ORIGINAL PAPER

Molecular modeling of wine polyphenols

V. Rastija · S. Nikolić · M. Medić-Šarić

Abstract   Due to wide range of health effects of wine polyphenols, it is important to investigate the relationship between their structure and physical properties (quantitative structure–property relationship, QSPR). We have investigated linear, nonlinear (polynomial), and multiple linear relationships between given topological indices and molecular properties of main pharmacological active components of wine, such as molecular weight (*MW*), van der Waals volume (*Vw*), molar refractivity (*MR*), polar molecular surface area (*PSA*) and lipophilicity (log *P*). Partition coefficient (log *P*) was calculated using three different computer program (CLOGP, ALOGPS and MLOGP). The best models were achieved using the MLOGP program. Topological indices used for correlation analysis include: the Wiener index, $W(G)$; connectivity indices, $\chi(G)$; the Balaban index, $J(G)$; information-theoretic index, $I(G)$; and the Schultz index, $MTI(G)$. QSPR was performed on the set of 19 polyphenols and, particularly, on the group of phenolic acids, and on the group of flavonoids with resveratrol. The connectivity index has been successfully used for describing almost all parameters. Significant correlations were achieved between the Wiener index and van der Waals volume, as well as molecular weight.

V. Rastija
Faculty of Agriculture, University J J. Strossmayer, Trg Sv. Trojstva 3, 31000 Osijek, Croatia
e-mail: vesna.rastija@pfos.hr

S. Nikolić (✉)
The Rugjer Bošković Institute, Bijenicka cesta 54, P.O. Box 180, 10002 Zagreb, Croatia
e-mail: sonja@irb.hr

M. Medić-Šarić
Department of Pharmaceutical Chemistry, Faculty of Pharmacy and Biochemistry,
University of Zagreb, A. Kovačića 1, 10000 Zagreb, Croatia
e-mail: bebamms@pharma.hr

1 Introduction

Polyphenols are secondary plant metabolites and they have been implicated in number
of varied roles including protection against ultraviolet light, pigmentation and defence
against pathogenic microorganisms [1]. Wines, especially red wines contain a wide
range of polyphenols that include phenolic acids, the trihydroxystilbene resveratrol,
the flavonols (e.g. quercetin and myricetin), flavanols (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. The processes of viticulture and vini-
fication determine the content and profile of polyphenols in wine. Vineyard factors
are: variety, quality, climate, ageing, geographical origin and diseases. Vinification
factors are length of grape skin contact and temperature [2]. White wines are usually
made from the free running juice, without the grape mash, having no contact with the
grape skins, and this is the main reason for relatively low phenolic content and lower
antioxidant activity of white wine in comparison to the red wine [3]. Polyphenols play
a major role in wine quality. They are responsible for sensory characteristics such as
colour and bitter flavour of wine. Most of them show beneficial physiological proper-
ties including cardioprotective, anticarcinogenic, anti-inflammatory, and antioxidant
activities [4,5]. Resveratrol is thought to be effective in lowering serum lipids and
inhibiting platelet aggregation [6].

In our research we used an axiom that molecular structure of polyphenols is the basis
of their molecular properties—from chemical and physical properties to the certain
biological activity [7]. Molecular structure is described with a number of molecular
structural parameters that can be calculated from molecular topology, e.g. topological
indices. The topological indices, as nonempirical structural parameters, are convenient
tools to formulate direct relationships between chemical structure and physical, chem-
ical and biological properties of molecules [8]. Several quantitative structure–property
relationship (QSPR) models based on topological indices for modeling the properties
of flavonoid glycosides isolated from *Paliurus spina-christi* Mill. were evaluated and
the best models were obtained using the connectivity index (χ) [9].

In this work we have carried out the quantitative structure–property relationship
to estimate five physico-chemical properties of wine polyphenols: molecular weight
(MW), the logarithm of n-octanol/water partition coefficient (log P), van der Waals
volume (Vw), molar refractivity (MR), and polar surface area (PSA).

Lipophilicity is a physico-chemical property of primary interest for the medici-
nal chemists determining pharmacokinetic and pharmacodinamic behaviour of drugs.
Therefore, quantitative descriptor, log P, is one of the most important pharmacoki-
netic parameter, which describes the oral absorption, cell uptake, protein binding,
blood-brain penetration, metabolism and toxicity (ADME/Tox processes) of bioac-
tive substances. In such a manner, some flavonoids are known to have low lipo-
philicity making it difficult for them to penetrate into the cells [10]. Accompanied
by traditionally "shake flask" method, frequently used experimental methods for the

determination of partition coefficient are chromatography and electrometric titration [11,12]. Moreover, a number of computer programs for prediction of lipophilicity, based on different theoretical approaches (CLOGP, ALOGPS, MLOGP, XLOGP, Log-Kow, etc.) have been developed [13].

