SUPPLEMENTARY MATERIAL

Title: A new anthraquinone and eight constituents from *Hedyotis caudatifolia* Merr. et Metcalf: isolation, purification and structural identification

Authors: Peng Luo^{a, b}, Jiale Su^a, Yilin Zhu^a, Jianhua Wei^a, Wanxing Wei^b and Weigao Pan^{a*}

Authors affiliations:

^aGuangxi University of Chinese Medicine, Nanning 530001, China

^bGuangxi University, College of Chemistry and Chemical Engineering, Nanning 530004, China

Correspondence:

Prof. Weigao Pan, Guangxi University of Chinese Medicine, Nanning 530001, China

E-mail: agao81706672@163.com

Abstract: *Hedyotis caudatifolia* Merr. et Metcalf. (HC), a folk medicine in Yao nationalities areas in China, was used to investigate the chemical constituents. Through silica gel and Sephadex LH-20 column chromatography, nine compounds were isolated and purified. By physical and chemical properties, IR, MS (EI-MS, high resolution EI-MS), 1D-NMR (¹H-NMR, ¹³C-NMR) and 2D-NMR (HSQC, ¹H-¹H COSY, HMBC), their structures were identified as β-sitosterol (**1**), stigmasterol (**2**), scopolin (**3**), 2-hydroxy-1,7,8-trimethoxyanthracene-9,10-dione (**4**), oleanolic acid (**5**), ursolic acid (**6**), methy barbinervate (**7**), β-daucosterol (**8**), p-Hydroxybenzoic acid (**9**). These compounds were isolated from HC for the first time, and **4** is a new anthraquinone whose biological activities is worth to be investigated in future. These compounds may contribution to the HC's pharmacological effects on treating diseases, and may be used as candidates for control index in establishing the quality control standard of HC.

---

*Corresponding author. Email: agao81706672@163.com

List of NMR data and chemico-physical data of the compounds

β -sitosterol (**1**): White needle crystal, m.p. 137~138°C, soluble in chloroform or benzene and insoluble in water or methanol.

$^1\text{H-NMR}$ (400MHz, CDCl_3) δ : 3.52 (1H, m, H-3), 2.23~2.31 (2H, m, H-4), 5.35 (1H, m, H-6), 0.67~0.69 (3H, m, H-18), 0.78~2.00 (42H, m). $^{13}\text{C-NMR}$ and DEPT (100MHz, CDCl_3) δ : 37.2 (C-1, CH_2), 31.6 (C-2, CH_2), 71.8 (C-3, CH), 42.3 (C-4, CH_2), 140.7 (C-5), 121.7 (C-6, CH), 31.9 (C-7, CH_2), 31.9 (C-8, CH), 50.1 (C-9, CH), 36.5 (C-10, C), 21.0 (C-11, CH_2), 39.7 (C-12, CH_2), 42.2 (C-13, C), 56.8 (C-14, CH), 24.3 (C-15, CH_2), 28.2 (C-16, CH_2), 56.7 (C-17, CH), 11.8 (C-18, CH_3), 19.4 (C-19, CH_3), 36.1 (C-20, CH), 18.9 (C-21, CH_3), 33.9 (C-22, CH_2), 26.0 (C-23, CH_2), 45.8 (C-24, CH), 29.1 (C-25, CH), 19.8 (C-26, CH_3), 19.0 (C-27, CH_3), 23.0 (C-28, CH_2), 12.0 (C-29, CH_3).

