Cell, Volume *187*

Supplemental information

Structure-guided discovery of bile acid derivatives for treating liver diseases without causing itch

Jun Yang, Tianjun Zhao, Junping Fan, Huaibin Zou, Guangyi Lan, Fusheng Guo, Yaocheng Shi, Han Ke, Huasheng Yu, Zongwei Yue, Xin Wang, Yingjie Bai, Shuai Li, Yingjun Liu, Xiaoming Wang, Yu Chen, Yulong Li, and Xiaoguang Lei

Methods S1. Chemical synthesis of compounds used in this study. Related to

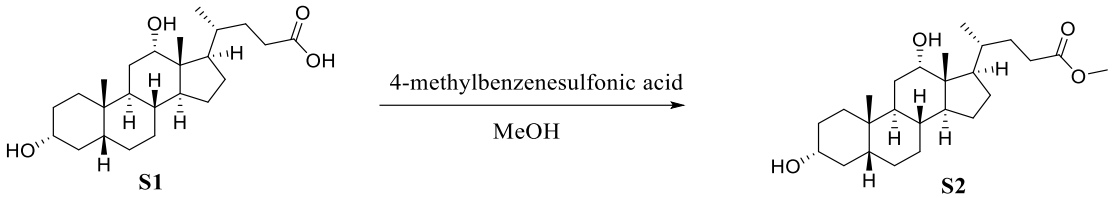
STAR Methods

General information

^{1}H NMR spectra were recorded on Bruker 400 MHz, 500 MHz or 600 MHz spectrometer at ambient temperature with $CDCl_3$, methanol-d_4 or DMSO-d_6 as the solvent unless otherwise stated. ^{13}C NMR spectra were recorded on Brucker 100 MHz, 125 MHz or 150 MHz spectrometer (with complete proton decoupling) at ambient temperature. Chemical shifts are reported in parts per million relative to $CDCl_3$, methanol-d_4 or DMSO-d_6 (^{1}H, δ 7.26 for $CDCl_3$, 3.31 for methanol-d_4, 2.50 for DMSO-d_6; ^{13}C, δ 77.16 for CDCl3, 49.00 for methanol-d_4, 39.52 for DMSO-d_6). Data for $_{1}$H NMR are reported as follows: chemical shift, integration, multiplicity (s = singlet, d = doublet, t = triplet, dd = doublet of doublets, m = multiplet) and coupling constants. High-resolution mass spectra were obtained at Peking University Mass Spectrometry Laboratory using a Bruker APEX Flash chromatography. Analytical thin layer chromatography was performed using 0.25 mm silica gel 60-F plates. Flash chromatography was performed using 200-300 mesh silica gel. Yields refer to chromatographically and spectroscopically pure materials unless otherwise stated. Tetrahydrofuran was distilled from sodium/benzophenone ketyl prior to use; the other solvents were distilled from calcium hydride unless otherwise noted. Reagents were purchased at the highest commercial quality and used without further purification unless otherwise stated. All reactions were carried out in oven-dried glassware under an argon atmosphere with dry solvents unless otherwise noted.

Synthesis of DCA-3S

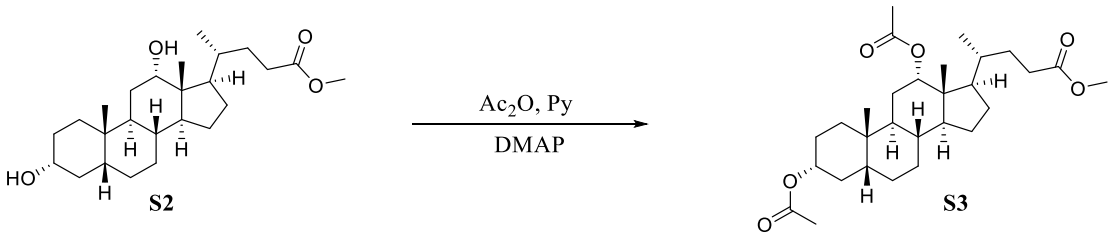
Step 1. Synthesis of compound S2



To a solution of compound S1 (1g, 2.56 mmol) in dry MeOH (10 mL) was added 4-methylbenzenesulfonic acid (263 mg, 1.5 mmol). The reaction mixture was stirred at 25℃ for 2 hours under N_2. The reaction mixture was concentrated. The organic solution was concentrated and purified by com flash (DCM / MeOH = 20:1) to afford product S2 (900 mg, 88%) as a white solid.

^{1}H NMR (400 MHz, Methanol-d_4) δ 3.99 (d, J = 3.0 Hz, 1H), 3.68 (s, 3H), 3.56 (m, 1H), 2.42 (m, 1H), 2.29 (m 1H), 1.03 (d, J = 6.3 Hz, 3H), 0.97 (s, 3H), 0.74 (s, 3H).

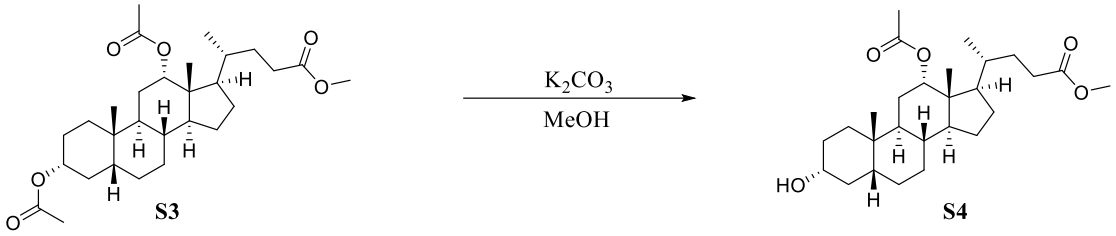
Step 2. Synthesis of compound S3



To a solution of S2 (200 mg, 0.5 mmol) in dry pyridine (5 mL) was added Ac_2O (0.6 mL, 2 mmol) and DMAP (40 mg, 0.02 mmol). The reaction mixture was stirred at 25℃ for 2 hours. Ice and ethyl acetate (20 mL) were added to the organic phase. The organic solution was washed twice with water and brine and dried over magnesium sulfate. The organic solution was concentrated and purified by com flash (petroleum ether / ethyl acetate = 4:1) to afford product S3 (180 mg, 75%) as a yellow oil.

^{1}H NMR (400 MHz, Chloroform-d) δ 5.15 – 5.01 (m, 1H), 4.71 (m, 1H), 3.67 (s, 3H), 2.36 (m, 1H), 2.21 (m, 1H), 2.11 (s, 3H), 2.05 (s, 3H), 0.92 (s, 3H), 0.82 (d, J = 6.4 Hz, 3H), 0.74 (s, 3H).

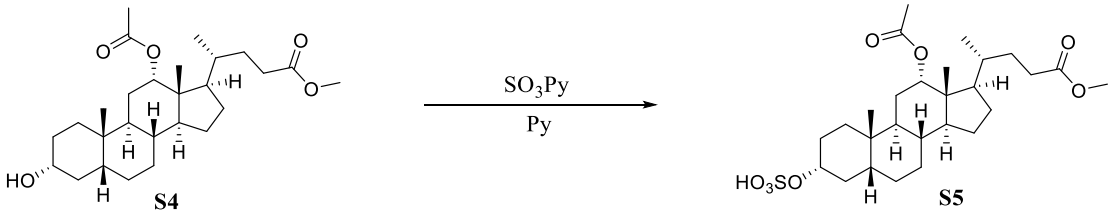
Step 3. Synthesis of compound S4



To a solution of S3 (200 mg, 0.41 mmol) in dry MeOH (5 mL) was added K_2CO_3 (138 mg, 1 mmol). The reaction mixture was stirred at 25°C for 2 hours. Acetic acid (0.5 mL) and ethyl acetate (20 mL) were added to the organic phase. The organic solution was washed twice with water and brine and dried over magnesium sulfate. The organic solution was concentrated and purified by prep-TLC (petroleum ether / ethyl acetate = 3:1) to afford product S4 (160 mg, 32%) as a yellow oil.

^{1}H NMR (400 MHz, Chloroform-d) δ 5.10 (s, 1H), 3.69 (s, 3H), 3.67 – 3.60 (m, 1H), 2.36 (m, 1H), 2.23 (m, 1H), 2.11 (s, 3H), 0.92 (s, 3H), 0.83 (d, *J* = 6.4 Hz, 3H), 0.75 (s, 3H).

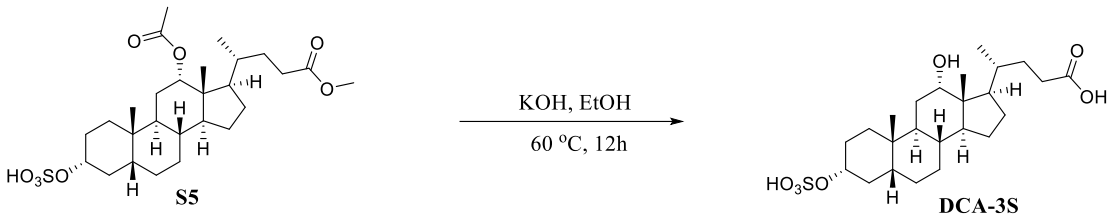
Step 4: Synthesis of compound S5



To a solution of S4 (44.8 mg, 0.1 mmol) in dry pyridine (3 mL), SO_3Py (45 mg, 0.3 mmol) was added. The reaction mixture was stirred at 25°C for 2 hours. The organic solution was filtered and concentrated to afford the crude product and purified by prep-TLC (DCM / MeOH = 10:1) to afford product S5 (44 mg, 85%) as a yellow oil.

^{1}H NMR (400 MHz, Methanol-d_4) δ 5.10 (s, 1H), 4.30 (d, *J* = 7.2 Hz, 1H), 3.68 (d, *J* = 0.9 Hz, 3H), 2.45 – 2.31 (m, 1H), 2.26 (m, 1H), 2.12 (d, *J* = 1.9 Hz, 3H), 0.98 (d, *J* = 2.0 Hz, 3H), 0.86 (d, *J* = 6.4 Hz, 3H), 0.80 (s, 3H).

Step 5. Synthesis of DCA-3S



To a solution of S5 (52.8 mg, 0.1 mmol) in EtOH (3 mL), KOH (48 mg, 0.8 mmol) was added in H_2O (1 mL). The reaction mixture was stirred at 85°C for 16 hours. The organic solution was concentrated and purified by prep-HPLC to afford product DCA-3S (22 mg, 46%) as a white solid.

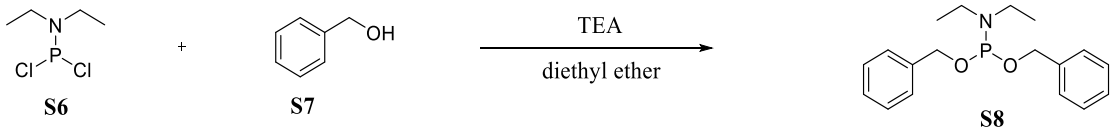
LC-HRMS: Rt=4.41, [M-H] =471.

^{1}H NMR (400 MHz, Methanol-d_4) δ 4.31 (m, 1H), 3.98 (d, J = 3.1 Hz, 1H), 2.37 (m, 1H), 2.22 (m, 1H), 1.04 (d, J = 6.3 Hz, 3H), 0.98 (d, J = 2.7 Hz, 3H), 0.74 (s, 3H).

