Supporting Information

Biocatalytic C-Glucosylation of Coumarins Using an Engineered C-Glycosyltransferase

Dawei Chen, Ridao Chen, Kebo Xie, Tian Yue, Xiaolin Zhang, Fei Ye, and Jungui Dai*

State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China.

*E-mail: jgdai@imm.ac.cn

Table of Contents

1. Experimental Procedures

1.1 Chemicals and Reagents

Chemical reagents were purchased from Sigma Aldrich (St. Louis, MO, USA), J&K Scientific Ltd. (Beijing, China) and InnoChem Science & Technology Co., Ltd. (Beijing, China). Coumarins **15**–**17** were chemically synthesized according to the literature.[1]

1.2 Analytical procedures

The conversion rates of the enzyme reactions were calculated from peak areas of glycosylated products and substrates as analyzed by HPLC. To facilitate inferential statistical analysis, three parallel assays were routinely performed; the means±SD from triplicate analyses are reported here. Enzymatic products were detected on an Agilent 1200 series HPLC system (Agilent Technologies, Germany) coupled with a LCQ Fleet ion trap mass spectrometer (Thermo Electron Corp., USA) equipped with an electrospray ionization (ESI) source. Compounds were characterized by ^{1}H NMR at 400 or 500 MHz and ^{13}C NMR at 125 or 150 MHz on Mercury-400 and Bruker AVIIIHD spectrometers. Chemical shifts (δ) were referenced to internal solvent resonances and given in parts per million (ppm). Coupling constants (*J*) were given in hertz (Hz).

1.3 Expression and purification of CGT mutants

The recombinant plasmid pET-28a-CGTs and empty pET-28a were transformed into *Escherichia coli* BL21 (DE3) (TransGen Biotech, China) for heterologous expression. The described *E. coli* strains were cultivated in 1 L shake flasks at 37 ℃ and 200 rpm using 200 mL LB-media containing 50 μg/mL kanamycin until they reached an optical density at 600 nm of 0.4-0.6. The recombinant N-terminal His_6-CGTs expression was achieved by induction with 0.5 mM isopropyl β-D-thiogalactoside (IPTG) for 16 h at 18 ℃. The cells were harvested by centrifugation at 10,000 g for 5 min at 4 ℃ and resuspended in 20 mL binding buffer (20 mM phosphate buffer, 0.5 M NaCl, 20 mM imidazole, pH 7.4) containing 1 mM phenylmethylsulfonyl fluoride (PMSF). Cells were disrupted by sonication in ice bath and the cell debris was removed by centrifugation at 10,000 g and 4 ℃ for 60 min. The soluble fraction was passed through a 0.45 μm syringe filter unit and the cleared supernatant was immediately applied to 1 mL of Ni-NTA resin (GE, USA) loaded in a column which was pre-equilibrated with binding buffer. The resin

was subsequently eluted with 5 mL (5 CVs) washing buffer (20 mM phosphate buffer, 0.5 M NaCl, 50 mM imidazole, pH 7.4). Elution was carried out with 5 mL (5 CVs) different elution buffer (20 mM phosphate buffer, 0.5 M NaCl, 75–250 mM imidazole, pH 7.4). The protein purification was performed under the flow rate of 1 mL/min and the temperature of 4 ℃. The proteins were concentrated and buffer exchanged to desalting buffer (50 mM NaH_2PO_4-Na_2HPO_4 buffer, 50 mM NaCl, 1 mM DTT, 5% glycerol, pH 8.0) using centrifugal concentrator with Amicon Ultra-30K (Millipore). Protein concentration for all studies was determined by the Protein Quantitative Kit (Bradford) (TransGen Biotech, China). The final protein was flash-frozen in liquid nitrogen and stored at –80 ℃.

1.4 Effects of pH, temperature and divalent metal ions

To study the optimal pH, the enzymatic reaction was performed in various reaction buffers with pH values in the range of 4.0–6.0 (Citric acid-sodium citrate buffer), 6.0–8.0 (Na_2HPO_4-NaH_2PO_4 buffer), 7.0–10.0 (Tirs-HCl buffer) and 9.0–11.0 (Na_2CO_3-$NaHCO_3$ buffer). To assay the optimal reaction temperature, the reactions were incubated at different temperatures (25–60 ℃). To test the necessity of divalent metal ions for MiCGTb-GAGM activity, $MgCl_2$, $BaCl_2$, $NiCl_2$, $MnCl_2$, $CaCl_2$, $CoCl_2$, $FeCl_2$, $ZnCl_2$, $CuCl_2$ and EDTA were used individually in the final concentration of 5 mM. All determations were performed with UDP-glucose (UDP-Glc) as a donor and 5,7-dihydroxycoumarin (**1**) as an acceptor. Three parallel assays were routinely carried out. Aliquots were quenched with ice cold MeOH and centrifuged at 15,000 g for 30 min. The supernatants were performed on a Shiseido capcellpak C18 MG III column (250 mm×4.6 mm I.D., 5 μm, Shiseido Co., Ltd., Tokyo, Japan) at a flow rate of 1 mL/ min at 30 ℃. The mobile phase was a gradient elution of solvents A (methanol) and B (0.1% formic acid aqueous solution). The gradient programs were used for the analyses of the reactions (Table S1).

1.5 Determination of kinetic parameters

For kinetic studies of MiCGTb-GAGM, a typical assay contained 50 mM Na_2HPO_4-NaH_2PO_4 (pH 8.0), saturating UDP-Glc (800 μM) and varying concentration of **1** (50–700 μM) or saturating **1** (800 μM) and varying concentration of UDP-Glc (50–400 μM) at 40 °C in a total volume of 100 μL. Aliquots were quenched with ice cold MeOH and centrifuged

at 15,000 g for 30 min. Supernatants were analyzed by analytical reverse-phase HPLC as described above. All experiments were performed in triplicate. The value of K_m was calculated with the method of Lineweaver-Burk plot.

1.6 HPLC-UV/MS based C-glycosylation activity assay

The reaction mixture contained 0.4 mM UDP-Glc, 0.2 mM coumarin and 100 μg purified enzymes in a final volume of 100 μL were incubated at the optimum condition for 12 h. The reactions were terminated by the addition of 200 μL ice cold MeOH. Subsequently, samples were centrifuged at 15,000 g for 30 min and analyzed by HPLC-UV/MS. For quantification, three parallel assays were routinely carried out. HPLC analyses were performed as described above.

1.7 Whole-cell bioconversion of coumarins

The whole-cell of MiCGTb-GAGM was prepared and induced as described above. After induction, cells were harvested by centrifugation and washed by M9 medium (pH 7.0, Na_2HPO_4 6.8 g L^{-1}; KH_2PO_4 3 g L^{-1}; NaCl 0.5 g L^{-1}; NH_4Cl 1 g L^{-1}; $MgSO_4$ 0.24 g L^{-1}; $CaCl_2$ 0.01 g L^{-1}). The engineered strains were then resuspended by M9 medium containing 2% glucose and the cell density was adjusted to an OD600 of 3.0. Coumarins were added to the reaction mixture at the final concentration of 0.1/0.2/0.4 mM in a total volume of 1 mL, and the bioconversion reactions were kept shaking at 30 ℃ for 24 h. The reactions (100 μL) were terminated by the addition of 200 μL ice cold MeOH. Subsequently, samples were centrifuged at 15,000 g for 30 min and analyzed by HPLC-UV/MS. For quantification, three parallel assays were routinely carried out. HPLC analyses were performed as described above.

1.8 Preparation and structural identification of C-glucosylated coumarins

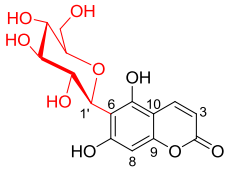
The preparative-scale reactions of coumarin C-glucosides (**1a**, **2a**, **4a**, **4b**, **8a**, **10a**, and **10b**) were carried out with purified MiCGTb-GAGM and UDP-Glc at the optimum condition for 12 h. According to the bioconversion experiment described above, five coumarins (**3**, **7**, **10**–**12**, and **17**) were added to the whole cells of MiCGTb-GAGM to prepare coumarin C-glucosides. All reactions were followed by 40 mL ethyl acetate to extract 5 times. The solvent was evaporated under reduced pressure; the residue was dissolved in 1.5 mL of MeOH and purified by reverse-phase semi-preparative HPLC. The obtained products were solved in CD_3OD-d_4 and

analyzed by ^{1}H NMR and ^{13}C NMR.

1.9 Inhibitory activity assay against SGLT2 for coumarin C-glucosides

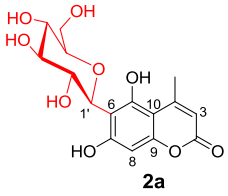
The inhibitory activity assay of SGLT2 for C-glucosides was measured as follows according to our previous work.[2] The plasmid encoding human SGLT2 was stably transfected into HEK293 cells. These cells cultured on 24-well plates were incubated at 37 ℃ in buffer containing 1-NBDG (100 μM with Na^{+}) with various concentrations of C-glycosides. Controls were given equivalent DMSO. After 4 h, the cells were washed twice with pre-chilled phosphate-buffered saline and incubated at room temperature in cell lysis buffer. Fluorescence intensity at 485/535 nm and protein concentration of cell lysis mixture were then measured to determine the inhibitory activity of SGLT2 for C-glucosides. All experiments were repeated in triplicate. IC_{50} values were performed using nonlinear regression with Prism5 software.

2. MS, ^{1}H and ^{13}C NMR data of C-glucosylated coumarins



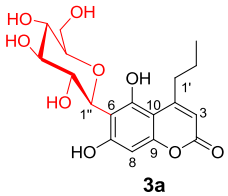
1a

6-(C-β-D-Glucosyl)-5,7-dihydroxycoumarin (1a) (10.7 mg; solid, white): ESI-MS *m/z* 339.09 $[M–H]^-$. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 8.11 (1H, d, *J* = 9.6 Hz, H-4), 6.33 (1H, s, H-8), 6.08 (1H, d, *J* = 9.6 Hz, H-3), 4.96 (1H, d, *J* = 9.8 Hz, H-1'), 3.89-3.81 (2H, m, H-6'), 3.65 (1H, dd, *J* = 8.9, 9.8 Hz, H-2'), 3.60-3.43 (3H, m, H-3'-5'); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 164.0 (C-2), 161.7 (C-7), 157.0 (C-5), 155.7 (C-9), 141.3 (C-4), 110.1 (C-3), 109.4 (C-6), 104.6 (C-10), 95.9 (C-8), 82.6 (C-5'), 79.4 (C-3'), 77.0 (C-1'), 74.4 (C-2'), 70.9 (C-4'), 61.8 (C-6').[3]



2a

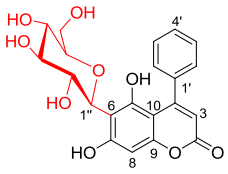
6-(C-β-D-Glucosyl)-4-methyl-5,7-dihydroxycoumarin (2a) (9.5 mg; solid, white): ESI-MS *m/z* 353.20 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{16}H_{19}O_9$ $[M+H]^+$: 355.1024; found: 355.1015. ^{1}H NMR (500 MHz, CD_3CD-d_4): δ = 6.32 (1H, s, H-8), 5.87 (1H, s, H-3), 4.99 (1H, d, *J* = 9.9 Hz, H-1'), 3.90-3.82 (2H, m, H-6'), 3.65 (1H, dd, *J* = 9.0, 9.9 Hz, H-2'), 3.58-3.45 (3H, m, H-3'-5'), 2.60 (3H, s, CH_3); ^{13}C NMR (125 MHz, CD_3CD-d_4): δ = 163.7 (C-2), 160.9 (C-7), 158.1 (C-4), 158.0 (C-5), 157.0 (C-9), 110.4 (C-3), 109.3 (C-6), 104.9 (C-10), 96.2 (C-8), 82.7 (C-5'), 79.3 (C-3'), 77.1 (C-1'), 74.0 (C-2'), 70.9 (C-4'), 61.8 (C-6'), 24.5 (CH_3).[4]



3a

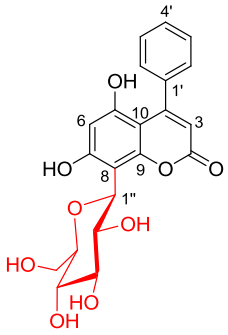
6-(C-β-D-Glucosyl)-4-propyl-5,7-dihydroxycoumarin (3a) (16.6 mg; solid, white): ESI-MS *m/z* 381.25 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{18}H_{23}O_9$ $[M+H]^+$: 383.1337; found: 383.1327. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 6.35 (1H, s, H-8), 5.89 (1H, s, H-3), 5.00 (1H, d, *J* = 9.9 Hz, H-1''), 3.91-3.82 (2H, m, H-6''), 3.67 (1H, dd, *J* = 8.9, 9.9 Hz, H-2''), 3.60-3.45 (3H, m, H-3'-5'), 3.05 (1H,

m, H-1'a), 2.86 (1H, m, H-1'b), 1.69 (2H, m, H-2'), 1.01 (3H, t, *J* = 7.4 Hz, H-3'); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 163.8 (C-2), 161.8 (C-7), 160.7 (C-4), 157.5 (C-5), 157.4 (C-9), 109.9 (C-3), 109.5 (C-6), 104.3 (C-10), 96.5 (C-8), 82.8 (C-5"), 79.3 (C-3"), 77.2 (C-1"), 74.1 (C-2"), 70.9 (C-4"), 61.8 (C-6"), 39.3 (C-1'), 24.1 (C-2'), 14.3 (C-3').[5]



4a

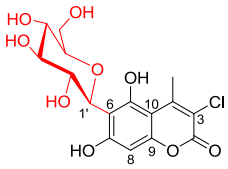
6-(C-β-D-Glucosyl)-4-phenyl-5,7-dihydroxycoumarin (4a) (10.4 mg; solid, white): ESI-MS *m/z* 415.35 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{21}H_{21}O_9$ $[M+H]^+$: 417.1180; found: 417.1171. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 7.38-7.34 (5H, m, H-2'-6'), 6.43 (1H, s, H-8), 5.83 (1H, s, H-3), 4.93 (1H, d, *J* = 9.8 Hz, H-1"), 3.79 (1H, dd, *J* = 1.7, 12.1 Hz, H-6"a), 3.70 (1H, dd, *J* = 4.2, 12.1 Hz, H-6"b), 3.59 (1H, m, H-2"), 3.48-3.40 (3H, m, H-3"-5"); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 163.2 (C-2), 161.5 (C-7), 158.9 (C-5), 157.4 (C-4), 157.1 (C-9), 141.2 (C-1'), 129.1 (C-2', 6'), 128.6 (C-3', 5'), 128.5 (C-4'),112.2 (C-3), 109.5 (C-6), 103.5 (C-10), 96.4 (C-8), 82.8 (C-5"), 79.4 (C-3"), 77.0 (C-1"), 73.9 (C-2"), 71.0 (C-4"), 61.9 (C-6").[5]



