http://dx.doi.org/10.5012/bkcs.2013.34.6.1643

Nonlinear QSAR Study of Xanthone and Curcuminoid Derivatives as α-Glucosidase Inhibitors

Youcef Saihi,[†,*] Khairedine Kraim,[†,‡] Fouad Ferkous,[†] Zeineddine Djeghaba,[†] Abdelkader Azzouzi,[§] and Sabrina Benouis[†]

[†]*Department of Chemistry, Faculty of Sciences, University of Badji Mokhtar, Annaba, 23000, Algeria*
[*]*E-mail: youcefchem@gmail.com*
[‡]*High School of Technologic Education (ENST) Azzaba-Skikda*
[§]*Department of Chemistry, Faculty of Sciences, University of Djelfa, 17000, Algeria*
Received November 26, 2012, Accepted March 6, 2013

A non linear QSAR model was constructed on a series of 57 xanthone and curcuminoide derivatives as α-glucosidase inhibitors by back-propagation neural network method. The neural network architecture was optimized to obtain a three-layer neural network, composed of five descriptors, nine hidden neurons and one output neuron. A good predictive determination coefficient was obtained (R^2_{Pset} = 86.7%), the statistical results being better than those obtained with the same data set using a multiple regression analysis (MLR). As in the MLR model, the descriptor MATS7v weighted by Van der Waals volume was found as the most important independent variable on the α-glucosidase inhibitory.

Key Words : α-Glucosidase, Inhibitors, Xanthone-curcuminoide derivatives, QSAR, Artificial neural networks

Introduction

α-Glucosidase are membrane-bound enzymes located at the epithelium of the small intestine[1] and the key enzymes of carbohydrate digestion. It specifically hydrolyzed the α-glucopyranoside bond, thereby releasing an α-glucose from the non reducing end of the sugar.[2]

Recently there had been wide spread interest in these enzymes, partly because of their potential as therapeutic targets, especially, the inhibition of α-glucosidase had been found to contribute to block the viral infection[3,4] with human immunodeficiency virus I (HIV-I)120. Also, α-glucosidase inhibitors can help to control postprandial blood glucose levels in diabetic patients,[5,6] thus clinical trials should that the α-glucosidase inhibitor improved long term glycemic control as measured by decreased hemoglobin in patients with type II diabetes and delay the development of type II diabetes in patients with impaired glucose tolerance.[7]

To date, many α-glucosidase inhibitors have been reported, such as acarbose and voglibose from microorganisms, and 1-deoxynojirimycin isolated from plants. However, they are confined to glucosidic derivatives.[8-11]

QSAR (Quantitative Structure-Activity Relationship) studies are a powerful method for the design of bioactive compounds and the prediction of activity according to the physical and chemical properties.[12,13] However, a limited number of studies have been devoted to the search for new α-glucosidase inhibitors using QSAR.[14-16] Many different chemometrics methods, such as Multiple Linear Regression (MLR), different types of Neural Networks (NN), Genetic Algorithms (GAs) can be employed to derive correlation models between molecular structures and properties.

In many cases, such as for biological properties, a simple linear relationship will not result in good predictive performance. In contrast, ANN (Artificial Neural Networks) are able to recognize highly non-linear relationships. The flexibility of ANN enables them to discover more complex relationships in experimental data, comparing with the MLR approach. Hence, the ANN provides proper analytical alternatives to conventional techniques and interesting approaches to the QSAR and QSPR studies.[17-20]

To the best of our knowledge, this is the first QSAR study using ANN for the prediction of pIC_{50} α-glucosidase inhibitors. Therefore, the aim of this study is to investigate the applicability of ANN in relating α-glucosidase inhibitory of 57 xanthone and curcuminoide derivatives to molecular descriptors and to compare the results with those obtained by a MLR using the same set of compounds. To fully investigate the performance of a selected descriptor subset, a variety of ANN architectures must be considered. This is carried out by developing models with the same set of input descriptors but varying architectures.

Experimental

Biological Data Set. The data set consists of 57 recently discovered xanthone and curcuminoid derivatives as α-glucosidase inhibitors and reported by Liu[21] and Du.[22] Their molecular structures and *in vitro* activities are listed in Chart 1, Chart 2 and Table 1. Activities were converted into the corresponding -logIC_{50} values (pIC_{50}), where IC_{50} is the effective concentration of compound required to achieve 50% of inhibition of α-glucosidase (experimental pIC_{50} values ranging from −2.371 to −0.204).

