**Acylated anthocyanins from sprouts of *Raphanus sativus* cv. Sango: isolation, structural elucidation and antioxidant activity**

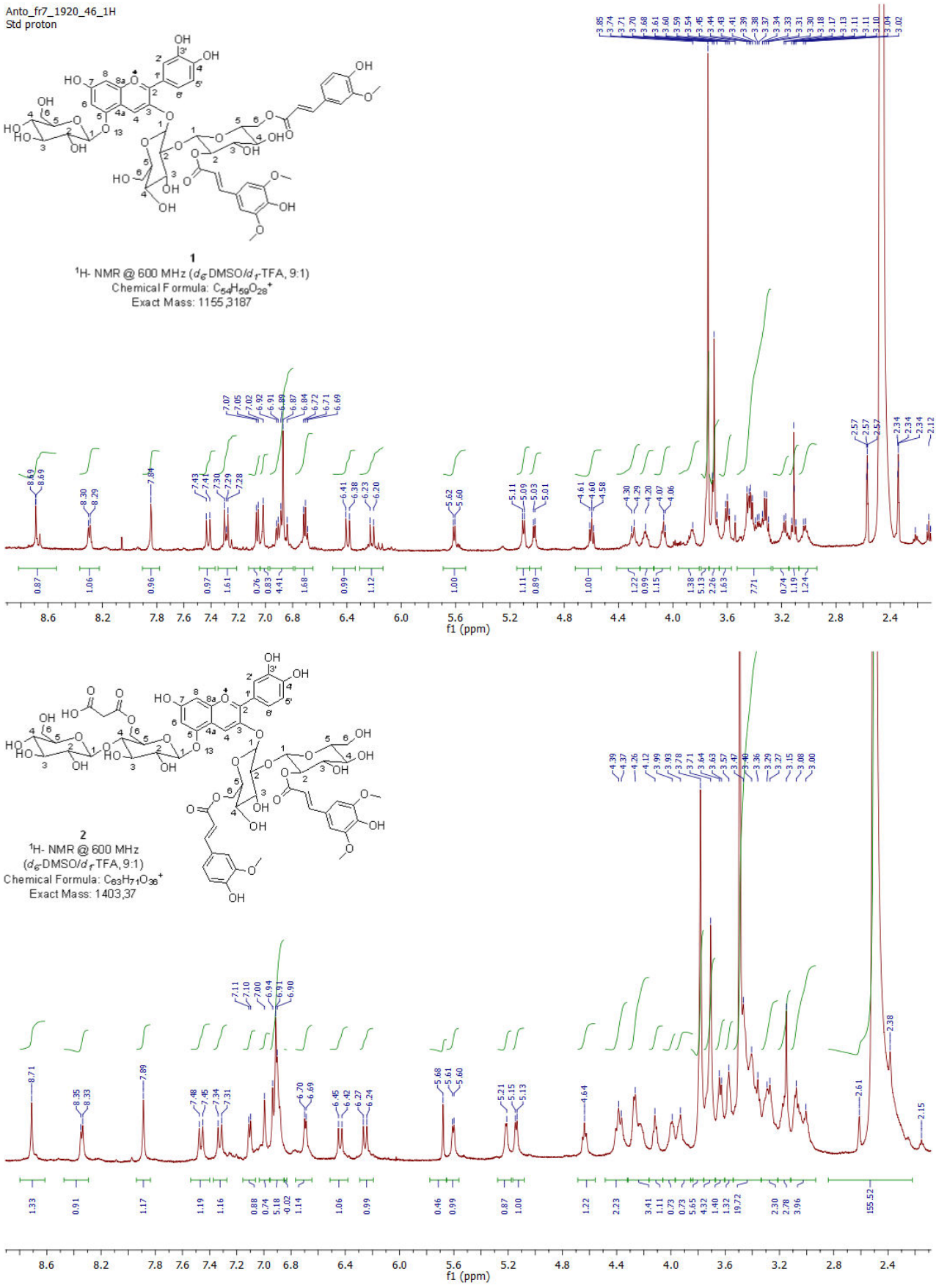
Riccardo Matera[a], Simone Gabbanini[a], Serena Berretti[b], Riccardo Amorati[b], Gina Rosalinda De Nicola[c], Renato Iori[c], Luca Valgimigli[b,*]

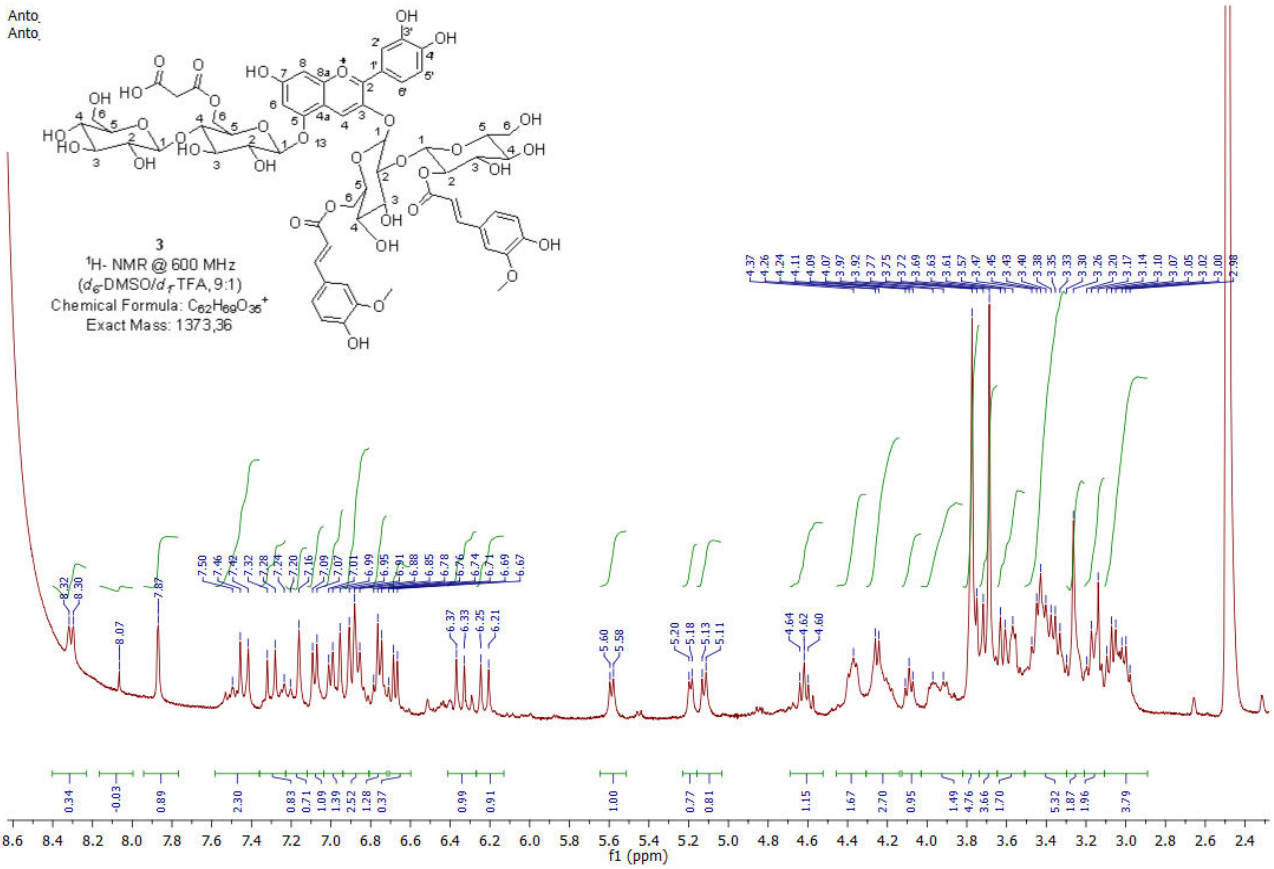
[a] *R&D Division, BeC S.r.l., Via C. Monteverdi 49, 47122 Forlì, Italy.* [b] *Department of Chemistry ''G. Ciamician'', University of Bologna, Via S. Giacomo 11, 40126 Bologna, Italy.* [c] *Industrial Crop Research Centre, Agricultural Research Council, (CRA-CIN), Via di Corticella, 133, 40128 Bologna, Italy*

# APPENDIX
(supplementary data)

## Content

**Figure S1.** $^{1}$H-NMR spectra of new anthocyanins (**1-4**)




Anto
Anto
3
¹H- NMR @ 600 MHz
($d_6$-DMSO/$d_1$TFA, 9:1)
Chemical Formula: $C_{62}H_{69}O_{35}^+$
Exact Mass: 1373,36
f1 (ppm)

