VIRTUAL SCREENING OF INDONESIAN HERBAL DATABASE FOR DNA METHYLTRANSFERASE INHIBITORS

I GUSTI NGURAH RAKA BHASKARA WILAPUTRAKA, AZMINAH AZMINAH, LINDA ERLINA, REZI RIADHI SYAHDI, ARRY YANUAR*

Department of Pharmacy, Faculty of Pharmacy, Universitas Indonesia, Depok, Indonesia. Email: arry.yanuar@ui.ac.id

ABSTRACT

Objective: DNA hypermethylation is an abnormal epigenetic process catalyzed by DNA methyltransferase 1 (DNMT1). It is also one of the factors that cause non-communicable diseases such as cancer, diabetes, and other metabolic diseases. DNA hypermethylation can be reversed by suppressing DNMT1 activity using a DNMT inhibitor. This study was conducted to seek out inhibitor candidates among natural products.

Methods: The search for potential inhibitors was conducted through a virtual screening of the Indonesian Herbal Database using AutoDockVina as docking software. Twenty-five compounds known for their inhibitory activity against DNMT1 were used as actives and as a reference for generating decoys, which was done using the Directory of Useful Decoys, Enhanced.

Results: The 12 compounds with binding energies below the cutoff value were cassiamin C (A1), procyanidin B2 (B2), ent-epicatechin-(4alpha->8)-ent-epicatechin (C3), epicatechin-(4beta->8)-epicatechin-3-O-gallate (D4), neorhusflavanone (E5), 3-O-galloylepicatechin-(4beta->6)-epicatechin-3-O-gallate (F6), withanolide (G7), 3-O-galloylepigallocatechin-(4beta->6)-epigallocatechin-3-O-gallate (H8), cyanidin 3-(6''-caffeylsophoroside)-5-glucoside (I9), epifriedelanol (J10), gallocatechin-(4alpha->8)-epicatechin (K11), and scutellarein 7-glucosyl-(1->4)-rhamnoside (L12). A1 had the lowest binding energy of −12.7 kcal/mol, whereas K11 had the highest of −11.5 kcal/mol.

Conclusions: The virtual screening yielded five potential DNMT1 inhibitors: Procyanidin B2, ent-epicatechin-(4alpha->8)-ent-epicatechin, epicatechin-(4beta->8)-epicatechin-3-O-gallate, neorhusflavanone, and cyanidin 3-(6''-caffeylsophoroside)-5-glucoside.

Keywords: Epigenetic, DNA methyltransferase inhibitor, Indonesian Herbal Database, Virtual screening, AutoDockVina.

INTRODUCTION

Epigenetic modification is one of the factors that cause non-communicable diseases. Living environment, lifestyle, exposure to toxic chemicals, and nutrition may influence an individual's epigenetic profile. Epigenetic modification can alter gene expression [1]. One type of epigenetic modification is DNA methylation. DNA methylation, catalyzed by DNA methyltransferases (DNMTs), is the addition of a methyl (CH3) group to the DNA strand, specifically to the fifth carbon atom of a cytosine ring. Methylated cytosine, called 5-methylcytosine, usually occurs to a cytosine base that lies adjacent to a guanine base, at what is called a CpG site. The reason for this name is that cytosine is bonded to guanine through a phosphate bond. These sites normally occur in clusters; an area rich in CpG sites is called a CpG island. CpG islands are usually located near a gene regulator such as a promoter. DNA methylation that affects a gene promoter may repress the transcription of said gene and ultimately its expression. This process of turning a gene on and off through methylation is part of normal eukaryotic cell function [2,3].

