Original article

Studies on log P, retention time and QSAR of 2-substituted phenylnitronyl nitroxides as free radical scavengers

Ming Zhao [a], Zheng Li [a], Yihui Wu [a], Yu-Rong Tang [c], Chao Wang [b],
Ziding Zhang [c,*], Shiqi Peng [a,**]

[a] *College of Pharmaceutical Sciences, Capital Medical University, Beijing 100069, PR China*
[b] *College of Pharmaceutical Sciences, Peking University, Beijing 100083, PR China*
[c] *College of Biological Sciences, China Agricultural University, Beijing 100094, PR China*

Abstract

As reported in our previous paper [Y. Wu, L. Bi, W. Bi, Z. Li, M. Zhao, C. Wang, J. Ju, S. Peng, Bioorg. Med. Chem. 14 (2006) 5711–5720.], a series of novel 2-substituted nitronyl nitroxides were synthesized and characterized to show enhanced free radical scavenger properties. Here, the log P values and retention time (alternatively represented as log K) of these 20 compounds were determined by the shake-flask method and HPLC analysis, respectively. We found that the prediction of log P directly from the measured log K value can result in a high accuracy, thus the determination of time-consuming log P can be replaced by the simple HPLC analysis. Using the conventional Hansch method, the QSAR equations of these nitronyl nitroxides were established to predict their capability of trapping free radicals such as $\bullet NO$, $\bullet H_2O_2$ and $\bullet OH$. Although the overall predictive accuracy is reasonable, it was found that the log P values for the compounds are not significantly related to their capability of trapping free radicals. Moreover, an improved QSAR analysis was performed by choosing new molecular descriptors generated from e-dragon server. The increased predictive accuracy suggests that these new QSAR models have practical application in screening new potential scavengers from the derivatives of 2-substituted phenylnitronyl nitroxides.
© 2007 Elsevier Masson SAS. All rights reserved.

Keywords: Nitronyl nitroxide derivatives; Free radical scavenger; log P; log K; QSAR; Molecular descriptors

1. Introduction

Under many physiological and pathological conditions, free radicals play important roles [1]. In general, any imbalance of the generation and scavenging of free radicals or excessive amounts of active radicals may contribute to disease development. Free radical reactions can result in deleterious modifications in membranes, proteins, enzymes and DNA [2], increase the risk of diseases such as cancer [3], Alzheimer's [4], Parkinson's [5], angiocardiopathy [6], arthritis [7], asthma [8],

diabetes [9], and degenerative eye disease [10]. Therefore, it is important to discover effective scavengers of free radicals for the prevention and treatment of such disorders.

It is commonly accepted that nitronyl nitroxides, synthesized more than 30 years ago [11], are chemically remarkable because of their capability to trap $\bullet NO$, $\bullet OH$, $\bullet H_2O_2$, and $\bullet O_2$, protecting endothelial cells from the attack of free radicals [12–15]. To explore the new inhibitors of Ach-induced vasorelaxation, 2-phenylnitronyl nitroxide was used as the lead compound [16]. Chemical modification on this scaffold resulted in a series of 2-substituted phenylnitronyl nitroxide derivatives **1–20** (see Scheme 1), which exhibit good or moderate radical scavenging activities as shown in the PC12 cell survival assay. The preliminary SAR analysis indicated that the free radical scavenging activities of **1–20** strongly
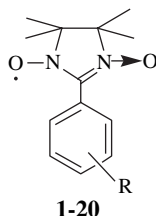
1-20

Scheme 1. Definition of 2-substituted phenylnitronyl nitroxides. In **1** R = H, **2** R = 2′-OH, **3** R = 3′-OH, **4** R = 4′-OH, **5** R = 4′-CH₃, **6** R = 4′-OCH₃, **7** R = 2′,4′-dimethoxyl, **8** R = 3′,4′-dimethoxyl, **9** R = 3′,4′-methylenedioxy, **10** R = 3′-OH−4′-OCH₃, **11** R = 4′-OH−3′-OCH₃, **12** R = 4′-N(CH₃)₂, **13** R = 2′-F, **14** R = 4′-Cl, **15** R = 4′-Br, **16** R = 2′,4′-dichloro, **17** R = 2′-NO₂, **18** R = 3′-NO₂, **19** R = 4′-NO₂, **20** R = 4′-OCH₂COOH.

depended on the character of the substituents and the substitution position in the phenyl ring. Consequently, establishing a practical QSAR model to direct the rational modification of 2-substituted phenylnitronyl nitroxide is needed.

For a receptor with unknown 3D structure, the QSAR analysis of the corresponding ligands is usually carried out using the CoMFA method [17]. It is well known that the free radical scavenging activity of 2-substituted phenylnitronyl nitroxide mainly depends on its direct interaction with the free radical and no receptor is directly involved. On the other hand, so far the CoMFA method is not very suitable for the QSAR analysis of the compounds bearing the characteristic of scavenging free radical. In this case, the QSAR analysis of 2-substituted phenylnitronyl nitroxides has to be carried out using classical QSAR methods (e.g. Hansch method [18]).

The log P value, defined as the logarithm of the partition coefficient between *n*-octanol and water, is an important parameter to judge a molecule's drug likeness [19]. Meanwhile, it has been widely utilized as an important parameter in QSAR analysis such as Hansch approach. Owing to the influence of the unpaired electrons there were no data available to calculate the exact log P values of 2-substituted phenylnitronyl nitroxides.

In the first step of the present study, the free radical scavengers **1−20** were prepared using the literature reported procedures [16], and their log P values were measured by the shake-flask method [20]. Considering the facts that the shake-flask method is time-consuming and the significant correlation between log P and retention time of HPLC, these compounds' retention times were recorded on an HPLC instrument following the routine procedure. Their log K values were calculated based on the retention time, and the correlation between log P and log K was then established.

In the second step, we focused our efforts on developing the QSAR models of compounds **1−20** as free radical scavenger. Firstly, the QSAR analysis was carried out using the conventional Hansch method. In the past decades, many molecular descriptors were developed to facilitate the QSAR study. For instance, a series of different algorithms for calculating descriptors have been compiled into e-dragon web-server, which allows to generate more than 1000 different descriptors for each molecule [21]. To make the current QSAR analysis practical for guiding the further development of potential free

radical scavengers, an improved QSAR analysis was performed using new molecular descriptors derived from the e-dragon server.

2. Methods and materials

2.1. Preparation of 2-substituted phenylnitronyl nitroxides **1−20**

The 2-substituted phenylnitronyl nitroxides represented by the formulae in Scheme 1 were synthesized according to the literature method [16], and the total yields of the compounds were comparable to those reported in this literature.

2.2. Determination of the log P values of 2-substituted phenylnitronyl nitroxides **1−20**

To a flask a solution of 5 mg 2-substituted phenylnitronyl nitroxide in 5 ml water saturated with 5 ml 1-octanol was added. The flask was shaken for 20 min. The mixed solution in the flask was converted into a centrifuge tube and centrifuged at 4000g for 15 min. Then the solution was allowed to stand for 30 min and separated into two layers. From the aqueous phase 80 μl sample was taken to analyze the retention time recorded using HPLC, and from the octanol phase 20 μl sample was also taken to analyze the retention time recorded using HPLC. The sample was injected into a Kromasil 4.6×150 mm C₁₈ reversed-phase column equipped with a 4×4 mm C₁₈ guard column for HPLC analysis. Gradient elution was performed with 40−80% acetonitrile in water at a flow rate of 1.0 ml/min. The tested 2-substituted phenylnitronyl nitroxide was monitored at its maximal absorbing wavelengths with Waters PDA 2996 UV detector. The content of 2-substituted phenylnitronyl nitroxide in a sample was represented by the peak area. The partition coefficients (log P) were calculated according to Eq. (1).

$$\log P = \log(C_{octanol}/C_{water}) = \log(S_{octanol,\ 20\ \mu l}/S_{water,\ 80\ \mu l} \times 4) \quad (1)$$

where $S_{octanol,\ 20\ \mu l}$ is the integral peak area of the tested nitronyl nitroxide in 20 μl of octanol solution sample, and $S_{water,\ 80\ \mu l}$ is the integral peak area of the tested nitronyl nitroxide in 80 μl of water solution of the same sample.