^{13}C NMR (101 MHz, Methanol-d_4) δ 183.17, 80.55, 74.08, 48.61, 48.45, 47.60, 43.77, 37.44, 37.27, 36.38, 36.09, 35.23, 34.74, 34.60, 33.90, 29.83, 28.72, 28.69, 28.32, 27.40, 24.89, 23.58, 17.75, 13.23.

Synthesis of DCA-3P

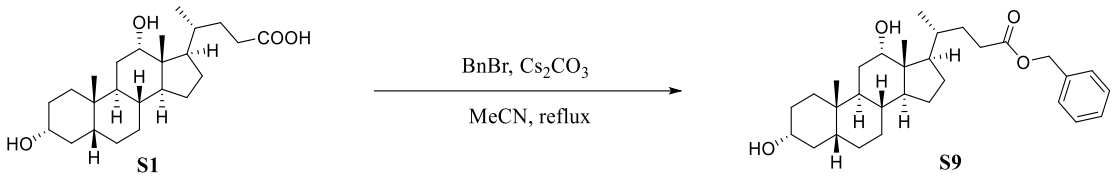
Step 1. Synthesis of compound S8



To a stirred solution of S6 (3 ml, 1.0 eq) in Et_2O (60 mL) was added a solution of S7 (4.27 mL, 2.0 eq) and TEA (6.30 mL, 2.2 eq) at 0 ℃ dropwise with vigorous stirring, and the mixture was warmed to room temperature and stirred at room temperature for 36 hours. The ammonium salt was removed by filtration, and solvents were removed under reduced pressure. The final product, S8, was afforded (3.13 g).

^{1}H NMR (400 MHz, Chloroform-d) δ 7.37 – 7.31 (m, 10H), 4.80 – 4.67 (m, 4H), 3.11 (m, 4H), 1.05 (t, J = 7.1 Hz, 6H).

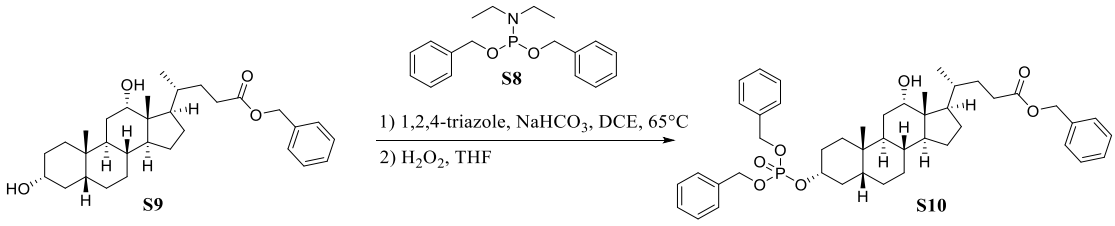
Step 2. Synthesis of compound S9



Cs_2CO_3 (2.44 g, 1.5 eq) and BnBr (3.0 mL, 5.0 eq) were added to a solution of S1 (1.96 g, 1.0 eq) in MeCN (32 mL), and the mixture was refluxed for 4 hours. After removal of the solvent, the residue was treated with saturated $NaHCO_3$ solution and extracted with ethyl acetate. The combined organic layers were washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. The crude product was purified by silica gel chromatography (0-1-2% MeOH in DCM) to give the desired product S9 (1.38 g).

^{1}H NMR (400 MHz, DMSO-*d*6) δ 7.41 – 7.28 (m, 5H), 5.07 (d, *J* = 3.8 Hz, 2H), 4.46 (d, *J* = 4.3 Hz, 1H), 4.20 (d, *J* = 4.0 Hz, 1H), 3.77 (d, *J* = 3.5 Hz, 1H), 3.36 (d, *J* = 4.7 Hz, 1H), 2.37 (m, 1H), 2.27 (m, 1H), 0.90 (d, *J* = 6.0 Hz, 3H), 0.84 (s, 3H), 0.56 (s, 3H).

Step 3. Synthesis of compound S10

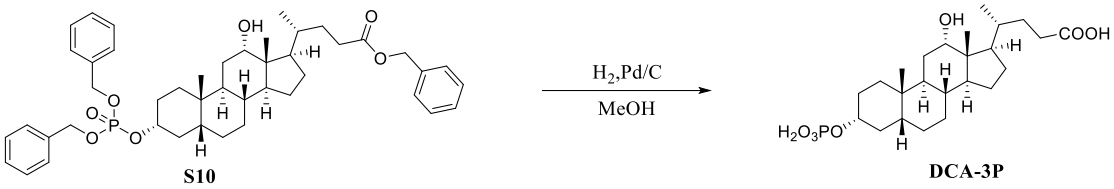


A stirred solution of S9 (748 mg, 1.0 eq), 1,2,4-triazole (225 mg, 2.1 eq), and $NaHCO_3$ (953 mg, 7.3 eq) in DCE (15 mL) was treated with S8 (0.6 mL, 1.02 eq), and the mixture was stirred at 65 °C overnight. After cooling in an ice bath, THF (6 mL) was added to the mixture, followed by a dropwise addition of 30% H_2O_2 (3 mL). After stirring for 5 min, saturated aq $Na_2S_2O_3$ (15 mL) was added slowly. (caution: highly exothermic reaction). The mixture was diluted with water (100 mL) and extracted with DCM. The combined organic layers were washed with brine, dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. The crude product was

purified by silica gel chromatography (20-30% EA in petroleum ether) to give the desired product S10 (920 mg).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.38 -7.30 (m, 15H), 5.11 (d, J = 3.2 Hz, 2H), 5.02 (m, 4H), 4.29 (m, 1H), 3.96 (d, J = 3.0 Hz, 1H), 2.40 (m, 1H), 2.33- 2.24 (m, 1H), 0.96 (d, J = 5.9 Hz, 3H), 0.88 (s, 3H), 0.64 (s, 3H).

Step 4. Synthesis of DCA-3P



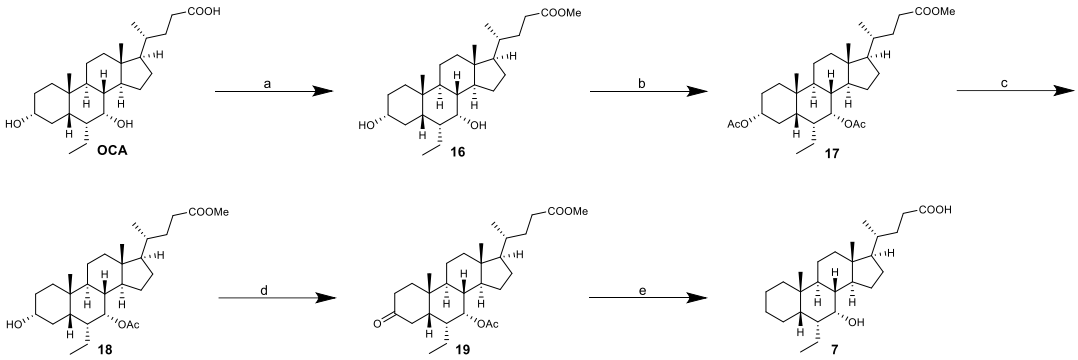
To a solution of S10 (342.6 mg, 1.0eq) in MeOH (15 mL), Pd/C (45 mg, 10%) was added. The flask was then evacuated, flushed three times with H_2, filled with H_2, and stirred at room temperature overnight. The reaction mixture was filtered through a pad of celite. The solvent was removed, and Na_2CO_3 (2.0 eq) dissolved in water was added, and the mixture was washed with ethyl acetate. The aqueous layer was separated, acidified with 1M HCl (pH⩽2), and extracted with $CHCl_3$ / i-PrOH (3:1). Water was added, and the solution was concentrated in vacuo until most of the organic solvent was removed. The resulting solution was lyophilized to afford the desired product as a white solid (302.4 mg).

LC-MS: Rt=5.317. [M-H] =471.

^{1}H NMR (400 MHz, Methanol-d_4) δ 4.18 (q, *J* = 4.8 Hz, ^{1}H), 3.97 (t, *J* = 3.0 Hz, 1H), 1.01 (d, *J* = 6.4 Hz, 3H), 0.94 (s, 3H), 0.71 (d, *J* = 2.5 Hz, 3H).

^{13}C NMR (101 MHz, Methanol-d_4) δ 178.26, 77.11, 77.06, 74.07, 48.12, 47.57, 43.68, 37.46, 36.70, 36.41, 35.77, 35.72, 35.23, 34.77, 32.32, 32.02, 29.90, 29.81, 29.77, 28.62, 28.36, 27.45, 24.87, 23.67, 17.56, 13.18.

Synthetic scheme for C7



Reagents and conditions: (a) p-toluenesulfonic acid monohydrate, MeOH, r.t., 2 h, quant.; (b) Ac_2O, DMAP, pyridine, 80 ℃, overnight; (c) conc.HCl, MeOH, r.t., overnight; (d) PCC, silica gel, DCM, r.t., overnight, 79% for 3 steps; (e) KOH, hydrazine hydrate, diethylene glycol, 120 ℃, 2 h, then 180 ℃, overnight, 78%.

Synthesis of compound 16

To a solution of OCA (10.00 g, 23.78 mmol) in MeOH (200 mL) was added p-toluenesulfonic acid monohydrate (1.80 g, 9.46 mmol), and the mixture was stirred at room temperature for 2 h. Most of the solvent was removed, and the residue was dissolved in DCM and washed with a saturated $NaHCO_3$ solution. The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo to give the desired product 16 (10.30 g, quant.). Spectroscopic data are by that previously reported[69].

Synthesis of compound 17

To a solution of 16 (10.30 g, 23.73 mmol) in pyridine (45 mL) was added Ac_2O (13.4 mL, 142.68 mmol) and 4-dimethylaminopyridine (289.0 mg, 2.37 mmol), and the mixture was stirred at 80 ℃ overnight. After cooling to room temperature, the solvent was removed, followed by the addition of saturated $NaHCO_3$ solution, and the mixture was extracted with EtOAc. The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo to give the desired product 17 (12.23 g), which was used in the next step without further characterization.

Synthesis of compound 18

To a solution of 17 (12.23 g, 23.61 mmol) in MeOH (120 mL) was added concentrated HCl (25 mL), and the mixture was stirred at room temperature overnight. The solvent was removed, followed by adding saturated $NaHCO_3$ solution, and the mixture was extracted with EtOAc. The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo to give the desired product 18 (11.25 g), which was used in the next step without further characterization.

Synthesis of compound 19

To a solution of 18 (11.25 g, 23.63 mmol) in DCM (400 mL) was added silica gel (21 g) and pyridinium chlorochromate (9.95 g, 46.16 mmol), and the mixture was stirred at room temperature overnight. The mixture was concentrated, and the crude product was purified by silica gel chromatography (10-20% EtOAc in petroleum ether) to give the desired product 19 (8.85 g, 79% for 3 steps).

