STRUCTURE-BASED VIRTUAL SCREENING OF INDONESIAN NATURAL PRODUCT COMPOUNDS AS EBOLA VIRUS VP30 PROTEIN INHIBITORS

Givan Andris Tio[1], Andrei Bernadette[1], Mochammad Arfin Fardiansyah Nasution[1], Puteri Aprilia Sitadevi[1], and Usman Sumo Friend Tambunan[1*]

[1]Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Indonesia

ABSTRACT: Indonesia has the second-highest biodiversity in the world. At least 9,600 out of 30,000 plant species exist in Indonesian tropical forests known to have medicinal properties. Hence lots of potentials still need to be explored including their abilities as an antiviral agent. Ebola virus (EBOV) continues as a major health threat worldwide with currently neither effective vaccine nor drug available. VP30 is one of the most important proteins for viral transcription activator of EBOV. Therefore, inhibiting this protein can be a viable choice for disturbing the life cycle of this virus. In this research, about 3,429 Indonesian natural product compounds were subjected into computational ADMET test using DataWarrior v4.7.2, while the molecular interaction and Gibbs free binding energy ($\Delta G_{rinding}$) value of the selected compounds were analyzed and calculated using MOE 2014.09 software. Finally, the oral bioavailability of the selected compounds was predicted using SwissADME software. Through this study, two compounds were selected to be potential VP30 inhibitors due to low $\Delta G_{rinding}$ value. They were acrylamide C and scoulerine, which have $\Delta G_{rinding}$ value of -9.7940 kcal/mol, and -7.3823 kcal/mol, respectively. Moreover, these two compounds did not possess any toxicity properties, and have high oral bioavailability, suggested it could highly be absorbed in the human body through oral administration. Thus, these compounds should be liable to be selected as the drug candidate of EBOV targeting VP30 and analyzed its antiviral activities further through molecular dynamics simulation and in vitro experiment.

Keywords: Indonesian natural products, EBOV, VP30, Molecular docking, ADMET test

1. INTRODUCTION

Ebola virus disease (EVD) is one of the fatal diseases around the globe, which infected no less than 28,642 people and killed 11,319 people globally [1]. Although extensive ongoing research has been performed, to date, there is no approved cure that has been affirmed for this malignant disease. EVD can be easily characterized by severe hemorrhagic fever, focal necrosis of the liver, kidney, and spleen, bleeding diathesis and sudden shock with a mortality rate of 90% [2]. EVD caused by Ebola virus (EBOV), a lipid-enveloped negatively stranded RNA virus affecting both human and non-human primates. There are five species of the virus been identified; include Zaire, Sudan, Ivory Coast, Bundibugyo and Reston EBOV with the highest fatality attributed to the Zaire species [3].

EBOV outbreak drew much attention to the drug discovery and development for this deadly pathogen. Currently, several approaches have been undergone to effectively combat EBOV, such as the development of peptides, application of monoclonal antibodies, small molecules inhibitors, recombinant DNA vectors as well as repurposing existing drugs [4], [5]. Therapeutics target and attack the virus at different stages of its life cycle, thereby halting virus replication and reducing destruction of the host immune system [6].

Due to the high mutable genome that resides in EBOV, the developments of EBOV drug are focused on the high-conserved regions. One of them is VP30, the transcription activator domain in EBOV. This virus had multiple functions, such as transcription activator and involved in nucleocapsid assembly. VP30 phosphorylation assumed to be a crucial regulatory factor in establishing if the protein is indeed involved in transcription or assembly of the virus. The N-terminal of EBOV VP30 presents with a Cys3-His zinc-binding domain native to M2-1 proteins of pneumoviruses and MARV VP30 [7]–[9].

Natural products have been considered as one of the potential lead compound sources because of their attractive bioactivities. Furthermore, many drugs have been sold on the market are either acquired or derived from natural sources compounds [10]. As the second-highest biodiversity country in the world, Indonesia has high potential to produce a new drug candidate from its natural sources, either from its plants or animals [11]. Thus, this research was aimed to obtain new drugs from Indonesian natural products for

combating EBOV by inhibiting in the VP30 protein through in silico drug design method.

