Iranian Journal of Analytical Chemistry
Volume 5, Issue 2, September 2018 (23-30)

مجله ایرانی شیمی تجزیه
دوره ۵، شماره ۲، مهر ماه ۱۳۹۷ (۳۰-۲۳)

مطالعه شبیه‌سازی: پارامترهای مدل‌سازی برای پیش‌بینی فعالیت مشتقات ایمینوکرومن

نصرت مددی ماهانی*، عذرا حرزاده
بخش شیمی، دانشگاه پیام نور، صندوق پستی ۳۶۹۷-۱۹۳۹۵ ، تهران، ایران
تاریخ دریافت: ۱۳ مرداد ۱۳۹۷ تاریخ پذیرش: ۲۸ شهریور ۱۳۹۷

Insilco study: Modeling Parameters for Prediction of Activity of Iminochromene Derivatives

Nosrat Madadi Mahani*, Azra Horzadeh
Department of Chemistry, Payame Noor University, 19395-4697 Tehran, Iran
Received: 4 August 2017 **Accepted:** 19 Sept 2018

چکیده
آنالیز رابطه فعالیت–ساختار کمی برای مجموعه‌ای از ۳۴ مشتق۸–هیدروکسی ۲–ایمینو کرومن با فعالیت بازداری آنزیم کربونیل ردکتاز۱ مورد بررسی قرار گرفت. محاسبات شیمی کوانتومی نیمه‌تجربی در سطح AM1 برای بدست آوردن ساختار بهینه ترکیبات استفاده گردید. کلیه توصیف‌گرها با استفاده از نرم‌افزار دراگون محاسبه گردید و زیرمجموعه‌ای از توصیف‌گرهای محاسبه‌شده از ۴۰۷ توصیف‌گر دراگون با استفاده از آنالیزهای رگرسیون خطی چند متغیره، حداقل مربعات جزئی و تحلیل مولفه‌های اصلی انتخاب گردید. نتایج نشان داد که رگرسیون خطی چند متغیره، فعالیت خوب و مناسبی را پیش‌بینی می‌نماید. بهترین مدل رگرسیون خطی چند متغیره با هفت توصیف‌گر انتخاب شد که این مدل ثبات بسیار خوبی نسبت به تغییرات داده‌ها برای روش‌های اعتبارسنجی نشان می‌دهد. مقادیر پیش‌بینی شده فعالیت این مشتقات با نتایج تجربی توافق مناسبی دارند. آنالیزها نشان داد که روش حداقل مربعات جزئی می‌تواند برای پیش-بینی فعالیت ترکیبات مورد مطالعه مفیدتر باشد. این مطالعه برای پیش‌بینی فعالیت ترکیبات دیگر از مشتقات همین گروه بسیار کاربردی و مهم می‌باشد.

واژه‌های کلیدی
رابطه فعالیت–ساختار کمی؛ ایمینوکرومن؛ رگرسیون خطی چند متغیره؛ حداقل مربعات جزئی؛ تحلیل مولفه‌های اصلی.

Abstract
A quantitative structure activity relationship analysis has been applied to a data set of 34 derivatives of 8-hydroxy-2-iminochromene with inhibitory activities for carbonyl reductase 1. Semi-empirical quantum chemical calculations at the AM1 level were used to find the geometry of the studied molecules. Whole numbers of descriptors were calculated with Dragon software, and a subset of calculated descriptors was selected from 407 Dragon descriptors with the multiple linear regression (MLR), partial least square and principal component analysis methods. Results displayed that the MLR method predicted of activity good enough. The best model of MLR, with seven descriptors was selected. Also it indicates very good consistency towards data variations for the validation methods. The predicted values of activities are in suitable agreement with the experimental results. The obtained results suggested that the PLS method could be more helpful to predict the biological activity of iminochromene derivatives. This study is be useful to predict the activity of other compounds in the same derivatives.

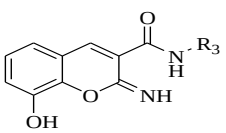
Keywords
Quantitative Structure-Activity; Relationship; Iminochromene derivatives; Multiple Linear Regression; Partial Least Square; Principal Component Analysis.

