Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Computational modeling predicts potential effects of the herbal infusion “horchata” against COVID-19

Eduardo Tejera\textsuperscript{a,b,*}, Yunieris Pérez-Castillo \textsuperscript{a,c}, Gisselle Toscano \textsuperscript{b}, Ana Lucía Noboa \textsuperscript{d}, Valeria Ochoa-Herrera \textsuperscript{d,e}, Francesca Giampieri \textsuperscript{f,g}, José M. Álvarez-Suarez \textsuperscript{b,h,i,j}

\textsuperscript{a} Grupo de Bio-Qumioinformática, Universidad de Las Américas, Quito, Ecuador
\textsuperscript{b} Facultad de Ingeniería y Ciencias Aplicadas, Universidad de Las Américas, Quito, Ecuador
\textsuperscript{c} Escuela de Ciencias Físicas y Matemáticas, Universidad de Las Américas, Quito, Ecuador
\textsuperscript{d} Colegio de Ciencias e Ingenierías, Instituto Biósfera, Universidad San Francisco de Quito, Quito, Ecuador
\textsuperscript{e} Departamento de Ingeniería en Alimentos, Colegio de Ciencias e Ingenierías, Universidad San Francisco de Quito, Quito, Ecuador
\textsuperscript{f} Department of Biochemistry, Faculty of Sciences, King Abdullah University, Jeddah, Saudi Arabia
\textsuperscript{g} Departamento de Ingeniería en Alimentos, Colegio de Ciencias e Ingenierías, Universidad San Francisco de Quito, Quito, Ecuador
\textsuperscript{h} Instituto de Investigaciones en Biomedicina iBioMed, Universidad San Francisco de Quito, Quito, Ecuador
\textsuperscript{i} King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

Keywords: Horchata infusion, Herbal infusion, SARS-CoV-2, COVID-19, Molecular dynamics, Docking, Gene ontology

ABSTRACT

Bioactive plant-derived molecules have emerged as therapeutic alternatives in the fight against the COVID-19 pandemic. In this investigation, principal bioactive compounds of the herbal infusion “horchata” from Ecuador were studied as potential novel inhibitors of the SARS-CoV-2 virus. The chemical composition of horchata was determined through a HPLC-DAD/ESI-MS\textsuperscript{a} and GC–MS analysis while the inhibitory potential of the compounds on SARS-CoV-2 was determined by a computational prediction using various strategies, such as molecular docking and molecular dynamics simulations. Up to 51 different compounds were identified. The computational analysis of predicted targets reveals the compounds’ possible anti-inflammatory (non-steroidal) and antioxidant effects. Three compounds were identified as candidates for M\textsuperscript{pro} inhibition: benzoic acid, 2-(ethyl-thio)-ethyl ester, L-Leucine-N-isobutoxycarbonyl-N-methyl-heptyl and isorhamnetin and for PL\textsuperscript{a,b}: isorhamnetin-3-O-(6-Orhamnosyl-galactoside), dihydroxy-methoxyflavanone and dihydroxyphenyl)-5-hydroxy-4-oxochromen-7-yl)oxy-3,4,5-trihydroxyxoxane-2-carboxylic acid. Our results suggest the potential of Ecuadorian horchata infusion as a starting scaffold for the development of new inhibitors of the SARS-CoV-2 M\textsuperscript{pro} and PL\textsuperscript{a,b} enzymes.

1. Introduction

Severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) was described in Wuhan, China, in late 2019 after several rare cases of pneumonia (Zhu \textit{et al.}, 2020). The disease caused by SARS-CoV-2, named COVID-19 (coronavirus disease 2019), spread rapidly around the world, and by March 11, 2020, the WHO determined that COVID-19 had reached the level of a pandemic (WHO, 2020). On March 11, 2021, exactly one year after the pandemic was declared, more than one hundred and eighteen million cases of COVID-19 and two million six hundred thousand deaths have been confirmed worldwide (CSSE Johns Hopkins, 2021). This complex situation has led to a race in the search for new treatments and actions of rapid application aimed at treating the disease and reducing its spread. Thus, the pursuit of new alternatives for treating and containing COVID-19 has been the objective of many research groups, where molecules derived from plants have emerged as candidates for therapies in the prevention and treatment of this disease (Diniz, Perez-Castillo, Elshabrawy, Bezerra Filho, & de Sousa, 2021; Sharma \textit{et al.}, 2021).

Many research efforts have focused on studying and validating proteins from the SARS-CoV-2 virus as targets for antiviral agents. Among these already validated molecular targets, the inhibition of the 3C-like
2. Material and methods

2.1. Sample preparation and identification of compounds

The horchata infusion used in this study was previously analyzed for its chemical composition (using HPLC-DAD-MS), antioxidant capacity, and inflammatory action (Guevara et al., 2020; 2019). In addition to the previous chemical composition using HPLC-DAD-MS, we also carried out a GC-MS analysis. For its preparation, first, a total of 23 plants (Table 1) were ground to a fine powder in liquid nitrogen using an analytical mill and stored at –20 °C until analysis. A mixture containing equal amounts of the fine powder of each plant was used to prepare the infusion of horchata following the traditional method. First, the water is heated to boiling point, then removed from the heat source. The plants are added to the water and allowed to cool to room temperature. The infusion was filtered through cellulose filter paper (25 μm), concentrated under a vacuum to obtain dried horchata infusion extract, and stored at –20 °C until analysis.

