QSRR Prediction of Immobilized Artificial Membrane Retention Factors of Some Drugs

Mohammad Hossein Fatemi[1]

Hoda Shamseddin[1]

Abstract

In this work multiple linear regression (MLR) and artificial neural network (ANN) were used to predict the retention factors of 40 basic and neutral drugs in immobilized artificial membrane liquid chromatography. Two separate models were developed for prediction of solute retention in two mobile phase compositions which were used five identical descriptors. The standard errors in ANN calculation of for training, internal and external test sets were 0.205, 0.3299 and 0.389, respectively, while these values are 0.280, 0.426 and 0.448, respectively for MLR model. Also the standard errors in ANN prediction of for training, internal and external test sets were 0.144, 0.596 and 0.557, respectively, while these values are 0.318, 0.613 and 0.453, respectively for MLR model. The validation and robustness of these ANN models were evaluated by cross-validation and Y-scrambling methods, which produce successful results.

Key Words: Artificial neural network● Molecular descriptor● QSRR●Immobilized artificial membrane chromatography

Introduction

The development of immobilized artificial membrane (IAM) chromatography unfolded new perspectives in the application of HPLC for the rapid evaluation of drug partitioning into cell membranes [1–3]. IAMs are monolayers of phospholipid molecules covalently bonded to the surface of silica particles. The functional groups of the bonded phospholipids are considered to play an important role in retention especially if charged molecules are analyzed, while for small neutral compounds the intermolecular forces resemble those underlying partitioning in octanol/water and retention in reversed-phase liquid chromatography [4, 5]. The pharmacokinetic behavior of drugs to access to their target sites are strongly dependent on the rate of their passive diffusion which was affected by the type and extent of their interactions with biological membranes. These behaviors can be accounted by solute retention in immobilized artificial membrane liquid chromatography {IAM-LC) [6, 7].

Therefore the investigation of drug retention in IAM-LC is very important in study of drug delivery and activities predictions. The retention of solute in chromatography depends on the interactions of solute with mobile and stationary phases. The type and extent of these interactions depends on solute structure. Therefore it is possible to relate the solute retention to its structural parameters, as was done in quantitative structure-retention relationships (QSRR) investigation. The results of this study can use in the prediction of the retention for new compounds as well as in further understanding of solute retention mechanism and also in prediction of drugs activities. There are several reports about QSRR prediction of solute retention in IAM-LC [8–16].

[1]Department of Chemistry, University of Mazandran, Babolsar, Iran

Li et al. used partial least squares regression (PLSR) for QSRR prediction of retention indices of 55 structurally diverse drugs in immobilized artificial membrane chromatography. The statistical parameters in prediction of logarithm of solute retention by their model were; regression coefficient of $R = 0.902$ and root mean square error ($RMSE$) of 0.400 [17]. Also Luco et al. represented a quantitative structure-retention relationship model for prediction of the capacity factors of 32 structurally diverse drugs in IAM-LC. They used kappa shape indices, the count of electron pairs on oxygen and nitrogen atoms, the count of O-H and N-H bonds, molar volume, molecular weight, total energy, heat of formation, energy of highest occupied molecular orbital, energy of lowest unoccupied molecular orbital, dipole moment, the most positive partial charge on a hydrogen atom and the most negative partial charge in the molecule as descriptors in their QSPR models [18]. Their orthogonal signal correction- partial least square model gives the statistics of $R = 0.979$, standard error =0.09 and Fisher statistics of F =537. Other quantitative structure–retention relationships studies in IAM-LC were done by using linear solvation energy relationships (LSER) parameters by Valko et al. [19, 20]. Moreover, Li et al. have introduced an electronic descriptor into the amended LSER parameters to describe accurately the retention of ionized solutes on IAM chromatography. They have obtained a satisfactory regression coefficient of R = 0.948 for all studied solutes compared with that of the original LSER without considering the electronic factor (R = 0.860) [21].

The main aim of the present work was to development of a quantitative structure retention relationships model using linear and non-linear techniques in modeling of retention factors of 40 structurally diverse neutral and basic drugs in immobilized artificial membrane chromatography. It is obvious that such model not only can used to predict the retention factor and biological activities of other drugs, but also can clarify the retention mechanism of solute in IAMC.

