J. Indian Chem. Soc.,
Vol. 92, November 2015, pp. 1747-1753

Hansch analysis of dihydro-pyrazolyl-thiazolinone derivatives as potential COX-2 inhibitors

Amrita Dwivedi, Ajeet Singh, Akansha Tripathi and A. K. Srivastava*

QSAR Research Laboratory, Department of Chemistry, University of Allahabad, Allahabad-211 002, Uttar Pradesh, India

E-mail : draksrivastava1@rediffmail.com

Manuscript received online 02 April 2015, accepted 07 May 2015

Abstract : Quantitative structure activity relationship (QSAR) study has been performed on the dihydro-pyrazolyl-thiazolinone derivatives. This study is based on modeling the COX-2 inhibitors of dihydro-pyrazolyl-thiazolinone derivatives using topological and physicochemical parameters. It has been demonstrated that steric, electronic and topological parameters along with indicator variables are significantly correlated with activity. The multiple regression analysis revealed that the four-parametric model containing χ_{eq}, Mv, I_1, I_2 are correlating parameters and best for modeling the activity of the compounds under present study which was able to explain 89.7% of the variance in the data. The QSAR models were tested for their statistical significance and reliability by using leave one out (LOO) cross validation method.

Keywords : COX inhibitors, QSAR, dihydro-pyrazolyl-thiazolinone derivatives, physicochemical parameters.

Introduction

The discovery and development of new drug is an expensive and time consuming process. Healing effects are hazardous to health and to estimate the property of a drug, series of experimental and *in vivo* tests are required. For experimental test, usage of animal model is often subject to ethical considerations. Therefore alternative methods are being developed to reduce the requirement of animal testing. *In silico* methods are often implemented due to their lower cost and high purity. Quantitative structure activity relationship (QSAR) has been widely used in the field of drug design and now a day most of the drug evolved by this technique. In pharmaceutical research, QSAR has a particular interest in the preclinical stages of drug discovery to replace the tedious and costly experiments, to avoid the uses of large chemical databases and to select specific drug candidates[1]. Historically, the primary objective of QSAR was to understand which properties are important to control a specific biological activity of a series of compounds. However, the main use of these techniques now a day is the prediction of novel compounds on the basis of previously synthesized molecules[2].

Non-steroidal and anti-inflammatory drugs (NSAIDs) provides relief from symptoms of inflammation and pain by inhibiting the cyclooxgenases (COX) enzyme. Cyclooxgenases enzyme are present in two isoform COX-1 and COX-2. COX enzyme produced prostaglandins, prostacyclin, and thromboxanes by transformation of arachidonic acid. The primary effect of the non-steroidal anti-inflammatory drugs (NSAIDs) is to inhibit cyclooxygenase (COX or prostaglandin synthase [PGHS]), thereby impairing the ultimate transformation of arachidonic acid to prostaglandins, prostacyclin, and thromboxanes[3]. COX-2 selective inhibitor is a form of non-steroidal anti-inflammatory drug (NSAID) that directly targets COX-2, an enzyme responsible for inflammation and pain. Targeting selectivity for COX-2 reduces the risk of peptic ulceration, and is the main feature of celecoxib, rofecoxib and other members of this class of drug. Therefore, the development of selective inhibitors of compounds COX-2 is an important target in order to reduce adverse side effects.

The pyrazole nucleus is a common scaffold which are used in many selective COX-2 inhibitors due to their anti-tumor[4], anti-bacterial[5], anti-inflammatory[6] pharmacologi-

cal activities. On the other hand thiazolinone is an important class of heterocyclic compounds, which exhibit anti-inflammatory, anti-proliferative, anti-viral and anti-bacterial activity[7,8]. On that basis, we select dihydro-pyrazolyl-thiazolinone derivatives for QSAR analysis as part of continuing effort for designing of various COX-2 inhibitors. For that purpose we have taken the activity of various derivative of dihydro-pyrazolyl-thiazolinone as pIC_{50} from literature[9].

