Structure-based identification of naphthoquinones and derivatives as novel inhibitors of main protease Mpro and papain-like protease PLpro of SARS-CoV-2


aDepartment of Biochemistry and Immunology, Federal University of Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brazil;
bDepartment of Oncology and Pneumonology, Internal Medicine VIII, University Hospital Tübingen, Otfried-Müller-Straße 10, DE72076, Tübingen, Germany;
cSchool of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, 70211, Kuopio, Finland.
dInstitute of Exact Sciences, Department of Chemistry, Federal University of Minas Gerais, Belo Horizonte, 31270-901, Minas Gerais, Brazil;
eSkaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0657, USA;
fInstitute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 16610 Prague, Czech Republic;
gDepartment of Medicine, Division of Infectious Diseases, University of California San Diego, La Jolla, CA 92093, USA.
hDepartment of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA 94143, USA.

Corresponding authors:
A. J. O’Donoghue, E-mail: ajodonoghue@health.ucsd.edu; E. N. da Silva Júnior, E-mail: eufranio@ufmg.br; R. S. Ferreira, E-mail: rafaelasf@icb.ufmg.br
Abstract: The worldwide COVID-19 pandemic caused by the coronavirus SARS-CoV-2 urgently demands novel direct antiviral treatments. The main protease (Mpro) and papain-like protease (PLpro) are attractive drug targets among coronaviruses due to their essential role in processing the polyproteins translated from the viral RNA. In the present work, we virtually screened 688 naphthoquinoidal compounds and derivatives against Mpro of SARS-CoV-2. Twenty-four derivatives were selected and evaluated in biochemical assays against Mpro using a novel fluorogenic substrate. In parallel, these compounds were also assayed with SARS-CoV-2 PLpro. Four compounds inhibited Mpro with half-maximal inhibitory concentration (IC₅₀) values between 0.41 µM and 66 µM. In addition, eight compounds inhibited PLpro with IC₅₀ ranging from 1.7 µM to 46 µM. Molecular dynamics simulations suggest stable binding modes for Mpro inhibitors with frequent interactions with residues in the S1 and S2 pockets of the active site. For two PLpro inhibitors, interactions occur in the S3 and S4 pockets. In summary, our structure-based computational and biochemical approach identified novel naphthoquinoidal scaffolds that can be further explored as SARS-CoV-2 antivirals.

Keywords: SARS-CoV-2, main protease, papain-like protease, naphthoquinoidal inhibitors, virtual screening, enzyme assays

Abbreviations:
SARS – severe acute respiratory syndrome; MERS – Middle East Respiratory Syndrome; Mpro – SARS-CoV-2 Main protease; PLpro – SARS-CoV-2 Papain-like protease; PCA – Principal Component Analysis; Hbond – Hydrogen bond; IC₅₀ – Half-maximal inhibitory concentration; MD – Molecular Dynamics; BL2 – Blocking Loop 2; RMSD – Root-Mean-Square Deviation
1. Introduction

COVID-19 is caused by a β-coronavirus that is related to the virus that was responsible for the severe acute respiratory syndrome (SARS) in 2003, and therefore designated SARS-CoV-2 [1]. In December 2019, the first cases of COVID-19 were reported in Wuhan, the capital of Hubei Province, China [2]. The new coronavirus showed a rapid geographical spread, associated with a high infection rate, and the World Health Organization (WHO) declared it as a pandemic on March 11, 2020 [3,4]. The rapid transmission from human to human is undoubtedly the main source of contagion, which occurs mainly through droplets, hand contact, or contact with contaminated surfaces [5]. To control the spread of this pandemic virus, biosecurity and hygiene measures are now worldwide applied [6]. Despite the rapid development and emergency authorization of vaccines, viral escape mutants have emerged, and SARS-CoV-2 infections remain a concern for the global community. Therefore, there is a continuing need to discover structural frameworks for drugs that can be employed against COVID-19 [7].

Drug development efforts have targeted the SARS-CoV-2 main protease (Mpro) also known as 3-chymotrypsin-like protease (3CLpro) or non-structural protein 5 (nsp5) [8,9]. Mpro is an essential cysteine protease that cleaves the precursor replicase polyprotein in a coordinated manner [10], to generate at least 11 non-structural proteins [11]. As a target, Mpro is conserved among other coronaviruses, and has no closely related human homolog [12–14]. Therefore, it has been intensively investigated as a drug target for SARS and Middle East Respiratory Syndrome (MERS) [15–18]. Several Mpro inhibitors with in vitro antiviral activity against SARS-CoV-2 have been reported [19–25], including peptidomimetic aldehydes (best IC$_{50}$ values ranging ~0.03–0.05 µM [19,21,23]), α-ketoamides (best IC$_{50}$ values ranging ~0.04–0.67 µM [20,22]), calpain inhibitors (best IC$_{50}$ values ranging ~0.45–0.97 µM [20,23]), nonpeptidic inhibitors (best IC$_{50}$ values ranging ~0.17–0.25 µM [24–26]). The binding modes of dozens of these inhibitors have been determined by crystallography [20,22–29]. Recently, a covalent reversible nitrile was reported as an orally bioavailable Mpro inhibitor with in vitro and in vivo antiviral activity [30], and shown to reduce hospitalizations in COVID-19 patients by 89% [31]. Coronaviruses also encode a second cysteine protease, PLpro, that plays an essential role in suppression of the host immune system [32–34]. PLpro can hydrolyze ubiquitin and interferon-induced gene 15 (ISG15) from host proteins which allows the virus to evade the host innate immune responses [10,35]. This enzyme also cleaves the
viral polypeptide to release the nsp1, nsp2 and nsp3 proteins [36]. SARS-CoV-2 PLpro inhibitors with antiviral efficacy have been described [37–40], including naphthalene-based (EC$_{50}$ values of 1.4 to 21 µM for antiviral activity and IC$_{50}$ values of 0.18 to 43.2 µM against the enzyme [37–39]) and 2-phenylthiophene-based (EC$_{50}$ values of 2.5 to 11.3 µM for antiviral activity and IC$_{50}$ values of 0.11 to 0.56 µM against the enzyme [40]) non-covalent compounds, and 21 crystallographic structures of this protease complexed with a ligand are available [8,37–42]. The crystal structures of Mpro and PLpro with bound ligands provided us with a structural basis to identify novel inhibitors.

Repurposing existing chemical libraries is a promising strategy to quickly discover novel therapies [43,44]. Several newly discovered therapies for treatment of COVID-19 infection are derived from approved drugs, clinical candidates, and other pharmacologically active compounds that were originally developed for other indications [45–48]. In addition, knowledge gained from previous outbreaks of SARS, MERS, and bat coronavirus (BatCoV-RaTG13) have facilitated the rapid discovery of SARS-CoV-2 drugs [2,6,49]. Remdesivir, a broad-spectrum viral RNA-dependent RNA polymerase (RdRp) inhibitor [50,51], was rapidly approved for treatment of hospitalized patients with COVID-19 [52], which has resulted in a more rapid recovery of patients and lower levels of airway infection [53]. Drugs that provide either symptom relief for patients or have not been scientifically proven to be effective are also being widely studied by the scientific community [54].

Embelin, a natural product with a quinone core, has antiviral activity against influenza and hepatitis B [55,56]. Recently, it was shown that Embelin may inhibit Mpro and therefore have potential to be used as a treatment of SARS-CoV-2 [57]. In addition to Embelin, other studies showed that molecules containing a quinoidal framework also had inhibitory activity against SARS-CoV-2 Mpro. These included celestrol, pristimerin, tingenone and iguesterin [58]. We have experience working with naphthoquinones and therefore searched for structures with potential activity against SARS-CoV-2. In this report, we outline an *in silico* screening of a library of 688 quinonoid compounds and derivatives against SARS-CoV-2 Mpro, from which 24 compounds were selected and tested against this protease. Based on this strategy, and on experimental screening against PLpro as well, we report novel naphthoquinoidal inhibitors of both SARS-CoV-2 proteases. In addition to biochemical validation, molecular dynamics (MD) simulations indicated the stability of the Mpro and PLpro quinoidal complexes binding modes, mediated by interactions that were also frequently found in crystallographic complexes.
of the proteases. The quinones are promising COVID-19 drug candidates to be further explored, while also offering valuable insights into Mpro and PLpro inhibition.

2. Results

2.1. Assembly of a chemical library for virtual screening against Mpro

To search for potential Mpro inhibitors, we retrieved a library of quinones and their derivatives, as detailed in Figures 1 and 2 (See the Supporting Information Figures S1-S40, for more structural information). Six hundred and eighty-eight compounds were considered for virtual screening by molecular docking. We divided the molecules into eight different groups as described in Figure 1 (Groups 1-4) and Figure 2 (Groups 5-8). The compounds listed in Group 1 are ortho-quinones with different substitution patterns. In general, we evaluated compounds containing arylamino [59–64], alcohol [60] and alkoxy groups [61,62], selenium and sulfur [65–67], the basic chalcone framework [68], among simple ortho-quinones [61,68–71].