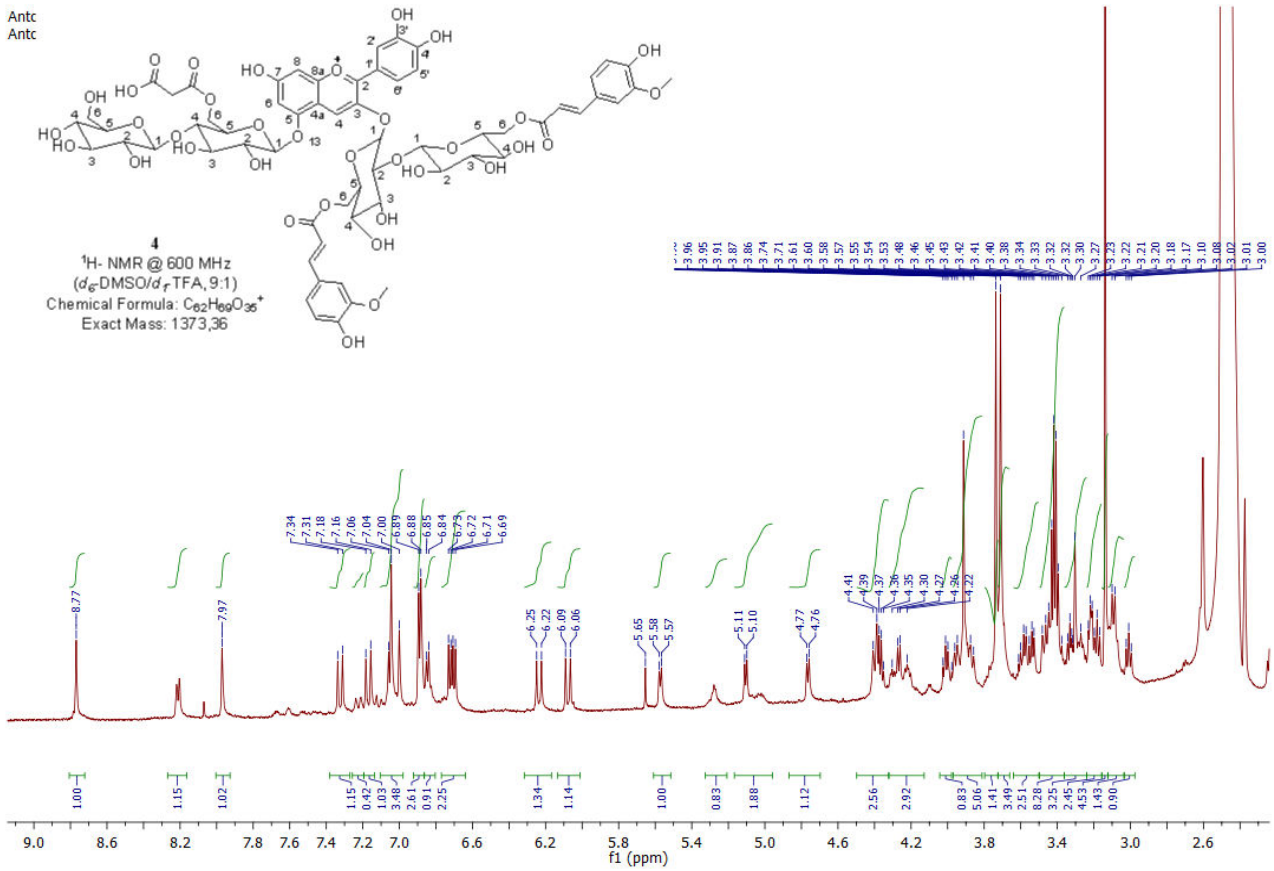

Antc
Antc
4
¹H- NMR @ 600 MHz
($d_6$-DMSO/$d_1$TFA, 9:1)
Chemical Formula: $C_{62}H_{69}O_{35}^+$
Exact Mass: 1373,36
f1 (ppm)

**Table S1.** UV-Vis spectroscopic data for anthocyanins isolated (**1-9**) from *R. sativus* cv. Sango.

| | Wavelength (nm) | | Absorbivity (μAbs)[a] | | | Acylation | Glycosylation |
|---|---|---|---|---|---|---|---|
| **id** | $\lambda_{acyl}$ **max** | $\lambda_{vis}$ **max** | $E_{acyl}$ | $E_{vis}$ | $E_{440}$ | **Ratio (%)[b]** | **Ratio (%)[c]** |
| **1** | 329 | 535 | 124516 | 118384 | 20397 | 105 | 17 |
| **2** | 330 | 535 | 98510 | 100200 | 22999 | 98 | 23 |
| **3** | 329 | 535 | 59400 | 60802 | 10960 | 98 | 18 |
| **4** | 328 | 534 | 31976 | 30016 | 4799 | 107 | 16 |
| **5** | 329 | 533 | 49976 | 504019 | 8591 | 99 | 17 |
| **6** | 334 | 527 | 32008 | 55050 | 9991 | 58 | 18 |
| **7** | 329 | 533 | 128900 | 100657 | 22218 | 128 | 22 |
| **8** | 327 | 522 | 63900 | 120657 | 22015 | 53 | 18 |
| **9** | 328 | 530 | 122049 | 110620 | 20065 | 110 | 18 |
| **C3G[d]** | / | 520 | | | | / | 29 |

[a] Evaluated by HPLC-DAD on Hypersil GOLD PFP column (5 μm, 175 Å, 4.6 ∅ × 250 mm) eluting with A (aq 0.12% v/v TFA) and B (MeCN + 0.12% v/v TFA). See gradient program on the text.
[b] Acylation ratio = $E_{acyl}/E_{vis}$, 59–63% indicated monoacylation, 98–128% indicates diacylation.
[c] Glycosylation ratio = $E_{440}/E_{vis}$, 15–24% indicated bioside, 29–35% indicates monoside.
[d] Cyanidin 3-glucoside (data from Fossen & Andersen[1])

## Structure elucidation of new anthocyanins (3-4)

The novel anthocyanin **3** ($C_{62}H_{69}O_{35}$) (Figure 1) shows a high percentage ratio of the UV-Vis absorbance ($E_{acyl}/E_{vis}$ = 98 %) indicating an anthocyanin bioside biacylated with two aromatic residues. The ESI(+)MS analysis shows *m/z* 1373 as molecular ion. $MS^2$ fragmentations of *m/z* 1373 revealed the three main ions (*m/z* 1329, *m/z* 963 and *m/z* 697) originating by loss of carbon dioxide (44 amu), malonoyl-diglucoside (410 amu) and the typical diglycosyl-diferuloyl (676 amu) residue, respectively. $MS^3$ fragmentation of each ion produced the cyanidin aglycone *m/z* 287 ion (Table 1). The anthocyanin **3** was presumed to be based on cyanidin-3-diferuloyl-diglucoside-5-malonoyl-diglucoside and was subjected to NMR characterization to assign the exact connectivity of glycosyl and acyl moieties (Tables 2-3 and $^1$H spectrum in Appendix). Four signals corresponding to the chemical shifts of acetalic groups formed by the anomeric carbons were detected and assigned to different sugars using previous reported data: the acetalic proton's signal shifted at the lowest field ($\delta$ 5.58) belongs to the $H_1$ of Glc A proton, whereas signals at higher field belong to Glc B ($\delta$ 5.12) and Glc C ($\delta$ 5.19), respectively. The doublet of acetalic proton shifted at the highest field ($\delta$ 4.26) belongs to Glc D. The exact connectivity pattern of sugar residues was obtained by $^1$H and COSY experiments. The typical diaxial protons coupling constants ($J$ = 7.0-8.4 Hz) of anomeric protons of sugar residues assigned to β-glucopyranoside, were observed in $^1$H-

[1] Fossen, T. & Andersen M. *Phytochemistry*, **1998,** *49*, 1065-1068

NMR experiments even though very weak HMBC correlation between $H_2$ of Glc A ($\delta$ 4.12) with $C_1$ of Glc B ($\delta$ 99.7) were detected to confirm the typical connectivity of β-1,2 glycosidic linkage of sophorose. As already described, sophorose is generally linked to $C_3$ position of cyanidin. Four couples of methylene protons belong to $C_6$ positions of each glucose as confirmed by HSQC together with an additional cross peak indicating methylene of malonic acid ($R_4$) ($\delta$ 3.30). The downfield signals belonging to geminal $H_{6a}$ e $H_{6b}$ of Glc A and Glc C refer to these positions acylated with ferulic and malonic acid, respectively, as confirmed also by weak HMBC correlation between ferulic ester carbonyl ($R_1$) ($\delta$ 166.7) and two geminal protons $H_{6a}$ e $H_{6b}$ of Glc A ($\delta$ 4.27 and 4.37), together with correlation between malonic ester carbonyl ($R_4$) ($\delta$ 167.7) and two protons $H_{6a}$ e $H_{6b}$ of Glc C ($\delta$ 4.20, 4.35). Low HMBC correlation between carbonyl ($\delta$ 166.3) of another ferulic acid ($R_3$) with $H_2$ of Glc B ($\delta$ 4.64) confirms this acylation site. The ferulic residues' configuration was assigned as *trans* due to the large coupling constants ($J$ = 15-16 Hz). The complete data set of anthocyanin is very similar to the compound **2** with the only difference related to ferulic acid instead of sinapic acid on Glc B**.** Due to low amount of material analyzed, no HMBC correlation were observed between Glc C and Glc D because of the intrinsic difficulty with observing $^3J_{CH}$ in flexible cycles like glycosides. We refer to similar acylated anthocyanins isolated by Mori and coworkers from *Raphanus sativus* based on β-1,4 diglucoside (Mori et al., 2006).[2] Therefore, the novel anthocyanin **3** was identified as cyanidin 3-*O*-[6-*O*-(*E*)-feruloyl-2-*O*-(2-*O*-(*E*)-feruloyl-β-D-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(4-*O*-β-D-glucopyranosyl-6-*O*-malonoyl-β-D-glucopyranoside).

The ESI(+)MS analysis of the novel anthocyanin **4** ($C_{62}H_{69}O_{35}$) (Figure 1) showed *m/z* 1373 as molecular ion. The percentage ratio of the UV-Vis absorbance at the maximum of the acyl unit and the absorbance at the visible maximum ($E_{acyl}/E_{vis}$) was 107 % suggesting an anthocyanin bioside biacylated with two aromatic residues. $MS^2$ fragmentation of *m/z* 1373 caused the ion *m/z* 1329 by loss of a carbon dioxide residue (44 amu), *m/z* 963 by loss of a malonoyl-diglycosyl residue (410 amu) and *m/z* 697 by loss of a diglycosyl-diferuloyl residue (676 = 324+352 amu). $MS^3$ fragmentation of *m/z* 963 and *m/z* 697 ion produced, respectively, the fragments *m/z* 625 by loss of glycosyl-feruloyl (338 amu) residue and the cyanidin ion at *m/z* 287 by losing the already mentioned 410 amu residue (Table 1). The anthocyanin **4** was presumed to be based on cyanidin-3-diferuloyl-diglucoside-5-malonoyl-diglucoside and was subjected to NMR characterization to assign the exact connectivity of glycosyl and acyl moieties using $^1$H, COSY experiments (Tables 2 and $^1$H spectrum in Appendix). Similarly to analogue anthocyanin **2** bearing four glucose moieties,

[2] Mori, M., Nakagawa, S., Maeschima, M., Niiura, S., & Yoshida, K. *Heterocycles* **2006**, *69*, 239-251.

four signals corresponding to the chemical shifts of acetalic groups formed by the anomeric carbons were detected and assigned to different sugars using previously reported data: the acetalic proton's signal shifted at the lowest field ($\delta$ 5.58) belong to the $H_1$ of Glc A, then the upfield signals belong to Glc B ($\delta$ 4.70) and Glc C ($\delta$ 5.07), respectively. The doublet of the acetalic proton shifted at the highest field ($\delta$ 4.26) belongs to Glc D. The exact connectivity pattern of sugar residues was obtained by $^1$H and COSY experiments. The typical diaxial protons coupling constants ($J$ = 7.2-7.8 Hz) of anomeric protons of sugar residues, as shown in $^1$H-NMR experiments, was assigned to β-glucopyranoside, and confirmed the typical connectivity of β-1,2 glycosidic linkage of sophorose, which is expected to be linked to $C_3$ position of cyanidin, as previously described. Among the four signals of methylene protons belonging to $C_6$ positions of each glucose, the most downfield shifted ones belong to $H_{6a}$ and $H_{6b}$ of Glc A ($\delta$ 4.21 and 4.30) and Glc B ($\delta$ 4.27), indicating the acylation sites with two ferulic acids ($R_1$ and $R_2$), whereas $H_6$ of Glc D ($\delta$ 3.70) suggests the acylation with the malonic moiety ($R_4$), upon comparing the data with similar acylated anthocyanins. The ferulic residues' configuration was elucidated as *trans* due to the large coupling constants ($J$ = 16.2 Hz). The $^1$H-NMR data set of anthocyanin is very similar to compound **9,** except for the signals related to the additional Glc D. In the absence of sufficiently intense $^{13}$C signals (not listed in table 3) decision about connectivity between Glc C and Glc D was based on the work by Mori and coworkers on *Raphanus sativus* (Mori et al., 2006)[3]. Therefore, the novel anthocyanin **4** was identified as cyanidin 3-*O*-[6-*O*-(*E*)-feruloyl-2-*O*-(6-*O*-(*E*)-feruloyl-β-D-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(4-*O*-β-D-glucopyranosyl-6-*O*-malonoyl-β-D-glucopyranoside).

