BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 6(4)

The importance of ARG513 as a hydrogen bond anchor to discover COX-2 inhibitors in a virtual screening campaign

Nunung Yuniarti[1], Zullies Ikawati[1], Enade Perdana Istyastono[2]*

[1]Department of Pharmacology and Clinical Pharmacy, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, Indonesia; [2]Molecular Modeling Division "MOLMOD.ORG", Pharmaceutical Technology Laboratory, Universitas Sanata Dharma, Yogyakarta, Indonesia; Enade Perdana Istyastono - Email: enade@usd.ac.id; Phone: +62-274-883037; Fax: +62-274-886529; *Corresponding author

Received March 15, 2011; Accepted April 25, 2011; Published May 07, 2011

Abstract:
Structure-based virtual screening (SBVS) protocols were developed to find cyclooxygenase-2 (COX-2) inhibitors using the Protein-Ligand ANT System (PLANTS) docking software. The directory of useful decoys (DUD) dataset for COX-2 was used to retrospectively validate the protocols; the DUD consists of 426 known inhibitors in 13289 decoys. Based on criteria used in the article describing DUD datasets, the default protocol showed poor results. However, having ARG513 as a hydrogen bond anchor increased the quality of the SBVS protocol. The modified protocol showed results that could be well considered, with a maximum enrichment factor (EF_{max}) value of 32.2.

Keywords: virtual screening, cyclooxygenase-2, docking

Background:
Molecular docking, used in structure-based virtual screening (SBVS) campaigns, is considered a powerful tool in drug discovery **[1-3]**. Some drugs on the market, e.g., Dorzolamide® (Merck & Co) **[4]**, Saquinavir® (Hoffmann-La Roche) **[5]**, Indinavir® (Merck & Co.) **[5]**, Ritonavir® (Abbott Laboratories) **[5]** and Nelfinavir® (Agouron Pharmaceuticals) **[5]**, were discovered or designed using this approach **[2]**. Some limitations of the molecular docking method have been identified and have led to fruitful discussions **[6, 7]**. The primary challenge was to improve the quality of the scoring functions, which determines the ranking of the compounds in SBVS **[6, 7]**. Most of the available scoring functions were developed as all-purpose models, which can presumably be applied in all protein-ligand complexes. However, despite the considerable progress in the development of scoring functions, they are still far from being universally accurate **[6, 7]**. Therefore, using prior knowledge and targeted scoring functions in the construction of the SBVS protocols can be a useful solution **[6]**. Cyclooxygenase-2 (COX-2) is an enzyme that plays an important role in inflammatory processes **[8]**. The dataset, which was compiled from the ZINC database (http://zinc.docking.org/) to retrospectively validate and compare SBVS protocols to find COX-2 inhibitors, is publicly available in the directory of useful decoys (http://dud.docking.org/r2/cox2.tar.gz). The enzyme has been considered a challenging target for SBVS campaigns **[9, 10]**. For the first SBVS campaign, a dataset from the directory of useful decoys (DUD) was used to find COX-2 inhibitors by employing the DOCK 3.5.54 docking software. The results showed a maximum enrichment factor (EF_{max}) value and an enrichment factor value at 20% of the database (EF_{20}) of 29.1 and 3.3, respectively **[9]**, which were considered as good results **[9]**. We had previously described the applicability of the Protein-Ligand ANT System (PLANTS) docking software to reproduce the crystal structure interaction of 1-phenylsulfonamide-3-trifluoromethyl-5-parabromophenyl-pyrazole (1) with COX-2 (pdb code: 6COX; http://dx.doi.org/10.2210/ pdb6cox/pdb) **[11, 12]**. In this article, we present the retrospective validation of a developed protocol to perform as a screening tool for the discovery of COX-2 inhibitors using SBVS campaigns. Additionally, the construction of a modified protocol is displayed by inserting an additional constraint to the hydrogen bond formed by ligands to ARG513 of the COX-2 since previously published studies indicate that this particular interaction is important for COX-2 inhibitors **[11, 13]**. Although the default protocol could reproduce the structural interactions from the crystal structure **[12]**, the results showed that the protocol showed poor results as an SBVS campaign to discover COX-2 inhibitors. Conversely, the modified protocol resulted in excellent EF_{max} and EF_{20} values **[9]**.

Methodology:
The COX-2 ligands and decoys (426 known inhibitors in 13289 decoys) were obtained from the DUD website (http://dud.docking.org/r2/) **[9]**. The target enzyme, virtual COX-2 (protein.mol2), and the PLANTS configuration file (plantsconfig) were obtained from a previous study **[12]**. The PLANTS docking software v1.1 was used as the SBVS docking tool (http://www.tcd.uni-konstanz.de/research/plants.php). The receiver operator characteristic (ROC) curves were calculated using R statistical computing software version 2.11.1 (http://www.r-project.org/). All computational simulations were performed on an HP-xw6600 workstation with Intel Xeon E5420/2.5 GHz Quadcores as the processors, 8 GB of RAM and a Linux version 2.6.26-2-amd64 (Debian 2.6.26-26lenny1) as the operating system. The compounds were virtually screened using the PLANTS docking software by employing a previously developed virtual COX-2 as the target **[12]**. Two independent simulations were performed: (i) original protocol (using previously described config file

