Supporting Information

Combinatorial Mutasynthesis of Flavonoid Analogues from Acrylic Acids in Microorganisms

Joseph A. Chemler, Yajun Yan, Effendi Leonard, Mattheos A. G. Koffas[†]

State University of New York at Buffalo, Department of Chemical and Biological Engineering, 303 Furnas Hall, North Campus, Buffalo, NY 14260
[†]email: mkoffas@eng.buffalo.edu

General. All commercial reagents were used as provided. The acrylic acids were dissolved in DMSO to a concentration of 0.5 M. ^{1}H NMR was performed on a Varian Mercury 400 MHz spectrometer. Proton chemical shifts were reported in ppm using TMS or residual acetone (2.05 ppm) as internal standards. HPLC was performed on an Agilent 1100 series instrument with a diode array detector and fraction collector. Analytical analysis was performed using a reverse phase ZORBAX SB-C18 column (4.6×150mm, Agilent Technologies) and purification fractions were collected when using a semi-preparative ZORBAX C18 column (9.6 x 250 mm, Agilent Technologies) maintained at 25°C. HPLC methods used linear gradients. Low and high resolution electron-spray ionization mass spectra were obtained at the Instrument Center at the University at Buffalo Department of Chemistry.

DNA manipulations. All DNA manipulations were performed by using standard procedures.[1] *Escherichia coli* strain TOP10 F' (Invitrogen, Carlsbad, CA) was used for cloning purposes and expression of protein for *in vitro* assays. Plasmids pYES2.1/V5-His-TOPO (Invitrogen) and pTrcHis2-TOPO (Invitrogen) were used for cloning purposes. Yeast strain INScV1 was used for fermentations. Construction of plasmid Yc4cc-181 (leucine selective marker) containing 4-comaroyl:CoA ligase (4CL) from parsley, and chalcone synthase (CHS) and chalcone isomerase (CHI) was previously described.[2] pYES2.1-MdFHT (uracil selective marker) was constructed in the same fashion as pTrcHis2-MdFHT[3] but pYES2.1/V5-His-TOPO was used as the T/A cloning vector. Vectors pTrcHis2-4CL, pTrcHis2-CHS, pTrcHis2-CHI-A were cloned into pTrcHis2-TOPO using primers previously described.[2] *E. coli* colonies were screened to find colonies containing plasmids with the inserted gene with the correct orientation for protein expression by restriction enzyme digests and gel electrophoresis. The cDNA of MATB was obtained by performing RT-PCR on mRNA isolated from *Rhizobium trifolii* using primers based on the DNA sequence available in the GenBank database (AAC83455). A QIAGEN RNeasy MiniKit was used for isolation of total RNA from the plant. The MATB gene was cloned into the *E. coli* expression vector pTrcHis2-TOPO to give pTrcHis2-MATB. Transformation of yeast was carried out according to manufacturer instructions. In all cases, the absence of undesired mutations introduced during PCR was verified by direct nucleotide sequencing.

Screening library. *p*-coumaric acid **(1a)**, *m*-coumaric acid **(1b)**, *o*-coumaric acid **(1c)**, *p*-fluorocinnamic acid **(1d)**, *o*-fluorocinnamic acid **(1e)**, *p*-aminocinnamic acid **(1f)**, trans,*trans*-muconic acid **(1g)**, ferulic acid (**1h**), 3,4-difluorocinnamic acid (**1i**), 3,4,5-trifluorocinnamic acid (**1j**), *p*-chlorocinnamic acid (**1k**), *o*-chlorocinnamic acid (**1l**), 3-(3-furyl)acrylic acid (**1m**), and 3-(3-thienyl)acrylic acid (**1n**).

4CL in vitro assay. An overnight Luria-Bertani broth preinoculum of *E. coli* strain TOP10 F' carry ampicillin resistant plasmid pTrcHis2-4CL, was added to 50 mL of fresh LB broth with 50 μg/mL ampicillin and placed in a horizontal shaker maintained at 37°C and 300 rpm. IPTG was added to a final concentration of 1mM when the OD_{600}0 reached 0.6. The culture was incubated for 4 hours in a horizontal shaker maintained at 30°C and 300 rpm. The culture was then cooled on ice and cells were harvested using centrifugation. The cell pellet was washed once and repelleted in 25 mL of ice-cold 0.1 M Tris-HCl, pH 7.8, with 20 mM $MgCl_2$. The cells were resuspended in 2 mL of fresh buffer and sonicated in 15 second intervals for up to 90 seconds at 50% amplitude. The lysed cells were removed by centrifugation and the protein concentration of the soluble protein fraction was determined using the bicinchoninic acid assay (Pierce Chemicals).

The final incubation mixture contained 0.25mM of the appropriate cinnamic acid analog, 20 mM of ATP (adjusted to pH 7-8 with 5N NaOH), 20 mM of $MgSO_4$, 0.1 M Tris-HCl, pH 7.8 and 0.5 mg/mL total protein in a total volume of 250 μL. The reaction was started with the addition of trilithium CoA to final concentration of 1.0 mM and carried out at 30°C. Control experiments contained no added CoA. A BioRad spectrometer was used to measure the change in absorption between 250–400 nm in 2 nm steps at 90 second intervals for 5 minutes.

4CL, CHS, CHI and MATB *in vitro* assay. *E. coli* strain TOP10F' was transformed separately with pTrcHis2-MATB, pTrcHis2-4CL, pTrcHis2-CHS and pTrcHis2-CHI. Each strain was induced, harvested and lysed as previously described. The final incubation mixture contained 0.25 mM of the appropriate cinnamic acid analog, 20mM of ATP (adjusted to pH 7-8 with 5 N NaOH), 20 mM of $MgSO_4$, 0.1 M Tris-HCl, pH 7.8, 1.0 mM trilitium coenzyme A, 1.0 mM sodium malonate and 0.5 mg/mL total protein from each recombinant *E. coli* strain in a total volume of 250 μL. Control incubation was carried out without total protein from the strain expressing CHS. Reactions were incubated at 30°C overnight and 75 μL of mixture was injected directly into the HPLC for analysis.

Yeast strains. *Saccharomyces cerevisiae* strain INScV1 Yc4cc-181 was previous constructed.[3] Yeast strain INScV1 Yc4cc-181 + pYES2.1-MdFHT was created by transforming INScV1 Yc4cc-181 with plasmid pYES2.1-MdFHT and then plating the transformation reaction onto SC minimal media plates deficient in leucine and uracil. Ten colonies were randomly selected and screened to find the colony with the highest conversion of *p*-coumaric acid into dihydrokaempferol.

