Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 44 (2009) 400–408
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://www.elsevier.com/locate/ejmech

Original article

QSAR study of antioxidant activity of wine polyphenols

Vesna Rastija [a,*], Marica Medić-Šarić [b]

[a] Faculty of Agriculture, Josip Juraj Strossmayer University of Osijek, Trg Sv. Trojstva 3, Osijek 31000, Croatia
[b] Department of Medicinal Chemistry, Faculty of Pharmacy and Biochemistry, University of Zagreb, A. Kovačića 1, Zagreb 10 000, Croatia

Received 21 December 2006; received in revised form 4 March 2008; accepted 4 March 2008
Available online 10 March 2008

Abstract

Quantitative structure activity relationships (QSAR) were obtained describing the antioxidant activity of the main pharmacologically active polyphenols of wine, using molecular properties and descriptors derived from the 2D and 3D representations of molecules. The best models for the prediction of the ability to scavenge the ABTS radical cation were obtained by polynomial regression analysis using zero-order connectivity index and molar refractivity. Statistically, significant models for lipid peroxidation inhibiting effects of flavonoids were obtained by polynomial and multiple regression using lipophilicity, Balaban index, Balaban-type index and 3D GETAWAY descriptor. The 3D descriptors possess the ability for discrimination of stereoisomers, like catechin and epicatechin. We demonstrated that a size and shape of molecules, as well as steric properties, play an important role in the antioxidant activity of polyphenols.

Keywords: Polyphenols; Wine; Antioxidant activity; QSAR

1. Introduction

Polyphenols are secondary plant metabolites, widely distributed in plants and foods of plant origin [1]. Wines contain a wide range of polyphenols that include phenolic acids, the trihydroxystilbene resveratrol, flavonols (e.g. quercetin and myricetin), flavan-3-ols (e.g. catechin and epicatechin), as well as polymers of the latter, defined as procyanidins and anthocyanins that are the pigments responsible for the colour of red wines. These dietary compounds have been reported to have multiple biological activities including vasodilatatory, anti-inflammatory, anticancerogenic, antiviral and antibacterial effects, and they are responsible for the healthy effects of moderate wine consumption. Health benefits of polyphenols arise from the antioxidative effects of these phytochemicals, which are based on their ability to scavenge different free radicals leading to the protection of biological molecules against oxidation [2,3]. Various measurements have been employed to determine the antioxidant activity of polyphenols, such as the determination of free radical scavenging activity against the active oxygen species (for example, peroxyl radical and hydroxyl radical) [4,5] and enzymatic or nonenzymatic inhibition activity against lipid peroxidation [6]. A commonly used method is the determination of the ability of hydrogen-donating antioxidants to scavenge the 2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) radical cation ($ABTS^{\bullet+}$) which is expressed in Trolox equivalent antioxidant capacity (TEAC) [7]. TEAC is defined as the concentration of Trolox solution with equivalent antioxidant potential to a 1 mmol L^{-1} concentration of compound under investigation and for an individual antioxidant it represents the number of the $ABTS^{\bullet+}$ radical-cations consumed per molecule of antioxidant [6,8]. Lipid peroxidation inhibitory effect is frequently monitored by the detection of the formation of lipid peroxidation products (such as malonodialdehyde) using thiobarbituric acid (TBA) [9].

The antioxidant activity of polyphenols can largely be predicted on the basis of their chemical structure. Antioxidant and antiradical activities of flavonoids are related to the presence of two neighboring hydroxyl groups on the B-ring, the number of free hydroxyl groups, a C2–C3 double bond in the C-ring, or the presence of a 3-hydroxyl group. For the antioxidant activity

* Corresponding author. Tel.: +385 31 224 226; fax: +385 31 207 017.
E-mail addresses: vrastija@pfos.hr (V. Rastija), bebamms@pharma.hr (M. Medić-Šarić).

0223-5234/$ - see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2008.03.001

of phenolic acids and their esters an important parameter is the number of hydroxyl groups in the molecule and the presence of steric hindrance [8,10]. Several structure–activity studies (SAR) of flavonoids and phenolic acids as antioxidants have been published, however, most of them have only been descriptive and they assayed flavonoids and phenolic acids separately [8,11,12]. Moreover, attention has been paid to the relationship between the structure and inhibition of lipid peroxidation. Recently, quantitative structure–activity relationships (QSAR) have been used extensively to develop models in order to estimate and predict antioxidant activity of flavonoids and phenolic acids using descriptors derived from chemical structure and various physicochemical parameters calculated with adequate programs [13–16]. There are many numerical descriptors available in chemistry, including topological indices, 3D descriptors, quantum chemical indices and physicochemical parameters associated with the molecular structure in QSAR researches [17–19]. Among them, topological indices are the most popular since they can effectively characterize molecular size, branching and variation in molecular shapes [20,21]. Hansch et al. [22] were the first who used electronic and steric parameters together with hydrophobic parameters based on *n*-octanol/water partition coefficient (log *P*) for developing QSAR models.

The main aim of this study is to develop a quantitative structure–activity relationship between "two-dimensional" (2D) topological indices, "three-dimensional" (3D) descriptors, calculated physicochemical parameters and antioxidant activity of polyphenols that are usually present in wine. Experimental data used in this study were taken from literature and consisted of antioxidant activity of 10 wine polyphenols determined using ABTS test expressed in TEAC (per mmol L^{-1}) values [7] and lipid peroxidation inhibitory effects of 8 flavonoids expressed as the concentration for 50% inhibition of lipid peroxidation (IC_{50}/μmol L^{-1}) [9]. The compounds and their activities reported in the literature were limited to polyphenols that had been identified previously in wine by different analytical methods (TLC, HPLC) [23,24]. The structures of all polyphenols used in this work are shown in Fig. 1.

