

## COMPUTER-AIDED DESIGN OF CHALCONE DERIVATIVES AS LEAD COMPOUNDS TARGETING ACETYLCHOLINESTERASE

Florentinus Dika Octa Riswanto, Maywan Hariono, Sri Hartati Yuliani, Enade Perdana Istyastono\*

Faculty of Pharmacy, Sanata Dharma University, Campus III Paingan, Maguwoharjo, Depok, Sleman, Yogyakarta 55282

Submitted: 27-12-2016

Revised: 25-01-2017

Accepted: 25-02-2017

\*Corresponding author  
Enade Perdana Istyastono

Email:  
enade@usd.ac.id

### ABSTRACT

One of well-established biological activities for chalcone derivatives is as acetylcholinesterase inhibitors, which can be developed for the therapy of Alzheimer's disease. Assisted by retrospectively validated structure-based virtual screening (SBVS) protocol to identify potent acetylcholinesterase inhibitors, 80 chalcone derivatives were designed and virtually screened. The *F*-measure value as the parameter of the predictive ability of the SBVS protocol developed in the research presented in this article was 0.413, which was considerably better than the original SBVS protocol (*F*-measure=0.226). Among the screened chalcone derivatives two were selected as potential lead compounds to design potent inhibitors for acetylcholinesterase: 3-[4-(benzyloxy)-3-methoxyphenyl]-1-(4-hydroxy-3-methoxy-phenyl)prop-2-en-1-one (**3k**) and 3-[4-(benzyloxy)-3-methoxyphenyl]-1-(4-hydroxyphenyl)prop-2-en-1-one (**4k**).

**Key words:** Computer-aided drug design, virtual screening, chalcone derivatives, acetylcholinesterase, Alzheimer's disease

### INTRODUCTION

Alzheimer's disease (AD), a progressive brain disorder, is a neurodegenerative disease which becomes symptomatic after brain changing happened over the years (Aggarwal *et al.*, 2015). The prevalence of dementia, as AD symptom were varied (Rizzi *et al.*, 2014) and the number of people living with AD have been predicted to be increased two times every two decades from 46.8 million by 2015 to 74.7 million people by 2030 (Prince *et al.*, 2015). The deficiency of the brain neurotransmitter acetylcholine (ACh) is often associated with pathogenesis of AD (Tabet, 2006). Acetylcholine plays important roles in the nervous system such as increasing neurotransmitter release, supporting synaptic transmission, inducing synaptic plasticity, and coordinating firing of groups of neurons (Picciotto *et al.*, 2012; Tsuda, 2012). The hydrolysis of ACh into choline and acetic acid, a reaction needed to allow the returning of cholinergic neuron to the resting state, was catalyzed by a family of enzymes called cholinesterase (Colovic *et al.*, 2013). Acetylcholinesterase (AChE), one of cholinesterase types found in many types of

conducting tissue, is a highly possible therapeutic target of Alzheimer disease (Mehta *et al.*, 2012).

Recently, research on AChE inhibitor has been rapidly developed due to the availability of supporting facilities for designing AChE inhibitor compounds in the treatment of AD (Singh *et al.*, 2013). Promising compounds to be developed as AChE inhibitor were chalcone derivatives (Sukumaran *et al.*, 2016; Tran *et al.*, 2016). Chalcones or 1,3-diphenyl-2-propene-1-one can be obtained both from the plants (Abdelwahab, 2013; Adewusi *et al.*, 2010) and from the synthetic way due to condensation reaction between substituted aromatic aldehyde with substituted acetophenones in alkaline condition (Jayapal and Sreedhar, 2010). Nevertheless, it is important to consider the structure-activity relationship and computational approach in designing more active AChE inhibitors (Andersson *et al.*, 2013; Ece *et al.*, 2015).

Some strategies to optimize predictive abilities of Structure-based Virtual Screening (SBVS) protocols have been developed (de Graaf *et al.*, 2011; Istyastono, 2015; Istyastono *et al.*, 2015a; Istyastono and Setyaningsih, 2015;

Radifar *et al.*, 2013<sup>b</sup>; Sirci *et al.*, 2012). Some of the strategies have been employed successfully in the prospective screenings to discover novel potent fragments for histamine H<sub>1</sub> (de Graaf *et al.*, 2011), H<sub>3</sub> (Sirci *et al.*, 2012) and H<sub>4</sub> receptors (Istyastono *et al.*, 2015<sup>a</sup>), and in the repurposing selective cyclooxygenase-2 inhibitor celecoxib as a potent ligand for estrogen  $\alpha$  receptor (Istyastono *et al.*, 2015<sup>b</sup>; Radifar *et al.*, 2013<sup>b</sup>). In 2011, by combining ChemPLP score (Korb *et al.*, 2009) and the post-docking scoring functions of the interaction fingerprint (IFP) program (Marcou and Rognan, 2007) as the cutoff values, de Graaf *et al.* could reach *F*-measure value of about 0.515 and discover 19 novel fragments with K<sub>i</sub> ranging from 10  $\mu$ M to 6 nM for histamine H<sub>1</sub> receptor after virtually screened about 13 million compounds and tested only 26 hits resulted from the prospective virtual screening (de Graaf *et al.*, 2011). In 2012, employing Fingerprint for Ligand and Protein (FLAP) software as the post-docking modeling strategy to identify active histamine H<sub>3</sub> fragments Sirci *et al.* could discover 18 novel fragments out of 29 tested hits resulted from prospective virtual screening of 156,090 compounds (Sirci *et al.*, 2012). In 2015, by employing several homology models with different reference ligands, using post-docking scoring functions of the IFP program (Marcou and Rognan, 2007), Istyastono *et al.* could optimize the *F*-measure value from 0.018 to 0.553 and discover 9 novel histamine H<sub>4</sub> fragments after virtually screen 43,326 fragments and tested 37 hits resulted from the prospective virtual screening (Istyastono *et al.*, 2015<sup>a</sup>). Fortunately the structure of histamine receptors was well-studied, which established ASP<sup>3.32</sup> as the anchor of the receptor-ligand binding (Istyastono *et al.*, 2011<sup>a</sup>; Jongejan *et al.*, 2008). These success stories employed the anchor interaction to filter the docking poses before proceeding to the subsequent step in the SBVS protocols (de Graaf *et al.*, 2011; Istyastono *et al.*, 2015<sup>a</sup>; Sirci *et al.*, 2012). However, the information of the anchor interaction is still rare for other targets in drug discovery. Moreover, instead of having anchor interaction, some targets accept more poses for their potent ligands (Istyastono, 2015). Several attempts to optimize SBVS protocols by taking

into account several poses of active ligands (Istyastono *et al.*, 2011<sup>b</sup>; Wang *et al.*, 2015) by decision trees construction have been performed targeting estrogen  $\alpha$  receptor and could optimize the *F*-measure value from 0.215 to 0.642 (Istyastono, 2015; Setiawati *et al.*, 2014). Notably, the same strategy could not be applied in this research since no decision tree could be constructed. Another strategy to optimize the predictive ability of the SBVS protocol in this research should be developed and tested.