Unnatural flavonoid production from *Saccharomyces cerevisiae* fermentations. A 50 mL culture of 2% glucose SC selective media (either deficient in leucine or leucine and uracil) was inoculated with either yeast strain INScV1 Yc4cc-181 or INScV1 Yc4cc-181 + pYES2.1-MdFHT and allowed to grow overnight at 30ºC with horizontal shaking. The following day, the supernatant was removed after centrifugation and resuspended in 2 mL of sterile water. To 200 mL of fresh 2% galactose SC selective media, enough of the preinoculum resuspension was added to reach an OD_{600} of 0.4. After 4 hours of incubation at 30ºC with horizontal shaking, aliquots of 10 mL were added to 250 mL flasks. Compounds **1a,b,c,f,g** were separately added to the yeast culture aliquots, in triplicates, to a final concentration of 1.0 mM. Compounds **1d,e** were separately added to the yeast culture aliquots, in triplicates, to a final concentration of 0.5 mM. The flasks were incubated at 30ºC with horizontal shaking over several days. Samples of 100 µL were withdrawn from each flask every 24 hours, centrifuged to remove cells and 25 µL of the supernatant was injected directly into the HPLC for analysis.

Large scale production followed the same procedure as above but the culture volumes were increase to 400 mL in 2 L flasks. After 2-3 days, cells were removed from the cultures by centrifugation and the supernatant was extract 2 times with equal volumes of ethyl acetate or ether. Extracts were dried by rotary evaporation and the residue was redissolved in a minimal volume of ethyl acetate. Samples were cleaned up using silica gel chromatography with a mobile phase of 2:1 ethyl acetate:hexanes. Samples were loaded onto the column and the first 400 mL fraction was collected, dried and resuspended in a minimal volume of acetonitrile. Final purification was performed by collecting fractions using HPLC with a semi-preparative ZORBAX SB-C18 column (9.6 x 250 mL). The mobile phase consisted of 0.1% formic acid in water and acetonitrile under isocratic conditions. The flowrate and the water:acetonitrile ratio varied for each sample to obtain optimal separation. Fractions were pooled and dried in preparation for chemical characterization.

HPLC Conditions:
Solvent A: 0.1% Formic acid in acetonitrile, Solvent B: 0.1% Formic acid in water
[ZORBAX SB-C18 column (4.6×150mm, Agilent Technologies); 1mL/min; 25°C]

Table 1. HPLC methods

Method 1			Method 2			Method 3		
Time (min)	% A	% B	Time (min)	% A	% B	Time (min)	% A	% B
0.00	10	90	0.00	10	90	0.00	10	90
10.00	40	60	10.00	40	60	20.00	10	90
15.00	60	40	15.00	60	40			
17.00	10	90	20.00	60	40			
			22.00	10	90			

Table 2. Retention times

Compound	Method	Time (min)
1a	1	7.49
1b	1	8.41
1c	1	9.33
1d	2	12.61
1e	2	12.59
1f	1	4.86
1g	3	4.04
4a	1	12.79
4b	1	13.45
4c	1	14.29
4d	2	16.79
4e	2	16.85
4f	1	9.37
4g	3	13.28
5a	1	9.67
5b	1	10.41
5c	1	11.05
5d	2	13.96
5e	2	13.66
5f	1	6.43

HPLC data of 4CL, CHS, CHI and MATB *in vitro* assay. Control assays (blue line) contained the appropriate acrylic acid and enzymes 4CL, CHI and MATB. Experimental assays (red line) contained the appropriate acrylic acid and enzymes 4CL, CHS, CHI and MATB. All HPLC graphs are at 290 nm.

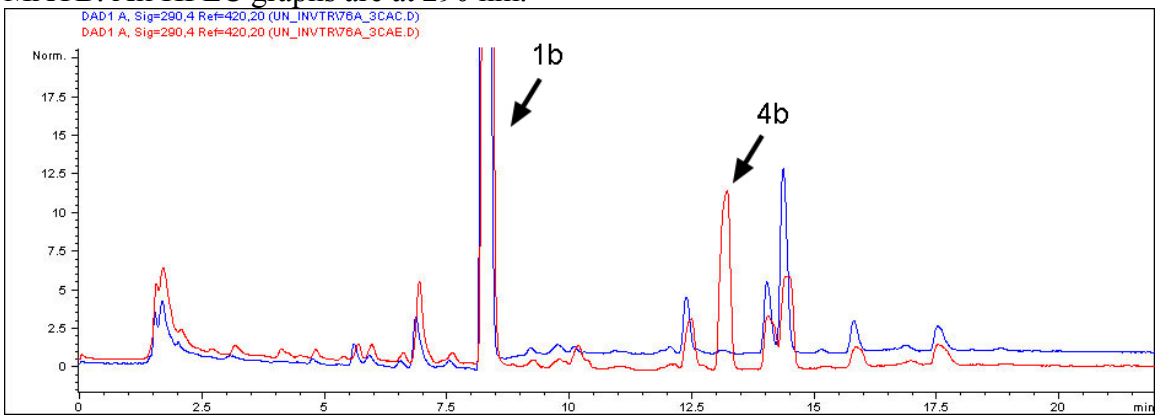

Figure 1. HPLC graph showing the *in vitro* enzymatic conversion of **1b** into **4b**.

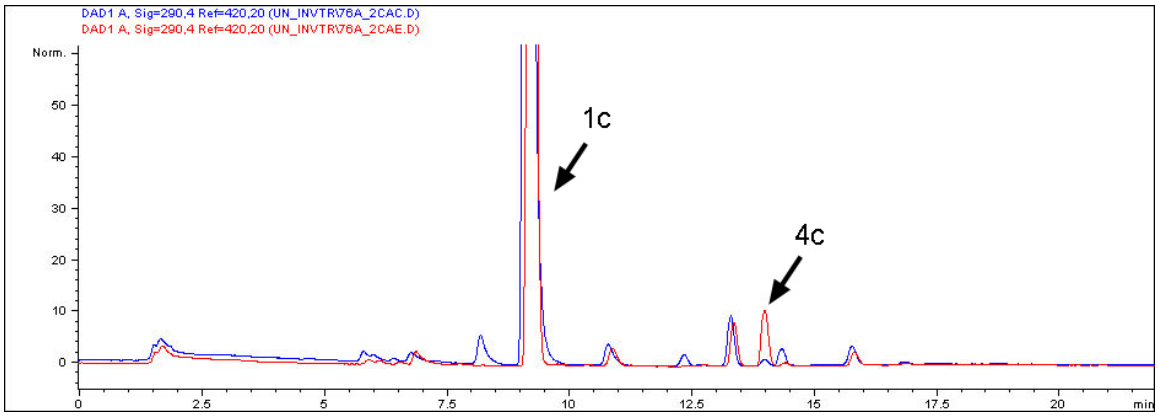

Figure 2. HPLC graph showing the *in vitro* enzymatic conversion of **1c** into **4c**.

