

Virtual Screening of Indonesian Herbal Database as HIV-1 Protease Inhibitor

Arry Yanuar^{1*}, Heru Suhartanto², Abdul Mun'im¹, Bram Hik Anugraha¹ & Rezi Riadhi Syahdi¹

¹Faculty of Pharmacy, Universitas Indonesia, Depok 16424, Indonesia; ²Faculty of Computer Sciences, Universitas Indonesia, Depok 16424, Indonesia; Arry Yanuar – Email: arry.yanuar@ui.ac.id; *Corresponding author

Received January 25, 2014; Accepted February 03, 2014; Published February 19, 2014

Abstract:

HIV-1 (Human immunodeficiency virus type 1)'s infection is considered as one of most harmful disease known by human, the survivability rate of the host reduced significantly when it developed into AIDS. HIV drug resistance is one of the main problems of its treatment and several drug designs have been done to find new leads compound as the cure. In this study, *in silico* virtual screening approach was used to find lead molecules from the library or database of natural compounds as HIV-1 protease inhibitor. Virtual screening against Indonesian Herbal Database with AutoDock was performed on HIV-1 protease. From the virtual screening, top ten compounds obtained were 8-Hydroxyapigenin 8-(2'',4''-disulfatoglucuronide), Isoscutellarein 4'-methyl ether, Amaranthin, Torvanol A, Ursonic acid, 5-Carboxypyranocyanidin 3-O-(6''-O-malonyl-beta-glucopyranoside), Oleoside, Jacoumaric acid, Platanic acid and 5-Carboxypyranocyanidin 3-O-beta-glucopyranoside.

Keywords: Indonesian herbal database, HIV-1 protease, molecular docking, virtual screening.

Background:

HIV (Human Immunodeficiency Virus) is still becoming global health major concern. More than 34 million of people were estimated living with HIV in 2011. Its infection could develop into AIDS (acquired immunodeficiency syndrome), marked as one of the most destructive disease in modern era. UNAIDS had estimated 1.7 million adult and child deaths due to AIDS in 2011 alone. HIV has high resistance characteristic due to its high rate of replication (10^9 - 10^{10} virions per person per day) and error-prone polymerase, and 5-10 mutations estimated to occur in each HIV life cycle [1,2]. Unless some steps in the cycle are interrupted by specific treatment, the infection will spread throughout the body leading to deterioration of the host's immune system [3].

Maturation is an essential step in HIV life cycle, triggered by proteolytic cleavage of polyproteins by viral protease [4]. The critical role of HIV protease in the virus cycle was demonstrated as early as 1988 and within a decade, the first inhibitor of the enzyme, saquinavir, was already licensed for use in the treatment of HIV infection [5, 6]. As for 2012, not less than 9 protease inhibitors were approved by FDA to be used as medication to HIV [7]. The development of those protease inhibitors were a notable example of structure aided drug design and suggest there are more to be attained from it. New lead of protease inhibitor could be developed using this

approach, including lead from natural product. Indonesia is recognized for its biodiversity, ranked second after Brazil. It consist more than 30,000 plant species out of 40,000 species known all over the world. At least 9,600 species have pharmacological activity [8]. Previous research has gathered some of the three dimensional data of Indonesian herbal compounds and published it as Indonesian Herbal Database (HerbalDB) (<http://herbaldb.farmasi.ui.ac.id>) [9].

In this research, we applied virtual screening as one of *in silico* structure aided drug design method to HerbalDB. This research employed structure-based screening (molecular docking) with AutoDock [10] and PyRx [11]. Top ten compounds obtained from the virtual screening of HerbalDB will be considered as hits and will be represented in ranks. The aim of this research is to find feasible AutoDock virtual screening parameter as HIV-1 protease inhibitor and to search the potential candidates from HerbalDB using the parameters obtained.

