



## OPEN ACCESS

## EDITED BY

Shruthi Viswanath,  
National Centre for Biological Sciences, India

## REVIEWED BY

Padhmanand Sudhakar,  
Kumaraguru College of Technology, India  
Sree Ganesh Balasubramani,  
University of California, San Francisco,  
United States

## \*CORRESPONDENCE

Nivedita Singh,  
✉ niv234@gmail.com,  
✉ s.nivedita@bhu.ac.in

RECEIVED 05 August 2024

ACCEPTED 17 September 2024

PUBLISHED 09 October 2024

## CITATION

Singh N and Singh AK (2024) Exploration of phytoconstituents of Medhya Rasayana herbs to identify potential inhibitors for *cerebroside sulfotransferase* through high-throughput screening.

*Front. Mol. Biosci.* 11:1476482.

doi: 10.3389/fmolb.2024.1476482

## COPYRIGHT

© 2024 Singh and Singh. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Exploration of phytoconstituents of Medhya Rasayana herbs to identify potential inhibitors for *cerebroside sulfotransferase* through high-throughput screening

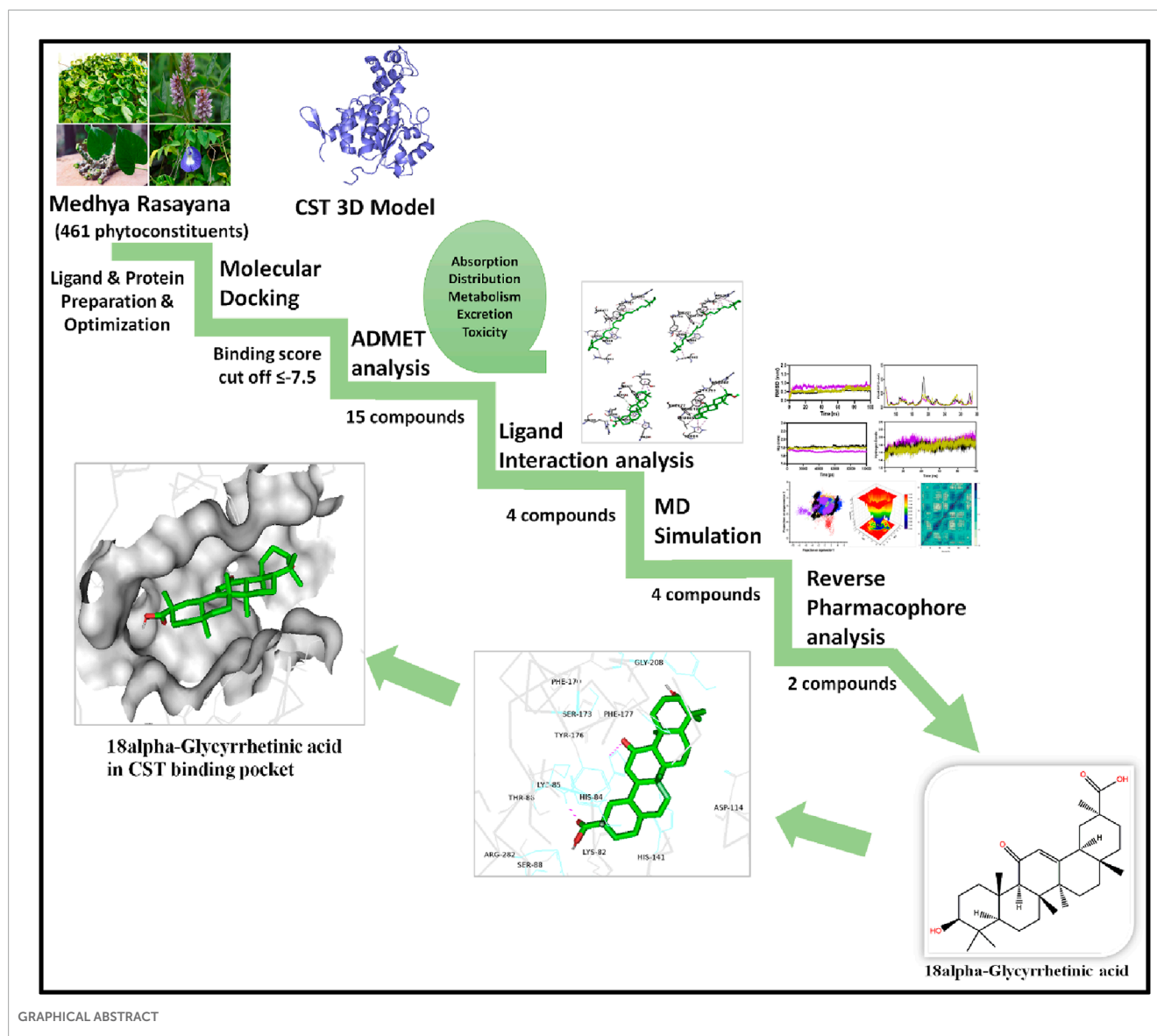
Nivedita Singh\* and Anil Kumar Singh

Department of Dravyaguna, Faculty of Ayurveda, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India

*Cerebroside sulfotransferase* (CST) is a key enzyme in sulfatide biosynthesis and regulation of the myelin sheath in the nervous system. To counter sulfatide accumulation with the deficiency of aryl sulfatase A, CST is considered a target protein in substrate reduction therapy in metachromatic leukodystrophy. In this study, 461 phytoconstituents from four herbs of Medhya Rasayana were screened using multi-pronged virtual screening methods including molecular docking, molecular dynamics (MD) simulation, and reverse pharmacophore analysis. The initial screening of the top 15 hits was based on the binding affinity of the compounds toward the CST substrate-binding site using the lowest free energy of a binding score cutoff of  $\leq -7.5$  kcal/mol, with the number of conformations in the largest cluster more than 75. The absorption, distribution, metabolism, and excretion (ADME) and toxicity-based pharmacokinetic analysis delivered the top four hits: 18alpha-glycyrrhetic acid, lupeol, alpha carotene, and beta-carotene, with high blood-brain barrier permeability and negligible toxicity. Furthermore, a 100-ns simulation of protein-ligand complexes with a trajectory analysis of structural deviation, compactness, intramolecular interactions, principal component analysis, free energy landscape, and dynamic cross-correlation analysis showed the binding potential and positioning of the four hits in the binding pocket. Thus, an in-depth analysis of protein-ligand interactions from pre- and post-molecular dynamics simulation, along with reverse pharmacophore mapping, suggests that 18alpha-glycyrrhetic acid is the most potent and specific CST inhibitor, while beta-carotene could be considered the second most potent compound for CST inhibition as it also exhibited overall stability throughout the simulation. Therefore, the computational drug screening approach applied in this study may contribute to the development of oral drugs as a therapeutic option for metachromatic leukodystrophy.

## KEYWORDS

*cerebroside sulfotransferase*, substrate reduction therapy, metachromatic leukodystrophy, inhibition, drug, active site



## 1 Introduction

Metachromatic leukodystrophy (MLD) is an autosomal recessive lysosomal storage neurodegenerative disorder, pathologically characterized by progressive motor and cognitive dysfunction (Miyake et al., 2021; Lamichhane and Cabrero, 2023; Penati et al., 2017; Singh and Singh, 2024b). MLD is caused by the deficiency of aryl sulfatase A (ARSA), which leads to the accumulation of sulfatide, the main glycolipid involved in stabilizing the myelin sheath equilibrium in the central and peripheral nervous systems (Sessa et al., 2016; Fumagalli et al., 2022). Sulfatide accumulation may cause various pathophysiological conditions, including progressive demyelination, neuroinflammation, communication gap, astrocyte dysfunction, developmental delay, speech disorder, and in severe cases it causes death within 5–6 years of the early onset of the disease (Miyake et al., 2021; Lamichhane and Cabrero, 2023; Fumagalli et al., 2022). Based on the age of onset, MLD is categorized into three major clinical types:

late-infantile ( $\leq 30$  months); juvenile, which is subdivided into early juvenile [30 months–6 years] and late juvenile [7–16 years]; and adult MLD ( $\geq 17$  years) (Miyake et al., 2021; Fumagalli et al., 2022; Chang et al., 2024). Worldwide, the prevalence of MLD is 1.4–1.8 in 100,000 or 1 in 40,000, bringing it to the loop of rare diseases of greater concern (Lamichhane and Cabrero, 2023; Chang et al., 2024; Shaimardanova et al., 2020).

To date, more than 280 mutations have been reported in the ARSA gene, which poses a major challenge in the success of existing therapeutic options, including gene therapy, hematopoietic stem cell therapy, enzyme replacement therapy, and chaperone therapy, thus making these options costly and, therefore, demanding a case-by-case approach that leads the treatment of this rare disease beyond the reach of the larger population (Shaimardanova et al., 2020; Kurtzberg, 2022; Eichler et al., 2022; Fernández-Pereira et al., 2021; Sevin and Deiva, 2021). Another major obstacle to the success of existing therapies is the blood–brain barrier (BBB) in delivering the therapeutic gene or enzyme to the

target site (Miyake et al., 2021). In contrast to existing therapies, which are mainly ARSA-dependent, a new approach of substrate reduction therapy (SRT) can be an alternate therapeutic option for metachromatic leukodystrophy. SRT has shown its success in various single-gene disorders including some key lysosomal storage disorders (Mistry et al., 2023; Komada et al., 2021; Istiti et al., 2022). SRT focuses on the development of oral drugs to treat pathophysiological conditions. Miglustat and eliglustat are the two FDA-approved oral drugs for Gaucher's disease (Mistry et al., 2023; Komada et al., 2021; Istiti et al., 2022).

In MLD, the therapeutic focus in substrate reduction therapy is the development of a specific, potent, and competitive inhibitor which targets the catalytic action of the rate-limiting enzyme, *cerebroside sulfotransferase*, that is involved in the biosynthesis of sulfatides (Singh and Singh, 2024b; Yaghootfam et al., 2007; Li et al., 2020). To date, the development in substrate reduction therapy has been at the nascent stage in MLD. Experimental data with available CST inhibitors are limited, even though the full-length CST sequence and cDNA are available (Li et al., 2020). The lack of structural information of proteins due to the unavailability of the three-dimensional structure of CST has been a major roadblock in initiating preliminary *in silico*-based drug discovery, which has great potential in screening drug-like candidates in the shortest possible time with less expenses. The development of the three-dimensional homology model of CST by our group, which used a multipronged modeling and validation approach, is a breakthrough in SRT research in MLD (Singh and Singh, 2024a). This study uses this model for the screening of phytoconstituents from Medhya Rasayana.

Medhya Rasayana is a group of Ayurvedic medicinal herbs with neuroprotective effects against various neurodegenerative diseases and known to improve cognitive function and neural tissue regeneration, retard brain aging, and balance mental health (Kulkarni et al., 2012; Sarokte and Rao, 2013; Rashmi et al., 2017). Medhya Rasayana comprises four major herbs: Mandukaparni (*Centella asiatica* Linn.), Yastimadhu (*Glycyrrhiza glabra* Linn.), Guduchi (*Tinospora cordifolia* Miers), and Shankhpushpi (*Convolvulus pluricaulis* Choisy) (Kulkarni et al., 2012). These herbs are rich in phytoconstituents with unique bioactive properties. This study focuses on identifying potent and specific bioactive compounds as inhibitors against CST through multipronged *in silico* studies. High-throughput virtual screening is the most common and reliable strategy used to identify the potent and specific inhibitors and potential lead molecules from the diverse dataset of small molecules at low cost and rapid pace using the recent advancement in the field of bioinformatics (Singh et al., 2018; Singh et al., 2017; Yasuo and Sekijima, 2019). With this perspective, the present study opens the door for many futuristic studies to achieve a marketed oral drug for metachromatic leukodystrophy that could bring this hereditary disease within the reach of a possible cost-effective treatment loop.

## 2 Methodology

### 2.1 Resources

The computational study was executed using the high-performance super-computing facility, PARAM Shivay, installed

at the Indian Institute of Technology, Banaras Hindu University, Varanasi, India, with a capacity of 837 TFLOPS with Intel(R) Xeon(R) Gold 6148 CPU @2.40 GHz and 40 CPUs per node. This study used key software programs including AutoDock 4.2, GROMACS 2023, Origin 2024, Chimera 1.17.3, PyMOL, Discovery Studio visualizer, VMD, and Bio3D tool with R package and key webtools including PharmMapper, pkCSM, SwissADME, and ProTox 3.0.

### 2.2 Protein structural analysis and receptor grid generation

For screening of bioactive phytoconstituents, the 3D model of CST developed in our earlier study was used as a receptor. The model comprises 69–336 amino acid residues of the full-length protein with 423 residues and covers the entire catalytic region of the protein with a sulfuryl acceptor substrate (galactocerebroside)-binding site and a sulfuryl donor co-substrate (PAPS)-binding site on a linear horizontal plane (Singh and Singh, 2024a; Singh and Singh, 2024b). The model protein was processed using AutoDock 4.2 by the removal of water molecules and heteroatoms and the addition of hydrogen atoms with proper assignment of atom type and Gasteiger charge to generate a PDBQT file of the protein. A grid box of 90 × 90 × 90 Å and spacing of 0.253 Å around the protein active site was generated considering the key residues LYS82, HIS84, LYS85, HIS141, PHE170, TYR176, PHE177, TYR203, and ARG202, and a grid file (.gpf) was created.

### 2.3 Ligand preparation

We used a set of compounds from the Indian Medicinal Plants, Phytochemistry, And Therapeutics (IMPPAT 2.0) online database (Vivek-Ananth et al., 2023; Mohanraj et al., 2018). The 3D (.mol2) structures of 81 compounds from *C. asiatica* Linn (Mandukaparni), 310 from *G. glabra* Linn (Yastimadhu), 52 from *T. cordifolia* Miers (Guduchi), and 18 from *C. pleuricaulis* Choisy (Shankhpushpi) were converted to .pdbqt files after energy minimization and assigning proper atom types using AutoDock Raccoon, a virtual screening file preparation tool (Forli et al., 2016).

Following the preparation of the protein, ligands, and grid files, AutoDock Raccoon was further used for the preparation of docking (.dpf) files for each ligand and the subsequent arrangement of all files in a single separate folder for each protein–ligand complex with the generation of a single virtual screening script file (.sh) for performing molecular docking-based virtual screening of all ligands simultaneously under the Linux environment. The parameters used for the generation of .dlg files for docking run were 100 GA run, a population size of 300, maximum number of generations of 27,000, and maximum number of evaluations of 25,000,000. AutoDock applied the Lamarckian genetic algorithm and a gradient-based local search method for protein–ligand interactions. The pKi and ligand efficiency are two major parameters for assessing the binding potential of ligands in the active site. The best-docked conformation was selected, processed using custom Python scripts, and visualized using .pdb visualization tools including PyMOL and Discovery

Studio Visualizer to analyze protein–ligand interactions and ligand-binding patterns in the active site.

