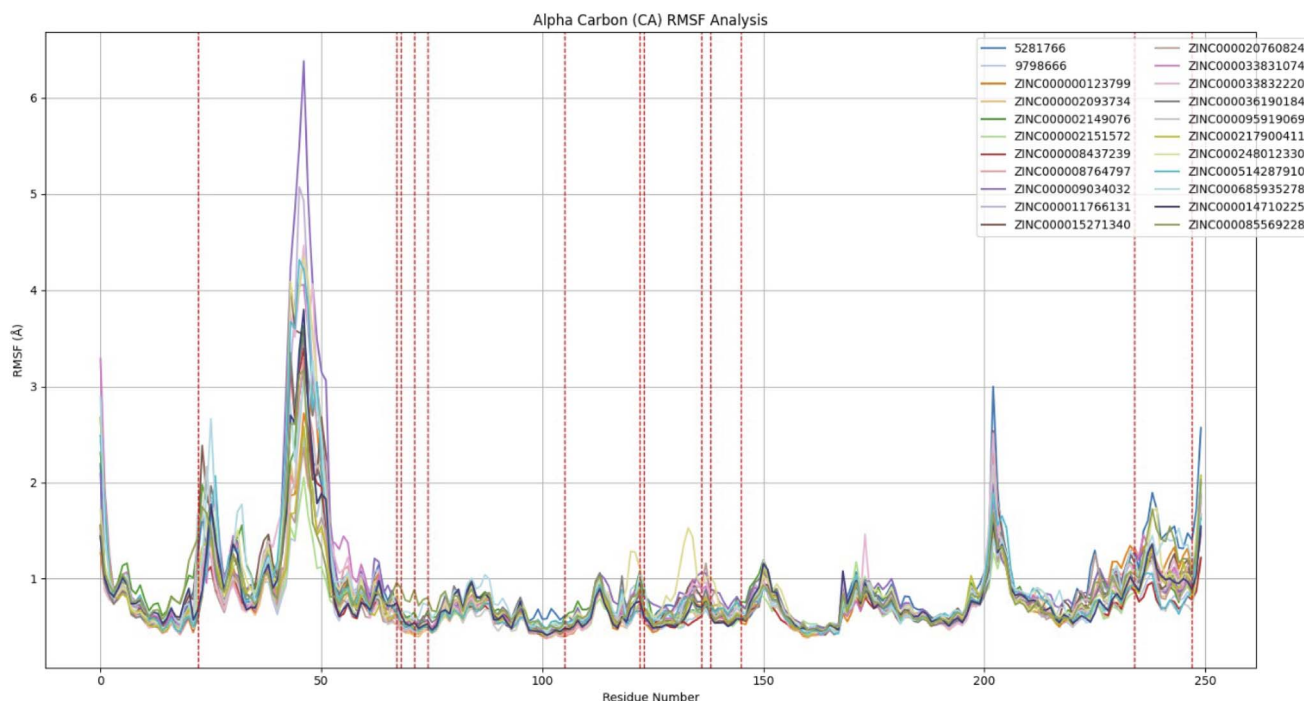
Fig. 4 RMSF data of the top ranking compounds following the short-term (10 ns) MD simulations. The red dashed lines correspond to the residues involved in the interactions with the reference agonists.

compounds showing strong binding affinity to the THR-β receptor. These plots serve as key indicators of the continuity, strength, and stability of the studied complexes over the simulation time.

In general, the studied compounds showed low RMSD values, indicating appropriate binding affinity. While the overall RMSD average of the compounds fluctuated in the range of 2–3 Å, some distinguished compounds showed relatively low RMSD values, compared to others, over the whole simulation time. These include compounds **1** (mean: 2.28 Å), **2** (mean: 2.28 Å), **5** (mean: 2.23 Å), **7** (mean: 2.23 Å), and **10** (mean: 2.28 Å). On the other hand, compound **22** showed the highest deviations in its RMSD values, suggesting relatively lower binding affinity to the receptor.

Examining the fluctuations in RMSF plots obtained following the binding of ligands to a certain protein provides valuable insights into the strength of their binding affinity towards that protein. Ligands showing low fluctuations with the amino acid residues in the active site are characterized by tight binding and strong affinity to the protein.[56] In this study, RMSF plots of the 10 ns MD simulations were thoroughly analyzed to identify the ligands that showed low fluctuations with the residues in the agonist site and high fluctuations in other regions of the protein structure. These ligands are expected to have strong agonist potential. As illustrated in Fig. 4, the red dashed lines correspond to the amino acid residues reported to interact with the reference agonists. The indexes of these residues are 67, 68, 71, 74, 105, 122, 123, 136, 138, 145, 227, 234, and 247. These indexes are located as residues 275, 276, 279, 282, 313, 330, 331, 344, 346, 353, 435, 442, and 455 in the protein structure, respectively.[31] Accordingly, ligands with

strong agonist potential should show minimal fluctuations with the amino acids in the region extending between indexes 65 and 150 and also in the region between indexes 200 and 250, while increased fluctuations should be observed elsewhere. Compounds **1, 2, 7, 11, 14**, and **18** showed this fluctuation pattern, in addition to their relatively low RMSD values, which indicate their strong THR-β agonist potential. Among these compounds, compound **2** is a caffeic acid derivative. Considering the favourable effects of caffeic acid reported in the literature in the treatment of non-alcoholic fatty liver disease, this result is quite consistent.[57] Another prominent compound, compound **18**, is a propanetriol (glycerol) derivative. The fact that there are reports in the literature that propanetriol derivatives also have a positive effect on non-alcoholic fatty liver disease shows that the information obtained from short-term molecular dynamics simulations is consistent.[58,59]

3.2.2. Long-term molecular dynamics simulations. Compounds **1, 2, 7, 11, 14**, and **18** that turned out to possess the most promising agonist potential towards THR-β in the short-term MD study were subjected to long-term (500 ns) MD simulations to verify their binding affinity to the target receptor. The cognate ligand and resmetirom (the reference agonists) were also employed in this study for comparison. The six compounds exhibited limited mobility, minimal displacement, and negligible conformational changes throughout the long-term simulations, which were consistent with their behavior in the short-term (10 ns) simulations. This could be attributed to the narrow binding site and the formation of salt bridges and hydrogen bonds with arginine residues in the agonist binding site.