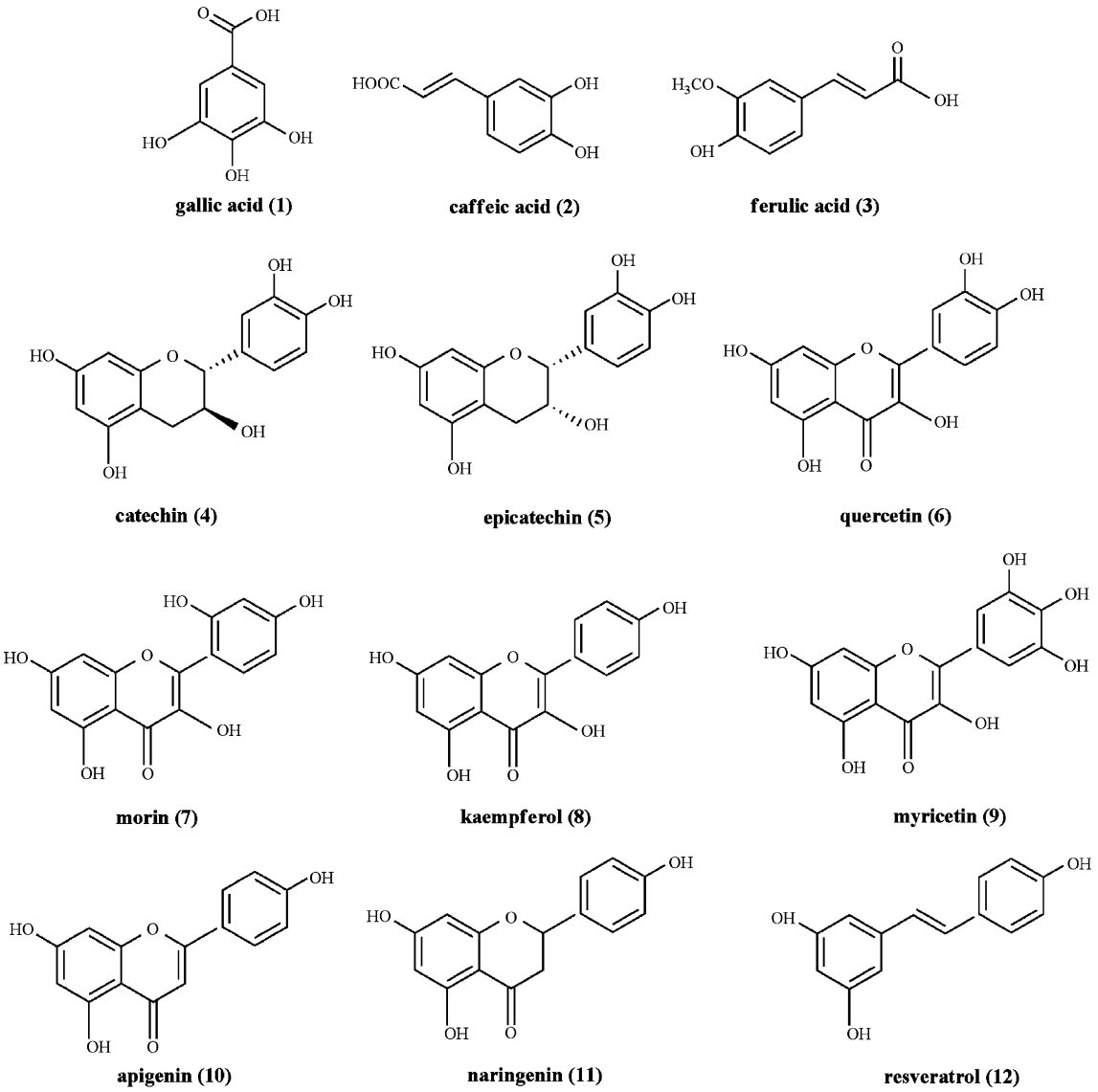

Fig. 1. Structural formulas of polyphenols used in the present study.

We have investigated linear, polynomial and multiple linear relationships between topological indices: Wiener index (W); connectivity index (χ); Balaban index (J); Balaban-type indices from atomic number-, from mass-, from van der Waals-, from electronegativity-, from polarizability-weighted distance matrix (J_Z, J_m, J_v, J_e, J_p); information-theoretic index (I) and Schultz index (MTI)), selected properties (molecular weight (M); n-octanol/water partition coefficient ($\log P$); van der Waals volume (V_w); molar refractivity (MR); polar surface area (PSA) of polyphenols and values of ABTS test and lipid peroxidation inhibitory effects. Calculated topological indices are shown in Table 1 and physicochemical properties for the studied compounds are shown in Table 2. Since some of the studied compounds contain two stereoisomers with different activities (catechin and epicatechin), it was necessary to use molecular descriptors calculated from three-dimensional representation of molecules [25]. Four groups of 3D descriptors have been used to generate QSAR models: geometrical; GETAWAY (geometry, topology, and atom weights assembly); 3D-MoRSE and RDF (radial distribution function) descriptors. Values of 3D descriptors selected within four different groups by best-subset regression are given in Table 3.

The present work should provide a better understanding of the beneficial effects of wine on human health, since the absorption and metabolism of polyphenols is influenced by their solubility and chemical structure [26].

2. Results and discussion

2.1. Quantitative structure–activity relationship of the antioxidant potential of polyphenols against radicals generated in the aqueous phase

QSAR models for the prediction of antioxidant activity of polyphenols determined by using $ABTS^{\bullet+}$ radical scavenging assay were developed. Models obtained by the linear and multivariate regression analysis were not statistically satisfactory. Best models were obtained by polynomial regression analysis with the zero-order connectivity index ($^0\chi$) and with molar refractivity (MR). Experimental and calculated TEAC values using the obtained models and associated 95% confidence intervals are given in Table 4.

Table 1
Topological indices calculated for polyphenols used in the present study

Comp. no.	N(G)	W(G)	$^0\chi$(G)	$^1\chi$(G)	J(G)	J_m(G)	I(G)	MTI
1	12	246	5.696	2.925	3.371	3.541	2.774	1330
2	13	417	6.336	3.201	3.079	3.032	3.453	2890
3	14	444	7.520	3.873	3.018	3.118	3.309	2398
4	21	1057	10.70	6.269	2.109	2.187	3.353	5828
5	21	1057	10.70	6.269	2.109	2.187	3.353	5828
6	22	1252	10.74	6.026	2.269	2.353	3.417	7174
7	22	1226	10.74	6.026	2.290	2.383	3.381	7042
8	21	1111	10.41	5.905	2.260	2.333	3.392	6488
9	23	1403	11.08	6.147	2.303	2.384	3.434	7874
10	20	1006	10.08	5.778	2.162	2.238	3.407	5992
11	20	952	10.29	6.033	2.105	2.177	3.341	5524
12	17	795	8.921	5.076	2.179	2.267	3.573	4689

W(G) = Wiener index, $^0\chi$(G) = zero-order connectivity index, $^1\chi$(G) = first-order connectivity index, J(G) = Balaban index, J_m(G) = Balaban-type index from mass-weighted distance matrix, I(G) = information-theoretic index, MTI = Schultz index.

