MARMARA
PHARMACEUTICAL JOURNAL

Research Article

www.marmarapharmj.com

QSAR studies of novel iminochromene derivatives as carbonyl reductase 1 (CBR1) inhibitors

Nosrat Madadi Mahani *, Alireaza Mohadesi Zarandi, Azra Horzadeh

Department of Chemistry, Payame Noor University, 19395-4697 Tehran, Iran

Correspondence: nmmadady@gmail.com (N.M.M.); ORCID No: 0000-0001-6967-5979.

Submitted: 9 July 2017 / Revised: 5 October 2017 / Accepted: 7 November 2017

ABSTRACT: A quantitative Structure-Activity Relationship (QSAR) model was applied to the prediction of the activity of iminochromene derivatives. The inhibition activity of 34 carbonyl reductase 1 (CBR1) inhibitors were modeled with the descriptors of quantum-chemical calculations with density functional theory (DFT) method at B3LYP/6–31G level. This study was conducted using the multiple linear regressions (MLR), the partial least square analysis (PLS) and the principal component analysis (PCA) method. Results displayed that the MLR method predicted of activity good enough. The best model, 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 PCA method could be more helpful to predict the biological activity of iminochromene derivatives. It is anticipated to be useful to predict the activity of other compounds in the same groups.

KEYWORDS: Quantitative Structure-Activity; Relationship (QSAR); Iminochromene derivatives; Multiple Linear Regression (MLR); Partial Least Square (PLS); Density Functional Theory (DFT); Principal Component Analysis (PCA).

1. INTRODUCTION

A branch of the short-chain dehydrogenase/reductase group is human carbonyl reductase 1 (CBR1) that decrease a variety of carbonyl compounds including therapeutic drugs [1]. CBR1 inhibitors are purposed to be suitable agents for adjutant therapy with twofold useful effect in prolongation the anticancer effect of the anthracyclines while reducing cardiotoxicity[2] . As well as, CBR1 decreases a wide variety of xenobiotic carbonyl alcohols, which are easier to be conjugated and omitted. Also, selective inhibitors of CBR1 may be helpful in clarifying the physiological functions of the enzyme and its portion to metabolism of lately developed carbonyl-containing drugs [3]. CBR1 is prevented by structurally various compounds, which are divided into natural and synthetic inhibitors. Some of synthetic inhibitors are 3-(1-tert-butyl-4-amino-1H-pyrazolo [3,4-d]pyrimidin-3-yl)phenol (hydroxy-PP), triclosan, ethacrynic acid, indomethacin and zearalenone analogues [4-6]. The natural inhibitors are resveratrol [7] and flavonoids [8-9] that a molecular docking study of four flavonoids in CBR1 proposed their different relative positions in its active site [10]. 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 [11]. 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 [11].

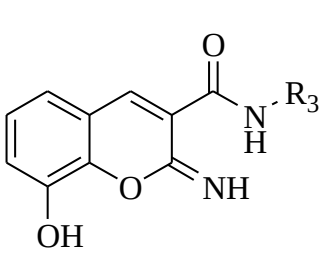
The QSAR methods are mathematical equations describing chemical structure to their target property (biological activity). These predictive models remit information that is effective for medicinal chemistry and drug discovery [12–14]. QSAR study on heterocyclic and aromatic sulfonamides compounds including 8-quinoline-sulfonyl carbonic anhydrase (CA) inhibitors with topical activity as antiglaucoma agents has been performed topologically with first-order valence connectivity index [15]. The 3D QSAR pharmacophore models for aldo-keto reductase family 1 B10 were generated using Density functional theory calculations [16]. 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 [17] .

How to cite this article: Madadi Mahani N, Mohadesi Zarandi A, Horzadeh A. QSAR studies of novel iminochromene derivatives as carbonyl reductase 1 (CBR1) inhibitors. Marmara Pharm J. 2018; 22 (2): 227-236.

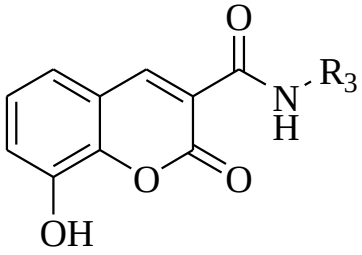
ISSN: 1309-0801

https://doi.org/10.12991/mpj.2018.60
Marmara Pharm J 2018; 22(2): 227-236

Table 1. Chemical structure and IC_{50} values of iminochromene derivatives [11].