## 2.4 *In silico* drug-likeness properties, ADME, and toxicity analysis

The selected top hits based on the binding score and number of conformations in the largest cluster were subjected to pharmacokinetic property analysis using pkCSM, SwissADME, and ProTox webservers (Banerjee et al., 2018; Daina et al., 2017; Pires et al., 2015). The canonical Simplified Molecular Input Line Entry System (SMILE) of the selected compounds was used as the entry data in these servers for their pharmacokinetic and drug-likeness property analysis through adsorption, distribution, metabolism, and excretion along with toxicity studies.

## 2.5 Molecular dynamics simulations

All atom molecular dynamic (MD) simulations were performed using the GROMACS 2023.1 software package using the CHARMM27 all-atom additive force field. For the simulation, a dodecahedron simulation box was created with a minimum distance of 1.2 Å from the box edge, and periodic boundary conditions were applied to minimize the edge effect. Each box was solvated with the TIP3P water solvation model, and the charges on the system were then neutralized by the addition of chloride (Cl<sup>-</sup>) ions, and thereafter, the solvated system was energy-minimized using the steepest descent algorithm. The energy-minimized system was then equilibrated to 1,000-ps (1 ns) NVT simulation with a time step of 2 fs at 300 K. Next, the system was equilibrated to 1,000-ps NPT simulations with a time step of 2.0 fs at 300 K. The LINCS algorithm was used to constrain the bond lengths. The final MD simulation run was performed for 100 ns, and trajectories were analyzed using different MD parameters including root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), hydrogen bonds, principal component analysis (PCA), and dynamic cross-correlation matrix (DCCM) analysis of the protein–ligand complexes.

## 2.6 Cross-target prediction

The PharmMapper server was used to identify a wide range of targets using an innovative reverse pharmacophore mapping approach (Wang et al., 2017; Liu et al., 2010). The best mapping poses of the submitted molecule (.mol2) were aligned against all target proteins available in PharmTargetDB. The algorithm used to perform the reverse pharmacophore matching protocol comprises a sequential combination of triangle hashing (TriHash) and genetic algorithm (GA) optimization (Liu et al., 2010). Based on the calculated highest fit-score (cutoff >5.0) between the small compound and the pharmacophore models, the probable protein targets were ranked. It was imperative to check the disease-causing potential of the targets.

# 3 Results

## 3.1 Molecular docking-based virtual screening

Virtual screening is a computational approach used to screen libraries of compounds available in databases against the target protein to identify a potential drug candidate for a targeted disease (Singh et al., 2018; Singh et al., 2021; Zare et al., 2024; Pirolli et al., 2023). In the CST-led research to develop substrate reduction therapy for metachromatic leukodystrophy, the state-of-the-art *in silico* approach could be the most promising and directional method for future studies. Toward this approach, the development of a homology model of the CST protein using various computational algorithms is an important step (Singh and Singh, 2024a). In the present study, this 3D model was utilized for screening specific and potent phytoconstituents of four major herbs of Medhya Rasayana. Out of a total of 461 bioactive compounds, 81 compounds from *C. asiatica* Linn., 310 from *G. glabra* Linn., 52 from *T. cordifolia* Miers, and 18 from *C. pleuricaulis* Choisy were screened against CST, which are detailed in [Supplementary Tables S1–S4](#). In the initial level of screening, using the lowest free energy of a binding cutoff of  $\leq -7.5$  kcal/mol and the number of conformation cutoff of  $\geq 75$  in the largest conformation cluster, the top 15 compounds were selected, i.e., 3 from *C. asiatica* and 12 from *G. glabra*. Compounds belonging to *T. cordifolia* and *P. Choisy* were found to be less potent and less specific toward CST and, hence, were not considered for further study. The docking score of these top 15 compounds falls between  $-7.57$  and  $-10.32$  kcal/mol. With a ligand efficiency of  $\leq 0.19$  per non-hydrogen atom, the top 15 compounds were considered promising for drug-likeness property analysis. Apart from  $\alpha$ - and  $\beta$ -carotene, all 13 compounds strictly followed the Lipinski rule of 5 parameters in terms of molecular mass ( $<500$ ), number of hydrogen bond donors ( $<5$ ), hydrogen bond acceptors ( $<10$ ), number of rotatable bonds, and topological surface area (TPSA). Despite failing at the Lipinski rule of five parameters,  $\alpha$ - and  $\beta$ -carotene were considered for further studies because of their wider therapeutic potential in neurodegenerative diseases, and these molecules have already been tested in the human body (Banerjee et al., 2023; Abrego-Guandique et al., 2023; Hira et al., 2019; Narisawa et al., 1996). Additionally, these compounds showed good lipophilicity with a  $\log p$ -value  $> 0$ , which is imperative for compounds to cross the lipophilic membrane and, thus, strengthen the movement of these molecules to the target site. Therefore, all 15 hits were considered for absorption, distribution, metabolism, and excretion (ADME) and toxicity analysis to eliminate false positives from this list, which is provided in [Table 1](#).

## 3.2 Absorption, distribution, metabolism, excretion, and toxicity profiling

After the administration of the drug in the body, it goes through absorption, distribution, metabolism, and excretion processes. In this process, the drug interacts with desirable or undesirable targets and causes a pharmacological impact (Zhang and Tang, 2018). Thus, the bioavailability of a drug depends on the rate of absorption, metabolism, and transportation of the drug to the target site (Zhang

TABLE 1. Binding score and physicochemical properties of the top 15 compounds.

Sl. no.	Compounds	Common name	Source	Binding affinity (kcal/mol)	Ligand efficiency (kcal/mol/non-h-atom)	pKi (nM)	Molecular weight	Hydrogen bond donor	Hydrogen bond acceptor	Rotational bond	Ring	LogP	TPSA (Å <sup>2</sup> )
1	IMPHY012226	18alpha-Glycyrrhethinic acid	<i>Glycyrrhiza glabra</i>	-10.32	-0.30	7.56	470.69	2	3	1	5	6.41	74.6
2	IMPHY002304	Liquoric acid	<i>Glycyrrhiza glabra</i>	-10.22	-0.29	7.49	484.68	2	4	1	6	5.4	83.83
3	IMPHY011842	11-Deoxoglycyrrhethinic acid	<i>Glycyrrhiza glabra</i>	-9.92	-0.30	7.27	456.71	2	2	1	5	7.23	57.53
4	IMPHY015702	Glabric acid	<i>Glycyrrhiza glabra</i>	-9.67	-0.28	7.09	486.69	3	4	1	5	5.38	94.83
5	IMPHY011880	Ursolic acid	<i>Centella asiatica</i>	-9.28	-0.28	6.80	456.71	2	2	1	5	7.09	57.53
6	IMPHY015709	28-Hydroxyglycyrrhethinic acid	<i>Glycyrrhiza glabra</i>	-9.21	-0.26	6.75	486.69	3	4	2	5	5.39	94.83
7	IMPHY010136	18α-Hydroxyglycyrrhethinic acid	<i>Glycyrrhiza glabra</i>	-9.03	-0.26	6.61	486.69	3	4	1	5	5.53	94.83
8	IMPHY005115	Ursolic acid lactone	<i>Centella asiatica</i>	-8.54	-0.26	6.26	456.71	1	3	0	6	6.76	46.53
9	IMPHY001586	Isoglabrolide	<i>Glycyrrhiza glabra</i>	-8.44	-0.25	6.18	468.68	1	4	0	6	6.01	63.6
10	IMPHY012473	Lupeol	<i>Glycyrrhiza glabra</i>	-8.35	-0.27	6.12	426.73	1	1	1	5	8.02	20.23
11	IMPHY006740	Glabrolide	<i>Glycyrrhiza glabra</i>	-8.23	-0.24	6.04	468.68	1	4	0	6	5.86	63.6
12	IMPHY011609	alpha-Carotene	<i>Centella asiatica</i>	-7.90	-0.20	5.79	536.89	0	0	10	2	12.46	0
13	IMPHY011707	beta-Carotene	<i>Glycyrrhiza glabra</i>	-7.64	-0.19	5.59	536.89	0	0	10	2	12.61	0

(Continued on the following page)

TABLE 1 (Continued) Binding score and physicochemical properties of the top 15 compounds.

Sl. no.	Compound	Common name	Source	Binding affinity (kcal/mol)	Ligand efficiency (kcal/mol/non-h-atom)	pKi (nM)	Molecular weight	Hydrogen bond donor	Hydrogen bond acceptor	Rotational bond	Ring	LogP	TPSA (Å <sup>2</sup> )
14	IMPHY011620	Lutein	<i>Glycyrrhiza glabra</i>	-7.59	-0.19	5.57	568.886	2	2	10	2	10.4	40.46
15	IMPHY001500	Hispaglabridin B	<i>Glycyrrhiza glabra</i>	-7.57	-0.19	5.55	390.479	1	4	1	5	5.48	47.92

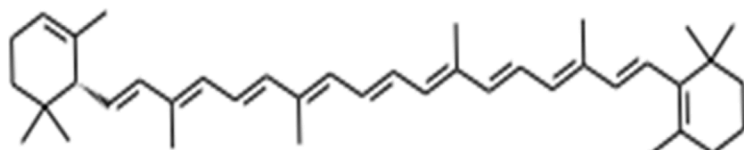
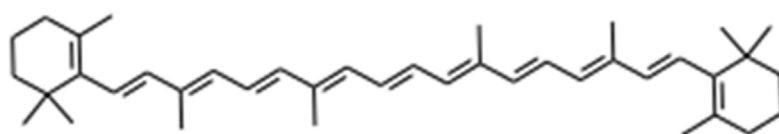
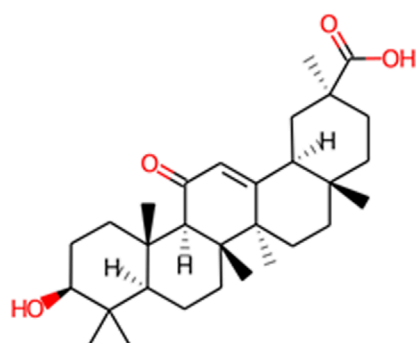
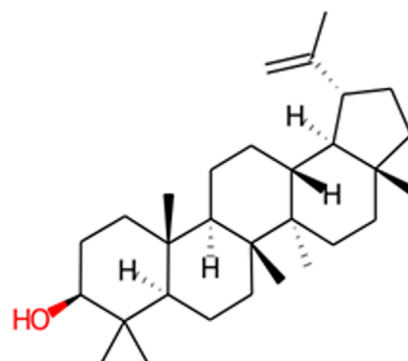
and Tang, 2018; Stielow et al., 2023). Human intestinal absorption and lipophilicity of compounds are other important parameters to be considered to ensure the easy diffusion of compounds. The cytochrome p450-based drug response to body metabolism exhibits genetic variability and plays a critical role in the detoxification of drugs and homeostasis (Zhao et al., 2021). In this study, CYP2C9, CYP2D6, and CYP3A4 inhibitors were used to analyze the drug response at the metabolism stage. Renal OCT2 is a renal transporter that plays an important role in determining the renal clearance or deposition of drugs (Zou et al., 2021; Wright, 2019). Except few, most of the selected compounds showed good intestinal absorption capacity, along with body metabolism and renal excretion potential. Since MLD is a brain disorder, the major deciding parameters for the potential drug candidates were BBB permeability, which screened four compounds—IMPHY012226 (18alpha-glycyrrhetic acid), IMPHY012473 (lupeol), IMPHY011609 (alpha-carotene), and IMPHY011707 (beta-carotene)—with considerable blood–brain barrier permeability. In order to select a potential lead molecule, *in silico* toxicity study becomes another crucial criterion due to its accuracy, accessibility, and rapidity in preclinical-level screening and providing a potential lead scaffold for further optimization (Raies and Bajic, 2016; Hemmerich and Ecker, 2020; Yang et al., 2018). pkCSM and ProTox 3.0 are freely available *in silico* toxicity study servers that use the SMILE of each compound to evaluate various toxicity parameters including Ames mutagenicity, hepatotoxicity, and cytotoxicity (Banerjee et al., 2018; Pires et al., 2015). These four phytoconstituents belonged to class “4” of the ProTox-predicted toxicity class based on their LD<sub>50</sub> range between 560 and 2,000 mg/kg, suggesting that the selected four compounds are mainly nontoxic. Table 2 provides the details of ADME and toxicity analysis of the top 15 hits, and Figure 1 shows the structural details of the best 4 compounds screened through ADME and toxicity analysis.

### 3.3 Pre-MD/docking-based protein-ligand interaction pose analysis

Protein–ligand interaction analysis was imperative to obtain an insight into the binding pattern of compounds in the substrate-binding site of the CST protein. The substrate-binding site comprises a left-end polar site dominated by Lys85, Phe170, Lys82, and Ser88; a middle part flanked on both sides by His84 and His141; and an aromatic site at the right end dominated by Tyr203 and Phe170. With the lowest free binding energy of -10.32 kcal/mol and 98 similar conformations in the largest cluster, 18alpha-glycyrrhetic acid was found to be the most potential ligand with the highest binding affinity among the 4 in the binding pocket of the CST protein. The compound consists of five steroidal rings with a carboxylic group attached at one end and a hydroxyl group at the other end. These ends determined the orientation of the compound toward the polar site based on the degree of polarity of the carboxylic group, which interacted with Lys85 by hydrogen bonding. One oxygen atom was attached by a double bond at the third ring in the middle of the compound positioned at the middle of the active site and was flanked on both sides by His84 and His141 and formed a hydrogen bond with His84. At the right side of the binding pocket, methyl groups located on the rightmost ring of the compound

TABLE 2 Pharmacological profile of the top 15 ligand molecules that were derived from pkCSM, SwissADME, and ProTox web servers.

