Journal of Chemistry

WILEY

RESEARCH ARTICLE OPEN ACCESS

Discovery of Triacylglycerol-Binding Proteins in *Cereus jamacaru* DC. Through Integrative In Silico Methods

M. I. O. Cardoso | J. Oliveira | V. Bassaneze | D. T. Amaral

Center for Natural and Human Sciences, Federal University of ABC (UFABC), Santo André, São Paulo, Brazil

Correspondence: D. T. Amaral (danilo.trabuco@ufabc.edu.br)

Received: 2 June 2025 | **Revised:** 14 November 2025 | **Accepted:** 7 January 2026

Academic Editor: Sohini Basu Roy

Keywords: bioactive proteins | bioinformatics | cardiovascular diseases | mandacaru | triacylglycerol

ABSTRACT

The Caatinga biome in Brazil harbors unique bioactive plants, such as *Cereus jamacaru* DC. Historically, it has been used in traditional medicine for treating kidney issues, diabetes, and cardiovascular conditions; however, these ethnobotanical reports do not constitute evidence of clinical efficacy. Here, we integrate transcriptomics, molecular docking, and molecular dynamics approaches to explore the potential of the predicted proteome of *C. jamacaru* to bind to triacylglycerol molecules, particularly targeting triacylglycerol formed by lauric, myristic, and palmitic acids. Transcriptome analysis identified 128,942 transcripts, with 14,739 homologous proteins screened for binding affinities. Molecular docking highlighted an isoform of *Banyan Peroxidase* as a versatile candidate, exhibiting strong binding energies across all triacylglycerols, particularly palmitic acid (−7.63 kcal/mol). *Xyloglucan Endotransglycosylase* demonstrated specificity for myristic acid (−7.75 kcal/mol), while *Nonspecific Lipid Transfer Protein* showed exceptional structural stability in dynamic simulations. The molecular dynamics simulations revealed key insights into protein stability and ligand interactions. *Banyan Peroxidase* displayed moderate flexibility, enhancing its adaptability to diverse triacylglycerol substrates. Conversely, *Xyloglucan Endotransglycosylase* exhibited compact stability, making it a strong candidate for future bioengineering efforts. These findings provide a computational starting point for future experimental validation of lipid-binding plant proteins as candidates for biotechnological applications. While promising, these results are preliminary and do not imply direct therapeutic relevance.

1 | Introduction

Brazil has a rich floral diversity, with the Caatinga biome being particularly notable for its variety of plant species adapted to water scarcity, such as cacti. This region provides valuable resources for local communities, especially in the food and herbal medicine sectors, due to the nutritional and bioactive potential of regional plants [1]. Among these species, the genus *Cereus* Mill. (Cereeae; Cactaceae), particularly *Cereus jamacaru* DC., known as mandacaru, stands out. Native to Brazil's semi-arid region, this plant is rich in macromolecules, with its spines and stems containing unsaturated fatty acids, such as oleic and linoleic acids, as well as saturated fatty acids, such as palmitic and stearic acids [2]. For decades, the root and stem of mandacaru have been used in infusions and teas by the populations of northeastern Brazil to treat a broader range of conditions, including diuresis, kidney issues, diabetes, and worm infections. It has also been employed as a laxative, antihypertensive, and antirheumatic agent [3, 4]. While *C. jamacaru* has a long record of traditional uses, these accounts are hypothesis-generating and do not demonstrate causality or clinical benefit. In this study, we therefore treat any potential "therapeutic" relevance as speculative and contingent upon future biochemical and in vivo validation. Our goal is to map in silico protein–triacylglycerol

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Copyright © 2026 M. I. O. Cardoso et al. *Journal of Chemistry* published by John Wiley & Sons Ltd.

interaction potential as a first step in bioprospecting, not to claim efficacy or mechanism in humans.

With advancements in omics sciences, research on the medicinal properties and macromolecular composition of natural sources, such as mandacaru, has increased [5]. The bioactive compounds found in plants form the basis for discovering pharmaceuticals and industrial enzymes through innovative techniques. This aligns with traditional practices of plant prospecting, which have long been integral to cultures worldwide, observing their medicinal, cosmetic, and culinary properties. It is estimated that 80% of the global population relies on traditional medicine, with 85% of this dependence rooted in the use of plants and their bioactive compounds [6, 7]. Mandacaru, with its diverse applications and bioactive potential, exemplifies the intersection of traditional knowledge and modern scientific exploration, reinforcing the importance of its study for both local and global health solutions. Similar in silico approaches have been employed to identify lipid-interacting proteins from medicinal plants or microbial sources, contributing to the understanding of molecular scaffolds with potential biotechnological applications (e.g., [8, 9]). However, no prior studies have explored such interactions in *C. jamacaru* or related cacti.

Recently, approaches to facilitate the bioprospecting of biomolecules have relied on computer-assisted (in silico) strategies, such as molecular docking, due to their advantages of quick access and low cost, allowing for large-scale application [10, 11]. Inverse virtual screening (IVS), which is based on molecular docking methods, involves predicting the binding mode and affinity of a ligand with a receptor. During the IVS process, molecular docking is used as a search step within a protein database. In this context, computational methods for protein modeling and molecular docking are highlighted. Both methods aim to predict protein structures and energetically assess the affinity between the protein and its substrate [12].

Fatty acids play a critical role in cardiovascular health, particularly long-chain saturated fatty acids, such as myristic and palmitic acids, which are associated with elevated low-density lipoprotein cholesterol (LDL-c) levels and increased cardiovascular risk. This risk arises through mechanisms, such as cholesterol esterification [13]. A significant association exists between reducing LDL-c levels and lower rates of cardiovascular events, including cardiovascular death, acute myocardial infarction (MI), coronary revascularization, and stroke [14]. LDL-c is generated via the chylomicron pathway [15]. This pathway begins with the digestion of triacylglycerols by gastric lipase in the stomach and pancreatic lipase in the intestine, followed by the absorption of fatty acids and the subsequent production of chylomicrons by enterocytes [15]. Lipase, an enzyme central to this process, is targeted by pharmacological agents, such as orlistat, which binds to the enzyme's active site [16]. This strategy promotes the elimination of triacylglycerols through the intestinal tract before digestion, thereby reducing cardiovascular risk. However, this approach indiscriminately eliminates all types of triacylglycerols, including beneficial unsaturated ones. Currently, there is no widely available or effective strategy that directly and specifically targets dietary triacylglycerols for human health without indiscriminate elimination.

Recent epidemiological and genetic studies have established that triacylglycerols are among the primary contributors to atherosclerotic cardiovascular disease (ASCVD; [17]). Obesity is a key underlying risk factor for ASCVD, as it facilitates other risk factors, such as hypertriglyceridemia and insulin resistance, both of which can trigger ASCVD. This constellation of factors constitutes metabolic syndrome, also referred to as metabolic risk factors [18]. The bioactive composition of *C. jamacaru* positions this plant as a promising source for therapeutic proteins capable of modulating triacylglycerol digestion. This unique biochemical profile, coupled with its historical use in traditional medicine, underscores its potential as a source of biomolecules for future biotechnological development aimed at addressing contemporary health challenges.

Although no protein-level or mechanistic assays have yet demonstrated lipid-modulating activity for *C. jamacaru*, several phytochemical and pharmacological studies report indirect evidence of metabolic and antioxidant effects that could influence lipid homeostasis. Ethanolic and aqueous extracts of *C. jamacaru* have shown hypoglycemic, anti-inflammatory, and hepatoprotective actions in rodent models [3, 4], and other members of the Cactaceae family exhibit lipid-lowering or hypocholesterolemic properties [19, 20]. These findings suggest that *C. jamacaru* may contain metabolites capable of modulating lipid pathways, providing indirect support for exploring its proteins in silico for triacylglycerol-binding potential.

Unlike previous studies that focus primarily on well-known medicinal plants or extensively studied enzymatic systems, this research uniquely explores the proteomic landscape of *C. jamacaru*, an underrepresented species from the Caatinga biome, to identify novel proteins with potential applications in triglyceride metabolism. By integrating transcriptomic data, molecular docking, and molecular dynamics simulations, this study provides the first in-depth computational characterization of *C. jamacaru* proteins in the context of cardiovascular health. The identification of proteins with significant triacylglycerol-binding affinities expands the current understanding of plant-derived bioactive molecules and opens new avenues for the identification of novel molecular targets with potential relevance in lipid metabolism. This approach bridges the gap between traditional ethnobotanical knowledge and cutting-edge computational strategies, reinforcing the importance of bioinformatics in modern drug discovery.

