Electronic Supplementary Material (ESI) for Chemical Communications.


# Electronic Supplementary Information

## UGT73F17, a new glycosyltransferase from *Glycyrrhiza uralensis*, catalyzes the regiospecific glycosylation of pentacyclic triterpenoids

Jun-bin He,[‡a] Kuan Chen,[‡a] Zhi-min Hu,[a] Kai Li,[a] Wei Song,[a] Li-yan Yu,[b] Chung-Hang Leung,[c] Dik-Lung Ma,[d] Xue Qiao*[a] and Min Ye*[a]

[a] *State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, 38 Xueyuan Road, Beijing 100191, China.*

[b] *Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China*

[c] *State Key Laboratory of Quality Research in Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Macau, China*

[d] *Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China*

*Email: qiaoxue@bjmu.edu.cn (X. Qiao), or yemin@bjmu.edu.cn (M. Ye).

[‡]J.H. and K.C. contributed equally to this work.

# Table of Contents

## 1. Experimental Procedures

### 1.1 General remarks

The known compounds 18*β*-glycyrrhizic acid (**1**), 18*α*-glycyrrhizic acid (**2**), 18*β*-glycyrrhetinic acid (**3**), 18*α*-glycyrrhetinic acid (**4**), licorice saponins G2 (**5**), **6**–**12**, **16**–**20**, **24**–**48**, **51**, and **52** in the acceptor library were isolated from *Glycyrrhiza uralensis*, *G. inflata*, *G. glabra*, *Antrodia cinnamomea* and the traditional Chinese drug ChanSu as we had previously reported.[1] Substrates **13**, **15**, **21**–**23**, **49**, **50** and **53**–**63** were purchased from Nanjing Zelang medical technology Co., Ltd. (Nanjing, China). Substrate **14** was chemically synthesized by hydrolysis reaction. Methanol and acetonitrile (Fisher Scientific, USA) were of HPLC grade. All other chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) and Beijing Chemical Corporation (Beijing, China) unless otherwise specified. The analyses of substrate specificity and the determinations of conversion rates were analyzed by HPLC on an Agilent 1260 instrument (Agilent Technologies, Waldbronn, Germany). The glycosylated products were isolated and purified by semi-preparative HPLC on an Agilent 1200 instrument (Agilent). LC/MS analysis was performed on an LCQ Advantage ion-trap mass spectrometer (Thermo Fisher, CA, USA). HR-ESI-MS spectra were obtained on a Bruker APEX IV FT-MS spectrometer. NMR spectra were recorded on a Bruker AVANCE III-400 instrument at 400 ($^{1}$H) and 100 ($^{13}$C) MHz, or a Bruker AVANCE III-600 instrument operating at 600 MHz for $^{1}$H and 150 MHz for $^{13}$C in pyridine-$d_5$, using TMS as reference. Chemical shifts ($\delta$) are given in parts per million (ppm). and coupling constants ($J$) are given in hert (Hz).

### 1.2 Plant materials

Newly dried seeds of *Glycyrrhiza uralensis* Fisch. were collected in autumn 2015 in Inner Mongonia Autonomous Region of China, and were placed in MS medium to obtain aseptic seedlings. Two week-old seedlings were identified by DNA barcoding, and were then used for total RNA extraction.

### 1.3 Transcriptome data assembly and candidate gene screening

All transcriptome data of *G*. *uralensis* used in this study were downloaded from the Sequence Read Archive (SRA) database (Table S1). The clean reads were *de novo* assembled using Trinity program (K-mer = 31) with default parameters,[2] and the unigenes were then compared with the reported nucleotide sequences of glycosyltransferases (GTs) using the local Basic Local Alignment Search (Blast) Tool (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST) with an e-value of $1e^{-21}$. The best alignment results were further compared with the public databases including the NCBI non-redundant nucleotide database (Nt), non-redundant protein database (Nr, http://www.ncbi.nlm.nih.gov/), Swiss-Protdatabase (http://www.expasy.ch/sprot) using BLASTN and BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi), respectively, with an e-values of less than $1e^{-5}$. Finally, 43 putative GT genes were screened out.

**1.4 Molecular cloning, heterologous expression and protein purification**

The total RNA of the aseptic seedling of *G*. *uralensis* was extracted by using the Tran*Zol™* kit (TransGen Biotech, China) and was reverse-transcribed to cDNA with the FastQuant RT Kit (Tiangen Biotech, China). The full-length cDNA of the *GuGTs* were amplified by PCR using the gene-specific primer pairs. The amplification product of *Gu*GT2 was inserted into pET-28a (+) vector (Invitrogen, USA) according to the Quick-change method. After verification of the sequences, the recombinant plasmid pET-28a-*GuGT*2 was transformed into *Escherichia coli* BL21 (DE3) (TransGen Biotech, China) for heterologous expression. *E*. coli cells were grown in 500 mL Luria-Bertani (LB) medium containing 50 μg/ml kanamycin at 37 °C with shaking (200 rpm). After $OD_{600}$ reached 0.4–0.6, the cells were induced with 0.1 mM Isopropyl β-D-thiogalactoside (IPTG) for 20 h at 16 °C. The cell pellets were collected by centrifugation (8000 rpm for 10 min at 4 °C) and washed with distilled water. Then the cells were resuspended in 15 mL of lysis buffer (pH 8.0, 50 mM $NaH_2PO_4$, 30 mM NaCl, 10 mM imidazole) and ruptured by sonication on ice. The cell debris were removed by centrifugation at 12000 rpm for 45 min at 4 °C. The supernatant was subsequently mixed with 4 mL Ni-NTA resin (TransGen Biotech, China) and incubated for 2 h in an ice bath. Then the mixture was applied to an Affinity Chromatographic Column (TransGen Biotech, China) which was pre-equilibrated with lysis buffer. The recombinant protein was eluted by 20 mL (5 column volumes) elution buffer containing 250 mM imidazole. After SDS-PAGE analysis (purity >95%, Figure S2), the purified protein was concentrated and desalted by a 30 kDa ultrafiltration tube (Solarbio, China) with storage buffer (50 mM Tris-HCl, 20% glycerol, pH 7.5). The protein concentration for all studies was determined using the Bradford method.

**1.5 Effects of reaction time, pH, temperature and divalent metal ions**

To characterize the enzymatic properties of *Gu*GT2 (UGT73F17), the reaction time of glycosylation was firstly determined for 30 min (Figure S3). To optimize the reaction pH, the enzymatic reactions were carried out in various reaction buffers ranged in pH values from 4.0–5.0 (Citric acid-sodium citrate buffer), 6.0–8.0 ($Na_2HPO_4$-$NaH_2PO_4$ buffer), 7.0–9.0 (Tirs-HCl buffer), and 9.0–11.0 ($Na_2CO_3$-$NaHCO_3$ buffer). To investigate the optimal reaction temperature, the enzymatic reactions were incubated at different temperatures (4–60 °C). To test the dependence of divalent metal ions for UGT73F17 activity, different divalent cations $Mg^{2+}$, $Ca^{2+}$, $Ba^{2+}$, $Fe^{2+}$, $Co^{2+}$, $Zn^{2+}$ and EDTA in the final concentration of 5 mM were used, individually. All enzymatic reactions were conducted with UDP-glucose (UDP-Glc) as the donor and glycyrrhizic acid (**1**) as the acceptor. All experiments were performed in triplicate and the mean value was used. The reactions were terminated with pre-cooled methanol (MeOH) and centrifuged at 12,000 rpm for 20 min for further HPLC analysis as described in general methods. The samples were analyzed on an Agilent Zorbax SB C18 column (250 mm×4.6mm, 5 μm) protected with a Zorbax Extend-C18 guard column (4.6 ×12.5 mm, 5 μm) at a flow rate of 1 mL $min^{-1}$. The column temperature was 30 °C and the enzymatic products were eluted with a linear gradient of 20% to 100% MeOH in $H_2O$ containing 0.1% formic acid in 20 min, and followed by 100% MeOH for 5 min. The conversion rates in

percent were calculated from peak areas of glycosylated products and substrates as analyzed by HPLC (Figure S3).

**1.6 Kinetic studies**

For kinetic studies of UGT73F17, enzymatic assays were performed in a final volume of 100 μL containing 50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer (pH 7.5), 1 μg of purified UGT73F17, 500 μM of saturated UDP-Glc, and varying concentrations (2–40 μM) of glycyrrhizic acid (**1**) and glycyrrhetinic acid (**3**). The reactions were conducted at 45 °C for 30 min, and were then quenched with 100 μL ice cold MeOH. Sample were centrifuged at 12,000 rpm for 20 min and analyzed by reversed-phase HPLC as described above. All experiments were performed in triplicate. The Michaelis-Menten constants ($K_m$) values were calculated by using the Lineweaver-Burk plot method (Figure S4).

**1.7 Glycosyltransferase activity assays**

To investigate the substrate promiscuity and specificity of UGT73F17, different acceptors including triterpenoids (**1**–**33**), diterpenoid (**34**), steroids (**35**–**37**), flavonoids (**38**–**46**), coumarin (**47**), benzofuran (**48**), anthraquinone (**49**), and aromatic compounds with −OH, −SH and $-NH_2$ groups (**50**–**63**) were tested (Figure 2 and Figure S5). The reaction mixtures were individually performed in a final volume of 100 μL containing 50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer (pH 7.5), 5 μg of purified UGT73F17, 0.1 mM aglycone, 0.5 mM UDP-Glc and 50 mM $Mg^{2+}$. All reactions were incubated at 45 °C for 12 h and quenched with 100 μL ice cold MeOH. The mixtures were then centrifuged at 12,000 rpm for 20 min and analyzed by reversed-phase HPLC as mentioned above and LC-MS as we had previously reported.[3] To explore the selectivity of UGT73F17 for other sugar donors, glycosylation reactions with UDP-glucuronic acid (UDP-GluA), UDP-xylose (UDP-Xyl), UDP-galactose (UDP-Gal) and UDP-arabinose (UDP-Ara)) were conducted using compounds **1**, **3** and **6** as acceptor, respectively. The reaction mixtures were individually performed in a final volume of 100 μL containing 50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer (pH 7.5), 2.5 μg of purified UGT73F17, 0.1 mM aglycone, 0.5 mM UDP-Glc and 50 mM $Mg^{2+}$. All the reactions were carried out as described above and then analyzed by LC/MS (Figures S6–S15).

**1.8 Preparative-scale reactions and structural identification of glycosylated products**

The substrates (25 μmol) were dissolved in 1 mL dimethyl sulfoxide (DMSO) and were diluted with reaction buffer (50 mM $NaH_2PO_4$-$Na_2HPO_4$, pH 7.5, 20 mL total volume). UDP-Glc (50 μmol) was added along with 3 mg of purified UGT73F17. The reactions were incubated at 45 °C for up to 12 h and then terminated by the addition of 80 mL of acetonitrile. The mixtures were then centrifuged at 12,000 rpm for 30 min, and the supernatants were concentrated and dissolved in 1 mL of 70% methanol. The glycosylated products were subsequently separated by reversed-phase semi-preparative HPLC and were characterized by HR-ESI-MS and NMR (Figures S16–S47).

**1.9 Anti-HIV Activity**

The anti-HIV activities of glycosylated products (**2a**–**5a**) were evaluated according to our previous work.[1a] 293T cells ($2 \times 10^5$) were co-transfected with 0.6 mg of pNL-Luc-E and 0.4 mg of pHIT/G. The VSV-G pseudo-typed viral supernatant (HIV-1) was harvested by filtration through a 0.45 μm filter after 48 h, and the concentration of viral capsid protein was measured by p24 antigen capture ELISA (Biomerieux). SupT1 cells were exposed to VSV-G pseudo-typed HIV-1 (MOI = 1) at 37.8 °C for 48 h in the presence of test products. Efavirenz was used as the positive control and all the experiments were conducted in triplicate. A luciferase assay system (Promega) was used to determine the inhibition rates.[4]

**2. HR-ESI-MS, $^1$H and $^{13}$C NMR data of glycosylated triterpenoids (2a–5a)**

**18*α*-glycyrrhizic acid 30-*O*-*β*-D-glucoside (2a):** 38% yield; white, amorphous powder (MeCN); $[\alpha]^{25}_D$ −128.00 (*c* 0.00625, MeCN); HRESIMS *m/z* 983.4492 $[M-H]^-$ (calcd. for $C_{48}H_{72}O_{21}$, 983.4488, Δ=0.4 ppm); $^1$H-NMR (400 MHz, pyridine-$d_5$) *δ*: 3.38 (1H, dd, *J* = 11.5, 4.1 Hz, H-3), 0.73 (1H, br d, *J* = 11.6 Hz, H-5), 2.21 (1H, s, H-9), 5.59 (1H, s, H-12), 6.41 (1H, d, *J* = 8.0 Hz, H-1'''); $^{13}$C-NMR (100 MHz, pyridine-$d_5$) *δ*: 39.9 (C-1), 27.2 (C-2), 89.7 (C-3), 40.4 (C-4), 55.9 (C-5), 18.2 (C-6), 34.4 (C-7), 44.4 (C-8), 61.4 (C-9), 37.5 (C-10), 199.4 (C-11), 123.5 (C-12), 165.6 (C-13), 45.5 (C-14), 27.3 (C-15), 38 (C-16), 36 (C-17), 40.8 (C-18), 32.2 (C-19), 43.6 (C-20), 29.5 (C-21), 36.5 (C-22), 28.6 (C-23), 17.4 (C-24), 19.1 (C-25), 21.1 (C-26), 21.3 (C-27), 16.5 (C-28), 22.3 (C-29), 177.8 (C-30), 107.4 (C-1'), 172.9 (C-6'), 107.4 (C-1''), 172.6 (C-6''), 96.8 (C-1'''), 65.3 (C-6''').

**18*β*-glycyrrhetinic acid 30-*O*-*β*-D-glucoside (3a):** 82% yield; white, amorphous powder (MeCN); HRESIMS *m/z* 631.3846 $[M-H]^-$ (calcd. for $C_{36}H_{56}O_9$, 631.3846, Δ=0.0 ppm); $^1$H-NMR (400 MHz, pyridine-$d_5$) *δ*: 3.51 (1H, dd, *J* = 11.6, 4.5 Hz, H-3), 0.89 (1H, br d, *J* = 11.8 Hz, H-5), 2.52 (1H, s, H-9), 6.03 (1H, s, H-12), 6.42 (1H, d, *J* = 8.1 Hz, H-1'); $^{13}$C-NMR (100 MHz, pyridine-$d_5$) *δ*: 40.3 (C-1), 27.2 (C-2), 80.0 (C-3), 40.3 (C-4), 55.8 (C-5), 18.4 (C-6), 33.5 (C-7), 46.0 (C-8), 62.7 (C-9), 38.1 (C-10), 200.2 (C-11), 129.3 (C-12), 169.7 (C-13), 43.9 (C-14), 27.3 (C-15), 28.7 (C-16), 32.6 (C-17), 48.7 (C-18), 42.0 (C-19), 44.8 (C-20), 31.8 (C-21), 38.3 (C-22), 28.4 (C-23), 17.1 (C-24), 17.3 (C-25), 19.3 (C-26), 23.8 (C-27), 29.0 (C-28), 29.2 (C-29), 176.3 (C-30), 96.5 (C-1'), 74.8 (C-2'), 78.4 (C-3'), 71.7 (C-4'), 79.3 (C-5'), 62.8 (C-6').[5]

**18*α*-glycyrrhetinic acid 30-*O*-*β*-D-glucoside (4a):** 45% yield; white, amorphous powder (MeCN); $[\alpha]^{25}_D$ −32.00(*c* 0.025, MeCN); HRESIMS *m/z* 677.3895 $[M-H+HCOOH]^-$ (calcd. for $C_{36}H_{56}O_9$, 677.3901, Δ=0.9 ppm); $^1$H-NMR (600 MHz, pyridine-$d_5$) *δ*: 3.53 (1H, m, H-3), 0.89 (1H, br d, *J* = 11.9 Hz, H-5), 2.39 (1H, s, H-9), 5.66 (1H, s, H-12), 2.26 (1H, d, *J* = 12.3 Hz, H-18), 6.43 (1H, d, *J* = 8.1 Hz, H-1'); $^{13}$C-NMR (150 MHz, pyridine-$d_5$) *δ*: 40.2 (C-1), 28.7 (C-2), 80.1 (C-3), 39.5 (C-4), 55.9 (C-5), 18.5 (C-6), 34.5 (C-7), 45.6 (C-8), 61.5 (C-9), 37.9 (C-10), 199.5 (C-11), 124.8 (C-12), 165.6 (C-13), 44.5 (C-14), 27.3 (C-15), 36.5 (C-16), 36.0 (C-17), 40.8 (C-18), 32.3 (C-19), 43.6 (C-20), 29.5 (C-21), 38.0 (C-22), 29.2 (C-23), 17.2 (C-24), 16.5 (C-25), 19.1 (C-26), 21.1 (C-27), 17.5 (C-28), 21.3 (C-29), 177.8 (C-30), 96.8 (C-1'), 74.9 (C-2'), 78.4 (C-3'), 71.5 (C-4'), 79.3 (C-5'), 62.7 (C-6').

**licorice-saponin G2 30-*O*-*β*-D-glucoside (5a):** 51% yield; white, amorphous powder (MeCN); HRESIMS *m/z* 999.4437 $[M-H]^-$ (calcd. for $C_{48}H_{72}O_{22}$, 999.4437, Δ=0.0 ppm); $^1$H-NMR (400 MHz, pyridine-$d_5$) $\delta$: 3.50 (1H, dd, $J$ = 11.5, 4.6 Hz, H-3), 2.51 (1H, s, H-9), 5.95 (1H, s, H-12), 6.40 (1H, d, $J$ = 8.1 Hz, H-1'''); $^{13}$C-NMR (100 MHz, pyridine-$d_5$) $\delta$: 39.5 (C-1), 23.8 (C-2), 90.1 (C-3), 44.8 (C-4), 56.4 (C-5), 17.1 (C-6), 33.7 (C-7), 45.0 (C-8), 62.4 (C-9), 37.4 (C-10), 199.9 (C-11), 129.2 (C-12), 169.7 (C-13), 43.8 (C-14), 27.2 (C-15), 27.3 (C-16), 32.6 (C-17), 46.0 (C-18), 40.0 (C-19), 42.0 (C-20), 30.4 (C-21), 38.3 (C-22), 22.2 (C-23), 62.7 (C-24), 16.0 (C-25), 19.0 (C-26), 23.5 (C-27), 28.4 (C-28), 28.9 (C-29), 176.2 (C-30), 105.2 (C-1'), 172.8 (C-6'), 105.9 (C-1''), 172.8 (C-6''), 96.5 (C-1'''), 65.3 (C-6''').

**Jaligonic acid 30-*O*-*β*-D-glucoside (14a):** 67% yield; white, amorphous powder (MeCN); HRESIMS *m/z* 679.3698 $[M-H]^-$ (calcd. for $C_{36}H_{56}O_{12}$, 679.3694, Δ=0.6 ppm); $^1$H-NMR (600 MHz, pyridine-$d_5$) $\delta$: 3.97 (1H, m, H-2), 3.72 (1H, d, $J$ = 10.4 Hz, H-3), 1.36 (3H, s, H-24), 1.54 (3H, s, H-25), 1.27 (3H, s, H-26), 1.05 (3H, s, H-27), 1.30 (3H, s, H-29), 6.39 (1H, d, $J$ = 7.9 Hz, H-1'); $^{13}$C-NMR (150 MHz, pyridine-$d_5$) $\delta$: 44.8 (C-1), 71.2 (C-2), 73.6 (C-3), 40.4 (C-4), 48.7 (C-5), 18.9 (C-6), 30.5 (C-7), 35.7 (C-8), 49.2 (C-9), 33.6 (C-10), 24.5 (C-11), 123.3 (C-12), 145.8 (C-13), 37.8 (C-14), 26.7 (C-15), 23.4 (C-16), 45.5 (C-17), 43 (C-18), 42.8 (C-19), 44.3 (C-20), 29.1 (C-21), 31.6 (C-22), 68.2 (C-23), 15.1 (C-24), 18.4 (C-25), 17.9 (C-26), 25.1 (C-27), 173.7 (C-28), 29 (C-29), 177.2 (C-30), 96.8 (C-1'), 74.6 (C-2'), 78.7 (C-3'), 72.2 (C-4'), 79.9 (C-5'), 62.8 (C-6').