Group 2 is composed of para-quinones. We studied compounds such as α-lapachones [60,72], arylamino derivatives [60,64,71], furanonaphthoquinones [73–75], and pyrrolonaphthoquinones [73,76], in addition to other derivatives based on para-quinones [64,75]. The selected compounds for this group exhibit a broad substitution pattern but, in general, arylamino and aryl groups are often observed. Compounds with antiviral activity containing the para-quinone core are frequently described in the literature [77–79].

Groups 3 and 4 consist of ortho- and para-quinones with a 1,2,3-triazole nucleus. Lapachone-based 1,2,3-triazoles have been studied because of their broad spectrum of biological activities. We studied compounds with aromatic and aliphatic substituents [61,71,80–86], the presence of selenium [64,87], BODIPY [88,89], and sugars [71], among other substituent groups in the present quinoid structure [45,90–93].

The phenazine form of the triazole compounds and quinones described in groups 1 and 3 were also evaluated in group 5 [94–97]. Group 6 is the most complete and diverse group addressed in this study, containing approximately two hundred 1,4-napthoquinones with broad substitution patterns in the benzenoid A-ring and B-ring. Compounds containing sulfur, as sulfoxides and sulfones [98,99], selenium [100], iodine [47], amines, bromine, hydroxyls, alkenes, among other substituent groups [46,101–107]
were studied and evaluated according to their potential to act as anti-SARS-CoV-2. Imine
derivatives were also targeted in our studies and were placed in group 6 [67].

Finally, groups 7 and 8 are formed by hydrazo, imidazole, and oxazole derivatives
[108–111]. The compounds in these groups were prepared from the quinones described
above and represent our attempt to study quinone-derived heterocyclic compounds with
biological activity against various microorganisms and their effectiveness against the
virus that causes COVID-19.
**Group 1: Ortho-quinones**

- 20 compounds
- 5 compounds
- X = Se, 11 compounds
- X = S, 5 compounds
- 4 compounds
- 6 compounds
- 9 compounds
- 7 compounds

**Group 2: Para-quinones**

- 9 compounds
- 1 compound
- X = O, 9 compounds
- X = N, 24 compounds
- 4 compounds
- 10 compounds

**Group 3: Ortho-quinone-based 1,2,3-triazoles**

- 44 compounds
- 21 compounds
- 2 compounds
- 8 compounds
- 1 compound
- 16 compounds

**Group 4: Para-quinone-based 1,2,3-triazoles**

- 8 compounds
- 2 compounds
- 17 compounds
- 1 compound
- X = N, 30 compounds
- X = O, 11 compounds
- 12 compounds

**Figure 1.** The basic structural framework of the compounds listed in Groups 1-4.
Figure 2. The basic structural framework of the compounds listed in Groups 5-8.
2.2. Available Mpro structures show conserved conformation, protein-ligand interactions, and location of waters molecules

As an initial investigation to support the virtual screening of the quinoidal library, we analyzed 72 Mpro structures with bound ligands that have a resolution of 1.3 Å to 2.5 Å (Supporting Information Figure S41B). Active Mpro forms a homodimer comprising two protomers [19], while its monomer is inactive [112]. Each protomer is formed by domains I, II, and III, binding to each other by an N-terminal finger between domains II and III [24] (Figure 3). The substrate-binding site is located in a cleft between domains I and II and covered by a loop connecting them. Also crucial to the formation of the active dimer, the N-terminal finger of one monomer extends to the other monomer, shaping and forming the active substrate-binding site [22]. The substrate-binding cleft is composed of four subsites S1’, S1, S2, and S4 [14,19], which features a non-canonical Cys-His catalytic dyad [19,24] (Figure 3).

Using principal component analysis (PCA) to assess conformational differences among the structures, we found a high similarity among the Mpro structures evaluated. Even for the four most divergent structures (PDB codes: 6M2N [113], 6W63 [114], 6LU7 [24], and 7BQY [24], Supporting Information Figure S41A), carbon alpha (Cα) root-mean-square deviation (RMSD) between the protease structures is less than 1.0 Å (Supporting Information Figure S41B), suggesting high overall conservation of the quaternary structure.

On the other hand, the Mpro active site is known for its high flexibility with conformational changes induced by ligand binding [24,115,116]. Thus, to evaluate possible differences in active site residue conformations, we superimposed six high-resolution Mpro structures (1.31 Å to 1.51 Å, PDB codes 5R82, 5RFW, 5RF6, 5RFV, and 5RF3 [117]) with four structures that were discovered to have lower structural similarity by PCA and had resolutions between 2.10 Å and 2.20 Å. The superposition of these structures reveals that most residues in the ligand-binding site adopt similar conformations (Supporting Information Figure S42), except for M49, N142, M165, and Q189, which were the most flexible among the other binding site residues.
Figure 3. Three-dimensional structure of Mpro dimer (PDB code: 5R82 [117]) and surface view of the active site. Protomer A is shown in gray, and protomer B is colored according to three domains: red for domain I, blue for domain II, and purple for domain III. The loop linking domains II and III, critical for the protein dimerization, is colored in pink and the N-terminal finger colored orange (A). The substrate-binding cleft is highlighted with the catalytic residues H$_{41}$ and C$_{145}$ displayed as sticks (B). In the close-up view of the active site, the residues and conserved water molecules are colored by the frequency they are involved in interactions with 72 crystallography ligands, according to an analysis using the program LUNA (https://github.com/keiserlab/LUNA).

To better understand the molecular recognition between Mpro and inhibitors, we assessed the common protein-ligand interactions found in the 72 experimentally determined crystal structures, of which 49 displayed covalent and 23 non-covalent ligands, using the program LUNA (https://github.com/keiserlab/LUNA). Within this comprehensive set, ligands interacted most frequently with the catalytic dyad and residues in the S1 and S2 pockets (Figure 3B). For the catalytic dyad, C$_{145}$ interacts with 78% of the ligands, forming hydrophobic and hydrogen bond interactions, while H$_{41}$ binds to 82% of the inhibitors, mostly through aromatic stacking, hydrophobic, cation-π, and weak hydrogen bond interactions (Figure 3B and 4A). Within the S1 pocket, polar protein-ligand interactions were enriched such as hydrogen bonds and hydrophobic interactions with G$_{143}$ (68% interaction frequency); hydrophobic interactions and weak hydrogen bonds with N$_{142}$ (65% interaction frequency); and cation-nucleophile, cation-π, and...
hydrogen bond interactions with H$_{163}$ (58% interaction frequency), S$_{144}$ (35%) and E$_{166}$ (29%) in the S1 pocket and D$_{187}$ (33%), and C$_{44}$ (21%) in S2 had lower frequency of interactions (Figure 3B and 4A). On the other hand, the S2 subsite is more hydrophobic.

The two residues with the highest interaction frequencies from this pocket were M$_{49}$ (65%) and H$_{164}$ (58%) which formed hydrophobic and weak hydrogen bond interactions, while M$_{165}$ (47%) and Q$_{189}$ (43%) interacted mainly by hydrophobic contacts with the ligands (Figure 3B and 4A). The high frequency of interactions with S1 and S2 residues showed that most of the ligands fill one or both pockets, conserving a more polar profile for S1, whereas the S2 retained a more aromatic and aliphatic profile as observed previously with the SARS-CoV Mpro [118] and in other studies with SARS-CoV-2 Mpro [119,120].

Additionally, two S1 residues, T$_{25}$ (39%) and T$_{26}$ (35%), displayed frequent hydrophobic and weak hydrogen bond interactions. Amino acids in more solvent-exposed pockets, such as S1 residues L$_{27}$ (13%) and T$_{24}$ (3%), and S4 residues retained few or no interactions (Figure 4A). Several of the hydrogen bond interactions found by Luna were mediated by water, meaning ligand and protein residue are bridged by a solvent molecule.

In Mpro, water molecules contribute to ligand stabilization by forming water-mediated hydrogen bonds [19,20,24] and act as a possible third element to the catalytic dyad [9,121,122]. Therefore, we investigated which waters are conserved among the chosen Mpro structures using the ProBiS H2O plugin [123]. We found four conserved water molecules (present in over 50% of the structures, Figure 3 and Supporting Information Table S1), that interacted with 20-45% of the ligands, displaying van der Waals, hydrogen bond, and weak hydrogen bond interactions (Figure 4A). Thus, these crystallographic conserved and buried water molecules might be important for ligand recognition.
Figure 4. Interaction analysis between Mpro binding site residues and ligands from crystallographic structures and docking with Glide and Vina. Frequency and interaction types between residues and 72 crystallographic ligands (A), the final 24 selected compounds for biochemical assays from both Glide (B) and Vina (C). Residues with (*) are from the other protomer. From docking results, the (**) highlight residues with no interactions, while (***) are residues that were not considered in the docking calculations.
2.3. Virtual screening of naphthoquinoidal compounds against SARS-CoV-2 main protease

Considering the high conservation within the Mpro crystal structure conformations, we performed initial molecular docking experiments with the highest-resolution structure (1.31 Å – PDB Code 5R82 [117]) from the most populated of the structural clusters, followed by a second round of flexible docking with compounds prioritized from rigid docking. Due to the importance of water molecules in the ligand binding site, we retained two of the four conserved water molecules, that may mediate hydrogen bonds with H41, C145, E166, and L167, for molecular docking (Figure 4A and Supporting Information Table S1). To account for the possibility of water displacement by ligands, a second Mpro preparation was also performed in the absence of water molecules. Both preparations were submitted to two distinct docking algorithms, Glide [124] and Autodock Vina [125].

