Production of recombinant SARS-CoV-2 3CL-protease: The key for the development of protease inhibitors screening kit in search of potential herb cure for COVID-19

Haniyya1, M Ulfah1, A Riswoko1, L Mulyawati1, T Ernawati2 and I Helianti1*
1Center for Bioindustrial Technology, National Agency of Research and Innovation (BRIN), Laboratorium of Bioindustrial Technology, 611/614 Bldg. LAPTIAB BPPT, Puspiptek - Serpong, Tangerang Selatan 15314, Indonesia
2Research Center for Chemistry, National Agency of Research and Innovation (BRIN), 452 Bldg, Jalan puspitek serpong Gate, Muncul, Kec. Serpong, Kota Tangerang Selatan, Banten 15314, Indonesia
*Corresponding author e-mail: is.helianti@bppt.go.id

Abstract. Coronavirus disease-19 (COVID-19) pandemic caused millions of deaths and socio-economics damage worldwide. Corresponding to this, many studies on antiviral drugs exploration are rising to investigate the potential of drug compounds that can repress the replication of the SARS-CoV-2 virus by targeting its main protease named 3-chymotrypsin like protease (3CLpro). Without 3CLpro splicing, polyproteins of SARS-CoV-2 will not function to form new virions. We conducted an in silico thermodynamic study, and found several already known compounds from natural sources such as lovastatin, quinidine, and quinine were potential in inhibiting 3CLpro. However, most of the findings still need further wet-lab experiments to confirm their activity against it. To facilitate rapid screening of protease inhibitors, we aim to develop a screening kit for researchers with focus studies on herbal sources for COVID-19 treatments. Hence, the development of 3CLpro as the target of inhibition screening is essential. Our current work focused on the expression of the SARS-CoV-2 3CL Pro gene harbored by pET-32b(+) in E. coli BL21(DE3) and its purification. The recombinant E.coli were cultivated in LB media at 37 °C with 5 hours IPTG-induced phase followed by crude protein lysate extraction by sonication, and then Ni-NTA purification. A band of recombinant 3CLpro protein (33.8 kDa) detected by SDS PAGE and Western blotting after Ni-NTA purification was specific and clear, showing the right size of our protein interest.
A final concentrated protein mixture obtained by 10-kDa membrane filtration displayed >90% purity with a total protein of 2.497 mg mL⁻¹. Further functional assays and inhibitory tests with several natural compounds showed that 3CLpro functioned well as a drug target.

Keywords: recombinant enzyme, inhibitor agent, pandemic disease, natural sources.

1. Introduction
The Covid-19 pandemic is a global health problem and caused huge damage on the social and economic aspects of most nations in the world. The cause of this disease, the Sars-coV-2 virus, has infected more than 235 million people, with the death of more than 4.8 million people in the world. In Indonesia, this virus has infected at least 4.2 million people which caused 142,000 people to die [1]. To date, the specific effective antiviral therapy for this Covid-19 has not yet been established and is still in the testing stage of clinical trials that require further study [2,3]. On the other aspect, Indonesia...
has rich biodiversity, especially herbal plants, which known as a resource for traditional medicine for many generations. These herbal resource is expected to be potential for the Covid-19 drug. Thus, screening the new antiviral drugs from the Indonesian herbal resources is very important, especially for the current pandemic.

The whole sequences of a single-stranded RNA SARS-CoV-2 have been identified and deposited. High throughput sequencing generated complete genome sequences that reported nearly 30,000 bp in length [4] and encoded 9860 amino acids [5]. RNA SARS-CoV-2 virus genome contains structural and non-structural proteins. The nucleocapsid (N), membrane (M) envelope (E), and Spike (S) was encoded the structural proteins, whereas non-structural encoded by two ORF regions located at the N-terminus, such as 3-chymotrypsin like protease (3CLpro), papain-like proteases (PLpro), and RNA-dependent RNA polymerase (RdRp) [6-8].