2. RESEARCH METHODOLOGY

2.1 Tools and Materials

This study was conducted using both offline and online software such as Molecular Operating Environment (MOE) 2014.09 [12], DataWarrior v4.7.2 [13], SWISS-MODEL [14], and SwissADME [15]. Moreover, the structure of Indonesian Natural Products that used as ligands were obtained from Herbal Database (HerbalDB), which developed by the Faculty of Pharmacy, Universitas Indonesia. This database can be accessed through an online website at (http://herbaldb.farmasi.ui.ac.id/v3/index.php) [11]. This research was conducted based on the research pipeline that has been previously performed by our research group [16]–[18].

2.2 Preparation and Optimization of Indonesian Natural Products

The molecular structures of Indonesian Natural Products were retrieved from HerbalDB in .mol file format. Followed by the computational ADME-Tox test using DataWarrior v4.7.2 software. All Indonesian natural product ligands were screened based on their drug-likeness and their toxicity properties, such as mutagenic, tumorigenic, irritant, and reproductive effect. Then, the remaining ligands were prepared and optimized by using the default protocol on MOE 2014.09 software.

2.3 Preparation of EBOV VP30

In this research, EBOV VP30 sequence was obtained from the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/) in FASTA file format. Then, the 3D-structure of EBOV VP30 was searched by using SWISS-MODEL then downloaded from the Research Collaboratory for Structural Bioinformatics - Protein Data Bank (RSCB-PDB) (http://rcsb.org/pdb/home/home.do) in .pdb file format. The downloaded structure was prepared and optimized by using 'LigX' protocol on MOE 2014.09 software, which slight modification has been made compared to the default protocol.

2.4 Molecular docking simulations

The docking simulations of EBOV VP30 protein was performed by using MOE 2014.09 software. Molecular docking simulations were divided into four stages. First, the prepared ligands underwent rigid docking with 30-1 retains. Then, these ligands were selected based on RMSD (Root Mean Square Deviation) and free binding energy calculation ($\Delta G_{rinding}$) values. Second, the selected ligands underwent rigid docking with 100-1 retains. Third, 100-1 retains flexible docking protocol was performed to decrease the number of selected ligands. Finally, these selected ligands underwent flexible docking with 300-3 retains.

2.5 Bioavailability and Pharmacokinetics Predictions

SwissADME online web service was utilized to predict the bioavailability and pharmacokinetic properties of the selected Indonesian natural products that have remained after the docking simulation. In the end, about two ligands that have the best properties were selected as best drug candidates for treating Ebola through the inhibition of EBOV VP30.

3. RESULTS AND DISCUSSIONS

3.1 Pre-docking Simulation

In this study, the Indonesian natural products molecular structures were retrieved from HerbalDB in .mol file format. The compounds itself were obtained through the literature study that mentioned any plants, fungi, or microbes that resides or can be found exclusively in Indonesia according to literature and extensive study [11]. Throughout this method, about 3,429 Indonesian natural product compounds were retrieved from this database. This step was followed up by a computational ADME-Tox test to determine their drug-likeness and toxicity properties. Any toxicity properties that owned by any ligand may lower their efficiency and effectivity when used as the drug. Hence, this step was compulsory to eliminate any compounds that have a low probability as drug candidates. According to the results from this step, about 3,128 Indonesian natural product compounds were omitted through this test because they possessed either toxicity properties, such as mutagenic, tumorigenic, reproductive effect, and irritant properties, or have a low possibility to be developed as a drug, according to their drug-likeness score. Thus, only 301 ligands have passed this stage and can be prepared further in the ligand optimization processes.

In this study, the optimization of the ligand structure was utterly achieved using MOE 2014.09 software. This step was conducted to obtain the ligand pose that resembles the molecular structure when treated in the real condition, especially in the human body. At first, the remaining 301 Indonesian natural product compounds, which saved earlier in .sdf file format, were opened and prepared

according to the default protocols in the respective software, including the 'Wash', 'Partial Charge' and 'Energy Minimization' features in Database Viewer, MOE 2014.09. The prepared ligands were later saved in the .mdb file format for the docking simulation processes.