Theoretical background :

In the present work, steric parameters such as molecular weight (Mw), molecular volume (Mv), molar refractivity (Mr), parachor (Pc), index of refraction (IOR), surface tension (st), density (D), molecular connectivity (χ); topological parameters such as Balaban indices (J), Wiener index (W), mean Wiener index (W_A), Balaban centric index (BAC), information theoretic index (ID); electronic parameters such as equalized electronegativity (χ_{eq}), polarizability (Pz); and hydrophobic parameter such as partition coefficient (log P) have been calculated for all the derivative of dihydro-pyrazolyl-thiazolinone presented in the series. Out of above descriptors molecular weight (Mw), molecular volume (Mv), molar refractivity (Mr), parachor (Pc), surface tension (st), density (D), index of refraction (IOR), and polarizability (Pz) were calculated by ACD Lab Chem. Sketch version 10 Software[10] where as molecular connectivity (χ), Balaban indices (J), Wiener index (W), mean Wiener index (W_A), Balaban centric index (BAC), were evaluated by E-Dragon Software[11]. The statistical significance of the models was determined by examining the regression coefficient, the standard deviation, the number of variables, the cross validation leave-one out statistics and the proportion between the cases and variables in the equation[12]. We have used Hansch analysis for developing the models[13]. The multiple regression analysis was used to derive the correlation by using SPSS 7.5 version program.

Parameter used :

Equalized electro-negativity (χ_{eq}) :

The χ_{eq} was evaluated using Sanderson's principle of electro-negativity equalization. When two or more atoms initially different in electronegativity combine chemically to form a molecule, they adjust to have the same intermediate electro-negativity within the compound.

$$\chi_{eq} = N/\Sigma\ (V/\chi)$$

where, N is the total number of atoms in the species formula, V is the number of atoms of a particular element in the species formula, χ is the electro-negativity of that particular atom.

Molar volume (Mv) :

The molar volume is the volume occupied by one mole of a substance (chemical element or chemical compound) at a given temperature and pressure.

$$Mv = \frac{M}{\rho}$$

Information theoretic index (ID) :

It is a numerical descriptor of molecular structure and is sensitive to size, shape symmetry and heterogeneity of atomic environment in the molecules.

$$ID(G) = \sum_{i=1}^{\chi} n_i \frac{d_i}{d} \log 2 \frac{d}{d_i}$$

n_i = number of different sets.

Partition coefficient (log P) :

The octanol-water partition coefficient, P, is a measure of the differential solubility of a neutral substance between the immiscible liquids and thereby, a descriptor of hydrophobicity (or the lipophilicity) of a neutral substance. It is typically used in its logarithmic form, log P.

Wiener index (W_A) :

The Wiener index (W_A) is a widely used topological index. It is based on the vertex-distances of the respective molecular graph. The Wiener index (W_A) was proposed in 1947 by Wiener and is defined as the sum of overall bonds of the product of the number of vertices on each side of the bond.

Indicator variables :

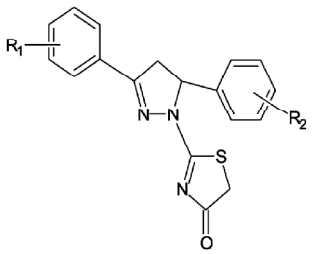
These are not QSAR parameters but are used to indicate the significance of any particular group or species at a particular substitution site in a given series of drugs.

Results and discussion

The success of QSAR studies mainly depends whether or not the molecular descriptors chosen are appropriate

to explain the biological activity. In an attempt to determine the role of structural features, which appear to influence the observed activity of reported compounds, QSAR studies were undertaken using linear free energy relationship (LFER) model proposed by Hansch and Fujita[14–19]. QSAR studies were performed on set of 20 derivatives of dihydro-pyrazolyl-thiazolinone derivatives. The biological activity data were correlated with different molecular descriptors such as equalized electronegativity (χ_{eq}), information theoretic index (ID), partition coefficient (log P), mean Wiener index (W_A), molecular volume (Mv) which are listed in Table 1. In this series COX-2 inhibitory activity has been expressed as IC_{50} values in micro molar units which represents the concentration of drug that inhibits 50% of COX-2 enzyme. The values were converted to negative logarithms (pIC_{50})[20] in order to reduce the skewness of the data set and obtain a linear relationship in the QSAR equations. The monoparametric models cannot be used for modeling the pIC_{50} because the quality of statistical data is not very good. Biparametric models were also discarded because of poor quality of statistical data. Hence, an attempt has been made to obtain multi-parametric models. In order to study the role of different substituents at different positions, indicator parameters I_1 for 4-H at R_2 position and I_2 for 2-CH_3 at R_2 position were introduced and are also listed in Table 1.