Following the long-term MD simulations, compounds **2**, **11**, **14**, and **18** exhibited mean RMSD values in the range of 2.91–3.12 Å, which was comparable to 2.79 Å and 3.02 Å for the cognate ligand and resmetirom, respectively. Compounds **2** and **18** exhibited the closest resemblance to the RMSD graphs of the reference agonists in terms of both average and trend. However, compounds **1** and **7** exhibited a deviation greater than 3.5 Å in the RMSD data and persisted in this condition for a period extended approximately to 200 ns. This could indicate that the protein–ligand complexes of these two compounds were
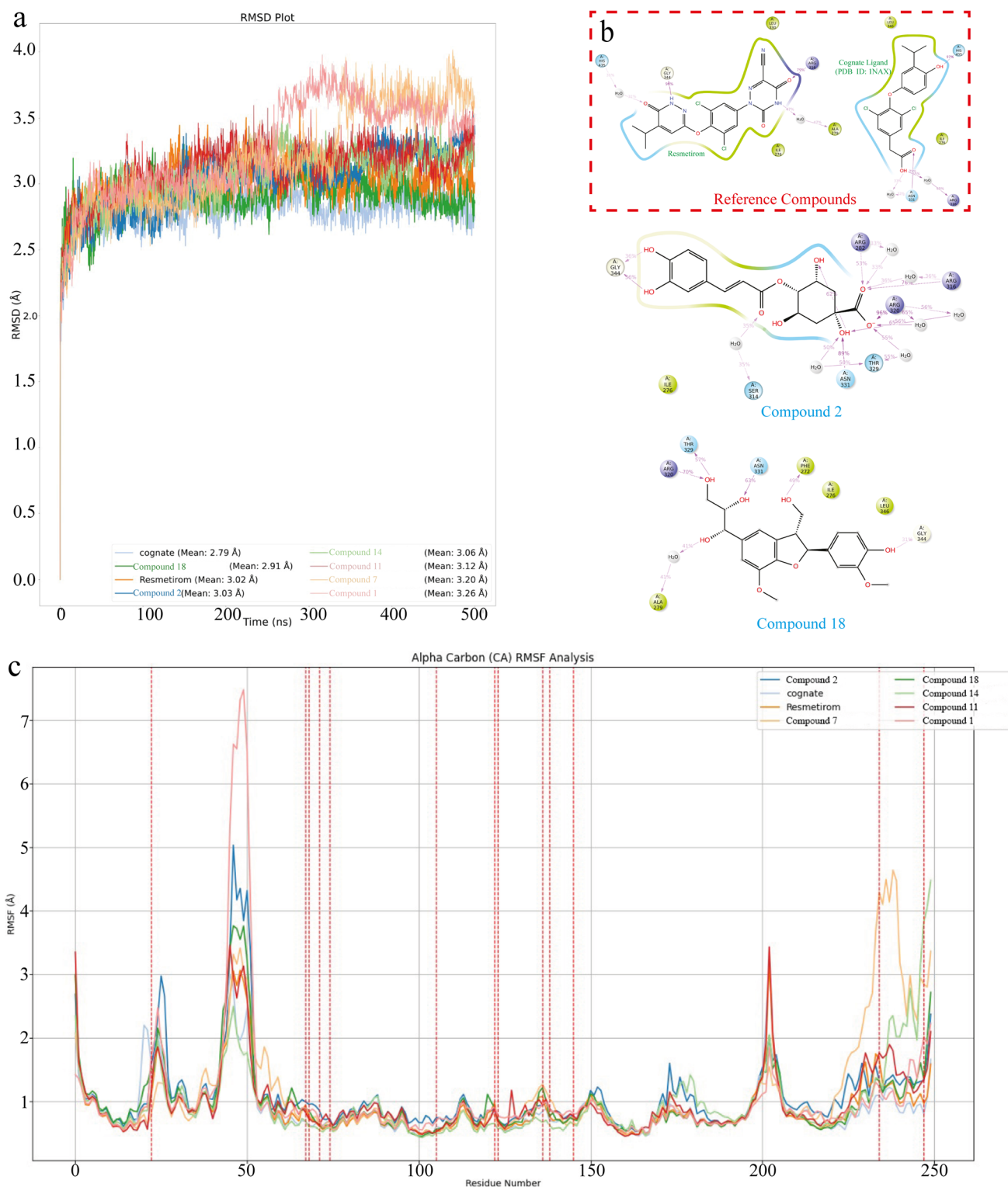

Fig. 5   RMSD data (a), interactions over 30% of the simulation time (b), and RMSF data (c) of the top ranking compounds and the reference agonists following the long-term (500 ns) MD simulations.

© 2025 The Author(s). Published by the Royal Society of Chemistry

unstable over the long term, as evident from their RMSD values, which reached up to approximately 4 Å towards the end of the simulations (Fig. 5a). This could also point out the relatively weak binding affinity of the two compounds and suggest that they were unlikely able to elicit an agonist signal upon binding to the receptor.

Analysis of the RMSF data obtained from the long-term MD simulations (Fig. 5c) revealed that compounds **2** and **18** exhibited distinctive fluctuations in the region far from the agonist binding site, around residue index 40 and highly reduced fluctuations in the regions of the agonist site, suggesting their high agonist potential to the receptor. Although the fluctuations observed by compound **1** were higher than those of other compounds in this range, this fluctuation pattern is sufficiently high to render the natural functioning of the protein structure uncertain. In view of the distinct RMSD and RMSF data sets of compound **1**, it might elicit specific effects, the nature of which is currently unclear. The fluctuations of compound **11** were significantly higher than the other compounds around the residue index 200. This information, when interpreted alone, indicates the appreciable agonist potential of the compound.

Compound **7** did not afford notable interactions with the crucial amino acid residues in the agonist site of the THR-β receptor in the region between indexes 200 and 250, as evidenced by the significant fluctuations it exhibited in this region. Although compound **14** had regular, satisfying, and plateauing RMSD data, it showed high fluctuation in the region of the agonist site residues, similar to the pattern observed with compound **7**.

Regarding the binding modes of the ligands employed during the long-term (500 ns) MD simulations, Fig. 5b illustrates the interactions between the six drug candidates and the amino acid residues in the agonist site of the target receptor, compared to the interactions afforded by the reference agonists over the simulation time.

In the agonist site, the reference agonists are engaged in crucial hydrogen bonding with Arg316 and His435 residues in addition to hydrophobic contacts with Ile276 and Leu330. Noteworthily, the nitrile group in resmetirom and the isopropyl group in the cognate ligand displayed a significant contribution to the drug–protein interactions.

Compound **2** formed both direct and indirect hydrogen bonds (through water bridges) with Arg316 that persisted for 76% and 36% of the simulation time, respectively, and also formed hydrogen bonding interaction with Gly344, similar to resmetirom. Compound **18** afforded hydrogen bonding interaction with Arg320 for 70% of the simulation time. However, no interactions with Arg316 were detected for more than 30% of the simulation time. Nevertheless, the compound showed hydrogen bonding interaction with Asn331, which was also afforded by the cognate ligand.

Compound **14** did not demonstrate any of the common interactions afforded by the reference agonists during the simulations. Compound **7** formed water bridges with Arg316, yet it did not exhibit direct interactions. In the same sense, the interactions afforded by compounds **1** and **11** were not as

effective, and their complexes with the protein did not show as high stability as that observed by other compounds such as **2** and **18** during the long-term MD simulations.

Conclusively, compounds **2** and **18** exhibited low RMSD values that were stable at the plateau levels, showed reduced fluctuations in the crucial regions of the receptor's agonist site with low RMSF values, and demonstrated similar binding modes observed by the reference agonists. Moreover, the ligand–protein complexes of the two compounds remained stable throughout the whole simulation time that indicated strong and robust adhesion to the residues in the agonist binding site. Moreover, compounds **2** and **18** were found to have a potential to enhance the signal carrier effect of the receptor by increasing the fluctuations in the regions away from the agonist binding site in a similar behavior to the reference agonists. Compound **2** has been reported as anti-diabetic, although its anti-inflammatory properties are predominantly mentioned in the literature.[60,61] Compound **18** is an anticancer compound with cytotoxic effects and is known to be effective in bone health and bone metabolism in very similar forms.[62,63]

3.3. Analyses of protein dynamics

Compounds **2** and **18** that showed the best performance in the long-term (500 ns) MD simulations had their impact on the dynamics of the THR-β receptor studied. The reference agonists were also included for comparison. Different analyses such as PCA, TICA-FES, and MSM were employed for this purpose. Consequently, the potential of these compounds to elicit effects analogous to those observed by the reference agonists could be evaluated with greater precision.

3.3.1. PCA. The PCA was applied using the data sets comprising the dynamic properties of the protein–ligand complexes of compounds **2** and **18** and the reference agonists to identify the most stable conformations and motions. The results, shown in Fig. 6, enabled the determination of the complexes' conformations and the states of energy barriers separating the conformational transitions.