Docking results were visually inspected and relevant poses were selected according to their overall binding site complementarity and specific protein-ligand interactions. Thus, we prioritized 70 compounds that interacted with the previously established high frequent residues, H41, M49, N142, G143, C145, M165, Q189, and water molecules for flexible docking approaches. Overall, Glide and Vina docking modes established contacts with S1’, S1, and S2 residues. However, a slight shift in interaction patterns was found. Compounds from the quinoidal library did not establish as many hydrogen bond interactions as the crystallographic ligands, giving a more hydrophobic nature to the interactions (Figures 4B/C and Supporting Information Figure S43).

In the second round of docking, we treated M49, N142, M165, and E189 as flexible residues, as these were most flexible within crystal structures analyzed and interacted with a high number of ligands (40-65%). Based on these results, we selected 24 (out of 70) compounds that matched the desired residue interactions (Figure 4B and C) and maintained good complementarity to the binding site (Supporting Information Figures S44 to S47), for experimental validation in biochemical assays. The compounds selected represent diverse scaffolds from our library, comprising ortho-quinone-based 1,2,3 triazoles (group 3), para-quinone-based 1,2,3 triazoles (group 4), 1,4-naphthoquinones (group 6), and hydrazo derivatives (group 7).
2.4. Design and validation of Mpro substrate

Prior to biochemically evaluating the compounds against Mpro, we designed a fluorescent-quenched peptide substrate with the sequence ATLQAIAS that corresponds to the P4 to P4' amino acids of the nsp7-nsp8 cleavage site and the dash representing the scissile bond. This substrate was chosen because the sequence most closely matches the consensus sequence for all 11 viral polypeptide cleavage sites (Figure 5A and B) [126]. ATLQAIAS was flanked by 7-methoxycoumarin-4-acetyl-L-lysine on the N-terminus, dinitrophenyl-L-lysine on the C-terminus. The peptide contains several non-polar amino acid residues and therefore two d-Arginine residues were added on the N-terminus to increase solubility. Using a concentration range of 3 µM to 250 µM, the Kᵢ for this substrate was calculated to be 52.1 µM ± 14.4 µM.

2.5. Validation of novel Mpro inhibitors

We evaluated the 24 hit compounds from our virtual screen in a biochemical assay using recombinant SARS-CoV-2 Mpro. The enzyme was pre-incubated with each compound at 10 µM and then assayed with the fluorogenic peptide substrate. To avoid detecting aggregators as false positives [127,128], our assay was performed in the presence of 0.01% Tween 20. Additionally, we evaluated the absorbance of MCA fluorescence by the compounds, to make sure the observed enzyme inhibition was not an artifact of fluorescence, another common cause of false positives in enzyme assays [129]. From this screen, three 1,4-naphthoquinones derivatives, 379, 382, and 415, fully inhibited Mpro, while two quinone-based 1,2,3 triazoles, 191 and 194, had 50% or more inhibition. 668 was insoluble in assay buffer and was therefore eliminated from further analysis, while the remaining compounds had inhibition profiles of less than 50% (Table 1). The most potent compounds were subsequently evaluated at a concentration range of 10 µM to 9.7 nM and the half-maximal inhibitory concentration (IC₅₀) was calculated to be 66 µM ± 22 for 191, an ortho-quinone-based 1,2,3 triazole, 5 µM ± 0.15 for 415, 0.63 µM ± 0.04 for 379, and 0.41 µM ± 0.015 for 382 (Table 1 and Figure 5).
Table 1. Percentage of inhibition at 10 µM and IC$_{50}$ for naphtoquinoidal compounds against SARS-CoV-2 Mpro and PLpro.

<table>
<thead>
<tr>
<th>Compound $^a$</th>
<th>Mpro</th>
<th>PLpro</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% inhibition at 10 µM $^b$</td>
<td>IC$_{50}$ (µM) $^c$</td>
</tr>
<tr>
<td>159</td>
<td>35 ± 1</td>
<td>ND</td>
</tr>
<tr>
<td>189</td>
<td>25 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>191</td>
<td>55 ± 3</td>
<td>66 ± 20</td>
</tr>
<tr>
<td>193</td>
<td>32 ± 1</td>
<td>ND</td>
</tr>
<tr>
<td>194</td>
<td>50 ± 7</td>
<td>ND</td>
</tr>
<tr>
<td>195</td>
<td>40 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>196</td>
<td>29 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>197</td>
<td>18 ± 4</td>
<td>ND</td>
</tr>
<tr>
<td>314</td>
<td>36 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>318</td>
<td>45 ± 13</td>
<td>ND</td>
</tr>
<tr>
<td>319</td>
<td>5 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>320</td>
<td>11 ± 6</td>
<td>ND</td>
</tr>
<tr>
<td>321</td>
<td>40 ± 8</td>
<td>ND</td>
</tr>
<tr>
<td>379</td>
<td>100 ± 0</td>
<td>0.63 ± 0.04</td>
</tr>
<tr>
<td>380</td>
<td>12 ± 3</td>
<td>ND</td>
</tr>
<tr>
<td>382</td>
<td>100 ± 0</td>
<td>0.41 ± 0.02</td>
</tr>
<tr>
<td>414</td>
<td>29 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>415</td>
<td>100 ± 0</td>
<td>5 ± 0.2</td>
</tr>
<tr>
<td>465</td>
<td>5 ± 5</td>
<td>ND</td>
</tr>
<tr>
<td>470</td>
<td>1 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>477</td>
<td>5 ± 2</td>
<td>ND</td>
</tr>
<tr>
<td>666</td>
<td>5 ± 5</td>
<td>ND</td>
</tr>
<tr>
<td>668</td>
<td>NT</td>
<td>NT</td>
</tr>
<tr>
<td>673</td>
<td>8 ± 3</td>
<td>ND</td>
</tr>
</tbody>
</table>

$^a$See Supporting Information File for all structures. $^b$Percentages of inhibition are reported as averages and standard deviation of the mean calculated from one experiment performed in triplicate. Compounds were pre-incubated with enzymes for 15 min before addition of the substrate. $^c$IC$_{50}$ values are reported as the averages and standard deviation of the mean, based on two independent experiments. Each IC$_{50}$ curve was determined based on at least 7 compound concentrations in triplicate. ND: not determined. NT: not tested.
Figure 5. Validation of SARS-CoV-2 Mpro inhibitors in enzyme assays. List of the 11 Mpro cleavage sites (A) and the design of a fluorescent-quenched peptide substrate (B). The ATLQAIAS substrate was chosen since it closely matches the consensus sequence of all 11 viral polypeptide cleavage sites. IC₅₀ curves for SARS-CoV-2 Mpro inhibitors (C). For each compound, two IC₅₀ curves are shown, corresponding to two independent experiments (data shown as spheres or squares for each experiment), in which the compounds were pre-incubated with Mpro prior to substrate addition. Each curve was determined based on at least 7 compound concentrations in triplicate.
To better understand the mechanism of Mpro inhibition by naphthoquinone-based derivatives, we evaluated whether compounds 382 and 415 were time-dependent inhibitors, a hallmark of covalent-acting molecules. First, enzyme inhibition after 15 min preincubation with the compounds was compared to activity without preincubation [130]. The IC$_{50}$ values observed in these two assay conditions were similar, with slightly lower IC$_{50}$ values upon preincubation (0.42 µM ± 0.02 upon incubation vs 0.80 µM ± 0.06 without incubation for 382 and 5.0 µM ± 0.2 upon incubation vs 16 µM ± 1 without incubation for 415) (Figure 6A and B), while for the positive control GC373 the IC$_{50}$ was ten-fold lower upon preincubation (0.003 µM ± 0.001). A dilution experiment was also performed, to check whether the compounds were irreversible. We incubated the inhibitors and Mpro at high concentrations and then diluted the incubation mixture, resulting in inhibitor concentrations 10-fold lower than their apparent IC$_{50}$. In this assay, an irreversible inhibitor will maintain approximately 10% of enzymatic activity, while a rapidly reversible inhibitor will dissociate from the enzyme to restore approximately 90% of enzymatic activity following the dilution event [130,131]. When this was performed with Mpro and GC373, a covalent Mpro inhibitor, the enzyme remained inhibited upon dilution. The same behavior was observed for compound 415 suggesting that this inhibitor is an irreversible covalent inhibitor (see Figure 7 for the proposed binding mechanism). However, when the same test was carried out with compound 382 enzyme activity returned after dilution (Figure 6C). This suggested that the inhibition by 382 is reversible.
**Figure 6.** Evaluation of time-dependence and reversibility of Mpro inhibition by compounds **382** and **415**. IC$_{50}$ curves for SARS-CoV-2 Mpro inhibitors **382** (A) and **415** (B). For each compound, two IC$_{50}$ curves are shown, corresponding to two independent experiments (data shown as spheres or squares for each experiment), in which the compounds were pre-incubated with Mpro prior to substrate addition (black) and without preincubation with the compounds (purple). Reversibility assay (C). After preincubation of Mpro with compounds, at higher concentrations, the sample was diluted, and product formation was monitored for 120 minutes. Compound **382** reduced the enzymatic reaction rate by 26% compared to vehicle control (red), while the compound **415** reduced product formation by 100%, and this activity was not restored over a 2h period post dilution, as observed for the covalent inhibitor GC373 (black).