There are several types of DNMTs. Based on their functions, they can be grouped into two categories: *De novo* DNMTs, which form new methylated cytosine, and maintenance DNMTs, which maintain existing epigenetic patterns from one generation of cells to its offspring. In mammalian cells, there are three active DNMTs: DNMT1, DNMT3L, and DNMT3A/3B. DNMT1 is a maintenance DNMT, whereas DNMT3A/3B and DNMT3L are *de novo* DNMTs [4]. DNMT uses S-adenosyl-L-methionine as a methyl donor and produces S-adenosyl-L-homocysteine (SAH) [5]. SAH is a strong endogenous inhibitor of DNMT. Hypermethylation is an aberrant methylation that occurs to normally unmethylated areas of the genetic sequence. Hypermethylation of

CpG islands in the promoter region of a gene may cause inappropriate gene silencing and may affect such genes as those responsible for tumor suppression. It is, therefore, associated with non-communicable diseases such as cancer and diabetes [6]. One approach to recovering the function of a silenced gene is inhibiting the enzyme that catalyzes the methylation process, that is, DNMT. DNMT inhibitors (DNMTi's) exist as nucleoside analogs and non-nucleoside analogs [7].

The development of non-communicable disease therapies that target the underlying epigenetic factors is paramount. Most of the existing therapies do not, and the limited number that do usually have severe side effects. These side effects arise because the existing epigenetic therapies use nucleoside analogs, which have similar structures to nucleosides and therefore interfere with normal cell function. The recent studies have focused on the finding non-nucleoside analog inhibitor candidates from natural sources [8,9]. Natural products are promising non-nucleoside DNMTi candidates because of their relatively low toxicity and high structural diversity [10,11]. The potential of several natural products has been studied; these products have included curcumin [12], epigallocatechin (EGCG) [13], genistein [14], caffeic acid [15], psammaplin A [16], mahanine [17], and laccaic acid [16]. However, the inhibitory effects of these natural compounds are relatively low [11]. Therefore, a search for more effective DNMTi's from natural compounds is necessary. Indonesia is recognized for its biodiversity, with at least 9600 species known to possess pharmacological activity [18,19], so a search for a DNMTi is promising. A previous study conducted by Yanuar *et al.* [20] collected three-dimensional data on several Indonesian herbal compounds and compiled a database, published as the Indonesian Herbal Database (HerbalDB). In this study, virtual screening was employed as an *in silico*

method of structure-based drug design. The virtual screening was done on the HerbalDB using AutoDock and AutoDockVina in the PyRx software [21-23]. The software was validated using decoys generated by the Directory of Useful Decoys, Enhanced (DUD-E) [24].

METHODS

This study was conducted using literature and virtual screening by molecular docking (structure-based virtual screening). The materials and methods described in this study were modified from the studies by Yanuar *et al.*, Syahdi *et al.*, and Maldonado-Rojas *et al.* [11,25,26]. A three-dimensional structure of DNMT1, positive and negative control compounds, and active herbal compounds contained in the HerbalDB were used in this study.

Preparing the target protein structure

The structure of the docking target, DNMT1, was obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank website at www.rcsb.org/pdb/home [27]. Several inclusion and exclusion criteria were used to select the DNMT structure. The selected structure was required to be present in humans (homo sapiens), include a complete amino acid sequence at the active site, and be in complex with a ligand. The structure could not be a mutant protein or have a resolution of more than 2.5 Å. UniProt, InterPro, and protein model portal were used to compare the sequences and lengths of the amino acid chains [28-30]. The selected structure was then downloaded and optimized. The optimization process included cleaning the structure of solutes and all chemical compounds other than the co-crystal ligand and the target protein, separating the protein from the co-crystal ligand, adding missing hydrogens and charges, and minimizing energy. The partial charge of the atoms was calculated using the Gasteiger-Marsili charge algorithm while the force field calculation and minimization were done using the MMFF 94 method [31,32]. The entire optimization process was done using the Python molecular viewer embedded into AutoDock Tools. The coordinates of the center of the search space were then determined based on the binding space occupied by the co-crystal ligand.