4b

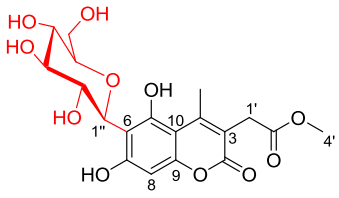
8-(C-β-D-Glucosyl)-4-phenyl-5,7-dihydroxycoumarin (4b) (1.0 mg; solid, white): ESI-MS *m/z* 415.26 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{21}H_{21}O_9$ $[M+H]^+$: 417.1180; found: 417.1173. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 7.37-7.29 (5H, m, H-2'-6'), 6.23 (1H, s, H-6), 5.83 (1H, s, H-3), 5.03 (1H, d, *J* = 9.7 Hz, H-1"), 3.89 (1H, dd, *J* = 2.1, 12.0 Hz, H-6"a), 3.75 (1H, dd, *J* = 5.0, 12.0 Hz, H-6"b), 3.53-3.48 (4H, m, H-2"-5"); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 163.2 (C-2), 162.7 (C-7), 159.4 (C-5), 158.4 (C-4), 158.3 (C-9), 141.4 (C-1'), 128.9 (C-2', 6'), 128.4 (C-3'-5'), 111.1

(C-3), 105.2 (C-8), 102.9 (C-10), 100.8 (C-6), 82.6 (C-5"), 80.1 (C-3"), 75.9 (C-1"), 73.2 (C-2"), 71.8 (C-4"), 62.9 (C-6").[5,6]



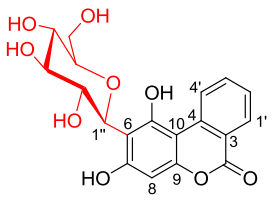
7a

6-(C-β-D-Glucosyl)-3-chloro-4-methyl-5,7-dihydroxycoumarin (7a) (9.4 mg; solid, white): ESI-MS *m/z* 387.29 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{16}H_{18}O_9Cl$ $[M+H]^+$: 389.0634; found: 389.0624. ^1H NMR (400 MHz, CD_3CD-d_4): δ = 6.33 (1H, s, H-8), 5.00 (1H, d, J = 9.9 Hz, H-1'), 3.91-3.82 (2H, m, H-6'), 3.66 (1H, dd, J = 8.8, 9.9 Hz, H-2'), 3.60-3.45 (3H, m, H-3'-5'), 2.79 (3H, s, CH_3); ^{13}C NMR (125 MHz, CD_3CD-d_4): δ = 160.8 (C-2), 159.2 (C-7), 157.7 (C-4), 154.7 (C-5), 152.6 (C-9), 116.1 (C-3), 109.8 (C-6), 104.7 (C-10), 96.1 (C-8), 82.7 (C-5'), 79.2 (C-3'), 77.1 (C-1'), 74.0 (C-2'), 70.9 (C-4'), 61.8 (C-6'), 20.8 (CH_3).[5]



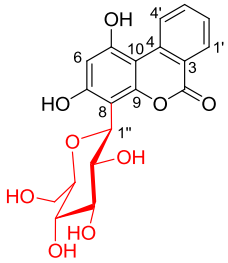
8a

6-(C-β-D-Glucosyl)-3-(methyl acetate)-4-methyl-5,7-dihydroxycoumarin (8a) (2.3 mg; solid, white): ESI-MS *m/z* 425.24 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{19}H_{23}O_{11}$ $[M+H]^+$: 427.1235; found: 427.1225. ^1H NMR (400 MHz, CD_3CD-d_4): δ = 6.33 (1H, s, H-8), 5.01 (1H, d, J = 9.9 Hz, H-1"), 3.90-3.81 (2H, m, H-6"), 3.67-3.48 (4H, m, H-2"-5"), 3.70 (3H, s, H-4'), 3.69 (2H, s, H-1'), 2.61 (3H, s, CH_3); ^{13}C NMR (125 MHz, CD_3CD-d_4): δ = 173.3 (C-2'), 163.8 (C-2), 160.3 (C-7), 158.0 (C-4), 155.7 (C-5), 154.4 (C-9), 115.1 (C-3), 109.6 (C-6), 105.2 (C-10), 96.0 (C-8), 82.8 (C-5"), 79.3 (C-3"), 77.2 (C-1"), 74.0 (C-2"), 70.9 (C-4"), 61.8 (C-6"), 52.6 (C-4'), 33.0 (C-1'), 20.1 (CH_3).[5,7]



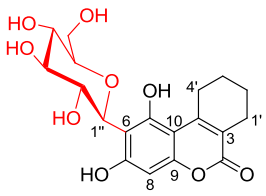
10a

6-(C-β-D-Glucosyl)-3,4-benzo-5,7-dihydroxycoumarin (10a) (12.3 mg; solid, white): ESI-MS *m/z* 389.18 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{19}H_{19}O_9$ $[M+H]^+$: 391.1024; found: 391.1010. ^1H NMR (400 MHz, CD_3CD-d_4): δ = 9.09 (1H, dd, *J* = 0.5, 8.5 Hz, H-1'), 8.27 (1H, dd, *J* = 1.5, 7.9 Hz, H-4'), 7.77 (1H, ddd, *J* = 1.5, 7.3, 8.5 Hz, H-2'), 7.47 (1H, ddd, *J* = 0.5, 7.3, 7.9 Hz, H-3'), 6.41 (1H, s, H-8), 5.07 (1H, d, *J* = 9.9 Hz, H-1"), 3.94-3.85 (2H, m, H-6"), 3.73 (1H, m, H-2"), 3.65-3.48 (3H, m, H-3"-5"); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 163.6 (C-2), 158.7 (C-7), 157.9 (C-5), 154.1 (C-9), 137.1 (C-4), 136.1 (C-3'), 130.7 (C-1'), 127.8 (C-2'), 127.6 (C-4'), 120.3 (C-3), 109.7 (C-6), 101.8 (C-10), 96.8 (C-8), 82.8 (C-5"), 79.4 (C-3"), 77.4 (C-1"), 74.2 (C-2"), 71.0 (C-4"), 61.8 (C-6").[8]



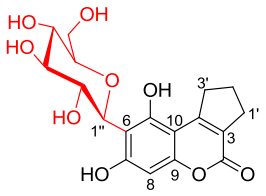
10b

8-(C-β-D-Glucosyl)-3,4-benzo-5,7-dihydroxycoumarin (10b) (7.2 mg; solid, white): ESI-MS *m/z* 389.36 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{19}H_{19}O_9$ $[M+H]^+$: 391.1024; found: 391.1016. ^1H NMR (400 MHz, CD_3CD-d_4): δ = 9.10 (1H, d, *J* = 8.6 Hz, H-1'), 8.28 (1H, dd, *J* = 1.6, 8.0 Hz, H-4'), 7.78 (1H, ddd, *J* = 1.6, 7.2, 8.6 Hz, H-2'), 7.47 (1H, ddd, *J* = 1.0, 7.2, 8.0 Hz, H-3'), 6.44 (1H, s, H-6), 5.08 (1H, d, *J* = 9.7 Hz, H-1"), 3.89 (1H, dd, *J* = 2.1, 12.0 Hz, H-6"a), 3.76 (1H, dd, *J* = 4.9, 12.0 Hz, H-6"b), 3.55-3.49 (4H, m, H-2"-5"); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 163.4 (C-2), 159.7 (C-7), 158.8 (C-5), 154.7 (C-9), 137.5 (C-4), 136.0 (C-3'), 130.4 (C-1'), 127.9 (C-2'), 127.5 (C-4'), 119.9 (C-3), 105.4 (C-8), 101.1 (C-10), 98.1 (C-6), 82.6 (C-5"), 80.0 (C-3"), 76.2 (C-1"), 73.3 (C-2"), 71.7 (C-4"), 62.8 (C-6").[6,8]



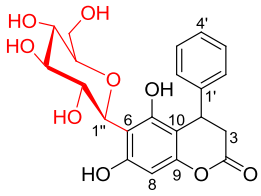
11a

6-(C-β-D-Glucosyl)-3,4-cyclahexyl-5,7-dihydroxycoumarin (11a) (3.4 mg; solid, white): ESI-MS *m/z* 393.22 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{19}H_{23}O_9$ $[M+H]^+$: 395.1337; found: 395.1332. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 6.30 (1H, s, H-8), 5.00 (1H, d, J = 9.9 Hz, H-1"), 3.90-3.81 (2H, m, H-6"), 3.66 (1H, J = 8.9, 9.9 Hz,, H-2"), 3.58-3.44 (3H, m, H-3"-5"), 3.18 (2H, m, H-4'), 2.45 (2H, m, H-1'), 1.74 (4H, m, H-2', 3'); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 164.1 (C-2), 159.4 (C-7), 157.5 (C-4), 155.2 (C-5), 153.1 (C-9), 118.4 (C-3), 109.3 (C-6), 105.2 (C-10), 96.0 (C-8), 82.7 (C-5"), 79.3 (C-3"), 77.2 (C-1"), 74.0 (C-2"), 71.0 (C-4"), 61.8 (C-6"), 31.0 (C-4'), 25.5 (C-1'), 23.3 (C-3'), 22.4 (C-2').[5]



12a

6-(C-β-D-Glucosyl)-3,4-cyclopentyl-5,7-dihydroxycoumarin (12a) (9.2 mg; solid, white): ESI-MS *m/z* 379.31 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{18}H_{21}O_9$ $[M+H]^+$: 381.1180; found: 381.1171. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 6.34 (1H, s, H-8), 4.96 (1H, d, J = 9.8 Hz, H-1"), 3.91-3.82 (2H, m, H-6"), 3.64 (1H, dd, J = 9.0, 9.8 Hz, H-2"), 3.59-3.44 (3H, m, H-3"-5"), 3.33(Overlap, H-3'), 2.68 (2H, t, J = 7.1 Hz, H-1'), 2.08 (2H, m, H-2'); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 162.8 (C-2), 160.2 (C-7), 159.4 (C-4), 157.1 (C-5), 156.3 (C-9), 122.4 (C-3), 109.1 (C-6), 104.3 (C-10), 96.0 (C-8), 82.7 (C-5"), 79.3 (C-3"), 77.2 (C-1"), 74.3 (C-2"), 71.0 (C-4"), 61.9 (C-6"), 37.2 (C-3'), 30.1 (C-1'), 23.7 (C-2').[5]



17a

6-(C-β-D-Glucosyl)-3,4-Dihydro-4-phenyl-5,7-dihydroxycoumarin (17a) (1.2 mg; solid, white): ESI-MS *m/z* 417.10 $[M–H]^-$; HRMS (ESI^+) calcd. for $C_{21}H_{23}O_9$ $[M+H]^+$: 419.1337; found: 419.1309. ^{1}H NMR (400 MHz, CD_3CD-d_4): δ = 7.27-7.08 (5H, m, H-2'-6'), 6.21 (1H, s, H-8), 4.92 (1H, d, *J* = 9.8 Hz, H-1"), 3.82-3.72 (2H, m, H-6"), 3.65-3.60 (1H, m, H-2"), 3.51-3.48 (3H, m, H-3"-5") , 4.60 (1H, dd, *J* = 1.6, 7.0 Hz, H-4) , 3.07 (1H, dd, *J* = 7.0, 15.8 Hz, H-3a) , 2.88 (1H, dd, *J* = 1.6, 15.8 Hz, H-3b); ^{13}C NMR (150 MHz, CD_3CD-d_4): δ = 169.9 (C-2), 157.5 (C-7), 155.4 (C-5), 154.0 (C-9), 143.5 (C-1'), 129.7 (C-2', 6'), 128.7 (C-4'), 128.0 (C-3', 5'), 109.3 (C-6), 106.2 (C-10), 96.7 (C-8), 82.7 (C-5"), 79.5 (C-3"), 77.3 (C-1"), 74.4 (C-2"), 70.9 (C-4"), 61.7 (C-6"), 38.6 (C-3), 36.0 (C-4).[1]

3. Supplementary Tables

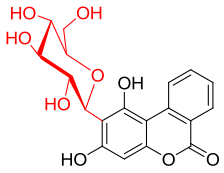
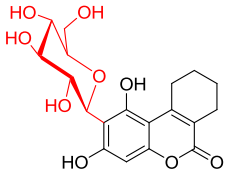
Table S1. HPLC methods used in this study

method	solvent A	solvent B	gradient	analysis substrates
A	Water containing 0.1% formic acid	MeOH	10–70% B, 25 min; 70–100% B, 5 min; 100% B, 5 min	**1**, **2**, **15**, **16**
B	Water containing 0.1% formic acid	MeOH	30–100% B, 25 min; 100% B, 5 min	**5**, **6**, **9**, **13**
C	Water containing 0.1% formic acid	MeOH	10–50% B, 25 min; 50–100% B, 5 min; 100% B, 5 min	**3**, **4**, **7**, **8**, **10**–**12**, **14**, **17**–**20**

Table S2. Conversion rates of O-glucosylated products catalyzed by purified enzyme or whole cells containing MiCGTb-GAGM

no.	enzyme (%)	whole cells (%)		
		0.1 mM	0.2 mM	0.4 mM
1	4.3	82.1	79.9	81.5
2	2.8	46.9	43.4	36.8
3	0	0	0	0
4	0	2.0	3.3	9.3
5	61.9	0	0	0
6	80.2	76.8	64.1	53.5
7	0	0	2.0	3.0
8	0	37.7	35.5	15.4
9	14.2	0	0	0
10	0	0	0	0
11	0	2.5	3.0	4.0
12	0	2.0	2.9	5.5
13	6.5	7.0	5.5	3.2
14	17.8	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	14.5	/	/	/
19	26.5	/	/	/
20	32.6	/	/	/

Table S3. *In vitro* SGLT2 inhibitory activity for **10a** and **11a**

compounds	structure	IC_{50} (10^{-6} M)
10a		2.2
11a		6.3

4. Supplementary Figures

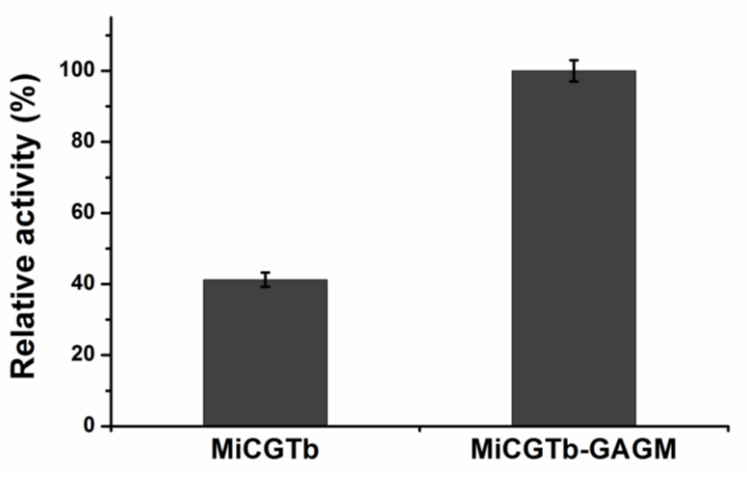


Figure S1. Relative activity of MiCGTb and MiCGTb-GAGM towards **1** with UDP-glucose.

