



OPEN

## Statine-based peptidomimetic compounds as inhibitors for SARS-CoV-2 main protease (SARS-CoV-2 Mpro)

Pedro Henrique R. de A. Azevedo<sup>1</sup>, Priscila G. Camargo<sup>2</sup>, Larissa E. C. Constant<sup>3</sup>, Stephany da S. Costa<sup>3</sup>, Celimar Sinézia Silva<sup>3</sup>, Alice S. Rosa<sup>4,5</sup>, Daniel D. C. Souza<sup>4,5</sup>, Amanda R. Tucci<sup>4,5</sup>, Vivian N. S. Ferreira<sup>4</sup>, Thamara Kelcya F. Oliveira<sup>4,5</sup>, Nathalia R. R. Borba<sup>4</sup>, Carlos R. Rodrigues<sup>2</sup>, Magaly G. Albuquerque<sup>6</sup>, Luiza R. S. Dias<sup>1</sup>, Rafael Garrett<sup>6</sup>, Milene D. Miranda<sup>4,5</sup>, Diego Allonso<sup>3,7</sup>, Camilo Henrique da S. Lima<sup>6</sup>✉ & Estela Maris F. Muri<sup>1</sup>✉

COVID-19 is a multisystemic disease caused by the SARS-CoV-2 airborne virus, a member of the *Coronaviridae* family. It has a positive sense single-stranded RNA genome and encodes two non-structural proteins through viral cysteine-proteases processing. Blocking this step is crucial to control virus replication. In this work, we reported the synthesis of 23 statine-based peptidomimetics to determine their ability to inhibit the main protease (Mpro) activity of SARS-CoV-2. Among the 23 peptidomimetics, 15 compounds effectively inhibited Mpro activity by 50% or more, while three compounds (7d, 8e, and 9g) exhibited maximum inhibition above 70% and  $IC_{50} < 1 \mu M$ . Compounds 7d, 8e, and 9g inhibited roughly 80% of SARS-CoV-2 replication and proved no cytotoxicity. Molecular docking simulations show putative hydrogen bond and hydrophobic interactions between specific amino acids and these inhibitors. Molecular dynamics simulations further confirmed the stability and persisting interactions in Mpro's subsites, exhibiting favorable free energy binding ( $\Delta G_{bind}$ ) values. These findings suggest the statine-based peptidomimetics as potential therapeutic agents against SARS-CoV-2 by targeting Mpro.

Coronavirus disease-19 (COVID-19) emerged in late 2019 as a massive pandemic with catastrophic effects on people's lives, accounting for more than 770 million cases and over 7 million deaths worldwide<sup>1</sup>. It is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), an enveloped airborne virus member of the *Coronaviridae* family (genus *Betacoronavirus*), which also contains other life-threatening viruses, such as SARS-CoV-1 and Middle East respiratory syndrome coronavirus (MERS-CoV)<sup>2</sup>. The genome of SARS-CoV-2 has a positive-sense single-strand RNA of 27–32 kb-long that encodes for two open reading frames (ORF1a and ORF1b), which are translated into large polyproteins (Pp1a and Pp1ab) further processed into 16 non-structural proteins (nsps) by two virus-encoded cysteine proteases: the chymotrypsin-like protease (3CLpro), also called main protease (Mpro), responsible for polyprotein cleavage at 11 sites, generating 13 nsps, and the papain-like protease (PLpro), which cleaves the polyprotein at three other sites, generating 4 nsps<sup>3</sup>. Pp1a and Pp1ab processing is crucial on the virus replicative cycle and, consequently, for proper virus infection. Therefore, blockage of polyprotein processing is an outstanding strategy to control virus replication. Mpro is an excellent

<sup>1</sup>Laboratório de Química Medicinal, Faculdade de Farmácia, Universidade Federal Fluminense, Niterói, RJ 24241-000, Brazil. <sup>2</sup>Faculdade de Farmácia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21941-853, Brazil. <sup>3</sup>Laboratório de Biotecnologia e Bioengenharia Tecidual, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21941-853, Brazil. <sup>4</sup>Laboratório de Morfologia e Morfogênese Viral, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, RJ 21040-900, Brazil. <sup>5</sup>Programa de Pós-Graduação em Biologia Celular e Molecular, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, RJ 21040-900, Brazil. <sup>6</sup>Programa de Pós-Graduação em Química, Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21941-853, Brazil. <sup>7</sup>Departamento de Biotecnologia Farmacêutica, Faculdade de Farmácia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21941-853, Brazil. ✉email: camilolima@iq.uff.br; estelamuri@id.uff.br

pharmacological target due to its relevance to the virus life cycle and the absence of a homologous enzyme in humans. Nowadays, there are several available vaccines against SARS-CoV-2 in the world, which exerted and still does an essential role for controlling severe cases and death. However, they failed to prevent mild and moderate infection, which is managed using antiviral or symptoms-relieving drugs. Despite the quick development of several vaccines, only a few antivirals targeting SARS-CoV-2 are available in the market, such as nirmatrelvir-ritonavir combination, molnupiravir, and remdesivir<sup>4</sup>, and most of them are not indicated for mild and moderate cases, which makes the development of new drugs urgent.

The scientific community has been trying to find active molecules against SARS-CoV-2 using strategies such as repurposing of existing drugs, bioinformatics, pharmacoinformatic approaches, and others<sup>5</sup>. One of the most important classes of molecules studied is the peptidomimetic compounds, including different analogs such as keto-glutamine, trifluoromethyl ketone,  $\alpha,\beta$ -unsaturated esters, and  $\alpha$ -keto amides<sup>6–11</sup>.

A class of unusual (non-canonical) amino acids<sup>12</sup> known as  $\gamma$ -amino- $\beta$ -hydroxy acids, such as statine (PubChem CID 123,915; (3*S*,4*S*)-4-amino-3-hydroxy-6-methylheptanoic acid; AHMHA) and the statine-like AHPPA<sup>13</sup> (PubChem CID 11,769,728; (3*S*,4*S*)-4-amino-3-hydroxy-5-phenylpentanoic acid) are mainly present as amino acid residues in several examples of natural peptides (e.g., pepstatin-A, ahpatinin-Ac)<sup>14,15</sup> and synthetic peptidomimetic inhibitors of aspartic proteases<sup>16</sup> and there is evidence that its activity against viral infections occurs by impairing virus protease activity<sup>17,18</sup>. Despite that, and as far as we know, there are no statine-based synthetic compounds described as cysteine inhibitors of SARS-CoV-2 Mpro.

Our research group has focused on protease inhibitors, and we have reported a new class of statine-based peptidomimetics as selective inhibitors of the human kallikrein 5 (KLK5) serine-protease activity<sup>18</sup>.

In the current study we repositioned, synthesized, and investigated our statine-based peptidomimetics as cysteine protease inhibitors against SARS-CoV-2 Mpro, including antiviral activity in cell model assay, cytotoxicity assays, and in silico evaluation.

## Materials and methods

### Chemistry

#### General information

Reagents were purchased from Sigma-Aldrich Co. All solvents were purchased as reagent grade, dried using standard conditions, and stored over molecular sieves. Purification of products was carried out using silica gel flash chromatography (Whatman 60, 230–400 mesh). NMR analyses were performed on a Varian Unity Plus-300 spectrometer. Melting points were obtained on a Thomas Hoover capillary melting point apparatus and are uncorrected. All compounds are >95% pure by high-resolution mass spectra (HRMS) that were performed on a Waters Micromass Q-ToF Micromass spectrometer equipped with a lock spray source. The IR spectra were obtained on a Perkin-Elmer spectrometer model Spectrum One in liquid film or KBr pellets. Optical rotation measurements were determined on a Perkin-Elmer 341 LC polarimeter.

#### General procedure for compounds 7a-h

To a 0 °C cooled mixture of compound **1** (0.2 g; 0.647 mmol) and the appropriate methyl ester hydrochloride (1.15 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (10 mL) were added EDC.HCl (0.186 g; 0.970 mmol), HOBt (0.13 g; 0.970 mmol) and *N*-methylmorpholine (0.21 mL; 1.94 mmol). The mixture was stirred at room temperature for 24 h, and the volatiles were removed under reduced pressure. The resulting residue was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (50 mL) and successively washed with 5% H<sub>3</sub>PO<sub>4</sub> (50 mL), 20% Na<sub>2</sub>CO<sub>3</sub> (50 mL), water (40 mL), and brine (50 mL) and dried with Na<sub>2</sub>SO<sub>4</sub> after which it was filtered and evaporated under reduced pressure. The products were purified by flash chromatography on silica gel using EtOAc/hexane as eluents.

#### General procedure for compounds 8a-g

The corresponding ester **7a-g** (5.0 mmol; 1 eq.) was solubilized in distilled dichloromethane (5.0 mL). Anhydrous pyridine (7.5 mmol; 0.6 mL; 1.5 eq.) and DMAP (5.0 mmol; 1 eq.) were added to the solution. The mixture was cooled to 0 °C and then acetic anhydride (7.5 mmol; 1.5 eq.) was added dropwise. The reaction was stirred at room temperature for 3 h until complete consumption of the starting material. The reaction was diluted with dichloromethane (50 mL) and extracted successively with water (50 mL) and brine (50 mL). The organic phase was dried over anhydrous sodium sulfate and then evaporated on a rotary evaporator. Purification by chromatographic column on silica gel afforded the products as solids.

#### General procedure for compounds 9a-g

The corresponding ester (**7a-h**) (5.0 mmol) was solubilized in a minimum volume of distilled dichloromethane. The resulting solution was cooled to 0 °C and then trifluoroacetic acid (7.50 mmol, 0.57 mL, 1.5 eq) was added dropwise. The reaction was stirred at room temperature for 3 h until complete consumption of the starting material and then completely evaporated in a rotary evaporator. The resulting product was purified by recrystallization from diethyl ether.

