

## Electronic Supplementary Information

### Functional Characterization of a C-Glycosyltransferase from *Pueraria lobata* with Dual-Substrate Selectivity

Yang-Oujie Bao,<sup>a,‡</sup> Meng Zhang,<sup>a,‡</sup> Xue Qiao,<sup>a,\*</sup> and Min Ye.<sup>a,\*</sup>

<sup>a</sup> State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, 38 Xueyuan Road, Beijing 100191, China

<sup>b</sup> Yunnan Baiyao International Medical Research Center, Peking University, 38 Xueyuan Road, Beijing 100191, China

Corresponding Author

\* E-mail: [yemin@bjmu.edu.cn](mailto:yemin@bjmu.edu.cn).

\* E-mail: [qiaoxue@bjmu.edu.cn](mailto:qiaoxue@bjmu.edu.cn).

‡ Y. B., M. Z. contributed equally to this work.

# Contents

## Experimental

1. General procedures
2. Plant materials
3. Molecular cloning
4. Expression and purification of PICGT
5. Enzyme activity assay
6. Phylogenetic analysis
7. Biochemical properties of PICGT
8. Determination of kinetic parameters
9. Sugar donor selectivity of PICGT
10. Substrates screening of PICGT
11. Scaled-up reactions
12.  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of glycosylated products
13. Protein structure prediction and molecular docking
14. Site-directed mutagenesis of PICGT and enzyme activity assay

## Tables

**Table S1.** CGT genes for BLAST analysis

**Table S2.** Primers for candidate genes

**Table S3.** GT genes for phylogenetic analysis

**Table S4.** Primers for site-directed mutagenesis

## Figures

**Figure S1.** Phylogenetic tree analysis of PICGT.

**Figure S2.** Sequence alignment between PICGT and PIUGT43.

**Figure S3.** SDS-PAGE analysis of purified recombinant PICGT.

**Figures S4 - S5.** NMR spectra of puerarin (**1a**) standard used for HPLC analysis.

**Figure S6.** Effects of pH (A), metal ions and EDTA (B), and temperature (C) on enzyme activity of PICGT.

**Figure S7.** Determination of kinetic parameters for PICGT against daidzein (**1**).

**Figure S8.** The sugar selectivity of PICGT.

**Figures S9 - S11.** LC/MS analysis of PICGT catalyzed products using different sugar donors.

**Figures S12 - S24.** LC/MS analysis of enzymatic reaction products (**1-13**) catalyzed by PICGT.

**Figures S25 - S42.** NMR spectra of glycosylated products (**2a, 3a, 6a**).

**Figures S43.** LC/MS analysis of enzymatic reaction products (**23**) catalyzed by PICGT.

**Figures S44 - S45.** NMR spectra of apigetrin (**23a**) standard used for HPLC analysis.

**Figures S46 - S47.** LC/MS analysis of enzymatic reaction products (**18-19**) catalyzed by PICGT.

**Figures S48 – S53.** NMR spectra of glycosylated products (**18a**).

**Figures S54 – S57.** NMR spectra of nothofagin (**19a**) and Trilobatin (**19b**) standard used for HPLC analysis.

**Figure S58.** Determination of kinetic parameters for PICGT against phloretin (**19**).

**Figure S59.** Conversion rate of PICGT using **1-23** substrates for short-time reactions.

**Figures S60-61.** HPLC analysis of PICGT and PIUGT43 catalyzing **1** and **19**.

**Figure S62.** Structure model of PICGT predicted by AlphaFold2.

**Figure S63 .** Conversion rate of mutants (T13A, L82A, V83A, L86A and F126A) using **1, 19, 23** as the substrates.

**Figures S68-S75.** HPLC analysis of mutants.

**Figure S76.** Conversion rate of PICGT mutants using **1, 19, 23** for short-time reactions.

**Figures S77.** Difference on the molecule docking between PICGT/phloretin and PIUGT43/phloretin.

## 1. General procedures

Compounds **1-3**, **5**, **6**, **9**, **10**, **12-15**, **18-27** were obtained from commercial sources (Solarbio, Beijing, China; Chengmust, Chengdu, China; Desite, Chengdu, China), and the other compounds listed in the substrate library were from the compound library of our laboratory. UDP-glucose, UDP-galactose, UDP-arabinose, UDP-xylose, UDP-glucuronic acid, and UDP-N-acetylglucosamine were purchased from Sigma-Aldrich (St. Louis, USA). HPLC-DAD-ESI-MS<sup>n</sup> (LC/MS) analysis was performed on an Agilent 1100 instrument coupled with a Finnigan LCQ advantage ion trap mass spectrometer (Thermo Fisher Scientific, Waltham, MA). Unless otherwise stated, samples were separated on a Zorbax SB-C18 column (4.6 × 150 mm, 5 μm, Agilent, USA) protected with a Zorbax Extend-C18 guard column (4.6 × 12.5 mm, 5 μm). The mobile phase consisted of methanol (B) and water containing 0.1% (v/v) formic acid (A). A linear gradient elution program was used: 0 min, 10% B; 12 min, 60% B; 16 min, 100% B. UV spectra were recorded by scanning from 200 to 400 nm. For MS analysis, the effluent was introduced into the ESI source of mass spectrometer at 0.2 mL/min via a T-union splitter. The mass spectrometer was operated in the (-)-ESI mode. Semi-preparative HPLC was performed on a Laballiance Series III instrument equipped with a Zorbax SB-C18 column (9.4×250 mm, 5 μm, Agilent, USA). HRESIMS spectra were obtained on a Waters Xevo G2 Q-TOF mass spectrometer (Waters, Milford, Massachusetts, USA). The NMR spectra were recorded on a Bruker AVANCE III-400 instrument (Bruker, Karlsruhe, Germany) at 400 MHz for <sup>1</sup>H and 100 MHz for <sup>13</sup>C, or a Bruker AVANCE III-600 instrument operated at 600 MHz for <sup>1</sup>H and 150 MHz for <sup>13</sup>C in DMSO-*d*<sub>6</sub> using TMS as the reference.

## 2. Plant materials

The seedlings of *Pueraria lobata* were purchased from Yunnan, China. After cultivation in the authors' Laboratory, the whole plants were washed with sterile deionized water, frozen in liquid nitrogen, and then stored at -80 °C before RNA

extraction.

### 3. Molecular cloning

To discover the candidate CGTs, BLAST search against cDNA library (built by Novogene, Beijing, China) of *Pueraria lobata* was applied. The reported CGT genes were used as templates for BLAST analysis (Table S1). A total of 13 candidate UGTs were screened.

The total RNA from *Pueraria lobata* was extracted using the TranZol™ kit (Transgen Biotech, China), and was reverse-transcribed (RT) to cDNA with SMARTer™ RACE cDNA Amplification kit (Clontech, USA). Primers were designed according to the cDNA fragments screened in BLAST analysis (Table S2). The target cDNA was amplified by PCR (3 min at 94 °C; 30 s at 94 °C, 30 s at 60 °C, 1.5 min at 68 °C, 36 cycles; 10 min at 68 °C) with TransStart® KD Plus DNA Polymerase kit (Transgen, China) using the designed primers, and inserted into pET 28a (+) vector (Invitrogen, USA) according to the Quick-change method. The recombinant plasmid was then introduced into *E. coli* BL21 (DE3) (Transgen Biotech, China) for heterologous expression.

### 4. Expression and purification of PICGT

*E. coli* cells containing the recombinant plasmid were grown in 500 mL LB medium with kanamycin (50 µg/mL) at 37°C. After OD<sub>600</sub> reached 0.5-0.7, the cells were induced with 0.5 mM IPTG at 16°C for another 18 h. Then cells were harvested by centrifugation (7,000 rpm, 3 min), and then resuspended in 15 mL lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 8.0, 300 mM NaCl, 10 mM imidazole). Then cells were disrupted by sonication in an ice bath, and the cell debris was removed by centrifugation at 8,000 rpm for 40 min at 4°C. The supernatant was collected and loaded onto a pre-equilibrated column (His Trap™ HP, 5mL, GE Healthcare) and then washed with two different mobile phases (A: 50 mM Tris, 10 mM NaCl; B: 50 mM Tris, 1 M NaCl). Elution was carried out with different concentrations of elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0,

300 mM NaCl, 25/50/100/150/200/250 mM imidazole). The target protein was eluted by the elution buffer containing about 100 mM imidazole. Subsequently, the fraction was further purified with a molecular sieve column (Superdex™ 200 increase 10/300 GL, GE Healthcare) and an ion exchange column (Resource™ Q, GE Healthcare). Finally, the purity of target protein was >90% according to SDS-PAGE analysis (Figure S3). The protein concentration was determined by the B-500 Biophotometer (METASH, China).

## 5. Enzyme activity assay

The function of PICGT was characterized by co-incubating 1 µg purified protein, 0.1 mM daidzein (**1**), and 0.5 mM UDP-Glc in 100 µL of 50 mM NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8.0, 37 °C, 2 h). Reactions were quenched with ice cold MeOH and centrifuged at 15,000 rpm for 10 min. The supernatants were analyzed by LC/MS.

## 6. Phylogenetic analysis

The phylogenetic tree was constructed using MEGA 6.0 Software with the Neighbor-Joining method based on ClustalW multiple alignments. (Figure S1)

## 7. Biochemical properties of PICGT

The biochemical properties of PICGT were investigated using **1** as the sugar acceptor and UDP-Glc as the sugar donor. To test the optional pH value for PICGT activity, assays were performed in different reaction buffers ranged in pH values from 4.0-6.0 (citric acid-sodium citrate buffer), 6.0-8.0 (Na<sub>2</sub>HPO<sub>4</sub>-NaH<sub>2</sub>PO<sub>4</sub> buffer), 7.0-9.0 (Tris-HCl buffer), and 9.0-11.0 (Na<sub>2</sub>CO<sub>3</sub>-NaHCO<sub>3</sub> buffer). To study the optimal reaction temperature for PICGT activity, the reactions were incubated at different temperatures (4, 18, 25, 37, 50, 70 °C). To determine the necessity of divalent metal ions for PICGT activities, FeCl<sub>2</sub>, BaCl<sub>2</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, and CoCl<sub>2</sub> were used

individually at a final concentration of 5 mM. To determine the effect of EDTA on the activity of PICGT, EDTA was added into the reaction mixture at 5 mM. For each condition, three parallel reactions were conducted (n=3). Reactions were quenched with ice cold MeOH and centrifuged at 15,000 rpm for 10 min. Supernatants were analyzed by HPLC as described above. (Figure S6)

## 8. Determination of kinetic parameters

Kinetic parameters of PICGT reaction were calculated using daidzein (**1**) as substrate. Assays were performed in a final volume of 100  $\mu$ L, consisting of 50 mM Na<sub>2</sub>HPO<sub>4</sub>-NaH<sub>2</sub>PO<sub>4</sub> (pH 8.0), 30 nM PICGT, 10 mM of saturated UDP-glucose, and different concentrations of daidzein (2.5, 5, 10, 20, 40, 60, 80  $\mu$ M). After incubating at 37°C for 10 min, the reactions were quenched with ice cold MeOH and centrifuged at 15,000 rpm for 10 min, and the supernatants were analyzed by HPLC. The samples were separated on a Zorbax SB-C18 column (4.6 $\times$ 150 mm, 5  $\mu$ m, Agilent, USA), with MeOH (B) and H<sub>2</sub>O containing 0.1% formic acid (v/v, A) as the mobile phase. A linear gradient HPLC elution program was used: 0 min, 10% B; 12 min, 60% B; 16 min, 100% B. All experiments were performed in triplicate. The  $K_m$  value was calculated using the method of Michaelis-Menten plot. (Figure S7)

## 9. Sugar donor selectivity of PICGT

To investigate the selectivity of PICGT for other sugar donors, glycosylation reactions with UDP-glucuronic acid (UDP-GluA), UDP-xylose (UDP-Xyl), UDP-galactose (UDP-Gal), UDP-arabinose (UDP-Ara) and UDP-*N*-acetylglucosamine (UDP-GlcNAc) were tested by using substrate **1** as acceptor, respectively. PICGT was purified, and put into a short-time or long-time reaction, respectively. Short-time reactions were constructed in a 100  $\mu$ L NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> buffer, and 0.5  $\mu$ g PICGT was added for 30 minutes. Long-time reactions were constructed in a 100  $\mu$ L NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub> buffer, and 1  $\mu$ g PICGT was added for 8 hours. All the reactions

were carried out as described above and analyzed by HPLC and LC/MS (Figures S8-S11).

## 10. Substrates screening of PICGT

PICGT was purified and put into a short-time or long-time reaction. Short-time reactions were constructed in a 100  $\mu\text{L}$   $\text{NaH}_2\text{PO}_4\text{-Na}_2\text{HPO}_4$  buffer, and 0.5  $\mu\text{g}$  PICGT was added for 30 minutes. Long-time reactions were constructed in a 100  $\mu\text{L}$   $\text{NaH}_2\text{PO}_4\text{-Na}_2\text{HPO}_4$  buffer, and 1  $\mu\text{g}$  PICGT was added for 8 hours. Reactions were incubated at 37  $^\circ\text{C}$  and were terminated by adding 100  $\mu\text{L}$  MeOH.

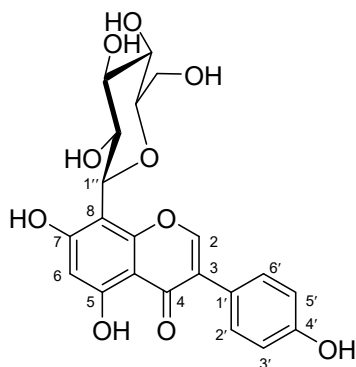
After centrifugation at 15,000 rpm for 10 min, the supernatants were collected, and aliquots were analyzed by LC/MS. (Figures S12-S24, S43, S46-S47) The NMR spectra of standards we used in HPLC analysis are shown in Figure S4-S5, S44-S45, S54-S57.

## 11. Scaled-up reactions

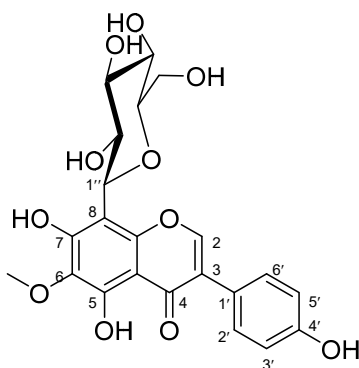
Scaled-up reactions were performed at 37  $^\circ\text{C}$  overnight using a total volume of 50 mL assay buffer solution (50 mM  $\text{Na}_2\text{HPO}_4\text{-NaH}_2\text{PO}_4$ , pH 8.0), containing 1 mM aglycone (~10 mg, dissolved in DMSO), 2-4 mM sugar donor, and 5 mL purified PICGT (10-15 mg of protein). The reaction was terminated by adding 50 mL MeOH, and centrifuged at 12,000 rpm for 20 min to obtain the supernatant. The organic solvent was removed under reduced pressure. Then the residue was dissolved in 1.0-1.5 mL of methanol and subjected to semi-preparative HPLC. The pure products were dissolved in  $\text{DMSO-}d_6$  to record the NMR spectra (Figures S25-S42, S48-S53).

## 12. $^1\text{H}$ and $^{13}\text{C}$ NMR data of glycosylated products



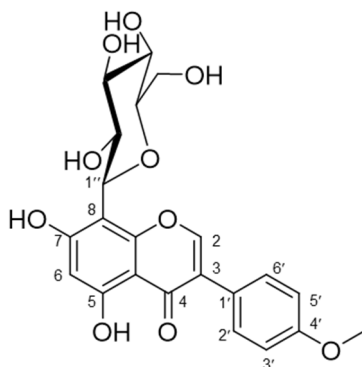


**Genistein 8-C- $\beta$ -D-glucoside (2a):** Yield 5.68 mg. HR-ESI-MS calcd. for  $C_{21}H_{20}O_{10}$  [M-H]<sup>-</sup>: 431.7. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 13.20 (H, s, 5-OH), 8.40 (H-2, s), 6.3 (H-6, s), 7.38 (2H, d, H-2', 6', J = 12 Hz), 6.81 (2H, d, H-3', 5', J = 12 Hz), 4.66 (H, d, H-1'', J = 8 Hz), 3.95 (H, m, H-2''), 3.22 (H, m, H-3''), 3.18 (H, m, H-4''), 3.19 (H, m, H-5''), 3.70/3.39 (2H, m, H-6''). <sup>13</sup>C NMR (150 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 153.9 (C-2), 122.0 (C-3), 180.5 (C-4), 161.1 (C-5), 98.6 (C-6), 163.1 (C-7), 104.2 (C-8, C-10), 156.1 (C-9), 121.1 (C-1'), 130.1 (C-2', C-6'), 115.1 (C-3', C-5'), 157.4 (C-4'), 73.2 (C-1''), 70.7 (C-2'', C-4''), 78.8 (C-4''), 81.7 (C-5''), 61.5 (C-6'').

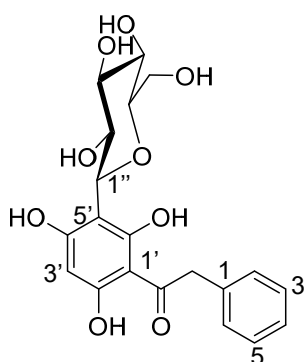


**Tectorigenin 8-C-glucoside(3a):** Yield 5.40 mg. HR-ESI-MS calcd. for  $C_{22}H_{22}O_{11}$  [M-H]<sup>-</sup>: 461.3. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 3.78 (3H, s, -CH<sub>3</sub>), 8.42 (H-2, s), 7.39 (2H, d, H-2', 6', J = 8.8 Hz), 6.82 (2H, d, H-3', 5', J = 8.4 Hz), 9.59 (H, s, 4'-OH), 13.40 (H, s, 5-OH), 4.70 (H, d, H-1'', J = 9.6 Hz), 3.98 (H, m, H-2''), 3.24 (H, m, H-3''), 3.16 (H, m, H-4''), 3.19 (H, m, H-5''), 3.71/3.40 (2H, m, H-6''). <sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 154.0 (C-2), 121.1 (C-3), 180.9 (C-4), 156.2 (C-5, C-7), 130.2 (C-6, C-2', C-6'), 104.3 (C-8, C-10), 152.3 (C-9), 60.2 (-CH<sub>3</sub>), 121.6 (C-1'), 115.1 (C-3', C-5'), 157.4 (C-

4'), 73.6 (C-1''), 70.6(C-2'', C-4''), 78.7 (C-3''), 81.1 (C-5''), 61.5 (C-6'').



**Biochanin A 8-C- $\beta$ -D -glucoside (6a)**: Yield 11.76 mg. HR-ESI-MS calcd. for  $C_{22}H_{22}O_{10}$  [M-H]<sup>-</sup>: 445.2. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 13.17 (H, s, 5-OH), 3.79 (3H, s, -CH<sub>3</sub>), 8.46 (H-2, s), 6.31 (H-6, s), 7.51 (2H, d, H-2', 6', J = 12 Hz), 7.00 (2H, d, H-3', 5', J = 8 Hz), 4.67 (H, d, H-1'', J = 8 Hz), 3.97 (H, m, H-2''), 3.19 (H, m, H-3''), 3.16 (H, m, H-4''), 3.19 (H, m, H-5''), 3.69/3.40 (2H, m, H-6''). <sup>13</sup>C NMR (150 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 154.2 (C-2), 121.7 (C-3), 180.4 (C-4), 161.1 (C-5), 98.9 (C-6), 163.2 (C-7), 104.3 (C-8), 104.6(C-10), 156.4(C-9), 122.9 (C-1'), 130.2 (C-2', C-6'), 113.7 (C-3', C-5'), 159.2 (C-4'), 55.2(-CH<sub>3</sub>), 73.2 (C-1''), 70.7(C-2'', C-4''), 78.8 (C-3''), 81.8 (C-5''), 61.4 (C-6'').



**2-Phenyl-2',4',6'-trihydroxyacetophenone 8-C- $\beta$ -D -glucoside (18a)**: Yield 7.9 mg. HR-ESI-MS calcd. for  $C_{20}H_{22}O_9$  [M-H]<sup>-</sup>: 406.1. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$ : 4.35 (2H, m, -CH<sub>2</sub>), 7.28 (2H, m, H-2,6), 7.22 (2H, m, H-3,5), 7.20 (H, m, H-4), 5.96 (H, s,

H-3'), 4.50 (H, d, H-1'', J = 8 Hz), 3.89 (H, m, H-2''), 3.14 (H, m, H-3''), 3.10 (H, m, H-4''), 3.11 (H, m, H-5''), 3.65/3.38 (2H, m, H-6''). <sup>13</sup>C NMR (100 MHz, DMSO-*d*<sub>6</sub>) δ: 136.1 (C-1), 129.6 (C-2,6), 128.1 (C-3,5), 126.3 (C-4), 49.0 (-CH<sub>2</sub>), 202.6 (C=O), 104.1 (C-1'), 161.5 (C-2'), 94.6 (C-3'), 164.0 (C-4'), 103.6 (C-5), 165.1 (C-6), 73.5 (C-1''), 70.5 (C-2''), 79.0 (C-3''), 70.6 (C-4''), 81.3 (C-5''), 61.3 (C-6'').

### 13. Protein structure prediction and molecular docking

The structure of PICGT was predicted by artificial intelligence-based tool AlphaFold2. The PICGT/UDP-Glc model was then constructed by alignment of PICGT model with previously reported structures TcCGT1 (PDB ID: 6JTD), *A. thaliana* UGT72B1 (PDB ID: 2VCE), *M. truncatula* UGT71G1 (PDB ID: 2ACW), and *V. vinifera* VvGT1 (PDB ID: 2C1Z). Subsequently, the predicted structure was used as target protein in Autodock<sup>1</sup>. Three different structural types of substrates (Daidzein (**1**), Phloretin (**19**), Apigenin (**23**)) used in the docking studies were firstly identified in the functional studies on PICGT as described in the paper. A total of 100 docking conformations was generated for each substrate. The substrates were sampled flexibly in the docking and no core or any constraints were applied to enable free, unbiased docking. The docked models were evaluated by docking scores and reasonability of the stereo-chemistry properties and the interactions between UDP-Glc and substrate. Three docking results were used in the end (Figure 3A-C).

To investigate the difference between PICGT and its homologous enzyme PIUGT43, the structure model of PIUGT43 was constructed using AlphaFold2. UDP-Glc was docked into the sugar donor binding pocket, and phloretin (**19**) was then docked into the substrate binding pocket using the same parameters as PICGT/phloretin.

Although most of the different amino acid residues were distributed in the non-substrate binding region, Phe11 and His 314 in PICGT were replaced by Ile11 and Gln314 in PIUGT43. Structure alignment of PICGT and PIUGT43 was then constructed (Figure S77A), and we could see the binding conformations of phloretin in the two enzymes were different. In PIUGT43, the carbonyl group of Gln314 showed a short

distance to 4-OH of phloretin, which might cause a twist of B ring. Accordingly, A ring of phloretin in PIUGT43 also rotated by a tiny angle. The greater distance between C-3' of phloretin and C-1 of Glc and between the carbonyl group of Asn16 and 7-OH of phloretin in PIUGT43 may cause the relatively weak activity of PIUGT43. (Figure S77C)

#### **14. Site-directed mutagenesis of PICGT and enzyme activity assay**

Site-directed mutants of PICGT, including F11A, T13A, N16H, N16Q, N16A, N16D, N16L, N16S, N16K, N16V, L82A, V83A, L86A, F126A and L187A, were constructed using a Fast Mutagenesis System kit (TransGen Biotech, China) according to the manufacturer's instructions. Primers used in mutagenesis were listed in Table S3. The corresponding degenerate primers designed to construct the site-directed mutants are listed in Table S3. After verification of the mutant sequences, the recombinant plasmids were transformed into *E. coli* BL21(DE<sub>3</sub>) for heterologous expression. The protein expression, purification and enzyme activity assays of mutants were performed under the same conditions for 4 hours when daidzein (**1**) was exactly catalyzed completely.

