Computational and Structural Biotechnology Journal 19 (2021) 3364–3371

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

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

Interaction of 8-anilinonaphthalene-1-sulfonate with SARS-CoV-2 main protease and its application as a fluorescent probe for inhibitor identification

Peerapon Deetanya [a,b], Kowit Hengphasatporn [c], Patcharin Wilasluck [a,b], Yasuteru Shigeta [c], Thanyada Rungrotmongkol [d,e], Kittikhun Wangkanont [a,b,*]

[a] Center of Excellence for Molecular Biology and Genomics of Shrimp, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[b] Molecular Crop Research Unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[c] Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan
[d] Program in Bioinformatics and Computational Biology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[e] Structural and Computational Biology Research Unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

ARTICLE INFO

Article history:
Received 2 April 2021
Received in revised form 29 May 2021
Accepted 30 May 2021
Available online 05 June 2021

Keywords:
8-Anilinonaphthalene-1-sulfonate
Fluorescent probe
Binding assay
SARS-CoV-2
Protease inhibitor
Flavonoids

ABSTRACT

The 3C-like main protease of SARS-CoV-2 ($3CL^{Pro}$) is responsible for the cleavage of the viral polyprotein. This process is essential for the viral life cycle. Therefore, $3CL^{Pro}$ is a promising target to develop antiviral drugs for COVID-19 prevention and treatment. Traditional enzymatic assays for the identification of $3CL^{Pro}$ inhibitors rely on peptide-based colorimetric or fluorogenic substrates. However, the COVID-19 pandemic has limit or delay access to these substrates, especially for researchers in developing countries attempting to screen natural product libraries. We explored the use of the fluorescent probe 8-anilinonaphthalene-1-sulfonate (ANS) as an alternative assay for inhibitor identification. Fluorescence enhancement upon binding of ANS to $3CL^{Pro}$ was observed, and this interaction was competitive with a peptide substrate. The utility of ANS-based competitive binding assay to identify $3CL^{Pro}$ inhibitors was demonstrated with the flavonoid natural products baicalein and rutin. The molecular nature of ANS and rutin interaction with $3CL^{Pro}$ was explored with molecular modeling. Our results suggested that ANS could be employed in a competitive binding assay to facilitate the identification of novel SARS-CoV-2 antiviral compounds.

1. Introduction

In addition to vaccines, antiviral drug development is another frontier to combat the Corona Virus Disease 2019 (COVID-19). COVID-19 is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus produces polyproteins that require further processing by proteases to liberate functional proteins that are required for the viral life cycle [1]. Inhibition of the 3C-like protease ($3CL^{Pro}$), or the main protease, is a known approach to inhibit SARS-CoV-1 and Middle East Respiratory Syndrome (MERS) virus replication [2]. Therefore, there has been significant interest in the scientific community to identify inhibitors of SARS-CoV-2 $3CL^{Pro}$, thereafter referred to as $3CL^{Pro}$, since the beginning of the outbreak [3,4]. Several research groups have reported promising results in using protease inhibitors to control SARS-CoV-2 infection [3–5]. Novel computational methodologies have also been developed to screen for inhibitor of $3CL^{Pro}$ [6]. Still, the discovery of novel inhibitors is necessary to overcome the high mutation rate of the RNA virus [7].

Library of compounds isolated from natural sources offers a valuable source for $3CL^{Pro}$ inhibitor discovery. We have been interested in the exploration of local natural products as potential $3CL^{Pro}$ inhibitors. Traditional assays for protease inhibitor identification employ colorimetric or fluorogenic peptide substrates. This approach has been quite successful in identification of $3CL^{Pro}$ inhibitors [3,4]. The peptide substrates could readily be synthesized by existing peptide synthesis methods. However, such services are not available in many developing countries with rich natural product resources, thus investigators are required to order synthetic peptides from abroad. With the COVID-19 pandemic, which results

* Corresponding author at: 254 Phayathai Road, Wangmai, Pathumwan, Bangkok 10330, Thailand.
E-mail address: kittikhun.w@chula.ac.th (K. Wangkanont).

https://doi.org/10.1016/j.csbj.2021.05.053

in laboratory and business closure, we have experienced a significant import delay of peptide substrates. Therefore, we seek to develop a new assay for $3CL^{Pro}$ inhibitor identification that could be used for natural product and chemical library screening. The enzyme inhibition activity could later be confirmed with a standard peptide-based assay. The assay should also be relatively inexpensive so that investigators with limited funding could employ the assay. Because the native cleavage sequences of $3CL^{Pro}$ (TSAVLQ-SGFRK and SGVTFQ-GKFKK, with - indicating the cleavage site)[8] contain a few hydrophobic amino acids and the active site of $3CL^{Pro}$ has positively charged histidine residues such as H41, H163, H164, and H172, we envision that the fluorescent probe 8-anilinonaphthalene-1-sulfonate (ANS) could be utilized in the desired assay.

ANS is a water-soluble molecule that exhibits increased fluorescence upon burial into a pocket in a protein [9]. ANS-based binding assay has been utilized to investigate protein–ligand interactions and protein unfolding processes [10–13]. ANS has previously been employed to characterize the folding process of SARS-CoV-1 $3CL^{Pro}$, in which ANS binds to the exposed hydrophobic region when the protein is unfolded [14]. However, the interaction of ANS and the native $3CL^{Pro}$ has not yet been explored. If ANS could bind in the active site of $3CL^{Pro}$, an increase in ANS fluorescence should be observed. Inhibitors of $3CL^{Pro}$ should be able to compete with ANS and resulted in a reduction in fluorescence. This equilibrium binding approach is also more convenient and requires less sophisticated equipment to perform than the time-sensitive enzyme kinetic screens or end point assays. In this work, we demonstrated that ANS could bind in the active site of $3CL^{Pro}$ by competition with the fluorogenic peptide substrate E(EDANS)TSAVLQSGFRK(DABCYL). In addition, ANS could compete with baicalein that is known to bind in the active site [15]. We also demonstrated the utility of the ANS-based binding assay for the identification of rutin as a $3CL^{Pro}$ inhibitor, as predicted by several computational studies [16–22]. The ANS-based competitive binding assay will be valuable in a cost-effective screen for novel $3CL^{Pro}$ inhibitors, especially when access to a peptide substrate is limited.

