The target of SARS-CoV-2: Analysis of N3 ligand binding energy vs. natural compounds (curcumin) using molecular dynamics of force fields CHARMM

Cite as: AIP Conference Proceedings **2663**, 030001 (2022); https://doi.org/10.1063/5.0108013
Published Online: 20 September 2022

Riyan Hidayat, Lutfi Rohman and Artoto Arkundato

The Target of SARS-CoV-2: Analysis of N3 Ligand Binding Energy vs. Natural Compounds (Curcumin) Using Molecular Dynamics of Force Fields CHARMM

Riyan Hidayat[1], Lutfi Rohman[1,a)], Artoto Arkundato[1]

Department of Physics, Faculty of Mathematics and Natural Sciences, University of Jember, East Java, Indonesia

[a)]Corresponding author: *el_rahman.fmipa@unej.ac.id*

Abstract. SARS-CoV-2 has been an endemic disease until now, but scientists are still trying to find an effective drug. Research has been carried out to find drugs that are effective against SARS-CoV-2. Mpro is used as the main target to be attacked. A natural compound test ligand (curcumin) was used anda compared with the N3 ligand. This research used molecular dynamics methods to identify and study inhibitor binding interactions. Molecular dynamics simulation using software Gromacs with CHARMM force field at 1000 kJ/mol and 300 K/1 bar conditions. This research will analyze the trajectory of the simulation results for the value of RMSD, RMSF, and Rg, which emphasizes the conditions of balance and compactness. The drug-binding affinity of the main protease SARS-CoV-2 was analyzed using the MM/PBSA method. An enormous free energy value will indicated the strength of the interaction between SARS-CoV-2 and the N3/curcumin ligand. From the results, it can be seen that the total interaction energy of the N3-protein ligand is -333.717 kJ/mol, and that of the curcumin-protein ligand is -0.023 kJ/mol. Meanwhile, based on the value of free energy on N3-protein is -28.566 kJ/mol and curucumin-protein 9.477 kJ/mol. Thus, this study showed that the N3 ligand was more effective than the curcumin ligand in binding to the main protease of SARS-CoV-2.

Keywords: SARS-CoV-2, Mpro, molecular dynamics, MM/PBSA, curcumin

INTRODUCTION

Covid-19 is a disease that has plagued worldwide and infected hundreds of millions of people with an increasing number of deaths. However, until now, scientists are still trying to find an effective drug for Covid-19. Today, with technological advances, the process of drug discovery through the world of simulation is getting closer and closer to accuracy and speed. A study conducted by Coundhary et al. (2020)[1]succeeded in discovering the crystal structure of Covid-19, which allowed identification of the inhibitor. The crystal structure is called Mpro or main protease. 3CLpro or Mpro through mass spectroscopic measurements are known to have a molecular mass of 33.797 kDa, with its catalytic efficiency on SARS-CoV-2 of 28,500 $M^{-1}S^{-1}$ [2][3]. Inhibition of the main protease (Mpro) can disrupt the replication and transcription of non-structural proteins that can cause the virus's death. According to Zhang et al. (2020)[3], It is known that humans do not have enzymes that are similar to this Mpro virus, so that the inhibitors of this Mpro will not cause toxic effects on the human body.

Research on the drug discovery process by targeting Mpro as a drug target has been carried out previously, including a study conducted by Ullrich and Nitsche (2020)[4] using N3 as the ligand. The N3 ligand was obtained by combining drug designs through high-speed virtual drug screening. This program focuses on identifying the existing drug load by targeting the main protease (Mpro). Another study conducted by Jin et al. (2020)[2] N3, in particular, has also shown the ability to inhibit Mpro of various coronaviruses, including SARS-CoV and MERS-CoV, has known effective antiviral activity against infectious bronchitis virus.

