

Self-Organizing Molecular Field Analysis on a New Series of COX-2 Selective Inhibitors: 1,5-Diarylimidazoles

Shunlai Li * and Yan Zheng

College of Science, Beijing University of Chemical technology, Beijing 100029, People's Republic of China

* Author to whom correspondence should be addressed; E-mail: lishunlai@yahoo.com.cn

Received: 12 June 2006 / Accepted: 21 July 2006 / Published: 26 July 2006

Abstract: Self-organizing molecular field analysis (SOMFA), a simple three-dimensional quantitative structure–activity relationship (3D-QSAR) method is used to study the correlation between the molecular properties and the anti-inflammatory biological activities of a new series of 1,5-Diarylimidazoles that act as selective COX-2 inhibitors. The statistical results, cross-validated r_{CV}^2 (0.507) and non cross-validated r^2 (0.546), show a satisfied predictive ability.

Keywords: NSAID design, selective COX-2 inhibitors, 3D-QSAR, SOMFA.

1. Introduction

Nonsteroidal anti-inflammatory drugs (NSAIDs) [1] display their anti-inflammatory actions primarily through the inhibition of cyclooxygenase (COX), which catalyzes converts arachidonic acid to prostaglandin (PG)₂ [2] and subsequently to a number of other prostaglandins which are potent mediators of inflammation. Cyclooxygenase exists in at least two different isoforms, namely, COX-1 and COX-2[3]. COX-1 is a constitutive enzyme [4], and COX-2 is an inducible isoform that leads to inflammation [5]. All classical NSAIDs, such as aspirin, ibuprofen, and indomethacin, can inhibit both COX-1 and COX-2, but bind more tightly to COX-1[6]. Selective COX-2 inhibitors are proving to have the same anti-inflammatory, anti-pyretic, and analgesic activities as do nonselective NSAID inhibitors, but with few or none of their gastrointestinal side-effects [7]. Nowadays, the search for novel and selective COX-2 inhibitors is increased due to their therapeutical potential in the treatment of inflammation. Recently, a new series of 1,5-diarylimidazoles compounds has been reported to

selectively inhibit COX-2[8].

The self-organizing molecular field analysis (SOMFA)[9] is a simple 3D-QSAR technique, which has been developed just recently by Robinson et al. The method has similarities to both comparative molecular field analysis (CoMFA)[10] and molecular similarity studies. Like CoMFA, a grid-based approach is used; however, no probe interaction energies need to be evaluated. Like the similarity methods it is the intrinsic molecular properties, such as the molecular shape and electrostatic potential, which are used to develop the QSAR models.

A SOMFA model could suggest a method of tackling the all-important alignment, which all 3D-QSAR methods have faced. The inherent simplicity of this method allows the possibility of aligning the training compounds as an integral part of the model derivation process and of aligning prediction compounds to optimize their predicted activities.

The purpose of this paper is to describe the application of self-organizing molecular field analysis, SOMFA, on this set of 1,5-diarylimidazoles, a novel class of selective COX-2 inhibitors. Thus our main objective is to provide some useful information by SOMFA analysis and design new specific inhibitors of COX-2 in the hope that these molecules may be further explored as powerful non-ulcerogenic anti-inflammatory agents.

2. Computational Methods

2.1. Data sets and biological activities

Twenty-nine 1,5-diarylimidazoles compounds are divided into two sets. The training set of 22 molecules with structures and their anti-inflammatory activities expressed as $-\log(\text{IC}_{50})$ are shown in Table 1. The predictive power of the models is evaluated using a test set of 7 molecules whose structures and activities are also shown in Table 1. Two sets of 29 molecules are selected in order to find some molecular descriptors and to elucidate convenient models for the predictive discrimination between these various activities. All compounds and their activities are processed as enantiomers in order to decrease the molecular alignment error derived from different configuration and increase the correlation of SOMFA models.

