Acta Pharmaceutica Sciencia
51: 313-321 (2009)

Quantitative structure pharmacokinetic relationship (QSPkR) of cox-2 inhibitors using artificial neural network

Sangita Saini* and S.K. Singh

Department of Pharmaceutical Sciences, Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India-125 001.

Abstract

The purpose of this study was to develop Quantitative Structure Pharmacokinetic Relationship (QSPkR) for selected Cox-2 inhibitors. Artificial neural network (ANN) was used for data pruning (leave one out strategy) and prediction of pharmacokinetic and construct the QSPkR model from molecular descriptors that were generated using the TSAR, Codessa and Dragon computer programme. The final equation estimating Cl, Vd, $t^{1/2}$, AUC and Tmax, all show good correlation but Vd,Cl and $t^{1/2}$ shows significant correlation (r: 0.9-0.7). This investigation will help in rationalizing the design of new molecule and allow more rapid progression of potential drug candidates to the markets.

Key words: Quantitative structure pharmacokinetic relationship (QSPkR), artificial neural network, pharmacokinetics, molecular modeling, descriptors.

Introduction

Knowledge of the pharmacokinetic properties is of critical importance for the understanding of their mode of action and for the validation of activity and bioavailability. The recognition and exploitation of the relationships between chemical structure and biological activity have risen to the discipline known as Quantitative Structure Activity Relationship (QSAR). The hilarious work of Hansch and co-workers in 1964 has started the new era of activity prediction from structure. QSAR has been used to predict the optimum molecular structure and to provide a means of understanding their mechanism of action and an important part of modern drug discovery. Hansch used physicochemical properties and correlated with biological activity using regression analysis. The result of treatment was an equation which describes in quantitative manner the relationship between biological compounds with its chemical structure (Hansh et al. 1964, Glen et al. 1987, Hyde et al. 1988, Lewis et al. 1990). However much less study has been devoted to quantitative structure pharmacokinetic relationship (QSPkR) because of limited use of substituent parameters as pioneered by Hansch and co-workers. The present study is to compute the pharmacokinetic properties of limited numbers of compounds using parameters which can be calculated or predicted and which do not require the compound to be synthesized and experimental measurements made. We have attempted to investigate which molecular descriptors are important in the pharmacokinetic prediction using the well known data pruning method Artificial Neural Network (ANN).

Artificial neural network are computational systems implemented in software or hardware that

***Corresponding author:** sangi_132@yahoo.com

attempt to simulate the neurological processing abilities of biological system particularly in brain.

ANN contains mainly three layers (1) Input layer (2) Hidden layer (3) Output layer. The layers that receive input from environment is input layer. The output layer generates the dependent variable and the hidden layer inter connect the input and output layer (Nestrov et al. 1999). It is based on human brain that in which neuron collect signal from other neurons or dendrites (Dondeti et al. 2003). When population pharmacokinetic data is analyzed using NONMEM and ANN, ANN show less predictive error than NONMEM (Chow et al. 1997). Neural network are most suitable to model the behavior of complex pharmacokinetic and pharmacodynamic study (Veng-pedersen et al. 1993). In the present study an attempt is to make to determine whether such an approach is possible. Modern drug discovery programmes particularly involve the synthesis and sequential functional group alteration of compounds aimed at producing specific pharmacological effects with appropriate efficacy and pharmacokinetic properties in humans. An important part in the generation of novel drug understands of how they are absorbed, distributed, metabolized and excreted (ADME Studies). In contrast to the increasing component of design of drug discovery, there is still higher degree of empiricism used in optimizing pharmacokinetic properties of the compounds. Much less efforts have been expended in modelling the pharmacokinetic properties. There is therefore a need to understand how combination of molecular physicochemical quantitativiely influences the pharmacokinetic properties of the drug.

Human pharmacokinetic of a compound play a key role in determining the suitability of New Chemical Entity (NCE) in further drug development. Sceening of ADME in vitro and from animal modeling are time consuming and expensive (Norris et al. 2000). Even results can not accurately effect the pharmacokinetic of the compounds. Need of more accurate results and due to financial pressure, focus is on the research of molecular modeling (Walter et al. 1998). Modeling provides a means to describe and understand data.

