850-2679-1-PB.pdf

by Istyastono_2 Enade

Submission date: 29-Aug-2017 04:32PM (UTC+0700)
Submission ID: 840860127
File name: 850-2679-1-PB.pdf (513.79K)
Word count: 3483
Character count: 18363

RETROSPECTIVE VALIDATION OF A STRUCTURE-BASED VIRTUAL SCREENING PROTOCOL TO IDENTIFY LIGANDS FOR ESTROGEN RECEPTOR ALPHA AND ITS APPLICATION TO IDENTIFY THE ALPHA-MANGOSTIN BINDING POSE

Agustina Setiawati[1], Florentinus Dika Octa Riswanto[2], Sri Hartati Yuliani[3], and Enade Perdana Istyastono[2,3,4,*]

[1]*Laboratory of Pharmacognosy-Phytochemistry, Faculty of Pharmacy, Sanata Dharma University, Paingan, Maguwoharjo, Depok, Yogyakarta 55284, Indonesia*

[2]*Laboratory of Pharmaceutical Chemistry, Faculty of Pharmacy, Sanata Dharma University, Paingan, Maguwoharjo, Depok, Yogyakarta 55284, Indonesia*

[3]*Laboratory of Pharmaceutical Technology, Faculty of Pharmacy, Sanata Dharma University, Paingan, Maguwoharjo, Depok, Yogyakarta 55284, Indonesia*

[4]*Center for Environmental Studies Sanata Dharma University (CESSDU), Soropadan, Condongcatur, Depok, Yogyakarta 55283, Indonesia*

Received April 23, 2014; Accepted June 6, 2014

ABSTRACT

The publicly available enhanced data of ligands and decoys for estrogen receptor alpha (ERα) which were recently published has made the retrospective validation of a structure-based virtual screening (SBVS) protocol to identify ligands for ERα possible. In this article, we present the retrospective validation of an SBVS protocol using PLANTS molecular docking software version 1.2 (PLANTS1.2) as the backbone software. The protocol shows better enrichment factor at 1% false positives ($EF_{1\%}$) value and the Area Under Curve (AUC) value of the Receiver Operator Characteristic (ROC) compared to the original published protocol. Moreover, in all 1000 iterative attempts the protocol could reproduce the co-crystal pose of 4-hydroxitamoxifen in ERα binding pocket. It shows that the protocol is not only able to identify potent ligands for ERα but also able to be employed in examining binding pose of known ligand. Hence, the protocol was successfully employed to examine the binding poses of α-mangostin, an ERα ligand found in the Garcinia mangostana, L. pericarp.

Keywords: *Structure-based virtual screening (SBVS); molecular docking; estrogen receptor alpha (ERα); α-mangostin*

ABSTRAK

Keberadaan data termutakhir ligan-ligan reseptor estrogen alfa (ERα) beserta pengecohnya memungkinkan dilakukan validasi retrospektif pada protokol-protokol Penapisan Virtual Berbasis Struktur (PVBS) yang dikembangkan untuk identifikasi ligan-ligan ERα. Artikel ini membahas validasi retrospektif protokol PVBS yang menggunakan aplikasi penambatan molekuler PLANTS versi 1,2 (PLANTS1.2) sebagai tulang punggung protokol tersebut dalam identifikasi ligan-ligan ERα. Hasil validasi retrospektif menunjukkan bahwa protokol yang dikembangkan memiliki nilai faktor pengayaan pada 1% false positives ($EF_{1\%}$) dan nilai di bawah kurva dari Receiver Operator Characteristic (ROC) yang lebih baik daripada protokol original yang dipublikasikan bersama data ligan dan pengecoh. Protokol tersebut juga divalidasi untuk melihat kemampuannya dalam menambatkan senyawa aktif dengan penambatan ulang 1000 kali ligan ko-kristal 4-hidroksitamoksifen pada kantung ikatan ERα. Dalam 1000 kali iterasi, keseluruhan simulasi menunjukkan kemampuan protokol dalam mereproduksi pose ko-kristal. Hal ini menunjukkan bahwa protokol yang dikembangkan memiliki kemampuan untuk identifikasi ligan-ligan poten pada ERα dan menambatkan ligan ERα dengan tepat di kantung ikatan. Oleh karena itu, protokol tervalidasi ini selanjutnya digunakan untuk meneliti pose-pose ikatan α-mangostin, senyawa aktif yang terdapat pada kulit manggis (Garcinia mangostana, L.).