The research presented in this paper was aimed to design chalcone derivatives as potent AChE inhibitors by utilizing retrospectively validated SBVS protocols constructed by employing PLANTS1.2 as the molecular docking software (Korb *et al.*, 2009, 2007) and PyPLIF as the Protein-Ligand Interaction Fingerprints (PLIF) identification software for re-scoring (Radifar *et al.*, 2013<sup>a</sup>; Radifar *et al.*, 2013<sup>b</sup>). New post PLIF identification descriptor or employing ensemble PLIF (abbreviated to *ensplif*) was introduced and employed here to construct and optimize SBVS protocols (Istyastono *et al.*, 2017), which was then assessed (Cannon *et al.*, 2007; de Graaf *et al.*, 2011; Desaphy *et al.*, 2013; Powers, 2011) and compared to the original SBVS accompanying the release of DUD-E (Mysinger *et al.*, 2012). The best SBVS protocol constructed in this research could outperform the predictive ability of the SBVS protocol of DUD-E. The protocol was subsequently employed to assist the selection of chalcone derivatives as lead compounds in the development of acetylcholinesterase inhibitors

## MATERIALS AND METHODS

The crystal structure of AChE obtained from the protein data bank (PDB) with PDB id of 1E66 was used as the reference structure (Dvir *et al.*, 2002; Mysinger *et al.*, 2012). Active AChE inhibitors (453) and the decoys (26,350) from DUD-E (Mysinger *et al.*, 2012) were employed to perform retrospective validations for the SBVS protocol. Chalcone derivatives inspired from Imran *et al.* (Imran *et al.*, 2015) which were designed manually (Table I) were used as entry points for lead compounds selection. All calculations and computational

simulations were performed on a Linux (Ubuntu 12.04 LTS Precise Pangolin) machine with Intel<sup>®</sup> Xeon<sup>®</sup> CPU E31220 (@ 3.10GHz) as the processors and 8.00GB of RAM. Computational medicinal chemistry applications employed in this research were SPORES (ten Brink and Exner, 2009), PLANTS1.2 (Korb *et al.*, 2009; Korb *et al.*, 2007), Open Babel 2.2.3 (O'Boyle *et al.*, 2011), PyPLIF0.1.1 (Radifar *et al.*, 2013<sup>a</sup>; Radifar *et al.*, 2013<sup>b</sup>), MarvinSketch 14.11.10.0 (ChemAxon, 2014) and PyMOL 1.2r1 (Lill and Danielson, 2011). Statistical analysis was performed by using R 3.3.0 (R Core Team, 2016).

### Computational methods

#### Virtual molecular target preparation

Previously published method to prepare cyclooxygenase-2 as virtual target (Istyastono, 2016) was adopted to prepare virtual target AChE. The crystal structure of AChE was downloaded from <http://www.rcsb.org/pdb/explore.do?structureId=1e66>. Then, the module *splitpdb* in SPORES was subsequently used to split the pdb file and to convert the splitted files into *mol2* files the virtual AChE (*protein.mol2*), the co-crystal ligand3-chloro-9-ethyl-6,7,8,9,10,11-hexahydro-7,11-methanocycloocta[B]quinolin-12amine (*ligand\_HUX803\_0.mol2*), and the water molecules. The *mol2* files were then ready to be employed in molecular docking simulation employing PLANTS1.2 docking software.

#### Ligands preparation for retrospective virtual screening

Similar to virtual target preparation, previously published methods for ligands preparation (Istyastono, 2016; Istyastono and Yuniarti, 2016) were adopted and modified. Known AChE active inhibitors and their decoys were downloaded in their SMILES format from DUD-E (Mysinger *et al.*, 2012). They were stored locally as *actives\_final.ism*, and *decoys\_final.ism*. Each compound in the files 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 SPORES was subsequently employed to properly check and assign the

protonation state of the *mol2* file into a proper *mol2* file ready to be docked by using PLANTS1.2 docking software.

#### Automated molecular docking and virtual screening

Similar to previously published procedures (Istyastono *et al.*, 2015<sup>b</sup>; Istyastono and Setyaningsih, 2015; Setiawati *et al.*, 2014), all virtual screenings were performed by docking program PLANTS1.2. For each compound, 50 poses were produced and clustered at root mean square deviation (RMSD) of 1.0 Å. The ChemPLP scoring function was used at speed setting 2. The binding pocket was defined by the coordinates of the center of the co-crystal *ligand\_HUX803\_0.mol2* and a radius of 5 Å. All other options of PLANTS1.2 were left at their default setting. Every compound was virtually screened five times independently. Seven different interaction types (negatively charged, positively charged, hydrogen bond (H-bond) acceptor, H-bond donor, aromatic face-to-edge, aromatic face-to-face, and hydrophobic interactions) were subsequently identified by employing PyPLIF for each docking pose (Radifar *et al.*, 2013<sup>a</sup>; Radifar *et al.*, 2013<sup>b</sup>).

#### Predictive ability of the SBVS protocols

The predictive ability of the SBVS protocols was determined by their *F*-measure value of the retrospective validation (Cannon *et al.*, 2007; Desaphy *et al.*, 2013). The SBVS protocols predict screened compound as active or inactive. Since the biology activity as AChE inhibitor of the compounds in the retrospective validation attempts using DUD-E has already known (Mysinger *et al.*, 2012), the confusion matrix could be constructed (Cannon *et al.*, 2007): (i) Active compounds which were predicted as active were assigned as true positives (TP); (ii) Active compounds which were predicted as inactive were assigned as false negatives (FN); (iii) Decoys which were predicted as active were assigned as false positives (FP); and (iv) Decoys which were predicted as inactive were assigned as true negatives (TN). The *F*-measure value (Cannon *et al.*, 2007; Desaphy *et al.*, 2013) was calculated as follows:  $(2 \times TP) / ((2 \times TP) + FN + FP)$ .

