

Supplementary Information

Achieving a Dramatic Blue Colour Stability in Anthocyanins Bearing Acylated Sugars in position 3',5'.

Joana Oliveira,*^a João T. S. Coimbra,^a Victor Freitas,^a Peiqing Yang,^{b,c} Nuno Basilio,^b
Fernando Pina *^b

Section 1 – Extraction, isolation and structural characterization of the tri-acylated anthocyanin from *Clitoria ternatea*.

Polyacylated anthocyanins were extracted from butterfly pea flowers (*Clitoria ternatea*) following the methodology described by Pereira *et al.* (2024)^[1]. The anthocyanin-rich extract was then purified by protein precipitation using trichloroacetic acid (TCA). For this, a 20% (w/v) TCA solution was prepared in acetone, and 50 mL of this solution was added to 50 mL of the anthocyanin extract, resulting in a final TCA concentration of 10% (w/v). The mixture was stored at -20 °C for 2 hours to allow protein precipitation, followed by centrifugation at 14,000 rpm for 10 minutes at 4 °C. The supernatant was collected, and acetone was removed using a rotary evaporator under reduced pressure at 37 °C. The resulting aqueous phase was subjected to C18 solid-phase extraction (SPE), where TCA, sugars, and other small polar compounds were eluted with water. Phenolic compounds were subsequently eluted with methanol acidified with a few drops of 2% HCl.

The purified extract was then fractionated by preparative HPLC using a Thermo Scientific Ultimate 3000 apparatus on a Luna reverse-phase pentafluorophenyl (PFP2) column (5 µm particle size, 100 Å pore size, 250 × 21.2 mm; Phenomenex, Torrance, CA, USA). Chromatographic separation was carried out at a flow rate of 10.0 mL/min over a 30-minute run. A binary solvent system was used: solvent A (4.5% formic acid in HPLC-grade water) and solvent B (HPLC-grade acetonitrile). The gradient started at 15% B and increased linearly to 30% B from 0 to 30 minutes^[2]. After collection, residual acetonitrile was removed under vacuum at 37 °C using a rotary evaporator. The resulting pigment fractions were further purified via C-18 SPE to remove excess formic acid. Finally, purified pigments were freeze-dried and stored at -20 °C until further analysis.

The structural elucidation of the tri-acylated anthocyanin (Pigment 1) was carried out using LC/DAD-MS and comprehensive NMR spectroscopy, including 1D (^1H) and 2D (gCOSY, TOCSY, NOESY, HSQC, and HMBC) experiments. Electrospray ionization mass spectrometry in positive ion mode revealed a molecular ion at m/z 1551, with fragment ions at m/z 1389, 1227, 1081, and 919. These fragments correspond to successive losses of glucose (-162 Da) and coumaroyl (-146 Da) moieties. Comparison with previously reported data suggested two possible structural isomers for this tri-acylated delphinidin derivative. To unambiguously determine the structure, detailed analysis using 2D NMR techniques was performed. The complete ^1H and ^{13}C chemical shift assignments are summarized in Table S1. Long-range correlations (HMBC) between the anomeric protons of each glucose moiety and carbons from either the flavylium cation core or the coumaroyl group enabled the differentiation of the glucose units and their specific positions within the anthocyanin molecule. Additionally, HMBC correlations involving the H-6 protons of the glucose moieties facilitated the identification of glucose units that were either free or esterified with coumaroyl groups. The TOCSY experiment was also crucial for assigning and distinguishing the proton signals of all glucose moieties.

Table S1 – ^1H and ^{13}C chemical shifts of pigment 1, determined in methanol d_4 .