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

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 [18]. QSAR and docking of 1, 4-Dihydropyridines as novel antitubercular agents has been studied [19]. Classical QSAR studies on chromene derivatives as lanosterol demethylase inhibitor have been performed by Vasanthanathan et al. [20]. Also, cytotoxic activity assessment and QSAR study of chromene-based chalcones have been investigated [21].

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 linear models by using chemometrics tools such as PCA, MLR and PLS that contain suitable descriptors for drug design.

2. RESULTS 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 2 shows the values of the calculated parameters calculated with DFT method at B3LYP 6–31G level of the studied molecules. Also, some values of molecular descriptors were obtained by using HyperChem that in table 3 are listed.

Table 2. Values of the descriptors gained at the B3LYP/6-31G level of studied CBR 1 inhibitors.

Name	HF Hartree	HOMO Hartree	LUMO Hartree	Gap Hartree	μ Hartree	η Hartree	σ Hartree	ω Hartree	DM Debye	ΔG kcal/mol	ΔH kcal/mol	C_V	ΔS	E_T	ZPE
1a	-951.8546	-0.2124	-0.0892	-0.1231	-0.1508	0.0616	16.2403	0.1847	3.3609	-951.6447	-951.5829	64.7100	130.0071	169.9290	0.2546
1b	-1027.0439	-0.2043	-0.0837	-0.1206	-0.1440	0.0602	16.5851	0.1719	2.6907	-1026.8320	-1026.7675	69.5010	135.6730	172.8630	0.2580
1c	-1027.0441	-0.2109	-0.0890	-0.1218	-0.1499	0.0609	16.4096	0.1846	1.9323	-1026.8322	-1026.7676	69.5201	135.7401	172.8850	0.2581
1d	-1027.0435	-0.1987	-0.0883	-0.1103	-0.1435	0.0552	18.1225	0.1866	2.9677	-1026.8315	-1026.7671	69.5220	135.5831	172.8820	0.2581
1e	-1051.0678	-0.2199	-0.0901	-0.1298	-0.1550	0.0649	15.4012	0.1851	2.2943	-1050.8675	-1050.8037	67.7590	134.2541	165.1340	0.2461
1f	-1051.0682	-0.2224	-0.0939	-0.1285	-0.1581	0.0642	15.5642	0.1946	4.5047	-1050.8682	-1050.8042	67.8471	134.7020	165.1270	0.2461
1g	-1051.0677	-0.2164	-0.0931	-0.1234	-0.1548	0.0617	16.2114	0.1942	5.4758	-1050.8676	-1050.8036	67.8381	134.5601	165.1360	0.2461
1h	-1411.4279	-0.2235	-0.0913	-0.1321	-0.1574	0.0660	15.1412	0.1876	2.4701	-1411.2302	-1411.1648	68.6070	137.5910	164.4910	0.2447
1i	-1411.4305	-0.2250	-0.0951	-0.1299	-0.1601	0.0649	15.3905	0.1972	4.9202	-1411.2328	-1411.1673	68.7210	137.5920	164.5270	0.2447
1j	-1411.4306	-0.2202	-0.0951	-0.1250	-0.1576	0.0625	15.9987	0.1988	6.2022	-1411.2327	-1411.16751	68.6920	137.2920	164.5410	0.2447
1k	-991.1533	-0.2321	-0.0843	-0.1478	-0.1582	0.07392	13.52814	0.1693	2.6641	-990.9186	-990.8514	69.5260	141.3881	188.8440	0.2833
1l	-1066.3366	-0.2155	-0.0775	-0.1380	-0.1465	0.0690	14.4885	0.1555	4.2573	-1066.0996	-1066.0300	74.370	146.3830	191.7940	0.2867
1m	-1066.3360	-0.2172	-0.0789	-0.1383	-0.1480	0.06915	14.4602	0.1584	5.7865	-1066.0995	-1066.0297	74.4141	146.7800	191.7270	0.2866