1. INTRODUCTION

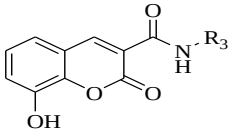
Human carbonyl reductase 1 (CBR1) is an emerging enzyme in drug discovery and development [1, 2]. CBR1 act as a major drug metabolizing enzyme in the metabolism of carbonyl-containing drugs [3, 4]. Also, CBR1 is potently inhibited by flavonoids with 7-hydroxylated chromene rings [5]. Using the chosen CBR1 inhibitor as the rector compound, 8-hydroxy-2-iminochromene (1a–1q) and 8-hydroxycoumarin (2a–2q) derivatives by displacing the pyridine moiety bound to the carboxamide of the chromene ring with substituted phenyl or benzyl rings have been synthesized by Hu and coworkers. These compounds are novel and potent inhibitors that are selective to human CBR1. The structures of the compounds and their inhibitory activities for CBR1 are summarized in Table 1 [6]. Quantitative structure–activity relationship

*Corresponding Author: nmmadady@gmail.com

Table 1. Chemical structure and IC_{50} of iminochromene derivatives [6].



1



2

R_3	Entry	IC_{50} (μM)	Entry	IC_{50} (μM)
Phenyl	1a	0.21 ± 0.012	2a	1.9 ± 0.16
2-Hydroxyphenyl	1b	0.33 ± 0.025	2b	0.47 ± 0.028
3-Hydroxyphenyl	1c	0.15 ± 0.011	2c	0.37 ± 0.030
4-Hydroxyphenyl	1d	0.88 ± 0.045	2d	1.3 ± 0.066
2-Fluorophenyl	1e	0.31 ± 0.037	2e	1.8 ± 0.42
3-Fluorophenyl	1f	0.37 ± 0.045	2f	2.5 ± 0.15
4-Fluorophenyl	1g	0.44 ± 0.086	2g	2.5 ± 0.11
2-Chlorophenyl	1h	0.034 ± 0.0035	2h	0.26 ± 0.037
3-Chlorophenyl	1i	0.12 ± 0.015	2i	1.5 ± 0.082
4-Chlorophenyl	1j	0.22 ± 0.015	2j	0.45 ± 0.0082
Benzyl	1k	0.33 ± 0.03	2k	0.92 ± 0.0091
2-Hydroxybenzyl	1l	0.35 ± 0.026	2l	1.3 ± 0.18
3-Hydroxybenzyl	1m	0.11 ± 0.0011	2m	1.1 ± 0.051
4-Hydroxybenzyl	1n	0.17 ± 0.022	2n	0.82 ± 0.013
2-Chlorobenzyl	1o	0.10 ± 0.013	2o	0.41 ± 0.019
3-Chlorobenzyl	1p	0.090 ± 0.00064	2p	1.1 ± 0.072
4-Chlorobenzyl	1q	0.26 ± 0.0072	2q	1.0 ± 0.0011

(QSAR) is widely accepted to be useful for the description of structural requirements of biologically active compounds [7] and has important role in drug design that relates structural features to biological activity [8]. The QSAR formulated mathematical models estimating the biological activities by based on the presumption that these are defined only by the molecular structures. The structures of molecule is translated into the so- called molecular descriptors, displaying some relevant feature of the molecules, with mathematical formulae gained from chemical graph theory, quantum mechanics and information theory [9]. The 3D QSAR pharmacophore models for aldo-keto reductase family 1 B10 were generated using Density functional theory calculations [10]. QSAR study on heterocyclic and aromatic sulfonamides compounds including 8-quinoline-sulfonyl carbonic anhydrase (CA) inhibitors with topical activity as antiglaucoma agents has been performed topologically with first-order valence connectivity index [11]. Molecular docking, Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) studies were investigated on a set of 4-azasteroidal human steroid 5α-reductase inhibitors [12]. The 2D, 3D QSAR and molecular docking studies on receptor antagonising thiazolo[3,2-a]pyrimidines as antipsychotic agents have been performed using VLife MDS3.5 software[13] .QSAR and docking of 1, 4-Dihydropyridines as novel antitubercular agents has been studied [14].Also, cytotoxic activity assessment and QSAR study of chromene-based chalcones have been investigated [15]. Classical QSAR studies on chromene derivatives as lanosterol demethylase inhibitor have been performed by Vasanthanathan et al. [16].