2. Results

The interaction of ANS with $3CL^{Pro}$ was first explored by monitoring the increased fluorescence in an ANS-$3CL^{Pro}$ mixture compared to the solution with no $3CL^{Pro}$ (Fig. 1A). When excited at 345 nm, the increased fluorescence with an emission maximum at 455 nm was observed. Therefore, ANS could be shielded from the aqueous solvent when interacting with $3CL^{Pro}$. The excitation and emission wavelengths at 345 and 455 nm, respectively, were used in subsequent experiments. The saturation binding curve suggested a simple binding interaction with the apparent dissociation constant (K_D) of 57 ± 7 μM (Fig. 1B).

To determine whether ANS could bind in the active site of $3CL^{Pro}$, inhibition experiments were performed using the fluorogenic peptide substrate E(EDANS)TSAVLQSGFRK(DABCYL) (Fig. 2). The non-linear fit of the kinetic data revealed the Michaelis constant (K_M) for the fluorogenic substrate to be 51 ± 9 μM. The addition of ANS and the non-linear fit of the resulting kinetic data to the competitive inhibition model gave the inhibitory constant (K_I) of 188 ± 24 μM.

To demonstrate the utility of ANS as a fluorescent probe to identify a $3CL^{Pro}$ inhibitor, we performed competitive binding and competitive inhibition experiments with baicalein and rutin (Fig. 3A). Baicalein is a flavonoid natural product known to inhibit $3CL^{Pro}$ [15]. X-ray crystallography revealed that baicalein binds in the active site of $3CL^{Pro}$. Thus, the reduction of ANS fluorescence by baicalein should be evidence that ANS binds in the active site of

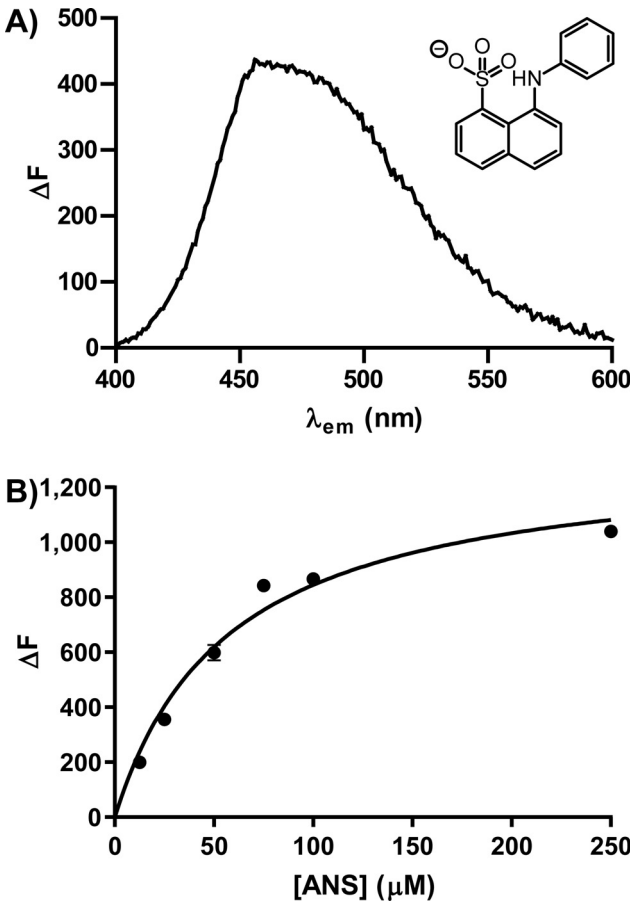

Fig. 1. A) Chemical structure and fluorescence emission spectrum of ANS (50 μM) bound to $3CL^{Pro}$. B) Saturation binding curve to ANS to $3CL^{Pro}$.

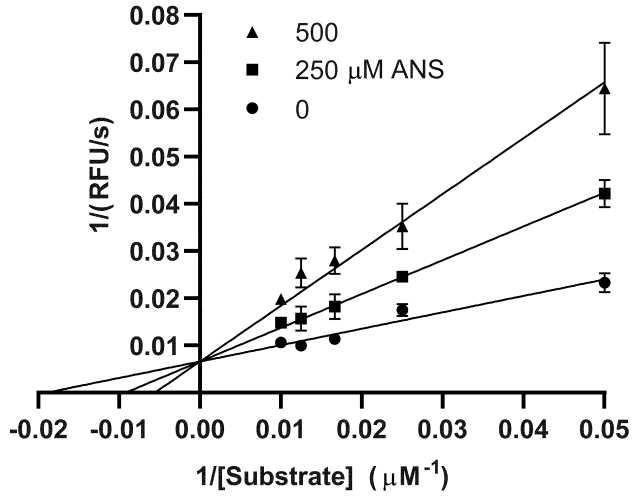

Fig. 2. Lineweaver–Burk plot of an initial rate enzyme kinetic experiment of $3CL^{Pro}$ at different substrate concentrations with ANS presented at 0, 250, and 500 μM. Trend lines were created with parameters derived from non-linear fit of the original data.

$3CL^{Pro}$. Indeed, the reduction of ANS fluorescence was observed as baicalein was titrated (Fig. 3B). For the ANS competitive binding assay, the half-maximal inhibitory concentration (IC_{50}) of baicalein was 42 ± 2 μM, and the corresponding K_I obtained from the Cheng-Prusoff equation was 15.2 ± 0.7 μM. Inhibition of $3CL^{Pro}$ by baicalein was also confirmed in an activity assay using the fluorogenic