**3. Supplementary Tables**

**Table S1.** Transcriptomes of *Glycyrrhiza uralensis* used in this study.

| No. | Run (accession) | Experiment (accession) | Release Date |
|---|---|---|---|
| 1 | SRR1783599 | SRX862639 | 2016/02/19 |
| 2 | SRR1783600 | SRX862639 | 2016/02/19 |
| 3 | SRR1783602 | SRX862640 | 2016/02/19 |
| 4 | SRR1811619 | SRX862640 | 2016/02/19 |
| 5 | SRR2537378 | SRX1295883 | 2016/02/02 |

**Table S2.** PCR primers used in this study.

| Primers | | Sequences (5' to 3') |
|---|---|---|
| GuGT2 | F-2 | GGGTCGCGGATCCGAATTCGAGCTCATGGACGCTGGTGAAGAACAGCCAC |
| | R-2 | CGAGTGCGGCCGCAAGCTTGTCGACTTAATCAAGTGGCTTACAGTCCCTC |

**Table S3.** Anomeric coupling constants for glycosylated products.

| Glycosylated products | Anomeric coupling constants (*J*, Hz) |
|---|---|
| **2a** | 8.0 |
| **3a** | 8.1 |
| **4a** | 8.1 |
| **5a** | 8.1 |
| **14a** | 7.9 |

**Table S4.** The anti-HIV activities of triterpenoid saponins **2a**–**5a** (100 μM).

| Products | Inhibition rate (%) |
|---|---|
| **2a** | 36.84 ± 7.08 |
| **3a** | 67.25 ± 5.52 |
| **4a** | — |
| **5a** | 20.46 ± 5.48 |

* The inhibition rate of positive control (Efavirenz, 4 nM) was 88.43 ± 3.21%.

## 4. Supplementary Figures

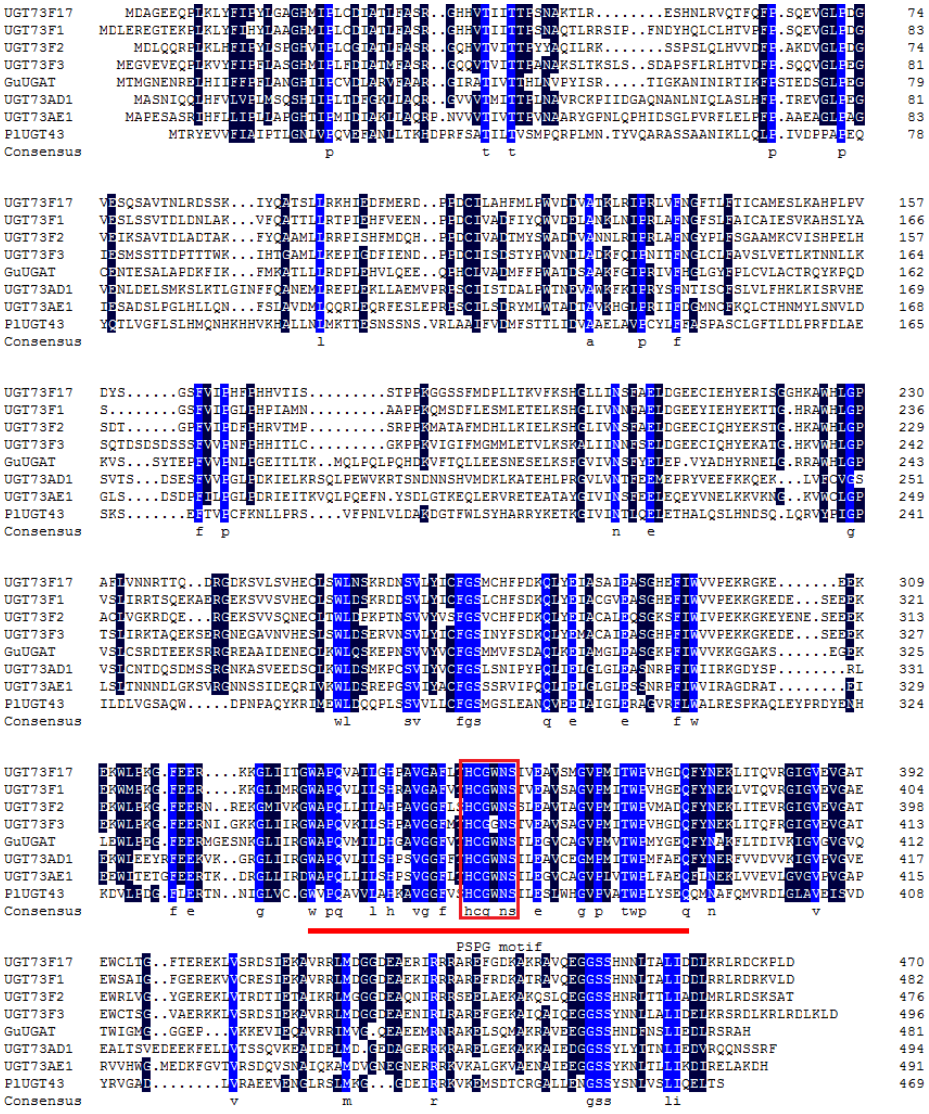


**Figure S1.** Alignments of the amino acid sequence of UGT73F17 and other plant GTs. The sequences were aligned using DNAMAN 8.0. UGT73F17 (in this study) is a pentacyclic triterpenoids glycosyltransferase; UGT73F1 (GenBank accession number BAC78438) is an isoflavonoid-specific glycosyltransferase from *Glycyrrhiza echinata*;[6] UGT73F2 (GenBank accession number BAM29362) is a triterpenoid saponins glycosyltransferase from *Glycine max*;[7] UGT73F3 (GenBank accession number ACT34898, from *Medicago truncatula*)[8] and UGT73AD1 (GenBank accession number ALD84259, from *Centella asiatica*)[9] are the 28-*O*-glycosyltransferase involved in saponin biosynthesis; *Gu*UGAT (GenBank accession number ANJ03631) is a glucuronosyltransferase from *G. uralensis*;[10] UGT73AE1 (GenBank accession number AJT58578, from *Carthamus tinctorius*) is a *O*-, *S*-, and *N*-glycosyltransferase;[11] PlUGT43 (GenBank accession number AMQ26115) is an isoflavonoid *C*-glycosyltransferase from *Pueraria lobata*.[12] Identical amino acids are shown in white on a blue background, and similar residues are marked in white on a mazarine background. The red line shows the conserved region of plant secondary product glycosyltransferases (PSPG motif).

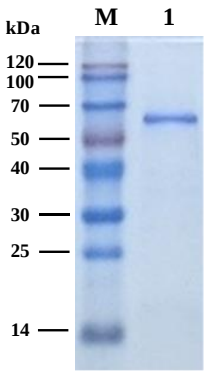


**Figure S2.** SDS-PAGE of the His-tagged UGT73F17 purified by Ni-NTA affinity chromatography. M: Protein Marker; 1: Recombinant UGT73F17 fused with His-tag (predicted M.W.: 58.3 kDa).

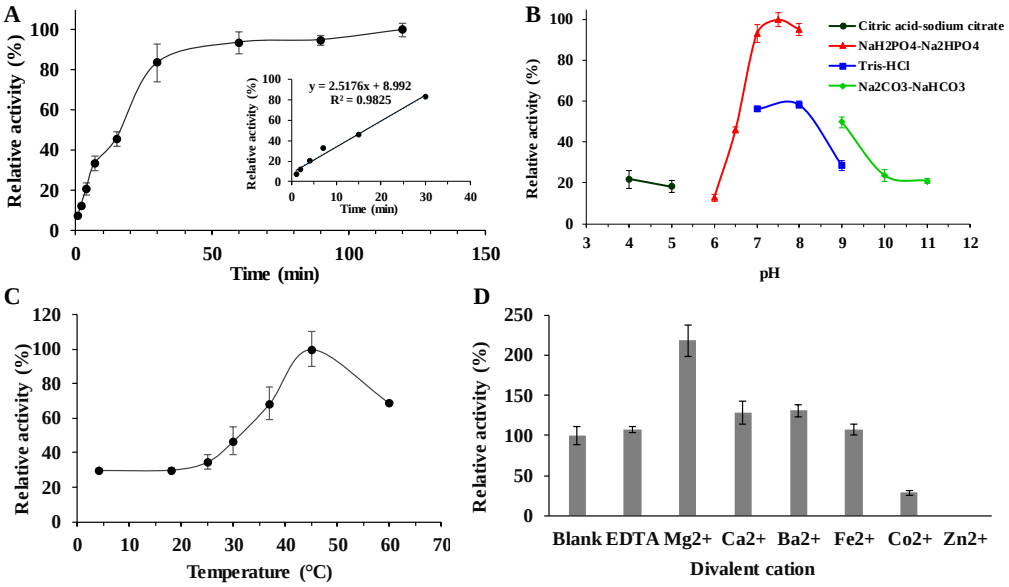


**Figure S3.** Effects of reaction time (A), reaction buffer (B), temperature (C), and divalent metal ions (D) on enzyme activity of UGT73F17. Compound **1** was used as the acceptor and UDP-Glc was used as the sugar donor. An optimized reaction time of 30 min was used. UGT73F17 exhibited its maximum activity at pH 7.5 (50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer) and 45 °C, and $Mg^{2+}$ could remarkably enhance the glycosylation activity.

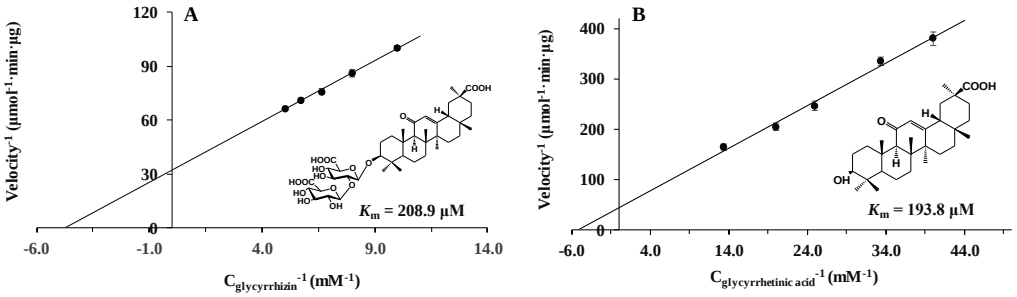


**Figure S4.** Determination of kinetic parameters for purified UGT73F17. The apparent $K_m$ values were determined with glycyrrhizin (**1**) and glycyrrhetinic acid (**3**) as the acceptor and UDP-Glc as the sugar donor at 45 °C and pH 7.5 for 30 min. The apparent $K_m$ values of **1** and **3** were 208.9 μM and 193.8 μM, respectively.

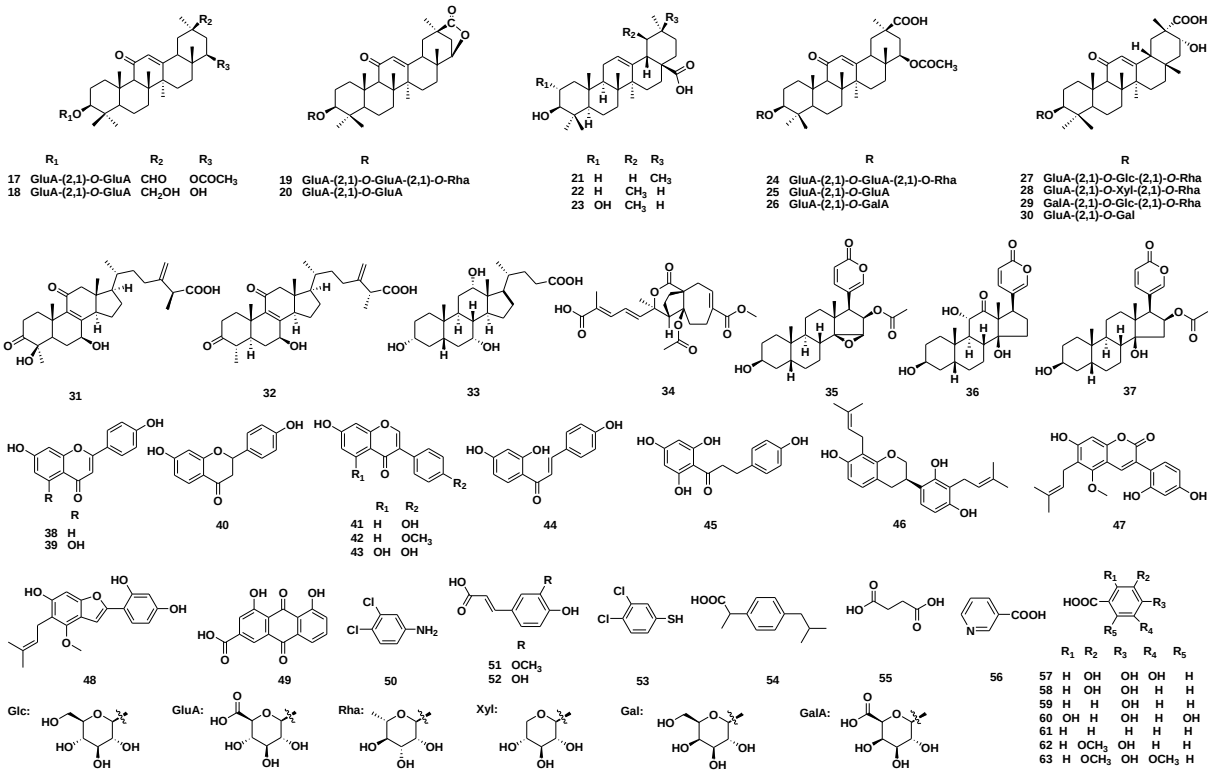


**Figure S5.** Substrates **17**–**63** that are not catalyzed by UGT73F17.

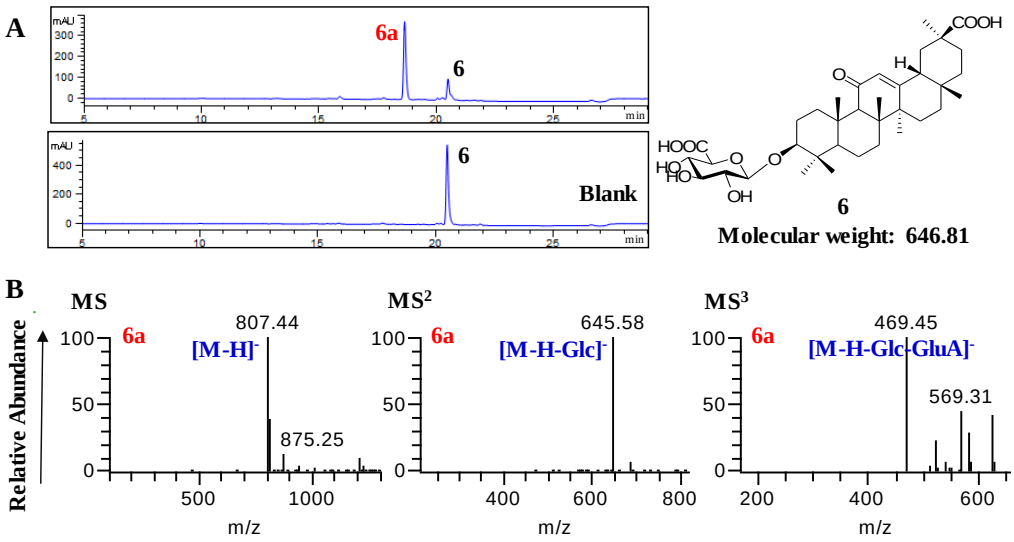


**Figure S6.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures for compound **6**. (A) HPLC chromatogram of **6** and enzymatic product **6a**. (B) (-)-ESI-MS and MS/MS spectra for **6a**. The analysis conditions were described in experimental procedures.

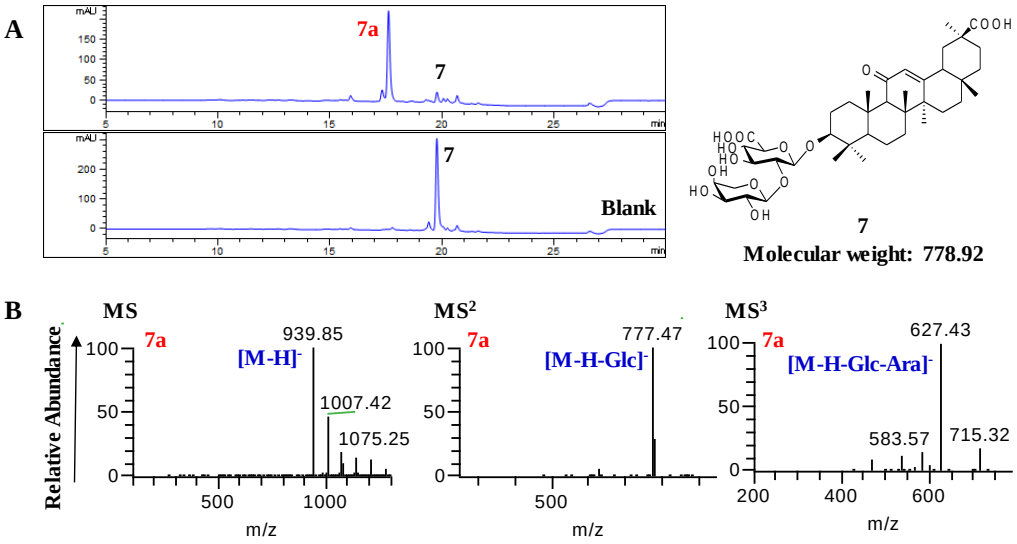


**Figure S7.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures for compound **7**. (A) HPLC chromatogram of **7** and enzymatic product **7a**. (B) (-)-ESI-MS and MS/MS spectra for **7a**. The analysis conditions were described in experimental procedures.

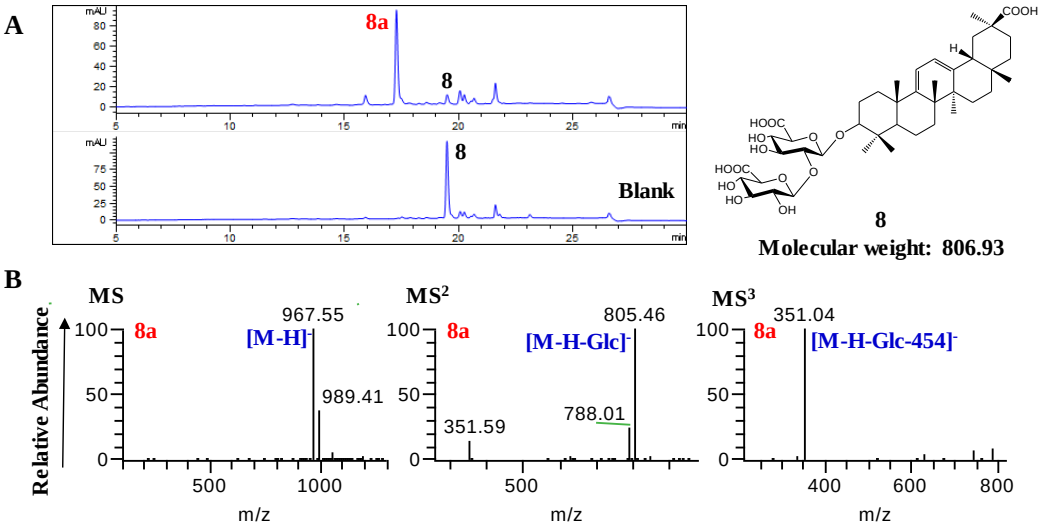


**Figure S8.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures for compound **8**. (A) HPLC chromatogram of **8** and enzymatic product **8a**. (B) (-)-ESI-MS and MS/MS spectra for **8a**. The analysis conditions were described in experimental procedures.

