Chemo-enzymatic Synthesis of Equisetin

Xiaojun Li, [a,†] Qingfei Zheng,[b,†] Jun Yin, [a] Wen Liu [b] * and Shuanhu Gao[a] *

a. Shanghai Key Laboratory of Green Chemistry and Chemical Processes, School of Chemistry and Molecular Engineering, East China Normal University, 3663N Zhongshan Road, Shanghai 200062

b. State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China.

† These authors contributed equally.

Supporting Information

Figure S1. Coomassie-stained SDS-PAGE analysis of the purified recombinant protein, Fsa2.

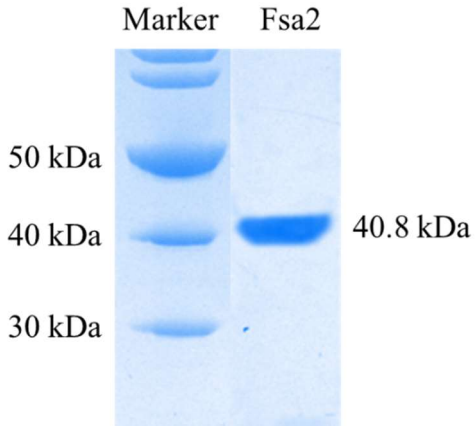


Figure S2. Condition optimization for the enzymatic reaction mediated by Fsa2. (A) pH-dependence of the Fsa2-mediated conversions from **11** to equisetin (**1**). (B) Temperature-dependence of the Fsa2-mediated conversions from **11** to equisetin (**1**). (C) Organic solvent tolerance of the Fsa2-mediated conversions from **11** to equisetin (**1**).

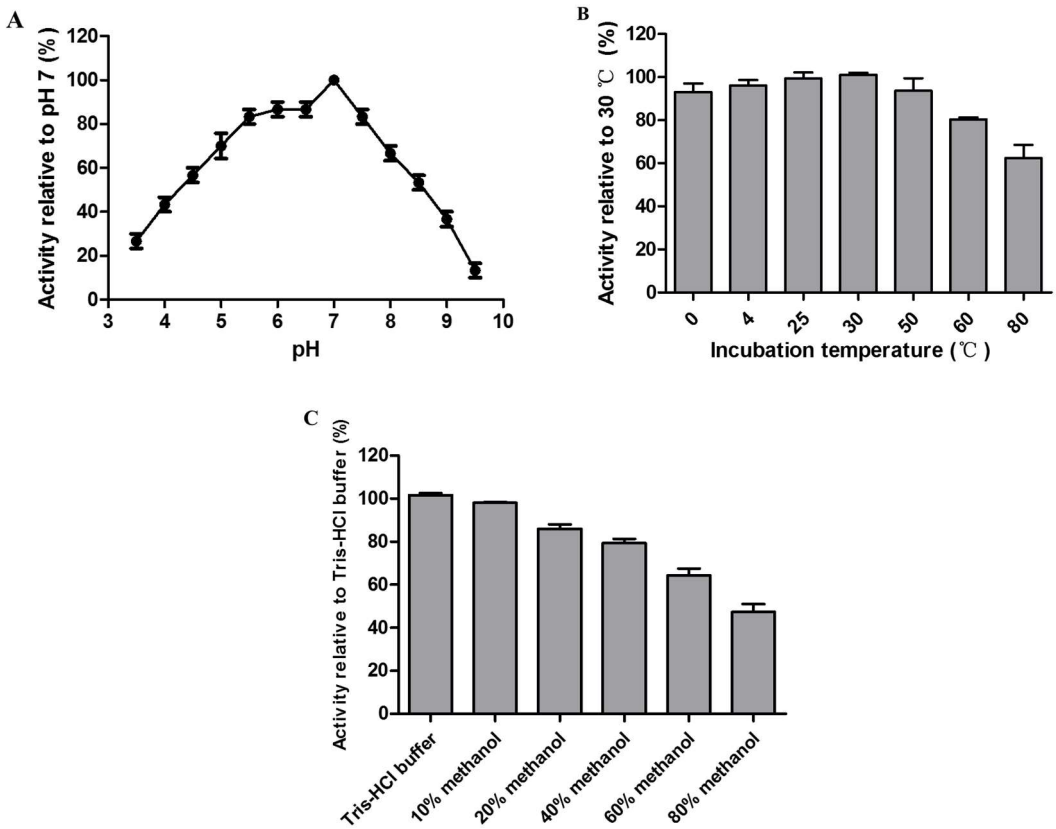
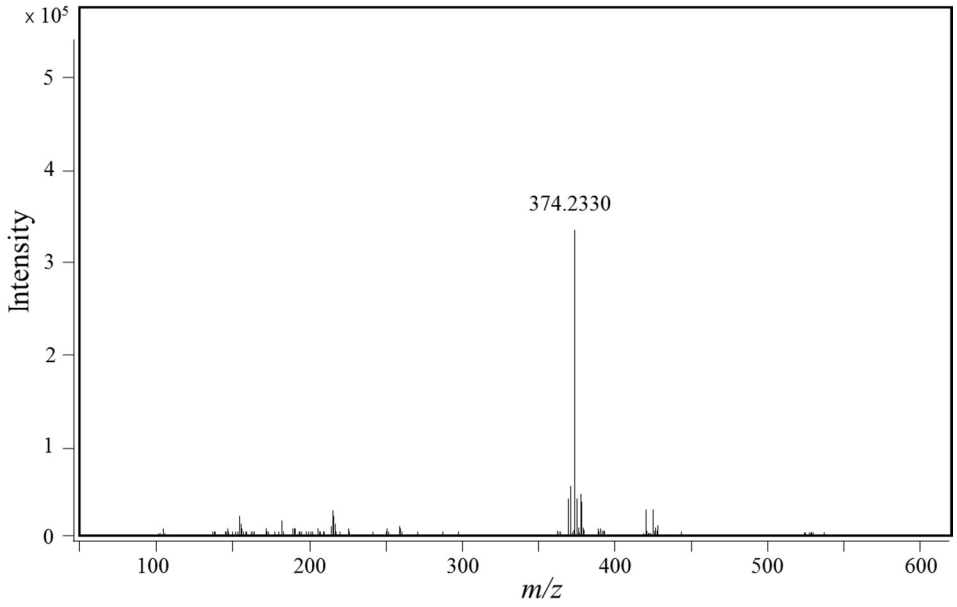
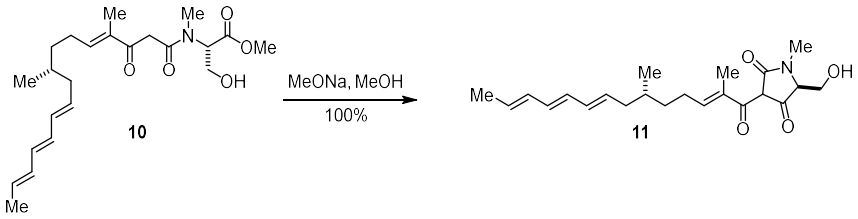