2 | Materials and Methods

2.1 | Data Collection and Transcriptome Assembly

The transcriptomes analyzed here correspond to roots and epidermis, which were selected for both biological and practical reasons. The epidermis plays a central role in lipid metabolism and protective barrier formation in plants, including the biosynthesis of cuticular waxes and lipid-based signaling molecules. Root tissues, however, are key sites of primary and secondary metabolism and have been widely used in ethnopharmacological preparations of *C. jamacaru* [3, 4]. Moreover, these were the tissues for which high-quality RNA-Seq data were already available [21]. Their combined analysis, therefore, provides complementary perspectives on lipid-associated gene expression across metabolically active and protective tissues.

The roots and epidermis were first washed with distilled water and 90% ethanol, dried on paper towels, and stored at −80°C. Approximately 50–90 mg of each tissue was ground under dry ice

using a mortar and pestle, and RNA was isolated using the RNeasy Plant Mini Kit (Qiagen) with RLC buffer. The pooled RNA sample (root + epidermis) was sent to BGI Americas (San Jose, CA, USA), where RNA integrity and concentration were assessed with an Agilent 2200 Bioanalyzer, yielding an RNA Integrity Number (RIN) of 7.8. Library preparation and sequencing were performed using the DNB-Seq platform (2×150 bp paired-end reads). *De novo* transcriptome assembly was carried out using the Trinity software (Version 2.5.11; [22]) and its associated packages. After assembly, transcripts were translated into amino acid sequences using the TransDecoder program (for more details, see [21]).

2.2 | IVS

The IVS analysis was performed using an automated workflow, BioProtIS pipeline ([23]; available at https://github.com/BBMDO/BioProtIS), developed in our laboratory. BioProtIS streamlines the entire workflow from sequence translation to docking analysis, allowing high-throughput screening of thousands of proteins from transcriptomic datasets. It integrates structural modeling, ligand preparation, and interaction profiling into a single automated framework, reducing manual steps and bias. Designed for plant and environmental "omics" applications, BioProtIS has been validated for scalability and reproducibility [23], enabling the efficient processing of nearly 15,000 *C. jamacaru* proteins in this study. This pipeline employs a set of Python programs to integrate "omics" data and stands out in the analysis of plant transcriptomics, detailed above:

2.2.1 | Three-Dimensional Protein Structure Modeling

The modeling process of the three-dimensional protein structures was carried out using MODELLER v.10.4 [24]. Subsequently, the identification of known three-dimensional structures in the PDB to be used as templates was performed. For this purpose, the predicted protein sequences from the transcriptome were compared to the protein sequences available in the PDB using the locally Blastp program. Proteins from *Cereus* with less than 50% homology to the template were promptly excluded, as low identity can also result in low-quality modeled protein structures. After identifying the model proteins in the PDB, the *batch_download.sh* script (available at: https://www.rcsb.org/scripts/batch_download.sh), developed by the PDB itself, was applied for the automated download of tertiary structures, resulting in 3,847 proteins. Using the protein sequences of *C. jamacaru* and the tertiary structure of the model sequence, the MODELLER program and its Python scripts were employed to perform sequence alignment between the target protein and the selected homologous protein. Based on this alignment, the initial model of the target protein was generated using the known structure of the homologous protein as a reference.

2.2.2 | Preparation of Inputs for Molecular Docking

After obtaining the models, we used the Open Babel 3.1.1 software [25] to prepare the protein structure. This included removing hydrogens, adding atomic types, establishing interactions and bonds, and centralizing the molecule to define the search grid (grid box). These operations were performed using the *pdb2gmx* and *editconf* functionalities, as proposed by Senra & Fonseca [11]. Additionally, the Open Babel 3.1.1 software [25] was applied to prepare the molecular structures of triacylglycerols [TAG (12:0/12:0/12:0), TAG(14:0/14:0/14:0), and TAG(16:0/16:0/16:0)] derived from lauric, myristic, and palmitic acids, with PubChem Compound Identifiers (CIDs) 10,851, 11,148, and 11,147, respectively.

The selection of lauric (C12:0), myristic (C14:0), and palmitic (C16:0) acids was guided by both biological and biomedical relevance. Palmitic acid is one of the most abundant saturated fatty acids present in members of the genus *Cereus* [2, 4], while lauric and myristic acids are representative medium- and long-chain saturated fatty acids widely used in cardiovascular and metabolic research. These three molecules are among the saturated fatty acids most strongly associated with dyslipidemia, inflammation, and cardiovascular disease in humans [26, 27]. Their inclusion enabled us to evaluate how chain length and metabolic relevance influence binding affinity to *C. jamacaru* proteins. Although lauric and myristic acids have not been explicitly quantified in *C. jamacaru*, their structural similarity to other fatty acids present in the species makes them suitable model substrates for comparative docking analyses.

2.2.3 | Molecular Docking

With the protein structures and their ligands prepared, we proceeded to the docking analysis, conducted using the blind docking strategy. In blind docking, the three triacylglycerol target substrates were virtually positioned within the docking grid without prior information about their binding sites, utilizing the AutoDock Vina program [28], which applies a Monte Carlo iterated local search algorithm coupled with a scoring function that estimates the Gibbs free energy of binding (kcal mol^{-1}), with default parameters. This method enabled the exploration of different conformations and orientations of the compounds within potential binding cavities of the target protein. This approach is particularly relevant as it allows the investigation of flexible conformations, the identification of allosteric binding sites, and functional implications in a more comprehensive manner [29]. As a result, proteins were selected based on their lowest binding energy (kcal/mol), indicating strong interactions between the protein and the substrate. This process aimed to identify potential targets for the functional and metabolic annotation of the pathways of interest. Docking was performed with an exhaustiveness value of 8, and the top 5 poses were retained per ligand. These parameters were selected after preliminary benchmarking with a subset of proteins using exhaustiveness values of 8, 16, and 32, which yielded no significant differences in predicted binding affinities ($\Delta < 0.2$ kcal mol^{-1}) or binding poses (RMSD < 1.5 Å). Given the large dataset (> 3000 protein–ligand combinations), an exhaustiveness of 8 provided an optimal balance between computational efficiency and convergence reliability. This value is within the range commonly adopted in large-scale docking studies using Vina [28, 30]. A cubic grid box of $40 \times 40 \times 40$ Å centered on the protein centroid was used for blind docking simulations. To assess interaction specificity, additional control dockings were conducted using unrelated small molecules (e.g., glucose and glycine) and random proteins from the same dataset. Protein–ligand interactions were further evaluated using Protein–Ligand Interaction Profiler

(PLIP; [31] available at: https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index to characterize hydrogen bonding, hydrophobic contacts, and salt bridges.

It is important to note that the docking results presented here are exploratory and serve as an initial screening to identify potential protein–ligand interactions rather than to establish definitive binding conclusions. Only AutoDock Vina was employed in this work, chosen for its balance between computational efficiency and accuracy in large-scale screenings. Although multi-engine or consensus docking approaches can further refine candidate selection, the scope of this study was to identify promising protein–triacylglycerol interactions for subsequent validation through molecular dynamics simulations.

2.3 | Functional Annotation

To obtain information about the function and metabolic pathways of the target proteins of *C. jamacaru*, functional annotation was performed using a combination of tools and public databases, including the *blastp* program and the KEGG, NCBI-nr, and UniProt databases. Protein sequences were submitted to PROSITE and InterProScan online tools to identify functional domains, protein families, and catalytic or binding residues. The data were analyzed to correlate structural properties with potential biological functions.

2.4 | Molecular Dynamics

The molecular dynamics simulation aimed to understand the binding potential and explore the structural and dynamic properties of the complex. This analysis was conducted using GROMACS v.2023 [32] in a multiparallel setup to leverage computational efficiency. Initially, the protein and ligand files, commonly outputted from docking software in PDBQT format, were converted to PDB format using Open Babel. The ligand topology was then prepared using ACPYPE v.2023.10.27 [33] with the all-atom representation of the GAFF2 force field, which assigns parameters tailored for small molecules, ensuring an accurate representation of the ligand's geometry and electrostatic properties for reliable simulation results.