In this study, the mechanism and interaction between Mpro and N3 and ligands will be investigated then N3 is replaced curcumin ligand. The curcumin ligand was chosen because it refers to the docking study conducted by Yasin et al. (2020)[5]in thousands of Indonesian herbal medicine bioactive compounds. That research found that

The 3rd International Conference on Physical Instrumentation and Advanced Materials (ICPIAM) 2021
AIP Conf. Proc. 2663, 030001-1–030001-7; https://doi.org/10.1063/5.0108013
Published by AIP Publishing. 978-0-7354-4393-8/$30.00

the curcumin compound had the most effective value among other test ligands. The compound curcumin has bond-free energy (ΔG) and the lowest inhibition constant value (Ki) at -6.9 kcal/mol and 2.99 μM. Therefore, it is necessary to conduct further studies to observe curcumin an inhibitor of SARS-CoV-2. Whether the results obtained are better or not compared to the N3 ligand.

There are many simulation methods for drug discovery, including the study conducted by Nayeem et al. (2021)[6] using the OPLS-AA force field molecular dynamics method, in continuity with the study of molecular mechanics Poisson-Boltzmann surface area method. Molecular dynamics is one of the most widely used methods in the drug discovery process. Molecular dynamics works based on the equations of Newton's laws and the laws of classical mechanics. The magnitude of the intermolecular forces is calculated explicitly, and the molecular movement is processed computationally by the integration method used to solve Newton's equations on the constituent atoms. The initial state of the molecule is described by the positions and velocities of the atoms. Based on Newton's perception, from the initial position, position and velocity can be calculated. In a small time interval and the force calculation at the new position can be repeated for a few moments, up to hundreds of times.[7]. Molecular dynamics plays a role in knowing and studying the binding interactions of Mpro inhibitor compounds or the interaction between ligands and protein proteases.

In molecular dynamics simulation, we need a force field. The force field is a computational method used to predict atomic forces occurrence in molecules and between molecules. In this study, the force field used is the CHARMM force field. The CHARMM force field was chosen because this force field has been widely used to study proteins, nucleic acids, lipids, and drug discovery to simulate complex chemical systems[8].

The result of the molecular dynamics process is in the form of a trajectory file. The obtained trajectories can be used for molecular mechanics studies. Molecular mechanics results in a calculation of the total energy of a compound without using any calculation of the wave function of electron density. The energy statement contains simple classical equations, such as the harmonic oscillator equation to describe the energy involved in the stretching, bending, and torsion of bonds, as well as intermolecular forces, such as van der Waals interactions and hydrogen bonds [9]. Simply put, In drug simulations, molecular mechanics is used to evaluating the binding affinity of compounds and identify residues important for binding Mpro.

MATERIALS AND METHODS

Material

The material used in this study is a 3-dimensional coordinate structure of proteins and ligands. The data is obtained from PDB (Protein Data Bank) with the address rscb.org [10]. The protein used is the Mpro structure with ID. 6LU7. The ligand structure used as a comparison was also obtained from the process of molecular anchoring studies and ADMET predictions conducted by Yasin et al. (2020)[5]. The data obtained previously is a complex protein that must be separated first to be simulated.

Preparation

Preparation and optimization of the ligand structure started by downloading the 3-dimensional structure through the PDB website in .pdb format, then the structure was examined, and unneeded hetero atoms were removed, especially for the N3 inhibitor ligand obtained from PDB ID. Because 6LU7 is a complex protein, it is necessary to separate the ligand and receptor using PyMOL software. After checking and obtaining the appropriate structure, the following process is to attach hydrogen atoms and change the format.

Molecular Dynamics Simulation

The molecular dynamics simulation of the Mpro protease with N3 and curcumin ligands was carried out according to the SOP (Standard Operating Procedure) contained in the Gromacs manual and also in the book written by Lemkul (2018)[11] using the Gromacs series 2021.1 programs with several stages, namely, preparation of input files, ligand preparation, topology, and coordinate creation, energy minimization, equilibration, production, and analysis.

Molecular Mechanics Simulation

The simulation results generated from the molecular dynamics simulation process produced a file in a trajectory. The resulting trajectory file was used as an input file for the molecular mechanic's simulation process. Simulations were carried out to determine the value of the resulting binding energy. The resulting compound's more negative indicated that the compound has a strong bond and is not easily separated.

