In silico molecular docking and molecular dynamics examination of Andrographis paniculata compounds of Andrographolide, Neoandrographolide, and 5-hydroxy-7,8,2',3'-tetramethoxyflavone inhibition activity to SARS-CoV-2 main protease

Raihan Kenji Rizqillah1, Jaka Fajar Fatriansyah1*, Fadilah2, Suhadi3, Siti Wahyuni3, Muhammad Arif Sudirman3, Helya Chafshoh Nafisah3, and Sukma Dewi Lestari3

1Department of Metallurgical and Materials Engineering, Faculty of Engineering, University of Indonesia, Depok, Jawa Barat, Indonesia, 16424
2Department of Medicine Chemistry, Faculty of Medicine, University of Indonesia, Salemba, Jakarta, Indonesia, 10430
3Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Semarang 50229

Abstract. In this work, Andrographis paniculata compounds of Andrographolide, Neoandrographolide, and 5-hydroxy-7,8,2',3'-tetramethoxyflavone inhibition activity to SARS-CoV-2 main protease were examined through in silico molecular docking and molecular dynamics simulation, with Remdesivir as control ligand. Docking score and MMGBSA were examined as well as molecular dynamics parameters: RMSD, RMSF and Protein ligand contact fraction. Our study found that Andrographis paniculata compounds of Andrographolide, Neoandrographolide, and 5-hydroxy-7,8,2',3'-tetramethoxyflavone have comparable inhibition activity to SARS-CoV-2 main protease in comparison to Remdesivir. 5-hydroxy-7,8,2',3'-tetramethoxyflavone has the lowest docking score, which was further validated by protein ligand contact fraction examination, although MMGBSA score is lowest for Remdesivir. Keywords: In silico molecular docking, molecular dynamics, SARS-CoV-2, Andrographis paniculata, Remdesivir.

1 Introduction

COVID-19 is an infectious disease caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) virus, a member of the family Coronaviridae family Nidovirales [1-3]. COVID-19 was officially declared as a global pandemic by the World Health Organization (WHO) in March 2020 [4]. SARS-CoV-2 has genetic similarity of 79 % to SARS-CoV which was first identified in 2002 and 50% with Middle East respiratory syndrome coronavirus (MERS-CoV) [5-6].

Epidemiologist determined that the COVID-19 viruses’ common mode of transmissions occurs through the released droplets during face-to-face exposure during talking, coughing, and sneezing [7]. Patient of COVID-19 can be distinguished as symptomatic and asymptomatic. Previous study showed that one in five of asymptomatic patients could transmit the virus, although less significantly than the symptomatic patients [8]. The most common symptoms experienced by symptomatic patients are: fever (70%-90%), dry cough (60%-86%), shortness of breath (53%-80%), fatigue (38%), muscle aches (15%-80%), nausea or diarrhea (15%-39%), headache and weakness (25%) and rhinorrhea (7%) [9].

Sambiloto (Andrographis paniculata) is a local plant from Indonesia which is produced up to 2290 ton per year. Various immunomodulatory and anthelmintic compounds such as andrographolide, saponin, tannin, etc. contained in sambiloto have hepatoprotective, antioxidant [10], anticancer, antiobesity and antidiabetic properties [11]. In traditional Chinese medicine, sambiloto has been long used to treat fever and various infections, such as: throat infection, pneumonia, tonsillitis, dysentery, gastroenteritis, and pyelonephritis [12-13].

Computer aided drug designing approach, molecular docking, and molecular dynamics, has been very effective in the past in drug discovery for various diseases [14]. Through in silico molecular docking, Rajagopal et al., found that Andrographolide has significant bond with the active site of SARS-CoV-2 main protease, PDB ID 5R82 and might produce significant activity [15]. Hiremath, et al. [16] also showed some compounds in Andrographis paniculata have potential as SARS-CoV-2 inhibitor. Main protease (Mpro) is viral protease that perform cleaving two polyprotein into individual non-structure protein. Mpro has important role in viral replication, thus can be target protein for inhibition [17]. Remdesivir has been reported to be a potential inhibitor of SARS-COV-2
Mpro and reported to be effective in treating COVID-19 [18-20]. In this paper, we investigate the inhibition activity of some *Andrographis paniculata* compounds (Andrographolide, Neandrographolide, and 5-hydroxy-7,8,2',3'-tetramethoxyflavone) to Mpro of SARS-CoV-2 through molecular docking, as well as examine dynamics interaction between compounds and Mpro through molecular dynamics simulation. Molecular docking and dynamic simulation of Remdesivir with Mpro was also performed to be positive control.