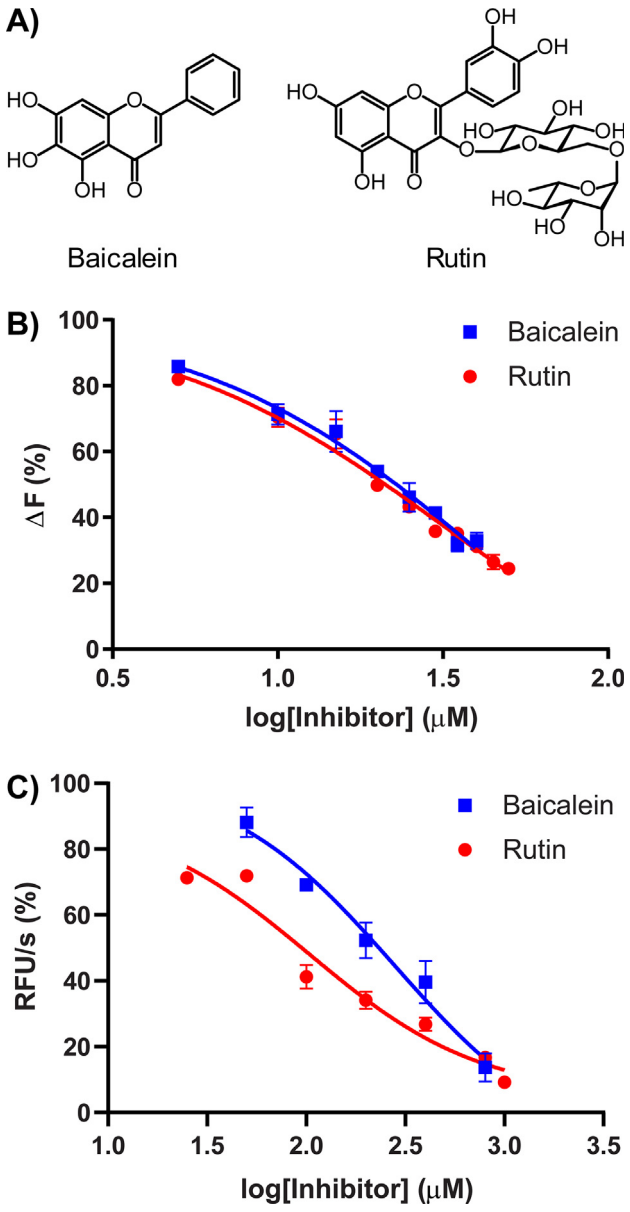

Fig. 3. A) Chemical structures of baicalein and rutin. B) Reduction in $3CL^{Pro}$-bound ANS fluorescence upon titration of baicalein and rutin. ANS concentration was at 100 μM. C) Inhibition of $3CL^{Pro}$ enzymatic activity by baicalein and rutin. The fluorogenic substrate concentration was at 25 μM.

peptide substrate E(EDANS)TSAVLQSGFRK(DABCYL) (Fig. 3C). For the enzyme activity inhibition assay, IC_{50} for baicalein was 280 ± 1 μM, and the corresponding K_I obtained from the Cheng-Prusoff equation was 188 ± 1 μM.

Similar to baicalein, rutin is a flavonoid previously predicted computationally to bind and inhibit $3CL^{Pro}$ [18–20,22,23]. We have also independently predicted binding of rutin within the active site of $3CL^{Pro}$ using a molecular docking (Fig. 4). The binding energy was −10.0 kcal/mol. Enhanced binding interactions were found for rutin binding (hydrophobic contacts with the 13 residues in chain A and 3 residues in chain B as well as four hydrogen bonds at the disaccharide rutinose (2.7–2.9 Å). This is somewhat inconsistent with the previous *in silico* studies using only one chain of protease (Table 1). Our results suggest that the termini of the other chain in the $3CL^{Pro}$ dimer are important in the interactions with rutin.

To gain further insight into the interaction between rutin and $3CL^{Pro}$ at the atomic level, the fragment molecular orbital calculation with a high level of theory (FMO-RIMP2/PCM) was applied on the rutin-$3CL^{Pro}$ complex. The total pair interaction energy (PIE^{total}) and energy contribution from the residues within the 7-Å sphere of rutin were plotted in Fig. 4, where the negative and positive values represent the ligand stabilization and destabilization, respectively. Only the residues with energy contribution $\leq$ -5.0 kcal/mol and $\geq$ 5.0 kcal/mol were discussed. The PIE^{total} value of −85.19 kcal/mol indicates that the complexation between rutin and dimeric $3CL^{Pro}$ was likely stable. Among the 44 fragments (amino acid residues in the FMO calculation), there are 12 residues in both chains (chain A: L27, H41, Y54, L141, C145, M165, E166, H172, D187, R188; and chain B: V303, F305), contributed to rutin binding with energy stabilization in a range of −5.2 to −25.5 kcal/mol. E166 provides most significantly to the high binding affinity of rutin in terms of electrostatic, charge transfer, dispersion, and solvation effect. The catalytic dyad H41 and C145 play a vital role in binding with the quercetin moiety of rutin. Although rutin is likely to be a non-covalent inhibitor, it shares corresponding binding patterns with the reported inhibitors of SARS-CoV-2 $3CL^{Pro}$ [24].

Despite predictions from other research groups, including us, no experimental demonstration of rutin inhibitory activity against $3CL^{Pro}$ has been reported. To experimentally confirm that rutin could bind $3CL^{Pro}$, we performed the ANS-based competitive binding assay. Titration of rutin into a solution of ANS-bound $3CL^{Pro}$ resulted in the reduction of ANS fluorescence. IC_{50} of rutin was 31 ± 1 μM, and the corresponding K_I obtained from the Cheng-Prusoff equation was 11.3 ± 0.4 μM for the ANS competitive binding assay (Fig. 3B). Enzyme inhibition was also confirmed using the fluorogenic peptide substrate (Fig. 3C). For the enzyme activity inhibition assay, IC_{50} for rutin was 104 ± 1 μM, and the corresponding K_I obtained from the Cheng-Prusoff equation was 69 ± 1 μM.

To obtain structural information for interaction between ANS and rutin, we attempted to crystallize these compounds with $3CL^{Pro}$. Crystals that diffracted to around 1.5 Å were readily obtained (Table 2). However, ANS or rutin was not observed in the active site, but a molecule of DMSO was present instead (Fig. 5A). We initially speculated that DMSO might interfere with ANS binding to a certain extent. However, the ANS saturation binding curve (Fig. 1B) was obtained in the presence of 1% DMSO as in all the assay buffers. The omission of DMSO for the saturation binding experiment yielded a virtually identical curve (data not shown). The crystal structure of baicalein is already available for comparison (Fig. 5B) [15]. To obtain an ANS-bound $3CL^{Pro}$ structure, we performed molecular docking (Fig. 5C). The binding energy was −6.9 kcal/mol. ANS showed hydrophobic contacts with 9 residues (H41, M49, F140, L141, C145, H164, M165, E166, and Q189) in chain A, including a hydrogen bond between its sulfonate group with catalytic H163 (3.1 Å), and with the N-terminal serine residue of chain B. The structure of the rutin-$3CL^{Pro}$ complex from docking mentioned previously was displayed in the same orientation for comparison (Fig. 5D). From our models and the crystal structure of the baicalein-$3CL^{Pro}$ complex, ANS, baicalein, and rutin occupy an overlapping space in the active site of $3CL^{Pro}$, which is also the space occupied by the peptidomimetic covalent inhibitor N3 (Fig. 6) [3].