2.2. Molecular modeling and docking

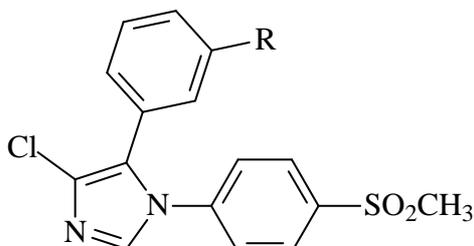
The three-dimensional structures of the 1,5-Diarylimidazoles are constructed with the CAChe worksystem pro evaluation[11], running on an AMD Athone XP 2400+ Processor/Microsoft Windows XP platform.

Unless otherwise indicated, parameters are default. Full geometry optimization are performed by PM5[11] semi-empirical method in the CAChe software. The final active conformations search are performed by dock into ActiveSite method which also in the CAChe software. The PDB entry of Cyclooxygenase receptors used in docking experiments is 6COX. For a example, the docked structure of ZA18 versus SC-558 in the active site of COX-2 are shown in Figure 1.

According to the docked structures or the alignment of the three cycle in the optimized geometries of 1,5-Diarylimidazoles, these compounds are then performed SOMFA analysis. The superposition of

1,5-Diarylimidazoles structures after docking are shown in Figure 2, the superposition of 1,5-Diarylimidazoles according to the optimized structures and alignment of the three cycle in 1,5-Diarylimidazoles are also shown in Figure 3. Using VEGA software [12], the final overlaid geometries are converted into CSSR file format, the only file format which SOMFA2 program can accept to process a SOMFA analysis.

Table 1. Chemical structures and corresponding biological activities of 1,5-Diarylimidazoles



compd ^a	R	-Log(IC ₅₀)	R	compd ^a	-Log(IC ₅₀)
ZA1	4-F	7.8539	ZA16	4-OMe-2-F	7.8239
ZA2	3-F	7.1871	ZA17	3,4-diCl	8.2218
ZA3	2-F	7.5528	ZA18	4-OMe-3-F	8.3979
ZA4	H	6.9101	ZA19	4-Me-3-F	8.2218
ZA5	4-Cl	7.7447	ZA20	4-OMe-3-Me	7.8861
ZA6	4-Me	7.7959	ZA21	4-Me-3-OMe	7.8239
ZA7	4-OMe	7.9586	ZA22	4-Cl-3-Me	7.9586
ZA8	4-OEt	8.3979	ZA23	4-NMe-3-Cl	7.5686
ZA9	4-Pr	7.0000	ZA24	4-OMe-3-Cl	8.0969
ZA10	4-Pr'	7.4089	ZA25	4-OEt-3-Cl	8.1549
ZA11	4-SMe	6.9586	ZA26	4-OEt-3-F	7.602
ZA12	4-SEt	7.2757	ZA27	4-F-3-OMe	7.7959
ZA13	4-NH ₂	6.7282	ZA28	4-OMe-3,5-diCl	8.1549
ZA14	4-NEt ₂	7.0969	ZA29	3,5-diF	7.1249
ZA15	2,4-diF	8.1549			

