Supplementary Information

Functional characterization of a UDP-xylose-preferring *C*-glycosyltransferase from *Lemna aequinoctialis*

Sharpkate Shaker, Zhi-Min Hu, Zi-Long Wang, Guo Ye, Jin-Chen Long, Xin-Fang Zhai, Xue Qiao* and Min Ye*

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

*Authors for correspondence:

Min Ye

Tel: +86 10 82801516

Email: yemin@bjmu.edu.cn

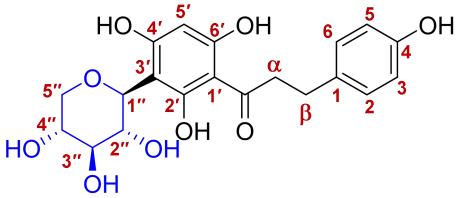
Xue Qiao

Tel: +86 10 82801516

Email: qiaoxue@bjmu.edu.cn

Table of Contents

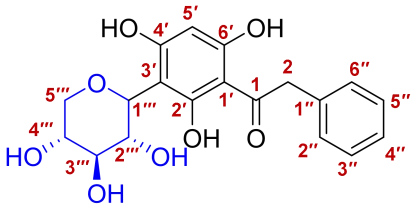
1. NMR data of compounds 1a, 5a, and 6a.



1a

Phloretin 3′-*C*-β-D-xyloside (**1a**):

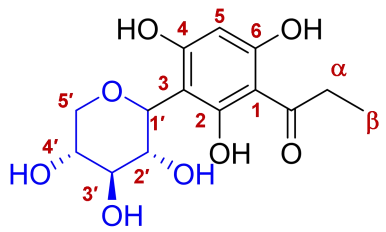
^{1}H NMR (400 MHz, methanol-d_4): Aglycone: δ = 7.06 (2H, d, J = 8.4 Hz, H-2, 6), 6.70 (2H, d, J = 8.4 Hz, H-3, 5), 5.95 (1H, s, H-5'), 3.36 (2H, m, 2H-α), 2.86 (1H, m, 2H-β). Xylosyl: δ = 4.74 (1H, d, J = 9.9 Hz, H-1"), 4.09 (1H, t, J = 9.3 Hz, H-2"), 3.97 (1H, m, H-5"a), 3.68 (1H, m, H-4"), 3.38 (1H, m, H-3") 3.28 (1H, m, H-5"b); ^{13}C NMR (100 MHz, methanol-d_4): Aglycone: δ = 206.4 (C=O), 166.3 (C-4'), 165.1 (C-2'), 162.9 (C-6'), 156.5 (C-4), 133.9 (C-9), 130.3 (C-2, 6), 116.1 (C-3, 5), 104.1 (C-3'), 103.8 (C-1'), 95.6 (C-5'), 47.4 (C-α), 31.4 (C-β). Xylosyl: δ = 80.3 (C-3"), 76.7 (C-1"), 72.7 (C-2"), 71.5 (C-4"), 64.4 (C-5").



5a

2-Phenyl-3′-(*C*-β-D-xylosyl)-2′,4′,6′-trihydroxyacetophenone (**5a**):

^{1}H NMR (400 MHz, methanol-d_4): Aglycone: δ = 7.20 (1H, m, H-4"), 7.18 (2H, m, H-2", H-4"), 7.09 (2H, m, H-3", H-5"), 5.85 (1H, s, H-5'), 4.31 (2H, m, H-2). Xylosyl: δ = 4.65 (1H, d, J = 9.9 Hz, H-1'"), 4.00 (1H, t, J = 9.2 Hz, H-2'"), 3.88 (1H, t, J = 5.5 Hz, H-5'"a), 3.55 (1H, m, H-4'"), 3.34 (1H, m, H-3'"), 3.19 (1H, m, H-5'"b); ^{13}C NMR (100 MHz, methanol-d_4): Aglycone: δ = 204.7 (C=O), 166.5 (C-4'), 165.4 (C-2'), 163.8 (C-6'), 137.4 (C-1"),130.7 (C-2", 6"), 129.1 (C-3", 5"), 127.4 (C-4"), 105.3 (C-3'), 104.2 (C-1'), 95.7 (C-5'). Xylosyl: δ = 80.3 (C-3'"), 76.7 (C-1'"), 72.7 (C-2'"), 71.5 (C-5'"), 71.5 (C-4'"). [1]



6a

Flopropione 3′-*C*-β-D-xyloside (**6a**):

^{1}H NMR (400 MHz, methanol-d_4): Aglycone: δ = 3.08 (2H, dd, J = 7.3 Hz, H-α), 1.13 (3H, t, J = 7.3 Hz, H-β), 5.92 (1H, s, H-5). Xylosyl: δ = 4.72 (1H, d, J = 9.9 Hz, H-1'), 4.05 (1H, t, J = 9.1 Hz, H-2'), 3.96 (1H, m, H-5'a), 3.63 (1H, m, H-4'), 3.37 (1H, m, H-3'), 3.25 (1H, m, H-5'b).[2]

2. Supplementary Tables

Table S1. PCR primers and nucleotide sequence used in this study.

PCR primers

Primers	Sequences (5' to 3')
LaCGT1-F	acagcccagatctgggtaccATGGCGTCAGTGATGTCGGCGGGA
LaCGT1-R	agtggtggtggtggtggtgctcgagGAGACGTCTAATCTTGAACTTTTCC
R271A-F	GCGACGGCGCTACCGCCGGAGCAGATAAGG
R271A-R	CGGTAGCGCCGTCGCGCTACCGAAGCTCAC
W357A-F	GCGAACTCCGTTACGGAGGCGGCGGTAAGC
W357A-R	CGTAACGGAGTTCGCGCCGCAATGGCACAG
D378A-F	GCGCAGAGGATCAACGCCATGGTGGTAGAG
D378A-R	GTTGATCCTCTGCGCACCGTGCCGCGGCCA
Q379A-F	GCGAGGATCAACGCCATGGTGGTAGAGAAG
Q379A-R	GGCGTTGATCCTCGCGTCACCGTGCCGCGG

Note: The vector sequences are shown in lower case, and the mutation sites are labeled in red.

Nucleotide sequence of LaCGT1

ATGGCGTCAGTGATGTCGGCGGGAGAGGCTCCTCACTTCGCGTTGTTGCCGAGCGCCGGCATGGGGCATCTCACCCCTTTCCTCCGCCTCGCCGCC
ATGCTAGCCTCTGGCGGCGCCACCATCACATTTATCTGCCCCACGCCGGTCGTTTCCGCCGCCGAGGCTGTCCATGTCGATGACTTGGTGTCGTTTT
CCCCGTCGTTTCGTCGGCTCGATTTTCCTCTCCCGCCGTTTGACGTTTCCACCGCCGCATCCAAGGATCCTTTCTTTCTCCAGTTCGAGAGCATCCG
GCGCGCAGCCTCCCTGCTGGGCCCCATCCTCTCGTCTGTGTCTCCTCCCCCGTCTGCCCTCATTCTCGACGTCACTCTCACCTCCGCCTTCGTCCCC
ATCGCCCGCGCCATCTCCCTTCCTGTATACATTATGTTCACCTCTTCTGCATGGATGCTCTCCCTCTTTCTCTCCTTTCCTAGAACTGCCAGCCGCGGC
GGCGCCATTGAGGTCCCTGGCCTGGCGGAGACTCTGCCCGCGAGCGCCCTCCCCCAGGCGCTGCGAGACCCCAATAATCTCTTCACCATTCAGTTC
TTCGAGAACGGGAAGGCCATGGCGGACGCCGACGGTATCATCGTCAACACATGGGAAGCACTGGAACCGACGACTCTGGCGGCGCTCAACGCCG
GCGAGGTCGTACCCGGCCTCCCGCCGGTGATCCCCGTTGGCCCGCTACATCCTTCTGAAAAATCTGGACCTTCCGCCCTACCTTGGCTCGACGCGC
AGCCGGATTCCTCGGTGATCTACGTGAGCTTCGGTAGCCGCACGGCGCTACCGCCGGAGCAGATAAGGGAGCTCGCCGCGGGACTGGAGAACAG
CGGGTGTAAATTCCTGTGGGTCCTGAAGACGAAGAAGGTCGACAAGGAAGAGCGGGAAGACGGCGAGCGAGCCATGAATGAGTTTCTTGGAGA
GGGTTTCTTTGAGAGGATGGAAGGGAGGGGCAAGGTGGTGAAGGGATGGGTGGACCAACAGGCAGTTCTGCGCCACCCAGCCGTAGGAGGATTT
CTGTGCCATTGCGGCTGGAACTCCGTTACGGAGGCGGCGGTAAGCGGGGTGAGGGTGCTGGCTTGGCCGCGGCACGGTGACCAGAGGATCAACG
CCATGGTGGTAGAGAAGAGCGGGCTGGGGAAGTGGCCAAAAGAATGGAGCTGGGACGGGGACGACGAGACAGTTCGCGGGGAAGAGATCAGTC
AGAGAATCAGGGCGCTCATGGCTGAGTCCGCTACCGCTGCCGCCAAAGTGAAAGAGCAGGCCATCGGCGCCGCATCCGACGGCGGTAGTTCGCAT
CGCCATTTGGAGCACCTTGTGGAAAAGTTCAAGATTAGACGTCTCTAA

