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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Inhibition of the SARS-CoV-2 3CL\textsuperscript{pro} main protease by plant polyphenols

Miha Bahun\textsuperscript{a,1}, Marko Jukič\textsuperscript{b,c,1}, Domen Oblak\textsuperscript{d}, Luka Kranjc\textsuperscript{a}, Gregor Bajč\textsuperscript{e}, Matej Butala\textsuperscript{e}, Krištof Bozovičar\textsuperscript{f}, Tomaž Bratkovič\textsuperscript{f}, Crtomir Podlipnik\textsuperscript{d,g}, Nataša Poklar Ulrih\textsuperscript{a,g,*}

\textsuperscript{a} Department of Food Science and Technology, Biotechnical Faculty, University of Ljubljana, SI-1000 Ljubljana, Slovenia
\textsuperscript{b} Laboratory of Physical Chemistry and Chemical Thermodynamics, Faculty of Chemistry and Chemical Engineering, University of Maribor, SI-2000 Maribor, Slovenia
\textsuperscript{c} Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, SI-6000 Koper, Slovenia
\textsuperscript{d} Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, SI-1000 Ljubljana, Slovenia
\textsuperscript{e} Department of Biology, Biotechnical Faculty, University of Ljubljana, SI-1000 Ljubljana, Slovenia
\textsuperscript{f} Department of Pharmaceutical Biology, Faculty of Pharmacy, University of Ljubljana, SI-1000 Ljubljana, Slovenia
\textsuperscript{g} Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins (CIPKeBiP), SI-1000 Ljubljana, Slovenia

ARTICLE INFO

Keywords:
Edible plants
Polyphenols
COVID-19
SARS-CoV-2
3CL\textsuperscript{pro}
Protease inhibitor

Chemical compounds studied in this article:
caffeic acid (PubChem CID: 689043)
chlorogenic acid (PubChem CID: 1794427)
curcumin (PubChem CID: 969516)
elgagic acid (PubChem CID: 5281855)
epigallocatechin gallate (PubChem CID: 65064)
ferulic acid (PubChem CID: 445858)
galangin (PubChem CID: 5281616)
4-hydroxy-3-nitrophenoxyacetic acid (PubChem CID: 447364)
isorhamnetin (PubChem CID: 5281654)
kaempferol (PubChem CID: 5280863)
kaempferol-rutinoside (PubChem CID: 531767)
phlorizin (PubChem CID: 6072)
protocatechuic acid (PubChem CID: 72)
quercetin (PubChem CID: 5280343)
quinic acid (PubChem CID: 6508)
resveratrol (PubChem CID: 445154)
sinapic acid (PubChem CID: 637775)
syringic acid (PubChem CID: 10742)
umbelllic acid (PubChem CID: 446611)

ABSTRACT

The abundance of polyphenols in edible plants makes them an important component of human nutrition. Considering the ongoing COVID-19 pandemic, a number of studies have investigated polyphenols as bioactive constituents. We applied \textit{in-silico} molecular docking as well as molecular dynamics supported by \textit{in-vitro} assays to determine the inhibitory potential of various plant polyphenols against an important SARS-CoV-2 therapeutic target, the protease 3CL\textsuperscript{pro}. Of the polyphenols in initial \textit{in-vitro} screening, quercetin, ellagic acid, curcumin, epigallocatechin gallate and resveratrol showed IC\textsubscript{50} values of 11.8 µM to 23.4 µM. \textit{In-silico} molecular dynamics simulations indicated stable interactions with the 3CL\textsuperscript{pro} active site over 100 ns production runs. Moreover, surface plasmon resonance spectroscopy was used to measure the binding of polyphenols to 3CL\textsuperscript{pro} in real time. Therefore, we provide evidence for inhibition of SARS-CoV-2 3CL\textsuperscript{pro} by natural plant polyphenols, and suggest further research into the development of these novel 3CL\textsuperscript{pro} inhibitors or biochemical probes.
1. Introduction

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which was first reported in December 2019. In the early months of 2020, COVID-19 spread worldwide, to cause a pandemic (Wang, et al., 2020). This first became a global health problem, and then later a socioeconomic problem, and it has claimed more than 4 million lives to date. Importantly, together with the relevant authorities, the medical and academic communities responded quickly, with intensive research campaigns that were accompanied by proposals for social guidelines, hygiene practices and the dissemination of quality information (Fry et al., 2020).

The SARS-CoV-2 virus belongs to the *Coronaviridae* family, and is named after its ‘crown’ serrations (Kahn & McIntosh, 2005). It is a single-stranded positive-sense RNA virus (Zhu et al., 2020), and its virulence ranges from a common cold to serious pathogenic potential, as demonstrated by the outbreaks of SARS in mainland China and Hong Kong in 2003, and of Middle East respiratory syndrome in Saudi Arabia, mainland China, United Arab Emirates and the Republic of Korea in 2012 (de Wit et al., 2016). In addition, swine acute diarrhoea syndrome coronavirus has been described in association with the plethora of animal diseases caused by coronaviruses, and it causes highly pathogenic severe acute porcine diarrhoea syndrome (Yang, Yu et al., 2020).

Therefore, the exploration of bioactive constituents in foods and the development of novel therapeutic options and drugs are essential for the future control of coronavirus infections (Sarkar et al., 2020; Khan, Umbreen et al., 2021).