## Structure elucidation of known anthocyanins (5-9)

The characterization of compound **5** is discusses in detail, showing 2D NMR spectra, in order to illustrate the general approach used for structural elucidation of any isolated anthocyanin **1-9**.

The known anthocyanin **5** ($C_{57}H_{61}O_{31}$) (Figure 1) shows a high percentage ratio of the UV-Vis absorbance ($E_{acyl}/E_{vis}$ = 99 %) indicating an anthocyanin bioside biacylated with two aromatic residues. The ESI(+)MS analysis showed *m/z* 1241 as molecular ion. $MS^2$ fragmentations of *m/z* 1241 revealed three main ions (*m/z* 1197, *m/z* 993 and *m/z* 535) originating by loss of carbon dioxide (44 amu), malonoyl-glucoside (248 amu) and the typical diglycosyl-feruloylsinapoyl (706 amu) residues, respectively. $MS^3$ fragmentation of *m/z* 1197 ion produced the main ions *m/z* 993 and *m/z* 491, by loss of already mentioned 706 amu residue and the acetyl-glucoside (204 amu). $MS^4$ fragmentation of each ion produced the cyanidin aglycone fragment *m/z* 287 (Table 1). The

[3] Mori, M., Nakagawa, S., Maeschima, M., Niiura, S., & Yoshida, K. *Heterocycles* **2006**, *69*, 239-251.

anthocyanin **5** was presumed to be based on cyanidin-3-feruloyl-sinapoyl-diglucoside-5-malonoyl-glucoside and was subjected to complete NMR characterization for assign the exact connectivity of glycosyl and acyl moieties using $^1$H (See Fig S2), COSY, HSQC and HMBC experiments (Table 2 and 3). Three signals corresponding to the chemical shifts of the acetalic groups formed by the anomeric carbons were detected and assigned to different sugars using previous reported data: the acetalic proton's signal shifted at the lowest field ($\delta$ 5.64) belong to the $H_1$ of Glc A proton, while signals at higher field belong to Glc B ($\delta$ 5.15), and Glc C ($\delta$ 5.13), respectively. The exact connectivity pattern of sugar residues was obtained by $^1$H and COSY experiments (See Fig S3), and by comparing with Literature data.[4] The typical diaxial protons coupling constants ($J$ = 7.2-8.4 Hz) of anomeric protons of sugar residues shown in $^1$H-NMR experiments, which were assigned to β-glucopyranoside, together with the HMBC correlation between $H_1$ of Glc B ($\delta$ 5.15) with $C_2$ of Glc A ($\delta$ 77.5) (See Fig S4), confirmed the typical connectivity of β-1,2 glicosidic linkage of sophorose which is linked to C3 position of cyanidin as confirmed by HMBC correlation between the most deshielded acetalic signal belong to $H_1$ of Glc A ($\delta$ 5.64) and cyanidin $C_3$ ($\delta$ 144.6) (See Fig S5). Three couples of methylene protons belong to $C_6$ positions of each glucose as confirmed by HSQC correlations together with an additional cross peak indicating methylene protons of malonic acid ($\delta$ 3.34). The most deshielded signals belong to Glc A and Glc C indicate these positions are acylated with ferulic and malonic acid, respectively, as confirmed also by HMBC correlation between ferulic ester ($R_1$) carbonyl ($\delta$ 166.2) and two methylene protons $H_{6a}$ e $H_{6b}$ of Glc A ($\delta$ 4.18 and 4.39) (See figures S6-S7) together with correlation between malonic ester ($R_4$) carbonyl ($\delta$ 167.0), and two methylene protons $H_{6a}$ e $H_{6b}$ of Glc C ($\delta$ 3.88, 4.36). Proton signals from sinapic acid residue ($R_3$) were observed at $\delta$ 7.46, 6.91, 6.44 which correlated with the singlet ($\delta$ 3.78) integrating for six protons. Furthermore, HMBC correlation exists between proton at $H_2$ of Glc B ($\delta$ 4.65) and carbon signal of carbonyl ($\delta$ 166.3) of sinapic acid. The cinnamic residues' configuration was elucidated as *trans*, based on the large coupling constants ($J$ = 16.2 Hz). Olefinic ferulic ($R_1$) protons at α and β position ($\delta$ 6.24 and 7.32, respectively) correlated via HMBC with proton signal of 3-$OCH_3$ ($\delta$ 3.69) and the $C_3$ signal ($\delta$ 148.4) of the feruloyl acid (See Fig S8-S12), whereas a correlation was observed between the 3-$OCH_3$ and 5-$OCH_3$ proton signal ($\delta$ 3.78) and the carbon signal at the 3-and 5-position ($\delta$ 148.4) of sinapoyl acid ($R_3$). Definitely, spectroscopic data were consistent with published data of the identical anthocyanin isolated from *Brassica campestris* var. *chinensis,*[5] therefore the known anthocyanin **5** was identified cyanidin 3-*O*-[6-*O*-(*E*)-feruloyl-2-*O*-(2-*O*-(*E*)-sinapoyl-β-D-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(6-*O*-malonoyl-β-D-glucopyranoside).

[4] Otsuki, T., Matsufuji, H., Takeda, M., Toyoda, M., & Goda, Y. *Phytochemistry*, **2002,** *60*, 79-87
[5] Suzuki, M., Nagata, T., Terahara, N. *Biosci. Biotech. Biochem.,* **1997**, *61*, 1929-1930.