Preparing the ligand structure

This study used three groups of ligands: Actives, decoys, and HerbalDB. The actives were positive controls or compounds that showed inhibitory activity against DNMT1 through *in vitro* or *in vivo* tests. Based on the previous research, 25 compounds were chosen as actives [11-17]. Their structures were obtained in SMILES format from the PubChem website at pubchem.ncbi.nlm.nih.gov [33]. These compounds are listed in Table 1. The second ligand group consisted of negative controls, chemical structures generated by the DUD-E at http://dude.docking.org based on the chemical structures of the 25 active compounds. For the virtual screening parameter optimization process, the libraries of actives and decoys were combined. The third ligand library contained

Table 1: 25 positive control compounds included in the actives library

Actives	
5,6-dihydro-5-azacytidine	**Laccaic acid A**
5-azacytidine	Mahanine
5-fluoro-2′-deoxycytidine	Nanaomycin A
Aza-AdoMet	Parthenolide
Caffeic acid	Procainamide
Chlorogenic acid	Procaine
Curcumin	Psammaplin A
Δ2-isoxazoline	RG 108
Decitabine	SAH
Deoxycytidine	SGI-110
EGCG	SGI-1027
Genistein	SFG
Hydralazine	

EGCG: Epigallocatechin, SFG: Sinefungin, SAH: S-adenosyl-L-homocysteine

1550 compounds, and all active metabolites found in the HerbalDB. The structures in all three groups were then converted to the mol2 file format using Open Babel [34].

Validating the docking method

The validation of AutoDock 4.2 and AutoDockVina in PyRx was done by redocking the co-crystal ligand of the selected target protein. The varied parameters were grid center, grid spacing, and number of grid points, which in turn affected the binding energy, inhibition constant, and root-mean-square deviation (RMSD) of the docking results. The best pose was selected by evaluating the lowest binding energy and the lowest RMSD value. The protein-ligand interaction of the resulting pose was then compared to that of the co-crystals original pose and analyzed. Validation of both docking programs was also done on the actives library through docking to the target protein. The resulting binding energies were then used as references to determine the binding energy cutoff in the virtual screening process.

Optimizing the virtual screening parameters

The parameters for virtual screening were optimized by docking the active and decoy libraries to the target protein. The results were then calculated and analyzed through a receiver operating characteristic (ROC) curve and enrichment factors (EF) to determine the optimum parameters for each program, that is, the parameters that accurately discriminated actives from decoys without compromising too much accuracy or sensitivity.

Virtual screening of the HerbalDB

The virtual screening was performed on all active compounds in the HerbalDB using AutoDockVina in PyRx. The optimum parameters obtained from the previous step were applied to the docking software. The screening results were ranked by their binding affinity values, and the actives' lowest binding energy from the docking validation step was used as the cutoff value to determine the hit compounds. Compounds with binding affinities lower than this cutoff value were considered hits. 12 compounds met this criterion and were analyzed further.

Analyzing target-ligand interaction

The ligand poses, resulting from the virtual screening, were then visualized using PyMOL and LigPlot+ and compared with the pose of the co-crystal ligand [35]. In the interaction between specific amino acids, the ligands, and the respective binding space, each ligand occupied was then analyzed and compared to that of the co-crystal ligand.

RESULTS AND DISCUSSION

The DNMT1 structure that met all the inclusion and exclusion criteria was therefore selected as the target protein was 3SWR. This structure is in complex with sinefungin (SFG). The ligand binding site was determined by the position of this co-crystal ligand. The coordinates were −5.235, −0.522, and −32.228 for the x-, y-, and z-axis, respectively. Analysis of the target-ligand interaction between SFG and 3SWR before the structure was optimized showed that the amino acids that interacted with DNMT1 were Phe1145, Leu1151, Glu1168, Met1169, Cys1191, Glu1266, and Val1580. SAH, which has one chemical group different from SFG, occupied the same hydrophobic pocket and showed the same amino acid interactions. Other research into inhibiting DNMT1 with SFG has shown similar interactions with these amino acids. The best redocking run of SFG to 3SWR with AutoDock yielded an RMSD value of 2.704 and a ΔG of −8.92 kcal/mol. The parameters used for this run were 60 × 60 × 60 grid points with a 0.375 Å space between points (Fig. 1).