For the protein topology, the all-atom representation of the OPLS-AA force field and the TIP3P water model were employed to describe system interactions, both selected for their robust performance in simulating biomolecular systems. A dodecahedron simulation box was generated around the protein with a buffer of 1.0 nm from the protein to the box edges, providing adequate space for solvation and interactions [32]. The system was solvated using the SPC216 water model, which balances computational efficiency and accuracy, and ions were added to neutralize the system and maintain a physiological ionic concentration of 0.1 M.

Following solvation and neutralization, the system underwent energy minimization to resolve steric clashes and bad atomic contacts. Equilibration was then performed in two phases. First, the system was equilibrated under constant volume and temperature (NVT ensemble) at 300K, using a V-rescale thermostat. Next, equilibration was conducted under constant pressure and temperature (NPT ensemble) using the Parrinello–Rahman barostat to maintain thermal and pressure equilibrium at 1 bar. Both equilibration stages were run for 100 ps each, allowing full relaxation of solvent and ion distributions before the production phase. These steps were critical to achieve thermodynamic stability and prepare the system for production dynamics.

The production molecular dynamics simulation was run for 100 ns (50,000,000 steps) with a 2 fs time step, recording energy, coordinates, and velocities every 50,000 steps. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method. Periodic boundary conditions (PBC) were applied in all directions to simulate an infinite system and prevent edge effects. Bond lengths involving hydrogen atoms were constrained using the LINCS algorithm, and the Verlet scheme was utilized for neighbor searching with a 1.0 nm cutoff for Coulomb and Van der Waals interactions. The simulations were conducted using an all-atom representation with the OPLS-AA force field for proteins and the GAFF2 force field for ligands, ensuring accurate treatment of bonded and nonbonded interactions. Nonbonded van der Waals interactions were truncated at 1.0 nm, and long-range electrostatics were computed with the PME method. In the final stages of analysis, the root mean square deviation (RMSD) was calculated to assess the structural deviation of the protein over time. In addition to RMSD, other parameters, such as radius of gyration (Rg), root mean square fluctuation (RMSF), and total potential energy, were computed across the 100 ns production run. Hydrogen bonds between the ligand and active-site residues were quantified throughout the trajectory using the GROMACS *hbond* utility, reporting the percentage of simulation time in which each H-bond was preserved. The solvent-accessible surface area (SASA) was also calculated to evaluate potential water penetration into the binding pocket, and the formation of new H-bonds or hydrophobic contacts not present in the initial docking pose was monitored during trajectory analysis. Interactions between the ligands and key residues were analyzed over time to assess complex stability.

3 | Results

3.1 | Transcriptome Assembly and Modeling

The transcriptome assembly of *C. jamacaru* yielded a total of 128,942 transcripts, which were translated into protein sequences. The IVS analysis of the *C. jamacaru* proteomic data began with ~120,000 translated amino acid sequences. Using the BioProtIS pipeline, the protein primary structures were screened for homologous sequences in the PDB, resulting in 14,739 proteins with identified similarities. The average sequence identity between *C. jamacaru* proteins and their respective templates was 62.4% ± 17.7%, with an average coverage of 78.5% ± 10.2% across aligned regions. All the models showed identities and covered above 50% and 70% of the target sequence, respectively (BioProtIS threshold). These values indicate robust homology modeling quality and are consistent with standard thresholds for functional inference [34]. This extensive dataset served as a foundational step for subsequent structural and functional analyses, enabling a robust exploration of the *C. jamacaru* proteome.

Next, the identified homologous sequences were subjected to three-dimensional structure modeling, generating 3,847 protein structures, which were used for molecular docking analysis. The execution time for each script was estimated based on the volume

of information generated and processed. On average, the processing pipeline completed all phases, producing results for each protein in approximately 280 hours. Considering the amount of data used and the methodologies employed, the cost–benefit ratio is favorable, potentially reducing experimental time in the laboratory.

3.2 | Molecular Docking

The binding affinities (kcal/mol) for lauric acid, myristic acid, and palmitic acid substrates indicate the interaction between the ligands and their respective protein targets. More negative binding energy values correspond to stronger ligand–protein interactions, while less negative values suggest weaker binding. The results consistently showed strong binding affinities across all ligands, reflecting robust data quality and validation according to standard binding score metrics.

A total of 3,847 proteins derived from *C. jamacaru* were subjected to molecular docking against triacylglycerol substrates composed of lauric acid, myristic acid, and palmitic acid (Table S1). Table 1 highlights the top 20 candidate proteins based on their binding energy scores (kcal/mol). These proteins represent the most promising targets for the identified ligands (Figure 1). For lauric acid, notable candidates include *Cytokinin Dehydrogenase* (*T89423*), which exhibited the strongest binding energy of −7.78 kcal/mol, followed by *Glycolate Oxidase* (*T57692*) and *Cinnamic Acid 4-Hydroxylase* (*T113801*). The presence of *Banyan Peroxidase* (*T97361*) among the top candidates underscores its potential as a versatile protein for triacylglycerol interactions, a feature also observed for other substrates.

For myristic acid, the protein *Xyloglucan Endotransglycosylase* (*T91586*) displayed the strongest binding energy (−7.75 kcal/mol), followed closely by *Nonspecific Lipid Transfer Protein* (*T6778*). It is important to note that this interaction was not exclusive to myristic acid. The same proteins also exhibited measurable affinities for lauric and palmitic acid triacylglycerols, albeit with less favorable binding energies. Therefore, the interaction can be interpreted as preferential rather than specific, suggesting a broader substrate compatibility typical of lipid-binding plant proteins. These results suggest their suitability for interactions with triacylglycerols containing this fatty acid. Palmitic acid, however, revealed significant involvement of *Banyan Peroxidase*, represented by three distinct transcripts (*T97361, T97368, and T97353*), with binding energies ranging from −7.63 to −7.48 kcal/mol. This consistent representation of *Banyan Peroxidase* across all three triacylglycerol types reinforces its potential as a key player in ligand binding, making it a prime target for further functional annotation and experimental validation. These findings provide a strong foundation for subsequent project phases, which will focus on elucidating the functional roles of these proteins.

To refine the analyses of the binding site potential (allosteric or not) of *C. jamacaru* proteins, transcripts that appeared repeatedly in the blind docking results for the three different substrates (lauric acid, myristic acid, and palmitic acid) were selected. For this purpose, variations in binding energies and the proteins most frequently observed among the three ligands simultaneously were analyzed (see Table 2 for the top five ligands; for an overview of the table, refer to Table S1). The results revealed a total of 59 proteins that appeared across all three substrates, representing a significant sample of proteins with potential for exploring therapeutic applications.

To further explore the interaction profiles of the selected proteins, we analyzed the number of hydrophobic contacts between each protein and the triacylglycerol ligands using PLIP. The results revealed that all proteins exhibited multiple hydrophobic interactions, supporting their ability to stably associate with the lipid substrates. *Banyan Peroxidase isoform* T97368 showed the highest number of hydrophobic contacts (21), followed by *T97361* with 20 contacts. T97361, T97353, T91586 (*Xyloglucan Endotransglycosylase*), and T6778 (*Nonspecific Lipid Transfer Protein*) each exhibited 16 interactions. These findings align with the molecular docking data and underscore the structural adaptability of *Banyan Peroxidase*, as well as the compact, stable interaction modes of *Xyloglucan Endotransglycosylase* and *nsLTP*.

3.3 | Functional Annotation

The analyses performed with PROSITE and InterProScan, along with the molecular docking results, reveal that the three identified and studied proteins possess structural and functional characteristics that make them promising candidates for interacting with triacylglycerols. The *nsLTP* belongs to a family of proteins recognized for their ability to bind and transport lipids, such as triacylglycerols. The presence of the *nsLTP1* domain and a signal peptide suggests that this protein is exported to extracellular environments, where it may play a role in lipid transport and modulation, as corroborated by its affinity for triacylglycerols formed by lauric, myristic, and palmitic acids, indicated by docking analyses. The *Banyan Peroxidase*, however, is a plant peroxidase that features conserved domains associated with calcium and heme binding, essential for its role in redox processes. Interestingly, its ability to interact with triacylglycerols suggests a possible additional role, potentially modulating or oxidizing triacylglycerols, which could expand its applicability in metabolic contexts. The *Xyloglucan Endotransglycosylase Protein*, known for its role in carbohydrate metabolism and cell wall modification in plants, exhibited an unexpected capacity to interact with triacylglycerols, which may be related to the flexibility of its catalytic domain and its predisposition for interactions with nonpolar molecules. These findings indicate that all three proteins possess functional affinity for triacylglycerols, broadening their biological roles.