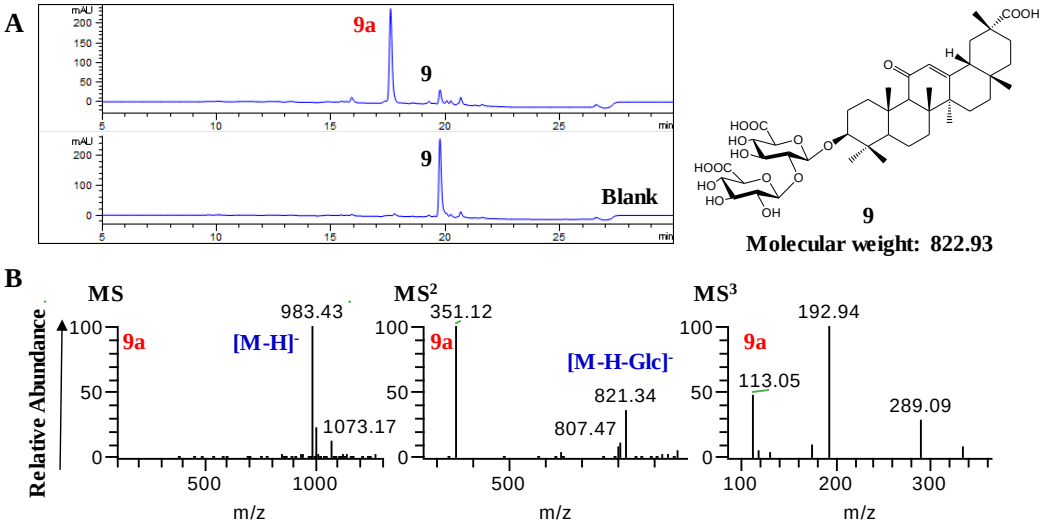


**Figure S9.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures for compound **9**. (A) HPLC chromatogram of **9** and enzymatic product **9a**. (B) (-)-ESI-MS and MS/MS spectra for **9a**. The analysis conditions were described in experimental procedures.

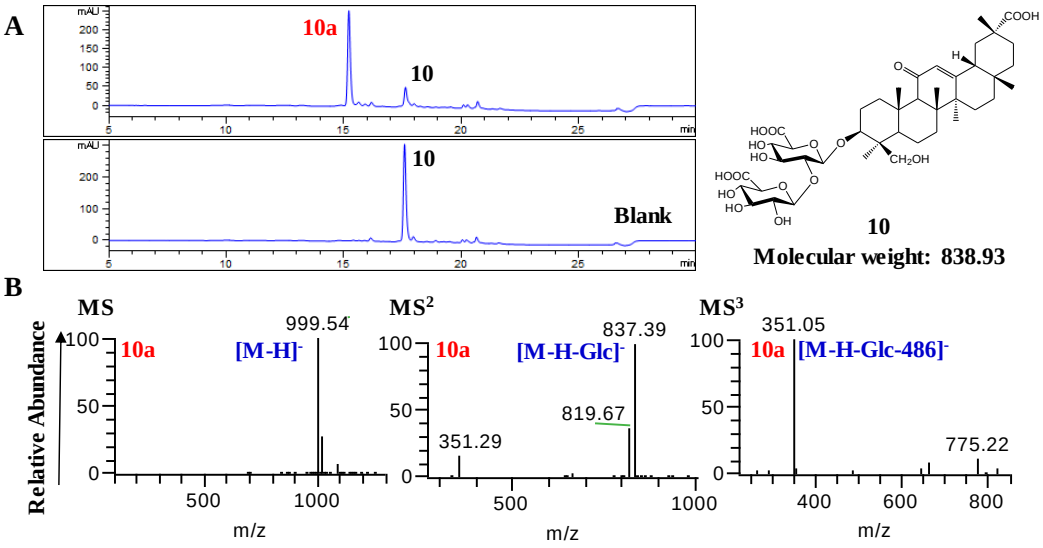


**Figure S10.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures for compound **10**. (A) HPLC chromatogram of **10** and enzymatic product **10a**. (B) (-)-ESI-MS and MS/MS spectra for **10a**. The analysis conditions were described in experimental procedures.

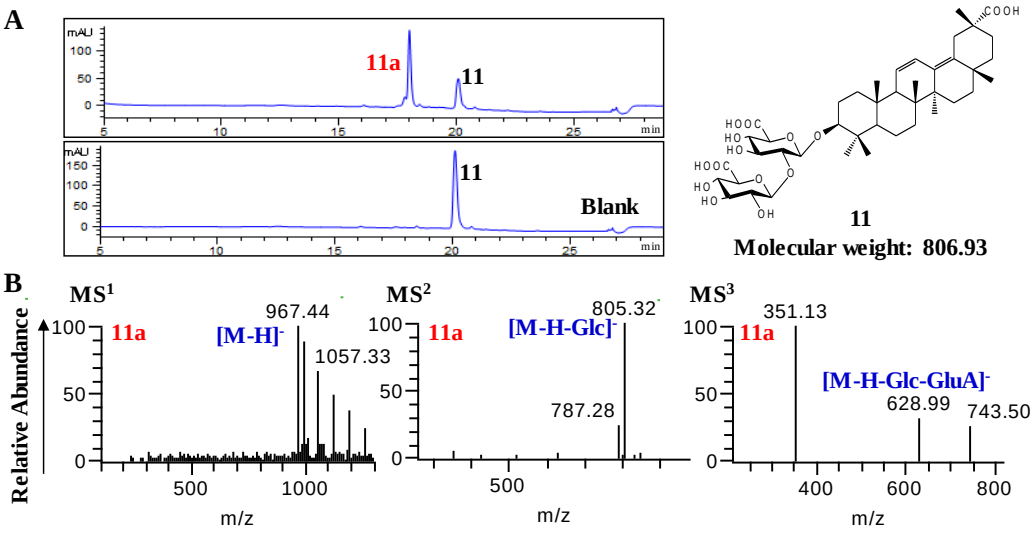


**Figure S11.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures for compound **11**. (A) HPLC chromatogram of **11** and enzymatic product **11a**. (B) (-)-ESI-MS and MS/MS spectra for **11a**. The analysis conditions were described in experimental procedures.

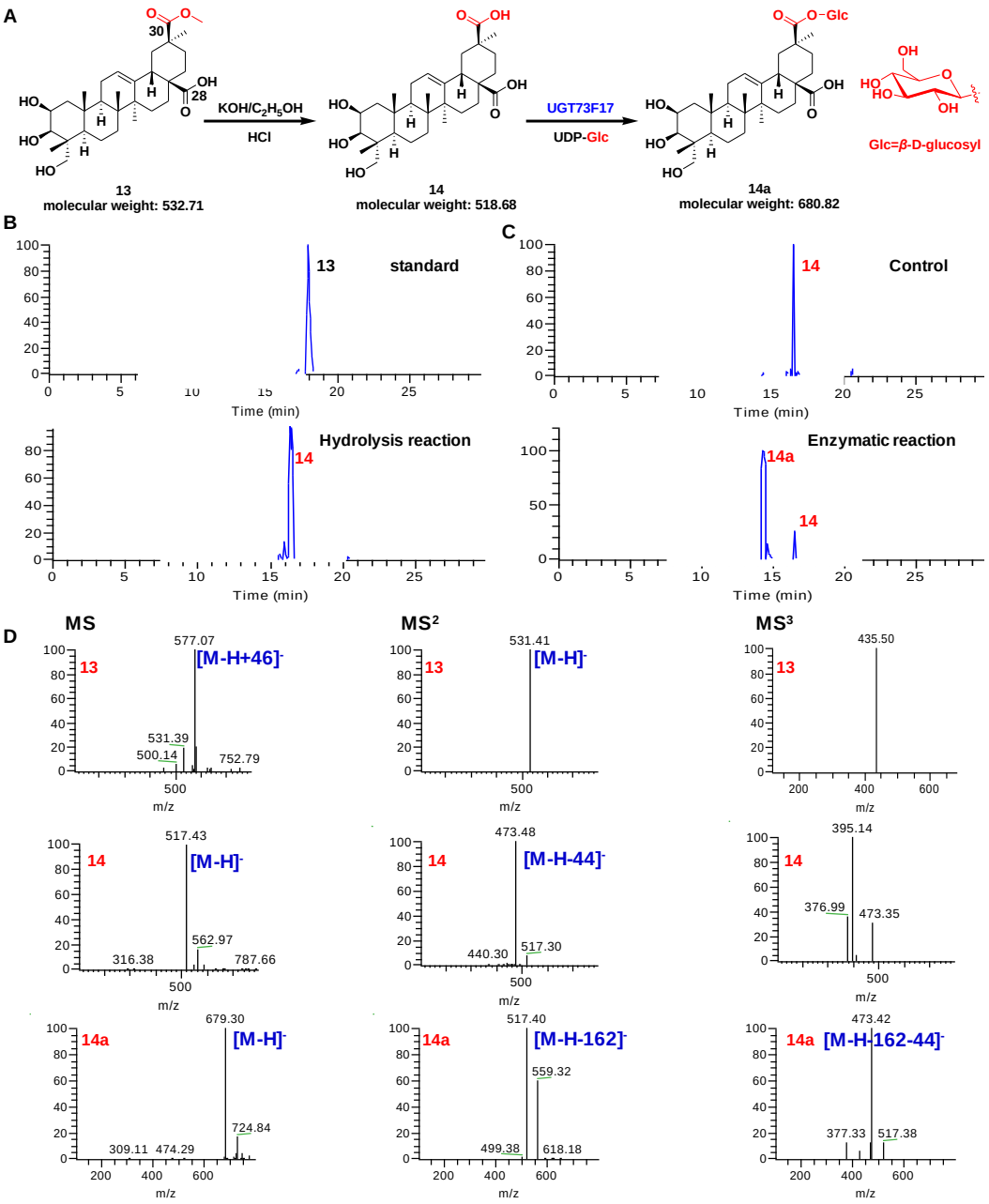


**Figure S12.** LC/MS analysis of UGT73F17 catalytic reaction mixtures for substrate **14**. (A) Hydrolysis reaction of **13** and enzymatic glycosylation of **14** by UGT73F17. (B) Extracted ion chromatograms of **13** and its hydrolysis product **14**. (C) Extracted ion chromatograms of **14** and its enzymatic product **14a**. (D) (-)-ESI-MS and MS/MS spectra for compounds **13**, **14**, and **14a**.

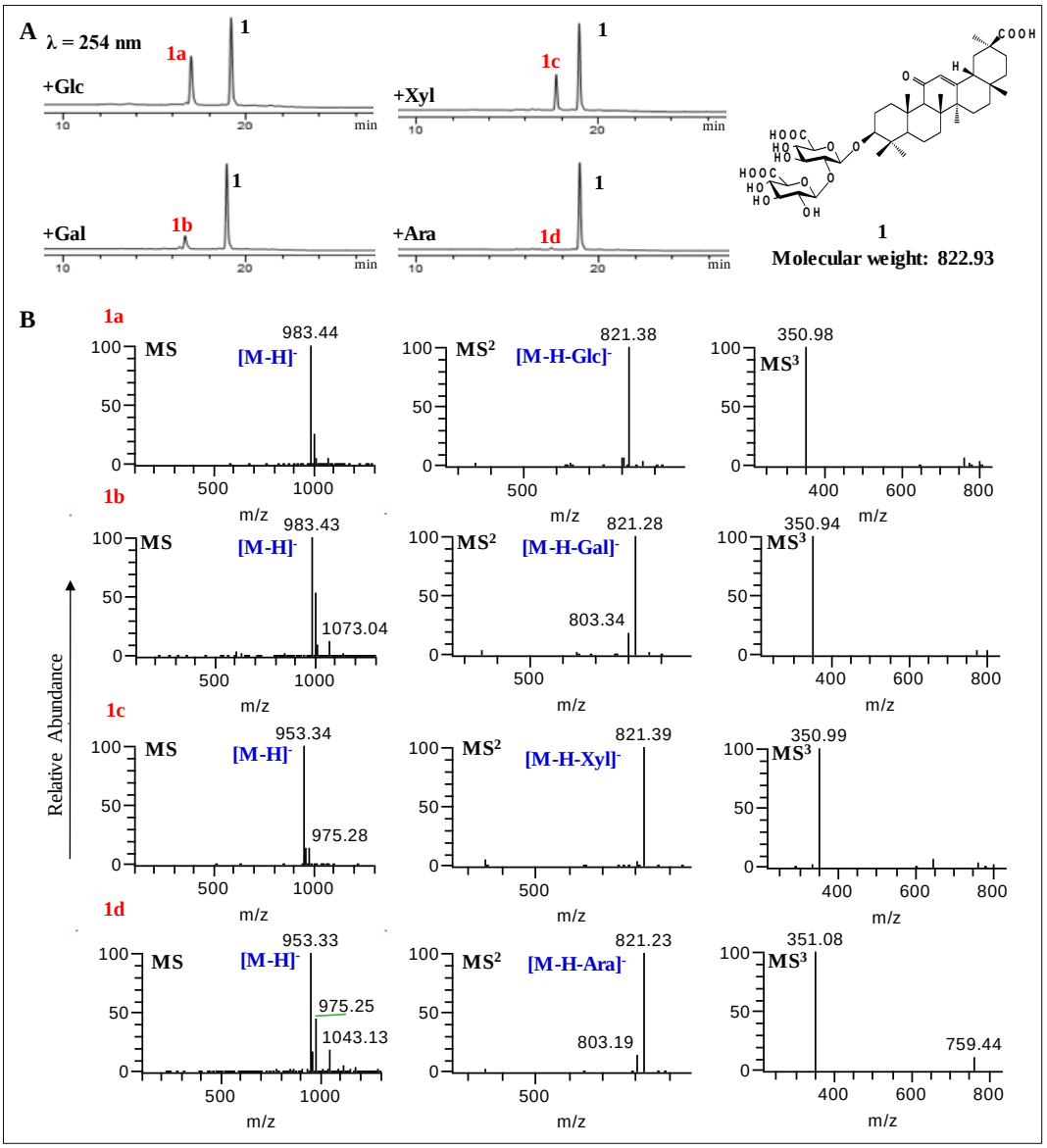


**Figure S13.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures using **1** as the substrate and UDP-Glc/Xyl/Gal/Ara as sugar donors. (A) HPLC chromatogram of **1** and enzymatic product **1a–1d**. (B) (-)-ESI-MS and MS/MS spectra for **1a–1d**. The reaction mixtures were individually performed in a final volume of 100 µL containing 50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer (pH 7.5), 2.5 µg of purified UGT73F17, 0.1 mM **1**, and 0.5 mM UDP-Glc in the presence of 50 mM $Mg^{2+}$. All reactions were incubated at 45 °C for 12 h. The analysis conditions were described in experimental procedures.

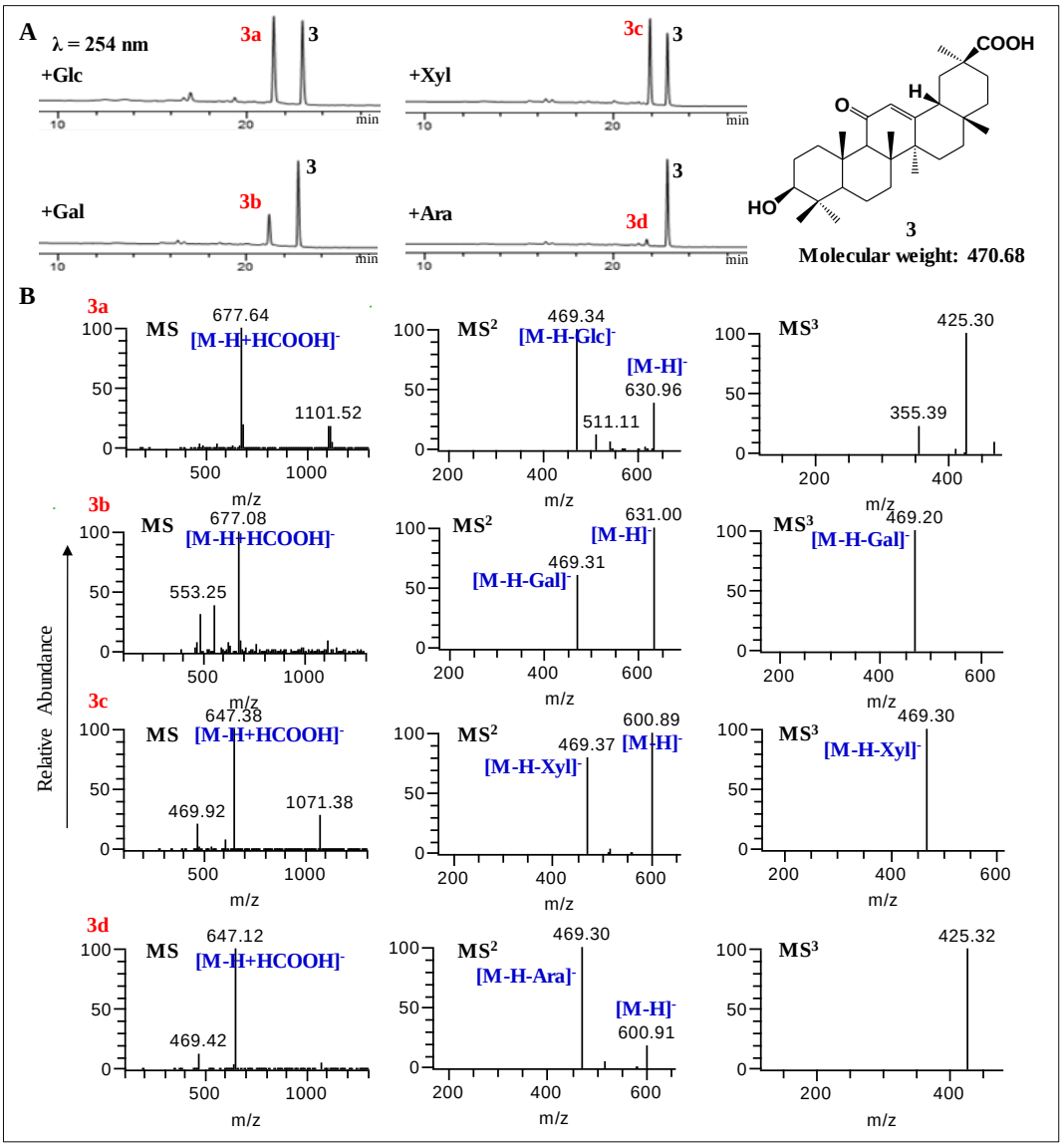


**Figure S14.** HPLC and LC/MS analysis of UGT73F17 catalytic reaction mixtures using **3** as the substrate and UDP-Glc/Xyl/Gal/Ara as sugar donors. (A) HPLC chromatogram of **3** and enzymatic product **3a–3d**. (B) (-)-ESI-MS and MS/MS spectra for **3a–3d**. The reaction mixtures were individually performed in a final volume of 100 µL containing 50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer (pH 7.5), 2.5 µg of purified UGT73F17, 0.1 mM **3**, and 0.5 mM UDP-Glc in the presence of 50 mM $Mg^{2+}$. All reactions were incubated at 45 °C for 12 h. The analysis conditions were described in experimental procedures.

