Journal of Tropical Pharmacy and Chemistry

Journal homepage: https://jtpc.farmasi.unmul.ac.id

Analysis of SARS-CoV-2 Spike Protein as the Key Target in the Development of Antiviral Candidates for COVID-19 through Computational Study

Taufik Muhammad Fakih*, Mentari Luthfika Dewi

Department of Pharmacy, Faculty of Mathematics and Natural Sciences, Universitas Islam Bandung, Bandung, Indonesia
*Corresponding author: taufikmuhammadf@gmail.com

Abstract

The recent public health crisis is threatening the world with the emergence of the spread of the new coronavirus 2019 (2019-nCoV) or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This virus originates from bats and is transmitted to humans through unknown intermediate animals in Wuhan, China in December 2019. Advances in technology have opened opportunities to find candidates for natural compounds capable of preventing and controlling COVID-19 infection through inhibition of spike proteins of SARS-CoV-2. This research aims to identify, evaluate, and explore the structure of spike protein macromolecules from three coronaviruses (SARS-CoV, MERS-CoV, and SARS-CoV-2) and their effects on Angiotensin-Converting Enzyme 2 (ACE-2) using computational studies. Based on the identification of the three spike protein macromolecules, it was found that there was a similarity between the active binding sites of ACE-2. These observations were then confirmed using a protein-docking simulation to observe the interaction of the protein spike to the active site of ACE-2. SARS-COV-2 spike protein has the strongest bond to ACE-2, with an ACE score of −1341.85 kJ/mol. Therefore, some of this information from the results of this research can be used as a reference in the development of competitive inhibitor candidates for SARS-CoV-2 spike proteins for the treatment of COVID-19 infectious diseases.

Keywords: spike protein, coronavirus, ACE-2, COVID-19, computational studies

Submitted: xxxxxxx **Accepted**: xxxxx **DOI**: https://doi.org/10.25026/jtpc.vxix.xxx

1 Introduction

The 2019 novel coronavirus (COVID-19) infectious disease caused by severe acute respiratory syndrome coronavirus (SARS-CoV-2) continues to spread from its origins in Wuhan City, Hubei Province, China to the rest of the world [1]. Till 10/04/2020 there are approximately 1.6 million cases in 2019

coronavirus (COVID-19), with 95.604 deaths and 355.671 recovered [2]. Indonesia has reported 4,241 cases to date. Because knowledge of this coronavirus is developing rapidly, researchers need to identify the characteristics of SARS-CoV-2 [3]. Thus, effective drug candidates for the prevention and control of this infectious disease can be found.

Coronavirus is a positive sensory RNA virus with diameters ranging from 60 nm to 140 nm belonging to the Coronaviridae family [4,5]. In general, this virus can cause mild respiratory infections in humans [6]. Several other coronavirus infections have previously caused deadly endemics, including SARS (Severe Acute Respiratory Syndrome) and MERS (Middle East Respiratory Syndrome) [7]. Both diseases are caused by zoonotic coronaviruses which belong to the genus Betacoronavirus in Coronaviridae [8]. SARS-CoV originated in southern China and was endemic in 2003, while MERS first occurred in Saudi Arabia in 2012.

Like SARS-CoV and MERS-CoV, SARS-CoV-2 has a genome size of about 30 kilobases which functions to encode several structural proteins, including spike protein (S), an envelope protein (E), membrane protein (M), and nucleocapsid protein (N) [9]. Preliminary studies also show that SARS-CoV-2 is identical to SARS-CoV based on the complete phylogenetic analysis of genomes. Because of the apparent similarity between the two viruses, it is necessary to further explore the constituent components of one structural protein, such as the spike protein from SARS-CoV and SARS-CoV-2 [10,11].

The entry of SARS-CoV-2 into cells is mediated by spike proteins found on the surface of the coronavirus and binds to Angiotensin-Converting Enzyme 2 (ACE-2) to infect host cells. SARS-CoV-2 involves ACE-2 with an affinity that is comparable to SARS-CoV and MERS-CoV [12]. A tight bond on ACE-2 can explain part of the efficient transmission of SARS-CoV-2 in humans, as happens in SARS-CoV and MERS-CoV [13,14]. Inhibition of the attachment of spike proteins from SARS-CoV-2 with ACE-2 can prevent COVID-19 infection.