2.3. Determination of the retention time of 2-substituted phenylnitronyl nitroxides **1−20**

2-Substituted phenylnitronyl nitroxides were dissolved in water to prepare sample solutions of 1.5 mM/L. The HPLC analysis was carried out on Waters 600E, the column was a reversed-phase Rainbow C18 column (Kromasil 4.6×150 mm). After 20 μl of the sample (1.5 mM/L) was loaded, the column was eluted with 40% solution of MeOH in phosphate buffer (0.001 M) as the mobile phase for 23 min. The flow rate was 0.25 ml/min. 2-Substituted phenylnitronyl nitroxide in the sample was monitored with UV detector (190−600 nm) at

275 nm and the retention time (t_R) of its peak was recorded. With the same HPLC conditions the retention time of acetone peak was recorded as t_0. In order to offset the influence of the solvent on the appearance time of the peak of 2-substituted phenylnitronyl nitroxide, the appearance time of acetone peak (t_0, 2.090 min) was used as a control. As an alternative representation of t_R, the parameter $\log K$ was defined based on Eq. (2).

$$\log K = \log[(t_R - t_0)/t_0] \tag{2}$$

2.4. The simple linear regression analysis of the $\log P$ and $\log K$ values of compounds 1–20

It has been well accepted that the $\log P$ and $\log K$ values are linearly correlated. To judge if the experimentally measured $\log K$ can be used to predict the $\log P$ value, we constructed the linear regression equation between the $\log P$ and $\log K$. In addition to the standard validations consisting of the resubstitution and Leave-One-Out (LOO) tests which will be described in the subsequent section, two novel compounds with tested $\log P$ and $\log K$ values were further used to validate the prediction of $\log P$ based on its linear relation with the $\log K$.

2.5. Validation of a prediction algorithm

To systematically assess a prediction algorithm, a reliable validation is required. Usually, a predictive method is evaluated by the predictive results for a training data set and testing data set, respectively. According to the statistical terminology, the former is called a test of resubstitution, and the latter is a test of cross-validation. By the test of resubstitution, the $\log P$ value of each compound in a training data set is predicted using the rules derived from the same set. Although this test gives somewhat optimistic error estimate because the same compounds are used to derive the prediction rules and to predict themselves, the resubstitution test is absolutely necessary due to its ability of reflecting the self-consistency of a given method. On the other hand, a cross-validation test for an independent testing data set is also needed because it can reflect the generalizing effectiveness of a predictive method. Of various cross-validation tests, the LOO test is thought to be a reliable one. By the LOO test, the $\log P$ value of each compound in the data set is predicted by the rules derived from all the other compounds except the predicted one. In other words, in the LOO test each compound is singled out as an independent sample and used to examine the predictive method. Both tests of resubstitution and LOO were used to evaluate the above prediction of $\log P$ based on experimentally measured $\log K$ value. In addition, the validation will be used for assessing the QSAR analysis in the current study.

In the resubstitution test, three statistical parameters (R, S and $\bar{e}$) were used to evaluate the performance. The regression coefficient R is defined by Eq. (3).

$$R = \sqrt{1 - \frac{\sum_{i=1}^{N}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2}} \tag{3}$$

where y_i, $\widehat{y}_i$ and $\overline{y}$ are predicted, actual, and mean values of the activity, respectively; N is the number of compounds in the resubstitution test. The root mean square of errors S is calculated with Eq. (4).

$$S = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \widehat{y}_i)^2}{N}} \tag{4}$$

To illustrate the predictive accuracy more explicitly, the $\bar{e}$ (absolute average error) is defined by Eq. (5).

$$\bar{e} = \frac{\sum_{i=1}^{N}|y_i - \widehat{y}_i|}{N} \tag{5}$$

In the LOO test, the above three statistical parameters, namely R_{LOO}, S_{LOO} and $\bar{e}_{LOO}$, were also calculated. As discussed above, the R_{LOO}, S_{LOO} and $\bar{e}_{LOO}$ are able to reflect the real predictive capability of the developed models.

2.6. QSAR analysis of 1–20 via Hansch approach

The classical Hansch approach was employed to build the QSAR equations for three different activities of compounds 1–20. The three frequently used parameters are the Hammett substituent parameter (σ), $\log P$ and molar refractivity (MR).

The Hammett substituent parameter is used as a concise measure of the electronic characteristics of the molecules. The $\sum \sigma$ is the sum of Hammett σ. The σ value of the substituent at the *meta* position of the phenyl is represented by σ_m, and the σ value of the substituent at the *para* position of the phenyl is represented by σ_p or σ_p^-. The positive σ values represent that the substituent withdraws electron from the phenyl, and the negative σ values represent that the substituent releases electron.

The molar refractivity MR is used to describe van der Waals interactions of the substituents. In order to make all parameters of the same order of magnitude, the MR values were divided by 10. All the parameters needed to calculate MR and $\sum \sigma$ were obtained from the literatures [22–24], and the calculated parameters for compounds 1–20 are summarized in Table 1.

As a suitable measure of relative hydrophobicity, the $\log P$ values for these 20 compounds were also chosen in the current QSAR analysis. In some cases, there is a parabolic dependence of the activity on $\log P$, thus the $(\log P)^2$ was also chosen as a parameter in our QSAR analysis.

For the QSAR analysis of the EC_{50} activity of scavenging three different free radicals in compounds 1–20, we used three different combinations of the above mentioned parameters to construct the linear equations, i.e. (1) two parameters ($\sum \sigma$, MR); (2) three parameters ($\sum \sigma$, MR and $\log P$); and (3) four parameters ($\sum \sigma$, MR, $\log P$ and $(\log P)^2$). Furthermore, the predictive accuracy based on different parameters was evaluated using the resubstitution and LOO tests and the optimal equation was established.

Table 1
The values for three molecular parameters in the Hansch approach and nine molecular descriptors used in the improved QSAR analysis