3.4 | Molecular Dynamics

The molecular dynamics results of the seven analyzed proteins revealed significant differences in structural and energetic stability, which may influence their potential as therapies for triacylglycerol binding and removal (Figure 2 and Table 3). The RMSD analysis demonstrated that T91586 (*Xyloglucan Endotransglycosylase*) exhibited the highest structural stability, with a range value of 5.0–5.4 nm and low variation (0.0656 nm), indicating good consistency over time. In contrast, T89423 (*Cytokinin Dehydrogenase*) showed higher variability in its RMSD, suggesting structural instability under dynamic conditions. Across all systems, the mean RMSD values ranged from 5.27 to 6.89 nm, reflecting moderate overall deviations typical of stable protein–ligand complexes under equilibrium conditions.

TABLE 1 | *C. jamacaru* transcripts, their respective binding energies, and annotations for the top 20 candidate protein targets for triacylglycerol formed by lauric acid, myristic acid, and palmitic acid (triglycerides), based on blind docking.

Transcripts	PDB ID	Binding energy (kcal/mol)	Target protein
Lauric acid triglyceride			
T89423	2q4w	−7782	*Cytokinin Dehydrogenase*
T57692	1gyl	−7721	*Glycolate Oxidase*
T113801	6vby	−7673	*Cinnamic Acid 4–Hydroxylase*
T97361	4cuo	−7603	*Banyan Peroxidase*
T64835	5nzv	−7579	*Coatomer Subunit Beta'*
Myristic acid triglyceride			
T91586	1un1	−7752	*Xyloglucan Endotransglycosylase*
T32032	4i94	−7563	*Probable Serine/Threonine–Protein Kinase*
T38514	4c44	−7524	*2-on-2 Hemoglobin*
T97366	4cuo	−7436	*Banyan Peroxidase*
T6778	7ksc	−7394	*Nonspecific Lipid Transfer Protein*
Palmitic acid triglyceride			
T97361	4cuo	−7632	*Banyan Peroxidase*
T19750	7vka	−7587	*Indole-3–Acetic Acid–Amido Synthetase*
T62011	7aaq	−7,57	*Sugar Transport Protein 10*
T97368	4cuo	−7561	*Banyan Peroxidase*
T97353	4cuo	−7484	*Banyan Peroxidase*

Note: RMSD values for all analyses were 0.0.

The RMSF analysis highlighted that *T6778* (*Nonspecific Lipid Transfer Protein*) exhibited the lowest residue fluctuations, indicating a compact and stable structure. Meanwhile, *T97361* (*Banyan Peroxidase*) showed greater flexibility in specific regions, such as loops and termini, which may facilitate dynamic interactions with ligands. The average RMSF values across all proteins ranged between 0.13 and 0.32 nm, suggesting that none of the proteins experienced excessive atomic displacement during the

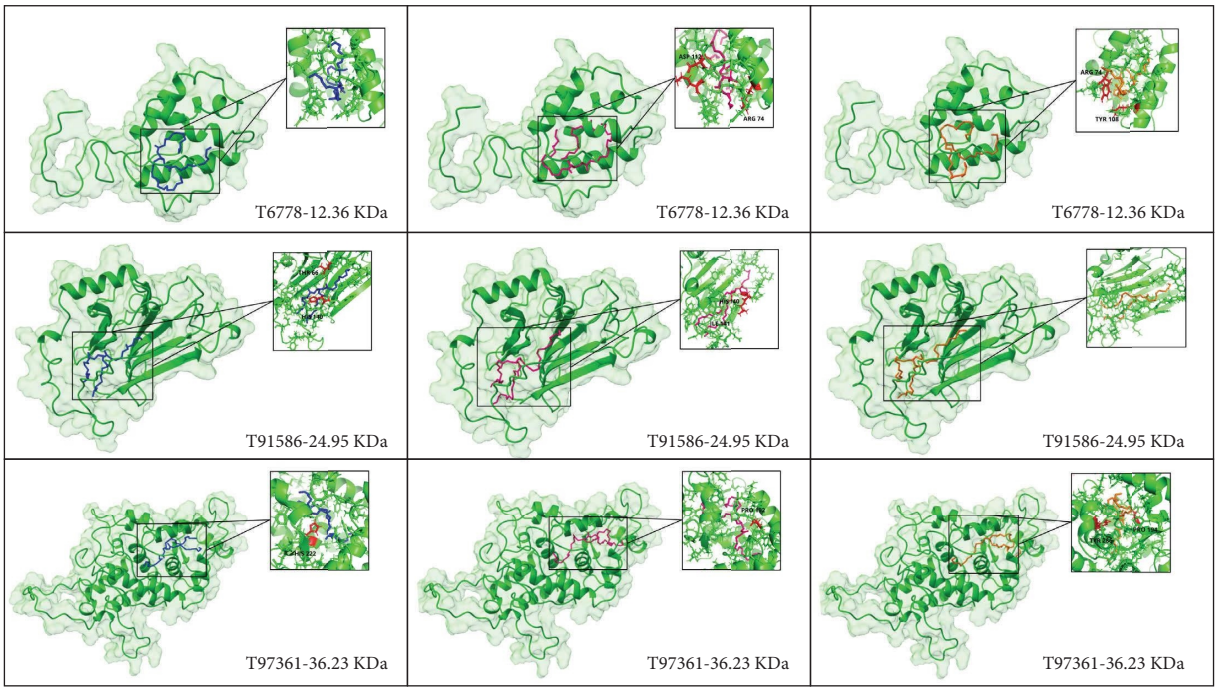


FIGURE 1 | Three-dimensional structure of the three main proteins with the best binding energy for the three substrates: T97361—*Banyan Peroxidase* (BP), T6778—*Nonspecific Lipid Transfer Protein* (NSLT), and T91586—*Xyloglucan Endotransglycosylase* (XE). (a) Docking for the anchoring of proteins with the lauric acid ligand. (b) Docking for the anchoring of proteins with the palmitic acid ligand. (c) Docking for the anchoring of proteins with the myristic acid ligand.

TABLE 2 | *C. jamacaru* transcripts; binding energies for triacylglycerol formed by lauric acid, myristic acid, and palmitic acid; and annotation of the 20 best candidate protein targets for the three substrates mentioned according to blind docking.

Transcripts	PDB ID	Binding energy of lauric acid (kcal/mol)	Binding energy of myristic acid (kcal/mol)	Binding energy of palmitic acid (kcal/mol)	Target protein
T97361	4cuo	−7603	−7178	−7632	*Banyan Peroxidase*
T6778	7ksc	−7475	−7394	−7479	*Nonspecific Lipid Transfer Protein*
T91586	1un1	−7102	−7752	−7304	*Xyloglucan Endotransglycosylase*
T97353	4cuo	−7425	−7063	−7484	*Banyan Peroxidase*
T97368	4cuo	−7386	−6933	−7561	*Banyan Peroxidase*

production phase. The Rg for *T6778* was also the most consistent, reinforcing its global stability. In contrast, *T32032* (*Probable Serine/Threonine-Protein Kinase*) displayed more pronounced variations, possibly linked to conformational transitions. On average, Rg values varied from 1.50 to 2.36 nm, confirming that the overall folding of the proteins remained stable throughout the simulation.

The analysis of SASA showed values between 79.6 and 234.6 nm^2, indicating distinct levels of solvent exposure among the complexes. Proteins, such as T89423 and T97368, exhibited larger SASA, consistent with more solvent-exposed hydrophobic residues that may play a role in substrate recognition. In addition, the number of hydrogen bonds fluctuated between 0.02 and 1.47 on average, with T89423 showing the highest H-bonding capacity, suggesting that hydrogen bond stability may contribute to its structural rigidity and interaction with the surrounding solvent.

Regarding total energy, *T6778* (*Nonspecific Lipid Transfer Protein*) stood out with the lowest average energy and reduced variation, indicating an energetically stable conformation. However, *T97368* (*Banyan Peroxidase*) showed the highest total energy (−7359.22 kJ/mol), reflecting a greater energetic cost to maintain its equilibrium under dynamic conditions. The energy profiles followed the same trend observed for Rg and RMSD, reinforcing the coherence between structural and energetic parameters.