3. Discussion

ANS could bind $3CL^{Pro}$, which resulted in an increased fluorescence at 455 nm. Although ANS has been used previously to study the unfolding process of SARS-CoV-1 $3CL^{Pro}$, our results demonstrated that ANS could bind to a pocket in the native $3CL^{Pro}$. To confirm that the ANS binding pocket is within the active site of $3CL^{Pro}$, we first showed that ANS is a competitive inhibitor of $3CL^{Pro}$. Using

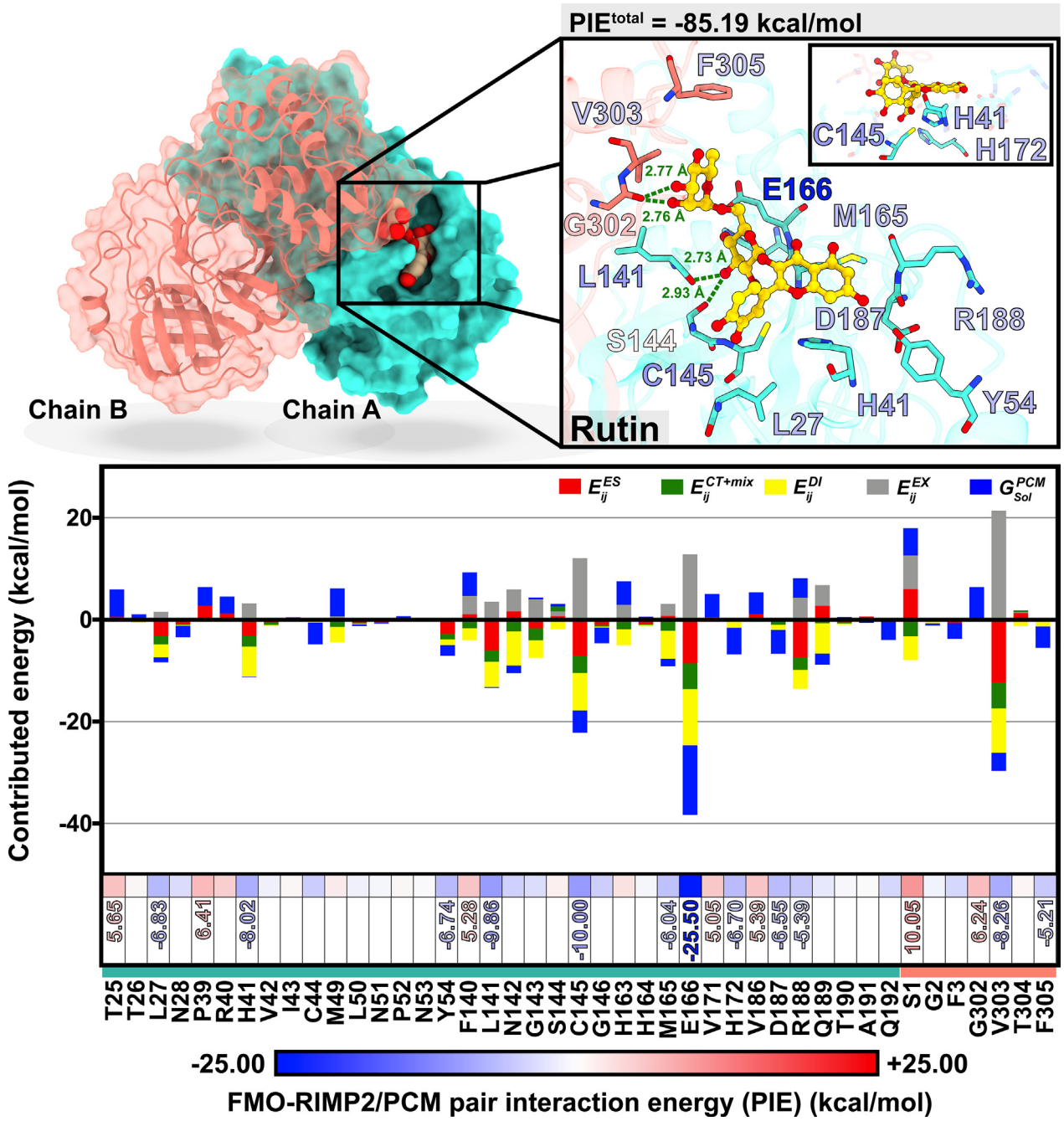

Fig. 4. FMO-RIMP2/PCM pair interaction energy (PIE^{total}) and energy components (electrostatic intyeraction (E_{ij}^{ES}), charge transfer with higher-order mixed terms energies (E_{ij}^{CT+mix}) dispersion (E_{ij}^{DI}), exchange-repulsion (E_{ij}^{EX}), and the PCM solvation effect (G_{Sol}^{PCM})) for rutin interacting with individual residues of dimeric SARS-CoV-2 $3CL^{Pro}$. Rutin orientation and interactions at the active site of SARS-CoV-2 $3CL^{Pro}$ are shown above, where the green dashed line represents the hydrogen bonding. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Predicted binding mode of rutin toward SARS-CoV-2 $3CL^{Pro}$.

Model	PDB ID	Software	Hydrogen bonding residues	Reference
dimer	6LU7	Autodock Vina	L141, S144, G302 (Chain B)	This study
monomer	6LU7	Autodock Vina	T26, Y54, L141, M165, E166	[18]
monomer	5R82	Autodock Vina	T45, G143, E166	[17]
monomer	6LU7	Autodock Vina	D178, R188, T190	[20]
monomer	6LU7	Autodock Vina	N.A.	[21]
monomer	6LU7	Glide	N142, G143, C145, T190	[22]
monomer	6LU7	LeDock	F140, E166, T26, L141, S144, C145, H163	[19]

Table 2
Crystallographic data collection and refinement statistics.