Methods

Activity/Pharmacokinetic data

The value of Total clearance (Cl), Half life ($t^{1/2}$), Volume of distribution (Vd), Time to maximal plasma concentration (Tmax), Area under the curve (AUC) were obtained from the literature of following therapeutic Non Steroidal anti inflammatory Drugs (NSAIDS): Celecoxib,Valdecoxib, Rofecoxib, Etoricoxib, Lumirocoxib, Deracoxib (Paulson et al. 2000, Chan et al. 1999, Rodrigues et al. 2002) and Drugs under clinical trial: DUP-697 , SC-57666, SC-58451, NS-398, JTE-522, SC-236, SC-58125 (Futaki et al. 1993, Isakson et al. 1994).

Molecular Modeling

The structures were sketched in the ISIS draw 2.4 (Standlone Software). The compounds were modeled into 3-D with the help of HYPERCHEM (Hypercube Inc.) and Corina modules. 3-D structure helps in the computation of 3-D descriptors. Files were saved as. Hin file and. Mol file as extension file and used as input file for the other software.

Generation of Molecular descriptors

The extension files .Hin file and. Mol files were imported into the TSAR 3.3 (Tool for structure activity relationship, Oxford Molecular Limited). These files were used as input file to calculate the molecular descriptors. The descriptors which were calculated: Molecular weight (MW), molecular surface area

(SA), molar volume (MV) Inertia moment (IM), Ellipsoidal volume (EV), Kier Chi (χ), Kappa shape (κ), shape flexibility index (ψ) and topological index (G). The electronic descriptors Heat of Formation (HEFO), energy of lowest unoccupied molecular orbital (LUMO), energy of highest occupied molecular orbital (HOMO), dipole moment (DIPO) and Ionization Potential (IOPO) were calculated using the Vamp which is semi-emperical package of Mopac software with a time limit of around 30 min. Finally lipophillic parameters (log P) were estimated by Hyperchem and Chem Sw (MM Plus). For evaluation purposes, estimated log P values were compared with available experimental estimates of all the coxibs. All the descriptors were cross checked by computing in other software (Codessa and Dragon).

Generation of pharmacokinetic data

Artificial neural network programme used were the Statistica neural network 6.0 (Statsoft inc.) and TSAR. All network of three layered feed forward back propagation (Multilayer perceptron) type, containing bias neurons in each layer and single neurons in output layer. Sigmoidal transfer function was employed in all neurons and weight adjustment was performed according to generalized delta rule (Kamperman et al. 2001). Connection weights were initialized with random values. Models were constructed using the training set of compounds. The validation subsets were used to provide an indication of model performance. All generated descriptors were included in initial model. Redundant descriptors were then pruned and systems were retrained until optimum model was achieved. True predictive ability was assessed using testing subset of compounds.

Statistical analysis

The data was analyzed by using software Instat (Graphpad software Inc.) Stepwise regression analysis was used to determine the most significant descriptors. The regression coefficient was obtained by multi-linear regression analysis. For each regression analysis following descriptors information was obtained: Number of observations used in the analysis (n), Square of Correlation (r^2), Correlation coefficient, standard error of estimate (S) and Fisher's criterion (F).

Result

A summary of activity and pharmacokinetic parameter values were used to develope the QSAR/QSPkR modelsis shown in Table 1.

Table 1. QSAR/QSPkR modelsis

PK properties	Celecoxib (Cele)	Valdecoxib (Val)	Deracoxib (Dera)	Rofecoxib (Rofe)	Lumirocoxib (Lumi)	Etoricoxib (Etori)
Tmax(hr)	3	3	2	1.0	2	2
AUC(ng/hr/ml)	705	1479	4018	3286	1872	300
Vd(L)	400	86	91	108	90	1.5
Cl(L/hr)	27.7	6	141	57	5.6	5
T1/2(hr)	8	11	17	22	3	19

Pharmacokinetic data obtained from literature

There is change in molecular structure and results in significant variability in activity/PK characteristics. For example $t^{1/2}$ value ranges from 3 h to 22 h for the selected compounds. Similarly other pharmacokinetic properties also demonstrate considerable variability. The large range of values in each property make this data well suited for QSAR/QSPkR analysis. The calculated physicochemical descriptors, obtained by molecular modeling in TSAR, Dragon and Codessa are listed in Table 2, 3 and 4.

Data analysis

The correlation matrix analyzed the descriptors and those showing good correlation (0.9, 0.8, 0.7) taken first for training artificial neural network.

Training and validation

All the descriptors generated were used to train the ANN, after this pruning was implemented. It is done manually or automatically large number of descriptors were removed in each cycle and root mean square (RMS) was calculated each time. RMS is a measure of performance. The clearance and Vd show less RMS error so these were good candidates for further analysis (Table 5 and 6).