2 Computational Method

2.1 Ligand and protein preparation

3D structure of sambiloto compounds and remdesivir as control, called as ligands, were obtained from PubChem database in the form of a 3D conformer with the .SDF format for later preparation using the Ligprep tool available on the Maestro Schrodinger software. The use of remdesivir as control ligand was determined as the study by McCready et al. showed that some of the randomized clinical trials suggested that remdesivir may improve recovery for individuals hospitalized with COVID-19, although inconclusive findings exist [21-23]. As for the preparation process, the ligands were desalted, and hydrogen atoms were added. All possible ionization states and tautomers were created using the Epic tool by loading the receptor file (.zip format) and the protein data bank (PDB ID: 7AOL) in .pdb format [24]. The typical structure of the proteins in this .pdb file cannot be used directly in in silico docking simulation calculations because they might consist heavy atoms and may include native ligands, water molecules, metal ions and crystallized cofactors. The proteins were prepared using the Protein Preparation Wizard in the Maestro Schrodinger software by adding missing hydrogen atoms, side chains, and loops and removing water molecules from the target. Next step is adjusting the ionization and tautomeric state of the hetero group using Epik. The final step is optimization of the hydrogen bonding and refining the structure by limiting the RMSD minimization to 0.3 Å.

2.2 Grid generation

Grid box in simulation was prepared using the Receptor Grid Generation tool. This was done to specify that the simulation process was carried out in a certain area. The grid is set at the centre of the inhibitor RQH-403 which is the native ligand of the PDB file 7AOL at the protein active site. The Van der Waals radius scale factor was set at 1.0 with a partial charge cut off 0.25. Then, the grid area is the centroid of workspace ligand, and the box size is 20 Å.

2.3 Molecular docking

Molecular docking was done using the Ligand Docking tool by loading the receptor file (.zip format) and the prepared ligand (.mae format) into working space. The Van der Waals radius scale factor was set at 1.0 with a partial charge cut off 0.25. Extra precision (XP) algorithm and flexible ligand sampling was used. The output of this docking process will be saved in .mae format.

2.4 Molecular Mechanics-Generalize Born Surface Area (MM-GBSA)

The results obtained from molecular docking did not show the affinity or free energy of the bond between the protein and the ligand. Therefore, it is necessary to calculate the MM-GBSA to determine the binding free energy. MM-GBSA can be calculated using Prime MM-GBSA tools on Maestro Schrodinger by loading the output file of molecular anchoring (.mae format). The solvation model in the option was set as VSGB and the sampling method on protein flexibility was minimized.

2.5 Molecular Dynamics

Prior to molecular dynamics simulation, it is necessary to prepare a system to determine parameter settings using the System Builder tool. Solvent model in the system was set as simple point charge (SPC) model with ensemble NPT. The simulation box shape was set to orthorhombic with distance a = 10.0 Å b = 10.0 Å c = 10.0 Å with minimized volume. Ion neutralization was carried out by adding Na+ ions. Salt was added with concentration of 0.15 M.

Molecular dynamics simulation was set for 10 ns to all ligands with the trajectory interval being of. Ensemble class uses NPT with a constant temperature of 310 K and a constant pressure of 1.0132 bar. The output obtained from this simulation was visualized for discussion.

3 Result and discussion

3.1 Molecular docking and MM-GBSA

In silico molecular docking involves conformation and orientation of the ligand when it binds the target. The docking score is a quantitative description of the bonding modes that occur in molecular binding and the score can distinguish whether ligands bind the target or not [25]. The lower the energy value indicates that the intermolecular bonds forces are stronger.

