Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/7/15/eabe7871/DC1

Supplementary Materials for

Discovery of a natural cyan blue: A unique food-sourced anthocyanin could replace synthetic brilliant blue

Pamela R. Denish, Julie-Anne Fenger, Randall Powers, Gregory T. Sigurdson, Luca Grisanti, Kathryn G. Guggenheim, Sara Laporte, Julia Li, Tadao Kondo, Alessandra Magistrato, Mícheál P. Moloney, Mary Riley, Mariami Rusishvili, Neda Ahmadiani, Stefano Baroni, Olivier Dangles, Monica Giusti, Thomas M. Collins, John Didzbalis, Kumi Yoshida*, Justin B. Siegel*, Rebecca J. Robbins*

*Corresponding author. Email: yoshidak@i.nagoya-u.ac.jp (K.Y.); jbsiegel@ucdavis.edu (J.B.S.); rebecca.robbins@effem.com (R.J.R.)

Published 7 April 2021, *Sci. Adv.* **7**, eabe7871 (2021)
DOI: 10.1126/sciadv.abe7871

This PDF file includes:

Supplementary Section 1: Violet Contribution

Supplementary Method 1.1. Calculating the violet contribution in a dye using UV visible spectrophotometry. Violet contribution to natural blue colors creates a challenge in obtaining a cyan blue. The violet contribution region of the visible light spectrum was defined as the absorbance from 500 nm to 600 nm. Increasing absorbance in the 500 nm to 600 nm range results in a more violet blue, due to the absorbance of green, yellow, and some orange light, whereas decreasing absorbance shifts the color more toward cyan blue. The violet contribution can be quantified using UV visible spectrophotometry. We did so by measuring the UV visible absorption spectrum of a solution, the area under the curve from 500 nm to 600 nm was calculated as the value representing the violet contribution. To ensure differences in absorption maxima, the absorbance at λ_{max} for each needs to be normalized. Normalization calculations are as follows:

$$\widehat{A_\lambda} = A_\lambda * \frac{A_{0\lambda max}}{A_{\lambda max}}, \text{ where}$$

- $\widehat{A_\lambda}$ is the normalized absorbance at wavelength λ
- A_λ is the measured absorbance at wavelength λ
- $A_{0\lambda max}$ is the measured absorbance at λmax for the reference spectrum (normalize other spectra to this spectrum)
- $A_{\lambda max}$ is the measured absorbance at λ_{max}

To compare the violet contribution for any colorant, the area under the visible absorption curve from 500-600 nm is calculated using the left Riemann sum for integration. See Supplementary Figure 1.1 which depicts an example area of the rectangle rule for integration. The width of the rectangles is defined as 1 nm, in other words the resolution of the spectrum. The height is defined as the absorbance

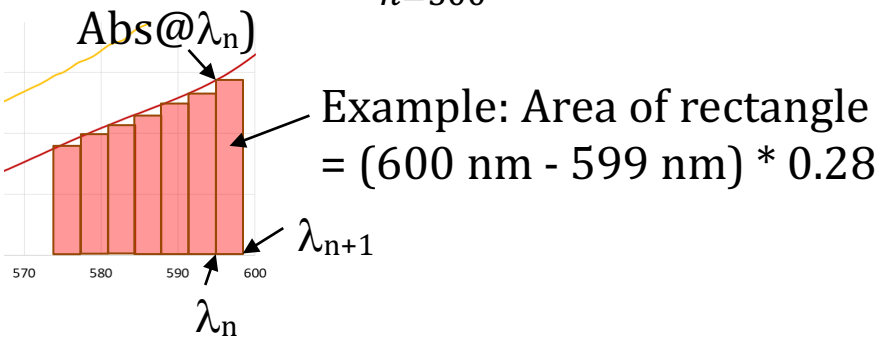
intensity at each wavelength from the normalized spectra. The total area underneath the curve between 500 and 600 nm was the sum of all the rectangles from the wavelength 500 nm to 600 nm. The midpoint rule was considered as an alternate integration method, however the difference in values was inconsequential.

To calculate the area under the curve from 500 nm to 600 nm, the following equation in Supplementary Figure 1.1 is used:

Sum the areas (width x height) of rectangles with:

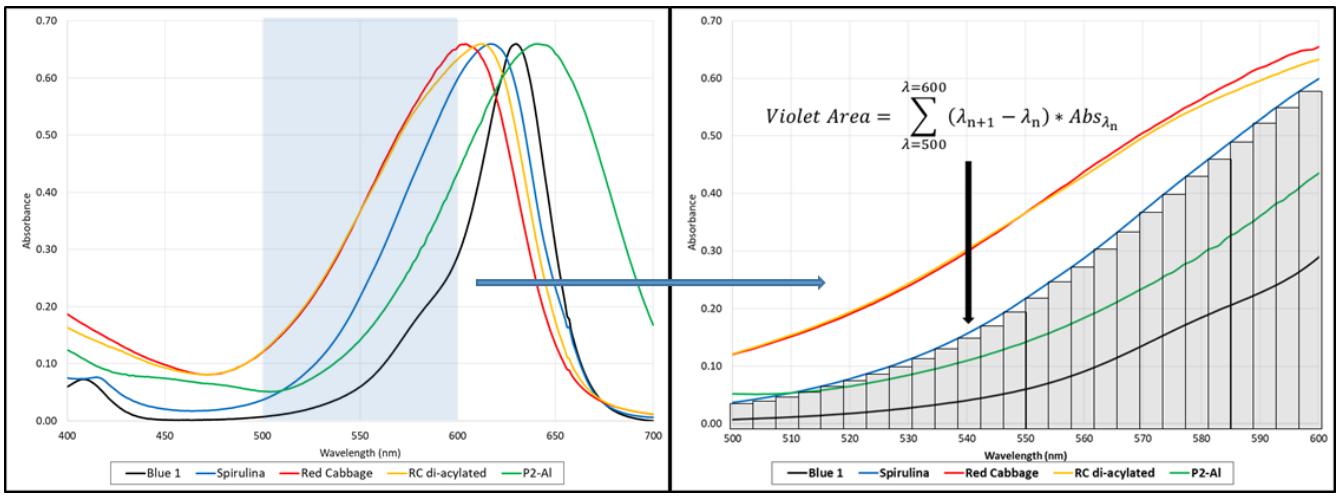
- Width = wavelength resolution, 1 nm
- Height = normalized absorbance at λ_n

$$Violet\ Area = \sum_{n=500}^{n=599} (\lambda_{n+1} - \lambda_n) * Abs(@\lambda_n)$$



Supplementary Figure 1.1. Calculating the violet contribution. The violet contribution of a blue colorant is defined to be the area under the visible absorption curve (AUC) between the wavelength ranges of 500 nm to 600 nm. The area is calculated using the left Riemann sum for integration where the height is the absorbance and width is the spectral resolution. The resolution is the numerical distance between recorded wavelengths (e.g. 600 nm – 599 nm = 1 nm resolution). The area of each rectangle, absorbance x resolution, is summed to produce the AUC, known as the violet contribution. In the example above, the resolution is 1, so the area is 1 * absorbance $\lambda_{n=599}$ = 1 * 0.28 = 0.28. With a

resolution of 1 nm, the AUC is the sum of the absorbance at each wavelength from 500 nm to 599 nm. For violet contribution comparisons, the spectra must be normalized at peak intensity, λ_{max}.

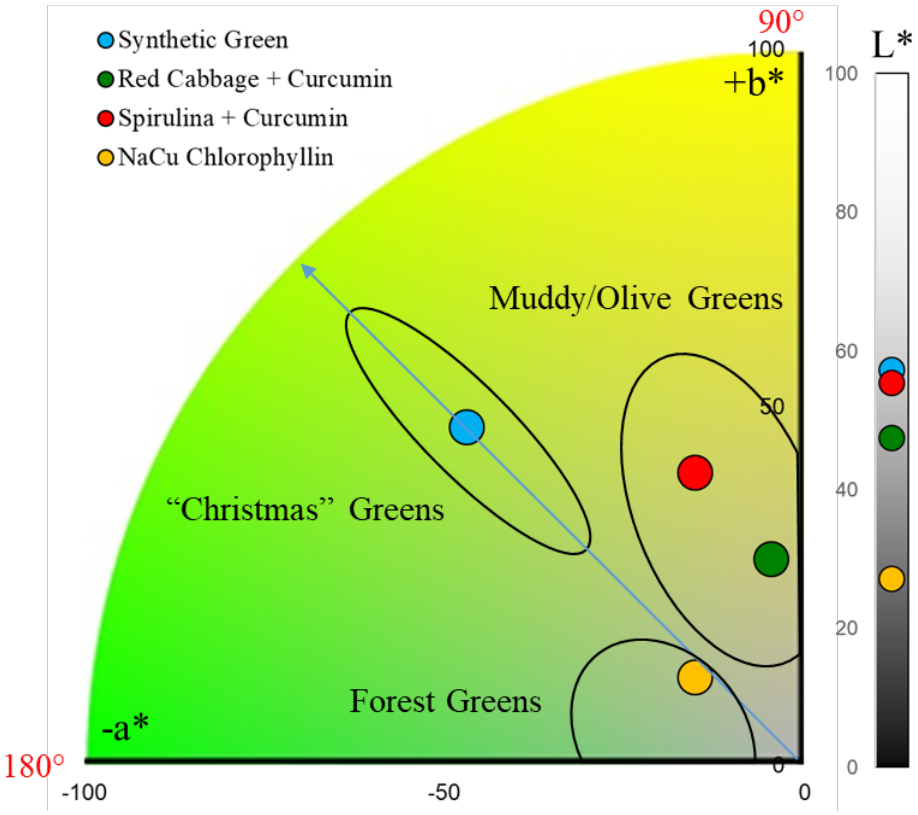


Supplementary Figure 1.2. Visual display and quantitative value of the Violet contribution (VC) of select blue colorants. The violet regions are shown as the area under the curves, for spirulina, di-acylated fraction of red cabbage (the bluer fractions – P6, P7, P8) and whole red cabbage (RC) at pH 8 (the conditions at which RC displays blue hues) and FD&C Blue No. 1, respectively. The calculated amount of the violet contribution exists in all blue colorants and are shown in Supplementary Table 1.1; however, it is the smallest in FD&C Blue No. 1 with a value of 9.2. The next smallest is P2 complex with 1/3 mol equivalents of Al^{3+} ($Al^{3+}(P2^{-})_3$) at pH 7 which has a value of 17.4.

Supplementary Table 1.1. The λ_{max} and the calculated violet contribution of select blue colorants.

Colorant	FD&C Blue No. 1	Spirulina	Red Cabbage pH 8	RC di-acylated pH 8	$Al^{3+}(P2^{-})_3$ pH 7
λ_{max} (nm)	630	617	604	608	640
Violet Contribution	9.2	25.4	37.3	36.9	17.4

Supplementary Section 2. Green a*b* quadrant and chlorophyll.



Supplementary Figure 2.1. Color wheel representation and undertones of select green colorants. In order to help visualize and provide clarity for the different types of green available in nature, this figure shows a portion (quadrant) of the a* b* color space (with L* plotted in the sidebar). The hue angle shown as an arrow depicts the control also labeled range of "Christmas" green with the ellipse. Other green blends shown with the red and green dots depict the approximate space for "muddy/olive" greens as shown in the ellipse. The yellow dot is the commercially available sodium copper chlorophyllin which falls into a general category called forest green primarily due to the lower L* value inherent in the color expression. Color expression data provided in Supplementary Table 2.1. Note: NaCu chlorophyllin is the only naturally sourced colorant is green without blending; the other sources of green are blends of blue and yellow including the control (blend of FD&C Blue No. 1 and FD&C Yellow 5)

Supplementary Table 2.1 Color expression data for varying green colorants

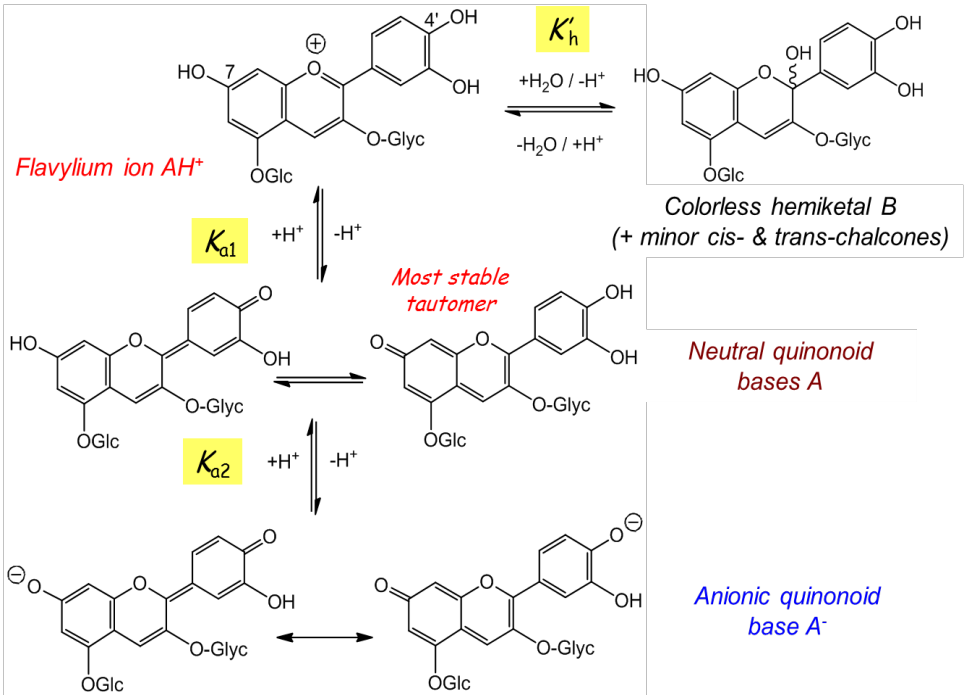
Name	L*	a*	b*	C*	h°
Synthetic Green	57.3	-46.8	47.2	66.5	134.8
Red Cabbage + Curcumin	47.4	-4.6	28.8	29.2	99.1
Spirulina + Curcumin	55.4	-15.2	40.9	43.6	110.4
NaCu Chlorophyllin	27.1	-15.2	12.3	19.5	141.2

Supplementary Discussion 2.1. Limits of chlorophyll as a source of natural green dye.

These naturally-sourced blues when blended with yellows, generate what we might refer to as a 'muddy or olive greens' due to the violet contribution also blending with the yellow. Currently, a green we might refer to as a 'Christmas green" is nearly impossible to achieve (Supplementary Figure 2.1). Additionally, stable and water-soluble natural greens from edible sources are non-existent *(12).* Although green is abundant in nature, the chlorophyll chromophore is not stable nor water soluble, and provides a 'Forest green' color expression which cannot be modified to a more desired green shade. Although chlorophyll is the most widely available pigment in the natural world, being present in all green plants, it has stability and solubility issues as a food colorant. It is prevalent in all green plants since chlorophyll is the key to the process of photosynthesis, absorbing light and releasing electrons that dissociate water and carbon dioxide to produce oxygen and carbohydrate. The green color of chlorophyll is determined by its chemical structure (porphyrin with a magnesium ion at its center). The magnesium ion at the heart of the structure is critical to stability and color expression. If this is removed, the remaining ring structure becomes purple.

Another area of concern is that naturally occurring chlorophyll is oil-soluble and of low stability. In order to generate a stable water-soluble green, the magnesium is replaced by a Cu generating Cu-Chlorophyllin (Cu-Chl - the E number or regulatory label is E-141ii colorant). Since it is perceived as a derived food colorant and is only available for use in specific applications for food, namely citrus beverages. Cu-Chl is commercially available as a complex mixture of different chlorin molecules. A green dye must have two absorption bands, one in the red region and one in the blue. While this characteristic is what makes it an efficient light absorber for the natural setting, this is difficult to achieve often without blending colors for food ingredients. Finally, Cu-Chl provide its own unique shade of green that generally has a low chroma at low concentrations and become very dark as concentration increases providing a green that might be referred to as a 'Forest Green' or 'Dark Forest Green'.

Supplementary Section 3. The structural transformations of anthocyanins in aqueous solution



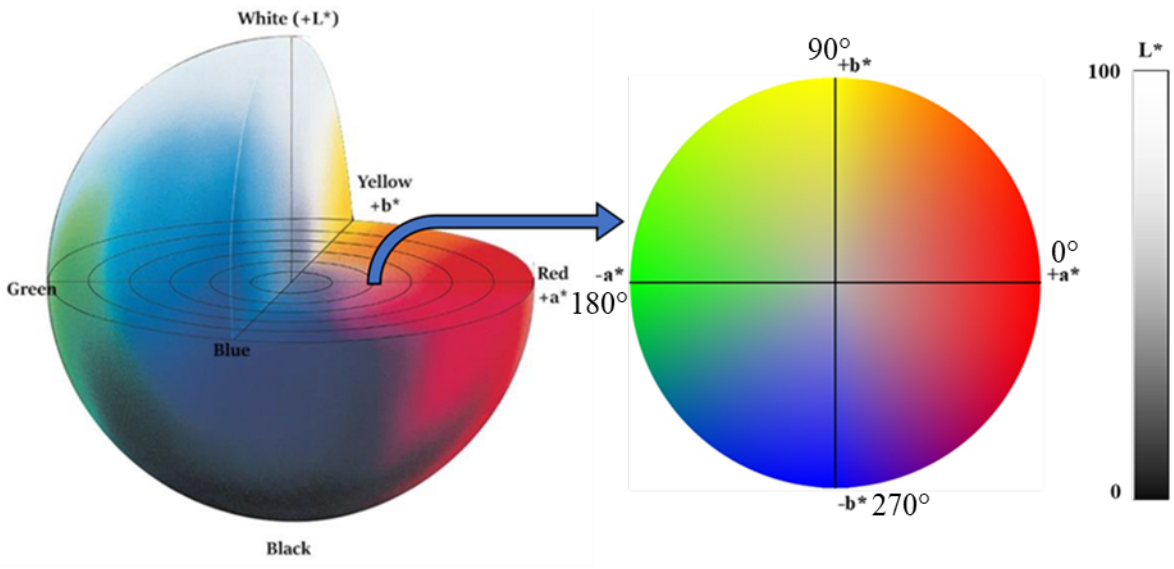
Supplementary Figure 3.1. The structural transformations of anthocyanins in acidic to neutral aqueous solution.

Supplementary Discussion 3.1. The structural transformations of anthocyanins in acidic to neutral aqueous solution. The deep red flavylium cation is stable in highly acidic solution only (pH < 2). When the pH is raised, a first deprotonation in mildly acidic solution (pK_{a1} ≈ 4) leads to a purple neutral quinonoid base, and a second deprotonation around neutrality (pK_{a2} ≈ 7) leads to a more intensely colored purple - blue anionic quinonoid base (Supplementary Figure 3.1) *(14-17)*. The hues expressed are modulated by the B-ring substitution and the glycosylation and acylation patterns. The first deprotonation (fast) is in competition with the nucleophilic attack of water (slow) onto the flavylium ion (hydration reaction), which leads to a colorless hemiketal, itself in equilibrium with usually minor concentrations of pale yellow *cis*- and *trans*-chalcones through C-ring opening and subsequent *cis-trans* isomerization (Supplementary Figure 3.1). The global equilibrium connecting the flavylium ion and the pool of colorless forms (hemiketal + chalcones) is characterized by the apparent hydration constant K'_h. With only a few exceptions, hydration is thermodynamically more favorable than the first proton loss, *i.e.* pK'_h < pK_{a1}. Hence, in equilibrated mildly acidic solution (pH 3 – 6), anthocyanins are mostly under the hemiketal form.

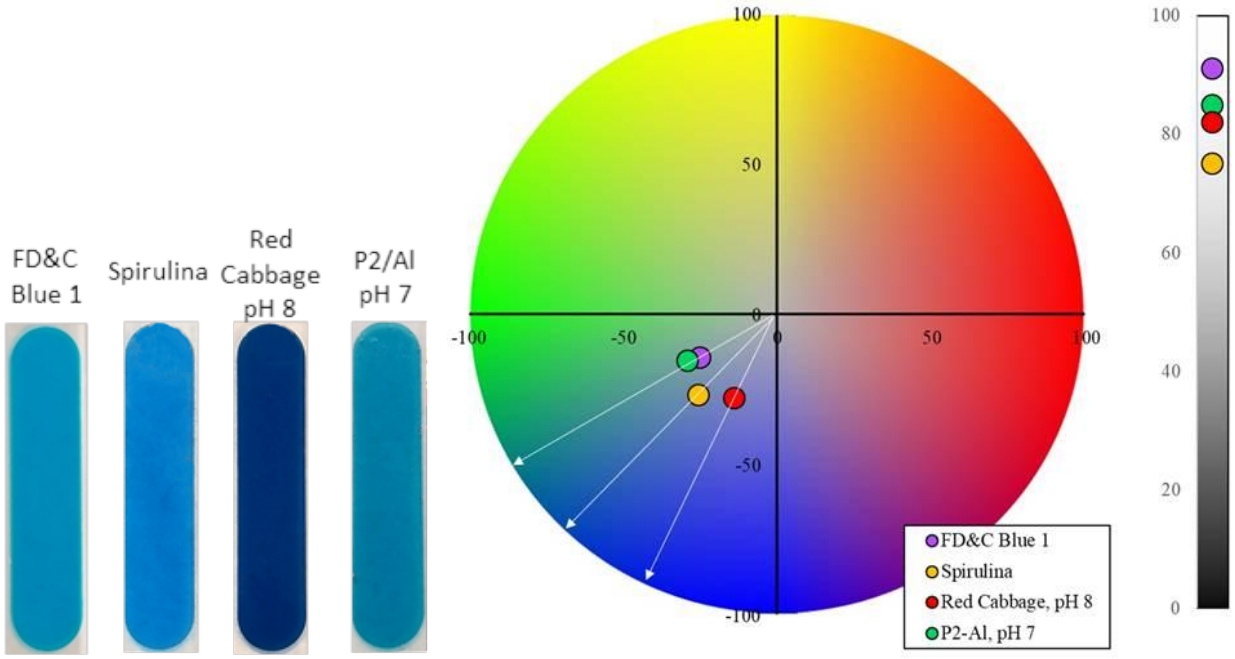
Anthocyanins acylated by hydroxycinnamic acid residues are prone to adopt folded conformations in which the chromophore and acyl groups develop π-stacking interactions *(5, 16, 17)*. Depending on the structure of acylated anthocyanins, self-association can also occur. These π-stacking interactions protect the flavylium ion against the nucleophilic attack of water at C2 (see Figure 1B for numbering). Thus, hydration is slower and less thermodynamically favored (higher pK'_h), so that the residual color of mildly acidic solutions is more intense. By contrast, the acidity constants are only moderately impacted by the acylation pattern, meaning that the π-stacking interactions are of similar intensity with all three colored forms (Supplementary Table 9.1). For the three most abundant diacylated anthocyanins in red cabbage

(P6, P7 and P8), hydration is slower and provides less colorless forms at equilibrium than for their monoacylated homologs (P3, P4, P5). This trend is general and is usually interpreted by a more abundant population of folded conformations for diacylated anthocyanins, including the possibility of sandwich-type conformations (one acyl stacked onto each face of the chromophore).

Supplementary Section 4. Color expression of blue solutions in the a*b* space.



Supplementary Figure 4.1. Graphical Color Representation – Dimension Reduction in Color Space. Color expression is plotted in three dimensions with the coordinates L*a*b*. 3D figures are difficult to represent in 2D space (i.e. documentation) so the L* and a*b* coordinates are separated, reduction of dimension, to allow for easier representation. C* and h° can be represented on the a*b* space using arrows for hue angle and chroma the distance from the center. In order to describe and depict color coordinates in 3-dimensional color space, the spherical graph was reduced to 2-dimensional by taking a cross-section (at an L* value of 50) to provide a* and b* in the plane of the page and 1-dimensional to the L* value is provided in the sidebar.



Supplementary Figure 4.2. Aqueous Solution Color Data for Four Blue Colorants. Aqueous solutions displayed in demountable 0.5 mm cuvettes of the four blue colorants: FD&C Blue No. 1, spirulina, red cabbage at pH 8, and P2-Al complex ($Al^{3+}(P2^{-})_3$) at pH 7. The colorimetry data are plotted in a*b* space. Measured values were plotted on a*b* space. Hue angles are shown by arrows and L* values in sidebar. P2-Al complex at pH 7 provides a very similar hue angle to FD&C Blue No. 1, 207.3 and 209.4 respectively. Whereas both spirulina and red cabbage colorants provide more violet-containing hue angles, 226.2 and 243.0 respectively. As can be seen, the P2-Al complex at pH 7 has a lower violet contribution. All of the colorimetry data is summarized in Supplementary Table 4.1 below. Photo Credit: Randall Powers, Mars Incorporated.

Supplementary Table 4.1. Color Data Collected in Aqueous Solution.

Name	L*	a*	b*	C*	h°
FD&C Blue No. 1	91.0	-25.0	-14.1	28.7	209.4
Spirulina	82.0	-25.6	-26.6	36.9	226.2
Red Cabbage, pH 8	75.0	-14.0	-27.5	30.9	243.0
P2-Al, pH 7	84.9	-29.2	-15.0	32.9	207.3

Supplementary Method 4.1. Conversion of UV visible spectra to L*a*b* values. UV visible spectra were measured from 190 nm – 1100 nm using an Agilent 8453 spectrophotometer and Jasco 0.5 mm demountable glass cuvettes. Colorimetric values L*a*b* were calculated using raw transmittance data. Conversion was established using the CIE color matching functions and published data for X, Y, and Z tristimulus values and data D65 illuminant (a 10° observer and 5 nm resolution).

Supplementary Section 5. Structural identification of P2.

Supplementary Method 5.1. Structural identification of P2 by instrumental analysis. For structural identification of P2 MS and NMR experiments were carried out. Electrospray ionization mass spectrometry (ESI-MS) of P2 (methanol solution containing 0.5% trifluoroacetic acid) was recorded with a Bruker compact instrument. The obtained data was analyzed the application purchased from Bruker Daltonics. NMR spectra were obtained with a Bruker Daltonics AVANCE III HD 600 with a TCI cryoprobe and BBO cryoprobe (^{1}H: 600 MHz and ^{13}C: 150 MHz) in a 5-mm i.d. tube at variable temperatures using 5% CF_3COOD-CD_3OD as a solvent. Chemical shifts were recorded as parts per

million (ppm) using the CD_2HOD resonance as a standard (3.31 ppm). Various 1D and 2D measurements were carried out for structure identification and assignment of signals using the previously reported irradiation experiments[20].

Supplementary Discussion 5.1. Structural identification of P2. High-resolution ESI-TOF MS analysis of P2 gave a molecular ion peak at 979.2709 (Supplementary Fig. 5.1), then, the molecular formula was determined to be $C_{44}H_{51}O_{25}$ (Calculated or Calcd. 979.2719). Although the structures of P2 are known *(22-25)*, complete assignment of ^1H NMR signals of the entire anthocyanin molecule and ^{13}C NMR were not reported. Especially no detailed assignment on sugar protons with multiplicity and coupling constant values were described, indicating that there was no solid evidence that the sugar species and connecting position of them. To fill the gaps in these positional assignment for P2, complete assignment of ^1H- and ^{13}C-NMR signals of the unique anthocyanin molecule were carried according to our previously published procedure (*21)*. A combination of 1D-TOCSY, COSY, NOESY, HSQC and HMBC experiments (Supplementary Fig. 5.2-5.9), all the signals were assigned as shown in Supplementary Table 5.1.

The connection of three sugars were determined by NOEs between each anomeric proton as shown in Supplementary Figure. 5.9. Anomeric proton of Glc-1 showed NOE with H4, that of Glc-2 showed NOEs with H-1 and H-2 of Glc-1, and that of Glc-3 showed NOE with H-6. The attaching position of the sinapoyl residue was determined to be 2-OH of the Glc-2, because the chemical shift of H-2 of Glc-2 was shifted at 4.88 ppm and this attachment was confirmed by HMBC correlation between H-2 and a-proton of sinapoyl residue via carbonyl carbon. The NOEs and HMBC correlations were indicated in Supplementary Figure 5.9, then, the structure of P2 was identified to be 3-*O*-(2-*O*-(2-*O*-(*E*)-sinapoyl-β-D-glucopyranosyl)- β-D-glucopyranosyl)-5-*O*-β-D-glucopyranosylcyanidin.

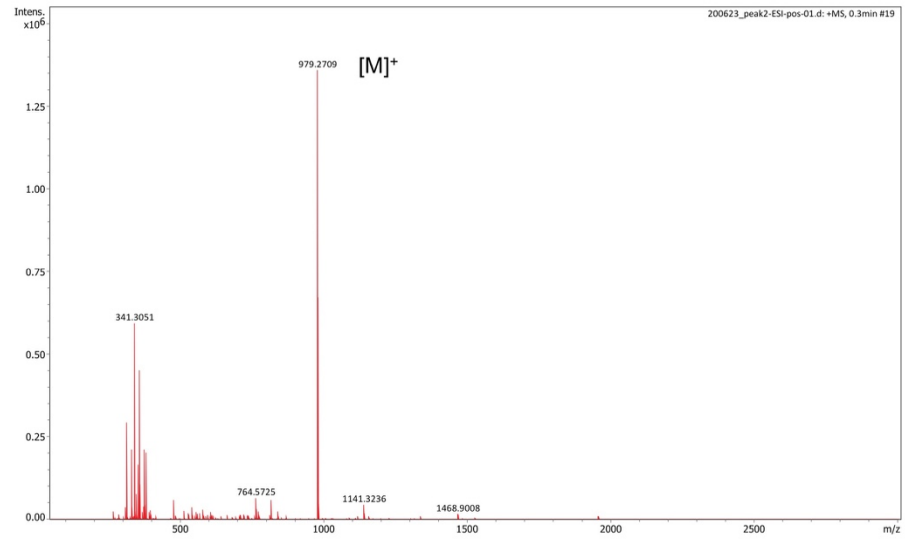
Supplementary Table 5.1. Complete Assignment of ^{1}H and ^{13}C NMR spectra of P2

Position		δ (ppm)	^{1}H multiple.	J (Hz)	^{13}C δ (ppm)
2					163.3
3					144.5
4		9.02	s		134.0
5					155.3
6		7.04	d	2.5	104.3
7					167.9
8		6.86	brd	2.5	95.9
9					155.2
10					111.9
1'					119.8
2'		7.95	d	2.5	117.0
3'					146.3
4'					155.3
5'		7.13	d	8.0	116.2
6'		8.29	dd	8.0, 2.5	128.2
Acyl-1	α	6.16	d	16.0	114.2
	β	7.35	d	16.0	145.5
	1				124.6
	2	6.61	s		105.2
	3				147.9
	4				138.3
	5				147.9
	6	6.61	s		105.2
	OCH_3	3.81	s		55.4

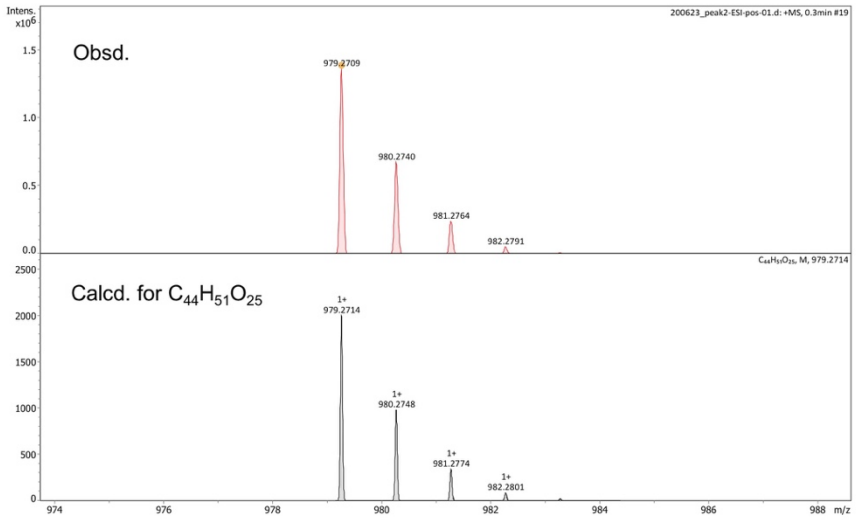
	C=O				166.6
Glc-1	1	5.68	d	8.0	5.61
(Glc_{int})	2	4.09	dd	9.0, 8.0	4.12
	3	3.85	t	9.0	3.86
	4	3.50	t	9.0	3.72
	5	3.64	ddd	9.0, 6.0, 2.0	3.92
	6a	3.69	dd	12.0, 6.0	4.41
	6b	3.93	dd	12.0, 2.0	4.62
Glc-2	1	5.40	d	8.0	4.70
(Glc_{ext})	2	4.88	dd	9.0, 8.0	3.29
	3	3.64	t	9.0	3.36
	4	3.43	t	9.0	3.27
	5	3.37	ddd	9.0, 6.0, 2.0	3.08
	6a	3.70	dd	12.0, 6.0	3.50
	6b	3.92	dd	12.0, 2.0	3.56
Glc-3	1	5.18	d	8.0	5.18
	2	3.74	dd	9.0, 8.0	3.77
	3	3.60	t	9.0	3.58
	4	3.49	t	9.0	3.46
	5	3.59	m	9.0, 6.0, 2.0	3.58
	6a	3.77	dd	12.0, 6.0	3.73
	6b	3.97	dd	12.0, 2.0	3.97

^{1}H: 600 MHz, 5% CF_3COOD-CD_3OD, rt.

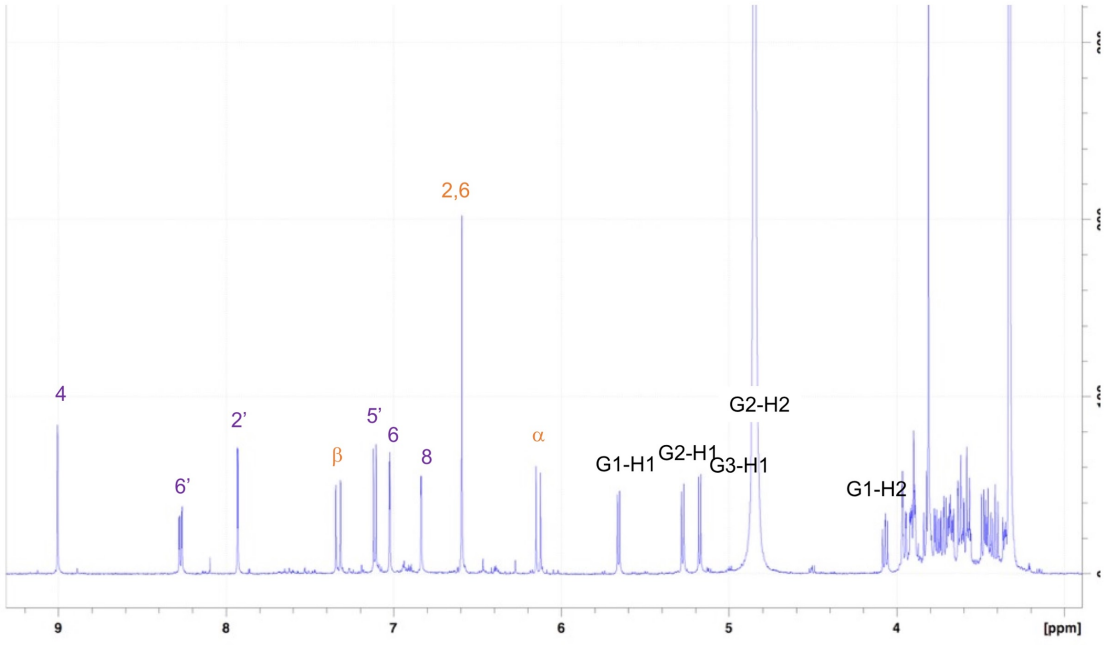
A)



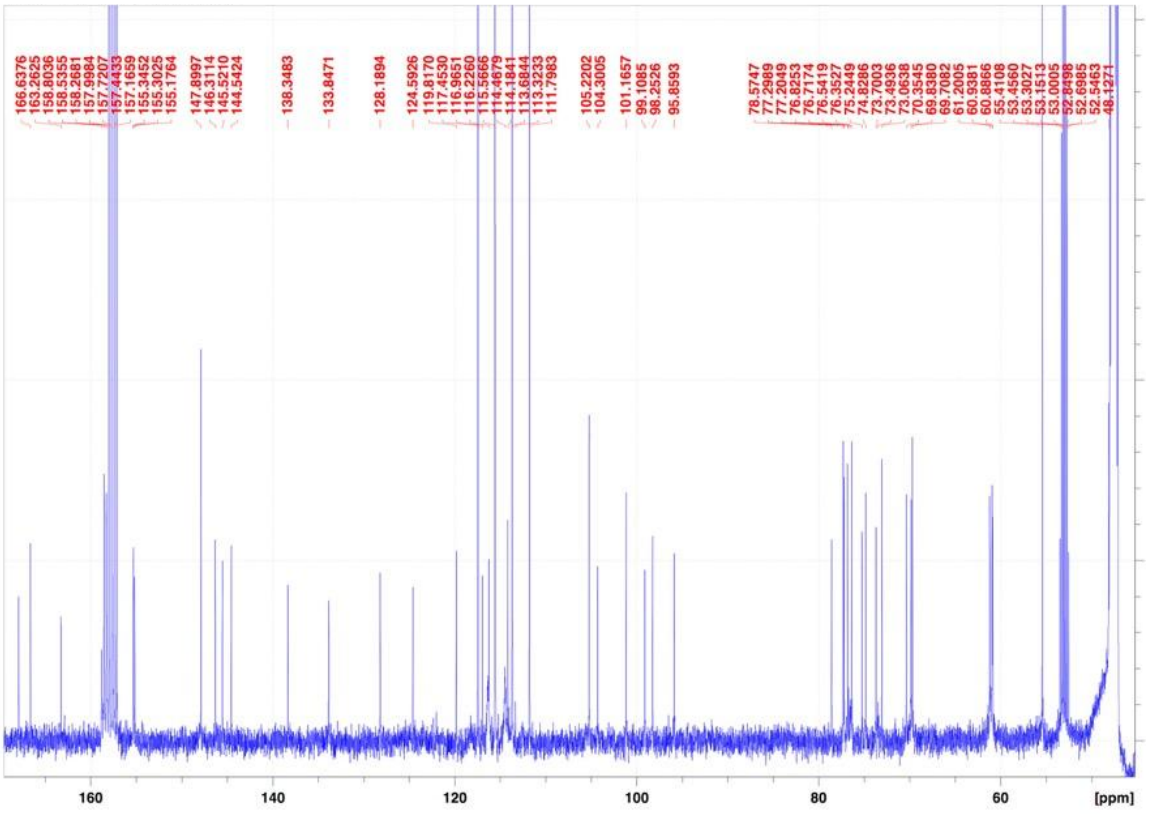
B)



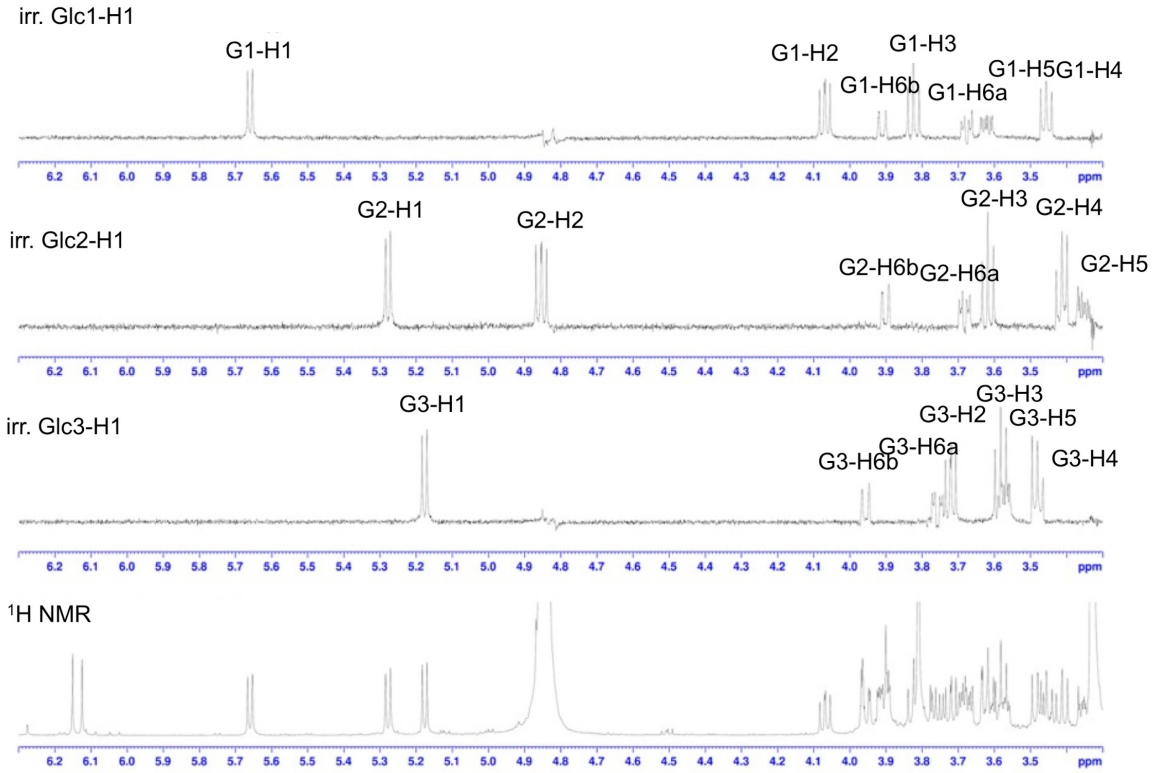
Supplementary Figure 5.1. Mass spectrum data of Peak 2. A) Mass spectrum of P2; B) Comparison of measured HR-ESI-TOF-MS of P2 (upper) and calculated spectrum for $C_{44}H_{51}O_{25}$ (lower).



