Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 43 (2008) 569–576
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://www.elsevier.com/locate/ejmech

Original article

QSAR study of neuraminidase inhibitors based on heuristic method and radial basis function network

W.J. Lü, Y.L. Chen, W.P. Ma, X.Y. Zhang, F. Luan, M.C. Liu*, X.G. Chen, Z.D. Hu

Department of Chemistry, Lanzhou University, Lanzhou 730000, Gansu Province, China

Received 9 January 2007; received in revised form 28 March 2007; accepted 19 April 2007
Available online 13 May 2007

Abstract

Neuraminidase (NA) is a critical enzyme of the influenza virus and many inhibitors targeting this enzyme are quite efficient anti-influenza agents. In this paper, quantitative structure–activity relationship (QSAR) method was used to predict the activity of different kinds of 46 NA inhibitors. Heuristic method (HM) and radial basis function network (RBFNN) were first used to build linear and nonlinear models, respectively. The prediction results were in agreement with the experimental value. The proposed model is simple and can be extended to other QSAR investigations.

Keywords: QSAR; Heuristic method; Radial basis function network; Neuraminidase inhibitors

1. Introduction

Influenza is a respiratory infection which has significant morbidity in the general population and mortality in elderly and high-risk patients. Over the last two decades, a number of classes of NA inhibitors have been developed and showed to be somewhat effective in controlling influenza infection in humans [1–8]. Zanamivir [9] and Oseltamivir [10] have been approved by FDA for the treatment and prevention of the influenza, but were recently questioned about recovery rate. Current therapeutic measures have only provided limited control of influenza. Vaccines are frequently ineffective because influenza viral antigens mutate rapidly. Therefore new influenza drugs with efficient activity are much needed.

Understanding the molecular and cell biology of influenza is highly important for suppressing this widely spread disease. Influenza is a RNA virus that contains two major surface glycoproteins, neuraminidase and hemagglutinin. The life process and the structure of the influenza virus provide several potential molecular targets for drug innovation. Among these potential targets, neuraminidase (NA) has been found to be a potential target to control influenza virus. NA is responsible for viral release from infected cells and viral transport through the mucus in the respiratory tract and destroys hemagglutinin receptor on host cells, thus allowing the emergence of progeny virus particles from infected cells. Therefore, compounds that inhibit NA can protect the host from viral infection and retard its propagation.

Structure-based drug design methods have played a critical role in the discovery of NA inhibitors. Among these methods, QSAR (quantitative structure–activity relationship) is very efficient. QSAR has been demonstrated to be an effective computational tool in understanding the correlation between the structure of molecules and their activities [11–13]. In a QSAR one seeks to find a mathematical relationship between the activity and one or more descriptors. Thus, this study can indicate the structural factors which may play an important role in the determination of the activity. And its advantage over other methods lies in the fact that the descriptors used to build the models can be calculated from the structure alone

* Corresponding author. Tel.: +86 931 8911486; fax: +86 931 8912582.
E-mail address: lvwj05@lzu.cn (M.C. Liu).

0223-5234/$ - see front matter © 2007 Published by Elsevier Masson SAS.
doi:10.1016/j.ejmech.2007.04.011

and are not dependent on any experimental properties. However, the main problems encountered in this kind of research are still the description of the molecular structure using appropriate molecular descriptors and selection of suitable modeling methods. At present, many types of molecular descriptors such as constitutional, topological, geometrical, electrostatic, and quantum chemical descriptors have been proposed to describe the structural features of molecules [14–16].

Previously different chemometric and chemoinformatic methods, such as multiple linear regression (MLR), heuristic method (HM), principal component regression (PCR) and partial least squares (PLS) were used. More recently, neural networks such as the back-propagation neural network (BPNN), the support vector machines (SVM) and the radial basis function network (RBFNN) are often used in QSAR due to their flexibility in modeling nonlinear cases. Especially, the radial basis function network (RBFNN) has the advantages of small training times and is guaranteed to reach the global minimum of error surface during training. The optimization of its topology and learning parameters are easy to implement. Many problems in chemistry and biology have been successfully solved by RBFNN [17–20].

Although several QSAR model based different methods have been built for some influenza virus NA inhibitors [21–23], most of these studies were confined to the same scaffold compounds. New information about structure–activity relationship based on more compounds with greater diversity in their structures and activities should be helpful in discovering new NA inhibitors for curing the disease. Therefore, in this work, a QSAR model based on compounds with different scaffolds and a larger data set was built to improve the predictable ability of the model.

In this investigation heuristic method and radial basis function network (RBFNN) were first used to predict the activity of 46 NA inhibitors together. These 46 NA inhibitors belong to different kinds of molecular structures including: carbocyclic derivatives, benzoic acids, cyclopentanes and pyrrolidines. The aim of this investigation was to establish a new quantitative structure–activity relationship (QSAR) model for predicting the activity of new compounds and to find the structural factors which affect the activity of NA inhibitors. The QSAR model was more exact and comprehensive because of being built based on different kinds of molecular structures of NA inhibitors and the prediction results of $-\log(IC_{50})$ were satisfactory.