Molecular Modeling. The data set is split into three sets for model building, validation and prediction purposes. They

Xanthone Derivatives

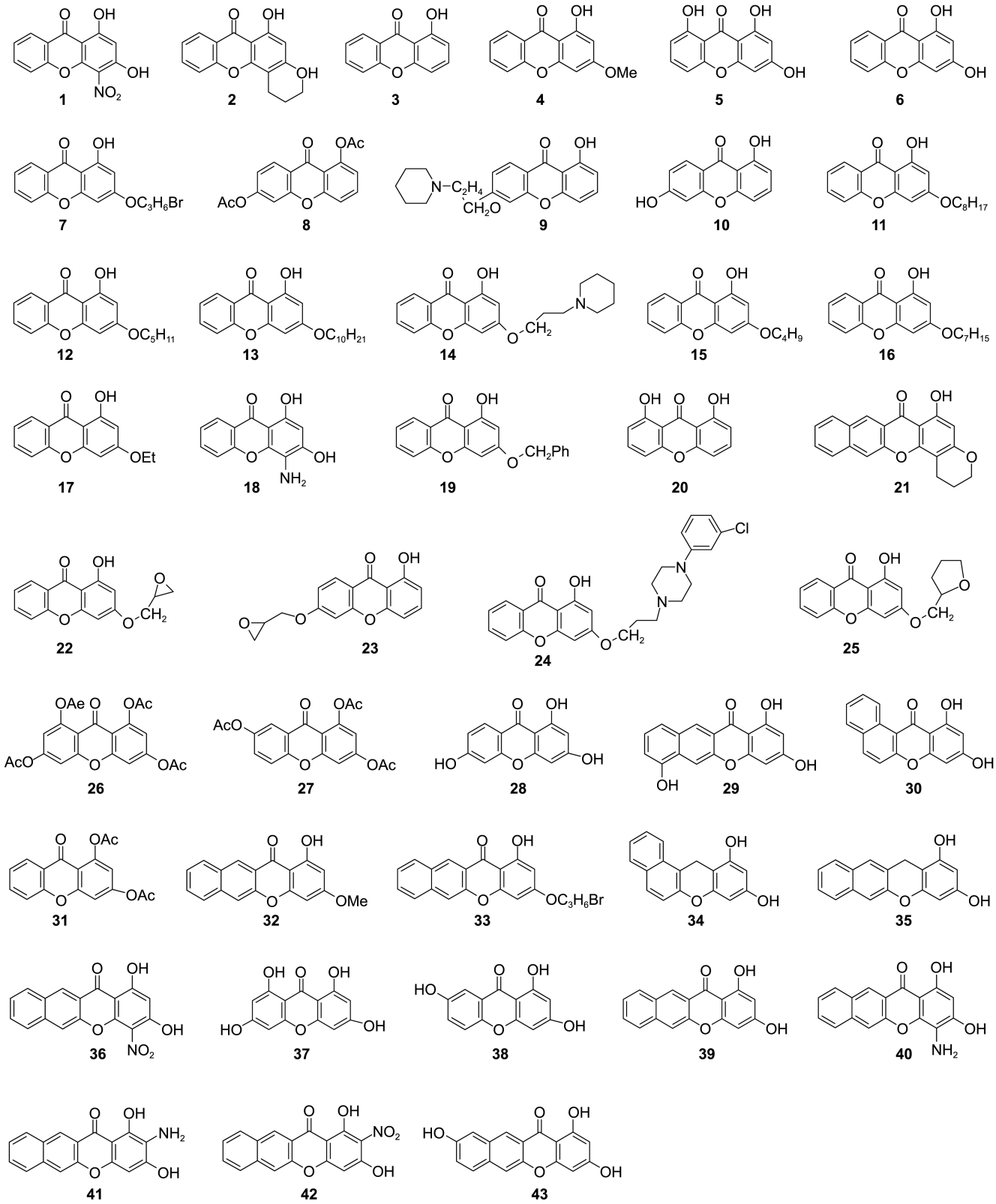

Chart 1

are the training set (TSET), validation set (VSET), and external prediction set (PSET). The TSET contains 50% of the compounds and is used for model building. For model validation, 25% of the compounds are assigned as the VSET. The remaining 25% of the compounds are assigned as the PSET. The following steps were performed in order to cover

