MATCH
Communications in Mathematical
and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **88** (2022) 67-78
ISSN: 0340-6253
doi: 10.46793/match.88-1.067M

Application of Changes in Atomic Charges Resulting from Different Electrochemical Oxidation Mechanisms for the Estimation of the First Oxidation Potential of Flavonoids

Ante Miličević

Institute for Medical Research and Occupational Health, Ksaverska cesta 2, HR-10 000 Zagreb, Croatia

antem@imi.hr

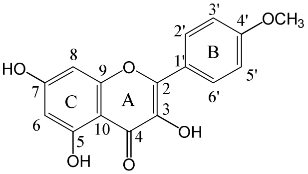
(Received December 15, 2021)

Abstract

In addition to our model based on the sum of the differences in the net atomic charges between a cation and a neutral flavonoid ($\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$) over the carbon atoms in the flavonoid skeleton, here I present two complementary models for the estimation of the first electrochemical oxidation potential, E_{p1}. The models were also based on the sum of differences in the net atomic charges, but between other species of a flavonoid; between a radical and an anion of a flavonoid, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$, and between a radical and a neutral flavonoid, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$. These three variables are connected with different mechanisms (SET-PT, SPLET and HAT) of electron loss during the electrochemical oxidation of a flavonoid. It was shown that the best model for the given set of 29 flavonoids was obtained by using mean values of all of the three variables as a variable ($R^2 = 0.974$, S.E. = 0.042 and S.E.$_{cv}$ = 0.45) which could mean that all three mechanisms are equally possible.

1 Introduction

In our last paper [1], we presented a new theoretical model for the estimation of the first oxidation potential, E_{p1}, of flavonoids. It was based on the sum (over carbon atoms in the skeleton of flavonoids; 1'-6' C atoms in B ring and 2-10 C atoms in CA moiety, Scheme 1) of differences in the net atomic charges between a cation (formed after electron abstraction) and a neutral flavonoid, $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$ (Eq. 1). Earlier [2-4] we developed a model based on the sum of atomic orbital spin populations over the carbon atoms in the skeleton of a flavonoid radical molecule, $\sum_{s(C)} AOSP_{Rad}$. A comparison of our models with the model using hydrogen bond dissociation energy (BDE), which is one of the most important variables for antioxidant activity modeling [5,6], showed inferiority of the BDE model for estimating E_{p1} values [1-4].



Kaempferide

Scheme 1. An example of a flavonoid molecule with numbered carbon atoms in the skeleton.

In this study, I used the same set of 29 flavonoids as in Ref. [1] and their measured E_{p1} (Table 1), and introduced two new models. The models were based on the sum of differences in the net atomic charges between a flavonoid radical and an anion of a flavonoid formed after the abstraction of a hydrogen proton from a neutral molecule, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$, and between a radical and a neutral flavonoid, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$, (Eqs. 2 and 3, respectively). I was able to use $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ variables for modeling E_{p1} because they are connected with the electron loss either by Eqs. 1, 2 or 3 corresponding to the mechanisms (or part of the mechanisms): single electron transfer-proton transfer (SET-PT), sequential proton loss electron transfer (SPLET) and hydrogen atom transfer (HAT), respectively:

$$R\text{-}OH \rightarrow R\text{-}OH^{\cdot +} + e^- \quad (1)$$

$$R\text{-}O^- \rightarrow R\text{-}O\cdot + e^- \quad (2)$$

$$R\text{-}OH \rightarrow R\text{-}O\cdot + H \quad (3)$$

Table 1. The values for the first oxidation potential, E_{p1}, for 29 flavonoids at pH = 3 and 7, active site (A site), sum of differences in the net atomic charges between cation and neutral flavonoid ($\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$), radical and anion ($\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$) and radical and neutral flavonoid ($\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$) calculated using the PM6 in water method and their mean values, and the number of OH groups.