Sl. no.	Compounds common name	Absorption		Distribution	Metabolism			Excretion		Toxicity	
		Gastrointestinal absorption	BBB permeation		CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	Renal OCT2 substrate	Ames test	Hepatotoxicity	Predicted toxicity class
1	18alpha-Glycyrrhetic acid	High	Moderate	No	No	No	No	No	No	No	4
2	Liquoric acid	High	Poor	No	No	No	No	No	No	No	3
3	11-Deoxyglycyrrhetic acid	Low	Poor	No	No	No	No	No	No	No	4
4	Glabric acid	High	Poor	No	No	No	No	No	No	No	4
5	Ursolic acid	Low	Poor	No	No	No	No	No	No	Yes	4
6	28-Hydroxyglycyrrhetic acid	High	Poor	No	No	No	No	No	No	No	4
7	18α-Hydroxyglycyrrhetic acid	High	Poor	No	No	No	No	No	No	No	6
8	Ursolic acid lactone	High	Moderate	No	No	No	No	No	No	Yes	5
9	Isoglabrolide	High	Moderate	No	No	No	No	No	No	No	4
10	Lupeol	Low	High	No	No	No	No	No	No	No	4
11	Glabrolide	High	Poor	No	No	No	No	No	No	No	5
12	alpha-Carotene	High	High	No	No	No	No	No	No	No	4
13	beta-Carotene	High	High	No	No	No	No	No	No	No	4
14	Lutein	Low	Poor	No	No	No	No	No	No	No	2
15	Hispaglabridin B	High	Poor	Yes	No	Yes	No	Yes	No	No	4

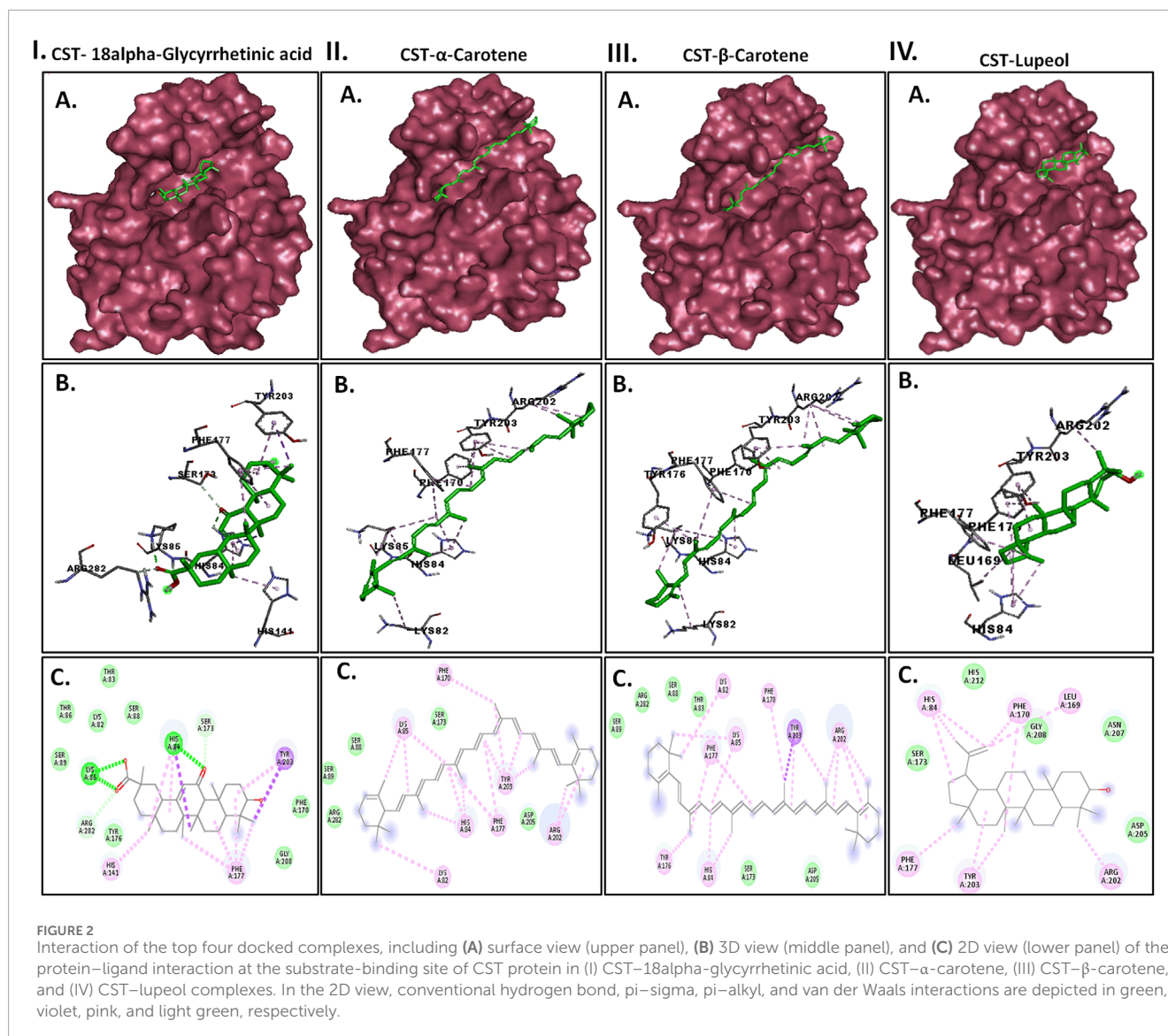
**A.  $\alpha$ -carotene****B.  $\beta$ -carotene****C. 18alpha-Glycyrrhetic acid****D. Lupeol**

**FIGURE 1**  
Top four selected compounds. (A)  $\alpha$ -carotene. (B)  $\beta$ -carotene. (C) 18alpha-glycyrrhetic acid. (D) lupeol.

interacted with the residues Tyr203 and Phe177 via pi-sigma and pi-alkyl bonds, respectively (Figure 2I). Both  $\alpha$ -carotene and  $\beta$ -carotene complexes shared similarity in the interaction pattern in the CST-binding pocket.  $\alpha$ -Carotene and  $\beta$ -carotene are isomers, differing in the position of the double bond on the ring at one end oriented toward the polar site in the binding pocket. Lys82 and Lys85 played a critical role in the interaction with the isomeric fragment of these compounds. In  $\alpha$ -carotene, the methyl group at the C2 position was oriented toward Lys85 and, thus, interacted with it, whereas in  $\beta$ -carotene, the methyl group at the C2 position was oppositely oriented and away from Lys85. In the CST- $\beta$ -carotene complex, the two methyl groups at the C6 position were oriented toward Lys85 and interacted via the pi-alkyl interaction, whereas in the CST- $\alpha$ -carotene complex, the methyl groups at the C6 carbon

were positioned away from Lys85 and interacted with Lys82. The aliphatic chain between the two aromatic rings interacted in a nearly similar fashion in both complexes with key residues Phe177, Tyr203, phe170, and His84, while Arg202 interacted with the aromatic ring on the right side of the binding pocket (Figures 2II, III). In contrast to the other three compounds that broadly occupied the active site core, in the CST-lupeol complex, the ligand predominantly occupied the right side of the binding pocket, which might be due to its small size and the missing of a polar carboxylic group at one end as it was in 18alpha-glycyrrhetic acid. The lack of occupancy of the polar site by lupeol is due to its five-carbon nonpolar ring that interacted strongly at the aromatic site of the binding pocket (Figure 2IV), while the six-membered polar ring in 18alpha-glycyrrhetic acid stretched the compound toward the





polar site and, thus, was strongly accommodated in the active site (Figure 2I). The details of the interaction types are given in Table 3. To further understand the protein–ligand interaction under a dynamic environment, all four compounds were considered for molecular dynamic simulations.

### 3.4 Molecular dynamic simulation

MD simulation is a computational approach used to analyze and optimize the overall stability of the protein–ligand complexes under atomistic simulation conditions with a dynamic aqueous environment (Guterres and Im, 2020; Kurniawan and Ishida, 2022; Khan et al., 2016). MD simulation provides a cumulative idea about the movement of every atom or atom in the protein over the simulation time span to study important biological processes, including the impact of ligand binding on the overall protein dynamics and the way the macromolecule responds at the atomic level with the binding or unbinding of the ligand

(Alrouji et al., 2023; Hollingsworth and Dror, 2018). In this study, to understand the in-depth dynamic behavior of the protein–ligand interaction, MD simulation was carried out for a time span of 100 ns to evaluate the strength and stability of the protein–ligand complexes through trajectory analysis with various parameters including RMSD, RMSE, Rg, SASA, and hydrogen bonding. We also carried out principal component analysis, free energy landscape analysis, and dynamic cross-correlation analysis to understand the dominant motions responsible for the binding pattern and stability of the protein–ligand complex.

#### 3.4.1 Structural deviation and flexibility with RMSD and RMSF analysis

The RMSD measures the conformational deviation of the protein structure from the initial docked conformation to the final conformation under the dynamic aqueous environment over the simulation time span to ensure the stability of the predicted protein–ligand complex after ligand binding (Zare et al., 2024; Aier et al., 2016). In this study, in a simulation time of 100 ns, the

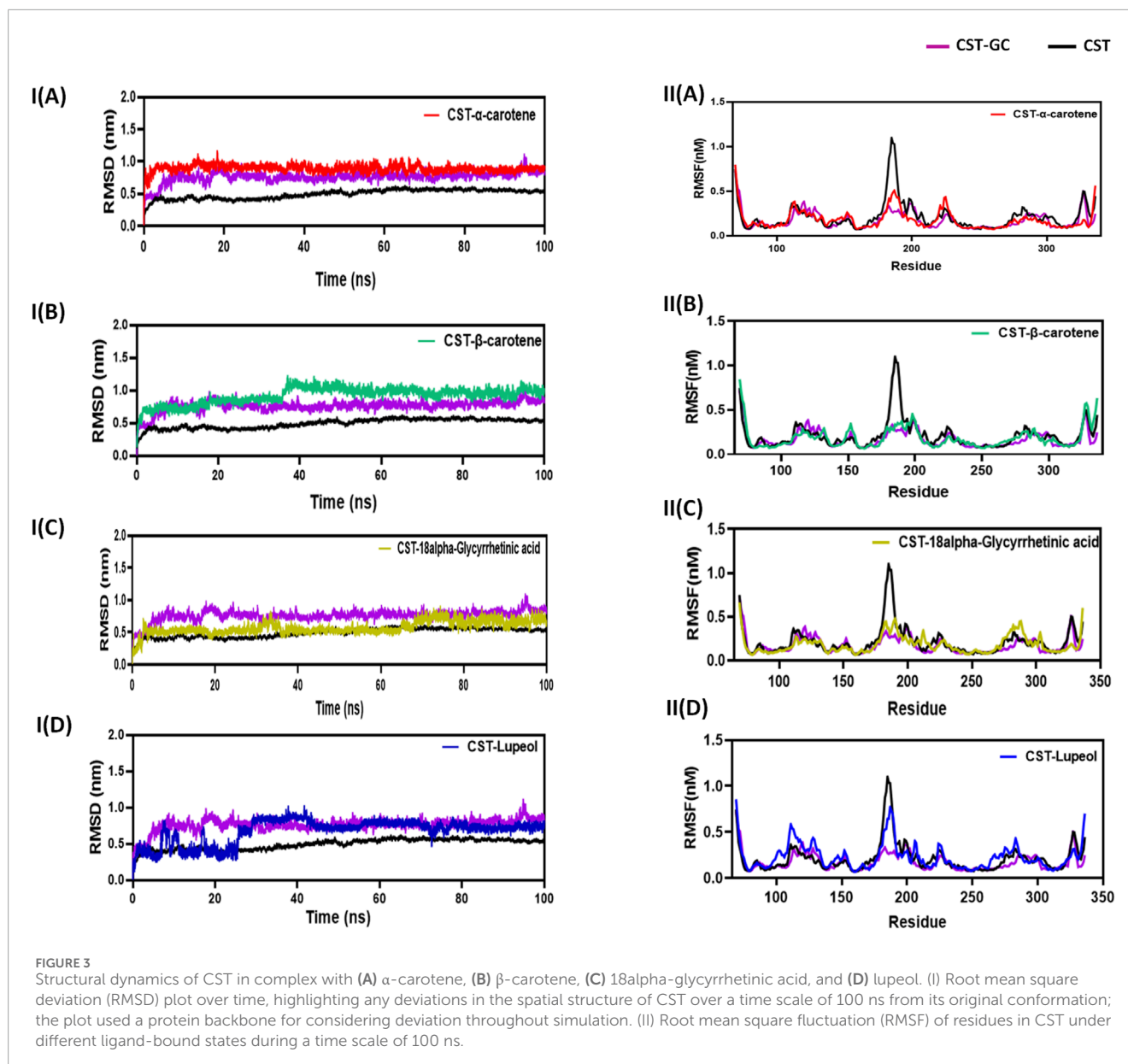
TABLE 3 Non-bonded interaction in docked protein–ligand conformation.

Sl. no.	Compound	Residue in contact	Bond type	Distance (Å)
1	18alpha-Glycyrrhetic acid	Lys85	Conventional hydrogen bond	1.60
		Lys85	Conventional hydrogen bond	2.10
		His84	Conventional hydrogen bond	2.20
		His84	Pi–sigma	4.25
		His84	Pi–alkyl	4.24
		His84	Pi–alkyl	5.30
		His141	Pi–alkyl	4.61
		Phe177	Pi–alkyl	5.37
		Phe177	Pi–alkyl	6.35
		Phe177	Pi–alkyl	6.63
		Phe177	Pi–alkyl	6.58
		Tyr203	Pi–sigma	4.60
		Tyr203	Pi–alkyl	5.10
		Ser173	van der Waals	3.48
		Arg282	van der Waals	4.33
2	α-Carotene	Lys82	Pi–alkyl	5.08
		Lys85	Pi–alkyl	3.12
		Lys85	Pi–alkyl	4.29
		Lys85	Pi–alkyl	6.93
		His84	Pi–alkyl	4.85
		His84	Pi–alkyl	4.96
		His84	Pi–Alkyl	5.76
		Phe170	Pi–alkyl	5.01
		Phe177	Pi–alkyl	5.36
		Phe177	Pi–alkyl	5.53
		Arg202	Pi–alkyl	3.37
		Arg202	Pi–alkyl	3.96

(Continued on the following page)

TABLE 3 (Continued) Non-bonded interaction in docked protein–ligand conformation.