The GC-MS analyses were performed using a GC/MS-QP 2010 Ultra gas chromatograph with an autosampler/injector AOC-20i/s for liquid samples (Shimadzu), and an Ultra Inert DB-5MS UI fused silica capillary column (Agilent). Helium gas was used as a carrier gas at a constant flow rate of 1.0 mL/min. Samples (2 μL) were injected at 250 °C in a split mode 1:10. The GC temperature program was as follows: initial temperature of 150 °C, ramp to 260 °C at a rate of 10 °C/min, and hold for 2 min; ramp to 280 °C at a rate of 10 °C/min and hold isothermally at 280 °C for 1 min. The mass spectrometer was operated in full scan mode and data were acquired under the following conditions: mass range: 200–600 m/z, scan time: 0.30 s, and solvent delay: 2 min. Compounds present in the horchata extract were detected and cross-checked by comparing their mass spectra fragmentation patterns with the stored database in the mass spectral libraries (NIST 11 database).

2.2. Modeling workflow

The modeling strategy herein employed is based on that described in our previous publications (Diniz et al., 2021; Tejera, Munteanu, López-Cortés, Cabrera-Andrade, & Pérez-Castillo, 2020). Briefly, the compounds under investigation were first docked into the active sites of the SARS-CoV-2 Mpro and PLpro enzymes. Different binding poses were predicted for each ligand-receptor pair. The highest-scoring conformers of each ligand docked into the receptors were selected for additional analyses. These included 10 ns of molecular dynamics simulations per predicted complex and the prediction of the free energies of binding. Finally, the top candidate molecules for the inhibition of the enzymes under investigation were selected and their predicted binding modes were analyzed in detail.

Table 1

| Common Name | Scientific Name |
|-------------|-----------------|
| Ataco       | Amaranthus hybridus |
| Albahaca    | Ocimum basilicum L. |
| Borrajín    | Borago officinalis L. |
| Cedrón      | Alyssia triphylla (L. Hér.) Britton |
| Culantrillo | Adiantum concinnum Humb. & Bonpl. ex Wild. |
| Cola de caballo | Equisetum hyemale Kunth |
| Congona     | Peperomia inaquiformis Ruiz & Pav. |
| Canelo      | Cinnamomum sp. |
| Escanuel    | Aerva sanguinolenta L. (Blume) |
| Stevia      | Stevia rebaudiana (Bertoni) Bertoni |
| Hierba luisa | Cymbopogon citratus (DC.) Stapf |
| Hoja de naranja | Citrus × aurantium L. |
| Llantén     | Plantago major L. |
| Malva esencia | Malva sp. |
| Malva olorosa | Pelargonium odoratissimum (L.) L’Hér. |
| Menta       | Mentha × piperita L. |
| Manzanilla  | Matricaria chamomilla L. |
| Malva blanca | Althaea officinalis L. |
| Oregano dulce | Origanum vulgare L. |
| Pena - pena | Fuchsia lonchocanthus |
| Pimpiñela   | Pimpinella aromatica M. Bieb. |
| Toronjil    | Melissa officinalis L. |
| Violeta     | Viola odorata L. |
2.3. Molecular docking

The crystal structures of the SARS-CoV-2 M\textsuperscript{PRO} (code 6Y2G) and PL\textsuperscript{PRO} (code 7JN2) enzymes were retrieved from the Protein Data Bank database and used for molecular docking calculations. One initial three-dimensional conformer for each compound was generated with OpenEye’s Omega (Hawkins, Skillman, Warren, Ellinong, & Stahl, 2010) and am1-bcc partial atomic charges were added to them with Molcharge (OpenEye Scientific Software, n.d.).

The Gold software was employed for molecular docking calculations (Jones, Willett, Glen, Leach, & Taylor, 1997). These proceeded following the same consensus scoring approach applied in previous publications (Aratojo, Pérez-Castillo, Oliveira, Nunes, & Sousa, 2020). In brief, the experimentally bound ligands were used to define the binding site of each protein. Residues with their side chains pointing toward the binding cavity were considered flexible during docking. The ChemPLP scoring function of Gold was chosen for primary scoring and 30 different binding poses were explored for each ligand, with the search efficiency parameter set to 200%. These 30 predicted binding modes were rescored with the ASP, ChemScore and GoldScore scoring functions of Gold. Finally, the aforementioned consensus scoring approach was applied and the highest-scoring ligand conformations were selected for additional analyses.

