Available online at www.tox.or.kr

Quantitative Structure-Activity Relationships for Radical Scavenging Activities of Flavonoid Compounds by GA-MLR Technique

Ae-Son Om[1], Jae-Chun Ryu[2] & Jae-Hyoun Kim[3]

[1]Department of Food & Nutrition, College of Human Ecology, Hanyang University, Seoul 133-791, Korea
[2]Laboratory of Cellular and Molecular Toxicology, Korea Institute of Science and Technology/University of Science and Technology, PO BOX 131, Cheongryang, Seoul 130-650, Korea
[3]Department of Health Science, School of Natural Science, Dongduk Women's University, Wolgok-dong 23-1, Seongbuk-gu, Seoul 136-714, Korea
Correspondence and requests for materials should be addressed to J. H. Kim (kjhyon@dongduk.ac.kr)

Accepted 18 March 2008

Abstract

The quantitative structure-activity relationship (QSAR) of a set of 35 flavonoid compounds presenting antioxidant activity was established by means of Genetic Algorithm-Multiple Linear Regression (GA-MLR) technique. Four-parametric models for two sets of data, the 1,1-diphenyl-2-picryl hydrazyl (DPPH) radical scavenging activity (R^2=0.788, Q^2_{cv}=0.699 and Q^2_{ext}=0.577) and scavenging activity of reactive oxygen species (ROS) induced by H_2O_2 (R^2=0.829, Q^2_{cv}=0.754 and Q^2_{ext}=0.573) were obtained with low external predictive ability on a mass basis, respectively. Each model gave some different mechanistic aspects of the flavonoid compounds tested in terms of the radical scavenging activity. Topological charge, H-bonding complex and deprotonation processes were likely to be involved in the radical scavenging activity.

Keywords: DPPH, Flavonoids, Radical scavenging activity, Reactive oxygen species (ROS), QSAR

Reactive oxygen species (ROS) include superoxide anion radical (O_2^-), singlet oxygen ($^1O_2$), hydrogen peroxide (H_2O_2), and the most highly reactive hydroxyl radical ($^{\cdot}OH$) is derived from the metabolic reduction of molecular oxygen to highly reactive and toxic species[1].

A number of flavonoids show positive biological or medicinal effects such as lipooxygenase inhibition[2], chemoprotective effects on colorectal cancer[3], anti-inflammatory activity[4] and antiparasitic activity[5]. Antioxidants can interfere with the oxidation process by reacting with free radicals, chelating catalytic metals, and also by acting as oxygen scavengers.

QSAR has been often used to find correlations between biological activities and physicochemical properties of compounds[6,7]. The ability of a compound to act as a free radical scavenger is related to its standard one-electron reduction potential (E°)[8-10], a measure of the reactivity of an antioxidant as hydrogen or electron donor. In another words, the electronegativity of OH/O^- groups depends on the strength of H-bonding in phenol/phenolate involved in H-bond complexation. The O-H BDE (bond dissociation enthalpy) is mainly governed by the resonance effect of electron-donating or electron-withdrawing group[11]. The singly -occupied molecular orbital (SOMO) of the 3-OH radical are localized on the O-atom, whereas SOMO of the 3′-OH and the 4′-OH radicals are delocalized, agreeing with high BDE for the first one and low BDE for the latter two[12]. Parameters such as ΔH_f, HOMO (lowest unoccupied molecular orbital) and LUMO (lowest unoccupied molecular orbital), and number of OH groups[6], or dipole moment and shape parameters[13] were applied for description of antioxidant activity of flavonoids and other phenolic compounds.

Recently, authors have intensively discussed the activity due to structural characteristics of flavonoid antioxidants[14-17]. The antioxidant activity of depends on the number of phenolic hydroxyl groups and the location of the hydroxyl groups. Moreover, the presence of 2,3-double bond in conjugation with a 4-oxo function in ring C, 3- and 5-OH groups, 3,5,7-trihydroxy, *ortho*-catechol group (3′,4′-OH), glycosylation in the ring and the dihedral angle between the B- and C-ring are additional determinants for a high antioxidant capacity.

An easy, rapid and sensitive method for the antioxidant screening is free radical scavenging assay using 1,1-diphenyl-2-picryl hydrazyl (DPPH) radical method. In the presence of an antioxidant, DPPH radical obtains one more electron as reactive hydrogen acceptor and the absorbance decreases[18,19]. QSAR analyses for DPPH radical scavenging activities of OH-benzlactones indicated that antioxidant and radical scavenging activities could be expressed by the physicochemical parameters and the hydrogen bonding[20]. However, there has been virtually no enough information for the QSAR associated with DPPH radical scavenging activities.