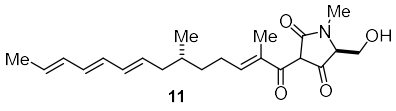
Figure S3. HR-LC-MS analysis of compound [(3*S*, 6*R*)-diastereomer of **1** (**1'**) + H]$^+$.



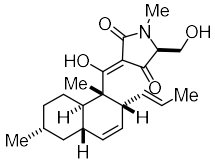
II Experimental Procedures and Spectroscopic Data of Compounds

General Experimental Procedures: Unless stated otherwise, all reactions were carried out under a nitrogen atmosphere. Dichloromethane, dimethylsulfoxide and dimethylformamide were distilled from calcium hydride. Tetrahydrofuran and 2-methyltetrahydrofuran was distilled from sodium-benzophenone in a continuous still under an atmosphere of nitrogen. Flash column chromatography was performed as described by Still (Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923–2925), employing Qingdao Haiyang silica gel 60 (200–300 mesh). TLC analyses were performed on EMD 250 μm Silica Gel $HSGF_{254}$ plates and visualized by quenching of UV fluorescence (λ_{max}= 254 nm), or by staining ceric ammonium molybdate, ammonium molybdate, or potassium permanganate. ^1H and ^{13}C NMR spectra were recorded on a Bruker-500, 400 spectrometer. Chemical shifts for ^1H and ^{13}C NMR spectra are reported in ppm (δ) relative to residue protium in the solvent ($CDCl_3$: δ 7.26, 77.0 ppm; CD_3OD: 3.31, 49.0 ppm; the multiplicities are presented as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. High-resolution mass spectra (HRMS) were acquired on a waters GCT premier, and Mass spectra at Agilent 5975C. The $[\alpha]_D$ was recorded using PolAAr 3005 High Accuracy Polarimeter. Reversed-phase HPLC was performed on an Agilent 1200 HPLC equipped with a DAD detector using either a semi-preparative (ZORBAX SB, 9.6 mm × 250 mm, 5 μm) C18 HPLC column or analytic (ZORBAX SB, 4.6 mm × 250 mm, 5 μm) C18 HPLC column. To analyze the species of the *in vitro* assay, reaction mixtures were analyzed by gradient elution of mobile phase A (H_2O supplemented with 0.1% formic acid) and mobile phase B (acetonitrile supplemented with 0.1% formic acid) with the flow rate of 1.0 mL/min: 0 to 5 min, 65% phase B; 5 to 15 min, 65% to 100% phase B; 15 to 20 min, 100% phase B; 20 to 25 min, 100% to 65% phase B; and 25 to 30 min, 65% phase B, λ at 220 nm). The gene of Fsa2 was optimized and synthesized at Shanghai Sangon Biotech Co. Ltd. (China). All the molecular subcloning experiments were conducted by using Phanta Max Super-Fidelity DNA polymerase (Vazyme Biotech Co. Ltd., China) and/or ClonExpress Entry One Step Cloning Kit (Vazyme Biotech Co. Ltd., China).



