

## Virtual screening for novel COX-2 inhibitors using the ZINC database

Kotaprolu Naga Sudha<sup>1,\*</sup>, Mohammed Shakira<sup>1</sup>, Paturi Prasanthi<sup>1</sup>, Nalla Sarika<sup>1</sup>, Ch Narasimha Kumar<sup>2</sup> and Padavala Ajay Babu<sup>2</sup>

<sup>1</sup>Department of Microbiology, Sir C. R Reddy College for Women, Vatluru, Eluru 534125, India; <sup>2</sup>ProGene Biosciences, Institute of Bioinformatics and Research Centre, 103, Bharat Towers, Dwaraka Nagar, Visakhapatnam 530016, India; Kotaprolu Naga Sudha\* - E-mail: subha\_vsk29@yahoo.co.in; \* Corresponding author

received March 20, 2008; revised April 24, 2008; accepted April 25, 2008; published May 13, 2008

### Abstract:

Cyclooxygenase-2 (COX-2) enzyme binds to arachidonic acid and releases metabolites that are used to induce pain and inflammation. COX-2 selective inhibitors such as celecoxib, rofecoxib and valdecoxib are currently used to reduce inflammatory response. However, they lack anti-thrombotic activity and hence lead to cardiovascular and renal liabilities apart from gastrointestinal irritation. Therefore, there is still a need to develop more potent COX-2 inhibitors. In this paper, we report the screening of various compounds from the ZINC database (contains 4.6 million small molecule compounds) using the eHiTS (electronic High Throughput Screening) software tool against the COX-2 protein. The strategy employed can be conveniently split into two categories, viz. screening and docking, respectively. Screening was performed using molecular constraints tool to filter compounds with physico-chemical properties similar to the 6COX bound ligand SC-558. The analysis resulted in 1042 Lipinski compliant hits which are docked and scored to identify structurally novel ligands that make similar interactions to those of known ligands or may have different interactions with other parts of the binding site. Our screening approach identified two molecules ZINC00663976 (eHiTS score of -7.135 kcal/mol) and ZINC02062094 (eHiTS score of -7.242 kcal/mol) from the ZINC database. Their energy scores are better than the 6COX bound co-crystallized ligand SC-558 with an eHiTS score of -6.559 kcal/mol. Both the ligands were docked within the binding pocket forming interactions with Leu352, Phe518, Met522, Val523, Ala527 and Ser353. Visual inspection suggested similar orientation and binding mode for ZINC02062094 with SC-558 ligand. The NH group of the ligand formed hydrogen bond interactions with the backbone NH of Ala527.

**Keywords:** cyclooxygenase; ZINC database; virtual screening; e-HiTS

### Background:

COX-2 is expressed after an inflammatory stimulus and releases metabolites that are used to induce pain and inflammation. During normal physiology, COX-2 levels are undetectable whereas during periods of acute and chronic inflammation, the level of COX-2 is significantly higher. NSAIDs (Non-Steroidal Anti-Inflammatory Drugs) exhibit their effect through inhibition of cyclooxygenase (COX) enzymes by blocking the synthesis of prostaglandins from arachidonic acid [1]. Conventional NSAIDs are profoundly used in the treatment of wide variety of inflammatory conditions and they act by the inhibition of cyclooxygenase (COX), the enzyme involved in the biosynthesis of prostaglandins, prostacyclins and thromboxanes from arachidonic acid [2, 3].

It has been shown recently that the COX enzyme exists in three isoforms - COX-1, COX-2 and COX-3. COX-1 primarily responsible for cytoprotection and COX-2, the inducible form is associated with inflammation, whereas, COX-3 has no importance in the development of inflammation. COX-1 enzyme is responsible for maintaining homeostasis (gastric and renal integrity), whereas COX-2 induces the inflammatory symptoms in response to inflammatory stimuli [4]. The two COX isoforms are about 60% homologous. The ability to inhibit one isoform selectively is attributed to the different amino acids at position 523, isoleucine in COX-1 and valine in

COX-2 [5]. Therefore, selective COX-2 inhibition would reduce the undesired side effects such as gastro-intestinal disorders, ulcers and renal failure [6].

