PRELIMINARY *IN SILICO* INVESTIGATION OF COX 2 SELECTIVE INHIBITORS[1]

Luminita Crisan, Liliana Pacureanu

Department of Computational Chemistry, Institute of Chemistry of Romanian Academy, Timişoara, Mihai Viteazul Avenue, 24, 300223 Timişoara, Romania
pacureanu@acad-icht.tm.edu.ro

Abstract
We report herein an attempt to generate QSAR models for a large number of structurally diverse compounds (1078 compounds) whose affinities for cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2) were experimentally determined. Initially, individual QSAR models for COX-1 (M1) and COX-2 (M2) for biological activity were developed. A selectivity QSAR model, M3 was then developed using as dependent variable Y the differences in pIC50 values between COX-1 and COX-2. The statistical results for all three models showed a satisfactory to good statistical parameters where the values for squared correlation coefficient (coefficient of determination) for the training set are: M1: 0.872, M2: 0.797 respectively M3: 0.739. The predicted values of affinity in the case of all three models selected M1, M2 and respectively M3, are very good 84.88%, 91.12%, 79.59% which lead to very small diffrences between observed and predicted biological activity/selectivity (less than 0.5 logarithimic units).

Keywords: COX-1, COX-2, QSAR

Introduction
Prostaglandin synthesis is promoted by the cyclooxygenase (COX), which during the '90 it was discovered that this enzyme exist in two isoforms: COX-1 which is responsible for protection of kidney and maintenance of gastric mucosal integrity function whereas COX-2 is implicated in the pathophysiological reactions including inflammation and pain[1]. Nonsteroidal anti-inflammatory drugs (NSAIDs) are the most frequently used drugs to treat pain, inflammation and cancers [2]. Cyclooxygenase inhibitors have been divided over time in (1) nonselective inhibitors, which shows a similar affinity for both COX-1 and COX-2, i.e. aspirin which block irreversibly the enzyme by acetylation of Ser530, (2) nonselective acting as competitive with arachidonic acid, i.e. diclofenac, indomethacin, ketoprofen, naproxen, ibuprofen, phenylbutazone, and meclofenamate [3], (3) nonselective, which inhibit preferentially COX-2, i.e. etodolac, meloxicam, nabumetone and nimesulide. Side effects resulting from the use of these inhibitors in the treatment of various diseases can lead to severe complications and increased costs of their relief [3]. In last years, it was demonstrated that COX-1 isoform, but not COX-2, is over-expressed in diverse human pathologies (ovarian, skin and colon cancer as well as in other cancer types) [4]. Thus, finding new drugs that are selective COX-2 and not inhibit COX-1 is a widely investigated topic which attracts a great interest nowadays. The present work reports the attempt to address the phenomenon of selectivity of COX-2 inhibitors by generating robust and predictive QSAR models for a large set of compounds in order to reliably predict novel selective COX-2 inhibitors.

[1] Dedicated to the 150th anniversary of the Romanian Academy

Material and Methods

Dataset selection and preparation

A number of 5900 inhibitors of cyclooxygenase-2 (COX-2; Assay ID: CHEMBL230) and 3763 inhibitors of cyclooxygenase-1 (COX-1; Assay ID: CHEMBL221) were downloaded from ChEMBL database [5]. 1145 Compounds were selected which have experimental activities for both enzymes COX-1 and COX-2. All these molecules were filtered using *BlockBuster software* available in FILTER module [6-13] from OpenEye package. Totally, 1078 compounds passed the filter criteria (HBA=0÷13, HBD=0÷9, MW=130÷781, RBN=0÷16, XLogP=-3.0÷6.85, 2dPSA=0÷205) and were used further in the QSAR analysis. The experimental activity for both protein is expressed in IC50 (nM) and for QSAR modeling it was transformed into negative logarithm of inhibitory concentrations, pIC50, and used later as a dependent variable. In order to obtain a selectivity QSAR model the Y values were obtained by taking the differences between pIC50 measured against COX-2 and COX-1. The LigPrep module [14] and ConfGen [15] from Schrödinger package [16] were used to generate the tautomers and ionization states in the pH range of 7.2±0.2, and conformational sampling.