Overall, the molecular dynamics data indicate that *T91586*, *T6778*, and *T97361* possess favorable characteristics that merit further exploration as potential scaffolds for lipid-binding biotechnological tools, either due to global stability or the moderate flexibility needed for specific interactions.

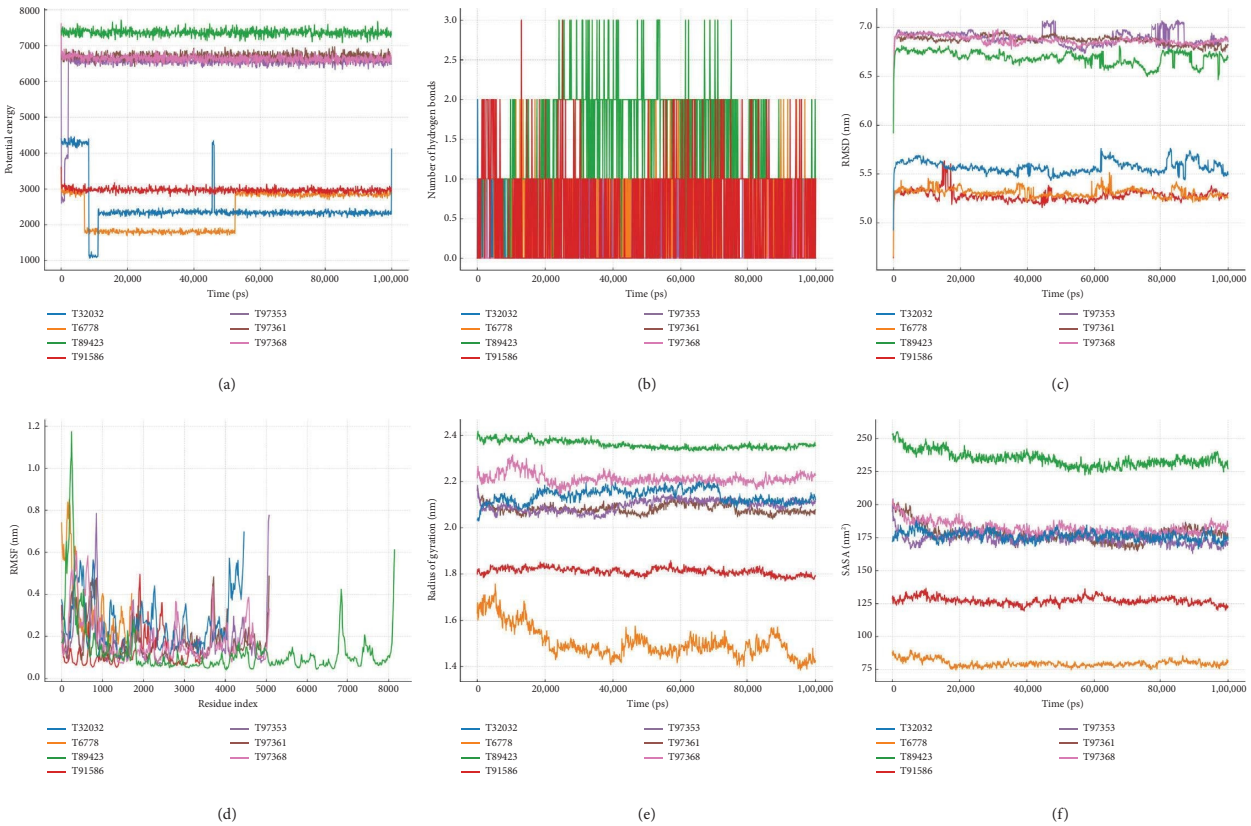


FIGURE 2 | Analysis of molecular dynamics parameters for candidate proteins, presenting the results for seven proteins derived from *C. jamacaru*. The panel includes (a) total energy; (b) hydrogen bonds; (c) RMSD over time; (d) RMSF per residue; (e) radius of gyration (Rg); and (f) SASA. Each line corresponds to a protein identified by its transcript code (T), with colors matching the legend provided.

TABLE 3 | Mean values of molecular dynamics descriptors for each transcript–protein complex.

Transcript	Energy	H-bonds	RMSD	RMSF	Radius of gyration	SASA
T32032	2477.114601	0.037962	5.569604	0.258935	2.138235	176.240820
T6778	2382.098837	0.471528	5.309108	0.327542	1.504125	79.650454
T89423	7359.220924	1.468531	6.690485	0.131493	2.360595	234.634966
T91586	2965.039146	0.541459	5.279295	0.133994	1.812111	127.093417
T97353	6494.788617	0.066933	6.895383	0.178783	2.097240	174.035349
T97361	6705.402738	0.024975	6.869549	0.177553	2.082102	177.839971
T97368	6634.013009	0.019980	6.869345	0.182011	2.213000	182.333305

Note: For each metric, potential energy ($kJ \cdot mol^{-1}$), number of hydrogen bonds (dimensionless), RMSD (nm), RMSF (nm), radius of gyration (nm), and SASA (nm^2), values correspond to the average over the entire production trajectory.

4 | Discussion

The therapeutic potential of natural products has been widely recognized, but transitioning traditional knowledge into modern applications poses challenges [35]. While current therapies for hypertriglyceridemia and cardiovascular diseases are often limited by accessibility and affordability, computational approaches, such as molecular docking, present a cost-effective avenue for exploring plant-derived proteins as potential therapeutic agents. Despite extensive studies on well-known medicinal plants, there is a notable gap in the proteomic exploration of native species from underrepresented ecosystems, such as the Caatinga. The unique adaptations of *C. jamacaru* to its environment likely yield specialized bioactive proteins, which remain largely untapped. We emphasize that ethnobotanical history motivated target selection but does not validate therapeutic effects. Accordingly, our results should be interpreted as computational leads rather than evidence of therapeutic efficacy. Translation to therapy would require rigorous biochemical assays, bioavailability and safety studies, and disease-relevant models.

The molecular docking was employed in this study to evaluate, for the first time, the binding interactions between triacylglycerols composed of myristic, lauric, and palmitic acids and proteins from *C. jamacaru*. The docking results revealed significant binding affinities, suggesting favorable molecular interactions between these ligands and several candidate proteins. Functional annotation through the PROSITE and InterProScan databases revealed conserved domains compatible with lipid metabolism or binding, providing contextual support rather than direct validation of the docking outcomes. These complementary analyses indicate that the three main identified proteins may possess structural features enabling interaction with triacylglycerols, a property not previously described for these candidates.

The patterns of hydrophobic interactions observed in the selected proteins offer important insights into their potential biological roles. Proteins with a larger number of contacts, such as specific isoforms of *Banyan Peroxidase*, may exhibit flexible binding surfaces that accommodate various lipid substrates, making them promising candidates for bioengineering approaches aimed at lipid sequestration or modulation. In contrast, proteins with a more conserved number of interactions, such as *nsLTP* and *Xyloglucan Endotransglycosylase*, likely possess compact binding pockets that favor specificity over versatility. This distinction could be crucial when selecting proteins for different biotechnological applications, whether aiming for broad-spectrum lipid-binding or targeted triacylglycerol interaction. Moreover, the structural differences observed among isoforms highlight the relevance of analyzing not only sequence similarity but also interaction dynamics and site architecture.

Banyan Peroxidase was first purified from the latex of *Ficus benghalensis* and underwent crystallographic analysis in 2012 [36]. While its crystallographic structure has been resolved, its specific biological functions remain largely unexplored, including any potential lipid-binding capabilities. This protein is generally recognized for its role in oxidative stress responses and the detoxification of reactive oxygen species in plants. *Xyloglucan Endotransglycosylase* has been intensely studied since its crystallization in 1992 [37] across several plant species. Its primary known function is the cutting and rejoining of interfibrillar xyloglucan chains, facilitating wall loosening required for plant cell expansion [37]. Despite its extensive study, no lipid-binding capacity has been demonstrated for this enzyme. This protein is primarily associated with plant growth and structural dynamics of the cell wall. The *Nonspecific Lipid Transfer Protein* was crystallized in 1995 from maize seedlings [38], and homologs have been identified in numerous plant species. In tobacco, this protein has been proposed to transfer phospholipids, contributing to plant defense against bacterial and fungal pathogens, and promoting cell wall extension [39]. It is widely recognized for its involvement in lipid transport and signaling within plant tissues, as well as its role in stress responses.

