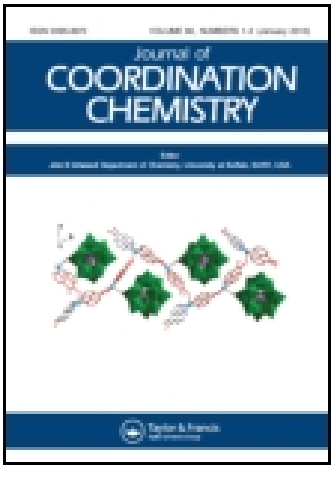


Journal of Coordination Chemistry
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gcoo20

QSAR study of flavonoid-metal complexes scavenging $O_2^{\cdot-}$

Jun-Zhen Qian[ab], Bo-Chu Wang[a], Ying Fan[c], Jun Tan[d] & Hong Ji Huang[e]

[a] Bioengineering College, Chongqing University, Chongqing, China

[b] Laboratory and Equipment Managing Section, Chongqing University, Chongqing, China

[c] Chongqing Telecom Planning & Designing Institute Co. Ltd, Chongqing, China

[d] Department of Biological and Chemical Engineering, Chongqing University of Education, Chongqing, China

[e] Department of Specific Diagnosis, Xinqiao Hospital, Third Military Medical University, Chongqing, China

Accepted author version posted online: 22 Aug 2014.Published online: 23 Sep 2014.

To cite this article: Jun-Zhen Qian, Bo-Chu Wang, Ying Fan, Jun Tan & Hong Ji Huang (2014) QSAR study of flavonoid-metal complexes scavenging $O_2^{\cdot-}$, Journal of Coordination Chemistry, 67:17, 2867-2884, DOI: 10.1080/00958972.2014.956663

To link to this article: http://dx.doi.org/10.1080/00958972.2014.956663

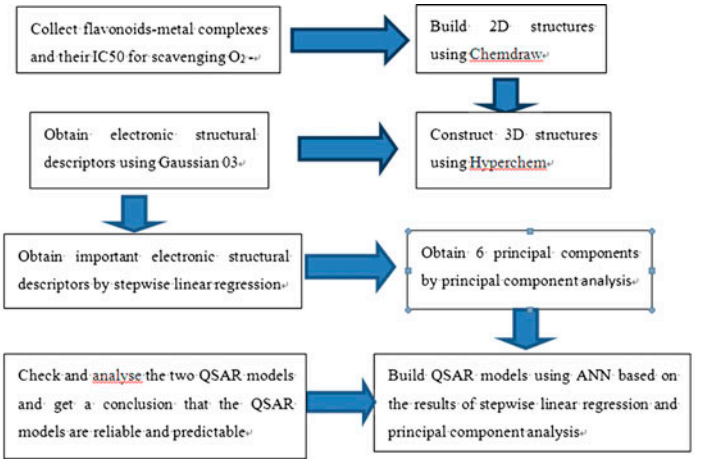
Journal of Coordination Chemistry, 2014
Vol. 67, No. 17, 2867–2884, http://dx.doi.org/10.1080/00958972.2014.956663

QSAR study of flavonoid–metal complexes scavenging $O_2^{\cdot -}$

JUN-ZHEN QIAN*†‡, BO-CHU WANG†, YING FAN§, JUN TAN¶ and HONG JI HUANG‖

†Bioengineering College, Chongqing University, Chongqing, China
‡Laboratory and Equipment Managing Section, Chongqing University, Chongqing, China
§Chongqing Telecom Planning & Designing Institute Co. Ltd, Chongqing, China
¶Department of Biological and Chemical Engineering, Chongqing University of Education, Chongqing, China
‖Department of Specific Diagnosis, Xinqiao Hospital, Third Military Medical University, Chongqing, China

(*Received 24 May 2014; accepted 21 July 2014*)



Flavonoid–metal complexes have antioxidant activities. However, the quantitative structure–activity relationship (QSAR) of flavonoid–metal complexes and their antioxidant activities is not known. On the basis of 31 structures of flavonoid–metal complexes and their antioxidant activities scavenging $O_2^{\cdot -}$, we optimized their structures using the density functional theory method, and subsequently calculated 21 quantum chemistry descriptors such as dipole, charge, and energy. Then, we chose several quantum chemistry descriptors that are very important to the IC_{50} of the antioxidant activities of flavonoid–metal complexes for scavenging $O_2^{\cdot -}$ through stepwise linear regression. We obtained six new variables through the principal component analysis. Finally, we built QSAR models based on those important quantum chemistry descriptors, the six new variables as independent variables, and the IC_{50} as the dependent variable using an artificial neural network. We validated the model using experimental data in references. These results show that the model in this article is reliable and predictive.

*Corresponding author. Email: xiaofei524249@163.com

Keywords: Flavonoid–metal complexes; Quantum chemistry descriptors; Scavenging $O_2^{\cdot-}$; Artificial neural network (ANN); Quantitative structure–activity relationship (QSAR); Antioxidant

1. Introduction

Flavonoids, classified as flavonoid, flavonol, flavonone, flavanonol, isoflavone, etc. according to different structural characteristics, can act as metallic chelators due to their structures [1]. In the most basic sense, flavonoids chelate metal ions to form complexes that have new pharmacological activities or enhance the intrinsic pharmacological activities focusing attention on flavonoid–metal complexes, especially their synthesis and pharmacological properties [2–14]. Many studies about the synthesis of flavonoid–metal complexes have been reported. The flavonoid groups that chelate with metals are 3–OH, 4C=O, 5–OH, or –OH on catechol. Current studies about flavonoid–metal complexes are mainly focused on their anticancer activities and antioxidant activities. For example, some studies show that the bonding constants of flavonoid–metal complexes such as morin-Zn, Cu, Co, rutin-Eu, Cu, Se, chrysin-La, and quercetin-Zn, Mn, Co, Ni, Cu, Pb, La with DNA are bigger than those of flavonoids alone. Some studies show that the IC_{50} of flavonoid–metal complexes such as morin-Mn, Co, Ni, Cu, Zn, quercetin-Zn, and hesperetin-Cu for tumor cells are smaller than those of flavonoids alone. Some studies show that the IC_{50} of flavonoid–metal complexes such as quercetin-Zn, Cu, Co, Ni, hesperetin-Cu, Zn, rutin-Cu, and luteolin-Zn for scavenging $O_2^{\cdot-}$, OH^-, and DPPH are smaller than those of flavonoids alone. These studies all show that the anticancer activities and antioxidant activities of flavonoid–metal complexes are better than those of flavonoids alone. However, activities of the synthesized flavonoid–metal complexes may be bad through a series of experiments according to some studies. Therefore, we constructed a model between the quantum chemistry descriptors of flavonoid–metal complexes and their activities for forecasting the activities of metal complexes, which is helpful for designing novel drugs. In this article, we built the model between the quantum chemistry descriptors of flavonoid–metal complexes and their antioxidant activities and then validate the reliable predictability of the model. Our results provide a solid foundation for the development of flavonoid–metal complexes and their antioxidant activities.

2. Methods

2.1. *Calculation of the electronic structure of flavonoid–metal complexes*

Ligands [2–14] including flavonol, flavonone, and isoflavone were calculated and analyzed. Their 2-D structures were built using Chemdraw and their 3-D structures were constructed and optimized using the MM+ method in Hyperchem. Full geometry optimization of the ligands was conducted with the density functional theory (DFT) method at the DFT/B3LYP theoretical level with the 6-31G*/LanL2DZ basis using the Gaussian 03 computational package [15]. A set of quantum chemistry descriptors for 31 flavonoid–metal complexes was obtained from their electronic structure properties. All calculated results are real frequencies to ensure stable configurations. The fundamental flavonoid structure is shown in figure 1. The structures of flavonoid–metal complexes containing ligands, metal, chelation sites, and ratio are listed in table 1.