Until now, the need to design effective antivirus candidates against SARS-CoV-2 has increased. This study aims to identify, evaluate, and explore the structure of spike protein macromolecules from SARS-CoV, MERS-CoV, and SARS-CoV-2, and their effects in inhibiting binding with ACE-2. Computational studies can be used to observe potential components of this coronavirus [15]. Specifically, the SARS-CoV-2 protein spike is considered a target because it is a major part of forming coronavirus characteristics. Therefore, through this research, information is expected to be used as a reference in the development of drug compounds for COVID-19 infectious diseases.

2 Materials and Methods

2.1 Spike Proteins Preparation

Macromolecules used in this study were the spike proteins of coronavirus (SARS-CoV, MERS-CoV, and SARS-CoV-2) obtained from Protein Data Bank (http://www.rcsb.org/pdb) with PDB ID 6NB6 [16], 5X5C [17], and 6VYB [18], respectively. The preparation of these three macromolecules was performed by removing water molecules and natural ligands using BIOVIA Discovery Studio 2020 [19].

2.2 Analysis of Three-Dimensional Conformation Spike Proteins

Spike protein macromolecules from SARS-CoV, MERS-CoV, and SARS-CoV-2 which were prepared then overlapped three-dimensional conformations with a representation of secondary structures to identify similarities and observe differences from the macromolecules of the two protein spikes. This process was performed using BIOVIA Discovery Studio 2020 [19] and Chimera 1.14 [20].

2.3 Identification of Spike Proteins Sequencing

Further identification was accomplished for sequencing of the SARS-CoV, MERS-CoV, and SARS-CoV-2 protein spike macromolecules using BIOVIA Discovery Studio 2020 [19] and Notepad++. Amino acid residues that play a role in the structure of spike protein macromolecules are then evaluated and explored.

2.4 Angiotensin-Converting Enzyme (ACE-2) Preparation

Macromolecules Angiotensin-Converting Enzyme (ACE-2) used in this study were obtained from Protein Data Bank (http://www.rcsb.org/pdb) with PDB ID 2AJF [21]. The preparation of this macromolecules was performed by removing natural ligands and water molecules using BIOVIA Discovery Studio 2020 [19].

2.5 Effect of Spike Protein Complexes against ACE-2 Binding

This simulation was accomplished between spike proteins of SARS-CoV, MERS-CoV, and SARS-CoV-2 against Angiotensin-Converting Enzyme (ACE-2) macromolecule. Complex types were selected as proteins with default clustering RMSD 4.0 Å. The representation of the Connolly dot surface of the molecule into different components including convex, concave, and flat patch was generated through the PatchDock algorithm [22]. PatchDock was optimized, refined, overhauled, and reselected the side chain interface from the top 10 candidate solutions. It also changes the orientation of the molecule relative by limiting flexibility in the side chains of the interacting surface and allowing the movements of small rigid-body. The suitability of the system was verified by visualization analysis using BIOVIA Discovery Studio 2020 [19].

3 Results and Discussion

The introduction of receptor macromolecular structures is the first step in recognizing the mechanism of viral infection against host cells and tissues. Increased binding affinity between the SARS-CoV spike protein and ACE-2 has been shown to correlate with increased virus transmission and disease severity in humans. In addition, the ability to involve ACE-2 from different animal species seems to reflect the host's susceptibility to SARS-CoV infection and facilitate the leap of viruses from animals to humans. SARS-CoV-2 uses ACE-2 as an entry receptor and recognizes it with an affinity similar to SARS-CoV and MERS-CoV.

In this study was performed identification, evaluation, and exploration of spike protein macromolecules from SARS-CoV, MERS-CoV, and SARS-CoV-2 to observe the structural characteristics of its components through computational study. Moreover, observations were also performed on the interactions between three spike proteins with ACE-2. Macromolecules that will be used in this study were prepared by removing water molecules and natural ligands using the BIOVIA Discovery Studio 2020. Macromolecular preparation was created to facilitate identification in subsequent procedures.