The aim of the research is to construct the predictive QSAR model for two different radical scavenging activity data on structurally diverse flavonoid compounds using GA-MRL analysis, and to understand the different mechanistic features for the radical scavenging activity.

QSAR of ROS Scavenging Activity Using Dragon Descriptors

We investigated the applicability of *Dragon* descriptors to radical scavenging activity data to construct the QSAR model. Following descriptor pruning procedures, four descriptors were selected to build the final QSAR model for the data set, consisting of 29 compounds. In order to investigate the possible existence of outliers from the 35 compounds which constitute the entire data sets. The LOO procedure indicated that chemicals, no. 7 (Astragalin), 17 (Apigenin) and 22 (Isoliquiritigenin), were influential outliers and need to be omitted from the procedure of regression analysis. Although no. 24 (Trifolirhizin) turned out to be a weak influential one that fell slightly outside from the warning leverage limit, we decided not to delete the compound from the training set.

On removal of outliers, four subsets of different descriptors used in the final model were chosen on the basis of the best values of R^2 and Q^2_{cv}, and the outlier-free data, from the 80 selected models. The chosen parameters in Eq. 5 were GATS2m (Geary autocorrelation-lag 2/weighted by atomic masses, 2D autocorrelations), JGI5 (mean topological charge index of order 5, 2D Galvez topological charge indices), JGI6 (mean topological charge index of order 6, 2D Galvez topological charge indices) and R6m (R autocorrelation of lag 6/weighted by atomic masses, 3D GETAWAY descriptors). The regression analysis of the reduced data set consisting of 32 compounds yielded a predictive model on mass basis (μg/mL) with significant statistics as below:

$$\log(1/\mathrm{ROS}) = 5.15(\pm 0.77)\mathrm{GATS2m} + 55.40(\pm 10.67)\mathrm{JGI5} - 43.71(\pm 8.40)\mathrm{JGI6} - 2.29(\pm 0.34)\mathrm{R6m} - 5.58(\pm 1.08) \quad \text{(Eq.1)}$$

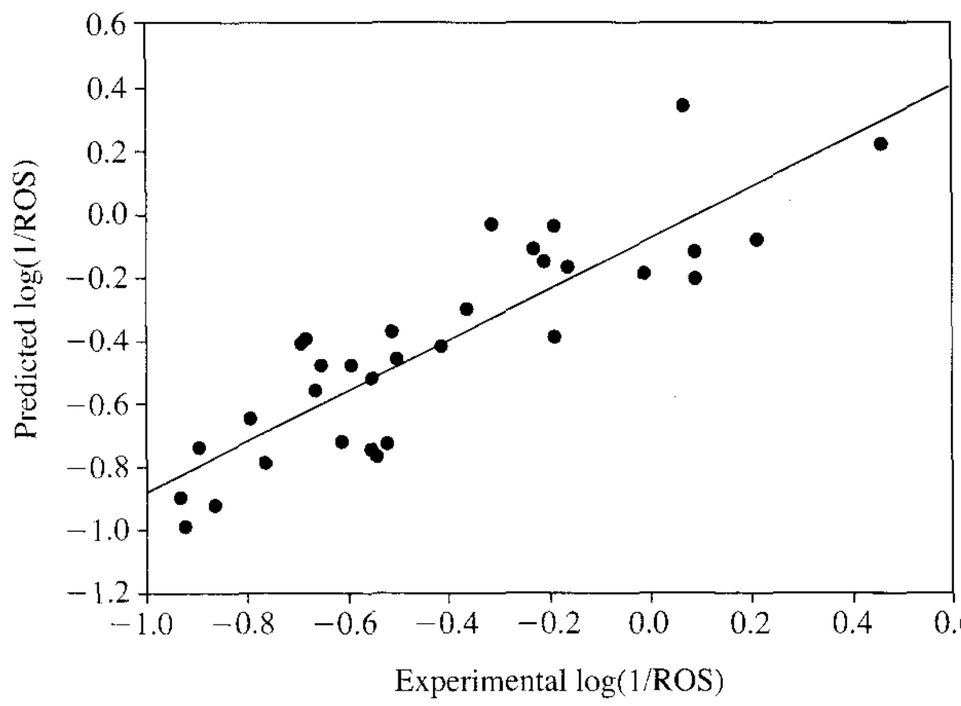

Figure 1. Leave-one-out cross-validated experimental *versus* predicted log(1/ROS) values (R^2=0.757) (Eq. 1).