In the previous work, QSAR analysis was performed on iminochromene derivatives with the descriptors of quantum-chemical by density functional theory (DFT) method at B3LYP/6-31G level [17]. In the present study, QSAR analysis was investigated for on 8-hydroxy-2-iminochromene derivatives to organized quantitative relationship between biological activity of derivatives and their structural and physicochemical properties. Therefore; we suggest new models by using PCA, MLR and PLS methods that contain suitable descriptors for drug design.

2. EXPERIMENTAL

The data sets of the inhibitory activity **(**IC50) of 8-hydroxy-2-iminochromene derivatives as selective and potent inhibitors of human carbonyl reductase 1(CBR1) were chosen from reference 6. The specific CBR activities of these compounds were expressed as the effective concentration, which causes the half maximal inhibitory concentration (IC_{50}) and then used for subsequent QSAR analysis as dependent variables. The structural features and biological activity of these

compounds are listed in Table 1. This set was divided into a train set and a test set, randomly. The chemical structure of the molecules was drawn by the HyperChem software. The optimization of the molecular structures was carried out by semi-empirical AM1 method. The resultant geometry was loaded into Dragon software [18] to calculate 1481 descriptors in 18 different classes.

The program contains scripts for generating descriptors of different types including: constitutional, topological, radial distribution functions(RDF), geometry, topology and atoms-weighted assembly(GETAWAY), functional groups, weighted holistic invariants(WHIM), Randic molecular profiles, 3D-molecular representation of structure based on electron diffraction(3D-Morse) etc [19]. Some of molecular descriptors employed in the present study are summarized in Table 2.

For each compound in the training sets, the correlation equation was obtained with the same descriptors. Then, the obtained equation was used to predict pIC_{50} values for the compounds from the corresponding test sets. In this study, two programs including SPSS and Minitab were used for multiple linear regression (MLR) , principal component analysis (PCA) and partial least square (PLS) . The objectives of this work are to develop predictive QSAR models for the pIC50 of our studied molecules.

3. RESULT AND DISCUSSION

A QSAR study was performed for a series of iminochromene derivatives, for characterizing a quantitative relationship between structure chemical and biological activity. The Table 3 shows the values of the calculated descriptors calculated. For the development of QSAR equations, four different methods were uszed: (i) stepwise multiple linear regression (MLR), (ii) principal component analysis (PCA) and (iii) partial least squares (PLS).

The multiple linear regression statistic technique is used to study the relation between one dependent variable and several independent variables. It is a mathematic technique that minimizes differences between actual and predicted values. The Linear Regression method related to a larger family of models called generalized linear models. The choice of the training set is one of the most significant stages in the QSAR modeling, since the confirmation and optimization of a QSAR model are based on this training set. Applicability and predictability of a QSAR model also relay on the training set selection. The data set (n = 34) was divided casually into two groups: train set (n = 24) and test set (n = 10). The Pearson correlation coefficients are listed in the following Table 4. The correlation coefficient (R^2) matrix for the descriptors used in different MLR equations shows that no significant correlation exists between pairs of descriptors. The acquired matrix gives information on the positive or negative correlation between variables.

Modeling of pIC_{50} values of all training iminochromene derivatives take away to the best value corresponding to the linear combination of the descriptors the resulting equation is:

$$pIC_{50} = -4.413 - 2.253 * MAXDN + 0.364 * ATS5m - 15.311 * R2v^{+} + 5.765 * MATS4e + 2.556 * Mor21m - 0.189 * RDF145u + 0.930 * RDF140v \quad (1)$$

We used the pIC_{50} of the iminochromene derivatives as the dependant variable, equation 1 was resulted from the total of calculated descriptors. This model with acceptable statistical quality ($R^2 = 0.914$, SE = 0.156) indicated that the inhibitory activity of compounds is influenced by descriptors topological (MAXDN), 2D-autocorrelation (ATS5m and MATS4e), GETAWAY (R2v), 2D autocorrelation (MATS4e), 3D-MoRSE (Mor21m), Radial Distribution Function (RDF145u and RDF140v).

Table 2. List of molecular descriptors involved in QSAR equations.

Descriptor definition	abbreviation
maximal electro-topological negative variation	MAXDN
Broto-Moreau autocorrelation of lag 5 (log function) weighted by mass	ATS5m
R maximal autocorrelation of lag 2 / weighted by van der Waals volume	R2v+
Moran autocorrelation of lag 4 weighted by Sanderson electronegativity	MATS4e
signal 21 / weighted by mass	Mor21m
Radial Distribution Function - 145 / un-weighted	RDF145u
Radial Distribution Function - 140 / weighted by van der Waals volume	RDF140v

Table 3. Values of descriptors used in the selected QSAR model.