Experimental

Data Set

The data set of retention factor in IAM-LC was taken from the values reported by Vrakas et al. [22] which is shown in Table1. The data set consists of the logarithm of retention factors of 40 basic and neutral drugs in two mobile phase compositions. Separation of these compounds were performed using 0.02M morpholinepropanesulfonic acid (MOPS) and also in 0.01M phosphate buffer saline (PBS) at pH 7.4 as the aqueous mobile phases and IAM.PC.DD2 stationary phase(the IAM.PC.DD2 column was filled with phosphatidylcholine (PC) residues covalently bonded to silica(Regis Technologies, Morton Grove, IL, USA)). Therefore two separate QSPR models were developed to investigate the retention of solutes in each mobile phase compositions. In each case the data set were sorted according to the solute retention factor and the training, internal and external test sets were chosen from this list with desired distance from each other. The training set was used to adjust the parameters of model, the internal test set was used to prevent the over fitting of model and the external test set was used to evaluate the prediction power of the constructed models.

Table 1 Data Set and Corresponding Observed and ANN and MLR Predicted Values of the $\log k^{IAM}$

		$\log k^{IAM}_{MOPS}$			$\log k^{IAM}_{PBS}$		
number	Name	Exp	MLR	ANN	Exp	MLR	ANN
1	Acyclovir	−0.62	0.80-	0.01	−1.15	-0.73	0.00
2	Haloperidol	3.57	3.89	3.54	2.65	3.37	2.68
3	Amitriptyline	3.46	3.67	3.67	2.99	2.96	2.72
4	Amlodipine	3.50	3.19	3.56	2.79	2.79	2.79
5	Atenolol	0.87 *ext*	0.97	0.10	0.51	0.76	0.51
6	Verapamil	3.40	3.65	3.41	2.76	1.33	2.76
7	Bromazepam	1.62	1.77	1.62	1.44	1.40	1.44
8	Diazepam	2.21	2.54	2.20	2.12	2.22	2.12
9	Dimethidene	3.26	3.19	3.25	2.62	2.60	2.63
10	Dipyridamol	3.23	2.86	3.23	3.21	2.44	3.21
11	Doxepine	3.37 *int*	3.55	3.75	2.50	2.84	2.78
12	Theophylline	0.06	0.12-	0.01	−0.08	-0.27	0.01
13	Imipramine	3.52	3.52	3.74	2.73	2.81	2.81
14	Clomipramine	4.05	3.99	3.97	3.29	3.39	3.33
15	Clopamide	1.18	1.59	1.20	0.97 *int*	1.15	0.97
16	Lidocaine	1.58	1.71	1.58	1.27	1.32	1.27
17	Lorazepam	2.27	1.94	2.28	2.03	1.97	2.03
18	Maprotiline	3.89	3.63	3.65	2.81	2.92	2.77
19	Metformine	0.19	0.71	0.20	−0.37 *int*	0.52	0.13
20	Midazolam	2.90	3.15	2.93	2.77	2.75	2.77
21	Nifedipine	1.55	1.94	1.56	1.66	1.58	1.66
22	Nortriptyline	3.78 *ext*	3.51	3.74	2.83	2.81	2.82
23	Norfluoxetine	4.10 *int*	3.18	3.74	3.04	2.48	3.02
24	Oxprenolol	2.15 *ext*	2.79	2.79	1.47 *int*	2.16	2.13
25	Paracetamol	0.25	0.57	0.25	0.18	0.29	0.18
26	Pindolol	2.10	2.39	2.09	1.47 *ext*	1.84	1.24
27	Piracetam	−0.70 *int*	0.04-	0.02	-	-	-
28	Prazepam	2.94 *int*	2.63	3.77	2.62 *int*	2.37	3.21
29	Promethazine	3.63	3.25	3.76	2.78 *ext*	2.56	3.02
30	Propranolol	2.95	2.91	2.96	2.33	2.28	2.33
31	Protriptyline	3.69	3.51	3.73	2.79	2.81	2.83
32	Pyrimethamine	2.38	2.10	2.36	1.87	1.79	1.87
33	Temazepam	1.96	1.69	1.93	1.76	1.58	1.72
34	Tioconazole	3.65	3.69	3.63	3.86 *ext*	3.56	3.33
35	Trimethoprim	1.51	1.45	1.50	1.05	1.13	1.05
36	Hydroxyzine	3.33	3.99	3.34	3.01 *ext*	3.44	2.76
37	Fluoxetine	4.08	3.39	3.77	2.98	2.69	2.90
38	Chlordiazepoxide	1.93	1.68	1.93	1.86	1.58	1.85
39	Chlorthalidone	1.49	1.30	1.48	1.37	1.22	1.40
40	Chlorpromazine	3.55 *ext*	3.72	4.03	3.33	3.14	3.33