^aCompound 1-22: training set; compounds 23-29: test set

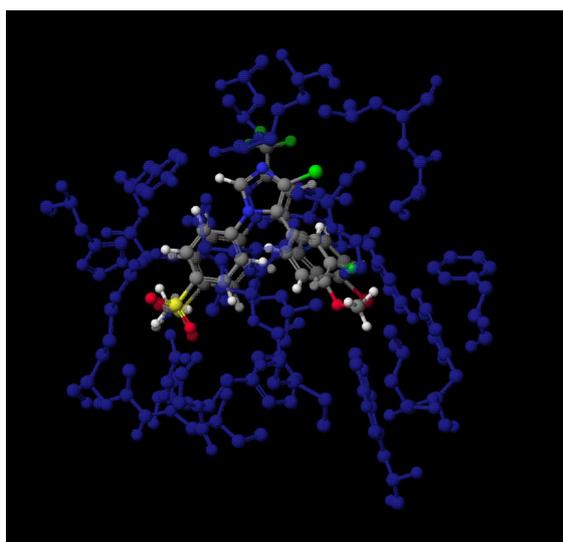


Figure 1. The docked structure of ZA18 versus SC-558 in the active site of COX-2

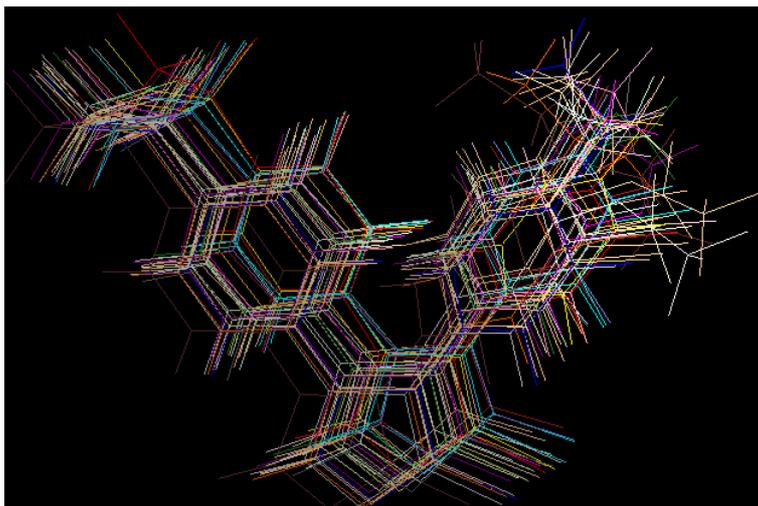


Figure 2. Superposition of 1,5-Diarylimidazoles structures after docking

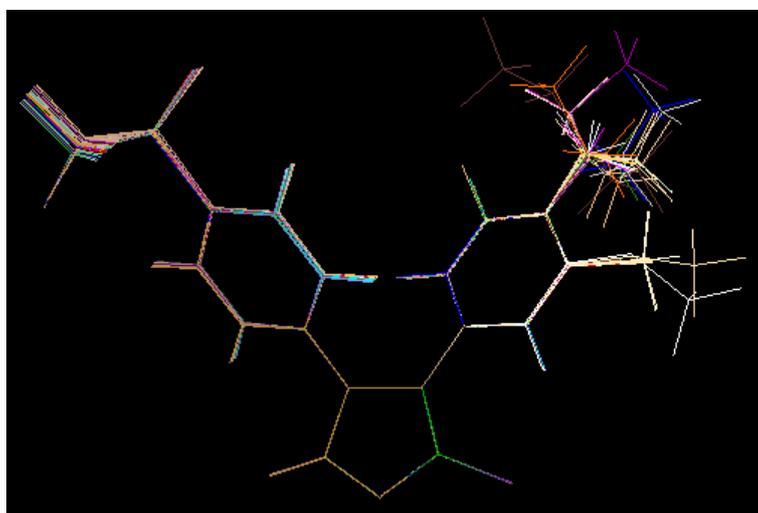


Figure 3. Superposition of 1,5-Diarylimidazoles according to the optimized structures and alignment of the three cycle in 1,5-Diarylimidazoles

2.3. SOMFA 3D-QSAR models

In the SOMFA study a 40x40x40 Å grid originating at (-20,-20,-20) with a resolution of 0.5 and 1 Å respectively, is generated around the aligned compounds. Table 2 reports 12 models using different alignment, charge and resolution of grid under exploration.

Table 2. Encoding 12 models for the training set used for the SOMFA investigations

Model No	Alignment	Charge	Resolution of grid (Å)
1	a	AM1	0.5
2	a	AM1	1
3	a	MNDO	0.5
4	a	MNDO	1
5	a	PM3	0.5
6	a	PM3	1
7	b	AM1	0.5
8	b	AM1	1
9	b	MNDO	0.5
10	b	MNDO	1
11	b	PM3	0.5
12	b	PM3	1

a: Alignment of the three cycle in optimized 1,5-Diarylimidazoles structures; b: original docked structures without alignment.