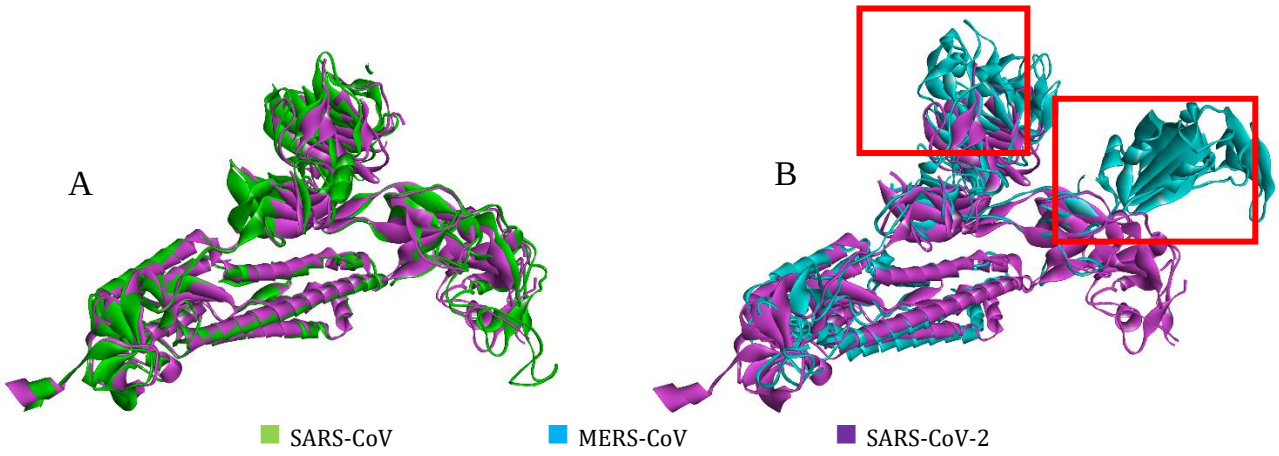

Figure 1. The three-dimensional conformation spike proteins overlay of (A) SARS-CoV against SARS-CoV-2 and (B) MERS-CoV against SARS-CoV-2

CPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLC
FSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATST
GNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGI
GYQPYRVVVLSFE

CPAGNSYTSFATYHTPATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEIL
EWFGITQTAQGVHLFSSRYVDLYGGNMFQFATLPVYDTIKYYSIIPHSIRSIQSDRKA
WAAFYVYKLQPLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFDVESGVYSVSSF
E

CPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCF
TNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDNYNYLY
RLFRKSNLKPFERDISTFPLQSYGFQPTNVGYQPYRVVVLSFE

Figure 2. The sequence of ACE-2 receptor binding domain (RBD) on the spike proteins of SARS-CoV (green), MERS-CoV (blue), and SARS-CoV-2 (purple)

Representations of spike protein macromolecules were visualized in the form of secondary structures to identify alpha-helix, beta-sheet, and loop sections as a whole. Figure 1 show that the protein spikes of SARS-CoV and SARS-CoV-2 have significant similarities. However, when compared to MERS-CoV, there are some differences, including the two sections located at the top (in the red box sections). Then the alpha-helix at the bottom of the SARS-CoV-2 spike protein macromolecule. This needs to be further proven by exploring the amino acid residues that make up the three spike protein macromolecules.

Interestingly, the sequences from the ACE-2 binding site of all spike proteins are relatively identical (Figure 2). The active site of SARS-CoV-2 was less that is only 180 amino acid residues compared to SARS-CoV and MERS-CoV. This phenomenon can be predicted that the protein spikes of SARS-CoV and MERS-CoV will bind more strongly on the surface of ACE-2. Nevertheless, the amino acid residue component forming receptor-binding domain (RBD) from SARS-CoV and SARS-CoV-2 have in common (in the yellow sections). There are 31 amino acid residues play an important role as RBD SARS-CoV and SARS-CoV-2, such as Cys336, Pro337, Phe338, Gly339, Glu340, Val341, Phe342, Asn343, Ala344, Thr345, Arg346, Phe347, Ala348, Ser349, Val350, Tyr351, Ala352, Trp353, Gly504, Tyr505, Gln506, Pro507, Tyr508, Arg509, Val510, Val511, Val512, Leu513, Ser514, Phe515, and Glu516. While MERS-CoV only 5 amino acid residues, including Cys336, Pro337, Ser514, Phe515, and Glu516.