1n	-1066.3368	-0.2121	-0.0793	-0.1328	-0.1457	0.0664	15.0625	0.1600	6.1877	-1066.0991	-1066.0301	74.3161	145.1471	191.8520	0.2869
1o	-1450.7221	-0.2351	-0.0830	-0.1520	-0.1590	0.0760	13.1561	0.1664	7.3984	-1450.4990	-1450.4288	73.5690	147.5731	183.4120	0.2733
1p	-1450.7227	-0.2350	-0.0834	-0.1516	-0.1592	0.0758	13.1934	0.16718	7.8534	-1450.5006	-1450.4295	73.6180	149.5360	183.390	0.2733
1q	-1450.7226	-0.2344	-0.0834	-0.1509	-0.1588	0.0755	13.2468	0.16721	7.2940	-1450.5005	-1450.4295	73.6180	149.3830	183.3720	0.2732
2a	-971.7255	-0.2161	-0.0964	-0.1196	-0.1562	0.0598	16.71961	0.2041	6.0681	-971.52805	-971.4662	63.9760	129.9801	162.0540	0.2421
2b	-1046.9153	-0.2070	-0.0908	-0.1162	-0.1489	0.0581	17.21022	0.1909	3.9944	-1046.7158	-1046.6513	68.7410	135.6331	165.0250	0.2456
2c	-1046.9146	-0.2137	-0.0962	-0.1174	-0.1549	0.0587	17.02562	0.2044	5.355	-1046.7152	-1046.6506	68.7710	135.7121	165.0120	0.2456
2d	-1046.9143	-0.2017	-0.0955	-0.1062	-0.1485	0.0531	18.8217	0.2077	6.5427	-1046.7148	-1046.6504	68.7870	135.5580	165.0040	0.2456
2e	-1070.9377	-0.2239	-0.0975	-0.1264	-0.1606	0.0632	15.8152	0.2042	6.0097	-1070.75	-1070.6861	67.0131	134.3310	157.2940	0.2337
2f	-1070.9386	-0.2261	-0.1010	-0.1250	-0.1635	0.0625	15.9910	0.2139	8.0893	-1070.751	-1070.6870	67.0930	134.6581	157.2621	0.2336
2g	-1070.9383	-0.2200	-0.1965	-0.0235	-0.2083	0.0117	85.0701	1.8454	8.2862	-1070.7507	-1070.6868	67.1110	134.5361	157.2630	0.2336
2h	-1431.2980	-0.2266	-0.0985	-0.1280	-0.1625	0.0640	15.6176	0.2063	6.3209	-1431.1127	-1431.0474	67.8380	137.4550	156.6630	0.2323
2i	-1431.3006	-0.2287	-0.1022	-0.1265	-0.16543	0.0632	15.8127	0.2163	8.5723	-1431.1154	-1431.0500	67.9441	137.4980	156.6741	0.2322
2j	-1431.3012	-0.2235	-0.1022	-0.1213	-0.1629	0.0606	16.4853	0.2187	9.0219	-1431.1158	-1431.0505	67.9531	137.2641	156.6710	0.2323
2k	-1011.0073	-0.2399	-0.0875	-0.1525	-0.1637	0.0762	13.1121	0.1757	5.8070	-1010.7852	-1010.7183	69.3980	140.5720	180.7610	0.2703
2l	-1086.2144	-0.2191	-0.0906	-0.1286	-0.1548	0.0643	15.5448	0.1862	3.8916	-1085.9898	-1085.9203	73.5321	146.1811	183.9970	0.2744
2m	-1086.2140	-0.2205	-0.0918	-0.1288	-0.1561	0.0644	15.5303	0.1893	5.4167	-1085.9896	-1085.9199	73.5901	146.5761	183.9210	0.2743
2n	-1086.2137	-0.2174	-0.0918	-0.1257	-0.1546	0.0628	15.9147	0.1902	6.4006	-1085.9901	-1085.9197	73.6041	148.0350	183.9250	0.2743
2o	-1470.5995	-0.2456	-0.0960	-0.1495	-0.1708	0.0747	13.3743	0.1951	7.0210	-1470.3888	-1470.3187	72.7410	147.4590	175.6121	0.2610
2p	-1470.6002	-0.2438	-0.0963	-0.1475	-0.1701	0.0737	13.5602	0.1959	8.0291	-1470.3902	-1470.3194	72.8030	148.9411	175.5790	0.2609
2q	-1470.6011	-0.2425	-0.0963	-0.1462	-0.1694	0.0731	13.6743	0.1962	7.1901	-1470.3898	-1470.3202	72.7091	146.3590	175.6771	0.2612

https://doi.org/10.12991/mpj.2018.60
Marmara Pharm J 2018; 22(2): 227-236

Table 3. Values of the chemical descriptors were obtained using HyperChem software.