**Table S1. CGT genes for BLAST analysis**

<b>GT name</b>	<b>GenBank (*) UniProtKB/Swiss-Prot(#)</b>	<b>Reference</b>
PIUGT43	A0A172J2G3.2 <sup>#</sup>	2
CuCGT	LC131334.1*	3
FeCGT	LC131333.1*	3
MiCGT	A0A0M4KE44.1 <sup>#</sup>	4
MiCGTb	AMM73095.1*	5
FeCGTa	A0A0A1HA03.1 <sup>#</sup>	6
FeCGTb	A0A0A1H7N4.1 <sup>#</sup>	6
UGT708D1	RDY12992.1*	7
PhCGT4	MK616591.1*	8
PpCGT1	MK616593.1*	8
OsCGT	C3W7B0.1 <sup>#</sup>	9
UGT708A6	A0A096SRM5.1 <sup>#</sup>	10
TcCGT1	QCZ42162.1*	11
GtUF6CGT1	A0A0B6VIJ5.1 <sup>#</sup>	12
WjGT1	BBI55602.1*	13

**Table S2. Primers for candidate genes**

Primers	
20738	F: GTGGACAGCAAATGGGTCGCGGATCCATGACTAGATATGAAGTGGTTTTC R: TGGTGGTGCTCGAGTGCGGCCGCAAGCTTACCTGTCAACTCTTGAATAA
31236	F: GTGGACAGCAAATGGGTCGCGGATCCATGGGTAGTAGTGAATATCATACT R: TGGTGCTCGAGTGCGGCCGCAAGCTTGTGGCTCAATGAACTTAACTCC
21261	F: ACAGCAAATGGGTCGCGGATCCATGTTCCGCCACCGCTTCCTCCTCATCAT R: TGGTGCTCGAGTGCGGCCGCAAGCTTTGGTCCAACCTCAACATCTTCCAA
13493	F: GTGGACAGCAAATGGGTCGCGGATCCATGGGTAACGAGAATCGCGAGCTT R: TGGTGCTCGAGTGCGGCCGCAAGCTTATGGGCACGCGACCTCAAATCCT
8573	F: GGTGGACAGCAAATGGGTCGCGGATCCATGGATGGGAAAGGGAAACTACA R: TGGTGGTGCTCGAGTGCGGCCGCAAGCTTATTTTGAAGGCATGCAATGGA
20823	F: TGGTGGACAGCAAATGGGTCGCGGATCCATGGGGTCTGAAGCTCCCATTCA R: TGGTGGTGCTCGAGTGCGGCCGCAAGCTTTATCTCGTTGACCCCGACAGCT
7967	F: TGGTGGACAGCAAATGGGTCGCGGATCCATGGGGAGCACTGAGATGAAGAA R: GTGGTGCTCGAGTGCGGCCGCAAGCTTGTTGCTTCCCATCATATTATCAAT
21094	F: GGTGGACAGCAAATGGGTCGCGGATCCATGAAGAAGAAAGCAGAGCTGAT R: TGGTGCTCGAGTGCGGCCGCAAGCTTATTGCTACCTGTCATAACATCAAT
14833	F: GGTGGACAGCAAATGGGTCGCGGATCCATGGATTCAACACCCCTTTGCAA R: TGGTGCTCGAGTGCGGCCGCAAGCTTGTTTTGGCCAAACATTTTGACAT
16969	F: GTGGACAGCAAATGGGTCGCGGATCCATGATTTTCCAAACAAACAATCCT R: TGGTGCTCGAGTGCGGCCGCAAGCTTAGCGCTACATGTAAATTGCGACTC
13128	F: GTGGACAGCAAATGGGTCGCGGATCCATGGCCCCGAAACCAATTCAATT R: GTGGTGCTCGAGTGCGGCCGCAAGCTTATCCTGTCTTGCCAACTGCGAG
13090	F: GGTGGACAGCAAATGGGTCGCGGATCCATGGATTCAACGGCCCTTAGCCA R: GTGGTGCTCGAGTGCGGCCGCAAGCTTGCTTTGGTCTAACCCCTTTGACAT
8472	F: TGGACAGCAAATGGGTCGCGGATCCATGTCATGCAGAAAGCTGAACTTGC R: GGTGCTCGAGTGCGGCCGCAAGCTTAGGATTGGGTGCTTTGGAAAGTTTG

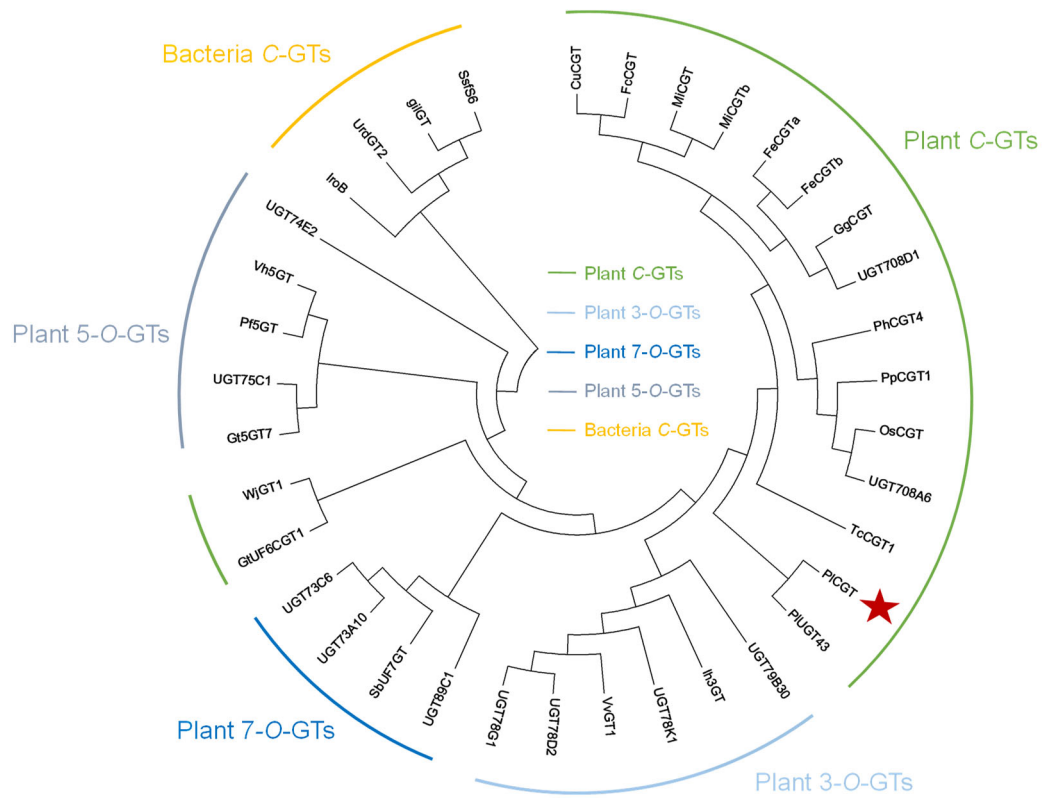
**Table S3. GT genes for phylogenetic analysis**

Functional type	GT name	GenBank (*) UniProtKB/Swiss-Prot( <sup>#</sup> )	Reference
Plant C-GT	CuCGT	LC131334.1*	3
	FcCGT	LC131333.1*	3
	MiCGT	A0A0M4KE44.1 <sup>#</sup>	4
	MiCGTb	AMM73095.1*	5
	FeCGTa	A0A0A1HA03.1 <sup>#</sup>	6
	FeCGTb	A0A0A1H7N4.1 <sup>#</sup>	6
	GgCGT	MH998596*	14
	UGT708D1	RDY12992.1*	7
	PhCGT4	MK616591.1*	8
	PpCGT1	MK616593.1*	8
	OsCGT	C3W7B0.1 <sup>#</sup>	9
	UGT708A6	A0A096SRM5.1 <sup>#</sup>	10
	TcCGT1	QCZ42162.1*	11
	PIUGT43	A0A172J2G3.2 <sup>#</sup>	2
	GtUF6CGT1	A0A0B6VIJ5.1 <sup>#</sup>	12
	WjGT1	BBI55602.1*	13
Bacterial C-GT	IroB	QBQ68863.1*	15
	UrdGT2	AAF00209.1*	16
	gilGT	AAP69578.2*	17
	SsfS6	ADE34512.1*	18
Plant 3-O-GT	UGT79B30	NM_001359019.1*	19
	Ih3GT	BAD83701.1*	20
	UGT78K1	ADC96620.1*	21
	VvGT1	NM_001397857.1*	22
	UGT78D2	NP_197207.1*	23
	UGT78G1	A6XNC6.1 <sup>#</sup>	24
Plant 7-O-GT	UGT89C1	Q9LNE6.1 <sup>#</sup>	25
	SbUF7GT	BAA83484.1*	26
	UGT73A10	BAG80536.1*	27
	UGT73C6	OAP07438.1*	28
Plant 5-O-GT	Gt5GT7	B2NID7.1 <sup>#</sup>	29
	UGT75C1	NM_001361345.1*	30
	Pf5GT	BAA36421.1*	31
	Vh5GT	BAA36423.1*	31
	UGT74E2	OAP17332.1*	28

**Table S4. Primers for site-directed mutagenesis**

<b>F11A</b>	F: GTGTTTATTGCAGCACCGACCCTGGGTAATCT R: TGCTGCAATAAACACAACCTTCATAACGGGTCA
<b>T13A</b>	F: ATCGCTTTCCCAGCACTTGGCAACCTCGTTCCA R: TGCTGGGAAAGCGATGAAAACCACTTCATATCT
<b>N16H</b>	F: TTTCCGACCCTGGGTCATCTGGTGCCGCAAGTG R: GACCCAGGGTCGGAAATGCAATAAACACAAC
<b>N16A</b>	F: TTTCCGACCCTGGGTCGACTGGTGCCGCAAGTG R: TGCACCCAGGGTCGGAAATGCAATAAACACAA
<b>N16K</b>	F: CCCAACCCCTTGGCAAACCTCGTTCCACAAGTTGA R: TTTGCCAAGGGTTGGGAAAGCGATGAAAACCAC
<b>N16D</b>	F: CCCAACCCCTTGGCGATCTCGTTCCACAAGTTGA R: ATCGCCAAGGGTTGGGAAAGCGATGAAAACCAC
<b>N16Q</b>	F: CCCAACCCCTTGGCCAGCTCGTTCCACAAGTTGA R: CTGGCCAAGGGTTGGGAAAGCGATGAAAACCAC
<b>N16L</b>	F: CCCAACCCCTTGGCCTGCTCGTTCCACAAGTTGA R: CAGGCCAAGGGTTGGGAAAGCGATGAAAACCAC
<b>N16V</b>	F: CCAACCCTTGGCGTTCTCGTTCCACAAGTTGAA R: AACGCCAAGGGTTGGGAAAGCGATGAAAACCAC
<b>N16S</b>	F: CCAACCCTTGGCAGCCTCGTTCCACAAGTTGAA R: CTGCCAAGGGTTGGGAAAGCGATGAAAACCAC
<b>L82A</b>	F: GCAGTACCAAACCGCAGTCGGCTTCCTCTCCCTC R: TGCGGTTTGGTACTGCTCGGGTGCGGGCGGGTC
<b>V83A</b>	F: GTACCAAACCTTGGCAGGCTTCCTCTCCCTCCA R: TGCCAAGGTTTGGTACTGCTCGGGTGCGGGCGGG
<b>L86A</b>	F: CTTGGTCGGCTTCGCATCCCTCCATATGCAAAA R: TGCGAAGCCGACCAAGGTTTGGTACTGCTCGGGT
<b>D124A</b>	F: GGCCATCTTCGTGCAATGTTCTCTACCACCCTC R: TGCGACGAAGATGGCCGCGAGTCGAACCGAGTT
<b>F126A</b>	F: CTTGTCGACATGGCATCTACCACCCTCATCGA R: TGCCATGTCGACGAAGATGGCCGCGAGTCGAAC
<b>L187A</b>	F: CGTTTTGCCGAACGCAGTGTTGGACGCGAAAGAC R: TGCGTTCGGCAAAAACGGAACGGGGCAAAAAGGTT

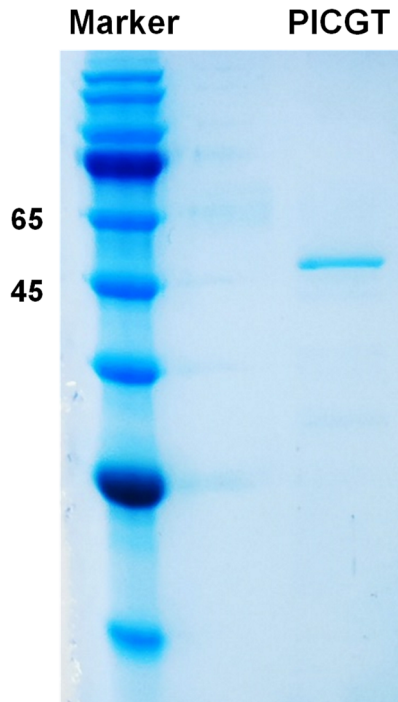




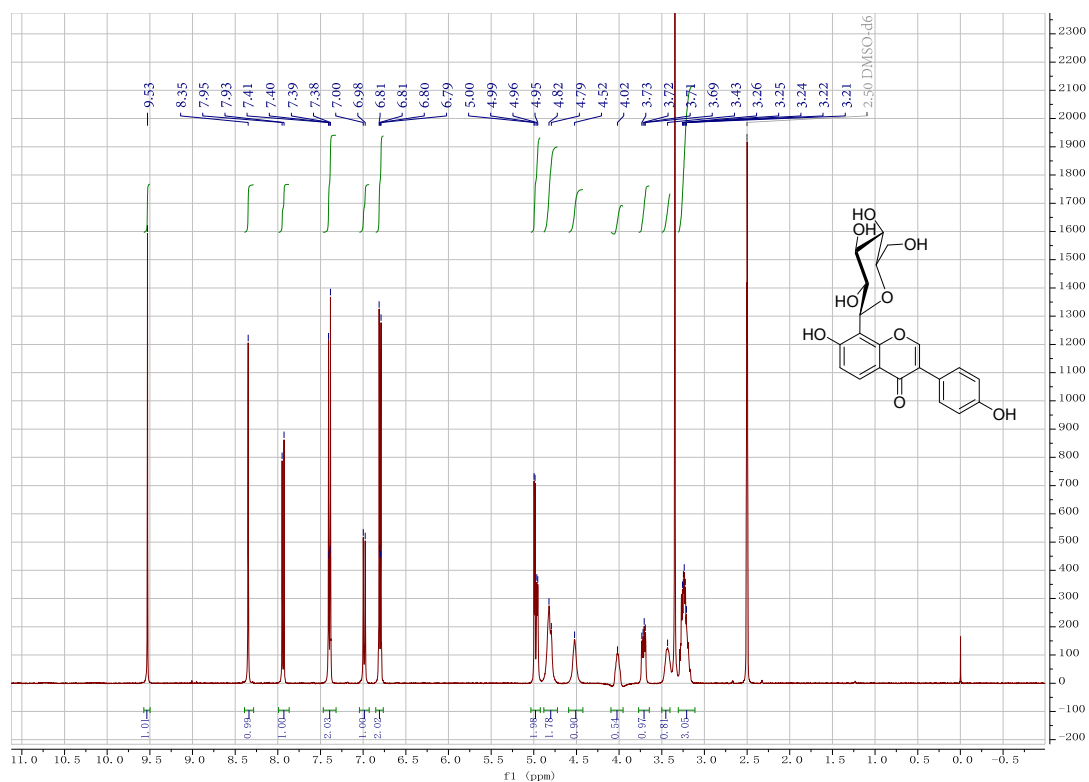
**Figure S1.** Phylogenetic tree analysis of PICGT1. GT genes used in the figure and their information were listed in [Table S3](#).

	1	10	20	30	40	50	60																																																					
PlCGT	M	T	R	Y	E	V	V	F	I	A	F	P	T	L	G	N	L	V	P	Q	V	E	F	A	N	L	L	T	K	H	D	P	R	F	S	A	T	I	L	T	V	S	M	P	Q	R	P	L	M	N	T	Y	V	Q	A	R	A	S	S	A
PlUGT43	M	T	R	Y	E	V	V	F	I	A	I	P	T	L	G	N	L	V	P	Q	V	E	F	A	N	L	L	T	K	H	D	P	R	F	S	A	T	I	L	T	V	S	M	P	Q	R	P	L	M	N	T	Y	V	Q	A	R	A	S	S	A
		70	80	90	100	110	120																																																					
PlCGT	A	N	I	K	L	L	Q	L	P	T	V	D	P	P	A	E	Q	Y	Q	T	L	V	G	F	L	S	L	H	M	Q	N	H	K	H	H	V	K	H	A	L	L	N	L	I	K	T	T	E	S	N	S	S	N	S	V	R	L	A	A	
PlUGT43	A	N	I	K	L	L	Q	L	P	I	V	D	P	P	A	E	Q	Y	Q	T	L	V	G	F	L	S	L	H	M	Q	N	H	K	H	H	V	K	H	A	L	L	N	L	M	K	T	T	E	S	N	S	S	N	S	V	R	L	A	A	
		130	140	150	160	170	180																																																					
PlCGT	I	F	V	D	M	F	S	T	T	L	I	D	V	A	T	E	L	A	V	P	C	Y	L	F	F	A	S	P	A	S	C	L	G	F	T	L	D	L	P	R	V	D	L	A	E	S	K	S	E	F	T	V	P	C	F	K	N	L	L	P
PlUGT43	I	F	V	D	M	F	S	T	T	L	I	D	V	A	A	E	L	A	V	P	C	Y	L	F	F	A	S	P	A	S	C	L	G	F	T	L	D	L	P	R	F	D	L	A	E	S	K	S	E	F	T	V	P	C	F	K	N	L	L	P
		190	200	210	220	230	240																																																					
PlCGT	R	S	V	L	P	N	L	V	L	D	A	K	D	G	T	F	W	L	S	Y	H	A	R	R	Y	K	E	T	K	G	V	V	I	N	T	L	Q	E	L	E	T	H	A	L	Q	S	L	H	N	S	Q	L	Q	R	V	Y	P	I	G	
PlUGT43	R	S	V	F	P	N	L	V	L	D	A	K	D	G	T	F	W	L	S	Y	H	A	R	R	Y	K	E	T	K	G	I	V	I	N	T	L	Q	E	L	E	T	H	A	L	Q	S	L	H	N	S	Q	L	Q	R	V	Y	P	I	G	
		250	260	270	280	290	300																																																					
PlCGT	P	I	L	D	L	V	G	S	A	Q	W	D	P	N	P	P	Q	Y	K	R	I	M	E	W	L	D	Q	Q	P	L	S	S	V	L	L	C	F	G	S	M	G	S	L	E	A	N	Q	V	E	E	I	A	I	G	L	E	R	A	G	
PlUGT43	P	I	L	D	L	V	G	S	A	Q	W	D	P	N	P	A	Q	Y	K	R	I	M	E	W	L	D	Q	Q	P	L	S	S	V	L	L	C	F	G	S	M	G	S	L	E	A	N	Q	V	E	E	I	A	I	G	L	E	R	A	G	
		310	320	330	340	350	360																																																					
PlCGT	V	R	F	L	W	A	L	R	E	S	P	K	A	H	L	E	Y	P	R	D	Y	E	N	H	K	D	V	L	P	D	G	F	L	E	R	T	N	N	I	G	L	V	C	G	W	P	Q	A	V	V	L	A	H	K	A	I	G	G	F	
PlUGT43	V	R	F	L	W	A	L	R	E	S	P	K	A	Q	L	E	Y	P	R	D	Y	E	N	H	K	D	V	L	P	D	G	F	L	E	R	T	N	N	I	G	L	V	C	G	W	P	Q	A	V	V	L	A	H	K	A	V	G	G	F	
		370	380	390	400	410	420																																																					
PlCGT	V	S	H	C	G	W	N	S	I	L	E	S	L	W	H	G	V	P	V	A	T	W	P	L	Y	S	E	Q	Q	M	N	A	F	Q	M	V	R	E	L	G	L	A	V	E	I	S	V	D	Y	R	V	G	A	D	L	V	R	A	E	E
PlUGT43	V	S	H	C	G	W	N	S	I	L	E	S	L	W	H	G	V	P	V	A	T	W	P	L	Y	S	E	Q	Q	M	N	A	F	Q	M	V	R	D	L	G	L	A	V	E	I	S	V	D	Y	R	V	G	A	D	L	V	R	A	E	E
		430	440	450	460																																																							
PlCGT	V	E	N	G	V	R	S	L	M	K	G	G	D	E	I	R	R	K	V	K	E	M	S	D	T	C	R	G	A	L	L	E	T	G	S	S	Y	S	Y	L	V	S	L	I	Q	E	L	T	S											
PlUGT43	V	E	N	L	R	S	L	M	K	G	G	D	E	I	R	R	K	V	K	E	M	S	D	T	C	R	G	A	L	L	E	N	G	S	S	Y	S	N	L	V	S	L	I	Q	E	L	T	S												

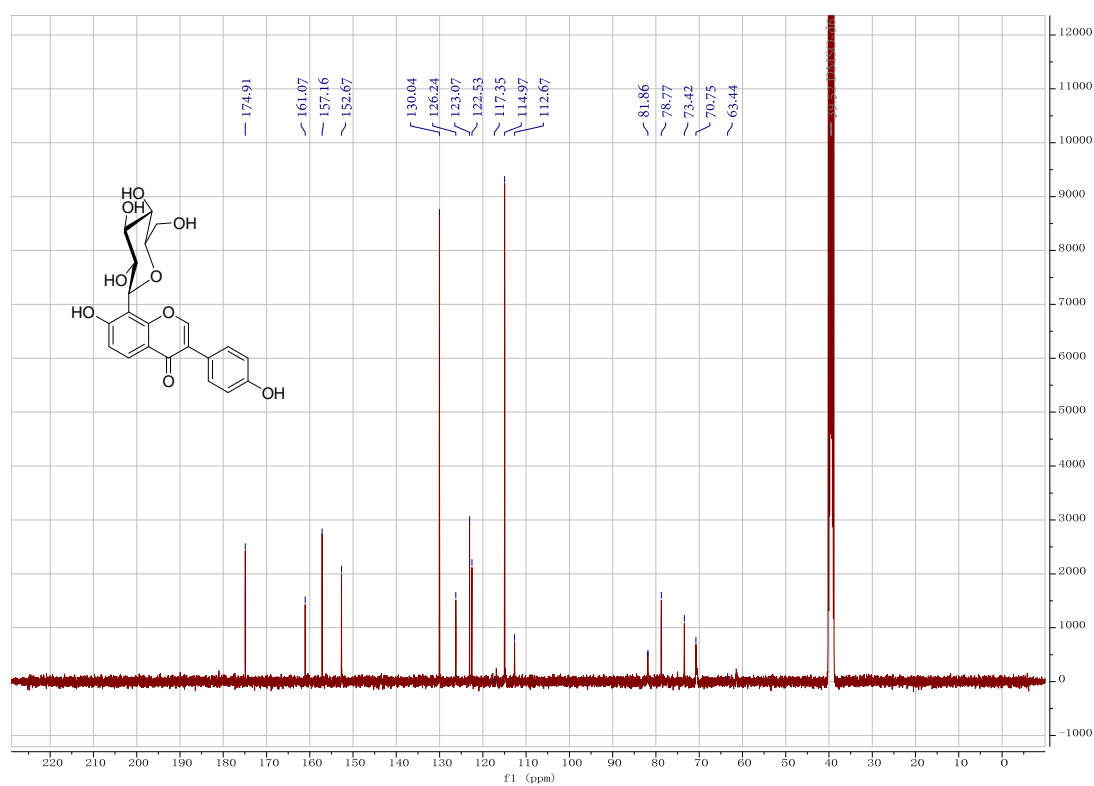
Figure S2. Sequence alignment of PlCGT and PlUGT43.



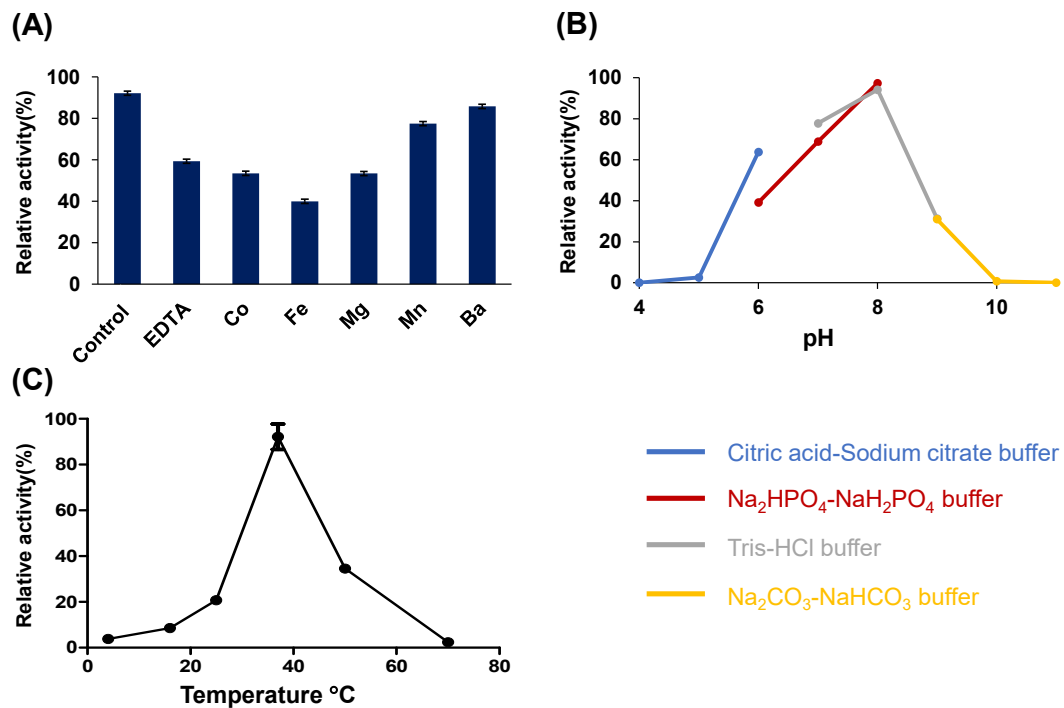
**Figure S3.** SDS-PAGE analysis of purified recombinant PICGT.



