Regular Article

ANALYTICAL
AND BIOANALYTICAL
CHEMISTRY
RESEARCH

Published by the
Iranian Chemical Society

Anal. Bioanal. Chem. Res., *Vol. 3, No. 2, 253-263, December 2016.*

The Correlation of Biological Activity and Chemical Structure of Quinolizidinyl Derivatives as Inhibitor of Alzheimer's Disease with Linear and Non-linear Models

G. Ghasemi[a,]* and A. Nemati Rashtehroodi[b]

[a]*Department of Chemistry, Rasht Branch, Islamic Azad University, Rasht, Iran*

[b]*Department of Chemistry, Payame Noor University, Sari Branch, Sari*

(Received 9 June 2016, Accepted 21 August 2016)

In this research, QSAR study has been carried out on quinolizidinyl derivatives as potent inhibitors of Acetyl and butyrylcholin esterase in Alzheimer's disease. Despite significant research efforts in industry and academia, there are currently no diseases modifying therapies available to treat this illness. Significant evidence suggests that the pathology of AD is linked to generation of β-amyloid peptides (Aβ) through proteolytic processing of amyloid precursor protein (APP).

Genetic algorithm (GA), Jack-Knife and stepwise multiple linear regressions (stepwise-MLR) were used to create non-linear and linear QSAR models. The root-mean square errors of the training set and the validation set for GA-ANN model using Jack-Knife method, were 0.1406, 0.2165 and R^2 was 0.90. Also, the R and R^2 values in the gas phase were obtained as 0.88 and 0.78 from GA-stepwise MLR model, respectively. Also, we suggest that compounds No. 4, 6, 10, 14, 24, 26 and 34 have the most appropriate structure for the design of drugs to pharmacists. Electronegativities, atomic polarizability and atomic van der Waals volumes were important descriptors in our study. Geometry optimization of compounds was carried out using the B3LYP method employing a 6-31G (d) basis set.

Keywords: QSAR model, Genetic Algorithm, Artificial Neural Network, Alzheimer's disease, Quinolizidinyl derivatives

INTRODUCTION

Alzheimer's disease (AD) has become the sixth leading cause of death in the United States and is the most common form of dementia in the elderly [1,2].

The inhibition of acetylcholinesterase (AChE), that is responsible for the breakdown of acetylcholine (Ach), has proven as a successful approach to relieve some cognitive and behavioral symptoms of AD [3,4]. In advanced AD, AChE levels in the brain are declining, but a progressive increase (up to 90%) of butyryl cholinesterase (BChE) is observed, which is able, even at lower rate, to hydrolyze Ach [5,6].

In computational chemistry, prediction of biological activity of compounds based on QSAR studies substantially increases the potential of work, avoiding time and consuming experiments. QSARs are among the important applications of chemometric tools with an objective of development of predictive models which can be used in different areas of chemistry including medicine, agriculture, environment, materials, *etc.* [7]. QSARs describe correlations between various physico-chemical properties of a chemical (usually known referred to as descriptors of molecular structures of chemicals) and their observed or predicted biological activities. QSARs generally assume a common mechanism behind the biological activity of a structurally/functionally related set of chemicals [8].

Molecular descriptors are numerical values obtained by the quantification of various structural and physicochemical characteristics of the molecule. It is envisaged that molecular descriptors quantify these attributes so as to determine the behavior of the molecule and the way the molecule interacts with a physiological system. Since the exact mechanism of drug activity is unknown in many

*Corresponding author. E-mail: ghasemi@iaurasht.ac.ir

cases, it is desirable to start with descriptors spanning as many attributes of the molecules as possible and then assess their ability to predict the desired activity/property.

THEORY

The chemical structure is represented by numerical entities called molecular descriptors, which are used to describe different characteristics of a certain structure to obtain yield information about the activity being studied [9]. We can also develop a kind of computer program with network topology called artificial neural networks (ANN) that after adequate training learns to predict target proteins for a given drug. It means, ANNs are network-like software that may use as inputs of topological indices and/or physico chemical parameters calculated in the previous steps to predict which molecular structures or network-like structures, present a desire property or not [10,11]. The ANN are known as a good method in expressing highly non-linear relationship between the input and output variables, hence, greater interests were attracted in applying them to the pattern classification of complex compounds. A genetic algorithm maintains a population of candidate solutions for the problem at hand, and makes it evolve by iteratively applying a set of stochastic operators. GAs are stochastic optimization methods that provide powerful means to perform directed random searches in a large problem space as encountered in chemometrics and drug design [12,13].

