Research Paper

Journal of Chemical Research

Journal of Chemical Research
1–9

DOI: 10.1177/1747519819875858
journals.sagepub.com/home/chl
SAGE

Synthesis and characterization of new impurities in obeticholic acid

Wei-Dong Feng[1], Song-Ming Zhuo[2] and Fu-Li Zhang[1,3]

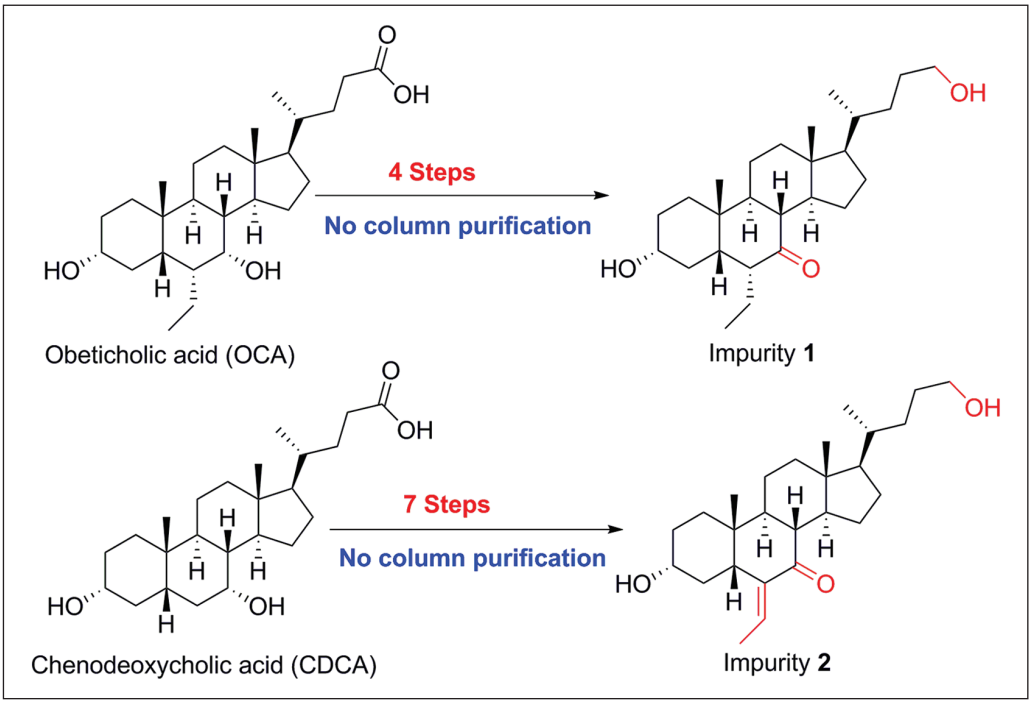
Abstract

Novel and efficient synthetic strategies are developed for the first synthesis of two new impurities found in obeticholic acid. The synthetic routes to the impurities are designed without column purification using 4-nitrobenzoyl chloride as a selective protecting group. The impurities, which are obtained in good yields and high purity, are identified and characterized using high-resolution mass spectrometry, Fourier transform infrared, one-dimensional nuclear magnetic resonance (^{1}H, ^{13}C, distortionless enhancement by polarization transfer), and two-dimensional nuclear magnetic resonance (Correlated Spectroscopy, heteronuclear single quantum coherence, heteronuclear multiple bond correlation, and rotating-frame Overhauser effect spectroscopy) techniques.

Keywords

bile acids, farnesoid X receptor, impurities, nuclear magnetic resonance, obeticholic acid

Date received: 26 March 2019; accepted: 21 August 2019



[1]Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, Zhejiang University of Technology, Hangzhou, P.R. China
[2]Department of Pharmaceutical Engineering, China Pharmaceutical University, Nanjing, P.R. China
[3]Shanghai Institute of Pharmaceutical Industry, China State Institute of Pharmaceutical Industry, Shanghai, P.R. China

Corresponding author:
Fu-Li Zhang, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, Zhejiang University of Technology, 18 Chaowang Road, Hangzhou 310032, P.R. China.
Email: zhangfuli01@126.com

Introduction

Obeticholic acid (OCA) is a farnesoid X receptor (FXR) agonist for the treatment of primary sclerosing cholangitis (PSC), primary biliary cholangitis (PBC), and non-alcoholic steatohepatitis (NASH).[1,2] OCA was approved by the Food and Drug Administration (FDA) under the brand name Ocaliva in 2016 for patients not responding to or who are intolerant to ursodeoxycholic acid (UDCA).[3]

The safety of a drug is dependent not only on the toxicological properties of the bulk drug itself but also on the impurities in the bulk drug, and the presence of these unwanted impurities, even in small quantities, may influence the efficacy and safety of the pharmaceutical product.[4] Thus, the detection, identification, synthesis, quantification, and control of these impurities, which originate in the manufacturing process, have become important elements of drug development to ensure product quality and ultimately patient safety.[5] According to the international conference on harmonization (ICH) guidelines, impurities in drug substances should be identified and controlled when present at a level of ⩾0.1% by isolating or synthesizing them in their pure form for analytical method development.[6] The process-related impurities of OCA have been previously reported in Douša et al.[7] During the analysis of laboratory batches of OCA, two unknown impurities of OCA at 0.13% and 0.10%, respectively, were detected by high-performance liquid chromatography (HPLC) using an evaporative light-scattering detector (ELSD). The impurities were isolated from the crystal mother liquor of OCA and identified using high-resolution mass spectrometry (HRMS) and nuclear magnetic resonance (NMR) techniques (Figure 1). To the best of our knowledge, impurities **1** and **2** have not been previously reported. We have successfully synthesized these impurities without column purification for the first time.

Results and discussion

Formation of impurities **1** *and* **2**

The synthetic route to OCA is shown in Scheme 1.[8,9] OCA is synthesized using a sequence of reactions involving oxidation, esterification, Mukaiyama aldol addition, hydrolysis, Pd/C catalytic reduction, and $NaBH_4$ reduction. The mechanisms of formation of impurities **1** and **2** involve the carboxyl moieties of unreacted intermediates **7** and **6**, respectively, being partially reduced by $NaBH_4$ in the last step.

Design of synthetic routes to impurities **1** *and* **2**

The synthetic routes to impurities **1** and **2** are shown in Schemes 2 and 3, respectively. The reduction of a carbonyl is far easier than that of a carboxyl under typical reaction

