

Comparative QSAR analysis of cyclo-oxygenase2 inhibiting drugs

Arumugam Mohanapriya^{1*} & Dayalan Achuthan²

¹Bioinformatics Division, School of Biosciences and Technology, VIT University; ²Department of General Surgery, Stanley Medical College and Hospital, Chennai; Mohanapriya Arumugam - Email: mohanapriyaa@vit.ac.in; Phone: +91 9843111639;

*Corresponding author

Received March 01, 2012; Accepted April 16, 2012; Published April 30, 2012

Abstract:

Cyclo-oxygenase 2 (COX2) inhibiting drugs were subjected to comparative quantitative structure activity relationship (QSAR) analysis with an attempt to derive and to understand the relationship between the biological activity and molecular descriptors by multiple regression analysis. The different drugs that inhibit cyclo-oxygenase 2 enzyme were compared instead of subjecting one drug and its derivatives to QSAR analysis. The study was conducted to look for the common structural features between the drugs which confer to a good biological activity. Based on the regression analysis the following descriptors were finalized as the components fitting best in the regression equations: Ss, SCBO, RBN, nN, SIC0, IC1, and H-055. These descriptors belong to constitution (Ss, SCBO, RBN, nN), information indices (SIC0, IC1) and atom centered fragments (H-055) category. Based on these descriptors QSAR models were generated and evaluated for best structure-activity correlation. The model generated from constitution and information indices descriptors corresponds to the essential structural features of the drugs and are found to have significant correlation with COX2 inhibiting activity. This study shall help in rational drug design and synthesis of new selective cyclo-oxygenase 2 inhibitors with predetermined affinity and activity.

Keywords: COX2, QSAR, biological activity, descriptors.

Background:

Quantitative structure -activity relationship (QSAR), is a type of analysis where some measures of chemical properties are correlated with biological activity to derive a mathematical illustration of the underlying structure activity relationship (SAR) [1]. Descriptors are numerical representations of specific molecular features. Such features can range from very simple ones such as the number of carbons to more complex and abstract features such as graph invariants of the molecular graph. Quantitative structure-activity relationship (QSAR) studies are unquestionably of great importance in modern chemistry and biochemistry. To get an insight into the structure-activity relationship we need molecular descriptors that can effectively characterize molecular size, molecular branching

or the variations in molecular shapes, and can influence the structure and its activities [2, 3]. A number of prospective studies show that COX enzyme gets increased in inflamed tissue and that COX is stimulated by interleukin-1 (IL-1) [4]. Cyclo-oxygenase (COX) converts arachidonic acid derived from cell membranes to prostaglandins, which have important signaling and housekeeping functions, particularly in platelets, the gastrointestinal tract, lungs, and kidneys. The two COX isoforms are the constitutive form (COX1) and the inducible form (COX2). Pharmacological inhibition of COX2 can provide relief from the symptoms of inflammation and pain [5]. COX2 inhibitors are a new class of non-steroidal anti-inflammatory drugs (NSAIDs). Because they selectively block the COX2 enzyme, blocking this enzyme obstructs the production of the chemical

messengers (prostaglandins) that cause the pain and swelling of arthritis inflammation. The COX2 inhibitor flurbiprofen is shown at its active site with COX2 (**Figure 1**). Selectivity for COX2 provides the anti-inflammatory action and also reduces the risk of peptic ulceration. COX2 inhibitors have recently been implicated in cancer prevention and the slowing of atherosclerosis [6, 7]. Dataset containing 15 COX 2 inhibiting drugs were formed and subjected to descriptor determination. At first a variety of molecular descriptors were calculated and 7 were finalized by multiple linear regression. The selected 7 molecular descriptors were found to confer the known biological activities of 15 COX2 inhibitors. Moreover, a regression model was hypothesized based on the results, which would be helpful for structural optimization of COX2 inhibitors.