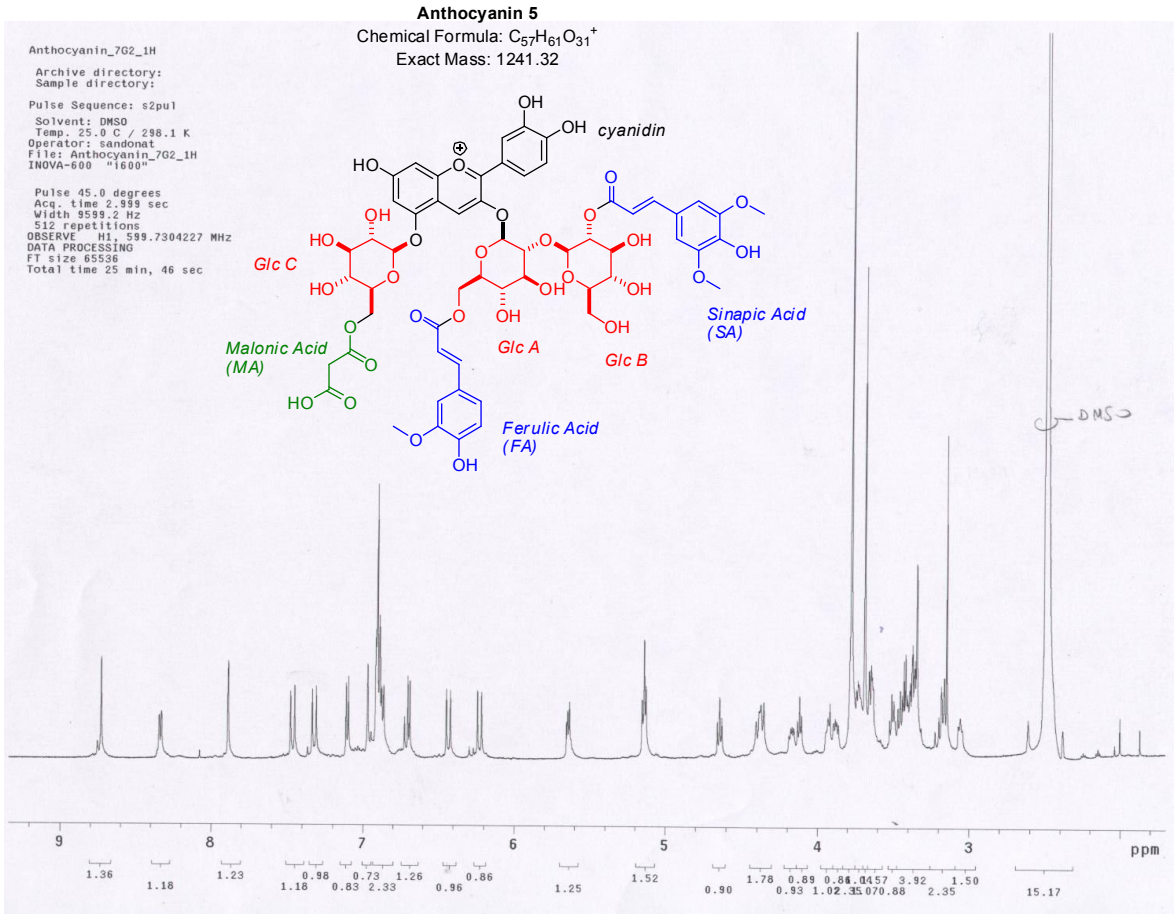



**Figure S2.** $^{1}$H-NMR of compound **5** (in **DMSO-$d_6$ : TFA-$d_1$** (9:1) @ 600 MHz)

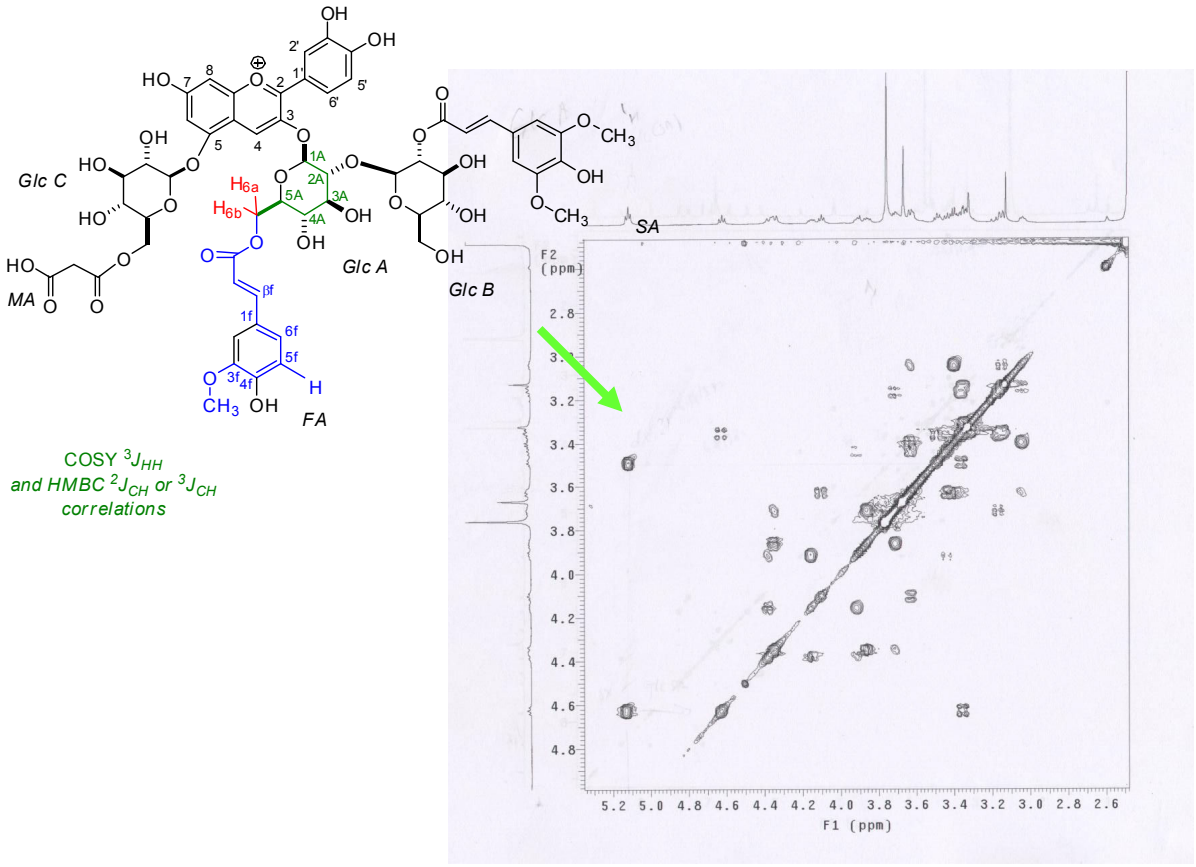



**Figure S3.** COSY plot of compound **5** (expansion of alifatic region)

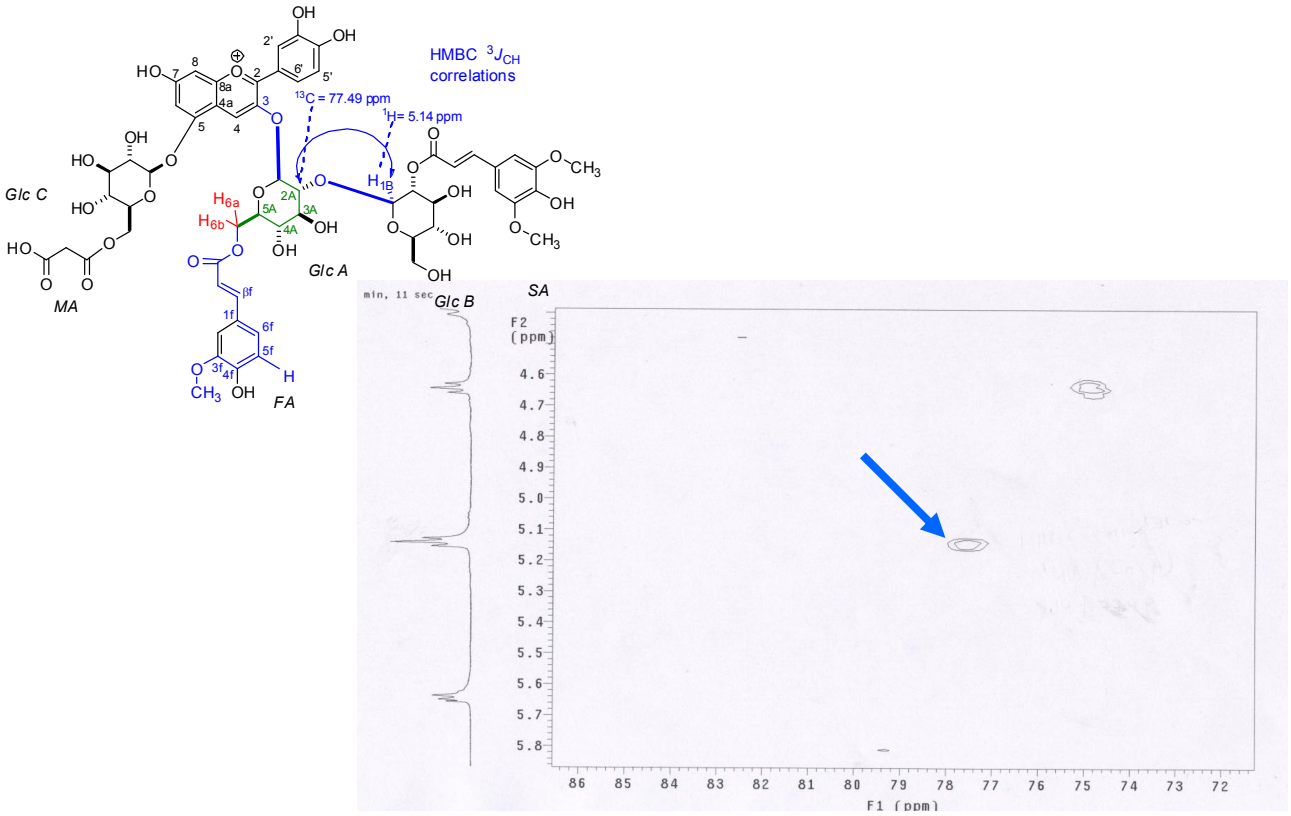



**Figure S4.** HMBC correlation between Glc A and Glc B for compound **5**.

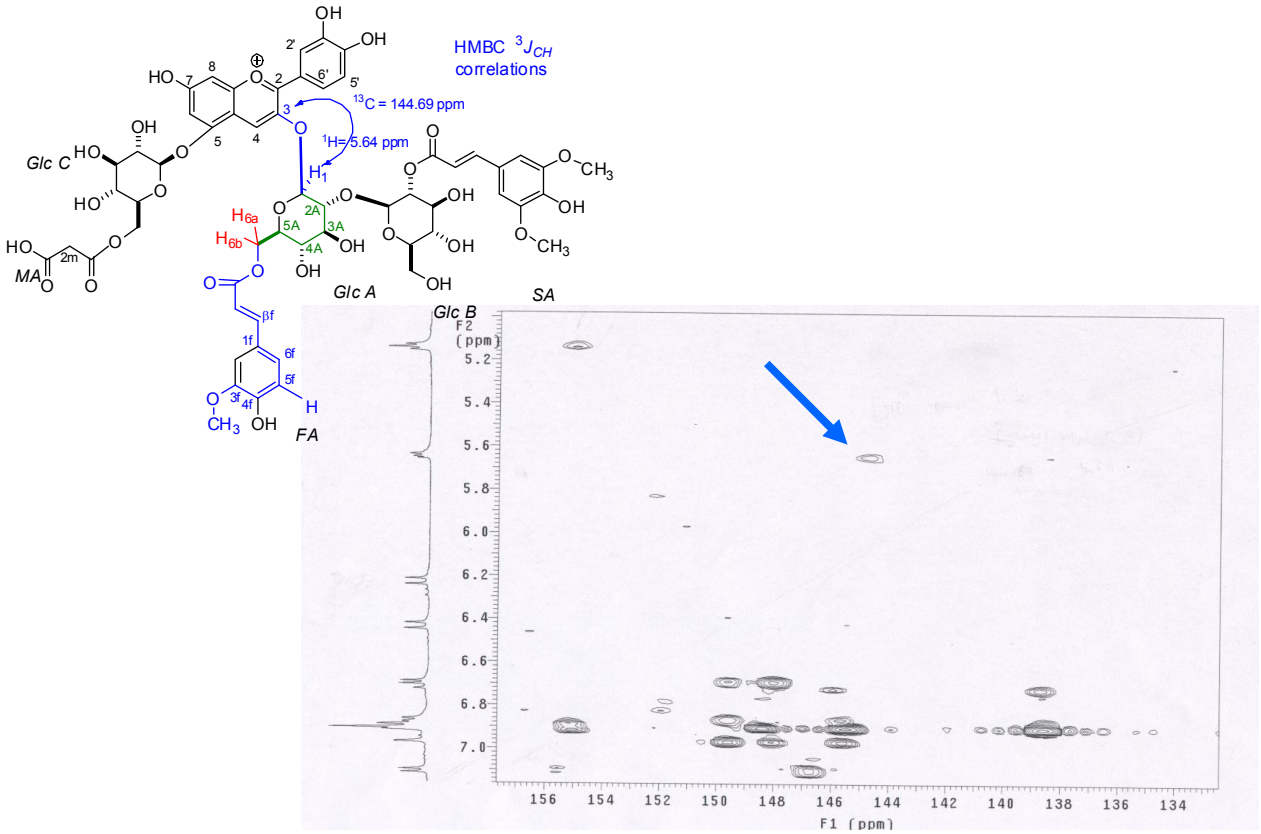



**Figure S5.** HMBC correlation between Glc A and $C_3$ of cyanidin, for compound **5**.

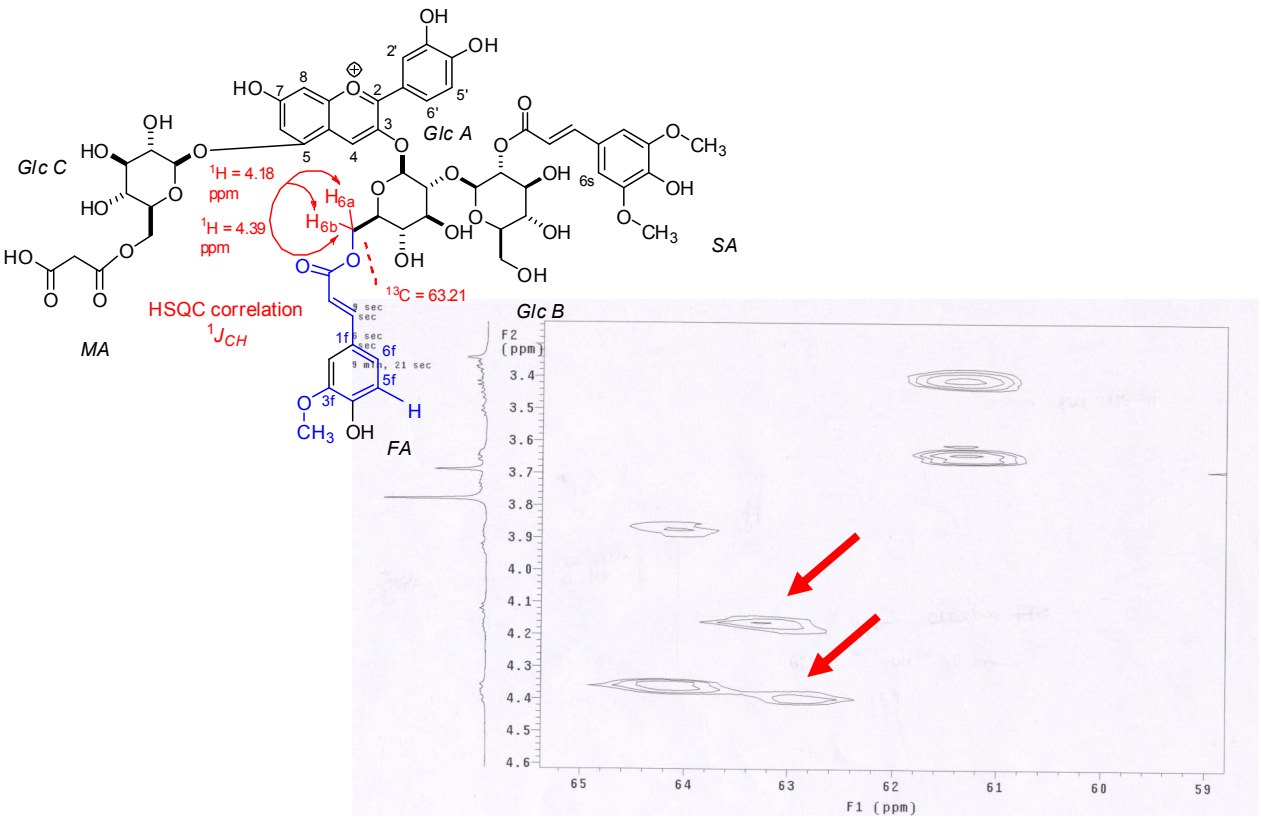



**Figure S6.** HSQC correlation of methylene protons of $H_6$ of Glc A for compound **5**.

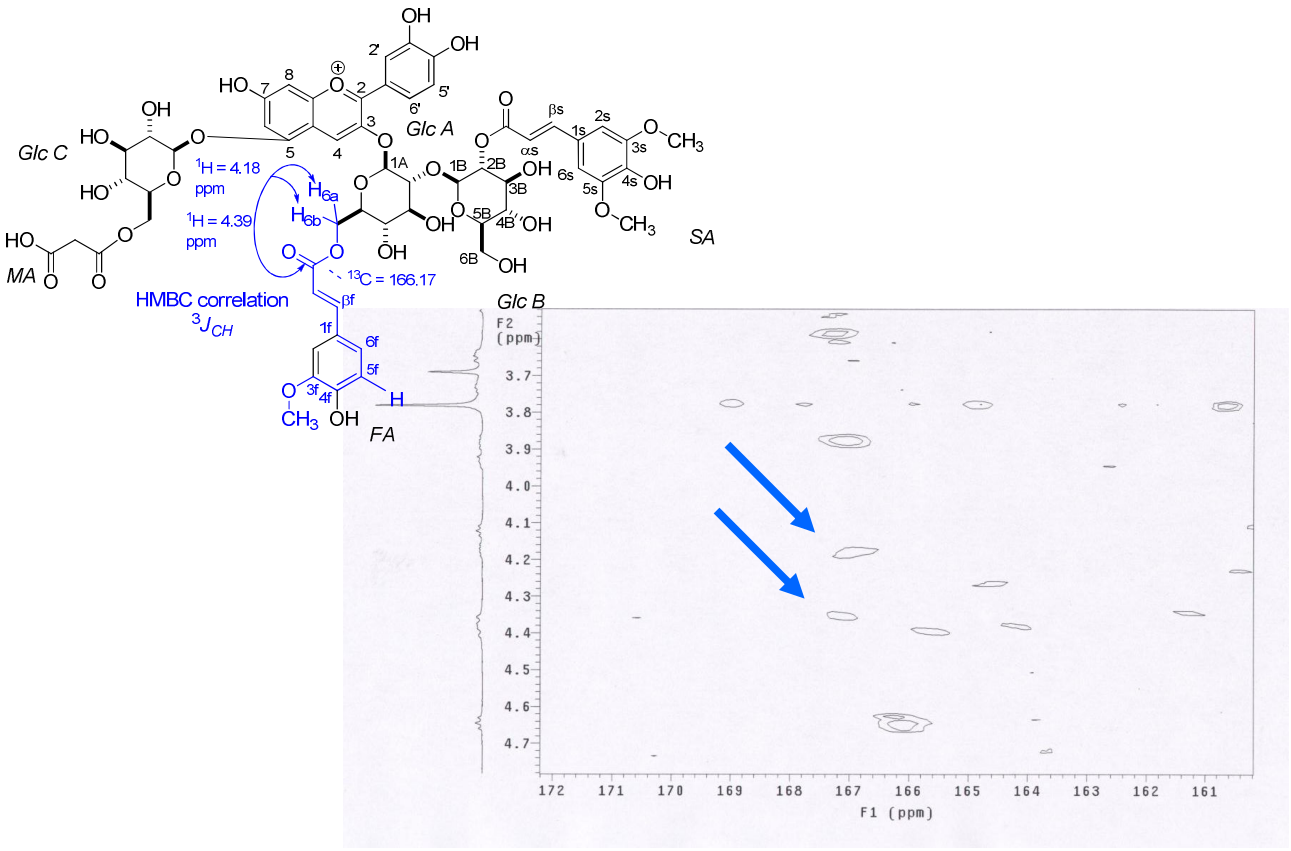



**Figure S7.** HMBC correlations of $H_6$ protons of Glc A and ferulic carbonyl for compound **5**.

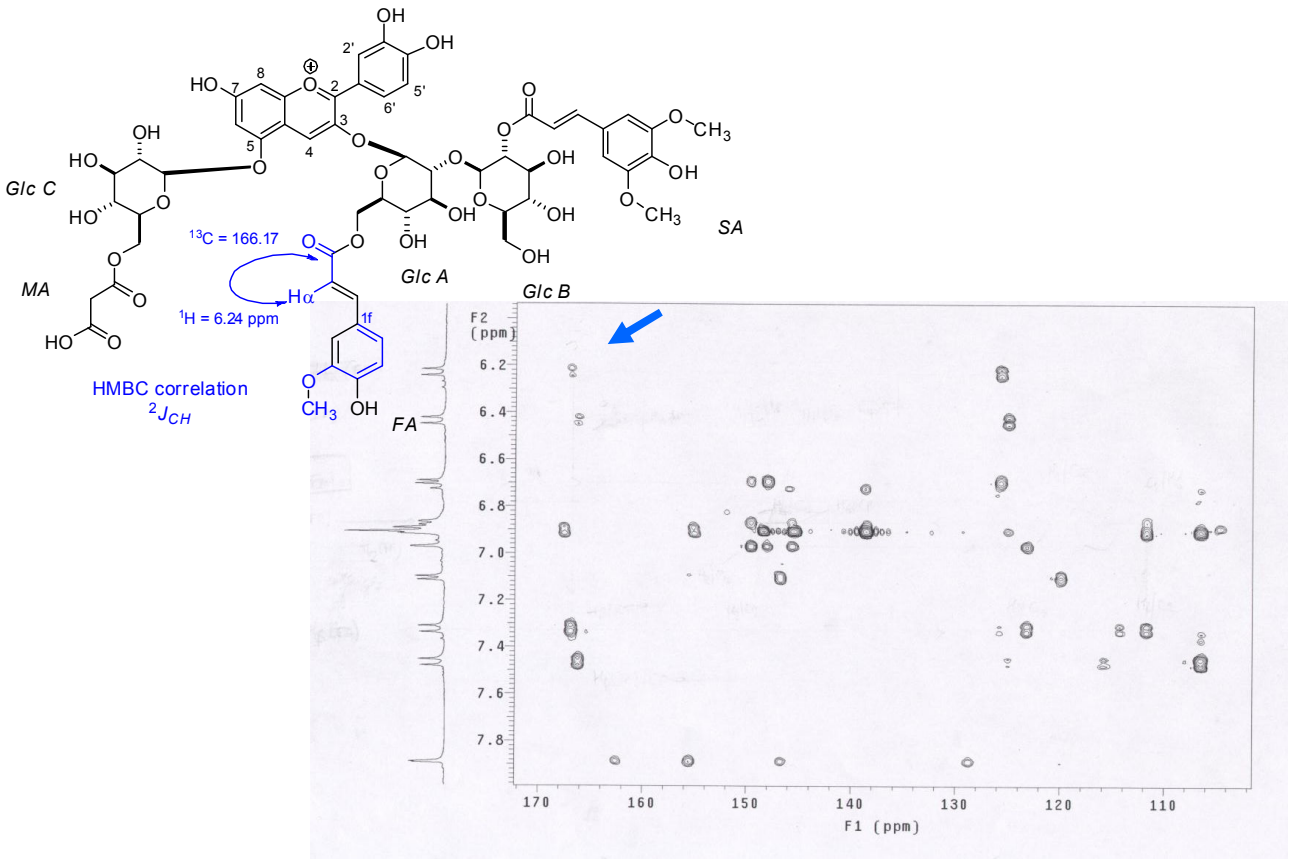



**Figure S8.** HMBC correlations of ferulic moiety for compound **5**.

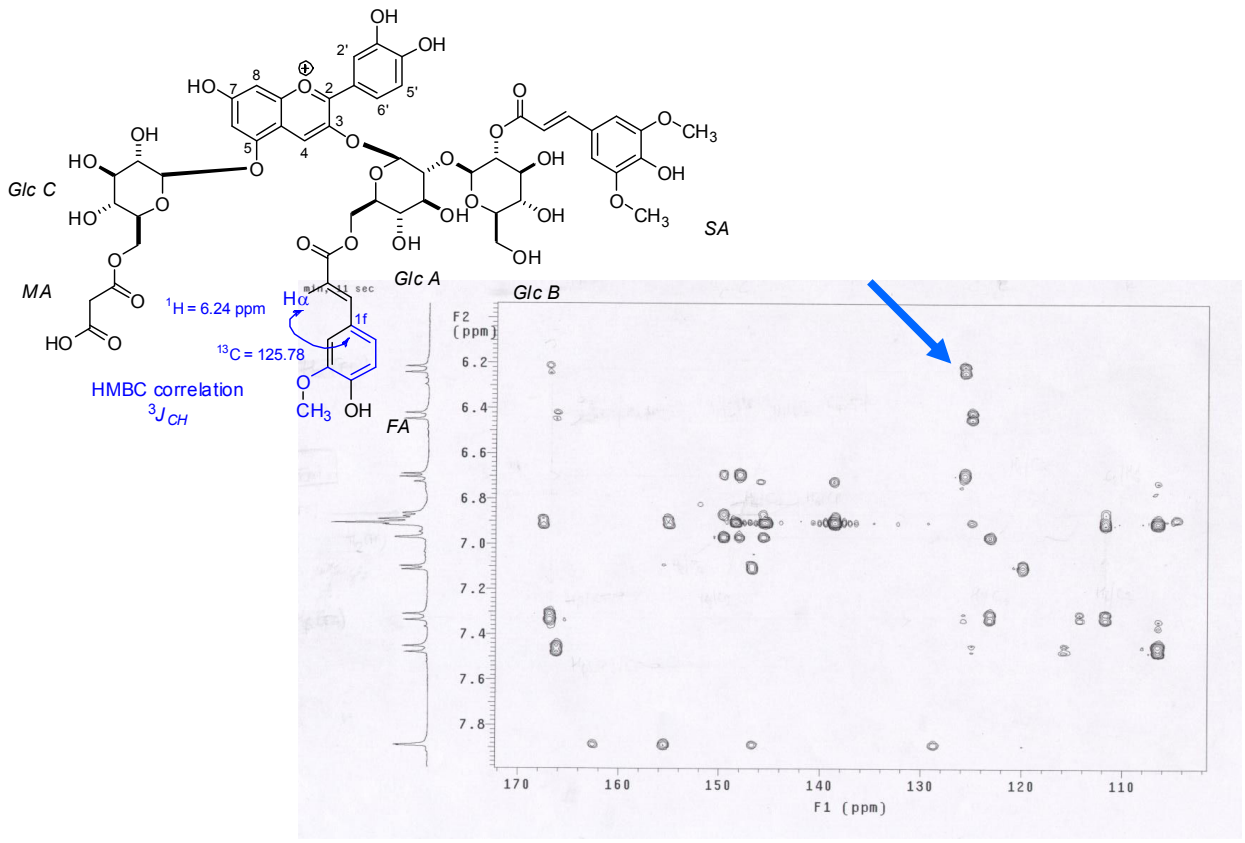



**Figure S9.** HMBC correlations of ferulic moiety for compound **5**.

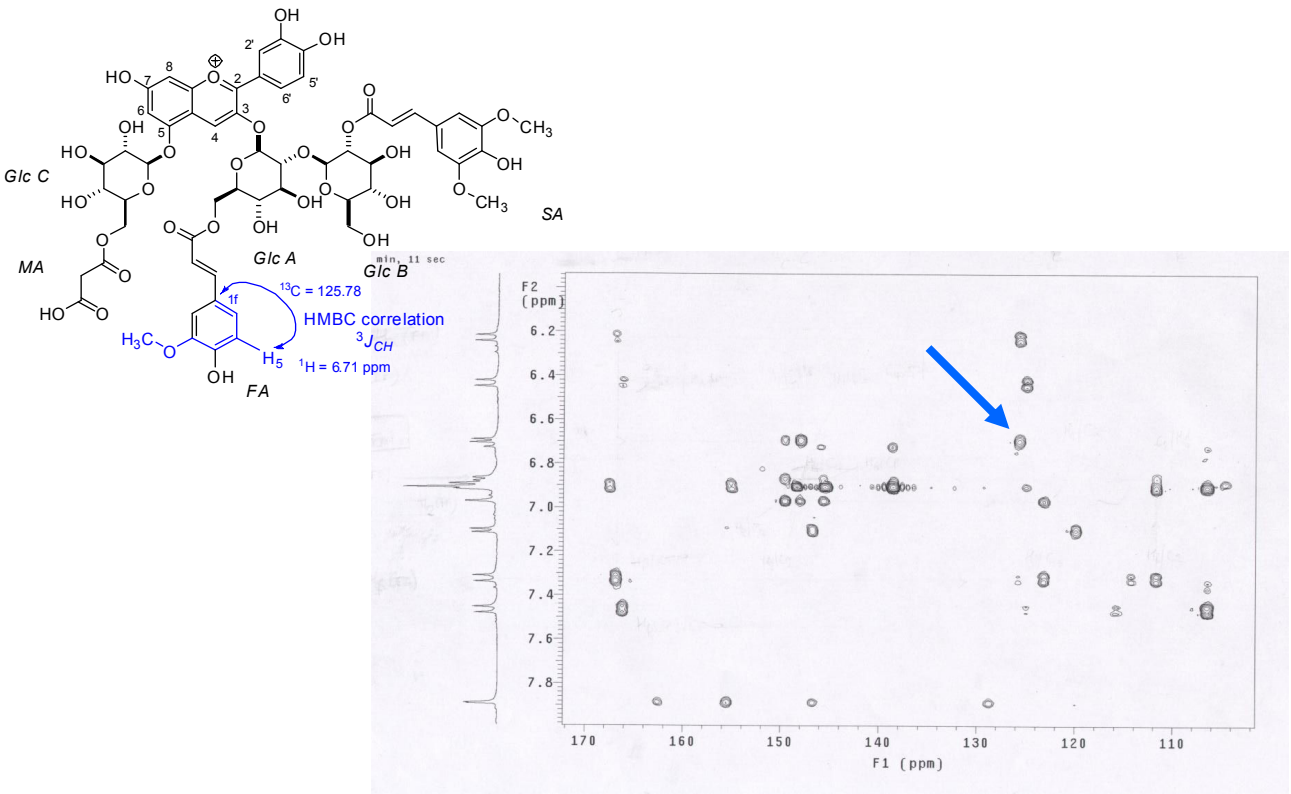



**Figure S10.** HMBC correlations of ferulic moiety for compound **5**.

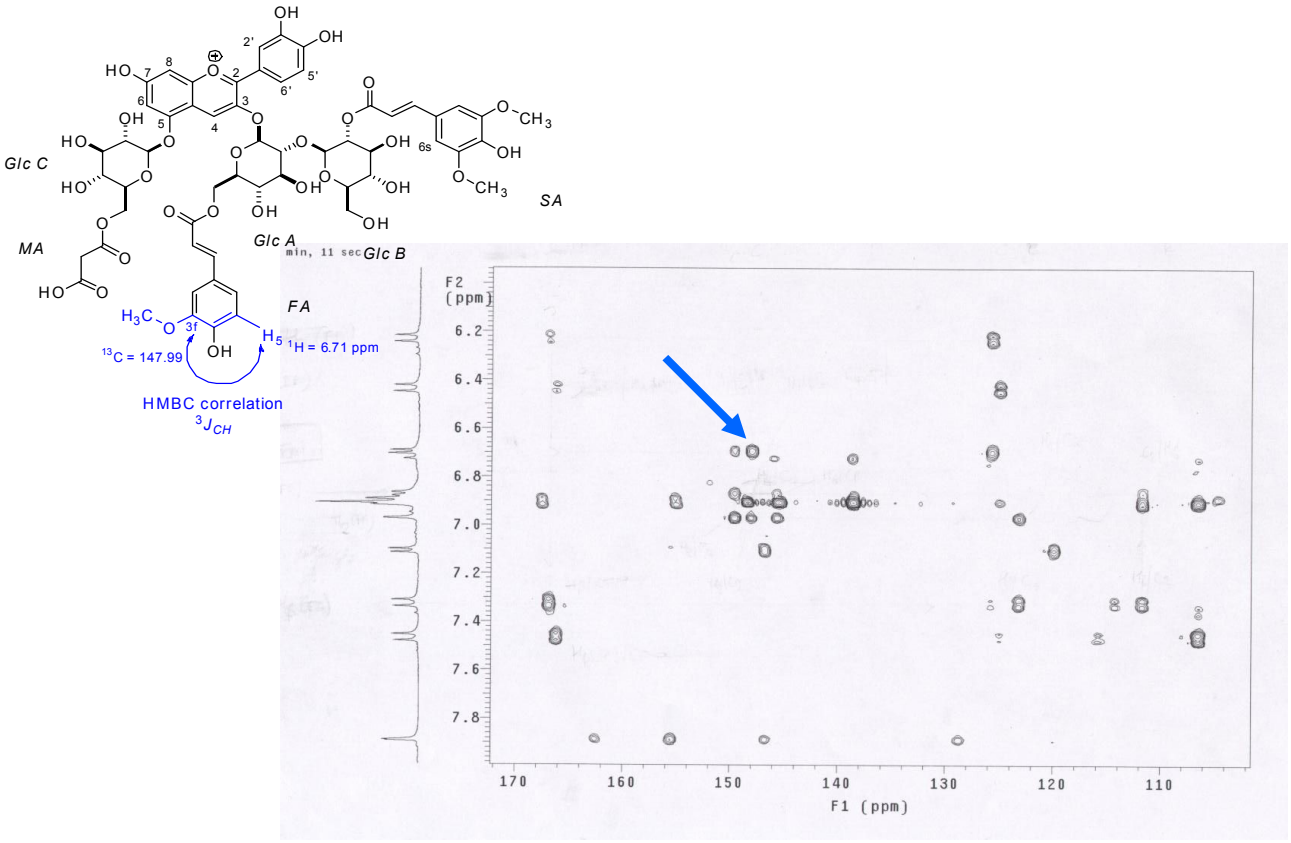



**Figure S11.** HMBC correlations of ferulic moiety for compound **5**.

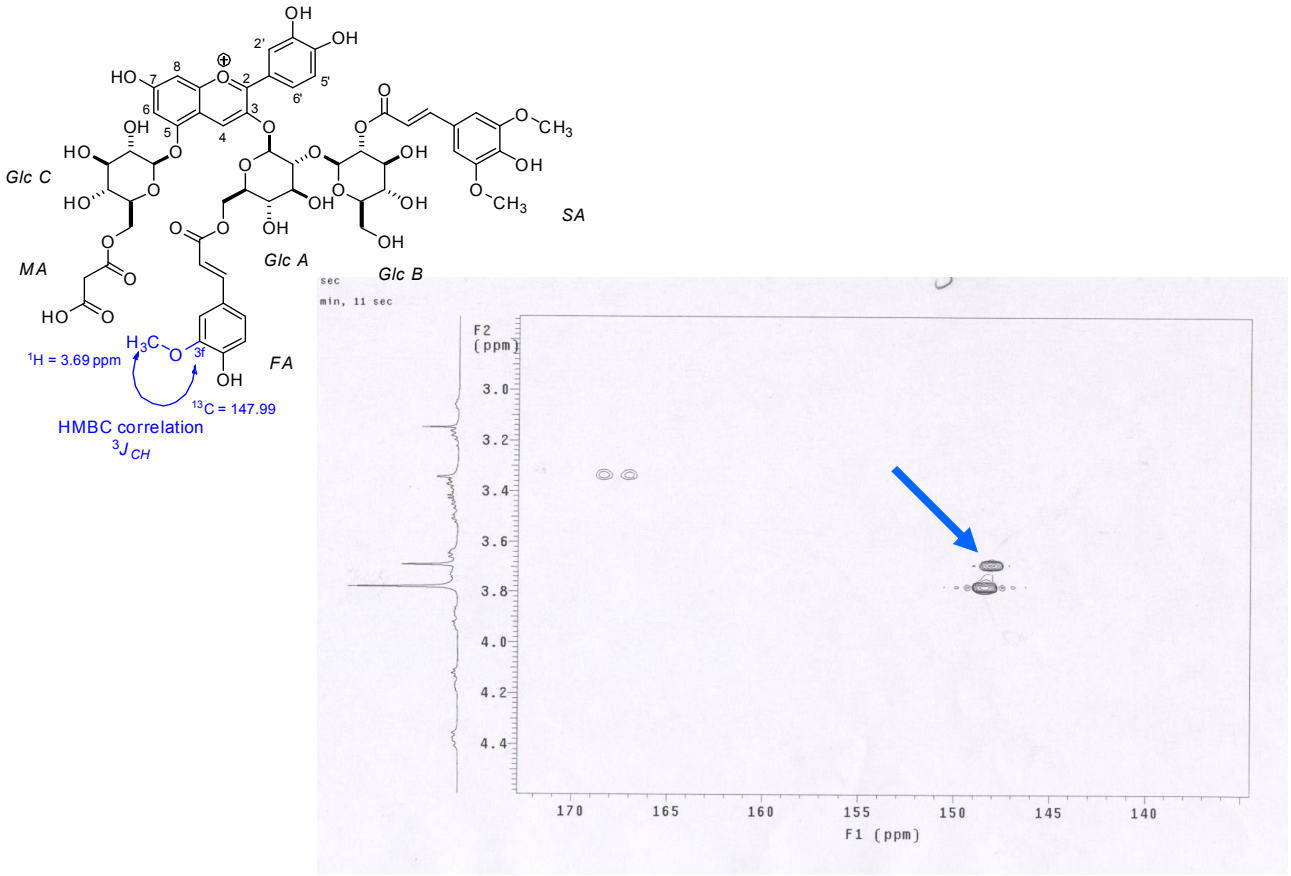



**Figure S12.** HMBC correlations of ferulic moiety for compound **5**.

The known anthocyanin **6** ($C_{44}H_{51}O_{25}$) (Figure 1) shows a percentage ratio of the UV-Vis absorbance at the maximum of the acyl unit over the absorbance at the visible maximum ($E_{acyl}/E_{vis}$) of 58%, suggesting an anthocyanin bioside monoacylated with one aromatic residue. The ESI(+)MS analysis showed *m/z* 979 as molecular