Further identification was performed through the protein-protein docking methods. The binding affinity of ACE-2 against each spike proteins was evaluated by the atomic contact energy (ACE) score integrated into the PatchDock algorithm. The purpose of this docking simulations was to examine the effect of the active binding sites of spike proteins on the surface of ACE-2. The large area of RBD in spike proteins is predicted to be able to facilitate the entry of coronavirus into cells because of the ability of SARS-CoV, MERS-CoV, and SARS-CoV-2 to reach ACE-2 and forward the signal. Therefore, it is also necessary to explore amino acid residues that play a role in the formation of molecular interactions between the spike proteins of coronavirus and ACE-2.

Table 1. The affinity of each spike proteins against ACE-2

Spike Proteins	Atomic Contact Energy (kJ/mol)
SARS-CoV	−1066.08
MERS-CoV	−1042.31
SARS-CoV-2	−1341.85

The results of the protein-protein docking simulations in Table 1 show that all three protein-protein complexes have a negative ACE score. This phenomenon can be caused by molecular interactions that form between spike proteins and ACE-2, both hydrogen bonds, hydrophobic interactions, and electrostatic interactions. Interestingly, the spike protein of SARS-CoV-2 was able to bind strongly to the

surface of ACE-2 macromolecules with an ACE score of −1341.85 kJ/mol. In designing an inhibitor for COVID-19, a natural compound that can inhibit SARS-CoV-2 is not easily bound to the active site of ACE-2, and stabilizes the structure of the receptor macromolecules and prevents the conformational changes needed to forward signals further.

Interactions between SARS-CoV-2 and ACE-2 consist of 16 hydrogen bonds (with Arg440, Tyr450, Pro476, Asp477, Asn487, Tyr489, Tyr495, Gly496, Tyr498, Gly502, and Tyr505), 7 hydrophobic interactions (with Tyr454, Leu486, Tyr489, Tyr498, and Tyr505), and 3 electrostatic interactions (with Arg440 and Asp441). Moreover, the molecular interactions that were formed are dominated by hydrogen bonds, especially those that act as donor hydrogen bonds and hydrogen bond acceptors. Most hydrogen bonds between these proteins are quite strong, with average bond lengths under 3 Å. In addition, there are also hydrophobic interactions and electrostatic interactions that contribute. It can be predicted that these bonds and interactions play an important role in stabilizing protein complexes.

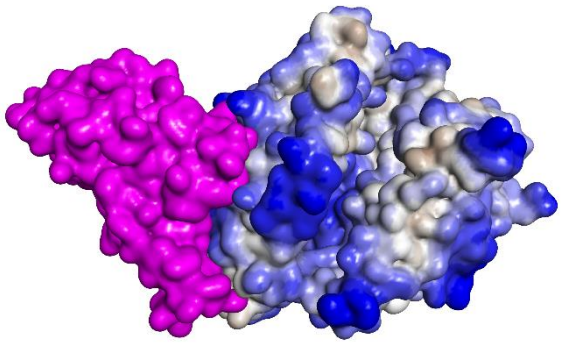

Figure 3. Protein-protein docking pose of SARS-CoV-2 spike proteins (purple) against ACE-2

Overall, it can be shown that the spike protein of SARS-CoV-2 binds to the non-polar patch of the ACE-2 surface (Figure 3). In developing COVID-19 inhibitor compounds, a candidate molecule that can prevent the attachment of SARS-CoV-2 to the active site is needed.