MLR analyses demonstrate that the best RSA model has the following statistical quality: n=32, R^2=0.829, Q^2_{cv}=0.754, R^2_{adj}=0.803, Q^2_{boot}=0.717, K_x=36.47, K_{xy}=44.57, F=32.66, SE=0.158, $PRESS$=0.970.

The calculated standardized regression coefficient values that contribute to the prediction of log(1/RSA) are as follows: GATS2m=0.840, JGI5=0.653, R6m =−0.597 and JGI6=−0.42104. Log(1/ROS) was found to be most influenced by GATS2m, and impacted by JGI5, JGI6 and R6m in a less degree. Scavenging activity decreases when GATS2m and JGI5 increase while the activity increase JGI6 and R6m are more negative.

The resultant final model obtained shows high flexibility and the predicting ability with mostly 2D descriptors, which is confirmed by the close values of R^2_{adj} and Q^2_{cv}. The predictive power, Q^2_{cv}, of the equations was confirmed by LOO cross-validation method. Table 1 shows a comparison of the experimental (Y-Exp), calculated (Y-Calc), LOO-predicted (Y-Pred), HAT and standardized error predicted values based on Eq. 1. A linear regression plot of LOO-cross-validated predictive *versus* experimental log(1/ROS) based on Eq. 1 is depicted in Figure 1.

In terms of applicability of domain, all data points fell within +/− 2 standard deviations from the mean although 24 (Trifolirhizin) was slightly out of critical warning leverage point. However, we decided not to delete it from the training set.

A plot of experimental *versus* predicted log(1/ROS)

values of 32 training set molecules shows good linearity as depicted in Figure 1.

GATS2m is the mass-weighted Geary graph spatial autocorrelation coefficient of the second lag. The Geary coefficient is a distance-type function varying from zero to infinity so that strong autocorrelation produces low values of this index[25]. The Galvez charge indices (mean topological charge indices: JGI5 and JGI6) played a role in antioxidant activity of flavonoids[26,27]. The Galvez charge indices contain important information on the relationship between the flavonoid structures and their activities by describing the molecular topology and the intramolecular charge transfer (intramolecular H-bonding). The presence of two charge descriptors suggests that the reaction between radical species and flavonoid molecule is basically radical addition, and that the deprotonated flavonoids obtain hydrogens after reaction. It is presumed from the model that positive sign of JGI5 indicates of proton addition, and intramolecular charge complex as hydroquinone for negative sign of JGI6. R6m belong to R-GETAWAY descriptors which combine the information provided by the molecular influence/distance matrix (MIM) with geometry matrix G[28].

No QSAR models with less than four descriptors were proposed because Q^2_{LOO} and Q^2_{ext} significantly decreased to 0.653 and 0.391 with three descriptors and 0.514 and 0.335 with two descriptors, indicating a poor internal and external predictivity.

For external validation (test set correlation validation), data set was divided into a training (n=22) and an external validation set (n=10) by 30% random splitting. The overall external prediction accuracy for the 94 modeling set ranged from 0.16 to 0.81 (average: 0.43) as measured. A training set selection of compounds led to Q^2_{LOO}=0.754 and Q^2_{ext}=0.573, with moderate internal and external predictivity. However, when we established the QSAR on molar basis (M/l) instead of mass basis (μg/mL), we could not validate overall predictivity of the model equations (Eq. 1 and 2) through the use of external validation.

QSAR of DPPH Radical Scavenging Activity Using Dragon Descriptors

Four outlier chemicals (Isoquercitrin (13), Maackianin (20), Isoliquiritigenin (22), Baicalin (28) were deleted from 35 compounds. The best predictive equation on mass basis (mg/L) in the four-parametric regression is given below:

$$\log(1/\text{DPPH}) = 0.440(\pm 0.098)\text{EEig11x} - 10.872(\pm 3.325)\text{G1u} + 3.987(\pm 0.993)\text{ISH}$$