The properties that can be obtained experimentally, *e.g.*, biological activity or toxicity are another expression of the molecular properties. These observed properties are related back to the intrinsic properties to predict the behavior of a molecule from its structure and physico-chemical properties. Construction of a quantitative/qualitative model that describes this relationship is the main goal of any quantitative/qualitative structure-activity relationship (QSAR) study. In this context, the chemometric calibration techniques are highly valued. Specifically principal component regression (PCR) and partial least squares regression (PLS) have become the usual methods of choice where a large number of descriptors are used. In addition to the quantitative issues, one is interested in identifying groups of molecules with similar properties as quantified by a set of molecular descriptors or by a certain observed property (*e.g.*, biological activity, toxicity). Cluster analysis and principal component analysis have proven to be excellent methods for the exploration and visualization of the huge numbers of descriptor data generated, whereas with classification and/or discriminant methods one can create logic rules for the classification of molecules.

The PLS method uses this data matrix to generate a QSAR model. PLS extracts principal component-like vectors (latent variables) from the matrices of independent and dependent variables. This method takes a matrix containing a large number of potentially useful structural descriptors, which can be highly inter-correlated, and offers a correlation using the latent variables. The optimum number of latent variables is determined by cross-validation [14-17]. The aim of this study is to assess QSAR models reliability, using GA-ANN methods for prediction of new anti-Alzheimer compounds.

The statistical model for multiple-regression is a extension of that for simple linear regression. The response variable, denoted by Y, is measured along with a set of predictor variables, denoted by $X_1, X_2,\ldots;X_p$ where *p* is the number of predictor variables. The formal statistical model is:

$$Y_i = \beta_0 + \beta_1 Xi_1 + \beta_2 Xi_2 + \ldots + \beta_p X_{ip} + \varepsilon_i \qquad (1)$$

where the unknown parameters are the set of β´s. The deviation between the observed value of Y and the predicted value from the regression equation, ε_i is distributed as a Normal distribution with a mean of 0 and an (unknown) variance of σ_2. This is often written using a short hand notation in many statistical packages as: $Y = X_1, X_2,\ldots X_p$ where the intercept (β_0) and the residual variation (ε) are implicit. This can also be written using matrices as

$$Y = X\,\beta + \varepsilon \qquad (2)$$

where Y is an $n \times 1$ column vector, X is an $n \times (p + 1)$ matrix [don't forget the intercept column] of the predictors, β is a $(p + 1) \times 1$ column vector (the intercept β_0, plus the p "slopes" $\beta_1;\ldots;\beta_p$), and ε is a $n \times 1$ vector of residuals that has a multivariate normal distribution with a mean of 0 and a covariance matrix of $I\sigma^2$ where *I* is the identity matrix.

The SVD (Singular Value Decomposition) algorithm is the most widely used algorithm to compute the estimated regression coefficients for MLR.

PLS has received a great amount of attention in the field of chemometrics. The algorithm has become a standard tool for processing a wide spectrum of chemical data problems. The success of PLS in chemometrics resulted in a lot of applications in other scientific areas including bioinformatics, food research, medicine, pharmacology, social sciences, physiology-to name but a few [18-21].

This part introduces the main concepts of PLS and provides an overview of its application to different data analysis problems. Our aim is to present a concise introduction, that is, a valuable guide for anyone who is concerned with data analysis.