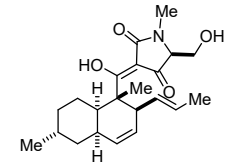


Diels-Alder precursor **11**: To a solution of **10** (12.2 mg, 0.03 mmol, 1.0 equiv.) in anhydrous d_4-methanol (0.5 mL) was added sodium d_4-methoxide (25 μL, 0.06 mmol, 2.4 mol/L, 2.0 equiv.) at 0 ºC. After stirring at 0 ºC for 10 minutes, the reaction mixture was stirred at 25 ºC for 2 hours. The mixture was directly detected by NMR and LC-MS (ESI), then the mixture was quenched with 1N HCl (0.1 mL) and stirred at 25 ºC for 5 minutes. The solvent was removed under reduced pressure, then ethyl acetate (5 mL) and water (2 mL) was added to the residue. The organic phase was separated and the aqueous phase was extracted with ethyl acetate (5 mL ×2). The combined organic phase was washed with water (5 mL) and brine (5 mL), then dried over anhydrous sodium sulfate, filtered, concentrated, to give crude **11** (11 mg, 100%) as light yellow oil. *R*f = 0.40 (100% ethyl acetate); $[\alpha]_D^{25}$ = -52.4 (*c* = 0.5 in MeOD); ^{1}H NMR (500 MHz, MeOD) δ 6.16 – 5.96 (m, 4H), 5.91 – 5.84 (m, 1H), 5.65 (ddt, *J* = 21.3, 14.1, 7.1 Hz, 2H), 4.98 (s, 1H), 3.89 (dd, *J* = 11.7, 3.3 Hz, 1H), 3.82 (dd, *J* = 11.7, 4.1 Hz, 1H), 3.45 (t, *J* = 3.6 Hz, 1H), 2.93 (s, 3H), 2.23 – 2.08 (m, 3H), 1.94 (dt, *J* = 20.4, 6.7 Hz, 1H), 1.81 (s, 3H), 1.74 (d, *J* = 6.7 Hz, 3H), 1.56 (td, *J* = 13.1, 6.6 Hz, 1H), 1.52 – 1.49 (m, 1H), 1.48 – 1.42 (m, 1H), 1.28 – 1.20 (m, 1H), 0.90 (d, *J* = 6.7 Hz, 3H). ^{13}C NMR (125 MHz, MeOD) δ 195.6, 193.7, 177.1, 138.7, 136.3, 133.4, 133.2, 132.1, 131.8, 129.37, 102.8, 67.9, 61.6, 49.9, 41.3, 36.7, 34.3, 27.6, 27.0, 19.9, 18.4, 13.5 ppm. MS (m/z): ESI [M] calcd for $C_{22}H_{31}NO_4$ 373.22, found 374.06 (M+1).



(-)-equisetin (**1**)

Equisetin (**1**) was obtained as a pale red foam based on our previous reports.[2] Rf = 0.14 (60% ethyl acetate-petroleum ether); [α]= - 286.7 (c = 0.79 in $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.40 (m, 2H), 5.31–5.10 (m, 2H), 4.03-4.01 (m, 1H), 3.89 (m, 1H), 3.64 (m, 1H), 3.35 (br, 1H), 3.04 (s, 3H), 2.04-1.86 (m, 2H), 1.86-1.74 (m, 4H), 1.53 (m, 3H), 1.24 (s, 3H), 1.16-0.99 (m, 3H), 0.91 (d, *J* = 6.5 Hz, 3H) ppm; ^{13}C NMR (100 MHz, CDCl3) δ 199.1, 190.6, 177.0, 130.9, 130.0, 127.1, 126.6, 66.7, 60.4, 48.7, 45.0, 42.2, 39.9, 38.6, 35.7, 33.5, 29.7, 28.2, 27.3, 22.5, 17.9, 14.1 ppm; HRMS (m/z): EI [M] calcd for $C_{22}H_{31}NO_4$ 373.2253, found 373.2254.



(3*S*, 6*R*)-diastereomer of **1** (**1'**)

The crude (3*S*, 6*R*)-diastereomer of **1** (**1'**) was obtained as a mixture, and the ratio of **1** and **1'** was 2:1 (Scheme 1, entry 3).

HRMS (m/z): ESI [M] calcd for $C_{22}H_{31}NO_4$ 373.2253, found 374.2330 (M+1) (Figure S3).