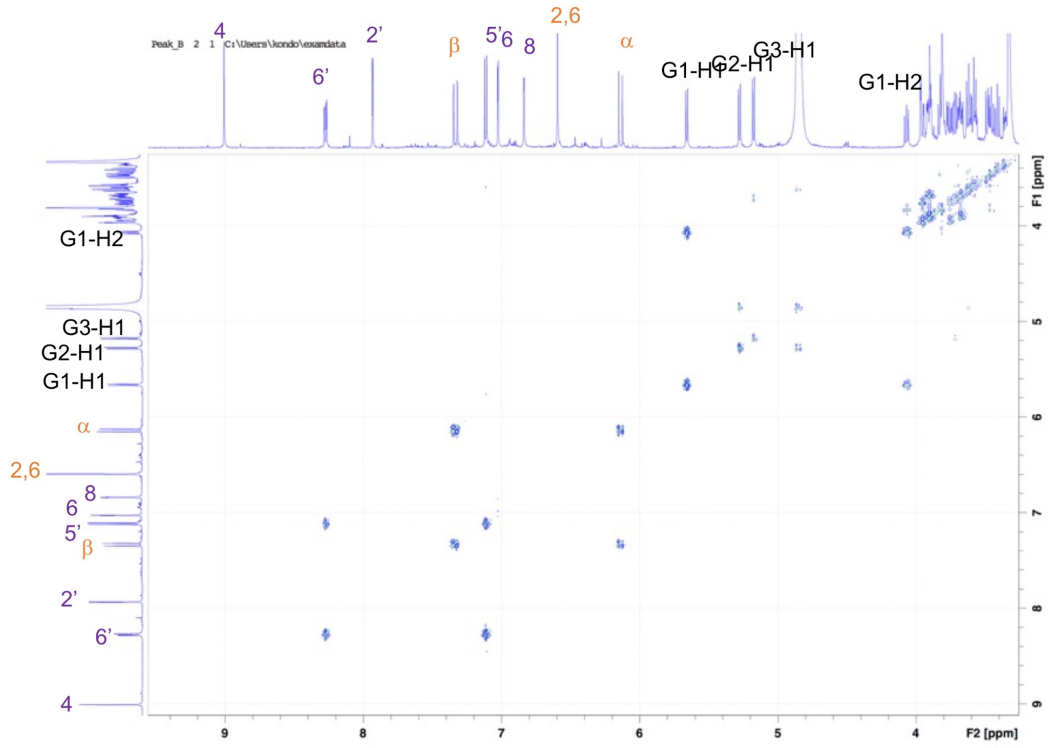
Supplementary Figure 5.2. ^{1}H NMR spectrum of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).



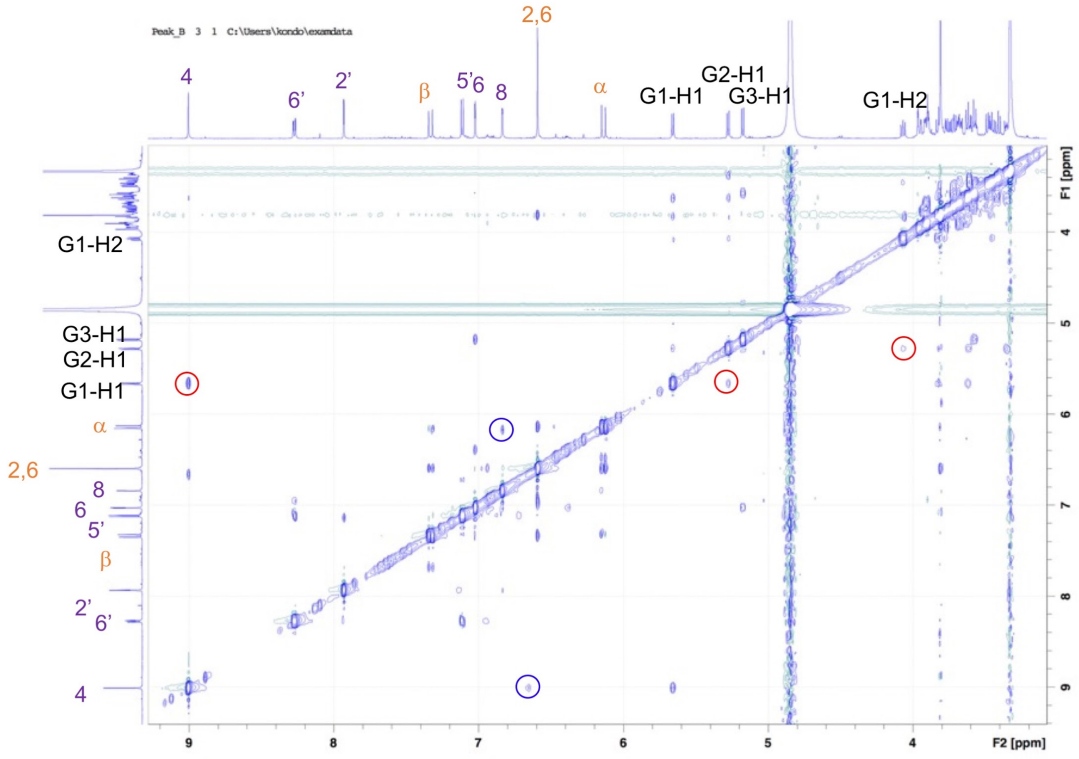
Supplementary Figure 5.3. ^{13}C NMR spectrum of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).



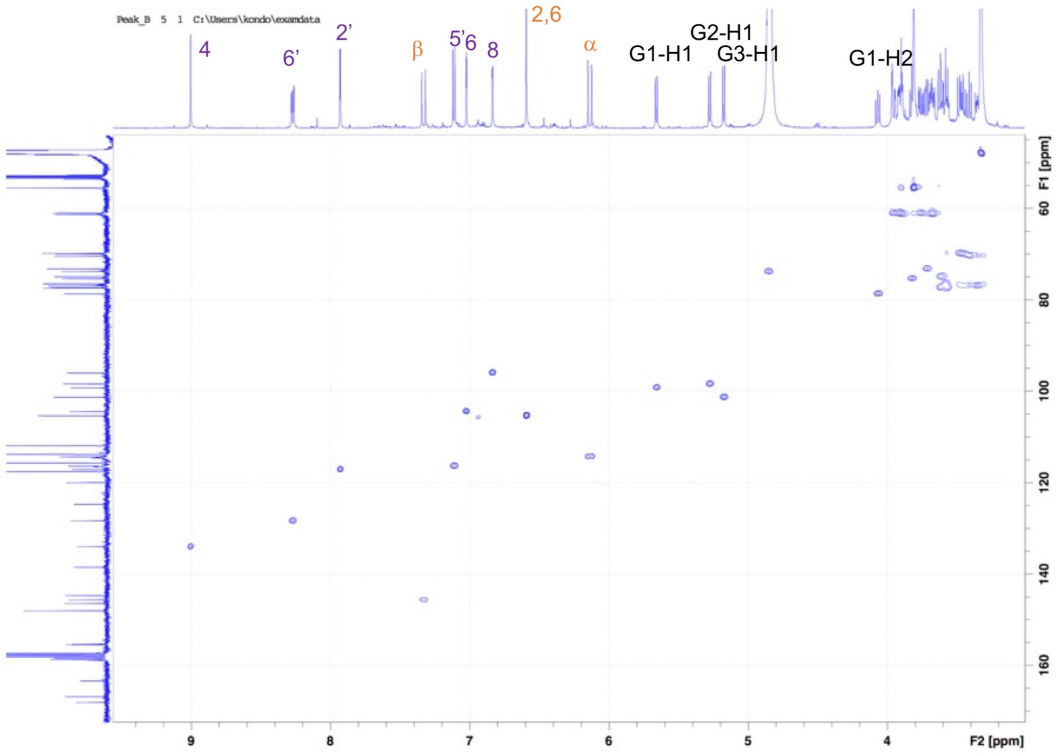
Supplementary Figure 5.4. 1D-TOCSY spectra of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).



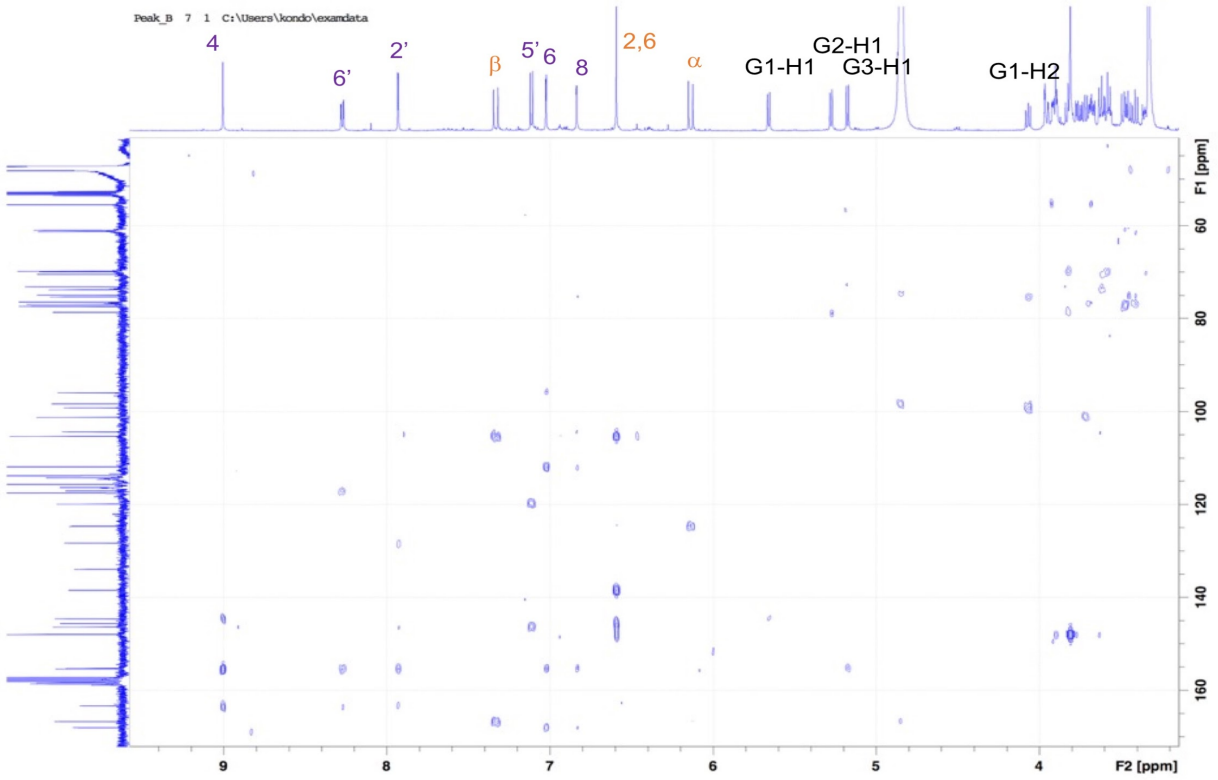
Supplementary Figure 5.5. COSY of P2 at rt (5% TFA-*d* CD_3OD, 600 MHz, rt).



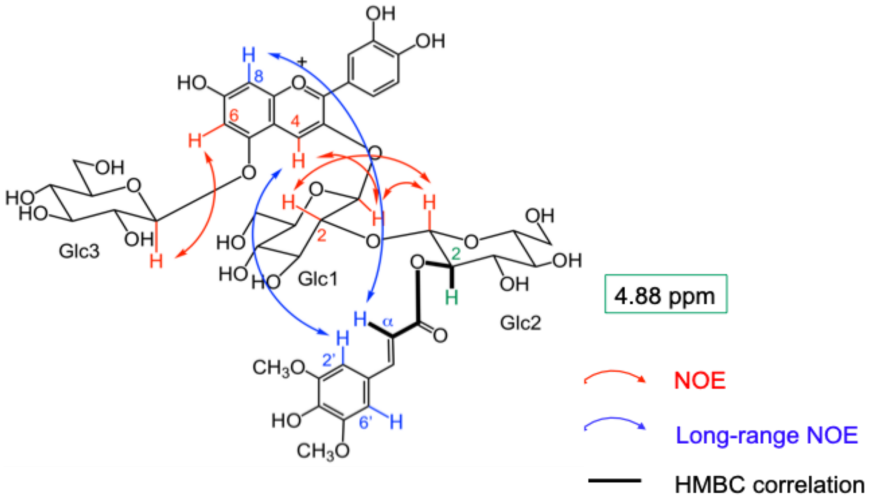
Supplementary Figure 5.6. NOESY of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).



Supplementary Figure 5.7. HMQC of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).



Supplementary Figure 5.8. HMBC of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).



Supplementary Figure 5.9. **Structure of P2**. Cyanidin-3-O-β-D-Glucose (2→1)-β-D-Glc with NOE (indicated with red arrows), HMBC (indicated with black bolded lines) and long-range NOE correlations

(indicated with blue arrows). (5% TFA-*d* CD_3OD, 600 MHz, rt). The hydrogens of the π-system of the sinapate moiety interact with the hydrogens of the cyanidin backbone (blue arrows).

Supplementary Section 6. UV visible spectra and stability of P2, P5 and P8 complexed with Al^{3+}.

Supplementary Method 6.1. UV visible spectroscopy analysis and stability of red cabbage anthocyanins (RCAs). The stability of aluminum complex of P2, P5, and P8 were analyzed according to a previously reported method with small revision *(21)*. Each stock solution of the anthocyanins (P2: 10 mM in H_2O same for P5 and P8) was diluted in 20 mM phosphate buffer at pH 7.0 to the final concentration of 0.05 mM, 0.5 mM, and 5 mM with/without aluminum ion (⅓ eq or 1 eq. to each anthocyanin) by addition of aqueous solution of ammonium aluminum sulfate dodecahydrate ($AlNH_4(SO_4)\cdot 12H_2O$). The obtained solutions were poured into a cuvette of 10 mm, 1 mm and 0.1 mm cell-length, respectively, then UV visible absorption spectra (200–800 nm) were recorded on a JASCO V-560 spectrophotometer. The solutions were kept at room temperature, protected from light, and measurements were carried out repeatedly. The stability was indicated by plotting the absorbance at λ_{max}.

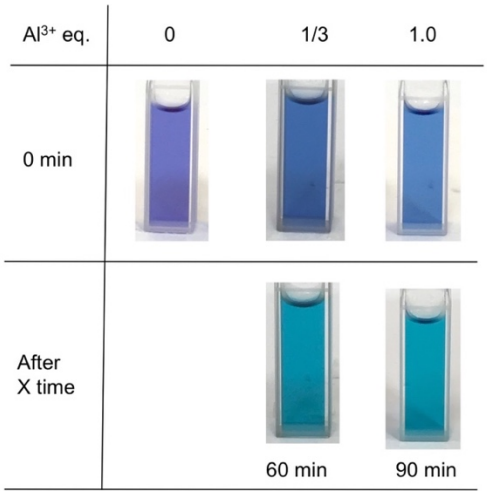
Supplementary Discussion 6.1. UV visible spectra and stability of P2, P5 and P8 complexed with aluminum ion. As a well-known phenomenon, polyacylated anthocyanins show remarkable stability in weakly acidic to neutral aqueous conditions even in diluted condition <0.05 mM, due to the intramolecular stacking of aromatic acyl moieties to anthocyanidin chromophore *(5, 55, 56)*. Monoacylated anthocyanins are usually unstable and only one example, alatanin C from purple yam, was reported to exhibit unusual stability by nested stacking *(57)*. We first analyzed stability of P2 at pH 7 (0.05 mM in 20 mM phosphate buffer) with plotting absorbance at λ_{max}. We already reported P5, a structural isomer of P2, and di-acylated anthocyanin, P8, therefore, the stability of them was compared. P8 was the most stable remaining 50% of

the color intensity after 8 days-storage under dark condition at rt. Contrasting to P8, the stability of P2 and P5 was 21% and 13%, respectively. These results indicate that mono-acylated anthocyanin in red cabbage showed usual properties (Figure 2A, Supplementary Table 6.1).

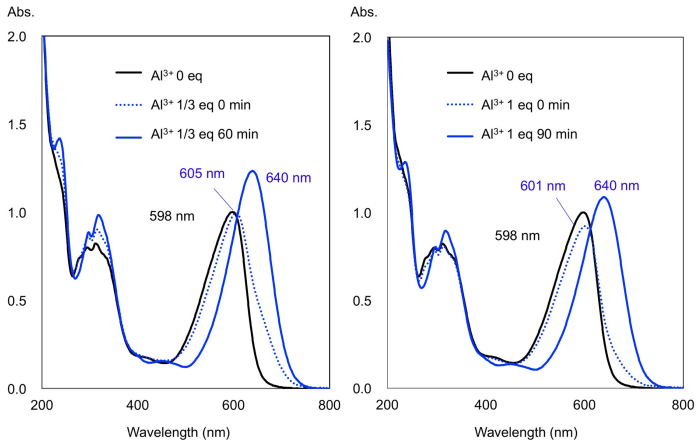
However, when aluminum ion was added to the solution, P2 showed very special characteristics: by addition of only 1/3 eq. of aluminum ion (Al^{3+}) to P2 the solution turned blue. Interestingly, it needed time to reach the equilibrium structure of $(Al^{3+}(P2^{-})_3$; as shown in Supplementary Figure 6.1 and 6.2, The color just after mixing was purplish blue and changed to blue after 60 min (Al^{3+}: 1/3 eq.) and 90 min (Al^{3+}: 1 eq.), then under an equilibrium state λ_{max} of the solutions moved more than 40 nm from 598 nm (0 eq) to 640 nm (1/3 eq, 1 eq.). Compared with this, P5 showed only 6 nm shift from 596 nm (0 eq) to 602 nm (1/3 eq) and P8 with 19 nm from 598 nm (0 eq) to 617 nm (1/3 eq) (Figure 2A, Supplementary Table 6.1). Among these three solutions, the color of P2 was the bluest and the most similar with that of FD&C Blue No. 1. When the equivalent of Al^{3+} to P2 increased to 1 eq, no difference was observed in the color indicating that stoichiometric complex, in which the ratio of P2 to Al^{3+} to be 3:1, might be formed (Figure 2A and Supplementary Figure 6.1). When the concentration of P2 increased to 0.5 mM and 5 mM, the similar blue solutions with λ_{max} of approximately 640 nm were obtained by addition of Al^{3+}. Furthermore, when the molar ratio of Al^{3+} to P2 increased to 5 eq. the color of the solution, Vis absorption spectrum and CD did not give remarkable change, indicating that the ratio of P2 to Al^{3+} might be 3 :1.

Supplementary Table 6.1. UV visible Data for Unique Stability of $Al^{3+}(P2^{-})_3$ - % Absorbance at λ_{max}

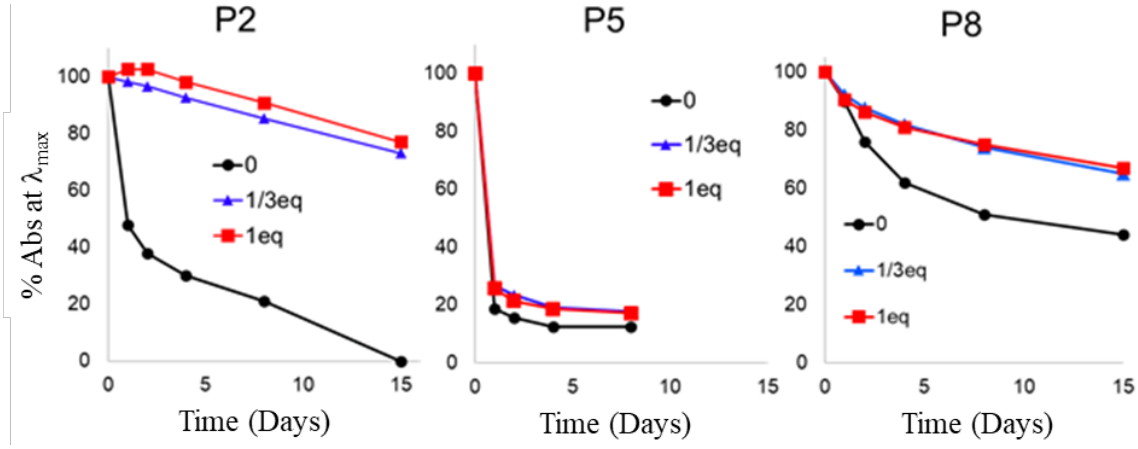
Name	λ_{max}	Day 0	Day 2	Day 4	Day 8	Day 15
P2	598	100	38	30	21	8
P2 + Al	640	100	97	93	85	73
P5	596	100	16	13	13	0
P5 + Al	602	100	24	19	18	0
P8	598	100	76	62	51	44
P8 + Al	617	100	88	82	74	65



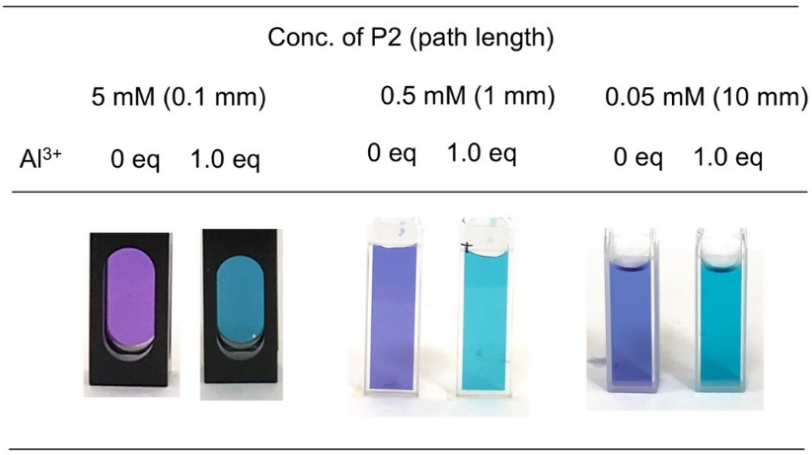
Supplementary Figure 6.1. Complex formation over time of $Al^{3+}(P2^{-})_3$. Cuvette pathlength was modified and the UV visible absorption spectrum was monitored over time after the addition of Al^{3+}, pH 7. Observations showed that complex formation takes longer in more dilute solutions. If the proposed structure $Al^{3+}(P2^{-})_3$ is correct, then building of the supramolecular assembly (three P2 moieties around the Al and interlocking of P2's) takes time to equilibrate. Photo Credit: Kumi Yoshida, Nagoya University.



Supplementary Figure 6.2. UV-vis spectral change in P2 following the addition Al^{3+}. In both cases the λ_{max} shift was slow, reaching completion after 60 min and 90 minutes respectively. Solutions were kept under dark condition at rt. P2: 0.05 mM in 20 mM phosphate buffer, pH 7. Cuvettes and color change are shown in Figure 2A and Supplementary Fig.6.1.

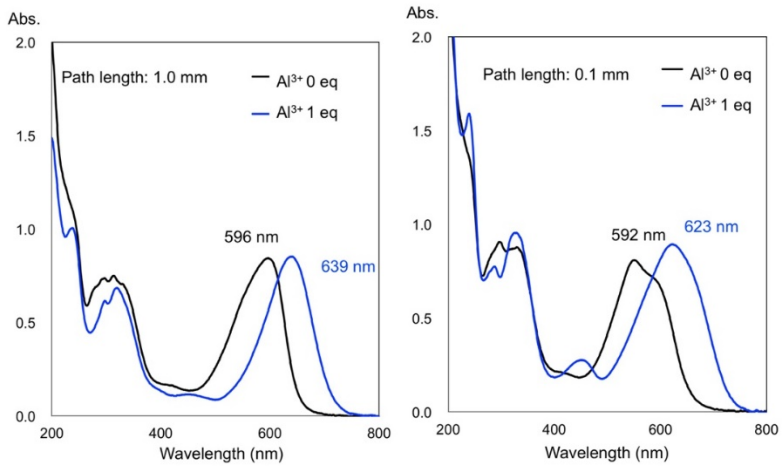


Supplementary Figure 6.3. Unique Stability of $Al^{3+}(P2^{-})_3$. P2, P5, and P8, are regio-isomers where the acyl substitution is sinapoyl attached to either Glc-1, Glc-2, or both. The addition of Al^{3+} enhances the stability of both P2 and P8, however has a minor effect on the stability of P5. P8 has an inherent additional stability due to the di-acylation but is still surpassed by the stability of $Al^{3+}(P2^{-})_3$. Once the $Al^{3+}(P2^{-})_3$ complex is formed, it demonstrates a dramatic increase in stability which is thought to stem from the intermolecular interlocking of the P2 moieties. Color code: 0, 1/3, 1 eq, of aluminum are black, blue, and red, respectively.



Supplementary Figrue 6.4

Supplementary Figure 6.4. The same colored complex is formed regardless of concentration of P2 and Al^{3+}. When the concentration of P2 increased to 0.5 mM and 5 mM, the same blue colored solutions were obtained indicating that $Al^{3+}(P2^{-})_3$ complex is formed. All solutions were at pH 7. Photo Credit: Kumi Yoshida, Nagoya University.



Supplementary Figure 6.5. The Vis absorption spectra of $Al^{3+}(P2^{-})_3$. The spectra of the solutions in Supplementary Figure 6.4 indicates that the λ_{max} of the solutions were around 640 nm. same blue colored solutions were obtained indicating that $Al^{3+}(P2^{-})_3$ complex is formed.

Supplementary Discussion 6.2 Comparing of the stability and behavior of P2, P5, P8 (sinapoyl-containing anthocyanin homologues). The stability of those blue colored solutions was analyzed, and the results were compared with those of P5 and P8 (Supplementary Figure 6.3). As mentioned above, without Al^{3+} diacylated anthocyanin, P8 was most stable and the color of the mono-acylated anthocyanins, P5 and P2 disappeared within a week. However, existence of Al^{3+} (1/3 eq) gave a remarkable stability to the solution and kept the color more than 70% after a half month. Compared with this, Al^{3+} (1/3 - 1eq) did not show any stabilizing effect on P5. P8 showed relative stability with co-existence of Al^{3+} (1/3 - 1eq), but the remaining color intensity was lower than that of P2 (1/3 eq: 65%, 1 eq: 67%).

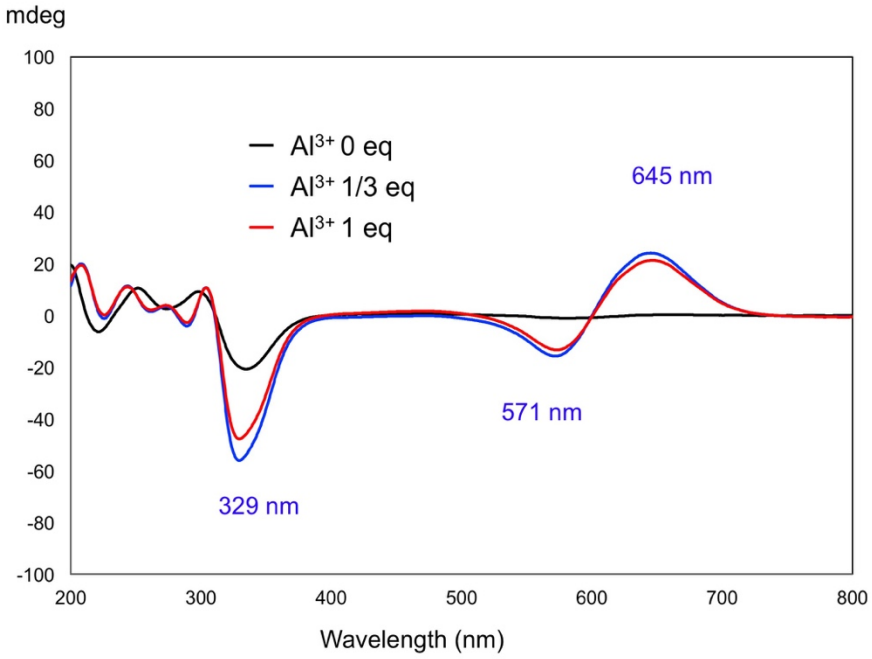
Supplementary Section 7. Structural analyses of the aluminum complex of P2, P5, and P8.

Supplementary Methods 7.1. Circular dichroism (CD) and ESI-TOF-MS analyses of P2, P5, and P8 complexes with aluminum. Circular dichroism (CD) was recorded from 200–800 nm with a JASCO J-720 spectrometer using the same solution for the measurement of stability described in Supplementary Section 6.1. To increase signal/noise ratio (S/N ratio) the scan was repeated four times, then, the data was averaged *(21)*.

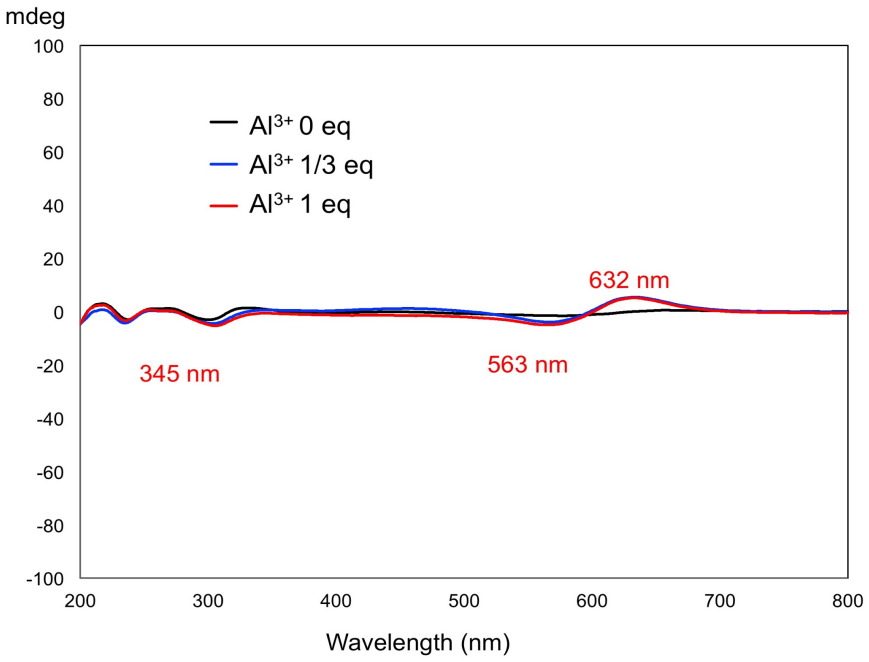
ESI-MS spectra of aluminum complex of P2, P5 and P8 solutions were recorded with a Bruker compact instrument and analyzed the application purchased from Bruker Daltonics. Stock solution of P2 (10 mM in H_2O) was diluted to 0.05 mM with phosphate buffer at pH 7.0, ammonium acetate buffer at pH 7.0 or

0.01-0.005% NH_4OH aq. The obtained sample solutions were infused into the MS instrument using microsyringe with 5 μL/min.

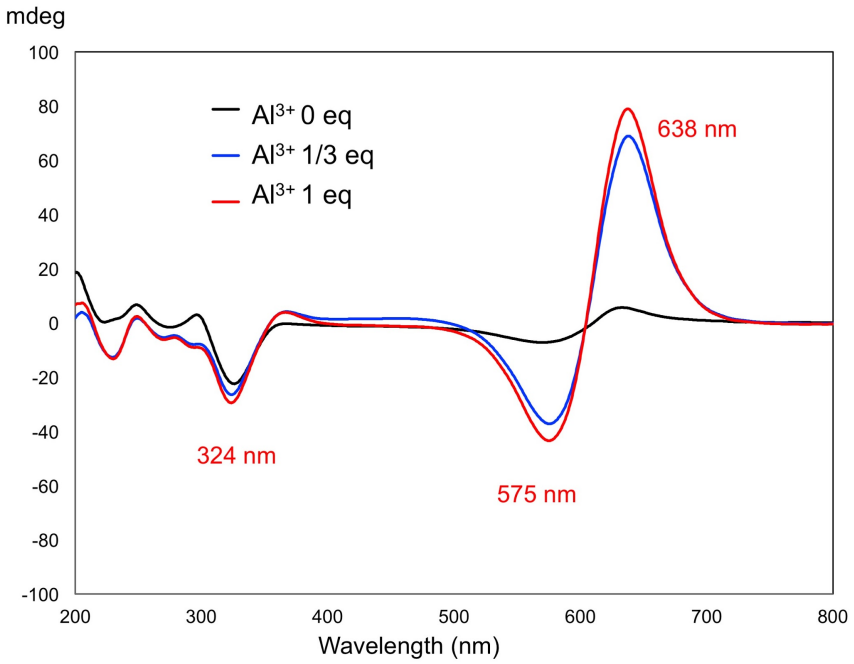
Supplementary Discussion 7.1. Structural characterization of blue $Al^{3+}(P2^-)_3$. To determine the structure of the blue-colored $Al^{3+}(P2^-)_3$, the circular dichroism (CD) of the solutions was recorded. The arrangement of the three cyanidin ligands around Al^{3+} resulted in coordination chirality (Δ(delta) and/or Λ(lamda) configurations) and an exciton-type Cotton effect was observed, as reported for the enantiomers of tris(phenanthroline)ruthenium(II) *(59).* By addition of aluminum ion, P2 showed an exciton-type positive Cotton effect around the λ_{max}. By contrast, no such Cotton effect was observed without Al^{3+} (Supplementary Figures 7.1-7.3). CD spectra showing exciton-type Cotton effect were also predicted by calculations (Supplementary Section 8.1). The 1:3 stoichiometry of the $Al^{3+}(P2^-)_3$ complex is consistent with the molecular ions detected by ESI-TOF-MS (Supplementary Discussion 7.4). No larger molecular ions (e.g. $Al^{3+}{}_2(P2^-)_6$) could be detected.