Table 1. Structural detail with biological activity and physicochemical data of dihydro-pyrazolyl-thiazolinone derivatives



Compd.	R_1	R_2	pIC_{50}	χ_{eq}	I_1	I_2	log P	ID	W_A	Mv
1	3,4-2Me	4-H	6.301	2.349	1	0	4.710	4.579	4.847	270.60
2	3,4-2Me	4-F	5.194	2.377	0	0	4.870	4.624	5.018	273.50
3	3,4-2Me	4-Cl	5.347	2.369	0	0	5.270	4.624	5.018	279.90
4	3,4-2Me	4-Br	5.745	2.364	0	0	5.540	4.624	5.018	283.20
5	3,4-2Me	4-Me	5.921	2.340	0	0	5.200	4.624	5.018	285.80
6	3,4-2Me	4-OMe	5.538	2.356	0	0	4.590	5.302	5.225	292.30
7	3,4-2Me	3-Cl	5.284	2.369	0	0	5.270	4.624	4.960	279.90
8	3,4-2Me	3-Me	5.569	2.340	0	0	5.200	4.624	4.960	285.80
9	3,4-2Me	2-Cl	5.143	2.369	0	0	5.270	4.624	4.902	258.80
10	3,4-2Me	2-Me	5.201	2.340	0	1	5.200	4.624	4.902	261.70
11	3,4-2Cl	4-H	5.569	2.419	1	0	4.860	4.579	4.847	268.10
12	3,4-2Cl	4-F	4.712	2.454	0	0	5.010	4.624	5.018	271.40
13	3,4-2Cl	4-Cl	4.931	2.444	0	0	5.410	4.624	5.018	274.00
14	3,4-2Cl	4-Br	5.119	2.437	0	0	5.680	4.624	5.018	280.50
15	3,4-2Cl	4-Me	5.328	2.403	0	0	5.340	4.624	5.018	268.10
16	3,4-2Cl	4-OMe	5.167	2.421	0	0	4.730	5.302	5.225	274.00
17	3,4-2Cl	3-Cl	4.793	2.444	0	0	5.410	4.624	4.960	268.10
18	3,4-2Cl	3-Me	5.143	2.403	0	0	5.340	4.624	4.960	274.00
19	3,4-2Cl	2-Cl	4.735	2.444	0	0	5.410	4.624	4.902	268.10
20	3,4-2Cl	2-Me	4.788	2.403	0	1	5.340	4.624	4.902	274.00

Multiple regression analysis of the data gave several regression models. The first step in obtaining a statistically significant model is to investigate whether or not any collinearity exists between the parameters used. This is achieved by obtaining correlation matrix; which are shown in Table 2.

The above model showed that parameter ID, log P, W_A and Mv have positive coefficient while χ_{eq} have negative coefficient, which indicate that more hydrophobic and bulkier group having more molar volume and less electro negativity should be preferred for future drug designing and enhances the activity of the drugs towards

Table 2. Correlation matrix demonstrating physicochemical parameters with indicator parameters

	pIC_{50}	χ_{eq}	I_1	I_2	Mv	log P	ID	W_A
pIC_{50}	1.000							
χ_{eq}	–0.723	1.000						
I_1	0.547	–0.071	1.000					
I_2	–0.234	–0.179	–0.111	1.000				
Mv	0.374	–0.313	–0.213	–0.274	1.000			
log P	–0.376	0.263	–0.463	0.102	–0.084	1.000		
ID	0.027	–0.027	–0.176	–0.103	0.359	–0.572	1.000	
W_A	–0.040	0.047	–0.476	–0.289	0.557	–0.295	0.834	1.000

Model 1

pIC_{50} = –7.730 (±2.126) χ_{eq} + 0.628 (±0.276) I_1 – 0.414 (±0.277) I_2 + 0.112 (±0.400) ID + 23.225

n = 20, R= 0.932, R^2 = 0.868, R^2_A = 0.833, S.E. = 0.168 $F_{(4–15)}$ = 24.694, Q = 5.548

Model 2

pIC_{50} = –8.004 (± 2.184) χ_{eq} + 0.666 (±0.299) I_1 – 0.436 (±0.273) I_2 + 0.127 (±0.326) log P + 23.744

n = 20, R = 0.933, R^2 = 0.871, R^2_A = 0.837, S.E. = 0.166 $F_{(4–15)}$ = 25.318, Q = 5.620

Model 3

pIC_{50} = –7.665 (±1.964) χ_{eq} + 0.742 (±0.296) I_1 – 0.336 (±0.275) I_2 + 0.760 (±0.935) W_A + 19.781