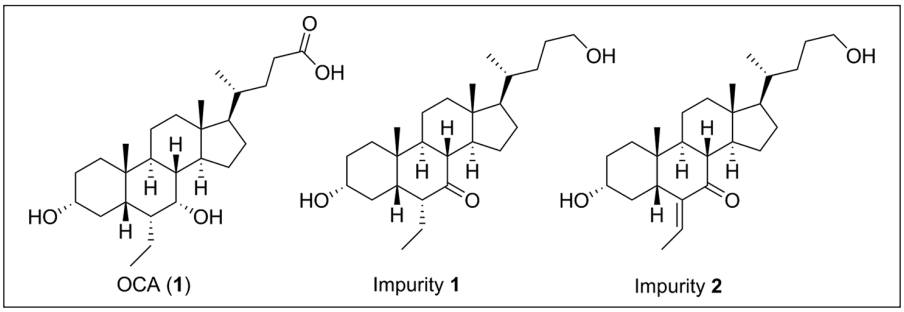
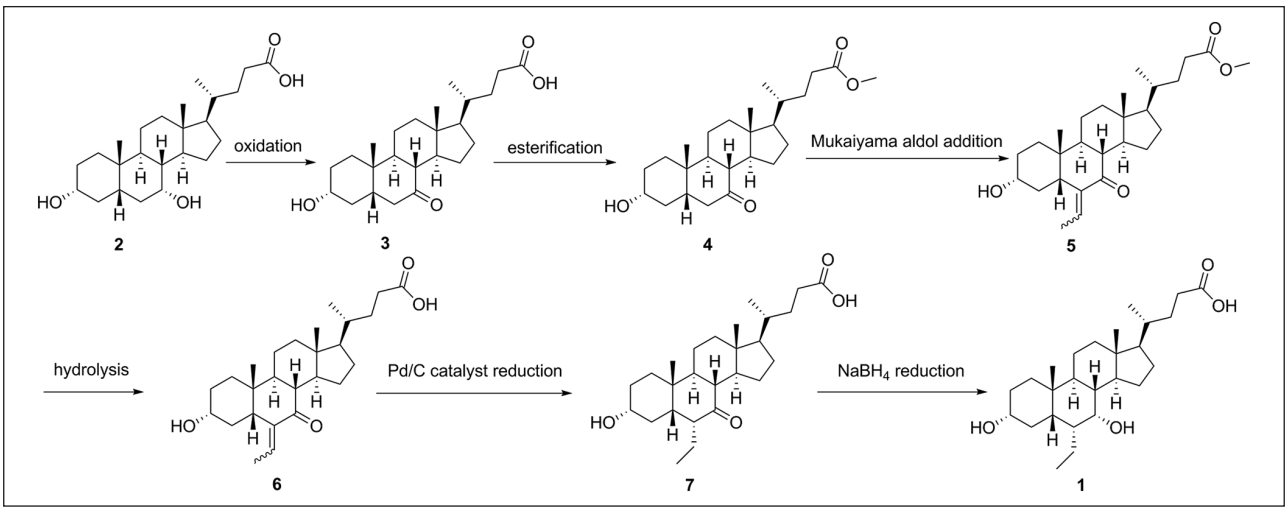
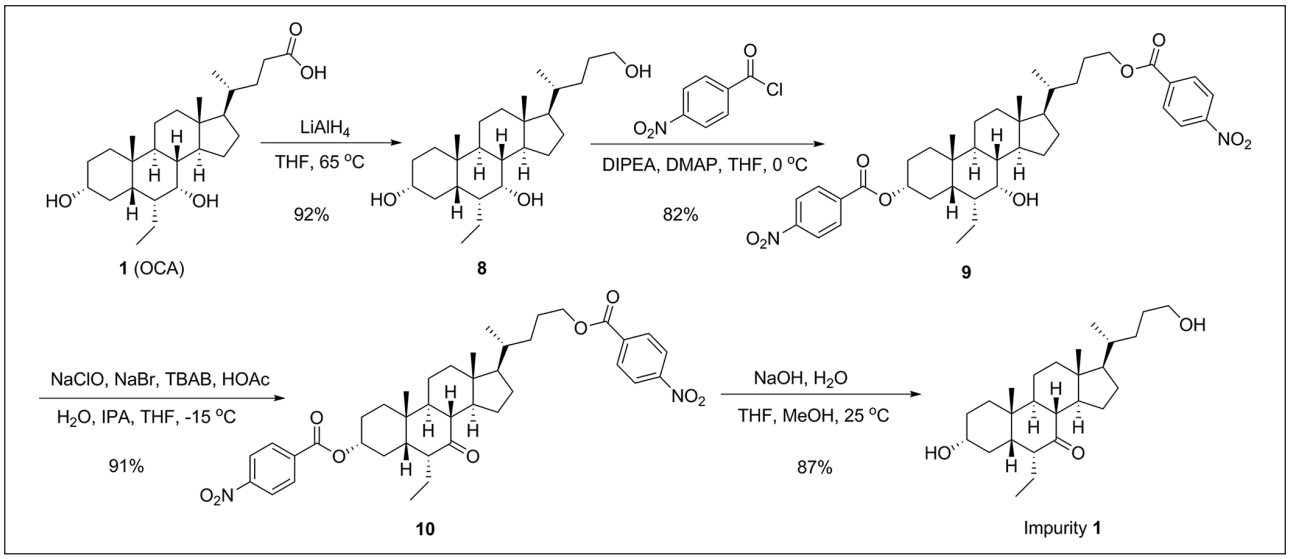


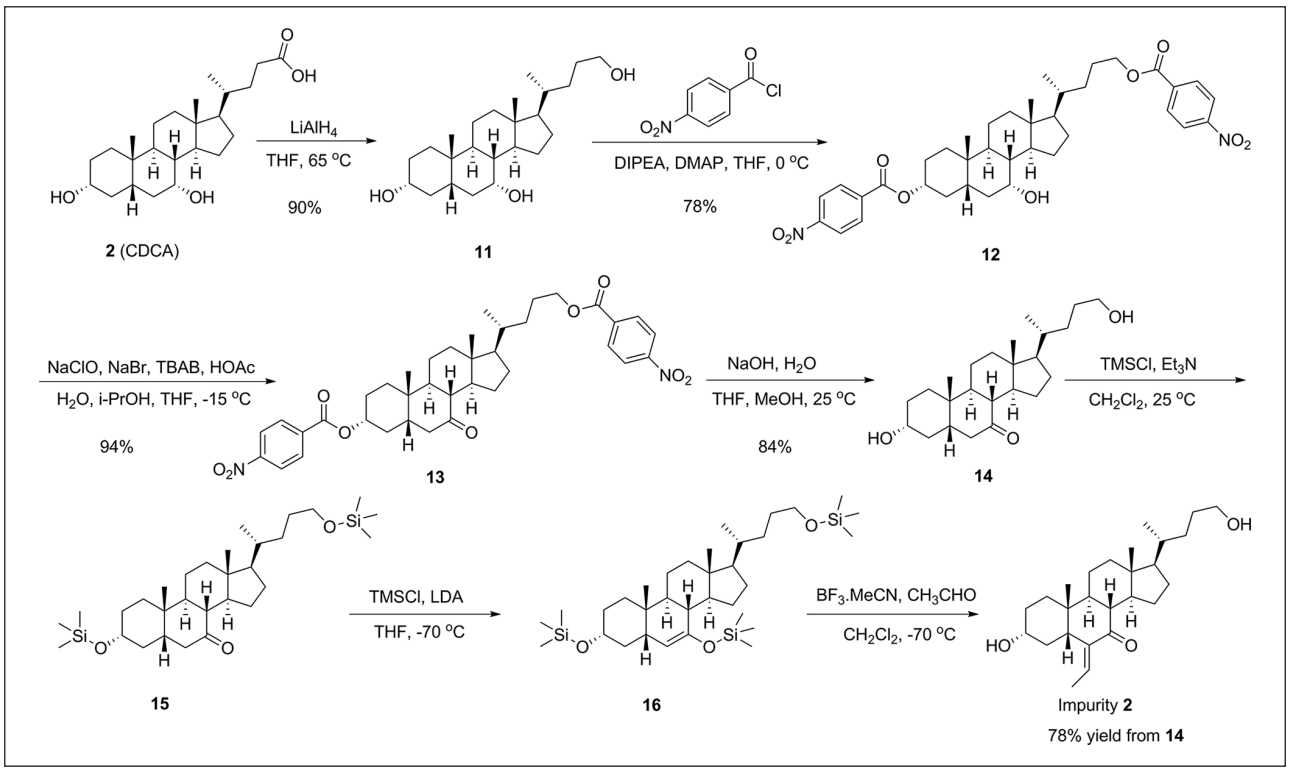
Figure 1. Structures of OCA, impurity **1**, and impurity **2**.



Scheme 1. Commercial process route to OCA (**1**).



Scheme 2. Synthetic route to impurity **1**.