2. Materials and methods

2.1. Data set

All data of the present investigation were obtained from the book of progress in medicinal chemistry [24]. The data set for this investigation consisted of 46 influenza virus NA inhibitors. The molecular structures of 46 NA inhibitors are shown in Fig. 1.

A complete list of the compound numbers and their experimental and predicted $-\log(IC_{50})$ by heuristic method and radial basis function network is shown in Table 1.

2.2. Molecular descriptor generation

To obtain a QSPR model, compounds are often represented by the molecular descriptors. The calculation process of the molecular descriptors was described as below.

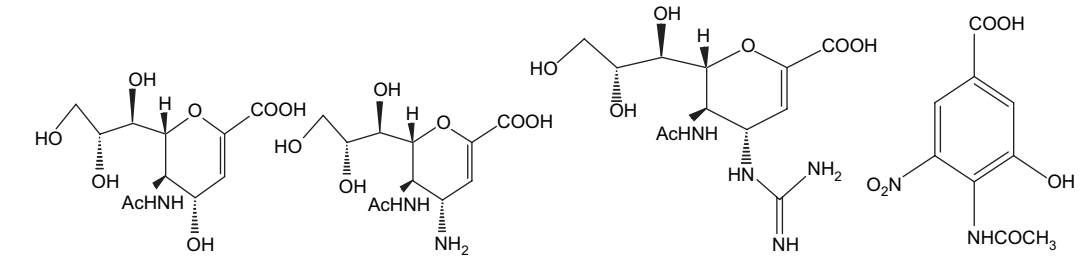
The two-dimensional molecular structures of 46 NA inhibitors were drawn by ChemDraw 6.0 and were geometrically optimized by the semi-empirical AM1 method in Hyperchem 5.0 software. Then calculated the optimized molecular structures in the software MOPAC [25], then the MOPAC files were transferred into software CODESSA to calculate all kinds of descriptors. The software CODESSA can calculate constitutional, topological, geometrical, electrostatic, and quantum chemical descriptors and has been successfully used in various QSPR researches [26–29].

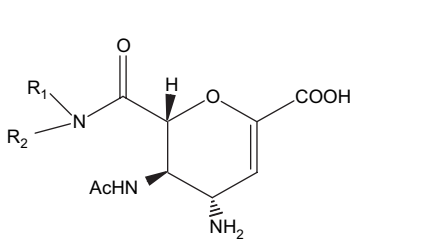
Constitutional descriptors are related to the number of atoms and bonds in each molecule. Topological descriptors include valent and non-valent molecular connectivity indices calculated from the hydrogen-suppressed formula of the molecule, encoding information about the size, composition, and the degree of branching of a molecule. The topological descriptors describe the atomic connectivity in the molecule. The geometrical descriptors describe the size of the molecule and require 3D-coordinates of the atoms in the given 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. In the present work, five classes of structural descriptors were obtained and about 434 descriptors were provided for every compound.

2.3. Theory of CODESSA

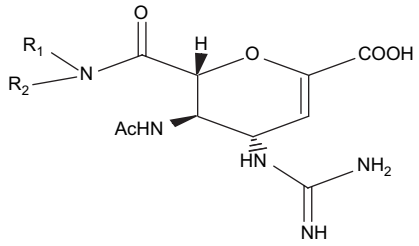
CODESSA includes two advanced procedures for systematic development of multi-linear QSAR/QSPR equations: (i) the heuristic method and (ii) the best multi-linear regression method. The heuristic method for descriptor selection proceeds with a preselection of descriptors by sequentially eliminating descriptors which do not match any of the following criteria: (i) the F-test greater than one unit; (ii) R value less than a value defined at the start (default 0.01); (iii) the student's t-test less than that defined (default 0.1) and (iv) duplicate descriptors having a higher squared inter-correlation coefficient than a predetermined level (usually 0.8). The next step involves correlation of the given property with (i) the top descriptor in the above list with each of the remaining descriptors, and (ii) the next one with each of the remaining descriptors, etc. The best pairs, as evidenced by the highest F-values in the two parameter correlations, are chosen and used for further inclusion of descriptors in a similar manner.

The goodness of the correlation is tested by the correlation coefficient (R), the F-test (F), and the squared standard error (s^2). The stability of the correlations was tested against the