To add discussion, we incorporated MM-GBSA (molecular mechanics-generalized Born and surface area) or binding energy, which is combination of the probability of free ligand-protein and the effect of the solution, composed of coulomb energy, covalent binding energy, Van der Waals energy, lipophilic energy, generalized Born electrostatic solvation energy, hydrogen binding energy, and π-π packing energy.
According to Table 1, most of the compounds contained in the sambiloto have lower docking score than remdesivir (-1.204 kcal/mol). 5-Hydroxy-ƍƍ-tetra-methoxyflavone compound was observed to has lowest docking score at -3.180 kcal/mol, while the lowest binding energy is Remdesivir at -45.99 kcal/mol.

![Fig. 1. 2D projection of Remdesivir interaction with Mpro](image1)

Difference between docking score and MMGBSA due to the different approach where MMGBSA allows for complex free energy decomposition into contributions from various groups of atoms and types of interaction [26]. The ligands-target interactions were shown in Figure 1-4 for Remdesivir, Andrographolide, Neoandrographolide, and 5-hydroxy-ƍƍ-tetramethoxyflavone respectively.

For Remdesivir-target interaction, Glu 178 is an acceptor that forms two hydrogen bonds with the hydroxyl and amine groups of Remdesivir (Fig.1). For Andrographolide-target interaction, they form two hydrogen bonds. One OH group is a donor to Phe 103, and another is acceptor with Lys 102 (Fig. 2). For Neoandrographolide-target interaction, they form three hydrogen bonds to target. donor to Gln 110, donor to Arg 105, and acceptor to Phe 103 (Fig. 3).

As observed, 5-Hydroxy-7,8,2',3'-tetramethoxyflavone has the highest docking score among all ligands, in agreement to the number of bonds that occurred between the proteins and ligand. For MMGBSA,
Neoandrographolide has the lowest score since the net of exchange energies between two residues (Arg 105 and Phe 103) in favour of stronger interaction.

Fig. 4. 2D projection of 5-Hydroxy-7,8,2',3'-tetramethoxy flavone interaction with Mpro

The difference between docking score and MMGBSA values exist, and it is not supposed to be compared. Ignjatović et al. [27], found there are almost no correlation between those two scoring functions, and both can be used independently as scoring function to predict binding affinity and ligands performance. According to our results, all of three sambiloto ligands, has comparable performance with Remdesivir, and in the case of docking score, all sambiloto ligands outperform Remdesivir.

3.2 Molecular Dynamics (MD)

Now we discuss about molecular dynamics simulation interaction between target and ligands. Molecular dynamics simulations were carried out to determine the stability of the ligands when interacting with proteins by considering the effects of the applied ensemble [28]. The results of the molecular dynamics simulation are the root mean square deviation (RMSD), root mean square fluctuation (RMSF) and protein-ligand contact [29].

RMSD shows whether the system has reached stability by observing the fluctuation until the value is not significant enough after reaching a specific period. If the RMSD is still fluctuating significantly until a specific time, then the system is not stable. The RMSD of the Cα and the side chain of the protein along simulation with sambiloto ligands were shown in Fig. 5-8. RMSD value is used is used to compare the docked conformation to reference conformation. The RMSD values below 2.0 Å with respect to the corresponding atomic structures are considered favourable [30]. In the beginning of the simulation, there was a sharp increase in the RMSD value for all compounds where the interaction attempted to reach equilibration. To reach a stable condition (equilibration), remdesivir took about 1.5 ns and during simulation time it reach up 2 Å, but only a glimpse. while andrographolide is less than 1 ns. In neo-andrographolide, stability occurred after 2 ns. while for the 5-hydroxy-7,8,2',3'-tetramethoxy flavone compound in sambiloto, the RMSD of Cα get higher than 2 Å and eventually reach stability under 2 Å after took 8.5 ns to reach stability. But none all the compounds have RMSD value of side chain below 2.0 Å while RMSD value of Cα were below 2.0 Å for all ligands. We had extended the time, but we could not find interesting features

Fig. 5. RMSD of Remdesivir MD simulation

Fig. 6. RMSD of Andrographolide MD simulation

Fig. 7. Protein RMSD of Neoandrographolide MD simulation

Fig. 8. RMSD of 5-hydroxy-7,8,2',3'-tetramethoxy flavone MD simulation

The stability of the ligand-protein interaction can be examined too through the RMSF values. In contrast to RMSD, RMSF is calculated for each residue to see how much fluctuation of each residue during the simulation [31]. In general, RMSF provides fluctuations of
individual atoms throughout the simulation [29]. RMSF of all ligands-protein simulation were shown in Fig. 9-12, with green line indicates interacting residue.