Scheme 3. Synthetic route to impurity **2**.

conditions, and therefore, impurity **1** is difficult to obtain directly by reduction of **7**. In Scheme 2, impurity **1** can be obtained directly by oxidation of **8**, but purification by recrystallization or column chromatography is challenging due to the production of byproducts resulting from the overoxidation of **8**. Hence, the primary hydroxy and hydroxy at C-3 of **8** should be protected to avoid oxidation. Furthermore, the protecting group should be chemoselective and not react with the C-7 hydroxy group. To our delight, *p*-nitrobenzoyl chloride is an appropriate hydroxy protecting group. The primary hydroxy and C-3 hydroxy groups of **8** were, thus, protected with *p*-nitrobenzoyl chloride to produce **9**, which was oxidized with sodium hypochlorite and acetic acid to afford **10**. This was followed by hydrolysis to yield impurity **1**. Intermediates **9** and **10** and impurity **1** are easily crystallized from appropriate solvents. The synthetic route to impurity **2** was designed as shown in Scheme 3. The hydroxy groups of intermediate **11** must also be selectively protected using *p*-nitrobenzoyl chloride to produce **12**, followed by oxidation to yield **13**, which is then hydrolyzed to give compound **14**. Subsequently, intermediate **14** was protected with trimethylsilyl chloride to afford **15**. Finally, impurity **2** was obtained from **15** using the Mukaiyama aldol addition

Table 1. ^{1}H NMR (CD_3OD, 600 MHz) and ^{13}C NMR (CD_3OD, 150 MHz) assignments for impurities **1** and **2**.

Position	^{1}H	^{13}C	DEPT	^{1}H	^{13}C	DEPT
1.	1.21 (m), 1.85 (m)	35.25	CH_2	1.93 (m), 1.20 (m)	35.41	CH_2
2.	1.22 (m), 1.63 (m)	30.53	CH_2	1.69 (m), 1.29 (m)	30.46	CH_2
3.	3.46 (sept, J = 4.8 Hz)	71.66	CH	3.59 (sept, J = 4.8 Hz)	70.96	CH
4.	0.79 (m), 1.64 (m)	32.57	CH_2	1.55 (m), 1.26 (m)	38.37	CH_2
5.	1.86 (m)	52.16	CH	2.68 (dd, J = 13.2, 4.2 Hz)	46.88	CH
6.	2.82 (q, J = 5.4 Hz)	53.29	CH	–	145.38	C
7.	–	215.48	C	–	207.64	C
8.	2.49 (t, J = 11.4 Hz)	51.19	CH	2.32 (t, J = 11.4 Hz)	50.18	CH
9.	1.79 (td, J = 12, 4.8 Hz)	45.36	CH	1.99 (m)	40.74	CH
10.	–	36.82	C	–	35.77	C
11.	1.53 (m)	22.97	CH_2	1.49 (m)	22.45	CH_2
12.	2.04 (dt, J = 13.2, 3 Hz), 1.14 (m)	40.43	CH_2	2.08 (dt, J = 13.2, 2.4 Hz), 1.16 (m)	40.32	CH_2
13.	–	43.76	C	–	44.66	C
14.	1.44 (m)	50.49	CH	1.42 (m)	52.03	CH
15.	0.95 (m), 2.12 (m)	25.62	CH_2	1.44 (m), 2.34 (m)	27.03	CH_2
16.	1.30 (m), 1.92 (m)	29.42	CH_2	1.28 (m), 1.89 (m)	29.53	CH_2
17.	1.15 (m)	56.51	CH	1.12 (m)	56.18	CH
18.	0.71 (s)	12.52	CH_3	0.69 (s)	12.49	CH_3
19.	1.25 (s)	23.92	CH_3	1.04 (s)	23.24	CH_3
20.	1.42 (m)	36.85	CH	1.41 (m)	36.87	CH
21.	0.96 (d, J = 6.6 Hz)	19.28	CH_3	0.96 (d, J = 6.6 Hz)	19.35	CH_3
22.	1.09 (m), 1.48 (m)	33.21	CH_2	1.07 (m), 1.48 (m)	33.24	CH_2
23.	1.41 (m), 1.63 (m)	30.26	CH_2	1.39 (m), 1.59 (m)	30.28	CH_2
24.	3.51 (m)	63.54	CH_2	3.51 (m)	63.55	CH_2
25.	1.11 (m), 1.70 (m)	20.07	CH_2	6.08 (q, J = 7.2 Hz)	130.34	CH
26.	0.81 (t, J = 7.2 Hz)	12.29	CH_3	1.71 (d, J = 7.2 Hz)	12.75	CH_3

NMR: nuclear magnetic resonance; DEPT: distortionless enhancement by polarization transfer.

reaction because this impurity could not be directly obtained from **13** via this reaction. Intermediates **12**, **13**, and **14** and impurity **2** can be purified by recrystallization. Finally, the synthetic routes to impurities **1** and **2** were successfully designed without the use of column chromatography purification and can be employed in the development of an HPLC method for OCA and as external standards with high purity to analyze the content of these impurities in OCA.

Structural elucidation of impurities **1** *and* **2**

The HRMS spectrum of impurity **1** provides a mass ion signal at *m*/*z* 405.3368 ($[M + H]^+$, Calcd. for $C_{26}H_{45}O_3$: 405.3369) in positive ion mode. This result indicates that the molecular formula of impurity **1** is $C_{26}H_{44}O_3$ and that the molecular weight is 404, which is 16 Da (O) less than that of OCA ($C_{26}H_{44}O_4$). The HRMS spectrum of impurity **2** shows a mass ion at *m*/*z* 403.3208 ($[M + H]^+$, Calcd. for $C_{26}H_{43}O_3$: 403.3212) in positive ion mode. This result indicates that the molecular formula of impurity **1** is $C_{26}H_{42}O_3$, and the molecular weight is 402, which is 18 Da (H_2O) less than that of OCA. The infrared (IR) absorptions of impurity **1** were observed for hydroxy (3269 cm^{-1}) and carbonyl (1705 cm^{-1}) functional groups. The IR spectrum of impurity **2** indicates the presence of hydroxy (3337 cm^{-1}), carbonyl (1690 cm^{-1}), and alkenyl (1636 cm^{-1}) functional groups, and the characteristic absorption of the carbonyl shifts to a lower number (<1700 cm^{-1}) due to the conjugated structure of the α,β-unsaturated ketone moiety of impurity **2**. Therefore, impurity **1** contains hydroxy and carbonyl functional groups, and impurity **2** contains an alkenyl group as well as hydroxy and carbonyl groups.

A series of one-dimensional (1D) and two-dimensional (2D) NMR experiments were performed on impurities **1** and **2**, including distortionless enhancement by polarization transfer (DEPT), H-H Correlated Spectroscopy (COSY), heteronuclear single quantum coherence (HSQC), and heteronuclear multiple bond correlation (HMBC) to assign all the ^{1}H and ^{13}C signals (Table 1). The H-H COSY and key HMBC data of impurities **1** and **2** are described in Figure 2.

The ^{13}C NMR and DEPT spectra of impurity **1** indicate the presence of 26 carbon atoms including 4 methyl carbon atoms, 11 methylene carbon atoms, 8 methine carbon atoms, and 3 quaternary carbon atoms. The quaternary carbon signal at 215.48 ppm indicates that C7 is a carbonyl carbon atom. The chemical shifts of H3 and H24 should be located in a lower field region (>3.0 ppm) in the ^{1}H NMR spectrum because the positions of 3 and 24 are adjacent to oxygen atoms. The chemical shifts of C3 and C24 were determined by the HSQC spectrum. The H18 and H19 methyl signals are single peaks, and the methyl signals of H21 and H26 appear as a double and triple, respectively. In addition, the chemical shift of H18 is lower than that of H19 in various bile acid compounds,[10] and the chemical shifts of C18, C19, C21, and C26 were determined from the HSQC spectrum. The chemical shift of the quaternary carbon C10 was determined from the HMBC spectrum via the correlations of C10 and H19. Similarly, the chemical shift