** In the above table *int* refer to internal test set and *ext* refer to external test set

Descriptors

To obtain a reliable QSRR model, the structural feature of molecules should be encoded by the molecular descriptors. In the first step the structures of compounds were drawn with Hyperchem (version 7.0) program [23] and exported in a file format suitable for Mopac (version 6.0) package [24], on the basis of the minimum energy molecular geometries optimized by *AM*1 semi empirical method. Then the Hyperchem and MOPAC output files were transferred into software CODESSA. This software can calculate constitutional, topological, geometrical, electrostatic, and quantum chemical descriptors and has been successfully used by various QSPR researches [25-31]. Constitutional descriptors are related to the number of atoms and bonds in each molecule. The topological descriptors describe the atomic connectivity in the molecule. The geometrical descriptors describe the size of the molecule. The electrostatic descriptors reflect characteristics of the charge distribution of the molecule. The quantum chemical descriptors offer information about binding and formation energies, partial atom charge, dipole moment, and molecular orbital energy levels. Some descriptors generated for each compound, encoded similar information about the molecule of interest, therefore, it was desirable to test each pair of descriptor and eliminate those that show high correlation (R>0.90) with each other. Subsequently, the method of stepwise multiple linear regression was performed on the training set to select the most relevant descriptors. Two separate QSRR models were developed to predict the retention of solute in two different mobile phase composition. These two MLR-QSRR models have identical descriptors but with different coefficients. The names of descriptors and the statistical parameters of constructed MLR models are shown in Table 2. These descriptors were used as inputs for the generated artificial neural networks.

Neural Network Generation

An ANN is a biologically inspired computer program designed to learn from data in a manner of emulating the learning pattern in the brain. Most ANN systems are very complex and high-dimension processing systems. Training of the ANN can be performed using the back-propagation algorithm. In order to train the network using the back-propagation algorithm, the differences between the ANN output and its desired value are calculated after each training iteration and the values of weights and biases modified using this error term. A detailed description of the theory behind a neural network has been adequately described elsewhere [32-34]. The program for the feed-forward neural network that was trained by the back-propagation algorithm was written in MATLAB 7.4. Descriptors that appeared in the selected MLR model were used as inputs for the generated ANN, and its output was the retention factor for the molecule of interest. Therefore this network has five nodes in the input layer and one node in the output layer. The number of nodes in the hidden layer would be optimized. The initial weights were randomly selected from a uniform distribution that ranged between-0.3 and 0.3. The initial bias values were set to be one. These values were optimized during the network training. The value of each input was divided into its mean value to bring them into the dynamic range of the sigmoid transfer function of the ANN. Before training, the network was optimized for the number of nodes in the hidden layer, learning rates, and momentum. Then, the network was trained using the training set to optimize the values of weights and biases. Finally in order to evaluate the prediction power of the ANN, a trained network was employed to calculate the retention factor for the external test set.

Result and Discussion

Linear Modeling

Two separate MLR models were constructed for QSRR modeling of IAM retention factors $\log k_{wPBS}^{IAM}$ and $\log k_{wMOPS}^{IAM}$ of selected drugs in different mobile phase compositions. Table 2 represented the specifications of these models.

