The Free Internet Journal
for Organic Chemistry

Paper

Archive for
Organic Chemistry

Arkivoc **2020**, part ii, 0-0
to be inserted by editorial office

Computational insights into substituent effects on the stability and reactivity of flavylium cation analogs of anthocyanins

Adilson A. Freitas,[a,*] Karina Shimizu,[a] Luis G. Dias,[b] and Frank H. Quina[c]

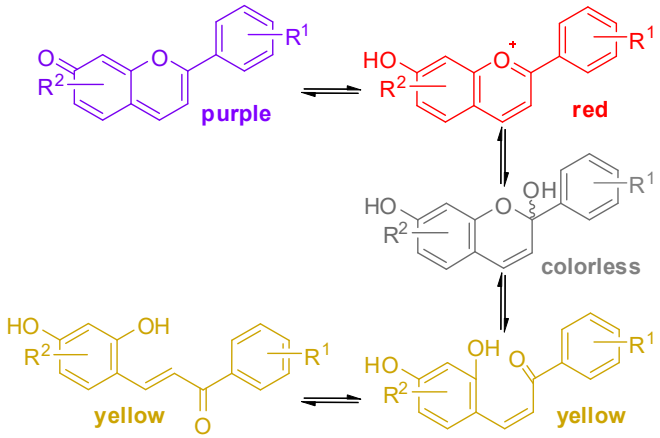
[a]*Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal*
[b]*Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brazil*
[c]*Departamento de Química Fundamental, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil*
Email: adilsondefreitas@tecnico.ulisboa.pt

**Dedicated to Prof. José Manuel Riveros Nigra on the occasion of his 80th Birthday
in appreciation of his life-long contributions to chemistry**

Received 12-13-2019 **Accepted** 05-08-2020 **Published on line** 05-22-2020

Abstract

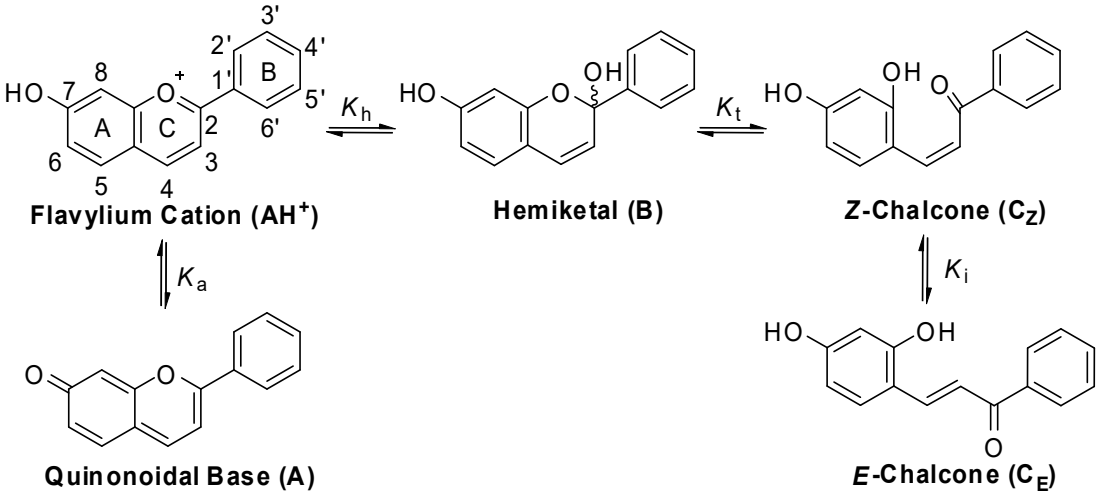
Synthetic flavylium salts and natural anthocyanins undergo hydration by water at physiological pHs, fading out the color. Equilibrium constants of hydration (K_h), tautomerization (K_t) and isomerization (K_i) of 21 substituted flavylium salts were rationalized based on molecular properties calculated by density functional theory (DFT). The good correlations between core-electron binding energies and equilibrium constants revealed an intricate balance of the substitution pattern on the stabilization of AH^+ and neutral species, indicating that the apparent pK_{ap} follows a non-linear relationship with Hammett substituent constants.



Keywords: Core-electron binding energy, electron affinity, density functional theory, flavylium cation, hemiketal, chalcone

Introduction

Anthocyanins are water-soluble pigments (derivatives of the 2-phenyl-benzopyrylium cation) found in many flowers, fruits, leaves and other plant tissues.[1] Distinct from chlorophylls and carotenoids, the absorption of light by anthocyanins can be tuned by substitution and the local environment over the entire visible spectrum (400 nm – 700 nm) as well as in the UV, thus playing an important role in the attraction of plant pollinators, repulsion of herbivores and protection of the photosynthetic apparatus.[2-5] In aqueous media, anthocyanins and flavylium salts can undergo a complex network of chemical reactions, as shown in Scheme 1. Below pH 2, the colored flavylium cation, AH^+, is the most stable form in solution. In mildly acidic or neutral solutions, two chemical transformations of AH^+ can take place: the formation of the colored quinonoidal base (A) in flavylium salts that have deprotonable OH groups and the hydration of AH^+ to give the colorless hemiketal B. Subsequently, the hemiketal produces the colorless *Z*-chalcone, C_Z, via ring-opening tautomerization, which can then isomerize to the *E*-chalcone C_E, typically the thermodynamically most stable form at physiological pH values.[6-11]



Scheme 1. Chemical transformations of hydroxyflavylium salts.

Much progress has been made in the last years in understanding the factors affecting the reactivity of flavylium salts, especially regarding their proton transfer[6-12] and charge transfer[4,13] reactions in water and at micelle-water interfaces. However, measurements of the thermodynamic equilibrium constants can be demanding because of the competitive formation of B and the chalcone isomers, as well as their deprotonated forms.[11] The one-electron reduction of AH^+ in aqueous solutions is irreversible, leading to large uncertainties in the redox potentials.[13,14] Consequently, the rationalization of the reactivity from a physical organic chemical point of view is difficult because there are relatively few compounds for which all of the experimental constants (either kinetic or thermodynamic) have been determined. For these reasons, procedures based on computational methods are an attractive alternative for estimating the reactivity of flavylium salts and anthocyanins.

