SUPPLEMENTARY MATERIAL

Isolation and characterization of secondary metabolites from *Trifolium vesiculosum* Savi and their antiproliferative activities

Dicle Çevik[a,*], Basak Aru[b], Sanem Karagoz[c], Niyazi Gurizi[c], Ozlem Demirkiran[a]

[a] *Department of Pharmacognosy, Faculty of Pharmacy, Trakya University, TR-22030, Balkan Campus, Edirne, Turkey*

[b] *Department of Immunology, Faculty of Medicine, Yeditepe University, TR-34755, Kayışdağı, İstanbul, Turkey*

[c] *Faculty of Pharmacy, Trakya University, TR-22030, Balkan Campus, Edirne, Turkey*

* Corresponding author: Dicle Çevik

E-mail address: dicledemir@windowslive.com (D. Çevik).

ABSTRACT

One unreported flavonol namely morin-7-*O*-methyl ether (**1**) along with seven known compounds were isolated from the aerial parts of *Trifolium vesiculosum* Savi which were elucidated by using extensive spectroscopic methods such as 1D and 2D NMR and HR-MS. According to the cell viability assay (MTS) on the purified compounds (**1**-**8**), quercetin-3-*O*-(6''-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**) revealed remarkable antiproliferative activity most particularly against breast cancer cells (IC_{50} = 2.90±0.25 μM in HCC1937 and 7.98±0.57 μM in MCF7) while moderate inhibitory activity (IC_{50} = 17.96±0.51-51.70±2.69 μM) on prostate, colorectal and liver cancer cell viability was observed. Further mechanistic examinations (Annexin V/PI staining, DNA content and detection of reactive oxygen species analyses) showed that compound **4** significantly induced apoptosis, enhanced mitochondrial reactive oxygen species (ROS) accumulation, and caused cell cycle arrest in cancer cells by increasing accumulation of cells at G_0/G_1 and/or G_2/M phases of the cell cycle.

Keywords: *Trifolium vesiculosum* Savi; Phytochemical isolation; Secondary metabolites; Antiproliferative activity; Apoptosis; Cell Cycle

CONTENTS

General experimental procedures..... 5
Isolation of compounds 1-8..... 5
Cell viability assay..... 6
Annexin V/Propidium Iodide staining..... 7
DNA content analysis..... 7
Detection of reactive oxygen species..... 7
Statistical analysis..... 8
Table S1. ^{1}H and ^{13}C NMR data for morin 7-O-methyl ether (**1**) (CD_3OD; ^{1}H: 400 MHz, ^{13}C: 100 MHz)[a]..... 8
Figure S1.. 9
Figure S2.. 10
Figure S3.. 11
Figure S4. ^{1}H NMR spectrum of morin-7-*O*-methyl ether (**1**) (CD_3OD, 400 MHz)..... 12
Figure S5. ^{13}C NMR spectrum of morin-7-*O*-methyl ether (**1**) (CD_3OD, 100 MHz)..... 12
Figure S6. NOESY spectrum of morin-7-*O*-methyl ether (**1**)..... 13
Figure S7. HSQC spectrum of morin-7-*O*-methyl ether (**1**)..... 13
Figure S8. HMBC spectrum of morin-7-*O*-methyl ether (**1**)..... 14
Figure S9. Key HMBC (arrow, H → C) correlations of **1**..... 14
Figure S10. HRESIMS spectrum of morin-7-*O*-methyl ether (**1**)..... 15
Figure S11. ^{1}H NMR spectrum of quercetin (**2**) (CD_3OD, 400 MHz)..... 15
Figure S12. ^{1}H NMR spectrum of hyperoside (**3**) (CD_3OD, 400 MHz)..... 16
Figure S13. ^{13}C NMR spectrum of hyperoside (**3**) (CD_3OD, 100 MHz)..... 16
Figure S14. COSY spectrum of hyperoside (**3**)..... 17
Figure S15. NOESY spectrum of hyperoside (**3**)..... 17
Figure S16. HSQC spectrum of hyperoside (**3**)..... 18
Figure S17. HMBC spectrum of hyperoside (**3**)..... 18
Figure S18. ^{1}H NMR spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**) (CD_3OD, 400 MHz)..... 19
Figure S19. ^{13}C NMR spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**) (CD_3OD, 100 MHz)..... 19
Figure S20. COSY spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**)..... 20
Figure S21. HSQC spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**)..... 20

Figure S22. HMBC spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**) ... 21
Figure S23. ^{1}H NMR spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) (CD_3OD, 400 MHz) ... 21
Figure S24. ^{13}C NMR spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) (CD_3OD, 100 MHz) ... 22
Figure S25. COSY spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) ... 22
Figure S26. HSQC spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) ... 23
Figure S27. HMBC spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) ... 23
Figure S28. ^{1}H NMR spectrum of daphetin (**6**) (CD_3OD, 400 MHz) ... 24
Figure S29. ^{1}H NMR spectrum of daphnin (**7**) (CD_3OD, 400 MHz) ... 24
Figure S30. ^{13}C NMR spectrum of daphnin (**7**) (CD_3OD, 100 MHz) ... 25
Figure S31. NOESY spectrum of daphnin (**7**) ... 25
Figure S32. HSQC spectrum of daphnin (**7**) ... 26
Figure S33. HMBC spectrum of daphnin (**7**) ... 26
Figure S34. ^{1}H NMR spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) (CD_3OD, 400 MHz) ... 27
Figure S35. ^{13}C NMR spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) (CD_3OD, 100 MHz) ... 27
Figure S36. COSY spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) ... 28
Figure S37. NOESY spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) ... 28
Figure S38. HSQC spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) ... 29
Figure S39. HMBC spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) ... 29

General experimental procedures

Chromatographic methods: Vacuum Liquid Chromatography (VLC) was conducted on LiChroprep® RP-18 (25-40 μm, Merck, Darmstadt, Germany). Polyamide (50-160 μm; Fluka Analytical, Sigma-Aldrich, St. Louis, MO, USA), Silica gel 60 (0.063-0.200 mm; Merck, Darmstadt, Germany) and Sephadex LH-20 (25-100 μm; Sigma-Aldrich, St. Louis, MO, USA) were used for Open Column chromatography (OCC) separations. Semi-preparative HPLC separation was performed on Waters Modular System (Waters Corporation, Massachusetts, USA), equipped with a binary HPLC pump (1525), autosampler (2707), a photodiode array detector (2998) and fraction collector III by using C_{18} column (YMC-Pack Pro C18, 5 μM, 250 x 20 mm). Sepacore® Flash Systems X10/X50 (Buchi Labortechnik AG, Flawil, Switzerland) was utilized for Medium-Pressure Liquid Chromatography (MPLC) on LiChroprep C_{18} RediSep Flash columns (Teledyne Isco, Lincoln, Nebraska, USA). Thin layer chromatography (TLC) analyses were carried out on aluminum plates layered with Silica gel 60 F_{254} plates (9.5 - 11.5 μm; Merck, Darmstadt, Germany), derivatization was performed by spraying 1% vanillin/sulphuric acid solution followed by heating at 105° for 2-3 min and visualized by using UV lamb at 254 and 366 nm.

Structure elucidation methods: UV spectra were measured on HP Agilent 8453 spectrophotometer (Agilent Technologies, Santa Clara, CA, USA); λmax in nm. IR Spectra were obtained by PerkinElmer 2000 FT-IR spectrometer (PerkinElmer, Waltham, Massachusetts, USA); Ʊ in cm^{-1}. ^{1}H (400 MHz), ^{13}C (100 MHz), COSY, HSQC, HMBC and NOESY NMR spectra were recorded on a Varian Mercury FT spectrometer (Palo Alto, CA, USA). Mass spectra were measured on HRESIMS 6200 series TOF/6500 series Q-TOF (Waters Corporation, Massachusetts, USA) in methanol (MeOH); positive-ion mode; in *m/z*.

Cell culture: Human breast (HCC1937 and MCF7), prostate (PC3 and DU145), colorectal (LoVo and HCT116) and liver (Hep3B and HepG2) cancer cell lines were purchased from the American Type Culture Collection (ATCC). Cells were cultured in either DMEM (MCF7, PC3, Hep3B and HepG2), DMEM/F12 (LoVo and HCT116) or RPMI (HCC1937 and DU145) culture media supplemented with 10% fetal bovine serum (FBS, Sigma Aldrich # F7524) and penicillin/streptomycin antibiotic solution (Sigma Aldrich #P4333).

Isolation of compounds 1-8

Fr. E_5 (81mg) was loaded into SiO_2-Open Column Chromatography (OCC) and eluted with DCM:MeOH (from 100:0 to 50:50) mobile phase system, yielding to subfractions including Fr $E_{5\text{-}A}$ (21 mg) which was separated by Sephadex LH-20 CC (10 g, MeOH) for the

purification of **6** (7.3 mg). Fr E_{12} (72 mg) was submitted to Sephadex LH-20 (30 g) CC to isolate compound **1** (5 mg). Semiprep-HPLC technique was used for the separation of fr. E_8 (450 mg) by using C_{18} column and linear gradient elution (0-10 min, from 45 to 45% MeOH; 10-40 min from 45 to 100% MeOH; flow rate: 9 mL/min) affording 15 subfractions (subfrs. $E_{8\text{-}A\text{-}O}$). Compound **3** (6.7 mg) was isolated from fr. $E_{8\text{-}F}$ (34 mg) by isocratic elution with MeOH on Sephadex LH-20 (35 g). Similarly, fr. $E_{8\text{-}O}$ was loaded into Sephadex LH-20 (35 g) with MeOH to afford compound **2** (3 mg). Fr 10 was submitted to OCC (SiO_2, 20 g) eluting with DCM:MeOH (0 → 20% MeOH) for fractionation (frs. $E_{10\text{-}A\text{-}L}$) as well as isolation of **2** (7.6 mg). Purification of compound **4** (8 mg) was achieved by submitting fr. $E_{10\text{-}K}$ (24 mg) into Gel Filtration Chromatography (Sephadex LH-20, 10 g) with MeOH.

Fr. B_1 (500 mg) was separated by OCC (SiO_2, 50 g) by using DCM:MeOH:H_2O (95:5:0 → 50:40:10) mobile phase systems and thus yielded subfractions of $B_{1\text{-}A\text{-}O}$. Fr. $B_{1\text{-}E}$ (39 mg) was submitted to Sephadex LH-20 (35 g) eluting with MeOH for the purification of compound **7** (9.5 mg). Separation of fr. B_5 (271 mg) by Polyamide CC (5 g, H_2O:MeOH; 10 → 100% MeOH) yielded to fr. $B_{5\text{-}D}$ (45 mg) which was further submitted to Sephadex LH-20 (40 g) CC with MeOH in order to isolate compound **3** (5 mg). Fr B_3 was applied to normal phase MPLC (80 g SiO_2) eluting with DCM:MeOH mixture (0 → 40% MeOH) to purify compound **5** (15 mg) as well as other fractions including Fr $B_{3\text{-}C}$ (125 mg) which was further fractionated by C_{18}-MPLC (43 g; 10 → 80% MeOH in H_2O). Yielded fr. $B_{3\text{-}C\text{-}1}$ (60 mg) was again loaded into reverse phase MPLC (15.5 g C_{18}) eluting with MeOH:H_2O (20 → 25% MeOH) for the isolation of compound **8** (6.5 mg).

Cell viability assay

Antiproliferative properties of the compounds on the breast (MCF7 and HCC1937), prostate (PC3 and DU145), colorectal (LoVo and HCT116) and liver (Hep3B and HepG2) cancer cell lines were evaluated with MTS assay. For this purpose, 5×10^3 cells were seeded as triplicates to 96-well cell culture plates and incubated overnight to allow attachment. Isolated compounds (**1**-**8**) were dissolved in DMSO at 100 mM final concentration to prepare stock solutions and further diluted in a complete culture medium to obtain 100, 50, 25, 12.5, 6.25 and 3.25 μM working concentrations. Cells were incubated with the compounds for 24 hours at the abovementioned doses for 48 hours. Untreated cells were used as positive control while culture media was used as blank. Viability was evaluated with MTS assay (Abcam, #ab197010) by measuring optical density at 490 nm wavelength and calculated according to the formula given below.

$$Viability\ (\%) = \frac{Sample\ OD - Blank\ OD}{Positive\ control\ OD - Blank\ OD} \times 100$$

Annexin V/Propidium Iodide staining

Apoptosis and viability rate were evaluated with Annexin V/Propidium Iodide (PI) staining which investigates the apoptotic marker phosphatidylserine translocation and increased membrane permeability simultaneously [15]. For this purpose, 5×10^5 cells were seeded to 60 mm tissue culture dishes as triplicates and treated with compound **4** at the respective IC_{50} doses determined for each cell line. At the end of the incubation period, cells were detached with trypsin-EDTA (Thermo Fisher Scientific, #25200056) solution. The pellet was collected by centrifuging tubes at 300 g for 3 minutes, followed by washing twice with ice-cold Annexin V Binding Buffer (Biolegend, #422201). Cells were labeled with 5 µL Annexin V-FITC (Biolegend, # 640906) by incubating tubes for 15 minutes at room temperature under dark conditions. 1 µL Propidium Iodide (diluted to 250 µg/mL in DPBS, Thermo Fisher Scientific, #P1304MP) was added 5 minutes prior to analysis to label dead cells. Tubes were acquired with DxFLEX flow cytometry system (Beckman Coulter Inc., USA). 2.5×10^4 cells were analysed for each tube. Data were evaluated with CytExpert software.

DNA content analysis

The effect of compound **4** on cell cycle progression as well as DNA fragmentation was investigated with DNA content analysis [16]. Cells were seeded to 60 mm tissue culture dishes as triplicates and treated with **4**, followed by detachment with trypsinization and fixation with 70% ethanol by incubating at 4°C for an hour. Ethanol was discarded by centrifuging cells at 350 g for 5 min. Cells were washed once with DPBS and stained with 500 µL Cell Cycle Kit reagent (Beckman Coulter, #C03551) by incubating tubes in the dark at room temperature for 40 min. For each tube, 5×10^4 cells were analysed. Evaluations were done with ModFit LT (Version 4.0) software.

Detection of reactive oxygen species

Mitochondrial reactive oxygen species (ROS) accumulation was determined with Dihydrorhodamine 123 (DHR123) staining. DHR123 is a non-fluorescent dye that passively diffuses into the cells where it is oxidized to rhodamine 123, a fluorescent probe accumulates into the mitochondria and can be detected with flow cytometry [17]. For evaluating ROS

accumulation, cells treated with **4** were collected with trypsinization and washed once with DPBS. The pellet was suspended in 500µl DPBS and incubated with 5 µM DHR123 for 40 minutes in the dark. Tubes were acquired with DxFLEX flow cytometry system. 2.5×10^4 cells were analysed for each tube. Data were evaluated with CytExpert software.

Statistical analysis

Statistical analyses were done with GraphPad Prism software (Version: 8.0). Viability, early apoptosis, late apoptosis and necrosis as well as DHR123 results were evaluated with unpaired t-test while the results of DNA content analysis were analysed with two-way analysis of variance (ANOVA) test followed by Sidak's multiple comparison test. All analyses were conducted in triplicates. P values lower than 0.05 were considered statistically significant.

Table S1. ^{1}H and ^{13}C NMR data for morin 7-*O*-methyl ether (**1**) (CD_3OD; ^{1}H: 400 MHz, ^{13}C: 100 MHz)[a]

C/H atom	δ_H (ppm), *J* (Hz)	δ_C (ppm)	
2	-	148.7	C
3	-	135.2	C
4	-	177.0	C
5	-	152.9	C
6	6.34 d (1.9)	99.1	CH
7	-	162.3	C
8	6.66 d (1.9)	88.7	CH
9	-	158.1	C
10	-	108.1	C
1′	-	109.1	C
2′	-	158.5	C
3′	6.50 d (2.1)	103.9	CH
4′	-	163.0	C
5′	6.51 dd (9.1,2.1)	108.6	CH
6′	7.45 d (9.1)	133.5	CH
7-OCH_3	3.83 s	56.2	CH_3

[a] Assignments are based on NOESY, HSQC and HMBC experiments.