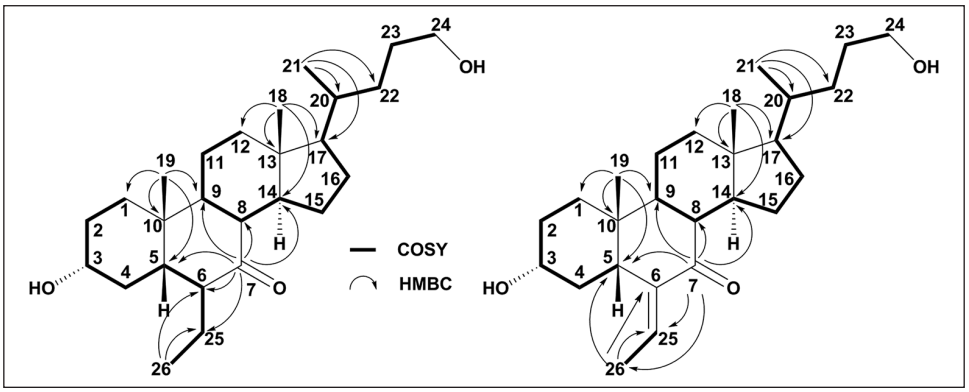


Figure 2. H-H COSY and key HMBCs of impurity **1** (left) and impurity **2** (right) (CD_3OD, 600 MHz).

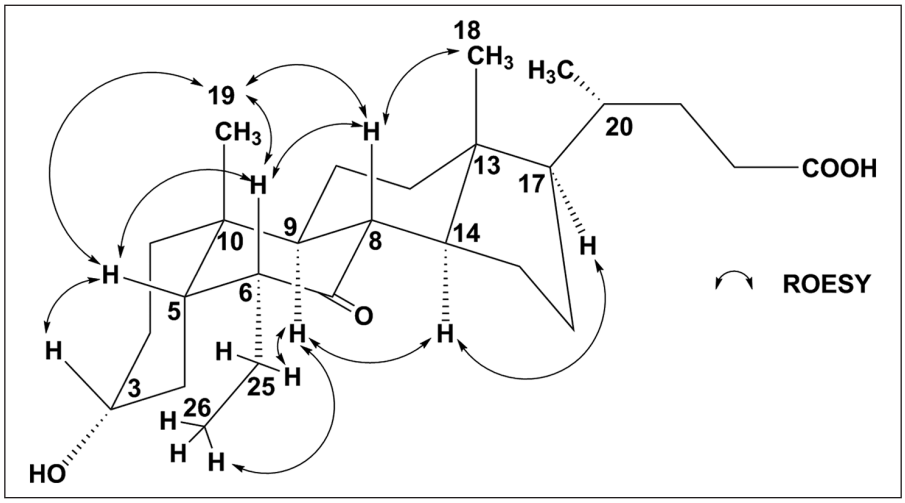


Figure 3. Key ROESY correlations for impurity **1** (CD_3OD, 600 MHz).

of the quaternary carbon C13 was also determined from the HMBC spectrum due to the correlations of C13 and H18. The chemical shift of H23 was determined from the H-H COSY spectrum due to the correlations of H23 and H24, and the corresponding chemical shift of C23 was determined by analysis of the HSQC spectrum. In addition, the chemical shifts of H25 and C25 were determined by the correlations between positions 25 and 26. H6 correlates with H5 and H25 in the H-H COSY spectrum and with positions 7 and 26 in the HMBC spectrum. Moreover, the chemical shifts of H6, H5, and C5 were deduced based on the H-H COSY and HSQC spectra. The methylene at C-1 correlates with positions 3, 5, 10, and 19 in the HMBC spectrum, and thus, the chemical shifts of H1 and C1 were determined. The chemical shifts of H4, C4, H2, and C2 were easily deduced based on the 2D NMR spectrum. The assignments of the ^{1}H and ^{13}C resonance signals of the B-, C-, and D-rings were determined based on an analogous analysis.

In the ^{1}H NMR spectrum of impurity **2**, the resonance signal of the alkenyl group was observed at 6.08 ppm, and the characteristic signals of the carbon atoms of the alkenyl group were observed at 145.38 and 130.34 ppm. These results indicate that impurity **2** contains an ethylene group. Interestingly, C5 and C7 were correlated to H26 in the HMBC spectrum, which correlates with long-range coupling of a propylene. The chemical shifts of impurity **2** were determined using NMR analysis analogous to that employed in the analysis of impurity **1**.

As shown in Scheme 1, intermediate **7** was synthesized from **6** in the presence of $Pd/C/H_2$ and sodium hydroxide in water and was obtained via tautomerism of the α-ketone of the 6β-ethyl intermediate under strong alkaline conditions. A small amount of the 6β-ethyl intermediate most likely remained in intermediate **7**. Therefore, the relative configuration of the ethyl moiety of impurity **1** was confirmed by the rotating-frame Overhauser effect spectroscopy (ROESY) experiment. The key ROESY correlation of impurity **1** is shown in Figure 3. The 6α-orientation of the ethyl moiety of impurity **1** was determined based on H6 being correlated with H5, H8, and H19, but not correlated with H9, as shown in Figure 4. In addition, the ROESY correlations of H9 and H25 or H26 also provided support for the relative configuration. In the ROESY spectrum, H5 is correlated with H3, H6, and H19, and H8 is correlated with H6, H18, and H19. Furthermore, H14 is correlated with H9 and H17, which indicates that the relative configurations of H3, H5, H6, H8, H18, and H19 have a β orientation. However, the relative configurations of H9, H14, and H17 have an α orientation.

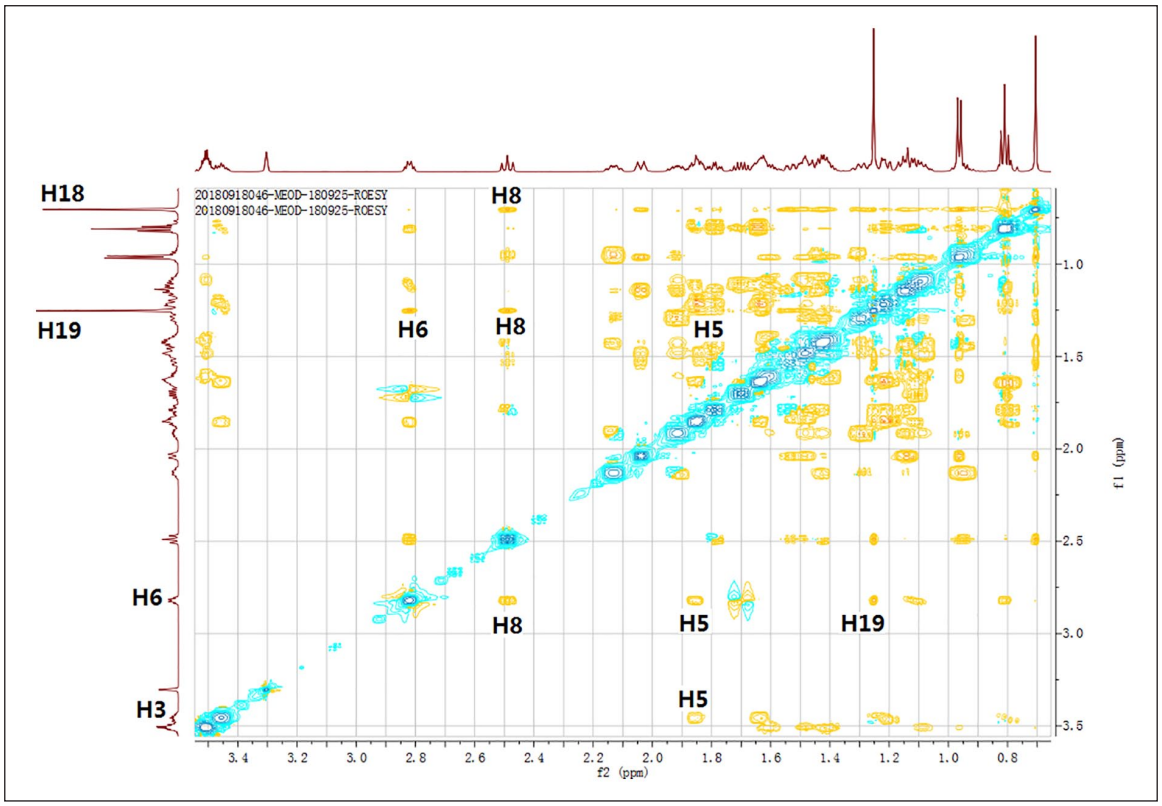


Figure 4. ROESY spectrum for impurity **I** (CD_3OD, 600 MHz).