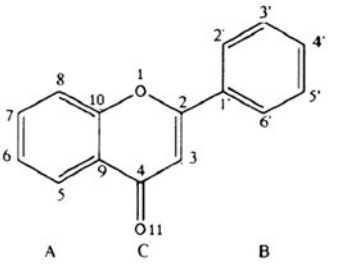


Figure 1. The structure of flavonoids.

2.2. *Stepwise linear regression and principal component analysis*

By studying the relationship between the 21 quantum chemistry descriptors and the IC_{50} of flavonoid–metal complexes for scavenging $O_2^{\cdot -}$ through stepwise linear regression, we can obtain important quantum chemistry descriptors as dependent variables.

Principal component analysis was used to compress a pool of descriptors into principal components as new variables and the new variables can affect the original variables. The model of principal components is given below:

$$F1 = e_{11}x_1 + e_{21}x_2 + \cdots + e_{p1}x_p$$

$$F2 = e_{12}x_1 + e_{22}x_2 + \cdots + e_{p2}x_p$$

$$\mathrm{Fp} = e_1 p_{x1} + e_2 p_{x2} + \cdots + e_{\mathrm{pp}}x_p$$

$x_1, x_2, \ldots, x_p$ are the original variables and $e = (e_{1\mathrm{i}}, e_{2\mathrm{i}}, \ldots, e_{\mathrm{pi}})$ is the coefficient matrix.

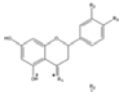
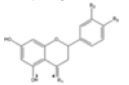
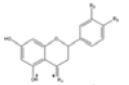
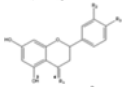
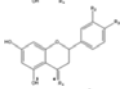
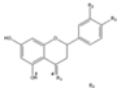
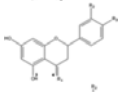
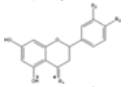
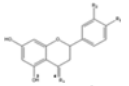
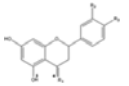
To ensure the integrity of all quantum chemistry descriptors, we used the method of principal component analysis to analyze 21 quantum chemistry descriptors from quantum chemistry calculations. We can further obtain new independent variables that consist of original variables through principal component analysis.

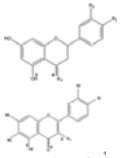
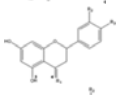
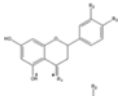
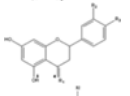
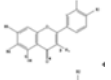
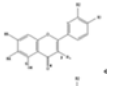
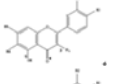
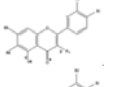
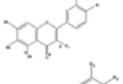
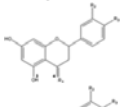
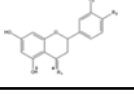
2.3. *Artificial neural network*

An artificial neural network (ANN) is a network model based on the principle of nervous system operation, which has the advantage of having a strong ability to approximate nonlinear functions, self-organization, self-learning, and so on. Back Propagation ANN is typical because this model has several advantages [14].

On the basis of the results of stepwise linear regression and principal component analysis, we chose important quantum chemistry descriptors and F as the independent variable, respectively, and the IC_{50} as the dependent variable, and variables were analyzed by ANN. This ANN program was coded in Matlab 7 for Windows. ANN networks were trained with the training sets and checked with the tested sets. Twenty-five sets of data were selected as training sets (unbolded values) by the back propagation strategy, while the other six data groups were the test sets (bolded values) from tables 2 and 3. The data of the training set and the tested set are randomly selected to ensure the applicability of the model. Based on these data, the models were built and checked with ANN. In the process, the important quantum chemistry descriptors are input parameters that build the vector matrix of 25 × 5,

Table 1. The chelate sites and ratio of flavonoid–metal complexes.

No.	Kind	Ligand	Metal	Basis structure	*R*1	*R*2	*R*3	*R*4	*R*5	Chelation	Ratio (metal/ligand)	Reference
1	Flavanone	Hesperetin derivatives	Zn(II)			–OH	–OCH	No	No	5–OH and 4=N	1 : 1	[2]
2	Flavanone	Hesperetin derivatives	Cu(II)			–OH	–OCH	No	No	5–OH and 4=N	1 : 1	[2]
3	Flavanone	Hesperetin derivatives	Ni(II)			–OH	–OCH	No	No	5–OH and 4=N	1 : 1	[2]
4	Flavanone	Hesperetin derivatives	Zn(II)			–OH	–OCH	No	No	5–OH and 4=N	1 : 1	[3]
5	Flavanone	Hesperetin derivatives	Cu(II)			–OH	–OCH	No	No	5–OH and 4=N	1 : 1	[3]
6	Flavanone	Naringenin derivatives	Zn(II)			No	–OH	No	No	5–OH and 4=N	1 : 1	[10]
7	Flavanone	Naringenin derivatives	Cu(II)			No	–OH	No	No	5–OH and 4=N	1 : 1	[10]
8	Flavanone	Naringenin derivatives	Ni(II)			No	–OH	No	No	5–OH and 4=N	1 : 1	[10]
9	Flavanone	Naringenin Schiff-base	Cu(II)			No	–OH	No	No	5–OH and 4=N	1 : 2	[4]
10	Flavanone	Naringenin Schiff-base	Ni(II)			No	–OH	No	No	5–OH and 4=N	1 : 2	[4]
11	Flavanone	Naringenin Schiff-base	Zn(II)			No	–OH	No	No	5–OH and 4=N	1 : 2	[4]

12	Flavonol	Quercetin	Ge(II)		–OH	–OH	–OH	No	–OH	3–OH and 4=O	1 : 1	[6]
13	Flavanone	Naringenin	Zn(II)		O	No	–OH	No	No	5–OH and 4=O	1 : 2	[8]
14	Flavanone	Naringenin	Cu(II)		O	No	–OH	No	No	5–OH and 4=O	1 : 2	[8]
15	Flavanone	Naringenin	Ni(II)		O	No	–OH	No	No	5–OH and 4=O	1 : 2	[8]
16	Flavonol	Quercetin	Zn(II)		–OH	–OH	–OH	No	–OH	5–OH and 4=O	1 : 2	[5]
17	Flavonol	Quercetin	Cu(II)		–OH	–OH	–OH	No	–OH	5–OH and 4=O, 3′–OH and 4′–OH	2 : 1	[11]
18	Flavonol	Quercetin	Ni(II)		–OH	–OH	–OH	No	–OH	5–OH and 4=O, 3′–OH and 4′–OH	2 : 1	[11]
19	Flavonol	Quercetin	Al(III)		–OH	–OH	–OH	No	–OH	3–OH and 4=O	1 : 3	[11]
20	Flavonol	Quercetin	Cu(II)		–OH	–OH	–OH	No	–OH	3–OH and 4=O, 3′–OH and 4′–OH	2 : 1	[11]
21	Flavanone	Naringenin derivatives	Zn(II)			No	–OH	No	No	5–OH and 4=O	1 : 1	[10]
22	Flavanone	Naringenin derivatives	Cu(II)			No	–OH	No	No	5–OH and 4=O	1 : 1	[10]

(*Continued*)

Table 1. (*Continued*).