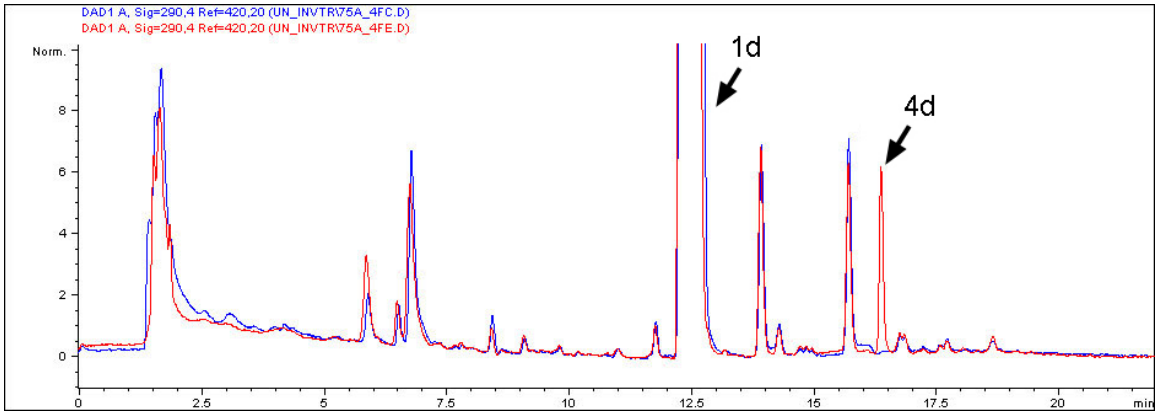

Figure 3. HPLC graph showing the *in vitro* enzymatic conversion of **1d** into **4d**.

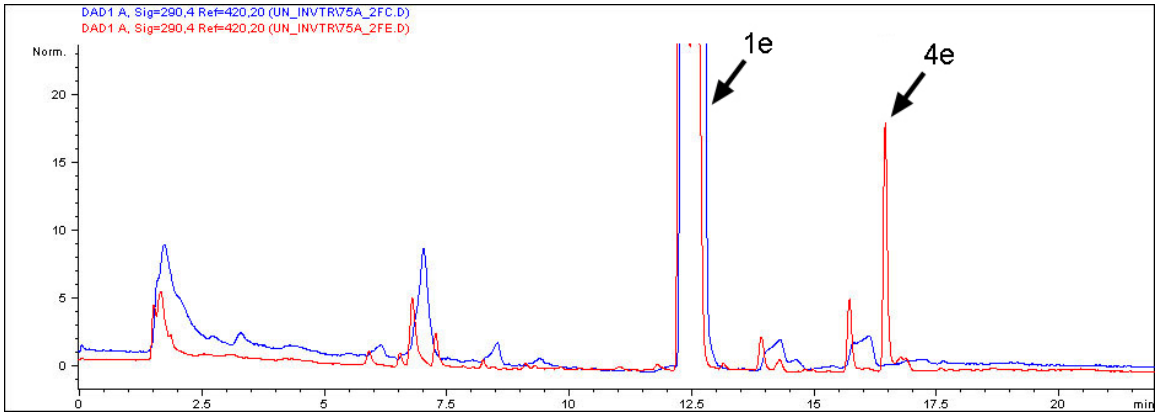

Figure 4. HPLC graph showing the *in vitro* enzymatic conversion of **1e** into **4e**.

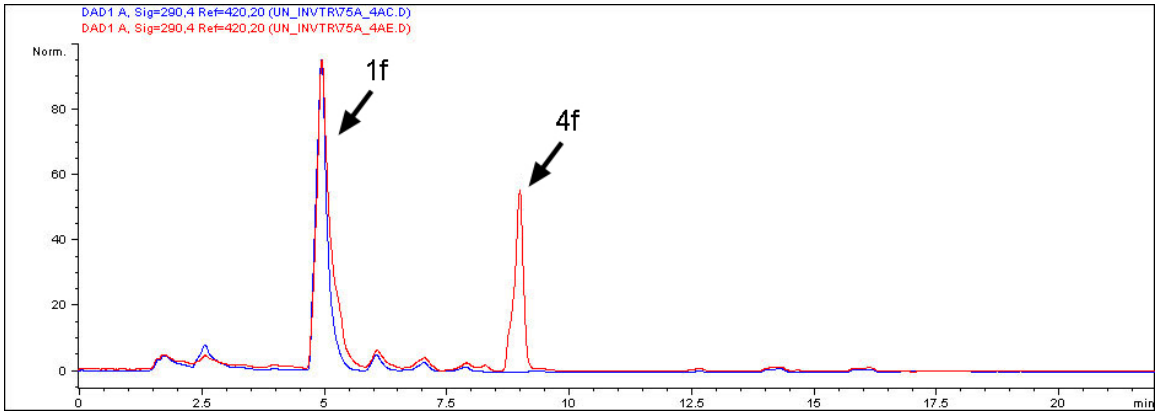

Figure 5. HPLC graph showing the *in vitro* enzymatic conversion of **1f** into **4f**.

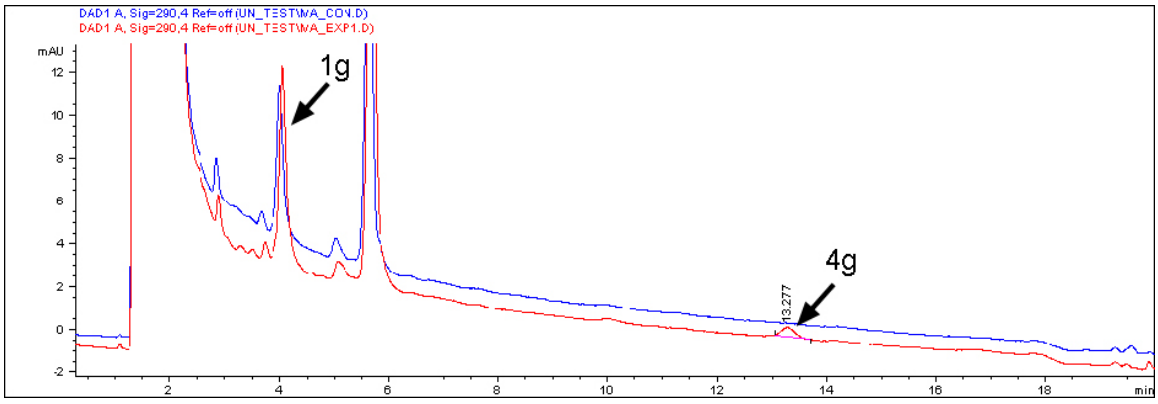

Figure 6. HPLC graph showing the *in vitro* enzymatic conversion of **1g** into **4g**.