Methodology:

This research was done using literature study and virtual screening by molecular docking (structure-based virtual screening). The experimental design for our experiment was as follow:

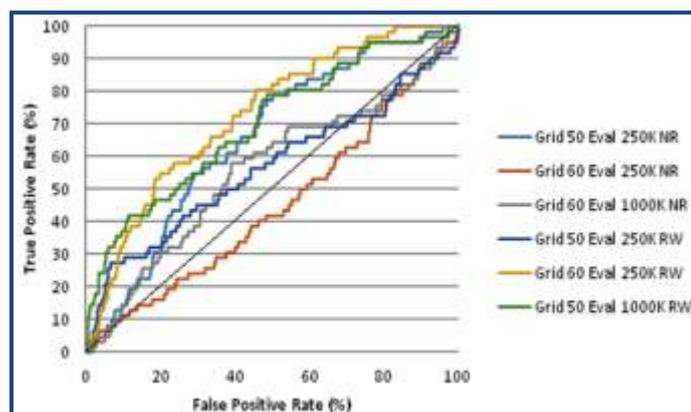


Figure 1: Receiver Operating Characteristic Curve of DUD Database Virtual Screening using AutoDock. K = kilo (thousand), NR = Not retaining water, RW = Retains water at binding site.

Preparing the HIV-1 PR Protein Structure

The macromolecule (HIV-1 PR) three dimensional structure were obtained from the Protein Data Bank website www.pdb.org [12]. The inclusion criteria for our target macromolecule were wild type / non-mutant protein and in a complex with ligand. The exclusion criteria were having resolution >2.5 Å and incomplete chain. Macromolecule was downloaded with the *.pdb format for further processes. The chosen macromolecule then has to be optimized for virtual screening preparation. The macromolecule cleaned from solute, marker and cofactor molecules. Gasteiger partial charges and AutoDock forcefield were applied.

Preparing Ligand Structure

Ligands used for this research were classified into two groups. The first group consists of ligands that were obtained from DUD [13]. We specifically downloaded controls for HIV-1 protease. It is consisted of 62 positive controls (simply called as DUD ligands) and 2038 negative controls (called as decoys). The data have *.mol2 format. The second ligands groups used in this research was obtained from HerbalDB. It contains known molecules from some plants that were used for herbal medicine in Indonesia. The molecules have *.mol and *.mol2 format. It were converted to *.pdbqt format so it will be available to use by AutoDock. All ligands (controls from DUD and ligands from HerbalDB) were converted from *.mol2 format into *.pdbqt format using Open Babel [14].

Optimizing the Virtual Screening Parameter

Validation of Virtual screening using DUD database ligands and decoys was done using AutoDock 4.2 in PyRx [10, 11]. The screening result will be tabulated and analyzed from its Receiver Operating Characteristic (ROC) curve and Enrichment Factors (EF) to search for feasible parameters which will gives accurate prediction without compensating too much times and computer costs.

Virtual Screening of Indonesian Herbal Database

After obtaining the parameter from the optimization, virtual screening was done to all structures of HerbalDB using AutoDock in PyRx. We sorted the screening result based on binding affinities and then ranked it to get top ten compounds that will be considered as hits in this reseach. Visualization

was also done using AutoDockTools and PyMol to see the interaction between the ligand and HIV-1 protease.

Result & Discussion:

The virtual screening target chosen for this research is wild HIV-1 protease with pdb identity 3OXC [15]. It meets our inclusion criteria as the virtual screening target. After it was downloaded from the protein data bank website, it was all cleaned from water molecules and other ligand chemical components such as saquinavir, formic acid and sulfate ion.

