UDC 541.6:547.13:546.562

QSAR STUDY OF FLAVONOID—METAL COMPLEXES AND THEIR ANTICANCER ACTIVITIES

J.-Z. Qian[1,2], B.-C. Wang[1], Y. Fan[3], J. Tan[4], X. Yang[5]

[1]*Bioengineering College, Chongqing University, Chongqing, China*
 E-mail: wangbc2000@126.com (B.-C. Wang)
[2]*Laboratory and Equipment Managing Section, Chongqing University, Chongqing, China*
[3]*Chongqing Telecom Planning and Designing Institute Co., Ltd, Chongqing, China*
[4]*Department of Biological and Chemical Engineering, Chongqing University of Education, Chongqing, China*
[5]*Chongqing Normal University, Chongqing, China*

Flavonoid-metal complexes have anticancer activities. However, the quantitative structure-activity relationship (QSAR) of flavonoid—metal complexes and their anticancer activities has not been known so far. Based on the 14 structures of flavonoid—metal complexes and their anticancer activities for HepG2 from the references, we optimised their structures using the density functional theory (DFT) method, and subsequently calculated 19 quantum chemical descriptors, such as dipole, charge, and energy. Then, we chose several quantum chemical descriptors that are very important for IC_{50} which represents the anticancer activities of flavonoid—metal complexes for HepG2 through the stepwise linear regression method. Meanwhile, we obtained six new variables through the principal component analysis. Finally, we built QSAR models based on those important quantum chemical descriptors, six new variables as independent variables, and IC_{50} as a dependent variable using an artificial neural network (ANN). At last, we validated the models using the experimental data from the references. The results show that models presented in this paper are accurate and predictive.

K e y w o r d s: flavonoid—metal complexes, quantum chemistry descriptors, anticancer, artificial neural network (ANN), quantitative structure-activity relationship (QSAR).

INTRODUCTION

Flavonoids (classified as flavonoid, flavonol, flavonone, flavanonol, isoflavone, etc. according to different structural characteristics) can act as metallic chelators due to their peculiar structures [1]. In general, flavonoids chelate with metal ions to form complexes that have new pharmacological activities and enhance the intrinsic pharmacological activities. More attention is focused on flavonoid—metal complexes, especially their synthesis and pharmacological properties [2—8]. Thus far, many studies about the synthesis of flavonoid—metal complexes have been reported. The groups of flavonoids that chelate with metals are 3-OH, 4C=O, 5-OH or —OH on catechol. Current studies about flavonoid—metal complexes mainly focus on their anticancer and antioxidant activities. For example, some studies show that the binding constants of flavonoid—metal complexes (such as morin-Zn, Cu, Co, rutin-Eu, Cu, Se, chrysin-La, quercetin-Zn, Mn, Co, Ni, Cu, Pb, and La with DNA) are larger than those of flavonoids alone. Some studies show that IC_{50} of flavonoid—metal complexes (such as morin-Mn, Co, Ni, Cu, Zn, quercetin-Zn, and hesperetin-Cu) for tumour cells are smaller than those of flavonoids alone. Some studies show that 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
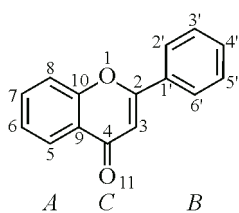
Fig. 1. Structure of the flavonoids

smaller than those of flavonoids alone. All these studies show that the anticancer and antioxidant activities of flavonoid—metal complexes are better than those of flavonoids alone. However, the activities of flavonoid—metal complexes may be not good through a serial of experiments according to some studies. Therefore, we built models based on the quantum chemical descriptors of flavonoid—metal complexes and their activities for predicting the activities of flavonoid—metal complexes, which are helpful for designing novel drugs. In the paper, we built two models based on the quantum chemical descriptors and F of flavonoid—metal complexes to analyze their anticancer activities. At last, we validated the accuracy and predictability of models. Our results provide foundation for the development of novel flavonoid—metal complexes and for predicting their anticancer activities.