The docking of the actives library to the target protein showed that there were natural, non-nucleoside analog compounds with lower binding energies than SAH, which was also included in this library. Some of these compounds were laccaic acid A, EGCG, mahanine, genistein, parthenolide, curcumin, and chlorogenic acid. From this result, it can be surmised that a virtual screening process would detect non-nucleoside analog compounds with better binding affinities than the endogenous compound (SAH). The compounds that produced the lowest binding

Table 2: Top 12 compounds from virtual screening with AutoDockVina in PyRx

Rank	Code	HerbalDB code	Ligand	ΔG (kcal/mol)
1	A	M00002800	Cassiamin C	−12.7
2	B	M00009077	Procyanidin B2	−12.3
3	C	M00009080	Ent-epicatechin-(4alpha->8)-ent-epicatechin	−12
4	D	M00002917	Epicatechin-(4beta->8)-epicatechin-3-O-gallate (D4)	−11.9
5	E	M00006467	Neorhusflavanone	−11.9
6	F	M00009207	3-O-galloylepicatechin-(4beta->6)-epicatechin-3-O-gallate	−11.7
7	G	M00032512	Withanolide	−11.6
8	H	M00009235	3-O-galloylepigallocatechin-(4beta->6)-epigallocatechin-3-O-gallate	−11.5
9	I	M00006843	Cyanidin 3-(6''-caffeylsophoroside)-5-glucoside	−11.5
10	J	M00030190	Epifriedelanol	−11.5
11	K	M00009112	Gallocatechin-(4alpha->8)-epicatechin	−11.5
12	L	M00004223	Scutellarein 7-glucosyl-(1->4)-rhamnoside	−11.5

HerbalDB: Herbal Database

Table 3: Target-ligand interactions of reference and virtual screening results

Amino acids	Reference			VS result				
	SFG co-crystal	SFG re-docked	SAH	B2	C3	D4	E5	I9
Glu698	-	-	-	D	-	-	D	D
Ala699	-	-	-	-	-	-	✓	A
Asp700	-	-	-	-	A	-	-	-
Phe1145	✓	✓	✓	✓	✓	✓	✓	D
Ser1146	-	-	O-H	-	D	-	-	-
Cys1148	-	-	-	-	-	A	-	-
Gly1149	-	-	-	-	-	A	-	-
Gly1150	-	-	-	-	-	A	-	-
Leu1151	H-O	H-O	H-O	-	-	A	-	-
Ile1167	-	-	-	-	-	-	-	D
Glu1168	O-H	O-H	O-H	D	-	-	D	-
Met1169	H-N	✓	✓	✓	A	✓	✓	✓
Trp1170	-	-	-	-	-	✓	-	✓
Ala1173	-	-	-	-	-	✓	-	-
Glu1189	-	-	-	D	-	-	D	-
Cys1191	H-N	H-N	H-N	✓	A	A	-	-
Gly1223	-	-	-	✓	-	-	-	-
Pro1224	-	-	-	-	-	D	-	-
Cys1226	-	-	-	-	-	-	-	A
Phe1247	-	-	-	-	✓	✓	✓	-
Glu1266	O-H	✓	-	D	D	-	D	-
Asn1267	-	-	-	-	-	D	D	D
Ala1579	-	-	-	-	✓	-	-	-
Val1580	H-O	✓	H-O	-	-	-	-	-
Asn1587	-	-	-	D	-	-	-	-
Val1768	-	-	-	-	-	-	D	-