Synthesis of C7

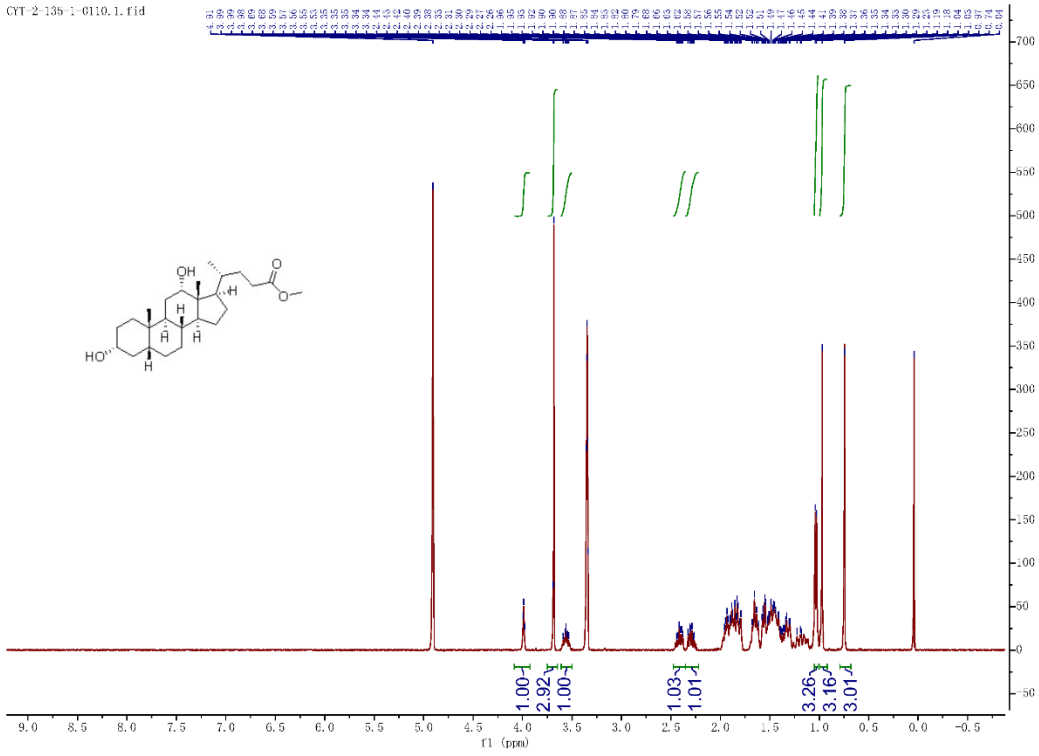
To a solution of 19 (1.50 g, 3.16 mmol) in diethylene glycol (60 mL) was added KOH (5.30 g, 94.46 mmol) and hydrazine hydrate (5.7 mL, 116.92 mmol), and the mixture was stirred at 120 ℃ for 2 h, followed by 180 ℃ overnight in a sealed tube. The mixture was cooled to room temperature, quenched with water, acidified by 1 M HCl, and extracted with EtOAc. The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. The crude product was purified by silica gel chromatography (30% EtOAc in petroleum ether) to give the desired product 7 (1.00 g, 78%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.72 – 3.68 (m, 1H), 2.44 – 2.34 (m, 1H), 2.30 – 2.20 (m, 1H), 0.95 – 0.86 (m, 9H), 0.65 (s, 3H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 180.35, 71.48, 55.93, 50.72, 47.11, 42.93, 41.52, 40.20, 39.83, 38.16, 36.47, 35.51, 33.51, 31.19, 30.96, 28.31, 27.81, 24.42, 24.15, 23.86, 22.36, 21.51, 20.92, 18.39, 11.94, 11.86.

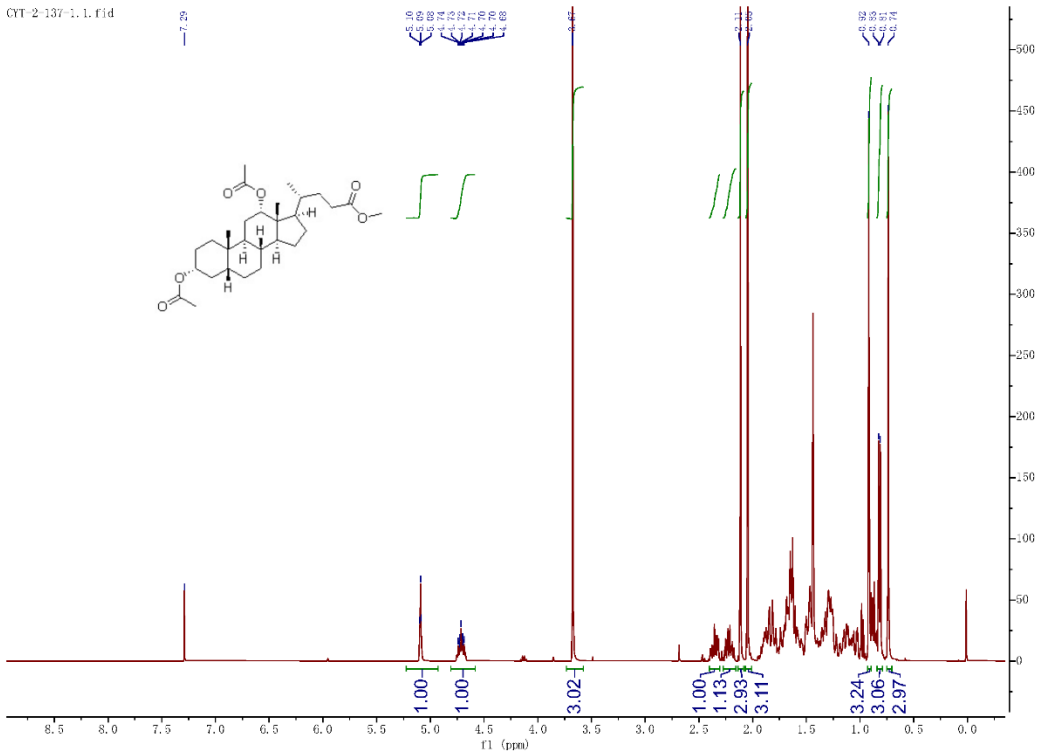
HRMS (ESI) [M + NH4]+ calculated for C26H48NO3: 422.3634, found: 422.3628.

NMR spectra

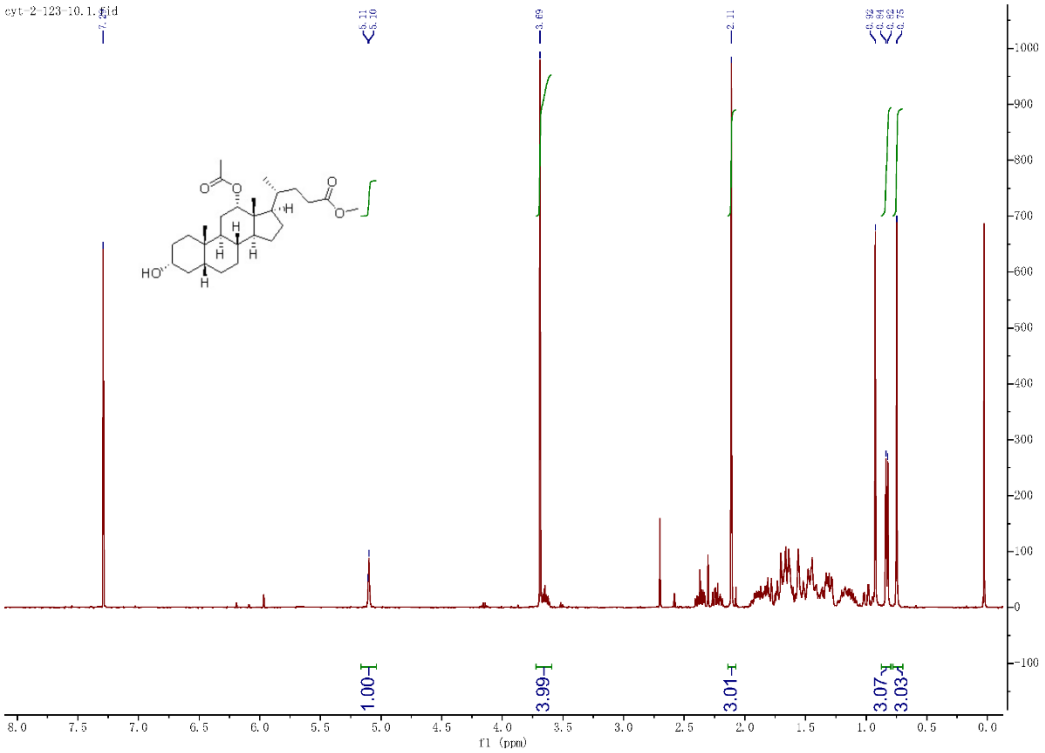
^{1}H-NMR of Compound S2 in Methanol-d_4



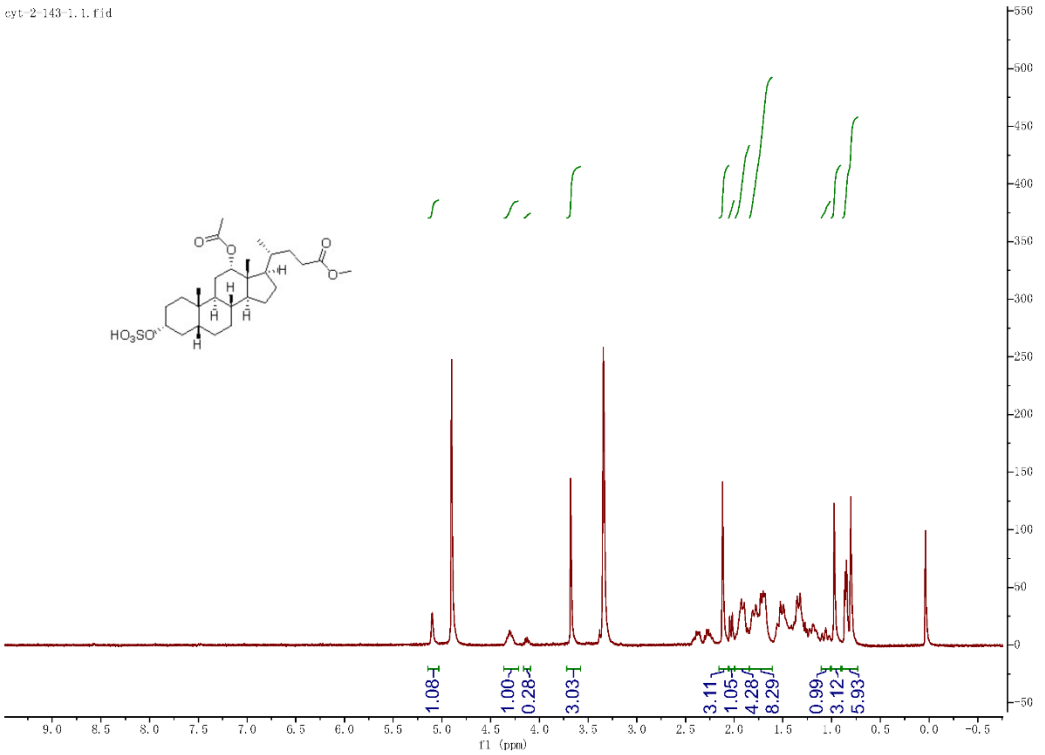
^{1}H-NMR of Compound S3 in Chloroform-d



