ENDEMIC PLANTS: EXPERIMENTAL AND THEORETICAL INSIGHTS INTO PROPERTIES OF BIOACTIVE METABOLITES WITH THERAPEUTIC POTENTIAL

EDITED BY: Carolina Otero, Jorge Ignacio Martínez-Araya and Jorge M. del Campo

PUBLISHED IN: Frontiers in Chemistry

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-872-6
DOI 10.3389/978-2-88971-872-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ENDEMIC PLANTS: EXPERIMENTAL AND THEORETICAL INSIGHTS INTO PROPERTIES OF BIOACTIVE METABOLITES WITH THERAPEUTIC POTENTIAL

Topic Editors:
Carolina Otero, Universidad Andrés Bello, Chile
Jorge Ignacio Martínez-Araya, Andres Bello University, Chile
Jorge M. del Campo, Universidad Nacional Autónoma de México, Mexico

Topic Editor Professor David Liebeskind is Consultant to Stryker, Imaging Core Lab. All other Topic Editors declare no competing interests with regards to the Research Topic subject.

Citation: Otero, C., Martínez-Araya, J. I., del Campo, J. M., eds. (2021). Endemic Plants: Experimental and Theoretical Insights Into Properties of Bioactive Metabolites With Therapeutic Potential. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-872-6

Table of Contents

frontiers in Chemistry

EDITORIAL
published: 25 October 2021
doi: 10.3389/fchem.2021.786865

Editorial: Endemic Plants: Experimental and Theoretical Insights Into Properties of Bioactive Metabolites With Therapeutic Potential

Carolina Otero[1], Jorge Ignacio Martínez-Araya[2], Jorge M. del Campo[3] and Felipe Gordillo-Fuenzalida[4]*

[1]School of Chemistry and Pharmacy, Faculty of Medicine, Universidad Andres Bello, Santiago, Chile, [2]Departamento de Ciencias Químicas, Facultad de Ciencias Exactas, Universidad Andrés Bello, Santiago, Chile, [3]Departamento de Física y Química Teórica, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México, Mexico City, Mexico, [4]Laboratorio de Microbiología Aplicada, Centro de Biotecnología de los Recursos Naturales, Facultad de Agronomía y Ciencias Forestales, Universidad Católica del Maule, Talca, Chile

Keywords: *in silico*, endemic plants, secondary metabolites, anti-inflammatory, anti-oxidant

Editorial on the Research Topic

Endemic Plants: Experimental and Theoretical Insights Into Properties of Bioactive Metabolites With Therapeutic Potential

INTRODUCTION

Plants are currently used in alternative and ancestral medicine and correspond to the best resource in drug discovery known so far. It has primarily been reported that infusion of herbs as aromatic drinks or tea can protect and heal our organism. From the scientific point of view, it is known that water-soluble substances (normally secondary metabolism) have beneficial effects on people's health. Researchers have taken advantage of plant metabolites as a fundamental part of the healing process, consuming them to treat diseases or relieve pathological conditions. Then, once consumed, these

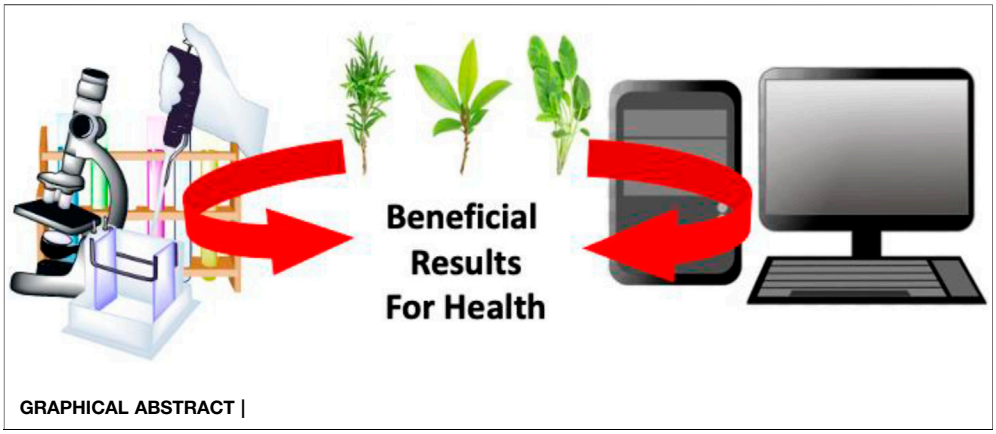
OPEN ACCESS

Edited and reviewed by:
Sam P. De Visser,
The University of Manchester,
United Kingdom

****Correspondence:***
Carolina Otero
maria.otero@unab.cl

Received: *30 September 2021*
Accepted: *04 October 2021*
Published: *25 October 2021*

Citation:
Otero C, Martínez-Araya JI, M. del Campo J and Gordillo-Fuenzalida F (2021) Editorial: Endemic Plants: Experimental and Theoretical Insights Into Properties of Bioactive Metabolites With Therapeutic Potential. Front. Chem. 9:786865. doi: 10.3389/fchem.2021.786865



GRAPHICAL ABSTRACT |

molecules can interact with different targets (including cell membranes), and in some cases, they can even be internalized by cells triggering intracellular responses. The continued search for potential compounds against therapeutic targets is still necessary. Fast screening *in silico* using bioinformatics tools has allowed finding new metabolite candidates derived from natural resources.

COMPUTATIONAL TOOLS

Informatics's role in biomedical science has become not a supplementary tool but a necessary support nowadays since it allows researchers to optimize their work in search of new molecules or new properties of well-known molecules. The maturity attained by Quantum Chemistry, Molecular Dynamics, and Molecular Mechanics provides helpful guidance for activities performed by experimentalist scientists, providing support to their scientific investigation that concerns the health arena. The present Research Topic exposes some examples of this theoretical and experimental combined work to reveal relevant properties that secondary metabolites might present.

CONCLUSION OF THIS TOPIC

Overall, this topic covers contributions that range from analytical chemistry to computational chemistry, applied to endemic Brazilian plants such as *Byrsonima intermedia* and *Serjania marginata* found to possess antiseptic, antimicrobial, anti-haemorrhagic, cicatrizant, and anti-inflammatory properties, computational peptidology applied to determination of the chemical reactivity and bioactivity properties of plant cyclopeptides isolated from *Rosaceae*, and antioxidant *in silico* prediction of moracin-C and iso-moracin-C isomers against the OOH free radical. *Annona muricata*, a plant that has been reported to demonstrate significant antiviral properties against the human immunodeficiency virus, herpes simplex virus, human papilloma virus, hepatitis C virus and dengue virus, to *in silico* search of GSK3-β inhibitors as a potential Alzheimer's disease treatment.

FUTURE PROSPECTS

Thanks to the studies shown in this section, we hope that more researchers will join the search for new properties of biomolecules extracted from different parts of endemic plants. In this sense, Latin America, particularly South America, has enormous potential due to its rich biodeversity that invites scientists to continue the search of these biomolecules with biomedical and biotechnological application.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

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.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 27 January 2021
doi: 10.3389/fchem.2020.624716

Evaluation of *Annona muricata* Acetogenins as Potential Anti-SARS-CoV-2 Agents Through Computational Approaches

Shashanka K. Prasad[1], Sushma Pradeep[1], Chandan Shimavallu[1], Shiva Prasad Kollur[2]*, Asad Syed[3], Najat Marraiki[3], Chukwuebuka Egbuna[4,5,13], Mihnea-Alexandru Gaman[7,8], Olga Kosakowska[9], William C. Cho[10], Kingsley Chukwuemeka Patrick-Iwuanyanwu[5,13], Joaquín Ortega Castro[11], Juan Frau[11], Norma Flores-Holguín[12] and Daniel Glossman-Mitnik[12]**

[1]Department of Biotechnology and Bioinformatics, School of Life Sciences, JSS Academy of Higher Education and Research, Mysuru, India, [2]Department of Sciences, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Mysuru, India, [3]Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, Saudi Arabia, [4]Department of Biochemistry, Faculty of Natural Sciences, Chukwuemeka Odumegwu Ojukwu University, Anambra State, Nigeria, [5]Department of Biochemistry, University of Port Harcourt, Choba, Nigeria, [6]Faculty of Medicine, Carol Davila University of Medicine and Pharmacy, Bucharest Romania, [7]Center of Hematology and Bone Marrow Transplantation, Fundeni Clinical Institute, Bucharest, Romania, [8]Center of Hematology and Bone Marrow Transplantation, Fundeni Clinical Institute, Bucharest, Romania, [9]Department of Vegetable and Medicinal Plants, Institute of Horticulture Sciences, Warsaw University of Life Sciences - SGGW, Warsaw, Poland, [10]Department of Clinical Oncology, Queen Elizabeth Hospital, Kowloon, Hong Kong SAR, China, [11]Departament de Química, Universitat de les Illes Balears, Palma de Malllorca, Spain, [12]Laboratorio Virtual NANOCOSMOS, Departamento de Medio Ambiente y Energía, Centro de Investigación en Materiales Avanzados, Chihuahua, Mexico, [13]World Bank Africa Centre of Excellence for Public Health and Toxicological Research (ACE-PUTOR), University of Port Harcourt, Nigeria*

Annona muricata, a tropical plant which has been extensively used in ethnomedicine to treat a wide range of diseases, from malaria to cancer. Interestingly, this plant has been reported to demonstrate significant antiviral properties against the human immunodeficiency virus, herpes simplex virus, human papilloma virus, hepatitis C virus and dengue virus. Additionally, the bioactive compounds responsible for antiviral efficacy have also shown to be selectively cytotoxic while inhibiting tumorigenic cell growth without affecting the normal cell growth. Annonaceous Acetogenins are a class of bioactive compounds exclusive to the Annonaceae family at which the plant *A. muricata* belongs. In the current study, we have created a library of Acetogenins unique to the plant, comprising of Annomuricin A, Annomuricin B, Annomuricin C, Muricatocin C, Muricatacin, *cis*-Annonacin, Annonacin-10-one, *cis*-Goniothalamicin, Arianacin and Javoricin, for *in silico* and theoretical evaluations against the SARS-CoV-2 spike protein in an attempt toward promotion of plant based drug development for the current pandemic of coronavirus disease 2019 (COVID-19). We found that all the Acetogenins showing *in silico* spike protein significantly docking with good binding affinities. Moreover, we envision *A. muricata* Acetogenins can be further studied by *in vitro* and *in vivo* models to identify potential anti-SARS-CoV-2 agents.

Keywords: Annonaceous acetogenins, COVID-19, *in silico*, molecular docking, molecular dynamics simulation study, *Annona muricata*

OPEN ACCESS

Edited by:
Jorge M. del Campo, National Autonomous University of Mexico, Mexico

Reviewed by:
Alejandro Ramirez-Solis, Universidad Autónoma del Estado de Morelos, Mexico
Carolina Mascayano, University of Santiago, Chile

****Correspondence:***
Chandan Shimavallu
chandans@jssuni.edu.in
Shiva Prasad Kollur
shivachemist@gmail.com
Daniel Glossman-Mitnik
daniel.glossman@cimav.edu.mx

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *01 November 2020*
Accepted: *18 December 2020*
Published: *27 January 2021*

Citation:
Prasad SK, Pradeep S, Shimavallu C, Kollur SP, Syed A, Marraiki N, Egbuna C, Gaman M-A, Kosakowska O, Cho WC, Patrick-Iwuanyanwu KC, Ortega Castro J, Frau J, Flores-Holguín N and Glossman-Mitnik D (2021) Evaluation of Annona muricata *Acetogenins as Potential Anti-SARS-CoV-2 Agents Through Computational Approaches. Front. Chem. 8:624716. doi: 10.3389/fchem.2020.624716*

1 INTRODUCTION

Annona muricata (L.), commonly called Soursop (English) or Lakshmanaphala (Kannada), is a tropical plant known for its wide range of applications in ethnomedicinal practices (Moghadamtousi et al., 2015; Gavamukulya et al., 2017; Prasad et al., 2019). Belonging to the Annonaceae family, *A. muricata* has grabbed the attention of global scientific communities toward investigating the medicinal significance of its constituent phytochemicals. Found exclusive to the Annonaceae family, the annonaceous Acetogenins are one class of phytochemicals that have been investigated thoroughly for their potent biomedical properties. Chemically, Acetogenins are the metabolic derivatives of long-chain fatty acids derived via the polyketide pathway (Prasad et al., 2019). Over 500 Acetogenins have been reported across the family (Moghadamtousi et al., 2015; Liaw et al., 2016). Acetogenins have been reported to show significant antiviral activities against herpes simplex virus - I (HSV-I) (Padma et al., 1998), herpes simplex virus - II (HSV-II) (Betancur-Galvis et al., 1999), human papillomavirus (HPV) (Donne et al., 2017), hepatitis C virus (HCV) (Apriyanto et al., 2018), dengue virus type 2 (DENV-2) (Wahab et al., 2018), human immunodeficiency virus - I (HIV-I) (van de Venter et al., 2014; Gavamukulya et al., 2017). The above observations confirm the broad-spectrum antiviral activity of *A. muricata* phytochemicals, which may also be responsible for the other biomedical properties. In this first combination of theoretical and computational studies on annonaceous Acetogenins, we have created a library of Annomuricin A, Annomuricin B, Annomuricin C, Muricatocin C, Muricatacin, *cis*-Annonacin, Annonacin-10-one, *cis*-Goniothalamicin, Arianacin and Javoricin, to investigate their *in silico* anti-SARS-CoV-2 activity. The above compounds are unique to *A. muricata* and are the only reported Acetogenins with potent anti-breast cancer activity. This study provides evidence for drug development exercises toward mitigation of the current pandemic, coronavirus disease 2019 (COVID-19) caused by the SARS-CoV-2 virus.

Chemoinformatics is a collection of different procedures for the study and organization of chemical information, it is a useful approach for the development of new medical drugs in the pharmaceutical industry. Chemoinformatics applications are valuable for the prediction of the molecular properties of many systems based on the knowledge of previously studied molecular structures and computational displaying of the same by considering the close relationship between biological data and organic structures and energies. This is the way that molecular descriptors can be related with molecular properties (Guha and Bender, 2012).

There are numerous *in silico* studies pertaining to the understanding of standard drug(s)/phytochemical(s) binding to SARS-CoV-2 surface proteins, most of which are considered a potential drug targets (Boopathi et al., 2020). Gupta et al. reported similar ADME properties apart from the 'Lipsinki Rule of 5′ for phytochemicals Belachinal, Macaflavanone E and Vibsanol B, all of which interacted mainly with the VAL25 and PHE26 aminThuso acids to facilitate the binding to SARS-CoV-2 E protein Gupta et al. (2020). In an alternative approach, Sarma et al. suggested the targeting of coronavirus nucleocapsid (N) protein RNA-binding N-terminal domain (NTD). In this study, Two suitable binders, one theophylline derivative and one 3,4-dihydropyrimidone class molecule, were identified to target the 2OFZ, an ultrahigh resolution crystal structure of RNA binding domain of CoV N protein of resolution 1.1 Å in monoclinic form Sarma et al. (2020). Meanwhile, another study focused on understanding the interaction of FDA approved drugs such as remdesivir, saquinavir and duranavir, as potent chymotrypsin-like protease ($3CL^{pro}$) of the SARS-CoV-2 virus in an attempt to provide for drug repurposing and expediting the drug discovery process (Khan et al., 2020b). Similarly, $3CL^{pro}$ as well as 2′-O-ribosemethyltransferase (2′-O-MTase) were reported druggable targets by Khan et al. for systematic drug repurposing of paritaprevir and raltegravir for $3CL^{pro}$ and dolutegravir and bictegravir for 2′-O-MTase (Khan et al., 2020a). The SARS-CoV-2 spike glycoprotein has been extensively reviewed as a potential target for COVID-19 therapeutics (Basit et al., 2020). The above findings laid the foundation for the current study. Other studies on the efficacy of phytochemicals against SARS-CoV-2 have been conducted for instance, Gupta et al. (2020) have analyzed the SARS-CoV-2 envelope (E) protein binding and inhibition efficiency of the plant-derived Belachinal, Macaflavanone E and Vibsanol B Gupta et al. (2020).

2 MATERIALS AND METHODS

2.1 Library and Macromolecule Preparation

Coronaviruses (CoVs) are single-stranded RNA viruses that infect a wide range of hosts to cause pulmonary complications which may range from mild to severe or fatal. The surface-exposed spike glycoprotein, or S protein, plays a major role in infecting the host cell membranes to release viral RNA molecules. Thereby making it a target of choice for COVID-19 therapeutics, where the vaccines or any small agents designed to cure or control the disease must primarily target and block or inhibit the activity of these spike proteins. In this study, ten Acetogenins unique to *A. muricata* have been evaluated for the SARS-CoV-2 inhibition potential, by means of targeting the spike protein. The *in silico* molecular docking results of ten Acetogenins against the SARS-CoV-2 spike protein were compared with the molecular docking results of dexamethasone.

In this regard, the crystal structure of SARS-CoV-2 spike receptor-binding domain bound with angiotensin converting enzyme 2 (ACE2) (PDB ID- 6M0J) (**Supplementary Figure S1**) was considered and downloaded from the biological structural database of proteins, i.e. Protein Data Bank (PDB). The 3-dimensional structure of the protein was checked for the presence of any pre-existing ligands and the water molecules were deleted or removed and the resultant protein structure was saved in .pdb format for the further validation.

The protein validation was carried out using RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) that gives the Ramachandran plot analysis of the protein predicting its amino acid residue details in both the allowed and the favored regions. The protein was found to have 94.6% number of residues in the favored region, with 1.8% of residues in the allowed region and thus was presumed suitable for docking.

The binding or the active site pocket residues where the ligands interact with the protein to inhibit its action were selected based on the literature survey and the grid box was generated around the binding pocket residues present in between the coronavirus spike protein and the ACE2 receptor.

2.2 Ligand Optimization

The 3D structural files of the ten Acetogenin ligand molecules were required to carry out the SARS-CoV-2 spike protein docking interaction. For which, an initial 2D structure was sketched and geometrically cleaned using the ChemSketch freeware software by the ACD labs, Canada. The final 2D structure was saved as chemical markup language files to be further converted to 3D structure files using the OpenBabel chemical toolbox, an open-collaborative project, by generating the 3D coordinates upon addition of hydrogen atoms into the structure. Finally, the resultant pdb files were analyzed for their structure, followed by the second geometrical clean-up to add up any leftover hydrogens and resaved in the. pdb format using the ArgusLab software (A molecular modeling, graphics and drug design program) before further interaction analysis.

To perform the comparative docking studies, the .sdf format file of the approved drug for COVID-19, dexamethasone (https://pubchem.ncbi.nlm.nih.gov/compound/Dexamethasone; PubChem CID: 5743), was downloaded from PubChem, a free online database of chemical molecules (https://pubchem.ncbi.nlm.nih.gov/). Similarly, the chemical structures, in. sdf file format, of Annomuricin A (https://pubchem.ncbi.nlm.nih.gov/compound/Annomuricin-A; PubChem CID: 157682), Annomuricin B (https://pubchem.ncbi.nlm.nih.gov/compound/Annomuricin-B; PubChem CID: 44575650), Annomuricin C (https://pubchem.ncbi.nlm.nih.gov/compound/Annomuricin-C; PubChem CID: 11758463), Muricatocin C (https://pubchem.ncbi.nlm.nih.gov/compound/Muricatocin-C; PubChem CID: 44584147), Muricatacin (https://pubchem.ncbi.nlm.nih.gov/compound/Muricatacin; PubChem CID: 3035657), *cis*-Annonacin (https://pubchem.ncbi.nlm.nih.gov/compound/cis-Annonacin; PubChem CID: 10698767), Annonacin-10-one (https://pubchem.ncbi.nlm.nih.gov/compound/Annonacin-10-one; PubChem CID: 180161), *cis*-Goniothalamicin (https://pubchem.ncbi.nlm.nih.gov/compound/cis-Goniothalamicin; PubChem CID: 10722235), Arianacin (https://pubchem.ncbi.nlm.nih.gov/compound/Arianacin; PubChem CID: 10698768) and Javoricin (https://pubchem.ncbi.nlm.nih.gov/compound/Javoricin; PubChem CID: 10326193), were obtained from PubChem. The downloaded files were converted to .pdb file by following the same steps as mentioned in the ligand preparation using the OpenBabel and ArgusLab software.

2.3 Molecular Docking and Interaction Studies

The molecular docking studies for the prepared ligands structural files were carried out to evaluate their inhibition potential, against the prepared and validated SARS-CoV-2 spike protein, using AutoDock Vina in PyRx an open-source virtual screening tool (https://pyrx.sourceforge.io/). Initially, the above obtained 3D structural files (.pdb) of the protein were uploaded alongside the ligand file (.pdb) of choice. The loaded .pdb files were generated into .pdbqt files comprising of whole protein molecule that was prepared and validated earlier with added hydrogen and charges along with the ligands in .pdbqt files. At this juncture, the binding site residues were selected and a grid box assigned to surround them in a manner that all the binding site residues were within the grid box. Finally, the docking process was run using the generic algorithm parameter.

The docking results were obtained in the form of conformations or poses, with eight different docking poses for every ligand, with the binding affinity values and their RMSD values against their respective protein or macromolecule. The best docked pose was saved in the .pdb format to check for their interaction against the protein.

2.4 Molecular Dynamics Simulation Studies

The ligand *cis*-Annonacin- SARS-CoV-2 spike protein complex that has obtained the highest binding affinity and has interacted very well was selected for the molecular dynamics (MD) simulation studies. The MD simulation was carried out using GROMACS, forced with CHARMM27 all-atom force field. The parameter files for the ligand required to run the simulation were generated using SwissParam online server. TIP3P water model was used to solvate the system and counter ions were added to neutralize the system. The steepest descent algorithm was used to minimize the structures and the system was equilibrated by applying position restrains by performing simulations using NVT ensembles followed by NPT. The equilibration was run for 20 picoseconds (ps) at 300 K temperature and 1 bar pressure. After equilibrating the system, the MD run was carried out for the system at 30 ns. The energies and coordinates were saved at every 10 ps for analysis purpose.

2.5 Computational Details

The determination of the conformers of the molecules considered in this study was performed by resorting to MarvinView 17.15 available from ChemAxon (Budapest, Hungary) by doing Molecular Mechanics calculations through the overall MMFF94 force field (Halgren, 1996a; Halgren, 1996b; Halgren, 1996c; Halgren and Nachbar, 1996; Halgren, 1999). This was followed by a geometry reoptimization by means of the Density Functional Tight Binding (DFTB) methodology (Frisch et al., 2016). The electronic properties of the studied molecules involved the use of MN12SX/Def2TZVP/H2O model chemistry on the previously optimized molecular structures because it has been shown that it allows the verification of the 'Koopmans in DFT' (KID) procedure (Flores-Holguín et al., 2019a; Flores-Holguín

TABLE 1 | Binding affinity values of the docked ligand and number of hydrogen bonds formed.

Compound	Binding affinity values (Kcal/mol)	Number of hydrogen bonds
Dexamethasone	−7.5	6
Annomuricin A	−6.8	7
Annomuricin B	−6.7	4
Annomuricin C	−5.9	9
Muricatocin C	−6.4	7
Muricatacin	−5.3	3
cis-Annonacin	−7.7	9
Annonacin-10-one	−6.5	6
cis-Goniothalamicin	−6.3	6
Arianacin	−6.7	3
Javoricin	−6.3	7

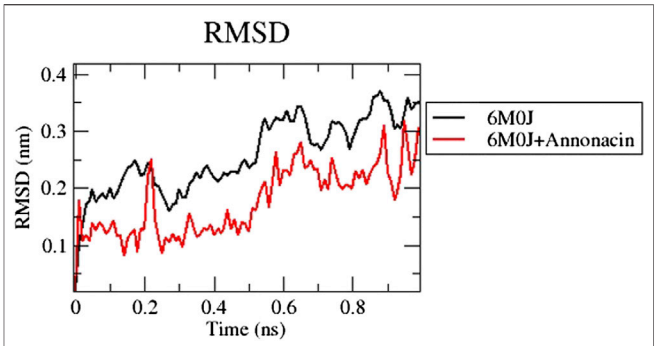


FIGURE 1 | Analysis of RMSD of 6M0J alone and complexed with the ligand over 10 ns simulation run.

et al., 2019b) with the aid of the Gaussian09 software (Frisch et al., 2016) and the SMD solvation model (Marenich et al., 2009).

3 RESULTS AND DISCUSSION

3.1 Molecular Docking Studies

To analyze the results, the best docked pose of the ligand and the protein were opened using PyMol and the obtained merged protein-ligand structures were saved in. pdb file to obtain a protein-ligand complex. The same thing was carried out for all the ten ligands including the standard Dexamethasone drug. The formation of bonded and non-bonded interactions of the ligands with the respective binding site amino acids of the protein were obtained using an online tool Protein-Ligand Interaction Profiler (https://projects.biotec.tu-dresden.de/plip-web/plip/index). The obtained PyMol session file was saved and the resultant interaction (hydrophobic interaction and hydrogen bonds) table was also saved, the dashed lines indicate hydrophobic interactions while the blue lines indicate hydrogen bonds formed between the amino acid residues of the protein with the ligands.

3.2 Ligands - Spike Protein Interaction

To explore the possible binding patterns of all the ligands selected for the study and interactions of the selected phytocompounds, Molecular Docking approaches were employed. The ten ligands along with the drug used for PyRx docking were evaluated on the basis of the obtained binding affinity results and their values interpreted on the basis of the strength which the ligand has bound to the macromolecule. The results were analyzed based on their interaction in the active site of the selected viral proteins. All the ten ligands were seen as interacting very well with the spike protein by forming binding affinity energies within the range of −5.3 to −7.7 kcal/mol whereas the drug molecule dexamethasone exhibited a binding affinity of −7.5 kcal/mol against the spike protein. The highest affinity of −7.7 kcal/mol interaction toward the protein was shown by *cis*-Annonacin which is better than the reference drug, while the lowest binding affinity was shown by Muricatacin. The compound *cis*-Annonacin showed the highest binding affinity of −7.7 kcal/mol with the binding site of the protein sharing five hydrogen bonds with PHE-390A, ARG-393A, GLN-409E, GLY-496E, TYR-505E and two each with ARG-403E and TYR-453E along with more hydrophobic interactions with ASN-33A, GLU-37A, PRO-389, ARG-403E, GLU-406E, LYS-417E, TYR-495E, PHE-497E and 2 with TYR-505E, while the lowest binding energy was shown by Muricatacin. The overall interaction of *cis*-Annonacin was shown to be comparatively better than the reference drug. The bond details of each ligand with the spike protein are shown in **Table 1** (**Supplementary Figures S2–S12**).

3.3 Molecular Dynamics Simulation Study

MD simulation study was performed for the protein-ligand complex to explain their dynamic behavior. The spike protein-*cis*-Annonacin ligand interaction having minimum binding energy (strong binding affinity) was considered for further evaluation using molecular dynamics simulation to check their stability in complexes under simulated conditions. The simulation study provides the analysis of root mean square deviation (RMSD), radius of gyration (Rg), solvent accessible surface area (SASA), number of hydrogen bonds maintained throughout the simulation time and variation of secondary structure pattern between the protein and their complexes. The simulation was performed at the time duration of 10 ns with the native protein alone and in complex with ligand Annonacin.

The RMSD plot in **Figure 1** shows that the protein-ligand complex reached equilibrium approximately at 0.3–0.4 ns time and remaining showed stable trajectory with minimal deviation in 0.1–0.15 nm RMSD range, the protein structural flexibility are being reserved while it is in free form in complex with ligand. After the initial fluctuations the ligand bound to the protein reached equilibrium.

The radius of gyration (Rg) considers the varied masses calculated to root mean square distances considering the central axis of rotation. The Rg plot (**Figure 2** considers the capability, shape and folding during the each time step of the whole trajectory throughout the simulation. The protein and their respective ligand complex exhibited the

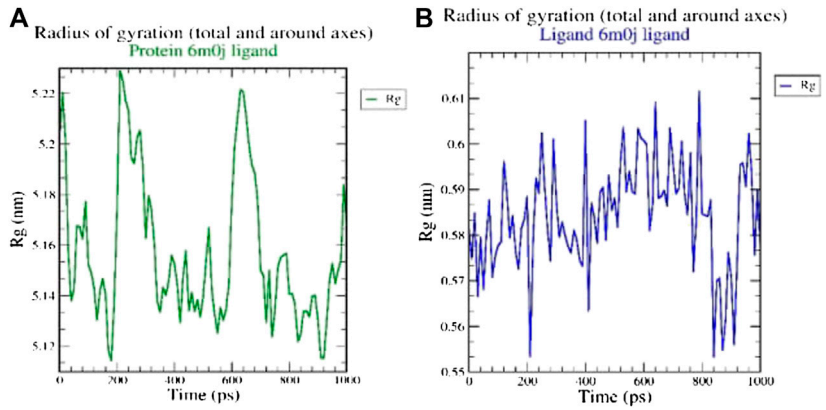


FIGURE 2 | Analysis of radius of gyration (Rg) of **(A)** 6M0J alone and **(B)** complex with the ligand over 10 ns simulation run.

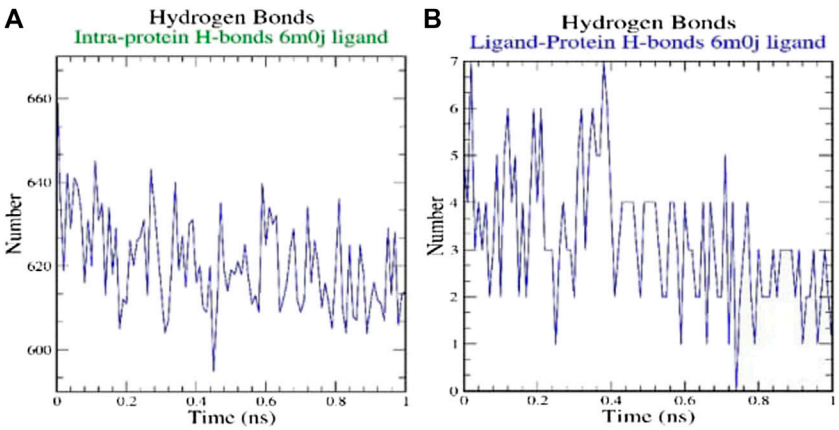


FIGURE 3 | Analysis of hydrogen bonds on **(A)** 6M0J alone and **(B)** complex with the ligand over 10 ns simulation run.

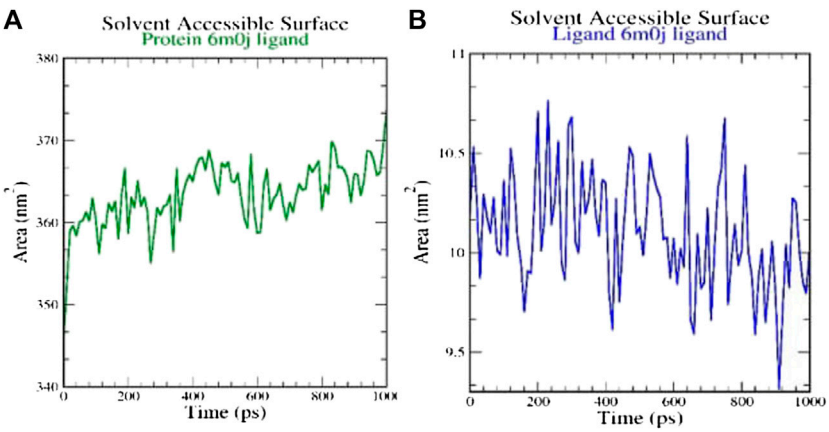


FIGURE 4 | Analysis of solvent accessible surface area on **(A)** 6M0J alone and **(B)** complex with the ligand over 10 ns simulation run.

similar pattern of Rg value with the deviation in the range of 0.55–5.22 nm.

H-bonds which appear during the Molecular Docking study being analyzed over the total simulation period. All the intermolecular H-bonds among ligands and protein were only considered during the analysis and plotted accordingly (**Figure 3**). It is evident from the plot that the number of H-bonds during the simulation runs remains consistent with Molecular Docking study, while few bonds were simultaneously broken and rebuilt.

SASA measures the area around hydrophobic core formed between protein-ligand complexes (**Figure 4**). Consistent SASA values were observed.

4 CONCLUSION

To sum up, for the A. muricata phytochemicals with known potent antioxidant and anti-inflammatory properties, in silico studies suggest that the Annonaceous Acetogenins show a good inhibition activity against the SARS-CoV-2 spike protein as per the Molecular Docking and MD simulation results obtained. Notwithstanding, *cis*-Annonacin was identified to be the most potent acetogenin with a low binding energy (indicative of the highest binding affinity) and greater hydrogen bond formation potential. This information could be useful in complementing the experimental data as the starting point for the development of new therapeutic drugs based on these molecules.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

SKP Research; SP: Research; CS: Research; SPK: Research and Writing the Manuscript; AS: Research; NM: Research; CE: Research; M-AG: Research; OK: Research; WC: Research; KCP-I: Research; JOC: Research; JF: Research; NF-H: Research; DG-M: Research and Writing the Manuscript.

FUNDING

Researchers Supporting Project number (RSP-2020/201), King Saud University, Riyadh, Saudi Arabia.

ACKNOWLEDGMENTS

KSP thankfully acknowledges the Director, Amrita Vishwa Vidyapeetham, Mysuru campus for infrastructure support. The authors extend their appreciation to the Researchers Supporting Project number (RSP-2020/201), King Saud University, Riyadh, Saudi Arabia for the financial assistance. SKP, CS and SP acknowledge the support and infrastructure provided by the JSS Academy of Higher Education and Research (JSSAHER), Mysuru, India. DG-M is a Visiting Professor at the Universitat de les Illes Balears from which support is gratefully acknowledged. Moreover, NF-H and DG-M want to thank CIMAV and CONACYT for partial support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2020.624716/full#supplementary-material.

REFERENCES

Apriyanto, D. R., Aoki-Utsubo, C., Hartati, S., Dewi, B. E., and Hotta, H. (2018). Potential of Indonesian plants (Annona muricata, Garcinia latissima, and Garcinia celebica) against hepatitis C virus. *Adv. Sci. Lett.* 24, 6807–6810. doi:10.1166/asl.2018.12850

Basit, A., Ali, T., and Rehman, S. U. (2020). Truncated human angiotensin converting enzyme 2; A potential inhibitor of SARS-CoV-2 spike glycoprotein and potent COVID-19 therapeutic agent. *J. Biomol. Struct. Dyn.* 2020, 1–10. doi:10.1080/07391102.2020.1768150

Betancur-Galvis, L., Saez, J., Granados, H., Salazar, A., and Ossa, J. (1999). Antitumor and antiviral activity of Colombian medicinal plant extracts. *Mem. Inst. Oswaldo Cruz.* 94, 531–535. doi:10.1590/s0074-02761999000400019

Boopathi, S., Poma, A. B., and Kolandaivel, P. (2020). Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and Rule out against its treatment. *J. Biomol. Struct. Dyn.* 2020, 1–10. doi:10.1080/07391102.2020.1758788

Donne, M. L., Lentini, M., Alibrandi, A., Salimbeni, V., Giuffre', G., Mazzeo, F., et al. (2017). Antiviral activity of ellagic acid and Annona muricata in cervical HPV related pre-neoplastic lesions: a randomized trial. *J. Funct. Foods.* 35, 549–554. doi:10.1016/j.jff.2017.06.006

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2019a). Computational peptidology assisted by conceptual density functional theory for the study of five new antifungal tripeptides. *ACS Omega*. 4, 12555–12560. doi:10.1021/acsomega.9b01463

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2019b). "Conceptual DFT as a helpful Chemoinformatics tool for the study of the clavanin family of antimicrobial marine peptides," in *Density functional theory calculations* (Rijeka, Croatia: IntechOpen), 1–11.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 09 revision E.01*. Wallingford CT: Gaussian Inc.

Gavamukulya, Y., Wamunyokoli, F., and El-Shemy, H. A. (2017). Annona muricata: is the natural therapy to most disease conditions including cancer growing in our backyard? A systematic review of its research history and future prospects. *Asian Pac. J. Trop. Med.* 10, 835–848. doi:10.1016/j.apjtm.2017.08.009

Guha, R., and Bender, A. (2012). *Computational approaches in cheminformatics and bioinformatics*. Hoboken, NJ: Wiley.

Gupta, M. K., Vemula, S., Donde, R., Gouda, G., Behera, L., and Vadde, R. (2020). In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus Envelope protein ion channel. *J. Biomol. Struct. Dyn.* 2020, 1–11. doi:10.1080/07391102.2020.1751300

Halgren, T. A. (1996a). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* 17, 490–519. doi:10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P

Halgren, T. A. (1996b). Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* 17, 520–552. doi:10.1002/(SICI)1096-987X(199604)17:5/6<520::AID-JCC2>3.0.CO;2-W

Halgren, T. A. (1996c). Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* 17, 616–641. doi:10.1002/(SICI)1096-987X(199604)17:5/6%3C616::AID-JCC5%3E3.0.CO;2-X

Halgren, T. A. (1999). MMFF VI. MMFF94s option for energy minimization studies. *J. Comput. Chem.* 20, 720–729. doi:10.1002/(SICI)1096-987X(199905)20:7%3C720::AID-JCC7%3E3.0.CO;2-X

Halgren, T. A., and Nachbar, R. B. (1996). Merck molecular force field. IV. Conformational energies and geometries for MMFF94. *J. Comput. Chem.* 17, 587–615. doi:10.1002/(SICI)1096-987X(199604)17:5/6<587::AID-JCC4>3.0.CO;2-Q

Khan, R. J., Jha, R. K., Amera, G. M., Jain, M., Singh, E., Pathak, A., et al. (2020a). Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2'-O-ribose methyltransferase. *J. Biomol. Struct. Dyn.* 2020, 1–14. doi:10.1080/07391102.2020.1753577

Khan, S. A., Zia, K., Ashraf, S., Uddin, R., and Ul-Haq, Z. (2020b). Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *J. Biomol. Struct. Dyn.* 2020, 1–10. doi:10.1080/07391102.2020.1751298

Liaw, C. C., Liou, J. R., Wu, T. Y., Chang, F. R., and Wu, Y. C. (2016). "Acetogenins from Annonaceae," in *Progress in the chemistry of organic natural products* (New York, NY: Springer International Publishing), 113–230. doi:10.1007/978-3-319-22692-7_2

Marenich, A., Cramer, C., and Truhlar, D. (2009). Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B.* 113, 6378–6396. doi:10.1021/jp810292n

Moghadamtousi, S., Fadaeinasab, M., Nikzad, S., Mohan, G., Ali, H., and Kadir, H. (2015). Annona muricata (Annonaceae): a review of its traditional uses, isolated acetogenins and biological activities. *Int. J. Mol. Sci.* 16, 15625–15658. doi:10.3390/ijms160715625

Padma, P., Pramod, N., Thyagarajan, S., and Khosa, R. (1998). Effect of the extract of Annona muricata and petunia nyctaginiflora on herpes simplex virus. *J. Ethnopharmacol.* 61, 81–83. doi:10.1016/s0378-8741(98)00013-0

Prasad, S. K., Varsha, V., and Devananda, D. (2019). Anti-cancer properties of Annona muricata (L.): a review. *Med. Plants Int. J. Phytomed. Rel. Indust.* 11, 123–134. doi:10.5958/0975-6892.2019.00016.9

Sarma, P., Shekhar, N., Prajapat, M., Avti, P., Kaur, H., Kumar, S., et al. (2020). In-silico homology assisted identification of inhibitor of RNA binding against 2019-nCoV N-protein (N terminal domain). *J. Biomol. Struct. Dyn.* 2020, 1–9. doi:10.1080/07391102.2020.1753580

van de Venter, M., Pruissen, M., Koekemoer, T., Sowemimo, A., and Govender, S. (2014). *In vitro* anti-HIV and -TB activities of Annona muricata and artemisia afra extracts. *Planta Med.* 80 (16) P1L29. doi:10.1055/s-0034-1394687

Wahab, N. Z. A., Ibrahim, N., Kamarudin, M. K. A., Lananan, F., H, J., Ghazali, A., et al. (2018). Cytotoxicity and antiviral activity of Annona muricata aqueous leaves extract against dengue virus type. *J. Fund. Appl. Sci.* 10, 580–589. doi:10.4314/jfas.v10i1s.41

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.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 21 April 2021
doi: 10.3389/fchem.2021.666647

Quantum Mechanical Predictions of the Antioxidant Capability of Moracin C Isomers

Angela Parise[1,2], *Bruna Clara De Simone*[1], *Tiziana Marino*[1], *Marirosa Toscano*[1] *and Nino Russo*[1]*

[1] *Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Rende, Italy,* [2] *Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France*

The antioxidant capability of moracin C and *iso*-moracin C isomers against the OOH free radical was studied by applying density functional theory (DFT) and choosing the M05-2X exchange-correlation functional coupled with the all electron basis set, 6-311++G(d,p), for computations. Different reaction mechanisms [hydrogen atom transfer (HAT), single electron transfer (SET), and radical adduct formation (RAF)] were taken into account when considering water- and lipid-like environments. Rate constants were obtained by applying the conventional transition state theory (TST). The results show that, in water, scavenging activity mainly occurs through a radical addition mechanism for both isomers, while, in the lipid-like environment, the radical addition process is favored for *iso*-moracin C, while, redox- and non-redox-type reactions can equally occur for moracin C. The values of pKa relative to the deprotonation paths at physiological pH were predicted in aqueous solution.

Keywords: moracin, antioxidants, DFT, kinetic constants, reaction mechanisms

OPEN ACCESS

Edited by:
Jorge Ignacio Martínez-Araya, Andres Bello University, Chile

Reviewed by:
José Pedro Cerón-Carrasco, Catholic University San Antonio of Murcia, Spain
Daniel Glossman-Mitnik, Advanced Materials Research Center (CIMAV), Mexico

***Correspondence:**
Nino Russo
nrusso@unical.it

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *10 February 2021*
Accepted: *25 March 2021*
Published: *21 April 2021*

Citation:
Parise A, De Simone BC, Marino T, Toscano M and Russo N (2021) Quantum Mechanical Predictions of the Antioxidant Capability of Moracin C Isomers. Front. Chem. 9:666647. doi: 10.3389/fchem.2021.666647

INTRODUCTION

In the last decades, 2–phenyl–benzofuran-containing molecules, found in a variety of plants (*Morus alba, Artocarpus champeden, Erythrina addisoniae, and Calpocalyx dinklagei*) (Hakim et al., 1999; Na et al., 2007; Naik et al., 2015; Kapche et al., 2017; Pel et al., 2017), have attracted considerable interest both for their massive use in pharmacology and for their ancient use in traditional medicine in Asia, Africa, and America (Fashing, 2001; Venkatesh and Seema, 2008; Kapche et al., 2009; Kuete et al., 2009). A rich source of natural products with a 2–phenyl–benzofuran basic scaffold is the *Moraceae* family (e.g., *M. alba, Morus mesozygia, Morus lhou,* and *Morus macroura*) (Sang-Hee et al., 2002), from which more than 24 molecules (moracin A–Z) have already been isolated and characterized (Nguyen et al., 2009). Many of them showed a variety of biological and pharmacological activities and were tested as potent antioxidants (Kapche et al., 2009; Seong et al., 2018), anti-cancer agents (Nguyen et al., 2009), anti-inflammatories, and anti-microbial agents (Kuete et al., 2009; Zelová et al., 2014; Lee et al., 2016). Furthermore, they were proven to act as cholinesterase (Delogu et al., 2016; Seong et al., 2018) and β-site amyloid precursor protein cleaving enzyme 1 (BACE1) (Jeon et al., 2007; Seong et al., 2018) inhibitors *in vitro*.