Curcuminoid Derivatives

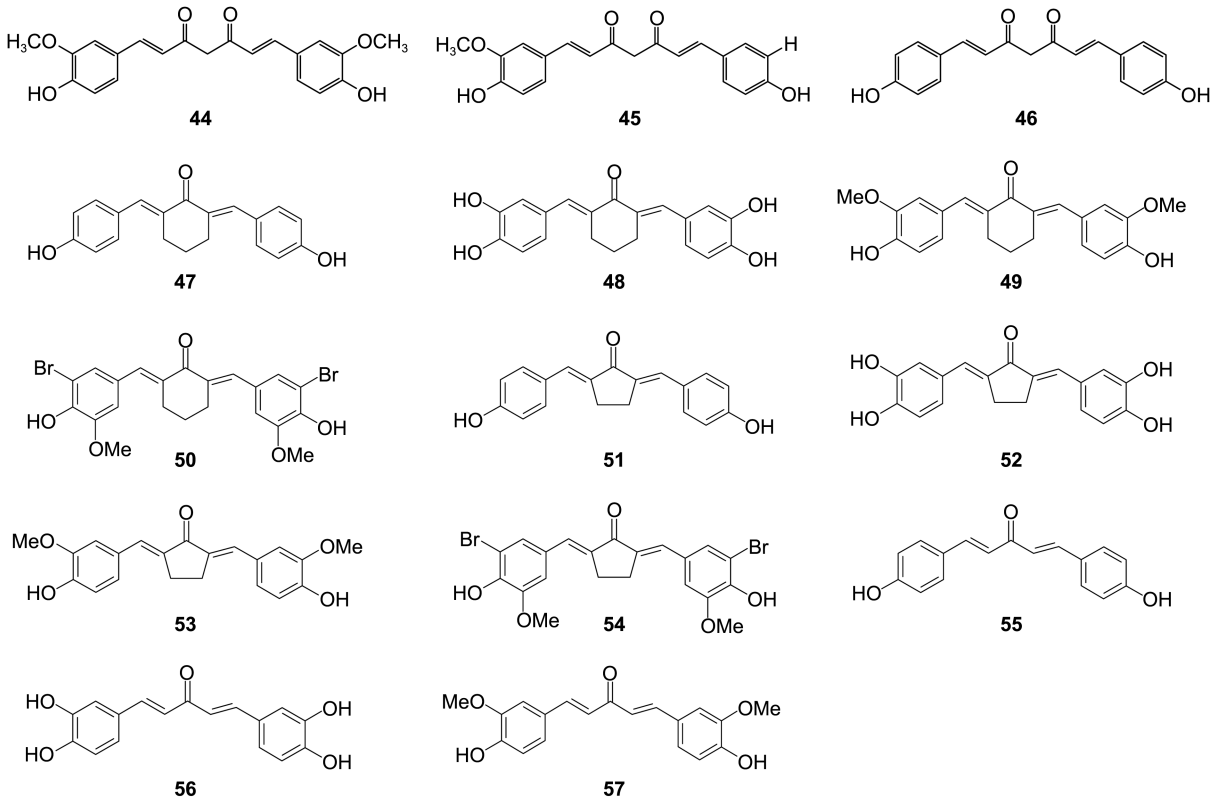

Chart 2

the same ranges of activities and the same distributions for the training and prediction data sets.

1. All compounds were ordered in ascending order of their pIC_{50} values.

2. Every second and fourth compounds were selected to become a part of the validation and test sets respectively.

3. All the remaining compounds formed the training set.

Nonlinear ANN Approach. Artificial neural networks are employed to solve numerous types of problems in engineering and science. The theory of ANN was described previously.[23] Briefly, artificial neural networks consist of layers of which outputs are connected to the other neurons. While there are many different artificial neural networks architecture, the most popular network used in QSAR is three layer feed-forward networks.[24-26]

The backpropagation (BP) algorithm is one of the most popularly applied to feed-forward training of neural networks, due to its simplicity, its capacity to extract useful information from samples and to store it implicitly as weights over their connections. In this study, the backpropagation neural network was applied and trained with Levenberg-Marquardt algorithm. Due to the nonlinear input-output dependency the transfer function was chosen sigmoid (1).

$$sf(input) = \frac{1}{1+e^{-input}} \quad (1)$$

That limits the neuron's output signal to values between 0 and 1. The output layer neurons usually have sigmoidal or linear transfer functions, depending on the application. The whole network represents a non-linear relationship which can be written for each output as (2):

$$\hat{y} = f(x) = \Sigma_h[sf(\Sigma_i x_i w_{ih-} \theta_i)]w_{h-}\theta_h \quad (2)$$

Where

$\hat{y}$: the predicted target value.

w_{ih} : is the connection weight between the input node i with the hidden node h.

w_h : are the connection weights between each hidden node h with the final output considered, y (desired output).

θ_i and θ_h are the biases corresponding to the input and hidden layers.

In this work, all units were fully interconnected and the input-to-output information flow was feed-forward (no feed-back connections). The number of neurons in the input layer was set equal to the number of descriptors, while the number of neurons in the output layer was set equal to 1 (Figure 1).

The weights and biases of all connections were randomly initialized, and then iteratively adjusted during training, according to Levenberg-Marquardt algorithm, to minimize the mean square error (the performance function) between outputs and target, that is given by the difference between $\hat{y}$ and y, which is subsequently back-propagated to modify the weights in order to attain the best fit (3).

Table 1. Experimental pIC_{50} values and calculated by ANN and MLR models for the 57 compounds listed in Chart 1 and Chart 2