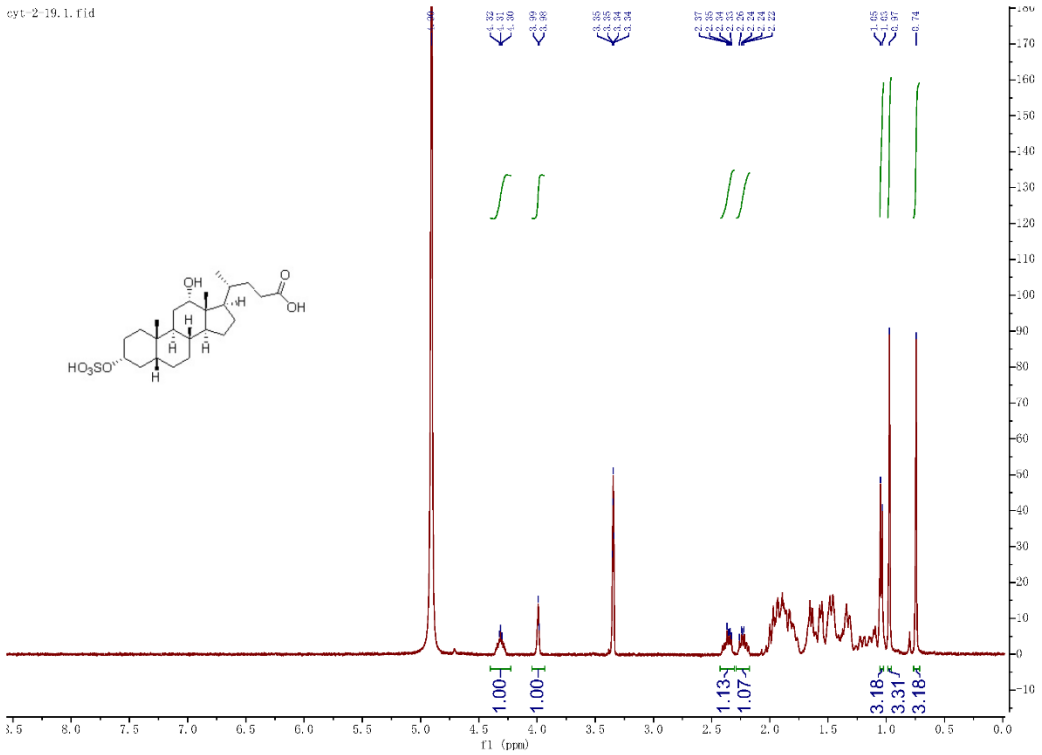
^{1}H-NMR of Compound S4 in Chloroform-*d*



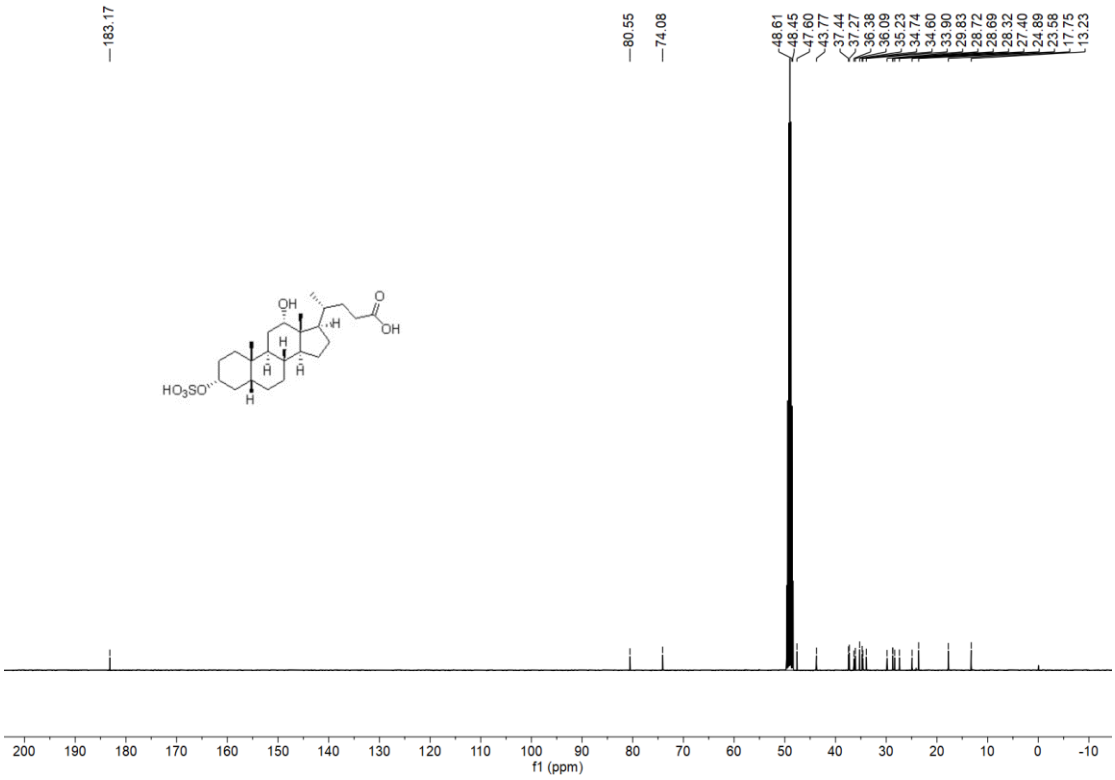
^{1}H-NMR of Compound S5 in Methanol-d_4



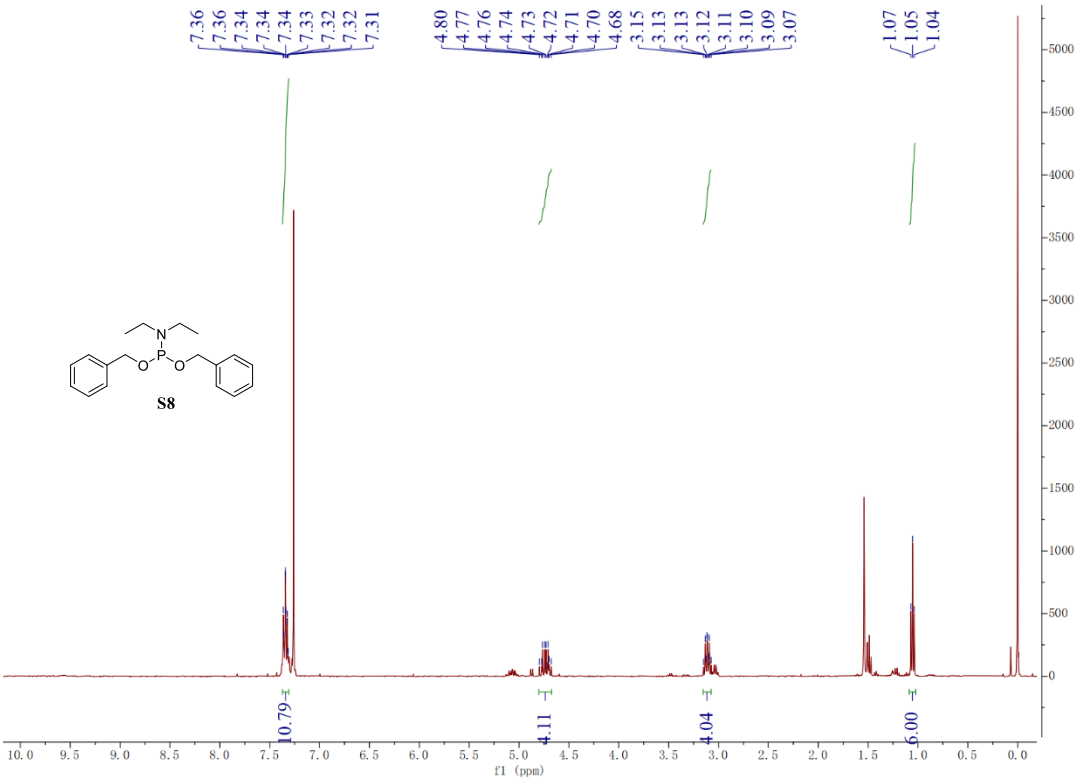
^{1}H-NMR of DCA-3S in Methanol-d_4



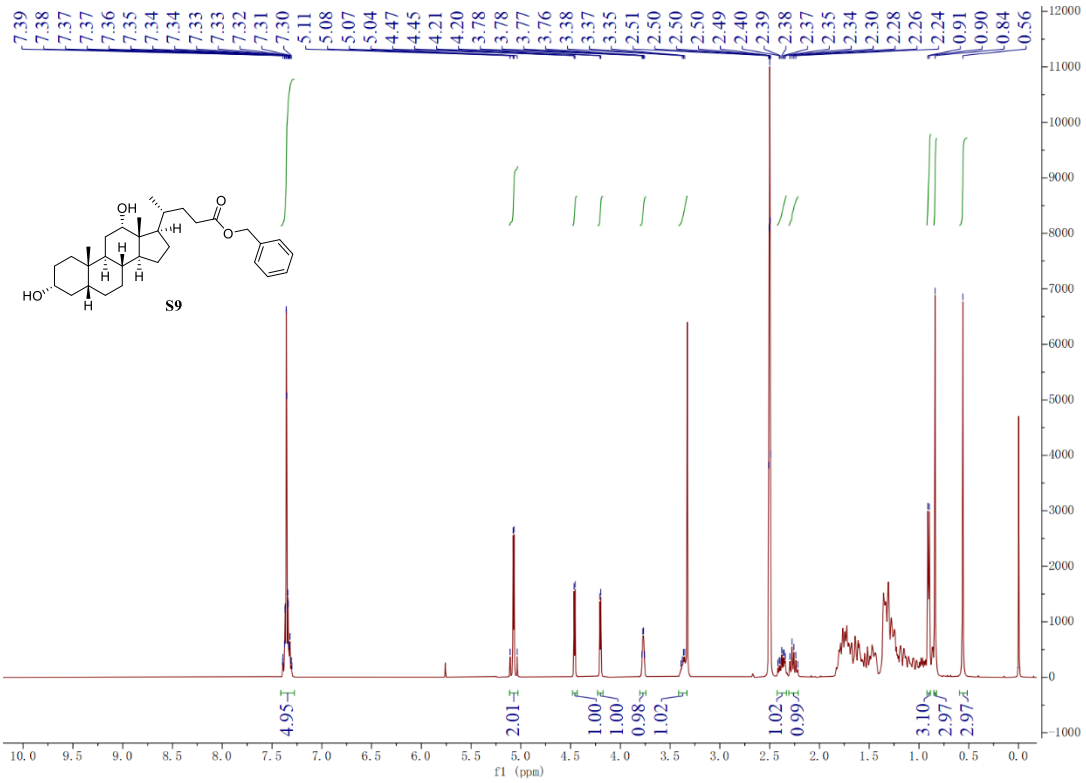
^{13}C-NMR of DCA-3S in Methanol-d_4



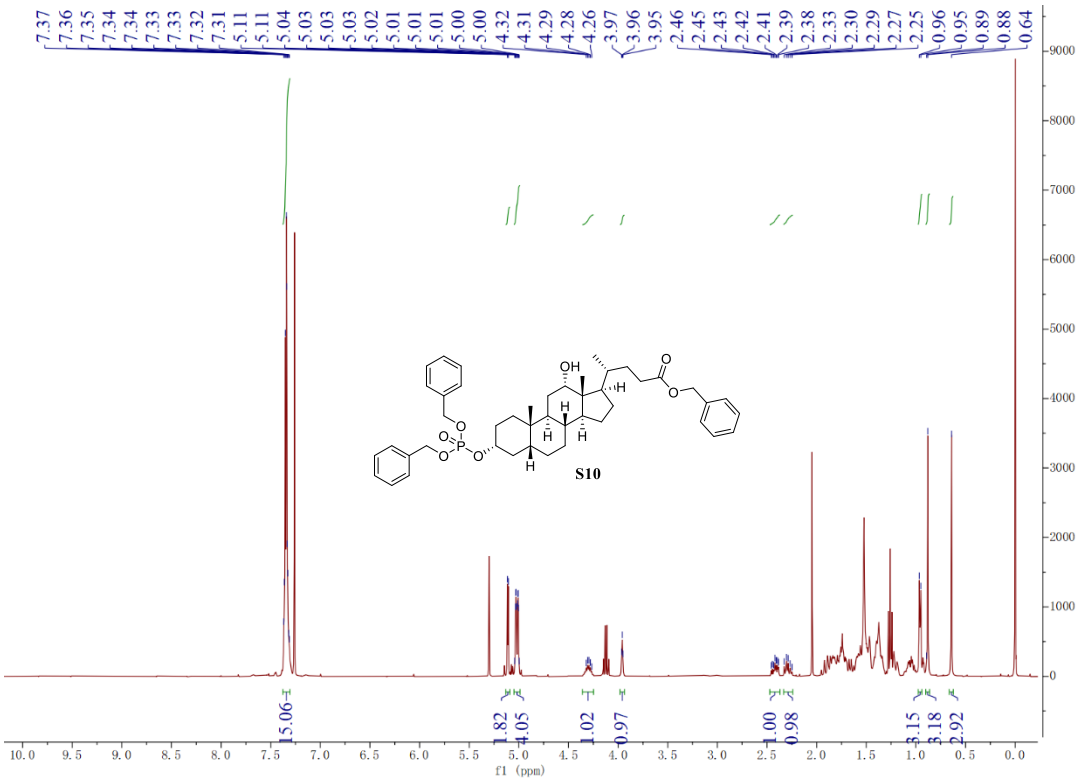
^{1}H-NMR of Compound S8 in Chloroform-*d*