Several classes of compounds having selective COX-2 inhibitory activity have been reported in the literature, for example, diaryl heterocyclics as oxazoles, thiophenes, pyrazoles, imidazoles [7], and common classical agents modified to have selective COX-2 inhibitory activity such as esters and amides of indomethacin and meclofenamic acid [4]. The classical NSAIDs produce their adverse effects via inhibition of COX-1 isoform, hence many investigations have been directed to find compounds able to act as selective COX-2 inhibitors such as 6COX bound SC-558, celecoxib, rofecoxib, valdecoxib and more recently, nitroxy substituted 1,5-diarylimidazoles [8], phenylazobenzenesulfonamides [9], respectively. However, evidence suggests that adverse reactions such as gastro-intestinal irritation or ulceration and renal liabilities are associated with prolonged use of COX-2 selective inhibitors [10]. These inhibitors are also known to suppress synthesis of prostacyclin, a potent vasodilator, gastroprotectant, and platelet inhibitor, via the inhibition of endothelial COX-2. The COX-2 selective inhibitors intrinsically lack anti-thrombotic activity and some cardiovascular liabilities have been found associated with them in preclinical studies [11]. Thus, there is still a need for novel, selective, and potent COX-2 inhibitors

with an improved profile compared to current COX-2 inhibitors.

Traditional synthesis of a series of new compounds utilizing combinatorial chemistry and high-throughput screening can be carried out at high cost and also are time consuming; whereas on the other hand, screening small molecule databases for novel compounds represents an alternative process. Docking various ligands to the protein of interest followed by scoring to determine the affinity of binding and to reveal the strength of interaction has become increasingly important in the context of drug discovery. Screening large databases of compounds can provide a feasible, alternative technique against high-throughput screening, but depends on the fast and accuracy of the docking algorithm. Hence, in this paper we report screening a library of compounds from ZINC database [12] against COX-2 protein, 6COX (PDB ID) with bound ligand SC-558 extracted from protein data bank, by utilizing a fast, exhaustive docking software eHiTS (electronic High Throughput Screening) [13].

Virtual screening is now established as an effective paradigm for filtering compounds for drug discovery process [14]. The technique employed is based on the prediction of binding modes and binding affinities of each compound in the dataset by means of docking to an X-ray crystallographic structure. Virtual screening utilizes docking and scoring of each compound from a dataset. Some recent studies [15] have focused on certain crucial factors such as the size and diversity of the ligand dataset, wide range of targets and the evaluation of docking programs. Various studies reported in literature stated the importance of dataset size such as 10,000 compounds using FlexX [16], 44,000 compounds using Surflex [17] and several others.

In general, it is important to visualize the docked poses of high-scoring compounds because many ligands are docked in different orientations and may often miss interactions that are known to be important for the target receptor. This sort of study becomes more difficult as the size of the dataset increases. Therefore, an alternative approach is to eliminate unpromising compounds before docking by restricting the dataset to drug-like compounds, by filtering the dataset based on appropriate property and sub-structural features and by performing diversity analysis. These approaches can be highly effective in reducing the dataset to be docked to the order of  $10^3$  to  $10^4$  compounds [18].

## Methodology:

### Receptor X-ray structure

The 3D coordinates of the crystal structure of COX-2 in complex with SC-558 inhibitor (PDB code: 6COX) was selected as the receptor model in virtual screening program. We used the chemical compound library, ZINC database and the docking program eHiTS for the study.

### Ligand ZINC database

ZINC, an acronym for 'ZINC is not commercial, a free database for virtual screening' contains over 4.6 million compounds in ready-to-dock, 3D formats, available at the URL <http://zinc.docking.org>. Molecules in ZINC are annotated by molecular property that include molecular weight, number of

rotatable bonds, calculated LogP, number of hydrogen-bond donors, hydrogen-bond acceptors, chiral centers, chiral double bonds (E/Z isomerism), polar and apolar desolvation energy (in kcal/mol), net charge and rigid fragments. The database contains 494,915 Lipinski compliant molecules and 202,134 'lead-like' molecules, having molecular weight in the range 150 to 350 with calculated LogP < 4, number of hydrogen-bond donors  $\leq 3$  three, and number of hydrogen-bond acceptors  $\leq 6$ . ZINC provides several search criteria such as molecular property constraint, ZINC codes, vendor based, and molecular substructure search.