The optimized gene sequence of Fsa2 (the restriction enzyme cutting sites for molecular subcloning are underlined and indicated in red):

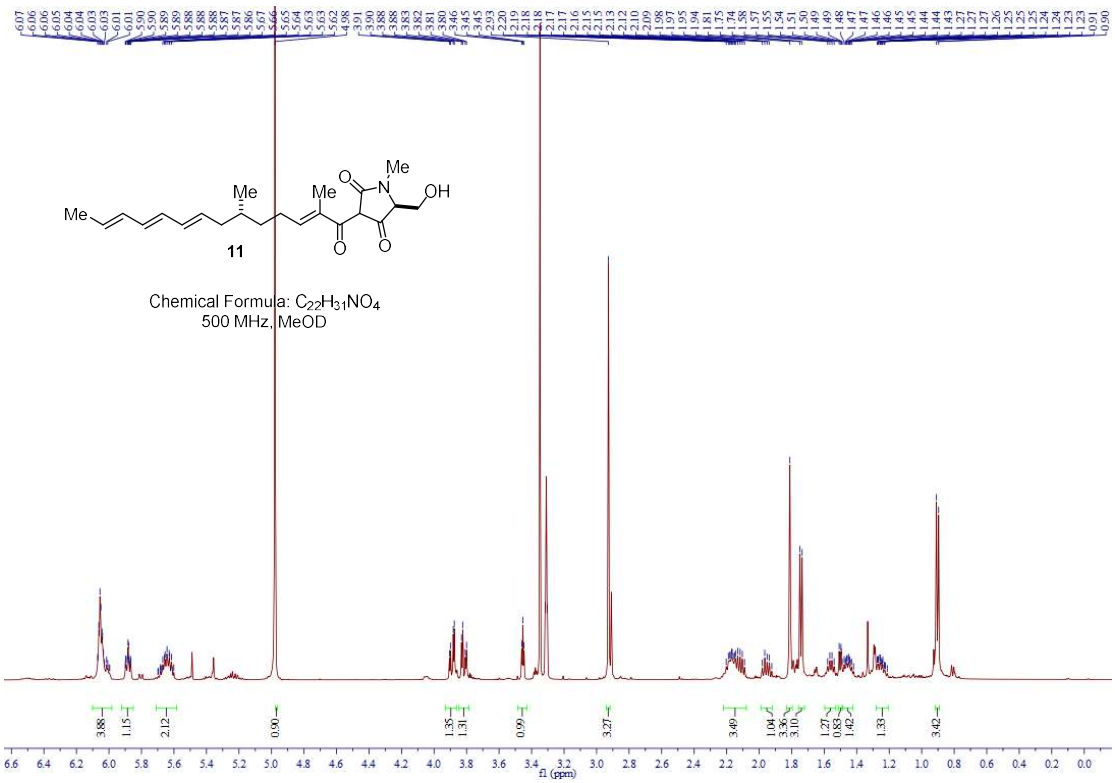
CATATGTCCAACGTGACCGTATCCGCATTCACGGTCGATAAGAGCATCAGCGAGGAGCATGT
CCTGCCTAGCAGCTTCATTCCGGGTAGCGGTAACATCTTCCCAAAGTTCACCTCCGCAATCC
CAAAGACGGCTTGGGAGCTGTGGTACTTTGACGGTATCAGCAAAGACGACAAAAGCTCCA
TCGTCATCGGTGTGACCCGTAACGCGGAGGGTCTGAAACATGGTGGTTTCAAAGTGCAGGT
CTTCGTGATCTGGGCAGACGAACGTACGTGGCACCGTGACCTGTTCTTCCCGGAATCCGTG