QSAR models generation

QSAR modeling was performed with AutoQSAR module from Schrödinger package [15]. The QSAR models are constructed using different machine learning algorithms and multiple automatically generated random training and test sets. As function of the quality of QSAR models on the training and test sets, a score rank have been implemented. The predicitivity performance of AutoQSAR generated models is close or better than previously published results but show the advantage of reduced time, costs and expertise.

For each individual QSAR model generated (M1 for COX1, M2 for COX2, and M3 for selectivity, modeling the difference between pIC_{50}COX-2– pIC_{50}COX-1, as dependent variable) multiple QSAR models were generated using different methods. These models are ranked based on high values of the statistics parameters fortraining and test set and the best ten models are listed.

Results and discussion

Linear regression models were built using 809 (75%) compounds in training set and 269 compounds in test set (25%). The independent variable matrix X contain binary fingerprints descriptors (radial, linear, dendritic, molprint2D) and canvasMolDescriptors, whereas the Y matrix contained experimental values for COX-2 affinity for the model M1, experimental affinities for COX-1 in the case of model M2 and the selectivity values for COX2 for M3 model. For each case ten models were generate and the best one was selected (see Table1 and Figures1, 2, 3).

Table1. Statistical results for the best linear QSAR models

	M1	M2	M3
# *	5	5	5
R^2	0.872	0.797	0.739
SD	0.461	0.464	0.689
Q^2	0.675	0.738	0.729
RMSE	0.705	0.685	0.841

* # = The number of PLS factors used in the partial least squares regression model; R^2 = R-squared value for the coefficient of determination for the training set; SD = Standard deviation of the model; Q^2 = the squared value of the regression coefficient for the test set;RMSE = Root-mean-square error for the test set predictions.

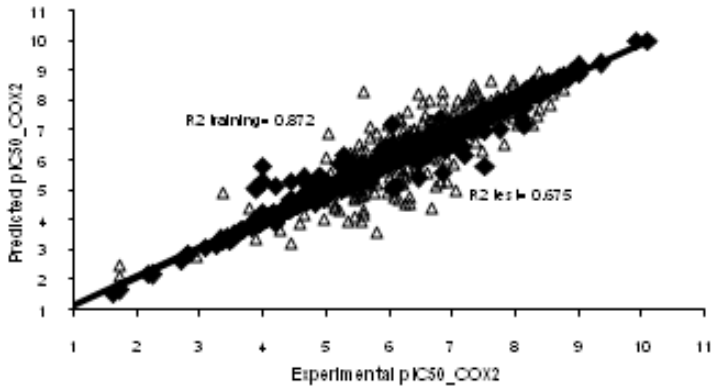

Figure 1. Plot of experimental versus predicted pIC_{50} values for model M1 (black squares – training setcompounds, white triangles – test compounds)

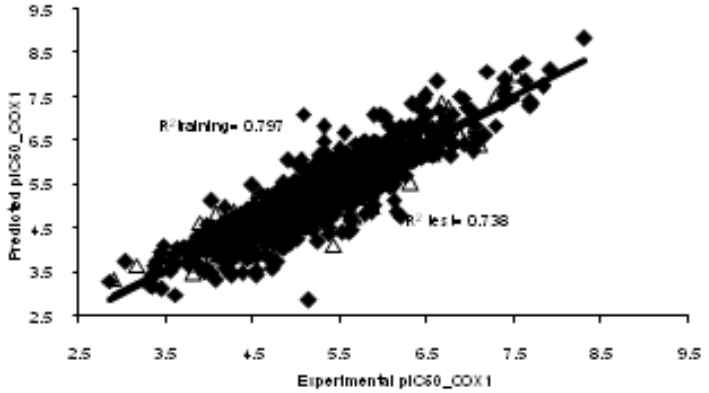

Figure 2. Plot of experimental versus predicted pIC_{50} values for model M2 (black squares – training compounds, white triangles – test compounds)

.