Supplementary Figure 7.1. CD spectra of P2 (0.05 mM) with Al^{3+} in buffered solution at pH 7.0 (path length: 10 mm, 20 mM phosphate buffer)



Supplementary Figure 7.2. CD spectra of P5 (0.05 mM) with Al^{3+} in buffered solution at pH 7.0 (path length: 10 mm, 20 mM phosphate buffer)



Supplementary Figure 7.3. CD spectra of P8 (0.05 mM) with Al^{3+} in buffered solution at pH 7.0 (path length: 10 mm, 20 mM phosphate buffer)

Supplementary Table 7.1. CD Data of P2-, P5- and P8-Al Complexes

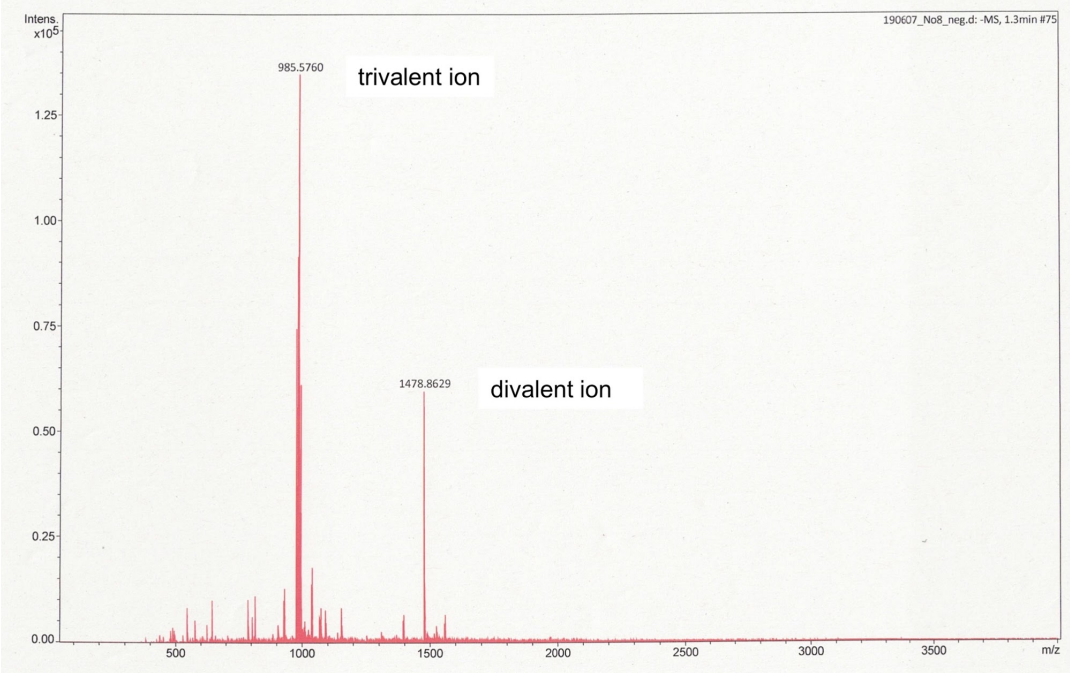
P2, pH 7			
1/3 eq Al^{3+}		**1 eq Al^{3+}**	
λ_{max} (nm)	Cotton (mdeg)	λ_{max} (nm)	Cotton (mdeg)
645	24.1	647	21.3
572	-15.7	573	-13.3
Intercept (nm)	ΔCotton (mdeg)	Intercept (nm)	ΔCotton (mdeg)
600	39.8	602	34.6
P5, pH 7			
1/3 eq Al^{3+}		**1 eq Al^{3+}**	

λ_{max} (nm)	Cotton (mdeg)	λ_{max} (nm)	Cotton (mdeg)
633	5.3	633	5.6
563	-4.8	567	-3.8
Intercept (nm)	ΔCotton (mdeg)	Intercept (nm)	ΔCotton (mdeg)
598	10.1	601	9.4
P8, pH 7			
1/3 eq Al^{3+}		**1 eq Al^{3+}**	
λ_{max} (nm)	Cotton (mdeg)	λ_{max} (nm)	Cotton (mdeg)
634	5.7	638	68.9
570	-7.2	575	-37.2
Intercept (nm)	ΔCotton (mdeg)	Intercept (nm)	ΔCotton (mdeg)
605	12.7	605	106.1

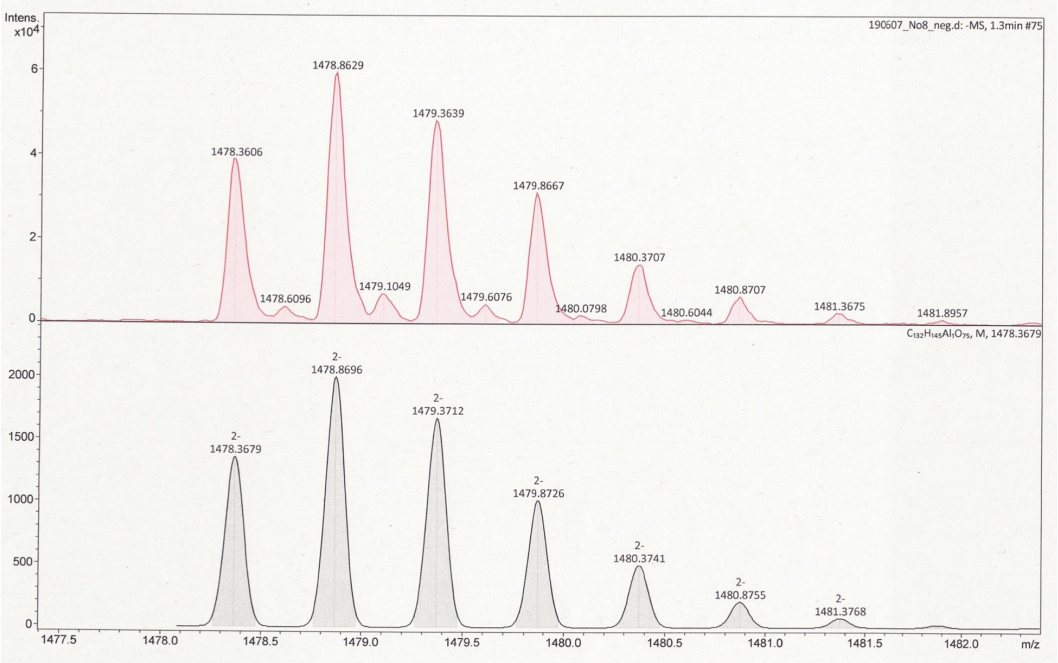
Supplementary Method 7.2. Circular dichroism (CD) was measured with a JASCO J-720 spectrometer using the same solution for the measurement of stability *(21)*.

Supplementary Discussion 7.2. ESI-TOF-MS analysis of blue solutions mixing P2 and Al^{3+}. To clarify the structure of the $Al^{3+}(P2^{-})_3$ complex, the blue solution constructed by mixing P2 and Al^{3+} in a buffer was analyzed by ESI-TOF-MS (Figure 2A and Supplementary Figure 6.1 and 6.2). The blue solution in 20 mM phosphate buffer did not give any molecular ions attributable to complex. Co-existence of concentrated salt usually suppress ionization of unstable metal complex, and in the measurement of

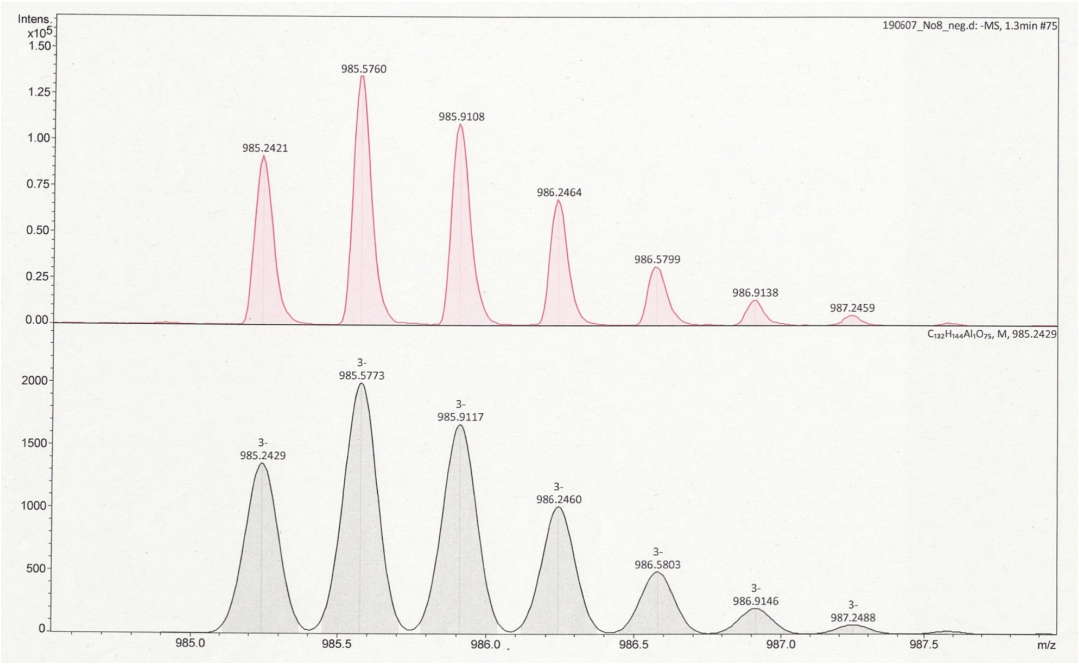
hydrangea blue-complex we re-constructed the complex in highly diluted buffer as 2 mM *(26)*. Furthermore, concentrated salines usually suppress the ionization of unstable metal complexes. Therefore, we mixed P2 and Al^{3+} in 5 mM AcOH-$AcNH_4$, and we could obtain the same colored blue solution. By addition of a very small portion of NH_4OH aq. blue color with the same UV-Vis and CD spectra could be reproduced. Therefore, ESI-TOF-MS analysis of these solutions were carried out. By negative detection, blue solution in 5 mM $AcONH_4$ buffer gave a small ion peak at *m/z* = 1478 attributable to the divalent ion of aluminum complex of P2 (the ratio is 3:1). The blue solution reconstructed in diluted NH_4OH aq. gave two peaks: one was observed at *m/z* = 1478.8629 and the other was at *m/z* = 985.5760. The former was attributable to divalent molecular ion peak of $[3 \text{ x } P2+Al-2H]^{2-}$ and the latter was trivalent molecular ion peak of $[3 \text{ x } P2+Al-3H]^{3-}$ (Supplementary Figure 7.4–7.6). The same results were obtained when the concentration of NH_4OH was reduced to 0.0005%. The molecular formula was confirmed by high resolution MS analysis to be $C_{132}H_{147}AlO_{75}$. These results confirmed that P2 gave a stoichiometric aluminum complex with the ratio to be 3:1.



Supplementary Figure 7.4. Negative detection ESI-TOF-MS of the blue solution obtained by mixing P2 (0.05 mM) with Al^{3+} (1 eq.) in 0.001% NH_4OH.

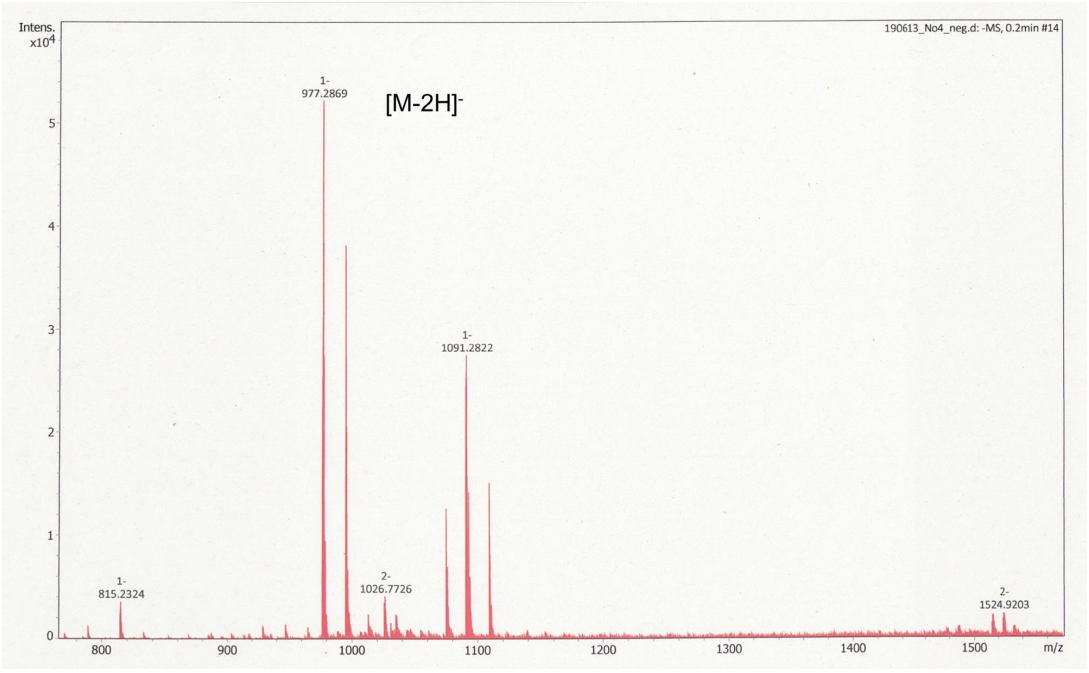


Supplementary Figure 7.5. Negative detection HR-ESI-TOF-MS of the divalent ion of the aluminum complex of P2. The upper: observed spectrum, the lower: calcd. spectrum for $C_{132}H_{145}Al_1O_{75}$.

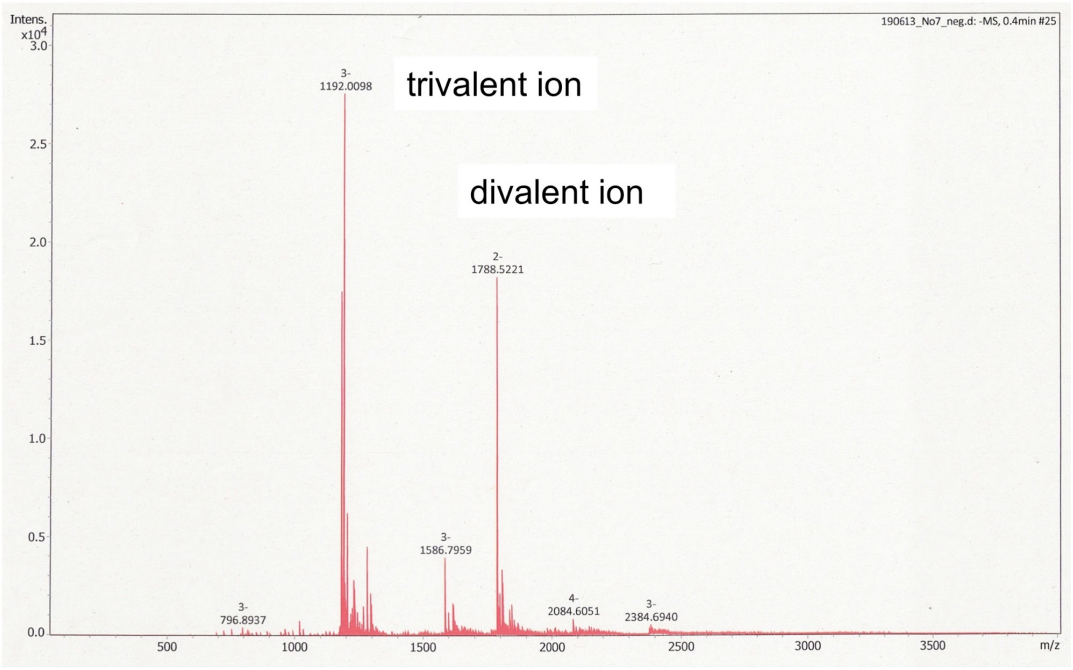


Supplementary Figure 7.6. Negative detection HR-ESI-TOF-MS of the trivalent ion of the aluminum complex of P2. (Top) observed spectrum, (Bottom) calculated spectrum for $C_{132}H_{145}Al_1O_{75}$.

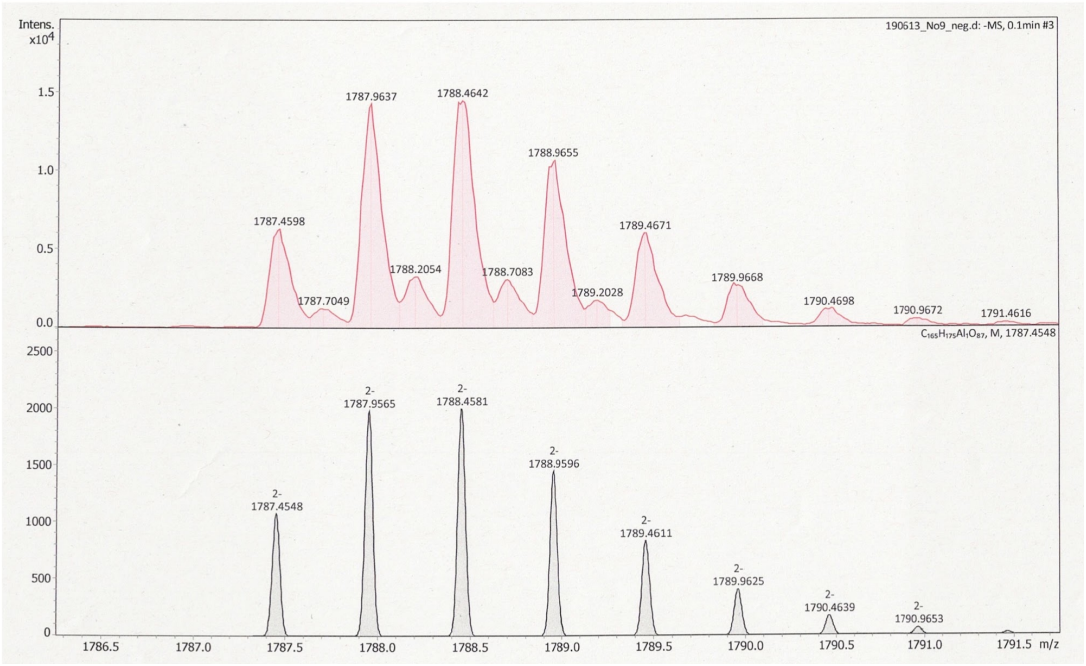
Supplementary Discussion 7.3. Mass spectrometry analysis of P5 and P8. Using P5 and P8 blue solutions were reconstructed under the same mixing condition, and MS analysis was carried out. P5 did not give any molecular ion of aluminum complex (Supplementary Figure 7.7), however, P8 gave both divalent (*m/z* =1788.5221: $[3xP8+Al-2H]^{2-}$) and trivalent (*m/z* =1192.0098: $[3xP8+Al-3H]^{3-}$) molecular ion peaks of metal complex (Supplementary Figures 7.8–7.10). These results indicated that the position of sinapoyl group should strongly affect the formation of an aluminum complex of red cabbage anthocyanin. P8 gave a stoichiometric aluminum complex with the ratio to be 3:1 as the same as that of P2, however it is thought that the steric hindrance by additional sinapoyl group at G-1 may prevent approach or bond formation with the Al and reduce bathochromic shift (618 nm for P8 with Al vs 640 nm for P2 with Al)



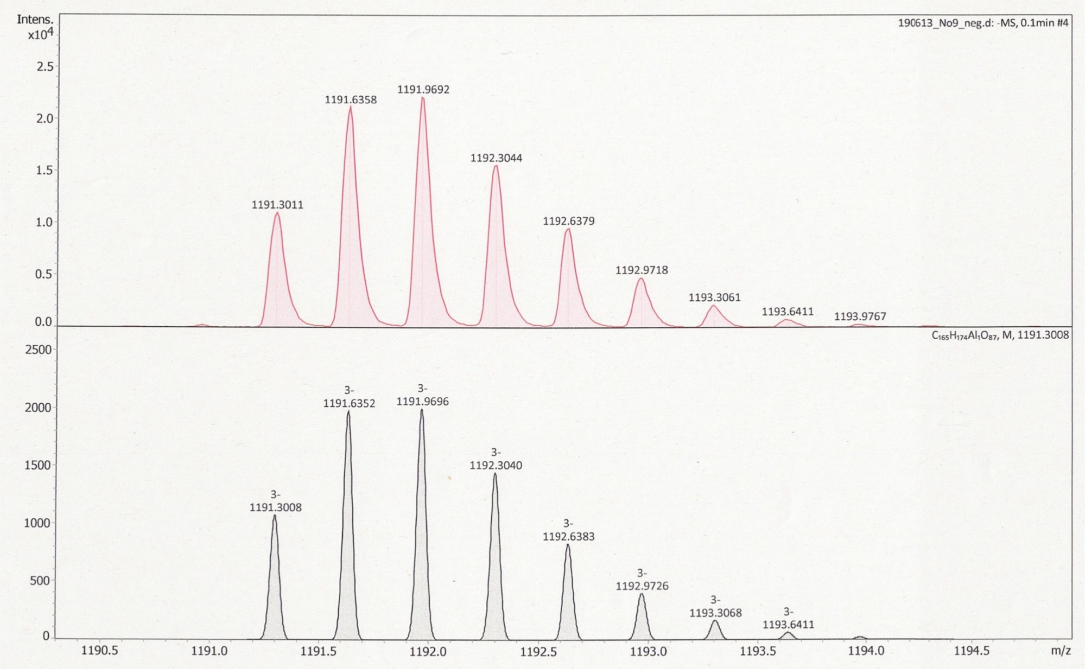
Supplementary Figure 7.7. Negative detection ESI-TOF-MS of P5 obtained by mixing P5 (0.05 mM) with Al^{3+} (1 eq.) in 0.01% NH_4OH.



Supplementary Figure 7.8. Negative detection ESI-TOF-MS of P8 obtained by mixing P8 (0.05 mM) with Al^{3+} (1 eq.) in 0.001% NH_4OH.



Supplementary Figure 7.9. Negative detection HR-ESI-TOF-MS of the divalent ion of the aluminum complex of P8. The upper: observed spectrum, the lower: calcd. spectrum for $C_{165}H_{175}Al_1O_{87}$.

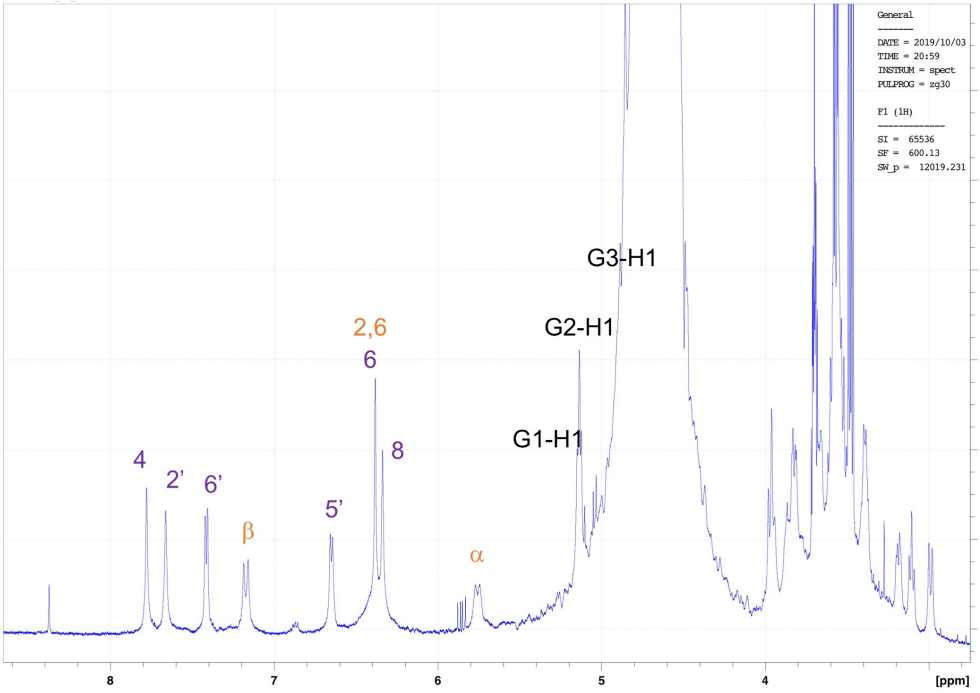


Supplementary Figure 7.10. Negative detection HR-ESI-TOF-MS of the tridivalent ion of the aluminum complex of P8. The upper: observed spectrum, the lower: calcd. spectrum for $C_{165}H_{175}Al_1O_{87}$.

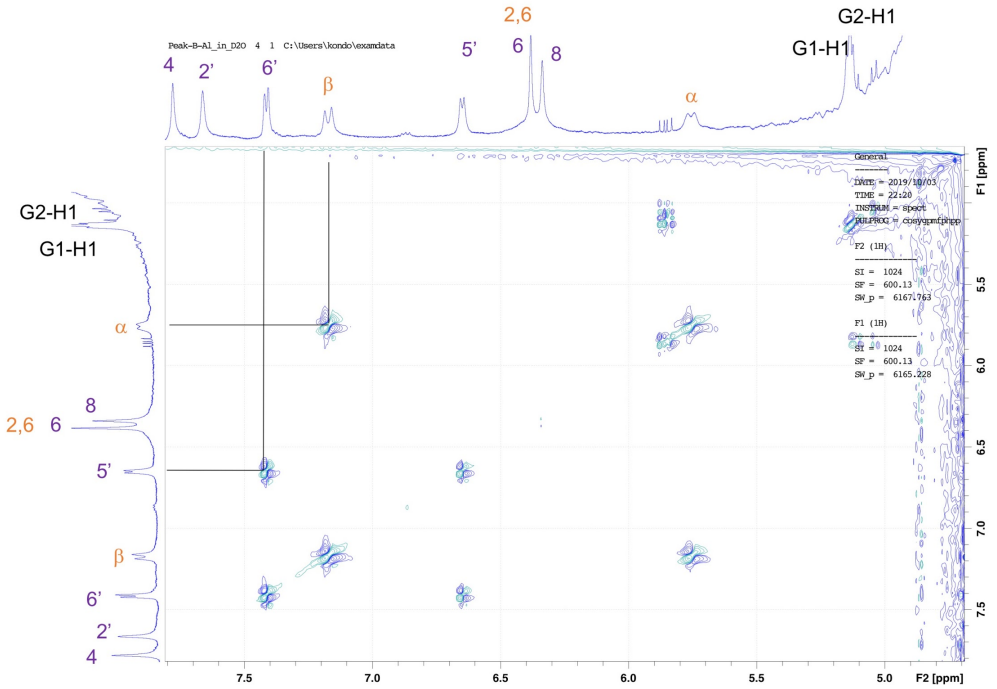
Supplementary Method 7.3. NMR analysis of P2-Al^{3+} complex. The blue solution mixed with P2 and Al^{3+} at the ratio of 3:1 was freeze dried, then, 15 mg of the blue powder was dissolved in D_2O (0.6 mL). NMR spectra were obtained with a Bruker Daltonics AVANCE III HD 600 with a TCI cryoprobe and BBO cryoprobe (^1H: 600 MHz and ^{13}C: 150 MHz) in a 5-mm i.d. tube at variable temperatures. Chemical shifts were recorded as parts per million (ppm) using the HOD resonance as a standard (4.67 ppm). Various 1D and 2D measurements were carried out.

Supplementary Discussion 7.4. NMR analysis of the P2-Al^{3+} complex. In aqueous solution the blue complex composed of P2 and Al^{3+} with the ratio of 3:1 was indicated to exist as a dimeric structure by CD. To obtain the intra and/or intermolecular stacking structure NMR studies were carried out. As shown

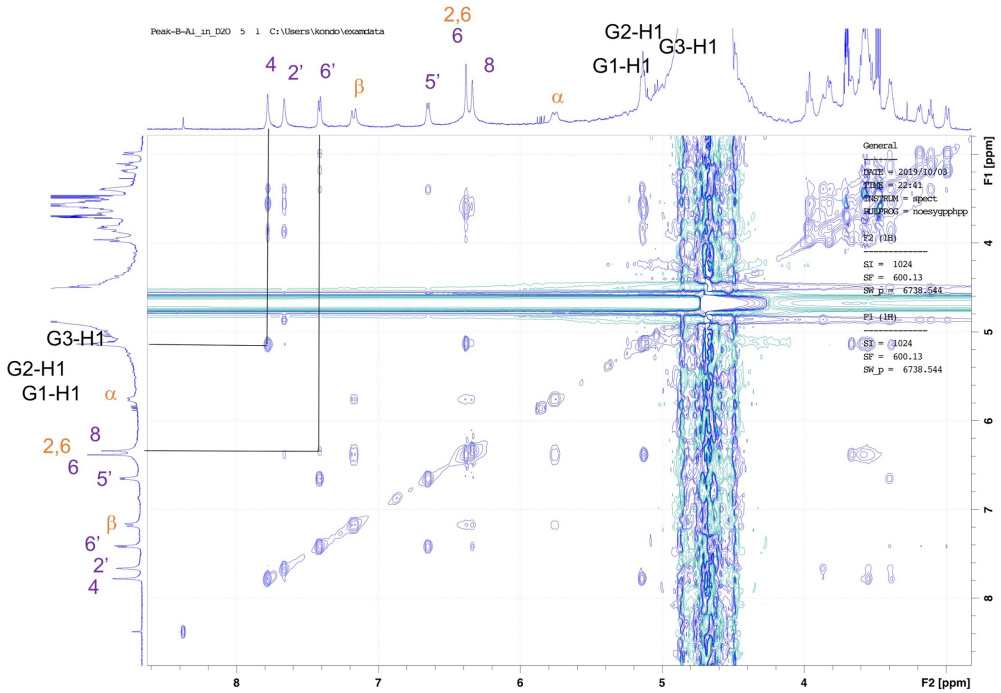
in Supplementary Figure 7.11 signals of ^{1}H NMR spectrum of the complex were broad but simple indicating that intra/intermolecular stacking might exist in the molecule and all the components were symmetrically distributed. By COSY, NOESY and HMQC measurements, several signals could be assigned (Supplementary Figure 7.11–7.14). Several signals of blue complex composed of P2 and Al^{3+} could be assigned, therefore, chemical shifts were compared with those of P2 in acidic CD_3OD (Supplementary Table 7.2). Strong and weak long-range NOEs were shown in Supplementary Figure 7.15. Analysis of these data indicated that the chemical shift of aromatic ring protons shifted toward higher field compared with those of P2 in flavylium cation form. In addition, long range NOEs were observed between signals of sinapoyl residue and cyanidin chromophore. These findings strongly suggest the intramolecular stacking of sinapoyl residue and cyanidin nucleus. As is observed with purple yam anthocyanins (*58*), it might be possible that the long-range NOEs are due to the intermolecular stacking of chromophore and acyl moieties belonging to different molecules.



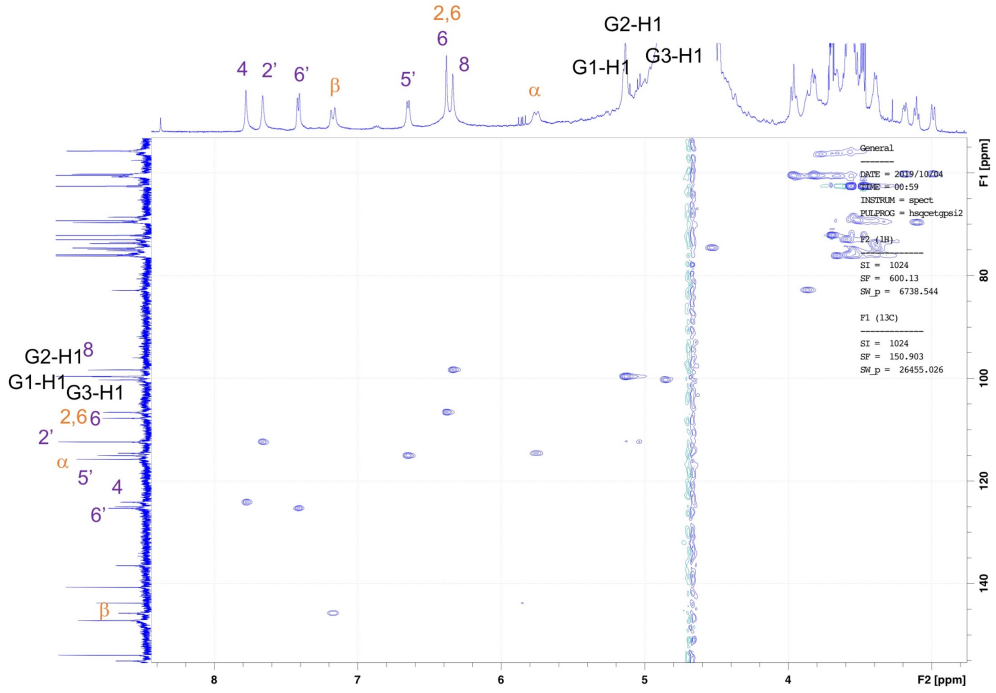
Supplementary Figure 7.11. ^{1}H NMR spectrum of the blue complex composed of P2 and Al^{3+} (D_2O, 600 MHz, rt).



Supplementary Figure 7.12. COSY of the blue complex composed of P2 and Al^{3+} (D_2O, 600 MHz, rt).



Supplementary Figure 7.13. NOESY of the blue complex composed of P2 and Al^{3+} (D_2O, 600 MHz, rt).

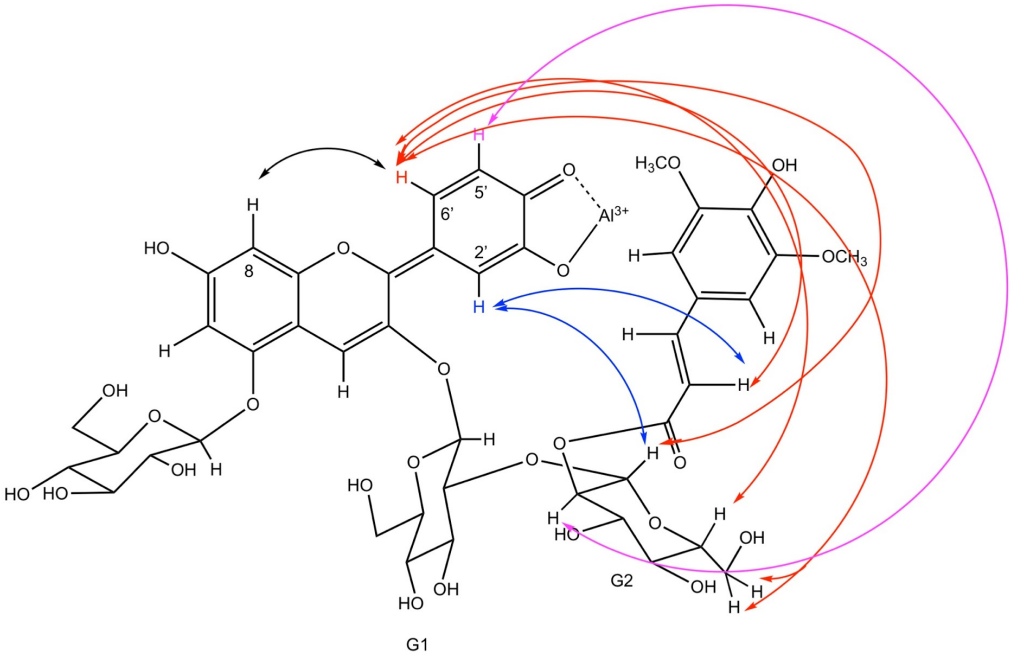


Supplementary Figure 7.14. HMQC of the blue complex composed of P2 and Al^{3+} (D_2O, 600 MHz, rt).

Supplementary Table 7.2. Comparison of chemical shift of signals of the blue complex composed of P2 and Al^{3+} (D_2O, 600 MHz, rt) with those of P2 (5% TFA-*d* CD_3OD, 600 MHz, rt).

		Δ (ppm)		**Difference (ppm)**
		$Al^{3+}(P2^-)_3$	**P2**	**P2 - $Al^{3+}(P2^-)_3$**
		D_2O	**TFA*d*-CD_3OD**	
A	4	7.87	9.02	1.15
	6	6.47	7.04	0.57
	8	6.42	6.86	0.44
B	2'	7.75	7.95	0.20
	5'	6.74	7.13	0.39

	6'	7.50	8.29	0.79
S-1	α	5.84	6.16	0.32
	β	7.26	7.35	0.09
	2	6.47	6.61	0.14
	6	6.47	6.61	0.14
Glc-1	1	5.23	5.68	0.45
Glc-2	1	5.21	5.40	0.19



Supplementary Figure 7.15. Strong and weak long-ranged NOE observed in ($Al^{3+}(P2^{-})_3$) complex. Long range NOEs of the blue complex composed of P2 and Al^{3+} (D_2O, 600 MHz, rt). Analysis of the chemical shift of aromatic ring protons (Supplementary Table 7.2) shifted toward higher field compared with those of P2 in flavylium cation form -indicating interactions shown in figure above. However, it might be possible that the long-range NOEs are due to intermolecular stacking of chromophore and acyl moiety belonging to different molecule.

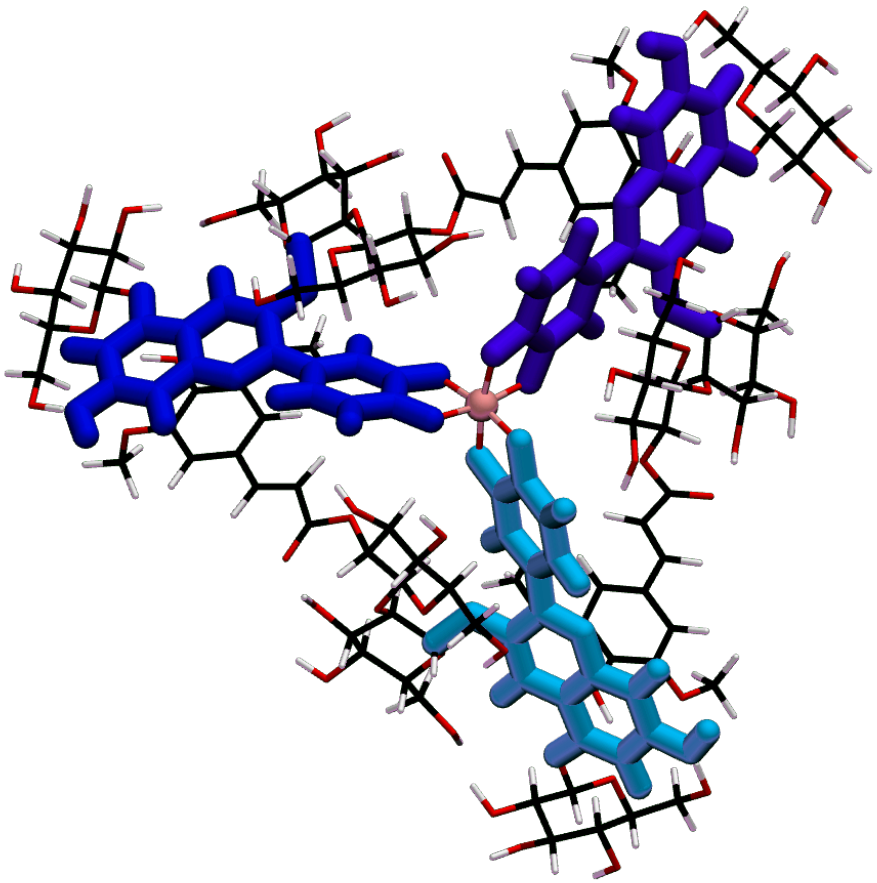
Supplementary Section 8. First-principles simulations of 3:1 complexes with aluminum result in distorted chromophores for P2.