The free energy surface plots obtained following the PCA of the ligand–protein complexes showed dark blue regions, representing low energy, surrounded by yellow areas that reflect relatively higher energy barriers. The energy landscapes of the analyzed complexes were remarkably similar. Specifically, compound **2** and resmetirom showed very comparable energy landscapes, and similarly compound **18** and the cognate ligand had very close energy surfaces. This is an evidence that the movement of the protein structure throughout the simulations was very similar for all systems. Noteworthily, the deep energy valleys appeared in the plots represent a pathway for conformational transitions, which is highly favorable for the generation of agonist signals.

3.3.2. TICA-FES analysis. TICA-FES provides useful information about the actual protein dynamics, the free energy landscape that governs these dynamics, and the signal transduction mechanism of the protein across disparate time scales. Moreover, TICA-FES is an excellent tool to construct protein dynamics Markov models.[23] In this study, TICA was conducted using lag values of 10, 100, and 500.
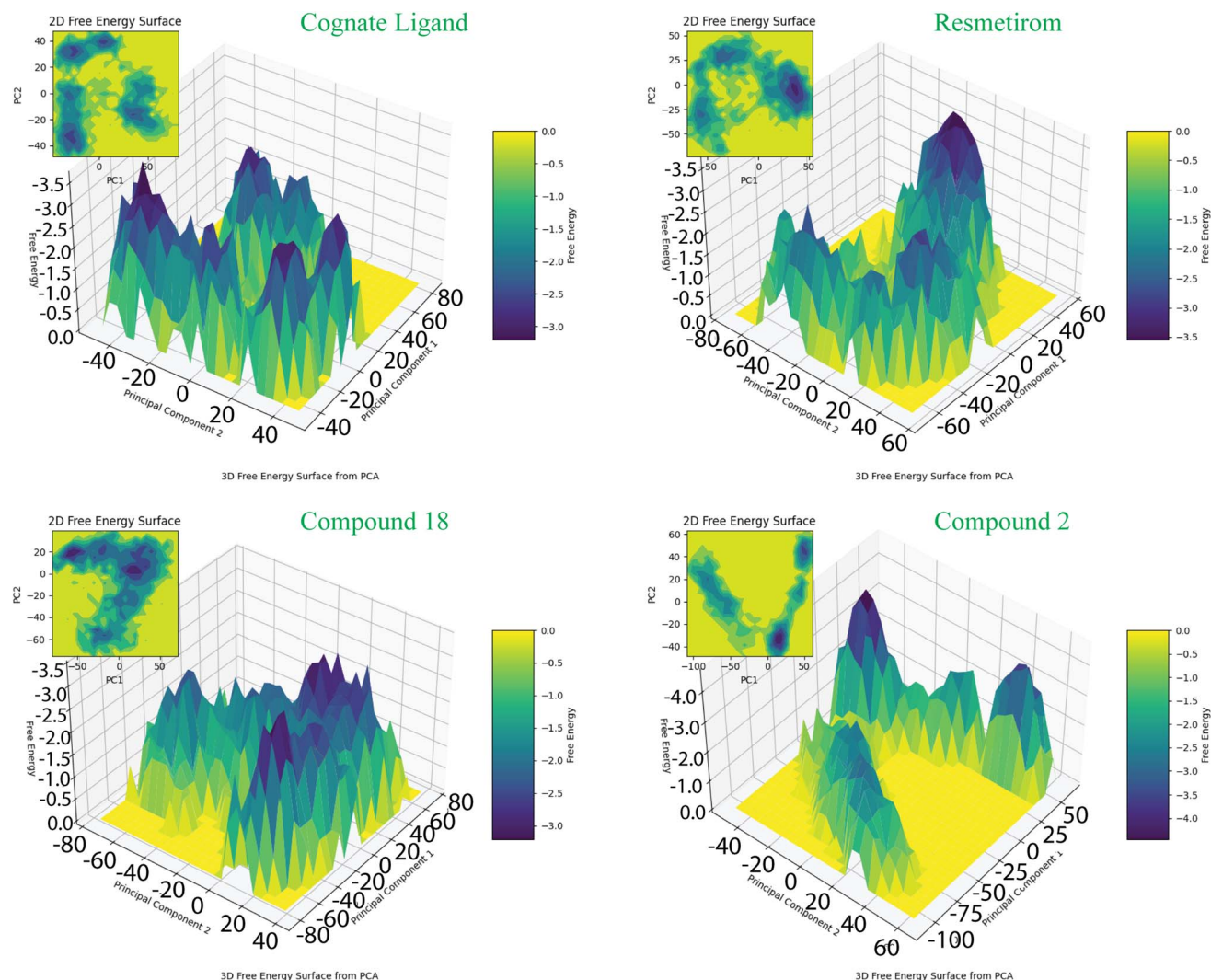
Fig. 6 PCA plots obtained upon analyzing the dynamic properties of the THR-β−ligand complexes of compounds **2** and **18** and the reference agonists. Yellow areas represent high energy, while dark blue areas represent low energy.

At a lag value of 10, resmetirom and the cognate ligand exhibited remarkably similar deep low energy valleys, whereas compounds **2** and **18** demonstrated a high degree of similarity to the energy landscape of the reference compounds. In all the systems, the protein was able to transition between different conformations by concentrating the deep energy valleys in specific regions. This points out that the studied compounds help stabilize certain conformations of the protein, making these states more likely to occur. However, when the lag value was set to 100, the energy profile similarity between resmetirom and the cognate ligand was undetectable, while the energy landscapes of the cognate ligand and the two hit compounds were still demonstrating a high degree of similarity. This observation underscores the potential for a pronounced agonist effect at this particular lag value. At a lag value of 500, the deep energy valleys exhibited highly analogous energy landscapes for all compounds, confirming similar agonist effects of the hit compounds to those of resmetirom and the cognate ligand over an extended timeframe (Fig. 7). This analysis at varying lag

values provides substantial evidence that the parallel agonist properties of both the hit compounds and the reference agonists towards THR-β are effective, particularly during the time interval when the lag value is 100, *i.e.*, around 50 ns.