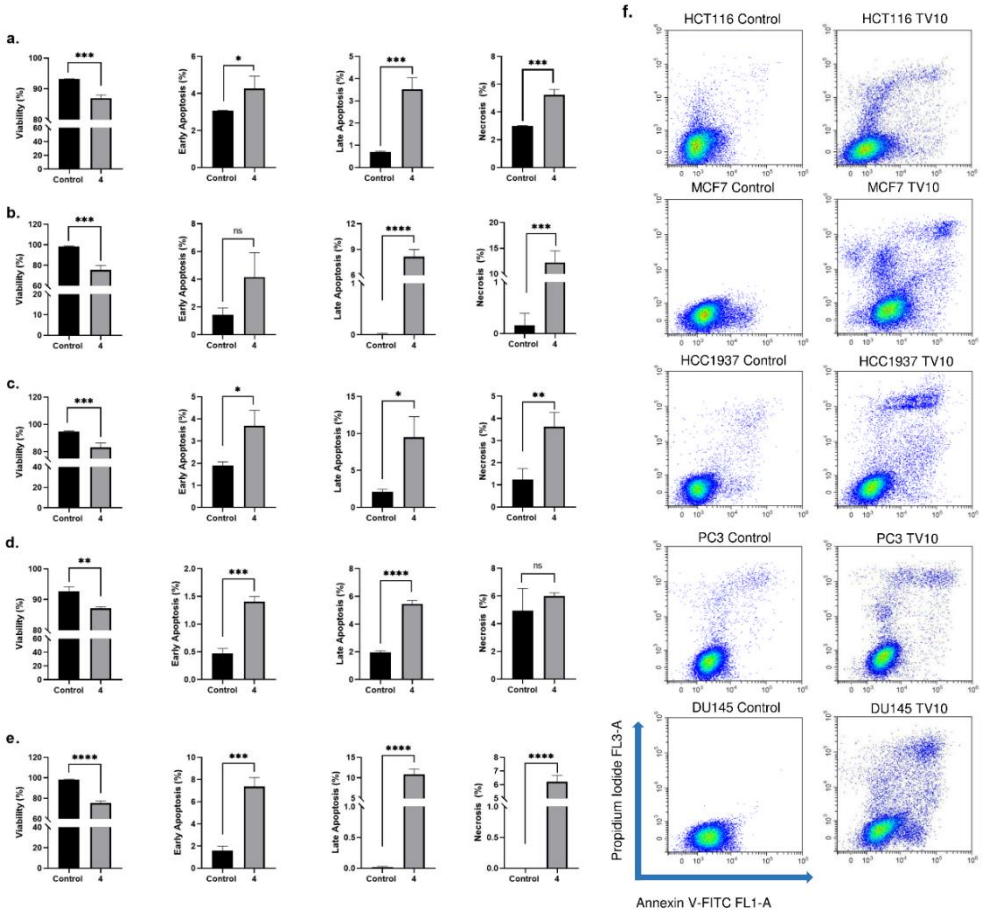


Figure S1. Compound **4** decreased viability while promoting apoptosis on all cell lines tested. **a.** Bar graphics indicating viability, early apoptosis, late apoptosis and necrosis ratios in colorectal cancer cell line HCT116. **b.** Bar graphics indicating viability, early apoptosis, late apoptosis and necrosis ratios in double positive MCF7 cell line. **c.** Bar graphics indicating viability, early apoptosis, late apoptosis and necrosis ratios in HCC1937 triple negative breast cancer cell line. **d.** Bar graphics indicating viability, early apoptosis, late apoptosis and necrosis ratios in PC3, androgen receptor irresponsive cell line. **E.** Bar graphics indicating viability, early apoptosis, late apoptosis and necrosis ratios in DU145, androgen receptor sensitive cell line. **f.** Representative flow cytometry dot plots. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

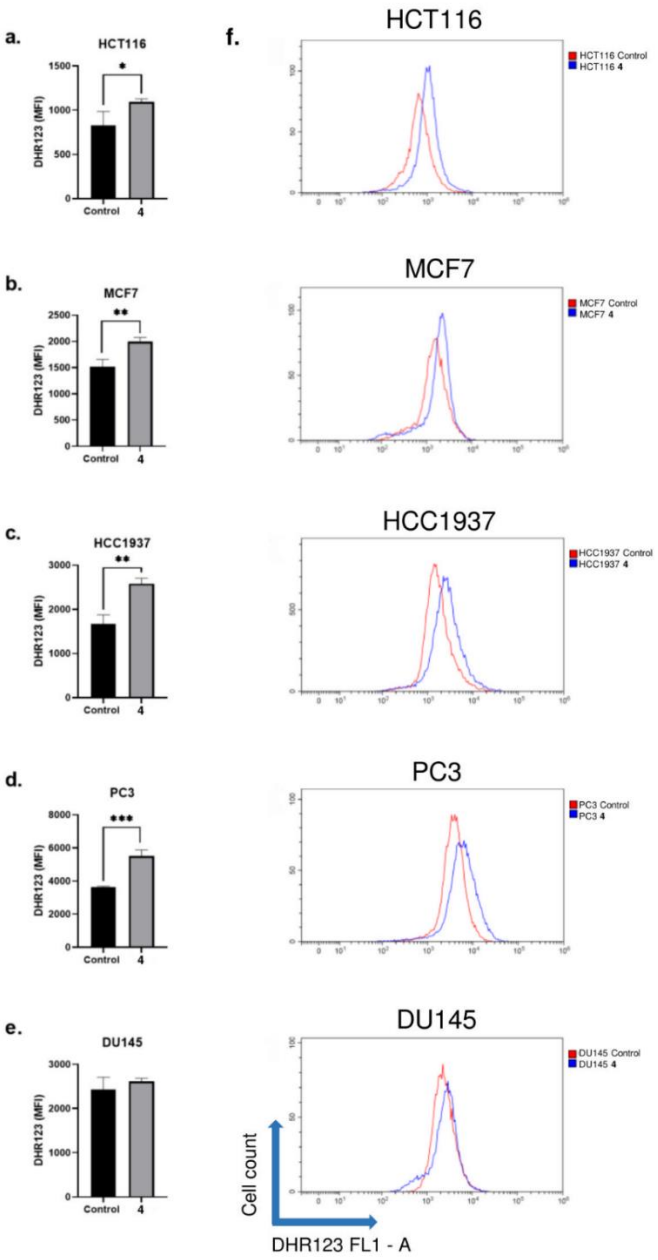


Figure S2. Compound **4** promoted reactive oxygen species' accumulation in all cell lines tested except DU145 cells. **a-e.** Bar graphics indicating differences between untreated and compound **4** treated cell lines DHR123 MFI levels. **f.** Representative flow cytometry histogram overlays. *p<0.05, **p<0.01, ***p<0.001.

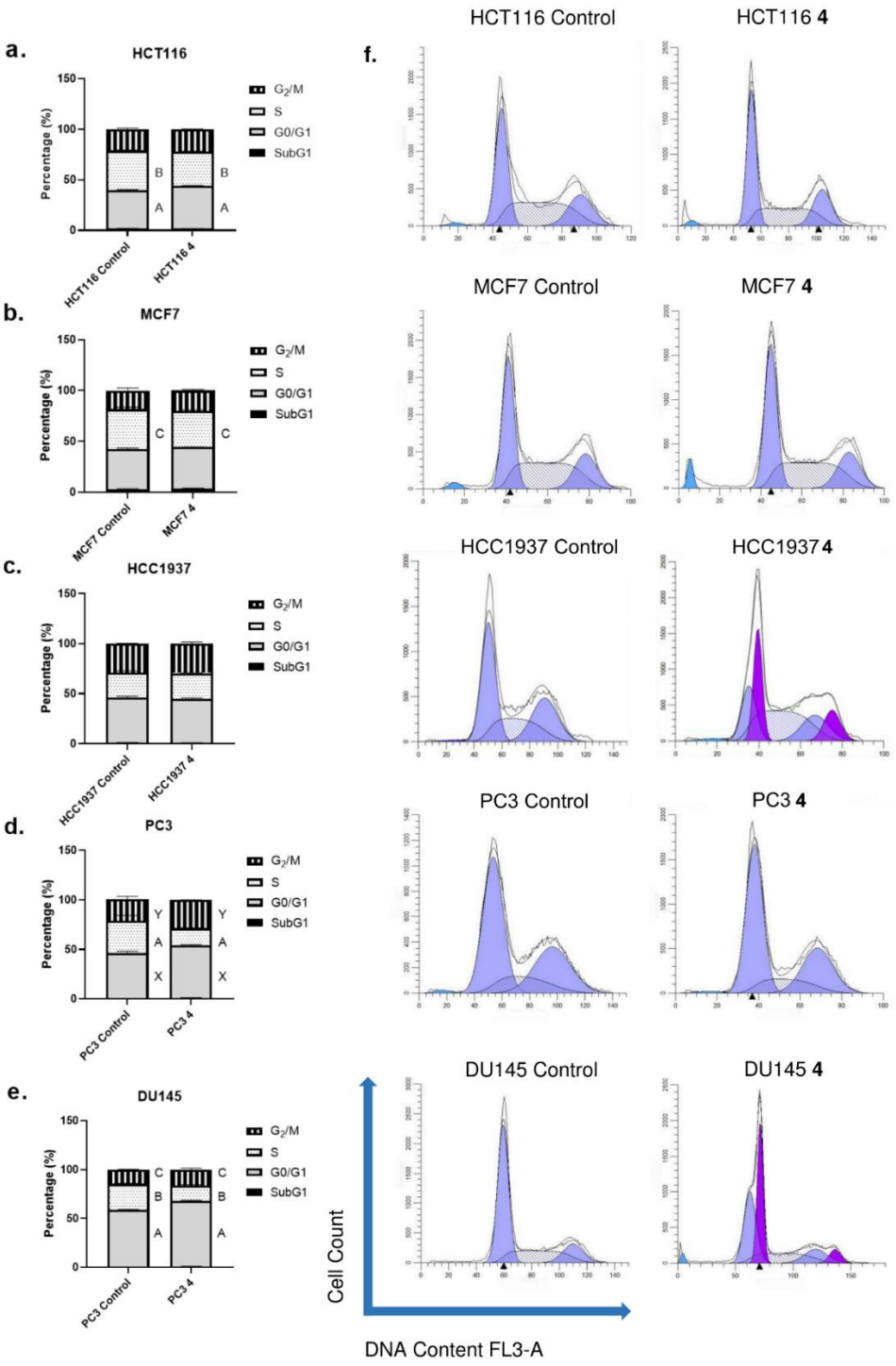


Figure S3. Compound **4** disrupted cell cycle progression in cancer cell lines evaluated in this study. **a.** Treatment led to an increase in G_0/G_1 phase which is accompanied by a decrease in S phase in HCT116 cells. **b.** Treatment decreased S phase in MCF7 cell line, but cell cycle arrest was not observed. **c. 4** did not increase accumulation in a cell division phase in HCC1937 cell line, but led to aneuploidy. **d.** Treatment promoted G_0/G_1 phase arrest which is accompanied by an increase in G_2/M phase in PC3 cell line. **e.** Treatment promoted G_0/G_1 phase arrest which is accompanied by an increase in G_2/M phase in DU145 cell line, moreover, aneuploidy upon treatment of **4** was observed. **f.** Representative flow cytometry histograms. A, B ****$p<0.0001$; C *$p<0.05$; X, Y $p<0.001$.

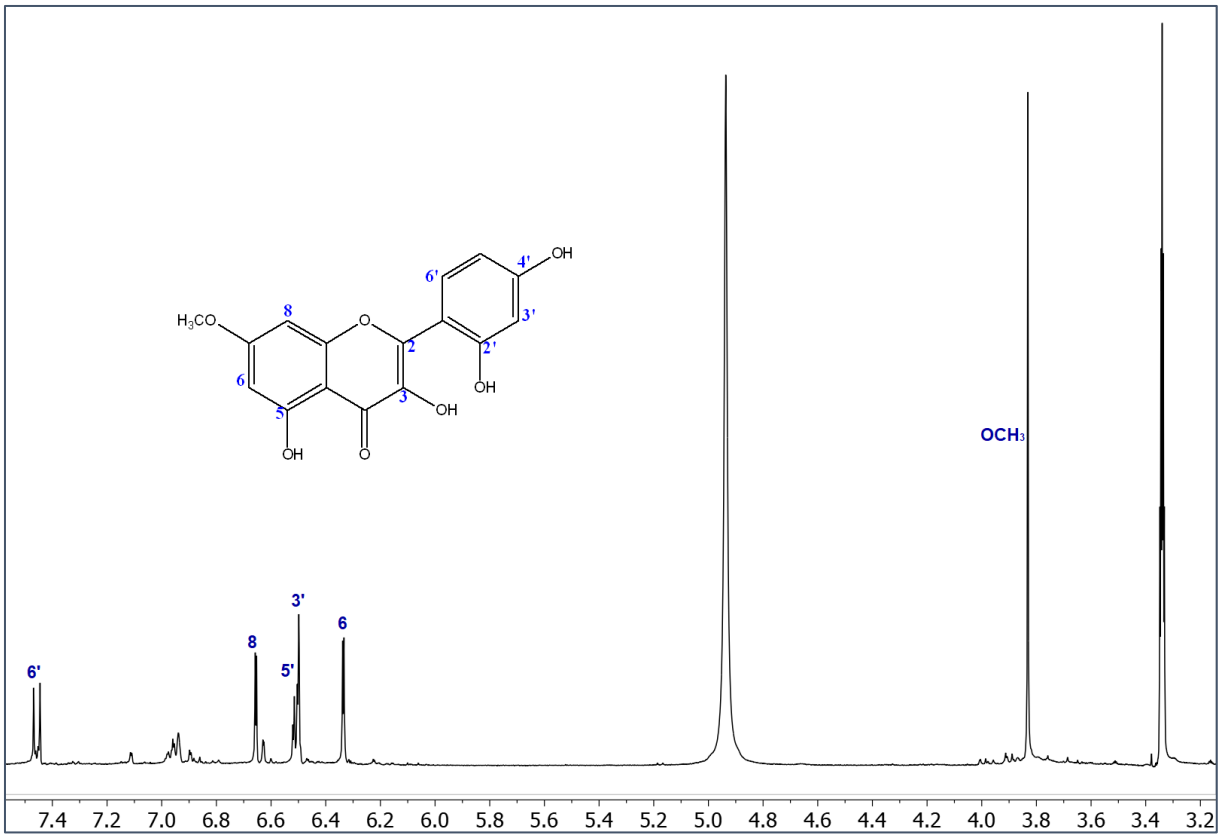


Figure S4. ^{1}H NMR spectrum of morin-7-*O*-methyl ether (**1**) (CD_3OD, 400 MHz)

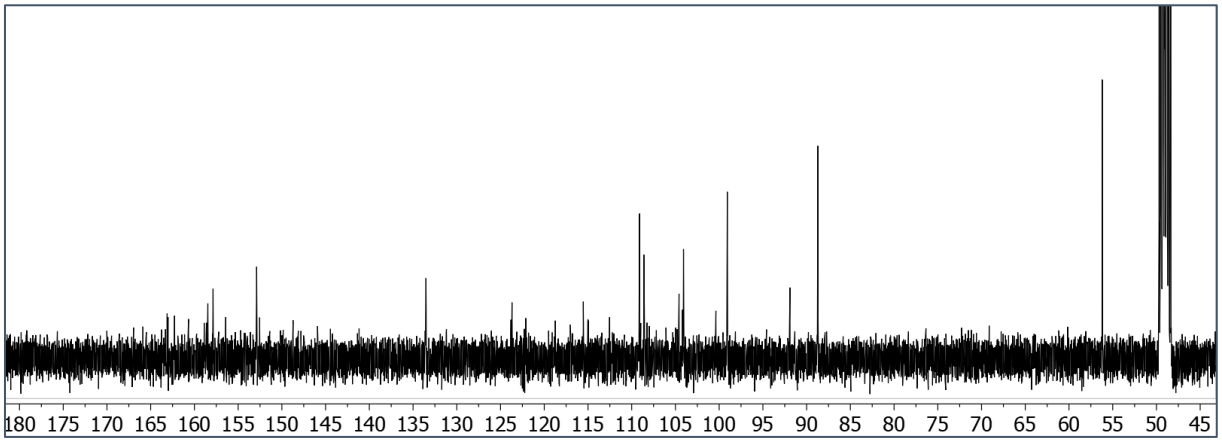


Figure S5. ^{13}C NMR spectrum of morin-7-*O*-methyl ether (**1**) (CD_3OD, 100 MHz)

