# Supplementary Material

# 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*

## Table of Contents

**Table S1.** Gas phase energies, solvation free energies (in kcal.mol$^{-1}$) and p$K_h$ equilibrium constants calculated at $T$ = 25 °C

| cpd | $\Delta G_g^*$ (AH$^+$)$^a$ | $\Delta G_{solv}^*$ (AH$^+$)$^b$ | $\Delta G_{solv}^*$ (AH$^+$)$^c$ | $\Delta G_{solv}^*$ (AH$^+$)$^d$ | $\Delta G_{solv}^*$ (AH$^+$)$^e$ | $\Delta G_{solv}^*$ (B)$^a$ | $\Delta G_{solv}^*$ (B)$^b$ | $\Delta G_{solv}^*$ (B)$^c$ | $\Delta G_{solv}^*$ (B)$^d$ | p$K_h$$^b$ | p$K_h$$^c$ | p$K_h$$^d$ | p$K_h$$^e$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | -144.5 | -46.9 | -40.3 | -45.9 | -39.2 | -6.82 | -6.43 | -9.19 | -5.98 | 3.12 | 2.91 | 2.67 | 3.39 |
| **2** | -141.5 | -50.2 | -46.3 | -50.5 | -44.3 | -11.0 | -12.3 | -14.7 | -11.1 | 4.70 | 5.20 | 4.20 | 5.63 |
| **3** | -141.0 | -50.1 | -46.2 | -50.1 | -43.5 | -11.0 | -12.6 | -14.7 | -10.9 | 4.90 | 5.21 | 4.24 | 5.44 |
| **4** | -137.7 | -48.7 | -44.7 | -48.9 | -42.6 | -11.1 | -11.5 | -14.2 | -10.9 | 6.21 | 7.30 | 6.14 | 7.23 |
| **5** | -139.1 | -46.5 | -39.7 | -44.7 | -38.2 | -8.64 | -8.05 | -10.6 | -7.16 | 5.45 | 5.25 | 4.74 | 5.81 |
| **6** | -142.6 | -45.7 | -38.9 | -44.2 | -38.0 | -6.51 | -6.16 | -8.79 | -5.72 | 3.85 | 3.47 | 3.11 | 4.12 |
| **7** | -140.3 | -53.4 | -44.2 | -51.9 | -44.4 | -16.2 | -12.8 | -18.4 | -14.3 | 4.12 | 4.14 | 3.38 | 4.21 |
| **8** | -140.8 | -51.6 | -47.5 | -52.2 | -45.2 | -14.1 | -15.0 | -18.5 | -13.1 | 3.94 | 4.62 | 3.12 | 5.30 |
| **9** | -138.7 | -49.1 | -44.9 | -48.8 | -43.1 | -10.9 | -12.1 | -14.2 | -10.8 | 6.05 | 6.38 | 5.40 | 7.01 |
| **10** | -152.0 | -60.7 | -54.2 | -60.1 | -51.4 | -13.8 | -14.9 | -17.5 | -12.7 | 2.57 | 1.33 | 1.47 | 1.89 |
| **11** | -129.5 | -47.0 | -43.5 | -46.6 | -39.0 | -13.4 | -14.0 | -16.1 | -11.5 | 9.34 | 10.7 | 9.14 | 10.2 |
| **12** | -134.9 | -47.3 | -41.0 | -45.9 | -42.0 | -12.4 | -10.2 | -13.8 | -10.9 | 6.38 | 7.70 | 6.30 | 8.87 |
| **13** | -142.4 | -50.9 | -46.9 | -51.2 | -44.0 | -11.1 | -12.6 | -14.8 | -10.9 | 4.45 | 4.75 | 3.99 | 4.81 |
| **14** | -137.1 | -50.4 | -46.1 | -49.7 | -43.4 | -12.7 | -13.8 | -15.8 | -11.8 | 6.75 | 7.17 | 6.03 | 7.67 |
| **MAD**$^f$ | | | | | | | | | | **0.64** | **0.68** | **0.55** | **0.99** |
| **15** | -138.2 | -53.9 | -52.4 | -55.0 | -48.4 | -15.2 | -18.4 | -19.9 | -15.5 | 6.67 | 7.59 | 6.10 | 7.75 |
| **16** | -139.3 | -57.9 | -61.5 | -60.7 | -55.4 | -22.9 | -28.1 | -29.8 | -23.4 | 3.22 | 6.34 | 2.18 | 6.34 |
| **17** | -134.0 | -51.4 | -49.3 | -51.7 | -46.1 | -16.5 | -18.5 | -20.4 | -16.1 | 6.99 | 8.30 | 6.34 | 8.70 |
| **18** | -131.4 | -50.4 | -48.7 | -50.3 | -44.8 | -16.2 | -17.9 | -19.6 | -15.0 | 8.37 | 10.4 | 7.86 | 10.8 |
| **19** | -136.1 | -50.6 | -49.0 | -51.2 | -45.4 | -14.2 | -16.2 | -18.2 | -13.8 | 6.59 | 8.27 | 6.15 | 8.44 |
| **20** | -139.8 | -46.8 | -39.6 | -45.0 | -38.2 | -8.42 | -7.57 | -10.4 | -6.85 | 5.29 | 5.00 | 4.59 | 5.52 |
| **21** | -136.5 | -55.4 | -56.4 | -57.5 | -51.8 | -18.9 | -23.3 | -24.3 | -19.9 | 6.32 | 8.19 | 5.87 | 8.22 |