Compound	MR	$\sum \sigma$	$\log P$	Sv	Mv	Ss	Ms	MSD	Har2	ZM1	nAT	Qmean
1	0.30	0.00	1.365	20.49	0.60	41.83	2.46	0.224	55.171	94	34	0.029
2	0.49	−0.37	1.169	21.01	0.60	47.50	2.64	0.216	60.388	100	35	0.029
3	0.48	0.12	1.141	21.01	0.60	47.50	2.64	0.222	59.976	100	35	0.029
4	0.48	−0.37	0.619	21.01	0.60	47.50	2.64	0.229	59.693	100	35	0.029
5	0.77	−0.17	1.897	22.09	0.60	43.50	2.42	0.229	59.693	100	37	0.027
6	0.99	−0.27	1.622	22.60	0.59	47.00	2.47	0.236	63.956	104	38	0.026
7	1.68	−0.54	0.770	24.71	0.59	52.17	2.48	0.217	74.744	114	42	0.024
8	1.68	−0.15	1.146	24.71	0.59	52.17	2.48	0.223	74.266	114	42	0.024
9	1.00	−0.16	1.508	22.52	0.61	49.67	2.48	0.221	70.769	116	37	0.027
10	1.17	−0.15	0.918	23.12	0.59	52.67	2.63	0.227	69.344	110	39	0.026
11	1.17	−0.25	1.135	23.12	0.59	52.67	2.63	0.222	69.552	110	39	0.026
12	1.75	−0.83	2.015	24.68	0.59	47.50	2.38	0.237	68.720	110	42	0.024
13	0.29	0.06	0.822	20.61	0.61	49.50	2.75	0.216	60.388	100	34	0.029
14	0.80	0.23	2.330	21.20	0.62	45.61	2.53	0.229	59.693	100	34	0.029
15	1.09	0.23	2.461	21.58	0.63	44.25	2.46	0.229	59.693	100	34	0.029
16	1.30	0.46	1.932	21.90	0.64	49.39	2.60	0.219	65.160	106	34	0.029
17	0.94	0.78	0.909	21.91	0.61	57.50	2.88	0.209	70.398	110	36	0.028
18	0.94	0.71	1.636	21.91	0.61	57.50	2.88	0.222	69.419	110	36	0.028
19	0.94	1.24	1.684	21.91	0.61	57.50	2.88	0.237	68.720	110	36	0.028
20	1.60	−0.24	−0.518	24.63	0.60	61.17	2.78	0.254	76.734	118	41	0.024

2.7. Improved QSAR analysis by selecting new molecular descriptors

From a series of molecular descriptors developed in the past decades the new molecular descriptors were selected to improve the current QSAR analysis. Firstly, the SMILES formats of compounds **1**−**20** were generated through JME server (http://www.molinspiration.com/cgi-bin/properties). Subsequently, the SMILES formats were further converted into the 3D structures with SDF file format with the assistance of CORINA (http://www2.chemie.uni-erlangen.de/software/corina/). Moreover, the 20 SDF files were submitted to e-dragon server (http://www.vcclab.org/lab/edragon/) to calculate the molecular descriptors. Finally, the 1664 different molecular parameters for each compound were obtained.

To make sure that the selected molecular descriptors have relatively intuitive physicochemical properties, only 210 parameters from four different types of descriptors (the constitutional descriptors, topological descriptors, charge descriptors and descriptors of molecular properties) remained for the further QSAR analysis. Since the number of the free radical scavengers under investigation is merely 20, the optimum number of molecular descriptors used in the linear QSAR equation is around 4. Here a two-step descriptor selection procedure was applied.

At the first step, the descriptors that are unable to provide useful statistical information were eliminated. For instance, the absolute value of the correlation coefficient between the selected descriptor (X_i) and the activity (Y) under investigation ($|\text{cor}(X_i, Y)|$) should be larger than 0.2. Since the activities of trapping $\bullet NO$, $\bullet OH$ and $\bullet H_2O_2$ are highly correlated, $-\log EC_{50}/\bullet OH$ was set as Y.

At the second step, the descriptor subset was optimized by using a genetic algorithm (GA) [25]. As a meta-heuristic method for getting an optimized solution of a given problem,

the central idea of GA is that a population of potential solutions is refined iteratively by using a strategy inspired by the Darwinist method or natural selection. With a preference to find "fittest" individuals, the selection method from a population of potential solutions has given these types of algorithms the name "genetic". The individuals in a population are frequently called "chromosomes", consisting of "genes" that represent the properties of the individual, and the function to optimize is called a "fitness" function. Each iteration is called a "generation". Concerning feature selection, for example, a chromosome contains a high number of genes (as many as the variables), each of them being just 1 bit long ("0" means variable absent, "1" indicates variable present). In this work, a GA program previously developed in our laboratory was used for the variable selection with the following configuration: (1) genetic operators: selection process of stochastic sampling with replacement [25], simple crossover and random mutation [26]; (2) fitness function: regression coefficient (R); (3) number of variables: 4; (4) population size: 500; (5) generation: 5000; (6) crossover probability: 80%; and (7) mutation probability: 10%.

Finally, the optimal QSAR equations were determined. The detailed values for the nine descriptors used in the above three equations are listed in Table 1, and the definitions of these descriptors are summarized in Table 2.

3. Results and discussion

3.1. Correlation between determined $\log P$ and retention time

In the present study, we have measured $\log P$ values and retention time (alternatively represented by $\log K$) for compounds **1**−**20** (cf. Table 3). To predict the $\log P$ value by

Table 2
The nine molecular descriptors reported in the improved QSAR analysis[a]

Abbreviation	Type	Definition
Ms	Constitutional descriptors	Mean electrotopological state
Ss	Constitutional descriptors	Sum of Kier–Hall electrotopological states
Sv	Constitutional descriptors	Sum of atomic van der Waals volumes (scaled on carbon atom)
Mv	Constitutional descriptors	Mean atomic van der Waals volume (scaled on carbon atom)
nAT	Constitutional descriptors	Number of atoms
ZM1	Topological descriptors	First Zagreb index M1
MSD	Topological descriptors	Mean square distance index (Balaban)
Har2	Topological descriptors	Square reciprocal distance sum index
Qmean	Charge descriptors	Mean absolute charge (charge polarization)

[a] The definitions of these nine descriptors were downloaded from the website of e-dragon (http://www.vcclab.org/lab/edragon/).

using the $\log K$ value, the least square regression was employed to generate Eq. (6).

$$\log P = 0.089 + 1.382 \log K$$
$$N = 20, \ R = 0.982, \ S = 0.123, \ \overline{e} = 0.088, \quad (6)$$
$$R_{LOO} = 0.978, \ S_{LOO} = 0.136, \ \overline{e}_{LOO} = 0.098$$

The detailed results of the resubstitution and LOO tests are listed in Table 3. In addition to high accuracy obtained in the above validation, Eq. (3) was used for predicting $\log P$ values of the free radical scavengers **21** and **22** (Table 4). Using the same procedure, the $\log P$ and $\log K$ values for **21** and **22**

were experimentally determined, the calculated $\log P$ values derived from Eq. (6) and the predictive errors are summarized in Table 4. In contrast, it was found that the errors of the predicted $\log P$ values of the free radical scavengers **21** and **22** were smaller than the absolute average error of the free radical scavengers **1–20** in the LOO test. Based on the above evaluations, we believe that Eq. (6) can be practically used for predicting $\log P$ values of the nitronyl nitroxides. Compared with the measurement of $\log K$, the determination of $\log P$ for a molecule is much more time-consuming. Therefore, the above results suggest that the time-consuming determination of $\log P$ can be replaced by the simple HPLC analysis.

3.2. QSAR analysis of 1–20 via Hansch approach

In our previous paper [16], the •NO, •H_2O_2 and •OH scavenging activities of compounds **1–20** were measured by PC12 cell survival assay and represented by EC_{50}. The corresponding EC_{50} values reported in our previous paper are listed in Table 5 and were used in the current QSAR study.