EXPERIMENTAL

Calculation of quantum chemical descriptors of flavonoid—metal complexes. Ligands [2—8], including flavonol, flavonone, isoflavone, and xanthone, were calculated and analysed. Their 2D structures were built using Chemdraw, and their 3D structures were constructed and optimised using the MM+ method in Hyperchem. Full geometry optimisation of the ligands was conducted with the DFT method at the DFT/B3LYP theoretical level with the 6-31G*/LanL2DZ basis set using Gaussian 03. A series of quantum chemical descriptors were obtained for 14 flavonoid—metal complexes from their electronic structure properties. All calculated results are real frequencies to ensure stable configurations. The fundamental flavonoid structure is shown in Fig. 1. The structures of flavonoid—metal complexes containing ligands, metal, chelation sites, and ratio are listed in Table 1.
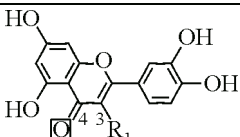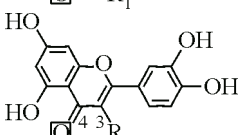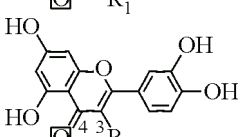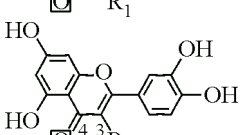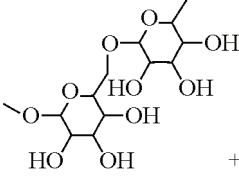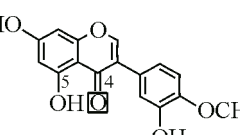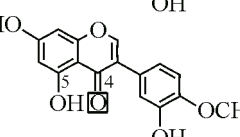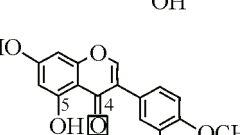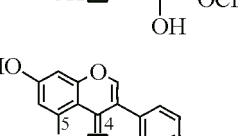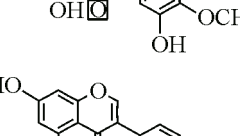
Stepwise linear regression and principal component analysis. By studying the relationship between the 19 quantum chemistry descriptors and IC_{50} of flavonoid—metal complexes for HepG2

T a b l e 1

Chelate sites and ratio of flavonoid-metal complexes

No	Kind	Ligand	Basis structure	R1	R2	Metal	Chelation	Ratio (metal/ligand)	Reference
1	2	3	4	5	6	7	8	9	10
1	Flavonoid derivative	Iso-euxanthone			no	Cu(II)	5-OH and 4=O	1:2	[2]
2	Flavonoid derivative	Iso-euxanthone			no	Cu(II)	5-OH and 4=O	1:2	[2]
3	Flavonoid derivative	Iso-euxanthone			no	Cu(II)	5-OH and 4=O	1:2	[2]
4	Flavonoid derivative	Iso-euxanthone		—CH₃	—OH	Cu(II)	5-OH and 4=O	1:2	[2]
5	Flavonoid derivative	Iso-euxanthone		—OH	—CH₃	Cu(II)	5-OH and 4=O	1:2	[2]

Continued Table 1

1	2	3	4	5	6	7	8	9	10
6	Flavo-nol	Quer-cetin	[structure: quercetin with R₁]	—OH	no	Cu(II)	3-OH and 4=O	1:2	[3]
7	Flavo-nol	Quer-cetin	[structure: quercetin with R₁]	—OH	no	Ni(II)	3-OH and 4=O	1:2	[4]
8	Flavo-nol	Quer-cetin	[structure: quercetin with R₁]	—OH	no	Zn(II)	3-OH and 4=O	1:2	[5]
9	Flavo-nol	Rutin	[structure: rutin with R₁]	[sugar structure] +	no	Se(IV)	3'-OH and 4'-OH	1:1	[6]
10	Flava-none	Hespe-retin	[structure: hesperetin, OCH₃, OH]	no	no	Cu(II)	5-OH and 4=O	1:2	[7]
11	Iso-flavo-nids	Iso-flavo-nids deri-vative	[structure, OCH₃, OH]	no	no	Cu(III)	5-OH and 4=O	1:2	[8]
12	Iso-flavo-nids	Iso-flavo-nids deri-vative	[structure, OCH₃, OH]	no	no	Mn(II)	5-OH and 4=O	1:2	[8]
13	Iso-flavo-nids	Iso-flavo-nids deri-vative	[structure, OCH₃, OH]	no	no	Ni(II)	5-OH and 4=O	1:2	[8]
14	Iso-flavo-nids	Iso-flavo-nids deri-vative	[structure, OCH₃, OH]	no	no	Zn(II)	5-OH and 4=O	1:2	[8]

through the stepwise linear regression method we obtained important quantum chemistry descriptors as dependent variables.