The ligand binding site was determined from the position of 3OXC ligand saquinavir. The XYZ coordinates is 5.192, -4.557, 14.799 for x-axis, y-axis, and z-axis, respectively. We varied gridbox parameter to see which one that gives the best EF1% and ROC AUC. We also compare the result of the screening if it was done completely without water molecule and retaining water molecules in the scope of macromolecule binding site. The enrichment and ROC AUC of DUD database virtual screening using AutoDock was shown in **Table 1** (see **supplementary material**). The best EF1% was obtained with the parameter as follow: retaining water molecules, maximum energy of evaluation 1,000,000 and gridbox 50 x 50 x 50 units (spacing 0.375 Angstrom per unit). The EF score is better than the EF shown in the DUD site, 11.9 vs 4.7, respectively. The best AUC was obtained from the 60x60x60 gridbox and 250,000 maximum energy of evaluation. Parameter that gives best AUC and best EF both have better AUC (0.7275 and 0.6992, respectively) compared to the AUC from previous research (0.69 for AutoDock). The screening using AutoDock also suggest us that retaining water molecules in binding site could give better Enrichment Factor. The ROC curve of AutoDock screening result was shown in **Figure 1**. We prefer to use the parameter that gives best EF rather than the parameter that gives the best AUC for virtual screening of herbal database. This is given that early EF will be more useful than AUC as the whole compounds of database was rarely synthetized or extracted. The compounds that were highlighted as the top ranks from the screening result will be synthetized and extracted instead.

Virtual screening of HerbalDB using AutoDock results were given on the **Table 2** (see **supplementary material**). The top ten ranked compound were listed as hits from this screening. The binding energy of the top ten ranked compounds were varies from -18.74 kcal/mol to -11.79 kcal/mol. The first and second compounds are flavone glycosides that come from the fruit of *Helicteres isora* [16, 17]. Otake and colleagues have done research to examine anti HIV-1 activity of 30 plants extract. They concluded that water extract of *Helicteres isora* fruit have anti HIV-1 activity, but the compound and mechanism of action wasn't determined from the research [18]. Amaranthin, the compound that was ranked third, is a lectin group pigment that could be found in *Amaranthus caudatus*, *A. tricolor*, *Celosia argentea* and *C. cristata* seed [16, 19]. Ursonic acid in the fifth rank is a member of pentacyclic triterpene that could be found in *Angelica ursica* Maxim, *Ficus macrocarpa*, *Cordia multispicata*, *Myrica rubra*, and *Lantara camara* Linn [16, 20, 21, 22, 23, 24]. From literature search, we haven't found any research about the activity of amaranthin and ursonic acid as protease inhibitor or anti HIV-1. 5-Carboxypyranocyanidin 3-O-(6''-O-malonyl-beta-glucopyranoside) and 5-Carboxypyranocyanidin

3-O-beta-glucopyranoside were in rank 6 and 10, respectively. Both are flavonoid glycoside from *Allium cepa* [16, 25]. *Allium cepa* has been patented to be used in HIV/AIDS treatment by Goldman *et al* [26]. Oleoside in the rank 5 is secoiridoid glucoside from Oleaceae familia plants [16, 27]. The leaf extract of this plant has been used to suppress the acute infection and HIV-1 cell to cell infection [28].

In the eighth rank is jacoumaric acid, a triterpene compound. It can be found in *Jacaranda caucana* and *Psidium guajava* [16, 29, 30]. There isn't any research for anti HIV activity for the compound yet. Platanic acid is found in *Syzigium claviflorum*, *Platanus occidentalis*, *Melaleuca ericifolia*, *Melaleuca leucadendron* [16, 31, 32, 33]. Fujioka and Kashiwada (1994) have researched the effect of betulinic acid and platanic acid as anti-HIV agent from *Syzigium claviflorum* plant. Platanic acid shows anti HIV activity with EC50 6.5 μ M [34].

From this research, we concluded that the optimum result for HIV virtual screening by AutoDock were given by the parameter gridbox 50 x 50 x 50 units, 0.375 Angstrom per unit, gridbox centre at 5.192, -4.557, 14.799 for x-axis, y-axis, and z-axis, respectively, 1,000,000 numbers of evaluation and retaining water molecule near the binding site. The parameter gives better EF1% than the benchmark EF1%. We also done virtual screening from HerbalDB and we obtained top ten compounds as hits. Some of the compound, such as 8-Hydroxyapigenin 8-(2''A''-disulfatoglucuronide), Isoscutellarein 4'-methyl ether, and oleoside came from the plant that have been studied before and shows anti-HIV activity but there isn't any research about the mechanism or the contributing compound. From this research, we could get some insight about the compound and they need to be assayed in vitro to search more about its potential.