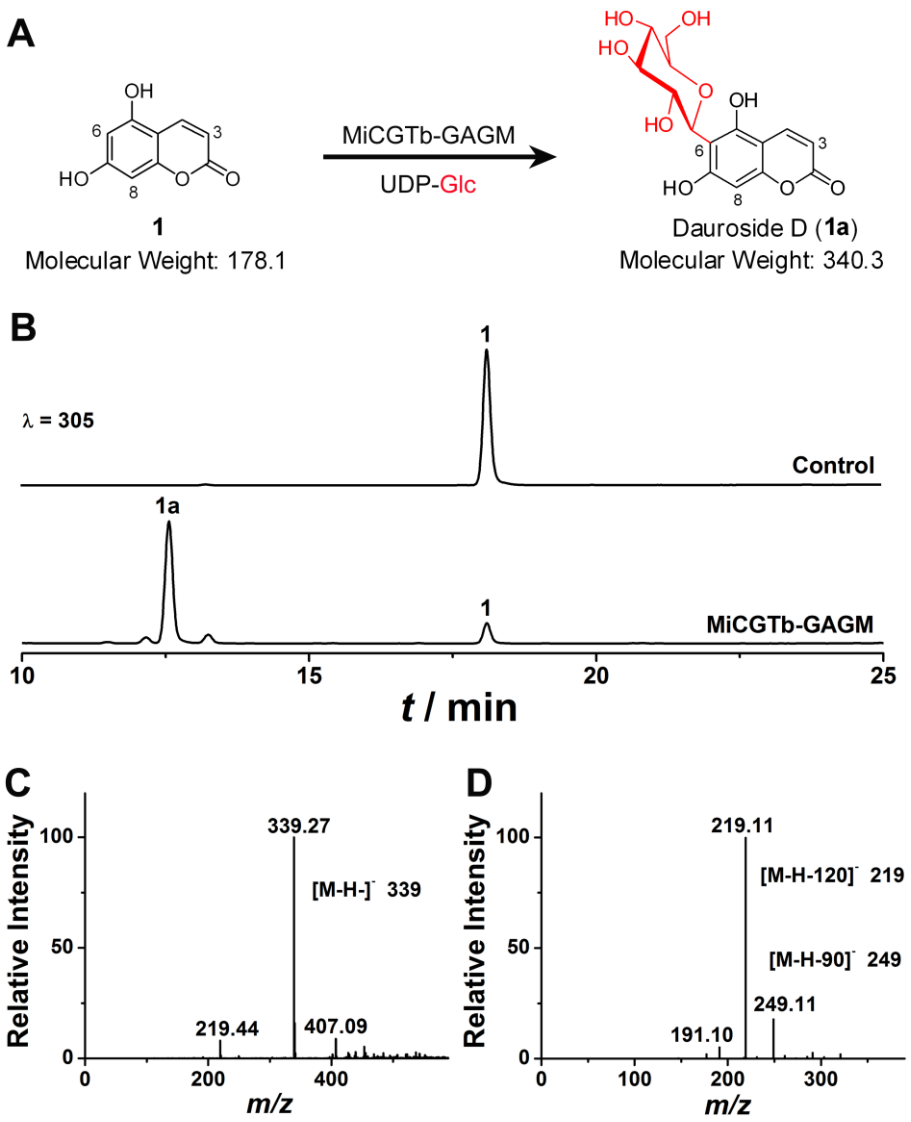


Figure S2. C-glucosylation of **1** catalyzed by an engineered C-glycosyltransferase, MiCGTb-GAGM. (A) The reaction catalyzed by MiCGTb-GAGM. (B) HPLC chromatograms of **1** and the enzyme product **1a**. (C) and (D) Typical negative-ion MS and MS^2 spectra for **1a**. The HPLC conditions are provided in Table S1 in the Supporting Information.

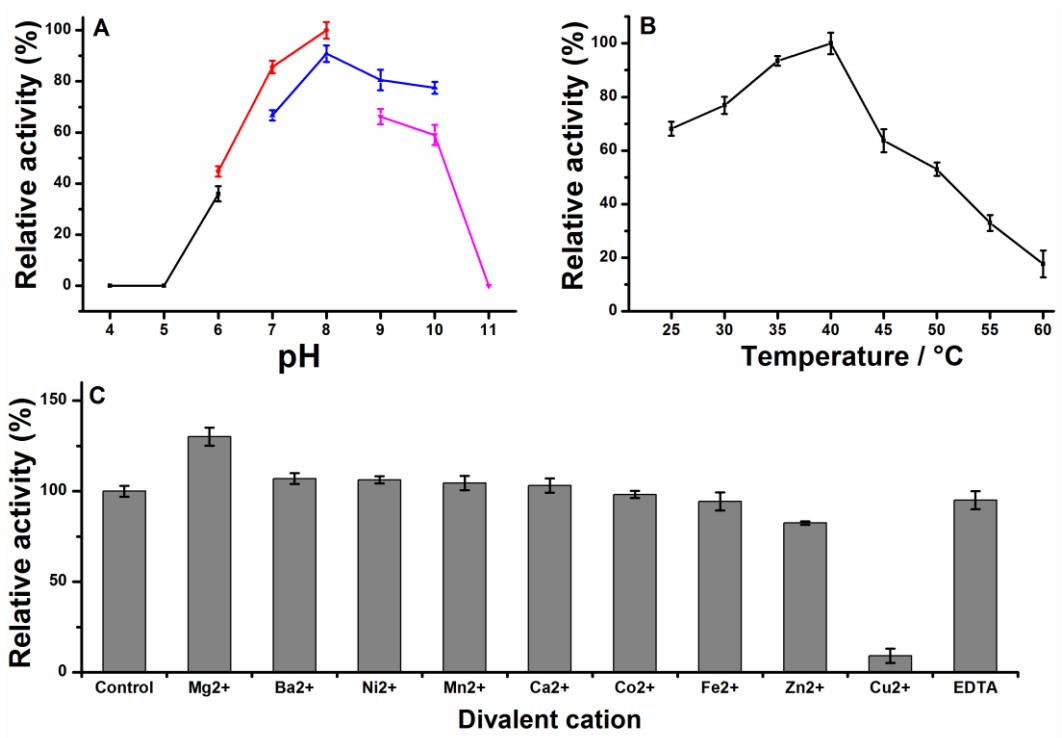


Figure S3. Effects of various temperature (A), pH (B), and divalent metal ions (C) on the C-glucosylation activity of MiCGTb-GAGM. UDP-Glc was used as a sugar donor and **1** was used as an acceptor. The optimum temperature and pH of purified MiCGTb-GAGM were 40 ℃ and 8.0, respectively. This enzyme was independent of metal ions, while Mg^{2+} could greatly enhance the activity.

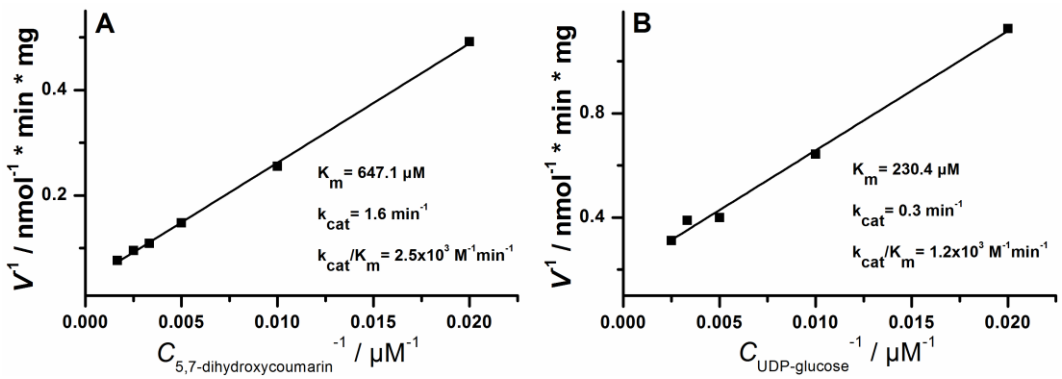


Figure S4. Determination of kinetic parameters for recombinant MiCGTb-GAGM. The kinetic parameters were detected with **1** (A) and UDP-Glc (B) at 40 ℃ and pH 8.0. The apparent K_m values of **1** and UDP-Glc are 647.1 μM and 230.4 μM, respectively. The k_{cat}/K_m values of **1** and UDP-Glc are 2.5 × 10^3 $M^{-1}min^{-1}$ and 1.2 × 10^3 $M^{-1}min^{-1}$, respectively.

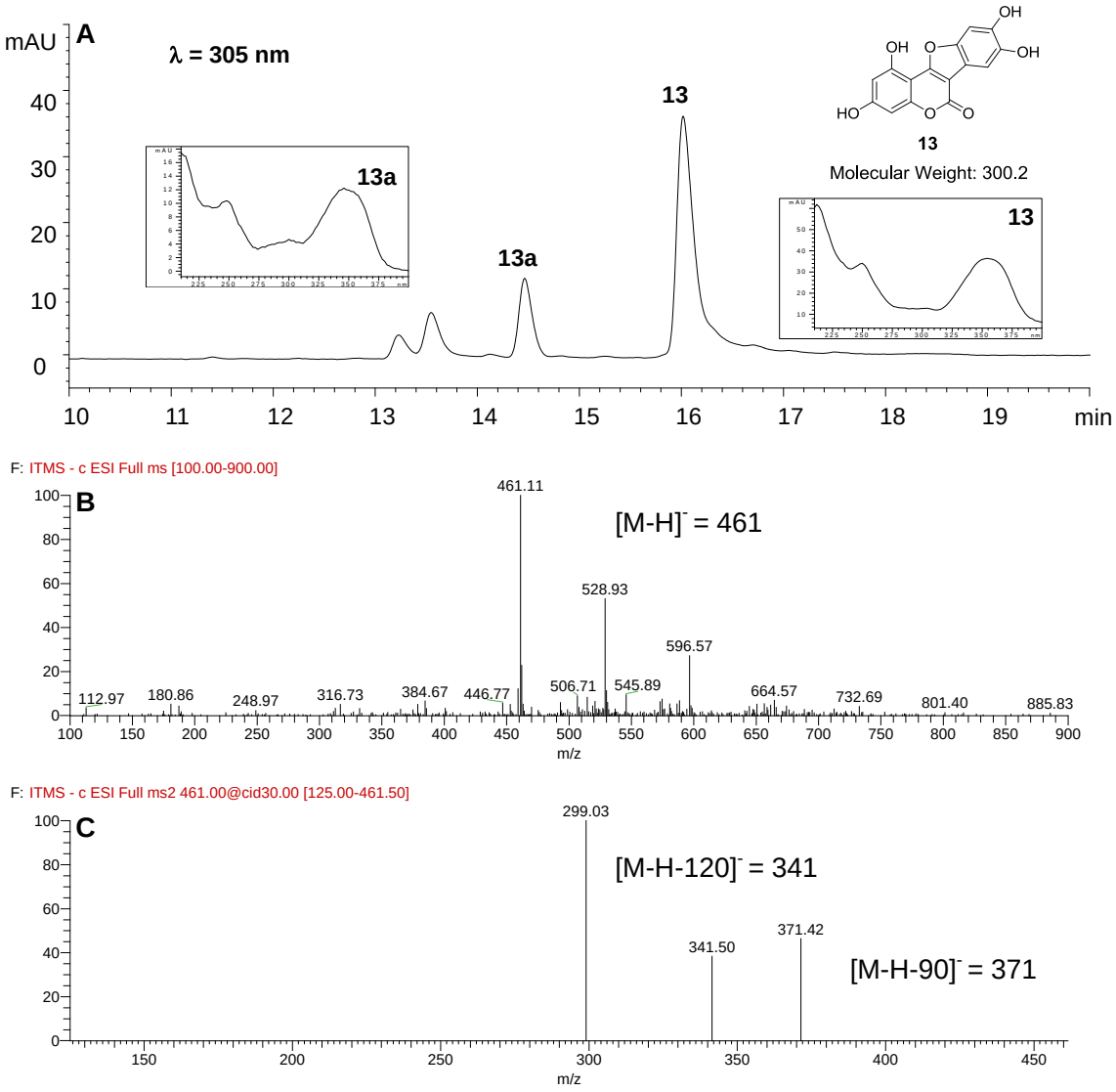


Figure S5. HPLC-DAD/ESI-MS analysis of MiCGTb-GAGM enzymatic product using **13** as an acceptor and UDP-Glc as a sugar donor. (A) HPLC chromatogram and UV spectra of **13** and enzyme product **13a**; (B) and (C) Typical negative ion MS and MS^2 spectra for **13a**. The HPLC conditions are available in Table S1.

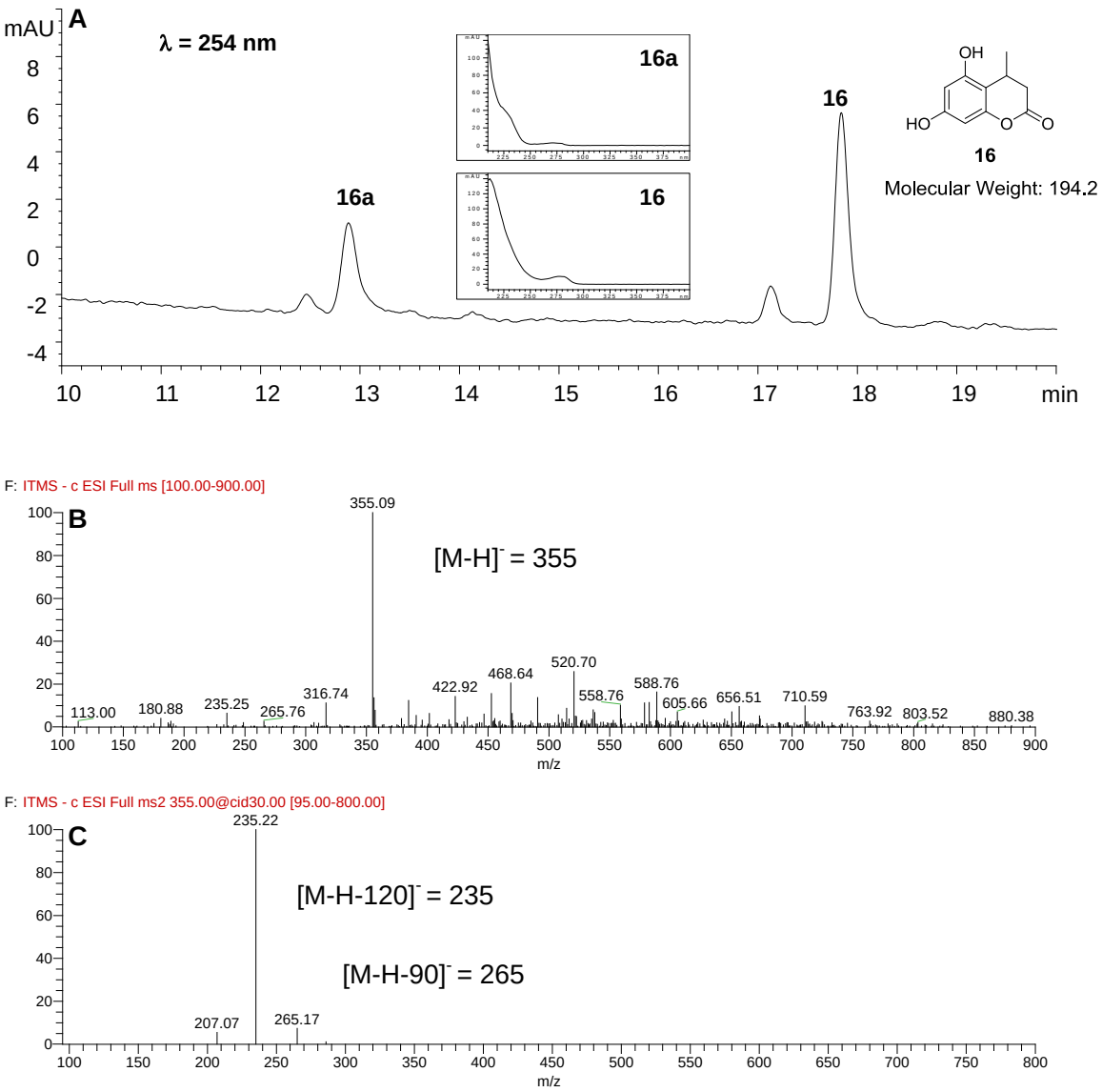


Figure S6. HPLC-DAD/ESI-MS analysis of MiCGTb-GAGM enzymatic product using **16** as an acceptor and UDP-Glc as a sugar donor. (A) HPLC chromatogram and UV spectra of **16** and enzyme product **16a**; (B) and (C) Typical negative ion MS and MS^2 spectra for **16a**. The HPLC conditions are available in Table S1.