For all of the studies, shape and electrostatic potential are generated. To sum up the predictive power of these two properties into one final model, we combine their individual predictions using a weighted average of the shape and electrostatic potential based QSAR, using a mixing coefficient (c_1) as illustrated in eq. 1[9].

$$\text{Activity} = c_1 \text{Activity}_{\text{shape}} + (1 - c_1) \text{Activity}_{\text{ESP}} \quad (1)$$

Clearly, multiproperty predictions could have been obtained through multiple linear regression. Using eq 1 instead gives greater insight into the resultant model by allowing the study of the variation in predictive power with different values of c_1 .

With the highest value of r^2 , the SOMFA models then are derived by the partial least squares (PLS), implemented in NoSA [13] with cross-validation.

The predictive ability of the model is quantitated in terms of r_{CV}^2 which is defined in eq. 2.

$$r_{\text{CV}}^2 = (\text{SD} - \text{PRESS}) / \text{SD} \text{ where } \text{PRESS} = \sigma(Y_{\text{pred}} - Y_{\text{actual}}) \text{ and } \text{SD} = \sigma(Y_{\text{actual}} - Y_{\text{mean}}) \quad (2)$$

SD is the sum of squares of derivations of the observed values from their meaning and PRESS is the prediction error sum of squares. The final models are constructed by a conventional regression analysis with the optimum value of mixing coefficient (c_1) equal to that yielding the highest r^2 and r_{CV}^2 value according to eq 2.

Table 3. Statistics of the various SOMFA models

Model	c_1	r^2	s	F	r_{cv}^2
1	0.685	0.478	0.351	24.711	0.455
2	0.679	0.464	0.356	23.342	0.439
3	0.765	0.458	0.358	22.858	0.434
4	0.752	0.446	0.361	21.736	0.421
5	0.633	0.503	0.342	27.316	0.473
6	0.658	0.479	0.351	24.835	0.451
7	0.500	0.525	0.335	29.820	0.486
8	0.514	0.532	0.333	30.680	0.500
9	0.512	0.546	0.327	32.515	0.507
10	0.506	0.544	0.328	32.149	0.500
11	0.644	0.498	0.344	26.826	0.458
12	0.645	0.497	0.344	26.716	0.464

r^2 , Non cross-validated correlation coefficient; s, standard error of estimate; F, F-test value; c_1 , mixing coefficient of SOMFA model; r_{cv}^2 , Cross validated correlation coefficient.

Table 4. Observed and predicted activities of 22 compounds in the training set

Compd	Observed	Predicted	Residual ^a	Compd	Observed	Predicted	Residual ^a
1	7.85	7.46	0.40	12	7.28	7.43	-0.16
2	7.19	7.57	-0.38	13	6.73	7.42	-0.70
3	7.55	7.70	-0.15	14	7.10	7.12	-0.02
4	6.91	7.11	-0.20	15	8.15	7.94	0.21
5	7.75	8.00	-0.25	16	7.82	7.70	0.13
6	7.80	7.57	0.23	17	8.22	7.79	0.43
7	7.96	7.69	0.27	18	8.40	8.15	0.25
8	8.40	7.76	0.64	19	8.22	8.21	0.01
9	7.00	7.18	-0.18	20	7.89	7.73	0.16
10	7.41	7.32	0.09	21	7.82	7.58	0.24
11	6.96	7.41	-0.45	22	7.96	8.02	-0.06

^a Residual=Observed-predicted.

Table 5. Observed and predicted activities of 7 compounds in the training set

Compd	Observed	Predicted	Residual ^a	Compd	Observed	Predicted	Residual ^a
1	7.59	8.02	-0.43	5	7.80	7.92	-0.13
2	8.10	8.11	-0.01	6	8.15	7.66	0.50
3	8.15	7.82	0.34	7	7.13	7.66	-0.53
4	7.60	7.83	-0.23				

^a Residual=Observed-predicted.