n = 20, R = 0.942, R^2 = 0.888, R^2_A = 0.858, S.E. = 0.155 $F_{(4–15)}$ = 29.603, Q = 6.077

Model 4

pIC_{50} = –6.884 (±2.07) χ_{eq} + 0.694 (±0.252) I_1 – 0.317 (±0.264) I_2 + 0.010 (±0.011) Mv + 18.855

n = 20, R = 0.947, R^2 = 0.897, R^2_A = 0.870, S.E. = 0.149 $F_{(4–15)}$ = 32.664, Q = 6.356

where, n = number of compounds in the data set, R = correlation coefficient, R^2 = coefficient of determination, R^2_A = adjusted coefficient of determination, S.E. = standard error of estimate, F = variance ratio[21,22] and Q = quality of fit[23,24].

COX-2 enzyme. The positive coefficient of indicator parameter I_1 suggest that presence of hydrogen (4-H) at R_2 position enhances the activity of the drugs and should be used at particular position in future drug modeling, while negative coefficient of indicator parameter I_2 suggest that presence of methyl group (2-Me) at R_2 position should be strictly avoided in the future drug modeling.

From the above all QSAR models, model (4) is the best model. According to model (4) compound having lowest value of equalized electro negativity, highest value of molar volume, presence of hydrogen (4-H) at R_2 position and absence of methyl group (2-Me) at R_2 position show highest inhibitory effect towards COX-2 enzyme. Focusing on Table 1 we find that compound **1** has maximum value of molar volume, minimum value of equalized electro negativity and hydrogen (4-H) is also present at R_2 position, which indicate that compound **1** inhibit COX-2 enzyme more potentially which is in good agreement with experimental finding Qiu *et al*. (2012). In order to confirm that the model with excellent statistics has also excellent prediction power too, we have evaluated quality factor Q. The predictive power as determined by the Pogliani Q parameter for the model (4) [Q = 6.356] confirms that this model has excellent statistics as well as excellent predictive power.

Predicted and residual values for the best model (4) are given in Table 3. Predicted values are the calculated activities of the equation and the residual values are the

Table 3. Comparison between observed and predicted activities and their residual values for model (4)

Sr. no.	pIC_{50} (Observed)	Model (4)	
		Predicted activity	Residual activity
1.	6.301	6.189	0.112
2.	5.194	5.332	–0.138
3.	5.347	5.454	–0.107
4.	5.745	5.523	0.222
5.	5.921	5.715	0.206
6.	5.538	5.672	–0.134
7.	5.284	5.454	–0.170
8.	5.569	5.715	–0.146
9.	5.143	5.235	–0.092
10.	5.201	5.147	0.054
11.	5.569	5.681	–0.112
12.	4.712	4.781	–0.069
13.	4.931	4.876	0.055
14.	5.119	4.992	0.127
15.	5.328	5.097	0.231
16.	5.167	5.035	0.132
17.	4.793	4.815	–0.022
18.	5.143	5.159	–0.016
19.	4.735	4.815	–0.080
20.	4.788	4.842	–0.054

difference between the observed biological activities and the calculated activities, which are found to be low.

In order to examine the relative potential of models, predictive correlation coefficient (R^2_{Pred}) were estimated by plotting graphs between observed and calculated pIC_{50} values obtained with the help of model (4), which are presented in Fig. 1. From the Fig. 1 R^2_{Pred} value obtained for model (4) is 0.897 and this is fairly high indicating the quality of the models.

Cross validation :

The cross validation analysis was performed using leave one out (LOO) method[25], in which one compound is removed from the data set and the activity is correlated using the rest of the data set. The cross-validated R^2 in each case was found to be very close to the value of R^2 for the entire data set and hence these models can be termed as statistically significant. Cross validation provides the values of PRESS, SSY, PSE, R^2_{CV} and R^2_A from which we can test the predictive power of the proposed model. These statistical parameters can be calculated from following equations :

