Structure and Activity of the *Camellia* Sapogenin Derivatives on Biofilm Inhibition of *Staphylococcus aureus* and *Escherichia coli*

Chunfang Zhu[1], Meng Zhang[1], Qiaoling Tang[1], Qian Yang[1], Yong Ye[1,2*]

[1] *Department of Pharmaceutical Engineering, School of Chemistry and Chemical Engineering, South China University of Technology, Guangzhou 510640, People's Republic of China*

[2] *Gannan Medical University Collaborative Innovation Center for Gannan Oil-tea Camellia Industrial Development, Guanzhou 341000, P R China*

**yeyong@scut.edu.cn*

Contents

Synthesis of key intermediates.

Synthesis of *Camellia* sapogenin derivative S-(1-19)

^{1}H NMR and ^{13}C NMR Spectrum for sapogenin (C_2D_6SO as solvent)

^{1}H NMR and ^{13}C NMR Spectrum for *Camellia* sapogenin derivative **S-(1-19)** (C_2D_6SO as solvent)

Synthesis of key intermediates. The desired raw material, *Camellia* sapogenin, was prepared by the purification of defatted *Camellia* oleifera seeds and acid-base hydrolysis. Sapogenin (10.1 g, 20 mmol) and triphenylchloromethane (5.57 g, 20 mmol) were dissolved in pyridine (40 mL) and reacted at room temperature for 24 h. Then the residue was washed with water, ethyl acetate and saturated sodium chloride solution, dried with sodium sulfate, and finally evaporated to remove the organic solvent. Intermediate **1a** was obtained as white solid; yield, 91.9%. Then intermediate **1a** (7.47 g, 10 mmol) was mixed with acetic anhydride (3.78 mL, 40 mmol) in DMF (40 mL), catalyzed by triethylamine (1.66 mL, 12 mmol) at room temperature for 24 h. The residue was washed, dried and evaporated in the same way to obtain Intermediate **2a** as white solid; yield, 86.0%. At last, methanol hydrochloride solution (5 mL, 10%, v/v) was added to the dichloromethane / methanol solution (1:1, v/v) of intermediate **2a** (4.58 g, 5 mmol), and the reaction was carried out at room temperature for 24 h. After washing and drying, it is purified by column chromatography to obtain intermediate 3a as white solid; yield, 82.3%.

Synthesis of *Camellia* sapogenin derivative S-(1-19).

*3β,16α,21β,22α,28-O-pentaacetyl-oleanane-12-Ene-23-aldehyde (**S-1**)*. DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in DMSO (10 mL) and reacted with acetic anhydride (115 μL, 1.2 mmol) at 40℃ for 24 h. After the reaction, the resulting mixture was diluted with distilled water and extracted with ethyl acetate. The residue was washed with hydrochloric acid, saturated sodium carbonate and saturated sodium chloride, dried with anhydrous sodium sulfate and evaporated to remove the organic solvent under reduced pressure, the final *Camellia* sapogenin derivative **S-1** was obtained as white solid (0.65 g, 90.92%). IR (KBr) 3360, 2933, 2861, 1739, 1657 cm^{-1}; ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.29 (s, 1H, CHO), 5.57 (d, J = 8.0 Hz, 1H, 21-H), 5.19 (t, 1H, 12-H), 5.13 (d, J = 12.1 Hz, 1H, 22-H), 4.35 (t, J = 5.1 Hz, 1H, 3-H), 4.19 (m, 1H, 16-H), 3.96 (d, J = 11.3 Hz, 1H, 28-H), 3.87 (d, J = 11.7 Hz, 1H, 28-H), 1.99 (s, 15H, CH_3), 1.92 (t, J = 6.6 Hz, 1H, 18-H), 1.85–1.30 (m, 16H, CH_2), 1.26 (s, 3H, CH_3, 24-H), 1.04–0.90 (s, 3H, 27-H), 0.90–0.73 (s, 12H, CH_3, 25,26,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 205.35, 170.33, 170.30, 170.20, 170.03, 157.07, 142.28, 123.28, 87.17, 77.45, 76.82, 73.18, 73.02, 54.28, 47.98, 46.71, 46.31, 46.13, 42.33, 41.49, 41.30, 38.24; 37.74, 35.63, 33.28, 31.56, 27.43, 25.81, 25.62, 23.46, 22.57, 21.49, 21.40, 21.26, 21.14, 20.97, 20.86, 17.11, 15.77, 9.66. HRMS (ESI^+), calcd for $C_{40}H_{59}O_{11}$ $(M+H)^+$, 715.4213; found, 715.4221.

*28-O-chloroacetyl-3β,16α,21β,22α-O-tetraacetyl-oleanone-12-Ene-23-aldehyde (**S-2**)*. DMAP (0.12

g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with chloracetyl chloride (90 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-1** and sapogenin derivative **S-2** was obtained as white solid (0.69 g, 92.08%). IR (KBr) 3327, 2926, 2853, 1726, 1622, 735 cm^{-1}. ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.32 (s, 1H, CHO), 5.56 (d, *J* = 8.0 Hz, 2H, 21,22-H), 5.32 (s, 1H, 12-H), 4.34 (t, 1H, *J* = 5.1 Hz, 3-H), 4.34 (s, 2H), 4.19 (s, 1H, 16-H), 3.96 (d, *J* = 11.3 Hz, 1H, 28-H), 3.87 (d, *J* = 11.7 Hz, 1H, 28-H), 1.99 (t, *J* = 6.2 Hz, 1H, 18-H), 1.72 (s, 12H, CH_3), 1.70–1.20 (m, 16H, CH_2), 1.10–0.80 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 204.46, 170.2, 169.98, 157.23, 157.08, 157.02, 144.13, 128.19, 86.46, 82.35, 76.82, 71.42, 65.45, 54.49, 47.98, 47.92, 46.98, 46.89, 45.18, 44.18, 42.86, 38.1, 36.66, 35.30, 29.48, 28.15, 25.92, 25.80, 25.81, 25.62, 25.03, 24.93, 24.92, 24.65, 24.02, 23.96, 23.82, 19.12, 19.01, 16.78. HRMS (ESI^+), calcd for $C_{40}H_{58}O_{11}Cl$ $(M+H)^+$, 749.3453; found, 746.3482.

*28-O-trifluoroacetyl-3β,16α,21β,22α-O-tetraacetyl-oleanone-12-Ene-23-aldehyde (**S-3**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in DMSO (10 mL) and reacted with trifluoroacetic anhydride (170 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-1** and sapogenin derivative **S-3** was obtained as white solid (0.69 g, 89.74%). IR (KBr) 3329, 2926, 2864, 1778, 1729, 1624, 1084 cm^{-1}. ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.26 (s, 1H, CHO), 5.57 (d, 1H, 21-H), 5.32 (d, *J* = 4.6 Hz, 1H, 22-H), 5.19 (t, 1H, 12-H), 4.22 (t, *J* = 6.5 Hz, 1H, 3-H), 4.16 (m, 1H, 16-H), 4.03–3.92 (d, 2H, 28-H), 1.99 (t, *J* = 10.2 Hz, 1H, 18-H), 1.86 (s, 12H, CH_3), 1.80–1.20 (m, 16H, CH_2), 1.13 (s, 3H, CH_3, 24-H), 1.06 (s, 3H, CH_3, 27-H), 1.01–0.79 (s, 12H, CH_3, 25,26,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.45, 171.05, 164.49, 157.08, 157.02, 150.07, 140.06, 129.11, 117.34, 88.16, 77.24, 74.26, 69.07, 63.56, 55.28, 47.98, 47.27, 47.01, 46.86, 42.36, 41.32, 39.30, 38.17, 34.13, 33.81, 31.97, 31.76, 30.48, 30.19, 29.49, 29.17, 29.03, 25.80, 25.60, 25.46, 25.32, 24.93, 22.56, 19.12, 14.39. HRMS (ESI^+), calcd for $C_{40}H_{56}O_{11}F_3$ $(M+H)^+$, 769.3562; found, 769.3530.

*28-maleic acid monoester-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-4**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in DMSO (10 mL) and reacted with maleic anhydride (0.12 g, 1.2 mmol) at 40℃ for 24 h. After the reaction, the resulting mixture was purified by the same method as **S-1** and then dissolved in dichloromethane / methanol (1: 1, v: v) to react with acetylchloride (11 μL, 0.15 mmol) at room temperature for 6 h. The solvent was removed and the product was extracted with ethyl acetate, and the ethyl acetate was removed under

reduced pressure. Sapogenin derivative **S-4** was obtained as white solid (0.5 g, 82.95%). IR (KBr) 3337, 2931, 2864, 1718, 1655, 720 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 12.48 (s, 1H, COOH), 9.38 (s, 1H, CHO), 6.94 (s, 1H, OH, 3-H), 6.44 (d, *J* = 2.1 Hz, 1H, CH), 6.19 (d, *J* = 2.1 Hz, 1H, CH), 5.32 (s, 2H, OH, 21,22-H), 5.18 (t, 1H, 12-H), 5.04 (s, 1H, 16-OH), 4.17 (d, *J* = 3.3 Hz, 1H, 28-H), 3.97 (m, 1H, 3-H), 3.89 (d, J = 6.5 Hz, 1H, 28-H), 3.35–3.30 (m, 3H,16,21,22-H), 1.99 (t, J = 12.0 Hz, 1H, 18-H), 1.90–1.30 (m, 16H, CH_2), 1.23 (s, 3H, CH_3, 24-H), 1.06 (s, 3H, 27-H), 0.95–0.82(s, 12H, CH_3, 25,26 29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.50, 164.35, 161.18, 144.01, 136.13, 129.94, 122.14, 77.30, 71.63, 71.56, 70.92, 67.60, 55.59, 47.41, 46.82, 46.65, 46.48, 44.06, 41.73, 41.71, 38.88, 36.72, 35.70, 33.99, 31.59, 27.47, 25.45, 25.42, 23.44, 22.57, 21.33, 16.83, 15.95, 9.35. HRMS (ESI^+), calcd for $C_{34}H_{51}O_9$ $(M+H)^+$, 603.3568; found, 603.3562.

*28-O-benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-5**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with benzoyl chloride (115 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-5** was obtained as light-yellow powder (0.51 g, 83.77%). IR (KBr) 3535, 3065, 2938, 2858, 1719, 1595, 1525, 1456, 1380, 711.5 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 9.42 (s, 1H, CHO), 7.99 (d, 2H), 7.76 (dd, J_1 = 8.5 Hz, J_2 = 1.7 Hz, 1H), 7.55 (m, *J* = 7.6 Hz, 2H), 6.74 (s, 1H, OH, 3-H), 5.58 (s, 1H, OH, 21-H), 5.55 (s, 1H, OH, 22-H), 5.07 (t, 1H, 12-H), 4.64 (s, 1H, OH, 16-H), 4.31 (d, *J* = 18.3 Hz, 1H, 28-H), 4.02 (d, *J* = 12.3 Hz, 1H, 28-H), 3.83 (t, 1H, 3-H), 3.70–3.51(m, 3H, 16,21,22-H), 1.91 (t, *J* = 4.5 Hz, 1H, 18-H), 1.87–1.30 (m, 16H, CH_2), 1.25 (s, 3H, CH_3, 24-H), 1.09–0.81 (s, 15H, CH_3, 25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 205.50, 165.79, 142.53, 133.91, 130.79, 129.78, 129.74, 128.96, 128.91, 123.41, 81.38, 74.70, 73.99, 68.63, 67.93, 54.56, 47.05, 46.69, 46.34, 46.15, 43.04, 41.79, 41.48, 38.18, 36.86, 35.70, 33.31, 31.77, 27.43, 25.07, 25.02, 23.47, 22.56, 20.53, 17.28, 15.78, 9.79. HRMS (ESI^+), calcd for $C_{37}H_{53}O_7$ $(M+H)^+$, 609.5747; found, 609.5750.