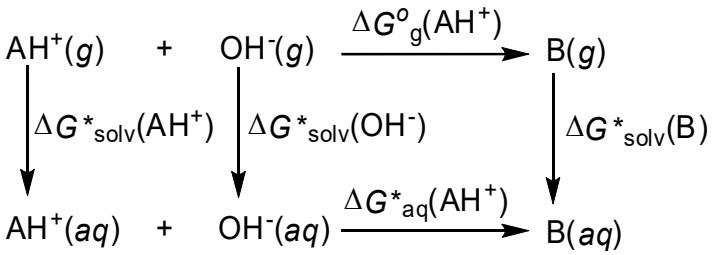
Previously we applied Density Functional Theory (DFT) with success to calculate the pK_as of a series of nine substituted flavylium salts, as well as the one-electron reduction potentials of five flavylium salts.[14] We also developed[15] Linear Free Energy Relationships (LSER) between experimental pK_h, K_t and K_i values for flavylium salts and Hammett substituent constants for substituents on both activated and non-activated

positions, resulting in a model for the prediction of the "apparent pK_a" ($K_{ap} = K_a + K_h + K_hK_t + K_hK_tK_i$) from structure.[1] Empirical Hammett relationships have been used for decades for the rationalization of organic reactions,[16] their mechanisms and other types of chemical phenomena, namely NMR shifts,[17] frontier orbital energies[18] or core-electron binding energy shifts.[19] Our objective in this study was to identify *ab initio* quantum molecular quantities that can be employed as surrogate Hammett parameters, that is, as one-parameter regression descriptors. Such descriptors should facilitate a more basic understanding of the underlying aspects responsible for the observed variations in the reactivity of flavylium cations and anthocyanins. Thus, going beyond earlier work that employed DFT to estimate pK_a values,[14,20] we have now expanded the methodology to compute the hydration equilibrium constants pK_h.

Results and Discussion

Assessment of the hydration equilibrium constants of flavylium salts

The flavylium salts employed in this study and the corresponding experimental equilibrium constants are listed in Table 1. The first step was the evaluation of calculated hydration equilibrium constants in order to ensure that the level of theory employed was adequate. Since the hydration equilibrium constant, K_h, can be written as $K_h = K'_hK_w/[H_2O]$, the alternative thermodynamic cycle shown in Scheme 2 was employed in the theoretical calculations of K'_h because it required only the calculation of the free energy of OH^- in the gas phase, rather than of the free energies of both water on the left side and H^+ on the right side of the equation (thus minimizing potential sources of error). The value of K_h was then obtained by using the experimental autoionization equilibrium constant of water, K_w. The calculated free energies are presented in the Supplementary Material (Table S1).



Scheme 2. Thermodynamic cycle for the calculation of the term K'_h.

Because the compounds studied are relatively large, with at least 16 non-hydrogen atoms, the four solvation models examined provided good results over a range of ca. 10 pK units (Figure 1). The solvation free energies calculated by the universal solvation model based on solute electron density (SMD), the Cramer-Truhlar solvation model (SM5.4P) and both polarized continuum models (PCM and PPM2) were particularly good, resulting in predicted pK_h values with a mean absolute deviation (MAD) of 0.55-0.68 relative to the experimental values.

The level of theory utilized in the gas phase calculations was consistent with that employed elsewhere.[14,21] The addition of diffuse functions to a double-ζ basis set considerably reduces errors in the calculation of reaction energies by DFT.[22] The use of a triple-ζ basis set should therefore have little effect on the calculated electronic energies because the valence basis set is already saturated.[23] Nonetheless, we tested the computation of pK_h employing single point electronic energies obtained with the triple-ζ basis set 6-

311++G(3df,3pd). However, in only one case, specifically compound **5**, did the calculated value show an improvement relative to the double-ζ basis set result. Compound **5** has a OCH_3 group at position C4', indicating that the description of the electron density needs to be enhanced when methoxy groups are present.

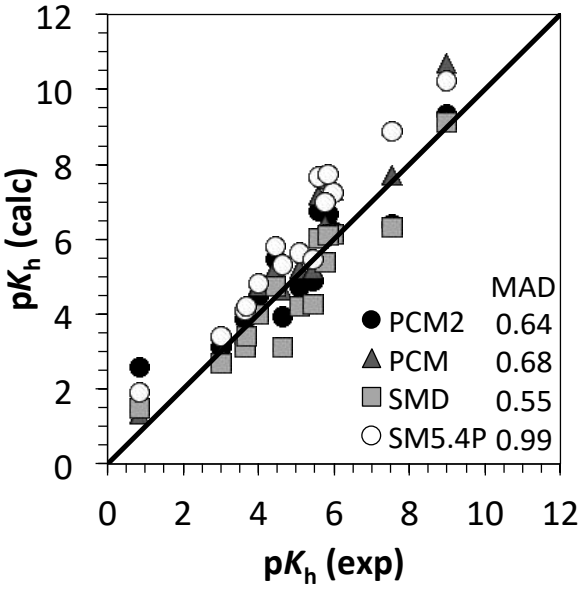


Figure 1. Calculated pK_h values for the flavylium salts studied and the respective mean absolute deviations (MAD) for the solvation models utilized. The straight line corresponds to pK_h calculated = pK_h experimental and is included to guide the eye.

Table 1. Substituents and experimental thermodynamic constants of flavylium salts at 25 °C (see the structures in Scheme 1); data compiled from reference 24

cpd	C3	C4	C5	C6	C7	C8	C4'	pK_{ap}	pK_h	pK_a	-log(K_t)	-log(K_i)
1	H	H	H	H	H	H	H	2.60	3.01		1.22	-2.60
2	H	H	H	H	OH	H	H	2.70	5.10	3.60		-2.70
3	H	H	H	H	H	H	OH	4.20	5.45	5.53	0.05	-3.54
4	H	Me	H	H	OH	H	H	4.40	5.97	4.40		
5	H	H	H	H	H	H	OMe	1.10	4.47		0.36	-3.70
6	H	H	H	H	H	H	Me	1.55	3.65		0.80	-2.88
7	H	H	H	H	H	H	NHCOMe	3.10	3.68		0.72	-1.15
8	H	H	H	H	OH	OH	H	2.30	4.64	3.30		-2.78
9	H	H	H	H	OH	H	Me	2.85	5.77	4.10	-0.68	-2.30
10	H	H	H	NO_2	H	H	OH	-0.60	0.86	5.50		
11	H	H	H	OH	H	H	$N(Me)_2$	6.40	8.98[a]	7.72	0.33	-1.40
12	H	H	H	H	H	H	NH_2	4.10	7.55		-2.33	-2.11
13	H	H	H	OH	H	H	H	2.80	4.00	6.20	0.24	-1.40
14	H	H	H	H	OH	H	OMe	3.14	5.60	4.10		-3.60
15	H	H	H	H	OH	H	OH	3.05	5.85	4.00	-1.31	-3.15
16	OH	H	OH	H	OH	H	OH	2.63				
17	H	H	OMe	H	OH	H	OH	3.92				
18	H	Me	OH	H	OH	H	OMe	4.10				
19	H	Me	OH	H	OH	H	H	4.38		4.35		
20	H	H	H	H	OMe	H	H	2.54				
21	H	H	OH	H	OH	H	OH	4.00		4.20		

[a] calculated at pH = 6 from $K_h = k_h[H^+]/(k_{-h}[H^+] + 0.48)$.