## Biological assays

#### In vitro inhibition of SARS-CoV-2 Mpro

Recombinant SARS-CoV-2 Mpro synthetic gene expressed in *E. coli* BL21(DE3)pLysS cells were used in a fluorescent resonance energy transfer (FRET) assay, using as substrate the DABCYL-AVLQ↓SGFRLL-EDANS peptide (Biomatik Corp., CA), and as positive control a synthetic dipeptide covalent inhibitor of SARS-CoV's Mpro, **GC-376** (PubChem CID 71,481,119). The enzyme concentration was fixed at 1.5  $\mu$ M, the substrate at 50  $\mu$ M and the compounds (statine-based peptidomimetics and **GC-376**) ranged from 0.001 to 1000  $\mu$ M. The enzyme

and compounds were incubated in 5 mM NaCl, 20 mM Tris.HCl pH 8.0, 5 mM DTT for 15 min at 37°C before starting with the substrate. The emission fluorescence of EDANS was monitored in the following parameters:  $\lambda_{exc} = 330$  nm,  $\lambda_{em} = 490$  nm, at 37 °C for 45 min. Fluorescence data (RFU) was converted into substrate cleavage-specific activity using fluorescent conversion factor (FEC) previously calculated based on the EDANS-DABCYL fluorophore pair. Maximum enzyme activity was considered in the situation with vehicle (DMSO), and the values were used to calculate the enzyme inhibition by the compounds. The concentration that inhibits 50% of the enzyme activity (i.e., the half-maximal inhibitory concentration,  $IC_{50}$ ) was calculated in the software GraphPad Prism 9.0.

#### *Antiviral and cytotoxicity assays*

We state that human participants are not involved in the study. We evaluated the compounds' biological activity in a cell model using Calu-3 cells, a human submucosal gland cell line. Calu-3 cell models are widely used as a preclinical model for respiratory disease drug screening due to their bronchiolar epithelium characteristics and ability to replicate viruses with higher titles, including SARS-CoV-2<sup>19–22</sup>.

The cytotoxicity assay consisted of interaction between the compounds **7d**, **8e**, and **9g**, at different concentrations (200, 100, 50, 25, and 12.5  $\mu$ M) with Calu-3 cells (kindly donated by the Farmanguinhos platform RPT11M) at a cell density of  $1 \times 10^4$  cells/well for 72 h. Afterward, the cells were submitted to viability evaluation by methylene blue assay. For this assay, cells were washed with PBS 1 $\times$  and stained with methylene blue solution (Hanks' solution (HBSS), 1.25% glutaraldehyde, and 0.6% methylene blue) for 1 h. Then, the cells were rewashed and elution solution (50% ethanol, 49% PBS 1 $\times$ , and 1% acetic acid) was added for 15 min. After that time, the absorbance was read at 660 nm in the spectrophotometer.

The antiviral effect was analyzed by Plaque Forming Units (PFU) assay. For that, Calu-3 cells ( $1 \times 10^4$  cells/well) infected with SARS-CoV-2 B.1 lineage isolate (GenBankMT710714, SisGen AC58AE2) at MOI 0.01 during 1 h at 37 °C and 5% CO<sub>2</sub>, were treated with the compounds **7d**, **8e**, and **9g**. The treatment was carried out with a semi-log curve of concentration (10, 3.16, 1, 0.316, and 0.1  $\mu$ M) for 24 h. Then, the supernatants were harvested for virus titer determination. Vero E6 cells (African green monkey kidney, ATCC CRL-1586) in a 96-well plate in  $1 \times 10^4$  cells/well density were incubated with different dilutions (1:100–1:12,800) of supernatants for 1 h at 37 °C and 5% CO<sub>2</sub>. After this period, carboxymethylcellulose medium (DMEM-Higher 10 $\times$ , 2.4% carboxymethylcellulose, and 2% fetal bovine serum) was added to the well at a ratio 1:1, and the cells were cultured for 72 h. Afterward, cells were fixed with formalin 5% for 3 h and posteriorly stained with crystal violet 0.04% for 1 h. The viral titers were determined by PFU/mL.

All the compounds used in vitro assays were resuspended in 100% DMSO (dimethyl sulfoxide), aliquoted, and stored at  $-20$  °C to avoid compound degradation<sup>23</sup>. The DMSO final concentrations for each molecule's test concentration were equal or lower than 1% (v/v) diluted in DMEM (Dulbecco's Modified Eagle Medium) not affecting the growth of the cells<sup>24,25</sup>. According to WHO guidelines<sup>26</sup>, all virus manipulation was realized at a biosafety level 3 (BSL3) multiuser facility.

#### *Statistical analysis*

The graphs were created using the GraphPad Prism 9.0 software and represent the middle of the results for each experiment realized with a minimum of three technical replicates. We determined the  $EC_{50}$  and  $CC_{50}$  values by Nonlinear regression of Log(inhibitor) or inhibitor vs. Normalized response of best curve generated ( $R^2$  values  $\geq 0.9$ ).

## Molecular modeling studies

#### *Protein and ligand structures preparation*

The molecular docking simulations were performed with the crystallographic structure of the SARS-CoV-2 Mpro enzyme obtained in a covalent complex with an irreversible peptide-like inhibitor named N3 (N-[(5-methylisoxazol-3-yl)carbonyl]alanyl-L-valyl-N ~ 1 ~ ((1R,2Z)-4-(benzyloxy)-4-oxo-1-[(3R)-2-oxopyrrolidin-3-yl]methyl}but-2-enyl)-L-leucinamide), available in the Protein Data Bank as PDB ID: 6LU7 (resolution = 2.16 Å)<sup>27</sup>. The missing residues were added using the CHARMM-GUI platform (<http://www.charmm-gui.org/>)<sup>28</sup>, defining the protonation state in physiological pH 7.4, which was predicted by pdb2pqr server (<https://server.poissonboltzmann.org/pdb2pqr>) and removing water molecules. The three-dimensional structures of the statine-like derivatives **7d**, **8e**, and **9g** were drawn using ChemDraw v. 20.0<sup>29</sup> considering their protonation state in physiological pH 7.4, and geometry optimization was performed using the MMFF94 force field available in the Spartan (v.10) software (Wavefunction, Inc. <https://www.wavefun.com>). Then the structures were converted to the pdbqt format using the Open Babel chemical toolbox<sup>30</sup>.

#### *Molecular docking*

The molecular docking simulations were performed using the AutoDock Vina 1.1.2 program<sup>31</sup> and prepared in the AutoDockTools (ADT) (v.1.5.6)<sup>32</sup> graphical interface according to the protocol and parameters previously described by our research group<sup>33</sup>, considering the physiological pH 7.4. The docking protocol was validated by redocking the N3 inhibitor as a non-covalent ligand. The N3 inhibitor was removed from the structure, and the binding orders were restored for the inhibitor interacting with Cys145 amino acid. The root-mean-square deviations (RMSD) calculations of the 10 pose results were carried out using the PyMOL (v. 3.5) software<sup>34</sup>, considering the best results RMSD < 2.0 Å. Ligand-Mpro complexes were analyzed for the main intermolecular interactions, such as hydrogen bond (H-bond) and hydrophobic interactions, with the PyMOL, and the images of the binding poses were composed with the Visual Molecular Dynamics (VMD) (v. 1.9.4) software<sup>35</sup>.

### Molecular dynamics

Molecular dynamics simulations were carried out in triplicate with the GROMACS 2022 package<sup>36</sup> using the CHARMM36 force field<sup>37</sup> with the top-ranked pose of each ligand-Mpro complex obtained by molecular docking applying the protocol described previously by our research group<sup>33</sup>. The ionization states of the protein's residues were adjusted to pH 7.4 using the pdb2gmX Python script. Ligand-protein complex was included in a periodic triclinic box (box dimensions: 5.416 × 4.538 × 4.348 nm and box volume: 854.88 nm<sup>3</sup>), solvated with the TIP3P model of water, and neutralized with 8 atoms of Na<sup>+</sup> ions to neutral ligands (**7d** and **8e**) and 7 Na<sup>+</sup> atoms to charged ligand (**9g**).

RMSD, RMSF, hydrogen bonding, and cluster analysis were made using gmX rms, gmX rmsf, gmX hbond, and gmX cluster modules available in the GROMACS package. The  $\Delta G_{\text{bind}}$  was calculated by the MM-PBSA method, applying g\_MMPBSA module v.5.1.2<sup>38</sup> considering the internal dielectric of the protein solute of 2. The energy contribution of residues was calculated with MmPbSaStat.py and MmPbSaDecomp.py scripts<sup>38</sup>. H-bonding frequencies were calculated with HbMap2Grace<sup>39</sup> software. Figures of the interactions and trajectories analysis were composed with the VMD<sup>35</sup> software.

## Results and discussion

### Chemistry

The statine-like core (AHPPA) contained in the final products was obtained by the stereoselective synthesis of the  $\beta$ -hydroxy- $\gamma$ -amino acid **1** synthesized from *L*-phenylalanine **2** (Fig. 1)<sup>40</sup>. The protected phenylalanine amino acid **3** and the freshly prepared Meldrum's acid were condensed to afford **4** and then it was refluxed in methanol affording the tetramic acid **5**. The  $\beta$ -hydroxypyrrrolidinone **6** was obtained from diastereoselective reduction of **5** using NaBH<sub>4</sub> and used as starting material to obtain the  $\beta$ -hydroxy- $\gamma$ -amino acid intermediate (*N*-Boc-(3*S*,4*S*)-AHPPA) **1**<sup>40–42</sup>.

The ester series of peptidomimetics **7a–h** (Table 1) was obtained by coupling reaction of statine-like compound **1** and several amino methyl ester hydrochlorides from the following natural amino acids *L*-serine (*L*-Ser-OMe.HCl), *L*-tyrosine (*L*-Tyr-OMe.HCl), *L*-threonine (*L*-Thr-OMe.HCl), *L*-isoserine (*L*-IsoSer-OMe.HCl), *L*-leucine (*L*-Leu-OMe.HCl), *L*-phenylalanine (*L*-Phe-OMe.HCl), *L*-proline (*L*-Pro-OMe.HCl) and *L*-valine (*L*-Val-OMe.HCl)<sup>43</sup>, by employing the classical *N*-(3-dimethylaminopropyl) *N*'-ethylcarbodiimide/1-hydroxy-benzotriazole/*N*-methylmorpholine (EDC/HOBt/NMM) protocol (Fig. 2)<sup>18,44</sup>.