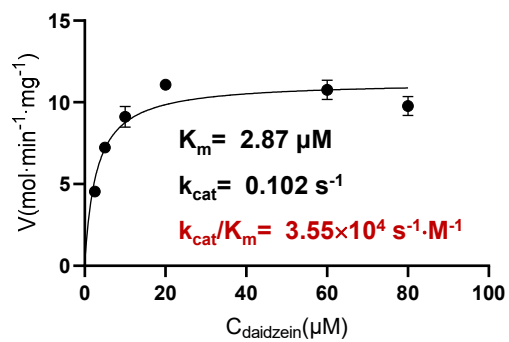
**Figure S4.** The  $^1\text{H}$  NMR spectrum of puerarin (**1a**) in  $\text{DMSO-}d_6$  (400 MHz)



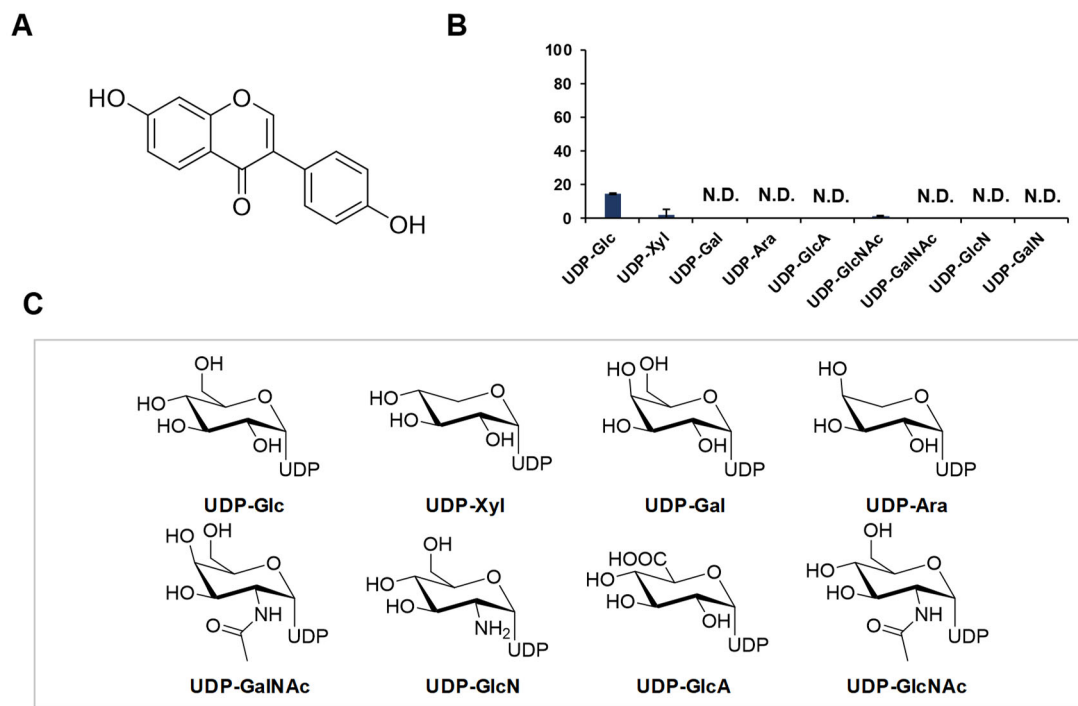
**Figure S5.** The  $^{13}\text{C}$  NMR spectrum of puerarin (**1a**) in  $\text{DMSO-}d_6$  (100 MHz)



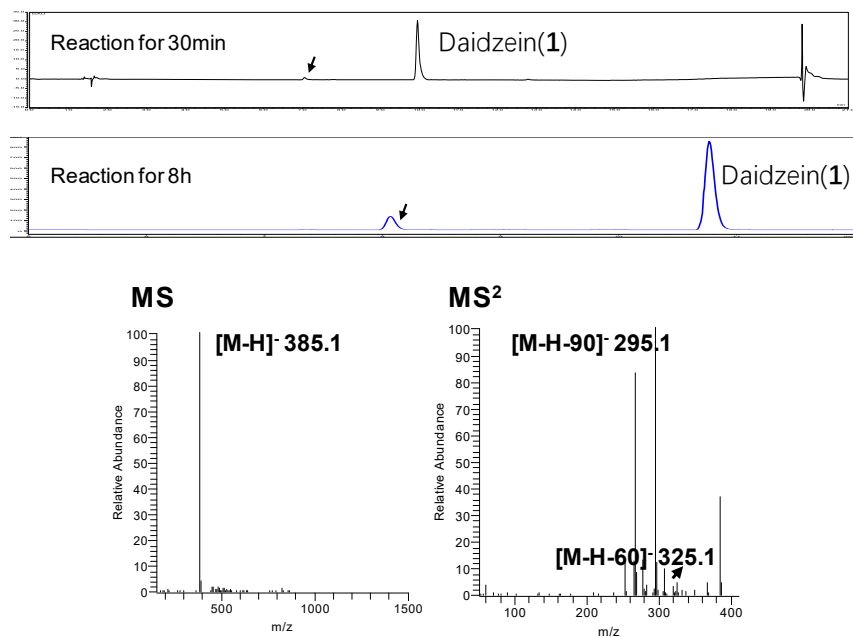
**Figure S6.** Effects of pH (A), divalent metal ions and EDTA (B), temperature (C) on enzyme activity of PICGT. UDP-Glc was used as the sugar donor and daidzein as acceptor. PICGT showed the optimal reaction temperatures at 37 °C, the maximum activity at pH 8.0, and was independent of divalent metal ions or EDTA.



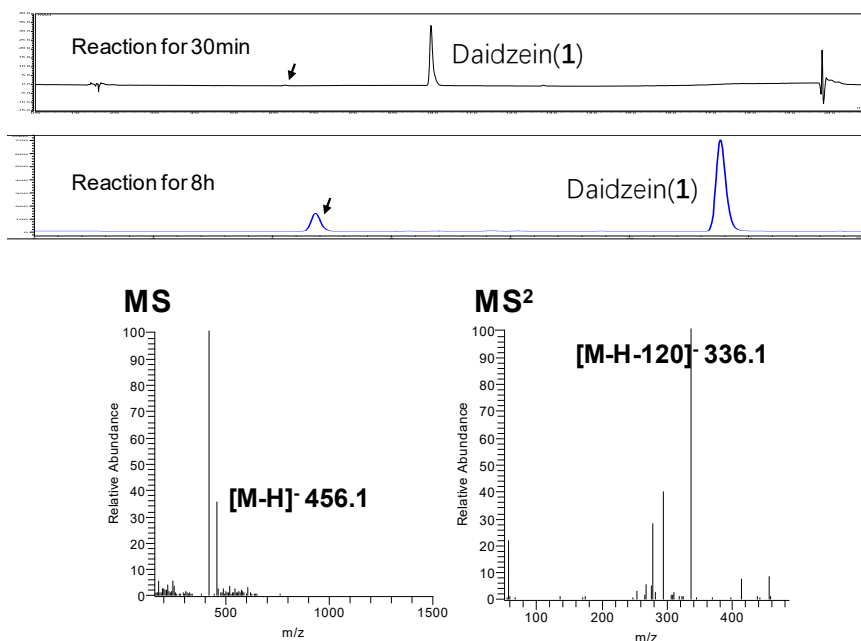
**Figure S7.** Determination of kinetic parameters for PICGT. Michaelis-Menten plot was fitted. The  $K_m$  value was calculated as  $2.87 \mu\text{M}$  using daidzein (**1**) as acceptor and UDP-Glc as donor.



**Figure S8.** The sugar selectivity of PICGT. (A) The structure of daidzein (**1**). (B) Conversion rates of **1** to its glucosides using different sugar donors. The reaction mixtures were incubated at 37 °C for 30min. (C) The structures of different sugar donors.

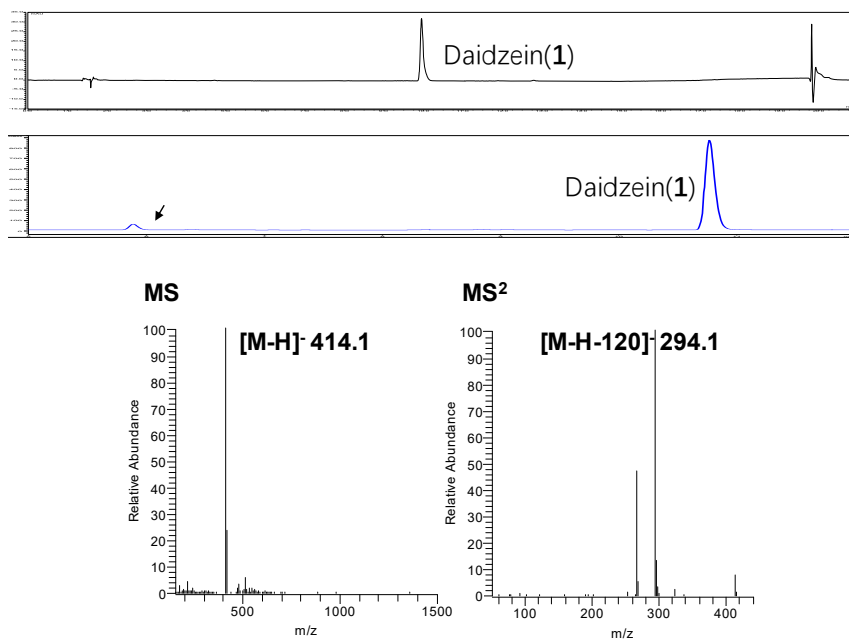


**Figure S9.** LC/MS analysis of PICGT catalyzed products using UDP-Xyl as the sugar donor.



**Figure S10.** LC/MS analysis of PICGT catalyzed products using UDP-GlcNAc as the sugar donor.





**Figure S11.** LC/MS analysis of PICGT catalyzed products using UDP-GlcN as the sugar donor.

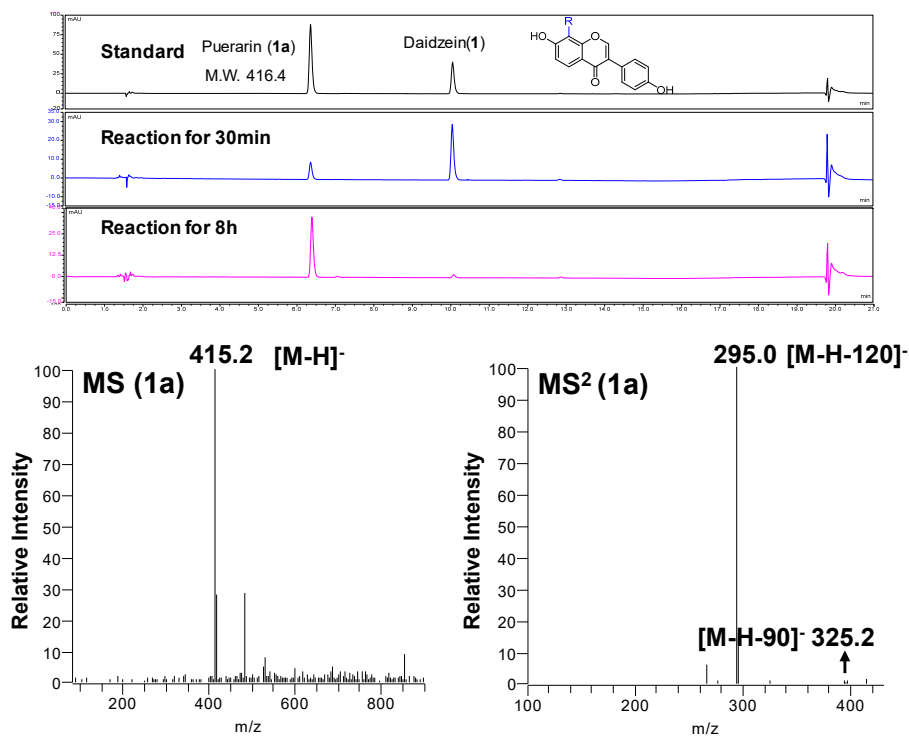


Figure S12. LC/MS analysis of PICGT catalyzed products using **1** as the substrate.

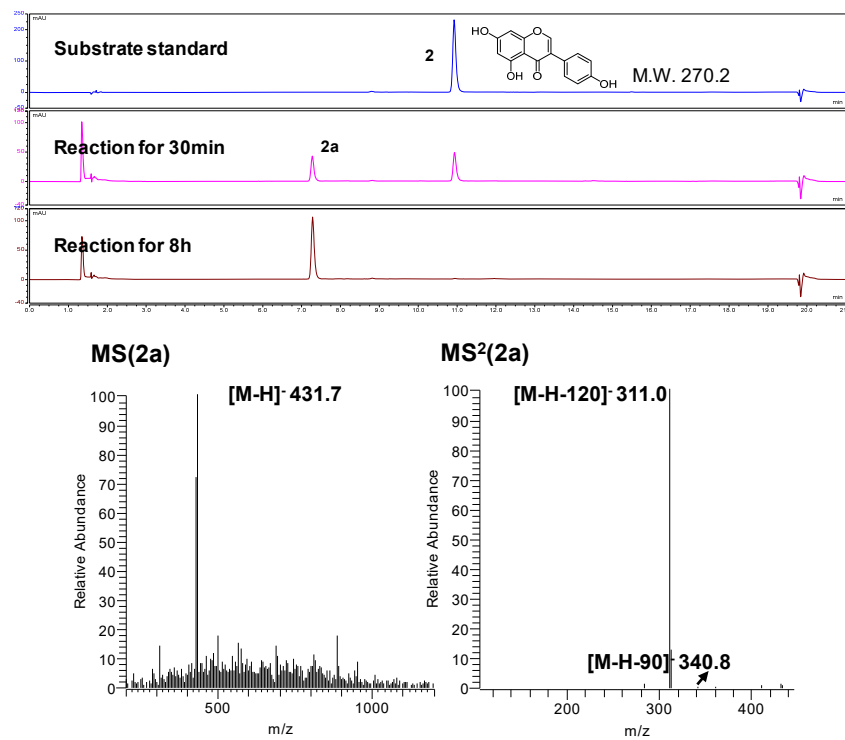


Figure S13. LC/MS analysis of PICGT catalyzed products using **2** as the substrate.

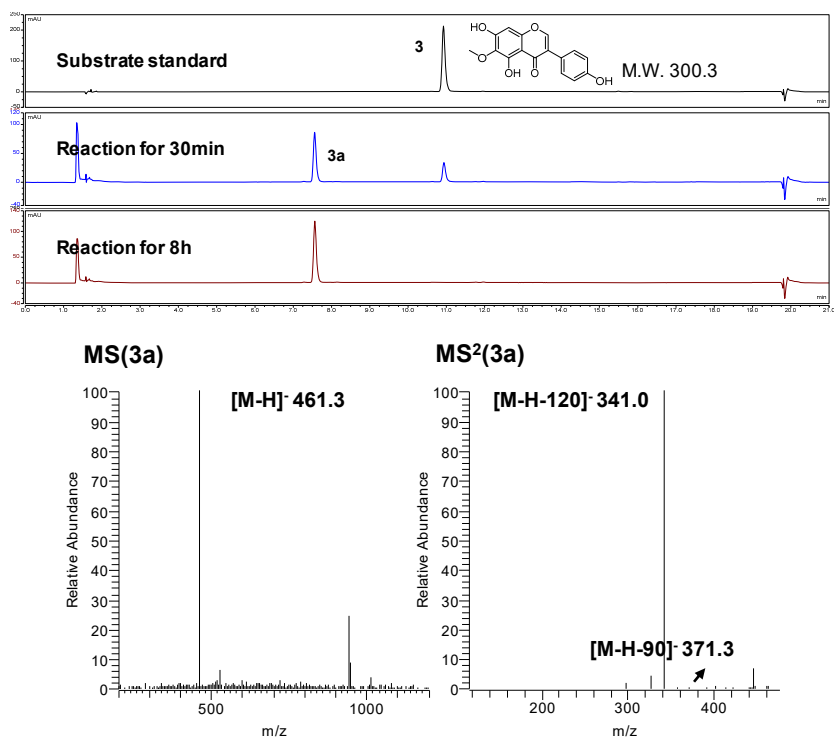


Figure S14. LC/MS analysis of PICGT catalyzed products using **3** as the substrate.

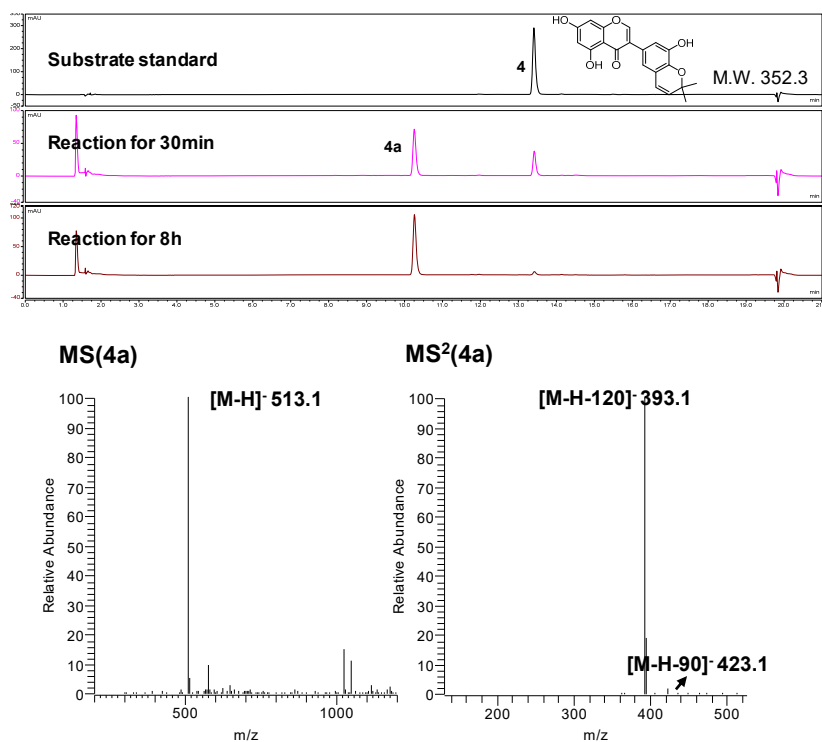


Figure S15. LC/MS analysis of PICGT catalyzed products using **4** as the substrate.

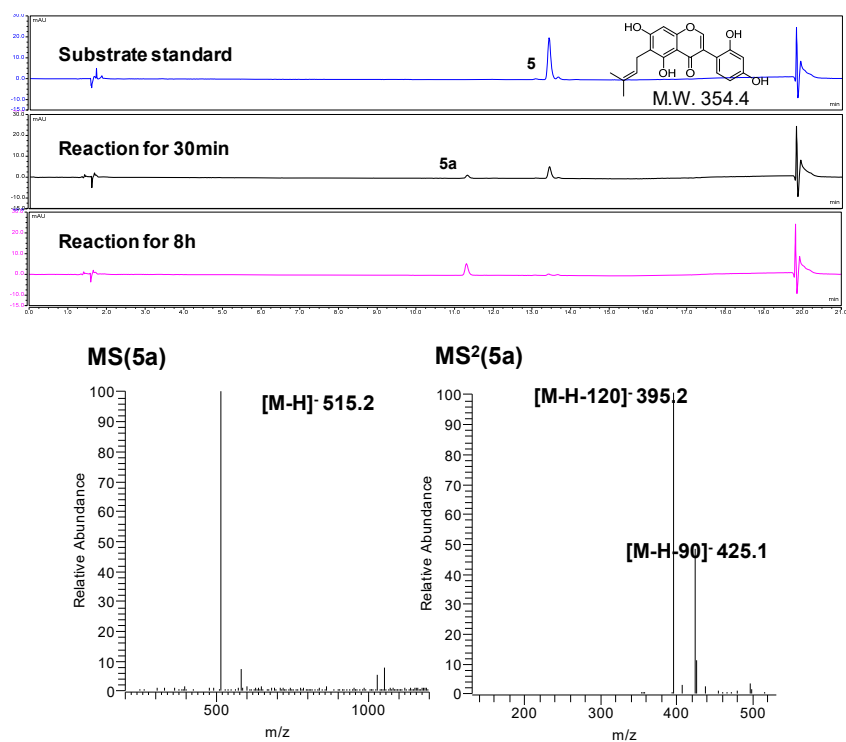


Figure S16. LC/MS analysis of PICGT catalyzed products using **5** as the substrate.

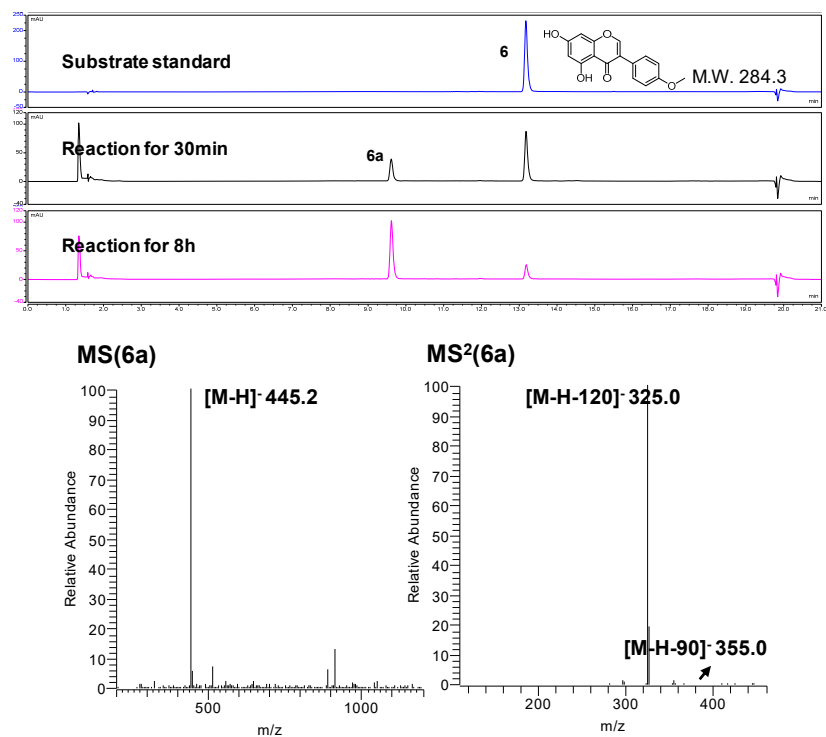


Figure S17. LC/MS analysis of PICGT catalyzed products using **6** as the substrate.

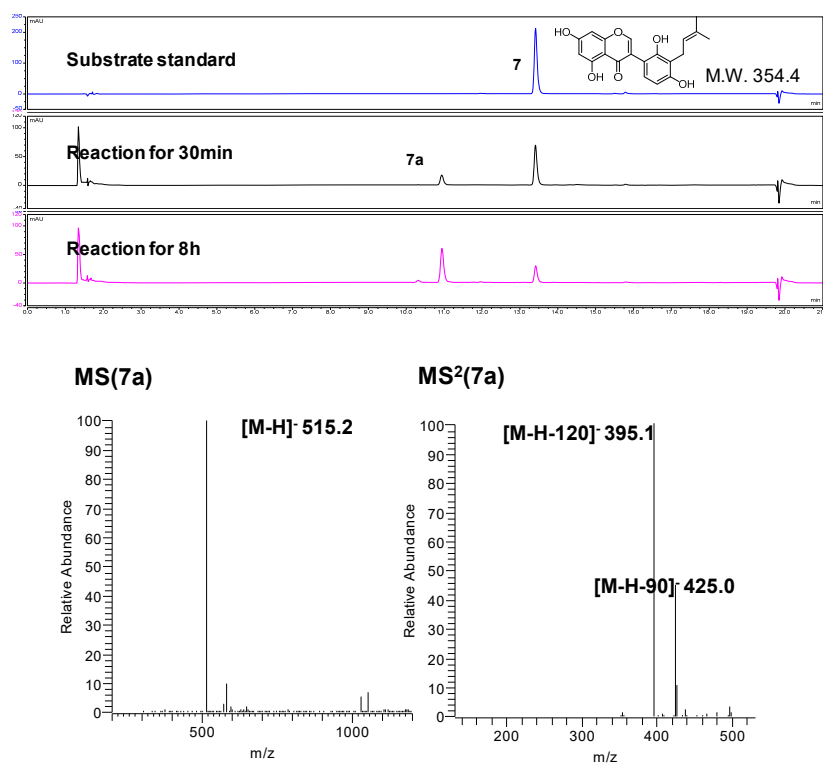


Figure S18. LC/MS analysis of PICGT catalyzed products using **7** as the substrate.