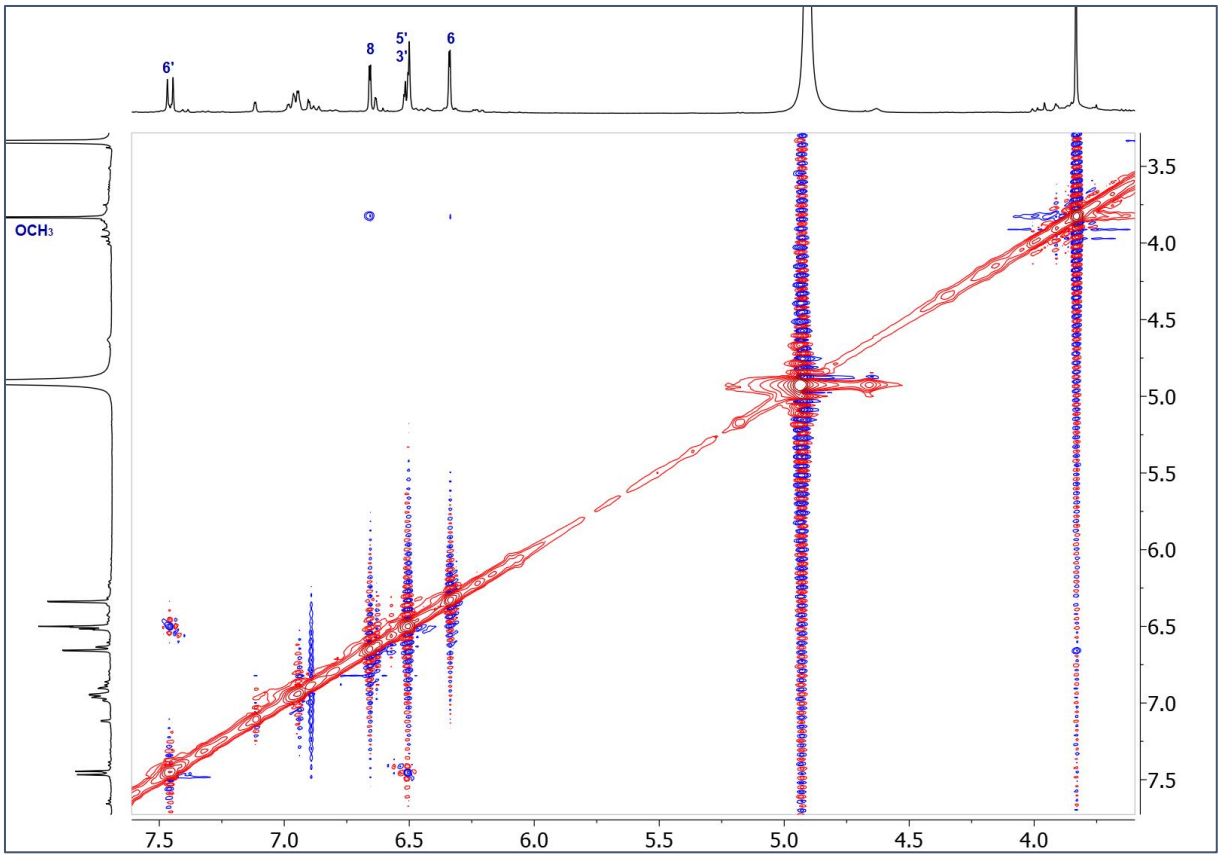


Figure S6. NOESY spectrum of morin-7-*O*-methyl ether (**1**)

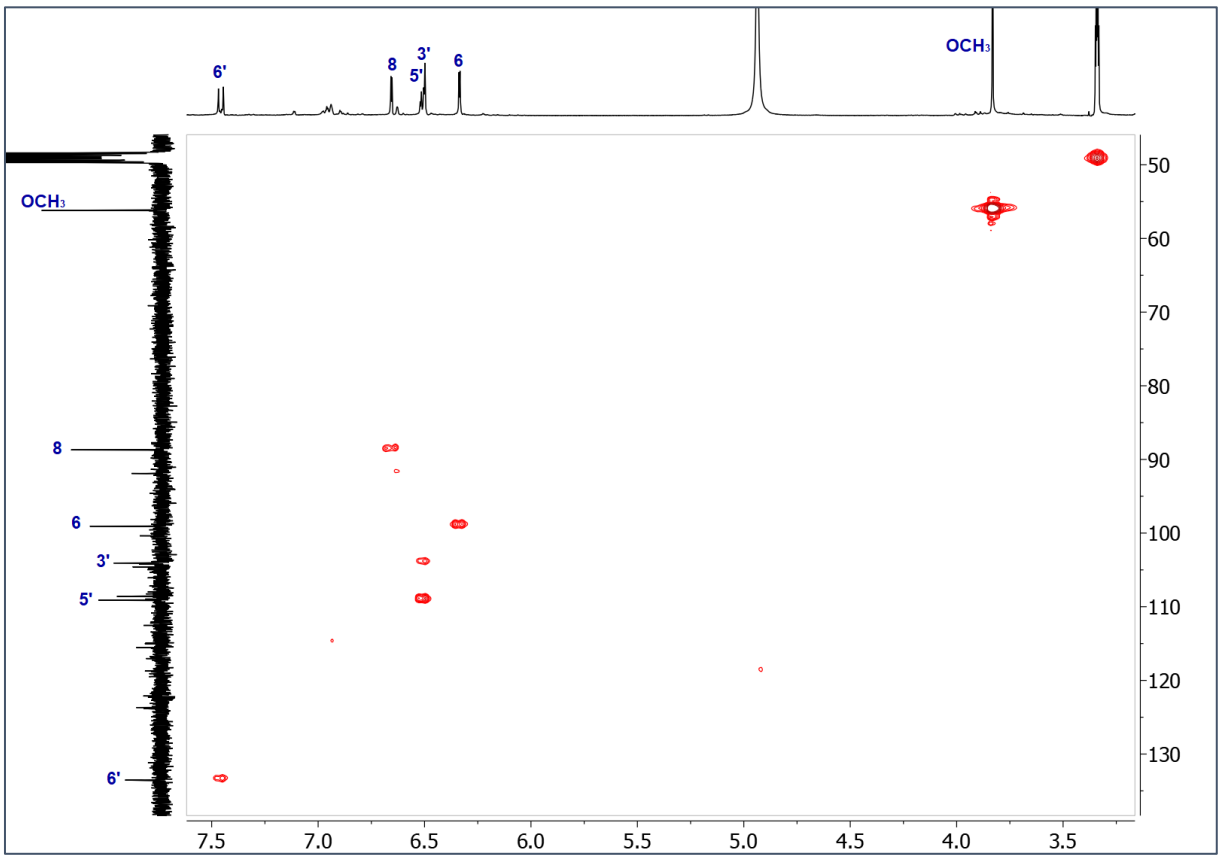


Figure S7. HSQC spectrum of morin-7-*O*-methyl ether (**1**)

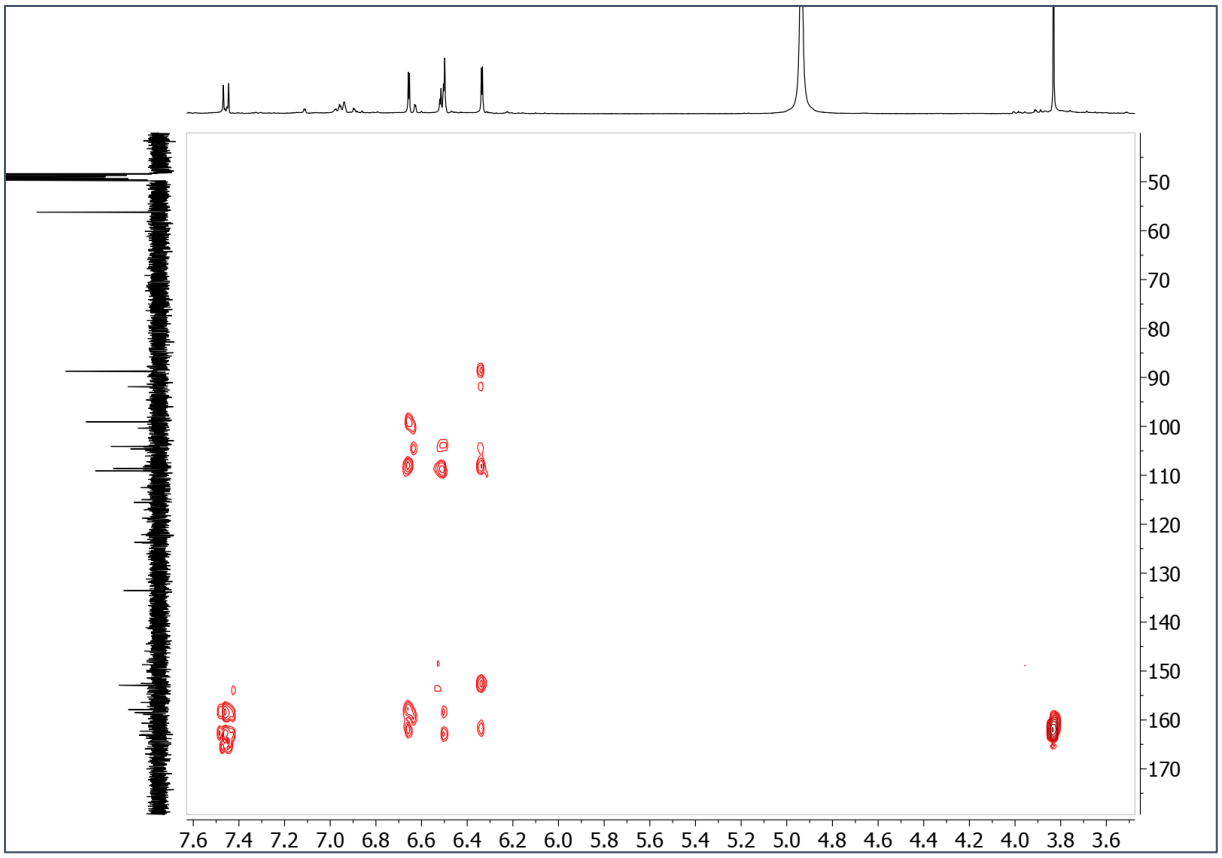


Figure S8. HMBC spectrum of morin-7-*O*-methyl ether (**1**)

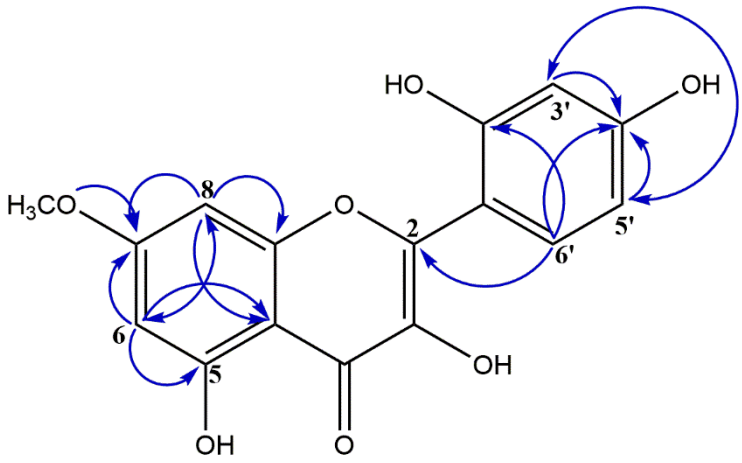


Figure S9. Key HMBC (arrow, H → C) correlations of **1**

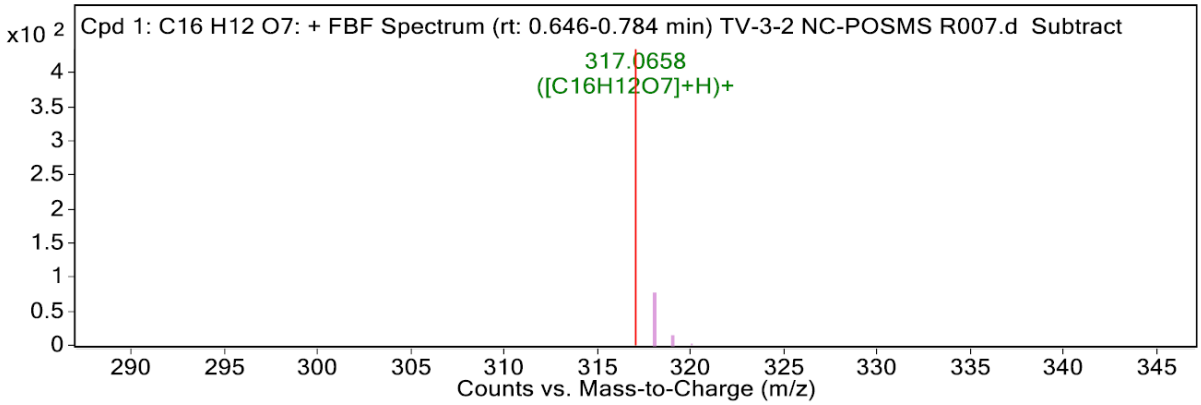


Figure S10. HRESIMS spectrum of morin-7-*O*-methyl ether (**1**)

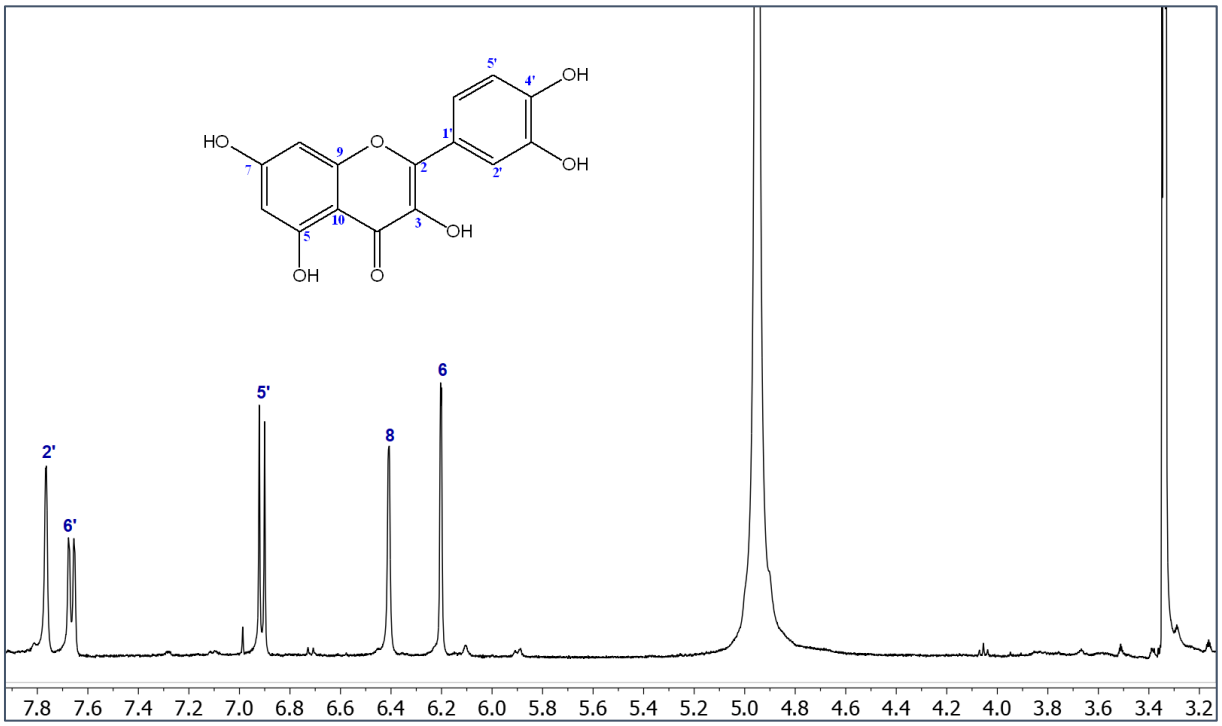


Figure S11. ^{1}H NMR spectrum of quercetin (**2**) (CD_3OD, 400 MHz)