Principal component analysis is 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 as follows:

$$F1 = e_{11}x_1 + e_{21}x_2 + \cdots + e_{p1}x_p$$
$$F2 = e_{12}x_1 + e_{22}x_2 + \cdots + e_{p2}x_p$$
$$Fp = e_{1p}x_1 + e_{2p}x_2 + \cdots + e_{pp}x_p$$

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

To ensure the integrity of all quantum chemical descriptors, we used the method of principal component analysis to analyse 19 quantum chemistry descriptors. We can further obtain new independent variables that consist of original variables through principal component analysis.

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 a strong ability to approximate nonlinear functions, self-organisation, self-learning and so on. Back propagation ANN (BP ANN) is typical because this model has several advantages, including simple simulations and easy achievement [9].

Based on the results of the stepwise linear regression and the principal component analysis, we chose important quantum chemical descriptors and F as the independent variables, respectively, and IC_{50} as the dependent variable. All variables were analysed 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. Eleven sets of data were selected as training sets (black label) by the back propagation strategy, while the other three sets of data as test sets (red label) in Tables 3 and 4. The training sets and the tested sets are a random selection to ensure the applicability of the model.

In the process, the important quantum chemical descriptors are input parameters that build the vector matrix of 11×4, and IC_{50} is the goal matrix. The ANN analysis contains three network layers: the first layer (named the input layer) including four neurons, the second layer (named the hidden layer) including six neurons, and the third layer (named the output layer) including one neuron.

To test the accuracy and predictability of the models, we calculated the standard error (S) between IC_{50} from the references (named the experimental results later) and IC_{50} from our QSAR model calculation (named the predicted results later) and the correlation coefficient (R^2) of the predicted and experimental results in test sets.

$$S = \sqrt{\frac{\sum (x_i - \overline{x})^2}{N-1}}.$$

R^2. Using the experimental results as the X axis and the predicted results as the Y axis in excel, then completing the linear analysis to obtain R^2 automatically using excel, we establish the correlation and the similarity between the experimental and predicted results and then verify the accuracy and predictability of the models.

RESULTS AND DISCUSSION

Quantum chemical descriptors of flavonoid—metal complexes. We calculated the electronic structure properties of flavonoid—metal complexes and obtained 19 quantum chemical descriptors of flavonoid—metal complexes including: dipole, net charge of the *A* ring (Q_a), net charge of the *B* ring (Q_b), net charge of the *C* ring (Q_c), net charge of the metal ion (Q_{metal}), net charge of the O atom at site 1 (Q_{O1}), net charge of the C atom at site 4 (Q_{C4}), net charge of the C atom at site 2 (Q_{C2}), net charge of the C atom at site 6 (Q_{C6}), net charge of the C atom at site 8 (Q_{C8}), net charge of the O or N atom at site 11 ($Q_{O/N11}$), net charge of the C atom at site 4′ ($Q_{C4'}$), net charge the O or N atom at site 13 ($Q_{O/N13}$), the two highest occupied molecular orbital energies (E_{HOMO}, E_{H-1}), the two lowest unoccupied molecular orbital energies (E_{LUMO}, E_{L-1}), and two LUMO-HOMO gaps (ΔE_{L-H1}, ΔE_{L-H}), which are shown in Table 2. Based on the frontier molecular orbital theory, the anticancer 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 favourable for accepting the electrons.
- The planarity area of the intercalative ligand because a larger planarity area is favourable for the activity.