Supplementary Method 8.1. Geometric optimization of putative P2 and P5 3:1 complexes with Aluminum. Optimizations of either P2 or P5 molecules in a 3:1 complex with aluminum were performed, starting from relevant structures of monomers obtained from extensive MD simulations and analysis, as detailed in Supplemental Section 10.1. These were arranged with a C3 symmetry around Al^{3+} and deprotonated at 3' and 4', and with Al^{3+} coordinated to the oxygens at 3' and 4'. The two chiral assemblies of the ligands, corresponding to left- and right-handed (labelled Λ and Δ respectively), were tried. We remark that the two isomers are diastereoisomers, as the ligand already presents several chiral centers in the glycosidic units. Relaxations were then performed in two steps: 1) a pre-optimization was run in vacuum with CP2K *(39)*, with DZVP-MOLOPT-SR-GTH basis set and GTH-BLYP pseudopotential *(40-42)* from the CP2K public repository (cutoff: 350 Ry) corrected with the D3(0) Grimme dispersion *(43)*. A cubic box sizes of 50 angstrom was used for the geometry minimizations performed using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) minimizer in CP2K *(60)*, combined with the Quickstep algorithm *(61)* (convergence 5e-7 a.u.). Final optimizations were run in Quantum Espresso *(44, 45)*, using ultrasoft pseudopotentials and PBE as exchange-correlation functionals *(46)* with implicit solvent (Environ module) *(47)* in periodic cubic box of 80 bohr. The following parameters were adopted for the wavefunction convergence: kinetic-energy cutoff of 35 Ry, charge density cutoff of 320 Ry. The Makov-Payne energy correction was employed to remove finite-box effects. Examples of the methods used to generate in silico data can be found at https://github.com/siegel-lab-ucd/blue-pigment-publication.git

Supplementary Method 8.2. Optical spectra from semi-empirical method. Optical properties for the optimized geometries were obtained with semi-empirical scheme ZINDO *(62)*, as implemented in the ORCA package v 4.2.0 *(63, 64)*. Transition accompanied by their i) oscillator strengths f (a dimensional, from transition electric dipole moments) and ii) rotational strength R (10 ^{40}*cgs unit) were convoluted to gaussian line shape with σ=0.07 eV to compute respectively i) absorption spectra and ii) circular dichroism (CD)spectra.

Supplementary Discussion 8.1. First-principles simulation results: geometries of P2 and P5 around Al. Minimizations starting from several geometries of the P2 ligand were performed, resulting from its available conformational states, following the computational details presented in Supplementary Method 10.1. From this large set of structures (approximately ten for P2), each resulting in complexes in Δ and Λ forms, several stable $Al^{3+}(P2^{-})_3$ structures were obtained in implicit water solvent, where the Al is coordinated with catecholic O(C3') and O(C4') on the B ring. From first-principles calculations, the potential energy of such complexes all falls within a range of ~ 1.5 eV of each other. Although thermal fluctuations and explicit solvation effects probably play a large role, we were able to select hypothetical geometries for $Al^{3+}(P2^{-})_3$. Interestingly, structures seem to receive additional stabilizations by the presence of internal hydrogen bonds interconnecting sugar moieties of the different units. This element and the crowding around the metal center induce large distortions of the chromophore, as shown previously to cause a blueing effect. Most stable structures obtained for $Al^{3+}(P2)_3$ and reported in Supplementary Fig. 8.1 is compatible with the positive Cotton Effect observed experimentally. The optical properties calculated with semi-empirical methods (see Supplementary Method 8.2) are reported in Supplementary Figure 8.5. Some geometries were also tried for P5, many of which were higher in energy, and are not reported here. Further details on lowest energy obtained structures and their

properties are reported in Supplementary Table 8.1. We note the large distortions obtained in the complexes involving P2, compared to much smaller distortions in P5. We propose the lower energy structure of $Al^{3+}(P2^{-})_3$ to be the most stable once in water, due to interlocking between acyl and chromophore moieties of neighboring molecules (Supplementary Fig. 8.2-8.4).



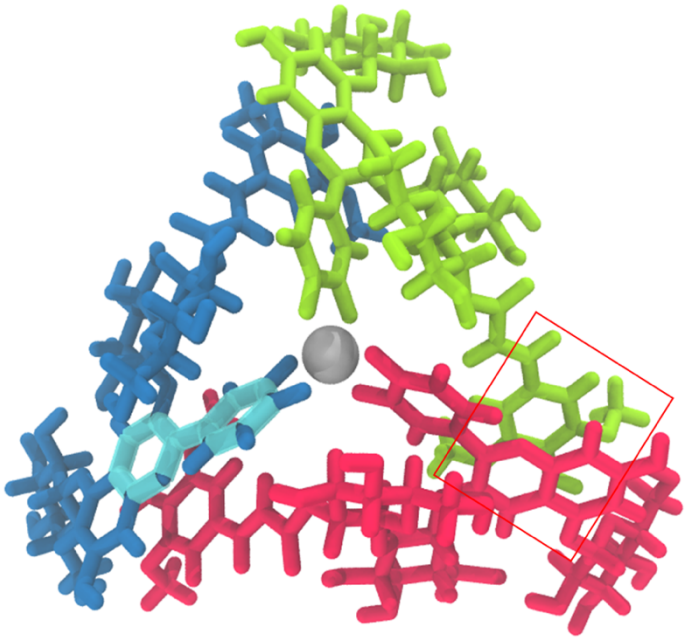
Supplementary Figure 8.1. Proposed structures for the complex of P2 around aluminum, $(Al(P2[odt])_3$ – last item in table above).

"Cyanidin units are highlighted with shades of blue and rendered with thick licorice models, while the rest of P2 molecules are thin and colored black, red and white for carbon, oxygen and hydrogen atoms, respectively. The central aluminum ion is represented as a pink sphere.

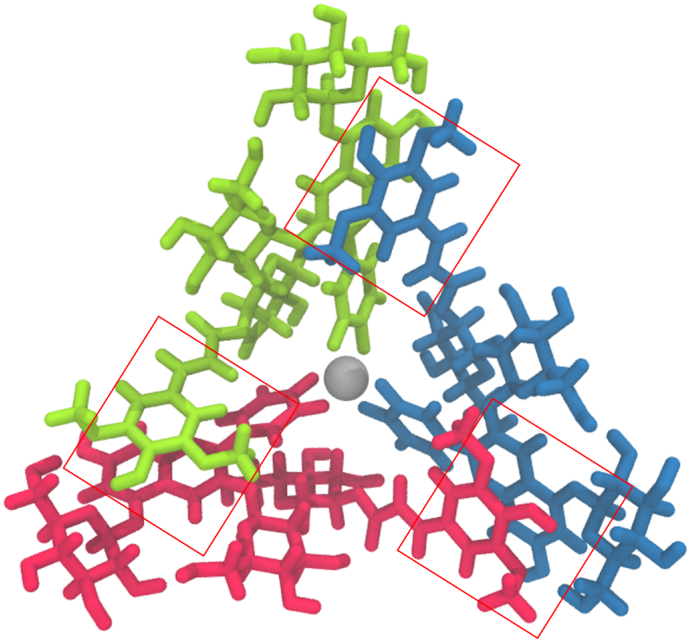
Supplementary Discussion 8.2. Molecular Modeling Confirmation of Observed Cotton Effects in aqueous solution.

In terms of electronic levels and the resulting excited states, the Al^{3+}(P2-)3 can be described in terms of three “strongly” interacting cyanidin chromophores, therefore determining excitonic features derived from 3-fold symmetry arrangements and responsible for the observed Cotton Effect and for a further red-shift of the main absorption peak (contributing to blueing). The sign of the Cotton Effect (positive or negative) is in turn the consequence of the orientation of the chromophores around the metal, namely i) chirality around the metal center (Δ or Λ) and ii) B-ring rotamers. Cotton Effect appears to be positive for Δ arrangement with P2 (or P5) conformations where the B-ring is oriented with O-catechol(3’) on the upper/external side (dihedral angle C3-C2-C1’C2’ ~ ±180 deg, equivalent to [xuy] conformers in Supplementary Table 8.1) – opposite to down orientation of B-ring, corresponding to dihedral angle C3-C2-C1’C2’ ~ ±0 deg,).. The relation between such structural properties and the calculated spectra, the Cotton Effect in CD spectra, is demonstrated by the four panel in Suppl. Figure 8.5. A more complete study of the interplay between dynamic, structural and electronic effects of metal coordination has been

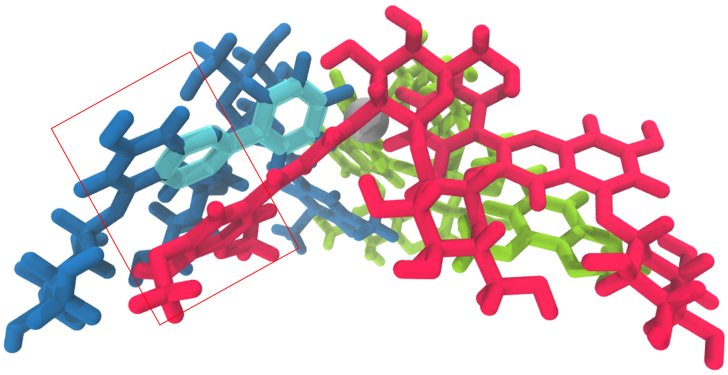
undertaken in our group in the smaller system of cyanin 3:1 with Al^{3+}, available soon in a separate publication.



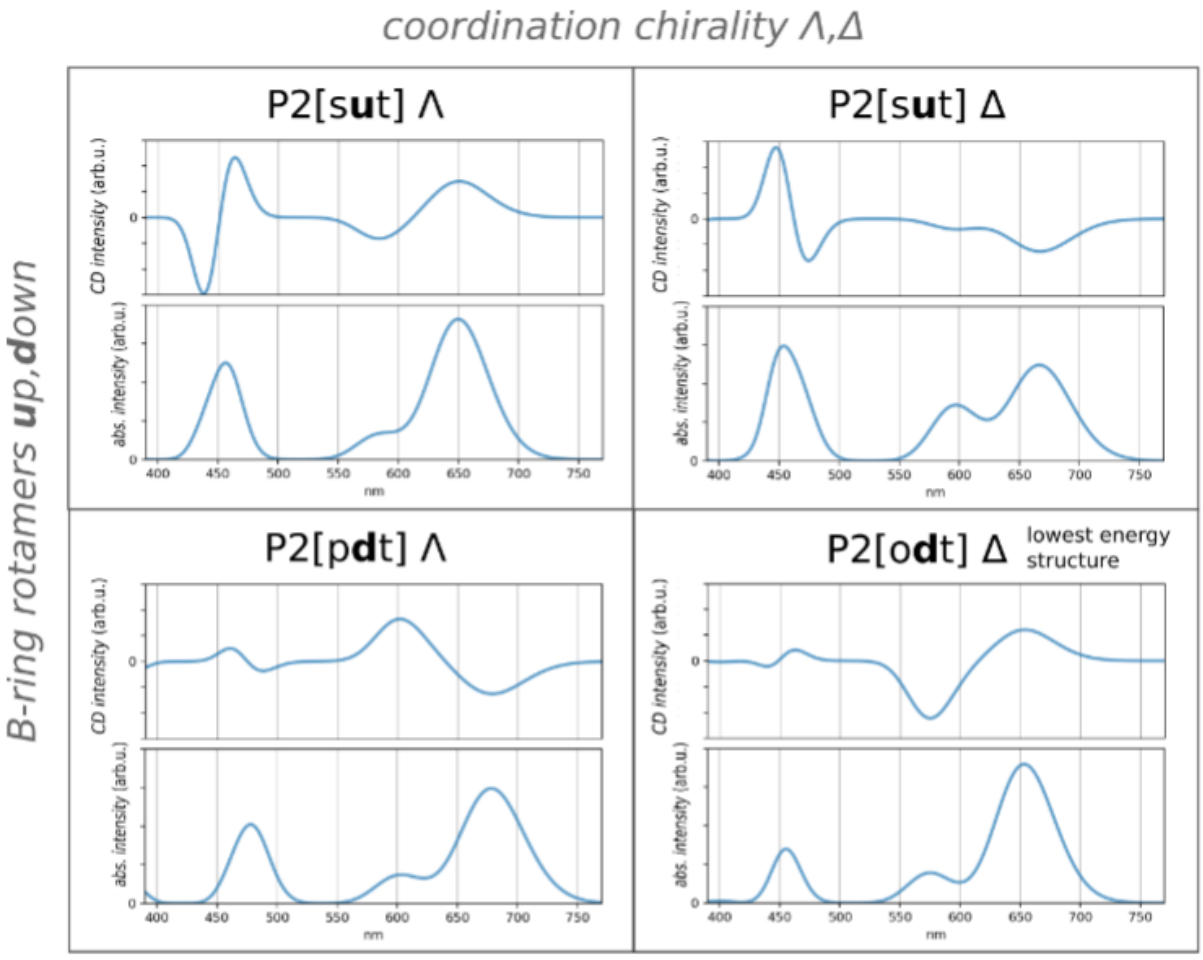
Supplementary Figure 8.2. Proposed 3D model of the P2:Al 3:1 complex. Proposed 3D model of the $Al^{3+}(P2^{-})_3$ complex structure where the highlighted cyan portion represents the torsional distortion angle between A and B ring of the anthocyanin (Also shown in Figure 3A θ angle). The atoms of each P2 unit are colored either in red, blue or green to differentiate between distinct molecules. The Al^{3+} ion is represented as a silver sphere, displaying octahedral coordination to the phenolic moieties of the chromophores. Acyl moieties (sinapoyl on P2) overlap in a parallel fashion (π-π orbital overlap) with adjacent chromophore of the neighboring unit (highlighted by the red box surrounding the sinapoyl of one P2 (green)) and cyanidin of another P2 (red) supporting the propeller-like arrangement.



Supplementary Figure 8.3. Intermolecular interactions in the 3D model of the $Al^{3+}(P2^{-})_3$ complex structure. Intermolecular (between P2 units) interlocking between structural units (highlighted by three red boxes) of the complex is forms the propeller-like arrangement around the Al^{3+}. This is another representation of complex depicted in Supplementary Fig. 8.2 to help with visualization.



Supplementary Figure 8.4. Proposed 3D model of the $Al^{3+}(P2^{-})_3$ complex structure viewed from the plane of the complex. This angle shows both the θ angle distortion that enhances the blueing as well as the interlocking between two distinct P2's enhancing the stability.



Supplementary Figure 8.5. Optical properties for a set of Al^{3+} $(P2^{-})^{3}$ complex structures in relation with their structures. The absorption and CD spectra calculated following Suppl. Method 8.2 are presented for four structures reported in the Suppl. Table 8.1 and corresponding to different combination of chirality (left vs. right) and B-ring rotamers (top vs. bottom). Lowest energy candidate P2[odt] (Δ) shows positive Cotton Effect as found experimentally.

Supplementary Table 8. 1. Selection of the lowest energy structures for P2 and P5, together with some of their properties. Distortions are described by θ (deg). Energies are listed relative to lowest energy structure, shown in Figure 8.1 and 8.2-8.4, which energy is set to 0. Cotton effects were estimated calculating optical spectra as detailed in Suppl. Method 8.2. Labels in square parentheses refer to conformation of the ligands (P2 or P5).

Ligand X: Al(X)$_3$	*chirality*	*θ(deg)*	*cotton effect*	*energy(eV)*
P5[puc]	Δ	4.0	-	0.18
P5[puc]	Λ	1.7	+	0.80
P2[pdt]	Δ	21.7	+ (weak)	1.17
P2[pdt]	Λ	27.0	-	0.58
P2[sut]	Δ	15.5	-	1.11
P2[sut]	Λ	8.9	+	1.55
P2[odt]*	Δ	24.1	+	**0**
*See Supplementary Figure 8.1				

Supplementary Discussion 8.3. Importance of the position of sinapoyl on regio-isomers P2 and P5 in enabling complex formation. The monoacylated anthocyanins P2 and P5 differ only in the position of the sinapoyl moiety, however this proves to lead to radical differences in their available conformations, stability, and metal complexation. Distortions of the chromophore hold a major role and are perhaps the main cause of blueing, with further blueing in some cases due to electronic interactions between chromophores and acyl groups. Our simulations show that plausible configurations result in distorted chromophores for P2, but not P5. Along with the compactness of these hypothetical structures,

this results in a complex which is a) distorted and thus blue, and b) resistant to nucleophilic attack at C2 and degradation through protection of the chromophore by the tightly stacked and interlocked acyl units.

Supplementary Section 9. Physicochemical properties of red cabbage anthocyanins in solution.

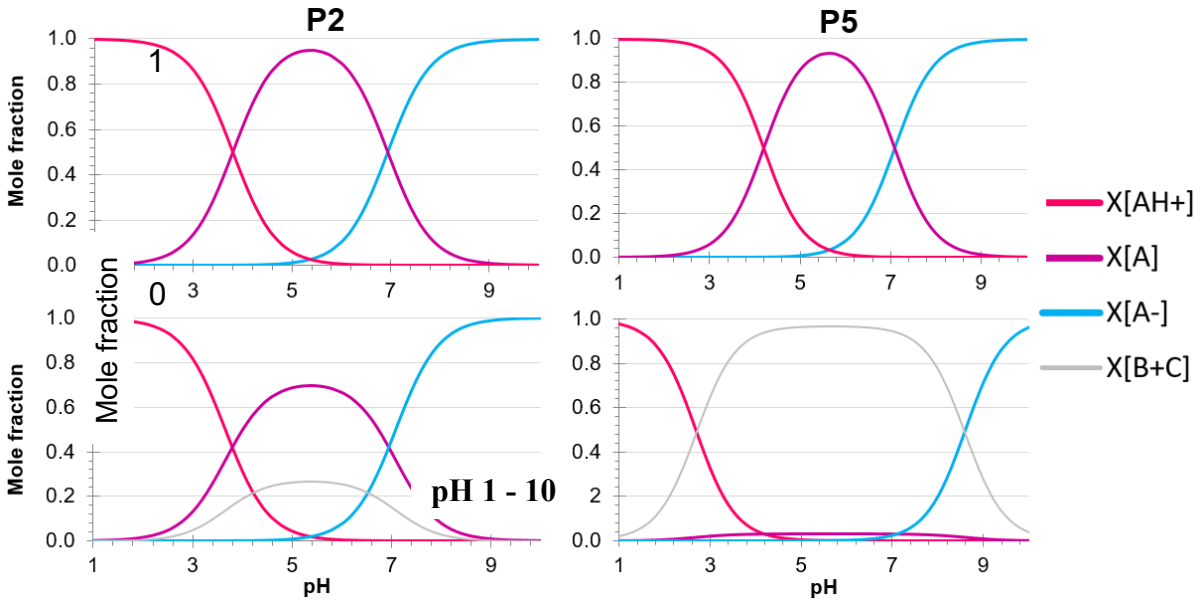
Supplementary Discussion 9.1. Determination of the physiochemical properties of the non-, mono- and di-acylated anthocyanins in red cabbage. As described in Supplementary Section 3, anthocyanins acylated by hydroxycinnamic acid (HCA) residues are prone to develop inter- and intramolecular π-stacking interaction *(5, 16)* that make the flavylium nucleus less accessible to water molecules, thus slowing down hydration and decreasing the percentage of colorless forms at equilibrium (lower K'_h, Supplementary Table 9.1). This improved color stability is modest with monoacylated P5 ($pK'_h = 2.7$, *vs*. 2.1 for nonacylated P1) but more significant with diacylated P8 ($pK'_h = 3.7$), possibly owing to P8 adopting sandwich-type conformations (intramolecular copigmentation) in which each sinapoyl stacks to each face of the anthocyanidin, thereby providing optimal protection. Most surprising is the observation that P2 ($pK'_h = 4.4$) is exceptionally resistant to water addition, despite its monoacylated status. Moreover, among the red cabbage anthocyanins, P2 specifically displays two remarkable long-range NOE correlations between cyanidin and the sinapoyl residue in CD_3OD/TFA-*d*, *i.e*. in a solvent that does not favor π-stacking interactions (Supplementary Fig. 5.2-5.3). Hence, it is proposed that the β-D-Glc(2-sinapoyl)-1,2-β-D-Glc-3-cyanidin sequence within monoacylated P2 is optimal for the development of strong π-stacking interactions between the HCA residue and the chromophore, thereby providing efficient protection of the flavylium ion against hydration. The simple model already developed to analyze intramolecular *copigmentation (16)* can be adapted to interpret the strong stabilization of the flavylium nucleus observed when a single sinapic residue is attached to the Glc-2 moiety of the 3-O-sophorosyl

group. In the model, it is assumed that the much lower K'_h value of P2 ($pK'_h{}^{P2} = 4.2$) compared to non-acylated P1 ($pK'_h{}^{P1} = 2.1$) is rooted in the lower chemical potential of the flavylium ion while that of the hemiketal is unchanged. In other words, the molecular interactions providing protection against water addition for P2 operate via the stabilization of the flavylium ion. Thus, the flavylium ion of P2 can be represented as a mixture of open (extended) and folded conformations with the former having the same propensity for water addition as the flavylium ion of P1. These assumptions can be translated into eq. 9.1 (x_{open}, x_{folded} = mole fractions of open and folded flavylium conformations):

$$\frac{K'^{P2}_h}{K'^{P1}_h} = \frac{[AH^+]_{open}}{[AH^+]_{open} + [AH^+]_{folded}} = x_{open} = 1 - x_{folded} \qquad \text{(Eq. 9.1)}$$

From the pK'_h values, it can be estimated that the open conformation of P2 is of the order of 1%, thus the anthocyanin chromophore is almost exclusively under its folded protected form.

The speciation diagrams of regio-isomers P2 and P5 (Supplementary Figure 9.1) demonstrate the critical impact of the acyl position on the efficiency of intramolecular copigmentation. The two top diagrams in Supplementary Figure 9.1 were constructed by only considering the fast proton transfer equilibria between the colored forms, *i.e.* hydration is neglected. The difference between P2 and P5 is slight, meaning that these equilibria are only weakly affected by intramolecular copigmentation. The two diagrams in Supplementary Figure 9.1 at the bottom include hydration and thus reflect the full equilibrium situation. Here, the contrast is impressive: water addition to the flavylium ion (slowly forming the colorless hemiketal) is clearly privileged over the deprotonation reaction with P5 ($K'_h > K_{a1}$) and the hemiketal accumulates under mildly acidic conditions (up to 90% at pH 5 – 6). On the contrary, with P2, the colored base remains dominant (up to 70% at pH 5 – 6), which is advantageous for a more intense purple to blue color over the range of pH 5 – 8.



Supplementary Figure 9.1. Speciation diagrams of P2 and P5. These diagrams were constructed from data in Supplementary Table 9.1. Mole fractions were calculated across pH 1-9. *Top*: hydration neglected. *Bottom*: hydration included (equilibrium).

Supplementary Table 9.1. Kinetic and thermodynamic parameters for the structural transformations of red cabbage anthocyanins in acidic to mildly alkaline solution (25°C).

Anthocyanin	Structure	pK'_a	pK_{a1}	pK'_h	k_h (s^{-1})	k'_{-h} (M^{-1} s^{-1})	pK_{a2}
1 a)	Cya-3-[Glc -2-Glc]-5-Glc	2.14 (± 0.02)	4.6 (± 0.2)	2.1	315 (± 19) x10^{-3}	84 (± 8)	7.18 (± 0.08)
2	Cya-3-[Glc -2-Glc(**2-Sp**]-5-Glc	3.66 (± 0.06)	3.8 (± 0.2)	4.2	96 (± 27) x10^{-3}	590 (± 93)	6.94 (± 0.05)
5 a)	Cya-3-[Glc(**6-Sp**) -2-Glc]-5-Glc	2.69 (± 0.03)	3.7 (± 0.1)	2.7	143 (± 13) x10^{-3}	154 (± 10)	7.09 (± 0.04)
8 a)	Cya-3-[Glc(**6-Sp**) -2-Glc(**2-Sp**]-5-Glc	3.57 (± 0.03)	4.2 (± 0.1)	3.7	6 (± 2) x10^{-3}	149 (± 42)	7.20 (± 0.06)

a) Values from *(21)*

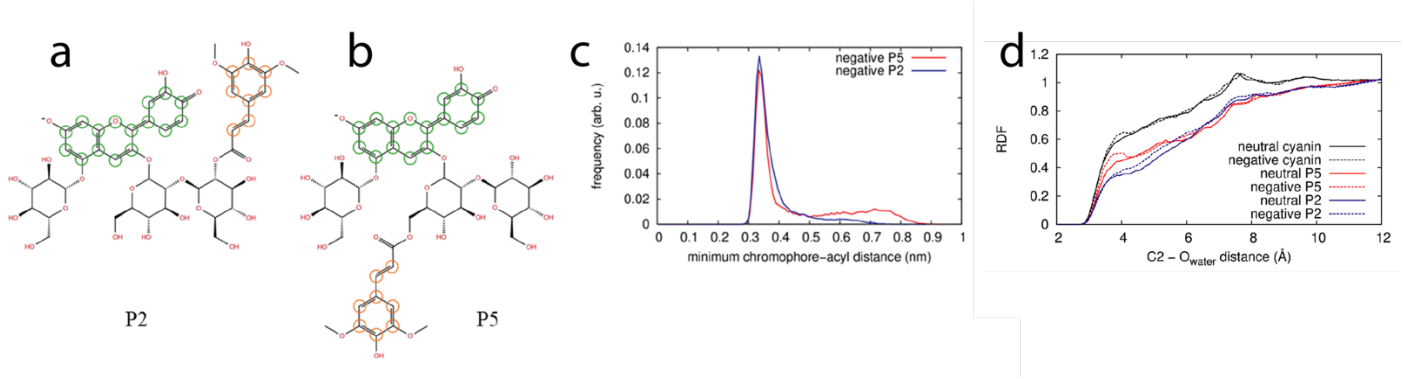
K_{a1}, K_{a2}: first and second acidity constants; K'_h: apparent hydration constant, $K'_a = K_{a1} + K'_h$: global acidity constant.

Supplemental Section 10. Molecular dynamics simulations reveal the conformational landscape of mono-acylated anthocyanins P5 and P2.

Supplementary Method 10.1. M**olecular dynamics (MD) simulations of P2 and P5.**

In the following, some results are presented for the negative charge state, deprotonated in 4' and 7, responsible for the blue hues of P5 and P2. Classical MD simulations of P5 and P2 were run using GROMACS 4.5.5 *(48)*, upon adequate equilibration with all bonds to hydrogen atoms constrained using LINCS *(49)* General AMBER Force Field (GAFF) parameters were assigned using the antechamber module of AmberTools13 *(50, 51)* with RESP charges at the HF/6-31G* level, calculated on DFT-B3LYP optimized geometries, 50/200 Ry basis set. For each species, ten replicas were simulated using Hamiltonian replica-exchange MD (HREMD) *(52)* as implemented in the PLUMED plugin *(53)*. The timestep was 1fs and replica exchange attempts were made every 100 timesteps, resulting in an average acceptance rate of 25% between adjacent replicas, for a total of 5 μs (500 ns per replica). The P2 and P5 molecules were fully solvated in TIP3P water, using cubic cells of 4 nm edge accommodating 2,130 water molecules, in periodic boundary conditions. The volume was kept fixed, and velocity rescaling *(65)* was used to keep the average temperature constant (NVT ensemble). In each case only the trajectory of the first replica, at 300 K, was retained for analysis. Example input files can be found at https://github.com/siegel-lab-ucd/blue-pigment-publication.git. To quantify the relative abundances of open to closed conformations, we track the minimum distance between atoms belonging to the chromophore moiety, those belonging to the acyl moiety. The distribution of this minimum distance enables us to determine how likely the molecules are to be in open (no stacking) or closed (-stacked) configurations. The results are shown in Supplementary Fig. 10.1c and reveal that P2 exists in a closed conformation over 90% of the time whereas P5 can have up to 30% open conformations. From these

trajectories we also identified relevant conformers by using a clustering algorithm on 10 key dihedrals of the molecules (*66)*.



Supplementary Figure 10.1. Distribution of minimum difference between acyl groups and chromophore for neutral bases. (a) P2 (b) P5 (c) Distribution of minimum distance between chromophore (green) and acyl (orange) moieties for anionic bases (removal of 2 protons) of P5 (red) and P2 (blue). Selecting a cut-off distance of 0.5 nm, closed structures represent 70% of the structures for P5, and 91% of the structures for P2. (d) Radial distribution function (RDF) of the carbon-oxygen distance between the C2 carbon and oxygen atoms from the water molecules

Supplementary Discussion 10.1. Acylation increases stability with respect to irreversible degradation by nucleophilic attack. Irreversible degradation of anthocyanins occurs by nucleophilic attack at C2, and subsequent decomposition through a network of side reactions. The MD simulations of P5 and P2 in full explicit solvent enable us to obtain a radial distribution function (RDF) between C2 and the water molecules, and thus infer trends in the likelihood of the initial nucleophilic attack occurring, depending on the species. RDF results are shown in Supplementary Fig. 10.1. Water molecules are less likely to come close to C2 in the acylated molecules than in cyanin, due to protection from the acyl unit, which is often stacked to the chromophore, as seen previously. This effect is more pronounced in P2 than in P5, because the folded/stacked conformations occur more often. P2 is thus

predicted to be more resistant to nucleophilic attack, and subsequent degradation, than both P5 and cyanin, as found experimentally.

Supplementary Section 11. Enzyme discovery and design for the enhancement of P2 from RCE.

Supplementary Method 11.1. Construction of the phylogenetic tree of esterases of interest. Sequeneces tested in this study (Supplementary Table 11.1) were aligned using Geneious 2017.10.1.3 using multiple sequence comparison by log-expectation (MUSCLE) alignment, with a maximum of 15 iterations. The distance between sequences was measured first by kmer6_6, and subsequently by pctid_kimura. The clustering method used was UPGMB. The tree rooting method used for all iterations was pseudo. The sequence weighting scheme for all iterations was CLUSTALW. Terminal gaps were assigned half penalty, and the objective score was calculated with sum of pairs scoring. The phylogenetic tree in Fig. 4C of the main text was constructed using the Geneious Consensus Tree Builder. The model utilized the Jukes-Cantor genetic distance model, and the neighbor-joining tree build method. The tree was resampled using bootstrap resampling, with 1,000 replicas. The consensus tree was created with a support threshold of 50%. The tree was visualized using GraPhlAn (https://github.com/siegel-lab-ucd/blue-pigment-publication.git) *(54).*

Supplementary Method 11.2. Protein purification and enzymatic assay of esterases. An *E. coli* codon optimized gene encoding each protein was purchased from Twist Biosciences and transferred into pET29b+ to encode a C-terminal hexahistidine tag. Mutants plasmids were produced by Kunkel mutagenesis *(55)*. Plasmids were incorporated into *Escherichia coli* BL21(DE3) via electroporation. 500 mL cultures were grown in Terrific Broth at 37 °C, induced with 1mM IPTG, and allowed to express at

18 °C for 24 hours. Cells were lysed via sonication and centrifuged at 4,700 RPM for 1 hour to remove cell debris. Our target proteins were purified with immobilized metal affinity chromatography and eluted 1 mL 50 mM HEPES buffer, pH 7.5, with 150mM imidazole. The purity of the proteins was determined with SDS-page, and the concentration by measuring the absorbance of the solution at 280 nm. Proteins were then diluted to 1 mg/mL for activity assays using 50 mM HEPES buffer, pH 7.5, with no imidazole. 90 μL of buffered enzyme was mixed with 10 μL Red Cabbage Extract (RCE) suspended in 1% HCl (100 mg/mL). To detect the presence of activity, assays were allowed to run for 24 hours at room temperature, after which they were quenched with 5% HCl. Assays to determine the rate of reactions were conducted using an enzyme concentration of 0.5 mg/mL. Reaction products were analyzed by high performance liquid chromatography (HPLC-DAD) on an Agilent 1100 Series. The column (Phenomenex Kinetix pentafluorophenyl propyl 2.6 μm particle size, 100 Å pore, 100 x 4.6mm) was held at room temperature. The mobile phase was a gradient mixture of solvent A (4.5% formic acid in Milli-Q water) and solvent B (acetonitrile) as follows: 100% A (0 min), 0% to 30% B (25 min), 30% to 95% B (1 min), 95% B (5 min), 95% to 0% B (1 min), 0% B (13 min). Products were monitored with DAD set to 530 nm. The experiment was repeated for enzymes that demonstrated activity on RCE anthocyanins. For these enzymes the reaction size was increased 5-fold, and a portion was withdrawn and quenched at timed intervals in order to evaluate the rate of the reactions.

Supplementary Method 11.3. Gram Scale Protein Production & Purification. The M73H point-mutant was created from the 1AUR (Sequence ID No. 46) WT plasmid via Kunkel Mutagenesis *(55)*. Then transformed into chemically competent *Escherichia coli* BLR (DE3) cells and plated on agar plates containing 50 μg/mL kanamycin. Single colonies were picked and used to inoculate 10 mL Terrific Broth (TB) in 50 mL Falcon tubes. These were incubated at 37°C while shaking at 300 RPM (10 RCF) for 12-

16 hours. For long term storage of the inoculated cells, a glycerol stock was made with 500 μL of the cell culture and 500 μL of 50% autoclaved glycerol in Milli-Q water and stored in -80ºC freezer. To begin protein production, the starting culture consisted of a 50 mL Terrific Broth (TB) with 50 μg/mL kanamycin and 15 μg/mL tetracycline hydrochloride in a 250 mL Erlenmeyer baffled flask. This was inoculated with a swab of cells from the glycerol stock and shaken at 250 RPM at 37ºC. After 14-16 hours, the starter culture was poured into a 1 L Erlenmeyer baffled flask containing 450 mL TB with 50 μg/mL kanamycin and 15 μg/mL tetracycline hydrochloride. These flasks were covered with breathable seals and shaken at 300 RPM (10 RCF), at 37ºC for 4-6 hours until the OD_{600} reached 0.7-1.0. The 500 mL cultures were then cooled at 4ºC for 30 minutes, and enough isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added for a final concentration of 1 mM and shaken again at 300 RPM (10 RCF), at 18ºC for 24 hours to induce expression of the M73H mutant enzyme. After expression, the cultures were pelleted by being spun down for 10 minutes at 4,700 RPM (2,474 RCF) at 4ºC. The supernatant was poured out and the pelleted cells were resuspended in 10 mL phosphate buffered saline (PBS) pH 7.4 (135 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 1.4 mM KH_2PO_4) and stored in 50 mL falcon tubes at -20ºC. For purification, frozen cells containing enzyme were thawed on ice and brought up to a volume of 40 mL with PBS. Cells were sonicated, and the lysed cells were spun down for 1 hour at 4,700 RPM (2,474 RCF), at 4ºC. The supernatant lysate containing active enzyme was collected and used for the reaction.

Supplementary Method 11.4. Gram Scale P2 Enzyme Reaction & Product Validation. To determine the yield of enzyme in the lysate, 500 μL of lysate was purified from every batch. A 1 mL column with 200 μL Cobalt beads was washed with 1mL PBS with 10 mM imidazole, pH 7.4. 500 μL of lysate was poured through the column, the His-tagged M73H mutant enzyme were immobilized on the beads via metal ion affinity chromatography and washed with 5 mL of Wash buffer. Enzyme was eluted with 250

μL PBS with 200 mM imidazole, pH 7.4. Protein yield was determined by A280 using a BioTek Epoch spectrophotometer. To setup the reaction, the substrate, lyophilized red cabbage extract, was dissolved in PBS pH 7.4. This was combined with an appropriate amount of lysate for a final ratio of 1 mg of enzyme for every 200 mg of RCE substrate. For example, 150 mL of lysate at 1 mg/mL of enzyme would be combined with 30 g RCE substrate dissolved in 600 mL of PBS. The assay reaction proceeded for 4 days in a 1 L Nalgene Amber bottle, while being gently agitated on a rocker. The reaction was then quenched by bring down the pH to 2.8 and frozen at -20°C. To verify reaction completion, 500 μL of product was combined with 500 μL of acetonitrile in a 2 mL centrifuge tube and spun down at 10,000 RPM (11,200 RCF) for 10 minutes at 4°C. 100 μL of the supernatant sampled to confirm reaction progress via the HPLC described in Supplementary Method 11.1.

