

Research Article

QSAR Investigation on Quinolizidinyl Derivatives in Alzheimer's Disease

Ghasem Ghasemi,¹ Sattar Arshadi,² Alireza Nemati Rashtehroodi,³ Mahyar Nirouei,⁴ Shahab Shariati,¹ and Zinab Rastgoo¹

¹ Department of Chemistry, Rasht Branch, Islamic Azad University, Rasht, Iran

² Department of Chemistry, Payame Noor University, Behshahr Branch, Behshahr, Iran

³ Department of Chemistry, Payame Noor University, Sari Branch, Sari, Iran

⁴ Department of Electrical Engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran

Correspondence should be addressed to Ghasem Ghasemi; ghasemi@iaurasht.ac.ir

Received 15 December 2012; Revised 24 March 2013; Accepted 7 April 2013

Academic Editor: Hon Keung Tony Ng

Copyright © 2013 Ghasem Ghasemi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sets of quinolizidinyl derivatives of bi- and tri-cyclic (hetero) aromatic systems were studied as selective inhibitors. On the pattern, quantitative structure-activity relationship (QSAR) study has been done on quinolizidinyl derivatives as potent inhibitors of acetylcholinesterase in alzheimer's disease (AD). Multiple linear regression (MLR), partial least squares (PLSs), principal component regression (PCR), and least absolute shrinkage and selection operator (LASSO) were used to create QSAR models. Geometry optimization of compounds was carried out by B3LYP method employing 6–31 G basis set. HyperChem, Gaussian 98 W, and Dragon software programs were used for geometry optimization of the molecules and calculation of the quantum chemical descriptors. Finally, Unscrambler program was used for the analysis of data. In the present study, the root mean square error of the calibration and R^2 using MLR method were obtained as 0.1434 and 0.95, respectively. Also, the *R* and R^2 values were obtained as 0.79, 0.62 from stepwise MLR model. The R^2 and mean square values using LASSO method were obtained as 0.766 and 3.226, respectively. The root mean square error of the calibration and R^2 using PLS method were obtained as 0.3726 and 0.62, respectively. According to the obtained results, it was found that MLR model is the most favorable method in comparison with other statistical methods and is suitable for use in QSAR models.

1. Introduction

Alzheimer's disease (AD) is a debilitating illness with unmet medical needs [1]. The number of people afflicted with the disease worldwide is expected to be triple up to the year 2050 [2]. The multifactorial pathogenesis of AD includes accumulation of aggregates of β -amyloid (A β) and tau protein and loss of cholinergic neurons with consequent deficit of the neurotransmitter acetylcholine (ACh) [3, 4]. In advancing AD, AChE levels in the brain are declining [5].

The well-known theory of the quantitative structureactivity relationships (QSARs) [6–8] is based on the hypothesis that the biological activity of a chemical compound is mainly determined by its molecular structure [6]. QSAR attempts to find consistent relationship between biological activity and molecular properties, so that these "rules" can be used to predict the activity of new compounds from their structures.

Today, QSARs are being applied in many disciplines with much emphasis on drug design. Over the years of development, many methods, algorithms, and techniques have been discovered and applied in QSAR studies [9, 10]. To date, QSARs are among the important applications of chemometric tools with the objective of development of predictive models which can be used in different areas of chemistry including medicinal, agricultural, environmental, and materials [11–13].

Drug discovery often involves the use of QSAR to identify chemical structures that could have good inhibitory effects on specific targets [15]. The aim of QSAR analysis is to investigate