Acknowledgement:

We are grateful for funding provided by the Directorate General of Higher Education, Ministry of Education and Culture Republic of Indonesia through the National Strategic Research Project 2013 to Arry Yanuar.

References:

- [1] www.fda.gov/downloads/Drugs/Guidances/ucm071173.pdf
- [2] Basavapathruni A & Anderson KS, *FASEB J.* 2007 **21**: 3795 [PMID: 17639073]
- [3] Brik A & Wong CH, *Org Biomol Chem.* 2003 **1**: 5 [PMID: 12929379]
- [4] Fun A *et al.* *Retrovirology* 2011 **8**: 70 [PMID: 21864346]
- [5] Kohl NE *et al.* *Proc Natl Acad Sci USA* 1988 **85**: 4686 [PMID: 3290901]
- [6] <http://archive.hhs.gov/news/press/1995pres/951207.html>
- [7] http://aidsinfo.nih.gov/contentfiles/ApprovedMedstoTreatHIV_FS_en.pdf
- [8] <http://perpustakaan.depkes.go.id:8180/bitstream//123456789/561/3/KMK381-0307-G.pdf>
- [9] Yanuar A *et al.* *Int J Comp Sci Issue.* 2011 **8**: 180 [http://ijcsi.org/papers/IJCSI-8-5-1-180-183.pdf]
- [10] Morris GM *et al.* *J Comput Chem.* 2009 **30**: 2785 [PMID: 19399780]
- [11] Wolf LK, *Chemical & Engineering News.* 2009 **87**: 31 [DOI: 10.1021/cen-v087n028.p031]
- [12] Berman HM *et al.* *Nucleic Acid Res.* 2000 **28**: 235 [PMID: 10592235]
- [13] Huang N *et al.* *J Med Chem.* 2006 **49**: 6789 [PMID: 17154509]
- [14] O'Boyle NM, *J Cheminform.* 2011 **3**: 33 [PMID: 21982300]
- [15] Tie Y *et al.* *Proteins* 2007 **67**: 232 [PMID: 17243183]
- [16] Afendi FM *et al.* *Plant Cell Physiol.* 2012 **53**: e1 [PMID: 22123792]
- [17] Kamiya K *et al.* *Phytochemistry* 2001 **57**: 297 [PMID: 11382247]
- [18] Otake T *et al.* *Phytother Res.* 1995 **9**: 6 [DOI: 10.1002/ptr.2650090103]
- [19] Schliemann W *et al.* *Phytochemistry* 2001 **58**: 159 [PMID: 11524126]
- [20] Bohlmann F *et al.* *Tetrahedron Lett.* 1968 **9**: 3947 [DOI: 10.1016/S0040-4039(00)72373-2]
- [21] Kuroyanagi M *et al.* *Chem Pharm Bull (Tokyo).* 2001 **49**: 954 [PMID: 11515584]
- [22] Lansky EP & Newman RA, *J Ethnopharmacol.* 2007 **109**: 177 [PMID: 17157465]
- [23] Morikawa T, *J Nat Med.* 2007 **61**: 112 [DOI: 10.1007/s11418-006-0105-8]
- [24] Begum S *et al.* *Chem Pharm Bull (Tokyo).* 2008 **56**: 1317 [PMID: 18758109]
- [25] Fossen T & Andersen ØM, *Phytochemistry* 2003 **62**: 1217 [PMID: 12648539]
- [26] Goldman SP *et al.* *US Patent* 2002.No: WO 2003059071 A1.
- [27] Jensen SR *et al.* *Phytochemistry* 2002 **60**: 213 [PMID: 12031440]
- [28] Lee-Huang S *et al.* *Biochem Biophys Res Commun.* 2003 **307**: 1029 [PMID: 12878215]
- [29] Ogura M *et al.* *Phytochem.* 1977 **16**: 286 [DOI: 10.1016/S0031-9422(00)86809-0]
- [30] Begum S *et al.* *Phytochemistry* 2002 **61**: 399 [PMID: 22375873]
- [31] Abdel Bar FM *et al.* *J Nat Prod.* 2008 **71**: 1787 [PMID: 18826277]
- [32] Lee CK, *J Nat Prod.* 1998 **61**: 375 [PMID: 9548878]
- [33] Ruzicka L & Lambertson AH, *Helv Chim Acta.* 1940 **23**: 1338 [DOI: 10.1002/hlca.194002301157]
- [34] Fujioka T *et al.* *J Nat Prod.* 1994 **57**: 243 [PMID: 8176401]