In particular, moracin C {2–[3′,5′-dihydroxy−4′-(3–methylbut−2–enyl)phenyl]−6–hydroxybenzofuran} and its *iso*–moracin C isomer {2–[3′,5′-dihydroxy−4′-(3–methylbut−1–enyl)phenyl]−6–hydroxybenzofuran} (see **Figure 1**), extracted from *M. alba* and *Artocarpus heterophyllus*, exhibit antioxidant capabilities (Li et al., 2018; Seong et al., 2018) and other biological

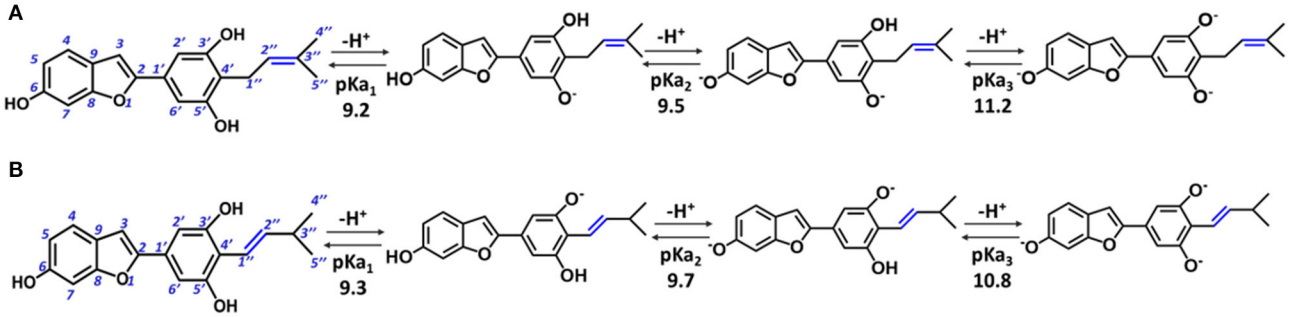


FIGURE 1 | Structure of moracin C and *iso*-moracin C, and pKa values of the relative deprotonation paths at physiological pH. **(A)** Moracin C. **(B)** Iso-moracin C.

functions correlated with oxidative stress (Zelová et al., 2014; Naik et al., 2015; Li et al., 2018; Seong et al., 2018). The only structural difference between the two isomers is the position of the C=C double bond in the methylbut–enyl moiety (see **Figure 1**). This apparent small structural difference may have significant consequences on the electronic and reactivity properties of the two isomers. In fact, when the C=C bond is close to the phenyl ring (as occurs in *iso* moracin), the electronic delocalization between the two groups increases, stabilizing accordingly the radical that is formed as a result of O–H abstraction reaction. On the contrary, the localization of the double bond in position 2″ prevents conjugation with the phenolic ring and, in principle, would favor radical attack reactions.

Very recently, in an accurate experimental study (Li et al., 2018), the authors attempted to correlate the estimated antioxidant properties with the position of the C=C bond in the two isomers, concluding that "[B]oth moracin C and *iso*–moracin C can inhibit ROS, likely through redox-related pathways (especially ET and H^+-transfer) and a non-redox-related RAF pathway. In the redox-related pathways, a double bond at the conjugation position can enhance the ET and H^+-transfer potential. However, in the non-redox-related pathway, the double bond position hardly affected the RAF potential."

We have conducted an accurate theoretical study on the thermodynamic and kinetic properties of moracin C and *iso*-moracin C when reacting with the OOH free radical by considering the following most common antioxidant scavenging reaction mechanisms (Leopoldini et al., 2011; Alberto et al., 2013; Mazzone et al., 2015; Galano et al., 2016; Markovic et al., 2016; Ahmadi et al., 2018; Castaneda-Arriaga et al., 2020; Romeo et al., 2020):

HAT: hydrogen atom transfer

$$\rightarrow H_3X + R^{\cdot} \rightarrow H_2X^{\cdot} + RH$$
$$\rightarrow H_2X^{-} + R^{\cdot} \rightarrow HX^{-\cdot} + RH$$

RAF: radical adduct formation

$$\rightarrow H_3X + R^{\cdot} \rightarrow [H_2X - RH]^{\cdot}$$
$$\rightarrow H_2X^{-} + R^{\cdot} \rightarrow [HX - RH]^{-\cdot}$$

SET: single electron transfer

$$\rightarrow H_3X + R^{\cdot} \rightarrow H_3X^{+\cdot} + R^{-}$$
$$\rightarrow H_2X^{-} + R^{\cdot} \rightarrow H_2X^{\cdot} + R^{-}$$

COMPUTATIONAL DETAILS

All calculations were performed with the Gaussian 09 code (Frisch et al., 2009) by applying the density functional theory. Following a well-consolidated protocol that was proven to be reliable in a large amount of antioxidant systems (Galano et al., 2016; Pérez-González et al., 2020), the M05-2X functional (Zhao et al., 2006) and the all electron basis set, 6-311++G(d,p) were chosen for all computations. Geometry optimization without any constraint was followed by frequency calculations to verify if the obtained structures were local minima (0 imaginary frequency) and transition states (TSs) (1 imaginary frequency) and to obtain zero-point energy corrections. Furthermore, for the TSs, it was verified that the imaginary frequency matched with the expected motion along the reaction coordinate. The solvation model based on density (SMD) (Marenich et al., 2009) was used to mimic the aqueous and lipid-like environments (water and pentyl ethanoate, respectively). Intrinsic reaction coordinate computations were performed to verify if the intercepted TSs properly connected to the relative minima in a given path.

Relative energies were computed with respect to the sum of separate reactants, and the thermodynamics corrections at 298.15 K were taken into account following the quantum mechanics-based test for the overall free radical scavenging activity (QM-ORSA) procedure (Galano and Alvarez-Idaboy, 2013, 2019). Rate constants, *k*, were determined by applying the conventional transition state theory (TST) at the 1M standard state (Truhlar et al., 1996). For the mechanism involving SETs, the barriers of reaction were computed using the Marcus theory (Marcus, 1957). For rate constants, close to the diffusion limit, the Collins–Kimball theory (Collins and Kimball, 1949) was applied.

RESULT AND DISCUSSION

For the study in water environment, knowledge of the acid-base equilibria under physiological conditions (pH = 7.4) is

TABLE 1 | pKa value and molar fractions (Mf) of the different acid–base species of moracin C and *iso*-moracin C, at physiological pH.

Molecule	pKa_1	pKa_2	pKa_3	Mf (H_3X)	Mf $(H_2X)^-$	Mf $(HX)^{2-}$	Mf $(HX)^{3-}$
Moracin *C*	9.2	9.5	11.2	9.8×10^{-1}	1.6×10^{-2}	1.2×10^{-4}	2.0×10^{-8}
Iso-moracin *C*	9.3	9.7	10.8	9.9×10^{-1}	1.2×10^{-2}	6.2×10^{-5}	2.5×10^{-8}

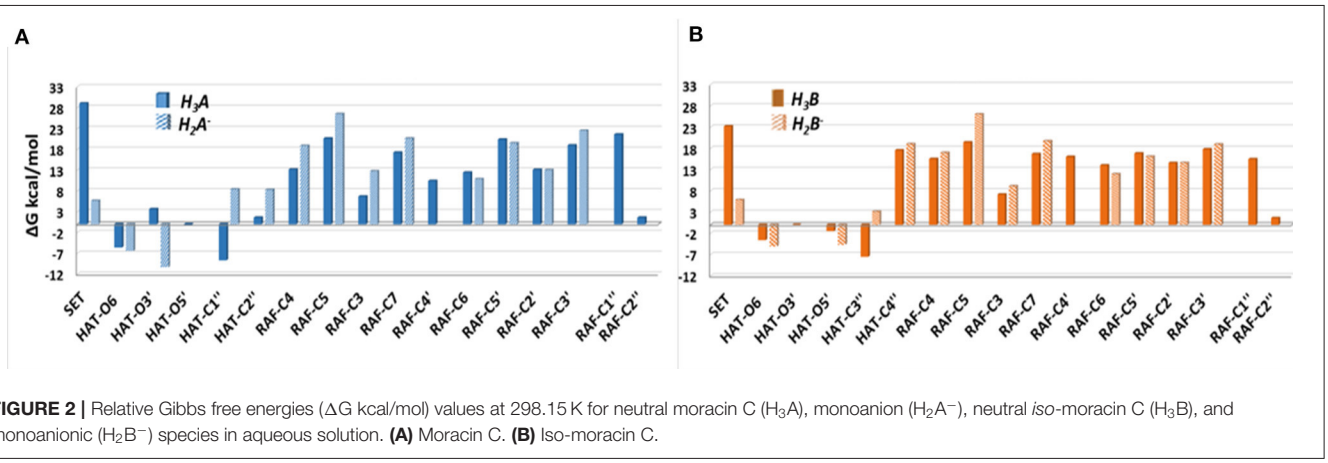


FIGURE 2 | Relative Gibbs free energies (ΔG kcal/mol) values at 298.15 K for neutral moracin C (H_3A), monoanion (H_2A^-), neutral *iso*-moracin C (H_3B), and monoanionic (H_2B^-) species in aqueous solution. **(A)** Moracin C. **(B)** Iso-moracin C.

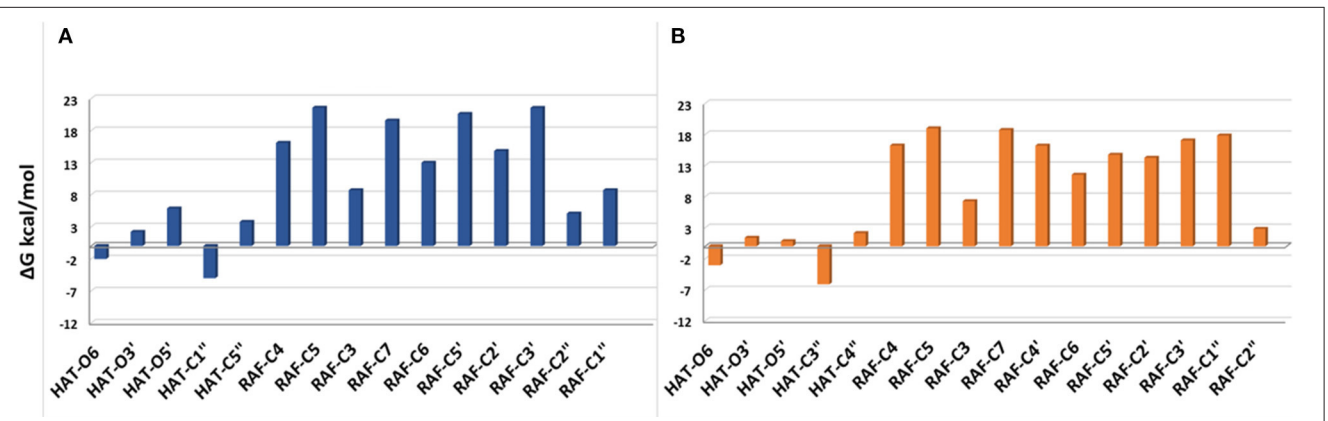


FIGURE 3 | Gibbs free energies of reaction (ΔG kcal/mol) at 298.15 K for neutral moracin C (H_3A) and *iso*-moracin C (H_3B) in pentyl ethanoate solvent. **(A)** Moracin C. **(B)** Iso-moracin C.

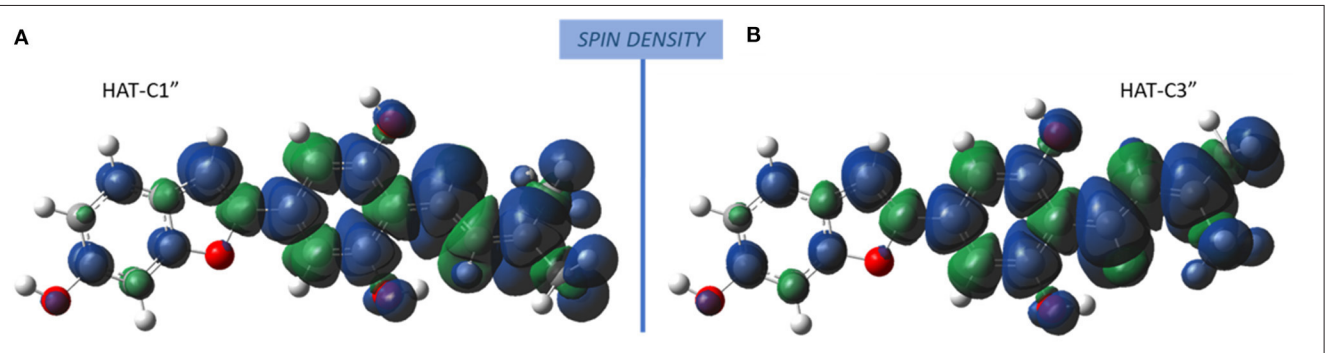


FIGURE 4 | Spin density distribution in the radical obtained after the HAT process at C1″ (moracin) and C3″ (*iso*-moracin) sites. **(A)** Moracin C. **(B)** Iso-moracin C.

very important. Because of lack of experimental information on both studied isomers, the relative pKa values were obtained (**Table 1**) using the parameters fitting method, which was previously proven to give results that are in good agreement with experimental data (Pérez-González et al., 2018). The deprotonation path of the two study systems is shown in **Figure 1**. The preferred deprotonation site in moracin C is the OH in the C5′ position, followed by those in C6 and C3′. On

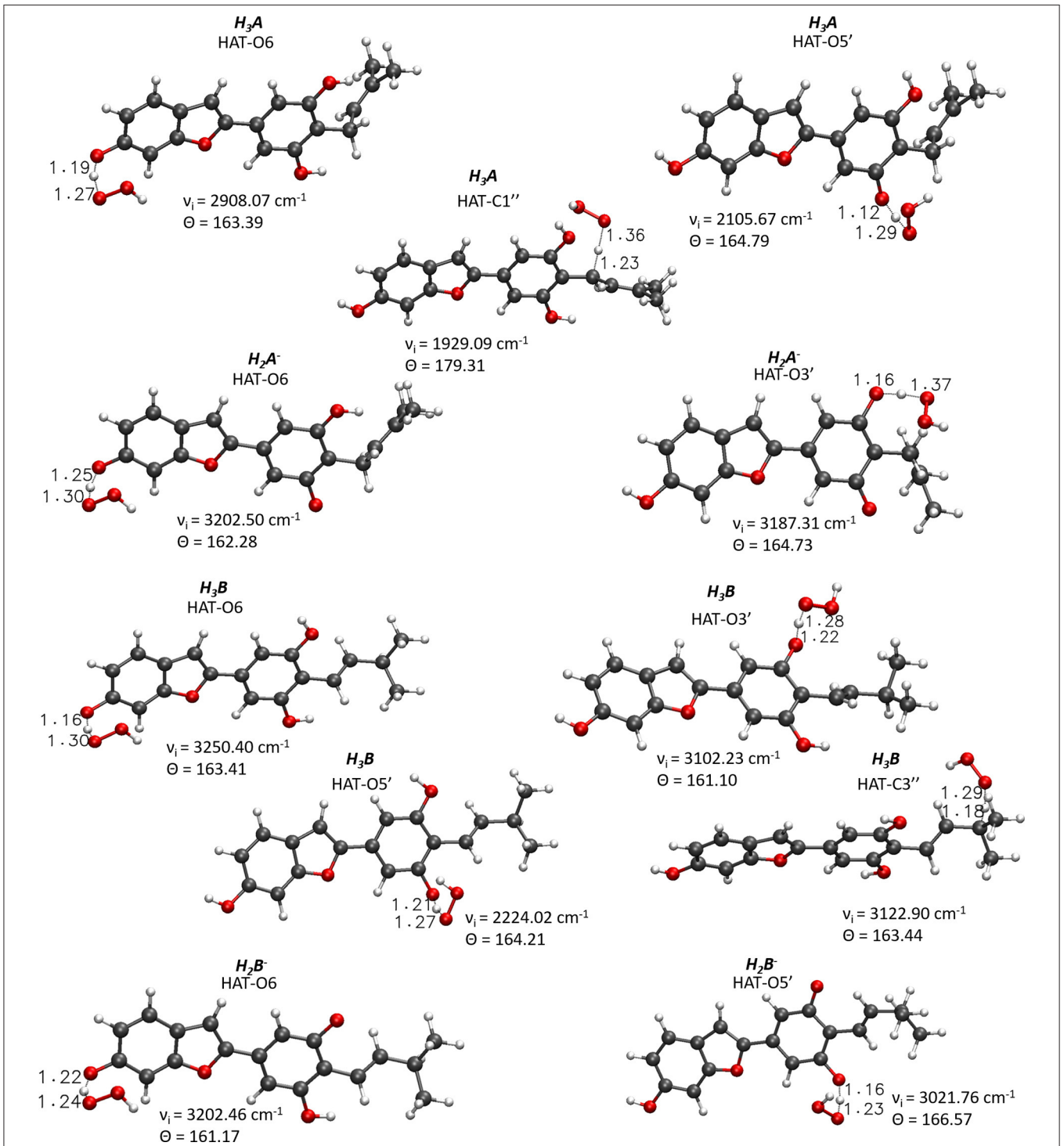


FIGURE 5 | Main geometrical parameters for the optimized TSs structures for the neutral and monoanionic species of moracin C and *iso*-moracin C involved in the HAT mechanism. Bond lengths are in Å, angles in degrees, and imaginary frequencies in cm^{-1}.

TABLE 2 | Gibbs free energies of reaction (ΔG) and activation (ΔG‡ kcal/mol) at 298.15 K in aqueous solution for neutral and monoanion moracin C and *iso*-moracin C species for the considered mechanisms.

Mechanism	H_3A		H_2A^-		H_3B		H_2B^-		H_3A^{PE}		H_3B^{PE}	
	ΔG	ΔG‡	ΔG	ΔG‡	ΔG	ΔG‡	ΔG	ΔG‡	ΔG	ΔG‡	ΔG	ΔG‡
SET	29.03		5.65		23.21		5.89					
HAT-O6	−5.62	19.78	−6.43	19.81	−3.62	21.04	−5.12	19.52	−0.02	17.23	−3.05	15.39
HAT-O3′	3.63		−10.32	17.04	0.01	19.86					0.84	17.23
HAT-O5′	−0.14	20.94			−1.81	26.27	−4.71	18.15				
HAT-C1″	−8.66	19.06							−5.02	6.08		
HAT-C3″					−7.52	17.85					−6.12	11.95

Apex PE refers to the neutral moracin C and iso-moracin C in pentyl ethanoate solvent.

the contrary, in *iso*-moracin C, the preferred deprotonation site is the OH in the C3′ position, while the second and the third ones involve sites C6 and C5′, respectively. In both conformers, all deprotonation sites are found in the benzene ring. A look at the molecular electrostatic potential, whose maps are reported in **Supplementary Figure 1**, shows that, in the case of *iso*-moracin C, the presence of the double bond in position C1′-C2′ increases the π delocalization as proven by a great negative charge on the oxygen of the hydroxyl group on C5′ position. The charge distribution reported in **Supplementary Table 1** further underlines that the localization of the double bond of methylbut–enyl moiety can influence the acid–base equilibrium of the two isomers. The calculated pKa values, at pH = 7.4 (see **Table 1**), indicate that, for both isomers, the neutral species are prevalent (molar fractions are 0.98 and 0.99 for moracin and *iso*-moracin, respectively). The monoanionic forms were not negligible in both isomers (see **Supplementary Figure 2**), so the H_3X and H_2X^- species were considered in the water environment study.

The Gibbs free energies of reaction (ΔG), computed for the two investigated mechanisms in water and lipid-like environments, are reported in **Figures 2**, **3**. As can be seen, for both molecules and environments, ΔG values for the RAF mechanism are all positive. However, since a recent experimental study (Li et al., 2018) suggested that this kind of mechanism might happen instead, we have also considered the addition of the OOH free radical to the C2″ sites, in which the obtained Gibbs reaction energies assume the less positive values.

Although the ΔG values obtained for SET are always positive, we have also considered this mechanism that was found active in several systems that had been previously studied (Galano et al., 2016; Castaneda-Arriaga et al., 2020; Romeo et al., 2020). From **Figure 2**, it is clear that HAT in the aqueous solution occurs preferentially at C1″, O6, and O5′ sites of the moracin C neutral form and O6 and O3′ sites of the corresponding monoanion. For *iso*-moracin, HAT is favored at C3″, O6, O5′, and O3′ sites of the neutral form and at O6 and O5′ sites of the monoanion one.

In the pentyl ethanoate solvent, where only the neutral species are present, the HAT process is favored, and the lowest ΔG values are obtained for the OOH attack at the C1″ site followed by O6 for moracin C and at C3″ and O6 for *iso*-moracin C.

The radicals obtained following the abstraction of the proton by OOH free radical have a spin density that is distributed over almost the entire molecular structure, as it is reported in **Figure 4** that the spin density plots of moracin C deprotonated on C1″ and *iso*-moracin C deprotonated on C3.″ In particular, due to the C=C double bond proximity to the phenyl ring, electron delocalization appears to be slightly more extended in *iso*-moracin C. In any case, this trend suggests good stability of the formed radical species for both molecules.

For the processes that show exergonic, almost isergonic, and moderate endergonic behaviors, we have computed the kinetic constants. To do this, it was necessary to locate the TSs to obtain the activation energies for the given reaction mechanism. The structures of all TSs obtained for the HAT process and their relative imaginary frequencies are shown in **Figure 5** for neutral and monoanionic species in the aqueous environment, while the structures of TSs of the neutral systems in the lipid-like environment are shown in **Supplementary Figure 3**. The obtained energy barriers ($\Delta G^{\#}$) are reported in **Table 2** together with the Gibbs free energies of reaction.

Inspection of the last Table reveals that, in aqueous solution, the $\Delta G^{\#}$ values of moracin C fall in the range of 19–21 kcal/mol for the neutral forms and become slightly lower for the charged ones. A similar behavior can be noted for *iso*-moracin C species. In the pentyl ethanoate solvent, the result is different, and for some sites, the barriers are sensibly smaller (e.g., 6.08 kcal/mol for the C1″ site). We would like to underline that the C–H bonds of the 3–methlbut−2-enyl in moracin C and in *iso*-moracin C are involved in the HAT process, making these two natural molecules interesting antioxidant agents.

All attempts to locate the relevant TS for the radical attack to the C2″ site for both molecules failed. However, this is not unusual since this type of radical attack often occurs without energy barriers. The structures derived from the OOH radical attack on the C2″ atom for both molecules are reported in **Supplementary Figure 4**. The C=C bond variation and atomic spin density for the moracin C–OOH and *iso*-moracin C–OOH radical adduct in both considered environments are shown in **Figure 6**, and the corresponding values are reported in **Supplementary Table 2**. An inspection of **Figure 6** shows that the addition of the OOH radical on

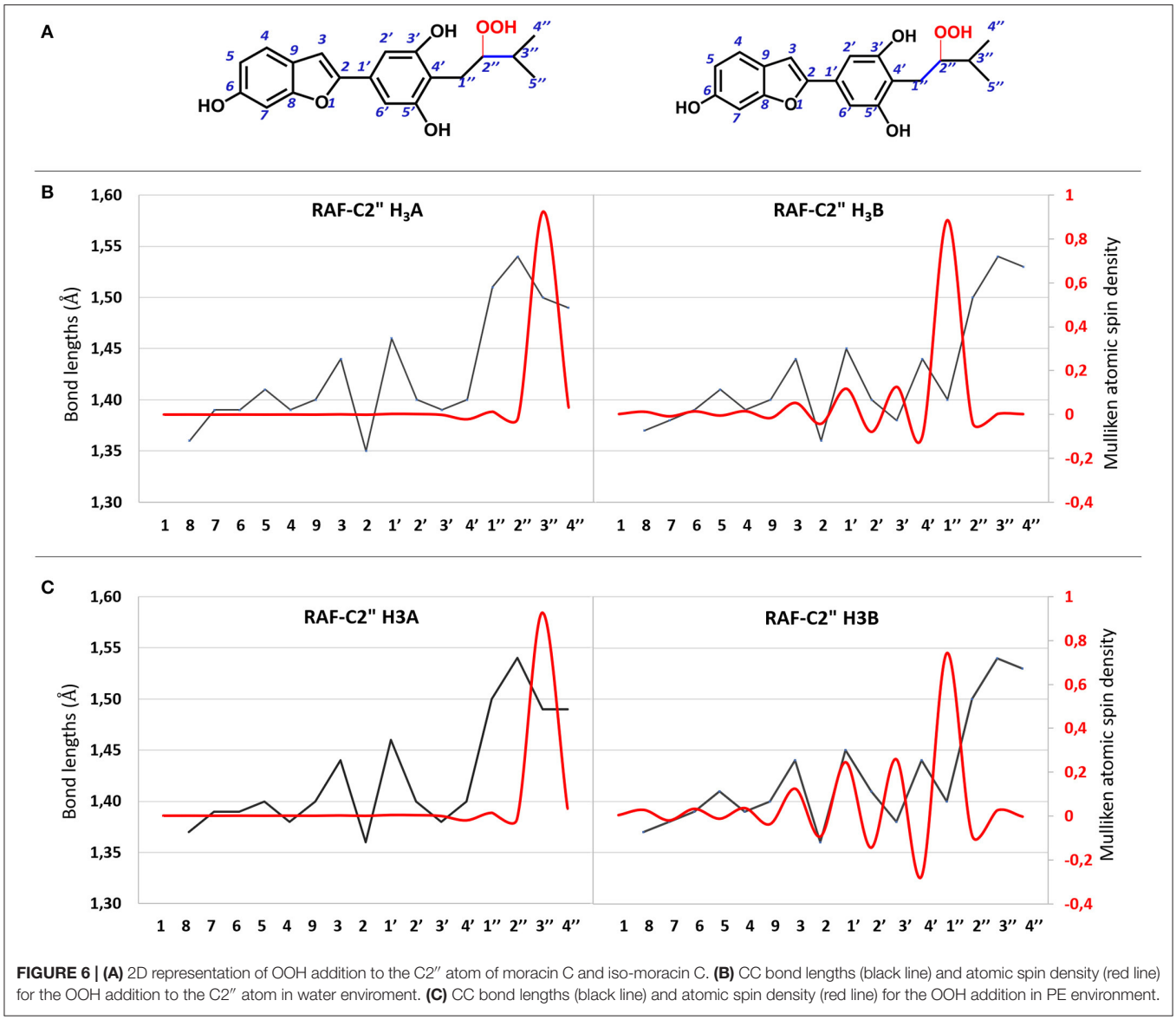


FIGURE 6 | (A) 2D representation of OOH addition to the C2″ atom of moracin C and iso-moracin C. **(B)** CC bond lengths (black line) and atomic spin density (red line) for the OOH addition to the C2″ atom in water enviroment. **(C)** CC bond lengths (black line) and atomic spin density (red line) for the OOH addition in PE environment.

the C2″ atom induces different effects in the two tautomers. In fact, in moracin C, in both the considered solvents, the spin density is essentially located at C3″ and the bond length results of C2″-C3″ needs to be elongated by assuming values close to those of a single bond (1.513 and 1.491Å in water and PE, respectively). In *iso*-moracin C, in the water environment, the addition of the radical in the same position induces a large spin density in the C1″ atom, a smaller but significant one in C3′ and C1′ atoms, and a very small density in C3 and C4 atoms. The C2″-C1″ distance is now 1.504 Å. This means that this radical, with a more extended spin density distribution result, would be more stable than the corresponding in moracin C. Similar relationships have previously been observed in other theoretical investigations on the antioxidant power of carotenoid derivatives (Ceron-Carrasco et al., 2010).

TABLE 3 | Ionization potential (IP), electron affinity (AE), electrodonating (ω-), and electroaccepting (ω+) indices of moracin C and *iso*-moracin C in water and PE (in parentheses) environments.

Molecule	IP	AE	ω^-	ω^+
Moracin C (H_3A)	5.04 (4.98)	1.32 (1.12)	4.54 (4.18)	1.36 (1.13)
Iso-moracin C (H_3B)	5.14 (4.79)	1.24 (1.25)	4.45 (4.31)	1.26 (1.29)

All values are in eV.

The computation of the electrodonating (ω^-) and electroaccepting (ω^+) values, as proposed by Gázquez et al. (2007), allows the verification of the possible correlation between these reactivity indices and the RAF antioxidant capability of the investigated molecule. Results are shown in **Table 3**. Since low values of ω^- indicate greater antioxidant activity, the

TABLE 4 | Rate constants ($M^{-1}s^{-1}$) and branching ratios (Γ) computed at the M05-2x level of theory at 298.15 K, **(A)** in aqueous and **(B)** in pentyl ethanoate solvent.

	H_3A		H_2A^-		H_3B		H_2B^-	
Mechanism	**k ($M^{-1}s^{-1}$)**	**Γ (%)**	**k ($M^{-1}s^{-1}$)**	**Γ (%)**	**k ($M^{-1}s^{-1}$)**	**Γ (%)**	**k ($M^{-1}s^{-1}$)**	**Γ (%)**
(A)								
SET	1.03×10^{-8}	∼0.00	1.83×10^{9}	100.0	1.08×10^{-9}	∼0.00	8.23×10^{8}	100.0
HAT-O6	4.57×10^{2}	∼0.00	4.45×10^{2}	∼0.0	2.45×10^{2}	∼0.00	9.74×10^{2}	∼0.0
HAT-O3'			7.49×10^{7}	∼0.0	2.79×10^{9}	∼0.00		
HAT-O5'	7.49×10^{1}	∼0.00			4.39×10^{-2}	∼0.00	3.97×10^{1}	∼0.00
HAT-C1″	2.99×10^{2}	∼0.00						
HAT-C3″					1.09×10^{3}	∼0.00		
RAF-C2″	2.15×10^{9}	100.00			2.15×10^{9}	100.00		
Total	2.15×10^{9}		1.83×10^{9}		2.15×10^{9}		8.23×10^{8}	
Overall	2.11×10^{9}		2.93×10^{7}		2.13×10^{9}		9.88×10^{6}	
	H_3A^{PE}		H_3B^{PE}					
Mechanism	**k ($M^{-1}s^{-1}$)**	**Γ (%)**	**k ($M^{-1}s^{-1}$)**	**Γ (%)**				
(B)								
HAT-O6	8.71×10^{2}	∼0.00	1.94×10^{4}	∼0.00				
HAT-O5'			3.42×10^{1}	∼0.00				
HAT-C1″	2.88×10^{9}	56.68						
HAT-C3″			1.57×10^{6}	0.07				
RAF-C2″	2.20×10^{9}	43.32	2.22×10^{9}	99.93				
Total	5.08×10^{9}		2.22×10^{9}					
Overall	4.98×10^{9}		2.20×10^{9}					

analysis of **Table 3** shows how *iso*-moracin C seems to have greater scavenging power in the aqueous environment. On the contrary, in the PE solvent, the antioxidant action of moracin C is greater. Considering the average of the values obtained in the two solvents as previously suggested by some authors (Ceron-Carrasco et al., 2012), values of ω^- being very close to each other are obtained (4.36 and 4.38 eV for moracin C and *iso*-moracin C, respectively), making it difficult to reliably predict their correlation with the antioxidant activity of the two molecules. The calculation of the kinetic constants can shed further light on the antioxidant activity of the two systems.

Using the data from **Table 2** and following the QM-ORSA computational protocol (Galano and Alvarez-Idaboy, 2013), we computed the individual, as well as the total kinetic, constants that are reported in **Table 4**.

For neutral moracin C (H_3A) in the water medium, the faster process is the RAF mechanism in the C2″ site ($k = 2.15 \times 10^9$ $M^{-1}s^{-1}$; branching ratio $\Gamma = 100\%$), while for the corresponding monoanion (H_2A^-), the SET mechanism is the faster process ($k = 1.83 \times 10^9$ $M^{-1}s^{-1}$, branching ratio $\Gamma = 100\%$). In the *iso*-moracin neutral system (H_3B), the calculated kinetic constants for RAF and HAT mechanisms are similar ($k = 2.79 \times 10^9$ $M^{-1}s^{-1}$ for HAT on the O3′ site and $k = 2.15 \times 10^9$ $M^{-1}s^{-1}$ for RAF on the C2″ atom), but the branching ratio for the former is 100%. For both the H2X- species, the SET mechanism is preferred ($k = 1.83 \times 10^9$ $M^{-1}s^{-1}$ for moracin and 8.23×10^8 $M^{-1}s^{-1}$ for *iso*-moracin).

In the PE environment, for H_3B, RAF is the preferred mechanism on the C2″ site with a k value of 2.22×10^9 $M^{-1}s^{-1}$, while, for H_3A, a competition between RAF on C2″ ($k = 2.20 \times 19^9$ $M^{-1}s^{-1}$; $\Gamma = 43.32\%$) and HAT on the C1″ site ($k = 2.88 \times 19^9$ $M^{-1}s^{-1}$; $\Gamma = 56.68\%$) mechanisms was found.

From the obtained individual and total kinetic constants, it is clear that the experimental mass spectrometric suggestion, according to which the RAF mechanism is best possible solution (Li et al., 2018), is theoretically confirmed. The data indicate that the scavenging activity in the water solution of both moracin C and *iso*-moracin neutral forms is carried out through the RAF mechanism. In the lipid-like environment, the situation appears to be different since mainly the non-redox RAF reaction on C3″ site can occur through the OOH attacking *iso*-moracin C, while moracin C can undergo the attack through both redox (SET)- and non-redox- (RAF) like reactions.

CONCLUSION

From the density functional computations on the antioxidant potential of moracin C and its isomer *iso*-moracin C, the following conclusions can be outlined:

- pKa calculations in water environment evidence that, for both systems, the neutral form is dominant with the

monoanionic species being present in lower percentage but not negligible;
- the preferred atomic sites for the different reaction mechanisms were established;
- the attack of the OOH free radical, for both the isomers, on the most abundant neutral species in a water solvent mainly occurs through a radical addition mechanism;
- for *iso*-moracin C, the radical addition process is favored in the lipid-like environment, while, for moracin C, both redox- and non-redox-type reactions can occur equally.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

AP, BD, TM, MT, and NR made equal contributions to the study and the publication of this work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation (Grant no. MAE00643232020-06-16).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.666647/full#supplementary-material

REFERENCES

Ahmadi, S., Marino, T., Prejanò, M., Russo, N., and Toscano, M. (2018). Antioxidant properties of the Vam3 derivative of resveratrol. *Molecules* 23:2446. doi: 10.3390/molecules23102446

Alberto, M. E., Grand, A., Russo, N., and Galano, A. (2013). A physicochemical examination of the free radical scavenging activity of Trolox: mechanism, kinetics and influence of the environment. *Phys. Chem. Chem. Phys*. 15:4642. doi: 10.1039/c3cp43319f

Castaneda-Arriaga, R., Marino, T., Russo, N., Alvarez-Idaboy, J. R., and Galano, A. (2020). Chalcogen effects on the primary antioxidant activity of chrysin and quercetin. *New J. Chem*. 44, 9073–9082. doi: 10.1039/D0NJ01795G

Ceron-Carrasco, J., Bastida, A., Requena, A., and Zuniga, J. (2010). A theoretical study of the reaction of bete-carotene with the nitrogen dioxide radical in solution. *J. Phys. Chem. B* 114, 4366–4372. doi: 10.1021/jp911846h

Ceron-Carrasco, J., Zuniga, J., Bastida, A., and Requena, A. (2012). Antioxidant properties of β-carotene isomers and their role in photosystems: insights from Ab initio simulations. *J. Phys. Chem. A* 116, 3498–3506. doi: 10.1021/jp301485k

Collins, F. C., and Kimball, G. E. (1949). Diffusion in chemical reaction processes and in the growth of colloid particles. *J. Colloid Sci*. 4:425. doi: 10.1016/0095-8522(49)90023-9

Delogu, G. L., Matos, M. J., Fanti, M., Era, B., Medda, R., Pieroni, E., et al. (2016). 2-Phenylbenzofuran derivatives as butyrylcholinesterase inhibitors: synthesis, biological activity and molecular modeling. *Bioorg. Med. Chem. Lett*. 26, 2308–2313. doi: 10.1016/j.bmcl.2016.03.039

Fashing, P. J. (2001). Feeding ecology of Guerezas in the Kakamega forest Kenya: the importance of Moraceae fruit in their diet. *Int. J. Primatol*. 22, 579–609. doi: 10.1023/A:1010737601922

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2009). *Guassian 09*. Pittsburgh, PA: Gaussian Inc.

Galano, A., and Alvarez-Idaboy, J. R. (2013). A computational methodology for accurate predictions of rate constants in solution: application to the assessment of primary antioxidant activity. *J. Comput. Chem*. 34, 2430–2445. doi: 10.1002/jcc.23409

Galano, A., and Alvarez-Idaboy, J. R. (2019). Computational strategies for predicting free radical scavengers' protection against oxidative stress: where are we and what might follow? *Int. J. Quantum. Chem*. 119:25665. doi: 10.1002/qua.25665

Galano, A., Mazzone, G., Alvarez-Diduk, R., Marino, T., Alvarez-Idaboy, J. R., and Russo, N. (2016). Food antioxidants: chemical insights at the molecular level. *Annu. Rev. Food Sci. Technol*. 7, 335–352. doi: 10.1146/annurev-food-041715-033206

Gázquez, J. L., Cedillo, A., and Vela, A. (2007). Electrodonating and electroaccepting powers. *J. Phys. Chem. A* 111, 1966–1970. doi: 10.1021/jp065459f

Hakim, E. H., Fahriyati, A., Kau, M. S., Achmad, S. A., Makmur, L., Ghisalberti, E. L., et al. (1999). Artoindonesianins A and B, two new prenylated flavones from the root of *Artocarpus champeden*. *J. Nat. Prod*. 62, 613–615. doi: 10.1021/np980279l

Jeon, S. Y., Kwon, S. H., Seong, Y. H., Bae, K., Hur, J. M., Lee, Y. Y., et al. (2007). β-secretase (BACE1)-inhibiting stilbenoids from Smilax Rhizoma. *Phytomedicine* 14, 403–408. doi: 10.1016/j.phymed.2006.09.003

Kapche, D., Lekane, N. M., Kulabas, S. S., Ipek, H., Tok, T. T., Ngadjui, B. T., et al. (2017). Aryl benzofuran derivatives from the stem bark of *Calpocalyx dinklagei* attenuate inflammation. *Phytochemistry* 141, 70–79. doi: 10.1016/j.phytochem.2017.05.007

Kapche, G. D. W. F., Fozing, C. D., Donfack, J. H., Fotso, G. W., Amadou, D., Tchana, A. N., et al. (2009). Prenylatedarylbenzofuran derivatives from *Morus mesozygia* with antioxidant activity. *Phytochemistry* 70, 216–221. doi: 10.1016/j.phytochem.2008.12.014

Kuete, V., Fozing, D. C., Kapche, W. F. G. D., Mbavengd, A. T., Kuiatea, J. R., Ngadjuib, B. T., et al. (2009). Antimicrobial activity of the methanolic extract and compounds from *Morus mesozygia* stem bark. *J. Ethnopharmacol*. 124, 551–555. doi: 10.1016/j.jep.2009.05.004

Lee, J. H., Ko, H. J., Woo, E. R., Lee, S. K., Moon, B. S., Lee, C. W., et al. (2016). Moracin M inhibits airway inflammation by interrupting the JNK/ c-Jun and NF-κB pathways *in vitro* and *in vivo*. *Eur. J. Pharmacol*. 783, 64–72. doi: 10.1016/j.ejphar.2016.04.055

Leopoldini, M., Russo, N., and Toscano, M. (2011). The molecular basis of working mechanism of natural polyphenolic antioxidants. *Food Chem*. 125, 288–306. doi: 10.1016/j.foodchem.2010.08.012

Li, X., Xie, H., Zhan, R., and Chen, D. (2018). Effect of double bond position on 2-phenyl-benzofuran antioxidants: a comparative study ofmoracin C and iso-moracin C. *Molecules* 23:754. doi: 10.3390/molecules23040754

Marcus, R. (1957). On the theory of oxidation-reduction reactions involving electron transfer. III. Applications to data on the rates of organic redox reactions. *J. Chem. Phys*. 26:872. doi: 10.1063/1.1743424

Marenich, A. V., Cramer, C. J., and Truhlar, D. G. (2009). Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* 113, 6378–6396. doi: 10.1021/jp810292n

Markovic, Z., Tošovic, J., Milenkovic, D., and Markovic, S. (2016). Revisiting the solvation enthalpies and free energies of the proton and electron in various solvents. *Comput. Theor. Chem*. 1077, 11–17. doi: 10.1016/j.comptc.2015.09.007

Mazzone, G., Malaj, N., Galano, A., Russo, N., and Toscano, M. (2015). Antioxidant properties of several coumarin– chalcone hybrids from theoretical insights. *RSC Adv*. 5, 565–575. doi: 10.1039/C4RA11733F

Na, M., Hoang, D. M., Njamen, D., Mbafor, J. T., Fomum, Z. T., Thuong, P. T., et al. (2007). Inhibitory effect of 2-arylbenzofurans from *Erythrina addisoniae*

on protein tyrosine phosphatase-1B. *Bioorg. Med. Chem. Lett.* 17, 3868–3871. doi: 10.1016/j.bmcl.2007.05.005

Naik, R., Harmalkar, D. S., Xu, X., Jang, K., and Lee, K. (2015). Bioactive benzofuran derivatives: moracins A-Z in medicinal chemistry. *Eur. J. Med. Chem.* 90, 379–393. doi: 10.1016/j.ejmech.2014.11.047

Nguyen, T. D., Jin, X., Lee, K., Hog, Y. S., Young, H. K., and Jung, J. L. (2009). Hypoxia-inducible factor-1 inhibitory benzofurans and chalcone-derived diels-alder adducts from Morus species. *J. Nat. Prod.* 72, 39–43. doi: 10.1021/np800491u

Pel, P., Chae, H. S., Nhoek, P., Kim, Y. M., and Chin, Y. W. (2017). Chemical constituents with proprotein convertase subtilisin/kexin type 9 mRNA expression inhibitory activity from dried immature *Morus alba* fruits. *J. Agric. Food Chem.* 65, 5316–5321. doi: 10.1021/acs.jafc.7b02088

Pérez-González, A., Castañeda- Arriaga, R., Verastegui, B., Carreón-González, M., Alvarez-Idaboy, J. R., and Galano, A. (2018). Estimation of empirically fitted parameters for calculating pK_a values of thiols in a fast and reliable way. *Theor. Chem. Acc.* 137:5. doi: 10.1007/s00214-017-2179-7

Pérez-González, A., García-Hernández, E., and Chigo-Anota, E. (2020). The antioxidant capacity of an imidazole alkaloids family through single-electron transfer reactions. *J. Mol. Model* 26:321. doi: 10.1007/s00894-020-04583-2

Romeo, I., Parise, A., Galano, A., Russo, N., Alvarez-Idaboy, J. R., and Marino, T. (2020). The antioxidant capability of higenamine: insights from theory. *Antioxidants* 9:358. doi: 10.3390/antiox9050358

Sang-Hee, L., Sang-Yoon, C., Hocheol, K., Jae-Sung, H., Byeong-Gon, L., Jian-Jun, G., et al. (2002). Mulberroside F isolated from the leaves of *Morus alba* inhibits melanin biosynthesis. *Biol. Pharm. Bull.* 25, 1045–1048. doi: 10.1248/bpb.25.1045

Seong, S. H., Ha, M. T., Min, B. S., Jung, H. A., and Choi, J. S. (2018). Moracin derivatives from Morus radix as dual BACE1 and cholinesterase inhibitors with antioxidant and anti-glycation capacities. *Life Sci.* 210, 20–28. doi: 10.1016/j.lfs.2018.08.060

Truhlar, D. G., Garrett, B. C., and Klippenstein, S. J. (1996). Current status of transition-state theory. *J. Phys. Chem.* 100:12771. doi: 10.1021/jp953748q

Venkatesh, K. R., and Seema, C. (2008). Mulberry: life enhancer. *J. Med. Plants Res.* 2, 271–278. doi: 10.5897/JMPR.9000005

Zelová, H., Hanáková, Z., Cermáková, Z., Šmejkal, K., Dall, A. S., Babula, P., et al. (2014). Evaluation of anti-inflammatory activity of prenylated substances isolated from *Morus alba* and morusnigra. *J. Nat. Prod.* 77, 1297–1303. doi: 10.1021/np401025f

Zhao, Y., Schultz, N. E., and Truhlar, D. G. (2006). Assessment of model chemistries for noncovalent interactions. *J. Chem. Theory Comput.* 2, 364–382. doi: 10.1021/ct0502763

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.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 06 July 2021
doi: 10.3389/fchem.2021.710025

UHPLC-(ESI)-HRMS and NMR-Based Metabolomics Approach to Access the Seasonality of *Byrsonima intermedia* and *Serjania marginata* From Brazilian Cerrado Flora Diversity

Ana C. Zanatta[1,2], *Wagner Vilegas*[2] *and RuAngelie Edrada-Ebel*[3]*

[1]Laboratory of Phytochemistry, Institute of Chemistry, Department of Biochemistry and Organic Chemistry, São Paulo State University (UNESP), Araraquara, Brazil, [2]Laboratory of Bioprospecting of Natural Products, Institute of Biosciences, São Paulo State University (UNESP), São Vicente, Brazil, [3]Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, United Kingdom

Seasonality is one of the major environmental factors that exert influence over the synthesis and accumulation of secondary metabolites in medicinal plants. The application of the metabolomics approach for quality control of plant extracts is essentially important because it helps one to establish a standard metabolite profile and to analyze factors that affect the effectiveness of the medicinal plants. The Brazilian Cerrado flora is characterized by a rich diversity of native plant species, and a number of these plant species have been found to have suitable medicinal properties. Some of these plant species include *Byrsonima intermedia* and *Serjania marginata*. To better understand the chemical composition of these plant species, we conducted a study using the state-of-the-art techniques including the HPLC system coupled to an Exactive-Orbitrap high resolution mass spectrometer with electrospray ionization interface UHPLC-(ESI)-HRMS and by NMR being performed 2D *J*-resolved and proton NMR spectroscopy. For the analysis, samples were harvested bimonthly during two consecutive years. UHPLC-(ESI)-HRMS data were preprocessed and the output data uploaded into an in-house Excel macro for peak dereplication. MS and NMR data were concatenated using the data fusion method and submitted to multivariate statistical analysis. The dereplication of LC-HRMS data helped in the annotation of the major compounds present in the extracts of the three plant species investigated allowing the annotation of 68 compounds in the extracts of *B. intermedia* (cinnamic acids, phenolic acids derived from galloyl quinic and shikimic acid, proanthocyanidins, glycosylated flavonoids, triterpenes and other phenols) and 81 compounds in the extracts of *S. marginata* (phenolic acids, saponins, proanthocyanidins, glycosylated flavonoids among other compounds). For a better assessment of the great number of responses, the significance of the chemical variables for the differentiation and correlation of the seasons was determined using the variable importance on projection (VIP) technique and through the application of the false discovery rate (FDR) estimation. The statistical data obtained showed that seasonal factors played an important role on the production of metabolites in each plant species.