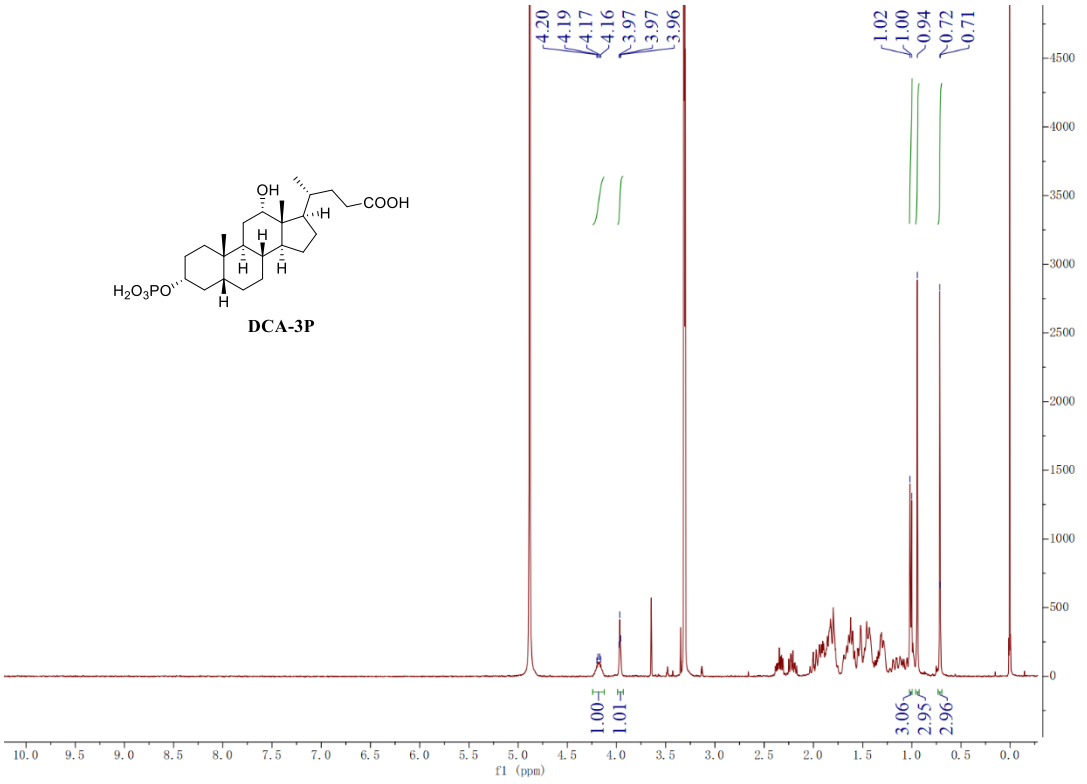
^{1}H-NMR of Compound S9 in DMSO-*d*6



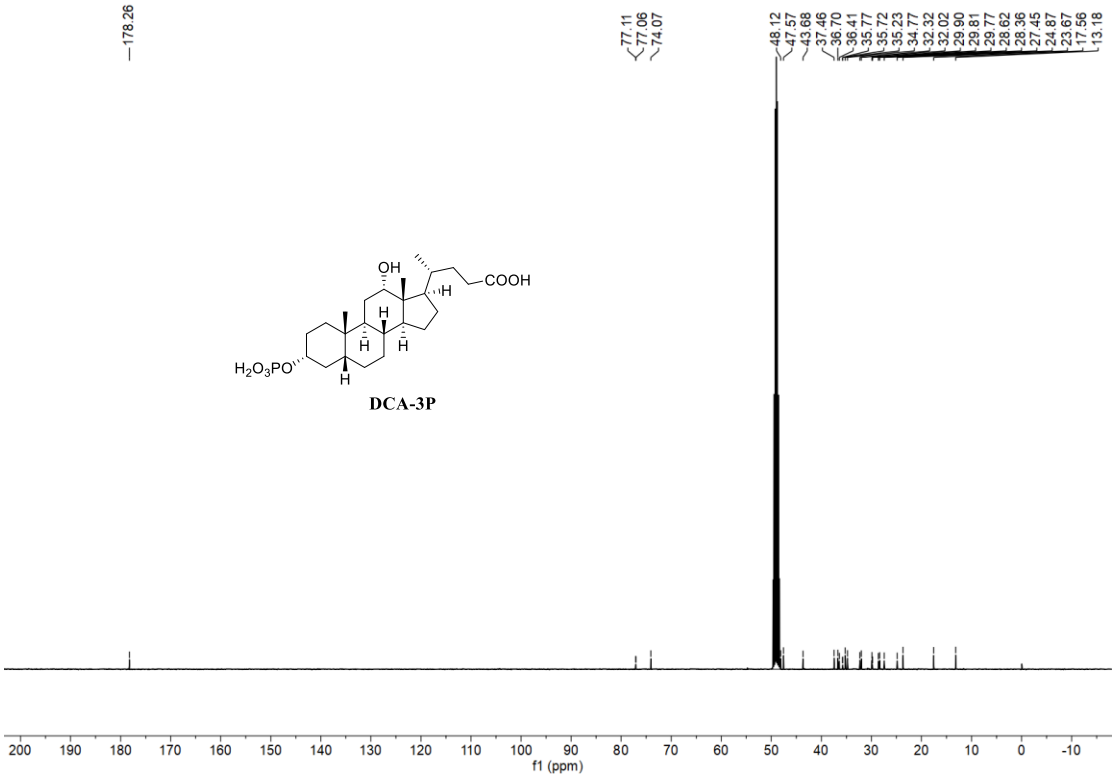
[1]H-NMR of Compound S10 in Chloroform-*d*

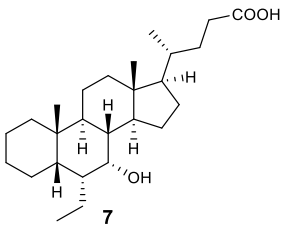


^{1}H-NMR of DCA-3P in Methanol-d_4



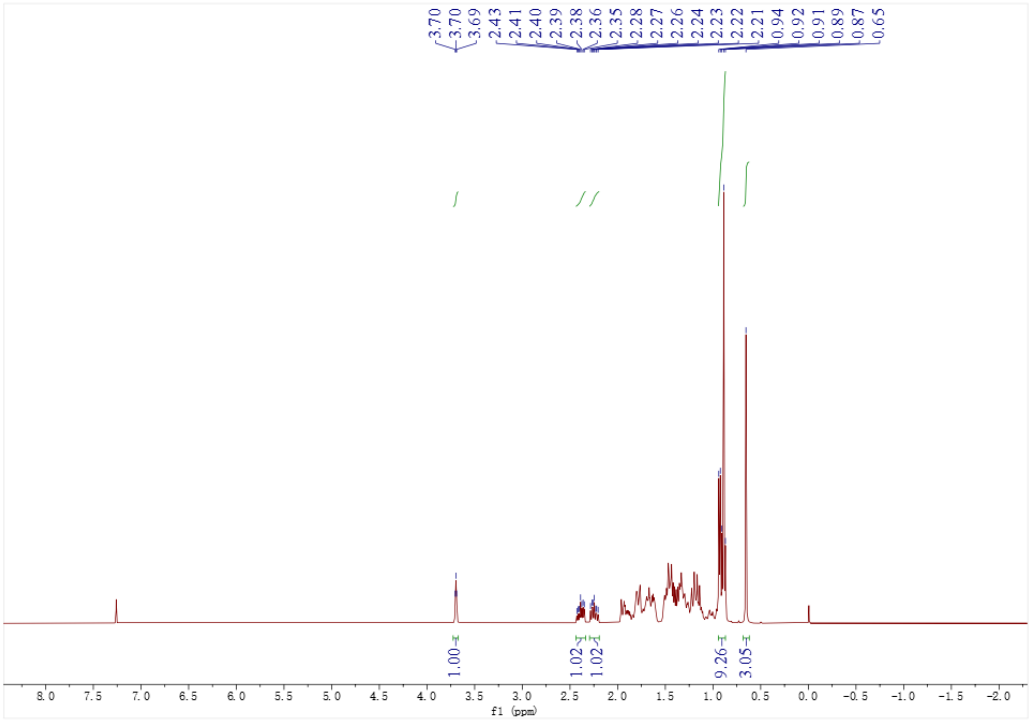
^{13}C-NMR of DCA-3P in Methanol-d_4





7

^{1}H NMR spectrum of C7:



^{13}C NMR spectrum of C7:

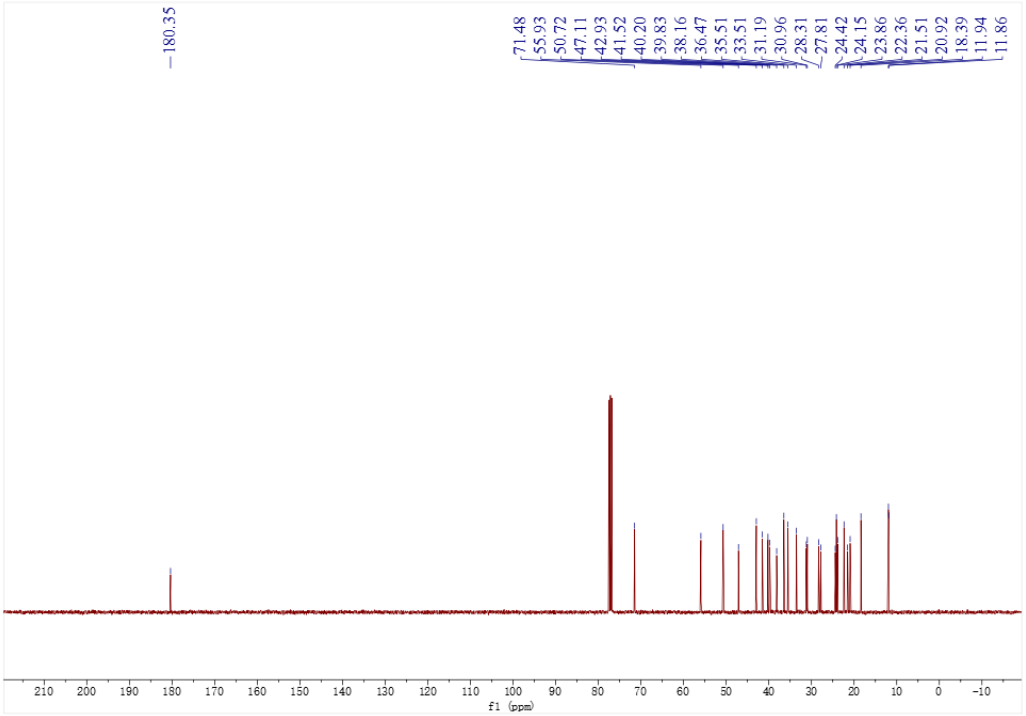
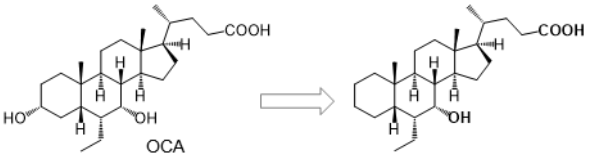


Table S1. The activity of OCA derivative C7. Related to Figure 5.