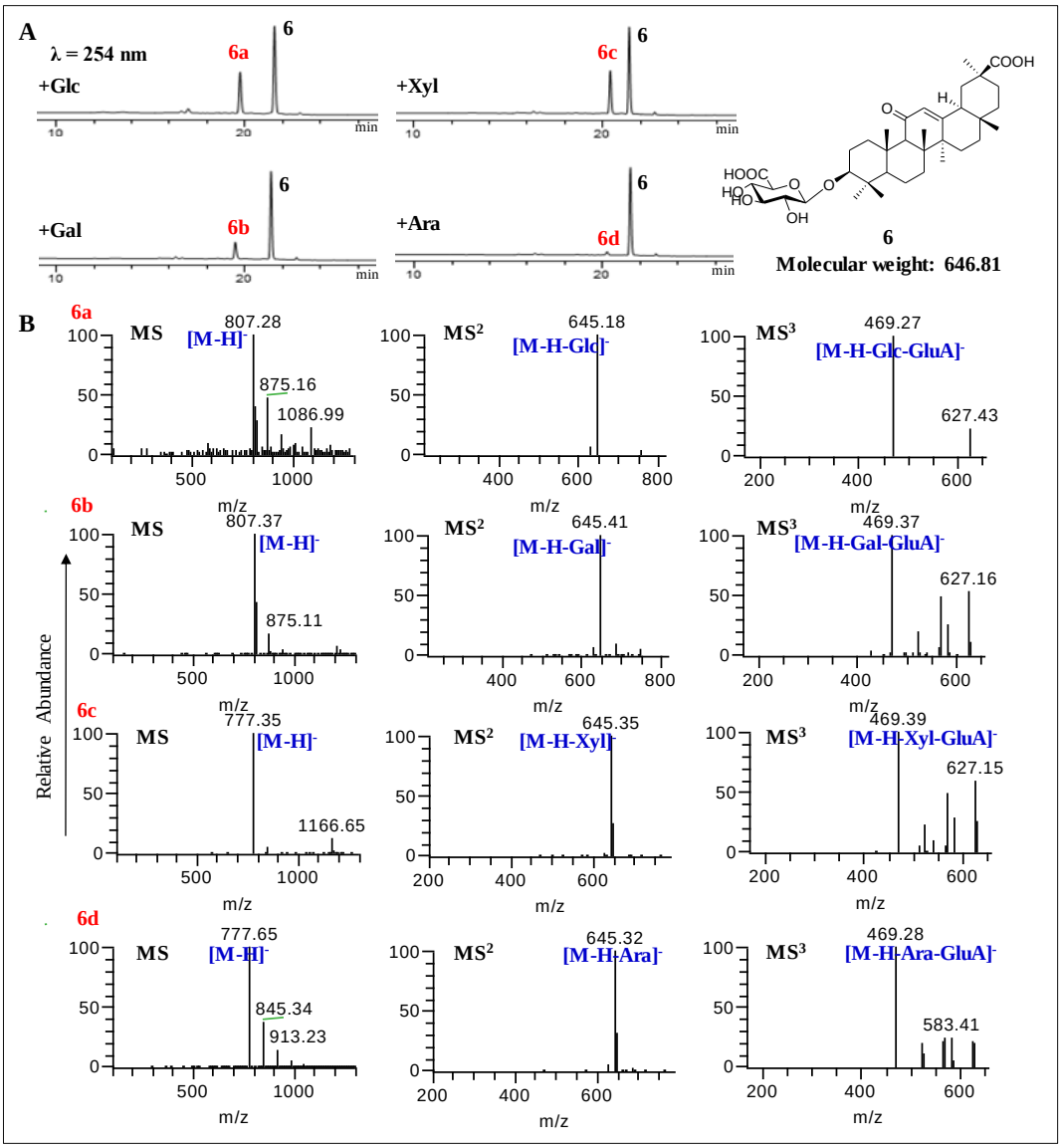


**Figure S15.** HPLC and LC-MS analysis of UGT73F17 catalytic reaction mixtures using **6** as the substrate and UDP-Glc/Xyl/Gal/Ara as sugar donors. (A) HPLC chromatogram of **6** and enzymatic products **6a–6d**. (B) (-)-ESI-MS and MS/MS spectra for **6a–6d**. The reaction mixtures were individually performed in a final volume of 100 µL containing 50 mM $NaH_2PO_4$-$Na_2HPO_4$ buffer (pH 7.5), 2.5 µg of purified UGT73F17, 0.1 mM **6**, and 0.5 mM UDP-Glc in the presence of 50 mM $Mg^{2+}$. All reactions were incubated at 45 °C for 12 h. The analysis conditions were described in experimental procedures.