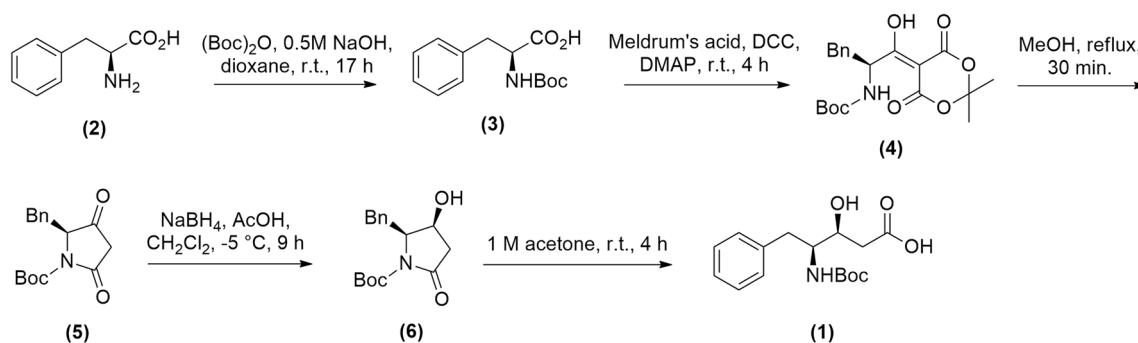
The acetylation reactions were carried out using acetic anhydride and DMAP in a basic medium affording the final compounds **8a–g**<sup>45</sup>. The *N*-Boc deprotection of **7a–h** compounds was made using trifluoroacetic acid generating peptidomimetics **9a–g**<sup>46</sup>. Finally, the hydrazide compound **10** was obtained from the corresponding ester **7 h** (R = -CH<sub>2</sub>Ph) and hydrazine hydrate in methanol (Fig. 2).

### Biological assays

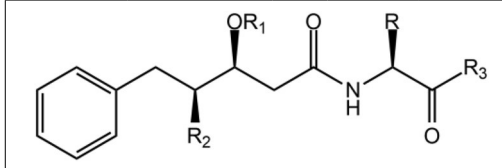
#### Inhibition of SARS-CoV-2 Mpro

The ability of compounds to inhibit the SARS-CoV-2 Mpro activity was assessed by an in vitro FRET-based assay. Table 2 shows % of maximum inhibition (efficacy) compared to negative control (vehicle, DMSO) and IC<sub>50</sub> (potency) values. The most promising compounds were selected based on efficacy (high maximum inhibition values) and potency (low half maximum inhibition, IC<sub>50</sub> values). As a positive control, we used the GC-376 (PubChem CID 71,481,119) compound, a small synthetic dipeptide previously identified as a covalent inhibitor of SARS-CoV's Mpro and used in several studies as gold standard for in vitro Mpro inhibition assay<sup>47,48</sup>. In our system, the maximum inhibition achieved by GC-376 was 75%, presenting an IC<sub>50</sub> = 0.541  $\mu$ M (Table 2). Sixteen out of 19 tested compounds inhibited Mpro activity by 50% or more (Table 2) and maximum inhibition was achieved with compound **9g**, which inhibited the enzyme activity by 80% compared to negative control (vehicle).

Only five compounds (**7d**, **8c**, **8d**, **8e**, and **9g**) were able to inhibit Mpro activity by 60% or more at low IC<sub>50</sub> values (Table 2), while all the other compounds only exerted inhibitory effects ( $\geq 50\%$ ) at high IC<sub>50</sub> values (IC<sub>50</sub> > 770  $\mu$ M). Therefore, we selected the three most promising compounds (**7d**, **8e**, and **9g**), which showed 70% or more inhibition than standard GC-376 and had an IC<sub>50</sub> of less than 1  $\mu$ M (Table 2). These compounds were further evaluated for their ability to inhibit virus replication.

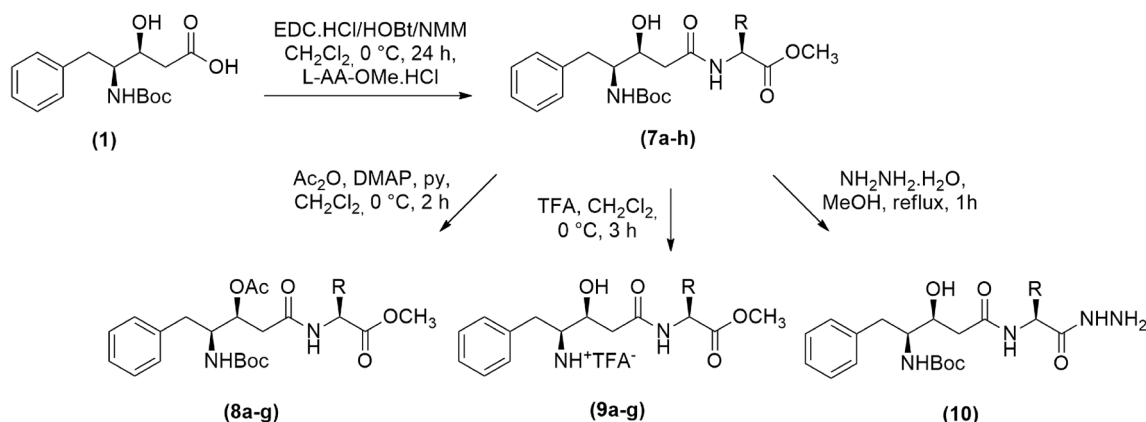


**Figure 1.** Synthesis of the  $\beta$ -hydroxy- $\gamma$ -amino acid intermediate (**1**).



Compound	R	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
7a	-CH <sub>2</sub> OH	-H	-NHBoc	-OCH <sub>3</sub>
7b	-CH <sub>2</sub> Ph-4-OH	-H	-NHBoc	-OCH <sub>3</sub>
7c	-CH(OH)CH <sub>3</sub>	-H	-NHBoc	-OCH <sub>3</sub>
7d	-CH <sub>2</sub> CH(OH)-	-H	-NHBoc	-OCH <sub>3</sub>
7e	-CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	-H	-NHBoc	-OCH <sub>3</sub>
7f	-CH <sub>2</sub> Ph	-H	-NHBoc	-OCH <sub>3</sub>
7g	-(CH <sub>2</sub> ) <sub>3</sub> -	-H	-NHBoc	-OCH <sub>3</sub>
7h	-CH(CH <sub>3</sub> ) <sub>2</sub>	-H	-NHBoc	-OCH <sub>3</sub>
8a	-CH <sub>2</sub> OAc	-Ac	-NHBoc	-OCH <sub>3</sub>
8b	-CH <sub>2</sub> Ph-4-OAc	-Ac	-NHBoc	-OCH <sub>3</sub>
8c	-CH(OAc)CH <sub>3</sub>	-Ac	-NHBoc	-OCH <sub>3</sub>
8d	-CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	-Ac	-NHBoc	-OCH <sub>3</sub>
8e	-CH <sub>2</sub> Ph	-Ac	-NHBoc	-OCH <sub>3</sub>
8f	-(CH <sub>2</sub> ) <sub>3</sub> -	-Ac	-NHBoc	-OCH <sub>3</sub>
8g	-CH(CH <sub>3</sub> ) <sub>2</sub>	-Ac	-NHBoc	-OCH <sub>3</sub>
9a	-CH <sub>2</sub> OH	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
9b	-CH <sub>2</sub> Ph-4-OH	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
9c	-CH(OH)CH <sub>3</sub>	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
9d	-CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
9e	-CH <sub>2</sub> Ph	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
9f	-(CH <sub>2</sub> ) <sub>3</sub> -	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
9g	-CH(CH <sub>3</sub> ) <sub>2</sub>	-H	-NH <sub>3</sub> <sup>+</sup>	-OCH <sub>3</sub>
10	-CH <sub>2</sub> Ph	-H	-NHNH <sub>2</sub>	-OCH <sub>3</sub>

**Table 1.** General structure of the synthesized statine-based peptidomimetics (**7a-h**, **8a-g**, **9a-g** and **10**) and their respective substituents (R, R<sub>1</sub>, R<sub>2</sub>, and R<sub>3</sub>).



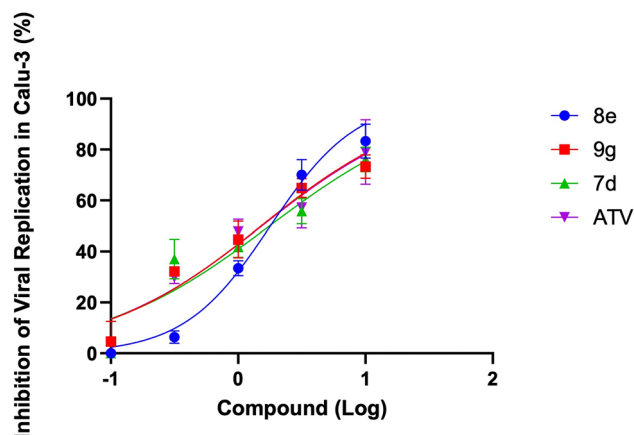
**Figure 2.** Synthesis of the statine-based peptidomimetics (**7a-h**, **8a-g**, **9a-g** and **10**).

#### Inhibition of SARS-CoV-2 virus replication

The compounds **7d**, **8e** and **9g** showed high efficacy, inhibiting around 80% of SARS-CoV-2 replication on Calu-3 cells at higher concentrations tested (Fig. 3). The EC<sub>50</sub> values of these compounds (**7d**, EC<sub>50</sub> = 1.76 μM; **8e**, EC<sub>50</sub> = 1.79 μM; and **9g**, EC<sub>50</sub> = 1.51 μM) are equivalent to atazanavir (ATV, EC<sub>50</sub> = 1.53 μM), an azapeptide antiretroviral drug approved for treating HIV (human immunodeficiency virus) patients (Table 3). Atazanavir is a HIV-1 aspartic protease inhibitor that can inhibit the SARS-CoV-2 Mpro enzyme, hampering therefore, the nsps processing and, consequently, the virus replication<sup>49,50</sup>. In addition, the compounds also presented a potent

Compound	Maximum inhibition (%)	IC <sub>50</sub> (μM)
7a	70	> 1000
7b	70	> 1000
7c	< 50	ND
7d	71	0.924
7e	< 50	ND
7f.	NT	ND
7g	NT	ND
7 h	< 50	ND
8a	50	1000
8b	65	> 1000
8c	60	2.449
8d	60	0.391
8e	70	0.606
8f.	NT	ND
8g	66	> 1000
9a	57	> 1000
9b	52	> 1000
9c	NT	ND
9d	58	777.8
9e	51	> 1000
9f.	68	> 1000
9g	80	0.617
10	61	> 1000
GC-376	75	0.541

**Table 2.** Inhibition of SARS-CoV-2 Mpro proteolytic activity by the statine-based peptidomimetics. *NT* not tested, *ND* not determined.