Name	Mass (amu)	Polar	Ref	LogP	HE(kcal/mol)	Vol	Sur_{Approx}
1a	280.08	30.78	76.18	2.33	-15.23	790.65	365.84
1b	296.08	31.42	77.88	2.05	-18.56	808.00	361.51
1c	296.08	31.42	77.88	2.05	-21.96	812.75	380.47
1d	296.08	31.42	77.88	2.05	-22.19	813.33	381.25
1e	298.08	30.69	76.4	2.47	-14.51	797.24	365.29
1f	298.07	30.69	76.4	2.47	-14.85	799.97	377.77
1g	298.07	30.69	76.4	2.47	-14.9	800.12	378.19
1h	314.04	32.71	80.99	2.85	-14.41	826.5	385.53
1i	314.04	32.71	80.99	2.85	-14.79	835.73	402.67
1j	314.04	32.71	80.99	2.85	-14.85	837.25	403.04
1k	294.10	32.62	81.02	2.43	-14.31	848.28	399.69
1l	310.09	33.25	82.71	2.14	-18.26	867.33	405.31
1m	310.09	33.25	82.71	2.14	-19.92	873.32	415.82
1n	310.09	33.25	82.71	2.14	-20.39	869.19	411.8
1o	328.06	34.55	85.82	2.95	-12.89	888.03	430.73
1p	328.07	34.55	85.82	2.95	-12.91	896.01	437.71
1q	328.07	34.55	85.82	2.95	-12.86	896.91	438.6
2a	281.07	29.93	74.13	2.16	-12.78	778.11	384.2
2b	297.06	30.57	75.82	1.88	-16.43	796.39	378.5
2c	297.06	30.57	75.82	1.88	-19.51	800.59	398.72
2d	297.06	30.57	75.82	1.88	-19.73	800.64	399.56
2e	299.04	29.84	74.34	2.30	-12.3	784.85	383.48
2f	299.04	29.84	74.34	2.30	-12.41	787.43	396
2g	299.04	29.84	74.34	2.30	-12.46	787.44	396.45
2h	315.03	31.86	78.93	2.68	-12.23	815.83	404.43
2i	315.03	31.86	78.93	2.68	-12.35	824.62	420.87
2j	315.03	31.86	78.93	2.68	-12.41	824.66	421.3
2k	295.08	31.77	78.96	2.26	-12.31	833.65	417.57
2l	311.08	32.41	80.65	1.97	-16.99	854.49	420.87
2m	311.08	32.41	80.65	1.97	-18.64	859.83	431.25
2n	311.08	32.41	80.65	1.97	-18.83	860.20	432.4
2o	329.04	33.70	83.77	2.77	-11.62	873.42	446.23
2p	329.045	33.70	83.77	2.77	-11.61	882.96	453.26
2q	329.0454	33.7	83.77	2.77	-11.69	879.77	449.37

2.1. Multiple Linear Regressions (MLR)

The multiple linear regression statistic method is used to study the relation between one dependent variable and several independent variables. Also, minimizes differences between experimental and predicted values. The MLR was acquired using the software SPSS to predict activities pIC_{50}. 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.

Table 4. Correlation coefficient (R^2) matrix for descriptors represented in multiple linear regression eq. 1.

Sur_{Approx}	HE	DM	ZPE	μ	η	Polar	
						1	Polar
					1	0.627	η
				1	0.611	0.154	μ
			1	0.402	0.479	0.671	ZPE
		1	-0.261	-0.548	-0.094	0.133	DM
	1	0.602	-0.438	-0.610	0.039	0.021	HE
1	0.293	0.606	0.391	-0.242	0.394	0.742	Sur_{Approx}

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:

$$pIC_{50} = -9.133 - 22.630\times\mu + 22.378\times\eta + 0.066\times DM - 8.045\times ZPE - 0.030\times Sur_{Approx} + 0.536\times Polar - 0.079\times HE \quad \text{(Eq. 1)}$$

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 = 0.836, SE = 0.215) indicated that the inhibitory activity of compounds is influenced by topological parameters (surface area), ZPE, HE and electronic chemical potential values as representative of quantum chemical descriptors. The positive relation of activity and "dipole moment, hardness, polarizibility" displays that increasing of these descriptors cases increasing inhibitory activity of compounds. The obtained descriptors demonstrating the electronic characteristic 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. The leave–one–out (LOO) approach was applied to carry out the cross-validated analysis. Q^2(cross–validated coefficient) is computed using the following equation [22-23]:

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

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:

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

Where n = number of compounds, y_i = experimental value, y_{ipred} = predicted value [24, 25]. 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.