[a] $\Delta G_g^* = \Delta G_g^o(AH^+) - \Delta G_g^{o\rightarrow *}$

[b] $\Delta G_{solv}^*$ calculated by the PCM2 model and $\Delta\Delta G_{solv}^*$ evaluated employing the experimental[46] $\Delta G_{solv}^*(OH^-)$= –104.6 kcal.mol$^{-1}$;

[c] $\Delta G_{solv}^*$ calculated by the IEFPCM HF(UAHF) model and $\Delta\Delta G_{solv}^*$ evaluated with $\Delta G_{solv}^*(OH^-)$ = –110.52 kcal.mol$^{-1}$, estimated at the same level;

[d] $\Delta G_{solv}^*$ calculated by the SMD model with $\Delta G_{solv}^*(OH^-)$ = –107.37 kcal.mol$^{-1}$, estimated at the same level;

[e] $\Delta G_{solv}^*$ calculated by the SM5.4 model and $\Delta\Delta G_{solv}^*$ evaluated with $\Delta G_{solv}^*(OH^-)$ = –111.87 kcal.mol$^{-1}$, estimated at the same level;

[f] Mean Absolute Deviation.

**Table S2.** Hammett $\sigma$ parameters for the substituents of the flavylium salts studied. The $\sigma$-values were compiled from reference 16

| compound | C3 ($\sigma_m$) | C4 ($\sigma_R$) | C5 ($\sigma_R$) | C6 ($\sigma_m$) | C7 ($\sigma_R$) | C8 ($\sigma_m$) | C4' ($\sigma_R$) | $\Sigma\sigma_X$ |
|---|---|---|---|---|---|---|---|---|
| **1** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 0 | 0 | 0 | -0.7 | 0 | 0 | -0.7 |
| **3** | 0 | 0 | 0 | 0 | 0 | 0 | -0.7 | -0.7 |
| **4** | 0 | -0.18 | 0 | 0 | -0.7 | 0 | 0 | -0.88 |
| **5** | 0 | 0 | 0 | 0 | 0 | 0 | -0.56 | -0.56 |
| **6** | 0 | 0 | 0 | 0 | 0 | 0 | -0.18 | -0.18 |
| **7** | 0 | 0 | 0 | 0 | 0 | 0 | -0.31 | -0.31 |
| **8** | 0 | 0 | 0 | 0 | -0.7 | 0.12 | 0 | -0.58 |
| **9** | 0 | 0 | 0 | 0 | -0.7 | 0 | -0.18 | -0.88 |
| **10** | 0 | 0 | 0 | 0.71 | 0 | 0 | -0.7 | 0.01 |
| **11** | 0 | 0 | 0 | 0.12 | 0 | 0 | -1.85[a] | -1.73 |
| **12** | 0 | 0 | 0 | 0 | 0 | 0 | -1.38[a] | -1.38 |
| **13** | 0 | 0 | 0 | 0.12 | 0 | 0 | 0 | 0.12 |
| **14** | 0 | 0 | 0 | 0 | -0.7 | 0 | -0.56 | -1.26 |
| **15** | 0 | 0 | 0 | 0 | -0.7 | 0 | -0.7 | -1.40 |
| **16** | 0.12 | 0 | -0.7 | 0 | -0.7 | 0 | -0.7 | -1.98 |
| **17** | 0 | 0 | -0.56 | 0 | -0.7 | 0 | -0.7 | -1.96 |
| **18** | 0 | -0.18 | -0.7 | 0 | -0.7 | 0 | -0.56 | -2.14 |
| **19** | 0 | -0.18 | -0.7 | 0 | -0.7 | 0 | 0 | -1.58 |
| **20** | 0 | 0 | 0 | 0 | -0.56 | 0 | 0 | -0.56 |
| **21** | 0 | 0 | -0.7 | 0 | -0.7 | 0 | -0.7 | -2.10 |

[a] through-resonance parameter.