RESULTS AND DISCUSSION

Molecular Dynamics

This section contains the results of molecular dynamics simulations, including analysis of RMSD, RMSF, Rg and interaction energy, and molecular mechanics analysis of the Poisson Boltzmann surface area method, which includes analysis of binding energy Gibbs energy, polar solvation, and nonpolar solvation. Molecular dynamics is a form of computer simulation in which atoms and molecules are allowed to interact over a period of time. This simulation is based on the equations of the laws of classical mechanics, one of which is Newton's law. The purpose of the molecular dynamics simulation is to test the stability of the interaction between ligands and protein with conditions that are made as close as possible to the physiology of the human body. The following are the results of the analysis of the molecular dynamics simulation:

RMSD (Root Mean Square Deviation)

Root Mean Square Deviation or the square root of the average deviation is a value that describes the condition or state of the protein in a stable and not denatured condition. RMSD is a measure that is often used in 3-dimensional geometry. Based on the graph presented in the simulation that lasted for 10 ns, each system at the beginning of the simulation immediately experienced an increase in RMSD, which indicated that the structure was starting to unfold.

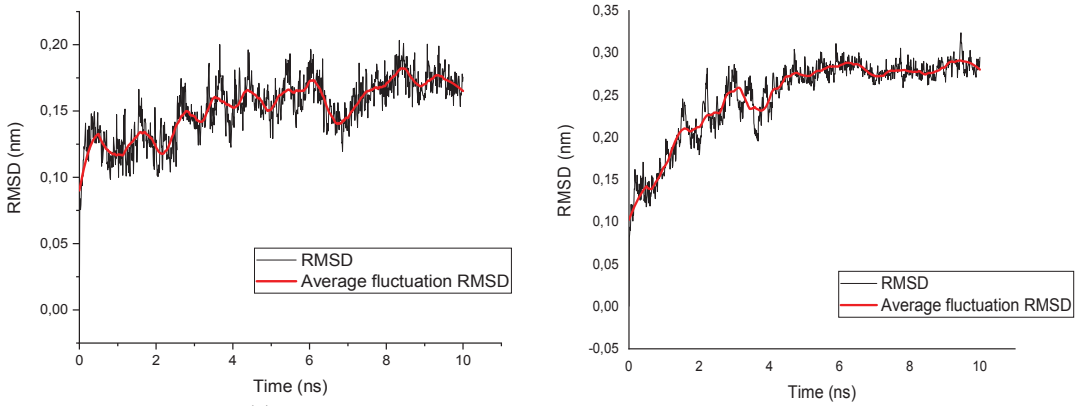


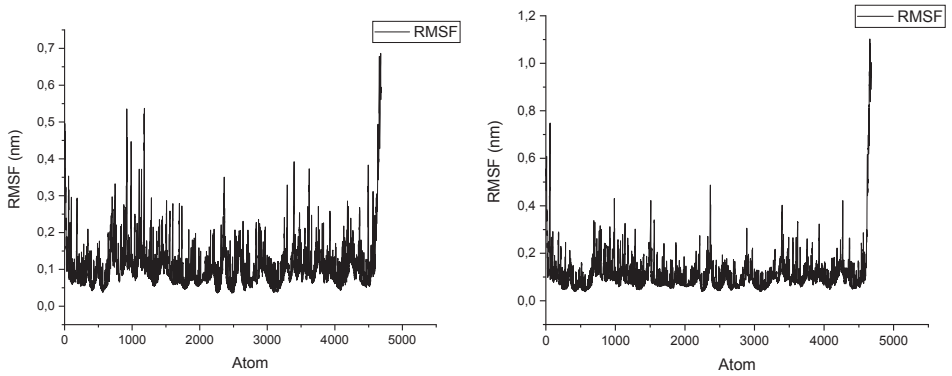
FIGURE 1. RMSD result graph MPro with (a) N3 ligand (b) curcumin ligand