Data collection statistics	
Wavelength (Å)	0.97872
Resolution range (Å)[a]	27.57–1.45 (1.47–1.45)
Space group	C 2
Unit cell dimensions	113.2, 53.0, 44.7 90.0, 103.1, 90.0
Total number of reflections	339,475 (16,721)
Number of unique reflections	43,879 (2,124)
Multiplicity	7.7 (7.9)
Completeness (%)	96.1 (93.8)
Mean I/σ(I)	15.3 (2.2)
Wilson B factor (Å^2)	12.21
R_{merge}	0.077 (1.035)
R_{meas}	0.082 (1.108)
R_{pim}	0.030 (0.394)
$CC_{1/2}$	0.999 (0.803)
Refinement Statistics	
Resolution range (Å)[a]	27.57–1.45 (1.50–1.45)
R-factor	0.1664 (0.2567)
R-free (5%)	0.2010 (0.2731)
Number of atoms	2,935
Protein	2,473
DMSO	16
Water	446
Number of protein residues	305
RMSD for bonds (Å)	0.005
RMSD for angles (deg)	0.822
Estimated coordinate error (ML, Å)	0.14
Ramachandran favored (%)	98.68
Ramachandran outliers (%)	0.33
Average isotropic B factor (Å^2)	20.96
Protein	18.86
DMSO	27.57
Water	32.37
PDB accession code	7DJR

[a] Statistics for the highest-resolution shell are given in parentheses.

the fluorogenic peptide substrate, we observed a reduction in catalytic activity in the presence of ANS that could be fitted with the competitive inhibition model. Therefore, ANS binding site overlaps with the substrate-binding site.

The binding affinity of ANS towards 3CL^{Pro} could be extracted from the saturation binding curve and the competitive inhibition experiment. However, the K_I value obtained from the competitive inhibition assay was higher than the K_D value obtained from the saturation binding curve of ANS to 3CL^{Pro}. Several factors might have resulted in the discrepancy, including the difference in assay conditions. The major factor could be the poor solubility of the peptide substrate that limits data points at high substrate concentrations in the ANS competitive inhibition experiment. The maximum velocity estimation and K_M may not be accurate. Another potential source of affinity discrepancy could be due to multiple ANS binding sites on 3CL^{Pro}. It is possible that 3CL^{Pro} possesses a high-affinity binding site for ANS that does not affect the 3CL^{Pro} catalytic activity. We cannot completely rule out this possibility with the current data. However, the saturation binding curve is evidently hyperbolic. Thus, even if there are multiple binding sites, the binding affinities are likely similar. Therefore, the K_D value obtained from the saturation binding curve might reflect the binding affinity of ANS to 3CL^{Pro} more accurately than the K_I value from the competitive inhibition experiment.

To confirm that ANS indeed binds in the substrate-binding site, we investigated whether baicalein could compete with ANS binding. Baicalein had already been crystallized in complex with 3CL^{Pro} [15]. Thus, the binding site is known. Baicalein had also been shown to inhibit 3CL^{Pro} in enzymatic inhibition assays [15] including one independently reported by us. Therefore, baicalein could be used to validate the ANS utility in inhibitor identification. Indeed, baicalein could compete with ANS for 3CL^{Pro} binding. The K_I value of baicalein from the ANS competitive binding assay (11.3 μM) is comparable with the K_D value previously obtained from an isothermal titration calorimetry experiment (4.03 μM) [15]. Therefore, the ANS-based competitive binding assay yielded similar results to other techniques used to identify 3CL^{Pro} inhibitors.

We next applied the ANS-based competitive binding assay to show that rutin is an inhibitor of 3CL^{Pro}. Rutin has been predicted computationally by multiple research groups to be a 3CL^{Pro} inhibitor (Table 1). We have shown that rutin is competitive with both ANS and the fluorogenic peptide substrate. Therefore, we have both demonstrated the inhibitory activity of rutin and also shown that ANS could be used to identify novel inhibitors.

To investigate the molecular details of the interaction between ANS and rutin with 3CL^{Pro}, we attempted to crystallized this compound with 3CL^{Pro}. However, only the DMSO-bound structure was obtained. Therefore, we employed computational methods to investigate the 3CL^{Pro}-ligand interactions. Both ANS and rutin binds in the active site overlapping the known baicalein[15] or N3 inhibitor binding site[3]. The sulfonate group of ANS interacts with a histidine residue, H163, as expected. H163 is not a part of the catalytic dyad but forms the S1 pocket for substrate binding and is also recognized by the γ-lactam moiety of the N3 inhibitor. Rutin interacts mainly with the S1 and S2 pockets, and also interact with residues from the neighboring 3CL^{Pro} subunit.

Compared with the crystal structure of DMSO-bound 3CL^{Pro}, other ligand-bound structures required movement of the M49 side chain and distortion of the corresponding helix. This required conformational change may explain why we did not observe any ligands in the DMSO-bound, tightly packed C2 crystal form. The baicalein-bound crystal was in the P1 space group, in which the M49-containing helix is more solvent-exposed. Therefore, further crystal engineering may be required to obtain new crystal forms for the ANS- and rutin-bound 3CL^{Pro} structures. In addition, because it appeared that the readily obtained C2 crystal form has ligand bias, it may be beneficial for future investigators to screen for other crystal forms to be used in other inhibitor discovery approaches.

In conclusion, ANS could be used in a competitive binding assay to identify ligands that interact with 3CL^{Pro} in the active site. Using baicalein as a model compound with a known binding location and 3CL^{Pro} inhibitory activity, we showed that ANS binding was competitive with baicalein. Thus, ANS likely binds in the active site of 3CL^{Pro}. We also applied this methodology to demonstrated that rutin, a compound predicted to bind and inhibit 3CL^{Pro}, was indeed binding competitively with ANS and could inhibit 3CL^{Pro} activity. These results demonstrated the utility of ANS in the identification of potential 3CL^{Pro} inhibitors. This inhibitor identification strategy could be employed when researchers have limited access to peptide-based protease assays.

4. Materials and methods

4.1. Protein expression, ANS binding assay, and enzyme kinetics

SARS-CoV-2 3CL^{Pro} with no tags at the termini was expressed, purified, and stored exactly as previously described for SARS-CoV-1 3CL^{Pro} [25]. The gene was synthesized and codon-optimized for expression in *Escherichia coli* based on the amino acid sequence in GenBank accession number NC_045512. 3CL^{Pro} concentrations were determined using the absorbance value at 280 nm and the extinction coefficient of 32,890 M^{-1} cm^{-1}.