These MLR models have the statistical parameters of $R = 0.967$, $SE = 0.280$ and $F = 420.785$and $R = 0.970$, $SE = 0.318$ and $F = 475.033$ for modeling of $\log k_{wPBS}^{IAM}$ and $\log k_{wMOPS}^{IAM}$, respectively. It can be seen from this table that five identical descriptors were appeared in these two models. These descriptors are: number of C atoms (nC), fractional hydrogen bonding acceptor ability of the molecule ($FHASA$), min (>0.1) bond order of N atom ($P_N^{\min}$), number of chlorine atoms (nCL) and average bond order of N atom ($\overleftarrow{N}$). The chemical meaning and the way of calculations of these descriptors were explained in the book of Molecular Descriptors by Todeschini et al. [35]. The correlations between these descriptors were calculated and are shown in Table 3. By inspection to these values, it was concluded that there is no significant correlation between selected descriptors. The calculated values of retention factors by these MLR equations were shown in Table 1.

Table2 Specifications of Multiple Linear Regression Models

		$\log k_{wPBS}^{IAM}$			$\log k_{wMOPS}^{IAM}$		
Descriptors	Notation	coefficient	Standard error	Mean effect	coefficient	Standard error	Mean Effect
Number of C atoms	nC	0.116	±0.020	0.932	0.120	±0.021	0.816
Number of chlorine atoms	nCl	0.585	±0.103	0.109	0.477	±0.088	0.098
Fractional H-bonding acceptor ability of the molecule	$FHASA$	-3.396	±0.787	-0.155	-4.463	±.0.884	-0.190
Min (>0.1) bond order of N atom	$P_N^{\min}$	1.243	±0.235	0.417	1.653	±0.257	0.459
Average bond order of N atoms	$\overline{N}$	-1.988	±0.535	-0.970	-2.696	±0.574	-1.126
Constant		1.376	±0.507		2.309	±0.582	

Table3 the Correlation Matrix between Selected Descriptor

	nC	nCl	$FHASA$	$P_N^{\min}$	$\overline{N}$
nC	1	-0.718	-0.726	0.424	-0.037
nCl		1	0.416	-0.071	0.041
$FHASA$			1	-0.368	0.040
$P_N^{\min}$				1	0.010
$\overline{N}$					1

Non-linear modeling

The selected descriptors can be used as inputs for generation of ANN models. The first step in the generation of a neural network was the optimization of its parameters. These parameters are; the number of nodes in the hidden layer, weights and biases learning rates and the momentum values. The procedure for optimization of these parameters is given in our previous works [36, 37]. Table 4 shows the architecture and specification of the optimized ANN models. Then the network was trained by using the training sets for the optimization of the weights and biases values by back propagation algorithm.

It is known that neural network can become over-trained. An over-trained network has usually learned perfectly the stimulus pattern it has seen but can not give accurate prediction for unseen stimuli, and it no longer able to generalize. There are several methods for overcoming this problem. One method is to use an internal test set to evaluate the prediction power of the network during its training. In this method after each 1000 training iteration the network was used to calculate retention factor of molecules included in the internal test set. To maintain the predictive power of the network at a desirable level, training was stopped when the value of errors for the internal test set started to increase. Since this error is not a good estimate of the generalization error, prediction potential of the model was evaluated on a third set of data, named external test set. Compounds in the external test set were not used during the training process and were reserved to evaluate the predictive power of the generated ANN.

Table 4 Architecture and Specification of the Generated ANNs

	$\log k_{wMOPS}^{IAM}$	$\log k_{wPBS}^{IAM}$
No. of nodes in the input layer	5	5
No. of nodes in the hidden layer	4	4
No. of nodes in the output layer	1	1
Weights learning rate	0.3	0.4
Bias learning rate	0.5	0.6
Momentum	0.5	0.5
Transfer function	Sigmoid	Sigmoid

Table5 The Statistical Parameters Obtained Using the MLR And ANN Models in Prediction of $\log k_{wPBS}^{IAM}$.