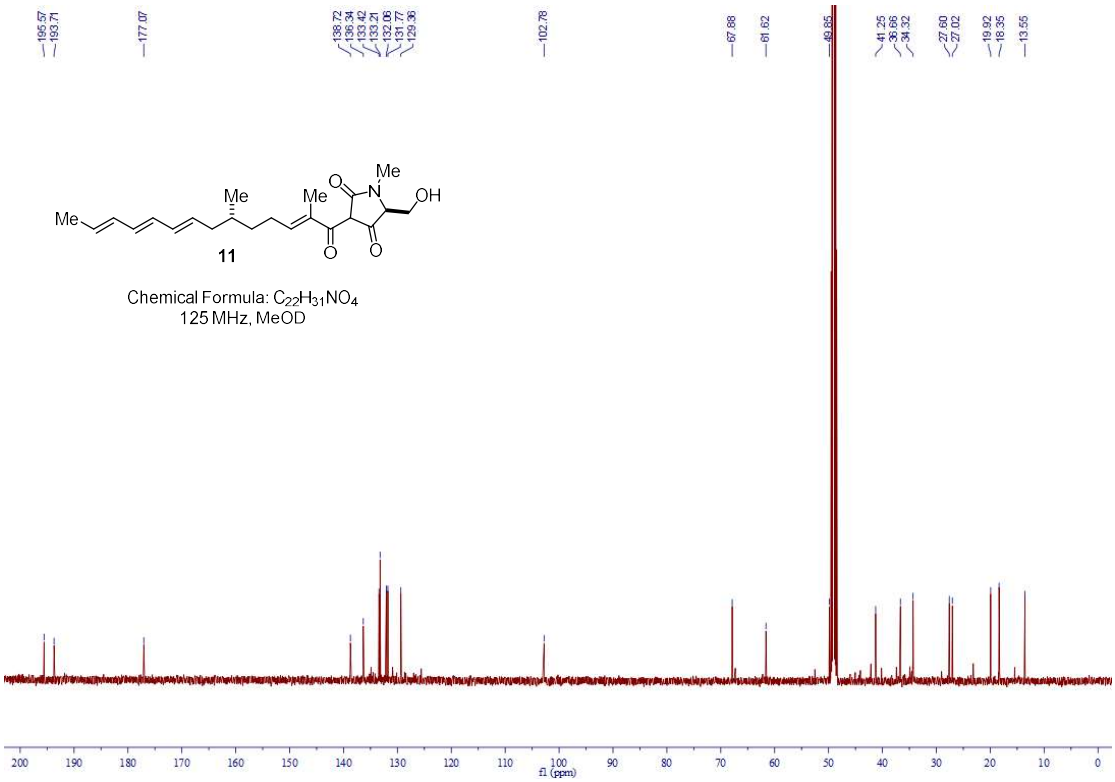
GTGTCTATCAACGAATCCGGTGTAACCGACGGTATTTGGAAAGACGCTACGAGCAACAGCT
CTATCAGCTTCAGCTGCGCTGGTGATCTGTCCAAAGCATCCCTGGTGTTCGACGTGCCGGG
CGTGGTACAGGGTGACATGCACCTGGAAGCTCTGCCAGGTGATACTGGTCTGGATACTGAC
GCACGTCTGGGTCCGTCTGTTTACTATGTTCGTCCGATCGGTCGTGCTTCTGTAAAAGCCCA
GCTGTCTCTGTACAGCTCCGACGCTACTGCGGCGGAACAATTTTCCCTGGGTACTTCCGCG
AATGGTGGTATGGATCGTGTTTGGTCCCCGCTGTCTTGGCCGCAAGTAATGACTGAATCTTA
CTACCTGCGCACTCAGGTTGGTCCGTACGCCATGCAAATCATGCGTATTTTTCCGCCGGCCG
GTTCTGAAGATCAGCCGTCTACCATGGCGCGTCTGTACCGCGAAGGCCAGCTGGTTTGTGT
TGCGCAGCACGTTGTAACTCGCGAAGATGCGCTGATGACCCACGATTCTCTGATTCTGTCTA
AACAGGATAACTCCGACTCTGAAGACGTTGTTACCGGCGGCTATCGTGATAAAAACACCGG
CTATACCGTTGAATTTGTTGAAAAAGGCAACGAGGGCCAGCGTTGGAAATTTCAGGTACGC
CACGAACGCATCATTTGGAACACCCCGACCTCTCGCCCTGGCCCGGATGCCACCGGCAATA
CCGGCTTCGTTGAAGTTCTGTGCGGCGGCACCATTGGCGAATCTTATGAAGGCGTTGGCAC
TGGCGGCCAGTGTGAACTGTCTtaaAAGCTT