Table S2. The transcriptome data of *Lemna aequinoctialis* from NCBI

No.	Run	Bases	Size	Experiment	Assay Type	LibrarySource	Release Date
1	SRR5688502	4.2Gbp	2.8Gb	SRX2922957	RNA-Seq	TRANSCRIPTOMIC	2017/6/15
2	SRR5688503	3.4Gbp	2.3Gb	SRX2922959	RNA-Seq	TRANSCRIPTOMIC	2017/6/15

3. Supplementary Figures

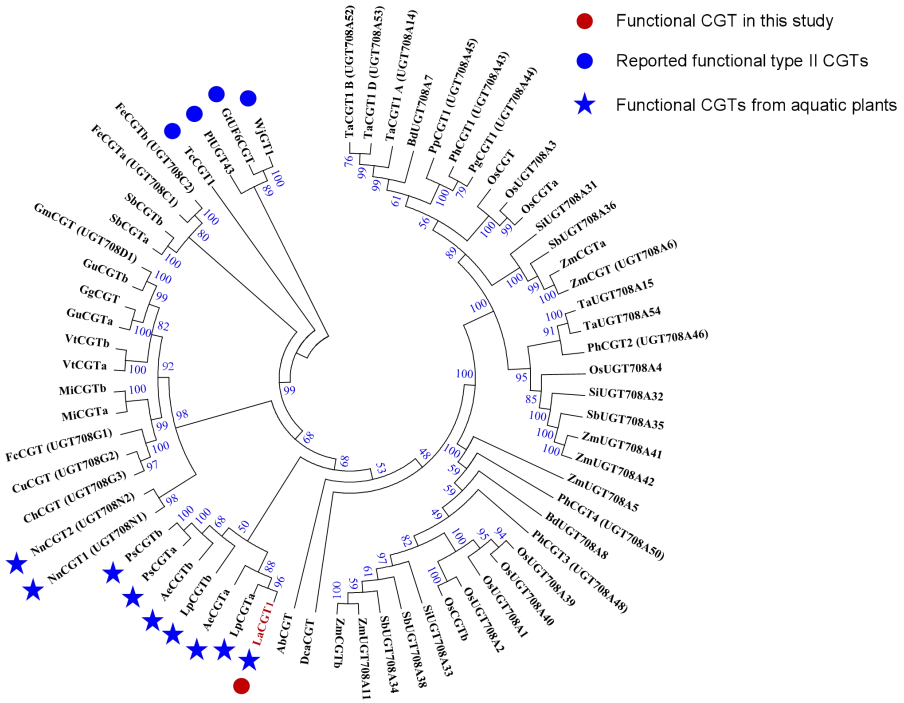
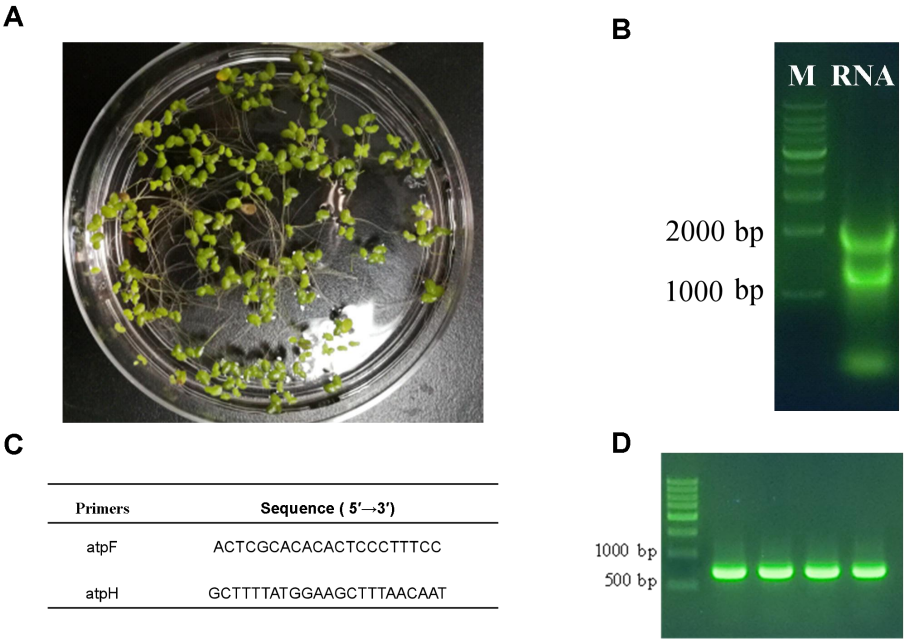


Figure S1. Phylogenetic analysis of LaCGT1 with 65 previously reported functional plant CGT genes.



Primers	Sequence (5′→3′)
atpF	ACTCGCACACACTCCCTTTCC
atpH	GCTTTTATGGAAGCTTTAACAAT

Figure S2. Fresh plant material of *Lemna aequinoctialis* (A) RNA electrophoresis of *Lemna aequinoctialis*. (B) Primers used in DNA barcoding and electrophoresis of *Lemna aequinoctialis* (C and D).

Figure S3. Sequence alignment of LaCGT1 with homologues.

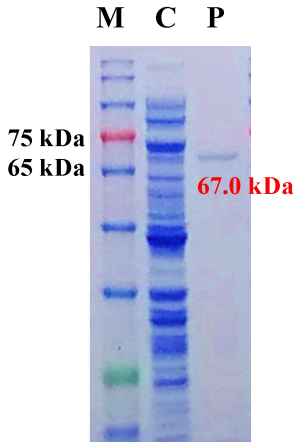


Figure S4. SDS-PAGE of His-tagged LaCGT1 purified by Ni-NTA affinity chromatography.

M, Protein marker; C, crude protein extract; P, purified LaCGT1.