3.3.3. MSM analysis. The conformational transition states of the lead compounds were analyzed by molecular dynamics simulations to construct the lag time–implied timescale (MSM) plots. Analyzing the transition states with different clustering numbers, compound **2** exhibited a high degree of fluctuation in the range of 90–100 ns. The lowest observed fluctuation was identified as clustering number 3; however, even within this specific clustering category, the inconsistency between 90 and 100 ns could not be eliminated. The behavior demonstrated by compound **2** within the 90–100 ns range might suggest the occurrence of rare events or the rapid development of substantial conformational changes over brief temporal intervals. Consequently, movements between 90 and 100 ns were excluded for all complexes. As shown in the MSM plot, the reference agonists possessed a high degree of conformational stability and reflected constant agonist potency over an
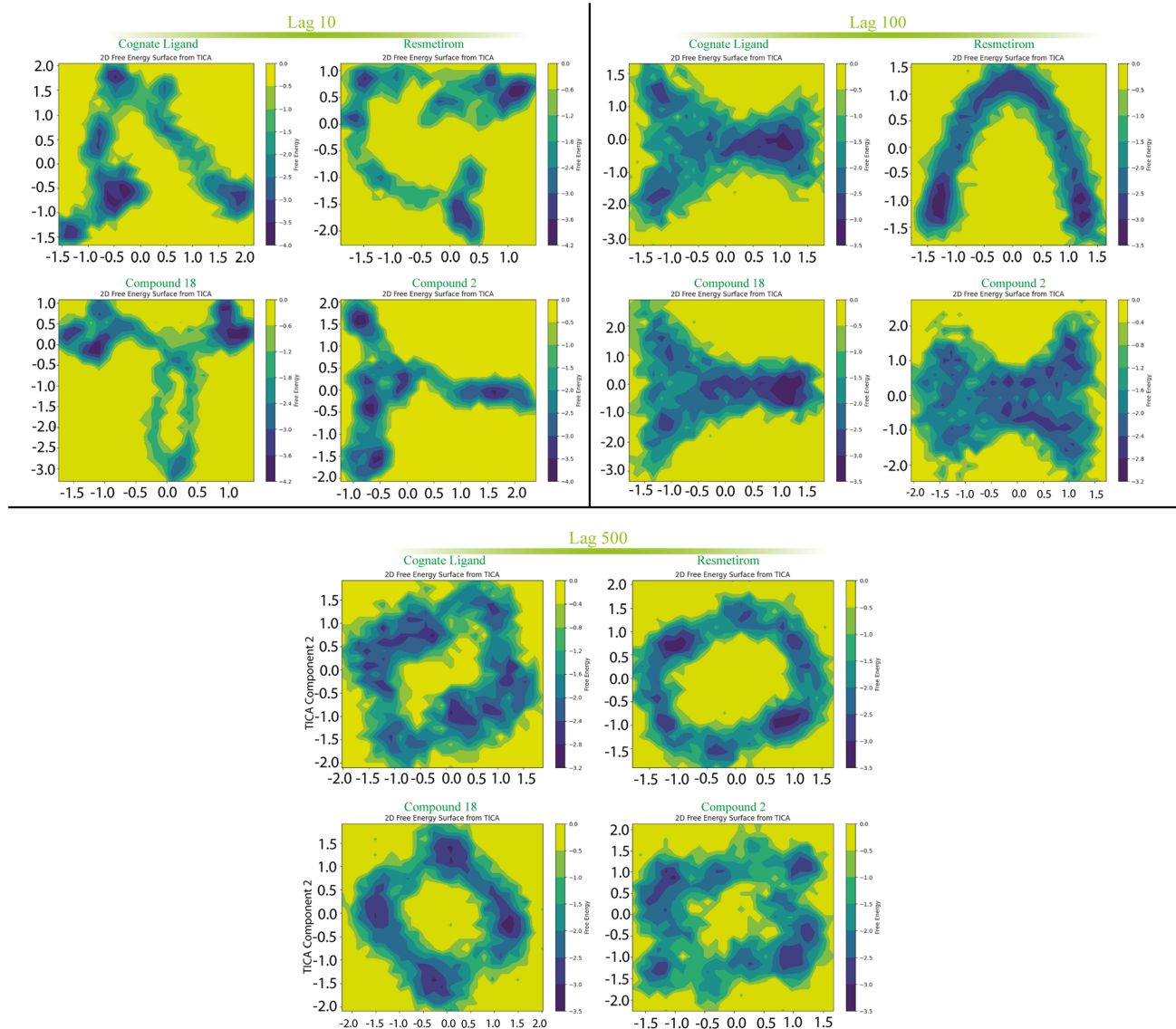
Fig. 7 TICA–FES analysis results of MD simulations at 10, 100 and 500 lag values.

extended time scale (approximately 70–100 ns), as indicated by the implied timescale data. Compound **2** exhibited a similar pattern to the reference compounds, indicating stable conformational states and high agonist potential by showing planar graphs. In contrast, compound **18** displayed a partial potential. It stabilized the receptor within a more rapid timescale. Further optimization may be necessary to achieve full agonism. Information on this can be seen in Fig. S4.

3.4. Molecular mechanics with generalized born surface area (MM/GBSA)

The free energy calculation utilising MM-GBSA (Fig. 8) was employed to evaluate the stability of the ligand–THR-β complexes formed by compounds **2** and **18** and the reference agonists based on the free energy values and the complexation slope throughout the long-term (500 ns) MD simulations.

The free energy data demonstrated that all the complexes exhibited low energy states, thereby confirming their stability at

all stages of the long-term MD simulations. The complex of compound **2** exhibited a slightly higher energy state than the others, reaching a maximum value of -65 kcal mol^{-1}. Subsequently, the free energy demonstrated a downward trend, declining below the threshold of -80 kcal mol^{-1}. Noteworthily, the other three complexes manifested comparable and proximate free energy values.

3.5. Evaluating the selectivity of the identified hits towards the THR-β receptor

The two hit compounds **2** and **18** as well as compounds **11** and **14** that demonstrated appreciable THR-β agonist properties in the MD simulations were selected to evaluate their selectivity towards the targeted beta isoform of the receptor over the alpha one. This was done by performing molecular docking and MD simulations on the THR-α receptor isoform, adopting the same protocols used previously throughout the study.
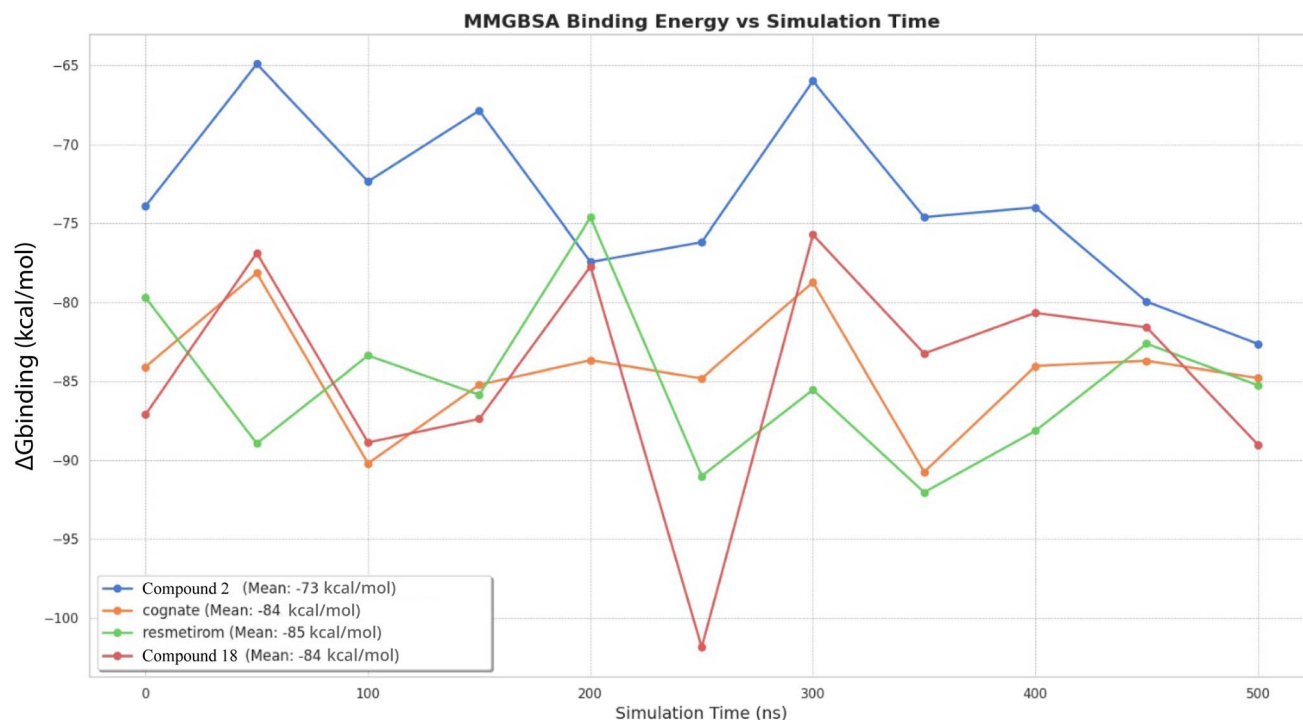
Fig. 8 MM/GBSA plots showing the change in free energies of the ligand–protein complexes of the reference agonists and compounds **2** and **18** during the long-term (500 ns) MD simulations.

