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 Hengphasatpornc, Patcharin Wilasluck a,b, Yasuteru Shigetac, Thanyada Rungrotmongkold,e, Kittikhun Wangkanonta,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 (3CLPro) is responsible for the cleavage of the viral polyprotein. This process is essential for the viral life cycle. Therefore, 3CLPro is a promising target to develop antiviral drugs for COVID-19 prevention and treatment. Traditional enzymatic assays for the identification of 3CLPro inhibitors rely on peptide-based colorimetric or fluorogenic substrates. However, the COVID-19 pandemic has limited or delayed 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 3CLPro was observed, and this interaction was competitive with a peptide substrate. The utility of ANS-based competitive binding assay to identify 3CLPro inhibitors was demonstrated with the flavonoid natural products baicalein and rutin. The molecular nature of ANS and rutin interaction with 3CLPro 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.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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 (3CLPro), 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 3CLPro, thereafter referred to as 3CLPro inhibitors, 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 3CLPro [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 3CLPro inhibitor discovery. We have been interested in the exploration of local natural products as potential 3CLPro inhibitors. Traditional assays for protease inhibitor identification employ colorimetric or fluorogenic peptide substrates. This approach has been quite successful in identification of 3CLPro inhibitors [3,4]. The peptide substrates could readily be synthesized by existing peptidase 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 in global shortages of healthcare commodities and supplies, the cost of such services will likely continue to increase.
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\textsuperscript{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\textsuperscript{pro} (TSAVLQ-SGFRK and SGVTFQ-GKFKK, with - indicating the cleavage site\cite{8}) contain a few hydrophobic amino acids and the active site of 3CL\textsuperscript{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\cite{9}. ANS-based binding assay has been utilized to investigate protein–ligand interactions and protein unfolding processes\cite{10–13}. ANS has previously been employed to characterize the folding process of SARS-CoV-1 3CL\textsuperscript{pro}, in which ANS binds to the exposed hydrophobic region when the protein is unfolded\cite{14}. However, the interaction of ANS and the native 3CL\textsuperscript{pro} has not yet been explored. If ANS could bind in the active site of 3CL\textsuperscript{pro}, an increase in ANS fluorescence should be observed. Inhibitors of 3CL\textsuperscript{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\textsuperscript{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\cite{15}. We also demonstrated the utility of the ANS-based binding assay for the identification of rutin as a 3CL\textsuperscript{pro} inhibitor, as predicted by several computational studies\cite{16–22}. The ANS-based competitive binding assay will be valuable in a cost-effective screen for novel 3CL\textsuperscript{pro} inhibitors, especially when access to a peptide substrate is limited.

2. Results

The interaction of ANS with 3CL\textsuperscript{pro} was first explored by monitoring the increased fluorescence in an ANS-3CL\textsuperscript{pro} mixture compared to the solution with no 3CL\textsuperscript{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\textsuperscript{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\textsubscript{D}) of 57 ± 7 \textmu M (Fig. 1B).

To determine whether ANS could bind in the active site of 3CL\textsuperscript{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\textsubscript{M}) for the fluorogenic substrate to be 51 ± 9 \textmu M. The addition of ANS and the non-linear fit of the resulting kinetic data to the competitive inhibition model gave the inhibitory constant (K\textsubscript{I}) of 188 ± 24 \textmu M.

To demonstrate the utility of ANS as a fluorescent probe to identify a 3CL\textsuperscript{pro} inhibitor, we performed competitive and competitive inhibition experiments with baicalein and rutin (Fig. 3A). Baicalein is a flavonoid natural product known to inhibit 3CL\textsuperscript{pro}\cite{15}. X-ray crystallography revealed that baicalein binds in the active site of 3CL\textsuperscript{pro}. Thus, the reduction of ANS fluorescence by baicalein should be evidence that ANS binds in the active site of 3CL\textsuperscript{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\textsubscript{50}) of baicalein was 42 ± 2 \textmu M, and the corresponding K\textsubscript{I} obtained from the Cheng–Prusoff equation was 15.2 ± 0.7 \textmu M. Inhibition of 3CL\textsuperscript{pro} by baicalein was also confirmed in an activity assay using the fluorogenic