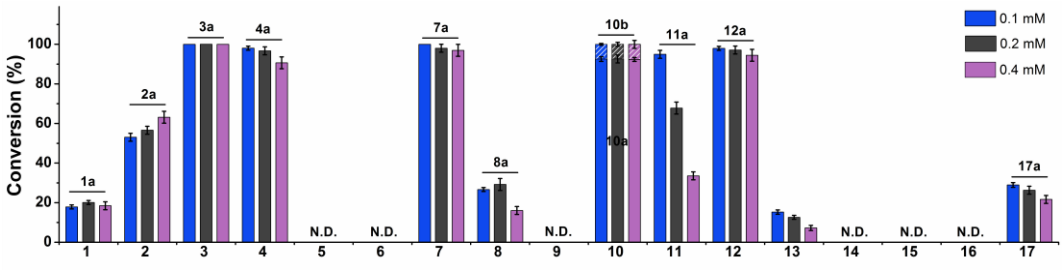


Figure S7. Conversion rates of coumarin C-glucosides using the whole cells containing MiCGTb-GAGM at the final concentration of 0.1/0.2/0.4 mM.

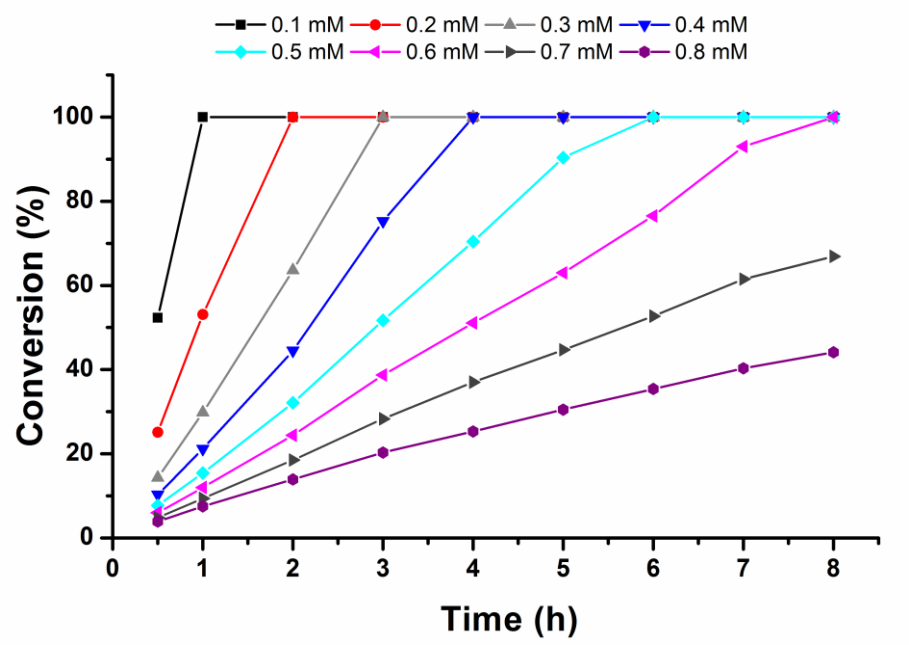


Figure S8. Time-course of **7a** by whole cells harboring MiCGTb-GAGM at different substrate concentration.

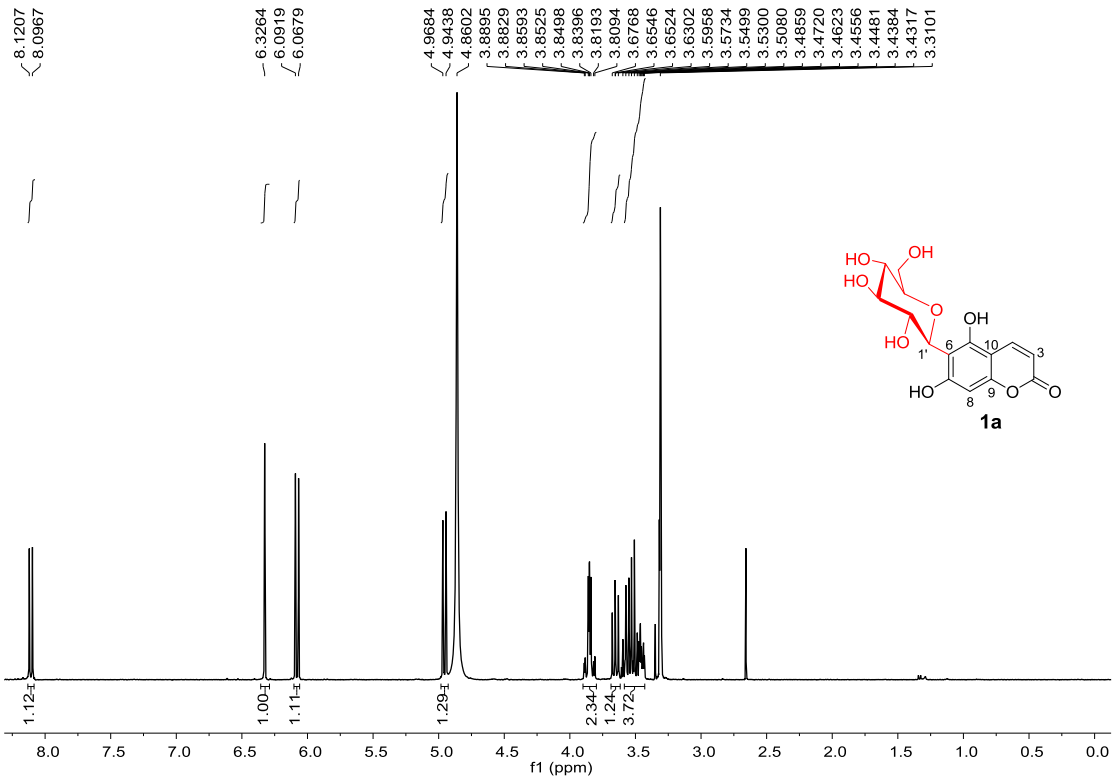


Figure S9. ^{1}H NMR spectrum of **1a** (CD_3OD, 400 MHz).

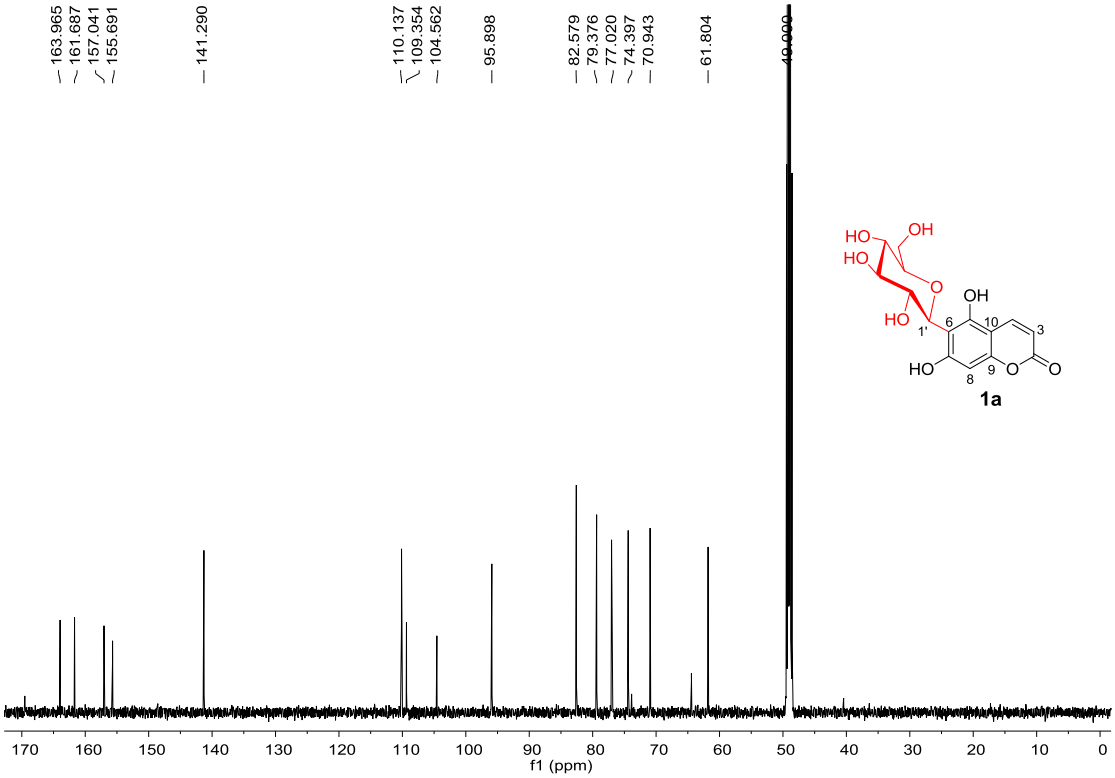


Figure S10. ^{13}C NMR spectrum of **1a** (CD_3OD, 150 MHz).

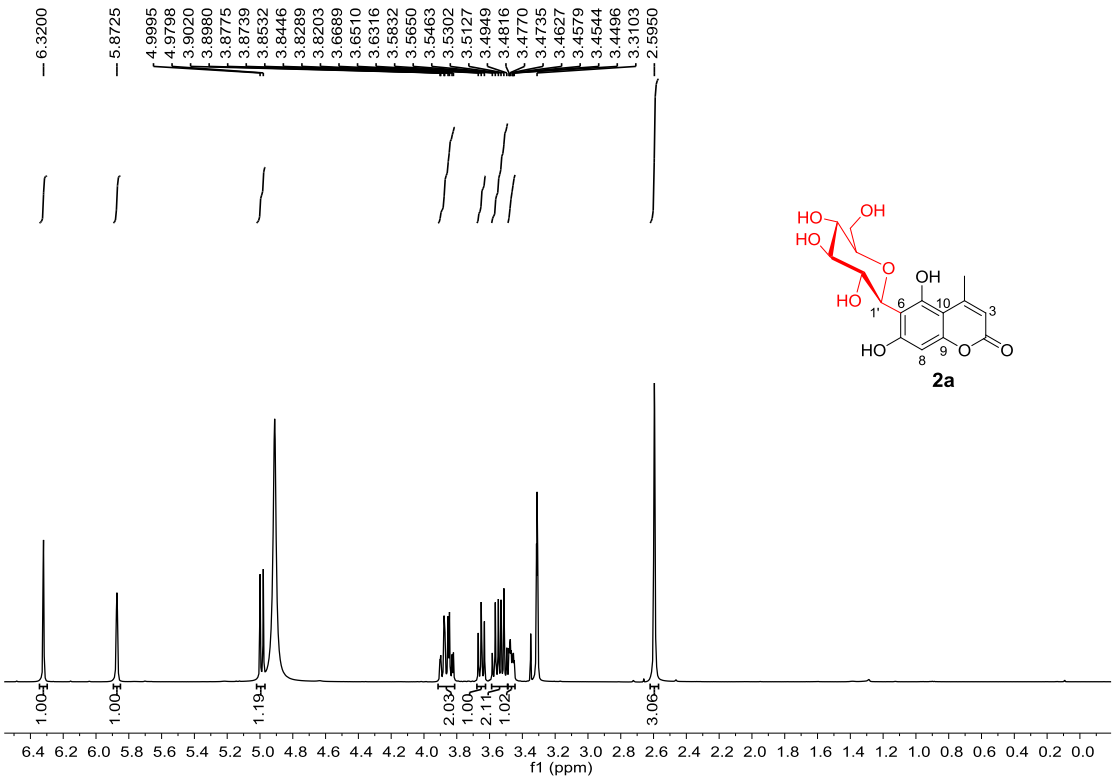


Figure S11. ^{1}H NMR spectrum of **2a** (CD_3OD, 500 MHz).

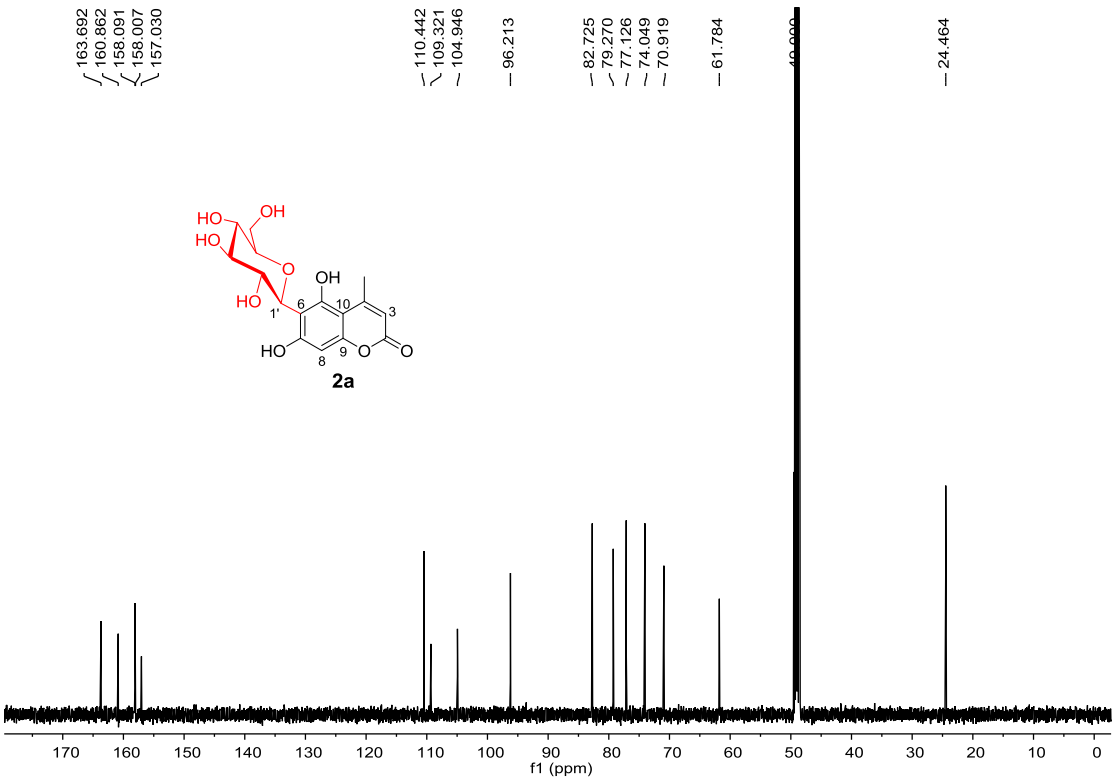


Figure S12. ^{13}C NMR spectrum of **2a** (CD_3OD, 125 MHz).