In FIGURE 1. Mpro with the N3 ligand showed that at the 1st time, the system started to stabilize, but at the 5th ns, the system decreased, but after that, the ligand was stable until it was finished. The ligand is stable if the fluctuation is not more than 0.1 nm (Lemkul, 2018)[11]. The graph also shows that the fluctuation of the RMSD graph on the N3 ligand is quite large compared to the curcumin ligand, but the shift in value is smaller than that of the curcumin ligand. The curcumin ligand showed that the ligand started to stabilize at the 4 ns to completion time. In the curcumin ligand, fluctuations from the RMSD graph were not too large compared to the N3 ligand, but the curcumin ligand at time 0 to time 4 ns increased large enough. This indicates that the structure begins to open, and the ligand begins to look for the binding site and coordinates corresponding to the protein. The ligand and protein

complexes have undergone maximum conformation after binding to each other, so they tend to maintain their position. In addition, the residues on the protein make the viral protein binding (Mpro) tend to maintain its structure.

RMSF (Root Mean Square Fluctuation)

Root Mean Square Fluctuation is a measure that describes the size of the deviation between the particle position and its reference position. In contrast to RMSD, RMSF is calculated for each protein residue, meaning that it will be seen how far each residue's fluctuations in the movement are during the simulation. Broadly speaking, the RMSF value describes the conformational shift of each amino acid residue that gives protein flexibility. High flexibility describes a residue that is unstable and easily separated from its structure.



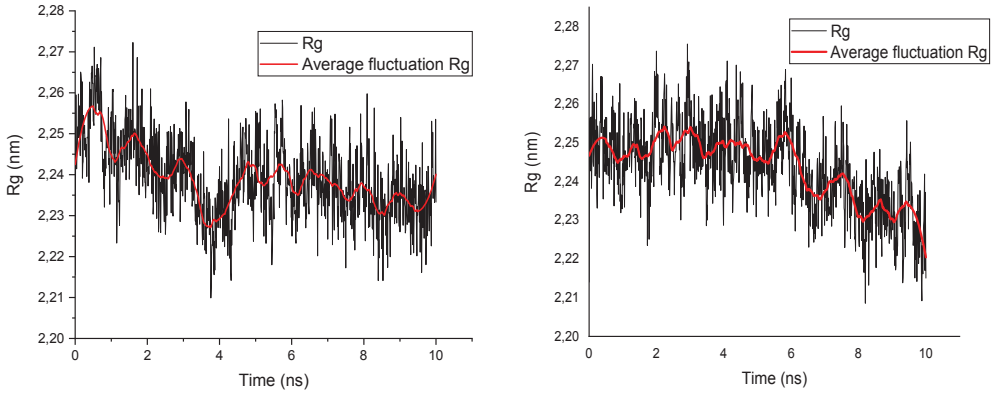
(a) (b)

FIGURE 2. RMSF results in graph Mpro with (a) N3 ligand (b) curcumin ligand

From the graph presented in what FIGURE 2, it can be seen that both the N3 and curcumin ligands have RMSF values that are not much different. However, the curcumin ligand at the beginning and the end of the simulation has a very large deviation value. This result showed that some of the protein residues have a high flexibility value, so that the structure tends to be loose or less stable. This phenomenon is because the molecules in the material have different speeds to form the velocity distribution. Statistically, it can be seen that most molecules are at a certain speed and they will decrease in the number of smaller molecules as they get further away from that speed.

Rg (Radius of gyration)

The radius of gyration measures the distance between the center of mass of a protein and its atomic system in a certain period. The Rg value represents the compactness of the protein and ligand during the simulation process. Protein-ligands that have good cohesiveness are indicated by small fluctuations, while very large fluctuations indicate that the protein-ligand structure has small cohesiveness values.



(a) (b)

FIGURE 3. Rg result graph Mpro with (a) N3 ligand (b) curcumin ligand

From the graph presented in what FIGURE 3, it can be seen that the N3-protein and curcumin-protein ligands have values that are not too different. From the graph in what FIGURE 3, it can be seen that the fluctuation of the curcumin ligand is smaller than that of the N3 ligand, but the deviation of curcumin is quite significant at the 6 ns time, and the graph starts to decrease until it is finished. While in the N3 ligand, although the fluctuation is more significant, the deviation value tends to be stable. This phenomenon indicates that at the time of simulation, the N3 ligand tends to have good cohesiveness.