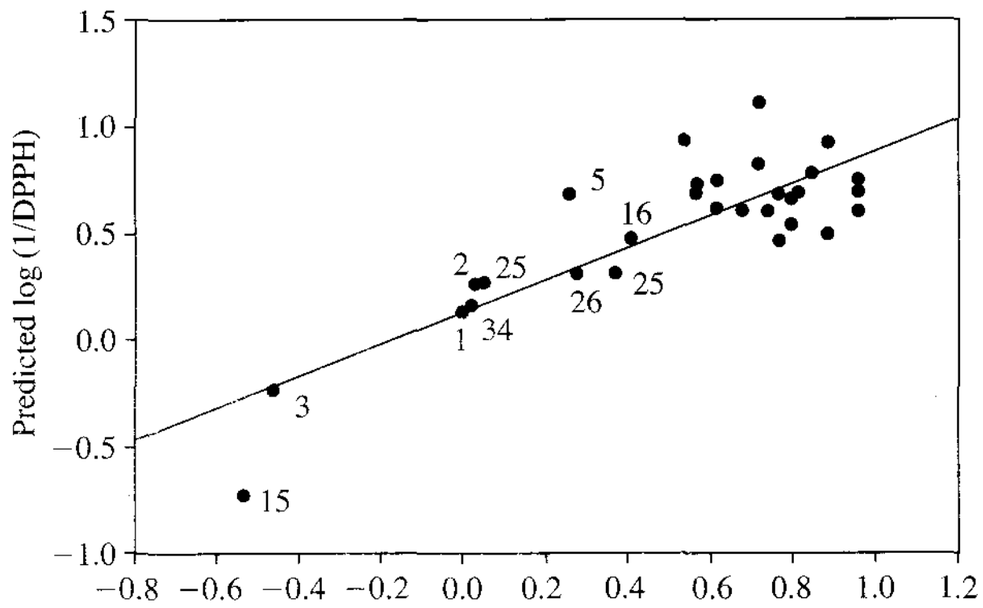

Figure 2. Leave-one-out cross-validated experimental *versus* predicted log(1/DPPH) values (R^2=0.702) (Eq. 2).

$$+10.583(\pm 1.938)\text{RPCG} - 3.027(\pm 1.157) \quad \text{(Eq. 2)}$$

n=31, R^2=0.788, Q^2_{cv}=0.699, R^2_{adj}=0.756, Q^2_{boot}=0.669, K_x=60.72, K_{xy}=66.14, F=24.194, SE=0.197

The calculated standardized regression coefficient values that contribute to the prediction of log(1/ROS) are as follows: EEig11x=0.9267, G1u=−0.4596, ISH=0.5391 and RPCG=0.8745. Log(1/ROS) was found to be most influenced by EEig11x (2D edge adjacency index) and RPCG (charge descriptor), and impacted by G1u (3D WHIM descriptor) and ISH (3D GETAWAY descriptor) in a less degree. The chemical domain of pplicability was verified by the leverage approach. All data points fell within +/− 2 standard deviations from the mean, and within warning leverage point (h*=0.484).

To establish a reliable QSAR model, an external test set consisting of ten molecules was used to measure the external predictivity. The overall external prediction accuracy for the 94 modeling set ranged from 0 to 0.66 (average: 38.49) as measured by 30% random splitting. A training set selection of compounds led to Q^2_{cv}=0.754 and Q^2_{ext}=0.577 for a test set, with moderate internal and external predictivity.

Figure 2 shows the plot of the predicted *versus* experimental log(1/DPPH) values of 31 training set compounds. As it can be seen in Figure 2, three different clusters of compounds are clearly separated in the plot. The compounds on numbers are evidently separated from a cluster in the top right corner.

RPCG is the charge of the most positively charged atom divided by sum of the total positive charge. It encodes various aspects of the molecular disposition to undergo electrostatic interactions[29]. EEig11x is re-

Table 1. A list of chemical names, experimental (Y-Exp), calculated (Y-Calc), LOO-predicted (Y-Pred) and residual values for both ROS and DPPH radical scavenging activities a mass basis (μg/mL).