Stigmasterol (**2**): White flake crystal, m.p. 169~171°C, soluble in chloroform. IR ν_{max} (cm^{-1} , KBr, w-weak, m-middle, s-strong, br-broad): 3735(m), 3450(s, br), 3425(s, br) [$\nu_{\text{O-H}}$]; 3029(w) [$\nu_{\text{C-H}}$]; 2959(s, br), 2868(s, br), 2852(s, br) [$\nu_{\text{C-H}}$]; 1665(w), 1658(w), 1641(w); 1464(s), 1382(s), 1332(w), 1241(w), 1192(m), 1133(m), 1108(m), 1054(s), 1022(m), 986(w), 970(m), 959(s), 926(w), 838(m), 801(s), 739(w), 627(w), 591(w). EI-MS m/z (%): 413[$\text{M}+1$] $^+$, 415(100), 396, 163, 135, 123, 109, 69. $^1\text{H-NMR}$ (400MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 3.86 (1H, m, H-3), 2.63 (2H, m, H-4), 5.42 (1H, m, H-6), 0.67~0.69 (3H, m, H-18), 5.20 (1H, m, H-22), 5.05 (1H, m, H-23), 6.22 (br, OH), 0.84~2.08 (38H, m). $^{13}\text{C-NMR}$ and DEPT (100MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 37.9 (C-1, CH_2), 32.2 (C-2, CH_2), 71.3 (C-3, CH), 43.6 (C-4, CH_2), 142.0 (C-5, C), 121.3 (C-6, CH), 32.7 (C-7, CH_2), 29.5 (C-8, CH), 50.5 (C-9, CH), 36.9 (C-10, C), 21.4 (C-11, CH_2), 40.0 (C-12, CH_2), 42.5 (C-13, C), 56.9 (C-14, CH), 24.6 (C-15, CH_2), 28.6 (C-16, CH_2), 56.2 (C-17, CH), 12.0 (C-18, CH_3), 19.2 (C-19, CH_3), 40.9 (C-20, CH), 19.7 (C-21, CH_3), 138.9 (C-22, CH), 129.5 (C-23, CH), 51.5 (C-24, CH), 32.3 (C-25, CH), 20.0 (C-26, CH_3), 19.1 (C-27, CH_3), 25.8 (C-28, CH_2), 12.2 (C-29, CH_3).

Scopolin (**3**): Light yellow needle crystal, m.p. 203~204°C. IR ν_{max} (cm^{-1} , KBr): 3830(w), 3650(w), 3339(s, br) [$\nu_{\text{O-H}}$]; 3008(w) [$\nu_{\text{C-H}}$]; 2948(w) [$\nu_{\text{C-H}}$]; 2485(s, br), 2346(w); 1704(s, br) [$\nu_{\text{C=O}}$]; 1627(m), 1607(s), 1565(s), 1510(s), 1454(m) [$\nu_{\text{C=C}}$]; 1435(m), 1375(m), 1290(s), 1263(s), 1219(m), 1190(m), 1163(w), 1141(s), 1099(m), 1019(s), 923(s), 857(s), 821(m), 746(m), 592(s). EI-MS m/z (%): 192(M^+ , 100), 177, 164, 149, 121, 69. $^1\text{H-NMR}$ (400MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 6.29(1H, d, $J=9.4\text{Hz}$, H-3), 7.67(1H, d, $J=9.4\text{Hz}$, H-4), 7.10 (1H, s, H-5), 7.02 (1H, s, H-8), 3.75 (3H, s, OCH_3). $^{13}\text{C-NMR}$ and DEPT (100MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 161.5 (C-2), 112.3 (C-3), 144.1 (C-4), 109.4 (C-5), 151.0 (C-6), 153.0 (C-7), 104.1 (C-8), 146.2 (C-9), 111.1 (C-10), 56.1 (OCH_3).

2-hydroxy-1,7,8-trimethoxyanthracene-9,10-dione (**4**): Light green featherlike crystal, m.p. 237~239°C, soluble in chloroform. IR ν_{max} (cm^{-1} , KBr, Figure S2): 3740(w), 3425(s, br) [$\nu_{\text{O-H}}$]; 2925(s, br), 2852(w) [$\nu_{\text{C-H}}$]; 1844(w); 1772(w), 1734(w) [$\nu_{\text{C=O}}$]; 1666(s), 1573(s), 1521(w), 1507(w), 1480(m), 1456(m) [$\nu_{\text{C=C}}$]; 1412(m), 1331(s), 1274(s), 1075(m), 1036(m), 978(m), 938(m), 840(w), 801(m), 741(m). EI-MS m/z (%) (Figure S1): 314(M^+ , 100), 299, 281, 253, 239, 225.