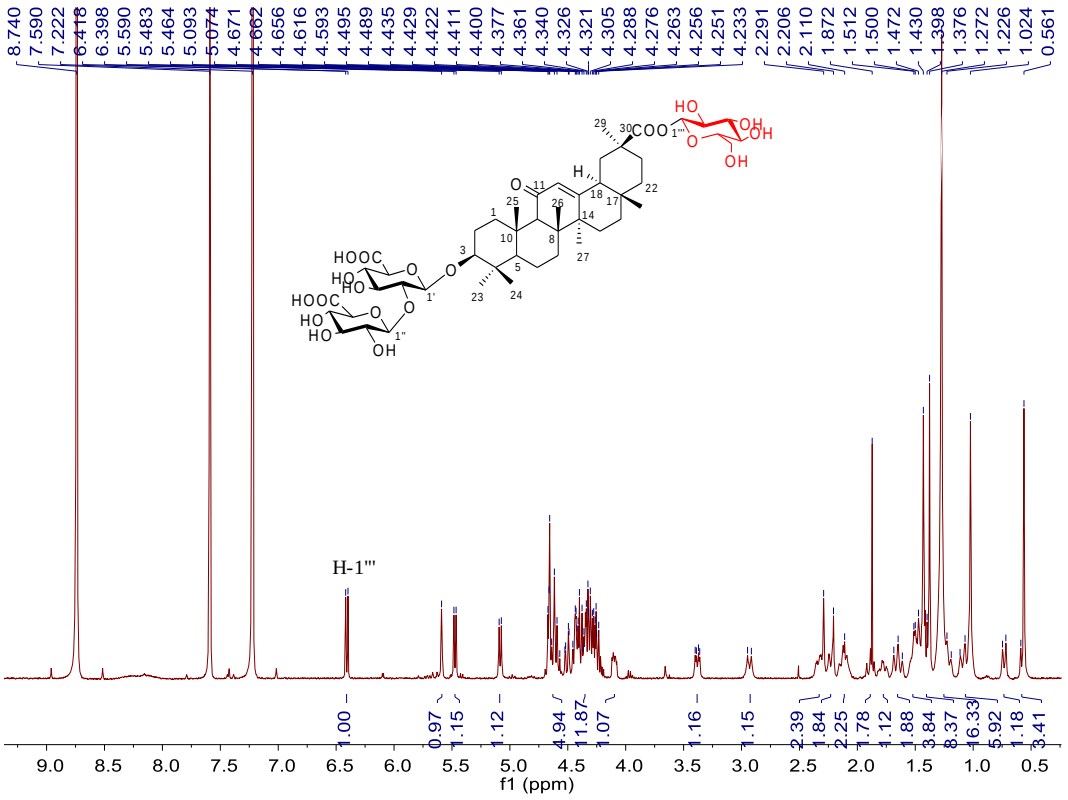


**Figure S16.** $^{1}$H NMR spectrum of **2a** in pyridine-$d_5$ (400 MHz).

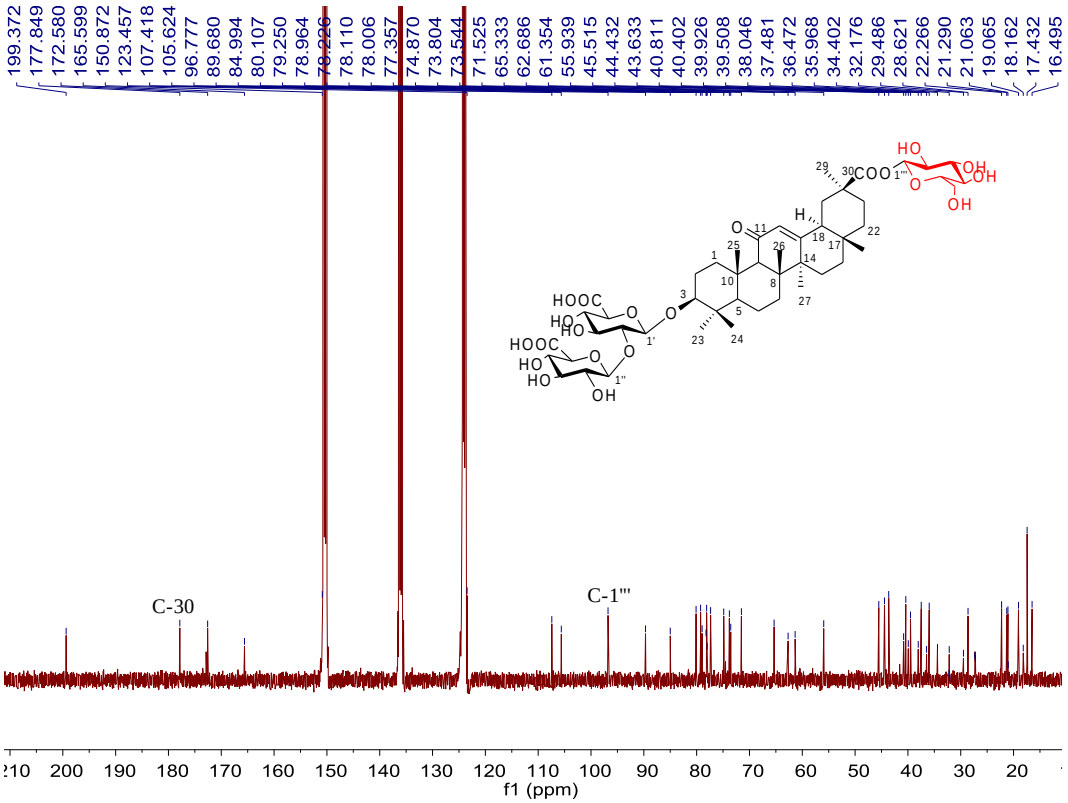


**Figure S17.** $^{13}$C NMR spectrum of **2a** in pyridine-$d_5$ (100 MHz).

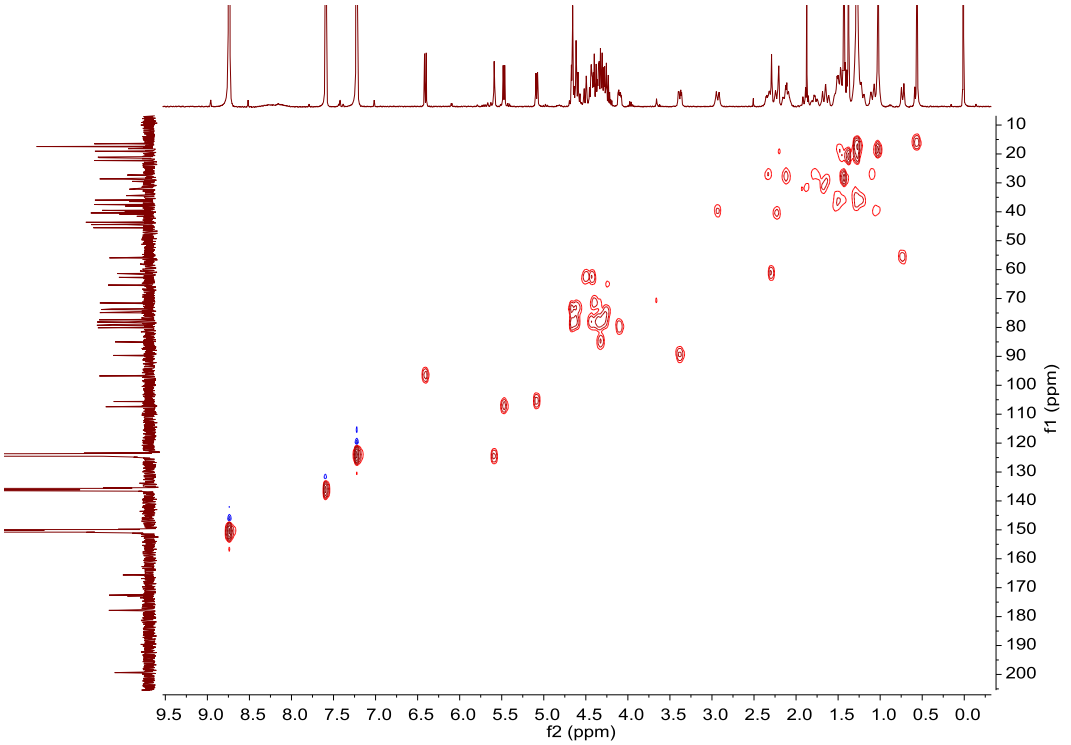


**Figure S18.** HSQC spectrum of **2a** in pyridine-$d_5$ (400 MHz).

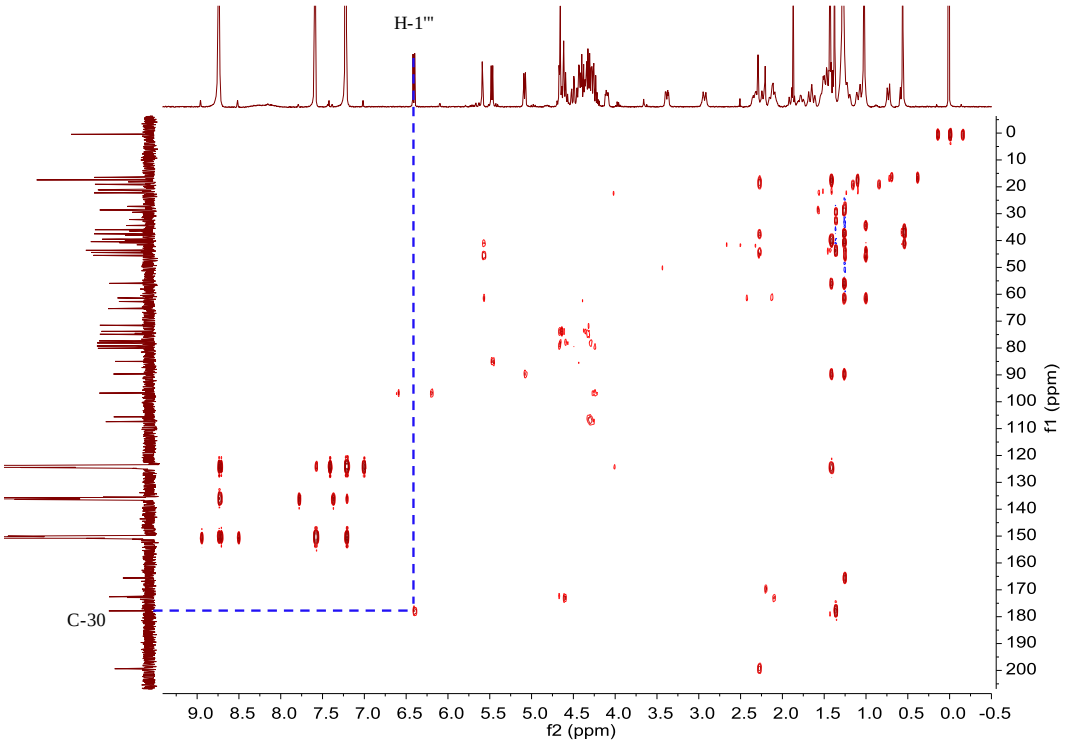


**Figure S19.** HMBC spectrum of **2a** in pyridine-$d_5$ (400 MHz).

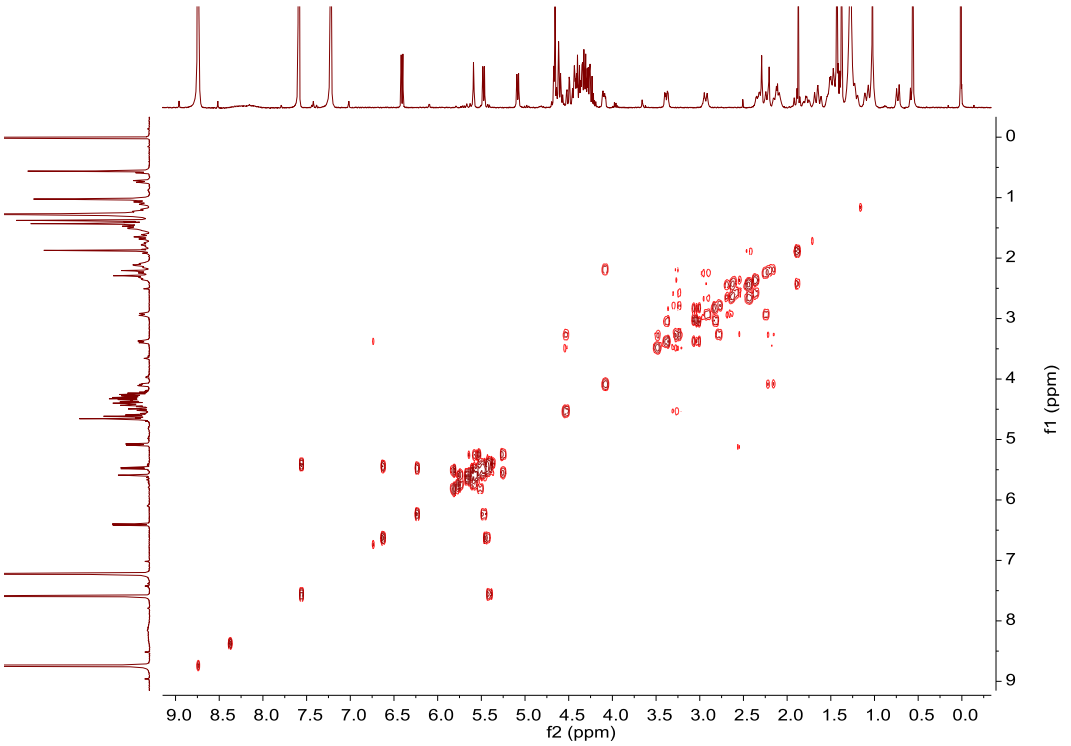


**Figure S20.** $^{1}$H-$^{1}$H COSY spectrum of **2a** in pyridine-$d_5$ (400 MHz).

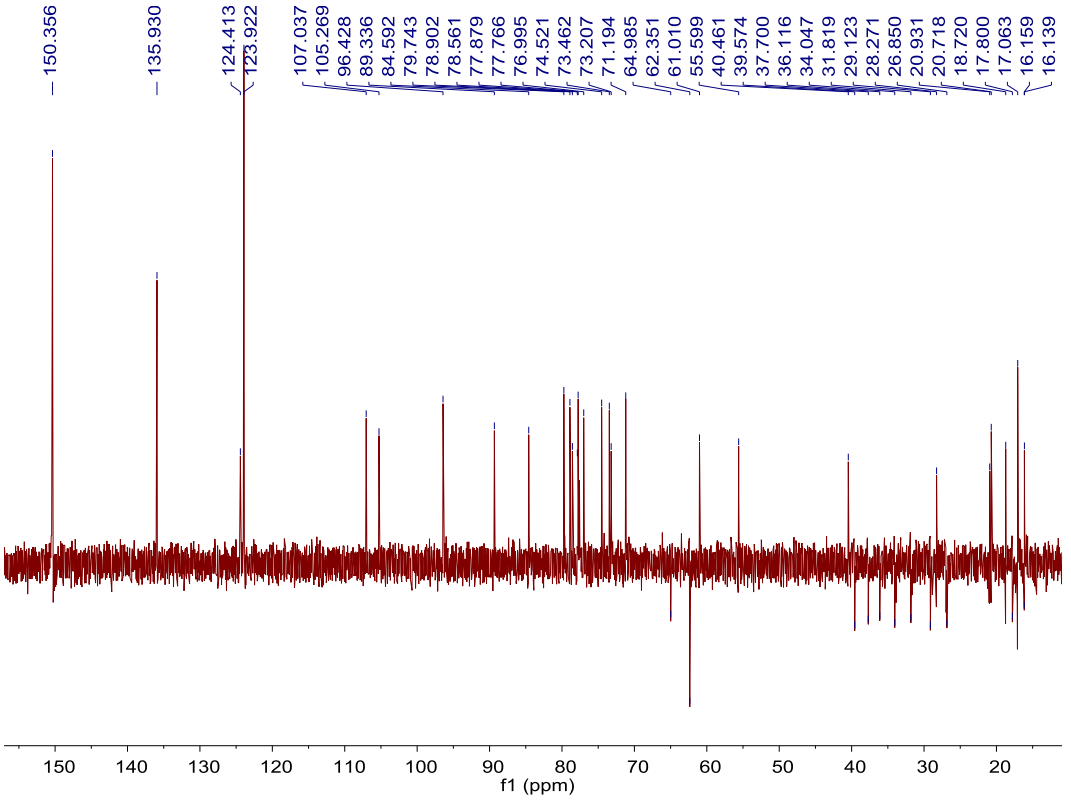


**Figure S21.** DEPT 135 spectrum of **2a** in pyridine-$d_5$ (100 MHz).

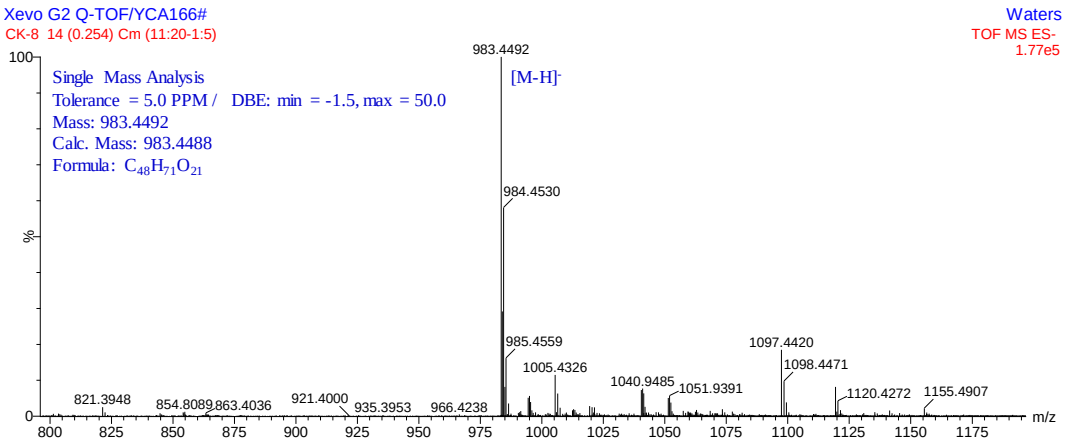


**Figure S22.** HR-ESI-MS spectrum of **2a**.

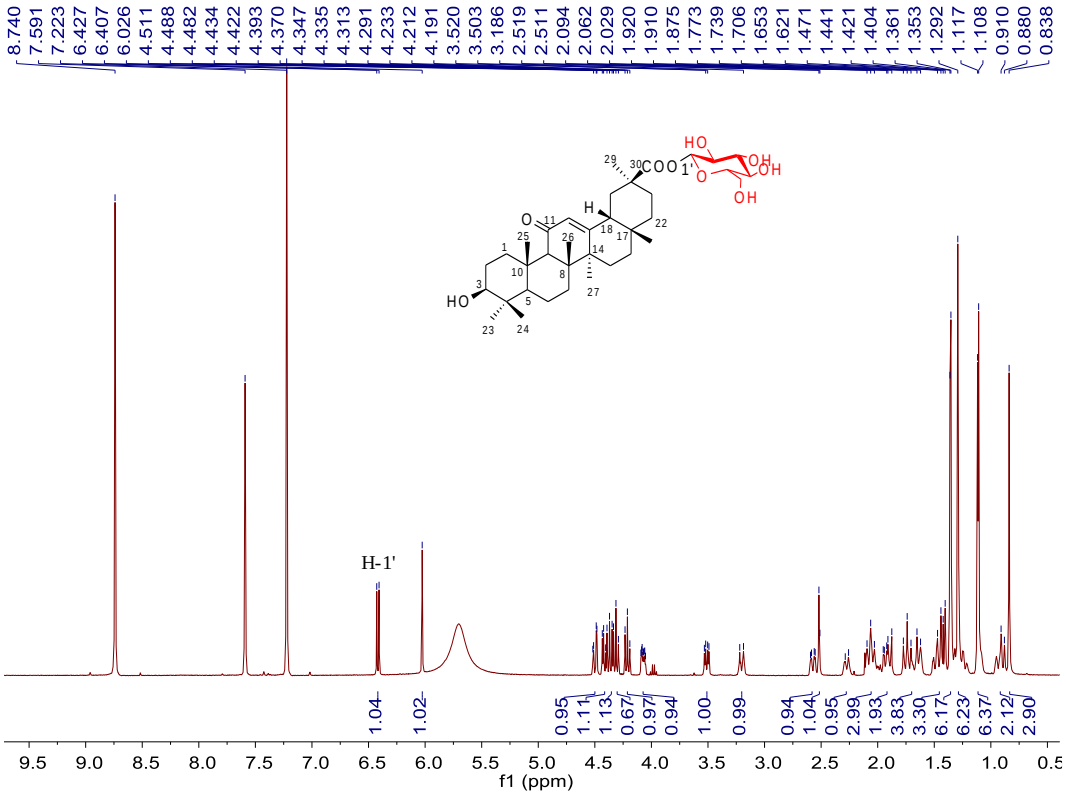


**Figure S23.** $^{1}$H NMR spectrum of **3a** in pyridine-$d_5$ (400 MHz).

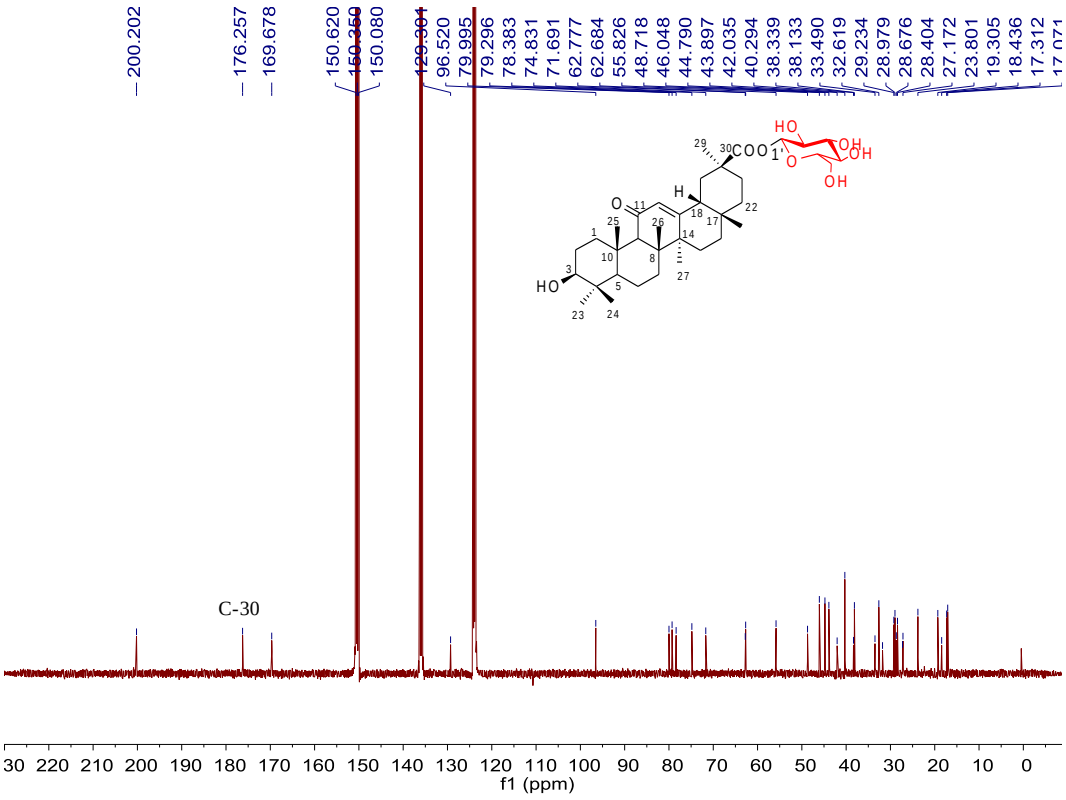


**Figure S24.** $^{13}$C NMR spectrum of **3a** in pyridine-$d_5$ (100 MHz).

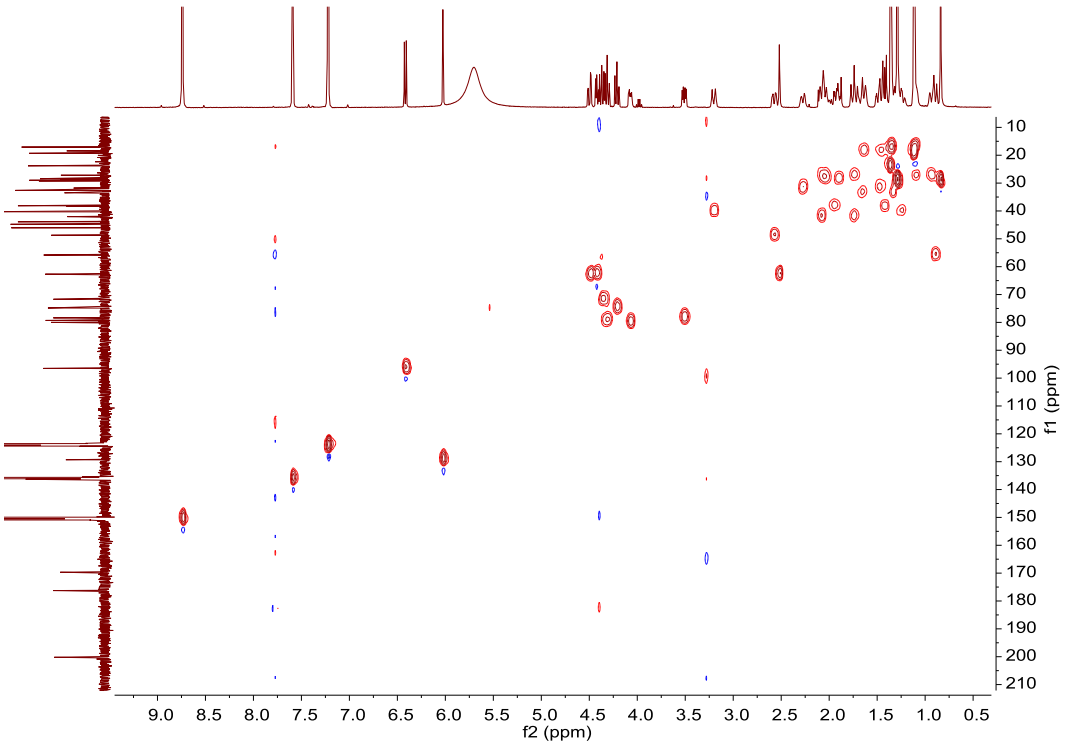


**Figure S25.** HSQC spectrum of **3a** in pyridine-$d_5$ (400 MHz).

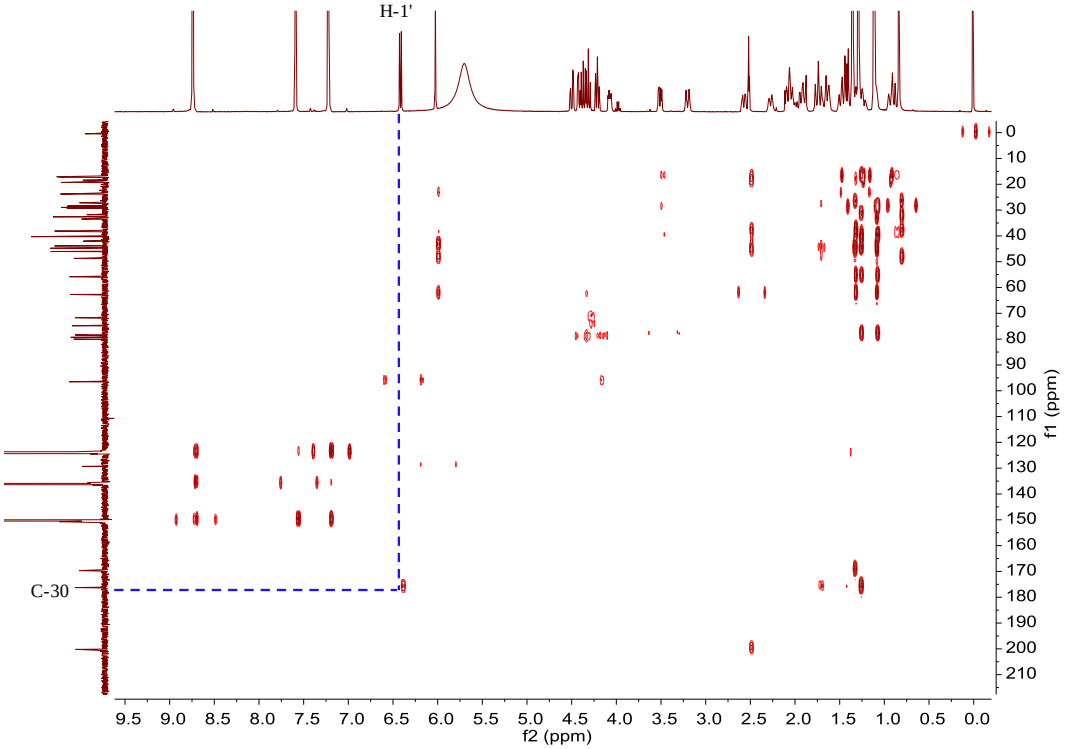


**Figure S26.** HMBC spectrum of **3a** in pyridine-$d_5$ (400 MHz).

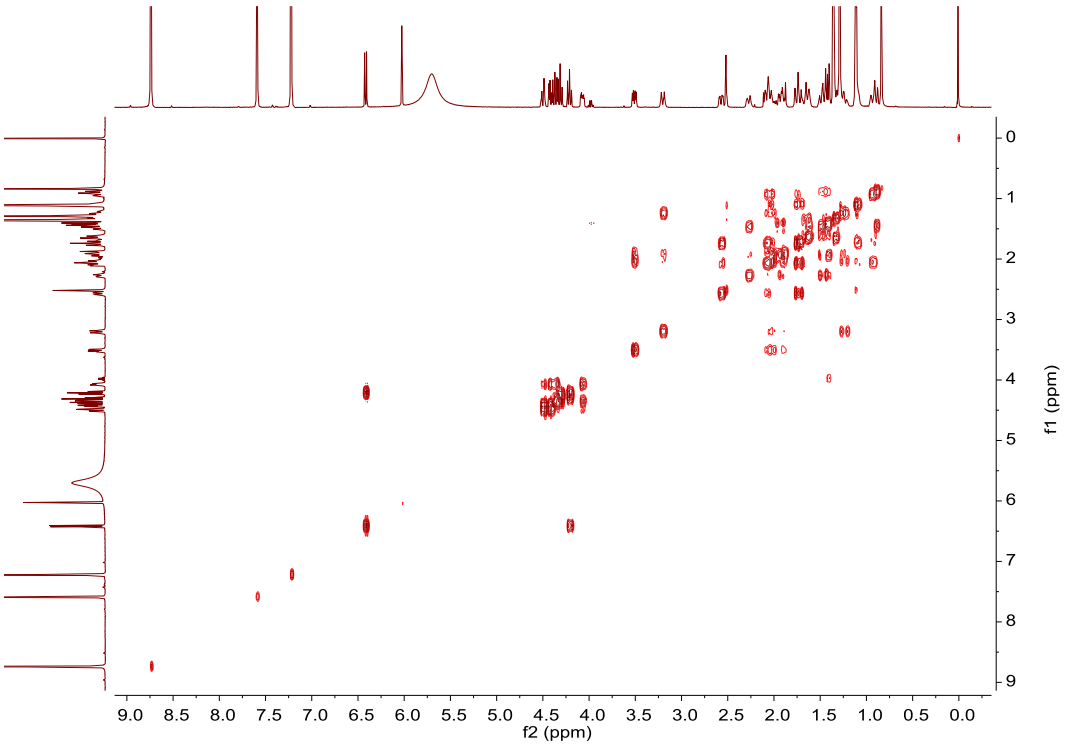


**Figure S27.** $^{1}$H-$^{1}$H COSY spectrum of **3a** in pyridine-$d_5$ (400 MHz).

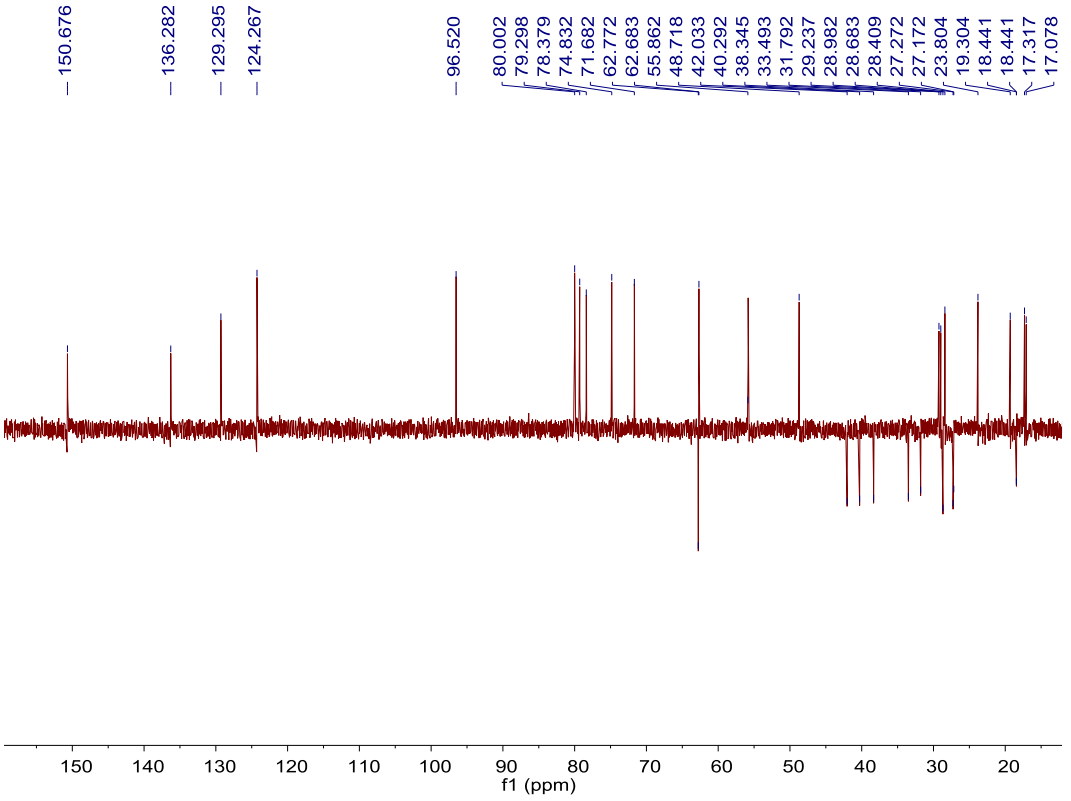


**Figure S28.** DEPT 135 spectrum of **3a** in pyridine-$d_5$ (100 MHz).

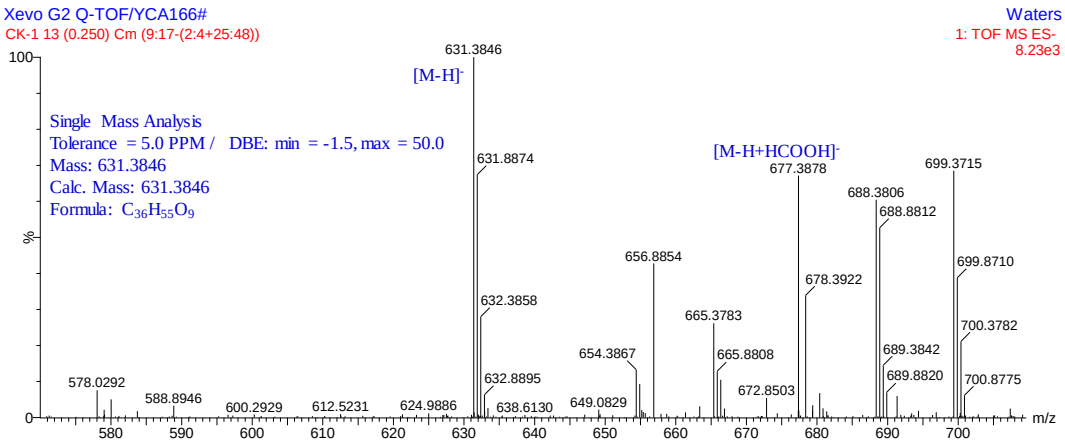