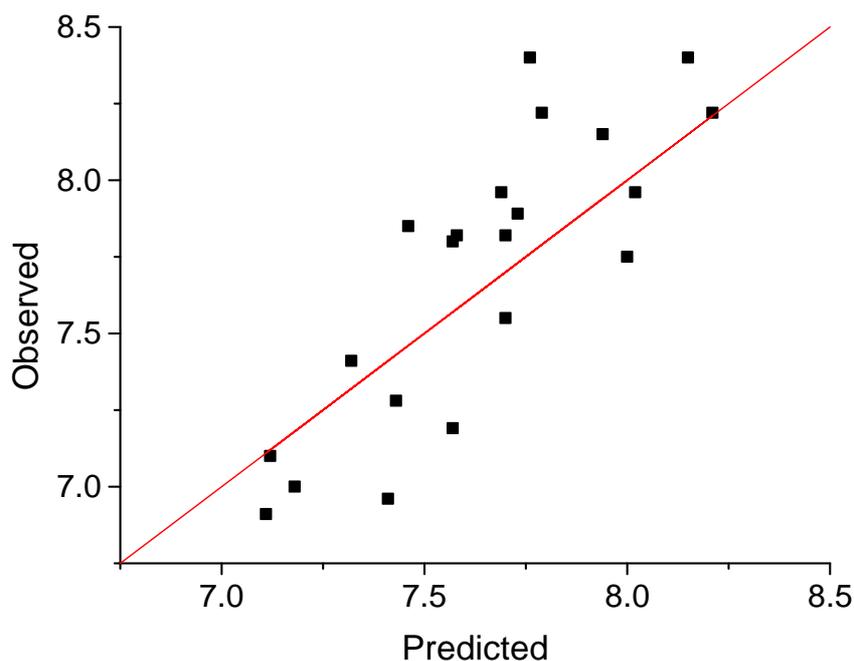


Figure 4. Observed versus predicted activities in the training set.

3. Results and Discussion

SOMFA, a novel 3D-QSAR methodology, is employed for the analysis with the training set composed of 29 various compounds, from which biological activities are known. Statistical results of 12 SOMFA models are summarized in Table 3.

A cross-validated value r_{cv}^2 which is obtained as a result of PLS analysis serves as a quantitative measure of the predictability of the SOMFA model. From the table we find that the result is less sensitive to resolution of grid and quantum chemistry charge but the model overlaid using original docked structures shows higher r_{cv}^2 values than using the model of alignment of the optimized 1,5-Diarylimidazoles structures.

Among the twelve models tested, the best predictive power is the ninth models from cross-validated. Good cross-validated correlation coefficient r_{cv}^2 values (0.507), moderate non cross-validated correlation coefficient r^2 values (0.546) proves a good conventional statistical correlation which have been obtained, and we also find that the resultant SOMFA model have a satisfied predictive ability.

During the SOMFA investigation, grid spacings of 1 and 0.5 Å were investigated. The 1 Å grid spacing produces a good correlation equal to 0.5 Å grid. This has improved marginally with the 0.5 Å spacing uses for the results presented here. Further increases in resolution has produced further small increases in model quality but not enough to warrant the extra computational time.

The observed and predicted activities of the training set are reported in Table 4. Figure 4 shows a satisfied linear correlation and moderate difference between observed and predicted values of molecules in the training set.

It's well known that the best way to validate a 3D-QSAR model is to predict biological activities for

some compounds of test set. The SOMFA analysis of the test set composed of 7 compounds is reported in Table 5. Most of compounds in test set show good correlation between observed and predicted values.

SOMFA calculation for both shape and electrostatic potentials are performed, then combined to get an optimal coefficient $c_1=0.512$ according to eq 1. The master grid maps derived from the best model is used to display the contribution of electrostatic potential and shape molecular field. The master grid maps give a direct visual indication of which parts of the compounds differentiate the activities of compounds in the training set under study. The master grid also offers an interpretation as to how to design and synthesis some novel compounds with much higher activities. The visualization of the electrostatic potential master grid and shape master grid of the best SOMFA model is showed in Figures 5 and 6, respectively, with compound 18 as the reference.