Regarding the Hansch equations of **1–20** as •NO scavenger, the QSAR Eq. (7) could give the highest predictive accuracy.

$$-\log EC_{50} = 4.135 + 0.310 MR - 0.376 \sum \sigma$$
$$N = 20, \ R = 0.787, \ S = 0.195, \ \overline{e} = 0.118, \quad (7)$$
$$R_{LOO} = 0.687, \ S_{LOO} = 0.231, \ \overline{e}_{LOO} = 0.140$$

This equation contains only the sum of Hammett substituent parameter ($\sum \sigma$) and molar refractivity (MR). As shown in Table 6, if the parameters $\log P$ and $(\log P)^2$ were involved the absolute average error ($\overline{e}$) in the resubstitution test

Table 3
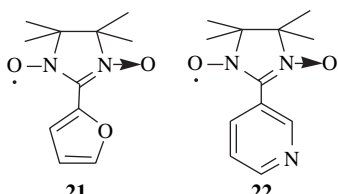The tested $\log P$, $\log K$ and predicted $\log P$ values for compounds **1–20**

Compound	Tested $\log P$[a]	Tested $\log K$	Resubstitution test		LOO test	
			Predicted $\log P$	Absolute error	Predicted $\log P$	Absolute error
1	1.365 ± 0.004	0.918	1.357	0.008	1.357	0.008
2	1.169 ± 0.009	0.729	1.096	0.073	1.092	0.077
3	1.141 ± 0.003	0.546	0.843	0.298	0.818	0.323
4	0.619 ± 0.002	0.483	0.756	0.137	0.769	0.150
5	1.897 ± 0.007	1.309	1.898	0.001	1.898	0.001
6	1.622 ± 0.005	1.101	1.610	0.012	1.609	0.013
7	0.770 ± 0.004	0.620	0.946	0.176	0.958	0.188
8	1.146 ± 0.004	0.905	1.339	0.193	1.350	0.204
9	1.508 ± 0.008	1.039	1.525	0.017	1.526	0.018
10	0.918 ± 0.004	0.534	0.827	0.091	0.819	0.099
11	1.135 ± 0.002	0.558	0.860	0.275	0.837	0.298
12	2.015 ± 0.023	1.328	1.924	0.091	1.915	0.100
13	0.822 ± 0.007	0.577	0.797	0.025	0.891	0.069
14	2.330 ± 0.011	1.600	2.300	0.030	2.294	0.036
15	2.461 ± 0.025	1.713	2.456	0.005	2.455	0.006
16	1.932 ± 0.019	1.347	1.950	0.018	1.952	0.020
17	0.909 ± 0.004	0.598	0.915	0.006	0.916	0.007
18	1.636 ± 0.012	1.157	1.688	0.052	1.691	0.055
19	1.684 ± 0.016	1.245	1.809	0.125	1.820	0.136
20	-0.518 ± 0.003	-0.372	-0.425	0.093	-0.360	0.158
Absolute average error ($\overline{e}$)[b]			0.088 ± 0.089		0.098 ± 0.101	

[a] $n = 3$, $\log P$ value for each compound was presented by $\overline{x} \pm SD$.
[b] Absolute average error ($\overline{e}$) was described as $\overline{x} \pm SD$.

Table 4
Chemical structures and measured $\log P$, $\log K$ and predicted $\log P$ values for compounds **21** and **22**

Compound	Tested $\log K$	Tested $\log P$	Predicted $\log P$	Absolute error
21	0.405	0.615	0.648	0.033
22	0.337	0.512	0.555	0.043

decreased, while the $\bar{e}_{\text{LOO}}$ increased. This phenomenon implied that an over-fitting may exist when the $\log P$ and $(\log P)^2$ are incorporated in the Hansch method. Therefore, Eq. (7) is the optimal Hansch equation for predicting •NO scavenging ability, and the detailed results based on the resubstitution and LOO tests are summarized in Table 7.

Regarding Hansch equations of **1–20** as •H$_2$O$_2$ scavenger, Eq. (8) is the optimal QSAR model for predicting •H$_2$O$_2$ scavenging activity.

$$-\log \text{EC}_{50} = 4.008 + 0.157\text{MR} - 0.236\sum\sigma$$
$$+ 0.286\log P - 0.101(\log P)^2$$
$$N = 20, R = 0.919, S = 0.063, \bar{e} = 0.053, \quad (8)$$
$$R_{\text{LOO}} = 0.827, S_{\text{LOO}} = 0.090, \bar{e}_{\text{LOO}} = 0.076$$

Eq. (8) contains four parameters ($\sum\sigma$, MR, $\log P$ and $(\log P)^2$). The detailed predictive results of compounds **1–20** are summarized in Table 8. In comparison to the QSAR

Table 5
The EC$_{50}$ (μm) of nitronyl nitroxides **1–20** as •NO, •H$_2$O$_2$ and •OH scavenger[a]

Compound	EC$_{50}$/•NO	EC$_{50}$/•H$_2$O$_2$	EC$_{50}$/•OH
1	67.61 ± 4.22	51.63 ± 2.41	66.51 ± 2.33
2	44.67 ± 0.82	52.83 ± 3.26	55.32 ± 2.77
3	61.66 ± 2.42	57.93 ± 2.86	69.64 ± 2.88
4	39.81 ± 1.52	48.18 ± 2.44	56.03 ± 2.47
5	33.11 ± 1.53	38.27 ± 1.91	64.99 ± 3.15
6	30.90 ± 1.28	42.94 ± 1.78	59.28 ± 2.54
7	9.33 ± 0.31	25.29 ± 0.72	27.73 ± 1.32
8	8.32 ± 0.25	30.40 ± 1.42	29.71 ± 1.24
9	19.50 ± 2.24	29.71 ± 1.46	47.08 ± 2.11
10	27.54 ± 1.17	40.08 ± 2.05	54.06 ± 2.48
11	25.70 ± 1.43	32.57 ± 1.63	42.94 ± 2.03
12	9.12 ± 0.14	29.03 ± 1.48	31.11 ± 1.47
13	44.67 ± 1.75	76.36 ± 3.82	85.68 ± 3.24
14	63.10 ± 3.57	58.88 ± 2.73	74.62 ± 2.56
15	46.77 ± 2.38	55.32 ± 2.05	60.66 ± 2.97
16	42.76 ± 1.97	64.99 ± 3.15	78.14 ± 3.85
17	47.86 ± 2.10	60.66 ± 3.03	74.62 ± 3.52
18	87.10 ± 2.81	62.07 ± 2.95	63.52 ± 3.02
19	114.82 ± 4.35	107.86 ± 4.37	167.06 ± 7.91
20	91.20 ± 2.23	74.62 ± 4.12	100.67 ± 4.56

[a] $n = 6$, EC$_{50} = \bar{x} \pm$ SD μm.

equation using two parameters ($\sum\sigma$ and MR) only, using Eq. (8) the $\bar{e}_{\text{LOO}}$ decreased from 0.088 to 0.076 (Table 6).

For Hansch equations of **1–20** as •OH scavenger, QSAR Eq. (9) containing two parameters ($\sum\sigma$ and MR) also resulted in an optimal predictive accuracy (cf. Table 9).

$$-\log \text{EC}_{50} = 4.109 + 0.119\text{MR} - 0.242\sum\sigma$$
$$N = 20, R = 0.776, S = 0.112, \bar{e} = 0.077, \quad (9)$$
$$R_{\text{LOO}} = 0.636, S_{\text{LOO}} = 0.137, \bar{e}_{\text{LOO}} = 0.093$$

Similar to the Hansch equations established for predicting •NO scavenging activity, incorporating the $\log P$ and $(\log P)^2$ terms into the regression equation the average error in the resubstitution test decreased, while the average error for the LOO test increased (Table 6).