**Table S3.** Electronic energies (in Hartrees) of the species in its normal oxidation state [E(*N*)], reduced state [E(*N*+1)] and oxidized state [E(*N*-1)], Vertical Ionization Potential (VIP) and Vertical Electron Affinity (VEA) in eV, calculated at the *m*PW1PW91/6-31G(d) level. Note that 1 Hartree = 27.2114 eV

| compound | E(*N*) | E(*N*+1) | E(*N*-1) | VEA | VIP |
|---|---|---|---|---|---|
| | | AH$^+$ form | | | |
| **1** | -653.1404 | -653.2778 | -652.8785 | 3.74 | 7.13 |
| **2** | -728.3548 | -728.4860 | -728.1082 | 3.57 | 6.71 |
| **3** | -728.3554 | -728.4860 | -728.1158 | 3.55 | 6.52 |
| **4** | -767.6664 | -767.7957 | -767.4192 | 3.52 | 6.73 |
| **5** | -767.6413 | -767.7732 | -767.3998 | 3.59 | 6.57 |
| **6** | -692.4511 | -692.5869 | -692.1954 | 3.70 | 6.96 |
| **7** | -861.1234 | -861.2572 | -860.8853 | 3.64 | 6.48 |
| **8** | -803.5541 | -803.6852 | -803.3214 | 3.57 | 6.33 |
| **9** | -767.6653 | -767.7951 | -767.4222 | 3.53 | 6.62 |
| **10** | -932.8043 | -932.9453 | -932.5579 | 3.84 | 6.70 |
| **11** | -862.2998 | -862.4224 | -862.0897 | 3.34 | 5.72 |
| **12** | -708.5036 | -708.6256 | -708.2874 | 3.32 | 5.88 |
| **13** | -728.3495 | -728.4847 | -728.1053 | 3.68 | 6.65 |
| **14** | -842.8553 | -842.9812 | -842.6217 | 3.43 | 6.35 |
| **15** | -803.5694 | -803.6942 | -803.3377 | 3.39 | 6.30 |
| **16** | -953.9798 | -954.1028 | -953.7615 | 3.35 | 5.94 |
| **17** | -918.0664 | -918.1889 | -917.8366 | 3.33 | 6.25 |
| **18** | -957.3749 | -957.4940 | -957.1438 | 3.24 | 6.29 |
| **19** | -842.8753 | -842.9986 | -842.6351 | 3.36 | 6.53 |
| **20** | -767.6411 | -767.7728 | -767.3923 | 3.58 | 6.77 |
| **21** | -878.7800 | -878.9018 | -878.5504 | 3.31 | 6.25 |