Linear correlations of calculated molecular properties with substituent effects

Previously, we demonstrated[15] that the Hammett resonance σ_R and meta-like σ_m parameters were the best descriptors for the effects of substituents at activated and non-activated positions, respectively, on the hydration equilibrium, pK_h, of flavylium salts. The electronic effects of multiple substituents at activated and non-activated positions attached to the flavylium moiety were found to be additive,[15] resulting in linear correlations with the total $\Sigma\sigma_X$ of the molecule (see Table S2 in Supplementary Material).

Initially, local descriptors such as interatomic distances, bond angles, dihedral angle between benzopyryl and phenyl rings, atomic charges and bond energies were examined. However, none of the geometrical parameters tested correlated with either $\Sigma\sigma_X$ or pK_h, as perhaps might be expected. Also, parameters such as the natural atomic charges of the O1 and C2 atoms of the AH^+ form and the corresponding bond energy performed only at a qualitative level. In conclusion, although the hydration reaction of flavylium salts involves large electrostatic interactions, the local charge values did not quantitatively describe the reactivity of flavylium cations. However, a qualitative analysis of the correlations indicated that electron-releasing groups do indeed enhance the charge density at the reaction center of AH^+, thereby increasing pK_h.

From the concepts of MO theory, the orbital energies $-\varepsilon$HOMO and $-\varepsilon$LUMO are related to the ionization potential (IP) and electron affinity (EA), respectively.[25] Within the Kohn-Sham (KS) framework of DFT, the eigenvalue of the highest-occupied KS orbital corresponds to the negative of the IP and the

eigenvalue of the lowest-unoccupied KS orbital to the negative of the EA.[26] Because the longest-wavelength absorption maxima of flavylium cations and quinonoidal bases are largely determined by the HOMO–LUMO electronic transition[14,27] and because IP – EA is simply the difference between εLUMO and εHOMO, we shall focus on the energies of these frontier orbitals in order to seek relationships linking molecular properties to the reactivity of flavylium salts. Furthermore, the HOMO–LUMO gap is related to the chemical hardness η[28] and the HOMO and LUMO energies are also associated with the absolute electronegativity χ, because χ = (IP + EA)/2. For cations, the EA effectively becomes the electronegativity,[25] since electron transfer from the cation to another molecule is unlikely. Because the exact functionals for exchange and correlation are unknown, the vertical ionization potential (VIP) and vertical electron affinity (VEA) of AH^+ were calculated based on the finite difference approximation employing the energies $E(N)$, $E(N\text{-}1)$ and $E(N\text{+}1)$, where N is the number of electrons in the molecule (*vide* Supplementary Material, Table S3).

We confirmed that there is indeed a good correlation linking VEA_{AH+} and $\Sigma\sigma_X$ ($\Sigma\sigma_X = 4.228\Delta VEA_{AH+}$, R^2 = 0.975, sd = 0.04, F = 779, n = 21) and that electron-releasing groups ($\sigma_X < 0$) decrease the value of VEA_{AH+}, as shown in Figure 2a. The fair correlation for VIP_{AH+} ($\Sigma\sigma_X = 1.501\Delta VIP_{AH+}$, R^2 = 0.844, sd = 0.25, F = 108, n = 21) indicates that it is much less sensitive to substituents given the cationic nature of the AH^+ form. The relationship between VEA_{AH+} and $\Sigma\sigma_X$ allows, for instance, the estimation of group contributions to the total $\Sigma\sigma_X$, the evaluation of whether to use the through-conjugation parameter σ_R^+ (rather than σ_R) for specific cases, or the estimation of theoretical σ-values for substituents with unknown Hammett parameters.

The good correlation found between $\Sigma\sigma_X$ and VEA_{AH+} led us to wonder how this approach might perform for the other multiequilibria of the flavylium salts in aqueous solution. According to Pearson's HSAB principle,[25,27] reactions with hard solvents such as water should become less favorable with the decrease of the absolute hardness of the molecule, which implies a decrease in the difference between the ionization potential and the electron affinity. Indeed, Lietz et al.[27] found two distinct equations correlating the apparent pK_{ap} and the chemical hardness η of 4,7-substituted and 7-OMe substituted flavylium cations. With regards to pK_h, however, there is no reasonable correlation with η, reflecting the fact that VIP_{AH+} is associated with the HOMO orbital energy, which in turn correlates poorly with pK_h (R^2 = 0.281). Because pK_h depends only weakly on VIP_{AH+}, one might expect that the hydration equilibrium is dependent on VEA_{AH+} alone. Indeed, a reasonable correlation was found between VEA_{AH+} and pK_h (Figure 3a, represented as the difference ΔVEA between the VEA of AH^+ and the corresponding value for compound **1**), pointing to the stabilization of AH^+ by electron-donor substituents, i.e., the pK_h of flavylium salts increases as the electron affinity of the flavylium cation decreases. This strongly supports the explanation given earlier by us[13] for the charge-transfer mediated stabilization of flavylium ions against hydration upon complexation with electron rich copigments.

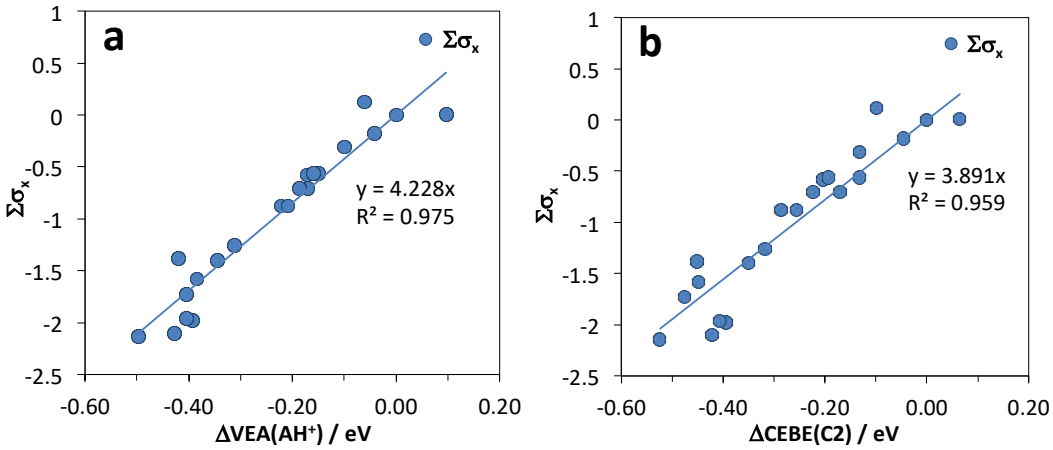


Figure 2. Correlation of the sum of the Hammett constants of substituents of flavylium cations with the: (a) vertical electron affinity and (b) core-electron biding energy of C2(1s) of the AH^+ form.