	ANN model			MLR model		
	R	F	SE	R	F	SE
Training	0.983	810.887	0.205	0.967	420.785	0.280
Internal	0.976	40.559	0.329	0.960	23.330	0.426
External	0.947	17.320	0.389	0.929	12.595	0.448
	$R^2_{cross-validation} = 0.946$				$R^2_{cross-validation} = 0.914$	
	$RMSE = 0.276$				$RMSE = 0.332$	
	$Q^2 = 0.938$				$Q^2 = 0.911$	

Model validation

For the evaluation of the prediction power of thenetwork, the trained ANN was used to predict theretention factors of the molecules included inprediction set. Table 1 represents the experimental and predicted values of retention factorsusing the generated ANN for the training, internal and external test sets. The statistical parameters obtained in calculation of $\log k_{wPBS}^{IAM}$ and $\log k_{wMOPS}^{IAM}$ by the ANN and MLR models are shown in Table 5 and Table 6. It can be seen from these tables that although descriptors appearing in the MLR models are used as inputs for the ANNs, the statistics of the latter show a large improvement.

The residuals of the ANN calculated values of the retention factors are plotted against their experimental values in Fig. 1. The propagation of the residuals in both sides of zero line indicates that no systematic error exists in the constructed QSPR model.

The leave many-out cross-validation method was used for the evaluation the prediction power of the obtained MLR and ANN models. The obtained statistical results of leave-five-out cross validation test is $Q^2 = 0.929$ and $Q^2 = 0.949$ for MLR and ANN model, respectively in prediction of $\log k_{wMOPS}^{IAM}$ and $Q^2 = 0.911$ and $Q^2 = 0.938$ for MLR and ANN models, respectively in prediction of $\log k_{wPBS}^{IAM}$. Another widely used approach to establish the model robustness is so called y-randomization test (randomization of response, i.e. in our case, retention) [38]. It consists of repeating the calculation procedure with randomized retention vector and subsequent probability assessment of the resultant statistics. It is expected that models obtained for the dataset with randomized retention should have low values of R^2. However, sometimes models based on the randomized data have high R^2 values due to chance correlation or structural redundancy [39]. The results of 30 times repetitions in randomization of Y vectors (retention factors) on MLR models were shown in Table 7. As can be seen from this table the random models were found to have significantly lower R^2 values ($\overline{R^2} = 0.129$ and $\overline{R^2} = 0.165$ in prediction of $\log k_{wPBS}^{IAM}$ and $\log k_{wMOPS}^{IAM}$, respectively) than the original model, which indicate that the good results in our original models are not due to the chance or structural dependency of the training set.

Table6 The Statistical Parameters Obtained Using The MLR and ANN Models in Prediction of $\log k_{wMOPS}^{IAM}$.

	ANN			MLR		
	R	F	SE	R	F	SE
Training	0.994	2438.826	0.144	0.970	475.033	0.318
Internal	0.974	36.590	0.596	0.970	34.523	0.613
External	0.942	15.699	0.557	0.962	24.705	0.453
	$R^2_{cross-validation} = 0.951$			$R^2_{cross-validation} = 0.931$		
	$RMSE = 0.300$			$RMSE = 0.351$		
	$Q^2 = 0.949$			$Q^2 = 0.929$		

Table7the results of Y-randomization test for $\log k_{wPBS}^{IAM}$ and $\log k_{wMOPS}^{IAM}$ for MLR models

	$\log k_{wPBS}^{IAM}$	$\log k_{wMOPS}^{IAM}$		$\log k_{wPBS}^{IAM}$	$\log k_{wMOPS}^{IAM}$
Number	R^2	R^2	Number	R^2	R^2
1	0.203	0.131	16	0.077	0.149
2	0.051	0.257	17	0.052	0.306
3	0.193	0.070	18	0.220	0.175
4	0.146	0.082	19	0.077	0.048
5	0.154	0.212	20	0.203	0.016
6	0.118	0.067	21	0.120	0.200
7	0.028	0.326	22	0.043	0.142
8	0.107	0.363	23	0.090	0.193
9	0.295	0.230	24	0.191	0.187
10	0.135	0.109	25	0.075	0.227
11	0.084	0.078	26	0.080	0.157
12	0.173	0.128	27	0.126	0.049
13	0.024	0.265	28	0.099	0.328
14	0.243	0.050	29	0.289	0.180
15	0.092	0.076	30	0.181	0.130