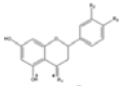
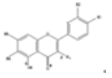
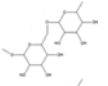
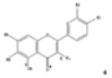
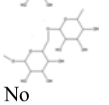
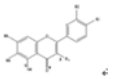
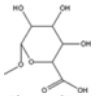
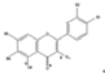
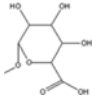
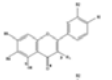
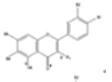
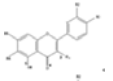
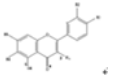
No.	Kind	Ligand	Metal	Basis structure	*R*1	*R*2	*R*3	*R*4	*R*5	Chelation	Ratio (metal/ligand)	Reference
23	Flavanone	Naringenin derivatives	Ni(II)			No	–OH	No	No	5–OH and 4=O	1 : 1	[10]
24	Flavonol	Rutin	Cu(II)			–OH	–OH	No	–OH	5–OH and 4=O	1 : 1	[12]
25	Flavonol	Rutin	Ge(II)			–OH	–OH	No	–OH	5–OH and 4=O	2 : 1	[9]
26	Flavonol	Baicalin	Cu(II)		No	No	No	–OH		5–OH and 4=O	1 : 2	[13]
27	Flavonol	Baicalin	Zn(II)		No	No	No	–OH		5–OH and 4=O	1 : 2	[13]
28	Flavonol	Quercetin	Pb(II)		–OH	–OH	–OH	No	–OH	3–OH and 4=O	1 : 2	[7]
29	Flavonol	Quercetin	Zn(II)		–OH	–OH	–OH	No	–OH	3–OH and 4=O	1 : 2	[7]
30	Flavonol	Quercetin	Ni(II)		–OH	–OH	–OH	No	–OH	3–OH and 4=O	1 : 2	[7]
31	Flavonol	Quercetin	Cu(II)		–OH	–OH	–OH	No	–OH	3–OH and 4=O	1 : 2	[7]

Table 2. Input parameters for ANN.

No.	IC_{50} (uM)	Q_{c7}	Dp	Q_{O11}	Q_a	$E_{L\text{-}1}$
1	1.930	0.35382	5.8508	−0.53593	0.13505	−1.199
2	0.420	0.35751	4.2776	−0.44743	0.14501	−1.991
3	4.460	0.35807	4.1803	−0.36992	0.14284	−1.9801
4	4.098	0.3484	10.3716	−0.47864	0.12756	−1.3257
5	0.401	0.34958	11.2834	−0.45308	0.13373	−1.4791
6	0.250	0.34236	2.7055	−0.52868	0.10693	−0.5135
7	1.190	0.34404	2.0445	−0.43033	0.10907	−0.5228
8	2.750	0.35099	12.5485	−0.38638	0.13843	−1.7775
9	1.568	0.34657	9.3646	−0.62764	0.10597	−1.3309
10	**4.820**	**0.34578**	**9.3194**	**−0.55392**	**0.10497**	**−1.3284**
11	9.063	0.34656	8.8007	−0.69055	0.10583	−1.2705
12	202.400	0.32507	5.494	−0.89525	0.07424	−0.352
13	**40.100**	**0.35693**	**3.6606**	**−0.74368**	**0.09166**	**−1.5539**
14	3.000	0.36199	5.7892	−0.69571	0.10805	−1.5305
15	6.160	0.36775	7.5578	−0.64124	0.11615	−2.2701
16	**83.500**	**0.34703**	**5.012**	**−0.72715**	**0.09313**	**−1.8724**
17	**48.660**	**0.3523**	**3.9059**	**−0.75457**	**0.12788**	**−0.6397**
18	**37.180**	**0.34431**	**6.3688**	**−0.75465**	**0.10797**	**−1.0382**
19	10.210	0.34843	6.157	−0.72594	0.14977	−1.9241
20	2.330	0.35069	2.1213	−0.75329	0.13131	−0.556
21	20.000	0.35429	2.7959	−0.5667	0.1334	−1.3641
22	1.500	0.35756	3.6137	−0.44718	0.14556	−1.4606
23	**29.000**	**0.35816**	**3.5745**	**−0.36985**	**0.14335**	**−1.9725**
24	11.000	0.34409	12.9896	−0.71046	0.10546	−0.1926
25	64.000	0.32526	12.4273	−0.89717	0.10137	−0.0753
26	100.000	0.29536	4.5203	−0.71061	0.64257	−2.4521
27	380.000	0.28595	4.5203	−0.74487	0.61348	−2.3302
28	70.600	0.35012	5.7394	−0.73706	0.12968	−2.1306
29	45.500	0.34775	3.1435	−0.77741	0.12263	−2.0729
30	54.400	0.35228	1.2753	−0.68466	0.15686	−2.4306
31	54.200	0.34961	3.1829	−0.74349	0.13429	−2.2062

Note: Unbolded data are training set and the bolded data are test set.

and IC_{50} is the goal matrix. ANN analysis contains three network layers: the first layer named the input layer includes five neurons, the second layer named the hidden layer includes 15 neurons, and the third layer named the output layer includes one neuron.

F is the input parameter used to build the vector matrix of 25 × 6 and the IC_{50} is the goal matrix; analysis contains three network layers: the first layer named the input layer includes six neurons, the second layer named the hidden layer includes 12 neurons, and the third layer named the output layer includes one neuron.

To understand the reliability and predictability of models, we calculate the standard error (S) between the IC_{50} from literature (named experimental results later) and the IC_{50} from our quantitative structure–activity relationship (QSAR) model calculation (named predicted results later) and the correlation coefficient (R^2) of predicted results and experimental results in the test set.

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N - 1}}$$

Using experimental results as X-axis and predicted results as Y-axis in excel, then complete linear analysis to obtain the R^2 automatically using excel; according to the R^2, we can understand the correlation and the similarity between the experimental results and predicted results, and then understand the reliability and predictability of models.

Table 3. Input parameters for ANN.

No.	IC_{50} (uM)	*F*1	*F*2	*F*3	*F*4	*F*5	*F*6
1	1.93000	−0.60509	−0.45395	0.39704	0.51516	0.32834	0.94034
2	0.42000	−0.52050	−1.06037	0.08071	0.45446	−0.60615	−0.95680
3	4.46000	−0.54375	−0.88491	−1.13323	0.43843	0.36771	−1.36224
4	4.09800	−0.65229	−0.37606	−0.20488	0.40438	1.29695	0.74014
5	0.40100	−0.61671	−0.63537	−0.38807	0.36061	1.36783	−0.14840
6	0.25000	−0.67615	−0.43137	1.49328	−0.27156	0.01595	0.69510
7	1.19000	−0.74432	−0.78989	1.36164	−0.28042	0.01396	−0.47480
8	2.75000	−0.92057	−0.31409	−1.45546	−0.39897	1.86197	−0.65948
9	1.56800	−0.94799	0.37887	0.21260	−0.31237	0.64111	0.44767
10	**4.82000**	**−0.90067**	**0.37763**	**−0.36787**	**−0.41943**	**1.21954**	**−0.42748**
11	9.06300	−0.87203	0.47067	0.38969	−0.25266	0.50853	1.56217
12	202.40000	1.83948	0.01798	1.50090	0.22164	1.66650	−0.31127
13	**40.10000**	**−0.97286**	**0.16349**	**0.78015**	**0.14594**	**−1.82282**	**1.59017**
14	3.00000	−0.87390	−0.12527	0.52014	0.17559	−1.42567	0.38298
15	6.16000	−1.07763	0.28112	−1.05893	0.01472	−0.80571	0.08281
16	**83.50000**	**0.55606**	**0.48522**	**−0.13582**	**0.72593**	**−0.96848**	**1.41255**
17	**48.66000**	**0.45756**	**0.61177**	**0.83264**	**0.71906**	**−1.08574**	**−1.11351**
18	**37.18000**	**−1.17953**	**3.98749**	**−0.27479**	**−0.58387**	**−0.42467**	**−1.37213**
19	10.21000	0.96722	1.29942	−2.27693	0.84516	0.74717	1.90333
20	2.33000	1.07950	−0.07000	0.78687	0.82694	−0.85974	−2.00743
21	20.00000	−0.69896	−0.68895	0.32131	0.02961	−0.12694	1.19773
22	1.50000	−0.75069	−1.32607	0.78996	0.00613	−0.47190	−0.37520
23	**29.00000**	**−0.66371**	**−1.23575**	**−0.87543**	**−0.03527**	**0.15401**	**−1.51488**
24	11.00000	0.21931	1.70116	1.04238	0.03996	0.19284	−1.13538
25	64.00000	1.79642	0.45378	1.85131	0.13380	2.14376	0.44863
26	100.00000	1.14332	−0.59510	−0.45086	−3.34450	−0.64522	−0.11070
27	380.00000	1.19462	−0.17883	−0.45356	−3.41254	−0.52934	0.41579
28	70.60000	1.25542	−0.18661	−0.63149	0.82479	−0.47963	0.16012
29	45.50000	1.28854	−0.10503	−0.30151	0.88587	−0.90866	0.58601
30	54.40000	1.21757	−0.45932	−1.98480	0.77228	−0.32484	−0.58610
31	54.20000	1.20236	−0.31165	−0.36698	0.77111	−1.04067	−0.00974