The output calculated each time and difference between actual and observed called error was propagated back to adjust the strengths of the connections or weights. Each iteration is called epochs. Number of maximum iteration in this study was 50,00,000. It is simply leave one out strategy. The final QSPkR equations along with their respective F probability, coefficient of regression (r), coefficient of determination (r^2). Values were reported in Table 7.

These relationships all appear to show significant correlation (R ranges from 0.8 to 0.7) and good predictive performance. The lipopillic parameter, shape index were advantageous since these were highly correlated with each other. Shape index parameters emerged as the primary determinant of the variability in area under curve, clearance, half life, Tmax and were also significant factors in volume of distribution (Vd). For Vd, lipophillic factors were common descriptors that appear to explain their broad range of values. Graphs showing observed versus predicted values for the pharmacokinetic parameters are shown in Figure 2 and all demonstrate good agreement about the line of identity.

Table 2. List of descriptors obtained from TSAR, Codessa and Dragon.

S.No	Name of molecule	Molecular mass(MW)	Molecular surface area(SA)	Molecular volume(mv)	Inertia moment (IM)	Ellipsoidal volume (EV)	Log p	Total lipole(TL)	Molar refractivity (MR)	Kier chi0 (atoms) Index(χ0)	Kier chi V0 (atoms) Index(χv0)	Kier chi1 (bonds) Index(χ1)
1.	Celecoxib	381.4	282.81	197.29	308.91	1374.7	4.2519	12.55	90.975	19.113	14.409	12.054
2.	Valdecoxib	314.38	246.68	167.67	174.42	739.04	2.8535	9.7086	84.707	15.742	12.736	10.465
3.	Rofecoxib	316.39	246.22	171.43	149.69	974.76	2.5511	5.7495	83.299	15.742	12.982	10.465
4.	Etoricoxib	358.86	251.99	186.45	270.78	829.26	3.5031	6.5922	95.104	17.320	14.893	11.342
5.	Deracoxib	318.32	260.06	180.41	208.9	812.06	4.4332	4.8322	79.566	16.397	12.210	11.097
6.	Lumiracoxib	293.74	207.78	146.46	115.77	311.35	3.9145	3.8567	75.619	14.698	11.600	9.4692
7.	Sc-57666	316.41	252.84	175.12	183.51	580.26	3.747	7.643	86.847	15.742	13.226	10.449
8.	Dup-697	411.31	262.27	186.18	304.70	961.85	4.425	10.792	96.541	16.613	15.473	10.842
9.	Sc-58451	342.45	280.76	194.05	272.96	566.14	4.006	7.2136	94.068	16.949	14.433	11.363
10.	Ns-398	314.39	241.19	175.44	232.81	341.58	2.0998	7.0673	77.914	15.458	12.624	9.8492
11.	Jte-522	320.44	264.88	185.26	178.59	755.89	2.5661	6.2711	84.750	15.742	13.643	10.449
12	Sc-236	401.81	279.76	198.86	411.22	590.13	4.3027	13.348	90.739	19.113	14.605	12.054
13.	Sc-558	446.26	280.96	201.56	591.94	636.07	4.5765	13.61	93.556	19.113	15.409	12.054
14.	Sc-58125	384.37	280.01	196.45	326.97	837.59	4.5495	12.501	87.747	19.113	14.210	12.054

Table 3. List of descriptors obtained by TSAR, Codessa and Dragon software.

