Topical Perspectives

Importance of hydrogen bonding and aromaticity indices in QSAR modeling of the antioxidative capacity of selected (poly)phenolic antioxidants

Svetlana Jeremić, Slavko Radenković, Miloš Filipović, Marija Antić, Ana Amić, Zoran Marković

A R T I C L E   I N F O

Article history:
Received 24 November 2016
Received in revised form 5 January 2017
Accepted 6 January 2017
Available online 10 January 2017

Keywords:
QSAR
Antioxidative activity
Phenolic compounds
Hydrogen bond energy
Thermodynamic parameters

A B S T R A C T

The quantitative structure–activity relationship (QSAR) models for predicting antioxidative capacity of 21 structurally similar natural and synthetic phenolic antioxidants was considered. The one-, two- and three-descriptor QSAR models were developed. For this purpose the literature data on the vitamin C equivalent antioxidative capacity (VCEAC) values were used as experimental descriptor of antioxidative capacity. Some thermodynamic and aromaticity properties, as well as the natural bond analysis (NBO) based quantities aimed at measuring the strength of intramolecular hydrogen bonds, were used as independent variables. It was examined whether a combination of these variables can yield a mathematical function that is in good correlation with the VCEAC values. It was shown that a combination of a certain thermodynamic descriptor (related to the single proton loss electron transfer mechanism) with the NBO-based quantities results in several two-descriptor models with the correlation coefficient greater than 0.950. Thus, a significant influence of internal hydrogen bonds on the antioxidative capacity of the studied molecules was confirmed. The best correlation with the VCEAC values was achieved within a three-descriptor QSAR model. This model was obtained by including a magnetic aromaticity index. It was found that aromaticity has only secondary effects on the antioxidative capacity.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Oxidative stress is a phenomenon characterized with increased amount of oxidizing species. These are most often reactive oxygen species, such as hydroxyl and peroxy free radicals, which can be overproduced under the influence of external or internal effects. Excessive amounts of free radicals can then lead to a change in the chemical structure of proteins, nucleic acids, and lipids. Depending on the intensity of oxidative stress, the final result may be a complete destruction of the cell without a possibility for its revitalization. This can be manifested in a form of a variety of diseases, such as cardiovascular, coronary and neurodegenerative, as well as many types of cancer [1–6].

Organisms have their own mechanisms to detoxify reactive intermediates and repair the resulting damage. Although the immune system significantly contributes to the prevention from harmful effects of free radicals, it is very important to protect the organism through nutrition rich in antioxidants. Natural antioxidants are generally polyphenolic molecules of various chemical structures. Most of these molecules can be found in the berries, roots, leaves or flowers of different fruits, vegetables, spices, and medicinal herbs. Because of positive effects to human health, many of the plants containing these compounds have been used in traditional medicine [7,8].

Natural antioxidants proved to be suitable additives in cosmetic products, foods, and food supplements. In addition to investigations devoted to natural sources of antioxidants, much effort has been focused towards finding suitable synthetic antioxidants able to protect industrial products from decay caused by oxidative stress. Synthetic antioxidant should be characterized with high scavenging potency and minor side effects. The influence of some synthesized antioxidants, such as butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), and 2-tert-butylhydroquinone (TBHQ) to human health is investigated [9].
It has been shown that experimentally obtained parameters related to the antioxidative capacity of the molecules mainly agree with their theoretically predicted behaviour. It is also known that biological activity and chemical reactivity of molecules depend on their structure [10,11]. Among the factors that influence the behaviour of molecules, the number and position of different substituents (hydroxyl, methoxy, amino, and methyl groups, double bond, etc.) are of crucial significance [12–15]. These functional groups exhibit strong resonance and inductive effects, thus enabling spin density and charge delocalization in the radical(s) issued from certain antioxidant molecule [13]. It has been shown that the ortho and para positions of the hydroxyl groups increase antioxidative capacity of the related molecule [16]. The vicinal O–H groups participate in the creation of intramolecular hydrogen bonds, stabilizing neutral molecule, as well as the corresponding reactive species [13,17–19].