**Figure 3.** The dose-dependent curves of the statine-based peptidomimetics **7d**, **8e**, and **9g** (and atazanavir, ATV, as positive control) against SARS-CoV-2. Calu-3 cells infected with SARS-CoV-2 MOI 0.01 were treated with the compounds **7d**, **8e**, **9g**, and ATV at a semi-log curve of concentration (10, 3.16, 1, 0.316, and 0.1 μM) for 24 h at 37 °C, 5% CO<sub>2</sub> (n = 3).

effect when compared with molnupiravir (estimated EC<sub>50</sub> of 1.97 μM), an antiviral drug authorized by Food Drug Administration (FDA) under an emergency use for COVID-19 treatment; and more potent than lopinavir/ritonavir combination (EC<sub>50</sub> = 8.2 μM), proposed as a treatment for COVID-19 during 2020<sup>51,52</sup>.

Furthermore, none of the compounds showed toxicity in Calu-3 cells, and their CC<sub>50</sub> values exceeded the highest concentration tested in the experiment (Table 3). The absence of toxicity (CC<sub>50</sub> > 200 μM) and the lower EC<sub>50</sub> make the compounds biological effect specific to the virus with SI (selectivity index) values more than 100 (Table 3). Thus, the compounds **7d**, **8e**, and **9g** are promising bioactive molecules for inhibiting SARS-CoV-2 replication.



Compound	EC <sub>50</sub> (μM)	CC <sub>50</sub> (μM)	SI
7d	1.76 ± 0.3	>200	>113.6
8e	1.79 ± 0.1	>200	>111.7
9g	1.51 ± 0.2	>200	>132.4
ATV	1.53 ± 0.3	>200	>130.7

**Table 3.** The EC<sub>50</sub>, CC<sub>50</sub>, and SI values of statine-based peptidomimetics (**7d**, **8e**, and **9g**) and atazanavir (ATV) as positive control in Calu-3 cells. EC<sub>50</sub> – Compound concentration necessary to obtain 50% of its effective inhibitory activity; CC<sub>50</sub>—Compound concentration required to reduce 50% of cell viability; SI – Selective Index (CC<sub>50</sub>/EC<sub>50</sub>).

### In silico molecular docking and dynamics simulations

The potential binding mode and main intermolecular interactions of the statine-based derivatives **7d**, **8e**, and **9g** into the active site of the SARS-CoV-2 Mpro were evaluated through molecular docking following the protocol previously reported by our research group<sup>33</sup>. The docking protocol used in our study was validated by redocking, considering the **N3** inhibitor as a non-covalent inhibitor. The bond between **N3** and the enzyme was broken, and its double bond was restored. Our analysis considered 14 rotatable bonds, as the **N3** has 4 amide bonds. After executing the protocol, we calculated the RMSD value, which was 1.91 Å (Figure S2), in comparison to the structure available in the crystal. According to the molecular docking simulations with compounds **7d**, **8e**, and **9g**, it was observed that statine-based peptidomimetics could interact with the Mpro enzyme at the substrate binding site by hydrogen bond (H-bond) and hydrophobic interactions. Mpro has several subsites (“S”) for substrate binding<sup>33</sup>, which are identified by the Schechter & Berger (1967) nomenclature (e.g., S5, S4, S3, S2, S1, S1', S2', S3', S4', S5'), including **S5** (Pro168, Thr190, and Ala191), **S4** (Leu167, Phe185, Gln189, and Gln192), **S2** (His41, Met49, Tyr54, Met165, and Asp187), **S1** (Phe140, Leu141, Asn142, Ser144, His163, Glu166, and His172), and **S1'** (Thr24 and Thr25)<sup>53</sup>.

Specifically, the top-ranked docking poses of the statine-based derivatives **7d**, **8e**, and **9g** in general shows H-bond and hydrophobic interactions with residues of the catalytic dyad (Cys145 and His41)<sup>27</sup> and the **S1** subsite (Leu141, Asn142, and Glu166) (Fig. 4a–c).

The pose of the statine-like **7d** shows putative H-bond interactions with Gly143 and catalytic Cys145 (these two residues constitute the oxyanion hole)<sup>54</sup>, and in addition hydrophobic interactions with amino acids residues from three subsites: **S1** (Asn142, Leu141, Phe140, Glu166, and His163), **S2** (His41 and Met165), and **S4** (Gln189) (Fig. 4a).

The pose of **8e** shows putative hydrophobic interactions with residues from all subsites of Mpro, particularly **S1** (Phe140, Leu141, and His163), **S2** (Met49, Tyr54, Met165, and Asp187), **S4** (Gln189), and **S5** (Thr190). It also presented H-bond interactions with catalytic dyad His41-Cys145 (Fig. 4b). Finally, **9g** pose shows putative H-bond interactions with four residues (Phe140, Gly143, Ser144, and Cys145) and hydrophobic interactions with residues of subsites **S1** (Leu141, Asn142, His163, and Glu166) and **S2** (His41, Met49, and Met165) (Fig. 4c).

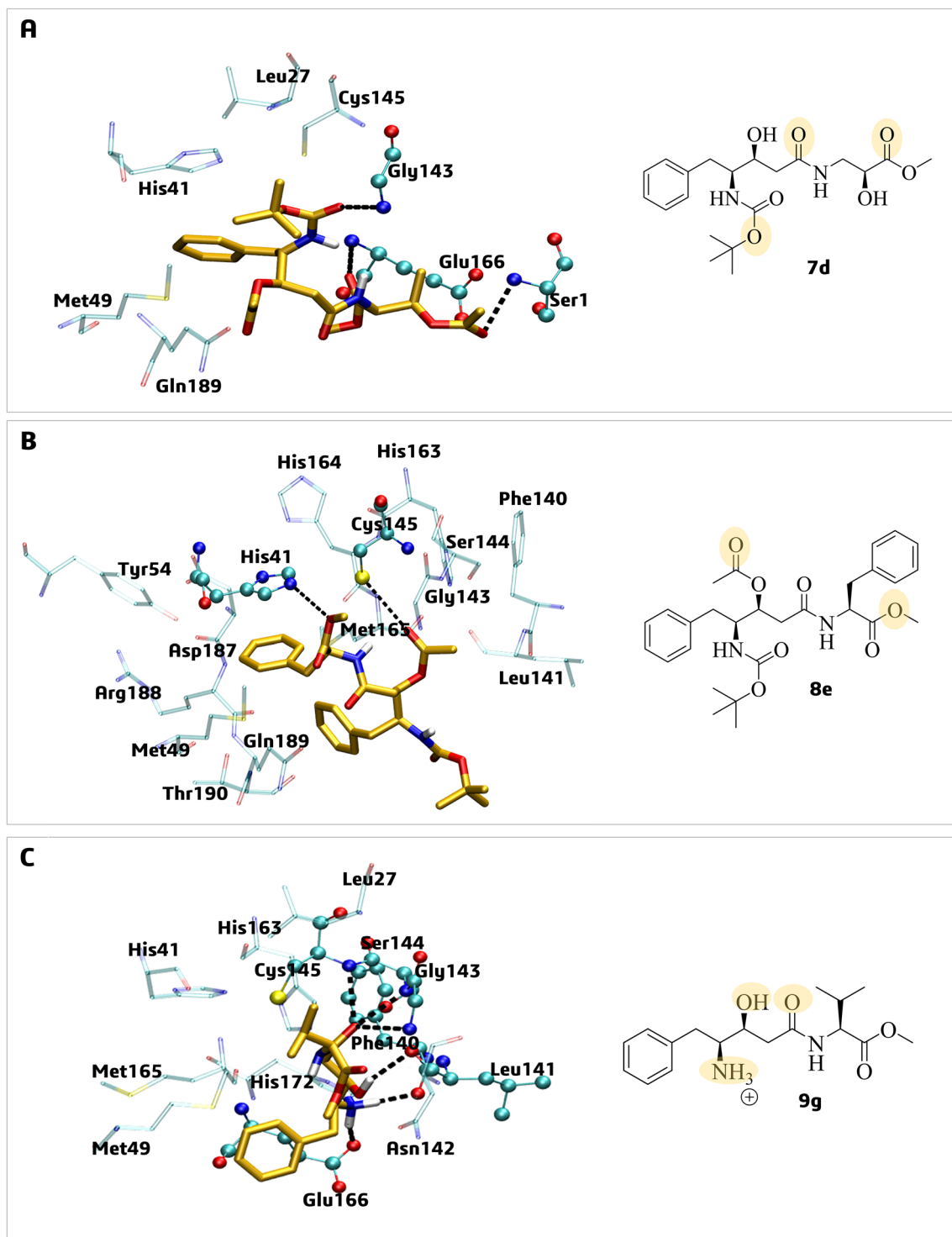
It is worth mentioning that the Mpro (PDB ID: 6LU7) co-crystallized inhibitor, **N3**, is an irreversible peptide-like inhibitor<sup>27</sup>. According to its redocking pose, **N3** shows similar H-bonding interactions with residues of subsites **S1** (Phe140), **S2** (His41 and Glu166), and **S4** (Gln189) (Figure S1), as seen in the statine-like derivatives proposed as non-covalent inhibitors, that shared similar binding mode, at least with two subsites as this inhibitor.

The molecular dynamics simulations (MD) were carried out in triplicate, starting with the top-ranked poses of **7d**, **8e**, and **9g** with SARS-CoV-2 Mpro (PDB ID: 6LU7), were performed to evaluate the behavior of these ligand–protein complexes in an aqueous system, during 200 ns, using the GROMACS software<sup>55</sup> with Charmm36 force field<sup>37</sup>.