Kata Kunci: *Penapisan virtual berbasis struktur (PVBS); penambatan molekuler; receptor estrogen alfa (ERα); α-mangostin*

* Corresponding author.
Email address : enade@usd.ac.id

INTRODUCTION

As the 1st leading cause of death in the developed countries and the 2nd in the developing countries, cancer is a major health problem in the world [1]. Among other cancers, breast cancer is ranked the 1st as leading cause of death for women [1-2]. In Yogyakarta, local data for cancer shows that in 2005 (data from 1998-2004), breast cancer was the highest prevalence cancer with 26% of the patients were less than 40 years old [3]. Among other molecular determinants in breast cancer development, estrogen receptor α (ERα) is one of molecular targets in the therapy [4]. Tamoxifen, one of drug of choice in breast cancer therapy [5], is targeting ERα. Tamoxifen itself is a ligand with high affinity for ERα, which is metabolized to 4-hydroxytamoxifen and *N-des*-methyl-4-hydroxo-tamoxifen with affinities circa 30 to 100 times stronger than tamoxifen for ERα [6]. Fortunately, the ERα crystal structure with 4-hydroxytamoxifen as its co-crystal ligand is publicly available in the Protein Data Bank (PDB) with the PDB code of 3ERT (Fig. 1) [7], which could be employed to be a virtual target to identify potential ERα fragments [8]. On the other hand, a public database of enhanced useful decoys (A database of useful decoys: Enhanced (DUD-e)) has recently published for 102 molecular drug targets, including ERα [9]. The article presenting DUD-e shows that employing ERα as the molecular target in a structure-based virtual screening (SBVS) campaign gave enrichment factor at 1% false positives ($EF_{1\%}$) value and the Area Under Curve (AUC) value of the Receiver Operator Characteristic (ROC) of 15.4 and 67.48%, respectively [9]. Some improvements in the virtual screening protocol are therefore required to have more convincing SBVS tools [10]. Therefore, development of SBVS protocols to discover drugs in order to cure or even to prevent the development of breast cancer by targeting ERα is of considerable and timely interest.

In this post genomics era, the development of computer technology is remarkably boosting and assisting the drug discovery and development [11]. One of developing technique in the field is structure-based drug design and discovery [10,12–15], which uses the availability of the three-dimensional (3D) structures of the protein targets [7,9,16-17] to identify or even to design novel ligands [14]. The availability of ERα structure and its ligands and decoys (which has been publicly available since 2006 in a database of useful decoys (DUD) [17]) has led some attempts to construct valid SBVS protocols to identify novel ligands for ERα [8-9,17-18]. By employing PLANTS docking software [19], Anita et al. [8] has developed and retrospectively validated SBVS protocol to identify ERα ligands. The protocol showed a better VS quality compared to the original protocol accompanying the publication of the