The non-structural 3-chymotrypsin like protease (3CLpro) of SARS-CoV-2 is known as essential in virus replication [8] and plays a role in processing polyprotein to mature protein [10]. The viral life cycle begins with proteolysis processing initiated by 3CLpro together with PLpro. The proteases cut the polyprotein into the slice fragments for folding and packaging new viruses, thereby increasing the spread of infection. Together with RdRp, these major enzymes 3CLpro and PLpro are essential for the proteolysis and replication of new viruses [11-12]. Since its urgent role, 3CLpro could be a potential drug target.

From the bioinformatic analysis of whole sequences of SARS-CoV-2 from Indonesia conducted until September 2020, the D614G mutation was the dominant one in 59% [13]. But nowadays, from the current data, the D614G is no longer a mutation but a reference. The development of this virus is very dynamic research on drugs to overcome this virus outbreak continues. Based on the pivotal role of 3CLpro as a drug target, in silico experimental inhibition of this enzyme by herbal compound has been conducted parallelly.

To facilitate rapid screening of protease inhibitors from herbal resources, we aim to develop a screening kit for researchers focused on herbal sources for COVID-19 treatments. Hence, the development of 3CLpro as the target of inhibition screening is essential. In this study, the recombinant plasmid harboring 3CLpro SARS-CoV-2 gene was constructed via synthetic DNA, and the transformation and verification of this recombinant plasmid were conducted and followed by expression in E. coli BL21 (DE3) host, purification, and functional analyses of the protease by Western Blot and activity assay. The recombinant 3CL protease will be used further to screen potential herbal bioactive compounds.

2. Materials and Methods

2.1 Bacterial samples and plasmids
A total of two types of E. coli, namely DH5α (for recombinant plasmid propagation) and BL21 or DE3 (for host expression), were used in this study. The synthetic DNA that encoding for the 3CLpro SARS-CoV-2 gene was ordered from GenScript USA Inc. and inserted between pET-32b(+) harboring 3CLpro SARS-CoV-2 between NdeI and XhoI sites. All strains were routinely grown on Luria Bertani (LB) medium (w/v) (1% Tryptone: 0.5% Yeast extract; 1% NaCl; 1.5% agar) containing ampicillin (100 µg mL⁻¹). Agar plates were incubated at 37 °C for 16 - 18 h and then stored at 4°C. For the long-term storage, 5.0 mL of an overnight culture of each recombinant strain was harvested by centrifugation and replaced with 1 mL of fresh LB, then mixed with 0.25 mL of 75% glycerol stock and stored at -80 °C.

2.2 Preparation of competent cell and transformation
Luria Bertani broth was used for the cultivation medium. TB (Transformation Buffer) containing: 10 mM PIPES, 15 mM CaCl₂, 2H₂O, and 250 mM KCl were mixed, pH adjusted to 6.7. Then 55 mM MnCl₂•2H₂O was dissolved into a mixture, then sterilized by 0.22 µL syringe filtration and stored at 4°C. These media and buffers were used to make E.coli competent cells. The preparation of competent cells was conducted by following standard procedures [14]. A single colony of each E. coli DH5α and
E. coli BL21 (DE3) was grown in LB broth at 37°C and shaken overnight at 250 rpm. As much as 1% of overnight culture was inoculated to LB 200 mL then incubated at 37°C to OD$_{600}$ 0.4-0.6. After reaching the desired OD, the culture was cooled on ice for 30 min. The culture was then harvested by centrifugation using HITACHI high-speed refrigerated centrifuge at 3000 rpm, 4 °C for 15 min. The pellet was washed by 16.75 mL cold TB and incubated on ice for 10 min. In the next step, pellet suspension was centrifuged at 3000 rpm, 4 °C for 10 min, then collected pellet was dissolved in 2 mL cold TB added by 0.3 ml dimethylsulfoxide (DMSO). The suspension was incubated on ice for 10 min, then aliquoted into microtube of each 50 μL, storage at -80 °C. These were E. coli competent cells.