To gain insights into the proposed binding mode of our Mpro inhibitors and guide future optimization efforts, we conducted docking and MD studies with compounds **382** and **415**, representatives from two inhibitor scaffolds discovered. Our simulations considered the **415** ligand covalently bound, given the proposed reaction mechanism (Figure 7A), to both monomers in the Mpro dimer, and **382** freely. For the free simulations, however, the loss of interactions with E166 resulted in ligands being expelled.
from one of the binding sites within a few nanoseconds (~200 ns) of simulation (Supporting Information **Table S2**). Our analysis is focused on the other binding site, that retained the ligand with stable interactions along the analyzed trajectory.

For both ligands, the most representative binding modes observed in the MD simulations (Figure 7A and 7B) retain key interactions proposed based on docking with Glide (Supporting Information **Figure S45**). However, compound 415 showed a more conserved binding mode throughout the trajectory, being well represented by a single pose, in which the nitro group interacts with the S1 pocket and the 1,4 naphthoquinone interacts with the catalytic H₄₁ and S2 subsite residues (Figure 7B). On the other hand, the higher variability in the orientation of compound 382 led to four clusters with frequency between 17.5 and 31.7% (Supporting Information **Figure S48**). Overall, the 1,4-naphthoquinone ring of ligand 415 occupies the S1 pocket, but fluctuations in the ring orientation reflect on varied positions for the phenyl substituents. In the most populated cluster (Figure 7C), the methoxyphenyl substituents occupy the S1’ and S2 subsites.

Compounds 382, and 415 display stable polar contacts (hydrogen bond and water bridges) with G₁₄₅ and S₁₄₄ in the S1 pocket and π-cation or π-π interactions with the sidechain of H₄₁. These interactions were more frequent in the covalent simulations. The ligands also display stable polar interactions with the main-chain nitrogen from E₁₆₆ and electrostatic contacts with its side-chain (Figure 7D and 7E), a residue that adopts a stable conformation due to an interaction between its sidechain and the S₁ from the other protomer (S₁*). Hydrophobic interactions to M₄₉ and M₁₆₅ from the S2 pocket, are seldomly observed for these inhibitors and frequent interactions with the side-chain of C₁₄₅ was seen for the covalent inhibitor.
**Figure 7.** Predicted binding modes of compounds 415 and 382 to SARS-CoV-2 Mpro.

The proposed mechanism of 415 covalent binding to C145 (A). Proposed binding modes from a representative frame in the MD simulation of compounds (B) 415 (orange) covalently bound to Mpro C145, and (C) 382 (pink) bound to Mpro, and the frequency of protein-ligand interactions for all simulations with ligands (D) 415 and (E) 382. Mpro
residues are colored according to the types of atoms in the interacting amino acid residues (protein carbon, light gray; nitrogen, blue; oxygen, red; sulfur, yellow), hydrogen bond interactions are represented as yellow dashed lines. Mean interaction frequency is represented, with standard error of the mean (N=5) interval depicted as error bars, each point displays the individual value for a particular simulation replica and each chain.

2.6. Validation of novel PLpro inhibitors

Although our virtual screening studies were focused solely on Mpro, we were also interested in testing the virtual screening hits against the second SARS-CoV-2 cysteine protease, PLpro, to determine if any of the molecules were dual inhibitors of the viral enzymes. PLpro cleaves three sites on the viral polypeptide but also acts as a de-ubiquitinase. Therefore, we identified a fluorogenic substrate for human de-ubiquitinases (Z-RLRGG-7-amino-4-methylcoumarin) as a substrate for SAR2-CoV-2 PLpro. Recombinant PLpro was incubated with 6 µM to 500 µM of this substrate and the $K_M$ value was calculated to be 376.6 ± 32.3 µM. PLpro enzyme was pre-incubated with the same set of 23 compounds at 10 µM and then assayed with the fluorogenic substrate. Compound 668 was again eliminated due to insolubility in the assay buffer. Surprisingly, a total of 12 compounds inhibited PLpro by > 50% and the top three (compounds 159, 189, and 191) inhibited at > 90%. These top compounds are ortho-quinone-based 1,2,3-triazoles derivatives, sharing a common scaffold. The IC$_{50}$ values were calculated to be 1.7 µM, 2.2 µM, and 3.1 µM for compounds 159, 189, and 191, respectively (Figure 8).

Among the compounds that caused lower PLpro inhibition, five are N-substituted analogs of these hits, compounds 193-197, and had IC$_{50}$ values between 20 and 46 µM (Table 1, Figure 7). These eight PLpro inhibitors share a tricyclic 1,2-naphthoquinone ring that seems important for enzyme inhibition, as its replacement by a para-tolyl sulfone abolished activity against PLpro (compare compounds 189 vs 319; 191 vs 321; 193 vs 314; and 197 vs 318, Table 1).
Figure 8. IC$_{50}$ curves for SARS-CoV-2 PLpro inhibitors. For each compound, two IC$_{50}$ curves are shown, corresponding to two independent experiments (data shown as spheres or squares for each experiment). Each curve was determined based on at least 11 compound concentrations in triplicate.
Since the PLpro inhibitors have a shared scaffold, we selected compounds 189 and 195 (N-substituted) for computational studies. The SARS-CoV-2 PLpro has similar folding to the homologous enzymes from other coronaviruses [132], with domains showing a “thumb-palm-fingers” pattern and an N-terminal ubiquitin-like (Ubl) domain (first 60 residues) (Figure 9A) [40]. As a cysteine protease, PLpro contains a canonical catalytic triad, Cys-His-Asp (C111, H272, and D286) [10] located in a solvent-exposed cleft at the interface of the palm and thumb domains [40]. Analysis of common protein-ligand interactions from the crystallographic structures showed little or no interaction with the catalytic triad, in agreement with very narrow S1 and S2 pockets, which have high specificity for glycine (Figure 9B and Supporting Information Figure S49). Only covalently bound peptidic inhibitors, containing glycines at P1 and P2, occupy these pockets [37,40]. Instead, the non-covalent ligands bind to a groove corresponding to the S3 and S4 subsites, approximately 8 Å from the catalytic cysteine [37]. This groove is created due to the blocking loop 2 (BL2 loop), a flexible substrate-binding loop (Gly266-Gly271) found adjacent to the active site (Figure 9A and B). The BL2 loop is found in an open conformation in unbound PLpro, while it closes upon substrate or inhibitor binding [37].

Figure 9. PLpro three-dimensional structure (PDB code: 7LBR [40]) and surface view of the active site. The four domains, fingers (purple), palm (green), thumb (red), and Ubl (gray) are showed in the cartoon representation (A). The BL2 loop (orange, Gly266-Gly271) is indicated by a line. The substrate binding cleft is highlighted with the catalytic triad C111, H272, and D286 displayed as sticks (pink) (B). In the close-up view of the active site,
the residues are colored by the frequency they are involved in interactions with 21
crystallography ligands, according to an analysis using the program LUNA.

As observed for the crystallography ligands, compounds 189 and 195 showed
docking predicted binding modes occupying the S3 and S4 subsites (Figures 9, 10 and
Supporting Information S49). To verify the stability of these proposed binding modes,
compounds 189 and 195 underwent MD simulations. The XR8-89 ligand (PDB code
7LBR [40]) was used as a positive control.

In simulations with XR8-89, the BL2 loop remained in the closed conformation,
and the ligand binding mode remained stable in all simulations, with its core structure
being stabilized by hydrogen bond interactions between the carbonyl group and the
backbone of Q269 (100% of the analyzed simulation time), as well as π-stacking
interactions with Y268 (79% of the analyzed simulation time) (Figures 10 and Supporting
Information S50). We also observed a water bridge between the nitrogen on the amide
and D164 (subsite S3, present on average 22% of the analyzed trajectory).

For 189, four of the five simulated replicas showed stable interactions, with the
initial pose changing dramatically from the initial coordinates after 500 ns in one of the
replicas. In terms of binding mode, the triazole and central amine groups of 189 stablished
hydrogen bond interactions with the Q269 (35% of the analyzed trajectory) and π-based
interactions with Y264. The carbonyl groups from the naphthalene-1,4-dione moiety
displayed water-mediated interactions with D164. The 1,2-naphthoquinone ring, shown to
be essential for protease inhibition in our biochemical assays, binds to the S4 pocket,
establishing hydrophobic interactions to P248 (Figure 10B).