OPEN ACCESS

Edited by:
Anna Napoli,
University of Calabria, Italy

Reviewed by:
Angela Amoresano,
University of Naples Federico II, Italy
Ravi Pratap Barnwal,
Panjab University, India

***Correspondence:**
RuAngelie Edrada-Ebel
ruangelie.edrada-ebel@strath.ac.uk

Specialty section:
This article was submitted to Analytical Chemistry, a section of the journal Frontiers in Chemistry

Received: *20 May 2021*
Accepted: *24 June 2021*
Published: *06 July 2021*

Citation:
Zanatta AC, Vilegas W and Edrada-Ebel R (2021) UHPLC-(ESI)-HRMS and NMR-Based Metabolomics Approach to Access the Seasonality of Byrsonima intermedia *and* Serjania marginata *From Brazilian Cerrado Flora Diversity. Front. Chem. 9:710025. doi: 10.3389/fchem.2021.710025*

Temperature conditions, drought and solar radiation were found to be the main factors that affected the variability of phenolic compounds in each species.

Keywords: metabolomic approach, specialized metabolites, seasonality, environmental factors, phenolic compounds, saponins, triterpenes

INTRODUCTION

Natural products derived from the secondary metabolism of plants have played an essentially important role in the treatment of diseases and illnesses throughout the life history of humans. The search for relief and cure of diseases through the ingestion of herbs and leaves may have been one of the earliest ways in which natural products were used (Cragg and Newman, 2013; Katz and Baltz, 2016). The diversity of molecular structures and nature's ability to provide molecules of structural complexity hardly imagined or elaborated by synthesis is the result of many different biosynthetic pathways that are involved in the production of a plant's secondary metabolites, and what also leads these metabolites to exhibit a wide range of biological activities (Harvey et al., 2015; Bernardini et al., 2018).

The production of specialized metabolites by plants is characterized by the specificity of the plant species and is regarded an essentially relevant adaptive response in terms of coping with environmental and/or external stimuli to which the plants are exposed (Ghorbanpour and Varma, 2017). In addition, these metabolites play a key role in regulating plants growth, development, and defense, as well as in mediating plant-environment interactions, including protection against several stressful and challenging environmental change conditions (Rai et al., 2017; Isah, 2019; Pandey and Senthil-Kumar, 2019; Erb and Kliebenstein, 2020).

In recent decades, there has been an increasingly growing scientific interest in gaining a comprehensive understanding regarding the changes in secondary metabolism in plants, and a considerable number of studies have been conducted aiming at investigating the effects of changes in environmental factors (e.g., temperature, climate change, precipitation, drought, salinity, UV radiation, light and humidity) and seasonality on the biosynthesis and accumulation of specialized metabolites in medicinal plants (Xiao et al., 2008; Sampaio et al., 2016; Berini et al., 2018; Yang et al., 2018). For purposes of illustration, Sehlakgwe et al. (2020) demonstrated the effect of seasonal variation on the chemical profile of plants and the differences in activity observed for the plant material harvested in different seasons; the authors showed that winter was the best harvest period for *Leucosidea sericea* leaves. In another study, Botha et al. (2018) investigated the variations in the accumulation of three secondary metabolites present in different parts of the *Euclea undulata* Thunb. var. *myrtina* plant and the correlation with the amount of rainfall during the dry and rainy seasons, as well as changes in temperature.

As aforementioned, due to several factors, the content of specialized metabolites can vary considerably in plants species; and since many of these metabolites are active principles, this variation can eventually change the therapeutic response of the medicinal plant when applied for particular pharmacological purposes. In view of that, studying the effects of factors that may determine or modify the yield of these compounds in plants, such as growing and/or harvesting conditions, season and time of day, is extremely important if one aims to obtain good quality herbal products with desirable concentrations of active ingredients (Kunle et al., 2012; Liebelt et al., 2019).

Conducting a rapid, high-quality analysis of the chemical constituents of a complex plant matrix is indispensable in the sense that it allows one to characterize the rationally active markers in the matrix and facilitates one's understanding of the relationship between the chemical composition and the possible efficacy, toxicity and therapeutic target of plant-based medicine (Wolfender et al., 2019).

Remarkably, a major obstacle to establishing a reliable chemical quality control framework of a plant extract is that plants have an extremely complex composition with a huge number of compounds, of which there is still very limited knowledge. The chemical diversity of compounds present in medicinal plants is directly associated with the high variability of the intrinsic physicochemical properties of natural products, and this makes the separation, detection and identification of the natural substrates of the plants technically challenging (Dunn et al., 2011; Wolfender et al., 2015).

Significant advances in analytical techniques along with new bioinformatics tools and multivariate statistical analysis have contributed meaningfully toward the advancement of chemical studies of complex samples. Metabolomics-based approaches which employ state-of-the-art techniques, such as high resolution mass spectrometry (HRMS) and one and two-dimensional NMR spectroscopy, is a fast and efficient way of maximizing the results of metabolic fingerprinting analysis of a large number of data sets (Kellogg et al., 2017; Han et al., 2018; Sut et al., 2019; Houriet et al., 2020). Furthermore, unlike their individual application, the combined application of MS and NMR (MS-NMR) provides more reliable identification results and unbiased assessment of quality control analyses of medicinal plants in addition to helping predict the bioactivity of the plants (Sampaio et al., 2016; Sut et al., 2019; Sehlakgwe et al., 2020).

The Brazilian Cerrado flora is widely known to be constituted by a rich diversity of plant species, and a number of these plant species have been found to possess suitable medicinal properties (Lahsen et al., 2016; Cortelo et al., 2021). Some of these plant species include *Byrsonima intermedia* and *Serjania marginata*, which present similar ethnopharmacological characteristics and have been found to be promising for future application in phytotherapy aimed at the treatment of inflammatory diseases. The hydroethanolic extracts obtained from the leaves of each of these species demonstrated potential gastroprotective (Arruda et al., 2009; Périco et al., 2015; dos Santos et al., 2019), anti-inflammatory

(Moreira et al., 2011; Salinas-Sánchez et al., 2017), antioxidant (Heredia-Vieira et al., 2015; Guilhon-Simplicio et al., 2017), analgesic (Di Stasi et al., 1988; Verdam et al., 2017) activities and absence of toxicity in the acute models evaluated (Santos et al., 2012; Périco et al., 2015; dos Santos et al., 2019).

The *B. intermedia* A. Juss species is popularly known as "murici-do-cerrado", "cajuzinho do cerrado" or "murici-mirim" - this name, which means a small tree in Tupi Guarani language, was given to the species by the indigenous people. The *B. intermedia* A. Juss species is native to the Brazilian Cerrado (the Brazilian tropical savanna region), and as the name implies, the plant species has a bushy-like characteristics and has been found to reach a maximum height of 1.60 m. The local population of this savanna uses the bark and leaves of *B. intermedia* as infusions; these infusions have been found to possess antiseptic, antimicrobial, anti-haemorrhagic, cicatrizant, and anti-inflammatory properties (Nogueira et al., 2004; Oliveira et al., 2007).

The *S. marginata* Casar. species is a climbing plant popularly known as "cipó-uva" or "cipó-timbó"; the plant species is used in folk medicine in the form of infusion or juice for the treatment of stomach pains (Arruda et al., 2009). The species is considered native to Brazil, Paraguay, Bolivia and Argentina. In Brazil, the *S. marginata* Casar. species has been found to be present in deciduous forests, floodable fields, and in the Chaquenha region of Mato Grosso do Sul, as well as in 18 forest fragments in the northwest of the State of São Paulo, and in the dense forest of the State of Pernambuco (Rodal and Nascimento, 2002; Moreira et al., 2013; Sprengel-Lima and Rezende, 2013).

Previous phytochemical studies conducted on the leaf extracts of the plant species mentioned above have shown that *B. intermedia* has phenolic acids, oligomeric proanthocyanidins, and flavonoids as its main constituents (Pereira et al., 2015; Fraige et al., 2018; Mannochio-Russo et al., 2020) and *S. marginata* has flavonoids, oligomeric proanthocyanidins and saponins as its main constituents (Heredia-Vieira et al., 2015).

In view of the chemical and biological potential of *B. intermedia* and *S. marginata* species, conducting qualitative chemical analyses of the leaves of these plants, based on a seasonal approach, is essentially relevant in the sense that it enables one to verify if there is chemical variation in the profile of the species harvested at different times of the year due to the environmental factors evaluated related to the season (temperature, humidity, solar radiation and rainfall), apart from helping to identify the biological potential of the plant matrix investigated. In the present work, state-of-the-art tools in metabolic profiling and data analysis methods based on liquid chromatography LC-HRMS and NMR techniques were employed for the conduct of qualitative chemical analyses of the aforementioned plant species.

MATERIALS AND METHODS

Chemicals

Methanol and formic acid (LC-MS grade) were obtained from Merck (Darmstadt, Germany) and ultrapurified water (Millipore®, United States) were used as mobile-phase components.

Plant Material

B. intermedia leaves were harvested in the Municipal Botanical Garden of Bauru (JBMB), Bauru, state of São Paulo, Brazil (22°20′36″S and 49°01′02″W, altitude 546 m). The voucher specimens were deposited at Herbarium JBMB under number JBMB 00013 and responsibility of Dr. Viviane Camila de Oliveira.

S. marginata leaves were harvested in a fragment of Cerrado located in Santa Madalena Farm, Dourados, Mato Grosso do Sul state, Brazil (22°08′05″S and 55°08′17″W, altitude of 452 m). The voucher specimens were deposited at the DDMS Herbarium, of the Federal Universiy of Grande Dourados (UFGD), under number 41054 and responsibility of Dr Emerson Silva.

The harvesting of the plant was always carried out at the same period between 9 and 10 am for two consecutive years (2017 and 2018) to study the seasonal variability of their secondary metabolites. The detailed data for each harvest are described in **Supplementary Table S1**.

The harvested data were added in the SisGen platform (National System of Management of Genetic Heritage and Associated Traditional Knowledge) as genetic patrimony with registration number A3476AF. For the transfer of material between São Paulo State University—UNESP and the University of Strathclyde, all samples were prepared in accordance with the Brazilian laws for access and shipment of genetic heritage material. The R0418CB shipment number was issued by SisGen and under authorization from the Genetic Heritage Management Council (CCGEN).

Hydroethanolic Extracts Preparation

The plant material (*B. intermedia* and *S. marginata* leaves) was washed and dried in an oven with air circulation at 40°C. The dry material was ground into an analytical mill (model IKA A11 basic). For the analytical analysis, 100 mg of powder was extracted with 1 ml of EtOH/H_2O 7:3 (v/v) in an ultrasound bath, three times for 20 min. After that, the resulting material was centrifuged at 13,000 rpm and the supernatant was filtered through a Millex® PTFE filter (0.22 µm, 25 mm). The extracted liquid was concentrated at a temperature below 40°C until removal of the organic solvent. The hydroethanolic extracts were frozen and lyophilized in a vacuum freeze dryer (BentchTop Pro SP Scientific) at 200 mT for 48 h at −60°C.

Climate Data

The climatic data for all the harvests of each plant were provided by the Meteorological Data Storage Section (SADMET) of the National Institute of Meteorology (INMET). The climatic data were temperature (°C), humidity (%), solar radiation (kJ/m^2) and rainfall (mm) for the years 2017 and 2018 provided in **Supplementary Tables S2 and 3**.

UHPLC-(ESI)-HRMS Analysis

The chromatographic chemical profile analysis of the extracts of each species was performed by UHPLC-(ESI)-HRMS Thermo Scientific® Accela using Thermo Scientific® UHPLC Accela

system coupled to an Exactive-Orbitrap high-resolution mass spectrometer Thermo Scientific®. To perform the analysis, the samples were suspended in methanol at a concentration of 1.0 mg/ml and filtered with Millex® PTFE filter (0.22 µm, 25 mm). LC separations were conducted through a C-18 column (ACE, 75 mm, id 3.0 mm, 5 µm). The injection volume and flow rate applied were 10 and 300 µL/min, respectively. For sample elution, a linear gradient with mobile phase composed of water (solvent A) and methanol (solvent B) both acidified with 0.1% formic acid, from 5 to 100% (B) was employed for 45 min. The acquisition range was *m/z* 150–2000 in both negative and positive ionization modes. The spray voltage applied was 4.5 kV for positive mode and 4.0 kV for negative mode and capillary temperature set at 280°C. The mass spectra were obtained and processed in Xcalibur software (version 3.0).

NMR Analysis

To conduct NMR analysis, samples were prepared by dissolving each sample in 650 µL of DMSO-d_6 (Sigma Aldrich®) to obtain the concentration of 5 mg/ml then transferred to 5 mm 7″ NMR tubes. *J*-resolved and proton (^{1}H NMR) experiments were performed using a Bruker® AVIII HD 500 (11.7 T) and the spectra were processed using MestReNova x64 software (version 14.1.2).

The data acquisition followed the parameters used in the previous work developed by Sampaio et al. (2016) using 16 scans in the pre-saturation pulse sequence for one-dimensional (1D) proton spectra (^{1}H NMR) and 32 scans and 64 increments per scan for two-dimensional (2D) ^{1}H–^{1}H *J*-resolved (*J*-res) NMR spectra; data points widths of 3.56 kHz for F2 (chemical shift axis) and 50 Hz for F1 (spin-spin coupling constant axis). The solvent signal was suppressed by the selective pre-saturation method.

Data Processing

The UHPLC-(ESI)-HRMS data were converted to mzML format in the MSConvert (ProteoWizard) software using the filter Peak Picking with the algorithm vendor checked. mzML files were processed in MZmine2 v.2.53 (http://mzmine.sourceforge.net/) (Katajamaa et al., 2006; Pluskal et al., 2010) using the following parameters: mass detector, Centroid and MS1 noise level, 1×10^3. For the construction of chromatogram, the parameters setting conditions were used: min time spam of 0.2 min; min height intensity of 1×10^4 and mass tolerance of 0.001 m/*z* or 5.0 ppm; for the deconvolution, the algorithm local minimum search was chosen using: chromatographic threshold of 5%; search miminum in RT range of 0.4 min; minimum absolute height of 1×10^4; min ratio of peak top/edge and peak duration range of 0.2–5 min. For data processing was applied deisotoping, filtering, alignment and gap filling steps. The identification of adducts and complexes were carried out. The molecular formulas were calculated by the Formula Prediction for both ionization modes ($[M+H]^+$ and $[M-H]^-$) within the mass tolerance window 0.001 m/*z* or 5.0 ppm and applying the element counts (C, H, N, O, P, S) from the detected *m/z* value considering all the heuristic rules (elemental ratios, RDBE restrictions and isotope pattern filter) (Pluskal et al., 2012). The output data (peak areas, exact mass, molecular formula and retention times) of each sample were exported as a CSV file and uploaded into an in-house Excel macro with the Dictionary of Natural Products (DNP) database for peak dereplication (Macintyre et al., 2014). A blank solvent was analyzed together with the extracts by LC-HRMS during data processing. Therefore, by macro Excel, the peaks from the blank were extracted and removed, applying an algorithm based on the intensity ratio of *m/z* between the samples and the blank.

The NMR data were processed using MestReNova x64 software version 14.1.2 (Mestrelab Research S.L.®). First, for processing the 2D *J*-res spectra the T1 noise reduction, 45° tilt and symmetrization by *J*-res sensitivity enhancement parameters were used; and at last, spectra in the one-dimensional projection were extracted.

The ^{1}H NMR and *J*-res projection spectra were previously stacked and then pre-processing according the following steps: baseline correction with the Whittaker Smoother option; apodization with Gaussian function of 1 GB [Hz]; normalization by the largest peak and smoothing Savitzky–Golay method. The processed NMR data were prepared using bin width of 0.04 ppm and the bin intensities method was average sum. Afterward, the spectral data were saved with the peak intensities. Chemical shifts (δ) values were established between 0.0 and 10.0 ppm for *S. marginata* samples data and 0.0–13.0 ppm for *B. intermedia*.

The data fusion method was performed for the concatenation of the processed MS and NMR data (Forshed et al., 2006; Sampaio et al., 2016). The data were pre-processed and organized in Excel® being MS and NMR data divided into two blocks. Each block were scaled according to the following procedure:

1) The standard deviation (SD) calculation of all the corresponding peak areas of each *m/z*;
2) The sum of all SD and then
3) The division of the areas of the original peaks by the sum of SD. MS and NMR scaled data were merged in a single data matrix which undergone to multivariate analysis.

Multivariate Analysis

The processed MS, NMR and MS-NMR fused data were submitted to multivariate statistical analyses to study the metabolites variation related to the climate data of each plant species *B. intermedia* and *S. marginata*. The data obtained from the analysis of the extracts were classified as chemical variables (primary variables ID) and the climatic data as environmental variables (secondary variables ID). The set of variables were scaled using the Pareto algorithm in the SIMCA-P software v. 15.0 (Umetrics®, Sweden). Partial Least Squares Discriminant Analysis regression (PLS-DA) and Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) were performed with at least three independent replicates for each season. Afterward, the model was validated by the permutation tests (repeated 100 times). The model fitting was given by parameter R2X and the expected variation by Q2X.

The correlation and differences in the metabolite profile for the seasons of each plant were screened by the PLS-DA loadings plot. To better interpret a large number of responses, the chemical

variables were classified according to their significance to the model by the parameter variable importance on projection (VIP). The 15 variables with the highest VIPs were selected for further analysis. Thereafter, the false discovery rate (FDR) method was applied as follows: ranking the *p*-values of the variables from low to high, multiplying each *p*-value by the number of variables ($N = 15$), and dividing by their order of rank (Benjamini and Hochberg, 1995; Wang et al., 2013). Finally, the resulting FDR were adjusted considering values ≤0.05 to be significant.

The web-based platform MetaboAnalyst 4.0 (https://www.metaboanalyst.ca/) (Chong et al., 2018) was also used in the PLS-DA and hierarchical analysis. For the purposes of processing, the MS-NMR fusion data was filtered by the mean intensity value (Hackstadt and Hess, 2009), the samples were normalized by sum and the set of variables were scaled using the Pareto algorithm. The normalized output data were exported as a CSV file and analyzed using GraphPad Prism 8.4.3 (GraphPad Software, San Diego, CA, United States). The statistical significance between results obtained for the seasonal groups were analyzed using one-way ANOVA followed by Tukey's test. All results are presented as mean ± standard deviation (SD).

RESULTS

Metabolite Profiling Annotation by NMR and MS

A meticulous analysis based on the negative and positive ESI-HRMS spectral data, ^{1}H NMR spectra, as well as 2D *J*-res spectra, allowed the annotation of the compounds present in the plant species *B. intermedia* and *S. marginata* extracts. MS and NMR features assignments were founded according to data previously reported in the literature for representative classes of metabolites or chemical markers for each genus and/or family studied. The comprehensive annotation of the compounds for each plant species are provided in the **Supplementary Tables S4 and 5** (based on an accurate mass and molecular formulas). **Supplementary Figures S1 to 8** shows the representative LC-HRMS chromatograms, the ^{1}H NMR spectra, 1D *J*-res NMR projection spectra and the 2D *J*-res NMR for each plant extract.

Due to the complexity of the plant matrices, the ^{1}H NMR spectra of the analyzed extracts show critical overlapping signals mainly as a result of the inherent signal splitting and occurrence of multiplet patterns. Therefore, 2D *J*-res NMR experiments proved to be a suitable choice to solve the problem of signal congestion. By projecting the *J*-res spectrum onto the chemical shift axis, the degree of spectral complexity was reduced, and the spectral resolution was increased (**Supplementary Figures S5 to 8**).

Considering the NMR identification, three main chemical shift regions were clearly distinguished in the spectra of *B. intermedia* and *S. marginata* samples, with peaks in the aliphatic (0.50–2.00 ppm), sugar and organic acids (3.00–6.00 ppm) and aromatic region (6.00–9.00 ppm) (**Supplementary Figures S5 to 8**). Regarding the carbohydrate portion of the compounds, the region of 4.00–6.00 ppm can be assigned to the signals of the anomeric protons of the sugars. In addition, the multiplicity and chemical shifts of these signals obtained in the 1D and 2D *J*-res spectra helped to propose the nature of the glycosides.

ESI-HRMS data annotation in positive and/or negative modes allowed to distinguish several classes of secondary metabolites in the three plant species.

B. intermedia extracts exhibited signals of cinnamic acids and derivatives (e.g., *m/z* 163.0396 $[M-H]^-$, *m/z* 199.0605 $[M+H]^+$, *m/z* 339.1078 $[M+H]^+$, *m/z* 355.1025 $[M+H]^+$, *m/z* 369.1183 $[M+H]^+$ and *m/z* 531.1503 $[M+H]^+$), galloylquinic acids (e.g., *m/z* 345.0819 $[M+H]^+$, *m/z* 497.0930 $[M+H]^+$, *m/z* 649.1040 $[M+H]^+$ and *m/z* 801.1151 $[M+H]^+$), galloylshikimic acids (e.g., *m/z* 327.0713 $[M+H]^+$, *m/z* 479.0825 $[M+H]^+$ and *m/z* 631.0937 $[M+H]^+$), proanthocyanidins [e.g., monomer (*m/z* 291.0864 $[M+H]^+$), dimer (*m/z* 577.1351 $[M+H]^+$, tetramer (*m/z* 1153.2609 $[M+H]^+$)], flavonoids derived from quercetin, such as monoglycosylated flavonoids (*m/z* 435.0928 $[M+H]^+$, *m/z* 449.1085 $[M+H]^+$ and *m/z* 465.1035 $[M+H]^+$), diglycosylated flavonoids (*m/z* 581.1511 $[M+H]^+$, *m/z* 597.1456 $[M+H]^+$ and *m/z* 611.1617 $[M+H]^+$), triglycosylated flavonoids (*m/z* 757.2195 $[M+H]^+$) and galloyl flavonoids (*m/z* 587.1041 $[M+H]^+$, *m/z* 601.1197 $[M+H]^+$ and *m/z* 617.1147 $[M+H]^+$), pentacyclic triterpenes with lupane and oleanane structures (e.g., *m/z* 427.3938 $[M+H]^+$, *m/z* 455.3546 $[M-H]^-$, *m/z* 487.3422 $[M+H]^+$ and *m/z* 489.3579 $[M+H]^+$), and other compounds (**Supplementary Table S4**).

S. marginata was characterized by having phenolic acids (e.g., *m/z* 153.0186 $[M-H]^-$ and *m/z* 315.0724 $[M-H]^-$), cinnamic acids (e.g., *m/z* 179.0345 $[M-H]^-$, *m/z* 199.0600 $[M-H]^-$ and *m/z* 343.0810 $[M+H]^+$), triterpenic saponins derived from oleanolic acid (e.g., *m/z* 733.4550 $[M-H]^-$, *m/z* 865.4971 $[M-H]^-$, *m/z* 895.5082 $[M-H]^-$, *m/z* 911.5031 $[M-H]^-$, *m/z* 941.5128 $[M-H]^-$, *m/z* 1011.5554 $[M-H]^-$, *m/z* 1027.5503 $[M-H]^-$, *m/z* 1043.5452 $[M-H]^-$, *m/z* 1057.5609 $[M-H]^-$, and so on), B-type proanthocyanidins [e.g., monomer (*m/z* 291.0860 $[M+H]^+$), dimer (*m/z* 579.1500 $[M+H]^+$), trimer (*m/z* 867.2117 $[M+H]^+$) and tetramer (*m/z* 1155.2719 $[M+H]^+$)] and A-type proanthocyanidins [e.g., dimer (*m/z* 577.1343 $[M+H]^+$), trimer *m/z* 863.1810 $[M+H]^+$), tetramer (*m/z* 1153.2602 $[M+H]^+$) and pentamer (*m/z* 1441.3228 $[M+H]^+$)], flavonoids glycosylated including *C*-glycosylated flavones (e.g., *m/z* 403.1023 $[M+H]^+$, *m/z* 417.1177 $[M+H]^+$ and *m/z* 419.0973 $[M+H]^+$), *m/z* 597.1451 $[M+H]^+$ and *m/z* 609.1479 $[M-H]^-$), *C,O*-glycosylated flavones (e.g., *m/z* 551.1394 $[M+H]^+$, *m/z* 561.1603 $[M+H]^+$, *m/z* 563.1761 $[M+H]^+$, *m/z* 565.1550 $[M+H]^+$, *m/z* 577.1551 $[M+H]^+$, *m/z* 579.1707 $[M+H]^+$ and *m/z* 591.1708 $[M+H]^+$), and *O*-glycosylated flavonols (e.g., *m/z* 433.1129 $[M+H]^+$, *m/z* 465.1027 $[M+H]^+$ and *m/z* 595.1658 $[M+H]^+$), among other compounds (**Supplementary Table S5**).

Seasonality Assessment From MS-NMR Fused Data

Multivariate analysis of HRMS and 2D *J*-res NMR data were evaluated to assess the seasonality of the *B. intermedia* and *S. marginata* harvests. MS and NMR data were evaluated separately and concatenated (**Figures 1**, **2**).

The PLS-DA plot in **Figure 1** shows that the *B. intermedia* harvests had a significant seasonal effect, with separation into

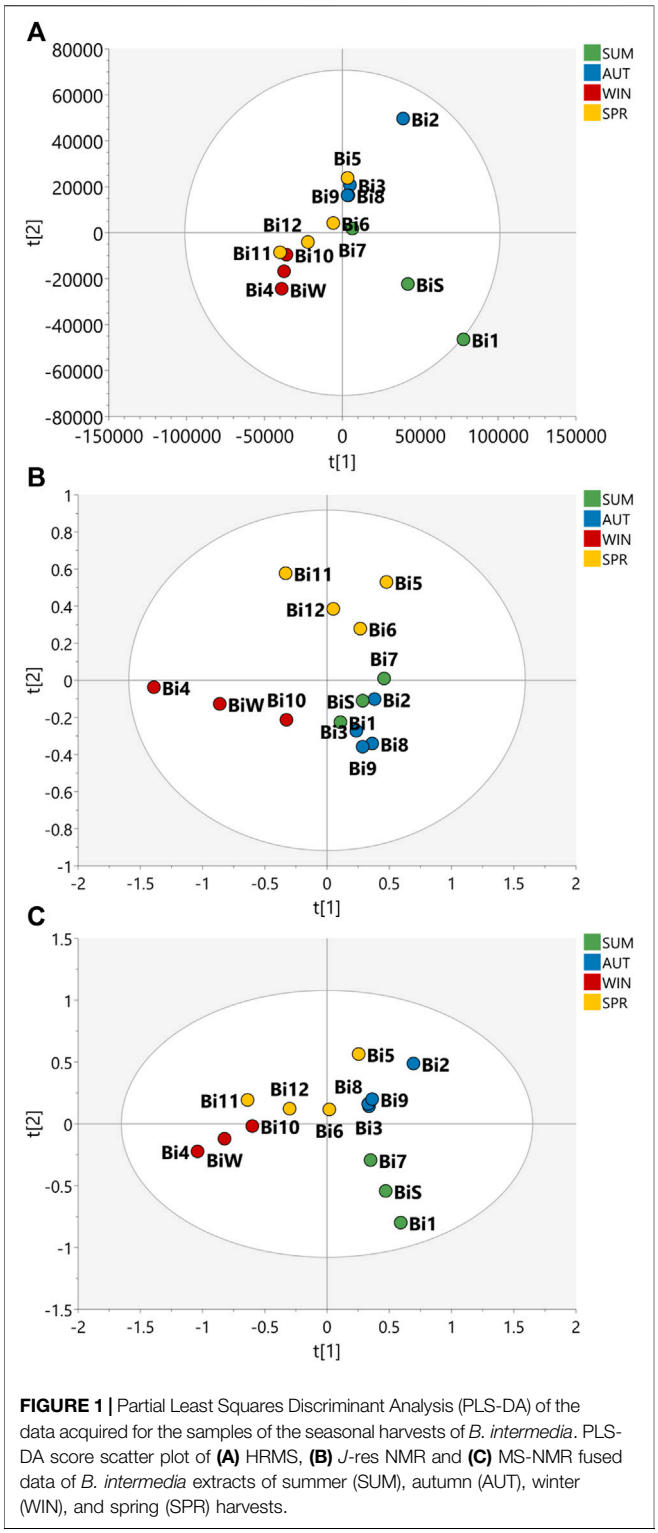


FIGURE 1 | Partial Least Squares Discriminant Analysis (PLS-DA) of the data acquired for the samples of the seasonal harvests of *B. intermedia*. PLS-DA score scatter plot of **(A)** HRMS, **(B)** *J*-res NMR and **(C)** MS-NMR fused data of *B. intermedia* extracts of summer (SUM), autumn (AUT), winter (WIN), and spring (SPR) harvests.

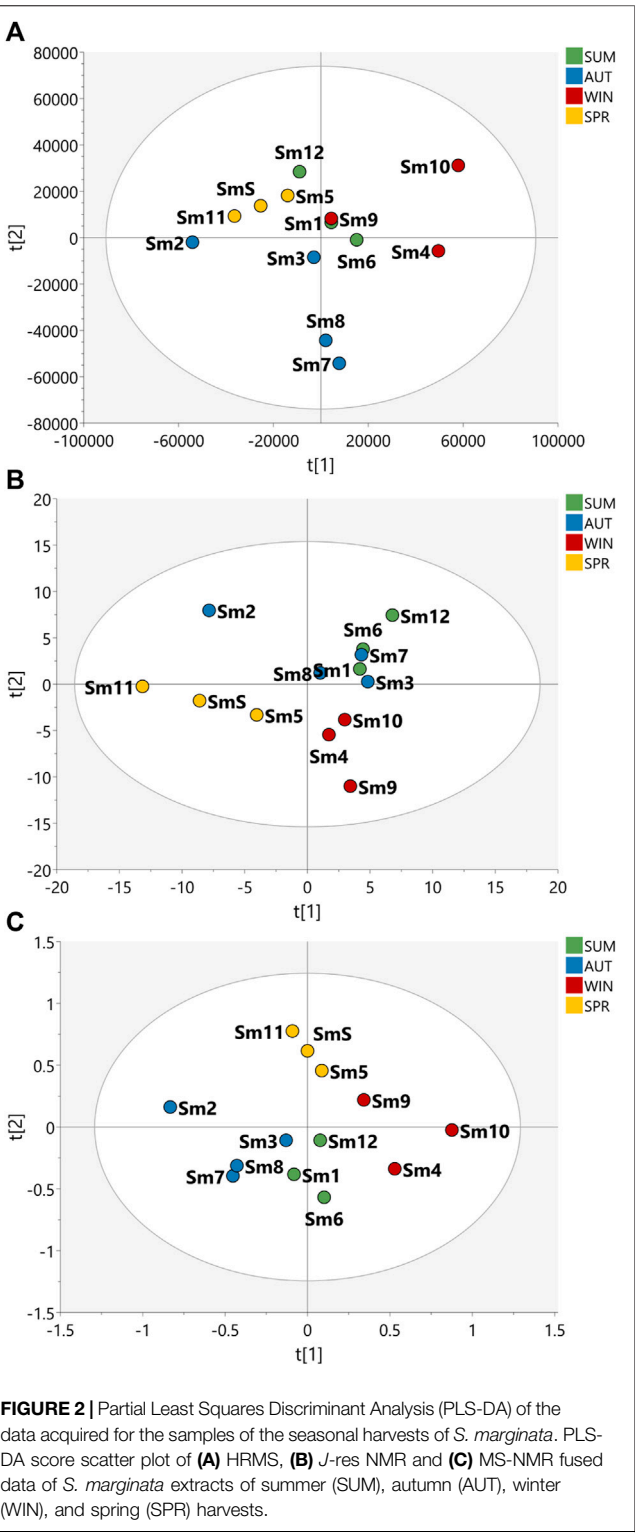


FIGURE 2 | Partial Least Squares Discriminant Analysis (PLS-DA) of the data acquired for the samples of the seasonal harvests of *S. marginata*. PLS-DA score scatter plot of **(A)** HRMS, **(B)** *J*-res NMR and **(C)** MS-NMR fused data of *S. marginata* extracts of summer (SUM), autumn (AUT), winter (WIN), and spring (SPR) harvests.

distinct groups, especially for summer and winter samples. In **Figure 1A**, clustering was performed according to HRMS data and it is apparent that the spring samples presented an intermediate MS profile between winter and autumn seasons; in **Figure 1B**, for the *J*-res data, summer and autumn had more similarity in their NMR profiles which contributed to the clustering of these samples and finally in **Figure 1C**, the concatenation of HRMS and *J*-res data resulted in an improved clustering of the harvests samples from the same season and more clearly represented the differentiation

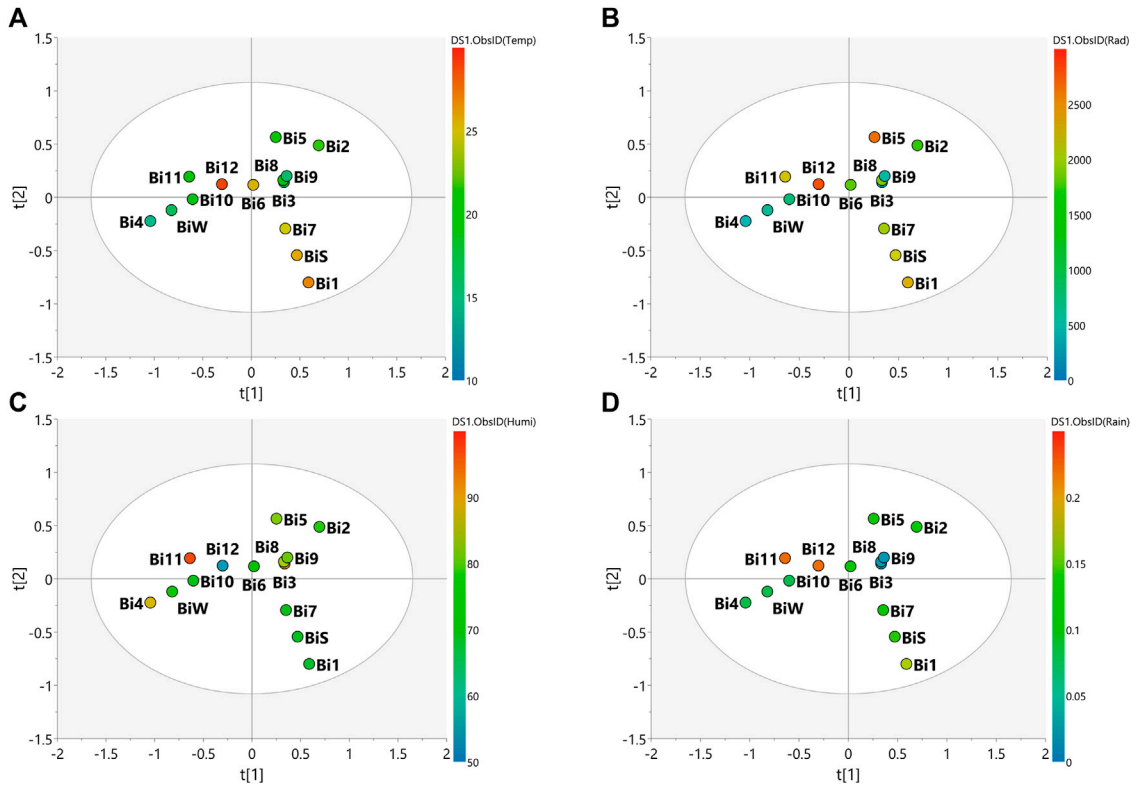


FIGURE 3 | PLS-DA score scatter plots of the MS-NMR fused data of *B. intermedia* seasonal harvests and the correlation with the environmental factors. Samples are categorized according to the values of observations for environmental factors. **(A)** Temp = temperature (°C); **(B)** Rad = Solar radiation (kJ/m^2); **(C)** Humi: relative humidity (%) and **(D)** Rain = rainfall (mm). Sample seasons: Summer (Bi1, Bi7 and BiS); Autumn (Bi2, Bi3, Bi8 and Bi9); Winter (Bi4, Bi10 and BiW) and Spring (Bi5, Bi6, Bi11 and Bi12).

in the sets of seasons according to the variation in their features.

With regard to *S. marginata* harvests, seasonality appeared not to have a significant influence on the differentiation of samples, demonstrating only a slight distinction of samples harvested during spring in comparison to the other seasonal groups. For the HRMS dataset in **Figure 2A**, the samples are quite scattered, with no clustering trends within seasons; in **Figure 2B**, the *J*-res data contributed to the separation of spring and winter into two distinct groups, while the autumn and summer seasons had closer clustering, and then finally, in **Figure 2C**, with the concatenated HRMS and *J*-res data, it can be observed that the samples are more scattered, that is, the sets of each season do not cluster tightly, moreover, summer showed to present a chemical profile with intermediate features between autumn and winter seasons.

The PLS-DA score scatter presented in **Figures 3**, **4** for the plant species samples, correlates the MS-NMR fused data of seasonal harvests plotted against the environmental factors. We choose to categorize the plant species according to the values of observations for main environmental factors observed in the region of the harvests (Temperature, Solar radiation, Relative humidity and Rainfall). Average values of weather characteristics of the local where each plant species was harvested are shown **Supplementary Tables S2 and 3**.

In view of the weather characteristics of the local where *B. intermedia* was harvested (**Supplementary Table S2**), it is possible to notice a lower rainfall amount during autumn and winter, and a higher amount during spring (**Figure 3D**). In addition, spring and summer periods had the warmest temperatures (**Figure 3A**) and the most intense incidence of solar radiation (**Figure 3B**). Visual inspection of the grouping of samples and the climate observations in the plot of **Figure 3** shows that the temperature is the climate variable which better explained the correlation between the samples, besides being the only variable that explained the correlation of the spring and winter samples and the influence on the chemical composition (**Figure 3A**). The samples harvested during winter (Bi4, Bi10 and BiW) proved to be the most homogeneous and are more tightly grouped, indicating minimal changes in the chemical profiles for the samples harvested in this season and this is due to the lower climatic variation during this period. Spring was the most heterogeneous season with the greatest dispersion of samples (Bi5, Bi6, Bi11, Bi12), indicating a greater variation in the chemical profiles of these samples. This season is intermediate between winter and summer and was influenced by the climatic conditions of late winter and early summer.