The van der Waals volume (*Vw*) is a widely used descriptor in modeling physico-chemical properties. Values of *Vw* are used in quantitative structure–activity relationship modeling to explain factors affecting the biological activity of molecule. Also, the complementary of the shape of drug molecule and receptor cavity is important for the selectivity. For this reasons, van der Waals volume plays a central role in drug design [14]. *Vw* of single molecule is calculated according to the method described by Moriguchi et al. [15]. Volumes of spheres are calculated using the atomic radii and the overlapping volumes are subtracted.

The molar refractivity (*MR*) is a constitutive-additive property used in QSPR/QSAR that is calculated by the Lorenz-Lorentz formula:

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

where *n* is the index of refraction, *M* is the molecular weight, and *d* is the density of a molecule. For a radiation of infinite wavelength, the molar refractivity represents the real volume of the molecules contained in one mole of the substance. *MR* as a molecular descriptor in QSAR studies correlates with lipophilicity, molar volume, and steric bulk [16]. Molar refractivity is related to the London dispersive forces that act in the drug–receptor interaction.

Polar surface area (*PSA*) of a molecule 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. *PSA* is useful parameter for prediction of drug transport properties, such as intestinal absorption [17] or blood-brain barrier penetration [18].

In this work, we have investigated linear, polynomial and multiple linear relationships between five topological indices (the Wiener index, connectivity index, the Balaban index, information-theoretic index and the Schultz index) and selected properties of polyphenols, main pharmacological active components of wine. Since that the physico-chemical properties of polyphenols determine their biological activities this work may be a preliminary study for the future QSAR investigation.

2 Materials and methods

2.1 Data set

Data set contains nineteen polyphenols that comprises eight phenolic acids, ten flavonoids, and one stillbene—resveratrol that were analyzed in Croatian wine by the thin-layer chromatography (TLC) and high-performance liquid chromatography (HPLC) [19]. List and structures of investigated polyphenols are given in Table 1. Calculated topological indices and selected physico-chemical properties for studied compounds are given in Table 2. Regression analysis was performed on a set of all nineteen compounds, then separately, on the set of eight phenolic acids, and on the set of ten flavonoids with resveratrol.

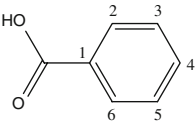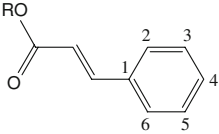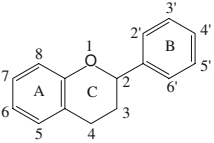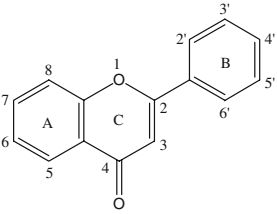**Table 1** Structures and full names of studies polyphenols
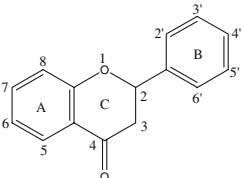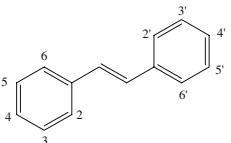
No.	Compounds	Substituent	Structures
	Hydroxybenzoic acid		
1	Vanillic acid	3-OCH$_3$, 4-OH	
2	Syringic acid	3,5-OCH$_3$, 4-OH	
3	Gallic acid	3,4,5-OH	
	Hydroxycinnamic acid		
4	*o*-Hydroxycinnamic acid	2-OH, R=H	
5	*p*-Hydroxycinnamic acid	4-OH, R=H	
6	Ferulic acid	3-OCH$_3$, 4-OH, R=H	
7	Caffeic acid	3,4-OH, R=H	
8	Chlorogenic acid	3,4-OH, R=quinic acid	
	Flavanols		
9	Catechin	3, 5, 7, 3', 4'-OH	
10	Epicatechin	3, 5, 7, 3', 4'-OH	
	Flavonols		
11	Quercetin	3, 5, 7, 3', 4'-OH	
12	Morin	3, 5, 7, 2', 4'-OH	
13	Kaempferol	3, 5, 7, 4'-OH	
14	Myricetin	3, 5, 7, 3', 4', 5'-OH	
15	Isorhamnetin	3, 5, 7, 4'-OH, 3'-OCH$_3$	
	Flavones		
16	Apigenin	5, 7, 4'-OH	
	Flavanones		
17	Flavanon		
18	Naringenin	5, 7, 4'-OH	
	Stilbene		
19	Resveratrol	3, 5, 4'-OH	