![Fig. 1. A) Chemical structure and fluorescence emission spectrum of ANS (50 \textmu M) bound to 3CL\textsuperscript{pro}. B) Saturation binding curve to ANS to 3CL\textsuperscript{pro}.](image)

![Fig. 2. Lineweaver–Burk plot of an initial rate enzyme kinetic experiment of 3CL\textsuperscript{pro} at different substrate concentrations with ANS presented at 0, 250, and 500 \textmu M. Trend lines were created with parameters derived from non-linear fit of the original data.](image)
Our results suggest that the termini of the other chain in the 3CL\textsuperscript{Pro} previous in silico ride rutinose (2.7–2.9 Å). This is somewhat inconsistent with the residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide 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 residues in chain B as well as four hydrogen bonds at the disaccharide binding (hydrophobic contacts with the 13 residues in chain A and 3 residues in chain B as well as four hydro...
Fig. 4. FMO-RIMP2/PCM pair interaction energy (PIE$_{\text{total}}$) and energy components (electrostatic interaction ($E_{\text{ES}}^{IJ}$), charge transfer with higher-order mixed terms energies ($E_{\text{CT+mix}}^{IJ}$), dispersion ($E_{\text{DI}}^{IJ}$), exchange-repulsion ($E_{\text{EX}}^{IJ}$)), and the PCM solvation effect ($G_{\text{PCM}}^{\text{Sol}}$) for rutin interacting with individual residues of dimeric SARS-CoV-2 3CLPro. Rutin orientation and interactions at the active site of SARS-CoV-2 3CLPro 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 3CLPro.

| 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 | Glide | N142, G143, C145, T190 | [22] |
| monomer | 6LU7 | LeDock | F140, E166, T26, L141, S144, C145, H163 | [19] |
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 3CLPro 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 $K_i$ value obtained from the saturation binding curve of ANS to 3CLPro. 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 3CLPro. It is possible that 3CLPro possesses a potential source of affinity discrepancy could be due to multiple binding sites. Indeed, ANS likely binds in the active site of 3CLPro and is also recognized by the $\gamma$-lactam moiety of the N3 inhibitor. Rutin interacts mainly with the S1 and S2 pockets, and also interact with residues from the neighboring 3CLPro subunit.

To investigate the molecular details of the interaction between ANS and rutin with 3CLPro, we attempted to crystallize this compound with 3CLPro. However, only the DMSO-bound structure was obtained. Therefore, we employed computational methods to investigate the 3CLPro-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 $\gamma$-lactam moiety of the N3 inhibitor. Rutin interacts with the S1 and S2 pockets, and also interact with residues from the neighboring 3CLPro subunit.

Compared with the crystal structure of DMSO-bound 3CLPro, 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 3CLPro 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 3CLPro in the active site. Using baicalein as a model compound with a known binding location and 3CLPro inhibitory activity, we showed that ANS binding was competitive with baicalein. Thus, ANS likely binds in the active site of 3CLPro. We also applied this methodology to demonstrated that rutin, a compound predicted to bind and inhibit 3CLPro, was indeed binding competitively with ANS and could inhibit 3CLPro activity. These results demonstrated the utility of ANS in the identification of potential 3CLPro 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 3CLPro with no tags at the termini was expressed, purified, and stored exactly as previously described for SARS-CoV-1 3CLPro[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. 3CLPro 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 $\mu$L. ANS binding assay was