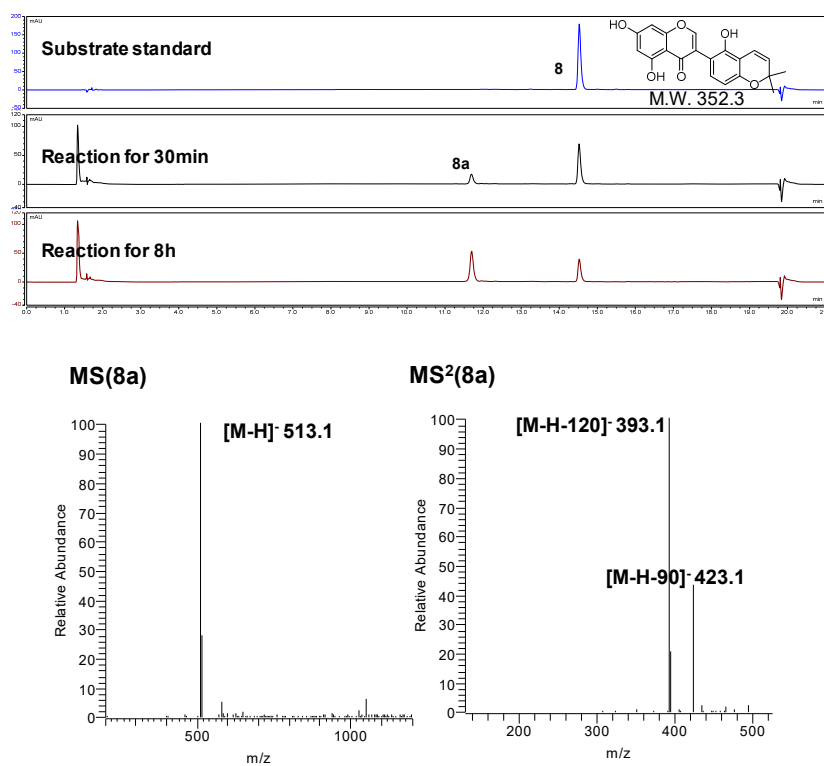
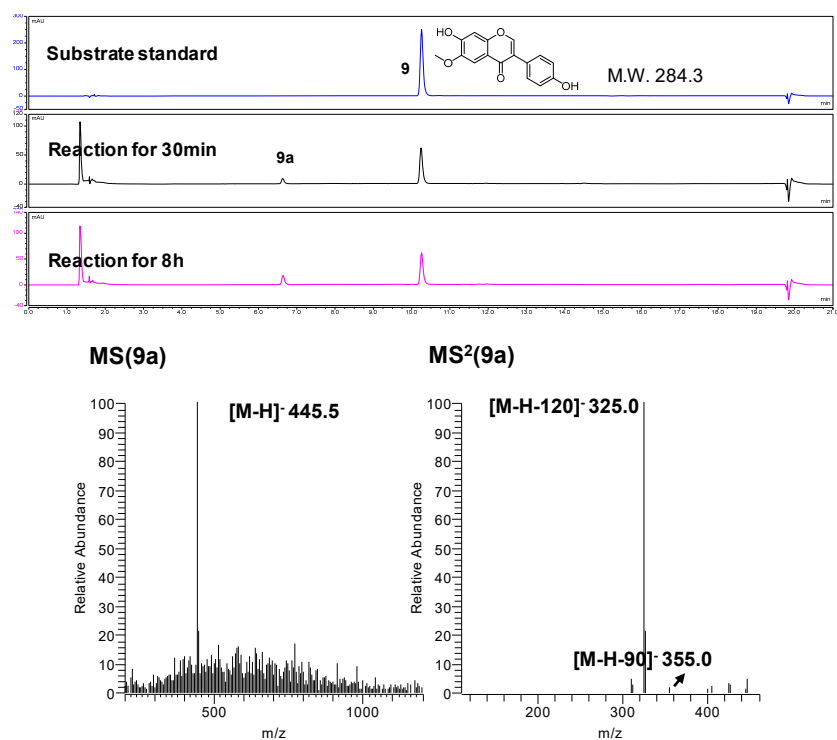
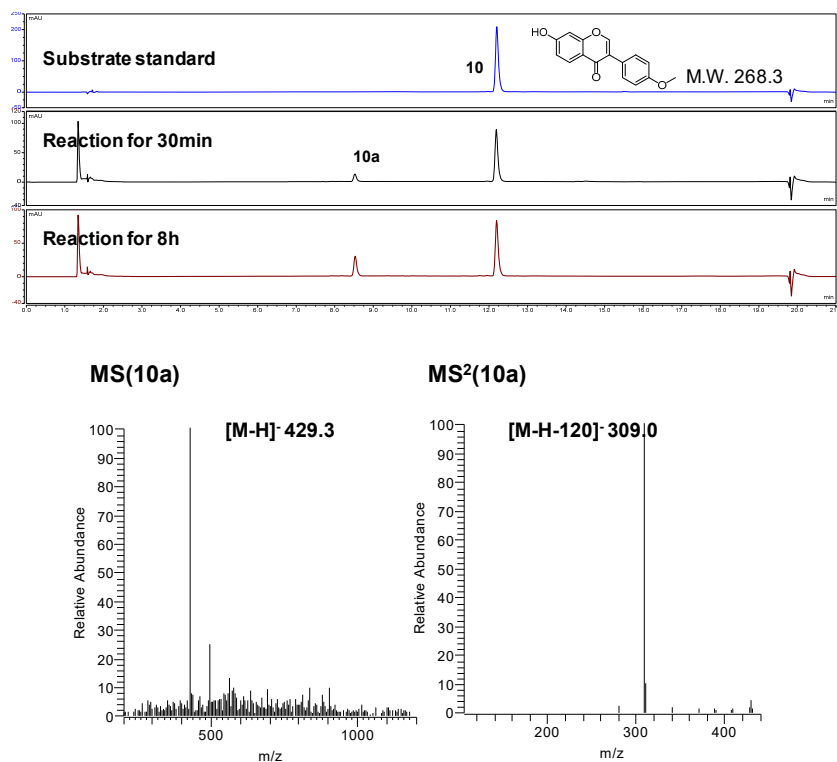


Figure S19. LC/MS analysis of PICGT catalyzed products using **8** as the substrate.



**Figure S20.** LC/MS analysis of PICGT catalyzed products using **9** as the substrate.



**Figure S21.** LC/MS analysis of PICGT catalyzed products using **10** as the substrate.

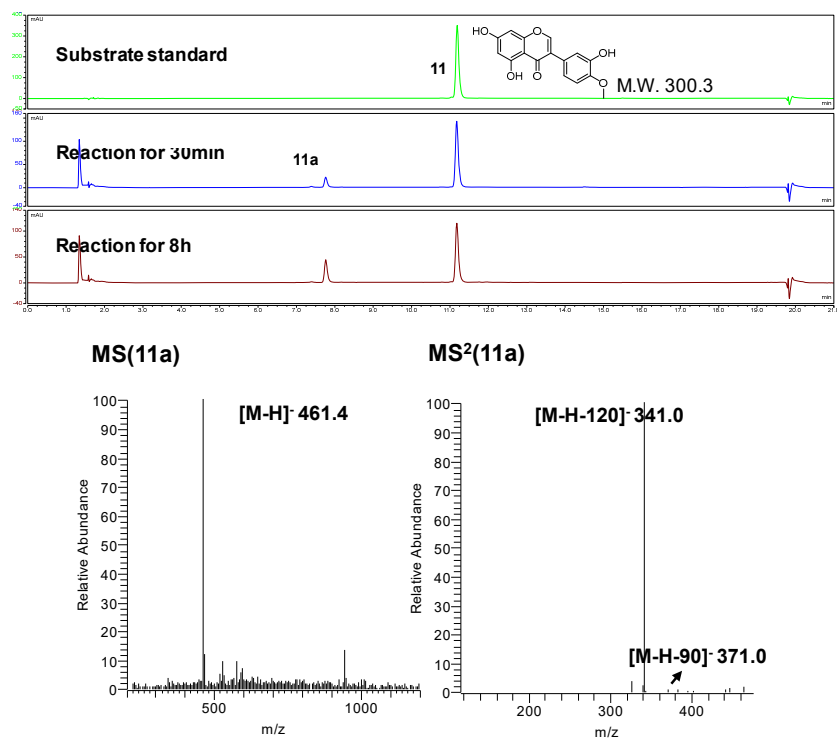


Figure S22. LC/MS analysis of PICGT catalyzed products using **11** as the substrate.

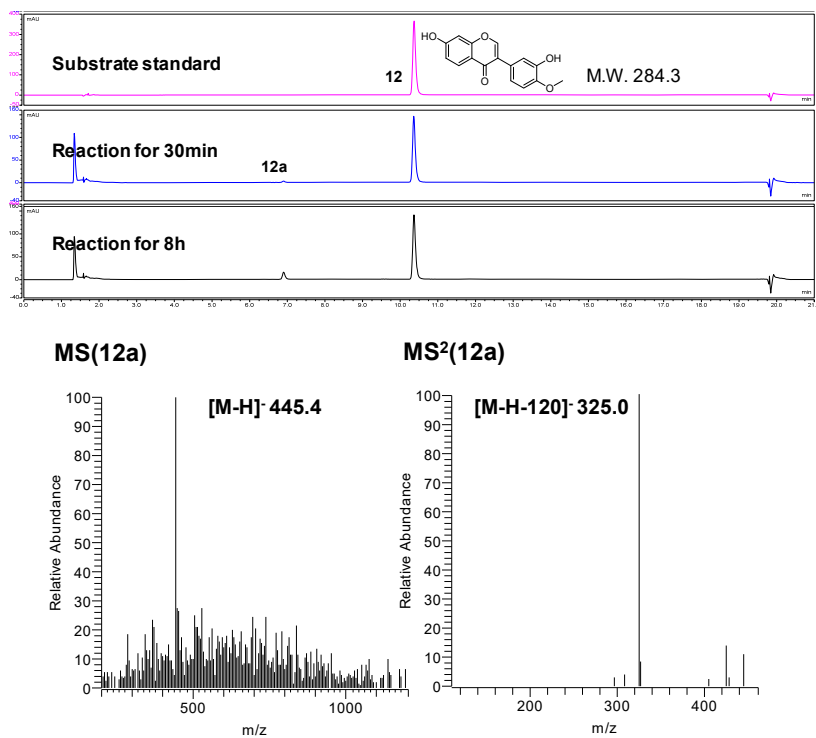
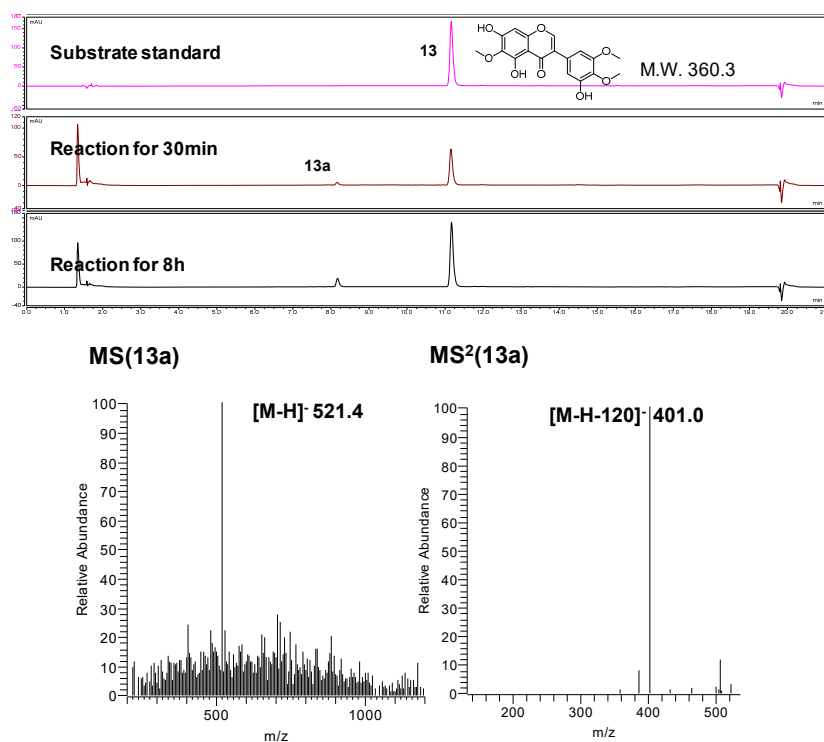
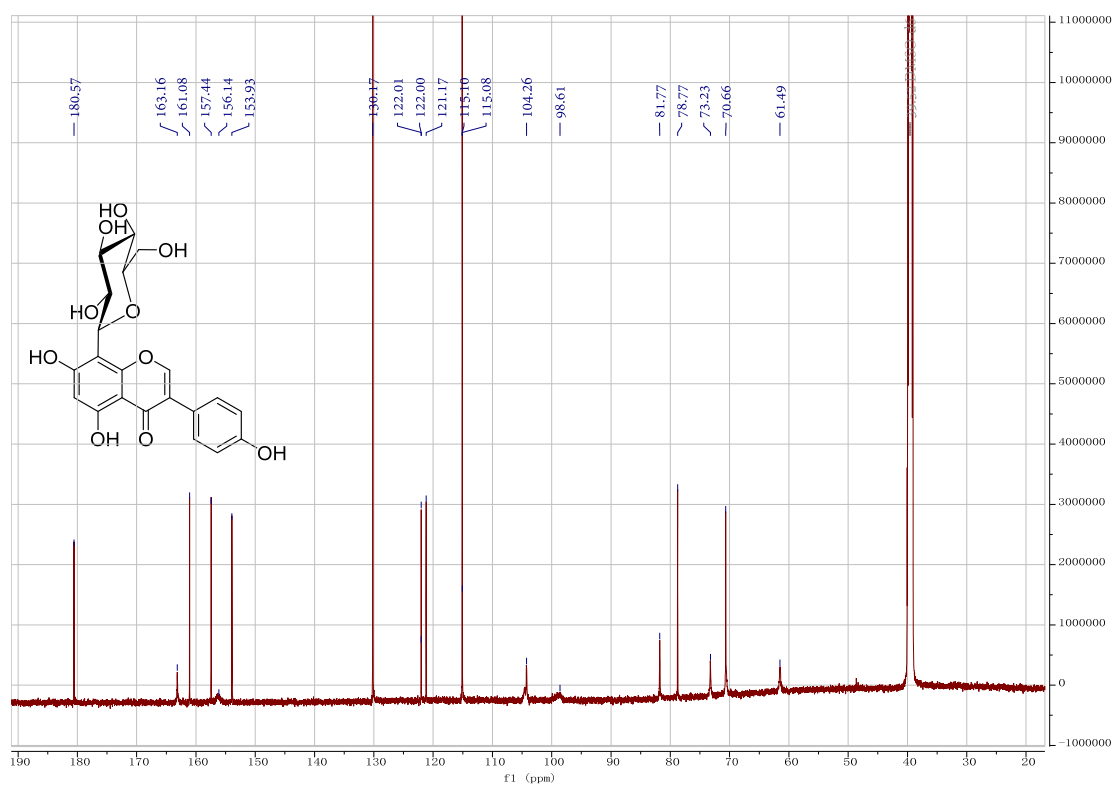
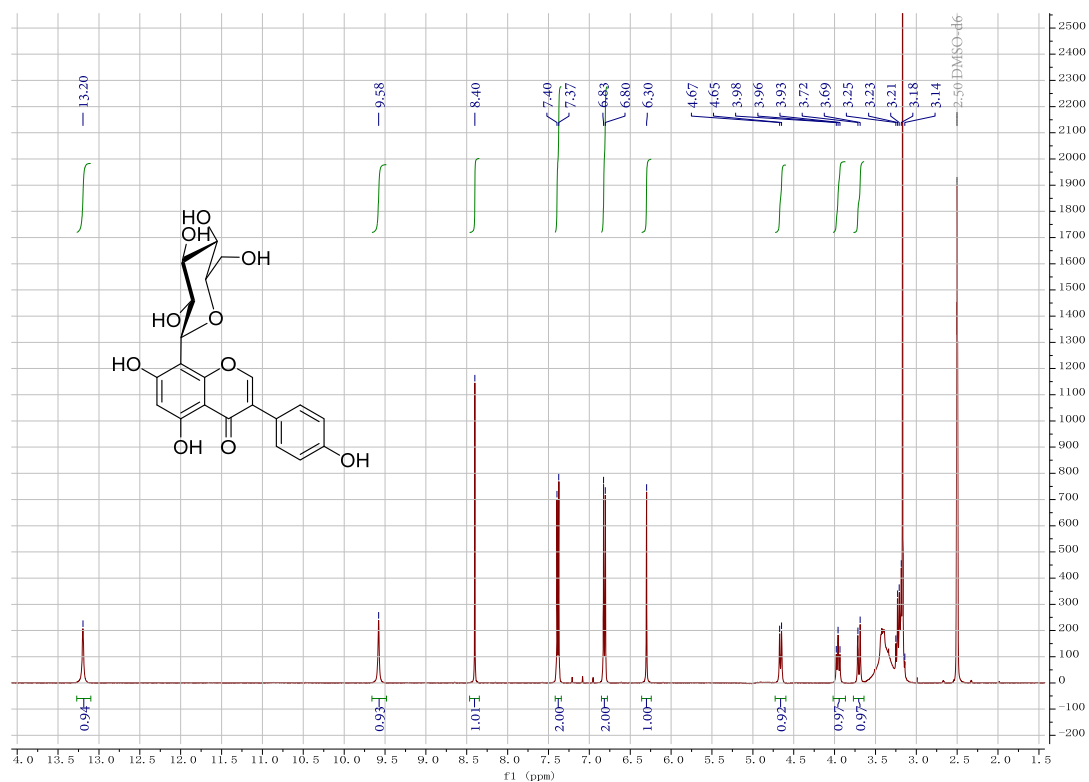


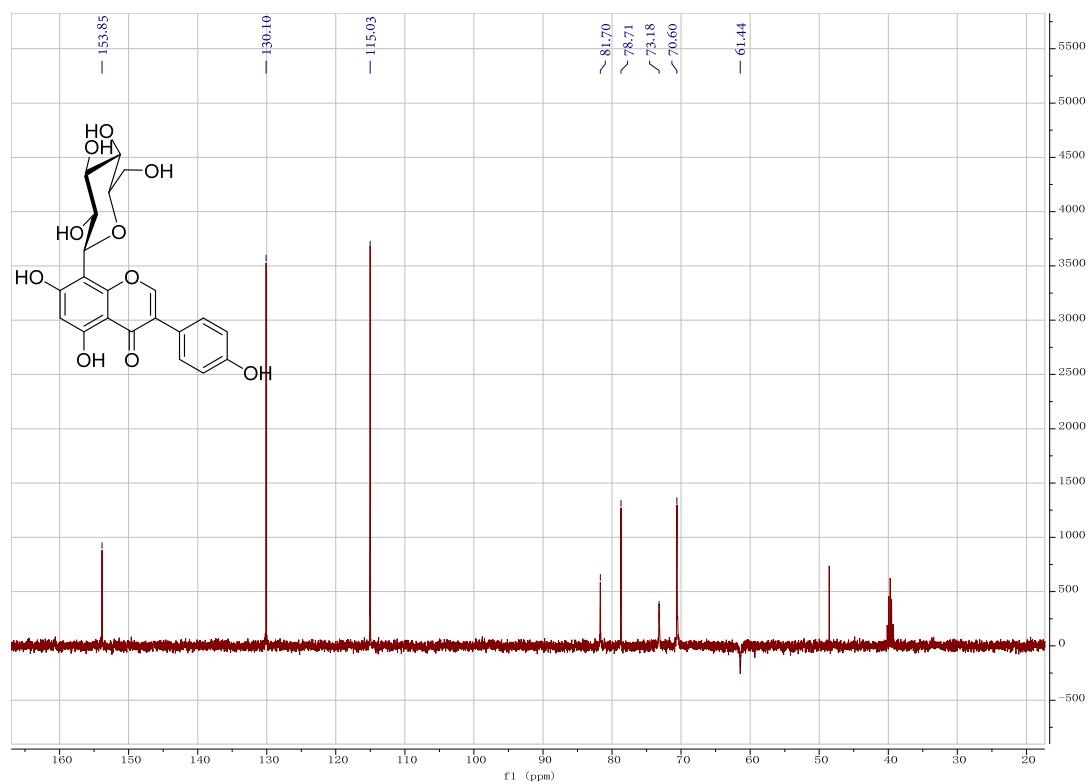
Figure S23. LC/MS analysis of PICGT catalyzed products using **12** as the substrate.



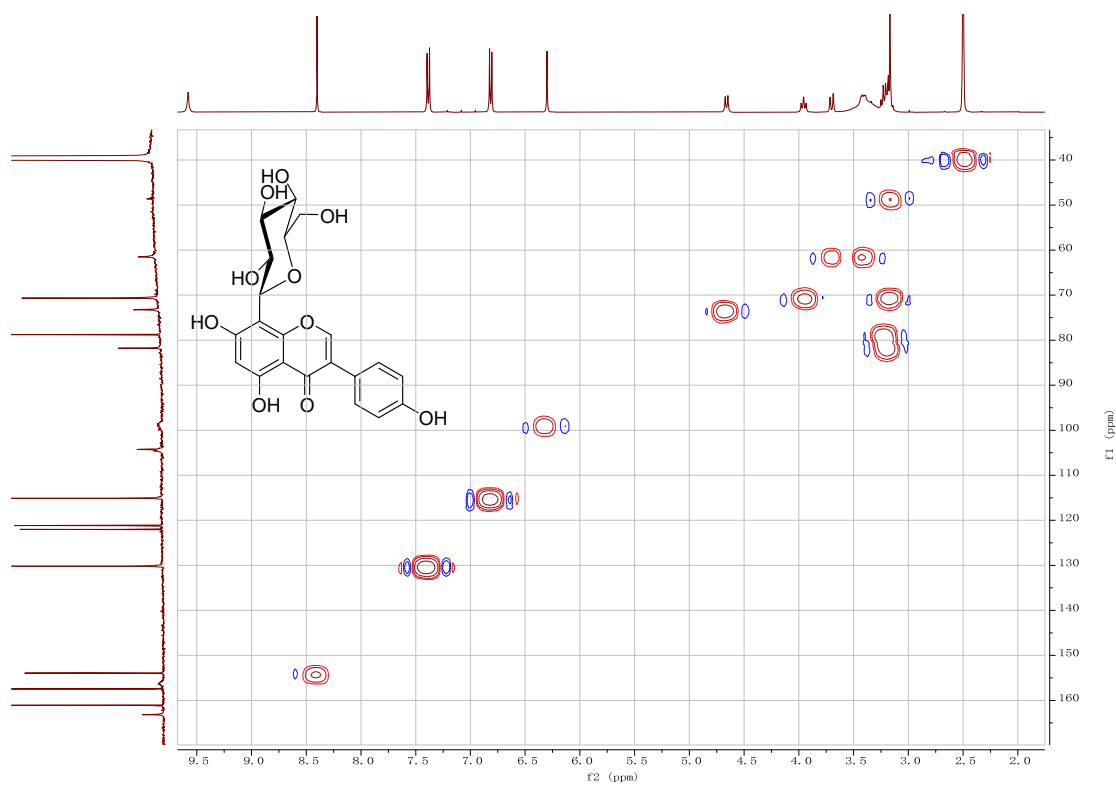
**Figure S24.** LC/MS analysis of PICGT catalyzed products using **13** as the substrate.



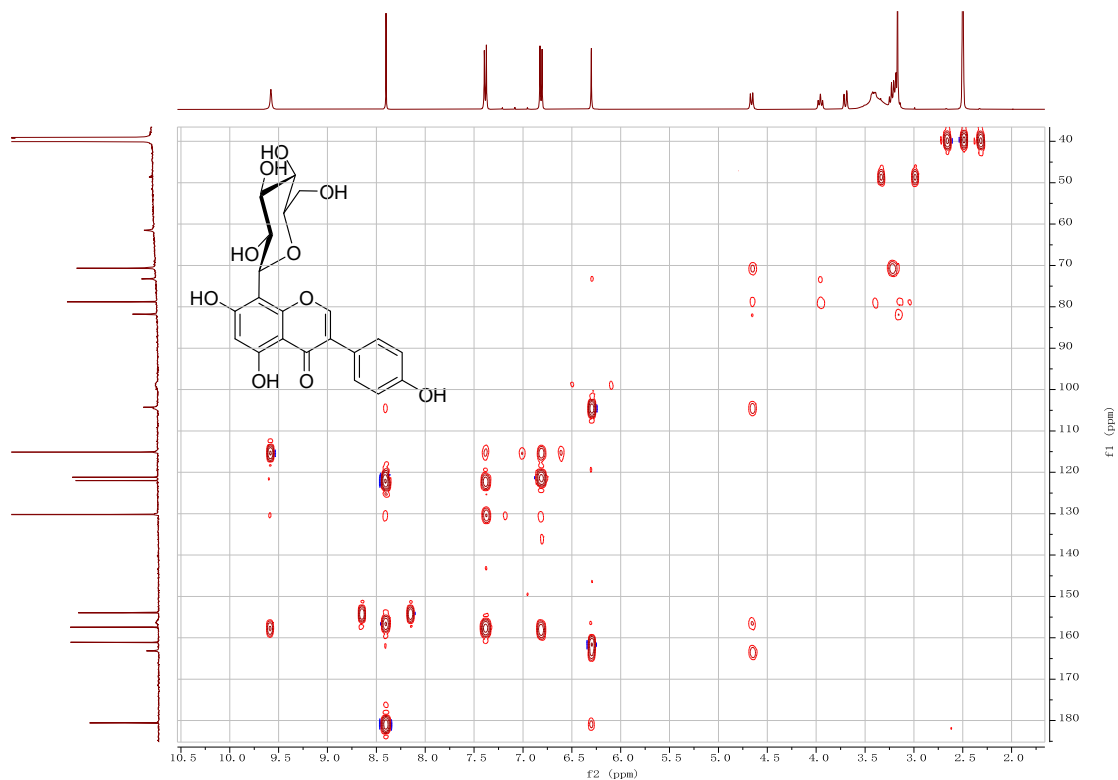




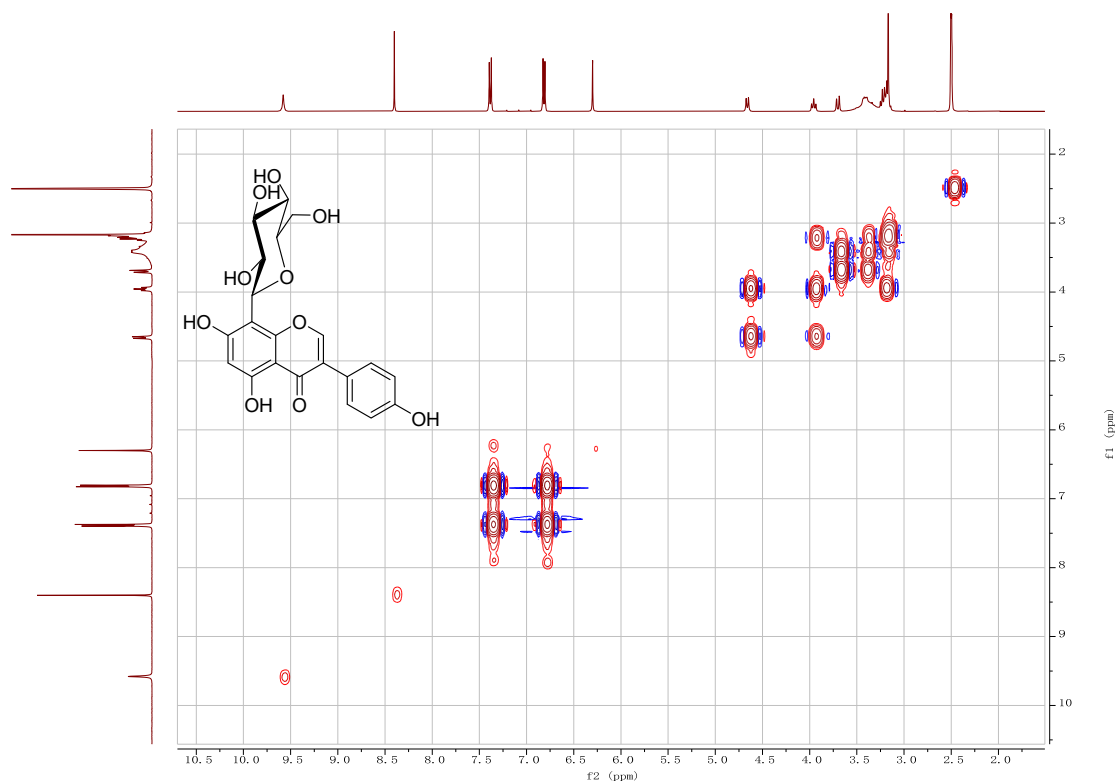
**Figure S27.** The DEPT135 spectrum of **2a** in DMSO-*d*<sub>6</sub> (100 MHz)