Name	MAXDN	ATS5m	R2v+	MATS4e	Mor21m	RDF145u	RDF140v
1a	2.097	30.029	0.028	0.068	-0.498	0.00	0.002
1b	2.250	33.234	0.026	-0.067	-0.526	0.00	0.002
1c	2.204	31.676	0.025	0.068	-0.533	0.00	0.031
1d	2.176	31.583	0.025	-0.01	-0.544	1.00	0.111
1e	2.33	33.648	0.028	-0.106	-0.526	0.00	0.002
1f	2.26	31.779	0.028	0.037	-0.52	0.00	0.002
1g	2.217	31.776	0.028	-0.054	-0.537	0.00	0.075
1h	2.174	36.958	0.032	-0.057	-0.741	0.00	0.003
1i	2.152	33.379	0.036	0.054	-0.69	0.00	0.002
1j	2.138	33.374	0.032	-0.019	-0.572	0.99	0.163
1k	2.053	30.602	0.037	0.047	-0.593	0.00	0.095
1l	2.16	32.271	0.033	-0.03	-0.556	0.00	0.034
1m	2.132	32.548	0.032	-0.001	-0.547	0.001	0.188
1n	2.114	31.949	0.028	-0.02	-0.382	0.00	0.00
1o	2.107	33.952	0.037	-0.033	-0.56	0.00	0.028
1p	2.094	34.671	0.041	-0.01	-0.621	1.87	0.000
1q	2.084	33.47	0.06	-0.027	-0.718	0.00	0.194
2a	2.461	30.375	0.029	0.092	-0.561	0.00	0.001
2b	2.54	33.581	0.027	-0.044	-0.527	0.00	0.002
2c	2.522	32.023	0.025	0.083	-0.601	0.00	0.023
2d	2.509	31.929	0.025	0.013	-0.591	1.00	0.199
2e	2.581	33.995	0.029	-0.09	-0.531	0.00	0.002
2f	2.553	32.125	0.029	0.047	-0.575	0.00	0.002
2g	2.534	32.123	0.029	-0.036	-0.601	0.00	0.085
2h	2.501	37.304	0.035	-0.033	-0.769	0.00	0.002
2i	2.492	33.725	0.038	0.07	-0.786	0.00	0.001
2j	2.486	33.72	0.032	0.005	-0.632	1.00	0.301
2k	2.445	30.949	0.036	0.077	-0.498	0.00	0.000
2l	2.506	32.617	0.033	-0.007	-0.595	0.00	0.060
2m	2.493	32.895	0.033	0.025	-0.627	0.00	0.100
2n	2.484	32.295	0.031	0.009	-0.608	0.00	0.195
2o	2.476	34.299	0.037	-0.008	-0.619	0.00	0.040
2p	2.47	35.017	0.043	0.018	-0.736	1.91	0.000
2q	2.465	33.816	0.059	0.003	-0.637	0.00	0.000

Table 4. Correlation Matrix for Model.

	RDF140v	RDF145u	Mor21m	MATS4e	R2v+	ATS5m	MAXDN
RDF140v	1						
RDF145u	0.135	1					
Mor21m	-0.112	0.338	1				
MATS4e	0.082	-0.509	0.002	1			
R2v+	-0.211	-0.643	0.516	-0.077	1		
ATS5m	-0.061	0.215	0.106	-0.011	-0.194	1	
MAXDN	-0.043	-0.123	0.042	-0.049	-0.082	0.236	1

The positive relation of activity and "ATS5m, MATS4e, Mor21m and RDF140v" displays that increasing of these descriptors cases increasing inhibitory activity of compounds. The obtained descriptors of the studied molecules listed on Table 5. pIC_{50} predicted of iminochromene derivatives by this model is partly like that observed. Values of pIC_{50} predicted in Table 6 are listed.

Also, Fig. 1 displays a very orderly distribution of pIC_{50}values based on the observed values.