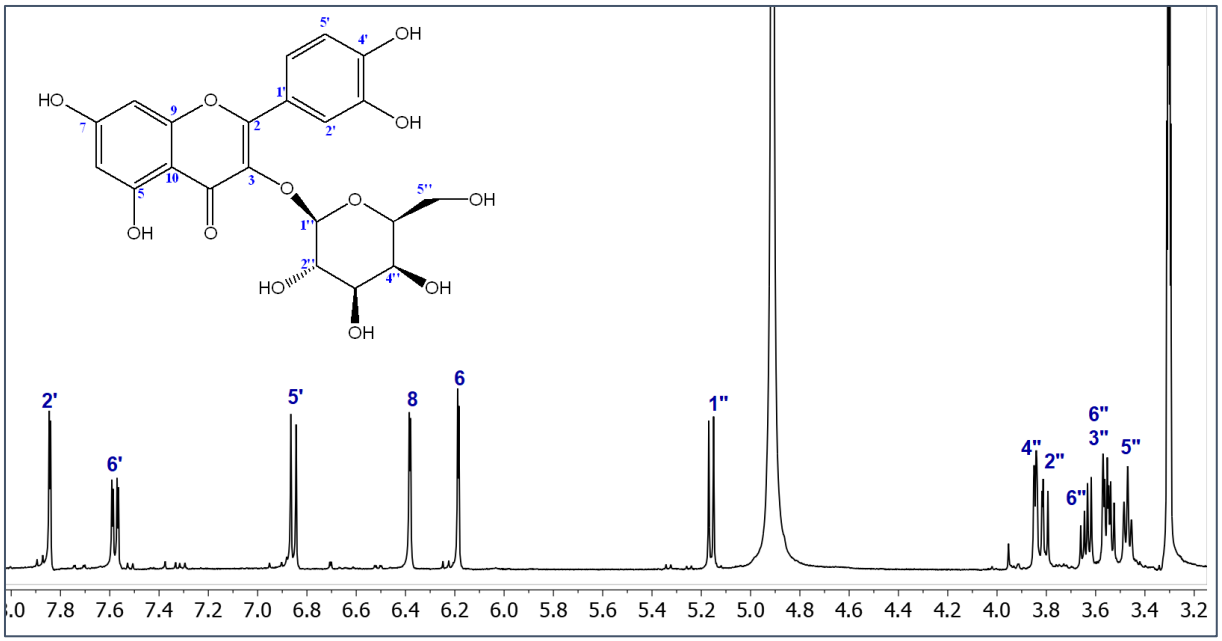


Figure S12. ^{1}H NMR spectrum of hyperoside (**3**) (CD_3OD, 400 MHz)

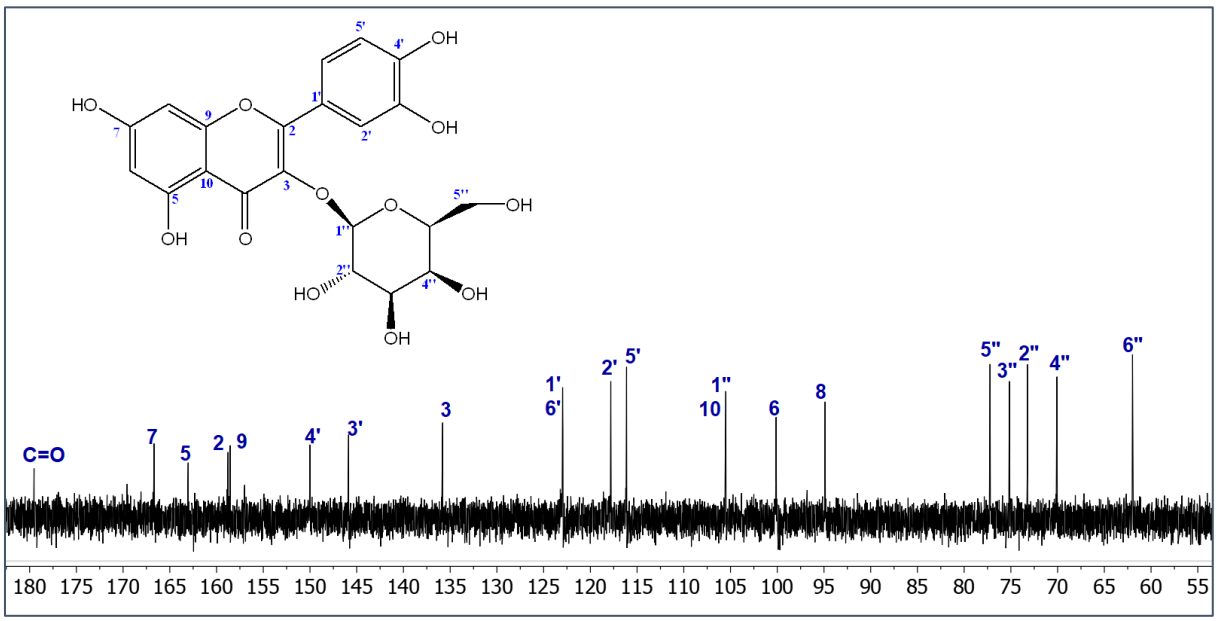


Figure S13. ^{13}C NMR spectrum of hyperoside (**3**) (CD_3OD, 100 MHz)

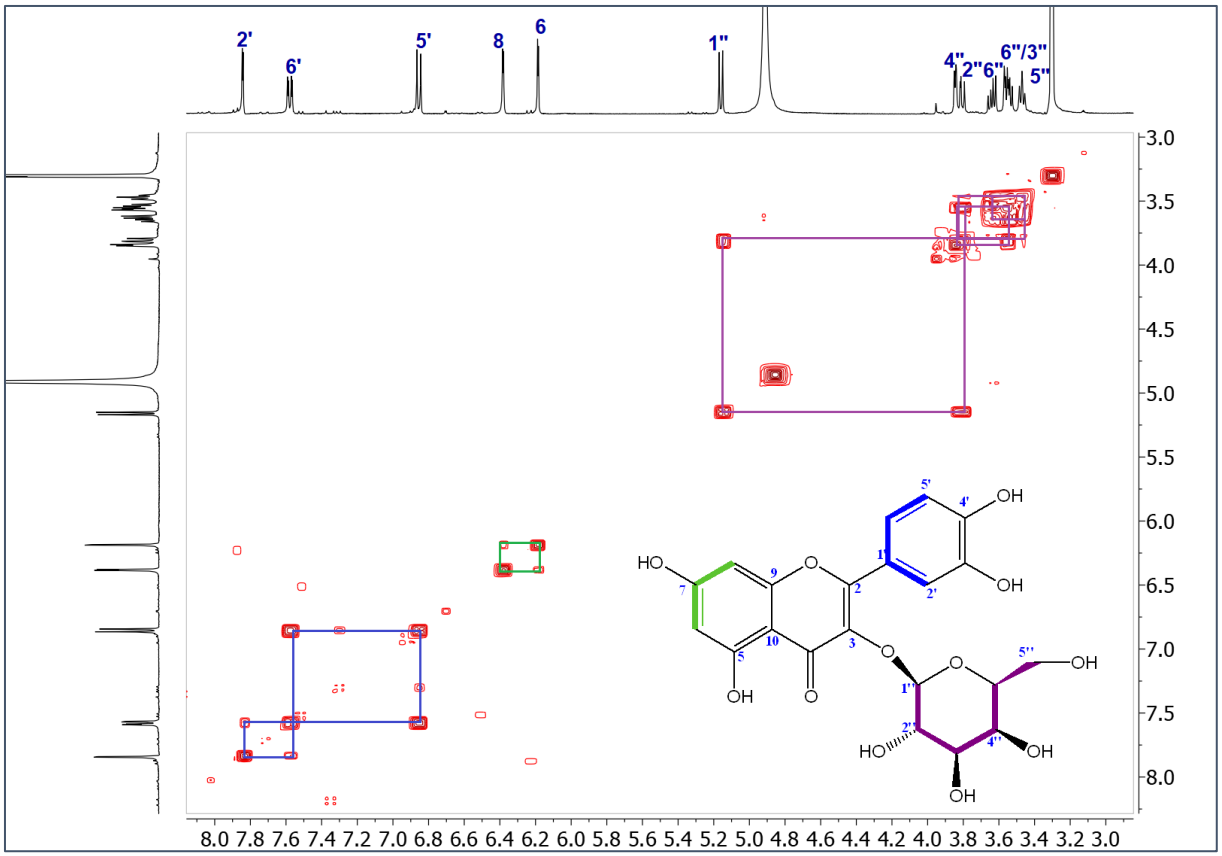


Figure S14. COSY spectrum of hyperoside (**3**)

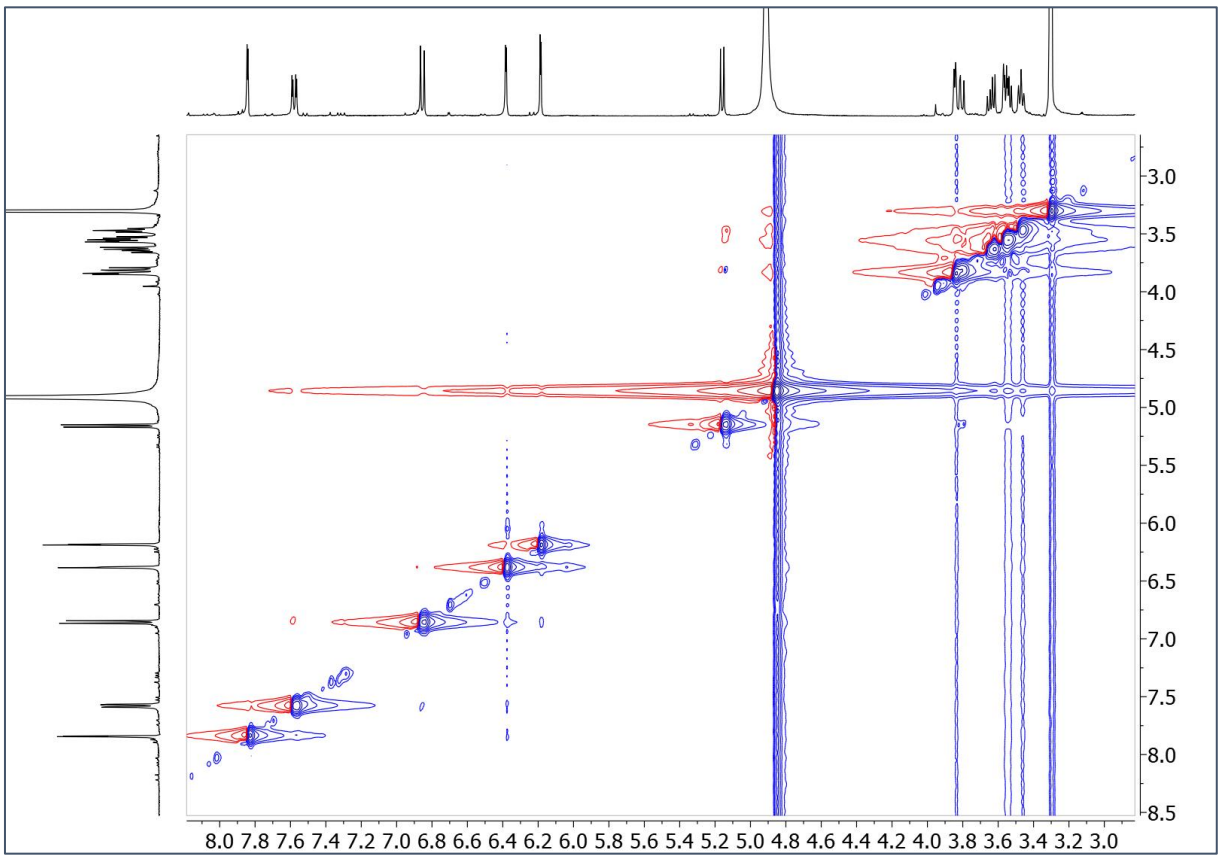


Figure S15. NOESY spectrum of hyperoside (**3**)

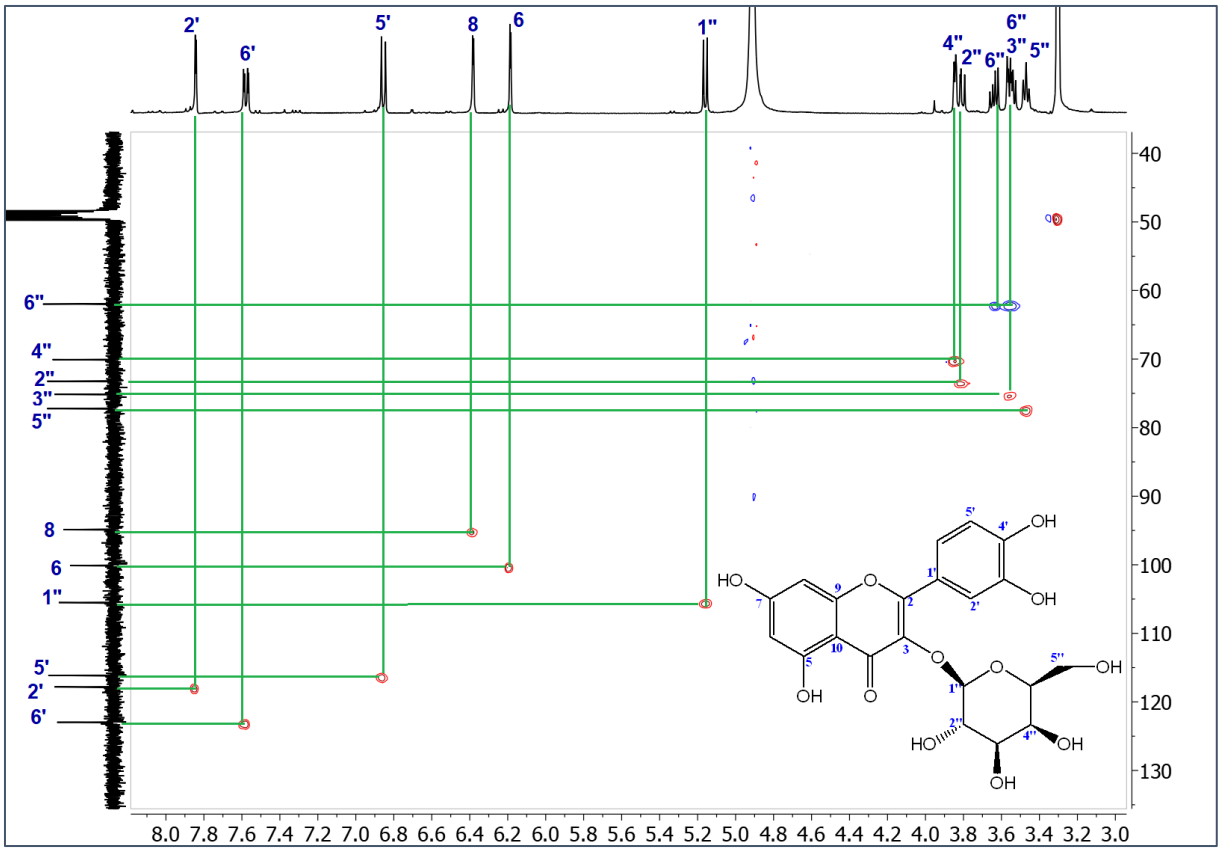


Figure S16. HSQC spectrum of hyperoside (**3**)

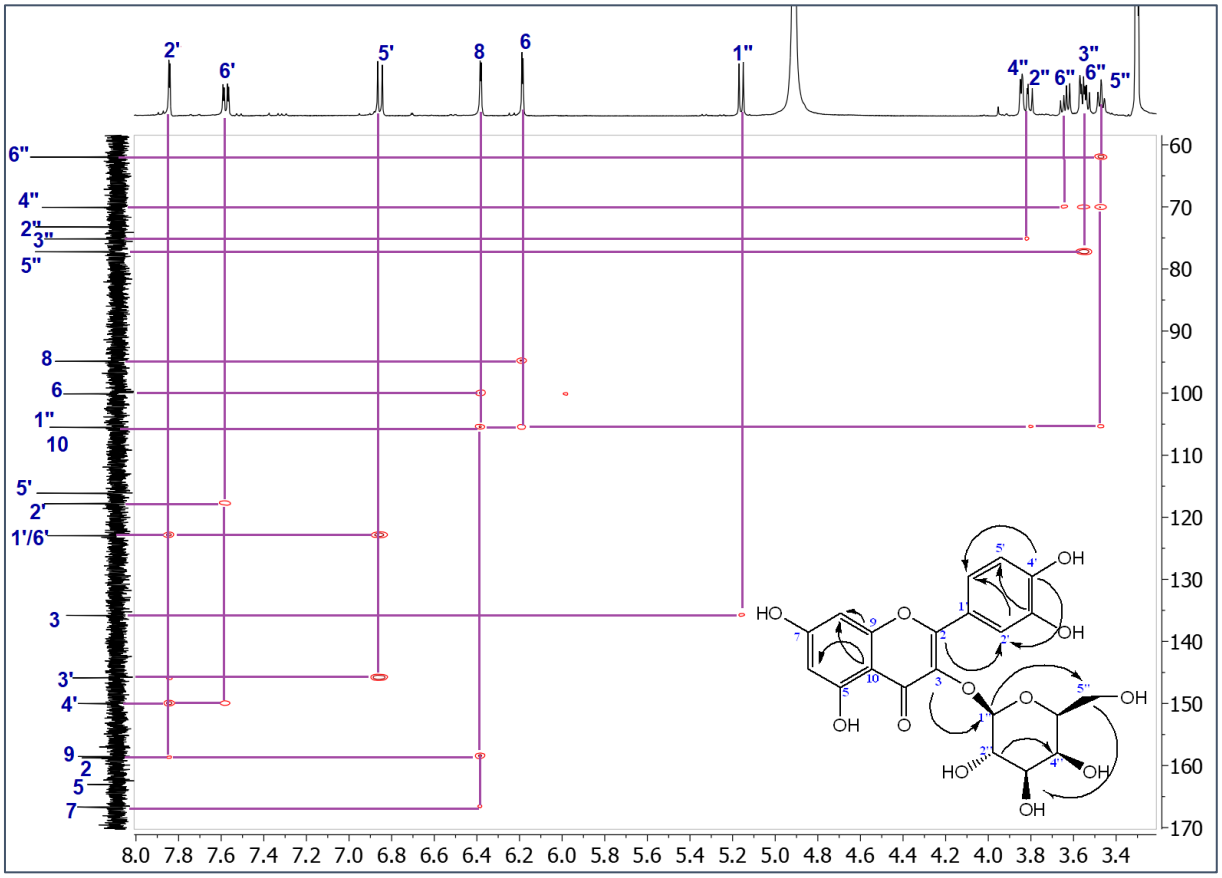


Figure S17. HMBC spectrum of hyperoside (**3**)