Conclusion

As per the stringent regulatory requirements that have been recommended by the ICH, we have developed efficient and practical approaches to deliver new impurities **1** and **2** that are found in OCA. In general, many bile acids and their derivatives are oils or foamy solids, which are difficult to purify via recrystallization. However, using the synthetic strategy outlined in this study, the crystallizability can be improved significantly by the formation of a nitrobenzoic ester at the 3-position hydroxy groups in these compounds. In addition, these compounds can be detected by HPLC with a typical UV–Vis detector rather than with ELSD or CAD (charged aerosol detector). The new impurities were characterized using HRMS, FTIR, 1D NMR and 2D NMR techniques. The relative configuration of impurity **1** was determined from the ROESY spectrum. This study provides useful reference information for organic, process, and analytical chemists in pharmaceutical companies and drug regulatory authorities.

Experimental

Materials and instruments

All solvents and reagents were purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China). OCA and its intermediates were prepared in-house. HPLC analyses were performed on a Dionex UltiMate 3000 HPLC (Thermo Fisher, Massachusetts, USA) instrument using a reverse-phase Agilent ZOBAX XDB-C18, 250 mm × 4.6 mm, and 5-μm particle size column (Agilent, California, USA). The melting points were measured on a WRS-1B apparatus (YiCe, Shanghai, China). HRMS measurements were performed on a Bruker Daltonics Solarix 7.0T (Bruker, Karlsruhe, DE, Germany). The IR spectra were recorded in the solid state on a Tracer-100 FTIR spectrometer (Shimadzu, Kyoto, Japan). The NMR spectra were recorded on a Bruker Avance III instrument (Bruker) operating at 400 or 600 MHz for ^{1}H NMR and 100 or 150 MHz for ^{13}C NMR with Me_4Si (tetramethylsilane (TMS)) as the internal standard. DEPT and 2D NMR (H-H COSY, HSQC, HMBC, and ROESY) experiments were also performed on the same instrument to assign the correlations.

Synthesis of impurity **I** *(Scheme 2)*

(3R,5S,6R,7R,8S,9S,10S,13R,14S,17R)-6-ethyl-17-((R)-5-hydroxypentan-2-yl)-10,13-dimethylhexadecahydro-1H-cyclopenta[a]phenanthrene-3,7-diol (**8**). To a mixture of **1** (10.0 g, 23.8 mmol) in dry tetrahydrofuran (100 mL), lithium aluminum hydride (1.8 g, 47.6 mmol) was added in batches at 25 °C under nitrogen followed by stirring at 65 °C for 4 h. The reaction mixture was cooled to 0–5 °C, and then, water (1.8 mL) followed by aqueous 10% sodium hydroxide solution (1.8 mL) were slowly added to the reaction mixture. After stirring for 30 min, the mixture was filtered and the filter cake was washed with tetrahydrofuran (50 mL), and the filtrate was dried using anhydrous sodium sulfate. The solvent was concentrated under reduced pressure to afford compound **8** as a white foamy solid (8.9 g, yield: 92%, purity: 98.3%). ^{1}H NMR (400 MHz, CD_3OD): δ 3.65 (m, 1H), 3.48–3.52 (m, 2H), 3.28–3.33 (m, 1H), 1.98–2.02 (m, 1H),

1.79–1.94 (m, 4H), 1.70–1.77 (m, 2H), 1.57–1.64 (m, 2H), 1.27–1.56 (m, 13H), 1.16–1.21 (m, 2H), 1.05–1.14 (m, 2H), 0.98–1.03 (m, 1H), 0.96 (d, J = 4.4 Hz, 3H), 0.91 (s, 3H), 0.90 (t, J = 4.8 Hz, 3H), 0.69 (s, 3H) (Supplemental material, Figure S2);^{13}C NMR (100 MHz, DMSO-d_6): δ 73.18, 71.19, 63.58, 57.56, 51.67, 46.95, 43.69, 43.16, 41.55, 41.08, 37.06, 36.79, 36.65, 34.52, 34.42, 33.23, 31.26, 30.28, 29.39, 24.59, 23.79, 23.50, 21.99, 19.25, 12.27, 12.06 (Supplemental material, Figure S3). HRMS: m/z [M + Na]$^+$ Calcd for $C_{26}H_{46}NaO_3$: 429.3345; found: 429.3336 (Supplemental material, Figure S1).

*(3R,5S,6R,7R,8S,9S,10S,13R,14S,17R)-6-ethyl-7-hydroxy-10,13-dimethyl-17-((R)-5-((4-nitrobenzoyl)oxy)pentan-2-yl) hexadecahydro-1H-cyclopenta[a]phenanthren-3-yl 4-nitrobenzoate (**9**).* To a mixture of **8** (8.0 g, 19.7 mmol), *N,N*-diisopropylethylamine (7.6 g, 59.0 mmol) and 4-dimethylaminopyridine (0.24 g, 2.0 mmol) in dry tetrahydrofuran (40 mL) was added to 4-nitrobenzoyl chloride (7.7 g, 41.4 mmol) in batches at 0–5 °C followed by stirring at 0 °C for 2 h. The reaction mixture was quenched with an aqueous solution of 20% citric acid (100 mL) and extracted with ethyl acetate (50 mL × 2). The combined organic layer was washed with 10% sodium bicarbonate solution (50 mL) and then dried over anhydrous sodium sulfate and concentrated under reduced pressure. The residue was recrystallized from ethyl acetate (40 mL) and heptane (80 mL) to yield compound **9** as a white solid (11.4 g, yield: 82%, purity: 98.7%). IR (KBr) ν_{max} (cm^{-1}): 3559 (OH), 3431, 2941, 2903, 2870, 1722 (C=O), 1607, 1528 (N=O), 1350 (N=O), 1277, 1119, 1103, 1015, 874, 719 (Supplemental material, Figure S4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.30 (dt, J = 6 Hz, 1.6 Hz, 2H), 8.27 (dt, J = 6 Hz, 1.2 Hz, 2H), 8.20–8.22 (m, 4H), 4.83–4.87 (m, 1H), 4.31–4.39 (m, 2H), 3.75 (s, 1H), 2.13 (q, J = 8.8 Hz, 1H), 2.01 (dt, J = 8 Hz, 2 Hz, 1H), 1.82–1.95 (m, 6H), 1.13–1.73 (m, 20H), 0.99 (d, J = 4.4 Hz, 3H), 0.96 (s, 3H), 0.92 (t, J = 4.8 Hz, 3H), 0.69 (s, 3H) (Supplemental material, Figure S5); ^{13}C NMR (100 MHz, $CDCl_3$): δ 164.91, 164.38, 150.64, 150.52, 136.52, 135.99, 130.83, 130.79, 123.71, 123.55, 76.75, 71.00, 66.69, 56.04, 50.71, 45.25, 42.92, 41.29, 40.12, 39.73, 35.76, 35.58, 35.28, 33.46, 32.15, 29.78, 28.44, 26.85, 25.34, 23.86, 23.27, 22.36, 20.95, 18.76, 11.97, 11.80 (Supplemental material, Figure S6). Anal. cacld for $C_{40}H_{52}N_2O_9$ (704.37): C, 68.16; H, 7.44; N, 3.97; found: C, 68.01; H, 7.28; N, 4.14.