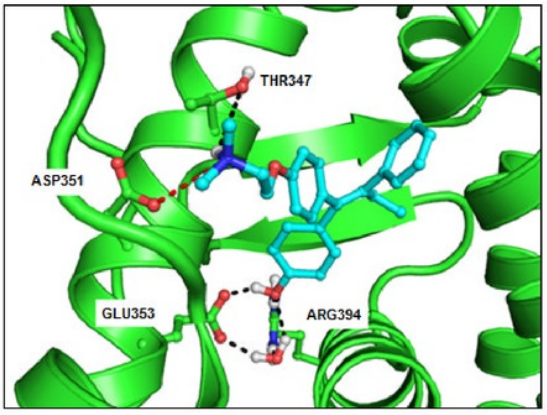

Fig 1. The co-crystal ligand 4-hydroxytamoxifen (carbon atoms are in cyan) in the ERα (carbon atoms are in green) binding pocket [7]. The ERα is presented in the cartoon mode, while the crystal structure pose is presented in the sticks mode. Only polar hydrogens (presented in white), residues (presented in sticks mode, carbon atoms are in green) with hydrogen bond interaction (presented in black dashes) and ionic interaction (presented in red dashes) to the ligand, and a conserved water molecule [8] are presented for the sake of clarity. Nitrogen and oxygen atoms are presented in blue and red, respectively

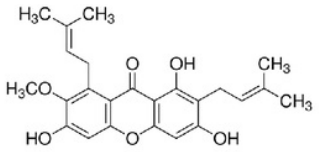

Fig 2. Structure α-mangostin [21]

ligands and decoys by Huang et al. [17]. Subsequently, in order to test the applicability of a software to identify protein-ligand interaction fingerprints PyPLIF in a SBVS campaign, Radifar et al. [18] re-scored and re-validated the results from the SBVS protocol constructed by Anita et al. [8]. PyPLIF-aided SBVS protocol showed a significant increase in quality by filtering on the hydrogen bond interactions of the ligands to the ASP351 of the ERα [18]. Interestingly, the co-crystal ligand 4-hydroxytamoxifen in the crystal structure of the ERα (3ERT.pdb) does not have the hydrogen bond interactions to the ASP351 of the ERα (Fig. 1) [7]. The SBVS protocol developed by Radifar et al. [18] is therefore not able to reproduce the pose of the co-crystal ligand in the crystal structure of the ERα. During the fine tuning to have a valid SBVS protocol that is able to identify ligands and to reproduce the co-crystal pose, a new enhanced database of useful

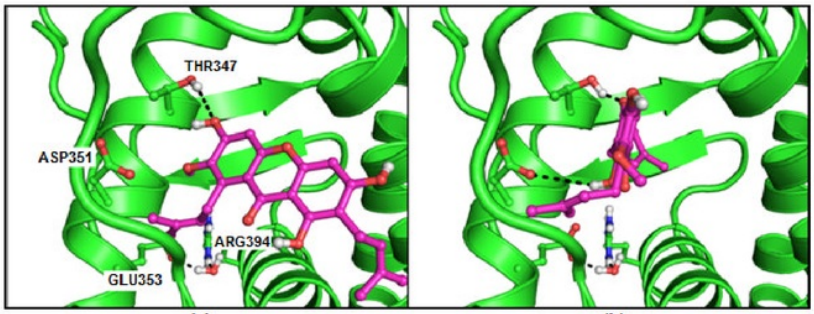

Fig 3. The representatives of the identified binding poses of α-mangostin (carbon atoms are in magenta) in the ERα (carbon atoms are in green) binding pocket [7] from cluster 1 **(a)** and cluster 2 **(b)**. The rendering is similar to Fig. 1

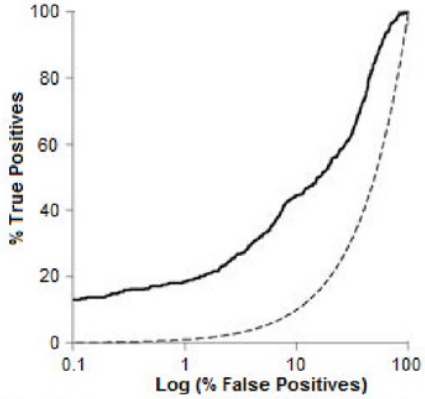

Fig 4. The ROC curves of the SBVS protocol developed in this research (solid line) and the random sampling (dashed line)

decoys DUD-e was published for 102 molecular drug targets, including ERα [9]. This database has more ligands and decoys compared to DUD [9,17].