No.	A site	Flavonoid	E_{p1}/V (pH 3)	E_{p1}/V (pH 7)	$\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$ (Var. 1)	$\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ (Var. 2)	$\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ (Var. 3)	Mean 1 and 2	Mean 1 and 3	Mean 2 and 3	Mean 1-3	N_{OH}
1	4'	3,3',4'THF	0.456^{b}	0.197^{b}	0.353	0.333	0.249	0.343	0.301	0.291	0.312	3
2	4'	3',4'DHF	0.513^{b}	0.283^{b}	0.373	0.387	0.272	0.380	0.322	0.329	0.344	2
3	3	3HF	0.751^{b}	0.566^{b}	0.428	0.440	0.239	0.434	0.333	0.339	0.369	1
4	5	5HF	1.164^{b}	0.909^{b}	0.516	0.493	0.358	0.505	0.437	0.425	0.456	1
5	8	7,8DHF	0.456^{b}	0.225^{b}	0.339	0.361	0.217	0.350	0.278	0.289	0.306	2
6	4'	Apigenin	0.928^{c}	0.696^{g}	0.467	0.460	0.335	0.464	0.401	0.398	0.421	3
7	5	Chrisin	1.162^{c}	0.956^{g}	0.508	0.493	0.375	0.501	0.442	0.434	0.459	2
8	3	Galangin	0.655^{c}	0.430^{b}	0.437	0.444	0.244	0.441	0.340	0.344	0.375	3
9	4'	Luteolin	0.513^{b}	0.288^{g}	0.366	0.380	0.266	0.373	0.316	0.323	0.337	4
10	4'	Quercetin	0.435^{c}	0.180^{g}	0.350	0.325	0.248	0.338	0.299	0.287	0.308	5
11	4'	Myricetin	0.351^{c}	0.089^{d}	0.281	0.253	0.229	0.267	0.255	0.241	0.254	6
12	4'	EGC	0.307^{e}	0.028^{e}	0.283	0.293	0.248	0.288	0.266	0.271	0.275	6
13	4'	EC	0.390^{f}	0.150^{f}	0.372	0.374	0.280	0.373	0.326	0.327	0.342	5
14	3	Morin	0.458^{c}	0.227^{g}	0.380	0.335	0.239	0.358	0.310	0.287	0.318	5
15	4'	EGCG	0.367^{c}	0.051^{e}	0.298	0.294	0.248	0.296	0.273	0.271	0.280	5
16	4'	ECG	0.477^{c}	0.162^{f}	0.362	0.374	0.276	0.368	0.319	0.325	0.337	4
17	4'	Naringenin	0.929^{c}	0.704^{h}	0.480	0.462	0.356	0.471	0.418	0.409	0.433	3
18	3	Kaempferid	0.584^{c}	0.369^{h}	0.414	0.407	0.233	0.411	0.323	0.320	0.351	3
19	4'	Dyhidromyricetin	0.354^{d}	0.098^{d}	0.305	0.302	0.245	0.304	0.275	0.273	0.284	6
20	4'	Rutin	0.504^{c}	0.267^{h}	0.361	0.367	0.271	0.364	0.316	0.319	0.333	4
21	3’	Hesperetin	0.737^{a}	0.510^{a}	0.423	0.429	0.322	0.426	0.372	0.375	0.391	3
22	4’	Daidzein	0.795^{a}	0.592^{a}	0.451	0.432	0.328	0.442	0.389	0.380	0.404	2
23	3	Kaempferol	0.498^{a}	0.235^{a}	0.419	0.409	0.234	0.414	0.326	0.321	0.354	4
24	5	Acacetin	1.174^{a}	0.952^{a}	0.509	0.491	0.374	0.500	0.442	0.433	0.458	2
25	4’	Naringin	0.959^{a}	0.732^{a}	0.466	0.463	0.348	0.465	0.407	0.405	0.426	2
26	3’	Neohesperidin	0.766^{a}	0.549^{a}	0.424	0.424	0.322	0.424	0.373	0.373	0.390	2

27	3’	Hesperidin	0.739[a]	0.542[a]	0.424		0.424		0.322		0.424	0.373	0.373	0.390	2
28	4’	Quercitrin	0.500[a]	0.270[a]	0.361		0.367		0.271		0.364	0.316	0.319	0.333	4
29	4’	Gossypin	0.416[a]	0.132[a]	0.349		0.328		0.244		0.339	0.297	0.286	0.307	5