The strategy allowed for the evaluation of the affinity of the modeled proteins for these substrates, providing a deeper understanding of their molecular interaction and functions. The results from this automated approach demonstrated effectiveness in identifying and modeling the proteins with the best binding energy models for the substrates. The discovered proteins may serve as molecular scaffolds to be optimized through protein engineering for potential biotechnological applications related to lipid metabolism.

Fatty acids are classified according to the length of their carbon chain and the number and configuration of their bonds. These chemical characteristics are associated with the plasma concentration of cholesterol and its distribution in lipoproteins. Fatty acids can be divided into saturated and unsaturated [13]. Long-chain fatty acids are found in solid form at room temperature, with the main ones being lauric (12:0), myristic (14:0), palmitic

(16:0), and stearic (18:0). Different saturated fatty acids may have diversified effects on the lipid profile and cardiovascular risk factors. Nonetheless, the definitive effect of a diet enriched with a specific fatty acid is still under investigation and debate.

The primary sources of dietary fat include butterfat and animal fat, which are rich in palmitic acid, as well as vegetable oils, such as soybean oil and palm oil. The latter contains a low concentration of myristic acid but is high in palmitic acid [40, 41]. Due to the stronger cholesterol-raising effect of myristic acid, attributed to the length of its carbon chain, earlier conclusions suggested that palmitic acid could serve as a favorable dietary substitute in some cases [42]. However, more recent studies indicate that palmitic acid derived from palm oil may not be entirely beneficial, as it has been associated with an elevation in LDL-c levels [43]. A proposed mechanism for this effect involves a decrease in hepatic LDL receptor activity combined with an increased production rate of LDL cholesterol [44]. Nevertheless, a systematic review reported that the overall effect of palm oil on lipid profiles remains inconclusive [45], potentially due to its additional effect of raising HDL-c level. This HDL-c-raising effect may counterbalance the LDL-c-raising effect, resulting in a neutral overall impact on cardiovascular health [46].

It is important to note that proteins, such as *Peroxidases* and *Xyloglucan Endotransglycosylases*, do not have established roles in human lipid metabolism. Their inclusion in this study was based on observed in silico interactions with triacylglycerols, not on prior biological evidence. As such, any suggestion of functional relevance should be interpreted cautiously and considered a starting point for future biochemical characterization and engineering processes.

4.1 | Integration of Molecular Docking and Molecular Dynamics

The integration of molecular docking and molecular dynamics results provides a comprehensive perspective on the structural and functional behavior of the proteins analyzed. Docking studies revealed the binding affinities of lauric acid, myristic acid, and palmitic acid to key proteins derived from *C. jamacaru*. Among these, the *Banyan Peroxidase* (*T97361*) consistently demonstrated strong binding energies across all ligands, with particularly high affinities for palmitic acid (−7.63 kcal/mol) and lauric acid (−7.60 kcal/mol). This versatility suggests its potential as a primary candidate for interacting with diverse triacylglycerol substrates (Figure 1). Incorporating molecular dynamics descriptors, *Banyan Peroxidase* exhibited a mean RMSD of 6.87 nm, an average RMSF of 0.18 nm, and a SASA of approximately 182 nm^2, indicating a moderately flexible yet solvent-exposed structure that may enhance substrate accommodation. The relatively low number of hydrogen bonds (0.02) also suggests transient but adaptable interactions with the surrounding medium, consistent with its role in catalysis and lipid accessibility.

The *Xyloglucan Endotransglycosylase* (*T91586*) also emerged as a significant candidate (Figure 1), particularly for myristic acid binding, with a binding energy of −7.75 kcal/mol. Molecular dynamics further supported its candidacy by highlighting its compact and stable structure, as evidenced by low RMSD values and consistent Rg. Quantitatively, this protein presented an average RMSD of 5.28 nm, one of the lowest RMSF values (0.13 nm), and a stable Rg around 1.81 nm, supporting minimal conformational fluctuation. The moderate SASA (127 nm^2) and intermediate hydrogen-bonding profile (0.54) indicate a balance between structural rigidity and solvent exposure, which could favor sustained ligand interactions without compromising folding integrity. This dual evidence of strong ligand affinity and structural robustness underscores its potential role in therapeutic applications targeting specific triacylglycerol types.

Meanwhile, the *Nonspecific Lipid Transfer Protein* (T6778), while showing slightly weaker docking results (Figure 1), stood out in dynamics analyses due to its superior energetic stability and minimal structural fluctuations, suggesting its suitability for scenarios requiring high stability under diverse conditions. The MD metrics reinforce this interpretation: T6778 displayed the lowest RMSF (0.32 nm) and one of the smallest Rg values (1.50 nm), alongside a low SASA (79.6 nm^2), confirming a tightly packed and stable conformation with minimal solvent exposure. Despite fewer hydrogen bonds (0.47), its high mean stability energy (∼2382 $kJ \cdot mol^{-1}$) suggests an energetically favorable core that resists unfolding during dynamic motion.

The integration of these datasets reveals nuanced roles for each protein in triacylglycerol interaction. *Banyan Peroxidase* emerges as the most versatile and robust candidate, suitable for general applications across multiple triacylglycerol types. *Xyloglucan Endotransglycosylase* excels in specific interactions with myristic acid, while *Nonspecific Lipid Transfer Protein* offers remarkable stability for experimental or therapeutic conditions that demand structural resilience. Beyond these three, other analyzed proteins, such as *Cytokinin Dehydrogenase* (T89423) and *Probable Serine/Threonine-Protein Kinase* (T32032), displayed higher RMSD and SASA values (6.69 nm and 234 nm^2, respectively), indicative of more flexible or solvent-exposed conformations that could be explored for dynamic substrate turnover rather than long-term structural rigidity. Together, these results provide a strong framework for prioritizing proteins for further experimental validation and functional annotation, bridging in silico findings with practical applications in the reduction of triacylglycerols and cholesterol.

The novelty of this study lies in its integration of in silico techniques to uncover previously uncharacterized proteins from an endemic plant species and assess their potential interactions with triacylglycerols. The results not only highlight the unique biochemical properties of *C. jamacaru* but also demonstrate how computational bioprospecting can efficiently identify promising candidates for metabolic intervention. The inclusion of quantitative MD parameters enhances the mechanistic understanding of protein stability, flexibility, and solvent interaction, providing a more detailed molecular picture than docking alone. This study represents a pioneering step toward understanding plant-derived bioactive molecules in a systematic and scalable manner, potentially guiding future experimental validations and biotechnological applications in lipid metabolism regulation. Nevertheless, it is important to recognize that plant-derived proteins identified here may face significant biological barriers for therapeutic application in humans. Issues, such as oral bioavailability, proteolytic degradation, immunogenicity, and delivery across biological membranes, must be addressed in future experimental phases. These findings, therefore, serve primarily to highlight molecular interaction potential and do not imply direct therapeutic applicability.

5 | Conclusion

The *Banyan Peroxidase* (*T97361*) stands out as a versatile candidate for future biotechnological development focused on modulating triacylglycerol interactions. Its combination of strong binding affinities and balanced molecular dynamics profile, moderate RMSD, low RMSF, and high SASA suggests a flexible yet solvent-accessible conformation capable of accommodating diverse lipid substrates. These preliminary in silico results suggest it may be a suitable target for further protein engineering studies. The *Xyloglucan Endotransglycosylase* (*T91586*) exhibited strong specificity for myristic acid, supported by its low RMSD, minimal atomic fluctuations, and consistent Rg, confirming its structural robustness during dynamics. Meanwhile, the *Nonspecific Lipid Transfer Protein* (*T6778*) demonstrated superior energetic stability, with compact folding and low solvent exposure, making it the most rigid and energetically favorable candidate among the set. These molecular descriptors reinforce that T91586, T6778, and T97361 remain stable throughout the entire simulation, validating their potential as triacylglycerol-binding scaffolds. Possibly, these candidates might form stable complexes in silico of triacylglycerols if appropriate conditions are offered.

From a global perspective, the increasing prevalence of cardiovascular diseases and obesity underscores the urgent need for innovative and cost-effective therapeutic solutions. These findings contribute to identifying plant proteins with molecular affinity to dietary triacylglycerols, offering a starting point for further functional and biochemical studies. While therapeutic application remains speculative, these results highlight the potential of *C. jamacaru* as a source of structurally interesting bioactive proteins for future exploration in biotechnology. It is important to emphasize that the results presented here are based solely on computational predictions. Further experimental validation is required to confirm these proteins' roles and evaluate their feasibility in real-world applications.