The weak correlation between ΔVEA_{AH+} and the negative logarithm of the tautomerization equilibrium $-\log(K_t)$ (R^2 = 0.598, n = 10) improves substantially (to R^2 = 0.935) if the outlier **11** is excluded from the correlation (Figure 3b). No relationship was found between ΔVEA of the *Z*- or *E*-chalcone forms and $-\log(K_i)$ (Figure 3c), in accordance with the absence of correlation between the redox potentials of substituted chalcones and Hammett constants.[29] Unfortunately, the set of experimental K_i data is too small to permit an independent treatment of ring effects in chalcones. Likewise, the pK_a values of flavylium compounds were found to be insensitive to VEA (or $\Sigma\sigma_X$) of the AH^+ or A forms (Figure 3d), reminiscent of the lack of correlation of the ground-state acidities of pyranoflavylium cations with Hammett sigma parameters.[30] Pyranoflavylium cations are synthetic analogs of the pyranoanthocyanins, which are formed during the maturation of red wine via condensation reactions between anthocyanins and copigments or yeast metabolic products.

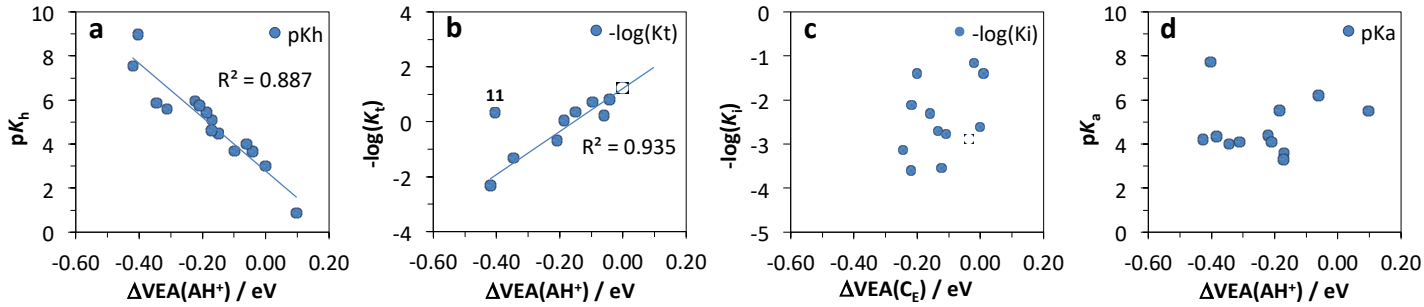


Figure 3. Representations of (a) pK_h and (b) $-\log(K_t)$ as a function of the vertical electron affinity of AH^+, (c) $-\log(K_i)$ as a function of the VEA of the C_E form and (d) pK_a as a function of the VEA of the AH^+ form. The VEAs of the flavylium salt species were calculated at the *m*PW1PW91/6-31G(d) level of theory.

In order to gain more insight into substituent effects on flavylium salt multiequilibria we focused on core-electron binding energies (CEBE), because CEBE shifts correlate nicely with Hammett constants.[19] A simple way to calculate the binding energy is provided by Koopman's theorem, which states that the negative of the orbital energy is equal to the ionization energy. This crude approximation does not take into account orbital relaxation effects, however. To make sure that the level of theory proposed was adequate one

calculated the 1s binding energies for O, N and C atoms in a series of organic compounds with diverse functional groups, totalizing 33 data points, and compared with the experimental data compiled from NIST X-ray Photoelectron Spectroscopy Database.[31] The results presented in Table S4 and Figure S1 (Supplementary Material) are quite good, with R^2 coefficients greater than 0.980 in all three cases. The dispersion in binding energies of O atom mostly reflect sensibility to the instrument setup, such as the kind of solid substrate were the compound is deposited, and experimental settings.

Regarding the flavylium salts, the calculated CEBE differences with respect to the unsubstituted compound **1**, ΔCEBE, were employed since there are no experimental data for flavylium salts for comparison. Table S5 (Supplementary Material) presents the core 1s orbital energies for the O1(1s) and C2(1s) atoms of the AH^+ form derived from the Natural Population Analysis. The relationship between $\Sigma\sigma_X$ and ΔCEBE for C2(1s), represented in Figure 2b ($\Sigma\sigma_X = 3.891\Delta CEBE_{C2(1s)}$, $R^2 = 0.959$, sd = 0.06, $F = 468$, $n = 21$), follows the same trend seen for VEA_{AH+} and indicates that electron-donor groups decrease $CEBE_{C2(1s)}$. Also, except for O1(1s) ($\Sigma\sigma_X = 4.421\Delta CEBE_{O1(1s)}$, $R^2 = 0.940$, sd = 0.09, $F = 312$, $n = 21$), no other atom of the AH^+ chromophoric moiety exhibited a meaningful correlation between the corresponding ΔCEBE and the Hammett constants. Surprisingly, pK_h ($R^2 = 0.910$, $n = 15$) and $-\log(K_t)$ ($R^2 = 0.967$, excluding again compound **11** as with VEA_{AH+}) correlated better with $\Delta CEBE_{C2(1s)}$ than with VEA_{AH+} (Figures 4a and 4b). For $-\log(K_i)$, using ΔCEBE of C2'(1s) or C6'(1s) of the *Z*- and *E*-chalcones resulted in reasonable correlations (after excluding compounds **13** and **7),** with the *Z*-isomer performing slightly better. The average values of ΔCEBE between C2'(1s) and C6'(1s) was employed for the C_Z species in Figure 4c because the two atoms are chemically equivalent in this series of chalcone isomers.

As seen earlier, the deprotonation equilibrium constant, pK_a, of flavylium compounds is insensitive to Hammett constants or VEA values calculated with frontier orbitals, leading us to focus our attention on core electrons in the search for possible correlations. In polyhydroxy flavylium compounds, the acidity of OH groups in the benzopyrylium part of the molecule (A and C rings) is higher than that of OH groups on the phenyl ring (B-ring) by ca. 2 pK_a units (*vide* compounds **2** and **3** in Table 1; refs. 32 and 33). For convenience, it was assumed that multiple OH groups on the benzopyryl moiety can be effectively treated as an average hydroxyl group and that OH groups on the phenyl ring will contribute to the pK_a only in the absence of an ionizable group on benzopyrylium moiety. The good correlation ($R^2 = 0.956$, excluding compound **10**) shown in Figure 4d again points to the stabilization of the AH^+ form by electron-donating substituents. The overall picture is that core-electron binding energies are indeed useful for rationalizing the reactivity of this class of compounds, being more sensitive to substituent-induced changes in reactivity than descriptors based on frontier orbitals such as Electron Affinities.