Many theoretical models have been used in pharmaceutical, medicinal, cosmetic, food, and related industries for developing and investigating new types of antioxidants. Quantitative structure–activity relationship (QSAR) modeling is a very useful statistical method for predicting antioxidative capacity of molecules. In our recent work it was shown that the energy based descriptors related to the antioxidative activity can be used to obtain reliable QSAR models [20]. In the present paper we employ the same set of molecules consisting of 21 structurally similar monomeric phenolic compounds (Fig. 1). Eighteen of them are natural benzoic acids, phenylenecarboxylic acids, and phenols. These compounds are constituents of edible plants in both aglycone and glycoside forms [21–23], and they participate in the normal growth, development, and defence of the plants against infection and injury [24]. Their ability to react with other simple molecules enables esterification or conjugation and formation of esters and polymers [25]. Synthetic antioxidants included in our model (BHA, BHT and TBHQ) are used as foodstuffs and cosmetic products to protect fats against oxidative rancidity. It has been shown that they inhibit production of cancerous luteoskyrin [26]. They do not exert mutagenic activity, but in high doses all three of them exhibit toxic effects [27].

Apparently, all the studied molecules possess aromatic 6-membered ring. This was our motivation to examine a possibility to relate different aromatic indices with the antioxidative capacity of the studied molecules. Finally, we examined how the intramolecular hydrogen bonding quantified within the NBO framework can improve the QSAR model based on the thermodynamic descriptors of the free radical scavenging processes [28].

2. Methodology section

2.1. DFT calculations

There are at least three well described mechanisms through which antioxidants form radical(s) [17,29–32]. According to the first one, radical can be obtained directly, via homolytic dissociation of hydroxyl O–H bond (Eq. (1)). This mechanism is named HAT (hydrogen atom transfer), and is described by a thermodynamic property BDE (bond dissociation enthalpy, Eq. (1.1)).

\[ \text{A} - \text{OH} \rightarrow \text{A} - \text{O}^* + \text{H}^* \] (1)

\[ \text{BDE} = H(\text{A} - \text{O}^*) + H(\text{H}^*) - H(\text{A} - \text{OH}) \] (1.1)

The second mechanism is a two-step process known as SET-PT (single-electron transfer followed by proton transfer). In the first step of this mechanism electron leaves neutral molecule forming radical-cation. Radical-cation then loses its proton in the second step of the reaction yielding the free radical (Eq. (2)). This mechanism is described by two thermodynamic properties: IP (ionization potential, Eq. (2.1)) and PDE (proton dissociation enthalpy, Eq. (2.2)).

\[ \text{A} - \text{OH} \rightarrow \text{A} - \text{OH}^* \rightarrow \text{A} - \text{O}^* \] (2)

\[ \text{IP} = H(\text{A} - \text{OH}^*) + H(e^-) - H(\text{A} - \text{OH}) \] (2.1)

\[ \text{PDE} = H(\text{A} - \text{O}^*) + H(\text{H}^+) - H(\text{A} - \text{OH}^*) \] (2.2)

The third mechanism is another two-step process known as SPLET (sequential proton loss electron transfer). The first step is a heterolytic cleavage of hydroxyl O–H bond, where the corresponding anion is formed. This anion loses an electron giving rise to a radical in the second step of the reaction (Eq. (3)). This mechanism is described by PA (proton affinity, Eq. (3.1)) and ETE (electron transfer energy, Eq. (3.2)) thermodynamic quantities.

\[ \text{A} - \text{OH} \rightarrow \text{A} - \text{O}^- \rightarrow \text{A} - \text{O}^* \] (3)

\[ \text{PA} = H(\text{A} - \text{O}^-) + H(\text{H}^+) - H(\text{A} - \text{OH}) \] (3.1)

\[ \text{ETE} = H(\text{A} - \text{O}^*) + H(e^-) - H(\text{A} - \text{O}^-) \] (3.2)

Geometries of all measured molecules, as well as corresponding charged and radical species, were optimized using the M05-2X functional in conjunction with the 6–311++G(d,p) basis set, as implemented in Gaussian 09 program package [33,34]. Combination of this functional with different basis sets has been widely used. The method was suggested by its creators as appropriate for investigations of thermodynamic and kinetic properties of the reactions that involve species with paired and unpaired spin [35–37].