2.4. Molecular dynamics simulations and prediction of the free energies of binding

Molecular dynamics (MD) simulations were run for all the ligand-receptor complexes selected from the molecular docking calculations. MD simulations were performed with Amber 18 (Case et al., 2018) following the previously described methodology (Diniz et al., 2021). All complexes underwent the same preparation, minimization, heating, equilibration, and production runs stages using an explicit solvent. Five simulations of 2 ns were run per complex, each one initialized with different random initial velocities. The free energies of binding were estimated with the MM-PBSA method, as implemented in Amber 18. A total of 100 MD snapshots were evenly extracted from the five production runs in order to estimate the free energies of binding.

2.5. Protein target predictions and consensus analysis

For each molecule, all their possible target proteins were predicted using the following algorithms (available online): MolTarPred (Peón, Naulaerts, & Ballester, 2017), Swisstarget Prediction (Daia, Michelin, & Zoete, 2019), Targetnet Schdd (Yao et al., 2016), Targetnet Schdd Ensemble (Yao et al., 2016), RF QSAR (Lee, Lee, & Kim, 2017), and PPB2 (Awaie & Reynold, 2019). With PPB2, we used the following algorithms: PPB2 - Extended Connectivity fingerprint ECP4p NN, PPB2 - Shape and Pharmacophore fingerprint Xfp NN, PPB2 - Molecular Quantum Numbers MQN NN, PPB2 - Extended Connectivity fingerprint ECP4 NNNB, PPB2 - Shape and Pharmacophore fingerprint Xfp NNNB, PPB2 - Molecular Quantum Numbers MQN NNNB, PPB2 - Extended Connectivity fingerprint ECP4 pNB and PPB2 - Extended Connectivity fingerprint DNN. Therefore, a total of 13 algorithms were used to predict the possible target proteins of each molecule.

In order to rank the relevance of the targets, we combined all target predictions using a consensus strategy previously used in other works (Tejera et al., 2020). The goal was to obtain a consensus between all methods used for target predictions. In brief, the consensus strategy consisted of the following:

1) For each compound “i”, all predicted scores for all proteins were normalized in each individual method. The final interaction score between the compound “i” and the protein “j” (S\textsubscript{i}) resulted from the average of the normalized value across all methods.

2) For each target, we computed the number of drugs predicted to interact with it (F\textsubscript{j}).

3) We computed the final score (FS\textsubscript{j}) for each target “j” as: FS\textsubscript{j} = \sqrt{\frac{1}{M} \sum_{i} S_{ij}}, where M is the number of molecules. This formulation basically corresponds with the geometrical means between the average score of all compounds interacting with a particular protein and the normalized number of compounds predicted to interact with the same protein “j”.

This final index (FS\textsubscript{j}) will be higher if a higher number of compounds interact with the same protein and if the average drug-protein interaction score is also high.

2.6. Enrichment analysis and integral metabolic network

Predicted protein targets were used for the enrichment analysis of biological processes using the Gene Ontology database (Carbon et al., 2021) and metabolic pathways using KEGG (Kanehisa, Sato, Furumichi, Morishima, & Tanabe, 2019). The enrichment analysis was carried out using David Bioinformatics Resources 6.8 (Huang, Sherman, & Lempicki, 2009). An adjusted p-value < 0.05, using the false discovery rate (FDR), was considered for statistical significance.

3. Results

From the previous analysis of the horchata extract using HPLC-DAD-MS, up to 41 compounds were identified (Guevara et al., 2020). From the 41 compounds, 34 had a direct chemical formulation while in the other 7, it was not possible to clearly establish the position of certain glycoside substituents. These 34 molecules were incorporated into this study. In addition to the previous analysis, we also included 7 molecules identified by the GC-MS analysis.

The mass spectra information from the compounds analyzed by GC-MS were compared to the stored database from NIST 11 mass library. Compound 45 with molecular ion at 208 m/z showed the fragmentation pattern in 134 m/z, 121 m/z and 177 m/z, identified as benzenebutanoic acid, 4-methoxy-, methyl ester in NIST 11 with 97% of similarity. Compound 46 with molecular ion at 329 m/z was fragmented in 200 m/z, 144 m/z, 100 m/z and 57 m/z. This fragmentation was identified as 1-Leucine, N-isobutoxycarbonyl-N-methyl-, heptyl ester with 92% of similarity. Compound 47 showed a molecular ion at 258 m/z, further fragmented in 243 m/z and 43 m/z. This pattern was matched to 7-Acetyl-1-ethyl-1,1,4,4-tetramethyltetralin with 89% of similarity. A molecular ion at 210 m/z corresponding to compound 48 was fragmented into 149 m/z, 136 m/z, 109 m/z, and 164 m/z. This pattern was identified as benzoic acid, 2-(ethylthio)-, ethyl ester with 85% of similarity. Compound 49, with molecular ion at 258 m/z, was fragmented into 243 m/z, 117 m/z, 73 m/z, and 129 m/z. Compound 50, with molecular ion at 272 m/z, was fragmented into 229 m/z and 169 m/z. Compound 51, with molecular ion at 260 m/z, was further fragmented into 205 m/z, 217 m/z, and 245 m/z. These compounds were identified as: undecanoic acid, trimethylsilyl ester (49), 1H- beta.-Carboline-3-carboxylic acid, 1-isopropyl-2,3,4,9-tetrahydro-, methyl ester (50), and Isopecunin (51), with similarities of 82, 85 and 75% respectively.

