BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 8(19)

Structure-based design of eugenol analogs as potential estrogen receptor antagonists

Yulia Anita[1], Muhammad Radifar[2], Leonardus BS Kardono[1], Muhammad Hanafi[1] & Enade P Istyastono[3]*

[1]Research Center for Chemistry, Indonesian Institute of Sciences, Komplek Puspiptek Serpong, Indonesia; [2]Graduate School of Biotechnology, GadjahMada University, Yogyakarta, Indonesia; [3]Pharmaceutical Technology Laboratory, Faculty of Pharmacy, Sanata Dharma University, Yogyakarta, Indonesia; Enade P Istyastono – Email: enade@usd.ac.id; Phone: +62-274-883037, Fax: +62-274-886529; *Corresponding author

Received September 08, 2012; Accepted September 13, 2012; Published October 01, 2012

Abstract:
Eugenol is an essential oil mainly found in the buds and leaves of clove (*Syzygium aromaticum* (L.) Merrill and Perry), which has been reported to have activity on inhibition of cell proliferation and apoptosis induction in human MCF-7 breast cancer cells. This biological activity is correlated to its activity as an estrogen receptor antagonist. In this article, we present the construction and validation of structure-based virtual screening (SBVS) protocols to identify the potent estrogen receptor α (ER) antagonists. The selected protocol, which gave acceptable enrichment factors as a virtual screening protocol, subsequently used to virtually screen eugenol, its analogs and their dimers. Based on the virtual screening results, dimer eugenol of 4-[4-hydroxy-3-(prop-2-en-1-yl)phenyl]-2-(prop-2-en-1-yl)phenol is recommended to be developed further in order to discover novel and potent ER antagonists.

Keywords: Virtual screening, Estrogen receptor, Docking, Eugenol

Background:
Eugenol (compound **1**) is an essential oil mainly found in the buds and leaves of clove (*Syzygium aromaticum* (L.) Merrill and Perry **[1].** This essential oil has some biological activities, e.g., antiinfective (i.e., antibacterial, anthelmintic, antifungal, antiplasmodial and antiviral), anti-inflammatory, analgesic, antioxidant, antimutagenic, antigenotoxic, modulatory and anticancer **[1, 2]**. As an anticancer, eugenol inhibits cell proliferation and induces apoptosis in human MCF-7 breast cancer cells **[2, 3].** This type of cancer is the most common form among women **[4].** Therefore, drug discovery efforts by exploring the potency of eugenol in order to develop novel and potent pharmaceuticals for breast cancer therapy are of considerable interest.

The biological activities of eugenol in human breast cancer cells can be correlated to its potential activity as an estrogen receptor α (ER) antagonist **[1–3].** Interestingly, the standard adjuvant for postmenopausal women with hormone-receptor–positive early breast cancer Tamoxifen (Astra Zeneca) is an ER antagonist **[5]**. Tamoxifen itself is a prodrug that is metabolized in the liver results in some active metabolites (e.g., 4-hydroxytamoxifen and *N*-desmethyl-4-hydroxy-tamoxifen), with 30-100 fold activity in the binding to ER compared to its original form **[6]**. On the other hand, compared to tamoxifen and its metabolites, eugenol can be considered as a small fragment that has a potency to be developed further in a direction guided by a computer-aided structure-based design **[7, 8]** to have compounds that have similar or even better affinities to ER than tamoxifen and its metabolites.

We described previously the construction and validation of the structure-based virtual screening (SBVS) protocols to discover cyclooxygenase-2 (COX-2) inhibitors **[9].** In this article, similar approaches were employed to construct and validate SBVS protocols to discover potent ER antagonist. Fortunately, similar