^{1}H NMR, UV and ESIMS data.

(2*S*)-Naringenin (**4a**): λ_{max} nm: 289

(2*S*)-3',5,7-Trihydroxyflavanone (**4b**): ^{1}H NMR (400 MHz, d_6-acetone): δ 12.149 (1H, s, 5-OH), 8.007 (1H, s, 3'-OH), 7.251 (1H, t, *J* = 7.6 Hz, H-5'), 7.017 (1H, m, H-6'), 7.000 (1H, m, H-2'), 6.848 (1H, dd, *J* = 2.2, 8.2 Hz, H-4'), 5.993 (1H, d, *J* = 1.6 Hz, H-6), 5.954 (1H, d, *J* = 2.0 Hz, H-8), 5.495 (1H, dd, *J* = 3.0, 12.6 Hz, H-2), 3.115 (1H, dd, *J* = 12.4, 17.2 Hz, H-3-*trans*), 2.785 (1H, dd, *J* = 2.8, 17.2 Hz, H-3-*cis*). UV λ_{max} nm: 287. MS: ESI^- $(M-H)^-$ *m/z* = 271.3 (calc. 271.26), $HRESI^+$ $(M+H)^+$ *m/z* = 273.0759 (calc. 273.0758).

(2*S*)-2',5,7-Trihydroxyflavanone (**4c**): ^{1}H NMR (400 MHz, d_6-acetone): δ 12.217 (1H, s, 5-OH), 8.015 (1H, s, 2'-OH), 7.516 (1H, d, *J* = 7.6 Hz, H-6'), 7.189 (1H, dd, J = 5.0, 11.0, H-4'), 6.928 (1H, m, H-5'), 6.907 (1H, m, H-3'), 5.961 (1H, d, *J* = 1.6 Hz, H-6), 5.910 (1H, d, *J* = 2.4 Hz, H-8), 5.754 (1H, dd, *J* = 2.6, 13.0 Hz, H-2), 3.026 (1H, dd, *J* = 12.8, 16.8 Hz, H-3-*trans*), 2.783 (1H, dd, *J* = 3.0, 17.0 Hz, H-3-*cis*). UV λ_{max} nm: 287. MS: ESI^- $(M-H)^-$ *m/z* = 271.2 (calc. 271.26), $HRESI^+$ $(M+H)^+$ *m/z* = 273.0756 (calc. 273.0758).

(2*S*)-4'-Fluoro-5,7-dihydroxyflavanone (**4d**): ^{1}H NMR (400 MHz, d_6-acetone): δ 12.172 (1H, s, 5-OH), 7.607 (2H, dd, *J* = 5.4, 8.4 Hz, H-2',6'), 7.196 (2H, t, *J* = 8.8 Hz, H-3',5'), 5.931 (1H, d, *J* = 2.4 Hz, H-6), 5.900 (1H, d, *J* = 2.0 Hz, H-8), 5.516 (1H, dd, *J* = 2.8, 12.8 Hz, H-2), 3.089 (1H, dd, *J* = 12.8, 17.2 Hz, H-3-*trans*), 2.735 (1H, dd, *J* = 3.2, 17.2 Hz, H-3-*cis*). UV λ_{max} nm: 290. MS: ESI^- $(M-H)^-$ *m/z* = 273.4 (calc. 273.25), $HRESI^+$ $(M+H)^+$ *m/z* = 275.0710 (calc. 275.0714).

(2*S*)-2'-Fluoro-5,7-dihydroxyflavanone (**4e**): ^{1}H NMR (400 MHz, d_6-acetone): δ 12.143 (1H, s, 5-OH), 7.692 (1H, t, *J* = 7.2 Hz, H-6'), 7.460 (1H, dd, *J* = 5.8, 13.4 Hz, H-4'), 7.302 (1H, t, *J* = 7.4 Hz, H-5'), 7.208 (1H, t, *J* = 9.4 Hz, H-3'), 5.981 (1H, d, *J* = 2.0 Hz, H-6), 5.927 (1H, d, J = 2.0 Hz, H-8), 5.7933 (1H, dd, *J* = 2.6, 13.0 Hz, H-2), 3.203 (1H, dd, *J* = 13.0, 17.4 Hz, H-3-*trans*), 2.779 (1H, dd, J = 3.2, 17.2 Hz, H-3-cis). UV λ_{max} nm: 288. MS: ESI^- $(M-H)^-$ *m/z* = 273.1 (calc. 273.25), $HRESI^+$ $(M+H)^+$ *m/z* = 275.0719 (calc. 275.0714).

(2*S*)-4'-Amino-5,7-dihydroxyflavanone (**4f**): ^{1}H NMR (400 MHz, d_6-acetone): δ 12.190 (1H, s, 5-OH), 7.228 (2H, d, *J* = 8.4 Hz, H-2',6'), 6.699 (2H, d, *J* = 8.4 Hz, H-3',5'), 5.927 (2H, s, H-6,8) 5.349 (1H, dd, *J* = 3.0, 13.0 Hz, H-2), 4.790 (2H, bs, 4'-NH_2), 3.175 (1H, dd, *J* = 13.0, 17.0 Hz, H-3-*trans*), 2.665 (1H, dd, *J* = 2.8, 17.2 Hz, H-3-*cis*). UV λ_{max} nm: 289. MS: ESI^- $(M-H)^-$ *m/z* = 270.7 (calc. 270.27), $HRESI^+$ $(M+H)^+$ *m/z* = 272.0923 (calc. 272.0917).

(2*E*)-[(2*S*)-5,7-Dihydroxy-4-chromanone]propenoic acid (**4g**): UV λ_{max} nm: 285. MS: ESI^- $(M-H)^-$ *m/z* = 248.9 (calc. 249.20).

(2*R*,3*R*)-Dihydrokaempferol (**5a**): λ_{max} nm: 291.