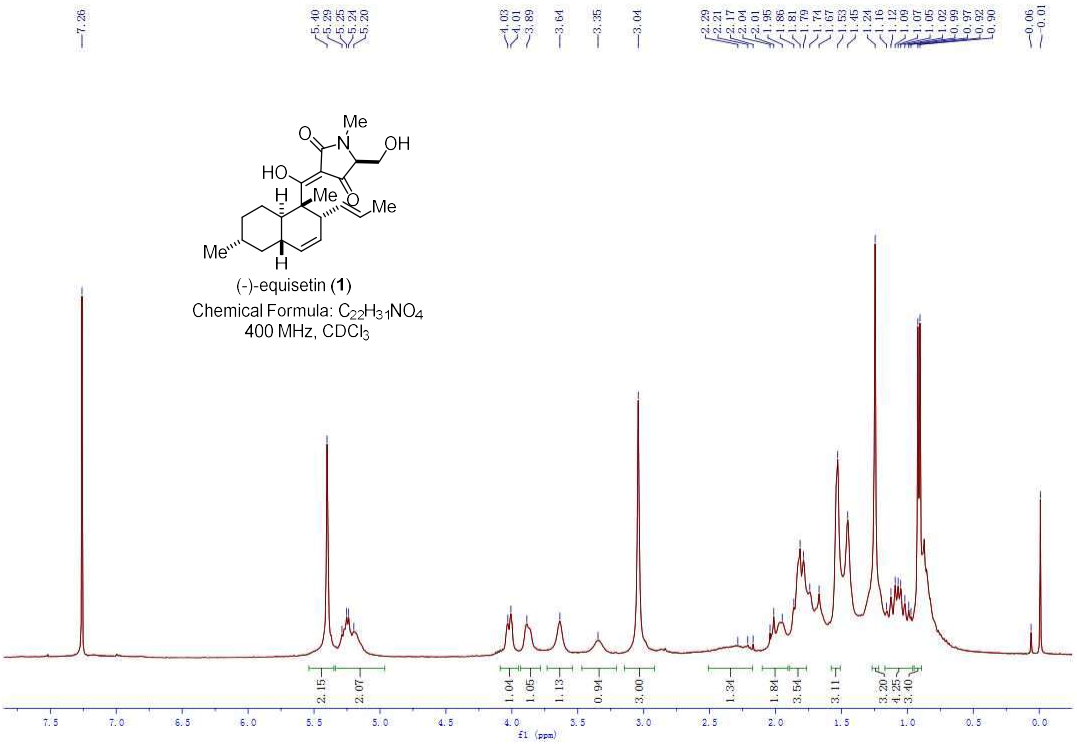
Protein expression and purification: The synthesized gene *fsa2* was purified, digested with *Nde*I and *Xho*I and then ligated into the vector pET-28a, which was digested with the same enzymes. The resulting recombinant plasmid was then transferred into *E. coli* BL21 (DE3) for protein overexpression. The culture of *E. coli* transformant was incubated in Luria-Bertani (LB) medium containing 50 μg/mL kanamycin at 37 °C and at 250 r.p.m. until the cell density reached 0.6-0.8 at OD600. Protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM, followed by further incubation for 40 h at 16 °C. The cells were harvested by centrifuging at 5,000 r.p.m. for 20 min at 4 °C and were resuspended in 30 mL of lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 10% glycerol, pH 7.0). After disruption by FB-110X Low Temperature Ultra-pressure Continuous Flow Cell Disrupter (Shanghai Litu Mechanical Equipment Engineering Co., Ltd, China), the soluble fraction was collected and subjected to purification of each target protein using a HisTrap FF column (GE Healthcare, USA). The desired protein fractions, as determined by SDS-PAGE, were concentrated and desalted using a PD-10 Desalting Column (GE Healthcare, USA) according to the manufacturer's protocols. The concentration of protein was determined by Bradford assay using bovine serum albumin (BSA) as the standard.

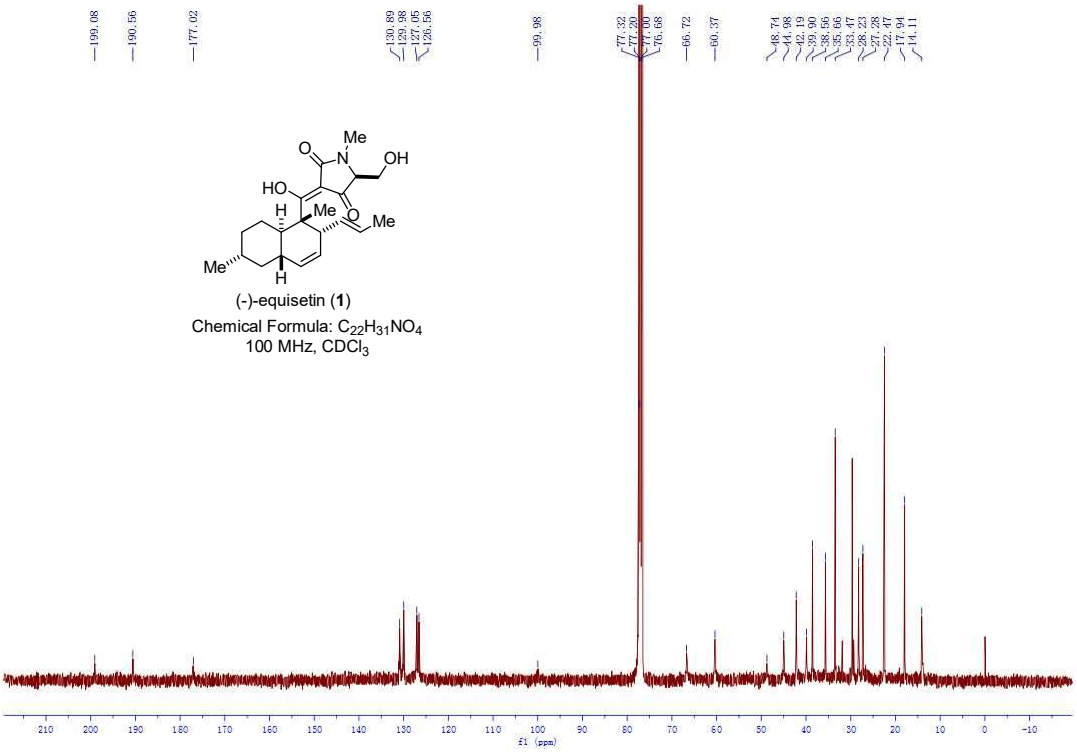
In vitro enzymatic assays: The assays were conducted at 30 °C in a mixture containing 200 μM substrate **11** in 50 mM Tris-HCl buffer (pH = 7.0). The enzyme Fsa2, at a final concentration of 5 μM for each, was added to the solution to initiate the reaction (the total volume was 100 μL). After 20 min of incubation, the reactions were quenched by being added a twofold volume of methanol. After centrifugation, the supernatant was subjected to HPLC or HPLC-ESI-MS analyses. For kinetic analysis, a time-course experiment was conducted to determine the initial rate conditions in each 20 μL reaction mixture that contained 200 μM **11**, 0.3 μM Fsa2 and 50 mM Tris-HCl buffer (pH 7.0). The reactions were initiated by adding Fsa2, incubated at 30 °C and then terminated with 40 μL of methanol at 0 min, 0.5 min, 1.0 min, 2.0 min, 5.0 min or 10.0 min. The production of equisetin, which was linear with respect to time from 0-3.0 min, was fitted to a linear equation to obtain the initial velocity. To determine the kinetic parameters, the reactions were performed at 30 °C in the solutions containing 0.3 μM PyrI4, 50 mM Tris-HCl buffer (pH 7.0) and varying concentrations of substrate **11** (0 μM, 25 μM, 50 μM, 100 μM, 200 μM, 400 μM, 600 μM or 800 μM). All assays were performed in triplicate. The resulting initial velocities were then fitted to the Michaelis-Menten equation using OriginPro 8 (Originlab, Co., USA) extract the parameters K_m and k_{cat}.