**Figure S28.** The HSQC spectrum of **2a** in DMSO-*d*<sub>6</sub> (400 MHz)

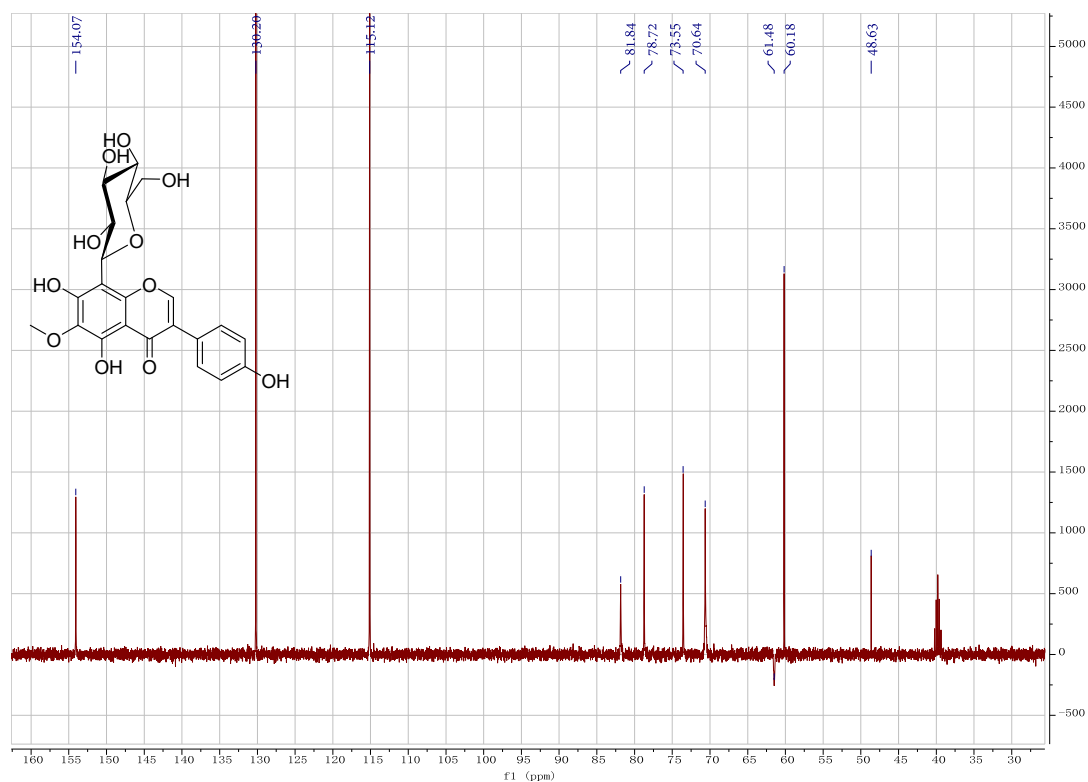


**Figure S29.** The HMBC spectrum of **2a** in DMSO-*d*<sub>6</sub> (400 MHz)

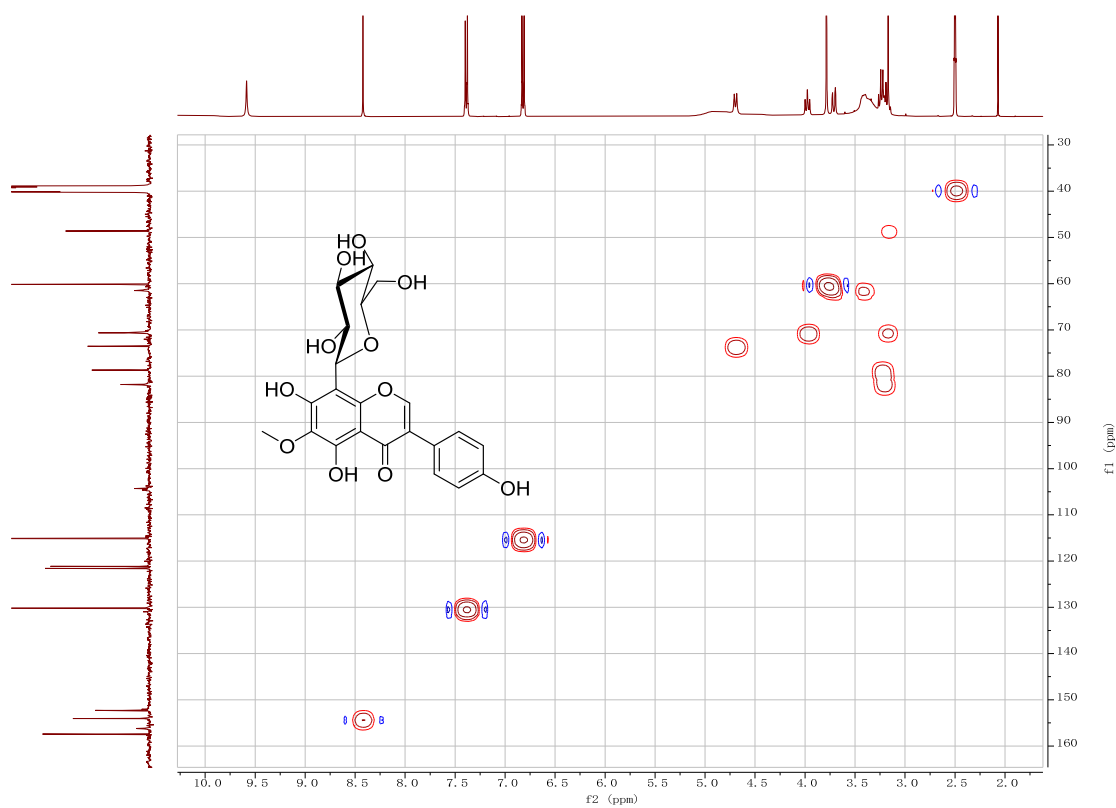


**Figure S30.** The <sup>1</sup>H-<sup>1</sup>H COSY spectrum of **2a** in DMSO-*d*<sub>6</sub> (400 MHz)





**Figure S33.** The DEPT135 spectrum of **3a** in DMSO-*d*<sub>6</sub> (100 MHz)



**Figure S34.** The HSQC spectrum of **3a** in DMSO-*d*<sub>6</sub> (400 MHz)

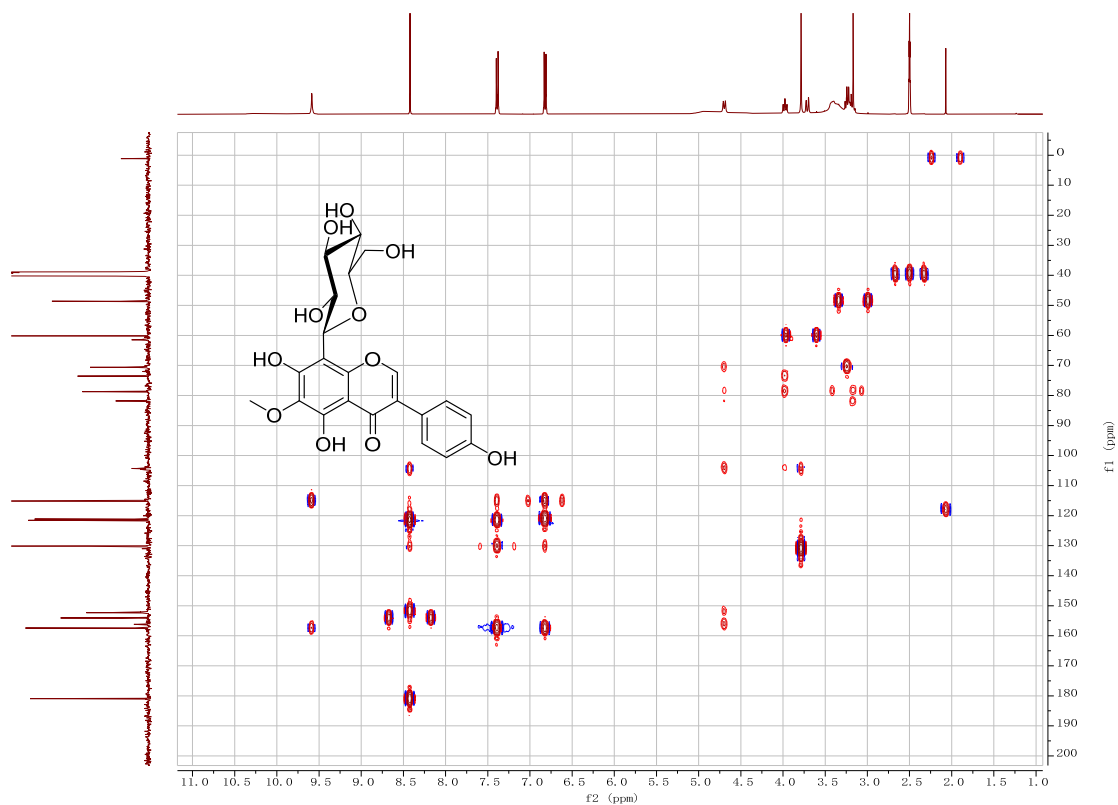


Figure S35. The HMBC spectrum of **3a** in DMSO-*d*<sub>6</sub> (400 MHz)

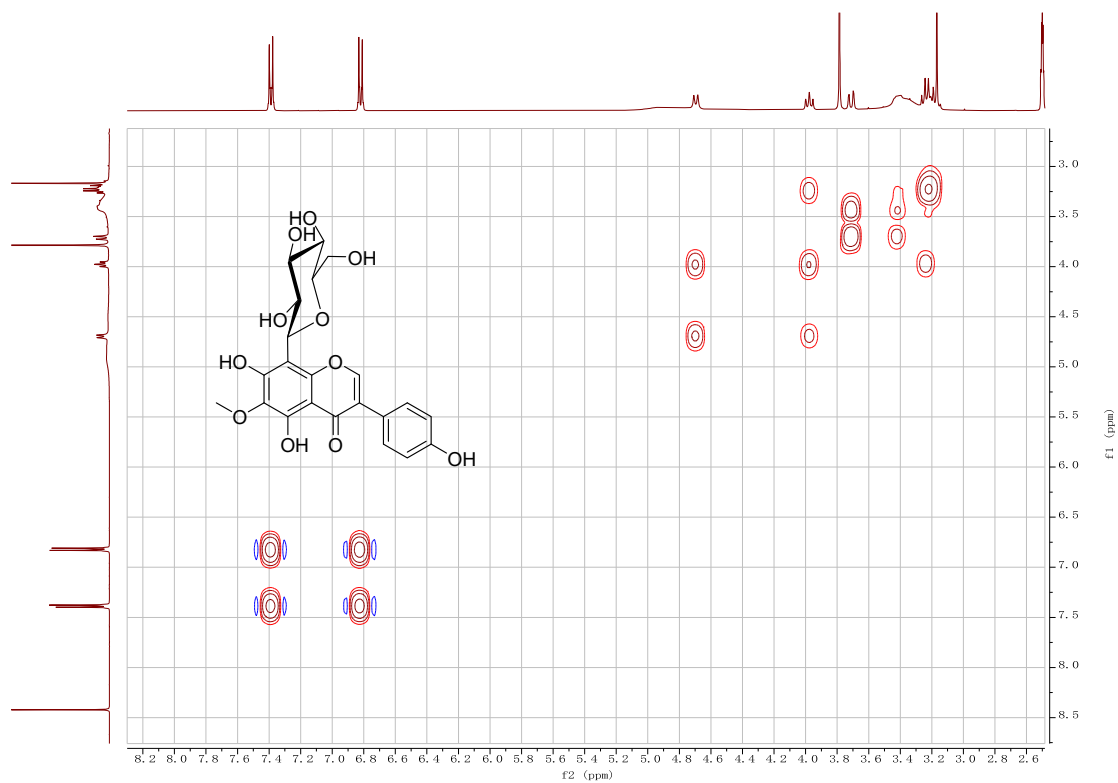
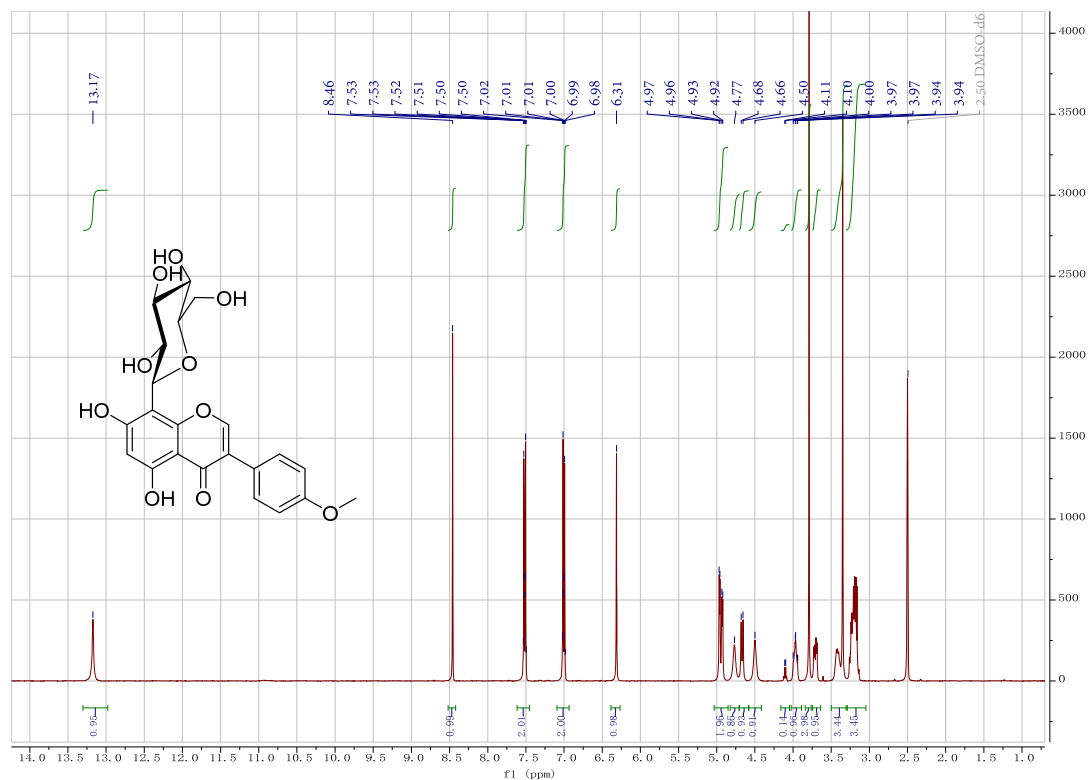
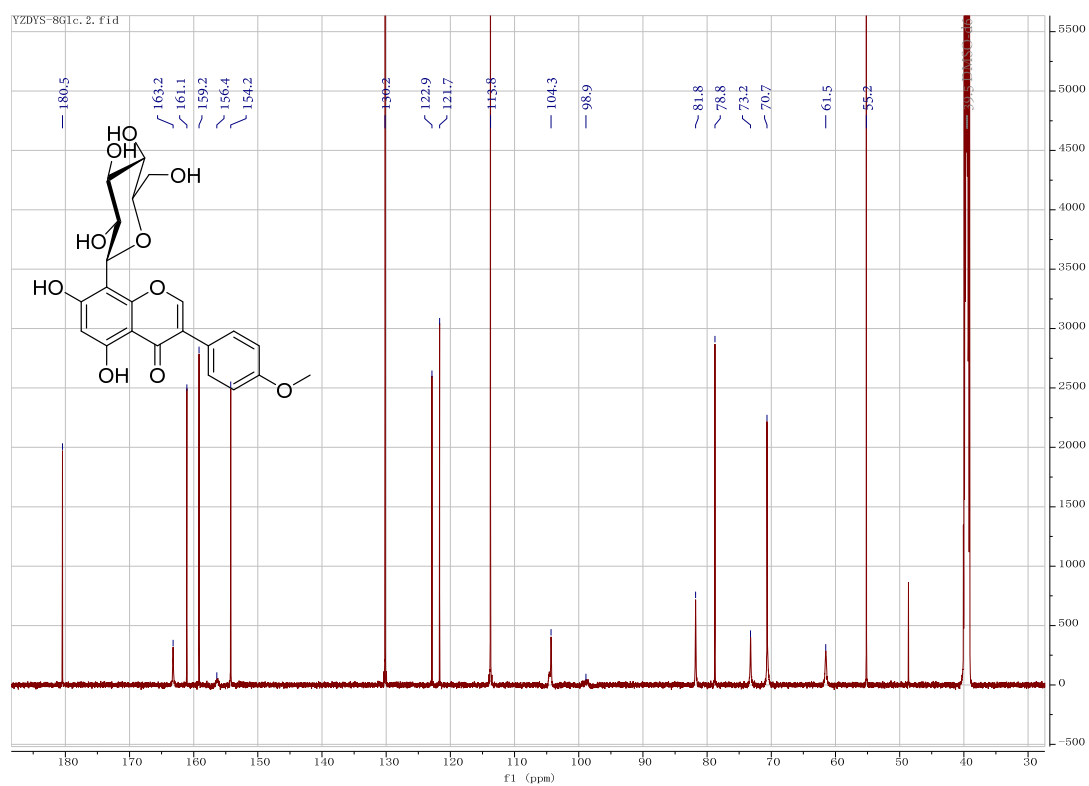


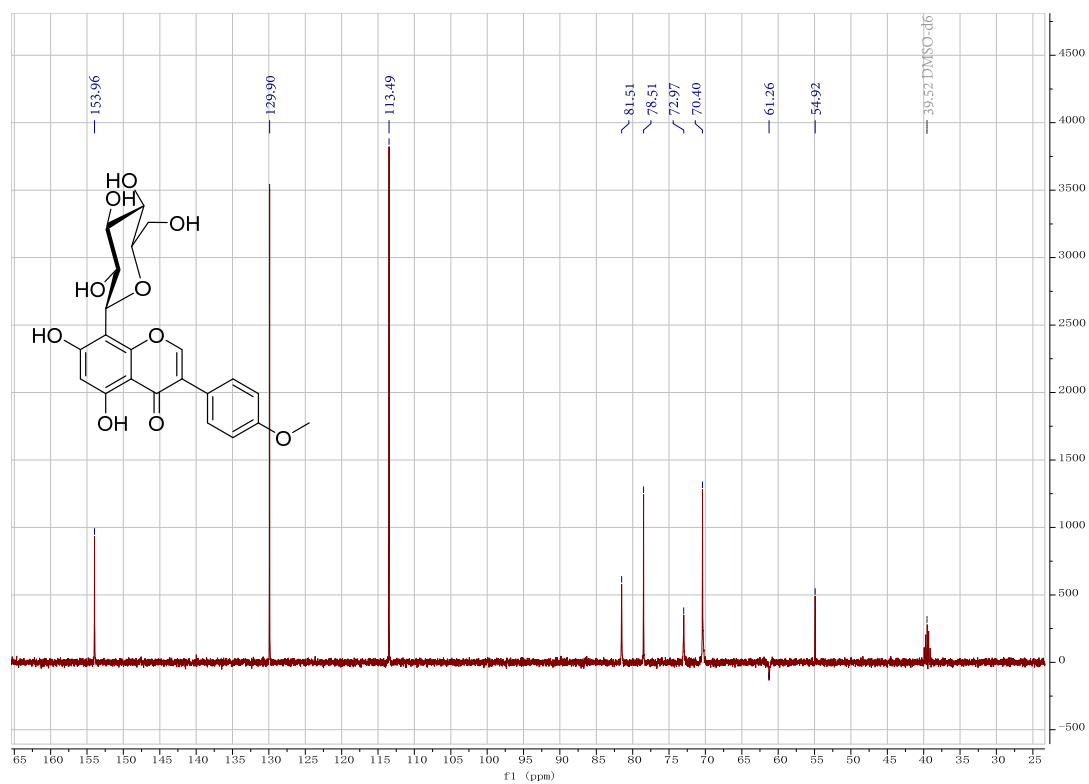
Figure S36. The <sup>1</sup>H-<sup>1</sup>H COSY spectrum of **3a** in DMSO-*d*<sub>6</sub> (400 MHz)



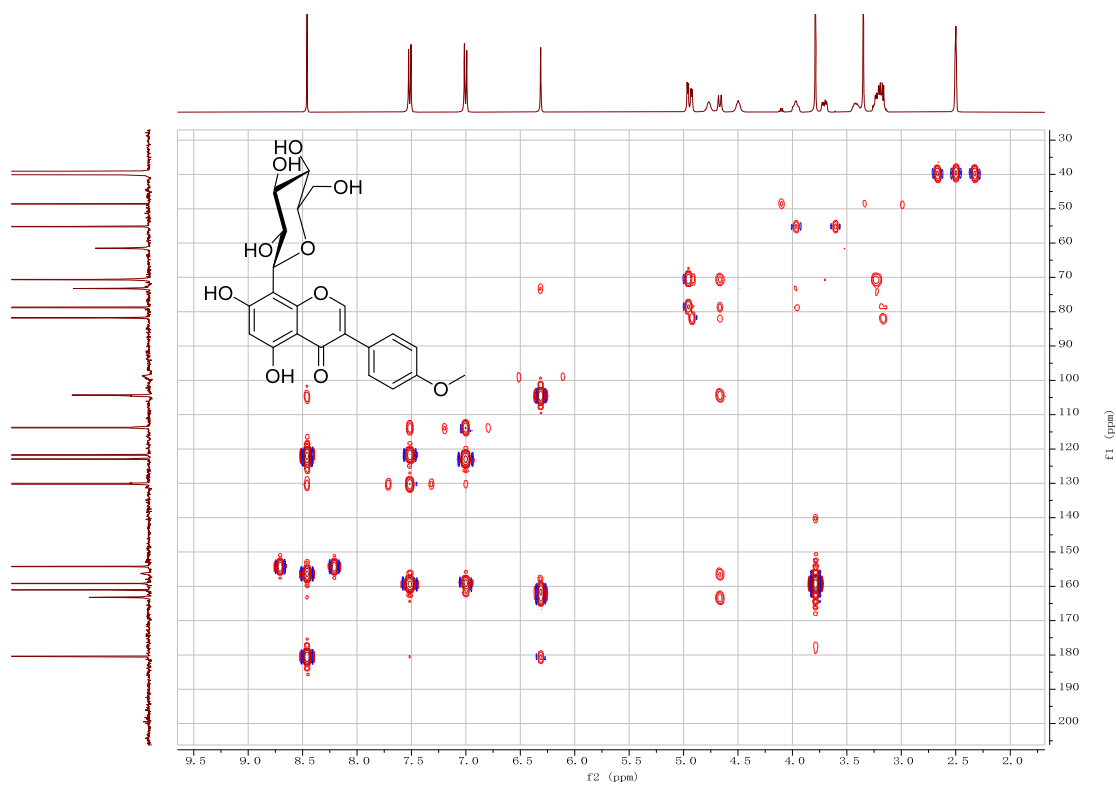
**Figure S37.** The <sup>1</sup>H NMR spectrum of **6a** in DMSO-*d*<sub>6</sub> (400 MHz)



