Trad. Med. J., May – August 2016
Vol. 21(2), p 69-76
ISSN : 1410-5918

Submitted : 11-04-2016
Revised : 02-06-2016
Accepted : 21-07-2016

CONSTRUCTION OF THREE DIMENSIONAL STRUCTURES OF PHYTOESTROGENS CONVERTED FROM SMILES STRING REPRESENTATIONS FOR SIMULATIONS USING PLANTS DOCKING SOFTWARE

KONSTRUKSI STRUKTUR TIGA DIMENSI FITOESTROGEN HASIL KONVERSI DARI *SMILES STRING REPRESENTATIONS* UNTUK SIMULASI PENAMBATAN MOLEKULER MENGGUNAKAN *SOFTWARE* PLANTS

Enade Perdana Istyastono[1,*] and Nunung Yuniarti[2]
[1]Division of Drug Design and Discovery, Faculty of Pharmacy, Sanata Dharma University, Yogyakarta 55282, Indonesia
[2]Department of Pharmacology and Clinical Pharmacy, Faculty of Pharmacy, Universitas Gadjah Mada, Sleman, Yogyakarta 55281, Indonesia

ABSTRACT

Phytoestrogens have some important biological effects and could be employed as medicinal resources, for example as cancer chemopreventive agents. Phytoestrogen is defined as a phytochemical that has estrogenic or anti estrogenic effects. However, there is no database providing comprehensive list of phytoestrogen structures. In computer-aided drug discovery, the database is required to perform virtual screening for retrospective validation and structure-based drug design. The research presented in this article attempted to collect a comprehensive list of phytoestrogen structures. Subsequently, the structures were prepared in their three dimensional structures using SPORES1.3 for molecular docking simulations using PLANTS1.2. The ready-to-dock structures were then stored online as phytoestrogens.zip and could be downloaded from http://molmod.org/phytoestrogens.zip. This database contains 30 ready-to-dock unique phytoestrogens with total of 53 different configurations.
Keywords: *Phytoestrogen, database, drug discovery, molecular docking*

ABSTRAK

Fitoestrogen merupakan senyawa memiliki aktivitas biologis penting bahkan seringkali berpotensi sebagai agen kemoprevensi kanker. Fitoestrogen didefinisikan sebagai senyawa-senyawa alami dari tanaman yang memiliki efek estrogenik atau anti estrogenik. Meskipun demikian, sejauh ini belum ditemui basis data yang menyediakan struktur-struktur fitoestrogen secara komprehensif. Hal ini dapat menghambat penemuan obat berbantukan komputer ketika ketersediaan basis data sangat penting untuk melakukan penapisan virtual guna validasi retrospektif atau perancangan obat berbasis struktur. Penelitian yang disampaikan di artikel ini mencoba mengumpulkan struktur-struktur fitoestrogen secara komprehensif dan melakukan preparasi untuk membuat struktur-struktur tersebut dalam bentuk tiga dimensi dengan aplikasi SPORES1.3 dan siap menjadi berkas masukan untuk simulasi penambatan molekuler PLANTS1.2. Struktur-struktur siap ditambatkan tersebut kemudian disimpan dalam jaringan sebagai phytoestrogens.zip dan bisa diunduh di http://molmod.org/phytoestrogens.zip. Database ini memuat 30 struktur fitoestrogen dengan total 53 konfigurasi yang berbeda yang siap menjadi input simulasi penambatan molekuler.
Kata kunci: *Fitoestrogen, basis data, penemuan obat, penambatan molekuler*

INTRODUCTION

Drugs discoveries have owed phyto-chemicals since they have become main resources for inspirations or even lead compounds in the early stages of the discovery and development (Brusotti *et al.*, 2014; Efferth and Koch, 2011). Currently, drug discovery targeting breast cancer has become of timely and considerable interest since there is no recent breakthrough in the therapy (Cardoso and Senkus, 2015). Screening campaigns on phytochemicals to discover breast cancer chemopreventive agents have been performed, since it is believed to be able to serve

Corresponding Author : Enade Perdana Istyastono
Email : enade@usd.ac.id

as starting points to discover novel breast cancer drugs (Balunas *et al.*, 2008; Helferich *et al.*, 2008; Istyastono, 2015; Setiawati *et al.*, 2014[a]; Shibata *et al.*, 2011; Vidhya and Devaraj, 2011). In the molecular level, one of the main targets to discover novel breast cancer chemopreventive agents is estrogen receptor α (ERα) (Ali and Coombes, 2000; Anita *et al.*, 2015, 2012; Istyastono, 2015; Radifar *et al.*, 2013; Setiawati *et al.*, 2014[a]; Shiau *et al.*, 1998).