[a]ref[1], [b]ref[2], [c]ref[9], [d]ref[10] [e]ref[11], [f]ref[12], [g][3], [h][4]

2 Calculations

MOPAC calculations

The geometries of 29 flavonoids and their cations, anions and radicals in water were optimized using the MOPAC2016™ PM6 method [7]. All of the initial structures were taken as planar, except hesperetin and its glycosides (**21**, **26** and **27**, Table 1), which we explained in our last paper [1]. The eigenvector following (EF) the optimization procedure was carried out with a final gradient norm under 0.01 kcal mol^{-1} $\mathrm{Å}^{-1}$. The solvent contribution to the enthalpies of formation was computed employing COSMO (Conductor-like Screening Model) calculations implemented in MOPAC2016™.

Regression calculations

Regression calculations, including the leave-one-out procedure (LOO) of cross-validation, were done using the CROMRsel program [8]. The standard error of the cross-validation estimate was defined as:

$$\mathrm{S.E._{cv}} = \sqrt{\sum_i \frac{\Delta X_i^{\,2}}{N}} \tag{4}$$

where ΔX and N denoted cv residuals and the number of reference points, respectively.

3 Results and discussion

After a very good regression based on a variable connected with the electron loss during the SET-PT mechanism, $\sum_{\mathrm{s(C)}} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}}$, on E_{p1} [1]:

$$E_{\mathrm{p1}} = a_1 \sum_{\mathrm{s(C)}} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}} + a_2\Big(\sum_{\mathrm{s(C)}} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}}\Big)^2 + b \tag{5}$$

$R^2 = 0.954$, S.E. = 0.055 and $\mathrm{S.E._{cv}} = 0.059$

I tried a quadratic regression using $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ on the same set of flavonoids. The regression using $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$, the sum of differences in the net atomic charges after a loss of an electron in a SPLET mechanism (Eq. 2), yielded very similar statistics; $R^2 = 0.955$, S.E. = 0.054 and S.E.$_{cv}$ = 0.059 (Fig. 1).

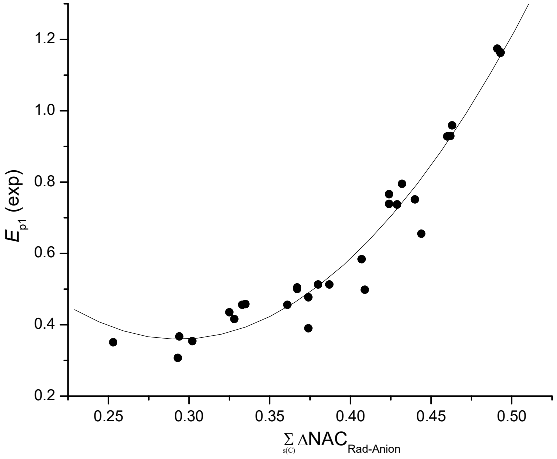


Figure 1. The dependence of experimental E_{p1} (pH = 3) on $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ for the set of 29 flavonoids. Quadratic regression yielded $R^2 = 0.955$, S.E. = 0.054 and S.E.$_{cv}$ = 0.059.

The mean of these two variables yielded even better results when regressed to E_{p1} (R^2 = 0.962, S.E. = 0.050 and S.E.$_{cv}$ = 0.053) which would speak in favor of the assumption that the mechanisms, both SET-PT and SPLET, are possible.

I also used $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$, the sum of differences in the net atomic charges between a flavonoid radical and a neutral flavonoid (based on the HAT mechanism), to correlate E_{p1}. Fig. 2 shows that flavonoids with an active hydroxyl group on the 3C (**3**, **8**, **14**, **18** and **23**) and 8C atom (**5**) are apart from the rest of the flavonoid in the set, and thus, the regression would yield poor statistics, $R^2 = 0.864$, S.E. = 0.094 and $S.E._{cv} = 0.103$. But, the regression on 23 flavonoids, without the aforementioned six flavonoids, yielded results ($R^2 = 0.961$, S.E. = 0.055 and $S.E._{cv} = 0.62$, Fig. 2) comparable to the results obtained by regressions using the other two variables on 29 flavonoids.