In the first instance, the compounds were docked into the Mpro's active site, and as mentioned before, they remained close to the Cys145-His41 catalytic dyad region. To confirm if this specific area would encourage favorable and persistent interactions, we conducted an RMSD (root-mean-square deviations) analysis of the ligands over a 200 ns simulation period.

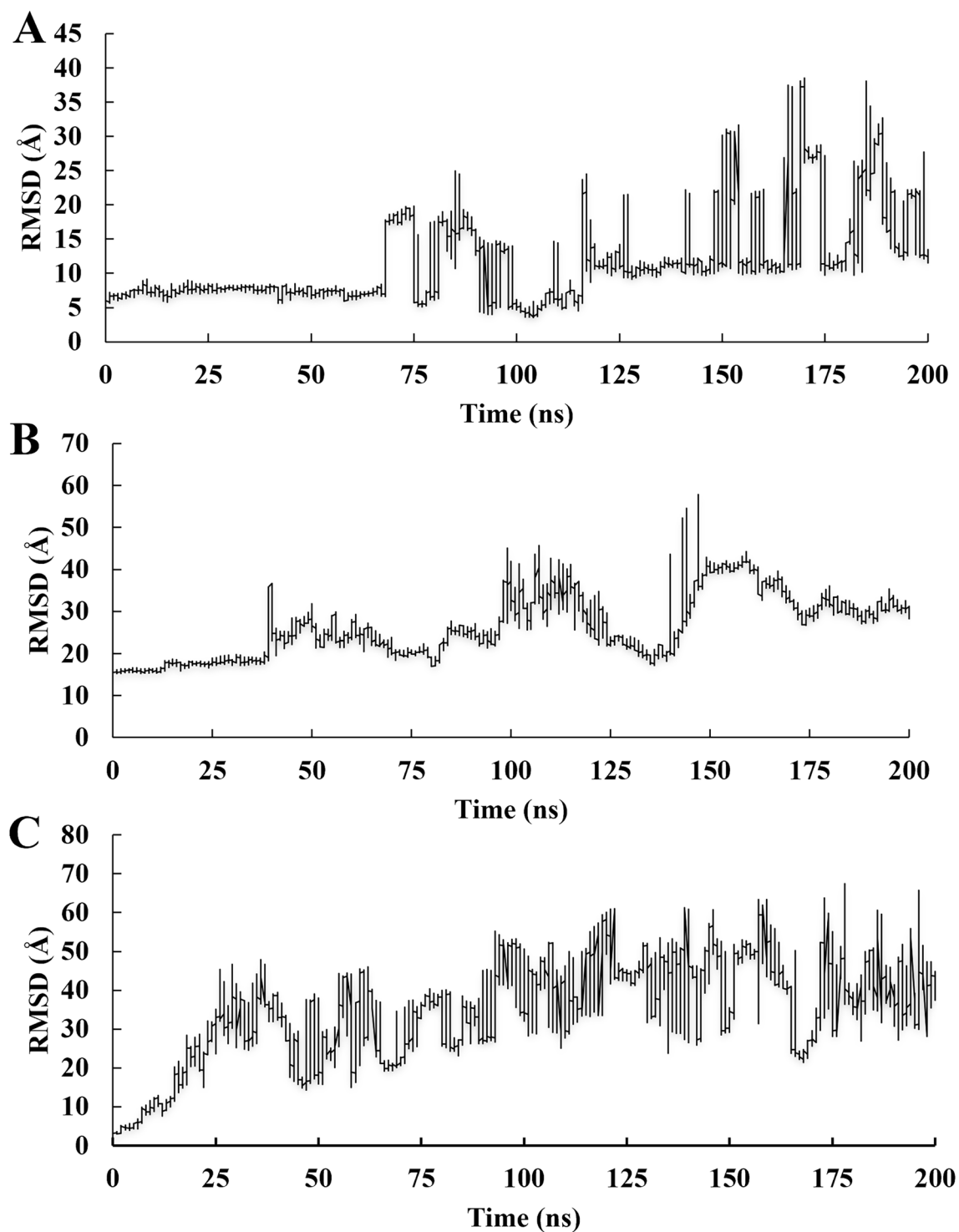
The RMSD analysis for compound **7d** showed a tendency to leave the active site after 70 ns of simulation, presenting an RMSD value of 11.6 ± 6.40 Å and a high standard deviation (Fig. 5a). Compound **8e** presented relative stability and persistence into the active site at the beginning of the simulation (0–40 ns); after that, left the active site presenting an RMSD value of 26.3 ± 7.34 Å (Fig. 5b). The derivative **9g** showed RMSD = 34.7 ± 12.8 Å, with low persistence in the active site (about 15 ns). Based on the 200 ns MDS analysis, it can be stated that inhibitor **7d** has the strongest interaction with the protein's active site. While the other inhibitors may exhibit good inhibition values, they cannot remain in contact with this region for an extended period.

Since Mpro is reported to have several subsites for substrate binding<sup>33</sup>, we evaluated the RMSD profiles of the Ca atoms per the subsites. The RMSD obtained from Ca atoms of the **7d-Mpro** complex showed instability of Ca atoms mainly to subsites **S2**, **S4**, and **S5** with RMSD values of 2.52 ± 0.35 Å, 1.48 ± 0.36 Å, and 1.30 ± 0.40 Å, respectively, showing more significant variations for **S4** and **S5**, even with low mean values. Moreover, it presented stability to **S1** + **S1'** subsites (0.85 ± 0.16 Å) for all MDs (Fig. 6a). On the other hand, **8e-Mpro** complex for subsites **S1** + **S1'** (1.09 ± 0.14 Å), **S4** (1.62 ± 0.42 Å), and **S5** (1.28 ± 0.38 Å) showed stability due to a minimal variation in standard deviation (sd) values < 1.0 Å<sup>56</sup>. At the same time, **S2** presented the high mean (2.16 ± 0.40 Å) (Fig. 6b). Although the mean and sd values observed are low, there appears to be some connection between the movements of the S2 and S4 subsites that affect the tendency for the ligand to escape from the binding site.



**Figure 4.** Best pose by molecular docking simulations of statine-like derivatives on the SARS-CoV-2 Mpro active site (PDB ID: 6LU7): (A) 7d; (B) 8e, and (C) 9g. The residues involved in H-bond interactions (dashed black lines) with the ligands are in ball-and-stick model (cyan color) and residues involved in hydrophobic interactions are in stick models (light green color). In 2D structures, the atoms (or groups) of the ligands involved in H-bond interactions are circled in yellow.