Each master grid map is colored in two different colors for favorable and unfavorable effects. In other words, the electrostatic features are red (more positive charge increases activity, or more negative charge decreases activity) and blue (more negative charge increases activity, or more positive charge decreases activity), and the shape feature are red (more steric bulk increases activity) and blue (more steric bulk decreases activity), respectively.

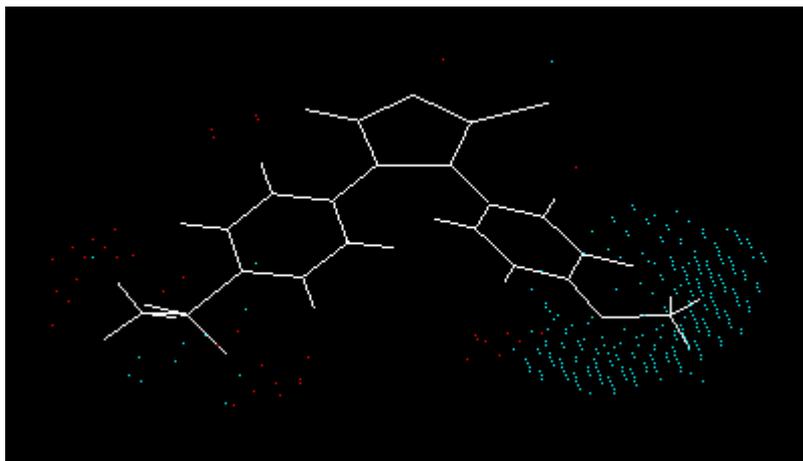


Figure 5. The electrostatic potential master grid with compound 18. Red represents areas where positive potential is favorable, or negative charge is unfavorable. Blue represents areas where negative potential is favorable, or positive charge is unfavorable.

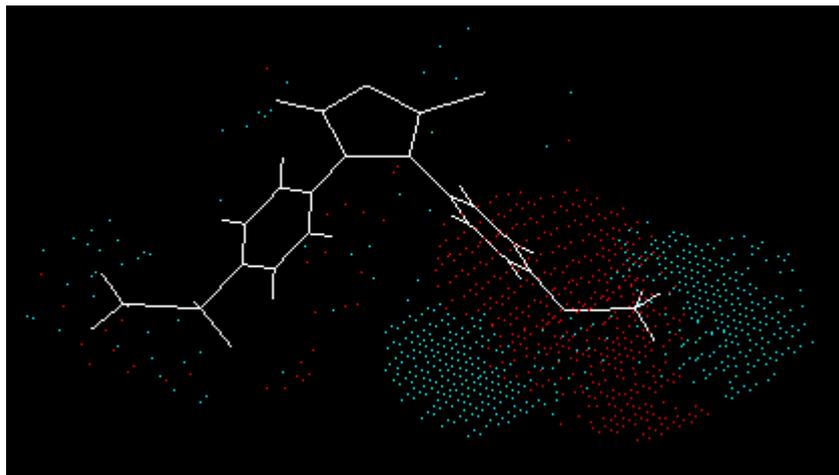


Figure 6. The shape master grid with compound 18. Red represents areas of favorable steric interaction. Blue represents areas of unfavorable steric interaction.

SOMFA analysis result indicates the electrostatic contribution is of a slightly low importance ($c_1=0.512$). The SOMAF electrostatic potential for the analysis is presented as master grid in Figure 5. In this map of important features, we find a high density of blue points around the substituent R_3 and R_4 at the second phenyl ring, which means some electronegative groups are favorable. Meanwhile, in the map of shape master grid, we can find a high density of red points around the substituent R_3 and R_4 at the second phenyl ring, which means a favorable steric interaction; simultaneously, we also find a high density of blue points outside the red regions which around substituent R_3 and R_4 at the second phenyl ring, where an unfavorable steric interaction may be expected to enhance activities. Generally, the medium-sized substituent R_3 and R_4 at the second phenyl increase the activity.

All analysis of SOMFA model may provide some useful information in the design of new 1,5-diarylimidazoles antagonists.

