Caspian J. Chem. 1(2012) 31-41

Caspian Journal of Chemistry

http://cjc.umz.ac.ir

Quantitative Structure-Pproperty Relationship Modeling of the Redox Potential for Some Phenolic Antioxidants

Mohammad Hossein Fatemi*, Zohreh Gharehchahi

Chemometrics Laboratory, Faculty of Chemistry, University of Mazandaran, Babolsar, Iran

* Corresponding author E-mail address: mhfatemi@umz.ac.ir

Received 16 May 2012|Received in revised form 14 August 2012|Accepted 4 September 2012

Abstract: In this work, quantitative structure-property relationship (QSPR) approaches were used to predict the redox potential of 42 phenolic antioxidants. The structures of all compounds optimized by the AM1 semi-empirical method and then a large number of molecular descriptors were calculated for each compound in the data set. Subsequently, stepwise multilinear regression was applied to select the most significant and relevant descriptors. The selected descriptors are; the highest occupied molecular orbital energy, the number of hydroxyl groups and harmonic oscillator model of aromaticity. These descriptors were used to develop the multiple linear regression (MLR) and artificial neural network (ANN) models. The values of root mean square error for ANN model were; 0.049, 0.075 and 0.043 for training, internal and external tests sets, respectively, while these values were; 0.061, 0.088 and 0.073, respectively for MLR model. Comparison between these values and other statistical parameters for these two models revealed the credibility of ANN in prediction of redox potential of phenolic antioxidants by using QSPR approaches.

Key words: Quantitative structure-property relationship, Artificial neural network, Molecular descriptor, Redox potential, Phenolic antioxidant.

®2012 Published by University of Mazandaran. All rights reserved.

1. Introduction

Oxidation is a chemical reaction that transfers electrons from a substance to an oxidizing agent. Although oxidation reactions are crucial for life, they can also be damaging by producing free radicals via chain reactions. These radicals are very reactive, rapidly attack and oxidize a variety of targets including DNA, proteins, enzymes, membranes and lipids [1-2]. Free radicals and some of oxygen-centered free radicals known as reactive oxygen species (ROS) (such as superoxide ($O_2^{\cdot -}$), peroxyl ($ROO^{\cdot}$), alkoxyl ($RO^{\cdot}$), hydroxyl ($HO^{\cdot}$) and nitric oxide ($NO^{\cdot}$)) have been recognized to be harmful and play an important role in the initiation and promotion

of various disease such as cancer, Alzheimer's, Parkinson's, arthritis, asthma, diabetes and degenerative eye disease [3-6]. An antioxidant is a molecule capable of inhibiting the oxidation of other molecules. Antioxidants are often reducing agents such as thiols, ascorbic acid (vitamin C), tocopherols (vitamin E), phenols and polyphenols [3]. The function of antioxidants is to suppressing the formation of free radicals by reacting with them at a rate faster than the organic substrate [2]. Recent studies on antioxidant reaction mechanisms indicated that the chain reaction was controlled mainly through the free radical-scavenging by phenolic hydroxyls of antioxidants [7-9]. These compounds characterized by at least one aromatic ring bearing one or more hydroxyl groups [8]. The process of phenolic antioxidants (ArOH) with a lipid peroxyl radical ($ROO^{\cdot}$), started with abstracting hydrogen atom from the phenolic hydroxyl group according to the following reaction [9]:

$$ROO^{\cdot} + ArOH \longrightarrow ROOH + ArO^{\cdot} \quad (1)$$

The rate of this reaction is much greater than the attack of the peyroxyl radicals on the organic substrate, that showed by the following:

$$ROO^{\cdot} + RH \longrightarrow ROOH + R^{\cdot} \quad (2)$$

The redox potential of phenolic antioxidants can be affected by their structural features. Therefore, it is possible to predict the antioxidant activities of these compounds from their theoretical derived molecular descriptors. Quantitative structure-activity/property relationship (QSAR/QSPR) is a powerful method in prediction and interpretation of physicochemical properties of chemicals. This method typically comprises of two main steps: (i) description of molecular structure and (ii) multivariate analysis for correlating molecular descriptors with observed property. Molecular descriptors are numerical values utilize to describe different characteristics about a certain structure in order to yield information about the property being studied. The key feature of the role of in silico techniques is that predictions can be made from molecular structure alone [10]. Consequently, such method would be an extremely beneficial tool for the development of novel therapeutic antioxidants that have not been synthesized or found.