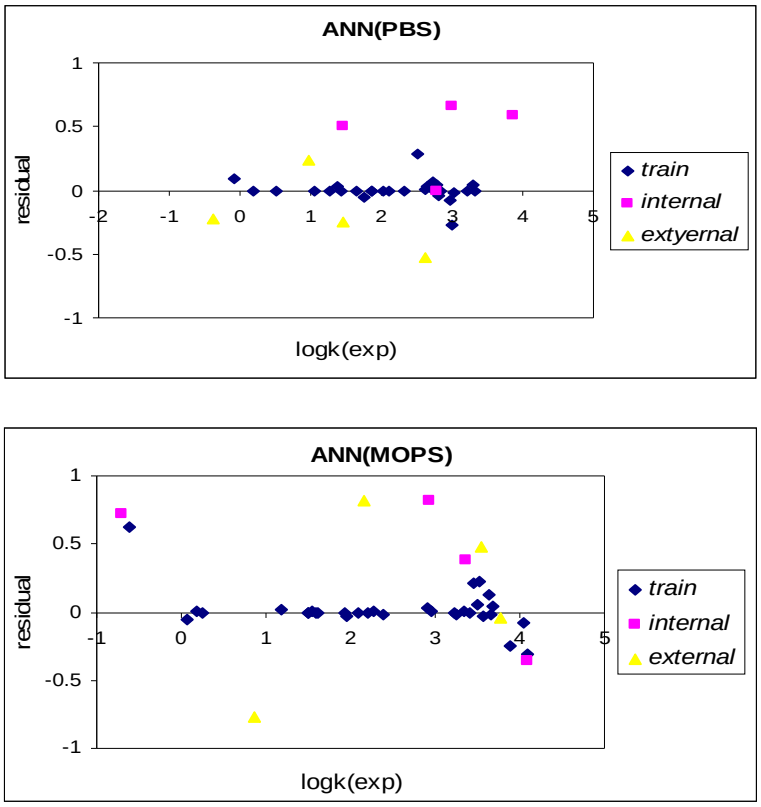

Fig. 1 Plot Of Residual Versus Experimental Values of Retention Factor For $\log k_{PBS}^{IAM}$ (A)and $\log k_{MOPS}^{IAM}$ (b).

Descriptor Interpretation

In order to obtain the relative importance and contribution of each descriptor in the models we calculated the values of mean effect (*ME*) for each descriptor from the following Eq. (1).

$$ME_j = \frac{\beta_j \sum_{i=1}^{i=n} d_{ij}}{\sum_j^m \beta_j \sum_i^n d_{ij}} \qquad (1)$$

Where$ME_j$is the mean effect for considered descriptor j, β_j is the coefficient of descriptor j in MLR equation and d_{ij} is the value of descriptor j for molecule i, and m is the number of descriptors in the model. The value of *ME* revealed the relative importance of a descriptor in comparison with other descriptors. The sign of mean effect showed the direction of influencing of descriptors on the value of retention factor. The calculated values of mean effects are indicated in the last column of Table 2 and also are shown in Fig. 2. Inspection to these values reveals that the order of importance of descriptors in two models is identical and is $\overline{N} > nC > P_N^{MIN} > FHASA > nCL$.

Most important descriptor with the highest mean effect is average bond order of N atoms. This descriptor can represent the electronic structure of molecules and can affect on the extent of electrostatic interactions between solute and mobile and stationary phases.