Table 2
Physicochemical properties for studied compounds

Comp. no.	M	ALOGPS	CLOGP	MLOGP	MR	PSA	V_w
1	170.1	1.17	0.425	0.029	37.90	17.07	1.405
2	180.2	1.67	0.975	1.116	43.60	17.07	1.569
3	194.2	1.58	1.421	1.419	47.02	26.30	1.702
4	290.3	1.02	0.534	0.246	70.24	9.23	2.688
5	290.3	1.02	0.534	0.246	70.24	9.23	2.489
6	304.3	1.07	0.771	−0.746	71.22	26.30	2.476
7	302.2	2.23	1.134	−0.746	71.22	26.30	2.476
8	286.2	1.99	1.368	0.014	69.52	26.30	2.374
9	318.2	1.66	0.837	−1.494	72.91	26.30	2.578
10	270.3	3.07	2.905	0.785	67.83	26.30	2.315
11	272.3	2.47	2.445	0.900	67.58	26.30	2.303
12	228.3	2.57	2.833	2.631	63.67	0.00	2.342

M = molecular weight, $\log P$ (ALOGP, CLOGP, MLOGP) = partition coefficients, MR = molar refractivity, PSA = polar surface area, V_w = van der Waals volume.

Among all the topological indices used, the zero-order connectivity index was found to correlate well with TEAC values of analyzed polyphenols. TEAC values can be expressed through the following equation:

$$\mathrm{TEAC} = 13.66(\pm 2.21) - 3.092(\pm 0.546)^0\chi + 0.178(\pm 0.032)(^0\chi)^2 \quad (1)$$
$$n = 10;\ R^2 = 0.833;\ S = 0.186;\ F = 17.47.$$

Since the connectivity index is closely related to the size of the molecule, the negative coefficient of $^0\chi$ in the model indicates that a bigger size of molecules reduces the ability to scavenge the $ABTS^{\bullet+}$ radical cation. Valence connectivity indices encode important structural features such as size, branching, unsaturation, cyclicity and heteroatom content, therefore are especially suitable for the QSAR of the different biological activities of flavonoids [27,28].

Polynomial regression resulted in the following statistically significant model using molar refractivity (MR) as descriptor:

$$\mathrm{TEAC} = 16.45(\pm 2.59) - 0.574(\pm 0.097)\mathrm{MR} + 0.005(\pm 0.001)\mathrm{MR}^2 \quad (2)$$
$$n = 10;\ R^2 = 0.845;\ S = 0.180;\ F = 19.05.$$

In order to confirm the found relationship between molar refractivity and antioxidant activity of polyphenols, TEAC values have been correlated with predicted TEAC values derived from Eq. (2). The obtained statistical parameters are: $R^2 = 0.845$; $S = 0.165$; $F = 43.54$. A scatter plot of the experimental data ($TEAC_{obs.}$) versus the predicted TEAC values ($TEAC_{calc.}$) using Eq. (2) is presented in Fig. 2.

Molar refractivity is related to the size and polarizability of substituents. A negative coefficient (−0.574) in Eq. (2) reflects stereochemical hindrance between substituents and free radical, and as a consequence, compounds with a smaller MR value possess higher activity [29]. This is specially expressed

Table 3
Values of 3D descriptors selected within four different groups (geometrical, GETAWAY, 3D-MoRSE, RDF) by best-subset regression

Comp. no.	^{3D}W	G(O···O)	DISPe	QXXe	Mor02u	Mor22u	Mor26u	H_7(u)	H_5(m)	HATS$_7$(m)	H_7(v)	H_7(p)	R_7^+(m)	RDF025u	RDF055e
1	555.9	45.05	0.074	48.96	14.37	0.211	−0.123	0.017	0.004	0.017	0.002	0.003	0.004	9.79	0.104
2	803.7	27.44	0.239	55.28	16.21	0.182	0.017	0.000	0.001	0.148	0.000	0.000	0.044	10.75	0.511
3	1129	27.44	0.237	78.16	18.65	−0.099	0.056	0.083	0.004	0.111	0.007	0.012	0.029	11.44	0.434
4	2990	98.28	0.095	124.4	27.06	0.299	0.146	0.197	0.127	0.135	0.022	0.031	0.027	24.53	0.893
5	2978	96.90	0.070	120.1	26.25	0.329	0.172	0.315	0.151	0.148	0.039	0.052	0.036	26.08	0.221
6	2575	129.9	0.154	122.0	26.46	0.189	−0.175	0.072	0.149	0.206	0.007	0.011	0.034	20.07	0.330
7	2529	120.8	0.201	120.2	24.43	−0.024	−0.073	0.102	0.155	0.176	0.013	0.018	0.027	19.76	0.468
8	2374	88.86	0.200	109.1	25.91	0.188	−0.077	0.089	0.145	0.157	0.009	0.013	0.027	20.27	0.553
9	2785	175.0	0.169	132.1	26.30	0.211	−0.192	0.072	0.163	0.230	0.007	0.010	0.033	19.80	0.829
10	2216	63.60	0.237	92.87	26.80	0.235	0.038	0.192	0.115	0.085	0.042	0.052	0.016	21.04	0.917
11	2461	63.38	0.219	101.6	26.12	0.271	0.017	0.121	0.050	0.124	0.012	0.018	0.018	22.38	0.158
12	2072	26.44	0.142	72.98	27.00	0.063	0.086	0.292	0.063	0.066	0.060	0.075	0.016	21.06	0.436

Molecular descriptors: ^{3D}W = 3D-Wiener number; G(O···O) = sum of geometrical distances between O···O; DISPe and QXXe = d COMMA2 and Qxx value/weighted by atomic Sanderson electronegativities; Mor02u, Mor22u and Mor26u = 3D-MoRSE/unweighted; H_7(u), H_5(m), H_7(v), H_7(p) and HATS$_7$(m) = H-GETAWEY; R_7^+(m) = R-GETAWEY; RDF025u and RDF055e = Radial Distribution Function/unweighted and weighted by atomic Sanderson electronegativities.

for phenolic acids. Verma and Hansch [30] studied the anti-HIV activity of caffeic acid and its derivates and found that the activity of investigated compounds increased with decreasing MR values. The substituent that causes a steric hindrance is a carboxylate group in benzoic acid because it has negative influence on the H-donating abilities of hydroxy benzoates. Hence, hydroxybenzoic acids are less effective than hydroxycinnamic acids with equal positions of hydroxyl groups in the ring [8]. The radical scavenging activity of flavonoids depends on molecular structure and the substitution pattern of free hydroxyl groups on the flavonoid skeleton [14].