The main weather conditions for *S. marginata* harvests (**Supplementary Table S3**) indicate higher temperatures

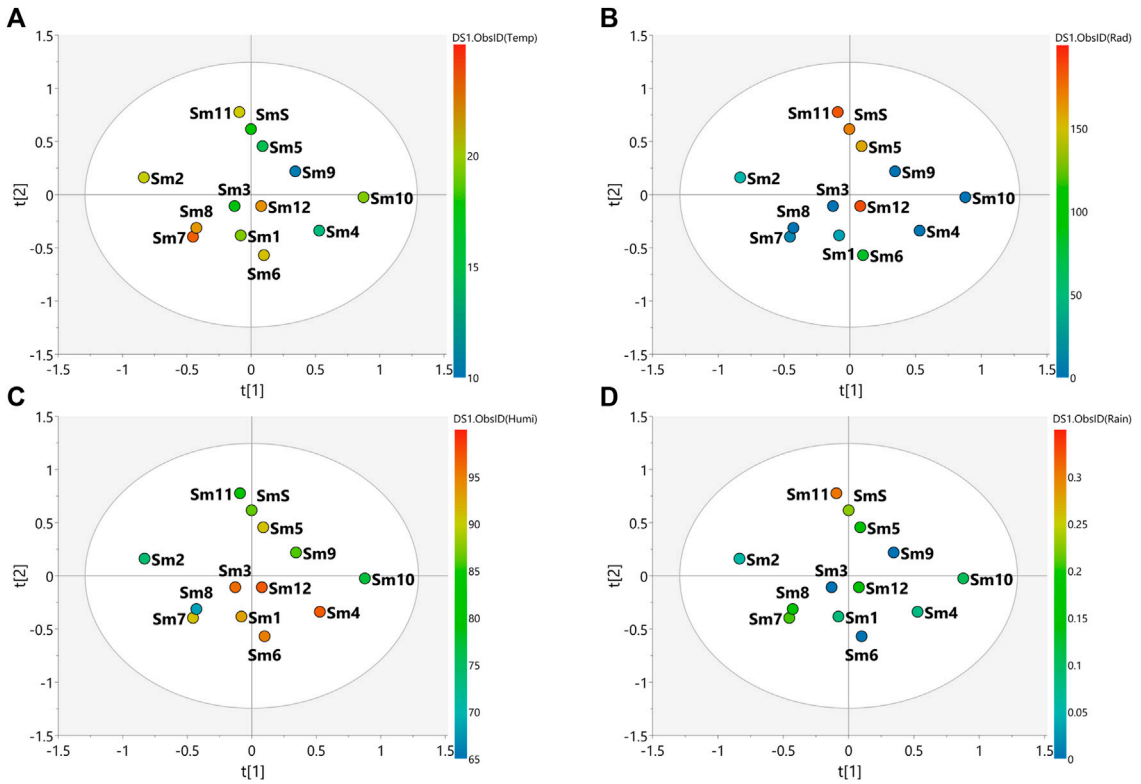


FIGURE 4 | PLS-DA score scatter plots of the MS-NMR fused data of *S. marginata* seasonal harvests and the correlation with the environmental factors. Samples are categorized according to the values of observations for environmental factors. **(A)** Temp = temperature (°C); **(B)** Rad = Solar radiation (kJ/m^2); **(C)** Humi: relative humidity (%) and **(D)** Rain = rainfall (mm). Sample seasons: Summer (Sm1, Sm6 and Sm12); Autumn (Sm2, Sm3, Sm7 and Sm8); Winter (Sm4, Sm9 and Sm10) and Spring (Sm5, Sm11 and SmS).

during the summer and autumn periods (**Figure 4A**), high humidity levels in the summer (**Figure 4C**) and higher rainfall amounts in the spring (**Figure 4D**). The spring season (Sm5, Sm11 and SmS) presented the most separated data in relation to the other groups, indicating differences in the chemical profile in relation to the other samples. In addition, the spring season showed the tightest clusters relative to the other seasons indicating minimal changes in chemical profiles for the samples harvested in this season. The autumn (Sm2, Sm3, Sm7 and Sm8), summer (Sm1, Sm6 and Sm12) and winter (Sm4, Sm9 and Sm10) samples were more scattered indicating a greater variation in the chemical profiles of these samples. Considering the environmental factors, it is more difficult to correlate the variation in chemical composition with the climate variables. In the PLS-DA plot, solar radiation incidence seems to be the variable with the greatest variation, especially with higher values for the spring season and lower values for autumn and winter, which may influence the chemical profile of these samples according to this environmental factor (**Figure 4B**).

To validate the model, OPLS-DA data analysis was used. For this purpose, the samples were separated into more scattered and more clustered groups. The summer season was the most scattered set of samples relative to the other season groups for *B. intermedia* and the spring season was most scattered for *S. marginata*. The Permutations plots (**Supplementary Figure S9**) show the model fit based on the R2 and Q2 values.

A total of 15 variables with the highest VIP scores were selected to validate the chemical features that are significant for group discrimination in the OPLS-DA model for the MS-NMR fused and HRMS data. Subsequently, after validation of the significant NMR regions contributing to seasonal group discrimination and putative annotation of functional groups and classes of compounds, the LC-HRMS data were analyzed for further insights into the annotation of the chemical structures of the discriminant compounds.

Figures 5, **6** show the scatter plots of the OPLS-DA score and the selected variables highlighted in the loading plots of the MS-NMR fused data and HRMS data, for *B. intermedia* and *S. marginata*, respectively. **Supplementary Tables S6 and 7** show the adjusted *p*-values and FDR estimation for each discriminant feature of the MS-NMR fused data and **Supplementary Figures S10 and 11** show the representative 1H NMR and 2D *J*-res NMR spectra with the top 15 peak signals according to the VIP scores. **Supplementary Tables S8 and 9** show the same analysis for the discriminant metabolites found with respect to the results obtained from the MS data. For the discriminating metabolites, the chemical formula was predicted from the exact mass, the most probable ring double bond formula (RDB) was calculated, and the compounds

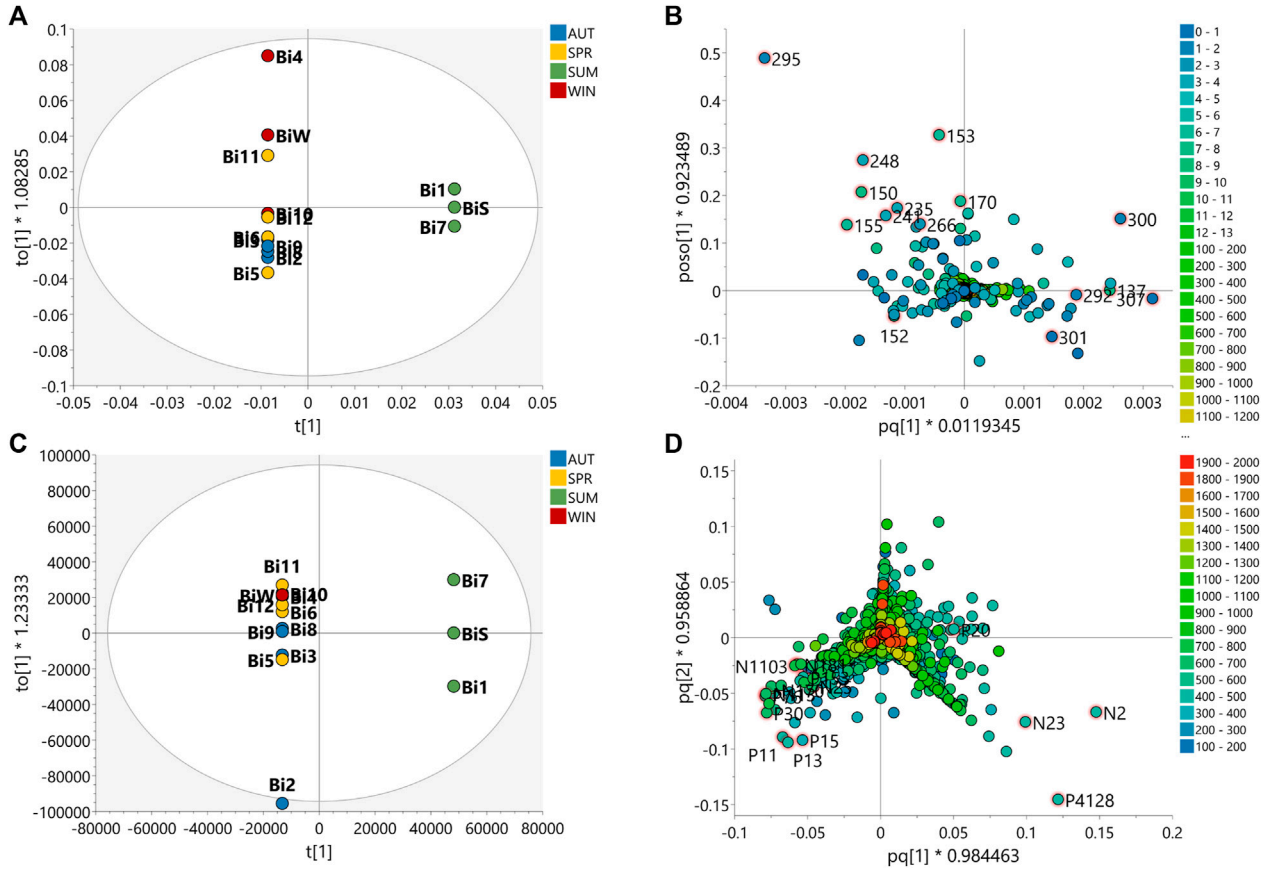


FIGURE 5 | Multivariate analyses of the data of *B. intermedia* species for prediction of the seasonally discriminating metabolites. Orthogonal Partial Least Square-Discriminant Analysis (OPLS-DA) score scatter plot of **(A)** MS-NMR fused data and **(C)** HRMS data with samples categorized according to the harvest season. OPLS-DA loadings plots with highlighted features of the 15 VIP metabolites of **(B)** MS-NMR fused data and **(D)** HRMS data. More scattered samples: Summer (Bi1, Bi7 and BiS); Less scattered samples: Autumn (Bi2, Bi3, Bi8 and Bi9); Winter (Bi4, Bi10 and BiW) and Spring (Bi5, Bi6, Bi11 and Bi12).

were annotated according to the hits in the DNP database. In addition, box-and-wisker plots in **Supplementary Figures S12 and 13** show the relative distribution of the discriminating metabolites according to seasonal harvest of *B. intermedia* and *S. marginata* samples, respectively.

OPLS-DA loadings plot for *B. intermedia*, show the selected VIP features, highlighting signals of chemical shifts in the aliphatic proton region (0.50–5.00 ppm) and aromatic region (6.00–8.00 ppm), and also molecular weights between 400 and 900 Da correlated to summer and autumn seasons (**Figure 5B**). The chemical shift signals at δ 7.53 (**137**), 6.90 (**153**), 1.35 (**292**), 2.39 (**266**), 7.04 (**150**), 3.12 (**248**), 3.36 (**241**), 6.21 (**170**), 3.62 (**235**), 0.77 (**307**) and 1.24 (**295**) presented FDR values lower than 0.05, representing the true positive annotation and indicating the discriminating metabolites associated with seasonal harvests (**Supplementary Table S6**). The signals at δ 7.53 (d, J = 2.3 Hz) and δ 6.21 (d, J = 2.1 Hz) indicate *meta*-couplings corresponding to the protons of the B and A rings of a flavonoid, respectively, which are consistent with the quercetin aglycone, while the signals at δ 3.12, δ 3.36 and δ 3.62 are characteristic of the protons of a saccharide unit, suggesting the presence of a glycosylated flavonoid (Harborne, 1994). The singlets at δ 7.04 and δ 6.90 were assigned to galloyl substituents, which may be attached to the structure of sugars or quinic acid by esterification of hydroxyl groups (Sannomiya et al., 2007). The chemical shifts of the shielded signals in the range δ 0.77–1.35 suggest methyl and methylene protons of triterpenoid compounds (Higuchi et al., 2008). Considering the loadings plot for the discriminant metabolites based on MS data (**Supplementary Figure S5D**), the flavonoids were annotated as quercetin-*O*-hexose (**P13**) and quercetin-*O*-(*O*-galloyl)-hexose (**N25**), the galoyl quinic acid derivatives as di-*O*-galloylquinic acid (**P76**) and tri-*O*-galloylquinic acid (**N113**) and the triterpenes as betulinic acid (**N2**), oleanolic acid (**N23**), β-amyrin (**P4128**) and 3-oxo-olean-12-en-28-al (**P22**) (**Supplementary Table S8** and **Figure S7**). In **Supplementary Figure S12**, it is possible to notice a progressive increase in the concentration of phenolic compounds, such as flavonoids and galloylquinic acid derivatives, since summer until reaching maximum values in winter for *B. intermedia* harvests. On the other hand, the signals related to triterpenes show that these compounds presented higher concentration during the summer and minimum values during the winter.

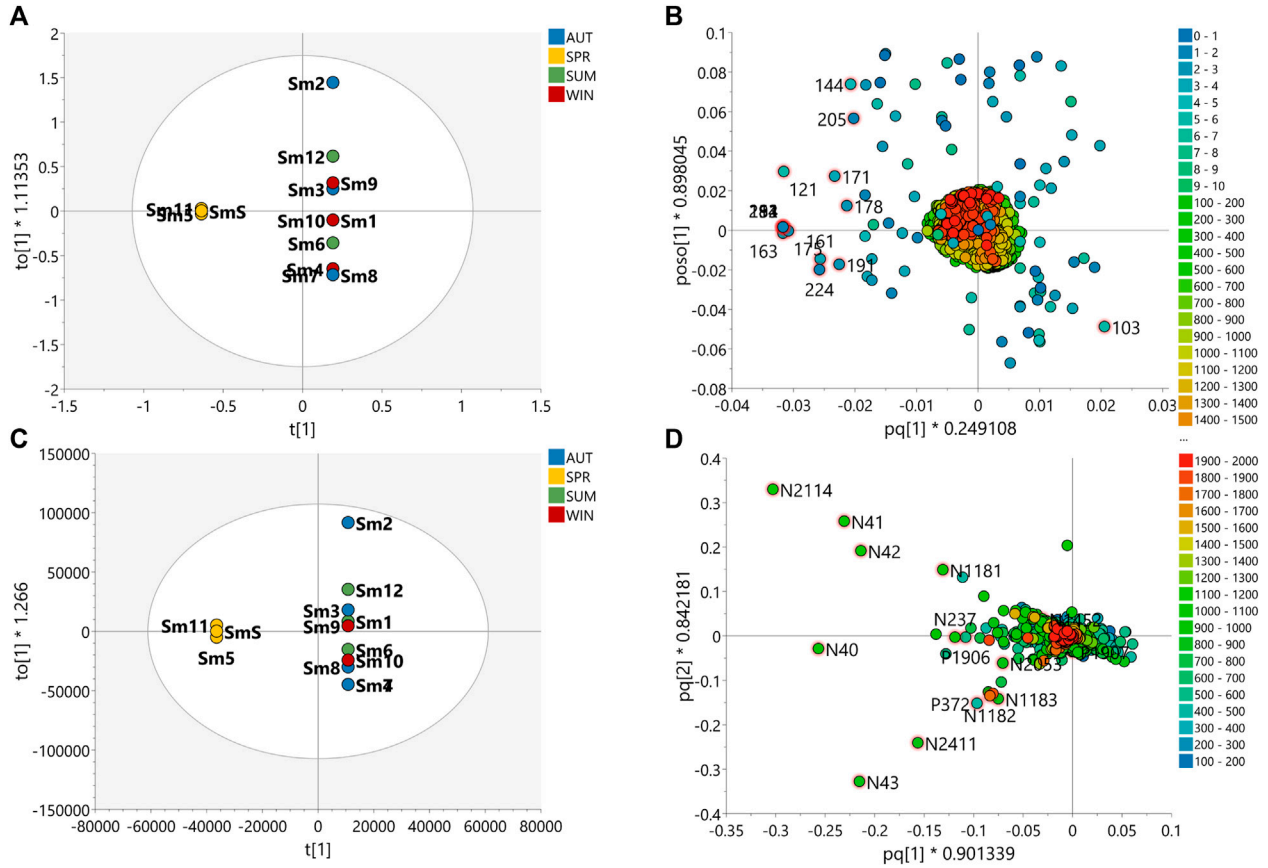


FIGURE 6 | Multivariate analyses of the data of *S. marginata* species for prediction of the seasonally discriminating metabolites. Orthogonal Partial Least Square-Discriminant Analysis (OPLS-DA) score scatter plot of **(A)** MS-NMR fused data and **(C)** HRMS data with samples categorized according to the harvest season. OPLS-DA loadings plots with highlighted features of the 15 VIP metabolites of **(B)** MS-NMR fused data and **(D)** HRMS data. More scattered samples: Spring (Sm5, Sm11 and SmS); Less scattered samples: Summer (Sm1, Sm6 and Sm12); Autumn (Sm2, Sm3, Sm7 and Sm8); Winter (Sm4, Sm9 and Sm10) and Spring (Sm5, Sm11 and SmS).

OPLS-DA loadings of the fused data for *S. marginata* shows compounds with high molecular weight (colored in red) associated with typical chemical shifts of sugar and methylene protons (colored in blue) correlated with spring season (**Figure 6B**). The chemical shift signals at δ 3.99 (**151**), 3.52 (**163**), 2.72 (**182**), 2.27 (**194**), 1.49 (**213**), 5.19 (**121**), 3.03 (**175**), 1.08 (**224**), 3.58 (**161**) and 3.18 (**171**) were significant (FDR ≤0.05) for the differentiation of the seasons (**Supplementary Table S7**). The signals with chemical shifts in the range of δ 1.82–3.18 suggest methylene protons of sapogenins portions of saponin molecules. Analyzing the loadings plot and the annotation of the discriminating metabolites considering the MS data (**Figure 6D**), it was obtained that the steroidal saponins (**N2053** and **N1452**), triterpenic saponins with hederagenin as aglycone (**N237**, **P1907**, **N43**, **N40** and **N2114**) and the flavones cassiaocidentalin A (**P1907**) and tetrastigma B (**P1906**) are the compounds that contribute the most to the separation of the samples (**Supplementary Table S9** and **Figure S7**). Samples harvested during spring had the highest amount of saponin, while winter samples showed the lowest amount. Flavones, on the other hand, had higher concentrations during winter, with a continuous increasing variation from summer to winter, followed by a decreasing in the concentration in the spring season (**Supplementary Figure S13**).

DISCUSSION

The analyses conducted for each set of samples using LC-HRMS and NMR showed that the extracts prepared from the leaves of the plant species *B. intermedia* and *S. marginata* had a comparable metabolite profile. Visual inspection of these data showed a variation in the intensity of the metabolites as a result of different harvest periods.

The computational analysis approach employed for the pre-processing of the untargeted LC-MS data was found to be very useful in terms of obtaining relevant information regarding the peak areas with a more accurate estimation of the relative ions abundance. Based on the application of the computational analysis approach, the present study was able to deconvolve

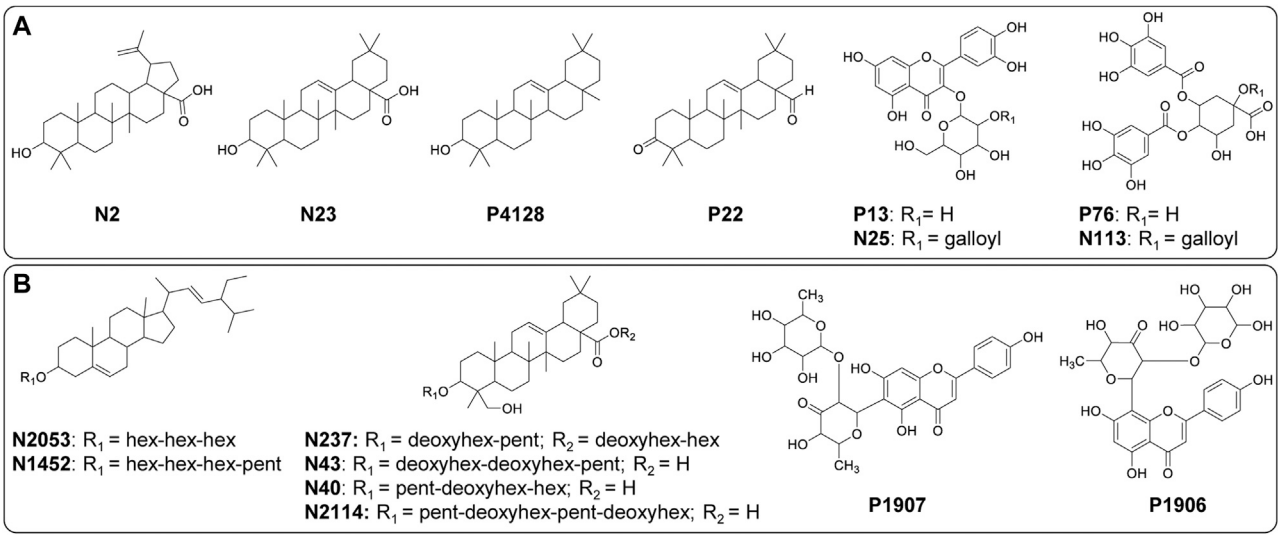


FIGURE 7 | Structures of significant metabolites (FDR ≤ 0.05) for the differentiation of (**A**) *B. intermedia* (**N2**, **N23**, **P4128**, **P22**, **P13**, **N25**, **P76** and **N113**) and (**B**) *S. marginata* (**N2053**, **N1452**, **N237**, **N43**, **N40**, **N2114**, **P1907** and **P1906**) harvests. Annotation of the metabolites is provided in the **Supplementary Tables 7 and 9**.

the co-eluted peak signals and predict the molecular formulas of individual features of the metabolites with a high degree of accuracy. The application of chromatographic analysis promoted a good separation of the main groups of metabolites and the electrospray ionization technique (ESI) led to the unequivocal formation of deprotonated and protonated molecules.

The dereplication of the ESI-HRMS data using the macro database containing the Natural Products Dictionary (DNP) allowed a rapid and effective annotation of 68 compounds in the extracts of *B. intermedia* and 81 compounds in the extracts of *S. marginata* (see **Supplementary Tables S4, S5**).

As mentioned previously in the results section, the metabolite profile of the hydroethanolic extracts of *B. intermedia* revealed that this plant is a rich source of cinnamic acids, phenolic acids derived from galloyl quinic and shikimic acid, proanthocyanidins, glycosylated flavonoids, triterpenes and other phenols (**Supplementary Table S4**). Glycosylated flavonoids were among the major classes of compounds found in the *B. intermedia* plant species. These compounds (glycosylated flavonoids) were characterized as quercetin derivatives which consisted of pentoses, deoxyhexoses and hexoses that are attached to the flavonoid aglycone by *O*-glycosidic linkages. The quinic and shikimic acid derivatives which contain galloyl as substituents were also among the major classes of compounds found in *B. intermedia*; these compounds contained between one to four gallic acid units per molecule (≈300–800 Da).

In the ^{1}H and 2D *J*-res NMR spectra (**Supplementary Figures S3, 5 and 6**), the main signals were observed in the aromatic region (δ 5.70–8.00) and can be assigned to the aromatic rings of the derivatives compounds of gallic acid, quercetin and catechin. The signals at δ 7.53 (d, J = 2.3 Hz), δ 7.67 (dd, J = 2.3; 8.5 Hz) and δ 6.83 (d, J = 8.5 Hz) were assigned to the protons of the trisubstituted aromatic ring. Two doublets with chemical shifts at δ 6.41 (d, J = 2.1 Hz); 6.21 (d, J = 2.1 Hz) indicate a *meta*-coupling corresponding to the protons of the A-ring flavonoid, which are consistent with the quercetin aglycone (Harborne, 1994). Whereas, the A-ring protons of the catechin derivatives are more shielded than those assigned to quercetin, and therefore were assigned to the doublet signals at δ 5.71 (d, J = 2.1 Hz) and δ 5.90 (d, J = 2.1 Hz). The singlet at δ 7.03 is characteristic of the two protons of the gallic acid/galoyl group aromatic ring.

In addition, the signals at δ 5.37 (d, J = 7.7 Hz), 5.27 (d, J = 3.8 Hz) and 4.55 (d, J = 1.4 Hz) were assigned to the anomeric protons of the sugar units glucose, arabinose and rhamnose, respectively and thereby confirming the presence of glycosylated flavonoids (**Supplementary Figures S3, 5 and 6**). The anomeric configurations for the sugar units were deduced from their $^{3}J_{\text{H-1/H-2}}$ coupling constants. The value of J for the signal at δ 5.37 indicates that the glycoside has a β configuration (J > 5.0 Hz), while for signals at δ 5.27 and 4.55 the configuration was assigned as α (J = 0–5.0 Hz) (Agrawal, 1992).

The signals of the compounds of the triterpene class can be observed in the aliphatic region of the 1H and 2D *J*-res NMR spectra; the signals in the range of δ 0.50–3.50 can be attributed to the methyl and methylene protons and some signals in the region of δ 3.50–5.00 to the oxymethynic protons of the triterpene structure (Higuchi et al., 2008).

The extracts of *S. marginata* showed mainly exhibited compounds derived from saponins, glycosylated flavonoids, catechins and oligomeric proanthocyanidins (**Supplementary Table S5**). The detection of saponins in *S. marginata* in this study helped to confirm the findings reported in the literature which pointed out that saponins are the major class of compounds in this plant species (Heredia-Vieira et al., 2015). Precisely, the results of our study showed that triterpenic saponins were prevalent in the extracts of *S. marginata*; these

compounds contained aglycones of oleanolic acid (olean-12-en-28-oic), hederagenin (3,23-dihydroxyolean-12-en-28-acid) and gypsogenin (3-hydroxy-23-oxo-olean-12-en-28-acid). The presence of glycosylated flavonoids was detected in three groups: *C*-glycosylated, *C,O*-glycosylated and *O*-glycosylated flavonoids (**Supplementary Table S5**). The compounds tetrastigma B (*m/z* 561.1603 $[M+H]^+$, R_t = 9.89) and cassiaocidentalin A (*m/z* 561.1602 $[M+H]^+$, R_t = 10.37) are examples of *C,O*-glycosylated flavonoids isomers with an interglycosidic linkage between a deoxyhexose and 6-dideoxyhexose sugars (Heredia-Vieira et al., 2015). Two types of proanthocyanidin compounds were found in *S. marginata*; these included B-type and A-type proanthocyanidins (**Supplementary Table S5**). B-type proanthocyanidins are characterized by a single interflavan linkage between monomeric units which can be formed by the equivalent units of (epi)catechin. A-type proanthocyanidins, on the other hand, contain double interflavan bonds.

The ^{1}H and 2D *J*-res NMR spectra recorded for the *S. marginata* extracts showed intense signals in the aromatic proton region between 5.70 and 8.10 ppm and an aliphatic region between 0.50 and 5.60 ppm, indicating characteristic signals corresponding to compounds of the flavonoid and saponin classes, respectively (**Supplementary Figures S4, 7 and 8**).

As observed in the mass spectrometry analyses, *S. marginata* has mainly flavonoids derived from apigenin and catechin. The most intense signals observed in the spectrum at δ 8.03–7.95 (d, J = 9.1 Hz) and other signals at δ 7.00–6.96 (d, J = 9.1 Hz) in the aromatic region can be attributed to the *ortho* coupling of the H-2′/H-6′ and H-3′/H-5′ protons, respectively, characteristic of a *para*-substituted aromatic ring like those observed for the aglycone derivatives of apigenin. The singlet at δ 6.55 was assigned to H-8 when ring A is substituted at position 6 and the singlet at δ 6.23 was assigned to H-6 when ring A is substituted at position 8. These signals corroborate the proposal that *C*-glycosylated isomers are found with the glycoside attached at the C-6 or C-8 position of the aglycone. Regarding the glycosides attached to flavonoids, the signals at δ 5.36 (d, J = 10.0 Hz) and δ 5.00 (d, J = 10.4 Hz) are consistent with the anomeric protons of a 6-deoxy-ribo-hex-3-ulose sugar units. Also, the doublets in δ 1.41 (d, J = 6.0 Hz) and δ 1.31 (d, J = 6.1 Hz), were assigned to the methyl groups of this sugar unit. Therefore, the data presented corroborate the presence of the compounds tetrastigma B and cassiaocidentalin A (Heredia-Vieira et al., 2015).

The signals in the ^{1}H and 2D *J*-res NMR spectra of double doublet at δ 7.29 (J = 8.3, 2.1 Hz), doublets with *meta* couplings at δ 7.07 (J = 2.3 Hz) and δ 7.03 (J = 2.5 Hz), doublets with *ortho* couplings δ 6.78 (J = 8.2 Hz), δ 6.74 (J = 8.3 Hz) and δ 6.68 (J = 8.2 Hz) and the singlets at δ 6.08 and δ 5.96 are some examples of signals which can be attributed to the rings of the (epi)catechin units of the proanthocyanidins (Kolodziej, 1992; Heredia-Vieira et al., 2015). Furthermore, the signals at δ 4.00–4.95 with coupling constants of 3.3–3.7 Hz are diagnostic features of the AB coupling systems of A-type proanthocyanidins (Jacques et al., 1974; Kolodziej, 1992; Lou et al., 1999).

The main signals observed for the saponins in the ^{1}H and 2D *J*-res NMR spectra were the signals at δ 5.33 (s), δ 5.20 (d, J = 3.7 Hz), δ 5.18 (d, J = 3.5 Hz) and δ 4.30 (d, J = 6.4 Hz), regarding anomeric protons and the presence of singlets between δ 0.73 and 1.31, characteristic of methyl protons attached to the triterpenoid nucleus (Heredia-Vieira et al., 2015).

A number studies published in the literature have shown that these compounds found in each plant species exhibit a wide range of biological activities. Phenolic compounds such as phenolic acids, tannins and flavonoid derivatives found in the plant species have been shown to have good antioxidant properties, which are essentially useful for the prevention and treatment of various inflammatory diseases including gastric ulcer, diabetes and obesity (Salinas-Sánchez et al., 2017; Bakoyiannis et al., 2019; Ma et al., 2019). Inflammatory diseases are often associated with the production and release of pro-inflammatory mediators, free radicals, ROS, and DNA-reactive aldehydes from lipid peroxidation (Anderson et al., 2014; Otimenyin, 2018). Compounds with antioxidant capacity are able to capture free radicals and reduce their creation (i.e. the creation of free radicals), chelate transition metals, and inhibit some enzymes involved in oxidative processes (Huang et al., 2005; Biswas et al., 2017; Tungmunnithum et al., 2018). Some previous studies attributed the anti-inflammatory action of glycosylated flavonoids to their ability to decrease the production of pro-inflammatory mediators and biomarkers of oxidative stress, by suppressing or alleviating the progression of inflammation (Jiang et al., 2017; Fraige et al., 2018; Dantas-Medeiros et al., 2021).

Proanthocyanidins can also act as antioxidants; in fact, according to studies reported in the literature, the bioavailability of these compounds is higher for monomers and small oligomers, such as those found in the extracts of the plant species investigated here (Cos et al., 2004; Rauf et al., 2019). Among the derivatives of gallic acid, such as galloylquinic acid and gallotannins derivatives, the galloyl group is found to possess antioxidant properties; in fact, reports in the literature have shown that an increment in the number of galloyl groups increases the antioxidant activity of the compound (Baratto et al., 2003; Karas et al., 2017). Interestingly, some studies have reported that the addition of a galloyl group to the structure of flavonoids glycosides and proanthocynidins leads to an increase in radical scavenging activity compared to non-galloylated compounds (Karas et al., 2017).

Other compounds that deserve mentioning are triterpenes and saponins. Triterpenes have been shown to have anti-inflammatory, anticarcinogenic, antidiabetic, antitubercular, hepatoprotective, antimicrobial, antimycotic, analgesic, immunomodulatory, and cardiotonic activities (Higuchi et al., 2008; Agra et al., 2015; Dinday and Ghosh, 2020). Saponins have also been found to possess antifungal, molluscicidal, antibacterial, hemolytic, anti-inflammatory, antiparasitic, antitumor, cytotoxic, antiviral, insecticidal, among other biological properties (Ekabo et al., 1996; Sparg et al., 2004; Biswas and Dwivedi, 2019). Some studies have shown that some structural features of triterpenes and saponins are related to an increase in their bioactivities; these features include glycosylation, and/or side-chain methylation or hydroxylation (Zhang et al., 2007; Biswas and Dwivedi, 2019).

Triterpenes are hydrophobic structures while saponins exhibit a hydrophobic aglycone and hydrophilic sugar groups; these differences in the structures of these compounds mean that

saponins, being less lipophilic, are more bioavailable than triterpenes. In addition, these characteristics of saponins structure with non-polar and polar portions, give them surface-active properties and therefore they are generally considered as haemolysing agents (Biswas and Dwivedi, 2019; Dinday and Ghosh, 2020).

Apart from their pharmacological properties, the specialized metabolites found in *B. intermedia* and *S. marginata* species have also been found to play important ecological roles and their chemically distinct groups act in different ways in the defense of these plants. As aforementioned, several studies published in the literature have described the chemical composition of *B. intermedia* and *S. marginata* species and have pointed out the various health benefits that these species bring to humans. However, knowledge is still very limited regarding the relationship between the chemical composition of these plant species and the variation observed in the metabolites when these species are harvested at different periods of the year. Thus, to have a better understanding of the influence of seasonality on the production of metabolites in *B. intermedia* and *S. marginata*, the HRMS and 2D *J*-res NMR data sets were evaluated using multivariate analysis.

Based on the PLS-DA plots (**Figures 1**, **2**) it was possible to visualize the differences in terms of clustering between the samples according to the features used; this shows that MS and NMR contribute uniquely to the analysis of statistical metabolomics analysis of the plant species. In addition, one can say that the data obtained from *J*-res NMR analysis provided highly accurate information, as it shows a greater contribution to the correlation and differentiation of the samples when the data is concatenated.

The permutation plot (**Supplementary Figure S9**) showed that the regression line (in blue) of the Q2-points intersected the vertical axis below zero; this strongly indicates that the model is valid and not overfitted. This result also demonstrates that the model shows good correlation (of data) and a reliable degree of predictability.

Based on the application of variable importance on projection (VIP) and false discovery rate (FDR) estimation techniques, we were able to conduct a better assessment of a great number of responses, and this enabled us to estimate with a high level of confidence, the significant and discriminant chemical variables for the differentiation and correlation of the seasons. The results of this study showed that the concentration of different compounds was directly affected by the harvest period. The differences observed between the samples may be related to the influence of the environmental conditions that are typically characteristic of each period, such as temperature, solar radiation, relative humidity, and rainfall.

B. intermedia and *S. marginata* are plant species that belong to the Brazilian savanna, known as Brazil's Cerrado biome, and for this reason, the harvest location of these species presented very similar weather conditions. The climatological characteristics of the Cerrado indicate the existence of two main periods/seasons: the long, dry and warm season [April/May (autumn) to August/September (winter)] and the short, rainy and hot season [September/October (spring) to February/March (summer)]. According to some studies reported in the literature, drought is the main environmental stress factor that affects plant yield and quality; this factor particularly affects the accumulation of bioactive compounds in the plants, leading to significant changes in the amount of phenolic compounds present in the chemical composition of these plants (Wang et al., 2016; Almeida et al., 2020; Lambers et al., 2020). Water deficiency in periods of drought enhances the formation of reactive oxygen species (ROS), and this leads the plant to activate its defense mechanism by changing the biosynthesis pathway in response to drought stress to produce antioxidant compounds such as, flavonol glycosides and galloylquinic acid derivatives for *B. intermedia* samples and *C*-glycosyl flavones for *S. marginata*, which have the ability to capture free radicals and reduce the oxidative damage (Paudel et al., 2016). Moreover, after a long period of drought stress, both *B. intermedia* and *S. marginata* experience a reduction in the content of phenolic compounds; this is particularly evident in the spring period (Ahmed et al., 2019).

Interestingly, a study carried out by Velasque and Del-Claro (2016), reported slightly different findings; according to this study, the production of new leaves in the species *B. intermedia* follows a seasonal pattern and occurs after rainy periods with predominant growth in dry periods. Thus, the increased levels of phenolic compounds during autumn and winter may also be associated with the defense mechanism of the plant which is related to the resistance of young leaf tissues to plant pathogens (Andrade et al., 2017).

The extremely high levels of triterpenes present in *B. intermedia* during the summer indicates that the accumulation of these compounds strongly depends on warm weather and higher light intensity, which are the predominant environmental factors during this season (Alqahtani et al., 2015). Higher temperatures can lead to higher evaporation rates and decreased water availability, and these factors cause negative effects on the development and growth of the plant (Paudel et al., 2016; de Araújo Silva et al., 2021). In this way, plants avoid dehydration by producing a lipophilic layer coating on the surface of the leaves (Buschhaus and Jetter, 2012). Studies in the literature, have shown that triterpenes are the major components of this cuticular layer present in many plants (Buschhaus and Jetter, 2011; Moses et al., 2014). Furthermore, this barrier is a plant-environment interface against abiotic stresses, and its functions include blocking the loss of water, deterring insects and pathogens, obstructing the penetration of UV rays, protecting the plants from high temperatures and other potential threats that the plant may be exposed to (González-Coloma et al., 2011; Buschhaus and Jetter, 2012; Diarte et al., 2021).

The different concentrations of saponins observed in different harvest periods of *S. marginata* showed that the plant recorded higher levels of saponins during the spring and autumn but had extremely low levels of saponins during the winter. Abiotic stress factors, such as humidity, high temperatures and solar radiation may have influenced the content of saponins in *S. marginata* during the autumn/spring seasons, while drought conditions during the winter may have contributed to the decrease in the content of saponins in the plant. In line with these observations, studies published in the literature have shown that, for some plant species, that

environmental stresses associated with increased saponin levels, are responsible for the mechanism of activation of some signaling agents that promote the biosynthesis of these compounds (Szakiel et al., 2011; de Costa et al., 2013; Moses et al., 2014).

It is worth noting that, in general, saponins possess antimicrobial, antifungal, antiparasitic and insecticidal properties, and these properties enable them to play a key role in the plant defense mechanism against pathogenic microbes, pests and herbivores (Sparg et al., 2004; Moses et al., 2014; Biswas and Dwivedi, 2019).

Other studies, reported that the presence of higher amounts of hederagenin-based saponins in some plants is associated with the plant defense response against herbivores; according to these authors the hydroxylation that occurs at the C-23 position of oleanolic acid and which promotes the formation of the hederagenin structure, is a modification in the chemical composition of the plant which enhances its defense mechanism (Liu et al., 2019).

However, the relationship between metabolite variation and the environmental factors need to be carefully analyzed, since there may be a gap in the physiological temporal response of the plant (Castro et al., 2017). In addition, bushfire (both natural and human) (Lambers et al., 2020), low soil nutrients (Lambers et al., 2020) and high rates of evapotranspiration (Almeida et al., 2020) are common factors that occur in the Cerrado (savanna) region; all these factors could affect the accumulation of compounds in *B. intermedia* and *S. marginata*.

This study brings novel findings to light regarding *B. intermedia* and *S. marginata* and provides relevant insights into the identification of specialized metabolites related to environmental conditions which may be useful for establishing standardization criteria for plant extracts aiming at improving the quality control of medicinal plants.

The results obtained from the present study also showed the relevant role played by the secondary metabolites in the defense mechanisms against stresses caused by environmental factors in each plant species. Drought, UV radiation and temperature were the main stress factors that were found to induce the formation of ROS and the accumulation of specialized metabolites with antioxidant capacity in the plants. The strong antioxidant activity of many of these metabolites was also found to have beneficial effects on human health. The findings of this study showed that the plant species investigated here have suitable medicinal properties which can be useful for the treatment of inflammatory diseases.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

AZ was responsible for Conceptualization, Validation, Formal Analysis, Investigation, Writing - Original draft and Visualization. WG was responsible for Conceptualization, Writing - Review and Editing, Supervision and Project Administration. RE-E was responsible for Conceptualization, Methodology, Investigation, Writing - Review and Editing, Supervision and Funding Acquisition. All authors contributed toward the elaboration of the manuscript and have given their approval of the final version.

FUNDING

The authors are grateful to the following Brazilian research funding agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the financial assistance provided in support of this research and for a grant to W.V.; and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the financial support granted to A.C.Z. (grant#2016/21044-4 and grant#2019/18673-8).