Author Contributions

M.I.O.C. and D.T.A conceived the study. M.I.O.C. and J.O. collected and analyzed the data, and D.T.A. and M.I.O.C. led the writing of the paper, while J.O., V.B., and D.T.A. contributed numerous conceptions and writing. All authors contributed to the intellectual development of the paper and made multiple revisions.

Acknowledgments

We acknowledge the support of São Paulo Research Foundation (FAPESP 2022/09910-9 to D.T.A. and 2023/06831-3 to M.I.O.C.) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 132515/2024-5 to M.I.O.C.).

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process. During the preparation of this work, the author(s) used ChatGPT 5.0 to improve English grammar. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

The study was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico, 132515/2024-5; Fundação de Amparo à Pesquisa do Estado de São Paulo, 2022/09910-9, 2023/06831-3.

Disclosure

A preprint of this manuscript has previously been published [47]. All authors approved the final draft.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets reused in this study are publicly available on NCBI under the project accession PRJNA1192998. The transcriptome assembly is available at FigShare: 10.6084/m9.figshare.30483716.

References

1. S. L. Cartaxo, M. M. de Almeida Souza, and U. P. de Albuquerque, "Medicinal Plants With Bioprospecting Potential Used in Semi-arid Northeastern Brazil," *Journal of Ethnopharmacology* 131, no. 2 (2010): 326–342, https://doi.org/10.1016/j.jep.2010.07.003.

2. M. V. Meiado, L. S. C. Albuquerque, E. A. Rocha, M. Rojas-Aréchiga, and I. R. Leal, "Seed Germination Responses of *Cereus jamacaru* D.C. Ssp. *Jamacaru* (Cactaceae) to Environmental Factors," *Plant Species Biology* 25, no. 2 (2010): 120–128, https://doi.org/10.1111/j.1442-1984.2010.00274.x.

3. M. D. S. L. Sales, L. V. Martins, I. Souza, M. S. M. D. Deus, and A. P. Peron, "Cereus jamacaru DE CANDOLLE (CACTACEAE), O MANDACARU DO NORDESTE BRASILEIRO. Publicatio UEPG," *Ciencias Biologicas e Da Saude* 20, no. 2 (2014): 135–142, https://doi.org/10.5212/publ.biologicas.v.20i2.0006.

4. L. F. C. R. Silva, L. de S. Valle, A. R. C. Nascimento, and M. F. T. Medeiros, "*Cereus jamacaru* DC. (Cactaceae): From 17th Century Naturalists to Modern Day Scientific and Technological Prospecting," *Acta Botanica Brasilica* 33, no. 2 (2018): 191–197, https://doi.org/10.1590/0102-33062018abb0352.

5. M. A. Jorquera, S. P. Graether, and F. Maruyama, "Editorial: Bioprospecting and Biotechnology of Xtremophiles," in *Frontiers in Bioengineering and Biotechnology*, 7 (Frontiers Media S.A, 2019), 204, https://doi.org/10.3389/fbioe.2019.00204.

6. J. Arora, S. Goyal, and K. G. Ramawat, "Biodiversity, Biology and Conservation of Medicinal Plants of the Thar Desert," in *Desert Plants Biology and Biotechnology*, ed. K. G. Ramawat (Berlin: Springer-Verlag, 2010), 3–36.

7. G. Yadav and M. Meena, "Bioprospecting of Endophytes in Medicinal Plants of Thar Desert: An Attractive Resource for Biopharmaceuticals," *Biotechnology Reports* 30 (2021): e00629, https://doi.org/10.1016/j.btre.2021.e00629.

8. P. Kumar, S. P. Banik, A. Goel, S. Chakraborty, M. Bagchi, and D. Bagchi, "Revisiting the Multifaceted Therapeutic Potential of Withaferin A (WA), a Novel Steroidal Lactone, W-ferinAmax Ashwagandha, from *Withania somnifera* (L) Dunal," *Journal of the American Nutrition Association* 43, no. 2 (2024): 115–130, https://doi.org/10.1080/27697061.2023.2228863.

9. Z. Shi, Z. J. Wang, H. L. Xu, et al., "Modeling, Docking and Dynamics Simulations of a Non-Specific Lipid Transfer Protein from *Peganum harmala* L," *Computational Biology and Chemistry* 47 (2013): 56–65, https://doi.org/10.1016/j.compbiolchem.2013.07.001.

10. X. Xu, M. Huang, and X. Zou, "Docking-Based Inverse Virtual Screening: Methods, Applications, and Challenges," *Biophysics Reports* 4, no. 1 (2018): 1–16, https://doi.org/10.1007/s41048-017-0045-8.

11. M. V. Senra and A. L. Fonseca, "Toxicological Impacts and Likely Protein Targets of Bisphenol A in *Paramecium caudatum*," *European Journal of Protistology* 88 (2023): 125958, https://doi.org/10.1016/j.ejop.2023.125958.

12. M. Xu and M. A. Lill, "Induced Fit Docking, and the Use of QM/MM Methods in Docking," *Drug Discovery Today: Technologies* 10, no. 3 (2013): e411–e418, https://doi.org/10.1016/j.ddtec.2013.02.003.

13. R. D. Santos, A. C. M. Gagliardi, H. T. Xavier, et al., "I Diretriz Sobre o Consumo de Gorduras e Saúde Cardiovascular," *Arquivos Brasileiros de Cardiologia* 100, no. 1 (2013): 1–40, https://doi.org/10.1590/S0066-782X2013000900001.

14. M. G. Silverman, B. A. Ference, K. Im, et al., "Association Between Lowering LDL-C and Cardiovascular Risk Reduction Among Different Therapeutic Interventions: A Systematic Review and meta-analysis," *JAMA* 316, no. 12 (2016): 1289–1297, https://doi.org/10.1001/jama.2016.13985.

15. M. M. Hussain, "Intestinal Lipid Absorption and Lipoprotein Formation," *Current Opinion in Lipidology* 25, no. 3 (2014): 200–206, https://doi.org/10.1097/mol.0000000000000084.

16. F. Carrière, C. Renou, S. Ransac, et al., "Inhibition of Gastrointestinal Lipolysis by Orlistat During Digestion of Test Meals in Healthy Volunteers," *American Journal of Physiology-Gastrointestinal and Liver Physiology* 281, no. 1 (2001): G16–G28, https://doi.org/10.1152/ajpgi.2001.281.1.g16.

17. U. Laufs, K. G. Parhofer, H. N. Ginsberg, and R. A. Hegele, "Clinical Review on Triglycerides," *European Heart Journal* 41, no. 1 (2020): 99–109, https://doi.org/10.1093/eurheartj/ehz785.

18. S. M. Grundy, "Obesity, Metabolic Syndrome, and Cardiovascular Disease," *Journal of Clinical Endocrinology and Metabolism* 89, no. 6 (2004): 2595–2600, https://doi.org/10.1210/jc.2004-0372.

19. E. Pimienta-Barrios, L. Mendez-Moran, B. C. Ramirez-Hernandez, J. E. Garcia de Alba-Garcia, and R. M. Dominguez-Arias, "Effect of Xoconostle (*Opuntia joconostle* Web.) Fruit Consumtion of Glucose and Seric Lipids," *Agrociencia* 42, no. 6 (2008): 645–653.

20. C. Gouws, R. Mortazavi, D. Mellor, A. McKune, and N. Naumovski, "The Effects of Prickly Pear Fruit and Cladode (*Opuntia* spp.) Consumption on Blood Lipids: A Systematic Review," *Complementary Therapies in Medicine* 50 (2020): 102384, https://doi.org/10.1016/j.ctim.2020.102384.

21. D. T. Amaral, I. A. Bonatelli, M. Romeiro-Brito, et al., "Comparative Transcriptome Analysis Reveals Lineage-and Environment-Specific Adaptations in Cacti From the Brazilian Atlantic Forest," *Planta* 260, no. 1 (2024): 4, https://doi.org/10.1007/s00425-024-04442-x.