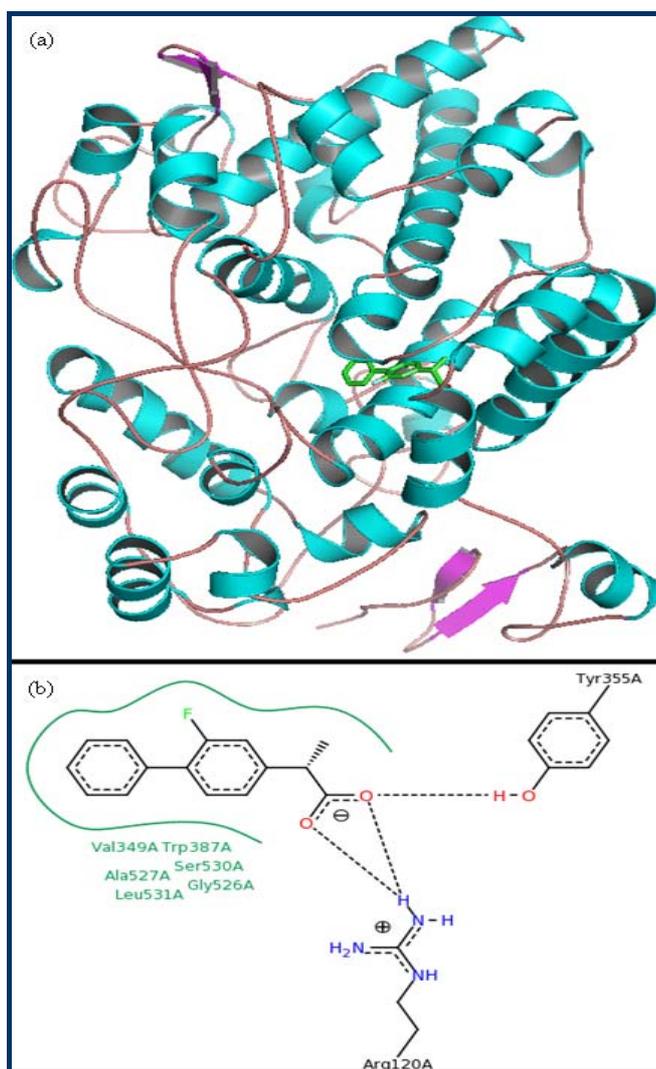


Figure 1: (a) Cyclo-oxygenase 2 complexes with inhibitor flurbiprofen. Selective inhibition of cyclo-oxygenase 2 enzyme by the inhibitor flurbiprofen [PDB: 3pgh]. The cyclo-oxygenase 2 enzyme is shown as cartoon and the inhibitor as the stick model in green color at its active site. (b) Interaction between the cyclo-oxygenase and flurbiprofen. Black dashed line - hydrogen bonds, salt bridges and metal interactions; green solid lines-hydrophobic interactions. (Image generated using PoseView software)

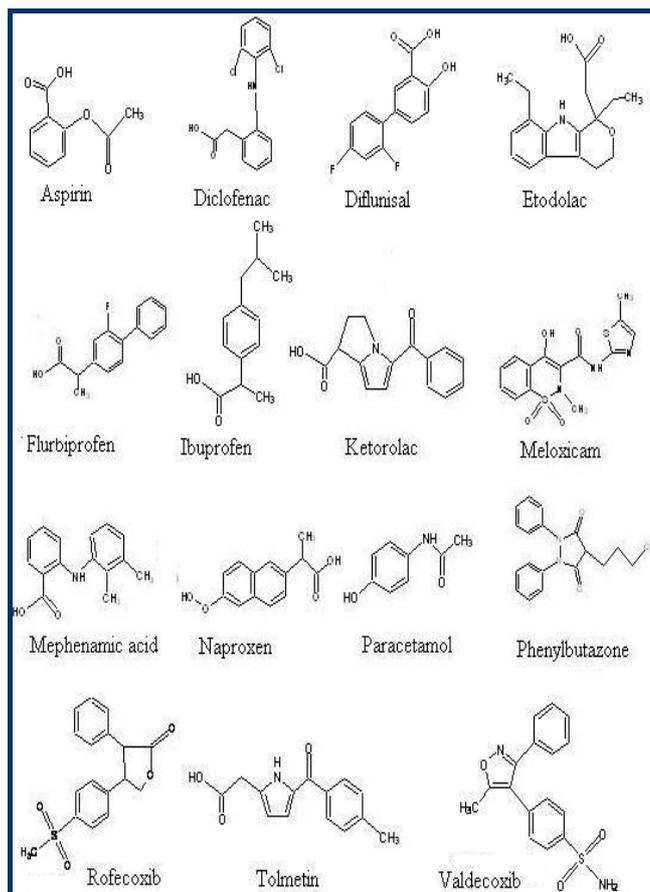


Figure 2: The drug dataset (15 drugs) taken for the study. The structures of the drugs were taken from Pubchem and drug bank databases. The structures were then drawn and optimized using ACD labs 3D optimizer.