In contrast to the observed for XR8-89 and 189, the tolyl substituent on the amine
of 195 prevented stable simulations on BL2 closed conformation. Thus, we performed
1 μs simulation initially, which displayed at first few interactions with the sidechain of
Q269 (less than 20% of simulation time) and later stable interactions with D164 (over 66%
of simulation time), while leading to the opening of BL2 and accommodating of the
ligand. The last frame of this simulation was used to generate a further five new replicas
(5 x 500 ns), to analyze the stability of this new binding mode, which was shown to be
stabilized by water bridges with the D264 (>75%), R166 (>40%) and Y273 (~ 40%) and
hydrophobic contacts with P248 (>40%) (Figure 10).
Figure 10. Proposed binding modes and protein-ligand interactions profile for PLpro inhibitors. Representative frames from the MD simulation describing the potential binding mode (A-C) for compounds **XR8-89** (green, from PDB code 7LBR [40]), **189** (cyan) and **195** (orange) bound to PLpro. PLpro residues are colored according to the types of atoms in the interacting amino acid residues (protein carbon, light gray; nitrogen, blue; oxygen, red; sulfur, yellow), hydrogen bond interactions are represented as yellow dashed lines. Frequency of protein-ligand interactions for all simulations with ligands **XR8-89** (D), **189** (E) and **195** (F). Mean interaction frequency is represented, with standard error of the mean (N=5) interval depicted as error bars, each point displays the individual value for a particular simulation replica.
2.7. Evaluation of hit compounds in a SARS-CoV-2 viral infection assay

We evaluated two Mpro hit compounds 382 and 415 and two PLpro hit compounds 189 and 191 in a SARS-CoV-2 viral infection assay of monkey-derived Vero E6 cells. The clinically approved RNA-polymerase inhibitor, remdesivir was used as a positive control. Remdesivir displayed antiviral efficacy in Vero E6 cells with EC$_{50}$ of 2.45 µM and no host cell toxicity at concentrations up to 20 µM. Under the same culture conditions the hit naphthoquinone compounds were tested in three concentrations (24 µM, 6 µM, and 1.4 µM) and showed no significant antiviral activity dissociated from host cytotoxicity (Supporting Information Figure S51). We decided to test some compounds in serial dilution starting at 1 µM and in infected human-derived HeLa cells expressing ACE2 in addition to infected Vero cells. For HeLa-ACE2 cells, remdesivir was more potent with EC$_{50}$ of 40 nM, however, cell cytotoxicity was also noted at concentrations above 2.4 µM. At lower concentrations the naphthoquinone compounds had no significant antiviral activity up to 1 µM (Supporting Information Figure S52). Therefore, it is important to further study the mechanism of action to understand the cytotoxicity and decouple from the direct antiviral activity.

3. Discussion

In this study, we used computational and biochemical approaches to evaluate a library of quinones to find inhibitors against SARS-CoV-2 proteases. The wealth of structural information from Mpro and PLpro allowed us to generate patterns of common protein-ligand interactions, which were helpful in two stages of our computational analysis. First, the selection of computational hits was guided by protein-ligand interactions frequently observed in Mpro crystallographic complexes. Thus, we prioritized compounds that interacted with conserved water molecules, S1 and S2 residues, filling one or more of the subsites with minimum solvent exposure. This strategy was successful as, among 24 compounds selected for inhibitory assays, three molecules with two different scaffolds were confirmed as Mpro inhibitors with low micromolar or submicromolar potency (compounds 382, IC$_{50}$ of 0.41 µM; 379, IC$_{50}$ of 0.63 µM; and 415, IC$_{50}$ of 5.0 µM), and another four compounds inhibited the enzyme by more than 50% at 10 µM. Additionally, for Mpro and PLpro inhibitors that were evaluated experimentally, we conducted MD simulations of the protein-ligands complexes.
Together with the observed stability of binding poses during the simulations, the fact that our inhibitors establish interaction patterns commonly observed in the crystal structures encourages the application of our results in structure-based optimization projects.

During the validation of Mpro and PLpro inhibitors, several precautions were taken to avoid artifactual inhibition. We were especially careful considering previous reports that indicate quinones as potential Pan Assay Interference Compounds (PAINS) [133]. To avoid common causes of artifacts [128–130], we conducted the assays in the presence of detergent, avoiding compound aggregation, and verified that compounds were not highly fluorescent. In addition, a comparison of the inhibition of both target enzymes by each of the compounds indicates that all inhibitors showed specificity to one of our targets, reducing the likelihood they would be promiscuous inhibitors. Furthermore, to assess if Mpro inhibitors were time-dependent and/or irreversible inhibitors, we determined IC$_{50}$ values of compounds 382 and 415 upon or without preincubation with the enzyme, and evaluated recovery of enzyme activity in a reversibility assays. Our results indicated that compound 382 is a reversible Mpro inhibitor, while 415 binds irreversibly to the target. This information was taken into account in our MD simulations, in which compound 415 was covalently bound to the enzyme, while compound 382 was simulated in a noncovalent complex.

An interesting pattern emerged in our MD simulations with Mpro complexes. For both compounds 382 and 415, the simulations suggested stability of the complexes via multiple intermolecular interactions, with H$_{41}$, G$_{143}$, and E$_{166}$. All three residues have reported key roles in the active site. As part of the catalytic dyad, H$_{41}$ serves as a base for nucleophilic attack performed by C$_{145}$ in peptide-bond cleavage [134], while G$_{143}$, an oxyanion hole residue, helps stabilize the tetrahedral intermediate of the peptide-bond cleavage [135]. Moreover, E$_{166}$ is essential for dimerization and its interactions with the other protomer N-finger also aid the correct orientation of H$_{163}$ and H$_{172}$ to form the S1 pocket [135,136]. However, long-lasting interactions were observed only from one of the protomers’ binding sites, while the ligand bound to the other protomer was expelled within a few nanoseconds. The instability in one of the protomers was observed as a reproducible pattern in most replicates of our MD simulations. The complete deletion of the N-finger (residues S1* – R4*) in SARS-CoV Mpro, reduces the extent of the dimerization and completely abolishes the enzymatic activity (<1%) [137]. Simulations of Mpro from SARS-CoV-2 with peptidomimetic inhibitors or substrate [138], suggest that a similar mechanism exists, where the N-finger conformation upon dimerization
exerts a direct influence on the oxyanion-loop motions and the stabilization of the catalytic conformation. 

Our initial focus was on Mpro inhibition, however, we also tested the 23 soluble compounds selected against PLpro, to possibly find dual inhibitors for both SARS-CoV-2 viral enzymes. Despite the numerous efforts to develop inhibitors of the SARS-CoV-2 proteases, reports of dual Mpro/PLpro SARS-CoV-2 inhibitors are still scarce [139]. In the current study, the only dual inhibitor found was 191, with modest Mpro inhibition (IC$_{50}$ = 66 µM) and more potent PLpro inhibition (IC$_{50}$ = 3.1 µM). Developing an effective dual inhibitor would require further optimization, but compound 191 is a candidate for such efforts. In addition, eight ortho-quinone-based 1,2,3, triazoles had IC$_{50}$ < 50 µM against PLpro, including three inhibitors with IC$_{50}$ in the single-digit micromolar range (compounds 159, IC$_{50}$ of 1.7 µM; 189, IC$_{50}$ of 2.2 µM; and 191, IC$_{50}$ of 3.1 µM). Considering the evidence supporting each SARS-CoV-2 protease as a therapeutic target [37,140,141], these compounds are interesting even in the absence of dual inhibition. MD simulations of 189 and 195 bound to PLpro were not as stable as the positive control XR8-89 (PDB code 7LBR [40]). The two scaffolds interacted, with low or moderate frequency, with residues in the S3 and S4 subsites, however, lacking long-lasting interactions with key residues, such as Y$_{268}$ and Q$_{269}$. These two residues form an unusual β-turn in the flexible β-hairpin BL2 loop that controls the access to the active site in the binding of host and viral proteins [40]. Thus, compound 189 and, particularly, 195 might not fully stabilize the closed conformation of BL2 loop as well as the potent XR8-89.