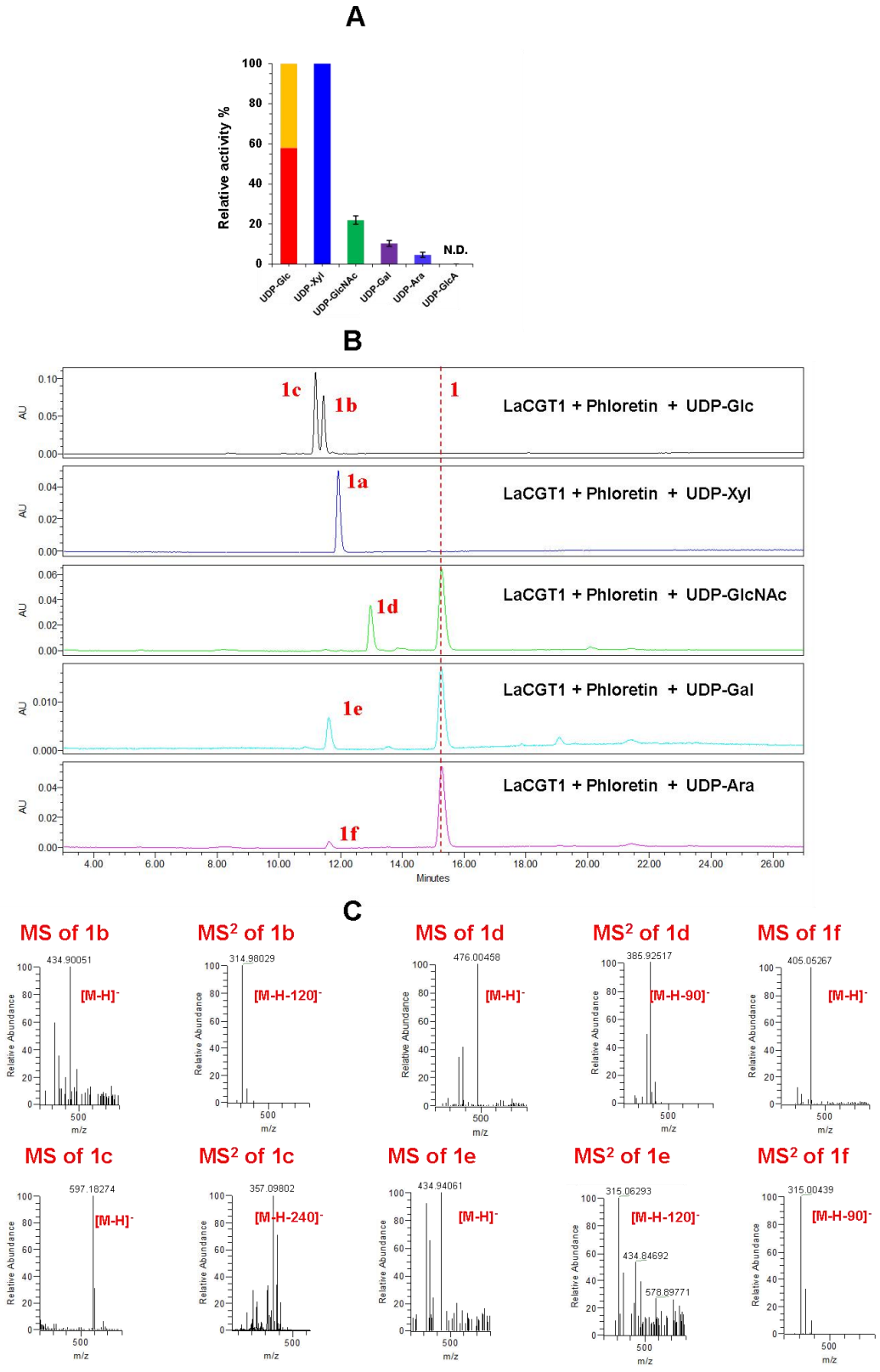


Figure S5. Sugar donor promiscuity of LaCGT1. (A) Relative conversion rates of *C*-glycosylated products using different sugar donors. LaCGT1 could glycosylate phloretin into nothofagin (red) and di-*C*-glycosylated product (yellow). (B) HPLC chromatograms of enzyme reactions of LaCGT1 with different sugar donors at high enzyme concentrations (50 μg/mL, 2 h). (C) LC/MS/MS analysis of glycosylated products in the negative ion mode.

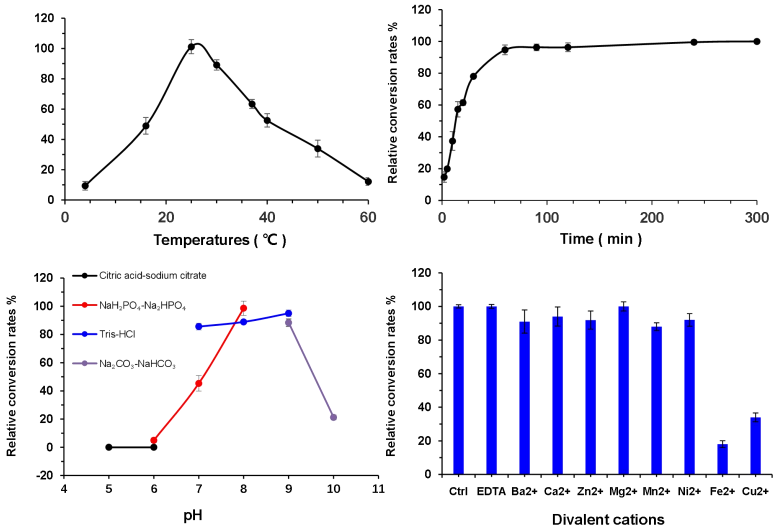


Figure S6. Effects of reaction temperature (A), time (B), reaction buffer (C), and divalent metal ions (D) on the activities of LaCGT1. Phloretin (**1**) was used as the acceptor and UDP-Xyl was used as the sugar donor. An optimized reaction time of 60 min was used. LaCGT1 exhibited its maximum activity at pH 8.0 (50 mM Na_2HPO_4-NaH_2PO_4, and 25°C.

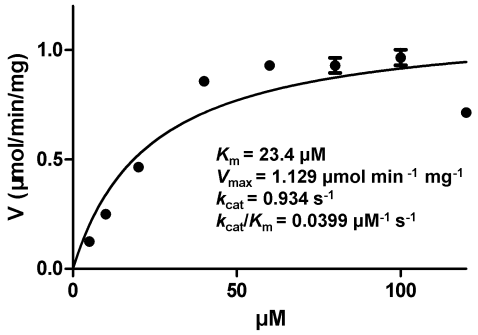


Figure S7. Determination of kinetic parameters for purified LaCGT1. The apparent K_m value was determined using phloretin (**1**) as the acceptor and UDP-Xyl as the sugar donor at 25°C and pH 8.0 for 10 min.

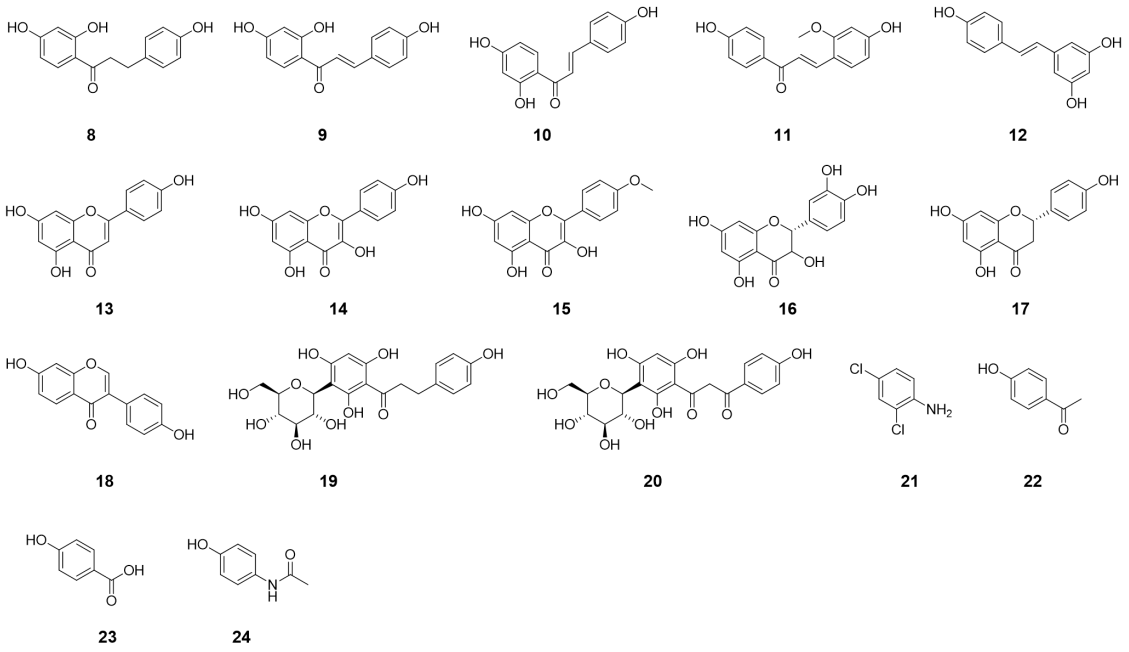


Figure S8. Substrates **8**-**24** that cannot be glycosylated by LaCGT1 using UDP-Xyl as the sugar donor.