The 3D structure of EBOV VP30 that used for the research was listed with PDB ID: 2I8B [19], which obtained from RSCB-PDB. The corresponding 3D structure was prepared using MOE 2014.09 software as well. First, the forcefield that used for the protein preparation was selected. In this research, Amber10: EHT with Gas Phase solvation was used. After that, EBOV VP30 was prepared and optimized using 'LigX', which was featured on MOE 2014.09. The preparation was done by removing the water molecules and tether the receptors and the optimization was done by fixing the atoms positions and hydrogen that will interact with EBOV VP30 position with the ligands. The optimized EBOV VP30 protein structure was later saved in .moe file format to retain its forcefield and structure, which can be used later on for the docking simulation.

3.2 Molecular Docking Simulation of EBOV VP30 and Indonesian Natural Products

In this study, the determination of binding sites of VP30 protein was obtained through 'Site Finder' feature in MOE 2014.09 software by searching the favorable binding sites based on several parameters, such as solvent's exposures and lipophilicity of the binding sites. The result shows that the binding sites of VP30 protein are Leu144, Arg179, Lys180, Phe181, Ser182, Lys183, Ser184, Gln185, Leu188, His215, Leu249, Pro250, Cys251, Glu252, and Ser253. These binding sites were selected due to their resemblance to the previous research [20], [21]. Thus, the molecular docking simulation would be performed on these binding sites.

The docking simulation of EBOV VP30 protein was done by using MOE 2014.09 software. The inhibition of EBOV VP30 protein can be achieved if the ligands capable of binding to a target receptor of the VP30 binding site. In this study, 301 compounds of Indonesian natural product, which already underwent ADME-Tox test, acted as "ligands" to inhibit the EBOV VP30 protein. These ligands went into molecular docking simulation. The simulation was done to determine the free binding energy calculations and the number of molecular interactions between the complexes of ligand and protein. $\Delta G_{rinding}$ is associated with the binding affinity between ligand and receptor and can be determined by the stability of ligand-protein complexes. To find a suitable candidate ligand for inhibiting the target receptor of EBOV VP30, that ligand must have a higher Gibbs free binding energy

and better molecular interaction with the binding site of EBOV VP30 pocket regarding binding energy and molecular interaction. The free binding energy value of the ligands was expected to have a negative value and lower than the standard ligands. So, this ligand can be applied as a novel compound for inhibiting EBOV VP30 protein [17].

Table 1 The docking results of EBOV VP30 and the best Indonesian natural product compounds, along with three standard ligands

No.	Ligand Code (Compound Name)	$\Delta G_{binding}$ value (RMSD (Å))
1.	IDNP-268 (Aglamide C)	-9.7940 kcal/mol (1.4500)
2.	IDNP-204 (3,10,11-Trihydroxydibenzo[b,e]oxonine-7,13(6H,8H)-dione)	-8.2686 kcal/mol (1.6453)
3.	IDNP-70 (Epiafzelechin)	-7.8768 kcal/mol (1.0326)
4.	IDNP-123 (Scoulerine)	-7.3823 kcal/mol (1.0382)
S1	Galidesivir / BCX4430	-7.3054 kcal/mol (0.9212)
S2	Gossypetin	-7.1857 kcal/mol (1.2655)
S3	Taxifolin	-7.8649 kcal/mol (0.8521)

In this study, BCX4430, gossypetin, and taxifolin were selected as the standard ligands. The first ligand was chosen because it has been utilized as a drug for combating Ebola through in vitro and in vivo experiments, while the latter two ligands were chosen because they have the highest binding affinity among all tested ligand through in silico experiment that performed in the previous research [22], [23]. According to the docking results from this research, these ligands have a $\Delta G_{rinding}$ value of -7.3054, -7.1857 and -7.8649 kcal/mol, respectively. Furthermore, the docking simulation results revealed about four ligands have better molecular interaction, and some of them have lower $\Delta G_{rinding}$ value compared to the standard ligands, as shown in Table 1. In this study, aglamide C was determined as the Indonesian natural product compounds which have the lowest $\Delta G_{rinding}$ value, sitting at -9.7940 kcal/mol. In addition, IDNP-206 (IUPAC: 3,10,11-trihydroxydibenzo[b,e]oxonine-7,13(6H,8H)-dione), epiafzelechin and scoulerine also have $\Delta G_{rinding}$ value as well, respectively at -8.2686, -7.8768, and -7.3823 kcal/mol. Except for scoulerine,