Table 2 Physico-chemical properties

Comp. no.	$N(G)$	$W(G)$	$^0\chi(G)$	$^1\chi(G)$	$J(G)$	$I(G)$	MTI	MW	ALOGPS	CLOGP	MLOGP	MR	PSA	V_w
1	12	251	6.365	3.213	3.232	2.830	1,356	168.15	1.71	1.355	0.882	39.624	26.30	1.436
2	14	368	7.696	3.742	3.498	2.872	1,870	198.18	1.55	1.068	0.680	44.737	35.53	1.783
3	12	246	5.696	2.925	3.371	2.774	1,330	170.12	1.17	0.425	0.029	37.899	17.07	1.405
4	12	292	6.189	3.350	2.861	3.135	1,696	164.16	1.90	1.572	1.655	41.904	17.07	1.467
5	12	307	6.189	3.344	2.771	3.255	1,768	164.16	1.74	1.572	1.655	41.904	17.07	1.467
6	14	444	7.520	3.873	3.018	3.309	2,398	194.16	1.58	1.421	1.419	47.017	26.30	1.702
7	13	417	6.336	3.201	3.079	3.453	2,890	180.16	1.67	0.975	1.116	43.598	17.07	1.569
8	25	1,960	12.707	7.225	2.215	3.786	9,976	354.31	0.16	−1.879	−0.043	79.811	43.37	2.418
9	21	1,057	10.698	6.269	2.109	3.353	5,828	290.27	1.02	0.534	0.246	70.243	9.23	2.688
10	21	1,057	10.698	6.269	2.109	3.353	5,828	290.27	1.02	0.534	0.246	70.243	9.23	2.489
11	22	1,252	10.744	6.026	2.269	3.417	7,174	304.26	1.07	0.771	−0.746	71.217	26.30	2.476
12	22	1,226	10.744	6.026	2.290	3.381	7,042	302.24	2.23	1.134	−0.746	71.217	26.30	2.476
13	21	1,111	10.413	5.905	2.260	3.392	6,488	286.24	1.99	1.368	0.014	69.523	26.30	2.374
14	23	1,403	11.075	6.147	2.303	3.434	7,874	318.24	1.66	0.837	−1.494	72.911	26.30	2.578
15	23	1,510	11.586	6.201	2.353	3.587	9,025	316.27	1.96	1.951	−0.496	74.636	35.53	2.340
16	20	1,006	10.082	5.778	2.162	3.407	5,992	270.25	3.07	2.905	0.785	67.829	26.30	2.315
17	17	612	9.297	5.682	2.052	3.162	3,873	224.26	3.10	3.475	2.795	62.501	26.30	2.000
18	20	952	10.290	6.033	2.105	3.341	5,524	272.27	2.47	2.445	0.900	67.583	26.30	2.303
19	17	795	8.921	5.076	2.179	3.573	4,689	228.25	2.57	2.833	2.631	63.670	0.00	2.342

MW molecular weight, log P partition coefficient, MR molar refractivity, PSA polar surface area and V_w van der Waals volume, $N(G)$ number of vertices and calculated topological indices, $(W(G)$ the Wiener index, $^0\chi(G)$ zero-order connectivity index, $^1\chi(G)$ first-order connectivity index, $J(G)$ the Balaban index, $I(G)$ information-theoretic index and (MTI) the Schultz index) for studied compounds

2.2 Calculation of topological indices

In this paper we have investigated whether five topological indices (the Wiener index, connectivity index, the Balaban index, information-theoretic index and the Schultz index) are applicable to QSPR studies of polyphenols from wines. All indices used in our work were calculated using TAM program [20]. 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.

2.3 Calculation of physico-chemical properties of polyphenols

2.3.1 Calculation of partition coefficient, log P

Molecular lipophilicity was calculated using three computer programs based on different theoretical approach (CLOGP, ALOGPS and MLOGP). In this work, SMILES [21] (Simplified Molecular Input Line Entry System) notation created by the structure drawing program CambridgeSoft's ChemDrawUltra was used as chemical structure input for all three programs.

a) ALOGPS 2.1

This program provides interactive on-line estimation of log P and aqueous solubility of compounds [22]. The program to predict lipophilicity was developed using the Efficient Partition Algorithm [23] and an Associative Neural Network (ASNN) approach [24]. This method combines electronic and topological characters to predict lipophilicity of the analyzed molecules.

b) CLOGP

Mannhold and van de Waterbeemd [25] described these programs 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. CLOGP values for studied compounds were calculated with the program accessible via Internet [26], working with the Hansch-Leo's, "fragment constant" method [27].

c) MLOGP

This method for predicting log P values was developed by Moriguchi et al. [28]. 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.