Consider the general setting of a linear PLS algorithm to model the relation between two data sets (blocks of variables). Denote by $X \subset R^N$ an N-dimensional space of variables representing the first block and similarly by $Y \subset R^M$ a space representing the second block of variables. PLS models the relations between these two blocks by means of score vectors. After observing n data samples from each block of variables, PLS decomposes the (n × N) matrix of zero-mean variables X and the (n × M) matrix of zero-mean variables Y into the form

$$X = T P^T + E$$

$$Y = UQ^T + F \qquad (3)$$

where the T, U are (n × p) matrices of the p extracted score vectors (components, latent vectors), the (N × p) matrix P and the (M × p) matrix Q represent matrices of loadings and the (n × N) matrix E and the (n × M) matrix F are the matrices of residuals. The PLS method, which in its classical form is based on the nonlinear iterative partial least squares (NIPALS) algorithm, finds weight vectors w, c such that

$$[\mathrm{cov}(t,u)]^2 = [\mathrm{cov}(Xw,Yc)]^2 = \max_{|r|=|s|=1}[\mathrm{cov}(Xr, Ys)]^2 \qquad (4)$$

where $\mathrm{cov}(t,u) = t^T u/n$ denotes the sample covariance between the score vectors t and u. The NIPALS algorithm starts with random initialisation of the Y-space score vector u and repeats a sequence of the following steps until convergence.

1) $w = X^T u/(u^T u)$

2) $\|w\| \rightarrow 1$

3) $t = X w$

4) $c = Y^T t/(t^T t)$

5) $\|c\| \rightarrow 1$

6) $u = Y c$

Note that u = y if M = 1, that is, Y is a one-dimensional vector that we denote by y. In this case the NIPALS procedure converges in a single iteration. It can be shown that the weight vector w also corresponds to the first eigenvector of the following eigenvalue problem

$$X^T Y Y^T X w = \lambda w \qquad (5)$$

The X- and Y-space score vectors t and u are then given as

$$t = X w \text{ and } u = Y c \qquad (6)$$

where the weight vector c is define in steps 4 and 5 of NIPALS. Similarly, eigenvalue problems for the extraction of t, u or c estimates can be derived [22].

The root mean square error is calculated for the prediction or validation samples (RMSEP) and for the calibration samples (RMSEC).

RMSEC (all validation methods):

$$RMSEC = \frac{1}{yWeight}\sqrt{\mathrm{Re}\,sYCalVar} \qquad (7)$$

RMSEP (leverage correction and test set validation):

$$RMSEP = \frac{1}{yWeight}\sqrt{\mathrm{Re}\,sYCalVar} \qquad (8)$$

RMSEP (cross validation):

Table 1. Structures of Quinolizidinyl Derivatives [26] of bi- and Tricyclic Systems Used for QSAR Model Building

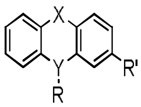
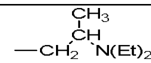
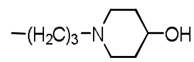
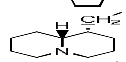
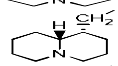
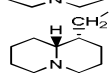
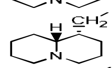
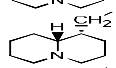
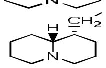
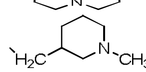
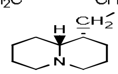
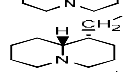
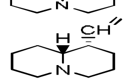
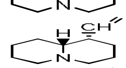
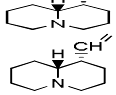
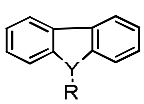
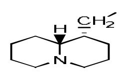
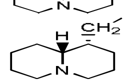
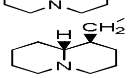
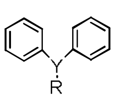
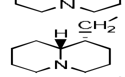
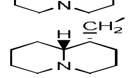
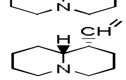
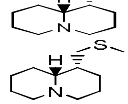
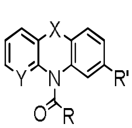
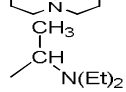
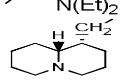
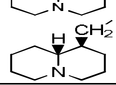
General structure	X	Y	R	R′	Nr
	S	N		H	1
	S	N	$-(H_2C)_3-N(Et)_2$	CF_3	2
	S	N		CN	3
	S	N		H	4
	S	N		CF_3	5
	O	N		H	6
	CH_2	N		H	7
	H_2C-CH_2	N		H	8
	$HC{=}CH$	N		H	9
	S	CH		H	10
	S	CH		H	11
	S	C−OH		H	12
	S	C		H	13
	H_2C-CH_2	C		H	14
	$HC{=}CH$	C		H	15
		N			16
		CH			17
		CH			18
		N			19
		CH			20
		C			21
		CH			22
	S	CH		H	23
	S	CH		H	24
	S	CH		H	25