Position	$\delta^1\text{H}$ (ppm); J (Hz)	$\delta^{13}\text{C}$ (ppm)	HMBC correlations	NOESY correlations
Flavylium cation core				
2C		159.3	H-4	
3C		144.3	H-4; H-1-1	
4C	8.35; s	134.6		H-1 Glc 1
4aA		112.4	H-4; H-6; H-8	
5A		157.3	H-4; H-6	
6A	6.54; <i>d</i> , 1.8	103.0		
7A		169.4	H-6; H-8	
8A	6.63; <i>d</i> , 1.3 (*)	94.3		
8aA		155.2	H-4; H-8	
1'B		n.a.		
2'B	7.93, bs	112.4		H-1 Glc 2
3'B		145.8	H-1-2	
4'B		n.a.		
5'B		145.8	H-1-4	
6'B	7.99, bs	n.a.		H-1 Glc 4
3 – glucose (moiety 1)				
1	4.73; <i>d</i> , 7.5	102.5		
2	3.71; *	76.1		
3	3.68; *	73.1		
4				
5	3.87; *	74.9	H-6a Glc 1	
6a	4.30; *	63.2		

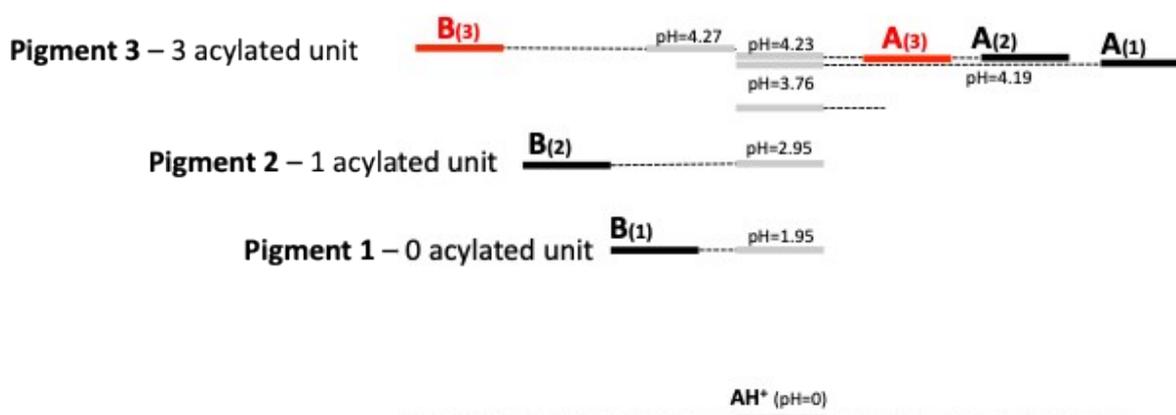
6b				
3' – glucose (moiety 2)				
1	5.27; <i>d</i> , 6.8	99.2		
2		73.2		
3				
4				
5				
6a	4.76; *			
6b	4.30; *			
3' – glucose (moiety 3)				
1	4.93; <i>d</i> , 7.8	99.8		
2		73.4		
3				
4				
5				
6a	4.66; <i>dd</i> , 11.8/2.4	63.4		
6b	4.26; *	63.4		
5' – glucose (moiety 4)				
1	5.32; <i>d</i> , 7.1	100.3		
2		73.2		
3				
4				
5				
6a	4.78; *	63.4		
6b	4.26; *	63.4		
5' – glucose (moiety 5)				
1	4.97; <i>d</i> , 7.4	99.8		
2	3.56;*	73.5		
3				
4				
5				
6a	4.05	61.1		
6b	3.82	61.1		
Coumaroyl unit (ring D)				
H- α	5.99; <i>d</i> , 15.7	115.7		
H- β	7.14; <i>d</i> , 15.7	143.7		
C=O		166.7	H-6a Glc 2	
1		127.5	H-2,6 (ring D)	
2,6	6.81; <i>d</i> , 8.4	129.2		
3,5	6.87; <i>d</i> , 8.4	116.5		
4		158.7	H-3,5 (ring D)	
Coumaroyl unit (ring E)				
H- α	6.20; <i>d</i> , 15.9	113.5		H-8 (flavylium core)
H- β	7.40; <i>d</i> , 15.9	145.3		
C=O		167.4	H-6a Glc 3	
1		125.4	H α , H- β (ring E)	
2,6	7.15; <i>d</i> , 8.5	129.7		H-8 (flavylium core)
3,5	6.65; <i>d</i> , 8.5	115.7		
4		160.0	H-3,5 (ring E)	
Coumaroyl unit (ring F)				
H- α	5.75; <i>d</i> , 15.7	115.5		
H- β	7.06; <i>d</i> , 15.7	143.7		
C=O		166.8	H-6a Glc 4	
1		125.9	H α , H- β , H-3,5 (ring	