Supplementary Table 11.1. Enzymes tested for cleavage of RCA acyl groups

ID#	EC #	UniProt ID	PDB	Enzyme Classification	Domain	Activity on RCE?	Database mining round
1	3.1.1.1	Q1QYJ5		Carboxylesterase	Bacteria	Yes	3
2	3.1.1.1	A0A1J8PUW1		Carboxylesterase	Bacteria	Yes	3
3	3.1.01.-	Q5NI32	4F21	Carboxylesterase/phospholipase	Bacteria	Yes	3
4	3.1.1.57	O87170	4DI9	2-pyrone-4,6-dicarbaxylate hydrolase	Bacteria	No	1
5	3.1.1.45	P0A114	1ZI8	Dienelactonase mutant	Bacteria	No	2
6	3.1.1.1	Q06174	1TQH	Carboxylesterase	Bacteria	Yes	2
7	3.1.3.1	B8Y562	4RGY	Alkaline esterase	Bacteria	Yes	3
8	3.1.1.91	F8QQ74		2-oxo-3-(5-oxofuran-2-ylidene)propanoate lactonase	Bacteria	No	1
9	3.5.2.6	F8CAF0		Beta-lactamase	Bacteria	Yes	3
10	3.1.1.27	Q988B9	3AJ3	4-pyridoxolactonase	Bacteria	No	1
11	3.1.1.92	A6XIG7		4-sulfomuconolactone hydrolase	Bacteria	No	1

12	3.5.2.6	B2BSN6		Beta-lactamase	Bacteria	No	1
13	3.5.2.6	P00811	2FFY	Beta-lactamase	Bacteria	No	1
14	3.1.1.65	A3LZU8		L-rhamnono-gamma-lactonase	Eukaryote	No	1
15	3.1.8.1	P15034	2V3Z	Xaa-Pro aminopeptidase	Bacteria	No	1
16	3.1.1.68	R0CVD2		Xylono-1,4-lactonase	Bacteria	No	1
17	3.1.1.1	G4RFI7	5DWD	Esterase	Bacteria	Yes	3
18	3.1.1.83	Q9EX73		Monoterpene epsilon-lactone hydrolase	Bacteria	No	1
19	3.1.1.15	Q1JUP5		L-arabinolactonase	Bacteria	No	1
20	3.2.1.8	P51584	1GKK	Feruloyl esterase	Bacteria	No	2
21	3.1.1.17	B0RN69	3DR2	Exported gluconolactonase	Bacteria	No	1
22	3.1.1.73	O42807	1USW	Ferulic acid esterase	Bacteria	yes	3
23	3.1.1.25	B2LYJ5		Levo-lactonase	Eukaryote	No	1
24	3.1.1.17	A9CPS8		Lactonase	Eukaryote	No	1
25	3.5.2.6	Q8GCU7		Metallo-beta-lactamase VIM-6	Bacteria	No	1
26	3.5.2.6	Q8KRJ3	2WHG	Beta-lactamase VIM-4	Bacteria	No	1
27	3.1.1.84	Q9L9D7	1JU3	Cocaine esterase	Bacteria	No	2
28	3.1.1.31	Q9GRG6	3EB9	6-phosphogluconolactonase	Eukaryote	No	1
29	3.1.1.24	Q13KT2	2XUA	3-oxoadipate enol-lactonase	Bacteria	No	1
30	3.1.1.1	A0A0M3KKY6	4UHC	Carboxylesterase	Bacteria	No	2
31	3.1.1.95	Q54528	1Q0R	Aclacinomycin methylesterase	Bacteria	No	2
32	3.1.1.81	P0CJ63	3DHA	N-acyl homoserine lactonase AiiA	Bacteria	No	1
33	3.1.1.11	P0C1A9	2NSP	Pectinesterase A	Bacteria	No	2
34	3.1.1.1	O28558	1JJI	Carboxylesterase	Archaea	No	2
35	3.1.1.4	A0A243LS46		Esterase	Bacteria	no	3
36	3.1.1.1	Q7SIG1	1EVQ	Carboxylesterase	Bacteria	no	3
37	3.1.1.2	B5BLW5	5L2P	Arylesterase	Archaea	Yes	1
38	3.1.1.1	F0NDQ1	5LK6	Lipase carboxylesterase	Bacteria	Yes	3
39	3.1.1.1	Q976W8	3AIK	Carboxylesterase	Bacteria	Yes	3
40	3.1.1.1	A0A139SSC5		Carboxylesterase	Bacteria	Yes	3

41	3.1.1.1	A0A2U2AP80		Carboxylesterase	Bacteria	Yes	3
42	3.1.1.1	L8MEL5		Carboxylesterase	Bacteria	No	3
43	3.1.1.1	A0A127MYW2		Carboxylesterase	Bacteria	Yes	3
44	3.1.1.1	Q9HXE7	3CN9	Carboxylesterase	Bacteria	Yes	2
45	3.1.1.1	A0A078LXQ1		Alpha/beta-hydrolase	Bacteria	Yes	3
46	3.1.1.1	Q53547	1AUR	Carboxylesterase	Bacteria	Yes	2

Supplementary Table 11.2. Amino acid sequences of enzymes tested in this study as used in Supplementary Method 11.1 (See Supplementary Table 11.1 for enzyme identification)

ID #1
MADSSPLIIEPRQARAADATVILLHGLGADGHDFEPLVPALPLAKDLAVR FVLPHAPRMPVTVNGGMEMPAWYDILDMNLGRRIDEAQLKASADMVHGLI DAEIARGIDSRRIIVAGFSQGGAVAYHAALTYPKPLGGLLALSTYFATAT SIEPSEANRALPIEVHHGSFDPVVPEALGHEGAERAEALGYAVTYRTYPM QHALCPEQIEDIGQWLNARLGAKEAGGSLEHHHHHH
ID #2
MTAPGELIIEPKDGQPADACVFIIHGLGADGHDFEPLVPALALPKDSRVR FIMPHAPRLPVTINGGMVMPAWYDILAMDLGRRVDERQLKQSAERIQALI QEQIDQGIDSQRIIVAGFSQGGAVAYHAALTFPAPLGGLLAMSTYFATAD NIDLAEANRQIPIEVQHGNFDPIVPESLGRSGADRLKEMGYAVNYRQYPM AHALCPQQVNDIGKWLSARLNGGSLEHHHHHH
ID #3
MNYELMEPAKQARFCVIWLHGLGADGHDFVDIVNYFDVSLDEIRFIFPHA DIIPVTINMGMQMRAWYDIKSLDANSLNRVVDVEGINSSIAKVNKLIDSQ VNQGIASENIILAGFSQGGIIATYTAITSQRKLGGIMALSTYLPAWDNFK GKITSINKGLPILVCHGTDDQVLPEVLGHDLSDKLKVSGFANEYKHYVGM QHSVCMEEIKDISNFIAKTFKIGGSLEHHHHHH
ID #4

MTNDERILSWNETPSKPRYTPPPGAIDAHCHVFGPMAQFPFSPKAKYLPR DAGPDMLFALRDHLGFARNVIVQASCHGTDNAATLDAIARAQGKARGIAV VDPAIDEAELAALHEGGMRGIRFNFLKRLVDDAPKDKFLEVAGRLPAGWH VVIYFEADILEELRPFMDAIPVPIVIDHMGRPDVRQGPDGADMKAFRRLL DSREDIWFKATCPDRLDPAGPPWDDFARSVAPLVADYADRVIWGTDWPHP NMQDAIPDDGLVVDMIPRIAPTPELQHKMLVTNPMRLYWSEEMGGSLEHH HHHH
ID #5
MLTEGISIQSYDGHTFGALVGSPAKAPAPVIVIAQEIFGVNAFMRETVSW LVDQGYAAVCPDLYARQAPGTALDPQDERQREQAYKLWQAFDMEAGVGDL EAAIRYARHQPYSNGKVGLVGYCLGGALAFLVAAKGYVDRAVGYYGVGLE KQLKKVPEVKHPALFHMGGQDHFVPAPSRQLITEGFGANPLLQVHWYEEA GHSFARTSSSGYVASAAALANERRLDFLAPLQSKKPGGSLEHHHHHH
ID #6
MKIVPPKPFFFEAGERAVLLLHGFTGNSADVRMLGRFLESKGYTCHAPIY KGHGVPPEELVHTGPDDWWQDVMNGYEFLKNKGYEKIAVAGLSLGGVFSL KLGYTVPIEGIVTMCAPMYIKSEETMYEGVLEYAREYKKREGKSEEQIEQ EMEKFKQTPMKTLKALQELIADVRDHLDLIYAPTFVVQARHDEMINPDSA NIIYNEIESPVKQIKWYEQSGHVITLDQEKDQLHEDIYAFLESLDWGGSL EHHHHHH
ID #7
MALFQCDFFSDVLGLSTSMTVILPQETTGQIGMAGGSERREHPTLFLLHG LSDDHTIWLRRTSIERYVAEMGLAVVMPAVHRSFYTDMAHGLQYWTFISE ELPALARSFFPLATAREDTFVAGLSMGGYGALKLGMRHPERFAAAASLSG ALDITFDPAEHIAMEDDVWVAEQRNIFGDLAALPGSDHDLFALAERMAQS DGPVPKLYQCCGTEDFLYEDNVRFRDHVRGLGLDFMYEESPGEHEWGYWD AQIQRVLAWLPLRPPGTAPAGGSLEHHHHHH
ID #8

MATETIAMDWVDIGTNGESRLAYLARPVVTGRLPAVIVMPAIHGINTYIK

DVAIDLAKAGFVALLIDIHSPEQEPDLSNAEKIQIAVETLDDRKVLKDVD

AAVRYLEQHAAVRADRLGILGFCVGGTYALLAARTPAIRVSVGFYGLLEY

QSRTDNKPVSPLDSVAQFTAPILFHVGDKDPWIDSKMLAEFTKRMQQHQK

SYELCIYRGAGHAFHEHFRDAYRPIAAQSAWNNTLIYLRWHLCGKRTVGG

SLEHHHHHH

ID #9

MNGLRWRLTGVVMAWVLVAPLAEAANVKQEVDRYISGFHQKGLFNGTVLV

ANERGILLKKGYGAANLEWKVPNAPDTKFRIGSITKSFTATVILQLAAEG

KLQLDDPITKHLPDYRKDTGDRVTITHLLNHTSGIPSYTSKPAIMKDADG

FESVAAFVKKACSDDLEFEPGTKYAYNNSGYFLLGAIIEKLTGQTYAEAV

QARILGPLGMKDTGYDVSATVLPKRASGYAQAPGGIVNAAWLDMNLPYAA

GSLYSTVEDLYRWERAFHGDTLLPAALKQKMLTPGLAHYGFGWVMSDMTL

HDGKTKLPGIFHTGGINGFSSILVRVPERKEAVILLDNMTHGGLQELAGG

VLSILHGLTPRPARMPIGNVMMESLGKGSVAQAIATYRTLKKTKQAEYDF

SERHLNTVGYHLLRSGRAADAIEVFKLNVEMFPEAANCHDSLGEAYAAHG

DKARAITSYRKALELAPKNEHAVKMLEQLEEPAAKRGGSLEHHHHHH

ID #10

MSDTKVYLLDGGSLVLDGYHVFWNRGPGGEVRFPVYSILIEHAEGRFLID

TGYDYDHVMKVLPFEKPIQEKHQTIPGALGLLGLEPRDIDVVVNSHFHFD

HCGGNKYFPHAKKICHRSEVPQACNPQPFEHLGYSDLSFSAEAAEARGAT

AQLLEGTTRANSTFEGIDGDVDLARGVKLISTPGHSIGHYSLLVEFPRRK

PILFTIDAAYTQKSLETLCQAAFHIDPVAGVNSMRKVKKLAEDHGAELMY

SHDMDNFKTYRTGTQFYGGGSLEHHHHHH

ID #11

MLPADQAGIPPCQGPRARSAPISFAIPKGAWDTHLHVFGPTAVFPYAEKR

PYTPPDSPLEDYLALMERLGIERGVCVHPNVHGIDNSVTIDAVERSDRRL

LGIIKPHRVMTFTELRDLKTRGVRGVRFAFNPQHGSGALDTELFERMHGW

CRELDWCINMHFAPDALEGLCDLIAGAETPIIIDHFGRVETAAGVNQLPF

KILRDLATLDHVWIKLTGADRISHSGVPYDDVVPFAHALSEIAPDRLLWG SDWPHSGYFDPKRMPDDGDLLNLVARFAPDVALRHKILVDNPARLFGVIG GSLEHHHHHH
ID #12
MKQRIALSLLALGPLLLVPRVYAAADEPMANIVEKAVQPLLEEYRIPGMA VAVLKEGKPHYFNYGVANRESGRRISERTLFEIGSVSKTFTATLGTYAVV KGGFRLDDKVSQHAPWLQNSAFDRVTMAQLATYSAGGLPLQFPDAVDSNE RMRQYYRQWSPLYAAGTHREYSNPSIGLFGHLAASTLGQPFRQLMSQTLL PKLDLQHTYLEVPDAAMVDYAYGYSKEDKPVRVNPGVLADEAYGIKTSAA DLIKFVGANMTGSGDKAVQQALAMTRTGFYSVGEMTQGLGWESYAYPVTE QALLAGNSPAVSFKANPVKPFVAPRVMGNERLYNKTGSTNGFGAYVVFVP ARGVGIVMLANRNYPIEARVKAAYAIMRHLAPGGSLEHHHHHH
ID #13
MFKTTLCALLITASCSTFAAPQQINDIVHRTITPLIEQQKIPGMAVAVIY QGKPYYFTWGYADIAKKQPVTQQTLFELGSVSKTFTGVLGGDAIARGEIK LSDPTTKYWPELTAKQWNGITLLHLATYTAGGLPLQVPDEVKSSSDLLRF YQNWQPAWAPGTQRLYANSSIGLFGALAVKPSGLSFEQAMQTRVFQPLKL NHTWINVPPAEEKNYAWGYREGKAVHVSPGALDAEAYGVKSTIEDMARWV QSNLKPLDINEKTLQQGIQLAQSRYWQTGDMYQGLGWEMLDWPVNPDSII NGSDNKIALAARPVKAITPPTPAVRASWVHKTGATGGFGSYVAFIPEKEL GIVMLANKNYPNPARVDAAWQILNALQGGSLEHHHHHH
ID #14
MSKYKILDSHIHLYSLANIPLLHWDEGNPLHGNRRLDEYIENSQSTQFDV EGVVWIECDAKIDLTQGLKGLENPIEEYLYICRNINGKLLPEEGVSTPFK RRLIKAMIPFAPMPLGSAGVEEYVKALKTRNSSEFHLVKGFRYLIQDKPP LTISDPHFVSSFQWLDSNGYVFDLGIDMRSGGLWQFKETLEVFKKVPNLK YIINHLTKPCLDFDPETIDSNPDFLSWKRLVTEMYITTPNSYMKLSGGFS EVEQDVALDVTSTSRHVYPWFKVVYELWGPERTIFASNWPVCAIPAGQNL TEKWFQVCETLFDSIGMDEDTRRKIYYSNAFKAYNIGGSLEHHHHHH

ID #15
MSEISRQEFQRRRQALVEQMQPGSAALIFAAPEVTRSADSEYPYRQNSDF WYFTGFNEPEAVLVLIKSDDTHNHSVLFNRVRDLTAEIWFGRRLGQDAAP EKLGVDRALAFSEINQQLYQLLNGLDVVYHAQGEYAYADVIVNSALEKLR KGSRQNLTAPATMIDWRPVVHEMRLFKSPEEIAVLRRAGEITAMAHTRAM EKCRPGMFEYHLEGEIHHEFNRHGARYPSYNTIVGSGENGCILHYTENEC EMRDGDLVLIDAGCEYKGYAGDITRTFPVNGKFTQAQREIYDIVLESLET SLRLYRPGTSILEVTGEVVRIMVSGLVKLGILKGDVDELIAQNAHRPFFM HGLSHWLGLDVHDVGVYGQDRSRILEPGMVLTVEPGLYIAPDAEVPEQYR GIGIRIEDDIVITETGNENLTASVVKKPEEIEALMVAARKQGGSLEHHHH HH
ID #16
MNAVTCVWDLKATLGEGPIWYDDSLWFVDIKSHKIHNYNPATDERFSFDA PEPVTFIAPLAPNARAGFVVGLKSGLHRFHPVMGGFKPLIQVESAELNNR PNDATVDHGGRLWFGTMHDDEEAKSGSLYRMDSTGVARMDKDICITNGPC VSPDGKTLYHTDTLEKIIWAYDLAEDGTLSNKRGFVNFQGENAVYPDGSV VDSEGYLWTALWGGFGVVRISPAGELVARIELPAPNVTKPCFGGPDLKTL YFTTARKGLSDETLAQYPLSGGLFGVRVDVAGQPQHEVRLVGGSLEHHHH HH
ID #17
MTEPVKLSGPMLPAVSGAAKSLVVLLHGYGSDGRDLIALGQFWRDSFPDT MFVAPNAPHVCGGNPFGYEWFPLDLERDRTLARLAGAETAHPVLDAFLAD LWAQTGLGPADTILVGFSQGAMMALYTGLRLPEPLKAIIAFSGLIVAPEK LEAEIASKPPVLLIHGDLDDVVPVIGSETALPKLIDLGIDARLHISQGSG HTIAQDGLDTATAFLREILGGSLEHHHHHH
ID #18
MSATDTARAKELLASLVSMPDATIDDFRALYEQVCATFELPDDAQVEPVD ANGADALWVSAPGVSADTVAVVVHGGGFTMGSAHGYRELGYRLSKSGNLR

ALVVDYRLAPESPFPAPVDDVVAAYRYARSLDGVENVFLVGDSAGGGIAM SALITLRDAGEQLPDAAVVLSPLVDLAGESPSLVDRAHLDPLPAAVLVNG MGGLYLNGLDVRHPVASPMHGDLTGLPATLVLVGTDEGLHDDSTRLVDKL KAADVEVQLEIGEGLPHIWPIFSFHPDAVAATDRIGEFLRSHVAAPRGGS LEHHHHHH
ID #19
MQQIHPAGQATLLADTRNTLGEGATWCDRTRALYWVDIEGAQLWRCRADG SDLTPWPMPERLACFALTDDPDVLLVGLATHLAFFDLRSGAFTRIVEVEP ELPTRLNDGRCDGSGAFVFGMKDEGAEPPRAVGGFYRLNADLTLERLALP PAAIANSIGFSPDGSKMYFCDSLVREIFVCDYRPGGEVANVRPFARLTDP DGDPDGSIVDRDGGLWNAQWGGRRVVRYGPDGVETDRVAVPTAQPSCTAL DGEGRLYVTSARVGLSDDALADDPHAGGVFVAQTRHAGMATARFAGTPRG GGSLEHHHHHH
ID #20
MKNKRVLAKITALVVLLGVFFVLPSNISQLYADYEVVHDTFEVNFDGWCN LGVDTYLTAVENEGNNGTRGMMVINRSSASDGAYSEKGFYLDGGVEYKYS VFVKHNGTGTETFKLSVSYLDSETEEENKEVIATKDVVAGEWTEISAKYK APKTAVNITLSITTDSTVDFIFDDVTITRKGMAEANTVYAANAVLKDMYA NYFRVGSVLNSGTVNNSSIKALILREFNSITCENEMKPDATLVQSGSTNT NIRVSLNRAASILNFCAQNNIAVRGHTLVWHSQTPQWFFKDNFQDNGNWV SQSVMDQRLESYIKNMFAEIQRQYPSLNLYAYDVVNEAVSDDANRTRYYG GAREPGYGNGRSPWVQIYGDNKFIEKAFTYARKYAPANCKLYYNDYNEYW DHKRDCIASICANLYNKGLLDGVGMQSHINADMNGFSGIQNYKAALQKYI NIGCDVQITELDISTENGKFSLQQQADKYKAVFQAAVDINRTSSKGKVTA VCVWGPNDANTWLGSQNAPLLFNANNQPKPAYNAVASIIPQSEWGDGNNP AGGGGGGKPEEPDANGYYYHDTFEGSVGQWTARGPAEVLLSGRTAYKGSE SLLVRNRTAAWNGAQRALNPRTFVPGNTYCFSVVASFIEGASSTTFCMKL QYVDGSGTQRYDTIDMKTVGPNQWVHLYNPQYRIPSDATDMYVYVETADD TINFYIDEAIGAVAGTVIEGPAPQPTQPPVLLGDVNGDGTINSTDLTMLK

RSVLRAITLTDDAKARADVDKNGSINSTDVLLLSRYLLRVIDKFPVAENP SSSFKYESAVQYRPAPDSYLNPCPQAGRIVKETYTGINGTKSLNVYLPYG YDPNKKYNIFYLMHGGGENENTIFSNDVKLQNILDHAIMNGELEPLIVVT PTFNGGNCTAQNFYQEFRQNVIPFVESKYSTYAESTTPQGIAASRMHRGF GGFSMGGLTTWYVMVNCLDYVAYFMPLSGDYWYGNSPQDKANSIAEAINR SGLSKREYFVFAATGSDHIAYANMNPQIEAMKALPHFDYTSDFSKGNFYF LVAPGATHWWGYVRHYIYDALPYFFHEGGSLEHHHHHH
ID #21
MDSHCRVRPAGPAVPADCDPPRITHAALAARLGDARLLTLYDQATWSEGP AWWEAQRTLVWSDLVGRRVLGWREDGTVDVLLDATAFTNGNAVDAQQRLV HCEHGRRAITRSDADGQAHLLVGRYAGKRLNSPNDLIVARDGAIWFTDPP FGLRKPSQGCPADPELAHHSVYRLPPDGSPLQRMADLDHPNGLAFSPDEQ TLYVSQTPEQGHGSVEITAFAWRDGALHDRRHFASVPDGLPDGFCVDRGG WLWSSSGTGVCVFDSDGQLLGHIPTPGTASNCTFDQAQQRLFITGGPCLW MLPLPGGSLEHHHHHH
ID #22
MKQFSAKYALILLATAGQALAASTQGISEDLYNRLVEMATISQAAYADLC NIPSTIIKGEKIYNAQTDINGWILRDDTSKEIITVFRGTGSDTNLQLDTN YTLTPFDTLPQCNDCEVHGGYYIGWISVQDQVESLVKQQASQYPDYALTV TGHSLGASMAALTAAQLSATYDNVRLYTFGEPRSGNQAFASYMNDAFQVS SPETTQYFRVTHSNDGIPNLPPADEGYAHGGVEYWSVDPYSAQNTFVCTG DEVQCCEAQGGQGVNDAHTTYFGMTSGACTWGGSLEHHHHHH
ID #23
MPSSISVLAAGILVPVLGAVAAKLPPTAQIIDQKSFNVLKDVPPPAVAND SLVFTWPGVTEESLVEKPFHVYDEEFYDVIGKDPSLTLIATSDTDPIFHE AVVWYPPTEEVFFVQNAGAPAAGTGLNKSSIIQKISLKEADEVRKGKKDE VKVTVVDSNPQVINPNGGTYYKGNIIFAGEGQGDDVPSALYLMNPLPPYN TTTLLNNYFGRQFNSLNDVGINPRNGDLYFTDTLYGYLQDFRPVPGLRNQ VYRYNFDTGAVTVVADDFTLPNGIGFGPDGKKVYVTDTGIALGFYGRNLS

SPASVYSFDVNQDGTLQNRKTFAYVASFIPDGVHTDSKGRVYAGCGDGVH VWNPSGKLIGKIYTGTVAANFQFAGKGRMIITGQTKLFYVTLGASGPKLY DGGSLEHHHHHH
ID #24
MRTLATVASQTDAWTGEGPVWCAARRCLYYVDLGDTRPGKLHVYHPERCV EEIHDLPAMTKDFTQVTAVTVVQNEPHRLAVATEAGVFLYDCQSGDLRRL TGELQPELPKGSYRSNDGKCDPRGRFLIGTMLFSADAPSGGLFSVAGSTI QQLLTGVTIGNGLAWSANGRTMYFIDSPLKRIDAFEYHLDAGTLGARRTA FDFADYFAQQAGWEEAAPDGMTIDAEGLLWVAIYGGGAALRVDPAKEEVV CRVDCPAKYTTSVALGGPARDTLYITSFRRGDAGPDAGAVFQCRAPAPGP PPAEFRLGGSLEHHHHHH
ID #25
MFKLLSKLLVYLTASIMAIASPLAFSVDSSGEYPTVSEIPVGEVRLYQIA DGVWSHIATRSFDGAVYPSNGLIVRDGDELLLIDTAWGAKNTAALLAEIE KQIGLPVTRAVSTHFHDDRVGGVDVLRAAGVATYASPSTRRLAEVEGSEI PTHSLEGLSSSGDAVRFGPVELFYPGAAHSTDNLVVYVPSASVLYGGCAI YELSRTSAGNVADADLAEWPTSIERIQQHYPEAQFVIPGHGLPGGLDLLK HTTNVVKAHTNRSVVEGGSLEHHHHHH
ID #26
MLKVISSLLVYMTASVMAVASPLAHSGEPSGEYPTVNEIPVGEVRLYQIA DGVWSHIATQSFDGAVYPSNGLIVRDGDELLLIDTAWGAKNTAALLAEIE KQIGLPVTRAVSTHFHDDRVGGVDVLRAAGVATYASPSTRRLAEAEGNEI PTHSLEGLSSSGDAVRFGPVELFYPGAAHSTDNLVVYVPSANVLYGGCAV HELSRTSAGNVADADLAEWPTSVERIQKHYPEAEVVIPGHGLPGGLDLLQ HTANVVKAHKNRSVAEGGSLEHHHHHH
ID #27
MVDGNYSVASNVMVPMRDGVRLAVDLYRPDADGPVPVLLVRNPYDKFDVF AWSTQSTNWLEFVRDGYAVVIQDTRGLFASEGEFVPHVDDEADAEDTLSW

ILEQAWCDGNVGMFGVSYLGVTQWQAAVSGVGGLKAIAPSMASADLYRAP WYGPGGALSVEALLGWSALIGTGLITSRSDARPEDAADFVQLAAILNDVA GAASVTPLAEQPLLGRLIPWVIDQVVDHPDNDESWQSISLFERLGGLATP ALITAGWYDGFVGESLRTFVAVKDNADARLVVGPWSHSNLTGRNADRKFG IAATYPIQEATTMHKAFFDRHLRGETDALAGVPKVRLFVMGIDEWRDETD WPLPDTAYTPFYLGGSGAANTSTGGGTLSTSISGTESADTYLYDPADPVP SLGGTLLFHNGDNGPADQRPIHDRDDVLCYSTEVLTDPVEVTGTVSARLF VSSSAVDTDFTAKLVDVFPDGRAIALCDGIVRMRYRETLVNPTLIEAGEI YEVAIDMLATSNVFLPGHRIMVQVSSSNFPKYDRNSNTGGVIAREQLEEM CTAVNRIHRGPEHPSHIVLPIIKRGGSLEHHHHHH
ID #28
MSFKPTISVHATPQELSAAGCRKIVEIIEASGSQQWPLSIALAGGSTPKM TYARLHDEHLNLLREKRALRFFMGDERMVPADSTDSNYNMAREVLLHDIP DDLVFPFDTSAVTPSAEATSADAMRVAEAYGKQLASLLPLKSVGEAGPKV PVFDVVLLGLGSDGHTASIFPGSQAEKETDGKVVVSVGFPSETMKPKVWR VTLSPATIMQARNVIVLATGAEKKWVVDGILADTAHKAPVARFLRGCEGN VSFLLDKEIAENLAKFGGSLEHHHHHH
ID #29
MPYAAVNGTELHYRIDGERHGNAPWIVLSNSLGTDLSMWAPQVAALSKHF RVLRYDTRGHGHSEAPKGPYTIEQLTGDVLGLMDTLKIARANFCGLSMGG LTGVALAARHADRIERVALCNTAARIGSPEVWVPRAVKARTEGMHALADA VLPRWFTADYMEREPVVLAMIRDVFVHTDKEGYASNCEAIDAADLRPEAP GIKVPALVISGTHDLAATPAQGRELAQAIAGARYVELDASHISNIERADA FTKTVVDFLTEQKGGSLEHHHHHH
ID #30
MAQRVKITTTATPGEIELAFEDTGTGLPVLLVHGFPLDRTMWKAQREELC DEFRVIVPDLRGFGESQVIPGVATMEAMADDLAGLCNHLGLTGKIVLGGL SMGGYVAFAFARKYRDRLAGLILCDTRARPDSPEAKENRRRVAERVRREG PGFIAEEMIPRLCCESTFRNHPEVIEKIRQMILSAPPEGVAAAALGMAER

PDSTDLLPALSCPTLVLVGQFDAISPPEEMEAMARTIPQSQFVVIPDAGH LPPMEQPERVTQAIREWLRKVHTEAGGSLEHHHHHH
ID #31
MSERIVPSGDVELWSDDFGDPADPALLLVMGGNLSALGWPDEFARRLADG GLHVIRYDHRDTGRSTTRDFAAHPYGFGELAADAVAVLDGWGVDRAHVVG LSMGATITQVIALDHHDRLSSLTMLLGGGLDIDFDANIERVMRGEPTLDG LPGPQQPFLDALALMNQPAEGRAAEVAKRVSKWRILSGTGVPFDDAEYAR WEERAIDHAGGVLAEPYAHYSLTLPPPSRAAELREVTVPTLVIQAEHDPI APAPHGKHLAGLIPTARLAEIPGMGHALPSSVHGPLAEVILAHTRSAAGG SLEHHHHHH
ID #32
MTVKKLYFIPAGRCMLDHSSVNSALTPGKLLNLPVWCYLLETEEGPILVD TGMPESAVNNEGLFNGTFVEGQILPKMTEEDRIVNILKRVGYEPDDLLYI ISSHLHFDHAGGNGAFTNTPIIVQRTEYEAALHREEYMKECILPHLNYKI IEGDYEVVPGVQLLYTPGHSPGHQSLFIETEQSGSVLLTIDASYTKENFE DEVPFAGFDPELALSSIKRLKEVVKKEKPIIFFGHDIEQEKSCRVFPEYI GGSLEHHHHHH
ID #33
MLKTISGTLALSLIIAASVHQAQAATTYNAVVSKSSSDGKTFKTIADAIA SAPAGSTPFVILIKNGVYNERLTITRNNLHLKGESRNGAVIAAATAAGTL KSDGSKWGTAGSSTITISAKDFSAQSLTIRNDFDFPANQAKSDSDSSKIK DTQAVALYVTKSGDRAYFKDVSLVGYQDTLYVSGGRSFFSDCRISGTVDF IFGDGTALFNNCDLVSRYRADVKSGNVSGYLTAPSTNINQKYGLVITNSR VIRESDSVPAKSYGLGRPWHPTTTFSDGRYADPNAIGQTVFLNTSMDNHI YGWDKMSGKDKNGNTIWFNPEDSRFFEYKSYGAGATVSKDRRQLTDAQAA EYTQSKVLGDWTPTLPGGSLEHHHHHH
ID #34
MLDMPIDPVYYQLAEYFDSLPKFDQFSSAREYREAINRIYEERNRQLSQH

ERVERVEDRTIKGRNGDIRVRVYQQKPDSPVLVYYHGGGFVICSIESHDA LCRRIARLSNSTVVSVDYRLAPEHKFPAAVYDCYDATKWVAENAEELRID PSKIFVGGDSAGGNLAAAVSIMARDSGEDFIKHQILIYPVVNFVAPTPSL LEFGEGLWILDQKIMSWFSEQYFSREEDKFNPLASVIFADLENLPPALII TAEYDPLRDEGEVFGQMLRRAGVEASIVRYRGVLHGFINYYPVLKAARDA INQIAALLVFDGGSLEHHHHHH
ID #35
MTLDLQVQSFLAQGGNLNTLTGEEHGEAKAVFKVEDFYIPVKDGEIKLRV YTPNEKESLPVFVYLHGGGWVAGDIQAVDTLCQNIAHDAECIVVAVEYRL APTHKFPVPLEDCYEATKWVAENSSMLNADKTKIAIGGDSAGGNIAAAVV IMAQKFNNLSLVAQVLVYPVVDLTLTFKAQSYRDNAEGYLLTTEGVFWAT QMYLRDEIDRYNVFASPSVNNELENLPPALIITAEYDPLRDDGAAYAKRL EAAGIPVEYKCYEGMVHGFFWMAGIMDKGLQARLQVSNYLKSVFVGKGGS LEHHHHHH
ID #36
MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQQSLFPPVKKEPVAEVRE FDMDLPGRTLKVRMYRPEGVEPPYPALVYYHGGGWVVGDLETHDPVCRVL AKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARIA VGGDSAGGNLAAVTSILAKERGGPALAFQLLIYPSTGYDPAHPPASIEEN AEGYLLTGGMMLWFRDQYLNSLEELTHPWFSPVLYPDLSGLPPAYIATAQ YDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVR IAEKLRDALAGGSLEHHHHHH
ID #37
MPLDPEVRNFLQVYYKANIIDFTKYQFQEIRQKVNELLAKAVPKDPVGET RDMKIKLEDYELPIRIYSPIKRTNNGLVMHFHGGAWILGSIETEDAISRI LSNSCECTVISVDYRLAPEYKFPTAVYDCFNAIVWARDNAGELGIDKDKI ATFGISAGGNLVAATSLLARDNKLKLTAQVPVVPFVYLDLASKSMNRYRK GYFLDINLPVDYGVKMYIRDEKDLYNPLFSPLIAEDLSNLPQAIVVTAEY DPLRDQGEAYAYRLMESGVPTLSFRVNGNVHAFLGSPRTSRQVTVMIGAL

LKDIFKGSSLEHHHHHH
ID #38
MMPLDPRIKELLESGFIVPIGKASVDEVRKIFRQLASAAPKVEVGKVEDI KIPGSEANINARVYLPKANGPYGVLIYLHGGGFVIGDVESYDPLCRAITN ACNCVVVSVDYRLAPEYKFPSAVIDSFDATNWVYNNLDKFDGKMGVAIAG DSAGGNLAAVVALLSKGKLNLKYQILIYPAVGFDSVSRSMIEYSDGFFLT REHIEWFGSQYLRSPADLLDFRFSPILAQDLSGLPPALIITAEYDPLRDQ GEAYANRLLQAGVPVTSVRFNNVIHGFLSFFPLIEQGRDAISLIGSVLRR TFYDKSGGSLEHHHHHH
ID #39
MIDPKIKKLLESTIQLPIGKASVEEIRSLFKQFSSLTPREEVGKIEDITI PGSETNIKARVYYPKTQGPYGVLVYYHGGGFVLGDIESYDPLCRAITNSC QCVTISVDYRLAPENKFPAAVVDSFDALKWVYNNSEKFNGKYGIAVGGDS AGGNLAAVTAILSKKENIKLKYQVLIYPAVSFDLITKSLYDNGEGFFLTR EHIDWFGQQYLRSFADLLDFRFSPILADLNDLPPALIITAEHDPLRDQGE AYANKLLQSGVQVTSVRFNNVIHGFVSFFPFIEQGRDAIGLIGYVLRKVF YGKGGSLEHHHHHH
ID #40
MTEPLIIEPSQPADSAVIWLHGLGADRFDFEPVARLLGQHLPSTRFILPQ APTRPVTFNMGHAMPSWYDILALDGSERAINPADLEASSETLIALINAQQ QSGIDSKRIVLAGFSQGGAVVLHTALLRFDEKLAGVLALSTYAPTFNAET QFAESKQNLPVLCMHGSEDAVLPISMGRAVYDKLSEQGIKANWRDYPMGH EVRPEQLRDILDWLKNTLPSLPGGSLEHHHHHH
ID #41
MDKPIILDPKQSADSAVIWLHGLGATKEDFLPVAQILQRDALPHTRFILP QAPVRPVTLNNGFPMPSWYDIIALTSPREIKLSELDESSQSIIALIEAEI EKGIPLERIILAGFSQGGAVVLHTAFIAYPKNVGGVMALSTYSATFDEAI TLDEKKKQIPTLHLHGSLDPVVKIELGRAAEQFLKAQGIDTRWHDYPMQH

EVINDELQDIAKWLIERLGGGSLEHHHHHH
ID #42
MSDTLILEPTHRADACVIWLHGLGADRYDFLPVAEALQDVLGTTRFVLPQ APTRAVTINGGWAMPSWYDILAMSPERAIDEAQLEASAQQVMALAQAQVD GGIEPRRIFLAGFSQGGAVVLHTAFLRWEDELGGVLALSTYGPTFTDGMT LPDAKRQLPVLCLHGTLDDVVLPAMGRAAHDRLAAAGVPVGWRDYPMAHE VLPQQVRDIGAWLVERLHSGGSLEHHHHHH
ID #43
MSQPLLLEPTQPADSCVIWLHGLGADRYDFEPVARMLQKVLPRTRFILPQ APTRPVTVFNGMPAPSWYDIKAMAPARAIDEAQLDASADAVIALIEGQLA EGIAQRRIVLAGFSQGGAVVLHTGYLRWPGELGGVMALSTYGPTFDDDLQ LPEAKKQQPALCLHGTYDDVVAPAMGRAAYDFLQRQGVAVQWRDYPMAHE VSNQEIADIAAWLRERLGGSLEHHHHHH
ID #44
MSEPLILDAPNADACIIWLHGLGADRTDFKPVAEALQMVLPSTRFILPQA PSQAVTVNGGWVMPSWYDILAFSPARAIDEDQLNASADQVIALIDEQRAK GIAAERIILAGFSQGGAVVLHTAFRRYAQPLGGVLALSTYAPTFDDLALD ERHKRIPVLHLHGSQDDVVDPALGRAAHDALQAQGVEVGWHDYPMGHEVS LEEIHDIGAWLRKRLGGSLEHHHHHH
ID #45
MTDPLIIEPAQTADSCVIWLHGLGADRYDFQPVAEMLQQRLLHTRFVLPQ APTRAVTINGGWAMPSWYDIQAMSPARAIDQAQLEQSAQTVIELIEQQRD SGIDPRRIFLAGFSQGGAVVYHTAFLRWAGPLGGVLALSTYAPTFGDDLK LSPLQAGLPVLCLHGSRDDVVPPAMGRAAHDCLQQNQVQTQWKEYPMAHE VQPTEIQDIGDWLASRLGGGSLEHHHHHH
ID #46
MTEPLILQPAKPADACVIWLHGLGADRYDFMPVAEALQESLLTTRFVLPQ APTRPVTINGGYEMPSWYDIKAMSPARSISLEELEVSAKMVTDLIEAQKR

TGIDASRIFLAGFSQGGAVVFHTAFINWQGPLGGVIALSTYAPTFGDELE
LSASQQRIPALCLHGQYDDVVQNAMGRSAFEHLKSRGVTVTWQEYPMGHE
VLPQEIHDIGAWLAARLGGGSLEHHHHHH

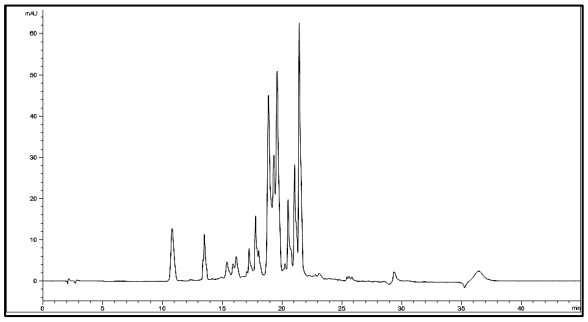
Supplementary Table 11.3. 1AUR mutants tested in this study. Examples of the methods used to generate *in silico* structural and enzyme data can be found at https://github.com/siegel-lab-ucd/blue-pigment-publication.git.