Table 1. Continued

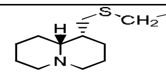
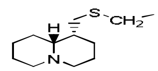
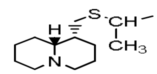
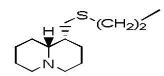
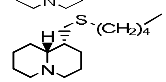
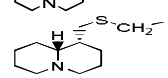
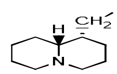
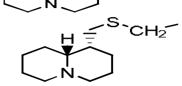
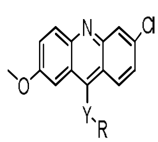
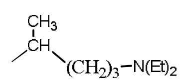
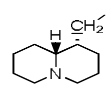
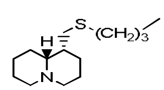
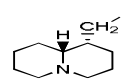
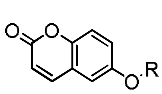
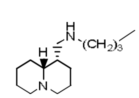
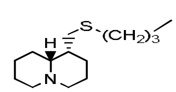
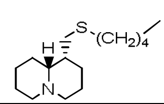
	S	CH		H	26
	S	CH		OCH_3	27
	S	CH		H	28
	S	CH		H	29
	S	CH		H	30
	S	CH		H	31
	H_2C-CH_2	CH		H	32
	HN—CO	N		H	33
	HN—CO	N		H	34
		NH			35
		NH			36
		NH			37
		NH			38
		S			39
					40
					41
					42

Table 2. The Models and Descriptors Selected by GA-Stepwise MLR Method

Model	R	R Square
1	0.638[a]	0.406
2	0.710[b]	0.504
3	0.782[c]	0.612
4	0.813[d]	0.662
5	0.837[e]	0.701
6	0.866[f]	0.750
7	0.883[g]	0.780

[a]Predictors: (Constant), HOMT. [b]Predictors: (Constant), HOMT, D/Dr06. [c]Predictors: (Constant), HOMT, D/Dr06, Mor28p. [d]Predictors: (Constant), HOMT, D/Dr06, Mor28p, R2e. [e]Predictors: (Constant), HOMT, D/Dr06, Mor28p, R2e, R8u+. [f]Predictors: (Constant), HOMT, D/Dr06, Mor28p, R2e, R8u+, nDB. [g]Predictors: (Constant), HOMT, D/Dr06, Mor28p, R2e, R8u+, nDB, R7v+. [h]Dependent Variable: Target.

$$RMSEP = \sqrt{\frac{1}{ltot}\sum_{S=1}^{Nseg}\frac{1}{yweights2}\sum_{i=1}^{ls}Fiys(i,j)} \quad (9)$$

MOLECULAR MODELING AND DESCRIPTORS GENERATION

The 3D structures of the investigated molecules were generated using the built optimum option of HyperChem software (version 6.0) and were shown in Table 1. Dragon program (version 3.0) was employed to calculate the molecular descriptors. Dragon software has been conceived to provide the user with a variety of molecular descriptors derived from different molecular representations, allowing the user to choose those molecular descriptors which are more suitable for his/her specific research.

All calculations were performed using Gaussian 03W program series. Geometry optimization of compounds was carried out by B3LYP method employing 6-31G basis set [23-25].

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.