Sl. no.	Compound	Residue in contact	Bond type	Distance (Å)
		Tyr203	Pi–alkyl	3.12
		Tyr203	Pi–alkyl	3.75
		Tyr203	Pi–alkyl	5.04
		Tyr203	Pi–alkyl	5.54
3	β-Carotene	Lys82	Pi–alkyl	4.51
		Lys85	Pi–alkyl	4.07
		Lys85	Pi–alkyl	5.63
		His84	Pi–alkyl	3.83
		His84	Pi–alkyl	4.68
		Phe170	Pi–alkyl	5.18
		Tyr176	Pi–alkyl	5.05
		Phe177	Pi–alkyl	5.87
		Phe177	Pi–alkyl	6.24
		Arg202	Pi–alkyl	3.74
		Arg202	Pi–alkyl	4.38
		Arg202	Pi–alkyl	4.51
		Arg202	Pi–alkyl	5.83
		Tyr203	Pi–sigma	3.86
		Tyr203	Pi–alkyl	4.21
		Tyr203	Pi–alkyl	4.38
		4	Lupeol	His84
His84	Pi–alkyl			5.84
His84	Pi–alkyl			5.85
Leu169	Pi–alkyl			5.32
Phe170	Pi–alkyl			3.74
Phe170	Pi–alkyl			5.50
Phe177	Pi–alkyl			6.31
Tyr203	Pi–alkyl			3.82
Tyr203	Pi–alkyl			4.29
Arg202	Pi–alkyl			4.18



CST-18alpha-glycyrrhetic acid and CST- $\alpha$ -carotene complexes were found to be the most stable with negligible fluctuation, while the CST- $\beta$ -carotene complex showed overall stability with little fluctuation at 40 ns. The CST-lupeol complex showed initial rapid fluctuation, but after 25 ns, it also achieved stability. The average deviation found for CST, CST-GC, CST-18alpha-glycyrrhetic acid, CST- $\alpha$ -carotene, CST- $\beta$ -carotene, and CST-Lupeol was 0.49, 0.76, 0.57, 0.89, 0.92, and 0.67, respectively. Thus, RMSD simulation showed that the four complexes of CST could maintain overall structural stability during the simulation and no significant conformational changes occurred in the protein structure after ligand binding as the RMSD was maintained throughout the simulation either between the RMSD of free CST and the CST-GC complex or closer to the RMSD of the CST-GC complex, which is a good indication toward achieving competitive inhibition (Figures 3IA–D) (Singh and Singh, 2024a).

Following the protein conformational deviation studies with RMSD, protein structural fluctuation at the residue level was analyzed with RMSF calculation of the protein–ligand complex (Martínez, 2015; Paul et al., 2022). In this study, the fluctuations in the residues of the CST protein in the presence of the selected ligand were compared with the RMSF of the free CST and CST-GC complex. The average RMSF of free CST, CST-GC, CST-18alpha-glycyrrhetic acid, CST- $\alpha$ -carotene, CST- $\beta$ -carotene, and CST-lupeol was 0.15, 0.17, 0.18, 0.176, 0.18, and 0.23, respectively. As depicted in (Fig 3IIA–D), a wider residual fluctuation was clearly observed in free CST between the amino acid residues 175 to 190, which represents the loop region in the substrate-binding site, while in the CST-GC complex, residual fluctuation was minimized significantly, as interaction with the ligand may provide less room for fluctuation (Singh and Singh, 2024a). Among the four test complexes,

CST-18alpha-glycyrrhetic acid and CST- $\beta$ -carotene showed relatively better stability between amino acid residues 180 and 190, while in the CST-lupeol complex, the fluctuation in this region was close to that of the free CST. Overall, the CST-18alpha-glycyrrhetic acid and CST- $\beta$ -carotene complexes showed residual fluctuation close to that of the CST-GC complex, which is again a good indication of the competitive inhibition of CST. The CST-lupeol complex was found to be least stable at the residual level.

### 3.4.2 Structural compactness with Rg, SASA, and hydrogen bonding analysis

The next level of screening was to evaluate changes in the size or structural compactness of the protein in the presence of ligands during the simulation by measuring Rg of the protein-ligand complex (Lobanov et al., 2008; Rampogu et al., 2022; Funari et al., 2022). Rg also estimates the flexibility of the protein under complex formation by comparing it with Rg of either the free protein or the substrate-bound protein. The larger Rg in the free CST indicated lesser compactness and greater flexibility, whereas the smaller Rg in the CST-GC complex indicated greater compactness and rigidity in the protein structure in the presence of the substrate. As depicted in Figure 4A, a similar trend was visible with inhibitor binding. Of the four, the Rg of the CST- $\beta$ -carotene complex was most closely related to the Rg of the protein-substrate complex. In the CST-18alpha-glycyrrhetic acid complex, structural compactness of the protein was between that of free CST and the CST-GC complex, suggesting that the presence of the ligand slightly compressed the protein structure. The Rg values of the CST- $\alpha$ -carotene and CST-lupeol complexes were found to be between the Rg of the free CST and the Rg of the CST-GC complex, also suggesting that the protein structure takes on stiffness and compactness in the presence of these ligands (Figure 4A).

The SASA measures the solvent-accessible surface area of the free protein or protein in the protein-ligand complex. It basically calculates the surface area of molecules that is exposed to solvent molecules (Alrouji et al., 2023; Durham et al., 2009; Ali et al., 2014; Savojardo et al., 2021). The SASA analyzes how parts of the protein come into contact with the solvent over the simulation time. The average SASA range for the selected compounds was between 170 and 180 nm<sup>2</sup>, which falls between the range of the SASA of free CST (180.1 nm<sup>2</sup>) and that of CST complexed with the substrate (173.07 nm<sup>2</sup>) (Singh and Singh, 2024a). The average SASA of the CST-18alpha-glycyrrhetic acid, CST- $\alpha$ -carotene, CST- $\beta$ -carotene, and CST-lupeol complex was 175.83 nm<sup>2</sup>, 180.89 nm<sup>2</sup>, 175.38 nm<sup>2</sup>, and 175.30 nm<sup>2</sup>, respectively (Figure 4B). Of the four hits, the SASA of CST-18alpha-glycyrrhetic acid, CST- $\beta$ -carotene, and CST-lupeol falls within the expected range of the free CST and CST-substrate complex, suggesting better compactness of the protein in the presence of these ligands with relatively lesser unwanted solvent accessibility. The SASA of the CST- $\alpha$ -carotene complex was close to the SASA of free CST, suggesting that ligand binding has a negligible impact on the solvent accessibility of the protein structure. Overall, the SASA of the protein-ligand complex of the selected compounds showed no major changes in the exposed protein structure after protein-ligand binding and maintained the natural structural integrity of the protein.

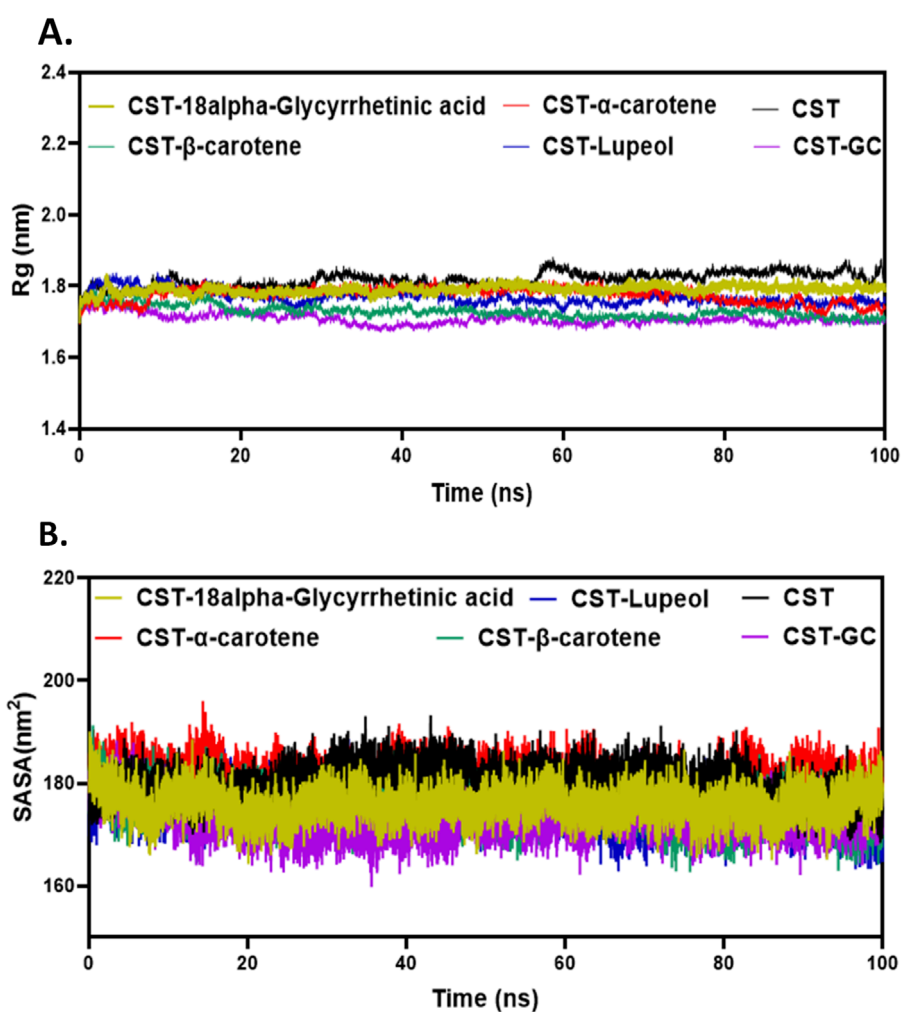
Furthermore, intramolecular hydrogen bonding analysis was vital for understanding the stability of the protein-ligand complex as

ligand binding impacts the overall protein intramolecular dynamics (Pace et al., 2014; Hubbard and Haider, 2010). Compared to the free CST protein, substrate binding to the substrate binding site in CST slightly increased the intramolecular hydrogen bonding, suggesting that when binding to the active site, the substrate pushed the protein structure inward, facilitating the proximity of atoms and, thus, facilitated additional contacts at the intramolecular level. Among the protein-ligand complexes of the selected compounds, the CST-18alpha-glycyrrhetic acid, CST- $\beta$ -carotene, and CST-lupeol complexes showed a negligible impact on the intramolecular hydrogen bond dynamics of the CST protein, suggesting the stability of protein intramolecular dynamics with ligand binding (Figures 5A, C, D). Probability distribution function (PDF) plots show good consistency of these three complexes. The CST- $\alpha$ -carotene complex was found to be the least stable and most deviated complex, and the protein structure in this complex even showed negatively more flexibility than that of the free CST (Figure 5B).

Thus, at the structural compactness level, the CST-18alpha-glycyrrhetic acid, CST- $\beta$ -carotene, and CST-lupeol complexes were found to be relatively more stable than the CST- $\alpha$ -carotene complex. In the presence of these three compounds, the CST protein underwent the expected natural stiffness and compactness phenomenon, as observed in the case of protein-substrate binding.

### 3.4.3 Principal component analysis

The specific catalytic action of each protein is executed through coordinated and collective atomic motion, which determines the stability of the protein (Zare et al., 2024; Yuan et al., 2024). PCA is widely used to analyze the atomic-level conformational changes through ligand binding. Basically, the principal components are the dominant mode of motion of the system that determine the structural and dynamic properties of the protein-ligand complex (Zare et al., 2024; Alrouji et al., 2023). In this study, principal components for the four selected CST-ligand complexes were largely determined by the first two (PC1 and PC2) and the first three eigenvectors (PC1, PC2, and PC3), which reflected the overall dynamics of the molecular subspace of the CST protein bound with the selected ligands (Figures 6A, B). With eigenvector 1 spanning between -3.3 and 4.5, eigenvector 2 between -3.07 and 3.09, and eigenvector 3 between -3.23 and 4.26, the CST-18alpha-glycyrrhetic acid complex showed better compactness among the four. In the CST- $\beta$ -carotene complex, eigenvector 1 covers -6.18 to 3.87, eigenvector 2 was between -3.3 and 2.7, and eigenvector 3 was between -1.7 and 3.4. With eigenvector 1 valued between -4.6 and 5.1 nm, eigenvector 2 from -7.3 to 4.1, and eigenvector 3 from -2.0 to 2.7 nm, the CST- $\alpha$ -carotene complex was found to be the most dispersed among the four, followed by the CST-lupeol complex, of which eigenvector 1 ranged between -4.3 and 5.9, eigenvector 2 from -2.6 to 3.8 nm, and eigenvector 3 from -3.6 to 4.6. In the first two eigenvectors, CST-18alpha-glycyrrhetic acid, CST- $\beta$ -carotene, and CST-lupeol complexes fell within the expected range of free CST and CST complexed with its substrate (GC) and showed relatively less motion and higher stability, while the CST- $\alpha$ -carotene complex was found to be the most diverted and less stable complex (Figure 6A). Subsequently, comparing the results of the first three principal components, we found that all four complexes primarily fall within the expected range of the free