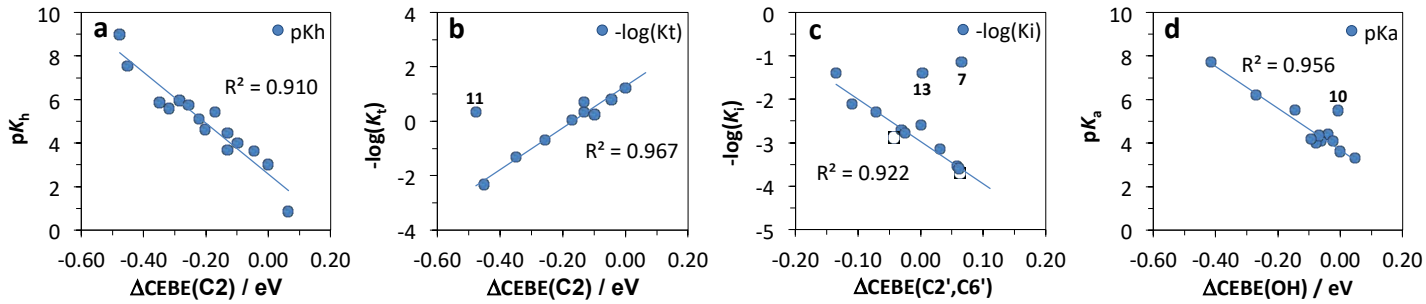


Figure 4. Representations of (a) pK_h and (b) –log(K_t) as a function of the difference between the core-electron binding energy of C2(1s) of AH^+ and the corresponding value for the unsubstituted compound **1**, ΔCEBE; (c) –log(K_i) as a function of the average ΔCEBE of C2'(1s) and C6'(1s) in the C_Z form; and (d) pK_a as a function of ΔCEBE of the O(1s) of the oxygen of the hydroxyl groups of the AH^+ species. The ΔCEBE values of the flavylium salt species were determined from Natural Population analysis at the *m*PW1PW91/6-31G(d) level of theory.

Implications for the prediction of the apparent pK_{ap} of flavylium salts

The stability of the AH^+ form can be regarded as an "apparent" acidity constant, K_{ap}, for the equilibrium between the flavylium cation and a set of "conjugate bases", defined as [CB] = [A]+[B]+[C_Z]+[C_E]. If the isomerization is slow and C_E has not yet been formed or if *E*-chalcone is not present, known as pseudo-equilibrium conditions, the "apparent pK_a" is given by pK_{ap1} = –log(K_a + K_h + K_hK_t). However, when significant amounts of C_E are present and the system has reached the final thermodynamic equilibrium composition, pK_{ap2} = –log(K_a + K_h + K_hK_t + $K_hK_tK_i$). These two definitions of the "apparent pK_a" coexisted in the past, generating some confusion in reported pK_{ap} values in the literature.[24]

Based on the correlations found with ΔCEBE, values of pK_{ap1} and pK_{ap2} were calculated and compared with experimental values (Figure 5a). The mixture of definitions of the "apparent pK_a" is evident here because part of the experimental pK_{ap} data lies close to the calculated values of pK_{ap1}, while the remainder falls in the domain of pK_{ap2} values. Using Hammett substituent constants as descriptors, a non-linear correlation with pK_{ap} can be expected given the complex nature of the mixture of conjugate bases, CB. Indeed, separating the experimental values into pK_{ap1} and pK_{ap2} values and plotting them as a function of $\Sigma\sigma_X$ (Figure 5b) indicated a change in reactivity, with a *plateau* near the limiting values of 4 and 3 for pK_{ap1} and pK_{ap2}, respectively, when $\Sigma\sigma_X < -1$. Both equilibria follow a similar trend with $\Sigma\sigma_X$ and the deviation of compound **11** from the general trend can be attributed to the unusual pH dependence of its reported experimental hydration constant, K_h, presumably reflecting significant cross-conjugative stabilization of the positive charge by the dimethylamino substituent.

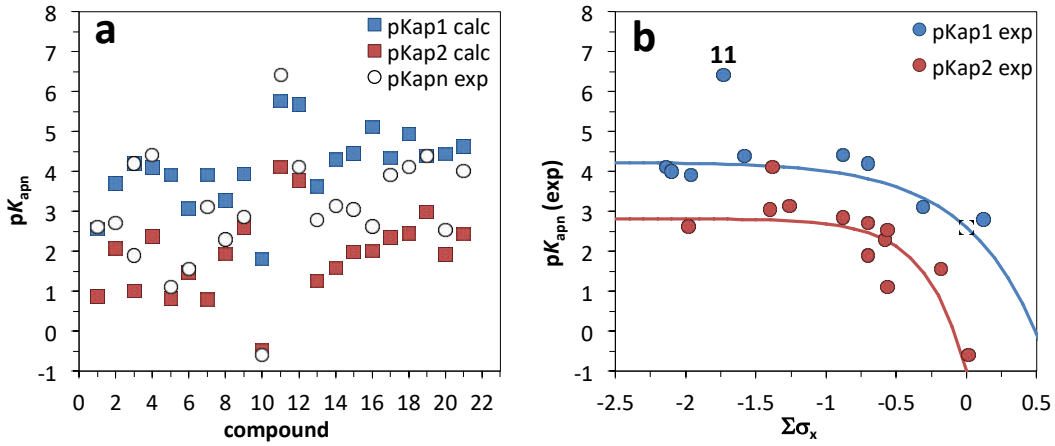


Figure 5. (a) Comparison between experimental and calculated apparent pK_{ap}s of the compounds studied and (b) plot of pK_{ap1} and pK_{ap2} as a function of the sum of Hammett substituent constants for the AH^+ form.

Behavior of anthocyanin pigments in nature

Table 2 lists experimental data for 10 of the most common anthocyanins found in nature and Figure S2 (Supplementary Material) presents the equilibrium constants as a function of the calculated CEBE differences (Table S6 in Supplementary Material). Based on the findings for the equilibrium constants and apparent pK_{ap} of synthetic flavylium salts one can conjecture that the chemical behavior of natural anthocyanin pigments should follow the same trends observed before, only displaced from the original correlations by an offset since these natural pigments are subjected to steric effects at position C3. This hypothesis is also supported by linear free energy relationships (LSER) based on Hammett correlations with synthetic flavylium salts and natural anthocyanins evaluated together.[15] Thus, using the correlations previously found and applying the corresponding offset for each process, the estimated pK_h, pK_a and pK_{ap2} from CEBE calculations gave mean absolute deviations of 0.39, 0.11 and 0.37 pK units, respectively.

The comparison of the ΔCEBE values of compounds in Tables 1 and 2, indicated that the bulky sugar units at C3, characteristic of anthocyanins, displaces the isomerization equilibrium towards the C_Z form to such an extent that the concentration of C_E species is very small to be detected in most cases.[34] The tautomerization equilibrium in anthocyanins, on the other hand, is displaced to the hemiketal form. Also, pK_a and pK_h decrease ca. 1.1 and 3.5 pK units, respectively, when comparing flavylium salts and anthocyanins with similar ΔCEBE values, even observing that the corresponding anthocyanins have lower $\Sigma\sigma_X$ values (Table S7 in Supplementary Material).