which has lower binding affinity than taxifolin, all of these ligands possessed higher binding affinity on EBOV VP30 protein on its binding site. Thus, these compounds, regarding Gibbs free binding energy, are suitable to be progressed as Ebola drug candidates.
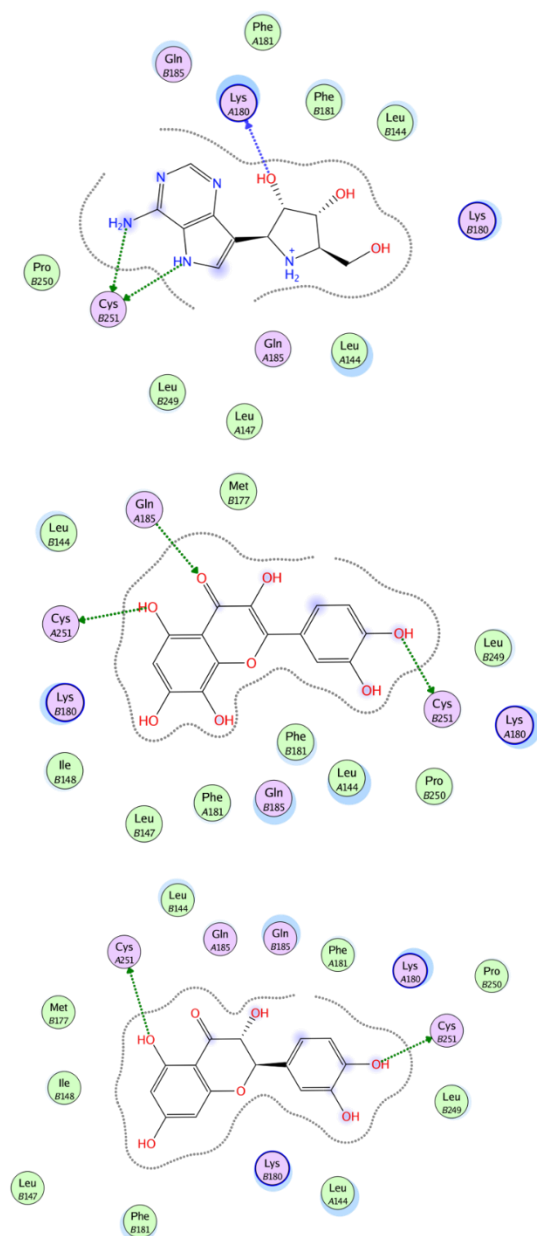

Fig. 1 The molecular interaction of BCX4430 (top), gossypetin (middle) and taxifolin (bottom) in the binding site of EBOV VP30.

In addition to free binding energy, the molecular interaction between the ligands and the binding site of EBOV VP30 protein is another important aspect of determining whether the ligand can be considered as a suitable inhibitor or not.

First, the molecular interactions of standard ligands were observed. As is displayed in Fig. 1. BCX4430 formed a molecular interaction with