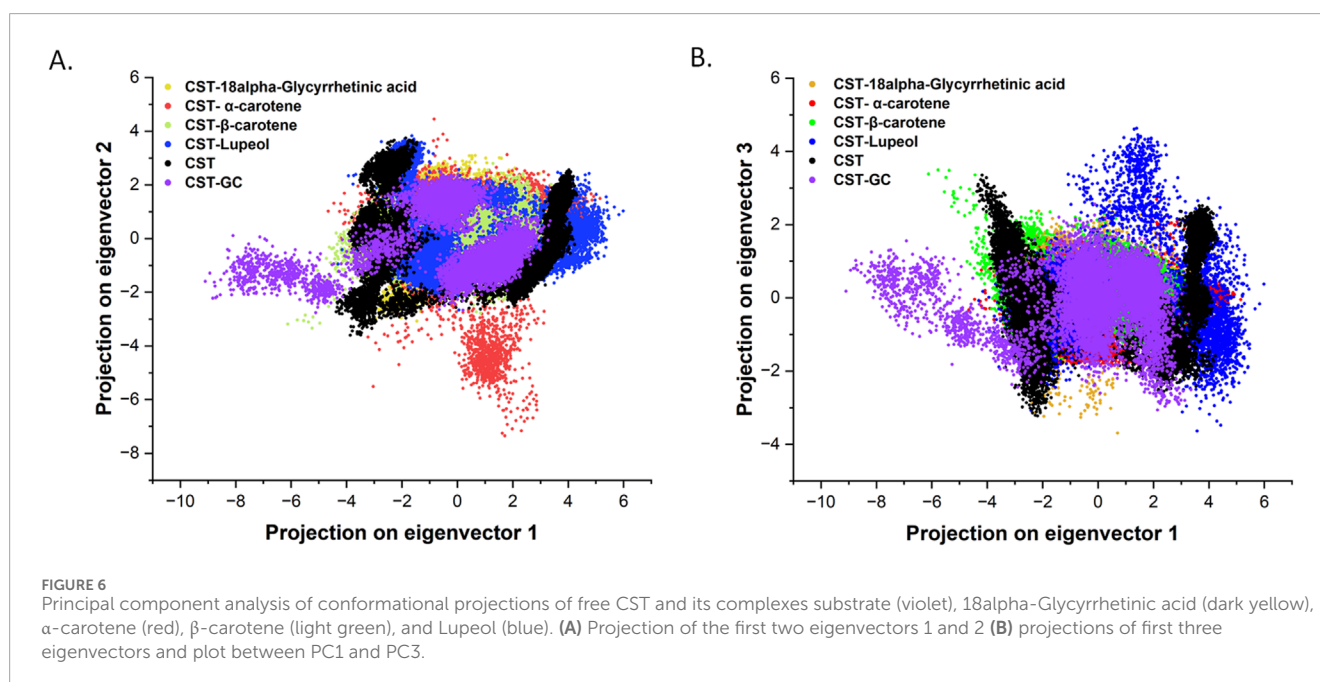
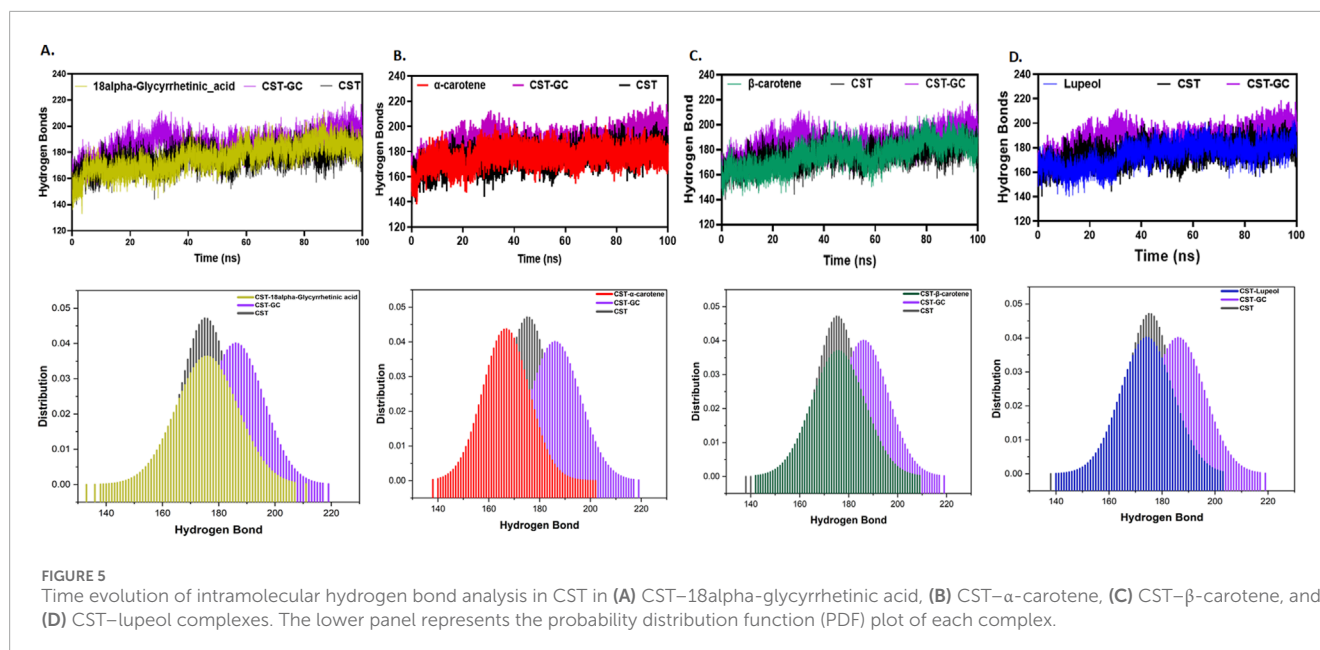
**FIGURE 4** Structural compactness analysis of CST in the CST–18alpha-glycyrrhetic acid, CST– $\alpha$ -carotene, CST– $\beta$ -carotene, and CST–lupeol complexes. Time evolution of (A) Rg values and (B) SASA values during the simulation.

CST and CST–substrate complex, with slight diversions observed in the CST–lupeol complex (Figure 6B). Thus, in both sets of eigenvectors, CST–18alpha-glycyrrhetic acid and CST– $\beta$ -carotene demonstrated relatively better stability because their principal components occupied a smaller subspace. Thus, these findings further strengthen the candidature of 18alpha-glycyrrhetic acid and  $\beta$ -carotene as inhibitors for CST.

### 3.4.4 Free energy landscape analysis

The free energy landscape (FEL) analysis is primarily applied to understand the folding pattern of proteins and the impact of ligand binding on the structure of the protein (Plattner and Noé, 2015; Lau and Roux, 2007; Khan et al., 2019; Maisuradze et al., 2010). In this study, the first two eigenvectors were used from the principal component analysis to generate the FEL plot of each protein–ligand complex based on the dominant and stable conformation during a simulation of 100 ns. In contrast to the free CST, which showed relatively increased conformational space with global minima of 16.60 kJ/mol for attaining a stable structure,

the conformational space of CST–GC complex slightly shifted with global minima of 17.10 kJ/mol (Singh and Singh, 2024a). In this study, in the spectrum of the free energy landscape, dark blue signified the most energetically favored region, and green depicted the moderately favored region, where the conformation of the protein could maintain its stability, while yellow represented the relatively unfavorable region and high energy state. According to the FEL plot, CST–18alpha-glycyrrhetic acid, CST– $\beta$ -carotene, and CST–lupeol complexes showed a wider blue zone with a global minima of 16.40 kJ/mol, 17.20 kJ/mol, and 15.10 kJ/mol, respectively (Figures 7A, C, D). The global minima of CST–18alpha-glycyrrhetic acid and CST– $\beta$ -carotene complexes were within the range of the global minima of the free CST (16.60 kJ/mol) and CST–GC complex (17.10 kJ/mol) (Singh and Singh, 2024a), thus showing potential for competitive inhibition in terms of energy requirement for binding. Among the four hits, the FEL of the CST– $\alpha$ -carotene complex with a global minima of 19.00 kJ/mol was found to have the least stable protein–ligand interaction (Figure 7B). Overall, the free energy landscape of the protein–ligand complexes provided a

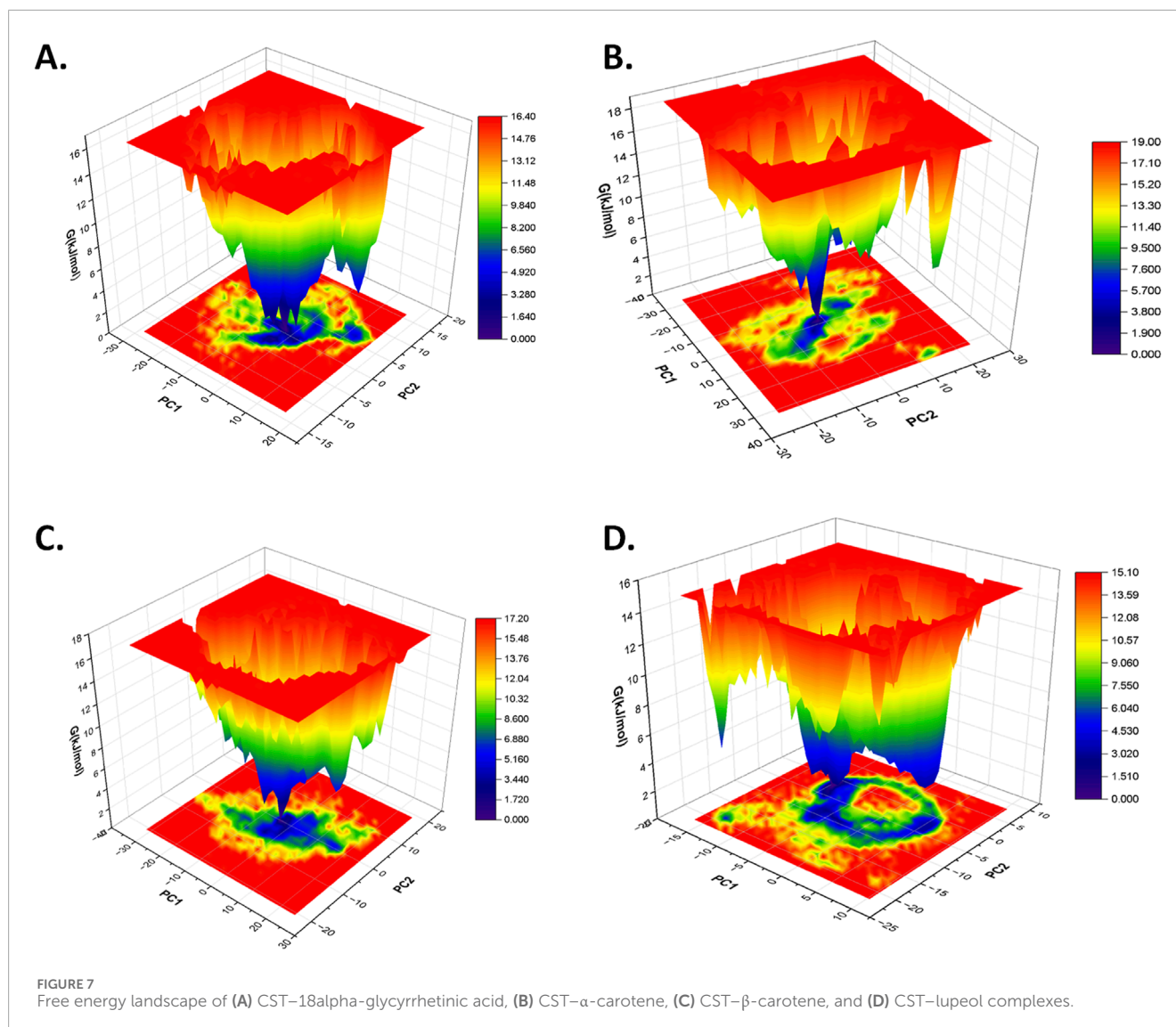


valuable insight into the interaction potential of the ligand and the stability of the protein in its bound form.

### 3.4.5 Dynamic cross-correlation matrix analysis

Using the MD trajectory data, a DCCM analysis was performed to understand the dominant correlation network of amino acid residues in the CST protein in the presence of ligands (Du et al., 2010; Avti et al., 2022). The dynamic correlation mapping of protein residues in the presence of different ligands indicated the impact of the ligand on the overall protein structure conformational stability. For the DCCM study of the four complexes, the MD simulation structure in the last 10 ns was considered to obtain an insight

into the protein–ligand complex. In the 2D matrix of DCCM analysis, the dark blue region represents the highly correlated motion of residues in the positive direction, the white region represents the highly anti-correlated motion of residues, and the no-color region indicates no correlation in the motion of the residue. Protein residues in the CST–18alpha-glycyrrhetic acid complex were considered highly correlated because they occupied a wider area of dark blue (Figure 8A). The protein residues in the CST– $\beta$ -carotene complex showed a relatively better residue correlation than that in the CST– $\alpha$ -carotene complex (Figures 8B, C). The residues in the CST–lupeol complex were found to have the least correlated residues as the dark blue region is relatively sparse

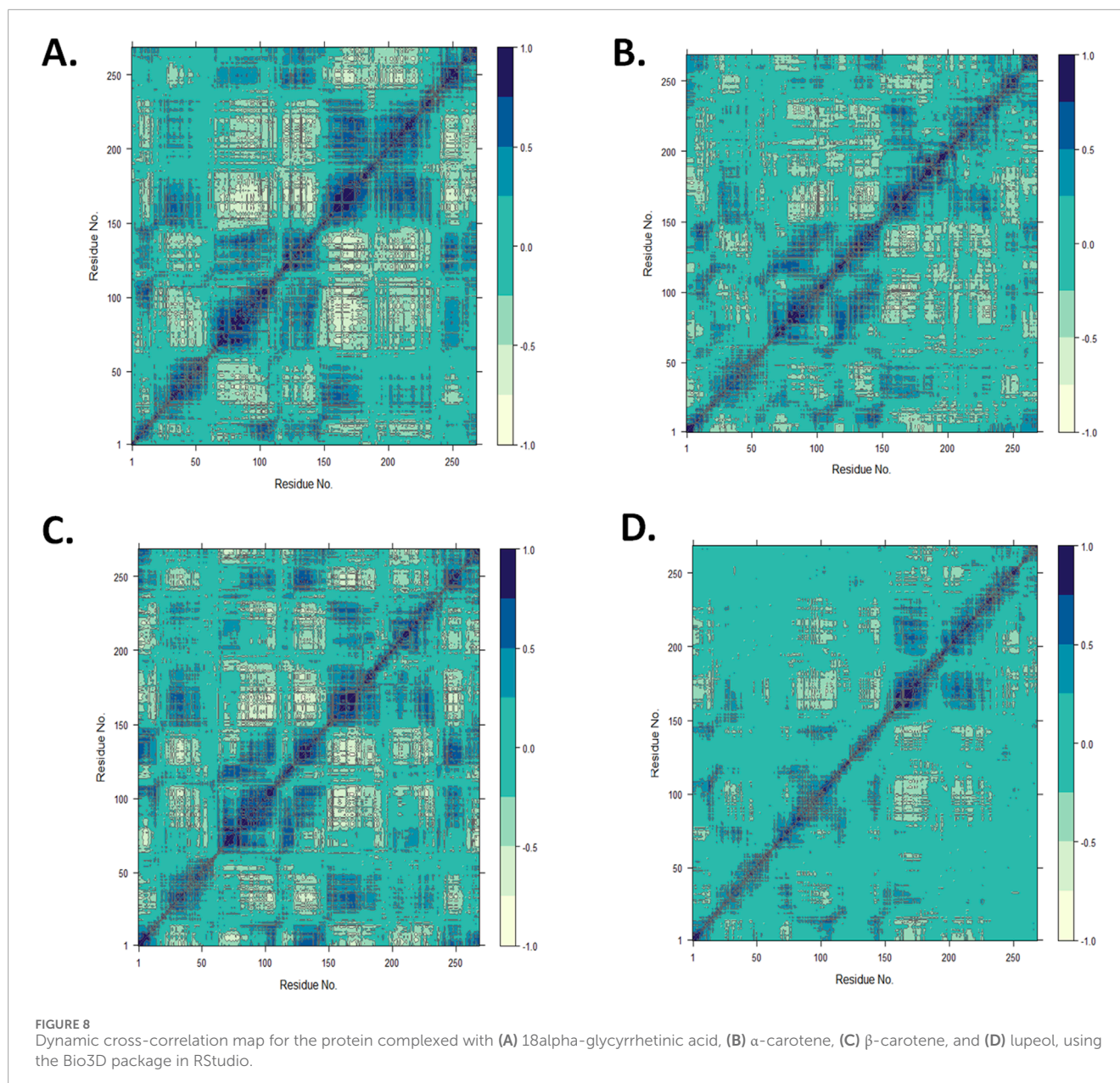


among the four protein–ligand complexes, and the large area in this complex is uncolored, indicating no significant correlation and reduced relation with the target protein (Figure 8D). Overall, the correlated motion of the residue analyzed in this study showed the quality of the protein–ligand complex. The DCCM analysis of residues bound to 18alpha-glycyrrhetic acid revealed the strongest correlation with the protein among all four complexes, thus indicating a strong protein–ligand interaction. In line with other MD trajectory analysis results,  $\beta$ -carotene also showed positive results in the DCCM analysis, showing a strong correlation with protein residues.