*28-O-m-chloro benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-6**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 3-chlorophenyl methyl chloride (155 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-6** was obtained as light-yellow powder (0.51 g, 79.28%). IR (KBr) 3440, 3070, 1642, 1572, 1532, 1462, 1120, 1093, 890, 798, 741 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 9.39 (s, 1H, CHO), 7.96 (s, 1H), 7.89 (d, *J* = 7.9 Hz, 1H), 7.52 (d, *J* = 7.8 Hz, 1H),

7.41 (m, J = 2.3 Hz, 1H), 6.20 (s, 1H, OH, 3-H), 5.77 (s, 1H, OH, 21-H), 5.75 (s, 1H, OH, 22-H), 5.45 (s, 1H, OH, 16-H), 5.14 (t, 1H, 12-H), 4.33 (d, J = 5.1 Hz, 1H, 28-H), 3.96 (d, J = 7.0 Hz, 1H, 28-H), 3.60–3.40 (m, 4H, 3,16,21,22-H), 2.18 (t, J = 7.2 Hz, 1H, 18-H), 1.90–1.20 (m, 16H, CH_2), 1.15–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 208.44, 166.89, 139.61, 134.00, 133.86, 131.92, 130.88, 129.28, 128.99, 126.05, 74.60, 73.56, 71.13, 65.50, 65.45, 54.12, 49.54, 48.69, 48.36, 48.15, 45.04, 43.79, 42.38, 38.14, 36.81, 35.69, 33.23, 30.49, 27.04, 25.57, 25.43, 23.44, 22.56, 19.12, 17.42, 13.99, 9.74. HRMS (ESI$^+$), calcd for $C_{37}H_{52}O_7Cl$ $(M+H)^+$, 643.3362; found, 643.3364.

*28-O-m-fluoro benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-7**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 3-fluorobenzophenyl chloride (145 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-7** was obtained as light-yellow powder (0.49 g, 78.18%). IR (KBr) 3442, 3074, 2928, 2858, 1720, 1644, 1589, 1535, 1480, 1201; 1084, 971, 888, 801, 751, 677 cm^{-1}. ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.40 (s, 1H, CHO), 7.95 (s, 1H), 7.79 (d, J = 7.2 Hz, 1H), 7.58 (s, J = 8.0 Hz, 1H), 7.29 (d, J = 8.1 Hz, 1H), 6.29 (s, 1H, OH, 3-H), 5.35 (s, 1H, OH, 21-H), 5.26 (s, 1H, OH, 22-H), 5.14 (t, 1H, 12-H), 4.61 (s, 1H, OH, 16-H), 4.34 (t, J = 5.0 Hz, 1H, 28-H), 4.03 (d, J = 5.0 Hz, 1H, 28-H), 3.69 (t, 1H, 3-H), 3.56–3.42 (m, 3H, 16,21,22-H), 1.92 (t, J = 12.9 Hz, 1H, 18-H), 1.90–1.20 (m, 16H, CH_2), 1.15–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 205.44, 167.06, 162.71, 139.96, 130.57, 130.52, 125.41, 123.40, 117.39, 114.51, 76.71, 75.70, 74.52, 68.35, 65.45, 54.08, 49.51, 47.18, 46.61, 46.11, 43.04, 41.64, 41.47, 35.69, 33.23, 31.67, 30.77, 30.48, 29.50, 25.57, 25.41, 23.45, 22.55, 19.12, 15.75, 13.99, 9.75. HRMS (ESI$^+$), calcd for $C_{37}H_{52}O_7F$ $(M+H)^+$, 627.4234; found, 627.4242.

*28-O-m-bromo benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-8**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 3-bromobenzoyl chloride (160 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-8** was obtained as gray powder (0.5 g, 72.71%). IR (KBr) 3523, 3068, 2928, 2858, 1724, 1594, 1510, 1469, 747, 809, 533 cm^{-1}. ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.40 (s, 1H, CHO), 8.05 (s, 1H), 7.94 (d, J = 7.8 Hz, 1H), 7.84 (d, J = 8.8 Hz, 1H), 7.50 (m, J = 7.9 Hz, 1H), 6.74 (s, 1H, OH, 3-H), 5.61 (s, 1H, OH, 21-H), 5.45 (s, 1H, OH, 22-H), 5.14 (t, 1H, 12-H), 4.59 (s, 1H, OH, 16-H), 4.33 (d, J = 5.1 Hz, 1H, 28-H), 3.96 (d, J = 7.0 Hz, 1H, 28-H), 3.57 (t, 1H, 3-H), 3.45 (m, 3H, 16,21,22-H), 1.92 (t, J = 9.4 Hz, 1H, 18-H), 2.0–1.20 (m, 16H, CH_2), 1.19–

0.80 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 205.96, 166.39, 144.53, 136.00, 132.20, 131.89, 131.32, 128.71, 122.32, 122.16, 77.26, 74.68, 70.95, 67.63, 60.20, 56.50, 47.97, 46.82, 46.65, 44.75, 43.09, 41.68, 41.50, 38.86, 36.89, 35.95, 33.23, 31.75, 27.40, 24.97, 23.49, 21.52, 21.50, 21.21, 19.12, 19.01, 14.54. HRMS (ESI^+), calcd for $C_{37}H_{52}O_7Br$ $(M+H)^+$, 688.4792; found, 688.4782.

*28-O-m-methyl benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-9**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 3-methylbenzophenyl chloride (160 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-9** was obtained as white powder (0.53 g, 85.09%). IR (KBr) 3517, 3065, 2938, 2858, 1714, 1595, 1525, 1456, 746, 806 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 9.42 (s, 1H, CHO), 7.82 (d, *J* = 8.6 Hz, 1H), 7.76 (s, 1H), 7.42 (m, *J* = 8.3 Hz, 1H), 7.18 (d, *J* = 7.8 Hz, 1H), 6.72 (s, 1H, OH, 3-H), 5.31 (s, 2H, OH, 21,22-H), 5.04 (t, *J* = 4.4 Hz, 1H, 12-H), 4.65 (s, 1H, OH, 16-H), 4.47 (d, *J* = 10.4 Hz, 1H, 28-H), 3.99 (d, *J* = 25.1 Hz, 1H, 28-H), 3.82 (t, 1H, 3-H), 3.42–3.30 (m, 3H,16,21,22-H), 2.36 (s, 3H, CH_3), 2.17 (t, *J* = 5.2 Hz, 1H, 18-H), 1.83–1.23 (m, 16H, CH_2), 1.20–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 204.61, 165.95, 142.20, 138.30, 133.86, 130.18, 129.99, 128.87, 126.90, 123.09, 77.28, 74.67, 70.68, 68.16, 59.55, 55.30, 47.05, 46.69, 46.56, 45.43, 43.98, 41.57, 41.52, 38.87, 36.91, 35.70, 33.69, 33.30, 27.42, 25.12, 24.97, 23.50, 22.57, 21.28, 21.05, 16.48, 15.77, 14.38. HRMS (ESI^+), calcd for $C_{38}H_{55}O_7$ $(M+H)^+$, 623.5902; found, 623.5913.

*28-O-p-nitro benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-10**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 4-nitrobenzoyl chloride (0.22 g, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-10** was obtained as light-yellow powder (0.53 g, 81.07%). IR (KBr) 3509, 3112, 2928, 2860, 1724, 1604, 1459, 1529, 1350, 840 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 9.43 (s, 1H, CHO), 8.37 (d, *J* = 14.1 Hz, 2H), 8.20 (d, *J* = 10.3 Hz, 2H), 6.74 (s, 1H, OH, 3-H), 5.33 (s, 2H, OH, 21,22-H), 5.05 (t, *J* = 4.4 Hz, 1H, 12-H), 4.70 (s, 1H, OH, 16-H), 4.33 (d, *J* = 4.6 Hz, 1H, 28-H), 4.02 (d, *J* = 12.2 Hz, 1H, 28-H), 3.85 (t, 1H, 3-H), 3.65 – 3.50 (m, 3H, 16,21, 22-H), 1.99 (t, *J* = 7.0 Hz, 1H, 18-H), 1.85–1.10 (m, 16H, CH_2), 1.05–0.70 (s, 18H, CH_3, 24, 25, 26, 27, 29, 30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 202.29, 164.49, 150.80, 142.22, 131.40, 130.08, 130.08, 124.44, 124.12, 123.93, 75.41, 70.89, 67.97, 65.83, 62.26, 55.29, 47.01, 46.69, 46.60, 45.05, 43.65,

41.57, 41.37, 38.87, 36.91, 34.13, 31.77, 27.05, 25.10, 24.96, 23.53, 22.56, 20.63, 19.12, 16.48, 15.84, 14.39. HRMS (ESI$^+$), calcd for $C_{37}H_{52}NO_9$ $(M+H)^+$, 654.6572; found, 654.6564.

28-O-p-methoxy benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde ***(S-11)***. DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 4-methoxybenzophenyl chloride (0.20 g, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-11** was obtained as white powder (0.49 g, 76.70%). IR (KBr) 3492, 3065, 2937, 2858, 1687, 1605, 1510, 1458, 1259, 1028, 746, 806 cm^{-1}. ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.25 (s, 1H, 23-H), 7.89 (d, *J* = 8.9 Hz, 2H), 7.02 (d, *J* = 8.9 Hz, 2H), 6.80 (s, 1H, OH, 3-H), 5.32 (s, 2H, OH, 21,22-H), 5.03 (t, *J* = 4.4 Hz, 1H, 12-H), 4.65 (s, 1H, OH, 16-H), 4.34 (d, *J* = 5.1 Hz, 1H, 28-H), 4.15 (d, *J* = 5.8 Hz, 1H, 28-H), 3.89 (t, 1H, 3-H), 3.83 (s, 3H, CH_3), 3.50–3.39 (m, 3H, 16,21,22-H), 2.09 (t, *J* = 3.6 Hz, 1H, 18-H), 1.90–1.20 (m, 16H, CH_2), 1.10–0.80 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.69, 167.44, 165.61, 143.32, 131.79, 131.41, 123.45, 122.66, 114.55, 114.24, 77.29, 75.42 , 71.66, 67.95, 64.28, 55.89, 55.86, 47.03, 46.80, 46.65, 46.47, 43.94, 41.67, 41.16, 38.87, 36.91, 35.69, 33.71, 31.76, 27.48, 25.33, 25.18, 23.49, 22.56, 19.01, 16.49, 14.39, 9.35. HRMS (ESI$^+$), calcd for $C_{38}H_{55}O_8$ $(M+H)^+$, 639.6842; found, 639.6849.

28-O-(2-thiophenoformyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde ***(S-12)***. DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 2-thiophenoformyl chloride (130 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-12** was obtained as yellow powder (0.53 g, 86.20%). IR (KBr) 3500, 3100, 2931, 2858, 1702, 1643, 1525, 744 cm^{-1}. ^{1}H NMR (400 MHz, C_2D_6SO) δ: 9.37 (s, 1H, CHO), 7.92 (d, *J* = 15.7 Hz, 1H), 7.80 (d, *J* = 14.5 Hz, 1H), 7.12 (t, *J* = 4.3 Hz, 1H), 6.76 (s, 1H, OH, 3-H), 5.30 (s, 2H, OH, 21,22-H), 5.19 (t, 1H, 12-H), 4.59 (s, 1H, OH, 16-H), 4.22 (d, *J* = 6.5 Hz, 1H, 28-H), 4.03 (d, *J* = 7.1 Hz, 1H, 28-H), 3.44 (m, 4H, 3,16,21,22-H), 1.98 (t, *J* = 7.6 Hz, 1H, 18-H), 1.90 – 1.20 (m, 16H, CH_2), 1.10 – 0.80 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 208.35, 170.74, 143.96, 131.93, 130.08, 129.10, 128.08, 122.02, 77.29, 74.46, 71.64, 67.61, 60.22, 55.32, 47.41, 46.98, 46.65, 45.18, 44.06, 41.65, 41.27, 38.88, 36.93, 34.13, 33.78, 31.77, 27.25, 25.44, 24.97, 23.45, 22.57, 21.49, 19.19, 16.88, 14.40. HRMS (ESI$^+$), calcd for $C_{35}H_{51}O_7S$ $(M+H)^+$, 615.4317; found, 615.4326.

28-O-(2-furfuroyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde ***(S-13)***.

DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with 2-furan formyl chloride (120 μL, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-13** was obtained as white powder (0.49 g, 81.84%). IR (KBr) 3520, 3132, 2930, 2858, 1713, 1642, 1390 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 9.36 (s, 1H, CHO), 7.94 (d, J = 17.3 Hz, 1H), 7.27 (d, J = 3.5 Hz, 1H), 6.67 (dd, J_1 = 3.4 Hz, J_2 = 1.7 Hz, 1H), 6.59 (s, 1H, OH, 3-H), 5.31 (s, 2H, OH, 21,22-H), 5.20 (t, J = 5.2 Hz, 1H, 12-H), 4.94 (s, 1H, OH, 16-H), 4.35 (d, J = 10.1 Hz, 1H, 28-H), 3.95 (d, J = 6.9 Hz, 1H, 28-H), 3.64 (t, 1H, 3-H), 3.57 (m, 3H, 16,21, 22-H), 1.98 (t, J = 6.4 Hz, 1H, 18-H), 1.85–1.20 (m, 16H, CH_2), 1.10–0.88 (s, 18H, CH_3, 24,25,26,27, 29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 205.45, 158.01, 148.06, 144.79, 143.95, 123.44, 118.32, 112.56, 74.25, 73.91, 68.00, 67.67, 67.64, 54.48, 47.00, 46.63, 46.12, 46.11, 42.95, 41.59, 41.46, 38.06, 36.83, 35.67, 33.26, 31.76, 27.46, 24.99, 23.43, 22.56, 20.46, 17.15, 15.77, 9.72. HRMS (ESI^+), calcd for $C_{35}H_{50}O_8$ $(M+H)^+$, 599.4232; found, 599.4248.

28-O-(2-hydroxybenzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde ***(S-14)***. Salicylic acid (0.17 g, 1.2 mmol) , EDC·HCl (0.19 g, 1.0 mmol) and DMAP (0.12 g, 1.0 mmol) were dissolved in pyridine (10 mL), activated in ice water bath for 2 h and reacted with intermediate **3a** at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-14** was obtained as white powder (0.39 g, 62.22%). IR (KBr) 3389, 2930; 2858, 1719, 1673, 1600, 1557, 1457, 757 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 10.45 (s, 1H, OH), 9.25 (s, 1H, CHO), 7.78 (d, J = 8.0 Hz, 1H), 7.52 (t, J = 7.7 Hz, 1H), 6.96 (m, J = 8.0 Hz, 2H), 6.73 (s, 1H, OH, 3-H), 5.56 (s, 2H, OH, 21,22-H), 5.32 (t, J = 4.7 Hz, 1H, 12-H), 4.64 (s, 1H, OH, 16-H), 4.32 (d, J = 4.4 Hz, 1H, 28-H), 3.93 (d, J = 7.4 Hz, 1H, 28-H), 3.70–3.60 (m, 4H, 3,16,21,22-H), 1.99 (t, J = 10.4 Hz, 1H, 18-H), 1.90–1.20 (m, 16H, CH_2), 1.10–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.50, 168.37, 160.19, 143.17, 136.05, 130.43, 122.03, 119.85, 118.22, 113.75, 77.29, 70.05, 69.03, 67.37, 62.96, 55.58, 47.99, 47.00, 46.62, 45.36, 43.58, 41.66, 41.55, 38.88, 36.92, 34.13, 33.68, 31.77, 27.05, 25.80, 24.96, 23.48, 22.56, 16.49, 14.39, 9.35. HRMS (ESI^+), calcd for $C_{37}H_{53}O_8$ $(M+H)^+$, 625.3642; found, 625.3645.

28-O-(5-bromo-2-hydroxybenzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde ***(S-15)***. 5-Bromosalicylic acid (0.26 g, 1.2 mmol), EDC·HCl (0.19 g, 1.0 mmol) and DMAP (0.12 g, 1.0 mmol) were dissolved in pyridine (10 mL), activated in ice water bath for 2 h and reacted with intermediate **3a** at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin

derivative **S-15** was obtained as white powder (0.42 g, 59.52%). IR (KBr) 3389, 2930, 2858, 1719, 1673, 1600, 1557, 1457, 829, 883, 532 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 10.49 (s, 1H, OH), 9.25 (s, 1H, 23-H), 7.86 (s, 1H), 7.67 (d, *J* = 5.7 Hz, 1H), 6.95 (d, *J* = 8.4 Hz, 1H), 6.73 (s, 1H, OH, 3-H), 5.56 (s, 2H, OH, 21,22-H), 5.23 (t, *J* = 6.7 Hz, 1H, 12-H), 4.62 (s, 1H, OH, 16-H), 4.19 (d, *J* = 6.4 Hz, 1H, 28-H), 3.91 (d, *J* = 11.0 Hz, 1H, 28-H), 3.70 – 3.60 (m, 4H, 3,16,21,22-H), 1.87 (t, *J* = 5.3 Hz, 1H, 18-H), 1.85–1.10 (m, 16H, CH_2), 1.10–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.51, 166.63, 157.09, 144.43, 138.19, 132.72, 122.79, 120.79, 120.44, 110.51, 75.83, 74.70, 70.92, 70.03, 67.22, 55.58, 47.99, 47.00, 46.61, 45.33, 43.53, 41.75, 41.60, 38.29, 36.91, 35.69, 33.66, 31.57, 29.50, 27.26, 25.80, 23.48, 19.06, 16.87, 15.92, 9.35. HRMS (ESI^+), calcd for $C_{37}H_{52}O_8Br$ $(M+H)^+$, 704.2767; found, 704.2760.

*28-O-(2-mercapto-4-methyl-5-thiazolyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-16**).* 2-Mercapto-4-methyl-5-thiazolylacetic acid (0.23 g, 1.2 mmol), EDC·HCl (0.19 g, 1.0 mmol) and DMAP (0.12 g, 1.0 mmol) were dissolved in pyridine (10 mL), activated in ice water bath for 2 h and reacted with intermediate **3a** at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-16** was obtained as yellow powder (0.44 g, 65.09%). IR (KBr) 3354, 2931, 2858, 2597, 1714, 1636 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 12.92 (s, 1H, SH), 9.25 (s, 1H, 23-H), 7.69 (s, 2H), 6.74 (s, 1H, OH, 3-H), 5.56 (s, 2H, OH, 21,22-H), 5.18 (t, *J* = 3.6 Hz, 1H, 12-H), 4.63 (s, 1H, OH, 16-H), 4.22 (d, *J* = 12.0 Hz, 1H, 28-H), 3.99 (d, *J* = 5.8 Hz, 1H, 28-H), 3.70–3.50 (m, 4H, 3,16,21,22-H), 2.20 (t, *J* = 6.4 Hz, 1H, 18-H), 2.0 (s, 3H, CH_3), 1.80–1.20 (m, 16H, CH_2), 1.0–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.55, 162.80, 157.10, 153.36, 144.03, 131.96, 121.84, 77.29, 71.50, 70.92, 67.60, 65.48, 55.60, 47.99, 47.42, 46.81, 45.18, 44.06, 41.74, 41.23, 38.88, 36.93, 35.71, 34.00, 32.03, 30.74, 28.68, 27.32, 25.80, 24.93, 23.43, 20.70, 16.84, 15.84, 14.01, 9.34. HRMS (ESI^+), calcd for $C_{36}H_{54}NO_7S_2$ $(M+H)^+$, 676.6276; found, 676.6274.

*28-O-(2-formic acid-benzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-17**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in DMSO (10 mL) and reacted with phthalic anhydride (0.18 g, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-17** was obtained as white powder (0.54 g, 82.72%). IR (KBr) 3448, 3063, 2932, 2863, 1715, 1650, 1601, 1505, 1454, 1170, 747 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 12.48 (s, 1H, COOH), 9.38 (s, 1H, CHO), 8.09–8.00 (m, 2H), 7.68–7.57 (m, 2H), 6.44 (s, 1H, OH, 3-H), 5.53 (s, 1H, OH, 21-H), 5.47 (s, 1H, OH, 22-H), 5.17 (t, 1H, 12-H), 4.65 (s, 1H, OH,

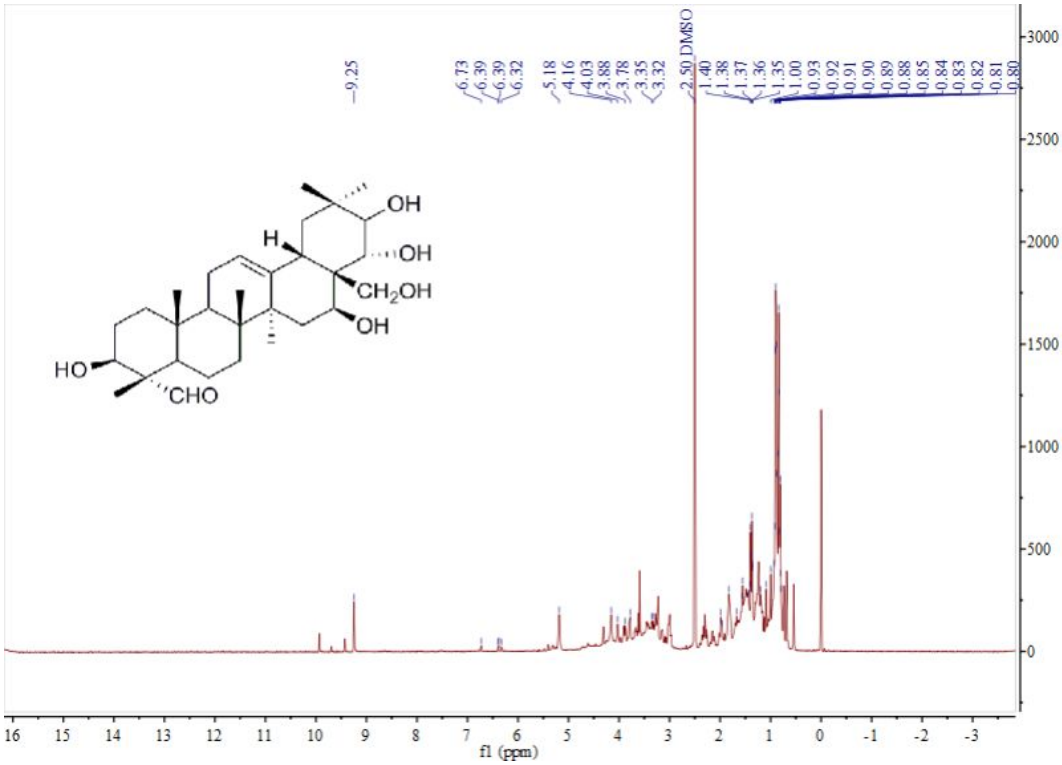
16-H), 4.30 (d, *J* = 11.6 Hz, 1H, 28-H), 3.97 (d, *J* = 6.8 Hz, 1H, 28-H), 3.70–3.55 (m, 4H, 3,16,21,22-H), 1.99 (t, *J* = 13.5 Hz, 1H, 18-H), 1.90–1.20 (m, 16H, CH_2), 1.0–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.42, 176.36, 164.35, 161.18, 159.65, 156.64, 147.27, 136.12, 129.95, 122.14, 77.30, 71.58, 70.92, 67.62, 60.32, 55.31, 47.42, 46.65, 45.18, 44.06, 43.04, 41.65, 41.24, 38.87, 36.92, 33.79, 32.82, 31.59, 28.67, 27.26, 25.45, 25.18, 23.44, 22.56, 16.88, 16.49, 15.83. HRMS (ESI^+), calcd for $C_{38}H_{53}O_9$ $(M+H)^+$, 653.4211; found, 653.4210.

