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.

Journal of Molecular Graphics and Modelling 110 (2022) 108050

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/jmgm

Small molecule interactions with the SARS-CoV-2 main protease: *In silico* all-atom microsecond MD simulations, PELE Monte Carlo simulations, and determination of *in vitro* activity inhibition

Julia Liang [a,b], Eleni Pitsillou [a,b], Katherine Ververis [a], Victor Guallar [c,d], Andrew Hung [b], Tom C. Karagiannis [a,e,*]

[a] Epigenomic Medicine, Department of Diabetes, Central Clinical School, Monash University, Melbourne, VIC, 3004, Australia
[b] School of Science, College of Science, Engineering & Health, RMIT University, VIC, 3001, Australia
[c] Barcelona Supercomputing Center, Jordi Girona 29, E-08034, Barcelona, Spain
[d] ICREA, Passeig Lluís Companys 23, E-08010, Barcelona, Spain
[e] Department of Clinical Pathology, The University of Melbourne, Parkville, VIC, 3052, Australia

ARTICLE INFO

Keywords:
Coronavirus
COVID-19
SARS-CoV-2
Main protease
GC376
Hypericin
Cyanidin-3-O-glucoside
Molecular modelling

ABSTRACT

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused the ongoing COVID-19 pandemic. With some notable exceptions, safe and effective vaccines, which are now being widely distributed globally, have largely begun to stabilise the situation. However, emerging variants of concern and vaccine hesitancy are apparent obstacles to eradication. Therefore, the need for the development of potent antivirals is still of importance. In this context, the SARS-CoV-2 main protease (M^{pro}) is a critical target and numerous clinical trials, predominantly in the private domain, are currently in progress. Here, our aim was to extend our previous studies, with hypericin and cyanidin-3-O-glucoside, as potential inhibitors of the SARS-CoV-2 M^{pro}. Firstly, we performed all-atom microsecond molecular dynamics simulations, which highlight the stability of the ligands in the M^{pro} active site over the duration of the trajectories. We also invoked PELE Monte Carlo simulations which indicate that both hypericin and cyanidin-3-O-glucoside preferentially interact with the M^{pro} active site and known allosteric sites. For further validation, we performed an *in vitro* enzymatic activity assay that demonstrated that hypericin and cyanidin-3-O-glucoside inhibit M^{pro} activity in a dose-dependent manner at biologically relevant (μM) concentrations. However, both ligands are much less potent than the well-known covalent antiviral GC376, which was used as a positive control in our experiments. Nevertheless, the biologically relevant activity of hypericin and cyanidin-3-O-glucoside is encouraging. In particular, a synthetic version of hypericin has FDA orphan drug designation, which could simplify potential clinical evaluation in the context of COVID-19.

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused the ongoing COVID-19 pandemic [1,2]. SARS-CoV-2 is an enveloped virus that contains a relatively large non-segmented, positive sense, single-stranded RNA genome [2–4]. The RNA genome is comprised of open reading frames (ORFs) that encode for non-structural proteins, the structural proteins, and accessory proteins [5]. ORF1a and ORF1b encode the polyproteins pp1a and pp1ab, which are processed to form 16 non-structural proteins that are essential for viral replication [6,7].

The main protease (M^{pro}), which is also known as the 3-chymotrypsin-like cysteine protease ($3CL^{pro}$), and the papain-like protease (PL^{pro}) cleave pp1a and pp1ab to generate the non-structural proteins [7]. The SARS-CoV-2 M^{pro} functions as a dimer and each active site consists of a cysteine-histidine (Cys145-His41) catalytic dyad that is involved in the hydrolysis of peptide bonds and the cleavage of polypeptide sequences [8]. The three-dimensional (3D) structure highlights the three domains that make up each protomer [9,10]. The substrate-binding site is located between domains I and II [9].

Studies have shown that M^{pro} cleaves polypeptides at the P2–P1 ↓ P1′ consensus sequence [11,12]. P1 is glutamine, P1′ is a residue with a

* Corresponding author. Head Epigenomic Medicine Program, Department of Diabetes, Central Clinical School, Monash University. Melbourne, VIC, 3004, Australia.
E-mail address: tom.karagiannis@monash.edu (T.C. Karagiannis).

https://doi.org/10.1016/j.jmgm.2021.108050
Received 18 June 2021; Received in revised form 16 September 2021; Accepted 8 October 2021
Available online 11 October 2021
1093-3263/© 2021 Elsevier Inc. All rights reserved.

small side chain (Ser/Ala/Gly/Asn), and P2 is a large, hydrophobic residue (Leu/Met/Phe/Val) [11,12]. The reaction mechanism of M^{pro} involves acylation and deacylation stages. In the acylation step, the peptide bond is broken and the P1' fragment of the substrate is released [8,13]. A covalent bond is formed between Cys145 of M^{pro} and the carbon atom of the P1 residue, and this produces an acyl-enzyme complex [8,13]. In the deacylation step, the acyl-enzyme complex is hydrolysed and the P1 fragment is released [8,13]. In addition to the active site, potential allosteric sites have been identified [14–19]. This includes the highly reactive binding pocket in the dimerization region at the domain III apex [20].

There has been an intense research focus on the development of potential antivirals that target the SARS-CoV-2 M^{pro}. Due to the cleavage site and substrate specificities of the M^{pro} being different from human proteases, inhibitors are predicted to be non-toxic to humans [7,10]. Potential inhibition of the SARS-CoV-2 M^{pro} has largely focussed on covalent inhibitors of the protease with particular interest in the α-ketoamide analogues which have potent activity in the low micromolar range [10,20,21]. A nucleophilic attack occurs between the sulfur atom of Cys145 and the ketoamide group of α-ketoamide inhibitors, forming a thiohemiketal [10,21]. In a study by Zhang et al., the hydroxyl (oxyanion) group of the thiohemiketal that was formed with the α-ketoamide inhibitor 13b and the catalytic Cys145 was stabilised by a hydrogen bond from His41 [10]. The amide oxygen of the inhibitor was found to accept a hydrogen bond from the residues Gly143, Ser144, and Cys145, which form the oxyanion hole [11,21,22]. The acyl-enzyme complex that covalent inhibitors form with M^{pro} cannot be hydrolysed and they consequently remain bound to the active site [8]. Other key compounds which have shown potential include N3 (covalent inhibitor), GC376 (cysteine protease covalent inhibitor), the indole 5h (reversible covalent inhibitor), the indoline GRL-1720 (irreversible covalent inhibitor), ebselen (inhibition of M^{pro} activity), and boceprevir (serine protease inhibitor that binds covalently) [9,23,24].

Further, there has been an enormous effort at screening large libraries, which have included a vast array of different structural classes of compounds to potentially identify novel inhibitors [25–29]. Our laboratory as well as many others have turned their attention to the utility of dietary compounds as M^{pro} inhibitors, and hypericin and cyanidin-3-O-glucoside have been identified as potential leads [20,30,31]. Hypericin and cyanidin-3-O-glucoside have been investigated as potential inhibitors against several SARS-CoV-2 structural and non-structural proteins including M^{pro}, PL^{pro}, and the spike protein [32–35]. In a study by Islam et al., hypericin and cyanidin-3-O-glucoside were docked to the M^{pro} and were selected as top-ranked candidates due to their strong binding affinity [35]. The stability of the protein-ligand complexes were evaluated further using MD simulations [35] Cyanidin-3-O-glucoside is a major anthocyanin that can be found in berries, while hypericin is an anthraquinone derivative that is found in the plant *Hypericum perforatum* (St. John's Wort) [36,37]. The antiviral and immunomodulating effects of these compounds have been reported [38–41]. Most notably, a synthetic form of hypericin (SGX301) has received Orphan Drug and Fast Track designations from the U.S Food and Drug Administration for photodynamic therapy [42].