STATISTICAL METHODS

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 step wise multiple linear regression (stepwise MLR) was used to select the most appropriate descriptor of all investigated descriptors. Totally 1498 descriptors were generated that were too many to be

Table 3. Descriptors Values by GA-Stepwise MLR Model for Correlation 0.3 to up

Molecule	HOMT	D/Dr06	Mor28p	R2e	R8u+	nDB	R7v+
1	11.524	127.462	-0.317	1.948	0.023	0.000	0.018
2	11.560	164.097	-0.122	2.213	0.013	0.000	0.010
3	11.945	236.297	-0.187	1.952	0.015	0.000	0.012
4	11.524	227.810	-0.365	2.059	0.017	0.000	0.013
5	11.560	280.777	-0.388	2.333	0.015	0.000	0.012
6	11.524	227.810	-0.238	2.093	0.020	0.000	0.014
7	11.436	227.810	-0.311	2.085	0.020	0.000	0.013
8	11.333	187.487	-0.139	2.101	0.019	0.000	0.012
9	11.142	187.487	-0.011	2.037	0.016	1.000	0.012
10	11.524	174.840	-0.273	1.981	0.021	0.000	0.012
11	11.699	227.810	-0.172	2.031	0.018	0.000	0.011
12	11.596	240.040	-0.386	2.079	0.019	0.000	0.012
13	11.436	227.810	-0.362	2.079	0.017	1.000	0.012
14	11.400	187.487	-0.394	2.109	0.016	1.000	0.012
15	11.070	187.487	-0.161	2.051	0.017	2.000	0.012
16	11.080	178.739	-0.250	2.100	0.021	0.000	0.014
17	11.452	178.740	0.029	2.063	0.020	0.000	0.010
18	11.452	178.740	-0.022	2.070	0.022	0.000	0.009
19	11.349	265.464	-0.152	2.079	0.019	0.000	0.013
20	11.627	265.464	-0.251	2.048	0.017	0.000	0.014
21	11.436	265.464	-0.379	2.052	0.018	1.000	0.013
22	11.611	289.601	-0.200	2.110	0.018	0.000	0.011
23	11.699	134.282	-0.216	1.894	0.021	1.000	0.013
24	11.699	265.182	-0.196	2.137	0.015	1.000	0.012
25	11.699	265.182	-0.264	2.128	0.015	1.000	0.012
26	11.699	315.021	-0.272	2.127	0.015	1.000	0.011
27	11.611	342.734	-0.386	2.202	0.012	1.000	0.008
28	11.575	329.478	-0.432	2.140	0.018	1.000	0.010
29	11.699	340.088	-0.215	2.118	0.017	1.000	0.011
30	11.699	365.251	-0.226	2.148	0.014	1.000	0.011
31	11.699	390.505	-0.274	2.132	0.014	1.000	0.011
32	11.452	256.524	-0.292	2.171	0.014	1.000	0.010
33	11.645	226.262	-0.125	2.166	0.016	2.000	0.012
34	11.645	266.607	-0.303	2.133	0.017	2.000	0.012
35	7.223	184.538	-0.235	2.068	0.018	4.000	0.010
36	7.647	291.649	-0.177	2.232	0.013	4.000	0.011
37	7.084	367.750	-0.088	2.265	0.011	4.000	0.011
38	6.945	393.298	-0.170	2.213	0.012	4.000	0.009
39	6.920	291.649	-0.345	2.232	0.016	4.000	0.010
40	4.978	299.653	-0.238	2.319	0.015	3.000	0.010
41	4.978	299.653	-0.177	2.298	0.013	3.000	0.008
42	4.978	320.813	-0.203	2.337	0.014	3.000	0.008

Table 4. Experimental and Predicted Values of log(1/IC_{50}) Using Jack-Knife, PCR and PLS Methods