In our previous work [16], we demonstrated that introducing 4-bromine, 2-fluorine, 4-dimethylamino, 2,4-dichloro, 3,4-dimethoxyl, 2,4-dimethoxyl, 3-methoxy-4-hydroxyl and 3-hydroxy-4-methoxyl into the phenyl resulted in high radical scavenging activity; introducing 4-methyl and 4-methoxyl into the phenyl resulted in moderate radical scavenging activity; and introducing 4-nitro, and 4-carboxymethyl into the phenyl resulted in lower radical scavenging activity. These observations imply that the electron character and solubility of the introduced substituent contribute a lot to the radical scavenging activity of **1–20**. Since **2–20** can be considered as homologous compounds of **1** and the substituent alternation in the phenyl results in no stereochemistry change, thus the QSAR analysis of **1–20** may be simply treated by the classical Hansch approach. The QSAR analysis mentioned above demonstrated that the overall accuracy for predicting the activity of •H$_2$O$_2$ and •OH scavenger is reasonable, although the model is still needed to be improved to obtain a satisfactory result for the prediction of scavenging •NO activity.

The above QSAR analysis also demonstrated that the $\log P$ value for individual compound, representing the solubility of the substituent in the phenyl to a certain level, does not significantly relate to its activity of trapping different free radicals, although introducing $\log P$ and $(\log P)^2$ had slightly improved the accuracy predicting •NO scavenging activity. The $\log P$ is a property that plays a relevant role when a cellular system is involved, because it can encode the ability of the compound in penetrating the cellular membrane, mimicking a passive transportation. In respect to the activity of a free radical scavenger, it is of low concern with the scavenger's ability to pass through a cell membrane. Therefore, the overall contribution for the parameter $\log P$ in the current QSAR analysis is weak.

Using Hansch approach, the prediction of compound **20** generates a much larger error in comparison with the prediction of other compounds (Tables 7–9). Indeed, the carboxyl group in compound **20** was converted into carboxylate ion at the pH of the cultured medium of PC12 cells, and this situation cannot be reflected in the parameter of $\sum\sigma$, resulting in a significant predictive error existing for compound **20**.

Table 6
The absolute average error for predicting the activity ($-\log EC_{50}$) of nitronyl nitroxides **1–20** as •NO, •H$_2$O$_2$ and •OH scavenger in Hansch equations by choosing different parameters

Parameters	Scavenging •NO		Scavenging •H$_2$O$_2$		Scavenging •OH	
	Resubstitution	LOO	Resubstitution	LOO	Resubstitution	LOO
$\sum \sigma$, MR	0.118	0.140	0.074	0.088	0.077	0.093
$\sum \sigma$, MR, log P	0.132	0.184	0.080	0.107	0.084	0.117
$\sum \sigma$, MR, log P, (log P)^2	0.095	0.167	0.053	0.076	0.062	0.108

3.3. Improved QSAR analysis

As described in Section 2, nine molecular descriptors calculated from the e-dragon web-server were finally used in the construction of the improved QSAR equations (cf. Tables 1 and 2). Of them, the Ms, Mv, Sv, Ss and nAT belong to the simple constitutional descriptors [27]. The ZM1, Har2 and MSD descriptors are grouped into the topological descriptors [28–30], while the Qmean parameter is one of the charge descriptors [31].

For the QSAR model about •NO scavengers, the four descriptors (Ms, Sv, Har2 and Ss) can result in regression equation (10) with the highest predictive accuracy.

$$-\log EC_{50} = -14.729 + 7.886Ms + 0.308Sv$$
$$+ 0.194Har2 - 0.419Ss$$
$$N = 20, R = 0.981, S = 0.062, \bar{e} = 0.050,$$
$$R_{LOO} = 0.971, S_{LOO} = 0.076, \bar{e}_{LOO} = 0.064$$

(10)

The detailed results of the resubstitution and LOO tests are listed in Table 7. Compared with the results based on the Hansch method (i.e. Eq. (7)), the predictive accuracy has been significantly improved.

In regard to the QSAR analysis of **1–20** as •H$_2$O$_2$ scavengers, the linear equation (11) was obtained.

$$-\log EC_{50} = 9.321 - 0.791Ms - 4.352Mv - 6.312MSD$$
$$+ 0.010ZM1$$
$$N = 20, R = 0.964, S = 0.043, \bar{e} = 0.034,$$
$$R_{LOO} = 0.941, S_{LOO} = 0.054, \bar{e}_{LOO} = 0.043$$

(11)

Compared with the Hansch method, the results have also been considerably improved (cf. Table 8).

For the QSAR equation of **1–20** as •OH scavengers, the regression equation (12) containing four descriptors (Ms, nAT, Qmean and MSD) can result in an optimal result.

Table 7
The validation of QSAR analysis for •NO scavenging activity of compounds **1–20** based on the Hansch approach and the improved QSAR method

Compound	$-\log EC_{50}$	Hansch method[a]				Improved QSAR method[b]			
		Resubstitution		LOO test		Resubstitution		LOO test	
		Predicted	Absolute error	Predicted	Absolute error	Predicted	Absolute error	Predicted	Absolute error
1	4.170	4.228	0.058	4.241	0.071	4.163	0.007	4.161	0.009
2	4.350	4.426	0.076	4.443	0.083	4.378	0.028	4.382	0.032
3	4.210	4.239	0.029	4.243	0.033	4.298	0.088	4.310	0.100
4	4.400	4.423	0.023	4.428	0.028	4.243	0.157	4.221	0.179
5	4.480	4.438	0.042	4.434	0.046	4.519	0.039	4.527	0.047
6	4.510	4.544	0.034	4.546	0.036	4.431	0.079	4.418	0.092
7	5.030	4.859	0.171	4.815	0.215[c]	5.088	0.058	5.111	0.081
8	5.080	4.713	0.367[c]	4.637	0.443[c]	4.995	0.085	4.967	0.113
9	4.710	4.505	0.205[c]	4.493	0.217[c]	4.690	0.020	4.638	0.072
10	4.560	4.554	0.006	4.554	0.006	4.522	0.038	4.518	0.042
11	4.590	4.592	0.002	4.592	0.002	4.563	0.027	4.559	0.031
12	5.040	4.990	0.050	4.969	0.071	5.078	0.038	5.107	0.067
13	4.350	4.203	0.147	4.170	0.180	4.284	0.066	4.257	0.093
14	4.200	4.297	0.097	4.304	0.104	4.227	0.027	4.233	0.033
15	4.330	4.387	0.057	4.391	0.061	4.362	0.032	4.368	0.038
16	4.370	4.366	0.004	4.365	0.005	4.472	0.102	4.481	0.111
17	4.320	4.134	0.186	4.092	0.228[c]	4.300	0.020	4.291	0.029
18	4.060	4.160	0.100	4.179	0.119	4.110	0.050	4.123	0.063
19	3.940	3.961	0.021	3.975	0.035	3.974	0.034	3.984	0.044
20	4.040	4.722	0.682[c]	4.839	0.799[c]	4.042	0.002	4.041	0.001
Absolute average error ($\bar{e}$)[d]		0.118 ± 0.160		0.140 ± 0.188		0.050 ± 0.037		0.064 ± 0.042	

[a] Eq. (4) was established by using $\sum \sigma$ and MR.
[b] Eq. (10) was established by using Ms, Sv, Har2 and Ss.
[c] Absolute error > 0.20.
[d] Absolute average error ($\bar{e}$) was described as $\bar{x} \pm SD$.