4. Conclusion

Here, we employed computational and biochemical assays to evaluate a quinones library, leading to the identification of 11 promising naphthoquinoidal inhibitors against the two SARS-CoV-2 viral proteases, Mpro and PLpro, with potency in the mid micromolar to nanomolar range. For all inhibitors experimentally characterized, we propose likely binding modes with good complementarity to the protease active sites, that closely resemble protein-ligand interaction patterns observed in crystallographic complexes and which were stable in MD simulations. Hence, the inhibitors presented here are novel scaffolds for further optimization to develop a treatment against SARS-CoV-2 infection.
5. Experimental Section

5.1. Compounds general experimental details

All chemicals were obtained from commercial sources and used without further purification. Melting points were obtained on a Thomas Hoover apparatus and are uncorrected. Column chromatography was performed on silica gel (Silica Flash G60 UltraPure 60-200 µm, 60 Å). Infrared spectra were recorded on a Shimadzu FTIR Spectrometer IR Prestige-21. $^1$H and $^{13}$C NMR were recorded at room temperature using a Bruker AVANCE DRX 200 and DRX 400 MHz, in the solvents indicated, with tetramethylsilane (TMS) as internal reference. Chemical shifts (δ) are given in parts per million (ppm) and coupling constants (J) in Hertz (Hz). The mass spectrometer was operated in the positive ion mode. A standard atmospheric pressure photoionization (APPI) source was used to generate the ions. The sample was injected using a constant flow (3 µL/min). The solvent was an acetonitrile/methanol mixture. The APPI-Q-TOF MS instrument was calibrated in the mass range of 50-3000 m/z using an internal calibration standard (low concentration tuning mix solution) supplied by Agilent Technologies. Data were processed employing Bruker Data Analysis software version 4.0. Compounds were named following IUPAC rules as applied by ChemBioDraw Ultra (version 12.0).

5.2. Synthesis of candidate inhibitors

*Ortho*-quinone-based 1,2,3-triazoles compounds 159, 189, 191-197, were prepared according to previously reported reports and their data are consistent with the literature [80,82,91,142]. *Para*-quinones-based 1,2,3-triazoles compounds 314, 318-321 were prepared as described in the literature [90]. *Para*-quinones and derivatives 379, 380, 382, 414, 415, 465, 470, 477, were synthesized following the previously published studies in the literature [76,99]. Hydrazo derivatives 666, 668 and 673 were prepared according to previously published reports and their data are consistent with the literature [68]. NMR spectra for all compounds have been previously published when they were originally described.

5.3. Comparison of available SARS-CoV-2 Mpro structures
All 72 crystallographic structures were downloaded from the PDB [143] (structures available in April/2020). Structural superposition was performed with program R [144]’s package, Bio3D [145], using the protein’s Ca. RMSD and PCA were also done with Bio3D package. As water molecules might play important roles in Mpro catalysis and ligand stabilization, we used the ProBiS H2O plugin [123]. This PyMol [146] plugin enables the identification of conserved water sites in proteins using experimental determined protein structures. The highest resolution structure, PDB code 5R82 [117], was used as reference to establish the water molecules position.

5.4. Analysis of protein-ligand interactions

The program LUNA (https://github.com/keiserlab/LUNA) was used to perform large-scale analysis of non-covalent interactions between the protein-ligands complexes of Mpro. With this program, it was possible identify frequently interacting residues between the ligands and Mpro active site. We submitted a list containing the of PDB ids of the 72 structures, discriminating chain A and the binding site ligands to be analyzed. After processing, we investigated the table (in .csv format) of the interacting frequencies by residues and ligands with program R.

5.5. Ligand and protein preparation

Three-dimensional ligand structures were generated with LigPrep (version 46013), using Epik to predict their protonation in pH 7.0 ± 2.0, and generating tautomers and diastereoisomers. The OPLS3e force-field was employed for structure generation. The SARS-CoV-2 Mpro protein structure was prepared from the PDB 5R82 [117], using the Protein Wizard Preparation tool, with standard options. Two Mpro receptor files were prepared for docking: one with all water removed and another containing waters 1189 and 1284 from the original PDB. The SARS-CoV-2 PLpro structure was prepared from the PDB 7LBR [40], using the same protocol as Mpro.

5.6. Molecular docking

Molecular docking was carried out with Glide SP (version 9.1) and Autodock Vina. For docking to Mpro with Glide [124], grids were centered at the central point of
the active site residues G\textsubscript{143}, C\textsubscript{145}, M\textsubscript{49} and H\textsubscript{41} (coordinates: 10.7313390385, -4.49000171154, 22.4985591538). Two docking grids were generated: one without waters and one containing the two conserved waters described in the protein preparation. Each compound was docked using both grids. In both cases, the dimensions of the inner box had 10 Å in each direction and the outer box had 30 Å in each direction. Whenever mentioned, covalent docking as performed using CovDock \cite{147} using the C\textsubscript{145} as anchor, nucleophilic addition to double bond as reaction type and generating up to 10 poses for each ligand. Poses were selected according to the docking score and relevant interactions.

For docking to Mpro with Autodock Vina \cite{125}, a grid box of size 22.5x24.5x22.5 Å was centralized in the geometrical center among the residues T\textsubscript{26}, M\textsubscript{49}, N\textsubscript{142}, and M\textsubscript{165}. All the experiments were done in triplicate starting from a random seed. Energy range, exhaustiveness, and the number of maximum modes parameters were set to 3 kcal/mol, 8, 9, respectively. Similar to docking using Glide, two experiments were done with and without conserved waters. For selected ligands, induced-fit docking was performed (with and without the conserved waters) by flexing the residues N\textsubscript{142}, E\textsubscript{189}, M\textsubscript{49}, and M\textsubscript{165}.

For docking to PLpro with Glide, using the Induced-Fit mode \cite{148}, a cubic grid box of size 12 Å was centralized in the geometric center of the co-crystallized ligand (PDB code 7LBR \cite{40}).

5.7. Molecular Dynamics simulations

Prepared SARS-CoV-2 Mpro and PLpro structures were simulated with the selected ligands. MD simulations were carried out by using the Desmond engine \cite{149} with the OPLS3e force-field \cite{150} according to a previously described protocol \cite{151}. In short, the system encompassed the protein-ligand/cofactor complex, a predefined water model (TIP3P \cite{152}) as a solvent and counterions (Na\textsuperscript{+} or Cl\textsuperscript{-} adjusted to neutralize the overall system charge). The system was treated in a cubic box with a periodic boundary condition (PBC) specifying the shape and the size of the box as 13 Å distance from the box edges to any atom of the protein. Short-range coulombic interactions were calculated using 1 fs time steps and 9.0 Å cut-off value, whereas long-range coulombic interactions were estimated using the Smooth Particle Mesh Ewald (PME) method \cite{153}. Each Mpro and PLpro systems were subjected to at least 5 μs simulations (five replicas of 1 μs each), with exception of PLpro – compound \textbf{195}, which had one preliminary 1 μs simulation, from which a stable conformation was selected for further shorter simulations. Atomic
interactions and distances were determined using the Simulation Event Analysis pipeline as implemented in Maestro 2020.2 (Schrödinger LCC).

Representative frames of the simulations were retrieved from clustering, which was performed with hierarchical clustering analyses. Trajectories were clustered using the script trj_cluster.py (implemented in Maestro 2021.2, Schrödinger LCC) using 2 Å as cut-off, which was chosen upon evaluating the RMSD of ligand’s heavy atoms. Trajectories where the ligand was expelled of the pocket were not considered for clustering or interaction analyses.

RMSD values of the protein backbone were used to monitor simulation equilibration and protein folding changes (Supporting Information Figure S50). All the trajectory and interaction data are available on the Zenodo repository (code: 10.5281/zenodo.5147951). MD trajectories were visualized, and figures produced by PyMol v.2.4 (Schrödinger LCC, New York, NY, USA).

5.8. Synthesis of Mpro substrate

A quenched fluorogenic peptide substrate with the sequence (D-Arg)-(D-Arg)-Lys(MCA)-Ala-Thr-Leu-Gln-Ala-Ile-Ala-Ser-Lys(DNP)-COOH (ATLQAIAS) was synthesized on a Biotage Syroll peptide synthesizer at room temperature through fluorenylmethyloxycarbonyl (Fmoc) solid-phase synthesis. The synthesis scale was 12.5 μmole with preloaded lysine(2-dinitrophenyl) Wang resin, where the DNP quencher was linked to the epsilon nitrogen of the lysine. For each coupling reaction, 4.9 equivalents of HCTU (O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate), 5 equivalents of Fmoc-amino acid-OH, and 20 equivalents of N-methylmorpholine (NMM) in 500 μL N,N-dimethylformamide (DMF) were used. The coupling reaction was carried out with shaking for 8 minutes. Each amino acid position was double coupled, and subsequent Fmoc deprotection was done with 500 μL of 40% 4-methylpiperidine in dimethyl formamide (DMF) for 10 minutes. Deprotection was followed by a wash with 500 μL of DMF for 3-minutes and the wash was repeated 6 times. The lysine amino acid, lysine (7-methoxycoumarin-4-acetic acid (MCA), was coupled where MCA was linked to the epsilon nitrogen of the lysine. The two final amino acid position couplings used d-Arginine to increase peptide solubility. The cleavage of the peptide from the Wang resin was carried out with a 500 μL of solution composed of
95% trifluoroacetic acid, 2.5% water, and 2.5% triisopropylsilane at room temperature for 1 hour with shaking. The crude peptide product was precipitated in 30 mL of a 1:1 mixture of cold diethyl ether and hexane. Product was then solubilized in a 1:1:1 mixture of DMSO, water and acetonitrile. The solubilized crude material was purified by high-performance liquid chromatography (HPLC) using an Agilent Pursuit 5 C18 column (5 mm bead size, 150 x 21.2 mm) on an Agilent PrepStar 218 series preparative HPLC. Mobile phase A was water + 0.1% TFA, and mobile phase B was acetonitrile + 0.1% TFA. The peptide product fractions were collected, combined, and had solvent removed under reduced atmosphere. The peptide substrate was solubilized in DMSO to a final concentration of 50 mM. Purity was confirmed by liquid chromatography-mass spectrometry and the stock was stored at -20°C.