The geometries of all investigated species in water were optimized by using the PCM/SMD solvation model, to approximate the solvation effects. The frequency calculations were included in all cases. The vibrational analysis has been done in order to verify that the obtained structures correspond to the structures with energy minima. All revealed stationary points with minima of energy showed the absence of imaginary frequencies. The thermodynamic parameters were calculated from the enthalpies of the optimized neutral and charged species. The enthalpies of the solvated proton and electron were taken from literature [38].

2.2. Aromaticity indices

The extent to which the electrons are delocalized within a given ring can be used as a measure of aromaticity. The multicentre delocalization indices quantify electron delocalization among a set of \( n \) atoms [39,40]. The six centre index (SCI) can be calculated as follows

\[ \text{SCI} = \sum_{\mu=1}^{1} \sum_{\nu} \sum_{\sigma} \sum_{C} \sum_{\xi} \sum_{\epsilon} \sum_{F} f_i \left( \text{PS}_{\mu \nu \sigma \xi \epsilon} \right) \] (4)

where \( P \) is the density matrix and \( S \) is the overlap matrix. The summation goes over all basis functions \( \mu, \nu, \sigma, ..., \xi \) centred on the atoms \( A–F \) involved in a given six-membered ring, and \( f_i \) is the permutation operator which for the given set of basis functions \( \mu, \nu, \sigma, ..., \xi \) produces all possible permutations (in total 6! permutations).

The evaluation of aromaticity can be performed by analysing its geometric manifestations. The harmonic oscillator model of aromaticity (HOMA) index is one of the most often used “geometry” aromaticity indices [41,42]. The HOMA index is calculated as

\[ \text{HOMA} = 1 - \frac{1}{n} \sum_{i=1}^{n} \alpha(R_{opt} - R_i)^2 \] (5)

where \( n \) is the number of bonds of the ring considered, \( \alpha \) is a normalization constant, \( R_{opt} \) is the optimal bond length for a fully delocalized \( \pi \)-electron system, and \( R_i \) stands for an actual bond length.
length. In the present work, we used the parameters needed for the HOMA calculations proposed by Krysowski [41].

The nucleus independent chemical shift (NICS) index is among the most popular aromaticity indices [43]. Originally, the NICS was defined as the negative value of the isotropic shielding constant calculated at the ring center [44]. The NICS is a tensor, and different components of the tensor calculated at different positions of a considered system can be used as appropriate indices of aromaticity [43]. In the present study, NICS calculated at the ring centre (NICS(0)), 1 Å above the ring centre (NICS(1)), as well as the out-of-plane component NICSzz(1) were employed.

2.3. NBO analysis

The strength of the hydrogen bonding in the studied molecules was assessed within the NBO framework [28]. In the NBO analysis delocalization effects are gained through the off-diagonal elements of the Fock matrix in the NBO basis. The extent of these delocalization interactions is quantified by means of the second order perturbation theory:

\[ E^{(2)} = q_i \frac{F^2(i,j)}{E(j) - E(i)} \]  

(6)

where \( q_i \) is the i-th donor orbital occupancy, \( E(i) \) and \( E(j) \) are the diagonal elements (energies of the orbitals involved in intramolecular hydrogen bond), and \( F(i,j) \) are the off-diagonal elements of the NBO Fock matrix. By analysing the charge transfer from Y lone pair as a donor to the non-bonding X-H orbital as electron acceptor one can obtain a measure of the strength of hydrogen bonding X-H-Y. In our QSAR model we also used the quantities \( \Delta[E(j)-E(i)] \) and \( \Delta F(i,j) \), which are equal to the differences in energies between the corresponding molecules and radicals. In the cases of the molecules with more than one hydrogen bond, all the NBO-based quantities were calculated as the sums of the corresponding contributions from individual hydrogen bonds.

3. Results and discussion

Here presented development of a QSAR model is based on finding correlation between the experimental parameters of the antioxidative activity on one side, and different thermodynamic descriptors, NBO-based quantities, and aromaticity indices on the other side. The antioxidative capacity of the examined compounds was described using the experimental VCEAC (vitamin C equivalent antioxidative capacity) values, which were taken from Kim and Lee [14]. Such approach is based on the fact that VCEAC is proportional to antioxidative potential. Namely, the VCEAC value greater than 100 mg L⁻¹ indicates that the tested molecule is more effective antioxidant than vitamin C. It was found that some of here investigated compounds are very efficient antioxidants (gallic acid, pyrogallol, homoprotocatechuic acid, and catechol), whereas some of them show very weak antioxidative capacity (salicylic acid and 4-hydroxybenzoic acid).