Lys180 via hydrogen bonds through its backbone, with other two molecular interaction with Cys251 were observed that act as sidechain donor and several other interactions within the binding pocket through van der Waals interaction. Moreover, gossypetin formed a molecular interaction with Gln185 as a sidechain acceptor, with two molecular interaction with Cys251 that act as sidechain donor. Finally, taxifolin has two molecular interaction with Cys251 as a sidechain donor.
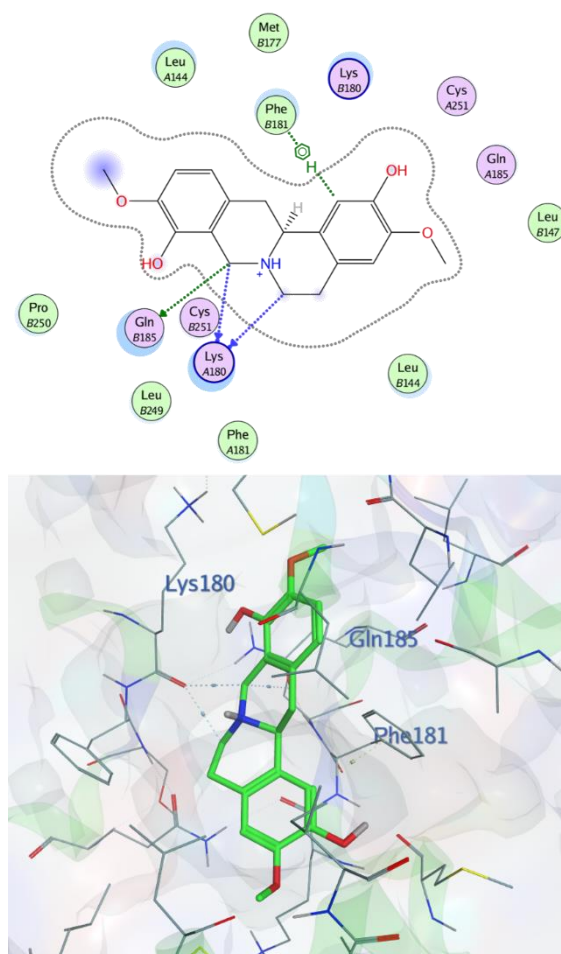

Fig. 2 The two-dimensional and three-dimensional molecular interaction of scoulerine in the binding site of EBOV VP30.

In this research, the molecular interaction between the Indonesian natural product compounds and EBOV VP30 were also compared as well. Even though scoulerine has the highest Gibbs free binding energy compared to the other four ligands, but this compound has better molecular interaction within the EBOV VP30 binding pocket. Scoulerine has four interesting molecular interactions formed in EBOV VP30 binding site, such as arene-H interaction with Phe181, hydrogen bonds with the side chain of Gln185, and two hydrogen bonds interaction with the backbone of Lys180. Other Indonesian natural product ligands such as IDNP-

206, aglamide C, and epiafzelechin only have one to two interactions within the binding pocket, which is not preferred, even though they have better Gibbs free binding energy. Thus, scoulerine has a better molecular interaction compared to any other Indonesian natural product compounds. The molecular interactions of scoulerine in the binding site of EBOV VP30 is displayed in Fig. 2.

3.3 Results of Bioavailability and Pharmacokinetics Predictions

In this study, the remaining four Indonesian natural product compounds, along with the three standard compounds, were subjected to bioavailability and pharmacokinetics prediction test. This test was done to measure the pharmacokinetic properties of the ligands, as well as the oral bioavailability of the selected ligands according to several parameters such as physicochemical properties (e.g., molecular weights, TPSA) that have been predicted through this test. This prediction was entirely made by using SwissADME online web service, which can be accessed through this website http://www.swissadme.ch/. Also, DataWarrior v4.7.2 software was also utilized to

measure the drug-likeness score and molecular properties that also determine the oral bioavailability of the compound.

Table 2 shows the results of the drug-likeness and molecular properties prediction tests. From the result, all of the Indonesian natural product ligands have a molecular weight less than 500 Daltons, logP less than 5.0, hydrogen bond acceptor less than 10 and hydrogen bond donor less than 5, which obey Lipinski's Rule of Five [24]. This means that all of the Indonesian natural product ligands have likable properties to be absorbed and administrated through the oral system in the human body. However, according to the same prediction test, only two out of four Indonesian natural product compounds, namely IDNP-206 and aglamide C, that has the positive drug-likeness value, which indicates these ligands have a likeness to become the drug compounds than any other ligands, based on their molecular fragments.

Also, the oral bioavailability prediction of Indonesian natural product compound was also conducted as well along with the molecular properties prediction using the same web service.