(2*R*,3*R*)-3',5,7-Trihydroxydihydroflavonol (5**b**): ^{1}H NMR (400 MHz, d_6-acetone): δ 11.711 (1H, s, 5-OH), 7.780 (1H, s, 3'-OH), 7.243 (1H, d, *J* = 8.0 Hz, H-5'), 7.068 (1H, m, H-6'), 7.048 (1H, m, H-2'), 6.872 (1H, d, *J* = 6.4 Hz, H-4'), 5.975 (1H, d, *J* = 7.6 Hz, H-6), 5.975 (1H, d, *J* = 7.6 Hz, H-8), 5.092 (1H, d, *J* = 11.2 Hz, H-2), 4.594 (1H, d, *J* = 11.2 Hz, H-3). UV λ_{max} nm: 287. MS: ESI$^-$ $(M-H)^-$ *m/z* = 287.3 (calc. 287.26), HRESI$^+$ $(M+Na)^+$ *m/z* = 311.0529 (calc. 311.0526).

(2*R*,3*R*)-2',5,7-Trihydroxydihydroflavonol (**5c**): ^{1}H NMR (400 MHz, d_6-acetone): δ 11.720 (1H, s, 5-OH), 8.465 (1H, s, 2'-OH), 7.529 (1H, d, *J* = 7.6 Hz, H-6'), 7.221 (1H, m, H-4'), 6.942 (2H, m, H-5',3'), 5.959 (1H, d, *J* = 2.4 Hz, H-6), 5.924 (1H, d, *J* = 2.4 Hz, H-8), 5.587 (1H, m, H-2), 4.838 (1H, d, *J* = 11.6 Hz, H-3). UV λ_{max} nm: 287. MS: ESI$^-$ $(M-H)^-$ *m/z* = 287.7 (calc. 287.26), HRESI$^+$ $(M+Na)^+$ *m/z* = 311.0538 (calc. 311.0526).

(2*R*,3*R*)-5,7-Dihydroxy-4'-fluorodihydroflavonol (**5d**): ^{1}H NMR (400 MHz, d_6-acetone): δ 11.722 (1H, s, 5-OH), 7.690 (2H, m, H-2',6'), 7.252 (2H, t, *J* = 8.4 Hz, H-3',5'), 6.032 (2H, d, *J* = 11.6 Hz, H-6,8), 5.245 (1H, d, *J* = 11.6 Hz, H-2), 4.870 (1H, s, 3-OH), 4.691 (1H, d, *J* = 11.6 Hz, H-3). UV λ_{max} nm: 290. MS: ESI$^-$ $(M-H)^-$ *m/z* = 289.7 (calc. 289.25), HRESI$^+$ $(M+Na)^+$ *m/z* = 313.0490 (calc. 313.0483).

(2*R*,3*R*)-5,7-Dihydroxy-2'-fluorodihydroflavonol (**5e**): ^{1}H NMR (400 MHz, d_6-acetone): δ 11.730 (1H, s, 5-OH), 9.844 (1H, s, 7-OH), 7.781 (1H, t, *J* = 7.2 Hz, H-6'), 7.523 (1H, d, *J* = 6.4, 13.6 Hz, H-4'), 7.355 (1H, t, *J* = 7.6 Hz, H-5'), 7.252 (1H, t, *J* = 10.4 Hz, H-3'), 6.063 (1H, d, *J* = 1.6 Hz, H-6), 6.016 (1H, d, J = 2.0 Hz, H-8), 5.538 (1H, d, *J* = 12.0 Hz, H-2), 4.945 (1H, s, 3-OH), 4.851 (1H, d, *J* = 12.0 Hz, H-3). UV λ_{max} nm: 288. MS: ESI$^-$ $(M-H)^-$ *m/z* = 289.6 (calc. 289.25), HRESI$^+$ $(M+Na)^+$ *m/z* = 313.0490 (calc. 313.0483).

(2*R*,3*R*)-4'-Amino-5,7-dihydroxydihydroflavonol (**5f**): ^{1}H NMR (400 MHz, d_6-acetone): δ 11.729 (1H, s, 5-OH), 7.267 (2H, d, *J* = 8.4 Hz, H-2',6'), 6.710 (2H, d, *J* = 8.0 Hz, H-3',5'), 5.983 (1H, s, H-6), 5.935 (1H, s, H-8), 5.001 (1H, d, *J* = 12.0 Hz, H-2), 4.799 (2H, bs, 4'-NH_2), 4.651 (1H, d, 11.6 Hz, H-3). UV λ_{max} nm: 289. MS: ESI$^-$ $(M-H)^-$ *m/z* = 286.3 (calc. 286.27).

UV Profiles.

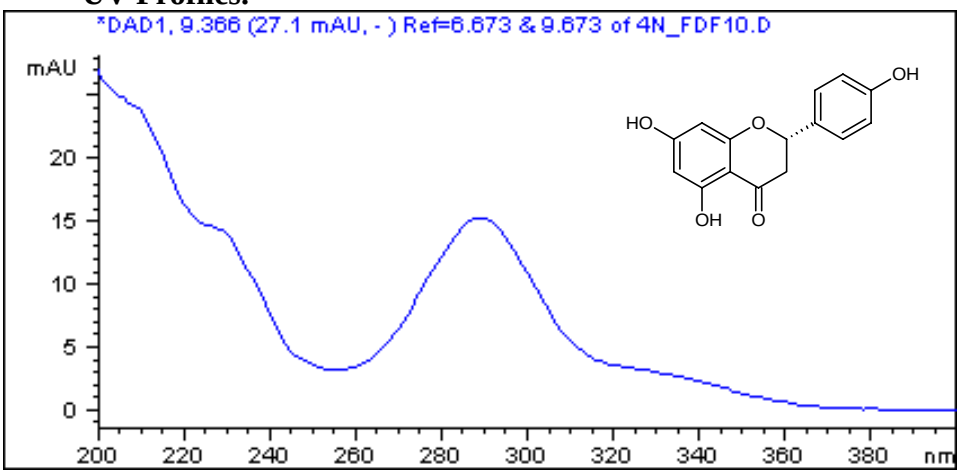

Figure 1. UV profile and structure of (2*S*)-naringenin (**1a**)