This research aimed to develop a robust computational medicinal chemistry tool in order to discover novel ERα ligands and to examine the binding poses of known ERα ligands. In this article, the retrospective validation of a modified SBVS protocol developed by Anita et al. [8] using the DUD-e database is presented. The validated protocol was subsequently examined to see its ability to reproduce the pose of the co-crystal ligand in the ERα binding pocket and to examine the binding pose of α-mangostin (Fig. 2), an ERα ligand found in the *Garcinia mangostana*, L. pericarp [20-21]. The modified SBVS protocol developed in this research showed better VS quality compared to the original SBVS protocol accompanying the publication of DUD-e [9] and was able to reproduce the pose of the co-crystal ligand [7]. Two distinct binding poses of α-mangostin in the ERα binding pocket (Fig. 3) were identified using the retrospectively validated protocol in this research.

METHODS

Materials

A dataset of ERα ligands (383 compounds) and their decoys (20,685 compounds) in file type of *.mol2* obtained from DUD-e [9]. The ERα crystal structure and its co-crystal ligand 4-hydroxytamoxifen (3ERT.pdb) downloaded from Protein Data Bank (PDB) submitted by Shiau et al. [7]. Configuration files to perform molecular docking simulations in order to perform SBVS to identify ligands for ERα using PLANTS docking software prepared by Anita et al. [8]: (i) The virtual target *protein.mol2*, (ii) the conserved water molecule *water.mol2*, and (iii) the configuration files to run PLANTS docking software *plants.config*.

Computation Details

Computational medicinal chemistry applications used in this research were: PLANTS docking software version 1.2 (PLANTS1.2) [19,22] to perform molecular docking simulations, R computational statistics software to calculate $EF_{1\%}$ and AUC and to perform statistical tests [23] and PyMOL [24, 25] to calculate the root mean square distances (RMSD) values and to generate molecular pictures. All calculations and computational simulations were performed on a Linux (Ubuntu 10.04 LTS Lucid Lynx) machine with Intel(R) Xeon(R) CPU E31220 (@ 3.10 GHz) as the processors and 8.00 GB of RAM.

Procedure

Every compound downloaded from DUD-e [9] were docked independently three times in the binding

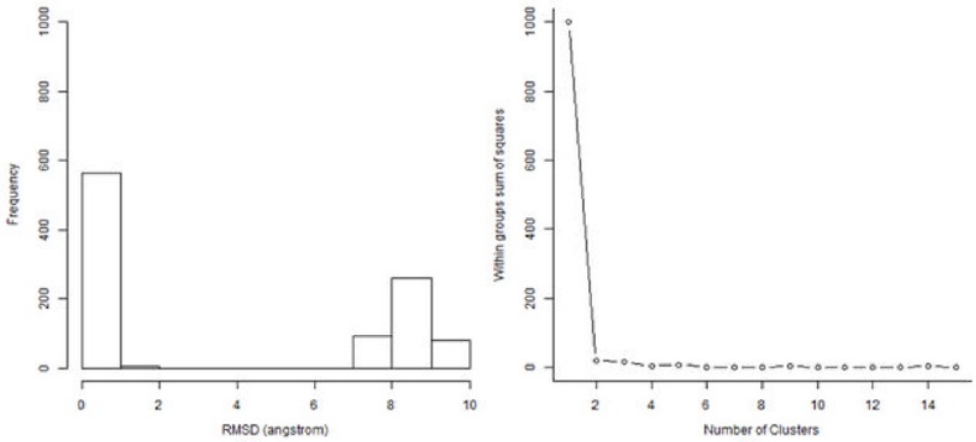

Fig 5. The histogram **(a)** and the Scree plot **(b)** of the RMSD values of α-mangostin docking poses compared to its binding pose with the best ChemPLP value

pocket of ERα by using configuration files from Anita et al. [8]. Every run resulted in 50 docked poses. Therefore, every compound had 150 poses. The best pose of each compound was selected as the pose with the best ChemPLP score [19]. The compounds were then ranked based on their ChemPLP score [10]. The $EF_{1\%}$ and the AUC values were then calculated [26] by using R computational statistics software [23].