For plasmid propagation and expression, 1 – 10 ng/μL of recombinant plasmid was added into thawed competent cell suspension by the heat-shock method. The mixture was incubated for 30 min on ice, then heated at 42°C for 1 min and 2 min reincubated on ice, then quickly added 0.5 mL of LB medium. The mixture was incubated in a shaker incubator at 37°C, 250 rpm for an hour. The cells were collected by centrifugation at 6000 rpm for 5 min at 4°C. As much as 100 - 200 μL mixture was resuspended by pipetting and plating in LB agar containing 100 μg/mL ampicillin. Plate incubated at 37°C for 16-18 hours.

A single colony of E. coli DH5α harboring recombinant plasmid pET-32b(+) harboring 3CLpro SARS-CoV-2 was cultivated in LB broth at 37°C, for 16-18 h. Culture then harvesting and the recombinant plasmid was extracted under Zippy Plasmid Kit protocol (Zymo Research). The extracted plasmid was confirmed by restriction enzymes Ndel and XhoI.

2.3 Protein expression
The expression of the gene target was conducted following the pET System Manual. The E. coli BL21 (DE3) harboring recombinant plasmid pET-32b(+) harboring 3CLpro SARS-CoV-2 gene was cultured in LB broth containing 100 ampicillin μg/ml. Shake flask fermentation was carried out in a 500 ml Erlenmeyer flask containing 100 ml culture. Medium inoculated with 2.5% (v/v) of overnight seed culture. The culture was shaken at 37°C for 250 rpm to OD$_{600}$ reached 0.4-0.6. The culture was induced by the addition of 1 mM Isopropyl β- d-l-thiogalactopyranoside (IPTG). Protein expression was conducted at a low temperature of 25°C, 200 rpm for 5 h. After that, the cell suspension was harvested by centrifugation at 8,000 rpm at 4°C for 15 min. The supernatant was stored at 4°C, while the pellet was kept at -20°C for further analysis.

2.4 Extraction of crude extract of recombinant 3CL protease
The sample was extracted based on the Probond Native method (Invitrogen). Cells pellet from 50 ml culture were dissolved in 8 ml of native binding buffer (0.05 M Na$_2$PO$_4$, 0.5 M, NaCl, 0.01 M imidazole, pH 8.0) containing 8 mg of lysozyme. The suspension was sonicated for 5 min (2 seconds on, 3 seconds off, 27 - 30% amplitude) on ice, then centrifuged at 3000 rpm, 4°C for 15 min, and the clear supernatant was collected as lysate. The lysate was kept for SDS-PAGE analysis and further purification steps.

2.5. SDS PAGE
Recombinant proteins were analysed by SDS-PAGE according to Bollag et al. [15]. The slab gel system consisted of 12.5% separating gel (2.4 mL ddH2O, 1.5 mL 1.5 M Tris-HCl buffer pH 8.8, 60 μL 10% SDS, 3 mL 30% polyacrylamide, 30 μL 10% ammonium persulfate (APS), 7 μL TEMED) and 5% stacking gel (1.5 mL ddH2O, 625 μL 0.5 M Tris-HCl buffer pH 6.8, 25 μL 10% SDS, 334 μL 30% polyacrylamide, 15 μL 10% ammonium persulfate (APS), 7 μL TEMED).

The cells pellet was resuspended in 5X protein sample buffer (0.25 M Tris-HCl pH 6.8 0.5 M DTT 10 % SDS 50 % Glycerol 0.5 % bromphenol blue) then boiled at 100 °C for 5 min then loaded into the gel. The Spectra™ Multicolor Broad Range Protein Ladder was used as the protein marker. Electrophoresis was run in 1x running buffer (1.92 M glycine, 1% (w/v) SDS, and 0.25 M Tris-HCl) at
room temperature for 90 min at constant voltage 120 v. After electrophoresis, the gel was stained following Xpress blue solution (Xpress Blue™ Protein Stain – Himedia) protocol.