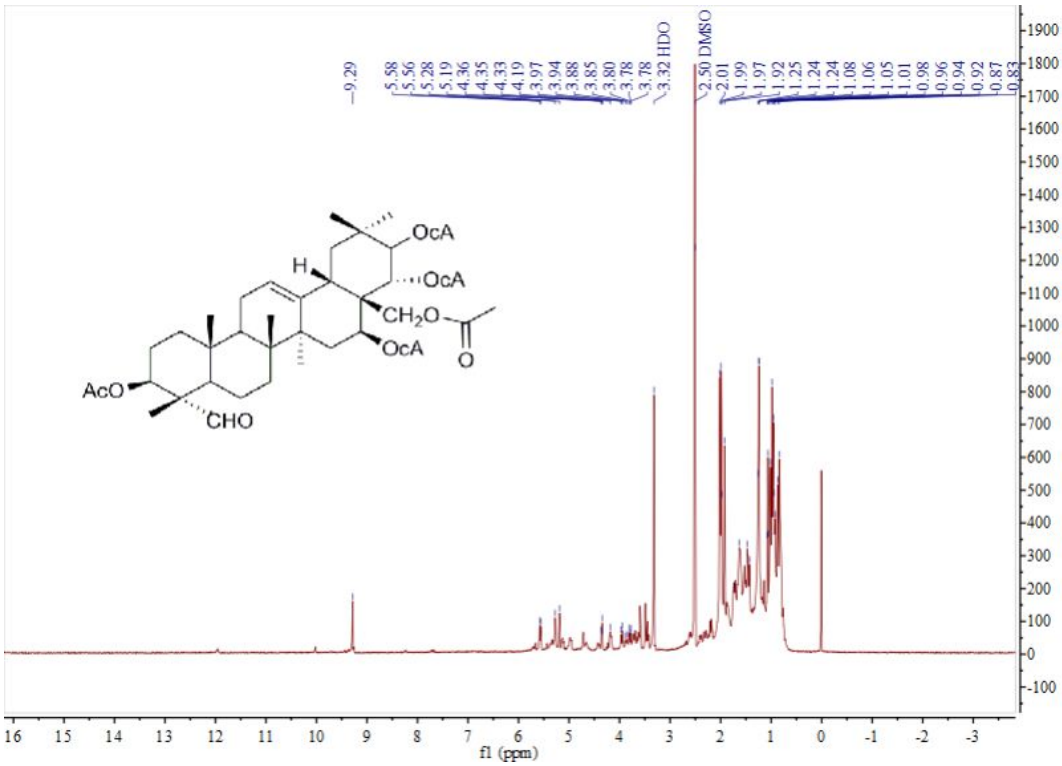
*28-O-(3,4,5,6-tetrachloro-2-formicacid-benzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-18**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in DMSO (10 mL) and reacted with tetrachloro phthalic anhydride (0.34 g, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-18** was obtained as white powder (0.6 g, 82.72%). IR (KBr) 3336, 2933, 2858, 1717, 1653, 1600, 1506, 1454, 1316, 1172, 763 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 12.48 (s, 1H, COOH), 9.38 (s, 1H, CHO), 6.45 (s, 1H, OH, 3-H), 6.20 (s, 2H, OH, 21,22-H), 5.19 (t, 1H, 12-H), 4.56 (s, 1H, OH, 16-H), 4.29 (d, 1H, 28-H), 3.93 (d, 1H, 28-H), 3.70–3.55 (m, 4H, 3,16,21,22-H), 2.0 (t, *J* = 13.1 Hz, 1H, 18-H), 1.95–1.20 (m, 16H, CH_2), 1.0–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.53, 164.35, 161.18, 144.00, 136.12, 129.95, 122.14, 122.05, 121.83, 77.30, 71.59, 70.92, 67.60, 67.38, 55.59, 47.41, 46.82, 46.64, 46.48, 44.06, 41.73, 41.24, 38.87, 36.92, 35.70, 33.99, 32.82, 27.32, 26.47, 25.45, 23.44, 20.70, 18.49, 15.95, 9.35. HRMS (ESI^+), calcd for $C_{38}H_{49}O_9Cl_4$ $(M+H)^+$, 790.8623; found, 790.8629.

*28-O-(4-acetylamino-benzenesulfonyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (**S-19**).* DMAP (0.12 g, 1.0 mmol) and intermediate **3a** (0.67 g, 1.0 mmol) were dissolved in pyridine (10 mL) and reacted with paracetamylbenzene sulfonyl chloride (0.28 g, 1.2 mmol) at 40℃ for 24 h. The post-processing was the same as that of **S-4** and sapogenin derivative **S-19** was obtained as white powder (0.55 g, 83.35%). IR (KBr) 3419, 3112, 2935, 2928, 1719, 1640, 1597, 1458, 842, 783 cm^{-1}. ^1H NMR (400 MHz, C_2D_6SO) δ: 10.50 (s, 1H, NH), 9.25 (s, 1H, 23-H), 7.86 (d, *J* = 5.7 Hz, 2H), 7.69 (d, *J* = 13.3 Hz, 2H), 6.73 (s, 1H, OH, 3-H), 6.39 (s, 2H, OH, 21,22-H), 5.18 (t, 1H, 12-H), 4.62 (s, 1H, OH, 16-H), 4.21 (d, *J* = 9.8 Hz, 1H, 28-H), 3.92 (d, *J* = 5.5 Hz, 1H, 28-H), 3.80–3.55 (m, 4H, 3,16,21,22-H), 2.10 (s, 3H, CH_3), 1.99 (t, *J* = 12.0 Hz, 1H, 18-H), 1.90 – 1.20 (m, 16H, CH_2), 1.0–0.70 (s, 18H, CH_3, 24,25,26,27,29,30-H). ^{13}C NMR (151 MHz, C_2D_6SO) δ: 207.51, 167.40, 146.96, 144.02, 143.64, 129.11, 122.08, 119.27, 77.30, 71.53, 70.81, 67.31, 65.47, 55.59, 47.62, 47.42, 46.64, 45.18,

42.35, 41.74, 41.24, 38.88, 36.63, 35.80, 35.71, 32.04, 28.68, 27.32, 25.46, 24.69, 23.44, 19.22, 16.83, 15.96, 9.35. HRMS (ESI$^+$), calcd for $C_{38}H_{56}NO_9S$ $(M+H)^+$, 702.5692; found, 702.5702.

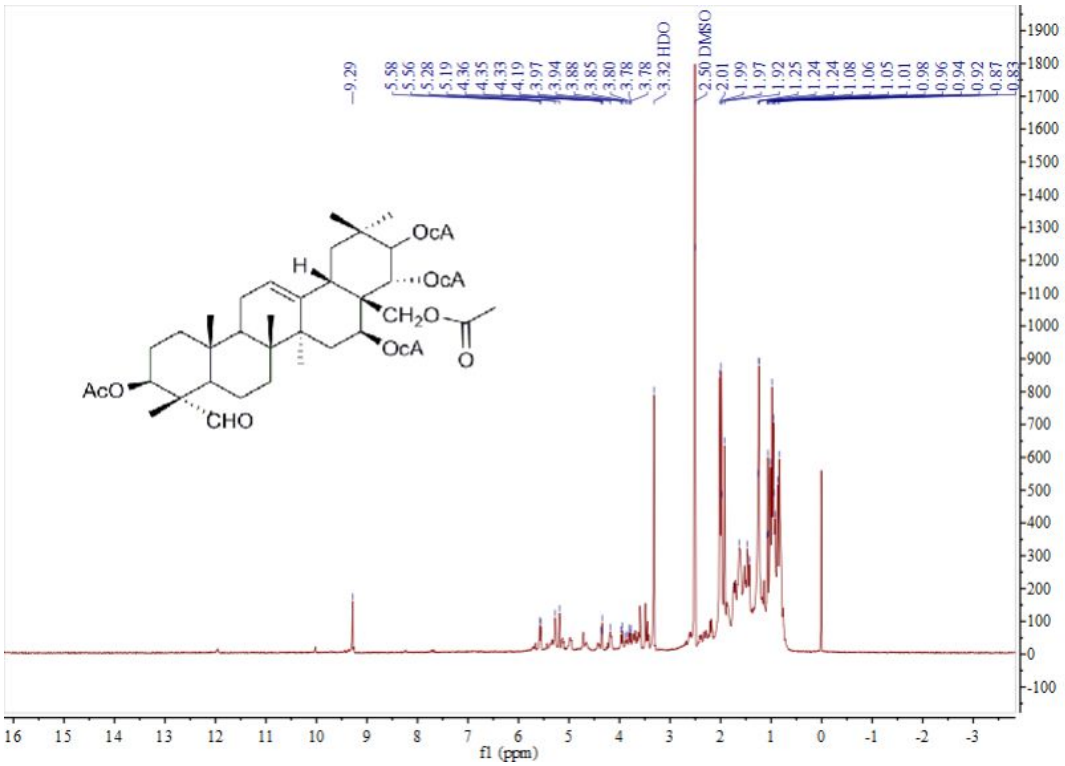
^{1}H NMR and ^{13}C NMR Spectrum for sapogenin (C_2D_6SO as solvent)