HR-EI-MS m/z (Figure S1): 314.0791. $^1\text{H-NMR}$ (600MHz, CDCl_3 , Figure S3) δ : 7.25~7.27 (2H, d, $J=12.0$ Hz, H-3), 7.99~8.01 (2H, d, $J=12.0$ Hz, H-4), 8.03~8.05 (2H, d, $J=12.0$ Hz, H-5), 7.18~7.20 (2H, d, $J=12.0$ Hz, H-6), δ_{H} 3.95 (3H, s, H-11), 3.92 (3H, s, H-12), 3.93 (3H, s, H-13), 6.54 (br, OH). $^{13}\text{C-NMR}$ and DEPT (150MHz, CDCl_3 , Figure S4) δ : 144.8 (C-1, C), 153.8 (C-2, C), 119.1 (C-3, CH), 124.5 (C-4, CH), 128.1 (C-4a, C), 124.1 (C-5, CH), 114.8 (C-6, CH), 157.7 (C-7, C), 148.0 (C-8, C), 125.4 (C-8a, C), 180.8 (C-9, C), 124.2 (C-9a, C), 180.9 (C-10, C), 127.2 (C-10a, C), 61.3 (C-11, OCH_3), 55.3 (C-12, OCH_3), 60.3 (C-13, OCH_3).

Oleanolic acid (**5**): White solid, m.p. 308~310°C, soluble in mixture solution of chloroform-methanol. IR ν_{max} (cm^{-1} , KBr): 3855(w), 3822(w), 3676(w), 3651(w), 3428(s, br) [$\nu_{\text{O-H}}$]; 2942(s, br), 2876(s, br) [$\nu_{\text{C-H}}$]; 2653(w), 2346(w), 1697(s, br), 1560(w), 1541(w), 1508(w), 1463(s), 1387(s), 1365(m), 1321(w), 1304(m), 1268(m), 1206(w), 1183(m), 1163(w), 1138(w), 1107(w), 1094(w), 1030(s), 1009(w), 996(m), 974(w), 950(w), 916(w), 884, 826(m), 816(w), 758(m), 680(w), 655, 642(w), 601(w), 565(m), 532(m). EI-MS $m/z(\%)$: 456(M^+), 438, 423, 410, 395, 300, 287, 257, 248(100), 233, 219, 203, 189, 175, 161, 147, 133, 119, 95, 81, 69, 55. $^1\text{H-NMR}$ (400MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 3.43 (1H, m, H-3), 5.49 (1H, brs, H-12), 3.31 (1H, m, H-18), 1.24(3H, s, H-23), 1.02(3H, s, H-24), 0.88(3H, s, H-25), 1.02(3H, s, H-26), 1.28(3H, s, H-27), 0.94(3H, s, H-29), 1.00(3H, s, H-30). $^{13}\text{C-NMR}$ and DEPT (100MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 39.0 (C-1, CH_2), 28.2 (C-2, CH_2), 78.1 (C-3, CH), 39.5 (C-4, C), 55.9 (C-5, CH), 18.9 (C-6, CH_2), 33.3 (C-7, CH_2), 39.8 (C-8, C), 48.2 (C-9, CH), 37.4 (C-10, C), 23.8 (C-11, CH_2), 122.6 (C-12, CH), 144.9 (C-13, C), 42.1 (C-14, C), 28.4 (C-15, CH_2), 23.9 (C-16, CH_2), 46.5 (C-17, C), 42.2 (C-18, CH), 46.7 (C-19, CH_2), 31.0 (C-20, C), 34.3 (C-21, CH_2), 33.3 (C-22, CH_2), 28.9 (C-23, CH_3), 15.6 (C-24, CH_3), 16.6 (C-25, CH_3), 17.5 (C-26, CH_3), 26.3 (C-27, CH_3), 180.3 (C-28, COOH), 33.4 (C-29, CH_3), 23.8 (C-30, CH_3).