Table 2. Substituents and experimental thermodynamic constants of common anthocyanins.

compound	C3	C5	C7	C3'	C4'	C5'	pK_{ap}	pK_h	pK_a	-log(K_t)	-log(K_i)	Ref.
malvin	OGl	OGl	OH	OMe	OH	OMe	1.7	1.85	3.88	0.43	0.27	6
cianin	OGl	OGl	OH	OH	OH	H	2.2		4.09			6
oenin[a]	OGl	OH	OH	OMe	OH	OMe	2.3	2.47	3.80	0.92	-0.30	6, 34
peonin	OGl	OGl	OH	OMe	OH	H	1.9					35
kuromanin[b]	OGl	OH	OH	OH	OH	H	2.5	2.51	3.80	0.92		34
callistephin[c]	OGl	OH	OH	H	OH	H	2.7	2.74	3.90	0.89		34
hibiscin[d]	OSb	OH	OH	OH	OH	OH	2.9		4.31			36
myrthillin[e]	OGl	OH	OH	OH	OH	OH	2.6	2.55	3.80	1.22		34
petunidin 3-*O*-glucoside	OGl	OH	OH	OH	OH	OMe	2.4	2.48	3.70	0.89		34
oxycoccicyanin[f]	OGl	OH	OH	OMe	OH	H	2.4	2.31	3.60	0.59		34

[a] malvidin 3-*O*-glucoside; [b] cianidin 3-*O*-glucoside or chrysanthemin; [c] pelargonidin 3-*O*-glucoside ; [d] delphinidin 3-*O*-sambubioside ; [e] delphinidin 3-*O*-glucoside; [f] peonidin 3-*O*-glucoside.

Conclusions

An interesting panorama emerges from the correlations of the core-electron binding energies with equilibrium constants. One strategy employed to protect the colored AH^+ species against the nucleophilic attack of water at carbon C2 is the inclusion of electron-donating substituents on the flavylium cation moiety. However, at the same time, this approach also leads to stabilization of the colorless Z-chalcone form. This intricate effect of substituents on the different pH-dependent multiequilibria of flavylium salts results in a non-linear relationship between descriptors of substituent effects and the apparent pK_{ap}, which levels off at values of ca. 4 and 3 for pK_{ap1} and pK_{ap2}, respectively, when $\Sigma\sigma_X < -1$. In nature, the chemical structures of the flavylium cation chromophore of the most abundant anthocyanin types (malvidin, delphinidin, pelargonin, cyanin, petunidin and peonin) differ essentially by the presence or absence of OH or methoxy groups at positions 3´and 5' (or in some case due to an additional glycosylation of the OH group at the 5-position) and have pK_{ap}s between 1.7 and 3.0.[24,35] Their $\Sigma\sigma_X$ values fall between –1.2 and –2.1, independent of the substitution pattern if the inductive/resonance effects of the glucosyl (small, similar to *t*-butyl[15]) or sambubiosyl units are not taken into account. The sugar units attached to the oxygen at the 3-position increase the solubility of anthocyanins, but their steric effect, which promotes hydration of the AH^+ form, is compensated by smaller $\Sigma\sigma_X$ values in nature. Rather than attempting to promote stability via the use of substituents, strategies such as the incorporation of flavylium cations in porous clays[37] may be a more useful way to increase the thermal and photochemical stability of the color of this class of compounds.

Experimental Section

Geometry optimization

Geometry optimizations of the flavylium salt structures were performed without any geometry constraint with the hybrid functional[38] *m*PW1PW91 in vacuum using the 6-31+G(d,p) basis set. The selection of this functional

was based on our previous study[14] showing that the electronic transitions for flavylium cations and quinonoidal bases calculated with this functional were in better agreement with experimental results than those calculated with B3LYP. Harmonic frequency calculations indicated that all stationary points found were minima on the electronic potential energy surface, i.e., no imaginary frequencies were found. In the condensed phase, the fully relaxed geometries were obtained at the *m*PW1PW91/6-31G(d) level and the Integral Equation Formalism for the Polarizable Continuum Model (IEFPCM)[39] described the implicit solvent. The united atom topological model,[40] UA0, was used to build the molecular cavity.

Natural population analysis

The natural population analysis (NPA) was carried out with the optimized geometries in condensed phase at the *m*PW1PW91/6-31G(d) level in order to obtain the natural charges, Q_n, and orbital energies of all species (AH^+, B, *Z*- and *E*-chalcones). NPA is a partitioning procedure that transforms a molecular wave function into a localized and orthogonal atomic form satisfying the Pauli exclusion principle. Atomic charge is not physical observable and its value depends on the choice of the basis set and the scheme adopted for the partitioning of the electron density matrix (schemes that do not come from first principles). The advantage of the NPA method is that it solves some shortcomings associated with the Mulliken population analysis procedure such as the basis set dependency problem.[41]

Evaluation of the vertical electron affinities (VEA)

The vertical electron affinities were evaluated as $VEA(_{AH+,\ B}) = E(_{AH+,\ B}) - E(_{AH+,\ B})^-$, where $E(_{AH+,\ B})$ is the electronic energy of the flavylium cation (or hemiketal) optimized at the mPW1PW91/6-31G(d) level in the condensed phase and $E(_{AH+,\ B})^-$ is the electronic energy of the corresponding reduced form calculated at the same geometry.

Calculation of the hydration equilibrium constant pK_h

The hydration equilibrium constant K_h can be written as $K_h = K'_h K_w/[H_2O]$, where the ionic product of water is $K_w = [H^+][OH^-] = 1.0 \times 10^{-14}$, $[H_2O] = 55.3$ mol·L^{-1} and $K'_h = [B]/[AH^+][OH^-]$. The equilibrium constants K'_h were calculated through the thermodynamic cycle shown in Scheme 2 together with the set of expression shown in Eqs. 1-5.

$$G_g^o(OH^-) = \Delta G_g^o(H_2O) + G_g^o(H_2O) - G_g^o(H^+) \quad (1)$$

$$\Delta G_g^o(AH^+) = G_g^o(B) - G_g^o(AH^+) - G_g^o(OH^-) \quad (2)$$

$$\Delta\Delta G_{solv}^* = \Delta G_{solv}^*(B) - \Delta G_{solv}^*(AH^+) - \Delta G_{solv}^*(OH^-) \quad (3)$$

$$\Delta G_{aq}^*(AH^+) = \Delta\Delta G_{solv}^* + \Delta G_g^o(AH^+) - \Delta G_g^{o\rightarrow *} \quad (4)$$

$$pK_h = \frac{\Delta G_{aq}^*(AH^+)}{2.303RT} - \log\left(\frac{1\times10^{-14}}{55.3}\right) \quad (5)$$