One of the first ever a QSAR model is that of Richardson (1869) where the narcotic effect of a series of alcohols was correlated with molecular weight. [8]. Bazoui and colleagues (2002) derived QSAR for 103 analogues of 1-(2-hydroxyethoxy) methyl-6-(phenylthio) thymine (HEPT), a potent inhibitor of the HIV-1 reverse transcriptase, by means of multiple linear regression (MLR) and artificial neural network (ANN) techniques. The results showed that the anti-HIV activity of HEPT derivatives was strongly dependent on hydrophobic character and also steric factors of substituent [9]. Guha and Jurs (2004) developed Quantitative Structure-Activity Relationship (QSAR) models to predict the biological activity of 179 artemisinin analogues. Multiple linear regression and computational neural network models are developed to link the structures to their reported biological activity [10]. Jain and Agrawal (2006) applied step-wise multiple regression analysis and reported lipophilicity and topological distance indices are important for COX2 inhibition [11]. Khan and colleagues (2006) studied quantitative structure activity relationship analysis on a series of N-alkyl imidazole analogues using combination of various thermodynamic electronic and spatial descriptors. The study revealed that the electronic property, i.e., dipole moment contributed positively, and spatial descriptor (principal moment of inertia at Y axis) contributed negatively to the biological activity [12]. Livia and colleagues (2007) conducted comparative molecular

field analyses (CoMFA) on a large set of estrogen receptor (ER) modulators. Results indicate that steric, electrostatic, hydrophobic (lipophilic) and hydrogen bond donor substituents play a significant role in COX2 inhibitory activity and selectivity of the compounds [13].

Methodology:

Molecular structure and optimization

The COX2 inhibiting drugs were collected from several public sources like drug bank and from literature (Figure 2) [14]. The structures of the drugs were obtained from PubChem. ACD-Chemsketch was used to draw the inhibitors (drugs) structure. The structures were then optimized using 3D optimizer of ACD Labs Software (Figure 2) and the structures were saved in the pdb format.

on its selective inhibition of cyclo-oxygenase 2 enzymes in human.

QSAR descriptors

The optimized structures of the 15 drugs have been used for calculating molecular descriptors using E-Dragon [15]. The descriptor types chosen were constitutional descriptors, information indices, functional group counts, topological descriptors, atom centered fragments and molecular properties. This analysis resulted with 142 individual descriptor values. This total descriptor values were shortlisted by regression analysis. The normalized values have then been used to generate QSAR models by the means of multiple linear regression analysis. The QSAR model was generated using Build QSAR module [16].

Discussion:

Multiple regression analysis

Each of the 15 drugs was analyzed for the molecular descriptors using E-dragon and the results were recorded. MLR is a method used for modeling linear relationship between a dependent variable Y (IC50) and independent variable X (2D descriptors). Out of 142 descriptors for the series of 15 COX2 inhibitors, 7 were found to be significant by multiple regression analysis using the SPSS statistical package Table 2 (see supplementary material).

Set of Final Descriptors

All the suitable descriptors were analyzed and finally the seven descriptors were found to suite the fifteen COX2 inhibitors. The E-Dragon software was used to calculate the descriptors for each molecule. The 7 Descriptors found significant were Ss, SCBO, RBN, nN, SIC0, IC1, and H-055 Table 2 (see supplementary material). The QSAR analysis was done using the Build QSAR module. These descriptors belong to constitution, information indices and atom centered fragments category. These values were used for further calculation and model generation. The graph has been plotted between the experimental and the predicted IC50 values.

QSAR Model Generation

Using the above-mentioned 7 descriptors the QSAR models were generated using Build QSAR model generator. The predicted activity value is the predicted IC50 value generated by the Build QSAR module, residual activity is the difference between the observed and predicted IC50 values. When the drug bioactivity data and descriptor data which fall under constitution, information indices and atom centered fragments category were subjected to multiple regression analysis the following significant equations were obtained.

$$n=015; r=0.966; r^2(R) = 0.933; s=2.314; F=18.711 - \text{Model 1}$$

This equation is obtained by analyzing the constitution (Ss, SCBO, RBN, nN) and information indices (SIC0, IC1) descriptors, where n is number of molecules under analysis, r is the correlation coefficient, $r^2(R)$ is the squared correlation coefficient, s is the standard deviation, F is the F statistical value. The cross validated squared correlation coefficient, Q2 is 0.495 and standard deviation of sum of square of