A=H acceptor, D=H donor, ✓: Unspecific bond or hydrophobic interaction, SFG: Sinefungin, SAH: S-adenosyl-L-homocysteine
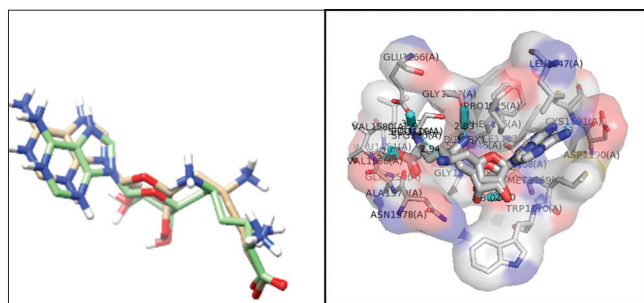

Fig. 1: Left: Superposition of re-docked sinefungin (SFG) (green) and SFG (brown). Right: Binding space of re-docked SFG and SFG

energies were SGI-1027 and laccaic acid A, with ΔG values of −9.04 and −11.1 kcal/mol for AutoDock and AutoDockVina, respectively. These values were then used as the cutoff to determine hits.

The optimization of the virtual screening parameters in PyRx showed that only AutoDockVina produced an acceptable validity. The grid box size used in AutoDock was 55 × 55 × 55, while in AutoDockVina, the grid box size varied from 70 × 70 × 70 through 35 × 35 × 35 with a 10-point reduction. As can be observed in Fig. 2, the validation parameter optimization in AutoDock failed to produce ROC values above 50%, indicating that the system discriminated actives from decoys randomly. Therefore, the virtual screening was not done in AutoDock. In AutoDockVina, the best ROC AUC and EF1% values were obtained from 60 × 60 × 60 grid points with 0.375 Å spacing (equivalent to 22.5 × 22.5 × 22.5 Å).

The results of the virtual screening of the HerbalDB using AutoDockVina are listed in Table 2. The 12 compounds with binding energies lower than the cutoff value were cassiamin C (A1), procyanidin B2 (B2), ent-epicatechin-(4alpha->8)-ent-epicatechin (C3), epicatechin-(4beta->8)-epicatechin-3-O-gallate (D4), neorhusflavanone (E5), 3-O-galloylepicatechin-(4beta->6)-epicatechin-3-O-gallate (F6), withanolide (G7), 3-O-galloylepigallocatechin-(4beta->6)-epigallocatechin-3-O-gallate (H8), cyanidin 3-(6''-caffeylsophoroside)-5-glucoside (I9), epifriedelanol (J10), gallocatechin-(4alpha->8)-epicatechin (K11), and scutellarein 7-glucosyl-(1->4)-rhamnoside (L12). The lowest binding energy was that of A1 at −12.7 kcal/mol, whereas the highest was that of K11 at −11.5 kcal/mol.
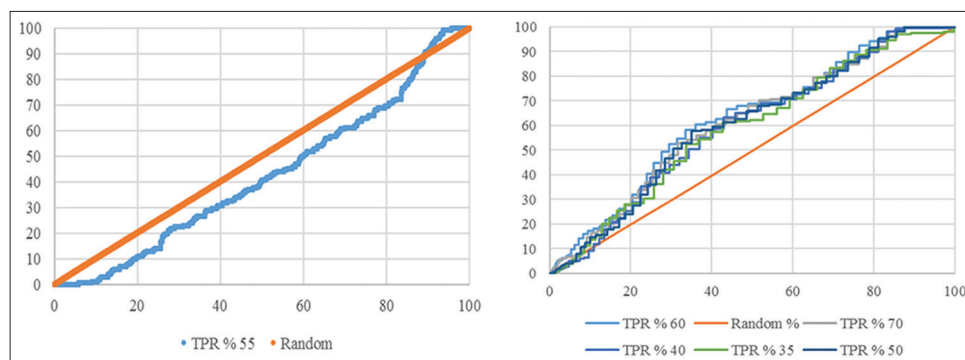
Fig. 2: Receiver operating characteristic curve of actives and decoys using AutoDock (left) and AutoDockVina (right)