The application of 3D descriptors for prediction of TEAC values by simple linear and polynomial regression did not result in statistically significant models, so we tried to build a better model with multiple regression. Independent variables for multiple regression were selected using the best-subset method. In order to obtain quality models as simple as possible, with as small a number of independent variables as possible with respect to the number of compounds used for model generation ($n = 10$), only two variables were included in the model. However, the best possible model obtained with two 3D descriptors (RDF055e and HATS$_7$(m)) had a poor statistics ($R^2 = 0.76$; $S = 0.220$; $F = 10.97$).

Exact comparison of our results with similar ones reported earlier is not possible because we used a different set of compounds and different parameters to obtain models for the prediction of the ability to scavenge the ABTS$^{\bullet+}$ radical cation.

2.2. Quantitative structure–activity relationship for the inhibitory effects of flavonoids

Models for the prediction of inhibitory effects (IC_{50}) with good statistical parameters were obtained by polynomial and multiple regressions. Experimental and calculated TEAC values using obtained models and associated 95% confidence intervals are shown in Table 5.

Among the three different computer programs for the calculation of partition coefficients (CLOGP, ALOGPS and MLOGP), the best results were achieved using the MLOGP program. The relation between the lipid peroxidation inhibitory effects of flavonoids and lipophilicity can be described by the following equation:

$$IC_{50} = 28.69(\pm 7.06) + 53.26(\pm 7.80)\text{MLOGP} + 30.62(\pm 8.95)\text{MLOGP}^2 \quad (3)$$
$$n = 8;\ R^2 = 0.904;\ S = 11.00;\ F = 23.55.$$

Table 4
Experimental (TEAC$_{obs.}$) and calculated (TEAC$_{calc.}$) TEAC values using models expressed by Eqs. (1) and (2), with residual (resid.) and associated 95% confidence intervals (cnf. int.)

Comp. no.	TEAC$_{obs.}$[a]/mmol L^{-1}	Model number 1			Model number 2		
		TEAC$_{calc.}$	Resid.	95% Cnf. int.	TEAC$_{calc.}$	Resid.	95% cnf. int.
1	1.98 ± 0.01	1.78	0.20	1.37, 2.23	1.89	0.09	1.45, 2.32
2	1.01 ± 0.00	1.18	0.17	0.89, 1.48	0.94	0.07	0.66, 1.22
3	0.23 ± 0.00	0.43	−0.20	0.10, 0.76	0.52	−0.29	0.20, 0.85
4	0.57 ± 0.01	0.87	−0.30	0.66, 1.07	0.81	−0.24	0.63, 1.00
5	0.99 ± 0.00	0.87	0.12	0.66, 1.07	0.81	0.18	0.63, 1.00
6	1.14 ± 0.09	0.90	0.24	0.69, 1.11	0.94	0.20	0.74, 1.15
7	0.83 ± 0.00	0.90	−0.07	0.69, 1.11	0.94	−0.11	0.74, 1.15
8	0.79 ± 0.08	0.68	0.11	0.49, 0.88	0.72	0.07	0.54, 0.91
9	1.02 ± 0.03	1.16	−0.14	0.889, 1.43	1.19	−0.17	0.92, 1.46
12	0.40 ± 0.01	0.18	0.22	−0.17, 0.53	0.18	0.22	−0.13, 0.50

[a] Experimental values were taken from Ref. [7]. Values represent means ± standard deviations of the duplicate of 5 or 6 concentrations within the linear interval.

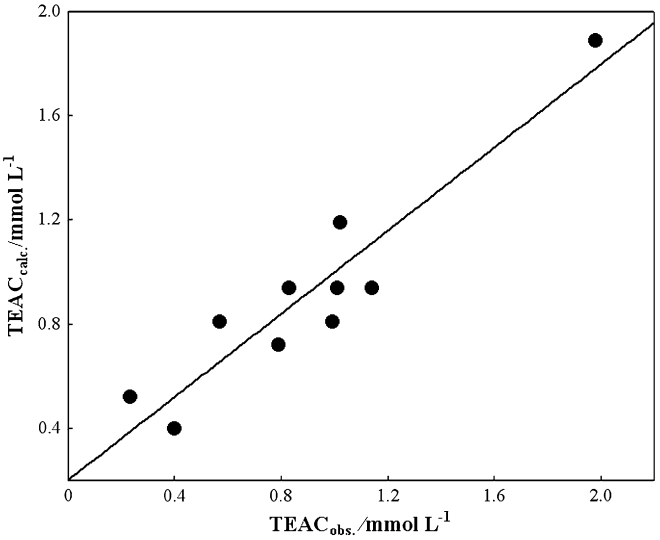

Fig. 2. Relationship of observed ($TEAC_{obs.}$) and calculated ($TEAC_{calc.}$) TEAC values derived from Eq. (2).

Yang et al. [9] obtained a QSAR model to describe the antioxidant activity of 23 flavonoids by multiparameter equation with the half-wave potential of the first oxidation step and calculated log P values as descriptors ($R = 0.852$). Our model includes a smaller set of substances, but it has a higher correlation coefficient ($R = 0.951$; $R^2 = 0.904$) and contains only one descriptor. The positive term of MLOGP in Eq. (3) suggests that the antioxidant activity of the investigated compounds increases with increasing lipophilicity. It means that flavonoids with higher lipophilicity have a better ability of interaction with lipid membranes and intercorporation into cells, thus exhibiting higher protective effect against lipid peroxidation.

The recent QSAR study [31] of inhibitory effect of phenolic antioxidants on lipid peroxidation indicates that the partition coefficient of phenolic compounds may define the antioxidant activities in lipophilic phase. QSAR model was obtained by multiple regression analysis with the 7 descriptors on the set of 15 compounds. Erlejman et al. [32] suggest that interaction of flavonoids with the polar head groups of phospholipids is the main factor that contributes to their antioxidant activity in liposomal membranes and that the increased lipophilicity of compounds is reflected on good antioxidant activity. Nevertheless, in a QSAR study of the same activity of 12 flavonoids with regard to the lipophilicity (CLOGP) Rackova et al. [15] obtained a poor model with $R = -0.580$. The model did not improve with the use of experimental values of log P. In the present study, we used the values of the partition coefficient calculated by a different program that is based on Moriguchi method (MLOGP). The polynomial regression with CLOGP as variable, gave a model with a smaller correlation coefficient ($R = 0.910$). Besides, our investigation was performed on a different set of compounds than the study of Rackova group [15].