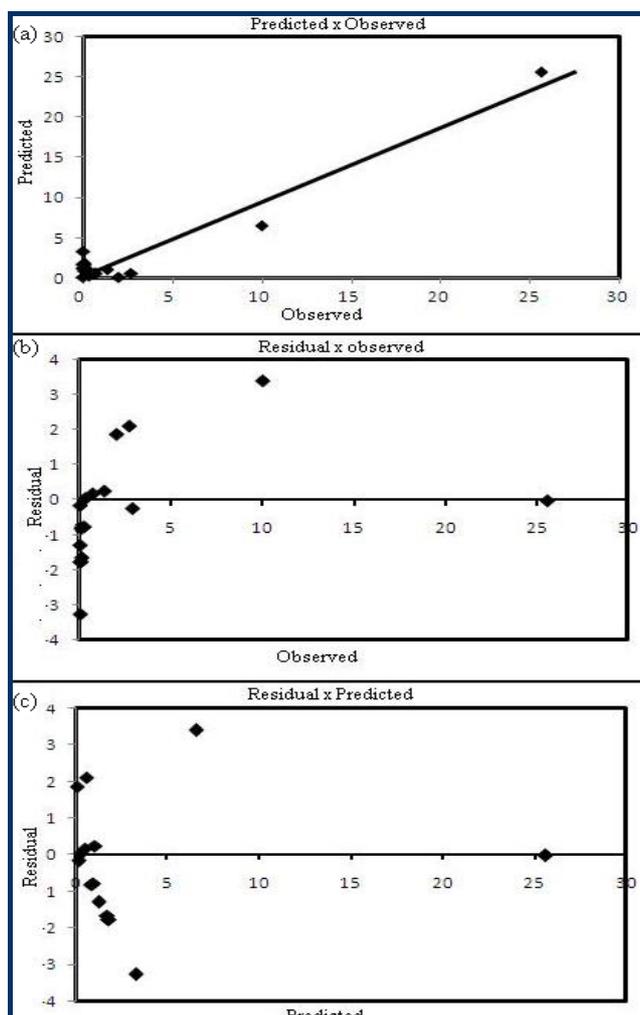


Figure 3: (a) The graph plotted between predicted and observed biological activity ($r=0.933$). The predicted activity shows linear relationship with observed activity because fit of the data to the regression line is good. (b) The graph plotted between observed and the residual biological activity. (c) Graph plotted between predicted and the residual biological activity.

The biological activity

The biological activity of the drugs i.e the IC50 values, were obtained from the Pubchem database Table 1 (see supplementary material). IC50 values were retrieved based

difference between predicted and observed values, SPRESS is 6.377.

$n=015$; $r=0.650$; $r^2(R) = 0.423$ $s=6.427$; $F=1.319$ – Model 2

This equation is obtained by analyzing the constitution (Ss, SCBO, RBN, nN) and atom centered fragments (H-055) descriptors with $Q2=1.354$ and $SPRESS=12.979$

$n=015$; $r=0.953$; $r^2(R) = 0.908$ $s=2.317$; $F=36.332$ – Model 3

This equation is obtained by analyzing the information indices (SIC0, IC1) and atom centered fragments (H-055) descriptors with $Q2=0.661$ and $SPRESS=4.454$