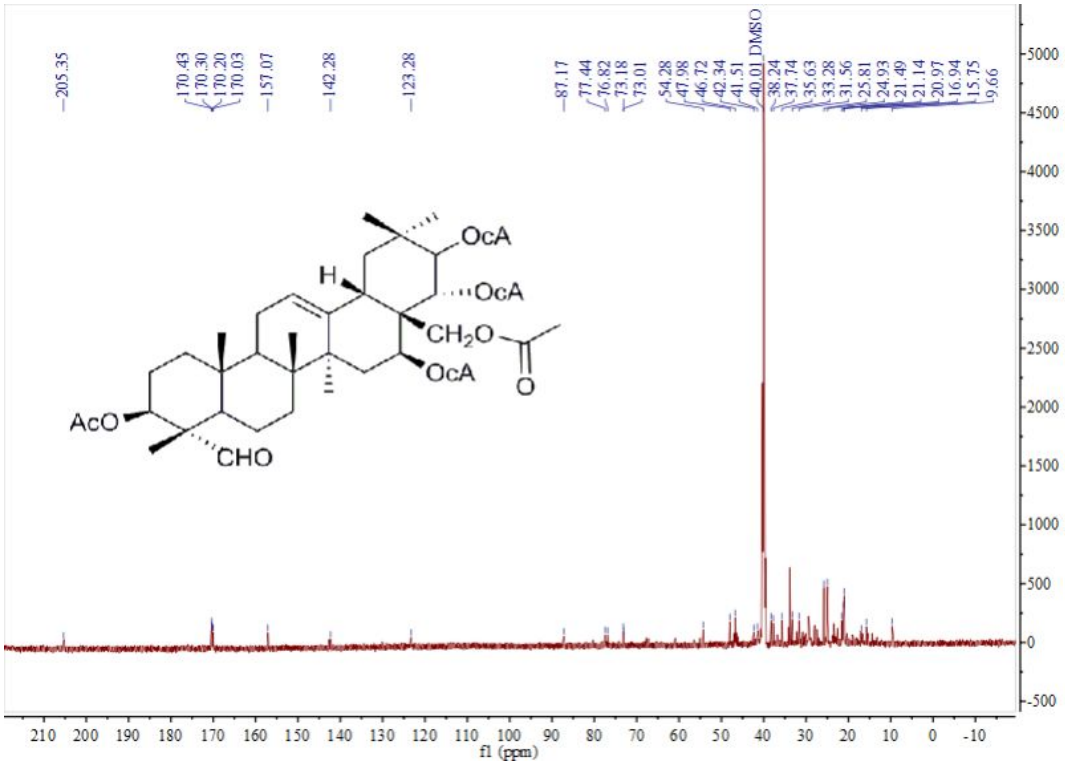




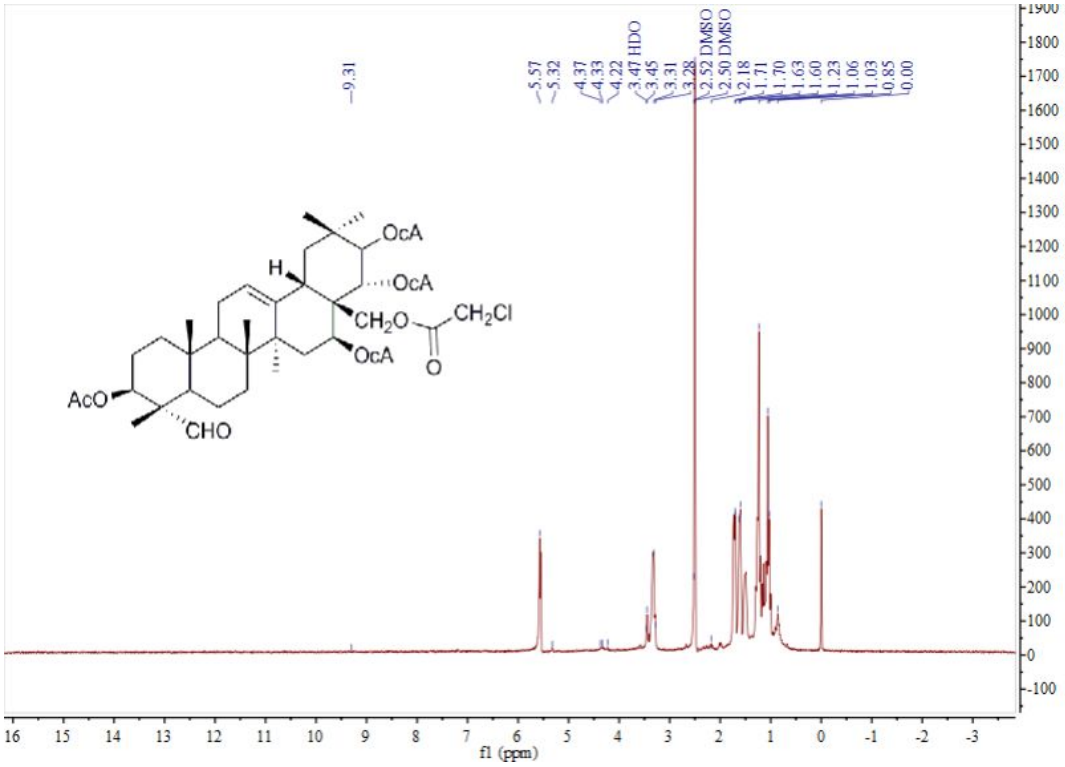
^{1}H NMR and ^{13}C NMR Spectrum for derivative S-(1-19) (C_2D_6SO as solvent)

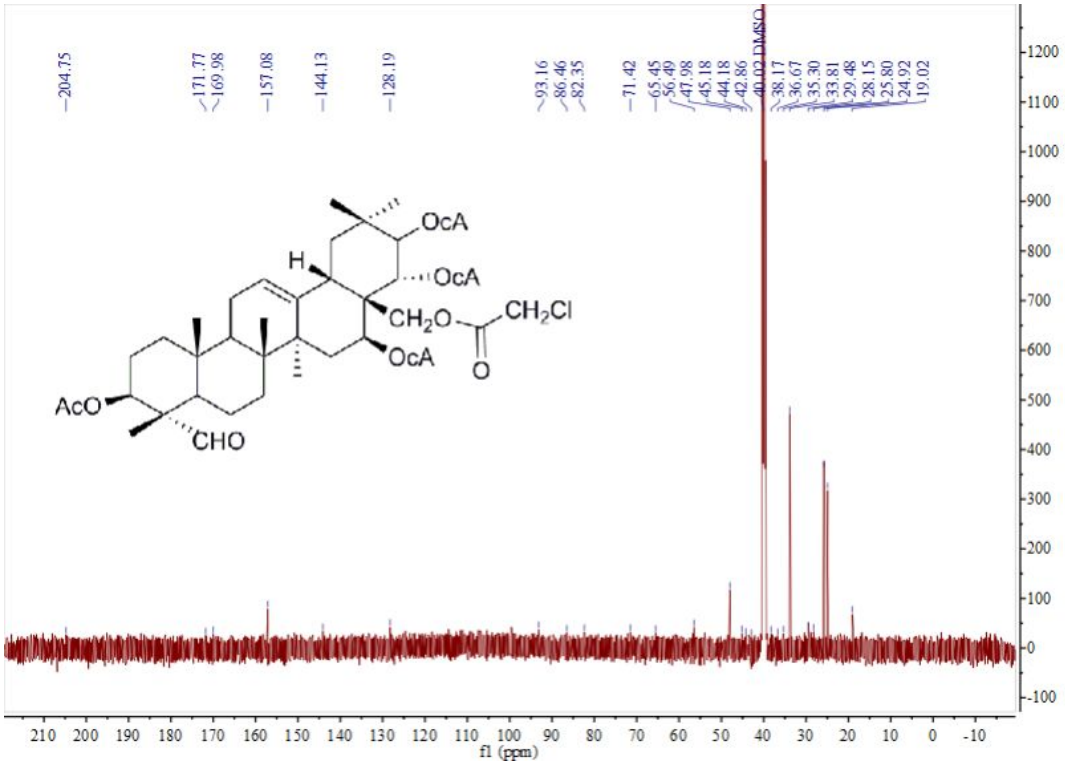
1) ***3β,16α,21β,22α,28-O-pentaacetyl-oleanane-12-Ene-23-aldehyde*** **(S-1)**



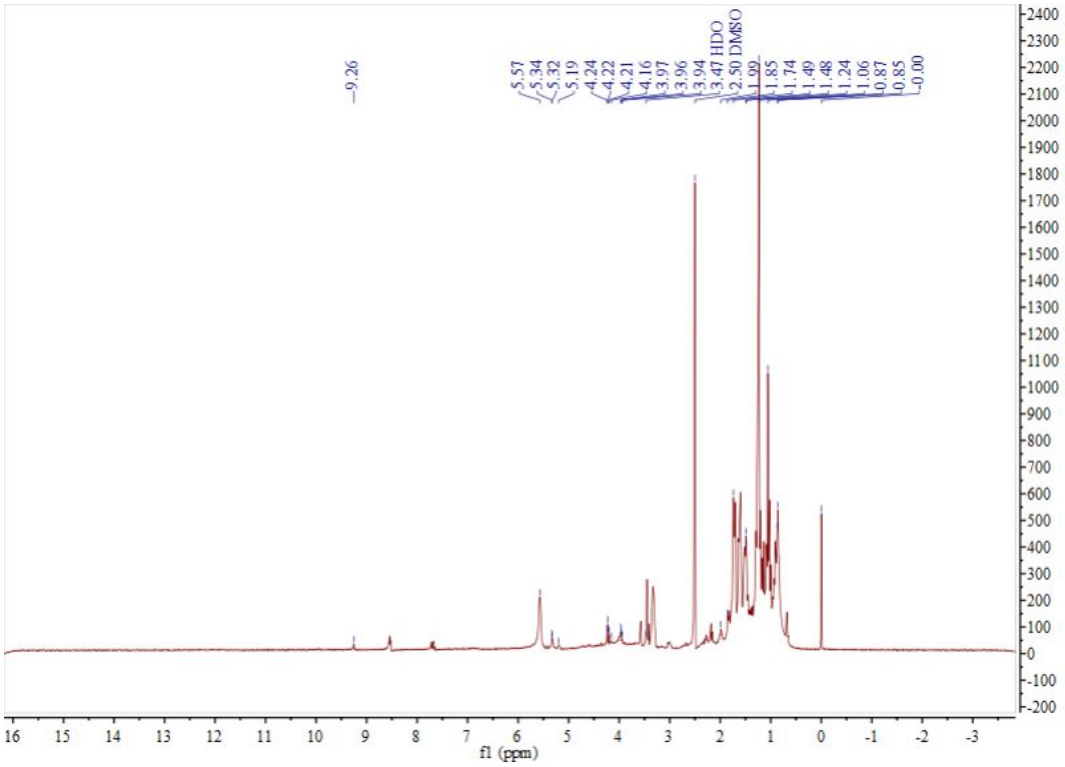


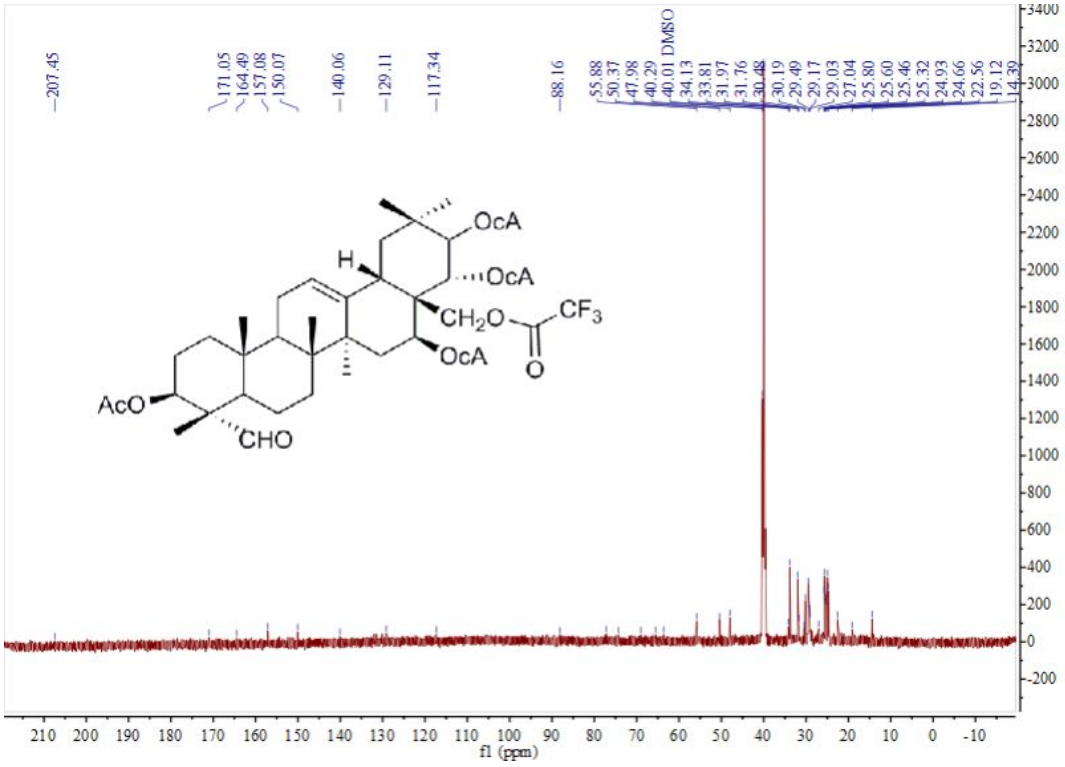
2) ***28-O-chloroacetyl-3β,16α,21β,22α-O-tetraacetyl-oleanone-12-Ene-23-aldehyde*** **(S-2)**