### Predictive ability optimization of the SBVS protocols

For comparison as the standard, the compounds in the retrospective SBVS were ranked based on the best ChemPLP score and the ChemPLP score of the 1% identified FP was used as the cutoff to predict as active or inactive (de Graaf *et al.*, 2011; Istyastono *et al.*, 2015<sup>a</sup>; Korb *et al.*, 2009; Mysinger *et al.*, 2012). The optimization of the SBVS protocols was performed by optimizing the post SBVS decision trees constructed in RPART package (Istyastono, 2015; R Core Team, 2016; Therneau *et al.*, 2015) by using *ensplif* as the descriptors. Besides considering all poses resulted in the SBVS, *ensplif* as descriptors were calculated by selecting poses by using ChemPLP scores systematically from -125 to 0. The algorithm of the *ensplif* calculation was as follows: (i) Poses selection based on the ChemPLP score as the cutoff; (ii) In every PLIF bitstring, all “on” interactions were counted; (iii) The numbers resulted from point (ii) were divided by 250 since the molecular docking resulted in 50 docking poses in each run and every compound was screened 5 times (see subsection Automated molecular docking and virtual screening).

### Computer-aided identification of chalcone derivatives as lead compounds to discover potent AChE inhibitors

Eighty chalcone derivatives were designed (Table I). Every structure in table I was built and converted to SMILES format by using module *Copy as smiles* in Marvin Sketch 14.11.10.0 (ChemAxon, 2014). The structure was then undergone ligand preparation process as presented in the subsection Ligands preparation for retrospective virtual screening. The best SBVS protocol to identify potent AChE inhibitors resulted from the subsection Predictive ability optimization of the SBVS protocols was used to virtually screen the designed compound. Based on the results of the virtual screening, some lead compounds were selected. The synthesis feasibility of the selected lead compounds was subsequently analyzed.

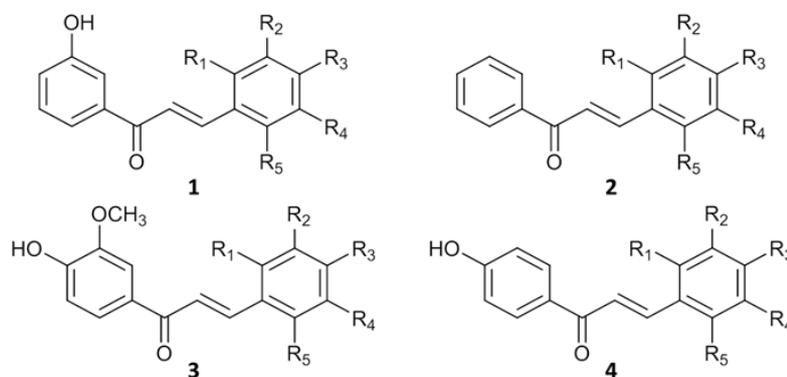
## RESULTS AND DISCUSSION

Aimed to initiate the discovery and development of potent AChE inhibitors, this research constructed and optimized the predictive ability of SBVS protocol to identify potent AChE inhibitors and employed the best protocol to select best potential inhibitors from 80 chalcone derivatives presented in Table I as lead compounds. The construction and optimization of the SBVS protocol to identify potent AChE inhibitors employed potent AChE inhibitors with IC<sub>50</sub> value of less than 1 μM and their decoys organized and stored in DUD-E (Mysinger *et al.*, 2012) as the retrospective compounds for the validation. Chalcone derivatives were selected since the synthesis expertise in the field is accessible for further investigation in the subsequent discovery and development process (Imran *et al.*, 2015).

### Construction and Predictive Ability Optimization of the SBVS Protocol

During molecular docking and PLIF identification phases in the retrospective SBVS campaign, nineteen ligands and 1,735 decoys could not pass the protocol, which were then predicted as inactive compounds. Therefore, the protocol has resulted in 19 FN and 1735 TN during the molecular docking and the PLIF identification phases. In every remaining result, the docking pose with the best ChemPLP score was collected and ranked based on the ChemPLP score. In this ranked database, at 1% FP (263 compounds) there were 52 TP, 401 FN, and 25,987 TN. Therefore the *F*-measure value of this ranked database was 0.135, which was corresponded to the enrichment factor value of 11.46. This predictive ability of the SBVS protocol by using 1% FP as the cutoff to predict the activity was much worse than the original protocol accompanying the release of DUD-E, which showed *F*-measure and enrichment factor values of 0.226 and 20.1, respectively (Mysinger *et al.*, 2012). Moreover, with this predictive ability, the SBVS protocol is not suggested to be used for prospective virtual screening campaigns (Chen, 2015; de Graaf *et al.*, 2011; Istyastono *et al.*, 2015<sup>a</sup>; Sirici *et al.*, 2012). Predictive ability optimization of this SBVS protocol was therefore required.

Table I. Designed and virtually screened chalcone derivatives in this research.