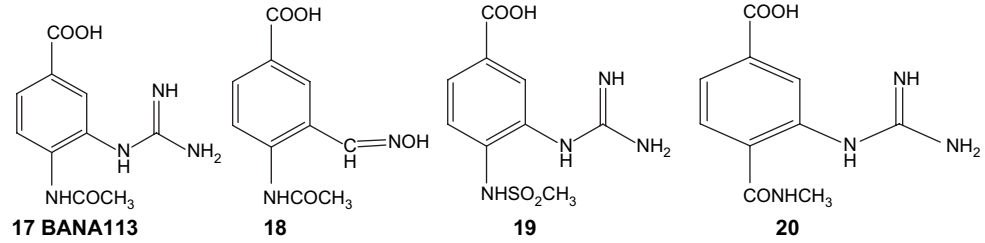


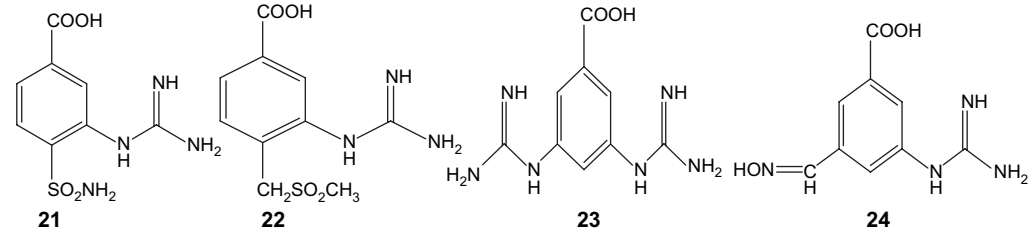


	R_1	R_2
4	$CH_2CH_2CH_3$	CH_3
5	$CH_2CH_2CH_3$	H
6	CH_3	CH_3
7	CH_2CH_3	CH_2CH_3
8	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$
9	$CH_2CH_2CH_2CH_3$	$CH_2CH_2CH_2CH_3$



	R_1	R_2
10	$CH_2CH_2CH_3$	H
11	$CH_2CH_2CH_3$	H
12	CH_3	CH_3
13	CH_2CH_3	CH_2CH_3
14	$CH_2CH_2CH_3$	CH_3
15	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$





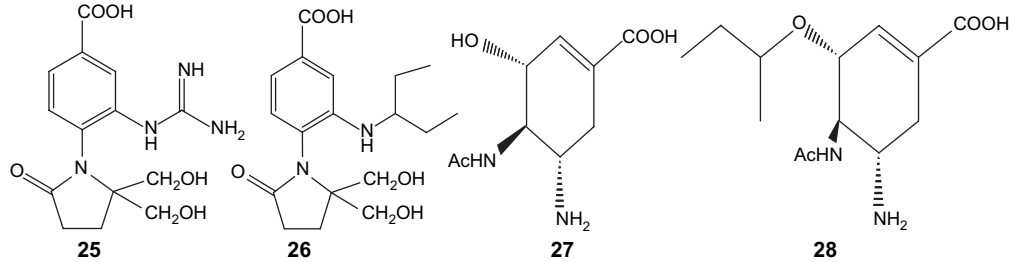


Fig. 1. Chemical structures of 46 NA inhibitors molecules.

cross-validated coefficient R^2_{cv} or adjusted squared correlation coefficient R^2_{adj}; these two parameters can avoid the over training of the model. The heuristic method's advantages are the high speed and no restrictions on the size of the data set. Besides, it will demonstrate descriptors which have bad or missing values, which are insignificant and which are highly intercorrelated. So, in this paper, the heuristic method was selected to build the best QSAR model.