Phytochemicals showed that estrogenic or anti estrogenic effects are, by definition, classed as phytoestrogens (Dixon, 2004; Helferich *et al.*, 2008). As ligands for estrogen receptors, these phytochemicals could show the estrogenic or anti estrogenic effects (Helferich *et al.*, 2008). Certain isoflavonoids, flavonoids, stilbenes, and lignans were identified as phytoestrogens (Dixon, 2004). Notably, the best studied phytoestrogen is the soy isoflavone genistein (Dixon, 2004; Helferich *et al.*, 2008; Varinska et al., 2015). The interests on phytoestrogen have emerged not only related to their potency as breast cancer chemopreventive agents (Balunas *et al.*, 2008; Helferich *et al.*, 2008; Setiawati *et al.*, 2014[a]; Setiawati *et al.*, 2014[b]), but also their roles in osteoporosis, cognitive function, cardiovascular disease, immunity and inflammation, and reproduction (Dixon, 2004). However, although many phytochemicals were identified as phytoestrogens (Dixon, 2004; Helferich *et al.*, 2008), the comprehensive list of phytoestrogen is not available yet. The lack of this list hinders the attempts to thoroughly study how they bind to the receptors especially to ERα (Istyastono, 2015; Pathania *et al.*, 2015; Setiawati *et al.*, 2014[a]), which in turn could efficiently assist the discovery of novel ligands for the receptors (Istyastono *et al.*, 2015[a]; Istyastono *et al.*, 2015[b]; Kufareva *et al.*, 2014; Pathania *et al.*, 2015).

This article presents the attempts to collect the structures of known phytoestrogens. The collected structures were virtually prepared using previously published ligand preparation procedures (Istyastono and Setyaningsih, 2015) to be ready for molecular docking simulations. This involved the generations of their three-dimensional (3D) structures (Istyastono and Setyaningsih, 2015; Setiawati *et al.*, 2014[a]). The structures in their ready-to-dock formats were subsequently stored online as phytoestrogens.zip and could be obtained from http://molmod.org/phytoestrogens.zip for further usages. Additionally, this article presents also the molecular docking studies of genistein as the best studied phytoestrogen in the ERα binding pocket. The atomic interactions of the resulted pose to the ERα binding pocket were then compared to the interactions of the pose of the co-crystal ligand 4-hydroxytamoxifen (4-OHT).

METHODOLOGY

Materials and Instrumentation

The phytochemicals mentioned in Dixon (2004) were used as the starting points. The ZINC and ChEMBL databases were the main references to obtain the phytoestrogen structure (Bento *et al.*, 2014; Irwin *et al.*, 2012). The crystal structure of ERα obtained from the protein data bank (PDB) with PDB id of 3ERT (Shiau *et al.*, 1998) was used as the reference structure for molecular docking simulations. Files **plants.config**, **protein.mol2**, and **water.mol2** to perform the simulations were obtained from Anita *et al.* (2012).

Computational medicinal chemistry applications employed in this research were SPORES version 1.3 or SPORES1.3 (ten Brink and Exner, 2009), PLANTS version 1.2 or PLANTS1.2 (Korb *et al.*, 2009; Korb *et al.*, 2007), Open Babel 2.2.3 (O'Boyle *et al.*, 2011), and PyMOL 1.2r1 (Lill and Danielson, 2011). All computations were performed in a Virtual Private Server (VPS) with internet protocol (IP) address of 103.247.10.66. The server has Ubuntu 12.04.3 LTS as the operating system, Intel(R) Xeon(R) E5630 Quad Core @ 2.53GHz as the processor and 1 GB of RAM.

Computational Details

The SMILES format of all compounds with estrogenic or anti estrogenic activity mentioned in Dixon (2004) were obtained from either ZINC database (Irwin *et al.*, 2012) or ChEMBL database (Bento *et al.*, 2014). Each compound was then subjected to Open Babel 2.2.3 conversion software to be converted in its three dimensional (3D) format as a **mol2** file. The ***reprot*** module in SPORES1.3 was subsequently employed to properly check and assign the **mol2** file into a proper **mol2** file with standard protonated state (ten Brink and Exner, 2009). This **mol2** file was ready to be docked by using PLANTS1.2 docking software. The ready-to-dock **mol2** files were compressed to a **zip** file and the **zip** file was subsequently uploaded to http://molmod.org/.

The file was subsequently downloaded to the VPS and unzipped. The **mol2** structures were then docked into ERα binding pocket employing PLANTS1.2 molecular docking software by following the previously published procedures by Anita *et al.* (2012) to examine if the structures were the ready-to-dock ones. The docking results were compressed to a **zip** file and the **zip** file was also uploaded to http://molmod.org/.

Table I. Name and structures (in SMILES format) of isoflavonoid phytoestrogens mentioned by Dixon (2004).

ZINC No.[a)]	Structure in SMILES format[a)]	Name[b)]
18847037	COc1ccc(cc1)c2coc3cc(cc(c3c2=O)O)O	biochanin A
1219	c1cc2c(cc1O)oc3c2c(=O)oc4c3ccc(c4)O	coumestrol
18847034	c1cc(ccc1c2coc3cc(ccc3c2=O)O)O	daidzein
4098610	c1cc(ccc1c2coc3cc(ccc3c2=O)O[C@H]4[C@@H]([C@H]([C@@H]([C@H](O4)CO)O)O)O)O	daidzin [c)]
388660	c1cc(ccc1[C@H]2Cc3ccc(cc3OC2)O)O	equol
388661	c1cc(ccc1[C@@H]2Cc3ccc(cc3OC2)O)O	equol
18847036	COc1ccc(cc1)c2coc3cc(ccc3c2=O)O	formononetin
18825330	c1cc(ccc1c2coc3cc(cc(c3c2=O)O)O)O	genistein
4097913	c1cc(ccc1c2coc3cc(cc(c3c2=O)O)O[C@H]4[C@@H]([C@H]([C@@H]([C@H](O4)CO)O)O)O)O	genistin[c)]
33831433	CC1(C=Cc2c(ccc(c2O)C3=Cc4ccc(cc4OC3)O)O1)C	glabrene
5963549	CC1(C=Cc2c(ccc(c2O1)C3=Cc4ccc(cc4OC3)O)O)C	glabrene
4098719	CC1(C=Cc2c(ccc3c2OC[C@H](C3)c4ccc(cc4O)O)O1)C	glabridin
899797	CC1(C=Cc2c(ccc3c2OC[C@@H](C3)c4ccc(cc4O)O)O1)C	glabridin
5999205	COc1cc2c(cc1O)occ(c2=O)c3ccc(cc3)O	glycitein
3872928	CC(C)Oc1ccc2c(=O)cc(oc2c1)c3ccccc3	ipriflavone
4016	CC(C)Oc1ccc2c(c1)occ(c2=O)c3ccccc3	ipriflavone
14727569	CC(=CCc1c(ccc(c1O)[C@@H]2Cc3c(cc(c(c3OC)CC=C(C)C)O)OC2)O)C	licoricidin
5854565	CC(=CCc1c(ccc(c1O)[C@H]2Cc3c(cc(c(c3OC)CC=C(C)C)O)OC2)O)C	licoricidin
59729510	CC(C)/C=C/c1c(ccc(c1O)[C@@H]2Cc3c(cc(c(c3O)CC=C(C)C)OC)OC2)O	licoricidin
59729511	CC(C)/C=C/c1c(ccc(c1O)[C@H]2Cc3c(cc(c(c3O)CC=C(C)C)OC)OC2)O	licoricidin
985403	c1cc(ccc1[C@@H]2CC(=O)c3ccc(cc3O2)O)O	liquiritigenin
156701	c1cc(ccc1[C@@H]2CC(=O)c3c(cc(cc3O2)O)O)O	naringenin
1785	c1cc(ccc1[C@H]2CC(=O)c3c(cc(cc3O2)O)O)O	naringenin
4098745	c1cc(ccc1c2coc3c(c2=O)ccc(c3[C@H]4[C@@H]([C@H]([C@@H]([C@H](O4)CO)O)O)O)O)O	puerarin [c)]

[a)]One compound with more than one possible 3D configuration could have more than one structure in SMILES format (Bento *et al.*, 2014; Irwin *et al.*, 2012); [b)]As mentioned by Dixon (2004); [c)]For glycosides, only one configuration is included in the list since they have more than 5 different 3D configurations.

Visual inspections of the best pose of genistein resulted from the molecular docking simulations using PyMOL (Lill and Danielson, 2011) was then performed to analyze the plausible molecular determinants of ERα-phytoestrogen binding.

RESULTS AND DISCUSSION

According to Dixon (2004), phytoestrogens were classified into isoflavonoid (Table I), flavonoid (Table II) and lignan (Table III) phytoestrogens. Beside those three main classes, there were other phytoestrogens, including chalcone and anthraquinone phytoestrogens (Table IV). Notably, there are still some other phytochemicals which have estrogenic or anti estrogenic effects but not listed in those tables (Tables I-IV), for example eugenol, alizarin, α-mangostin (Matsuda *et al.*, 2001; Shibata *et al.*, 2011). Nevertheless, to construct a comprehensive phytoestrogen structure database requires starting points (Pathania *et al.*, 2015) and the review on phytoestrogens by Dixon (2004) could serve as the starting points since it has covered most phytoestrogen classes and discussed the potent ones (Dixon, 2004; Helferich *et al.*, 2008; Mysinger *et al.*, 2012).

In total, there were 30 unique phytoestrogens presented in Tables I-IV. Since one compound could have more than one 3D configuration, there were 53 structural configurations indicated by the SMILES formats in Tables I-IV. ChEMBL database (Bento et al., 2014) was subsequently inspected to append the information with the biological activities of the listed phytoestrogens at ERα. The results are presented in Table V. Only 9 phytoestrogens with biological activities at ERα are listed (Table V and Figure 1).

Table II. Name and structures (in SMILES format) of flavonoid phytoestrogens mentioned by Dixon (2004).

ZINC No.[a)]	Structure in SMILES format[a)]	Name[b)]
3873159	CC(=CCc1c(c(c2c(c1O)C(=O)C[C@H](O2)c3ccc(cc3)O)CC=C(C)C)O)C	6,8-diprenylnaringenin
4098363	CC(=CCc1c(cc2c(c1O)C(=O)C[C@H](O2)c3ccc(cc3)O)O)C	6-prenylnaringenin
5854068	CC(=CCC/C(=C/Cc1c(cc(c2c1O[C@@H](CC2=O)c3ccc(cc3)O)O)O)/C)C	8-geranylnaringenin
39451	CC(=CCc1c(cc(c2c1O[C@H](CC2=O)c3ccc(cc3)O)O)O)C	8-prenylnaringenin
39452	CC(=CCc1c(cc(c2c1O[C@@H](CC2=O)c3ccc(cc3)O)O)O)C	8-prenylnaringenin
12353732	c1cc(ccc1/C=C\c2cc(cc(c2)O)O)O	resveratrol[c)]
6787	c1cc(ccc1/C=C/c2cc(cc(c2)O)O)O	resveratrol[c)]
2008845	CC(=CC[C@@H](Cc1c(cc(c2c1O[C@H](CC2=O)c3ccc(cc3O)O)O)O)C(=C)C)C	sophoraflavanone_G
2008847	CC(=CC[C@H](Cc1c(cc(c2c1O[C@H](CC2=O)c3ccc(cc3O)O)O)O)C(=C)C)C	sophoraflavanone_G
2008848	CC(=CC[C@@H](Cc1c(cc(c2c1O[C@@H](CC2=O)c3ccc(cc3O)O)O)O)C(=C)C)C	sophoraflavanone_G
2008850	CC(=CC[C@H](Cc1c(cc(c2c1O[C@@H](CC2=O)c3ccc(cc3O)O)O)O)C(=C)C)C	sophoraflavanone_G

[a)]One compound with more than one possible 3D configuration could have more than one structure in SMILES format (Bento *et al.*, 2014; Irwin *et al.*, 2012); [b)]As mentioned by Dixon (2004); [c)]Resveratrol was classified here since the biosynthesis was related to flavonoids (Dixon, 2004).

Table III. Name and structures (in SMILES format) of lignan phytoestrogens mentioned by Dixon (2004).

ZINC No.[a)]	Structure in SMILES format[a)]	Name[b)]
1360	c1cc(cc(c1)O)C[C@@H]2COC(=O)[C@H]2Cc3cccc(c3)O	enterolactone
388657	c1cc(cc(c1)O)C[C@H]2COC(=O)[C@H]2Cc3cccc(c3)O	enterolactone
388658	c1cc(cc(c1)O)C[C@H]2COC(=O)[C@@H]2Cc3cccc(c3)O	enterolactone
388659	c1cc(cc(c1)O)C[C@@H]2COC(=O)[C@@H]2Cc3cccc(c3)O	enterolactone
4098820	COc1cc(ccc1O)C[C@H]2CO[C@@H]([C@H]2CO)c3ccc(c(c3)OC)O	lariciresinol
1595956	COc1cc(ccc1O)C[C@H]2COC(=O)[C@H]2Cc3ccc(c(c3)OC)O	matairesinol
1595957	COc1cc(ccc1O)C[C@H]2COC(=O)[C@@H]2Cc3ccc(c(c3)OC)O	matairesinol
1595958	COc1cc(ccc1O)C[C@@H]2COC(=O)[C@@H]2Cc3ccc(c(c3)OC)O	matairesinol
900187	COc1cc(ccc1O)C[C@@H]2COC(=O)[C@H]2Cc3ccc(c(c3)OC)O	matairesinol
4098921	COc1cc(ccc1O)[C@@H]2[C@H]3CO[C@@H]([C@H]3CO2)c4ccc(c(c4)OC)O	pinoresinol
2020114	COc1cc(ccc1O)C[C@@H](CO)[C@@H](Cc2ccc(c(c2)OC)O)CO	secoisolariciresinol
14435411	COc1cc(cc(c1O)OC)[C@@H]2[C@@H]3CO[C@H]([C@H]3CO2)c4cc(c(c(c4)OC)O)OC	syringaresinol
18283077	COc1cc(cc(c1O)OC)[C@@H]2[C@@H]3CO[C@@H]([C@H]3CO2)c4cc(c(c(c4)OC)O)OC	syringaresinol
5998957	COc1cc(cc(c1O)OC)[C@@H]2[C@@H]3CO[C@@H]([C@@H]3CO2)c4cc(c(c(c4)OC)O)OC	syringaresinol
5998959	COc1cc(cc(c1O)OC)[C@@H]2[C@@H]3CO[C@H]([C@@H]3CO2)c4cc(c(c(c4)OC)O)OC	syringaresinol

[a)]One compound with more than one possible 3D configuration could have more than one structure in SMILES format (Bento *et al.*, 2014; Irwin *et al.*, 2012); [b)]As mentioned by Dixon (2004).

Table IV. Name and structures (in SMILES format) of other classes (chalcone and anthraquinone) phytoestrogens mentioned by Dixon (2004).

ZINC No.[a)]	Structure in SMILES format[a)]	Name[b)]
3860201	c1cc2c(cc1O)C(=O)c3ccc(cc3C2=O)O	2,6-dihydroxyanthraquinone
3824868	Cc1cc2c(c(c1)O)C(=O)c3c(cc(cc3O)O)C2=O	emodin
3869608	c1cc(ccc1/C=C/C(=O)c2ccc(cc2O)O)O	isoliquiritigenin

[a)]One compound with more than one possible 3D configuration could have more than one structure in SMILES format (Bento *et al.*, 2014; Irwin *et al.*, 2012); [b)]As mentioned by Dixon (2004).

Table V. Biological activities (IC_{50}) at ERα of some listed phytoestrogens (Bento *et al.*, 2014).

Name	ChEMBL[a)]		IC_{50} (nM)[b)]
	Compound Identity	Assay Identity	
Coumestrol	CHEMBL30707	CHEMBL827788	76
Daidzein	CHEMBL8145	CHEMBL1219560	450
Genistein	CHEMBL44	CHEMBL891640	395
Ipriflavone	CHEMBL165790	CHEMBL1909145	Not active
Naringenin	CHEMBL9352	CHEMBL909843	75,302
8-prenylnaringenin	CHEMBL460647	CHEMBL909843	57
Resveratrol	CHEMBL165	CHEMBL1909145	Not active
2,6-Dihydroxyanthraquinone	CHEMBL298398	CHEMBL679893	1100
Emodin	CHEMBL289277	CHEMBL679893	2700

[a)]Ref: (Bento *et al.*, 2014); [b)]If there were more than one biological activities listed in the database, the most recent one was selected.

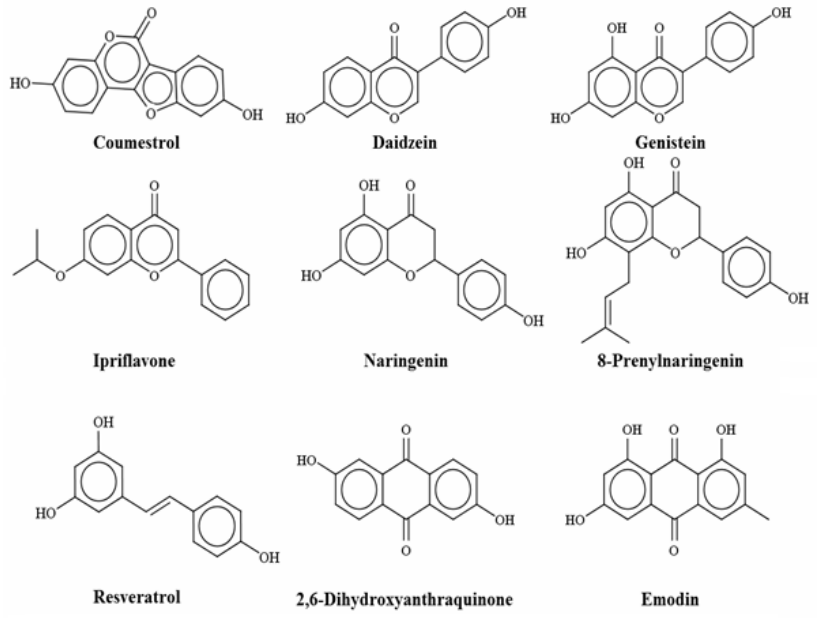

Figure 1. Stuctures (in two-dimensional format converted from the SMILES format) of some phytoestrogens (Table V).

Remarkably, all listed active phytoestrogens shared similar pharmacophore, *i.e.* they had hydroxyl phenolic and carbonyl moieties separated by 5 heavy atoms (Table V and Figure 1).

For the purpose of creating the database of the ready-to-dock phytoestrogens, an **smi** file named **phytoestrogens.smi** containing all listed structures in Tables I-IV was created. The structure of the file is as follows:

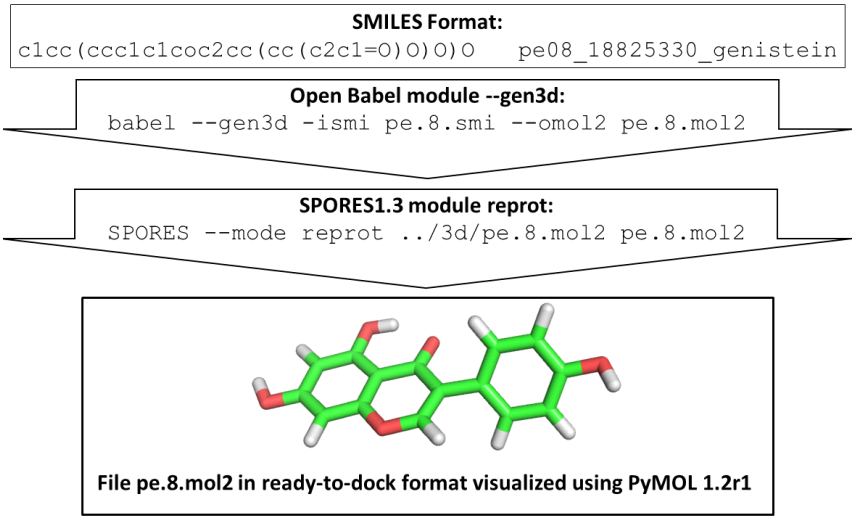

Figure 2. Flowchart of the ligand preparation process suggested in this article with genistein as the representative.

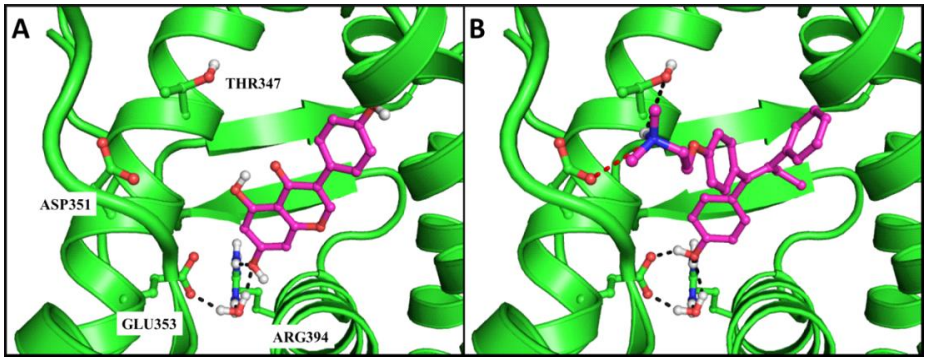

Figure 3. Binding poses genistein (A) and the co-crystal ligand 4-hydroxy tamoxifen (B) in the ERá binding pocket (Anita *et al.*, 2012; Shiau *et al.*, 1998). The ERá is presented in the cartoon mode with carbon atoms are in green, while the ligands are presented in the sticks mode with carbon atoms are in magenta. 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 (Anita *et al.*, 2012) are presented for the sake of clarity. Nitrogen and oxygen atoms are presented in blue and red, respectively. These figures adopted the similar figures published by Setiawati *et al.* (2014)

[SMILES] molmodID_zincID_name. For example, biochanin A (Table I) was written in the first line of **phytoestrogens.smi** as follows: COc1ccc(cc1)c2coc3cc(cc(c3c2=O)O)O pe01_18847037_biochaninA. Following the procedures aforementioned resulted in a **zip** file with size of 70 kilobytes. This file has been uploaded and attached to http://molmod.org/ and could therefore be obtained from http://molmod.org/phytoestrogens.zip. In order to check the readiness of the uploaded ready-to-dock **mol2** file, further molecular docking simulations employing previously published procedure by Anita et al. (2012) were performed. The docking simulations ran without giving any error messages. This confirmed that the uploaded input files were the ready-to-dock structures. For further analysis by public, the docking results were zipped in a **zip** file with size of 12 megabytes and the **zip** file was uploaded to http://molmod.org/ and can be obtained from http://molmod.org/phytoestrogens.PLANTS.zip.

The best studied phytoestrogen genistein (Dixon, 2004; Helferich *et al.*, 2008; Varinska *et al.*,

2015) was selected as the representative for visual inspections. The process of transforming genistein from its SMILES format to the ready-to-dock 3D format is depicted in Figure 2. The best docking pose of genistein in the ERα binding pocket is presented in Figure 3.

Based on Figures 2 and 3, it could be concluded that the resulted **mol2** files were ready to be docked using PLANTS1.2. Moreover, by examining the binding poses of genistein and 4-OHT in the ERα binding pocket (Figure 3), the important binding interactions could be identified: the hydrogen bond network between genistein, GLU353, ARG394, and the conserved water molecule (Anita *et al.*, 2012; Shiau *et al.*, 1998). This is in line with the identified pharmacophore by examining Figure 1. Notably, the aromatic moieties of genistein align very similarly to the aromatic moieties of 4-OHT when they bind to ERα. Thus GLU353, ARG394, and the aromatic residue around genistein in the ERα binding pocket are highly suggested as the molecular determinants of ERα-ligand binding. The most plausible aromatic residue as the molecular determinant in the ERα binding pocket is TRP383 (Istyastono, 2015).

CONCLUSION

The database of ready-to-dock 3D structures of phytoestrogens resulted in this research could be employed further to study the atomic level of their estrogenic or anti estrogenic activities. Genistein as the representative phytoestrogen employed further in molecular docking simulations into ERα has unraveled some the molecular determinants in the ERα-ligand binding: GLU353, TRP383 and ARG394.

ACKNOWLEDGMENT

This work was financially supported by the Hibah Kompetensi Research Grant from the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia. The authors thank Molecular Modeling Center "MOLMOD.ORG" for providing space in its website http://molmod.org/ to upload the database resulted in this research and C.V. Digikom Multimedia Pratama (http://digikom.co.id/) for providing instruments used in the research.

REFERENCES

Ali, S. and Coombes, R.C. 2000. Estrogen receptor alpha in human breast cancer: occurrence and significance. *J. Mammary Gland Biol. Neoplasia* 5: 271–281.

Anita, Y., Radifar, M., Kardono, L., Hanafi, M. and Istyastono, E.P. 2012. Structure-based design of eugenol analogs as potential estrogen receptor antagonists. *Bioinformation* 8: 901–906.

Anita, Y., Sundowo, A., Dewi, N.L.P., Filailla, E., Mulyani, H., Risdian, C., Banjarnahor, S., Hanafi, M. and Istyastono, E.P. 2015. Biotransformation of eugenol to dehydroeugenol catalyzed by *Brassica juncea* peroxidase and its cytotoxicity activities. *Procedia Chem.* 16: 265–271.

Balunas, M.J., Su, B., Brueggemeier, R.W. and Kinghorn, A.D. 2008. Xanthones from the botanical dietary supplement mangosteen (*Garcinia mangostana*) with aromatase inhibitory activity. *J. Nat. Prod.* 71: 1161–1166.

Bento, A.P., Gaulton, A., Hersey, A., Bellis, L.J., Chambers, J., Davies, M., Krüger, F.A., Light, Y., Mak, L., McGlinchey, S., Nowotka, M., Papadatos, G., Santos, R. and Overington, J.P. 2014. The ChEMBL bioactivity database: An update. *Nucl. Acids Res.* 42: 1083–1090.

Brusotti, G., Cesari, I., Dentamaro, A., Caccialanza, G. and Massolini, G. 2014. Isolation and characterization of bioactive compounds from plant resources: The role of analysis in the ethnopharmacological approach. *J. Pharm. Biomed. Anal.* 87: 218–228.

Cardoso, F. and Senkus, E., 2015. Breast cancer in 2014: A call back to reality! *Nat. Rev. Clin. Oncol.* 12, 67–68.

Dixon, R.A. 2004. Phytoestrogens. *Annu. Rev. Plant Biol.* 55: 225–261.

Efferth, T. and Koch, E. 2011. Complex interactions between phytochemicals. The multi-target therapeutic concept of phytotherapy. *Curr. Drug Targets* 12: 122–132.

Helferich, W.G., Andrade, J.E. and Hoagland, M.S. 2008. Phytoestrogens and breast cancer: a complex story. *Inflammopharmacology* 16: 219–26.

Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. and Coleman, R.G., 2012. ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model.* 52: 1757–1768.

Istyastono, E.P. 2015. Employing recursive partition and regression tree method to increase the quality of structure-based virtual screening in the estrogen receptor alpha ligands identification. *Asian J. Pharm. Clin. Res.* 8: 21–24.

Istyastono, E.P., Kooistra, A.J., Vischer, H., Kuijer, M., Roumen, L., Nijmeijer, S., Smits, R., de Esch, I., Leurs, R. and de Graaf, C., 2015[a]. Structure-based virtual screening for fragment-like ligands of the G protein-coupled histamine H_4 receptor. *Med. Chem.*

Commun. 6: 1003–1017.

Istyastono, E.P., Riswanto, F.D.O., and Yuliani, S.H. 2015[b]. Computer-aided drug repurposing: a cyclooxygenase-2 inhibitor celecoxib as a ligand for estrogen receptor alpha. *Indones. J. Chem.* 15: 274–280.

Istyastono, E.P. and Setyaningsih, D., 2015. Construction and retrospective validation of structure-based virtual screening protocols to identify potent ligands for human adrenergic β_2 receptor. *Indones. J. Pharm.* 26: 20–28.

Korb, O., Stützle, T. and Exner, T.E., 2007. An ant colony optimization approach to flexible protein–ligand docking. *Proc. IEEE Swarm Intell. Symp.* 1: 115–134.

Korb, O., Stützle, T. and Exner, T.E., 2009. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.* 49: 84–96.

Kufareva, I., Katritch, V., Stevens, R.C. and Abagyan, R. 2014. Advances in GPCR modeling evaluated by the GPCR DOCK 2013 assessment: Meeting new challenges. *Structure* 22: 1120–1139.

Lill, M.A. and Danielson, M.L. 2011. Computer-aided drug design platform using PyMOL. *J. Comput. Aided Mol. Des.* 25: 13–19.

Matsuda, H., Shimoda, H., Morikawa, T. and Yoshikawa, M. 2001. Phytoestrogens from the roots of *Polygonum cuspidatum* (Polygonaceae): Structure-requirement of hydroxyanthraquinones for estrogenic activity. *Bioorg. Med. Chem. Lett.* 11: 1839–1842.

Mysinger, M.M., Carchia, M., Irwin, J.J. and Shoichet, B.K. 2012. Directory of Useful Decoys, Enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.* 55: 6582–6594.

O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., and Hutchison, G.R. 2011. Open Babel: An open chemical toolbox. *J. Cheminform.* 3: 33–47.

Pathania, S., Ramakrishnan, S.M. and Bagler, G. 2015. Phytochemica: A platform to explore phytochemicals of medicinal plants. *Database* 10.1093/database/bav075: 1–8.

Radifar, M., Yuniarti, N., Istyastono, E.P. 2013. PyPLIF: Python-based protein-ligand interaction fingerprinting. *Bioinformation* 9: 325–328.

Setiawati, A., Riswanto, F.D.O., Yuliani, S.H. and Istyastono, E.P. 2014[a]. 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. *Indo. J. Chem.* 14: 103–108.

Setiawati, A., Riswanto, F.O.D., Yuliani, S.H. and Istyastono, E.P. 2014[b]. Anticancer activity of mangosteen pericarp dry extract against MCF-7 breast cancer cell line though estrogen receptor-α. *Indones. J. Pharm.* 25: 119–124.

Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A. and Greene, G.L. 1998. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 95: 927–937.

Shibata, M.-A., Iinuma, M., Morimoto, J., Kurose, H., Akamatsu, K., Okuno, Y., Akao, Y. and Otsuki, Y. 2011. α-Mangostin extracted from the pericarp of the mangosteen (*Garcinia mangostana* Linn) reduces tumor growth and lymph node metastasis in an immunocompetent xenograft model of metastatic mammary cancer carrying a p53 mutation. *BMC Med.* 9: 69.

ten Brink, T. and Exner, T.E., 2009. Influence of protonation, tautomeric, and stereoisomeric states on protein-ligand docking results. *J. Chem. Inf. Model.* 49: 1535–1546.

Varinska, L., Gal, P., Mojzisova, G., Mirossay, L. and Mojzis, J. 2015. Soy and breast cancer: Focus on angiogenesis. *Int. J. Mol. Sci.* 16: 11728–11749.

Vidhya, N. and Devaraj, S.N. 2011. Induction of apoptosis by eugenol in human breast cancer cells. *Indian J. Exp. Biol.* 49: 871–878.