General structure	Х	Y	R	R 🗆	Nr
	S	N	CH ₃ CH CH ₂ N(Et) ₂	Н	1
	S	Ν	$-(H_2C)_3 - N(Et)_2$	CF ₃	2
	S	Ν	$-(H_2C)_3-N$ $-OH$	CN	3
	S	Ν	$ \begin{array}{c} H \\ \underline{C} \\ \underline{H}_{2} \\ \underline{V}_{N} \\ V$	Н	4
	S	Ν	$ \begin{array}{c} H \stackrel{\underline{C}H_2}{\longleftarrow} \\ N \stackrel{\underline{C}H_2}{\longleftarrow} \\ N \stackrel{\underline{C}H_2}{\longleftarrow} \\ \end{array} $	CF ₃	5
	Ο	Ν	$ \begin{array}{c} H \stackrel{\underline{C}H_2}{\longleftarrow} \\ N \\ \end{array} $	Н	6
X	CH ₂	Ν	$\underbrace{\begin{array}{c} H_{2} \\ H_{1} \\ N \end{array}}_{N}$	Н	7
Y R'	H ₂ C-CH ₂	Ν	H CH2 N	Н	8
	HC=CH	Ν	H CH ₂	Н	9
	S	СН	H ₂ C N _{CH3}	Н	10
	S	СН	$ \begin{array}{c} H \stackrel{\underline{C}H_2}{\longleftarrow} \\ N \\ \end{array} $	Н	11
	S	С-ОН	$ \begin{array}{c} H \stackrel{\underline{C}H_{2}}{\longleftarrow} \\ N \stackrel{\underline{C}H_{2}}{\longleftarrow} \\ N \stackrel{\underline{C}H_{2}}{\longleftarrow} \\ N \stackrel{\underline{C}H_{2}}{\longleftarrow} \\ N \stackrel{\underline{C}H_{2}}{\longrightarrow} \\ \end{array} $	Н	12
	S	С		Н	13
	H ₂ C-CH ₂	С		Н	14

TABLE 1: Structures of quinolizidinyl derivatives of bi- and tricyclic systems used for QSAR model building [14].

Journal of Computational Medicine

		TABLE I,	Continued.		
General structure	Х	Y	R	R 🗆	Nr
	HC=CH	С		Н	15
		Ν	$ \begin{array}{c} H \stackrel{CH_2}{\stackrel{\cdot}{\overset{\cdot}}} \\ N \stackrel{\cdot}{\overset{\cdot}} \end{array} $		16
Y R		СН	$\underbrace{\bigvee_{N}^{H \stackrel{\underline{C}H_{2}}{\cdot}}}_{N}$		17
		СН	H H H H H H H H H H		18
		Ν	$\underbrace{\bigvee_{N}^{H \stackrel{\underline{C}H_{2}}{\cdot}}}_{N}$		19
Y R R		СН	$ \begin{array}{c} H \stackrel{CH_2}{\stackrel{\cdot}{\overset{\cdot}}} \\ N {\overset{\prime}} \\ \end{array} $		20
		С	$\underbrace{\bigvee_{N}}_{c}^{H}$		21
		СН			22
	S	СН	CH ₃ CH N(Et) ₂	Н	23
	S	СН	$ \begin{array}{c} H \stackrel{\underline{C}H_2}{\stackrel{\cdot}{\cdot}} \\ N \\ \end{array} $	Н	24
$ \begin{array}{c} \overbrace{V}^{X} \\ Y \\ O \\ R \end{array} $	S	СН		Н	25
	S	СН	H = S CH ₂	Н	26
	S	СН	$ \begin{array}{c} H = S CH_2 \\ \hline N \\ \end{array} $	OCH ₃	27

TABLE 1: Continued.