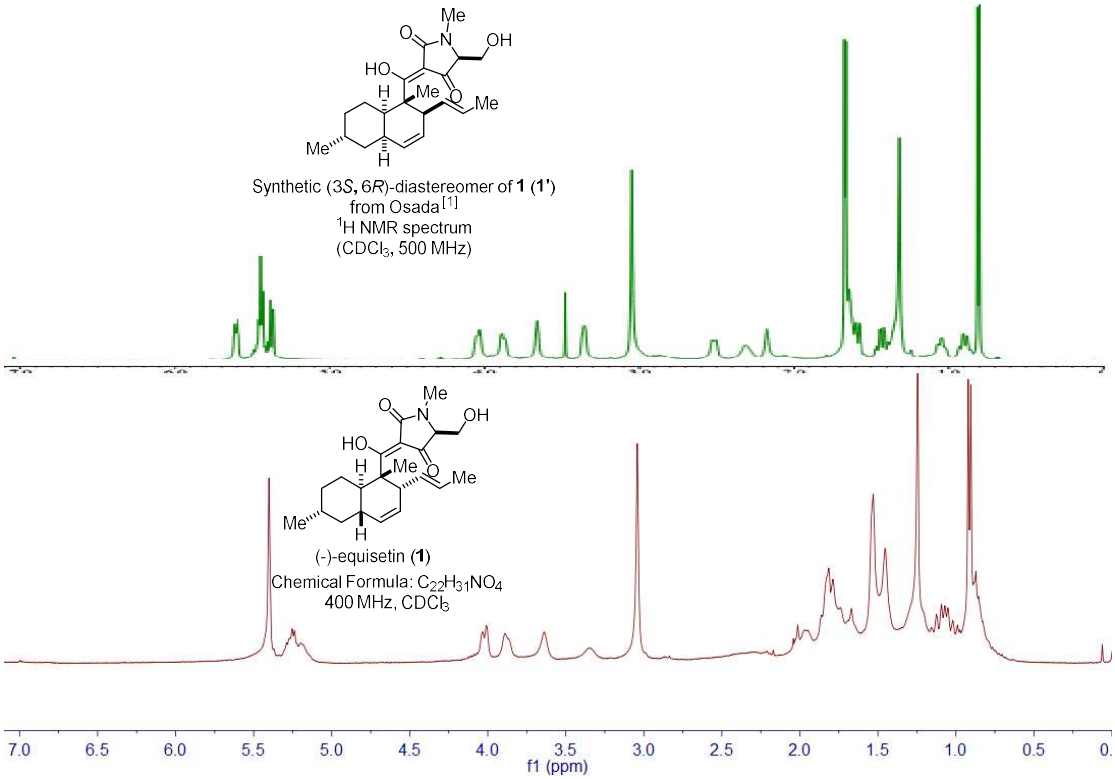
III ^{1}H and ^{13}C NMR Spectra of Compounds