1AUR WT (Sequence ID No. 46) nucleotide sequence
ATGACCGAACCGCTGATTCTGCAACCGGCAAAACCGGCAGATGCATGTGT
TATTTGGCTGCATGGTCTGGGTGCAGATCGTTATGATTTTATGCCGGTGG
CCGAAGCACTGCAAGAAAGCCTGCTGACCACCCGTTTTGTTCTGCCGCAG
GCACCGACCCGTCCGGTTACAATTAATGGTGGTTATGAAATGCCGAGCTG
GTATGATATTAAAGCAATGAGTCCGGCACGTAGCATTAGCCTGGAAGAAC
TGGAAGTTAGCGCAAAAATGGTTACCGATCTGATTGAAGCACAGAAACGT
ACCGGTATTGATGCAAGCCGTATTTTTCTGGCAGGTTTTAGCCAGGGTGG
TGCAGTTGTTTTTCATACCGCCTTTATTAACTGGCAGGGTCCGCTGGGTG
GTGTTATTGCACTGAGCACCTATGCACCGACCTTTGGTGATGAACTGGAA
CTGAGCGCCAGCCAGCAGCGTATTCCGGCACTGTGTCTGCATGGCCAGTA
TGATGATGTTGTTCAGAATGCAATGGGTCGTAGCGCATTTGAACATCTGA
AAAGCCGTGGTGTTACCGTTACCTGGCAAGAATATCCGATGGGTCATGAA
GTGCTGCCGCAAGAAATTCATGATATTGGTGCATGGCTGGCAGCACGTCT
GGG
1AUR WT amino acid sequence
MTEPLILQPAKPADACVIWLHGLGADRYDFMPVAEALQESLLTTRFVLPQA
PTRPVTINGGYEMPSWYDIKAMSPARSISLEELEVSAKMVTDLIEAQKRTG
IDASRIFLAGFSQGGAVVFHTAFINWQGPLGGVIALSTYAPTFGDELELSA
SQQRIPALCLHGQYDDVVQNAMGRSAFEHLKSRGVTVTWQEYPMGHEVLPQ
EIHDIGAWLAARLGLEHHHHHH

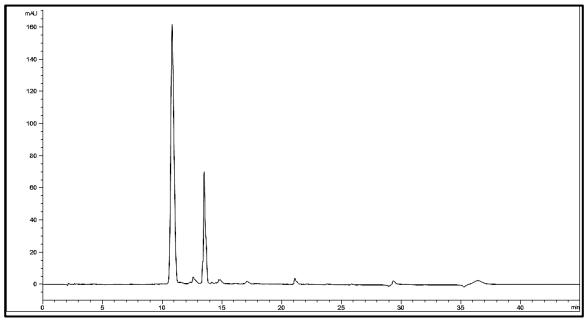
1AUR point mutations		
A25N	Y28T,M73V	A25G,I70L,M73V
A25D	A25G,M73V	I70L,M73V,E200S
A25G	A25G,Y28T	A25G,Y28T,I70L,M73V
D29G	M73V,F113A	E63D
I70L	M73V,F113Y	E200T,V201I
I70F	M73V,E200S	E200T
M73Q	I70L,M73A	F30H
M73V	I70L,M73L	L138V,S139N
M73L	I70L,M73V	M31K
M73H	I70L,M73F	M31R
F113A	I70L,M73S	P65SM73H
F113Y	I70T,M73F	M73H
V170N	I70T,M73A	A25G,I70L,M73V
V170T	I70T,M73L	I70L,M73V,E200S
E200S	I70T,M73V	Q166G,Y167F,V170I,V171A,Q172S, N173P,A174D,M175A,S178A
	I70Q,M73V	Q166G,Y167F,V170I,V171A,Q172S, N173P,A174D,M175A,S178A,E200T,V20I
	I70Q,M73L	W19C,Q166G,Y167F,V170I,V171A, Q172S,N173P,A174D,M175A,S178A, E200T,V201I
	I70Q,M73I	A25G,Y28T,D29S,F30H,M31R
	I70Q,M73F	A25G,Y28T,D29S,F30H,M31R, E200T, V201I
	I70Q,M73S	Q166G,Y167F,V170I,V171A,Q172S,N173P,A174D,M175A,S178A
	I70Q,M73A	Q166G,Y167F,V170I,V171A,Q172S,N173P,A174D,M175A,S178A,E200T,V201I

Supplementary Figure 11.1. HPLC Chromatograms of enzymes in Supplementary Table 11.1 mixed with RCAs . Multiple controls included to account for experimental variability. Peaks occurring after 30 minutes are an artifact of the enzyme reaction matrix and are not part of the RCA solution. See Supplementary Method 11.1 for experimental conditions

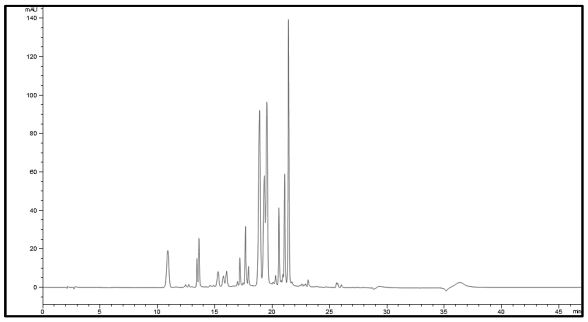
Control (no enzyme)



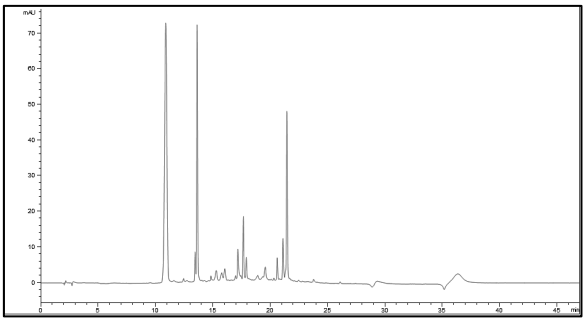
Sequence ID No. 1: Uniprot ID Q1QYJ5



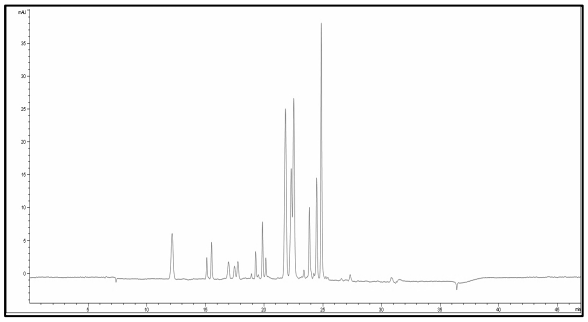
Control (no enzyme)



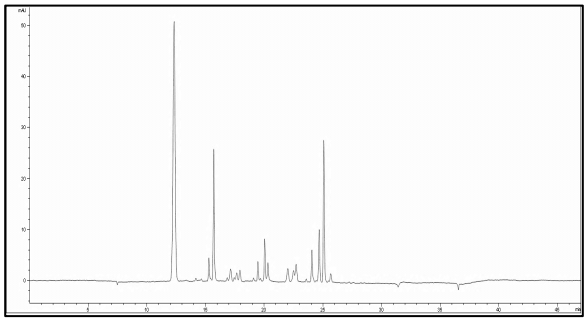
Sequence ID No. 2: Uniprot ID A0A1J8PUW1



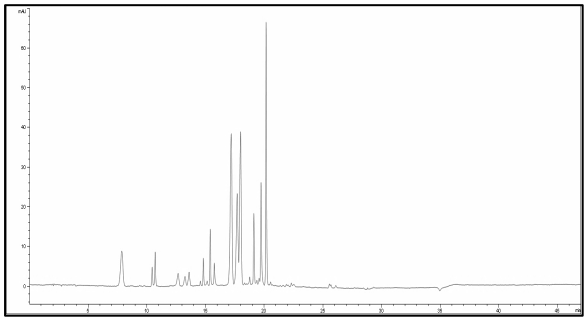
Control (no enzyme)



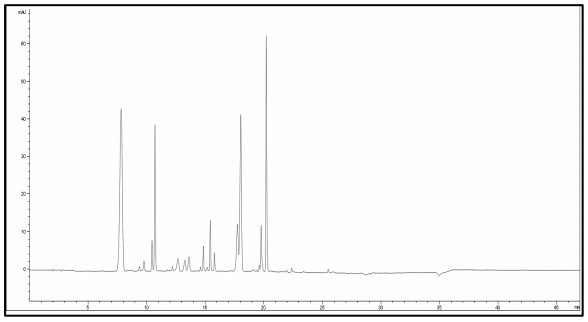
Sequence ID No. 3: PDB 4F21



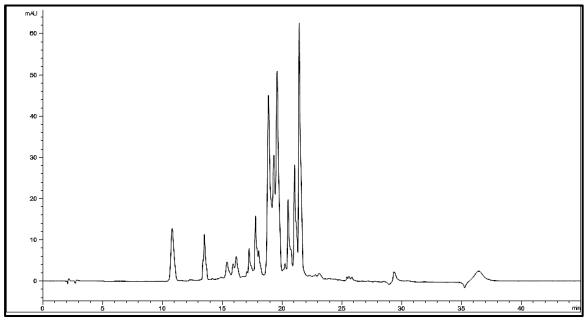
Control (no enzyme)



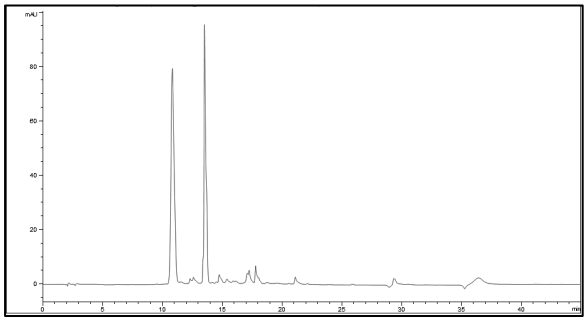
Sequence ID No. 4: PDB 5L2P



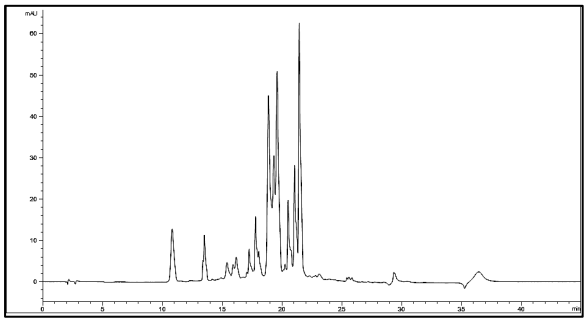
Control (no enzyme)



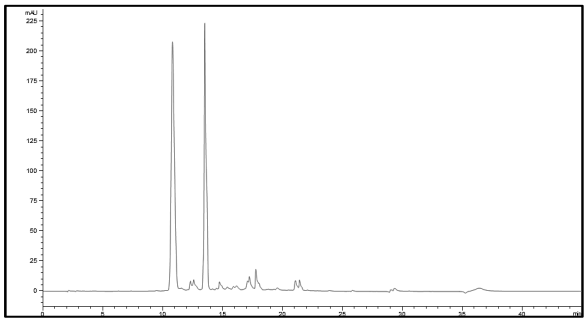
Sequence ID No. 6: PDB 1TQH



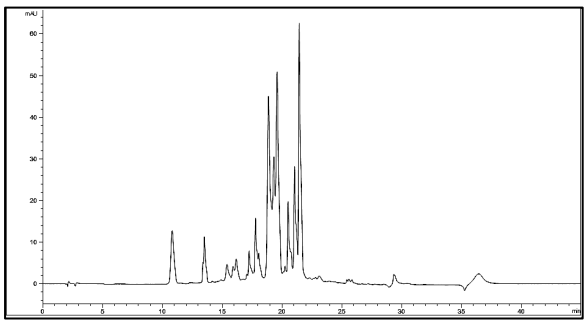
Control (no enzyme)



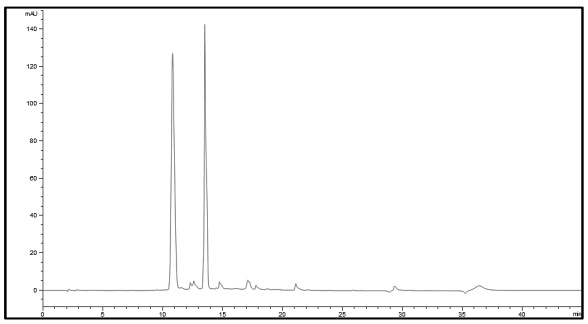
Sequence ID No. 7: PDB 4RGY



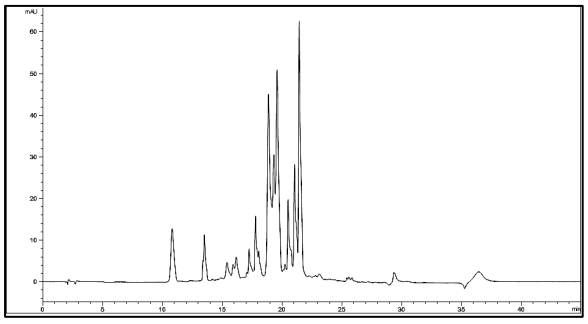
Control (no enzyme)



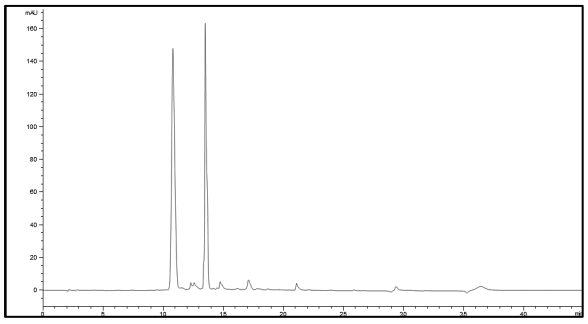
Sequence ID No. 8: Uniprot ID F8QQ74



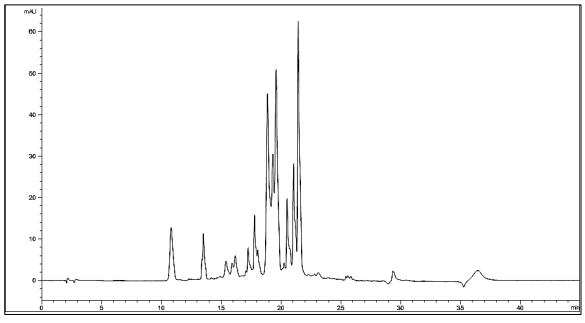
Control (no enzyme)



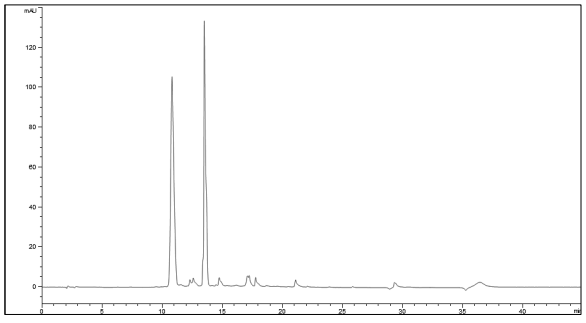
Sequence ID No. 9: Uniprot ID F8CAF0



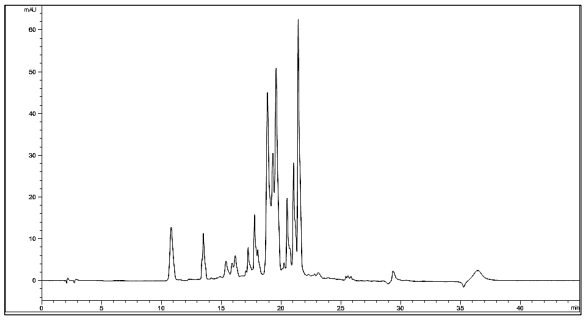
Control (no enzyme)



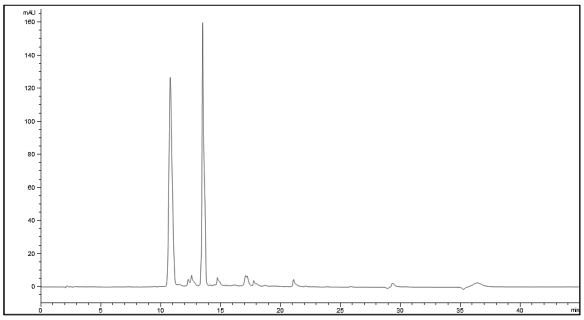
Sequence ID No. 10: PDB 3AJ3



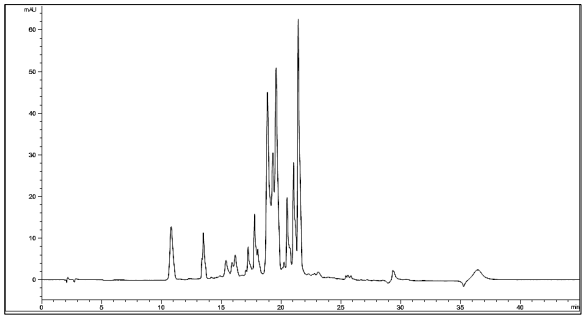
Control (no enzyme)



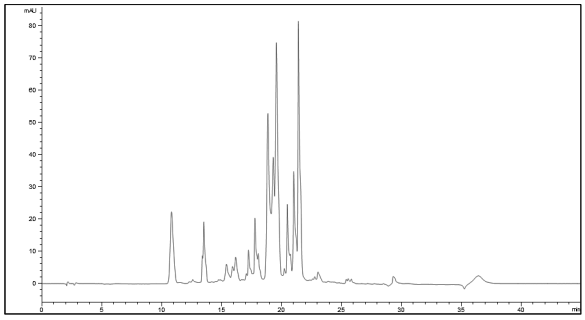
Sequence ID No. 11: Uniprot ID A6XIG7



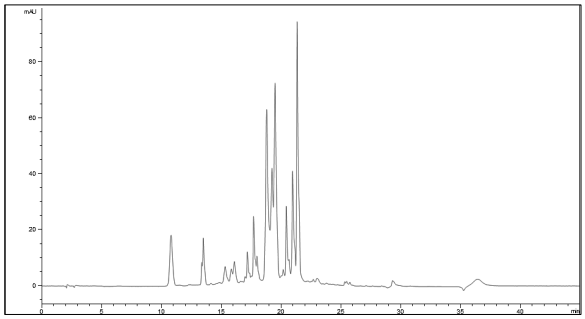
Control (no enzyme)



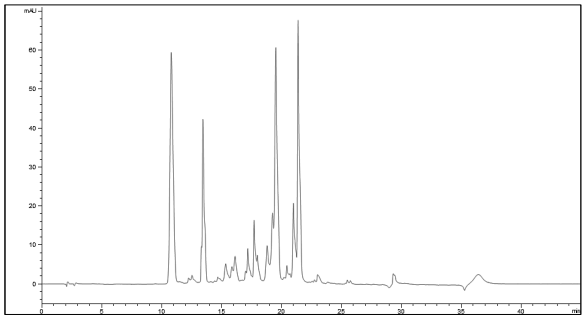
Sequence ID No. 12: Uniprot ID B2BSN6



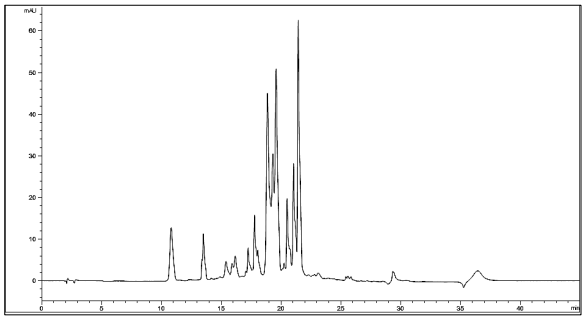
Control (no enzyme)



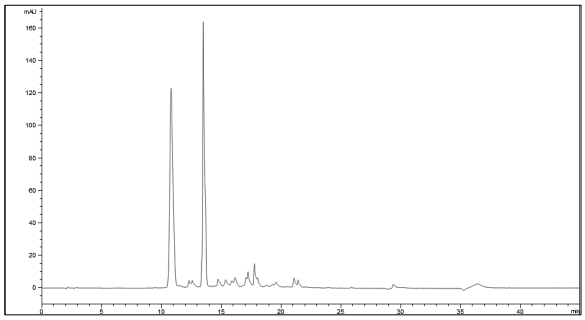
Sequence ID No. 13: PDB 2FFY



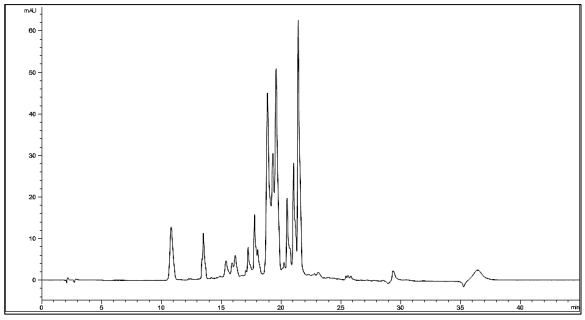
Control (no enzyme)



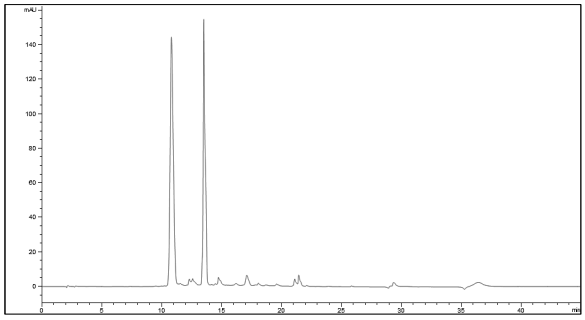
Sequence ID No. 14: Uniprot ID A3LZU8



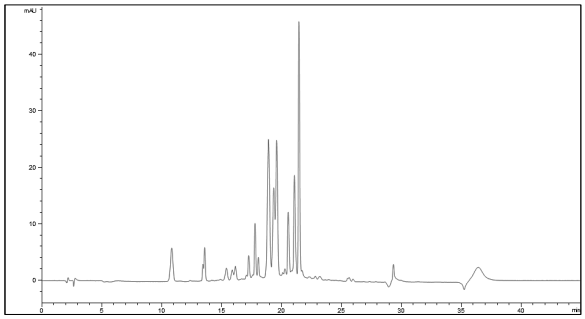
Control (no enzyme)



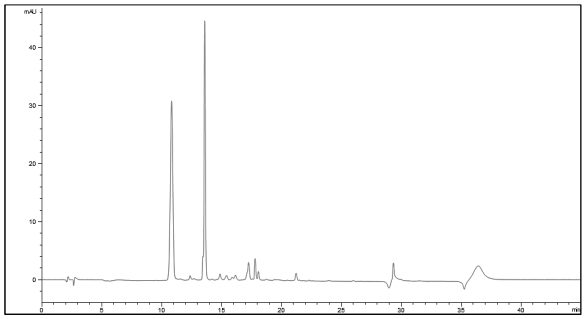
Sequence ID No. 15: PDB 2V3Z



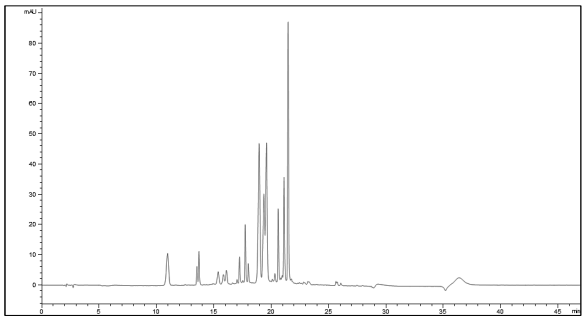
Control (no enzyme)



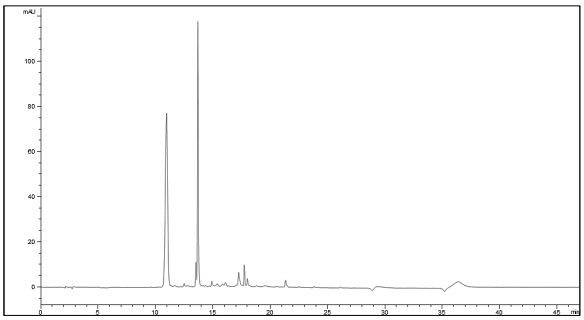
Sequence ID No. 16: Uniprot ID R0CVD2



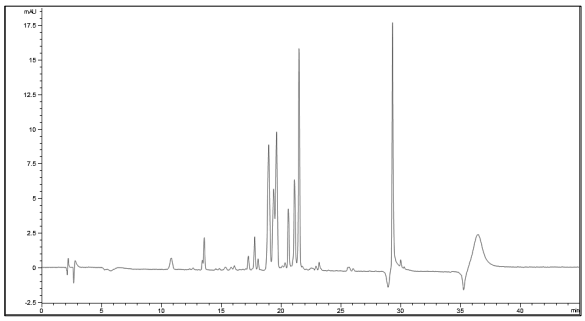
Control (no enzyme)



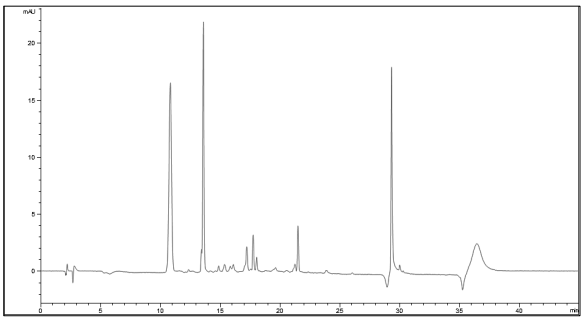
Sequence ID No. 17: PDB 5DWD



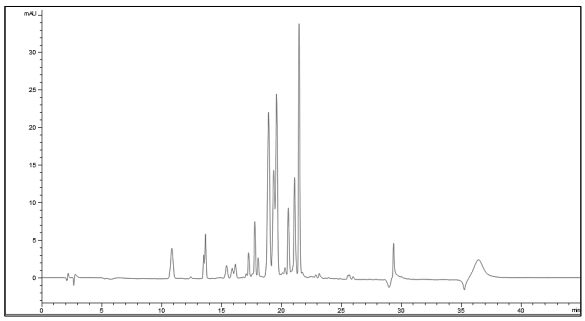
Control (no enzyme)



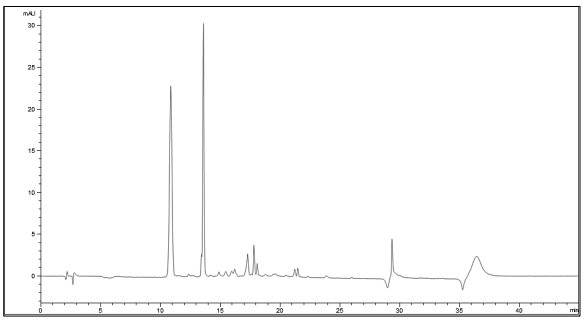
Sequence ID No. 18: Uniprot ID Q9EX73



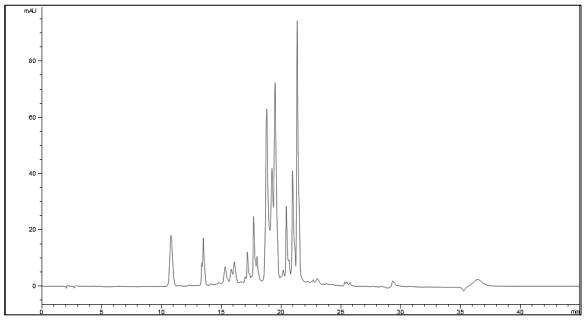
Control (no enzyme)



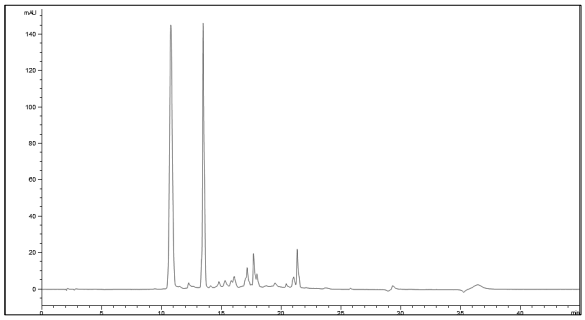
Sequence ID No. 19: Uniprot ID Q1JUP5



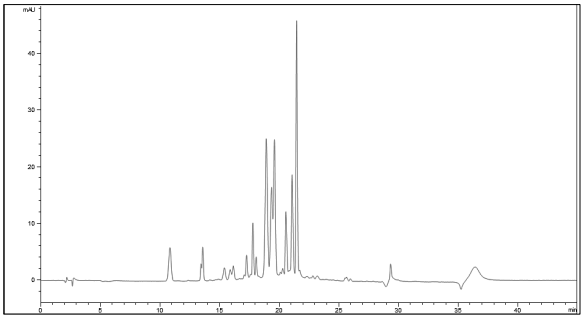
Control (no enzyme)



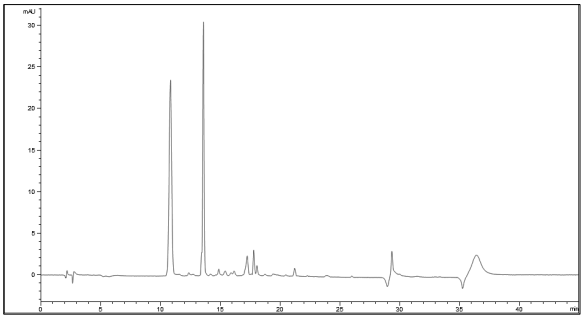
Sequence ID No. 20: PDB 1GKK



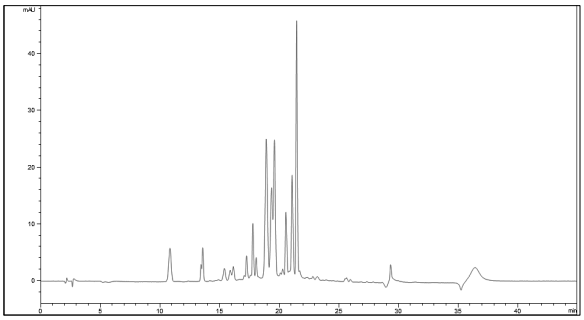
Control (no enzyme)



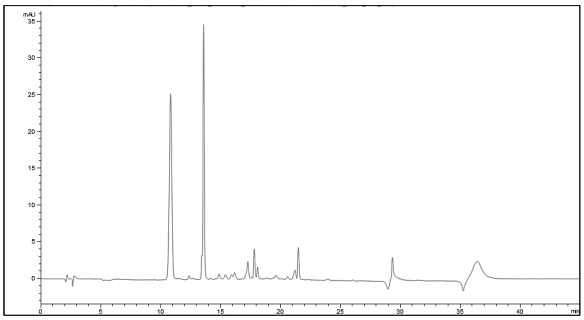
Sequence ID No. 21: PDB 3DR2



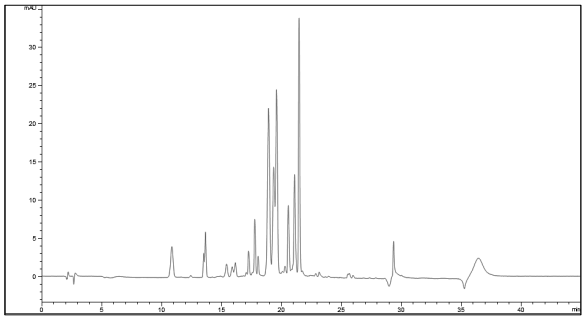
Control (no enzyme)



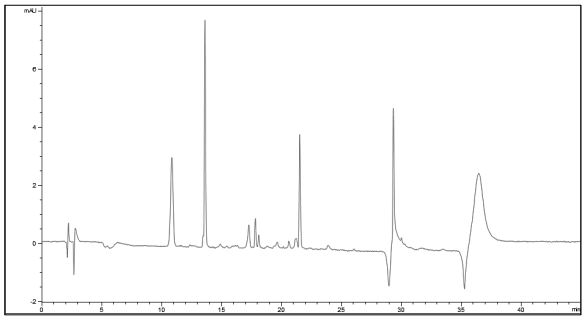
Sequence ID No. 22: PDB 1USW



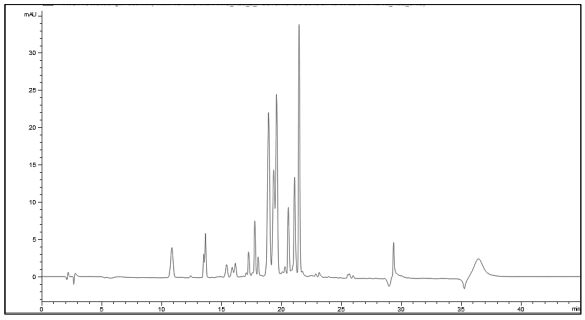
Control (no enzyme)



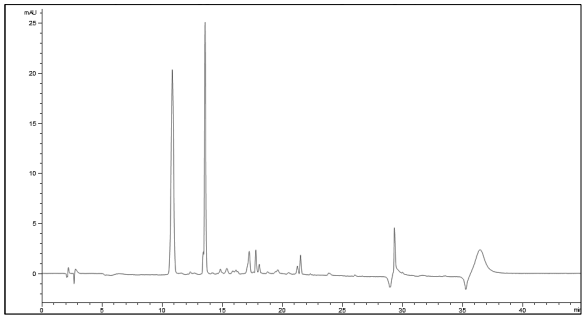
Sequence ID No. 23: Uniprot ID B2LYJ5



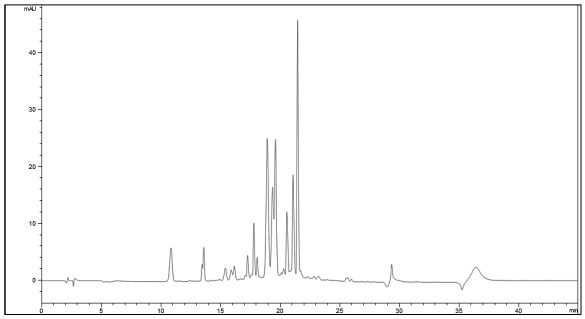
Control (no enzyme)



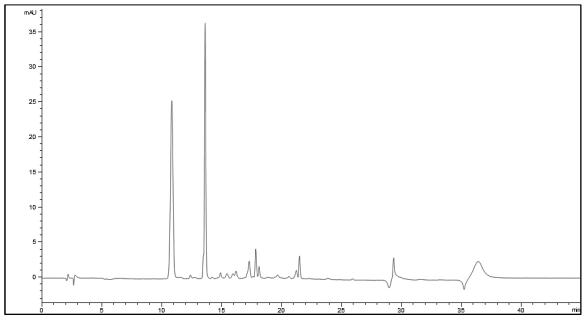
Sequence ID No. 24: Uniprot ID A9CPS8



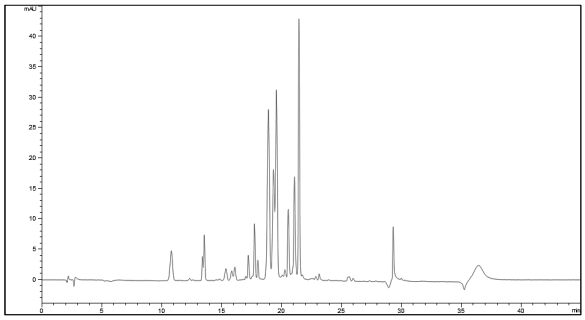
Control (no enzyme);



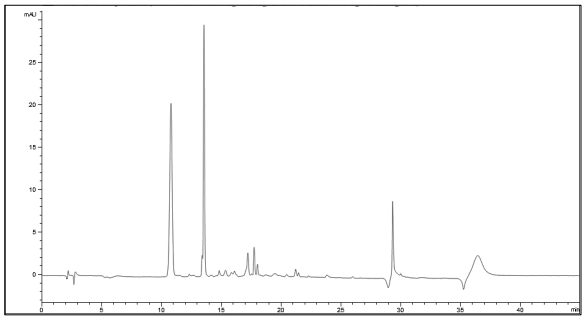
Sequence ID No. 25: Q8GCU7



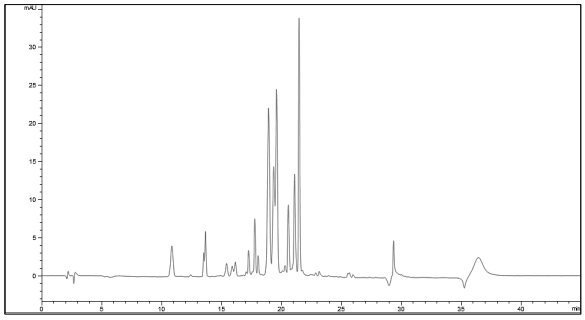
Control (no enzyme)



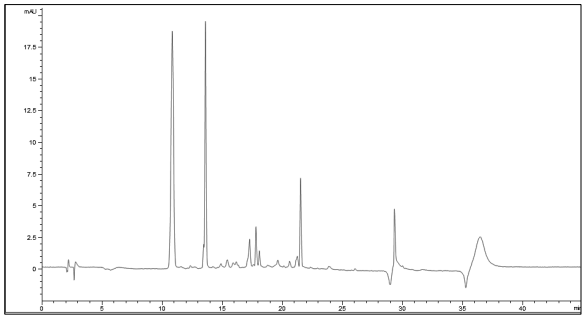
Sequence ID No. 26: PDB 2WHG



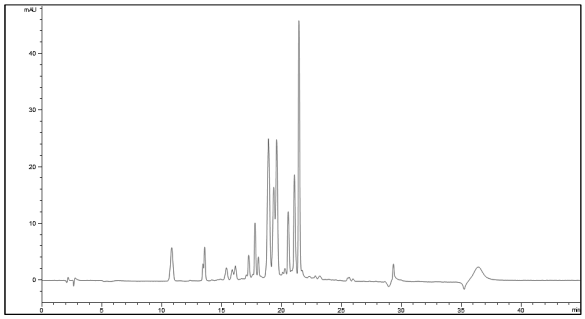
Control (no enzyme)



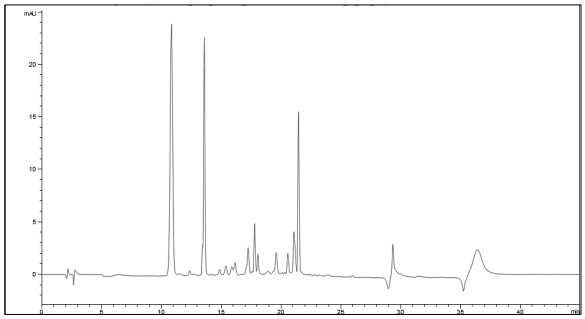
Sequence ID No. 27: PDB 1JU3



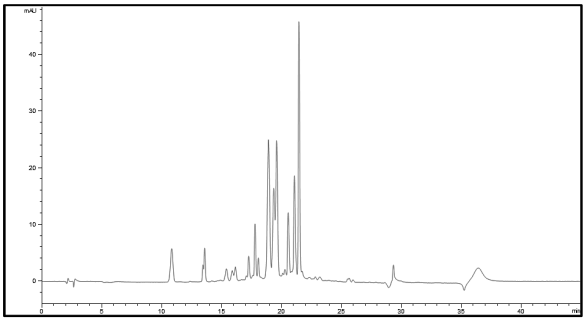
Control (no enzyme)



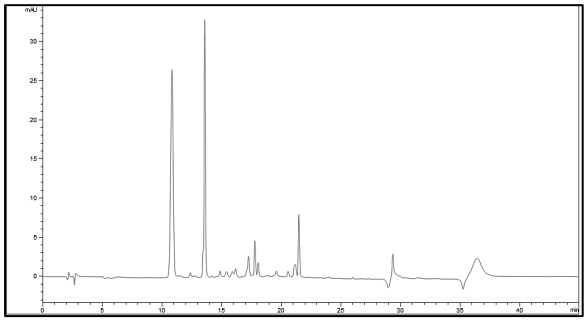
Sequence ID No. 28: PDB 3EB9



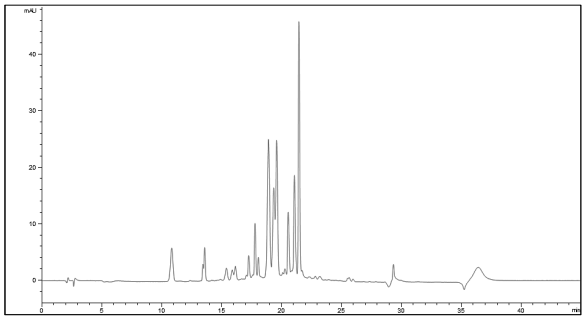
Control (no enzyme)



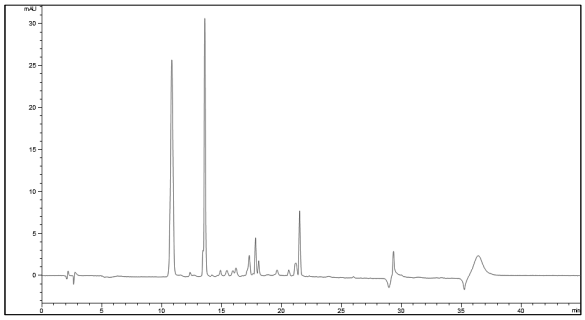
Sequence ID No. 29: PDB 2XUA



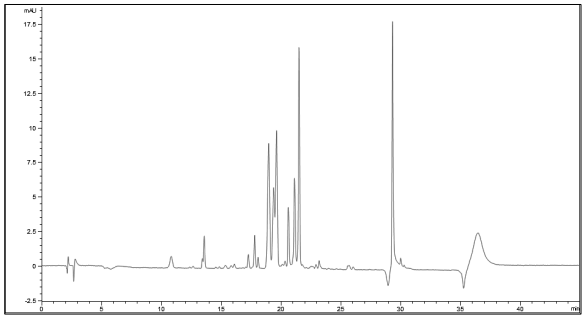
Control (no enzyme)