Compound	Observed pIC_{50}	Predicted pIC_{50}	
		Multilinear model	Nonlinear ANN model
1	-2.371	-2.071	-2.134
2	-2.297	-2.086	-2.306
3	-2.249	-1.812	-2.059
4	-2.238	-2.162	-2.306
5	-2.206	-2.024	-2.065
6	-2.166	-1.974	-2.03
7	-2.143	-2.026	-1.941
8	-2.123	-1.94	-2.03
9	-2.119	-2.06	-2.072
10	-2.114	-2.012	-2.052
11	-2.092	-2.105	-2.089
12	-2.082	-2.139	-2.052
13	-2.063	-2.131	-2.073
14	-2.062	-2.182	-2.131
15	-2.056	-2.218	-2.156
16	-2.045	-2.130	-2.131
17	-2.01	-2.163	-2.155
18	-1.992	-1.609	-1.511
19	-1.961	-2.042	-2.057
20	-1.913	-2.035	-1.511
21	-1.828	-1.707	-1.827
22	-1.823	-1.738	-1.589
23	-1.803	-1.925	-1.694
24	-1.791	-1.724	-1.589
25	-1.724	-1.495	-1.658
26	-1.696	-1.564	-1.578
27	-1.667	-1.571	-1.556
28	-1.618	-2.026	-1.73
29	-1.601	-1.190	-1.612
30	-1.543	-1.589	-1.489
31	-1.504	-1.758	-1.799
32	-1.496	-1.718	-1.337
33	-1.473	-1.744	-1.714
34	-1.444	-1.215	-1.276
35	-1.303	-1.009	-1.149
36	-1.233	-1.288	-1.246
37	-1.167	-1.991	-1.149
38	-0.968	-1.442	-1.184
39	-0.919	-1.401	-1.168
40	-0.903	-1.058	-0.867
41	-0.799	-0.873	-1.168
42	-0.771	-1.052	-0.785
43	-0.763	-0.723	-0.721
44	-1.571	-1.371	-1.489
45	-1.63	-1.231	-1.612
46	-1.362	-1.529	-1.471
47	-1.575	-1.618	-1.553
48	-0.447	-0.694	-0.469
49	-1.672	-1.764	-1.713
50	-1.467	-1.434	-1.471
51	-1.512	-1.236	-1.337
52	-0.415	-0.556	-0.721
53	-1.723	-1.735	-1.713
54	-1.53	-1.479	-1.506
55	-1.338	-1.492	-1.271
56	-0.204	-0.530	-0.224
57	-1.568	-1.384	-1.482

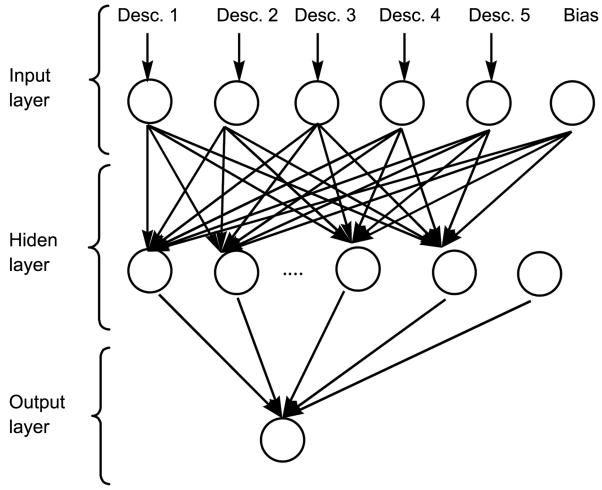

Figure 1. Schematic diagram of artificial neural network used in this work.

$$E = \frac{1}{2}\Sigma_{p=1}^{p}\Sigma_{j}^{J}(t_{pj-}\ o_{pj})^2 \quad (3)$$

Where t_{pj} is the target value (desired output: y) and o_{pj} is the actual output (predicted target value: $\hat{y}$).

All statistical calculations and ANN implementation were done on personal computer using Minitab (version 15.1.0.0)[27] and Matlab[28] version 7.0 software respectively.

Results and Discussion

In our previous work, we established the structure-inhibitory activity (pIC_{50}) of 57 Xanthone and Curcuminoid derivatives against α-glucosidase using Multiple Linear Regression analysis (MLR) coupled with genetic algorithms as features selection approach. Five descriptors were selected as important factors for α-glucosidase inhibitory. Between these descriptors there are three, MATS7v, R4e+ and nArOR are weighted by atomic Van der Waals volumes, atomic Sanderson electro-negativities and number of aromatic ethers, respectively, and two descriptors unweighted (Mor15u and H5u) Table 2.

The equation that describes the model relative to five variables selected using GA-MLR technique is shown with its statistical parameters (4).

$$pIC_{50} = -2.01 + 1.17\ MATS7v + 0.485\ Mor15u + 0.414\ H5u + 6.05\ R4e^{+} - 0.480\ nArOR \quad (4)$$

Table 2. Glossary of the descriptors reported

Descriptor symbol	Meaning
MATS7v	Moran autocorrelation of lag7 / weighted by atomic Van Der Waals volumes.
Mor15u	3D MoRSE-signal15 / unweighted.
H5u	H autocorrelation of lag5 / unweighted.
R4e+	R maximal autocorrelation of lag4 / weighted by Sanderson lcectronegativities.
nArOR	Number of ethers (aromatic)

N = 44; R^2 = 0.857; S = 0.197; F = 45.53; P < 10^{-4}; $Q^2_{cv\text{-}loo}$ = 0.815; $s_{cv\text{-}loo}$ = 0.208; Q^2_{Pset} = 0.66.