Protein-Ligan Interaction Energy

The interaction energy is a value that describes the strength of the interaction unbound between ligands and proteins. The interaction energy is the contribution of the total energy caused by the interaction of compounds, in this case, are ligands and proteins. The results of the interaction energy can be seen from the image presented in what FIGURE 4 below.

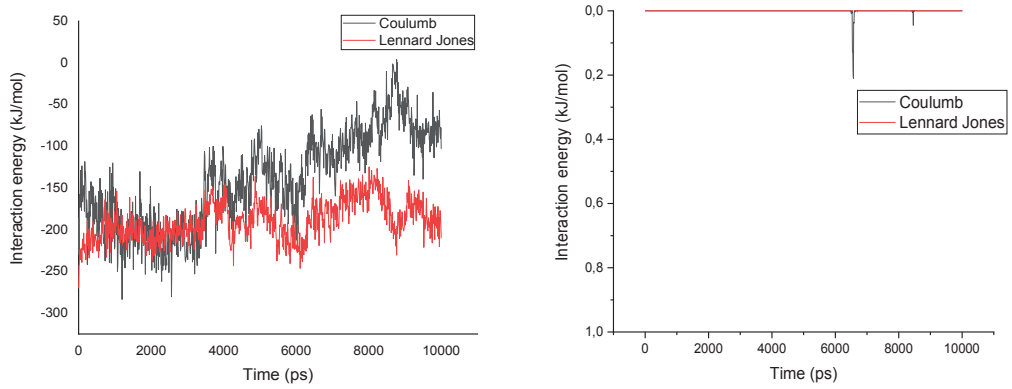
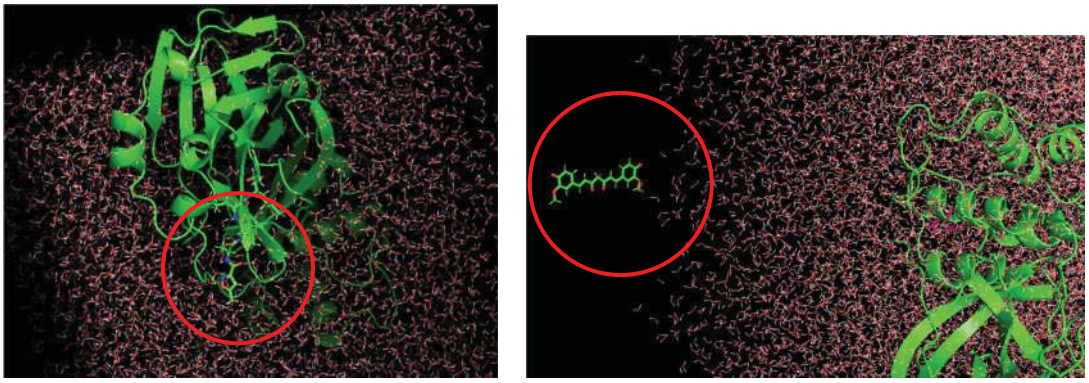


FIGURE 4. Interaction energy yield graph Mpro with (a) N3 ligand (b) curcumin ligand

The graph presented in what FIGURE 4 shows that the interaction energy is caused by two potential energies, namely the Coulomb potential and the Lennard-Jones potential. Coulomb potential is also known as electric potential energy. Electrical potential energy is produced by conservative Coulombic forces and is associated with many point charges in a system. The Lennard-Jones potential is a function of the distance between the centers of two particles. When an infinite distance separates two non-bonding particles, their potential energy is assumed to be zero. However, as the separation distance decreases as the interaction progresses, it causes the interaction to increase and the binding potential to decrease from zero to a negative quantity. As long as the particles are bound, the distance between their centers decreases until the particles reach equilibrium, which is determined by the separation distance at which the minimum potential energy is reached. Coulomb potential energy describes the intense energy between charges, while the Lennard-Jones energy describes the neutral energy between molecules or weak energy. The N3 ligand has an average Coulomb interaction energy value of -149.748 kJ/mol, and the energy of the Lennard-Jones interaction is -189.969 kJ/mol, so the total interaction energy of the N3 ligand-protein is -339.717 kJ/mol. The curcumin ligand has an average Coulomb interaction energy value of -0.00178011 kJ/mol. The energy of the Lennard-Jones interaction is -0.0215839 kJ/mol, so the total interaction energy of the curcumin-protein ligand is -0.02336401 kJ/mol. It can be seen that the curcumin ligand, based on the analysis of its total binding energy, has a lower value than the N3 ligand, and this indicates that curcumin has a bond structure with viral proteins that is more easily broken.