The full list of the 34 molecules previously identified by HPLC-DAD-MS in horchata extract, as well as the 7 identified by GC-MS, are listed in Table 2. In addition to these molecules, we included in the analysis a possible modification (i.e. o-glycoside rupture), which lead to a total of 51 compounds, as presented in Table 2.

3.1. Docking and molecular dynamics

The detailed results of the consensus docking of the 51 compounds to the SARS-CoV-2 M\textsuperscript{PRO} and PL\textsuperscript{PRO} proteases are provided in Tables TS1 and TS2 (supplementary material SM1). Compound 49 was excluded from our analyses due to parametrization problems in Amber 18 (see below), resulting from the presence of a silicon atom. A total of 102 and 100 valid docking poses to M\textsuperscript{PRO} and PL\textsuperscript{PRO} were predicted for the remaining
been shown that the post-processing of docking predicted complexes employing free energy calculations based on MD data can improve the accuracy performance of virtual screening strategies (Poli, Granchi, & Tuccinardi, 2020). To get further insights into the potential binding of the studied molecules to both receptors, the 202 predicted complexes were subjected to MD simulations and their free energies of binding were predicted with the MM-PBSA method from the MD-derived snapshots. The results of these calculations are presented in Tables TS3 and TS4 (supplementary material SM1) for the Mpro and PLpro enzymes, respectively, and summarized in Fig. 1.

Fig. 1 shows that the studied compounds present a better binding profile to PLpro (Fig. 1B) than to Mpro (Fig. 1A). Moreover, the predicted free energies of binding of five compounds to Mpro are lower than ~5 kcal/mol, while the same amount of chemicals achieves the same criterion for PLpro. Notably, compound 47 was predicted among the top five candidate inhibitors for both targets and hence it can be considered as a potential dual Mpro-PLpro inhibitor. According to the obtained results, compounds 48, 46 and 31 were selected as the most probable inhibitors of the Mpro enzyme. On the other hand, chemicals 30, 26 and 40 were predicted as the most probable inhibitors of the SARS-CoV-2 PLpro enzyme. The coordinates of the predicted complexes of compounds 48, 46 and 31 with Mpro and of chemicals 30, 26 and 40 with PLpro are provided as supplementary materials (SM3-SM8) in PDB format.

The predicted binding modes of 48, 46 and 31 to Mpro, as well as the diagrams of interactions between the ligands and the receptor, are represented in Fig. 2. The complex snapshot selected for depiction corresponds to the centroid of the most populated cluster, resulting from clustering the 100 MD snapshots employed for MM-PBSA calculations. The molecular graphics and analyses were performed with UCSF Chimera (Pettersen et al., 2004) and the interaction diagrams were depicted with LigPilot+ (Laskowski & Swindells, 2011), while the networks of interactions were analyzed with Cytoscape (Shannon et al., 2003). Only interactions observed in at least 50% of the MD snapshots used for free energy calculations are represented in Fig. 2.

Interestingly, the top candidate for Mpro inhibition (48) is predicted to form no hydrogen bonds with the receptor. In this complex, the compound is predicted to bind mainly at the S2 subpocket of the enzyme (see Zhang et al., 2020) for the definition of the Mpro pocket regions), with its aromatic ring buried within this region. The phenyl ring of compound 48 is predicted to interact with C44, M49, S46, Y54, and D187, while its ethanethiol moiety makes contacts with C44, T45, and S46. Furthermore, the ethyl formate substituent of 48 projects from S2 to part of the S3 region of Mpro, positioning in front of the catalytic H41 residue.

On the other hand, compound 46 is predicted to span the S1 and S3 subsites of the enzyme, with its isobutyl formate group blocking the access to S2 through interactions with T25. The carbonyl-methylamino group of compound 46 is predicted to form a hydrogen bond with the receptor. In this complex, the compound is predicted to bind mainly at the S2 subpocket of the enzyme. On the other hand, chemicals 30, 26 and 40 are predicted to form free energy calculations are represented in Fig. 2.

Note: (*) Because it is possible the hydrolysis of the glycoside part of several of those molecules (especially in the acidic gastric environment), we decided to consider those molecules (especially in the acidic gastric environment), we decided to consider separately. (**) Molecules identified by GC-MS analysis.

50 compounds, respectively. In all cases, the predicted binding conformations were inside the receptors’ active sites or orientated to block the access to the catalytic residues. According to the same consensus scoring protocol followed for the selection of the docking poses, compounds 10, 25 and 41 were predicted as the best candidates for Mpro binding. In the same way, chemicals 40, 39 and 25 were the highest-scoring docking molecules for PLpro binding.