*(3R,5S,6R,8S,9S,10S,13R,14S,17R)-6-ethyl-10,13-dimethyl-17-((R)-5-((4-nitrobenzoyl)oxy)pentan-2-yl)-7-oxohexadecahydro-1H-cyclopenta[a]phenanthren-3-yl 4-nitrobenzoate (**10**).* To a mixture of **9** (10.0 g, 14.2 mmol), sodium bromide (73 mg, 0.71 mmol), and tetrabutylammonium bromide (229 mg, 0.71 mmol) in a mixture of water (5 mL), tetrahydrofuran (20 mL), isopropanol (30 mL), and acetic acid (10 mL), was added slowly to an aqueous solution consisting of 5% sodium hypochlorite (42.3 g, 28.4 mmol) at −15 °C followed by stirring at −15 °C for 2 h. The reaction mixture was quenched with 10% sodium bisulfite solution (50 mL) and extracted with ethyl acetate (50 mL × 2). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to afford a residue, which was recrystallized from methyl tert-butyl ether (50 mL) and heptane (50 mL) to yield compound **10** as a white solid (9.1 g, yield: 91%, purity: 97.8%). IR (KBr) ν_{max}/cm^{-1}: 3426 (OH), 2957, 2874, 1722 (C=O), 1607, 1530 (N=O), 1350 (N=O), 1277, 1119, 1103, 1015, 874, 719 (Supplemental material, Figure S7). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.29 (dd, J = 10.8 Hz, 6 Hz, 4H), 8.19 (dd, J = 16 Hz, 6 Hz, 4H), 4.94 (hept, J = 3.2 Hz, 1H), 4.31–4.39 (m, 2H), 2.77 (q, J = 8.8 Hz, 4 Hz, 1H), 2.42 (t, J = 7.6 Hz, 1H), 2.17–2.22 (m, 1H), 2.03 (dt, J = 8.8 Hz, 2.4 Hz, 1H), 1.91–1.98 (m, 3H), 1.83–1.88 (m, 4H), 1.66–1.77 (m, 2H), 1.46–1.55 (m, 6H), 1.30–1.34 (m, 1H), 1.28 (s, 3H), 1.09–1.27 (m, 6H), 0.98 (d, J = 4.4 Hz, 3H), 0.91–0.96 (m, 1H), 0.82 (t, J = 4.8 Hz, 3H), 0.69 (s, 3H) (Supplemental material, Figure S8). ^{13}C NMR (100 MHz, $CDCl_3$): δ 212.83, 164.76, 164.07, 150.50, 135.92, 135.85, 130.69, 130.65, 123.56, 123.48, 75.05, 66.51, 54.94, 52.05, 50.63, 50.01, 48.99, 43.92, 42.66, 38.98, 35.78, 35.29, 33.84, 32.05, 28.44, 27.81, 26.08, 25.21, 24.57, 23.53, 21.98, 18.88, 18.71, 12.11, 12.06 (Supplemental material, Figure S9). Anal. cacld for $C_{40}H_{50}N_2O_9$ (702.35): C, 68.36; H, 7.17; N, 3.99; found: C, 68.11; H, 6.97; N, 4.24.

*(3R,5S,6R,8S,9S,10S,13R,14S,17R)-6-ethyl-3-hydroxy-17-((R)-5-hydroxypentan-2-yl)-10,13-dimethylhexadecahydro-7H-cyclopenta[a]phenanthren-7-one (impurity **1**).* To a mixture of **10** (9.0 g, 12.8 mmol) in a solution consisting of water (6 mL), tetrahydrofuran (30 mL), and methanol (15 mL), sodium hydroxide (1.3 g, 32 mmol) was added at 25 °C followed by stirring at 25 °C for 1 h. The reaction mixture was diluted with water (50 mL) and extracted with ethyl acetate (50 mL × 2). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to afford a residue, which was recrystallized from dichloromethane (90 mL) and heptane (45 mL) to yield impurity **1** as a white solid (4.5 g, yield: 87%, purity: 99.3%). m.p.: 108–110 °C. IR (KBr) ν_{max}/cm^{-1}: 3269 (OH), 2945, 2874, 1705 (C=O), 1464, 1452, 1377, 1061, 1013, 735 (Supplemental material, Figure S10). ^{1}H NMR (600 MHz, CD_3OD): δ 3.49–3.52 (m, 2H), 3.46 (sept, J = 4.8 Hz, 1H), 2.82 (q, J = 12.6 Hz, 5.4 Hz, 1H), 2.49 (t, J = 11.4 Hz, 1H), 2.10–2.16 (m, 1H), 2.04 (dt, J = 13.2 Hz, 3 Hz, 1H), 1.89–1.95 (m, 1H), 1.83–1.88 (m, 2H), 1.79 (td, J = 12.0 Hz, 4.8 Hz, 1H), 1.70 (quin, J = 7.2 Hz, 1H), 1.56–1.66 (m, 3H), 1.53 (dd, J = 12.6 Hz, 3.6 Hz, 1H), 1.46–1.51 (m, 2H), 1.38–1.45 (m, 3H), 1.26–1.33 (m, 1H), 1.25 (s, 3H), 1.06–1.23 (m, 6H), 0.96 (d, J = 6.6 Hz, 3H), 0.92–0.95 (m, 1H), 0.81 (t, J = 7.2 Hz, 3H), 0.77–0.80 (m, 1H), 0.71 (s, 3H) (Supplemental material, Figure S12); ^{13}C NMR (150 MHz, CD_3OD) δ 215.48, 71.66, 63.54, 56.51, 53.29, 52.16, 51.19, 50.49, 45.36, 43.76, 40.43, 36.85, 36.82, 35.25, 33.21, 32.57, 30.53, 30.26, 29.42, 25.62, 23.92, 22.97, 20.07, 19.28, 12.52, 12.29 (Supplemental material, Figure S13). HRMS: m/z [M + H]$^+$ Calcd for $C_{26}H_{45}O_3$: 405.3369; found: 405.3368 (Supplemental material, Figure S11).

Synthesis of impurity **2** *(Scheme 3)*

*(3R,5S,7R,8R,9S,10S,13R,14S,17R)-17-((R)-5-hydroxypentan-2-yl)-10,13-dimethylhexadecahydro-1H-cyclopenta[a]phenanthrene-3,7-diol (***11***).* To a mixture of **2** (20.0 g, 50.9 mmol) in dry tetrahydrofuran (200 mL), lithium aluminum hydride (3.9 g, 101.8 mmol) was added in batches at 25 °C under nitrogen followed by stirring at 65 °C for 3 h. The reaction mixture was cooled to 0–5 °C, and then, water (3.9 mL) followed by 10% sodium hydroxide solution (3.9 mL) were slowly added to the reaction mixture. After 30 min, the mixture was filtered, and the filter cake was washed with tetrahydrofuran (100 mL), and the filtrate was dried over anhydrous sodium sulfate. The solvent was concentrated under reduced pressure to afford compound **11** as a white foamy solid (17.3 g, yield: 90%, purity: 98.6%). ^{1}H NMR (400 MHz, CD_3OD): δ 3.82 (q, J = 1.6 Hz, 1H), 3.49–3.56 (m, 2H), 3.37–3.42 (m, 1H), 2.28 (q, J = 8.4 Hz, 1H), 2.03 (dt, J = 6.4 Hz, 2.4 Hz, 1H), 1.97–2.01 (m, 1H), 1.86–1.95 (m, 3H), 1.73–1.78 (m, 1H), 1.60–1.69 (m, 3H), 1.29–1.56 (m, 11H), 1.18–1.24 (m, 2H), 1.07–1.16 (m, 2H), 1.00–1.03 (m, 1H), 0.99 (d, J = 4.4 Hz, 3H), 0.95 (s, 3H), 0.72 (s, 3H) (Supplemental material, Figure S21); ^{13}C NMR (100 MHz, CD_3OD): δ 72.86, 69.08, 63.59, 57.54, 51.54, 43.64, 43.18, 41.08, 40.77, 40.46, 37.06, 36.57, 36.23, 35.90, 34.05, 33.23, 31.36, 30.28, 29.36, 24.65, 23.44, 21.81, 19.26, 12.21 (Supplemental material, Figure S22). HRMS: *m/z* [M + H]$^+$ Calcd for $C_{24}H_{43}O_3$: 379.3212; found: 379.3207 (Supplemental material, Figure S20).