### 3.5 Post-MD protein–ligand interaction pose analysis

Post-simulation interaction analysis was an attempt to understand the binding pattern of the selected compounds under a dynamic aqueous environment, where the protein was free to

take its most stable state. Of the four hits, 18alpha-glycyrrhetic acid successfully retained its interaction with key residues, as well as its orientation and positioning in the binding pocket of the protein. In the CST–18alpha-glycyrrhetic acid complex, as in the docked conformation, C=O of the carboxylic group formed a hydrogen bond with Lys85 in the polar site. In the middle of the binding pocket, His84 and His141 flank the compound on both sides, and hydrogen bond formation took place with His84, while Tyr203 at the aromatic side interacted with the compound through two pi-alkyl bonds (Figure 9A). Additionally, Ser173, as in the docked conformation, retained van der Waals interaction with the C=O group of the third aromatic ring of the compound. Thus, the CST–18alpha-glycyrrhetic acid complex showed a strong protein–ligand interaction. During simulation,  $\beta$ -carotene also largely maintained a similar orientation in the binding pocket, with a slight change in the interaction pattern. Lys85 and Tyr176 in the left polar region, His84 and Phe177 in the middle, and Tyr203, Phe170, and Arg202 at the right end of the binding pocket determined the positioning of  $\beta$ -carotene in the binding pocket (Figure 9C). In



contrast to the similarity in the docked conformation of CST- $\alpha$ -carotene and CST- $\beta$ -carotene complexes,  $\alpha$ -carotene was found to be the least stable in the binding pocket during the simulation. The isomeric aromatic ring of  $\alpha$ -carotene shifted outward from the polar region away from Lys85, which was a key interacting residue in the docked conformation (Figure 9B). The simulation-led structural shift is shown in pre- and post-MD aligned complexes (Figure 10B). As in the docked conformation, the CST-lupeol complex maintained its position in the binding pocket, occupying mostly the right side of the binding pocket. Due to its relatively small size, lupeol left major room for ligand flexibility, which might not be suitable for competitive inhibition (Figure 9D). The distance of most of the pi-interactions was relatively short in both the CST-18alpha-glycyrrhetic acid and CST- $\beta$ -carotene-simulated complexes than their respective docked complexes, suggesting the strongest binding

potential of 18alpha-glycyrrhetic acid and the second strongest binding potential of  $\beta$ -carotene in the protein-binding pocket in the dynamic environment (Tables 3 and 4). Therefore, based on the interaction pattern analysis and alignment of pre- and post-MD complexes, 18alpha-glycyrrhetic acid and  $\beta$ -carotene were found to be suitable candidates for further reverse pharmacophore analysis.

### 3.6 Reverse pharmacophore mapping and cross-target identification

Reverse pharmacophore mapping for potential cross-target identification of the CST bioactive compounds (18alpha-glycyrrhetic acid and  $\beta$ -carotene) was performed through PharmMapper, which compared the pharmacophore of the active



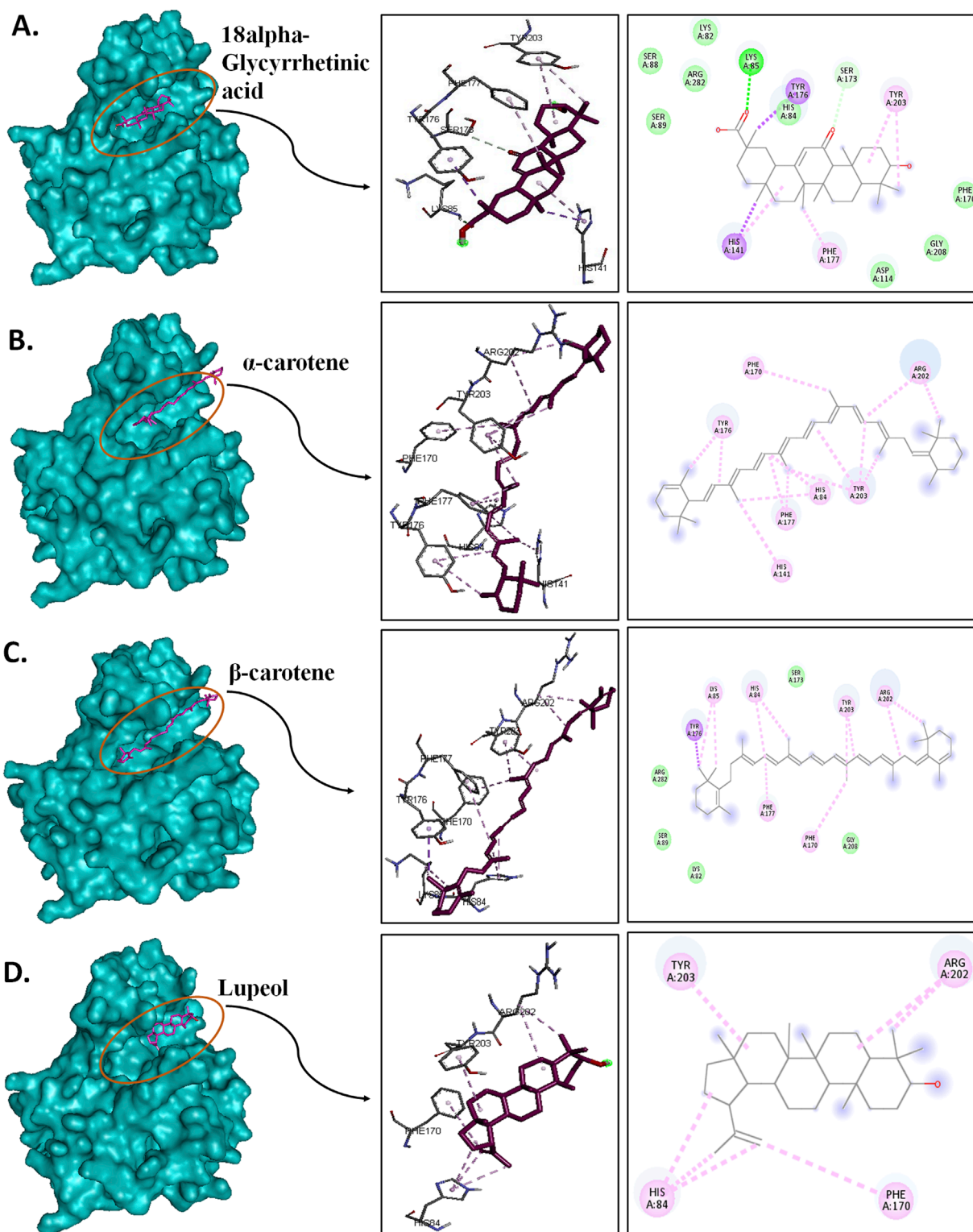


FIGURE 9

Post-MD interaction analysis of protein–ligand complexes with a surface view; 3D and 2D protein–ligand interaction of (A) CST–18alpha-glycyrrhetic acid, (B) CST– $\alpha$ -carotene, (C) CST– $\beta$ -carotene, and (D) CST–lupeol. Dark green, violet, light pink, and light green represent conventional hydrogen bonds, pi–sigma bonds, pi–alkyl bonds, and van der Waals interactions, respectively.

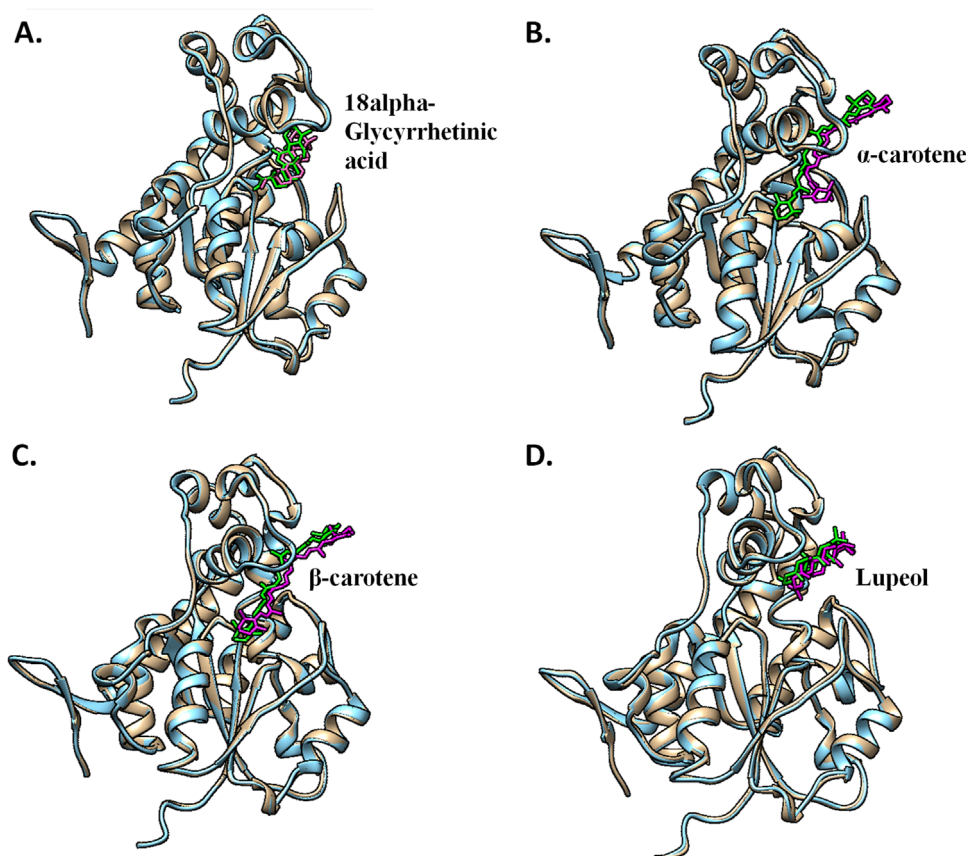


FIGURE 10

Pre- and post-MD structural alignment of (A) CST–18alpha-glycyrrhetic acid, (B) CST– $\alpha$ -carotene, (C) CST– $\beta$ -carotene, and (D) CST–lupeol complexes. Brown indicates the pre-MD CST complex with green ligands, and blue depicts the post-simulation protein complex with magenta ligands.

compounds with the pharmacophore models of 300 proteins deposited in the database (Wang et al., 2017; Liu et al., 2010). Based on their fitness score cutoff  $\geq 5.0$ , the key cross-targets were identified. Table 5 provides the details of the cross-target interactions of 18alpha-glycyrrhetic acid and  $\beta$ -carotene. The two potential targets of 18alpha-glycyrrhetic acid were corticosteroid 11-beta-dehydrogenase isozyme 1 and estradiol 17-beta-dehydrogenase 1, and positively, none of these are associated with the brain, while in the case of  $\beta$ -carotene, potential cross-targets are transthyretin, cellular retinoic acid-binding protein 2, and retinol-binding protein 4 (RBP4), among which transthyretin and RBP4 are associated with the brain.

## 4 Discussion

Metachromatic leukodystrophy is one of the critical neuropathological conditions that arise due to the accumulation of sulfatides over neurons, which leads to the development of white matter abnormalities and, thus, interrupt the proper processing of communication signals throughout the central and peripheral nervous systems. In the era of bioinformatics advancement, the option of developing an oral drug as a part of substrate reduction therapy is emerging as a new strategy in single-gene disorders

including MLD. To counter the major existing challenge in SRT, i.e., the lack of the availability of the three-dimensional structure of the target protein, *cerebroside sulfotransferase*, our group successfully developed a 3D homology model of the protein, which was used as a receptor for screening of compounds in this study. In the present study, the substrate-binding site of the protein was used for grid generation and protein–ligand interaction study. For this study, four potential herbs of Medhya Rasayana were considered for screening their phytoconstituents as these herbs are well-regarded as memory enhancers. Initially, based on a binding score cutoff of  $\leq -7.5$  kcal/mol, the top 15 compounds were selected for performing pharmacokinetics studies to eliminate false positives. Along with showing no significant toxicity under ADMET analysis, the four potential compounds—18alpha-glycyrrhetic acid,  $\alpha$ -carotene,  $\beta$ -carotene, and lupeol—showed high blood–brain barrier permeability, which was considered a critical parameter for the selection of inhibitors as the target’s physical location is the brain, and it was important to select those compounds that can cross the blood–brain barrier smoothly.

The interaction analysis of these four compounds in pre- and post-MD showed interesting outcomes, which became a basis for the identification of the most potent compounds for CST among the four compounds. In the pre-MD protein–ligand complex, the protein was rigid, and flexible ligands attempted to adjust in the active

TABLE 4 Non-bonded interaction in the simulated protein–ligand complex.

Sl. no.	Compound	Residue in contact	Bond type	Distance (Å)
1	18alpha-Glycyrrhetic acid	Lys85	Conventional hydrogen bond	1.90
		His84	Conventional hydrogen bond	3.50
		Tyr176	Pi-sigma	5.03
		His141	Pi-sigma	4.23
		Ser173	Van der Waals	4.24
		Phe177	Pi-alkyl	5.86
		Phe177	Pi-alkyl	5.30
		Tyr203	Pi-alkyl	5.81
		Tyr203	Pi-alkyl	7.31
2	$\alpha$ -Carotene	His84	Pi-alkyl	5.08
		His84	Pi-alkyl	6.02
		His141	Pi-alkyl	5.80
		Tyr176	Pi-alkyl	6.93
		Tyr176	Pi-Alkyl	5.93
		Phe170	Pi-alkyl	6.23
		Phe177	Pi-alkyl	4.98
		Phe177	Pi-alkyl	6.78
		Arg202	Pi-alkyl	4.01
		Tyr203	Pi-alkyl	4.80
		Tyr203	Pi-alkyl	5.19
3	$\beta$ -Carotene	Lys85	Pi-alkyl	4.21
		Lys85	Pi-alkyl	4.79
		His84	Pi-alkyl	4.72
		His84	Pi-alkyl	4.83
		Phe170	Pi-alkyl	5.91
		Tyr176	Pi-sigma	4.84
		Phe177	Pi-alkyl	6.53

(Continued on the following page)

TABLE 4 (Continued) Non-bonded interaction in the simulated protein–ligand complex.