**Figure S38.** The <sup>13</sup>C NMR spectrum of **6a** in DMSO-*d*<sub>6</sub> (150 MHz)

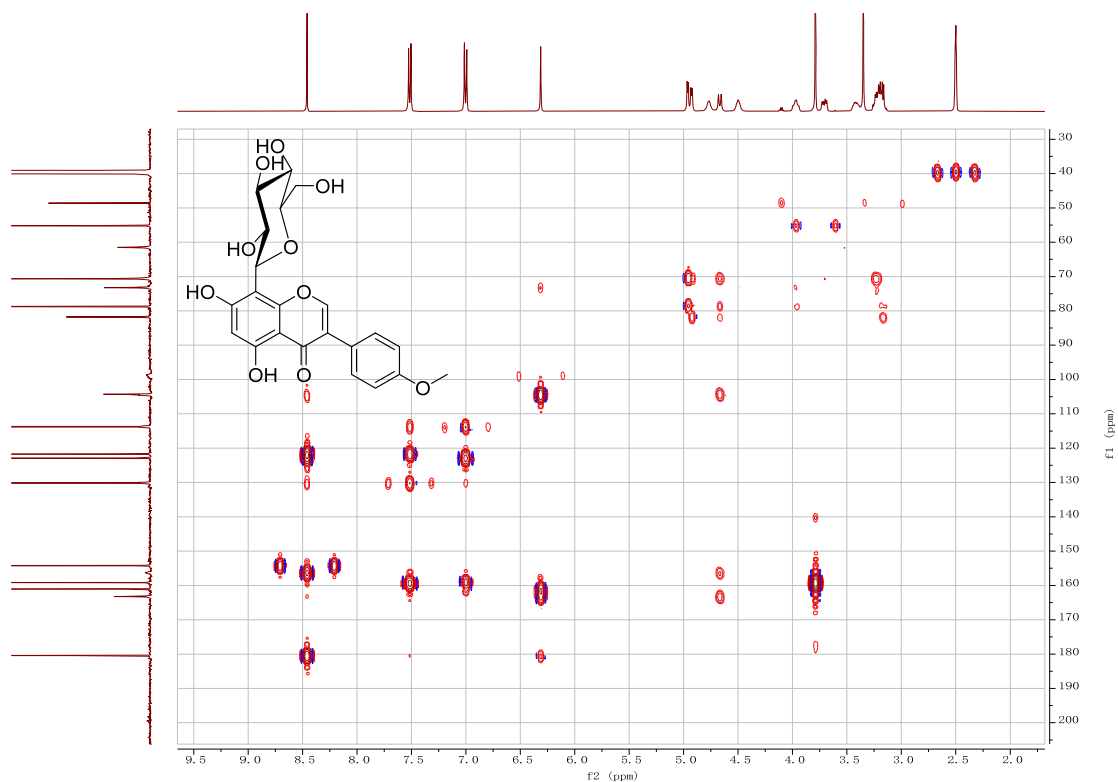


**Figure S39.** The DEPT135 spectrum of **6a** in DMSO-*d*<sub>6</sub> (100 MHz)

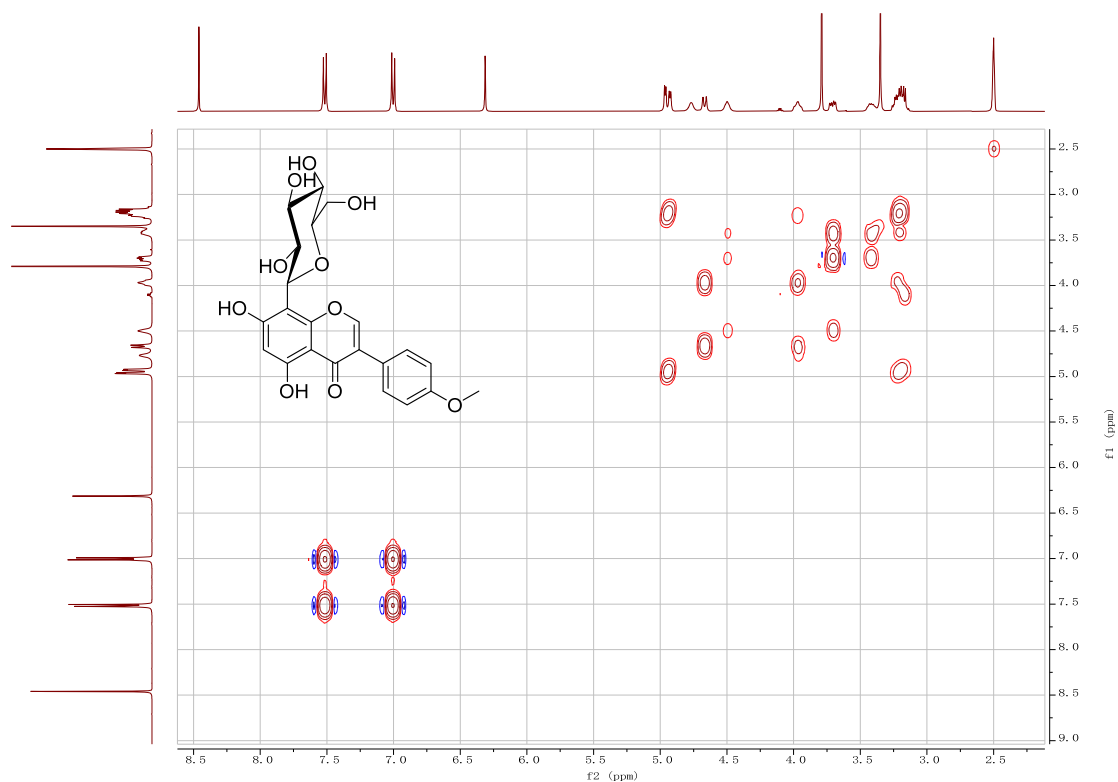


**Figure S40.** The HSQC spectrum of **6a** in DMSO-*d*<sub>6</sub> (400 MHz)

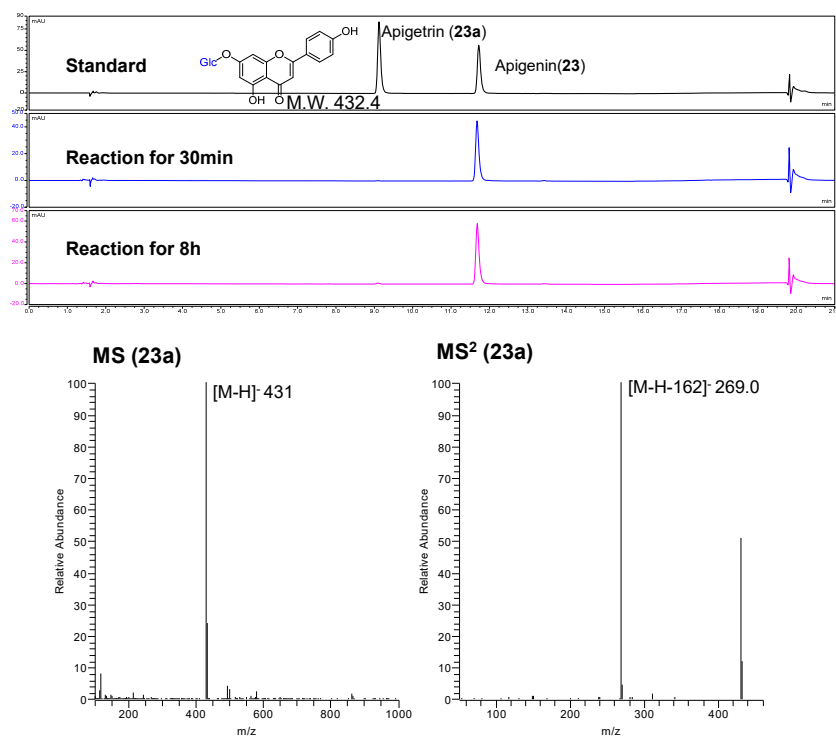




**Figure S41.** The HMBC spectrum of **6a** in DMSO-*d*<sub>6</sub> (400 MHz)



**Figure S42.** The <sup>1</sup>H-<sup>1</sup>H COSY spectrum of **6a** in DMSO-*d*<sub>6</sub> (400 MHz)



**Figure S43.** LC/MS analysis of PICGT catalyzed products using **23** as the substrate.

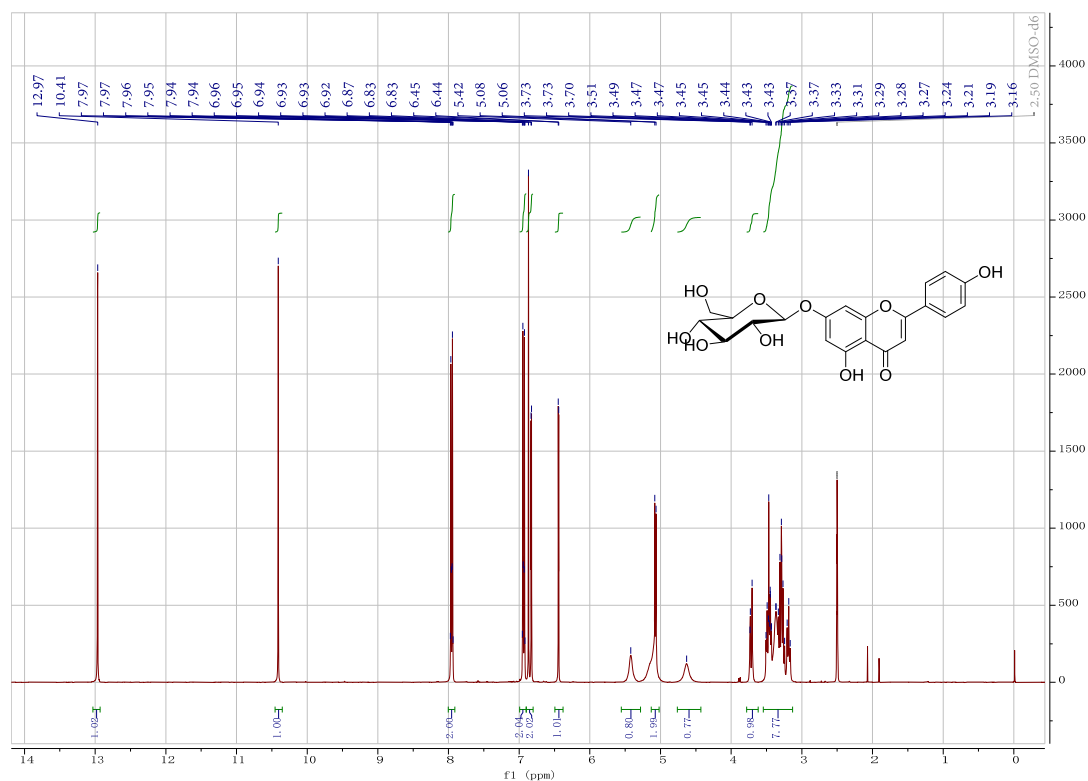


Figure S44. The  $^1\text{H}$  NMR spectrum of Apigetrin (**23a**) in  $\text{DMSO-}d_6$  (400 MHz)

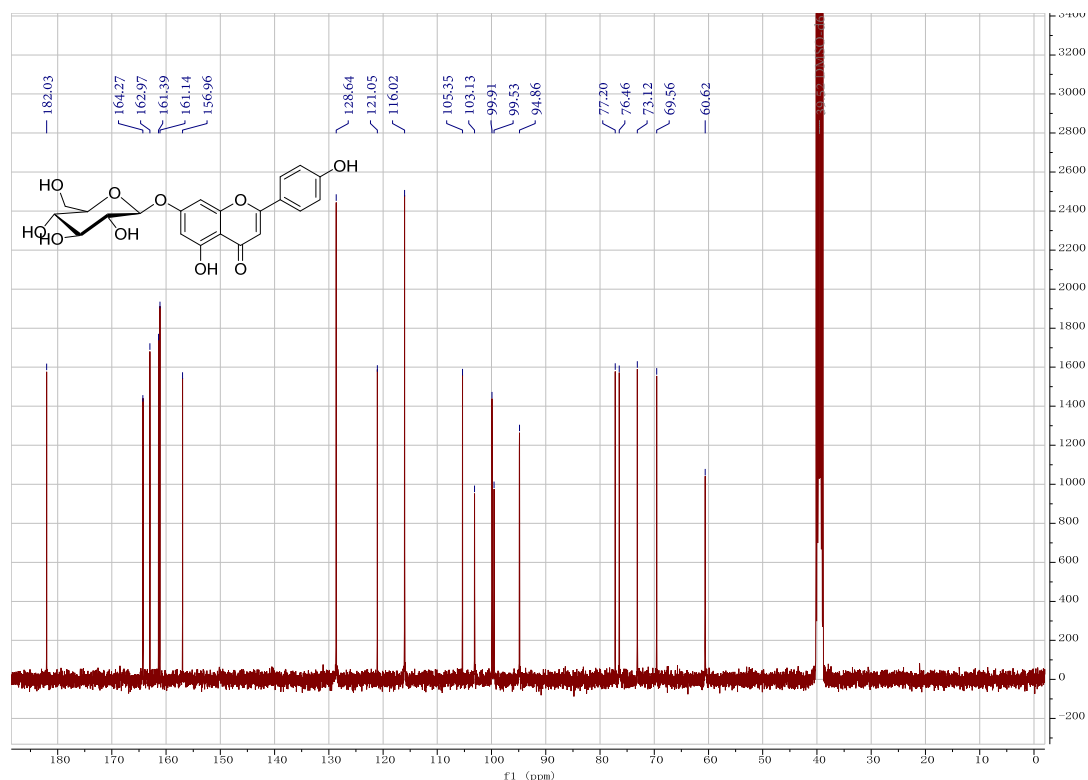
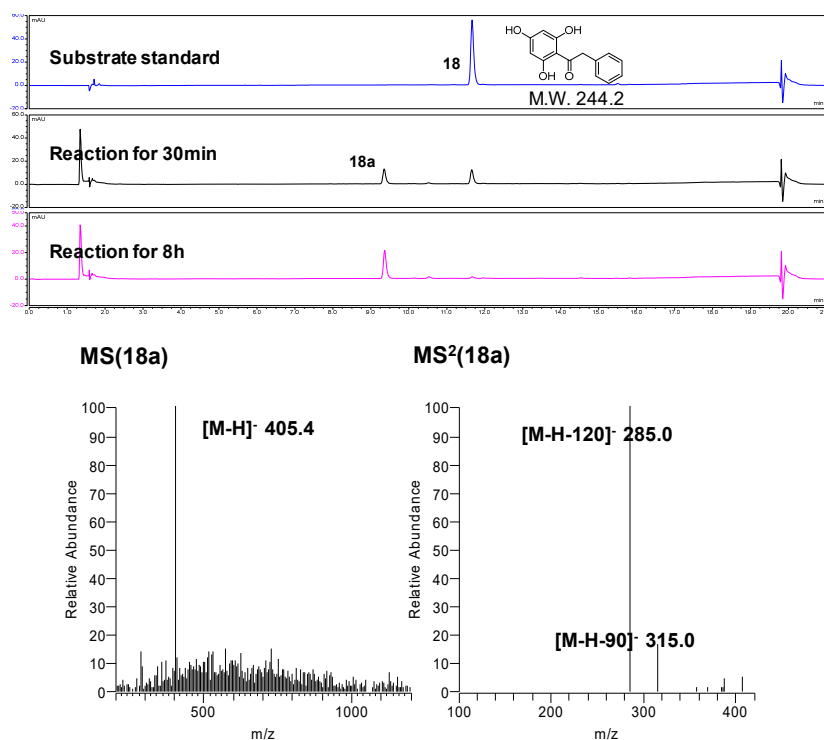
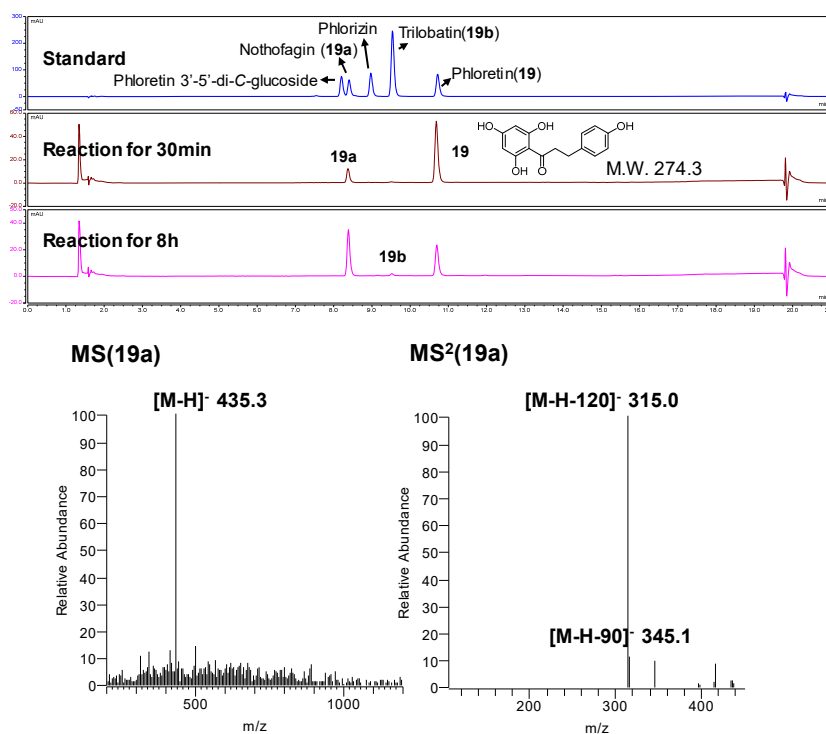


Figure S45. The  $^{13}\text{C}$  NMR spectrum of Apigetrin (**23a**) in  $\text{DMSO-}d_6$  (100 MHz)



**Figure S46.** LC/MS analysis of PICGT catalyzed products using **18** as the substrate.



**Figure S47.** LC/MS analysis of PICGT catalyzed products using **19** as the substrate. Peak identification: phloretin 3'-C-glucoside (**19a**, nothofagin), phloretin 4'-O-glucoside (**19b**, trilobatin), phloretin 2'-O-glucoside (phlorizin), and phloretin 3',5'-di-C-glucoside.

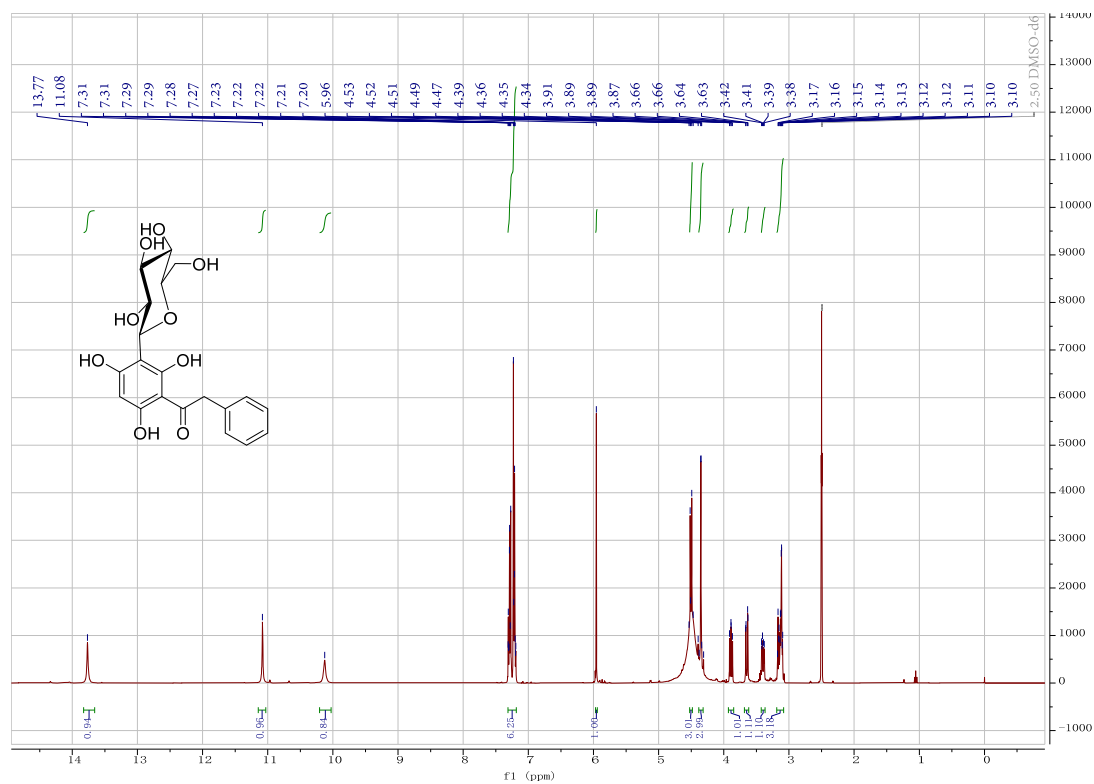


Figure S48. The  $^1\text{H}$  NMR spectrum of **18a** in  $\text{DMSO-}d_6$  (400 MHz)

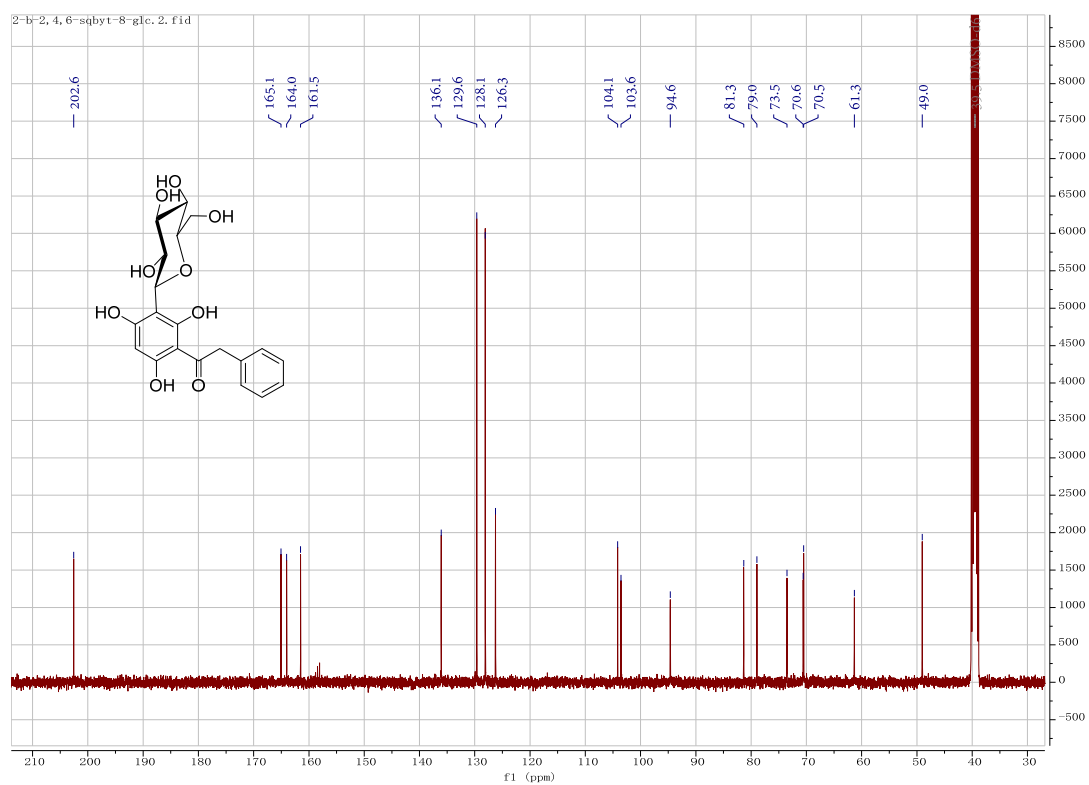
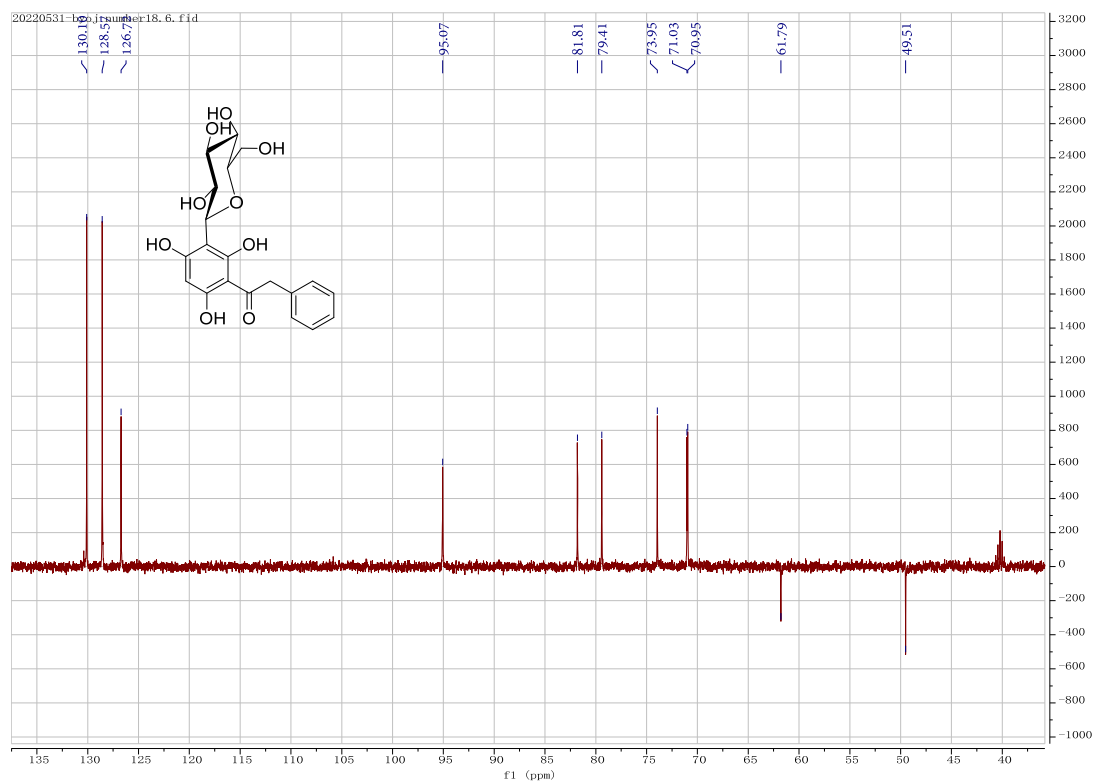
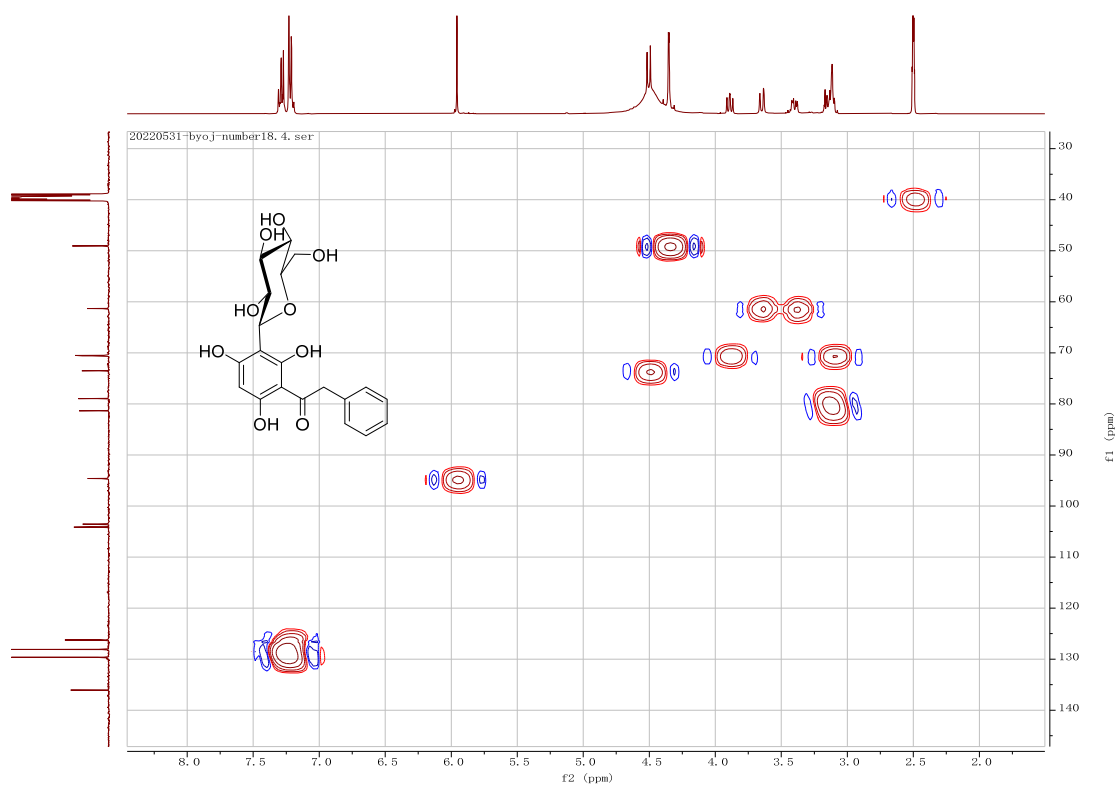