### eHiTS software

eHiTS (<http://www.simbiosys.ca>) has a novel flexible ligand docking method that is exhaustive on the conformations and poses that avoid severe steric clashes between receptor and ligand. The algorithm generates all major docking modes that are compatible with the steric and chemistry constraints. First, the binding pocket is determined by building a steric grid for the whole receptor, dividing regions into separate pockets and identifying the possible interaction sites. Then, a cavity description is built that consists of thousands of geometric shapes (polyhedra). The ligand is divided into rigid fragments and connecting flexible chains. The eHiTS docks all rigid fragments to all possible places in the cavity independently of each other.

### Screening

An exhaustive matching of compatible rigid fragment pose sets is performed by a rapid hyper-graph clique detection algorithm. This yields a few hundred to several million acceptable combinations of poses. This approach makes the search exhaustive, i.e., it finds all possible solutions. The accuracy is set to the highest (accuracy = 6) during docking. eHiTS has three different scoring functions: (a) a simple and fast chemical flag based statistical scoring function; (2) empirical-scoring function and (3) a third scoring function that combines both statistical and empirical components along with grid based geometrical terms, entropy loss estimation and receptor surface area. To visualize different poses, CheVi software, a graphical user interface was utilized.

Before screening ZINC database, the docking protocol was validated. 6COX protein bound ligand SC-558 was docked into the binding pocket to obtain the docked pose and the RMSD (Root Mean Square Deviation) of all atoms between these two conformations is 0.51 Å indicating that the parameters for docking simulation are good in reproducing the X-ray crystal structure. Therefore, the ZINC database was screened for compounds similar to SC-558 structural features (structure based search) and by providing molecular constraints (property based search). The physico-chemical properties such as log P value, H-bond donors, H-bond acceptors, molecular weight and rotational bonds, for SC-558 ligand are calculated using the ADME/Tox filter server [19].

### Discussion:

We searched the ZINC database using structural features that are similar to SC-558 and no hit was obtained. We then repeated the search based on property and identified 1042 hits that are Lipinski [20] compliant. We docked these 1042 compounds using eHiTS and evaluated binding compatibility

with the receptor based on docked energy (in kcal/mol). The docking tool generated 32 conformers for each docked molecule in about 1-2 minutes of CPU time. The virtual screening technique employed in this study identified diverse, yet specific ligands that bind in a comparable manner similar to SC-558 binding for COX-2. This approach identified two compounds ZINC00663976 and ZINC02062094 from the ZINC database for compatible binding with COX-2 (Table 1 in supplementary material).

SC-558, a diaryl heterocyclic inhibitor has 1900-fold selectivity for COX-2 over COX-1. It has a central pyrazole ring and a sulphonamide substituent attached to one of the aryl rings. The bromophenyl ring of SC-558 is bound in a hydrophobic cavity formed by Phe381, Leu384, Tyr385, Trp387, Phe518, Met522, Val523 and Ser530. The trifluoromethyl group is bound in an adjacent pocket formed by Met113, Val116, Val349, Tyr355, Leu359 and Leu531. The sulphonamide group extends into a relatively polar region and interacts with His90, Gln192, Leu352 and Ser353. One of the oxygen atoms of SC-558 forms a hydrogen bond to NE2 atom of His90 and the amide nitrogen forms three hydrogen bonds to OE1 of Gln192, and carbonyl oxygens of Leu352 and Ser353, respectively. These features of SC-558 binding are quite important in studying the binding modes of various selective and non-selective COX-2 inhibitors.

### Binding modes

The surface point types of two ligands along with their docked energies are given in Table 1 (see supplementary material). Active site of COX-2 offers many different binding modes for these compounds as they are strongly dependent on the attached substituent. Both the ligands were docked deeply within the binding pocket region forming interactions with Leu352, Phe518, Met522, Val523, Ala527 and Ser353, respectively. In Table 1 (under supplementary material), the binding affinities and the possible number of interactions based on chemical property identifiers (surface point types) are reported. Our screening approach identified two molecules ZINC00663976 (eHITS score of -7.135 kcal/mol) and ZINC02062094 (eHITS score of -7.242 kcal/mol) from the ZINC database. Their energy scores are better than the 6COX bound co-crystallized ligand SC-558 with an eHITS score of -6.559 kcal/mol.