*(3R,5R,7R,8R,9S,10S,13R,14S,17R)-7-hydroxy-10,13-dimethyl-17-((R)-5-((4-nitrobenzoyl)oxy)pentan-2-yl)hexadecahydro-1H-cyclopenta[a]phenanthren-3-yl 4-nitrobenzoate (***12***).* To a mixture of **11** (16 g, 42.3 mmol), *N,N*-diisopropylethylamine (16.4 g, 126.8 mmol), and 4-dimethylaminopyridine (0.51 g, 4.2 mmol) in dry tetrahydrofuran (100 mL), 4-nitrobenzoyl chloride (16.5 g, 88.8 mmol) was added in batches at 0–5 °C followed by stirring at 0 °C for 2 h. The reaction mixture was quenched with aqueous 20% citric acid (200 mL) and extracted with ethyl acetate (100 mL × 2). The combined organic layer was washed with 10% sodium bicarbonate solution (100 mL). The organic layer was dried using anhydrous sodium sulfate and concentrated under reduced pressure to afford a residue, which was recrystallized from ethyl acetate (80 mL) and heptane (160 mL) to yield compound **12** as a white solid (22.3 g, yield: 78%, purity: 98.0%). IR (KBr) ν_{max}/cm^{-1}: 3564 (OH), 2940, 2868, 1722 (C=O), 1607, 1528 (N=O), 1350 (N=O), 1277, 1119, 1103, 1015, 874, 719 (Supplemental material, Figure S23). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.28 (dd, J = 12.4 Hz, 6 Hz, 4H), 8.20–8.22 (m, 4H), 4.88 (sept, J = 4.8 Hz, 1H), 4.33–4.39 (m, 2H), 3.89 (s, 1H), 2.49 (t, J = 8.8 Hz, 1H), 2.00–2.05 (m, 2H), 1.83–1.95 (m, 6H), 1.60–1.72 (m, 3H), 1.11–1.55 (m, 15H), 0.99 (d, J = 4.4 Hz, 3H), 0.97 (s, 3H), 0.69 (s, 3H) (Supplemental material, Figure S24); ^{13}C NMR (100 MHz, $CDCl_3$): δ 164.75, 164.24, 150.49, 150.37, 136.35, 135.85, 130.67, 130.63, 123.56, 123.41, 76.19, 68.52, 66.52, 55.90, 50.50, 42.72, 41.24, 39.59, 39.36, 35.43, 35.25, 35.13, 34.94, 34.46, 32.91, 32.00, 28.26, 26.75, 25.19, 23.72, 22.74, 20.64, 18.59, 11.81 (Supplemental material, Figure S25). Anal. cacld for $C_{38}H_{48}N_2O_9$ (676.34): C, 67.44; H, 7.15; N, 4.14; found: C, 67.21; H, 7.07; N, 4.31.

*(3R,5S,8R,9S,10S,13R,14S,17R)-10,13-dimethyl-17-((R)-5-((4-nitrobenzoyl)oxy)pentan-2-yl)-7-oxohexadecahydro-1H-cyclopenta[a]phenanthren-3-yl 4-nitrobenzoate (***13***).* To a mixture of **12** (20.0 g, 29.6 mmol), sodium bromide (152 mg, 1.48 mmol), and tetrabutylammonium bromide (477 mg, 1.48 mmol) in a solution consisting of water (10 mL), tetrahydrofuran (40 mL), isopropanol (60 mL), and acetic acid (20 mL), a 5% sodium hypochlorite solution (88.1 g, 59.2 mmol) was slowly added at −15 °C followed by stirring at −15 °C for 2 h. The reaction mixture was quenched with aqueous 10% sodium bisulfite (100 mL) and extracted with ethyl acetate (100 mL × 2). The organic layer was dried using anhydrous sodium sulfate and concentrated under reduced pressure to afford a residue, which was recrystallized from methyl tert-butyl ether (100 mL) and heptane (100 mL) to yield compound **13** as a white solid (18.8 g, yield: 94%, purity: 98.2%). IR (KBr) ν_{max}/cm^{-1}: 2949, 2874, 1717 (C=O), 1607, 1526 (N=O), 1348(N=O), 1287, 1117, 1105, 1015, 872, 716 (Supplemental material, Figure S26). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.25–8.30 (m, 4H), 8.15–8.21 (m, 4H), 4.98 (sept, J = 4.8 Hz, 1H), 4.29–4.39 (m, 2H), 2.91 (dd, J = 12.4 Hz, 5.6 Hz, 1H), 2.44 (t, J = 11.2 Hz, 1H), 2.15–2.23 (m, 1H), 1.46–2.05 (m, 16H), 1.14–1.34 (m, 6H), 1.25 (s, 3H), 0.91–1.04 (m, 1H), 0.97 (d, J = 6.8 Hz, 3H), 0.68 (s, 3H) (Supplemental material, Figure S27); ^{13}C NMR (100 MHz, $CDCl_3$): δ 212.04, 164.88, 164.19, 150.62, 136.02, 135.98, 130.78, 123.68, 123.61, 74.85, 66.62, 55.03, 49.68, 49.04, 46.09, 45.39, 43.12, 42.75, 39.07, 35.40, 35.37, 33.88, 33.22, 32.16, 28.53, 26.25, 25.36, 24.88, 23.18, 21.93, 18.84, 12.21 (Supplemental material, Figure S28). Anal. cacld for $C_{38}H_{46}N_2O_9$ (674.32): C, 67.64; H, 6.87; N, 4.15; found: C, 67.46; H, 6.58; N, 4.38.

*(3R,5S,8R,9S,10S,13R,14S,17R)-3-hydroxy-17-((R)-5-hydroxypentan-2-yl)-10,13-dimethylhexadecahydro-7H-cyclopenta[a]phenanthren-7-one (***14***).* To a mixture of **13** (18.0 g, 26.7 mmol) in a solution consisting of water (13 mL), tetrahydrofuran (60 mL), and methanol (30 mL), sodium hydroxide (2.7 g, 66.8 mmol) was added at 25 °C followed by stirring at 25 °C for 1 h. The reaction mixture was diluted with water (100 mL) and extracted with ethyl acetate (100 mL × 2). The organic layer was dried over anhydrous sodium sulfate and concentrated under reduced pressure to afford a residue, which was recrystallized from ethyl acetate (180 mL) and heptane (110 mL) to yield compound **14** as a white solid (8.4 g, yield: 84%, purity: 96.2%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.56–3.64 (m, 3H), 2.85 (dd, J = 12.4 Hz, 6 Hz, 1H), 2.38 (t, J = 11.2 Hz, 1H), 1.77–2.02 (m, 6H), 1.60–1.72 (m, 3H), 1.35–1.51 (m, 7H), 1.05–1.33 (m, 9H), 1.19 (s, 3H), 0.88–0.99 (m, 1H), 0.93 (d, J = 6 Hz, 3H), 0.65 (s, 3H) (Supplemental material, Figure S30); ^{13}C NMR (100 MHz, $CDCl_3$): δ 212.16, 71.10, 63.73, 55.11, 49.69, 49.06, 46.24, 45.56, 42.90, 42.77, 39.13, 37.57, 35.65, 35.31, 34.30, 32.00, 30.06, 29.59, 28.57, 25.00, 23.21, 21.86, 18.89, 12.21 (Supplemental material, Figure S31). HRMS:

m/z [M + H]$^+$ Calcd for $C_{24}H_{41}O_3$: 377.3056; found: 377.3050 (Supplemental material, Figure S29).