Here, our overall aim was to extend our previous studies to gain further insight into the potential antiviral activity of our lead compounds [20,29,43]. We performed all-atom molecular dynamics (MD) simulations (1 μs), PELE Monte Carlo simulations, and *in vitro* assays to further evaluate hypericin and cyanidin-3-O-glucoside as potential inhibitors of the SARS-CoV-2 M^{pro}.

2. Materials and methods

2.1. Molecular dynamics simulations

A homomeric complex of the SARS-CoV-2 M^{pro} (PDB ID: 6LU7) was assembled using the PDBePISA (Proteins, Interfaces, Structures and Assemblies) server [44]. Cyanidin-3-O-glucoside, hypericin and its isomer were docked using the quantum-mechanics polarised ligand docking (QPLD) protocol of the Schrödinger suite to the active site of each M^{pro} protomer, as previously described [29]. Docked ligands served as starting structures for simulations, with protein-ligand complexes for simulations containing two ligands – one bound to each protomer of M^{pro}. Ligand topologies were generated using SwissParam [45]. Molecular dynamics (MD) simulations were performed using GROMACS 2018.2 software [46,47] with the CHARMM27 force field [48,49]. Protein-ligand complexes were solvated using TIP3P water [50] in a dodecahedral box, with a minimum distance of 2.0 nm between any protein atom to the box edge. The solvated system was neutralized with sodium ions. Energy minimisation was performed using the steepest-descent gradient method. Systems were restrained using a canonical (NVT) ensemble followed by an isothermal-isobaric ensemble (NPT) for 100 ps. Temperature was maintained at 310 K with a modified Berendsen thermostat [51], and pressure at 1.0 bar with the Parrinello-Rahman barostat [52]. Bond lengths were constrained using the LINCS algorithm [53], and long-range electrostatic forces were calculated using the particle-mesh Ewald scheme (PME) [54] (grid spacing 0.16 nm). Cut-off ratios of 1.2 nm for Coulomb and van der Waals potentials were used for the calculation of short-range nonbonded interactions. Simulations were carried out for 1000 ns or 100 ns in triplicate with a time-step of 2 fs.

Simulated trajectories were visualised using Visual Molecular Dynamics 1.9.3 [55]. Root mean square deviation (RMSD), radius of gyration (Rg), and root mean square fluctuation (RMSF) analysis tools included in Gromacs 2018.2 were utilised. The number of contacts between residues of M^{pro} and ligands throughout the entire trajectory was calculated using gmx mindist, with a threshold of 0.45 nm to define a contact between the ligand and protein residue [46,47]. Using principal component analysis (PCA) to examine large-scale concerted motions of the protein in each system, essential dynamics analysis was employed [56]. A covariance matrix was constructed from the C-alpha atoms of the protein using the gmx covar tool within Gromacs, which was diagonalized to obtain a set of eigenvectors and associated eigenvalues. This was performed for the 1000 ns trajectories and for concatenated trajectories of the triplicate runs. The gmx anaeig tool was used to calculate the 2D projections with respect to the selected eigenvector components.

2.2. Protein energy landscape exploration (PELE) Monte Carlo simulations

For a more realistic and reliable determination of the binding of hypericin and cyanidin-3-O-glucoside to the SARS-CoV-2 M^{pro}, we performed PELE Monte Carlo simulations according to Lecina et al. [57]. For all systems we performed a global search and a local refinement. The global search is intended for a blind binding site search. The ligand is placed in randomly (around 40) initial poses at the protein surface. Here the ligand is fully solvated, that is with no direct contact with the protein. Then 256 computing cores start an adaptive-PELE Monte Carlo (MC) search. In the global sampling we combine large and short, ~3 and ~1 Å, ligand translations. The adaptive procedure included 100 epochs of 10 MC steps each. Thus, in total we have around ~256 000 PELE steps. Moreover, in the adaptive procedure, in each epoch we aim for exploring those regions less explored, resulting in an effective exploration of the entire surface [57].

The local search used as initial structures the best ones from the global search, those local minima with lower interaction energy. Then a shorter simulation using smaller translations and rotations is used. It involved about 20–30 computing cores per minima and around 10 epochs of 24 MC steps each. For the analysis, we used PELE's protein-ligand interaction energies. These are obtained by subtracting the receptor and ligand energies from the complexes one at a given geometry, using the OPLS-AA force field with a generalised surface Born solvent model. Thus, they are not intended to effectively discriminate among