![Fig. 9. Protein RMSD of Remdesivir MD simulation](image)

![Fig. 10. RMSD of Andrographolide MD simulation](image)

![Fig. 11. Protein RMSD of Neoandrographolide MD simulation](image)

![Fig. 12. Protein RMSD of 5-hydroxy-7,8,2',3'-tetramethoxy flavone MD simulation](image)

3.3 Protein-ligand contact fraction

In this subsection, protein-ligand contact was investigated to further validate docking scores and MM-GBSA values. Protein-Ligand contact during the simulation was recorded and processed as a simulation interaction diagram and timeline representation. These results inform the type, fraction, which residue is involved. Interaction fraction value inform fraction of simulation time that a specific interaction occurs. For instance, interaction fraction 0.6 means that a residue interacted with ligand for 60% of simulation time, or in this study of 10ns molecular dynamic equal to 6ns. Value of more than 1.0 is possible for a single residue as it may contact multiple groups in a ligand.

![Fig. 13. Protein-ligand contact of Remdesivir MD simulation](image)

In remdesivir (Fig. 13), there were 15 contact residue interactions between the ligand and target: Asn 84, Phe 103, Val 104, Arg 105, Ile 106, Gln 107, Phe 134, Asp 176, Glu 178, Asn 180, Phe 181, Gly 183, Pro 184, Phe 185. The highest fraction interaction is Arg105 at 0.8 because it has water bridges, hydrophobic, hydrogen bonds interactions to the target.

![Fig. 14. Protein-ligand contact of Andrographolide MD simulation](image)

Figure 14 shows the presence of 22 contact residues with andrographolide namely Gln 83, Tyr 101, Lys 102, Phe 103, Val 104, Arg 105, Ile 106, Gln 107, Phe 134, Asp 176, Glu 178, Asn 180, Phe 181, Gly 183, Pro 184, Phe 185. The highest fraction interaction is Arg105 at 0.8 because it has water bridges, hydrophobic, hydrogen bonds interactions to the target.

Based on the RMSF evaluation, some of the protein residues of remdesivir and sambiloto MD simulation showed fluctuations higher than 2.5 Å. Further observation shows that the residue fluctuated more than 2.5 Å were not the interacting residue. Whereas all interacting residues, indicated by green line, generated RMSF below 2.5 Å, thus it is considered acceptable [32] indicating that the interaction of ligands with proteins are stable.
In Figure 15, there are 15 residual contacts of Neoandrographolide, namely Tyr 101, Lys 102, Phe 103, Val 104, Arg 105, Gln 107, Gln 110, Ile 152, Asp 153, Tyr 154, Ser 158, Phe 294, Val 297, Arg 298, Ser 30. Arg 298 has the highest fraction at 1.6 produced from hydrogen bonding and relatively stable travel times from 3 ns to 10 ns.

Most of ligands interact with Arg 105 residues through various interactions except for Neoandrographolide with its dominant interaction with Arg 298 residue. However, it is for 5-hydroxy-7,8,2',3'-tetramethoxyflavone which has high frequency of hydrophobic interactions, distributed to Arg 105, Phe 134, and pro 184 residues, these hydrophobic interactions which distributed, are more frequent in high-efficiency ligands [33], thus validates the lowest docking score of 5-hydroxy-7,8,2',3'-tetramethoxyflavone.