Ursolic acid (**6**): White solid, m.p. 274~276°C, soluble in mixture solution of chloroform-methanol. IR ν_{max} (cm^{-1} , KBr): 3739(w), 3425(s, br) [$\nu_{\text{O-H}}$]; 2966(s, br), 2926(s, br), 2871(s, br) [$\nu_{\text{C-H}}$]; 1703(s, br), 1692(s, br), 1635(w), 1554(m), 1454(s), 1388(m), 1378(m), 1311(m), 1274(m), 1237(m), 1187(w), 1140(w), 1092(w), 1029(s), 997(m), 974(w), 828(w), 807(w), 761(m), 662(m), 532(m). EI-MS m/z : 457($\text{M}+1$), 458, 439, 411, 396, 248(100), 249, 207, 203, 133, 119. $^1\text{H-NMR}$ (400MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 3.45 (1H, dd, $J=9.5, 6.3$ Hz, H-3), 5.49 (1H, d, $J=3.2$ Hz, H-12), 2.63 (1H, d, $J=11.2$ Hz, H-18), 2.32 (2H, m, H-15), 2.11 (2H, m, H-16), 1.22 (3H, s, H-23), 1.02 (3H, s, H-24), 1.04 (3H, s, H-25), 0.87 (3H, s, H-26), 1.24 (3H, s, H-27), 0.94 (3H, s, $J=6.1$ Hz, H-29), 0.99 (3H, s, $J=6.4$ Hz, H-30). $^{13}\text{C-NMR}$ and DEPT (100MHz, $\text{C}_5\text{D}_5\text{N}$) δ : 39.1 (C-1, CH_2), 28.1 (C-2, CH_2), 78.1 (C-3, CH), 39.4 (C-4, C), 55.8 (C-5, CH), 18.8 (C-6, CH_2), 33.6 (C-7, CH_2), 40.0 (C-8, C), 48.1 (C-9, CH), 37.5 (C-10, C), 23.6 (C-11, CH_2), 125.7 (C-12, CH), 139.3 (C-13, C), 42.5 (C-14, C), 28.7 (C-15, CH_2), 24.9 (C-16, CH_2), 48.1 (C-17, C), 53.6 (C-18, CH), 39.5 (C-19, CH), 39.4 (C-20, CH), 31.1 (C-21, CH_2), 37.3 (C-22, CH_2), 28.8 (C-23, CH_3), 16.6 (C-24, CH_3), 15.7 (C-25, CH_3), 17.5 (C-26, CH_3), 23.9 (C-27, CH_3), 179.9 (C-28, COOH), 17.6 (C-29, CH_3), 21.4 (C-30, CH_3).

Methy barbinervate (7): Colorless needle crystal, m.p. 321~322°C. IR $\nu_{\max}$ (cm⁻¹, KBr): 3855(w), 3822(w), 3752(w), 3568(s, br), 3448(s, br) [$\nu_{\text{O-H}}$]; 2970(s, br), 2935(s, br), 2878(s, br) [$\nu_{\text{C-H}}$]; 2346(w), 1689(s, br), 1560(w), 1542(w), 1508(w), 1459(s), 1379(s), 1271(m), 1235(m), 1157(m), 1091(w), 1068(m), 1027(m), 1004(m), 957(w), 931(m), 910(w), 865(m), 804(w), 766(m), 690(w), 653(m), 601(m), 534(m). EI-MS $m/z(\%)$: 471, 443, 370, 264, 246, 231, 218, 206(100), 187, 175, 146, 119. ¹H-NMR (400MHz, C₅D₅N) δ : 4.44 (1H, brs, H-3), 5.60 (1H, brs, H-12), 3.03 (1H, s, H-18), 1.61(3H, s, H-23), 1.01(3H, s, H-25), 1.09(3H, s, H-26), 1.69(3H, s, H-27), 1.64(3H, s, H-29), 1.11(3H, s, H-30), 3.59(3H, s, OCH₃). ¹³C-NMR and DEPT (100MHz, C₅D₅N) δ : 34.0 (C-1, CH₂), 26.5 (C-2, CH₂), 70.0 (C-3, CH), 43.9 (C-4, C), 50.2 (C-5, CH), 19.2 (C-6, CH₂), 34.0 (C-7, CH₂), 40.6 (C-8, C), 47.8 (C-9, CH), 37.5 (C-10, C), 24.2 (C-11, CH₂), 128.1 (C-12, CH), 139.9 (C-13, C), 42.1(C-14, C), 29.3 (C-15, CH₂), 26.4 (C-16, CH₂), 48.3 (C-17, C), 54.6 (C-18, CH), 72.7 (C-19, C), 42.4(C-20, CH), 27.0 (C-21, CH₂), 38.5 (C-22, CH₂), 23.6 (C-23, CH₃), 65.7 (C-24, CH₂), 16.0 (C-25, CH₃), 17.1 (C-26, CH₃), 24.6 (C-27, CH₃), 180.7 (C-28, COOR), 27.1(C-29, CH₃), 16.8 (C-30, CH₃), 49.7 (OCH₃).