ID	Object	ROS scavenging activity[a]					DPPH scavenging activity[b]				
		Y-Exp.	Y-Calc	Y-Pred	Hat	Std. Err. Pred.	Y Exp.	Y-Calc	Y-Pred	Hat	Std. Err. Pred.
1	Quercetin	−0.89	−0.76	−0.74	0.157	1.08	0.02	0.1	0.16	0.389*	0.88
2	Rutin	−0.86	−0.91	−0.92	0.168	−0.39	0.03	0.2	0.26	0.27	1.4
3	(+)-Catechin	−0.93	−0.91	−0.9	0.288	0.22	−0.46	−0.32	−0.24	0.361*	1.42
4	Robinin	−0.5	−0.47	−0.46	0.151	0.31	0.74	0.62	0.6	0.109	−0.78
5	Genistein	−0.19	−0.06	−0.04	0.131	1.02	0.26	0.64	0.68	0.105	2.29
6	Vitexicarpin	−0.69	−0.44	−0.41	0.124	1.91	0.89	0.92	0.92	0.113	0.19
7	Astragalin						0.68	0.6	0.6	0.056	−0.43
8	Isorhamnetin	−0.61	−0.71	−0.72	0.123	−0.77	0.8	0.67	0.66	0.093	−0.74
9	Linarin	−0.23	−0.12	−0.11	0.097	0.76	0.54	0.88	0.93	0.13	2.12
10	Chrysin	−0.21	−0.16	−0.15	0.096	0.42	0.62	0.73	0.74	0.146	0.69
11	Naringenin	−0.31	−0.06	−0.03	0.131	1.9	0.96	0.73	0.69	0.149	−1.47
12	Acacetin	0.46	0.28	0.22	0.254	−1.73	0.57	0.65	0.68	0.222	0.62
13	Isoquercitrin	−0.79	−0.67	−0.65	0.079	0.87					
14	Icariin	−0.19	−0.38	−0.39	0.047	−1.32	0.89	0.54	0.49	0.114	−2.12
15	Epicatechin	−0.92	−0.97	−0.99	0.271	−0.51	−0.53	−0.63	−0.73	0.470*	−1.43
16	Hesperidin	−0.36	−0.3	−0.3	0.093	0.44	0.41	0.46	0.47	0.09	0.3
17	Apigenin						0.77	0.53	0.46	0.222	−1.77
18	Genistin	−0.16	−0.17	−0.17	0.061	−0.06	0.72	0.81	0.82	0.112	0.51
19	Apigenin 7-O-neohesperidoside	−0.01	−0.17	−0.19	0.097	−1.19	0.77	0.69	0.68	0.081	−0.46
20	Maackianin	0.07	0.24	0.34	0.379*	2.16					
21	Sophoflavescenol	−0.55	−0.7	−0.75	0.221	−1.42	0.72	1.06	1.11	0.13	2.11
22	Isoliquiritigenin										
23	Vitexin	−0.66	−0.57	−0.56	0.126	0.74	0.85	0.79	0.78	0.077	−0.38
24	Trifolirhizin	−0.65	−0.56	−0.48	0.487**	1.52	0.8	0.63	0.54	0.363*	−1.64
25	Naringin	−0.59	−0.49	−0.48	0.105	0.71	0.37	0.32	0.31	0.106	−0.28
26	Poncirin	−0.41	−0.42	−0.42	0.084	−0.09	0.28	0.31	0.31	0.117	0.15
27	Diosmetin 7-O-glucoside	0.21	−0.05	−0.08	0.101	−1.95	0.62	0.61	0.61	0.144	−0.05
28	Baicalin	−0.76	−0.78	−0.79	0.102	−0.16					
29	Eupatilin	0.09	−0.17	−0.2	0.097	−1.89	0.96	0.77	0.75	0.094	−1.14
30	Liquiritigenin	0.09	−0.1	−0.12	0.111	−1.4	0.82	0.71	0.69	0.149	−0.72
31	Isoxanthohumol	−0.51	−0.38	−0.37	0.083	0.98	0.68	0.6	0.59	0.133	−0.46
32	Spinosin	−0.54	−0.73	−0.77	0.168	−1.54	0.57	0.72	0.73	0.102	0.89
33	Sophoraflavanone G	−0.52	−0.68	−0.73	0.224	−1.51	0.96	0.65	0.6	0.138	−1.95
34	Sophoraflavanone D	−0.68	−0.47	−0.4	0.231	2.06	0	0.12	0.13	0.112	0.73
35	Kenusanone I	−0.55	−0.52	−0.52	0.115	0.18	0.05	0.25	0.27	0.1	1.2

Y-Exp. represents the negative logarithm of % ROS scavenging activity per μg/mL.
[a]3*HAT=0.469; [b]3*HAT=0.484

lated to topological structure and encodes the substituent effect on the acidity of substituted phenol. ISH denotes standard information content taking into account equivalence of H matrix elements. G1u is the 1st component symmetry directional WHIM index/ unweighted, encoding molecular symmetry that extracts the global symmetry information[30].