The inhibition activity of studied compounds was also found to have a good polynomial correlation with the van der Waals volume, which is expressed by Eq. (4):

$$\mathrm{IC}_{50} = 1.112\times10^{4}\left(\pm2.381\times10^{3}\right) - 8.789\times10^{3} \times\left(\pm1.920\times10^{3}\right)V_{\mathrm{w}} + 1.738\times10^{3}\left(\pm3.87\times10^{2}\right)V_{\mathrm{w}}^{2} \quad (4)$$
$$n = 8;\ R^2 = 0.858;\ S = 13.40;\ F = 15.09.$$

Previous structure activity relationship (SAR) studies have shown that 3′-C hydroxyl group determines the inhibitory effects of flavonoids in lipophilic phases as well as a 2,3-double bond in conjugation with 4-oxo and 3-hydroxyl groups [8,33]. These structural features of flavonoids are essential for the stabilization of phenoxyl radical by electron delocalization across the aromatic ring.

The van der Waals volume is widely used as the parameter for the description of size and steric contribution to the behavior of biologically active compounds. Therefore, a QSAR model for the antibacterial activity of 10 benzaldehydes obtained by multiple linear regressions includes van der Waals volume as a steric parameter together with some electronic–steric parameters. Molar refractivity could be used instead of van der Waals volume in the same model [34]. Beside the partition coefficient (log P) of phenolic compounds, the number of physicochemical parameters such as the enthalpy of phenoxyl radical formation (ΔH) and hydration energy (E_{HYDR}) may define the antioxidant activities in the lipophilic phase [15,31]. However, there are no literature data for the contribution of van der Waals volume to the lipid peroxidation inhibitory effects of flavonoids. The present study indicates that this

Table 5
Experimental (obs.) and calculated (calc.) IC_{50} values using models expressed by Eqs. (3)–(6) with residual (resid.) and associated 95% confidence intervals (cnf. int.)

Comp. no.	$IC_{50obs.}$[a]/ mmol L^{-1}	Model number 3			Model number 4			Model number 5			Model number 6		
		$IC_{50calc.}$	Resid.	95% Cnf. int.	$IC_{50calc.}$	Resid.	95% Cnf. int.	$IC_{50calc.}$	Resid.	95% Cnf. int.	$IC_{50calc.}$	Resid.	95% Cnf. int.
4	51.0	43.65	7.36	28.09, 59.20	53.03	−2.03	13.53, 92.52	37.99	13.01	20.41, 55.57	42.90	8.10	23.60, 60.20
5	25.0	43.65	−18.65	28.09, 59.20	11.15	13.85	−9.80, 32.09	37.99	−12.99	20.41, 55.57	23.27	1.73	8.49, 38.06
6	8.50	6.00	2.50	−11.78, 23.78	13.20	−4.70	−7.544, 33.95	10.51	−2.01	−2.05, 23.07	21.11	−12.61	9.52, 32.70
7	23.0	6.00	17.00	−11.78, 23.78	13.20	9.80	−7.54, 33.95	17.13	5.87	3.74, 30.53	18.94	4.06	7.87, 30.02
8	19.0	29.44	−10.44	12.25, 46.63	49.72	−30.72	31.36, 68.08	30.87	−11.87	19.49, 42.26	16.78	2.22	3.23, 30.33
9	10.5	17.47	−6.97	−14.40, 49.34	12.86	−2.36	−8.23, 33.95	5.15	5.35	−9.58, 19.88	14.62	−4.12	−3.19, 32.42
10	100.0	89.37	10.63	68.59, 10.16	87.35	12.65	61.37, 113.3	99.58	0.42	80.06, 119.1	110.6	−10.60	90.33, 130.8
11	100.0	101.4	−1.4	76.45, 26.41	96.49	3.51	67.45, 125.5	97.79	2.21	80.55, 115.0	88.79	11.21	71.16, 106.4

[a] Experimental values were taken from Ref. [9].

parameter is quite satisfactory for the prediction of antioxidant activities of flavonoids.

For the prediction of IC_{50} a multiple regression was also applied. A statistically significant model is obtained by employing CLOGP and Balaban index (J):

$$\begin{aligned} IC_{50} &= 480.1(\pm 103.6) - 217.4(\pm 46.4)J \\ &\quad +30.84(\pm 4.54)\mathrm{CLOGP} \\ n &= 8;\ R^2 = 0.945;\ S = 8.290;\ F = 43.33. \end{aligned} \tag{5}$$

The best model that includes 3D descriptor H_7(p) in combination with Balaban-type index from mass-weighted distance matrix (J_m) was chosen by the best-subset regression:

$$\begin{aligned} IC_{50} &= 320.3(\pm 247.9) + 17.46(\pm 1.92)J_m \\ &\quad +19.63(\pm 2.59)H_7(\mathrm{p}) \\ n &= 8;\ R^2 = 0.95;\ S = 7.87;\ F = 47.66. \end{aligned} \tag{6}$$

The obtained models demonstrate the significance of Balaban and Balaban-type indices for the modeling of biological activities of flavonoids. The Balaban index is a good descriptor for the shape of the molecule. Its large and negative coefficient in model 5 indicates that a bigger size and high branching of molecules reduce the lipid peroxidation inhibitory effect of flavonoids. On the contrary, the positive sign of Balaban-type index from mass-weighted distance matrix (J_m) in model 6 indicates that molecules with larger mass, flavonoid molecules with more hydroxyl groups, are better inhibitors of lipid peroxidation. Despite the fact that examples of QSAR studies based on Balaban index and Balaban-type indices are rare in literature, a few recent studies [35,36] and the present study show that this index can be successfully used for this purpose.

H_7(p) descriptor belongs to the group of H-GETAWAY descriptors that have been calculated from the molecular influence matrix H. These descriptors are sensitive to significant conformational changes and to the bond lengths that account for atom types and bond multiplicity. H_7(p) is an autocorrelation descriptor calculated for 3D-spatial molecular geometry based on *lag* (topological distance) and weighted by atomic polarizabilities [37]. Considering that the H_7(p) descriptor has a positive coefficient in Eq. (6), we may conclude that enhanced values of polarizability are favourable for lipid peroxidation inhibitory effects of flavonoids. GETAWAY descriptors have shown great potential as powerful variables in QSAR modeling of different biological activities because they encode information about molecular shape, size and atom distribution [38,39]. Considering that GETAWAY descriptors were developed recently, there has been no literature evidence on their application in QSAR study of antioxidant activity until now.