Where N is the number of compounds included in the model, s is the standard deviation of the regression, R^2 is the determinationd coefficient, F is the Fischer ratio, $s_{cv\text{-}loo}$ and $Q^2_{cv\text{-}loo}$ are the cross-validation standard deviation, and determination coefficient respectively, Q^2_{Pset} is the external determination coefficient.

In the current study, an ANN-QSAR model is presented for pIC_{50} of 57 xanthone and curcuminoid derivatives against α-glucosidase enzyme. The networks were generated using the five descriptors appeared in the MLR model as their inputs and pIC_{50} as their output target. Before starting, data set was separated into three groups: training, validation and test sets. The training set, consisted of 29 molecules, was used for the model generation. However the validation set, consisted of 14 molecules, was used to take care of the overtraining. The prediction set, consisted of 14 molecules, was used to evaluate and generalize the generated model.

As can be seen from Figure 2, the above given procedure ensures that the training, validation, and prediction data sets have similar data distribution. Where the response histogram (in the training, validation, and test sets) highlighting that the entire range of response is covered in the three sets in a way that is quite well balanced.

The range of the response in the training set is −2.375 to −0.125 log unit (pIC_{50} mean = −1.606), while the range of validation set response is −2.3 to −0.7 log unit (pIC_{50} mean = −1.667), and the range of test set response is −2.125 to −0.375 log unit (pIC_{50} mean = −1.589).

The ANN architecture used in this study is made up of three layers: input layer, hidden layer, and output layer comprising five neurons (MATS7v, Mor15u, H5u, R4e+, and nArOR) and a bias in the input layer and one output neuron pIC_{50}. Signals are sent from the input layer through the hidden layer and finally to the output layer in a feed-forward manner.

This is followed by readjustment of weights according to the prediction error. The connection weights were randomly initialized and modified during the training process using the backpropagation algorithm.

Table 3. Determination coefficients values with different numbers of hidden neurons

N° Neurons	MSE	R^2tr	R^2val	R^2_{Pset}
1	0.028	0.882	0.755	0.769
2	0.024	0.893	0.835	0.769
3	0.013	0.945	0.748	0.760
4	0.002	0.986	0.702	0.787
5	0.008	0.927	0.835	0.728
6	0.000	0.994	0.726	0.832
7	0.008	0.947	0.783	0.803
8	0.000	1.000	0.755	0.731
9	**0.004**	**0.941**	**0.869**	**0.867**
10	0.003	0.960	0.843	0.711
11	0.003	0.962	0.766	0.783
12	0.000	0.990	0.712	0.704

MSE: mean square error; R^2tr, R^2val, R^2_{Pset}: a determination coefficient for training, validation, and prediction sets respectively.

One major problem in neural networks is how to determine the number of neurons in the hidden layer. Though there is no rigorous rule to rely on,[29,30] some authors[31] have proposed a parameter ρ, to determine the number of hidden units,[31] defined in Eq. (5).

$$\rho = \frac{\text{Number of data points in the training set}}{\text{Number of connections in the network}} \quad (5)$$

A second practical way is to use a trial and error procedure (used in this work). The approach consists in using either a bottom-up strategy, starting with too few neurons and then adding more if need be, or a top-down. The number of hidden neurons was optimized from 1 to 12 by this approach. The results are depicted respectively in Table 3 and Figure 3.

During training the weights and biases of the network are iteratively adjusted to minimize the performance function MSE (mean square error) between the network outputs and the target outputs for the training set. As shown in Table 2, this fitness function yields models which are optimal for the

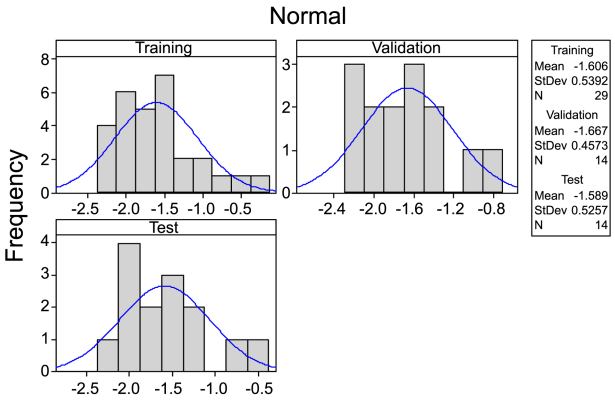

Figure 2. Plot of the distribution of pIC_{50} values for training, validation, and prediction sets.

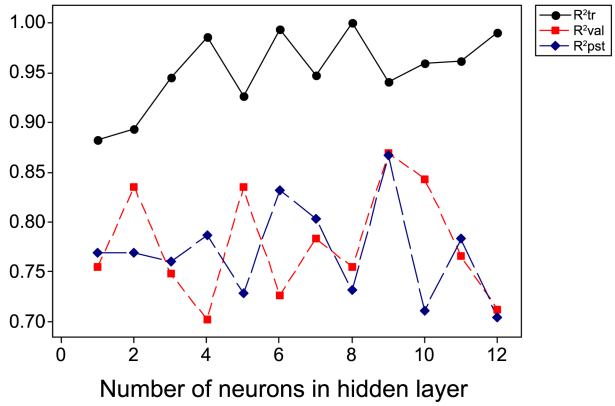

Figure 3. Influence of the number of neurons in hidden layer on R^2 (for training, validation, and prediction sets).