(plantsconfig)) **[12]** and (ii) modified protocol (original protocol with a constraint to atom number 7808 (Hydrogen bond to ARG513; mod_plantsconfig). The compounds were then ranked based on their scores, and ROC curves were calculated and generated. The quality of the screening procedures was judged using criteria introduced by Huang *et al.* **[9]**. The paired *t*-test of the true positive rate values was performed to analyze the importance of hydrogen bonding at ARG513, which would aid in the discovery of COX-2 inhibitors. The configuration files used in these screenings are available in **Supplementary material**.

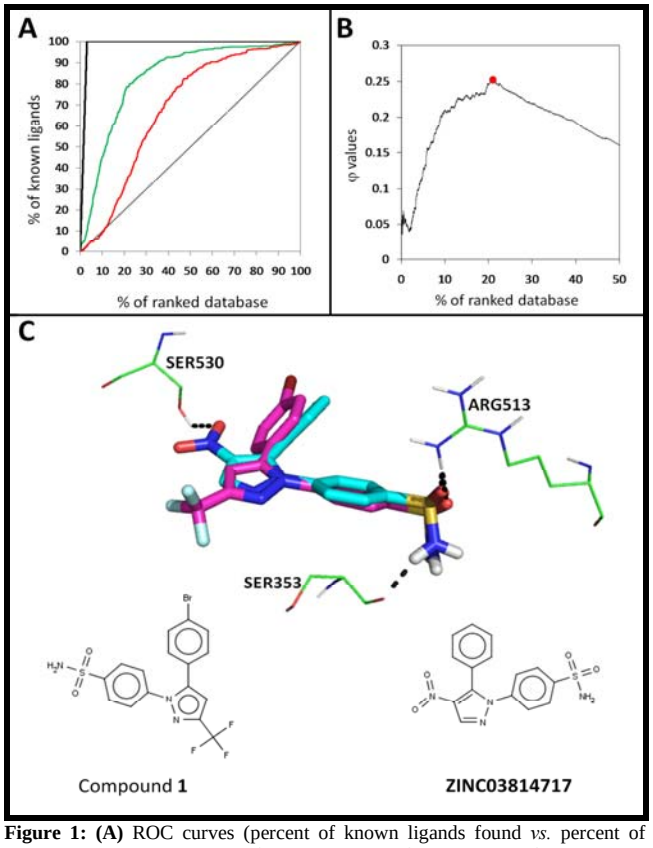

Figure 1: (A) ROC curves (percent of known ligands found *vs.* percent of ranked database). The results from the ideal, default, and modified protocols and random picking are presented in solid black, green, red and dashed black curves, respectively. **(B)** Curve of φ values *vs.* percent of ranked database. A red point indicates the coordinate of the suggested reference compound, ZINC03814717. **(C)** Compound ZINC03814717 (the 2D structure is shown as an inset) in the binding pocket of COX-2 together with the crystal structure of compound 1 (the 2D structure is shown as an inset) **[11]**. Only residues forming hydrogen bonds are shown here. The residues are presented as lines, and the ligands are presented as sticks. Carbon atoms are presented in green for COX-2, magenta for compound 1 and cyan for compound ZINC03814717. Polar hydrogen atoms are presented in white, fluorine atoms are presented in light blue, nitrogen atoms are presented in blue, oxygen atoms are presented in red, and sulfur atoms are presented in yellow. Dashed black lines indicate the hydrogen bonds. The 3D figure was created using PyMOL 1.2 (http://www.pymol.org/).

Results and Discussion:

The ROC curves are presented in **Figure 1A**. The figure shows that the modified protocol provided better results compared to the results from the default protocol. The EF_{max} and EF_{20} values of the SBVS using the default protocol were 1.8 and 1.5, respectively. However, the modified protocol showed EF_{max} and EF_{20} of 32.2 and 3.7, respectively. Referring to standards determined by Huang *et al.* **[9]**, the quality of the SBVS default protocol has been considered poor, while the modified protocol can be considered very good. Moreover, the paired t-test of true positive values provided a significance value of less than 0.05. This showed that the modified protocol at a 95% level of confidence differed significantly compared to the default method for the screening COX-2 inhibitors. This data supports previous hypotheses that the hydrogen bond formation between ligands and ARG513 is important for COX-2 inhibition **[11, 13]**. This also confirmed that inputting prior knowledge into the protocol could increase the quality of the SBVS **[6]**. Furthermore, this result provided a valid protocol that could be used to perform virtual screening to design and discover COX-2 inhibitors. The computational time needed to perform each retrospective validation is about 80 hours. It is not efficient to perform similar validation to avoid calculation instability caused by differences in the computer systems used in the screening. Therefore, we suggest selecting a single compound as the reference compound. The compound can be selected by calculating the Matthews correlation coefficient (φ) values **(Figure 1B)**. The reference compound was the compound with the highest φ value, which indicated the best discrimination ability compared to other known inhibitors. The modified protocol showed that the selected reference compound was ZINC03814717 (shown as an inset in **Figure 1C**), with an enrichment factor value of 3.7 ($EF_{21.01}$). The accuracy and sensitivity of the SBVS for $EF_{21.01}$ were 80.7% and 78.4%, respectively. Compound ZINC03814717 has been reported as a COX-2 inhibitor with an IC_{50} value of 290 nM or pIC_{50} of 6.54 **[14]**. **Figure 1C** shows the docking pose of ZINC03814717, using the modified protocol, in the COX-2 binding pocket together with the crystal structure of compound **1**. Similar to the pose of compound **1** (shown as an inset in Figure 1C), the sulfonamide moiety formed a hydrogen bond to ARG513. The sulfonamide moiety showed an additional hydrogen bond at the backbone of SER353. Remarkably, the nitro group of ZINC03814717 showed a hydrogen bond to SER530, which is known as a binding anchor of COX-2 ligands **[15]**. This interaction was not observed in the binding pose of compound **1** to COX-2 **[11, 12]**. This indicated that there were other important interactions that could be used to increase the quality of the SBVS. Further experiments were suggested to optimize the SBVS by varying the constraint weight and the binding residues.

Conclusion:

A retrospective SBVS validation using the DUD dataset for COX-2 inhibitors showed that, although the docking protocol, i.e., the default setting on the PLANTS docking software, was able to accurately reproduce the crystal structure interaction between compound **1** and COX-2 **[12]**, it needed additional constraints in the protocol to perform better SBVS. This study confirmed the importance of interactions between COX-2 inhibitors and ARG513. Using criteria introduced by Huang *et al.* **[9]**, the quality of the SBVS approach increased.

Acknowledgements:

The authors thank Dr. Chris de Graaf, Dr. Iwan de Esch, Dr. Luc Roumen and Albert Kooistra at the Medicinal Chemistry Division, VU University Amsterdam for their helpful discussion. This work was supported by the Directorate of Higher Education, of the Department of Education of the Government of the Indonesian Republic, through the *Hibah Kompetensi* 2010 Research Grant.

References:

[1] Huang SY & Zou X. *Int J Mol Sci.* 2010 **11**: 3016 [PMID: 21152288]
[2] Kubinyi H. *J Recept Signal Transduct Res.* 1999 **19**: 15 [PMID: 10071748]
[3] Kitchen DB *et al. Nat Rev Drug Discov.* 2004 **3**: 935 [PMID: 15520816]
[4] Greer J *et al. J Med Chem.* 1994 **37**: 1035 [PMID: 8164249]
[5] Vacca JP & Condra JH. *Drug Discov Today.* 1997 **2**: 261
[6] Seifert MH *Drug Discov Today.* 2009 **14**: 562 [PMID: 19508918]
[7] Warren GL *et al. J Med Chem.* 2006 **49**: 5912 [PMID: 17004707]
[8] Agarwal S *et al. Expert Rev Clin Immunol.* 2009 **5**: 145 [PMID: 20477063]
[9] Huang N *et al. J Med Chem.* 2006 **49**: 6789 [PMID: 17154509]
[10] Sudha KN *et al. Bioinformation* 2008 **2**: 325 [PMID: 18685718]
[11] Kurumbail RG *et al. Nature* 1996 **384**: 644 [PMID: 8967954]
[12] Prasojo SL *et al. Indonesian Journal of Chemistry* 2010 **10**: 348
[13] Levita J *et al. ITB Journal of Science* 2009 **41 A**: 110
[14] Penning TD *et al. J Med Chem.* 1997 **40**: 1347 [PMID: 9135032]
[15] Rowlinson SW *et al. J Biol Chem.* 2003 **278**: 45763 [PMID: 12925531]

Edited by P Kangueane
Citation: Yuniarti *et al.* Bioinformation 6(4): 164-166 (2011)

Supplementary material:

plantsconfig:

```
# scoring function and search settings
scoring_function chemplp
search_speed speed1


# input
protein_file protein.mol2
ligand_file cox2_all_dud.mol2

# output
output_dir results

# write single mol2 files (e.g., for RMSD calculation)
write_multi_mol2 0

# binding site definition
bindingsite_center 23.6651 23.3126 47.8265
bindingsite_radius 11.3323

# cluster algorithm
cluster_structures 10
cluster_rmsd 2.0
```

mod_plantsconfig:

```
# scoring function and search settings
scoring_function chemplp
search_speed speed1


# input
protein_file protein.mol2
ligand_file cox2_all_dud.mol2

# output
output_dir results_mod

# write single mol2 files (e.g., for RMSD calculation)
write_multi_mol2 0

# binding site definition
bindingsite_center 23.6651 23.3126 47.8265
bindingsite_radius 11.3323

#constraint
chemplp_protein_hb_constraint 7808 10

# cluster algorithm
cluster_structures 10
cluster_rmsd 2.0
```