2.6. Purification of recombinant 3CLpro SARS-CoV
Purification was performed by using ProBond Purification System [Invitrogen]. Purification of native protein was conducted using Native Binding Buffer (0.05 M NaH2PO4, 0.5 M NaCl, 0.01 M imidazole, pH 8.0), Native Wash Buffer (0.05 M NaH2PO4, 0.5 M NaCl, 0.02 M imidazole, pH 8.0), and Native Elution Buffer (0.05 M NaH2PO4, 0.5 M NaCl, 0.25 M imidazole, pH 8.0).

2.7. Concentration of recombinant 3CLprotease
Recombinant 3CLprotease was concentrated by using Amicon Ultra-15 10 kDa Centrifugal Filter Devices [Merck]. The column was washed twice using 2 mL milli-Q water and then centrifuged at 3800 rpm, 20 min, 4°C. After that, a total of 9 mL of protein samples were added into the column and centrifuged at 3800 rpm, 30 min 4°C. The 1 mL of retentate that remained was collected in a tube.

2.8. Western blot analysis
After separating by SDS PAGE, proteins were transferred to a PVDF sheet membrane, using a semidry tank system, blotting apparatus (Power Blotter System-Thermo Scientific). A sandwich of filter paper, SDS PAGE gel, and membrane PVDF were soaked in transfer buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, 20% methanol). The PVDF membrane was immersed in methanol for 15 minutes before use. Furthermore, the transfer process to the PVDF membrane was carried out on 2.5 mA, 21 V for 30 minutes. Afterward, PVDF membrane with transferred protein was incubated in blocking buffer (5% skim milk blotting grade-blocker dissolved in phosphate buffer saline (PBS) containing 0.05% Tween 20) for 1 hour at 37°C with gentle shaking. After blocking, the membrane was washed in PBS-T (PBS-Tween 20) buffer with shaking and the buffer was replaced by a new one every 10 min for 1 h. The following step blocking, the membrane was probed with a primary antibody (mouse 6x His Tag Monoclonal Antibody HIS H8, Invitrogen) that dissolved in 5% skim milk in PBS-T with a 1:3.000 ratio. The membrane was incubated at 4°C for 16 - 18 h, thereafter continued at 37°C for 1 hour with gentle shaking. After that membrane was washed with fresh PBS-T buffer every 10 min for 1 h.

The next probing step was secondary antibody (goat anti-mouse IgH (H+L) horseradish peroxide conjugate, Invitrogen) was dissolved in 5% skim milk in PBS-T buffer with a 1:100,000 ratio. The membrane was incubated at 37°C for 1 h with gentle shaking and washed with fresh PBS-T buffer every 10 min for 1 h. The next step was to visualize the probing results by treatment a PVDF membrane to the chemiluminescent substrate (Super Signal™ West Femto, Thermo Scientific). The incubation process was carried out in a dark room for 5 min. To visualize the protein target, both PVDF membranes treated with Femto substrate were incubated with film (high-performance chemiluminescence film, Amersham Hyperfilm™ ECL) for 1 - 15 min. Visualization was conducted by staining film with developer and fixer solution (Carestream Kodak autoradiography GBX fixer and developer), staining of the developer, water, fixer, and water again, each from 30 seconds to 1 min.

2.9. Molecular docking of some bioactive compound
Some bioactive compounds such as Methyl trans-cinnamate, Cinnamic acid, Piperine, Quinine, Quinidine, Lovastatin, Butyro lactone, Citronellol, Isoeugenol acetate, etc., were screened for 3CL protease inhibitor by using computational prediction program Autodock.