| Code | R <sub>1</sub>    | R <sub>2</sub>    | R <sub>3</sub>                                      | R <sub>4</sub>    | R <sub>5</sub>    |
|------|-------------------|-------------------|---|-------------------|-------------------|
| a    | -H                | -OH               | -H  | -OH               | -H                |
| b    | -H                | -H                | -F  | -H                | -H                |
| c    | -H                | -NO <sub>2</sub>  | -H  | -H                | -H                |
| d    | -H                | -OCH <sub>3</sub> | -OH   | -H                | -H                |
| e    | -H                | -H                | -OH   | -H                | -H                |
| f    | -H                | -H                | -OCH <sub>3</sub>                                   | -H                | -H                |
| g    | -H                | -OCH <sub>3</sub> | -OCH <sub>3</sub>                                   | -OH               | -H                |
| h    | -OCH <sub>3</sub> | -OCH <sub>3</sub> | -OCH <sub>3</sub>                                   | -H                | -H                |
| i    | -OCH <sub>3</sub> | -H                | -OCH <sub>3</sub>                                   | -OCH <sub>3</sub> | -H                |
| j    | -H                | -H                | -O-(CH <sub>2</sub> ) <sub>3</sub> -CH <sub>3</sub> | -H                | -H                |
| k    | -H                | -OCH <sub>3</sub> | -O-CH <sub>2</sub> -C <sub>6</sub> H <sub>5</sub>   | -H                | -H                |
| l    | -NO <sub>2</sub>  | -OH               | -H  | -H                | -H                |
| m    | -OCH <sub>3</sub> | -H                | -OCH <sub>3</sub>                                   | -H                | -H                |
| n    | -OCH <sub>3</sub> | -H                | -OCH <sub>3</sub>                                   | -H                | -OCH <sub>3</sub> |
| o    | -OCH <sub>3</sub> | -CH <sub>3</sub>  | -OCH <sub>3</sub>                                   | -H                | -H                |
| p    | -OCH <sub>3</sub> | -H                | -OCH <sub>3</sub>                                   | -H                | -CH <sub>3</sub>  |
| q    | -OCH <sub>3</sub> | -H                | -OCH <sub>3</sub>                                   | -OCH <sub>3</sub> | -H                |
| r    | -H                | -OCH <sub>3</sub> | -OCH <sub>3</sub>                                   | -OCH <sub>3</sub> | -H                |
| s    | -H                | -OH               | -OH   | -H                | -H                |
| t    | -OH               | -OH               | -H  | -H                | -H                |

Equipped with the facts that some ligands could interact with their protein targets in more than one pose (Istyastono *et al.*, 2011<sup>b</sup>; Wang *et al.*, 2015) and inspired by the lock-and-key theory (Koshland, 1994; Stoddard and Koshland, 1992), we introduce a novel post PLIF identification technique to optimize the predictive ability of the SBVS protocol by using *ensplif*. This strategy was successfully utilized to slightly increase the SBVS protocol to identify potent ligands for estrogen  $\alpha$  receptor (Istyastono, 2015; Istyastono *et al.*, 2017). Therefore, we were tempted to apply the same technique in this research since employing the

technique used by Istyastono (2015) could not result in any constructed decision tree in this research. Taking into account all docking poses in every retrospective screened compound in this research resulted in a post SBVS decision tree using *ensplif* as the descriptors with *F*-measure value of 0.371, which was better than the ChemPLP-based ranked database at 1% FP and to the original protocol accompanying the release of DUD-E (*F*-measure = 0.226; Mysinger *et al.*, 2012). Notably, with this *F*-measure value the SBVS protocol has already accepted for further prospective virtual screening. However, the protocol could still be

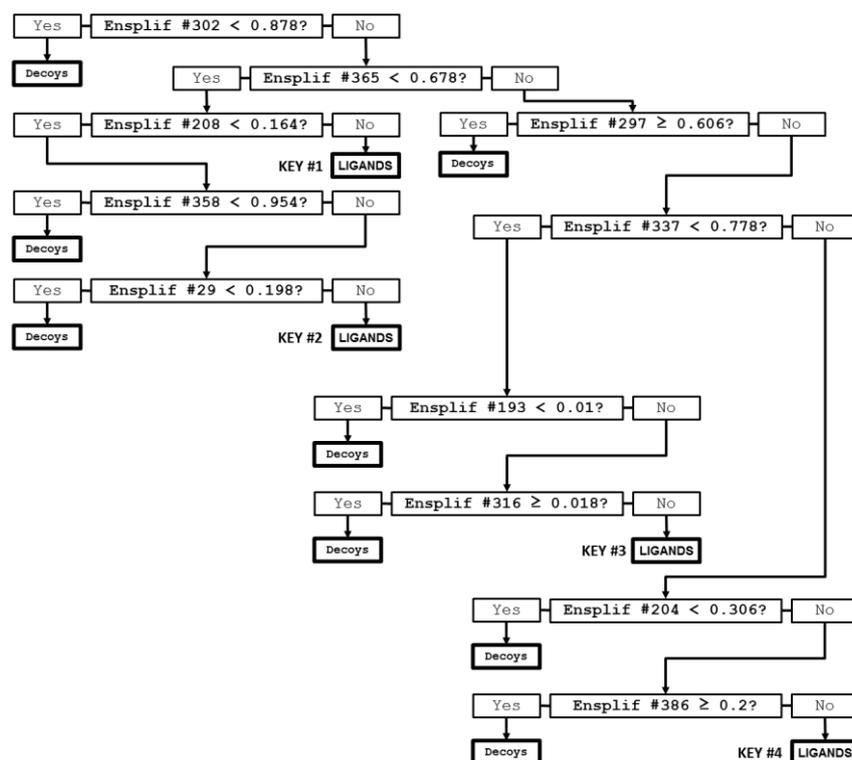


Figure 1. The best decision tree employing ensplif as descriptors to identify potent ligands for AChE. Four types of how ligands bind to AChE or “key” are identified.

optimized by selecting the more plausible docking poses with a certain ChemPLP score as the cutoff (de Graaf *et al.*, 2011).

The optimization by systematically assessing the predictive ability of employing ChemPLP scores from -125 to 0 as the cutoff for pose selection resulted in *F*-measure values ranging from 0.103 to 0.413. The best SBVS protocol with *F*-measure value of 0.413 was comprised of molecular docking simulation using PLANTS1.2 (Korb *et al.*, 2009), PLIF identification of the docking poses using PyPLIF (Radifar *et al.*, 2013<sup>a</sup>; Radifar *et al.*, 2013<sup>b</sup>), and the decision tree constructed using RPART (Therneau *et al.*, 2015) with ensplif calculated from PLIF of docking poses with ChemPLP score  $\leq -33$  (Figure 1).

The RPART method to construct decision trees used in the research is one of supervised machine learning methods which are susceptible for overfitting (Cannon *et al.*, 2007; Gabel *et al.*, 2014) and chance correlation (Lim *et*

*al.*, 2009; Tarcsay *et al.*, 2013; Zambre *et al.*, 2007). A model is overfitting if the ratio of the cross-validation error rate over the training error rate is more than 1.5 (Cappel *et al.*, 2015; Istyastono, 2016). Since the ratio of the cross-validation error rate over the training error rate of the best decision tree (Figure 1) was 1.158, this model was not overfitting. In order to ensure if the best decision tree was not a chance correlation, 1000 times *y*-randomization were performed (Lim *et al.*, 2009; Smits *et al.*, 2010) and there was no *y*-randomized model with *F*-measure higher than the *F*-measure of the best decision tree (Figure 1). Therefore, there was no evidence of overfitting and chance correlation of the SBVS protocol comprising the best decision tree constructed using RPART with ensplif calculated from PLIF of docking poses with ChemPLP score  $\leq -33$  as the descriptors.