Molecular docking algorithms are designed for the conformational exploration and scoring of large amounts of molecules in a short space of time. This calculation speed comes at the cost of simplifying and neglecting important factors involved in molecular recognition. It has
divided into four subsites: S1-S4 (Gao et al., 2021). Compound 30 is predicted to orient one of its oxane rings toward S4, a positioning that is stabilized by hydrogen bonds to the side chain of D164, R166, Y268, Y273 and D302, and additional contacts with S245, A246, and T301. The rest of the compound is exposed to the solvent, with the chromone ring stacking parallel to Y268 and the methoxyphenyl ring interacting with P247 and P248, and hydrogen bonding to the later amino acid.

Compound 26 is predicted to bind PLpro perpendicular to the predicted orientations of 30 and 40, spanning the S3 and S4 regions. This chemical is predicted to hydrogen bond the side chain of R166 and the backbone of Q269. The chromone nucleus of 26 settles between R166, P247, P248, and Y268, while its methoxybenzyl substituent interacts with Q269, Y264, Y268, and Y273. Finally, compound 40 is predicted to orient on PLpro similarly to 30, with its hydroxyphenyl moiety occupying the S4 subpocket and the rest of it either interacting with the receptor surface or completely exposed to the solvent. The anchoring at S4 is stabilized by hydrogen bonds to the side chains of R166 and D302, and interactions with D164, V165, M208, A246, Y273, and T301. In this case, the chromone ring locates itself between P247, P248, and Y268.

3.2. System biology analysis

A total of 1075 possible protein targets were predicted in Homo sapiens (see SM2, supplementary materials). However, only the top 150 proteins were used for further analysis (around 14% of the total). The selection of these proteins was conducted based on the fact that most of them were predicted to interact with the compounds under study with a high probability across almost all methods used. Moreover, beyond that cut-off point, the decrement in the consensus score starts to decrease almost linearly. The enriched biological processes (Table 3) clearly indicate an emphasis on the oxidation–reduction process, as well as oxidative stress and the involvement of one-carbon metabolism. It is connected with the carbonic anhydrase enzymes family that were identified as potential targets. In fact, from the targets obtained through consensus analysis, the CA12, CA12, CA4, CA9, CA14 and CA6 were obtained in the top 10 candidates (full list in SM2).

Besides the biological processes involved in oxidative stress and carbon metabolism, those involved in inflammation and even immune response are also relevant. The proteins: CYP19A1, MIF, PGH1, ESR1, ESR2, ALDR1, TLR9 were also found ith high score in the target consensus analysis. Some of the biological processes found are also clearly presented in the enriched metabolic pathway analysis (Table 3).

A selection of the enriched metabolic pathways is presented in Table 3. The full list can be consulted in SM2 (supplementary material). The comparative analysis between biological processes and metabolic pathways shows the involvement of the arachidonic pathways and their
Fig. 2. Predicted binding modes of compounds 48, 46 and 31 to the SARS-CoV-2 M\(^{\text{pro}}\) enzyme. In the lefthand figures, oxygen atoms are depicted as red, nitrogen atoms blue, sulfur atoms yellow, and the carbon atoms of the ligands orange. The same coloring scheme was applied to the interaction diagrams (right), while carbon atoms are colored black on them. These diagrams depict all atoms for the receptor residues forming hydrogen bonds with the ligands. Only receptor residues interacting with the ligands in at least 50% of the analyzed MD snapshots are labeled in the lefthand figures and represented in the interaction diagrams. The compounds are: 7-Acetyl-6-ethyl-1,1,4,4-tetramethyltetralin (48), l-Leucine, N-isobutoxycarbonyl-N-methyl-, heptyl ester (46), and Isorhamnetin (31). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 3. Predicted binding modes of compounds 30, 26 and 40 to the SARS-CoV-2 PLₜₚ enzymo. In the lefthand figures, oxygen atoms are colored red, nitrogen atoms blue, sulfur atoms yellow, and the carbon atoms of the ligands orange. The same coloring scheme was applied to the interaction diagrams (right), while carbon atoms are colored black on them. These diagrams depict all atoms for the receptor residues forming hydrogen bonds with the ligands. Only receptor residues interacting with the ligands in at least 50% of the analyzed MD snapshots are labeled in the lefthand figures and represented in the interaction diagrams. The compounds are: Isorhamnetin-3-O-(6-O-rhamnosyl-galactoside) (30), 5,7-dihydroxy-4'-methoxyflavanone (26), and (2S,3S,4S,5R,6S)-6-[2-(3,4-dihydroxyphenyl)-5-hydroxy-4-oxochromen-7-yl]oxy-3,4,5-trihydroxyoxane-2-carboxylic acid (40). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Table 3 Selected enriched biological processes and metabolic pathways.