Polyphenols are a structurally diverse group of aromatic compounds that contain multiple hydroxyl groups (Abbas et al., 2017). These naturally occurring compounds are produced through secondary metabolism in plants, where they are involved in defence mechanisms against pathogens and UV radiation (Treutter, 2006). Due to the abundance of polyphenols in edible plants, they constitute an important part of the human diet (Williamson, 2017). In the past, several studies have suggested positive effects of plant polyphenols on human health, as different polyphenols have demonstrated antioxidant, anti-inflammatory, antitumour, anti-allergic and antimicrobial activities (Rasouli et al., 2017).

Polyphenols have also been shown to have antiviral activities and even studied in the context of SARS-CoV-2 main protease (Khan, Heng et al., 2021). Different mechanisms of attenuation of viral infections by polyphenols have been demonstrated, such as impairment of the viral replication cycle (Kim et al., 2010) and prevention of virus entry through blocking attachment of viral particles to host cells (Calland et al., 2015). In the light of the current COVID-19 pandemic, a number of studies have considered the polyphenols as possible natural remedies to combat SARS-CoV-2 infection (Méhany et al., 2021, Paraiso et al., 2020, El-Missiry et al., 2021). However, the vast majority of studies that have investigated the interactions of polyphenols with SARS-CoV-2 proteins have only applied in-silico methods. Nevertheless, selected polyphenols have been shown to interact with different SARS-CoV-2 proteins in-vitro, including the receptor binding domain of the Spike glycoprotein (Goc et al., 2021), and the 3CL\textsuperscript{pro} which are among the most studied SARS-CoV-2 drug targets.

3CL\textsuperscript{pro} (cysteine protease; EC 3.4.22.69) in particular is crucial for the cleavage of coronavirus polyproteins to form mature non-structural proteins that are themselves essential for viral replication mechanisms. Proteolysis itself occurs via a catalytic dyad defined by Cys145 and His41 and this enzyme has been successfully targeted by a multitude of medicinal chemistry approaches (Dai et al., 2020). In this work we have conducted a multitude of molecular dynamics experiments on a complete set of in-house available polyphenols supported by in-vitro biological evaluation.

The polyphenols that have shown in-vitro SARS-CoV-2 3CL\textsuperscript{pro} inhibitory activity to date include epigallocatechin gallate (EGCG) (Du et al., 2021, Chiou et al., 2021, Ni et al., 2021), quercetin (Rizzuti et al., 2021), naringenin and kaempferol (Du et al., 2021). In the present study, we screened several additional plant polyphenols for inhibition of SARS-CoV-2 3CL\textsuperscript{pro} using a combination of in-vitro and in-silico approaches (Fig. 1).

---

**Fig. 1.** Structures of the polyphenols used in this study. The polyphenols are divided into seven groups, according to their structures: phenolic acids, hydroxycinnamic acids, flavonoids, curcuminoids, stilbenes and other polyphenols.
2. Materials and methods

2.1. Polyphenol library preparation and molecular docking

A selection of 19 plant polyphenols that we study in-house and have available as physical samples (Fig. 1) was obtained via the PubChem Compound Database as SMILES strings. They were checked for structural errors, and their three-dimensional (3D) structures were calculated using the Schrödinger ligand preparation workflow and the Ligprep software (release 2021–1; Schrödinger LLC, New York, NY, USA). The final 3D conformations included enumeration of chiral centres and tautomeric structures, ionisation at pH 7.4, and minimisation using the OPLS3e force field. In the end, a focused library of 39 polyphenols was obtained, with these expanded in the subsequent docking and consensus scoring to increase the ranking power and to gain information on the predicted binding modes (Houston & Walkinshaw, 2013).

The docking software GlideXP (Friesner et al., 2006) was used for the consensus scoring (release 2021–1; Schrödinger LLC, New York, NY, USA), along with AutoDock VINA (Trott & Olson, 2010) implemented in the Yasara structure (release 20.12.24) and CmDock (version 0.1.3; https://gitlab.com/Jukic/cmdock). The 3D structure of the SARS-CoV-2 3CLpro was obtained from the Protein Data Bank (PDB ID: 6M2N). The structure of the 3D protein complex was solved at 2.20 Å using X-ray diffraction, as a homodimer with an inhibitor (3WL; polyphenolic; 5,6,7-trihydroxy-2-phenyl-4H-chromen-4-one) bound to the active sites (one binding site per subunit). The crystal structure active site residue positioning has been carefully checked against all available similar structures and the HOLO structure was chosen in order to provide accessible “open” conformation of the catalytic active site. The multistep Schrödinger protein preparation workflow (release 2021-1; Schrödinger LLC, New York, NY, USA) was used for the final preparation of the molecular docking with the receptor. The prepared protein structure was free of any ligand and water molecules, hydrogens were added, the H-bonding network was optimised, and the final structure was edited with restrained minimisation of heavy atoms towards 0.3 Å in the OPLS3e force field. For the CmDock docking package, a docking receptor was generated using CmCavity (Ruiz-Carmona et al., 2014).

Initially, to evaluate all possible binding sites, additional blind docking was performed using AutoDock VINA implemented in the Yasara structure (release 20.12.24) (200 runs). As expected, the regions of the catalytic site in the vicinity of the Cys145 residue of the SARS-CoV-2 3CLpro were most favoured for the ligand binding. Therefore, a cubic simulation cell was constructed with 10 Å extensions along each axis from the centre of the catalytic site, and docking into the catalytic site was repeated (200 runs, 32 exhaustiveness parameters). Next, molecular docking was performed using both the GlideXP software (GlideScore version XP5.0; 3 poses per ligand) and the CmDock software (https://gitlab.com/Jukic/cmdock; 100 runs in exhaustive docking mode).