to the retrospective validation of the SBVS to discover COX-2 inhibitors, the dataset to retrospectively validate SBVS protocols to discover potent ER antagonist has been made publicly available in the directory of useful decoys (DUD; http://dud.docking.org/r2/er_antagonist.tar.gz) **[10]**. The validated protocol has a better enrichment factor in 1% false positive ($EF_{1\%}$) compared to the first SBVS campaign using DUD to retrospectively identify ER antagonists **[10]**. Moreover, the $EF_{1\%}$ value of the validated protocol constructed here is significantly higher than the average value (17.3) resulted from the first SBVS campaign of 40 targets employing DUD and can therefore be considered as acceptable **[10]**. The validated protocol was subsequently employed to virtually screen eugenol (compound 1), its analogues (compounds **2**-**7**) and their dimers (8-14). None of the compounds show better docking score as compared to the threshold compound of the $EF_{1\%}$ value. However, instead of being considered as drug-like compounds, the screened compounds are considered as fragments that can be developed further **[7, 11]**. Therefore, by employing docking score ligand efficiency (DSLE; the absolute value of docking score divided by number of heavy atoms) value these initial results guide us to select dimer 11 (4-[4-hydroxy-3-(prop-2-en-1-yl)phenyl]-2-(prop-2-en-1-yl)phenol) as the most potential fragment to be developed further in order to discover novel and potent ER antagonists.

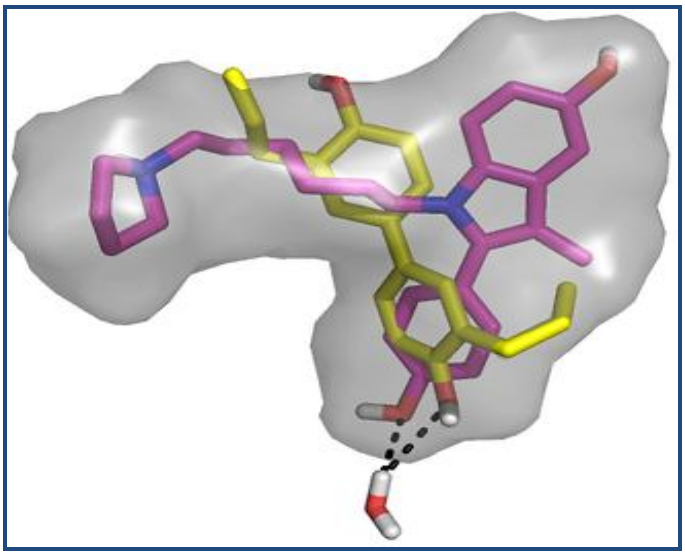


Figure 1: The superposition of the docked poses of compounds 11 (yellow carbon atoms) and 15 (magenta carbon atoms). The surface was generated based on the docked pose of compound 15. The conserved water molecule is also showed here for clarity. The hydrogen bonds are indicated by dashed black lines. The 3D figure was created using PyMOL 1.2 (http://www.pymol.org/).

Methodology:

Molecular docking protocol construction and internal validation

The crystal structure of human ER bound to 4-hydroxytamoxifen (PDB code: 3ERT; http://www.rcsb.org/pdb/files/3ERT.pdb.gz) was used as the reference target **[10, 12].** By employing SPORES **[13]** subjected to the reference target, the virtual target file (*protein.mol2*) was prepared. The binding pocket of ER was defined by the coordinates of the co-crystalized 4-hydroxytamoxifen in the 3ERT structure and a radius of 12.8 Å, which is the maximum distance from the center defined by a 5 Å **[8]** radius around 4-hydroxytamoxifen. In the visual inspection of the 3ERT structure, a water molecule in the binding pocket was observed. As the representative of the water molecule, a file named *water.mol2* was then prepared. Two PLANTS **[14]** configuration files were then prepared: (i) the configuration that ignores the conserved water molecule (*plantsconfig*) and (ii) the configuration that involves the conserved water molecule (*water_plantsconfig*). For each configuration the internal validation was performed by redocking the co-crystalized ligand 4-hydroxytamoxifen into the virtual target using the docking software PLANTS1.2 **[14]** and subsequently compared the docking pose to the original crystal structure pose **[12, 15]**. In order to avoid bias, instead of using the co-crystalized ligand as the starting point, the optimized state of the lowest energy conformer of the co-crystalized ligand 4-hydroxytamoxifen was used as the input ligand. By employing Open Babel 2.2.3 **[16]**, hydrogen atoms in pH 7.4 was added to the input ligand by module *babel –p 7.4* and followed by module *obconformer* to perform conformer search using Monte Carlo simulations with maximum 250 conformers and followed by energy optimization using steepest descent method with maximum 100 steps.