The protective role of several classes of antioxidants has been widely studied by means of QSPR [11-14]. For instance, Bosque and Sales developed a QSPR model to predict the O-H bond dissociation energy for a set of 78 phenols [15]. They proposed a multiple linear regression, contains seven descriptors namely εHOMO, topographic electronic index, minimum atomic state energy for a C atom, RPCS relative positive charged SA, minimum bond order of a C atom, ZX shadow, FPSA-3 fractional PPSA. The statistical parameters for obtained QSPR model, including the correlation coefficient and standard error for training and prediction sets were 0.8978, 6.62, 0.9076 and 4.26, respectively.

Rasulev et al. has used a QSAR technique to predict the inhibition of lipid peroxidation (LPO) effects for 27 flavonoids [16]. Their results showed that the position of the OH groups, the magnitude of dipole moment and the shape of molecule play an important role in inhibition of LPO by flavonoids. Recently Abreu et al. modeling the radical scavenger activity (RSA) of 26 di(hetero)arylamine's derivatives of

benzo[b]thiophenes by using the partial least squares (PLSs) projection of latent structures method [17]. Nevertheless, redox potential had been applied seldom to build QSPR model. For example, Lien et al. derived a linear relationship between the one-electron redox potential and calculated parameters such as heat of formation, number of OH groups and energy of lowest unoccupied molecular orbital for 31 phenolic antioxidants.

They have reported correlation coefficient of R^2=0.914 and standard error of SE=0.074 for all studied phenolic antioxidants. Reis and coworkers compared four different computational methods including density functional theory (DFT), hartree-fock, AM1 and PM3 to predict redox potential of phenolic antioxidants by using simple and multiple linear regression models [18]. Their investigation indicated that among these methods, DFT had the best function to describe the properties of phenolic antioxidants. The vertical ionization potentials (IPvs) and the charge on oxygen atom 7, reported to be the significant molecular descriptors in this work. The main aim of the present work is to development of a nonlinear model to study the relationship between redox potential of phenolic antioxidants and their theoretical calculated molecular descriptors.

2. Experimental

2.1. Data set

The values of redox potential of the substituted phenols in water at pH 7 (E_7) were taken from References [18-19] and were used as data set. The collection of this data set consists of the redox potential for 42 phenol derivatives.

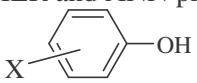
Table 1 shows the structure of chemicals in data set and their corresponding redox potential, in V. The redox potential fall in the range of 0.41 to 1.23 V for 4-NH_2 and 4-NO_2 phenol substitutions, respectively. The data were sorted according to their redox potential and the training, internal and external test sets were chosen from this list with desired distance from each other. The training, internal and external test sets consist of 32, 5 and 5 redox potential data, respectively. The training set was used to adjust the parameters of the model, while the internal and external test sets were applied to prevent overfiting of the model and evaluate the model predictivity, respectively.

2.2. Descriptors calculation

Molecular descriptors can be defined as a useful numbers, which are derived from molecular features of the interested molecules. These molecular descriptors were mainly computed by using the CODESSA software [20-22] and DRAGON (Ver. 3.0) package [23]. In order to calculate these descriptors, all structures of the compounds were drawn with Hyperchem (Ver. 7.0) program (Hyper (2002) release 7.0 for windows, Hypercube), optimized by the AM1 semi-empirical method and exported in a file format suitable for MOPAC (Ver. 6.0) package [24].

The Hyperchem and MOPAC output files were transferred into CODESSA software that can calculate constitutional, topological, geometrical, electrostatic and quantum chemical descriptors.