Molecular Dynamics Results Visualization

The characteristics of the interaction between curcumin and N3 ligands with Mpro compounds can be observed by visualizing the results of the trajectories resulting from the production process in molecular dynamics simulations. The results of the visualization are shown in FIGURE 5.



(a) **(b)**

FIGURE 5. Results of visualization of protein interactions with (a) N3 ligand (b) curcumin ligand

In FIGURE 5. (a), it is known that the final result of the simulation process on the N3 ligand shows that the ligand remains in the predetermined position (a red circle shows the ligand). This happens because the intermolecular and intramolecular binding energies of the N3 ligand are pretty large so that when the system is simulated, ligands and proteins remain bound. In FIGURE 5. (b), it is known that the final result is shown when the system is simulated shows that the curcumin ligand is separated or separated from the protein. This is due to the low binding energy between molecules in curcumin and protein so that when the simulation is carried out, the ligand is separated from the protein. Furthermore, it is reviewed based on the RMSD analysis.

Molecular Mechanics

Molecular mechanics coupled with the Poisson-Boltzmann (MM/PBSA) method is a popular approach for estimating the free energy of binding ligands and proteins. Molecular mechanics is based on the trajectory of the molecular dynamics processes of protein-ligand complexes. In complex biochemical systems, the classical approach is more often used because of its better efficiency and accuracy, therefore as an additional analytical tool to strengthen the results of this simulation, an approach with molecular mechanics Poisson-Boltzmann Surface Area (MM/PBSA) method is carried out. The results of the MM/PBSA can be tabulated as follows:

TABLE 1. Results of MM/PBSA SARS-CoV-2+N3 and SARS-CoV-2+Curcumin

System	vdW (kJ/mol)	elect (kJ/mol)	PS (kJ/mol)	SASA (kJ/mol)	G (kJ/mol)
SARS-CoV-2+N3	-95.518	-44.184	122.166	-10.981	-28.566
SARS-CoV-2+Curcumin	-0.029	-0.013	9.479	0.009	9.477

In general, when compared, based on the MM/PBSA approach, Mpro Bonds withN3 ligand has a better value. A good value of free energy is indicated by a negative value and the largest number; the higher the negative value, the stronger the bond that occurs. The graph also shows the van der Walls energy; the van der Walls energy refers to the intermolecular forces. Intermolecular forces generally refer to the forces arising from molecule polarization to dipole. The van der Walls force is a relatively weak electric attraction due to permanent or induced molecular polarity. Permanent polarity occurs due to the polarity within the molecule.

In contrast, non-permanent polarity occurs due to the molecule being induced by other charged particles so that the molecule is spontaneously polar for a moment. So based on the table, it can be seen that the bond N3 ligand-protein virus has a better polarization, so it has a stronger attractive force than the bond protein virus-curcumin ligand. This is also indicated by the PS (polar solvation) value, which shows that the bond-N3 ligand-protein virus has a higher polar solvation value than in bond curcumin ligand-protein virus. Next is the electrostatic force; this force occurs because of the attractive force between the atomic nucleus and the electrons of other atoms, called the electrostatic attraction (Coulomb force), which is generally found in polar compounds. The table also shows that the value of the electrostatic force on bond N3 ligand-protein virus is better than proper bond N3 ligands-protein virus,

so generally based on molecular mechanics studies method *Poisson-Boltzmann Surface Area* bond N3 ligand-protein virus has a better value than bond protein virus-curcumin ligand.