In these equations, $\Delta G_g^o(H_2O)$ is the free energy of dissociation of water in the gas phase, $G_g^o(i)$ is the standard free energy of species "*i*" in the gas phase, $\Delta G_{solv}^*(i)$ is the solvation free energy of "*i*" and $\Delta G_{aq}^*(AH^+)$ is the free energy change for the proton transfer reaction in the aqueous phase. The open circle

superscript (°) represents free energies that use a standard-state gas-phase pressure of 1 atm and the superscript asterisk (*) denotes free energies that use an aqueous phase standard-state of 1 mol·L^{-1}. The term $\Delta G_g^{o\rightarrow *} = -RTln(24.46)$ was added in the $\Delta G_g^o(AH^+)$ expression to take into account the change of concentration units from the gas phase reference state to the liquid phase reference state (1 atm to 1 mol·L^{-1}). At 298.15 K $\Delta G_g^{o\rightarrow *}$ is −1.89 kcal·mol^{-1}. Note that this term cancels out in reactions where the same number of moles of reactants and products are transferred from the gas phase to solution.[42,43]

The thermal contributions to the gas phase free energies of the molecules were carried out within the framework of statistical thermodynamics at the *m*PW1PW91/6-31+G(d,p) level. The electronic contribution to the gas phase free energy was acquired by single-point calculations with the 6-311+G(2d,2p) basis set. The Gibbs free energy of OH^- in the gas phase was calculated as follows: first, the H_2O molecule was optimized in vacuum at the *m*PW1PW91/6-31+G(d,p) level and the thermal contributions were carried out at the *m*PW1PW91/6-311+G(2d,2p) level to calculate the term $G_g^o(H_2O)$. The gas phase free energy for OH^- was then calculated from the equation $\Delta G_g^o(H_2O) = G_g^o(H^+) + G_g^o(OH^-) - G_g^o(H_2O)$, taking $G_g^o(H^+)$ = –6.28 kcal·mol^{-1} from the Sackur–Tetrode equation[21] and the experimental[44] value for $\Delta G_g^o(H_2O)$ = 383.7 kcal·mol^{-1}. We employed the experimental value for the free energy of OH^- in the gas phase for the purpose of minimizing potential sources of error.

Since the solvation model directly affects the calculation of pK_h, the solvation free energies were evaluated through four different approaches: by IEFPCM single point at the HF/6-31G(d) level with UAHF radii,[45] by IEFPCM single point reparameterized with Bondi[46] radii (PCM2) at the HF/6-31G(d) level, by solvation model SM5.4P at the PM3 level[47] and by the SMD model.[48] The experimental value[49] of $\Delta G_{solv}^*(OH^-)$ = –104.6 kcal·mol^{-1} was utilized in the calculation of the pK_h with the PCM2 method, given that this model is not parameterized for ions.

The calculations of the terms in the thermodynamic cycle, natural population analysis, vertical electron affinities (VEA) and solvation energies were performed with the Gaussian 03 and 09 packages[50,51] and AMSOL.[47]

Acknowledgements

A.A.F and K.S. thank Fundação para a Ciência e Tecnologia, FCT/MCTES (Portugal) for financial support through CEEC contracts (IST-ID/93/2018 to A.A.F. and IST-ID/100/2018 to K.S.) and through projects PTDC/QUI-QFI/29527/2017, PTDC/QEQ-EPR/5841/2014 and UID/QUI/00100/2019. L.G.D. is the recipient of a CNPq research fellowship and thanks Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, for financial support (projects 2013/08166-5 and 2017/03204-7). F.H.Q. thanks the CNPq for a research productivity fellowship and NAP-PhotoTech, INCT-Catálise (CNPq 465454/2014-3 and CNPq 444061/2018-5), and CNPq (FHQ Universal grant 408181/2016-3) for funding.

Supplementary Material

Supplementary data associated with this article can be found in the online version.

References

1. Brouillard, R. In *Anthocyanins as Food Colours*; Markakis, P. Ed.; Academic Press: New York, 1982, ch. 9. https://doi.org/10.1016/B978-0-12-472550-8.X5001-X
2. Quina, F. H.; Moreira Jr., P. F.; Vautier-Giongo, C.; Rettori, D.; Rodrigues, R. F.; Freitas, A. A.; Ferreira da Silva, P.; Maçanita, A. L. *Pure Appl. Chem.* **2009**, *81*, 1687-1694. https://doi.org/10.1351/PAC-CON-08-09-28
3. Costa, D. A.; Galvão, M.; Di Paolo, R. E.; Freitas, A. A.; Lima, J. C.; Quina, F. H.; Maçanita, A. L. *Tetrahedron* **2015**, *71*, 3157-3162. https://doi.org/10.1016/j.tet.2014.06.092
4. Ferreira da Silva, P.; Paulo, L.; Barbafina, A.; Eisei, F.; Quina, F. H.; Maçanita, A. L. *Chem. Eur. J.* **2012**, *18*, 3736-3744. https://doi.org/10.1002/chem.201102247
5. Silva, V. O.; Freitas, A. A.; Maçanita, A. L.; Quina, F. H. *J. Phys. Org. Chem.* **2016**, *29*, 594-599. https://doi.org/10.1002/poc.3534
6. Lima, J. C.; Vautier-Giongo, C.; Melo, E.; Lopes, A.; Quina, F. H.; Maçanita, A. L. *J. Phys. Chem. A* **2002**, *106*, 5851-5859. https://doi.org/10.1021/jp014081c
7. Paulo, L.; Freitas, A. A.; Ferreira da Silva, P.; Shimizu, K.; Quina, F. H.; Maçanita, A. L. *J. Phys. Chem. A* **2006**, *110*, 2089-2096. https://doi.org/10.1021/jp054038f
8. Freitas, A. A.; Paulo, L.; Maçanita, A. L.; Quina, F. H. *Langmuir* **2006**, *22*, 7986-7993. https://doi.org/10.1021/la0608314
9. Freitas, A. A.; Quina, F. H.; Fernandes, A. C.; Maçanita, A. L. *J. Phys. Chem. A* **2010**, *114*, 4188-4196. https://doi.org/10.1021/jp100281u
10. Freitas, A. A.; Quina, F. H.; Maçanita, A. L. *J. Phys. Chem. A* **2011**, *115*, 10988-10995. https://doi.org/10.1021/jp2069754
11. Pina, F.; Petrov, V.; Laia, C. A. T. *Dyes Pigm.* **2012**, *92*, 877-889. https://doi.org/10.1016/j.dyepig.2011.03.033
12. Freitas, A. A.; Quina, F. H.; Maçanita, A. L. *Photochem. Photobiol. Sci.* **2013**, *12*, 902-910. https://doi.org/10.1039/C3PP25445C
13. Ferreira da Silva, P.; Lima, J. C.; Freitas, A. A.; Shimizu, K.; Quina, F. H.; Maçanita, A. L. *J. Phys. Chem. A* **2005**, *109*, 7329-7338. https://doi.org/10.1021/jp052106s
14. Freitas, A. A.; Shimizu, K.; Dias, L.G.; Quina, F. H. *J. Braz. Chem. Soc.* **2007**, *18*, 1537-1546. http://dx.doi.org/10.1590/S0103-50532007000800014
15. Freitas, A. A.; Dias, L. G.; Maçanita, A. L.; Quina, F. H. *J. Phys. Org. Chem.* **2011**, *24*, 1201-1208. https://doi.org/10.1002/poc.1847
16. Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91*, 165-195. https://doi.org/10.1021/cr00002a004
17. Kopylovich, M. N.; Mahmudov, K. T.; Guedes da Silva, M. F. C.; Martins, L. D. R. S.; Kuznetsov, M. L.; Silva, T. F. S.; Fraústo da Silva, J. T.; Pombeiro, A. J. L. *J. Phys. Org. Chem.* **2011**, *24*, 764-773. https://doi.org/10.1002/poc.1824
18. Mutai, T.; Sawatani, H.; Shida, T.; Shono, H.; Araki, K. *J. Org. Chem.* **2013**, *78*, 2482-2489.