**Table S3.** (continued)

| compound | E(*N*) | E(*N*+1) | E(*N*-1) | VEA | VIP |
|---|---|---|---|---|---|
| | | B form | | | |
| **1** | -729.1045 | -729.1456 | -728.8868 | 1.12 | 5.92 |
| **2** | -804.3137 | -804.3469 | -804.1101 | 0.90 | 5.54 |
| **3** | -804.3136 | -804.3532 | -804.0990 | 1.08 | 5.84 |
| **4** | -843.6233 | -843.6544 | -843.4195 | 0.85 | 5.55 |
| **5** | -843.6006 | -843.6407 | -843.3851 | 1.09 | 5.86 |
| **6** | -768.4135 | -768.4540 | -768.1966 | 1.10 | 5.90 |
| **7** | -937.0838 | -937.1251 | -936.8675 | 1.12 | 5.89 |
| **8** | -879.5128 | -879.5451 | -879.3111 | 0.88 | 5.49 |
| **9** | -843.6227 | -843.6551 | -843.4194 | 0.88 | 5.53 |
| **10** | -1008.774 | -1008.865 | -1008.551 | 2.48 | 6.06 |
| **11** | -938.2485 | -938.2876 | -938.0578 | 1.06 | 5.19 |
| **12** | -784.4544 | -784.4928 | -784.2548 | 1.05 | 5.43 |
| **13** | -804.3105 | -804.3521 | -804.1071 | 1.13 | 5.53 |
| **14** | -918.8096 | -918.8414 | -918.6078 | 0.87 | 5.49 |
| **15** | -879.5226 | -879.5536 | -879.3212 | 0.84 | 5.48 |
| **16** | -1029.937 | -1029.951 | -1029.753 | 0.36 | 5.01 |
| **17** | -994.0181 | -994.0477 | -993.8189 | 0.81 | 5.42 |
| **18** | -1033.324 | -1033.348 | -1033.125 | 0.67 | 5.41 |
| **19** | -918.8280 | -918.8534 | -918.6285 | 0.69 | 5.43 |
| **20** | -843.6008 | -843.6354 | -843.3961 | 0.94 | 5.57 |
| **21** | -954.7312 | -954.7591 | -954.5327 | 0.76 | 5.40 |

**Table S3.** (continued)

| compound | E(*N*) | E(*N*+1) | E(*N*-1) | VEA | VIP |
|---|---|---|---|---|---|
| $C_Z$ form | | | | | |
| **1** | -729.0942 | -729.1700 | -728.8777 | 2.06 | 5.89 |
| **2** | -804.3046 | -804.3791 | -804.0990 | 2.03 | 5.59 |
| **3** | -804.3055 | -804.3751 | -804.0902 | 1.89 | 5.86 |
| **4** | -843.6123 | -843.6845 | -843.4059 | 1.96 | 5.62 |
| **5** | -843.5914 | -843.6635 | -843.3747 | 1.96 | 5.90 |
| **6** | -768.4038 | -768.4771 | -768.1878 | 2.00 | 5.88 |
| **7** | -937.0749 | -937.1484 | -936.8589 | 2.00 | 5.88 |
| **8** | -879.5042 | -879.5784 | -879.2972 | 2.02 | 5.63 |
| **9** | -843.6139 | -843.6873 | -843.4088 | 2.00 | 5.58 |
| **10** | -1008.767 | -1008.856 | -1008.537 | 2.43 | 6.27 |
| **11** | -938.2414 | -938.3069 | -938.0447 | 1.78 | 5.35 |
| **12** | -784.4482 | -784.5133 | -784.2436 | 1.77 | 5.57 |
| **13** | -804.2994 | -804.3746 | -804.0999 | 2.05 | 5.43 |
| **14** | -918.8022 | -918.8727 | -918.5984 | 1.92 | 5.55 |
| **15** | -879.5155 | -879.5849 | -879.3119 | 1.89 | 5.54 |
| **16** | -1029.927 | -1029.987 | -1029.739 | 1.63 | 5.13 |
| **17** | -994.0112 | -994.0797 | -993.8086 | 1.86 | 5.51 |
| **18** | -1033.317 | -1033.382 | -1033.112 | 1.75 | 5.57 |
| **19** | -918.8195 | -918.8896 | -918.6141 | 1.91 | 5.59 |
| **20** | -843.5914 | -843.6666 | -843.3853 | 2.05 | 5.61 |
| **21** | -954.7246 | -954.7926 | -954.5223 | 1.85 | 5.51 |