Discussion

Table 2 summarizes statistical results of the two MLR models. The regression analyses showed that the final QSAR models for two different types of scavenging activities produced statistically significant predictivities with four Dragon descriptors. The appearance of different outliers and descriptors in both types of models implies that each model may be acting by an obviously different mechanism or elicit changes in the scavenging activity.

We obtained two four-parametric QSAR models and descriptors for two different training sets (ROS and DPPH) in the generation of each model with a low predictivity. An additional study however, showed that the 3-D QSAR result seems not better than that derived from the 2-D QSAR, which may arise

Table 2. A summary of descriptors and statistics of GA-MLR models obtained from log(1/ROS) and log(1/DPPH).

Descriptors	*n*	R^2	Q^2_{LOO}	Q^2_{boot}	Q^2_{ext}	*SE*	*F*	*PRESS*
ROS scavenging activity GATS2m, JGI5, JGI6, R6m	32	0.829	0.754	0.717	0.573	0.158	32.66	0.970
DPPH radical scavenging activity EEig11x, G1u, ISH, RPCG	31	0.788	0.699	0.669	0.577	0.197	24.19	1.428

GATS2m (Geary autocorrelation-lag 2/weighted by atomic masses), JGI5 (mean topological charge index of order 5), JGI6 (mean topological charge index of order 6), R6m (R autocorrelation of lag 6/weighted by atomic masses), EEig11x (Eigenvalue 11 from edge adjacent matrix weighted by edge degrees), G1u (1st component symmetry directional WHIM index/unweighted), ISH (standardized information content on the leverage equality) and RPCG (relative positive charge)

from the fact that the radicals, the targets of antioxidant, are small molecules, hence the radical-scavenging activity has little relevance to their 3-D structures, possibly due to the calculation of data on a mass basis instead of molar basis. The different selection of descriptors appears to be largely due to the inherent variability between cell experiment and experiment with solid DPPH. The QSAR model by first ROS data set demonstrated a more homogeneous distribution compared to the model by DPPH data set (Figure 2). Three different classes of compounds are clearly separated in the plot of DPPH study, where (+)-Catechin and Epicatechin had high leverage values and are to be considered influential (strong scavenging activity). Therefore, it is interesting to note that a biggest cluster consisting of 21 compounds includes mostly lower molecular weight polyphenolic compounds. A separate QSAR model for each cluster may be needed to reveal other selective responses for future study. It is shown that constitutive descriptors (*e.g.* number of OH groups in the molecule) have no significant role in QSAR models, which is consistent with a previous finding[26]. It is clear that topological charge, H-bonding complex and it's deprotonation processes are involved in the radical scavenging activity.

In this study, it was possible to consider some mechanistic aspects of the structure of the flavonoid compounds studied that are associated with the free radical scavenging activity of ROS and DPPH radical scavenging activity on a mass basis. However, MLR regression models performed with external validation on a molar basis by a random sampling method yielded no prediction power, due to high multi-collinearity in both data sets.

There is a limitation that needs to be acknowledged and addressed to elucidate their underlying overall aspects of mechanisms of both ROS and DPPH radical scavenging activities. Our previous study works[31] revealed that DPPH QSAR models used a set of 3D-Morse, WHIM and GETAWAY descriptors to interpret the radical scavenging mechanisms of flavonoids. This difference must be attributable to the calculation of the data set based on mass basis rather than molar basis, showing a partial mechanism of whole spectrum of the radical scavenging activity.

Additionally, there is the need for additional high-quality quantitative data on structurally diverse chemicals for further investigation. Much higher QSAR correlation coefficients (*R*) for flavonoid end-point antioxidant activity in all 3 assay systems using the frontal polygon (FP) method (evaluated in the biochemical, enzymatic, and whole cells assay systems, respectively) were obtained in a recent QSAR study[32]. Hence, the combination of different assay systems improves the accuracy of data even though DPPH evidently offers a convenient and accurate antioxidant activity.

Methods

Training Data Sets

Two types of data sets of experimental DPPH-RSA and scavenging activity of intercellular ROS for 35 polyphenolic antioxidants were taken from the literature[21]. The names of the flavonoids, ROS and DPPH scavenging activity values of training set compounds are listed in Table 1.

Descriptor Calculation

By using Hyperchem software 7.0 (Hypercube, Inc., USA), 200 chemicals were drawn and named by CAS -number. Molecular mechanic force field (MM+) was selected for the geometry optimization using Polak-Ribiere algorithm with a maximum cycle (10000) and a convergence limit of the 0.005 kcal/mol. After optimizing chemical structures from the Hyperchem 7 software, Dragon 5.4 (TALETE srl-Milano, Italy) is employed for the calculation of molecular descriptors[22].