ion. Fragmentation of *m/z* 979 produce *m/z* 817 ion by loss of a glycosyl residue (162 amu) and *m/z* 449 by loss of a diglycosyl-sinapoyl residue (530 = 324+206 amu). Fragmentation of both *m/z* 817 and *m/z* 449 ions in $MS^3$ spectrum give *m/z* 287 ion corresponding to the cyanidin aglycone, by losing the already mentioned 530 and 162 amu residues, respectively (Table 1). The anthocyanin **6** was preliminarily identified as cyanidin-3-sinapoyl-diglucoside-5-glucoside and was subjected to NMR characterization for assign the connectivity of glycosyl and acyl moieties by $^1H$ and COSY (Table 2). The $^1H$ spectrum of anthocyanin **1** was similar to the already described compound **1** except for the signals of $H_6$ of Glc B ($\delta$ 3.40) that were shifted to higher field than the ones belonging to the analogue **1** ($\delta$ 4.25 and 4.32) indicating that the 6-OH of Glc B was not acylated. On the other hand, the single aromatic residue (*i.e.* sinapic acid, $R_3$) is linked at 2-OH position of Glc B as revealed by his downfield signal ($\delta$ 4.60) similarly to the analogue **1**. No $^{13}C$-NMR signals were obtained due to the exiguity of material neither from HSQC nor HMBC experiments, however the complete set of $^1H$ signals of the known compound **6** is consistent with $^1H$-NMR published data of the identical compound obtained by partial hydrolysis of anthocyanins from *Iberis umbellata*[6] and is recognized as cyanidin 3-*O*-[2-*O*-(2-*O*-(*E*)-sinapoyl-β-D-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(β-D-glucopyranoside).