**Table S3.** (continued)

| compound | E(*N*) | E(*N*+1) | E(*N*-1) | VEA | VIP |
|---|---|---|---|---|---|
| | | $C_E$ form | | | |
| **1** | -729.1064 | -729.1924 | -728.8869 | 2.34 | 5.97 |
| **2** | -804.3178 | -804.3990 | -804.1101 | 2.21 | 5.65 |
| **3** | -804.3178 | -804.3993 | -804.1002 | 2.22 | 5.92 |
| **4** | -843.6148 | -843.6915 | -843.4069 | 2.09 | 5.66 |
| **5** | -843.5765 | -843.6854 | -843.3842 | 2.96 | 5.23 |
| **6** | -768.4160 | -768.5008 | -768.1970 | 2.31 | 5.96 |
| **7** | -937.0871 | -937.1723 | -936.8688 | 2.32 | 5.94 |
| **8** | -879.5171 | -879.5991 | -879.3083 | 2.23 | 5.68 |
| **9** | -843.6273 | -843.7075 | -843.4201 | 2.18 | 5.64 |
| **10** | -1008.777 | -1008.870 | -1008.548 | 2.53 | 6.23 |
| **11** | -938.2540 | -938.3326 | -938.0581 | 2.14 | 5.33 |
| **12** | -784.4604 | -784.5384 | -784.2567 | 2.12 | 5.54 |
| **13** | -804.3115 | -804.3979 | -804.1078 | 2.35 | 5.54 |
| **14** | -918.8155 | -918.8934 | -918.6099 | 2.12 | 5.59 |
| **15** | -879.5289 | -879.6059 | -879.3230 | 2.10 | 5.60 |
| **16** | -1029.936 | -1030.007 | -1029.742 | 1.94 | 5.28 |
| **17** | -994.0235 | -994.0987 | -993.8204 | 2.05 | 5.53 |
| **18** | -1033.319 | -1033.389 | -1033.112 | 1.91 | 5.62 |
| **19** | -918.8214 | -918.8952 | -918.6138 | 2.01 | 5.65 |
| **20** | -843.6043 | -843.6864 | -843.3960 | 2.23 | 5.67 |
| **21** | -954.7372 | -954.8117 | -954.5342 | 2.03 | 5.52 |

**Table S4.** Core-Electron Binding Energies (in eV) of some selected compounds compiled from NIST X-ray Photoelectron Spectroscopy Database[a] and values estimated based on the Koopmans' theorem, calculated at the *m*PW1PW91/6-31G(d) level. The asterisk (*) in XPS formula indicates the specific atom which is being observed

| Compound | XPS Formula | Experimental C(1s) | Calculated C(1s) |
|---|---|---|---|
| benzene | (C*H) | 284.8 | 274.8 |
| toluene | $C_6H_5C^*H_3$ | 284.7 | 274.9 |
| toluene | $C_6H_5CH_3$ (C*-CH$_3$) | 285.1 | 274.9 |
| toluene | $C_6H_5CH_3$ (C*H) | 285.0 | 274.8 |
| naphthalene | (C*H) | 284.3 | 274.8 |
| trimethoxymethane | HC*(OCH$_3$)$_3$ | 289.7 | 279.8 |
| benzonitrile | (C*H) | 285.4 | 275.5 |
| acetonitrile | C*H$_3$CN | 286.3 | 276.4 |
| acetonitrile | CH$_3$C*N | 287.2 | 276.8 |
| methanol | C*H$_3$OH | 286.8 | 276.7 |
| phenol | $C_6H_5OH$ (C*H) | 284.9 | 274.7 |
| phenol | $C_6H_5OH$ (C*-OH) | 286.7 | 276.7 |
| aniline | $C_6H_5NH_2$ (C*H) | 284.8 | 274.5 |
| aniline | $C_6H_5NH_2$ (C*-NH$_2$) | 285.8 | 275.8 |
| hexafluorobenzene | (C*F) | 288.7 | 278.9 |
| anisole | $C_6H_5OCH_3$ (C*H) | 284.9 | 274.9 |
| acetophenone | $C_6H_5C(=O)CH_3$ (C*H) | 285.1 | 275.1 |
| nitrobenzene | $C_6H_5NO_2$ (C*H) | 285.3 | 275.5 |
| **Compound** | **XPS Formula** | **Experimental O(1s)** | **Calculated O(1s)** |
| methanol | CH$_3$O*H | 539.1 | 518.4 |
| 2-propanol | (O*H) | 538.5 | 517.9 |
| cyclohexanol | (O*H) | 538.4 | 517.8 |
| formic acid | HC(=O)O*H | 540.7 | 519.8 |
| trifluoroacetic acid | CF$_3$C(=O)O*H | 541.3 | 520.3 |
| phenol | C$_6$H$_5$O*H | 539.2 | 518.6 |
| methyl formate | HC(=O)O*-CH$_3$ | 539.9 | 519.8 |
| anisole | (C-O*-C) | 538.7 | 518.7 |
| acetophenone | (C=O*) | 532.3 | 516.1 |
| nitrobenzene | C$_6$H$_5$NO*$_2$ | 533.6 | 518.9 |
| **Compound** | **XPS Formula** | **Experimental N(1s)** | **Calculated N(1s)** |
| benzonitrile | (CN*) | 399.2 | 385.4 |
| acetonitrile | (CN*) | 399.4 | 385.3 |
| aniline | C$_6$H$_5$N*H$_2$ | 400.3 | 387.6 |
| nitrobenzene | C$_6$H$_5$N*O$_2$ | 405.9 | 394.8 |
| trimethylamine N-oxide | (CH$_3$)$_3$N*O | 403.0 | 391.5 |

[a] NIST X-ray Photoelectron Spectroscopy Database, NIST Standard Reference Database Number 20, National Institute of Standards and Technology, Gaithersburg MD, 20899 (2000).