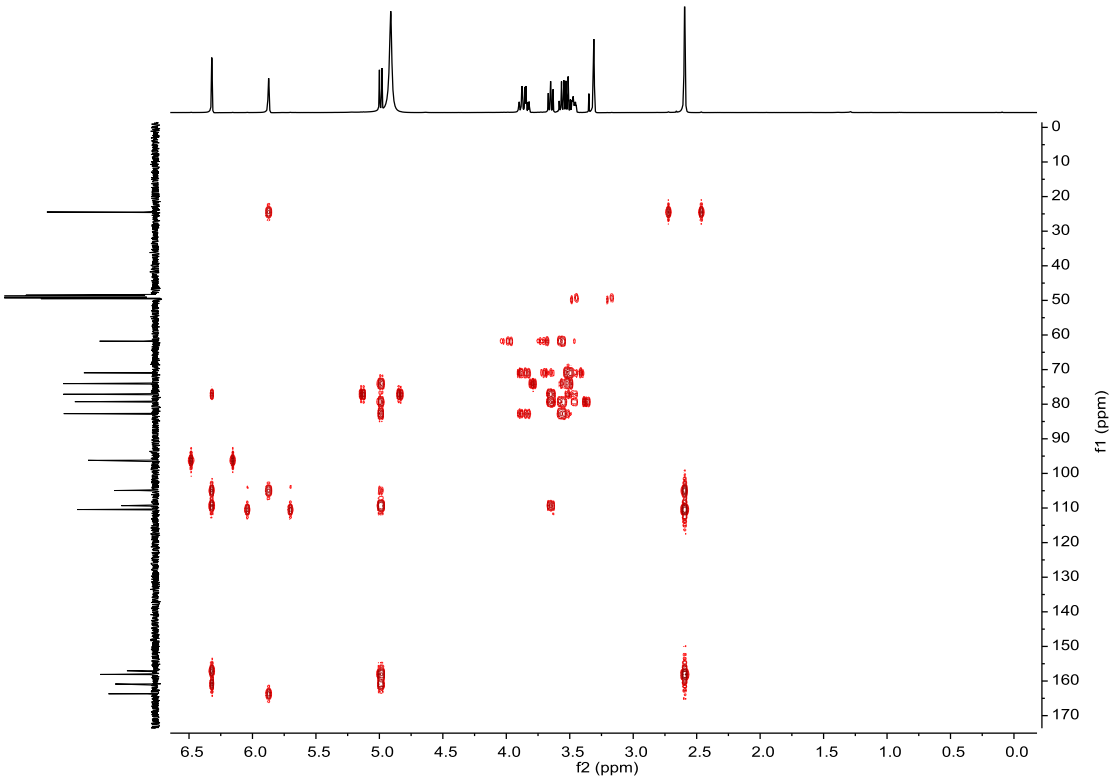


Figure S13. HMBC spectrum of **2a** (CD_3OD).

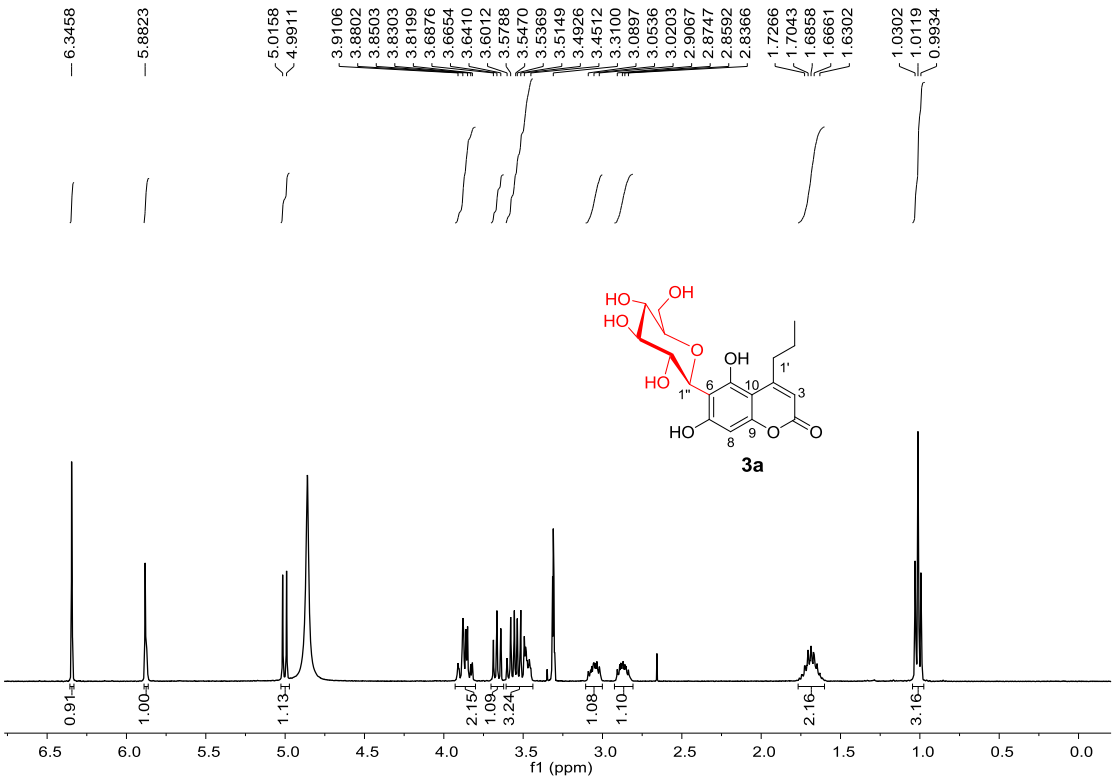


Figure S14. ^{1}H NMR spectrum of **3a** (CD_3OD, 400 MHz).

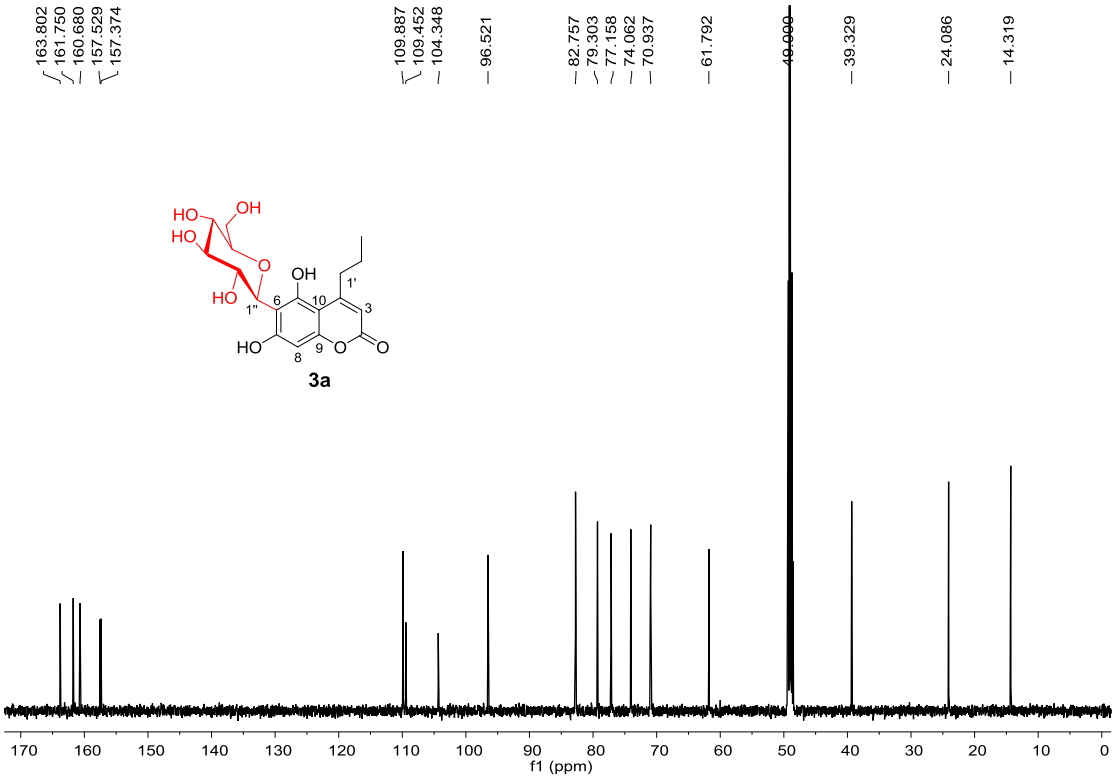


Figure S15. ^{13}C NMR spectrum of **3a** (CD_3OD, 150 MHz).

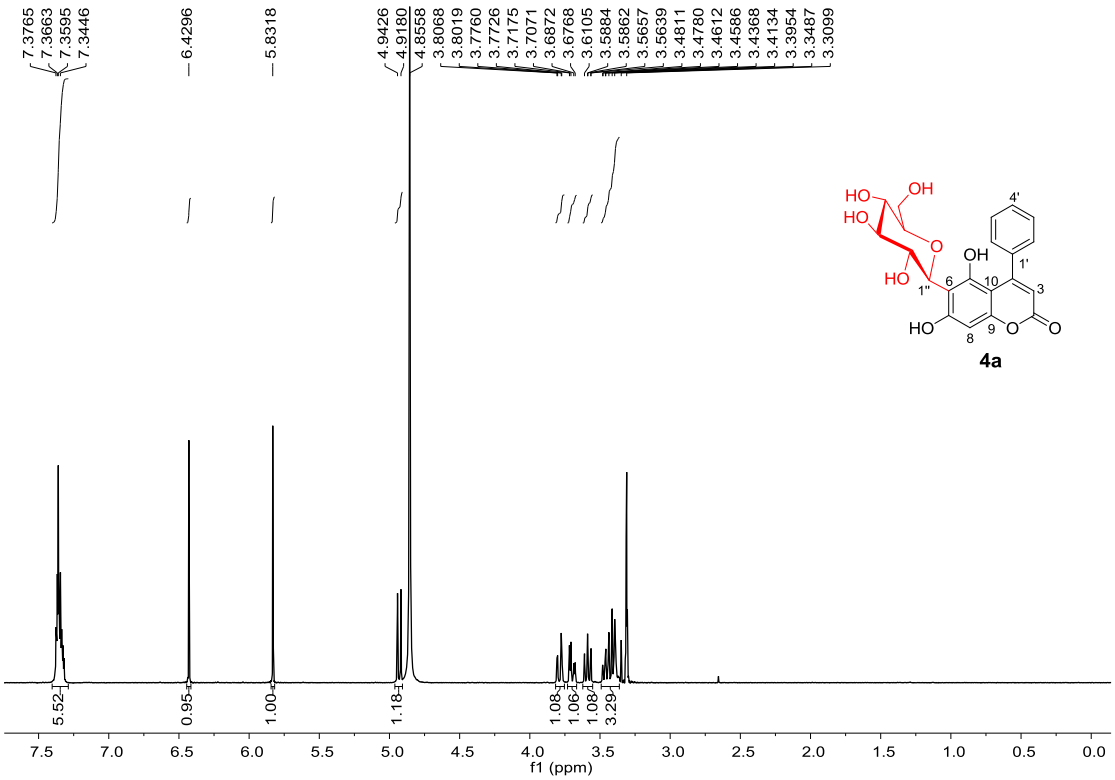


Figure S16. ^{1}H NMR spectrum of **4a** (CD_3OD, 400 MHz).

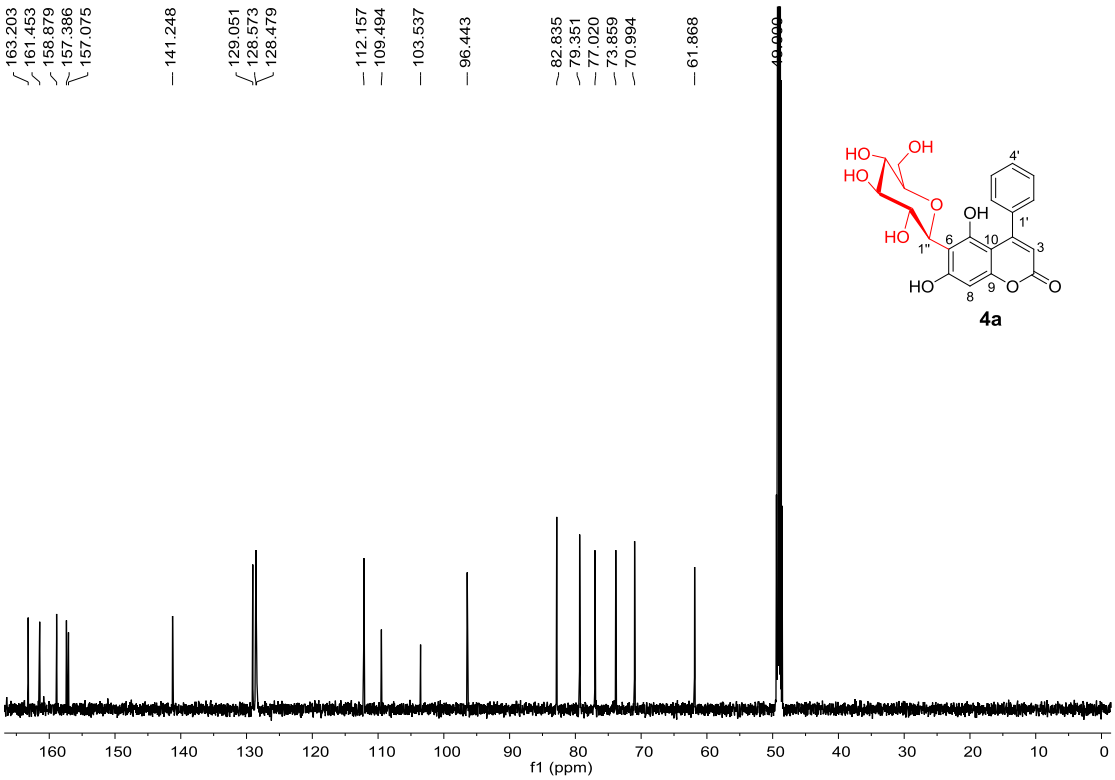


Figure S17. ^{13}C NMR spectrum of **4a** (CD_3OD, 150 MHz).

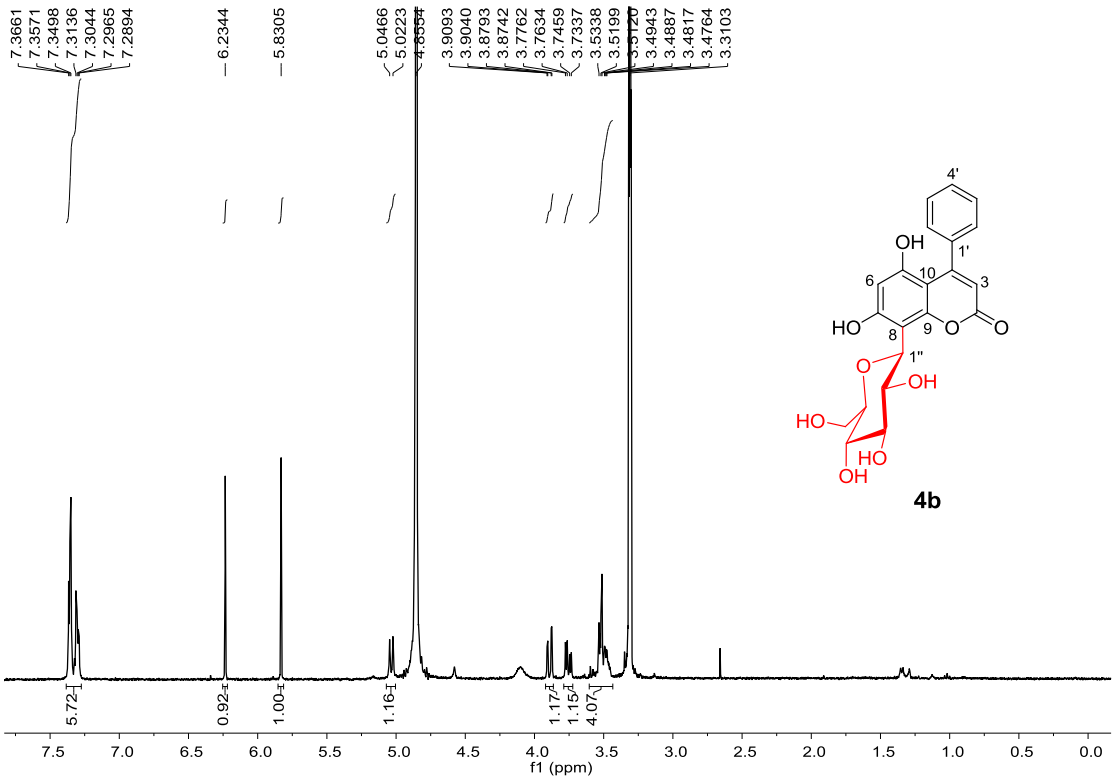


Figure S18. ^{1}H NMR spectrum of **4b** (CD_3OD, 400 MHz).