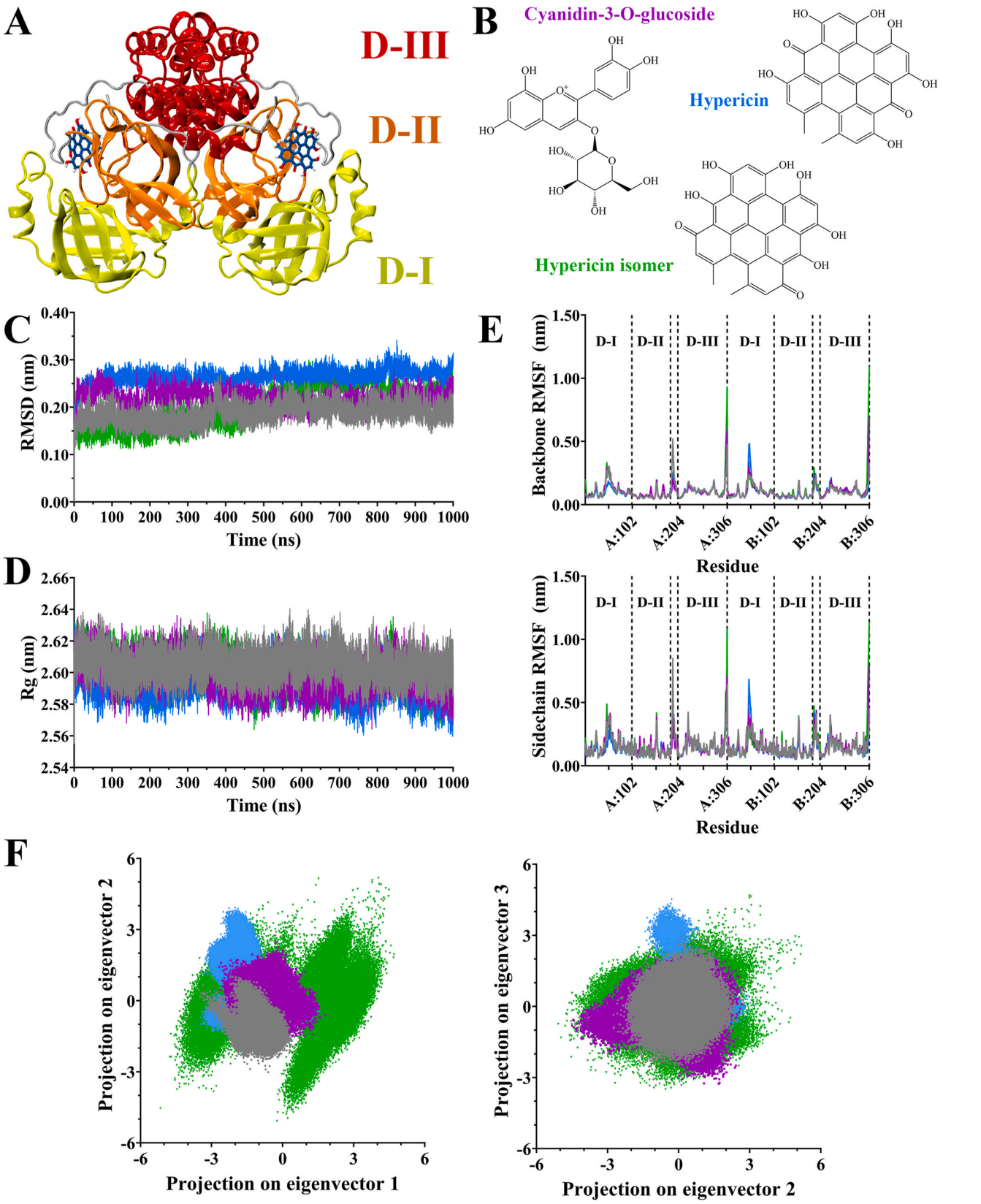

Fig. 1. Classical MD simulations with ligands bound to the M^{pro} active site. A) The SARS-CoV-2 M^{pro} is consisted of three domains. Systems consisted of the M^{pro} dimer with a single ligand bound to the active site of each protomer, with simulations carried out for 1000 ns. B) Chemical structures of the ligands studied. C) Root mean square deviation (RMSD) of M^{pro} protein backbone. D) Radius of gyration of M^{pro} protein backbone. E) Root mean square fluctuation (RMSF) of the M^{pro} backbone and sidechains over the trajectory. F) 2D projection of M^{pro} C-alpha atoms along the first two eigenvectors, and second and third eigenvectors. Data for the apo protein is shown in grey, cyanidin-3-O-glucoside-bound protein in purple, hypericin-bound protein in blue, and hypericin isomer-bound protein in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

ligands but between poses from a given ligand.

2.3. Fluorogenic M^{pro} inhibition assay

A commercially available fluorogenic M^{pro} protease assay kit (BPS Bioscience, San Diego, CA, USA), was used to investigate the *in vitro* inhibition properties of small molecules. According the manufacturer's protocol, a fluorogenic substrate (excitation wavelength = 360 nm) was detected at 460 nm, and the positive internal control GC376 was used at a final concentration of 50 μM. The small molecule inhibitors hypericin (89%, HWI pharma services GmbH, Germany), and cyandin-3-O-glucoside (reference standard, PhytoLab, Germany), were prepared as 20 mM stocks and stored at −80 °C until use. Doubling dilutions to achieve final concentrations in the range of 0.2–200 μM of hypericin and cyandin-3-O-glucoside were prepared in assay buffer. The test inhibitors were assayed in triplicate; six determinations were made for the background, total M^{pro} protease activity, and the GC376 positive control. Absolute fluorescence intensity values at 460 nm were measured, and % protease inhibition and IC_{50} values were calculated.

3. Results and discussion

3.1. Protein effects in response to ligand binding to the M^{pro} active site

Classical all atom MD simulations were carried out for 1 μs with ligands bound to the protein dimer (Fig. 1A). Ligands examined were cyanidin-3-O-glucoside and hypericin, as well as an isomer of hypericin (Fig. 1B), which we previously identified as potential leads for inhibition of M^{pro} [20]. MD data indicated that ligand binding did not greatly influence the overall structure of the M^{pro} dimer. Average RMSD values for the trajectory were similar between all systems in the microsecond trajectories, with values of 0.19 nm for the apo protein, 0.22 nm for cyanidin-3-O-glucoside, and 0.20 for hypericin isomer bound to M^{pro} (Fig. 1C). A slightly higher RMSD of 0.26 nm was observed for hypericin-bound M^{pro}. Triplicate 100 ns trajectories had similar RMSD values following equilibration of 0.19 nm for apo and 0.20 nm for cyanidin-3-O-glucoside- and hypercin-bound M^{pro} (Fig. S1). Rg values were almost identical for all systems, with an average of 2.60 nm (Fig. 1D and Fig. S1). Principal components analysis (PCA) of M^{pro} in the microsecond trajectories along the first two eigenvectors showed the least variation in collective motion for the apo protein compared to ligand-bound M^{pro} (Fig. 1F). Motions towards positive and negative values were observed for the hypericin analogues, whereas cyanidin-3-O-glucoside was more similar to the apo form, with a slight positive shift in eigenvector 2 (Fig. 1F). Projection along eigenvectors 2 and 3 were more similar between the systems, with more overlap indicating similar protein motions between apo- and ligand-bound M^{pro} (Fig. 1F). Similar trends are observed in the triplicate 100 ns trajectories (Fig. S1).

RMSF analysis indicated that the C-terminal tail of each protomer had the greatest flexibility in the protein (Fig. 1E). Regions within the protein showing flexibility included residues 46–52 in domain I, and residues 187–191 of the connecting loop between domains II and III. The flexible domain I residues are in proximity to the hydrophobic S2 subsite involved in substrate binding [10]. Residues 46–52 are particularly flexible for all systems, with a maximum backbone RMSF of approximately 0.30 nm in protomer A for all systems studied except hypericin-bound M^{pro}, with a value of 0.15 nm. For these residues in protomer B, backbone RMSF was lowest in the apo protein (~0.20 nm), with slightly higher values for the cyanidin-3-O-glucoside and hypericin isomer-bound (~0.30 nm), and the highest backbone RMSF for hypericin-bound M^{pro} (0.48 nm). The flexible connecting loop residues also displayed some differences in fluctuation in response to ligand binding for protomer A, with a higher backbone RMSF of 0.52 nm for Gln189 in the apo protein, compared to 0.20, 0.24, and 0.25 nm for cyanidin-3-O-glucoside, hypericin, and hypericin isomer-bound M^{pro},