β -daucosterol (8): White amorphous powder, m.p. 286~287°C. positive reaction in Liebermann-Burchard detection or Molish detection. IR $\nu_{\max}$ (cm⁻¹, KBr): 3739(w), 3571(s), 3451(s, br), 3405(s, br) [$\nu_{\text{O-H}}$]; 2959(s, br), 2934(s, br), 2869(s, br) [$\nu_{\text{C-H}}$]; 1659(w), 1641(w), 1631(w), 1463(s), 1436(w), 1419(w), 1379(s), 1367(s), 1254(w), 1197(w), 1164(m), 1105(s), 1074(s), 1025(s), 960(w), 800(m), 620(w), 601(w). EI-MS $m/z(\%)$: 396(100), 383, 255, 161, 135. ¹H-NMR (400MHz, C₅D₅N) δ : 3.96 (1H, m, H-3), 2.70(1H, m, H-4a), 2.46(1H, m, H-4b), 5.33 (1H, m, H-6), 0.65 (3H, m, H-18), 0.84~2.20 (42H, m), 5.04~5.06 (1H, d, J=7.6Hz, H-1'), 4.06~4.07 (1H, m, H-2'), 4.28~4.30 (2H, m, H-3', 4'), 3.97 (1H, m, H-5'), 4.55~4.58 (1H, m, H-6a'), 4.39~4.43 (1H, m, H-6b'). ¹³C-NMR and DEPT (100MHz, C₅D₅N) δ : 37.5 (C-1, CH₂), 30.3 (C-2, CH₂), 78.5 (C-3, CH), 40.0 (C-4, CH₂), 140.9 (C-5), 122.0 (C-6, CH), 32.2 (C-7, CH₂), 32.1 (C-8, CH), 50.4 (C-9, CH), 36.9 (C-10, C), 21.4 (C-11, CH₂), 39.4 (C-12, CH₂), 42.5 (C-13, C), 56.8 (C-14, CH), 24.5 (C-15, CH₂), 28.6 (C-16, CH₂), 56.3 (C-17, CH), 12.0 (C-18, CH₃), 19.5 (C-19, CH₃), 36.4 (C-20, CH), 19.0 (C-21, CH₃), 34.2 (C-22, CH₂), 26.4 (C-23, CH₂), 46.1 (C-24, CH), 29.5 (C-25, CH), 19.2 (C-26, CH₃), 20.0 (C-27, CH₃), 23.4 (C-28, CH₂), 12.2 (C-29, CH₃), 102.6(C-1', CH), 75.4 (C-2', CH), 78.6 (C-3', CH), 71.7 (C-4', CH), 78.1 (C-5', CH), 62.8 (C-6', CH₂).