$$PRESS = \Sigma(X_{obs} - X_{cal})^2 \quad \text{(i)}$$

$$SSY = \Sigma(X_{obs} - X_{mean})^2 \quad \text{(ii)}$$

$$PSE = \sqrt{PRESS/n} \quad \text{(iii)}$$

$$R^2_{cv} = 1 - \frac{PRESS}{SSY} \quad \text{(iv)}$$

$$R^2_A = 1 - (r^2)\left(\frac{n-1}{n-p-1}\right) \quad \text{(v)}$$

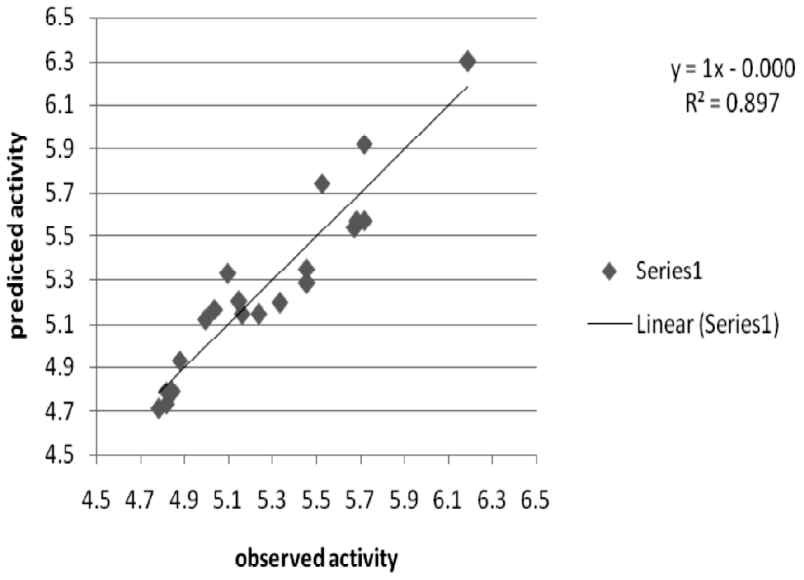


Fig. 1. A plot showing comparison between observed activity and predicted activity of model (4).

Table 4. Cross validated parameters and predictive error of coefficient of correlation (PE) for the proposed model

Model no.	n	PRESS	SSY	PRESS/SSY	R^2_{CV}	PSE	R	1-R^2	PE	6PE
1	20	0.425	2.796	0.152	0.848	0.154	0.932	0.132	0.021	0.126
2	20	0.415	2.805	0.147	0.853	0.152	0.933	0.129	0.020	0.120
3	20	0.362	2.858	0.127	0.873	0.142	0.942	0.112	0.018	0.108
4	20	0.332	2.889	0.115	0.885	0.136	0.947	0.103	0.016	0.096

It is argued that PRESS, is a good estimate of the real predictive error of the model and if it is smaller than SSY the model predicts better than chance and can be considered statistically significant. Furthermore, the ratio of PRESS/SSY can be used to calculate approximate confidence intervals of prediction of a new compound. To be a reasonable QSAR model PRESS/SSY should be smaller than 0.4. Also, if PRESS value is transformed in a dimension less term by relating it to the initial sum of squares, we obtain R^2_{CV} i.e. the complement to the traces on of unexplained variance over the total variance. The PRESS and R^2_{CV} have good properties. However, for practical purposes of end users the use of square root of PRESS/n, which is called predictive square error (PSE), is more directly related to the uncertainty of the predictions. The PSE values also support our results. The calculated cross-validated parameters confirm the validity of the models. All the requirements for an ideal model have been fulfilled by model (4), that's why, we have considered model (4) as the best model. R^2_A takes into account the adjustment of R^2. R^2_A is a measure of the percentage explained variation in the dependent variable that takes into account the relationship between the number of cases and the number of independent variables in the regression model, whereas R^2 will always increase when an independent variable is added. R^2_A will decrease if the added variable does not reduce the unexplained variable enough to offset the loss of decrease of freedom.

Predictive error of coefficient of correlation (PE) :

The predictive error of coefficient of correlation (PE)[26] is yet another parameter used to decide the predictive power of the proposed models. We have calculated PE value of all the proposed models and they are reported in Table 4. It is argued that if the values R < PE, then such correlation is not significant; however if values are R > PE in several times (at least three times), then values are correlated. However, if values are R > 6PE, then mathematically the correlation is unquestionably good. For all the models developed the condition R > 6PE is satisfied and hence they can be said to have a good predictive power.

Conclusions

The current study was performed to examine the applicability of the various empirical parameters in QSAR analysis for studying the biological activity of a series of dihydro-pyrazolyl-thiazolinone derivatives as potential COX-2 inhibitors. The calculated QSAR results based on empirical parameter demonstrate that more hydrophobic and bulkier group having more molar volume and less electro negativity should be preferred in future drug designing and enhances the activity of the drugs towards COX-2 enzyme. Indicator parameter suggest that the presence of hydrogen (4-H) and absence of methyl group (2-Me) at R_2 position enhances the activity of the drugs towards COX-2 enzyme.