		TABLE 1	: Continued.		
General structure	Х	Y	R	R 🗆	Nr
	S	СН	$ \begin{array}{c} H = S \\ H = CH \\ CH_{3} \end{array} $	Н	28
	S	СН	$ \begin{array}{c} H = S \\ (CH_2)_2 \\ N \\ N \end{array} $	Н	29
	S	СН	$ \begin{array}{c} H = \\ N \\ N \end{array}^{N} $ (CH ₂) ₃	Н	30
	S	СН	$ \begin{array}{c} H = S \\ (CH_2)_4 \\ N \\ N \end{array} $	Н	31
	H ₂ C-CH ₂	СН	H = S CH ₂	Н	32
	HN-CO	Ν	$\underbrace{\bigvee_{N}^{H \stackrel{CH_2}{\cdot}}}_{N}$	Н	33
	HN-CO	Ν	H = S CH ₂	Н	34
		NH	CH ₃ CH (CH ₂) ₃ –N(Et) ₂		35
, N. A. Cl		NH	$ \begin{array}{c} H \stackrel{CH_2}{\cdot} \\ N \\ \end{array} $		36
VO VI VI R		NH	$ \begin{array}{c} H = S \\ (CH_2)_2 \\ V \\ N \\ \end{array} $		37
		NH	H = S(CH ₂) ₃		38
		S	$\underbrace{\bigvee_{N}^{H \stackrel{\underline{C}H_{2}}{\cdot}}}_{N}$		39

4



TABLE 2:	The mean	of selected	descriptors.
----------	----------	-------------	--------------

Descriptor symbol	Descriptor group	Meaning
G (N⋯O)	Geometrical descriptors	Sum of geometrical distances between N····O
ARR	Constitutional descriptors	Aromatic ratio
Te	WHIM descriptors	T total size index/weighted by atomic Sanderson electronegativities
MATS6e	2D autocorrelations	Moran autocorrelation—lag 6/weighted by atomic Sanderson electronegativities
Mor31m	3D-MoRSE descriptors	3D-MoRSE— signal 31/weighted by atomic masses
Mor18m	3D-MoRSE descriptors	3D-MoRSE—signal 18/weighted by atomic masses

TABLE 3: The statistical parameters of different constructed QSAR models.

Mathad	RM	ISE	R^2		
Method	Calibration	Prediction	Calibration	Prediction	
PLS	0.372616	0.466533	0.624241	0.426009	
PCR	0.372537	0.484057	0.624401	0.407646	
LASSO	_	_	0.7	66	

the correlation between activity, generally, biological activity, and the physicochemical properties of a set of molecules [16].

PLS regression technique is especially useful in quite common case where the number of descriptors (independent variables) is comparable to or greater than the number of compounds (data points), and/or there exist other factors leading to correlations between variables. In this case, the solution of classical least squares problem does not exist or is unstable and unreliable. On the other hand, PLS approach leads to stable, correct, and highly predictive models even for correlated descriptors [17].

PCR is a combination of principal component analysis (PCA) and MLR. The first step in PCR is to decompose a spectral data matrix using PCA. Generally, there are two types of decomposition techniques. The first technique is by computing eigenvectors and eigenvalues. We used singular value decomposition (SVD) to decompose the spectral data matrix. This is because SVD is generally accepted as the most stable and numerically accurate technique [18, 19].

LASSO translates each coefficient by a constant factor truncating at zero. This is called soft thresholding. Best subset selection drops all variables with coefficients smaller than the $M_{\rm th}$ largest. This is a form of hard thresholding.

2. Computational Details

The 3D structures of the molecules were drawn using the built optimum option of Hyperchem software (version 8.0). Then, the structures were fully optimized based on the ab initio method, using DFT level of theory. Hyperchem (version 3.0) and Dragon (version 3.0) programs were employed to calculate the molecular descriptors. All calculations were performed using Gaussian 98 W program series. Geometry optimization of compounds was carried out by B3LYP method employing 6–31 G basis set [20].

In this study, the independent variables were molecular descriptors, and the dependent variables were the actual half maximal inhibitory concentration (IC_{50}) values. More than 1498 theoretical descriptors were selected and calculated. These descriptors can be classified into several groups including: (i) constitutional, (ii) topological, (iii) molecular walk counts, (iv) BCUT, (v) Galvez topological charge indices, (vi) autocorrelations, (vii) charge, (viii) aromaticity indices, (ix) randic molecular profiles, (x) geometrical, (xi) RDF, (xii) MoRSE, (xiii) WHIM, (xiv) GETAWAY, (xv) functional groups, (xvi) atom-centred, (xvii) empirical, and (xviii) properties descriptors. Finally, Unscrambler (version 9.7) program was used for analysis of data and statistical calculation.