			F)	
2,6	6.96; d, 8.4	129.0		H-8 (flavylium core)
3,5	6.82; d, 8.4	116.2		H-1 Glc 5; H-6 (flavylium core)
4		158.9	H-3,5 (ring F)	

s, singlet; d, doublet; bs, broad singlet; n.a., not assigned; *, unresolved.

Section 2

In Scheme S1, the relative energy level diagram of three anthocyanins extracted from *Ipomoea tricolor*, as a function of the number of acylated units, is presented.^[3] An identical behaviour was reported for the acylated anthocyanins extracted from *Pharbitis ipomoea nil*.^[4]



Scheme S1. Energy level diagrams of anthocyanins extracted from *Ipomoea tricolor* in acidic medium: black, 3 acylated units (see scheme 2 of the main text); blue, 2 acylated units; red, no acylated unit. The energy level of Ct cannot be measured with the necessary accuracy due to some decomposition in HBA 2 and HBA 3 for longer reaction times.

The energy level of the quinoidal base of the *Ipomoea tricolor* derivative with none ($pK_a=4.19$) or two acylated units ($pK_a=4.23$), Scheme 1, does not change significantly (falls within the experimental error). In the case of the derivative bearing three acylated units, there is a slight stabilization. In contrast, the colourless hemiketal (like in the case of *Pharbitis ipomoea nil*) increases significantly its energy level with the number of acylated units, with inversion of the relative stability for the species bearing 3 acylated units. Both effects, the decrease of $pK_{A/A-}$ (commonly identified by pK_a) and the increase of pK_h (referring to the equilibrium between flavylum cation and hemiketal, see Scheme 1), contribute to the increase of the quinoidal base colour intensity and its extension to higher pHs. Nevertheless, the most significant effect is the decrease in the stability of the hemiketal.

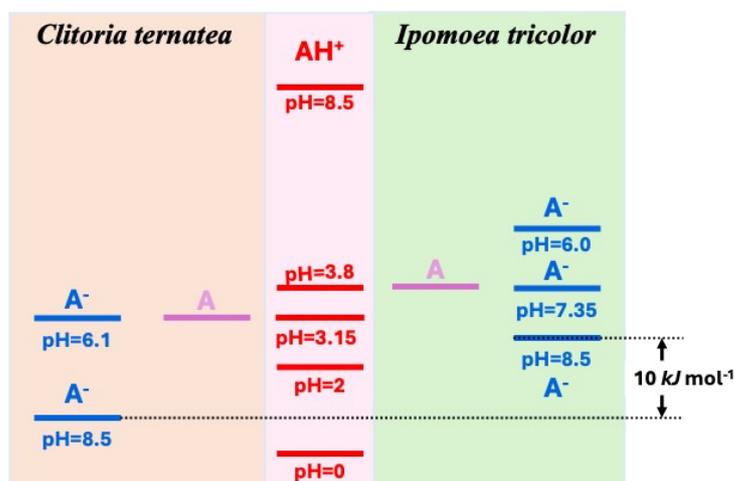
This thermodynamic effect is related to the kinetics, Table S2; and increasing the pK_h corresponds to the decrease of the respective equilibrium constant (for hemiketal $K_h=k_h/k_{-h}$). In fact, in the case of three acylated units k_h decrease while k_{-h} increase. The hydration increase is explained by the protection effect to the water attack. On the other hand, the Gibbs free energy (ΔG^0) for the dehydration becomes more favourable, explaining the increase of k_{-h} . This effect

presupposes that the activation energy does not change significantly with the increasing number of acylated units.