Retrospective SBVS validation

The ER antagonists and decoys were obtained from DUD website (http://dud.docking.org/r2/) **[10]**. The compounds were treated similar to the co-crystal ligand in the input ligand preparation described in the previous subsection. The prepared input ligands were subsequently screened using PLANTS1.2 **[14]**. For each configuration, a retrospective SBVS campaign was performed independently. The compounds were then ranked based on the scores and the $EF_{1\%}$ values were calculated. The quality of the screening procedures was judged by comparing the $EF_{1\%}$ value to $EF_{1\%}$ of the first retrospective SBVS campaign on ER antagonist using DUD (12.1) **[10]**.

SBVS on eugenol analogs

Eugenol (compound **1)**, its analogues (compounds 2-7) and their dimers (8-14) were virtually screened using the selected SBVS protocols. Their docking scores were then compared to the docking score of the compound located in the $EF_{1\%}$ in the ranked dataset resulted from the selected SBVS validation. Additional objective function called docking score ligand efficiency (DSLE = |docking score/number of heavy atom|) **[17]** was used to rank the potency of eugenol and its analogs to be developed further **[7]**.

All computational simulations were performed on a Dell Power Edge 1900 server with Intel Xeon 2.66 GHz dual core as the processors and 3 GB of RAM and Linux version 2.6.32-30-generic (Ubuntu 10.04 Lucid) as the operating system.

Results and Discussion:

The aim of this research was to construct a validated SBVS protocol to discover potent ER antagonist and subsequently use the protocol to virtually screen small fragments eugenol and its analogs in order to develop novel and potent ER antagonists. Potential small fragments with low potency but high ligand efficiency recognized in a SBVS campaign can successfully lead to high affinity ligand after structure-based optimization **[7, 8, 17].**

The crystal structure with PDB code of 3ERT, which was used in the reference SBVS protocol using DUD, was selected as the reference target **[10, 12]**. This crystal structure has an acceptable resolution (1.90 Å) and the ER in this crystal structure was co-crystalized with ligand 4-hydroxytamoxifen, a high affinity ER antagonist with binding affinity (K_i) value in the range of nanomolar concentrations **[12]**. By visual inspection of the crystal structure 3ERT, the optimal protein-ligand interactions can be studied. There were 70 residues recognized as the binding pocket residues: LEU327, TYR328, SER329, GLU330, SER341, MET342, MET343, GLY344, LEU345, LEU346, THR347, ASN348, LEU349, ALA350, ASP351, ARG352, GLU353, LEU354, VAL355, MET357, LEU379, GLU380, CYS381, ALA382, TRP383, LEU384, GLU385, ILE386, LEU387, MET388, ILE389, GLY390, LEU391, VAL392, ARG394, SER395, LEU402, LEU403, PHE404, ALA405, LEU408, LEU410, GLY415, VAL418, GLU419, GLY420, MET421, VAL422, GLU423, ILE424, PHE425, LEU428, ILE514, HIS516, MET517, SER518, ASN519, LYS520, GLY521, MET522, GLU523, HIS524, LEU525, TYR526, SER527, MET528, LYS529, CYS530, LEU536, and LEU539. Interestingly, one water molecule was observed in the binding pocket and this water molecule can be considered as conserved **[18]**. Two hydrogen bonds networks were observed during the visual inspection: (i) the 4-hydroxy moiety of the 4-hydroxytamoxifen with the conserved water molecule and residues GLU353 and ARG394, and (ii) the (2-hydroxyethyl) dimethylazanium moiety of the 4-hydroxytamoxifen with residues THR347 and ASP351. Proper constraints can lead to the increase of SBVS quality significantly **[8, 9]**. In this SBVS construction, however, no constrain has introduced yet. Instead, the default configuration that ignores the conserved water molecule and the configuration that involves the conserved water molecule were constructed and compared.