The MLOGP program is included in the DRAGON program [29], the software for the calculation of a large number of molecular descriptors. In our work we used a DRAGON Evaluation version— Software version 5.3 downloaded from Internet [30].

2.3.2 Calculation of other physico-chemical properties

Molecular weight (MW/g mol^{-1}), molar refractivity (MR/m^3 mol^{-1}), and polar surface area (PSA/Å^2) were calculated using DRAGON Evaluation version. Van der Waals volume (Vw/Å^3) was calculated according to the method described by Moriguchi et al. [15].

2.4 Regression analysis

The statistical analysis was performed using STATISTICA 6.0 (StatSoft, Inc.) and CROMRsel (Rugjer Bošković Institute, Zagreb) [31]. We have investigated linear, nonlinear (polynomial) and multivariate relationships between given topological indices and selected properties of polyphenols. To test the quality and accuracy of derived models, following statistical parameters were used: correlation coefficient (R), standard deviation of regression (S) and F-test. Standard deviation of regression (S) was calculated using a following equation:

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

where n denoted the total number of cases (molecules); y_i and $y_i^{'}$ denote the calculated value and value obtained by regression model. The validity of the QSPR model was tested in a leave-one-out cross-validation procedure and marked by the cross-validation correlation coefficient (R_{cv}) and standard error of estimate (S_{cv}). The best possible QSPR models were presented in this paper.
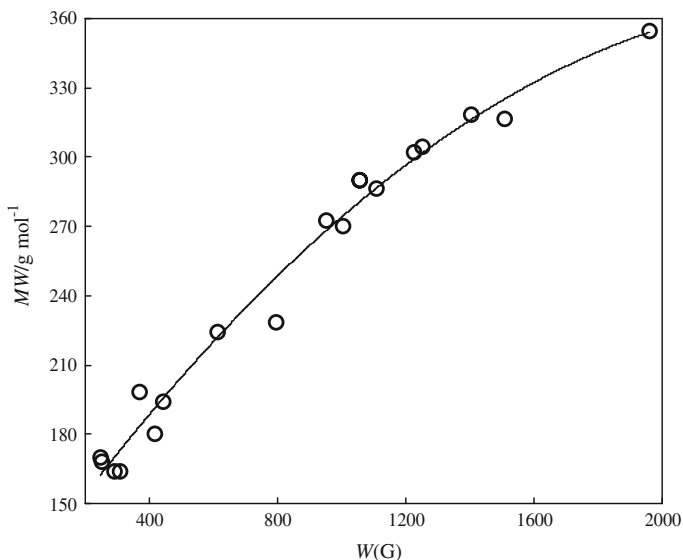
3 Results and discussion

3.1 Molecular weight, MW

The best polynomial model for the estimation of molecular weight of all 19 polyphenols, obtained by regression analysis analysis contains the Wiener index (W):

$$MW = 116.321\,(\pm 6.813) + 0.196\,(\pm 0.016)\,W - 3.794 \times 10^{-5}W^2 \tag{3}$$
$$n = 19;\ R = 0.991;\ S = 8.155;\ F = 438.626;\ R_{cv} = 0.989;\ S_{cv} = 8.874$$

The corresponding graph of MW versus Wiener index is given in Fig. 1.

Fig. 1 Relationship between molecular weight (*MW*) of 19 polyphenols and Wiener index (*W*(G)) (second order polynomial regression, Eq. 3)

Also, good model was obtained by simple linear correlation with the Wiener index for the set of eight phenolic acids

$$MW = 140.601(\pm 4.712) + 0.109(\pm 0.006)\,W \qquad (4)$$
$$n = 8;\ R = 0.991;\ S = 8.214;\ F = 313.2;\ R_{cv} = 0.987,\ S_{cv} = 11.240$$

The Wiener index, which is the sum of all the edges between all pairs of vertices in chemical graph, is highly related with molecular weight as additive properties that can be expressed as a sum of atomic contributions.

For a set of compounds that includes 10 flavonoids and resveratrol, the best model was obtained by linear correlation of *MW* with the zero-order connectivity index ($^0\chi$):

$$MW = -136.282\ (\pm 37.006) + 40.175\ (\pm 3.545)\,^0\chi \qquad (5)$$
$$n = 11;\ R = 0.967;\ S = 7.717;\ F = 128.429;\ R_{cv} = 0.939,\ S_{cv} = 11.558$$

3.2 Partition coefficient

Between the three different computer programs for calculation of partition coefficient (CLOGP, ALOGPS and MLOGP), the best models were achieved using the MLOGP program.