Analysis of all target-ligand interactions showed that out of the 12 compounds, B, C, D, E, and I interacted with 4 out of the 7 amino acids that SFG and SAH interacted with Phe1145, Met1169, Cys1191, and Glu1266. The other compounds interacted with adjacent amino acids and occupied a different binding space. Phe1145 and Met1169 interacted hydrophobically, either with SFG and SAH or with the VS compounds. In either SFG or SAH, Cys1191 interacted with N1 on the adenine ring, while in the VS compounds, Cys1191 interacted with the hydroxyl group on the B ring of catechin or epicatechin. Therefore, it can be concluded that these five compounds are possible DNMT1 inhibitors. Further, *in vitro* study is needed to confirm this premise. The results of all target-ligand analysis can be observed in Table 3.

CONCLUSION

AutoDockVina in PyRx yields the best virtual screening result with the following parameters: A grid box center at −5.325, −0.522, and −32.228 and a grid box size at 22.5 × 22.5 × 22.5 Å (equivalent to grid point numbers of 60, 60, and 60 for the x-, y-, and z-axes, respectively, with 0.375 Å between points). A virtual screening conducted with these parameters produced 12 compounds with binding energy values lower than −11.1 kcal/mol. An analysis of protein-ligand interactions showed that five of these compounds interacted with many of the same amino acids as the reference compounds did. These five compounds were procyanidin B2, ent-epicatechin-(4alpha->8)-ent-epicatechin, epicatechin-(4beta->8)-epicatechin-3-O-gallate, neorhusflavanone, and cyanidin 3-(6''-caffeylsophoroside)-5-glucoside. An *in vitro* study must be conducted to explore the DNMT1 inhibitory potential of these five compounds.