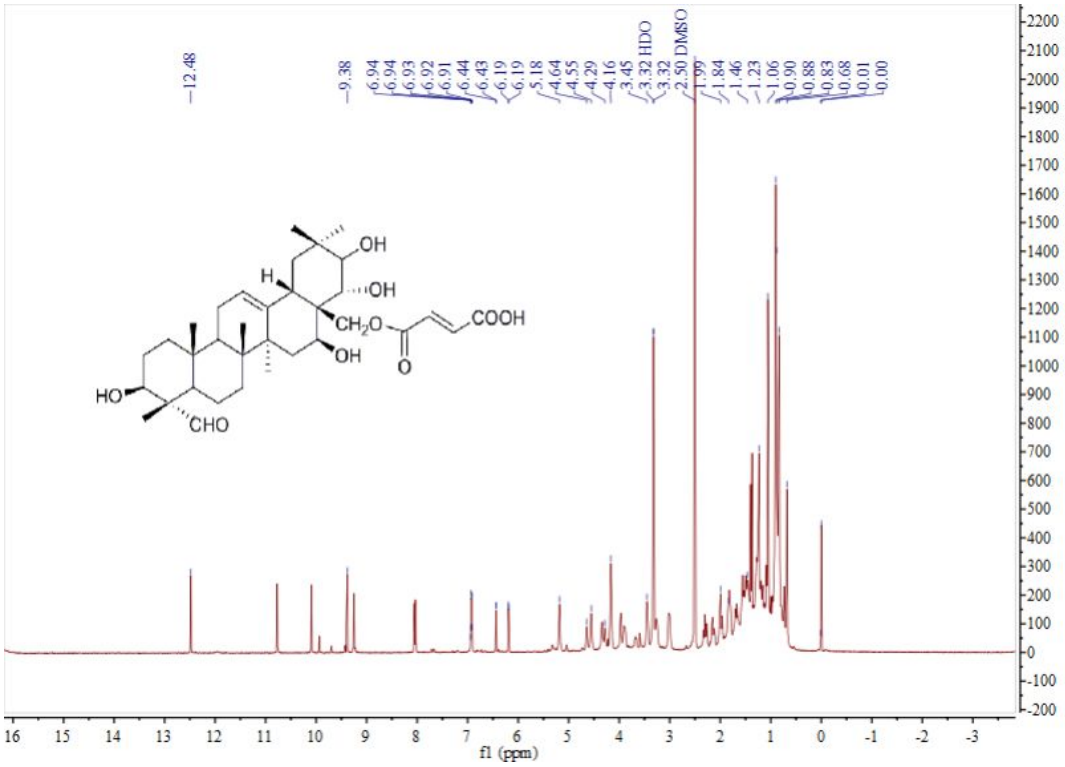


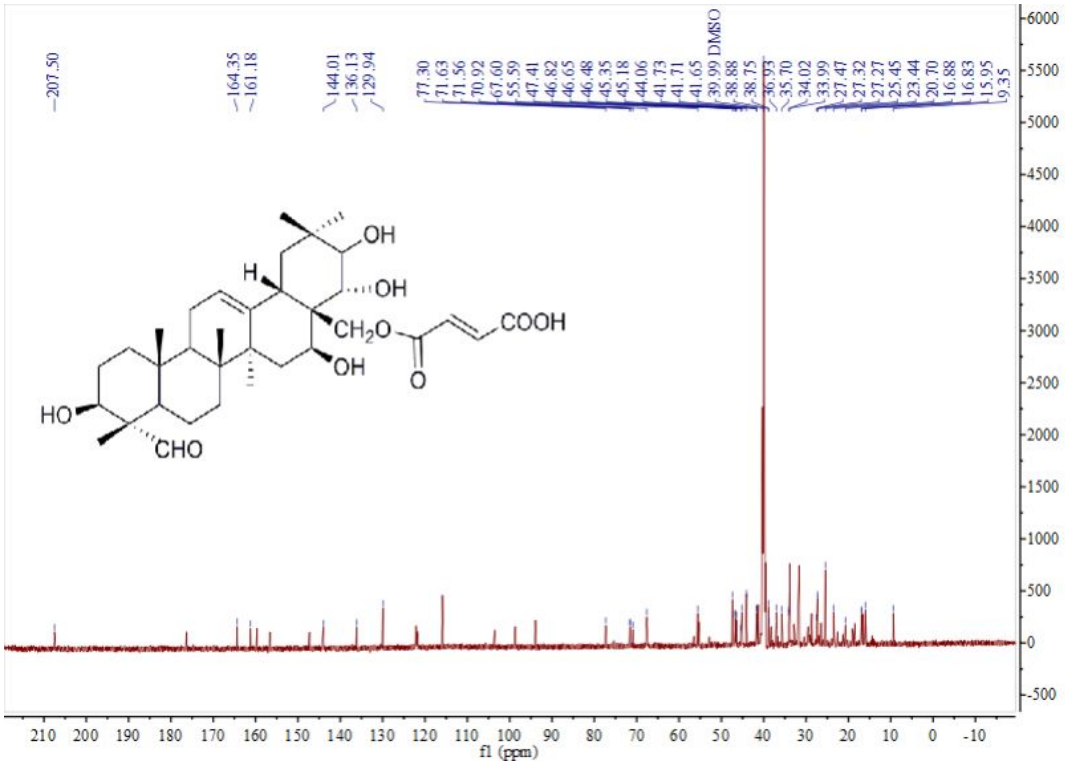
3) ***28-O-trifluoroacetyl-3β,16α,21β,22α-O-tetraacetyl-oleanone-12-Ene-23-aldehyde*** **(S-3)**



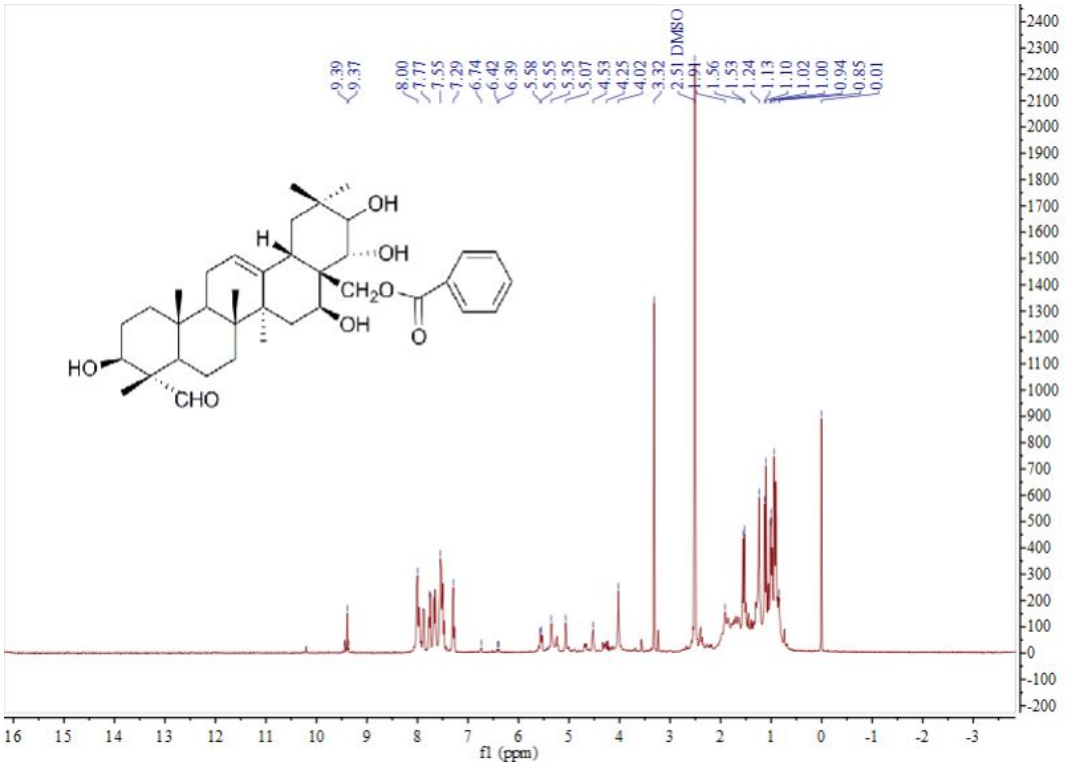


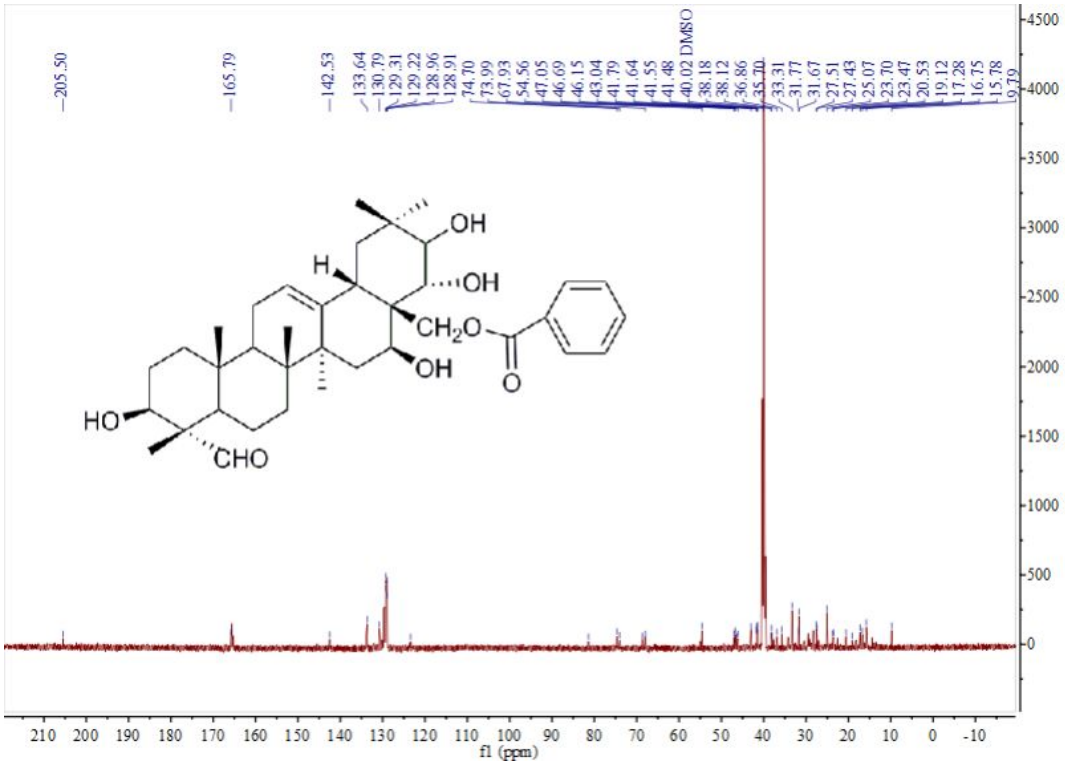
4) ***28-maleic acid monoester-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-4)**