---

**Table 2** Crystallographic data collection and refinement statistics.

| Data collection statistics |  |
|---------------------------|--|
| Wavelength (Å)            | 0.97872 |
| Resolution range (Å)      | 27.57–1.45 (1.47–1.45) |
| Space group               | C 2    |
| Unit cell dimensions      | 113.2, 53.0, 44.7 |
| Total number of reflections | 339,475 (16,721) |
| Number of unique reflections | 43,879 (2,124) |
| Multiplicity              | 7.7 (7.9) |
| Completeness (%)          | 96.1 (91.8) |
| Mean I/σ(I)               | 15.3 (2.2) |
| Wilson B factor (Å$^2$)   | 12.21  |
| Rmerge                    | 0.077 (1.035) |
| Residuals                 | 0.082 (1.106) |
| Rfree (5%)                | 0.030 (0.394) |
| CC1/2                     | 0.999 (0.803) |

**Refinement Statistics**

| Resolution range (Å$^2$) | 27.57–1.45 (1.50–1.45) |
|-------------------------|--|---|
| R-factor                | 0.1664 (0.2567) |
| R-free (% C)            | 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 |

* Statistics for the highest-resolution shell are given in parentheses.
performed with 3CLPro 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 mM of 3CLPro. 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

3CLPro 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. Microseedling 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 3CLPro 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

![Fig. 5. A) Crystal structure of 3CLPro crystallized in the presence of 5 mM ANS but has DMSO (green) bound instead of ANS. B) Crystal structure of 3CLPro with baicalein bound (PDB ID 6M2N). C-D) Conformations of ANS and rutin binding to 3CLPro 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.)](image)

![Fig. 6. Position of ANS and rutin in the active site compared to the peptidomimetic N3 inhibitor.](image)
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-B3LYP/6–31 g(d) basis set using gaussian 16 [35]. The 3CLPro protein covalently bonded with the N3 inhibitor was retrieved from the protein databank (PDB ID 6LU7) [3]. From our previous studies [24,36], to construct the initial structure of the ligand–3CLPro 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 3CLPro, the pair interaction energy (PIE) calculation using the second-order Møller–Plesset perturbation theory (MP2) with the resolution-of-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:

$$\text{PIE} = \Delta E_{g}^{\text{IE}} + \Delta E_{g}^{\text{IE}-\text{mix}} + \Delta E_{q}^{\text{IE}} + \Delta E_{q}^{\text{IE}-\text{mix}}$$

(1)

where $\Delta E_{g}^{\text{IE}}$ is electrostatic interaction, $\Delta E_{g}^{\text{IE}-\text{mix}}$ is charge transfer with higher-order mixed terms energies $\Delta E_{q}^{\text{IE}}$, dispersion, $\Delta E_{g}^{\text{IE}-\text{mix}}$ is exchange-repulsion, and $G_{\text{PCM}}$ is the polarizable continuum model (PCM) solvation effect [44–46].

CRediT authorship contribution statement

Peerapon Deetanya: Methodology, Investigation, Writing - review & editing. Kantirat Hengphasatporn: Methodology, Investigation, Writing - original draft, Writing - review & editing, Visualization. Patcharin Wilasluck: Methodology, Investigation, Validation. Yasutera Shigeta: Methodology, Investigation, Resources, Funding acquisition. Thanaya Rungrutmongkol: Methodology, Investigation, Writing - review & editing, Resources, Funding acquisition. Kittikun 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 6407203008-1). 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. Warin-thorn 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 3CLPro and performed ANS binding assay and enzyme kinetics, PD and KW crytalized and determined the 3CLPro 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] Jing Z, Li X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M(3)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, Sauerhaling 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] Bocas R, Jones RM, Anson BJ, Arenson D, Aschenbrenner L, Bakowska 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 2020.2020.12.29.49489.
[6] Jukić M, Janičić 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] Schönbrunn E, Eschenburg S, Luger K, Kabsch W, Amrein N. Structural basis for the interaction of the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS) with the antibiotic target MraY. PNAS 2000;97(12):6354–9.
[11] Kane CD, Bernioli DA. A simple assay for intracellular lipid-binding proteins using displacement of 1-anilinonaphthalene 8-sulfonic acid. Anal Chem 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] Bhattacharyee C, Das KP. Thermal unfolding and refolding of beta-lactoglobulin. An intrinsic and extrastrandary 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 guandimine chloride. Biophys J 2007;92(4):1374–83.