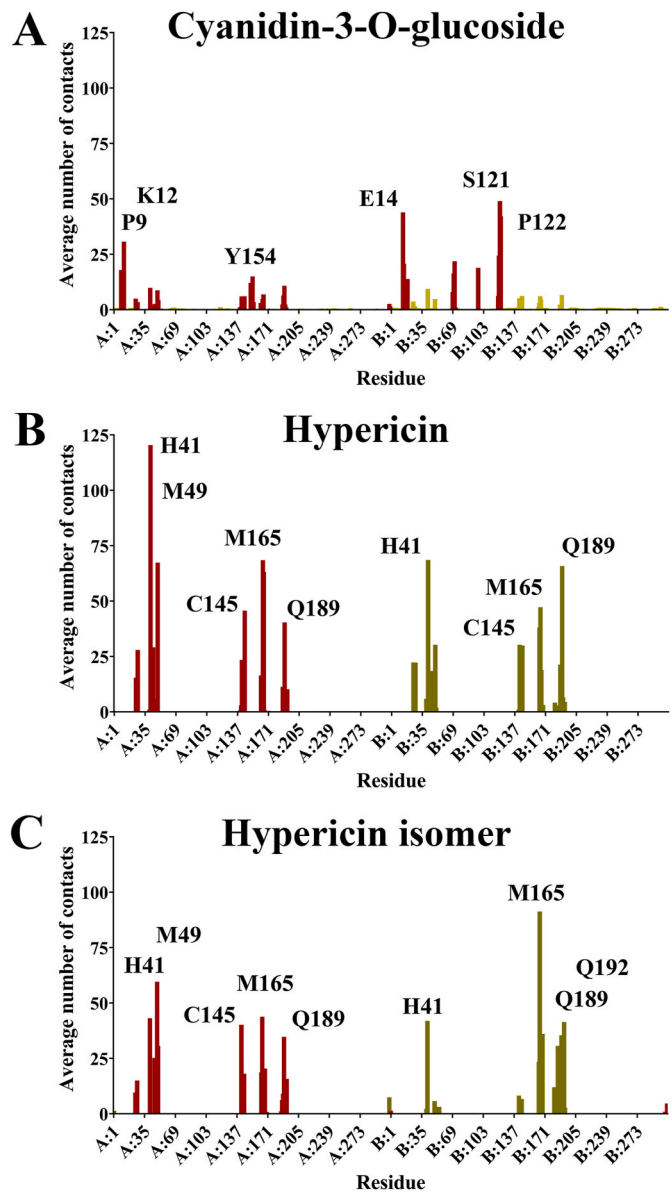

Fig. 2. Number of contacts between ligands bound to SARS-CoV-2 M^{pro}. The average number of contacts between A) cyanidin-3-O-glucoside, B) hypericin, and C) hypericin isomer initially bound to protomer A (red) and protomer B (yellow) of the SARS-CoV-2 M^{pro} dimer was calculated over a 1000 ns trajectory. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

respectively. In protomer B, differences between the systems were less obvious, with values of approximately 0.20 nm for all. Similar trends are observed with sidechain RMSF calculations, with a fluctuation of 0.85 nm for Gln189 in protomer A for the apo protein, compared to values of 0.44, 0.37, and 0.52 nm for cyanidin-3-O-glucoside, hypericin, and its isomer, respectively. For Gln189 in protomer B, the sidechain fluctuation was approximately 0.35 nm for all systems. As these connecting loop residues form part of the substrate binding site, it is suggested that ligand binding may stabilise these residues in protomer A of the M^{pro} dimer. These differences in the flexibility of binding site residues may account for the observed preference for binding of ligands to protomer A over protomer B, as discussed below.

3.2. Stability of compounds bound to the M^{pro} active site

Visual inspection of the trajectories show that M^{pro} remains mostly

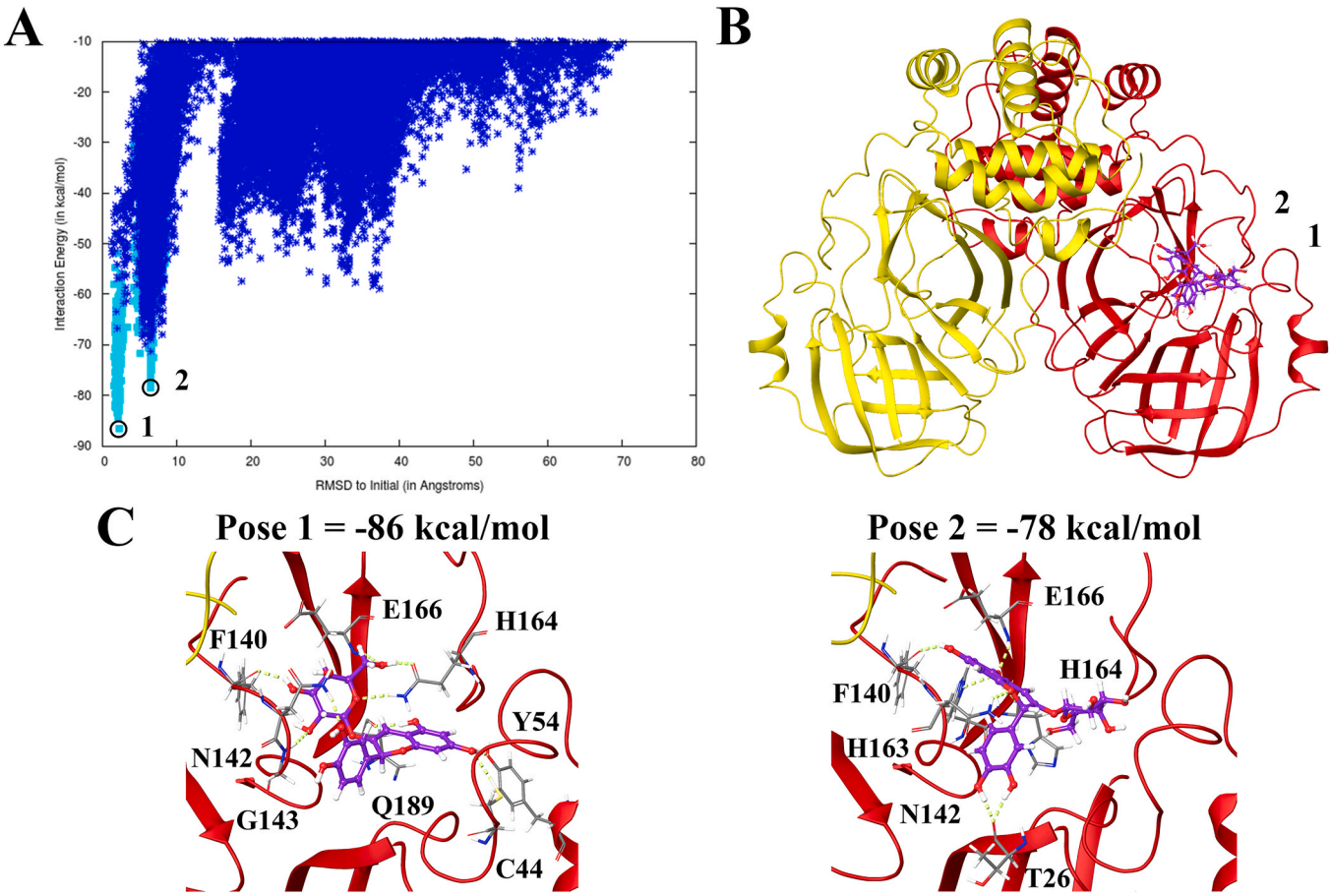

Fig. 3. PELE binding site search of SARS-CoV-2 M^{pro} dimer for cyanidin-3-O-glucoside. A) Interaction energy plot vs RMSD distance to initial position of cyanidin-3-O-glucoside in the substrate binding site of protomer A of SARS-CoV-2 M^{pro}. The global search is shown in dark blue, and local refined poses are shown in cyan. B) The best poses following local refinement are numbered 1 and 2. C) Interaction between ligand poses and protein residues, with hydrogen-bonds indicated by dashed lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

stable in both its apo and ligand-bound forms (Movie S1-4). Hypericin and its isomer remain bound to the active site of both protomers of the M^{pro} for the entire trajectory (Movies S3 and S4). On the other hand, ligand unbinding of cyanidin-3-O-glucoside occurs from both substrate binding sites. At 120 ns, the ligand initially bound to protomer B begins to loosen from the substrate binding site, before completely detaching from the protein at 200 ns (Movie S2). While the cyanidin-3-O-glucoside initially bound to protomer A also unbinds at 120 ns, at approximately 275 ns it is observed to re-attach to M^{pro} at the interface of domains I and II between both protomers (Movie S2). The ligand remains bound to this site for the remainder of the trajectory, indicating the presence of a potential allosteric binding site on the M^{pro} dimer.