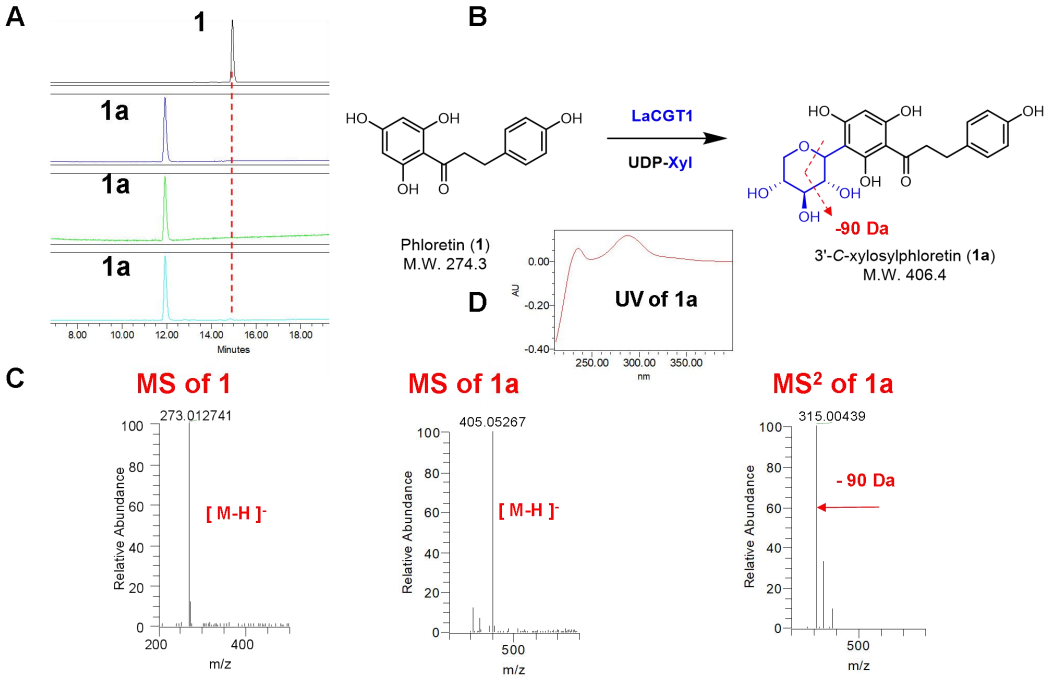


Figure S9. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 1 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **1** and enzymatic product **1a** (3′-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **1** and **1a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **1a**.

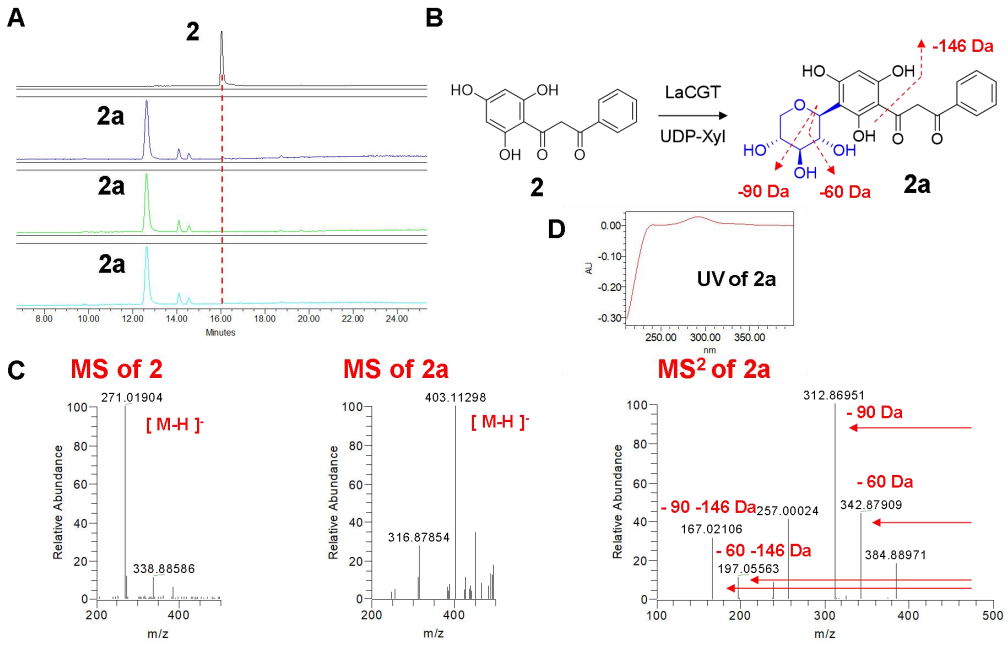


Figure S10. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 2 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **2** and enzymatic product **2a** (3′-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **2** and **2a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **2a**.

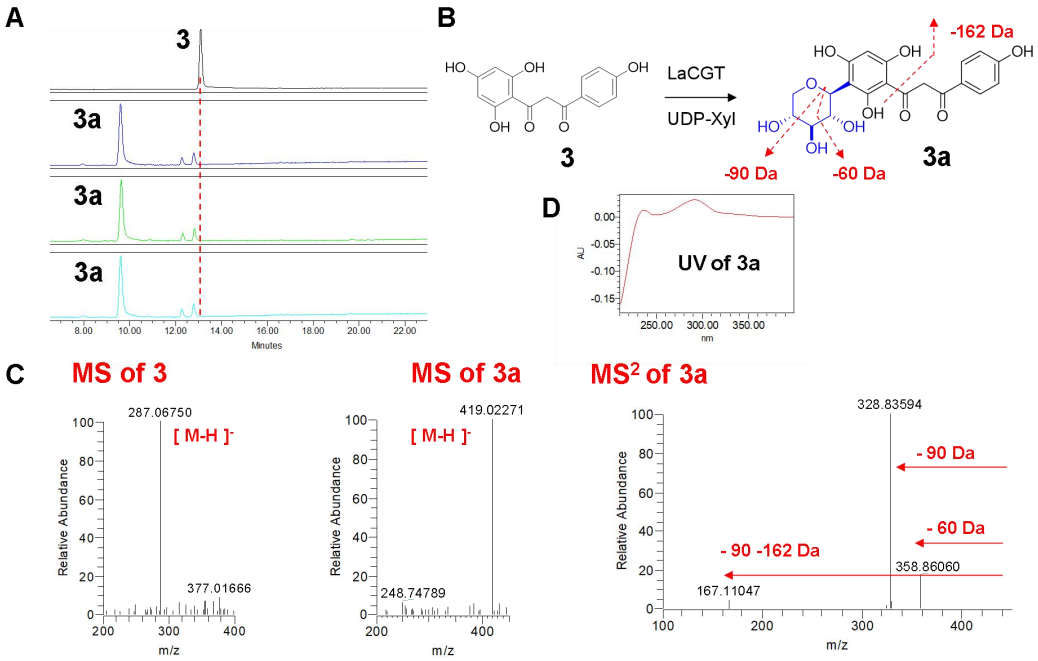


Figure S11. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 3 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **3** and enzymatic product **3a** (3′-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **3** and **3a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **3a**.