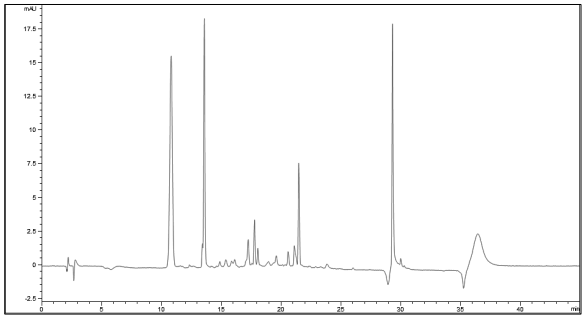
Sequence ID No. 30: PDB 4UHC



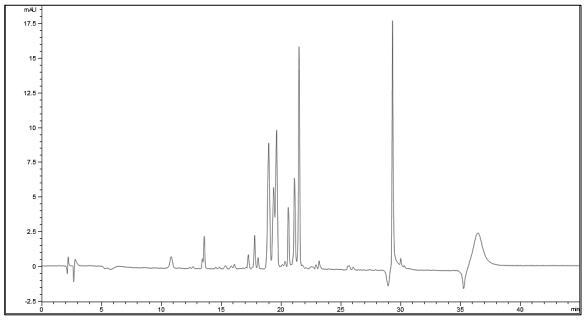
Control (no enzyme)



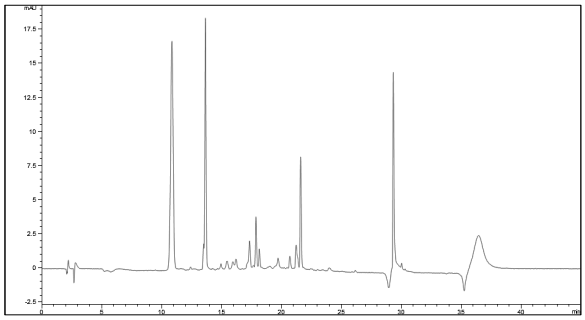
Sequence ID No. 31: PDB 1Q0R



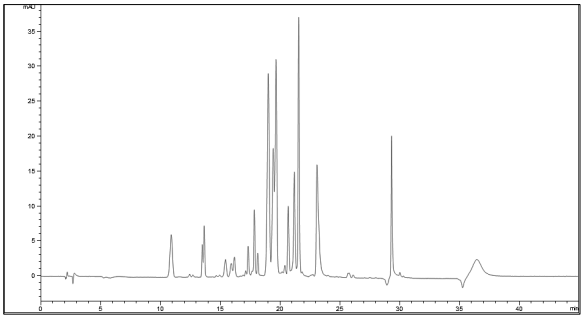
Control (no enzyme)



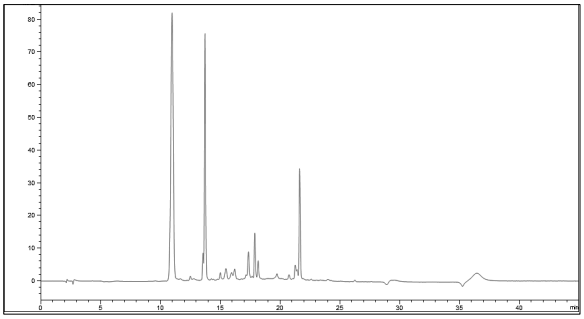
Sequence ID No. 32: PDB 3DHA



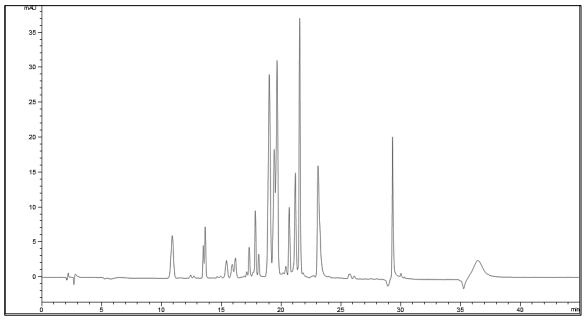
Control (no enzyme)



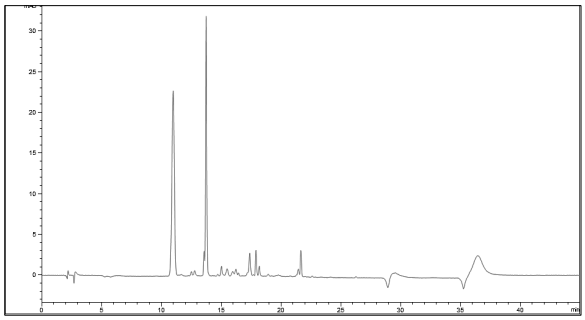
Sequence ID No. 33: PDB 2NSP



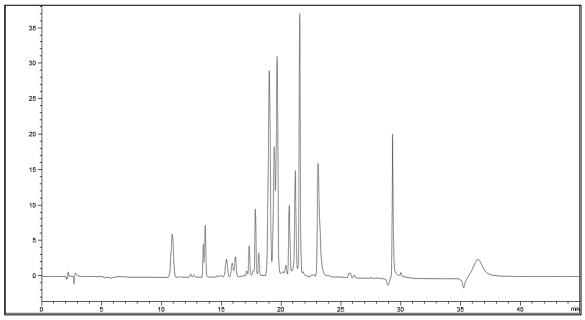
Control (no enzyme)



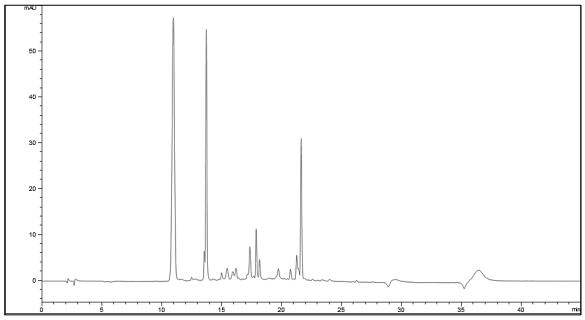
Sequence ID No. 34: PDB 1JJI



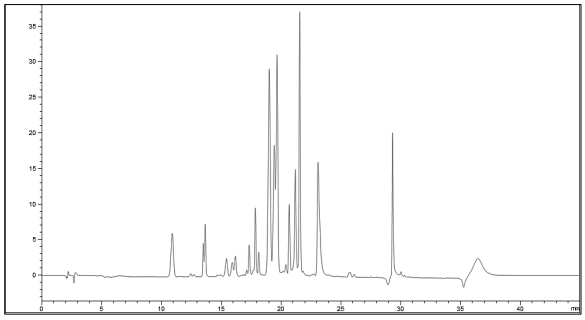
Control (no enzyme)



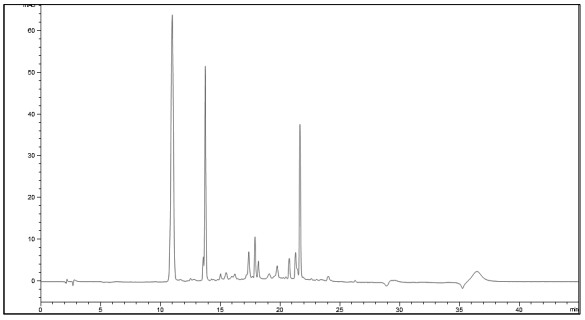
Sequence ID No. 35: Uniprot ID A0A243LS46



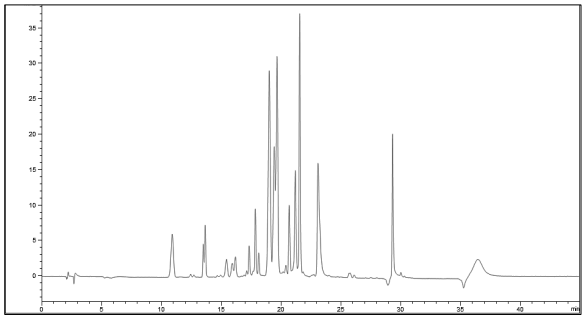
Control (no enzyme)



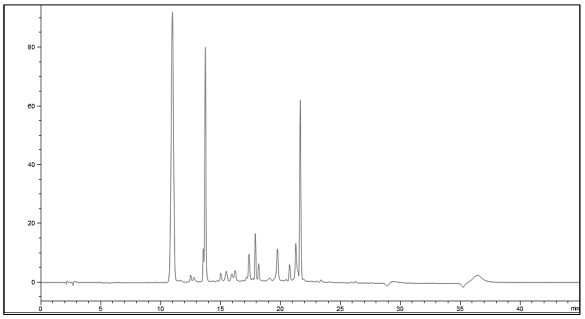
Sequence ID No. 36: PDB 1EVQ



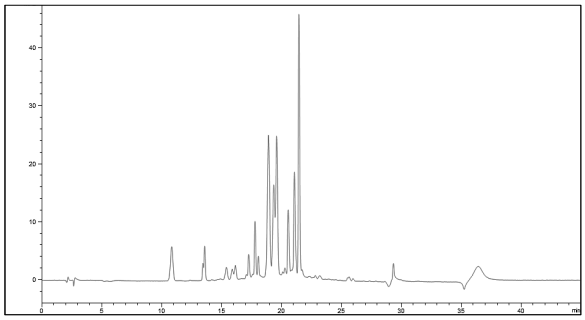
Control (no enzyme)



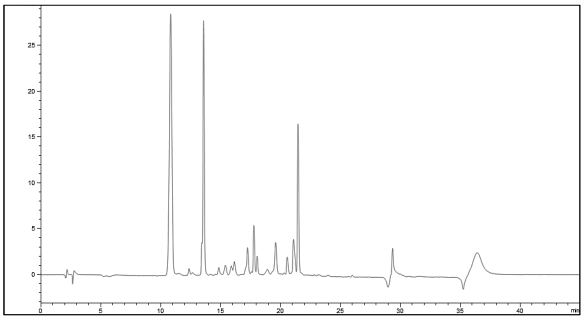
Sequence ID No. 37: PDB 5L2P



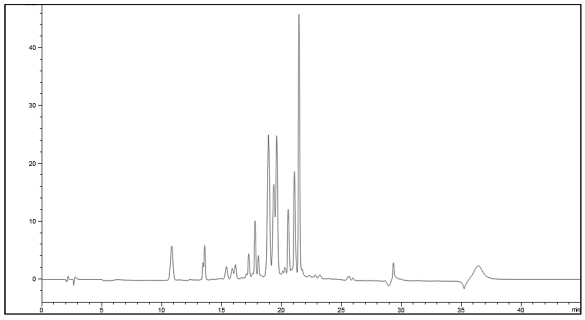
Control (no enzyme)



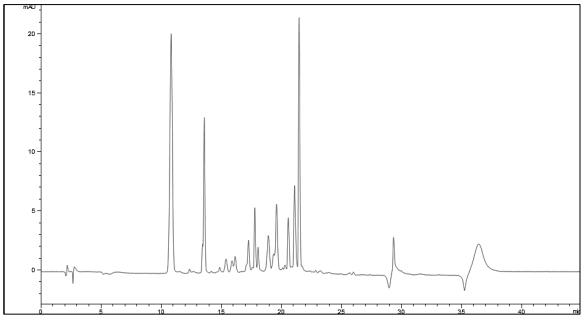
Sequence ID No. 38: PDB 5LK6



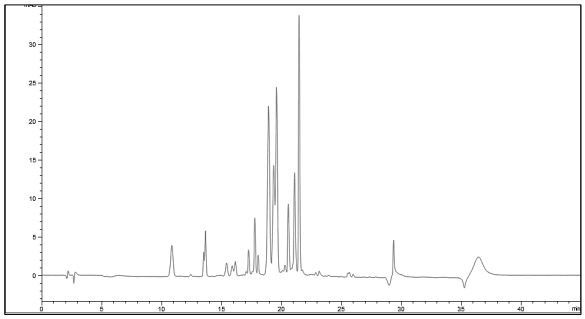
Control (no enzyme)



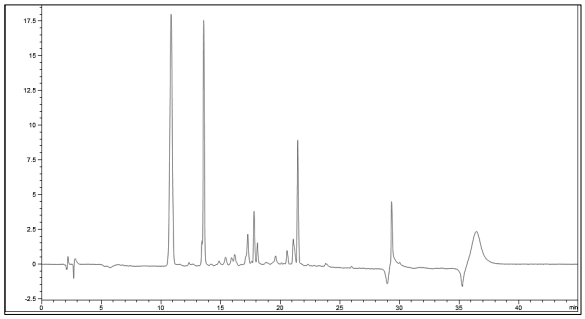
Sequence ID No. 39: PDB 3AIK



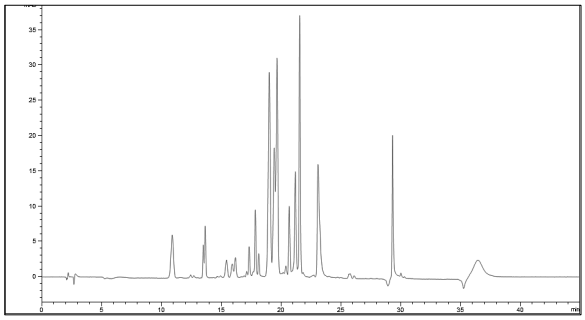
Control (no enzyme)



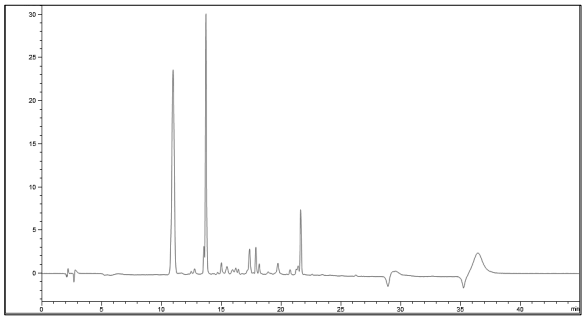
Sequence ID No. 40:Uniprot ID A0A139SSC5



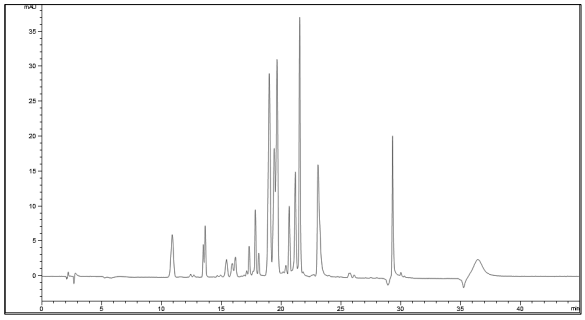
Control (no enzyme)



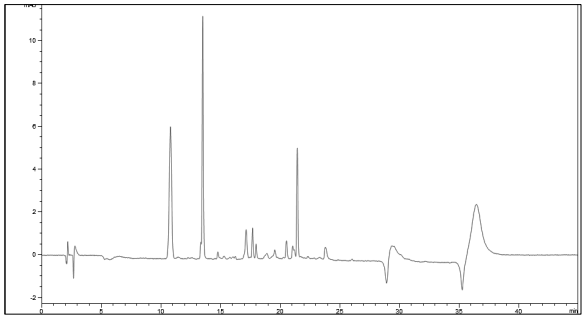
Sequence ID No. 41: Uniprot ID A0A2U2AP80



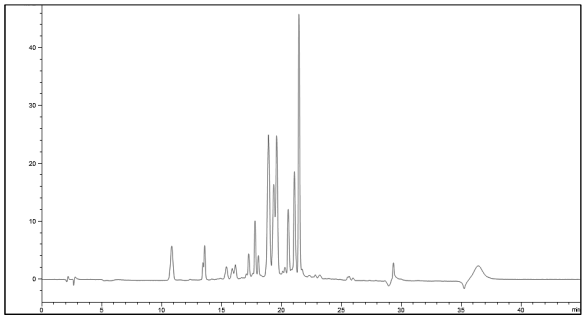
Control (no enzyme)



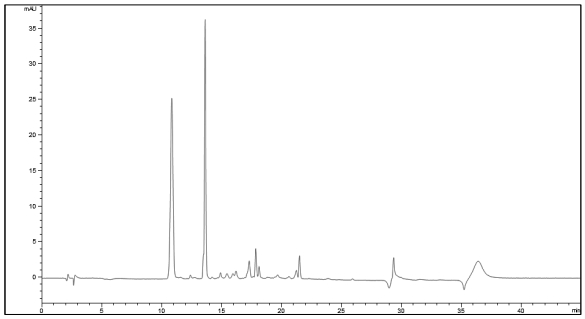
Sequence ID No. 42: Uniprot ID L8MEL5



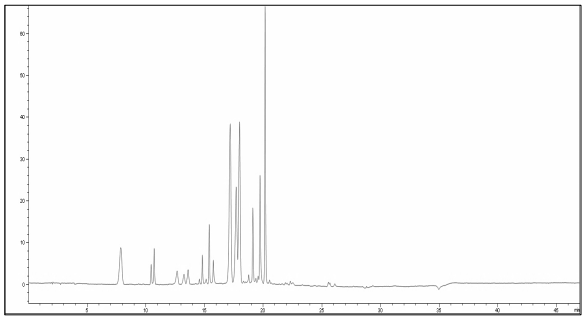
Control (no enzyme)



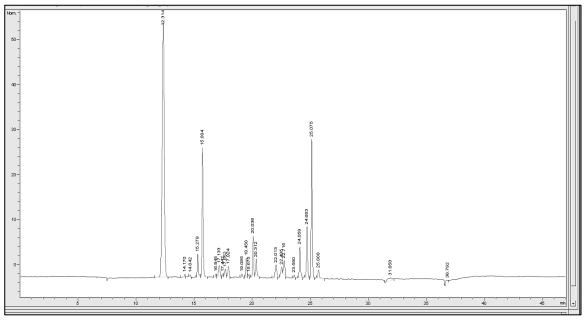
Sequence ID No. 43: Uniprot ID A0A127MYW2



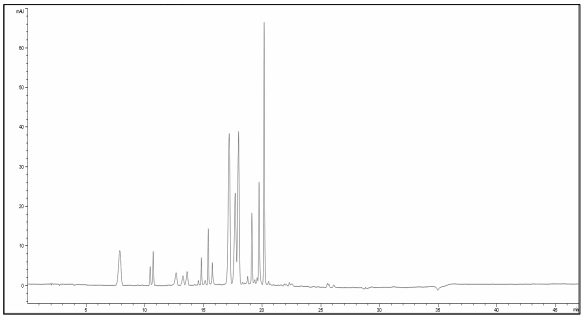
Control (no enzyme)



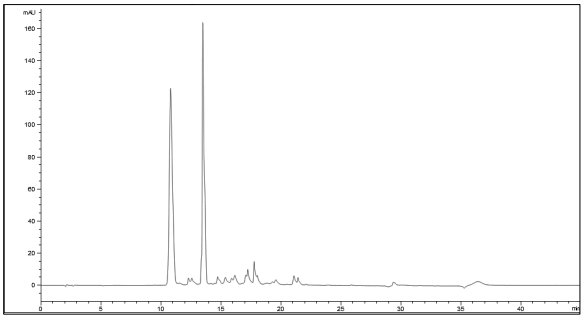
Sequence ID No. 44: PDB 3CN9



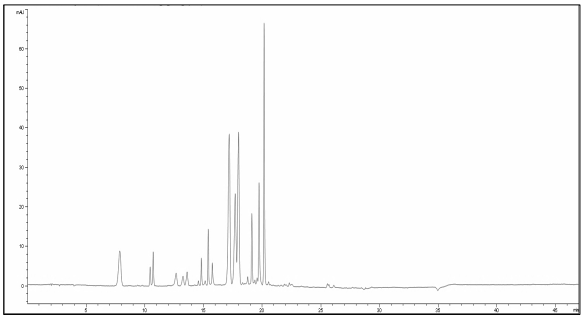
Control (no enzyme)



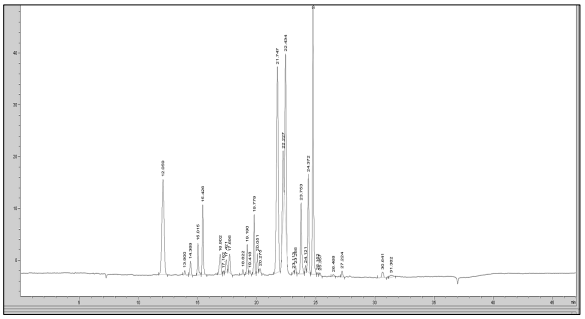
Sequence ID No. 45: A0A078LXQ1



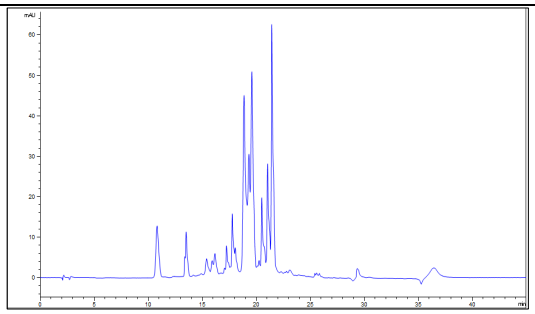
Control (no enzyme)



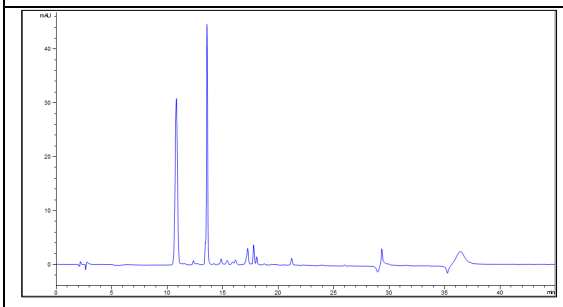
Sequence ID No. 46: PDB 1AUR



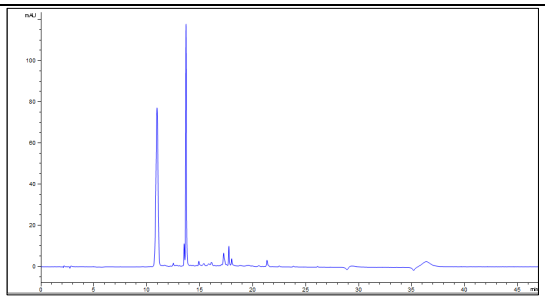
Supplementary Figure 11.2. HPLC Chromatograms of novel enzymes based on enzyme ID No. 46: PDB 1AUR (Supplementary Table 11.2). See Supplementary Method 11.1 for experimental conditions. Peaks occurring after 30 minutes are an artifact of the enzyme reaction matrix and are not part of the RCA solution. Reactions contained 1.0-1.5 mg/mL enzyme and 10 mg/mL RCE, and were run for 24 hours prior to quenching with 2% HCl



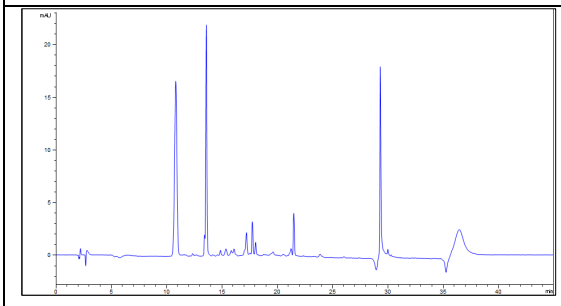
GFP (negative control)



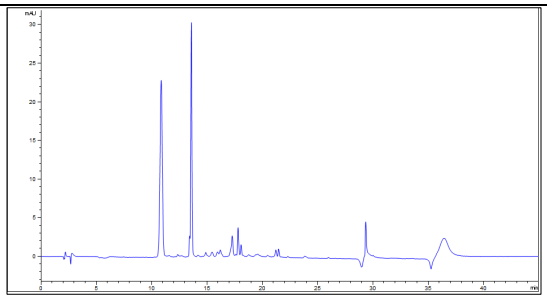
M73H



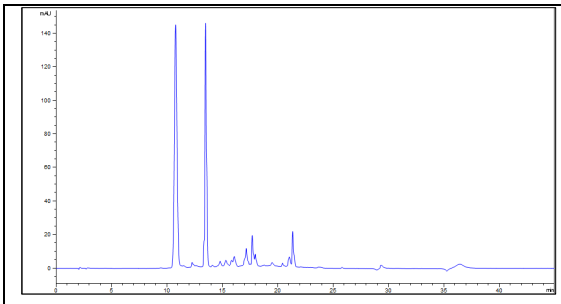
D29G



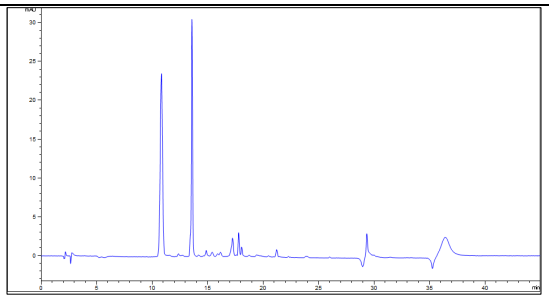
M73V



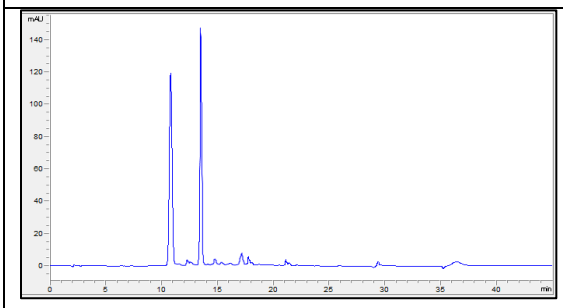
E200S



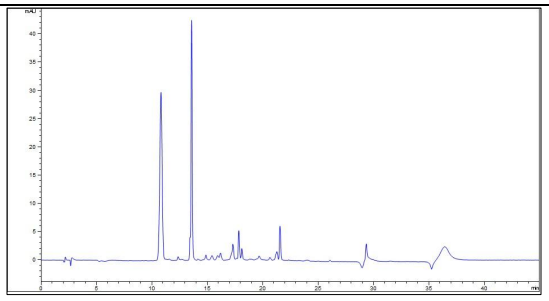
E200T



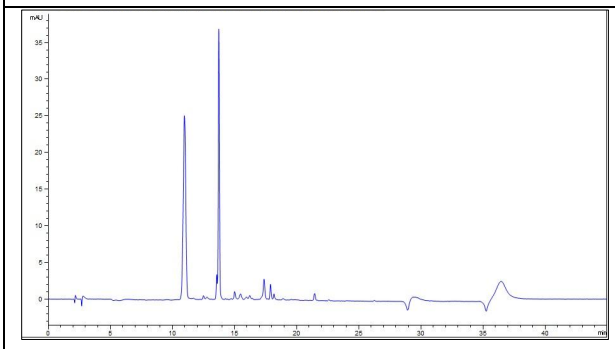
P65S;M73H



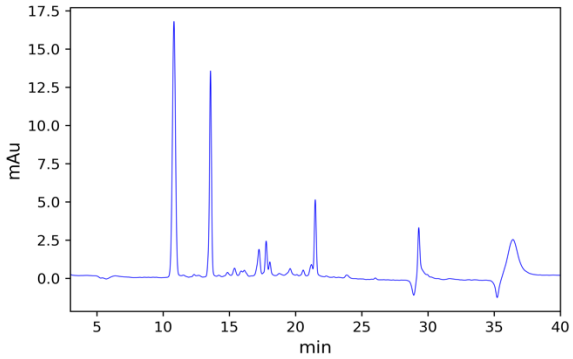
I70L;M73V;E200S



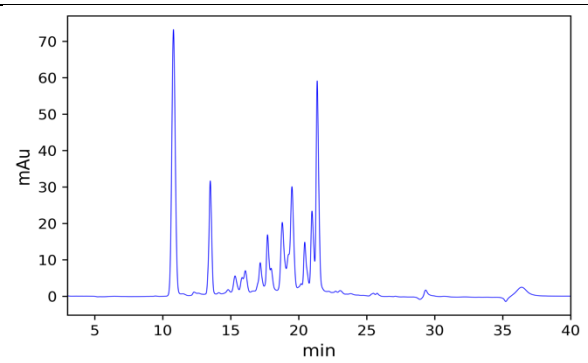
E200T;V201I



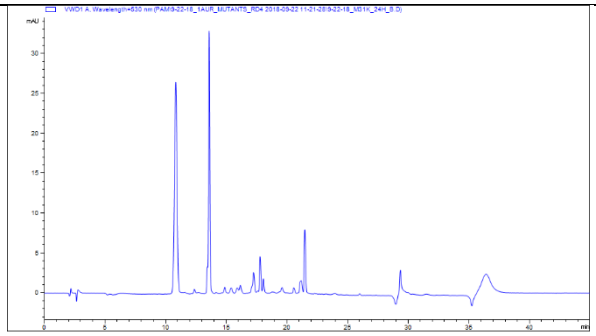
M73V;E200S



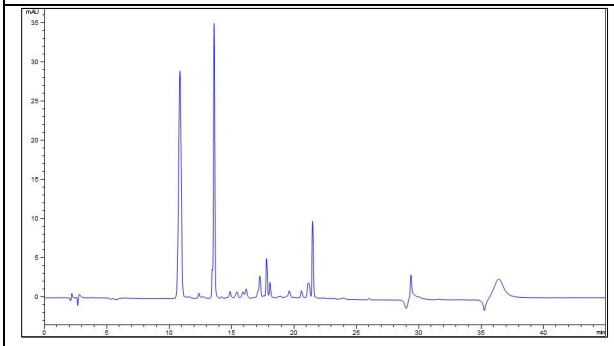
A25G;Y28T;I70L;M73V



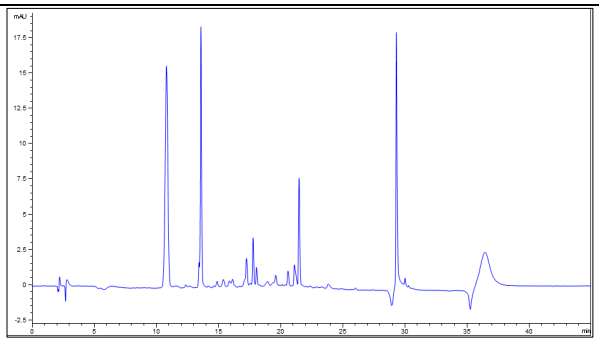
A25G;Y28T;D20S;F30H;M31R



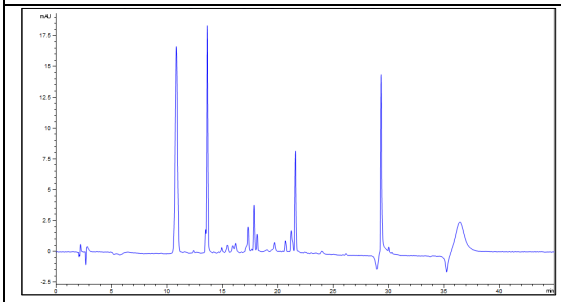
M31K



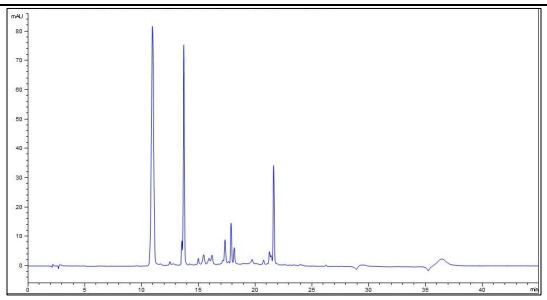
M31R



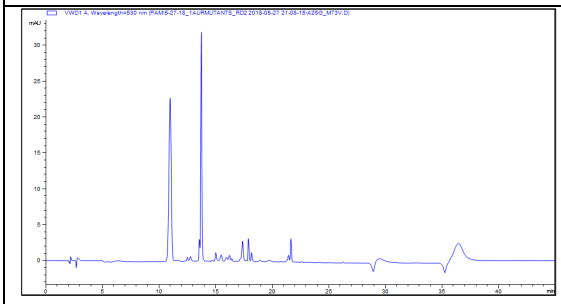
I70L



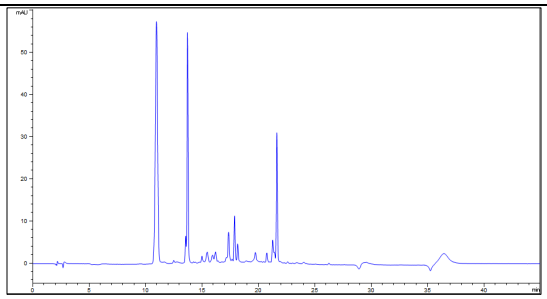
M73L



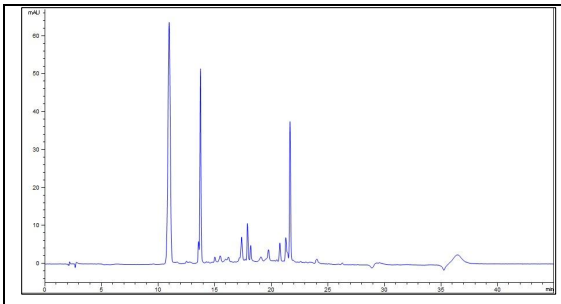
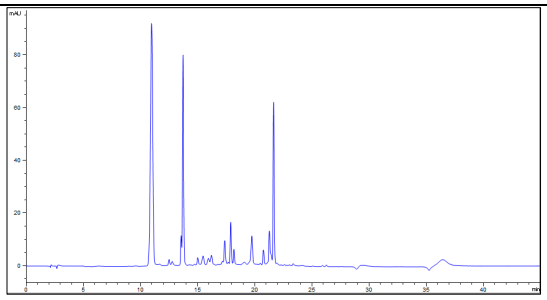
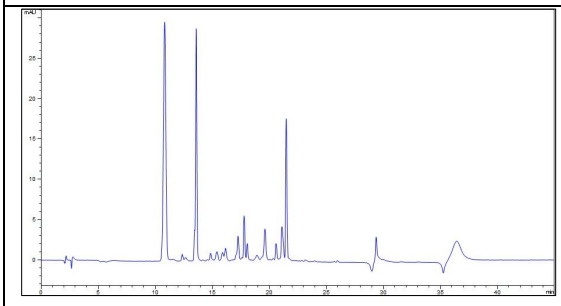
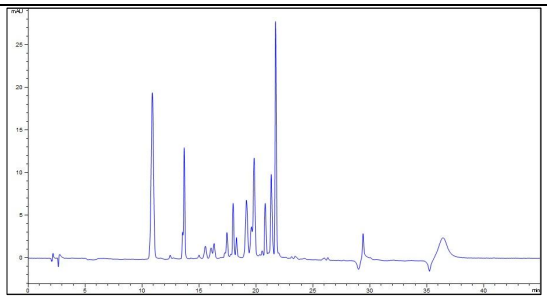
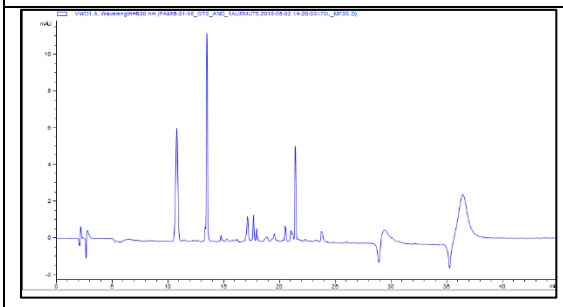
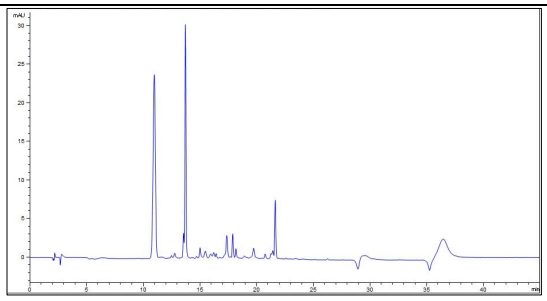
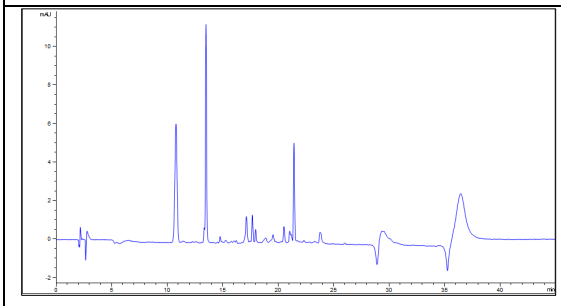
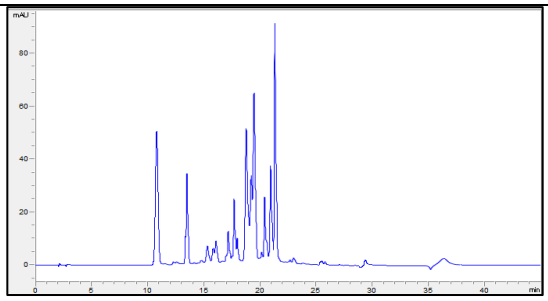
Y28T;M73V

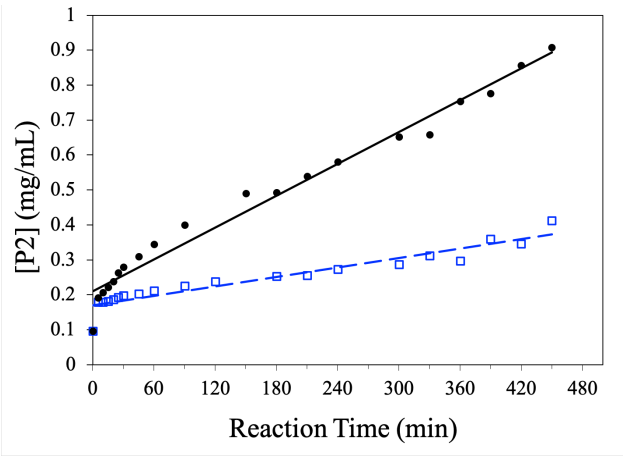


A25G;M73V



A25G;Y28T

	
I70L;M73A	I70L;M73L
	
F30H	L138V;S139N
	
A25G	I70L;M73V
	
I70L;M73S	A25G;Y28T;D20S;F30H;M31R;E200T;V201I



Supplementary Figure 11.3. Rate of formation of P2 via enzymatic treatment with enzyme ID No. 46: PDB 1AUR WT *(blue squares)* and 1AUR M73H *(black circles)*. Experimental set-up is described in Supplementary Method 11.1

Supplementary Section 12. Methodological details of the purification of anthocyanin P2 and the subsequent experimental details for the formation of the P2-Al complex ($Al^{3+}(P2^{-})_3$).