3. Conclusion

Simple and significant QSAR models have been obtained by polynomial regression for the prediction of antioxidant activity of wine polyphenols. Zero-order connectivity index ($^0\chi$) and molar refractivity (MR) are found to be useful parameters for modeling free radical scavenging activity of polyphenols belonging to different groups (phenolic acids and flavonoids – flavans, flavonols and stilbene). The results of modeling lipid peroxidation inhibitory effects of flavonoids indicate that lipophicity and van der Waals volume (V_w) are significant molecular descriptors for the prediction of antioxidant activity of flavonoids in the lipophilic phase. We demonstrated that the antioxidant activity of polyphenols, as hydrogen-donating free radical scavengers, is closely related to their chemical structure, especially with the number and the arrangement of free hydroxyl groups on the flavonoid skeleton, or on the phenol ring of phenolic acids. Moreover, considerable effect on antioxidant activity of polyphenols has a size, mass and shape of molecules, as well as, steric properties, which have been described by Balaban index, Balaban-type index from mass-weighted distance matrix, molar refractivity, van der Waals volume, as well as the descriptor derived from three-dimensional molecular structure, GETAWAY descriptor H_7(p). Although 3D descriptors possess the ability to discriminate between stereoisomers, some models with 2D topological descriptors are of better quality.

All the above-mentioned parameters could be used for future QSAR investigations of wine polyphenols, extended to a larger number of substances, including anthocyanins and dimeric procyanidins as the most active polyphenols in wine [40].

4. Calculation

4.1. Antioxidant activities

The antioxidant activities (TEAC and IC_{50}) for the set of polyphenols used in the present study were taken from literature [7,9].

4.2. Calculation of topological indices

Five topological indices (Wiener index, connectivity index, Balaban index, information-theoretic index and Schultz index) [41,42] used in our work were calculated using TAM [43] program. The total number of vertices, N(G), in the molecular graph was considered as a topological parameter. It is identical to the number of atoms in the hydrogen-depleted molecular structure. The calculation of Balaban-type indices is described in Section 4.3.

4.2.1. Wiener index, W(G)

The Wiener index W of a structure G can be simply obtained from the distance matrix D of the corresponding hydrogen-depleted chemical graph G as half-sum of the elements of D [44]:

$$W(\mathrm{G}) = \frac{1}{2}\sum_{i,j}(D)_{ij} \tag{7}$$

where $(\mathrm{D})_{ij}$ represent off-diagonal elements of D(G) which stand for the shortest distance in terms of the number of bonds between atoms i and j in G. The distance i–j can also be weighted in the case of heteroatomic systems [45]. In the

case studied, all the structures considered are depicted by weighted graphs.

4.2.2. Valence connectivity index, $\chi^{v}(G)$

The connectivity index, when it includes heteroatoms, is called valence connectivity index and is denoted by $\chi^{v}(G)$. This index is defined as [46]

$$\chi^{v}(G) = \sum_{i,j} [\delta(i)\delta(j)]^{-1/2} \tag{8}$$

where $\delta(i)$ and $\delta(j)$ are weights (valence delta values) of vertices (atoms) i and j making up i–j edge (bond) in a vertex-weighted graph G_{vw}.

Depending on the number of atoms included in computation, the connectivity index may be extended on a set of indices:

(a) The zero-order connectivity index ($^{0}\chi$) is defined as

$$^{0}\chi = \sum_{i=1}^{N} [a(i)]^{-1/2} \tag{9}$$

where N is the total number of vertices in G.

(b) The first-order connectivity index ($^{1}\chi$) is defined as

$$^{1}\chi = \sum_{i,j=1}^{E} [a(i)a(j)]^{-1/2} \tag{10}$$

where E is the total number of edges in G.

4.2.3. Balaban index, J(G)

Balaban index can be described as the average distance sum connectivity [47]:

$$J(G) = \frac{E}{\mu + 1} \sum_{\text{edges}} \left(ds_i\, ds_j\right)^{-1/2} \tag{11}$$

where E is the number of edges in G, μ is the cyclomatic number of G and ds_i ($i = 1, 2, \ldots, N$; N is the number of vertices in G) is a distance sum.

The distance sum, ds_i, for a vertex i represents the sum of all entries in the corresponding row (or column) of the distance matrix D:

$$ds_i = \sum_{j=1}^{N} (D)_{ij} \tag{12}$$

Balaban-type indices are obtained by weighting the contributions of atoms and bonds with parameters based on atomic number (J_Z), mass (J_m), van der Waals (J_v), electronegativity (J_e) and polarizability (J_p) [42].

4.2.4. Information-theoretic index, I(G)

The information-theoretic index I(G) was calculated by the application of information-theoretic formalism on the chemical graph, by means of the total information content or by the modification of Shannon's relation [48]:

$$I(G) = -\sum_{i=1}^{n} \frac{2N_1}{N(N-1)} \log_2 \frac{2N_1}{N(N-1)} \tag{13}$$

where n is the number of different sets of elements, N_i is the number of elements in the i-th set of elements and the sum is over all sets of elements.

4.2.5. Schultz index, MTI

This index has been introduced by Schultz [49] and is denoted by MTI (molecular topological index). The Schultz index is based on the adjacency ($N \times N$) matrix (A), the distance ($N \times N$) matrix (D), and the valency ($I \times N$) matrix (v) of an alkane tree. The sum of elements e_i ($I = 1,2,\ldots N$) of the row matrix $v[A + D]$ gives the Schultz index:

$$\text{MTI} = \sum_{i=1}^{N} e_i \tag{14}$$

4.3. Calculation of molecular descriptors derived from 3D structure and Balaban-type indices

The 2D structures of all compounds were drawn using ISIS/Draw 2.3 [50]. The 3D structures of 12 polyphenols were optimized using the quantum chemical semi-empirical method AM1 applying the HyperChem 8.0 Evaluation software package [51]. The Balaban-type indices (J_Z, J_m, J_v, J_e, J_p) and the 3D molecular descriptors used in this study have been calculated applying the on-line software Parameter Client (PCLIENT) [52]. PCLIENT is an extension of E-Dragon, on-line version of Dragon program and makes it possible to generate more than 3000 indices. SMILES notations, created on-line by Demo-CONVERT Interactively [53], were used as chemical structure input for PCCLIENT program.

We calculated descriptors of the geometrical, GETAWAY, 3D-MoRSE and RDF series to generate the QSAR models. Different kinds of conformationally dependent descriptors based on molecular geometry, belong to the group of geometrical descriptors. In this work we used a few geometrical descriptors: 3D-Wiener (^{3D}W), 3D-Balaban (^{3D}J), 3D-Harary index (^{3D}H), average distance/distance degree (ADDD) and spherosity (SPH) [54]. GETAWAY descriptors are derived by the use of geometrical information given by the molecular matrix (H-GETAWAY) or the combination of this information with geometric interatomic distances in molecules (R-GETAWAY) [37]. The 3D-MoRSE (molecule representation of structures based on electron diffraction) descriptors are developed from equation used in electron diffraction studies and represent the 3D structure of molecules by a fixed (constant) number of values [55]. The RDF descriptors are obtained by radial basis functions centered on different interatomic distances. This function can be interpreted as a probability to find an atom in spherical volume of radius r [19,56].