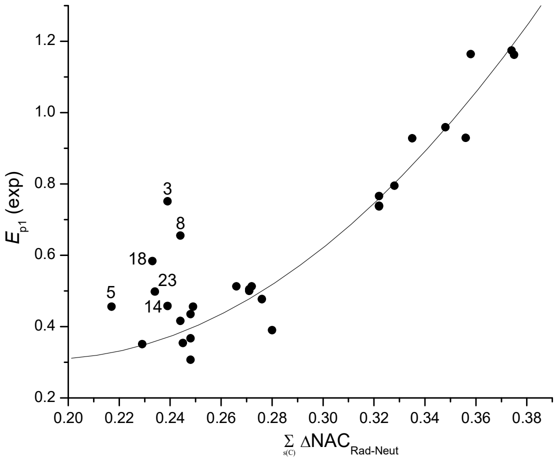


Figure 2. The dependence of experimental E_{p1} (pH = 3) on $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ for the set of 29 flavonoids. Quadratic regression yielded $R^2 = 0.961$, S.E. = 0.055 and $S.E._{cv} = 0.62$ when flavonoids **3**, **5**, **8**, **14**, **18** and **23** were excluded from the regression. (The regression on the full set would yield $R^2 = 0.864$, S.E. = 0.094 and $S.E._{cv} = 0.103$).

When I took the average values of all three variables, $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$, to correlate E_{p1} on 29 flavonoids, a significantly better regression ($R^2 = 0.974$, S.E. = 0.042 and $S.E._{cv} = 0.45$ (Fig. 3)) was obtained than the regressions calculated using any of the variables solely. This regression was also better than any regression using an averaging of a pair of variables; mean of $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ yielded S.E. = 0.54, mean of $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ yielded S.E. = 0.50, mean of $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ yielded S.E. = 0.53.

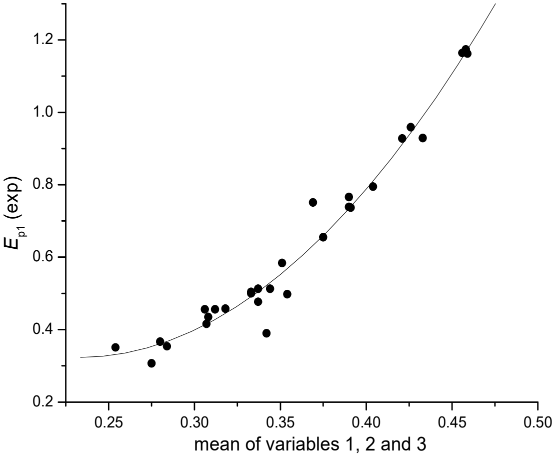


Figure 3. The dependence of experimental E_{p1} (pH = 3) on the mean values of $\sum_{s(C)}$ $\Delta NAC_{Cat\text{-}Neut}$, $\sum_{s(C)}$ $\Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)}$ $\Delta NAC_{Rad\text{-}Neut}$ (variables 1, 2 and 3, Table 1) for the set of 29 flavonoids. Quadratic regression yielded $R^2 = 0.974$, S.E. = 0.042 and S.E.$_{cv}$ = 0.45.

On the other hand, the regressions on 23 flavonoids, without the flavonoids **3**, **5**, **8**, **14**, **18** and **23**, using any of the presented variables resulted in far better S.E. than regressions on a full set of flavonoids; 0.038, 0.036 and 0.055 *vs.* 0.055, 0.054 and 0.094 for $\sum_{s(C)}$ $\Delta NAC_{Cat\text{-}Neut}$, $\sum_{s(C)}$ $\Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)}$ $\Delta NAC_{Rad\text{-}Neut}$, respectively. The regression using mean values of all three variables on E_{p1} ($N = 23$) yielded just a little improvement (S.E. = 0.034) regarding the results of the best two variables, $\sum_{s(C)}$ $\Delta NAC_{Cat\text{-}Neut}$ and $\sum_{s(C)}$ $\Delta NAC_{Rad\text{-}Anion}$, and the same S.E. as the regression using their mean values.