Molecule	G (N · · · O)	ARR	Te	MATS6e	Mor31m	Mor18m
1	0.000	0.500	13.546	0.045	-0.099	0.381
2	0.000	0.429	17.814	0.106	0.018	1.122
3	20.250	0.448	18.480	0.007	-0.294	1.186
4	0.000	0.414	15.923	0.007	-0.181	1.383
5	0.000	0.364	17.411	0.090	0.014	1.615
6	9.600	0.414	15.799	0.011	-0.113	0.104
7	0.000	0.414	16.071	0.022	-0.084	0.469
8	0.000	0.400	16.312	0.034	0.002	1.199
9	0.000	0.400	16.108	0.019	-0.191	1.233
10	0.000	0.480	14.345	0.088	0.225	1.547
11	0.000	0.414	15.271	0.119	0.106	2.132
12	4.490	0.400	15.838	-0.029	0.055	1.896
13	0.000	0.414	16.287	0.090	0.085	2.410
14	0.000	0.400	16.628	0.107	0.291	1.618
15	0.000	0.400	16.514	0.098	0.091	1.494
16	0.000	0.464	15.920	0.012	-0.039	0.330
17	0.000	0.429	14.869	0.129	0.083	1.143
18	0.000	0.429	14.841	0.129	0.083	0.946
19	0.000	0.444	13.992	0.003	-0.167	1.393
20	0.000	0.444	15.531	0.110	0.085	0.358
21	0.000	0.444	15.794	0.086	0.151	0.396
22	0.000	0.429	16.962	-0.002	0.004	0.488
23	2.850	0.480	14.983	-0.219	-0.126	1.175
24	3.590	0.387	14.711	-0.111	0.050	1.531
25	4.490	0.387	15.677	-0.111	0.210	1.747
26	6.090	0.364	19.718	-0.059	0.174	1.143
27	19.660	0.343	22.226	0.086	0.082	1.374
28	5.930	0.353	20.024	-0.050	-0.035	1.639
29	6.420	0.353	25.383	-0.101	0.140	1.918
30	7.620	0.343	28.305	-0.041	-0.007	1.792
31	9.500	0.333	36.586	-0.050	0.071	2.232
32	6.150	0.353	20.780	-0.113	-0.036	0.525
33	25.410	0.364	14.812	-0.051	-0.164	0.240
34	34.830	0.343	25.376	0.009	-0.292	0.729
35	19.340	0.400	24.194	0.080	-0.158	0.643
36	17.760	0.394	17.479	0.071	-0.267	1.192
37	19.490	0.333	26.550	0.079	-0.329	1.476
38	19.170	0.324	29.353	0.066	-0.033	1.304
39	12.740	0.333	19.215	0.087	0.048	1.531
40	43.630	0.200	17.259	-0.009	0.062	0.817
41	29.170	0.200	23.440	0.002	0.058	0.737
42	24.170	0.194	20.222	0.024	-0.122	0.665

For each compound in the training sets, the correlation equation was derived with the same descriptors. Then, the obtained equation was used to predict log $(1/IC_{50})$ values for the compounds from the corresponding test sets. In the present work, the method of stepwise multiple linear regression (stepwise MLR) was used in order to select the most appropriate descriptor of all descriptors. Totally, 1498 descriptors were generated. In this study, two programs

including SPSS (version 19) and Unscrambler were used for MLR, PLS, PCR, and LASSO.

3. Results and Discussions

The structures of the quinolizidinyl derivatives used in this study were shown in Table 1. Since, the variation in

TABLE 5: Experimental and predicted values of log $(1/IC_{50})$ using PCR and PLS methods.