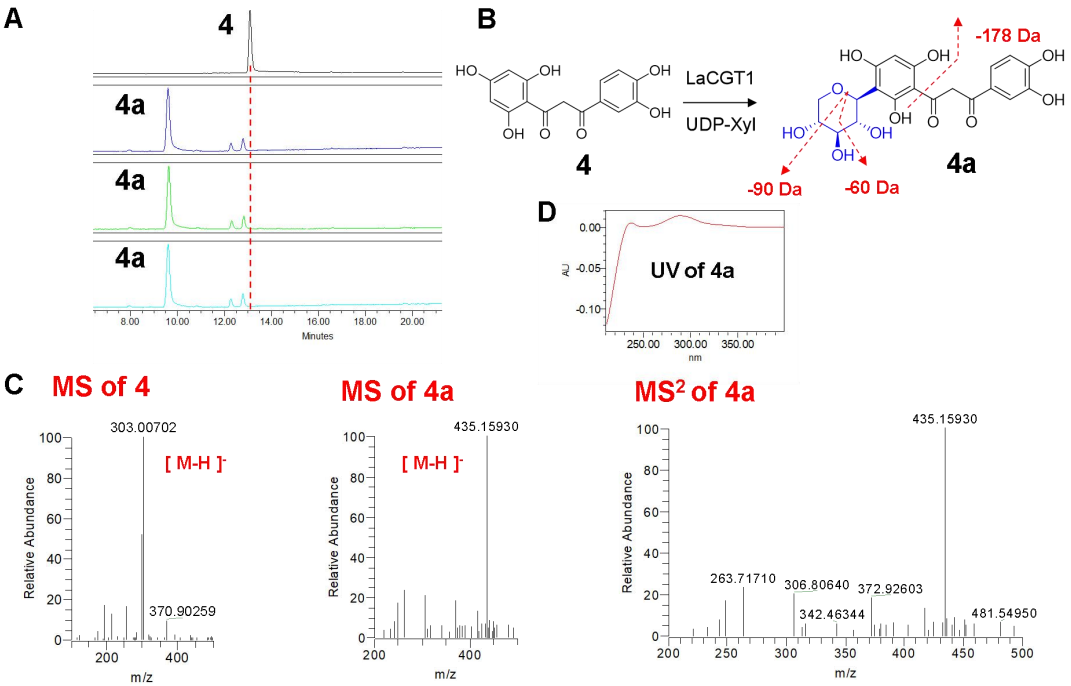


Figure S12. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 4 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **4** and enzymatic product **4a** (3′-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **4** and **4a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **4a**.

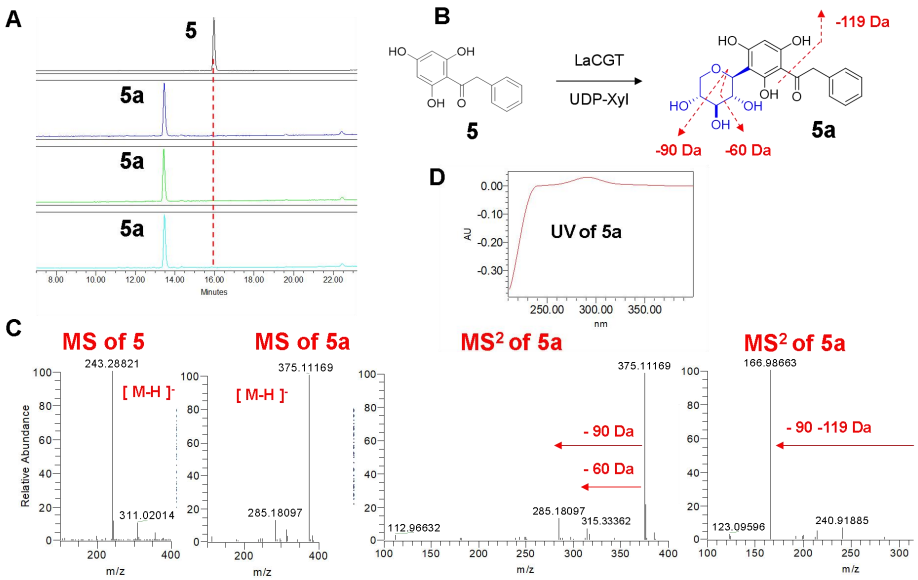


Figure S13. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 5 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **5** and enzymatic product **5a** (3-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **5** and **5a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **5a**.

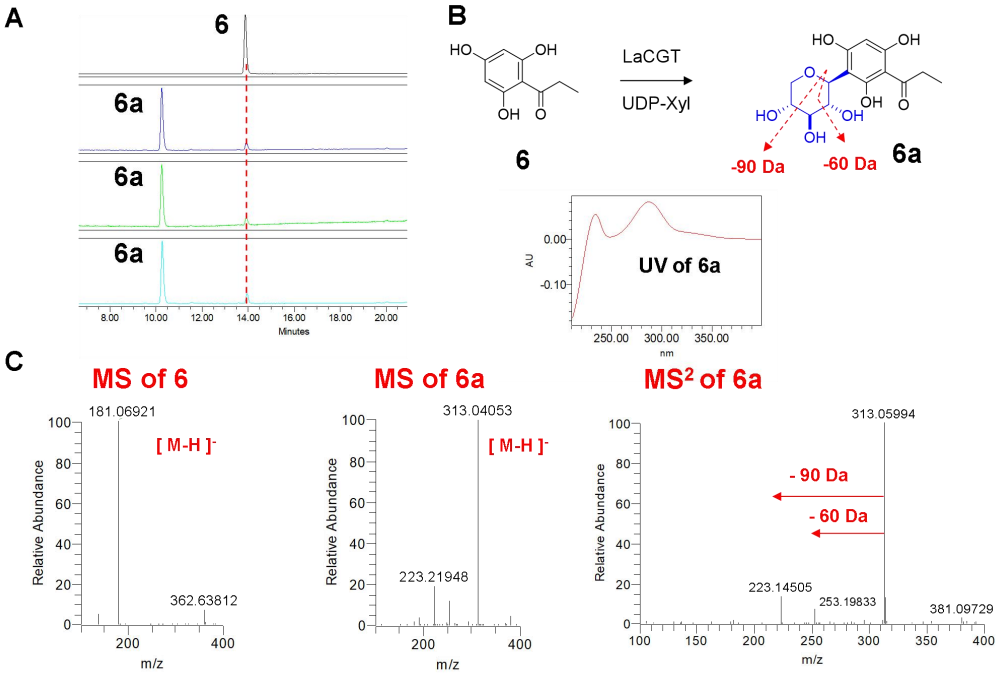


Figure S14. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 6 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **6** and enzymatic product **6a** (3-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **6** and **6a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **6a**.

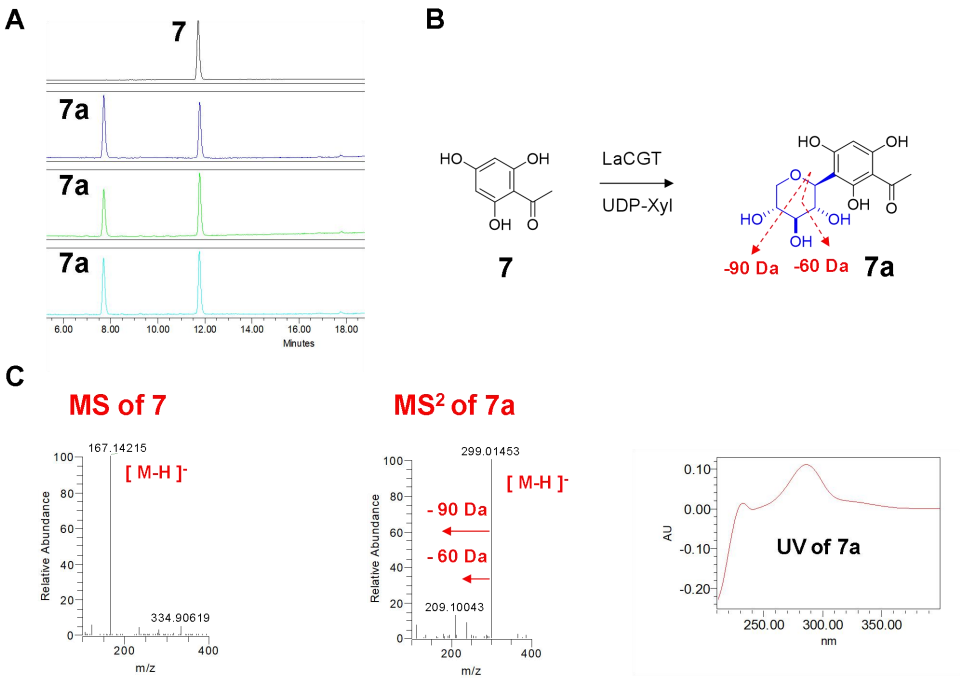


Figure S15. HPLC and LC/MS analysis of LaCGT1 catalytic reaction mixtures for compound 7 with UDP-Xyl. The experiments were performed in triplicate**.** (A) HPLC chromatogram of **7** and enzymatic product **7a** (3′-*C*-Xyl). (B) The structure of substrate and enzymatic product. (C) MS/MS spectra for **7** and **7a**. The analysis conditions are described in experimental procedures. (D) UV spectrum of **7a**.