The internal validation was aimed to examine whether the docking simulation used by the SBVS protocols can reproduce the pose of the co-crystal ligand **[15]**. The objective function used in the internal validation was the root mean square distance (RMSD) value between the heavy atoms of the docked pose and the crystal structure pose. The default configuration resulted in the RMSD value of 1.670 Å, while the configuration that took into account the conserved water molecule resulted in the RMSD value of 1.403 Å. Although the configuration considering the conserved water molecule gave a slightly better RMSD value, since a protocol is acceptable if the RMSD value is less than 2.0 **[15]**, both protocols can be considered as acceptable. Interestingly, the docked poses resulted from both protocols still maintain the hydrogen bonds networks with residues THR347, ASP351, GLU353, and ARG394, though the default protocol did not involve the conserved water molecule.

The reference retrospective SBVS campaign using DUD showed $EF_{1\%}$ value of 12.7 **[10]**. Moreover, the most recent retrospective SBVS campaign using enhanced DUD showed $EF_{1\%}$ value of 15 **[19]**. Remarkably, the independent retrospective SBVS campaigns using DUD dataset employing PLANTS1.2 described here showed that the default protocol resulted in $EF_{1\%}$ value of 15.9 and the protocol that involved the conserved water molecule resulted in $EF_{1\%}$value of 21.2. Both values give better $EF_{1\%}$ value compared to the reference SBVS campaigns. Notably, the SBVS protocol that involved the conserved water molecule gave significantly higher $EF_{1\%}$value compared to others. This indicates that the conserved water molecule plays an important role in the SBVS campaigns to identify ER antagonists. The $EF_{1\%}$value of the validated protocol constructed here (21.2) is above the average value (17.3) resulted from the first SBVS campaign of 40 targets employing DUD **[10]**. Thus, the SBVS protocol that involved the conserved water molecule is therefore acceptable and selected for further SBVS campaign in subsequent prospective efforts. Using the $EF_{1\%}$value, a reference compound that can be used as the threshold compound in the prospective SBVS was recognized. The compound is ZINC01914469 (compound **15**), an ER antagonist with IC_{50} value of 69.23 nM **[20].**

The prospective screening results of eugenol (**1**), its analogues (2-7) and their dimers (8-14) together with compound **15** as the reference compound are presented in **Table 1 (see supplementary material)**. None of the screened compound shows a better ChemPLP score as compared to compound **15**. However, in order to rank the small fragments 1-14 to be developed further, another objective value named DSLE is introduced. This value is a modified ligand efficiency **[17]** which uses docking score instead of the observed affinity.

According to **Table 1**, eugenol and its analogs in this research resulted in higher DSLE values than the reference compound **15**. This indicates that compounds 1-7 can serve as good starting points in the development of novel and potent ER antagonists. In order to narrow the degree of freedom in the further development, initial design by dimerization (compounds 8-14) was proposed. The prospective SBVS campaign showed that the success of the strategy was monomer dependent since compounds 8, 10, 13 and 14 were shown significant decrease in the DSLE values, which were lower than the DSLE value of the reference compound. Notably, the dimer 11 4-[4-hydroxy-3-(prop-2-en-1-yl) phenyl]-2-(prop-2-en-1-yl) phenol showed the highest DSLE value among the dimers and therefore has been suggested to be developed further. The superposition of the docked poses of compounds 11 and 15 is presented in **(Figure. 1**). Based on **(Figure 1)**, the phenolic moieties nearest to the conserved water molecule of both compound 11 and 15 are located very similar. This creates the hydrogen bonds network to residues GLU353 and ARG394.However, compound 11 lacks of basic moiety that can bind to residue ASP351. Therefore the recommended design strategy to develop compound 11 is to add at least a basic moiety in the similar position to basic moiety of compound 15**.** Subsequently, another phenol moiety to fulfill the hydrophobic pocket possessed by compound **15** can be added to increase the affinity **(Figure 1).**