Docking results showed that only compounds **2**, **11**, and **14** were able to dock to THR-α. However, they exhibited minimal interactions with the amino acid residues in the agonist site of the receptor (Fig. S5). Interestingly, compound **2** utilized water molecules as a bridge for the majority of its hydrogen bonding interactions with the protein. While water molecules contribute to the interaction energies, they must be meticulously examined in the context of static interactions due to their inherent mobility in dynamic processes. Compound **18** was not able to dock to THR-α, indicating high selectivity potential towards the β isoform of the receptor. The high degree of selectivity observed is attributable to the presence of the Ser277 residue in the α isoform at the ligand binding sites, despite the absence of the Asn331 residue in the β isoform. This amino acid residue difference has resulted in alterations to the hydrogen bond interaction pattern, thereby creating conditions conducive to selectivity. Furthermore, α-helices, free ends and helical structure differences, as demonstrated in Fig. 1, significantly distinguish the protein dynamics between the two structures. These helical structure differences are primarily observed in regions near the backbone residues within the binding site in the α isoform, where folding and length variations are evident.

RMSD and RMSF data of compound **11** showed minimal deviation values, suggesting robust stability of the ligand–protein complex and hence an appreciable affinity towards THR-α. Noteworthily, the RMSF plot of compound **11** displayed a notable indication of a partial agonist effect. This was manifested as an additional fluctuation in the highly fluctuating region around residue index 50. The RMSD and RMSF data of compounds **2** and **14** exhibited elevated deviation values,

suggesting very low affinity and a minimal impact on the protein structure (Fig. S5).

The PCA data of compound **11** demonstrated the greatest similarity to the cognate ligand's effects on the protein dynamics. The free energy surface of this compound contained two distinct low-energy regions, analogous to those observed in the cognate ligand's surface. Compound **2**, on the other hand, exhibited inconsistencies in its probability density representation of the agonist effect, with low-energy regions distributed across a wide range. Compound **14** appeared to restrict protein dynamics, exhibiting a single low-energy region (Fig. S5).

When all the data are considered collectively, it becomes evident that compounds **2**, **14**, and **18** evinced minimal interest in the alpha isoform. Compound **11**, however, partly demonstrated compatible and consistent protein dynamics data with those of the cognate ligand, which is known to be an agonist. The RMSD and RMSF data of this compound could suggest a significant partial agonistic potential. Compound **14** exerted a constrained effect on the protein dynamics and thus could have the potential to demonstrate a direct antagonist effect towards the THR-α isoform. Compound **2** did not exhibit any evidence of agonistic effect and displayed markedly distinct protein dynamics. Compound **18**, which did not dock to THR-α, could have a significant preference to the beta isoform with high selectivity and considerable agonistic potential.

3.6. Drug likeness, pharmacokinetics (ADME), and toxicity profile analyses

We finally utilized the SwissADME online platform to evaluate the drug likeness potential, determine the ADME properties,

and investigate the toxicity profiles of the top performing compounds **2** and **18** along with the reference agonists (Table 2).

Lipinski's rule of five states that compounds with molecular weights less than 500, partition coefficients ($\log P$) below 5, fewer than 5 hydrogen bond donors, and fewer than 10 hydrogen bond acceptors are more likely to show good oral bioavailability.[64,65] Veber's rule considers two additional criteria, which are a total polar surface area (TPSA) of less than 140 Å^2 and less than 10 rotatable bonds.[66] The fraction of the sp^3-hybridized carbon atoms relative to the total count of carbon (C_{sp^3}) is another component that influences drug likeness. It reflects on the complexity of molecules and indicates their carbon saturation. Studies showed that compounds with zero C_{sp^3} tend to have low absorption and bioavailability.[67] In the same manner, water solubility ($\log S$) significantly influences drug absorption across biological membranes. Drug candidates with $\log S$ values of -4 or higher are expected to have good oral bioavailability[68]

Compound **2** showed one violation to each of the Lipinski and Veber criteria, low absorption, and exhibited PAINS and Brenk alerts by virtue of its catechol and α,β-unsaturated carbonyl moieties, respectively. Indeed, drug candidates with catechol moieties demonstrate PAINS alert as they often nonspecifically bind enzymes and receptors, either *via* metal chelation or covalent bonding, other than the target one leading to false positive assay results. Moreover, catechols are possibly oxidized to quinones, generating reactive oxygen species that lead to oxidative stress and cause damage to DNA, lipids, and proteins. For instance, nordihydroguaiaretic acid that was used to manage asthma conditions was recalled due to liver and kidney toxicity linked to its catechol moiety. However, the catechol derivatives dopamine and epinephrine are safely used therapeutically, in a wide range, since they are very well managed by the anti-oxidant defense mechanisms of the body as long as the dosage limits and the route of administration (IV, injection) are properly considered.[69]

The α,β-unsaturated carbonyls are electrophilic groups that react with nucleophiles, such as the SH of cysteine residues in proteins, *via* the Michael addition reaction. This could impair protein functions and trigger geno- and cytotoxicity at high levels. The reported neurotoxicity and oxidative damage of acrolein and crotonaldehyde are mainly linked to their α,β-unsaturated carbonyl groups. That said, the presence of this group is still beneficial for the action of some drugs, such as dimethyl fumarate that is used in psoriasis and multiple sclerosis. The drug utilizes its α,β-unsaturated carbonyl to react with the cysteine residues of cellular proteins, forming adducts that influence signaling pathways involved in inflammation and oxidative stress such as the Nrf2 pathway.[70] Moreover, curcumin is a natural α,β-unsaturated chalcone that is used safely as a potent antioxidant and anti-inflammatory agent[71]

In contrast, compound **18** and the cognate ligand did not show any violation to either Lipinski or Veber drug likeness criteria, showed good absorption with no predicted interactions with the cytochromes or toxicity alerts. However, resmetirom

Table 2 Drug likeness, ADME, and toxicity profile analysis of the two top ranking compounds and the reference agonists

Compound number	Molecular weight (g mol^{-1})	$\log P$	H-bond acceptors	H-bond donors	TPSA (Å^2)	Rotatable bonds	C_{sp^3}	$\log S$	GI absorption	BBB permeant	Cytochrome substrate	Toxicity alert
2	354.31	−0.37	9	6	164.7	5	0.38	−1.62	Low	No	No	PAINS alert Brenk alert
18	392.40	0.84	8	5	128.8	7	0.4	−2.13	High	No	No	No alerts
Cognate ligand	355.21	2.82	4	2	66.8	5	0.24	−5.21	High	No	CYP1A2, CYP2C9, and CYP2C19	No alerts
Resmetirom	435.2	2.22	7	2	146.5	4	0.18	−4.42	Low	No	No	No alerts

showed a similar profile with low absorption and one Veber violation (Table 2).

Our comprehensive *in silico* study has pointed out that compounds **2** and **18** are promising hits to be considered as potential candidates for NAFLD management. Interestingly, 4-*O*-caffeoylquinic acid (compound **2**) has been reported to display a wide array of pharmacological activities and strong therapeutic potential against several ailments. It was able to improve memory and cognition functions and prevent neuronal loss in mouse models of Alzheimer's disease, most likely due to its support of mitochondrial health alongside its powerful antioxidant activity.[64] This compound also demonstrated potent anti-diabetic activity *in vitro* by inhibiting the enzyme α-glucosidase, thereby altering carbohydrate absorption and by inhibiting protein tyrosine phosphatase 1B, a negative regulator in insulin signalling.[72,73] In the HepG2 lipid-accumulation model, 4-*O*-caffeoylquinic acid lowered the triglyceride levels effectively, surpassing simvastatin *in vitro*, which suggested a promising potential in treating dyslipidemia and fatty liver[74]