Observed log (1/IC50)	Predicted PCR	Predicted PLS
1.531	1.534	1.426
1.653	1.656	1.656
0.854	1.340	1.327
1.591	1.679	1.675
1.771	1.713	1.893
1.568	1.647	1.661
1.699	1.657	1.660
1.74	1.482	1.401
0.919	1.429	1.351
1.634	1.572	1.489
0.845	1.684	1.677
1.623	2.260	2.324
1.763	1.630	1.625
1.663	1.423	1.341
0.919	1.282	1.203
1.613	1.560	1.602
1.653	1.574	1.614
1.681	1.574	1.614
0.949	1.047	1.080
0.826	1.042	1.073
1.544	1.068	1.100
1.653	0.919	0.956
1.653	1.767	1.701
1.69	1.693	1.726
1.477	1.689	1.722
1.505	1.358	1.370
1.672	1.212	1.275
1.602	1.467	1.503
1.672	1.160	1.153
0.833	0.962	0.919
0.756	0.733	0.630
1.532	1.161	1.076
1.623	1.579	1.584
1.462	1.193	1.171
1.69	1.491	1.387
0.863	0.521	0.581
-0.076	0.275	0.276
-0.658	0.022	-0.026
1.756	1.010	1.062
0.82	0.417	0.497
-0.456	0.431	0.495
0.079	0.273	0.337

the chemical structure of the considered compounds is low, the selection of chemical descriptors, which can encode small variations between structures of molecules in data set, is very important. In this way, GETAWAY descriptors are very informative 3D descriptors that can encode structural features of molecules. The four most significant descriptors which were selected are as follows [14, 20]:

G (N···O), ARR, Te, MATS6e, Mor31m, and Mor18m.

The mean values of selected descriptors are shown in Table 2. As can be seen from this table, atomic masses and electronegativities were important descriptors in our study.

The selected descriptors through these methods were used to construct some linear models using PCR and PLS methods. Statistical parameters of different constructed QSAR models are shown in Table 3. R^2 and RMSE values for calibration in MLR method are better than the two other methods. In the present study, the root mean square error of the calibration and R^2 using MLR method were obtained as 0.1434 and 0.95, respectively.

Considering the experimental error, the overall prediction of the log $(1/IC_{50})$ values was quite satisfactory. The results of MLR method were much better than the two other methods.

In the present study, linear variable selection methods were used to select the most significant descriptors (stepwise MLR) (Table 4).

The performance of the QSAR model to predict log (IC₅₀) value was also estimated using the internal cross-validation method. The resulted predictions of the log ($1/IC_{50}$) using PLS and PCR methods in gas phase were given in Table 5.

4. Conclusion

In our study, the linear methods were used to select the most significant descriptors. The stepwise MLR, MLR, PLS, and PCR were used to construct a quantitative relation between the activities of quinolizidinyl derivatives and their calculated descriptors. MLR has been successfully used for finding a QSAR model for quinolizidinyl derivatives. It provides the best results in comparison with other studied methods. Our present attempt to correlate the log ($1/IC_{50}$) with theoretically calculated molecular descriptors has led to a relatively successful QSAR model that relates these derivatives. The results obtained from stepwise MLR method were suitable for drug design and classification.

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgment

The authors thank the Research vice Presidency of Islamic Azad University, Rasht Branch, for their encouragement, permission, and financial support.