ACKNOWLEDGMENTS

The authors would like to express their sincerest gratitude to the Prof. Dr Silvia Cristina Heredia-Vieira for the harvesting, and to the Prof. Dr Maria do Carmo Vieira for the identification of the leaves of the species *S. marginata*.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.710025/full#supplementary-material

REFERENCES

Agra, L. C., Ferro, J. N. S., Barbosa, F. T., and Barreto, E. (2015). Triterpenes with Healing Activity: A Systematic Review. *J. Dermatol. Treat.* 26, 465–470. doi:10.3109/09546634.2015.1021663

Agrawal, P. K. (1992). NMR Spectroscopy in the Structural Elucidation of Oligosaccharides and Glycosides. *Phytochemistry* 31, 3307–3330. doi:10.1016/0031-9422(92)83678-R

Ahmed, S., Griffin, T. S., Kraner, D., Schaffner, M. K., Sharma, D., Hazel, M., et al. (2019). Environmental Factors Variably Impact Tea Secondary Metabolites in the Context of Climate Change. *Front. Plant Sci.* 10, 939. doi:10.3389/fpls.2019.00939

Almeida, T., Pinto, G., Correia, B., Gonçalves, S., Meijón, M., and Escandón, M. (2020). In-depth Analysis of the Quercus suber Metabolome under Drought Stress and Recovery Reveals Potential Key Metabolic Players. *Plant Sci.* 299, 110606. doi:10.1016/j.plantsci.2020.110606

Alqahtani, A., Tongkao-On, W., Li, K. M., Razmovski-Naumovski, V., Chan, K., and Li, G. Q. (2015). Seasonal Variation of Triterpenes and Phenolic Compounds in australianCentella asiatica(L.) Urb. *Phytochem. Anal.* 26, 436–443. doi:10.1002/pca.2578

Anderson, D. R., Duryee, M. J., Shurmur, S. W., Um, J. Y., Bussey, W. D., Hunter, C. D., et al. (2014). Unique Antibody Responses to Malondialdehyde-Acetaldehyde (MAA)-Protein Adducts Predict Coronary Artery Disease. *PLoS One* 9, e107440. doi:10.1371/journal.pone.0107440

Andrade, B. S., Matias, R., Corrêa, B. O., Oliveira, A. K. M., Guidolin, D. G. F., and Roel, A. R. (2017). Phytochemistry, Antioxidant Potential and Antifungal of Byrsonima Crassifolia on Soil Phytopathogen Control. *Braz. J. Biol.* 78, 140–146. doi:10.1590/1519-6984.166532

Arruda, B. N., Coelho, R. G., Honda, N. K., Ferrazoli, C., Pott, A., and Hiruma-Lima, C. A. (2009). Gastroprotective Effect of Serjania Erecta Radlk (Sapindaceae): Involvement of Sensory Neurons, Endogenous Nonprotein Sulfhydryls, and Nitric Oxide. *J. Med. Food* 12, 1411–1415. doi:10.1089/jmf.2008.0269

Bakoyiannis, I., Daskalopoulou, A., Pergialiotis, V., and Perrea, D. (2019). Phytochemicals and Cognitive Health: Are Flavonoids Doing the Trick? *Biomed. Pharmacother.* 109, 1488–1497. doi:10.1016/j.biopha.2018.10.086

Baratto, M. C., Tattini, M., Galardi, C., Pinelli, P., Romani, A., Visioli, F., et al. (2003). Antioxidant Activity of Galloyl Quinic Derivatives Isolated from P. Lentiscus Leaves. *Free Radic. Res.* 37, 405–412. doi:10.1080/1071576031000068618

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57, 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Berini, J. L., Brockman, S. A., Hegeman, A. D., Reich, P. B., Muthukrishnan, R., Montgomery, R. A., et al. (2018). Combinations of Abiotic Factors Differentially Alter Production of Plant Secondary Metabolites in Five Woody Plant Species in the Boreal-Temperate Transition Zone. *Front. Plant Sci.* 9. 1257. doi:10.3389/fpls.2018.01257

Bernardini, S., Tiezzi, A., Laghezza Masci, V., and Ovidi, E. (2018). Natural Products for Human Health: an Historical Overview of the Drug Discovery Approaches. *Nat. Product. Res.* 32, 1926–1950. doi:10.1080/14786419.2017.1356838

Biswas, S., Das, R., Das, R., and Ray Banerjee, E. (2017). Role of Free Radicals in Human Inflammatory Diseases. *AIMS Biophys.* 4, 596–614. doi:10.3934/biophy.2017.4.596

Biswas, T., and Dwivedi, U. N. (2019). Plant Triterpenoid Saponins: Biosynthesis, *In Vitro* Production, and Pharmacological Relevance. *Protoplasma* 256, 1463–1486. doi:10.1007/s00709-019-01411-0

Botha, L. E., Prinsloo, G., and Deutschländer, M. S. (2018). Variations in the Accumulation of Three Secondary Metabolites in Euclea Undulata Thunb. Var. Myrtina as a Function of Seasonal Changes. *South Afr. J. Bot.* 117, 34–40. doi:10.1016/j.sajb.2018.04.016

Buschhaus, C., and Jetter, R. (2012). Composition and Physiological Function of the Wax Layers Coating Arabidopsis Leaves: β-Amyrin Negatively Affects the Intracuticular Water Barrier. *Plant Physiol.* 160, 1120–1129. doi:10.1104/pp.112.198473

Buschhaus, C., and Jetter, R. (2011). Composition Differences between Epicuticular and Intracuticular Wax Substructures: How Do Plants Seal Their Epidermal Surfaces? *J. Exp. Bot.* 62, 841–853. doi:10.1093/jxb/erq366

Castro, A. H. F., Alvarenga, A. A. d., Barbo, J. P. R. A. D., Mansur, T. D. O. F., and Paula, A. C. C. F. F. d. (2017). Seasonal Patterns of Phenylalanine Ammonia-Lyase Activity and Total Phenol and Tannin Contents in Stryphnodendron Adstringens (Mart.) Coville. *Ciênc. Florest.* 27, 1037. doi:10.5902/1980509828679

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., et al. (2018). MetaboAnalyst 4.0: towards More Transparent and Integrative Metabolomics Analysis. *Web Serv. Issue Publ.* 46, W486–W494. online 46. doi:10.1093/nar/gky310

Cortelo, P. C., Demarque, D. P., Dusi, R. G., Albernaz, L. C., Braz-Filho, R., Goncharova, E. I., et al. (2021). A Molecular Networking Strategy: High-Throughput Screening and Chemical Analysis of Brazilian Cerrado Plant Extracts against Cancer Cells. *Cells* 10, 691. doi:10.3390/cells10030691

Cos, P., Bruyne, T., Hermans, N., Apers, S., Berghe, D., and Vlietinck, A. (2004). Proanthocyanidins in Health Care: Current and New Trends. *Curr. Med. Chem.* 11, 1345–1359. doi:10.2174/0929867043365288

Cragg, G. M., and Newman, D. J. (2013). Natural Products: A Continuing Source of Novel Drug Leads. *Biochim. Biophys. Acta* 1830, 3670–3695. doi:10.1016/j.bbagen.2013.02.008

Dantas-Medeiros, R., Furtado, A. A., Zanatta, A. C., Torres-Rêgo, M., Guimarães Lourenço, E. M., Ferreira Alves, J. S., et al. (2021). Mass Spectrometry Characterization of Commiphora Leptophloeos Leaf Extract and Preclinical Evaluation of Toxicity and Anti-inflammatory Potential Effect. *J. Ethnopharmacol.* 264, 113229. doi:10.1016/j.jep.2020.113229

de Araújo Silva, M. M., Ferreira, L. T., de Vasconcelos, F. M. T., Willadino, L., Camara, T. R., dos Santos, D. Y. A. C., et al. (2021). Water Stress-Induced Responses in the Growth, Cuticular Wax Composition, Chloroplast Pigments and Soluble Protein Content, and Redox Metabolism of Two Genotypes of Ricinus communis L. *J. Plant Growth Regul.* 40, 342–352. doi:10.1007/s00344-020-10103-6

de Costa, F., Yendo, A. C. A., Fleck, J. D., Gosmann, G., and Fett-Neto, A. G. (2013). Accumulation of a Bioactive Triterpene Saponin Fraction of Quillaja Brasiliensis Leaves Is Associated with Abiotic and Biotic Stresses. *Plant Physiol. Biochem.* 66, 56–62. doi:10.1016/j.plaphy.2013.02.003

Di Stasi, L. C., Costa, M., Mendaçolli, S. L. J., Kirizawa, M., Gomes, C., and Trolin, G. (1988). Screening in Mice of Some Medicinal Plants Used for Analgesic Purposes in the State of SãO Paulo. *J. Ethnopharmacol.* 24, 205–211. doi:10.1016/0378-8741(88)90153-5

Diarte, C., Xavier de Souza, A., Staiger, S., Deininger, A.-C., Bueno, A., Burghardt, M., et al. (2021). Compositional, Structural and Functional Cuticle Analysis of Prunus Laurocerasus L. Sheds Light on Cuticular Barrier Plasticity. *Plant Physiol. Biochem.* 158, 434–445. doi:10.1016/j.plaphy.2020.11.028

Dinday, S., and Ghosh, S. (2020). "Triterpenoids: Structural Diversity, Biosynthetic Pathway, and Bioactivity," in *Studies in Natural Products Chemistry*. Editor A. Rahman (Amsterdam: Elsevier), 411–461. doi:10.1016/B978-0-12-819483-6.00012-6

dos Santos, R. de. C., Bonamin, F., Périco, L. L., Rodrigues, V. P., Zanatta, A. C., Rodrigues, C. M., et al. (2019). Byrsonima Intermedia A. Juss Partitions Promote Gastroprotection against Peptic Ulcers and Improve Healing through Antioxidant and Anti-inflammatory Activities. *Biomed. Pharmacother.* 111, 1112–1123. doi:10.1016/j.biopha.2018.12.132

Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., and Griffin, J. L. (2011). Systems Level Studies of Mammalian Metabolomes: The Roles of Mass Spectrometry and Nuclear Magnetic Resonance Spectroscopy. *Chem. Soc. Rev.* 40, 387–426. doi:10.1039/b906712b

Ekabo, O. A., Farnsworth, N. R., Henderson, T. O., Mao, G., and Mukherjee, R. (1996). Antifungal and Molluscicidal Saponins fromSerjaniasalzmanniana. *J. Nat. Prod.* 59, 431–435. doi:10.1021/np960208r

Erb, M., and Kliebenstein, D. J. (2020). Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. *Plant Physiol.* 184, 39–52. doi:10.1104/pp.20.00433

Forshed, J., Idborg, H., and Jacobsson, S. P. (2007). Evaluation of Different Techniques for Data Fusion of LC/MS and 1H-NMR. *Chemometrics Intell. Lab. Syst.* 85, 102–109. doi:10.1016/j.chemolab.2006.05.002

Fraige, K., Dametto, A. C., Zeraik, M. L., de Freitas, L., Saraiva, A. C., Medeiros, A. I., et al. (2018). Dereplication by HPLC-DAD-ESI-MS/MS and Screening for Biological Activities of Byrsonima Species (Malpighiaceae). *Phytochem. Anal.* 29, 196–204. doi:10.1002/pca.2734

M. Ghorbanpour and A. Varma (Editors) (2017). *Medicinal Plants and Environmental Challenges*. (Cham: Springer International Publishing). doi:10.1007/978-3-319-68717-9

González-Coloma, A., López-Balboa, C., Santana, O., Reina, M., and Fraga, B. M. (2011). Triterpene-based Plant Defenses. *Phytochem. Rev.* 10, 245–260. doi:10.1007/s11101-010-9187-8

Guilhon-Simplicio, F., Machado, T. M., do Nascimento, L. F., Souza, R. d. S., Koolen, H. H. F., da Silva, F. M. A., et al. (2017). Chemical Composition and Antioxidant, Antinociceptive, and Anti-inflammatory Activities of Four Amazonian Byrsonima Species. *Phytother. Res.* 31, 1686–1693. doi:10.1002/ptr.5884

Hackstadt, A. J., and Hess, A. M. (2009). Filtering for Increased Power for Microarray Data Analysis. *BMC Bioinformatics* 10, 11. doi:10.1186/1471-2105-10-11

Han, Y., Zhang, A.-H., Zhang, Y.-Z., Sun, H., Meng, X.-C., and Wang, X.-J. (2018). Chemical Metabolomics for Investigating the Protective Effectiveness ofAcanthopanax senticosusHarms Leaf against Acute Promyelocytic Leukemia. *RSC Adv.* 8, 11983–11990. doi:10.1039/C8RA01029C

Harvey, A. L., Edrada-Ebel, R., and Quinn, R. J. (2015). The Re-emergence of Natural Products for Drug Discovery in the Genomics Era. *Nat. Rev. Drug Discov.* 14, 111–129. doi:10.1038/nrd4510

Heredia-Vieira, S. C., Simonet, A. M., Vilegas, W., and Macías, F. A. (2015). Unusual C,O-Fused Glycosylapigenins fromSerjania marginataLeaves. *J. Nat. Prod.* 78, 77–84. doi:10.1021/np500715x

Higuchi, C. T., Pavan, F. R., Leite, C. Q. F., Sannomiya, M., Vilegas, W., Leite, S. R. D. A., et al. (2008). Triterpenes and Antitubercular Activity of Byrsonima Crassa. *Quím. Nova* 31, 1719–1721. doi:10.1590/S0100-40422008000700023

Houriet, J., Allard, P.-M., Queiroz, E. F., Marcourt, L., Gaudry, A., Vallin, L., et al. (2020). A Mass Spectrometry Based Metabolite Profiling Workflow for Selecting Abundant Specific Markers and Their Structurally Related Multi-Component Signatures in Traditional Chinese Medicine Multi-Herb Formulae. *Front. Pharmacol.* 11, 578346. doi:10.3389/fphar.2020.578346

Huang, D., Ou, B., and Prior, R. L. (2005). The Chemistry behind Antioxidant Capacity Assays. *J. Agric. Food Chem.* 53, 1841–1856. doi:10.1021/jf030723c

Isah, T. (2019). Stress and Defense Responses in Plant Secondary Metabolites Production. *Biol. Res.* 52, 39. doi:10.1186/s40659-019-0246-3

Jacques, D., Haslam, E., Bedford, G. R., and Greatbanks, D. (1974). Plant Proanthocyanidins. Part II. Proanthocyanidin-A2 and its Derivatives. *J. Chem. Soc. Perkin Trans. 1* 1, 2663–2671. doi:10.1039/p19740002663

J. B. Harborne (1994). *The Flavonoids: Advances in Research Since 1986* (London: Chapman & Hall). Available at: https://pubs.acs.org/doi/abs/10.1021/ed072pA73.11.

Jiang, W.-J., Daikonya, A., Ohkawara, M., Nemoto, T., Noritake, R., Takamiya, T., et al. (2017). Structure-activity Relationship of the Inhibitory Effects of Flavonoids on Nitric Oxide Production in RAW264.7 Cells. *Bioorg. Med. Chem.* 25, 779–788. doi:10.1016/j.bmc.2016.11.055

Karas, D., Ulrichová, J., and Valentová, K. (2017). Galloylation of Polyphenols Alters Their Biological Activity. *Food Chem. Toxicol.* 105, 223–240. doi:10.1016/j.fct.2017.04.021

Katajamaa, M., Miettinen, J., and Oresic, M. (2006). MZmine: Toolbox for Processing and Visualization of Mass Spectrometry Based Molecular Profile Data. *Bioinformatics* 22, 634–636. doi:10.1093/bioinformatics/btk039

Katz, L., and Baltz, R. H. (2016). Natural Product Discovery: Past, Present, and Future. *J. Ind. Microbiol. Biotechnol.* 43, 155–176. doi:10.1007/s10295-015-1723-5

Kellogg, J. J., Graf, T. N., Paine, M. F., McCune, J. S., Kvalheim, O. M., Oberlies, N. H., et al. (2017). Comparison of Metabolomics Approaches for Evaluating the Variability of Complex Botanical Preparations: Green Tea (Camellia Sinensis) as a Case Study. *J. Nat. Prod.* 80, 1457–1466. doi:10.1021/acs.jnatprod.6b01156

Kolodziej, H. (1992). "1H NMR Spectral Studies of Procyanidin Derivatives: Diagnostic 1H NMR Parameters Applicable to the Structural Elucidation of Oligomeric Procyanidins," in *In* Plant Polyphenols - Synthesis, Properties, Significance. Editors R. W. Hemingway and P. E. Laks (New York: Plenum Press)), 295–319. doi:10.1007/978-1-4615-3476-1_17

Kunle, O. F., Egharevba, H. O., and Ahmadu, P. O. (2012). Standardization of Herbal Medicines - A Review. *Int. J. Biodvers. Conserv.* 4, 101–112. doi:10.5897/IJBC11.163

Lahsen, M., Bustamante, M. M. C., and Dalla-Nora, E. L. (2016). Undervaluing and Overexploiting the Brazilian Cerrado at Our Peril. *Environ. Sci. Pol. Sustain. Develop.* 58, 4–15. doi:10.1080/00139157.2016.1229537

Lambers, H., de Britto Costa, P., Oliveira, R. S., and Silveira, F. A. O. (2020). Towards More Sustainable Cropping Systems: Lessons from Native Cerrado Species. *Theor. Exp. Plant Physiol.* 32, 175–194. doi:10.1007/s40626-020-00180-z

Liebelt, D. J., Jordan, J. T., and Doherty, C. J. (2019). Only a Matter of Time: the Impact of Daily and Seasonal Rhythms on Phytochemicals. *Phytochem. Rev.* 18, 1409–1433. doi:10.1007/s11101-019-09617-z

Liu, Q., Khakimov, B., Cárdenas, P. D., Cozzi, F., Olsen, C. E., Jensen, K. R., et al. (2019). The Cytochrome P450 CYP72A552 Is Key to Production of Hederagenin-based Saponins that Mediate Plant Defense against Herbivores. *New Phytol.* 222, 1599–1609. doi:10.1111/nph.15689

Lou, H., Yamazaki, Y., Sasaki, T., Uchida, M., Tanaka, H., and Oka, S. (1999). A-type Proanthocyanidins from Peanut Skins. *Phytochemistry* 51, 297–308. doi:10.1016/S0031-9422(98)00736-5

Ma, X., Moilanen, J., Laaksonen, O., Yang, W., Tenhu, E., and Yang, B. (2019). Phenolic Compounds and Antioxidant Activities of tea-type Infusions Processed from Sea Buckthorn (Hippophaë Rhamnoides) Leaves. *Food Chem.* 272, 1–11. doi:10.1016/j.foodchem.2018.08.006

Macintyre, L., Zhang, T., Viegelmann, C., Martinez, I., Cheng, C., Dowdells, C., et al. (2014). Metabolomic Tools for Secondary Metabolite Discovery from Marine Microbial Symbionts. *Mar. Drugs* 12, 3416–3448. doi:10.3390/md12063416

Mannochio-Russo, H., Bueno, P. C. P., Bauermeister, A., de Almeida, R. F., Dorrestein, P. C., Cavalheiro, A. J., et al. (2020). Can Statistical Evaluation Tools for Chromatographic Method Development Assist in the Natural Products Workflow? A Case Study on Selected Species of the Plant Family Malpighiaceae. *J. Nat. Prod.* 83, 3239–3249. doi:10.1021/acs.jnatprod.0c00495

Moreira, L. Q., Vilela, F. C., Orlandi, L., Dias, D. F., Santos, A. L. A., da Silva, M. A., et al. (2011). Anti-inflammatory Effect of Extract and Fractions from the Leaves of Byrsonima Intermedia A. Juss. In Rats. *J. Ethnopharmacol.* 138, 610–615. doi:10.1016/j.jep.2011.10.006

Moreira, R. P. de. M., Augusto, C., Batista, S., and Neto, G. G. (2013). "Check List" de Angiospermas da vegetação marginal da estrada Santo Antônio de Leverger - Mimoso, Pantanal de Mato Grosso. Available at: https://periodicoscientificos.ufmt.br/ojs/index.php/flovet/article/view/1524 (Accessed November 20, 2020).

Moses, T., Papadopoulou, K. K., and Osbourn, A. (2014). Metabolic and Functional Diversity of Saponins, Biosynthetic Intermediates and Semi-synthetic Derivatives. *Crit. Rev. Biochem. Mol. Biol.* 49, 439–462. doi:10.3109/10409238.2014.953628

Nogueira, R. C., Paiva, R., Castro, A. H. d., Vieira, C. V., Abbade, L. C., and Alvarenga, A. A. (2004). Germinação *In Vitro* de murici-pequeno (Byrsonima intermedia A. Juss.). *Ciênc. Agrotec.* 28, 1053–1059. doi:10.1590/S1413-70542004000500012

Oliveira, M. I. B., Polido, C. do. A., Costa, L. C., and Fava, W. S. (2007). Sistema reprodutivo e polinização de Byrsonima intermedia A. Juss. (Malpighiaceae) em Mato Grosso Do Sul. Brasil. *Rev. Bras. Biociências* 5, 756–758. Available at: http://www.ufrgs.br/seerbio/ojs/index.php/rbb/article/viewFile/802/650 (Accessed November 20, 2020).

Otimenyin, S. O. (2018). Antiinflammatory Medicinal Plants. *Nat. Prod. Drug Discov.*, 411–431. doi:10.1016/B978-0-08-102081-4.00015-0

Pandey, P., and Senthil-Kumar, M. (2019). Plant-pathogen Interaction in the Presence of Abiotic Stress: What Do We Know about Plant Responses?. *Plant Physiol. Rep.* 24, 541–549. doi:10.1007/s40502-019-00483-7

Paudel, G., Bilova, T., Schmidt, R., Greifenhagen, U., Berger, R., Tarakhovskaya, E., et al. (2016). Osmotic Stress Is Accompanied by Protein Glycation inArabidopsis Thaliana. *Exbotj* 67, 6283–6295. doi:10.1093/jxb/erw395

Pereira, V. V., Borel, C. R., and Silva, R. R. (2015). Phytochemical Screening, Total Phenolic Content and Antioxidant Activity ofByrsonimaspecies. *Nat. Product. Res.* 29, 1461–1465. doi:10.1080/14786419.2014.1002407

Périco, L. L., Heredia-Vieira, S. C., Beserra, F. P., de Cássia dos Santos, R., Weiss, M. B., Resende, F. A., et al. (2015). Does the Gastroprotective Action of a Medicinal Plant Ensure Healing Effects? An Integrative Study of the Biological Effects of Serjania Marginata Casar. (Sapindaceae) in Rats. *J. Ethnopharmacol.* 172, 312–324. doi:10.1016/j.jep.2015.06.025

Pluskal, T., Castillo, S., Villar-Briones, A., and Orešič, M. (2010). MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* 11. doi:10.1186/1471-2105-11-395

Pluskal, T., Uehara, T., and Yanagida, M. (2012). Highly Accurate Chemical Formula Prediction Tool Utilizing High-Resolution Mass Spectra, MS/MS Fragmentation, Heuristic Rules, and Isotope Pattern Matching. *Anal. Chem.* 84, 4396–4403. doi:10.1021/ac3000418

Rai, A., Saito, K., and Yamazaki, M. (2017). Integrated Omics Analysis of Specialized Metabolism in Medicinal Plants. *Plant J.* 90, 764–787. doi:10.1111/tpj.13485

Rauf, A., Imran, M., Abu-Izneid, T., Haq, I.-U., Patel, S., Pan, X., et al. (2019). Proanthocyanidins: A Comprehensive Review. *Biomed. Pharmacother.* 116, 108999. doi:10.1016/j.biopha.2019.108999

Rodal, M. J. N., and Nascimento, L. M. d. (2002). Levantamento florístico da floresta serrana da reserva biológica de Serra Negra, microrregião de Itaparica, Pernambuco, Brasil. *Acta Bot. Bras.* 16, 481–500. doi:10.1590/S0102-33062002000400009

Salinas-Sánchez, D., Jiménez-Ferrer, E., Sánchez-Sánchez, V., Zamilpa, A., González-Cortazar, M., Tortoriello, J., et al. (2017). Anti-Inflammatory Activity of a Polymeric Proanthocyanidin from Serjania Schiedeana. *Molecules* 22, 863. doi:10.3390/molecules22060863

Sampaio, B. L., Edrada-Ebel, R., and Da Costa, F. B. (2016). Effect of the Environment on the Secondary Metabolic Profile of Tithonia Diversifolia: a Model for Environmental Metabolomics of Plants. *Sci. Rep.* 6, 29265. doi:10.1038/srep29265

Sannomiya, M., Cardoso, C. R. P., Figueiredo, M. E., Rodrigues, C. M., dos Santos, L. C., dos Santos, F. V., et al. (2007). Mutagenic Evaluation and Chemical Investigation of Byrsonima Intermedia A. Juss. Leaf Extracts. *J. Ethnopharmacol.* 112, 319–326. doi:10.1016/j.jep.2007.03.014

Santos, R. C., Kushima, H., Rodrigues, C. M., Sannomiya, M., Rocha, L. R. M., Bauab, T. M., et al. (2012). Byrsonima Intermedia A. Juss.: Gastric and Duodenal Anti-ulcer, Antimicrobial and Antidiarrheal Effects in Experimental Rodent Models. *J. Ethnopharmacol.* 140, 203–212. doi:10.1016/j.jep.2011.12.008

Sehlakgwe, P. F., Lall, N., and Prinsloo, G. (2020). 1H-NMR Metabolomics and LC-MS Analysis to Determine Seasonal Variation in a Cosmeceutical Plant Leucosidea Sericea. *Front. Pharmacol.* 11, 1–11. doi:10.3389/fphar.2020.00219

Sparg, S. G., Light, M. E., and van Staden, J. (2004). Biological Activities and Distribution of Plant Saponins. *J. Ethnopharmacol.* 94, 219–243. doi:10.1016/j.jep.2004.05.016

Sprengel-Lima, C., and Rezende, A. A. (2013). Sapindaceae Do noroeste paulista: lista de espécies e chave de identificação baseada em caracteres vegetativos. *Biota Neotrop.* 13, 270–282. doi:10.1590/S1676-06032013000200026

Sut, S., Dall'Acqua, S., Uysal, S., Zengin, G., Aktumsek, A., Picot-Allain, C., et al. (2019). LC-MS, NMR Fingerprint of Potentilla Argentea and Potentilla Recta Extracts and Their *In Vitro* Biopharmaceutical Assessment. *Ind. Crops Prod.* 131, 125–133. doi:10.1016/j.indcrop.2019.01.047

Szakiel, A., Pączkowski, C., and Henry, M. (2011). Influence of Environmental Abiotic Factors on the Content of Saponins in Plants. *Phytochem. Rev.* 10, 471–491. doi:10.1007/s11101-010-9177-x

Tungmunnithum, D., Thongboonyou, A., Pholboon, A., and Yangsabai, A. (2018). Flavonoids and Other Phenolic Compounds from Medicinal Plants for Pharmaceutical and Medical Aspects: An Overview. *Medicines* 5, 93. doi:10.3390/medicines5030093

Velasque, M., and Del-Claro, K. (2016). Host Plant Phenology May Determine the Abundance of an Ecosystem Engineering Herbivore in a Tropical savanna. *Ecol. Entomol.* 41, 421–430. doi:10.1111/een.12317

Verdam, M. C. D. S., Guilhon-Simplicio, F., Andrade, K. C. d., Fernandes, K. L. M., Machado, T. M., Da Silva, F. M. A., et al. (2017). Analgesic, Anti-Inflammatory, and Antioxidant Activities of Byrsonima Duckeana W. R. Anderson (Malpighiaceae). *Scientific World J.* 2017, 1–8. doi:10.1155/2017/8367042

Wang, B., Shi, Z., Weber, G. F., and Kennedy, M. A. (2013). Introduction of a New Critical P Value Correction Method for Statistical Significance Analysis of Metabonomics Data. *Anal. Bioanal. Chem.* 405, 8419–8429. doi:10.1007/s00216-013-7284-4

Wang, W., Xin, H., Wang, M., Ma, Q., Wang, L., Kaleri, N. A., et al. (2016). Transcriptomic Analysis Reveals the Molecular Mechanisms of Drought-Stress-Induced Decreases in Camellia Sinensis Leaf Quality. *Front. Plant Sci.* 7, 385. doi:10.3389/fpls.2016.00385

Wolfender, J.-L., Marti, G., Thomas, A., and Bertrand, S. (2015). Current Approaches and Challenges for the Metabolite Profiling of Complex Natural Extracts. *J. Chromatogr. A* 1382, 136–164. doi:10.1016/j.chroma.2014.10.091

Wolfender, J.-L., Nuzillard, J.-M., Van Der Hooft, J. J. J., Renault, J.-H., and Bertrand, S. (2019). Accelerating Metabolite Identification in Natural Product Research: Toward an Ideal Combination of Liquid Chromatography-High-Resolution Tandem Mass Spectrometry and NMR Profiling,in SilicoDatabases, and Chemometrics. *Anal. Chem.* 91, 704–742. doi:10.1021/acs.analchem.8b05112

Xiao, C., Dai, H., Liu, H., Wang, Y., and Tang, H. (2008). Revealing the Metabonomic Variation of Rosemary Extracts Using 1H NMR Spectroscopy and Multivariate Data Analysis. *J. Agric. Food Chem.* 56, 10142–10153. doi:10.1021/jf8016833

Yang, L. Y., Yang, S. L., Li, J. Y., Ma, J. H., Pang, T., Zou, C. M., et al. (2018). Effects of Different Growth Temperatures on Growth, Development, and Plastid Pigments Metabolism of Tobacco (Nicotiana Tabacum L.) Plants. *Bot. Stud.* 59, 5. doi:10.1186/s40529-018-0221-2

Zhang, Z.-K., Ouyang, M.-A., Wu, Z.-J., Lin, Q.-Y., and Xie, L.-H. (2007). Structure-activity Relationship of Triterpenes and Triterpenoid Glycosides against Tobacco Mosaic Virus. *Planta Med.* 73, 1457–1463. doi:10.1055/s-2007-990263

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.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 12 August 2021
doi: 10.3389/fchem.2021.708364

In Silico Pharmacokinetics, ADMET Study and Conceptual DFT Analysis of Two Plant Cyclopeptides Isolated From Rosaceae as a Computational Peptidology Approach

Norma Flores-Holguín[1], *Juan Frau*[2] *and Daniel Glossman-Mitnik*[1]*

[1]*Laboratorio Virtual NANOCOSMOS, Departamento de Medio Ambiente y Energía, Centro de Investigación en Materiales Avanzados, Chihuahua, Mexico,* [2]*Departament de Química, Universitat de les Illes Balears, Palma de Mallorca, Spain*

This research presents the outcomes of a computational determination of the chemical reactivity and bioactivity properties of two plant cyclopeptides isolated from Rosaceae through the consideration of Computational Peptidology (CP), a protocol employed previously in the research of similar molecular systems. CP allows the prediction of the global and local descriptors that are the integral foundations of Conceptual Density Functional Theory (CDFT) and which could help in getting in the understanding of the chemical reactivity properties of the two plant cyclopeptides under study, hoping that they could be related to their bioactivity. The methodology based on the Koopmans in DFT (KID) approach and the MN12SX/Def2TZVP/H2O model chemistry has been successfully validated. Various Chemoinformatics tools have been used to improve the process of virtual screening, thus identifying some additional properties of these two plant cyclopeptides connected to their ability to behave as potentially useful drugs. With the further objective of analyzing their bioactivity, the CP protocol is complemented with the estimation of some useful parameters related to pharmacokinetics, their predicted biological targets, and the Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) parameters related to the bioavailability of the two plant cyclopeptides under study are also reported.

Keywords: plant cyclopeptides, pharmacokinetics, ADMET, conceptual density functional theory, Koopmans in density functional theory

OPEN ACCESS

Edited by:
Jorge M. del Campo, Universidad Nacional Autónoma de México, Mexico

Reviewed by:
Ramon Miranda Quintana, University of Florida, United States
Tanmoy Chakraborty, Sharda University, India

***Correspondence:**
Daniel Glossman-Mitnik
daniel.glossman@cimav.edu.mx

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: *11 May 2021*
Accepted: *08 July 2021*
Published: *12 August 2021*

Citation:
Flores-Holguín N, Frau J and Glossman-Mitnik D (2021) In Silico Pharmacokinetics, ADMET Study and Conceptual DFT Analysis of Two Plant Cyclopeptides Isolated From Rosaceae as a Computational Peptidology Approach. Front. Chem. 9:708364. doi: 10.3389/fchem.2021.708364

1 INTRODUCTION

Plant-based bioactive compounds have drawn attention of all communities around the world due to their unique biochemical activities and health benefits. Research studies have confirmed the safeguarding effects of certain plant-based diets on cardiovascular diseases, obesity, cancer, diabetes, etc. (Guha et al., 2021). Plant-based drugs from secondary metabolites constitute more than 25% of approved new drugs during last 30 years. Also, 50% of the commercially successful medicinal components were developed based on knowledge from plant secondary metabolites and their structures (Chaudhari and Chakraborti, 2021).

Bioactive peptides are organic substances formed by amino acids joined by covalent bonds known as amide or peptide bonds. Although peptides can exist free in terrestrial plants and marine sources, the vast majority of known bioactive peptides are enclosed within the structure

of the proteins and can be released using enzymatic processes. Bioactive peptides play a significant role in human health by affecting the digestive, endocrine, cardiovascular, immune, and nervous systems. The increasing interest in bioactive peptides has incentivized the scientific community in the exploration and development of new therapeutic drugs based on these peptides (Sánchez and Vázquez, 2017).

Cyclic peptides can be considered as an alternative scaffold. The smaller size and several functional groups of peptides help to make the contact area large enough to provide good selectivity. Their ability to form several hydrogen bonds make easier to obtain strong binding affinity. Moreover, the cyclization of peptides helps in the generation of structural and functional features that are considered to be critical for their use as pharmaceutical drugs, including resistance to degradation by blood proteases. Also, the cyclization of the peptides facilitates the passage through the cell membrane. Because of such favorable features, many cyclic peptides from terrestrial plant and marine sources and their derivatives have been considered for drug design and development (Gang et al., 2018). Besides these biological features, cyclopeptides have smaller sizes than proteins and reduced flexibilities, exhibiting lower conformations than their linear counterparts, thus making easier and affordable the DFT calculations of their structures and properties.

By considering that the knowledge of the chemical reactivity properties of a given molecule is essential for the development of new therapeutic drugs, we are currently researching on new families of cyclopeptides obtained from terrestrial plants and marine sources hoping that the obtained information could be of help for the design of pharmaceutical based on these peptides (Kim, 2013). The objective of the present work is to report the global and local chemical reactivity descriptors of two plant cyclopeptides, Pashinintides A and B, that have been isolated from Rosacea (Cai et al., 2014) by making use of the Conceptual DFT (CDFT) methodology (Parr and Yang, 1989; Chermette, 1999; Geerlings et al., 2003; Toro-Labbé, 2007; Chattaraj, 2009; Geerlings et al., 2020; Chakraborty and Chattaraj, 2021). A recent review has highlighted the basic electronic structure principles and various reactivity descriptors defined within the premise of CDFT (Chakraborty and Chattaraj, 2021). The study is complemented by considering the report of some additional properties of potential application in Structure Activity Relationships (SAR) research for the development of therapeutic drugs, and also with the bioactivity radars related to the drug-like behavior of the studied peptides, their predicted biochemical targets and the values associated with Pharmacokinetics and ADMET properties (Daina et al., 2017; Pires et al., 2015; Daina et al., 2019) through standard Chemoinformatics procedures (Begam and Kumar, 2012; González-Medina et al., 2017). By considering this integrative strategy, called Conceptual DFT-based Computational Peptidology as a branch of Computational Chemistry dedicated to the study of peptides and cyclopeptides, the current research represents an extension of our recent studies on the properties of some families of therapeutic peptides of marine origin (Frau et al., 2018; Flores-Holguín et al., 2019a; Frau et al., 2019; Flores-Holguín et al., 2019c; Flores-Holguín et al., 2020; Flores-Holguín et al. 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021).

2 MATERIALS AND METHODS

2.1 In Silico Pharmacokinetics Analysis and Absorption, Distribution, Metabolism, Excretion and Toxicity Study

The starting molecular structures of the two plant cyclopeptides to be studied, shown in **Figure 1**, lwere obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov), which is an open chemistry database.

As a first step, the SMILES (Simplified Molecular Input Line Entry Specification) notation of every studied compound, which was obtained by accessing ChemDoodle 11.3.0 software, was fed into the online program Chemicalize, a software developed by ChemAxon (http://www.chemaxon.com), which was used for naming, molecular finger prints, structure generation and the prediction of several properties related to Chemoinformatics (http://chemicalize.com/) (accessed March 2021).

The similarity searches in the chemical space of compounds with molecular structures similar to those that are being studied was accomplished using the online available Molinspiration software from Molinspiration Cheminoformatics (https://www.molinspiration.com/) (accessed, March 2021) which was used for the prediction of the bioactivity scores for different drug targets.

A Webtool named SwissTargetPrediction for efficient prediction of protein targets of small molecules has been considered for the determination of the potential bioactivity of the two terrestrial plant cyclopeptides considered in this study (Daina et al., 2019). The associated website allows the estimation of the most probable macromolecular targets of a small molecule, assumed as bioactive. During the process of development of a new therapeutic drug, it is of the most importance to adcquire a knowledge of the fate of the pharmacokinetics, that is, the fate of a compound in the organism. This is usually performed by through individual indices that are called Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) parameters. These parameters are generally obtained using computer models as an alternative to the experimental procedures for their determination. In this research, some ADME parameters were estimated with the aid of Chemicalize and the online available SwissADME software (Daina et al., 2017).Additional information about the Pharmacokinetics parameters and the ADMET properties were obtained by resorting to pkCSM (Pires et al., 2015), a software for the prediction of small-molecule pharmacokinetic properties

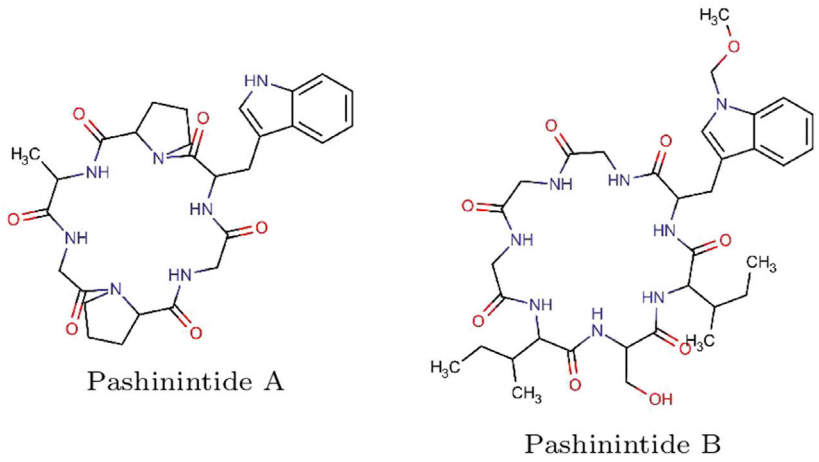


FIGURE 1 | Graphical sketches of the molecular structures of the Pashinintides A and B plant cyclopeptides.

using SMILES (https://biosig.unimelb.edu.au/pkcsm/) (accessed, March 2021).

2.2 Density Functional Theory Calculations

The goodness of a given density functional can be determined through a comparison of the results that it renders with the experimental values or with the results that can be obtained by means of high-level calculations. However, the lack of experimental results for the molecular systems under study or the large size of the molecules that made computationally impractical the use of some accurate methodologies. Kohn-Sham (KS) methodology includes the determination of the molecular energy, the electronic density and the orbital energies of a given system, related to the frontier orbitals including the Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) (Young, 2001; Lewars, 2003; Cramer 2004; Jensen, 2007). This methodology is convenient when thinking of quantitative qualities related with Conceptual DFT descriptors (Parr and Yang, 1989; Chermette, 1999; Geerlings et al., 2003; Toro-Labbé, 2007; Chattaraj, 2009, Geerlings et al., 2020; Chakraborty and Chattaraj, 2021). Range-separated (RS) exchange-correlation density functionals are of extraordinary concern in Kohn-Sham DFT calculations (Iikura et al., 2001; Yanai et al., 2004; Heyd and Scuseria, 2004; Chai and Head-Gordon, 2008; Stein et al., 2009a; Stein et al., 2009b; Stein et al., 2010; Karolewski et al., 2011; Kuritz et al., 2011; Ansbacher et al., 2012; Kronik et al., 2012; Stein et al., 2012). A methodology called KID (Koopmans in DFT) has been established by our research group (Frau et al., 2018; Flores-Holguín et al., 2019a; Frau et al., 2019; Flores-Holguín et al., 2019c; Flores-Holguín et al., 2020; Flores-Holguín et al., 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021), t for the validation of a given density functional in terms of its internal coherence. Several descriptors associated with the results of the HOMO and LUMO calculations are related to the results obtained from the estimation of the vertical I and A following the ΔSCF procedure, where SCF refers to the Self-Consistent Field technique. It has been demonstrated that there is a relationship between the KID descriptors and the Koopmans' theorem or the Ionization Energy theorem, which is its equivalent within the Generalized Kohn-Sham (GKS) version of DFT, by connecting ϵ_H to -I, ϵ_L to -A, and their actions by defining the HOMO – LUMO gap $J_I = |\epsilon_H + E_{gs}(N-1) - E_{gs}(N)|$, $J_A = |\epsilon_L + E_{gs}(N) - E_{gs}(N+1)|$, and $J_{HL} = \sqrt{J_I{}^2 + J_A{}^2}$. It should be noticed that the J_A descriptor consists of an approximation which is only valid if the HOMO of the radical anion (the SOMO) resembles the LUMO of the neutral system. For this reason, another descriptor ΔSL has been designed by our research group (Frau et al., 2018; Flores-Holguín et al., 2019a; Frau et al., 2019; Flores-Holguín et al., 2019c; Flores-Holguín et al., 2020; Flores-Holguín et al. 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021), to help in the verification of the accuracy of the approximation.

Taking into account the KID methodology considered in the previous research being integrated into the finite difference approximation (Frau et al., 2018; Flores-Holguín et al., 2019a; Frau et al., 2019; Flores-Holguín et al., 2019c; Flores-Holguín et al., 2020; Flores-Holguín et al., 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021), the following definitions can be used for the global descriptors that help in the understanding of the chemical reactivity of the molecular systems (Parr and Yang, 1989; Chermette, 1999; Geerlings et al., 2003; Gázquez et al., 2007; Chattaraj et al., 2009; Chakraborty and Chattaraj, 2021): Electronegativity as $\chi = -\frac{1}{2}(I + A) \approx \frac{1}{2}(\epsilon_L + \epsilon_H)$, Global Hardness as $\eta = (I - A) \approx (\epsilon_L - \epsilon_H)$, Electrophilicity as $\omega = \mu^2/2\eta = (I + A)^2/4(I - A) \approx (\epsilon_L + \epsilon_H)^2/4(\epsilon_L - \epsilon_H)$, Electrodonating Power as $\omega^- = (3I + A)^2/16(I - A) \approx (3\epsilon_H + \epsilon_L)^2/16\eta$, Electroaccepting Power as $\omega^+ = (I + 3A)^2/16(I - A) \approx (\epsilon_H + 3\epsilon_L)^2/16\eta$ and Net Electrophilicity as $\Delta\omega^{\pm} = \omega^+ - (-\omega^-) = \omega^+ + \omega^-$, being ϵ_H and ϵ_L the HOMO and LUMO energies associated with each of the peptides considered in this work. It is worth to mention that for

TABLE 1 | Names, identifiers, molecular fingerprints and basic properties of the studied molecular systems.

Property	Value
Common name	Pashinintide A
PubChem CID	122386973
Molar mass	565.631 g/mol
Exact mass	565.264881875 Da
Formula	$C_{28}H_{35}N_7O_6$
Composition	C (59.46%), H (6.24%), N (17.33%), O (16.97%)
IUPAC name	3-[(1H-indol-3-yl)methyl]-18-methyl-1,4,7,13,16,19-hexaazatricyclo[19.3.0.0^{9},^{13}] tetracosane-2,5,8,14,17,20-hexone
Traditional name	3-(1H-indol-3-ylmethyl)-18-methyl-1,4,7,13,16,19-hexaazatricyclo[19.3.0.0^{9},^{13}] tetracosane-2,5,8,14,17,20-hexone
SMILES	CC1NC(=O)C2CCCN2C(=O)C(CC2=CNC3=CC=CC=C23)NC(=O)CNC(=O) C2CCCN2C(=O)CNC1=O
InChI	InChI=1/C28H35N7O6/c1-16-25(38)31-15-24(37)34-10-4-8-21(34)26(39) 30-14-23(36)33-20(12-17-13-29-19-7-3-2-6-18(17)19)28(41)35-11-5-9-22 (35)27(40)32-16/h2-3,6-7,13,16,20-22,29H,4-5,8-12,14-15H2,1H3,(H,30,39) (H,31,38)(H,32,40)(H,33,36)
InChIKey	MKXJIZUYLVDQCC-UHFFFAOYNA-N
IUPAC condensed	cyclo[Ala-Gly-Pro-Gly-Trp-Pro]
Common name	Pashinintide B
PubChem CID	122386974
Molar mass	714.821 g/mol
Exact mass	714.370075222 Da
Formula	$C_{34}H_{50}N_8O_9$
Composition	C (57.13%), H (7.05%), N (15.68%), O (20.14%)
IUPAC name	6,12-bis(butan-2-yl)-9-(hydroxymethyl)-3-[1-(methoxymethyl)-1H-indol-3-yl] methyl-1,4,7,10,13,16,19-heptaazacyclohenicosane-2,5,8,11,14,17,20-heptone
Traditional name	9-(hydroxymethyl)-3-[1-(methoxymethyl)indol-3-yl]methyl-6,12-bis(sec-butyl)
Traditional name	-1,4,7,10,13,16,19-heptaazacyclohenicosane-2,5,8,11,14,17,20-heptone
SMILES	CC1NC(=O)C2CCCN2C(=O)C(CC2=CNC3=CC=CC=C23)NC(=O)CNC(=O) C2CCCN2C(=O)CNC1=O
InChI	InChI=1/C28H35N7O6/c1-16-25(38)31-15-24(37)34-10-4-8-21(34)26(39)30-14 23-23(36)33-20(12-17-13-29-19-7-3-2-6-18(17)19)28(41)35-11-5-9-22(35)27(40) 32-16/h2-3,6-7,13,16,20-22,29H,4-5,8-12,14-15H2,1H3,(H,30,39)(H,31,38) (H,32,40)(H,33,36)
InChIKey	MKXJIZUYLVDQCC-UHFFFAOYNA-N
IUPAC condensed	cyclo[Gly-Gly-Gly-xille-Ser-xille-Trp(MeOMe)]

the global indices the chemical power is directly related with the electronic density as well as the corresponding Hohenberg-Kohn functional (Putz, 2011).