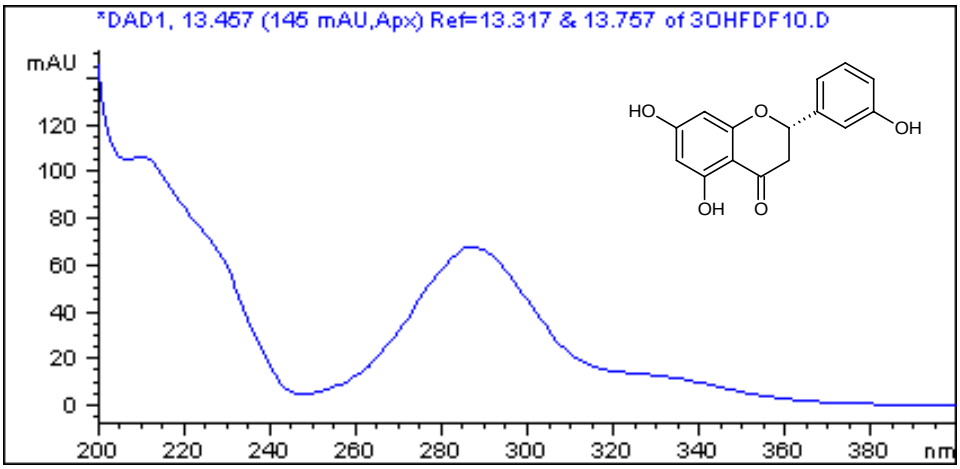

Figure 2. UV profile and structure of (2*S*)-3',5,7-trihydroxyflavanone **(4b)**

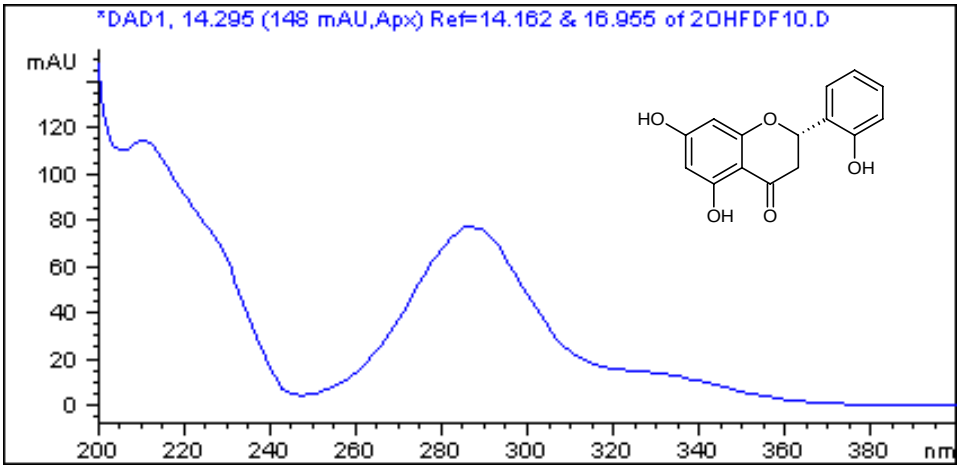

Figure 3. UV profile and structure of (2*S*)-2',5,7-trihydroxyflavanone **(4c)**

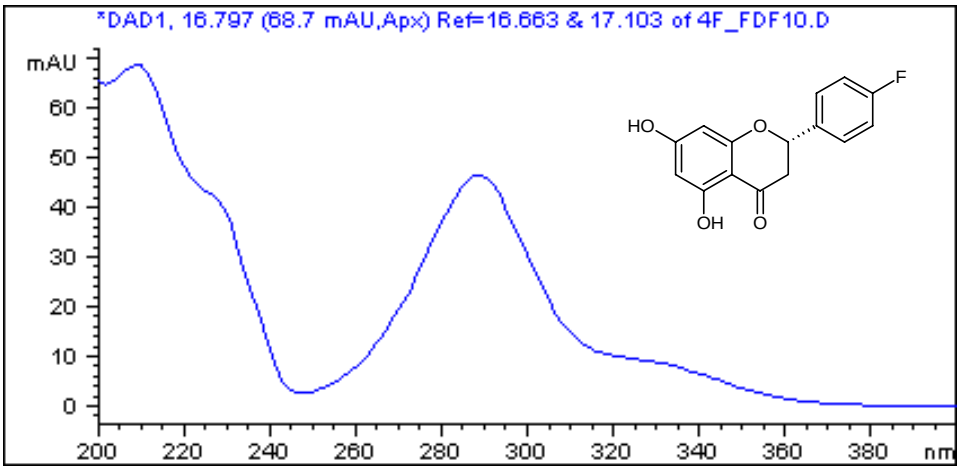

Figure 4. UV profile and structure of (2*S*)-5,7-dihydroxy-4'-fluoroflavanone **(4d)**

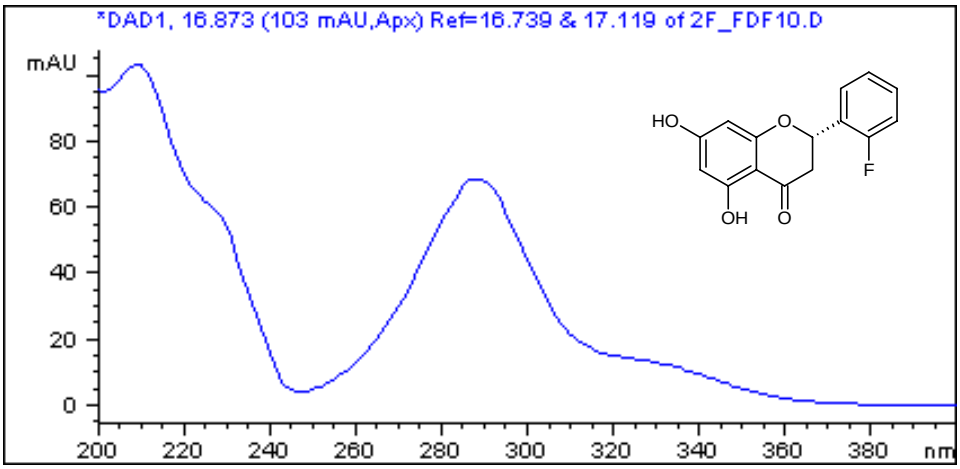

Figure 5. UV profile and structure of (2*S*)-5,7-dihydroxy-2'-fluoroflavanone **(4e)**