The similar procedure was performed to dock co-crystal ligand 4-hidroxytamoxifen in the ERα binding pocket 1000 times to see the ability of the protocol to reproduce the pose of the co-crystal ligand. The best poses collected in every run were compared to the pose of the co-crystal ligand and the RMSD values were calculated by using PyMOL [24-25,27]. The protocol was considered as acceptable in reproducing the co-crystal ligand pose if resulted in the RMSD value of less than 2.0 Å [28].

The same procedure to dock co-crystal ligand 4-hidroxytamoxifen in the ERα binding pocket was performed to dock α-mangostin which resulted in 1000 poses. Those poses were compared to the pose with the best ChemPLP score and the RMSD values were calculated [27]. Based on the RMSD values, the poses were clustered by employing *k*-means clustering [29] in R computational statistics software [27].

RESULT AND DISCUSSION

This research aimed to retrospectively validate an SBVS protocol in order to develop a tool to identify ERα ligands and to examine the binding poses of identified ERα ligands. The SBVS protocol used in the research was initially developed by Anita et al. [8] and has been modified in this research. The modification in the SBVS protocol was using three independent molecular docking simulations for each compound instead of one as performed in the initial SBVS protocol [8]. One of the advantages of using some more independent simulations is that we can sample more converged docking poses for every compounds although it increases the computing cost [27,30-31].

Fig. 4 presents the ROC of the % true positives (%TP) and % false positives (%FP) results from the retrospective validations to identify ERα ligands by employing the DUD-e dataset [9,26]. The retrospective validation showed that the $EF_{1\%}$ value of the modified protocol was 18.54 (Fig. 4). This $EF_{1\%}$ value is higher than the $EF_{1\%}$ value of original SBVS protocol (15.4) [9]. The $EF_{1\%}$ represents the early enrichment results from the protocol. The higher the $EF_{1\%}$ value, the better the protocol in the identification of known ERα ligands is [9,26]. It means that in the first 1% of the ranked database the protocol can identify known ligands and put them in the higher rank compared to their decoys [9,26]. Based on Fig. 4, the AUC value could be calculated by using pROC package in R computational statistics software [23]. The AUC value resulted from the retrospective validations was 76.41% with 95% confidence interval of 74.05%-78.78%. This value is also better than the AUC value of the original protocol (67.48%) [9]. The ideal value of the AUC is 100% [26], which indicates that all known ligands are ranked higher than their decoys. In random sampling, the AUC value is 50% [32]. The $EF_{1\%}$ value represents the early enrichment of the protocols, while the AUC value represents the global enrichment [26,32]. Since the $EF_{1\%}$ and AUC values of the SBVS protocol develop in

this research are better than the original protocol [9], we are confident that the protocol is more robust to identify ERα ligands.

The developed protocol was intended to be employed also in the examination of the binding pose of known ERα ligands. The protocol was then challenged to redock the co-crystal ligands 4-hydroxytamoxifen in the ERα binding pocket [7] for 1000 times [27,31]. Remarkably, in all redocking simulations the protocol showed its ability to reproduce the co-crystal pose with RMSD values of < 2.00 Å. The developed protocol in this research is therefore able to identify ERα ligands and to examine their binding pose in the ERα binding pocket.