The known anthocyanin **7** ($C_{54}H_{59}O_{28}$) (Figure 1) shows a high percentage ratio of the UV-Vis absorbance ($E_{acyl}/E_{vis}$ = 128%) indicating an anthocyanin bioside biacylated with two aromatic residues. The ESI(+)MS analysis showed *m/z* 1155 as molecular ion. $MS^2$ processes of *m/z* 1155 revealed the same fragmentation pattern obtained for compound **1**, *m/z* 993 and *m/z* 449 ions were derived by loss of a glycosyl residue (162 amu) and the already mentioned 706 amu residue (diglycosyl-feruloylsinapoyl), respectively. The $MS^3$ residue fragmentation of both *m/z* 993 and *m/z* 449 ions in $MS^3$ spectrum produced the cyanidin aglycone fragment *m/z* 287 (Table 1). The anthocyanin **7** was tentatively identified cyanidin-3-feruloyl-sinapoyl-diglucoside-5-glucoside and was subjected to full NMR characterization for assign the exact connectivity of glycosyl and acyl moieties. The olefinic protons of ferulic acid ($R_1$) ($\delta$ 6.24 and 7.33) and sinapic acid ($R_3$) (δ 6.46 and 7.47) respectively, were assigned by $^1H$-NMR (Table 2) and homonuclear correlation experiments (COSY) with their high coupling constant values ($J$ = 15-16 Hz) indicated *trans*

[6] Saito, N., Tatsuzawa, F., Suenaga, E., Toki, K., Shinoda, K., Shigihara, A., & Honda, T. *Phytochemistry*, **2008**, *69*, 3139-3150.

configuration. Data obtained by HMBC and COSY allowed to assign the 3-$OCH_3$ ($\delta$ 3.74) group of ferulic acid which correlated through $^3J_{CH}$ at ester carbonyl ($\delta$ 166.9). Crossing data from HMBC, HSQC (Table 3) and COSY allowed to distinguish all signals related to ferulic and sinapic acid. $^1$H signals of acetalic protons were identified by comparing Literature[7] and previous data: the most deshielded acetalic proton ($\delta$ 5.66) is referred to Glc A, then signals at $\delta$ 5.14 and 5.06 ppm belong to Glc B and Glc