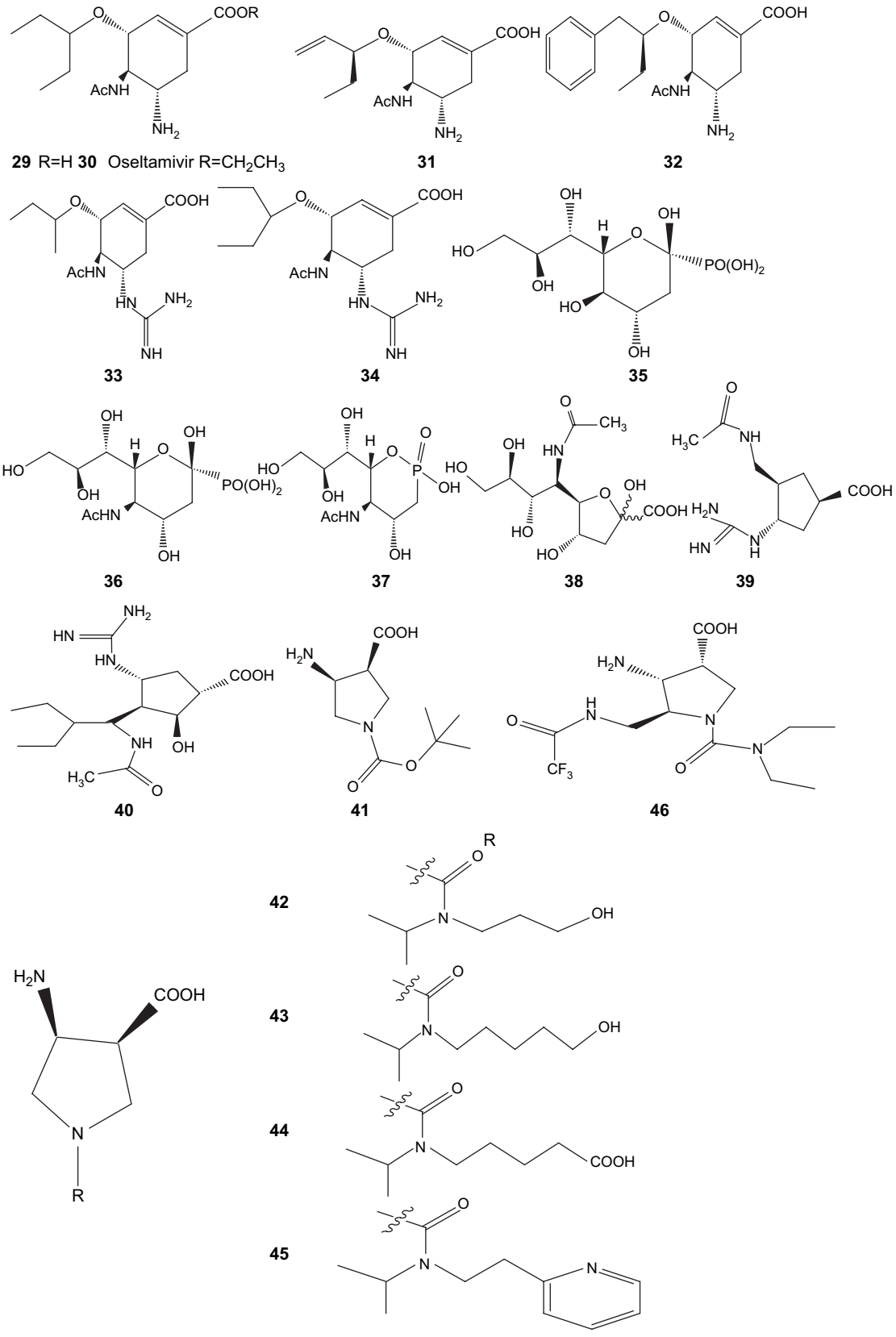


Fig. 1 (*continued*).

Table 1
The NA inhibitors and the predicted $-\log(IC_{50})$

No.	Experimental $-\log(IC_{50})$	Calculated$_{HM}$[a]	Abs. error[b]	Calculated$_{RBFNN}$[c]	Abs. error[b]
1	−0.9345	−1.62752	−0.6930	−1.1531	−0.2186
2[d]	0.4949	−0.88267	−1.3775	−0.7967	−1.2916
3	2.3010	1.171453	−1.1296	2.2581	−0.0429
4	0.7447	1.546316	0.8016	1.7567	1.0120
5	−1.2789	−0.45603	0.8229	−0.7173	0.5616
6[d]	−0.3802	0.499802	0.8800	0.378	0.7582
7	2.5229	1.326212	−1.1967	1.4505	−1.0724
8	1.9208	2.010127	0.0893	2.4371	0.5163
9	0.3468	0.141571	−0.2052	0.277	−0.0698
10[d]	−1.3010	0.570805	1.8718	0.3152	1.6162
11	0.3010	1.591982	1.2910	0.6991	0.3981
12	1.6210	1.674595	0.0536	1.6749	0.0539
13	3.0000	2.729901	−0.2701	2.7881	−0.2119
14[d]	2.3979	1.9835	−0.4144	1.9818	−0.4161
15	2.6990	2.946438	0.2474	2.5319	−0.1671
16	−2.8751	−1.89237	0.9827	−2.9483	−0.0732
17	−0.3979	−0.66806	−0.2702	−0.1937	0.2042
18[d]	−3.7404	−2.33763	1.4027	−1.0842	2.6562
19	−2.0000	−1.67679	0.3232	−1.536	0.4640
20	−0.6990	−0.61034	0.0887	−1.1074	−0.4084
21	−0.9542	−1.72616	−0.7720	−1.542	−0.5878
22[d]	−2.1461	−0.79919	1.3469	−1.2089	0.9372
23	−3.1761	−2.89729	0.2788	−3.2507	−0.0746
24	−2.6990	−2.8962	−0.1972	−2.1571	0.5419
25	2.3013	2.945432	0.6441	2.2461	−0.0552
26[d]	4.3188	3.041656	−1.2771	2.8833	−1.4355
27	−0.7993	−1.04773	−0.2484	−0.6591	0.1402
28	2.0000	0.859149	−1.1409	1.4637	−0.5363
29	3.0000	2.199078	−0.8009	2.4589	−0.5411
30[d]	3.0969	2.043737	−1.0532	2.9812	−0.1157
31	3.0000	3.295569	0.2956	3.6475	0.6475
32	3.5229	3.861088	0.3382	3.3933	−0.1296
33	3.3010	3.94301	0.6420	3.5328	0.2318
34[d]	3.3010	4.490019	1.1890	4.306	1.0050
35	−2.3010	−1.27876	1.0222	−2.3354	−0.0344
36	−1.7782	−1.55832	0.2199	−1.7944	−0.0162
37	−1.6021	−0.37787	1.2242	−1.4748	0.1273
38[d]	−2.0607	−0.87753	1.1832	−1.6385	0.4222
39	1.0000	1.001445	0.0014	0.8216	−0.1784
40	5.0000	3.78315	−1.2169	4.7276	−0.2724
41	−1.6990	−2.03935	−0.3404	−1.5785	0.1205
42[d]	−0.3222	0.731869	1.0541	0.2257	0.5479
43	−0.3010	0.133876	0.4349	0.0807	0.3817
44	−0.1139	−1.35923	−1.2453	−0.9358	−0.8219
45	−0.1139	0.716927	0.8308	−0.2382	−0.1243
46[d]	0.6990	1.221497	0.5225	−0.0774	−0.7764

[a] Predicted $-\log(IC_{50})$ by HM.
[b] Absolute value of calculated − experimental.
[c] Predicted $-\log(IC_{50})$ by RBFNN.
[d] Compounds in the test set.