The protocol was then employed to examine the binding pose of α-mangostin (Fig. 2) resulted in 1000 selected poses from 1000 different iterations of the protocol. The compound α-mangostin, which can be found in the pericarp of *Garcinia mangostana*, L. [21,33] is a known ligand for ERα [21,33]. *Garcinia mangostana*, L. has recently gained its popularity [34] due to its applications as herbal medicines to treat inflammation and bacterial infections [33] as well as its application in cancer chemoprevention [35]. Therefore, it is of interest to select α-mangostin as the lead compound in the structure-based drug design in this research. By examining the RMSD values of the poses compared to the pose with the best ChemPLP value presented in a histogram and a Scree plot in Fig. 5, two distinct poses of α-mangostin were identified (Fig. 3). Interestingly, both poses shows that α-mangostin located in the ERα binding pocket (Fig. 3) only in the subpocket where the co-crystal ligand 4-hydroxytamoxifen interacts to THR347 and ASP351 (Fig. 1). α-mangostin could not go to the subpocket where 4-hydroxytamoxifen interacts to GLU353 and ARG394 (Fig. 1). This indicates that referring to 4-hydroxytamoxifen as the co-crystal ligand, α-mangostin interacts to ERα as an allosteric ligand.

CONCLUSION

The SBVS protocol developed in this research is a robust computational tool to identify ERα ligands and to examine their poses in the binding pocket ERα. The application of the SBVS protocol to examine the binding poses of a known ERα ligand α-mangostin resulted in two distinct binding poses. The binding poses of α-mangostin indicate that it interacts in the allosteric site of ERα.

ACKNOWLEDGEMENT

This work was supported by Sanata Dharma University Special Internal Research Grant 2013 (No. 083/Panel/LPPM USD/SP/X/2013).

REFERENCES

1. Siegel, R., Naishadham, D., and Jemal, A., 2012, *CA Cancer J. Clin.*, 62 (1), 10–29.
2. Jemal, A., Bray, F., and Ferlay, J., 2011, *CA Cancer J. Clin.*, 61 (2), 69–90.
3. Purnomosari, D., 2006, *PhD Thesis*, VU University, Amsterdam.
4. Harvey, J.M., Clark, G.M., Osborne, C.K., and Allred, D.C., 1999, *J. Clin. Oncol.*, 17 (5), 1474–1481.
5. Mouridsen, H., Giobbie-Hurder, A., Goldhirsch, A., Thürlimann, B., Paridaens, R., Smith, I., Mauriac, L., Forbes, J.F., Price, K.N., Regan, M.M., Gelber, R.D., and Coates, A.S., 2009, *N. Engl. J. Med.*, 361 (8), 766–776.
6. Desta, Z., Ward, B.A., Soukhova, N.V., and Flockhart, D.A., 2004, *J. Pharmacol. Exp. Ther.*, 310 (3), 1062–1075.
7. Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A., and Greene, G.L., 1998, *Cell*, 95 (7), 927–937.
8. Anita, Y., Radifar, M., Kardono, L., Hanafi, M., and Istyastono, E.P., 2012, *Bioinformation*, 8 (19), 901–906.
9. Mysinger, M.M., Carchia, M., Irwin, J.J., and Shoichet, B.K., 2012, *J. Med. Chem.*, 55 (14), 6582–6594.
10. De Graaf, C., Kooistra, A.J., Vischer, H.F., Katritch, V., Kuijer, M., Shiroishi, M., Iwata, S., Shimamura, T., Stevens, R.C., de Esch, I.J.P., and Leurs, R., 2011, *J. Med. Chem.*, 54 (23), 8195–8206.
11. Van Gunsteren, W.F., Bakowies, D., Baron, R., Chandrasekhar, I., Christen, M., Daura, X., Gee, P., Geerke, D.P., Glättli, A., Hünenberger, P.H., Kastenholz, M.A., Oostenbrink, C., Schenk, M., Trzesniak, D., van der Vegt, N.F.A., and Yu, H.B., 2006, *Angew. Chem. Int. Ed.*, 45 (25), 4064–4092.
12. Kooistra, A.J., Leurs, R., de Esch, I.J.P., and de Graaf, C., 2014, "From Three-Dimensional GPCR Structure to Rational Ligand Discovery" in *G Protein-Coupled Receptors - Modeling and Simulation.,* Ed. Filizola, M., Springer Netherlands, 129–157.
13. De Graaf, C., Oostenbrink, C., Keizers, P.H.J., van der Wijst, T., Jongejan, A., and Vermeulen, N.P.E., 2006, *J. Med. Chem.*, 49 (8), 2417–2430.
14. Istyastono, E.P., Nijmeijer, S., Lim, H.D., van de Stolpe, A., Roumen, L., Kooistra, A.J., Vischer, H.F., de Esch, I.J.P., Leurs, R., and de Graaf, C, 2011, *J. Med. Chem.*, 54 (23), 8136–8147.
15. Sirci, F., Istyastono, E.P., Vischer, H.F., Kooistra, A.J., Nijmeijer, S., Kuijer, M., Wijtmans, M., Mannhold, R., Leurs, R., de Esch, I.J.P., and de

