Investigation of the potential covid-19 inhibitor from *Cassia alata* and *Dendrophthoe petandra* using computational approach

Teni Ernawati*, Marissa Angelina, Rizna Triana Dewi, Sofa Fajriah, Rokip

Research Center for Chemistry - Indonesian Institute of Sciences (LIPI), Kawasan Puspiptek, 452 Building, Tangerang Selatan, Banten 15314, Indonesia

*Corresponding author’s email: teni.ernawati001@gmail.com

Abstract. COVID-19 is a recent coronavirus outbreak caused by SARS-CoV-2 beginning in early December 2019 in Wuhan China raising global health problems. The SARS coronavirus main peptidase (SARS-CoV M(pro)) plays an essential role in the life-cycle of the virus and is a primary target for the development of anti-SARS agents. Also, the viral 3-chymotrypsin-like cysteine protease 3 (3CLpro) enzyme controls coronavirus replication and is essential for its life cycle. Therefore, herein, we performed in silico virtual screening of the compounds from *Cassia alata* and *Dendrophthoe petandra* such as; aloe-emodin, emodin, kaempferol, quercetin and quercitrin against SARS-CoV-2 M<sup>pro</sup> and also 3CL<sup>pro</sup> through molecular docking approach using Autodock 4.2, with the Lamarckian Genetic Algorithm. The binding energies obtained from the result of molecular docking of SARS-CoV-2 M<sup>pro</sup> and also 3CL<sup>pro</sup> quercetin, kaempferol and aloe-emodin appeared to have the best potential to act as COVID-19 inhibitors.

1. Introduction
In the past ten years, many new coronaviruses have been identified. They infect large numbers of hosts from mammals to birds and closely related coronaviruses have been identified [1-5]. Corona viruses (CoV) belong to a group of viruses that can infect humans and vertebrate animals. CoV infection affects the respiratory, digestive, liver, and central nervous systems of humans and animals [6-8]. This research focuses on proteases in CoV (Mpro/3CLpro), especially PDB ID 6LU7,2GTB, 3M3V and also through Homology modeling as a potential target protein for COVID-19 treatment [9-12]. Protein is taken from a protein data bank that can be accessed by the public since early February 2020 and there are also those who use homology modeling using Swiss-Model [13-17]. The use of proteins that have similarities with SARS-Cov is used to tether the active compounds contained in *Cassia alata* and *Dendrophthoe petandra* plants namely aloe-emodin, emodin, kaempferol, quercetin and quercitrin. Molecular docking was carried out in this study to predict the activity of the five compounds in the SARS-CoV active site protein. In addition, bioactivities of the five compounds and commercial antiviral compounds were compared to see the extent to which the antiviral activity of the compounds contained in *C. alata* and *D. petandra*. SARS-CoV protein is very important for viral proteolytic maturation and has been examined as a potential target protein to prevent the spread of infection by inhibiting the division of the polyprotein virus [13,18,19]. This molecular docking study provides a great opportunity to identify potential drug candidates for antiviral treatment.
2. Materials and Methods

The 3D structures of the tested compounds were built by Chem Draw 12.0. The crystal structures of main protease covid-19 (PDB ID: 6LU7, 2GTB, 3M3V, and homology docking) were used as enzyme models and processed by AutoDock 4.2. The hydrogen bond network was refined by means of the optimize option. The grid boxes were centered on x, y, and z coordinates -11.683, 14.686, 65.363 (6LU7); 18.216, -12.873, 11.035 (2GTB); -27.887, -20.584, 25.196 (3M3V) and also for homology modelling 14.313, 14.502, 26.864. The homology modeling template using SWISS-MODEL, application of molecular visualization using UCSF-Chimera 1.9, and Autodock 4.2 RAMPAGE for testing Ramachandran plot. In this paper, the active compound from C. alata and D. petandra such as aloemodin, emodin, kaempferol, quercetin and quercitrin were selected as ligand to dock into the main protease covid-19 (PDB ID: 6LU7, 2GTB, 3M3V, and homology docking). And also, the commercial antiviral medicines were used as ligand to dock into the same main protease covid-19 for comparative.