The lignan derivative, dihydroxydehydrodiconiferyl alcohol (compound **18**), on the other hand, demonstrated a therapeutic value in bone disorders such as osteoporosis as it showed agonistic effect on estrogen receptors, promoting osteoblast differentiation. It also exhibited strong anti-inflammatory activity through inactivating the NF-Kb signaling pathway and altered adipogenesis and lipid accumulation in some *in vitro* studies, indicating its promising potential against obesity and dyslipidemia related diseases.[75,76]

4. Conclusion

Using a rigorous, multi-tiered *in silico* workflow—encompassing high-retention molecular docking of 47 199 ZINC15 natural products, 500 ns all-atom molecular dynamics simulations, MM/GBSA free-energy estimates, and advanced protein-dynamics analyses (PCA, TICA-FES, and MSM)—we identified 4-*O*-caffeoylquinic acid (compound **2**) and dihydroxydehydrodiconiferyl alcohol (compound **18**) as compelling thyroid hormone receptor-β (THR-β) agonist candidates for the management of non-alcoholic fatty liver disease (NAFLD). Both ligands established persistent electrostatic and hydrogen-bonding interactions with the key residues Arg316 and Arg320, with their carboxylate and hydroxyl moieties complementing the positively charged guanidinium groups and thereby reinforcing an active-like receptor conformation. Consistent with these contacts, MD trajectories reveal microstate transition rates and collective motions that closely mirror those induced by the reference agonist resmetirom, indicating a comparable capacity to trigger THR-β signal transduction. *In silico* ADMET profiling further predicts favorable oral bioavailability, metabolic stability, and safety margins for both compounds. Collectively, these findings position compounds **2** and **18** as tractable chemical leads for THR-β-targeted NAFLD therapy and underscore the power of integrative computational pipelines to accelerate the discovery of mechanism-informed therapeutics. Subsequent biochemical and *in vivo* studies are now warranted to validate their efficacy and safety profiles.

Conflicts of interest

There are no conflicts to declare.

Data availability

All data, code, and materials needed to evaluate, reproduce, and extend the findings of this study are openly available as detailed below.

Custom code & workflow: the scripts and workflow specifications used for database mining, structure preparation, docking/MD analyses, post-processing, and figure generation are archived on Zenodo and mirrored on GitHub. Please cite the exact archived version used in this article and refer readers to the most up-to-date version *via* the Zenodo record:

● Version used in this article (archived, immutable): **https://doi.org/10.5281/zenodo.17194142** (MDScripts Molecular Dynamics Simulation Analysis Package (MDSAP), v1.0; license: MIT).

● Most recent version: the latest citable version is accessible *via* the "Versions" panel of the same Zenodo record; the live development repository is available at **https://github.com/cannabinoid13/MDScripts**.

Availability during peer review and at publication: referees had access to the above code and data during peer review. The Zenodo record listed here is public and provides a persistent identifier for the archived version used; readers can retrieve any newer versions *via* the Zenodo "Versions" view.

Natural compounds from the ZINC15 database are publicly accessible at **http://zinc15.docking.org**. Any additional data related to this article are available from the corresponding author upon reasonable request.

All relevant data used in this study are provided in the supplementary information (SI). Supplementary information: details of molecular modelling, docking, pharmacophore, and ADME/toxicity analyses, including raw and processed data. See DOI: **https://doi.org/10.1039/d5dd00146c**.

References

1 N. A. ElSayed, G. Aleppo, V. R. Aroda, R. R. Bannuru, F. M. Brown, D. Bruemmer, B. S. Collins, K. Cusi, M. E. Hilliard, D. Isaacs, E. L. Johnson, S. Kahan, K. Khunti, J. Leon, S. K. Lyons, M. L. Perry, P. Prahalad, R. E. Pratley, J. J. Seley, R. C. Stanton, Z. Younossi and R. A. Gabbay, 4. Comprehensive Medical Evaluation and Assessment of Comorbidities: Standards of Care in Diabetes—2023, *Diabetes Care*, 2023, **46**, s49–s672.

2 S. Pouwels, N. Sakran, Y. Graham, A. Leal, T. Pintar, W. Yang, R. Kassir, R. Singhal, K. Mahawar and D. Ramnarain, Non-alcoholic fatty liver disease (NAFLD): a review of pathophysiology, clinical management and effects of weight loss, *BMC Endocr. Disord.*, 2022, **22**, 63.

3 R. E. Powell, F. Zaccardi, C. Beebe, X. M. Chen, A. Crawford, J. Cuddeback, R. A. Gabbay, L. Kissela, M. L. Litchman, R. Mehta, L. Meneghini, K. M. Pantalone, S. Rajpathak, P. Scribner, J. W. Skelley and K. Khunti, Strategies for

overcoming therapeutic inertia in type 2 diabetes: A systematic review and meta-analysis, *Diabetes, Obes. Metab.*, 2021, **23**, 2137–2154.

4 W. Guo, Y. Song, Y. Sun, H. Du, Y. Cai, Q. You, H. Fu and L. Shao, Systemic immune-inflammation index is associated with diabetic kidney disease in Type 2 diabetes mellitus patients: Evidence from NHANES 2011-2018, *Front. Endocrinol.*, 2022, **13**, 1071465.

5 J.-K. Long, W. Dai, Y.-W. Zheng and S.-P. Zhao, miR-122 promotes hepatic lipogenesis via inhibiting the LKB1/AMPK pathway by targeting Sirt1 in non-alcoholic fatty liver disease, *Mol Med*, 2019, **25**, 26.

6 Y. Zhou, Y.-L. Ding, J.-L. Zhang, P. Zhang, J.-Q. Wang and Z.-H. Li, Alpinetin improved high fat diet-induced non-alcoholic fatty liver disease (NAFLD) through improving oxidative stress, inflammatory response and lipid metabolism, *Biomed. Pharmacother.*, 2018, **97**, 1397–1408.

7 T. Yan, N. Yan, P. Wang, Y. Xia, H. Hao, G. Wang and F. J. Gonzalez, Herbal drug discovery for the treatment of nonalcoholic fatty liver disease, *Acta Pharm. Sin. B*, 2020, **10**, 3–18.

8 P. J. Davis, J. L. Leonard and F. B. Davis, Mechanisms of nongenomic actions of thyroid hormone, *Front. Neuroendocrinol.*, 2008, **29**, 211–218.

9 F. Flamant, S.-Y. Cheng, A. N. Hollenberg, L. C. Moeller, J. Samarut, F. E. Wondisford, P. M. Yen and S. Refetoff, Thyroid Hormone Signaling Pathways: Time for a More Precise Nomenclature, *Endocrinology*, 2017, **158**, 2052–2057.

10 M. O. Ribeiro, Effects of thyroid hormone analogs on lipid metabolism and thermogenesis, *Thyroid*, 2008, **18**, 197–203.

11 T. Jakobsson, L.-L. Vedin and P. Parini, Potential Role of Thyroid Receptor β Agonists in the Treatment of Hyperlipidemia, *Drugs*, 2017, **77**, 1613–1621.

12 W. Zhou, S. Rahimnejad, K. Lu, L. Wang and W. Liu, Effects of berberine on growth, liver histology, and expression of lipid-related genes in blunt snout bream (Megalobrama amblycephala) fed high-fat diets, *Fish Physiol. Biochem.*, 2019, **45**, 83–91.

13 S. J. Keam, Resmetirom: First Approval, *Drugs*, 2024, **84**, 729–735.

14 Z. Li, Q. Cao, H. Chen, J. Yang, Z. Wang, X. Qu, Y. Yao, Z. Zhou and W. Zhang, Dual Phytochemical/Activity-Guided Optimal Preparation and Bioactive Material Basis of Orthosiphon Stamineus Benth. (Shen Tea) against Nonalcoholic Fatty Liver Disease, *J. Agric. Food Chem.*, 2024, **72**, 18561–18572.