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jmgm.2021.108050.

The average number of contacts between the bound ligand and each residue of the protein throughout the duration of the trajectory was calculated (Fig. 2). It was apparent that there was a greater number of contacts with protomer A across all systems. Due to unbinding events observed with cyanidin-3-O-glucoside, a lesser number of contacts is observed compared to hypericin. The protomer A-bound cyanidin-3-O-glucoside is shown to have contacts with Pro9 (18 contacts), Lys12 (31 contacts) and Tyr154 (15 contacts) in protomer A, and residues Glu14 (44 contacts), Ser121 (49 contacts), and Pro122 (42 contacts) in protomer B. As both hypericin and its isomer remained tightly bound to the substrate binding site in both chains of the M^{pro}, a greater average number of contacts with key active site residues is observed. The catalytic dyad residues were among the residues with the greatest number of contacts with these ligands. His41 in protomer A had 120 contacts and 69 contacts in protomer B with hypericin, suggesting a preference for more favourable binding to protomer A. His41 made 43 and 42 contacts for its isomer bound to protomer A and B, respectively. For Cys145, hypericin had 46 and 30 contacts, while its isomer had 18 and 7 contacts in protomers A and B, respectively. Other notable key residues within the substrate binding site making contacts with these two ligands included Met49, Met165, and Gln189 across both protomers. This reaffirms the active site of M^{pro} as a favourable binding region for hypericin.

3.3. PELE highlights the active site and known allosteric sites

Adaptive-PELE Monte Carlo simulations were performed with cyanidin-3-O-glucoside and hypericin to identify binding modes on the M^{pro} dimer. Local refinement of the best structures from global exploration revealed distinct minima for both ligands with strong interaction energies for the active site. Particularly for cyanidin-3-O-glucoside, the two best poses were in the active site of protomer A (Fig. 3). Following refinement, the pose closest to the starting position (pose 1) had the lowest interaction energy of −86 kcal/mol. In this position, binding of cyanidin-3-O-glucoside in the active site is stabilised by an extensive hydrogen bond network, with half of these being formed with the glucoside group of the ligand. Key active site residues forming hydrogen bonds were Phe140, Asn142, Glu166 of the S1 subsite, and Tyr54 of the S2 subsite. The alternative orientation of the second pose also had a strong interaction energy (−78 kcal/mol) and was stabilised by

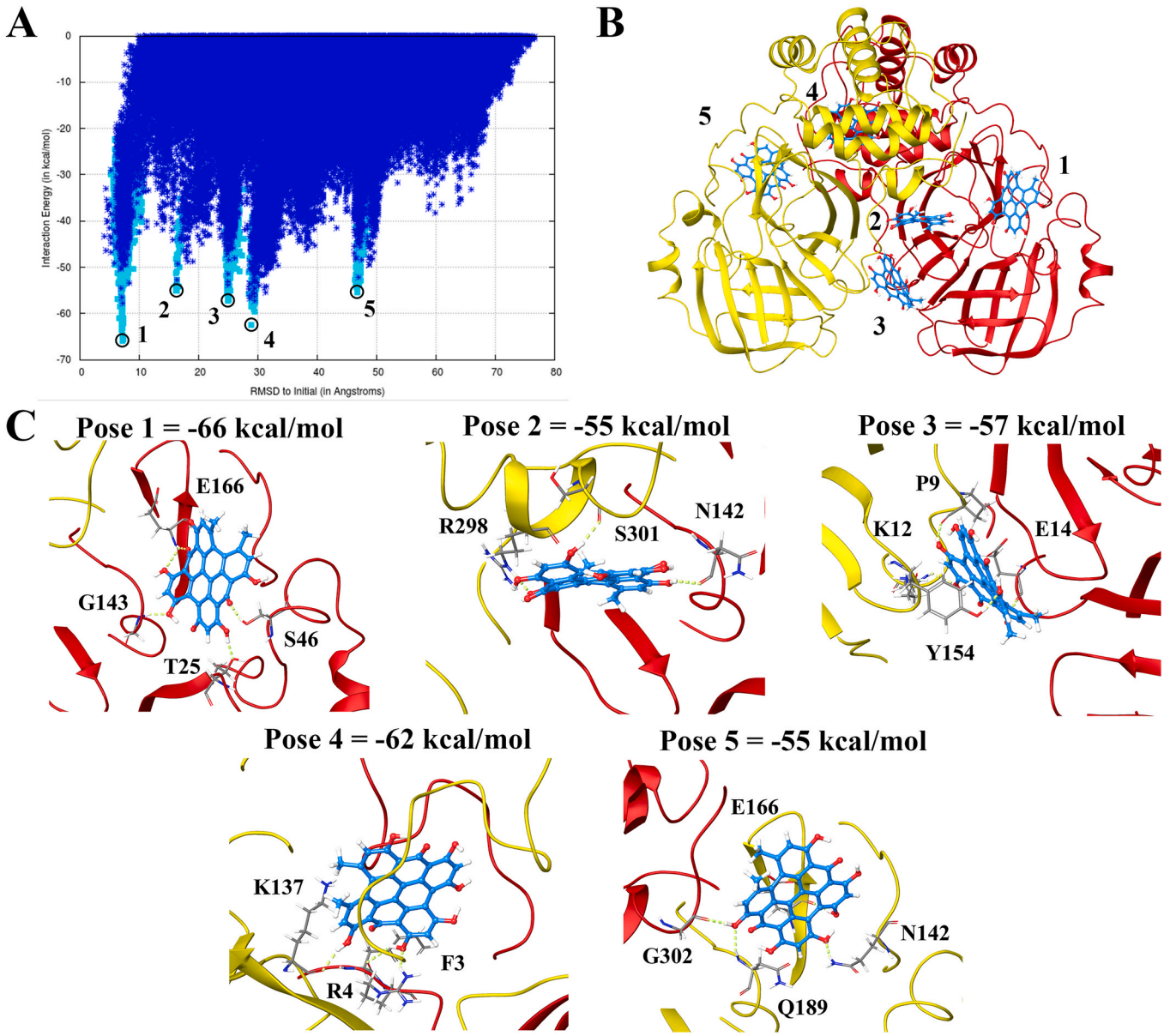

Fig. 4. PELE binding site search of SARS-CoV-2 M^{pro} dimer for hypericin. A) Interaction energy plot vs RMSD distance to initial position of hypericin in the substrate binding site of protomer A of SARS-CoV-2 M^{pro}. The global search is shown in dark blue, while local refined poses are shown in cyan. B) The best poses following local refinement are numbered from 1 to 5. C) Interaction between ligand poses and protein residues, with hydrogen-bonds indicated by dashed lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

hydrogen bond interactions with the same S1 subsite residues, with the addition of His163. In this pose (pose 2), the hydrogen bonds were all formed with the hydroxyl groups of cyanidin. For hypericin, the pose with the lowest interaction energy following refinement (−66 kcal/mol, pose 1) was also closest to the initial position in the active site of protomer A (Fig. 4). The active site on protomer B was also found to be a binding site for hypericin shown in pose 5, albeit with a weaker interaction energy (−55 kcal/mol), further demonstrating a preference for substrate binding to protomer A.