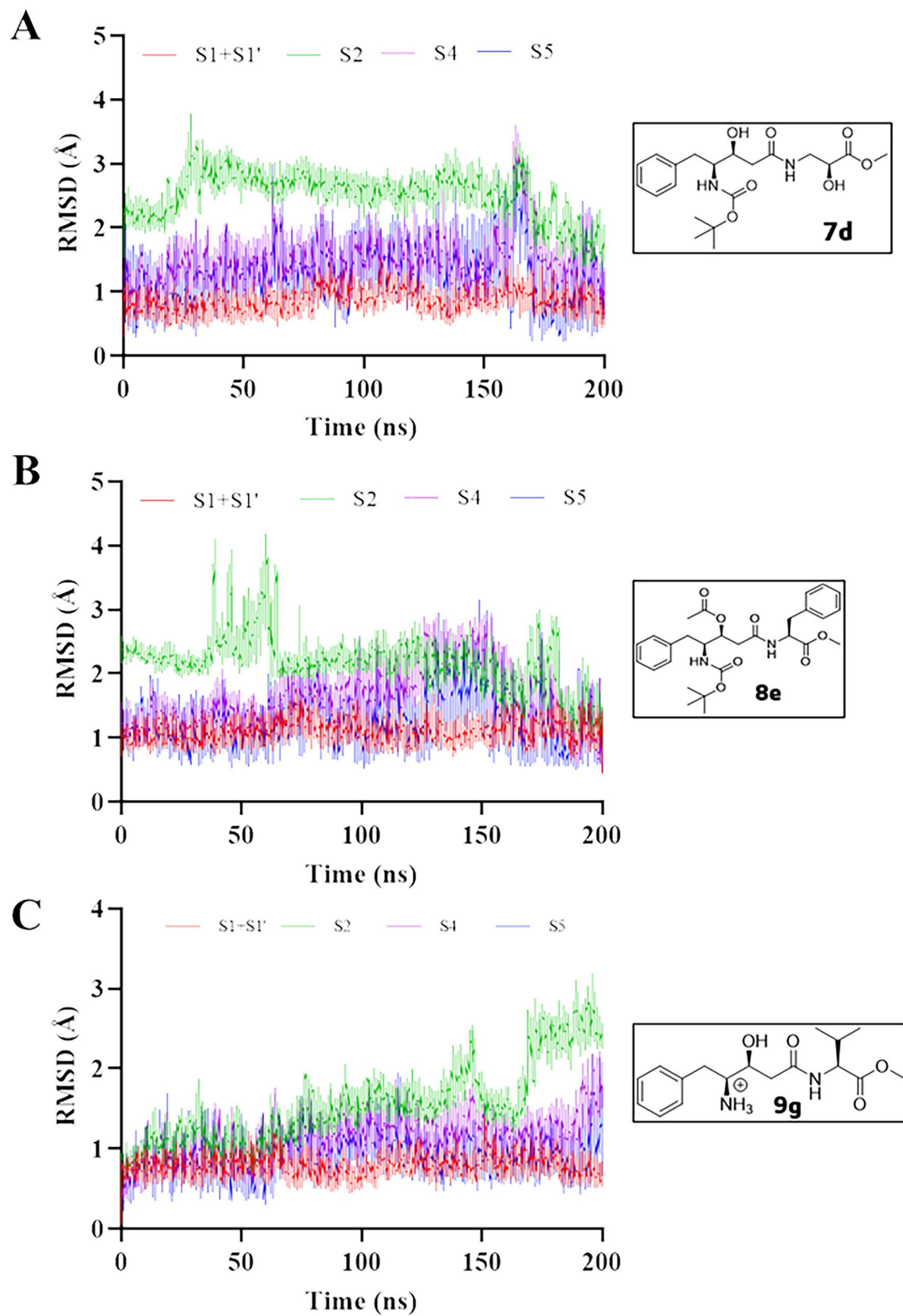


**Figure 5.** RMSD analysis of the **7d** (A), **8e** (B), and **9g** (C) relative to the Mpro-Ca atoms.

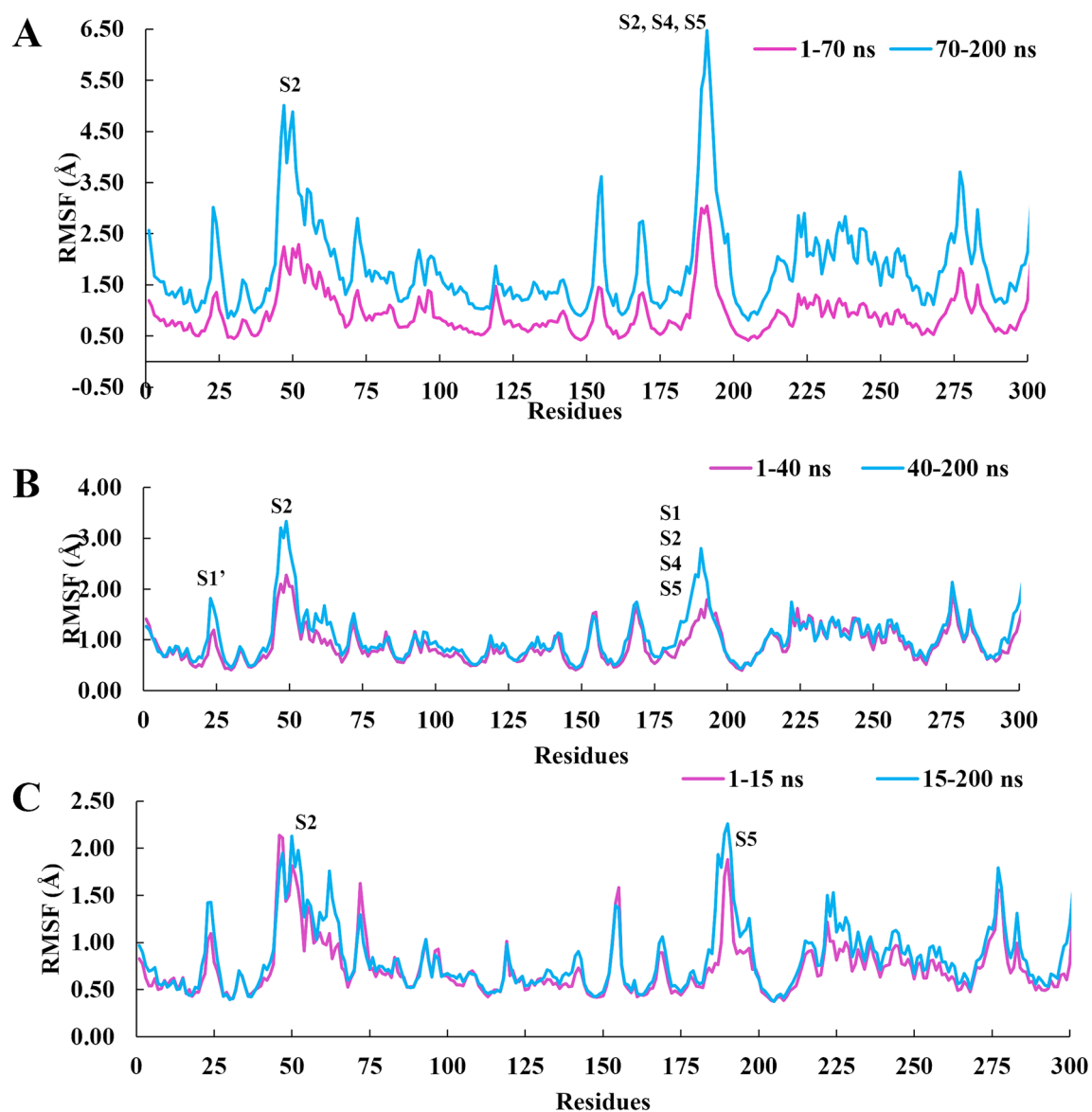
Notably, subsites S2 and S4 contain mainly hydrophobic residues that can interact with the phenyl groups of the ligand in these regions.

Finally, **9g-Mpro** complex showed instability of Ca atoms to S2 subsite from 150 ns of MD simulation presenting RMSD =  $1.54 \pm 0.50$  Å, while presenting stability for other subsites with  $0.78 \pm 0.11$  Å (S1 + S1'),  $1.13 \pm 0.24$  Å (S4), and  $0.95 \pm 0.21$  Å (S5) for all period (Fig. 6c).

It is interesting to note that the root-mean-square-fluctuation (RMSF) calculation for **7d** indicates that, even when this ligand is inside and out of the binding site cavity, the residues have mobility greater than 2.0 Å for Ca atoms of Asp187(S2), Gln189(S4), and Thr190(S5) (Fig. 7a). Considering the movement **8e** during the 200 ns of simulation, we evaluated the difference in the RMSF of the Ca atoms in two time intervals: 1–40 ns and 40–200 ns (Fig. 7b). In general, the ligand induces a gain of 0.98 to 2.10 Å in the mobility of residues belonging



**Figure 6.** Ca-RMSD analysis per Mpro subsites S1 + S1' (red line), S2 (green line), S4 (purple line), and S5 (blue line) that are relative to the simulations by molecular dynamics of the ligand-Mpro complex. Ligands: 7d (A), 8e (B), and 9g (C).



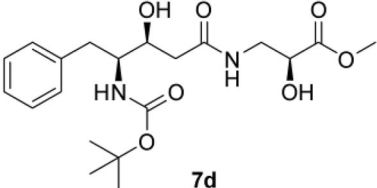
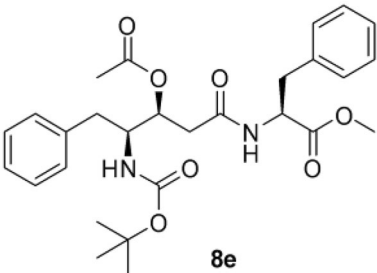
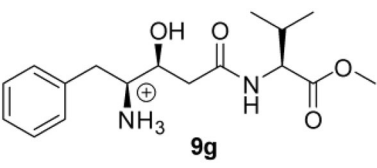
**Figure 7.**  $\text{Ca}$ -RMSF analysis at different times relative to the simulations by molecular dynamics of the ligand-Mpro complex. Ligands: **7d** (A), **8e** (B) and **9g** (C).

to **S1**, **S2**, **S4**, and **S5** subsites, especially for: (i) Thr24(**S1'**) and Thre25(**S1'**) (RMSF = 0.98 Å); (ii) Met49(**S2**) (RMSF = 2.10 Å); (iii) Glu166(**S1**), His172(**S1**), Met165(**S2**), Leu167(**S4**), and Pro168(**S5**) (RMSF = 1.49 Å); (iv) Asp187(**S2**), Gln189(**S4**), Thr190(**S5**), and Ala191(**S5**) (RMSF = 1.58 Å). The active site of Mpro is formed by catalytic dyad Cys145 and His41 and other residues, such as Leu141, Asn142, Pro168, Thr190, and Ala191<sup>57</sup>, which the last two presented fluctuations in the RMSF results. In addition, Glu166(**S1**) plays a critical role in establishing the interaction point between the two monomers of the Mpro<sup>58</sup>.

Regarding the statine-like derivative **9g**, the RMSF values were analyzed considering two-time intervals (1–15 ns, 15–200 ns) that indicated fluctuations above 2 Å for residues belonging to subsites **S2** (His41, Tyr54, and Met49) and **S5** (Thr190) predominantly (Fig. 7c).

The analysis of intermolecular interactions via hydrogen bonding of ligand **7d** exhibits a hydrogen bonding interaction with the residue Glu166(**S1**), His163(**S1**), and Ser144, with a great lifetime observed between 18 and 30%. In addition to H-bond interactions with catalytic residues Cys145 and His41 with low persistence, 5.19 and 8.97 respectively (Table 4).

The analysis of the **8e-Mpro** complex revealed those with the longest duration involving the ligand and the following residues: Glu166(**S1**) (43.8%) and Gln189(**S4**) with short lifetime from 6.26 to 15% (Table 4). Furthermore, FDA-approved Paxlovid™ (nirmatrelvir + ritonavir) was the first oral antiviral for mild to moderate COVID-19 cases in adults on May 25, 2023<sup>59</sup>. Nirmatrelvir inhibits viral replication by bonding to Cys145 catalytic residue from Mpro and forming hydrogen bonds with catalytic His164, Glu166(**S1**), and Gln189(**S4**)<sup>60</sup>. This finding supported the acetylated statine-like derivative **8e**, which demonstrates the potential of binding

Ligand	Donor	Acceptor	Lifetime (%)
 7d	Lig-N	Cys145-N	5.19
	Thr304-O	Lig-O	6.79
	Lig-O	Glu166-N	7.00
	Lig-O	Ser144-N	7.47
	Asn142-N	Lig-O	7.66
	His41-N	Lig-O	8.36
	Lig-N	His41-N	8.97
	Lig-N	Glu166-N	12.0
	Glu166-N	Lig-O	18.9
	Lig-N	Glu166-O	22.3
	Lig-O	His163-N	23.6
	Lig-N	Glu166-O	27.4
	Lig-O	Ser144-N	30.0
	Lig-O	Glu166-O	30.3
Lig-O	Glu166-O	30.4	
 8e	Lig-N	Gln189-O	6.26
	Gln189-N	Lig-N	7.93
	Gln189-N	Lig-O	15.0
	Glu166-N	Lig-O	13.5
	Glu166-N	Lig-O	43.8
 9g	Lig-N	Ser46-O	5.07
	Lig-O	Asp248-O	5.63
	Lig-N	Gln189-O	5.98
	Lig-N	Asp245-O	7.68
	Lig-O	Glu166-N	8.12
	Lig-N	Asp245-O	8.38
	Lig-N	Asp248-O	9.51
	Lig-O	Asp248-O	11.2
	Lig-N	Gln306-O	13.9
	Lig-N	Gln306-O	15.9

**Table 4.** Donor and acceptor atoms from ligands **7d**, **8e**, and **9g** involving interactions with respective amino acids and their lifetime (in percentage).