There are 11 selected ensplif in the decision tree (Table II). Notably, hydrophobic

Table II. The ensplif as selected descriptors in the decision tree (Figure 1) and their corresponding interactions in the AChE binding pocket

| Ensplif No. | Corresponding Interactions in the AChE Binding Pocket |   |
|-------------|---|---|
|             | Amino Acid Residue                                    | Interaction Type <sup>a)</sup>          |
| 29          | GLN <sup>74</sup>                                     | hydrophobic interaction                 |
| 193         | TYR <sup>130</sup>                                    | hydrogen bond (protein as the donor)    |
| 204         | SER <sup>200</sup>                                    | hydrophobic interaction                 |
| 208         | SER <sup>200</sup>                                    | hydrogen bond (protein as the acceptor) |
| 297         | PHE <sup>330</sup>                                    | aromatic edge-to-face interaction       |
| 302         | PHE <sup>331</sup>                                    | hydrophobic interaction                 |
| 316         | LEU <sup>333</sup>                                    | hydrophobic interaction                 |
| 337         | TRP <sup>432</sup>                                    | hydrophobic interaction                 |
| 358         | HIS <sup>440</sup>                                    | hydrophobic interaction                 |
| 365         | GLY <sup>441</sup>                                    | hydrophobic interaction                 |
| 386         | ILE <sup>444</sup>                                    | hydrophobic interaction                 |

<sup>a)</sup>Ref: (Radifar *et al.*, 2013b)

interaction to PHE<sup>331</sup> is essential for AChE ligand since the ligand should have ensplif #302  $\geq 0.878$  to be predicted active as potent ligand for AChE, otherwise it will be predicted as inactive (Figure 1). Hydrophobic interactions dominate the selected interactions. This interaction type is however responsible only for the affinity but not for the selectivity because of its lack of directionality (Bissantz *et al.*, 2010; Desaphy *et al.*, 2013; Marcou and Rognan, 2007). Only three out of these 11 interactions have directionality (Bissantz *et al.*, 2010; Desaphy *et al.*, 2013; Marcou and Rognan, 2007; Radifar *et al.*, 2013b): (i) Ensplif#193 (hydrogen bond to TYR<sup>130</sup> with tyrosine as the donor); (ii) Ensplif#208 (hydrogen bond to SER<sup>200</sup> with serine as the acceptor); and (iii) Ensplif#297 (aromatic edge-to-face interaction to PHE<sup>330</sup>). These three interactions are responsible not only for the ligand affinity but also for the ligand selectivity (Bissantz *et al.*, 2010; Istyastono *et al.*, 2011a, 2011b). Hydrogen bond and aromatic interaction have been identified as the main factors in activity cliffs with at least 100-fold activity increase (Furtmann *et al.*, 2015). Based on the decision tree presented in Figure 1, there are 4 types of “key” that can bind to the “lock” AChE. Notably, all branches leading to active compound in the decision tree (Figure 1) or the “keys” involve either the hydrogen bond (ensplif#208 or #193) or the aromatic interaction (ensplif#297). Oddly, in

contrary with the hydrogen bonds, which are favorable, the aromatic interaction here is unfavorable.

In the retrospective attempts, the optimized protocol resulted in *F*-measure value of 0.413 from 124 TP; 329 FN; 26,226 TN and 24 FP. By employing the retrospective results, the sensitivity (true positive rate) and the specificity (true negative rate) can be calculated (Cannon *et al.*, 2007) and get the following results: 0.274 and 0.999, respectively. This indicates that the optimized protocol tends to correctly predict inactive compounds. A predicted active compound using the optimized protocol is therefore highly probable to be an active AChE inhibitors but it is highly suggested to further investigate a predicted inactive compound especially the one that has ensplif number 302 of more than or equal to 0.878 (Figure 1).

### Virtual Screening on Designed Chalcone Derivatives Employing the Optimized Protocol

Eighty chalcone derivatives presented in Table I were virtually screened employing the optimized SBVS protocol (Figure 1). However, none of these compounds were predicted as active. Therefore, selection of lead compounds instead of identification of potent AChE inhibitors from the chalcone derivatives (Table I) was the subsequent step. As mentioned in the previous subsection,

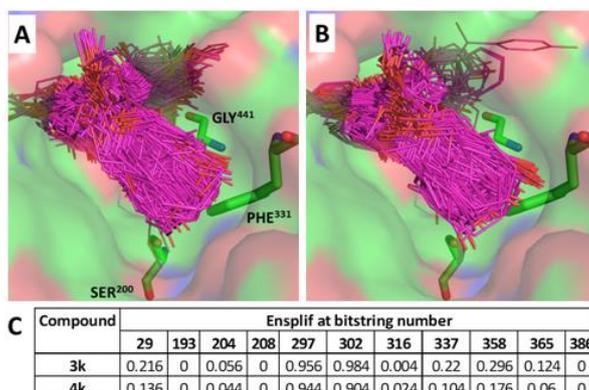


Figure 2. Docking poses of **3k**(A) and **4k**(B) which show hydrophobic interaction to PHE<sup>331</sup> but do not show hydrophobic interaction to GLY<sup>441</sup> and the corresponding important enspif values (C). Carbon atoms of compounds **3k** and **4k** are presented in magenta. Carbon atoms of AChE are presented in green. Oxygen and nitrogen atoms are presented in red and blue, respectively. Hydrogen atoms are not shown for clarity. The ligands are in lines mode, while the enzyme is in surface mode with 0.5 transparencies. Important residues following path of key #1 in Figure 1 are also shown in stick mode.