2.10. Verification of function of purified recombinant 3CL protease
For the assay of recombinant 3CL protease activity, the mixture of 50 µL of the substrate (2.5, 5.0, 20, 80, 160 µM) and 50 µL of 50 nM enzymes was incubated at 37°C for 50 minutes (with 10 minutes pre-heated), and the fluorescence was observed. Inhibition assay was conducted by mixing enzyme reaction with Quinidine and Lovastatin, respectively. The fluorescence intensity was measured every 5 minutes at Ex = 485 nm and Em = 535 nm.
3. Results and Discussion

3.1 Re-transformation plasmid into E. coli DH5 and BL21 (DE3) and plasmid verification

The synthetic recombinant DNA of SARS-CoV-2 3CLpro gene in expression plasmid pET32b(+) in frame with 6xHis tag sequence at C-terminal was constructed (Figure 1). For verification, whether the recombinant plasmid was constructed correctly or not, digestion by restriction endonucleases was applied. The digestion of pET32b(+) – 3CLpro SARS-CoV-2 with Ndel/XhoI resulted from two fragments of 945 bp and 5.4 kb (Figure 2).

![Figure 1. The map of recombinant pET32b(+) – 3CLpro SARS-CoV-2 with Ndel/XhoI.](image1)

![Figure 2. Restriction pattern of pET32b(+) – 3CLpro SARS-CoV-2 digested by NdeI and XhoI (A) Control undigested plasmid; (B) Plasmid digested by NdeI and XhoI.](image2)

3.2 SDS PAGE Analysis

The expression 3CLpro under T7promoter was visualized after being induced by 1 mM IPTG. A soluble protein with the intensity of protein band located between 25 and 35 kDa was increased compared to an un-induced fraction (Figure 3). It has corresponded with the theoretical molecular weight of 3CLpro at approximately 33.8 kDa [16].
3.3. Purification of recombinant 3CLpro SARS-CoV
Purification of soluble recombinant 3CLpro was successfully carried out. An almost single band of protein at 33 kDa was detected in elution fraction. These protein bands were confirmed as 3CLpro by Western Blot Assay.

3.4. Molecular docking of some bioactive compound
Some bioactive compounds such as Methyl trans-cinnamate, Cinnamic acid, Piperine, Quinine, Quinidine, Lovastatin, Butyrolactone, Citronellol, Isoleugenol acetate, etc., were screened for 3CL protease inhibitor by using computational prediction program AutoDock. The compounds with ΔG less than -8.0 Kcal/mol were selected, namely Quinine, Quinidine, Cinchonine, and Lovastatin (Table 1).

3.5 Verification of 3CL protease’s function.
The concentrated recombinant 3CL Pro (2.497 mg/mL) was used for verification of recombinant 3CLpro function in protease inhibition assay. Our recombinant 3CL pro functioned well as an active protease by showing its activity against LGSAVLQ-Rhodamin 110 substrate (Figure 5A). The presence of Lovastatin (20 mM) and Quinidine (100 mM) inhibited the protease activity (Figure 5B & C).

The inhibition of 3CL protease activity in vitro by Lovastatin is according to a study reported that statin (including Lovastatin) has potential therapy effect, poses no additional risk to individuals exposed to SARSCoV-2 and that some statins may have a mild beneficial effect on COVID-19 outcome [17]. Derivatives of quinolone, such as quinidine, quinine and hydroxychloroquine (HCQ, hydroxylated form of aminoquinone), have been used for decades to treat different diseases, including malaria, rheumatological diseases, cardiac arrhythmias and, most recently COVID-19 [18]. Thus, it is
supported our results. The confirmation of purified 3 CL Pro by these compounds suggested that the produced protease could be used as a drug target for Covid19 drug screening from herbal resources.