PLS is an extension of regression, which can apply data with strongly correlated and⁄or numerous X variables. It gives diminished solution, which is statistically stronger than MLR. The linear PLS model finds 'new variables' (latent variables or X scores) that are linear combination of the basic variables. To avoid overfitting, a harsh test for the significance of each successive PLS component is needful and then stopping when the components are non-significant. Cross-validation is a practical and reputable method of

Table 5. The statistical parameters of different created QSAR models.

	R		R^2		SE	
	Train	Test	Train	Test	Train	Test
MLR	0.956	0.887	0.914	0.787	0.156	0.131
PLS	0.992	0.977	0.984	0.955	0.067	0.06
PCA	0.963	0.911	0.928	0.829	0.143	0.117

Table 6. The predicted activity (by MLR, PLS and PCA) for pIC_{50} of iminochromene derivatives.

Name	IC50	pIC50	MLR	PLS	PCR
1a	0.21	0.678	0.485	0.555	0.574
1b*	0.33	0.481	0.488	0.554	0.911
1c	0.15	0.824	0.827	0.889	0.836
1d	0.88	0.056	0.264	0.186	0.173
1e*	0.31	0.509	0.203	0.512	0.726
1f	0.37	0.432	0.520	0.352	0.484
1g*	0.44	0.357	0.116	0.357	0.519
1h	0.034	1.469	1.432	1.500	1.288
1i	0.12	0.921	0.887	0.871	0.813
1j	0.22	0.658	0.821	0.694	0.87
1k	0.33	0.481	0.378	0.370	0.635
1l*	0.35	0.456	0.399	0.449	0.511
1m	0.11	0.959	0.912	0.998	0.855
1n	0.17	0.770	0.933	0.727	0.565
1o	0.10	1.000	1.036	1.041	1.207
1p	0.09	1.046	0.862	1.069	0.992
1q*	0.26	0.585	0.346	0.583	0.709
2a	1.90	-0.279	-0.248	-0.189	-0.277
2b	0.47	0.328	0.076	0.312	0.604
2c	0.37	0.432	0.142	0.397	0.107
2d*	1.30	-0.114	-0.266	-0.241	-0.481
2e	1.80	-0.255	-0.172	-0.208	-0.284
2f	2.50	-0.398	-0.112	-0.432	-0.309
2g	2.50	-0.398	-0.538	-0.319	-0.332
2h*	0.26	0.585	0.841	0.528	0.497
2i	1.50	-0.176	0.062	-0.137	-0.21
2j	0.45	0.347	0.275	0.282	0.293
2k*	0.92	0.036	-0.036	0.108	-0.34
2l	1.30	-0.114	-0.201	-0.159	-0.025
2m	1.10	-0.041	0.073	-0.007	-0.073
2n*	0.82	0.086	-0.050	0.077	-0.325
2o	0.41	0.387	0.331	0.308	0.332
2p	1.10	-0.041	-0.027	-0.016	-0.025
2q*	1.00	0.000	-0.172	0.050	0.078

testing this significance [20]. Application of PLS thus allows the setting up of larger QSAR equations while still avoiding overfitting and eliminating most variables.

The obtained parameters explaining the electronic aspect of the investigated molecules listed on Table 5. The resulted predictions of the pIC_{50} using PLS method in was given in Table 6. pIC_{50} predicted of iminochromene derivatives by PLS method is little similar to that observed. Fig. 1 shows a normal distribution of pIC_{50}values based on the observed values. The resulted parameters describing the electronic aspect of the studied molecules are: R^2=0.984 and SE= 0. 067.

PLS is normally used in combination with cross-validation to obtain the optimum number of components. Leave-one-out cross-validation procedure was used to obtain the optimum number of factors based on the Haaland and Thomas F-ration criterion.