Graaf, C., 2012, *J. Chem. Inf. Model.*, 52 (12), 3308–3324.
16. Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S., and Coleman, R.G., 2012, *J. Chem. Inf. Model.*, 52 (7), 1757–1768.
17. Huang, N., Shoichet, B.K., and Irwin, J.J., 2006, *J. Med. Chem.*, 49 (23), 6789–6801.
18. Radifar, M., Yuniarti, N., and Istyastono, E.P., 2013, *Bioinformation*, 9 (6), 325–328.
19. Korb, O., Stützle, T., and Exner, T.E., 2009, *J. Chem. Inf. Model.*, 49, 1, 84–96.
20. Hopert, A.C., Beyer, A., Frank, K., Strunck, E., Wünsche, W., and Vollmer, G., 1998, *Environ. Health Perspect.*, 106 (9), 581–586.
21. Shibata, M.A., Iinuma, M., Morimoto, J., Kurose, H., Akamatsu, K., Okuno, Y., Akao, Y., and Otsuki, Y., 2011, *BMC Med.*, 9 (69), 1–18.
22. Korb, O., Stützle, T., and Exner, T.E., 2007, *Swarm Intell.*, 1 (2), 115–134.
23. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., and Müller, M., 2011, *BMC Bioinf.*, 12 (77), 1–8.
24. Lill, M.A., and Danielson, M.L., 2011, *J. Comput. Aided Mol. Des.*, 25 (1), 13–19.
25. The PyMOL Molecular Graphics System, Version 1.2r1, Schrödinger, LLC.
26. De Graaf, C., and Rognan, D., 2008, *J. Med. Chem.*, 51 (16), 4978–4985.
27. Istyastono, E.P., 2012, *Indo. J. Chem.*, 12 (2), 141–145.
28. Marcou, G., and Rognan, D. 2007, *J. Chem. Inf. Model.*, 47 (1), 195–207.
29. Shi, W., and Zeng, W., 2013, *Int. J. Environ. Res. Public Health*, 10 (6), 2578–2595.
30. Wijtmans, M., de Graaf, C., de Kloe, G., Istyastono, E.P., Smit, J., Lim, H., Boonnak, R., Nijmeijer, S., Smits, R.A., Jongejan, A., Zuiderveld, O., de Esch, I.J.P., and Leurs, R., 2011, *J. Med. Chem.*, 54 (6), 1693–1703.
31. Radifar, M., Yuniarti, N., and Istyastono, E.P., 2013, *Indo. J. Chem.*, 13 (3), 283–286.
32. Jain, A.N., and Nicholls, A., 2008, *J. Comput. Aided Mol. Des.*, 22 (3-4), 133–139.
33. Yodhnu, S., Sirikatitham, A., and Wattanapiromsakul, C., 2009, *J. Chromatogr. Sci.*, 47 (3), 185–189.
34. Purwanto, A., 2008, *Skripsi*, Universitas Sebelas Maret Surakarta
35. Balunas, M.J., Su, B., Brueggemeier, R.W., and Kinghorn, A., 2008, *J. Nat. Prod.*, 71 (7), 1161–1166.

850-2679-1-PB.pdf

ORIGINALITY REPORT

0%	0%	0%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

Exclude quotes	On	Exclude matches	< 25%
Exclude bibliography	On		