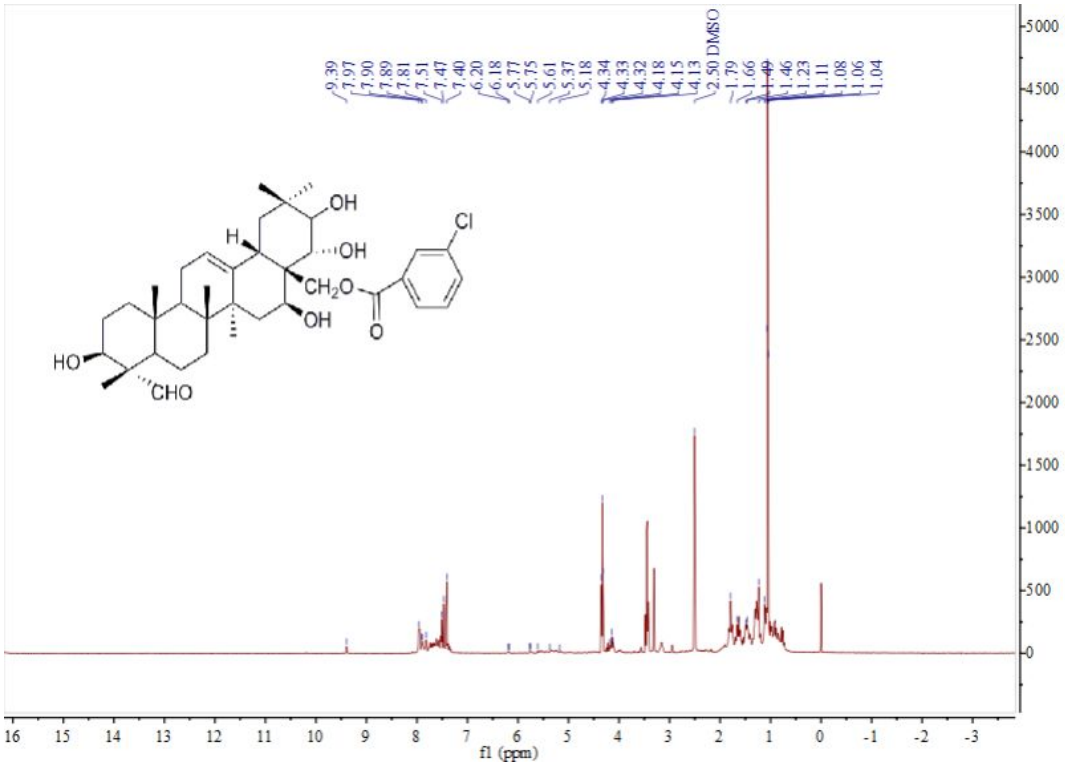


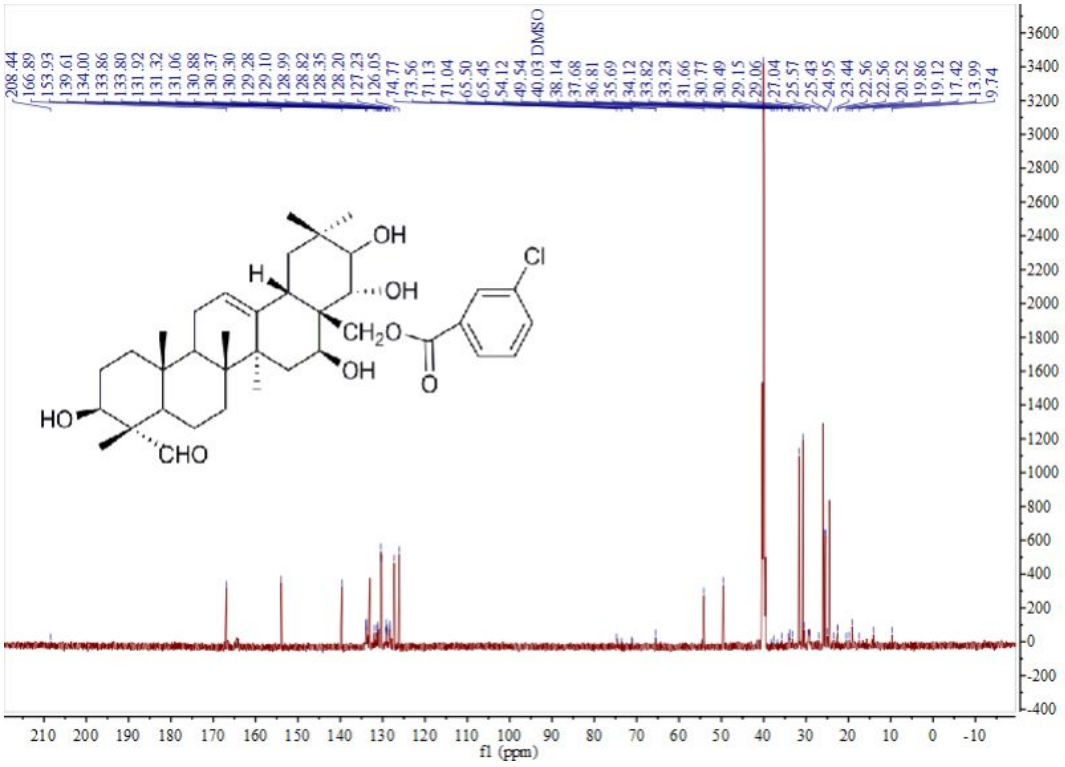
5) ***28-O-benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-5)**



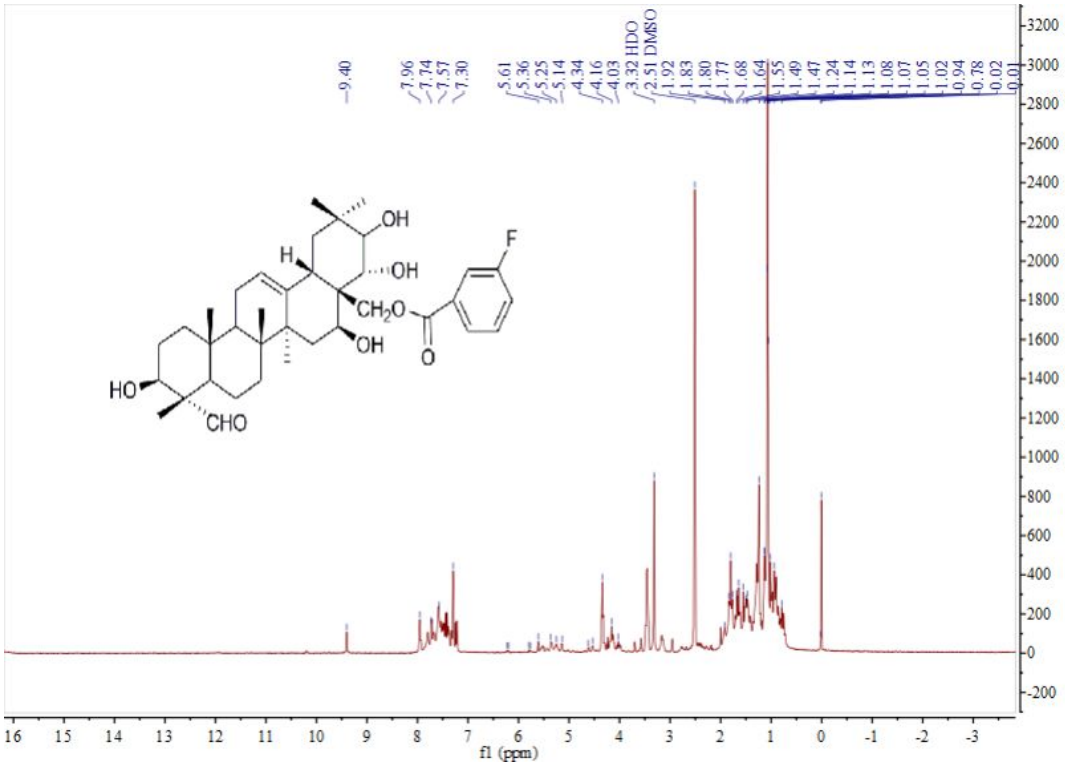


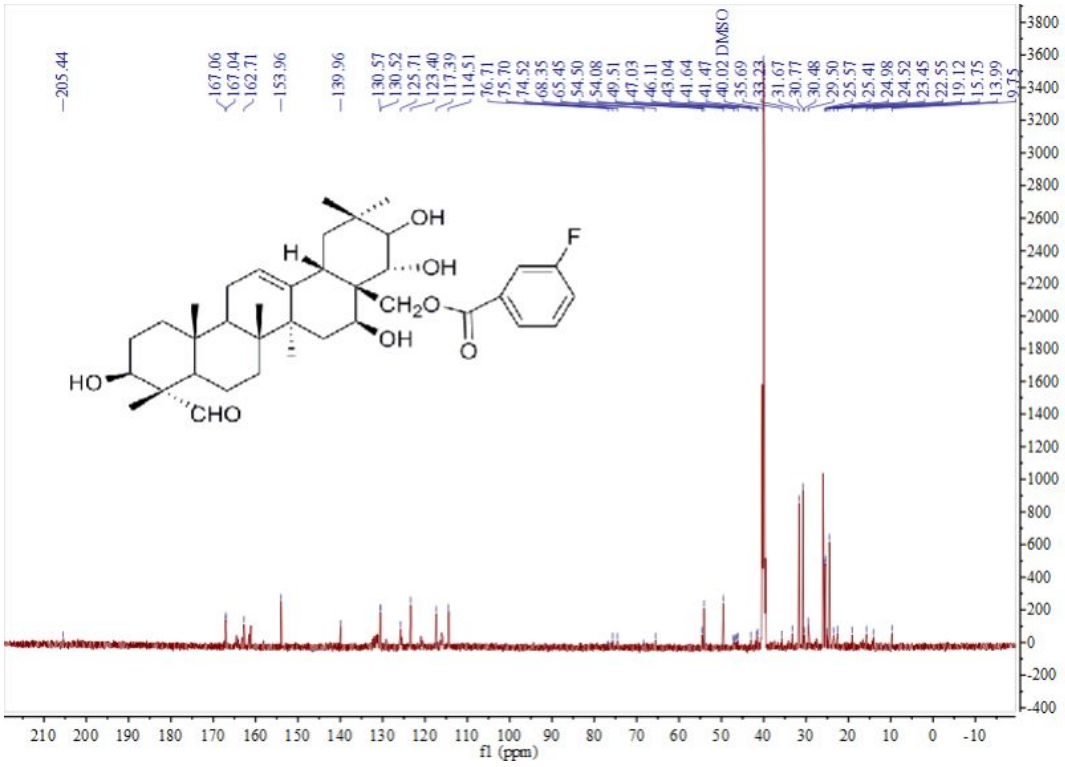
6) ***28-O-m-chloro benzoyl-3β, 16α, 21β, 22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-6)**



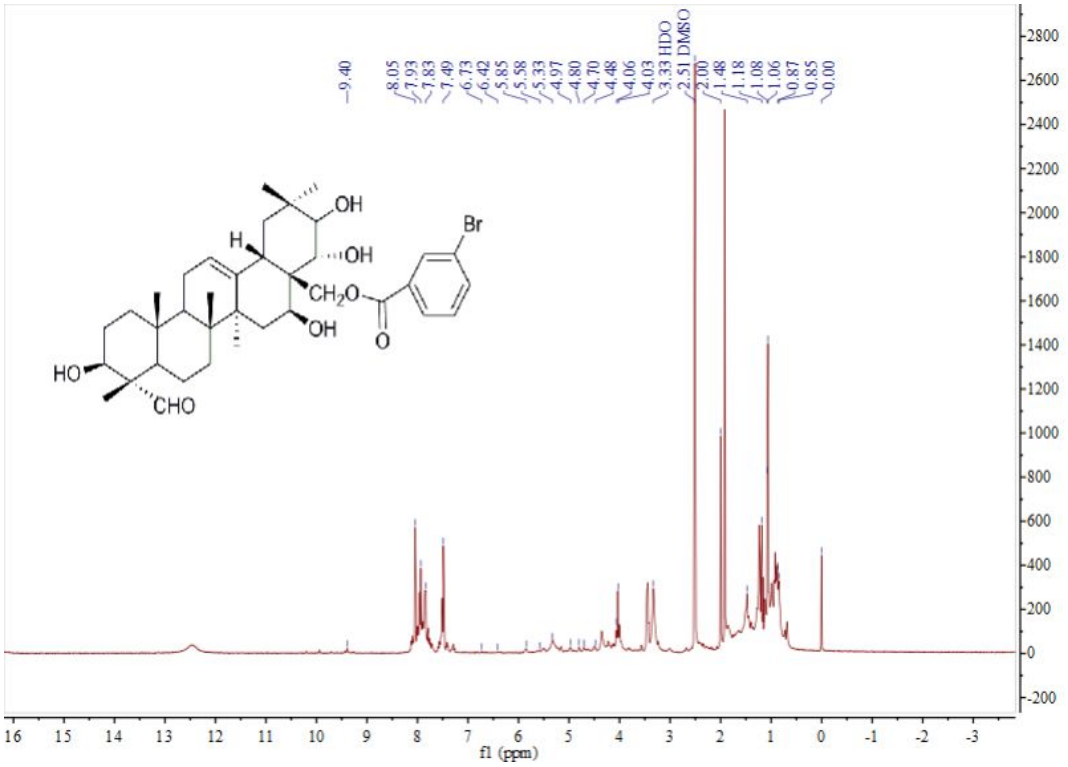


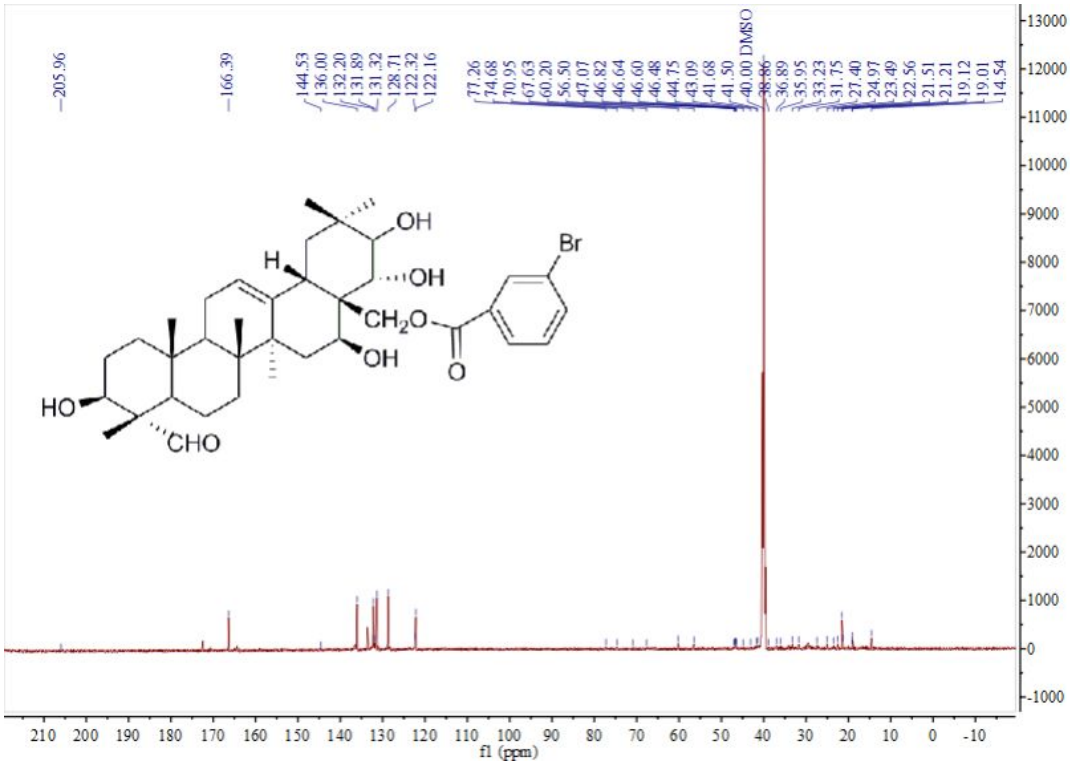
7) ***28-O-m-fluoro benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-7)**