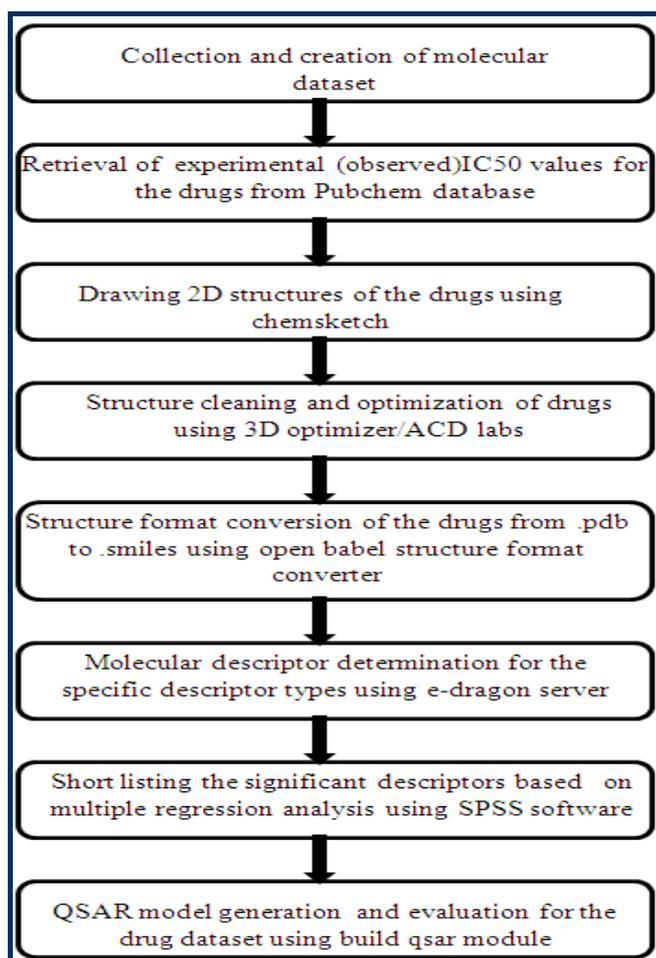


Figure 4: Steps followed for QSAR analysis for COX2 inhibitors. Dataset contains a total of 15 cyclo-oxygenase2 inhibiting drugs. Totally seven descriptors were subjected to QSAR model generation and validation.

Evaluating the model

Among the above three models, model 1 has produced high statistical quality equation ($n=015$; $r=0.966$; $r^2(R) = 0.933$; $s=2.314$; $F=18.711$) which was obtained by pooling constitution (Ss, SCBO, RBN, nN) and information indices (SIC0, IC1) descriptors. It is seen that model 1 has highest r and R value i.e. 0.966 and 0.933 respectively and lowest standard deviation ($s = 2.314$) when compared to other two models. The $Q2$ value is low than the other two, however, based on the correlation coefficient and the standard deviation model 1 can be considered as the best one among

the generated models. It is seen that the average observed activity in model 1 is 2.876; average predicted activity is 3.123467; and the residual activity value does not deviate beyond ± 1 and its average being 0.248 **Table 3** (see **supplementary material**). Hence the chosen descriptors best fit for the COX2 inhibition. The auto correlation between the descriptors studied was analyzed **Table 4** (see **supplementary material**). The highest correlation is observed between the descriptors RBN and SIC0 with the value 0.277.

Graphical Analysis

The graphical analysis has been performed and the graph is shown in **(Figure 3)**. The graph has been plotted between the predicted and observed IC_{50} values **(Figure 3a)**. The predicted activity shows linear relationship with observed activity because fit of the data to the regression line is good. The higher the value for r^2 , less likely the relationship is due to chance. This QSAR investigation indicates that the descriptors namely Ss, SCBO, RBN, nN, SIC0, IC1 for the set of cyclo-oxygenase inhibitors studied were found to have a great deal to positively contribute to biological activity. The graph is plotted for observed activity versus residual **(Figure 3b)** and predicted activity versus residual **(Figure 3c)**. The finalized descriptors were found to be the members of constitutional and information indices. The workflow for the present comparative QSAR study is shown in **(Figure 4)**. Thus, this work sets sights on to identify the associated molecular properties of COX2 inhibiting drugs and exploit it to optimize COX2 inhibiting activity.

Conclusion:

COX2 plays a significant role in the development of various symptoms and conditions, including fever, inflammation, and pain. The overall result of the conducted comparative QSAR analysis explicates the exploration of common structural descriptors of the COX2 inhibiting drugs in explaining its bioactivity against the cyclo-oxygenase enzyme. These descriptors indicate simple structural features like constitution property, information indices and atom centered fragments which confer to better bioactivity. This provides insight into the physicochemical nature of the activity under consideration. The results of this study might pave the way towards rationalizing the design and discovery of novel COX2 inhibitor.

Acknowledgement:

Authors wish to thank Dr. Prof.V. Umashankar, Reader and Head, Centre for Bioinformatics, Sankar Nethralaya, Chennai for his support and AM is grateful to VIT University for providing laboratory facility.

References:

- [1] Testa B & Kier LB, *Med Res Rev.* 1991 **11**: 35 [PMID: 1994153]
- [2] Karol Kozak *et al. Adaptive and Natural Computing Algorithms.* 2007 **4432**: 327
- [3] Abhilash T *et al. Archive for Organic Chemistry.* 2004 **14**: 36
- [4] Raz A *et al. J Biol Chem.* 1988 **263**: 3022 [PMID: 3125173].