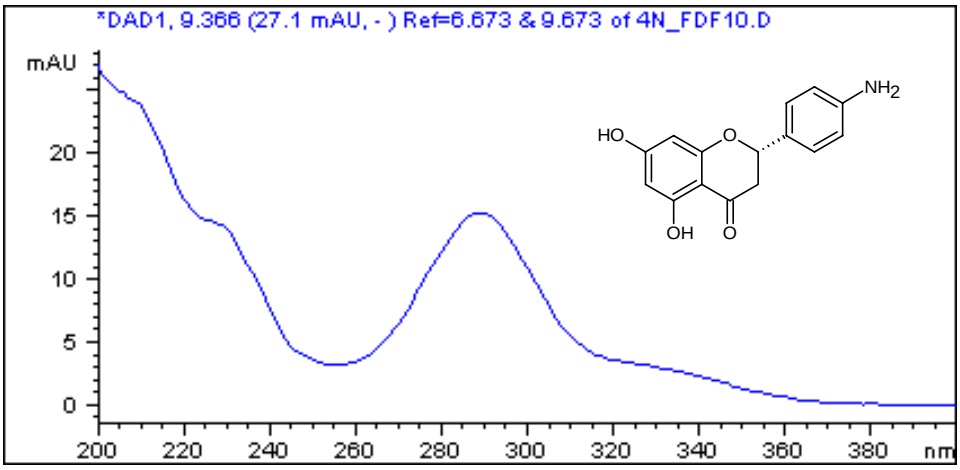

Figure 6. UV profile and structure of (2*S*)-4'-amino-5,7-dihydroxyflavanone **(4f)**

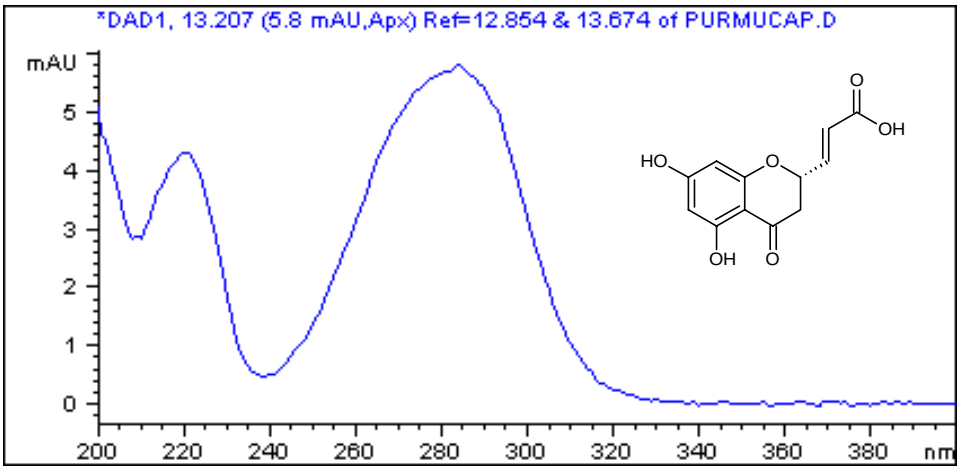

Figure 7. UV profile and structure of (2*E*)-[(2*S*)-5,7-Dihydroxy-4-chromanone]propenoic acid (**4g**)

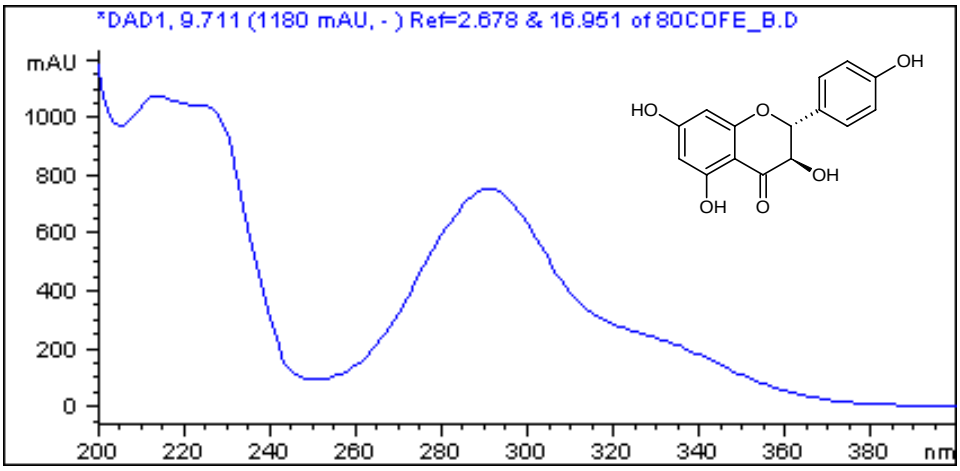

Figure 8. UV profile and structure of (2*R*,3*R*)-dihydrokaempferol **(5a)**

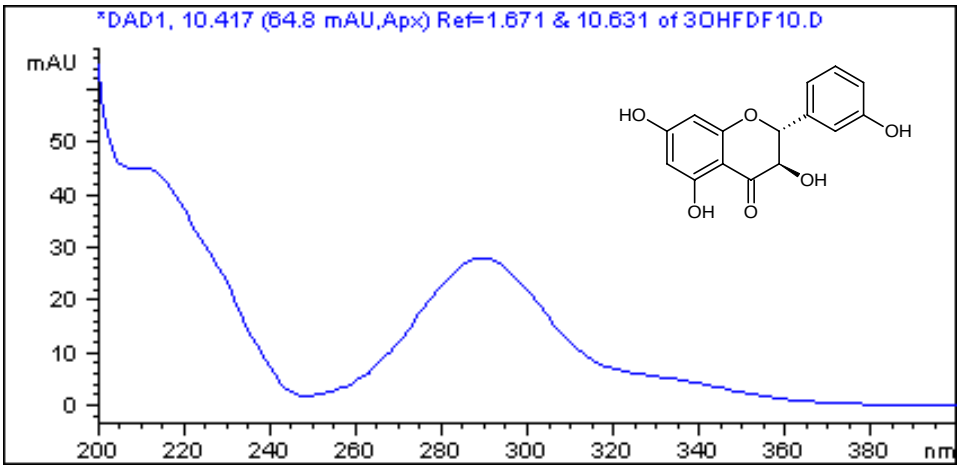

Figure 9. UV profile and structure of (2*R*,3*R*)-3',5,7-trihydroxydihydroflavonol **(5b)**