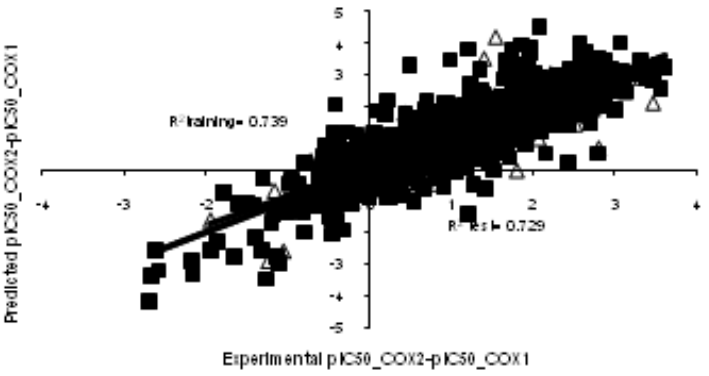

Figure 3. Plot of experimental versus predicted pIC_{50} values for model M3 (black squares – training compounds, white triangles – test compounds)

The best statistical results were obtained for the PLS model M1 while the best predictive power we observed in the case of the PLS models M2 and M3 models. The predicted values of affinity/selectivity of all QSAR models are very good: 84.88%, 91.12%, and 79.59% of inhibitors of the M1, M2 and respectively M3 models are predicted very well with a diffrence smaller than 0.5 logarithimic units between observed and predicted biologiocal activity/selectivity.

Conclusions

A set of structurally diverse 1078 compounds downloaded from ChEMBL have been engaged to develop quantitative structure - activity, respectively selectivity QSAR models to correlate the structural features with the affinity/selectivity against COX-2 enzyme. The statistical performances of the QSAR models are reasonably satisfactory. The correlation coefficient (R^2) is higher than 0.800 for model M1 and higher than 0.700 for models M2 and M3.

Acknowledgements

This project was financially supported by Project 1.2 of the Institute of Chemistry of the Romanian Academy. Access to the OpenEye Ltd. software are greatly acknowledged by the authors.

References

[1] W.L. Smith, D.L. Dewitt, Adv. Immunol. 62 (1996) 167.
[2] A.R. Chopade, F.J. Sayyad, N.S. Naikwade, 1st ed. LAP Lambert Academic Publishing; 2011.
[3] D. L. Simmons, R. M. Botting, T. Hla, Pharmacol. Rev. 56 (2004) 387.
[4] M.G. Perrone, P. Malerba, J. Uddin, P. Vitale, A. Panella, B.C. Crews, C.K. Daniel, K. Ghebreselasie, M. Nickels, M.N. Tantawy, H.C. Manning, L.J. Marnett, A. Scilimati, Eur. J. Med. Chem. 80 (2014) 562.
[5] A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, J. P. Overington, Nucleic Acids Res.42 (2014) D1083.
[6] R. Wang, F. Ying, L. Lai, J. Chem. Inf. Comput. Sci. 37 (1997) 615.
[7] P. Ertl, B. Rohde, P. Selzer, J. Med. Chem.43 (2000) 3714.
[8] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Adv Drug Deliv Rev23 (1997) 3.
[9] W. J. Egan, K. M. Merz, J. J. Baldwin, J. Med. Chem.43(2000) 3867.
[10] D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, K. D. Kipple,J. Med. Chem.45 (2002) 2615.
[11] Y. C. Martin, J. Med. Chem.48 (2005) 3164.
[12] T. Fawcett, Pattern Recognit. Lett. 27 (2006) 861.
[13] FILTER version 2.0.2: OpenEye Scientific Software, Santa Fe, NM. www.eyesopen.com.
[14] LigPrep, version 3.9, Schrödinger, LLC, New York, NY, 2016.
[15]ConfGen, version 3.7, Schrödinger, LLC, New York, NY, 2016.
[16] Schrödinger, LLC, New York, NY, 2016-3.