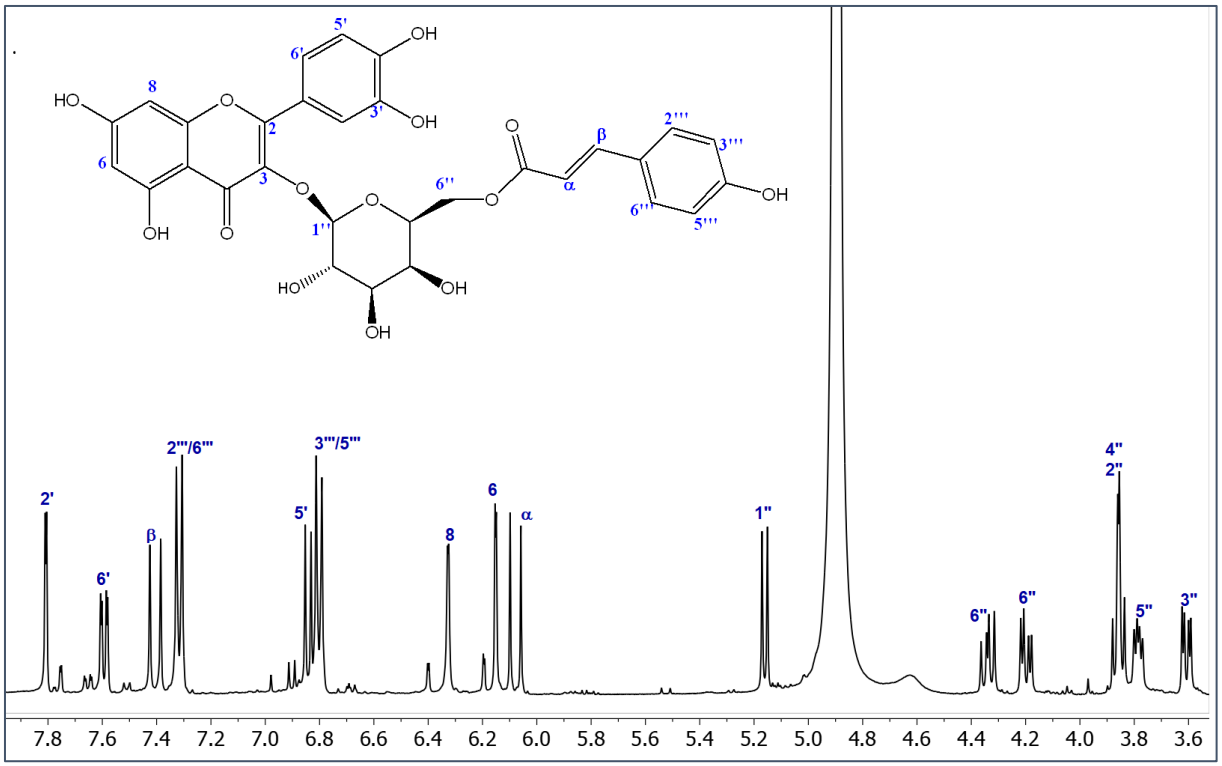


Figure S18. ^{1}H NMR spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**) (CD_3OD, 400 MHz)

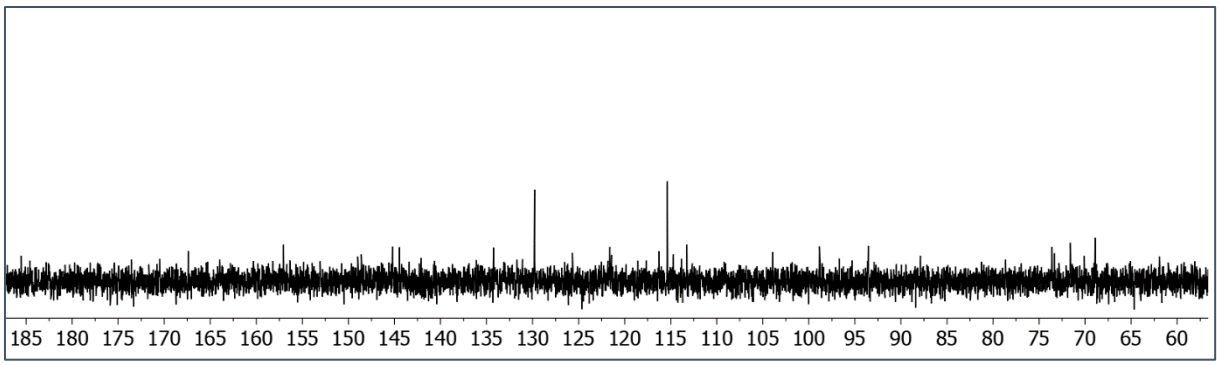


Figure S19. ^{13}C NMR spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**) (CD_3OD, 100 MHz)

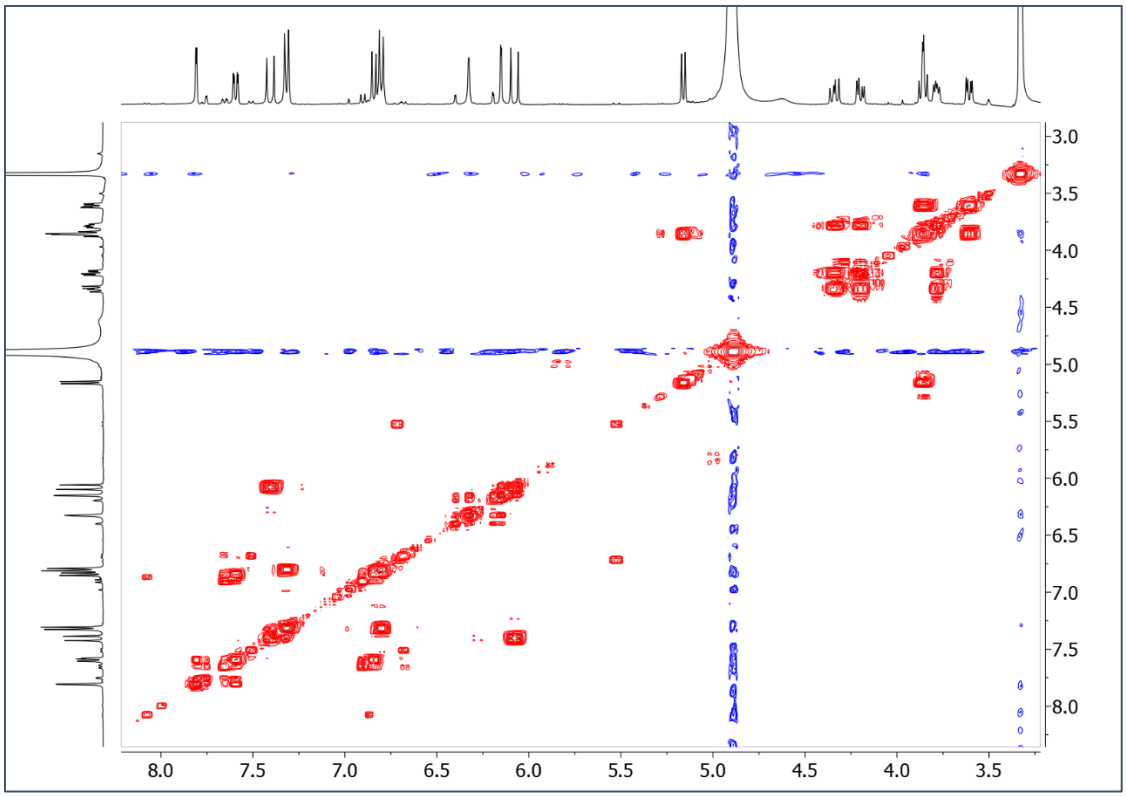


Figure S20. COSY spectrum of quercetin-3-*O*-(6''-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**)

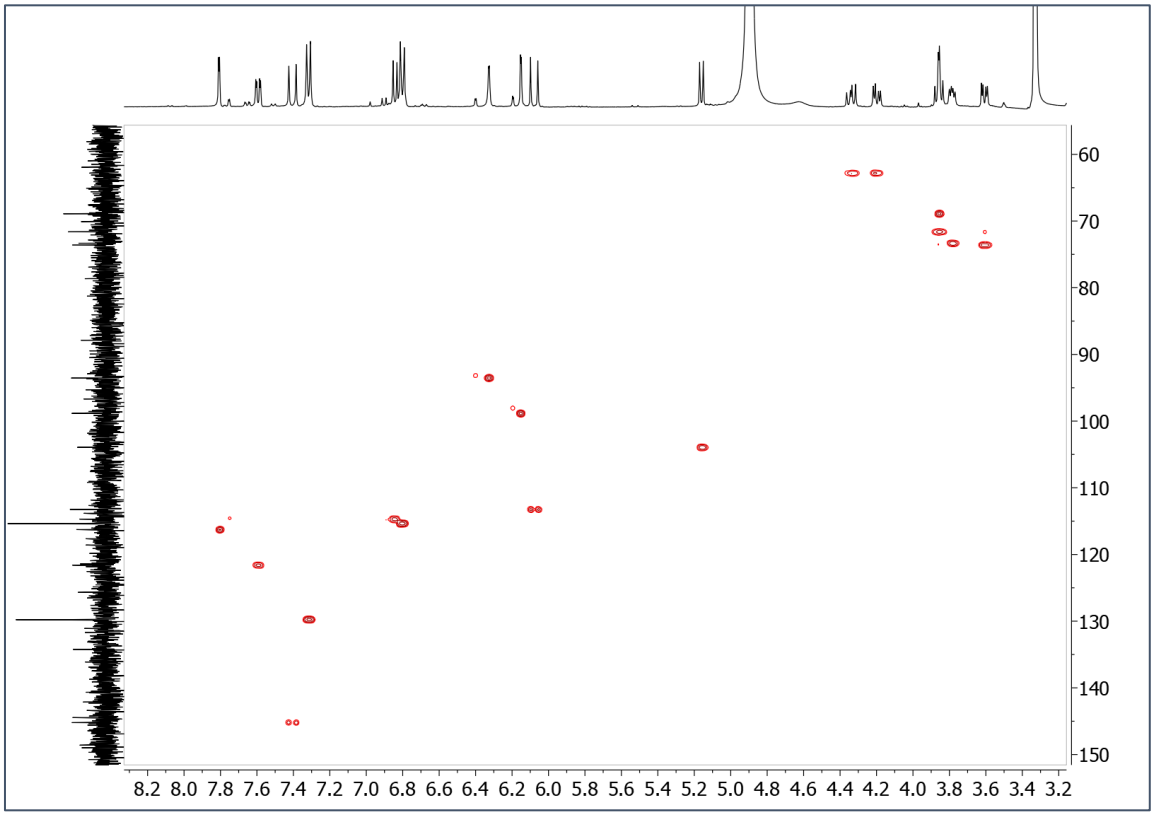


Figure S21. HSQC spectrum of quercetin-3-*O*-(6''-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**)

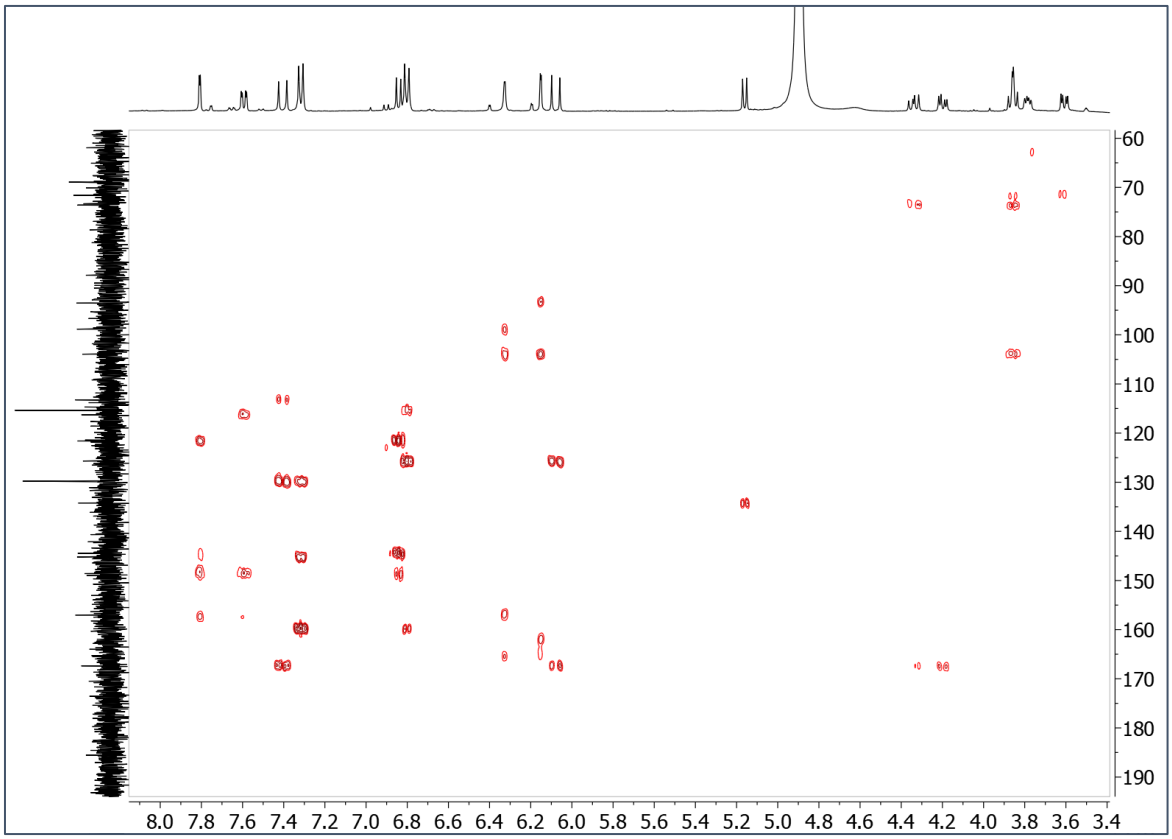


Figure S22. HMBC spectrum of quercetin-3-*O*-(6"-*trans*-*p*-coumaroyl)-*β*-galactoside (**4**)