2.2. Molecular dynamics

The most favoured complexes from the molecular docking were prepared using the Yasara structure software (release 20.12.24). The proteins were capped (N-terminal acetyl group; C-terminal N-methylamide group) and residue ionisation was assigned at pH 7.4 (Krieger & Vriend, 2015). A cubic box (10 Å away from all atoms) was solvated with the TIP3P water model and 0.9% NaCl to define a physiological solution. After removing bumps/clashes using steepest descent minimisation, annealing minimisation was used to reach a stable local energy minimum. From here, the simulation was started by assigning random initial velocities and slowly heating the system to 298 K. A molecular dynamics simulation was started using the AMBER14 force field for the solute, GAFF and AM1BCC charges for the ligands, and TIP3P for the water. The temperature was set to 298 K at a pressure of 1 atm (NPT ensemble, Berendsen barostat and thermostat, coupled to time-averaged temperature and density), with snapshots saved every 100 ps. Non-bonded long-range interactions were calculated using the Particle Mesh Ewald algorithm. The SHAKE algorithm was not used, and the energy of the system was stable throughout the production run, as was the root-mean-square deviation of the protein backbone for all of the systems (details in Supplementary Materials Fig. S1). The production run was performed over a period of 110 ns, with the first 10 ns serving to equilibrate the system.

2.3. Cloning, expression and purification of recombinant 3CLpro

The codon-optimised gene encoding SARS-CoV-2 3CLpro with an N-terminal 6 × His tag and a TEV protease cleavage site was obtained (Integrated DNA Technologies, USA) and cloned into the pET-28c(-) plasmid between the Ncol and NotI restriction sites. The resulting expression vector was used to transform chemically competent Escherichia coli Nco21(DE3) (New England Biolabs, USA). The transformed bacteria were cultured at 37 °C in lysogeny broth supplemented with 50 µg/mL kanamycin, with agitation at 250 rpm. When an OD600 of ~ 1.8 was reached, the cultures were cooled to 16 °C and 3CLpro expression was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside. The cultures were then incubated at 16 °C with agitation at 250 rpm for 24 h. The cells were harvested by centrifugation (3000 × g for 20 min at 4 °C), and the cell pellets were stored at −80 °C until purification of the 3CLpro.

For purification of the recombinant product, the cell pellets were resuspended in buffer A (20 mM Tris, pH 7.5, 0.05 mM ethylendiaminetetraacetic acid (EDTA), 2.5 mM dithiothreitol (DTT), 10% glycerol). The cells were lysed by sonication on ice, and the insoluble cellular debris was removed by centrifugation (15000 × g for 10 min at 4 °C repeated twice). The clarified supernatant was then filtered to a 100-kDa molecular weight cut-off (Amicon Ultra-15 centrifugal filter units; Merck, Germany). The filtrate was combined with 0.5 M (NH4)2SO4 (final concentration) and loaded onto a 1 mL Hitrap Phenyl HP column (Cytiva, USA) that was pre-equilibrated with buffer B (50 mM Tris, pH 7.5, 0.5 M (NH4)2SO4, 0.05 mM EDTA, 2.5 mM DTT, 10% glycerol). The column was then washed with 15 mL buffer B and the 3CLpro protein was eluted from the column using a linear gradient of 0% to 100% buffer A in buffer B. The 3CLpro protein eluted as a broad peak at ~ 70% buffer A. The eluted 3CLpro was concentrated with a 30-kDa molecular weight cut-off (Ultra-4 centrifugal filter units; Amicon), snap-frozen in liquid nitrogen, and stored at −80 °C. The purity of the isolated 3CLpro was assessed by SDS-PAGE using 12% polyacrylamide gels. The concentration of 3CLpro was determined using absorbance at 280 nm and an extinction coefficient of 34380 M−1 cm−1 (using the ProtParam bioinformatic tool).

2.4. 3CLpro inhibition assays

The inhibitory activities of the polyphenols against 3CLpro were determined using the fluorescence resonance energy transfer (FRET) peptide substrate MCA-AVLQSFGF-Lys(Dnp)-Lys-NH2 (Isea Biochemicals, UK). Stock solutions of 8 mM polyphenols were prepared in 100% dimethylsulphoxide (DMSO), and were diluted in 50 mM Tris (pH 7.3), 1 mM EDTA for the analyses. The same buffer was used for dilution of the substrate and the 3CLpro.

For the initial screening, 10 µL of the different polyphenols at 40 µM were pre-mixed with 10 µL of the FRET substrate at 64 µM. The reactions were initiated by addition of 20 µL 3CLpro at 0.2 µM, to the final reaction volume of 40 µL. The final concentrations in the reaction mixtures were thus 0.1 µM 3CLpro, 16 µM substrate and 10 µM polyphenol. For the blanks, 20 µL buffer was added to the reaction mixtures instead of the 3CLpro. All of the reactions contained 1.25% (v/v) DMSO, including the control reactions without any polyphenols added. The reactions were carried out in black 384-well plates. Fluorescence signals were recorded continuously every 15 s for 10 min after addition of 3CLpro (Spark microplate reader, Tecan, Switzerland), at 320 nm excitation and 405 nm emission.
To determine the concentrations of the selected polyphenols required to inhibit 50% of the 3CL\textsuperscript{pro} activity (IC\textsubscript{50}), 10 μl of polyphenols at different concentrations were mixed with 20 μl 3CL\textsuperscript{pro} and incubated at 25 °C for 30 min, followed by addition of 10 μl substrate. The final concentrations in the reaction mixtures were again 0.1 μM 3CL\textsuperscript{pro} and 16 μM substrate. For the blanks, the polyphenols were mixed with 20 μl buffer instead of 3CL\textsuperscript{pro} before addition of the substrate. These reaction mixtures were incubated in black 384-well plates at 25 °C. Two hours after addition of the substrate, the fluorescence signals were recorded, as above.