As a complement of these global reactivity descriptors that arise from Conceptual DFT (Parr and Yang, 1989; Chermette, 1999; Geerlings et al., 2003; Gázquez et al., 2007; Chattaraj et al., 2009; Chakraborty and Chattaraj, 2021), Domingo and his collaborators (Domingo et al., 2008; Jaramillo et al., 2008; Domingo and Sáez, 2009; Domingo and Perez, 2011; Domingo et al., 2016) have proposed a Nucleophilicity index N through the consideration of the HOMO energy obtained through the KS scheme with an arbitrary shift of the origin taking the molecule of tetracyanoethylene (TCE) as a reference.

The determination of the conformers of the molecules considered in the current study was performed by resorting to MarvinView 17.15 available from ChemAxon (http://www.chemaxon.com) by doing Molecular Mechanics calculations through the overall MMFF94 force field (Halgren, 1996a, Halgren, 1996b, Halgren, 1999; Halgren and Nachbar, 1996; Halgren, 1996c). This was followed by a geometry optimization and frequency calculation by means of the Density Functional Tight Binding (DFTBA) methodology (Frisch et al., 2016). This last step was required for the verification of the absence of imaginary frequencies as a check for the stability of the optimized structures as being a minimum in the energy landscape. The electronic properties and the chemical reactivity descriptors of the studied molecules involved the use of MN12SX/Def2TZVP/H2O model chemistry (Weigend and Ahlrichs, 2005; Weigend, 2006; Peverati and Truhlar, 2012) on the optimized molecular structures due to is ability in the verification of the "Koopmans in DFT" (KID) protocol (Frau and Glossman-Mitnik, 2018a; Frau and Glossman-Mitnik, 2018b; Frau and Glossman-Mitnik, 2018c; Frau and Glossman-Mitnik, 2018d; Frau and Glossman-Mitnik, 2018e; Frau and Glossman-Mitnik, 2018f; Flores-Holguín et al., 2019a; Flores-Holguín et al., 2019b; Flores-Holguín et al., 2019d; Frau et al., 2019; Flores-Holguín et al., 2019c, Flores-Holguín et al., 2020; Flores-Holguín et al., 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021) using Gaussian 16 (Frisch et al., 2016) and the SMD model for the simulation of the solvent (Marenich et al., 2009). This model chemistry considers the MN12SX screened-exchange density functional (Peverati and Truhlar, 2012) together with the Def2TZVP basis set (Weigend and Ahlrichs, 2005; Weigend, 2006) and in all cases the charge of the

TABLE 2 | Geometrical and structural properties of the studied molecular systems.

Property	Pashinintide A	Pashinintide B
Atom count	76	101
Non-hydrogen atom count	41	51
Asymmetric atom count	4	6
Rotatable atom count	2	9
Ring count	5	3
Aromatic ring count	2	2
Hetero ring count	4	2
FSP3	0.50	0.56
Hydrogen bond donor count	5	8
Hydrogen bond acceptor count	6	9
Formal charge	0	0
Van der Waals volume (Å^3)	498.20	654.63
Van der Waals surface area (Å^2)	797.07	1,071.58
Solvent accesible surface area (Å^2)	688.91	729.09
Topological polar surface area (Å^2)	172.81	238.09
Polarizability (Å^3)	57.51	72.58
Molar refractivity (cm^3/mol)	145.88	182.79

molecules is equal to zero while the radical anion and cation have been considered in the doublet spin state.

3 RESULTS AND DISCUSSION

3.1 Physicochemical Properties, Bioactivity Scores and Biological Targets

The names, identifiers, molecular fingerprints and basic properties of the two Pashinintides A and B plant cyclopeptides are presented in **Table 1**, while their geometrical and structural properties are displayed in **Table 2**.

This information could it be of interest for future SAR studies based on these and other peptides as well as for potential derivatives designed for therapeutical purposes using Peptidomimetics.

TABLE 3 | Bioactivity scores for the Pashinintides A and B.

Property	Pashinintide A	Pashinintide B
GPCR Ligand	0.40	−0.54
Ion channel modulator	−0.20	−1.64
Nuclear receptor ligand	−0.03	−1.15
Kinase inhibitor	−0.20	−1.42
Protease inhibitor	0.53	−0.16
Enzyme inhibitor	0.04	−0.92

The effect of the geometrical and structural properties on the bioavailability of the Pashinintides A and B presented in **Table 2** can be better visualized considering the Bioavailability Radars displayed in **Figure 2**.

It can be appreciated that the two more important properties that could prevent the use of the Pashinintides A and B as therapeutic drugs are their molecular size and their polar character. Although these cyclic peptides violate some of the limits traditionally considered to be important for oral bioavailability of drug-like small molecules, it can be expected that the reduced flexibility could ease oral absorption (Nielsen et al., 2017). However, it must be remarked that these ideal values are based on the Lipinski's Rule of Five (Lipinski et al., 2001), which is not always applicable to peptides (Zhang and Wilkinson, 2007; Doak et al., 2014; Santos et al., 2016; Nielsen et al., 2017; Sable et al., 2017).

The Bioactivity Scores for the Pashinintides A and B are shown in **Table 3**.

It can be seen from the results on **Table 3** that while the bioactivity of Pashinintide B towards the different targets is considered to be low, for Pashinintide A, its interactions as a GPCR Ligand and a Protease Inhibitors could be of importance for its consideration as a potential therapeutic drug. The same conclusion can be extracted by checking visually the predicted biological targets for these plant cyclopeptides shown in **Figure 3**.

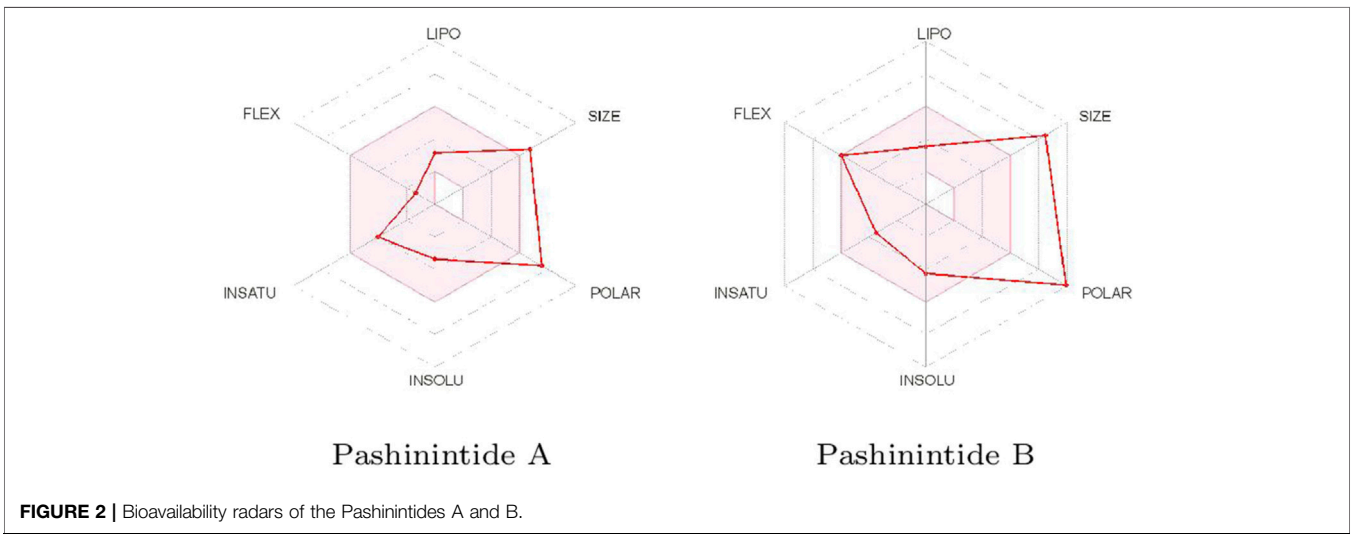


FIGURE 2 | Bioavailability radars of the Pashinintides A and B.

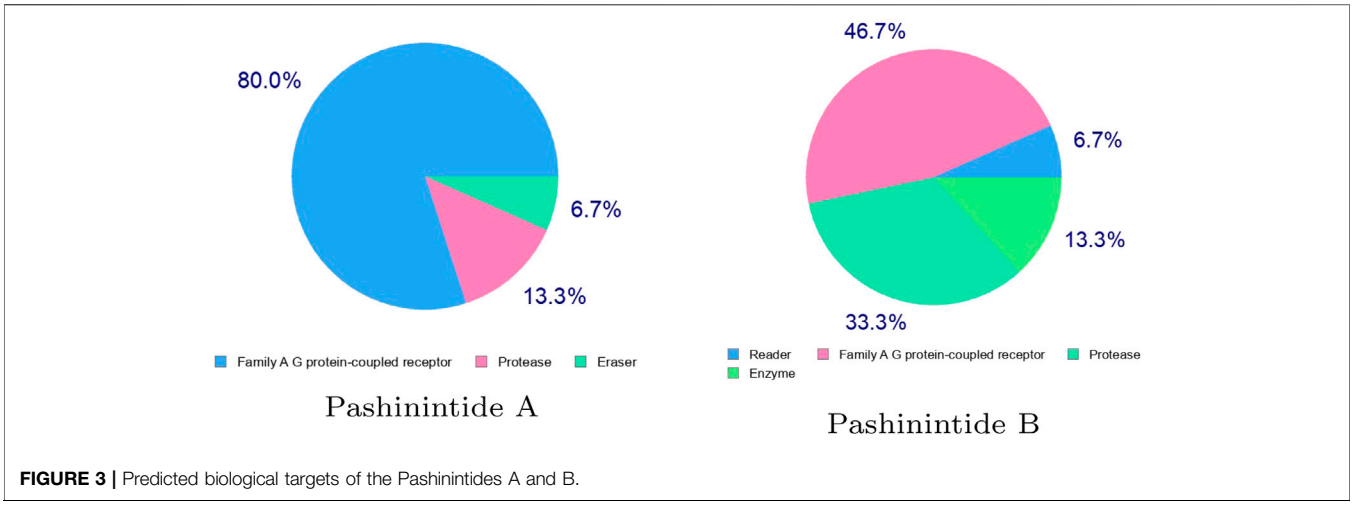


FIGURE 3 | Predicted biological targets of the Pashinintides A and B.

TABLE 4 | Absorption properties of the Pashinintides A and B.

Property	Pashinintide A	Pashinintide B
Water solubility	−3.197	−3.377
Caco-2 permeability	0.814	0.803
Intestinal absorption	52.552	31.004
Skin permeability	−2.736	−2.735
P-glycoprotein substrate	Yes	Yes
P-glycoprotein I inhibitor	No	Yes
P-glycoprotein II inhibitor	No	No

TABLE 5 | Distribution properties of the Pashinintides A and B.

Property	Pashinintide A	Pashinintide B
VD	−0.055	−0.769
Fraction unbound	0.394	0.364
BBB permeability	−0.299	−0.789
CNS permeability	−3.921	−4.459

3.2 Absorption, Distribution, Metabolism, Excretion and Toxicity Study

An ADMET study is the assessment of pharmacokinetics of a drug which stands for Absorption, Distribution, Metabolism, Excretion and Toxicity. The prediction of the fate of a drug and the effects caused by a drug inside the body, such as how much drug is absorbed if administered orally and how much is absorbed in the gastrointestinal tract, is an indispensable part of drug discovery. In a similar way, if the absorption is poor, its distribution and metabolism would be affected, which can lead to causing neurotoxicity and nephrotoxicity. Ultimately, the study is to understand the disposition of a drug molecule within an organism. Thus, ADMET study is one of the most essential parts of computational drug design.

3.2.1 Absorption

A compound can reach a tissue, if it is taken into the bloodstream. Usually, a drug is administered often through mucous surfaces such as the digestive tract, i.e., intestinal absorption before it is taken up by the target cells. Factors like poor compound solubility, intestinal transit time, gastric emptying time, inability to permeate the intestinal wall and chemical instability in the stomach are responsible for reducing the extent of drug absorption after oral administration. Critically, absorption determines the bioavailability of a compound. Drugs with poor absorption are less desirable for oral administration, such as by inhalation or intravenously (Pires et al., 2015; Jujjavarapu et al., 2019).

The computed absorption properties of the Pashinintides A and B are presented in **Table 4**.

The water solubility of a compound (logS) reflects the solubility of the molecule in water at 25°C. The predicted water solubility of a compound is given as the logarithm of the molar concentration (log mol/L) being their values very similar for both cyclopeptides. A compound is considered lo have a high Caco-2 permeability has a Papp $>8 \times 10^8$ cm/s. Thus, high Caco-2 permeability would translate in predicted values >0.90, presenting the Pashinintides A and B values which are a bit lower than the ideal one. The Intestine is normally the primary site for absorption of a drug from an orally administered solution. A molecule with an Intestinal Absorption of less than 30% is considered lo be poorly absorbed. From **Table 4**, both plant cyclopeptides will be highly absorbed. The P-glycoprotein has the function of a biological barrier by extruding toxins out of cells. The model predicts whether a given compound is likely to be a substrate of P-glycoprotein or not. The prediction is in the positive direction in both cases. Thus, the study predicts that both cyclopeptides will not act as P-glycoprotein II inhibitors, but Pashinintide A will not be a P-glycoprotein I inhibitor while Pashinintide B is likely to act in that way. Also, it can be predicted whether a given compound is likely to be skin permeable. A compound is considered to have a relatively low skin permeability if it has a log Kp > −2.5. It means that both cyclopeptides could be of interest for the development of transdermal drug delivery (Pires et al. (2015)).

TABLE 6 | Metabolism properties of the Pashinintides A and B.

Property	Pashinintide A	Pashinintide B
CYP2D6 substrate	No	No
CYP3A4 substrate	No	No
CYP1A2 inhibitor	No	No
CYP2C19 inhibitor	No	No
CYP2C9 inhibitor	No	No
CYP2D6 inhibitor	No	No
CYP3A4 inhibitor	No	No

TABLE 7 | Excretion properties of the Pashinintides A and B.

Property	Pashinintide A	Pashinintide B
Total clearance	0.495	0.856
Renal OCT2 substrate	No	No

3.2.2 Distribution

The computed distribution properties of the Pashinintides A and B are presented in **Table 5**.

VD is the theoretical volume required by a drug to be uniformly distributed in blood. The higher the VD is, the more of a drug is distributed in tissue rather than plasma. From **Table 5** and the usual standards, it can be said that VD for Pashinintide A is low and it is high for Pashinintide B. The Fraction Unbound parameter predicts the fraction that will be unbound in plasma resulting in the values shown in **Table 5**. The knowledge of the ability of a drug to cross into the brain is an important parameter that may help to reduce side effects and toxicities. A logBBB (for Blood-Brain Barrier) >−0.3 for a given drug is considered to easily cross the BBB while molecules with logBBB >−1 are poorly distributed to the brain, being predicted that both cyclopeptides have a BBB Permeability of the first case. Another measurement is the blood-brain permeability-surface area product or CNS Permeability where compounds with a logPS >−2 will be able to enter the Central Nervous System (CNS), while those with logPS <−3 will be unable to penetrate the CNS. For the current study, both cyclopeptides are predicted to do not penetrate the CNS (Pires et al., 2015).

3.2.3 Metabolism

The computed metabolism properties of the Pashinintides A and B are presented in **Table 6**.

Cytochrome P450 is an important detoxification enzyme in the body. Many drugs are deactivated by the cytochrome P450 isoforms while some can be activated by it. As can be seen from **Table 6**, both cyclopeptides are predicted as not being P450 inhibitors for any isoform. It is al important to know if a given compound is likely to be a cytochrome P450 substrate. The predictions indicate that this will be not the case for any of the cyclopeptides (Pires et al., 2015).

3.2.4 Excretion

The computed excretion properties of the Pashinintides A and B are presented in **Table 7**.

TABLE 8 | Toxicity properties of the Pashinintides A and B.

Property	Pashinintide A	Pashinintide B
AMES toxicity	No	No
MRTD	0.136	0.590
hERG I inhibitor	No	No
hERG II inhibitor	No	Yes
ORAT	2.914	2.901
ORCT	3.043	4.044
Hepatotoxicity	Yes	Yes
Skin sensitisation	No	No
T. Pyriformis toxicity	0.285	0.285

Drug clearance occurs as a combination of hepatic clearance and renal clearance (excretion via the kidneys) which is related to bioavailability. The predicted Total Clearance of the Pashinintides A and B are given in log(ml/min/kg) being the value for the former about 55% of the later. OCT2 is a renal uptake transporter that plays an important role in disposition and renal clearance of drugs. In this case, it is predicted that neither of the cyclopeptides will behave as OCT2 substrates (Pires et al., 2015).

3.2.5 Toxicity

The computed excretion properties of the Pashinintides A and B are presented in **Table 8**.

AMES Toxicity s a widely employed methodology considered to check the mutagenic potential of a given drug using bacteria, thus indicating that when the results is positive, the studied compound will be mutagenic and could behave as a carcinogen. From **Table 8**, the predictions are negative for both cyclopeptides under study. The maximum recommended tolerated dose (MRTD) provides an estimate of the toxic dose threshold of chemicals in humans. A low value for Pashinintide A and high value for Pashinintide B are found from the results in **Table 8**. Also, the predictions indicate that both cyclopeptides are unlikely to be hERG I inhibitors, but for the case of hERG II, the behavior will be different: Pashinintide A will not be a hERGII inhibitor while Pashinintide B will. The lethal dosage values (LD50) are a standard measurement of acute toxicity and is defined as the amount of a compound that causes the death of 50% of a group of test animals and are measured through the ORAT and ORCT indices where the predicted values are given in mol/kg.Drug-induced liver injury is a major safety concern for drug development. Hepatoxicity is associated with disrupted normal function of the liver and the predicted values for both cyclopeptides are positive. On the other hand, the predicted values for Skin Sensitisation are negative. T. Pyriformis is a protozoa bacteria whose toxicity is often used as a toxicity test. The predicted values for this parameter are the same for both cyclopeptides (Pires et al., 2015).

3.3 Conceptual Density Functional Theory Studies

The optimized molecular structures of the Pashinintides A and B are displayed in **Figure 4**.

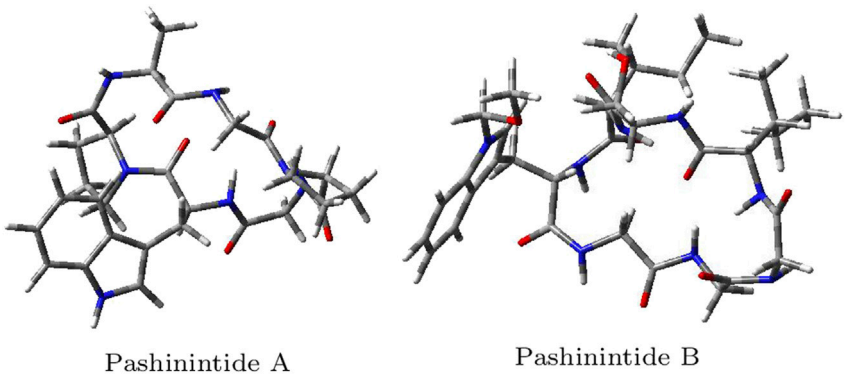


FIGURE 4 | Optimized molecular structures of the Pashinintides A and B.

TABLE 9 | HOMO, LUMO and SOMO orbital energies, HOMO-LUMO gap and the KID descriptors (all in eV) tested in the verification of the Koopmans-like behavior of the MN12SX/Def2TZVP/H2O model chemistry for the Pashinintides A and B.

Molecule	HOMO	LUMO	SOMO	H-L gap	J_I	J_A	J_{HL}	ΔSL
Pashinintide A	−5.5424	−0.9796	−1.1113	4.5628	0.018	0.054	0.057	0.132
Pashinintide B	−5.6589	−1.2210	−1.3641	4.4379	0.028	0.070	0.076	0.143

TABLE 10 | Global reactivity descriptors (in eV) for the Pashinintides A and B.

Molecule	χ	η	ω	S	N	ω^-	ω^+	$\Delta\omega^{\pm}$
Pashinintide A	3.2610	4.5628	1.1653	0.2192	3.2501	4.2463	0.9853	5.2316
Pashinintide B	3.4399	4.4379	1.3332	0.2253	3.1336	4.6637	1.2238	5.8875

Although the Koopmans-complaint behavior of the MN12SX density functional has been proved previously for the case of marine peptides (Frau et al., 2018; Flores-Holguín et al., 2019a; Frau et al., 2019; Flores-Holguín et al., 2019c; Flores-Holguín et al., 2020; Flores-Holguín et al., 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021), we are now performing a further validation for the plant cyclopeptides considered in the present study. This determination has been done by resorting to the in-house developed CDFT software tool and the resulting values are shown in **Table 9**.

As can be seen from the values presented in **Table 9**, the KID descriptors are all very close to zero meaning that the chosen MN12SX density functional displays a Koopmans-complaint behavior. This in agreement with our previous studies on peptides (Frau et al., 2018; Flores-Holguín et al., 2019a; Frau et al., 2019; Flores-Holguín et al., 2019c; Flores-Holguín et al., 2020; Flores-Holguín et al., 2020a; Flores-Holguín et al., 2020b; Flores-Holguín et al., 2021), thus justifying the adequacy of the MN12SX/Def2TZVP/H2O model chemistry for the purpose of this research.

The results for the global reactivity indices were estimated by making use of the mentioned CDFT tool and are presented in **Table 10**.

As the global hardness η can be regarded as a direct measure of the deformation of the electron density and of the chemical reactivity being related to the HOMO-LUMO gap, it can be seen that Pashininitide A will be slightly more reactive than the other cyclopeptide. The electrodonating ability ω^- is more important that its electroaccepting power ω^+ for both cyclopeptides because of their molecular structures. However, after a comparison of the values of ω^- and ω^+ for each molecule, it can be concluded that there are not important differences between them. The electrophilicity ω index encompasses the balance between the tendency to acquire an extra amount of electron density by an electrophile and the resistance of a molecule to exchange electron density with the environment Domingo et al. (2016). By studying the electrophilicities of a series of reagents involved in Diels-Alder reactions (Domingo et al., 2002a; Domingo and Sáez, 2009; Pérez et al., 2003), an electrophilicity ω scale for the classification of organic molecules as strong, moderate or marginal electrophiles was proposed being $\omega > 1.5$ eV for the first case, $0.8 < \omega < 1.5$ eV for the second case and $\omega < 0.8$ eV for the last case (Domingo et al., 2002a; Domingo and Sáez, 2009; Pérez et al., 2003). By inspection of **Table 10**, it can be said that both peptides considered in this study may be regarded as moderate electrophiles. Notwithstanding, the overall chemical reactivity is about the same for both cyclopeptides. This information could be of interest for future studies on the potential therapeutic ability of these compounds.

Besides global reactivity descriptors, their local counterparts have been developed to get an idea of the differences in chemical reactivity between the atoms within the molecule. Among these

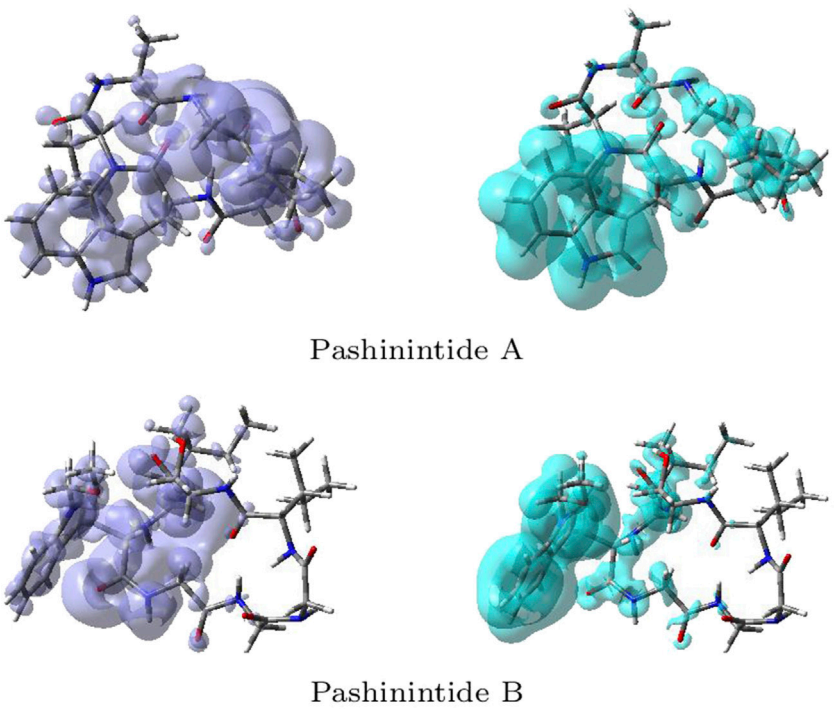


FIGURE 5 | Graphical representations of the dual descriptor DD of the Pashinintides A and B. Left: DD > 0, right: DD < 0.

local reactivity descriptors are the Fukui functions (Parr and Yang, 1989; Chermette, 1999; Geerlings et al., 2003) and the Dual Descriptor (Toro-Labbé, 2007; Morell et al., 2005, Morell et al., 2006; Martínez-Araya, 2012a; Martínez-Araya, 2012b; Martínez-Araya, 2015), which have been defined as: Nucleophilic Fukui Function (NFF) = $f^{+}(\mathbf{r}) = \rho_{N+1}(\mathbf{r}) - \rho_{N}(\mathbf{r})$, Electrophilic Fukui Function (EFF) = $f^{-}(\mathbf{r}) = \rho_{N}(\mathbf{r}) - \rho_{N-1}(\mathbf{r})$, and Dual Descriptor (DD) = $\Delta f(\mathbf{r}) = (\partial f(\mathbf{r})/\partial N)_{v(\mathbf{r})}$, relating the electronic densities of the neutral, positive and negative species.

The NFF, $f^{+}(\mathbf{r})$, is associated with the sites within a molecular system which are prone to nucleophilic attacks while the EFF, $f^{+}(\mathbf{r})$, describes those sites that are more susceptible to electrophilic attacks. Although the NFF and the EFF have been used successfully for the identification of reactive sites, the Dual Descriptor $\Delta f(\mathbf{r})$ or DD, can describe unambiguously nucleophilic and electrophilic sites within a molecule (Martínez-Araya (2015)). A graphical representation of the DD for the Pashinintides A and B cyclopeptides is displayed in **Figure 5** showing the zones where DD > 0 and DD < 0.

Although these graphical representations allowed to distinguish the regions within the molecules where the Dual Descriptor will be greater or smaller than zero, it can appreciated that there is some overlap between them. Thus, for a better estimation of these reactivity areas it is worth to determine the values of the Condensed Dual Descriptor (Δf_k) (Morell et al., 2008; Frau and Glossman-Mitnik, 2018d) over the all the atoms (excluding H) in comparison with the condensed versions of the Electrophilicity, that is, the Condensed Electrophilicity (ω_k) (Domingo et al., 2002b), and of the Nucleophilicity, being the Condensed Nucleophilicity N_k (Pérez et al., 2009). The resulting values are displayed in **Tables 11** and **12** for the Pashinintides A and B, respectively.

Even if every atom within the peptides cannot be graphically individualized due to the large size of the molecules, it is clear from the results in **Tables 11** and **12** about which are the sites prone to electrophilic and nucleophilic attacks based on the agreement between the values for the Condensed Electrophilicity ω_k with the positive values of the Condensed Dual Descriptor Δf_k, for one side, and with the values of the Condensed Nucleophilicity N_k and the negative results for the Condensed Dual Descriptor Δf_k.

From **Table 11** it can be seen that for Pashinintide A the maximum values for the Condensed Electrophilicity ω_k (shown in bold) are located over the C (27) and O (4) which correlate well with the maximum positive results for the Condensed Dual Descriptor Δf_k over those atoms. The same situation is found for the case of the Condensed Nucleophilicity N_k, whose maximum values over the C (29) and C(35) correlate with the maximum negative values Condensed Dual Descriptor Δf_k localized on those atoms. The closeness between C(27) and C (29) explains the overlap between the two regions within the graphical representation of the Dual Descriptor.

For the case of Pashinintide B, it can be appreciated from **Table 12** that the maximum values are located over the O (1), C (24), O (5) and C (37) (in that order) correlating in agreement with the greatest positive results for the Condensed Dual Descriptor Δf_k, while for the Condensed Nucleophilicity N_k the order of reactivity will be C (27) > C (36) > C (40) > C (45) > N (13) being the same as for those derived from the Condensed Dual Descriptor Δf_k. As for the case of the other

TABLE 11 | Comparison of several reactivity descriptors: condensed electrophilicity ω_k, condensed nucleophilicity N_k and condensed dual descriptor Δf_k, over the atoms of Pashinintide A. H atoms are not shown.

Atom	ω_k	N_k	Δf_k
O (1)	0.0247	0.0307	0.0121
O (2)	0.0132	0.0127	0.0075
O (3)	0.0210	0.0028	0.0168
O (4)	**0.1613**	0.0031	**0.1339**
O (5)	0.0129	0.0204	0.0051
O (6)	0.0257	0.0050	0.0201
N (7)	0.0097	0.0045	0.0093
N (8)	0.0576	0.0002	0.0481
N (9)	0.0052	0.0039	0.0033
N (10)	0.0016	0.0129	–0.0023
N (11)	0.0103	0.0020	0.0080
N (12)	0.0146	0.0034	0.0112
N (13)	0.0018	**0.2820**	**–0.0773**
C (14)	0.0035	0.0058	0.0013
C (15)	0.0029	0.0056	0.0009
C (16)	0.0028	0.0099	–0.0005
C (17)	0.0035	0.0071	0.0009
C (18)	0.0191	0.0007	0.0157
C (19)	0.0101	0.0007	0.0082
C (20)	0.0123	0.0004	0.0102
C (21)	0.0233	0.0007	0.0193
C (22)	0.0312	0.0083	0.0237
C (23)	0.0059	0.0337	–0.0045
C (24)	0.0091	0.0046	0.0063
C (25)	0.0076	0.0624	–0.0111
C (26)	0.0094	0.0013	0.0075
C (27)	**0.1956**	0.0006	**0.1632**
C (28)	0.0046	0.0012	0.0036
C (29)	0.0003	**0.3993**	**–0.1114**
C (30)	0.0416	0.0022	0.0341
C (31)	0.0061	0.0052	0.0036
C (32)	0.0081	0.0128	0.0032
C (33)	0.0009	0.0893	–0.0242
C (34)	0.0180	0.0019	0.0145
C (35)	0.0033	**0.3996**	**–0.1089**
C (36)	0.0042	0.0016	0.0031
C (37)	0.0016	**0.1249**	**–0.0336**
C (38)	0.0027	**0.2805**	**–0.0761**
C (39)	0.0028	**0.2514**	**–0.0679**
C (40)	0.0026	**0.1608**	**–0.0428**
C (41)	0.0019	**0.2900**	**–0.0794**

TABLE 12 | Comparison of several reactivity descriptors: condensed electrophilicity ω_k, condensed nucleophilicity N_k and condensed dual descriptor Δf_k, over the atoms of Pashinintide B. H atoms are not shown.

Atom	ω_k	N_k	Δf_k
O (1)	**0.1768**	0.0238	**0.1213**
O (2)	0.0181	0.0091	0.0105
O (3)	0.0034	0.0003	0.0024
O (4)	0.0030	0.0032	0.0012
O (5)	**0.1346**	0.0291	**0.0891**
O (6)	0.0012	0.0430	–0.0115
O (7)	0.0013	0.0006	0.0008
O (8)	0.0150	0.0037	0.0098
O (9)	0.0016	0.0011	0.0015
N (10)	0.0072	0.0009	0.0049
N (11)	0.0539	0.0162	0.0344
N (12)	0.0020	0.0009	0.0012
N (13)	0.0057	**0.2153**	**–0.0580**
N (14)	0.0003	0.0003	0.0001
N (15)	0.0589	0.0121	0.0392
N (16)	0.0009	0.0008	0.0005
N (17)	0.0059	0.0018	0.0038
C (18)	0.0050	0.0014	0.0032
C (19)	0.0196	0.0014	0.0138
C (20)	0.0038	0.0011	0.0024
C (21)	0.0006	0.0003	0.0003
C (22)	0.0344	0.0361	0.0145
C (23)	0.0009	0.0003	0.0005
C (24)	**0.1587**	0.0090	**0.1123**
C (25)	0.0176	0.0647	–0.0059
C (26)	0.0068	0.0012	0.0046
C (27)	0.0072	**0.3991**	**–0.1100**
C (28)	0.0041	0.0016	0.0025
C (29)	0.0005	0.0001	0.0003
C (30)	0.0130	0.0098	0.0066
C (31)	0.0082	0.1020	–0.0235
C (32)	0.0043	0.0008	0.0029
C (33)	0.0037	0.0010	0.0024
C (34)	0.0005	0.0002	0.0003
C (35)	0.0053	**0.1143**	**–0.0292**
C (36)	0.0180	**0.3832**	**–0.0976**
C (37)	**0.1253**	0.0105	**0.0878**
C (38)	0.0026	0.0026	0.0011
C (39)	0.0004	0.0002	0.0002
C (40)	0.0128	**0.2814**	**–0.0720**
C (41)	0.0024	0.0502	–0.0128
C (42)	0.0096	**0.2328**	**–0.0603**
C (43)	0.0006	0.0001	0.0004
C (44)	0.0074	**0.1523**	**–0.0386**
C (45)	0.0114	**0.2742**	**–0.0709**
C (46)	0.0173	0.0022	0.0119
C (47)	0.0006	0.0003	0.0003
C (48)	0.0106	0.0017	0.0072
C (49)	0.0026	0.0009	0.0016
C (50)	0.0009	0.0003	0.0005
C (51)	0.0010	0.0223	–0.0057

peptide, the partial overlapping between the different reactive areas could be attributed to nearness between C (24) and C (27).

4 CONCLUSION

Two cyclic peptides, Pashinintides A and B, isolated from a terrestrial plant have been studied by resorting to some techniques of common use in the process of drug discovery and development through our proposed Computational Peptidology methodology showing that these kind of molecules can be regarded as potential therapeutic drugs.

With the further objective of analyzing their bioactivity, some useful parameters for future Structure Activity Relationships (SAR) research for the development of therapeutic drugs, their predicted biological targets, and the ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) parameters related to the bioavailability and pharmacokinetics of the two plant cyclopeptides under study were predicted and analyzed.

The chemical reactivities of the studied cyclopeptides have been exhaustively analyzed through the optimization of their structures using the DFTBA methodology and the estimation of their electronic properties making use of the MN12SX/Def2TZVP/H2O model chemistry already considered in previous research for the study of

peptides, thus verifying its usefulness for this kind of calculations and supplemented with the calculation the Conceptual DFT-derived global and local reactivity descriptors, allowing to identify the preferred reactivity atoms within the molecules.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

NF-H and JF: Research and data analysis; DG-M: Research, data analysis and writing of the manuscript.

ACKNOWLEDGMENTS

NFH and DGM are researchers of CIMAV and CONACYT and want to thank both institutions for partial support.

REFERENCES

Ansbacher, T., Srivastava, H. K., Stein, T., Baer, R., Merkx, M., and Shurki, A. (2012). Calculation of Transition Dipole Moment in Fluorescent Proteins-Towards Efficient Energy Transfer. *Phys. Chem. Chem. Phys.* 14, 4109–4117. doi:10.1039/c2cp23351g

Begam, B. F., and Kumar, J. S. (2012). A Study on Cheminformatics and its Applications on Modern Drug Discovery. *Proced. Eng.* 38, 1264–1275. doi:10.1016/j.proeng.2012.06.156

Cai, L., Zhao, M., Liu, S., Yin, T., Zhou, H., Dong, J., et al. (2014). Pashinintide A, the First Plant Cyclopeptide from Rosaceae, Included a Sucrose, Suggests a New Natural Receptor for Saccharide. *Tetrahedron Lett.* 55, 6231–6235. doi:10.1016/j.tetlet.2014.09.063

Chai, J.-D., and Head-Gordon, M. (2008). Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* 10, 6615–6620. doi:10.1039/b810189b

Chakraborty, D., and Chattaraj, P. K. (2021). Conceptual Density Functional Theory Based Electronic Structure Principles. *Chem. Sci.* 12, 6264–6279. doi:10.1039/d0sc07017c

Chattaraj, P. K., Chakraborty, A., and Giri, S. (2009). Net Electrophilicity. *J. Phys. Chem. A.* 113, 10068–10074. doi:10.1021/jp904674x

P. K. Chattaraj (Editors) (2009). *Chemical Reactivity Theory - A Density Functional View* (Boca Raton, FL: CRC Press. Taylor & Francis Group).

Chaudhari, R., and Chakraborti, D. (2021). "Bioactive Compounds from In-Vitro Culture of Swertia Chirayita (Roxb. Ex Flem.) Karsten: Identification and Quantification," in *Plant-based Functional Foods and Phytochemicals: From Traditional Knowledge to Present Innovation*. Editors M. R. Goyal, A. Nath, H. Suleria, and F. L. Palm Bay (USA Burlington, ON, Canada Boca Raton, FL, USA Abingdon, Oxon, UK: Apple Academic Press Inc. CRC Press), 201–239.

Chermette, H. (1999). Chemical Reactivity Indexes in Density Functional Theory. *J. Comput. Chem.* 20, 129–154. doi:10.1002/(sici)1096-987x(19990115)20:1<129::aid-jcc13>3.0.co;2-a

Cramer, C. (2004). *Essentials of Computational Chemistry - Theories and Models*. 2nd edn. Chichester, England: John Wiley & Sons.

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, 42717. doi:10.1038/srep42717

Daina, A., Michielin, O., and Zoete, V. (2019). SwissTargetPrediction: Updated Data and New Features for Efficient Prediction of Protein Targets of Small Molecules. *Nucleic Acids Res.* 47, W357–W364. doi:10.1093/nar/gkz382

Doak, B. C., Over, B., Giordanetto, F., and Kihlberg, J. (2014). Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* 21, 1115–1142. doi:10.1016/j.chembiol.2014.08.013

Domingo, L. R., and Pérez, P. (2011). The Nucleophilicity N Index in Organic Chemistry. *Org. Biomol. Chem.* 9, 7168–7175. doi:10.1039/c1ob05856h

Domingo, L. R., and Sáez, J. A. (2009). Understanding the Mechanism of Polar Diels-Alder Reactions. *Org. Biomol. Chem.* 7, 3576–3583. doi:10.1039/b909611f

Domingo, L. R., Aurell, M. J., Pérez, P., and Contreras, R. (2002a). Quantitative Characterization of the Global Electrophilicity Power of Common diene/Dienophile Pairs in Diels-Alder Reactions. *Tetrahedron* 58, 4417–4423. doi:10.1016/s0040-4020(02)00410-6

Domingo, L. R., Aurell, M. J., Pérez, P., and Contreras, R. (2002b). Quantitative Characterization of the Local Electrophilicity of Organic Molecules. Understanding the Regioselectivity on Diels–Alder Reactions. *J. Phys. Chem. A.* 106, 6871–6875. doi:10.1021/jp020715j

Domingo, L. R., Chamorro, E., and Pérez, P. (2008). Understanding the Reactivity of Captodative Ethylenes in Polar Cycloaddition Reactions. A Theoretical Study. *J. Org. Chem.* 73, 4615–4624. doi:10.1021/jo800572a

Domingo, L., Ríos-Gutiérrez, M., and Pérez, P. (2016). Applications of the Conceptual Density Functional Theory Indices to Organic Chemistry Reactivity. *Molecules* 21, 748. doi:10.3390/molecules21060748

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2019a). Chemical-Reactivity Properties, Drug Likeness, and Bioactivity Scores of Seragamides A-F Anticancer Marine Peptides: Conceptual Density Functional Theory Viewpoint. *Computation* 7, 52. doi:10.3390/computation7030052

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2019b). Computational Peptidology Assisted by Conceptual Density Functional Theory for the Study of Five New Antifungal Tripeptides. *ACS Omega* 4, 12555–12560. doi:10.1021/acsomega.9b01463

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2019c). Computational Prediction of Bioactivity Scores and Chemical Reactivity Properties of the Parasin I Therapeutic Peptide of Marine Origin Through the Calculation of Global and Local Conceptual DFT Descriptors. *Theor. Chem. Acc.* 138, 78. doi:10.1007/s00214-019-2469-3

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2019d). "Conceptual DFT as a Helpful Chemoinformatics Tool for the Study of the Clavanin Family of Antimicrobial Marine Peptides," in *Density Functional Theory Calculations* (Rijetia: IntechOpen), 1–11.

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2020). A Fast and Simple Evaluation of the Chemical Reactivity Properties of the Pristinamycin Family of Antimicrobial Peptides. *Chem. Phys. Lett.* 739, 137021. doi:10.1016/j.cplett.2019.137021

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2020a). Conceptual DFT-Based Computational Peptidology of Marine Natural Compounds: Discodermins A-H. *Molecules* 25, 4158. doi:10.3390/molecules25184158

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2020b). Virtual Screening of Marine Natural Compounds by Means of Chemoinformatics and CDFT-Based Computational Peptidology. *Mar. Drugs* 18, 478. doi:10.3390/md18090478

Flores-Holguín, N., Frau, J., and Glossman-Mitnik, D. (2021). "Conceptual DFT as a Helpful Chemoinformatics Tool for the Study of the Clavanin Family of Antimicrobial Marine Peptides," in *Density Functional Theory*. Editors S. R. De Lazaro, L. H. Da Silveira Lacerda, and R. A. Pontes Ribeiro (London, UK: IntechOpen), 57–67. chap. 3. doi:10.5772/intechopen.88657

Frau, J., and Glossman-Mitnik, D. (2018a). Blue M2: An Intermediate Melanoidin Studied via Conceptual DFT. *J. Mol. Model.* 24, 1–13. doi:10.1007/s00894-018-3673-0

Frau, J., and Glossman-Mitnik, D. (2018b). Chemical Reactivity Theory Applied to the Calculation of the Local Reactivity Descriptors of a Colored Maillard Reaction Product. *Chem. Sci. Int. J.* 22, 1–14. doi:10.9734/csji/2018/41452

Frau, J., and Glossman-Mitnik, D. (2018c). Computational Study of the Chemical Reactivity of the Blue-M1 Intermediate Melanoidin. *Comput. Theor. Chem.* 1134, 22–29. doi:10.1016/j.comptc.2018.04.018

Frau, J., and Glossman-Mitnik, D. (2018d). Local Molecular Reactivity of the Colored Dansylglycine in Water and Dioxane Studied through Conceptual DFT. *J. Chem.* 2018, 1–7. doi:10.1155/2018/3172412

Frau, J., and Glossman-Mitnik, D. (2018e). Conceptual DFT Study of the Local Chemical Reactivity of the Dilysyldipyrrolones A and B Intermediate Melanoidins. *Theor. Chem. Acc.* 137, 1210. doi:10.1007/s00214-018-2244-x

Frau, J., and Glossman-Mitnik, D. (2018f). Molecular Reactivity and Absorption Properties of Melanoidin Blue-G1 through Conceptual DFT. *Molecules* 23, 559. doi:10.3390/molecules23030559

Frau, J., Flores-Holguín, N., and Glossman-Mitnik, D. (2018). Chemical Reactivity Properties, pKa Values, AGEs Inhibitor Abilities and Bioactivity Scores of the Mirabamides A-H Peptides of Marine Origin Studied by Means of Conceptual DFT. *Mar. Drugs* 16, 302–319. doi:10.3390/md16090302

Frau, J., Flores-Holguín, N., and Glossman-Mitnik, D. (2019). Chemical Reactivity Theory and Empirical Bioactivity Scores as Computational Peptidology Alternative Tools for the Study of Two Anticancer Peptides of Marine Origin. *Molecules* 24, 1115. doi:10.3390/molecules24061115

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 16 Revision C.01*. Wallingford CT: Gaussian Inc. [Dataset].