In this work one-, two-, and three-descriptor QSAR models were developed using linear regression analysis. One-descriptor QSAR models were based on the dependence of different parameters (thermodynamic descriptors, hydrogen bond related parameters, and aromaticity indices), implying that linear correlation was separately examined for each parameter.

The BDE values, and sums (IP + PDE) and (PA + ETE) were examined as possible thermodynamic descriptors of antioxidative activity of the investigated compounds. Their values calculated for water solution are summarized in Table 1. It can be seen that (IP + PDE) is equal to (PA + ETE) [31]. Only the compounds 8, 16 and 17 deviate from this regularity, because in these cases the proton leaves different O–H groups when the reaction follows the SET-P or SPLET mechanism. It is well-known that SPLET is preferable mechanism for the most antioxidants in polar solutions [45]. Bearing these facts in mind it was concluded that the (PA + ETE) values are of the greatest practical importance [13,17,45–47]. For this reason (PA + ETE) was selected as the most suitable thermodynamic descriptor.
Table 1: Experimental and some theoretical indicators of antioxidative activity of investigated molecules: BDE, (IP + PDE) and (PA + ETE) stand for thermodynamic descriptors (kJ mol⁻¹), whereas VCEAC denotes vitamin C equivalent antioxidative capacity (mg L⁻¹) and R denotes correlation coefficient.

<table>
<thead>
<tr>
<th>Compound</th>
<th>VCEAC²</th>
<th>BDE² (IP + PDE)</th>
<th>(PA + ETE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Salicylic acid</td>
<td>1.4</td>
<td>400.7</td>
<td>581.8</td>
</tr>
<tr>
<td>2 4-Hydroxybenzoic acid</td>
<td>4.8</td>
<td>392.4</td>
<td>573.4</td>
</tr>
<tr>
<td>3 3-Hydroxybenzoic acid</td>
<td>53.7</td>
<td>383.2</td>
<td>564.3</td>
</tr>
<tr>
<td>4 Carvacrol</td>
<td>58.0</td>
<td>363.1</td>
<td>544.2</td>
</tr>
<tr>
<td>5 Butylated hydroxytoluene</td>
<td>77.4</td>
<td>333.7</td>
<td>514.8</td>
</tr>
<tr>
<td>6 Syringic acid</td>
<td>80.4</td>
<td>351.7</td>
<td>532.8</td>
</tr>
<tr>
<td>7 4-Hydroxyphenylactic acid</td>
<td>82.8</td>
<td>371.0</td>
<td>552.1</td>
</tr>
<tr>
<td>8 2-tert-butylhydroquinone</td>
<td>83.9</td>
<td>338.2</td>
<td>519.3</td>
</tr>
<tr>
<td>9 Thymol</td>
<td>85.3</td>
<td>399.6</td>
<td>540.7</td>
</tr>
<tr>
<td>10 Homogentisic acid</td>
<td>87.8</td>
<td>340.0</td>
<td>521.1</td>
</tr>
<tr>
<td>11 Gentisic acid</td>
<td>90.8</td>
<td>358.5</td>
<td>539.6</td>
</tr>
<tr>
<td>12 3-Hydroxyphenylactic acid</td>
<td>91.6</td>
<td>374.1</td>
<td>555.2</td>
</tr>
<tr>
<td>13 2-Hydroxyphenylactic acid</td>
<td>95.1</td>
<td>364.9</td>
<td>546.1</td>
</tr>
<tr>
<td>14 Butylated hydroxynasone</td>
<td>97.6</td>
<td>337.0</td>
<td>518.1</td>
</tr>
<tr>
<td>15 Vanillic acid</td>
<td>117.2</td>
<td>371.0</td>
<td>552.1</td>
</tr>
<tr>
<td>16 Protocatechlic acid</td>
<td>163.2</td>
<td>361.0</td>
<td>542.1</td>
</tr>
<tr>
<td>17 2,3-Dihydroxybenzoic acid</td>
<td>169.6</td>
<td>362.1</td>
<td>543.2</td>
</tr>
<tr>
<td>18 Catechol</td>
<td>253.1</td>
<td>346.7</td>
<td>527.8</td>
</tr>
<tr>
<td>19 Homoprotopocatechlic acid</td>
<td>316.7</td>
<td>316.5</td>
<td>497.6</td>
</tr>
<tr>
<td>20 Gallic acid</td>
<td>324.3</td>
<td>345.7</td>
<td>526.8</td>
</tr>
<tr>
<td>21 Pyrogallol</td>
<td>331.2</td>
<td>332.1</td>
<td>513.2</td>
</tr>
</tbody>
</table>