According to the calculation results, the *A* and *C* rings 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 at 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

Table 2

Structure parameters and IC50 *of flavonoid-metal complexes for* HepG2

No	IC$_{50}$	Dp	Qa	Qb	Qc	Q_{metal}	Q_{O1}	Q_{C2}	Q_{C4}	Q_{C6}	Q_{C8}
1	123uM	1.7039	−0.31471	−0.30238	0.28991	1.36933	−0.46753	0.34090	0.50027	−0.30914	−0.32005
2	62uM	2.6414	−0.32102	−0.30169	0.29772	1.37169	−0.46859	0.35130	0.50586	−0.31126	−0.32122
3	131uM	0.0160	−0.31549	−0.26556	0.31173	1.37868	−0.46811	−0.16286	0.50811	−0.30993	−0.32106
4	131uM	2.4933	−0.31296	−0.08532	0.25099	1.38209	−0.46664	0.36334	0.50697	−0.31245	−0.32157
5	36uM	4.8122	−0.31296	−0.0697	0.30379	1.34511	−0.46648	0.36064	0.49792	−0.30820	−0.32032
6	68.93uM	3.1829	0.13429	−0.34508	0.60065	1.38445	−0.46910	0.31649	0.41231	−0.37563	−0.34878
7	14.44uM	1.2753	0.15686	−0.32876	0.62873	1.06796	−0.46518	0.33362	0.41208	−0.37105	−0.34619
8	10.67uM	3.1435	0.12263	−0.35345	0.5846	1.68666	−0.47115	0.30606	0.40970	−0.37806	−0.35059
9	0.56uM	6.8590	0.11973	−0.36914	0.63163	0.95846	−0.46769	0.35125	0.45244	−0.37586	−0.34621
10	20uM	8.4532	0.10226	−0.35331	−0.41051	1.35195	−0.53223	0.07169	0.52753	−0.37435	−0.39153
11	52uM	4.5655	0.13148	−0.7679	0.16652	1.37196	−0.45662	−0.16011	0.49409	−0.38320	−0.35472
12	9uM	0.0086	0.07995	−0.77753	0.09842	1.10046	−0.46301	−0.16392	0.46295	−0.39418	−0.36344
13	11uM	2.9861	0.12138	−0.77827	0.16797	1.07481	−0.45307	−0.15322	0.48619	−0.38591	−0.35093
14	8uM	2.0217	0.1003	−0.76925	0.13702	1.70583	−0.45957	−0.16588	0.48906	−0.39382	−0.36159

No	IC$_{50}$	$Q_{C4'}$	Q_{O11}	Q_{O13}	E_{H-1}	E_{HOMO}	E_{LUMO}	E_{L-1}	ΔE_{L-H1}	ΔE_{L-H}
1	123uM	0.34090	−0.70039	−0.52317	−5.6062	−5.5243	−2.0128	−2.0008	3.6054	3.5115
2	62uM	0.33786	−0.69439	−0.51586	−5.7528	−5.3989	−1.8635	−1.7938	3.959	3.5354
3	131uM	0.31872	−0.68644	−0.53239	−5.7011	−5.6296	−2.4148	−2.3889	3.3122	3.2148
4	131uM	0.02415	−0.74206	−0.69221	−5.6905	−5.5288	−2.0778	−2.0517	3.6388	3.451
5	36uM	0.34939	−0.71037	−0.67832	−5.7947	−5.6908	−2.1545	−2.0212	3.7735	3.5363
6	68.93uM	0.29188	−0.74349	−0.68530	−5.3579	−5.276	−2.2519	−2.2062	3.1517	3.0241
7	14.44uM	0.29695	−0.68466	−0.68249	−5.6864	−5.4291	−2.6811	−2.4306	2.9985	2.748
8	10.67uM	0.28874	−0.77741	−0.68728	−5.2632	−5.1294	−2.1066	−2.0729	3.1903	3.0288
9	0.56uM	0.24749	−0.67253	−0.66651	−5.9734	−5.3312	−2.253	−1.8782	4.0952	3.0782
10	20uM	0.27506	−0.69284	−0.50417	−5.7705	−5.6103	−2.1045	−1.6924	4.0781	3.5058
11	52uM	0.30696	−0.67932	−0.52170	−5.6367	−5.5178	−1.8869	−1.6717	3.965	3.6309
12	9uM	0.30586	−0.63569	−0.52187	−4.8489	−4.5946	−1.5137	−1.3657	3.4832	3.0809
13	11uM	0.30823	−0.64010	−0.51981	−5.7049	−5.6163	−2.161	−2.0958	3.6091	3.4553
14	8uM	0.30648	−0.73987	−0.52188	−5.3739	−5.3391	−1.5556	−1.5156	3.8583	3.7835