The physico-chemical properties of polyphenols determine their in vivo characteristics of absorption and distribution. Only the molecules of appropriate lipophilicity can diffuse across the phospholipids membrane. The total size of molecule, molecular weight, the three-dimensional shape, the number of functional groups, and the
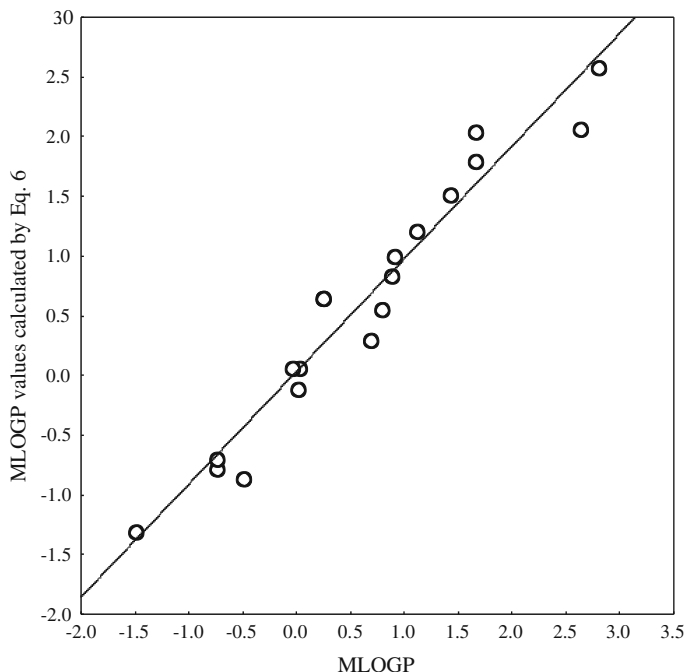
Fig. 2 A plot of the log P values of studied 19 polyphenols calculated using MLOGP method against their values obtained by Eq. 6

number of ring substituents affect on lipophilicity. Among the studied polyphenols, the lowest log P value has myricetin (MLOGP $= -1.494$) and the highest value has flavanon (MLOGP $= 2.795$). p-Hydroxycinnamic acid and o-hydroxycinnamic acid are the most lipophilic phenolic acids (MLOGP $= 1.655$), whereas chlorogenic acid has the lowest lipophilicity (MLOGP $= -0.043$).

Multiple regression analysis found that the tri-parametric model containing the number of atoms (N), the information-theoretic index $I(G)$, and first-order connectivity index ($^{1}\chi$) gives the best results for the estimation of MLOGP values of 19 polyphenols studied

$$\text{MLOGP} = -1.44 \, (\pm 1.023) - 0.923 \, (\pm 0.065) \, N + 2.116 \, (\pm 0.382) \, I$$
$$+ \, 2.291 \, (\pm 0.208)^{1}\chi \tag{6}$$
$$n = 19; \; R = 0.972; \; S = 0.259; \; F = 84.604; \; R_{\text{cv}} = 0.953; \; S_{\text{cv}} = 0.374$$

The plot of log P values calculated by MLOGP program versus MLOGP values calculated by Eq. 6 for studied 19 polyphenols is shown in Fig. 2.

Polynomial regression resulted in the following statistically significant model for the partition coefficient of phenolic acids using the information-theoretic index:

$$\text{MLOGP} = -58.747\,(\pm 8.104) + 37.170\,(\pm 5.031)\,I - 5.727\,(\pm 0.774)\,I^2 \quad (7)$$
$$n = 8;\ R = 0.957;\ S = 0.182;\ F = 27.4;\ R_{\text{cv}} = 0.837;\ S_{\text{cv}} = 0.544.$$

The regression analysis, performed over the set of ten flavonoids with resveratrol, yielded the following bi-parametric model that contains the zero-order connectivity index ($^0\chi$) in combination with the number of atoms (N):

$$\text{MLOGP} = 9.762\,(\pm 0.847) - 1.101\,(\pm 0.088)\,N + 1.281\,(\pm 0.238)^0\chi \quad (8)$$
$$n = 11;\ R = 0.996;\ S = 0.119;\ F = 462.679;\ R_{\text{cv}} = 0.986;\ S_{\text{cv}} = 0.253$$

The obtained models demonstrate the significance of connectivity index for the modeling of lipophilicity. Among the existing topological indices, the connectivity index is one of the most commonly used and it has found wide applications in QSPR/QSAR studies [32,33]. Derived directly from the structural formula, given index encode important structural features such as size, branching, unsaturation, cyclicity, and heteroatom content. The valence connectivity index has been successfully used for describing of lipophilicity in sets of closely related compounds [34] and heterogeneous compounds [35]. In study of Medić-Šarić et al. [9], application of this descriptor to data set of 15 flavonoids and flavonoid glycosides resulted in excellent correlation with log $P(R = 0, 993;\ F = 926.046)$.