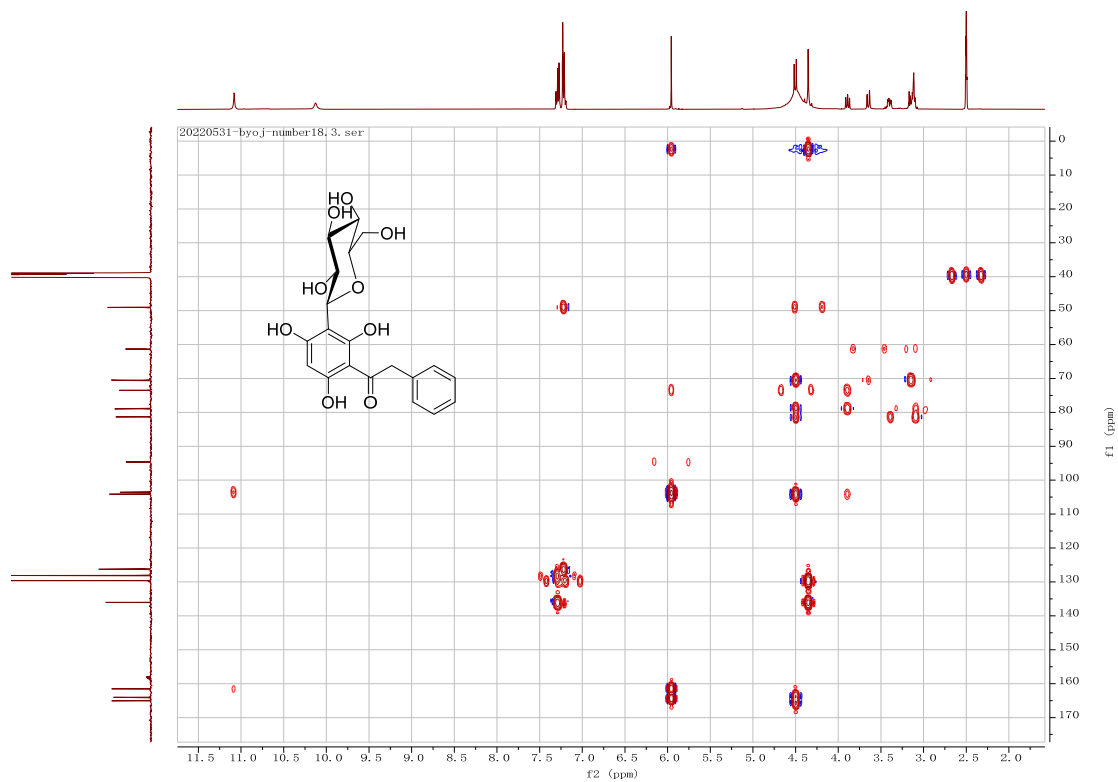
Figure S49. The  $^{13}\text{C}$  NMR spectrum of **18a** in  $\text{DMSO-}d_6$  (150 MHz)



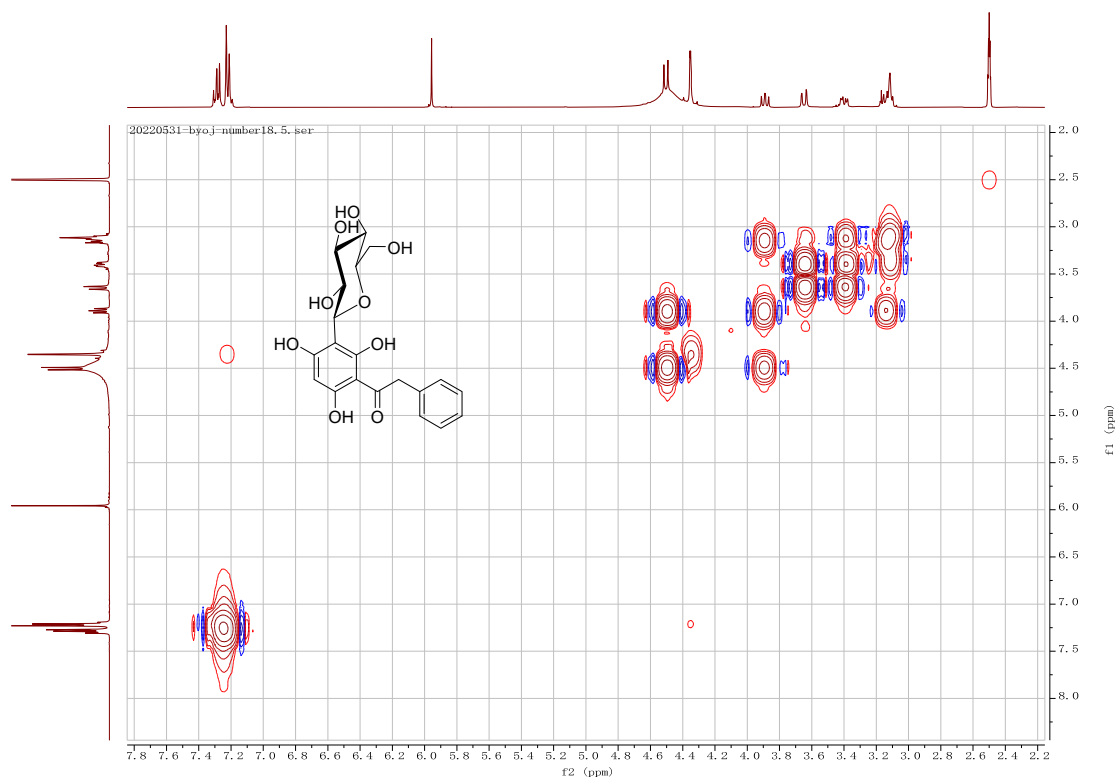
**Figure S50.** The DEPT135 spectrum of **18a** in DMSO- $d_6$  (400 MHz)



**Figure S51.** The HSQC spectrum of **18a** in DMSO- $d_6$  (400 MHz)



**Figure S52.** The HMBC spectrum of **18a** in DMSO- $d_6$  (400 MHz)



**Figure S53.** The  $^1\text{H}$ - $^1\text{H}$  COSY spectrum of **18a** in DMSO- $d_6$  (400 MHz)

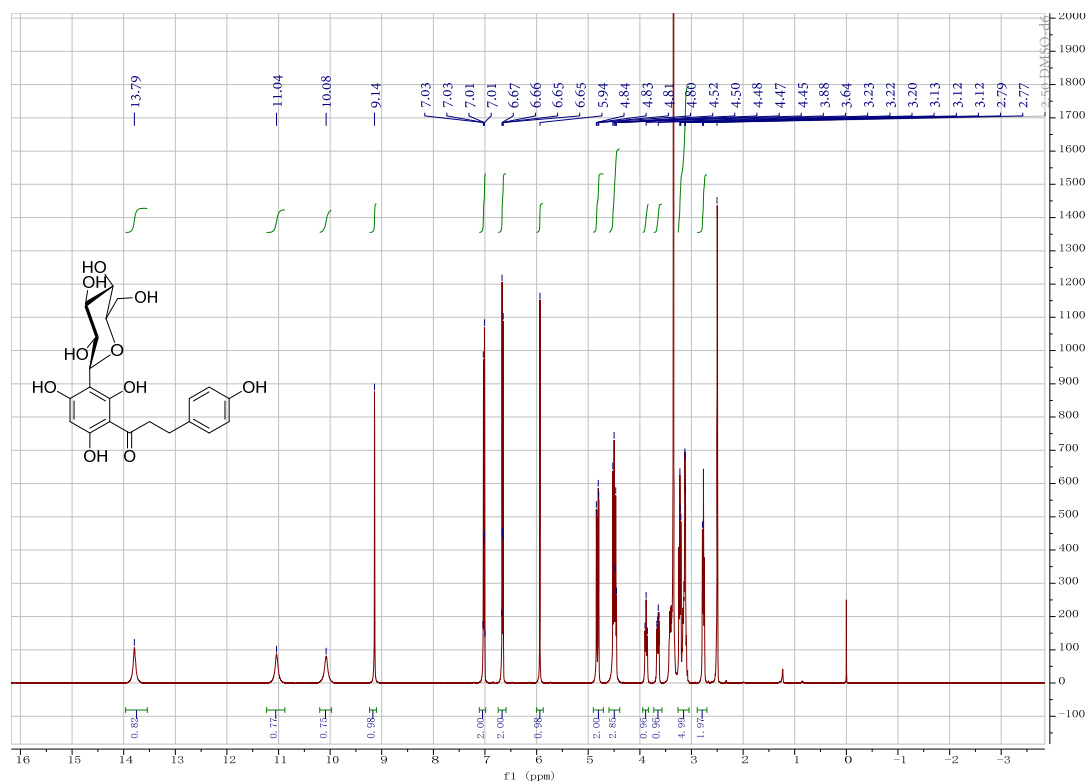


Figure S54. The  $^1\text{H}$  NMR spectrum of nothofagin (**19a**) in  $\text{DMSO-}d_6$  (400 MHz)

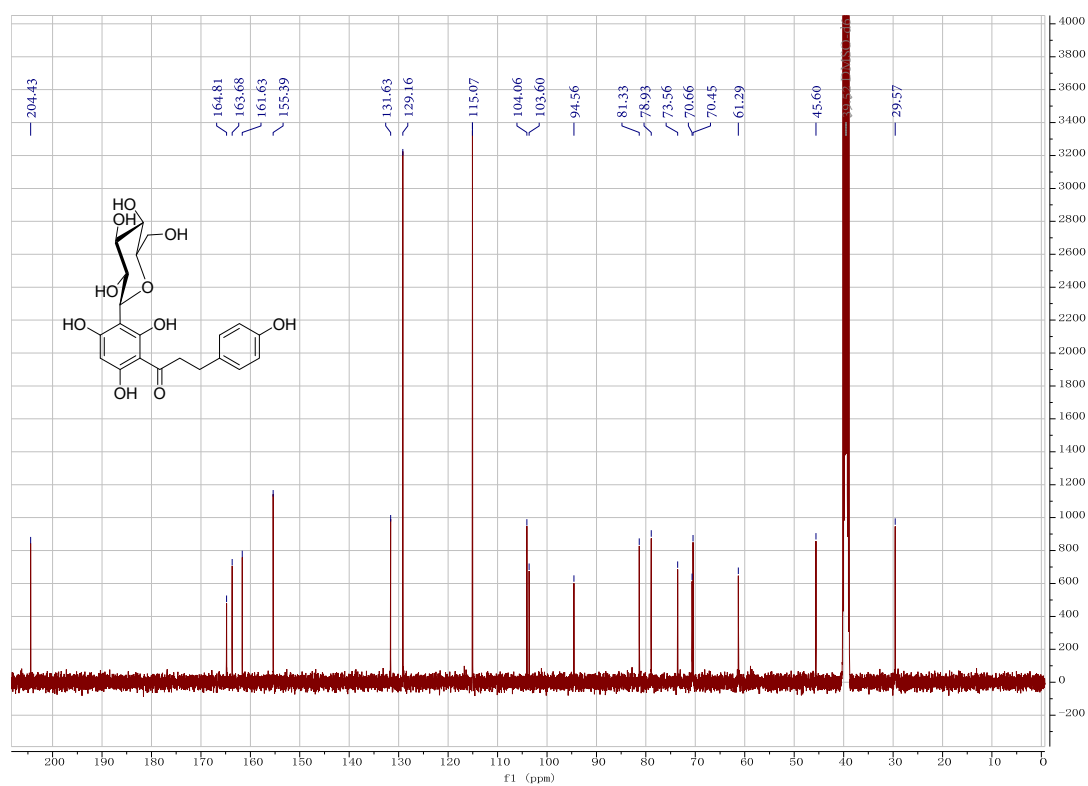
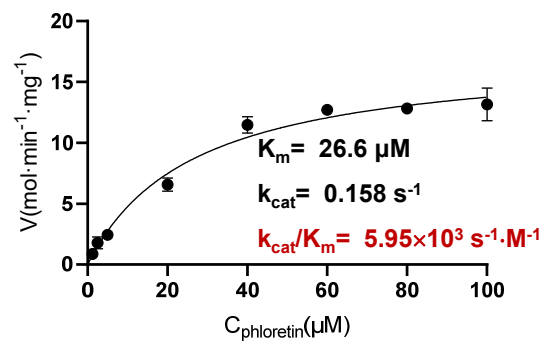


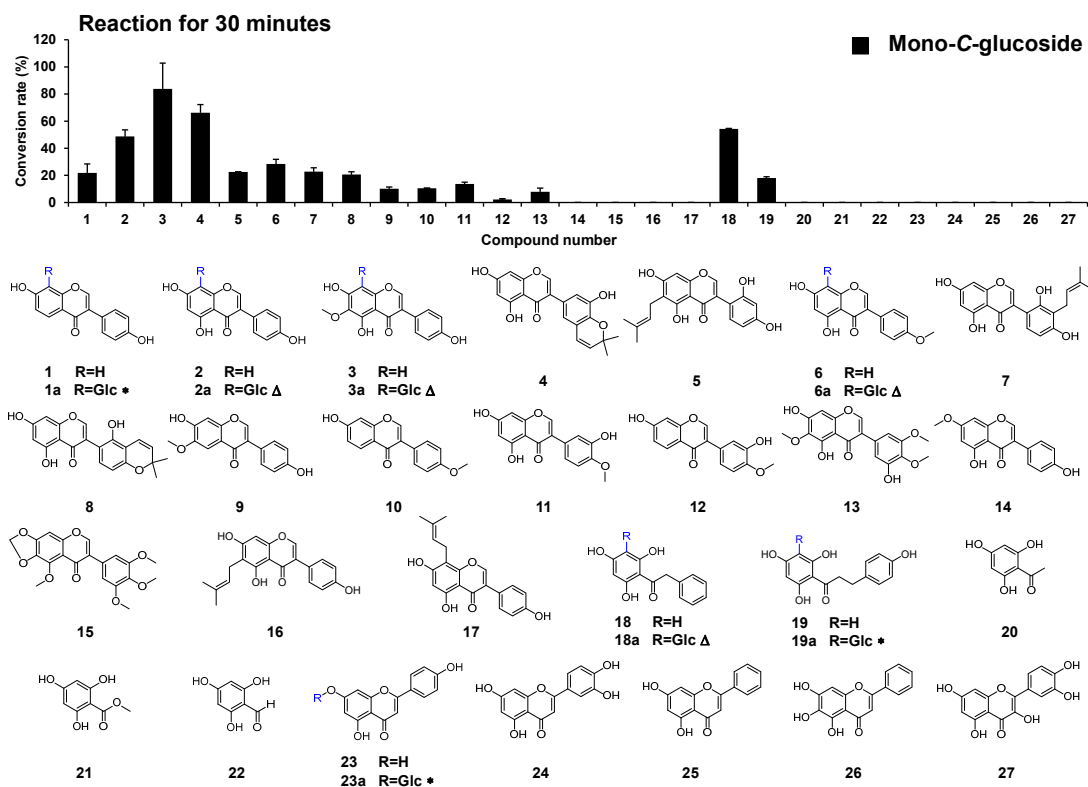
Figure S55. The  $^{13}\text{C}$  NMR spectrum of nothofagin (**19a**) in  $\text{DMSO-}d_6$  (100 MHz)



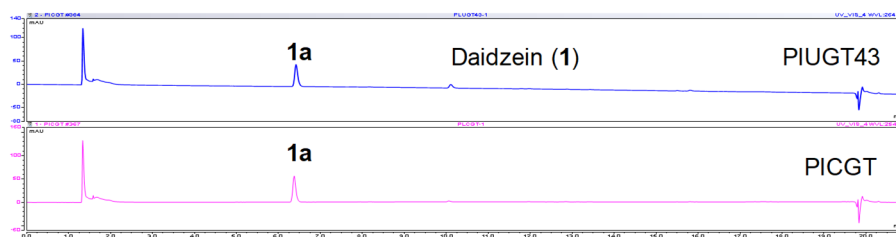




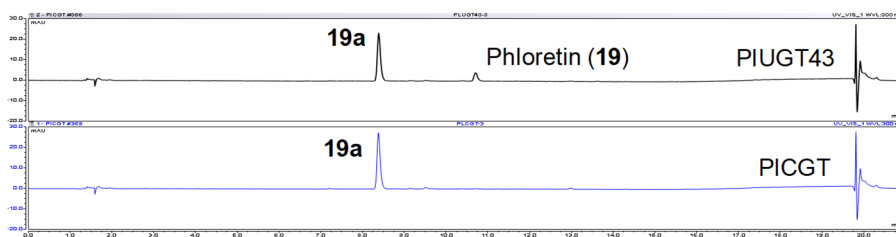
**Figure S58.** Determination of kinetic parameters for PICGT. Michaelis-Menten plot was fitted. The  $K_m$  value was calculated as  $26.6 \mu\text{M}$  using phloretin (**19**) as acceptor and UDP-Glc as donor.



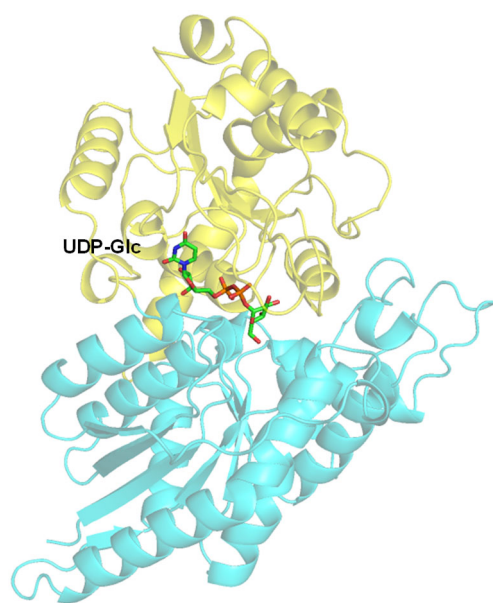
**Figure S59.** Conversion rates of PICGT using substrates **1-23** for short-time reactions. Reactions for 30 minutes with 0.5  $\mu$ g PICGT.



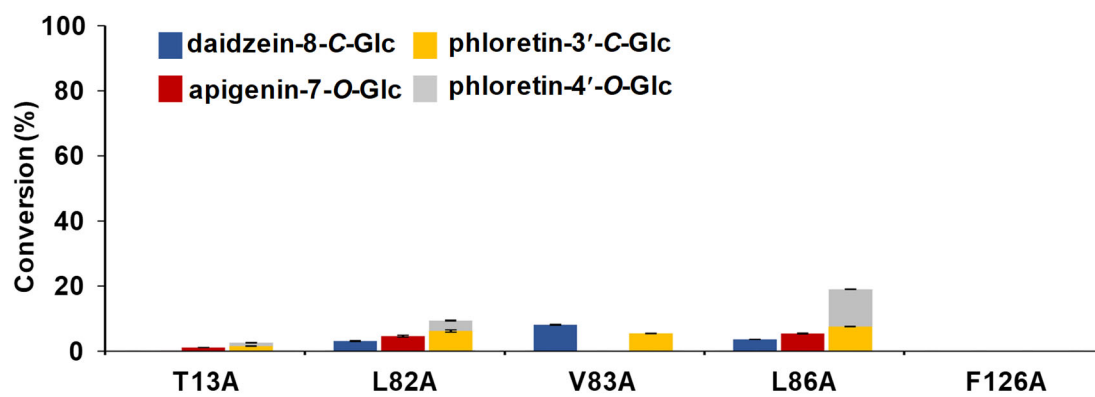
**Figure S60.** HPLC analysis of PICGT and PIUGT43 catalyzing **1** to **1a** in the same condition (37°C, overnight).



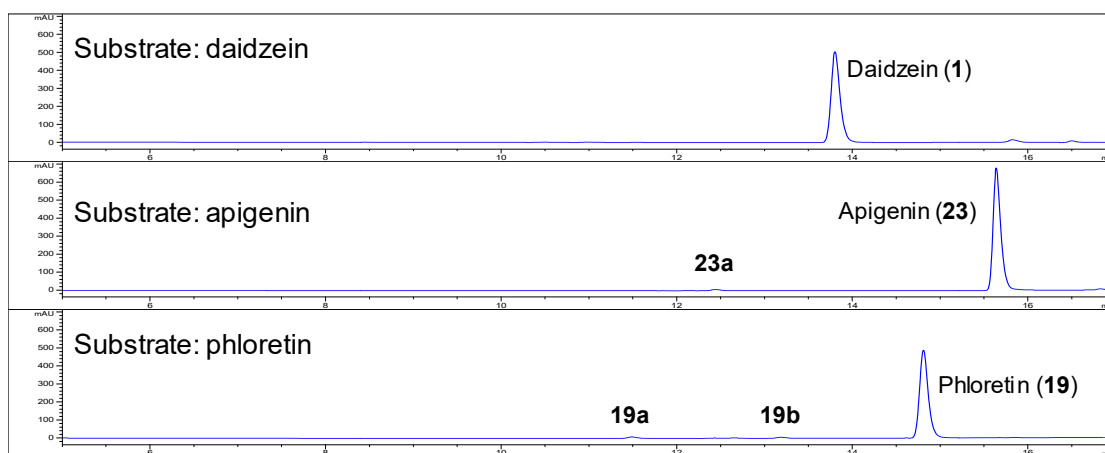
**Figure S61.** HPLC analysis of PICGT and PIUGT43 catalyzing **19** to **19a** in the same condition (37°C, overnight).



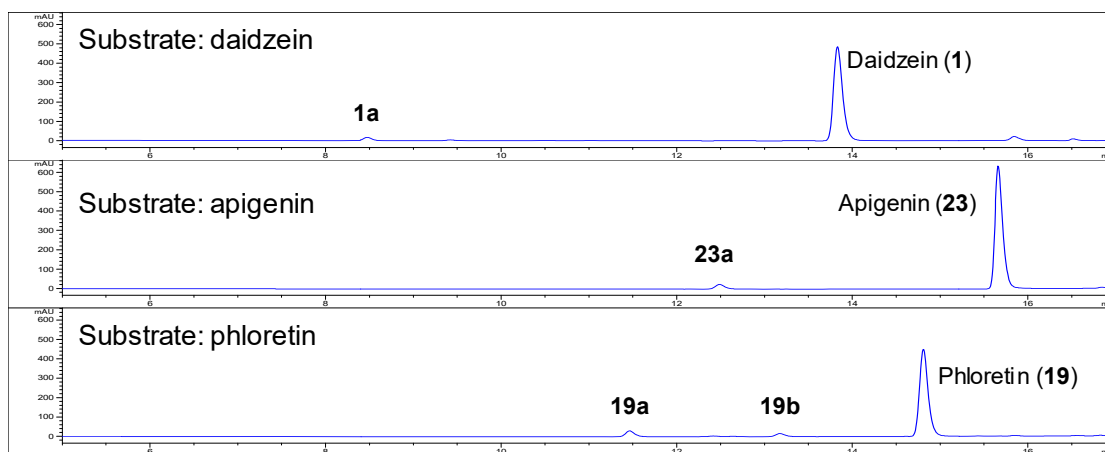
**Figure S62.** Structure model of PICGT predicted by AlphaFold2.



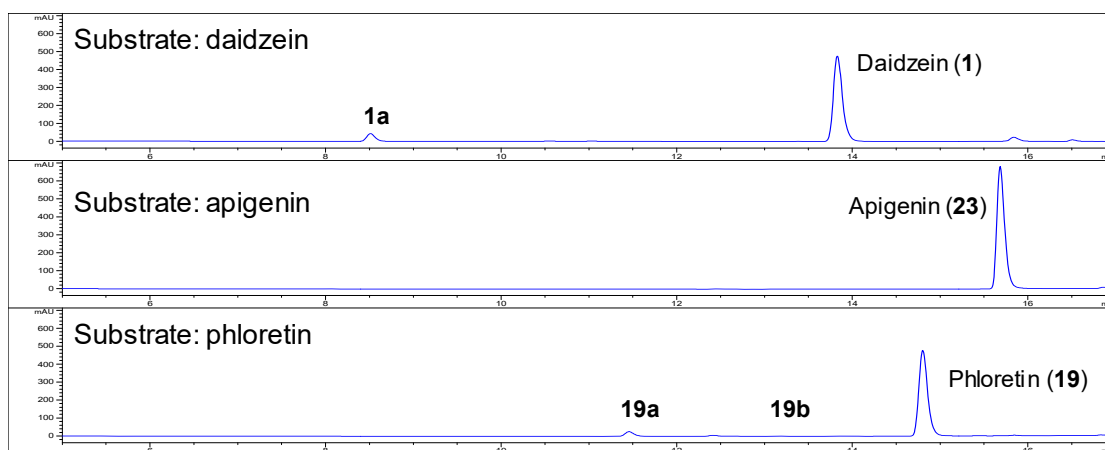
**Figure S63.** Conversion rate (%) of mutants using **1**, **19**, **23** as the substrates.