| GO-ID | GO-Name | FDR |
|-------|---------|-----|
| GO:0006730 | one-carbon metabolic process | 9.47E-11 |
| GO:0015701 | bicarbonate transport | 9.47E-11 |
| GO:0006954 | inflammatory response | 9.09E-09 |
| GO:0043066 | negative regulation of apoptotic process | 3.26E-08 |
| GO:0030168 | platelet activation | 1.45E-06 |
| GO:0043401 | steroid hormone mediated signaling pathway | 6.25E-06 |
| GO:0043406 | positive regulation of MAP kinase activity | 7.51E-06 |
| GO:0055114 | oxidation-reduction process | 8.99E-06 |
| GO:0048010 | vascular endothelial growth factor receptor signaling pathway | 2.41E-05 |
| GO:0000187 | activation of MAPK activity | 3.93E-05 |
| GO:0050852 | T cell receptor signaling pathway | 3.47E-04 |
| GO:00010507 | negative regulation of autophagy | 8.33E-04 |
| GO:0043065 | positive regulation of apoptotic process | 0.002795 |

PATH-ID Pathway Name FDR

Hsa00910 | Nitrogen metabolism | 1.10E-13 |
Hsa04071 | Sphingolipid signaling pathway | 2.66E-07 |
Hsa04080 | Neuroactive ligand-receptor interaction | 1.27E-06 |
Hsa04917 | Prolactin signaling pathway | 1.75E-06 |
Hsa04370 | VEGF signaling pathway | 3.07E-06 |
Hsa04660 | T cell receptor signaling pathway | 4.25E-06 |
Hsa04722 | Neurotrophin signaling pathway | 4.25E-06 |
Hsa04024 | cAMP signaling pathway | 9.13E-06 |
Hsa00140 | Steroid hormone biosynthesis | 9.71E-06 |
Hsa04151 | P53-Akt signaling pathway | 6.66E-05 |
Hsa04621 | NOD-like receptor signaling pathway | 3.21E-05 |
Hsa04923 | Regulation of lipolysis in adipocytes | 3.21E-05 |
Hsa04068 | FoxO signaling pathway | 7.74E-04 |
Hsa04662 | B cell receptor signaling pathway | 0.00101 |
Hsa04014 | Ras signaling pathway | 0.001144 |
Hsa04510 | Focal adhesion | 0.001551 |
Hsa04915 | Estrogen signaling pathway | 0.00196 |
Hsa04015 | Rap1 signaling pathway | 0.001716 |
Hsa04150 | mTOR signaling pathway | 0.00195 |
Hsa00590 | Arachidonic acid metabolism | 0.002496 |
Hsa04064 | NF-kappaB signaling pathway | 0.003018 |

Note: 1) gene ontology identification number, 2) gene ontology name for the biological process, 3) adjusted p-value of the enrichment analysis using the false discovery rate (FDR), 4) KEEG pathway identification number and 5) pathway name.

4. Discussion

From the results obtained from the docking and MD analysis, we focused on three molecules that are potential inhibitors of the SARS-CoV-2 M\textsuperscript{pro} and three other potential inhibitors of PL\textsuperscript{pro} enzymes. Altogether, our results suggest that natural compounds contained in the herbs used to produce horchata have the potential of providing compounds capable of inhibiting the SARS-CoV-2 virus by targeting its M\textsuperscript{pro} and PL\textsuperscript{pro} enzymes. As mentioned before, the top three candidate compounds for PL\textsuperscript{pro} inhibition are chromone derivatives. One chromone is also predicted as a potential inhibitor of M\textsuperscript{pro}. The chemical variations of natural chromone are wide, as seen in our own results. However, several of these natural chemical diverse chromones have been previously described as antiviral agents (Hu et al., 2012). We consider that the chromones contained within these plants deserve, in particular, more attention and experimentation as potential SARS-CoV-2 inhibitors.

In our predicted molecules,isorhamnetin (31) had been previously found to be active against influenza virus A/PR/8/34(H1N1) (Doyem, Choi, Kim, & Cho, 2015), and other authors found it in the glycosylated form (similar to form 30 and 32) through computational analysis (docking without molecular dynamics) as a potential inhibitor of the SARS-CoV-2 M\textsuperscript{pro} enzyme (Owis et al., 2020). Similarly, through docking analysis, isorhamnetin had also been proposed to bind M\textsuperscript{pro} enzyme of the SARS-CoV-2 virus (J. Xu, Gao, Liang, & Chen, 2021) by other authors. To the best of our knowledge, other molecules herein proposed, had not been previously reported elsewhere as potential inhibitors of M\textsuperscript{pro} and PL\textsuperscript{pro}.

SARS-CoV-2 infection is related to a release of pro-inflammatory cytokines and uncontrolled inflammation that induce the accumulation of intra-alveolar fibrin, leading to pulmonary damage. Moreover, in the progression of COVID-19, the involvement of oxidative stress is also well-known (Beltran-Garcia et al., 2020). Therefore, not only a virus-targeted strategy is desirable but anti-inflammatory and antioxidant effects are also desirable for treating COVID-19. The system biology analysis clearly indicates the possible action of horchata on the inflammation and immune response’s biological processes and pathways. The anti-inflammatory and antioxidant effects of horchata had been previously described (Guevara et al., 2020; 2019), The involvement of one-carbon and bicarbonate metabolisms are consequences of the involvement of potential interactions with several carbonic anhydrase proteins. The inhibition of these proteins by phenolic compounds (such as those present in horchata) regulating oxidative stress is well known (Gulcin & Beydemir, 2012), as is its possible use in COVID-19 treatment (Lammi & Arnoldi, 2021).