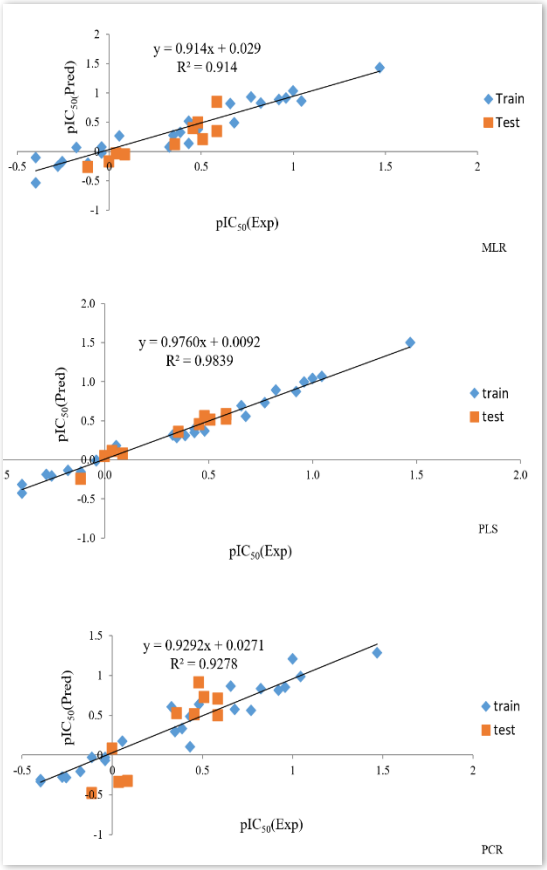


Fig. 1. Correlation of predicted vs. experimental pIC50.

The PLS have two goals: to estimated the matrix X of molecular structure descriptors to the matrix Y of dependent variables and to maximize the correlation between them. The leave–one–out (LOO) method was used to do the cross validated analysis. The cross–validated coefficient, Q2, is calculated using the following equation [21]:

$$Q^2 = 1 - \frac{\Sigma(y_i - y_{ipred})^2}{\Sigma(y_i - y_{imean})^2} \quad (2)$$

Where y_i is the i_{th} experimental pIC_{50}value, y_{ipred} is the i_{th} predicted pIC_{50} and y_{mean} is the mean of the experimental pIC_{50}. The accuracy of the model was mostly estimated by Root Mean Square Error (RMSE) that calculated using the following equation:

$$RMSE = \sqrt{\frac{y_i - y_{ipred}}{n}} \quad (3)$$

Where n = number of compounds, y_i = experimental value, y_{ipred} = predicted value [22]. The RMSE and Q^2 of the calibration using MLR method were obtained as 0.207 and 0.835, respectively. For suitable anticipation model the RMSE values should be low <0.3 and Q^2 is used as a criterion of both validity and predictive ability of the model.

To assess the predictive ability, predictions for all the test objects should be assessed independently of test set composition which can be random or dependent upon the size and distribution of the new data. Validation of models by means of objects whose data have not taken part in the process of model expansion is generally introduced to as external validation. Two different phrases for calculation of external validation Q^2, that is, $Q2$ based on predictions for external test compounds, were evaluated [23]. These expressions are:

$$Q^2_{F1} = 1 - \frac{\Sigma_{i=1}^{n_{EXT}}(y_i - \hat{y}_i)^2}{\Sigma_{i=1}^{n_{EXT}}(y_i - \bar{y}_{TR})^2} \quad (4)$$

$$Q^2_{F2} = 1 - \frac{\Sigma_{i=1}^{n_{EXT}}(y_i - \hat{y}_i)^2}{\Sigma_{i=1}^{n_{EXT}}(y_i - \bar{y}_{EXT})^2} \quad (5)$$

Where $\bar{y}_{TR}$ and $\bar{y}_{EXT}$ indicate the response means of the training set and the external test set, respectively. Also, the external predictive ability can be calculated as the following definition [24]:

$$Q^2_{F3} = 1 - \frac{\left[\Sigma_{i=1}^{n_{EXT}}(y_i - \hat{y}_i)^2\right]/n_{EXT}}{[\Sigma_{i=1}^{n_{EXT}}(y_i - \bar{y}_{TR})^2]/n_{TR}} \quad (6)$$

Values from function Q^2_{F1}, Q^2_{F2}, and Q^2_{F3}, R^2 values and root-mean-square error over the external evaluation set (RMSE) for eight data sets are listed in Table 7.

All the three functions Q^2_{F1}, Q^2_{F2}, and Q^2_{F3} in data sets of 2, 3 and 5 give suitable approximations of the model fit when test objects are identically distributed and cover the whole range of the training set. Also, 2 data set has the smallest RMSE which corresponds to the largest value of Q^2_{F3}(RMSE=0.168 and Q^2_{F3}=0.875).

Table 7. Model fit estimates (cross- validation) for iminochromene derivatives data sets.