**Table S5.** Core-Electron Binding Energies (in eV) of the most relevant centers in flavylium salts (see the structure in Scheme 1), taken from Natural Population Analysis calculated at the *m*PW1PW91/6-31G(d) level

| | | $AH^+$ form | | $C_Z$ form | |
|---|---|---|---|---|---|
| compound | OH(1s) | O1(1s) | C2(1s) | C2'(1s) | C6'(1s) |
| **1** | — | 520.99 | 279.09 | 274.91 | 275.01 |
| **2** | 519.50 | 520.82 | 278.87 | 274.88 | 274.98 |
| **3** | 519.35 | 520.79 | 278.92 | 274.95 | 275.08 |
| **4** | 519.46 | 520.75 | 278.81 | 274.87 | 274.97 |
| **5** | — | 520.83 | 278.96 | 274.95 | 275.09 |
| **6** | — | 520.94 | 279.05 | 274.86 | 274.97 |
| **7** | — | 520.84 | 278.96 | 274.98 | 275.06 |
| **8** | 519.55 | 520.86 | 278.89 | 274.88 | 274.98 |
| **9** | 519.48 | 520.78 | 278.84 | 274.83 | 274.94 |
| **10** | 519.49 | 521.02 | 279.16 | 275.01 | 275.13 |
| **11** | 519.08 | 520.54 | 278.62 | 274.77 | 274.87 |
| **12** | — | 520.51 | 278.64 | 274.79 | 274.90 |
| **13** | 519.23 | 520.97 | 278.99 | 274.91 | 275.01 |
| **14** | 519.44 | 520.73 | 278.78 | 274.95 | 275.09 |
| **15** | 519.42 | 520.69 | 278.74 | 274.92 | 275.05 |
| **16** | 519.33 | 520.69 | 278.70 | 274.97 | 275.04 |
| **17** | 519.44 | 520.65 | 278.69 | 274.91 | 275.04 |
| **18** | 519.37 | 520.54 | 278.57 | 274.94 | 275.08 |
| **19** | 519.43 | 520.63 | 278.65 | 274.87 | 274.98 |
| **20** | — | 520.89 | 278.90 | 274.88 | 274.98 |
| **21** | 519.40 | 520.63 | 278.67 | 274.91 | 275.04 |

**Table S6.** Core-Electron Binding Energies (in eV) of the most relevant centers in natural anthocyanins (see the structure in Scheme 1), taken from Natural Population Analysis calculated at the *m*PW1PW91/6-31G(d) level

| | $AH^+$ form | | | $C_Z$ form | |
|---|---|---|---|---|---|
| compound | OH(1s) | O1(1s) | C2(1s) | C2'(1s) | C6'(1s) |
| malvin | 519.38 | 520.84 | 278.82 | 274.98 | 274.98 |
| cianin | 519.33 | 520.79 | 278.75 | 274.88 | 274.91 |
| oenin | 519.39 | 520.81 | 278.80 | 274.98 | 274.91 |
| peonin | 519.38 | 520.82 | 278.77 | 274.97 | 274.97 |
| kuromanin | 519.37 | 520.76 | 278.73 | 274.89 | 274.91 |
| callistephin | 519.38 | 520.77 | 278.73 | 274.99 | 275.03 |
| hibiscin | 519.32 | 520.79 | 278.82 | 274.91 | 274.91 |
| myrthillin | 519.38 | 520.77 | 278.73 | 274.78 | 274.78 |
| petunidin 3-*O*-glucoside | 519.38 | 520.79 | 278.76 | 274.86 | 274.84 |
| oxycoccicyanin | 519.38 | 520.79 | 278.75 | 274.97 | 274.97 |