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

RMSE	Q^2	SE		R^2		R		Method
Train	Train	Test	Train	Test	Train	Test	Train	
0.207	0.835	0.146	0.215	0.734	0.836	0.857	0.914	MLR
0.199	0.846	0.115	0.205	0.835	0.851	0.914	0.923	PLS
0.155	0.908	0.153	0.161	0.710	0.908	0.843	0.953	PCA

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

Name AME	MLR	PLS	PCA	Name AME	MLR	PLS	PCA
1a	0.56	0.485	0.5392	2a	-0.28	-0.211	-0.116
1b*	1.04	0.853	1.0013	2b	0.15	0.058	0.1945
1c	0.84	0.843	0.9812	2c	0.03	0.184	0.2697
1d	0.63	0.551	0.2354	2d*	-0.17	-0.073	-0.4931
1e*	0.64	0.415	0.3773	2e	-0.1	-0.118	-0.2241
1f	0.49	0.334	0.1944	2f	-0.28	-0.265	-0.3964
1g*	0.41	0.235	-0.01	2g	-0.4	-0.400	-0.4038
1h	1.21	1.248	1.4088	2h*	0.44	0.621	0.7295
1i	0.92	0.807	0.9226	2i	0.15	0.170	0.2931
1j	0.89	0.703	0.6739	2j	0.06	0.019	0.0592
1k	0.62	0.590	0.6093	2k*	-0.04	0.063	0.1928
1l*	0.81	0.639	1.0201	2l	-0.05	-0.192	-0.0778
1m	0.76	0.692	0.9535	2m	-0.1	-0.129	-0.1062
1n	0.83	0.682	0.7777	2n*	-0.12	0.043	-0.1042
1o	1.07	1.000	1.0407	2o	0.36	0.448	0.303
1p	0.89	0.866	0.8962	2p	0.18	0.202	0.0552
1q*	0.81	0.714	0.8282	2q*	0.22	-0.015	-0.0793

* test set

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 [26,27]. These expressions are:

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

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

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 [28]:

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

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.

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

Data set	R^2	Q^2_{F1}	Q^2_{F2}	Q^2_{F3}	RMSE
1	0.836	0. 110	0.022	0.758	0.251
2	0.798	0.628	0.620	0.815	0.210
3	0.799	0.727	0.726	0.696	0.245
4	0.817	0.292	0.241	0.538	0.324
5	0.817	0.643	0.643	0.786	0.222
6	0.808	-1.205	-1.632	-0.618	0.601
7	0.869	-1.269	-1.270	0.449	0.379
8	0.871	-0.966	-0.978	0.587	0.331

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=0.210 and Q^2_{F3}=0.815).

2.2. Partial Least Square analysis (PLS)

PLS technique is a generalization of regression, which can apply data with forcefully correlated and/or numerous independent variables. The linear PLS model detects new variables that are linear combination of the principal variables. To eschewing over fitting, a formidable test for the significance of each successive PLS component is essential and then pausing when the components are non-significant. The PLS have two purposes: to estimate the matrix *X* of molecular structure descriptors to the matrix Y of dependent variables and to maximize the relationship between them [29]. We presented the data matrix organized clearly from the descriptors offered by MLR corresponding to the molecules, to the PLS. This method used the coefficients *R*, R^2, and the SE values to chose the best regression performance.

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 gas phase was given in table 6. pIC_{50} predicted of iminochromene derivatives by PLS method is little similar to that observed. Figure 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=0.851 and RMSE= 0.155. Cross-validation is a practical and validity method for testing the significance. PLS is usually used in merging with cross-validation to gain the optimum number of components.

2.3. Principal Components Analysis (PCA)

The PCA is a helpful statistical technique for summing up all the information coded in the structures of compounds and very useful for identifying the link between the different variables [30]. 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 gas phase were given in Table 6. Values of pIC_{50} predicted of iminochromene derivatives by PCA method is almost similar to that observed. Figure1 shows a very adequate distribution of pIC_{50}values based on the observed values. The obtained parameters defining the electronic aspect of the studied molecules are: R^2=0.908 and RMSE= 0.155. It corroborates that the PCA results were the best to creating the quantitative structure activity relationship models.

3. 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 PCR 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 sanity of the three created models used in this study has good consistency and great predictive power. By defining the molecular descriptors in the regression model, we finalize that the decreased surface area, HE, ZPE and electronic chemical potential as well as the increased magnitude of dipole moment, hardness and polarizibility are reliable for the larger activity of the investigated compounds. Eventually, the accuracy and predictability of the suggested models were demonstrated by evaluating essential statistical indexs such, as Q^2, R2 and RMSE of different models using different statistical models and descriptors, as shown in Table 5.