Conclusion:

The construction and the retrospective validation of SBVS protocols to identify ER antagonists have successfully provided a valid tool to screen potential ER antagonists virtually. The validated protocol has an EF1% of 21.2, which is considered as acceptable. The validation processes have also revealed that the conserved water molecule in the binding pocket of the crystal structure 3ERT plays an important role in the quality of the SBVS protocol. Subsequent prospective screen on eugenol, its analogs and their dimers has suggested dimer 11 4-[4-hydroxy-3-(prop-2-en-1-yl)phenyl]-2-(prop-2-en-1-yl)phenol to be developed further in order to discover novel and potent ER antagonists.

Acknowledgement:
The authors thank our colleagues at Natural Products, Food and Pharmaceuticals Division, Research Centre of Chemistry Serpong (Euis Filailla, et al.) for their technical assistances. This work was supported by Indonesian Institute of Sciences through Competitive Research Block Grant 2012.

References:

[1] Kamatou GP *et al. Molecules.* 2012 **17**: 6953 [PMID: 22728369]
[2] Vidhya N & Devaraj SN, *Indian J Exp Biol.* 2011 **49**: 871 [PMID: 22126019]
[3] Brooks SC *et al. J Biol Chem.* 1973 **248**: 6251 [PMID: 4353636]
[4] Joulaee A *et al. Int Nurs Rev*. 2012 **59**: 362 [PMID: 22897187]
[5] Mouridsen H *et al. N Engl J Med.* 2009 **361**: 766 [PMID: 19692688]
[6] Desta Z *et al. J Pharmacol Exp Ther*. 2004 **310**: 1062 [PMID: 15159443]
[7] Rognan D, *Top Curr Chem.* 2012 **317**: 201 [PMID: 21710380]
[8] de Graaf C *et al. J Med Chem.* 2011 **54**: 8195 [PMID: 22007643]
[9] Yuniarti N *et al. Bioinformation.* 2011 **6**: 164 [PMID: 21572885]
[10] Huang N *et al. J Med Chem.* 2006 **49**: 6789 [PMID: 17154509]
[11] de Kloe *et al. Drug Discov Today.* 2009 **14**: 630 [PMID: 19443265]
[12] Shiau AK *et al. Cell.* 1998 **95**: 927 [PMID: 9875847]
[13] ten Brink T *et al. J Chem Inf Model.* 2009 **49:** 1535 [PMID: 19453150]
[14] Korb O *et al. J Chem Inf Model*. 2009 **49**: 84 [PMID: 19125657]
[15] Marcou G & Rognan D, *J Chem Inf Model.* 2007 **47**: 195 [PMID: 17238265]
[16] O'Boyle NM *et al. J Cheminform.* 2011 **3**: 33 [PMID: 21982300]
[17] Schultes S *et al. Drug Discov Today Technol.* 2010 **7**: e157
[18] Gust R *et al. J Med Chem.* 2001 **44**: 1963 [PMID: 11384241]
[19] Mysinger MM *et al. J Med Chem*. 2012 **55**: 6582 [PMID: 22716043]
[20] Miller CP *et al. J Med Chem.* 2001 **44**: 1654 [PMID: 11356100]