**Table S7.** Hammett $\sigma$ parameters for the substituents of the natural anthocyanins studied. The $\sigma$-values were compiled from reference 16

| compound[a] | C3<br>($\sigma_m$) | C5<br>($\sigma_R$) | C7<br>($\sigma_R$) | C3’<br>($\sigma_m$) | C4’<br>($\sigma_R$) | C5'<br>($\sigma_m$) | $\Sigma\sigma_X$ |
|---|---|---|---|---|---|---|---|
| malvin | -0.1 | -0.18 | -0.7 | 0.12 | -0.7 | 0.12 | -1.44 |
| cianin | -0.1 | -0.18 | -0.7 | 0.12 | -0.7 | 0 | -1.56 |
| oenin | -0.1 | -0.7 | -0.7 | 0.12 | -0.7 | 0.12 | -1.96 |
| peonin | -0.1 | -0.18 | -0.7 | 0.12 | -0.7 | 0 | -1.56 |
| kuromanin | -0.1 | -0.7 | -0.7 | 0.12 | -0.7 | 0 | -2.08 |
| callistephin | -0.1 | -0.7 | -0.7 | 0 | -0.7 | 0 | -2.20 |
| hibiscin | -0.1 | -0.7 | -0.7 | 0.12 | -0.7 | 0.12 | -1.96 |
| myrthillin | -0.1 | -0.7 | -0.7 | 0.12 | -0.7 | 0.12 | -1.96 |
| petunidin 3-*O*-glucoside | -0.1 | -0.7 | -0.7 | 0.12 | -0.7 | 0.12 | -1.96 |
| oxycoccicyanin | -0.1 | -0.7 | -0.7 | 0.12 | -0.7 | 0 | -2.08 |

[a] *O*-glucoside and *O*-sambubioside units were represented with electronic and steric parameters ($E_S$) of the *t*-butyl group ($\sigma_R$ = -0.18; $\sigma_m$ = -0.1; $E_S$ = -1.78)

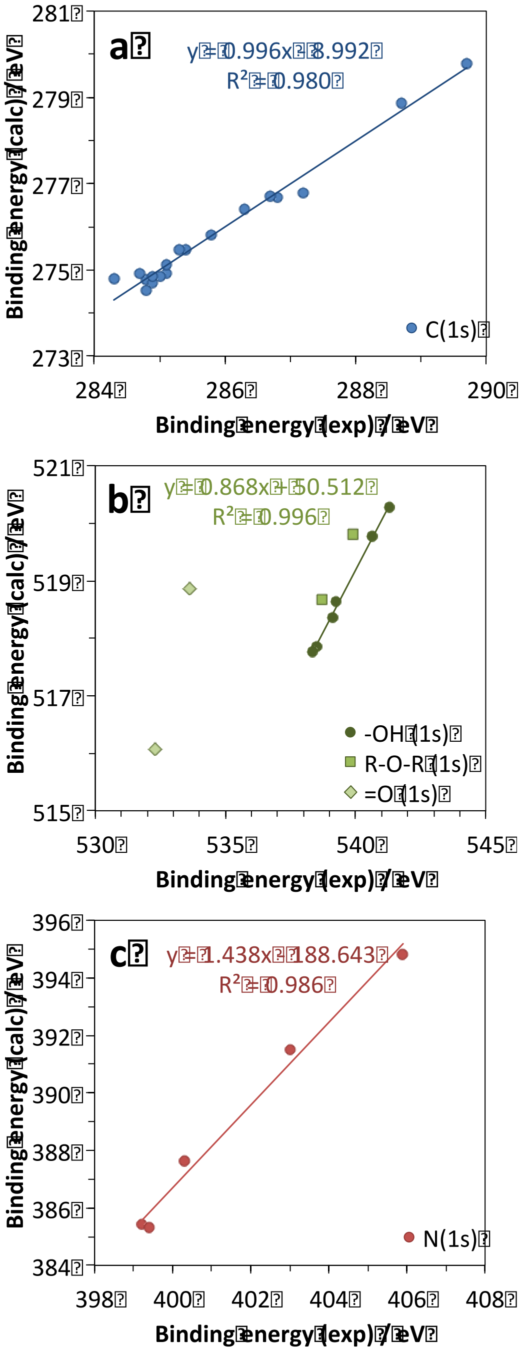

**Figure S1.** Experimental binding energies (orbital 1s) and the corresponding calculated values at the *m*PW1PW91/6-31G(d) level with solvent effects for (a) carbon, (b) oxygen and (c) nitrogen data on Table S4.