Compound	hX4		FXR	
	EC_{50} (μM)[a]	**Efficacy (%)[a]**	**EC_{50} (μM)[b]**	**Efficacy (%)[b]**
OCA	26±1.8	85±7.9	0.8±0.7	341±26
C7	>100	5.9±4.5	2.6±0.1	280±16

[a] FLIPR assay results, efficacy of DCA=100%; [b] NanoLuc assay results, efficacy of CDCA=100%. All data (mean ± sem, n=3)

Table S2. The primer list for qPCR. Related to STAR Methods.

Gene	Forward primers	Reverse primers
Human		
Shp	GAATATGCCTGCCTGAAAGG	TCCAGGACTTCACACAGCAC
Cyp7a1	GAGAAGGCAAACGGGTGAAC	GGTATGACAAGGGATTTGTGATGA
Mrp3	TCTACTTCCTCTGGCAGAACCT	TGATTAACTGGGGCAGCATGT
Bsep	GGGCCATTGTACGAGATCCTAA	TGCACCGTCTTTTCACTTTCTG
Abcg5	GATTGTCGTCCTCCTGGTGGAA	TCTCCGAAGCTCAGGATGGCAA
Abcg8	GGGCCACTCCCCAGGATAC	TGCTGGTCAGGTCCACATAGA
Fgf19	GGAGATCAAGGCAGTCGCTC	AGAGAACATGTCAGATTCCAAG
Lpin2	GAGTCCTGAGATCCAAAGAGA	CTCCGTTATCACCCAACTTC
Elov15	CCTTGGGCTAAAAGGTTTTCAA	GACCGTGATCTGGTGGTTGT
Gapdh	CAGCCTCAAGATCATCAGCA	GGTCATGAGTCCTTCCACGA
Rat		
Shp	TGGTACCCAGCTAGCCAAGG	TGTTCTTGAGGGTGGAAGCC
Cyp7a1	TGGATCAAGTGCAACTGAATGAC	GCACTGGAAAGCCTCAGAGC
Cyp8b1	CCAGATGCTGCACGTAGCC	GCATGGCCCGGTTGAG
Bsep	GCAAATTCCGCTGCCTATAGA	CCCTGAAAACGTGGCTGAA
Mdr2	TCCGAGCTCAACTTGGCATT	GAGACACGACACGGCTGTTGT
Cxcl1	CCAAACCGAAGTCATAGCCAC	ACGCCATCGGTGCAATCTATC
Cxcl2	CCAGACAGAAGTCATAGCCACTC	CCCAGGTCAGTTAGCCTTGC
Acta2	GCTCCATCCTGGCTTCTCTA	TAGAAGCATTTGCGGTGGAC
Col3a1	ACACCTGCTCCTGTCATTCC	AAGACCAGGGTCGCCATTTC

Timp1	TGTGGGAAATGCCACAGGTT	TTCCGTTCCTTAAACGGCCC
Acaca	CCTCCAACCTCAACCACTAC	AGCCTGTCATCCTCAATATCG
Gpat3	GGAGGATGAAGTGACCCAGA	CCAGTTTTTGAGGCTGCTGT
Lpin1	CCATTCACAGCGAGTCTTCA	TGGAAGGGGAATCTGACTTG
Gapdh	CCCTCAAGATTGTCAGCAATG	AGTTGTCATGGATGACCTTGG

Data S1. Bile acids' structure used in this study, information of cholestatic patients and plasma concentrations of bile acids in cholestatic patients with or without pruritus. Related to Figure 1 and S1.

Exhibit 1: The chemical structure of bile acids was analyzed in patient plasmas and used in activity determination, related to Figure 1.

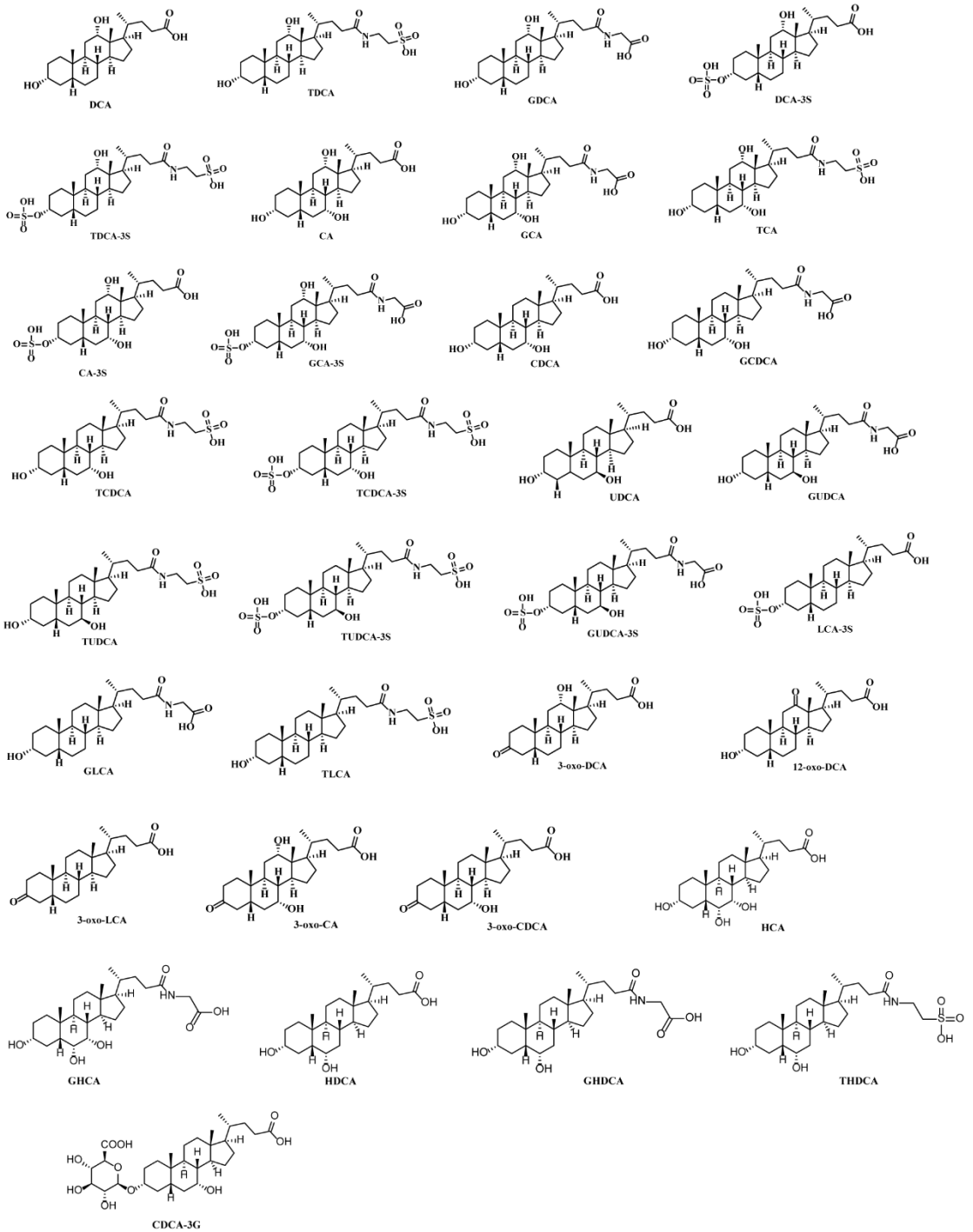


Exhibit 2: Information of cholestatic patients and the visual analogue scale (VAS) is related to Figure 1.

	Gender	Age	Underlying disease	Visual analogue scale (VAS)
Itch				
1	Male	62	Hepatitis B cirrhosis	7
2	Female	57	Hepatitis B cirrhosis	5
3	Female	69	Primary biliary cholangitis	7
4	Female	47	Liver cirrhosis	6
5	Male	70	Drug-induced liver injury	4
6	Male	42	IgG4 associated cholangitis	7
7	Male	36	Drug-induced liver injury	6
8	Male	70	Acute viral hepatitis E	7
9	Female	45	Drug-induced liver injury	6
10	Male	54	Liver transplantation	9
11	Male	43	Alcoholic cirrhosis	6
12	Male	76	Hilar cholangiocarcinoma	8
13	Male	51	Acute viral hepatitis E	7
14	Male	60	Alcoholic cirrhosis	5
Non-Itch				
1	Male	49	Hepatitis B cirrhosis	2
2	Male	45	Hepatitis B cirrhosis	1

3	Male	28	Chronic hepatitis B	0
4	Male	53	Hepatitis B cirrhosis	1
5	Male	44	Alcoholic cirrhosis	1
6	Male	68	Acute viral hepatitis E	2
7	Male	58	Hepatitis B cirrhosis	0
8	Female	68	Primary biliary cholangitis	3
9	Male	59	Drug-induced liver injury	1
10	Male	76	Drug-induced liver injury	2
11	Male	41	Hepatitis B cirrhosis	0
12	Female	72	Hepatitis B cirrhosis	2
13	Male	43	Hepatitis B cirrhosis	1
14	Male	48	Alcoholic cirrhosis	2
15	Male	63	Alcoholic cirrhosis	3
16	Male	58	Alcoholic cirrhosis	3
17	Male	56	Hepatitis B cirrhosis	2
18	Male	19	Cholangitis	3
19	Female	80	Drug-induced liver injury	2
20	Female	68	Primary biliary cholangitis	3

Exhibit 3: Plasma concentrations of bile acids in cholestatic patients with or without pruritus, related to Figure 1 and S1.