C, respectively, with typical coupling constant ($J$ = 7-8 Hz) of β-*O*-glicosidic linkage. Among three signals of methylene protons in HSQC belonging to $C_6$ position of each glucose, the most deshielded ones are assigned to geminal $H_{6a}$ and $H_{6b}$ of Glc A ($\delta$ 4.23 and 4.34) indicating this position is acylated with ferulic acid, as confirmed also by HMBC correlation with ester carbonyl ($\delta$ 166.9) of ferulic moiety. As revealed by $^1$H and COSY analysis, the signal of $H_2$ of Glc B is shifted at lower field in analogy to similar compound **1** even though the evidence of HMBC correlation between $H_2$ of Glc B ($\delta$ 4.67) with ester carbonyl ($\delta$ 166.2) of sinapic acid is lacking due to the limited amount of available compound. The evident HMBC correlation between $H_2$ of Glc A ($\delta$ 4.12) with $C_1$ of Glc B ($\delta$ 99.88) confirmed the typical connectivity of β-1,2 glicosidic linkage of sophorose. Literature data revealed that sophorose is always linked to C3 position of cyanidin even though very weak HMBC correlation between the most deshielded acetalic signal belong to $H_1$ of Glc A ($\delta$ 5.66) and cyanidin C3 ($\delta$ 144.7) were observed. On the other hand, HMBC correlation between $H_1$ of Glc C ($\delta$ 5.06) and cyanidin C5 ($\delta$ 155.2) indicated that Glc C is linked at 5-position of cyanidin. However, spectroscopic data were consistent with $^1$H published data of the identical anthocyanin isolated from *Brassica oleracea* by Idaka et al.[8] Definitely, known anthocyanin **7** was identified as (cyanidin 3-*O*-[6-*O*-(*E*)-feruloyl-2-*O*-(2-*O*-(*E*)-sinapoyl-β-d-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(β-D-glucopyranoside).