Data Set	R^2	Q^2f_1	Q^2f_2	Q^2f_3	RMSE
1	0.787	0.491	0.444	0.862	0.189
2	0.81	0.807	0.794	0.875	0.169
3	0.789	0.659	0.659	0.797	0.216
4	0.908	0.807	0.770	0.859	0.178
5	0.740	0.395	0.397	0.874	0.183

Principal component analysis (PCA) as one of most widely used data-reduction techniques[25] is an important and useful algebraic tool in drug design and discovery. PCA, in a typical quantitative structure–activity relationship (QSAR) study, analyzes an original data matrix in which molecules are described by several intercorrelated quantitative dependent variables. The molecules of 8-hydroxy-2-iminochromene derivatives were studied by statistical method based on the PCA. The obtained parameters from PCA analysis of the studied molecules are listed on Table 5. The resulted predictions of the pIC_{50} using PCA method in were given in Table 6. Values of pIC_{50} predicted of iminochromene derivatives by PCA method is almost similar to that observed. Fig. 1 shows a very sufficient distribution of pIC_{50} values based on the observed values. The resulted parameters describing the electronic aspect of the studied molecules are: R^2=0.928 and SE= 0.143. It confirmed that the PCA results were the best to creating the quantitative structure activity relationship models.

4. CONCLUSION

In this work, we have studied the QSAR models to predict the activity of iminochromene derivatives. The study of the MLR, PCR and PLS models show that the PLS method has substantially better predictive capability than the other methods. With considering the error, the prediction of the pIC_{50} values was quite satisfactory and the performance of the QSAR model to predict pIC_{50}value was also calculated using the internal cross-validation method. The resonableness of the three created models used in this study has good agreement and great predictive power. By defining the molecular descriptors in the regression model, we finalize that the decreased MAXDN, $R2v^+$ and RDF145u as well as the increased values of ATS5m, MATS4e, Mor21m and RDF140v are valid for the larger activity of the studied compounds. Finally, the accuracy and predictability of the suggested models were demonstrated by evaluating essential statistical indexs such, as Q^2, R^2 and RMSE of different models using different statistical models and descriptors, as shown in Table 5.

ACKNOWLEDGMENTS

We gratefully thank Payam Noor University for financial support. The authors would like to express their appreciations to department of chemistry this university for providing research facilities.

REFERENCES

[1] O.S. Bains, T.A. Grigliatti, R.E. Reid and K.W. Riggs, Naturally occurring variants of human aldo-keto reductases with reduced in vitro metabolism of daunorubicin and doxorubicin, *J. Pharmacol. Exp. Ther.* 335 (2010) 533-345.

[2] F. Hoffmann and E. Maser, Carbonyl reductases and pluripotent hydroxysteroid dehydrogenases of the short-chain dehydrogenase/reductase superfamily, *Drug Metab. Rev.* 39 (2007) 87-144.

[3] T. Matsunaga, S. Shintani and A. Hara, Multiplicity of mammalian reductases for xenobiotic carbonyl compounds, *Drug Metab. Pharmacokinet.* 21 (2006) 1-18.

[4] P. Malatkova, E. Maser and V. Wsol, Human carbonyl reductases, *Curr. Drug Metab*. 11 (2010) 639-658.

[5] J.A. Doorn, E. Maser, A. Blum, D.J. Claffey and D.R. Petersen, Human carbonyl reductase catalyzes reduction of 4-oxonon-2-enal, *Biochemistry* 43 (2004) 13106-13114.

[6] D. Hu, N. Miyagi, Y. Arai, T. Oguri, T. Miura, T. Nishinaka T, et al. Synthesis of 8-hydroxy-2-iminochromene derivatives as selective and potent inhibitors of human carbonyl reductase 1, *Org. Biomol. Chem.* 13 (2015) 7487-7499.

[7] E. Arkan, M. Shahlaei, A. Pourhossein, K. Fakhri and A. Fassihi, Validated QSAR analysis of some diaryl substituted pyrazoles as CCR2 inhibitors by various linear and nonlinear multivariate chemometrics methods, *Eur. J. Med. Chem.* 45 (2010) 3394-3406.

[8] M. Shahlaei, A. Madadkar-Sobhani, L. Saghaie and A. Fassihi, Application of an

expert system based on Genetic Algorithm–Adaptive Neuro-Fuzzy Inference System (GA–ANFIS) in QSAR of cathepsin K inhibitors, *Expert Syst. Appl.* 39 (2012) 6182-6191.