2.5. Surface plasmon resonance

For the surface plasmon resonance (SPR) assays, 1.45 mg/mL 3CL\textsuperscript{pro} was dialysed against 10 mM HEPES (pH 7.3), 1 mM DTT, and then diluted in 10 mM sodium acetate buffer (pH 5.0) to the final concentration of 100 μg/mL 3CL\textsuperscript{pro}. The assays were performed on a SPR system (Biacore T200; GE Healthcare, USA) at 25 °C. The 3CL\textsuperscript{pro} protein was immobilised via the free amino groups using CMS sensor chip coated with carboxymethylated dextran (Cytiva, USA). Before the covalent immobilisation of 3CL\textsuperscript{pro}, the chip surface was activated with 0.1 M 1-ethyl-3-(3-dimethylaminopropyl)-carboimide hydrochloride and 0.1 M N-hydroxy succinimide. The injection of 3CL\textsuperscript{pro} was adjusted to obtain 9000 response units of immobilised 3CL\textsuperscript{pro} on the chip surface, and the uncoupled carboxymethyl groups were blocked with ethanolamine. The running buffer for the SPR assays was 10 mM HEPES (pH 7.4), 1 mM EDTA, 150 mM NaCl, 0.005% PEG and 5% DMSO. The polyphenol samples were prepared by diluting the polyphenol stock solutions in running buffer. The final concentration of DMSO in all of the polyphenol samples was 5%. Different concentrations of the polyphenols were injected over the chip surface for 60 s (250 s for ellagic acid). Dissociation of the polyphenols from the chip surface was then monitored for 60 s (200 s for ellagic acid). The polyphenols that remained bound to the chip surface were removed after the dissociation phase by a 6 s pulse of 2 mM NaOH. The buffer flow rate was set at 30 μL/min for all of the phases of the SPR assays. Sensograms were corrected for the corresponding responses of flow-cell 1 without the immobilised 3CL\textsuperscript{pro}. The Biacore T200 evaluation software was used to determine the equilibrium dissociation constants (K\textsubscript{D}) by fitting the data to a steady-state affinity model or to two-state kinetics. All of the experiments were performed in duplicate.

2.6. Circular dichroism

The 5 mM stock solutions of polyphenols for the circular dichroism (CD) measurements were prepared in 96% ethanol. Tris buffer (50 mM, pH 7.3) with 1 mM EDTA was used to dilute the 3CL\textsuperscript{pro} and polyphenol stocks to the final concentrations of 0.1 mg/mL and 100 μM, respectively. For the CD measurements, quartz cuvettes with an optical path of 1 mm were used. The CD spectra were scanned (J-1500CD spectrometer; Jasco, Japan), from 250 nm to 200 nm. The scanning speed was set at 20 nm/min, with a bandwidth of 1.0 nm. All of the CD spectra were collected at 20 °C. The spectra of the blank samples (buffer with polyphenols) were subtracted from the corresponding 3CL\textsuperscript{pro} spectra. The mean residue weight of 111 g/mol was used to calculate the molar ellipticity of 3CL\textsuperscript{pro}.

3. Results and discussion

3.1. In-silico screening of a polyphenol focused library

The top scoring 18 plant polyphenols obtained were subjected to cluster analysis to obtain representative top scoring binding conformations, and the results of the consensus docking experiment are shown in Table 1.

Polyphenols are often known as ‘promiscuous’ binders, and we fully acknowledge their inclusion in lists of pan-assay interference compounds (Bael & Holloway, 2010). Nevertheless, we felt compelled to investigate whether they have the potential to bind to the active site of SARS-CoV-2 3CL\textsuperscript{pro}. Furthermore, we advocate their potential use as scaffolds or hits ideal for scaffold hopping, similarity searches and further molecular optimisation. To this end, we performed additional blind pre-docking analyses in which the polyphenols were docked as potential ligands to the entire 3CL\textsuperscript{pro} protein without prior knowledge of the binding sites. Fig. 2A shows that ligands can generally interact with several different regions on the 3CL\textsuperscript{pro} dimer structure, although the catalytic site regions produced the most favourably scored binding poses. If focusing on the active site of 3CL\textsuperscript{pro}, Fig. 2B demonstrates the

| Polyphenol group | Polyphenol | MM/GBSA binding affinity (kcal/mol) | Consensus docking score | Interacting residues
|-----------------|-----------|-------------------------------------|------------------------|------------------------|
| Flavonoids      | Epigallocatechin gallate | -79.3 | -9.59 | 41, 49, 165, 166, 187, 189 |
|                 | Galangin  | -50.1 | -6.75 | 26, 41, 48, 143, 166 |
|                 | Isorhamnetin | -46.7 | -7.58 | 25, 27, 143, 145 |
|                 | Kaempferol | -54.0 | -7.43 | 49, 165, 166, 187, 189 |
|                 | Kaempferol- rutinoside | -44.0 | -9.30 | 25, 27, 41, 46, 49, 145, 187, 189 |
|                 | Phlorizin  | -40.9 | -8.36 | 41, 165, 166, 187, 189 |
| Curcuminoids    | Quercetin  | -50.9 | -8.05 | 48, 49, 187 |
|                 | Curcumin  | -60.2 | -6.82 | 49, 165, 166, 189, 192 |
| Phenylacetic acids | 4-Hydroxy-3-nitrophenoxyacetic acid | n.a. | n.a. | 41, 164, 187 |
| Stilbenes       | Resveratrol | n.a. | n.a. | 41, 165, 166, 187, 189 |
| Other polyphenols | Chlorogenic acid | -56.1 | -6.97 | 24, 25, 41, 44, 49, 187, 189 |
|                 | Ellagic acid | 51.6 | -7.42 | 142, 165, 166 |
| Phenolic acids  | Protocatechuic acid | -17.2 | -4.97 | 41, 48, 164, 187 |
|                 | Quinic acid | -27.7 | -5.33 | 26, 41, 142, 143 |
|                 | Syringic acid | -17.4 | -4.57 | 143, 166, 189 |
| Hydroxycinnamic acids | Caffeic acid | -24.5 | -5.05 | 41, 49, 54, 165, 187, 189 |
|                 | Ferulic acid | -25.3 | -5.55 | 25, 41, 44, 49, 189 |
|                 | Sinapic acid | -25.8 | -4.51 | 41, 164, 187, 189 |
|                 | Trans ferulic acid | -25.6 | -4.69 | 41, 48, 166 |
|                 | Umbellifer | -24.2 | -4.74 | 41, 48, 166, 187 |