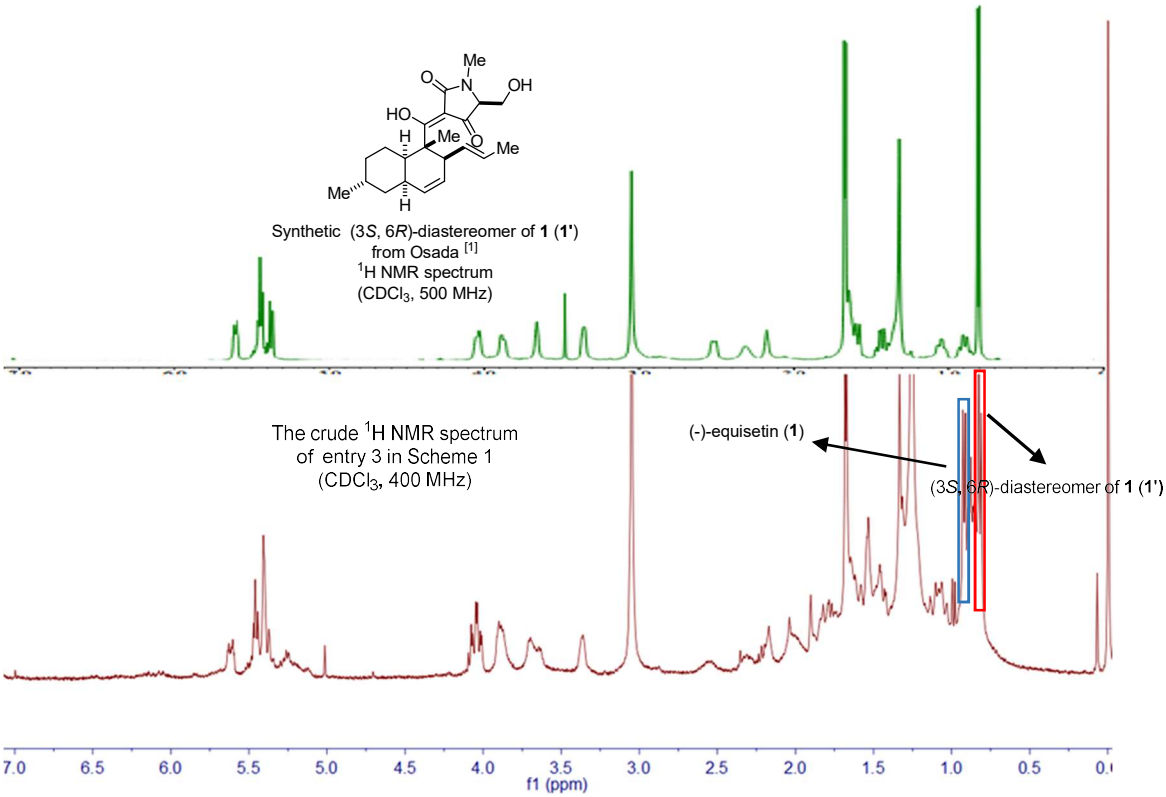





7.26
5.40
5.29
5.25
5.24
5.20
4.03
4.01
3.89
3.64
3.35
3.04
2.29
2.21
2.17
2.04
2.01
1.95
1.86
1.81
1.79
1.74
1.67
1.53
1.45
1.24
1.16
1.12
1.09
1.07
1.05
1.02
0.99
0.97
0.92
0.90
0.06
0.01
Me
O
N
OH
HO
H
Me
O
Me
Me
H
(-)-equisetin (1)
Chemical Formula: C22H31NO4
400 MHz, CDCl3
2.15
2.07
1.04
1.05
1.13
0.94
3.00
1.34
1.84
3.54
3.11
3.20
4.25
3.40
f1 (ppm)


199.08
190.56
177.02
130.89
129.98
127.05
126.95
99.98
77.32
77.20
77.00
76.68
66.72
60.37
48.74
41.98
42.19
39.90
38.56
35.66
33.47
28.23
27.28
22.47
17.94
14.11
Me
O
N
OH
HO
H
Me
O
Me
Me
H
(-)-equisetin (1)
Chemical Formula: C22H31NO4
100 MHz, CDCl3
f1 (ppm)


Synthetic (3S, 6R)-diastereomer of 1 (1')
from Osada [1]
1H NMR spectrum
(CDCl3, 500 MHz)
(-)-equisetin (1)
Chemical Formula: C22H31NO4
400 MHz, CDCl3
7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)


Synthetic (3S, 6R)-diastereomer of 1 (1')
from Osada [1]
1H NMR spectrum
(CDCl3, 500 MHz)
The crude 1H NMR spectrum
of entry 3 in Scheme 1
(CDCl3, 400 MHz)
(-)-equisetin (1)
(3S, 6R)-diastereomer of 1 (1')
7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)

IV Supplementary references

1 N. Kato, T. Nogawa, H. Hirota, J. H. Jang, S. Takahashi, J. S. Ahn and H. Osada, *Biochem. Biophys. Res. Commun.* **2015**, *460*, 210.

2 J. Yin, L. Kong, C. Wang, Y. Shi, S. Cai and S. Gao, *Chem. - Eur. J.* **2013**, *19*, 13040.