S.No.	Kier chi v1(bonds)	Kier chi 2 (path) index	Kier chi v2(path)	Kier chi3 (cluster) index	Kappa1 index	Kappa 2 index	Kappa 3 index	KAlpha2 index	KAlpha 3 index	Randic topological index	Balban topological index	Wiener topological index	polarity
1.	9.1028	12.860	7.4174	3.9149	20.727	7.788	4.8994	6.5713	4.0296	12.054	1.6693	1654	4.83
2.	8.2873	10.172	6.5859	2.3266	16.844	6.8571	3.7351	5.7230	3.0232	10.465	1.5973	1057	3.92
3.	9.0272	10.172	7.6454	2.3266	16.844	6.5281	4.7456	5.6646	2.9863	10.654	1.6051	1061	9.97
4.	9.6886	11.342	8.1856	2.6932	18.781	7.7091	4.6012	6.8184	3.9833	11342	1.6243	1351	7.56
5.	6.8734	9.992	4.8518	1.5079	17.811	7.9213	3.9837	6.6048	3.1851	11.097	1.6732	1159	4.72
6.	6.4035	8.831	4.8737	1.6222	16.372	7.3199	4.4963	6.1343	3.6424	9.4692	1.8855	840	7.16
7.	9.2794	10.256	7.7934	2.4045	16.844	6.3451	3.9256	5.6879	3.1454	10.449	1.5476	1099	2.75
8.	10.441	10.902	9.8750	2.6932	17.811	7.0869	4.1588	6.2619	3.5948	10.842	1.5816	1219	8.19
9.	10.194	11.734	9.2934	3.1116	17.416	6.3017	3.3600	5.2808	2.7243	11.363	1.4053	1340	2.93
10.	8.8237	9.8418	7.080	2.5981	17.355	7.5130	5.6055	6.6087	4.850	9.8492	1.9677	936	4.91
11.	9.2269	10.280	7.5397	2.4045	16.844	6.3434	3.9256	6.1855	3.4743	10.449	1.5867	1067	7.40
12.	9.2004	12.860	7.5301	2.9149	21.772	7.7921	4.7896	6.7360	4.146	11.234	1.6693	1654	4.74
13.	9.6028	12.860	7.9948	3.9149	18.899	7.8976	3.5674	6.8446	4.2231	10.231	1.5674	1543	10.64
14.	9.3374	12.860	7.7766	1.9148	19.123	7.4536	4.9272	6.5543	4.0177	12.789	1.0970	1678	10.71

Table 4. List of electronic descriptors obtained by the Vamp modules.

S. No.	VAMP total energy	VAMP Heat of formation	VAMP Ionization potential	VAMP LUMO	VAMP Nuclear energy	VAMP HOMO	VAMP total dipole
1.	-5278.6	-99.116	9.8269	-1.2021	29265	-9.8269	8.1386
2.	-3807.2	19.46	9.6482	-0.98752	21577	-9.6482	3.6129
3.	-3871.6	-89.306	9.9511	-0.94579	22991	-9.9511	4.4254
4.	-4131.9	24.442	9.5112	-1.2078	24524	-9.5112	3.5664
5	-4542.3	-52.048	9.0572	-0.54826	23477	-9.0572	3.3352
6	-3806.4	-96.841	8.57	-0.38196	18870	-8.57	2.2307
7	-3857.6	-53.325	9.0215	-1.0314	21857	-9.0215	6.5206
8	-4208	-24.824	9.4145	-1.1199	22692	-9.4145	5.2196
9	-4140.1	-23.139	8.9932	-1.0054	24999	-8.9932	6.3731
10	-4119.7	-99.988	9.759	-1.3269	23131	-9.759	4.4045
11	-3893	-68.484	9.1631	-1.0814	24531	-9.1631	4.5666
12	-5482.7	-98.155	9.8233	-1.3871	29139	-9.833	3.1955
13	-5462.2	-86.009	9.8564	-1.4038	29071	-9.8564	3.1817
14	-5528.3	-128.25	9.8465	-1.4204	26754	-9.8465	2.761

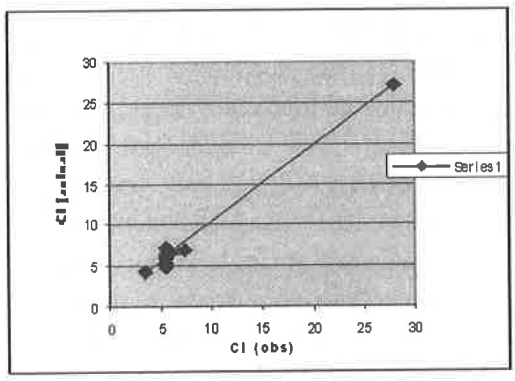
Table 5, 6. Showing the training and architecture of the data.

	Architecture Input-hidden-output	RMS error Root mean square
Clearance	52-2-1	0.36
	28-2-1	0.18
	13-1-1	0.44
	9-1-1	0.08
	5-1-1	0.04

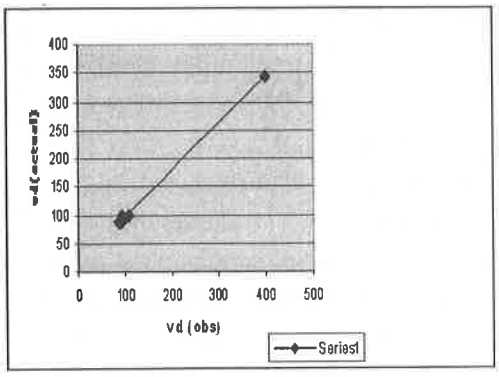
	Architecture Input-hidden-output	RMS error Root mean square
Vd	52-11-1	0.5
	34-7-1	0.4
	17-5-1	0.5
	6-4-1	0.08
	4-2-1	0.02

Table 7. Cox-2 inhibitors QSPkR (best correlated) equations.