Molecule	Observed log(1/IC_{50})	Predicted Jack-Knife	Predicted PCR	Predicted PLS	ΔΘ Jack-Knife	ΔΘ GA-PCR	ΔΘ GA-PLS
1	1.531	1.345	1.835	1.813	0.186	-0.304	-0.282
2	1.653	1.621	2.011	1.981	0.032	-0.358	-0.328
3	0.854	0.819	1.118	1.004	0.035	-0.264	-0.150
4	1.591	1.617	1.432	1.467	-0.026	0.159	0.124
5	1.771	1.647	1.566	1.621	0.124	0.205	0.150
6	1.568	1.633	1.359	1.384	-0.065	0.209	0.184
7	1.699	1.608	1.313	1.311	0.091	0.386	0.388
8	1.740	1.633	1.683	1.572	0.107	0.057	0.168
9	0.919	1.597	1.665	1.582	-0.678	-0.746	-0.663
10	1.634	1.666	1.577	1.666	-0.032	0.057	-0.032
11	0.845	1.489	1.465	1.500	-0.644	-0.620	-0.655
12	1.623	1.795	1.596	1.738	-0.172	0.027	-0.115
13	1.763	1.534	1.477	1.523	0.229	0.286	0.240
14	1.663	1.705	1.687	1.623	-0.042	-0.024	0.040
15	0.919	1.428	1.491	1.412	-0.509	-0.572	-0.493
16	1.613	1.817	1.404	1.353	-0.204	0.209	0.260
17	1.653	1.328	1.264	1.284	0.325	0.389	0.369
18	1.681	1.450	1.583	1.564	0.231	0.098	0.117
19	0.949	1.144	1.173	1.119	-0.195	-0.224	-0.170
20	0.826	1.208	1.330	1.282	-0.382	-0.504	-0.456
21	1.544	1.265	1.334	1.294	0.279	0.210	0.250
22	1.653	1.276	1.161	1.170	0.377	0.492	0.483
23	1.653	1.512	1.748	1.861	0.141	-0.095	-0.208
24	1.690	1.568	1.522	1.565	0.122	0.168	0.125
25	1.477	1.406	1.627	1.700	0.071	-0.150	-0.223
26	1.505	1.639	1.378	1.448	-0.134	0.127	0.057
27	1.672	1.626	1.236	1.292	0.046	0.436	0.380
28	1.602	1.743	1.438	1.450	-0.141	0.164	0.152
29	1.672	1.619	1.151	1.108	0.053	0.521	0.564
30	0.833	1.021	0.988	0.981	-0.188	-0.155	-0.148
31	0.756	0.862	1.007	0.932	-0.106	-0.251	-0.176
32	1.532	1.772	1.456	1.424	-0.240	0.076	0.108
33	1.623	1.253	1.801	1.862	0.370	-0.178	-0.239
34	1.462	1.635	1.539	1.526	-0.173	-0.077	-0.064
35	1.690	1.288	1.103	1.107	0.402	0.587	0.583
36	0.863	1.230	0.772	0.804	-0.367	0.091	0.059
37	-0.076	-0.234	0.379	0.381	0.158	-0.455	-0.457
38	-0.658	-0.289	0.133	0.121	-0.369	-0.791	-0.779
39	1.756	1.219	0.571	0.663	0.537	1.185	1.093
40	0.820	0.735	0.331	0.244	0.085	0.489	0.576
41	-0.456	-0.669	0.136	0.109	0.213	-0.592	-0.565
42	0.079	-0.131	0.348	0.347	0.210	-0.269	-0.268

$\Delta\Theta = \Theta_{Observed} - \Theta_{Calculated}$

fitted in our models. So, it was necessary to reduce the number of descriptors through an objective feature selection which was performed in three steps. First, descriptors that had the same value for at least 70% of compounds within the dataset were removed. In next step, descriptors with correlation coefficients less than 0.3 with the dependent variable were regarded redundant and removed. Since highly correlated descriptors provide approximately identical information, a pair wise correlation was performed. When their correlation coefficient exceeded 0.90, one of two descriptors was randomly removed. Finally, Unscrambler program (version 9.7) was used for analysis of data and statistical methods.

RESULTS AND DISCUSSION

In the present study, two linear and non-linear variable selection methods were used to select the most significant descriptors. Based on the types of variable selection method and also the types of the feature mapping technique, these models can be shown as MLR-ANN, GA-MLR and GA-ANN. The models and descriptors selected by GA-MLR method are shown in Table 2. Also, the values for each molecular descriptors selected by GA-MLR method are shown in Table 3.

The prediction of log(1/IC_{50}) using a Jack Knife, PCR and PLS methods and also the difference between observed and predicted values are given in Table 4. The GA-ANN was used to construct a quantitative relation between activities of quinolizidinyl derivatives analogues and their calculated descriptors (Fig. 1). The main selected descriptors using GA-MLR models are shown in Table 5. The statistical parameters of the QSAR models with different methods are compared in Table 6. According to the results of the Jack-Knife method, the following compounds have the smallest difference between the observed and predicted values exist and are proposed for drug design.