Table 8
The validation of QSAR analysis of compounds **1**–**20** as $\cdot H_2O_2$ scavenger based on the Hansch approach and the improved QSAR method

Compound	$-\log EC_{50}$	Hansch method[a]				Improved QSAR[b]			
		Resubstitution		LOO test		Resubstitution		LOO test	
		Predicted	Absolute error	Predicted	Absolute error	Predicted	Absolute error	Predicted	Absolute error
1	4.290	4.257	0.033	4.249	0.041	4.332	0.042	4.347	0.057
2	4.280	4.368	0.088	4.387	0.107	4.303	0.023	4.307	0.027
3	4.240	4.249	0.009	4.251	0.011	4.265	0.025	4.269	0.029
4	4.320	4.309	0.011	4.306	0.014	4.221	0.099	4.204	0.116
5	4.420	4.346	0.074	4.335	0.085	4.395	0.025	4.390	0.030
6	4.370	4.424	0.054	4.429	0.059	4.397	0.027	4.402	0.032
7	4.600	4.559	0.041	4.543	0.057	4.614	0.014	4.618	0.018
8	4.520	4.501	0.019	4.495	0.025	4.576	0.056	4.591	0.071
9	4.530	4.403	0.127	4.393	0.137	4.522	0.008	4.518	0.012
10	4.400	4.404	0.004	4.405	0.005	4.390	0.010	4.388	0.012
11	4.490	4.444	0.046	4.439	0.051	4.421	0.069	4.410	0.080
12	4.540	4.642	0.102	4.713	0.173	4.525	0.015	4.519	0.021
13	4.120	4.206	0.086	4.230	0.110	4.173	0.053	4.184	0.064
14	4.230	4.195	0.035	4.179	0.051	4.221	0.009	4.219	0.011
15	4.260	4.214	0.046	4.185	0.075	4.233	0.027	4.218	0.042
16	4.190	4.277	0.087	4.295	0.105	4.205	0.015	4.217	0.027
17	4.220	4.148	0.072	4.122	0.098	4.219	0.001	4.217	0.003
18	4.210	4.185	0.025	4.179	0.031	4.136	0.074	4.116	0.094
19	3.970	4.057	0.087	4.116	0.146	4.042	0.072	4.071	0.101
20	4.130	4.141	0.011	4.275	0.145	4.141	0.011	4.157	0.027
Absolute average error $(\bar{e})$[c]		0.053 ± 0.036		0.076 ± 0.050		0.034 ± 0.027		0.046 ± 0.033	

[a] Eq. (5) was established by using $\sum\sigma$, MR, $\log P$ and $(\log P)^2$.
[b] Eq. (11) was established by using Ms, Mv, MSD and ZM1.
[c] Absolute average error $(\bar{e})$ was described as $\bar{x} \pm SD$.

Table 9
The validation of QSAR analysis of compounds **1**–**20** as $\cdot OH$ scavenger based on the Hansch approach and the improved QSAR method

Compound	$-\log EC_{50}$	Hansch method[a]				Improved QSAR method[b]			
		Resubstitution		LOO test		Resubstitution		LOO test	
		Predicted	Absolute error	Predicted	Absolute error	Predicted	Absolute error	Predicted	Absolute error
1	4.180	4.145	0.035	4.138	0.042	4.195	0.015	4.198	0.018
2	4.260	4.258	0.002	4.257	0.003	4.261	0.001	4.260	0.000
3	4.160	4.138	0.022	4.135	0.025	4.214	0.054	4.229	0.069
4	4.250	4.257	0.007	4.258	0.008	4.158	0.092	4.114	0.136
5	4.190	4.243	0.053	4.247	0.057	4.296	0.106	4.312	0.122
6	4.230	4.293	0.063	4.298	0.068	4.205	0.025	4.197	0.033
7	4.560	4.441	0.119	4.411	0.149	4.581	0.021	4.592	0.032
8	4.530	4.346	0.184	4.308	0.222[c]	4.534	0.004	4.535	0.005
9	4.330	4.268	0.062	4.264	0.066	4.320	0.010	4.318	0.012
10	4.270	4.286	0.016	4.287	0.017	4.291	0.021	4.293	0.023
11	4.370	4.310	0.060	4.305	0.065	4.331	0.039	4.326	0.044
12	4.510	4.520	0.010	4.524	0.014	4.489	0.021	4.476	0.034
13	4.070	4.130	0.060	4.143	0.073	4.070	0.000	4.070	0.000
14	4.130	4.149	0.019	4.150	0.020	4.110	0.020	4.106	0.024
15	4.220	4.184	0.036	4.181	0.039	4.156	0.064	4.139	0.081
16	4.110	4.153	0.043	4.160	0.050	4.144	0.034	4.150	0.040
17	4.130	4.032	0.098	4.010	0.120	4.157	0.027	4.175	0.045
18	4.200	4.049	0.151	4.021	0.179	4.055	0.145	4.016	0.184
19	3.780	3.921	0.141	4.012	0.232[c]	3.937	0.157	4.013	0.233[c]
20	4.000	4.359	0.359[c]	4.420	0.420[c]	3.976	0.024	3.929	0.071
Absolute average error $(\bar{e})$[d]		0.077 ± 0.084		0.093 ± 0.103		0.044 ± 0.046		0.060 ± 0.062	

[a] Eq. (6) was established by using $\sum\sigma$ and MR.
[b] Eq. (12) was established by using Ms, nAT, Qmean and MSD.
[c] Absolute error > 0.20.
[d] Absolute average error $(\bar{e})$ was described as $\bar{x} \pm SD$.

$$-\log EC_{50} = -0.039 - 0.648Ms + 0.119nAT + 122.092$$
$$\times Qmean - 7.880MSD$$
$$N = 20, \; R = 0.936, \; S = 0.063, \; \bar{e} = 0.044,$$
$$R_{LOO} = 0.874, \; S_{LOO} = 0.086, \; \bar{e}_{LOO} = 0.060$$

(12)

In comparison with the results obtained from Hansch method, the predictive accuracy is also improved with an average error in the LOO test decreasing from 0.093 to 0.060 (cf. Table 9).

Since the fitness function of the GA program used in this study is the regression coefficient R, and the LOO test was performed after the parameters were selected, the LOO results may not really represent the predictive ability of the established models. To clarify this, we did two additional computational experiments. Firstly, the GA-based variable selection was performed again by using the R_{LOO} as the fitness function. It turned out that the molecular descriptors used in Eqs. (10)–(12) remained to be ranked as the top solutions. In the second test, we added a Leave-Two-Out test to assess the performance

of these improved QSAR models. The resulted $\bar{e}$ values (absolute average errors) are 0.062, 0.060 and 0.067, respectively. Compared with the corresponding values in the LOO test (cf. Eqs. (10)–(12)), the $\bar{e}$ values only slightly increased. Based on the above two tests, it can be envisaged that the improved QSAR model based on these descriptors should expect a similar predictive accuracy when it is tested for an independent set of the derivatives of 2-substituted phenylnitronyl nitroxides.