As previously stated [1-4], an improvement of all of the models was obtained by adding the number of OH groups (N_{OH}) on a flavonoid skeleton as a variable to quadratic regressions. The statistics for all of the models mentioned were significantly improved, but again, the model using mean values of all three variables yielded by far the best results, $R^2 = 0.985$, S.E. = 0.031 and S.E.$_{cv}$ = 0.037 ($N = 29$). The introduction of pH as a variable [1-4] allowed the estimation of E_{p1} values at pHs of both 3 and 7 ($N = 58$), and the statistics obtained by mean variable was of similar quality; $R^2 = 0.982$, S.E. = 0.039 and S.E.$_{cv}$ = 0.043 (Fig. 4).

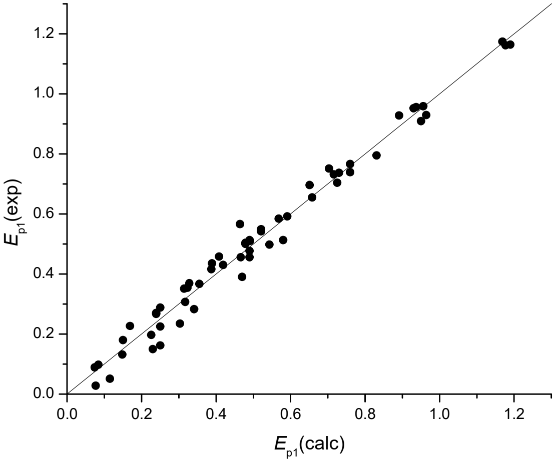


Figure 4. Correlation of experimental (for the set of 29 flavonoids at pHs 3 and 7) *vs.* theoretical E_{p1} values calculated using the mean values of $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ variables, number of OH groups (N_{OH}) and pH ($E_{p1} = a_1 \sum_{s(C)} \text{mean} + a_2(\text{mean})^2 + a_3N_{OH} + a_4\text{pH} + b$;). The correlation yielded $r = 0.991$, S.E. = 0.039 and S.E.$_{cv}$ = 0.043 ($N = 58$).

Due to a different mode of B ring bonding in the isoflavone daidzein (the B ring is bonded on the 3C instead on 2C atom as is the case with other flavonoids in the set), all carbon atoms were not taken into the summation term for $\sum_{s(C)} \Delta NAC_{Cat\text{-}Neut}$, $\sum_{s(C)} \Delta NAC_{Rad\text{-}Anion}$ and $\sum_{s(C)} \Delta NAC_{Rad\text{-}Neut}$ calculation of daidzein. Only the carbon atoms in the B ring (1'-6') and atoms 2C and 3C in the C ring were taken into account, which enabled us to obtain much better statistics for all of the models. For example, the best model (S.E. = 0.042), using mean values of all three variables, would yield much worse results if this restriction for daidzein was not taken into account (S.E. = 0.055).

The case of daidzein is also related to the disruption of the aromaticity similarly as in the case of flavonoids without a C2=C3 double bond (**12**, **13**, **15**-**17**, **19**, **21**, **25**-**27**). They were supposed to be explained together in our last paper [1], but daidzein was inadvertently dropped.

4 Conclusion

By showing that a quadratic regression using $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Anion}}$ is as good model (S.E. = 0.054, Fig. 1) for the estimation of E_{p1} as the same kind of regression using $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}}$ (Eq. 5, S.E. = 0.055), we can conclude that during the oxidation of the flavonoids both the SET-PT and SPLET mechanisms are equally possible. Although the $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Neut}}$ quadratic regression yielded poor statistics (S.E. = 0.094, Fig. 2), by regression using the mean values of all three variables, $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}}$, $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Anion}}$ and $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Neut}}$, on E_{p1} I obtained the best statistics (S.E. = 0.042, Fig. 3) regarding regressions using any of the variables alone. These were also significantly better than the standard errors obtained on the same set of flavonoids using $\sum_{s(C)} \mathrm{AOSP}_{\mathrm{Rad}}$ (S.E. = 0.063 [1]) or BDE (S.E. = 0.085 [1]) which is usually considered as the best variable for the estimation of antioxidant activity [5,6].