**Figure S29.** HR-ESI-MS spectrum of **3a**.

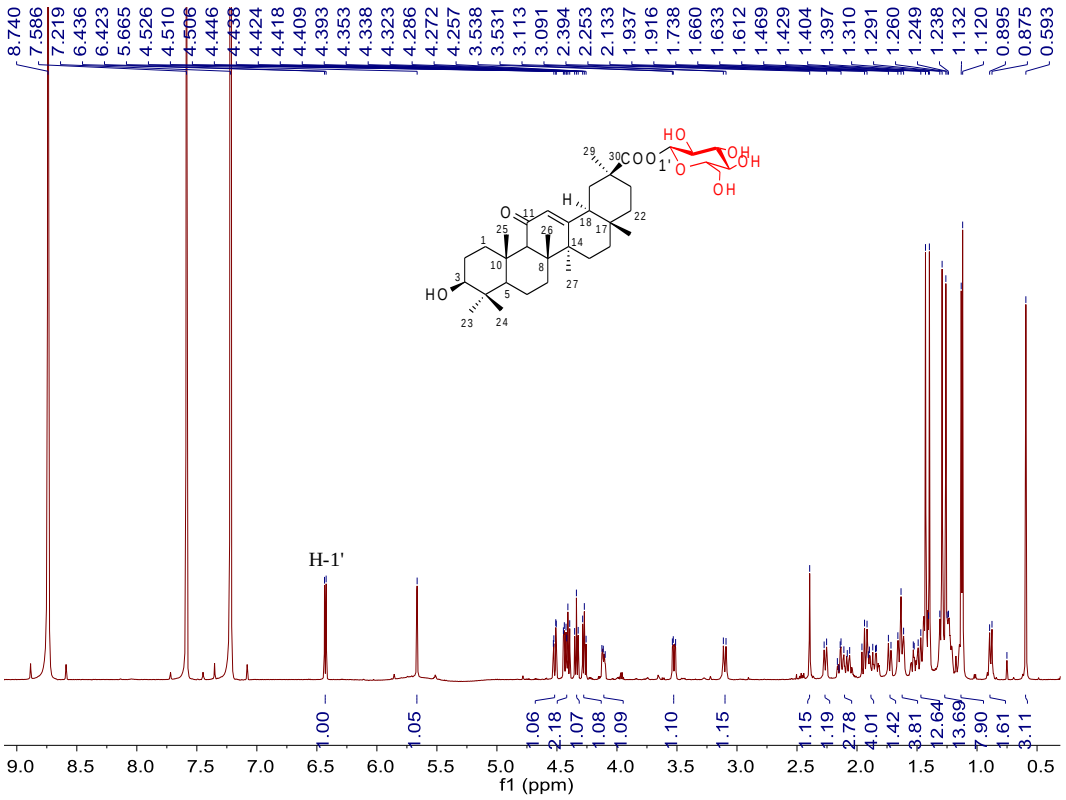


**Figure S30.** $^{1}$H NMR spectrum of **4a** in pyridine-$d_5$ (600 MHz).

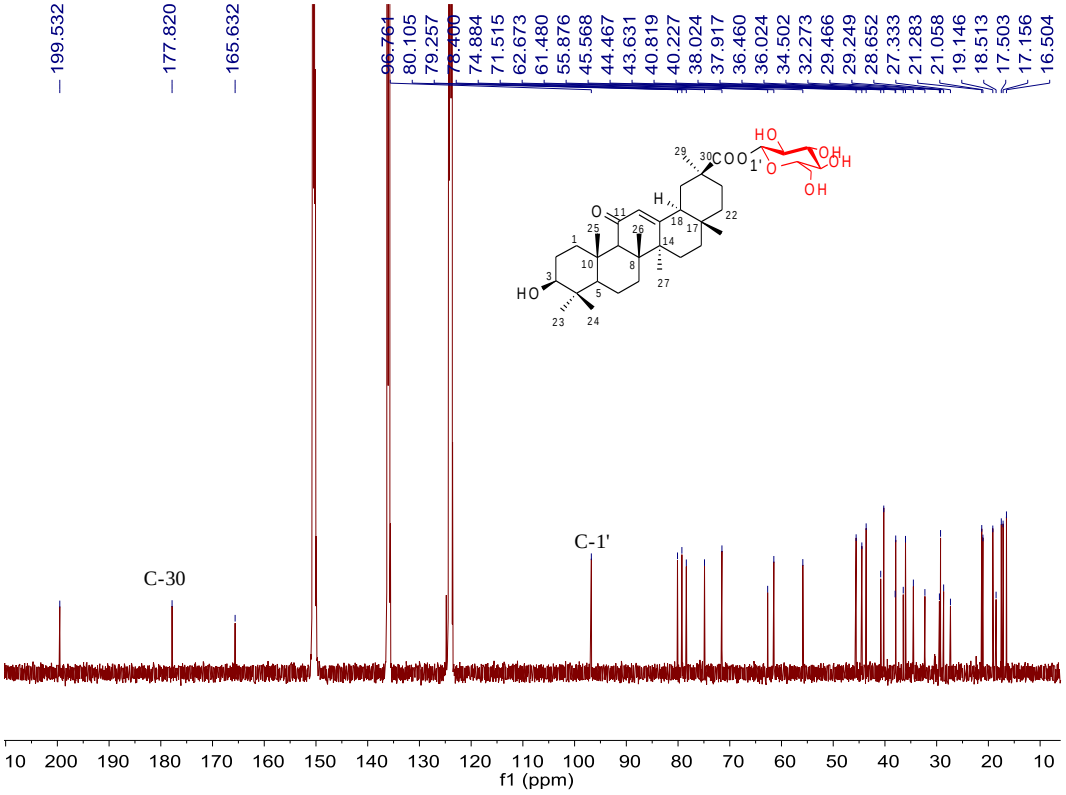


**Figure S31.** $^{13}$C NMR spectrum of **4a** in pyridine-$d_5$ (150 MHz).

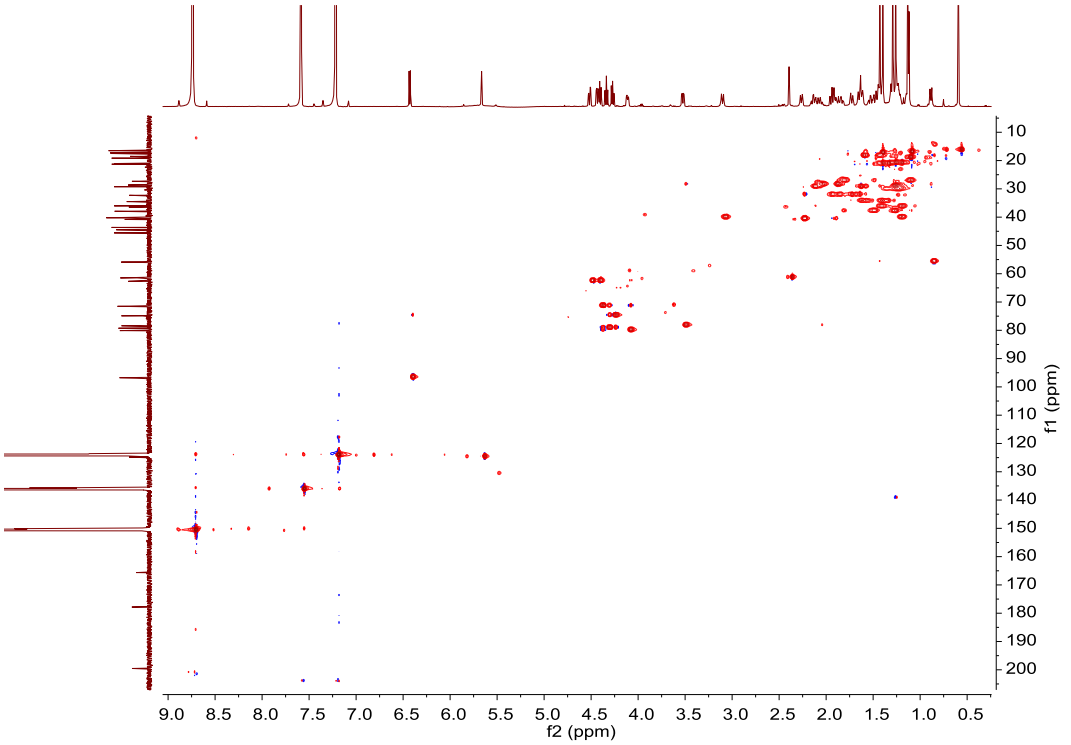


**Figure S32.** HSQC spectrum of **4a** in pyridine-$d_5$ (600 MHz).

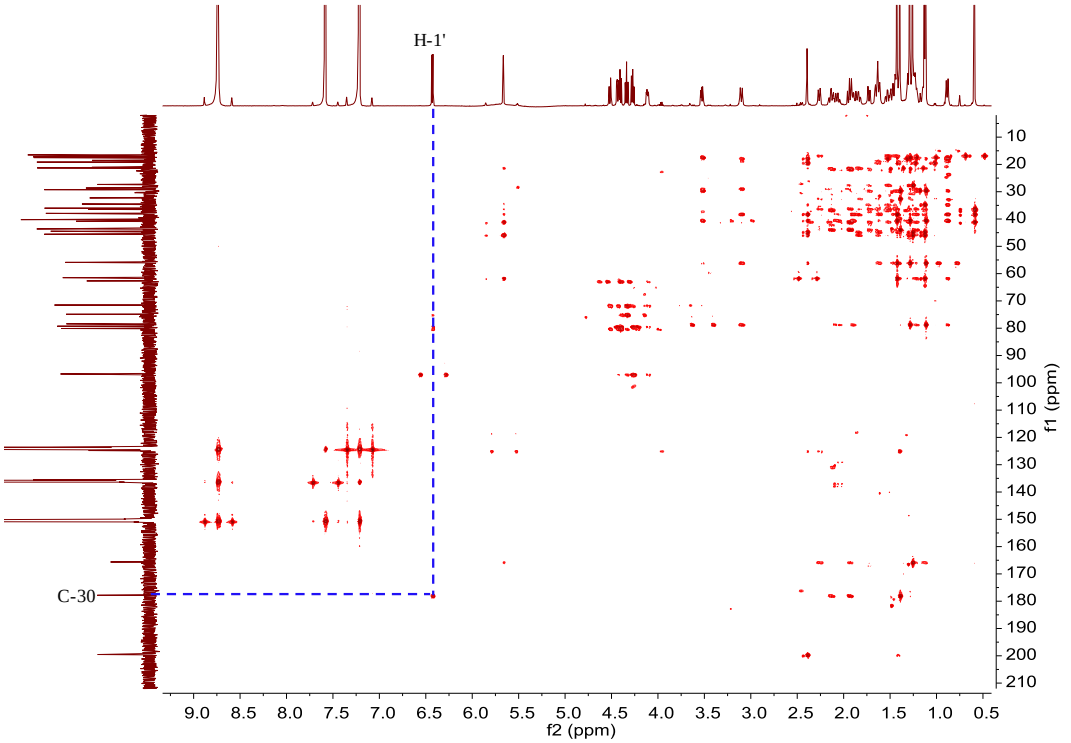


**Figure S33.** HMBC spectrum of **4a** in pyridine-$d_5$ (600 MHz).

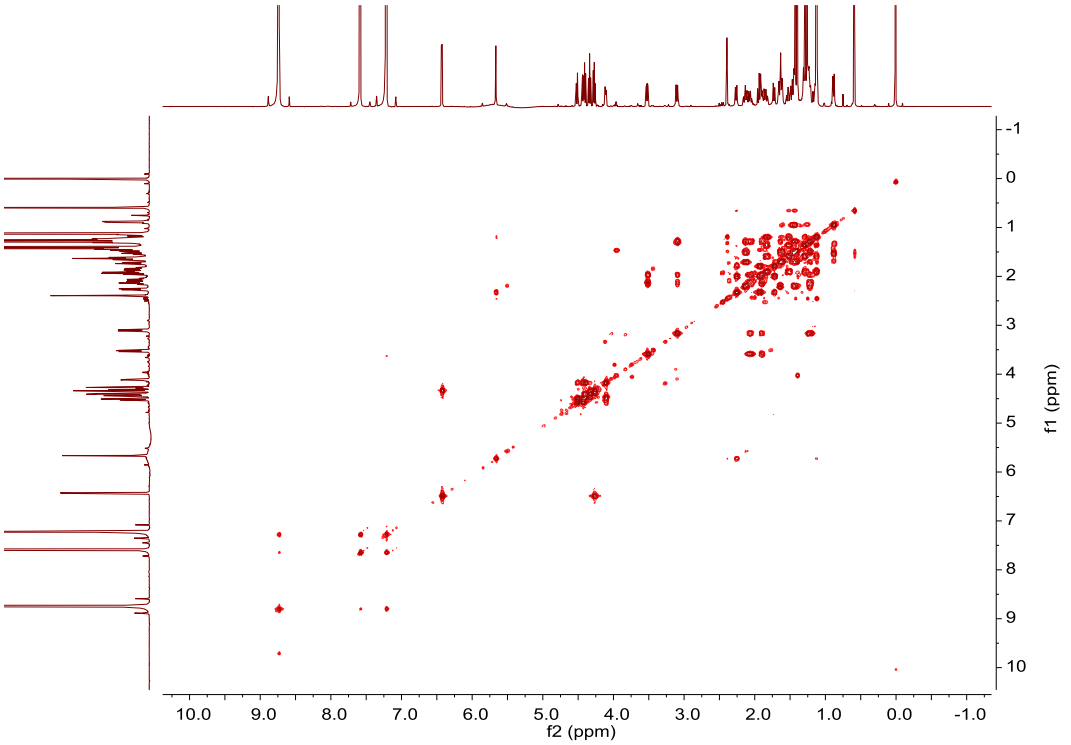


**Figure S34.** $^{1}$H-$^{1}$H COSY spectrum of **4a** in pyridine-$d_5$ (600 MHz).

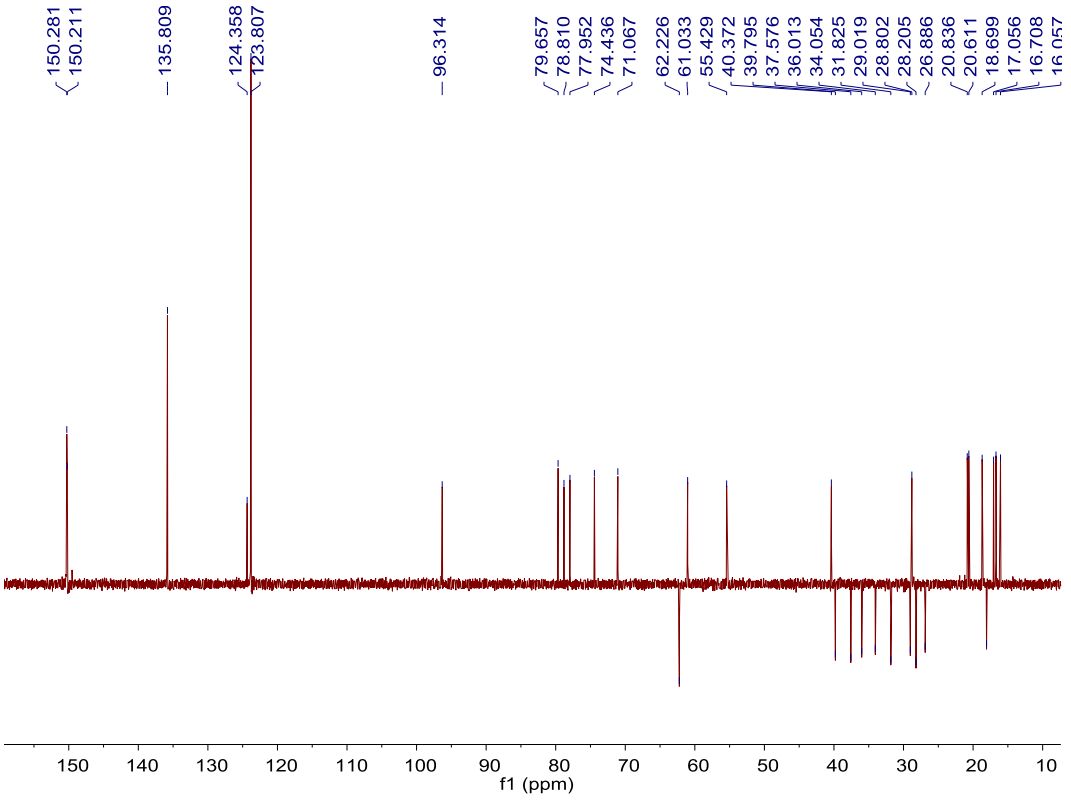


**Figure S35.** DEPT 135 spectrum of **4a** in pyridine-$d_5$ (150 MHz)

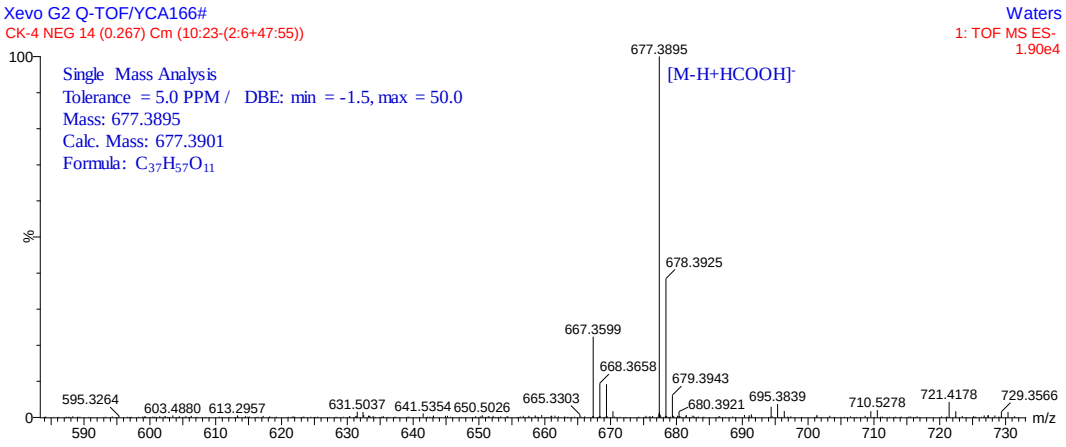


**Figure S36.** HR-ESI-MS spectrum of **4a**.

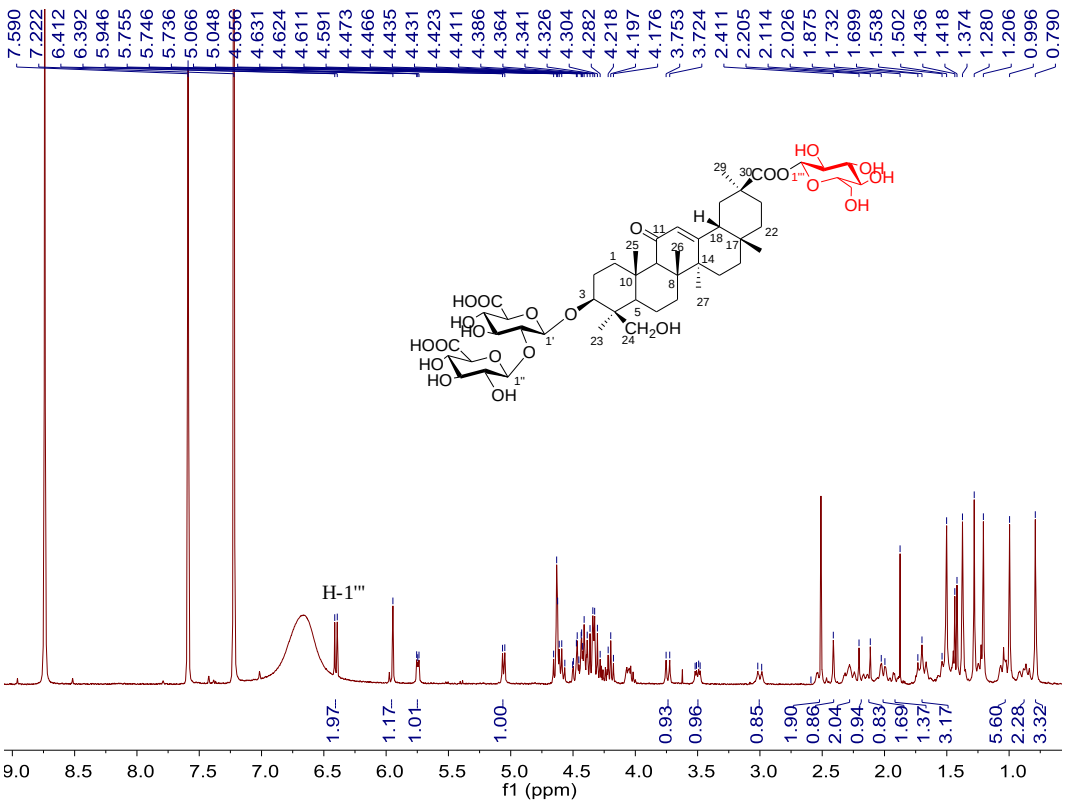


**Figure S37.** $^{1}$H NMR spectrum of **5a** in pyridine-$d_5$ (400 MHz).

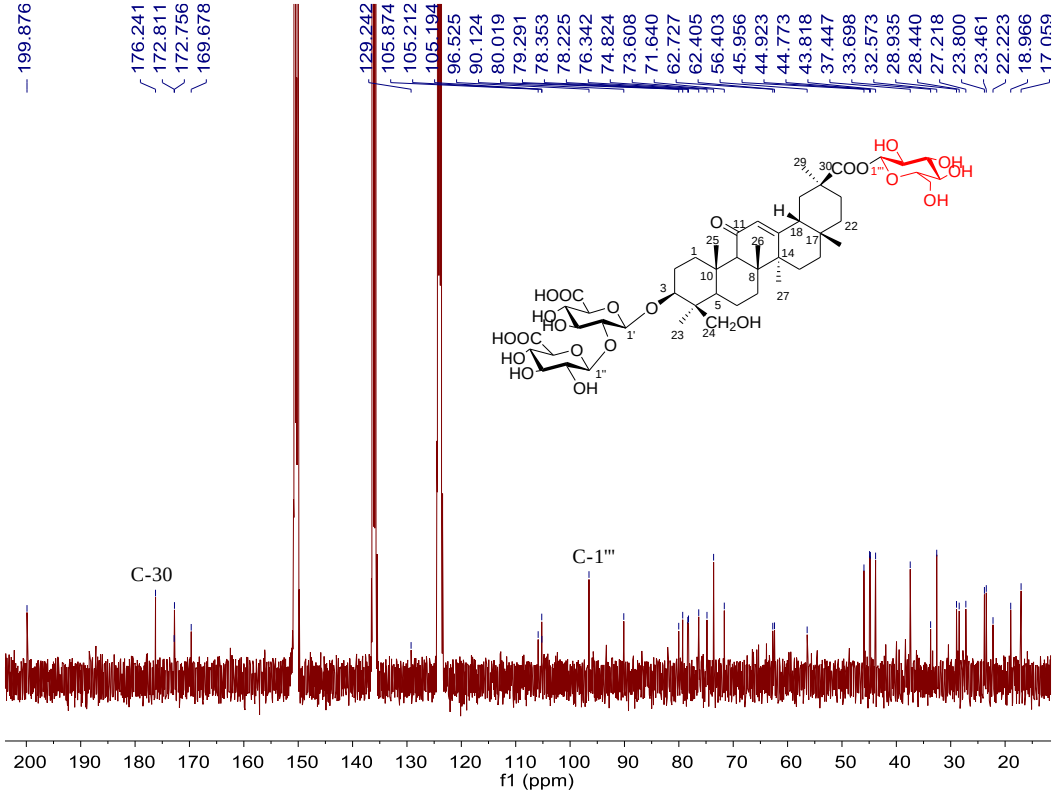


**Figure S38.** $^{13}$C NMR spectrum of **5a** in pyridine-$d_5$ (100 MHz).

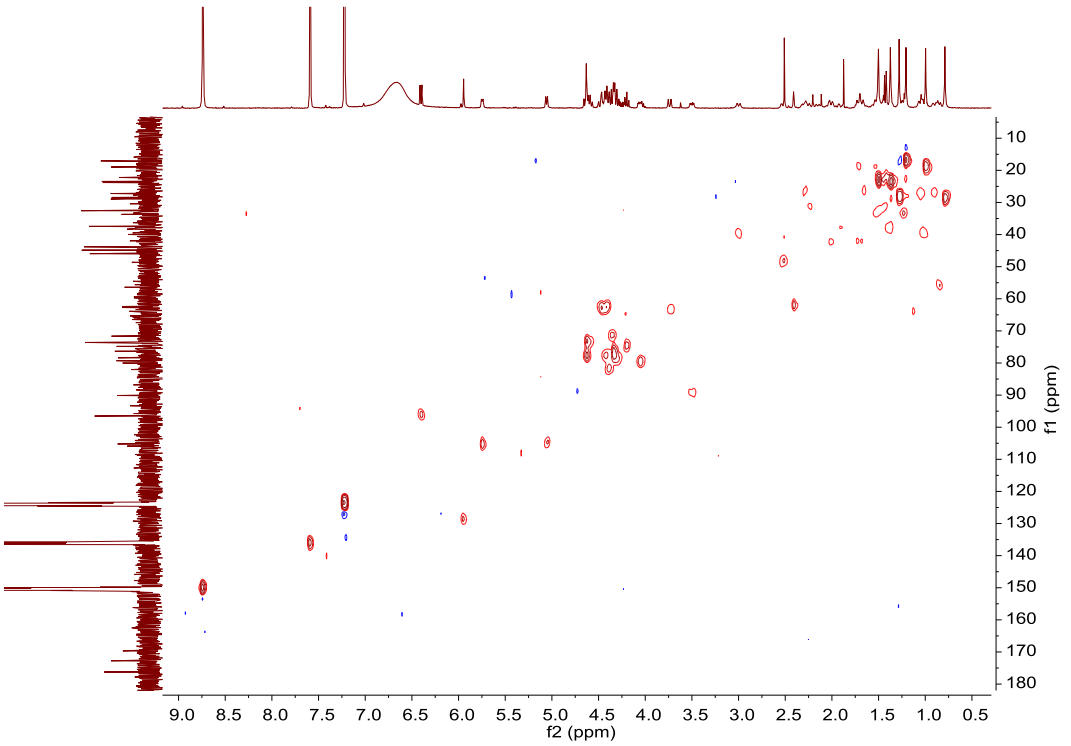


**Figure S39.** HSQC spectrum of **5a** in pyridine-$d_5$ (400 MHz).

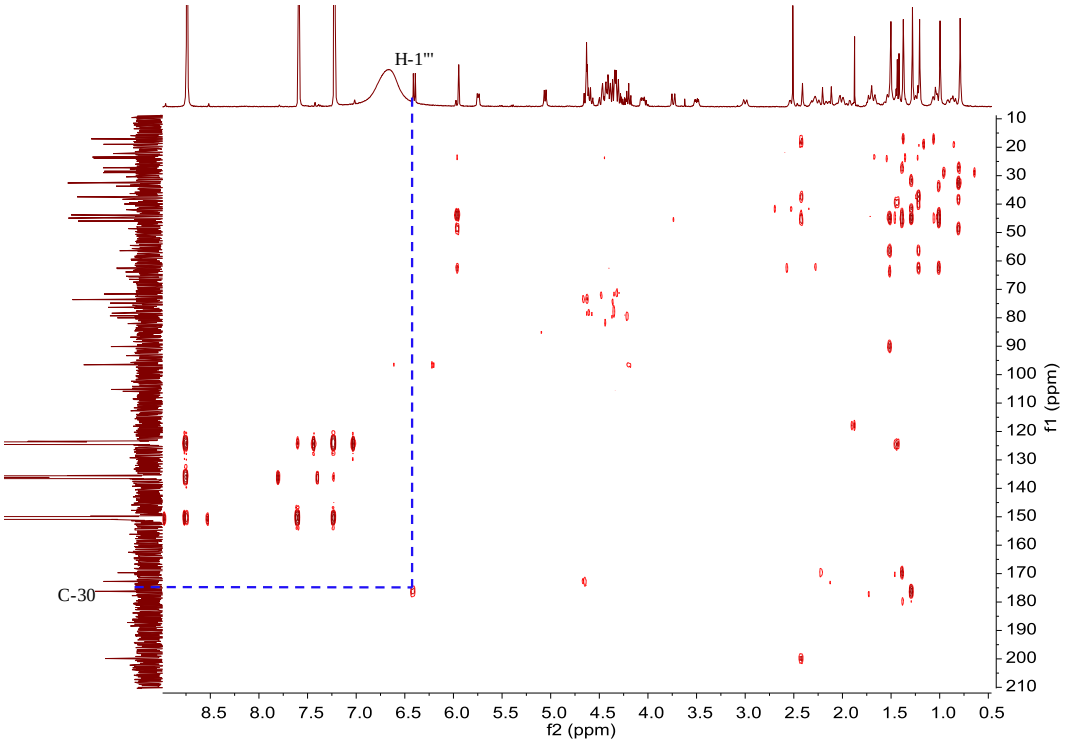


**Figure S40.** HMBC spectrum of **5a** in pyridine-$d_5$ (400 MHz).

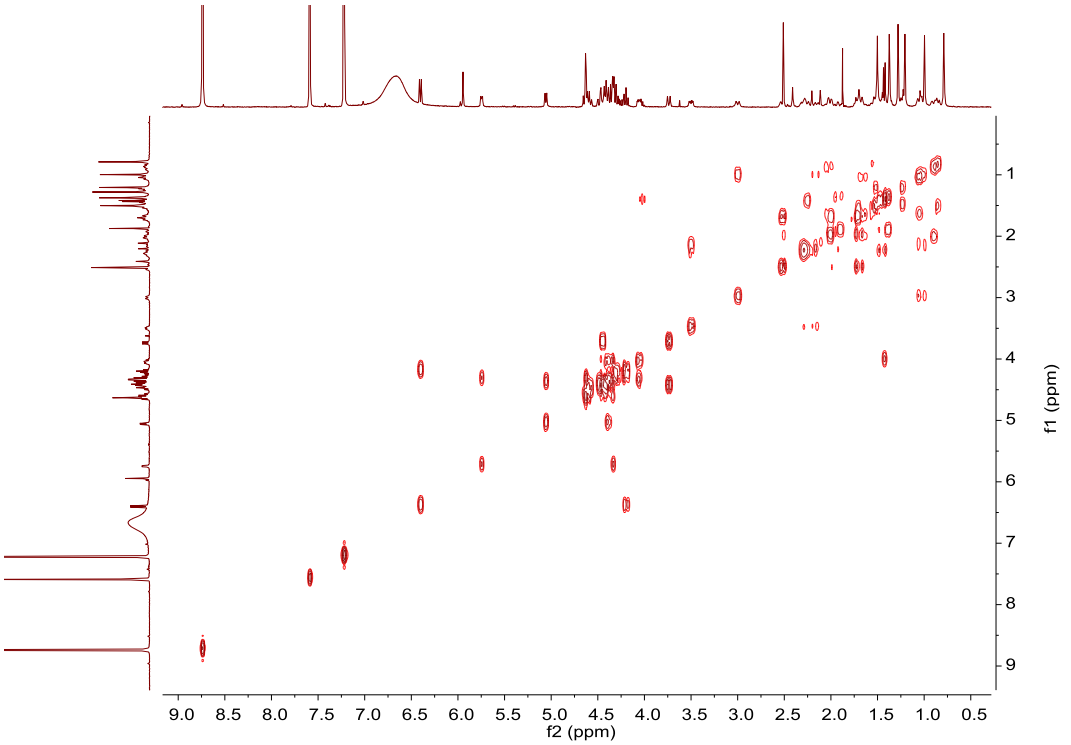


**Figure S41.** $^{1}$H-$^{1}$H COSY spectrum of **5a** in pyridine-$d_5$ (400 MHz).

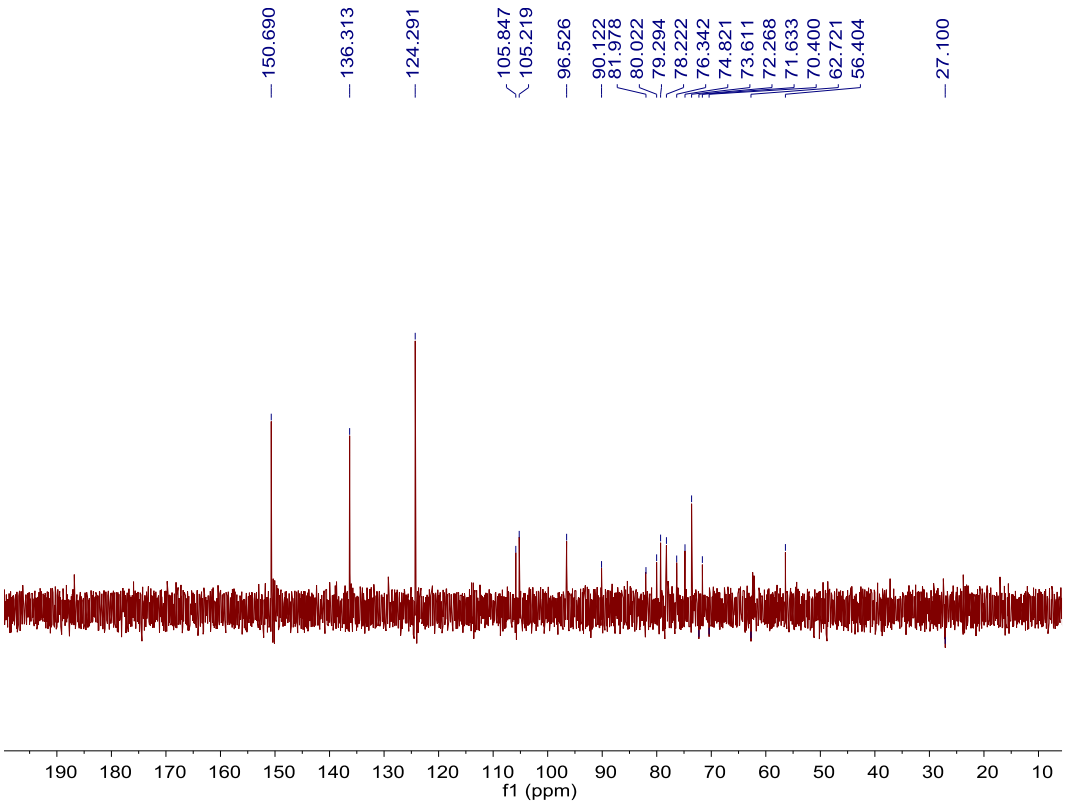


**Figure S42.** DEPT 135 spectrum of **5a** in pyridine-$d_5$ (100 MHz).

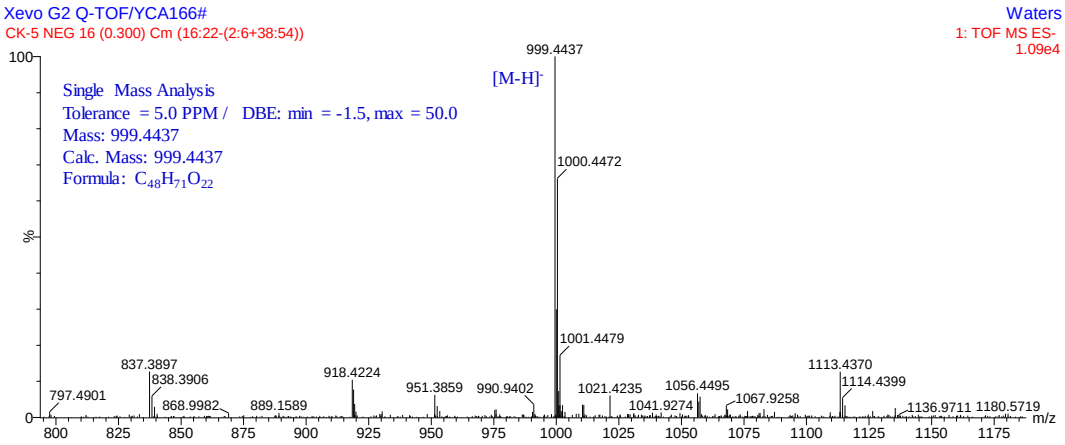