https://doi.org/10.1021/jo302711t
19. Takahata,Y.; Chong, D. P. *Int. J. Quantum Chem.* **2005**, *103*, 509-55. https://doi.org/10.1002/qua.20533
20. Busch, M. S.; Knapp, E. -W. *Chemphyschem* **2004**, *5*, 1513-1522. https://doi.org/10.1002/cphc.200400171
21. Saracino, G. A. A.; Improta, R.; Barone, V. *Chem. Phys. Lett.* **2003**, *373*, 411-415. https://doi.org/10.1016/S0009-2614(03)00607-9
22. Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384-1388. https://doi.org/10.1021/jp021590l
23. Grimme, S.; Steinmetz, M.; Korth, M. *J. Org. Chem.* **2007**, *72*, 2118-2126. https://doi.org/10.1021/jo062446p
24. Pina, F.; Melo, M. J.; Laia, C. A. T.; Parola, A. J.; Lima, J. C. *Chem. Soc. Rev.* **2011**, *41*, 869-908. https://doi.org/10.1039/C1CS15126F
25. Pearson, R. G. *J. Am. Chem. Soc.* **1986**, *108*, 6109-6114. https://doi.org/10.1021/ja00280a002
26. Parr, R. G.; Yang, W.; *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.
27. Lietz, H.; Haucke, G.; Czerney, P.; John, B. *J. Prakt. Chem.* **1996**, *338*, 725-730. https://doi.org/10.1002/prac.199633801142
28. Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512-7516. https://doi.org/10.1021/ja00364a005
29. Alston, J. Y.; Fry, A. J. *Electrochim. Acta* **2003**, *49*, 455-459. https://doi.org/10.1016/j.electacta.2003.08.028
30. Freitas, A. A.; Silva, C. P.; Silva, G. T. M.; Maçanita, A. L.; Quina, F. H. *Photochem. Photobiol.* **2018**, *94*, 1086-1091. https://doi.org/10.1111/php.12944
31. NIST X-ray Photoelectron Spectroscopy Database, NIST Standard Reference Database Number 20, National Institute of Standards and Technology, Gaithersburg MD, 20899 (2000), (retrieved April 2020). http://dx.doi.org/10.18434/T4T88K
32. Costantino, L.; Rastelli, G.; Rossi, M. C.; Albasini, A. *J. Chem. Soc., Perkin Trans. 2* **1995**, 227. https://doi.org/10.1039/P29950000227
33. Alejo-Armijo, A.; Basílio, N.; Freitas, A. A.; Maçanita, A. L.; Lima, J. C.; Parola, A. J.; Pina, F. *Phys. Chem. Chem. Phys.* **2019**, *21*, 21651-21662. https://doi.org/10.1039/C9CP04917G
34. Leydet, Y.; Gavara, R.; Petrov, V.; Diniz, A. M.; Parola, A. J.; Lima, J. C.; Pina, F. *Phytochemistry* **2012**, *83*, 125-135. https://dx.doi.org/10.1016/j.phytochem.2012.06.022
35. Abe, K.; Sakaino, Y.; Kakinuma, J.; Kakisawa, H. *Nippon Kagaku Kaishi* **1977**, 1197-1204. https://doi.org/10.1246/nikkashi.1977.1197
36. Vidot, K.; Achir, N.; Mertz, C.; Sinela, A.; Rawat, N.; Prades, A.; Dangles, O.; Fulcrand, H.; Dornier, M. *J. Agric. Food Chem.* **2016**, *64*, 4139-4145. https://doi.org/10.1021/acs.jafc.6b00701
37. Silva, G. T. M.; Silva, C. P.; Gehlen, M. H.; Oake, J.; Bohne, C.; Quina, F. H. *App. Clay Sci.* **2018**, *162*, 478-486. https://doi.org/10.1016/j.clay.2018.07.002

38. Adamamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664-675. https://doi.org/10.1063/1.475428
39. Mennucci, B.; Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1998**, *109*, 2798-2807. https://doi.org/10.1063/1.476878
40. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024-10035. https://doi.org/10.1021/ja00051a040
41. Jensen, F.; *Introduction to Computational Chemistry*; John Wiley & Sons: New York, 1999.
42. Pliego Jr, J. R.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622-1627. https://doi.org/10.1039/B109595A
43. Pliego Jr, J. R.; Riveros, J. M. *WIREs Comput Mol Sci.* **2020**, 10, e1440. https://doi.org/10.1002/wcms.1440
44. Palascak, M. W.; Shields, G. C. *J. Phys. Chem. A* **2004**, *108*, 3692-3694. https://doi.org/10.1021/jp049914o
45. Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210-3221. https://doi.org/10.1063/1.474671
46. Shimizu, K.; Freitas, A. A.; Farah, J. P. S.; Dias, L. G. *J. Phys. Chem. A* **2005**, *109*, 11322-11327. https://doi.org/10.1021/jp054673l
47. Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chambers, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Zhu, T.; Thompson, J. D.; Winget, P.; Lynch, B. J.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G.; *AMSOL-version 7.0*, 2003.
48. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378-6396. https://doi.org/10.1021/jp810292n
49. Camaioni, D. M.; Schwerdtfeger, C. A. *J. Phys. Chem. A* **2005**, *109*, 10795-10797. https://doi.org/10.1021/jp054088k
50. Gaussian 03, Revision B.04, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. A.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Peterson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Know, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, Ö.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh PA, 2003.
51. Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma,

K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.