	k_h / s^{-1}	$k_{-h} / \text{M}^{-1} \text{s}^{-1}$
3 acylated	0.01	377
2 acylated	0.12	145
0 acylated	0.35	43

Regarding the relative stability of the blue anionic quinoidal base (the dominant species at neutral and moderately basic solutions), it is necessary to extend the energy level diagram to this species.^[5]

In Scheme S2, the relative energy level of the coloured species is shown. There are two sources of stability of the blue anionic quinoidal base in *Clitoria ternatea* compared with *Ipomoea tricolor*. On one hand, the energy level of the quinoidal base is lower in the first and on the other hand, deprotonation to give the anionic quinoidal base is easier. For example, at pH =8.5 the energy level of *Clitoria ternatea* is 10 kJ·mol⁻¹ more stable than the corresponding energy level of *Ipomoea tricolor*.



Scheme S2. Relative energy level diagram of the neutral and anionic quinoidal bases of *Clitoria ternatea* and *Ipomoea tricolor* in comparison to the respective flavylum cation. The relative energy levels of the two flavylum cations were superimposed to allow comparison between the relative stability of the quinoidal bases.

In Fig. S1, the spectral variations taken immediately after direct pH jumps (addition of base to equilibrated solutions of the flavylum cation) are presented. It is worth of note an increasing on the absorption at 380 nm in Fig. S1d that could be attributed to the acyl deprotonation.

The evolution of the blue chromophore ($\lambda_{\max} \approx 620$ nm) towards less vivid hues at higher pH, as shown in Fig. S1, is accompanied by an increase in absorbance near 380 nm, indicating the progressive deprotonation of the terminal acyl group (ring E). This deprotonation likely disrupts π - π stacking interactions involving the acyl moieties, as both the terminal acyl unit and the interacting phenolic rings acquire negative charges. The observation of long-range NOE correlations between ring A and rings E and F in CD₃OD confirms that pigment 1 adopts compact folded conformations even in the absence of a hydrophobic effect, further supporting the role of intramolecular copigmentation in stabilising the chromophore.

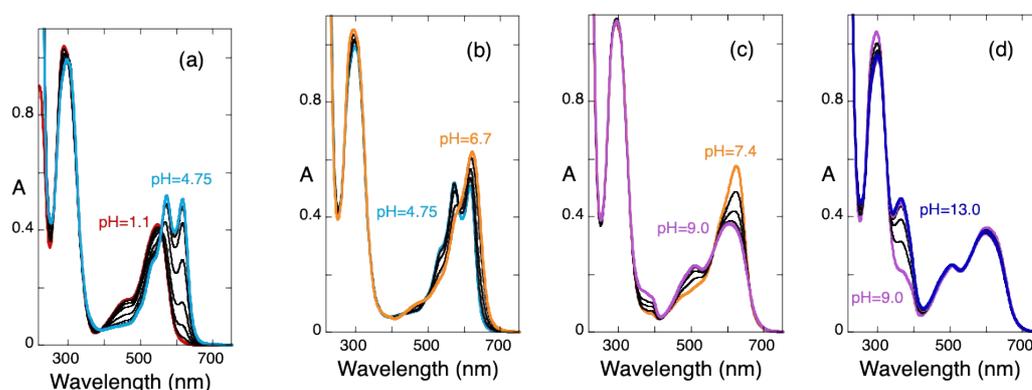


Figure S1. Absorption spectra taken immediately after a direct pH jump of solutions of pigment 1 versus pH. All experiments were carried out in the dark.

Section 3 – Conformational Analysis of Tri-Acylated Anthocyanins: An *In Silico* Study

CREST conformational search and optimization

CREST^[6] was used to generate conformers for two anthocyanin molecules – pigment 1 from *Clitoria ternatea* and pigment 2 from *Ipomoea tricolor*^[7]. Their structure was generated using Open Babel^[8] and GaussView,^[9] because no 3D molecular structure was available. For the sugar moieties in both molecules, a β -D-glucose configuration was considered.