4. Conclusion

We have developed predictive SOMFA 3D-QSAR models for 1,5-diarylimidazoles as anti-inflammatory agents. The master grid obtained for the various SOMFA models electrostatic potential contributions can be mapped back onto structural features relating to the trends in activities of the molecules. On the basis of the spatial arrangement of the various electrostatic potential contributions, novel molecules are being designed with improved activity.

Acknowledgements

We are grateful for the help of Mr Mingyong Li, a PhD candidate from Department of Medicinal Chemistry, China Pharmaceutical University. We also gratefully acknowledge Dr Daniel Robinson and the Computational Chemistry Research Group (Oxford University, UK) for the use of the SOMFA software.

References and notes

1. Vane, J. R.; Botting, R. M. Mechanism of action of anti-inflammatory drugs. *Scand. J. Rheumatol. Suppl.* **1996**, 102, 9-21.
2. Dannhardt, G.; Kiefer, W. Cyclooxygenase inhibitors—current status and future prospects. *Eur. J. Med. Chem.* **2001**, 36(2), 109-126
3. Picot, D.; Loll, P. J.; Garavito, R. M. The X-ray crystal structure of the membrane protein prostaglandin H₂ synthase-1. *Nature* **1994**, 367, 243-249
4. Crofford, L. J. COX-1 and COX-2 Tissue Expression: Implications and Predictions. *J. Rheumatol.* **1997**, 24 (Suppl 49), 15-19
5. Marnett, L. J. Cyclooxygenase mechanisms. *Curr. Opin. Chem. Biol.* **2000**, 4(5), 545-552
6. Laneuville, O.; Breuer, D. K.; Dewitt, D. L.; Hla, T.; Funk, C. D.; Smith, W. L. Differential inhibition of human prostaglandin endoperoxide H synthases-1 and -2 by nonsteroidal anti-inflammatory drugs. *J. Pharmacol. Exp. Ther.* **1994**, 271(2), 927-934
7. Marnett, L. J.; Kalgutkar, A. S. Design of selective inhibitors of cyclooxygenase-2 as nonulcerogenic anti-inflammatory agents. *Curr. Opin. Chem. Biol.* **1998**, 2(4), 482-490
8. Carmen A.; Jose A.; Alberto F. et al. Synthesis and structure-activity relationship of a new series of COX-2 selective inhibitors: 1,5-diarylimidazoles. *J. Med. Chem.* **2003**, 46(11), 3463-3475
9. (a) Robinson, D. D.; Winn, P. J.; Lyne, P. D.; Richards, W. G. Self-Organizing Molecular Field Analysis: A Tool for Structure-Activity Studies. *J. Med. Chem.* **1999**, 42(4), 573-583
(b) SOMFA2 v2.0.0, a novel suite of programs which can carry out an advanced 3-D QSAR analysis of a series of molecules, can be downloaded from author's homepage: <http://bellatrix.pcl.ox.ac.uk>. (c) Robinson, D. D. SOMFA2 Instruction Manual; New Chemistry Laboratory: Oxford University, UK, 1999.
10. Cramer, R. D., III; Patterson, D. E.; Bunce, J.D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, 110(18), 5959-5967
11. CAChe 6.1.1. Evaluation; Beaverton, USA: CAChe Group, Fujitsu Limited, 2005
12. (a) Pedretti, A.; Villa, L.; Vistoli, G. VEGA: A Versatile Program to Convert, Handle and Visualize Molecular Structures on Windows-based PCs. *J. Mol. Graph. Model.* **2002**, 21(1), 47-49.
(b) VEGA ZZ Release 2.0.5.52, a versatile program to convert, handle and visualize molecular structure on Windows-based PCs; its executable and the source code can be free downloaded from: <http://www.ddl.unimi.it/vega/index2.htm>
13. NoSA V2.30 Build 2003.02.11, a statistics software with simple interface, which can be downloaded from its homepage: <http://nosa.diy.myrice.com/> as a shareware.