Note: Unbolded data are training set and the bolded data are test set.

3. Results and discussion

3.1. *Electronic structure of flavonoid–metal complexes*

We calculated the electronic structure properties of flavonoid–metal complexes and obtained 21 quantum chemistry descriptors of flavonoid–metal complexes including: IC_{50} (from literature), dipole, net charges of the A ring (Q_a), the net charge of B ring (Q_b), the net charge of C ring (Q_c), the net charge of metal ion (Q_{metal}), the net charge of O at site 1 (Q_{O1}), the net charge of C at site 4 (Q_{c4}), the net charge of C at site 2 (Q_{c2}), the net charge of C at site 5 (Q_{c5}), the net charge of O or N at site 11 ($Q_{O/N11}$), the net charge of O at site 12 (Q_{O12}), the net charge of C at site 7 (Q_{c7}), the three highest occupied molecular orbital energies (E_{HOMO}, $E_{H\text{-}1}$, $E_{H\text{-}2}$), the three lowest unoccupied molecular orbital energies (E_{LUMO}, $E_{L\text{-}1}$, $E_{L\text{-}2}$), and the three LUMO–HOMO gaps ($\Delta E_{L\text{-H2}}$, $\Delta E_{L\text{-H1}}$, $\Delta E_{L\text{-}H}$), which are shown in table 4. On the basis of the frontier molecular orbital theory, the antioxidant activity of the complex relies on the following two factors:

- The LUMO (the lowest unoccupied molecular orbital) energy of the complex because a lower LUMO energy of the complex is favorable for accepting electrons.
- The planarity area of the intercalative ligand because the larger planarity area is favorable to activity.

Table 4. The structure parameters and IC_{50} of flavonoid–metal complexes.

No.	IC_{50}(uM)	Dp	Q_a	Q_b	Q_c	Q_{metal}	Q_{O1}	Q_{c4}	Q_{c2}	Q_{c5}	Q_{O11}	Q_{O12}
1	1.93	5.8508	0.13505	−0.35778	−0.57904	1.65705	−0.53996	0.32257	0.07619	0.41900	−0.53593	−0.89183
2	0.42	4.2776	0.14501	−0.34482	−0.55425	1.21170	−0.53142	0.32841	0.07245	0.41130	−0.44743	−0.77171
3	4.46	4.1803	0.14284	−0.34447	−0.55381	1.00466	−0.53049	0.32707	0.07152	0.40856	−0.36992	−0.70856
4	4.098	10.3716	0.12756	−0.3619	−0.62688	1.64857	−0.54437	0.28928	0.07888	0.4279	−0.47864	−0.85411
5	0.401	11.2834	0.13373	−0.36185	−0.61956	1.30335	−0.54081	0.28859	0.08012	0.42297	−0.45308	−0.73923
6	0.25	2.7055	0.10693	−0.78901	−0.60062	1.64094	−0.53985	0.29428	0.07536	0.40581	−0.52868	−0.92492
7	1.19	2.0445	0.10907	−0.78786	−0.59478	1.25454	−0.53905	0.29462	0.07415	0.40347	−0.43033	−0.81288
8	2.75	12.5485	0.13843	−0.79415	−0.61351	1.03710	−0.53718	0.28259	0.08047	0.41249	−0.38638	−0.68270
9	1.568	9.3646	0.10597	−0.79364	−0.5866	1.27226	−0.53915	0.31664	0.07777	0.42113	−0.62764	−0.74826
10	4.82	9.3194	0.10497	−0.79342	−0.58937	0.96632	−0.53802	0.31001	0.07781	0.41816	−0.55392	−0.67235
11	9.063	8.8007	0.10583	−0.79297	−0.59383	1.66425	−0.54310	0.31884	0.07581	0.41960	−0.69055	−0.86309
12	202.4	5.4940	0.07424	−0.41836	0.29666	1.33260	−0.49614	0.23916	0.26015	0.34594	−0.89525	−0.71105
13	40.1	3.6606	0.09166	−0.79033	−0.42094	1.69804	−0.53704	0.53209	0.07163	0.43234	−0.74368	−0.86207
14	3	5.7892	0.10805	−0.78738	−0.40493	1.39537	−0.53310	0.53406	0.07111	0.42680	−0.69571	−0.77521
15	6.16	7.5578	0.11615	−0.78277	−0.40361	1.07377	−0.52942	0.52512	0.07082	0.42191	−0.64124	−0.73796
16	83.5	5.0120	0.09313	−0.36729	0.55851	1.70450	−0.47515	0.44539	0.28659	0.41200	−0.72715	−0.86437
17	48.66	3.9059	0.12788	−0.15027	0.59837	0.77123	−0.46995	0.41902	0.31345	0.42463	−0.75457	−0.82049
18	37.18	6.3688	0.10797	−0.3305	0.56648	0.72470	−0.47232	0.40755	−0.24223	0.41972	−0.75465	−0.80879
19	10.21	6.1570	0.14977	−0.35725	0.6297	2.07425	−0.46643	0.41825	0.30662	0.38305	−0.72594	−0.73940
20	2.33	2.1213	0.13131	−0.17247	0.62481	0.74972	−0.46280	0.41912	0.28912	0.38393	−0.75329	−0.69011
21	20	2.7959	0.1334	−0.77903	−0.56353	1.64522	−0.53215	0.33082	0.07375	0.41032	−0.56670	−0.90102
22	1.5	3.6137	0.14556	−0.7777	−0.55298	1.21174	−0.53036	0.32849	0.07225	0.41137	−0.44718	−0.77182
23	29	3.5745	0.14335	−0.77718	−0.55257	1.00478	−0.52938	0.32720	0.07125	0.40865	−0.36985	−0.70861
24	11	12.9896	0.10546	−0.34076	0.56648	0.78912	−0.47991	0.42565	0.34185	0.42409	−0.71046	−0.82298
25	64	12.4273	0.10137	−0.09592	0.39507	1.34080	−0.50278	0.28084	0.29150	0.34842	−0.89717	−0.90574
26	100	4.5203	0.64257	−1.2007	0.15118	1.38379	−0.46428	0.47601	0.39246	0.34176	−0.71061	−0.82443
27	380	4.5203	0.61348	−1.20442	0.12209	1.71125	−0.46802	0.47266	0.38707	0.34418	−0.74487	−0.88487
28	70.6	5.7394	0.12968	−0.34195	0.60553	1.49222	−0.46919	0.41137	0.31707	0.38307	−0.73706	−0.69007
29	45.5	3.1435	0.12263	−0.35345	0.5846	1.68666	−0.47115	0.40970	0.30606	0.38379	−0.77741	−0.69318
30	54.4	1.2753	0.15686	−0.32876	0.62873	1.06796	−0.46518	0.41208	0.33362	0.38901	−0.68466	−0.68336
31	54.2	3.1829	0.13429	−0.34508	0.60065	1.38445	−0.46910	0.41231	0.31649	0.38526	−0.74349	−0.69058

(*Continued*)

Table 4. (*Continued*).