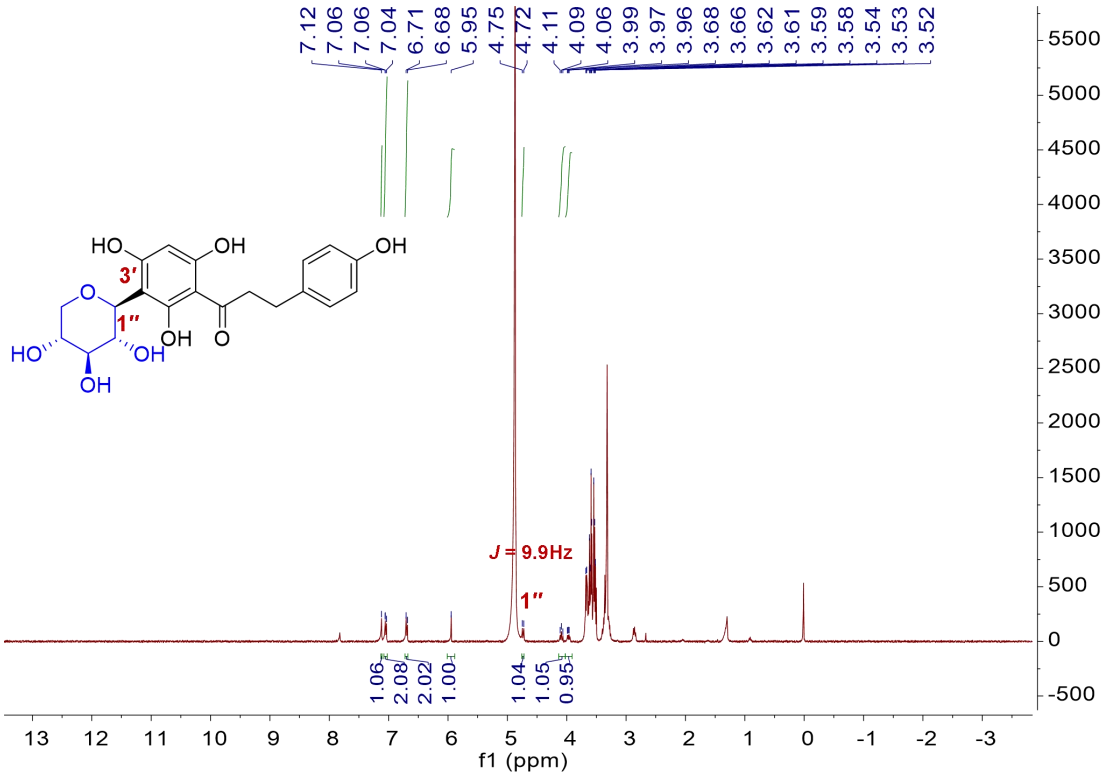


Figure S16. [1]H NMR spectrum of **1a** in Methanol-d_4, 400 MHz.

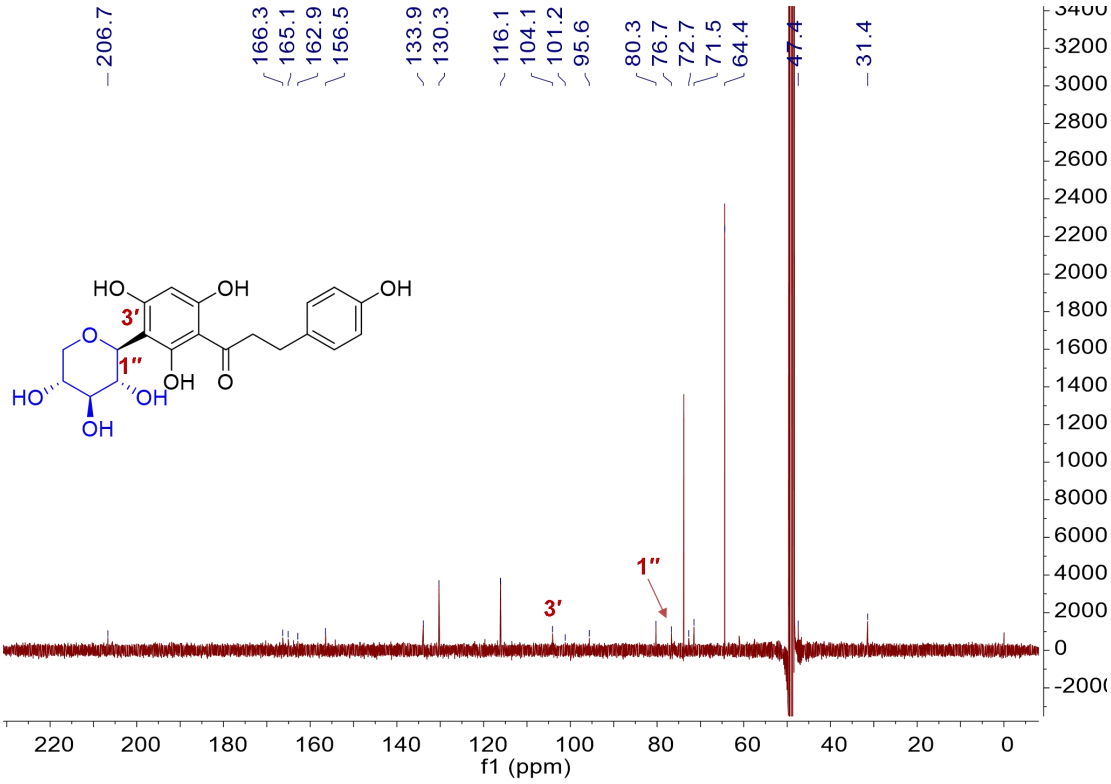


Figure S17. [13]C NMR spectrum of **1a** in Methanol-d_4, 100 MHz.

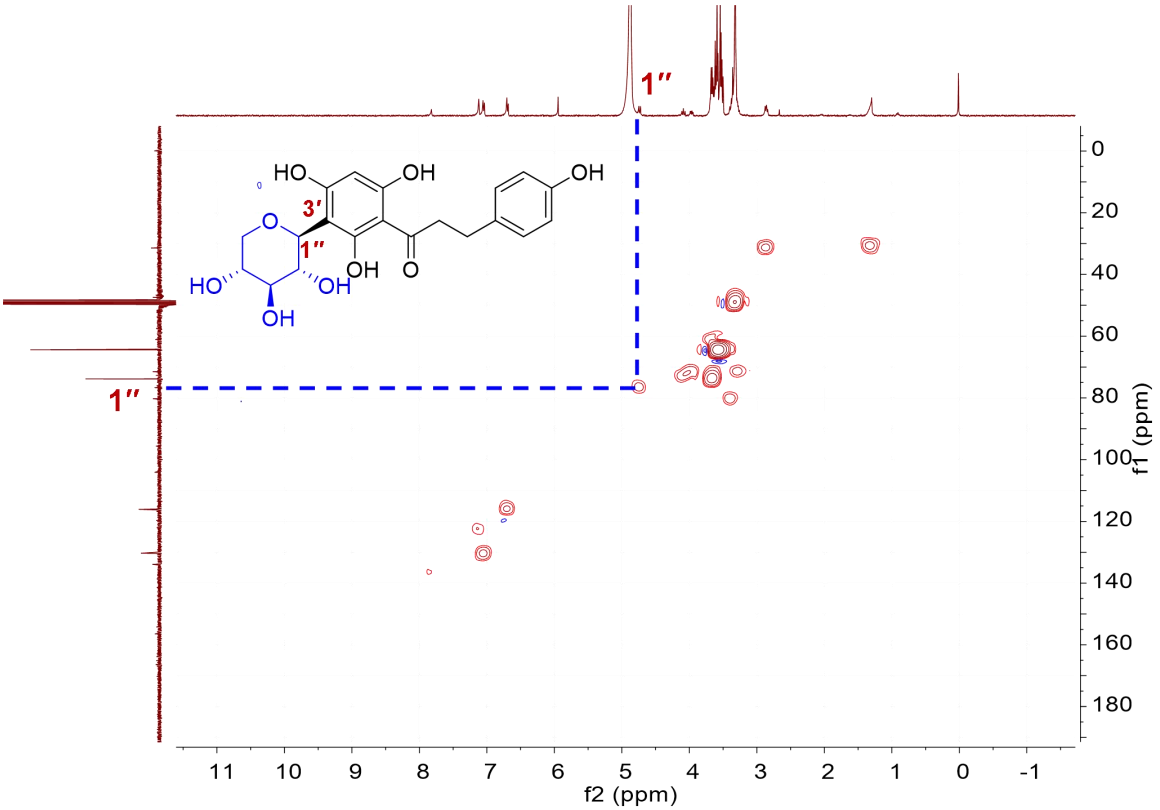


Figure S18. HSQC spectrum of **1a** in Methanol-d_4, 400 MHz.