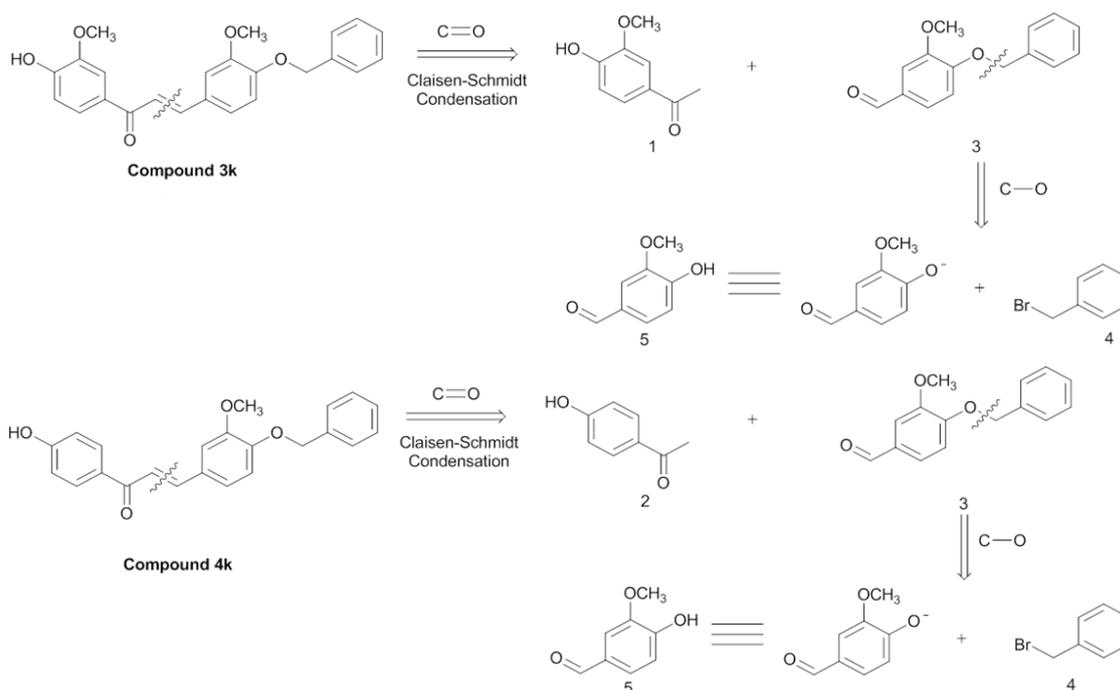


Figure 3. Retrosynthetic analysis of compound **3k** and **4k**. Reagents: (1) acetovanillone; (2) 3'-hydroxyacetophenone; (3) o-benzylvanillin; (4) benzyl bromide; (5) vanillin.

all predicted active compounds must have enspif #302  $\geq$  0.878 (Figure 1). Among the virtually screened chalcone derivatives (Table I), only compounds **3k** and **4k** showed enspif

#302  $\geq$  0.878. The enspif #302 of compounds **3k** and **4k** were 0.956 and 0.944 (Figure 2). Therefore, compounds **3k** and **4k** were selected as lead compounds for further investigation in

this quest to discover potent AChE inhibitors. Since compounds **3k** and **4k** showed  $\text{ensplif} \#365 < 0.678$  (Figure 2), the shortest path to design potent AChE inhibitor by employing these lead compounds is the path of key #1 (Figure 1). Additional hydrogen bond to SER<sup>200</sup> is therefore required (Figures 1 and 2). Based on visual inspection using PyMOL in docking poses of **3k** and **4k** which show hydrophobic interaction to PHE<sup>331</sup> but do not show hydrophobic interaction to GLY<sup>441</sup> (Figure 2), this could be achieved by appending or substituting the R<sub>2</sub> part of the selected lead compounds (Table I) with functional groups which have possibility to be hydrogen bond donors, e.g., guanidine, amide and isothiourea.

Synthesis and further verification using *in vitro* analysis should be done in the near future to provide evidences that compounds **3k** and **4k** could serve as lead compounds. Retrosynthetic analysis or synthon disconnection approaches of compounds **3k** and **4k** (Figure 3). We propose here *o*-benzylvanillin (4-(benzyloxy)-3-methoxybenzaldehyde) to be used as starting material which could be synthesized using vanillin (4-hydroxy-3-methoxybenzaldehyde) and benzylbromide (bromomethylbenzene). Further, *o*-benzylvanillin could be reacted with acetovanillone (1-(4-hydroxy-3-methoxyphenyl) ethanone) for yielding compound **3k** and 3'-hydroxyacetophenone (1-(4-hydroxyphenyl) ethanone) for yielding compound **4k**.

## CONCLUSIONS

Compounds **3k** and **4k** have been identified as potent lead compounds in the drug discovery and development for AD therapy targeting AChE. The identification of these lead compounds employed a retrospectively validated SBVS protocol with a decision tree employing  $\text{ensplif}$  resulted from molecular docking PLANTS1.2 and PLIF identification PyPLIF. In the retrospective validation, the SBVS protocol showed *F*-measure value of 0.413, which outperformed the original SBVS accompanying the database used for the retrospective validation.

## ACKNOWLEDGEMENT

This research was financially supported by the Internship Grant for Lecturer 2016,

Faculty of Pharmacy Sanata Dharma University (No. Far/079/X/2016/ST/D).

## REFERENCES

- Abdelwahab, SI., 2013, *In vitro* inhibitory effect of Boeserngin A on human acetylcholinesterase: Understanding its potential using in silico ADMET studies. *Journal of Applied Pharmaceutical Science*, 3, 30-35.
- Adewusi, EA., Moodley, N., Steenkamp, V., 2010, Medicinal plants with cholinesterase inhibitory activity: A Review. *Afr. J. Biotechnol.*, 9, 8257-8276.
- Aggarwal, NT., Shah, RC., Bennett, DA., 2015, Alzheimer's disease: Unique markers for diagnosis and new treatment modalities. *Indian J. Med. Res.*, 142, 369-382.
- Andersson, CD., Forsgren, N., Akfur, C., Allgardsson, A., Berg, L., Engdahl, C., Qian, W., Ekstrom, F., Linusson, A., 2013, Divergent structure-activity relationships of structurally similar acetylcholinesterase inhibitors. *J. Med. Chem.*, 56, 7615-7624.
- Bissantz, C., Kuhn, B., Stahl, M., 2010, A medicinal chemist's guide to molecular interactions. *J. Med. Chem.*, 53, 5061-5084.
- Cannon, EO., Amini, A., Bender, A., Sternberg, MJE., Muggleton, SH., Glen, RC., Mitchell, JBO., 2007, Support vector inductive logic programming outperforms the naive Bayes classifier and inductive logic programming for the classification of bioactive chemical compounds. *J. Comput. Aided Mol. Des.*, 21, 269-280.
- Cappel, D., Dixon, SL., Sherman, W., Duan, J., 2015, Exploring conformational search protocols for ligand-based virtual screening and 3-D QSAR modeling. *J. Comput. Aided Mol. Des.*, 29, 165-182.
- ChemAxon, 2014. MarvinSketch 14.11.10.0. <https://www.chemaxon.com/products/marvin/marvinsketch>.
- Chen, Y., 2015, Beware of docking! *Trends Pharmacol. Sci.*, 36, 78-95.
- Colovic, MB., Krstic, DZ., Lazarevic-Pasti, TD., Bondzic, AM., Vasic, VM., 2013,