2.4. Theory of RBFNN

The theory of RBFNN has been adequately presented in some references [30,31]. Here only a brief description of the RBFNN principle is given. The RBFNN consists of three layers: an input layer, a hidden layer and an output layer. The input layer does not process the information; it only distributes the input vectors to the hidden layer. Each neuron on the hidden layer employs a radial basis function as a nonlinear transfer function to operate on the input data. In general, there are several radial basis functions (RBF): linear, cubic, thin plate spline, Gaussian, multi-quadratic and inverse multi-quadratic. The most often used RBF is a Gaussian function that is characterized by a center (c_j) and width (r_j). In this study, the Gaussian was selected as the radial basis function. The nonlinear transformation with RBF in the hidden layer is given as follow:

$$h_j(x) = \exp\left(\frac{-\|x - c_j\|^2}{r_j^2}\right)$$

where, h_j is the notation for the output of the jth RBF unit, c_j and r_j are the center and width of the jth RBF, respectively.

The operation of the output layer is linear, which is given as below:

$$y_k(x) = \sum w_{kj} h_j(x) + b_k$$

where, y_k is the kth output unit for the input vector x, w_{kj} is the weight connection between the kth output unit and the jth hidden layer unit, and b_k is the bias.

The training procedure when using RBF involves selecting centers, width and weights. In this paper, the forward subset selection routine was used to select the centers from training set samples. The adjustment of the connection weight between the hidden layer and the output layer was performed using a least squares solution after the selection of centers and widths of radial basis functions.

The overall performance of RBFNN was evaluated in terms of root-mean-square (RMS) error which was defined as below:

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{n_s}(y_k - \hat{y}_k)^2}{n_s}}$$

where, y_k is the desired output, $\hat{y}_k$ the actual output of the network, and n_s is the number of compounds in training set.

2.5. RBFNN implementation and computation environment

All calculation programs implementing RBFNN were written in M-file based on basis MATLAB script for radial basis function neural networks. All programs were operated on a system with AMD 1700+ PC with 256MB RAM.

3. Results and discussion

3.1. Result of the HM

To find the optimum number of descriptors describing $-\log(IC_{50})$ for current set of structures, the heuristic correlations, performed for the training set, chose nine descriptors to provide the optimal seven equations for different numbers of descriptors from one to seven. Fig. 2 shows that R rises as the number of parameters increasing from one to seven steadily, although after fifth the increase was not very evident, but R^2_{adj} decreases at the sixth point. To avoid the ''over parameterization''

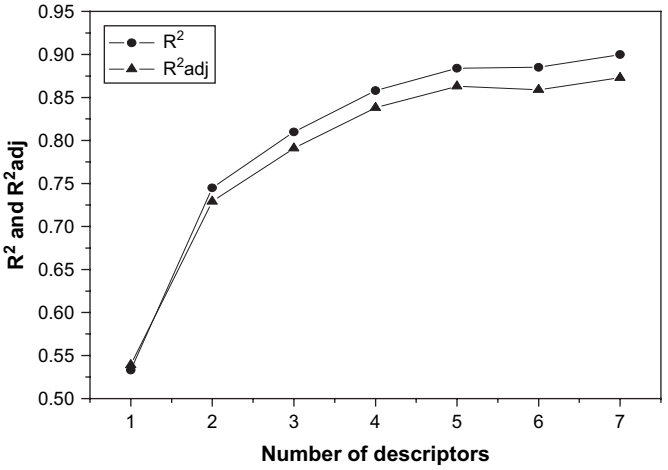


Fig. 2. Influence of the number of descriptors on R and R^2_{adj} of the regression models.

of the model, an increase of the R values of less than 0.02 or the point where R^2_{adj} begin to decrease was chosen as the breakpoint criterion [32]. It can be seen that five descriptors appear to be sufficient for a successful regression model. A detailed description of the linear model based on compounds in the training set is summarized in Table 2.

The obtained model has a correlation coefficient $R = 0.884$, $F = 42.559$, $s^2 = 0.638$ and an adjusted squared correlation coefficient $R^2_{adj} = 0.863$.

This model contains three electrostatic (WPSA3, Qmin, HDSA1), one topological (IC1) and one quantum chemical (MNRIO) descriptors. These descriptors encoded different aspects of the molecular structure. The combination of these descriptors, comprising shape, electronic and bond information about molecules, adequately represents hydrophobic, steric and stability effects on the activity of a molecule. By interpreting the descriptors in the regression model, it is possible to gain some insight into factors that are likely to govern the activity of influenza virus NA inhibitors.