Table 1. Molecular docking of some active compounds with the lowest AG.

| No | Compounds | Chemical Structure | AG (kcal/mol) | Kd (uM) | Interaction |
|----|-----------|--------------------|---------------|---------|-------------|
| 5  | Quinine   | ![Quinine chemical structure](image1) | -8.24 | 0.916 | ![Quinine interaction](image2) |
| 6  | Quinidine | ![Quinidine chemical structure](image3)  | -8.2 | 0.979 | ![Quinidine interaction](image4) |
| 7  | Cinchonine| ![Cinchonine chemical structure](image5)   | -8.04 | 1.29 | ![Cinchonine interaction](image6) |
| 20 | Lovastatin| ![Lovastatin chemical structure](image7) | -8.98 | 0.260 | ![Lovastatin interaction](image8) |

Figure 5. The activity of 3CL protease (A) and the inhibition of protease activity by Lovastatin (20 mM) (B) and Quinidine (C) (100 mM) expressed by fluorescence.
4. Conclusion

Induction of the E.coli BL21 (DE3) cells by IPTG resulted from increased intensity of protein band of 33.8 kDa as detected by SDS-PAGE and Western Blot Analyses. After Ni-NTA purification a single protein of the same size was detected in elution fraction. A final concentrated 3CL protease protein obtained by 10-kDa membrane filtration displayed >90% purity with a total protein of 2.497 mg/mL. Further functional assays and inhibitory tests resulted that our 3CLpro showing activity against Rhodamine peptide as substrate. The inhibition of 3CL protease activity by Lovastatin and Quinidine was confirmed, which was according to the molecular docking approach data. These results showed that our recombinant 3CL Pro was expressed as a soluble fraction and successfully purified into almost homogeneity. Furthermore, the recombinant 3CL protease was functioned well and gave protease activity against peptide Rhodamin and could be used as a drug target for Covid19 drug screening from herbal resources.

Acknowledgments
This work was funded by Research Consortium and Innovation Covid-19. Research Incentives Program of the Ministry of Research and Technology - Indonesia Endowment Fund for Education (LPDP), Ministry of Finance 2020 - 2021 granted to IH.

References
[1] https://www.worldometers.info/coronavirus/
[2] Hoffmann M, et al. 2020. Cell 181(2) 271-280
[3] Utomo RY, et al. 2020. Preprints. DOI: 10.20944/preprints202002.0431.v1
[4] Kumar S, Nyodu, R, Maurya V K and Saxena S K 2020 Morphology, Genome Organization, Replication, and Pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-Coronavirus Disease 2019 (COVID-19) 23–31
[5] Chen L, Liu W, Zhang Q, Xu K, Ye G, Wu W, et al. 2020. Emerg. Microbes. Infect. 9 313–319
[6] Huang Y, Yang C, Xu X 2020. Acta Pharmacol. Sin. 41 1141–1149
[7] Naqvi AAT, Fatima K, Mohammad T, Fatima U, Singh A, Atif SM, Hariprasad G, Hasan GM, Hassan MI 2020. Biochim. Biophys. Acta Mol. Basis Dis. 1 1866
[8] Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S 2020. Emerg. Microbes. Infect. 9 221-236
[9] Gao X, Qin B, Chen Pu, Zhu K, Hou P, Wojiyla JA, Wang M, Cui S 2021 Crystal structure of SARS-CoV-2 papain-like protease Acta Pharm Sin. B 11 237-245
[10] Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y 2020. Acta. Pharm. Sin. B 10 766-788
[11] Shang J, Ye G, Shi K, Yan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F 2020. Nature 581 221–224
[12] Tong TR 2009. Infect. Disord. Drug Targets 9(2) 223-45
[13] Ulfa H, Helianti I 2021. Iran J. Microbiol 13(2) 145-155
[14] Inoue H, Nojima H and Okayama H 1990. Gene 96 23–28
[15] Bollag DM, Rozyczki MD and Edelstein SJ 1997 Protein methods (2nd ed.). New York, USA: John Wiley-Liss, Inc.
[16] https://web.expasy.org/compute_pi/
[17] Moeller, Zapatero-Belinchón, Lasswitz et al. 2020 Effect of statins on SARS-CoV-2 infection. preprint version –www.biorxiv.org/
[18] Yetkin E., Yalta K, Waltenberger J. 2020. Neth Heart J 28 437–438