15 S. Shen, K. Wang, Y. Zhi, W. Shen and L. Huang, Gypenosides improves nonalcoholic fatty liver disease induced by high-fat diet induced through regulating LPS/TLR4 signaling pathway, *Cell Cycle*, 2020, **19**, 3042–3053.

16 F. J. García-Alonso, R. González-Barrio, G. Martín-Pozuelo, N. Hidalgo, I. Navarro-González, D. Masuero, E. Soini, U. Vrhovsek and M. J. Periago, A study of the prebiotic-like effects of tomato juice consumption in rats with diet-induced non-alcoholic fatty liver disease (NAFLD), *Food Funct.*, 2017, **8**, 3542–3552.

17 J.-H. Yan, B.-J. Guan, H.-Y. Gao and X.-E. Peng, Omega-3 polyunsaturated fatty acid supplementation and non-alcoholic fatty liver disease: A meta-analysis of randomized controlled trials, *Medicine*, 2018, **97**, e12271.

18 H. Dib, M. Abu-Samha, K. Younes and M. A. O. Abdelfattah, Evaluating the Physicochemical Properties–Activity Relationship and Discovering New 1,2-Dihydropyridine Derivatives as Promising Inhibitors for PIM1-Kinase: Evidence from Principal Component Analysis, Molecular Docking, and Molecular Dynamics Studies, *Pharmaceuticals*, 2024, **17**, 880.

19 S. Sel, T. Tunç, A. B. Ortaakarsu, S. Mamaş, N. Karacan and M. S. Karacan, Acetohydroxyacid Synthase (AHAS) Inhibitor-Based Commercial Sulfonylurea Herbicides as Glutathione Reductase Inhibitors: in Vitro and in Silico Studies, *ChemistrySelect*, 2022, **7**, e202202235.

20 A. B. Ortaakarsu and H. Medetaliİeyoğlu, Computational drug repurposing effort for identifying novel hits for the treatment of diseases such as endometriosis, uterine fibroids, and prostate cancer, *Turk. J. Chem.*, 2024, **48**, 402–421.

21 A. Daina, O. Michielin and V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, *Sci. Rep.*, 2017, **7**, 42717.

22 A. Kitao, Principal Component Analysis and Related Methods for Investigating the Dynamics of Biological Macromolecules, *J*, 2022, **5**, 298–317.

23 S. Schultze and H. Grubmüller, Time-Lagged Independent Component Analysis of Random Walks and Protein Dynamics, *J. Chem. Theory Comput.*, 2021, **17**, 5766–5776.

24 X. Yang, Y. Gao, F. Cao and S. Wang, Molecular Dynamics Simulations Combined with Markov Model to Explore the Effect of Allosteric Inhibitor Binding on Bromodomain-Containing Protein 4, *Int. J. Mol. Sci.*, 2023, **24**, 10831.

25 *Schrödinger Release 2023-4*, Maestro, Schrödinger, LLC, New York, NY, 2023.

26 T. Sterling and J. J. Irwin, ZINC 15 – Ligand Discovery for Everyone, *J. Chem. Inf. Model.*, 2015, **55**, 2324–2337.

27 S. Petta, G. Targher, S. Romeo, U. B. Pajvani, M.-H. Zheng, A. Aghemo and L. V. C. Valenti, The first MASH drug therapy on the horizon: Current perspectives of resmetirom, *Liver Int.*, 2024, **44**, 1526–1536.

28 *Schrödinger Release 2023-4*, Epik, Schrödinger, LLC, New York, NY, 2023.

29 *Schrödinger Release 2023-4*, LigPrep, Schrödinger, LLC, New York, NY, 2023.

30 C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner and E. D. Harder, OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space, *J. Chem. Theory Comput.*, 2021, **17**, 4291–4300.

31 L. Ye, Y.-L. Li, K. Mellström, C. Mellin, L.-G. Bladh, K. Koehler, N. Garg, A. M. Garcia Collazo, C. Litten, B. Husman, K. Persson, J. Ljunggren, G. Grover, P. G. Sleph, R. George and J. Malm, Thyroid Receptor

Ligands. 1. Agonist Ligands Selective for the Thyroid Receptor β1, *J. Med. Chem.*, 2003, **46**, 1580–1588.

32 A. Zamanos, G. Ioannakis and I. Z. Emiris, HydraProt: A New Deep Learning Tool for Fast and Accurate Prediction of Water Molecule Positions for Protein Structures, *J. Chem. Inf. Model.*, 2024, **64**, 2594–2611.

33 *Schrödinger Release 2023-4: Protein Preparation Wizard*, Epik, Schrödinger, LLC, New York, NY, 2023; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2023.

34 *Schrödinger Release 2023-4: Glide*, Schrödinger, LLC, New York, NY, 2023.

35 *Schrödinger Release 2023-4: Induced Fit Docking protocol*, Glide, Schrödinger, LLC, New York, NY, 2023, Prime, Schrödinger, LLC, New York, NY, 2023.

36 M. K. Ucuncu, M. Y. Ucuncu, N. Topcuoglu, E. Kitin, O. Yazicioglu, A. B. Ortaakarsu, M. Aydın and A. Erol, The impact of a-tomatine on shear bonding strength in different dentin types and on cariogenic microorganisms: an in vitro and in silico study, *BMC Oral Health*, 2024, **24**, 1220.

37 D. J. Price and C. L. Brooks, A modified TIP3P water potential for simulation with Ewald summation, *J. Chem. Phys.*, 2004, **121**, 10096–10103.

38 The Nose–Hoover thermostat | The Journal of Chemical Physics | AIP Publishing, **https://pubs.aip.org/aip/jcp/article-abstract/83/8/4069/219065/The-Nose-Hoover-thermostatThe-Nose-Hoover?redirectedFrom=fulltext**, (accessed 1 January 2024).

39 A. B. Ortaakarsu, Ö. B. Boğa and E. B. Kurbanoğlu, Cardaria draba subspecies Shalepensis exerts in vitro and in silico inhibition of α-glucosidase, TRP1, and DLD-1 proliferation, *Sci. Rep.*, 2025, **15**, 10402.

40 G. J. Martyna, D. J. Tobias and M. L. Klein, Constant pressure molecular dynamics algorithms, *J. Chem. Phys.*, 1994, **101**, 4177–4189.

41 F. Sittel, A. Jain and G. Stock, Principal component analysis of molecular dynamics: On the use of Cartesian vs. internal coordinates, *J. Chem. Phys.*, 2014, **141**, 014111.

42 G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis and F. Noé, Identification of slow molecular order parameters for Markov model construction, *J. Chem. Phys.*, 2013, **139**, 015102.

43 L. Molgedey and H. G. Schuster, Separation of a mixture of independent signals using time delayed correlations, *Phys. Rev. Lett.*, 1994, **72**, 3634–3637.

44 Y. Li and Z. Dong, Effect of Clustering Algorithm on Establishing Markov State Model for Molecular Dynamics Simulations, *J. Chem. Inf. Model.*, 2016, **56**, 1205–1215.

45 F. Qin, Restoration of Single-Channel Currents Using the Segmental k-Means Method Based on Hidden Markov Modeling, *Biophys. J.*, 2004, **86**, 1488–1501.

46 J. D. Chodera and F. Noé, Markov state models of biomolecular conformational dynamics, *Curr. Opin. Struct. Biol.*, 2014, **25**, 135–144.

47 C. R. Harris, K. J. Millman, S. J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. Van Kerkwijk, M. Brett, A. Haldane, J. F. Del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke and T. E. Oliphant, Array programming with NumPy, *Nature*, 2020, **585**, 357–362.

48 N. Michaud-Agrawal, E. J. Denning, T. B. Woolf and O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations, *J. Comput. Chem.*, 2011, **32**, 2319–2327.

49 M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz and F. Noé, PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models, *J. Chem. Theory Comput.*, 2015, **11**, 5525–5542.