Chemometric Methods

For the validation of the model, the combination GA-MLRA technique was utilized to select. the ap-

propriate descriptors and to generate different QSAR models as implemented in the Mobydigs program[23]. This algorithm allows to construct the models with the following statistical characteristics: high squared correlation coefficient R^2, low standard deviation S, the least number of descriptors, Fisher coefficient F, Predictive Residual Error Sum of Squares (*PRESS*), Standard error of calibration (*SEC*), standard error of prediction (*SEP*), explained variance (R^2_{adj}) as additional statistical parameters. The predictive stability of the models was verified applying internal Leave-One-Out (*LOO*). The robustness of the model and its predictivity was guaranteed by both the stability of Q^2_{LOO}, and bootstrap (Q^2_{boot}). The model obtained on the first chemicals selected is used to predict the values for the excluded sample, then Q^2_{LOO} is calculated for each model. The proposed model was also checked for reliability and robustness by permutation testing: new models were recalculated for randomly reordered response (Y-scrambling). These chemometric validation techniques are well described in literature[24].

The collinearity among descriptors must be checked, and correlation between the X block and the Y response verified (the QUIK rule of Todeschini based on K: K_{xy} must be significantly higher than K_x).

For appilicability of domain, the developed model was checked for the chemical domain to verify the prediction reliability. In the Williams plot, or OLS outlier and leverages (hat diagonals; $h^*=3p'/n$), standardized cross-validated (or jackknifed) residuals are also calculated.

References

1. Waris, G. & Ahsan, H. Reactive oxygen species: role in the development of cancer and various chronic conditions. *J Carcinog* **5**:1-8 (2006).
2. Schneider, I. & Bucar, F. Lipoxygenase inhibitors from natural plant sources. I. Medicinal plants with inhibitory activity on arachidonate 5-lipoxygenase and 5-lipoxygenase/cyclooxygenase. *Phytother Res* **19**:81-102 (2005).
3. Theodoratou, E. *et al.* Dietary flavonoids and the risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev* **16**:684-693 (2007).
4. Kim, H. P., Son, K. H., Chang, H. W. & Kang, S. S. Anti-inflammatory plant flavonoids and cellular action mechanisms. *J Pharmacol Sci* **96**:229-245 (2004).
5. Tasdemir, D. *et al.* Antitrypanosomal and antileishmanial activities of flavonoids and their analogues: *In vitro, In vivo*, structure-activity relationship, and quantitative structure-activity relationship studies. *Antimic Agents Chemother* **50**:1352-1364 (2006).
6. Cheng, Z., Ren, J., Li, Y., Chang, W. & Chen, Z. Establishment of a quantitative structure-activity relationship model for evaluating and predicting the protective potentials of phenolic antioxidants on lipid peroxidation. *J Pharm Sci* **92**:475-484 (2003).
7. Cabrera, M. *et al.* Synthetic chalcones, flavanones, and flavones as antitumoral agents: Biological evaluation and structure-activity relationships. *Bioorg Med Chem* **15**:3356-3367 (2007).
8. Heijnen, C. G., Haenen, G. R., Vekemans, J. A. & Bast, A. Protection of flavonoids against lipid peroxidation: the structure activity relationship revisited. *Free Radic Res* **36**:575-581 (2002).
9. Kontogiorgis, A. C., Pontiki, A. E. & Hadjipavlou-Litina, D. A review on quantitative structure-activity relationships (QSARs) of natural and synthetic antioxidants compounds. *Mini Rev Med Chem* **5**:563-574 (2005).
10. Weber, K. C., Honorio, K. M., Bruni, A. T., Andricopulo, A. D. & da Silva, A. B. F. A partial least squares regression study with antioxidant. flavonoid compounds. *Struct Chem* **17**:307-313 (2006).
11. Zhang, H. Y., Sun, Y. M. & Chen, D. Z. O-H bond dissociation energies of phenolic compounds are determined. by field/inductive effect or resonance effect? A DFT study and its implication. *Quant Struct-Act Relat* **20**:148-152 (2001).
12. Trouillas, P. *et al.* A theoretical study of the conformational behavior and electronic structure of taxifolin correlated with the free radical-scavenging activity. *Food Chem* **88**:571-582 (2004).
13. Rasulev, B. F., Abdullaev, N. D., Syrov, V. N. & Leszczynski, J. A quantitative structure-activity relationship (QSAR) study of the antioxidant activity of flavonoids. *QSAR Comb Sci* **24**:1056-1065 (2005).
14. van Zanden, J. J. *et al.* Quantitative structure activity relationship studies on the flavonoid. *Biochem Pharm* **69**:699-708 (2005).
15. Reis, M., Lobato, B., Lameira, J., Santos, A. S. & Alves, C. N. A theoretical study of phenolic compounds with antioxidant properties. *Eur J Med Chem* **42**:440-446 (2007).
16. Khlebnikov, A. I., Schepetkin, I. A., Domina N. G., Kirpotina, L. N. & Quinn, M. T. Improved quantitative structure-activity relationship models to predict antioxidant activity of flavonoids in chemical, enzymatic, and cellular systems. *Bioorg Med Chem* **15**: 1749-1770 (2007).
17. Amic, D. *et al.* SAR and QSAR of the antioxidant activity of flavonoids. *Curr Med Chem* **14**:827-845 (2007).
18. Koleva, I. I., Van Beek, T. A., Linssen, J. P. H., de Groot, A. & Evstatieva, L. N. Screening of plant extracts for antioxidant activity: a comparative study on three testing methods. *Phytochem Anal* **13**:8-17 (2002).
19. Sanchez-Moreno, C. Methods used to evaluate the free radical scavenging activity in foods and biological systems. *Food Sci Technol Intl* **8**:121-137 (2002).
20. Yamagami, C. *et al.* Quantitative structure-activity relationship analyses of antioxidant and free radical