As well as the active sites, PELE also identified additional binding sites for hypericin with comparable interactions energies (Fig. 4). Interestingly, pose 3 identified by PELE simulations (−57 kcal/mol) overlapped with the cyanidin-3-O-glucoside following MD simulations. Pose 3 formed hydrogen bonds with prominent residues identified from contacts analysis Pro9, Lys12, Glu14, and Tyr154 (Fig. 2), suggesting the presence of an alternative binding site on M^{pro}. Pose 4 had an interaction energy of −62 kcal/mol, binding within the domain III interface of the dimer. This pose overlaps with a potential allosteric site we previously identified, forming hydrogen bonds with residues Phe3, Arg4, and Lys137 [15]. Domain III is involved in dimerization of M^{pro}, which is regulated by a salt-bridge interaction between E290 and R4 of opposite protomers [58]. Pose 2 is also located within the dimerization domain and had an interaction energy of −55 kcal/mol. This pose corresponds to an allosteric site identified by Günther et al. in a large scale x-ray crystallography screen and cell-based antiviral assays [14]. In their study, pelitinib bound in this pocket and demonstrated strong antiviral activity ($EC_{50} = 1.25$ μM, $CC_{50} = 13.96$ μM). Pose 2 of hypericin binds in a similar fashion, interacting with Ser301 from one protomer and Asn142 from the opposing protomer within the dimer. It has been reported that the integrity of this pocket is essential for enzyme activity [59]. Pose 2 also forms a hydrogen bond with Arg298, a vital residue for dimerization and stability of the S1 pocket [60].

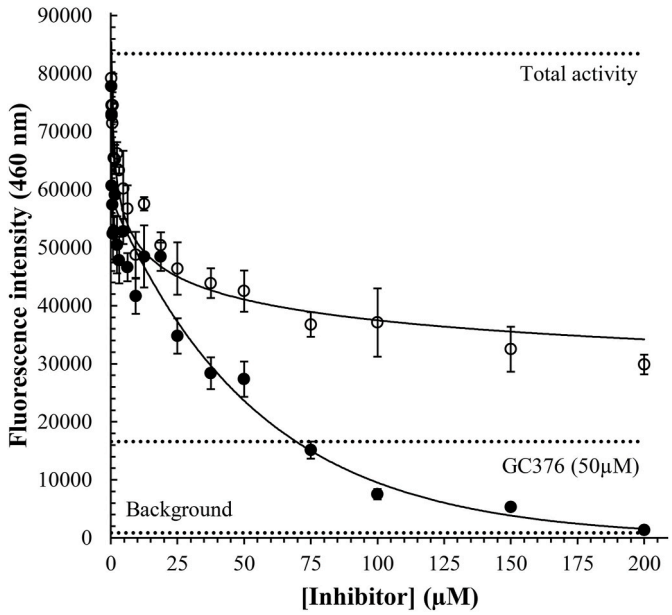

Fig. 5. Inhibition of the SARS-CoV-2 M^{pro} protease activity by hypericin and cyanidin-3-O-glucoside. A commercially available fluorogenic M^{pro} protease assay kit (BPS Bioscience, San Diego, CA, USA) consisting of a proprietary fluorogenic substrate (emission wavelength at 460 nm), was used to determine inhibition activity *in vitro*. Both test inhibitors, hypericin (closed circles), and cyanidin-3-O-glucoside (open circles), were shown to result in a concentration-dependent inhibition of M^{pro} activity (0.2 μM–200 μM concentration range was investigated). Average values ± SEM from triplicate determinations are shown. The average background (n = 6), total M^{pro} activity (n = 6), and inhibition by the positive control GC376 at 50 μM (n = 6), are highlighted (horizontal dotted lines).

Table 1
Inhibition of M^{pro} protease activity by GC376, hypericin, and cyanidin-3-O-glucoside. Percentage inhibition at a ligand concentration of 50 μM and IC_{50} values from *in vitro* fluorogenic M^{pro} protease assays.

Compound	n[a]	IC_{50} (μM)[b]	% M^{pro} inhibition at 50 μM ligand[b]
GC376	2	0.46[a]	92.6 ± 6.7
Hypericin	3	43.6 ± 12.2	59.9 ± 8.4
Cyanidin-3-O-glucoside	2	63.2 ± 15.4	23.2 ± 5.5

[a] n = number of independent experiments; IC_{50} for GC376 according to assay manufacturer (BPS Bioscience, San Diego, CA, USA).
[b] Average ± SEM.

3.4. Hypericin and cyanidin-3-O-glucoside inhibit protease activity *in vitro*

We have previously shown the *in vitro* inhibition activity of hypericin and cyanidin-3-O-glucoside against the SARS-CoV-2 M^{pro} [20]. In accordance with our previous experiments we observed a concentration-dependent inhibition of the M^{pro} with hypericin showing greater activity than cyanidin-3-O-glucoside (Fig. 5, Table 1). Our findings highlight that both of our test compounds are much less potent than the well-known covalent GC376 M^{pro} inhibitor (IC_{50} = 0.46 μM, Table 1), which was used as a positive control in our experiments. Here we extended our previous work and calculated collated IC_{50} values for hypericin and cyanidin-3-O-glucoside (IC_{50} = 43.6 and 63.2 μM, respectively, Table 1); although, not as potent as covalent inhibitors, these are within a biologically (μM) relevant range, suggesting that further evaluation of these compounds as potential M^{pro} inhibitors may be warranted.

4. Conclusion

MD simulations and PELE Monte Carlo simulations have highlighted that hypericin and cyanidin-3-O-glucoside bind to active and allosteric sites in the SARS-CoV-2 M^{pro}. MD simulation trajectories spanning a microsecond have shown that hypericin preferentially binds in the substrate binding site in both protomers of the dimeric M^{pro}, as well as known allosteric sites identified through PELE simulations. Finally, an *in vitro* enzymatic assay highlighted that both hypericin and cyanidin-3-O-glucoside inhibit M^{pro} activity in a concentration-dependent manner, albeit with much less potency than known covalent M^{pro} inhibitors such as GC376 used in our experiments. Nevertheless, potentially favourable bioactivity, toxicity and pharmacokinetic profiles may not completely rule out further investigation of these compounds. In this context, despite the checkered clinical history of hypericin, a synthetic version (SGX301, Soligenix, Inc.), has received an orphan drug designation from the US Food and Drug Administration for photodynamic therapy in cutaneous T-cell lymphoma. This clinical experience will be helpful for establishing protocols for potential evaluation of hypericin in a COVID-19 clinical setting.