Gang, D., Kim, D., and Park, H.-S. (2018). Cyclic Peptides: Promising Scaffolds for Biopharmaceuticals. *Genes* 9, 557. doi:10.3390/genes9110557

Gázquez, J. L., Cedillo, A., and Vela, A. (2007). Electrodonating and Electroaccepting Powers. *J. Phys. Chem. A.* 111, 1966–1970. doi:10.1021/jp065459f

Geerlings, P., De Proft, F., and Langenaeker, W. (2003). Conceptual Density Functional Theory. *Chem. Rev.* 103, 1793–1874. doi:10.1021/cr990029p

Geerlings, P., Chamorro, E., Chattaraj, P. K., De Proft, F., Gázquez, J. L., Liu, S., et al. (2020). Conceptual Density Functional Theory: Status, Prospects, Issues. *Theor. Chem. Acc.* 139, 36. doi:10.1007/s00214-020-2546-7

González-Medina, M., Naveja, J. J., Sánchez-Cruz, N., and Medina-Franco, J. L. (2017). Open Chemoinformatic Resources to Explore the Structure, Properties and Chemical Space of Molecules. *RSC Adv.* 7, 54153–54163. doi:10.1039/c7ra11831g

Guha, M., Das, N., and Savarimuthu, X. (2021). "Natural Phytobiactives: Let's Eat Smart!," in *Plant-based Functional Foods and Phytochemicals: From Traditional Knowledge to Present Innovation*. Editors M. R. Goyal, A. Nath, H. Suleria, and F. L. Palm Bay (FL, USA Burlington, ON, Canada Boca Raton, FL, USA Abingdon, Oxon, UK: Apple Academic Press Inc. CRC Press), 201–239.

Halgren, T. A., and Nachbar, R. B. (1996). Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94. *J. Comput. Chem.* 17, 587–615. doi:10.1002/(sici)1096-987x(199604)17:5/6<587::aid-jcc4>3.0.co;2-q

Halgren, T. A. (1996a). Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* 17, 490–519. doi:10.1002/(sici)1096-987x(199604)17:5/6<490::aid-jcc1>3.0.co;2-p

Halgren, T. A. (1996b). Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comput. Chem.* 17, 520–552. doi:10.1002/(sici)1096-987x(199604)17:5/6<520::aid-jcc2>3.0.co;2-w

Halgren, T. A. (1996c). Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data, and Empirical Rules. *J. Comput. Chem.* 17, 616–641. doi:10.1002/(sici)1096-987x(199604)17:5/6<616::aid-jcc5>3.0.co;2-x

Halgren, T. A. (1999). MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* 20, 720–729. doi:10.1002/(sici)1096-987x(199905)20:7<720::aid-jcc7>3.0.co;2-x

Heyd, J., and Scuseria, G. E. (2004). Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd-Scuseria-Ernzerhof Screened Coulomb Hybrid Functional. *J. Chem. Phys.* 121, 1187–1192. doi:10.1063/1.1760074

Iikura, H., Tsuneda, T., Yanai, T., and Hirao, K. (2001). A Long-Range Correction Scheme for Generalized-Gradient-Approximation Exchange Functionals. *J. Chem. Phys.* 115, 3540–3544. doi:10.1063/1.1383587

Jaramillo, P., Domingo, L. R., Chamorro, E., and Pérez, P. (2008). A Further Exploration of a Nucleophilicity Index Based on the Gas-Phase Ionization Potentials. *J. Mol. Struct. THEOCHEM* 865, 68–72. doi:10.1016/j.theochem.2008.06.022

Jensen, F. (2007). *Introduction to Computational Chemistry*. 2nd edn. Chichester, England: John Wiley & Sons.

Jujjavarapu, S. E., Dhagat, S., and Yadav, M. (2019). *Computer-Aided Design of Antimicrobial Lipopeptides as Prospective Drug Candidates*. Boca Raton: CRC Press LLC.

Karolewski, A., Stein, T., Baer, R., and Kümmel, S. (2011). Communication: Tailoring the Optical Gap in Light-Harvesting Molecules. *J. Chem. Phys.* 134, 151101–151105. doi:10.1063/1.3581788

Kim, S.-K. (2013). *Marine Proteins and Peptides - Biological Activities and Applications*. Chichester, UK: Wiley-Blackwell.

Kronik, L., Stein, T., Refaely-Abramson, S., and Baer, R. (2012). Excitation Gaps of Finite-Sized Systems from Optimally Tuned Range-Separated Hybrid Functionals. *J. Chem. Theor. Comput.* 8, 1515–1531. doi:10.1021/ct2009363

Kuritz, N., Stein, T., Baer, R., and Kronik, L. (2011). Charge-Transfer-Like $\pi \rightarrow \pi^*$ Excitations in Time-Dependent Density Functional Theory: A Conundrum and its Solution. *J. Chem. Theor. Comput.* 7, 2408–2415. doi:10.1021/ct2002804

Lewars, E. (2003). *Computational Chemistry - Introduction to the Theory and Applications of Molecular and Quantum Mechanics*. Dordrecht: Kluwer Academic Publishers.

Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings 1PII of Original Article: S0169-409X(96)00423-1. The Article Was Originally Published in Advanced Drug Delivery Reviews 23 (1997) 3-25. 1. *Adv. Drug Deliv. Rev.* 46, 3–26. doi:10.1016/s0169-409x(00)00129-0

Marenich, A. V., Cramer, C. J., and Truhlar, D. G. (2009). Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* 113, 6378–6396. doi:10.1021/jp810292n

Martínez-Araya, J. I. (2015). Why Is the Dual Descriptor a More Accurate Local Reactivity Descriptor Than Fukui Functions?. *J. Math. Chem.* 53, 451–465. doi:10.1007/s10910-014-0437-7

Martínez-Araya, J. I. (2012a). Explaining Reaction Mechanisms Using the Dual Descriptor: A Complementary Tool to the Molecular Electrostatic Potential. *J. Mol. Model.* 19, 2715–2722. doi:10.1007/s00894-012-1520-2

Martínez-Araya, J. I. (2012b). Revisiting Caffeate's Capabilities as a Complexation Agent to Silver Cation in Mining Processes by Means of the Dual Descriptor-A Conceptual DFT Approach. *J. Mol. Model.* 18, 4299–4307. doi:10.1007/s00894-012-1405-4

Morell, C., Grand, A., and Toro-Labbé, A. (2005). New Dual Descriptor for Chemical Reactivity. *J. Phys. Chem. A.* 109, 205–212. doi:10.1021/jp046577a

Morell, C., Grand, A., and Toro-Labbé, A. (2006). Theoretical Support for Using the Δf(r) Descriptor. *Chem. Phys. Lett.* 425, 342–346. doi:10.1016/j.cplett.2006.05.003

Morell, C., Hocquet, A., Grand, A., and Jamart-Grégoire, B. (2008). A Conceptual DFT Study of Hydrazino Peptides: Assessment of the Nucleophilicity of the Nitrogen Atoms by Means of the Dual Descriptor Δf(r). *J. Mol. Struct. THEOCHEM* 849, 46–51. doi:10.1016/j.theochem.2007.10.014

Nielsen, D. S., Shepherd, N. E., Xu, W., Lucke, A. J., Stoermer, M. J., and Fairlie, D. P. (2017). Orally Absorbed Cyclic Peptides. *Chem. Rev.* 117, 8094–8128. doi:10.1021/acs.chemrev.6b00838

Parr, R., and Yang, W. (1989). *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press.

Pérez, P., Domingo, L. R., José Aurell, M., and Contreras, R. (2003). Quantitative Characterization of the Global Electrophilicity Pattern of Some Reagents Involved in 1,3-Dipolar Cycloaddition Reactions. *Tetrahedron* 59, 3117–3125. doi:10.1016/s0040-4020(03)00374-0

Pérez, P., Domingo, L. R., Duque-Noreña, M., and Chamorro, E. (2009). A Condensed-To-Atom Nucleophilicity Index. An Application to the Director Effects on the Electrophilic Aromatic Substitutions. *J. Mol. Struct. THEOCHEM* 895, 86–91. doi:10.1016/j.theochem.2008.10.014

Peverati, R., and Truhlar, D. G. (2012). Screened-Exchange Density Functionals with Broad Accuracy for Chemistry and Solid-State Physics. *Phys. Chem. Chem. Phys.* 14, 16187–16191. doi:10.1039/c2cp42576a

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, 4066–4072. doi:10.1021/acs.jmedchem.5b00104

Putz, M. V. (2011). Chemical Action Concept and Principle. *Math. Comput. Chem.* 66, 35–63.

Sable, R., Parajuli, P., and Jois, S. (2017). Peptides, Peptidomimetics, and Polypeptides from Marine Sources: A Wealth of Natural Sources for Pharmaceutical Applications. *Mar. Drugs* 15, 124. doi:10.3390/md15040124

Sánchez, A., and Vázquez, A. (2017). Bioactive Peptides: A Review. *Food Qual. Saf.* 1, 29–46. doi:10.1093/fqs/fyx006

Santos, G. B., Ganesan, A., and Emery, F. S. (2016). Oral Administration of Peptide-Based Drugs: Beyond Lipinski's Rule. *ChemMedChem* 11, 2245–2251. doi:10.1002/cmdc.201600288

Stein, T., Kronik, L., and Baer, R. (2009a). Prediction of Charge-Transfer Excitations in Coumarin-Based Dyes Using a Range-Separated Functional Tuned from First Principles. *J. Chem. Phys.* 131, 244119. doi:10.1063/1.3269029

Stein, T., Kronik, L., and Baer, R. (2009b). Reliable Prediction of Charge Transfer Excitations in Molecular Complexes Using Time-Dependent Density Functional Theory. *J. Am. Chem. Soc.* 131, 2818–2820. doi:10.1021/ja8087482

Stein, T., Eisenberg, H., Kronik, L., and Baer, R. (2010). Fundamental Gaps in Finite Systems from Eigenvalues of a Generalized Kohn-Sham Method. *Phys. Rev. Lett.* 105, 266802–266804. doi:10.1103/physrevlett.105.266802

Stein, T., Autschbach, J., Govind, N., Kronik, L., and Baer, R. (2012). Curvature and Frontier Orbital Energies in Density Functional Theory. *J. Phys. Chem. Lett.* 3, 3740–3744. doi:10.1021/jz3015937

A. Toro-Labbé (Editors) (2007). *Theoretical Aspects of Chemical Reactivity* (Amsterdam: Elsevier Science).

Weigend, F., and Ahlrichs, R. (2005). Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* 7, 3297–3305. doi:10.1039/b508541a

Weigend, F. (2006). Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* 8, 1057–1065. doi:10.1039/b515623h

Yanai, T., Tew, D. P., and Handy, N. C. (2004). A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3lyp). *Chem. Phys. Lett.* 393, 51–57. doi:10.1016/j.cplett.2004.06.011

Young, D. (2001). *Computational Chemistry - A Practical Guide for Applying Techniques to Real-World Problems*. New York: John Wiley & Sons.

Zhang, M.-Q., and Wilkinson, B. (2007). Drug Discovery beyond the 'rule-Of-Five'. *Curr. Opin. Biotechnol.* 18, 478–488. doi:10.1016/j.copbio.2007.10.005

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.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 07 October 2021
doi: 10.3389/fchem.2021.709932

Inhibition of GSK_3β by Iridoid Glycosides of Snowberry (*Symphoricarpos albus*) Effective in the Treatment of Alzheimer's Disease Using Computational Drug Design Methods

Marzieh Eskandarzadeh[1], *Parastou Kordestani-Moghadam*[2], *Saeed Pourmand*[3], *Javad Khalili Fard*[4,5], *Bijan Almassian*[6] and *Sajjad Gharaghani*[7]*

[1]Research Committee of Faculty of Pharmacy, Lorestan University of Medical Science, Khorramabad, Iran, [2]Social Determinants of Health Research Center, Lorestan University of Medical Sciences, Khorramabad, Iran, [3]Department of Chemical Engineering, Faculty of Chemical and Petroleum Engineering, University of Tabriz, Tabriz, Iran, [4]Razi Herbal Medicines Research Center, Lorestan University of Medical Sciences, Khorramabad, Iran, [5]Department of Pharmacology and Toxicology, Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran, [6]CaroGen Corporation, Farmington, CT, United States, [7]Laboratory of Bioinformatics and Drug Design, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

OPEN ACCESS

Edited by:
Jorge Ignacio Martínez-Araya,
Andres Bello University, Chile

Reviewed by:
Imtaiyaz Hassan,
Jamia Millia Islamia, India
Madhu Sudhana Saddala,
Sri Venkateswara University, India

***Correspondence:**
Sajjad Gharaghani
S.gharaghani@ut.ac.ir

Specialty section:
This article was submitted to Theoretical and Computational Chemistry, a section of the journal Frontiers in Chemistry

Received: 14 May 2021
Accepted: 31 August 2021
Published: 07 October 2021

Citation:
Eskandarzadeh M, Kordestani-Moghadam P, Pourmand S, Khalili Fard J, Almassian B and Gharaghani S (2021) Inhibition of GSK_3β by Iridoid Glycosides of Snowberry (Symphoricarpos albus) Effective in the Treatment of Alzheimer's Disease Using Computational Drug Design Methods. Front. Chem. 9:709932. doi: 10.3389/fchem.2021.709932

The inhibition of glycogen synthase kinase-3β (GSK-3β) activity prevents tau hyperphosphorylation and binds it to the microtubule network. Therefore, a GSK-3β inhibitor may be a recommended drug for Alzheimer's treatment. In silico methods are currently considered as one of the fastest and most cost-effective available alternatives for drug/design discovery in the field of treatment. In this study, computational drug design was conducted to introduce compounds that play an effective role in inhibiting the GSK-3β enzyme by molecular docking and molecular dynamics simulation. The iridoid glycosides of the common snowberry (*Symphoricarpos albus*), including loganin, secologanin, and loganetin, are compounds that have an effect on improving memory and cognitive impairment and the results of which on Alzheimer's have been studied as well. In this study, in the molecular docking phase, loganin was considered a more potent inhibitor of this protein by establishing a hydrogen bond with the ATP-binding site of GSK-3β protein and the most negative binding energy to secologanin and loganetin. Moreover, by molecular dynamics simulation of these ligands and GSK-3β protein, all structures were found to be stable during the simulation. In addition, the protein structure represented no change and remained stable by binding ligands to GSK-3β protein. Furthermore, loganin and loganetin have higher binding free energy than secologanin; thus, these compounds could effectively bind to the active site of GSK-3β protein. Hence, loganin and loganetin as iridoid glycosides can be effective in Alzheimer's prevention and treatment, and thus, further *in vitro* and *in vivo* studies can focus on these iridoid glycosides as an alternative treatment.

Keywords: loganin, Alzheimer's disease, GSK_3 beta, docking, *Symphoricarpos*

INTRODUCTION

Alzheimer's disease (AD) is the most prevalent form of dementia (Lin, Jones et al., 2020). This health problem results in malignant neurological disorder along with cognitive, functional, and behavioral changes progressively and irreversibly. According to statistics, by 2050, this age-associated disease will affect 1 out of 85 people globally (Brookmeyer et al., 2007), posing a social and economic burden on the future community (Association 2019).

AD has a multifactorial etiology, with the most prominent ones associated with extracellular deposition of amyloid-β (Aβ) plaques, aggregation of tau protein, inflammation, oxidative stress, and declined levels of acetylcholine (De Strooper and Karran 2016; Hall et al., 2019). Nowadays, some evidence suggests that Aβ deposition is related to aging abnormalities. Also, tau (τ) protein is probably a higher target than Aβ because several therapeutic techniques concerning Aβ have not been highly effective in the disease treatment (Kametani and Hasegawa 2018). Tau protein is a member of the microtubule-related protein family that promotes microtubule stability and cytoskeleton formation (Kolarova et al., 2012). Intracellular tau protein aggregation in nerve fiber nodes is a neuropathological feature of AD. In the brain of people with AD, tau protein undergoes hyperphosphorylation because of some factors such as, formation of Aβ and environmental factors etc (Gong and Iqbal 2008). Thus, it is separated from microtubules (Mandelkow et al., 2007). Consequently, it results in neurofibrillary tangles (NFTs), ultimately leading to synapse loss and nerve cell destruction (Lauretti et al., 2020). Different kinases and phosphatases regulate the degree of phosphorylation of tau protein (Kolarova et al., 2012). In people with AD, tau protein is hyperphosphorylated, and it causes tauopathy due to degranulation and imbalance between them (Martin et al., 2013). One of the major kinases involved in tau protein phosphorylation is the glycogen synthase kinase_3 (GSK-3) enzyme (Jouanne et al., 2017). This serine–threonine kinase has two isoforms of alpha and beta (GSK3-α and GSK3-β) (Bagyinszky et al., 2014), which are encoded by two different genes (Lee et al., 2006).

In the central nervous system of people with AD, the GSK3-β enzyme is hyperactive, and evidence supports its role in AD pathology (Llorens-Marítin et al., 2014). GSK3-β phosphorylates many tau protein sites among people with AD and results in hyperphosphorylation of tau protein. Moreover, GSK3-β hyperactivity has been associated with impaired neurogenesis, microglia activation, amyloid-beta (Aβ) deposition, and inhibition of long-term potentiation (LTP) hippocampus and memory impairment (Llorens-Marítin et al., 2014).

This protein has two lobes, i.e., n-terminal and c-terminal, with residues 25 to 134 located in the N-terminal region and 35 to 380 in the c-lobe. Furthermore, these two lobes are held together by the hinge area such that residues 133 to 139 are considered a part of this area. Activation LOOP (A-Loop) is one of the most important loops of this enzyme where residues of A-LOOP are phosphorylated and have a key role in enzyme regulation. Among the residues of this loop, the DFG motif (ASP200-GLY202) is vital in enzyme regulation (ter Haar, Coll et al., 2001; Elangovan, Dhanabalan et al., 2020).

GSK3-β maintains an ATP-binding site between the two lobes. An important region of the ATP-binding site is the hinge region containing the residues Asp133, TYR134, and Val135. Some studies revealed that VAL135 and AP133 are very important residues for the binding of small molecules to the ATP-binding site (Gyurak et al., 2012; Pandey and DeGrado 2016; Davies 2019).

Hence, given the various roles of GSK3-β in AD pathology, the therapeutic potentials for its inhibition have been extensively examined in different studies.

Currently, there are no definitive treatments for AD (Eftekharzadeh et al., 2018), and the FDA-approved medicines for the disease are only for symptomatic treatment to improve behavioral changes and delay performance decline. These medicines include acetylcholinesterase inhibitors (e.g., donepezil, rivastigmine, and galantamine) and the N-methyl-D-aspartate receptor (NMDAR) (memantine). To date, several attempts have been made to find new inhibitors for AD, such as finding GSK3-β inhibitors.

The common snowberry plant (with the scientific name *Symphoricarpos albus*) has a fruit rich in phenolic acids, carbohydrates, and iridoids. The most important iridoid glycosides of this plant are loganin, loganetin, and secologanin (Gilbert 1995; Makarevich et al., 2009). Many biological activities like antioxidant, anti-inflammatory, anticancer, antidiabetic, antimicrobial, antiviral, anti-prosthetic, and neuroprotective activities have been reported for iridoid glycosides (Dinda et al., 2007; Dinda et al., 2019). In this respect, a study showed that iridoid glycosides effectively improve memory and cause nerve cell survival by increasing the expression of synaptophysin and neurotrophic factors in Alzheimer's mice with cholinergic defects. Additionally, these compounds enhance cognitive impairment and inhibit hyperphosphorylation of tau protein in the hippocampus and striatum of SAMP8 mice (Ma et al., 2016).

Another study indicated that iridoid glycosides could increase pp2a activity, decrease GSK3-β activity, and finally inhibit tau hyperphosphorylation. Therefore, these treatment options can effectively improve cognitive and behavioral disorders in AD patients (Wang et al., 2020; Zhang et al., 2020). In this respect, we have used computational drug design methods, which are one of the fastest and most cost-effective ways available for drug design and finding new inhibitors of the GSK3- β enzyme. Moreover, we examined the inhibitory potential of the GSK3-β enzyme by loganin, loganetin, and secologanin with the help of molecular docking methods and molecular dynamic (MD) simulation.

MATERIALS AND METHODS

Molecular Docking

Validation

Before molecular docking, we need to validate the software processing. Validation is of vital significance as it will eliminate false results. Since GSK3-β inhibition can be a

treatment for AD, many inhibitors have been examined; many of them are competitive inhibitors with the ATP site and target its binding site in GSK3-β. The co-crystal structures of these inhibitors can be obtained through the Protein Data Bank (PDB) site. Codes related to the GSK3-β protein as a complex with the inhibitor are 1Q41 (Bertrand et al., 2003), 1Q3D (Bertrand et al., 2003), 1Q4L (Bertrand et al., 2003), 2OW3 (Zhang et al., 2007), 1R0E (Allard. 2004), 1Q3W (Bertrand et al., 2003), 3GB2 (Saitoh et al., 2009), 3F88 (Saitoh et al., 2009), 3F7Z (Saitoh et al., 2009), 3I4B (Aronov, Tang et al., 2009), 1UV5 (Meijer et al., 2003), 3Q3B (Coffman et al., 2011), 3L1S (Arnost et al., 2010), 1Q5K (Bhat et al., 2003), and 2O5K (Shin et al., 2007).

After downloading the PDB file corresponding to each of these codes, the structure of the GSK3-β protein and the cognate ligand of each code were prepared and optimized in ViewerLite 5.0 (Lite, 1998) software. All water molecules and nonpolar hydrogen of protein and nonpolar hydrogen of inhibitors were removed. In the next step, we performed redocking for each complex using both AutoDock4.2 (Forli et al., 2012) and AutoDock Vina (Trott and Olson 2010) software. We have used the two software to select the most appropriate software during the validation phase. Moreover, 100 runs were considered for each complex. Afterward, the best ligand conformation with the most negative ΔG and the lowest root mean square deviation (RMSD) was considered for each complex. Eventually, using Visual Molecular Dynamics (VMD) 1.9.3(http://www.ks.uiuc.edu) software, we obtained the RMSD between the selected conformations for the ligand and the ligand in the obtained crystalline structure from the PDB site.

In the next step, the obtained RMSDs of redock using two software were compared to each other for choosing the best software. Between these two tools, whichever showed the lower RMSD than the other, we chose it for loganin, secologanin, and loganetin docking with GSK3-β because it means that it can better predict the active site of the enzyme than the other.

Additionally, after selecting the software, the code with the lowest RMSD was used as the selected code for protein–ligand docking.

Docking Loganin, Secologanin, and Loganetin With GSK3-β Protein

Ligand Structure Preparation

First, the two-dimensional structure of each ligand was prepared in the ChemDraw Ultra 2d 8.0 (Mendelsohn 2004) tool and then was transmitted into ChemDraw Ultra 3d 8-0 (Mendelsohn 2004) software for preparing the three-dimensional structure. Then, the structures of ligands were subjected to energy minimization with the default of the MOPAC program in ChemDraw Ultra 3d 8.0 and then were saved as in the format of .mol. So, we converted them to the PDB format by ViewerLite 5.0 and used them as inputs of selected software for molecular docking.

The Gasteiger–Marsili procedure was used for calculating the partial charges of atoms (Shahlaei et al., 2011). Nonpolar hydrogens were deleted and then rotatable bonds were determined. All rotatable bonds of ligands were assumed to be flexible.

Protein Structure Preparation

We used the protein structure related to a selected complex in the validation step that it had lowest RMSD. And then, all water molecules and co-crystallized ligands were removed by the ViewerLite 5.0 tool, and just chain A of the protein was considered for molecular docking. Afterward, all hydrogens were added, and nonpolar hydrogens were removed by the selected tool. Then, Kolman charges were determined for the protein.

Docking Procedure

Molecular docking was performed by selected software in the validation stage. We used the Lamarckian genetic algorithm (LGA) method in AuotoDock4.2 based on the previous studies because they have revealed that other approaches (simulated annealing and genetic algorithm) are less effective than LGA (Morris et al., 1998). A total of 100 independent runs were considered for each ligand.

The dimensions of the center grid box for loganin, loganetin, and secologanin were considered 94.63 $A^{\circ} \times 68.1680\ A^{\circ} \times 9.7880\ A^{\circ}$, and the software was allowed to search the entire volume of the protein. After finishing the docking, the two-dimensional figures of the results (e.g., the binding site and interactions between GSK3-β and inhibitors) were exhibited using LigPlot^{+} (Laskowski and Swindells 2011) software.

Molecular Dynamic Simulation

We carried out MD simulation in two steps. First, the structure of GSK3-β from a selected complex in the validation step was simulated in a water box. Second, to investigate the effects of ligand binding on GSK3-β conformation, the lowest binding energy conformation from the docking step of ligands was imported into the MD simulation.

For simulation, we used GROMAX 2019.1 (https://doi.org/10.5281/zenodo.2564761) software. At first, the protein was modified using the Swiss_PDB viewer (SPDBV) (Kaplan and Littlejohn 2001) program since the protein had the missing atom. Then, all the heteroatoms and water molecules were removed from the protein. In the next stage, topology and coordinate files were prepared. CHARMM27 was used as a force field, with the intermolecular (nonbonded) potential demonstration as a sum of Lennard–Jones (LJ) force and pairwise Coulomb interaction, and the long-range electrostatic force was defined by the particle mesh Ewald (PME) method. The velocity Verlet algorithm was applied for the numerical intermixture (Gharaghani et al., 2013).

Under periodic boundary conditions, the system was placed in a cubic water box of extended simple point-charge (SPC) water molecules with dimensions of $9.0465 \times 9.06465 \times 9.06465\ nm^3$ with 1 nm away from the wall on each side.

Moreover, the system has to be examined for electrostatic neutrality, and anions or cations should be added if necessary. Here, 6 chloride ions were added to neutralize the system, and the entire system was included 5,579 atoms of GSK3-β, 6 Cl^{-}, and 22,560 solvent atoms. In the next stage, the energy was minimized

in 50,000 steps for 2 fs. Then, the system was equilibrated at a constant temperature (NVT) of 300 k using the Berendsen thermostat. The cutoff radius was considered as 1.2 nm. Then, the system was balanced in the same way at a constant pressure (NVT) of 1 bar.

The goal of balancing the system at constant pressure and temperature for solvent molecules was to reach their best arrangement around the solute. Ultimately, the system was placed in the main run of the simulation for 100 ns at 300 K temperature and 1 bar pressure. To simulate the GSK3-β–ligand complex, first, the ligand-related topology file was generated using the SWISSPARAM server for the CHARMM27 force field. Afterward, the topology parameters and ligand coordinates were added to the topology parameters and coordinates of GSK3-β. Molecular dynamic simulation of the protein–ligand complex was performed similar to protein simulation for 100 ns. Ultimately, all the analyses related to the simulation were carried out in GROMACS 2019.1 software. These analyses include the evaluation of intermolecular hydrogen bonds, RMSD, Rg, RMSF, and free binding energy using the MM–GBSA method. The MD simulation was carried out on Ubuntu 18.04 Linux on an Intel Core 12 Quad 6800K 3.6 GHz, Gpu = gtx 1080ti NVIDIA, and 16 GB RAM.

Binding Free Energy Calculation

The binding free energy is a critical step of an in silico drug design approach which determines the binding affinity of inhibitors to the receptor. The binding free energy was calculated for the ligand receptor of loganin, loganetin, and secologanin complexes using MM/PBSA which defines the binding free energy of the protein and ligand as

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

Here, $\Delta G_{complex}$ demonstrates the total MMPBSA energy of the protein–ligand complex; $\Delta G_{protein}$, and ΔG_{ligand} are solution free energies of the individual protein and ligand, respectively. The free energy of the individual existence can be expressed as follows:

$$\Delta G = E_{MM} + G_{solvation} - TS \quad (2)$$

E_{MM} shows the average molecular mechanical's potential energy in the vacuum; $G_{solvation}$ defines the free energy of solvation. T and S explain the temperature and entropy, respectively, and together TS exhibits the entropic contribution to the free energy in vacuum. In addition, the E_{MM} contains both bonded and nonbonded interactions of the molecules, including the bond angle, torsion, and electrostatic (E_{elec}) and van der Waals (E_{vdw}) interactions. Last, the free energy of solvation and $G_{solvation}$ include both electrostatic and nonelectrostatic (G_{polar} and $G_{nonpolar}$) components. The binding free energies of the loganin_GSK3-β, loganetin_GSK3-β, and secologanin_GSK3-β complexes were calculated for 200 snapshots obtained from the last 20 ns of the trajectories (Padhi et al., 2021).

Drug Likeness Prediction and Toxicity

The OSIRIS property explorer (http://www.organic-chemistry.org/prog/peo/) was used to investigate the likeness of drugs. Pharmacokinetic properties like Log S calculation, TPSA, Clog P calculation, molecular mass, and toxicities like mutagenicity, tumorigenicity, irritation force, and hazard of the reproductive force of three compounds have been established.

TABLE 1 | Result of redock by AutoDock4.2 and AutoDock Vina software.

Code	RMSD of AutoDock4.2	RMSD of AutoDock Vina
1Q4L	2.4015	2.9621
1Q3D	3.011	2.4863
1Q41	6.2644	6.816
1Q3W	1.2191	5.0456
1R0E	4.27	6.02693
2OW3	4.0532	3.9890
1UV5	0.6679	0.5720
3I4B	7.3559	7.1482
3F7Z	4.7133	5.2366
3F88	5.916	5.7504
3GB2	3.3699	1.8433
1Q5K	6.7193	8.9578
2O5K	6.3006	10.6104
3L1S	4.115	5.2689
3Q3B	6.8446	6.77

RESULTS AND DISCUSSION

Molecular Docking

In the molecular docking stage, it was seen that the obtained RMSD by AutoDock4.2 in most codes is less than the obtained results by AutoDock Vina software (**Table 1**). Since the AutoDock4.2 software can accurately predict the active site of GSK3-β protein, it was selected to dock our ligands. Moreover, according to the RMSD results of all codes by this software, the code 1UV5 with RMSD = 0.6679 had the lowest RMSD compared to other codes. 1UV5 contains the 6-Bromine dirubio-3-oxime inhibitor (BRW1383), which inhibits GSK3-β protein competitively with the ATP active site.

Among the most important amino acids with a crucial role in regulating this kinase, one can name ASP133, ASP200, and VAL135. As shown in **Figure 1A**, interacted BRW1383 in AutoDock 4.2 software with a −9.56 Kcal/mol binding energy could be matched on the BRW1383 from X-ray crystallography well and could establish the same bonds as those reported in the X-ray crystallographic structure (**Figure 1B**).

Thus, the docking method used in this study could effectively identify how the inhibitor binds to the enzyme. The average binding energy was, respectively, −7.15, −5.43, and −4.98 kcal/mol in the molecular docking of loganin, loganetin, and secologanin with GSK3-β protein. As can be seen, the binding energy of loganin was positive compared to that of BRW1383 but negative compared to that of other ligands. Docking results showed that loganin had a hydrogen bond with the amino acids VAL135 and ASP200 and a hydrophobic bond with ASP133. The details of the interaction between loganin and GSK3-β provided by Ligplot software are given in **Figure 2B**. As already stated, these three amino acids are the key amino acids in the regulation of GSK3-β protein (**Figure 2A**). However,

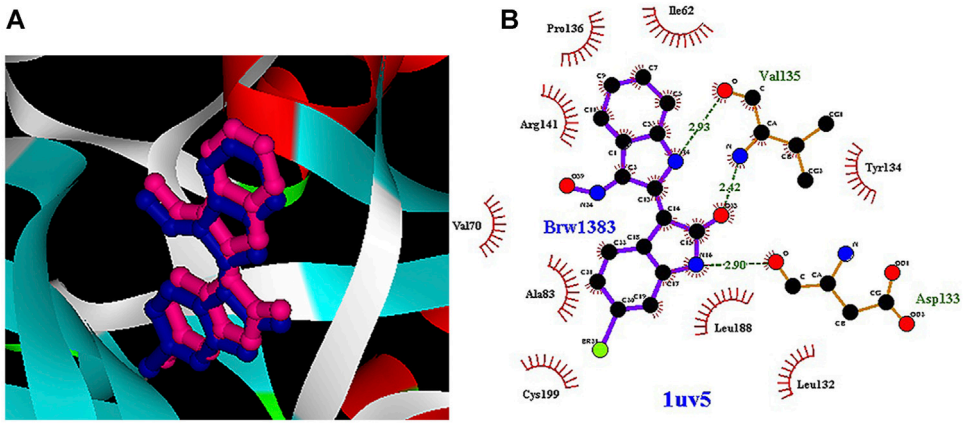


FIGURE 1 | (A) Interacted BRW1383 (pink) and BRW1383 from X-ray crystallography (blue). **(B)** Interactions of BRW1383 with GSK3_β protein obtained from X-ray crystallography.

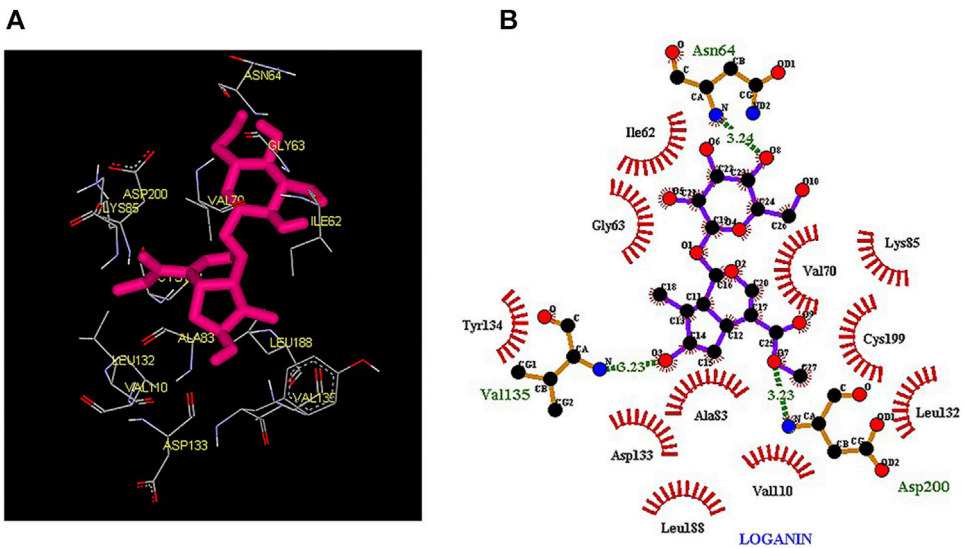


FIGURE 2 | (A) Three-dimensional view of loganin at the GSK_3β protein binding site. **(B)** Interactions of loganin with amino acids of the GSK3_β protein binding site.

BRW1383 has better binding energy vs. loganin, as shown in **Table 2** which demonstrates that loganin interacts with all of the key amino acids in the active site, but the reference inhibitor interacts with just ASP133 and VAL135. As loganin is a natural compound, it can be a choice for inhibiting GSK3-β protein compared with BRW1383 which is a chemical compound. According to **Figure 3**, loganetin has established a hydrogen bond with VAL135 and a hydrophobic bond with ASP133. Additionally, secologanin has established two hydrogen bonds with ASN64 and ARG141 and several hydrophobic bonds outside the active site (**Figure 4**).

The results in **Table 2** show that loganin binds more strongly to the active site of the GSK3-β protein compared to other ligands. To ensure the stability of the docked compound, the distance of hydrogen bonds between the compounds and the active site residues was investigated. According to Jeffrey's study (Jeffrey 1997), the acceptable hydrogen bond distance between the donor and acceptor should be from 2.7 to 3.3 A°. The 2.2–2.5 Å distance is considered strong and covalent, 2.5–3.2 Å is considered medium and electrostatic bonds, and 3.2–4 Å is considered weak electrostatic bonds. Moreover, Ippolito and colleagues (1990) and Raschka and colleagues (2018) have revealed that the average distance between the protein and ligand from 2.4 to 3.5 Å could be considered hydrogen bonding.

Thus, the hydrogen bond interaction of our ligands with the protein was considered the medium hydrogen bond type. Therefore, in this study, the moderate hydrogen bond seems to be enough to inhibit the GSK3-β protein.

TABLE 2 | Docking results of loganin, loganetin, secologanin, and BRW1383 with GSK3-β protein.

Ligand	Two-dimensional structure	Binding energy (Kcal.mol)	Hydrogen bonds	Bond length (A°)	Hydrophobic bonds
Loganin	fx1	−7.15	ASN64	3.24	Cys199 Asp133 Ala83
			Asp200	3.23	Val110 Leu188 Leu132
			Val135	3.23	Val70 Lys85 Ile62 Gly63
Loganetin	fx2	−5.43	Val135	2.95	Asp133 Leu132 Ala83
			Cys199	3.27	Val110 Tyr134 Thr138 Ile62 Val70 Leu188
Secologanin	fx3	−4.98	Asn64	3.34	Tyr140 Gln185 Cys199 Asp200
			Arg141	2.93	Asn186 Gly63 Val70 Ile62 Thr138
BRW1383	fx4	−9.53	Val135	2.93 2.42	Ile62, Val70 Ala83, Leu132 Tyr134
			Asp133	2.90	Thr138, Arg141, Leu188 Cys199

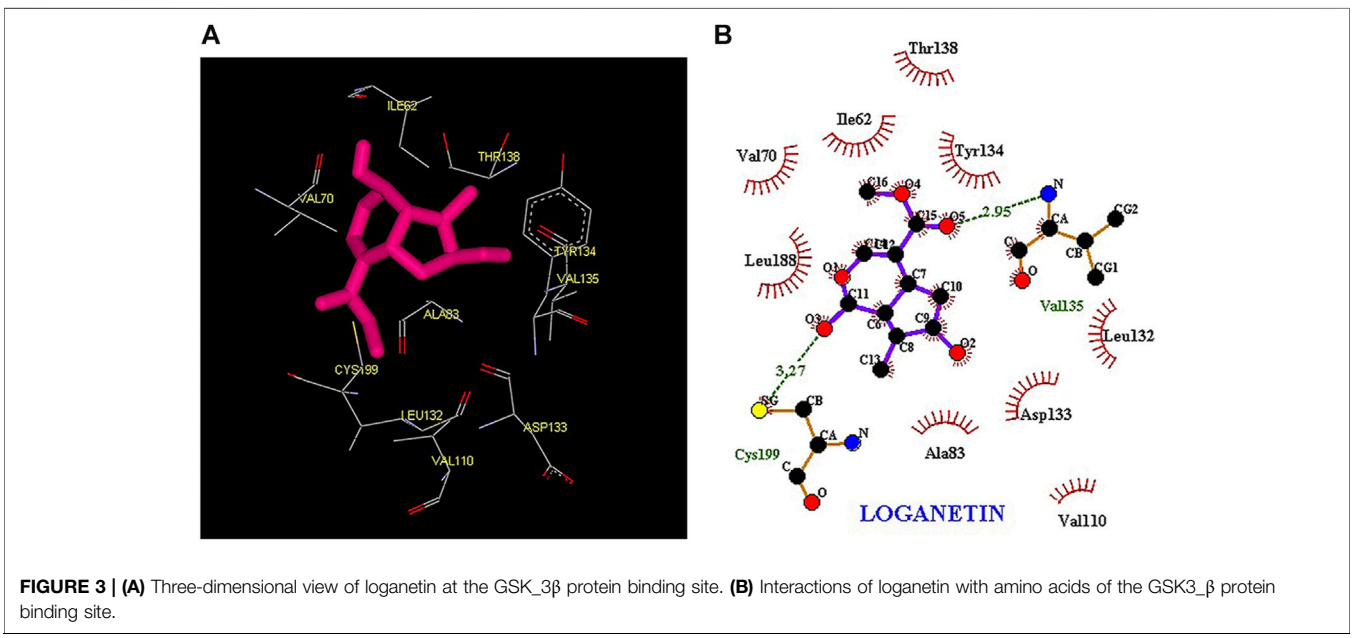


FIGURE 3 | (A) Three-dimensional view of loganetin at the GSK_3β protein binding site. **(B)** Interactions of loganetin with amino acids of the GSK3_β protein binding site.

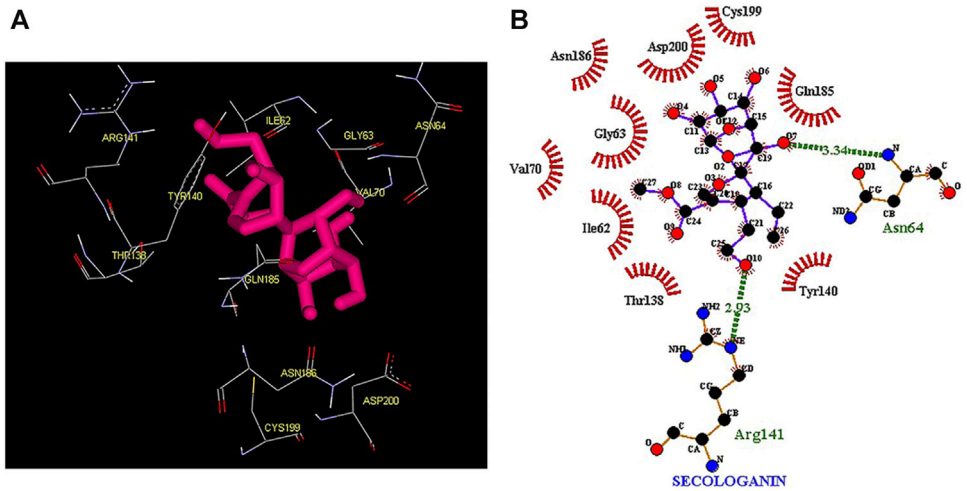


FIGURE 4 | (A) Three-dimensional view of secologanin at the GSK_3β protein binding site. **(B)** Interactions of secologanin with amino acids of the GSK3_β protein binding site.

TABLE 3 | The before and after simulation of protein–ligand interaction residues.