p-Hydroxybenzoic acid (9): White solid, m.p. 217~219°C, soluble in methanol. EI-MS $m/z(\%)$: 138(M⁺). IR $\nu_{\max}$ (cm⁻¹, KBr): 3854(w), 3394(s, br) [$\nu_{\text{O-H}}$]; 3199(w) [$\nu_{\text{C-H}}$]; 2964(s, br), 2835(s, br) [$\nu_{\text{C-H}}$]; 2669(s, br), 2519(s, br), 2226(w), 2055(m), 1925(m), 1890(w), 1794(w), 1750(w); 1685(s, br), 1608(s, br), 1509(s), 1448(s) [$\nu_{\text{C=C}}$]; 1425(s), 1364(s), 1318(s), 1292(s), 1242(s), 1170(s), 1128(m), 1103(m), 1063(w), 1046(w), 1014(s), 931(s), 855(s), 770(s), 641(m), 639(s), 619(s), 548(s), 504(s). ¹H-NMR (400MHz, CD₃OD) δ : 5.34 (2H, d, J=12.0Hz, H-3,5), 6.40 (2H, d, J=8.0Hz, H-2,6). ¹³C-NMR and DEPT (100MHz, CD₃OD) δ : 122.7 (C-1), 116.0 (C-2,6), 133.0 (C-3, 5), 163.3 (C-4), 170.1 (COOH).

Table S1. 1D-NMR data and 2D-NMR correlations of compound 4 (in CDCl₃ at 600/150 MHz to TMS)

Position(carbon)*	δ_C	DEPT	δ_H^{**}	¹ H- ¹ H COSY	HMBC*** (proton)
1	144.8	C			3, 11
2	153.8	C			4, (3)
3	119.1	CH	7.25~7.27 (d, J=12.0 Hz)	4	
4	124.5	CH	7.99~8.01 (d, J=12.0 Hz)	3	
4a	128.1	C			3
5	124.1	CH	8.03~8.05 (d, J=12.0 Hz)	6	
6	114.8	CH	7.18~7.20 (d, J=12.0 Hz)	5	
7	157.7	C			5, (6), 13
8	148.0	C			6, 12
8a	125.4	C			5
9	180.8	C			
9a	124.2	C			4
10	180.9	C			5, 4
10a	127.2	C			6
11	61.3	CH ₃	3.95(s)		
12	55.3	CH ₃	3.92(s)		
13	60.3	CH ₃	3.93(s)		

*Serial numbers see Fig 1; ** Assign ¹H to ¹³C by HSQC; *** Two-bond correlations were shown in brackets.

14:44:28 24-Nov-2014

M141124E-14 22 (1.618) Cn (Top,2, Ht); Cm (22-9:19)

KIB

Autospec Premier

P776

Magnet EI+
1.53e4

Elemental Composition Report

Single Mass Analysis

Tolerance = 10.0 PPM / DBE: min = -10.0, max = 120.0

Selected filters: None

Monoisotopic Mass, Odd and Even Electron Ions

14 formula(e) evaluated with 1 results within limits (up to 51 closest results for each mass)

Elements Used:

C: 0-200 H: 0-400 O: 5-7

15:57:02 17-Apr-2015

Voltage EI+

KIB
M150417EA-02AFAMM 17 (1.561)
314.0791

Autospec Premier
P776
661

Minimum: -10.0
Maximum: 200.0 10.0 120.0

Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	Formula
314.0791	314.0790	0.1	0.3	11.0	5546337.5	C17 H14 O6

Figure S1 EI-MS and HREI-MS spectra of compound 4

Figure S2 IR spectra of compound 4

Bruker AV-400MHz;

H cdc13

Figure S3 ^1H -NMR spectra (400MHz or 600MHz) of compound 4

c13 and dept 100MHz

Figure S4 ^{13}C -NMR and DEPT spectra (100MHz or 150MHz) of compound 4

Figure S5 ^1H - ^1H COSY spectra of compound 4

HSQC CDC13 BRUKER AVANCE 600 (600/150MHz)

Figure S6 HSQC spectra of compound 4

Figure S7 HMBC spectra of compound 4