It is well known that the interaction between medicine and the acceptors mainly includes the interaction between molecules, hydrophobicity, electrostatic effect and hydrogen bond. Among the five descriptors we chose, information content (order 1) (IC1) describe the size, shape and branching information of the molecules and give some information about the steric effects on the activity of a molecule. WPSA3 encodes features responsible for polar interactions between molecules and is calculated as a combination of the contributions of atomic partial charges to the total molecular solvent-accessible surface area. Qmin reflects characteristics of the charge distribution of the molecule. The empirical partial charges in the molecule are calculated using the approach based on the Sanderson electronegativity scale and uses the concept which represents the molecular electronegativity as a geometric mean of atomic electronegativities. HDSA1 is hydrogen donor charged solvent-accessible surface area; this descriptor belongs to charged partial surface area (CPSA) descriptors, and represents the sum of solvent-accessible surface area of the H-bonding donor atoms. It describes the hydrogen bond donation ability of the molecule. As the HDSA1 increases, the proportion of the sum of solvent-accessible surface area of hydrogen donors among the total molecular surface area also increases and the formation of the hydrogen bond becomes easier. Thus, HDSA1 describes the effect of hydrogen bond donation ability of NA inhibitors. Maximum nucleophilic reactivity index of an O atom (MNRIO) belongs to charge distribution-related descriptors of quantum chemical descriptors group. It estimates the relative reactivity of the O atoms in the molecule for a given series of compounds and is related to the activation energy of the corresponding chemical reaction. The positive coefficient in the model implies that increasing the value of this descriptor can lead to the lower activity of NA inhibitors.

Table 2
Descriptors, coefficients, standard error, and t-test values for the linear model[a]

Descriptor	Chemical meaning	Coefficient	Error	t-test
Constant	Intercept	−16.835	2.171	−7.755
IC1	Information content (order 1)	0.088	0.007	12.033
WPSA3	WPSA3 weighted PPSA (PPSA3 × TMSA/1000) [Zefirov's PC]	−0.907	0.123	−7.380
Qmin	Min. partial charge	−55.491	11.992	−4.627
HDSA1	HA dependent HDSA1/TMSA [Zefirov's PC]	10.508	2.933	3.583
MNRIO	Max. nucleoph. react. index for a O atom	53.907	21.657	2.489

[a] $R = 0.884$, $F = 42.559$, $s^2 = 0.638$, $R^2_{adj} = 0.863$, $n = 34$.

From the above discussion, it can be seen that all the descriptors involved in the model have physical meaning, and these descriptors can account for the structural features responsible for the activities of the NA inhibitors: HDSA1 reflects the hydrogen bond donor ability of NA inhibitors; IC1 represents the steric effects and interactions between molecules; MNRIO reflects the activation energy of the corresponding chemical reaction; WPSA3 describes the hydrophobicity of NA inhibitors and Qmin reflect the electrostatic effect between medicine and NA.

The calculated $-\log(IC_{50})$ by HM are given in Table 1, and a plot of the calculated vs. experimental $-\log(IC_{50})$ for all the 46 NA inhibitors studied including the training set and the test set is shown in Fig. 3. The model gave RMS of 0.7351 for the training set, 1.1926 for the test set, and 0.8716 for the whole set, and the corresponding correlation coefficients (R) were 0.940, 0.8970, and 0.9206, respectively.

3.2. Result of RBFNN method

Table 2 and Fig. 3 show that the result of the multiple linear regression was not sufficiently accurate and the factors influencing the activity of these compounds were complex and not all of them had absolute linear correlation with the activity. In order to compare with the heuristic method and improve the prediction of model, RBFNN were used to build the nonlinear prediction model for further discussion on the correlation

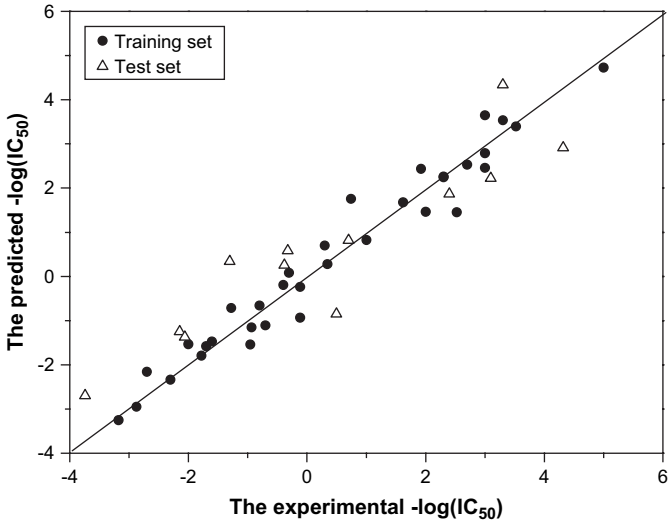


Fig. 3. The predicted vs. experimental $-\log(IC_{50})$ (HM).

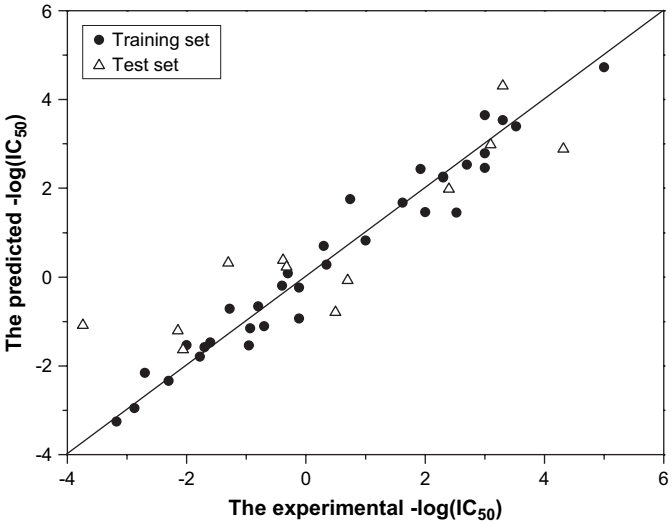


Fig. 5. The predicted vs. experimental $-\log(IC_{50})$ (RBFNN).