scavenging activities for hydroxybenzalacetones. *Bioorg Med Chem Lett* **14**:5629-5633 (2004).
21. Kang, K. A. *et al.* Screening of antioxidant and anticancer effects from flavonoids. *Cancer Prev Res* **11**: 235-239 (2006).
22. Todeschini, R., Consonni, V., Mannhold, R., Kubinyi H. & Timmerman, H. *HANDBOOK of MOLECULAR DESCRIPTORS*. Weinheim: Wiley-VCH, 2000.
23. Todeschini, R., Consonni, V. & Pavan, M. MobyDigs: *SOFTWARE for REGRESSION and CLASSIFICATION MODELS by GENETIC ALGORITHMS*. In *Nature-inspired methods in chemometrics: Genetic algorithms and artificial neural networks*, edited by Leardi, R., Elsevier, 2003.
24. Gramatica, P. Principles of QSAR models validation: internal and external. *QSAR Comb Sci* **26**:694-701 (2007).
25. Fortin, M. J., Dale, M. R. T. & ver Hoef, J. *SPATIAL ANALYSIS in ECOLOGY*. Volume 4, pp. 2051-2058. In *Encyclopedia of Environmetrics* (ISBN 0471 899976), edited by Abdel H, El-Shaarawi and Walter W. Piegorsch. John Wiley & Sons, Ltd, Chichester, 2002.
26. Farkas, O., Jakus, J. & Héberger, K. Quantitative structure-antioxidant activity relationships of flavonoid compounds. *Molecules* **9**:1079-1088 (2004).
27. Fernández, M., Caballero, J., Helguera, A. M., Castro, E. A. & González, M. P. Quantitative structure-activity relationship to predict differential inhibition of aldose reductase by flavonoid compounds. *Bioorg Med Chem* **13**:3269-3277 (2005).
28. Katritzky, A. R., Mu, L., Lobano, V. S. & Karelson, M. Correlation of boiling points with molecular structure: I. A training set of 298 diverse organics and a test set of 9 simple inorganics. *J Phys Chem* **100**: 10400-10407 (1996).
29. Stanton, D. T. & Jurs, P. C. Development and use of charge partial surface-area structural descriptors in. computer-assisted quantitative structure property relationship studies. *Anal Chem* **62**:2323-2329 (1990).
30. Todeschini, R. & Gramatica, P. *The WHIM THEORY*: New 3D-molecular descriptors for QSAR in environmental modelling. *SAR QSAR Environ Res* **7**:89-115 (1997).
31. Om, A. & Kim, J. H. A quantitative structure-activity relationship model for radical scavenging activity of flavonoids. *J Med Food* **11**:29-37 (2008).
32. Khlebnikov, A. I. *et al.* Improved quantitative structure-activity relationship models to predict antioxidant activity of flavonoids in chemical, enzymatic, and cellular systems. *Bioorg Med Chem* **15**:1749-1770 (2007).