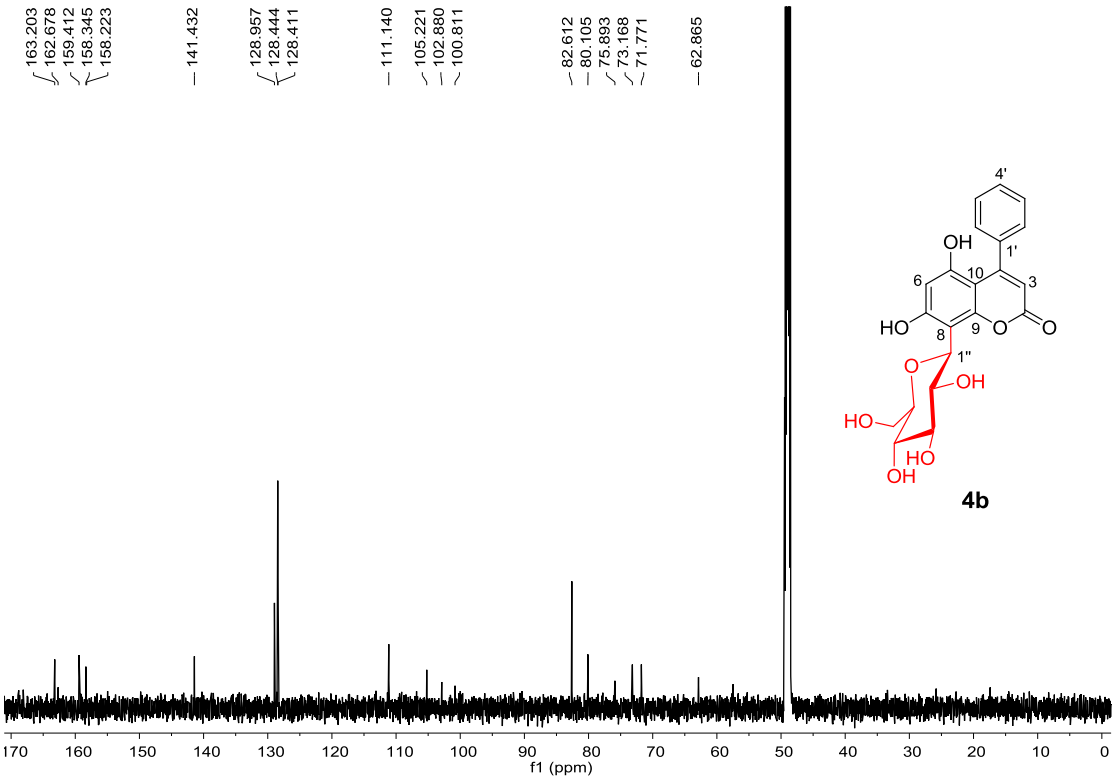


Figure S19. ^{13}C NMR spectrum of **4b** (CD_3OD, 150 MHz).

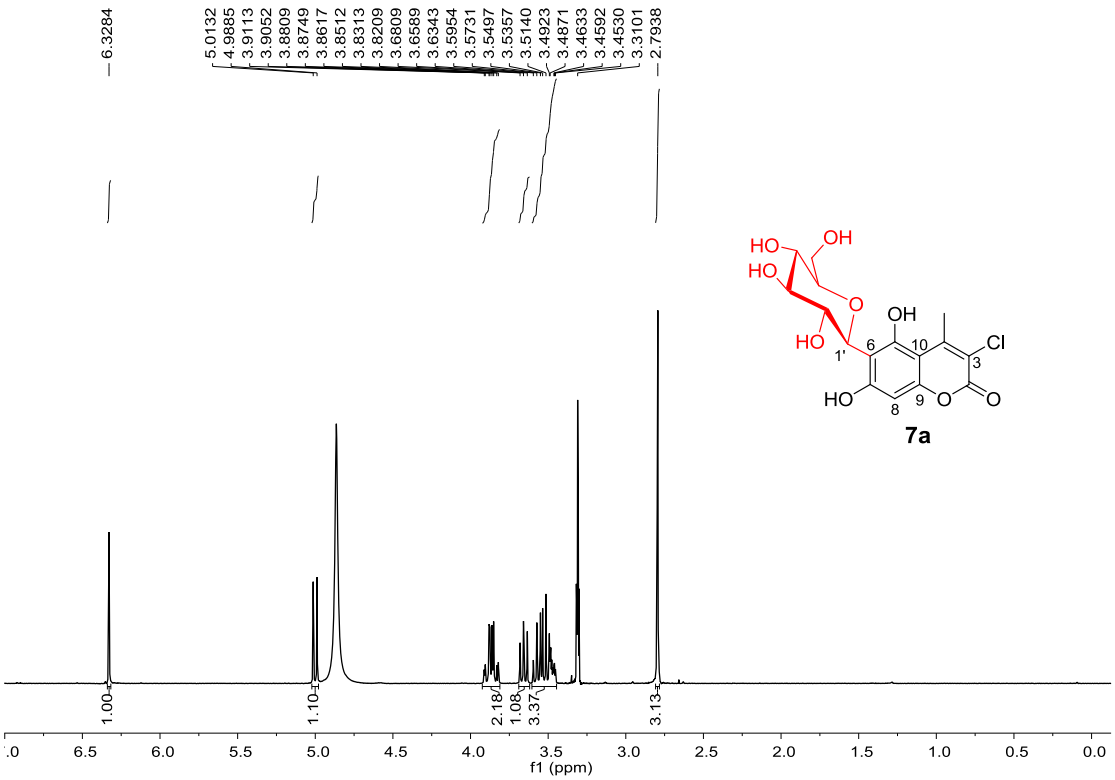


Figure S20. ^{1}H NMR spectrum of **7a** (CD_3OD, 400 MHz).

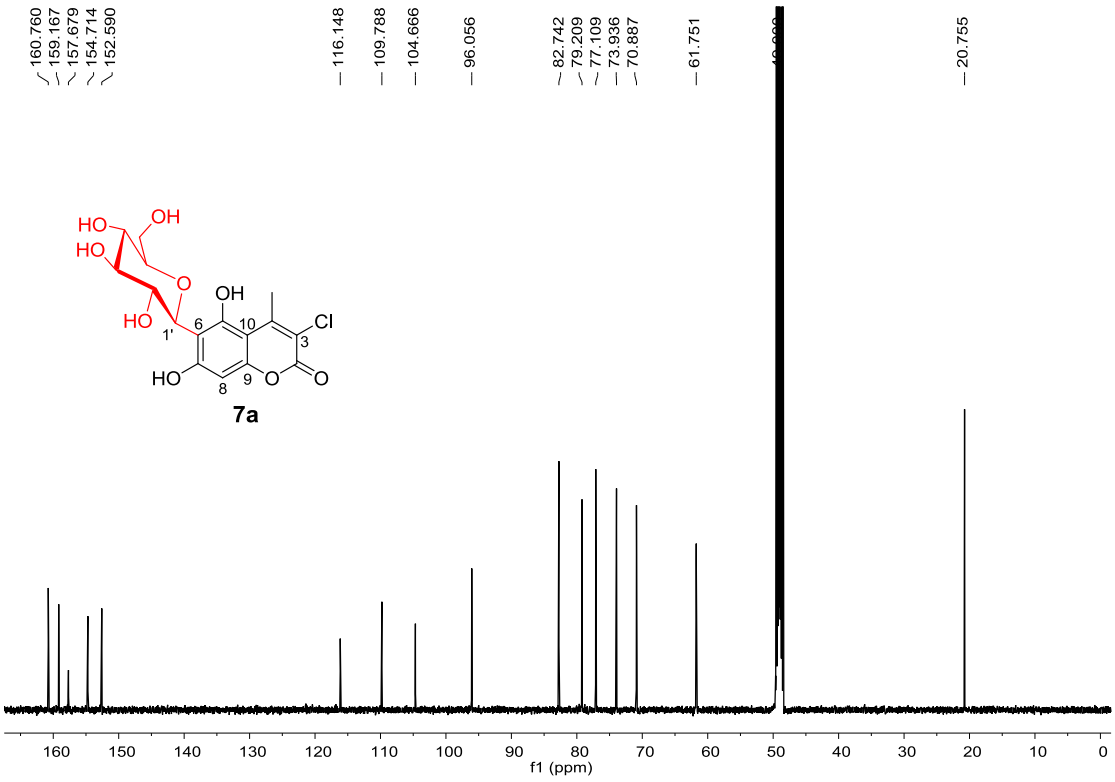


Figure S21. ^{13}C NMR spectrum of **7a** (CD_3OD, 150 MHz).

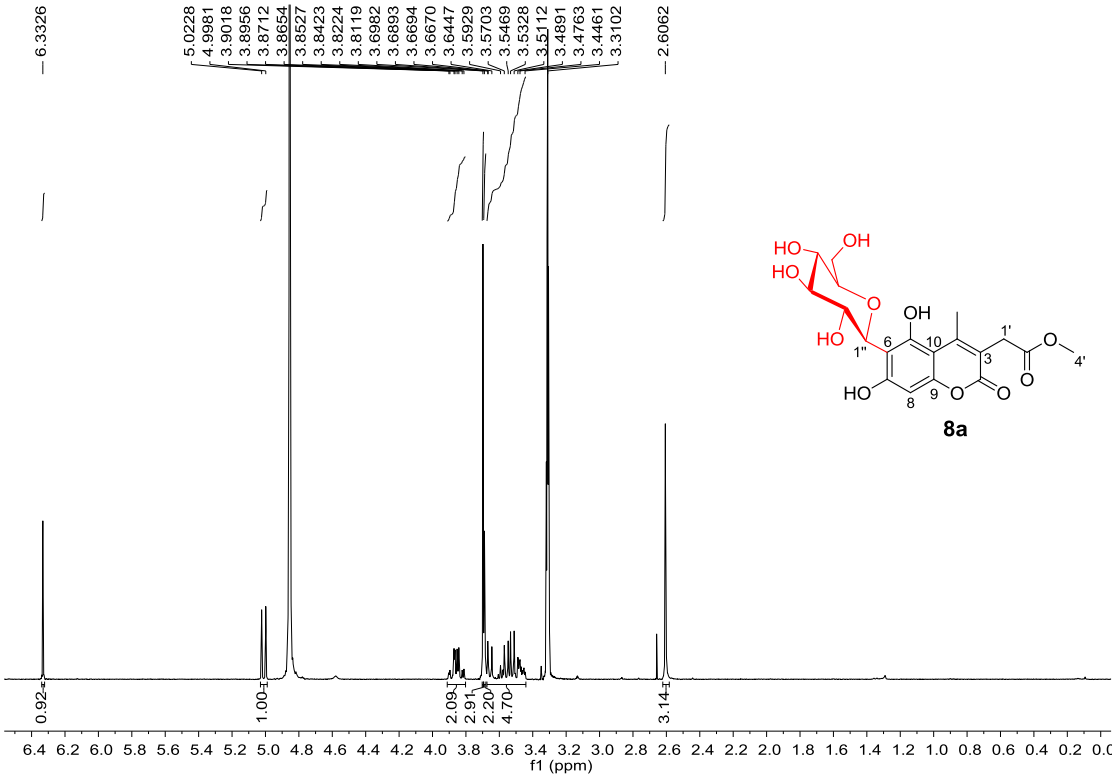


Figure S22. ^{1}H NMR spectrum of **8a** (CD_3OD, 400 MHz).

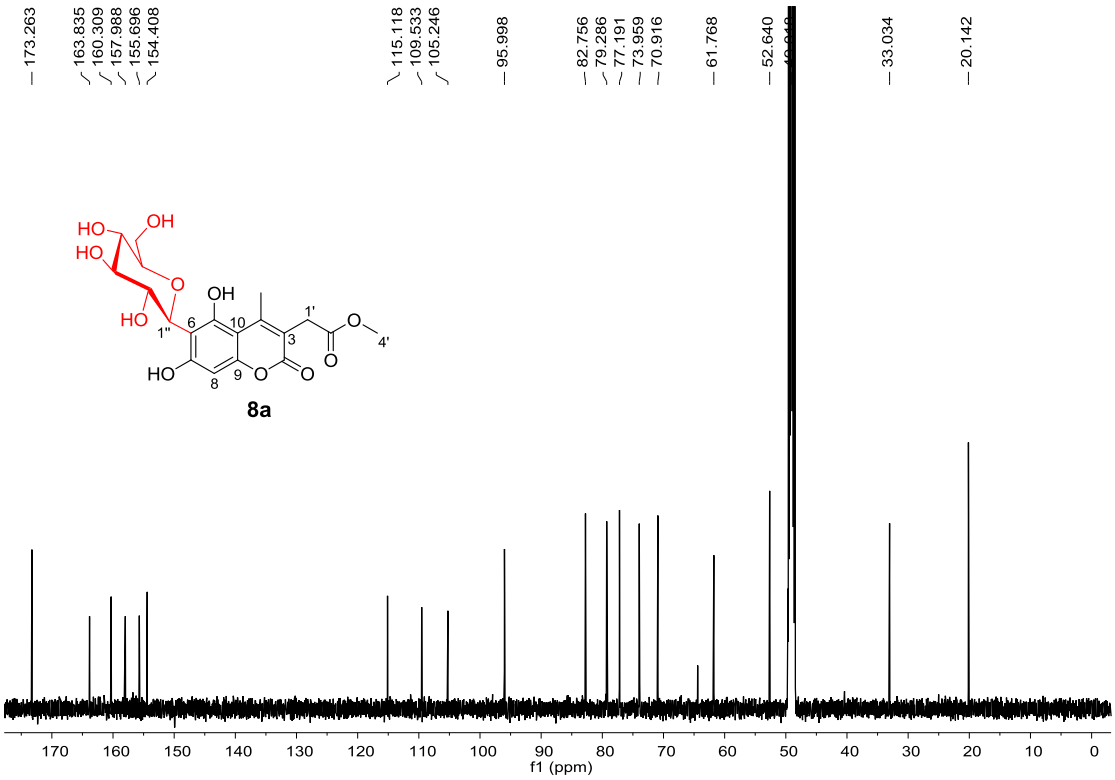


Figure S23. ^{13}C NMR spectrum of **8a** (CD_3OD, 125 MHz).