between the molecular structure and the absolute activity based on the same subset of descriptors. Such RBFNN can be designed as 5-n_k-1 net to indicate the number of the units in the input, the hidden, and the output layer, respectively. The optimal width was selected by experimenting with a number of trials and selecting the one most favored by the model selection criterion: varying the width indicates that the width has little effect on the performance of RBFNN, if it exceeds 3. So we selectively computed the width for every 0.1from 0 to 3. Each minimum error on LOO cross-validation was plotted vs. the width (Fig. 4) and the minimum was chosen as the optimal condition. In this case: $r = 1.5$ and $n_k = 14$.

From the best network, the inputs in the test set were presented, and the results with RBFNN were obtained. They are shown in Table 1 and Fig. 5. The model gave RMS of 0.4266 for the training set, 1.1949 for the test set, and 0.7119 for the whole set, and the corresponding correlation coefficients (R) were 0.9796, 0.8856, and 0.9472, respectively.

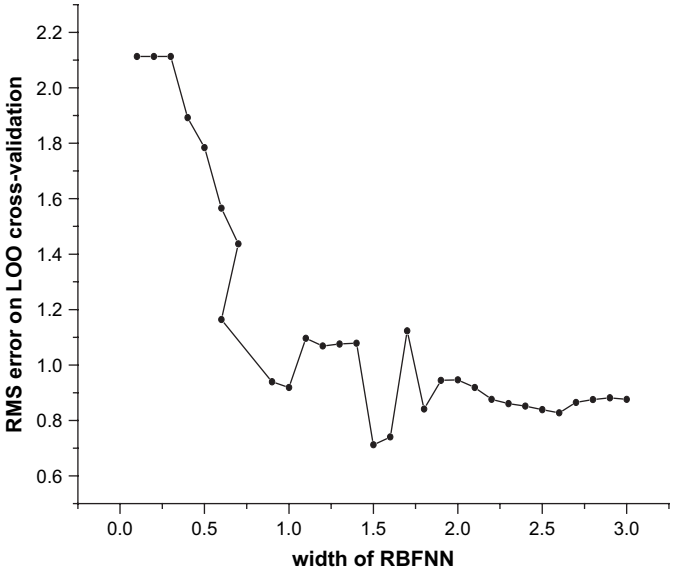


Fig. 4. The width of RBFNN vs. RMS error on LOO cross-validation.

4. Conclusion

QSAR models for the prediction of activity of NA inhibitors using HM and RBFNN based on descriptors calculated from molecular structure alone have been developed. Satisfactory results were obtained with the proposed method. By Analyzing the obtained results, the linear model indicated that the interaction between molecules, hydrophobicity and hydrogen bond interactions play important role in the activity of NA inhibitors. Additionally, the nonlinear RBFNN models produced better results with good predictive ability than linear model. So we can conclude that (1) the proposed models in this paper could identify and provide some insight into which structural features are related to the activity of NA inhibitors and help to get more useful information about exploring or synthesis of new NA inhibitors; (2) nonlinear relationship can describe accurately the relationship between the structural parameter and the activity factors of the studied NA inhibitors. Additionally, compared with the previous work, the data set used in our investigation was more diverse; the compounds studied in this work cover almost all types of NA inhibitors. The models we developed are more general and applicable for the prediction of the activity of NA inhibitors. Furthermore, the proposed method can also be extended to other QSAR investigations.

Acknowledgement

The authors thank the National Science Foundation of China (No. 20275014) for supporting this work.

References

[1] V.R. Atigadda, W.J. Brouillette, F. Duarte, S.M. Ali, Y.S. Babu, S. Bantia, P. Chand, N. Chu, J.A. Montgomery, D.A. Walsh, E.A. Sudbeck, J. Finley, M. Luo, G.M. Air, G.M. Laver, J. Med. Chem. 42 (1999) 2332–2343.