**Figure S43.** HR-ESI-MS spectrum of **5a**.

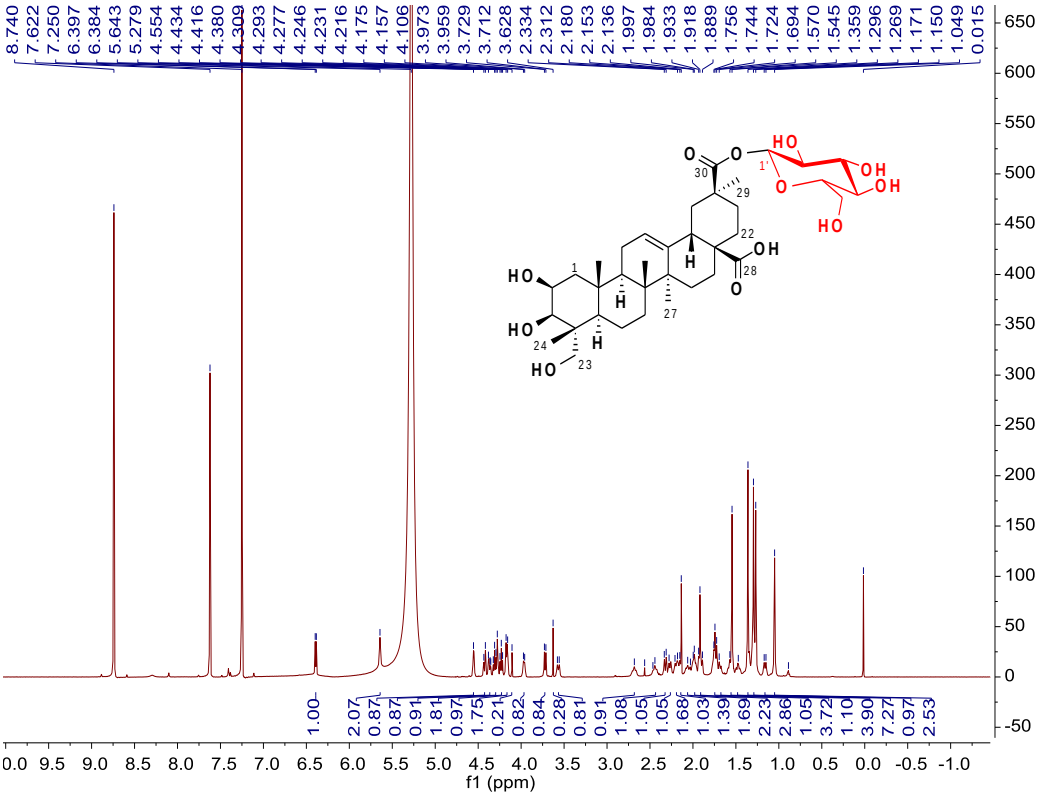

**Figure S44.** $^{1}$H NMR spectrum of **14a** in pyridine-$d_5$ (600 MHz).

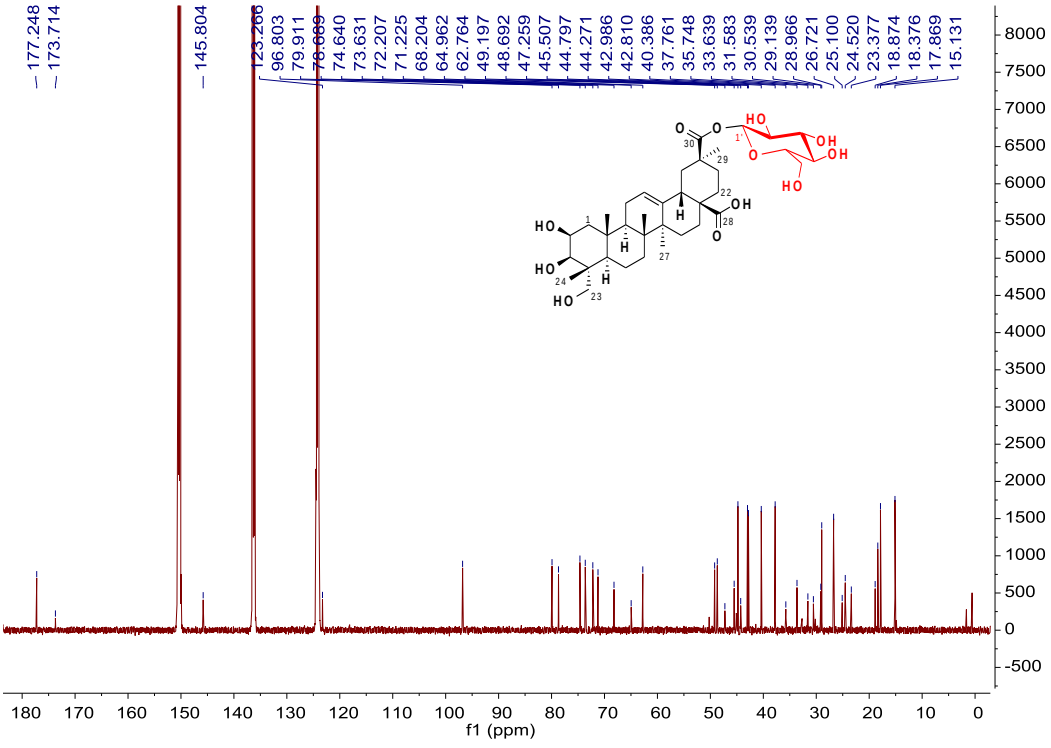

**Figure S45.** $^{13}$C NMR spectrum of **14a** in pyridine-$d_5$ (150 MHz).

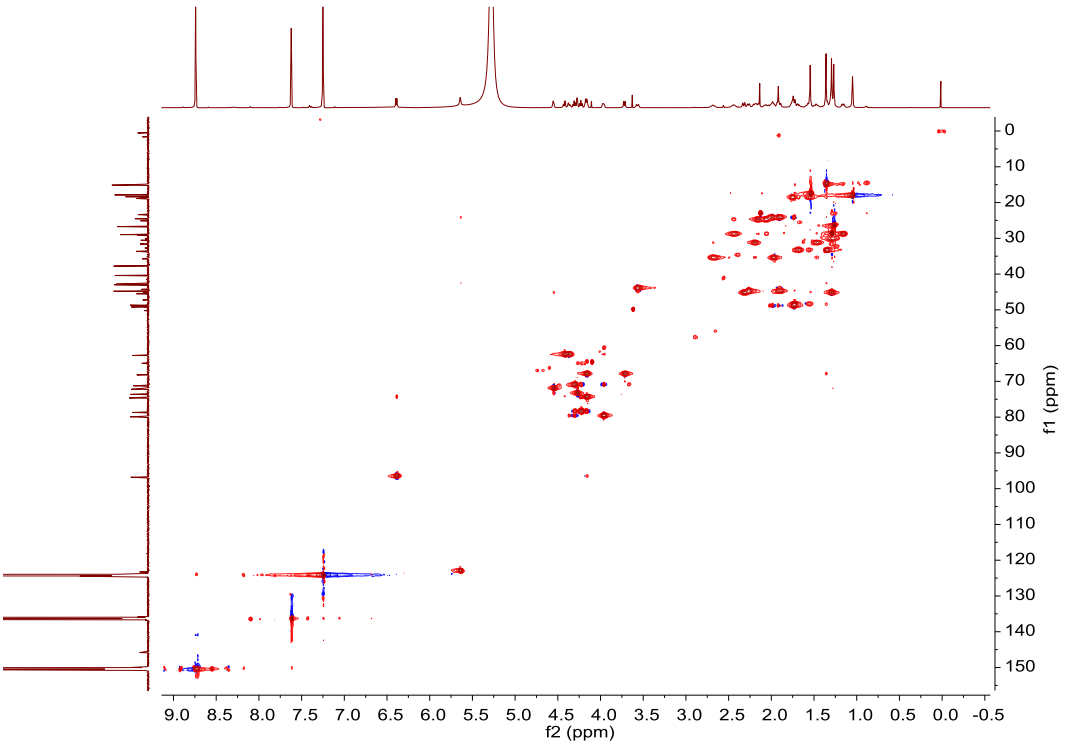


**Figure S46.** HSQC spectrum of **14a** in pyridine-$d_5$ (600 MHz).

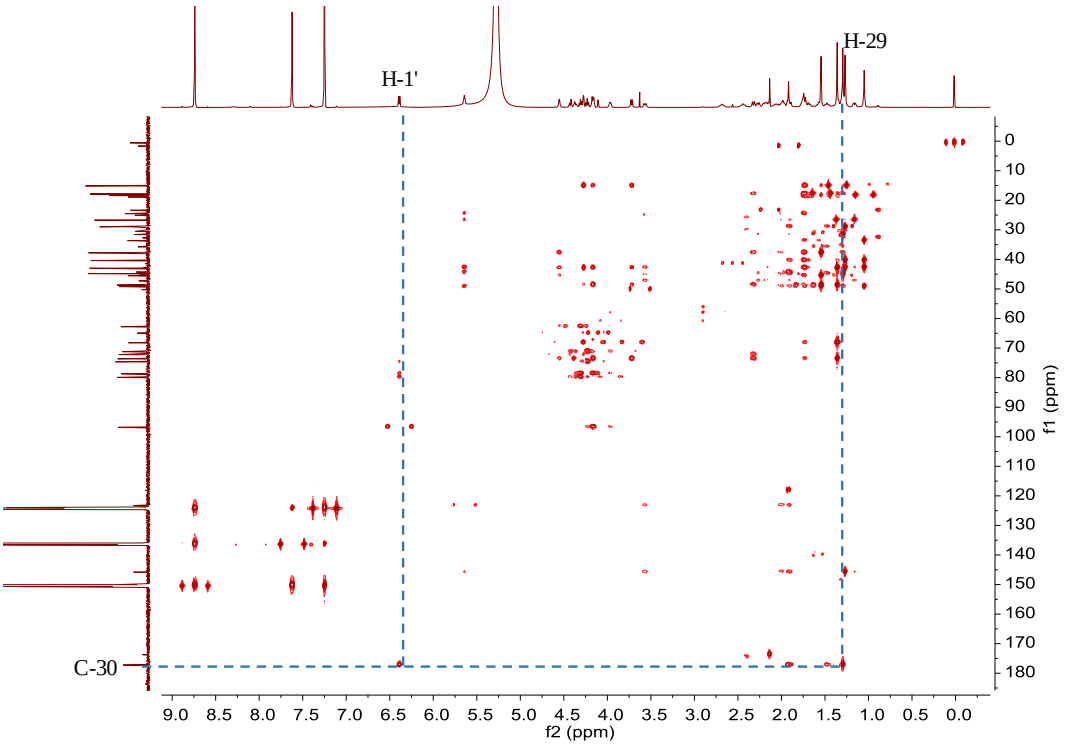


**Figure S47.** HMBC spectrum of **14a** in pyridine-$d_5$ (600 MHz).

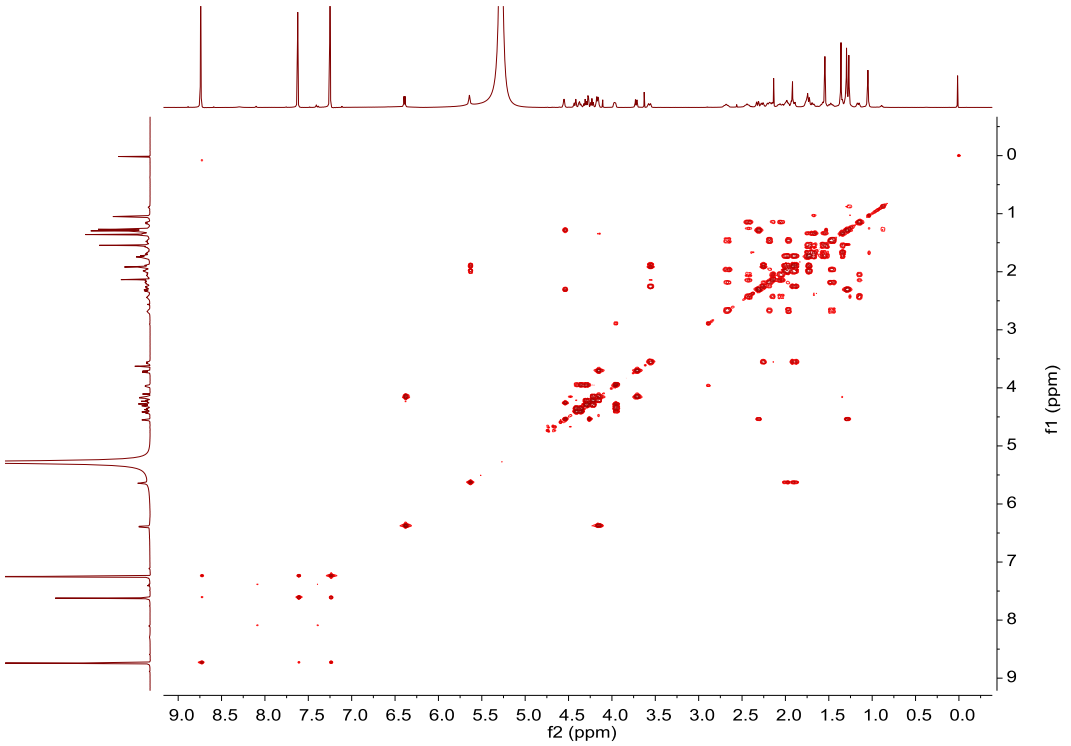


**Figure S48.** $^{1}$H-$^{1}$H COSY spectrum of **14a** in pyridine-$d_5$ (600 MHz).

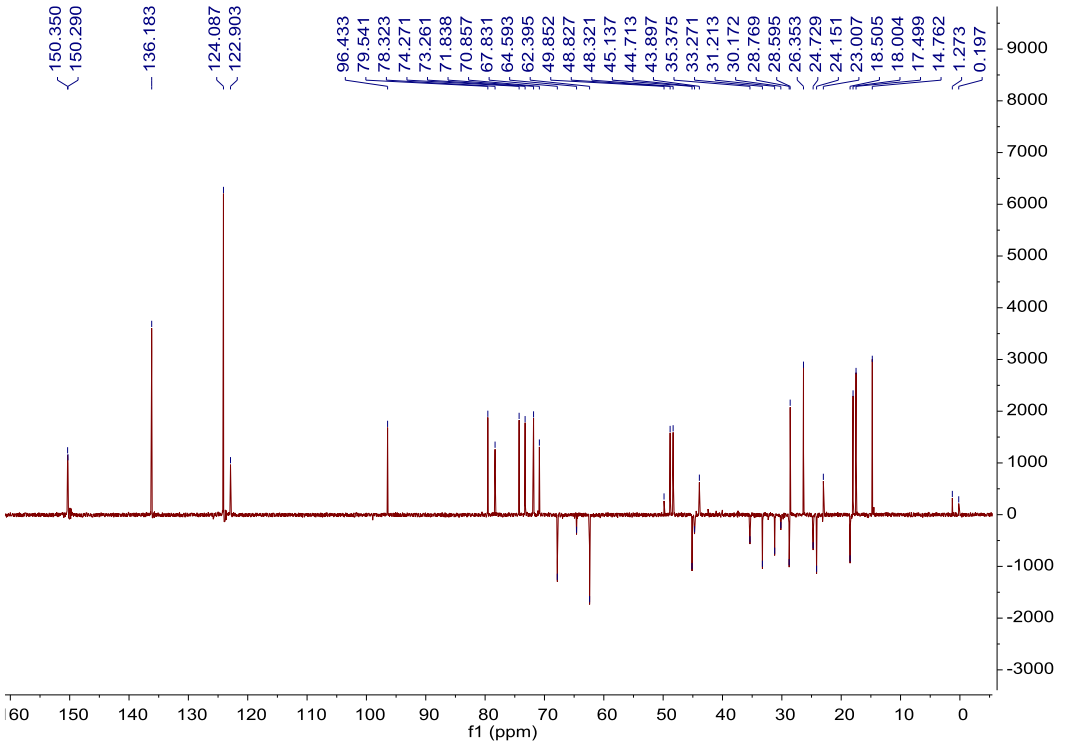


**Figure S49.** DEPT 135 spectrum of **14a** in pyridine-$d_5$ (150 MHz)

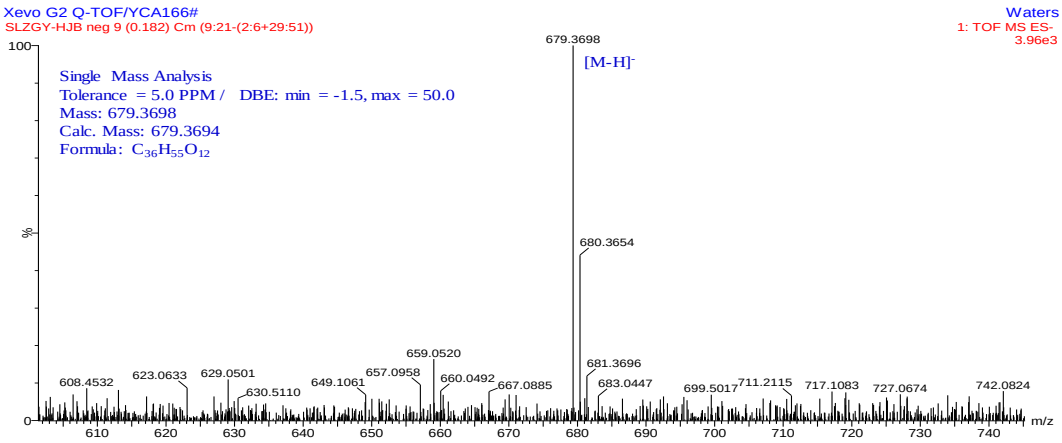


**Figure S50.** HR-ESI-MS spectrum of **14a**.

**References**

(1) (a) Song, W.; Si, L.; Ji, S.; Wang, H.; Fang, X.-m.; Yu, L.-y.; Li, R.-y.; Liang, L.-n.; Zhou, D.; Ye, M. *J. Nat. Prod.* **2014,** *77*, 1632-1643. (b) Song, W.; Qiao, X.; Chen, K.; Wang, Y.; Ji, S.; Feng, J.; Li, K.; Lin, Y.; Ye, M. *Anal. Chem.* **2017,** *89*, 3146-3153. (c) Huang, Y.; Lin, X.; Qiao, X.; Ji, S.; Liu, K.; Yeh, C.-t.; Tzeng, Y.-m.; Guo, D.; Ye, M. *J. Nat. Prod.* **2014,** *77*, 118-124. (d) Ye, M.; Han, J.; Tu, G. Z.; An, D. G.; Guo, D. *J. Nat. Prod.* **2005,** *68*, 626-628.

(2) Grabherr, M. G.; Haas, B. J.; Yassour, M.; Levin, J. Z.; Thompson, D. A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q. D.; Chen, Z. H.; Mauceli, E.; Hacohen, N.; Gnirke, A.; Rhind, N.; di Palma, F.; Birren, B. W.; Nusbaum, C.; Lindblad-Toh, K.; Friedman, N.; Regev, A. *Nat. Biotechnol.* **2011**, *29*, 644-652.