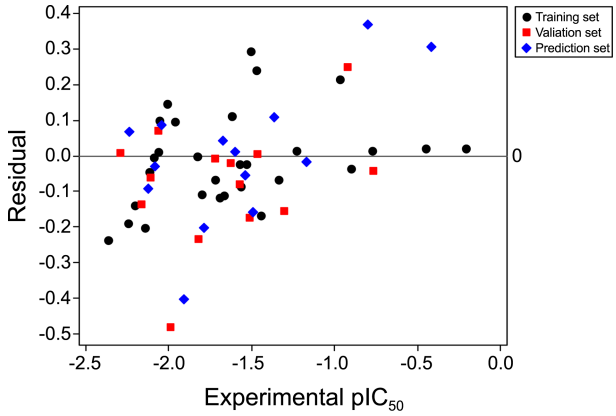

Figure 4. Residual versus experimental values of pIC_{50} for non-linear model.

training data (*e.g.* for ANN 5-8-1: $R^2tr=1$ with MSE=0) but they do not have good predictive abilities (R^2val and R^2test are respectively 0.755 and 0.731) with the same ANN architecture. For this reason, we used the validation and extern determination coefficients as a relative measure of the predictive performance, which are defined as the degree of correlation between the predicted and experimental pIC_{50} values for validation and predictions sets respectively. As shown in the Figure 3, the optimal model corresponds to 9 neurons in the hidden layer, where the determination coefficient for validation set and prediction set go above 86%. This non linear model (5-9-1) is able to account 94.1% of the variance of training set by the five descriptors with 0.004 as MSE value.

The residuals of the ANN calculated values of the α-glucosidase inhibitory are plotted against the experimental pIC_{50} values in Figure 4. On the both sides the distribution of the residuals on zero line indicates that no systematic error exists in the constructed ANN-QSAR model.

To show that the results obtained by the ANN model were not due to chance correlation, a randomizing experiment was performed. The dependent variables pIC_{50} (TSET, VSET, and PSET) were randomly shuffled keeping the ANN parameters (weights, biases, and input matrix) fixed (Supplementary data). This operation is repeated ten times and the

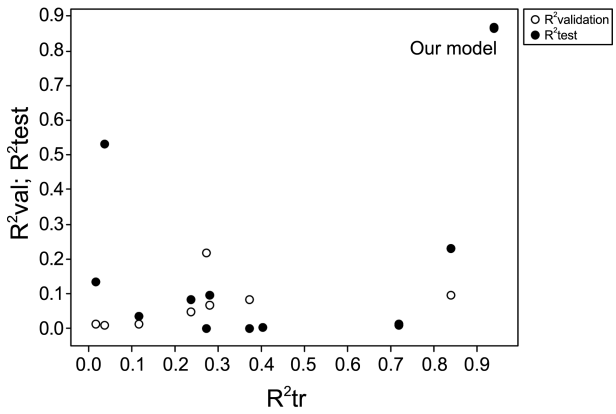

Figure 5. Determination coefficients for validation and test values against training values in y-randomization test.

Table 4. Determination coefficients and root means squares errors obtained by MLR and ANN models

	MLR		ANN	
	RMSE	R^2	RMSE	R^2
Training set	0.156	0.857	0.129	0.941
Prediction set	0.194	0.660	0.188	0.867

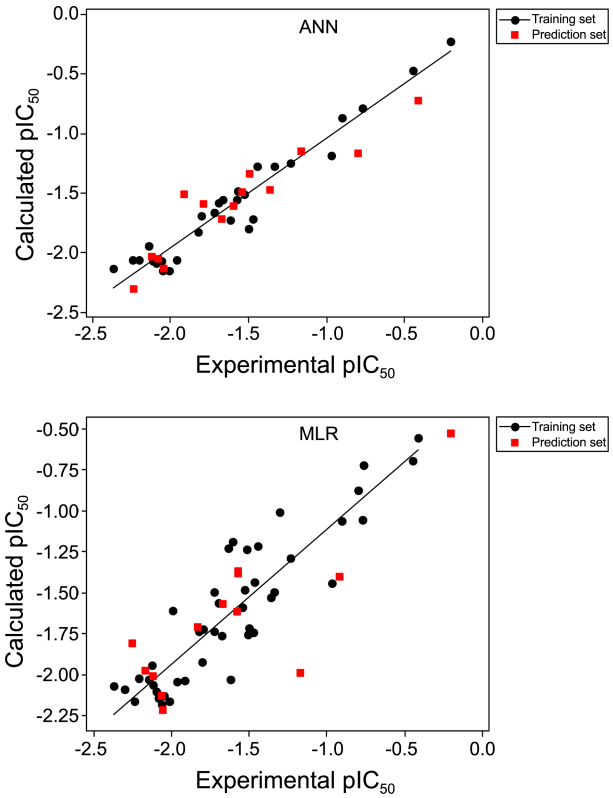

Figure 6. Calculated versus experimental pIC_{50} for α-glucosidase inhibition by xanthone and curcuminoide derivatives, obtained by ANN and MLR.

results are depicted in Figure 5. The results obtained expressed by the low values of R^2 (for training, validation, and prediction sets) show that the best model was extremely unlikely to have been found because of chance correlation effects or structural dependency of the training set.