Concerning our findings of the enrichment of biological processes related to inflammation response, a closer look points toward the possible interaction with PTGDR (prostaglandin D2 receptor), PTG51 (prostaglandin G/H synthase 1), and PL2G1B (Phospholipase A2). These findings are consistent with previous research regarding several dietary phenolic compounds (Yoon & Baek, 2005), and also consistent with a possible non-steroidal anti-inflammatory effect. Molecules like isorhamnetin, quercetin, luteolin, ferulic acid and others which are major components of horchata are also predominant in the Reduning injection, which was recommended in China for the treatment of pneumonia induced by COVID-19 (X. Xu et al., 2021). In addition, a similar list of compounds (also including acacetin (26) and kaempferol) can be found in other traditional medicines used in China to treat COVID-19-induced pneumonia (Luo et al., 2020).

It must be stressed that our results point to the potential inhibition of the SARS-CoV-2 M\textsuperscript{pro} and PL\textsuperscript{pro} enzymes by compounds present in the horchata infusion. Further research is required to test our inhibition hypotheses, which could lead to the identification of hits compounds capable of entering the drug development pipeline. This means that any experimentally confirmed inhibitor will not automatically become a drug for the treatment of COVID-19. Instead, these hit molecules must undergo a process whereby their inhibitory activity and pharmacokinetic properties must be jointly optimized by performing structural modifications until a desirable drug candidate is obtained.

5. Conclusions

The potential use of horchata infusion as a source of inhibitors of SARS-CoV-2 enzymes has been explored in our presented work by combining the chemical characterization of the herbal infusion and several computational approaches. We found a total of six potential inhibitors: 7-Acetyl-6-ethyl-1,1,4,4-tetramethyltetralin (48), l-Leucine, N-isobutoxycarbonyl-N-methyl-, heptyl ester (46) and Isorrhamnetin (31) for the M\textsuperscript{pro} enzyme, and isorrhamnetin-3-O-(6-O-rhamnosyl-galactoside) (30), 5,7-di-hydroxy-4-methoxyflavanone (26) and (2S,3S,4S,5S,6S)-6-[2-(3,4-Dihydroxyphenyl)-5-hydroxy-4-oxochromen-7-yl]oxy-3,4,5-trihydroxyxane-2-carboxylic acid (40) for the PL\textsuperscript{pro} enzyme. Our analysis of all possible targets interacting with the identified molecules also points toward the possible anti-inflammatory and antioxidant effect of the compounds contained in the herbal infusion. These effects could be mediated by interaction with several targets involved in the arachidonic acid metabolism and the carbonic anhydrase enzymes. Even when our results support the hypothesis that horchata could provide a source of important bioactive compounds that could enter the drug development pipeline against COVID-19, further
experimental studies are needed to validate our findings and to elucidate possible mechanisms of action.

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

Acknowledgements
This study was funded by the Universidad de Las Américas (UDLA) (grant number AGRJAVA.18.05), Quito, Ecuador. Access to plant genetic resources was granted by means of the “FRAMEWORK AGREEMENT FOR ACCESS TO PLANT GENETIC RESOURCES: MAE-DNB-CM-2017-0072” celebrated between the Environment Ministry of Ecuador (MAE) and the Universidad de Las Américas (UDLA), Ecuador. We acknowledge financial assistance from the Universidad San Francisco de Quito through POLI grant 11162 to Valeria Ochoa-Herrera.

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2021.130589.