Then, a conformational search and sorting was accomplished with the CREST software, using GFN-FF level of theory.^[10] GFN-FF is a completely automated, partially polarizable generic force-field for the accurate description of structures and dynamics of large molecules, combining force-field speed with quantum mechanical accuracy. Conformational search was performed using default settings in methanol implicit solvent to mimic experimental conditions, using Generalized-Born Surface Area (GBSA). Ensemble sorting was performed using an energy threshold of 6 kcal·mol⁻¹, a root-mean-square deviation (RMSD) threshold of 0.25 Å, and an energy threshold of 0.1 kcal·mol⁻¹.

The generated ensemble structures were then reoptimized at the GFN2-xTB level of theory^[11] with GBSA implicit solvation for methanol. GFN2-xTB is a semiempirical tight-binding method, designed for the fast computation of molecular energies of systems encompassing tens or hundreds of atoms. The structures that better reproduced the long-range NOE signals of pigment 1 and pigment 2, were then further analyzed in terms of solvent accessibility at the C2 position of the flavylium ion, and for the overall molecular stacking of the anthocyanin molecules. Solvent-accessible surface area (SASA) calculations were performed using PyMOL's `get_area` command, with `dot_solvent` set to 1, `dot_density` set to 3, and a solvent radius of 1.4 Å, to ensure accurate surface representation. For molecular visualization, image rendering and analysis, we used Open-Source PyMOL.^[12]

The conformational analysis was based on computer simulations, which cannot fully capture all factors influencing molecular structure and dynamics under experimental conditions. Limitations include the absence of explicit solvent molecules (with methanol modelled implicitly). While CREST is well-suited for extensive conformational sampling, it may still miss relevant conformations. Nonetheless, the integration of NMR data provides additional confidence in the reliability of the *in silico* results. It is important to note that the analysis is based on representative conformations, whereas in solution, the molecule likely exists as an ensemble of interconverting structures.

References

- [1] A. R. Pereira, V. C. Fernandes, C. Delerue-Matos, V. de Freitas, N. Mateus, J. Oliveira, *Food Chem.* 2024, *461*, 140945.
- [2] G. T. Sigurdson, R. J. Robbins, T. M. Collins, M. M. Giusti, *Food Chem.* 2017, *221*, 1088-1095.
- [3] J. Mendoza, N. Basilio, F. Pina, T. Kondo, K. Yoshida, *J. Phys. Chem. B* 2018, *122*, 4982-4992.
- [4] aO. Dangles, N. Saito, R. Brouillard, *J. Am. Chem. Soc.* 1993, *115*, 3125-3132; bN. B. L. Cruz, N. Mateus, V. de Freitas, F. Pina, *Chemical Reviews* 2022, *122*, 1416-1481.
- [5] aF. Pina, M. J. Melo, M. Maestri, R. Ballardini, V. Balzani, *J. Am. Chem. Soc.* 1997, *119*, 5556-5561; bN. Basilio, F. Pina, *Molecules* 2016, *21*, 25.
- [6] aP. Pracht, F. Bohle, S. Grimme, *Physical Chemistry Chemical Physics* 2020, *22*, 7169-7192; bP. Pracht, S. Grimme, C. Bannwarth, F. Bohle, S. Ehlert, G. Feldmann, J. Gorges, M. Müller, T. Neudecker, C. Plett, S. Spicher, P. Steinbach, P. A. Wesolowski, F. Zeller, *The Journal of Chemical Physics* 2024, *160*.
- [7] T. Goto, T. Kondo, *Angewandte Chemie International Edition in English* 1991, *30*, 17-33.
- [8] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, *Journal of Cheminformatics* 2011, *3*, 33.
- [9] V. GaussView, Roy Dennington, Todd Keith, and John Millam, Semichem Inc., Shawnee Mission, KS, 2019.
- [10] S. Spicher, S. Grimme, *Angewandte Chemie International Edition* 2020, *59*, 15665-15673.
- [11] C. Bannwarth, S. Ehlert, S. Grimme, *Journal of Chemical Theory and Computation* 2019, *15*, 1652-1671.
- [12] v. S. The PyMOL Molecular Graphics System, LLC., 2019.