3. Results and Discussion

The five compounds from C. alata and D. petandra as well as five commercial antiviral compounds were subjected to docking studies using AutoDock 4.2. The proteins used in this docking simulation are the 2 main SAR-CoV-2 proteases obtained from protein bank data and also the 3CL\textsuperscript{pro} homology modeling using the Swiss Model. These four SAR-CoV-2 proteins are used to calculate the binding affinity and free energy of the SAR-CoV-2 receptor binding. Table 1 was shown the structures and amino acids found in the active site pockets of 2 main Protease M\textsuperscript{pro} (6LU7, 2GTB) and 3CL\textsuperscript{pro} (3M3V and its homology model). 6LU7 is the main protease (M\textsuperscript{pro}) found in COVID-19, which been structured and repositioned in PDB and can be accessed by the public, as of early February 2020. 2GTB is the main protease found in the CoV associated with the severe acute respiratory syndrome (SARS), which can be accessed in PDB and was suggested to be a potential drug target for 2019-nCov. Xu et al mentioned that the main protease in 2019-nCov shares 96% similarity with that in SARS [1]. Ligands and several drug candidates compounds have been previously selected, based on adherence to Lipinski’s rule of five. The selected ligands that did not incur more than 2 violations of Lipinski’s rule could be used in molecular docking experiments with the target protein.

The results of molecular docking of chemical substituents from C. alata and D. petandra on several SAR-CoV-2 proteins are shown in Table 1. The less free energy generated from the results of the inhibition of compounds with the protein SAR-CoV-2, the more active the prediction of its compounds according to the docking simulation theory. Likewise, the smaller the value of the inhibition, the faster the kinetics of the required inhibition of the compound against the SAR-CoV-2 protein. Figure 1 showed the docking score data for all active compounds in the sample and commercial antiviral compounds showed that among the 10 simulated chemical compound structures; nelfinavir as a commercial antiviral drug had the lowest docking score while for the sample compound (almost all the SAR-CoV proteins used, except 3M3V); quercetin which is an active compound contained in D. petandra had the lowest docking score (for SAR-CoV-2 6LU7 and 3M3V); kaempferol as active compound from C. alata had the lowest docking score (SAR-CoV-2 2GTB and 3CL\textsuperscript{pro} model Homology).

The binding of hydrogen with Cys-145 and His-163 plays an important role in the inhibition of the SAR-CoV-2 protein. The binding of hydrogen with Cys-145 and His-163 plays an important role in the inhibition of SAR-CoV because cysteine and histidine residues are the main constituents of the SAR-CoV-2 receptor. In addition to the main residues that play an important role, there are also related amino acid residues that are no less important in producing small inhibitory energy so that the SAR-CoV-2 molecules are inhibited, among others; Pro168 Glu166 Leu167 Met165 Gln189 Gln192 Met49 Cys44 Asp187 Arg188. From the results of docking simulations of 10 chemical structures, it can be concluded that those that have consistent active predictions on all the SAR-CoV-2 proteins tested are: quercetin, kaempferol and aloemodin for chemical substituents from C. alata and D. petandra compounds as well as nelfinavir, hydroxychloroquine and remdesivir for commercial antiviral compounds.
Table 1. Result of docking score of chemical substituents from *C. alata* and *D. petandra*.