References
Araújo, M. O., Pérez-Castillo, Y., Oliveira, L. H. G., Nunes, F. C., & Sousa, D. P. (2020). Larvicidal activity of cinnamic acid derivatives: Investigating alternative products for Aedes aegypti L. control. Molecules, 26(1), 61. https://doi.org/10.3390/ molecules26010061.
Awale, M., & Reymond, J. L. (2019). Polypharmacology browser PP2R: target prediction combining nearest neighbors with machine learning. Journal of Chemical Information and Modeling, 59(1), 10–17. https://doi.org/10.1021/acs.jcim.8b00524.
Beltrán-García, J., Osca-Verdegal, R., Pallardó, F. V., Ferreres, J., Rodríguez, M., Mulet, S., García-Giménez, J. L. (2020). Oxidative stress and inflammation in covid-19-associated sepsis: The potential role of anti-oxidative therapy in avoiding disease progression. Antioxidants, 9(10), 936. https://doi.org/10.3390/antioxid9100936.
Carbon, S., Douglass, E., Good, B. M., Unni, D. R., Harris, L. N., Mungall, C. J., … Ebel, J. (2021). The gene ontology resource: Enriching a Qold mine. Nucleic Acids Research, 49(D1), D225–D334. doi: 10.1093/nar/gkaa1113.
Case, D. A., Ben-Shalom, T. Y., Brozell, S. R., Gerutti, D. S., Cheatham III, T. E., Cuzziere, V. W. D., … Gilson, M. K. (2018). Amber 2018. 2018. University of California, San Francisco. p. 1. Retrieved from http://ambermd.org/.
Chen, X., Wu, Y., Chen, C., Gu, Y., Zhu, C., Wang, S., … Wu, C. (2021). Identifying potential anti-COVID-19 pharmacological compounds of traditional Chinese medicine Liuqianwengying capsule based on human exposure and ACE2 biochromatography screening. Acta Pharmaceutica Sinica B, 11(2), 222–236. https://doi.org/10.1016/j.apsb.2020.10.002.
CSSE Johns Hopkins. (2021). COVID-19 Map – Johns Hopkins coronavirus resource center. Retrieved March 11, 2021, from Johns Hopkins Coronavirus Resource Center website: https://coronavirus.jhu.edu/map.html.
Daina, A., Michielin, O., & Zotei, V. (2019). SwissTargetPrediction: Updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Research, 47(W1), W357-W364. https://doi.org/10.1093/nar/gkx382.
Deymen, A. A., Choi, H. Y., Kim, Y. B., & Cho, S. G. (2015). Antiviral effect of methylated flavonol isorhamnetin against influenza. PLoS ONE, 10(3), e0121610. https://doi.org/10.1371/journal.pone.0121610.
Diniz, L. R. L., Perez-Castillo, Y., Elshabrawy, H. A., Bezerra Filho, C. da S. M., & de Sousa, D. P. (2021). Article bioactive terpenes and their derivatives as potential SARS-CoV-2 protease inhibitors from molecular modeling studies. Biomolecules, 11 (1), 1–19. https://doi.org/10.3390/biom11010074.
Gao, X., Qin, B., Zhou, G., Liao, X., You, Z., Wang, F., … Zhu, K. (2020). Crystal structure of SARS-CoV-2 papain-like protease. Acta Pharmaceutica Sinica B, 11(1), 237–245. https://doi.org/10.1016/j.apsb.2020.08.014.
Guevara, M., Pintoa, A., Tejera, E., Ballesteros, I., Sánchez, M. E., Granja-Albuja, M. G., … Alvarez-Sanchez, J. M. (2020). Protective effect of the medicinal herb infusion “horchata” against oxidative damage in cigarette smokers: An ex vivo study. Food and Chemical Toxicology, 143, 111538. https://doi.org/10.1016/j.fct.2020.111538.
Guevara, M., Tejera, E., Inirralde, G. A., Jaramillo-Vivanco, T. G., Granja-Albuja, M. G., Granja-Albuja, S., … Alvarez-Sanchez, J. M. (2019). Anti-inflammatory effect of the medicinal herb mixture infusion, Horchata, from southern Ecuador against LPS-induced cytokotoxicity in macrophages. Food and Chemical Toxicology, 131, 110594. https://doi.org/10.1016/j.fct.2019.110594.
Gulcin, I., & Beydjem, S. (2013). Phenolic Compounds as antioxidants: Carbonic anhydrase isoenzymes inhibitors. Mini Reviews in Medicinal Chemistry, 13(3), 408–430. https://doi.org/10.2174/13895571380499874.
Xu, J., Gao, L., Liang, H., Chen, & Dong, S. (2021). In silico screening of potential anti-COVID-19 bioactive natural constituents from food sources by molecular docking. *Nutrition*, 82, Article 111049. https://doi.org/10.1016/j.nut.2020.111049.

Xu, X., Zhang, J., Zheng, W., Yang, Z., Zhao, X., Wang, C., … Liu, Q. (2021). Efficacy and safety of reduning injection in the treatment of COVID-19: A randomized, multicenter clinical study. *Annals of Palliative Medicine*, 10(5), 5146–5155.

Yao, Z. J., Dong, J., Che, Y. J., Zhu, M. F., Wen, M., Wang, N. N., … Cao, D. S. (2016). TargetNet: A web service for predicting potential drug-target interaction profiling via multi-target SAR models. *Journal of Computer-Aided Molecular Design*, 30(5), 413–424. https://doi.org/10.1007/s10822-016-9915-2.

Yoon, J. H., & Baek, S. J. (2005). Molecular targets of dietary polyphenols with anti-inflammatory properties. *Yonsei Medical Journal*, 46(5), 585–596. https://doi.org/10.3349/ymj.2005.46.5.585.

Zhang, L., Lin, D., Sun, X., Carlb, U., Drosten, C., Sauerhering, L., … Hilgenfeld, R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors. *Science*, 368(6489), 409–412. https://doi.org/10.1126/science.abb3405.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., … Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, 382(8), 727–733. https://doi.org/10.1056/NEJMoa2001017.