Supplementary Method 12.1. P2 Purification from Enzyme-treated Red Cabbage. The enzyme-treated red cabbage material contained a mixture of P2, phenolic acid and other anthocyanin impurities, buffer salts, and the enzyme used for production of P2. To achieve the cyan blue color expressed by $Al^{3+}(P2^{-})_3$ at pH 7, the material was purified using several techniques to remove impurities: enzyme precipitation, solid-phase extraction, and preparatory HPLC. The enzyme was precipitated by adjusting the pH of the solution to pH 1.1 – 1.5 with trace metal hydrochloric acid, and then left to sit in an ice bath for 2 hours. The slurry was transferred to centrifuge tubes followed by centrifugation at 3540 RCF for 5 minutes. The supernatant was decanted and subsequently vacuum filtered through a Buchner funnel and Whatman 1 filter paper. Buffer salts and other polar solutes were removed from the filtered impure P2 solution (enzyme now removed) by solid phase extraction. A Phenomenex Strata C18 SPE cartridge was

activated with two column volumes of 190 proof ethanol followed by two column volumes of acidified water (0.01% trace metal HCl). The material was then loaded onto the cartridge, washed with 2 column volumes of acidified water, then eluted the anthocyanins using acidified 190 proof ethanol (0.01% HCl). The solution was placed in a nitrogen evaporator with the water bath at 34℃ and a steady stream of nitrogen gas passing over to evaporate the ethanol. As the ethanol evaporated, an equal volume of Milli-Q 18.2 MΩ deionized water was added to the sample.

The P2 anthocyanins were purified using reverse phase preparatory HPLC-DAD. The HPLC-DAD parameters are as follows: 250 mm L x 50 mm D Phenomenex Luna C18 (2) preparatory LC column with 10 μm particle size and 100 Å pore size; mobile phases A: 5% glacial acetic acid in deionized water and B: 5% glacial acetic acid in 190 proof ethanol; flow rate of 60 mL/min; wavelength detection set to 280 nm and 520 nm. The gradient method used was: (1) B: 0%-12%, 0-30 min, (2) B: 12%-50%, 30-35 min, (3) B: 50% isocratic, 35-45 min. (4) B: 50%-0%, 45-46 min. Fractions were collected manually based on the elution of P2 observed using visible signal, 520 nm. In order to lyophilize the P2 material, the volatile ethanol was removed using rotary evaporation. The water bath was set to 35℃ and pulled under vacuum with a resulting pressure of 60 mbar. The P2 solution was transferred to a polypropylene container and immersed in liquid nitrogen until fully frozen then placed into the lyophilizer with vacuum set to 0.03 mbar and the condenser to -80 ℃.

Supplementary Method 12.2 Anthocyanin – Metal Complex ($Al^{3+}(P2^{-})_3$) Formation Details. 18.2 MΩ Milli-Q deionized water was added to lyophilized P2 and gently swirled to dissolve. Aluminum (III) salt stock solution was added to the P2 solution (see example calculation below), adjusted to pH 6-7 with 2 M trace metal NaOH, and let stand for 2 hours protected from light. Before lyophilization, the pH was

adjusted to 7.0 with NaOH. The P2 solution was transferred to a polypropylene container and immersed in liquid nitrogen until fully frozen then placed into the lyophilizer with vacuum set to 0.03 mbar and the condenser to -80 °C.

$$mass\ (g)\ of\ AlK(SO_4)_2\ for \frac{1}{3} mol\ eq\ to\ P2 = \left(\frac{mass\ of\ P2\ powder\ g}{979\ g\ P2\ mol^{-1}}\right)\left(\frac{1\ mol\ AlK(SO_4)_2}{3\ mol\ P2}\right)\left(\frac{258.21\ g\ AlK(SO_4)_2}{1\ mol}\right)$$

For use with safflower in a green powdered blend, safflower was dissolved in 18.2 MΩ Milli-Q deionized water, adjusted to pH 6-7 with trace metal sodium hydroxide, then added powdered $Al^{3+}(P2^{-})_3$. After dissolution, the pH was adjusted to 7.0 with sodium hydroxide. The green solution was transferred to a polypropylene container and immersed in liquid nitrogen until fully frozen then placed into the lyophilizer with vacuum set to 0.03 mbar and the condenser to -80 °C.

For comparison purposes a table of aluminum content in common food sources is provided (Supplementary Table 12.1). The aluminum content in a finished lentil is calculated using colorant use rate at 0.7 ppm.

Supplementary Table 12.1 Aluminum Content in Common Food Sources

Product	Al Content (ppm)
Flour *(67)*	4
Baking Premix *(67)*	51
Pasta *(67)*	10
Herb Tea *(67)*	40
Fruit juice and fruit juice drinks *(67)*	3
Baking powder *(68)*	18,000 - 28,000
Non-dairy creamer *(68)*	110 - 590

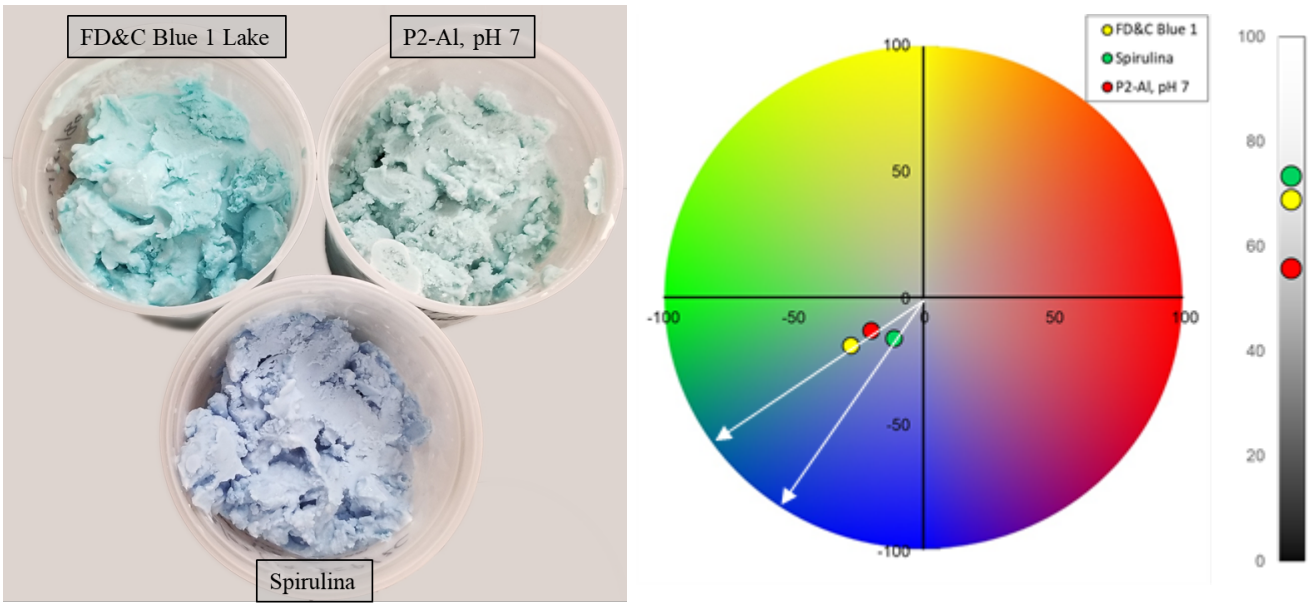
Lentil with P2-Al complex[†]	0.7

[†] Content in colorant only

Discussion 12.1. Safety considerations of anthocyanins as a colorant for human food products

Anthocyanins from red cabbage are recognized by the FDA and the EU as an approved source of color in food processing (*36*). Additionally, aluminum sulfate dodecahydrate, which was used in this study, is approved for use in coated confections, such as the candies and donuts sampled (Fig. 5B, C, Supplementary Fig. 13.2) *(67, 68)*. As such, we anticipate that this colorant will be classified as a low-level concern, and that general toxicity tests in bacteria and short-term toxicity tests in rodents should confirm the safety of the aluminum oxide-anthocyanin complex for use in food products (*37*). The enzymatic processing used to isolate P2 is analogous to approved protocols used in beer, bread, and cheese production (*38).* Furthermore, before human consumption it will be critical for the appropriate toxicology studies to be conducted.

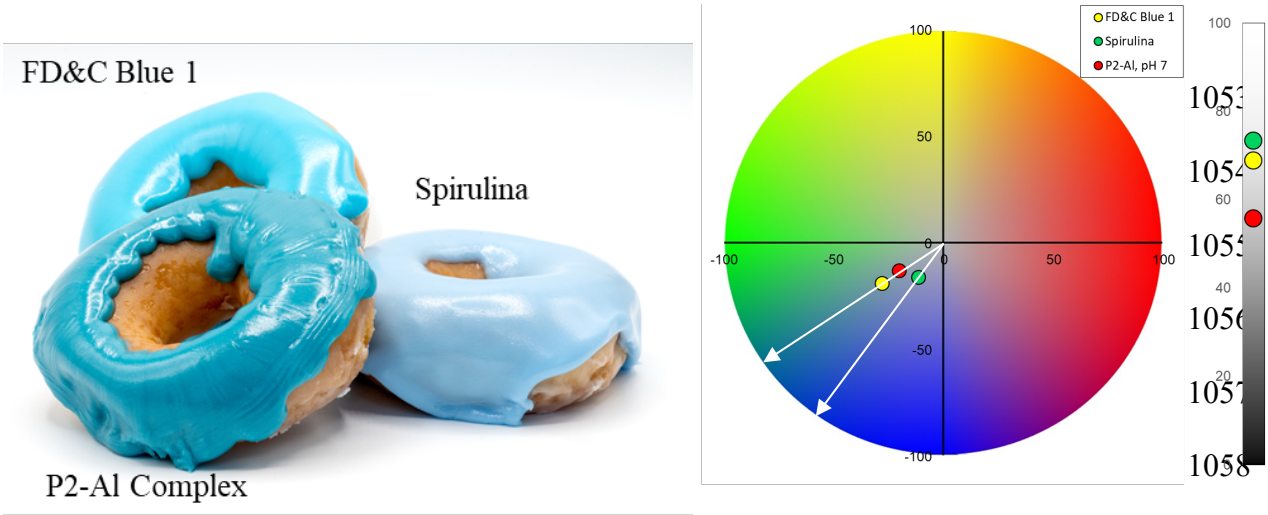
Supplementary Section 13. Prototypes, Application Examples of $Al^{3+}(P2^{-})_3$, and Stability Data.



Supplementary Figure 13.1. **Application of blue colorants in ice cream.** In a metal mixing bowl, heavy whipping cream, milk, and sugar were combined. The ingredients were mixed with a wooden spoon until the sugar dissolved, then added in the powdered colorant continued to mix until the colorant dissolved or dispersed completely. Liquid nitrogen was poured in slowly while continuously mixing until the blend became solid and firm. Samples were taken and measured for color expression and plotted (see Figure 13.1) Values also provided in Table 13.1 below. Hue angles for FD&C Blue No. 1 and ($Al^{3+}(P2^{-})_3$), pH 7 are a near match: 205.8 and 203.1 respectively. The ice cream was thawed until it could be pressed into a 30 mm diameter x 1 cm height optical glass petri dish with no air pockets. Before measuring, the petri dish was checked for condensation. The L*a*b* values were measured using a Konica Minolta CM-5 colorimeter with 30 mm mask, 10-degree observer, D65 lighting, and Spectral Component Excluded (SCE). Measured values were plotted on a*b* space. Hue angles are shown by arrows and L* values in sidebar. Photo Credit: Randall Powers, Mars Incorporated

Supplementary Table 13.1. Colorimetry Data Collected for Three Blue Ice Cream Prototypes

Name	L*	a*	b*	C*	h°
FD&C Blue No. 1	77.50	-18.15	-8.76	20.15	205.77
Spirulina	76.87	-10.58	-12.24	16.18	229.13
$Al^{3+}(P2^{-})_3$, pH 7	78.30	-11.23	-4.79	12.21	203.12



1059

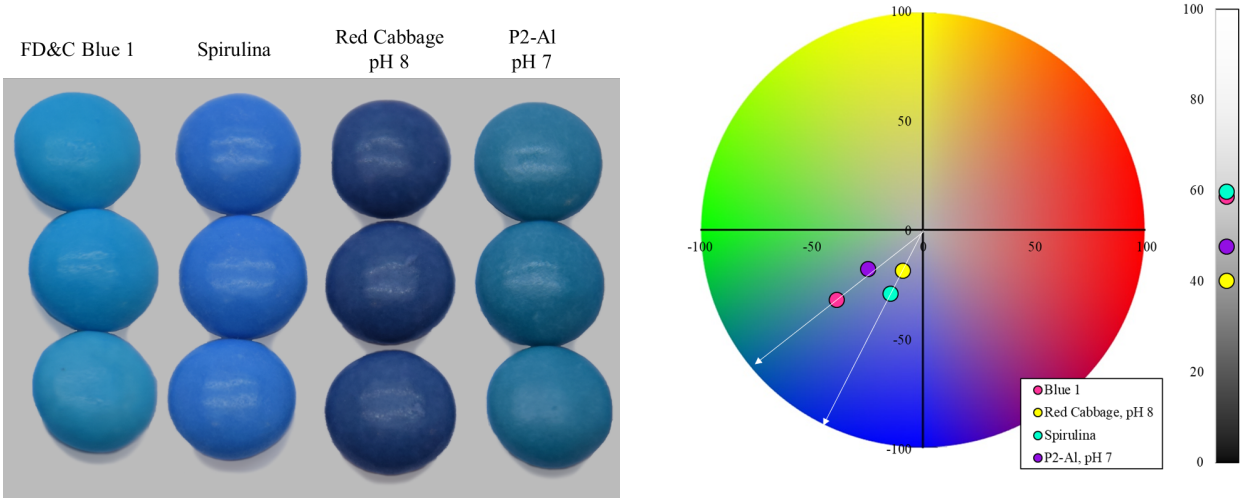
Supplementary Figure 13.2. Doughnut icing application of three blue prototypes. The icing was prepared with a typical recipe using water, powdered sugar, light corn syrup, the addition of calcium carbonate to increase opacity, and colorant. Measured values were plotted on a*b* space. Hue angles are shown by arrows and L* values in sidebar. Data was measured using a VeriVide DigiEye color imaging system. The image was captured with a calibrated Nikon D7000 DSLR camera and diffuse D65 lighting. L*a*b* measurements were taken at several areas on the surface of the icing and averaged. The data is summarized in Supp. Table 13.2. Photo Credit: Randall Powers, Mars, Incorporated.

Supplementary Table 13.2. Colorimetric data collected for blue icing prototypes

Name	L*	a*	b*	C*	h°
FD&C Blue No. 1	68.8	-27.9	-18.9	33.7	214.1
Spirulina	73.3	-11.6	-16.0	19.7	234.0
$Al^{3+}(P2^{-})_3$, pH 7	55.7	-20.2	-12.7	23.8	212.3

Supplementary Discussion 13.1. P2-Al complex ($Al^{3+}(P2^{-})_3$) stability in saturated sugar syrup.

Visible absorption spectra of $Al^{3+}(P2^{-})_3$ at pH 7 (solid lines) and red cabbage at pH 8 (dashed lines) in saturated sugar syrup monitored over time is shown in Figure 4A. Over the course of ten days, the absorbance at λ_{max} of red cabbage anthocyanins drops by 57% indicating loss of cyanidin chromophore. Whereas, $Al^{3+}(P2^{-})_3$ shows remarkable stability with only 14% loss of absorbance at λ_{max} over the course of 55 days. With this remarkable stability, several applications for the colorant were attempted.

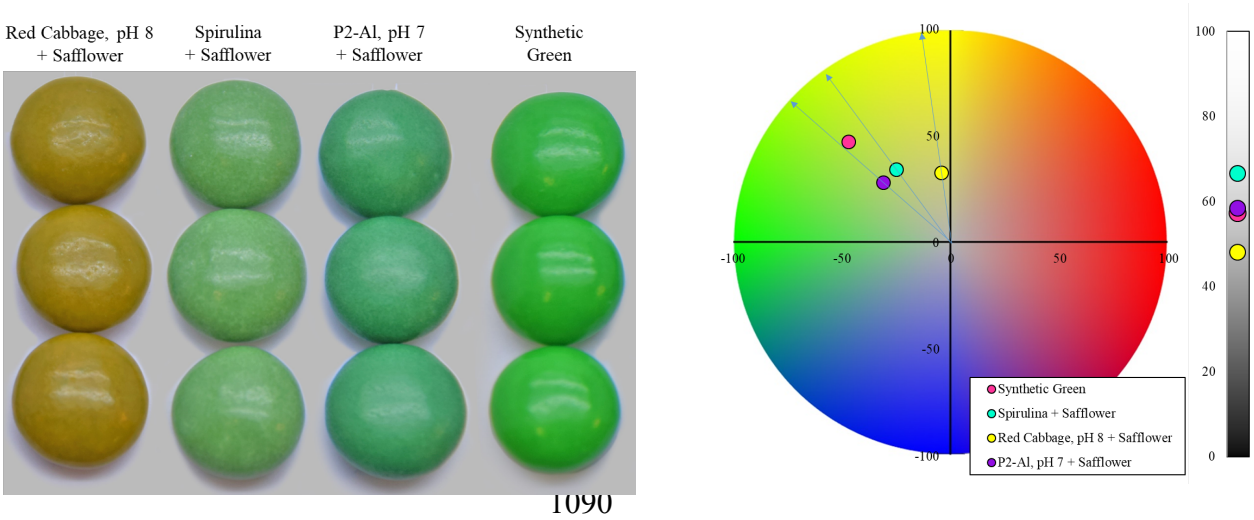


Supplementary Figure 13.3. Lentils of four blue colorants were obtained under typical panning conditions. The colorants used were FD&C Blue No. 1, spirulina, red cabbage at pH 8, and $Al^{3+}(P2^{-})_3$ at

pH 7. The colorimetry data are plotted in a*b* space. Hue angles are shown by arrows and L* values in sidebar. $Al^{3+}(P2^{-})_3$ at pH 7 provides a similar hue angle to FD&C Blue No. 1, 215.5 and 219.3 respectively. Whereas both spirulina and red cabbage colorants provide more violet-containing hue angles, 243.2 and 243.7 respectively. As can be seen, the $Al^{3+}(P2^{-})_3$ at pH 7 has a lower violet contribution. The colorimetry data is summarized in Supplementary Table 13.3 below. The L*a*b* values were measured using a Konica Minolta CM-5 colorimeter with 8 mm mask, 10-degree observer, D65 lighting, and SCE. Measured values were plotted on a*b* space. Hue angles are shown by arrows and L* values in sidebar. Photo Credit: Randall Powers, Mars, Incorporated.

Supplementary Table 13.3. Colorimetric data collected for blue lentil prototypes

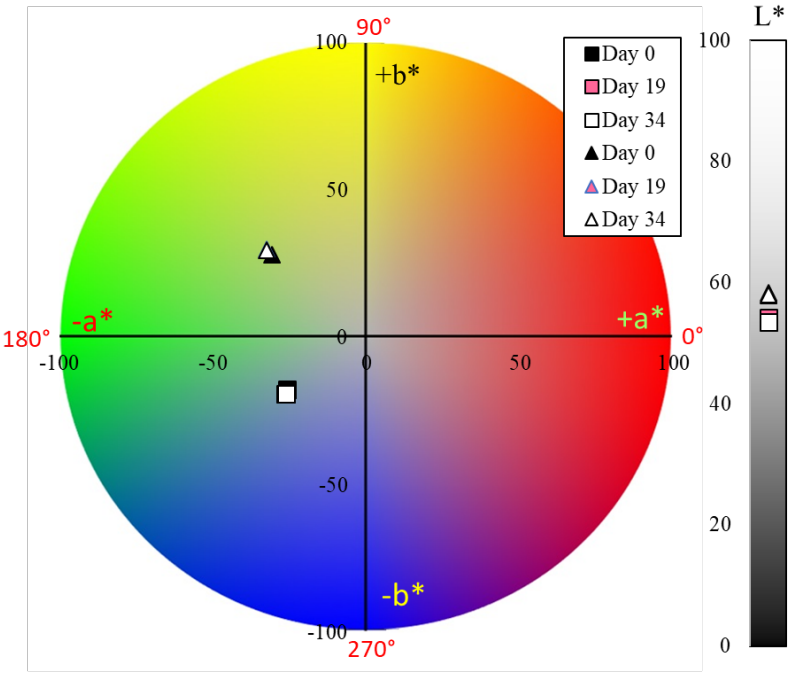
Name	L*	a*	b*	C*	h°
FD&C Blue No. 1	58.64	-38.68	-31.67	49.99	219.31
Spirulina	59.74	-14.58	-28.80	32.28	243.15
Red Cabbage, pH 8	40.09	-9.12	-18.47	20.60	243.72
$Al^{3+}(P2^{-})_3$, pH 7	47.55	-24.79	-17.69	30.45	215.51



1090

Supplementary Figure 13.4. Lentils of four green colorants were obtained under typical panning conditions. The colorant blends used were synthetic green, spirulina, red cabbage at pH 8, and $Al^{3+}(P2^{-})_3$ at pH 7 with the latter three mixed with safflower as the yellow component. The L*a*b* values were measured using a Konica Minolta CM-5 colorimeter with 8 mm mask, 10-degree observer, D65 lighting, and SCE. Measured values were plotted on a*b* space. Hue angles are shown by arrows and L* values in sidebar. The red cabbage at pH 8 and spirulina, having higher violet contributions of 37.3 and 24.4, respectively prevent achieving the desired green hue angle of 135. For example, the red cabbage and safflower blend at pH 8 has a hue angle of 97.6 and the spirulina and safflower blend has a hue angle of 126.2. The $Al^{3+}(P2^{-})_3$ at pH 7 has a lower violet contribution of 17.4 (see Supplementary Table 2.1) As the violet contribution decreases, the hue angle of the green aligns closer to the synthetic green

(provided in Supplementary Table 13.4). The $Al^{3+}(P2^{-})_3$ at pH 7 has a hue angle of 137.7 whereas synthetic control is 134.8. Photo Credit: Randall Powers, Mars, Incorporated.



Supplementary Table 13.4. Colorimetric data collected for green lentil prototypes

Name	L*	a*	b*	C*	h°
Synthetic Green	57.27	-46.84	47.15	66.46	134.81
Spirulina + Safflower	66.55	-25.00	34.15	42.32	126.20
Red Cabbage, pH 8 + Safflower	47.98	-4.35	32.72	33.00	97.57
$Al^{3+}(P2^{-})_3$, pH 7 + Safflower	58.31	-30.91	28.14	41.80	137.69

Supplementary Figure 13.5. Shelf-life of the $Al^{3+}(P2^{-})_3$, pH 7 on finished blue and green candy. The candies were stored at ambient conditions and measured over time to assess color stability. The data is shown in Supplementary Table 13.5. The data in a solid sugar shell matrix shows good stability. In the figure, blue candies are represented by a square and green candies by a triangle. As can be seen, the L*,

a*, and b* remain nearly constant. The L*a*b* values were measured using a Konica Minolta CM-5 colorimeter with 8 mm mask, 10-degree observer, D65 lighting, and SCE. Measured values were plotted on a*b* space. L* values are shown in sidebar.

Table 13.5. Raw colorimetric data for finished lentil color stability study

Prototype	Time	L*	a*	b*	C*	h°
(square) $Al^{3+}(P2^{-})_3$ Blue	Day 0	53.61	-26.06	-17.66	31.48	214.13
	Day 19	54.16	-26.64	-19.17	32.82	215.74
	Day 34	53.45	-26.35	-19.08	32.53	215.91
	Day 310*	53.30	-26.87	-19.78	33.37	216.36
(triangle) $Al^{3+}(P2^{-})_3$ + Safflower = Green	Day 0	58.31	-30.91	28.14	41.80	137.69
	Day 19	57.94	-32.60	29.82	44.18	137.55
	Day 34	57.95	-32.59	29.61	44.03	137.75
	Day 310*	58.90	-32.89	30.01	44.52	137.62
*Data not shown in Supplementary Figure 13.5 for clarity since change was minor						

REFERENCES AND NOTES

1. F. M. Clydesdale, Color as a factor in food choice. *Crit. Rev. Food Sci. Nutr.* **33**, 83–101 (2009).
2. D. Asioli, J. Aschemann-Witzel, V. Caputo, R. Vecchio, A. Annunziata, T. Næs, P. Varela, Making sense of the "clean label" trends: A review of consumer food choice behavior and discussion of industry implications. *Food Res. Int.* **99**, 58–71 (2017).
3. K. Kupferschmidt, In search of blue. *Science* **364**, 424–429 (2019).
4. B.C. Freitas-Dörr, C. O. Machado, A. C. Pinheiro, A. B. Fernandes, F. A. Dörr, E. Pinto, M. Lopes-Ferreira, M. Abdellah, J. Sá, L. C. Russo, F. L. Forti, L. C. P. Gonçalves, E. L. Bastos, A metal-free blue chromophore derived from plant pigments. *Sci. Adv.* **6**, eaaz0421 (2020).
5. A. E. Smith, H. Mizoguchi, K. Delaney, N. A. Spaldin, A. W. Sleight, M. A. Subramanian, Mn^{3+} in trigonal bipyramidal coordination: A new blue chromophore. *J. Am. Chem. Soc.* **131**, 17084–17086 (2009).
6. M. Buchweitz, Natural solutions for blue colors in food, in *Handbook on Natural Pigments in Food and Beverages: Industrial Applications for Improving Food Color*, R. Carle, R. M. Schweiggert, Eds. (Elsevier, 2016), chapter 17, pp. 355–384.
7. P. Nabais, J. Oliveira, F. Pina, N. Teixeira, V. de Freitas, N. F. Brás, A. Clemente, M. Rangel, A. M. S. Silva, M. J. Melo, A 1000-year-old mystery solved: Unlocking the molecular structure for the medieval blue from *Chrozophora tinctoria*, also known as folium. *Sci. Adv.* **6**, eaaz7772 (2020).
8. A. G. Newsome, C. A. Culver, R. B. van Breemen, Nature's palette: The search for natural blue colorants. *J. Agric. Food Chem.* **62**, 6498–6511 (2014).
9. L. Jespersen, L. D. Strømdahl, K. Olsen, L. H. Skibsted, Heat and light stability of three natural blue colorants for use in confectionary beverages. *Eur. Food Res. Technol.* **220**, 261–266 (2004).
10. A. Chaovanalikit, M. M. Thompson, R. E. Wrolstad, Characterization and quantification of anthocyanins and polyphenolics in blue honeysuckle (*Lonicera caerulea* L.). *J. Agric. Food Chem.* **52**, 848–852 (2004).
11. G. T. Sigurdson, P. Tang, M. M. Giusti, Natural colorants: Food colorants from natural sources. *Annu. Rev. Food Sci. Technol.* **8**, 261–280 (2017).
12. I. Viera, A. Pérez-Gálvez, M. Roca, Green natural colorants. *Molecules* **24**, 154 (2019).
13. O. Dangles, J.-A. Fenger, The chemical reactivity of anthocyanins and its consequences in food science and nutrition. *Molecules* **23**, 1970 (2018).

14. F. Pina, M. J. Melo, C. A. T. Laia, A. J. Parola, J. C. Lima, Chemistry and applications of flavylium compounds: A handful of colours. *Chem. Soc. Rev.* **42**, 869–908 (2012).
15. P. Trouillas, J. C. Sancho-García, V. De Freitas, J. Gierschner, M. Otyepka, O. Dangles, Stabilizing and modulating color by copigmentation: Insights from theory and experiment. *Chem. Rev.* **116**, 4937–4982 (2016).
16. O. Dangles, N. Saito, R. Brouillard, Kinetic and thermodynamic control of flavylium hydration in the pelargonidin-cinnamic acid complexation. Origin of the extraordinary flower color diversity of Pharbitis nil. *J. Am. Chem. Soc.* **115**, 3125–3132 (1993).
17. K. Yoshida, M. Mori, T. Kondo, Blue flower color development by anthocyanins: From chemical structure to cell physiology. *Nat. Prod. Rep.* **26**, 884–915 (2009).
18. J.-A. Fenger, M. Moloney, R. J. Robbins, T. M. Collins, O. Dangles, The influence of acylation, metal binding and natural antioxidants on the thermal stability of red cabbage anthocyanins in neutral solution. *Food Funct.* **10**, 6740–6751 (2019).
19. G. T. Sigurdson, R. J. Robbins, T. M. Collins, M. M. Giusti, Molar absorptivities (ε) and spectral and colorimetric characteristics of purple sweet potato anthocyanins. *Food Chem.* **271**, 497–504 (2019).
20. N. Ahmadiani, R. J. Robbins, T. M. Collins, M. M. Giusti, Molar absorptivity (ε) and spectral characteristics of cyanidin-based anthocyanins from red cabbage. *Food Chem.* **197**, 900–906 (2016).
21. M. Moloney, R. J. Robbins, T. M. Collins, T. Kondo, K. Yoshida, O. Dangles, Red cabbage anthocyanins: The influence of D-glucose acylation by hydroxycinnamic acids on their structural transformations in acidic to mildly alkaline conditions and on the resulting color. *Dyes Pigments* **158**, 342–352 (2018).
22. K. Ikeda, Structure of two acylated anthocyanins from red cabbage (*Brassica oleracea*). *Chem. Express* **2**, 563–566 (1987).
23. N. Saito, F. Tatsuzawa, E. Suenaga, K. Toki, K. Shinoda, A. Shigihara, T. Honda, Tetra-acylated cyanidin 3-sophoroside-5-glucosides from the flowers of *Iberis umbellata* L.(*Cruciferae*). *Phytochemistry* **69**, 3139–3150 (2008).
24. R. Matera, S. Gabbanini, G. R. De Nicola, R. Iori, G. Petrillo, L. Valgimigli, Identification and analysis of isothiocyanates and new acylated anthocyanins in the juice of *Raphanus sativus* cv. Sango sprouts. *Food Chem.* **133**, 563–572 (2012).

25. R. Matera, S. Gabbanini, S. Berretti, R. Amorati, G. R. De Nicola, R. Iori, L. Valgimigli, Acylated anthocyanins from sprouts of *Raphanus sativus* cv. Sango: Isolation, structure elucidation and antioxidant activity. *Food Chem.* **166**, 397–406 (2015).
26. T. Ito, K.-i. Oyama, K. Yoshida, Direct observation of hydrangea blue-complex composed of 3-*O*-glucosyldelphinidin, Al^{3+} and 5-*O*-acylquinic acid by ESI-mass spectrometry. *Molecules* **23**, 1424 (2018).
27. G. T. Sigurdson, R. J. Robbins, T. M. Collins, M. M. Giusti, Spectral and colorimetric characteristics of metal chelates of acylated cyanidin derivatives. *Food Chem.* **221**, 1088–1095 (2017).
28. K. Takeda, Blue metal complex pigments involved in blue flower color. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **82**, 142–154 (2006).
29. T. Kondo, M. Ueda, H. Tamura, K. Yoshida, M. Isobe, T. Goto, Composition of protocyanin, a self-assembled supramolecular pigment from the blue cornflower, *Centaurea cyanus*. *Angew. Chem. Int.* **33**, 978–979 (1994).
30. M. Rusishvili, L. Grisanti, S. Laporte, M. Micciarelli, M. Rosa, R. J. Robbins, T. Collins, A. Magistrato, S. Baroni, Unraveling the molecular mechanisms of color expression in anthocyanins. *Phys. Chem. Chem. Phys.* **21**, 8757–8766 (2019).
31. O.B. Malcıoğlu, A. Calzolari, R. Gebauer, D. Varsano, S. Baroni, Dielectric and thermal effects on the optical properties of natural dyes: A case study on solvated cyanin. *J. Am. Chem. Soc.* **133**, 15425–15433 (2011).
32. L. Jeske, S. Placzek, I. Schomburg, A. Chang, D. Schomburg, BRENDA in 2019: A European ELIXIR core data resource. *Nucleic Acids Res.* **47**, D542–D549 (2018).
33. K. K. Kim, H. K. Song, D. H. Shin, K. Y. Hwang, S. Choe, O. J. Yoo, S. W. Suh, Crystal structure of carboxylesterase from *Pseudomonas fluorescens*, an α/β hydrolase with broad substrate specificity. *Structure* **5**, 1571–1584 (1997).
34. R. Kleffner, J. Flatten, A. Leaver-Fay, D. Baker, J. B. Siegel, F. Khatib, S. Cooper, Foldit Standalone: A video game-derived protein structure manipulation interface using Rosetta. *Bioinformatics* **33**, 2765–2767 (2017).
35. J. K. Leman, B. D. Weitzner, S. M. Lewis, J. Adolf-Bryfogle, N. Alam, R. F. Alford, M. Aprahamian, D. Baker, K. A. Barlow, P. Barth, B. Basanta, B. J. Bender, K. Blacklock, J. Bonet, S. E. Boyken, P. Bradley, C. Bystroff, P. Conway, S. Cooper, B. E. Correia, B. Coventry, R. Das, R. M. De Jong, F. D. Maio, L. Dsilva, R. Dunbrack, A. S. Ford, B. Frenz, D. Y. Fu, C. Geniesse, L.

Goldschmidt, R. Gowthaman, J. J. Gray, D. Gront, S. Guffy, S. Horowitz, P.-S. Huang, T. Huber, T. M. Jacobs, J. R. Jeliazkov, D. K. Johnson, K. Kappel, J. Karanicolas, H. Khakzad, K. R. Khar, S. D. Khare, F. Khatib, A. Khramushin, I. C. King, R. Kleffner, B. Koepnick, T. Kortemme, G. Kuenze, B. Kuhlman, D. Kuroda, J. W. Labonte, J. K. Lai, G. Lapidoth, A. Leaver-Fay, S. Lindert, T. Linsky, N. London, J. H. Lubin, S. Lyskov, J. Maguire, L. Malmström, E. Marcos, O. Marcu, N. A. Marze, J. Meiler, R. Moretti, V. K. Mulligan, S. Nerli, C. Norn, S. Ó'Conchúir, N. Ollikainen, S. Ovchinnikov, M. S. Pacella, X. Pan, H. Park, R. E. Pavlovicz, M. Pethe, B. G. Pierce, K. B. Pilla, B. Raveh, P. D. Renfrew, S. S. Roy Burman, A. Rubenstein, M. F. Sauer, A. Scheck, W. Schief, O. Schueler-Furman, Y. Sedan, A. M. Sevy, N. G. Sgourakis, L. Shi, J. B. Siegel, D.-A. Silva, S. Smith, Y. Song, A. Stein, M. Szegedy, F. D. Teets, S. B. Thyme, R. Y.-R. Wang, A. Watkins, L. Zimmerman, R. Bonneau, Macromolecular modeling and design in Rosetta: Recent methods and frameworks. *Nat. Methods* **17**, 665–680 (2020).

36. R. E. Wrolstad, C. A. Culver, Alternatives to those artificial FD&C food colorants. *Annu. Rev. Food Sci. Technol.* **3**, 59–77 (2012).
37. U.S. Food and Drug Administration (FDA), *Summary of Color Additives for Use in the United States in Foods, Drugs, Cosmetics, and Medical Devices* (EPA, 2015); https://www.fda.gov/industry/color-additive-inventories/summary-color-additives-use-united-states-foods-drugs-cosmetics-and-medical-devices.
38. M. W. Pariza, E. M. Foster, Determining the safety of enzymes used in food processing. *J. Food Prot.* **46**, 453–468 (1983).
39. J. Hutter, M. Iannuzzi, F. Schiffmann, J. V. Vondele, CP2K: Atomistic simulations of con-densed matter systems. *WIREs Comput. Mol. Sci.* **4**, 15–25 (2014).
40. S. Goedecker, M. Teter, J. Hutter, Eparable dual-space gaussian pseudopotentials. *Phys.Rev. B* **54**, 1703–1710 (1996).
41. A. D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
42. C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785 (1988).
43. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **143**, 154104 (2010).

44. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, R. M. Wentzcovitch, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J. Condens. Matter Phys.* **21**, 395502 (2009).

45. P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. Di Stasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, S. Baroni, Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys. Condens. Matter* **29**, 465901 (2017).

46. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

47. O. Andreussi, I. Dabo, N. Marzari, Revised self-consistent continuum solvation in electronic-structure calculations. *J. Chem. Phys.* **136**, 064102 (2012).

48. S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).

49. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).

50. J. Wang, W. Wang, P. A. Kollman, D. A. Case, Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).

51. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general AMBER force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

52. G. Bussi, Hamiltonian replica exchange in GROMACS: A flexible implementation. *Mol. Phys.* **112**, 379–384 (2014).

53. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
54. F. Ascinar, G. Weingart, T. L. Tickle, C. Huttenhower, N. Segata, Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **3**, e1029 (2015).
55. T. A. Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
56. K. Yoshida, T. Kondo, T. Goto, Intramolecular stacking conformation of gentiodelphin, a diacylated anthocyanin from *Gentiana makinoi*. *Tetrahedron* **42**, 4313–4326 (1992).
57. K. Yoshida, K.-i. Oyama, T. Kondo, Structure of polyacylated anthocyanins and their UV protective effect. *Rec. Adv. Polyphen. Res.* **5**, 171–192 (2017).
58. K. Yoshida, T. Kondo, T. Goto, Unusually stable monoacylated anthocyanin from purple yam *Dioscorea alata*. *Tetrahedron Lett.* **32**, 5579–5580 (1991).
59. J. K. Barton, A. Danishefsky, J. Goldberg, Tris (phenanthroline) ruthenium (II): Stereoselectivity in binding to DNA. *JACS* **106**, 2172–2176 (1984).
60. R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* **16**, 1190–1208 (1995).
61. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
62. J. Ridley, M. Zerner, An intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and the azines. *Theor. Chim. Acta* **32**, 111–134 (1973).
63. F. Neese, The ORCA program system. *Wiley Interidscip. Rev. Comput. Mol. Sci.* **2**, 73–78 (2012).
64. F. Neese, Software update: The ORCA program system, version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1327 (2017).
65. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
66. A. Rodriguez, A. Laio, Clustering by fast search and find of density peaks. *Science* **344**, 1492–1496 (2014).
67. T. Stahl, H. Taschan, H. Brunn, Aluminum content of selected foods and food products. *Environ. Sci. Eur.* **23**, 37 (2011).

68. S. M. Sayied, R. A. Yokel, Aluminum content of some foods and food products in the USA, with aluminum food additives. *Food Addit. Contam.* **22**, 234–244 (2005).