[2] M.A. Williams, W. Lew, D.B. Mendel, C.Y. Tai, P.A. Escarpe, W.G. Laver, R.C. Stevens, C.U. Kim, Bioorg. Med. Chem. Lett. 7 (1997) 1837–1842.
[3] C.U. Kim, W. Lew, M.S. Williams, H. Wu, L. Zhang, X. Chen, P.A. Escarpe, D.B. Mendel, W.G. Laver, R.C. Stevens, J. Med. Chem. 41 (1998) 2451–2460.
[4] W. Lew, H.W. Wu, D.B. Mendel, P.A. Escarpe, X.W. Chen, G. Laver, B.J. Graves, C.U. Kim, Bioorg. Med. Chem. Lett. 8 (1998) 3321–3324.
[5] P. Chand, Y.S. Babu, S. Bantia, S. Rowland, A. Dehghani, P.L. Kotian, T.L. Hutchison, S. Ali, W. Brouillette, Y. El-Kattan, T.H. Lin, J. Med. Chem. 47 (2004) 1919–1929.
[6] G.T. Wang, Y. Chen, S. Wang, R. Gentles, T. Sowin, W. Kati, S. Muchmore, V. Giranda, K. Stewart, H. Sham, D. Kempf, W.G. Laver, J. Med. Chem. 44 (2001) 1192–1201.
[7] Y.S. Babu, P. Chand, S. Bantia, P. Kotian, A. Dehghani, Y. El-Kattan, T.H. Lin, T.L. Hutchison, A.J. Elliott, C.D. Parker, S.L. Ananth, L.L. Horn, G.W. Laver, J.A. Montgomery, J. Med. Chem. 43 (2000) 3482–3486.
[8] P. Chand, Y.S. Babu, S. Bantia, N. Chu, L.B. Cole, P.L. Kotian, W.G. Laver, J.A. Montgomery, V.P. Pathak, S.L. Petty, D.P. Shrout, D.A. Walsh, G.M. Walsh, J. Med. Chem. 40 (1997) 4030–4052.
[9] N.R. Taylor, M.V. Itzstein, J. Med. Chem. 37 (1994) 616–624.
[10] C.U. Kim, W. Lew, M.A. Williams, H.T. Liu, L.J. Zhang, S. Swaminathan, N. Bischofberger, M.S. Chen, D.B. Mendel, C.Y. Tai, W.G. Laver, R.C. Steven, J. Am. Chem. Soc. 119 (1997) 681–690.
[11] Y.Y. Niu, L.M. Yang, H.Z. Liu, Y.Y. Cui, L. Zhu, J.M. Feng, J.H. Yao, H.Z. Chen, B.T. Fan, Z.N. Chen, Y. Lu, Bioorg. Med. Chem. Lett. 15 (2005) 4814–4818.
[12] D. Mandloi, S. Joshi, P.V. Khadikar, N. Khosla, J. Inorg. Biochem. 99 (2005) 575–583.
[13] G. Chen, X.M. Luo, W.L. Zhu, C. Luo, H. Liu, C.M. Puah, K.X. Chen, H.L. Jiang, Bioorg. Med. Chem. 12 (2004) 2409–2417.
[14] M. Karelson, Molecular Descriptors in QSAR/QSPR, Wiley, New York, 2000.
[15] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim, Germany, 2000.
[16] J. Devillers, A.T. Balaban, Topological Indices and Related Descriptors in QSAR and QSPR, Gordon and Breach, Amsterdam, 1999.
[17] Q.F. Li, X.J. Yao, X.G. Chen, M.C. Liu, R.S. Zhang, X.Y. Zhang, Z.D. Hu, Analyst 125 (2000) 2049–2053.
[18] Y.W. Wang, X.Y. Zhang, X.J. Xiao, Y.H. Gao, M.C. Liu, Z.D. Hu, B.T. Fan, Anal. Chim. Acta 463 (2002) 89–97.
[19] Y.W. Wang, X.Y. Zhang, X.J. Xiao, R.S. Zhang, M.C. Liu, Z.D. Hu, B.T. Fan, Talanta 57 (2002) 641–652.
[20] Y.H. Gao, Y.W. Wang, X.J. Xiao, X.Y. Zhang, M.C. Liu, Z.D. Hu, B.T. Fan, Talanta 59 (2003) 229–237.
[21] R.P. Verma, C. Hansch, A QSAR study on influenza neuraminidase inhibitors, Bioorg. Med. Chem. 14 (2006) 982–996.
[22] X.W. Zhang, Y.L. Yap, J. Mol. Struct.: THEOCHEM 681 (2004) 137–141.
[23] X. Yi, Z.R. Guo, F.M. Chu, Bioorg. Med. Chem. 11 (2003) 1465–1474.
[24] X. Cong, Z.J. Yao, L.M. Wu, Q.J. Liao, in: S.X. Peng (Ed.), Progress in Medicinal Chemistry, vol. 3, Chemical Industrial Publication, Beijing, 2004, pp. 24–53.
[25] J.P.P. Stewart, MOPAC, v.6.0, Quantum Chemistry Program Exchange QCPE, No. 455, Indiana University, Bloomington, IN, 1989.
[26] E.J. Delgado, J.B. Alderete, G.A. Jana, J. Chem. Inf. Comput. Sci. 43 (2003) 1928–1932.
[27] R. Hiob, M. Karelson, J. Chem. Inf. Comput. Sci. 40 (2000) 1062–1071.
[28] R. Bosque, J. Sales, J. Chem. Inf. Comput. Sci. 43 (2003) 637–642.
[29] A.R. Katritzky, R. Petrukhin, R. Jain, M. Karelson, J. Chem. Inf. Comput. Sci. 41 (2001) 1521–1530.
[30] X.J. Yao, Y.W. Wang, X.Y. Zhang, R.S. Zhang, M.C. Liu, Z.D. Hu, B.T. Fan, Chemom. Intell. Lab. Syst. 62 (2002) 217–225.
[31] X.J. Yao, X.Y. Zhang, R.S. Zhang, M.C. Liu, Z.D. Hu, B.T. Fan, Talanta 57 (2002) 297–306.
[32] J.D. Dyekjar, S.O. Jonsdottir, Ind. Eng. Chem. Res. 42 (2003) 4241–4259.