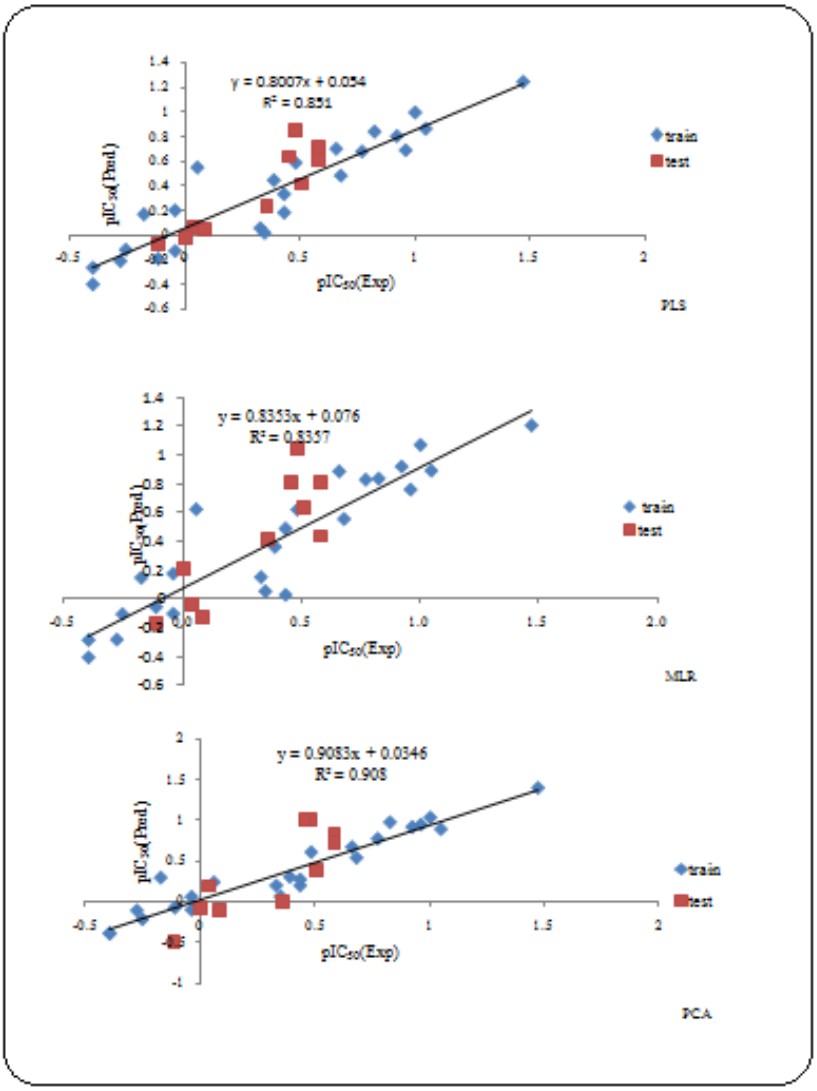


Figure 1. Correlation of predicted vs. experimental pIC_{50}

4. MATERIALS AND METHODS

The inhibitory activity of 8-hydroxy-2-iminochromene derivatives was used. The chemical structures and biological activity of these compounds are listed in Table 1. 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 QSAR analysis as dependent variables [11]. A complete geometry optimization was carried out with GAUSSIAN 03 program [31] taking the most general conformations as outset geometries. Density functional theory (DFT) calculations [32] of the structures were performed in B3LYP/6-31G level [33]. The molecular descriptors were obtained using HyperChem and GAUSSIAN Packages. Some of quantum chemical descriptors including dipole moment (DM), lowest unoccupied molecular orbital (LUMO) and highest occupied molecular orbital (HOMO) energies, hardness (η), softness (σ),entropy(ΔS) , enthalpy(ΔH), Gibbs free energy(ΔG), thermal energy(ET), zero point energy (ZPE), hardness ($\eta = \frac{1}{2(E_{LUMO}-E_{HOMO}}$) , electronic chemical potential($\mu = {}^{1}/_{2}\,(E_{HOMO} + E_{LUMO})$, global electrophilicity index ($\omega = \frac{\mu^2}{2\eta}$) and mass [34]. Some chemical parameters including molecular volume (Vol), molecular surface area (Sur), hydrophobicity (logP), polarizability, refractivity (Ref) and hydration energy (HE) were calculated using Hyperchem software.