3.3 Molar refractivity, *MR*

The simple linear correlation obtained between the values of the first-order connectivity index ($^1\chi$) and values of the molar refractivity (*MR*) for the data set of 19 polyphenols gave the excellent results in accordance with the following expression:

$$MR = 8.360\,(\pm 1.885) + 10.17\,(\pm 0.359)^1\chi \quad (9)$$
$$n = 19;\ R = 0.990;\ S = 2.010;\ F = 800.712;\ R_{\text{cv}} = 0.987;\ S_{\text{cv}} = 0.2342$$

Corresponding graphs of linear correlations between the molar refractivity (*MR*) and first-order connectivity index ($^1\chi$) calculated for above-mentioned data set is given in Fig. 3.

The best QSPR model for a data set of phenolic acids is also based on the first-order connectivity index:
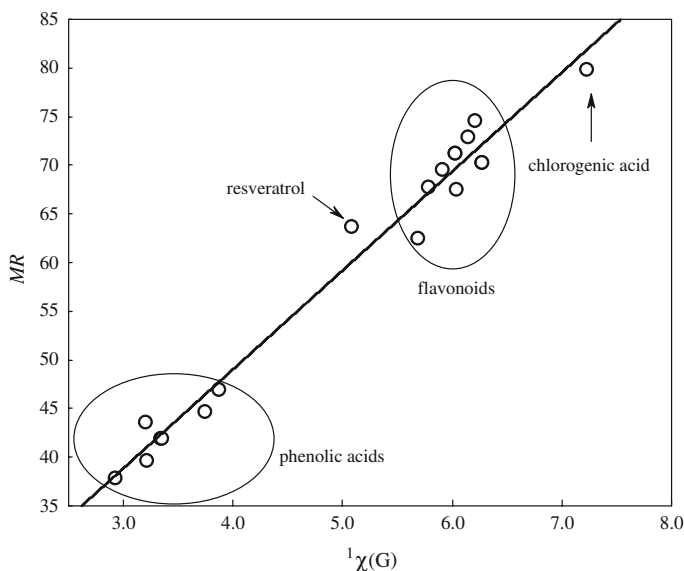
$$MR = 9.742\,(\pm 1.525) + 9.67\,(\pm 0.374)^1\chi \quad (10)$$
$$n = 8;\ R = 0.996;\ S = 1.195;\ F = 667.529;\ R_{\text{cv}} = 0.992;\ S_{\text{cv}} = 1.831$$

Significant correlation was established between zero-order connectivity index ($^0\chi$) and molar refractivity of 11 flavonoids

$$MR = 20.434\,(\pm 3.681) + 4.686\,(\pm 0.353)^0\chi \quad (11)$$
$$n = 11;\ R = 0.975;\ S = 0.768;\ F = 176.564;\ R_{\text{cv}} = 0.946;\ S_{\text{cv}} = 1.163$$

Fig. 3 Linear correlation between molar refractivity (MR) and the first-order connectivity index ($^1\chi$) for 19 polyphenols (Eq. 9)

Strong linear relationships between the molar refractivity and connectivity index have been reported formerly. For example, Yang et al. obtained good QSPR models by connectivity index for alkanes, alkenes, and alcohols with one-variable linear equation [32]. Due to strong correlation between *MR* and connectivity index, a novel hybrid descriptor ($MR\chi$) for molecular modeling was developed. The index is based on a molar refractivity partition using Randić-type graph-theoretical invariant [36].