larger contributions, which are shown in Table 2. The charge density distributions of these atoms have an important effect on their activities. The results show that Q_{O1}, Q_{O11}, Q_{C6}, Q_{C8}, and Q_{O13} have more negative charges. The Q_{O1} charges range from −0.45307 to −0.53223; the Q_{O11} charges range from −0.64010 to −0.77741; the Q_{C6} charges range from −0.30820 to −0.39382; the Q_{C8} charges range from −0.32005 to −0.39153; and the Q_{O13} charges range from −0.50417 to −0.69221. Q_{C4}, $Q_{C4'}$, and Q_{metal} have larger positive charges. The Q_{C4} charges range from 0.40970 to 0.52753; the $Q_{C4'}$ charges range from 0.02415 to 0.34939, and the Q_{metal} charges range from 0.95846 to 1.70583. ΔE_{L-H} varies from 2.748 to 3.7835, and ΔE_{L-H1} varies from 2.9985 to 4.0952.

Results of the stepwise linear regression and the principal component analysis. According to the study of the relationship between the 19 quantum chemical descriptors and IC$_{50}$ of the flavonoid—metal complexes for HepG2 through the of stepwise linear regression method, only three quantum chemical descriptors are important for IC$_{50}$ of the flavonoid—metal complexes for HepG2. They are the net charges of the *A* ring (Qa), net charges of the *B* ring (Qb), LUMO1-HOMO1 gap(ΔE_{L-H1}), and

net charge of the C atom at site 4′ ($Q_{C4'}$). The predictable equation is as follows:

$$IC_{50} = 212.105 - 193.616Q\text{a} - 33.586Q\text{b} - 168.837Q_{C4'} - 37.703\Delta E_{L\text{—H1}}.$$

Based on the results of the principal component analysis 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 93.270 %, which largely represents the original variances. The principal component equation is:

$$F1 = 0.12D_p - 0.04Q\text{a} - 0.07Q\text{b} - 0.01Q\text{c} - 0.007Q_{metal} + 0.133Q_{O1} + 0.049Q_{C2} + 0.106Q_{C4} -$$
$$- 0.007Q_{C6} + 0.05Q_{C8} - 0.124Q_{C4'} + 0.003Q_{O11} - 0.001Q_{O13} - 0.043E_{H\text{-}1} +$$
$$+ 0.057E_{Homo} + 0.258E_{Lumo} + 0.267E_{L\text{-}1} + 0.303\Delta E_{L\text{—H1}} + 0.214\Delta E_{L\text{—H}};$$

$$F2 = -0.207D_p - 0.304Q\text{a} + 0.163Q\text{b} - 0.083Q\text{c} - 0.019Q_{metal} - 0.032Q_{O1} + 0.025Q_{C2} + 0.166Q_{C4} +$$
$$+ 0.278Q_{C6} + 0.199Q_{C8} - 0.037Q_{C4'} + 0.065Q_{O11} + 0.083Q_{O13} + 0.051E_{H\text{-}1} + 0.019E_{Homo} +$$
$$+ 0.018E_{Lumo} - 0.043E_{L\text{-}1} - 0.073\Delta E_{L\text{—H1}} + 0\Delta E_{L\text{—H}};$$

$$F3 = -0.234D_p - 0.115Q\text{a} + 0.055Q\text{b} - 0.085Q\text{c} - 0.025Q_{metal} - 0.083Q_{O1} + 0.02Q_{C2} -$$
$$0.015Q_{C4} + 0.059Q_{C6} - 0.004Q_{C8} - 0.127Q_{C4'} + 0.032Q_{O11} + 0.017Q_{O13} + 0.347E_{H\text{-}1} +$$
$$+ 0.37E_{Homo} + 0.185E_{Lumo} + 0.147E_{L\text{-}1} - 0.143\Delta E_{L\text{—H1}} - 0.158\Delta E_{L\text{—H}};$$

$$F4 = -0.188D_p + 0.03Q\text{a} - 0.178Q\text{b} - 0.095Q\text{c} + 0.12Q_{metal} + 0.043Q_{O1} - 0.324Q_{C2} + 0.066Q_{C4} -$$
$$- 0.027Q_{C6} - 0.031Q_{C8} + 0.333Q_{C4'} + 0.088Q_{O11} + 0.277Q_{O13} + 0.01E_{H\text{-}1} - 0.194E_{Homo} -$$
$$- 0.108E_{Lumo} - 0.173E_{L\text{-}1} - 0.1749\Delta E_{L\text{—H1}} + 0.072\Delta E_{L\text{—H}};$$