CONCLUSION

It can be concluded that based on molecular dynamics analysis and molecular mechanics analysis, curcumin ligands are less potential to be developed into SARS-CoV-2 drugs. Although curcumin ligands tend to be stable as shown in FIGURE 1, they are more easily released see FIGURE 5. This is indicated by lower values of interaction energy and binding energy compared to N3 ligand. From these results, it can be seen that the total interaction energy of the N3 protein ligand is -333.717 kJ/mol, and the interaction energy of the curcumin protein ligand is -0.023 kJ/mol. Meanwhile, based on the value of free energy in protein N3 is -28.566 kJ/mol and protein curucumin 9.477 kJ/mol. PaTo use curcumin ligand as a suitable inhibitor, its structure must be improved or synthesized with other compounds in order to have a better binding energy value.

ACKNOWLEDGMENTS

The author would like to thank the University of Jember for providing research facilities with a supportive computing environment. All the lecturers who have helped in this research.

REFERENCE

[1] MI Choudhary, M. Shaikh, Atia-Tul-Wahab, and Atta-Ur-Rahman, “In silico identification of potential inhibitors of key SARS-CoV-2 3CL hydrolase (Mpro) via molecular docking, MMGBSA predictive binding energy calculations, and molecular dynamics simulation,” PLoS One, vol. 15, no. 7 July, pp. 1–15, 2020, doi:10.1371/journal.pone.0235030.

[2] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, and Y. Zhao, “Structure of M pro from SARS-CoV-2 and discovery of its inhibitors,” Nature, vol. 582, no. February, 2020, doi: 10.1038/s41586-020-2223-y.

[3] L. Zhang, D. Lin, X. Sun, K. Rox, and R. Hilgenfeld, “X-ray Structure of Main Protease of the Novel Coronavirus SARS-CoV-2 Enables Design of -Ketoamide Inhibitors,” pp. 1–14, 2020, doi:10.1101/2020.02.17.952879.

[4] S. Ullrich and C. Nitsche, “The SARS-CoV-2 main protease as drug target,” Bioorg. Med. Chem. Lett., vol. 30, no. 17, p. 127377, Sept. 2020, doi: 10.1016/j.bmcl.2020.127377.

[5] SA Yasin, A. Azzahra, NE Ramadhan, and ..., "A Study of Molecular Anchoring and Prediction of ADMET Bioactive Compounds of Several Indonesian Herbal Medicines against SARS-CoV-2 Main Protease (Mpro)," Berk. science. …, vol. 7, no. 2, pp. 24-41, 2020, [Online]. Available: https://bimfi.e-journal.id/bimfi/article/view/45.

[6] SM Nayeem, EM Sohail, G. Ridhima, and MS Reddy, "Target SARS-CoV-2: computation of binding energies with drugs of dexamethasone/umifenovir by molecular dynamics using OPLS-AA force field," Res. Biomed. Eng., 2021, doi: 10.1007/s42600-020-00119-y.

[7] AD Astuti and AB Mutiara, "Molecular dynamics simulation of proteins with the application of gromacs," Teh. Information. and Ind., vol. 1, no. 2, pp. 1–9, 2011.

[8] ND Malau and M. Sianturi, “Analysis of hydrophobic interactions on the thermal stability of the Aspergillus niger xylanase enzyme,” … J. Pendidikan, Mat. and …, vols. 3, no. 2, pp. 215–227, 2019, [Online]. Available: http://repository.uki.ac.id/2614/.

[9] Taufik, FM 2014. What and How Computational Chemistry. http://www.komput.lipi.go.id/utama.cgi?article&1420077026[Accessed on June 18, 2021]

[10] Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, Duan Y, Yu J, Wang L, Yang K, Liu F, Jiang R, Yang X, You T, Liu X, Yang X, Bai F, Liu H, Liu X, Guddat LW, Xu W, Xiao G, Qin C, Shi Z, Jiang H, Rao Z, Yang H. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020 Jun;582(7811):289-293. doi: 10.1038/s41586-020-2223-y. Epub 2020 Apr 9. PMID: 32272481.

[11] J.A. Lemkul.(2018).From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package, v1.0" Living J. Comp. Mol. Sci. In Press