- [5] Saqib A & Karigar CS, *The Internet Journal of Pharmacology*. 2009 **7**: 1
- [6] Thealdi Mitchell & Elsabé van der Merwe, SA *Pharmaceutical Journal*. 2010
- [7] Ho L & Pasinetti GM, *Drugs Today (Barc)*. 2001 **37**: 181 [PMID:12783090]
- [8] Richardson BJ, *Med Times and Gazzette*. 1869 **2**: 703
- [9] Bazoui H *et al. SAR and QSAR in Environmental Research*. 2002 **13**: 567
- [10] Guha R & Jurs PC, *J Chem Inf Comput Sci*. 2004 **44**: 1440 [PMID:15272852]
- [11] Jain HK & Agrawal RK, *J Mol Des*. 2006 **5**: 224
- [12] Khan N *et al. Indian J Pharm Sci*. 2006 **68**: 341
- [13] Salum Lde B *et al. J Mol Graph Model*. 2007 **26**: 434 [PMID: 17349808]
- [14] Fe'lix MA *et al. Pharmacoepidemiology and drug safety*. 2006 **15**: 861
- [15] Tetko IV *et al. J Comput Aided Mol Des*. 2005 **19**: 453 [PMID: 16231203].
- [16] DeOliveira DB & Gaudio AC, *Quant Struct -Activ Relat*. 2001 **19**: 599

Edited by P Kanguane

Citation: Mohanapriya & Achuthan, *Bioinformation* 8(8): 353-358 (2012)

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:

Table 1: Drugs and their IC50 values.

Sl.No	Drug	IC50 μ M
1	Aspirin	0.30
2	Diclofenac	0.01
3	Diflunisal	2.70
4	Etodolac	2.00
5	Flurbiprofen	0.01
6	Ibuprofen	0.10
7	Ketorolac	0.23
8	Meloxicam	0.70
9	Mephenamic acid	1.33
10	Naproxen	0.06
11	Paracetamol	25.58
12	Phenylbutazone	10.00
13	Rofecoxib	0.02
14	Tolmetin	0.09
15	Valdecoxib	0.01

The IC50 were retrieved from the PubChem database. This activity is selective to cyclo-oxygenase 2 enzyme in human cells.

Table 2: Final Descriptors selected by MLR. Ha - The superscript represents the formal oxidation number. The formal oxidation number of a carbon atom equals the sum of the conventional bond orders with electronegative atoms; the C--N bond order in pyridine may be considered as 2 while we have one such bond and 1.5 when we have two such bonds; the C..X bond order in pyrrole or furan may be considered as 1.

S.No	Name	Description	Type of Descriptors
1	Ss	Sum of Kier - Hall electropological states	Constitutional
2	SCBO	Sum of conventional bond orders	Constitutional
3	RBN	Number of rotatable bonds	Constitutional
4	nN	Number of nitrogen atoms	Constitutional
5	SIC0	Structural Information Content (Neighborhood Symmetry Of 0-Order)	Information indices
6	IC1	Information Content Index (Neighborhood Symmetry Of 1-Order)	Information indices
7	H-055	H ^a attached to C ⁰ (sp ³) with 4X attached to next C	Atom Centered Fragments

Table 3: The observed and the predicted biological activity of the drugs.

Compounds	Observed activity	Predicted activity	Residual
Aspirin	0.300	0.26	0.040
Diclofenac	0.005	1.305	-1.300
Diflunisal	2.700	0.59	2.110
Etodolac	2.000	0.128	1.872
Flurbiprofen	0.010	1.791	-1.781
Ibuprofen	0.100	1.756	-1.656
Ketorolac	0.230	1.002	-0.772
Meloxicam	0.700	0.528	0.172
Mephenamic acid	1.330	1.081	0.249
Naproxen	0.060	1.836	-1.776
Paracetamol	25.580	25.607	-0.027
Phenylbutazone	10.000	6.6	3.400
Rofecoxib	0.020	3.293	-3.273
Tolmetin	0.090	0.901	-0.811
Valdecoxib	0.010	0.174	-0.164
Average	2.876	3.124	-0.248

The average observed, predicted and residual activity being 2.876, 3.124 and -0.248 respectively.

Table 4: The auto correlation matrix between the descriptor values

	Ss	SCBO	RBN	nN	SIC0	IC1
Ss	1	0.264	0.128	0.171	0.003	0.087
SCBO		1	0.033	0.031	0.014	0
RBN			1	0.002	0.277	0.188
nN				1	0.013	0.09
SIC0					1	0.638
IC1						1

The highest correlation is observed between the descriptors RBN (number of rotatable bonds) and SIC0 (structural information content).