**References**

1. C.S.G. of the International, Nat. Microbiol., 5,4 (2020)
2. H. Harapan, N. Itoh, A. Yufika, W. Winardi, S. Keam, H. Te, D. Megawati, Z. Hayati, A.L. Wagner, M. Mudatsir, J. Infect. Public. Heal., 13, 5 (2020)
3. N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, N. England J. M., (2020)
4. D. Zhang, M. Hu, Q. Ji, Finance Res. Lett. 36, 101528 (2020).
5. A. A. Rabaan, S. H. Al-Ahmed, S. Haque, R. Sah, R. Tiwari, Y. S. Malik, A. J. R. Morales, Infez. Med., 28, 2 (2020).
6. Z. Abdelrahman, M. Li, X. Wang, Front. Immunol., 11, 2309 (2020)
7. W.J. Wiersinga, A. Rhodes, A.C. Cheng, S.J. Peacock, H.C. Prescott, Jama, 324, 8 (2020)
8. B. Nogradry, Nature, 587, 7835 (2020)
9. R. Mao, Y. Qiu, J.S. He, J.Y. Tan, X.H. Li, J. Liang, J. Shen, L.R. Zhu, Y. Chen, M. Iacucci, S.C. Ng, Lancet Gastroenterol. Hepatol., 5, 7 (2020)
10. Niranjan, S.K. Tewari, A. Lehri, Indian. J. Nat. Prod. Resour., 1, 2 (2010)
11. Y. Dai, S.R. Chen, L. Chai, J. Zhao, Y. Wang, Y. Wang, Crit. Rev. Food. Sci. Nutr., 59, 1 (2019)
12. K.V. Ergil, E.J. Kramer, A.T. Ng, West. J. Med., 176, 4 (2002)
13. W.L. Deng, R.J. Nie, J.Y. Liu, Chin. Pharm. J., 17, 1 (1982)
14. M. Alazmi, O. Motwalli, J. Biomol. Struct. Dyn., 1, 11 (2020)
15. K. Rajagopal, P. Varakumar, A. Baliwada, G. Byran, Future J. Pharm. Sci., 6, 1 (2020)
16. S. Hiremath, H.V. Kumar, M. Nandan, M. Mantesh, K.S. Shankarappan, V. Venkatatarvanappan, C.J. Basha, C.L. Reddy, 3 Biotech, 11, 2 (2021)
17. J. Qiao, Y. S. Li, R. Zeng, F. L. Liu, R. H. Luo, C. Huang, ... & S. Yang., Sci., 371,6536 (2021)
18. A. M. da Fonseca, A. L. G. Cavalcante, R. M. M. Carvalho, J. F. do Amaral, R. P. Colares, E. S. Marinho, M. M. Neto, Int. J. Res, 8,11, (2020)
19. H.L. Nguyen, N.Q. Thai, D.T. Truong, M.S. Li, J. Phys. Chem. B, 124,50 (2020)
20. F.I. Khan, T. Kang, H. Ali, D. Lai, Front. Pharmacol., 12, (2021)
21. E. K. McCready, D. C. Angus, Jama, 324,11 (2020)
22. L. Y. Wang, J. J. Cui, Q. Y. Ouyang, Y. Zhan, C. X. Guo, J. Y. Yin, Lancet, 396,10256 (2020)
23. M. M. Ignjatovic, O. Caldararu, G. Dong, C. Munoz-Gutierrez, F. Adasme-Carreno, U. Ryde, J. Comput. Aided, 30,9 (2016)
24. J. D. Norrie, Lancet, 395,10236 (2020)
25. S. Günther, P. Y. Reinke, Y. Fernández-García, J. Lieske, T. J. Lane, H. M. Ginn, A. Meents, Science, 372,6542 (2021)
26. S. Genheden, O. Kuhn, P. Mikulskis, D. Hoffmann, U. Ryde, J. Chem. Inf. Model., 52,8 (2012)
27. T. Hou, J. Wang, Y. Li, W. Wang, J. Chem. Inf. Model., 51,1 (2011)
28. Singh, A. Mishra, J. Biomol. Struct. Dyn. J. Biomol. Struct. Dyn., (2020)
29. K.A. Peele, C.P. Durthi, T. Srihansa, S. Krupanidhi, V.S. Ayyagari, D.I. Babu, M. Indira, A.R. Reddy, T.C. Venkateswarulu, Inform. Med. Unlocked, 19 (2020)
30. J. M. Yang, C. C. Chen, Proteins., 55,2 (2004)
31. M.S. Zubair, S. Maulana, A. Mukaddas, Galenika Journal of Pharmacy, 6,1 (2020)
32. R. K. Pandey, D. Sharma, T. K. Bhatt, S. Sundar, V. K. Prajapati, J. Biomol. Struct. Dyn., 33,12 (2015)
33. R. F. de Freitas, M. Schapira, Medchemcomm., 8,10 (2017)