REFERENCES

1. Balbus JM, Barouki R, Birnbaum LS, Etzel RA, Gluckman PD Sr, Grandjean P, *et al*. Early-life prevention of non-communicable diseases. Lancet 2013;381(9860):3-4.
2. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology 2013;38(1):23-38.
3. Dhe-Paganon S, Syeda F, Park L. DNA methyl transferase 1: Regulatory mechanisms and implications in health and disease. Int J Biochem Mol Biol 2011;2(1):58-66.
4. Ling C, Del Guerra S, Lupi R, Rönn T, Granhall C, Luthman H, *et al*. Epigenetic regulation of PPARGC1A in human Type 2 diabetic islets and effect on insulin secretion. Diabetologia 2008;51(4):615-22.
5. Wegner M, Neddermann D, Piorunska-Stolzmann M, Jagodzinski PP. Role of epigenetic mechanisms in the development of chronic complications of diabetes. Diabetes Res Clin Pract 2014;105(2):164-75.
6. Mikeska T, Craig JM. DNA methylation biomarkers: Cancer and beyond. Genes (Basel) 2014;5(3):821-64.
7. Lyko F, Brown R. DNA methyltransferase inhibitors and the development of epigenetic cancer therapies. J Natl Cancer Inst 2005;97(20):1498-506.
8. Komashko VM, Farnham PJ. 5-azacytidine treatment reorganizes genomic histone modification patterns. Epigenetics 2010;5(3):229-40.
9. Poirier S, Samami S, Mamarbachi M, Demers A, Chang TY, Vance DE, *et al*. The epigenetic drug 5-azacytidine interferes with cholesterol and lipid metabolism. J Biol Chem 2014;289(27):18736-51.
10. Medina-Franco JL. Advances in computational approaches for drug discovery based on natural products. Rev Latinoam Quim 2013;41(2):95-110.
11. Maldonado-Rojas W, Olivero-Verbel J, Marrero-Ponce Y. Computational fishing of new DNA methyltransferase inhibitors from natural products. J Mol Graph Model 2015;60:43-54.
12. Liu Z, Xie Z, Jones W, Pavlovicz RE, Liu S, Yu J, *et al*. Curcumin is a potent DNA hypomethylation agent. Bioorg Med Chem Lett 2009;19(3):706-9.
13. Fang MZ, Wang Y, Ai N. Tea polyphenol (−)-epigallocatechin-3-gallate inhibits DNA methyltransferase and reactivates methylation-silenced genes in cancer cell lines reactivates methylation-silenced genes in cancer cell lines. Cancer Res 2003;63(2):7563-70.
14. Xie Q, Bai Q, Zou LY, Zhang QY, Zhou Y, Chang H, *et al*. Genistein inhibits DNA methylation and increases expression of tumor suppressor genes in human breast cancer cells. Genes Chromosomes Cancer 2014;53(5):422-31.
15. Lee WJ, Zhu BT. Inhibition of DNA methylation by caffeic acid and chlorogenic acid, two common catechol-containing coffee polyphenols. Carcinogenesis 2006;27(2):269-77.
16. García J, Franci G, Pereira R, Benedetti R, Nebbioso A, Rodríguez-Barrios F, *et al*. Epigenetic profiling of the antitumor natural product psammaplin A and its analogues. Bioorg Med Chem 2011;19(12):3637-49.
17. Agarwal S, Amin KS, Jagadeesh S, Baishay G, Rao PG, Barua NC, *et al*. Mahanine restores RASSF1A expression by down-regulating DNMT1 and DNMT3B in prostate cancer cells. Mol Cancer 2013;12(1):99.
18. Caniago I, Stephen FS. Medicinal plant ecology, knowledge and conservation in Kalimantan, Indonesia. Econ Bot 1998;52(3):229-50.
19. Slik JW, Raes N, Aiba S, Brearley FQ, Cannon CH, Meijaard E, *et al*. Environmental correlates for tropical tree diversity and distribution patterns in Borneo. Divers Distrib 2009;15(3):523-32.
20. Yanuar A, Mun'im A, Lagho AB, Syahdi RR, Rahmat M, Suhartanto H. Medicinal plants database and three dimensional structure of the chemical compounds from medicinal plants in Indonesia. Int J Comput Sci Issu 2011;8(5):180-3.
21. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, *et al*. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. J Comput Chem 1998;19:1639-62.
22. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31(2):455-61.
23. Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Methods Mol Biol 2015;1263:243-50.
24. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. J Med Chem 2012;55(14):6582-94.
25. Yanuar A, Suhartanto H, Munim A, Anugraha BH, Syahdi RR. Virtual screening of indonesian herbal database as HIV-1 protease inhibitor. Bioinformation 2014;10(2):52-5.
26. Syahdi RR, Mun'im A, Suhartanto H, Yanuar A. Virtual screening of Indonesian herbal database as HIV-1 reverse transcriptase inhibitor. Bioinformation 2012;8(24):1206-10.
27. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, *et al*. The protein data bank. Nucleic Acids Res 2000;28(1):235-42.

28. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bansal P, Bridge AJ, *et al.* UniProtKB/Swiss-prot, the manually annotated section of the UniProt knowledge base: How to use the entry view. Methods Mol Biol 2016;1374:23-54.
29. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, *et al.* The InterPro protein families database: The classification resource after 15 years. Nucleic Acids Res 2015;43:D213-21.
30. Haas J, Roth S, Arnold K, Kiefer F, Schmidt T, Bordoli L, *et al.* The protein model portal-a comprehensive resource for protein structure and model information. Database (Oxford) 2013;2013:bat031.
31. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity-a rapid access to atomic charges. Tetrahedron 1980;36(22):3219-28.
32. Halgren TA. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 1996;17:490-519.
33. Weininger DS. A chemical language and information system. Introduction to methodology and encoding rules. J Chem Inf Model 1998;28(1):31-6.
34. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open babel: An open chemical toolbox. J Cheminform 2011;3:33.
35. Laskowski RA, Swindells MB. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model 2011;51(10):2778-86.