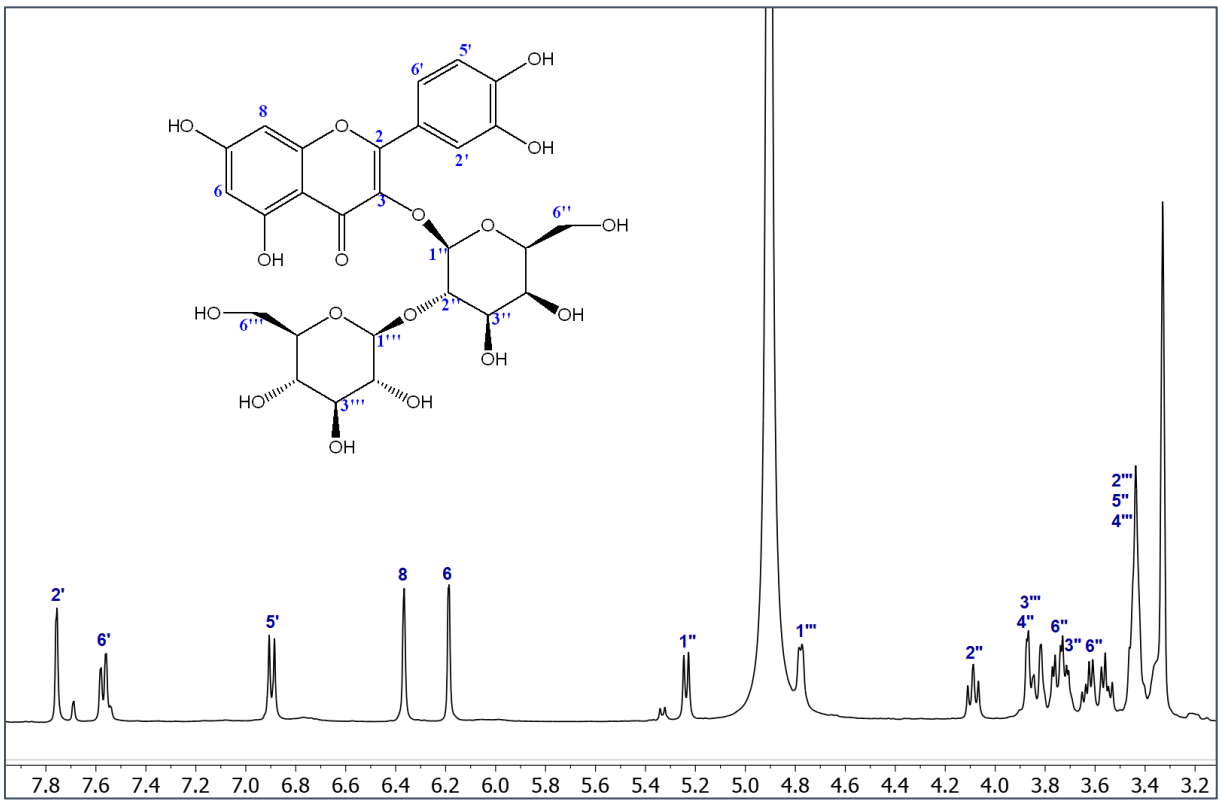


Figure S23. ^{1}H NMR spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) (CD_3OD, 400 MHz)

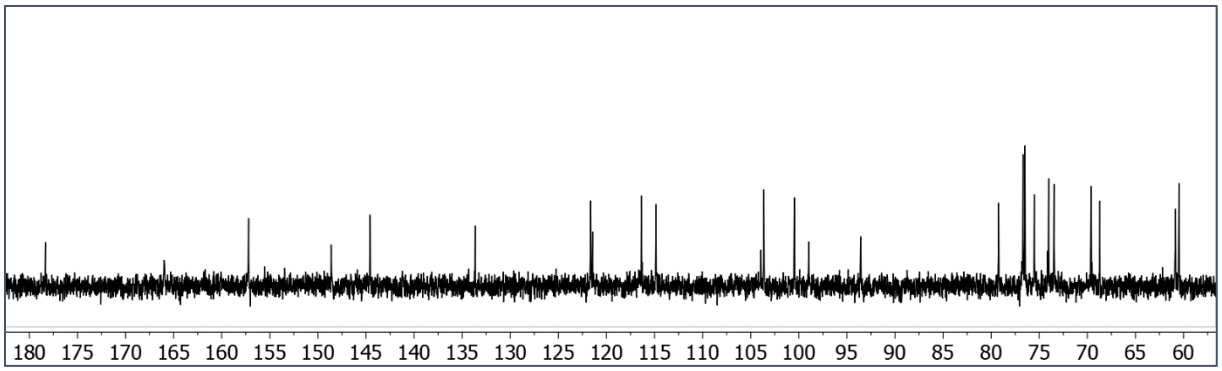


Figure S24. ^{13}C NMR spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**) (CD_3OD, 100 MHz)

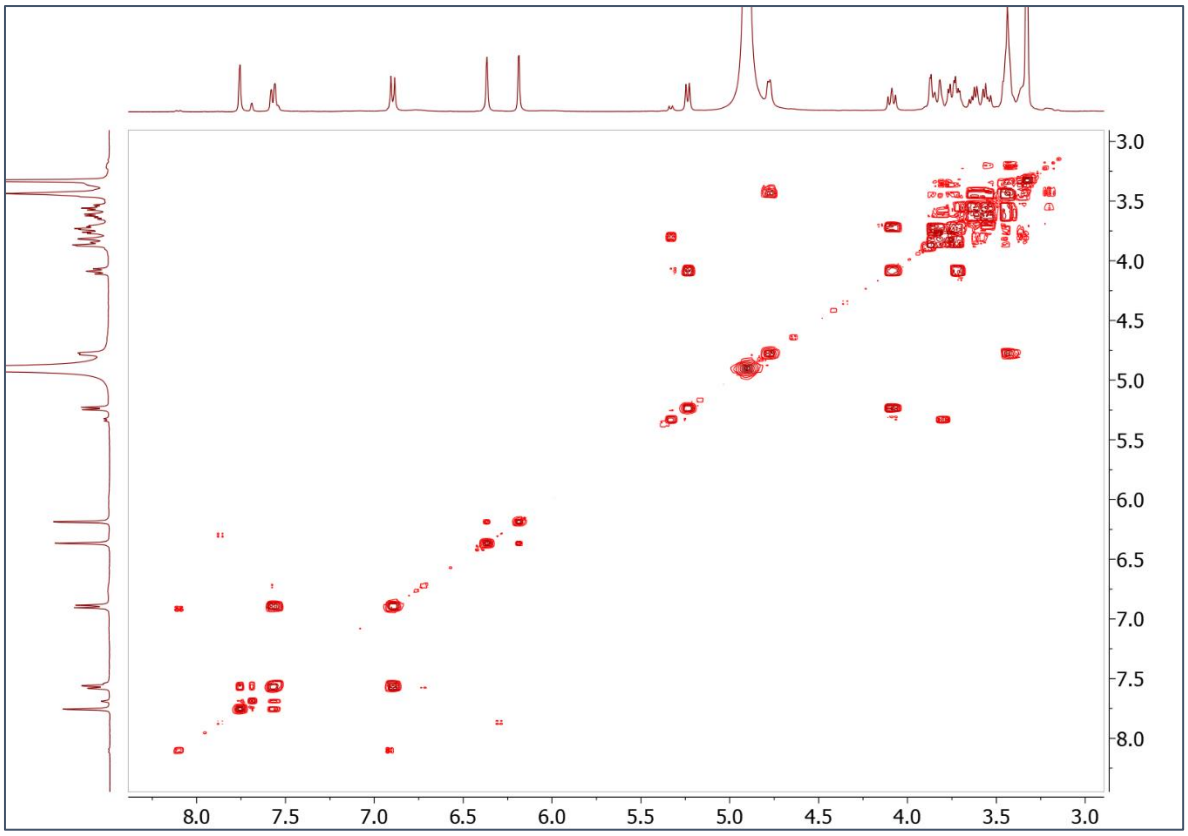


Figure S25. COSY spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**)

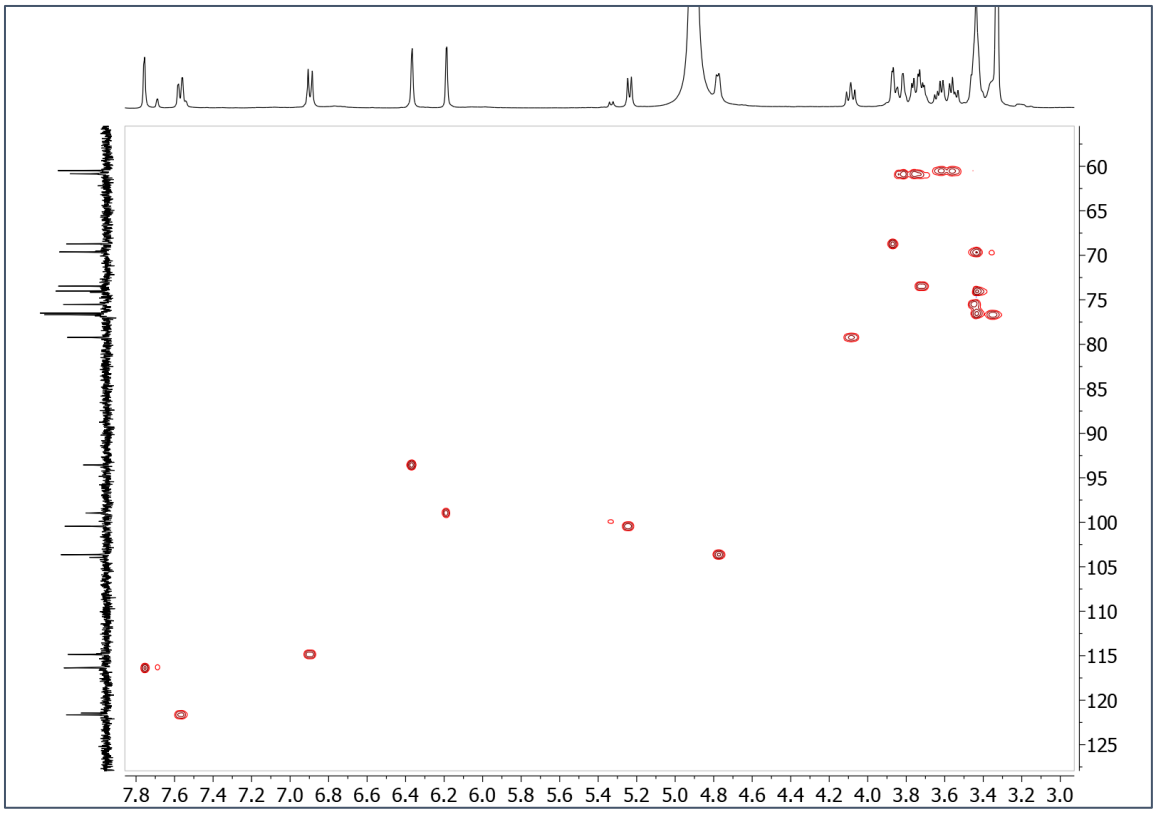


Figure S26. HSQC spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**)

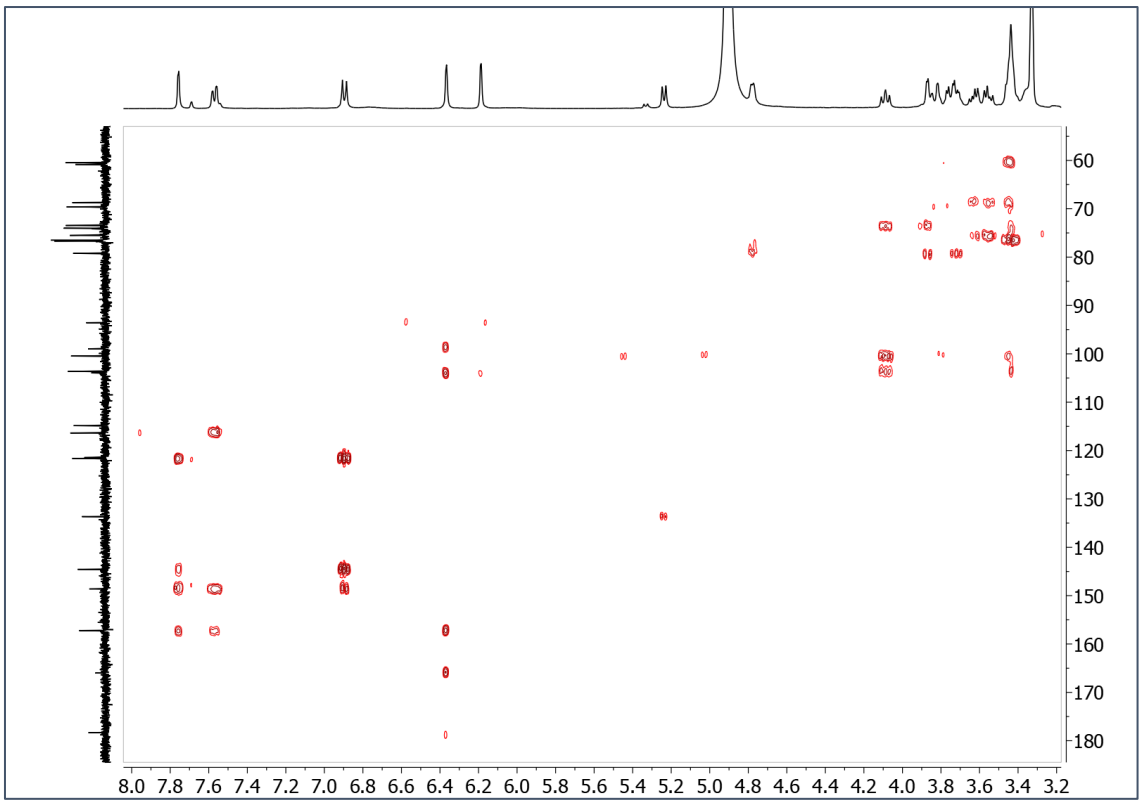


Figure S27. HMBC spectrum of quercetin-3-*O*-(2"-*O*-*β*-glucopyranosyl)-*β*-galactopyranoside (**5**)

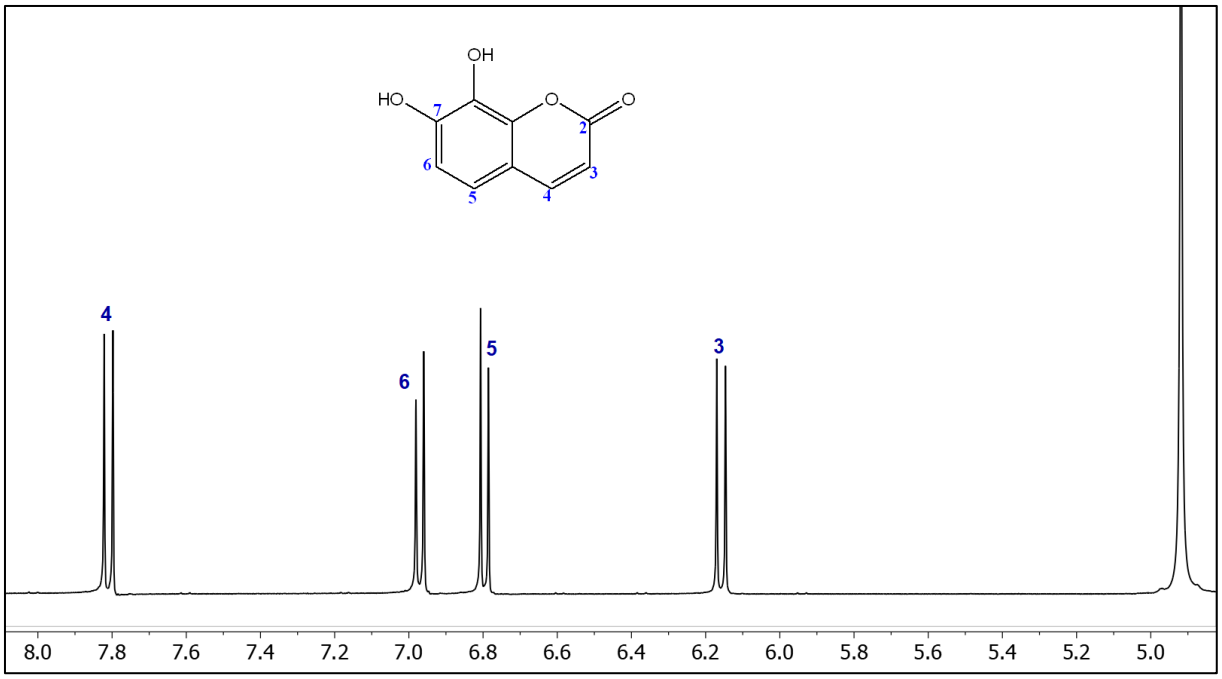


Figure S28. ^{1}H NMR spectrum of daphetin (**6**) (CD_3OD, 400 MHz)