Sl. no.	Compound	Residue in contact	Bond type	Distance (Å)
		Arg202	Pi-alkyl	3.84
		Arg202	Pi-alkyl	5.54
		Tyr203	Pi-alkyl	3.98
		Tyr203	Pi-alkyl	4.93
4	Lupeol	His84	Pi-alkyl	5.47
		His84	Pi-alkyl	5.68
		His84	Pi-alkyl	5.98
		Phe170	Pi-alkyl	5.67
		Tyr203	Pi-alkyl	6.29
		Arg202	Pi-alkyl	4.64
		Arg202	Pi-alkyl	7.69

site. However, during MD, the protein was subjected to an aqueous environment, which provided enough flexibility to the protein to adjust itself in the best possible way to provide maximum stability to its complex formation with the ligand. Among the four complexes, 18alpha-glycyrrhetic acid showed a better interaction with the CST active site while maintaining its position throughout the simulation (Figure 10A). The carboxylic acid end of the compound facilitated the orientation of compounds toward the polar region dominated by Lys85, Lys82, and Tyr176. The middle part of the compound was positioned exactly in the middle of the active site and flanked by His84 and His141 on both sides. Additionally, the oxygen atom at the middle of the compound facilitated its tight interaction in the active site. Phe170, Phe177, and Tyr203 were key residues in the aromatic site and interacted with the compound via a pi-alkyl interaction. In the pre-MD interaction study,  $\alpha$ -carotene and  $\beta$ -carotene showed a similar interaction pattern, with minor differences at their isomeric ends. However, after MD,  $\beta$ -carotene maintained its position in the active site (Figure 10C), while  $\alpha$ -carotene deviated from its initial position toward outside, as shown in Figure 10B. Based on active site occupancy criteria, lupeol, despite maintaining its position in the active site throughout the simulation, was regarded as less potent as it left enough space in the active site for other ligands to either cross-contaminate the active site or replace it if the other ligand is available in sufficient concentration (Figure 10D). Therefore, lupeol was considered less suitable to satisfy the competitiveness criteria.

The in-depth trajectory analysis of the selected compounds provided sufficient information for understanding the above structural changes. The RMSD of four ligands showed stability in terms of overall structural deviation within the stipulated

TABLE 5 Reverse pharmacophore mapping for identifying potential targets for the best-performing compounds.

Compounds	Proteins identified in PharmMapper	PDB ID	Disease	No. of pharmacophore features	Fitness score
18alpha-Glycyrrhetic acid	Corticosteroid 11-beta-dehydrogenase isozyme 1	2BEL	Polycystic ovary syndrome (PCOS)	10	5.383
	Estradiol 17-beta-dehydrogenase 1	1JTV	None	8	5.142
$\beta$ -Carotene	Transthyretin	1RLB	Amyloidosis type 1 (AMYL1)	10	5.937
	Cellular retinoic acid-binding protein 2	1CBS	None	10	5.517
	Retinol-binding protein 4	1RBP	Night vision problems	8	5.081

simulation time. The residual-level fluctuation was sharpened in the RMSF study, where the fluctuation was more pronounced in lupeol, followed by  $\alpha$ -carotene. The CST-18alpha-glycyrrhetic acid and CST- $\beta$ -carotene complexes showed better residual stability. Regarding structural compactness parameters including Rg and SASA, the protein in all four complexes largely maintained its structural compactness within the range of the free CST and CST-substrate complex, while at intramolecular hydrogen-bond formation-based structural compactness criteria, the protein in the CST- $\alpha$ -carotene complex showed least compactness among the four complexes. Furthermore, in PCA, the deviation of  $\alpha$ -carotene and lupeol was more pronounced, while CST-18alpha-glycyrrhetic acid and CST- $\beta$ -carotene were relatively better confined with occupying minimal subspace. A similar trend was also observed in the free energy landscape analysis; among the four, only 18alpha-glycyrrhetic acid and  $\beta$ -carotene showed global minima between the range of the free CST and CST-substrate complex, while  $\alpha$ -carotene was far above the range, and lupeol was found far below the range. In the dynamic cross-correlation map, the CST-lupeol complex was found to be least correlated. In the correlation matrix, the CST-18alpha-glycyrrhetic complex was found to be the best correlated complex among the four, followed by the CST- $\beta$ -carotene complex. Thus, the overall trajectory analysis represented the CST- $\alpha$ -carotene and CST-lupeol complexes as lesser stable complexes, hence excluded from further studies.

Reverse pharmacophore mapping was carried out to rule out the possibilities of cross-contamination with other nervous system proteins. With no potential neurodegenerative disease link target, 18alpha-glycyrrhetic acid can be considered the most potent and specific drug candidate against CST, while in the case of  $\beta$ -carotene, transthyretin and RBP4 may act as cross-targets in the brain (Pinheiro et al., 2022; Ishii et al., 2019). Cross-binding of  $\beta$ -carotene to transthyretin may cause  $\beta$ -amyloid deposition, leading to neuropathological conditions. In the central nervous system, RBP4 binds and inhibits transthyretin and, thus, regulates the breakdown of amyloid- $\beta$ , whereas the cross-interaction of  $\beta$ -carotene with RBP4 may reduce the  $\beta$ -amyloid deposition by restricting the interaction of RBP4 with transthyretin (Pinheiro et al., 2022). To understand the merit of the existing discrepancies with the cross-target interactions

of  $\beta$ -carotene with other proteins, a strong recommendation is made to further validate the result through experimental data.

Thus, the combinatorial high-throughput *in silico* screening procedures conducted in this study suggested the overall stability of 18alpha-glycyrrhetic acid and  $\beta$ -carotene inside the active site pocket of the CST protein. Although 18alpha-glycyrrhetic acid was found to be the best performer throughout the screening while maintaining the overall integrity of the protein conformation during complex formation in a dynamic simulation environment,  $\beta$ -carotene can also be considered for further studies as the second-best drug candidate as it also showed strong binding affinity in the binding site pocket of CST. Thus, the application of molecular docking, ADMET analysis, and simulation-led trajectory analysis provided a thorough understanding of the overall stability of the protein-ligand interaction. However, despite the multiple levels of stringent screening done in this study, there is no ruling out the possibility of failure of the proposed drug candidates at the experimental level. Therefore, the potency and specificity determined in this study need further validation through *in vitro* enzymatic, cell line studies and subsequent *in vivo* validation using animal models to complete the preclinical drug discovery validation process to obtain a potent drug candidate for the clinical trial of future marketed drugs. Additionally, since these drug candidates are originated from medicinal plants, they can also open the door for combinatorial substrate reduction therapy where potent drugs can combine with the hydroalcoholic extract of the whole plant to minimize any form of side effects of potent compounds while maintaining the potency via optimizing the dosage combination. Thus, the idea is to develop the safest possible oral medication for metachromatic leukodystrophy, which is one of the most fatal genetic diseases in children.

## 5 Conclusion

*Cerebroside sulfotransferase* is identified as an attractive target protein to develop substrate reduction therapy as a new therapeutic intervention in the field of metachromatic leukodystrophy. This study aimed to identify a potent and selective inhibitor for CST

inhibition using a multipronged virtual screening approach using a library of phytoconstituents of popular herbs of Medhya Rasayana. The screening of key neuroprotective herbs of Medhya Rasayana is an important strategy to identify potential neuroprotective inhibitors of CST. The phytoconstituents are natural compounds and are more adaptable to the natural system. Identifying potent drug candidates will further strengthen the potency of Ayurvedic herbs by applying the combinatorial strategy of treatment while minimizing any negative impact of the high concentration of drug candidates. In this study, 18 $\alpha$ -glycyrrhetic acid was identified as the most potent and specific inhibitor of CST, while  $\beta$ -carotene was found to be the second-most potent inhibitor of CST. Thus, the virtual screening approach applied in this study is a targeted approach that saved time and resource compared to the traditional approach of testing random chemicals through enzyme technology or applying a chemical synthesis approach for testing new drug entities as inhibitors. The MD simulation provided an insight into the binding pattern and orientation of compounds in the active site by providing knowledge about the most dominant mode of the motion responsible for determining the potency and selectivity of compounds. This information would be beneficial for designing new compounds based on the amino acid composition of the active site pocket. Thus, the findings of the present study narrow down the large pool of datasets and provide a direction for further *in vitro* and *in vivo* studies.

## Data availability statement

All data generated or analyzed during this study can be found in this published article and its supplementary file, and are also available from the corresponding author upon reasonable request.

## Author contributions

NS: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration,

## References

- Abrego-Guandique, D. M., Bonet, M. L., Caroleo, M. C., Cannataro, R., Tucci, P., Ribot, J., et al. (2023). The effect of beta-carotene on cognitive function: a systematic review. *Brain Sci.* 13 (10), 1468. doi:10.3390/BRAINS13101468
- Aier, I., Varadwaj, P. K., and Raj, U. (2016). Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Sci. Rep.* 6, 34984. doi:10.1038/SREP34984
- Ali, S., Hassan, Md., Islam, A., and Ahmad, F. (2014). A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. *Curr. Protein Pept. Sci.* 15 (5), 456–476. doi:10.2174/1389203715666140327114232
- Alrouji, M., Benjamin, L. S., Alhumaydhi, F. A., Al Abdulmonem, W., Baesa, S. S., Rehan, M., et al. (2023). Unlocking potential inhibitors for bruton's tyrosine kinase through in-silico drug repurposing strategies. *Sci. Rep.* 13 (1), 17684. doi:10.1038/S41598-023-44956-0
- Avti, P., Chauhan, A., Shekhar, N., Prajapat, M., Sarma, P., Kaur, H., et al. (2022). Computational basis of SARS-CoV 2 main protease inhibition: an insight from molecular dynamics simulation based findings. *J. Biomol. Struct. Dyn.* 40 (19), 8894–8904. doi:10.1080/07391102.2021.1922310
- Banerjee, P., Eckert, A. O., Schrey, A. K., and Preissner, R. (2018). ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* 46 (Web Server issue), W257–W263. doi:10.1093/NAR/GKY318
- Banerjee, S., Baghel, D., Pacheco de Oliveira, A., and Ghosh, A. (2023).  $\beta$ -Carotene, a potent amyloid aggregation inhibitor, promotes disordered  $\alpha\beta$  fibrillar structure. *Int. J. Mol. Sci.* 24 (6), 5175. doi:10.3390/ijms24065175
- Chang, S. C., Bergamasco, A., Bonnin, M., Bisonó, T. A., and Moride, Y. (2024). A systematic review on the birth prevalence of metachromatic leukodystrophy. *Orphanet J. Rare Dis.* 19 (1), 80. doi:10.1186/S13023-024-03044-W
- Daina, A., Michielin, O., and Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7 (1), 42717–42813. doi:10.1038/srep42717
- Du, Q. S., Wang, C. H., Liao, S. M., and Huang, R. B. (2010). Correlation analysis for protein evolutionary family based on amino acid position mutations and application in PDZ domain. *PLoS One* 5 (10), e13207. doi:10.1371/JOURNAL.PONE.0013207
- Durham, E., Dorr, B., Woetzel, N., Staritzbichler, R., and Meiler, J. (2009). Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *J. Mol. Model* 15 (9), 1093–1108. doi:10.1007/S00894-009-0454-9

resources, software, supervision, validation, visualization, writing—original draft, and writing—review and editing. AS: supervision, validation, and writing—review and editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Nivedita Singh acknowledges the Institute of Eminence, Banaras Hindu University, Government of India, for providing the fellowship and research grant for the ongoing project under the Malaviya Post-Doctoral Fellowship scheme with Grant ID R/Dev/G/6031/IOE/MPDFs/61698.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2024.1476482/full#supplementary-material>