1 MM/GBSA, molecular mechanics/generalised Born model and solvent accessibility.
2 As per PDB ID: 6M2N; n.a., not available.
identified ellagic acid calculated binding mode in the vicinity of the catalytic Cys145. Polyphenol thus makes key six hydrogen bonds with the active site residues, namely: Thr25, Gly143, Ser144, Cys145 and Glu166. Further \( \pi-\pi \) interaction can be observed between ligand aromatic system and His41 where ligand occupies \( P_1 \) and \( P_1' \) subsites extending towards \( P_2 \) pocket of the active site (Fig. 2).

After calculation of the consensus docking scores as the mean of the AutoDock VINA, Glide XP and CmDock scores, additional re-scoring was carried out using the Prime molecular mechanics/generalised Born model and solvent accessibility (MM/GBSA) methodology, as implemented in Schrödinger SMD (release 2021-1; Schrödinger, LLC, New York, NY, USA). Here, MM/GBSA \( \Delta G \) binding was estimated via the Prime energy of the optimised complex minus the Prime energy of the free ligand and the free receptor, all in the OPLS3e force field with implicit solvent. The consensus docking scores correlated well with the MM/GBSA \( \Delta G \) binding estimates, as well as with the Glide XP scores, which suggests that our top scorers might be promising binders of SARS-CoV-2 3CL\(^{pro}\). Upon inspection of the top-scoring binding poses, we observed a general similar binding motif to the crystal complex (PDB ID: 6M2N) of the polyphenolic ligand (5,6,7-trihydroxy-2-phenyl-4H-chromen-4-one), where all of the screened polyphenols were positioned at the active site or at the active site gorge entrance.

### 3.2. Molecular dynamics contact analysis

To gain further insight into the compound interaction with the 3CL\(^{pro}\) protein, we performed a plethora of molecular dynamics simulations on calculated binding poses of examined polyphenols. We were focusing on the stability of the complexes (see Supplementary Materials for details of all of the complexes examined) and we therefore examined the conformational stability of the examined polyphenols during the molecular dynamics production runs. We prepared an in-house script which recorded the contact maps of each polyphenol during the production runs, with a total of 12 polyphenols at 100 ns each (1.2 \( \mu \)s production simulation time). From the contact maps, it was observed that the major contacts that persevered through the simulation production time of all examined polyphenols with the 3CL\(^{pro}\) protein were towards the Met49, Glu166, Asp187 and Gln189, placing the compounds at the entrance or in the \( P_2 \) pocket of the 3CL\(^{pro}\) active site. Further observed major contacts were towards His41, Met165 and Glu166 at the \( P_1 \) and \( P_1' \) 3CL\(^{pro}\) active site pockets. The identified contacts (contact maps can be found in Supplementary Materials)
therefore place the polyphenols at the active site during the majority of the production runs, mainly at the P1–P1′ pockets, reaching also towards P2 subsite. Only two polyphenols: ferulic acid and resveratrol dissociated at the beginning of molecular dynamics simulations at 10 and 12 ns respectively (see Supplementary Materials for further details).

3.3. In-vitro screening for 3CLpro inhibition by the polyphenols

Initially, 19 different polyphenols (Fig. 1) were screened for inhibitory activity against 3CLpro at 10 μM, using the synthetic FRET substrate MCA-AVLQSGFR-Lys(Dnp)-Lys-NH2. In these assays, the substrate and tested polyphenols were pre-mixed, and the reactions were initiated by addition of 3CLpro. The hydrolysis of the substrate was followed continuously (see Supplementary Materials Fig. S2 for further details), and the inhibitory activities of the polyphenols were determined after 10 min. Half of the polyphenols showed < 25% inhibition of 3CLpro, with syringic acid showing no inhibition at all (Fig. 3). The majority of these weakly active polyphenols were those composed of a single aromatic ring; these are classified as phenolic acids (quinic, protocatechuic acids), hydroxycinnamic acids (sinapic, ferulic, caffeic, umbelllic acids), phenylacetic acids (4-hydroxy-3-nitrophenylacetic acid) and phenolic acid esters (chogenic acid) (Fig. 1). Greater inhibitory activities were seen for the polyphenols of the flavonoid class (isorhamnetin, phlorizin, galangin, kaempferol), which showed 25% to 50% inhibition of 3CLpro activity (Fig. 3). Of note, the glycosylated kaempferol analogue, kaempferol–rutinoside, showed significantly less 3CLpro inhibition (18%) compared to kaempferol (45%). This implies that the sugar moiety of kaempferol-rutinoside decreases its binding to the 3CLpro active site. Five of the polyphenols showed > 50% 3CLpro inhibition: quercetin, ellagic acid, curcumin, EGCG and resveratrol (Fig. 3). Interestingly, these compounds differ significantly in their structures and belong to distinct polyphenol subclasses; i.e., the flavonols (quercetin), flavonoids (EGCG), hydroxybenzoic acid derivatives (ellagic acid), curcuminoids (curcumin) and stilbenes (resveratrol) (Fig. 1).