Table 1. Data set and corresponding observed, MLR and ANN predicted values of redox potential in V



No.	Substituent X	$E_{7(exp)}$	$E_{7(MLR)}$	$Resid_{(MLR)}$	$E_{7(ANN)}$	$Resid_{(ANN)}$
1	4-NO_2	1.23	1.26	0.03	1.22	-0.01
2	3,5-Cl_2	1.15	1.09	-0.06	1.10	-0.05
3	4-CF_3	1.13	1.19	0.06	1.18	0.05
4	3-NO_2	1.13_{ext}	1.27	0.14	1.21	0.08
5	4-PhCO	1.12	1.01	-0.11	1.04	-0.08
6	3-CN	1.11	1.14	0.03	1.12	0.01
7	4-I	1.09	1.00	-0.09	0.99	-0.1
8	4-COOH	1.04	0.98	-0.06	1.03	-0.01
9	3-CH_3CO	0.98	1.07	0.09	1.05	0.07
10	4-H	0.97	0.96	-0.01	0.93	-0.04
11	4-Br	0.96_{ext}	0.97	0.01	0.96	0.00
12	4-Cl	0.94	0.95	0.01	0.93	-0.01
13	4-F	0.93	0.88	-0.05	0.89	-0.04
14	Tyrosine	0.89	0.81	-0.08	0.87	-0.02
15	3-OH, 4-$COCH_3$	0.89_{int}	0.83	-0.06	0.82	-0.07
16	4-CH_3	0.87	0.86	-0.01	0.83	-0.04
17	3,5-$(CH_3O)_2$	0.85	0.81	-0.04	0.82	-0.03
18	3-CH_3	0.85	0.90	0.05	0.88	0.03
19	3-OH, 5-OCH_3	0.84_{ext}	0.78	-0.06	0.80	-0.04
20	3,5-$(CH_3)_2$	0.84	0.9	0.06	0.86	0.02
21	4-Ph	0.84	0.74	-0.10	0.73	-0.11
22	2-CH_3	0.82	0.89	0.07	0.86	0.04
23	3-OH	0.81_{int}	0.83	0.02	0.89	0.08
24	2-OCH_3	0.77	0.74	-0.03	0.73	-0.04
25	4-OCH_3	0.73_{ext}	0.75	0.02	0.75	0.02
26	3,4-$(CH_3O)_2$	0.67	0.68	0.01	0.67	0.00
27	3,4,5-$(CH_3O)_3$	0.66	0.57	-0.09	0.56	-0.10
28	Sesamol	0.62_{int}	0.67	0.05	0.64	0.02
29	2-OH, 4-COOH	0.60	0.65	0.05	0.61	0.01
30	2,6-$(CH_3O)_2$	0.58	0.63	0.05	0.59	0.01
31	2,3-$(OH)_2$	0.58	0.52	-0.06	0.52	-0.06
32	2,3-$(OH)_2$, 5-$COOCH_3$	0.56_{ext}	0.62	0.06	0.59	0.03
33	3,4-Dihydrocynnamic acid	0.54	0.53	-0.01	0.54	0.00
34	2-OH	0.53	0.63	0.10	0.57	0.04
35	2-OH, 4-CH_3	0.52	0.58	0.06	0.49	-0.03
36	4-OH	0.46_{int}	0.64	0.18	0.59	0.13
37	4-NH_2	0.41	0.41	0.00	0.44	0.03
38	4-CN	1.17	1.09	-0.08	1.09	-0.08
39	4-$COCH_3$	1.06_{int}	1.06	0.00	1.06	0.00
40	4-t-Bu	0.80	0.87	0.07	0.84	0.04
41	2,6-$(CH_3)_2$	0.77	0.84	0.07	0.82	0.05
42	2-OCH_3, 4-CH_3	0.68	0.70	0.02	0.71	0.03

The subscripts of int and ext refer to the internal and external test set, respectively. The substituent X is belonged to the compounds 1-13, 15-27, 29-32 and 34-42

In addition, Hyperchem output files were used by DRAGON to calculate some topological descriptors. The number of descriptors that were calculated by DRAGON and CODESSA software was 947 ones. In pre-screening step, constant or near constant values of descriptors were omitted from further investigation. The relative standard deviation below 1% was selected as criteria for elimination of constant and near constant variables. Since some descriptors that generated for each compound encoded similar information about the molecule of interest, the correlations of descriptors with each other were examined and those that showed high correlation (R>0.90) were eliminated. Subsequently, stepwise multilinear regression method has been performed to select the most relevant descriptors by SPSS statistics software (Ver. 17.0) [25]. The selected descriptors are: the highest occupied molecular orbital energy (εHOMO), the number of hydroxyl groups (nOH) and harmonic oscillator model of aromaticity (HOMA). These descriptors were used as inputs for developing of multiple linear regressions (MLR) and artificial neural network (ANN) models.