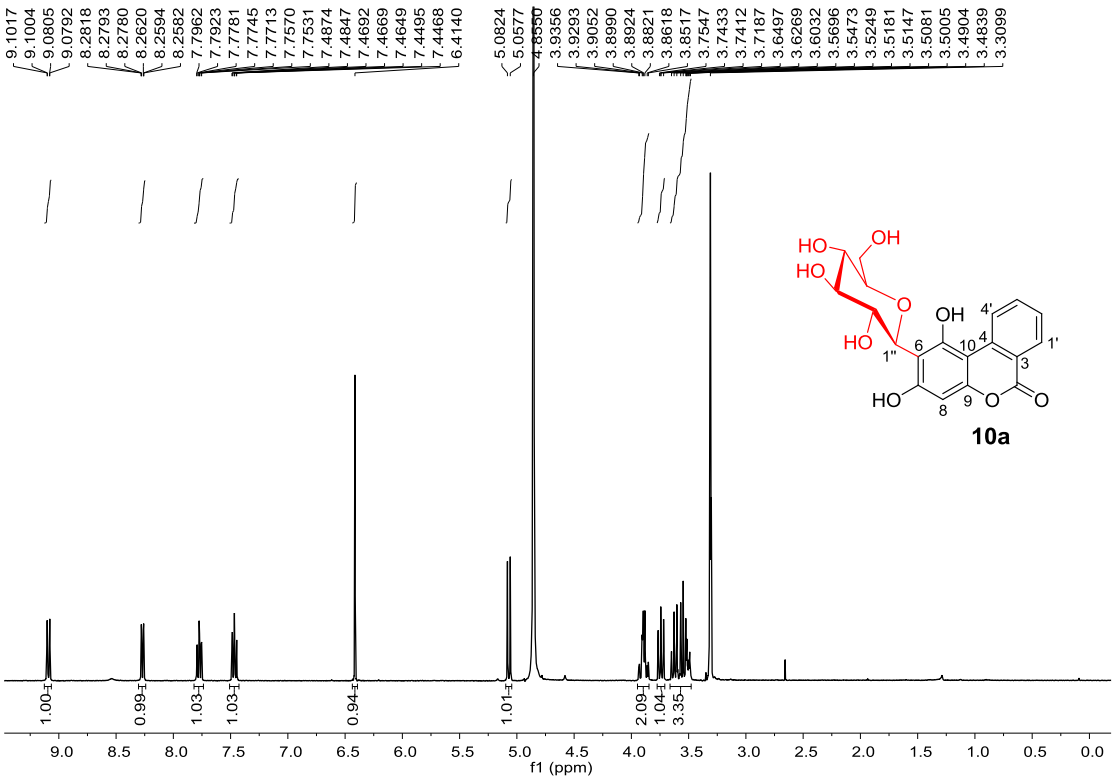


Figure S24. ^{1}H NMR spectrum of **10a** (CD_3OD, 400 MHz).

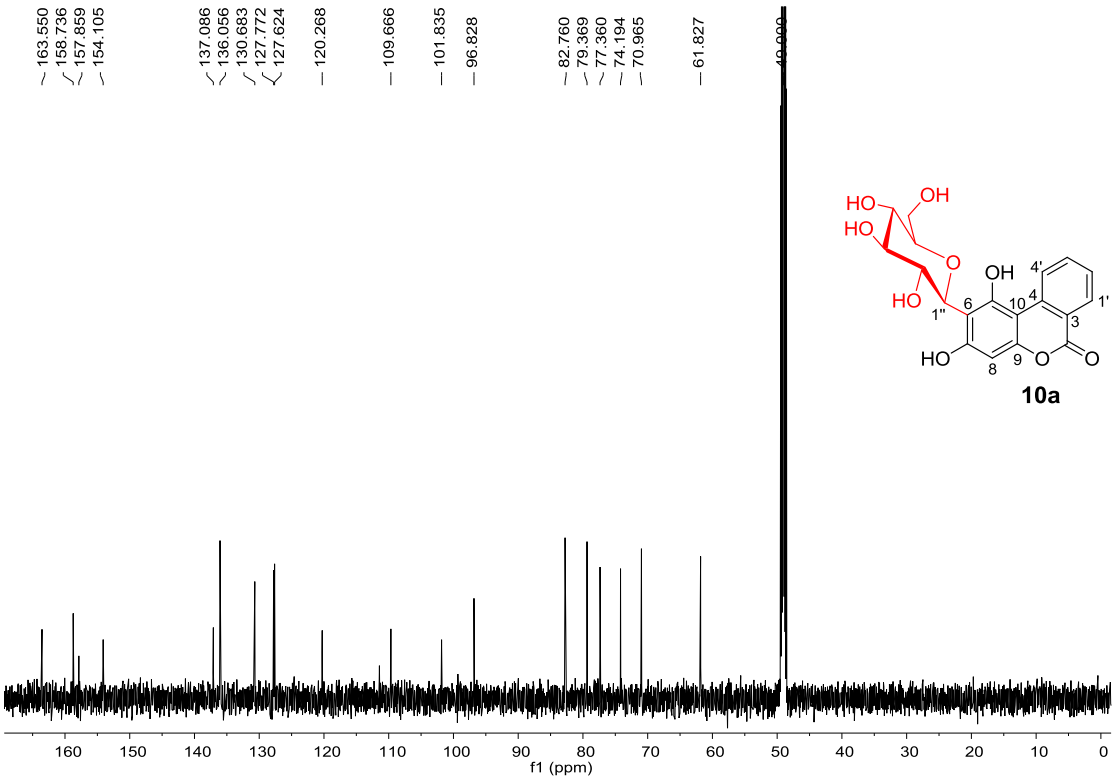


Figure S25. ^{13}C NMR spectrum of **10a** (CD_3OD, 150 MHz).

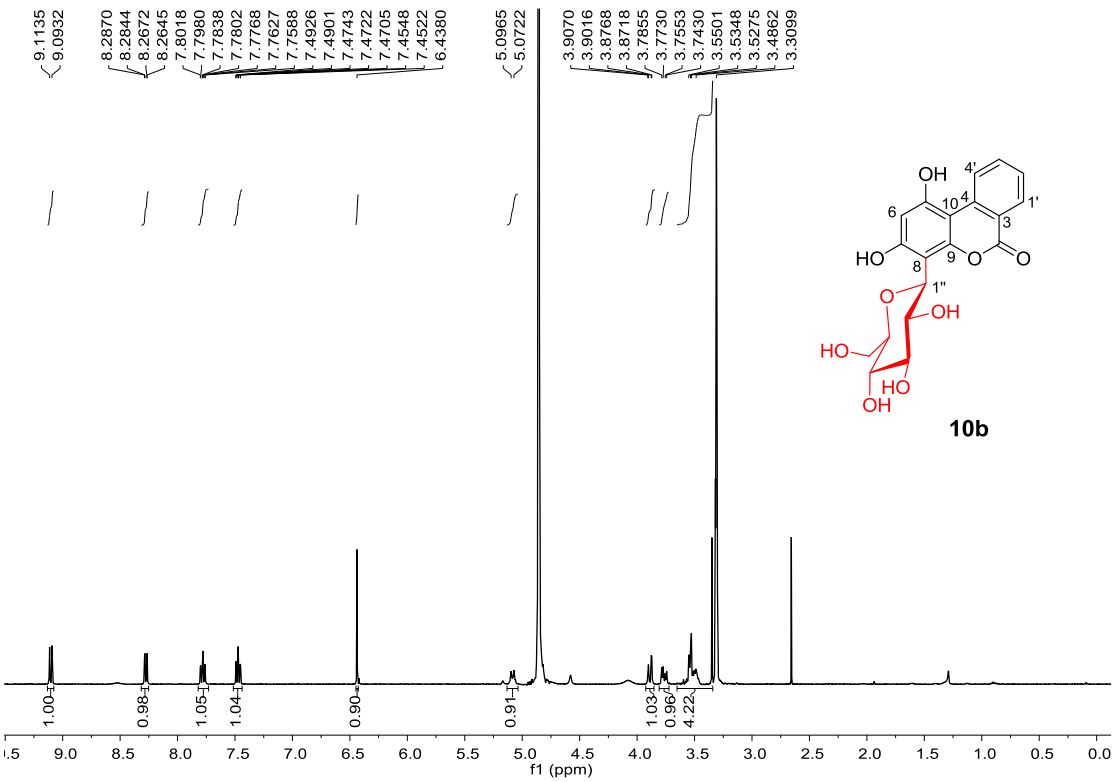


Figure S26. ^{1}H NMR spectrum of **10b** (CD_3OD, 400 MHz).

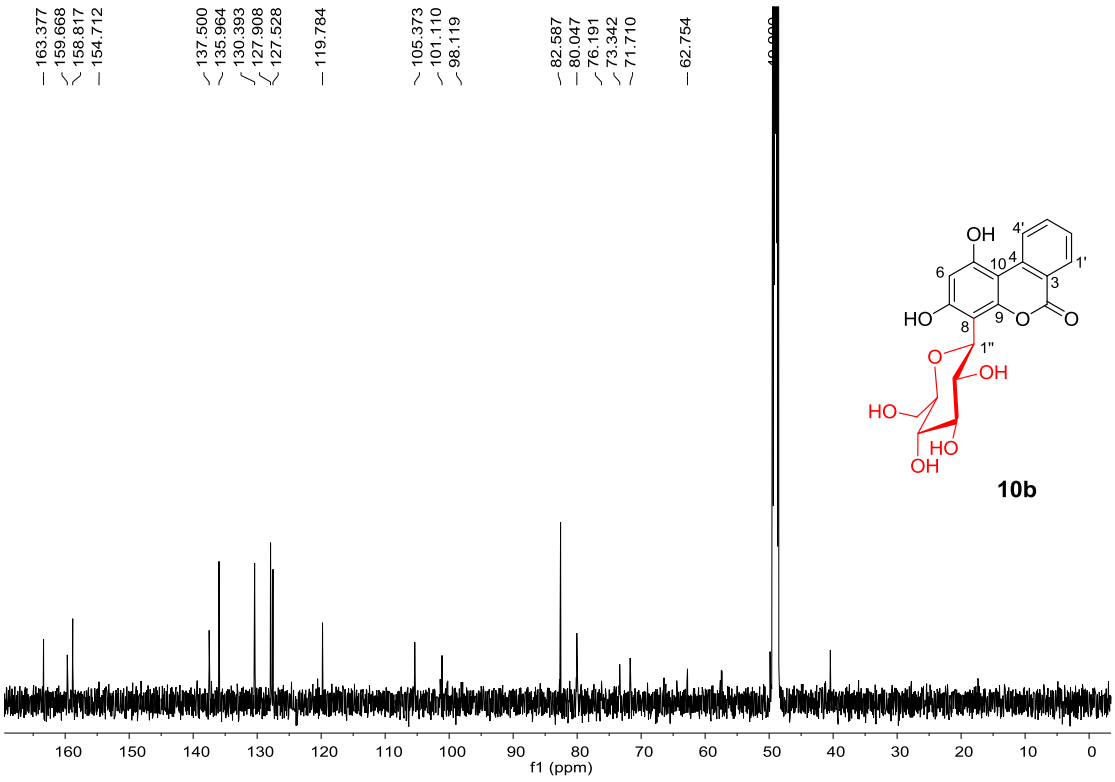


Figure S27. ^{13}C NMR spectrum of **10b** (CD_3OD, 150 MHz).

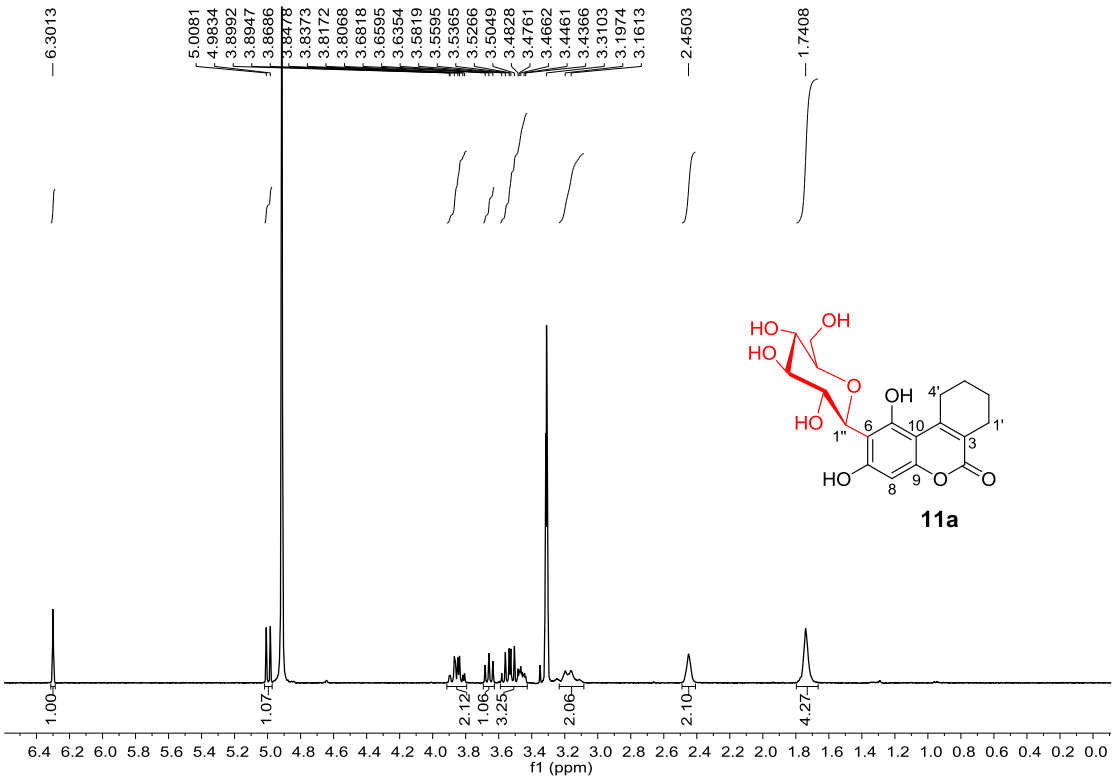


Figure S28. ^{1}H NMR spectrum of **11a** (CD_3OD, 400 MHz).