In comparison of statistical results, obtained by ANN model and same obtained with MLR model (Table 4), we are plotted, the experimental versus calculated values of pIC_{50}, obtained by ANN and MLR models, are plotted (for the training and prediction sets) Figure 6.

Based on the statistical results given by linear and non-linear models and according to the results shown in Figure 6 we observe that the ANN model has a reliability and predictive capability significantly better than those given by the MLR model. The standard errors of calculation are lower and the determination coefficients are higher with ANN than with MLR. We therefore consider that the nonlinear approach is best suited to analyze the structure-inhibitory activity relationship of Xanthone and Curcuminoid derivatives

Table 5. The contribution C_i of each descriptor in ANN model

	ΔpIC_{50}				
	MATS7v	Mor15u	H5u	R4e+	NArOR
Δmi	0.299	0.238	0.194	0.235	0.172
Ci	**26.271**	20.904	17.046	20.664	15.114

Δmi : the mean of the absolute deviation values; *Ci:* the contribution of each descriptor.

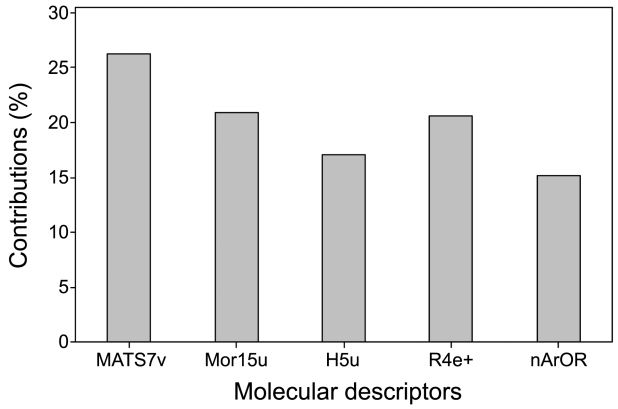

Figure 7. Contribution of the selected five descriptors for the nonlinear model.

against α-glucosidase.

Analysis of Descriptor's Contribution in ANN Model: As a final step, we studied the effect of the five descriptors on the α-glucosidase inhibitory. For this purpose and for each descriptor, we removed the descriptor under study from the input matrix and reinitialized the model keeping the weights matrix fixed, then we calculate the mean of the absolute deviation values Δm_i between the observed and predicted value for all compounds. Finally, the contribution C_i of each descriptor is given by (6).[32-34]

$$C_i = \frac{100x\Delta m_i}{\Sigma \Delta m_i} \quad (6)$$

The results are depicted in Table 5 and Figure 7. As shown in Table 5, the five descriptors are ordered according to their contribution as follow: MATS7v > Mor15u > R4e+ > H5u > nArOR. As in MLR model,[15] the descriptor coded by Van der Waals volume exerts the major effect on the α-glucosidase inhibitory. In second and third order we find two descriptors: Mor15u and R4e+belonging to 3D-MoRSE and 2D-autocorrelation blocks respectively. The last two descriptors (H5u, nArOR) have minor contribution against α-glucosidase inhibitory in comparison with MATS7v Figure 7.

Conclusion

The aim of this study was to present another approach than the regression analysis to develop a QSAR; therefore the five descriptors obtained with GA-MLR method were used in ANN. Our study demonstrated that the three layer back-propagation NN provided us a promising QSAR model for α-glucosidase inhibitors. Based on the ANN results, ANN model bring more reliable statistics according to good predictive determination coefficients, and the five descriptors (MATS7v, Mor15u, H5u, R4e+, and nArOR) matched well with the activity of interest; this means that the descriptors used in MLR model include non-linear relationships.

The contributions of descriptors in the α-glucosidase inhibitory activity are in accordance with those obtained by MLR analysis, where the effect of Van der Waals volume is predominant.

Acknowledgments. The authors are thankful to ATRST (Agence Thématique de Recherche en Sciences & Technologie ex-ANDRU) for the financial support of this work. And the publication cost of this paper was supported by the Korean Chemical Society.