3.4. Comparison of the Hansch equations and the dragon descriptor-based QSAR models

In comparison to the Hansch approach, the QSAR analysis involving new molecular descriptors has significantly improved the performance. The results are illustrated in Fig. 1. In addition to the high predictive accuracy, the improved QSAR models also shed some light on the relationship between the structure and the activity. Among these nine
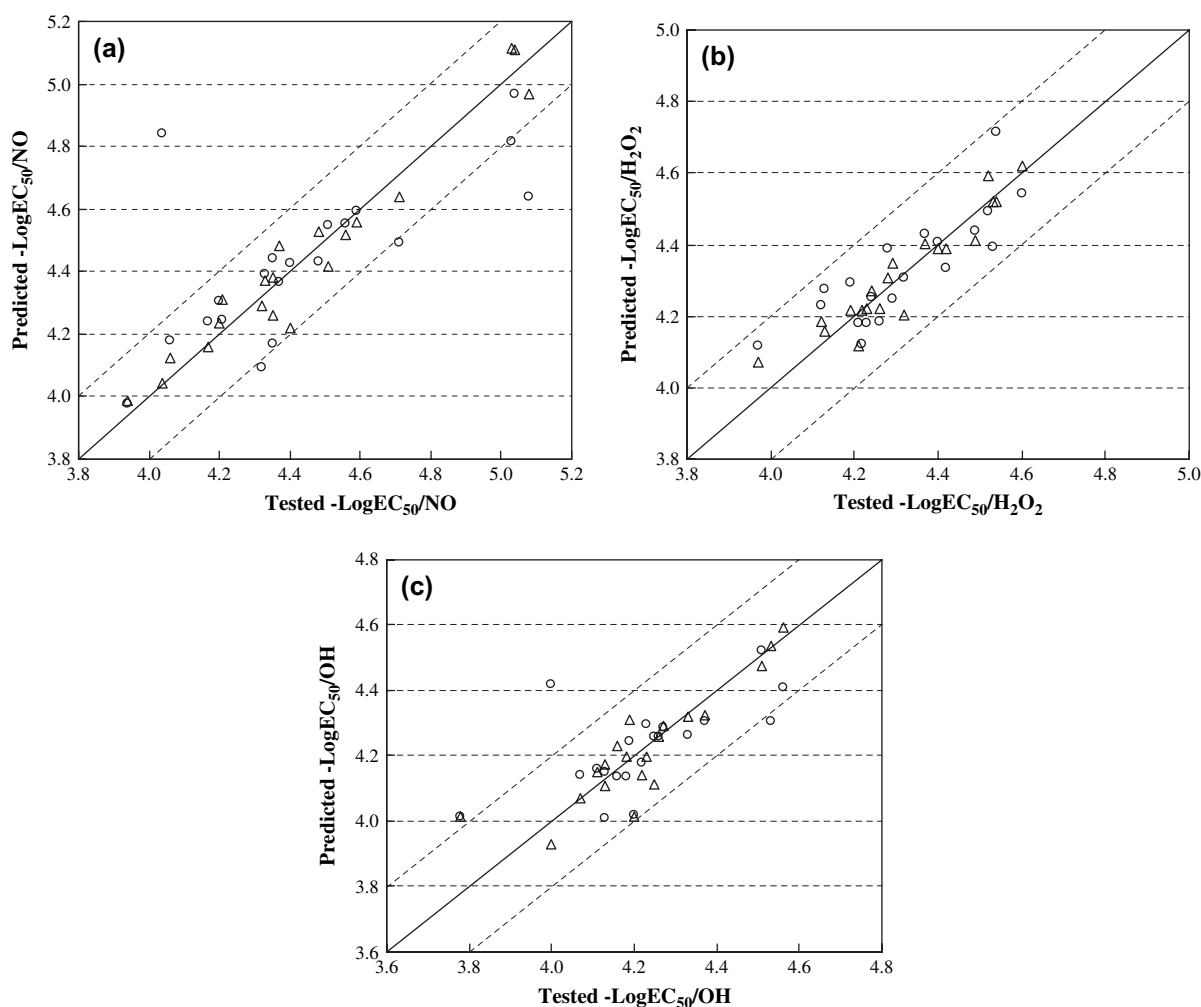

Fig. 1. Comparison of the activity of compounds 1–20 predicted by the LOO test with the observed activity. The data points represented as circles are the results from the Hansch approach, while those represented as triangles are the results from the improved QSAR analysis. The points situated at the zone between the two diagonal-dash lines correspond to the compounds, in which the prediction errors (i.e. absolute errors) are less than 0.20. (a) The QSAR analysis of compounds 1–20 as •NO scavenger; (b) the QSAR analysis of compounds 1–20 as •H_2O_2 scavenger; (c) the QSAR analysis of compounds 1–20 as •OH scavenger.

descriptors, Ms, Ss and Qmean are directly related to a molecule's electrostatic property. Especially, the Ms descriptor, representing the mean electrotopological state of a molecule, is consistently used in three QSAR equations. The other six descriptors are related to the volume or topology of a molecule. Therefore, the activities of free radical scavengers **1–20** can be reasonably explained by their electrostatic and steric effects, which is in line with the preliminary finding based on the Hansch approach.

To illustrate the quantitative relationship of the $•NO$, $•H_2O_2$ and $•OH$ trapping activities, the three parameters in Hansch equations and the nine molecular descriptors used in the improved QSAR models, the correlation matrix of these 15 parameters was calculated (cf. Table 10). The three activities are shown to be highly correlated with an average correlation coefficient around 0.90, implying the molecular mechanisms of scavenging three free radicals should be similar. Here, we defined that two parameters sharing a correlation coefficient ≥ 0.5 or ≤ -0.5 are significantly correlated. Regarding the parameters in Hansch equations (i.e. Eqs. (7)–(9)), the MR and $\sum\sigma$ share significant correlation with the activities, and the log P is poorly correlated with the activities. As discussed above, the activities can be then reasonably predicted by the MR and $\sum\sigma$, while the log P has minor contribution in the Hansch models. Concerning the four parameters (Ms, Sv, Har2 and Ss) selected in predicting $•NO$ scavenging activity (Eq. (10)), the Ms and Sv share significant correlation with the activity. Furthermore, the Ms is significantly correlated with $\sum\sigma$, while Sv and Har2 are significantly correlated with MR. Therefore, the information used in the improved QSAR model is overlapped with that in the Hansch model (Eq. (7)), but more delicate parameters in Eq. (10) resulted in a higher predictive accuracy. The similar phenomenon was also observed in analyzing the parameters in Eqs. (11) and (12). Although the different parameters were used in Eqs. (10)–(12), the molecular information embedded in these three equations are quite similar. Taking the comparison of Eqs. (10) and (12) as an example, the Ms descriptor is consistently used in two equations. Additionally, the Sv and Har2 in Eq. (10) are highly correlated with the nAT and Qmean in Eq. (12). The similar phenomenon can be also found when Eq. (11) was compared with Eqs. (10) and (12). Therefore, the above correlation matrix quantitatively revealed that the molecular mechanisms of compounds **1–20** as scavengers of different radicals should be similar. Either in the Hansch equations or in the improved QSAR models, the physicochemical information embedded in these selected parameters is close. The electrostatic and steric effects of a compound have major contributions to its activity. Compared with the Hansch equations, more delicate descriptors were employed in the improved QSAR models and thus a higher predictive accuracy was reached.

Thanks to the community of chemoinformatics, some web-servers such as e-dragon are well compiled to calculate the different molecular descriptors with the input of a molecule's structural information. Note that in the current analysis we only focused on the 210 descriptors from four different

Table 10
The correlation matrix between the parameters of QSAR models, including three activities, three parameters in Hansch models and nine descriptors in improved QSAR equations[a]