The orientation of the docked pose of ZINC00663976 suggests that the binding site was partially unoccupied due to lack of substitutions on the molecule. The trifluoro substitution of ZINC00663976 is oriented towards bromophenyl ring region and the halogen substituted benzene ring occupied polar region. The other molecule, ZINC02062094 occupied all the

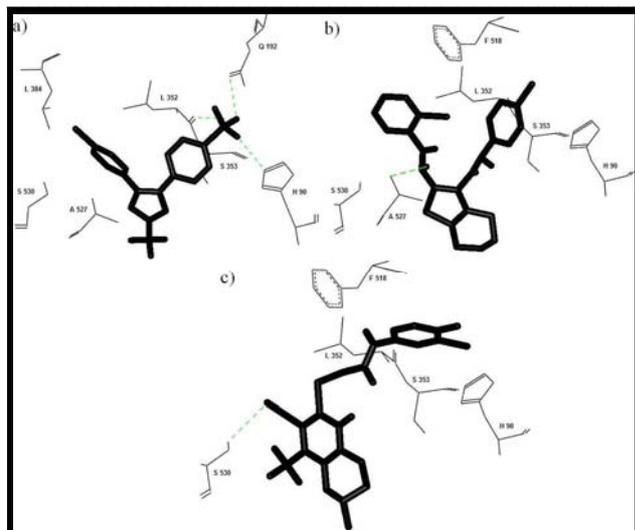
three pocket regions and represented similar binding modes as observed with SC-558. But the NH group of ligand formed one hydrogen bond interaction with backbone NH of Ala527. Apart from this, scoring system in eHITS generated 26 interactions between protein and ligand. The scoring in eHITS is a statistically derived empirical scoring function and is based on interaction of 23 distinct surface point [13] types between receptor and ligand. In other words, the interaction between ligand surface points and receptor surface points determine the score of each fragment and on summation delivers the final score for the entire ligand.

COX-2 co-crystallized SC-558 ligand resulted in a e-HITS dock score of -6.559 kcal/mol. Therefore, any molecule from the dataset which shows a score lower than -6.559 kcal per mol would be regarded as much interest. Table 1 (see supplementary material) gives two such ligands having lower binding energies than SC-558. In other words, these two zinc ligands would act as inhibitors against COX-2 protein and such screening analysis forms the basis when millions of ligands are available in compound libraries such as the ZINC database.

Interaction of all the 3 compounds (ZINC00663976 and ZINC02062094 and co-crystallized SC-558 ligand molecules) with the active site residues are presented in Figure 1. The specific interactions between active site residues and ligand atoms are represented as dotted lines.

### Conclusion:

Virtual screening methods are routinely and extensively used to reduced cost and time of drug discovery. It has been clearly demonstrated that the approach utilized in this study is successful in finding two novel COX-2 inhibitors from the ZINC database. ZINC02062094, in particular, showed high binding affinity with a e-HITS dock score of -7.242 kcal/mol against 6COX. The docked pose resembles similar orientation as observed with SC-558 ligand. The ligand was docked deeply within the binding pocket region forming interactions with Leu352, Phe518, Met522, Val523, Ala527 and Ser353, respectively. Therefore, this study states the importance of small molecule libraries and their use to enhance drug discovery process prior synthesis. This approach to screen novel compounds as COX-2 inhibitors from ZINC database depends on various parameters such as Lipinski's rule of 5, pharmacophoric groups attached on the ligand, size of the dataset and compound libraries among others. Further, work can be extended to study the receptor-ligand interactions experimentally and evaluation of their biological activity would help in designing compounds based on virtual screening techniques.



**Figure 1:** Interaction of all the 3 compounds with the active site residues of 6COX a) SC-558 ligand displaying 3 H-bond interactions b) ZINC02062094 ligand showing the H-bond interaction between NH group of ligand and backbone NH of Ala527 c) ZINC00663976 ligand showing one H-bond interaction between cyano group of ligand and hydroxyl group of Ser530.