Compound	Active site amino acids before simulation	Active site amino acids after simulation
Loganin	ASP133, VAL135, ASP200, TYR134	VAL135,ASP133,TYR134
Loganetin	VAL135,ASP133,TYR134	VAL135,ASP133, TYR134
Secologanin	ASP200	ASP200

Molecular Dynamic Simulation

Although the docking results provide good information on the quality and interactions of the ligand and protein, they cannot predict how the protein conformations will change after ligand binding. Thus, to examine the changes and dynamics of the protein in interaction with three selected ligands, simulations were performed throughout for 100 ns

The before and after MD simulation of protein–ligand interaction residues are compared in **Table 3** (Saddala and Adi 2018).

The active site amino acids of all compounds in the molecular docking were almost similar to the active site residues in the MD simulation. Thus, these compounds have stayed stable in the active site after simulation and did not change the protein structure. Although, secologanin before and after simulation could not interact with active site residues well and just could interact with ASP200. Consequently, secologanin was not considered a GSK3-β inhibitor. Also, we have prepared 3d and 2d demonstrations of all compounds after MD simulation as shown in **Figures 5**–**7**. As can be seen, the most important functional groups of loganin and loganetin that participate in hydrogen bond interactions with active site residues are related to the sugar part of these iridoid glycosides.

In molecular docking, only the ligand has a flexible state. However, both the protein and ligand were considered flexible in MD simulation. Therefore, during MD simulation, we could evaluate all observed interactions in molecular docking and find potential and new effects regarding the conformational

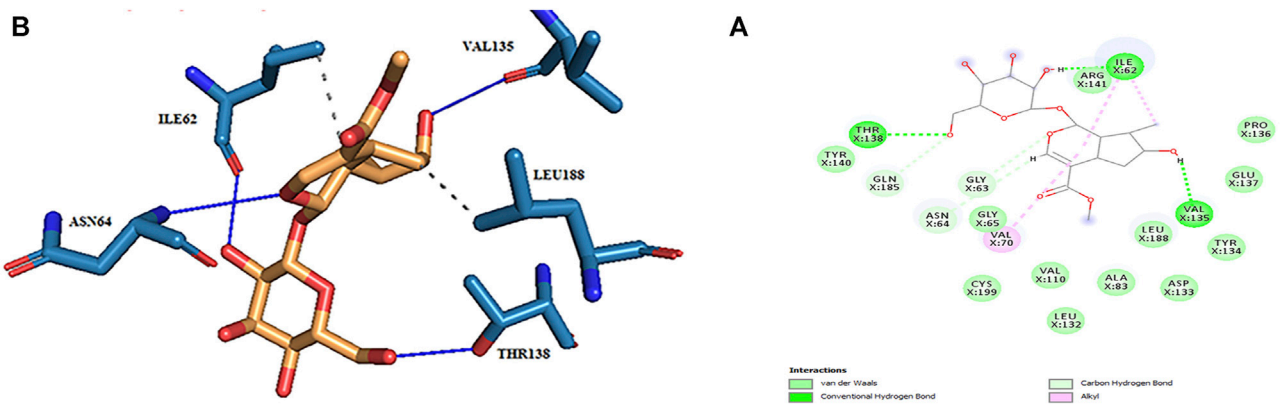


FIGURE 5 | (A) 2d and **(B)** 3d view of the GSK_3β_loganin complex after MD simulation.

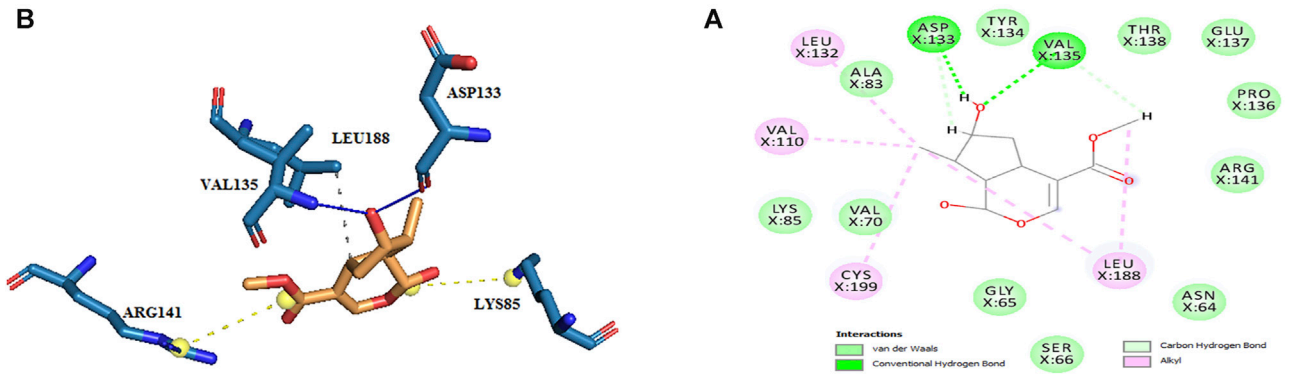


FIGURE 6 | (A) 2d and **(B)** 3d view of the GSK_3β_loganetin complex after MD simulation.

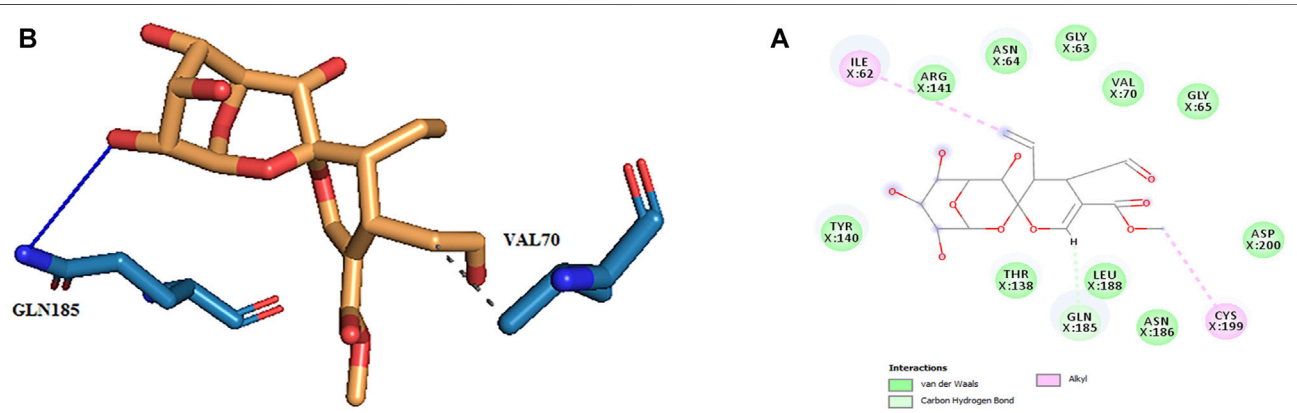


FIGURE 7 | (A) 2d and **(B)** 3d view of the GSK_3β_secologanin complex after MD simulation.

changes of the ligand and the enzyme-binding site. As shown in **Table 4**, in hydrogen bond examinations of the GSK3-β_loganin complex, the most stable interactions have been established between O val135 and H20 atoms with 97% occupancy and HG1 THR138 with O4 loganin with 94% occupancy. Here, val135 is the active site amino acid and is involved in the regulation of GSK3-β protein. Among the hydrogen bonds that loganin has established with the protein, only the hydrogen bond with val135 is identical to the docking results. Although the protein is considered rigid in docking, in MD simulation, both the protein and ligand are considered flexible; because of that, MD results are more reliable. It is also noteworthy that only two hydrogen bonds were predicted for secologanin in docking, but several hydrogen bonds have been predicted for MD. As **Table 4** shows, the occupancy listed for each of the secologanin hydrogen bonds is nonsignificant, and none of these bonds are established in the protein active site.

Loganetin has two other hydrogen bonds, where its H17 atom is bonded to oxygen ASP133, and its O2 atom is bonded to HN VAL135, whereas these two amino acids are of the residues of the active site of proteins; however, the occupancy of the loganin hydrogen bonds is greater. Thus, loganin binds more strongly to the protein compared to the other two ligands.

TABLE 4 | Analysis of intermolecular hydrogen bonds.

Ligand	Donor–acceptor	Occupancy (%)
Loganin	383LIG (H20)–135VAL(O)	80.1
	383LIG (H15)–62 ILE (O)	22.4
	138THR (HG1)–383LIG (O4)	94.5
	64ASN(HN)–383LIG (O8)	42.3
Loganetin	383LIG (H17)–133ASP(O)	84
	135VAL(HN)–383LIG (O2)	54.7
Secologanin	186ASN(D21)–383LIG (O4)	13.2
	183LYS(HZ1)–383LIG (O6)	15
	183LYS(HZ1)–383LIG (O4)	22.4
	66SER(HN)–383LIG (O5)	13
	64ASN(HN)–383LIG (O5)	14.9

Root Mean Square Deviation

The first critical analysis of MD simulation is the RMSD, which shows the stability and structural changes during the simulation. RMSD is a parameter that shows the deviation of the particle position from the original structure at any point in time (Czelen 2017; Sargsyan et al., 2017).

The lower RMSD value indicates less fluctuation during the simulation, suggesting that the stability of the protein is high

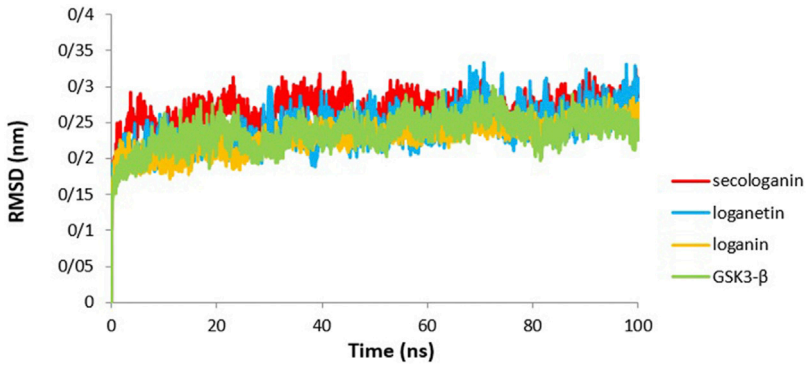


FIGURE 8 | RMSD analysis GSK3-β protein (green), loganin-GSK3-β complex (yellow), loganetin-GSK3-β complex (blue), and secologanin-GSK3-β complex (red).

(Elangovan et al., 2020). Therefore, the system will be balanced. **Figure 6** presents the variation in RMSD values of GSK3-β protein comparing with the three complexes. The RMSD value revealed that all pathways reached equilibrium after 10 ns Thus, RMSF, Rg, SASA, and hydrogen bonds were analyzed at an equilibrium state.

As **Figure 8** shows for the GSK3-β protein (green diagram), the mean RMSD has remained at about 0.19 nm with minimal fluctuation. Consequently, the protein does not undergo severe deformation and denaturation while simulating.

As can be seen from the GSK3-β_secologanin complex, the RMSD has higher values and fluctuations (0.23 nm) relative to the free protein and compared to loganin and loganetin about 0.19 and 0.21, respectively. Thus, the protein–secologanin complex is more unstable than other structures. Here, the protein–loganin RMSD compared to GSK3-β protein is more stable than the other two complexes. Also, the enzyme structure has not changed significantly in the presence of loganin, and the system is stable during the simulation.

Root Mean Square Fluctuation

Another used parameter to measure stability and flexibility is the RMSF during 90 ns. The RMSF estimates the average deviation of a particle (like protein residues) from a reference position (typically the average particle position) over time (Reddy et al., 2015).

Therefore, RMSF analyzes those parts of the structure with the most and least fluctuations from their dependent structure, and this parameter describes how ligand binding can result in conformational changes at the residual level (Elangovan et al., 2020; Saravanan et al., 2020).

Overall, the peaks seen in the RMSF diagram showed the most unstable target protein residues.

Figure 9 illustrates the high residue fluctuations at the start and end of the RMSF chart.

Moreover, from **Figure 9**, loganin has less fluctuation compare with loganetin and secologanin, and it has the most adaption with the RMSF GSK3-β protein diagram.

Moreover, amino acids' RMSF of the active site in each of the three ligands reveals that binding of the ligand to the protein in the active site does not cause the fluctuation of residues, and the structure remains stable. The mean RMSF values for GSK3-β protein, loganin, loganetin, and secologanin were 2.16, 2.16, 2.17, and 2.15, respectively. These values suggest that ligand binding does not change the original conformation of the residues. Also, it is inferred that the compounds do not fluctuate much, they seem consistent with protein fluctuations, and the complexes seem stable.

Radius of Gyration

The gyration radius is a variable that is studied to examine the compression changes during MD simulation. Protein compression during interactions with the ligand is affected by protein chains, in which the flexibility between the ligand and the protein depends on it (Shukla et al., 2019; Ghosh et al., 2020).

The calculated Rg value for the three complexes (**Figure 10**) was 2.16, which is associated with the protein size as well. The Rg of GSK3-β protein was 2.16, indicating that GSK3-β protein remains compressed during MD simulation and binding to all three ligands.

Solvent-Accessible Surface Area

The SASA value is used to analyze the magnitude and significance of ligand binding to the receptor and the changes in protein conformation due to ligand binding (Shukla et al., 2019; Elangovan et al., 2020).

As shown in **Figure 11**, all complexes show a similar magnitude of protein conformation that interacts with the ligand compared to the GSK3-β protein.

On the other hand, SASA changes are similar to the changes in the Rg diagram, confirming the accuracy of the molecular dynamic simulations. Thus, one can infer that all protein–ligand complexes are stable in the SASA analysis.

The Analysis of Hydrogen Bonds

Different interactions like hydrogen bonds, hydrophobic interactions, and ionic interactions stabilize the protein–ligand complex. Between them, the hydrogen bonds are more important and specific transient interactions than others for protein–ligand stabilization (Shukla et al., 2019). The number of hydrogen bonds was calculated for the last 90 ns trajectory. The number of hydrogen bonds vs. time is revealed for each complex in

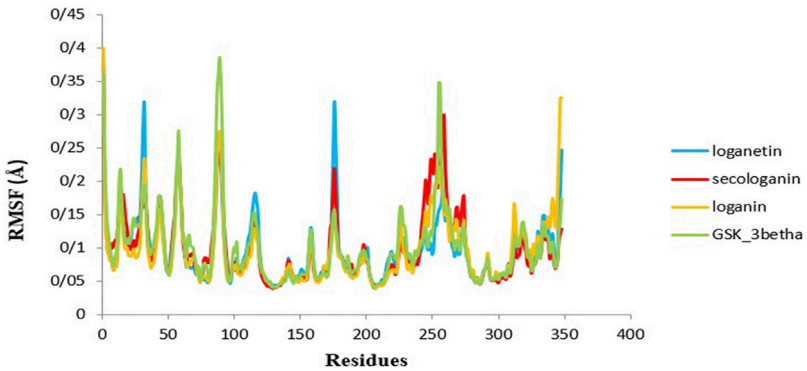


FIGURE 9 | RMSF analysis GSK3-β protein (green), loganin-GSK3-β complex (yellow), loganetin-GSK3-β complex (blue), and secologanin-GSK3-β complex (red).

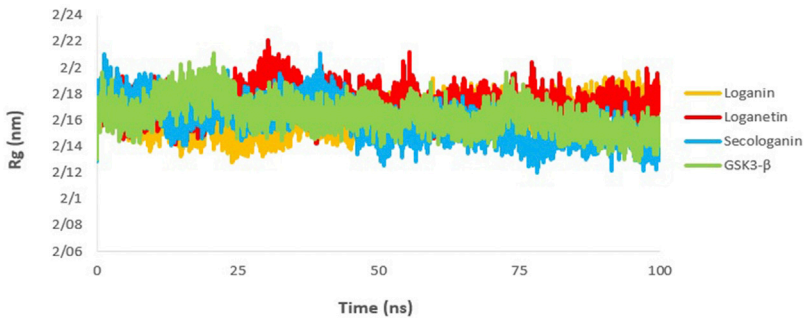


FIGURE 10 | Rg analysis GSK3-β protein (green), loganin-GSK3-β complex (yellow), loganetin-GSK3-β complex (blue), and secologanin-GSK3-β complex (red).

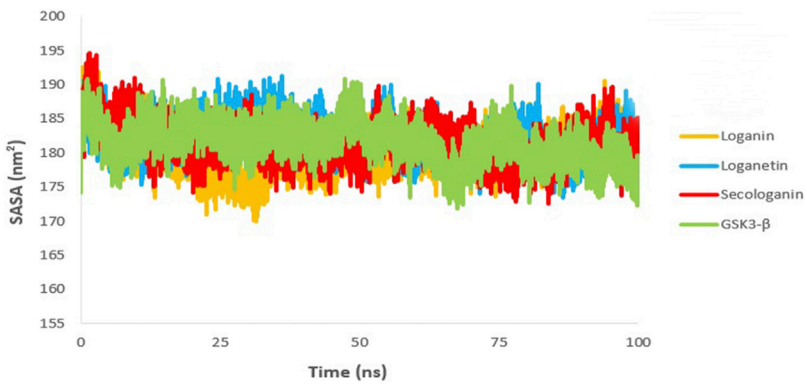


FIGURE 11 | SASA analysis GSK3-β protein (green), loganin-GSK3-β complex (yellow), loganetin-GSK3-β complex (blue), and secologanin-GSK3-β complex (red).

Figure 12. The average number of hydrogen bonds for loganin_GSK3-β (**Figure 12A**) and loganetin_GSK3-β was (**Figure 12B**) 3 and 2, respectively. But the hydrogen bonds of secologanin_GSK3-β are variable.

As can be seen, a partial change in the hydrogen bond formation between the ligand and protein for loganin_GSK3-β and loganetin_GSK3-β was observed. So, these complexes were stable for most parts of the simulation trajectory compared to secologanin_GSK3-β. These results are approximately similar to the results that are presented in **Table 4**.

Binding Free Energy Calculation

The binding free energy is a critical step of the in silico drug design approach which determines the binding affinity of

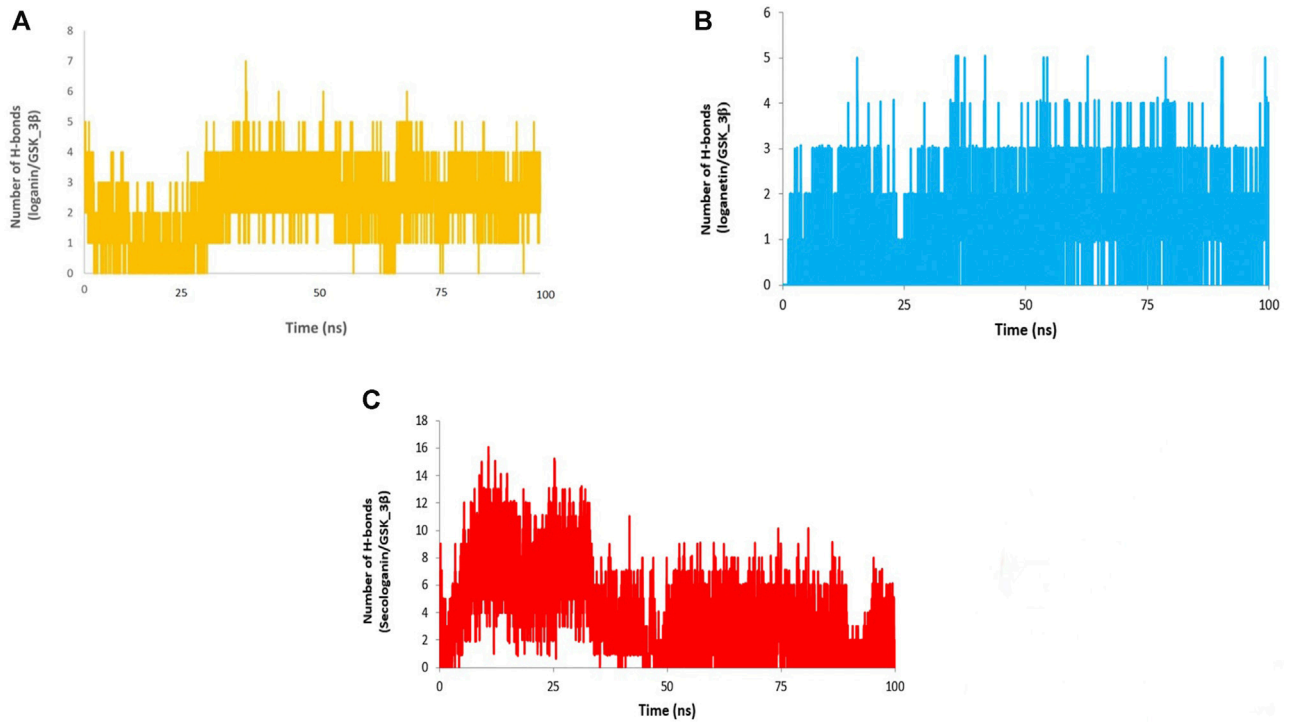


FIGURE 12 | H-bond analysis **(A)** loganin, **(B)** loganetin, and **(C)** secologanin.

TABLE 5 | Table represents the Van der Waals, electrostatic, polar solvation, SASA, and binding energy in kJ.mol_1 for secologanin, loganetin, and loganin.

Compound	van der Waals energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy
Secologanin	−150.661 ± 11.942	−21.726 ± 12.326	187.567 ± 826.391	−17.050 ± 1.152	−1.870 ± 827.549
Loganetin	−105.747 ± 8.947	−30.152 ± 8.113	95.579 ± 16.494	−12.559 ± 0.718	−52.878 ± 11.941
Loganin	−157.429 ± 12.579	−44.508 ± 12.625	168.159 ± 15.026	−18.644 ± 0.932	−52.421 ± 14.661

inhibitors to the receptor. The binding free energy was calculated for the ligand receptor of loganin, loganetin, and secologanin complexes using the MM/PBSA method. The nonpolar and polar solvation energy was analyzed in the electrostatic interaction, van der Waals energy, and SASA energy. The major desirable portions of the stereoisomer binding were the van der Waals and electrostatic energies for all complexes. Besides, the terms polar solvation free energy and SASA energy are unfavorable for binding in three complexes. The total free binding energy ($\Delta G_{Binding}$) was calculated for each compound in **Table 5**.

Loganin and loganetin have an almost similar amounts of binding energy, respectively, −52.421 and −52.878 KJ/mol, but they are much higher than that of secologanin (−1.870 KJ/mol). We have concluded that loganin and loganetin are more energetically favorable than secologanin, and these results are adopted with docking conclusions.

Drug Likeness Prediction and Toxicity

Pharmacokinetic properties and toxicities were predicted for each compound used by the OSIRIS property explorer server, and the results are revealed within **Table 6**. The mutagenicity, tumorigenicity, irritation force, and hazard of the reproductive force were predicted for toxicity verification.

The higher cLog P (logarithm of compounds' partition coefficient between water and n-octanol) value indicates lower permeation and absorption due to lower hydrophilicity. The solubility and molecular weight (MW) affect the absorption rate; thus, high solubility and lower MW increase absorption. The topological polar surface area (TPSA) reveals the surface of polar atoms in the compounds. An increased TPSA value is relevant to the least permeability of the membrane. So, the compounds with a larger TPSA value will be a better substrate for p-glycoprotein which is amenable for the efflux of a drug from the cell, and thus, the diminished TPSA was useful for the drug-like property. Some investigations also predicted that a molecule with a better CNS penetration should have a lesser value of the TPSA (Reddy et al., 2015; Srivastava et al., 2020).

Furthermore, the drug score mixes drug likeness, clog P, TPSA, MW, and toxicity risk parameters to reveal a compound as a drug candidate. Based on **Table 6**, loganetin

TABLE 6 | Drug likeness prediction and toxicity of loganin, loganetin, and secologanin.

Properties	Loganin	Loganetin	Secologanin
Mutagenic	NO	NO	NO
Tumorigenic	NO	NO	NO
Irritant	NO	NO	High risk fragment indicating irritating effects
Reproductive effective	NO	NO	NO
C Log p	−1.97	−0.13	−1.85
Solubility	−1.16	−1.46	−1.23
MW	390	228	388
TPSA	155.1	76.99	151.9
Drug likeness	−3.92	−0.11	−10.16
Drug score	0.45	0.71	0.27

has a higher score (0.71) than loganin (0.45) and secologanin (0.27) because loganin has a lower TPSA, MW, and cLog P and higher solubility than other compounds.

CONCLUSION

This study aimed at finding potential and available inhibitors for GSK3-β protein using computational drug design methods like molecular docking and MD simulation.

We performed molecular analysis of the iridoid glycosides in the Common snowberry plant, including loganin, secologanin, and loganetin, which have beneficial effects on improving memory.

Molecular docking analysis revealed that loganin and loganetin bind exactly to the ATP-binding site. Loganin established hydrogen bonding with VAL135 and ASP200 and hydrophobic bonding with ASP133, and loganetin established hydrogen bonding with VAL135 and hydrophobic bonding with ASP133, but the binding affinity of loganin was more than that of loganetin.

RMSD, RMSF, Rg, SASA, hydrogen bond analyses, and MMPBSA were carried out in the molecular dynamic simulation. The results of the RMSD value indicated that the loganin–GSK3 complex has more stability than the other two complexes.

The RMSF diagram indicated that ligand binding does not increase fluctuations, and the complexes stay stable following the binding.

The SASA analysis was equal with variations of the Rg value, and equality of the Rg value of the GSK3-β protein with complexes indicates that the structure remains compact during MD simulation.

Loganin and loganetin have an almost similar amount of binding energy, and loganetin has a better drug score than others.

Consequently, both loganin and loganetin may effectively inhibit GSK3-β because these compounds established strong hydrogen bonds with the active site residues before and after MD simulation. Although the binding affinity of loganin and loganetin was slightly less than that of the BRW1383 inhibitor in molecular docking, they can be considered to prevent tauopathy since those are available and nontoxic herbal compounds. Nonetheless, this study requires more examination, and it is better to carry out laboratory studies along with molecular studies to prove the effects of iridoid glycosides on AD prevention and treatment.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

ME and SP analyzed data and were responsible for manuscript writing. PK was the advisor and assisted in gathering information and research. BA assisted in research and writing the primary version of the manuscript. SG and JK were the research supervisors. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We gratefully appreciate the support of the Faculty of Pharmacy of Lorestan University of Medical Sciences for providing the facilities to carry out this study.

REFERENCES

Allard, J. (2004). "From Genetics to Therapeutics: the Wnt Pathway and Osteoporosis." 1r0e PDB entry.

Arnost, M., Pierce, A., Haar, E. t., Lauffer, D., Madden, J., Tanner, K., et al. (2010). 3-Aryl-4-(arylhydrazono)-1H-pyrazol-5-ones: Highly Ligand Efficient and Potent Inhibitors of GSK3β. *Bioorg. Med. Chem. Lett.* 20 (5), 1661–1664. doi:10.1016/j.bmcl.2010.01.072

Aronov, A. M., Tang, Q., Martinez-Botella, G., Bemis, G. W., Cao, J., Chen, G., et al. (2009). Structure-guided Design of Potent and Selective Pyrimidylpyrrole Inhibitors of Extracellular Signal-Regulated Kinase (ERK) Using Conformational Control. *J. Med. Chem.* 52 (20), 6362–6368. doi:10.1021/jm900630q

Association, A. s. (2019). 2019 Alzheimer's Disease Facts and Figures. *Alzheimer's Demen.* 15 (3), 321–387. doi:10.1016/j.jalz.2019.01.010

Bagyinszky, E., Youn, Y. C., An, S., and Kim, S. (2014). The Genetics of Alzheimer's Disease. *Cia* 9, 535. doi:10.2147/cia.s51571

Bertrand, J. A., Thieffine, S., Vulpetti, A., Cristiani, C., Valsasina, B., Knapp, S., et al. (2003). Structural Characterization of the GSK-3β Active Site Using Selective and Non-selective ATP-Mimetic Inhibitors. *J. Mol. Biol.* 333 (2), 393–407. doi:10.1016/j.jmb.2003.08.031

Bhat, R., Xue, Y., Berg, S., Hellberg, S., Ormö, M., Nilsson, Y., et al. (2003). Structural Insights and Biological Effects of Glycogen Synthase Kinase 3-specific Inhibitor AR-A014418. *J. Biol. Chem.* 278 (46), 45937–45945. doi:10.1074/jbc.m306268200

Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the Global burden of Alzheimer's Disease. *Alzheimer's Demen.* 3 (3), 186–191. doi:10.1016/j.jalz.2007.04.381

Coffman, K., Brodney, M., Cook, J., Lanyon, L., Pandit, J., Sakya, S., et al. (2011). 6-Amino-4-(pyrimidin-4-yl)pyridones: Novel Glycogen Synthase Kinase-3β Inhibitors. *Bioorg. Med. Chem. Lett.* 21 (5), 1429–1433. doi:10.1016/j.bmcl.2011.01.017

Czeleń, P. (2017). Inhibition Mechanism of CDK-2 and GSK-3β by a Sulfamoylphenyl Derivative of Indoline—A Molecular Dynamics Study. *J. Mol. Model.* 23 (8), 1–11.

Davies, M. P. (2019). *Silico Screening and* in Vitro *Evaluation of GSK-3β Type I and II Inhibitors: Potential Treatment for Alzheimer's Disease?* University of Central Lancashire.

De Strooper, B., and Karran, E. (2016). The Cellular Phase of Alzheimer's Disease. *Cell* 164 (4), 603–615. doi:10.1016/j.cell.2015.12.056

Dinda, B., Debnath, S., and Harigaya, Y. (2007). Naturally Occurring Secoiridoids and Bioactivity of Naturally Occurring Iridoids and Secoiridoids. A Review, Part 2. *Chem. Pharm. Bull.* 55 (5), 689–728. doi:10.1248/cpb.55.689

Dinda, B., Dinda, M., Kulsi, G., Chakraborty, A., and Dinda, S. (2019). Therapeutic Potentials of Plant Iridoids in Alzheimer's and Parkinson's Diseases: A Review. *Eur. J. Med. Chem.* 169, 185–199. doi:10.1016/j.ejmech.2019.03.009

Eftekharzadeh, B., Daigle, J. G., Kapinos, L. E., Coyne, A., Schiantarelli, J., Carlomagno, Y., et al. (2018). Tau Protein Disrupts Nucleocytoplasmic Transport in Alzheimer's Disease. *Neuron* 99 (5), 925–940. doi:10.1016/j.neuron.2018.07.039

Elangovan, N. D., Dhanabalan, A. K., Gunasekaran, K., Kandimalla, R., and Sankarganesh, D. (2020). Screening of Potential Drug for Alzheimer's Disease: a Computational Study with GSK-3 β Inhibition through Virtual Screening, Docking, and Molecular Dynamics Simulation. *J. Biomol. Struct. Dyn.*, 1–15. doi:10.1080/07391102.2020.1805362

Forli, S., and Olson, A. J. (2012). A Force Field with Discrete Displaceable Waters and Desolvation Entropy for Hydrated Ligand Docking. *J. Med. Chem.* 55 (2), 623–638. doi:10.1021/jm2005145

Gharaghani, S., Khayamian, T., and Ebrahimi, M. (2013). Molecular Dynamics Simulation Study and Molecular Docking Descriptors in Structure-Based QSAR on Acetylcholinesterase (AChE) Inhibitors. *SAR QSAR Environ. Res.* 24 (9), 773–794. doi:10.1080/1062936x.2013.792877

Ghosh, R., Chakraborty, A., Biswas, A., and Chowdhuri, S. (2020). Identification of Polyphenols from Broussonetia Papyrifera as SARS CoV-2 Main Protease Inhibitors Using In Silico Docking and Molecular Dynamics Simulation Approaches. *J. Biomol. Struct. Dyn.*, 1–14. doi:10.1080/07391102.2020.1802347

Gilbert, O. L. (1995). Symphoricarpos Albus (L.) S. F. Blake (S. Rivularis Suksd., S. Racemosus Michaux). *J. Ecol.* 83 (1), 159–166. doi:10.2307/2261160

Gong, C.-X., and Iqbal, K. (2008). Hyperphosphorylation of Microtubule-Associated Protein Tau: a Promising Therapeutic Target for Alzheimer Disease. *Cmc* 15 (23), 2321–2328. doi:10.2174/092986708785909111

Gyurak, A., Goodkind, M. S., Kramer, J. H., Miller, B. L., and Levenson, R. W. (2012). Executive Functions and the Down-Regulation and Up-Regulation of Emotion. *Cogn. Emot.* 26 (1), 103–118. doi:10.1080/02699931.2011.557291

Hall, A., Pekkala, T., Polvikoski, T., van Gils, M., Kivipelto, M., Lötjönen, J., et al. (2019). Prediction Models for Dementia and Neuropathology in the Oldest Old: the Vantaa 85+ Cohort Study. *Alzheimers Res. Ther.* 11 (1), 11–12. doi:10.1186/s13195-018-0450-3

Ippolito, J. A., Alexander, R. S., and Christianson, D. W. (1990). Hydrogen Bond Stereochemistry in Protein Structure and Function. *J. Mol. Biol.* 215 (3), 457–471. doi:10.1016/s0022-2836(05)80364-x

Jeffrey, G. (1997). *An Introduction to Hydrogen Bonding*. New York: Oxford University PressGoogle Scholar There is no corresponding record for this reference, 220–225.

Jouanne, M., Rault, S., and Voisin-Chiret, A.-S. (2017). Tau Protein Aggregation in Alzheimer's Disease: an Attractive Target for the Development of Novel Therapeutic Agents. *Eur. J. Med. Chem.* 139, 153–167. doi:10.1016/j.ejmech.2017.07.070

Kametani, F., and Hasegawa, M. (2018). Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease. *Front. Neurosci.* 12, 25. doi:10.3389/fnins.2018.00025

Kaplan, W., and Littlejohn, T. G. (2001). Swiss-PDB Viewer (Deep View). *Brief. Bioinformatics* 2 (2), 195–197. doi:10.1093/bib/2.2.195

Kolarova, M., García-Sierra, F., Bartos, A., Ricny, J., and Ripova, D. (2012). Structure and Pathology of Tau Protein in Alzheimer Disease. *Int. J. Alzheimers Dis.* 2012, 731526. doi:10.1155/2012/731526

Laskowski, R. A., and Swindells, M. B. (2011). *LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery*. European Bioinformatics Institute.

Lauretti, E., Dincer, O., and Praticò, D. (2020). Glycogen Synthase Kinase-3 Signaling in Alzheimer's Disease. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1867 (5), 118664. doi:10.1016/j.bbamcr.2020.118664

Lee, S. J., Chung, Y. H., Joo, K. M., Lim, H. C., Jeon, G. S., Kim, D., et al. (2006). Age-related Changes in Glycogen Synthase Kinase 3β (GSK3β) Immunoreactivity in the central Nervous System of Rats. *Neurosci. Lett.* 409 (2), 134–139. doi:10.1016/j.neulet.2006.09.026

Lin, R., Jones, N. C., and Kwan, P. (2020). Unravelling the Role of Glycogen Synthase Kinase-3 in Alzheimer's Disease-Related Epileptic Seizures. *Ijms* 21 (10), 3676. doi:10.3390/ijms21103676

Llorens-Marítin, M., Jurado, J., Hernández, F., and Avila, J. (2014). GSK-3β, a Pivotal Kinase in Alzheimer Disease. *Front. Mol. Neurosci.* 7, 46.

Ma, D., Zhu, Y., Li, Y., Yang, C., Zhang, L., Li, Y., et al. (2016). Beneficial Effects of Cornel Iridoid Glycoside on Behavioral Impairment and Senescence Status in SAMP8 Mice at Different Ages. *Behav. Brain Res.* 312, 20–29. doi:10.1016/j.bbr.2016.06.008

Makarevich, I. F., Kovalenko, S. N., Gusarova, T. D., and Gubin, Y. I. (2009). Iridoids from Symphoricarpos Albus. *Chem. Nat. Compd.* 45 (1), 40–44. doi:10.1007/s10600-009-9257-6

Mandelkow, E., von Bergen, M., Biernat, J., and Mandelkow, E.-M. (2007). Structural Principles of Tau and the Paired Helical Filaments of Alzheimer? s Disease. *Brain Pathol.* 17 (1), 83–90. doi:10.1111/j.1750-3639.2007.00053.x

Martin, L., Latypova, X., Wilson, C. M., Magnaudeix, A., Perrin, M.-L., Yardin, C., et al. (2013). Tau Protein Kinases: Involvement in Alzheimer's Disease. *Ageing Res. Rev.* 12 (1), 289–309. doi:10.1016/j.arr.2012.06.003

Meijer, L., Skaltsounis, A.-L., Magiatis, P., Polychronopoulos, P., Knockaert, M., Leost, M., et al. (2003). GSK-3-selective Inhibitors Derived from Tyrian Purple Indirubins. *Chem. Biol.* 10 (12), 1255–1266. doi:10.1016/j.chembiol.2003.11.010

Mendelsohn, L. D. (2004). ChemDraw 8 Ultra, Windows and Macintosh Versions. *J. Chem. Inf. Comput. Sci.* 44, 2225–2226. doi:10.1021/ci040123t

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* 19 (14), 1639–1662. doi:10.1002/(sici)1096-987x(19981115)19:14<1639:aid-jcc10>3.0.co;2-b

Padhi, A. K., Seal, A., Khan, J. M., Ahamed, M., and Tripathi, T. (2021). Unraveling the Mechanism of Arbidol Binding and Inhibition of SARS-CoV-2: Insights from Atomistic Simulations. *Eur. J. Pharmacol.* 894, 173836. doi:10.1016/j.ejphar.2020.173836

Pandey, M. K., and DeGrado, T. R. (2016). Glycogen Synthase Kinase-3 (GSK-3)-Targeted Therapy and Imaging. *Theranostics* 6 (4), 571–593. doi:10.7150/thno.14334

Raschka, S., Wolf, A. J., Bemister-Buffington, J., and Kuhn, L. A. (2018). Protein-ligand Interfaces Are Polarized: Discovery of a strong Trend for Intermolecular Hydrogen Bonds to Favor Donors on the Protein Side with Implications for Predicting and Designing Ligand Complexes. *J. Comput. Aided Mol. Des.* 32 (4), 511–528. doi:10.1007/s10822-018-0105-2

Reddy, S. V. G., Reddy, K. T., Kumari, V. V., and Basha, S. H. (2015). Molecular Docking and Dynamic Simulation Studies Evidenced Plausible Immunotherapeutic Anticancer Property by Withaferin A Targeting

Indoleamine 2,3-dioxygenase. *J. Biomol. Struct. Dyn.* 33 (12), 2695–2709. doi:10.1080/07391102.2015.1004834

Saddala, M. S., and Adi, P. J. (2018). Discovery of Small Molecules through Pharmacophore Modeling, Docking and Molecular Dynamics Simulation against Plasmodium Vivax Vivapain-3 (VP-3). *Heliyon* 4 (5), e00612. doi:10.1016/j.heliyon.2018.e00612

Saitoh, M., Kunitomo, J., Kimura, E., Hayase, Y., Kobayashi, H., Uchiyama, N., et al. (2009). Design, Synthesis and Structure-Activity Relationships of 1,3,4-oxadiazole Derivatives as Novel Inhibitors of Glycogen Synthase Kinase-3β. *Bioorg. Med. Chem.* 17 (5), 2017–2029. doi:10.1016/j.bmc.2009.01.019

Saitoh, M., Kunitomo, J., Kimura, E., Iwashita, H., Uno, Y., Onishi, T., et al. (2009). 2-{3-[4-(Alkylsulfinyl)phenyl]-1-benzofuran-5-yl}-5-methyl-1,3,4-oxadiazole Derivatives as Novel Inhibitors of Glycogen Synthase Kinase-3β with Good Brain Permeability. *J. Med. Chem.* 52 (20), 6270–6286. doi:10.1021/jm900647e

Saravanan, K., Hunday, G., and Kumaradhas, P. (2020). Binding and Stability of Indirubin-3-Monoxime in the GSK3β Enzyme: a Molecular Dynamics Simulation and Binding Free Energy Study. *J. Biomol. Struct. Dyn.* 38 (4), 957–974. doi:10.1080/07391102.2019.1591301

Sargsyan, K., Grauffel, C., and Lim, C. (2017). How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *J. Chem. Theor. Comput.* 13 (4), 1518–1524. doi:10.1021/acs.jctc.7b00028

Shahlaei, M., Madadkar-Sobhani, A., Fassihi, A., and Saghaie, L. (2011). Exploring a Model of a Chemokine Receptor/ligand Complex in an Explicit Membrane Environment by Molecular Dynamics Simulation: the Human CCR1 Receptor. *J. Chem. Inf. Model.* 51 (10), 2717–2730. doi:10.1021/ci200261f

Shin, D., Lee, S.-C., Heo, Y.-S., Lee, W.-Y., Cho, Y.-S., Kim, Y. E., et al. (2007). Design and Synthesis of 7-Hydroxy-1h-Benzoimidazole Derivatives as Novel Inhibitors of Glycogen Synthase Kinase-3β. *Bioorg. Med. Chem. Lett.* 17 (20), 5686–5689. doi:10.1016/j.bmcl.2007.07.056

Shukla, R., Munjal, N. S., and Singh, T. R. (2019). Identification of Novel Small Molecules against GSK3β for Alzheimer's Disease Using Chemoinformatics Approach. *J. Mol. Graphics Model.* 91, 91–104. doi:10.1016/j.jmgm.2019.06.008

Srivastava, V., Yadav, A., and Sarkar, P. (2020). Molecular Docking and ADMET Study of Bioactive Compounds of Glycyrrhiza Glabra against Main Protease of SARS-CoV2. *Mater. Today Proc.* doi:10.1016/j.matpr.2020.10.055

ter Haar, E., Coll, J. T., Austen, D. A., Hsiao, H.-M., Swenson, L., and Jain, J. (2001). Structure of GSK3β Reveals a Primed Phosphorylation Mechanism. *Nat. Struct. Biol.* 8 (7), 593–596. doi:10.1038/89624

Trott, O., and Olson, A. J. (2010). AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi:10.1002/jcc.21334

Lite, V. (1998). Version 5.0. Accelrys Inc 9685.

Wang, C., Gong, X., Bo, A., Zhang, L., Zhang, M., Zang, E., et al. (2020). Iridoids: Research Advances in Their Phytochemistry, Biological Activities, and Pharmacokinetics. *Molecules* 25 (2), 287. doi:10.3390/molecules25020287

Zhang, H.-C., Boñaga, L. V. R., Ye, H., Derian, C. K., Damiano, B. P., and Maryanoff, B. E. (2007). Novel Bis(indolyl)maleimide Pyridinophanes that Are Potent, Selective Inhibitors of Glycogen Synthase Kinase-3. *Bioorg. Med. Chem. Lett.* 17 (10), 2863–2868. doi:10.1016/j.bmcl.2007.02.059

Zhang, Y., Huang, N., Lu, H., Huang, J., Jin, H., Shi, J., et al. (2020). Icariin Protects against Sodium Azide-Induced Neurotoxicity by Activating the PI3K/Akt/GSK-3β Signaling Pathway. *PeerJ* 8, e8955. doi:10.7717/peerj.8955

Conflict of Interest: BA is the president and CEO at CaroGen Corporation in Farmington, CT United States of America.

The remaining 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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership