The in Silico Study of Nutmeg Seeds (Myristica fragrans Houtt) as Peroxisome Proliferator Activated Receptor Gamma Activator Using 3D-QSAR Pharmacophore Modelling

M. Muchtaridi1*, Karen Low1, Keri Lestari2
1Department of Pharmaceutical Analysis and Medicinal Chemistry, Faculty of Pharmacy Universitas Padjadjaran, Bandung, Indonesia. 2Department of Pharmacology and Clinical Pharmacy, Faculty of Pharmacy Universitas Padjadjaran, Bandung, Indonesia.

ARTICLE INFO
Article history:
Received on: 19/01/2016
Revised on: 14/04/2016
Accepted on: 17/06/2016
Available online on: 26/09/2016

Key words:
Anti-diabetic, Hypogen, pharmacophore modelling, PPARγ.

ABSTRACT
In this study, we created a pharmacophore models from a dataset of agonists for PPAR gamma receptor using the Catalyst/Hypogen module. A training set consists of 22 compounds activity range between 0.1 to 3,500 nM, were carefully selected. In previous study, molecular docking of macelignan against PPARγ binding pocket showed a free energy binding of -11.07 kJ/mol, interaction with the hydrophobic pocket (diphenyl pocket)(Celik et al., 2007), and a hydrogen bond network (His323, Tyr473, His449 and Ser289). The pharmacophore model (Hypo1), consisting of 5 features, i.e. one hydrogen bond acceptor (HBA), negative ionizable (NI), ring aromatic (RA) and two hydrophobics (HY) features, and one excluded volume. Hypo1 has the lowest total cost value (92.055), the highest cost difference (40.9316), the lowest RMSD (0.591049), and the best correlation coefficient (0.972949). Fourteen natural substances reported from nutmeg seeds (Myristica fragrans HOUTT.) were then mapped against Hypo1, and macelignan shows a fair fit value of 7.00102 with an estimated value of 1271.990 nM. This concludes, macelignan in nutmeg might have antidiabetic properties via PPARγ receptor activation.

INTRODUCTION
Peroxisome Proliferator-Activated Receptors (PPARs) are members of the nuclear receptor super-family that is involved in protein gene expression for energy, glucose and lipid metabolism, proliferation and differentiation of adipocyte, as well as insulin sensitivity (Arck et al., 2010; Hiukka et al., 2010; Farce et al., 2009). Three isoforms of these core receptors have been identified so far: PPARα, PPARβ/δ dan PPARγ. Each of them has different selectivity towards type of tissue, type of ligand, and finally, unique biological response. PPARγ is mostly expressed in adipose tissue, colon and macrophages, and has been proven for its role in regulating carbohydrate metabolism as well as fatty acid storage. Finding a new and poten ligand for PPARγ is the main focus in research for treating patients with Type II diabetes mellitus, as it can restore insulin sensitivity. PPARγ agonist has made a big attraction in the clinical management of cardiovascular risk factors associated with metabolic syndrome and Type 2 diabetes mellitus (Pearson, 2009). A few classes of PPARγ agonist have been reported to possess anti-diabetic properties, such as: thiazolidindiones, dihydrobenzofurans, dihydro-benzopyrans, benzofuran benoxazoles and α-amino-β-phenylpropanoic acid derivatives (Henke, 2004b). Thiazolidinediones (TZDs) is an important class of PPARγ synthetic agonist (Feldman et al., 2008).

TZDs are anti-diabetic agents that work on adipose tissue and increase insulin sensitivity. At this point, it is still used in the effort to treat Type II diabetes mellitus (Pourcet et al., 2006). Despite the clinical benefit of TZDs, they are related with side effects such as an increased body weight, an increased adipogenesis, kidneys fluid retention, and possible increase in cardiovascular disorder occurrences (Jones, 2010).
Thus, new PPARγ activator with an increased therapeutic efficacy and lesser side effects is necessary. Partial PPAR agonist is recently popular as a promising and new generation of PPAR activator. Partial agonist induces alternative receptor confirmation and as such, recruits different co-activator, which results in a different transcription effect when compared with TZDs. Therefore, it is indicated that such partial agonists will probably offer the required effectiveness and reduce side effects (Chang et al., 2007; Yumuk, 2006). Natural products are an important and promising source in the discovery of new potent drugs (Newman and Cragg, 2007). Here, we created a pharmacophore models from a dataset of agonists for PPARγ receptor using the Catalyst/Hypogon module for predicting the nutmeg seeds compounds. In the previous study, Najjar et al. (2011) generated the pharmacophoric space of PPARg using seven diverse sets of activators (Al-Najjar et al., 2011). However, the compounds screened was not natural compounds. Thus, the purpose of this study was to identify the functional groups or features that activate the human PPARγ as agonist. The approach with identification and modelling of pharmacophore at 2-aryloxy-3-phenyl-propanoic acid has been done with in silico method.

MATERIAL AND METHOD

Selecting Training Set Compound

Choosing the training set is important for generating a hypothesis in the Hypogon and has to follow certain ground rules, such as a minimum of 16 varieties of structural compound has to be chosen to avoid by chance of correlation. Activity data was limited at least 4 magnitude order. The selected compounds should have simple and clear information. Any redundancy (extreme) in the structural features or in the activity data must be avoided. Inactive compound due to steric factor is unsuitable to be used in the training set because Catalyst is limited to accommodate that such cases.

Data Set mining

The number of 117 ligands were collected from some literatures as dataset that substraction as training set and external set that has activity against PPARγ (Martres et al., 2008; Henke, 2004a; Koyama et al., 2004; Shibata et al., 2012; Oon Han et al., 2007). The database were chosen based on similar bioassay method as shown in Fig. 5. Activity values were presented as Effective Concentration, or EC₅₀.

Dataset was then divided into training set and test set. The Most Active (MA) compound gives an important information for pharmacophore modeling and a few Moderately Active (M) compound as well as Less Active (L) compound are also important in spreading the range for activity value as large as possible. A number of 20 ligands that fulfilled the criteria were used as training set and the remaining 25 compounds were selected as a test set based on the formula of 1.1 and 1.2. Training set was selected using formula (1), while the external set was chosen beside of training set. For the training set data, attribute of uncertainty was set up to show uncertain value with a score of 3 on the spreadsheet. This means, the activity of the actual ligand in the training set would be in the range of 1/3 to 3 times from the ligand’s activity.

Activity range of PPARγ compound is classified as followed: Most Active (<0.9nM), Moderate (0.9nM-316nM) and Less Active (>316nM). Compound’s activity is determined by calculation with the formula:

\[ \text{MA} \times \text{UncMA} - \frac{A}{\text{UncA}} > 0.0 \]  

(1.1)

\[ \log (A) - \log (\text{MA}) > 3.5 \]  

(1.2)

Description: MA = compound with highest activity (lowest EC₅₀); UncMA=uncertainty of activity is measured and A= activity from active compound (Abu Hammad and Taha, 2009).

Pharmacophore Modeling

Generation of Conformation Library of Bioactive Compounds

For the training and test sets molecules, conformational models representing their available conformational space were calculated. All molecules were built using the 2D and 3D sketcher of Hyperchem 7.0, and optimized using MM2 in Hyperchem 7.0. A conformational set was generated for each molecule using the poling algorithm and the best energy option, based on CHARMmm force field from Discovery Studio 2.5 (Brooks et al., 1983; Musfiroh et al., 2013). The molecules associated with their conformational models were mapped onto the pharmacophore model using the “best fit” option to obtain the bioactive conformation of each molecule.

Pharmacophore Models

A number of 255 Conformers was chosen to be minimized as best conformation, and 20 kcal/mol was set as energy threshold as global energy minimum for conformational searching (Wang et al., 2008), this protocol is available in DS 2.5 packages. The best pharmacophore models was validated according to Deng et al. (Deng et al., 2008) in terms of cost functions and other statistical parameters which were calculated by HypoRefine module during hypothesis generation. A good pharmacophore model should have a high correlation coefficient, lowest total cost and RMSD values, and the total cost should be close to the fixed cost and away from the null cost. The best pharmacophore model was further validated by external test set and Fischer’s randomization test (Abu Hammad and Taha, 2009).

Molecular Docking

This study used LBD (ligand binding domain) of peroxisome proliferator-activated receptor gamma or PPARγ(PDB id : 3HOD) (Fracciolla et al., 2009) structure with 2.1 Å resolution. The crystal structure were selected according to best resolution or lower resolution value, and also have R-free and R-value lower than 0.25 (Muchtaridi et al., 2014). The 3D structures of lignan derivatives compounds were constructed using Hyperchem 7, then were optimized using Austin Model 1 (AM1). Docking results have been saved in a dlg file to be analyzed to get
information about ligand orientation, binding energy value and $K_i$ value (Ikram et al., 2015). Molecular docking method was validated by re-docking of the ligand of crystal as control docking against PPARγ (PDB id: 3HOD). Autodock 4.0 was employed to know the interaction between ligand and receptor (Muchtaridi et al., 2014).

RESULTS

3D-QSAR Pharmacophore Model

The Hypogen that embedded in DS 2.5. was generated 100 models. The results of the pharmacophore hypothesis from the 3D QSAR Pharmacophore were presented in Table 1. The results showed that hypothesis 1, run 9 (Hypo1) was the best hypothesis results in this study. This is based on the value of cost difference, which is the highest value of 40.9316, lowest root mean square error with value of 0.591049, and highest correlation coefficient with value of 0.972949.

Table 1: Hypothesis result of 3D QSAR pharmacophore for training set.

<table>
<thead>
<tr>
<th>Run no.</th>
<th>Total cost</th>
<th>Cost diff.</th>
<th>RMSD (Å)</th>
<th>Correlation (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>89.4434</td>
<td>40.7476</td>
<td>0.628438</td>
<td>0.96755</td>
</tr>
<tr>
<td>5</td>
<td>89.6173</td>
<td>40.5737</td>
<td>0.683979</td>
<td>0.962129</td>
</tr>
<tr>
<td>6</td>
<td>93.0029</td>
<td>37.1881</td>
<td>0.788700</td>
<td>0.949402</td>
</tr>
<tr>
<td>7</td>
<td>91.2994</td>
<td>38.8916</td>
<td>0.799089</td>
<td>0.947947</td>
</tr>
<tr>
<td>9</td>
<td>89.2594</td>
<td>40.9316</td>
<td>0.591049</td>
<td>0.972949</td>
</tr>
</tbody>
</table>

(Null cost-total cost), null cost = 130.191. All cost values are in bits.

This proved that there was more than 97% correlation between estimated activity and actual activity from the training set compounds. Fixed cost, total cost, and null cost are 84.8473, 89.2594 and 130.191, respectively. Ideal difference between null cost and fixed cost are the value in between 70-100 bit. The difference between null cost and fixed cost that lies between 40-60 bits indicates that correlation probability from Hypo1 is 75%-90% (Accelrys, 2002). The bigger difference and the higher probability could identify a good pharmacophore model. Configuration cost value has to be less than 17 and configuration value obtained is 16.4502 (Sutter et al., 2000). The 3D space and distance constraints of these pharmacophore features are shown in Fig. 2B. Pharmacophore features of Hypo1 was presented in Figure 1 (A), consisting of 5 pharmacophore features, employing HBA, NI, RA, two HY and one excluded volume in Figure 1(B), shows the distance between features in Amstrong (Å) units. The distance between features of HY1 and HY2 was 6.6885 Å. HY1 to HBA was 4.220 Å, HBA to NI was 12.096 Å, HBA to RA, 7.401 Å and RA to NI was 5.880 Å. This distance happens to be a combination of electrical and steric features of compounds that is necessary to ensure an optimal interaction of molecules with PPARγ to trigger its biological response (agonist) (Manetti et al., 2003). Figure 1c showed a lot of hydrophobic interactions in the binding pocket PPARγ (hydrophobic character). HBA interacted with the amino acid residue HIS449 (2.4 Å), HIS323 (3.0 Å) and TYR473 (2.8 Å). This interaction explained that the ligands should have hydrophobic interactions and hydrogen bonding, to ensure their biological activity (PPARγ agonist). Mapping results of the training set compound on Hypo1 with the activity (experimental and estimated) and their corresponding error.

Fig. 1: Feature of Best Pharmacophore with Validation by Hyporefine Run in DS 2.5. (a) The best HypoRefine pharmacophore model, Hypo1. (b) 3D spatial relationship and geometric parameters of Hypo1. (c) Mapping of Hypo1 to MA 6 from training set (d) Mapping of Hypo1 to LA 105 from training set. Pharmacophore features are color coded; green: hydrogen-bond acceptor (HBA), cyan – hydrophobic feature (Hy), dark blue – negative ionizable (N), orange – aromatic ring (AR), and grey – excluded volume.
All most active (MA) compounds have been correctly predicted by Hypo1. All MA compounds mapped all pharmacophore features covering HBA, NI, AR and HY. One of MA compounds, i.e. compound 6 with EC50 value of 0.10 nM shows a fit value of 10.85 (Experiment: 0.1nM; Prediction: 0.18nM). Fit value indicates how well the pharmacophore features overlay chemical features in the molecule (Guner, 2005). A higher fit value is better [32]. Figure (3A), showed the compound 6 with HBA feature fitted with nitrogen sp/sp2 hybridization that has free electrons and smaller or no charge. The NI feature fitted against the carboxylate functional group, AR mapped to the benzene and both HY with methyl and phenyl respectively.

**Validation of Pharmacophore Model**

**Fischer Randomization test**

Fischer Randomization is embedded in the DS 2.5 package to evaluate the best pharmacophore model. CatScramble module within Catalyst was performed to evaluate the models.

![Graph showing correlation and total cost of hypotheses](image)

**Fig. 2**: The difference in total cost (b) and correlation (a) of hypotheses between the initial spreadsheet and 19 random spreadsheets after CatScramble run.

This statistical program in CatScramble mixes up activity values of all training set compounds to check whether there are strong correlation between the structure and activity.

Cross validation test explained that Hypo1 has a significant value of 95% and the results. It could be seen in Figure 2a that the correlation value (r) on all pharmacophore models (Random 1 – Random 19) were smaller than the Hypo1 correlation, which showed highest correlation between experimental activity and predictive value. Fischer randomization test was done to assess the predictive quality of Hypo1. COST value indicated that the difference between estimation values and experiment value. COST data obtained from running of 3D QSAR from every hypothesis was created a mold to test for Fischer randomization. Compounds were taken at random from every hypothesis, and then compared with the other hypothesis. Based on the results of Fischer randomization (Figure 2b), it was shown that data cost from training set compound was taken randomly by Hypo1 that showed the lowest COST. Both COST value and correlation showed that Hypo1 was the best hypothesis.

**Test set methods**

Subsequently, all the test set molecules were prepared by the same way as that for the training set molecules. Hypo1 was applied to map the 25 test set compounds which gave a correlation coefficient of between experimental and estimated activities as shown in estimated Fig 4.

The results showed that all ligands with MA activity were predicted accurately, except for ligand 6 and 10, where by the ligand is underestimated became M.

From all the M ligands, only ligand 45 was overestimated to have MA activity. LA ligands were all predicted accurately as well, except for ligand 63, in which it is overestimated into MA. Hypo1 shows a correlation as large as 0.9011, which means that there are good relationship between experimental value and estimated value. A linear regression showing the similarity of experimental activity and estimated activity is shown in Figure 3.

![Graph showing correlation and estimated activity](image)

**Fig. 3**: Plot of the correlation (r) between the experimental activity and the predicted activity by Hypo1 for the test set molecules and the training set molecules.

Figure 4 showed that the character of the PPARγ binding pocket using software Ligand Explorer incorporated in the Protein Data Bank (PDB) website. Binding mdeof the lead compound was shown by the interaction between the amino acid residues with the binding site. Figure 1c shows a lot of hydrophobic interactions in the binding pocket PPARγ (hydrophobic character). Figure 1c and Figure 3 showed that HBA interacted with the amino acid residue...
HIS 449 (2.4 Å), HIS 323 (3.0 Å) and TYR 473 (2.8 Å). This interaction suggested that the ligands should have hydrophobic interactions and hydrogen bonding, to ensure their biological activity (PPARγ agonist).

![Image](65x313 to 279x478)

**Fig. 4:** Character of the PPARγ binding pocket using software Ligand Explorer incorporated in the Protein Data Bank (PDB). (a) Hydrophobic interaction in PPARγ binding pocket, (b) Hydrogen bond interaction in PPARγ binding pocket

**Activity Prediction of Nutmeg Seed Compounds**

Dihydro-di-isoeugenol (DHDE) in nutmeg (*Myristica fragrans* Hout) is recently developed as an antiabetic medicine with predicted mechanism of action as a dual agonist, i.e. PPAR α and γ (Muchtaridi et al., 2014). In this study, 14 compounds derived from nutmeg plant, thiazolidinediones and lead compound 2-arloxy-3-phenyl propanoic acid (3HOD) which is expected active, was mapped by Hypo1 model.

Reason for mapping is to approach that even with different chemotypes, it can have good mapping results with Hypo1. From a total of 22 compounds, Malabaricone B from nutmeg extract shows highest fit value (8.10119) with estimated activity value of 100,997 nM. However, this compound still doesn’t have any research on its antidiabetic activity towards PPARγ receptor *in vitro* or *in vivo*. Antidiabetic currently being used in Indonesia now is Pioglitazone, with a fit value of 6.87839. Hypo1 predicted EC50 1686.99nM where experimental activity value of 1280nM (Casimiro-Garcia et al., 2013). Rivoglitazone shows fit value of 7.08281 (Kong et al., 2011; Schimke and Davis, 2007) with estimated activity value of 1053.62 nM.

Although Malabaricone B shows highest fit value, it exists in a small quantity in nutmeg seeds which complicates the process being a marker and for extraction. Perhaps in future study, a new lead compound can be found or a better extraction technique for the nutmeg contents.

**CONCLUSIONS**

Active pharmacophoric features for PPAR gamma agonist are Hydrogen Bond Acceptor (HBA) and Hydrophobic (HY). The compouds sof nutmeg seeds (*Myristica fragrans* Hout) have the highest fit value (8.10119) with a total estimated of activity value 100,997 nM.

**Financial support and sponsorship:** Nil.

**Conflicts of interest:** There are no conflicts of interest.

**REFERENCES**


Feldman, PL, Lambert, MH, and Henke, BR. PPAR modulators and PPAR pan agonists for metabolic diseases: the next generation of


Newman, DJ, and Cragg, GM. Natural Products as Sources of New Drugs over the Last 25 Years.L. Journal of Natural Products, 2007; 70 (3): 461-477.


Wang, HY, Cao, ZX, Li, LL, Jiang, PD, Zhao, YL, Luo, SD, Yang, L, Wei, YQ, and Yang, SY. Pharmacophore modeling and virtual screening for designing potential PLK1 inhibitors. Bioorg Med Chem Lett, 2008; 18 (18): 4972-4977.