Edited by P Kanguane

Citation: Yanuar *et al.* Bioinformation 10(2): 052-055 (2014)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Table 1: Enrichment Factor and Area Under Curve of HIV-1 Protease DUD Database Virtual Screening using AutoDock. The best EF1% and AUC was indicated by bold text.

HIV-1 Condition	Protease	Macromolecule	Grid (units)	Box	Maximum Evaluations	Number of Energy	EF 1%	EF 10%	EF 20%	AUC
Without water molecule on binding site			50x50x50		250.000		0	1,45	1,53	0.6551
					1.000.000		6,45	1,77	1,53	0.6152
			60x60x60	250.000		1,61	1,13	0,81	0.4492	
				1.000.000		1,61	1,45	1,45	0.5574	
Retaining water molecules on binding site			50x50x50		250.000		3,23	2,74	1,61	0.5654
					1.000.000		11,29	3,55	2,34	0.6992
			60x60x60	250.000		1,61	3,23	2,66	0,7275	
				1.000.000		1,61	1,13	1,21	0,6055	

Table 2: Top ten compounds of HIV-1 Protease virtual screening result using AutoDock

Rank	Name	ΔG (kcal/mol)	N	SD	CV (%)	Plant(s) Source
1	8-Hydroxyapigenin (2'',4''-disulfatoglucuronide)	-18,74	5	2,89	15,41	<i>Helicteres isora</i>
2	Isoscutellarein 4'-methyl ether 8-(2'',4''-disulfatoglucuronide)	-17,83	5	1,20	6,72	<i>Helicteres isora</i>
3	Amaranthin	-17,72	5	1,58	8,91	<i>Amaranthus caudatus</i> , <i>Amaranthus tricolor</i> , <i>Celosia argentea</i> var.cristata (L.) Kuntze, <i>Celosia argentea</i> var.plumosa (Burvenich) Voss, <i>Celosia cristata</i> , <i>Atriplex hortensis</i> , <i>Chenopodium amaranticolor</i>
4	Torvanol A	-14,58	5	0,59	4,04	<i>Solanum torvum</i>
5	Ursonic acid	-12,46	5	0,03	0,28	<i>Angelica ursina</i> Maxim, <i>Cordia multispicata</i> , <i>Myrica rubra</i> , <i>Lantana camara</i> LINN.
6	5-Carboxypyranocyanidin 3-O-(6''-O-malonyl-beta-glucopyranoside)	-12,39	4	0,35	2,82	<i>Allium cepa</i>
7	Oleoside	-12,04	5	0,46	3,79	<i>Chionanthus retusus</i> Lindl.& Paxton, <i>Forestiera acuminata</i> (Michx.) Poir., <i>Fraxinus</i> sp., <i>Jasminum</i> sp., <i>Ligustrum</i> sp., <i>Nestegis sandwicensis</i> (Gray) O.& I.Deg.& Johnson, <i>Olea</i> sp., <i>Osmanthus</i> sp., <i>Phillyrea latifolia</i> L., <i>Picconia excelsa</i> (Aiton) DC., <i>Syringa</i> sp
8	Jacoumaric acid	-11,95	4	0,06	0,47	<i>Jacaranda caucana</i> , <i>Psidium guajava</i>
9	Platanic acid	-11,91	5	0,19	1,56	<i>Platanus occidentalis</i>
10	5-Carboxypyranocyanidin 3-O-beta-glucopyranoside	-11,79	3	0,55	4,63	<i>Allium cepa</i>