22. M. Grabherr, B. Haas, M. Yassour, et al., "Full-Length Transcriptome Assembly from RNA-Seq Data Without a Reference Genome," *Nature Biotechnology* 29, no. 7 (2011): 644–652, https://doi.org/10.1038/nbt.1883.

23. G. S. Virgens, J. Oliveira, M. I. O. Cardoso, J. A. Teodoro, and D. T. Amaral, "Bioprotis: Streamlining Protein-Ligand Interaction Pipeline for Analysis in Genomic and Transcriptomic Exploration," *Journal of Molecular Graphics and Modelling* 128 (2024): 108721, https://doi.org/10.1016/j.jmgm.2024.108721.

24. B. Webb and A. Sali, "Comparative Protein Structure Modeling Using Modeller," *Current Protocols in Bioinformatics* 54, no. 1 (2016): 5.6.1–-5.6.37, https://doi.org/10.1002/cpbi.3.

25. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open Babel: An Open Chemical Toolbox," *Journal of Cheminformatics* 3, no. 1 (2011): 33, https://doi.org/10.1186/1758-2946-3-33.

26. C. E. Annevelink, P. A. Sapp, K. S. Petersen, G. C. Shearer, and P. M. Kris-Etherton, "Diet-Derived and Diet-Related Endogenously Produced Palmitic Acid: Effects on Metabolic Regulation and Cardiovascular Disease Risk," *Journal of Clinical Lipidology* 17, no. 5 (2023): 577–586, https://doi.org/10.1016/j.jacl.2023.07.005.

27. P. W. Siri-Tarino, Q. Sun, F. B. Hu, and R. M. Krauss, "Saturated Fatty Acids and Risk of Coronary Heart Disease: Modulation by Replacement Nutrients," *Current Atherosclerosis Reports* 12, no. 6 (2010): 384–390, https://doi.org/10.1007/s11883-010-0131-6.

28. O. Trott and A. J. Olson, "Autodock Vina: Improving the Speed and Accuracy of Docking With a New Scoring Function, Efficient Optimization, and Multithreading," *Journal of Computational Chemistry* 31, no. 2 (2009): 455–461, https://doi.org/10.1002/jcc.21334.

29. B. Iorga, D. Herlem, E. Barré, and C. Guillou, "Acetylcholine Nicotinic Receptors: Finding the Putative Binding Site of Allosteric Modulators Using the "Blind Docking" Approach," *Journal of Molecular Modeling* 12, no. 3 (2006): 366–372, https://doi.org/10.1007/s00894-005-0057-z.

30. S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell, and A. J. Olson, "Computational Protein–Ligand Docking and Virtual Drug Screening With the Autodock Suite," *Nature Protocols* 11, no. 5 (2016): 905–919, https://doi.org/10.1038/nprot.2016.051.

31. P. Schake, S. N. Bolz, K. Linnemann, and M. Schroeder, "PLIP 2025: Introducing Protein–Protein Interactions to the Protein–Ligand Interaction Profiler," *Nucleic Acids Research* (2025): gkaf361.

32. M. J. Abraham, T. Murtola, R. Schulz, et al., "GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism From Laptops to Supercomputers," *SoftwareX* (2015): 19–25, https://doi.org/10.1016/j.softx.2015.06.001.

33. A. W. Sousa da Silva and W. F. Vranken, "ACPYPE-Antechamber Python Parser Interface," *BMC Research Notes* 5 (2012): 1–8, https://doi.org/10.1186/1756-0500-5-367.

34. Z. Xiang, "Advances in Homology Protein Structure Modeling," *Current Protein and Peptide Science* 7, no. 3 (2006): 217–227, https://doi.org/10.2174/138920306777452312.

35. G. Porras, F. Chassagne, J. T. Lyles, et al., "Ethnobotany and the Role of Plant Natural Products in Antibiotic Drug Discovery," *Chemical Reviews* 121, no. 6 (2020): 3495–3560, https://doi.org/10.1021/acs.chemrev.0c00922.

36. A. Sharma, G. J. Palm, M. Kumari, S. Panjikar, M. V. Jagannadham, and W. Hinrichs, "Purification, Crystallization and Preliminary Crystallographic Analysis of Banyan Peroxidase," *Acta Crystallographica, Section F: Structural Biology and Crystallization Communications* 68, no. 8 (2012): 931–934, https://doi.org/10.1107/s1744309112026097.

37. S. C. Fry, R. C. Smith, K. F. Renwick, D. J. Martin, S. K. Hodge, and K. J. Matthews, "Xyloglucan Endotransglycosylase, a New Wall-Loosening Enzyme Activity From Plants," *Biochemical Journal* 282, no. 3 (1992): 821–828, https://doi.org/10.1042/bj2820821.

38. D. H. Shin, J. Y. Lee, K. Y. Hwang, K. K. Kim, and S. W. Suh, "High-Resolution Crystal Structure of the Non-Specific Lipid-Transfer Protein from Maize Seedlings," *Structure* 3, no. 2 (1995): 189–199, https://doi.org/10.1016/s0969-2126(01)00149-6.

39. N. J. Wang, C. C. Lee, C. S. Cheng, et al., "Construction and Analysis of a Plant Non-Specific Lipid Transfer Protein Database (Nsltpdb)," *BMC Genomics* 13, no. S1 (2012): 1–9, https://doi.org/10.1186/1471-2164-13-s1-s9.

40. E. Fattore, C. Bosetti, F. Brighenti, C. Agostoni, and G. Fattore, "Palm Oil and Blood Lipid–Related Markers of Cardiovascular Disease: A Systematic Review and Meta-Analysis of Dietary Intervention Trials," *The American Journal of Clinical Nutrition* 99, no. 6 (2014): 1331–1350, https://doi.org/10.3945/ajcn.113.081190.

41. A. Mancini, E. Imperlini, E. Nigro, et al., "Biological and Nutritional Properties of Palm Oil and Palmitic Acid: Effects on Health," *Molecules* 20, no. 9 (2015): 17339–17361, https://doi.org/10.3390/molecules200917339.

42. P. L. Zock, J. H. M. de Vries, and M. B. Katan, "Impact of Myristic Acid Versus Palmitic Acid on Serum Lipid and Lipoprotein Levels in Healthy Women and Men," *Arteriosclerosis, Thrombosis, and Vascular Biology* 14, no. 4 (1994): 567–575, https://doi.org/10.1161/01.atv.14.4.567.

43. Y. Sun, N. Neelakantan, Y. Wu, R. Lote-Oke, A. Pan, and R. M. van Dam, "Palm Oil Consumption Increases LDL Cholesterol Compared with Vegetable Oils Low in Saturated Fat in a meta-analysis of

Clinical Trials," *The Journal of Nutrition* 145, no. 7 (2015): 1549–1558, https://doi.org/10.3945/jn.115.210575.

44. R. J. Nicolosi, "Dietary Fat Saturation Effects on low-density-lipoprotein Concentrations and Metabolism in Various Animal Models," *The American Journal of Clinical Nutrition* 65, no. 5 (1997): 1617S–1627S, https://doi.org/10.1093/ajcn/65.5.1617s.

45. S. R. Ismail, S. K. Maarof, S. Siedar Ali, and A. Ali, "Systematic Review of Palm Oil Consumption and the Risk of Cardiovascular Disease," *PLoS One* 13, no. 2 (2018): e0193533, https://doi.org/10.1371/journal.pone.0193533.

46. C. Unhapipatpong, P. C. Shantavasinkul, V. Kasemsup, et al., "Tropical Oil Consumption and Cardiovascular Disease: An Umbrella Review of Systematic Reviews and Meta Analyses," *Nutrients* 13, no. 5 (2021): 1549, https://doi.org/10.3390/nu13051549.

47. M. I. Cardoso, J. Oliveira, V. Bassaneze, and D. T. Amaral, "*In Silico* Approach for the Prospecting of Molecules from *Cereus jamacaru* (Cereeae) in the Treatment of Obesity and Cardiovascular Diseases," *Research Square* (2024): https://doi.org/10.21203/rs.3.rs-5961003/v1.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. *(Supporting Information)*

Table S1. Complete list of the 3,847 modeled protein structures from *C. jamacaru* used in the molecular docking analysis. The table includes transcript IDs, corresponding PDB IDs, binding energies (kcal/mol) for triacylglycerols composed of lauric, myristic, and palmitic acids, and functional annotations of each protein. Proteins that appeared across all three ligands were highlighted and ranked by average binding affinity.