R = -0.641 - 0.641 - 0.637

a Kim and Lee [14];

b Filipović et al. [20].

Linear correlations between the VCEAC values and (PA + ETE), as well as between the VCEAC values and aromaticity indices HOMA, SCI, and NICS are characterized with small correlation coefficients (Tables 1, S1 and S2). Only the correlations where (PA + ETE) and NICS(0) are involved are reasonably satisfactory. On the other hand, the NBO-based quantities related to the strength of intramolecular hydrogen bonds correlate well with the experimental VCEAC values (Table 2). It should be emphasized that only five out of 21 examined antioxidants have intramolecular hydrogen bonds. Several parameters show correlation coefficients larger than 0.900: Δ[E(j)-E(i)], m-[E(j)-E(i)], r-[E(j)-E(i)], and m-F(i), implying that these NBO-based quantities exert significant influence to the antioxidative capacity of the investigated molecules (Figs. 2, S1 – S13). By including the Δ[E(j)-E(i)] term in the best one-descriptor QSAR model was obtained (Eq. (7)):

VCEAC = 75.05(±10.08) + 0.15(±0.02) × Δ[E(j) − E(i)]

(7)

N = 21, R = 0.919, r cv = 0.902, s = 39.8, s cv = 44.7, F = 102.7.

In the above and in subsequent equations, N represents the number of compounds, R is the correlation coefficient, s is the standard error of estimate, and F is Fisher’s F-value. Regression coefficients and the corresponding errors of regression coefficients were computed using the least-square fit procedure. The stability and validity of the model, i.e. the model predictive ability, was tested by the leave-one-out cross-validation (LOO CV) technique. The LOO CV correlation coefficient (r cv) and standard error of fit (s cv) are included. In Fig. 2 the plot of the experimental VCEAC values versus those calculated as a function of Δ[E(j)−E(i)] is presented.

Fig. 2: Scatter plot of experimental VCEAC values obtained by ABTS radical versus calculated Δ[E(j)−E(i)] values by Eq. (7).

Table 2: Values of energies obtained from the perturbation theory energy analysis (kJ mol⁻¹). All here presented energies refer to the energy of hydrogen bond in investigated moieties; prefixes m and r refer to the molecule and radical; prefix Δ refers to the differences in energies between molecule and radical form; R is correlation coefficient.

<table>
<thead>
<tr>
<th>Compound</th>
<th>m-E(2)</th>
<th>m-[E(j)-E(i)]</th>
<th>m-F(1)</th>
<th>r-E(2)</th>
<th>r-[E(j)-E(i)]</th>
<th>r-F(1)</th>
<th>ΔE(2)</th>
<th>Δ[E(j)-E(i)]</th>
<th>ΔF(1)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 Protocatechlic acid</td>
<td>3.9</td>
<td>3150.6</td>
<td>78.8</td>
<td>5.2</td>
<td>2257.9</td>
<td>110.3</td>
<td>-1.3</td>
<td>892.7</td>
<td>-31.5</td>
<td></td>
</tr>
<tr>
<td>17 2,3-Dihydroxybenzoic acid</td>
<td>4.1</td>
<td>3150.6</td>
<td>81.4</td>
<td>7.4</td>
<td>2231.7</td>
<td>160.2</td>
<td>-3.4</td>
<td>918.9</td>
<td>-78.8</td>
<td></td>
</tr>
<tr>
<td>18 Catechol</td>
<td>3.4</td>
<td>3176.9</td>
<td>73.5</td>
<td>6.0</td>
<td>2205.4</td>
<td>118.1</td>
<td>-2.6</td>
<td>971.4</td>
<td>-44.6</td>
<td></td>
</tr>
<tr>
<td>19 Homoprotopocatechlic acid</td>
<td>3.4</td>
<td>3176.9</td>
<td>73.5</td>
<td>5.5</td>
<td>2231.7</td>
<td>112.9</td>
<td>-2.1</td>
<td>945.2</td>
<td>-39.4</td>
<td></td>
</tr>
<tr>
<td>20 Gallic acid</td>
<td>5.4</td>
<td>6327.5</td>
<td>131.3</td>
<td>7.0</td>
<td>4568.4</td>
<td>183.8</td>
<td>-1.5</td>
<td>1759.1</td>
<td>-52.5</td>
<td></td>
</tr>
<tr>
<td>21 Pyrogallol</td>
<td>5.4</td>
<td>6335.7</td>
<td>128.6</td>
<td>7.6</td>
<td>4568.4</td>
<td>189.0</td>
<td>-2.3</td>
<td>1785.3</td>
<td>-60.4</td>
<td></td>
</tr>
</tbody>
</table>