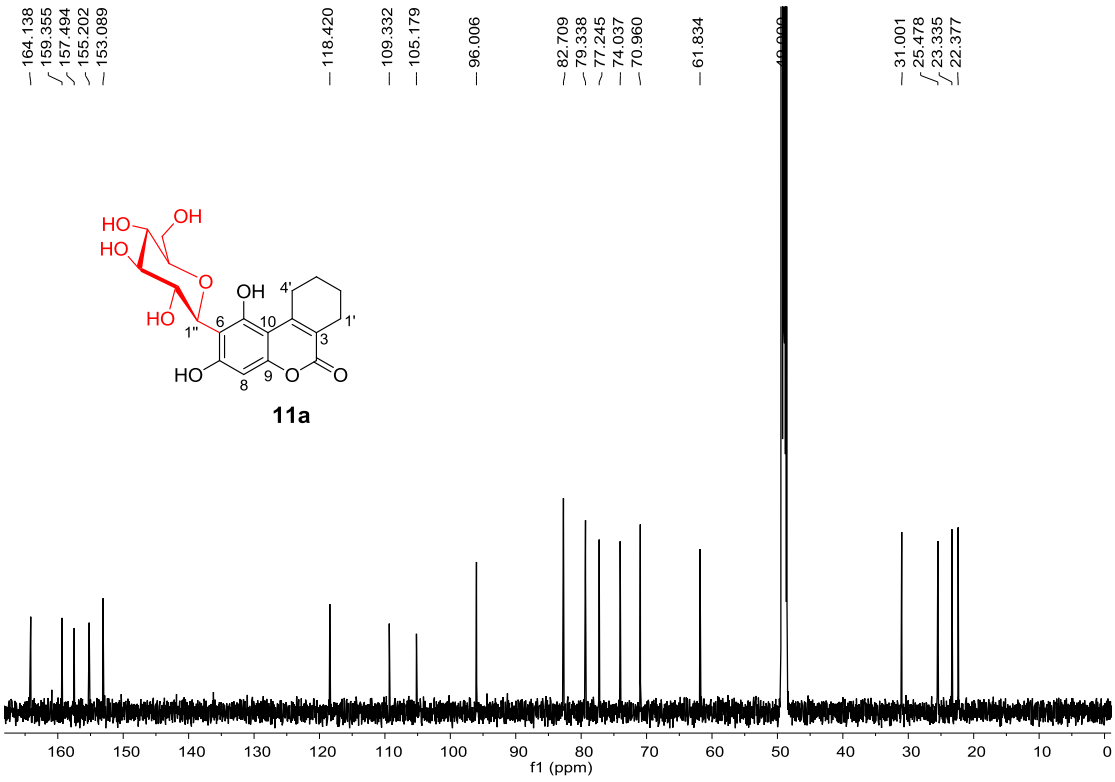


Figure S29. ^{13}C NMR spectrum of **11a** (CD_3OD, 150 MHz).

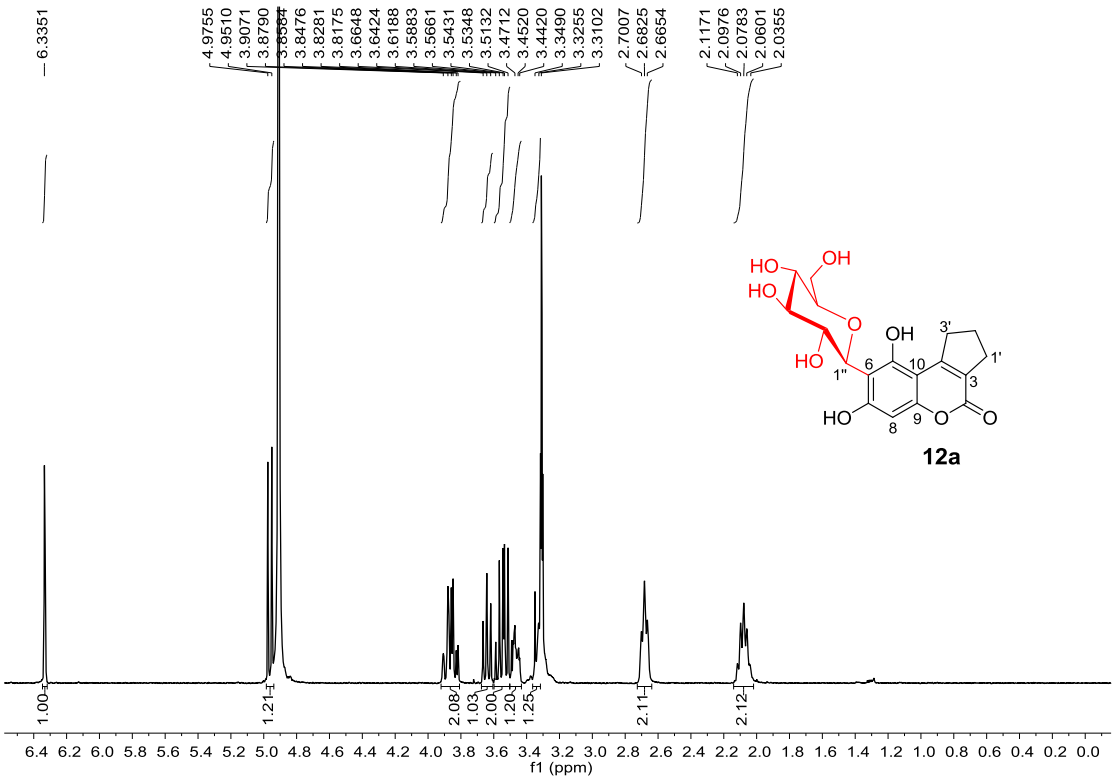


Figure S30. ^{1}H NMR spectrum of **12a** (CD_3OD, 400 MHz).

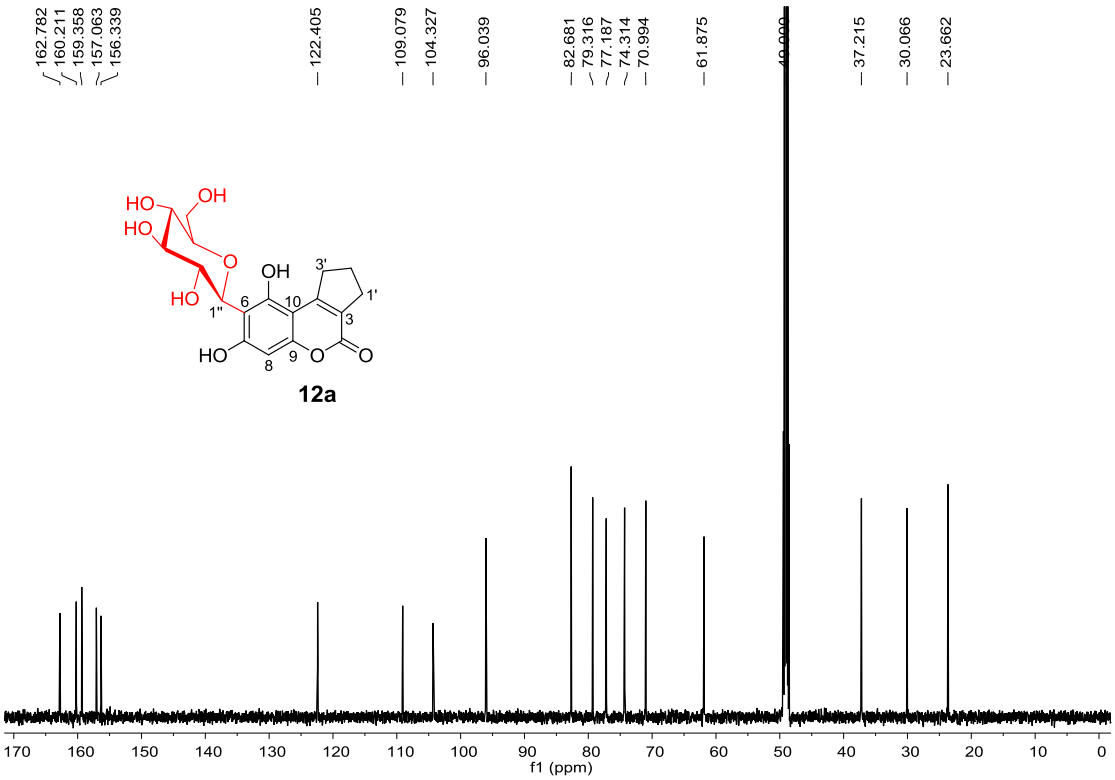


Figure S31. ^{13}C NMR spectrum of **12a** (CD_3OD, 150 MHz).

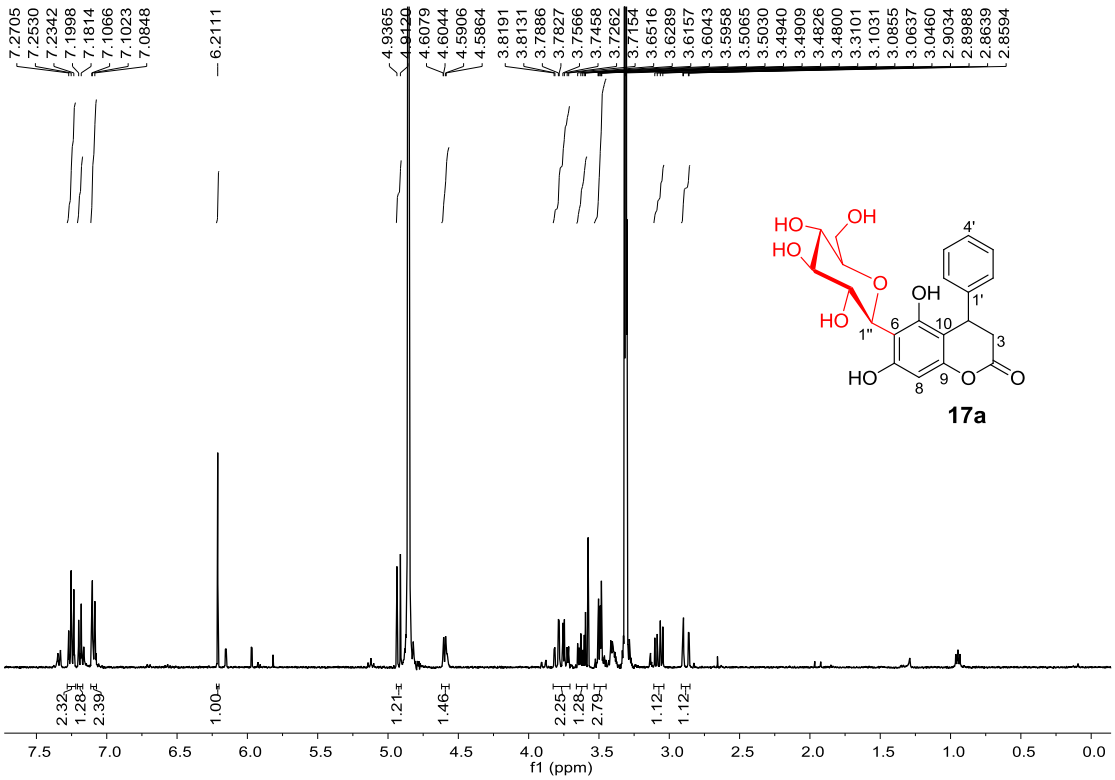


Figure S32. ^{1}H NMR spectrum of **17a** (CD_3OD, 400 MHz).

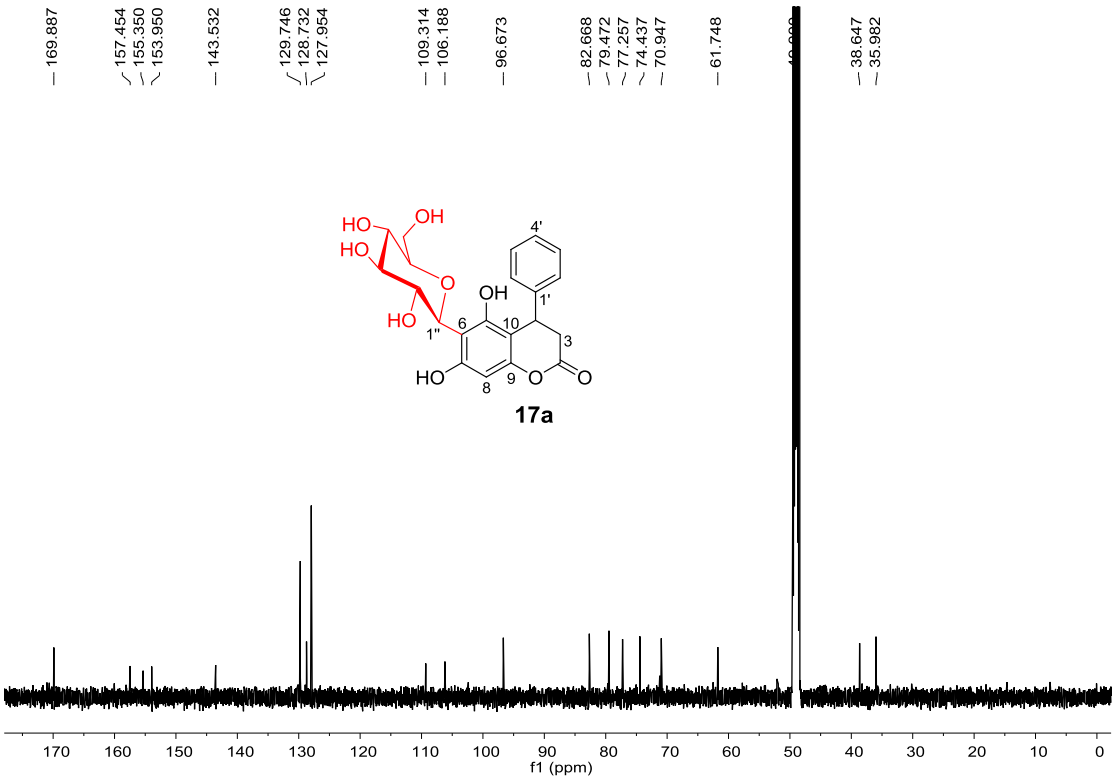


Figure S33. ^{13}C NMR spectrum of **17a** (CD_3OD, 150 MHz).

References

(1) (a) Kumar, A.; Singh, B. K.; Tyagi, R.; Jain, S. K.; Sharma, S. K.; Prasad, A. K.; Raj, H. G.; Rastogi, R. C.; Watterson, A. C.; Parmar, V. S. *Bioorg. Med. Chem.* **2005**, *13*, 4300–4305. (b) Bennardi, D. O.; Ruiz, D. M.; Romanelli, G. P.; Baronetti, G. T.; Thomas, H. J.; Autino, J. C. *Lett. Org. Chem.* **2008**, *5*, 607–615.

(2) Chen, D.; Chen, R.; Wang, R.; Li, J.; Xie, K.; Bian, C.; Sun, L.; Zhang, X.; Liu, J.; Yang, L.; Ye, F.; Yu, X.; Dai, J. *Angew. Chem. Int. Ed.* **2015**, *54*, 12678–12682.

(3) (a) Vdovin, A. D.; Batsurén, D.; Batirov, É. K.; Yagudaev, M. R.; Malikov, V. M. *Chem. Nat. Compd.* **1983**, *19*, 413–416. (b) Chen, C.; Wang, P.; Zou, L.; Yang, L.; Fan, Y.; Hu, W.; Ge, G. *Chem. Res. Chin. Univ.* **2017**, *33*, 194–199.

(4) (a) Keri, R. S.; Hosamani, K. M.; Reddy, H. R. *Catal. Lett.* **2009**, *131*, 321–327. (b) Madhav, J.; Kuarm, B.; Someshwar, P.; Rajitha, B.; Reddy, Y.; Crooks, P. *J. Chem. Res.* **2008**, *4*, 232–234.

(5) Li, G.; Wang, D.; Sun, M.; Li, G.; Hu, J.; Zhang, Y.; Yuan, Y.; Ji, H.; Chen, N.; Liu, G. *J. Med. Chem.* **2010**, *53*, 1741–1754.

(6) Mahling, J.-A.; Schmidt, R. R. *Liebigs Ann.* **1995**, *3*, 467–469.

(7) Garazd, M. M.; Muzychka, O. V.; Voyk, A. I.; Nagorichna, I. V.; Ogorodniichuk, A. S. *Chem. Nat. Compd.* **2007**, *43*, 19–23.

(8) Krzeszewski, M.; Vakuliuk, O.; Gryko, D. T. *Eur. J. Org. Chem.* **2013**, *25*, 5631–5644.