(3) Feng, J.; Zhang, P.; Cui, Y.; Li, K.; Qiao, X.; Zhang, Y.-T.; Li, S.-M.; Cox, R. J.; Wu, B.; Ye, M.; Yin, W.-B. *Adv. Synth. Catal.* **2017,** *359*, 995-1006.

(4) Zhang, Q.; Liu, Z.; Mi, Z.; Li, X.; Jia, P.; Zhou, J.; Yin, X.; You, X.; Yu, L.; Guo, F.; Ma, J.; Liang, C.; Cen, S. *Antiviral Res.* **2011,** *91*, 321-329.

(5) (a) Hayashi, H.; Fukui, H.; Tabata, M. *Phytochemistry* **1990,** *29*, 2149-2152. (b) Orihara, Y.; Furuya, T. *Phytochemistry* **1990,** *29*, 3123-3126.

(6) Nagashima, S.; Inagaki, R.; Kubo, A.; Hirotani, M.; Yoshikawa, T. *Planta* **2004,** *218*, 456-459.

(7) Sayama, T.; Ono, E.; Takagi, K.; Takada, Y.; Horikawa, M.; Nakamoto, Y.; Hirose, A.; Sasama, H.; Ohashi, M.; Hasegawa, H.; Terakawa, T.; Kikuchi, A.; Kato, S.; Tatsuzaki, N.; Tsukamoto, C.; Ishimoto, M. *Plant Cell* **2012,** *24*, 2123-2138.

(8) Naoumkina, M. A.; Modolo, L. V.; Huhman, D. V.; Urbanczyk-Wochniak, E.; Tang, Y.; Sumner, L. W.; Dixon, R. A. *Plant Cell* **2010,** *22*, 850-866.

(9) de Costa, F.; Barber, C. J. S.; Kim, Y. B.; Reed, D. W.; Zhang, H. X.; Fett-Neto, A. G.; Covello, P. S. *Plant Sci.* **2017,** *262*, 9-17.

(10) Xu, G. J.; Cai, W.; Gao, W.; Liu, C. S. *New Phytol.* **2016,** *212*, 123-135.

(11) Xie, K.; Chen, R.; Li, J.; Wang, R.; Chen, D.; Dou, X.; Dai, J. *Org. Lett.* **2014,** *16*, 4874-4877.

(12) Wang, X.; Li, C. F.; Zhou, C.; Li, J.; Zhang, Y. S. *Plant J.* **2017,** *90*, 535-546.