- Acetylcholinesterase inhibitors: Pharmacology and toxicology. *Curr. Neuropharmacol.*, 11, 315-335.
- de Graaf, C., Kooistra, AJ., Vischer, HF., Katritch, V., Kuijer, M., Shiroishi, M., Iwata, S., Shimamura, T., Stevens, RC., de Esch, IJP., Leurs, R., 2011. Crystal structure-based virtual screening for fragment-like ligands of the human histamine H<sub>1</sub> receptor. *J. Med. Chem.*, 54, 8195-8206.
- Desaphy, J., Raimbaud, E., Ducrot, P., Rognan, D., 2013, Encoding protein-ligand interaction patterns in fingerprints and graphs. *J. Chem. Inf. Model.*, 53, 623-637.
- Dvir, H., Wong, DM., Harel, M., Barril, X., Orozco, M., Luque, FJ., Munoz-Torrero, D., Camps, P., Rosenberry, TL., Silman, I., Sussman, JL., 2002, 3D Structure of *Torpedo californica* acetylcholinesterase complexed with Huprine X at 2.1 Å resolution: Kinetic and molecular dynamic correlates. *Biochemistry*, 41, 2970-2981.
- Ece, A., Pejin, B., Ece, A., Pejin, B., 2015, A computational insight into acetylcholinesterase inhibitory activity of a new lichen depsidone. *J. Enzym. Inhib. Med. Chem.*, 30, 528-532.
- Furtmann, N., Hu, Y., Gütschow, M., Bajorath, J., 2015, Identification and analysis of the currently available high-confidence three-dimensional activity cliffs. *RSC Adv.*, 5, 43660-43668.
- Gabel, J., Desaphy, J., Rognan, D., 2014, Beware of machine learning-based scoring functions-on the danger of developing black boxes. *J. Chem. Inf. Model.*, 54, 2807-2815.
- Imran, S., Taha, M., Ismail, N.H., Kashif, S.M., Rahim, F., Jamil, W., Hariono, M., Yusuf, M., Wahab, H., 2015, Synthesis of novel flavone hydrazones: *In-vitro* evaluation of  $\alpha$ -glucosidase inhibition, QSAR analysis and docking studies. *Eur. J. Med. Chem.*, 105, 156-170.
- Istyastono, EP., 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, EP., 2016, Optimizing structure-based virtual screening protocol to identify phytochemicals as cyclooxygenase-2 inhibitors. *Indones. J. Pharm.*, 27, 163-173.
- Istyastono, EP., de Graaf, C., de Esch, IJP., Leurs, R., 2011<sup>a</sup>, Molecular determinants of selective agonist and antagonist binding to the histamine H<sub>4</sub> receptor. *Curr. Top. Med. Chem.*, 11, 661-679.
- Istyastono, EP., Kooistra, AJ., Vischer, H., Kuijer, M., Roumen, L., Nijmeijer, S., Smits, R., de Esch, I., Leurs, R., de Graaf, C., 2015<sup>a</sup>, Structure-based virtual screening for fragment-like ligands of the G protein-coupled histamine H<sub>4</sub> receptor. *Med. Chem. Commun.*, 6, 1003-1017.
- Istyastono, EP., Nijmeijer, S., Lim, HD., van de Stolpe, A., Roumen, L., Kooistra, AJ., Vischer, HF., de Esch, IJP., Leurs, R., de Graaf, C., 2011<sup>b</sup>. Molecular determinants of ligand binding modes in the histamine H<sub>4</sub>receptor: Linking ligand-based three-dimensional quantitative structure-activity relationship (3D-QSAR) models to *in silico* guided receptor mutagenesis studies. *J. Med. Chem.*, 54, 8136-8147.
- Istyastono, EP., Yuniarti, N., Hariono, M., Yuliani, SH., Riswanto, FDO., 2017, Binary quantitative structure-activity relationship analysis in retrospective structure-based virtual screening campaigns targeting estrogen receptor alpha, (Manuscript in preparation).
- Istyastono, EP., Riswanto, FDO., Yuliani, SH., 2015<sup>b</sup>, Computer-aided drug repurposing: A cyclooxygenase-2 inhibitor celecoxib as a ligand for estrogen receptor alpha. *Indones. J. Chem.*, 15, 274-280.
- Istyastono, EP. and Setyaningsih, D., 2015, Construction and retrospective validation of structure-based virtual screening protocols to identify potent ligands for human adrenergic  $\beta_2$  receptor. *Indones. J. Pharm.*, 26, 20-28.
- Istyastono, EP. and Yuniarti, N., 2016, Construction of three dimensional structures of phytoestrogens converted