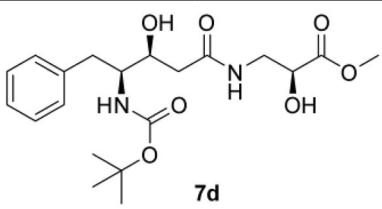
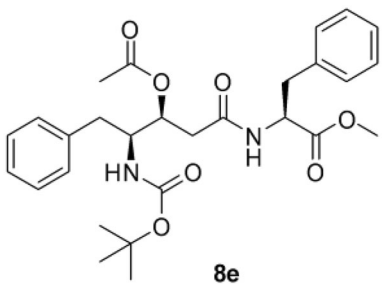
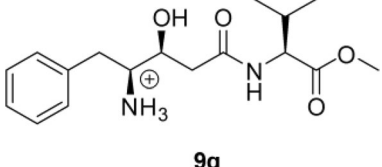
with at least two subsites, which is important to maintain the ligand in the active site, especially for proteases as Mpro<sup>53</sup>, highlighting these statine-based peptidomimetics to act as promising inhibitors.

Finally, for **9g** the most considerable hydrogen bonds observed presented a short lifetime with Glu166(S1) (8%), Asp245 and 248 (5.63–11.2%), Gln306 (13–15%), and Ser46 (5%) (Table 4).

In contrast to the docking results, the molecular dynamics simulations revealed that statine-like derivatives, like **8e** and **7d**, with large and hydrophobic substituents, including phenyl and *N*-Boc groups, had a greater influence on the number of interactions with amino acids belonging to at least two subsites due to their increased hydrophobicity. On the other hand, **9g**, which has weaker bonds, had less impact on the interactions.

Finally, since the RMSD analysis of the statine-like ligands indicated two constant regions before and after leaving the center of the active site, we calculated the binding free energy ( $\Delta G_{\text{bind}}$ ) for both periods of time (Table 5). Considering the period during which the ligands remained bound to the enzyme active site, it is more evident that statine-like derivative **9g** is promising for inhibition of Mpro since it presented the best  $\Delta G_{\text{bind}}$  value of  $-57.7$  kcal/mol, while statine-based derivative **7d** showed a lower  $\Delta G_{\text{bind}}$  value of  $-37.8$  kcal/mol and statine **8e** almost twice less energy value observed for **9g** with  $\Delta G_{\text{bind}} = -31.1$  kcal/mol (Table 5).

Based on our previous discussions, the binding free energy values also indicate the favorable hydrophobicity contribution of these statine-based compounds, probably due the hydrophobic characteristic of several residues on the Mpro subsites, such as S1 (Phe140, Leu141), S2 (Met49, Tyr54, Met165), S4 (Leu167, Phe185), and S5 (Ala191)<sup>53</sup>. As a result, the **8e** and **7d** derivatives with multiple hydrophobic substituents (i.e., two phenyl rings, an acetyl, and an *N*-Boc group) have stronger interactions with these pockets. Finally, the **9g** derivative has the most hydrophilic groups (i.e., a positive charged amine group and a free hydroxyl group) in the evaluated series,

 <b>7d</b>					
Time (ns)	$\Delta E_{\text{vdw}}$	$\Delta E_{\text{elect}}$	$\Delta E_{\text{solv}}$	$\Delta E_{\text{sasa}}$	$\Delta G_{\text{bind}}$
1–70	$-35.4 \pm 4.93$	$-16.7 \pm 12.6$	$18.8 \pm 11.4$	$-4.57 \pm 0.41$	$-37.8 \pm 22.6$
70–200	$-26.1 \pm 22.8$	$-8.01 \pm 9.72$	$7.88 \pm 5.29$	$-3.13 \pm 2.60$	$-29.3 \pm 28.4$
 <b>8e</b>					
Time (ns)	$\Delta E_{\text{vdw}}$	$\Delta E_{\text{elect}}$	$\Delta E_{\text{solv}}$	$\Delta E_{\text{sasa}}$	$\Delta G_{\text{bind}}$
1–48	$-35.6 \pm 3.10$	$-6.90 \pm 2.00$	$15.9 \pm 14.1$	$-4.60 \pm 0.20$	$-31.1 \pm 9.80$
48–200	$-33.6 \pm 1.60$	$-5.60 \pm 1.50$	$12.7 \pm 11.0$	$-4.20 \pm 0.40$	$-30.6 \pm 10.5$
 <b>9g</b>					
Time (ns)	$\Delta E_{\text{vdw}}$	$\Delta E_{\text{elect}}$	$\Delta E_{\text{solv}}$	$\Delta E_{\text{sasa}}$	$\Delta G_{\text{bind}}$
1–15	$-20.2 \pm 4.60$	$-78.5 \pm 22.2$	$44.5 \pm 41.6$	$-3.24 \pm 0.51$	$-57.7 \pm 14.8$
15–200	$-4.21 \pm 7.27$	$-60.7 \pm 73.0$	$24.4 \pm 28.4$	$-0.96 \pm 1.62$	$-41.5 \pm 55.2$

**Table 5.** The binding free energy ( $\Delta G_{\text{bind}}$ ) terms of the ligand-Mpro complexes calculated for **7d**, **8e**, and **9g** with the MM-PBSA method (mean  $\pm$  standard deviation energies; kcal/mol): van der Waals ( $\Delta E_{\text{vdw}}$ ), electrostatic ( $\Delta E_{\text{elect}}$ ), solvation ( $\Delta E_{\text{solv}}$ ), and solvent accessible surface area ( $\Delta E_{\text{sasa}}$ ).

resulting in the highest energy cost of desolvation of the binding site (44.5 kcal/mol) when compared to the other compounds ( $\sim 20$  kcal/mol) (Table 5).

It was noted that  $\Delta G_{\text{bind}}$  energy analysis of the ligands before moving out from the binding site was consistent, as previously discussed. Although derivative **9g** remained in the active site for less time than **8e** and **7d**, this interaction was sufficient to cause inhibition of the enzyme, which resulted in its best binding free energy value of  $-57.7$  kcal/mol observed (Table 5).

## Conclusions

Targeting the SARS-CoV-2 main protease (Mpro), 23 statine-based peptidomimetics were synthesized and tested for their ability to inhibit the Mpro activity. The three most effective compounds (**7d**, **8e**, and **9g**) could inhibit the Mpro enzyme activity in the sub-micromolar range. These compounds have been found to be non-cytotoxic and can suppress about 80% of the replication of the SARS-CoV-2 virus. In silico studies have also shown that these compounds are stable and have persistent interactions with the Mpro active site, indicating their potential as inhibitors. By blocking the activity of the main protease, which is essential for viral replication, these compounds have the potential to inhibit virus replication with low micromolar  $EC_{50}$ . Finally, we found new hit compounds that could lead to promising drug candidates against the COVID-19 disease.

## Data availability

The authors confirm all data generated and analyzed during this study are available in the article and in the supplementary information.