5.9. Assays against Mpro

Recombinant SARS-CoV-2 Mpro was expressed and purified as described previously in Mellot et al. [48]. Mpro activity was measured using the fluorogenic substrate, ATLQAIAS, on a Biotek® Synergy HTX plate reader. All assays were performed in black flat-bottom 384-well plates, in 30 µL of 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.01% Tween-20 using 50 nM Mpro and 10 µM of FRET substrate. Initial screening was performed at 10 µM. Prior to addition of the substrate, enzyme was incubated with the compounds for 15 minutes. Following the substrate addition proteolysis was measured at 320/420 nm (excitation/emission) at 25 °C. Percent inhibition was calculated relative to control reactions containing a maximum of 0.5% DMSO. Two independent experiments were performed in triplicate wells. Half-maximal inhibitory concentration (IC50) was determined by nonlinear regression analysis of the velocity vs. inhibitor concentration plot using GraphPad Prism 6 (GraphPad Prism, version 6.00, La Jolla, California, USA). At least seven inhibitor concentrations were used to build each curve. DMSO was used as negative control. The hit compounds 382 and 415 were also tested without incubation to investigate the time-dependency behavior.

5.10. Reversibility assay

Mpro at 100-fold its final assay concentration was incubated with the hits at 10-fold its respective IC50 value for 30 min in a volume of 2 µL. This mixture was diluted
100-fold with an assay buffer containing 10 μM ATLQAIAS substrate to a final volume of 30 μL, resulting in a standard concentration of Mpro and 0.1 times the IC\textsubscript{50} value of hits [130,131]. Fluorescence intensities were monitored continuously during substrate hydrolysis on Synergy 2 (BioTek\textsuperscript{®}) plate reader for 120 minutes.

5.11. Assays against PLpro

Recombinant SARS-CoV-2 PLpro was purchased from Acro Biosystems, PAE-C5184. Proteolytic activity was measured using Z-Arg-Leu-Arg-Gly-Gly-AMC substrate (Bachem, 369 I1690) as described previously in Ashhurst et al. [154] The release of fluorescent 7-amido-4-methylcoumarin was measured at 360 nm/460 nm wavelengths for excitation/emission, on a Biotek\textsuperscript{®} Synergy HTX. All assays were performed in 384-well black plate at 25 °C, in a final volume of 30 μL of 50 mM HEPES pH 6.5, 150 mM NaCl, 0.1 mM DTT, 0.01% Tween-20, 50 nM enzyme and 50 μM of substrate. Enzymatic activity was calculated by comparison to initial rates of reaction of a DMSO control.

Initial screening was performed at 10 μM of each compound in triplicate wells. Compounds that inhibited by 75% or more of the PLpro activity in the initial screen had their IC\textsubscript{50} determined. At least two independent experiments were performed, each involving at least eleven compound concentrations in triplicates. IC\textsubscript{50} curves were obtained by non-linear regressions analysis of the velocity vs. inhibitor concentration using GraphPad Prism 6 (GraphPad Prism, version 6.00, La Jolla, California, USA). Reported IC\textsubscript{50} values refer to the mean values and the standard error of the mean.

5.12. Antiviral activity

Antiviral assays were performed according to the protocol previously described in Mellot et al. [48]. Compounds 189, 191, 382, and 415 were evaluated in a SARS-CoV-2 viral infection assay of monkey-derived Vero E6 cells and human-derived HeLa cells that overexpress ACE2. Remdesivir was employed as a positive control. Each compound was evaluated in ten concentrations, in two-fold dilutions, from 20 μM to 39 nM in the case of remdesivir and from 1.0 μM to 1.9 nM for all other compounds, in triplicates.

6. Author Contributions
L.H.S. analyzed PDB structures and interaction patterns; L.H.S., R.E.O.R., and R.S.F performed and analyzed virtual screening results; T.K. and A.P. performed and analyzed molecular dynamics simulations; R.G.A. and E.N.S.J. designed the chemical library for virtual screening; C.B., A.L.L. and C.S.C. designed and synthesized the Mpro substrate. D.S., E.B.S. and A.J.O. performed Mpro and PLpro assays; J.C.O., L.V.B. and T.K. performed docking studies; P.F., D.S., L.M.P., J.H.M. and A.J.O. expressed and purified Mpro; M.A.G., B.W., and J.L.S.N. performed antiviral assays. All authors were involved in experiment design and analyses; L.H.S., T.K., R.G.A., E.B.S., E.N.S.J., and R.S.F wrote the manuscript, with revisions and contributions from all authors. E.N.S.J. and R.S.F conceived the overall design of the study.

7. Declaration of Competing Interest

L. H. Santos, R. G. Almeida, E. Barbosa da Silva, A. O'Donoghue, E. N. da Silva Júnior and R. S. Ferreira are inventors on a pending patent related to technology described in this work.

8. Acknowledgments

The authors would like to thank CNPq, CAPES (Finance Code 001) and FAPEMIG for the financial support and scholarships. E. N. da Silva Júnior acknowledges funding from CNPq (PQ 309774/2020-9), Fapemig (Rede de Pesquisa e Inovação para Bioengenharia de Nanossistemas-RED-00282-16 and PPM-00635-18), Return Fellowship of the Alexander von Humboldt Foundation (AvH) and the Royal Society of Chemistry for the research fund grant (R19-9781). R.S.F. received funding from CAPES (grant CAPES-EPIDEMIAS-0688/2020), FAPEMIG (Rede Mineira de Imunobiologicos grant # REDE-00140-16) and holds a CNPq Researcher Scholarship (Bolsa de Produtividade em Pesquisa, 306606/2017-8). J.C.O. and L.V.B. received scholarships from CAPES (processes 88887.508402/2020-00 and 88887.518393/2020-00, grant CAPES-EPIDEMIAS - Programa Estratégico Emergencial de Prevenção e Combate a Surtos, Endemias, Epidemias e Pandemias). C.B., A.L.L. and C.S.C. were supported by National Institutes of Health grant P50AI150476. The SARS-CoV-2 Mpro plasmid was provided by Rolf Hilgenfeld, University of Lübeck, Germany. Authors would also like to thank the CSC-Finland for the generous computational resources provided.
9. Appendix A. Supplementary data

Supplementary Information that contains general information of the assembly of the chemical library, structural information of all virtual screened compounds, detailed information of the computational procedures is available for this manuscript, and SARS-CoV-2 viral infection assay.

10. References


N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W.
Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan,
A Novel Coronavirus from Patients with Pneumonia in China, 2019, N. Engl. J.

Maltseva, S. Chen, V. Nicolaescu, S. Dvorkin, K. Furlong, R.S. Kathayat, M.R.
Firpo, V. Mastrodomenico, E.A. Bruce, M.M. Schmidt, R. Jedrzejczak, M.Á.
Muñoz-Alia, B. Schuster, V. Nair, K. Han, A. O’Brien, A. Tomatisidou, B.
Baker, B.C. Mounce, N.S. Heaton, W.E. Severson, K.E. Palmer, B.C. Dickinson,
A. Joachimiak, G. Randall, S. Tay, Masitinib is a broad coronavirus 3CL
inhibitor that blocks replication of SARS-CoV-2, Science (80-. ). (2021)
eabg5827. https://doi.org/10.1126/science.abg5827.

T.T. Huang, M. Bekes, M. Drag, S.K. Olsen, Activity profiling and crystal
structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework
https://doi.org/10.1126/sciadv.abd4596.

S. Ullrich, C. Nitsche, The SARS-CoV-2 main protease as drug target,

Y.M. Báez-Santos, S.E. St. John, A.D. Mesecar, The SARS-coronavirus papain-
like protease: Structure, function and inhibition by designed antiviral compounds,

C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li,
X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-
2 and discovery of potential drugs by computational methods, Acta Pharm. Sin.

Y. Kim, H. Liu, A.C. Galasiti Kankanamalage, S. Weerasekara, D.H. Hua, W.C.
Groutas, K.O. Chang, N.C. Pedersen, Reversal of the Progression of Fatal
Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease
https://doi.org/10.1371/journal.ppat.1005531.
https://doi.org/10.1021/acs.jmedchem.9b01828.

https://doi.org/10.1371/journal.pbio.0030324.

https://doi.org/10.1093/emboj/cdf327.


https://doi.org/10.1126/science.abb4489.
Zhang, Y. Hu, N. Kitamura, B. Hurst, B. Tarbet, M.T. Marty, A. Kolocouris, Y.
Xiang, Y. Chen, J. Wang, Structure and inhibition of the SARS-CoV-2 main
protease reveal strategy for developing dual inhibitors against Mpro and
https://doi.org/10.1126/sciadv.abe0751.