- from smiles string representations for simulations using PLANTS docking software. *Trad. Med. J.*, 21, 69-76.
- Jayapal, MR. and Sreedhar, NY., 2010, Synthesis and characterization of 4-hydroxy chalcones by aldol condensation using  $\text{SOCl}_2/\text{EtOH}$ , *Int. J. Curr. Pharm. Res.*, 2, 2-4.
- Jongejan, A., Lim, HD., Smits, RA., de Esch, IJP., Haaksma, E., Leurs, R., 2008, Delineation of agonist binding to the human histamine  $\text{H}_4$  receptor using mutational analysis, homology modeling, and *ab initio* calculations. *J. Chem. Inf. Model.*, 48, 1455-1463.
- Korb, O., Stützel, T., Exner, TE., 2007, An ant colony optimization approach to flexible protein-ligand docking. *Proc. IEEE Swarm Intell. Symp.*, 1, 115-134.
- Korb, O., Stützel, T., Exner, TE., 2009, Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model.*, 49, 84-96.
- Koshland, DE., 1994, The key-lock theory and the induced fit theory. *Angew. Chem. Int. Ed. Engl.*, 33, 2375-2378.
- Lill, MA. and Danielson, ML., 2011, Computer-aided drug design platform using PyMOL. *J. Comput. Aided Mol. Des.*, 25, 13-19.
- Lim, HD., Istyastono, EP., van de Stolpe, A., Romeo, G., Gobbi, S., Schepers, M., Lahaye, R., Menge, WMBP., Zuiderveld, OP., Jongejan, A., Smits, RA., Bakker, RA., Haaksma, EEJ., Leurs, R., de Esch, IJP., 2009, Clobenpropit analogs as dual activity ligands for the histamine  $\text{H}_3$  and  $\text{H}_4$  receptors: synthesis, pharmacological evaluation, and cross-target QSAR studies. *Bioorg. Med. Chem.*, 17, 3987-3994.
- Marcou, G. and Rognan, D., 2007, Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.*, 47, 195-207.
- Mehta, M., Adem, A., Sabbagh, M., 2012, New acetylcholinesterase inhibitors for Alzheimer's disease. *Int. J. Alzheimers Dis.*, 2012, 1-8.
- Mysinger, MM., Carchia, M., Irwin, JJ., Shoichet, BK., 2012, Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.*, 55, 6582-6594.
- O'Boyle, NM., Banck, M., James, CA., Morley, C., Vandermeersch, T., Hutchison, GR., 2011, Open Babel: An open chemical toolbox. *J. Cheminform.*, 3, 33-47.
- Picciotto, MR., Higley, MJ., Mineur, YS., 2012, Acetylcholine as a neuromodulator: Cholinergic signaling shapes nervous system function and behavior. *Neuron*, 76, 116-129.
- Powers, D., 2011. Evaluation: From precision, recall and *F*-measure to ROC, informedness, markedness and correlation. *J. Mach. Learn. Tech.*, 2, 37-63.
- Prince, M., Wimo, A., Guerchet, M., Ali, G., WU, Y., Prina, M., 2015, *World Alzheimer Report 2015. The Global Impact of Dementia. An Analysis of Prevalance, Incidence, Cost and Trends.* Alzheimer's Disease International, London.
- R Core Team, 2016. R: A language and environment for statistical computing. Vienna. <http://www.r-project.org>.
- Radifar, M., Yuniarti, N., Istyastono, EP., 2013<sup>a</sup>. PyPLIF-assisted redocking indomethacin-(R)-alpha-ethyl-ethanolamide into cyclooxygenase-1. *Indones. J. Chem.*, 13, 283-286.
- Radifar, M., Yuniarti, N., Istyastono, EP., 2013<sup>b</sup>. PyPLIF: Python-based protein-ligand interaction fingerprinting. *Bioinformatics*, 9, 325-328.
- Rizzi, L., Rosset, I., Roriz-Cruz, M., 2014. Global epidemiology of dementia: Alzheimer's and vascular types. *Biomed Res. Int.*, 2014, 1-8.
- Setiawati, A., Riswanto, FDO., Yuliani, SH., Istyastono, EP., 2014, 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.
- Singh, M., Kaur, M., Kukreja, H., Chugh, R., Silakari, O., Singh, D., 2013, Acetylcholinesterase inhibitors as Alzheimer therapy: From nerve toxins to neuroprotection. *Eur. J. Med. Chem.*, 70, 165-188.

- Sirci, F., Istyastono, EP., Vischer, HF., Kooistra, AJ., Nijmeijer, S., Kuijer, M., Wijtmans, M., Mannhold, R., Leurs, R., de Esch, IJP., de Graaf, C., 2012, Virtual fragment screening: Discovery of histamine H<sub>3</sub> receptor ligands using ligand-based and protein-based molecular fingerprints. *J. Chem. Inf. Model.*, 52, 3308-3324.
- Smits, RA., Adami, M., Istyastono, EP., Zuiderveld, OP., van Dam, CME., de Kanter, FJJ., Jongejan, A., Coruzzi, G., Leurs, R., de Esch, IJP., 2010, Synthesis and QSAR of quinazoline sulfonamides as highly potent human histamine H<sub>4</sub> receptor inverse agonists. *J. Med. Chem.*, 53, 2390-2400.
- Stoddard, BL., Koshland, DE., 1992. Prediction of the structure of a receptor-protein complex using a binary docking method. *Nature*, 358, 774-776.
- Sukumaran, S., Chee, C., Viswanathan, G., Buckle, M., Othman, R., Rahman, NA., Chung, L., 2016, Synthesis, biological evaluation and molecular modelling of 2'-hydroxychalcones as acetylcholinesterase inhibitors. *Molecules*, 21, 1-10.
- Tabet, N., 2006, Acetylcholinesterase inhibitors for Alzheimer's disease: Anti-inflammatories in acetylcholine clothing! *Age Ageing*, 35, 336-338.
- Tarcsay, A., Paragi, G., Vass, M., Jójárt, B., Bogár, F., Keserű, GM., 2013, The impact of molecular dynamics sampling on the performance of virtual screening against GPCRs. *J. Chem. Inf. Model.* 53, 2990-2999.
- ten Brink, T., Exner, TE., 2009, Influence of protonation, tautomeric, and stereoisomeric states on protein-ligand docking results. *J. Chem. Inf. Model.*, 49, 1535-1546.
- Therneau T., Atkinson B., Ripley B., 2015, rpart: Recursive partitioning and regression trees. R package version 4.1-9. <https://CRAN.R-project.org/package=rpart>
- Tran, T., Nguyen, T., Nguyen, N., Nguyen, D., Nguyen, T., Le, M., Thai, K., 2016, Synthesis of novel chalcones as acetylcholinesterase inhibitors. *Appl. Sci.*, 6, 1-11.
- Tsuda, K., 2012. Renin-angiotensin system and sympathetic neurotransmitter release in the central nervous system of hypertension. *Int. J. Hypertens.*, 2012, 1-8.
- Wang, A., Stout, CD., Zhang, Q., Johnson, EF., 2015. Contributions of ionic interactions and protein dynamics to cytochrome P450 2D6 (CYP2D6) substrate and inhibitor binding. *J. Biol. Chem.*, 290, 5092-5104.
- Zambre, AP., Ganure, AL., Shinde, DB., Kulkarni, VM., 2007, Perspective assessment of COX-1 and COX-2 selectivity of nonsteroidal anti-inflammatory drugs from clinical practice: Use of genetic function approximation. *J. Chem. Inf. Model.*, 47, 635-643.