Table 2 The physicochemical properties of the selected Indonesian natural products and the standard ligands

Ligand	Drug-likeness					
	MW	clogP	H-acceptor	H-donor	TPSA (Å^2)	Drug-likeness
3,10,11-Trihydroxydibenzo[b,e]oxonine-7,13(6H,8H)-dione	300.26	1.16	6	3	104.06	0.29065
Scoulerine	328.38	3.07	4	3	63.36	-1.4498
Aglamide C	300.40	2.86	2	1	49.41	2.9194
Epiafzelechin	274.27	1.51	5	4	90.15	0.3153
BCX4430	266.28	0.12	5	6	144.89	-0.2121
Gossypetin	318.24	1.33	8	6	151.59	-0.0083
Taxifolin	304.25	0.71	7	5	127.45	0.4448

Table 3 The oral bioavailability prediction of the selected Indonesian natural products and the standard ligands

Ligands	GI	Lipinski's	Ghose's	Veber's	Egan's	Bioavailability Score
3,10,11-Trihydroxydibenzo[b,e]oxonine-7,13(6H,8H)-dione	High	0	0	0	0	0.55
Scoulerine	High	0	0	0	0	0.55
Aglamide C	High	0	0	0	0	0.55
Epiafzelechin	High	0	0	0	0	0.55
BCX4430	Low	1	1	1	1	0.55
Gossypetin	Low	1	0	1	1	0.55
Taxifolin	High	0	0	0	0	0.55

The result can be seen in Table 3. According to the results, BCX4430 and gossypetin have low oral bioavailabilities compared to other ligands, these mainly due to the number of H-bond donor (six H-bond donors in both compounds) which existed in the molecules, which violating the Lipinski's RO5. Additionally, both molecules have high TPSA values, which violating Veber's Rule. Surprisingly, four Indonesian natural product compounds (IDNP-206, scoulerine, aglamide C and epiafzelchin), including taxifolin as the standard ligand, did not break any of these rules, hence these ligands have higher gastrointestinal absorption than BCX4430 and gossypetin.

4. CONCLUSION

In this research, about 3,429 Indonesian natural product compounds were subjected into the computational ADME-Tox screening test, molecular docking simulation, and bioavailability prediction test to determine the best compounds that can be utilized as the EBOV VP30 inhibitors. In the end, after the series of docking simulations and ADME-Tox prediction test, two Indonesian natural product compounds, namely, scoulerine and aglamide C was selected as the best Indonesian natural product compounds to inhibit EBOV VP30 due to their binding affinity, in terms of Gibbs free binding energies and molecular interactions, towards the binding site of EBOV VP30, lacks of toxicity properties and high oral bioavailability through gastrointestinal system. Thus, these compounds can be selected as the drug candidate of EVD targeting EBOV VP30 after the series of molecular dynamics simulations were performed to determine the stability of both EBOV VP30-scoulerine and EBOV VP30-aglamide C complexes under real environments.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] CDC, Ebola (Ebola Virus Disease), 2014.

[2] De Clercq E., Ebola virus (EBOV) infection: Therapeutic strategies, Biochem. Pharmacol., Vol. 93, No. 1, pp. 1–10, 2015.

[3] Gebre Y., Gebre T., and Peters A., The Ebola virus: a review of progress and development in research, Asian Pac. J. Trop. Biomed., Vol. 4, No. 12, pp. 928–936, 2014.

[4] Haque A., Hober D., and Blondiaux J., Addressing therapeutic options for Ebola virus infection in current and future outbreaks, Antimicrob. Agents Chemother., Vol. 59, No. 10, pp. 5892–5902, 2015.

[5] Reward S. and Mirdha D., Transmission of Ebola virus disease: An overview, Annals of Global Health, Vol. 80, No. 6. pp. 444–451, 2014.

[6] Nyakatura E. K., Frei J. C., and Lai J. R., Chemical and Structural Aspects of Ebola Virus Entry Inhibitors, ACS Infectious Diseases, Vol. 1, No. 1. pp. 42–52, 2015.