| No | Protein (PDB ID) | Ligand     | ΔG(kkal/mol) | Ki (uM) | Amino acid residue interactions                                                                                                                                 |
|----|------------------|------------|--------------|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Mpro 6LU7        | Aloeemodin | -7.84        | 1.78   | Val3 Leu4 His163 Phe140 Leu141 Asn142 Ser144 Cys145 Leu27 Gly143 Thr26                                                                                       |
|    |                  | Emodin     | -7.30        | 4.47   | His172 Glu166 Val3 Leu4 His163 Phe140 Leu141 Asn142 Gly143 Cys145 Leu27                                                                                      |
|    |                  | Kaempferol | -7.26        | 4.80   | Glu166 Val3 Leu4 Phe140 His163 Leu141 Asn142 Gly143 Ser144 Cys145 Thr26                                                                                     |
|    |                  | Quercetin  | -8.18        | 1.01   | Thr26 Gly143 Leu27 Cys145 Ser144 Leu141 Asn142 Phe140 His163 Val3 Glu166                                                                                     |
|    |                  | Quercitin  | -5.49        | 93.8   | Thr26 Gly143 Leu27 Ser144 Leu141 Leu4 Phe140 His163 Val3 Glu169                                                                                               |
|    |                  | Hidroxychloroquinin | -8.23     | 0.930  | Leu4 Met49 Thr45 Ser144 Cys145 Leu27 Gly143                                                                                                                   |
|    |                  | Lopinavir  | -5.56        | 0.83   | Thr25 Asn142 Leu141 Glu189 Met165 Leu4 Met49 Val3 Glu166 Ser46 Thr45 His163 Asn142 Cys145 Leu27 Gly143 Thr26 Thr25 Thr24 |
|    |                  | Nelfinavir | -8.79        | 0.360  | Ser46 Thr45 Met49 Val3 Leu4 His163 Cys145 Leu27 Thr25 Gly143 Thr26 Thr25 Asn142                                                                         |
|    |                  | Remdesivir | -6.81        | 0.01   | Gln189 Met49 Met165 His172 Glu166 Val3 Leu4 His163 Phe140 Leu141 Asn142 Gly143 Thr26                                                                          |
|    |                  | Ritonavir  | -7.53        | 0.003  | Asn142 Leu141 Gly143 Leu27 Cys145 His167 Met165 Leu4 Met49 Val3 Glu166 Gln189                                                                               |
| 2  | Mpro 2GTB        | Aloeemodin | -7.25        | 4.85   | Phe185 Val186 Ala173 Met165 Gln189 Thr190 Gln192 Leu167 Pro188                                                                                        |
|    |                  | Emodin     | -6.37        | 21.57  | Met165 Gln189 Ala173 Thr190 Phe185 Gln192 Ala191 Leu167 Pro168                                                                                        |
|    |                  | Kaempferol | -7.87        | 1.69   | Tyr54 His164 His41 Cys44 Cys145 Ser144 Asp187 Met165 Met49 Gln168 Gln189                                                                                  |
|    |                  | Quercetin  | -7.59        | 2.72   | Gln166 Gln189 Met165 Met49 Asp187 His164 His41 Tyr54 Ser144 Cys44 Cys145                                                                               |
|    |                  | Quercitin  | -5.98        | 41.52  | Pro168 Leu167 Ala191 Thr190 Gln166 Met165 Gln189 His164                                                                                        |
|    |                  | Hidroxychloroquinin | -9.45     | 0.117  | Asn142 Ser144 Cys145 His163 Met165 Met49 His41 His164 Arg188                                                                                  |
|    |                  | Lopinavir  | -10.7        | 0.014  | Leu141 Asn142 Pro168 Gln166 His163 Met165 Gln189 Thr190 Met49 Cys44 His41 His164 Asp187 Arg188                                                          |
|    |                  | Nelfinavir | -11.64       | 0.002  | Pro168 Gln166 Leu167 His163 Met165 Gln189 Cys145 Gln192 Met49 Cys44 Asp187 Arg188                                                                      |
|    |                  | Remdesivir | -8.54        | 0.546  | Cys145 Cys44 His141 His154 Thr54 Asp187 Met49 Met165 His163 Asn142 Gln166 Leu167 Arg188 Gln192                                                     |
|    |                  | Ritonavir  | -7.77        | 0.002  | Thr25 Cys145 Ser144 Gly143 Asn142 Gln166 Met165 His41 Met49 Gln189                                                                                            |
| 3  | 3CLpro 3M3V      | Aloeemodin | -4.66        | 386.9  | Arg40 Glu55 Tyr54 Asn53 Pro52 Arg188 Asn51                                                                                        |
|    |                  | Emodin     | -4.27        | 744.4  | Gln192 Ala191 Thr190 Gln189 Leu167 Glu166 Pro268                                                                                                           |
|    |                  | Kaempferol | -4.85        | 278.5  | Ala191 Thr190 Pro168 Gln189 Glu166 Leu50 Glu47                                                                                                           |
Figure 1. Histogram showing molecular docking results Sample compounds from *C. alata* and *D. petandra*
Proteases represent potential targets for the inhibition of CoV replication, and the protein sequences of the SARS-CoV M\textsubscript{pro}/3CL\textsubscript{pro} and the 2019-nCoV M\textsubscript{pro}/3CL\textsubscript{pro} are 96% identical, and the active sites in both proteins remain free from mutations. The disruption of protease activity can lead to various diseases; thus, commonly, host proteases can be used as potential therapeutic targets. In many viruses, proteases play essential roles in viral replication; therefore, proteases are often used as protein targets during the development of antiviral therapeutics.