Yen, T.T. Peng, S.E. Chang, C.T. Huang, Y.R. Tsai, S.H. Wu, S.J. Lee, J.H. Lin,
B.S. Liu, W.C. Sung, S.R. Shih, C.T. Chen, J.T.A. Hsu, Discovery of M protease

Rox, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a
basis for design of improved a-ketoamide inhibitors, Science (80-. ). 368 (2020)

[23] C. Ma, M.D. Sacco, B. Hurst, J.A. Townsend, Y. Hu, T. Szeto, X. Zhang, B.
Tarbet, M.T. Marty, Y. Chen, J. Wang, Boceprevir, GC-376, and calpain
inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral
main protease, Cell Res. 30 (2020) 678–692. https://doi.org/10.1038/s41422-
020-0356-z.

Peng, Structure of M pro from SARS-CoV-2 and discovery of its inhibitors,

Isaacs, B. Lindenbach, S.J. Miller, K.S. Anderson, W.L. Jorgensen, Potent
Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular
Sculpting of the Drug Perampanel Guided by Free Energy Perturbation
Calculations, ACS Cent. Sci. 7 (2021) 467–475.
https://doi.org/10.1021/acscentsci.1c00039.

Shang, C. Ke, L. Feng, X. Jiang, J. Shen, G. Xiao, H. Jiang, L. Zhang, Y. Ye, Y.
Xu, Identification of pyrogallol as a warhead in design of covalent inhibitors for
https://doi.org/10.1038/s41467-021-23751-3.


https://doi.org/10.1016/j.ejmech.2017.05.011.


[60] E.N. da Silva Júnior, M.C.B.V. de Souza, M.C. Fernandes, R.F.S. Menna-
Barreto, M. do C.F.R. Pinto, F. de Assis Lopes, C.A. de Simone, C.K.Z.

Andrade, A. V. Pinto, V.F. Ferreira, S.L. de Castro, Synthesis and anti-

Trypanosoma cruzi activity of derivatives from nor-lapachones and lapachones,


https://doi.org/10.1016/j.bmc.2008.03.032.


https://doi.org/10.1021/jm900865m.


https://doi.org/10.1016/j.ejmech.2015.06.044.


1226 Cavalcanti, C. Pessoa, I.N.N. Namboothiri, E.N. da Silva Júnior, Quinonoid
1227 compounds via reactions of lawsone and 2-aminonaphthoquinone with α-
1228 bromonitroalkenes and nitroallylic acetates: Structural diversity by C-ring
1231 [74] H. Suginome, A. Konishi, H. Sakurai, H. Minakawa, T. Takeda, H. Senboku, M.
1232 Tokuda, K. Kobayashi, Photoinduced molecular transformations. Part 156. New
1233 photoadditions of 2-hydroxy-1,4-naphthoquinones with naphthols and their
1235 4020(94)01026-V.
1237 Pereira, K. Salomão, R.F.S. Menna-Barreto, I.N.N. Namboothiri, J.F. Bower,
1238 E.N. da Silva Júnior, Strategies towards potent trypanocidal drugs: Application of
1239 Rh-catalyzed [2 + 2 + 2] cycloadditions, sulfonyl phthalide annulation and
1240 nitroalkene reactions for the synthesis of substituted quinones and their
1241 evaluation against Trypanosoma cruzi, Bioorganic Med. Chem. 28 (2020)
1244 Rogge, P.M.S. Costa, C. Pessoa, C. Jacob, L. Ackermann, Ruthenium(II)-
1245 Catalyzed Double Annulation of Quinones: Step-Economical Access to Valuable
1247 https://doi.org/10.1002/chem.202001434.
1249 Synthesis and biological evaluation of novel 1,4-naphthoquinone derivatives as
1253 Wills, S. Carpenter, G.A. Kraus, M.S. Gordon, J.W. Petrich, Synthesis of
1254 hydroxy and methoxy perylene quinoines, their spectroscopic and computational


V. Grum-Tokars, K. Ratia, A. Begaye, S.C. Baker, A.D. Mesecar, Evaluating the

https://doi.org/10.1016/j.virusres.2007.02.015.


L.S. Franco, R.C. Maia, E.J. Barreiro, Identification of LASSBio-1945 as an inhibitor of SARS-CoV-2 main protease (MPRO) through in silico screening
supported by molecular docking and a fragment-based pharmacophore model,


[120] J. Gossen, S. Albani, A. Hanke, B.P. Joseph, C. Bergh, M. Kuzikov, E. Costanzi,
1463 C. Manelfi, P. Storici, P. Gribov, A.R. Beccari, C. Talarico, F. Spyrokis, E.
1464 Lindahl, A. Zaliani, P. Carloni, R.C. Wade, F. Musiani, D.B. Kokh, G. Rossetti,
1465 A Blueprint for High Affinity SARS-CoV-2 Mpro Inhibitors from Activity-
1466 Based Compound Library Screening Guided by Analysis of Protein Dynamics,
1468 https://doi.org/10.1021/acsptscsi.0c00215.
1469

1470 assessment of the important residues involved in the dimerization and catalysis of
1472 https://doi.org/10.1371/journal.pone.0144865.

[122] J. Ziebuhr, Molecular biology of severe acute respiratory syndrome coronavirus,
1476 https://doi.org/10.1016/j.mib.2004.06.007.
1477

1479 3103. https://doi.org/10.1021/acs.jcim.7b00443.
1480

[124] T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard,
1481 J.L. Banks, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2.
1483 https://doi.org/10.1021/jm030644s.

[125] O. Trott, A.J. Olson, AutoDock Vina: Improving the speed and accuracy of
1484 docking with a new scoring function, efficient optimization, and multithreading,

1488 of host cellular protein substrates of sars-cov-2 main protease, Int. J. Mol. Sci. 21

[127] S.L. McGovern, B.T. Helfand, B. Feng, B.K. Shoichet, A specific mechanism of
1492 https://doi.org/10.1021/jm030266r.

[128] B.Y. Feng, B.K. Shoichet, A detergent-based assay for the detection of


Shen, Severe acute respiratory syndrome coronavirus 3C-like proteinase N
terminus is indispensable for proteolytic activity but not for enzyme
dimerization: Biochemical and thermodynamic investigation in conjunction with
https://doi.org/10.1074/jbc.M408211200.

[138] D. Suárez, N. Díaz, SARS-CoV-2 main protease: a molecular dynamics study, J.

D. Das, B.J. Anson, E.K. Lendy, Y. Takamatsu, N. Takamune, N. Kishimoto, K.
Murayama, K. Hasegawa, M. Li, D.A. Davis, E.N. Kodama, R. Yarchoan, A.
Wlodawer, S. Misumi, A.D. Mesecar, A.K. Ghosh, H. Mitsuya, A small
molecule compound with an indole moiety inhibits the main protease of SARS-
https://doi.org/10.1038/s41467-021-20900-6.

[140] W.M. Singh, J.B. Baruah, Synthesis of mixed aryl 2,3-diarylsulphonyl-1,4-
https://doi.org/10.1080/00397910802528951.

Duan, Q.C. Zou, H.L. Zhang, W. Qu, Y.H.P. Long, M.H. Li, R.C. Yang, X. Yang,
J. Zou, L. Li, Y. Hu, G.W. Lu, W.M. Li, Y.Q. Wei, Y.T. Zheng, J. Lei, S. Yang,
SARS-CoV-2 Mpro inhibitors with antiviral activity in a transgenic mouse
https://doi.org/10.1126/science.abf1611originally.

[142] D.J.B. Lima, R.G. Almeida, G.A.M. Jardim, B.P.A. Barbosa, A.C.C. Santos,
W.O. Valença, M.R. Scheide, C.C. Gatto, G.G.C. de Carvalho, P.M.S. Costa, C.
Pessoa, C.L.M. Pereira, C. Jacob, A.L. Braga, E.N. da Silva Júnior, It takes two
to tango: synthesis of cytotoxic quinones containing two redox active centers
https://doi.org/10.1039/d1md00168j.

Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000)
[144] Core R Team, A Language and Environment for Statistical Computing, R Found.

[145] B.J. Grant, A.P.C. Rodrigues, K.M. ElSawy, J.A. McCammon, L.S.D. Caves,
Bio3d: an R package for the comparative analysis of protein structures,


Docking covalent inhibitors: A parameter free approach to pose prediction and
https://doi.org/10.1021/ci500118s.

553. https://doi.org/10.1021/jm050540c.

Klepeis, I. Kolossvary, M.A. Moraes, F.D. Sacerdoti, J.K. Salmon, Y. Shan, D.E.
Shaw, Scalable Algorithms for Molecular Dynamics Simulations on Commodity
43. https://doi.org/10.1109/sc.2006.54.

[150] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J.Y. Xiang, L. Wang, D.
Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L.
Jorgensen, R. Abel, R.A. Friesner, OPLS3: A Force Field Providing Broad

Poso, SARS-COV-2 Mpro conformational changes induced by covalently bound

[152] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein,
Comparison of simple potential functions for simulating liquid water, J. Chem.

https://doi.org/10.1063/1.464397.