3.4 Polar surface area, *PSA*

Statistically significant model for the polar surface area was obtained only for the set of phenolic acids by multiple regression analysis. *PSA* of these molecules were found out to correlate with zero-order connectivity index ($^0\chi$) and Schultz index (*MTI*) as shown in Eq. 12

$$PSA = -33.590\,(\pm 8.387) + 9.957\,(\pm 1.604)^0\chi - 0.005\,(\pm 0.001)\,MTI \quad (12)$$
$$n = 8; \; R = 0.972; \; S = 2.191; \; F = 43.295; \; R_{cv} = 0.936; \; S_{cv} = 4.187$$

Also, good model for the same set of phenolic acids was obtained by multiple regression with Wiener index and quadratic values of zero-order connectivity index ($^0\chi^2$)

$$PSA = 0.404\,(\pm 2.675) - 0.523\,(\pm 0.009)\,W + 0.901\,(\pm 0.124)^0\chi^2 \quad (13)$$
$$n = 8; \; R = 0.981; \; S = 1.801; \; F = 65.200; \; R_{cv} = 0.961; \; S_{cv} = 2.708$$

Very weak correlation with other topological indices refers to the fact that *PSA* depends on conformation and possible internal hydrogen bonding [37].

3.5 Van der Waals volume

The best polynomial model for the estimation of van der Waals volume (*Vw*) was obtained by regression analysis based on Wiener index (*W*) for a set of all 19 polyphenols:

$$Vw = 0.893\,(\pm 0.087) + 0.002\,(\pm 0.0002)\,W - 8 \times 10^{-7}W^2 \qquad (14)$$

$$n = 19;\ R = 0.971;\ S = 0.104;\ F = 133.662;\ R_{cv} = 0.937;\ S_{cv} = 0.156$$

Good correlation between the van der Waals volume and Wiener index could be rationalized by the fact that the Wiener index is convenient measure of compactness of the molecule [6,7]. Namely, the Wiener index is roughly proportional to the van der Waals surface area of the respective molecule [38].

Linear regression analysis was found suitable to describe the relation between the van der Waals volume (*Vw*) and zero-order connectivity index ($^0\chi$) for the set of phenolic acids:

$$Vw = 0.585\,(\pm 0.068) + 0.146\,(\pm 0.009)^0\chi \qquad (15)$$

$$n = 8;\ R = 0.989;\ S = 0.046;\ F = 271.585;\ R_{cv} = 0.978;\ S_{cv} = 0.108$$

The best relation for the estimation the van der Waals volumes of flavonoids and resveratrol was obtained by multiple linear regression analysis using Wiener index (*W*) and Schultz index (*MTI*) as descriptors:

$$Vw = 2.118\,(\pm 0.099) + 3.9 \times 10^{-3}(\pm 5.9 \times 10^{-4})W$$
$$-6.3 \times 10^{-4}(\pm 1.1 \times 10^{-4})MTI \qquad (16)$$
$$n = 11;\ R = 0.937;\ S = 0.059;\ F = 29.026;\ R_{cv} = 0.854;\ S_{cv} = 0.089$$

4 Conclusions

QSPR studies are powerful tool for estimation the physical, pharmacological, and toxicological properties of chemical compounds. They have been frequently used in physical, organic, analytical, pharmaceutical, and medicinal chemistry. The main advantage of QSPR techniques is possibility of estimating the properties of compounds when the experimental determination is very complex and expensive. The main hypothesis in the QSPR and QSAR approach is that all properties of a chemical substance are statistically related to its molecular structure. Since the topological indices can be derived directly from the molecular structure, without any experimental effort, they have received great attention in QSPR/QSAR studies.

In this paper, significant regression equations were obtained by linear, polynomial and multiple linear regression analysis for examined physico-chemical properties of

wine polyphenols. Models established are statistically stable. To achieve a better QSPR models for some properties, each group of polyphenols (flavonoids and phenolic acid) was investigated separately.

The best regression correlations were based on the following descriptor: the Wiener index (W), connectivity index (χ), the Schultz index (MTI), and information theoretic index (I). Good correlations were obtained between Wiener index and van der Waals volume, as well as, molecular weight by linear and polynomial regression analysis. Strong relationship of connectivity index and molar refractivity was confirmed by simple linear regression for all three different set of substances.

We calculated the log P values of studied polyphenols using three different computer programs and correlated them with topological indices. The best model for lipophilicity (MLOGP) was obtained for the group of flavonoids with resveratrol, using the zero-order connectivity index in combination with the number of atoms in molecule.

Lipophilicity, van der Waals volume, molar refractivity and polar surface area are useful parameters for structure–activity analysis that can be easily calculated for a polyphenols from wine and used for the modeling of their pharmacological properties, so, here we give a preliminary study for the future QSAR investigation.