Acknowledgement

The author A. K. Srivastava gratefully acknowledges the University Grants Commission, New Delhi, India for the grant, Grant No. 40-70/2011 (SR). Author A. Dwivedi is grateful to UGC, New Delhi, for the award of Junior Research Fellowship.

References

1. T. Puzyn, J. Leszcynski, M. T. Cronin and F. Bajot, "Recent Advances in QSAR Studies", 2010, Chap. 9, pp. 261-162.
2. T. Puzyn, J. Leszcynski, M. T. Cronin and W. Sippl, "Recent Advances in QSAR Studies", 2010, Chap. 4, pp. 123.
3. J. R. Vane, *Nature New Biology*, 1971, **232**, 231.
4. H. A. Park, K. Lee, S. J. Park, B. Ahn, J. C. Lee, H. Y. Cho and K. I. Lee, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 3307.
5. A. Tanitame, Y. Oyamada, K. Ofuji, M. Fujimoto, N. Iwai, Y. Hiyama, K. Suzuki, H. Ito, M. Wachi and J. Yamagishi, *J. Med. Chem.*, 2004, **47**, 3693.

6. G. Szabo, J. Fischer, A. Kis-Varga and K. Gyires, *J. Med. Chem.*, 2008, **51**, 142.
7. R. Ottana, R. Maccari, M. L. Barreca, G. Bruno, A. Rotondo, A. Rossi, G. Chiricosta, R. Di Paola, L. Sautebin, S. Cuzzocrea and M. G. Vigorita, *Bioorg. Med. Chem.*, 2005, **13**, 4243.
8. V. Gududuru, E. Hurh, J. T. Dalton and D. D. Miller, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5289.
9. K. M. Qiu, R. Yan, M. Xing, H. H. Wang, H. E. Cui, H. B. Gong and H. L. Zhu, *Bioorg. Med. Chem.*, 2012, **20**, 6648.
10. Acd-Lab software for calculating the referred physico-chemical parameters. Chem Sketch www.acdlabs.com/acdlabs-rss-feed.xml.
11. IV Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk and V. V. Prokopenko, *J. Comput. Aid. Mol. Des.*, 2005, **19**, 453.
12. E. A. Coats, C. S. Genther, C. D. Selassie, C. D. Strong and C. Hansch, *J. Med. Chem.*, 1985, **28**, 1910.
13. C. Hansch and A. Leo, "Exploring QSAR – Fundamentally and Application in Chemistry and Biology", American Chemical Society, Washington, DC, 1995.
14. C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, 1964, **86**, 1616.
15. A. K. Srivastava, Archana, A. A. Khan and M. Shakil, *Proc. Nat. Acad. Sci. India*, 2004 **74(A)**, 115.
16. A. K. Srivastava, Archana and M. Jaiswal, *J. Indian Chem. Soc.*, 2007, **84**, 260.
17. A. K. Srivastava, V. K. Pathak, Archana, M. Jaiswal and V. K. Agarwal, *Med. Chem. Res.*, 2011, **20**, 1713.
18. A. Dwivedi, A. Singh and A. K. Srivastava, *J. Saudi. Chem. Soc.*, 2013 http://dx.doi.org/10.1016/j.jscs.2013.01.007.
19. A. Dwivedi, A. K. Srivastava and A. Singh, *Med. Chem. Res.*, 2014, **23**, 1865.
20. X. Hui-Ding, *Chin. J. Struct. Chem.*, 2009, **28**, 621.
21. D. Bikash, G. Shovanlal, B. Subrata, S. Soma and J. Tarun, *Bioorg. Med. Chem.*, 2003, **11** 5493.
22. M. V. Diudea, "QSPR/QSAR studies for molecular descriptors", Nova Science and Huntington, New York, 2001.
23. L. Pogliani, *Amino Acids*, 1994, **6**, 141.
24. L. Pogliani, *J. Phys. Chem.*, 1996, **100**, 18065.
25. R. D. Cramer III, J. D. Bunce, D. E. Patterson and I. E. Frank, *Quant. Struct. Act. Relat.*, 1988, **7**, 18.
26. S. Chaterjee, A. S. Hadi and B. Price, "Regression Analysis by Example", 3rd ed., Wiley VCH, New York, 2000.