The known anthocyanin **8** ($C_{46}H_{51}O_{27}$) (Figure 1) shows a medium percentage ratio of the UV-Vis absorbance ($E_{acyl}/E_{vis}$ = 53 %) indicating an anthocyanin bioside monoacylated with one cinnamic acid. The ESI(+)MS analysis showed *m/z* 1035 as molecular ion. $MS^2$ fragmentations of *m/z* 1035 revealed two main ions (*m/z* 787 and *m/z* 535) originating by loss of malonoyl-glucoside (248 amu) and the feruloyl-diglycoside (500 amu), respectively. $MS^3$ fragmentation of each ions produced the cyanidin aglycone fragment *m/z* 287 (Table 1). The anthocyanin **8** was presumed to be based on cyanidin-3-feruloyl-diglucoside-5-malonoyl-glucoside and was analyzed by NMR characterization to assign the exact connectivity of glycosyl and acyl moieties using $^1$H, COSY, HSQC and HMBC

[7] Otsuki, T., Matsufuji, H., Takeda, M., Toyoda, M., & Goda, Y. *Phytochemistry*, **2002,** *60*, 79-87
[8] Idaka, E., Yamakita, H., Ogawa, T., Kondo, T., Yamamoto, M., Goto, T., *Chem. Lett.* **1987**, 1213-1216.

experiments (Table 2 and 3). Three signals corresponding to the chemical shifts of acetalic groups formed by the anomeric carbons were detected and assigned to different sugars using previous reported data: the acetalic proton's signal shifted at the lowest field ($\delta$ 5.60) belong to the $H_1$ of Glc A proton, whereas signals at higher field belong to Glc B ($\delta$ 4.61), and Glc C ($\delta$ 5.10), respectively. The exact connectivity pattern of sugar residues was obtained by $^1$H and COSY experiments. The typical diaxial protons coupling constants ($J$ = 7.2-8.4 Hz) of anomeric protons of sugar residues assigned to β-glucopyranoside, as shown in $^1$H-NMR experiments, together with the HMBC correlation between $H_2$ of Glc A ($\delta$ 4.06) with $C_1$ of Glc B ($\delta$ 99.1) confirmed the typical connectivity of β-1,2 glicosidic linkage of sophorose which is linked to $C_3$ position of cyanidin as confirmed by HMBC correlation between the most deshielded acetalic signal belong to $H_1$ of Glc A ($\delta$ 5.61) and cyanidin C3 ($\delta$ 144.9). Glc C is linked at 5-position of cyanidin as revealed by HMBC correlation between $H_1$ of Glc C ($\delta$ 5.10) and cyanidin C5 ($\delta$ 155.1). Three couples of methylene protons belong to $C_6$ positions of each glucose as confirmed by HSQC correlations together with an additional cross peak indicating methylene of malonic acid ($\delta$ 3.33). The most deshielded methylene signals belong to Glc A and Glc C indicating these positions are acylated with ferulic ($R_1$) and malonic acid ($R_4$), respectively, as confirmed also by HMBC correlation between ferulic ester carbonyl ($\delta$ 166.9) and two protons $H_{6a}$ e $H_{6b}$ of Glc A ($\delta$ 4.28 and 4.44) together with correlation between malonic ester ($R_4$) carbonyl ($\delta$ 167.0) and two protons $H_{6a}$ e $H_{6b}$ of Glc C ($\delta$ 3.88 and 4.35). The ferulic residue configuration was elucidated as *trans*, based on the large coupling constants ($J$ = 16.0 Hz). Olefinic ferulic protons at α and β position ($\delta$ 6.25 and 7.34, respectively) correlated via HMBC with proton signal of 3-$OCH_3$ ($\delta$ 3.68) and the $C_3$ signal ($\delta$ 148.0) of the ferulic acid. Definitely, spectroscopic data were consistent with published data of the identical anthocyanin isolated from *R. sativus* var. *Benikanmi* by Tatsuzawa et al.,[9] therefore the known anthocyanin **8** was identified as cyanidin 3-*O*-[2-*O*-(6-*O*-(*E*)-feruloyl-β-D-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(6-*O*-malonoyl-β-D-glucopyranoside).

The ESI(+)MS analysis of the known anthocyanin **9** ($C_{56}H_{59}O_{30}$) (Figure 1) showed *m/z* 1211 as molecular ion. The percentage ratio of the UV-Vis absorbance at the maximum of the acyl unit over the absorbance at the visible maximum ($E_{acyl}/E_{vis}$) was 110 % suggesting an anthocyanin bioside biacylated with two aromatic residues. $MS^2$ fragmentation of *m/z* 1211 caused three main ions at *m/z* 1167, *m/z* 963 and *m/z* 535 by loss of carbon dioxide (44 amu), a malonoyl-glycosyl residue (248 amu) and a diglycosyl-diferuloyl residue (676 = 324+352 amu), respectively. Fragmentation of

[9] Tatsuzawa, F., Saito, N., Toki, K., Shinoda, K., Shigihara, A. and Honda T. *J. Japan. Soc. Hort. Sci.*, **2010**, *79*, 103-107.

*m/z* 963 and 535 ions in $MS^3$ spectrum give the fragments *m/z* 287 corresponding to the cyanidin aglycone, by losing the already mentioned 676 or 248 amu residues (Table 1). The anthocyanin **9** was presumed to be based on cyanidin-3-diferuloyl-diglucoside-5-malonoyl-glucoside and was subjected to full NMR characterization for assign the exact connectivity of glycosyl and acyl moieties using $^1H$, COSY, HSQC and HMBC experiments (Tables 2 and 3). Three signals corresponding to the chemical shifts of acetalic groups formed by the anomeric carbons were detected and assigned to different sugars using previous reported data for similar anthocyanins. The most downfield proton's signal at $H_1$ of Glc A ($\delta$ 5.62), followed by those of Glc B ($\delta$ 4.78), and Glc C ($\delta$ 5.06). The typical diaxial protons coupling constants ($J$ = 7.3-7.9 Hz) of anomeric protons of sugar residues assigned to β-glucopyranoside, as shown in $^1H$-NMR experiments. Among three couples of methylene proton found in HSQC spectrum, the downfield signals ($\delta$ 4.20 and 4.41) and ($\delta$ 3.89 and 4.01) belong to $H_{6a}$ and $H_{6b}$ positions of Glc A and Glc B respectively, which are acylated with two ferulic acids ($R_1$ and $R_2$). The methylene protons downfield shifted refer to $H_{6a}$ and $H_{6b}$ of Glc C ($\delta$ 3.96, 4.35) acylated with malonic acid ($R_4$) to whom belong the last methylene proton signal shifted at higher field ($\delta$ 3.33). The ferulic residue configuration was elucidated as *trans*, based on the large coupling constants ($J$ = 15-16 Hz). Due to the exiguity of material analyzed, no HMBC correlation were observed between ferulic ester carbonyls and two protons $H_{6a}$ e $H_{6b}$ of Glc A and Glc B even though the shifting at downfield of $H_6$ proton confirm of the suggested connectivity pattern also in accordance with the characterization previous compounds. No $^{13}C$ signals are given for compound **9**. As already described, malonic acylation occurs at $C_6$ position of Glc C as confirmed by consistent data of published $^1H$-NMR (not complete) characterization of the identical anthocyanin isolated from *Ajuga pyramidalis* cell cultures,[10] therefore the known anthocyanin **9** was identified cyanidin 3-*O*-[6-*O*-(*E*)-feruloyl-2-*O*-(6-*O*-(*E*)-feruloyl-β-d-glucopyranosyl)-β-D-glucopyranoside]-5-*O*-(6-*O*-malonoyl-β-D-glucopyranoside).

[10] Madhavi, D.L., Juthangkoon, S., Lewen, K., Berber-Jimenez, M.D. and Smith M.A.L. *J. Agric. Food Chem.* **1996**, *44*,1170-1176.

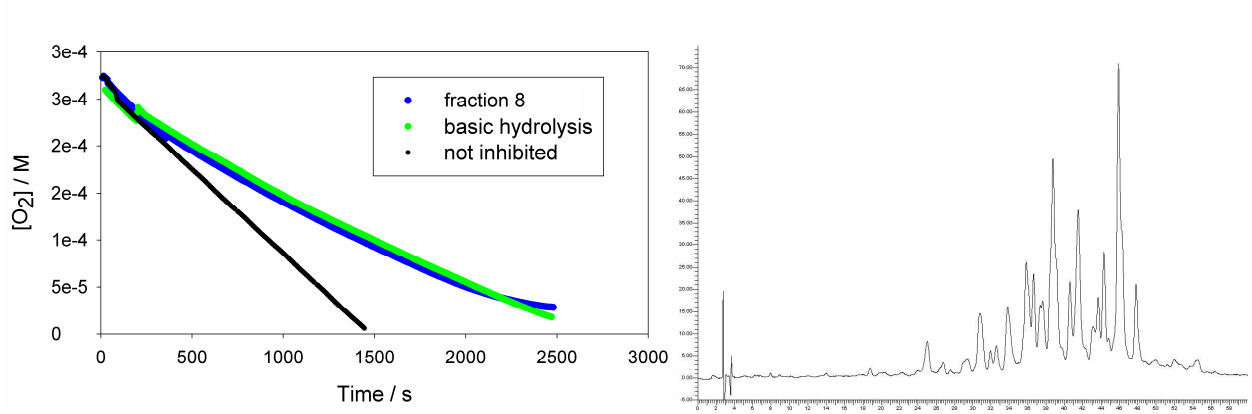



**Figure S13.** (left) Oxygen consumption during the autoxidation of 30 mM linoleic acid in aqueous Triton-X100 (0.16M) micelles ar 37°C, initiated by 8.3 mM AAPH, in the presence of 1% formic acid, with or without 4 μM antioxidant. Antioxidant was whole fraction 8 (average MW 1200 amu) or fraction 8 after basic hydrolysis. (right) HPLC-UV profile of fraction 8 recorded on Polar-RP column as described in the manuscript (section 2.1).

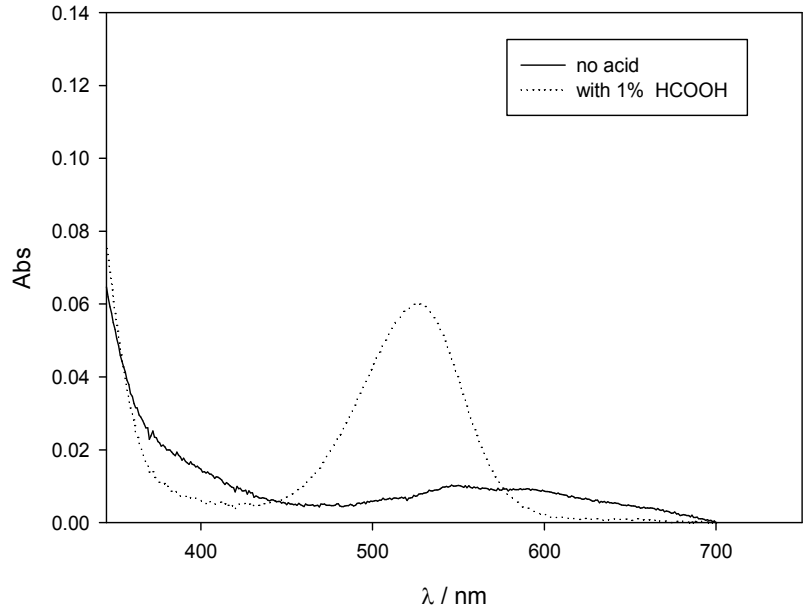



**Figure S14.** UV-vis spectrum of anthocyanin **5** in the same solvent mixture used to perform the autoxidation studies (acetonitrile containing 2.1 M styrene) in the absence or presence of 1% v/v formic acid.