References

1. Din, Z. Y.; Liu, R. R.; Shao, W. Y.; Mao, X. Q.; Ma, L.; Quan, Gu, L.; Huang, Z. S.; Chan, A. S. *Eur. J. Med. Chem.* **2006**, *41*, 213.
2. Melo, E. B.; Gomes, A. S.; Carvalho, I. *Tetrahedron* **2006**, *62*, 10277.
3. Mehta, A.; Zitzmann, N.; Rudd, P. M.; Block, T. M.; Dwek, R. A. *FEBS Lett*. **1998**, *430*, 17.
4. Fischer, P. B.; Karlsson, G. B.; Butters, T. D.; Dwek, R. A.; Platt, F. M. *J. Virol*. **1996**, *70*, 7143.
5. Hwangseo, P.; Kyo, Y. H.; Young, H. K.; Kyung, H. O.; Jae, Y. L.; Keun, K. *Bioorg. Med. Chem. Lett*. **2008**, *18*, 3711.
6. Blicklé, J. F.; Andres, E.; Brogard, J. M. *Rev. Méd. Interne.* **1999**, *20*, 379.
7. Chiasson, J. L.; Josse, R. G.; Gomis, R.; Hanefeld, M.; Karasik, A.; Laakso, M. *Lance.* **2002**, *359*, 2072.
8. Luo, J. G.; Ma, L.; Kong, L. Y. *Bioorg. Med. Chem*. **2008**, *16*, 2912.
9. Yan, L.; Lin, M.; Wen-Hua, C.; Bo, W.; Zun-Le, X. *Bioorg. Med. Chem*. **2007**, *15*, 2810.
10. Gao, H.; Kawabata, J. *Bioorg. Med. Chem*. **2005**, *13*, 1661.
11. Muraoka, O.; Yoshikai, K.; Takahashi, H.; Minematsu, T.; Lu, G. X.; Tanabe, G.; Wang, T.; Matsuda, H.; Yoshikawa, M. *Bioorg. Med. Chem*. **2006**, *14*, 500.
12. Varnek, A.; Tropsha, A. *Chemoinformatics Approaches to Virtual Screening*; Royal Society of Chemistry: United Kingdom., 2008; pp 18-88.
13. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; Wiley-Interscience: New York, 2000; pp 19-71.
14. Liu, Y.; Ke, Z.; Cui, J.; Chen, W. H.; Ma, L.; Wang, B. *Bioorg. Med. Chem*. **2008**, *16*, 7185.
15. Kraim, K.; Khatmi, D.; Saihi, Y.; Ferkous, F.; Brahimi, M. *Chemom. Intell. Lab. Syst*. **2009**, *97*, 118.
16. Jun, X.; Sichao, H.; Haibin, L.; Guoji, L.; Jiaolin, B.; Shaohui, C.; Yuqiang, W. *Int. J. Mol. Sci*. **2010**, *11*, 880.
17. Devillers, J. *Neural Networks in QSAR and Drug Design*; Academic Press: Limited. United Kingdom, 1996; pp 45-62.
18. Sattwa, A. M.; Roy, K. *Eur. J. Med. Chem*. **2009**, *44*, 1509.
19. Yaobin, M.; Bin-Le, L. *Ecotoxicol. Environ. Safe*. **2008**, *71*, 172.
20. Yangali-Quintanilla, V.; Verliefde, A.; Kim, T.-U.; Sadmani, A.; Kennedy, M.; Amy, G. J. *Membr. Sci*. **2009**, *342*, 251.
21. Liu, Y.; Zou, L.; Ma, L.; Chen, W.-H.; Wang, B.; Xu, Z.-L. *Bioorg. Med. Chem*. **2006**, *14*, 5683.
22. Du, Z. Y.; Liu, R. R.; Shao, W. Y.; Mao, X. P.; Ma, L.; Gu, L. Q.; Huang, Z. S.; Chan, A. S. C. *Eur. J. Med. Chem*. **2006**, *41*, 213.
23. Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*; Wiley-VCH: Weinheim., 1999; pp 15-19.
24. Winkler, D. A. *Molecular Biotech*. **2004**, *27*, 139.

25. Nantasenamat, C.; Naenna, T.; Isarankura, C. A.; Prachayasittikul, V. J. *Comput. Aided Mol. Des*. **2005**, *19*, 509.
26. Schneider, G.; Wrede, P. *Prog. Biophys. Mol. Biol*. **1998**, *70*, 175.
27. MINITAB, version 15.1.0.0, Minitab Inc. (http://www.minitab.com).
28. MATLAB, version 7.0, The MathWorks Inc., 2004. (http://www.mathworks.com).
29. Zahouily, M.; Rakik, J.; Lazar, M.; Bahlaoui, M. A.; Rayadh, A.; Komiha, N. *ARKIVOC.* **2007**, *xiv*, 245.
30. Golbraikh, A.; Tropsha, A. *J. Mol. Graph. Model*. **2002**, *20*, 269.
31. Andrea, T. A.; Kalayeh, H. *J. Med. Chem*. **1991**, *34*, 2824.
32. Guha, R.; Jurs, P. C. *J. Chem. Inf. Model*. **2005**, *45*, 800.
33. Zheng, F.; Bayram, E.; Sumithran, S. P.; Ayers, J. T.; Zhan, C. G.; Schmitt, J. D.; Dwoskina, L. P.; Crooks, P. A. *Bioorg. Med. Chem*. **2006**, *14*, 3017.
34. Cherqaoui, D.; Esseffer, M.; Villemin, D.; Cence, J. M.; Chastrette, M.; Zakarya, D. *New J. Chem*. **1998**, *22*, 839.