2.3. Artificial neural network

Artificial neural networks are mathematical methods that inspired by studies on biological nervous systems. One of the most significant advantages of ANNs is that these networks can learn from examples and make predictions for new situations. They have seen an explosion of interest over the last few years and are being successfully applied across an extraordinary range of complex and nonlinear problem domains such as geology, chemistry, medicine, physics and weather forecasting [26-30]. Detailed accounts of ANNs theory has been extensively described elsewhere [31-34]. Here, only a brief description of ANNs principle was given. ANNs are composed of simple processing elements (nodes or neurons) operating in parallel which are organized in some layers. Among many types of ANN, the multi-layer perceptron (MLP) associated with the back-propagation algorithm, showed the best performance. The back-propagation network receives a set of inputs, which are multiplied by each node and then a nonlinear transfer function is applied. The MLP should have at least three layers including input, hidden and output layers. The input layer simply serves to introduce the values of the input variables, which are selected descriptors by stepwise variable selection procedure in this investigation. Therefore, this network had three nodes in the input layer. The number of nodes in the output layer was set to be one. This layer handles the output from the network, which is redox potential in this study. The most important stage in designing MLP is to optimize the number of nodes in the hidden layer. After a series of trials, a MLP with four nodes in hidden layer is found to have the best performance. Then the network was trained using the training set. During the training, the internal test set was used to prevent of over training. The goal of training is to change the weights between the layers in a direction to minimize the output errors. In order to evaluate the performance of network, the root mean square error (RMSE) in prediction for training and internal test sets were examined and monitored during the

learning epochs. Then the performance of the network can be tested with a verification set.

3. Results and discussion

3.1. Linear modeling

Multiple linear regressions have been found as a useful technique in modeling especially within the areas of QSPR. A set of approximately 2000 molecular descriptors were calculated by means of CODESSA and DRAGON softwares. Prior to model generation, the correlation between computed descriptors was obtained and the numbers of significant descriptors reduced according to the criteria mentioned in the Section 2.2.

In order to acquire a reliable equation for structure and property relationship, feature selection based on stepwise multiple linear regression was performed on each different groups of descriptors (constitutional, topological, geometrical, electrostatic and quantum chemical descriptors) to choose the high informative ones.

Afterwards the total selected descriptors were transferred to a spreadsheet to select the ultimate descriptors with the application of stepwise variable selection method.

Table 2 Specifications of MLR model

Descriptors	Notation	Coefficient	Standard error
Highest occupied molecular orbital energy	εHOMO	-0.361	0.032
Number of hydroxyl groups	nOH	-0.145	0.018
Harmonic oscillator model of aromaticity	HOMA	1.891	0.568
Constant		-4.016	0.452

n=32, R=0.960, F=108.318, SE=0.065

Table 2 shows the name of selected descriptors as well as specifications of the developed MLR model. As can be seen from this table three descriptors appeared in the MLR model, which are the highest occupied molecular orbital energy (εHOMO), the number of hydroxyl groups (nOH) and harmonic oscillator model of aromaticity (HOMA) index. These descriptors encode different topological and quantum-chemical aspects of molecular structure. Comprehensive details about the chemical meaning and the calculation procedure of these descriptors were explained adequately in Molecular Descriptors for Chemoinformatics by Todeschini et al. [35]. The assessment of the generated model was performed by applying leave-one-out cross validation test, which produce the statistical parameters of Q^2=0.625 and SPRESS=0.0113. These values reveal the predictivity and robustness of constructed linear model. Moreover, the y-scrambling test was carried out on the data set to investigate any chance correlation in our modeling. The mean value of R^2 after 30 times reiteration and modeling was 0.386, which suggests that the well-founded model is not merely the result of pure chance or structural dependency in data set.

3.2. Non-linear modeling

Artificial neural network was applied as a nonlinear feature mapping technique to investigate probable nonlinear relationship between selected molecular descriptors and E_7 of phenolic antioxidants. The input vectors were descriptors that appeared in the MLR model and the signal of the output node represented the redox potential of the interested compounds. Thus, this network has three nodes in the input layer and one node in the output layer. For the purpose of the present application, a MLP-NN structure has been trained with a back-propagation training procedure by means of STATISTICA software (Ver .7.0) [36]. Before training the network the parameters of the nodes in the hidden layer, learning rate and momentum were optimized. The procedure for optimization of these parameters was explained in our previous works [37-38]. The architecture and specifications of the optimized network were summarized in Table 3.