The second important descriptor according to its mean effect is number of C atoms. As the number of carbons in a molecule increases the hydrophobicity of molecule increases, therefore their retention increases. The next descriptor is min (>0.1) bond order of N atoms which is a quantum chemical descriptor. This descriptor relate to the strength of intramolecular bonding interactions and characterize the stability of the molecules. The positive sign for the mean effect of this descriptor indicates that an increasing in the value of this descriptor causes an increasing in retention factor. The forth descriptor in the model is fractional hydrogen bonding acceptor ability of the molecule ($FHASA$). This descriptor defines as below:

$$FHASA = \frac{HASA1}{TMSA} \qquad (2)$$

where $HASA1$ is the hydrogen bonding acceptor ability, and $TMSA$ is the total molecular surface area. When the $FHASA$ increases, the hydrophilicity of molecules increase and its retention was decreases. The last descriptor in the model is the number of Cl atoms. By increasing this value the size and molecular weight of solute increases. All of these descriptors can encode different aspects of solute which affected on hydrophobic and steric and electrostatic interactions which control the solute retention in liquid chromatography.

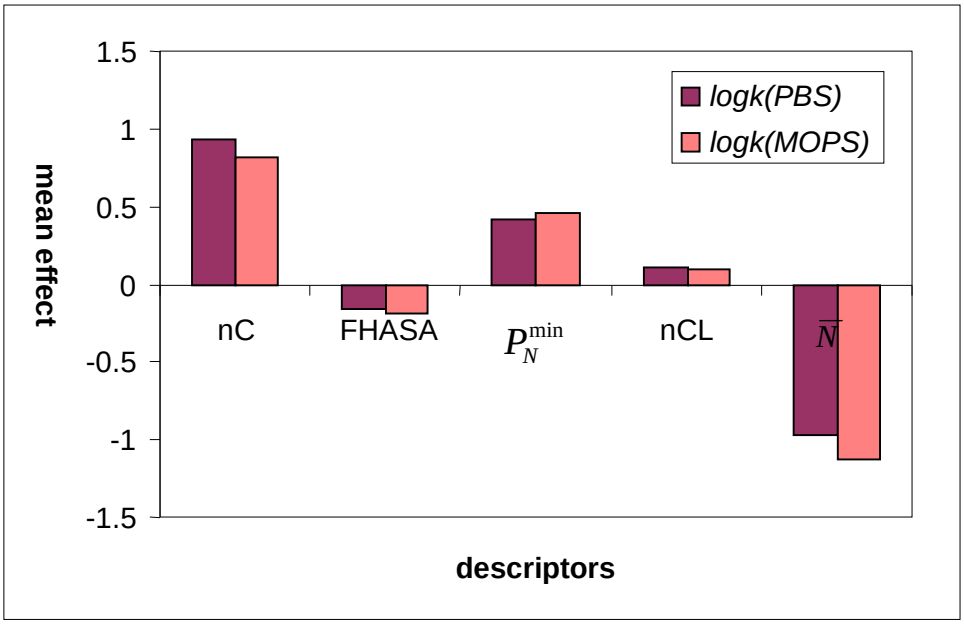

Fig. 2 Plot of Mean Effect Values versus Descriptors

Conclusions

Results of this study reveal that MLR and ANN can be used successfully in developing of a QSRR model to predict the solute retention factors in IAMC, in different aqueous mobile phases compositions. Descriptors that were appeared in the models are constitutional and quantum chemical types which can encode features of molecules that are responsible in steric, and lipophilicity interactions of molecules. The statistical parameters of ANN model are few better than MLR one, which illuminate that there are some non-linear relations between molecular descriptors and solute retention in IAM-LC. The validation and robustness of these ANN models were evaluated by cross-validation and Y-scrambling methods, which produce successful results.