4. Conclusion
Virtual screening with docking simulation in this study was carried out on five chemical compounds contained in \textit{C. alata} and \textit{D. petandra} on the SAR-CoV-2 protein and also five commercial antiviral compounds. Molecular docking analysis produces ligand interactions with the binding site. The interaction of ligand binding with SAR-CoV-2 shows that the interaction of amino acid residues plays an important role in maintaining functional conformation. The mechanism of enzyme and ligand interaction in this study is useful for understanding the prospective mechanism of the interaction of SAR-CoV-2 and ligands. Two parameters of the virtual filtering simulation that can predict SAR-CoV-2 inhibitory activity are low free energy Gibbs and low kinetic inhibition value. Gibbs free energy (\(\Delta G\)) resulting from docking simulation were shown quercetin, kaempferol and aloemodin have lower free energy Gibbs and low kinetic inhibition compared to other sample compounds. Nelfinavir, lopinavir and hydroxychloroquine as commercial antivirals show good results as antivirals compared to other commercial compounds.
Acknowledgments
We are grateful for the funding supported by The Deputy of Engineering Sciences-LIPI. We are also very thankful for all the support and help from Research Center for Chemistry – Indonesian Institute of Sciences in facilitating the research. The first author who is also the correspondent author is the main contribution to this work.

References

[1] Xu X, Chen P, Wang J, Feng J, Zhou H, Li X et al 2020 Sci. China Life Sci. 63 457-60
[2] Ji W, Wang W, Zhao X, Zai J and Li X 2020 J. Med. Virol. 92 433-40
[3] Zhu N, et al. 2020 N. Engl. J. Med. 382 727-33
[4] Shu Y and McCauley J J E 2017 Euro Surveill. 22(13) 30494
[5] Chen Y, Liu Q and Guo D 2020 J. Med. Virol. 92 418-23
[6] Zhou P, et al. 2020 Nature 579 270–73
[7] Wu F, et al 2020 Nature 579 265-69
[8] Needle D, Lountos G T and Waugh D S 2015 Acta Crystallogr. D Biol. Crystallogr. 71 1102-111
[9] Anand K, Ziebuhr J, Wadhwani P, Mesters JR and Hilgenfeld R 2003 Science 300 1763-67
[10] Ghosh AK, et al. 2005 J. Med. Chem. 48 6767-71
[11] Kumar V, Tan KP, Wang YM, Lin SW and Liang PH 2016 Bioorg. Med. Chem. 24 3035-42
[12] Pillaiyar T, Manickam M, Namasivayam V, Hayashi Y and Jung SH 2016 J. Med. Chem. 59 6595-628
[13] Qamar M T, et al. 2019 Sci. Rep. 9 1-16
[14] Notredame C, Higgins DG, Heringa J and Coffee T 2000 J. Mol. Biol. 302 205-17
[15] Gouet P, Courcelle E, Stuart DI and Metoz F 1999 Bioinformatics 15 305-8
[16] Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, et al. 2006 Bioinformatics 5(1) 5-6.1-5.5.30
[17] Johnson M, Zaretskaya I, Raytselis Y, Meregzhuk Y, McGinis S and Madden TL 2008 Nucleic Acids Res. 36 5-9
[18] Gasteiger E, Gattiker A, Hoogland C, Ivanyi I Appel RD and Bairoch A 2003 Nucleic Acids Res. 31 3784-88
[19] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. 2004 J. Comput. Chem. 25 1605-12