Epigallocatechin gallate, ellagic acid, curcumin, resveratrol and quercetin were subjected to further dose–response studies, to determine the concentrations required to inhibit 3CLpro activity by 50% (IC50). Those polyphenols were selected on the basis of their relatively high inhibitory activities in the initial in-vitro screening (Fig. 3). For determination of the IC50 values, increasing concentrations of these polyphenols were incubated with 3CLpro before the addition of the FRET substrate, and the fluorescence signals were determined 2 h later. The IC50 values for EGCG, ellagic acid, curcumin, resveratrol and quercetin were 13.9 μM, 11.8 μM, 11.9 μM, 16.9 μM and 23.4 μM, respectively (Fig. 4). Thus, the potencies of these selected polyphenols for 3CLpro inhibition were similar in this study. Note, none of these polyphenols perturbed the secondary structure of 3CLpro to any great extent (see Supplementary Materials Fig. S3 for further details). Therefore, the inhibition of 3CLpro by these polyphenols appears to be due to their block of the 3CLpro active site, and not due to disruption of the conformation of 3CLpro. These IC50 values are comparable to, or lower than, those of the compounds that were identified as the top five SARS–CoV–2 3CLpro inhibitors among the 3987 US Food and Drug Administration approved drugs in a study by Mody et al. (2021).

Of these five polyphenols that showed the greatest inhibition of SARS-CoV-2 3CLpro in the present study, only quercetin and EGCG had been examined previously in in-vitro inhibition analyses of this protease. Quercetin was previously identified as the most potent inhibitor of SARS–CoV–2 3CLpro among 150 different compounds screened in-vitro, with an inhibition constant (Kic) estimated at 7.4 μM (Abian et al., 2020). This is in a similar range to the quercetin IC50 determined in the present study (23.4 μM) (Fig. 4). An IC50 for EGCG that was reported recently (0.056 μM) (Du et al., 2021) was considerably lower than that in the present study (13.9 μM) (Fig. 4), and in other recent studies, at 4.2 μM (Chiu et al., 2021) and 8.8 μM (Ni et al., 2021). Interestingly, EGCG appears to inhibit SARS-CoV-2 3CLpro more strongly than SARS-CoV 3CLpro, which was demonstrated previously (Chiu et al., 2021).

Resveratrol and curcumin were among these five strongest SARS–CoV-2 3CLpro inhibitors in the present study (Fig. 3), and these are also widely known natural compounds with antiviral activities (Campagna & Rivas, 2010, Mathew & Hsu, 2018). In the two recent studies, it was demonstrated that resveratrol inhibited proliferation of SARS–CoV-2 in mammalian cell cultures, with EC50 values from 5 μM to 10 μM (Yang, Wei et al., 2020, Pasquerane et al., 2021). Curcumin was previously shown to inhibit SARS-CoV 3CLpro in-vitro (Ryu et al., 2010), and was also predicted to bind stably into the SARS–CoV-2 3CLpro active site in-silico (Huyhn et al., 2020). To the best of our knowledge, no previous studies have investigated the inhibition of SARS–CoV-2 3CLpro by curcumin and resveratrol in-vitro. Ellagic acid is a tetracyclic polyphenol that is abundant in berries and nuts (Daniel et al., 1989), and it was also previously predicted to inhibit SARS-CoV-2 3CLpro, although only in-silico (Khalifa et al., 2020, Ni et al., 2021). The present study confirms the strong inhibitory activity of ellagic acid against this protease in-vitro, with an IC50 in the micromolar range.

![Fig. 3. 3CLpro inhibitory activities of the polyphenols, ordered according to increasing inhibitory activities. Data are means ± standard deviation of three experiments. The dotted line marks 50% 3CLpro activity. HNPA, 4-hydroxy-3-nitrophenylacetic acid; EGCG, epigallocatechin gallate.](image-url)
3.4. Interactions of polyphenols with 3CL\textsubscript{pro}

The binding of the selected polyphenols to 3CL\textsubscript{pro} was further studied using SPR spectroscopy, which allowed measurement of their interactions in real time. Here, the interactions were examined for three polyphenols that showed high 3CL\textsubscript{pro} inhibitory activities, as EGCG, ellagic acid, and resveratrol, and for two polyphenols that showed weak or no 3CL\textsubscript{pro} inhibitory activities in the initial \textit{in-vitro} screening (Fig. 3), as syringic acid and caffeic acid. Although curcumin and quercetin were also among the most potent 3CL\textsubscript{pro} inhibitors identified here, they were not examined by SPR due to their low solubilities.