4.4. Calculation of physicochemical properties of polyphenols

Molecular lipophilicity was calculated using three computer programs based on different theoretical approaches.

ALOGPS was developed using the Efficient Partition Algorithm [57] and an associative neural network (ASNN) approach [58]. This method combines electronic and topological characters to predict lipophilicity of the analyzed molecules.

MLOGP program is based on the method developed by Moriguchi et al. [59]. The method begins with a straightforward counting of lipophilic atoms (all carbons and halogens with a multiplier rule for normalizing their contributions) and hydrophilic atoms (all nitrogen and oxygen atoms). The Moriguchi method then applies 11 correction factors, four that increase the hydrophobicity, and seven that increase the lipophilicity. ALOGP and MLOGP values were calculated using Parameter Client.

Mannhold and van de Waterbeemd [60] described the CLOGP program as substructure approaches where the final $\log K_{ow}$ is determined by summing the single-atom or fragment contributions. The calculation result is accompanied by the picture of chemical structure as generated by the DEPICT algorithm. Software-predicted lipophilicity of the compounds was calculated on-line with the CLOGP program that is available at the Daylight Website [61].

Molecular weight (M/g mol^{-1}), molar refractivity (MR/m^3 mol^{-1}) and polar surface area (PSA/Å^2) were also calculated using PCLIENT program. The molar refractivity is a constitutive–additive property used in QSPR/QSAR and in the radiation of infinite wavelength represents the real volume of the molecules contained in 1 mol of the substance. MR is calculated from the Lorentz–Lorentz equation and is described as follows:

$$\mathrm{MR} = \frac{(n^2-1)}{(n^2+2)}\left(\frac{M}{d}\right) \tag{15}$$

where M is the molecular weight, and d is the density of the substance [62].

Polar surface area is defined as the area of its van der Waals surface that arises from oxygen or nitrogen atoms or hydrogen atoms attached to the oxygen or nitrogen atoms [63]. van der Waals volume (V_w/Å^3) was calculated according to the method described by Moriguchi et al. [64]. The volumes of spheres are calculated using the atomic radii and the overlapping volumes are subtracted.

4.5. Statistical analysis

The statistical analysis was performed using STATISTICA 6.0 [65].

Relationship between 2D and 3D descriptors, physicochemical properties and antioxidant activities of polyphenols was investigated by simple linear, polynomial (quadratic) and multiple regression analysis. The selection of predictor variables for multiple regression was performed by best-subset method in which regression equation is fitted to every subset of independent variables. The criterion used to determine "best" is based on the R^2 values of analyzed models.

To test the quality and accuracy of derived models, the following statistical parameters were used: squared correlation coefficient (R^2), standard deviation of regression (S) and Fisher ration values (F). Standard deviation of regression (S) was calculated using the following equation:

$$S = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i')^2}{n}} \tag{16}$$

where n denotes the total number of cases (molecules); y_i and y_i' denote the calculated value and value obtained by regression model. The best possible QSAR models, which are presented in this paper, were selected on the basis of the highest correlation coefficients and F-test and lowest standard deviations.

For the best models we also gave a 95% confidence interval for predicted values of each compound (Tables 4 and 5). This interval specifies the ranges of values that will contain the value of the regression function with a pre-specified probability.

Acknowledgments

This work was supported in part by Grants No. 006-0061117-1237 (M. M.-Š.) awarded by the Ministry of Science, Education and Sport of the Republic of Croatia. We are thankful to the reviewers for their constructive comments.

References

[1] C. Manach, A. Scalbert, C. Morand, C. Rémésy, L. Jiménez, Am. J. Clin. Nutr. 79 (2004) 727–747.
[2] J. Burns, P.T. Gardner, D. Matthews, G.G. Duthie, M.E.J. Lean, A. Crozier, J. Agric. Food Chem. 49 (2001) 5797–5808.
[3] M.A. Soobrattee, V.S. Neergheen, A. Luximon-Ramma, O.I. Aruoma, T. Bahorun, Mutat. Res. 579 (2005) 200–213.
[4] Y. Kono, K. Kobayashi, S. Tagawa, K. Adachi, A. Ueda, Y. Sawa, H. Shibata, Biochim. Biophys. Acta 1335 (1997) 335–342.
[5] Z. Cheng, J. Ren, Y. Li, W. Chang, Z. Chen, Bioorg. Med. Chem. 10 (2002) 4067–4073.
[6] V. Roginsky, E.A. Lissi, Food Chem. 92 (2005) 235–254.
[7] D. Villaño, M.S. Fernández-Pachón, A.M. Troncoso, M.C. García-Parrilla, Anal. Chim. Acta 538 (2005) 391–398.
[8] C.A. Rice-Evans, N.J. Miller, G. Paganga, Free Radic. Biol. Med. 20 (1996) 933–956.
[9] B. Yang, A. Kotani, K. Arai, F. Kusu, Anal. Sci. 17 (2001) 599–604.
[10] S. Burda, W. Oleszek, J. Agric. Food Chem. 49 (2001) 2774–2779.
[11] K.E. Heim, A.R. Tagliaferro, D.J. Bobilya, J. Nutr. Biochem. 13 (2002) 572–584.
[12] N. Nenadis, L.-F. Wang, M. Tsimidou, H.-Y. Zhang, J. Agric. Food Chem. 52 (2004) 4669–4674.
[13] E. Sergediené, K. Jönsson, H. Szymusiak, B. Tyrakowska, I.M.C.M. Rietjens, N. Čénas, FEBS Lett. 462 (1999) 393–396.
[14] D. Amić, D. Davidović-Amić, D. Bešlo, N. Trinajstić, Croat. Chem. Acta 76 (2003) 55–61.
[15] L. Rackova, S. Firakova, D. Kostalova, M. Stefek, E. Sturdik, M. Majekova, Bioorg. Med. Chem. 13 (2005) 6477–6488.