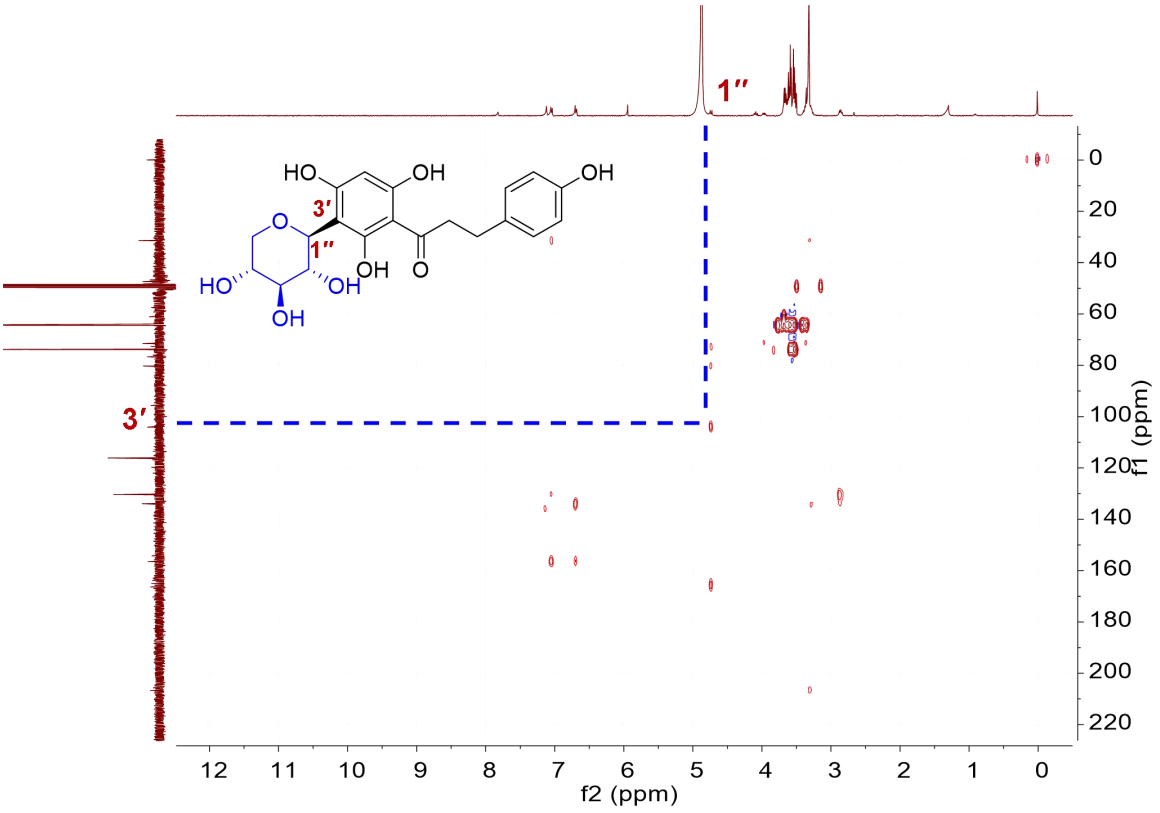


Figure S19. HMBC spectrum of **1a** in Methanol-d_4, 400 MHz.

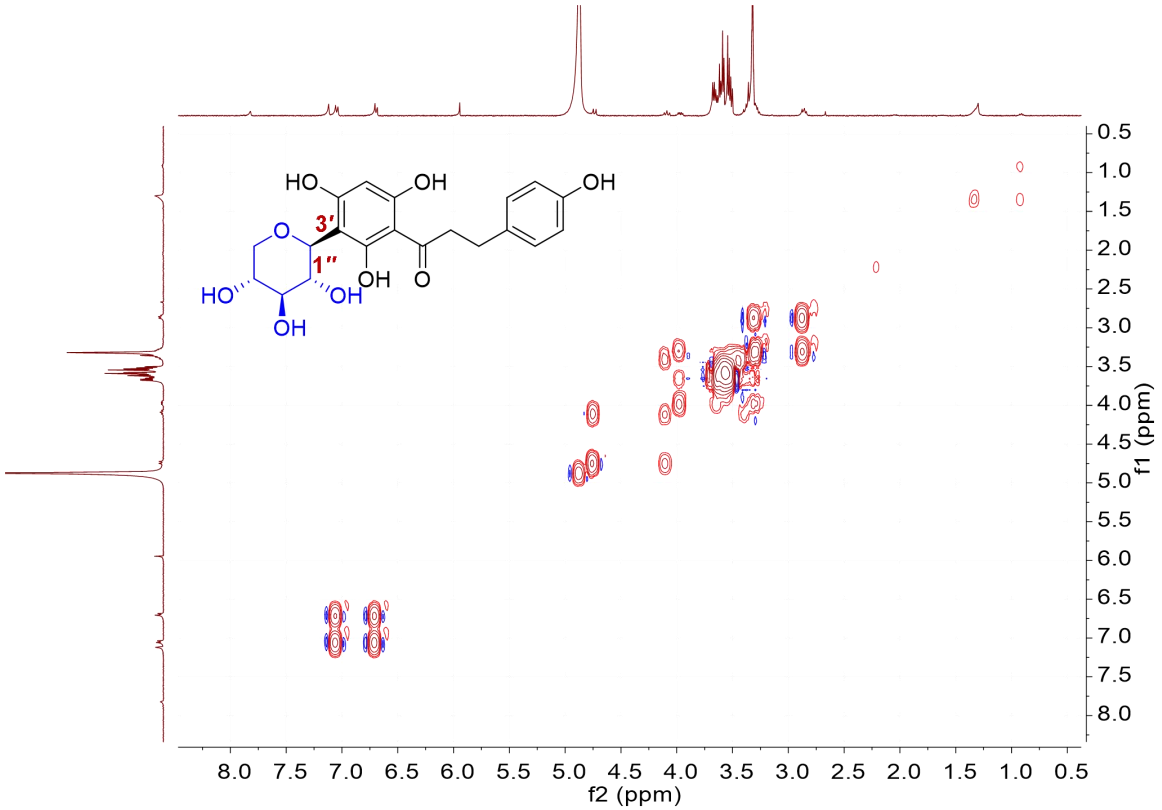


Figure S20. ^{1}H-^{1}H COSY spectrum of **1a** in Methanol-d_4, 400 MHz.