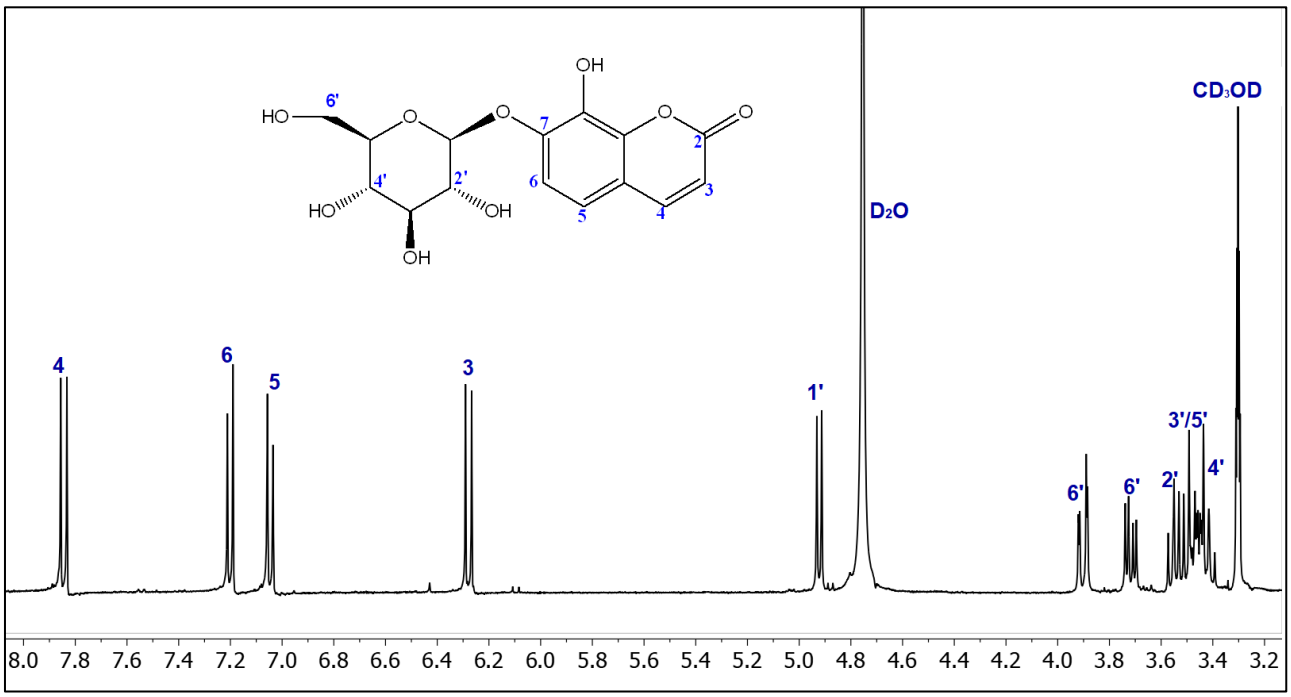


Figure S29. ^{1}H NMR spectrum of daphnin (**7**) (CD_3OD, 400 MHz)

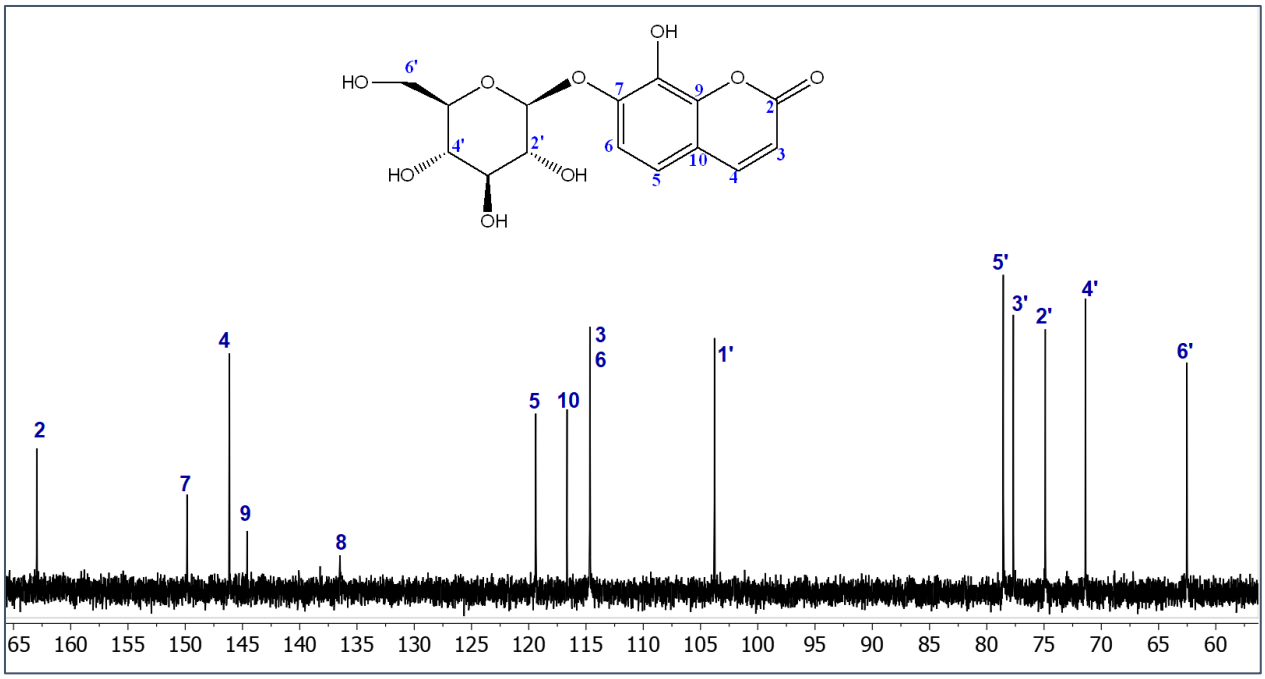


Figure S30. ^{13}C NMR spectrum of daphnin (**7**) (CD_3OD, 100 MHz)

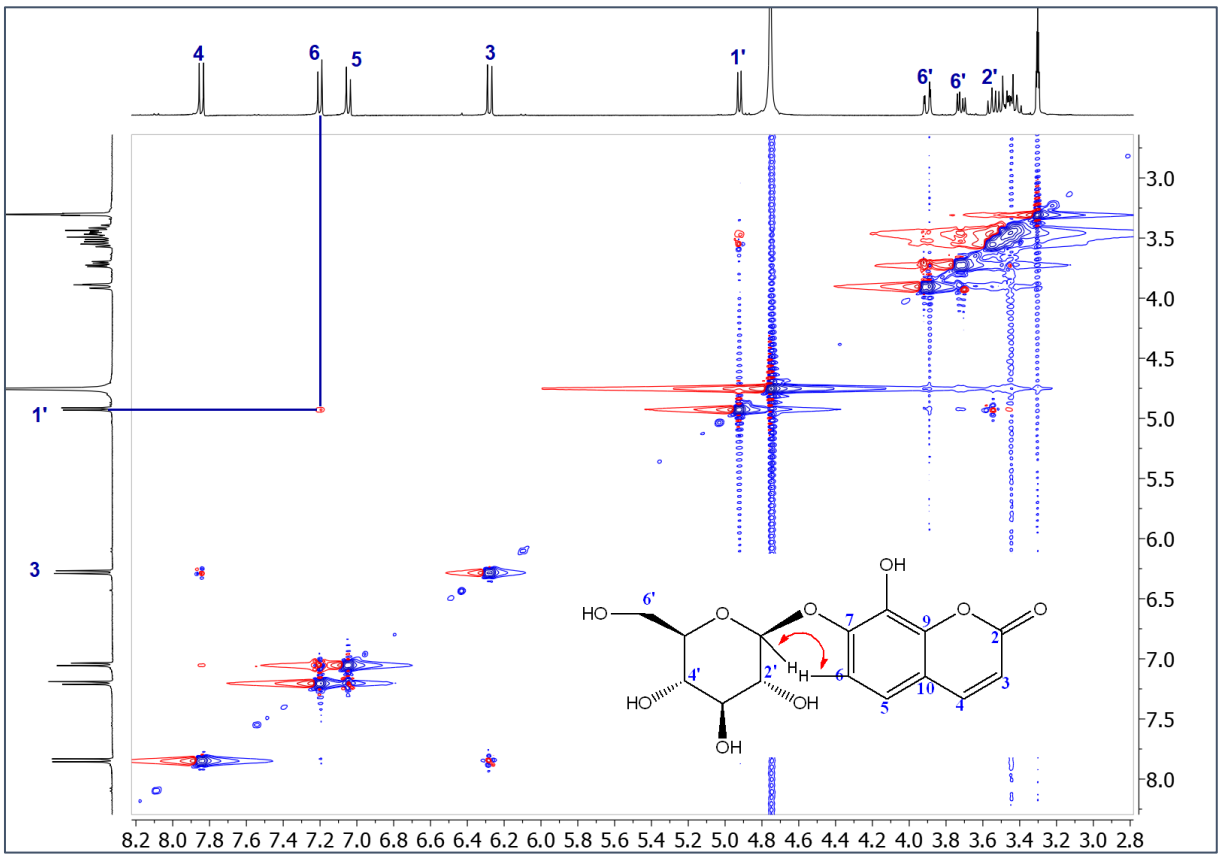


Figure S31. NOESY spectrum of daphnin (**7**)

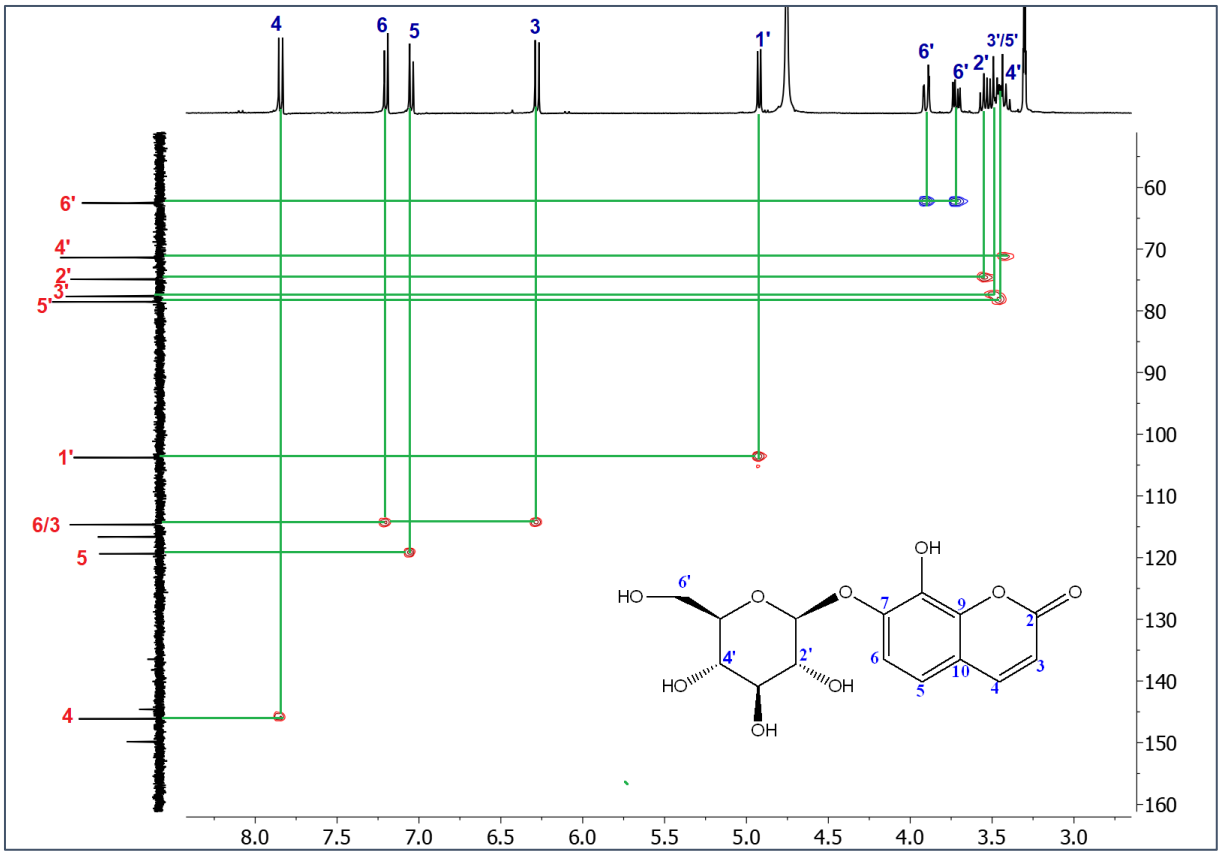


Figure S32. HSQC spectrum of daphnin (**7**)

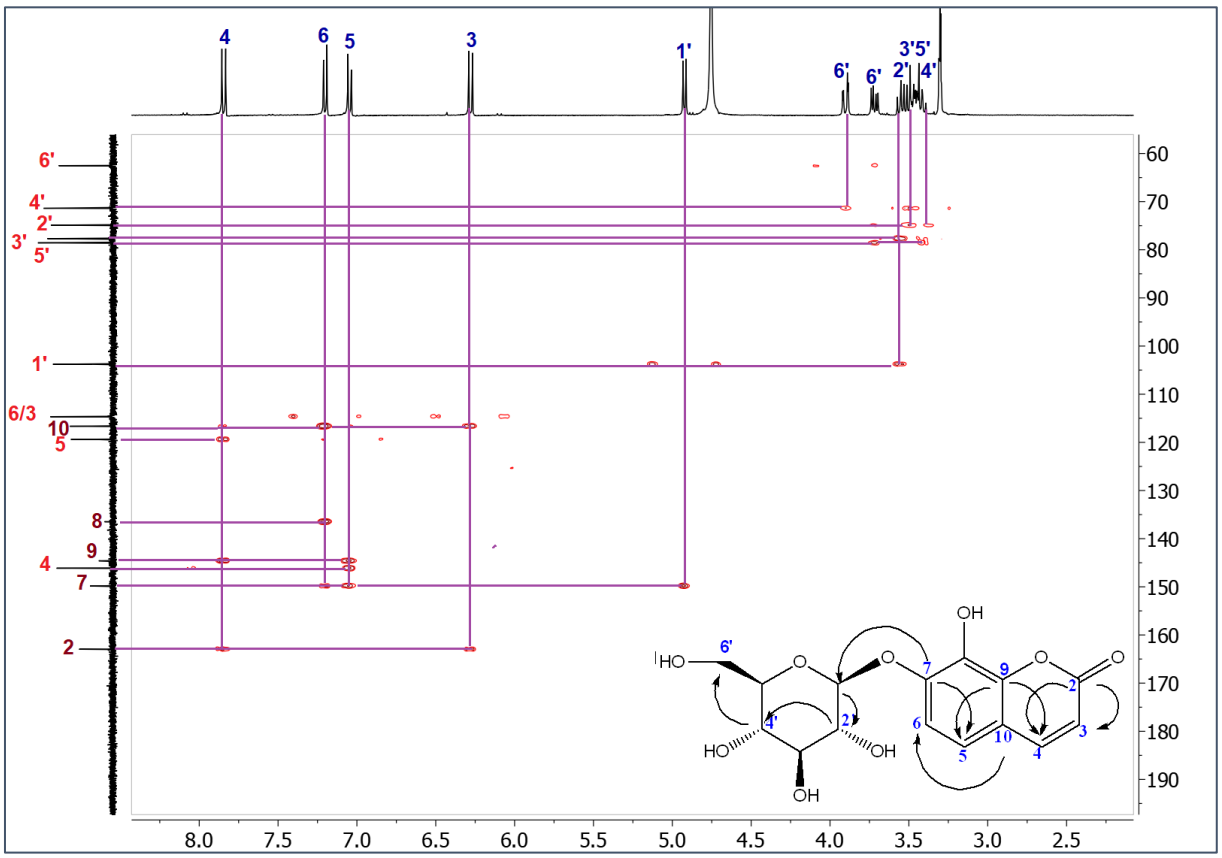


Figure S33. HMBC spectrum of daphnin (**7**)

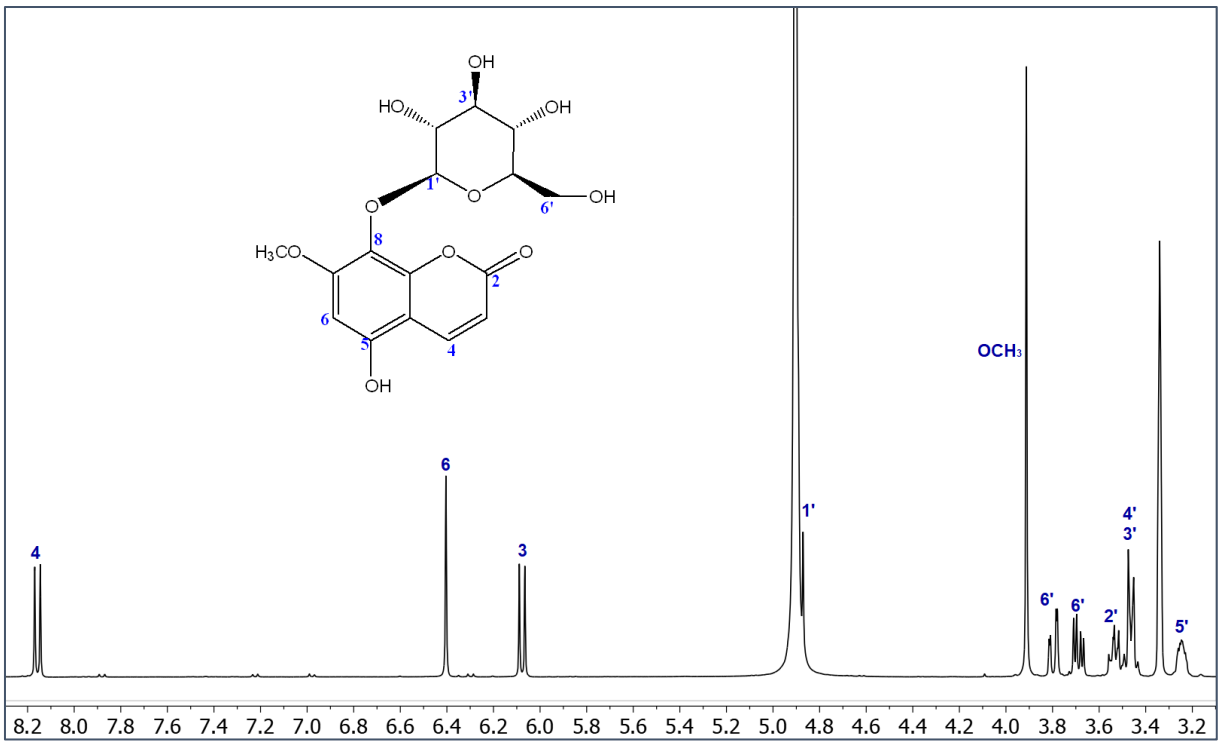


Figure S34. ^{1}H NMR spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) (CD_3OD, 400 MHz)