Table 3 The specifications and topology of optimized ANN

Parameter	Value
No. of nodes in the input layer	3
No. of nodes in the hidden layer	4
No. of nodes in the output layer	1
Learning rate	0.01
Momentum	0.4
Transfer functions	Sigmoid

The training continues until the differences between predicted E_7 and the target values were minimized. The iterative training algorithm progresses through a number of 12000 epochs but the best network which has the lowest error on internal test set was obtained at 5073 epoch. Then the trained network was applied to predict the E_7 for compounds in external test set as well as training and internal test sets.

The predicted values of E_7 for these sets were shown in Table 1. Root mean square error value is a fine validity criterion to estimate the performance of the obtained models. The values of RMSE for MLP-NN model were; 0.049, 0.075 and 0.043 for training, internal and external tests sets respectively, while these values were; 0.061, 0.088 and 0.073 respectively for MLR model. Other statistical parameters including correlation coefficient, average

absolute error (AAE), average error (AE) and standard deviation (SD) for these models were shown in Table 4. Comparison between these parameters reveals that, nonlinear MLP-NN model produced better results with good predictive ability than linear model.Fig. 1 shows the plot of the predicted values versus the experimental values of E_7 for all molecules in the data set, which indicate the good correlation between estimated and experimental E_7. In addition, the residuals of this calculation are plotted against the experimental values of E_7 in Fig. 2. The propagation of the residuals on both sides of zero line indicates that no systematic error exists in developed neural network model.

Table 4 The statistical parameters of developed MLR and ANN models

Parameter	**MLR**			**ANN**		
	Training set	Internal set	External set	Training set	Internal set	External set
RMSE	0.061	0.088	0.073	0.049	0.075	0.043
AAE	0.053	0.062	0.058	0.04	0.06	0.034
AE	0.001	-0.038	-0.034	0.013	-0.032	-0.01
SD	0.062	0.098	0.082	0.05	0.084	0.048
R	0.960	0.959	0.962	0.975	0.954	0.985

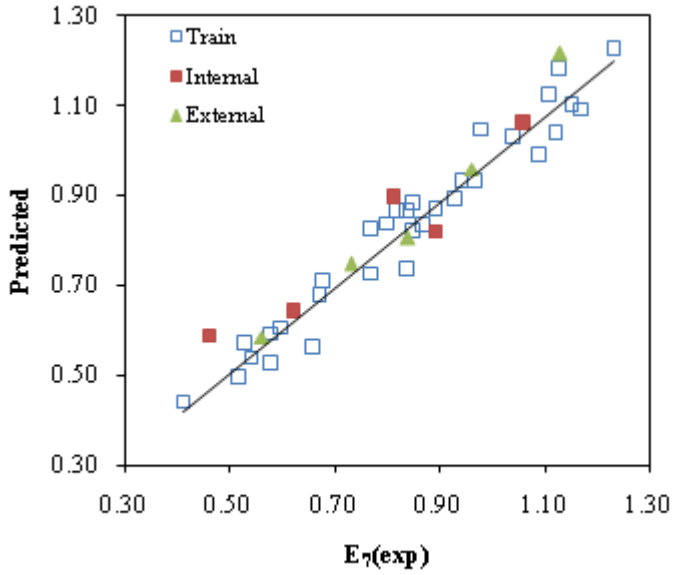


Fig. 1. Plot of predicted vs. experimental values of E_7 by ANN model

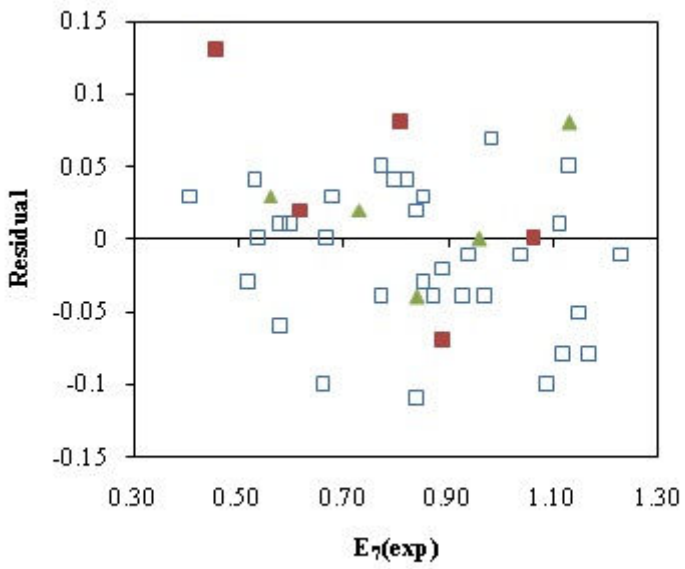


Fig. 2. Plot of residual vs. experimental values of E_7 by ANN model (Legends as Fig. 1)