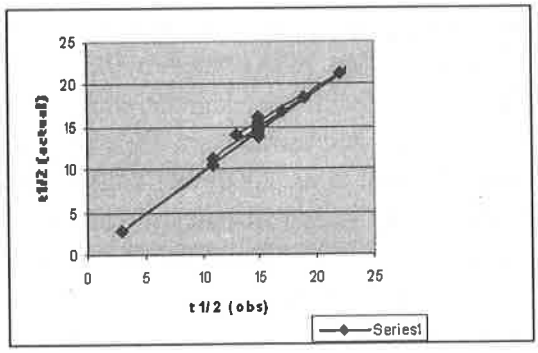
Property	Equations	R	R2
Clearance	Cl=170.22* κ1-1212.5* κ2+7301.1	0.921	0.77
Vd	Vd=0.005*HOF+0.07+0.36*IOP-1.67	0.852	0.65
$T^{1/2}$	$T^{1/2}$=0.014*EP+0.048*1+3.03LP-3.76	0.716	0.53



Clearance



Vd



$T^{1/2}$

Figure 1. Comparisons of actual vs observed pharmacokinetic parameters (best correlated) with the line of identity (dashed line).

Discussion

NSAIDS are important agent for the treatment of acute and chronic inflammatory disorder such as osteoarthritis and rheumatoid arthritis. The chronic use of these agents is often limited by some common side effects like gastrointestinal hemorrhage and ulceration (Silkari et al. 2001). Selective Cox-2 inhibitors are more advantageous than conventional NSAIDS (Wilkerson et al. 1995). Then Silkari et al. studied the QSAR analysis of cyclooxygenase-2 inhibitory activity. However, activity alone cannot predict whether a compound is suitable for further development. Subsequent pharmacokinetic studies may determine that most active or potent agent has an unfavorable *in vivo* time course of drug concentration or response. Thus, QSAR analysis may serve to guide further drug development, while our current understanding of primary determinants of cox-2 inhibitors pharmacokinetics.

In present study QSPkR models have been developed based on molecular properties obtained from calculations of molecular quantum mechanics. MLR are used to arrive at the primary property determinants with minimal number of independent variables or descriptors. In this manner, most molecular properties are identified, giving direction for structural optimization and

a basis for understanding fundamental of pharmacokinetic of the drugs. The equations finally drawn from this study are advantageous however, one must always be having knowledge of limitations of QSAR/QSPKR modeling (Hansch et al. 1964) and main is linear regression. First a small number of compounds were included. Artificial network has been used to avoid the problems. First only a small number of compounds were included and molecular descriptors were obtained previously used to train the network. Firstly the data was trained with available pharmacokinetic parameters and this is called trained data, when RMS become constant then stops training. After training of network, test data was used and calculated the values of remaining pharmacokinetic parameters. In this entire data was selected as single identity and a set was chosen randomly at each iteration. Small numbers of compounds were replaced and a new diversity value was computed. Final the QSPkR equations should be devoid of highly correlated descriptors to avoid chance correlations and to improve the significance of the resulting relationships. This limitation has been avoided for all parameters.

One of the most important considerations in establishing and interpreting QSAR/QSPkR models is the nature of activity/pharmacokinetic parameters. Although parameters may be carefully selected to represent the most typical values, issue related to inter and intra subject variability and nonlinear parameters can not be considered. For example complex pharmacokinetic behaviour arising from the plasma protein binding and metabolism. Celecoxib pharmacokinetic parameters can also be affected by gender (Paulson et al. 2000) and pathophysiological conditions, but in case of valdecoxib it is not affected by gender difference. The parameter values for this study were chosen from selected pharmacokinetic reviews that represent the most typical values in healthy subjects, as well as those resulting from low doses so that the complexities introduced by dose dependent pharmacokinetic phenomena could be kept minimum so artificial neural network has been used. ANN is valuable and robust tool for development and drug discovery. The utility of ANN in pharmaceutics was demonstrated by successful construction of a number of QSPkR. They have been shown to be fast and reliable method for the prediction of human pharmacokinetic parameters. ANNs have the potential to aid in drug discovery and development by providing a tool to complement existing screening techniques. It is not proposed that ANN will replace current in vitro and in vivo screen tools. Rather if they are appropriately incorporated into overall drug design and development process they provide considerable savings in resources and allow more rapid progression of potential drug candidates to the markets The results demonstrate that ANNs provide a valuable modeling tool that may be useful in drug discovery and development.