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. Totally, 22 descriptors were created. Two programs including SPSS and Minitab were used for MLR, PCA and PLS.

Acknowledgements: We gratefully thank Payam Noor University for financial support.

Author contributions: Concept – N.M.M. ; Design – N.M.M., A.M.Z., A.H.; Supervision – N.M.M.; Resource – N.M.M.; Materials –A.H.; Data Collection and/or Processing – N.M.M., A.H.; Analysis and/or Interpretation - N.M.M., A.H.; Literature Search – N.M.M., A.H.; Writing – N.M.M., A.H.; Critical Reviews – N.M.M., A.M.Z., A.H.; Final Proof – N.M.M., A.M.Z., A.H.

Conflict of interest statement: The authors declared no conflict of interest in the manuscript.

REFERENCES

[1] Hoffmann F, Maser E. Carbonyl reductases and pluripotent hydroxysteroid dehydrogenases of the short-chain dehydrogenase/reductase superfamily. Drug Metab Rev. 2007; 39(1): 87-144.

[2] Oppermann U. Carbonyl reductases: the complex relationships of mammalian carbonyl- and quinone-reducing enzymes and their role in physiology. Annu Rev Pharmacol Toxicol.2007; 47: 293-322.

[3] Malatkova P, Maser E, Wsol V. Human carbonyl reductases. Curr Drug Metab. 2010; 11(8): 639-658.

[4] Wermuth B. Purification and properties of an NADPH-dependent carbonyl reductase from human brain. Relationship to prostaglandin 9-ketoreductase and xenobiotic ketone reductase. J BiolChem. 1981; 256(3):1206-1213.

[5] Tanaka M, Bateman R, Rauh D, Vaisberg E, Ramachandani S, Zhang C, Hansen KC,Burlingame AL, Trautman JK, Shokat KM, Adams CL. An unbiased cell morphology-based screen for new, biologically active small molecules. PLoS Biol.2005; 3(5):128-135.

[6] Zimmermann TJ, Niesen FH, Pilka ES, Knapp S, Oppermann U, Maier ME. Discovery of a potent and selective inhibitor for human carbonyl reductase 1 from propionate scanning applied to the macrolide zearalenone. Bioorg Med Chem. 2009; 17(2):530-536.

[7] Ito Y, Mitani T, Harada N, Isayama A, Tanimori S, Takenaka S, Nakano Y, Inui H, Yamaji R. Identification of carbonyl reductase 1 as a resveratrol-binding protein by affinity chromatography using 4'-amino-3,5-dihydroxy-trans-stilbene. J Nutr Sci Vitaminol. (Tokyo) 2013; 59(4): 358-364.

[8] Huang W, Ding L, Huang Q, Hu H, Liu S, Yang X, Hu X, Dang Y, Shen S, Li J, Ji X, Jiang S, Liu JO, Yu L. Carbonyl reductase 1 as a novel target of (-)-epigallocatechin gallate against hepatocellular carcinoma. Hepatology2010; 52(2):703-714.

[9] Gonzalez-Covarrubias V, Kalabus JL, Blanco JG. Inhibition of polymorphic human carbonyl reductase 1 (CBR1) by the cardioprotectant flavonoid 7-monohydroxyethyl rutoside (monoHER). Pharm Res. 2008; 25(7):1730-1734.

[10] Carlquist M, Frejd T, Gorwa-Grauslund MF. Flavonoids as inhibitors of human carbonyl reductase 1. Chem Biol Interact. 2008; 174(2):98-108.

[11] Hu D, Miyagi N, Arai Y, Oguri T, Miura T, Nishinaka T, Terada T, Gouda H, El-Kabbani O, Xia S, Toyooka N, Hara A, Matsunaga T, Ikari A, Endo S. Synthesis of 8-hydroxy-2-iminochromene derivatives as selective and potent inhibitors of human carbonyl reductase 1. Org Biomol Chem. 2015;13(27): 7487-7499.

[12] Hadjipavlou-Litina D. Review: reevaluation and new results in quantitative structure-activity studies of anticonvulsants. Med Res Rev.1998; 18(2): 91-119.

[13] Gramatica P, Papa E. QSAR modeling of bioconcentration factor by theoretical molecular descriptors. QSAR Comb Sci. 2003; 22(3):374-385.

[14] Hansch C, Kurup A, Garg R, Gao H. Chem-bioinformatics and QSAR: a review of QSAR lacking positive hydrophobic terms. Chem Rev. 2001; 101(3): 619-672.