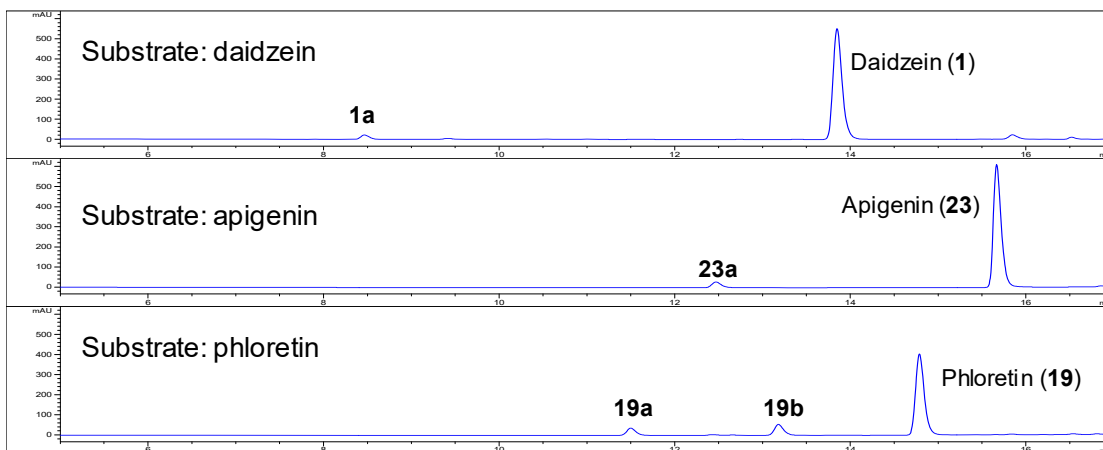
**Figure S64.** HPLC data for mutant T13A using daidzein (1), apigenin (23), and phloretin (19) as the substrates.



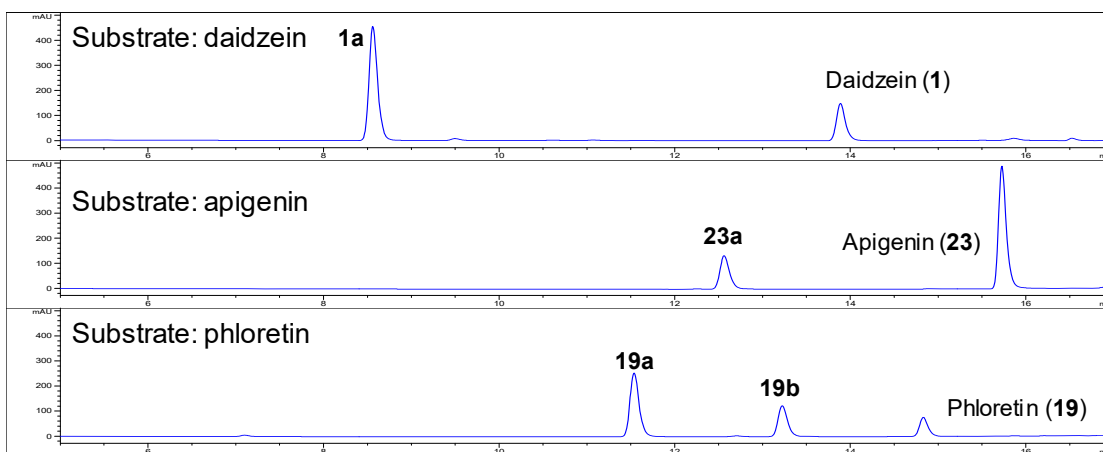
**Figure S65.** HPLC data for mutant L82A using daidzein (1), apigenin (23), and phloretin (19) as the substrates.



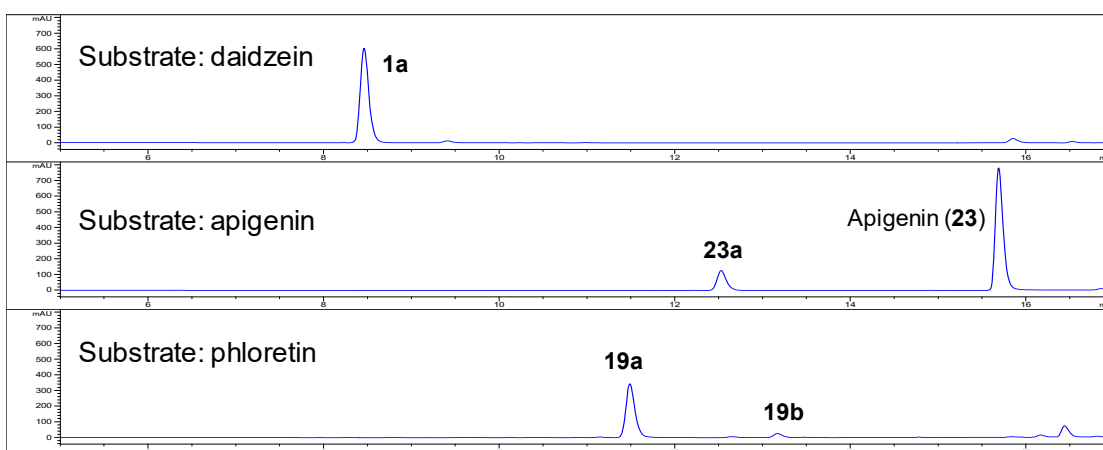
**Figure S66.** HPLC data for mutant V83A using daidzein (1), apigenin (23), and phloretin (19) as the substrates.



**Figure S67.** HPLC data for mutant L86A using daidzein (**1**), apigenin (**23**), and phloretin (**19**) as the substrates.



**Figure S68.** HPLC data for mutant F11A using daidzein (**1**), apigenin (**23**), and phloretin (**19**) as the substrates.

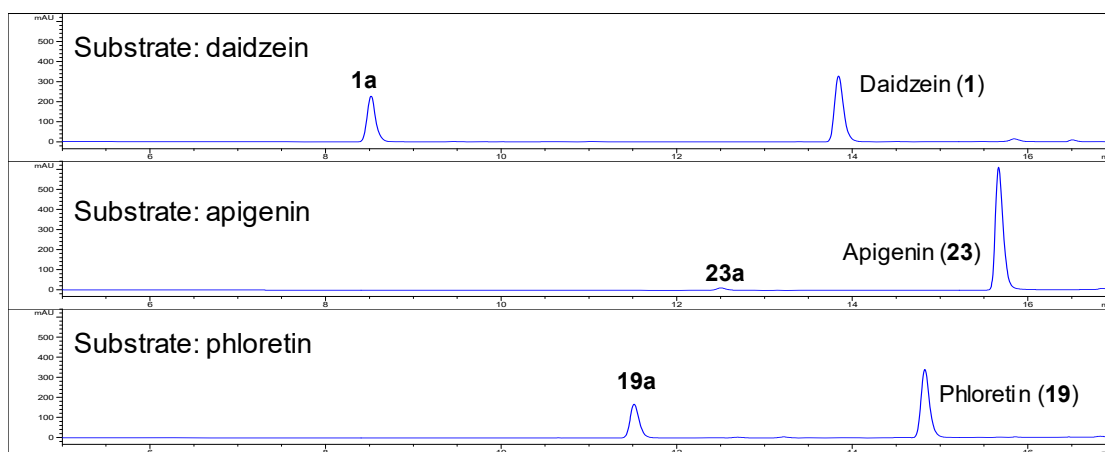


**Figure S69.** HPLC data for mutant L187A using daidzein (**1**), apigenin (**23**), and phloretin (**19**) as the substrates.

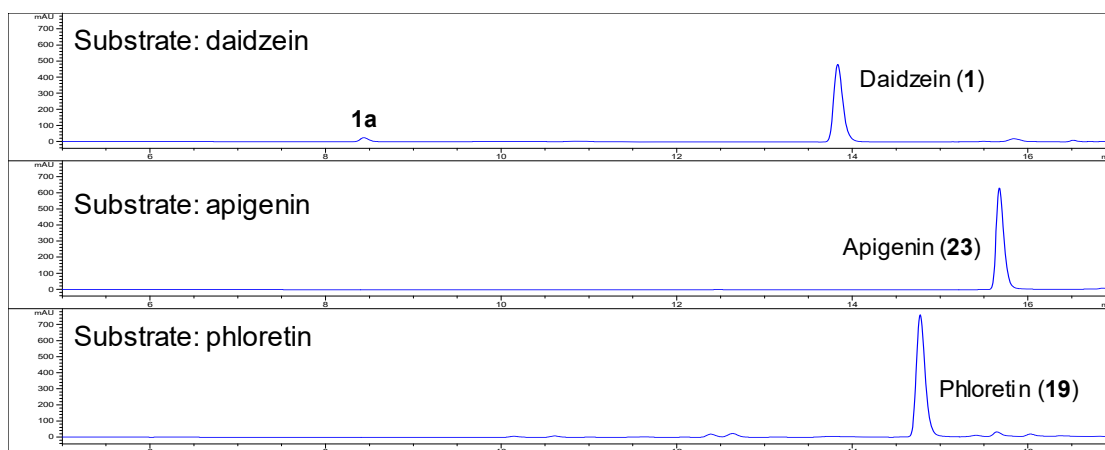




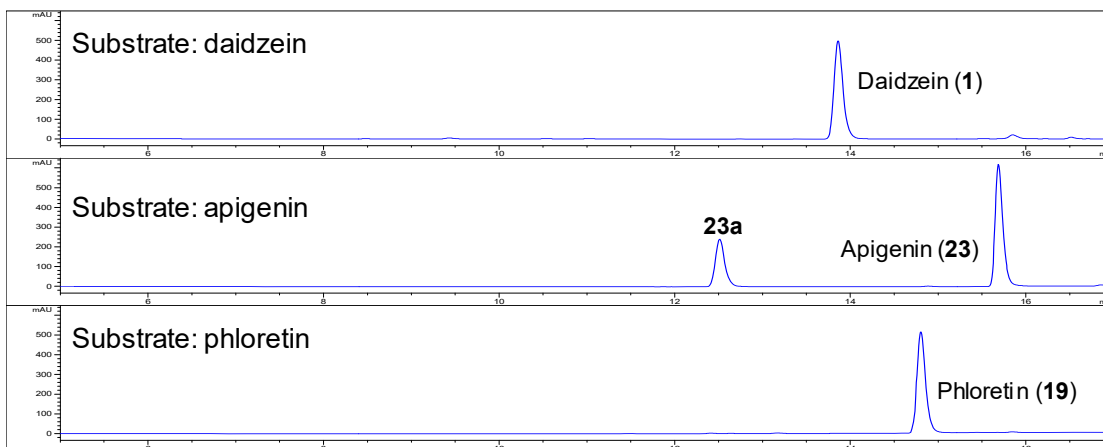
**Figure S70.** HPLC data for mutant N16A using daidzein (1), apigenin (23), and phloretin (19) as the substrates.



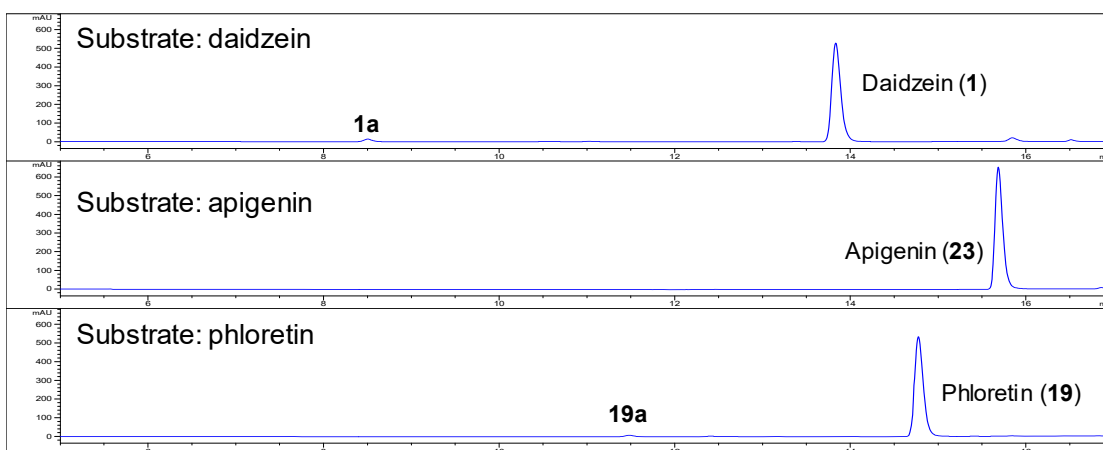
**Figure S71.** HPLC data for mutant N16H using daidzein (1), apigenin (23), and phloretin (19) as the substrates.



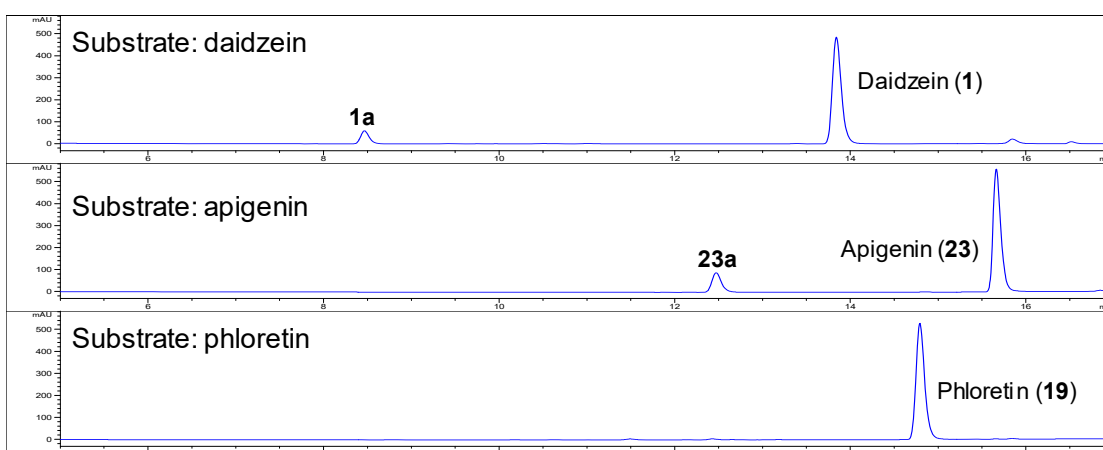
**Figure S72.** HPLC data for mutant N16D using daidzein (1), apigenin (23), and phloretin (19) as the substrates.



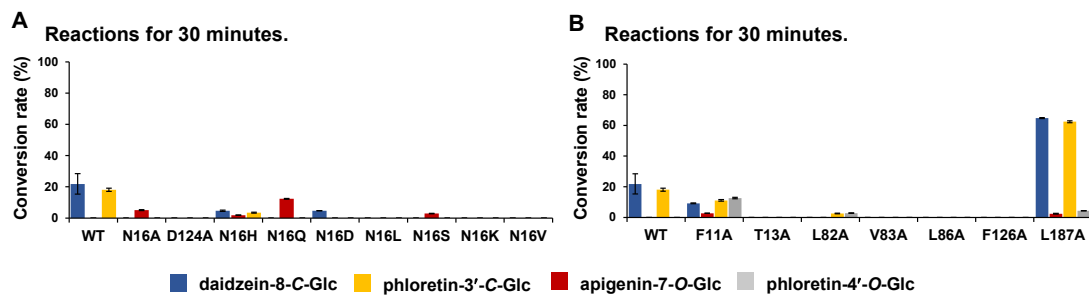
**Figure S73.** HPLC data for mutant N16Q using daidzein (**1**), apigenin (**23**), and phloretin (**19**) as the substrates.



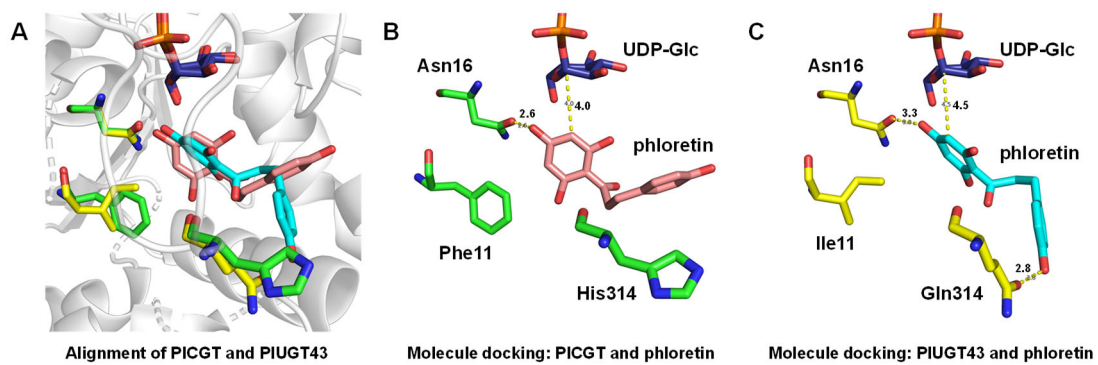
**Figure S74.** HPLC data for mutant N16L using daidzein (**1**), apigenin (**23**), and phloretin (**19**) as the substrates.



**Figure S75.** HPLC data for mutant N16S using daidzein (**1**), apigenin (**23**), and phloretin (**19**) as the substrates.



**Figure S76.** Conversion rates of PICGT mutants using **1**, **19**, **23** for short-time reactions. Reactions for 30 minutes with about 1  $\mu\text{g}$  enzymes.



**Figure S77.** Difference on the molecular docking between PICGT/phloretin and PIUGT43/phloretin.

## References

- 1 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2009, **31**, 455–461.
- 2 X. Wang, C. Li, C. Zhou, J. Li and Y. Zhang, *Plant J.*, 2017, **90**, 535–546.
- 3 T. Ito, S. Fujimoto, F. Suito, M. Shimosaka and G. Taguchi, *Plant J.*, 2017, **91**, 187–198.
- 4 D. Chen, R. Chen, R. Wang, J. Li, K. Xie, C. Bian, L. Sun, X. Zhang, J. Liu, L. Yang, F. Ye, X. Yu and J. Dai, *Angew. Chem. Int. Ed.*, 2015, **54**, 12678–12682.
- 5 D. Chen, L. Sun, R. Chen, K. Xie, L. Yang and J. Dai, *Chem. - Eur. J.*, 2016, **22**, 5873–5877.
- 6 Y. Nagatomo, S. Usui, T. Ito, A. Kato, M. Shimosaka and G. Taguchi, *Plant J.*, 2014, **80**, 437–448.
- 7 Y. Hirade, N. Kotoku, K. Terasaka, Y. Saijo-Hamano, A. Fukumoto and H. Mizukami, *FEBS Lett.*, 2015, **589**, 1778–1786.
- 8 Y. Sun, Z. Chen, J. Yang, I. Mutanda, S. Li, Q. Zhang, Y. Zhang, Y. Zhang and Y. Wang, *Commun. Biol.*, 2020, **3**, 110.
- 9 M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel and R. Edwards, *J. Biol. Chem.*, 2009, **284**, 17926–17934.
- 10 M. L. Falcone Ferreyra, E. Rodriguez, M. I. Casas, G. Labadie, E. Grotewold and P. Casati, *J. Biol. Chem.*, 2013, **288**, 31678–31688.
- 11 J. He, P. Zhao, Z. Hu, S. Liu, Y. Kuang, M. Zhang, B. Li, C. Yun, X. Qiao and M. Ye, *Angew. Chem. Int. Ed.*, 2019, **58**, 11513–11520.
- 12 N. Sasaki, Y. Nishizaki, E. Yamada, F. Tatsuzawa, T. Nakatsuka, H. Takahashi and M. Nishihara, *FEBS Lett.*, 2015, **589**, 182–187.
- 13 K. Mashima, M. Hatano, H. Suzuki, M. Shimosaka and G. Taguchi, *Plant Cell Physiol.*, 2019, **60**, 2733–2743.
- 14 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 and M. Ye, *J. Am. Chem. Soc.*, 2020, **142**, 3506–3512.
- 15 M. McClelland, K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston and R. K. Wilson, *Nature*, 2001, **413**, 852–856.
- 16 B. Faust, D. Hoffmeister, G. Weitnauer, L. Westrich, S. Haag, P. Schneider, H. Decker, E. Künzel, J. Rohr and A. Bechthold, *Microbiology*, 2000, **146**, 147–154.
- 17 C. Fischer, F. Lipata and J. Rohr, *J. Am. Chem. Soc.*, 2003, **125**, 7818–7819.
- 18 L. B. Pickens, W. Kim, P. Wang, H. Zhou, K. Watanabe, S. Gomi and Y. Tang, *J. Am. Chem. Soc.*, 2009, **131**, 17677–17689.
- 19 S. Di, F. Yan, F. R. Rodas, T. O. Rodriguez, Y. Murai, T. Iwashina, S. Sugawara, T. Mori, R. Nakabayashi, K. Yonekura-Sakakibara, K. Saito and R. Takahashi, *BMC Plant Biol.*, 2015, **15**, 126.
- 20 N. Yoshihara, T. Imayama, M. Fukuchi-Mizutani, H. Okuhara, Y. Tanaka, I. Ino and T. Yabuya, *Plant Sci.*, 2005, **169**, 496–501.

- 21 N. Kovinich, A. Saleem, J. T. Arnason and B. Miki, *Phytochemistry*, 2010, **71**, 1253–1263.
- 22 Y. Yin, G. Borges, M. Sakuta, A. Crozier and H. Ashihara, *Plant Physiol. Biochem.*, 2012, **55**, 77–84.
- 23 Kazusa DNA Research Institute, The Cold Spring Harbor and Washington University Sequencing Consortium, The European Union Arabidopsis Genome Sequencing Consortium, and Institute of Plant Genetics and Crop Plant Research (IPK), *Nature*, 2000, **408**, 823–826.
- 24 L. V. Modolo, J. W. Blount, L. Achnine, M. A. Naoumkina, X. Wang and R. A. Dixon, *Plant Mol. Biol.*, 2007, **64**, 499–518.
- 25 A. Theologis, J. R. Ecker, C. J. Palm, N. A. Federspiel, S. Kaul, O. White, J. Alonso, H. Altafi, R. Araujo, C. L. Bowman, S. Y. Brooks, E. Buehler, A. Chan, Q. Chao, H. Chen, R. F. Cheuk, C. W. Chin, M. K. Chung, L. Conn, A. B. Conway, A. R. Conway, T. H. Creasy, K. Dewar, P. Dunn, P. Etgu, T. V. Feldblyum, J. Feng, B. Fong, C. Y. Fujii, J. E. Gill, A. D. Goldsmith, B. Haas, N. F. Hansen, B. Hughes, L. Huizar, J. L. Hunter, J. Jenkins, C. Johnson-Hopson, S. Khan, E. Khaykin, C. J. Kim, H. L. Koo, I. Kremenetskaia, D. B. Kurtz, A. Kwan, B. Lam, S. Langin-Hooper, A. Lee, J. M. Lee, C. A. Lenz, Joycelyn. H. Li, Y. Li, X. Lin, S. X. Liu, Z. A. Liu, J. S. Luros, R. Maiti, A. Marziali, J. Militscher, M. Miranda, M. Nguyen, W. C. Nierman, B. I. Osborne, G. Pai, J. Peterson, P. K. Pham, M. Rizzo, T. Rooney, D. Rowley, H. Sakano, S. L. Salzberg, J. R. Schwartz, P. Shinn, A. M. Southwick, H. Sun, L. J. Tallon, G. Tambunga, M. J. Toriumi, C. D. Town, T. Utterback, S. Van Aken, M. Vaysberg, V. S. Vysotskaia, M. Walker, D. Wu, G. Yu, C. M. Fraser, J. C. Venter and R. W. Davis, *Nature*, 2000, **408**, 816–820.
- 26 M. Hirota, R. Kuroda, H. Suzuki and T. Yoshikawa, *Planta*, 2000, **210**, 1006–1013.
- 27 A. Noguchi, N. Sasaki, M. Nakao, H. Fukami, S. Takahashi, T. Nishino and T. Nakayama, *J. Mol. Catal. B Enzym.*, 2008, **55**, 84–92.
- 28 L. Zapata, J. Ding, E.-M. Willing, B. Hartwig, D. Bezdán, W.-B. Jiao, V. Patel, G. Velikkakam James, M. Koornneef, S. Ossowski and K. Schneeberger, *Proc. Natl. Acad. Sci.*, , DOI:10.1073/pnas.1607532113.
- 29 T. Nakatsuka, K. Sato, H. Takahashi, S. Yamamura and M. Nishihara, *J. Exp. Bot.*, 2008, **59**, 1241–1252.
- 30 Y. Sun, K. Ji, B. Liang, Y. Du, L. Jiang, J. Wang, W. Kai, Y. Zhang, X. Zhai, P. Chen, H. Wang and P. Leng, *Plant J.*, 2017, **91**, 574–589.
- 31 M. Yamazaki, Z. Gong, M. Fukuchi-Mizutani, Y. Fukui, Y. Tanaka, T. Kusumi and K. Saito, *J. Biol. Chem.*, 1999, **274**, 7405–7411.