Epigallocatechin gallate, ellagic acid and resveratrol bound to the 3CL\textsubscript{pro} immobilised on the chip (Fig. 5). Sensorgrams of EGCG and resveratrol indicated rapid association and dissociation of these polyphenols. The responses of EGCG and resveratrol did not reach plateaus even with these polyphenols at 200 \( \mu \text{M} \). Together with the fitting of the sensorgrams to steady-state kinetics (see Supplementary Materials Fig. S4 for more details), this SPR analysis suggested non-specific binding of EGCG and resveratrol to 3CL\textsubscript{pro}. In contrast, the two-state kinetics model provided a fit for the ellagic acid–3CL\textsubscript{pro} interaction, which indicated a \( K_D \) of 311 \( \pm \) 69 \( \mu \text{M} \). The sensorgrams suggested that the association of ellagic acid with 3CL\textsubscript{pro} is slower than...
that of EGCG and resveratrol, and it can be noted that the dissociation of ellagic acid from 3CL\textsubscript{pro} appeared to occur in two steps, as an initial rapid decrease in the SPR response that was followed by a slow dissociation. This is likely to be the consequence of multiple ellagic acid binding sites on 3CL\textsubscript{pro} with different affinities. Syringic and caffeic acids only bound weakly to 3CL\textsubscript{pro}, with EGCG, ellagic acid and resveratrol showing 10-fold higher RU values (Fig. 5). This is consistent with the low inhibitory activity of syringic and caffeic acids towards 3CL\textsubscript{pro}.

4. Conclusion

We initially developed a consensus docking approach that resulted in a library enrichment of 2.5-fold, which was determined by the biological evaluation in the following step. IC\textsubscript{50} values were obtained for EGCG, curcumin, resveratrol, quercetin and ellagic acid of 13.9 µM, 11.9 µM, 16.9 µM, 23.4 µM and 11.8 µM, respectively. Further in-silico molecular dynamics studies indicated labile interactions for resveratrol, while they produced stable binding conformations at the active site of 3CL\textsubscript{pro} for the other four polyphenols. EGCG showed hydrophobic interactions with 3CL\textsubscript{pro} residues Met49, Met165 and Gln189, and hydrogen bonds with residues Asp187 and Glu166. Curcumin formed hydrogen bonds with
3CL\textsuperscript{pro} residues Gln192 and Arg188. Ellagic acid occupied a binding mode at the active site of 3CL\textsuperscript{pro}, with hydrogen bonds with residues Cys44, Asn142, His164 and Gln189, and it formed hydrophobic interactions with residues Asn142, Cys145, Met165 and Met166, and π-π stacking with His41. The binding of ellagic acid to 3CL\textsuperscript{pro} was confirmed by SPR using a two-state kinetics model, with a measured K\textsubscript{D} of 311 ± 69 μM. We have therefore demonstrated that ellagic acid as a low micromolar SARS-CoV-2 3CL\textsuperscript{pro} inhibitor that is suitable for further crystallographic studies. This study also enables the use of ellagic acid and similar polyphenols for scaffold hopping and similarity searches and optimisation, with a view to new potential SARS-CoV-2 inhibitors or novel molecular probes.

Declaration of Competing Interest

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

Acknowledgements

The authors thank Rita Podzuna from Schrödinger LLC for her support, and for further support from Schrödinger LLC. They also thank Ole Tange for making wonderful GNU tools.

Funding

This work was supported by the Slovenian Ministry of Science and Education infrastructure grant HPC-RIVR, grant OP20.04342, and by the Slovenian Research Agency (ARRS) programme and project grants P1-0201, P2-0046, P4-0121, J1-2471 and P1-0207.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2021.131594.