No.	IC_{50} (uM)	Q_{c7}	$E_{H\text{-}2}$	$E_{H\text{-}1}$	E_{HOMO}	E_{LUMO}	$E_{L\text{-}1}$	$E_{L\text{-}2}$	$\Delta E_{L\text{-}H2}$	$\Delta E_{L\text{-}H1}$	$\Delta E_{L\text{-}H}$
1	1.93	0.35382	−5.9481	−5.9106	−5.5308	−1.9464	−1.1990	−0.6003	5.3478	4.7116	3.5844
2	0.42	0.35751	−6.1105	−5.8831	−5.8208	−2.9422	−1.991	−0.2821	5.8284	3.8921	2.8786
3	4.46	0.35807	−6.1241	−5.8893	−5.7441	−2.9422	−1.9801	−1.4865	4.6376	3.9092	2.8019
4	4.098	0.3484	−5.9924	−5.9677	−5.4620	−2.2070	−1.3257	−0.8089	5.1835	4.642	3.255
5	0.401	0.34958	−6.0588	−5.9897	−5.5948	−2.2859	−1.4791	−0.8378	5.221	4.5106	3.3089
6	0.25	0.34236	−6.0414	−5.5684	−5.2376	−1.4582	−0.5135	−0.2105	5.8309	5.0549	3.7784
7	1.19	0.34404	−6.0561	−5.6271	−5.3535	−1.5567	−0.5228	−0.3283	5.7278	5.1043	3.7968
8	2.75	0.35099	−6.145	−5.7528	−5.7076	−2.4891	−1.7775	−1.5961	4.5489	3.9753	3.2185
9	1.568	0.34657	−5.7479	−5.4065	−5.2578	−1.4495	−1.3309	−0.4281	5.3198	4.0756	3.8083
10	4.82	0.34578	−5.848	−5.3641	−4.9958	−1.3992	−1.3284	−1.0573	4.7907	4.0357	3.5966
11	9.063	0.34656	−5.8676	−5.3802	−5.2888	−1.4261	−1.2705	−0.3979	5.4697	4.1097	3.8627
12	202.4	0.32507	−5.8602	−5.3758	−3.4016	−0.3958	−0.352	−0.2214	5.6388	5.0238	3.00581
13	40.1	0.35693	−6.1358	−5.6116	−5.5050	−1.6233	−1.5539	−0.3166	5.8192	4.0577	3.8817
14	3	0.36199	−6.0797	−5.882	−5.4065	−1.9715	−1.5305	−0.4292	5.6505	4.3515	3.435
15	6.16	0.36775	−6.0098	−5.7688	−5.644	−2.3215	−2.2701	−1.3597	4.6501	3.4987	3.3225
16	83.5	0.34703	−5.7922	−5.3935	−5.2493	−1.9056	−1.8724	−0.5875	5.2047	3.5211	3.3437
17	48.66	0.35230	−5.486	−5.0976	−4.8636	−1.8763	−0.6397	−0.4855	5.0005	4.4579	2.9873
18	37.18	0.34431	−4.0283	−3.9331	−3.5833	−1.3469	−1.0382	−0.3204	3.7079	2.8949	2.2364
19	10.21	0.34843	−5.1756	−5.1468	−5.2286	−2.1869	−1.9241	−1.8254	3.3502	3.2227	3.0417
20	2.33	0.35069	−5.4541	−5.4036	−4.8462	−2.127	−0.556	−0.4885	4.9656	4.8476	2.7192
21	20	0.35429	−6.275	−5.9405	−5.7863	−1.827	−1.3641	−0.9651	5.3099	4.5764	3.9593
22	1.5	0.35756	−6.3496	−6.0991	−5.8148	−1.9829	−1.4606	−0.3906	5.959	4.6385	3.8319
23	29	0.35816	−6.3857	−6.1127	−5.7365	−2.9338	−1.9725	−1.4797	4.906	4.1402	2.8027
24	11	0.34409	−5.0127	−4.7954	−4.0232	−1.437	−0.1926	0.0699	5.0826	4.6028	2.5862
25	64	0.32526	−5.9364	−5.3682	−3.3064	−0.3171	−0.0753	−0.0612	5.8752	5.2929	2.9893
26	100	0.29536	−6.6950	−5.7699	−5.6622	−2.2366	−2.4521	−1.1769	5.5181	3.3178	3.4256
27	380	0.28595	−6.5525	−5.5823	−5.4454	−2.3631	−2.3302	−1.0673	5.4852	3.2521	3.0823
28	70.6	0.35012	−6.0599	−5.4664	−5.187	−2.3264	−2.1306	−0.7442	5.3157	3.3358	2.8606
29	45.5	0.34775	−6.0591	−5.2632	−5.1294	−2.1066	−2.0729	−0.5084	5.5507	3.1903	3.0288
30	54.4	0.35228	−6.395	−5.6864	−5.4291	−2.6811	−2.4306	−2.3264	4.0686	2.9985	2.748
31	54.2	0.34961	−6.1638	−5.3579	−5.276	−2.2519	−2.2062	−0.6389	5.5249	3.1517	3.0241

According to the calculation results, the A ring and C ring in these complexes have excellent planarity, and the dihedral angles between the B ring and the above plane are also very small, resulting in a large conjugated system. The reaction often occurs on the sites that have larger electronic densities when the drug interacts with receptor molecules. The active sites of the drug are those atoms that have larger contributions, which are shown in table 4. The charge density distributions of these atoms have important influence on their activities. The results show that O_1, O_{11}, and O_{12} have larger negative charges; the negative charges of O_1 are from −0.46280 to −0.54437, the negative charges of O_{11} are from −0.36992 to −0.89717, the negative charges of O_{12} are from −0.67235 to −0.90574, while C_7, C_4, and C_5 and metal ions have larger positive charges. The positive charges of C_4 are from 0.23916 to 0.53406, the positive charges of C_5 are from 0.34176 to 0.43234, the positive charges of C_7 are from 0.28595 to 0.36775, and the positive charges of metal ion are from 0.74972 to 2.07425. $\Delta E_{L\text{-}H}$ are from 2.2364 to 3.8817, $\Delta E_{L\text{-}H1}$ are from 2.9985 to 5.1043, and $\Delta E_{L\text{-}H2}$ are from 3.3502 to 5.8309.

3.2. *The results of stepwise linear regression and principal component analysis*

By studying the relationship between the 21 quantum chemistry descriptors and IC_{50} of flavonoid–metal complexes for scavenging $O_2^{\cdot-}$ through stepwise linear regression, we obtain important quantum chemistry descriptors as dependent variables (shown in table 5), the correlation coefficient (shown in table 6), and the regression coefficient (shown in table 7).

The results of principal component analysis for table 2 are shown in tables 8 and 9. The total variance explained through principal component analysis is in table 6. We chose the top six components as principal components for the ANN analysis because the percentage of cumulative variance of the top six components is 86.939%, which largely represents the original variances. According to the principal component analysis, the principal component coefficient matrix is used to calculate F values, as shown in table 9.