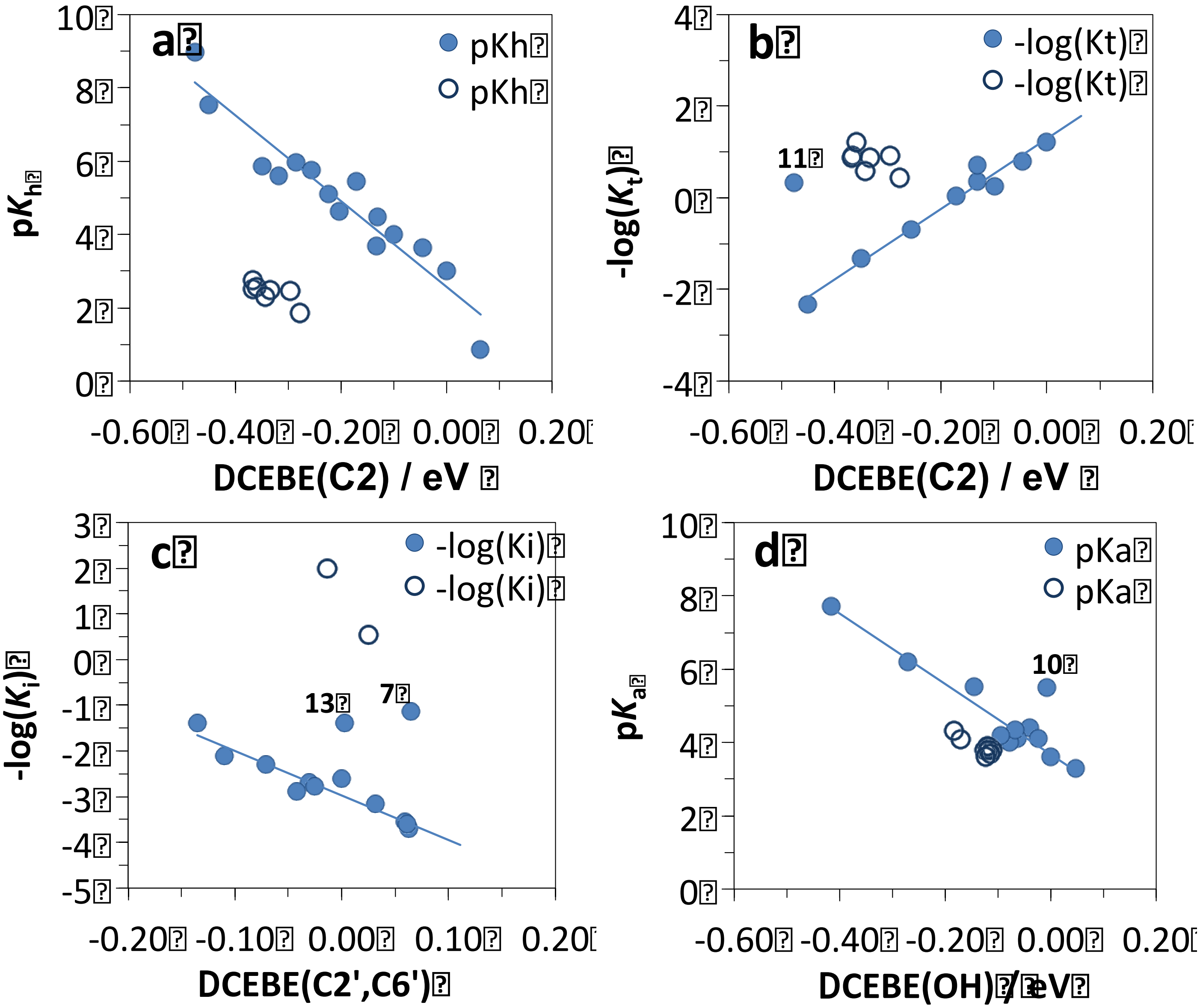

**Figure S2.** Representations of (a) $pK_h$ and (b) $–log(K_t)$ as a function of the difference between the core-electron binding energy of C2(1s) of $AH^+$ and the corresponding value for the unsubstituted compound **1**, ΔCEBE; (c) $–log(K_i)$ as a function of the average ΔCEBE of C2'(1s) and C6'(1s) 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 (full symbols) and anthocyanins (open symbols) were determined from Natural Population analysis at the *m*PW1PW91/6-31G(d) level of theory.