References

1. E.G.W.M. Schijlen, C.H.Ric de Vos, A.J. van Tunen, A.G. Bovy, Phytochemistry **65**, 2631 (2004)
2. J. Burns, P.T. Gardner, D. Matthews, G.G. Duthie, M.E.J. Lean, A. Crozier, J. Agric. Food Chem. **49**, 5797 (2001)
3. B. Fuhrman, N. Volkova, A. Suraski, M. Aviram, J. Agric. Food Chem. **49**, 3164 (2001)
4. G. Mazza, L. Fukumoto, P. Delaquis, B. Girard, B. Ewert, J. Agric. Food Chem. **47**, 4009 (1999)
5. N. Al-Awwadi, J. Azay, P. Poucheret, G. Cassanas, M. Krosniak, C. Auger, F. Gasc, J.-M. Mouanet, G. Cros, P.-L. Teissedre, J. Agric. Food Chem. **52**, 1008 (2004)
6. C.R. Pace-Asciak, S. Hahn, E.P. Diamandis, G. Soleas, D.M. Goldberg, Clin. Chim. Acta **235**, 207 (1995)
7. N. Trinajstić, S. Nikolić, B. Lučić, D. Amić, Acta Pharm. **45**, 24 (1996)
8. N. Trinajstić, *Chemical Graph Theory* (CRC Press, Boca Raton, 1992)
9. M. Medić-Šarić, Ž. Maleš, S. Šarić, A. Brantner, Croat. Chem. Acta **69**, 1603 (1996)
10. M. Shichijo, N. Yamamoto, H. Tsujishita, M. Kimata, H. Nagal, T. Kokubo, Biol. Pharm. Bull. **26**, 1685 (2003)
11. K. Valko, C.M. Du, C. Bevan, D.P. Reynolds, M.H. Abraham, Curr. Med. Chem. **8**, 1137 (2001)
12. R.G. Franz, AAPS PharmSci **3** (2001) article 10
13. M. Medić-Šarić, A. Mornar, T. Badovinac-Črnjević, I. Jasprica, Croat. Chem. Acta **77**, 367 (2004)
14. S.P. Gupta, M.S. Babu, S. Sowmya, Bioorg. Med. Chem. **6**, 2185 (1998)
15. I. Moriguchi, Y. Kanada, K. Komatu, Chem. Pharm. Bull. **24**, 1799 (1976)
16. R. Carrasco, J.A. Padrón, J. Gálvez, J. Pharm. Pharm. Sci. **7**, 19 (2004)
17. D.E. Clark, J. Pharm. Sci. **88**, 807 (1999)
18. D.E. Clark, J. Pharm. Sci. **88**, 815 (1999)
19. V. Rastija, *Chromatographic analysis of polyphenols in wines from Croatia*. Ph.D. Thesis, University of Zagreb (2007)
20. M. Vedrina, S. Marković, M. Medić-Šarić, N. Trinajstić, Comput. Chem. **21**, 355 (1997)
21. D.J. Weininger, Chem. Inf. Comput. Sci. **28**, 31 (1988)

22. http://146.107.217.178/lab/alogps/
23. I.V. Tetko, A.E.P. Villa, Neural Networks **10**, 1361 (1997)
24. I.V. Tetko, V.Yu. Tanchuk, J. Chem. Inf. Comput. Sci. **42**, 1136 (2002)
25. R. Mannhold, H. van de Waterbeemd, J. Comput. Aided Mol. Des. **15**, 337 (2001)
26. www.daylight.com/
27. C. Hansch, A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology* (Wiley, New York, 1979)
28. I. Moriguchi, S. Hirono, Q. Liu, Y. Nakagome, Y. Matsushita, Chem. Pharm. Bull. **40**, 127 (1992)
29. R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors* (Wiley-VCH, Weinheim, 2000)
30. http://www.talete.mi.it/download.htm
31. B. Lučić, N. Trinajstić J. Chem. Inf. Comput. Sci. **39** 121 (1999)
32. C. Yang, C. Zhong, J. Chem. Inf. Comput. Sci. **43**, 1998 (2003)
33. S. Nikolić, N. Trinajstić, I. Baučić, J. Chem. Inf. Comput. Sci. **38**, 42 (1998)
34. L.B. Kier, L.H. Hall, *Molecular Connectivity in Structure–Activity Analysis* (Research Studies Press/Wiley, New York, 1986)
35. M. Šoškić, D. Plavšić, J. Chem. Inf. Comput. Sci. **45**, 930 (2005)
36. J.A. Padrón, R. Carrasco, R.F. Pellón, J. Pharm. Pharm. Sci. **5**, 258 (2002)
37. T. Österberg, U. Norinder, J. Chem. Inf. Comput. Sci. **40**, 1408 (2000)
38. J. Zheng, S.K. Han, K.H. Row, Bull. Korean Chem. Soc. **26**, 629 (2005)