R = 0.895 0.913 0.907 0.879 0.910 0.890 -0.789 0.919 -0.816

a All energy values for the compounds 1–15 are equal to 0.01 kJ mol⁻¹ because they lack hydrogen bonds.
molecular descriptor statistically significant improvement of the model was achieved. Thus, the best three-descriptor QSAR model for antioxidative activity of the examined phenolic compounds was obtained by using (PA + ETE), $\Delta[E(j) - E(\bar{i})]$ and m-NICS(1) as molecular descriptors, Eq. (9). The plot of the experimental VCEAC values versus those calculated by means of Eq. (9) is presented in Fig. 4.

$$\text{VCEAC} = 655.58 ( \pm 191.00 ) - 1.44 ( \pm 0.30 ) \cdot (\text{PA} + \text{ETE})$$

$$+ 0.13 ( \pm 0.01 ) \cdot \Delta[E(j) - E(\bar{i})] - 19.43 (8.84) \cdot m\text{-NICS}(1) \quad (9)$$

$$N = 21, R = 0.971, r_{CV} = 0.956, s = 25.4, s_{CV} = 32.1, F = 93.9$$

Fig. 4 Our investigation reveals that the magnetic index of aromaticity NICS(1) in combination with thermodynamic and NBO parameters provides very accurate QSAR model. It should be emphasized that the magnetic properties of the examined molecules quantified through the NICS(1) values only in combination with the thermodynamic and NBO-based properties can result in acceptable QSAR models.

4. Conclusions

In this paper, the quantitative structure–activity relationship (QSAR) models for antioxidative activity of 21 selected phenolic molecules were developed. The proposed QSAR models are based on the correlation between the VCEAC values as experimental parameters of antioxidative activity, and different parameters for describing the thermodynamic and magnetic properties of the molecules, as well as the NBO parameters for quantifying the strength of intramolecular hydrogen bonding. All applied independent variables are simple and clearly defined descriptors.

It was found that thermodynamic parameter related to the SPLLET mechanism (PA + ETE) in combination with some NBO parameters provides very good correlation with the experimental VCEAC values. In this way the best two-descriptor model, based on (PA + ETE) and $\Delta[E(j) - E(\bar{i})]$, was obtained. Although it was proved that aromaticity plays less important role in assessing the VCEAC values, the m-NICS(1) aromaticity index in combination with (PA + ETE) and $\Delta[E(j) - E(\bar{i})]$ yields the three-descriptor QSAR model that is best correlated with the VCEAC values.

Acknowledgement

This work was supported by the Ministry of Science of the Republic of Serbia (Projects Nos. 172015, 174028 and 174033), and by the Foundation of the Croatian Academy of Sciences and Arts (Projects No. 10–102/244–1–2016).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, at the online version, at http://dx.doi.org/10.1016/j.mmgm.2017.01.011.

References

Santos-Buelga, 2014
Food


Frisch, M.J., Trucks, G.W., Schlegel, et al., Gaussian 09, Revision B.01, Gaussian Inc., Wallingford, CT, 2009.