	$-\log EC_{50}/•NO$	$-\log EC_{50}/•H_2O_2$	$-\log EC_{50}/•OH$	MR	$\sum\sigma$	log P	Ms	nAT	Qmean	Sv	Mv	MSD	Har2	Ss	ZM1
$-\log EC_{50}/•NO$	1.000														
$-\log EC_{50}/•H_2O_2$	0.895	1.000													
$-\log EC_{50}/•OH$	0.901	0.935	1.000												
MR	0.569	0.442	0.455	1.000											
$\sum\sigma$	-0.668	-0.737	-0.722	-0.246	1.000										
log P	0.045	0.057	0.039	-0.066	0.235	1.000									
Ms	-0.639	-0.711	-0.632	-0.209	0.673	-0.419	1.000								
nAT	0.680	0.618	0.589	0.823	-0.522	-0.361	-0.236	1.000							
Qmean	-0.649	-0.591	-0.542	-0.837	0.513	0.343	0.264	-0.989	1.000						
Sv	0.644	0.554	0.534	0.937	-0.428	-0.278	-0.228	0.966	-0.972	1.000					
Mv	-0.470	-0.551	-0.495	-0.184	0.595	0.452	0.189	-0.691	0.644	-0.484	1.000				
MSD	-0.199	-0.172	-0.269	0.326	-0.162	-0.135	-0.078	0.333	-0.394	0.367	-0.165	1.000			
Har2	0.371	0.269	0.259	0.810	-0.034	-0.432	0.244	0.784	-0.780	0.843	-0.295	0.172	1.000		
Ss	-0.179	-0.280	-0.245	0.425	0.376	-0.554	0.762	0.392	-0.380	0.438	-0.092	0.155	0.807	1.000	
ZM1	0.356	0.274	0.231	0.788	-0.031	-0.377	0.215	0.741	-0.743	0.811	-0.237	0.203	0.983	0.777	1.000

[a] Ms, Sv, Har2 and Ss were used to establish the QSAR model of **1–20** as $•NO$ scavengers (i.e. Eq. (10)). Ms, Mv, MSD and ZM1 were used to establish the QSAR model of **1–20** as $•H_2O_2$ scavengers (i.e. Eq. (11)). Ms, nAT, Qmean and MSD were used to establish the QSAR model of **1–20** as $•OH$ scavengers (i.e. Eq. (12)).

types of e-dragon descriptors. Selecting other descriptors from the available 1664 e-dragon descriptors may further improve the current predictive accuracy. However, the physicochemical information embedded in these descriptors may be more complicated and thus it remains a challenge to provide further insight in the relationship between structure and activity. In addition to the multiple linear regression method, state-of-the-art machine learning methods may also be helpful in establishing a better QSAR model. Judged from the overall performance obtained in the above improved QSAR analysis, however, we argue that the current QSAR model has reached a level to be used for screening novel scavengers from the derivatives of 2-substituted phenylnitronyl nitroxides.

4. Conclusions

In this study we measured the $\log P$ values and retention time of 20 2-substituted phenylnitronyl nitroxides by the shake-flask method and HPLC analysis, respectively. It was found that the highly accurate $\log P$ can be predicted from the tested $\log K$ by using the linear regression equation between $\log P$ and $\log K$. Consequently, the results suggested that the determination of time-consuming $\log P$ can be replaced by measuring $\log K$ based on the simple HPLC analysis. Using classical Hansch approach, some QSAR equations were obtained for predicting the activities of scavenging free radicals ($\cdot$NO, $\cdot H_2O_2$ and $\cdot$OH) with reasonable accuracy. It was also found that the contribution of $\log P$ to the activity is negligible, which may be considered due to the fact that the $\cdot$NO, $\cdot H_2O_2$ and $\cdot$OH injury occurred mainly on the surface of the cells and thus the procedure of scavenging free radicals may have less concern with a scavenger's ability to pass through a cellular membrane. To improve the QSAR analysis new molecular descriptors generated from e-dragon server were used. Compared with Hansch approach, the predictive results have been considerably improved. Thus, the QSAR equations based on the new molecular descriptors can be practically used for directing the modification of the free radical scavenger of 2-substituted phenylnitronyl nitroxide.

Acknowledgements

References

[1] G. Vendemiale, I. Grattagliano, E. Altomare, Int. J. Clin. Lab. Res. 29 (1999) 49−55.
[2] M. Valko, C.J. Rhodes, J. Moncol, M. Izakovic, M. Mazur, Chem. Biol. Interact. 160 (2006) 1−40.
[3] M.S. Cooke, M.D. Evans, M. Dizdaroglu, J. Lunec, FASEB J 17 (2003) 1195−1214.
[4] R. Terranova, R. Sorace, A. Romeo, C.D. Mauro, R. Romeo, S. Luca, Minerva Med. 92 (2001) 405−410.
[5] C.W. Olanow, Ann. Neurol. 32 (Suppl.) (1992) S2−S9.
[6] D.D. Bankson, M. Kestin, N. Rifai, Clin. Lab. Med. 13 (1993) 463−480.
[7] S. Jaswal, H.C. Mehta, A.K. Sood, J. Kaur, Clin. Chim. Acta 338 (2003) 123−129.
[8] J.C. Mak, M.M. Chan-Yeung, Curr. Opin. Pulm. Med. 12 (2006) 7−11.
[9] S.S. Sheu, D. Nauduri, M.W. Anders, Biochim. Biophys. Acta 1762 (2006) 256−265.
[10] R.E. Anderson, F.L. Kretzer, L.M. Rapp, Adv. Exp. Med. Biol. 366 (1994) 73−86.
[11] J.H. Osiecki, E.F. Ullman, J. Am. Chem. Soc. 90 (1968) 1078−1079.
[12] T. Akaike, M. Yoshida, Y. Miyamoto, K. Sato, M. Kohno, K. Sasamoto, K. Miyazaki, S. Ueda, H. Maeda, Biochemistry 32 (1993) 827−832.
[13] D.E. Koshland, Science 258 (1992) 1861.
[14] R.F. Haseloff, S. Zollner, I.A. Kirilyuk, I.A. Grigor'ev, R. Reszka, R. Bernhardt, K. Mertsch, B. Roloff, I.E. Blasig, Free Radic. Res. 26 (1997) 7−17.
[15] I.E. Blasig, K. Mertsch, R.F. Haseloff, Neuropharmacology 43 (2002) 1006−1014.
[16] Y. Wu, L. Bi, W. Bi, Z. Li, M. Zhao, C. Wang, J. Ju, S. Peng, Bioorg. Med. Chem. 14 (2006) 5711−5720.
[17] R.D. Cramer III, D.E. Patterson, J.D. Bunce, Prog. Clin. Biol. Res. 291 (1989) 161−165.
[18] C. Hansch, Acc. Chem. Res. 2 (1969) 232−239.
[19] C.A. Lipinski, J. Pharmacol. Toxicol. Methods 44 (2001) 235−249.
[20] T. Fujita, J. Iwasa, C. Hansch, J. Am. Chem. Soc. 86 (1964) 5175−5180.
[21] I.V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V.A. Palyulin, E.V. Radchenko, N.S. Zefirov, A.S. Makarenko, V.Y. Tanchuk, V.V. Prokopenko, J. Comput. Aided Mol. Des. 19 (2005) 453−463.
[22] C. Hansch, A. Leo, S.H. Unger, K.H. Kim, D. Nikartani, E.J. Lien, J. Med. Chem. 16 (1973) 1207−1216.
[23] C. Hansch, S.D. Rockwell, P.Y.C. Jow, A. Leo, E.E. Steller, J. Med. Chem. 20 (1977) 304−306.
[24] T. Fujita, T. Nishioka, Prog. Phys. Org. Chem. 12 (1976) 49−89.
[25] J.H. Holland, Adaptation in Natural and Artificial Systems, second ed. MIT Press, Cambridge, USA, 1992.
[26] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer, New York, USA, 1992.
[27] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim, Germany, 2000.
[28] I. Gutman, B. Ruscic, N. Trinajstic, C.F. Wilcox Jr., J. Chem. Phys. 62 (1975) 3399−3405.
[29] A.T. Balaban, Pure Appl. Chem. 55 (1983) 199−206.
[30] O. Ivanciuc, T.-S. Balaban, A.T. Balaban, J. Math. Chem. 12 (1993) 309−318.
[31] M. Karelson, V.S. Lobanov, A.R. Katritzky, Chem. Rev. 96 (1996) 1027−1043.