[7] Balmith M., Faya M., and Solaiman M. E. S., Ebola virus: A gap in drug design and discovery - experimental and computational perspective, Chem. Biol. Drug Des., Vol. 89, No. 3, pp. 297–308, 2017.

[8] Feldmann H. and Geisbert T. W., Ebola haemorrhagic fever, Lancet, Vol. 377, No. 9768, pp. 849–862, 2011.

[9] Groseth A. et al., The Ebola virus ribonucleoprotein complex: A novel VP30-L interaction identified, Virus Res., Vol. 140, No. 1–2, pp. 8–14, 2009.

[10] Patrick G. L., Introduction of Medicinal Chemistry, 5th Ed. Oxford: Oxford University Press, 2013.

[11] Yanuar A., Mun'im A., Lagho A. B. A., Syahdi R. R., Rahmat M., and Suhartanto H., Medicinal Plants Database and Three Dimensional Structure of the Chemical Compounds from Medicinal Plants in Indonesia, Int. J. Comput. Sci., Vol. 8, No. 5, pp. 180–183, Nov. 2011.

[12] Vilar S., Cozza G., and Moro S., Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery., Curr. Top. Med. Chem., Vol. 8, No. 18, pp. 1555–1572, Jan. 2008.

[13] Sander T., Freyss J., Von Korff M., and Rufener C., DataWarrior: An open-source program for chemistry aware data visualization and analysis, J. Chem. Inf. Model., Vol. 55, No. 2, pp. 460–473, 2015.

[14] Biasini M. et al., SWISS-MODEL: modeling protein tertiary and quaternary structure using evolutionary information., Nucleic Acids Res., Vol. 42, No. Web Server Issue, pp. W252--8, Jul. 2014.

[15] Daina A., Michielin O., and Zoete V., SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci.

Rep., Vol. 7, 2017.

[16] Tambunan U. S. F., Zahroh H., Parikesit A. A., Idrus S., and Kerami D., Screening analogs of the β-OG pocket binder as fusion inhibitor of dengue virus 2, Drug Target Insights, Vol. 9, pp. 33–49, 2015.

[17] Tambunan U. S. F., Parikesit A. A., Adam V. C., Nasution M. A. F., Puspitasari R. D., and Kerami D., Virtual Screening of Commercial Cyclic Peptides as β--OG Pocket Binder Inhibitor in Dengue Virus Serotype 2, Int. J. GEOMATE, Vol. 13, No. 37, pp. 60–68, 2017.

[18] Tambunan U. S. F. and Nasution M. A. F., Identification of novel Ebola virus (EBOV) VP24 inhibitor from Indonesian natural products through in silico drug design approach, AIP Conf. Proc., Vol. 030091, pp. 1–9, 2017.

[19] Hartlieb B., Muziol T., Weissenhorn W., and Becker S., the Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association, Proc. Natl. Acad. Sci., Vol. 104, No. 2, pp. 624–629, 2007.

[20] Balmith M. and Soliman M. E. S., Potential Ebola drug targets – filling the gap: a critical step forward towards the design and discovery of potential drugs, Biologia (Bratisl)., Vol. 72, No. 1, pp. 1–13, 2017.

[21] Setlur A. S., Naik S. Y., and Skariyachan S., Herbal Lead as Ideal Bioactive Compounds Against Probable Drug Targets of Ebola Virus in Comparison with Known Chemical Analogue: A Computational Drug Discovery Perspective., Interdiscip. Sci., 2016.

[22] Janeba Z., Development of Small-Molecule Antivirals for Ebola, Med. Res. Rev., Vol. 35, No. 6, pp. 1175–1194, 2015.

[23] Raj U. and Varadwaj P. K., Flavonoids as Multi-target Inhibitors for Proteins Associated with Ebola Virus: In Silico Discovery Using Virtual Screening and Molecular Docking Studies, Interdiscip. Sci. Comput. Life Sci., Vol. 8, No. 2, pp. 132–141, 2016.

[24] C.A. Lipinski P. J., Lombardo F., Dominy B.W., and Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting, Adv. Drug Deliv. Rev., Vol. 64, pp. 4–17, 2012.