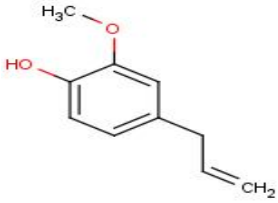
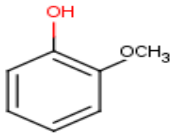
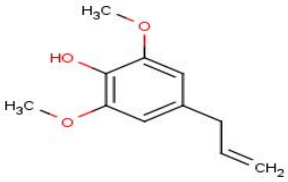
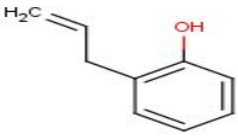
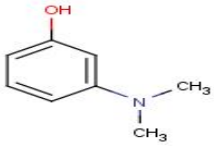
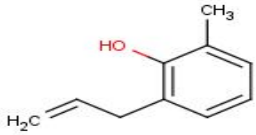
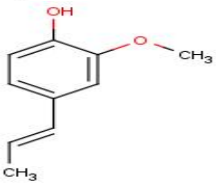
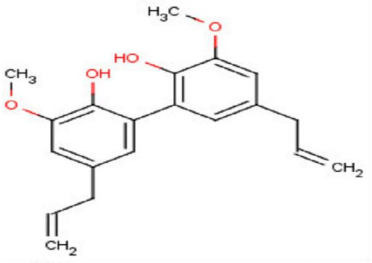
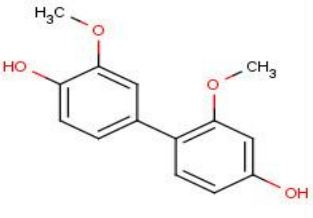
Edited by P Kangueane
Citation: Anita *et al*. Bioinformation 8(19): 901-906 (2012)

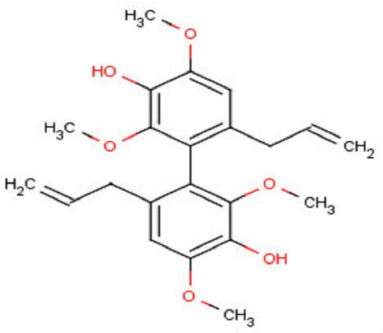
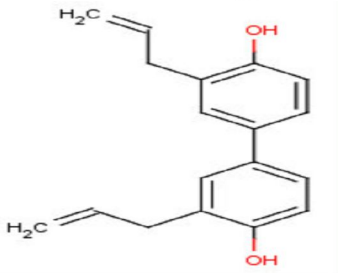
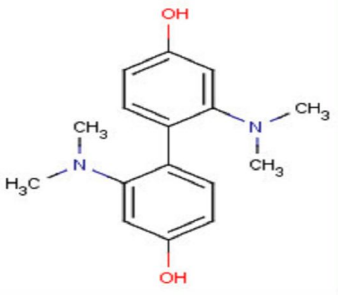
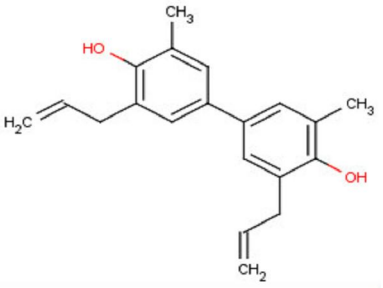
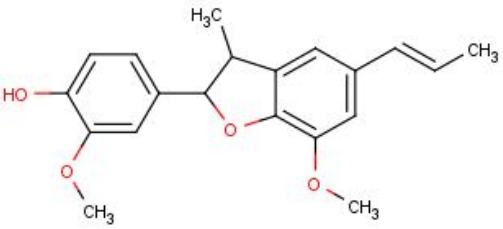
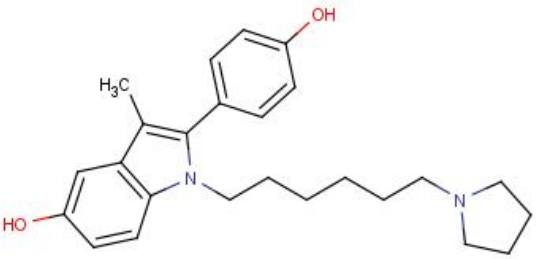
BIOINFORMATION

open access

Supplementary material:

Table 1: The ChemPLP score and the DSLE of eugenol, its analogs, and their dimers resulted from the prospective SBVS campaign.

Sl. no	Compound	ChemPLP Score	DSLE
1		-64.083	5.340
2		-63.075	5.256
3		-65.362	4.668
4		-66.773	6.677
5		-70.628	7.063
6		-64.559	5.869
7		-67.270	5.606
8		-80.487	3.354
9		-73.073	4.060

10		-69.992	2.500
11		-84.502	4.225
12		-82.041	4.102
13		-87.782	3.990
14		-78.046	3.252
15		-117.508	4.052