4 Conclusions

Based on research that has been done, the results show that there are some similarities between the structures of the spike protein macromolecules of several coronaviruses. SARS-CoV-2 spike protein has a fairly strong interaction with the binding site of ACE-2, with an ACE score of −1341.85 kJ/mol. Therefore, natural compounds that act as SARS-CoV-2 spike protein inhibitors are needed for the development of candidate molecules that can prevent and control infectious diseases of COVID-19.

5 Acknowledgments

The authors thank the LPPM (Institute for Research and Community Service), Universitas Islam Bandung, for the research financially supported by the Special Research Grant Program 2020, No.039/B.04/LPPM/IV/2020.

6 Conflicts of Interest

The authors declare that they have no conflicts of interest with the contents of this article.

7 Author Contributions

TMF conceived and designed the experiments, performed the experiments, collected the data, analyzed the data, and corrected manuscript; MLD performed the experiments, collected the data, and drafted the manuscript.

8 References

[1] Wang C, Horby PW, Hayden FG, Gao, GF, 2020. A novel coronavirus outbreak of global health concern, *Lancet*, **395**, (10223), 470-473.

[2] Richman DD, Whitley RJ, Hayden FG, 2016. *Clinical Virology, 4th ed*, Washington: ASM Press.

[3] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al, 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet*, **395**, (10223), 497-506.

[4] Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al, 2020. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany, *The New England Journal of Medicine*, **382**, (10), 970-971.

[5] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al, 2020. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia, *The New England Journal of Medicine*, **382**, (13), 1199-1207.
[6] Heymann DL, 2020. Data sharing and outbreaks: Best practice exemplified, *Lancet*, **395**, (10223), 469-470.
[7] Liu X, Wang XJ, 2020. Potential inhibitors for 2019-nCoV coronavirus M protease from clinically approved medicines, *Journal of Genetics and Genomics*, **47**, (2), 119-121.
[8] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al, 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature*, **579**, 270-273.
[9] Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al, 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding, *Lancet*, **6736**, (10224), 1-10.
[10] Letko M, Munster V, 2020. Functional assessment of cell entry and receptor usage for lineage B β-coronaviruses, including 2019-nCoV. *Nature Microbiology*, **5**, 562-569.
[11] Hoffmann M, Kleine-Weber H, Kruger N, Muller M, Drosten C, Pohlmann S, 2020. The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells.
[12] Graham RL, Becker MM, Eckerle LD, Bolles M, Denison MR, Baric RS, 2020. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease, *Nature Medicine*, **18**, (12), 1820-1826.
[13] Ng OW, Chia A, Tan AT, Jadi RS, Leong HN, Bertoletti A, et al, 2016. Memory T cell responses targeting the SARS coronavirus persist up to 11 years post-infection, *Vaccine*, **34**, (17), 2008-2014.
[14] Liu WJ, Zhao M, Liu K, Xu K, Wong G, Tan W, et al, 2017. T-cell immunity of SARS-CoV: Implications for vaccine development against MERS-CoV, *Antiviral Research*, **137**, 82-92.
[15] Kumar S, Maurya VK, Prasad AK, Bhatt MLB, Saxena SK, 2020. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV), *Virusdisease*, **31**, (3), 13-21.
[16] Walls AC, Xiong X, Park YJ, Tortorici MA, Snijder J, Quispe J, et al, 2019. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion, *Cell*, **176**, (5), 1026-1039.
[17] Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, et al, 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains, *Nature Communications*, **8**, 15092.
[18] Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D, 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein, *Cell*, **180**, (2), 1-12.
[19] Dassault Systemes BIOVIA, Discovery Studio Modeling Environment, Release 2020, Dassault Systemes: San Diego, CA, USA.
[20] Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE, 2006. Tools for Integrated Sequence-Structure Analysis with UCSF Chimera, *BMC Bioinformatics*, **7**, 339.
[21] Li F, Li W, Farzan M, Harrison SC, 2005. Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor, *Science*, **309**, (5742), 1864-1868.
[22] Prabhu DS, Rajeswari VD, 2016. In silico docking analysis of bioactive compounds from Chinese medicine Jinqi Jiangtang Tablet (JQJTT) using Patch Dock, *Journal of Chemical and Pharmaceutical Research*, **5**, (8), 15-21.