References

Abbas, M., Saeed, F., Anjum, F. M., Afzal, M., Tufail, T., Bashir, M. S., et al. (2017). Natural polyphenols: An overview. International Journal of Food Properties, 20(8), 1689-1699. https://doi.org/10.1080/10942912.2016.1220939
Ahn, O., Ortega-Alarcon, D., Jimenez-Alesanco, A., Ceballos-Laita, L., Vega, S., Reyburn, H., et al. (2020). Structural stability of SARS-CoV-2 3CLpro and identification of quercetin as an inhibitor by experimental screening. International Journal of Biological Macromolecules, 164, 1693-1703. https://doi.org/10.1016/j.ijbiomac.2020.07.235
Baell, J. B., & Holloway, G. A. (2010). New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. Journal of Medicinal Chemistry, 53(7), 2719-2740. https://doi.org/10.1021/jm901137v
Calland, N., Sahar, M.-E., Belouzard, S., Pène, V., Bonnafous, P., Mesalam, A. A., et al. (2015). Polyphenols inhibit hepatitis C virus entry by a new mechanism of action. Journal of Virology, 89(9), 10553-10563. https://doi.org/10.1128/JVI.01732-15
Campana, M., & Rivas, C. (2010). Antiviral activity of resveratrol. Revista de Beverages. Scientific Transactions, 38(1), 50-53. https://doi.org/10.1042/BSST0380050
Choi, W.-C., Chen, J.-C., Chen, Y.-T., Yang, J.-M., Hwang, L.-H., Lyu, Y.-S., et al. (2021). The inhibitory effects of PGG and EGCG against the SARS-CoV-2 3CL\textsuperscript{pro}. Biochemical and Biophysical Research Communications. https://doi.org/10.1016/j.bbrc.2020.12.106
Dai, W., Zhang, B., Jiang, X. M., Su, H., Li, J., Zhao, Y., et al. (2020). Structure-based in-silico mutagenesis-based remodelling of SARS-CoV-1 peptide (ATLQAIAS) to inhibit SARS-CoV-2: structural-dynamics and free energy calculations. Interdisciplinary Sciences: Computational Life Sciences, 13(3), 521-534. https://doi.org/10.1007/s12539-021-00447-2
Kahn, S., & Melchnosh, K. (2005). History and recent advances in coronavirus discovery. The Pediatric Infectious Disease Journal, 24(11), S223-S227. https://doi.org/10.1097/01.pid.0000181667.17240.64
Khalifa, I., Zhu, W., Mohammed, H. H., Dutta, K., & Li, C. (2020). Tannins inhibit SARS-CoV-2 through binding with catalytic dyad residues of 3CL\textsuperscript{pro} in a mimic of viral receptor with 19 structural different hydrolysable tannins. Journal of Food Biochemistry, 44(10), Article e13432. https://doi.org/10.1111/jfbc.13432
Kim, Y., Narayanan, S., & Chang, K. O. (2010). Inhibition of influenza virus replication by plant-derived isocoumarin. Antiviral Research, 88(2), 227-235. https://doi.org/10.1016/j.antiviral.2010.08.016
Krieger, E., & Friend, G. (2015). New ways to boost molecular dynamics simulations. Journal of Computational Chemistry, 36(13), 996–1007. https://doi.org/10.1002/jcc.23899
Mathew, D., & Hsu, W. (2018). Antiviral potential of curcumin. Journal of Functional Foods, 40, 692-699. https://doi.org/10.1016/j.jff.2017.12.017
Mehany, T., Khalifa, I., Barakat, H., Alhawwab, S. A., Alharbi, Y. M., & El-Shemy, S. C. (2021). Polyphenols as promising biologically active substances for preventing SARS-CoV-2: A review with research evidence and underlying mechanisms. Food Research International, 140, 108891. https://doi.org/10.1016/j.foodres.2021.108891
Mody, V., Ho, J., Wills, S., Mawri, A., Lawson, L., Ebert, M. C., et al. (2021). Identification of 3-chromopyrin like protein (3CL\textsuperscript{pro}) inhibitors as potential anti-SARS-CoV-2 agents. Communications Biology, 4(1), https://doi.org/10.1038/s42003-020-01577-z
Ni, W., Wei, X.-N., Wei, S.-Y., Lan, L.-L., Qiu, R.-J., Ling, Y.-P., et al. (2021). Study on the mechanism of active components of Lithop tea on 3CL\textsuperscript{pro} based on HPLC-DAD fingerprint and molecular docking technique. Journal of Food Biochemistry, 45(5), https://doi.org/10.1111/jfs.13170
Parizzo, I. L., Revel, J. S., & Stein, J. F. (2020). Potential use of polyphenols in the battle against COVID-19. 2020 Current Opinion in Food Science, 32, 149–155. https://doi.org/10.1016/j.cofs.2020.08.004
Pascuereau, S., Nebme, Z., Haidar Ahmad, S., Daudou, F., Van Assche, J., Wallet, C., et al. (2021). Resveratrol inhibits HCoV-229E and SARS-CoV-2 coronavirus replication in vitro. Viruses, 13(2), 354. https://doi.org/10.3390/v13020354
Passali, H., Farzaei, M. H., & Khodarahmi, R. (2017). Polyphenols and their benefits: A review. International Journal of Food Properties, 20(2), 1700-1741. https://doi.org/10.1080/10286018.2017.1354017
Rizzuti, B., Grande, F., Conforti, F., Jimenez-Alesanco, A., Ceballos-Laita, L., Ortega-Alarcon, D., et al. (2021). Rutin is a low micromolar inhibitor of SARS-CoV-2 main protease 3CL\textsuperscript{pro}: Implications for drug design of quercetin analogs. Biomolecules, 9(10). https://doi.org/10.3390/biom9100676
Ruiz-Carmona, S., Alvarez-Garcia, D., Foloppe, N., Garmendia-Doval, A. B., Juhos, S., Paraiso, I. L., Revel, J. S., & Stevens, J. F. (2020). Potential use of polyphenols in the future perspectives. Frontiers in Pharmacology, 11. https://doi.org/10.3389/fphar.2020.572870
Ryu, Y. B., Park, S.-J., Kim, Y. M., Lee, J.-Y., Seo, W. D., Chang, J. S., et al. (2010). SARS-CoV 3CL\textsuperscript{pro} inhibitory effects of quinone-methide triterpenes from Triglycerum regelii. Bioorganic and Medicinal Chemistry Letters, 20(6), 1873-1876. https://doi.org/10.1016/j.bmcl.2010.01.152
Sarkar, C., Mondal, M., Torequl Islam, M., Martorell, D., Docaa, A. O., Marowy, A., et al. (2020). Potential therapeutic options for COVID-19: Current status, challenges, and future perspectives. Frontiers in Pharmacology, 11. https://doi.org/10.3389/fphar.2020.729870
Tretutter, D. (2006). Significance of flavonoids in plant resistance: A review. Environmental Chemistry Letters, 4(3), 147-157. https://doi.org/10.1007/s10311-006-0068-8
Trott, O., & Olson, A. J. (2010). AutoDock VINA: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multitreading.
M. Bahun et al.

Journal of Computational Chemistry, 31(2), 455–461. https://doi.org/10.1002/jcc.21133

Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. The Lancet, 395(10223), 470–473. https://doi.org/10.1016/S0140-6736(20)30185-9

Williamson, G. (2017). The role of polyphenols in modern nutrition. Nutrition Bulletin, 42(3), 226–235. https://doi.org/10.1111/nbu.12278

Yang, M., Wei, J., Huang, T., Lei, L., Shen, C., Lai, J., et al. (2020). Resveratrol inhibits the replication of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in cultured Vero cells. Phytotherapy Research, 35(3), 1127–1129. https://doi.org/10.1002/ptr.6916

Yang, Y.-L., Yu, J.-Q., & Huang, Y.-W. (2020). Swine enteric alphacoronavirus (swine acute diarrhea syndrome coronavirus): An update three years after its discovery. Virus Research, 285, 198024. https://doi.org/10.1016/j.virusres.2020.198024

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