According to the analysis results, only five quantum chemistry descriptors are important to the IC_{50} of flavonoid–metal complexes scavenging $O_2^{\cdot-}$. They are the Dp, the net charge of C at site 7 (Q_{c7}), net charge of C at site 11 (Q_{O11}), net charge of the A ring (Q_a), and the lower unoccupied molecular orbital energy $E_{L\text{-}1}$. The predictable equation is given below:

$$IC_{50} = 1626.087 - 4686.216Q_{C7} - 3.981D_p - 69.941Q_{O11} - 35.257E_{L-1} - 232.15Q_a$$

Table 5. Variables entered/removed.

Model	Variables entered	Variables removed	Method
1	Q_{c7}	.	Stepwise (criteria: probability-of-F-to-enter ≤ 0.200, probability-of-F-to-remove ≥ 0.300)
2	Dp	.	Stepwise (criteria: probability-of-F-to-enter ≤ 0.200, probability-of-F-to-remove ≥ 0.300)
3	Q_{O11}	.	Stepwise (criteria: probability-of-F-to-enter ≤ 0.200, probability-of-F-to-remove ≥ 0.300)
4	$E_{L\text{-}1}$	.	Stepwise (criteria: probability-of-F-to-enter ≤ 0.200, probability-of-F-to-remove ≥ 0.300)
5	Q_a	.	Stepwise (criteria: probability-of-F-to-enter ≤ 0.200, probability-of-F-to-remove ≥ 0.300)

Note: Dependent variable: IC_{50}.

Table 6. Model summary.

Model	R	R^2	Adjusted R^2	Std. error of the estimate
1	0.787	0.620	0.607	47.540364
2	0.810	0.657	0.632	45.969215
3	0.826	0.682	0.647	45.047442
4	0.842	0.710	0.665	43.871861
5	0.854	0.730	0.676	43.171956

Table 7. Coefficients.

Model		Unstandardized coefficients		Standardized coefficients		
		B	Std. error	Beta	t	Sig.
1	(constant)	1247.478	175.598		7.104	0.000
	Q_{c7}	−3487.125	507.257	−0.787	−6.874	0.000
2	(constant)	1288.567	171.435		7.516	0.000
	Q_{c7}	−3530.905	491.140	−0.797	−7.189	0.000
	Dp	−4.358	2.509	−0.193	−1.737	0.093
3	(constant)	1130.485	199.513		5.666	0.000
	Q_{c7}	−3234.499	521.882	−0.730	−6.198	0.000
	Dp	−4.205	2.461	−0.186	−1.709	0.099
	Q_{O11}	−86.567	58.934	−0.173	−1.469	0.153
4	(constant)	1056.180	199.984		5.281	0.000
	Q_{c7}	−3155.432	510.750	−0712	−6.178	0.000
	Dp	−3.104	2.497	−0.137	−1.243	0.225
	Q_{O11}	−106.092	58.727	−0.212	−1.807	0.082
	$E_{L\text{-}1}$	−19.280	12.277	−0.176	−1.570	0.128
5	(constant)	1626.087	462.931		3.513	0.002
	Q_{c7}	−4686.216	1232.621	−1.058	−3.802	0.001
	Dp	−3.981	2.541	−0.176	−1.567	0.130
	Q_{O11}	−69.941	63.610	−0.140	−1.100	0.282
	$E_{L\text{-}1}$	−35.257	16.850	−0.322	−2.092	0.047
	Q_a	−232.150	170.686	−0.392	−1.360	0.186

According to the above analysis, the principal component equations are:

$$F1 = -0.038\mathrm{D_p} - 0.011Q_a + 0.182Q_b + 0.182Q_c + 0.114Q_{metal} + 0.157Q_{O1} - 0.088Q_{C4} + 0.285Q_{C2} - 0.244Q_{C5} - 0.136Q_{O/N11} + 0.09Q_{O12} - 0.07Q_{C7} - 0.122E_{H-2} - 0.071E_{H-1} + 0.075E_{\mathrm{HOMO}} + 0.027E_{\mathrm{LUMO}} + 0E_{L-1} - 0.045E_{L-2} + 0.057\Delta E_{L-H2} + 0.039\Delta E_{L-H1} + 0.079\Delta E_{L-H}$$

$$F2 = 0.153\mathrm{D_p} + 0.012Q_a - 0.067Q_b + 0.04Q_c + 0.018Q_{metal} + 0.03Q_{O1} + 0.123Q_{C4} - 0.152Q_{C2} + 0.11Q_{C5} - 0.106Q_{O/N11} - 0.115Q_{O12} - 0.043Q_{C7} + 0.276E_{H-2} + 0.28E_{H-1} + 0.12E_{\mathrm{HOMO}} + 0.117E_{\mathrm{LUMO}} + 0.015E_{L-1} + 0.019E_{L-2} - 0.204\Delta E_{L-H2} - 0.155\Delta E_{L-H1} - 0.02\Delta E_{L-H}$$

$$F3 = -0.156\mathrm{D_p} - 0.008Q_a - 0.011Q_b - 0.004Q_c - 0.142Q_{metal} - 0.008Q_{O1} + 0.04Q_{C4} + 0.001Q_{C2} + 0.004Q_{C5} - 0.047Q_{O/N11} - 0.116Q_{O12} - 0.019Q_{C7} - 0.02E_{H-2} - 0E_{H-1} + 0.074E_{\mathrm{HOMO}} + 0.112E_{\mathrm{LUMO}} + 0.206E_{L-1} + 0.302E_{L-2} + 0.288\Delta E_{L-H2} + 0.213\Delta E_{L-H1} + 0.043\Delta E_{L-H}$$

Table 8. The cumulative percentage of component.

	Initial eigenvalues			Extraction sums of squared			Rotation sums of squared		
Component	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	5.678	27.039	27.039	5.678	27.039	27.039	4.211	20.051	20.051
2	4.842	23.057	50.096	4.842	23.057	50.096	3.944	18.780	38.831
3	3.640	17.332	67.427	3.640	17.332	67.427	3.844	18.303	57.134
4	1.616	7.697	75.124	1.616	7.697	75.124	2.517	11.986	69.120
5	1.477	7.032	82.156	1.477	7.032	82.156	1.927	9.174	78.294
6	1.004	4.783	86.939	1.004	4.783	86.939	1.816	8.646	86.939
7	0.751	3.574	90.513						
8	0.710	3.380	93.894						
9	0.538	2.560	96.453						
10	0.310	1.476	97.930						
11	0.181	0.862	98.792						
12	0.101	0.480	99.271						
13	0.057	0.273	99.544						
14	0.035	0.166	99.710						
15	0.033	0.159	99.869						
16	0.018	0.087	99.956						
17	0.008	0.039	99.995						
18	0.001	0.004	99.999						
19	0	0.001	100.000						
20	6.366E−7	3.032E−6	100.000						
21	7.960E−17	3.790E−16	100.000						

Note: Extraction method: principal component analysis.

Table 9. The component score coefficient matrix.