Author contributions statement

TCK, AH, and VG conceptualized the aims and methodology, were involved in supervision, and production of the first draft of the manuscript. VG performed data analysis and curated data. KV performed and analysed the *in vitro* protease inhibition assay. JL and EP were involved in data analysis and curation, and in production of the first draft of the manuscript. All authors contributed to editing and reviewing the manuscript.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Epigenomic Medicine Program (TCK) is supported financially by McCord Research (Iowa, USA), which may have a financial interest in dietary compounds described in this work. However, there is no conflict of interest with respect to the inhibition of the SARS-CoV-2 main protease. The remaining co-authors also have no conflicts of interest.

Acknowledgements

We would like to acknowledge intellectual and financial support by McCord Research (Iowa, USA). JL is supported by an Australian Government Research Training Program Scholarship. We are indebted to Alfonso Perez Escudero and the team at Crowdfight COVID-19 for enabling access to supercomputing facilities. We thank the National Computing Infrastructure (NCI), and the Pawsey Supercomputing Centre in Australia (funded by the Australian Government). Further, we thank the Spartan High Performance Computing service (University of Melbourne), and the Partnership for Advanced Computing in Europe (PRACE) for awarding the access to Piz Daint, hosted at the Swiss National Supercomputing Centre (CSCS), Switzerland.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmgm.2021.108050.

References

[1] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (7798) (2020) 270–273.
[2] A.E. Gorbalenya, S.C. Baker, R.S. Baric, R.J. de Groot, C. Drosten, A.A. Gulyaeva, et al., The species Severe acute respiratory syndrome-related coronavirus: classifying