3.3. Descriptors interpretation

For inspection of the relative importance and contribution of each descriptor in the model, sensitivity analysis was conducted on the inputs of neural network [39].

Sensitivity analysis rates variables according to the deterioration in modeling performance. The basic sensitivity figure is the error ratio, which tells how sensitive the output is to a perturbation of the input. The reported error ratio is the ratio of the error with the variable unavailable (equal to zero) to the ratio with its available. Important variables have a high error ratio, indicating that the network performance deteriorates badly if they are not present. If the error ratio is one or lower, then making the variable unavailable either has no effect on the performance of the network, or actually enhances it. These steps performed on the developed network. According to the results of this test, the relative importance rank of descriptors is: εHOMO > nOH > HOMA. Therefore, the εHOMO descriptor with the maximum error ratio has the major role to correlate chemical structure of phenolic antioxidants to E_7.

The εHOMO is the energy of the highest occupied molecular orbital and encodes quantum chemical aspect of molecular structure. Molecules with higher εHOMO values can donate their electrons more easily compared to molecules with lower εHOMO values, and hence are more reactive. In agreement with the Koopmans theorem, the εHOMO descriptor is related to the ionization potential (IP= -εHOMO) and has significant effect in oxidation/reduction process, in particular for radical reactions [35, 40]. The second significant descriptor is the number of hydroxyl group, which is a topological descriptor and reflecting the molecular composition. As stated in the introduction section, abstracting hydrogen atom from the phenolic hydroxyl group is the first step to initialize the reaction of phenolic antioxidants with $ROO^{\cdot}$ radical. Therefore increasing the number of hydroxyl groups enhances this reaction. Harmonic oscillator model of aromaticity index, which is a

delocalization degree index, has the third level of importance and accounts the π-electron mobility in a molecule and therefore can affects on electron donation by a molecule. The appearance of these descriptors in developed QSPR models reveals that electronic aspects of phenolic derivations have an effective role in their antioxidants activities.

4. Conclusion

In the present study, MLR and ANN were used as linear and non-linear feature mapping techniques to correlate the redox potential of some phenolic antioxidants to their calculated molecular descriptors. The calculated statistical parameters of these models revealed that ANN was better than MLR, which means that there are some non-linear relations between selected molecular descriptors and redox potential of phenolic antioxidants. Descriptors, which appeared in these models, were quantum chemical and topological types and are highly informative about E_7 of phenolic antioxidants. Besides, the successful results of this investigation signified that QSPR method based on semi-empirical calculated molecular descriptors is appropriate enough to predict the redox potential of phenolic antioxidants.

5. References

[1] M. S. Cooke, M. D. Evans, M. Dizdaroglu, J. Lunec, FASEB J. 17 (2003) 1195.

[2] J. S. Wright, E. R. Johnson, G. A. DiLabio, J. Am. Chem. Soc. 123 (2001) 1173.

[3] W. Bors, W. Heller, C. Michel, M. Saran, Methods in Enzymol. 186 (1990) 343.

[4] B. Halliwell, Drugs Aging 18 (2001) 685.

[5] M. Valko, D. Leibfritz, J. Moncol, M. T. D. Cronin, M. Mazur, J. Telser, Int. J. Biochem. Cell. Biol. 39 (2007) 44.

[6] H. Sies, Exp. Physiol., 82 (1997) 291.

[7] W. J. Xin, B. L. Zhao, X. J. Li, J. W. Hou, Res. Chem. Intermed. 14 (1990) 171.

[8] S. Tafazoli, J. S. Wright, P. J. O'Brien, Chem. Res. Toxicol. 18 (2005) 1567.