[15] Agrawal VK, Bano S, Supuran CT, Khadikar PV. QSAR study on carbonic anhydrase inhibitors: aromatic/heterocyclic sulfonamides containing 8-quinoline-sulfonyl moieties, with topical activity as antiglaucoma agents. Eur J Med Chem.2004; 39(7): 593-600.

[16] Kumar R, Son M, Bavi R, Lee Y, Park C, Arulalapperumal V, Cao GP, Kim H, Suh J, Kim Y, Kwon YJ, Lee KW. Novel chemical scaffolds of the tumor marker AKR1B10 inhibitors discovered by 3D QSAR pharmacophore modeling. Acta Pharmacol Sin. 2015; 36(8): 998-1012.

[17] Sawant RL, Ramdin SS, Wadekar JB. Synthesis, QSAR and docking studies of 5HT2Areceptor antagonising thiazolo[3,2-a]pyrimidines as antipsychotic agents.Marmara PharmJ. 2014;18(3): 109-119.

[18] Kumar R, Malla P, Verma A, Kumar M. Design of potent human steroid 5α-reductase inhibitors: 3D-QSAR CoMFA, CoMSIA and docking studies. Med Chem Res. 2013; 22(10): 4568-4582.

[19] Chaudhari RY, Bhise SB, Yadav A, Sonawane T. QSAR, Synthesis and Docking Study of 1, 4-DHP as Novel Antitubercular Agents. J Pharm Res Clin Pract. 2016;6(1): 1-9.

[20] Vasanthanathan P, Lakshmi M, Babu MA, Kaskhedikar SG. Classical QSAR study on chromene derivatives as lanosterol 14alpha- demethylase inhibitor: a non azole antifungal target. Med Chem. 2006; 2(4):363-367.

[21] Firoozpour L, Edraki N, Nakhjiri M,Emami S, Safavi M, Ardestani SK, Khoshneviszadeh M, Shafiee A, Foroumadi A.Cytotoxic activity evaluation and QSAR study of chromene-based chalcones. Arch Pharm Res.2012;35(12):2117-2125.

[22] Consonni V, Ballabioand D, Todeschini R. Comments on the definition of the Q2 parameter for QSAR validation. J Chem Inf Model 2009; 49(7):1669-1678

[23] Consonni V, Ballabioand D, Todeschini R. Evaluation of model predictive ability by external validation techniques.J Chemometrics 2010; 24(3-4): 194-201.

[24] Lee PY, Chen CY. Toxicity and quantitative structure–activity relationships of benzoic acids to Pseudokirchneriella subcapitata. J Hazard Mater. 2009; 165(1-3): 156-161.

[25] Jing G, Zhou Z, Zhuo J. Quantitative structure-activity relationship (QSAR) study of toxicity of quaternary ammonium compounds on Chlorella pyrenoidosa and Scenedesmus quadricauda. Chemosphere 2012;86(1): 76-82.

[26] Shi LM, Fang H, Tomg W, Wu J, Perkins R, Blair RM, Branham WS, Dial SL, Moland CL, Sheenan DM. QSAR Models Using a Large Diverse Set of Estrogens. J Chem Inf Comput Sci. 2001; 41(1): 186–195.

[27] Hawkins DM. The Problem of Overfitting. J Chem Inf Comput Sci.2004; 44(1): 1–12.

[28] Consonni V, Ballabio D, Todeschini R. Evaluation of model predictive ability by external validation techniques. J Chemometrics.2010; 24(3-4):194–201.

[29] Wold S, Ericksson L. 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.

[30] VairaS, MantovaniVE, Robles JC, SanchisJC, GoicoecheaHC. Use of Chemometrics: Principal Component Analysis (PCA) and Principal Component Regression (PCR) for the Authentication of Orange Juice. J Anal Lett. 1999; 32(15): 3131-3141.

[31] Frisch MJ. Gaussian 03, Revision B, 01, Gaussian, Inc, Pittsburgh, PA. 2003.

[32] Parr R G, Pearson R G. Absolute hardness: companion parameter to absolute electronegativity. J Am Chem Soc. 1983; 105(26):7512-7516.

[33] Lee C, Yang Wand W, Parr RG. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. Phys Rev. B 1988; 37(2):785-789.

[34] Mushinski A, Nightingale MP. Many- body trial wave functions for atomic systems and ground states of small noble gas clusters. J Chem Phys. 1994; 101(5): 8831-8840.