2019-nCoV and naming it SARS-CoV-2, Nature Microbiology 5 (4) (2020) 536–544.
[3] A.R. Fehr, S. Perlman, Coronaviruses: an overview of their replication and pathogenesis, Methods Mol. Biol. 1282 (2015) 1–23.
[4] B. Hu, H. Guo, P. Zhou, Z.-L. Shi, Characteristics of SARS-CoV-2 and COVID-19, Nat. Rev. Microbiol. 19 (3) (2021) 141–154.
[5] F. Benedetti, G.A. Snyder, M. Giovanetti, S. Angeletti, R.C. Gallo, M. Ciccozzi, et al., Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1, J. Transl. Med. 18 (1) (2020) 329.
[6] A.G. Harrison, T. Lin, P. Wang, Mechanisms of SARS-CoV-2 transmission and pathogenesis, Trends Immunol. 41 (12) (2020) 1100–1115.
[7] S. Yan, G. Wu, Potential 3-chymotrypsin-like cysteine protease cleavage sites in the coronavirus polyproteins pp1a and pp1ab and their possible relevance to COVID-19 vaccine and drug development, Faseb. J. 35 (5) (2021) e21573.
[8] C.A. Ramos-Guzmán, J.J. Ruiz-Pernía, I. Tuñón, Unraveling the SARS-CoV-2 main protease mechanism using multiscale methods, ACS Catal. 10 (21) (2020) 12544–12554.
[9] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, et al., Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, Nature 582 (7811) (2020) 289–293.
[10] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, et al., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors, Science 368 (6489) (2020) 409.
[11] J. Lee, L.J. Worrall, M. Vuckovic, F.I. Rosell, F. Gentile, A.-T. Ton, et al., Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site, Nat. Commun. 11 (1) (2020) 5877.
[12] M.D. Sacco, C. Ma, P. Lagarias, A. Gao, J.A. Townsend, X. Meng, et al., Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against M^{pro} and cathepsin L, Science Advances (50) (2020) 6, eabe0751.
[13] Mengist HM, Dilnessa T, Jin T. Structural basis of potential inhibitors targeting SARS-CoV-2 main protease. 2021;9(7).
[14] S. Günther, P.Y.A. Reinke, Y. Fernández-García, J. Lieske, T.J. Lane, H.M. Ginn, et al., X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease, Science 372 (6542) (2021) 642–646.
[15] J. Liang, C. Karagiannis, E. Pitsillou, K.K. Darmawan, K. Ng, A. Hung, et al., Site mapping and small molecule blind docking reveal a possible target site on the SARS-CoV-2 main protease dimer interface, Comput. Biol. Chem. 89 (2020) 107372.
[16] T.J. El-Baba, C.A. Lutomski, A.L. Kantsadi, T.R. Malla, T. John, V. Mikhailov, et al., Allosteric inhibition of the SARS-CoV-2 main protease: insights from mass spectrometry based assays**, Angew. Chem. Int. Ed. 59 (52) (2020) 23544–23548.
[17] Jiménez-Avalos G, Vargas-Ruiz AP, Delgado-Pease NE, Olivos-Ramirez GE, Sheen P, Fernández-Díaz M, et al. Comprehensive virtual screening of 4.8 k flavonoids reveals novel insights into allosteric inhibition of SARS-CoV-2 MPRO. Sci. Rep.. 2021;11(1):15452.
[18] Menéndez Cintia A, Byléhn F, Perez-Lemus Gustavo R, Alvarado W, de Pablo Juan J. Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease. Science Advances.6(37):eabd0345.
[19] M. Yuce, E. Cicek, T. Inan, A.B. Dag, O. Kurkcuoglu, F.A. Sungur, Repurposing of FDA-approved drugs against active site and potential allosteric drug-binding sites of COVID-19 main protease, Proteins: Structure, Function, and Bioinformatics 89 (11) (November 2021) 1425–1441.
[20] E. Pitsillou, J. Liang, C. Karagiannis, K. Ververis, K.K. Darmawan, K. Ng, et al., Interaction of small molecules with the SARS-CoV-2 main protease in silico and in vitro validation of potential lead compounds using an enzyme-linked immunosorbent assay, Comput. Biol. Chem. 89 (2020), 107408.
[21] L. Zhang, D. Lin, Y. Kusov, Y. Nian, Q. Ma, J. Wang, et al., α-Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: structure-based design, synthesis, and activity assessment, J. Med. Chem. 63 (9) (2020) 4562–4578.
[22] K. Świderek, V. Moliner, Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 Mpro by QM/MM computational methods, Chem. Sci. 11 (39) (2020) 10626–10630.
[23] Hattori S-i, N. Higashi-Kuwata, H. Hayashi, S.R. Allu, J. Raghavaiah, H. Bulut, et al., A small molecule compound with an indole moiety inhibits the main protease of SARS-CoV-2 and blocks virus replication, Nat. Commun. 12 (1) (2021) 668.
[24] L. Fu, F. Ye, Y. Feng, F. Yu, Q. Wang, Y. Wu, et al., Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease, Nat. Commun. 11 (1) (2020) 4417.
[25] M.I. Choudhary, M. Shaikh, A. Tul-Wahab, A. Ur-Rahman, In silico identification of potential inhibitors of key SARS-CoV-2 3CL hydrolase (Mpro) via molecular docking, MMGBSA predictive binding energy calculations, and molecular dynamics simulation, PLoS One (7) (2020) 15, e0235030-e.
[26] Y. Kumar, H. Singh, C.N. Patel, In silico prediction of potential inhibitors for the main protease of SARS-CoV-2 using molecular docking and dynamics simulation based drug-repurposing, J Infect Public Health 13 (9) (2020) 1210–1223.
[27] P.A. Nogara, F.B. Omage, G.R. Bolzan, C.P. Delgado, M. Aschner, L. Orian, et al., In silico studies on the interaction between Mpro and PLpro from SARS-CoV-2 and ebselen, its metabolites and derivatives, Molecular Informatics 40 (8) (August 2021), 2100028.
[28] Teli DM, Shah MB, Chhabria MT. In silico screening of natural compounds as potential inhibitors of SARS-CoV-2 main protease and spike RBD: targets for COVID-19. 2021;7(429).
[29] J. Liang, E. Pitsillou, C. Karagiannis, K.K. Darmawan, K. Ng, A. Hung, et al., Interaction of the prototypical α-ketoamide inhibitor with the SARS-CoV-2 main protease active site in silico: molecular dynamic simulations highlight the stability of the ligand-protein complex, Comput. Biol. Chem. 87 (2020), 107292.
[30] M.A.A. Ibrahim, K.A.A. Abdeljawaad, A.H.M. Abdelrahman, M.-E.F. Hegazy, Natural-like products as potential SARS-CoV-2 Mpro inhibitors: in-silico drug discovery, J. Biomol. Struct. Dyn. (2020) 1–13.
[31] R.R. Narkhede, A.V. Pise, R.S. Cheke, S.D. Shinde, Recognition of natural products as potential inhibitors of COVID-19 main protease (Mpro): in-silico evidences, Natural Products and Bioprospecting 10 (5) (2020) 297–306.
[32] C. S, S. DK, V. Ragunathan, P. Tiwari, P.B.D. A S, Molecular docking, validation, dynamics simulations, and pharmacokinetic prediction of natural compounds against the SARS-CoV-2 main-protease, J. Biomol. Struct. Dyn. (2020) 1–27.
[33] S. Yalçın, S. Yalçınkaya, F. Ercan, Determination of potential drug candidate molecules of the Hypericum perforatum for COVID-19 treatment, Curr Pharmacol Rep (2021) 1–7.
[34] A. Romeo, F. Iacovelli, M. Falconi, Targeting the SARS-CoV-2 spike glycoprotein prefusion conformation: virtual screening and molecular dynamics simulations applied to the identification of potential fusion inhibitors, Virus Res. 286 (2020), 198068.
[35] R. Islam, M.R. Parves, A.S. Paul, N. Uddin, M.S. Rahman, A.A. Mamun, et al., A molecular modeling approach to identify effective antiviral phytochemicals against the main protease of SARS-CoV-2, J. Biomol. Struct. Dyn. 39 (9) (2021) 3213–3224.
[36] P. Mohammadi Pour, S. Fakhri, S. Asgary, M.H. Farzaei, J. Echeverría, The signaling pathways, and therapeutic targets of antiviral agents: focusing on the antiviral approaches and clinical perspectives of anthocyanins in the management of viral diseases, Front. Pharmacol. 10 (2019) 1207.
[37] Jendželovská Z, Jendželovský R, Kuchárová B, Fedoročko P. Hypericin in the light and in the dark: two sides of the same coin. 2016;7(560).
[38] D.F. Birt, M.P. Widrlechner, K.D. Hammer, M.L. Hillwig, J. Wei, G.A. Kraus, et al., Hypericum in infection: identification of anti-viral and anti-inflammatory constituents, Pharm. Biol. 47 (8) (2009) 774–782.
[39] N.D. Weber, B.K. Murray, J.A. North, S.G. Wood, The antiviral agent hypericin has in vitro activity against HSV-1 through non-specific association with viral and cellular membranes, Antiviral Chem. Chemother. 5 (2) (1994) 83–90.
[40] Chen H, Muhammad I, Zhang Y, Ren Y, Zhang R, Huang X, et al. Antiviral activity against infectious bronchitis virus and bioactive components of Hypericum perforatum L. 2019;10(1272).
[41] Y. Fu, E. Zhou, Z. Wei, W. Wang, T. Wang, Z. Yang, et al., Cyanidin-3-O-β-glucoside ameliorates lipopolysaccharide-induced acute lung injury by reducing TLR4 recruitment into lipid rafts, Biochem. Pharmacol. 90 (2) (2014) 126–134.
[42] D. van Straten, V. Mashayekhi, H.S. de Bruijn, S. Oliveira, D.J. Robinson, Oncologic photodynamic therapy: basic principles, current clinical status and future directions, Cancers 9 (2) (2017) 19.
[43] E. Pitsillou, J. Liang, K. Ververis, A. Hung, T.C. Karagiannis, Interaction of small molecules with the SARS-CoV-2 papain-like protease: in silico studies and in vitro validation of protease activity inhibition using an enzymatic inhibition assay, J. Mol. Graph. Model. 104 (2021), 107851.
[44] E. Krissinel, K. Henrick, Protein interfaces, surfaces and assemblies service PISA at European Bioinformatics Institute, J. Mol. Biol. 372 (2007) 774–797.
[45] V. Zoete, M.A. Cuendet, A. Grosdidier, O. Michielin, SwissParam: a fast force field generation tool for small organic molecules, J. Comput. Chem. 32 (11) (2011) 2359–2368.
[46] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1) (1995) 43–56.
[47] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, et al., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, Software 1–2 (2015) 19–25.
[48] P. Bjelkmar, P. Larsson, M.A. Cuendet, B. Hess, E. Lindahl, Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models, J. Chem. Theor. Comput. 6 (2) (2010) 459–466.
[49] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, et al., CHARMM General Force Field (CGenFF): a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, J. Comput. Chem. 31 (4) (2010) 671–690.
[50] 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. Phys. 79 (2) (1983) 926–935.
[51] H.J.C. Berendsen, J.P.M. Postma, WFv Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (8) (1984) 3684–3690.
[52] M. Parrinello, A. Rahman, Crystal structure and pair potentials: a molecular-dynamics study, Phys. Rev. Lett. 45 (14) (1980) 1196–1199.
[53] B. Hess, H. Bekker, H.J. Berendsen, J.G. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (12) (1997) 1463–1472.
[54] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (12) (1993) 10089–10092.
[55] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1) (1996) 33–38, 27-8.
[56] A. Amadei, A.B. Linssen, H.J. Berendsen, Essential dynamics of proteins, Proteins 17 (4) (1993) 412–425.
[57] D. Lecina, J.F. Gilabert, V. Guallar, Adaptive simulations, towards interactive protein-ligand modeling, Sci. Rep. 7 (1) (2017) 8466.

[58] J. Shi, J. Song, The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain, FEBS J. 273 (5) (2006) 1035–1045.
[59] J. Tan, K.H.G. Verschueren, K. Anand, J. Shen, M. Yang, Y. Xu, et al., pH-dependent conformational flexibility of the SARS-CoV main proteinase (Mpro) dimer: molecular dynamics simulations and multiple X-ray structure analyses, J. Mol. Biol. 354 (1) (2005) 25–40.
[60] J. Shi, J. Sivaraman, J. Song, Mechanism for controlling the dimer-monomer switch and coupling dimerization to catalysis of the severe acute respiratory syndrome coronavirus 3C-like protease, J. Virol. 82 (9) (2008) 4620–4629.