	Component					
	1	2	3	4	5	6
Dipole	−0.038	0.153	−0.156	−0.079	0.436	0.152
Q_a	−0.011	0.012	−0.008	−0.387	−0.031	−0.147
Q_b	0.182	−0.067	−0.011	0.379	0.074	0.032
Q_c	0.182	0.040	−0.004	0.116	−0.091	0.015
Q_{metal}	0.114	0.018	−0.142	0.184	0.130	0.648
Q_{O1}	0.157	0.030	−0.008	0.011	−0.123	−0.062
Q_{c4}	−0.088	0.123	0.040	−0.062	−0.457	0.033
Q_{c2}	0.285	−0.152	0.001	0.107	0.032	0.062
Q_{c5}	−0.244	0.110	0.004	0.094	−0.179	0.035
Q_{O11}	−0.136	−0.106	−0.047	−0.062	0.089	−0.198
Q_{O12}	0.090	−0.115	−0.116	0.125	0.089	−0.215
Q_{c7}	−0.070	−0.043	−0.019	0.349	−0.149	0.057
$E_{H\text{-}2}$	−0.122	0.276	−0.020	−0.009	−0.016	0
$E_{H\text{-}1}$	−0.071	0.280	0	−0.089	−0.057	0.002
E_H	0.075	0.120	0.074	−0.023	0.138	−0.052
E_L	0.027	0.117	0.112	−0.023	0.134	0.158
$E_{L\text{-}1}$	0	0.015	0.206	0.009	0.058	−0.125
$E_{L\text{-}2}$	−0.045	0.019	0.302	−0.008	−0.242	−0.084
$\Delta E_{L\text{-}H2}$	0.057	−0.204	0.288	0	−0.204	−0.076
$\Delta E_{L\text{-}H1}$	0.039	−0.155	0.213	0.058	0.091	−0.126
$\Delta E_{L\text{-}H}$	−0.079	−0.020	0.043	0.005	−0.025	0.303

$$F4 = -0.079\mathrm{D_p} - 0.387Q_a + 0.379Q_b + 0.116Q_c + 0.184Q_{metal} + 0.011Q_{O1} - 0.062Q_{C4} + 0.107Q_{C2} + 0.094Q_{C5} - 0.062Q_{O/N11} + 0.125Q_{O12} + 0.349Q_{C7} - 0.009E_{H-2} - 0.089E_{H-1} - 0.023E_{\mathrm{HOMO}} - 0.023E_{\mathrm{LUMO}} + 0.009E_{L-1} - 0.008E_{L-2} + 0\Delta E_{L-H2} + 0.058\Delta E_{L-H1} + 0.005\Delta E_{L-H}$$

$$F5 = 0.436\mathrm{D_p} - 0.031Q_a + 0.074Q_b - 0.091Q_c + 0.13Q_{metal} - 0.123Q_{O1} - 0.457Q_{C4} + 0.032Q_{C2} - 0.179Q_{C5} + 0.089Q_{O/N11} + 0.089Q_{O12} - 0.149Q_{C7} - 0.016E_{H-2} - 0.057E_{H-1} + 0.138E_{\mathrm{HOMO}} + 0.134E_{\mathrm{LUMO}} + 0.058E_{L-1} - 0.242E_{L-2} - 0.204\Delta E_{L-H2} + 0.091\Delta E_{L-H1} - 0.025\Delta E_{L-H}$$

$$F6 = 0.152\mathrm{D_p} - 0.147Q_a + 0.032Q_b + 0.015Q_c + 0.648Q_{metal} - 0.062Q_{O1} + 0.033Q_{C4} + 0.062Q_{C2} + 0.035Q_{C5} - 0.198Q_{O/N11} - 0.215Q_{O12} + 0.057Q_{C7} - 0E_{H-2} + 0.002E_{H-1} - 0.052E_{\mathrm{HOMO}} + 0.158E_{\mathrm{LUMO}} - 0.125E_{L-1} - 0.084E_{L-2} - 0.076\Delta E_{L-H2} - 0.126\Delta E_{L-H1} + 0.303\Delta E_{L-H}$$

3.3. *The results of the ANN*

Based on the results of stepwise linear regression, we choose Q_{c7}, Q_{O11}, Dp, Q_a, and $E_{L\text{-}1}$ as the independent variables and the IC_{50} as the dependent variable to analyze for the ANN. The data are shown in table 2.

On the basis of the results of principal component analysis, we choose F as the independent variables and the IC_{50} as the dependent variable to analyze for the ANN, as shown in table 3.

We obtained a QSAR model with the appointed convergence accuracy based on Q_{c7}, Q_{O11}, Dp, Q_a, $E_{L\text{-}1}$, and IC_{50} after 130 training times, table 10. Calculation results based on table 2 using ANN show the predicted results of our QSAR model. In tables 2 and 10, the 25 data groups (unbolded values) were used as the training set and the other six data groups (bolded values) were the test set. The six data groups (bolded values) were used to check the predictability of our model. For ensuring the applicability of the model, we selected the training sets and the tested sets randomly. The standard error between experimental results and predicted results was 2.262. Figure 2 shows a comparison between the experimental results and predicted results. Figure 2 shows that the calculated values agree well with the experimental results and the residual is very small. The R^2 of predicted and experimental results in test set (bolded values) is 0.958. Our results suggested that our QSAR model is reliable and predictive.

We obtained a QSAR model with the appointed convergence accuracy based on F and IC_{50} after 136 training times (table 11). Calculation results based on table 3 using ANN show the predicted results of our QSAR model. In tables 9 and 11, the 25 data groups (unbolded values) were used as the training set and the other six data groups (bolded values) were the test set used to check the predictability of our model. For ensuring the applicability of the model, we selected the training sets and the tested sets randomly. The standard error between experimental results and predicted results is 1.554. Figure 3 shows a comparison between the experimental results and predicted results. The predicted results agree well with the experimental results, and the residual is very small. The R^2 of predicted

Table 10. Calculation results based on table 2 using ANN.

No.	Initial IC_{50} (uM)	Calculate IC_{50} (uM)	Error
1	1.93000	1.90408	0.02592
2	0.42000	0.43968	−0.01968
3	4.46000	4.45917	0.00083
4	4.09800	3.35406	0.74394
5	0.40100	2.25627	−1.85527
6	0.25000	0.24975	0.00025
7	1.19000	1.18991	0.00009
8	2.75000	1.63781	1.11219
9	1.56800	1.56845	−0.00045
10	**4.82000**	**2.81966**	**2.00034**
11	9.06300	9.06399	−0.00099
12	202.40000	202.39990	0.00010
13	**40.10000**	**37.45999**	**2.64001**
14	3.00000	2.99934	0.00066
15	6.16000	6.16687	−0.00687
16	**83.50000**	**90.40076**	**−6.90076**
17	**48.66000**	**40.0406**	**8.6194**
18	**37.18000**	**39.18781**	**−2.00781**
19	10.21000	10.20382	0.00618
20	2.33000	2.32990	0.00010
21	20.00000	19.99986	0.00014
22	1.50000	1.50033	−0.00033
23	**29.00000**	**32.3751**	**−3.3751**
24	11.00000	10.99987	0.00013
25	64.00000	63.99999	0.00001
26	100.00000	100.00237	−0.00237
27	380.00000	379.99701	0.00299
28	70.60000	70.59698	0.00302
29	45.50000	45.50051	−0.00051
30	54.40000	54.40000	0.00000
31	54.20000	54.19950	0.00050

Note: Unbolded data are training set and the bolded data are test set.

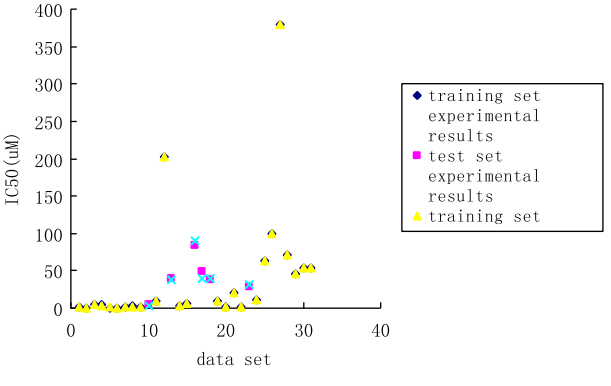


Figure 2. Calculated and experimental results of IC_{50} in training and test sets based on table 10.

and experimental results in the test set (bolded values) is 0.981. Our results suggested that our QSAR model is reliable and predictive.