[16] O. Farkas, J. Jakus, K. Héberger, Molecules 9 (2004) 1079–1088.
[17] A. Thakur, M. Thakur, N. Kakani, A. Joshi, S. Thakur, A. Gupta, ARKIVOC xiv (2004) 36–43.
[18] M. Karelson, V.S. Lobanov, A.R. Katritzky, Chem. Rev. 96 (1996) 1027–1043.
[19] M. Vračko, J. Gasteiger, Internet Electron. J. Mol. Des. 1 (2002) 527–544.
[20] N. Trinajstić, Chemical Graph Theory, CRC Press, Boca Raton, 1992.
[21] A.M. Ponce, S.E. Blanco, A.S. Molina, R. García-Domenech, J. Gálvez, J. Chem. Inf. Comput. Sci. 40 (2000) 1039–1045.
[22] C. Hansch, A. Kurupa, R. Garg, H. Gao, Chem. Rev. 101 (2001) 619–672.
[23] V. Rastija, A. Mornar, I. Jasprica, G. Srečnik, M. Medić-Šarić, J. Planar Chromatogr. Mod. TLC 17 (2004) 26–31.
[24] S. Kallithraka, E. Tsoutsouras, E. Tzourou, P. Lanaridis, Food Chem. 99 (2006) 784–793.
[25] A.T. Balaban (Ed.), From Chemical Topology to Three-Dimensional Geometry, Plenum press, New York, 1997.
[26] I.E. Dreosti, Nutrition 16 (2000) 692–694.
[27] D. Amić, D. Davidović-Amić, A. Jurić, B. Lučić, N. Trinajstić, J. Chem. Inf. Comput. Sci. 35 (1995) 1034–1038.
[28] E. Estrada, J.A. Quincoces, G. Patlewicz, Mol. Divers. 8 (2004) 21–33.
[29] J. Souza, F.A. Molfetta, K.M. Honório, R.H.A. Santos, A.B.F. Silva, J. Chem. Inf. Comput. Sci. 44 (2004) 1153–1161.
[30] R.P. Verma, C. Hansch, ChemBioChem 5 (2004) 1188–1195.
[31] Z. Cheng, J. Ren, Y. Li, W. Chang, Z. Chen, J. Pharm. Sci. 92 (2003) 475–484.
[32] A.G. Erlejman, S.V. Verstraeten, C.G. Fraga, P.I. Oteiza, Free Radic. Res. 38 (2004) 1311–1320.
[33] H.Y. Zhang, Y.M. Sun, X.L. Wang, Chem.—Eur. J. 9 (2003) 502–508.
[34] M.E. Ramos-Nino, C.A. Ramirez-Rodriguez, M.N. Clifford, M.R. Adams, J. Appl. Microbiol. 84 (1998) 207–212.
[35] A. Thakur, M. Thakur, P.V. Khadikar, C.T. Supuran, P. Sudele, Bioorg. Med. Chem. 12 (2004) 789–793.
[36] A.T. Balaban, P.V. Khadikar, C.T. Supuran, A. Thakur, M. Thakur, Bioorg. Med. Chem. Lett. 15 (2005) 3966–3973.
[37] V. Consonni, R. Todeschini, M. Pavan, J. Chem. Inf. Comput. Sci. 42 (2002) 682–692.
[38] M.P. González, C. Terán, M. Teijeira, M.J. González-Moa, Eur. J. Med. Chem. 40 (2005) 1080–1086.
[39] L. Saíz-Urra, M.P. González, Y. Fall, G. Gómez, Eur. J. Med. Chem. 42 (2007) 64–70.
[40] B. Baderschneider, P. Winterhalter, J. Agric. Food Chem. 49 (2001) 2788–2798.
[41] M. Karelson, Molecular Descriptors in QSAR/QSPR, Wiley, New York, 2000.
[42] J. Devillers, A.T. Balaban (Eds.), Topological Indices and Related Descriptors in QSAR and QSPR, Gordon and Breach, Amsterdam, 1999.
[43] M. Vedrina, S. Marković, M. Medić-Šarić, N. Trinajstić, Comput. Chem. 21 (1997) 355–361.
[44] H. Wiener, J. Am. Chem. Soc. 69 (1947) 17–20.
[45] M. Barysz, G. Jasari, R.S. Lall, V.K. Srivastava, N. Trinajstić, R.B. King (Eds.), Chemical Applications of Topology and Graph Theory, Elsevier, Amsterdam, 1983.
[46] L.B. Kier, L.H. Hall, J. Pharm. Sci. 65 (1976) 1806–1809.
[47] A.T. Balaban, Chem. Phys. Lett. 89 (1982) 399–404.
[48] C. Shannon, W. Weaver, Mathematical Theory of Communication, University of Illinois, Urbana, 1949.
[49] H.P. Schultz, J. Chem. Inf. Comput. Sci. 29 (1989) 227–228.
[50] ISIS/Draw 2.3, MDL Information System, Inc, 2000.
[51] HyperChem 8.0 Evaluation, Hypercube Inc., http://www.hyper.com, 2007.
[52] I.V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V.A. Palyulin, E.V. Radchenko, N.S. Zefirov, A.S. Makarenko, V.Y. Tanchuk, V.V. Prokopenko, J. Comput.-Aided Mol. Des. 19 (2005) 453–463. http://146.107.217.178/lab/pclient/.
[53] Demo-CONVERT Interactively, Molecular Networks GmbH, http://www.molecular-networks.com/online_demos/convert_demo.html, 2006.
[54] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim, 2000.
[55] J.H. Schuur, P. Selzer, J. Gasteiger, J. Chem. Inf. Comput. Sci. 36 (1996) 334–344.
[56] C.M. Hemmer, V. Steinhauer, J. Gasteiger, J. Vibrat. Spect. 19 (1999) 151–164.
[57] I.V. Tetko, A.E.P. Villa, Neural Networks 10 (1997) 1361–1374.
[58] I.V. Tetko, V.Y. Tanchuk, J. Chem. Inf. Comput. Sci. 42 (2002) 1136–1145.
[59] I. Moriguchi, S. Hirono, I. Nakagome, H. Hirano, Chem. Pharm. Bull. 42 (1994) 976–978.
[60] R. Mannhold, H. van de Waterbeemd, J. Comput.-Aided Mol. Des. 15 (2001) 337–354.
[61] CLOGP, Daylight Chemical Information Systems, Inc.; http://www.daylight.com/daycgi/clogp, 2006.
[62] R. Carrasco, J.A. Padrón, J. Gálvez, J. Pharm. Sci. 7 (2004) 19–26.
[63] D.E. Clark, J. Pharm. Sci. 88 (1999) 807–814.
[64] I. Moriguchi, Y. Kanada, K. Komatsu, Chem. Pharm. Bull. 24 (1976) 1799–1806.
[65] Statistica 6.0, StatSoft, Inc, 2001.