All assays were performed with BioTek Synergy H1 microplate reader using PBS with 1 mM DTT and 1% DMSO as the reaction buffer. The volume was fixed at 100 μL. ANS binding assay was

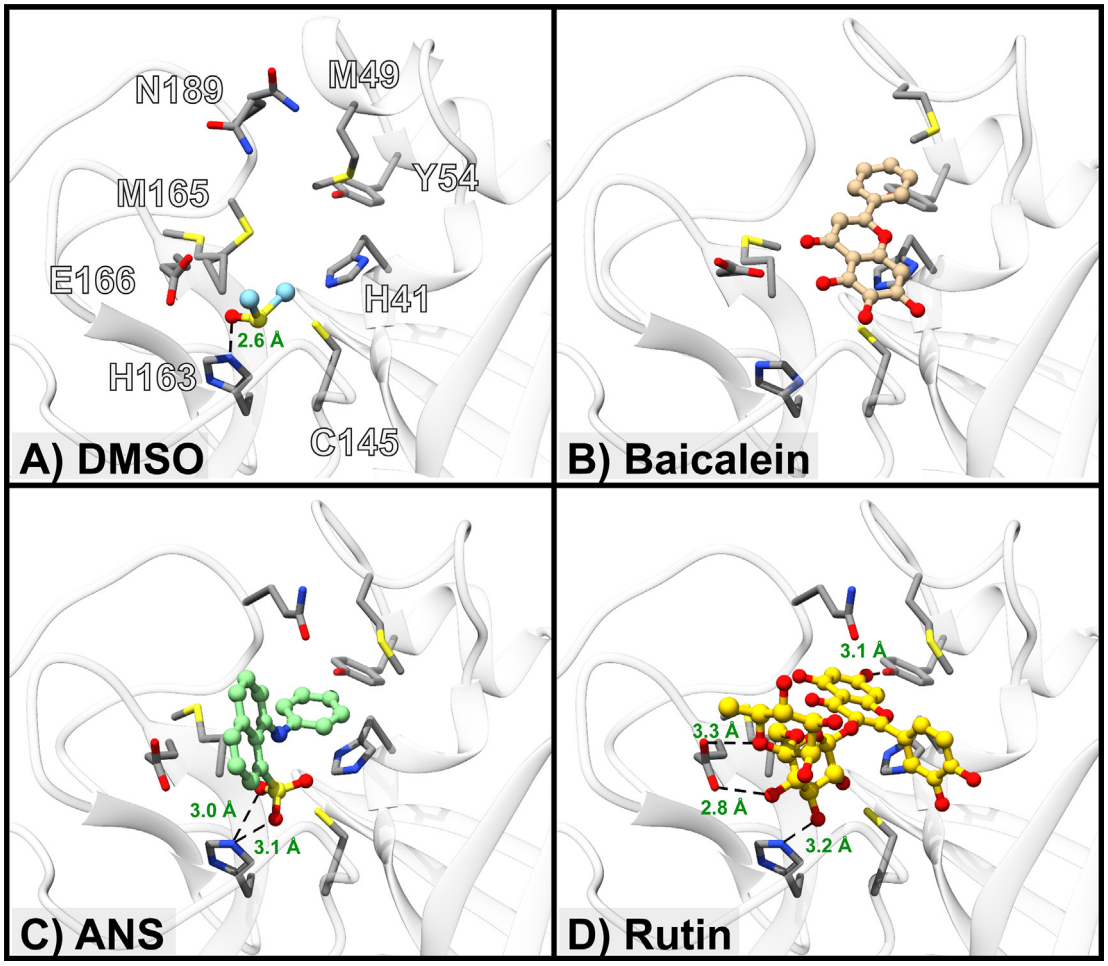

Fig. 5. A) Crystal structure of 3CL^{Pro} crystallized in the presence of 5 mM ANS but has DMSO (green) bound instead of ANS. B) Crystal structure of 3CL^{Pro} with baicalein bound (PDB ID 6M2N). C-D) Conformations of ANS and rutin binding to 3CL^{Pro} by molecular docking. Potential non-covalent interactions are shown as dash lines with distances in Ångstrom (Å). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

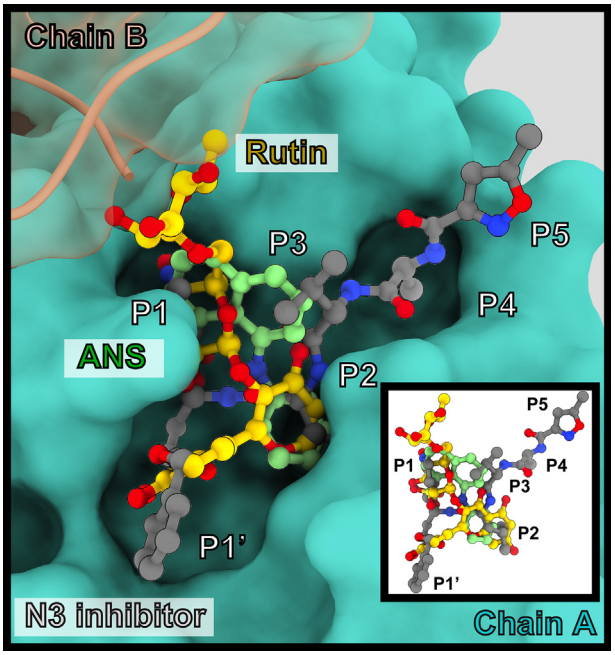

Fig. 6. Position of ANS and rutin in the active site compared to the peptidomimetic N3 inhibitor.

performed with 3CL^{Pro} at 5 μM. The excitation and emission wavelengths used were 345 and 455 nm, respectively. For enzyme kinetics, the fluorogenic substrate E(EDANS)TSAVLQSGFRK(DABCYL) (Biomatik) was used with 0.2 μM of 3CL^{Pro}. The excitation and emission wavelengths employed were 340 and 490 nm, respectively. GraphPad Prism 8 (San Diego, California USA, www.graphpad.com) was used for graphing and non-linear fit. [26]

4.2. X-ray crystallography

3CL^{Pro} was crystallized as described previously (100 mM MES pH 6.5, 15% PEG 4,000, and 5% DMSO)[27] but in the presence of 5 mM ANS. Microseeding was required to obtain crystals suitable for data collection. Crystals were cryoprotected with the crystallization buffer with 5 mM ANS but the PEG 4,000 concentration raised to 35%. Diffraction data were collected at the Life Sciences Collaborative Access Team beamline 21-ID-F (Advanced Photon Source, Argonne National Laboratory). The data were indexed and integrated with XDS [28]. Space group determination and scaling were performed with AIMLESS [29]. Molecular replacement phasing was accomplished with Phaser [30] using a previously reported SARS-CoV-2 3CL^{Pro} crystal structure (PDB ID 5RE9) as a search model. Refinement and model adjustments were performed with phenix.refine [31,32] and COOT [33] respectively. Structure figures were created with UCSF Chimera [34]. The crystal structure

and the associated experimental data were deposited at the Protein Data Bank under the accession code 7DJR.