*(3R,5S,8R,9S,10S,13R,14S,17R)-10,13-dimethyl-3-((trimethylsilyl)oxy)-17-((R)-5-((trimethylsilyl)oxy)pentan-2-yl) hexadecahydro-7H-cyclopenta[a]phenanthren-7-one (***15***).* To a mixture of **14** (8 g, 21.2 mmol) and triethylamine (6.4 g, 63.6 mmol) in dichloromethane (80 mL), trimethylsilyl chloride (6.9 g, 63.6 mmol) was added slowly at 0 °C followed by stirring at 25 °C for 2 h. The reaction mixture was diluted with water (50 mL). The organic layer was separated, dried over anhydrous sodium sulfate, and concentrated under reduced pressure to afford **15** as yellow oil (14.2 g), which was not purified prior to use in the next step. A small analytical sample of **15** was purified by column chromatography. IR (KBr) ν_{max}/cm^{-1}: 2955, 2866, 1707 (C=O), 1466, 1381, 1252, 1098, 1078(Si–O), 1015, 885, 841(Si–C), 756, 746 (Supplemental material, Figure S32). ^{1}H NMR (600 MHz, $CDCl_3$): δ 3.51–3.57 (m, 3H), 2.81 (dd, *J* = 12.6 Hz, 6 Hz, 1H), 2.34 (t, *J* = 11.4 Hz, 1H), 2.15–2.22 (m, 1H), 1.98 (dt, *J* = 12.6 Hz, 3 Hz, 1H), 1.78–1.92 (m, 5H), 1.55–1.61 (m, 3H), 1.35–1.46 (m, 7H), 1.19–1.26 (m, 2H), 1.17 (s, 3H), 1.03–1.16 (m, 4H), 0.87–0.95 (m, 4H), 0.63 (s, 3H), 0.11 (s, 9H), 0.09 (s, 9H) (Supplemental material, Figure S33); ^{13}C NMR (100 MHz, $CDCl_3$): δ 211.92, 71.45, 63.41, 55.12, 49.65, 49.04, 46.25, 45.60, 42.73, 42.47, 39.10, 37.76, 35.66, 35.26, 34.49, 32.04, 30.46, 29.54, 28.54, 25.05, 23.19, 21.77, 18.84, 12.19, 0.30, -0.27 (Supplemental material, Figure S34). Anal. cacld for $C_{30}H_{56}O_3Si_2$ (520.38): C, 69.17; H, 10.84; found: C, 68.89; H, 10.61.

*(((3R,5S,8S,9S,10S,13R,14S,17R)-10,13-dimethyl-17-((R)-5-((trimethylsilyl)oxy)pentan-2-yl)-2,3,4,5,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthrene-3,7-diyl)bis(oxy)) bis(trimethylsilane) (***16***).* To a mixture of lithium diisopropylamide (28.8 mL, 57.6 mmol, 2 mol L^{-1}) and trimethylsilyl chloride (5.2 g, 48.0 mmol) in tetrahydrofuran (50 mL), a tetrahydrofuran (30 mL) solution containing **15** (10 g, 19.2 mmol) was added slowly at −70 °C followed by stirring for 2 h. The reaction mixture was diluted with ethyl acetate (100 mL) and quenched with an aqueous solution (50 mL) of citric acid monohydrate (4 g, 19.2 mmol). The organic layer was separated, dried over anhydrous sodium sulfate, and concentrated under reduced pressure to afford **16** as yellow oil (13.5 g), which was not purified prior to use in the next step. Intermediate **16** was not analyzed due to its inherent instability.

(3R,5R,8S,9S,10R,13R,14S,17R,E)-6-ethylidene-3-hydroxy-17-((R)-5-hydroxypentan-2-yl)-10,13-dimethylhexadecahydro-7H-cyclopenta[a]phenanthren-7-one (impurity **2***).* To a mixture of **16** (13 g, 21.9 mmol) and acetaldehyde (1.9 g, 43.8 mmol) in dichloromethane (60 mL), an acetonitrile solution containing 19% boron trifluoride (19.5 g, 54.8 mmol) was added slowly at −70 °C followed by stirring for 1 h and then warming to −10 °C for 1 h. The reaction mixture was diluted with dichloromethane (50 mL) and quenched with 50 mL of aqueous sodium hydroxide solution (2.6 g, 65.8 mmol). The organic layer was separated, dried over anhydrous sodium sulfate, and concentrated under reduced pressure to afford a residue, which was recrystallized from dichloromethane (80 mL) to yield impurity **2** (4.5 g, 78% yield from **14**, purity: 98.8%). m.p.: 106–107 °C. IR (KBr) ν_{max}/cm^{-1}: 3337 (OH), 2940, 2872, 1690 (C=O), 1636 (C=C), 1458, 1375, 1065, 1011, 735, 669 (Supplemental material, Figure S35). ^{1}H NMR (600 MHz, CD_3OD): δ 6.08 (q, *J* = 7.2 Hz, 1H), 3.59 (sept, *J* = 4.8 Hz, 1H), 3.48–3.54 (m, 2H), 2.68 (dd, *J* = 13.2 Hz, 4.2 Hz, 1H), 2.35–2.38 (m, 1H), 2.32 (t, *J* = 11.4 Hz, 1H), 2.08 (dt, *J* = 13.2 Hz, 3 Hz, 1H), 1.89–2.00 (m, 3H), 1.71 (d, *J* = 7.2 Hz, 3H), 1.69 (m, 1H), 1.46–1.65 (m, 5H), 1.39–1.45 (m, 3H), 1.27–1.35 (m, 3H), 1.06–1.24 (m, 5H), 1.04 (s, 3H), 0.96 (d, *J* = 6.6 Hz, 3H), 0.69 (s, 3H) (Supplemental material, Figure S37); ^{13}C NMR (150 MHz, CD_3OD): δ 207.64, 145.38, 130.34, 70.96, 63.55, 56.18, 52.03, 50.18, 46.88, 44.66, 40.74, 40.32, 38.37, 36.87, 35.77, 35.41, 33.24, 30.46, 30.28, 29.53, 27.03, 23.24, 22.45, 19.35, 12.75, 12.49 (Supplemental material, Figure S38). HRMS: *m/z* [M + H]$^+$ Calcd for $C_{26}H_{43}O_3$: 403.3212; found: 403.3208 (Supplemental material, Figure S36).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received the following financial support for the research, authorship, and/or publication of this article: Financial support was provided by the Zhejiang University of Technology and the China State Institute of Pharmaceutical Industry.

Supplemental material

Supplemental material for this article is available online.

References

1. Aguilar MT and Carey EJ. *Clin Liver Dis* 2018; 22: 613.
2. Gawrieh S and Chalasani N. *Clin Liver Dis* 2018; 22: 189.
3. Floreani A and Mangini C. *Eur J Intern Med* 2018; 47: 1.
4. Feng WD, Zhuo SM and Zhang FL. *J Pharm Biomed Anal* 2019; 165: 325.
5. Zhao Y, Li XL, Liu H, et al. *Steriods* 2015; 95: 7.
6. ICH harmonised tripartite guideline: impurities in new drug substances Q3A (R2) (Current Step 4 version), 25 October 2006.
7. Douša M, Slavíková M, Kubelka T, et al. *J Pharm Biomed Anal* 2018; 149: 214.
8. Ferrari M and Pellicciari R. PCT Patent 2006/122977A3, 2006.
9. Steiner A, Waenerlund Poulsen H, Jolibois E, et al. PCT Patent application 20130345188A1, 2013.
10. Yu d, Mattern DL and Forman BM. *Steroids* 2012; 77: 1335.