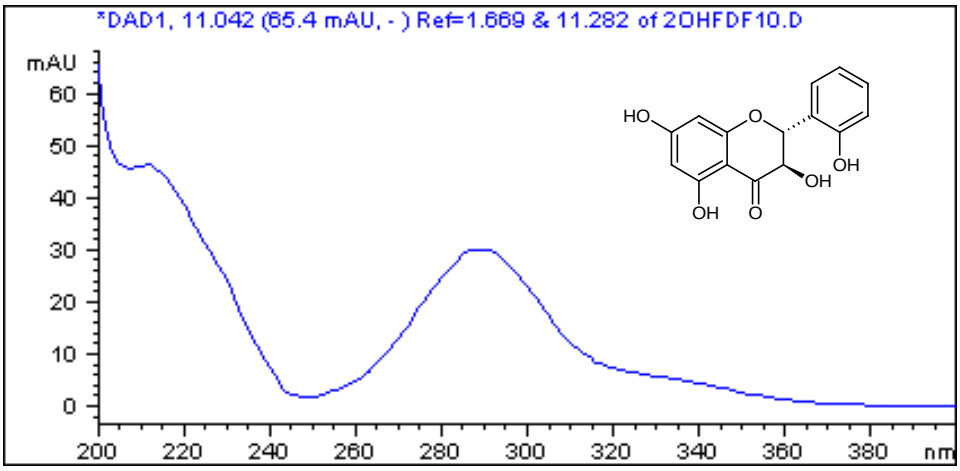

Figure 10. UV profile and structure of (2*R*,3*R*)-2',5,7-trihydroxydihydroflavonol **(5c)**

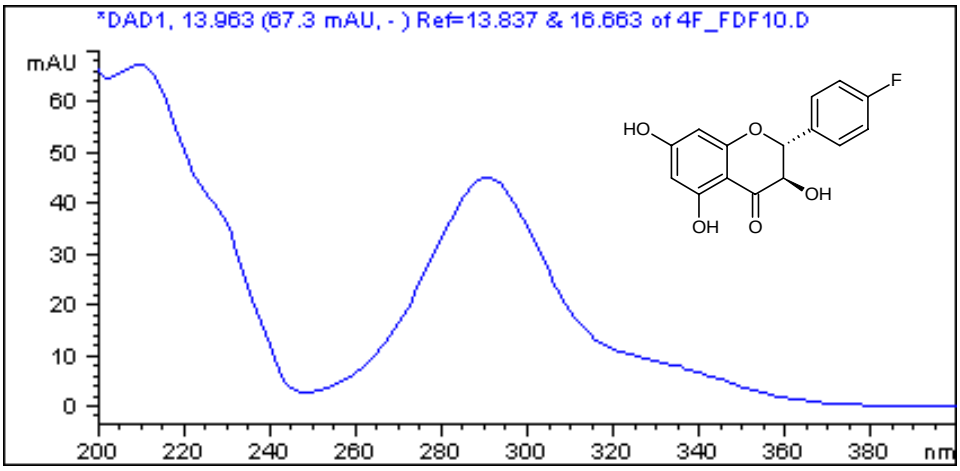

Figure 11. UV profile and structure of (2*R*,3*R*)-5,7-dihydroxy-4'-fluorodihydroflavonol **(5d)**

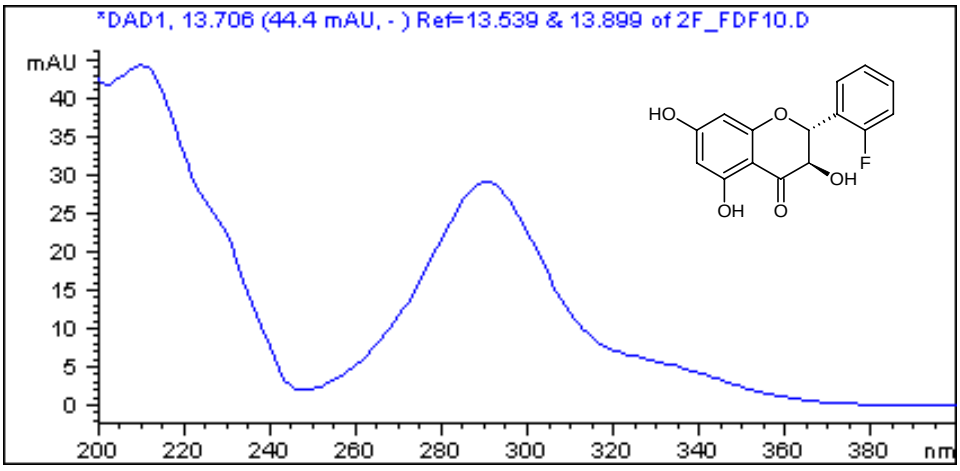

Figure 12. UV profile and structure of (2*R*,3*R*)-5,7-dihydroxy-2'-fluorodihydroflavonol **(5e)**

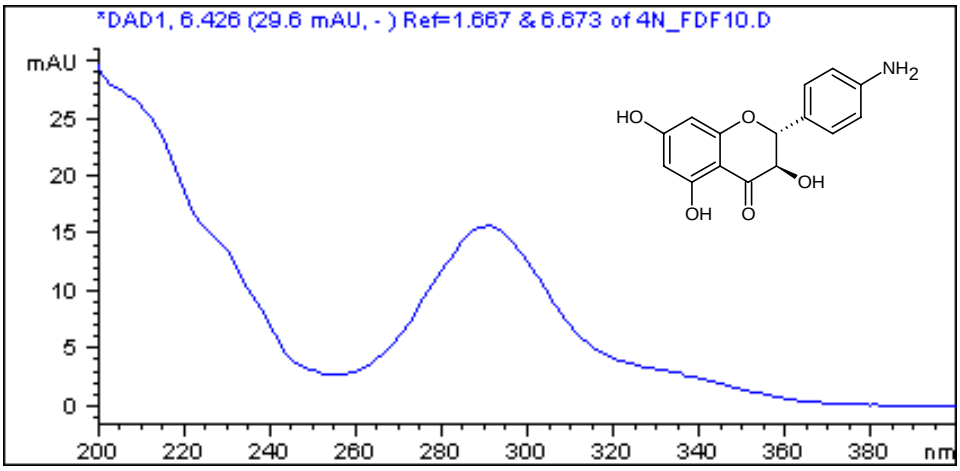

Figure 13. UV profile and structure of (2*R*,3*R*)-4'-amino-5,7-dihydroxydihydroflavonol **(5f)**

References.

(1) Sambrook, J.; Fritsch, E. F.; Maniatis, T., *Molecular Cloning: a Laboratory Manual*. 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., **1989**.
(2) Yan, Y.; Kohli, A.; Koffas, M. A. *Appl Environ Microb* **2005,** *71*, 5610-5613.
(3) Yan, Y.; Chemler, J.; Huang, L.; Martens, S.; Koffas, M. A. *Appl. Environ. Microb.* **2005,** *71*, 3617-3623.