50 J. D. Hunter, Matplotlib: A 2D Graphics Environment, *Comput. Sci. Eng.*, 2007, **9**, 90–95.

51 M. Hoffmann, M. Scherer, T. Hempel, A. Mardt, B. de Silva, B. E. Husic, S. Klus, H. Wu, N. Kutz, S. L. Brunton and F. Noé, Deeptime: a Python library for machine learning dynamical models from time series data, *Mach. Learn.: Sci. Technol.*, 2021, **3**, 015009.

52 *Schrödinger Release 2023-4: Prime*, Schrödinger, LLC, New York, NY, 2023.

53 J. Li, R. Abel, K. Zhu, Y. Cao, S. Zhao and R. A. Friesner, The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 2794–2812.

54 F. Saponaro, S. Sestito, M. Runfola, S. Rapposelli and G. Chiellini, Selective Thyroid Hormone Receptor-Beta (TRβ) Agonists: New Perspectives for the Treatment of Metabolic and Neurodegenerative Disorders, *Front Med*, 2020, **7**, 331.

55 K. Wang, F. Chen, J. Wang and H. Liu, Drug discovery targeting thyroid hormone receptor β (THRβ) for the treatment of liver diseases and other medical indications, *Acta Pharm. Sin. B*, 2025, **15**, 35–51.

56 Z. Chen, S. Dong, F. Meng, Y. Liang, S. Zhang and J. Sun, Insights into the binding of agonist and antagonist to TAS2R16 receptor: a molecular simulation study, *Mol. Simul.*, 2018, **44**, 322–329.

57 H.-N. Mu, Q. Zhou, R.-Y. Yang, W.-Q. Tang, H.-X. Li, S.-M. Wang, J. Li, W.-X. Chen and J. Dong, Caffeic acid prevents non-alcoholic fatty liver disease induced by a high-fat diet through gut microbiota modulation in mice, *Food Res. Int.*, 2021, **143**, 110240.

58 E. S. Jin, J. D. Browning, R. E. Murphy and C. R. Malloy, Fatty liver disrupts glycerol metabolism in gluconeogenic and lipogenic pathways in humans, *J. Lipid Res.*, 2018, **59**, 1685–1694.

59 M. Zhao, Z. Jiang, H. Cai, Y. Li, Q. Mo, L. Deng, H. Zhong, T. Liu, H. Zhang, J. X. Kang and F. Feng, Modulation of the Gut Microbiota during High-Dose Glycerol Monolaurate-Mediated Amelioration of Obesity in Mice Fed a High-Fat Diet, *mBio*, 2020, **11**, e00190–20.

60 X. Ma, S. K. Okyere, L. Hu, J. Wen, Z. Ren, J. Deng and Y. Hu, Anti-Inflammatory Activity and Mechanism of

Cryptochlorogenic Acid from Ageratina adenophora, *Nutrients*, 2022, **14**, 439.

61 Y. Zhou, The Protective Effects of Cryptochlorogenic Acid on β-Cells Function in Diabetes in vivo and vitro via Inhibition of Ferroptosis, *Diabetes Metab Syndr Obes*, 2020, **13**, 1921–1931.

62 T. Esposito, F. Sansone, S. Franceschelli, P. Del Gaudio, P. Picerno, R. P. Aquino and T. Mencherini, Hazelnut (Corylus avellana L.) Shells Extract: Phenolic Composition, Antioxidant Effect and Cytotoxic Activity on Human Cancer Cell Lines, *Int. J. Mol. Sci.*, 2017, **18**, 392.

63 J. Lee, D. Kim, J. Choi, H. Choi, J.-H. Ryu, J. Jeong, E.-J. Park, S.-H. Kim and S. Kim, Dehydrodiconiferyl Alcohol Isolated from Cucurbita moschata Shows Anti-adipogenic and Anti-lipogenic Effects in 3T3-L1 Cells and Primary Mouse Embryonic Fibroblasts, *J. Biol. Chem.*, 2012, **287**, 8839–8851.

64 W. P. Walters, Going further than Lipinski's rule in drug design, *Expert Opin. Drug Discovery*, 2012, **7**, 99–107.

65 C. M. Chagas, S. Moss and L. Alisaraie, Drug metabolites and their effects on the development of adverse reactions: Revisiting Lipinski's Rule of Five, *Int. J. Pharm.*, 2018, **549**, 133–149.

66 D. F. Veber, S. R. Johnson, H.-Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, Molecular Properties That Influence the Oral Bioavailability of Drug Candidates, *J. Med. Chem.*, 2002, **45**, 2615–2623.

67 W. Wei, S. Cherukupalli, L. Jing, X. Liu and P. Zhan, Fsp3: A new parameter for drug-likeness, *Drug Discovery Today*, 2020, **25**, 1839–1845.

68 M. C. Sorkun, A. Khetan and S. Er, AqSolDB, a curated reference set of aqueous solubility and 2D descriptors for a diverse set of compounds, *Sci. Data*, 2019, **6**, 143.

69 J. L. Bolton and T. Dunlap, Formation and Biological Targets of Quinones: Cytotoxic versus Cytoprotective Effects, *Chem. Res. Toxicol.*, 2017, **30**, 13–37.

70 M. C. Egbujor, B. Buttari, E. Profumo, P. Telkoparan-Akillilar and L. Saso, An Overview of NRF2-Activating Compounds Bearing α,β-Unsaturated Moiety and Their Antioxidant Effects, *Int. J. Mol. Sci.*, 2022, **23**, 8466.

71 T. Nakayachi, E. Yasumoto, K. Nakano, S. R. M. Morshed, K. Hashimoto, H. Kikuchi, H. Nishikawa, M. Kawase and H. Sakagami, Structure-Activity Relationships of α, β-Unsaturated Ketones as Assessed by their Cytotoxicity against Oral Tumor Cells, *Anticancer Res.*, 2004, **24**, 737–742.

72 D. Xu, Q. Wang, W. Zhang, B. Hu, L. Zhou, X. Zeng and Y. Sun, Inhibitory Activities of Caffeoylquinic Acid Derivatives from Ilex kudingcha C.J. Tseng on α-Glucosidase from Saccharomyces cerevisiae, *J. Agric. Food Chem.*, 2015, **63**, 3694–3703.

73 J. Zhang, T. Sasaki, W. Li, K. Nagata, K. Higai, F. Feng, J. Wang, M. Cheng and K. Koike, Identification of caffeoylquinic acid derivatives as natural protein tyrosine phosphatase 1B inhibitors from *Artemisia princeps*, *Bioorg. Med. Chem. Lett.*, 2018, **28**, 1194–1197.

74 Y. Tian, X.-X. Cao, H. Shang, C.-M. Wu, X. Zhang, P. Guo, X.-P. Zhang and X.-D. Xu, Synthesis and In Vitro Evaluation of Caffeoylquinic Acid Derivatives as Potential Hypolipidemic Agents, *Molecules*, 2019, **24**, 964.

75 W. Lee, K. R. Ko, H.-K. Kim, S. Lim and S. Kim, Dehydrodiconiferyl alcohol promotes BMP-2-induced osteoblastogenesis through its agonistic effects on estrogen receptor, *Biochem. Biophys. Res. Commun.*, 2018, **495**, 2242–2248.

76 J. Lee, D. Kim, J. Choi, H. Choi, J.-H. Ryu, J. Jeong, E.-J. Park, S.-H. Kim and S. Kim, Dehydrodiconiferyl Alcohol Isolated from Cucurbita moschata Shows Anti-adipogenic and Anti-lipogenic Effects in 3T3-L1 Cells and Primary Mouse Embryonic Fibroblasts, *J. Biol. Chem.*, 2012, **287**, 8839–8851.