$$F5 = -0.036D_p + 0.057Q\text{a} - 0.199Q\text{b} + 0.355Q\text{c} - 0.011Q_{metal} + 0.563Q_{O1} + 0.016Q_{C2} -$$
$$- 0.124Q_{C4} - 0.073Q_{C6} + 0.236Q_{C8} + 0.053Q_{C4'} - 0.02Q_{O11} - 0.113Q_{O13} - 0.121E_{H\text{-}1} -$$
$$- 0.052E_{Homo} + 0.105E_{Lumo} + 0.034E_{L\text{-}1} + 0.139\Delta E_{L\text{—H1}} + 0.159\Delta E_{L\text{—H}};$$

$$F6 = 0.001D_p + 0.044Q\text{a} - 0.004Q\text{b} + 0.016Q\text{c} + 0.514Q_{metal} - 0.011Q_{O1} - 0.035Q_{C2} - 0.055Q_{C4} -$$
$$- 0.036Q_{C6} - 0.038Q_{C8} + 0.082Q_{C4'} - 0.441Q_{O11} - 0.032Q_{O13} + 0.071E_{H\text{-}1} - 0.14E_{Homo} +$$
$$+ 0.043E_{Lumo} - 0.062E_{L\text{-}1} - 0.093\Delta E_{L\text{—H1}} + 0.179\Delta E_{L\text{—H}}.$$

ANN results. Based on the results of the stepwise linear regression, we chose Qa, Qb, $Q_{C4'}$, and $\Delta E_{L\text{—H1}}$ as the independent variables and IC_{50} as the dependent variable to analyse by ANN. Then we obtained a QSAR model with the appointed convergence accuracy based on Qa, Qb, $Q_{C4'}$, $\Delta E_{L\text{—H1}}$ and IC_{50} after 72 training runs. Table 3 shows the predicted results of our QSAR model. In Tables 3, 11 data groups (black label) were used as the training sets and the other three data groups (red label) were the test sets. The Table 3 data groups (red label) were used to check the predictability of our model. The standard error between the experimental and predicted results was 0.4799. The predicted results agree well with the experimental values, and the residual is very small. The results suggest that the QSAR model is accurate and predictable.

Based on the results of the principal component analysis, we chose F as the independent variable and IC_{50} as the dependent variable to analyse ANN. Then we obtained a QSAR model with the appointed convergence accuracy based on F and IC_{50} after 134 training runs. Table 4 shows the predicted results of our QSAR model. In Table 4, the 11 data groups (black label) were used as the training sets and the other three data groups (red label) were the test sets. The Table 3 data groups (red label) were used to check the predictability of our model. The standard error between the experimental and predicted results is 0.4349. The predicted results agree well with the experimental values, and the residual is very small. Our results suggest that our QSAR model is accurate and predictable.

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

CONCLUSIONS

According to the results of the quantum chemical descriptors of flavonoid—metal complexes, the A, B, and C rings in these complexes have excellent planarity, resulting in a large conjugated system. Q_{O1}, Q_{O11}, Q_{C6}, Q_{C8}, and Q_{O13} have more negative charges. Q_{C4}, $Q_{C4'}$, and Q_{metal} have more positive charges. This result is the main reason for the charge distribution.

T a b l e   3

Input parameters for ANN *and the predicted results based on the stepwise linear regression*

No	Qa	Qb	$Q_{C4'}$	ΔE_{L-H1}	Experimental results IC50 uM	Predicted results IC50 uM	Error
1	−0.31471	−0.30238	0.3409	3.6054	123	122.9940	−0.0060
2	−0.32102	−0.30169	0.33786	3.959	62	62.0057	0.0057
3	−0.31549	−0.26556	0.31872	3.3122	131	131.0028	0.0028
4	−0.31296	−0.08532	0.02415	3.6388	131	130.9978	−0.0022
5	−0.31296	−0.0697	0.34939	3.7735	36	36.0044	0.0044
6	0.13429	−0.34508	0.29188	3.1517	68.93	68.9299	−0.0001
7	0.15686	−0.32876	0.29695	2.9985	14.44	14.4402	0.0002
8	0.12263	−0.35345	0.28874	3.1903	10.67	10.6702	0.0002
9	0.11973	−0.36914	0.24749	4.0952	0.56	0.5604	0.0004
10	0.10226	−0.35331	0.27506	4.0781	20	19.9992	−0.0008
11	0.13148	−0.7679	0.30696	3.965	52	52.0000	0.0000
12	0.07995	−0.77753	0.30586	3.4832	9	10.3006	1.3006
13	0.12138	−0.77827	0.30823	3.6091	11	11.9013	0.9013
14	0.1003	−0.76925	0.30648	3.8583	8	7.2999	−0.7001