Bile Acids	Non-Itch (n=20) (nM)	Itch (n=14) (nM)	*p-value*	Significance
DCA	31.1 (0, 218.1)	76.4 (3.2, 675.8)	0.314	ns
TDCA	192.4 (41.5, 601.7)	279.7 (71.9, 529.1)	0.089	ns
GDCA	16.0 (1.7, 29.5)	16.2 (5.5, 26.3)	0.964	ns
DCA-3S	21.5 (0, 97.7)	12.6(0.8, 57.9)	0.242	ns
TDCA-3S	0 (0, 0)	0 (0, 0)	-	-
CA	105.2 (6.7, 418.7)	150.7 (13.8, 871.1)	0.443	ns
GCA	20316 (94.5, 32333.8)	21008 (252.1, 35131.2)	0.859	ns
TCA	53953 (7719.0, 242852.0)	65410 (6329.5, 235763.1)	0.593	ns
CA-3S	36.5 (5.3, 84.3)	46.34 (11.0, 89.3)	0.250	ns
GCA-3S	10.2 (0, 48.2)	26.7 (0, 84.8)	0.0725	ns
CDCA	984.4 (62.8, 4632.8)	820.2 (85.4, 4151.9)	0.745	ns
GCDCA	24768 (268.1, 39655.8)	25278 (878.3, 37293.8)	0.907	ns
TCDCA	72458 (401.8, 191932.8)	85031 (454.5, 192550.4)	0.0309	*
TCDCA-3S	9023 (3042.5, 15482.9)	12816 (7841.0, 29433.9)	0.018	*
UDCA	249.1 (24.8, 1132.6)	408.2 (57.0, 1402.1)	0.176	ns
GUDCA	111624 (4833.9, 340624.7)	251685 (4983.7, 776492.7)	0.0276	*
TUDCA	4846 (587.9, 11098.7)	7553 (2265.9, 28754.3)	0.116	ns
TUDCA-3S	1369 (403.2, 3402.9)	2739 (249.7, 8934.1)	0.037	*
GUDCA-3S	980.9 (38.2, 2963.7)	3804 (4.8, 11385.4)	0.003	**
LCA-3S	11.2 (0, 82.3)	6.5 (0, 34.1)	0.411	ns
GLCA	115.2 (2.6, 382.5)	125.5 (10.3, 396.0)	0.815	ns
TLCA	9.0 (2.1, 35.0)	12.0 (3.3, 40.2)	0.393	ns
HCA	5.3 (0.0, 20.7)	4.2 (0.0, 18.1)	0.664	ns
GHCA	117.2 (23.5, 420.6)	121.1 (33.3, 489.5)	0.354	ns
HDCA	1.2 (0.0, 10.2)	0.7 (0.0, 4.7)	0.572	ns
GHDCA	0.3 (0.0, 3.4)	0.6 (0.0,4.2)	0.783	ns

THDCA	3547.3 (217.5, 7862.1)	4295.9 (342.7, 8894.4)	0.231	ns
CDCA-3G	30.6 (0.1, 78.2)	40.7 (0, 57.3)	0.458	ns
Total BAs-3S	11449 (4022.5, 17335.3)	19445 (8401.7, 43166.2)	0.002	**
Total BAs	303571 (149619.7, 633918.7)	477856 (328554.4, 1108346.5)	0.034	*

Data S2. Single particle reconstruction, cryo-EM data validation statistics and structure alignment of $X4^{DCA\text{-}3P}$ with other agonists bound GPRs. Related to Figure 2, 3 and STAR Methods.

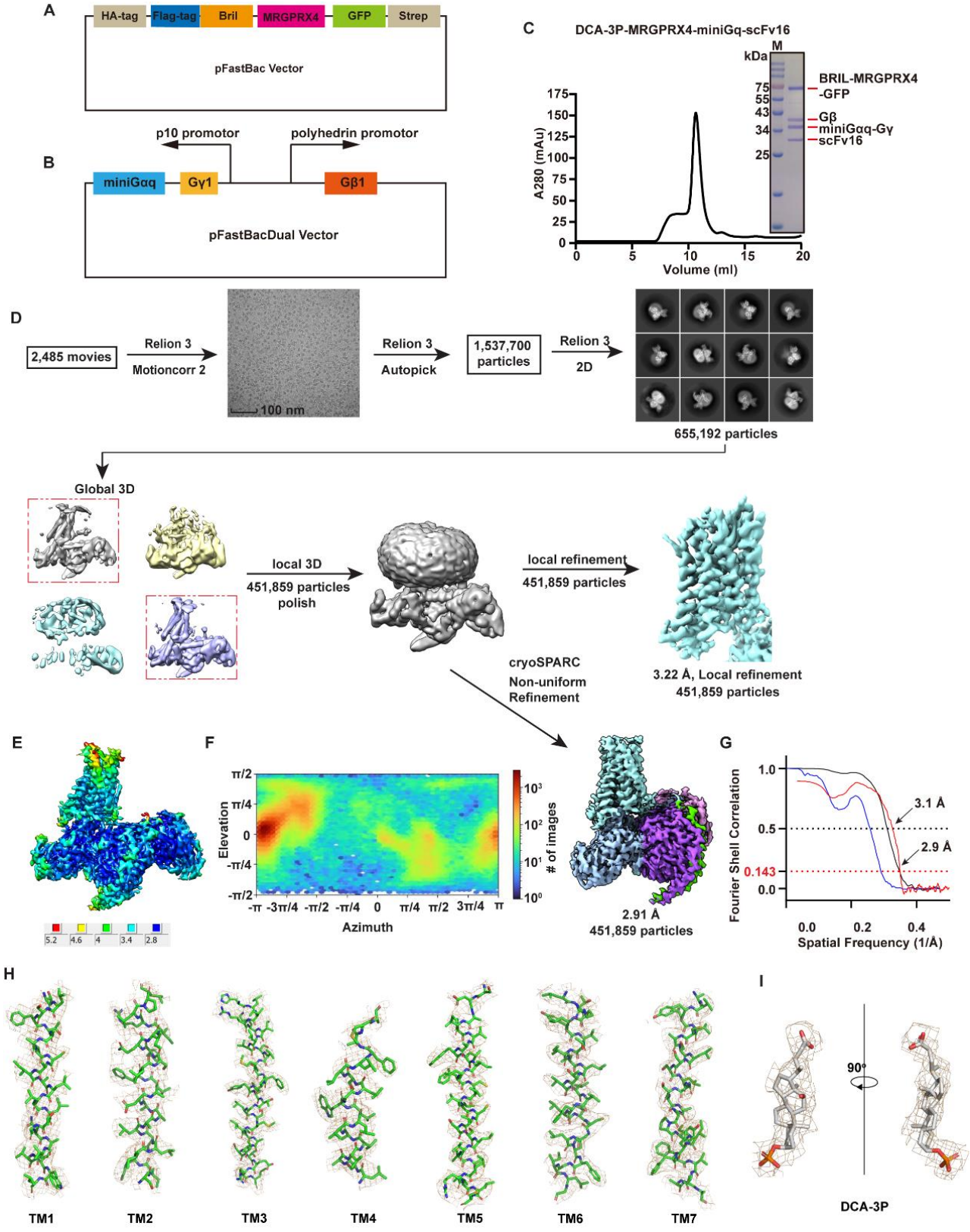


Exhibit 1: Cryo-EM images and single particle reconstruction are related to Figure 2.

(A-B) Constructs of hX4 **(A)** and G protein **(B)**.

(C) Purification of $X4^{DCA\text{-}3P}$ coupled with G protein.

(D) Flow chart of cryo-EM data processing of $X4^{DCA\text{-}3P}$.

(**E)** Local resolution map of the $X4^{DCA\text{-}3P}$.

(F) Particle angular distribution calculated in cryoSPARC for the final reconstruction.

(G) Fourier Shell Correlations (FSC) of the final full map of the $X4^{DCA\text{-}3P}$, calculated between two independently refined half-maps before (blue) and after (black) post-processing, overlaid with an FSC curve calculated between the cryo-EM density map and the structural model shown in red.

(H-I) Electron microscopy density maps of the TM helices of hX4 **(H)** and ligand DCA-3P **(I)**.

Exhibit 2: Cryo-EM data collection, model refinement, and validation statistics are related to Figure 2.

	MRGRPX4-DCA-3P (8K4S)
Data collection and processing	
Magnification	130,000
Voltage (kV)	300
Electron exposure (e-/Å2)	60
Defocus range (μm)	-1.1-2.2
Pixel size (Å)	1.04
Symmetry imposed	C1
Initial particle projections (no.)	1,537,700
Final particle projections (no.)	451,859
Map resolution (Å)	2.91
FSC threshold	0.143
Map resolution range (Å)	2.88-3.40
Refinement	
Initial model used (PDB code)	7S8P
Model resolution (Å)	3.1
FSC threshold	0.5
Model resolution range (Å)	3.07-3.22
Map sharpening B factor (Å2)	98.1
Model composition	
Non-hydrogen atoms	8196
Protein residues	1094
Ligand	32
B factors (Å2)	
Protein	45.8
Ligand	107
R.m.s. deviations	
Bond lengths (Å)	0.004
Bond angles (°)	1.032
Validation	
MolProbity score	2.43
Clashscore	11.51
Rotamer outliers (%)	4.46
Ramachandran plot	
Favored (%)	94.79
Allowed (%)	5.21
Disallowed (%)	0.00

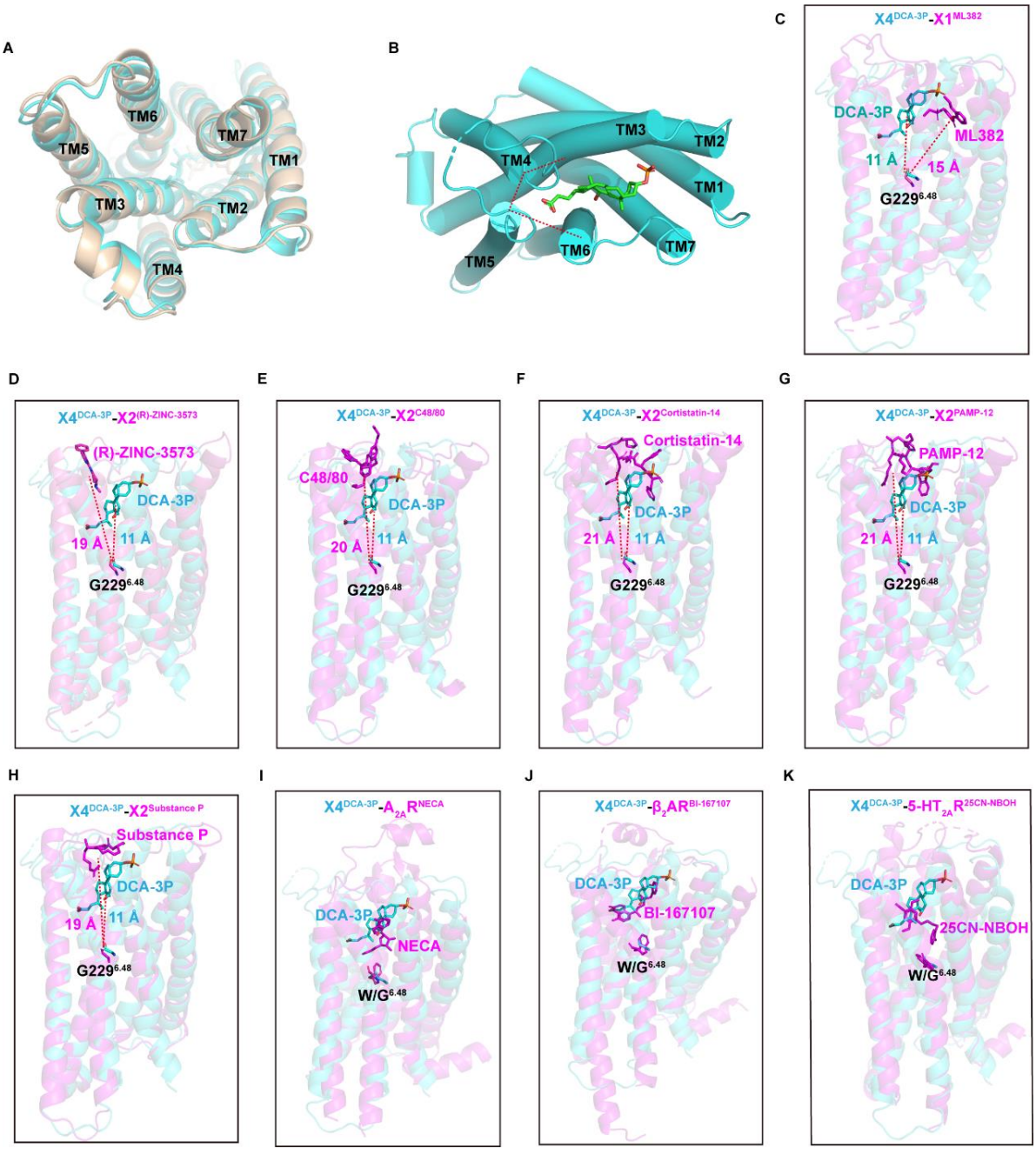


Exhibit 3: Structural alignment of X4^{DCA-3P} and other agonists bound GPRs, related to Figure 2.