- Eichler, F., Sevin, C., Barth, M., Pang, F., Howie, K., Walz, M., et al. (2022). Understanding caregiver descriptions of initial signs and symptoms to improve diagnosis of metachromatic leukodystrophy. *Orphanet J. Rare Dis.* 17 (1), 370. doi:10.1186/S13023-022-02518-Z
- Fernández-Pereira, C., Millán-Tejado, B. S., Gallardo-Gómez, M., Pérez-Márquez, T., Alves-Villar, M., Melcón-Crespo, C., et al. (2021). Therapeutic approaches in lysosomal storage diseases. *Biomolecules* 11, 1775. doi:10.3390/biom11121775
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* 11 (5), 905–919. doi:10.1038/NPROT.2016.051
- Fumagalli, F., Calbi, V., Natali Sora, M. G., Sessa, M., Baldoli, C., Rancoita, P. M. V., et al. (2022). Lentiviral haematopoietic stem-cell gene therapy for early-onset metachromatic leukodystrophy: long-term results from a non-randomised, open-label, phase 1/2 trial and expanded access. *Lancet* 399 (10322), 372–383. doi:10.1016/S0140-6736(21)02017-1
- Funari, R., Bhalla, N., and Gentile, L. (2022). Measuring the radius of gyration and intrinsic flexibility of viral proteins in buffer solution using small-angle X-ray scattering. *ACS Meas. Sci. au* 2 (6), 547–552. doi:10.1021/ACSMEASURES.2C00048
- Guterres, H., and Im, W. (2020). Improving protein-ligand docking results with high-throughput molecular dynamics simulations. *J. Chem. Inf. Model* 60 (4), 2189–2198. doi:10.1021/ACS.JCIM.0C00057
- Hemmerich, J., and Ecker, G. F. (2020). *In silico* toxicology: from structure–activity relationships towards deep learning and adverse outcome pathways. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 10 (4), e1475. doi:10.1002/WCMS.1475
- Hira, S., Saleem, U., Anwar, F., Sohail, M. F., Raza, Z., and Ahmad, B. (2019).  $\beta$ -Carotene: a natural compound improves cognitive impairment and oxidative stress in a mouse model of streptozotocin-induced alzheimer's disease. *Biomolecules* 9 (9), 441. doi:10.3390/Biom9090441
- Hollingsworth, S. A., and Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron* 99 (6), 1129–1143. doi:10.1016/J.NEURON.2018.08.011
- Hubbard, R. E., and Haider, M. K. (2010). Hydrogen bonds in proteins: role and strength. *Encycl. Life Sci.* doi:10.1002/9780470015902.A0003011.PUB2
- Ishii, M., Kamel, H., and Iadecola, C. (2019). Retinol binding protein 4 levels are not altered in preclinical alzheimer's disease and not associated with cognitive decline or incident dementia. *J. Alzheimers Dis.* 67 (1), 257–263. doi:10.3233/JAD-180682
- Istaiti, M., Becker-Cohen, M., Dinur, T., Revel-Vilk, S., and Zimran, A. (2022). Real-life experience with oral Eliglustat in patients with gaucher disease previously treated with enzyme replacement therapy. *J. Clin. Med.* 11 (21), 6265. doi:10.3390/jcm11216265
- Khan, M. T., Khan, A., Rehman, A. U., Wang, Y., Akhtar, K., Malik, S. I., et al. (2019). Structural and free energy landscape of novel mutations in ribosomal protein S1 (RpsA) associated with pyrazinamide resistance. *Sci. Rep.* 9 (1), 7482–7512. doi:10.1038/s41598-019-44013-9
- Khan, S., Farooq, U., and Kurnikova, M. (2016). Exploring protein stability by comparative molecular dynamics simulations of homologous hyperthermophilic, mesophilic, and psychrophilic proteins. *J. Chem. Inf. Model* 56 (11), 2129–2139. doi:10.1021/ACS.JCIM.6B00305
- Komada, N., Fujiwara, T., Yoshizumi, H., Ida, H., and Shimoda, K. (2021). A Japanese patient with gaucher disease treated with the oral drug Eliglustat as substrate reducing therapy. *Case Rep. Gastroenterol.* 15 (3), 838–845. doi:10.1159/000519005
- Kulkarni, R., Girish, K. J., and Kumar, A. (2012). Nootropic herbs (Medhya Rasayana) in Ayurveda: an update. *Pharmacogn. Rev.* 6 (12), 147–153. doi:10.4103/0973-7847.99949
- Kurniawan, J., and Ishida, T. (2022). Protein model quality estimation using molecular dynamics simulation. *ACS Omega* 7 (28), 24274–24281. doi:10.1021/acsomega.2c01475
- Kurtzberg, J. (2022). Gene therapy offers new hope for children with metachromatic leukodystrophy. *Lancet* 399 (10322), 338–339. doi:10.1016/S0140-6736(22)00057-5
- Lamichhane, A., and Cabrero, F. R. (2023). *Metachromatic leukodystrophy. Frontiers in lysosomal storage diseases (LSD) treatments*, 153–162. doi:10.1177/0883073816656401
- Lau, A. Y., and Roux, B. (2007). The free energy landscapes governing conformational changes in a glutamate receptor ligand-binding domain. *Structure* 15 (10), 1203–1214. doi:10.1016/J.STR.2007.07.015
- Li, W., Guillaume, J., Baqi, Y., Wachsmann, I., Gieselmann, V., Van Calenbergh, S., et al. (2020). Synthesis and structure-activity relationships of cerebroside analogues as substrates of cerebroside sulphotransferase and discovery of a competitive inhibitor. *J. Enzyme Inhib. Med. Chem.* 35 (1), 1503–1512. doi:10.1080/14756366.2020.1791841
- Liu, X., Ouyang, S., Yu, B., Liu, Y., Huang, K., Gong, J., et al. (2010). PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. *Nucleic Acids Res.* 38 (Web Server issue), W609–W614. doi:10.1093/NAR/GKQ300
- Lobanov, M. Y., Bogatyreva, N. S., and Galzitskaya, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. *Mol. Biol.* 42 (4), 623–628. doi:10.1134/s0026893308040195
- Maisuradze, G. G., Liwo, A., and Scheraga, H. A. (2010). Relation between free energy landscapes of proteins and dynamics. *J. Chem. Theory Comput.* 6 (2), 583–595. doi:10.1021/CT9005745
- Martínez, L. (2015). Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS One* 10 (3), e0119264. doi:10.1371/JOURNAL.PONE.0119264
- Mistry, P. K., Kishnani, P. S., Balwani, M., Charrow, J. M., Hull, J., Weinreb, N. J., et al. (2023). The Two Substrate Reduction Therapies for Type 1 Gaucher Disease Are Not Equivalent. Comment on Hughes et al. Switching between Enzyme Replacement Therapies and Substrate Reduction Therapies in Patients with Gaucher Disease: Data from the Gaucher Outcome Survey (GOS). *J. Clin. Med.* 2022, 11, 5158. *J. Clin. Med.* 12 (9), 3269. doi:10.3390/JCM12093269
- Miyake, N., Miyake, K., Sakai, A., Yamamoto, M., Suzuki, H., and Shimada, T. (2021). Treatment of adult metachromatic leukodystrophy model mice using intrathecal administration of type 9 AAV vector encoding arylsulfatase A. *Sci. Rep.* 11 (1), 20513. doi:10.1038/S41598-021-99979-2
- Mohanraj, K., Karthikeyan, B. S., Vivek-Ananth, R. P., Chand, R. P. B., Aparna, S. R., Mangalapati, P., et al. (2018). IMPPAT: a curated database of Indian medicinal plants, Phytochemistry and therapeutics. *Sci. Rep.* 8 (1), 4329–4417. doi:10.1038/s41598-018-22631-z
- Narisawa, T., Fukaura, Y., Hasebe, M., Ito, M., Aizawa, R., Murakoshi, M., et al. (1996). Inhibitory effects of natural carotenoids, alpha-carotene, beta-carotene, lycopene and lutein, on colonic aberrant crypt foci formation in rats. *Cancer Lett.* 107 (1), 137–142. doi:10.1016/0304-3835(96)04354-6
- Pace, C. N., Fu, H., Fryar, K. L., Landua, J., Trevino, S. R., Schell, D., et al. (2014). Contribution of hydrogen bonds to protein stability. *Protein Sci.* 23 (5), 652–661. doi:10.1002/PRO.2449
- Paul, S. K., Saddam, M., Rahaman, K. A., Choi, J. G., Lee, S. S., and Hasan, M. (2022). Molecular modeling, molecular dynamics simulation, and essential dynamics analysis of grancalcin: an upregulated biomarker in experimental autoimmune encephalomyelitis mice. *Heliyon* 8 (10), e11232. doi:10.1016/J.HELIYON.2022.E11232
- Penati, R., Fumagalli, F., Calbi, V., Bernardo, M. E., and Aiuti, A. (2017). Gene therapy for lysosomal storage disorders: recent advances for metachromatic leukodystrophy and mucopolysaccharidosis I. *J. Inherit. Metab. Dis.* 40 (4), 543–554. doi:10.1007/S10545-017-0052-4
- Pinheiro, F., Pallarès, I., Peccati, F., Sánchez-Morales, A., Varejão, N., Bezerra, F., et al. (2022). Development of a highly potent transthyretin amyloidogenesis inhibitor: design, synthesis, and evaluation. *J. Med. Chem.* 65 (21), 14673–14691. doi:10.1021/acs.jmedchem.2c01195
- Pires, D. E. V., Blundell, T. L., and Ascher, D. B. (2015). pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* 58 (9), 4066–4072. doi:10.1021/acs.jmedchem.5b00104
- Pirollo, D., Righino, B., Camponeschi, C., Ria, F., Di Sante, G., and De Rosa, M. C. (2023). Virtual screening and molecular dynamics simulations provide insight into repurposing drugs against SARS-CoV-2 variants spike protein/ACE2 interface. *Sci. Rep.* 13 (1), 1494. doi:10.1038/S41598-023-28716-8
- Plattner, N., and Noé, F. (2015). Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and markov models. *Nat. Commun.* 6 (1), 7653–7710. doi:10.1038/ncomms8653
- Raies, A. B., and Bajic, V. B. (2016). *In silico* toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 6 (2), 147–172. doi:10.1002/WCMS.1240
- Rampogu, S., Lee, G., Park, J. S., Lee, K. W., and Kim, M. O. (2022). Molecular docking and molecular dynamics simulations discover curcumin analogue as a plausible dual inhibitor for SARS-CoV-2. *Int. J. Mol. Sci.* 23 (3), 1771. doi:10.3390/IJMS23031771
- Rashmi, P., Anubha, C., and Gaurav, S. (2017). Concept of Medhya Rasayana in Ayurveda: an overview neelam. *Int. J. Res. Ayurveda Pharm.* 8 (2). doi:10.7897/2277-4343.08287
- Sarokte, A. S., and Rao, M. V. (2013). Effects of Medhya Rasayana and yogic practices in improvement of short-term memory among school-going children. *Ayu* 34 (4), 383–389. doi:10.4103/0974-8520.127720
- Savojarjo, C., Manfredi, M., Martelli, P. L., and Casadio, R. (2021). Solvent accessibility of residues undergoing pathogenic variations in humans: from protein structures to protein sequences. *Front. Mol. Biosci.* 7, 626363. doi:10.3389/fmolb.2020.626363
- Sessa, M., Lorioli, L., Fumagalli, F., Acquati, S., Redaelli, D., Baldoli, C., et al. (2016). Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *Lancet* 388 (10043), 476–487. doi:10.1016/S0140-6736(16)30374-9
- Sevin, C., and Deiva, K. (2021). Clinical trials for gene therapy in lysosomal diseases with CNS involvement. *Front. Mol. Biosci.* 8, 624988. doi:10.3389/FMOLB.2021.624988
- Shaimardanova, A. A., Chulpanova, D. S., Solovyeva, V. V., Mullagulova, A. I., Kitaeva, K. V., Allegrucci, C., et al. (2020). Metachromatic leukodystrophy: diagnosis, modeling, and treatment approaches. *Front. Med. (Lausanne)* 7, 576221. doi:10.3389/FMED.2020.576221

- Singh, N., Malik, A. H., Iyer, P. K., and Patra, S. (2021). Diversifying the xanthine scaffold for potential phosphodiesterase 9A inhibitors: synthesis and validation. *Med. Chem. Res.* 30 (6), 1199–1219. doi:10.1007/s00044-021-02722-9
- Singh, N., Patra, S., and Patra, S. (2018). Identification of xanthine derivatives as inhibitors of phosphodiesterase 9A through *in silico* and biological studies. *Comb. Chem. High. Throughput Screen* 21 (7), 476–486. doi:10.2174/1386207321666180821100713
- Singh, N., Saravanan, P., Thakur, M. S., and Patra, S. (2017). Development of xanthine based inhibitors targeting phosphodiesterase 9A. *Lett. Drug Des. Discov.* 14 (10), 1122–1137. doi:10.2174/1570180813666161102125423
- Singh, N., and Singh, A. K. (2024a). *In silico* structural modeling and binding site analysis of *cerebroside sulfotransferase* (CST): a therapeutic target for developing substrate reduction therapy for metachromatic leukodystrophy. *ACS Omega* 9, 10748–10768. doi:10.1021/acsomega.3c09462
- Singh, N., and Singh, A. K. (2024b). A comprehensive review on structural and therapeutical insight of *cerebroside sulfotransferase* (CST) - an important target for development of substrate reduction therapy against metachromatic leukodystrophy. *Int. J. Biol. Macromol.* 258, 128780. doi:10.1016/j.ijbiomac.2023.128780
- Stielow, M., Witczyńska, A., Kubryń, N., Fijałkowski, Ł., Nowaczyk, J., and Nowaczyk, A. (2023). The bioavailability of drugs—the current state of knowledge. *Molecules* 28 (24), 8038. doi:10.3390/MOLECULES28248038
- Vivek-Ananth, R. P., Mohanraj, K., Sahoo, A. K., and Samal, A. (2023). IMPPAT 2.0: an enhanced and expanded phytochemical atlas of Indian medicinal plants. *ACS Omega* 8 (9), 8827–8845. doi:10.1021/acsomega.3c00156
- Wang, X., Shen, Y., Wang, S., Li, S., Zhang, W., Liu, X., et al. (2017). PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.* 45 (W1), W356–W360. doi:10.1093/NAR/GKX374
- Wright, S. H. (2019). Molecular and cellular physiology of organic cation transporter 2. *Am. J. Physiol. Ren. Physiol.* 317 (6), F1669–F1679. doi:10.1152/AJPRENAL.00422.2019
- Yaghootfam, A., Sorkalla, T., Häberlein, H., Gieselmann, V., Kappler, J., and Eckhardt, M. (2007). *Cerebroside sulfotransferase* forms homodimers in living cells. *Biochemistry* 46 (32), 9260–9269. doi:10.1021/BI700014Q
- Yang, H., Sun, L., Li, W., Liu, G., and Tang, Y. (2018). Corrigendum: *in silico* prediction of chemical toxicity for drug design using machine learning methods and structural alerts. *Front. Chem.* 6, 129. doi:10.3389/fchem.2018.00129
- Yasuo, N., and Sekijima, M. (2019). Improved method of structure-based virtual screening via interaction-energy-based learning. *J. Chem. Inf. Model* 59 (3), 1050–1061. doi:10.1021/ACS.JCIM.8B00673
- Yuan, F., Li, T., Xu, X., Chen, T., and Cao, Z. (2024). Identification of novel PI3Ka inhibitor against gastric cancer: QSAR-molecular docking-and molecular dynamics simulation-based analysis. *Appl. Biochem. Biotechnol.* doi:10.1007/S12010-024-04898-3
- Zare, F., Ataollahi, E., Mardaneh, P., Sakhteman, A., Keshavarz, V., Solhjo, A., et al. (2024). A combination of virtual screening, molecular dynamics simulation, MM/PBSA, ADMET, and DFT calculations to identify a potential DPP4 inhibitor. *Sci. Rep.* 14 (1), 7749. doi:10.1038/S41598-024-58485-X
- Zhang, Z., and Tang, W. (2018). Drug metabolism in drug discovery and development. *Acta Pharm. Sin. B* 8 (5), 721–732. doi:10.1016/J.APSB.2018.04.003
- Zhao, M., Ma, J., Li, M., Zhang, Y., Jiang, B., Zhao, X., et al. (2021). Cytochrome P450 enzymes and drug metabolism in humans. *Int. J. Mol. Sci.* 22 (23), 12808. doi:10.3390/IJMS222312808
- Zou, W., Shi, B., Zeng, T., Zhang, Y., Huang, B., Ouyang, B., et al. (2021). Drug transporters in the kidney: perspectives on species differences, disease status, and molecular docking. *Front. Pharmacol.* 12, 746208. doi:10.3389/FPHAR.2021.746208