References

 D. Selkoe, "Alzheimer's disease: genes, proteins, and therapy," *Physiological Reviews*, vol. 81, pp. 741–766, 2001.

Journal of Computational Medicine

- [2] H. H. Griffiths, I. J. Morten, and N. M. Hooper, "Emerging and potential therapies for Alzheimer's disease," *Expert Opinion on Therapeutic Targets*, vol. 12, no. 6, pp. 693–704, 2008.
- [3] J. R. Roland and H. Jacobsen, "Alzheimer's disease: from pathology to therapeutic approaches," *Angewandte Chemie*, vol. 48, no. 17, pp. 3030–3059, 2009.
- [4] A. Gella and N. Durany, "Oxidative stress in Alzheimer disease," *Cell Adhesion and Migration*, vol. 3, no. 1, pp. 88–93, 2009.
- [5] M. Mesulam, A. Guillozet, P. Shaw, and B. Quinn, "Widely spread butyrylcholinesterase can hydrolyze acetylcholine in the normal and Alzheimer brain," *Neurobiology of Disease*, vol. 9, no. 1, pp. 88–93, 2002.
- [6] C. Hansch and A. Leo, Exploring QSAR. Fundamentals and Applications in Chemistry and Biology, American Chemical Society, Washington, DC, USA, 1995.
- [7] H. Kubinyi, QSAR: Hansch Analysis and Related Approaches, Wiley-Interscience, New York, NY, USA, 2008.
- [8] T. Puzyn, J. Leszczynski, and M. T. Cronin, *Recent Advances in QSAR Studies: Methods and Applications*, Springer, New York, NY, USA, 1st edition, 2009.
- [9] L. He and P. C. Jurs, "Assessing the reliability of a QSAR model's predictions," *Journal of Molecular Graphics and Modelling*, vol. 23, no. 6, pp. 503–523, 2005.
- [10] D. V. Eldred, C. L. Weikel, P. C. Jurs, and K. L. E. Kaiser, "Prediction of fathead minnow acute toxicity of organic compounds from molecular structure," *Chemical Research in Toxicology*, vol. 12, no. 7, pp. 670–678, 1999.
- [11] Q. S. Du, P. G. Mezey, and K. C. Chou, "Heuristic molecular lipophilicity potential (HMLP): a 2D-QSAR study to LADH of molecular family pyrazole and derivatives," *Journal of Computational Chemistry*, vol. 26, no. 5, pp. 461–470, 2005.
- [12] Q. S. Du, R. B. Huang, Y. T. Wei, L. Q. Du, and K. C. Chou, "Multiple field three dimensional quantitative structureactivity relationship (MF-3D-QSAR)," *Journal of Computational Chemistry*, vol. 29, no. 2, pp. 211–219, 2008.
- [13] Q. S. Du, R. B. Huang, and K. C. Chou, "Recent advances in QSAR and their applications in predicting the activities of chemical molecules, peptides and proteins for drug design," *Current Protein & Peptide Science*, vol. 9, no. 3, pp. 248–259, 2008.
- [14] B. Tasso, M. Catto, O. Nicolotti et al., "Quinolizidinyl derivatives of bi- and tricyclic systems as potent inhibitors of acetyl- and butyrylcholinesterase with potential in Alzheimer's disease," *European Journal of Medicinal Chemistry*, vol. 46, no. 6, pp. 2170–2184, 2011.
- [15] E. K. Freyhult, K. Andersson, and M. G. Gustafsson, "Structural modeling extends QSAR analysis of antibody-lysozyme interactions to 3D-QSAR," *Biophysical Journal*, vol. 84, pp. 2264–2272, 2003.
- [16] C. Karthikeyan, N. S. H. Moorthy N, and P. Trivedi, "QSAR study of substituted 2-pyridinyl guanidines as selective urokinase-type plasminogen activator (uPA) inhibitors," *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 24, pp. 6–13, 2009.
- [17] A. Höskuldsson, "PLS regression methods," *Journal of Chemometrics*, vol. 2, no. 3, pp. 211–228, 1988.
- [18] P. J. Gemperline, "Principal component analysis," in *Practical Guide to Chemometrics*, pp. 69–104, CRC Press, 2nd edition, 2006.
- [19] D. J. Livingstone and D. W. Salt, "Variable Selection—Spoilt for Choice?" *Reviews in Computational Chemistry*, vol. 21, pp. 287– 348, 2005.

[20] G. Ghasemi, M. Nirouei, S. Shariati, P. Abdolmaleki, and Z. Rastgoo, Arabian Journal of Chemistry. In press.



The Scientific World Journal



Gastroenterology Research and Practice





Journal of Diabetes Research



Disease Markers



Immunology Research





International Journal of Endocrinology



BioMed **Research International**





Computational and Mathematical Methods in Medicine





Behavioural Neurology



Complementary and Alternative Medicine











Oxidative Medicine and Cellular Longevity