(A) No noticeable conformation changes were observed in the intracellular side of X4^{DCA-3P} (cyan) and X4^{MS47134} (wheat).

(B) P-Head and C-Tail of DCA-3P were surrounded by TM1-TM2-TM3 and TM4-TM5-TM6 of MRGPRX4, respectively.

(C-H) DCA-3P is bound in a completely different pocket closed to G2296.48 in hX4,

while other ligand pockets in X1ML382 (C), X2(R)-ZINC-3573 (D), X2C48/80 (E), X2Cortistatin-14 (F), X2PAMP-12 (G) and X2Substance-P (H) are closed to the solvent.

(I-K) DCA-3P occupied a similar pocket with $A_{2A}R^{NECA}$ **(I)**, $\beta_2AR^{BI167107}$ **(J)**, and 5-$HT_{2A}R^{25CN\text{-}NBOH}$ **(K)**.

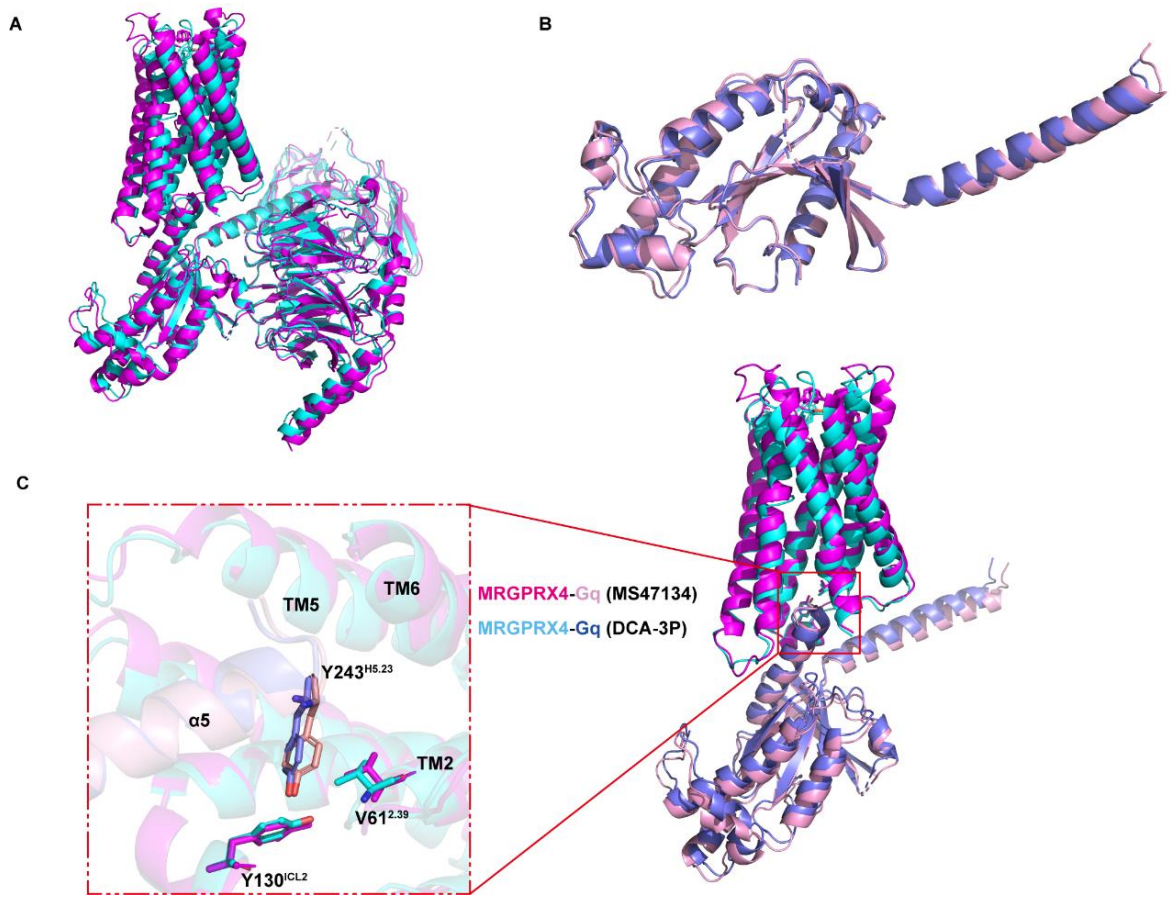


Exhibit 4: The structural alignment of $X4^{DCA\text{-}3P}$ with X4MS1347 is related to Figure 3.

No noticeable conformation changes were observed in overall structure **(A),** Gq protein **(B),** and interactions between Gq and X4 **(C)** of $X4^{DCA\text{-}3P}$ with $X4^{MS1347}$.

Data S3. EC_{50} and E_{max} of DCA-3P to MRGPRX4 or its variants and MM/GBSA binding affinity calculation results. Related to Figure 3.

Exhibit 1: EC_{50} and E_{max} of DCA-3P to MRGPRX4 or its variants, related to Figure 3.

Variants	EC_{50} (μM)	E_{max} ± s.e.m. (% WT)
WT	0.044	100 ± 1.7
R82A	17.9	25.2 ± 4.0
R86A	3.85	65.0 ± 1.7
R95A	36.2	15.4 ± 1.2
K96A	60.2	29.5 ± 1.6
V99A	1.79	35.2 ± 2.5
T103A	0.63	87.4 ± 4.9
I239A	0.14	41.2 ± 1.6
L235A	14.7	121.8 ± 5.9
E157A	12.1	28.7 ± 1.9
F232A	30.8	17.8 ± 1.7
L246A	20.9	17.0 ± 1.9
Y254A	41.5	122.3 ± 1.8
R241A	2.3	36.3 ± 1.5

Exhibit 2: MM/GBSA binding affinity calculation results are related to Figure 3.

Residues	MM/GBSA binding affinity (kcal/mol)					
	DCA-3P			DCA-3S		
	1	2	3	1	2	3
R95	-12.91	-15.70	-11.39	-4.43	-0.16	-4.32
R86	-13.97	-1.00	-0.05	-1.03	-4.16	-0.16
R82	-9.40	-14.68	-6.97	-1.23	-3.77	-6.19
R241	-9.68	-8.59	-7.83	-0.02	-0.01	-6.34
G236	-2.48	-2.19	-2.24	-1.76	-0.72	-1.42
L235	-2.53	-1.50	-1.66	-0.76	-0.87	-1.12
F232	-1.99	-0.62	-0.77	-1.19	-1.60	-0.67
V99	-1.57	-1.43	-1.50	-1.91	-1.52	-1.36
Y240	-1.11	-2.68	-2.40	-0.13	-0.05	-2.12
Total	-72.20	-73.10	-44.50	-29.60	-31.40	-54.70

Data S4. Mouse PK data of C7 and OCA and transportation of OCA and C7 by ASBT, NTCP and OSTα/β. Related to Figure 6-7.

Exhibit 1: Mouse PK data of C7 and OCA, related to Figures 6 and 7.

Compound	Intravenous [a]			Oral gavage [b]		
	CL (mL/h/kg)	V_d (mL/kg)	$T_{1/2}$ (h)	AUC (ng/mL*h)	Cmax (ng/mL)	F (%)
C7	0.69±0.03	0.86±0.09	0.86±0.05	6273±1045	4160±525	16.6±1.3
OCA	3.67±1.09	23.9±9.92	4.24±0.51	420±97	385±44	4.8±0.37

[a] Intravenous: 5 mg/ kg in saline with 10% DMSO, 10% solutol. [b] Oral gavage: 30 mg/ kg in water with 0.5% CMC-Na, 0.1% Tween80.

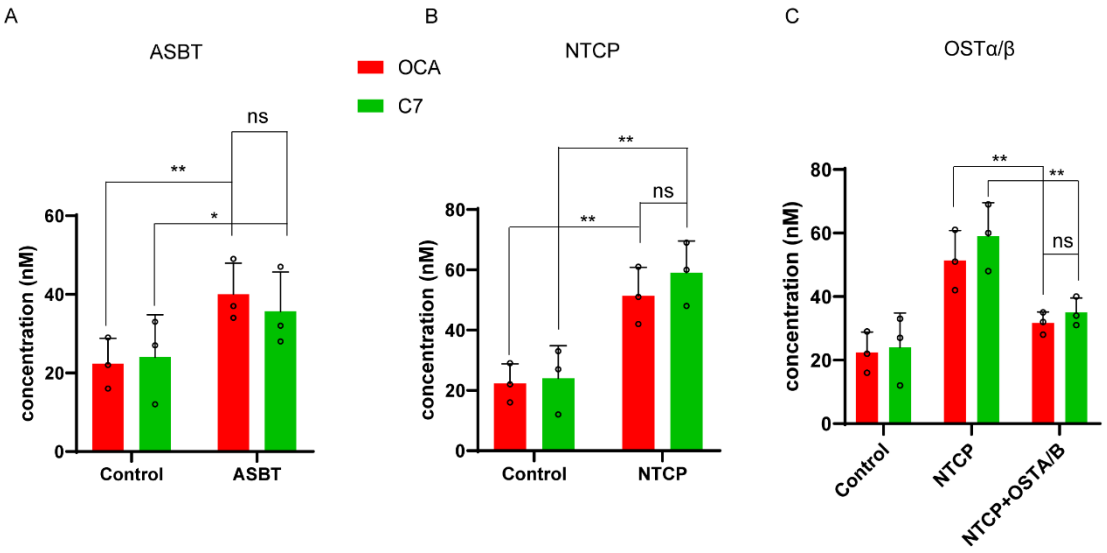


Exhibit 2: Transportation of OCA and C7 by ASBT (A), NTCP (B) and OSTα/β (C), related to Figure 6.

(A-B) OCA and C7 can be transported from media to cells by ASBT and NTCP. HEK293T cells were plated into 10 cm plates and transfected with plasmids pEG2-GFP as control or pEG2-ASBT-GFP or pEG2-NTCP-GFP. OCA or C7 was added to the cells at the final concentration of 10 μM. After 40 min. The cells were collected and the concentration of OCA or C7 was detected using HPLC-MS/MS. n=3 per group.

(C) OCA and C7 can be transported from cells to media by OSTα/β. HEK293T cells were plated into 10 cm plates and transfected with plasmids pEG2-GFP as control or co-transfected with plasmids pEG2-NTCP-GFP and pEG2- OSTα/β-GFP. OCA or C7 was added to the cells at the final concentration of 10 μM. After 40 min. The cells were collected and the concentration of OCA or C7 was detected using HPLC-MS/MS. n=3 per group.