The presented results show that $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}}$ and $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Anion}}$ are better variables for the estimation of E_{p1} than $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Neut}}$, either on the set of 29 or 23 (without **3**, **5**, **8**, **14**, **18** and **23**) flavonoids. This led me to the conclusion that SET-PT and SPLET are more preferable mechanisms for the electrochemical oxidation of the flavonoids than the HAT mechanism. However, regressions using mean values of pairs and all three variables showed that the HAT mechanism is also important, especially when flavonoids with an active hydroxyl group on the 3C or 8C atoms are present in the set. The best model, on a full set of flavonoids ($N = 29$), was obtained under the assumption that all mechanisms equally contribute (mean of $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Cat\text{-}Neut}}$, $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Anion}}$ and $\sum_{s(C)} \Delta\mathrm{NAC}_{\mathrm{Rad\text{-}Neut}}$) to the electrochemical oxidation of the flavonoids.

Acknowledgment: This work was supported by the Croatian Ministry of Science and Education.

References

[1] A. Miličević, I. Novak Jovanović, The relationship between the first oxidation potential and changes in electronic structures upon the electrochemical oxidation of flavonoids: Approach to *O*-glycosyl, galloyl and methoxy substituents, *J. Mol. Liq.* **335** (2021) 116223.

[2] A. Miličević, I. Novak Jovanović, G. I. Miletić, Changes in electronic structures of flavonoids upon electrochemical oxidation and a theoretical model for the estimation of the first oxidation potential, *Electrochim. Acta.* **284** (2018) 742–750.

[3] A. Miličević, G. I. Miletić, I. Novak Jovanović, Electrochemical oxidation of flavonoids: PM6 and DFT for elucidating electronic changes and modelling oxidation potential, *J. Mol. Liq.* **285** (2019) 551–556.

[4] A. Miličević, G. I. Miletić, I. Novak Jovanović, Electrochemical oxidation of flavonoids: PM6 and DFT for elucidating electronic changes and modelling oxidation potential (part II), *J. Mol. Liq.* **295** (2019) #111730.

[5] J. S. Wright, E. R. Johnson, G. A. DiLabio, Predicting the activity of phenolic antioxidants: theoretical method, analysis of substituent effects, and application to major families of antioxidants, *J. Am. Chem. Soc.* **123** (2001) 1173–1183.

[6] D. Amić, B. Lučić, G. Kovačević, N. Trinajstić, Bond dissociation enthalpies calculated by the PM3 method confirm activity cliffs in radical scavenging of flavonoids, *Mol. Divers.* **13** (2009) 27–36.

[7] MOPAC2016, HTTP://OpenMOPAC.net/, Version nn.nnn, J. J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, (n.d.). http://openmopac.net/.

[8] B. Lučić, N. Trinajstić, Multivariate regression outperforms several robust architectures of neural networks in QSAR modeling, *J. Chem. Inf. Comput. Sci.* **39** (1999) 121–132.

[9] I. Novak Jovanović, A. Miličević, A new, simplified model for the estimation of polyphenol oxidation potentials based on the number of OH groups, *Arch. Ind. Hyg. Toxicol.* **68** (2017) 93–98.

[10] Š. Komorsky–Lovrić, I. Novak, Abrasive stripping voltammetry of myricetin and dihydromyricetin, *Electrochim. Acta.* **98** (2013) 153–156.

[11] I. Novak, M. Šeruga, Š. Komorsky–Lovrić, Electrochemical characterization of epigallocatechin gallate using square-wave voltammetry, *Electroanalysis* **21** (2009) 1019–1025.

[12] I. Novak, M. Šeruga, Š. Komorsky–Lovrić, Square–wave and cyclic voltammetry of epicatechin gallate on glassy carbon electrode, *J. Electroanal. Chem.* **631** (2009) 71–75.