References

Ong S, Liu H, Qiu X, Pidgeon C, (1996) J Chromatogr A 728:113.
Salminen T, Pulli A, Taskinen (1997) J Pharm Biomed Anal 15: 469.
Reichel A, Begley D, (1998) J Pharm Res 15: 1270.
Taillardat-Bertschinger A, Barbat F, Quercia MT, Carrupt PA, Reist M, La Rotonda MI, Testa B, (2002) HelvChim Act 85: 519.
Taillardat-Bertschinger A, Martinet CAM, Carrupt PA, Reist M, Caron G, FrutteroR, Testa B, (2002) Pharm Res 19: 729.
Vrakas D, Tsantili-Kakoulidou A, Hadjipavlou-Litina D, (2003) Quant StructAct Relat 22: 622.
Vrakas D, Hadjipavlou-Litina D, Tsantili-Kakoulidou A, (2005) J Pharm Biomed Ana 39: 908
Barbato F, La Rotonda M.I, Quaglia F, (1997) Pharm Res 14: 1699.
Sun J, Deguchi Y, Chen JM, Zhang RH, Morimoto K, (2002) ActaPharmacol Sin5: 430.
Barbato F, Quaglia F, Quercia MT, La Rotonda MI, (2000) HelvChimActa83: 767.
Barbato F, La Rotonda MI, Quaglia F, (1996) Eur J Med Chem 31: 311.
Abraham MH, Gola JMR, Kumarsingh R, Cometto-Muniz JE, Cain WS, (2000) J Chromatogr B 745: 103.
Barbato F, di Martino G, Grumetto L, La Rotonda MI, (2004) Eur J Pharm Sci22: 261.
Barbato F, di Martino G, La Grumetto MI, (2005) Eur J Pharm Sci 25: 379.
Sun J, Deguchi Y, Chen J, Zhang R, Morimoto K, (2002) Pharmazie 57: 497.
Sun J, Deguchi Y, Chen J, Zhang R, Morimoto K, (2002) Acta Pharm Sin 23: 430.
Li J, Sun J, He Zh, (2007) JChromatogr A 1140:174.
Luco JM, Salinas AP, Torriero AAJ, V´azquez RN, Raba J, Marchevsky E, (2003) J ChemInfComputSci 43: 2129.
Valko K, Du CM, Bevan CD, Reynolds DP, Abraham MH,(2000) J Pharm Sci89: 1085.
Lepont C, Poole CF, (2002) J Chromatogr A946: 107.
Li J, Sun J, Cui SM, He ZG, (2006) J Chromatogr A 1132: 174.
Vrakas D, Giaginis C, Tsantili-Kakoulidou A, (2006) J Chromatography A 1116: 158.
Hyper release 7.0 for windows, Hypercube Inc (2002), http://www.hyper.com
Stewart JPP, MOPAC 6.0, Quantum chemistry program exchang, (1990).
Katritzky AR, Lobanov VS, Karelson M, CODESSA: Training Manual,University of Florida, Gainesville, FL, (1995).
Katritzky AR, Lobanov VS, Karelson M, CODESSA: Reference Manual, University of Florida, Gainesville, FL, (1994) *QCPE*, No. 455, India University, Blooming, IN (1989).
Hiob R, Karelson M, (2000) J ChemInfComputSci40: 1062.
Katritzky AR, Petrukhin R, Jain R, Karelson M, (2001) J ChemInfComputSci41: 1521.
Delgado EJ, Alderete JB, Jana GAA, (2003) J ChemInfComputSci 43: 1928.
Bosque R, Sales J, (2003) J ChemInfComputSci43: 637.
Hu RJ, Liu, HX, Zhang RS, Xue CX, Yao XJ, Liu MC, Hu ZD, Fan BT, (2005)Talanta 68(1): 31
Haykin S, (1994) Neural Network. Prentice-Hall, Englewood Cliffs, NJ 145-187.
Beal MT, Hagan HB, Demuth M, (1996) Neural Network Design. PWS, Boston, MA 75-92.
Bose NK, liang P, (1996) Neural Network Fundamentals. MCGraw-Hill, New York 241-250.http://www.Disat.unimib.it/chm/Dragon.htm
Jalali-Heravi M, Fatemi MH, (2000) Anal Chim Act 415: 95.
Jalali-Heravi M, Fatemi MH, (2000) J Chromatogr A 897: 227.
Wold S, Eriksson L, (1995) Statistical Validation of QSAR Results.In: Waterbeemd, H.van de (ed.), Chemometrics Methods in Molecular Design, VCH 309.
Clark RD, Sprous DG, Leonard JM, Validating Models Based on Large Dataset. In: Höltje, H-D, Sippl W (eds.) Rational Approaches to Drug Design. Proceedings of the 13th European Symposium on Quantitative Structure-Activity Relationships. 27Aug – 1Sept, (2000), Duesseldorf, Germany.Prous Science, (2001) 475.