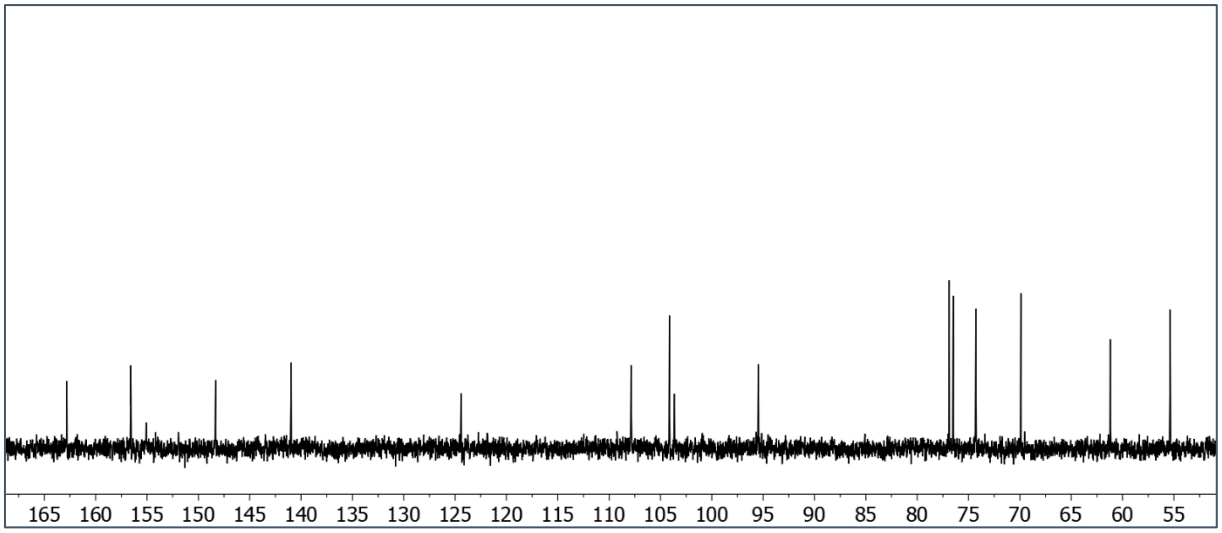


Figure S35. ^{13}C NMR spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**) (CD_3OD, 100 MHz)

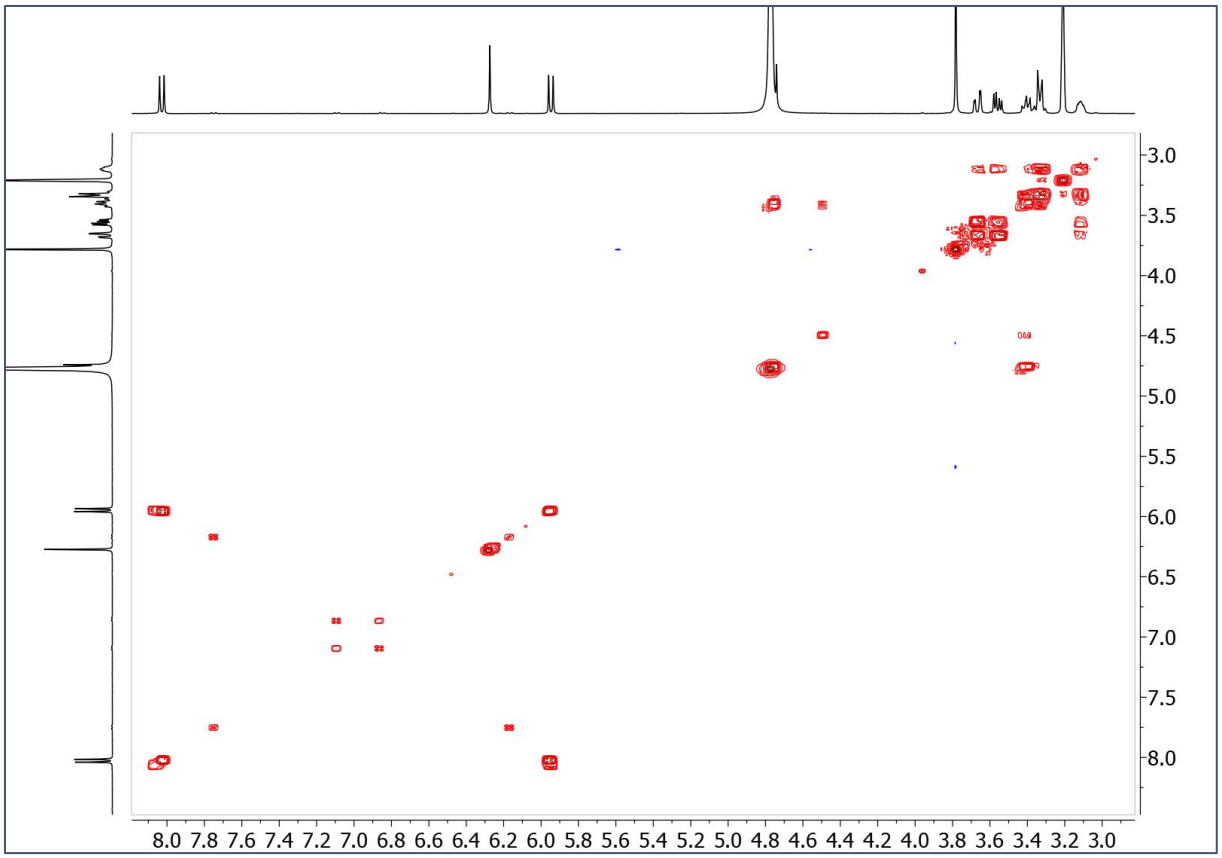


Figure S36. COSY spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**)

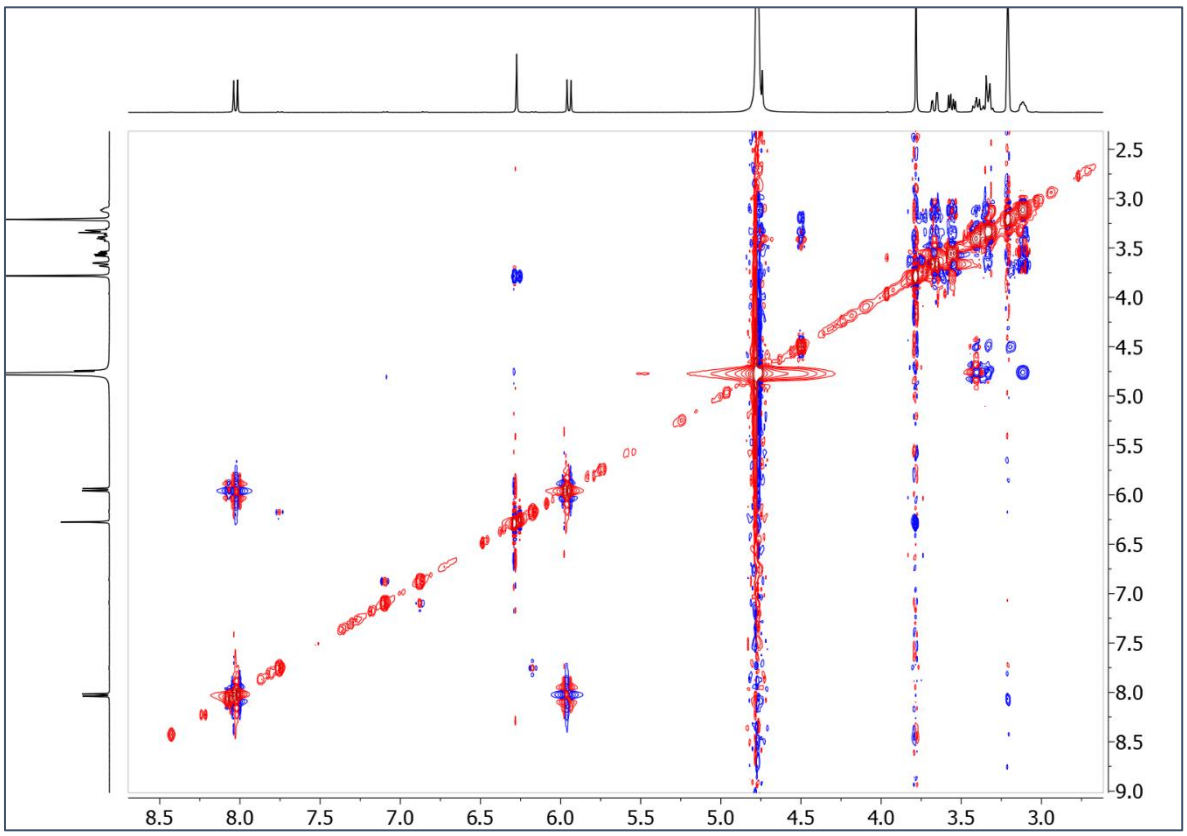


Figure S37. NOESY spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**)

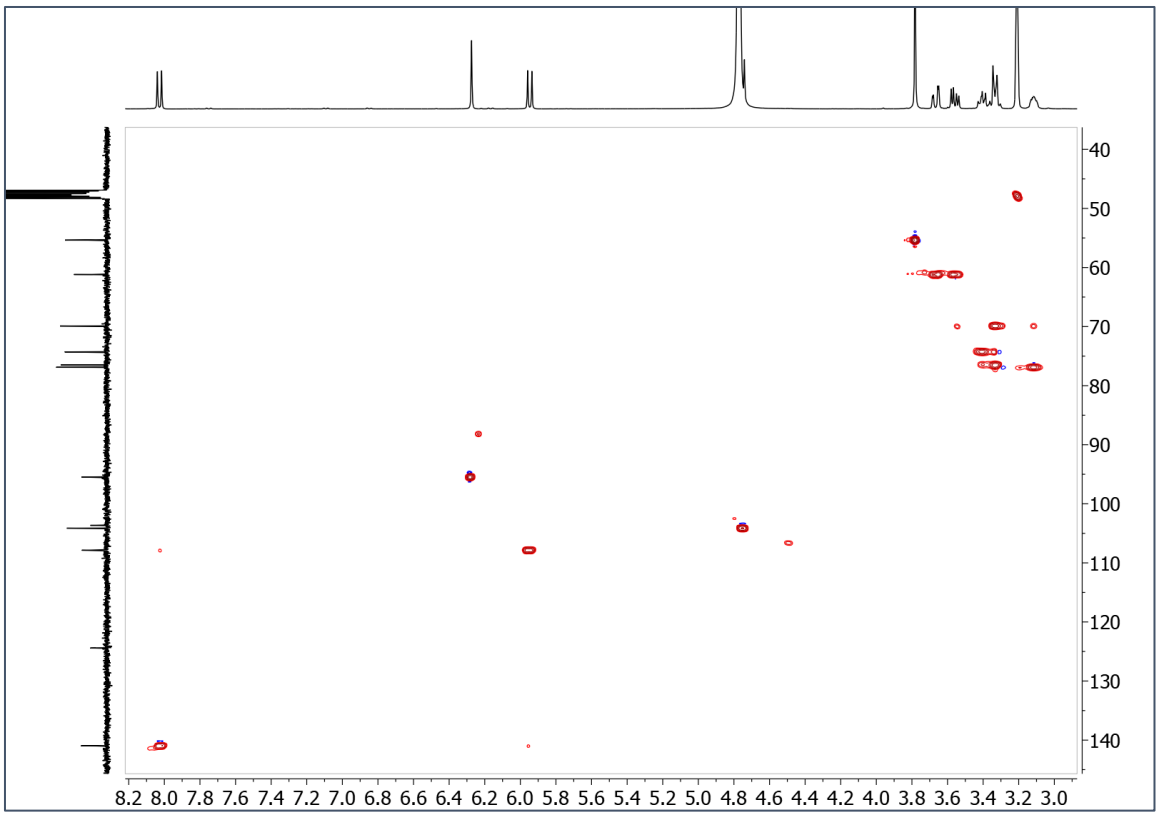


Figure S38. HSQC spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**)

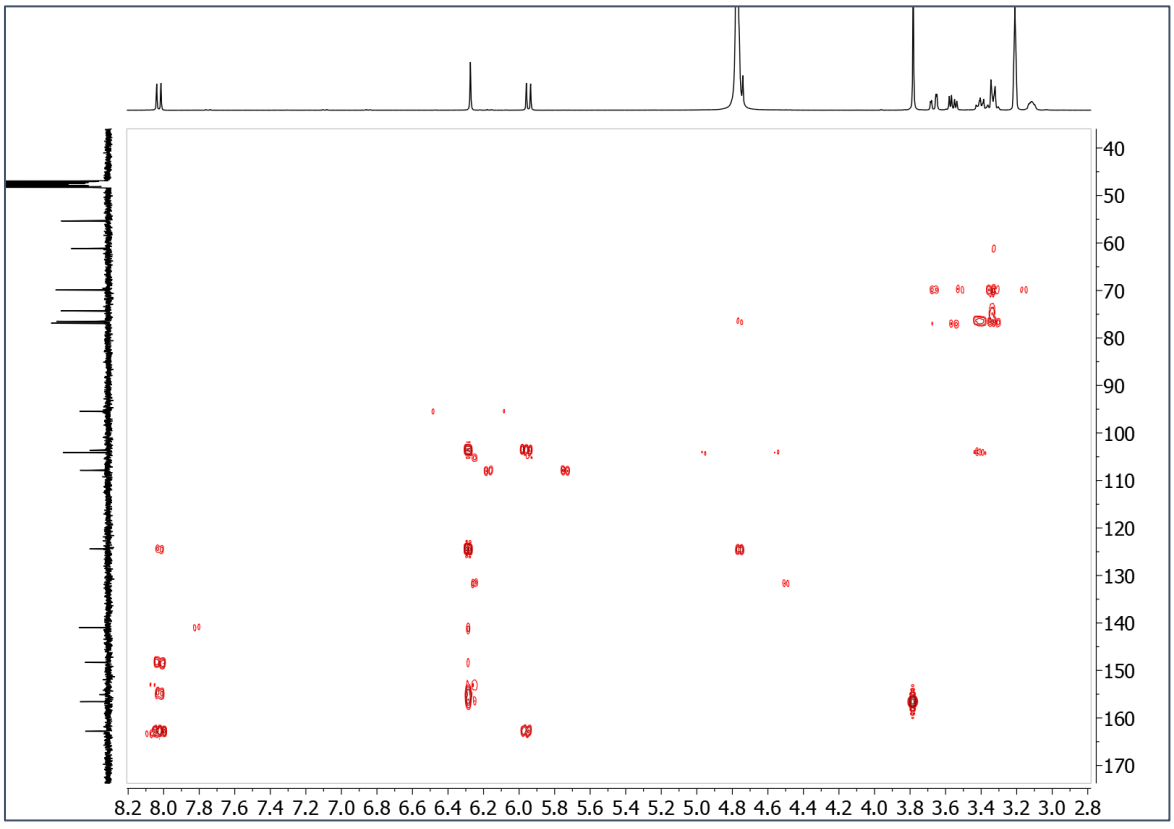


Figure S39. HMBC spectrum of 5-hydroxy-7-methoxy-8-*O*-*β*-glucosyl coumarin (**8**)