4.3. Prediction of ligand–protein binding and interactions

The 3D structures of the ANS and rutin were built by GaussView 6.0.16 and then optimized with DFT-B3LPY/6–31 g(d) basis set using gaussian 16 [35]. The $3CL^{Pro}$ protein covalently bonded with the N3 inhibitor was retrieved from the protein databank (PBD ID 6LU7) [3]. The prepared protein structure without inhibitor binding was taken from our previous studies [24,36]. To construct the initial structure of the ligand-$3CL^{Pro}$ complex, the optimized structures of deprotonated ANS and rutin were separately docked into the binding site of N3 inhibitor using AutoDock Vina 1.1.2 according to the standard procedure [37]. The ligand orientation with the highest binding affinity was chosen for ligand–protein analysis and further *ab initio* fragment molecular orbital (FMO) calculation [38–40].

To provide a detailed insight into the binding of rutin to $3CL^{Pro}$, the pair interaction energy (PIE) calculation using the second-order Møller–Plesset perturbation theory (MP2) with the resolution-of-the-identity (RI) approximation and polarizable continuum model (PCM) solvation effect (FMO-RIMP2/PCM) was carried out by GAMESS software [41,42] with the 32 cores computer cluster using a generalized distributed data interface (GDDI). The complex was divided into small fragments (one residue/ligand per fragment) called a monomer, and then all fragments were performed in a parallel manner by the molecular orbital (MO) calculation [43]. Each pair interaction of monomers (I and J) was computed the PIE by a summation of several energy contributions among the clustered residues to identify the essential interacted residue for ligand binding by the following equation:

$$\mathrm{PIE} = \Delta E_{IJ}^{ES} + \Delta E_{IJ}^{CT+mix} + \Delta E_{IJ}^{DI} + \Delta E_{IJ}^{EX} + \Delta G_{Sol}^{PCM} \quad (1)$$

where E_{IJ}^{ES} is electrostatic interaction, E_{IJ}^{CT+mix} is charge transfer with higher-order mixed terms energies E_{IJ}^{DI} dispersion, E_{IJ}^{EX} is exchange-repulsion, and G_{Sol}^{PCM} is the polarizable continuum model (PCM) solvation effect [44–46].

CRediT authorship contribution statement

Peerapon Deetanya: Methodology, Investigation, Writing - review & editing. **Kowit Hengphasatporn:** Methodology, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Patcharin Wilasluck:** Methodology, Investigation, Validation. **Yasuteru Shigeta:** Methodology, Investigation, Resources, Funding acquisition. **Thanyada Rungrotmongkol:** Methodology, Investigation, Writing - review & editing, Resources, Funding acquisition. **Kittikhun Wangkanont:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition, Supervision, Resources.

Declaration of Competing Interest

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

Acknowledgements

This research used resources of the Advanced Photon Source, a U. S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817). This project is partly funded by the National Research Council of Thailand. KW is partially supported by the Institute for the Promotion of Teaching Science and Technology (IPST) under the Research Fund for DPST Graduate with First Placement [Grant no. 08/2559] and the Chulalongkorn University grant to the Center of Excellence for Molecular Biology and Genomics of Shrimp (GCE 6302823006-1), and to the Molecular Crop Research Unit (GRU 6407023008-1). PD is partially supported by and the 90th Anniversary of the Chulalongkorn University Scholarship. KH and TR are grateful for computational resources supported by NSTDA Supercomputer Center (ThaiSC). YS and KH are supported by the Japan Agency for Medical Research and Development (AMED) under Grant Number JP20ae0101047h0001. KH and YS are also partly supported by CREST program "Precise arrangement toward functionality", Grant No. JPMJCR20B3. The authors would like to thank Dr. Warinthorn Chavasiri for the gift of baicalein, and Dr. Supot Hannongbua for comments and criticism on the manuscript.

Author Contributions

KW conceived of the study and was in charge of the overall direction. KW and TR procured funding for this study. PD and PW purified $3CL^{Pro}$ and performed ANS binding assay and enzyme kinetics. PD and KW crystallized and determined the $3CL^{Pro}$ structure. KH, YS, and TR performed docking and computational characterization. The manuscript was written with contributions from all authors. All authors have given approval to the final version of the manuscript.

References

[1] V'Kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 2020:1–16.
[2] Hilgenfeld R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. FEBS J 2014;281(18):4085–96.
[3] Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. Nature 2020;582(7811):289–93.
[4] Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. Science (New York, NY). 2020;368(6489):409–12.
[5] Boras B, Jones RM, Anson BJ, Arenson D, Aschenbrenner L, Bakowski MA et al. 2020. Discovery of a Novel Inhibitor of Coronavirus 3CL Protease as a Clinical Candidate for the Potential Treatment of COVID-19. bioRxiv.2020.2009.2012.293498.
[6] Jukič M, Janežič D, Bren U. Ensemble docking coupled to linear interaction energy calculations for identification of coronavirus main protease (3CL(pro)) non-covalent small-molecule inhibitors. Molecules (Basel, Switzerland) 2020;25(24).
[7] Kaushal N, Gupta Y, Goyal M, Khaiboullina SF, Baranwal M, Verma SC. Mutational Frequencies of SARS-CoV-2 genome during the beginning months of the outbreak in USA. Pathogens (Basel, Switzerland) 2020;9(7).
[8] Lee J, Worrall LJ, Vuckovic M, Rosell FI, Gentile F, Ton A-T, et al. Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. Nat Commun 2020;11(1):5877.
[9] Stryer L. Fluorescence spectroscopy of proteins. Science (New York, NY). 1968;162(3853):526–33.
[10] Schonbrunn E, Eschenburg S, Luger K, Kabsch W, Amrhein N. Structural basis for the interaction of the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS) with the antibiotic target MurA. PNAS 2000;97(12):6345–9.
[11] Kane CD, Bernlohr DA. A simple assay for intracellular lipid-binding proteins using displacement of 1-anilinonaphthalene 8-sulfonic acid. Anal Biochem 1996;233(2):197–204.
[12] Pearce MC, Rubin H, Bottomley SP. Conformational change and intermediates in the unfolding of alpha 1-antichymotrypsin. J Biol Chem 2000;275(37):28513–8.
[13] Bhattacharjee C, Das KP. Thermal unfolding and refolding of beta-lactoglobulin. An intrinsic andextrinsic fluorescence study. Eur J Biochem 2000;267(13):3957–64.
[14] Chang HP, Chou CY, Chang GG. Reversible unfolding of the severe acute respiratory syndrome coronavirus main protease in guanidinium chloride. Biophys J 2007;92(4):1374–83.