Received: 26 October 2023; Accepted: 10 April 2024

Published online: 18 April 2024

## References

- WHO; World Health Organization. WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/> (2023).
- V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: Implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
- Hilgenfeld, R. From SARS to MERS: Crystallographic studies on coronaviral proteases enable antiviral drug design. *FEBS J.* **281**, 4085–4096 (2014).
- NIH; National Institutes of Health. COVID-19 Treatment Guidelines. 2023 <https://www.covid19treatmentguidelines.nih.gov/about-the-guidelines/whats-new/>.
- Mohapatra, S. *et al.* Repurposing therapeutics for COVID-19: Rapid prediction of commercially available drugs through machine learning and docking. *PLoS One* **15**, e0241543 (2020).
- Jain, R. P. *et al.* Synthesis and Evaluation of Keto-Glutamine Analogues as Potent Inhibitors of Severe Acute Respiratory Syndrome 3CL pro. *J. Med. Chem.* **47**, 6113–6116 (2004).
- Shie, J.-J. *et al.* Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic  $\alpha$ ,  $\beta$ -unsaturated esters. *Bioorg. Med. Chem.* **13**, 5240–5252 (2005).
- Shao, Y.-M. *et al.* Design, synthesis, and evaluation of trifluoromethyl ketones as inhibitors of SARS-CoV 3CL protease. *Bioorg. Med. Chem.* **16**, 4652–4660 (2008).
- Zhang, L. *et al.*  $\alpha$ -Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: Structure-based design, synthesis, and activity assessment. *J. Med. Chem.* **63**, 4562–4578 (2020).
- Konno, S. *et al.* 3CL protease inhibitors with an electrophilic arylketone moiety as anti-SARS-CoV-2 agents. *J. Med. Chem.* **65**, 2926–2939 (2022).
- Liu, H. *et al.* Development of optimized drug-like small molecule inhibitors of the SARS-CoV-2 3CL protease for treatment of COVID-19. *Nat. Commun.* **13**, 1891 (2022).
- Goettig, P., Koch, N. G. & Budisa, N. Non-canonical amino acids in analyses of protease structure and function. *Int. J. Mol. Sci.* **24**, 14035 (2023).
- Omura, S. *et al.* Ahpatinins, new acid protease inhibitors containing 4-amino-3-hydroxy-5-phenylpentanoic acid. *J. Antibiot. (Tokyo)* **39**, 1079–1085 (1986).
- Marciniszyn, J., Hartsuck, J. A. & Tang, J. Pepstatin Inhibition Mechanism. in 199–210 (1977). [https://doi.org/10.1007/978-1-4757-0719-9\\_12](https://doi.org/10.1007/978-1-4757-0719-9_12).
- Sun, Y., Takada, K., Nogi, Y., Okada, S. & Matsunaga, S. Lower Homologues of Ahpatinin, Aspartic Protease Inhibitors, from a Marine *Streptomyces* sp. *J. Nat. Prod.* **77**, 1749–1752 (2014).
- Fehrentz, J. A. *et al.* HIV-1 protease inhibitors containing statine : Inhibitory potency and antiviral activity. *Biochem. Biophys. Res. Commun.* **188**, 865–872 (1992).
- Khare, S. *et al.* Antitrypanosomal treatment with benznidazole is superior to posaconazole regimens in mouse models of Chagas disease. *Antimicrob. Agents Chemother.* **59**, 6385–6394 (2015).
- Hoelz, L. V. B. *et al.* Synthesis, biological evaluation and molecular modeling of pseudo-peptides based statine as inhibitors for human tissue kallikrein 5. *Eur. J. Med. Chem.* **112**, 39–47 (2016).
- Kumar, S. *et al.* Clinically relevant cell culture models and their significance in isolation, pathogenesis, vaccine development, repurposing and screening of new drugs for SARS-CoV-2: A systematic review. *Tissue Cell* **70**, 101497 (2021).
- Jureka, A., Silvas, J. & Basler, C. Propagation, inactivation, and safety testing of SARS-CoV-2. *Viruses* **12**, 622 (2020).
- Sousa, F. & Castro, P. Cell-based in vitro models for nasal permeability studies. In *Concepts and Models for Drug Permeability Studies* 83–100 (Elsevier, 2016). <https://doi.org/10.1016/B978-0-08-100094-6.00006-7>
- Zhu, Y., Chidekel, A. & Shaffer, T. H. Cultured human airway epithelial cells (Calu-3): a model of human respiratory function, structure, and inflammatory responses. *Crit. Care Res. Pract.* **2010**, 1–8 (2010).
- Kozikowski, B. A. *et al.* The effect of freeze/thaw cycles on the stability of compounds in DMSO. *SLAS Discov.* **8**, 210–215 (2003).
- Dludla, P. V. *et al.* A dose-dependent effect of dimethyl sulfoxide on lipid content, cell viability and oxidative stress in 3T3-L1 adipocytes. *Toxicol. Rep.* **5**, 1014–1020 (2018).
- Nguyen, S. T., Nguyen, H. T.-L. & Truong, K. D. Comparative cytotoxic effects of methanol, ethanol and DMSO on human cancer cell lines. *Biomed. Res. Ther.* **7**, 3855–3859 (2020).
- World Health Organization (WHO). Laboratory biosafety guidance related to coronavirus disease (COVID-19). <https://www.who.int/publications/i/item/WHO-WPE-GIH-2021.1> (2023).
- Jin, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
- Lee, J. *et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).
- Evans, D. A. History of the Harvard ChemDraw Project. *Angew. Chem. Int. Ed.* **53**, 11140–11145 (2014).
- O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminform.* **3**, 33 (2011).
- Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, NA-NA (2009).
- Morris, G. M. *et al.* Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
- de Santiago-Silva, K. M. *et al.* In silico approach identified benzoylguanidines as SARS-CoV-2 main protease (Mpro) potential inhibitors. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2022.2123396> (2022).
- DeLano, W. L. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. protein Crystallogr.* **40**, 82–92 (2002).
- Humphrey, W., Dalke, A. & Schulten, K. V. M. D. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
- Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).
- Huang, J. & MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
- Kumari, R., Kumar, R. & Lynn, A. g\_mmpbsa—A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).
- Gomes, D. E. B., da Silva, A. W., Lins, R. D., Pascutti, P. G. & A., S. HbMap2Grace. Software for mapping the hydrogen bond frequency. *Lab. Mol. Model. Dyn.* (2009).
- Courcambeck, J., Bihel, F., De Michelis, C., Quéléver, G. & Kraus, J. L. Design of potential new HIV protease inhibitors: Enantio-convergent synthesis of new pyrrolidin-3-ol, and pyrrolidin-3-one peptide conjugates. *J. Chem. Soc. Perkin Trans. 1*, 1421–1430. <https://doi.org/10.1039/b101584m> (2001).
- Jouin, P., Castro, B. & Nisato, D. Stereospecific synthesis of N-protected statine and its analogues via chiral tetramic acid. *J. Chem. Soc. Perkin Trans. 1*, 1177. <https://doi.org/10.1039/p19870001177> (1987).

42. Pinheiro, S. *et al.* A general approach for the synthesis of 5-substituted-4-amino-pyrrolidin-2-ones and 5-substituted-4-amino-3-pyrrolin-2-ones. *Tetrahedron Lett.* **50**, 2402–2404 (2009).
43. Houšteká, R. *et al.* Biomimetic macrocyclic inhibitors of human Cathepsin D: Structure-activity relationship and binding mode analysis. *J. Med. Chem.* **63**, 1576–1596 (2020).
44. Abrahim-Vieira, B. *et al.* Novel isomannide-based peptide mimetics containing a tartaric acid backbone as serine protease inhibitors. *Med. Chem. Res.* **23**, 5305–5320 (2014).
45. Roncaglia, F., Parsons, A. F., Bellesia, F. & Ghelfi, F. Acetic anhydride/Et 3 N/DMAP: An effective acetylating system for hemiacetals. *Synth. Commun.* **41**, 1175–1180 (2011).
46. Liu, J. *et al.* Design, synthesis and biological evaluation of tasiamide B derivatives as BACE1 inhibitors. *Bioorg. Med. Chem.* **23**, 1963–1974 (2015).
47. Fu, L. *et al.* Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **11**, 4417 (2020).
48. Sharun, K., Tiwari, R. & Dhama, K. Protease inhibitor GC376 for COVID-19: Lessons learned from feline infectious peritonitis. *Ann. Med. Surg.* **61**, 122–125 (2021).
49. Chaves, O. A. *et al.* Atazanavir is a competitive inhibitor of SARS-CoV-2 Mpro, impairing variants replication in vitro and in vivo. *Pharmaceuticals* **15**, 21 (2021).
50. Fintelman-Rodrigues, N. *et al.* Atazanavir, alone or in combination with ritonavir, inhibits SARS-CoV-2 replication and proinflammatory cytokine production. *Antimicrob. Agents Chemother.* **64** (2020).
51. Li, P. *et al.* SARS-CoV-2 Omicron variant is highly sensitive to molnupiravir, nirmatrelvir, and the combination. *Cell Res.* **32**, 322–324 (2022).
52. Sacramento, C. Q. *et al.* In vitro antiviral activity of the anti-HCV drugs daclatasvir and sofosbuvir against SARS-CoV-2, the aetiological agent of COVID-19. *J. Antimicrob. Chemother.* **76**, 1874–1885 (2021).
53. Singh, E. *et al.* A comprehensive review on promising anti-viral therapeutic candidates identified against main protease from SARS-CoV-2 through various computational methods. *J. Genet. Eng. Biotechnol.* **18**, 69 (2020).
54. Lee, J. *et al.* Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. *Nat. Commun.* **11**, 5877 (2020).
55. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
56. de Souza, A. S. *et al.* 3-Acyltetramic acids as a novel class of inhibitors for human kallikreins 5 and 7. *Bioorg. Med. Chem. Lett.* **29**, 1094–1098 (2019).
57. Kneller, D. W. *et al.* Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat. Commun.* **11**, 3202 (2020).
58. Ferreira, J. C., Fadl, S. & Rabeh, W. M. Key dimer interface residues impact the catalytic activity of 3CLpro, the main protease of SARS-CoV-2. *J. Biol. Chem.* **298**, 102023 (2022).
59. Food and Drug Administration. Authorization for Paxlovid for Treatment of COVID-19. <https://www.fda.gov/media/155052/download>.
60. Noske, G. D. *et al.* An in-solution snapshot of SARS-COV-2 main protease maturation process and inhibition. *Nat. Commun.* **14**, 1545 (2023).

## Acknowledgements

The authors would like to thank the UFF multiuser laboratories (LaReMN of the Chemistry Institute and the Analytical Center of the Faculty of Pharmacy) for supporting the NMR and IR data, respectively. Thanks are also due to Biosafety Level 3 (BSL3) laboratory facility in Pavilhão Leonidas Deane, Instituto Oswaldo Cruz, Fiocruz, RJ; and Andre Sampaio from Farmanguinhos, platform RPT11M, for kindly donating the Calu-3 cells.

## Author contributions

Conceptualization, P.H.R.A.A., E.M.F.M.; methodology, P.H.R.A.A., P.G.C., A.S.R., D.D.C.S., A.R.T.; software, C.R.R., C.H.S.L.; formal analysis, P.H.R.A.A., P.G.C., L.E.C.C., S.S.C., C.S.S., A.S.R., D.D.C.S., A.R.T., V.N.S.F., T.K.F.O., N.R.R.B.; investigation, P.H.R.A.A., P.G.C., L.E.C.C., S.S.C., C.S.S., A.S.R., D.D.C.S., A.R.T., V.N.S.F., T.K.F.O., N.R.R.B.; writing-original draft preparation, P.H.R.A.A., E.M.F.M., P.G.C.; writing-review and editing, C.R.R., M.G.A., L.R.S.D., M.D.M., D.A., C.H.S.L., E.M.F.M.; resources, E.M.F.M., D.A., C.H.S.L.; supervision, M.G.A., L.R.S.D., R.G., M.D.M., D.A., C.H.S.L., E.M.F.M.; project administration, M.G.A., L.R.S.D., R.G., M.D.M., D.A., C.H.S.L., E.M.F.M.; All authors have read and agreed to the published version of the manuscript. All the authors are giving consent to publish.

## Funding

This research was funded by the following Brazilian governmental agencies: “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES) [Grant Numbers 88881.506711/2020-01, 88887.648435/2021-00, 88887.636136/2021-00, 88887.717861/2022-00, 88887.694990/2022-00, 88887.719751/2022-00], “Conselho Nacional de Desenvolvimento Científico e Tecnológico” (CNPq) [Grant Number 312027/2022-2], “Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro” (FAPERJ) [Grant Numbers E-26/201.426/2022, E-26/201.574/2021, E-26/210.068/2021, E-26/205.966/2022, E-26/205.967/2022, SEI-260003/019723/2022] and “Instituto Oswaldo Cruz, FIOCRUZ, FIOTEC (IOC)” [Grant Number IOC-023-FIO-18-2-58].

## Competing interests

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.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-59442-4>.

**Correspondence** and requests for materials should be addressed to C.H.S.L. or E.M.F.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024