#### Acknowledgement:

The authors thank Dr. V. Jaya Lakshmi, Head of the Department of Microbiology, Sir CRR College for Women, Eluru for helpful comments and suggestions.

#### References:

- [01] G. S. Bleumink, *et al.*, *Drugs*, 63: 52 (2003) [PMID: 12656651]
- [02] T. Grosser, *et al.*, *J Clin Invest.*, 116: 4 (2006) [PMID: 16395396]
- [03] D. A. Kujubu, *et al.*, *J. Biol. Chem.*, 266: 12866 (1991) [PMID: 1712772]
- [04] D. L. Simmons, *Thromb Res.*, 110: 265 (2003) [PMID: 14592545]
- [05] S. Kalgutkar, *et al.*, *Bioorg. Med. Chem. Lett.*, 12: 521 (2002) [PMID: 11844663]
- [06] R. N. DuBois, *et al.*, *FASEB J*, 12: 1063 (1998) [PMID: 9737710]
- [07] I. K. Khanna, *et al.*, *J. Med. Chem.*, 43: 3168 (2000) [PMID: 10956225]
- [08] K. Chegaev, *et al.*, *J Med Chem.*, 50: 1449 (2007) [PMID: 17335184]
- [09] W. J. Tsai, *et al.*, *Bioorg. Med. Chem. Lett.*, 16: 4440 (2006) [PMID: 16814546]
- [10] M. M. Wolfe, *et al.*, *N. Engl. J. Med.*, 340: 1888 (1999) [PMID: 10369853]
- [11] G. de Gaetano, *et al.*, *Trends Pharmacol. Sci.*, 24: 245 (2003) [PMID: 12767724]
- [12] J. J. Irwin & B. K. Shoichet, *J Chem Inf Model.*, 45: 177 (2005) [PMID: 15667143]
- [13] Z. Zsoldos, *et al.*, *J Mol Graph Model*, 26: 198 (2007) [PMID: 16860582]
- [14] M. Jalaie and V. Shanmugasundaram, *Mini Rev. Med. Chem.*, 6: 1159 (2006) [PMID: 17073716]
- [15] G. L. Warren, *et al.*, *J. Med. Chem.*, 49: 5912 (2006) [PMID: 17004707]
- [16] Rummey, *et al.*, *Bioorg. Med. Chem. Lett.*, 16: 1405 (2006) [PMID: 16321524]
- [17] E. Kellenberger, *et al.*, *J. Med. Chem.*, 50: 1294 (2007) [PMID: 17311371]
- [18] Waszkowycz *Drug Discov Today*, 13: 219 (2008) [PMID: 18342797]
- [19] <http://bioserv.rpbs.jussieu.fr/Help/FAFDugs.html>
- [20] A. Lipinski, *J Pharmacol Toxicol Methods.*, 44: 235 (2000) [PMID: 11274893]

Edited by P. Kanguane

Citation: Sudha *et al.*, *Bioinformatics* 2(8): 325-329 (2008)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

### Supplementary material

S. No	Compound	eHiTS score (kcal/mol)	vdWaals score	H-bond score	No of H-bonds	Steric clash	Lipophilicity score	No. of interactions	Interacting residues
1	SC-558 ligand	-6.559	-11.452	-12.431	3	---	-51.712	24	Val349, Leu359 Val523, Ala527 Leu352, Ser530 Leu384, Val116 Ser353, Gln192 His90
2	ZINC00663976	-7.135	-10.247	-4.143	1	---	-58.235	26	Leu384, Met522 Leu352, Val523 Phe518, Ser353 Ala527, Leu531 Val116, Met113 Val349
3	ZINC02062094	-7.242	-11.846	-3.742	1	---	-64.971	26	Gly526, Met522 Leu352, Ser530 Val349, Leu359 Val116, Met113 Ala527, Ser353 Phe518, Val523

**Table 1:** Interaction parameters of COX-2 with the predicted ligand ZINC00663976 and ZINC02062094 and co-crystallized SC-558 ligand molecules.