[15] Su HX, Yao S, Zhao WF, Li MJ, Liu J, Shang WJ, et al. Anti-SARS-CoV-2 activities in vitro of Shuanghuanglian preparations and bioactive ingredients. Acta Pharmacol Sin 2020;41(9):1167–77.
[16] Agrawal PK, Agrawal C, Blunden G. Rutin: A potential antiviral for repurposing as a SARS-CoV-2 main protease (Mpro) inhibitor. Nat Prod Commun 2021;16 (4). 1934578X21991723.
[17] Al-Zahrani AA. Rutin as a promising inhibitor of main protease and other protein targets of COVID-19. Silico Study 2020;15(9). 1934578X20953951.
[18] Cherrak SA, Merzouk H, Mokhtari-Soulimane N. Potential bioactive glycosylated flavonoids as SARS-CoV-2 main protease inhibitors: A molecular docking and simulation studies. PLoS ONE 2020;15(10):e0240653.
[19] Hu X, Cai X, Song X, Li C, Zhao J, Luo W, et al. Possible SARS-coronavirus 2 inhibitor revealed by simulated molecular docking to viral main protease and host toll-like receptor. Future Virol 2020. https://doi.org/10.2217/fvl-2020-0099.
[20] Huynh T, Wang H, Luan B. Structure-based lead optimization of herbal medicine rutin for inhibiting SARS-CoV-2's main protease. PCCP 2020;22 (43):25335–43.
[21] Peterson L. COVID-19 and flavonoids. In silico molecular dynamics docking to the active catalytic site of SARS-CoV and SARS-CoV-2 main protease. SSRN 2020. https://doi.org/10.2139/ssrn.3599426.
[22] Xu Z, Yang L, Zhang X, Zhang Q, Yang Z, Liu Y, et al. Discovery of potential flavonoid inhibitors against COVID-19 3CL proteinase based on virtual screening. Strategy. 2020;7(247).
[23] Rahman F, Tabrez S, Ali R, Alqahtani AS, Ahmed MZ, Rub A. Molecular docking analysis of rutin reveals possible inhibition of SARS-CoV-2 vital proteins. J Tradition Complem Med 2021.
[24] Somboon T, Mahalapbutr P, Sanachai K, Maitarad P, Lee VS, Hannongbua S, et al. Computational study on peptidomimetic inhibitors against SARS-CoV-2 main protease. J Mol Liq 2021;322:114999.
[25] Xue X, Yang H, Shen W, Zhao Q, Li J, Yang K, et al. Production of authentic SARS-CoV M(pro) with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. J Mol Biol 2007;366(3):965–75.
[26] Dowd JE, Riggs DS. A comparison of estimates of michaelis-menten kinetic constants from various linear transformations. J Biol Chem 1965;240:863–9.
[27] Douangamath A, Fearon D, Gehrtz P, Krojer T, Lukacik P, Owen CD, et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. Nat Commun 2020;11(1):5047.
[28] Kabsch W. XDS. Acta Crystallogr Sect D Biol Crystallogr 2010;66(Pt 2):125–32.
[29] Evans PR, Murshudov GN. How good are my data and what is the resolution?. Acta Crystallogr D Biol Crystallogr 2013;69(Pt 7):1204–14.
[30] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40(Pt 4):658–74.
[31] Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 2010;66(Pt 2):213–21.
[32] Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr 2012;68(Pt 4):352–67.
[33] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66(Pt 4):486–501.
[34] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 2004;25(13):1605–12.
[35] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR et al. 2016. Gaussian 16 Rev. C.01. Wallingford, CT.
[36] Nutho B, Mahalapbutr P, Hengphasatporn K, Pattaranggoon NC, Simanon N, Shigeta Y, et al. Why are lopinavir and ritonavir effective against the newly emerged coronavirus 2019? Atomistic insights into the inhibitory mechanisms. Biochemistry 2020;59(18):1769–79.
[37] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31(2):455–61.
[38] Kitaura K, Sugiki S-I, Nakano T, Komeiji Y, Uebayasi M. Fragment molecular orbital method: analytical energy gradients. Chem Phys Lett 2001;336 (1):163–70.
[39] Pham BQ, Gordon MS. Development of the FMO/RI-MP2 fully analytic gradient using a hybrid-distributed/shared memory programming model. J Chem Theory Comput 2020;16(2):1039–54.
[40] Tsukamoto T, Mochizuki Y, Watanabe N, Fukuzawa K, Nakano T. Partial geometry optimization with FMO-MP2 gradient: Application to TrpCage. Chem Phys Lett 2012;535:157–62.
[41] Fedorov DG. The fragment molecular orbital method: theoretical development, implementation in GAMESS, and applications. WIREs Comput Mol Sci 2017;7 (6):e1322.
[42] Steinmann C, Jensen JH. Effective Fragment Molecular Orbital Method. Fragmentation John Wiley & Sons; 2017.
[43] Fedorov DG, Kitaura K. Pair interaction energy decomposition analysis. J Comput Chem 2007;28(1):222–37.
[44] Boonyasuppayakorn S, Saelee T, Visitchanakun P, Leelahavanichkul A, Hengphasatporn K, Shigeta Y, et al. Dibromopinocembrin and dibromopinostrobin are potential anti-dengue leads with mild animal toxicity. Molecules (Basel, Switzerland). 2020;25(18).
[45] Maruyama K, Sheng Y, Watanabe H, Fukuzawa K, Tanaka S. Application of singular value decomposition to the inter-fragment interaction energy analysis for ligand screening. Comput Theor Chem 2018;1132:23–34.
[46] Tokiwa T, Nakano S, Yamamoto Y, Ishikawa T, Ito S, Sladek V, et al. Development of an analysis toolkit, AnalysisFMO, to visualize interaction energies generated by fragment molecular orbital calculations. J Chem Inf Model 2019;59(1):25–30.