T a b l e   4

Input parameters for ANN *and the predicted results based on the principal component analysis*

No	F1	F2	F3	F4	F5	F6	Experimental results IC50 uM	Predicted results IC50 uM	Error
1	0.12139	1.33678	0.03226	0.33018	0.11273	0.12294	123	122.9999	0.0001
2	0.85144	1.17011	−0.02770	−0.15176	0.29711	−0.14584	62	61.9999	0.0001
3	−1.16909	1.55579	−0.20329	1.51843	−0.35057	−0.01800	131	130.9999	0.0001
4	0.51548	1.29734	0.28504	−1.83070	−0.09636	0.27827	131	130.9999	0.0001
5	0.22394	0.94994	−0.90363	−0.53358	0.29496	0.23344	36	35.9999	0.0001
6	−1.22112	−0.73429	0.31933	−0.51897	0.19729	0.76721	68.93	68.9300	0.0000
7	−2.15048	−0.45402	−0.20049	−0.04588	0.16869	−0.78657	14.44	14.4399	0.0001
8	−0.96387	−0.83171	0.72165	−0.66430	0.12835	1.80315	10.67	10.6700	0.0000
9	0.39616	−1.02745	−0.96334	−1.60001	0.77955	−1.55055	0.56	0.5600	0.0000
10	0.52852	−0.75077	−0.78069	0.01046	−3.24871	0.00439	20	20.0108	0.0108
11	1.05420	−0.86328	−0.68535	0.85165	0.74114	0.08020	52	52.0000	0.0000
12	0.54603	−0.29063	2.93710	0.31785	−0.28962	−1.47257	9	7.7970	1.2030
13	−0.03114	−0.55645	−0.82287	1.42624	0.61949	−0.99215	11	10.2500	0.7500
14	1.29852	−0.80136	0.29198	0.89039	0.64596	1.67608	8	7.3299	0.6701

Based on the results of the stepwise linear regression analysis and the principal component analysis, we chose the important quantum chemistry descriptors, F as the independent variables, and IC$_{50}$ as the dependent variable analysed by ANN. We built two QSAR models and checked them with data from the references. Our results show that the two models in this paper have reliable predictability. The predictability of the model based on the principal component analysis is better. Therefore, we conclude that the two models can offer a theoretical instruction for predicting the activities of flavonoid—metal complexes and designing new drugs.

REFERENCES

1. *Zhen-Xue Z., Ke L., Shou-Yu W. et al.* // Chinese traditional and herbal drugs. – 1996. – **27**, N 3. – P. 179 – 182.
2. *Hui-Fan W.* LanZhou University, Doctor′s degree thesis.
3. *Jun T., Bochu W., Liancai Z.* // J. Biol. Inorg. Chem. – 2009. – **14**. – P. 727 – 739.
4. *Jun T., Bochu W., Liancai Z.* // Biometals. – 2010. – **23**. – P. 1075 – 1084.
5. *Jun T., Bochu W., Liancai Z.* // Bioorg. Med. Chem. – 2009. – **17**. – P. 614 – 620.
6. *Peng Z., Jian-Zhang W., Pei-Hong C. et al.* // Chin. Pharm. – 2007. – **24**, N 42. – P. 1905 – 1907.
7. *Mingxiong T., Jinchan Z., Yingming P. et al.* // Bioinorg. Chem. Appl. – 2009. – Article ID 347872, 9.
8. *Xiang C., Li-Jun T., Yu-Na S. et al.* // J. Inorg. Biochem. – 2010. – **104**. – P. 379 – 384.
9. *Ji-Qian F., Ying L.* The introduction of BP ANN model. – Beijing, Peoples′s medical publishing, 2002.