The seven most significant descriptors which were selected by GA-stepwise MLR are as follows:

HOMT, D/Dr06, Mor28p, R2e, R8u+, nDB, R7v+.

MoRSE descriptors (3D Molecule Representation of Structures based on Electron diffraction) are derived from Infrared spectra simulation using a generalized scattering function.

Harmonic Oscillator Model of Aromaticity index (HOMA) index is based on the degree of alternation of single and double bonds, measuring the bond length deviations from optimal lengths attributed to the typical aromatic state.

Appropriate descriptors were obtained dusing GA-ANN model, which describes the characteristics of atomic mass and Randic shape index for quinolizidinyl derivatives, and are as follows:

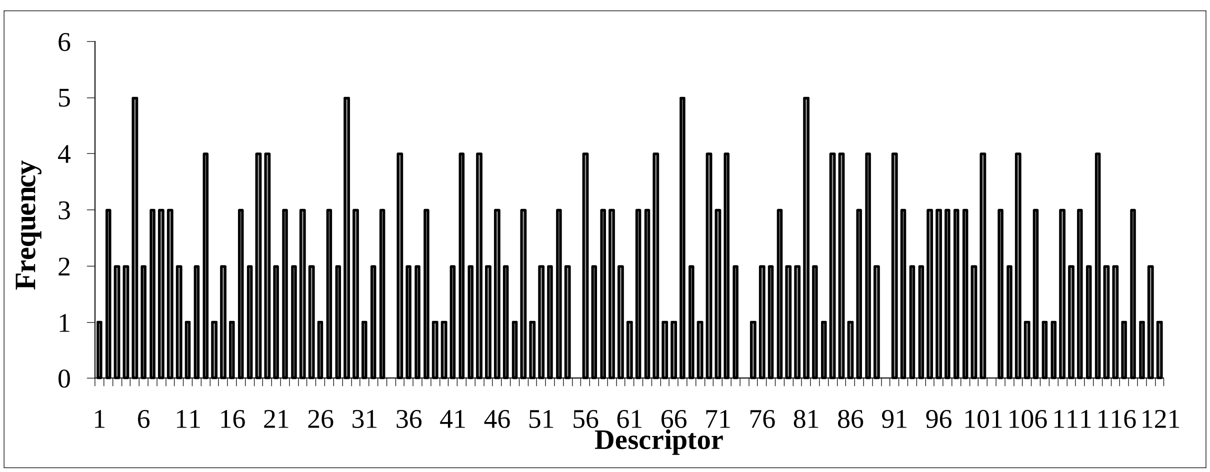


Fig. 1. The results of GA-ANN for correlation0.3 to up.

Table 5. The Mean of Selected Descriptors

Descriptor symbol	Descriptor group	Meaning
HOMT	Geometrical (3D)	HOMA total (trial)
D/Dr06	Topological (2D)	Distance/detour ring index of order 6
Mor28p	3D MoRSE	3D-MoRSE-signal 28/weighted by atomic polarizabilities
R2e	GETAWAY (3D)	R autocorrelation of lag 2/weighted by atomic Sanderson electronegativities
R8u+	GETAWAY (3D)	R maximal autocorrelation of lag 8/unweighted
nDB	Constitutional (0D)	Number of double bonds
R7v+	GETAWAY (3D)	R maximal autocorrelation of lag 7/weighted by atomic van der Waals volumes
PW2	Topological (2D)	path/walk 2-Randic shape index
MATS4m	2D autocorrelations	Moran autocorrelation -lag 4/weighted by atomic masses
Mor15m	3D-MoRSE	3D-MoRSE-signal 15/weighted by atomic masses
G1u	WHIM	1st component symmetry directional WHIM index/unweighted

Table 6. The statistical Parameters of Different Constructed QSAR Models

Method	Calibration	Prediction	Calibration	Prediction
GA-PCR	0.3979	0.6407	0.5715	0.4687
GA-PLS	0.3841	0.4744	0.6007	0.4370
GA-ANN Jack-Knife	0.1406	0.2165	0.8969	

PW2, MATS4m, Mor15m, G1u.