[9] A.R. Katritzky, V.S. Lobanov and M. Karelson, QSPR: the correlation and quantitative prediction of chemical and physical properties from structure, *Chem. Soc. Rev.* 24 (1995) 279-287.

[10] R. Kumar, M. Son, R. Bavi, Y. Lee, C. Park, V. Arulalapperumal, et al. Novel chemical scaffolds of the tumor marker AKR1B10 inhibitors discovered by 3D QSAR pharmacophore modeling, *Acta Pharmacol. Sin.* 36 (2015) 998-1012.

[11] V.K. Agrawal, S. Bano, C.T. Supuran and P.V.Khadikar , QSAR study on carbonic anhydrase inhibitors: aromatic/heterocyclic sulfonamides containing 8-quinoline-sulfonyl moieties, with topical activity as antiglaucoma agents, *Eur. J. Med .Chem.* 39 (2004) 593-600.

[12] R. Kumar, P. Malla, A. Verma and M. Kumar, Design of potent human steroid 5α-reductase inhibitors: 3D-QSAR CoMFA, CoMSIA and docking studies. *Med. Chem. Res.* 22 (2013) 4568-4582.

[13] R.L. Sawant, S.S.Ramdin and J.B. Wadekar, Synthesis, QSAR and docking studies of 5HT2A receptor antagonising thiazolo[3,2-a]pyrimidines as antipsychotic agents, *Marmara Pharmaceut. J.* 18 (2014) 109-116.

[14] R.Y. Chaudhari, S.B. Bhise, A. Yadav and T. Sonawane, QSAR, Synthesis and Docking Study of 1, 4-DHP as Novel Antitubercular Agents, *J. Pharm. Res. Clin. Pract.* 2 (2016) 1-9.

[15] L. Firoozpour , N. Edraki , M. Nakhjiri , S. Emami , M. Safavi , S.K.Ardestani , et al. Cytotoxic activity evaluation and QSAR study of chromene-based chalcones. *Arch. Pharm. Res.* 35 (2012) 2117-2125.

[16] P. Vasanthanathan , M. Lakshmi ,M.A. Babu and S.G. Kaskhedikar , Classical QSAR study on chromene derivatives as lanosterol 14alpha- demethylase inhibitor: a non azole antifungal target, *Med. Chem.* 2 (2006) 363-367.

[17] N. Madadi Mahani, A. Mohadesi Zarandi and A. Horzadeh, QSAR studies of novel iminochromene derivatives as carbonyl reductase 1 (CBR1) inhibitors, Marmara *Pharm. J.* 22 (2018) 227-236.

[18] Talete srl, *Dragon* (ver. 5.4), Milano, Italy. Web site: www.talete.mi.it/products/ software. htm

[19] R.Todeschini and V.Consonni *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim (2000) pp. 667-690.

[20] S. Wold and L. Ericksson, Partial least squares projections to latent structures (PLS) in chemistry. *In Encyclopedia of computational chemistry*, Ragu & Schleyer, P. (ed.), John Wiley & Sons, ltd. Chichester(2002).

[21] V. Consonni, D. Ballabioand and R. Todeschini, Comments on the definition of the Q^2 parameter for QSAR validation, *J. Chem. Inf. Model*. 49 (2009) 1669-1678.

[22] V. Consonni, D. Ballabioand and R. Todeschini , Evaluation of model predictive ability by external validation techniques, *J. Chemometrics.* 24 (2010) 194-201.

[23] L. M. Shi, H .Fang, W.Tomg, J. Wu, R. Perkins, R.M. Blair, W.S.Branham, S.L.Dial, C.L. Moland and D.M. Sheenan, QSAR Models Using a Large Diverse Set of Estrogens, *J. Chem. Inf. Comput .Sci.* 41 (2001) 186–195.

[24] D.M. Hawkins, the Problem of Overfitting, *J. Chem. Inf. Comput. Sci.* 44 (2004) 1–12.

[25] S. Vaira , V.E. Mantovani , J.C. Robles, J.C. Sanchis and H.C. Goicoechea , Use of Chemometrics: Principal Component Analysis (PCA) and Principal Component Regression (PCR) for the Authentication of Orange Juice, *J. Anal. Lett.* 32 (1999) 3131-3141.