References

Chan, C.C., Boycc, S. and Ricndcau, D. (1999). Rofecoxib: a potent and orally active cyclooxygenase-2 inhibitors: pharmacological and biochemical profiles. *J. Pharmacol. and Experi. Thera.* 290: 551-560.

Chow, H.H., Tolle, K.M., Roe, D.J., Elsberry, V. and Chen, H.C. (1997). Application of neural networks to population pharmacokinetic data analysis. *J. Pharm .Sci.* 86: 840-845.

Dondeti, S., Kannan, K. and Manavalan, R. (2004). Insight into artificial neural networks and its implications for pharmacy- a tutorial review: part-1 *Indian J. Pharm. Educ.* 38: 123-129.

Futaki, N., Arai, I., Hamasaka, Y., Takahashi, S., Higuchi, S. and Otomo, S. (1993). Selective inhibition of NS-398 on prostanoid production in inflamed tissue in rat carrageenan-air-pouch inflammation. *J. Pharm. Pharmacol.* 45:753-755.

Glen, R.C. and Rose, V.S. (1987). A computer programme suite for the calculation, storage and manipulation of molecular property and activity descriptors. *J. Mol. Graph.* 5: 79-86.

Gregory, A.S., Isakson, C.P. and Anderson G. (2002). Immunosuppressive effects of administration of a cyclooxygenase-2 inhibitor and a 5-lipoxygenase inhibitor. *US Patent* 6376528.

Hansch, C, Fujita, T. (1964). Rho sigma-pi analysis. A method for correlation of biological activity and chemical structure. *J. Am Chem. Soc.* 86: 1616-1626.

Hyde, R.M. and Livingstone, D.J. (1988). Perspectives in QSAR: computer chemistry and pattern recognition. *J. Comput. Aid. Mol. Design.* 2:145-155.

Kamperman, G.J., Dommerhelt, F.J. (2001). A computational model to predict clatheration of molecule with cephadrinr *J. Chem. Soc.* 2: 981-987.

Lewis, D.F.V. (1990). MO-QSARs: a review of molecular orbital-generated quantitative structure-activity relationships. *Prog. Drug Metab.* 12: 205-255.

Nestrov, I.S., Hadjitodorov, S.T., Petrov, I. and Rowland, M. (1999). Emperical versus mechanistic modeling: Comparison of an artificial neural network to a mechanistically based model for quantitative structure pharmacokinetic relationship of homologous series of barbiturates. *AAPS Pharm. Sci.* 1: E17.

Norris, D.A., Leesman, G.D., Sinko, P.J. and Grass, G.M. (2000). Development of predictive pharmacokinetic stimulation models for drug discovery. *J.Control Release* 65: 55-62.

Paulson, S.K., Zhang, J.Y., Breau, A.P. (2000). Pharmacokinetics, tissue distribution, metabolism and excretion of celecoxib in rats.*Drug Meta. and Dispost.* 514-521.

Rodrigues, D.A. and Halpin, R.A. (2002). Absorption, Metabolism and Excretion of Etoricoxib-a potent and selective cyclooxygenase-2 inhibitors in healthy male volunteers. *Drug Meta. Dispost.* 31: 224-232.

Silkari, O., Dixit, A., Kohli, D.V., Chaturvedi, S.C. (2001). QSAR analysis of 4,5 diaryl pyrrole with cyclooxygenase-2 inhibitors activity. *Indian J.Pharm.Sci.* 518-521.

Veng-pedersen, P, Modi, N.B. (1993). Application of neural network to pharmacodynamics *J.Pharm.Sci* 82: 918-925.

Walters, W.P., Stahl, M.T. and Murcko, M.A. (1998). Virtual screening-an overview. *Drug Discovery Today* 3: 160-178.

Wilkerson, W.W., Copeland, R.A., Covington, M. and Trzaskos, J.M. (1995). Antiinflammatory 4,5-diarylpyrroles. II: Activity as a function of cyclooxygenase-2 inhibition. *J. Med. Chem.* 38: 3895-3901.

Received: 11.09.2008
Accepted: 24.10.2009