[9] W. M. Indrasena, C. J. Barrow, in: C. Alasalvar, F. Shahidi, K. Miyashita, U. Wanasundara (Eds.), Handbook of Seafood Quality, Safety and Health Applications, Wiley-Blackwell, Oxford, 2010.

[10] M. T. D. Cronin, in: T. Puzyn, J. Leszczynski, M. T. D. Cronin (Eds.), Recent Advances in QSAR Studies Methods and Applications, Springer, London, 2010.

[11] N. Singh, R. J. Loader, P. J. O'Malley, P. L. A. Popelier, J. Phys. Chem. A 110 (2006) 6498.

[12] A. I. Khlebnikov, I. A. Schepetkin, N. G. Domina, L. N. Kirpotina, M. T. Quinn, Bioorg. Med. Chem. 15 (2007) 1749.

[13] V. Rastija, M. Medic-Saric´, Eur. J. Med. Chem. 44 (2009) 400.

[14] Z. Cheng, J. Ren, Y. Li, W. Chang, Z. Chen, Bioorg. Med. Chem. 10 (2002) 4067.

[15] R. Bosque, J. Sales, J. Chem. Inf. Comput. Sci., 43 (2003) 637.

[16] B. F. Rasulev, N. D. Abdullaev, V. N. Syrov, J. Leszczynski, QSAR Comb. Sci. 24 (2005) 1056.

[17] R. M. V. Abreu, I. C. F. R. Ferreira, M. J. R. P. Queiroz, Eur. J. Med. Chem., 44 (2009) 1952.

[18] M. Reis, B. Lobato, J. Lameira, A. S. Santos, C. N. Alves, Eur. J. Med. Chem., 42 (2007) 440.

[19] E. J. Lien, S. Ren, H-H. Bui, R. Wang, Free Radic. Biol. Med. 26 (1999) 285.

[20] A. R. Katritzky, V. S. Lobanov, M. Karelson, CODESSA: Comprehensive Descriptors for Structural and Statistical Analysis, Reference

Manual, Version 2.0, 1994.

[21] A. R. Katritzky, V. S. Lobanov, M. Karelson, Chem. Soc. Rev. 24 (1995) 279.

[22] A. R. Katritzky, V. S. Lobanov, M. Karelson, Pure Appl. Chem., 69 (1997) 245.

[23] R. Todeschini, V. Consonni, A. Mauri, M. Pavan, DRAGON Software, Version 3.0, Milano Chemometrics and QSAR Research Group, Milano, Italy, 2003.

[24] J. J. P. Stewart, J. Comput.-Aided Mol. Des. 4 (1990) 1.

[25] SPSS Statistics Software, Version 17.0, SPSS Inc. Chicago, 2008.

[26] H. Dai, C. MacBeth, Geophys. J. Int., 120 (1995) 758.

[27] H. Golmohammadi, Z. Dashtbozorgi, J. Struct. Chem. 51 (2010) 833.

[28] P. J. G. Lisboa, Neural Networks 15 (2002) 11.

[29] N. P. Barradas, A. Vieira, Phys. Rev. E 62 (2000) 5818.

[30] M. N. French, W. F. Krajewski, R. R. Cuykendall, J. Hydrol. 137 (1992) 1.

[31] S. Haykin, Neural Network, Prentice-Hall, Englewood Cliffs, NJ, 1994.

[32] N. K. Bose, P. Liang, Neural Network Fundamentals, McGraw-Hill, New York, 1996.

[33] M. T. Beal, H. B. Hagan, M. Demuth, Neural Network Design, PWS Pub. Co., Boston, 1996.

[34] M. Jalali-Heravi, M. H. Fatemi, J. Chromatogr. A 825 (1998) 161.

[35] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors for Chemoinformatics, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2009.

[36] STATISTICA (Data Analysis Software System), Version 7.0, StatSoft. Inc., Tulsa, 2004.

[37] M. Jalali-Heravi, M. H. Fatemi, J. Chromatogr. A 897 (2000) 227.

[38] M. Jalali-Heravi, M. H. Fatemi, Anal. Chim. Acta 415 (2000) 95.

[39] J. Mira, F. Sandoval, From Natural to Artificial Neural Computation, Vol. 930, Springer, Berlin, 1995.

[40] P. Politzer, F. Abu-Awwad, Theor. Chem. Acc., 99 (1998) 83.