By comparing the above two results, we conclude that the predictability of the obtained model based on principal component analysis is better.

Table 11. Calculation results based on table 3 using ANN.

No.	Initial IC_{50} (uM)	Calculate IC_{50} (uM)	Error
1	1.93000	1.92547	0.00453
2	0.42000	0.42124	−0.00124
3	4.46000	4.46006	−0.00006
4	4.09800	4.10157	−0.00357
5	0.40100	0.40105	−0.00005
6	0.25000	0.25007	−0.00007
7	1.19000	1.18964	0.00036
8	2.75000	2.74857	0.00143
9	1.56800	1.56841	−0.00041
10	**4.82000**	**3.01943**	**1.80057**
11	9.06300	9.06289	0.00011
12	202.40000	202.40030	−0.00030
13	**40.10000**	**42.3074**	**−2.20740**
14	3.00000	2.99989	0.00011
15	6.16000	6.15980	0.00020
16	**83.50000**	**88.9837**	**−5.48370**
17	**48.66000**	**43.76010**	**4.89990**
18	**37.18000**	**35.18012**	**1.99988**
19	10.21000	10.21011	−0.00011
20	2.33000	2.32901	0.00099
21	20.00000	19.99999	0.00001
22	1.50000	1.50000	0.00000
23	**29.00000**	**31.50201**	**−2.50201**
24	11.00000	10.99985	0.00015
25	64.00000	63.99999	0.00001
26	100.00000	100.00071	−0.00071
27	380.00000	379.99930	0.00070
28	70.60000	70.60032	−0.00032
29	45.50000	45.50014	−0.00014
30	54.40000	54.40725	−0.00725
31	54.20000	54.20007	−0.00007

Note: Unbolded data are training set and the bolded data are test set.

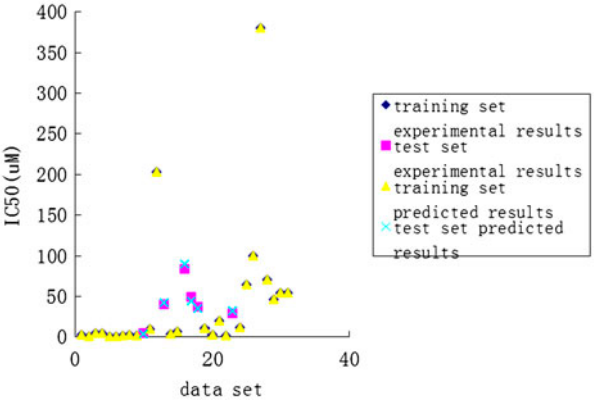


Figure 3. Calculated and experimental results of IC_{50} in training and test sets based on table 11.

4. Conclusion

According to the calculation results of the electronic structures of flavonoid–metal complexes, the A ring, B ring, and C ring in these complexes have excellent planarity, resulting in a large conjugated systems. O_1, O_{11}, and O_{12} have larger negative charges, the negative charges of O_1 are from −0.46280 to −0.54437, the negative charges of O_{11} are from −0.36992 to −0.89717, the negative charges of O_{12} are from −0.67235 to −0.90574, while C_7, C_4, C_5, and metal ion have larger positive charges, the positive charges of C_4 are from 0.23916 to 0.53406, the positive charges of C_5 are from 0.34176 to 0.43234, the positive charges of C_7 are from 0.28595 to 0.36775, and the positive charges of metal ion are from 0.74972 to 2.07425. $\Delta E_{L\text{-}H}$ are from 2.2364 to 3.8817, $\Delta E_{L\text{-}H1}$ are from 2.9985 to 5.1043, and $\Delta E_{L\text{-}H2}$ are from 3.3502 to 5.8309. This result is the main reason for the distribution of charge.

According to the results of stepwise linear regression, the IC_{50} negatively correlates to the net charges of C_7, Q_{11}, the net charges of ring A, the Dp, and the LUMO1–HOMO1 gap ($\Delta E_{L\text{-}H1}$). These five descriptors are most important to the IC_{50} of flavonoid–metal complexes scavenging $O_2^{\cdot -}$. According to the results of principal component analysis, we found that the IC_{50} is also related to calculated quantum chemistry descriptors. So we built two models based on the results of stepwise linear regression and the results of principal component analysis.

Based on the results of stepwise linear regression analysis and principal component analysis, we chose important quantum chemistry descriptors and the linear combination of quantum chemistry descriptors, named F as independent variables and the IC_{50} as the dependent variables to analyze for the ANN. We built two QSAR models and checked them with data. Our results show that the two models, built in this article, have reliable predictability. The predictability of the model based on principal component analysis is better, but the model based on stepwise linear regression analysis is more helpful in designing new drugs. Therefore, we conclude that the two models we built can forecast the activities of flavonoid–metal complexes for designing new drugs.

Funding

This work was financially supported by Project No. CDJXS11232241; the Fundamental Research Funds for the Central Universities; Chongqing University Postgraduates' Innovative Team Building Project, Team Number: 201105A1001; National Natural Science Foundation of China [grant number 20901086]; Natural Science Foundation Project of CQ CSTC [grant number 2011BB5109]; Visiting Scholar Foundation of Key Laboratory of Biorheological Science and Technology (Chongqing University); Ministry of Education [CQKLBST-2012-007]; Universities Innovation Team Development Program of Chongqing.

References

[1] Z.-x. Zhang, K. Liu, S.-y. Wang. *Chin. Traditional Herbal Drugs*, **27**, 179 (1996).
[2] Y. Li, Z.-Y. Yang, M.-F. Wang. *J. Fluoresc.*, **20**, 891 (2010).
[3] Y. Li, Z.-y. Yang. *Inorg. Chim. Acta*, **362**, 4823 (2009).
[4] Y. Li, Z.-y. Yang, T.-r. Li. *Chem. Pharm. Bull.*, **56**, 1528 (2008).
[5] Z. Xinshen. Study on relation of molecular structure bioactivity and cooperative action for flavonoid coordination complexes. Doctor's degree thesis, SiChuan University (2004).
[6] W. Junjie, H. Chunxiang. *J. Sci. Teach. College. Univ.*, **30**, 71 (2010).
[7] J. Zhou, L. Wang, J. Wang, N. Tang. *Transition Met. Chem.*, **26**, 57 (2001).
[8] H.-L. Wang, Z.-Y. Yang, B.-d. Wang. *Transition Met. Chem.*, **31**, 470 (2006).
[9] B. Zhijuan, C. Qiu'e, Z. Yongming. *Nat. Prod. Res. Dev.*, **3**, 263 (2006).
[10] T. Li. Synthesis, charaeterization and properties of aryacylhydrazone ligands and their metal complexes. Doctor's degree thesis, LanZhou University (2010).

[11] L. Jiang. Study on synthesis, activity, and electronic structure of metallic complex of quercetin. Master's degree thesis, SiChuan University (2005).
[12] B. Zhijuan, F. Yunshan, D. Zhongtao, C. Qiu-E. *Chin. J. Spectrosc. Lab.*, **22**, 253 (2005).
[13] F. Yu, H. Daodao, L. Xiaojun. *Chin. Biochem. J.*, **7**, 753 (1991).
[14] F. Ji-qian, L. Ying. *The Introduction of BP ANN Model*, Peoples's Medical Publishing, Beijing (2002).
[15] J.B. Foresman, A.E. Frisch. *Exploring Chemistry with Electronic Structure Methods*, 2nd Edn, Gaussian, Pittsburgh, PA (1996).