Molecular Path/Walk Index is defined as the average sum of atomic path/walk indices of equal length.

CONCLUSIONS

QSAR study has been carried out on quinolizidinyl derivatives as potent inhibitors of Acetyl and Butyrylcholin esterase in Alzheimer's disease. The GA-stepwise MLR selected descriptors were used as inputs for the construction of ANN model. As can be seen from Table, electronegativities, atomic polarizability and atomic van der Waals volumes were important descriptors in our study.

In the present study, GA-ANN (Jack-knife) as the most appropriate method for the correlation, because it has the highest value of $R^2$and lowest RMSE. We suggest that compounds No.4, 6, 10, 14, 24, 26 and 34 have the most appropriate structure.

ACKNOWLEDGMENTS

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

REFERENCES

[1] Y. An, C. Zhang, S. He, J. Neuron. 53 (2007) 477.
[2] G.G. Glenner, C.W. Wong, Biochem. Biophys. Res. Commun. 16 (1984) 885.
[3] M. Citron, Nat. Rev. Drug Discov. 9 (2010) 387.
[4] M.N. Sabbagh, Am. J. Geriatr Pharmacother. 7 (2009) 167.
[5] M.M. Mesulam, A. Guillozet, P. Shaw, B Quinn Neurobiol. Dis. 9 (2002) 88.
[6] Q.S. Du, P.G. Mezey, K.C Chou, J. Comput. Chem. 26 (2005) 461.
[7] Q.S. Du, R.B. Huang, Y.T. Wei, L.Q. Du, K.C. Chou, J. Comput. Chem. 29 (2008) 211.
[8] S.C. Basak, S. Bertelsen, G.D. Grunwald, Toxicol. Lett. 79 (1995) 239.
[9] A.R. Katritzky, V.S. Lobanov, M. Karelson, Chem. Soc. Rev. 24 (1995) 279.
[10] H. Gonzalez-Diaz, I. Bonet, C.E. Teran, Eur. J. Med. Chem. 42 (2007) 580.
[11] S. Vilar, L. Santana, E. Uriarte, J. Med. Chem. 49 (2006) 1118.
[12] S. Handschuh, J. Gasteiger, J. Mol. Model. 6 (2000) 358.
[13] T. Kimura, J. Chem. Inf. Comput. Sci. 38 (1998) 276.
[14] J. Huuskonen, J. Chem. Inf. Comput. Sci. 40 (2000) 773.
[15] G. Schneider, J. Med. Chem. 42 (1999) 5072.
[16] F.R. Burden, D.A. Winkler, J. Med. Chem. 42 (1999) 3183.
[17] F.R. Burden, J. Chem. Inf. Comput. Sci. 40 (2000) 1423.
[18] K.J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston, A.C. Evans. Hum. Brain Mapp. 4 (1996) 58.
[19] J. Nilsson, S. de Jong, A.K. Smilde, J. Chemom. 11 (1997) 511.
[20] S. Hulland, Strategic Manage. J. 20 (1999) 195.
[21] N.J. Lobaugh, R. West, A.R. McIntosh, Psychophysiology 38 (2001) 517.
[22] A. Hoskuldsson, J. Chemom. 2 (1988) 211.
[23] G. Ghasemi, M. Nirouei, S. Shariati, P. Abdolmaleki, Z. Rastgoo, Arabian J. Chem. (Article in press) (2011).
[24] G. Ghasemi, S. Arshadi, A. Nemati Rashtehroodi, M. Nirouei, S. Shariati, Z. Rastgoo, J. Comp. Med (2013) 1.
[25] M. Nirouei, G. Ghasemi, P. Abdolmaleki, A. Tavakoli, S. Shariati, Indian J. Biochem. Biophy. 49 (2012) 202.
[26] B. Tasso, M. Catto, O. Nicolotti, J. Med. Chem. 46 (2011) 2170.