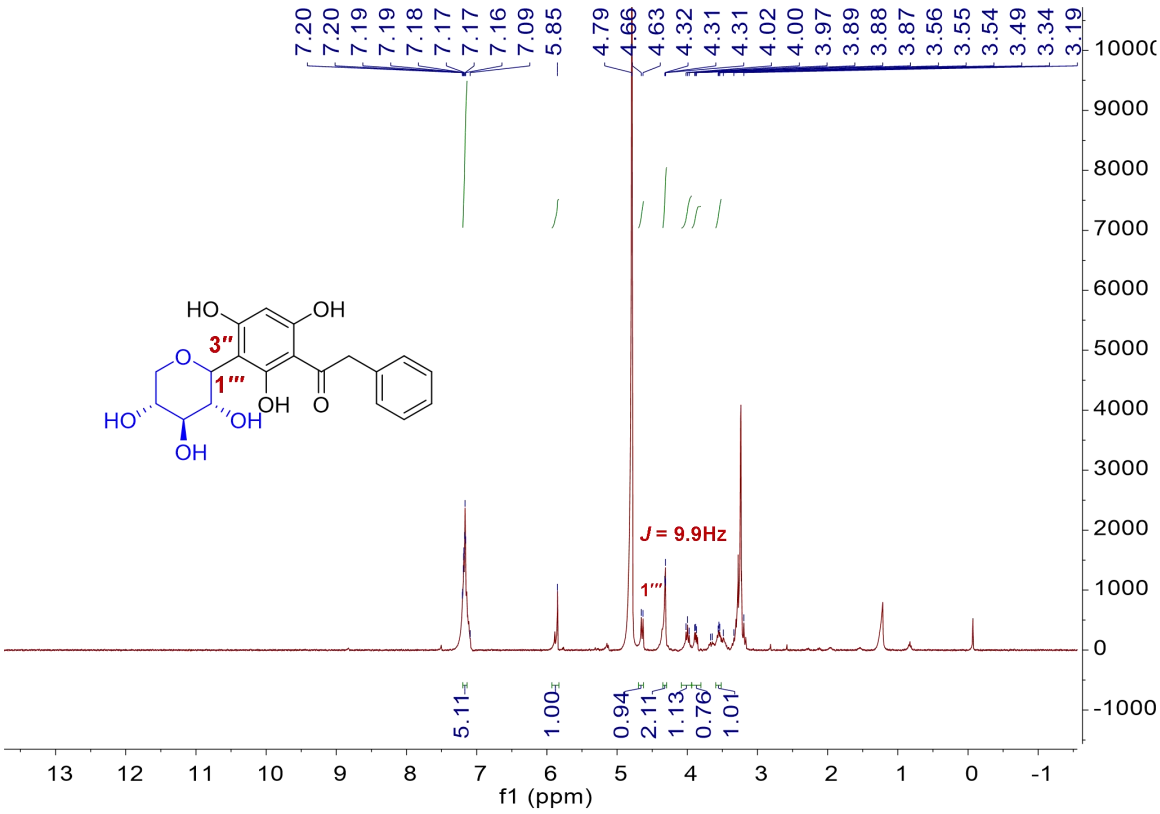


Figure S21. ^{1}H NMR spectrum of **5a** in Methanol-d_4, 400 MHz.

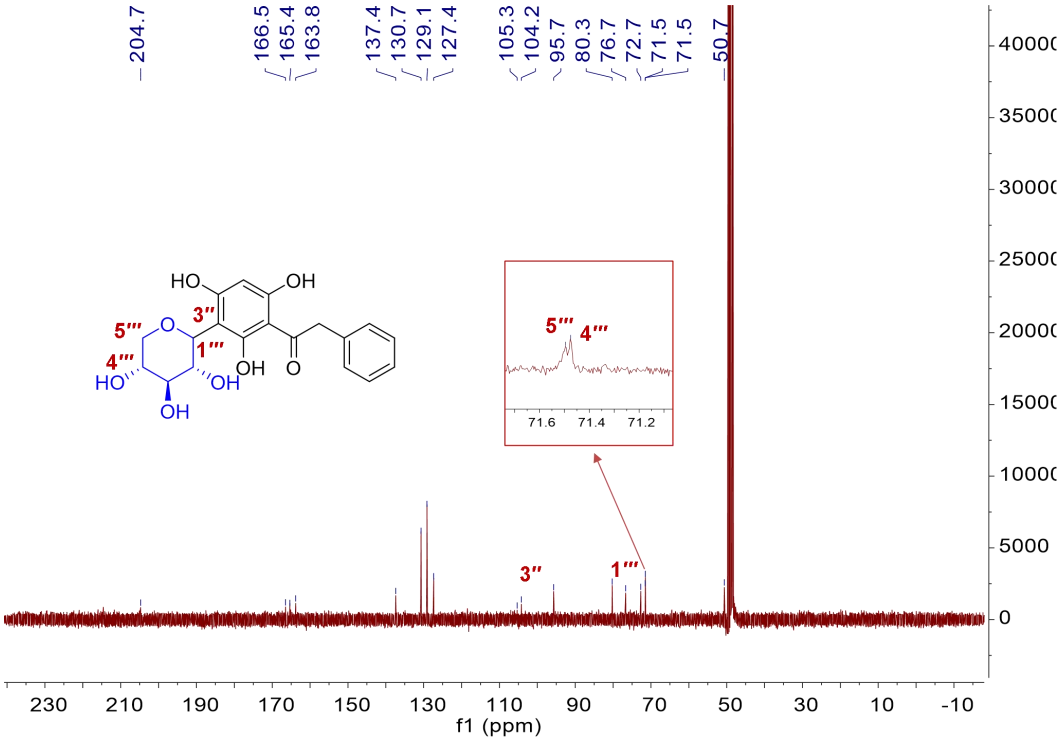


Figure S22. ^{13}C NMR spectrum of **5a** in Methanol-d_4, 100 MHz.

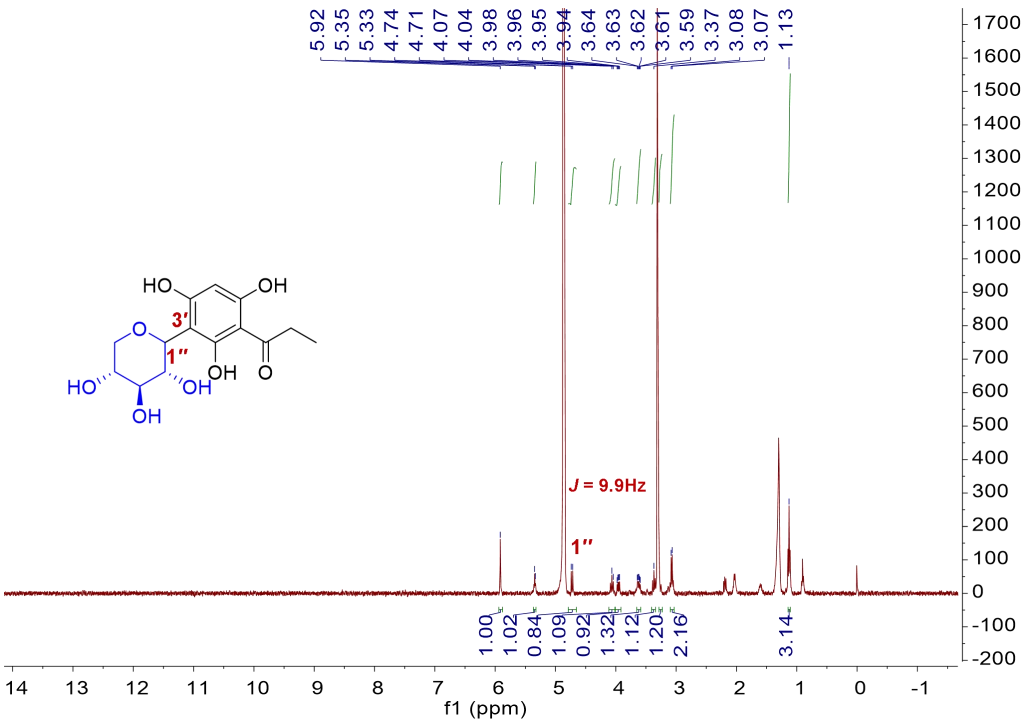


Figure S23. ^{1}H NMR spectrum of **6a** in Methanol-d_4, 400 MHz.

References

[1] D. Chen, SH. Fan, R. Chen, K. Xie, S. Yin, L. Sun, J. Liu, L. Yang, J. Kong, ZH. Yang, J. Dai, *ACS Catal*. 2018, *8*, 4917-4927.

[2] M. Zhang, F. D. Li, K. Li, Z. L. Wang, Y. X. Wang, J. B. He, H. F. Su, Z. Y. Zhang, C. B. Chi, X. M. Shi, C. H. Yun, Z. Y. Zhang, Z. M. Liu, L. R. Zhang, D. H. Yang, M. Ma, X. Qiao, M. Ye, *J. Am. Chem. Soc.* 2020, *142*, 3506-3512.