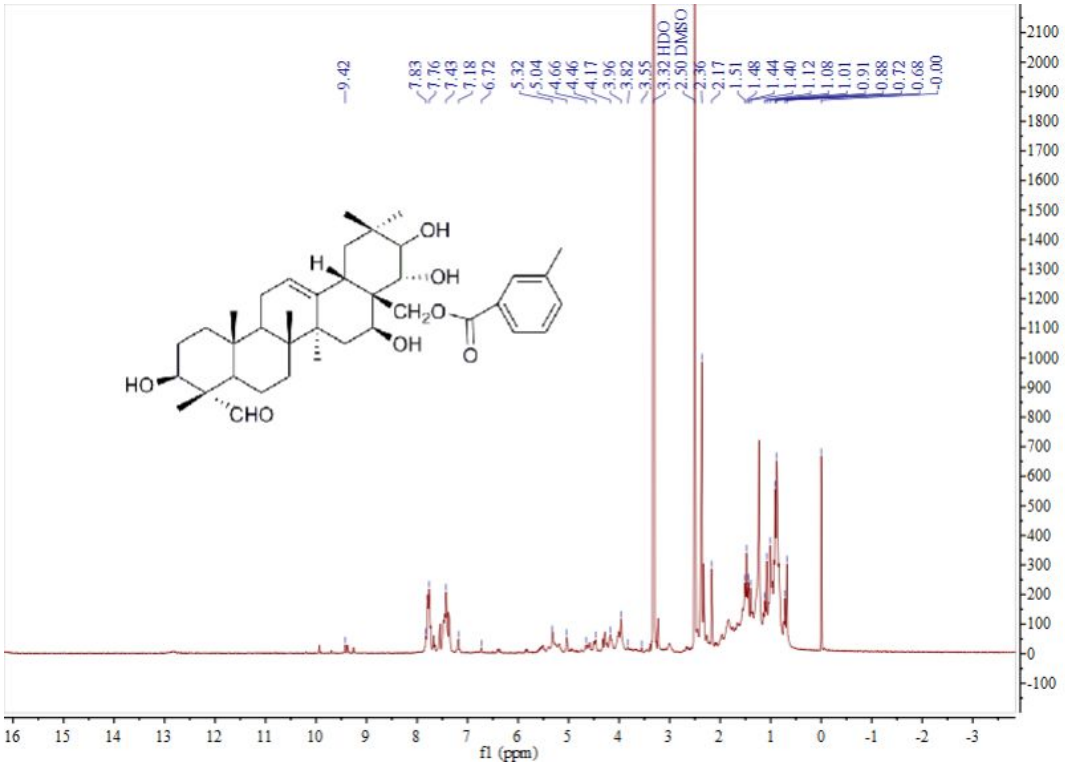


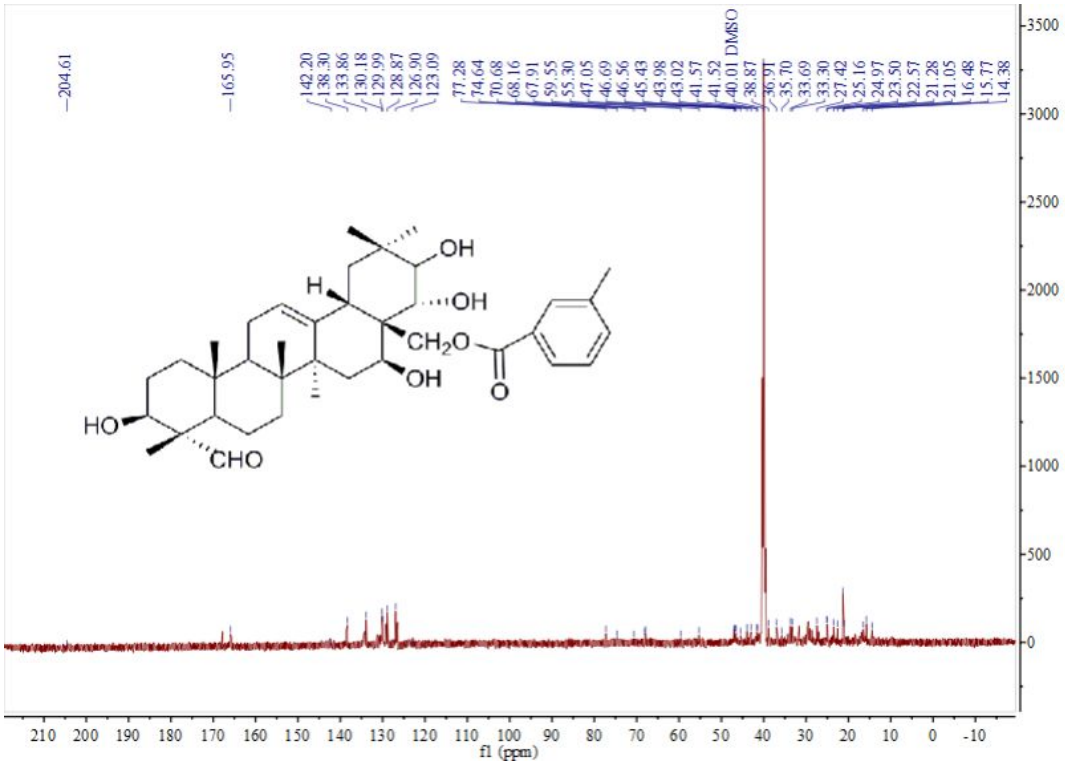
8) ***28-O-m-bromo benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-8)**



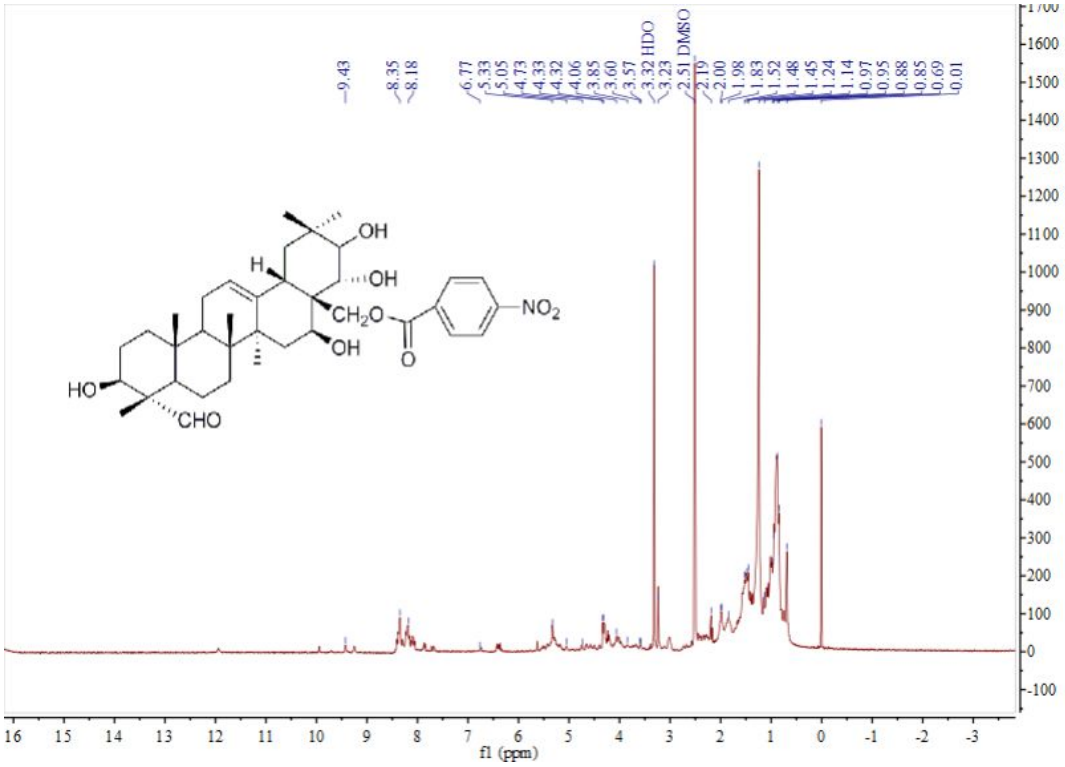


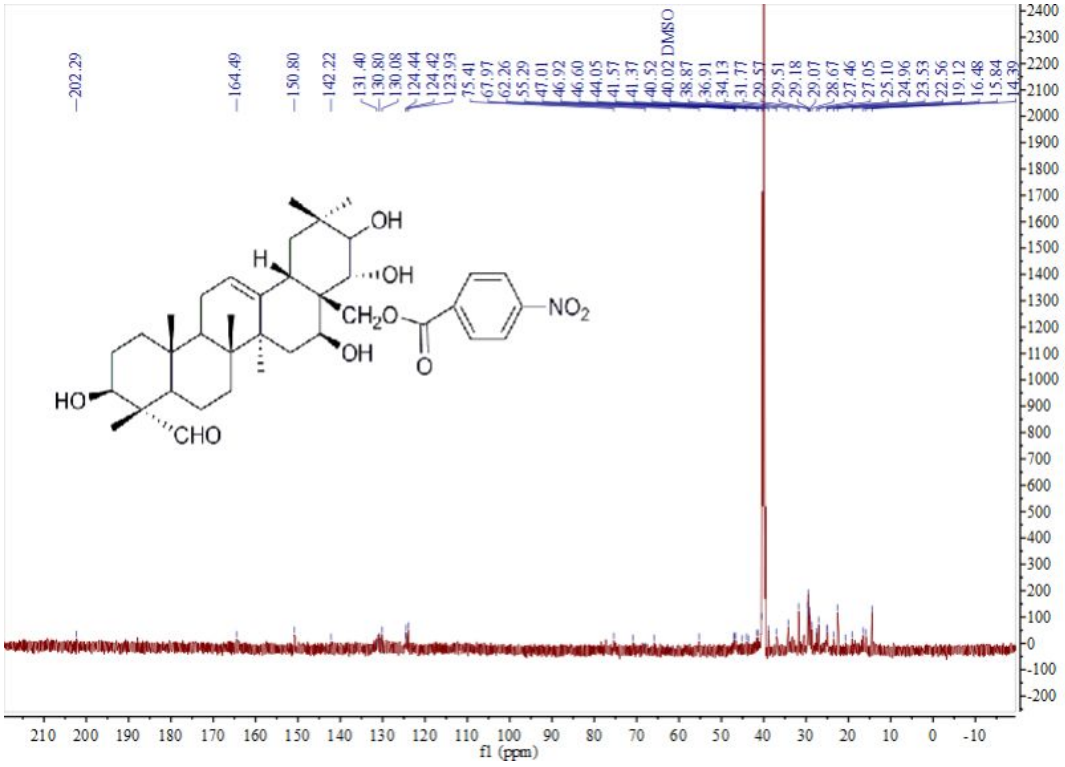
9) ***28-O-m-methyl benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-9)**



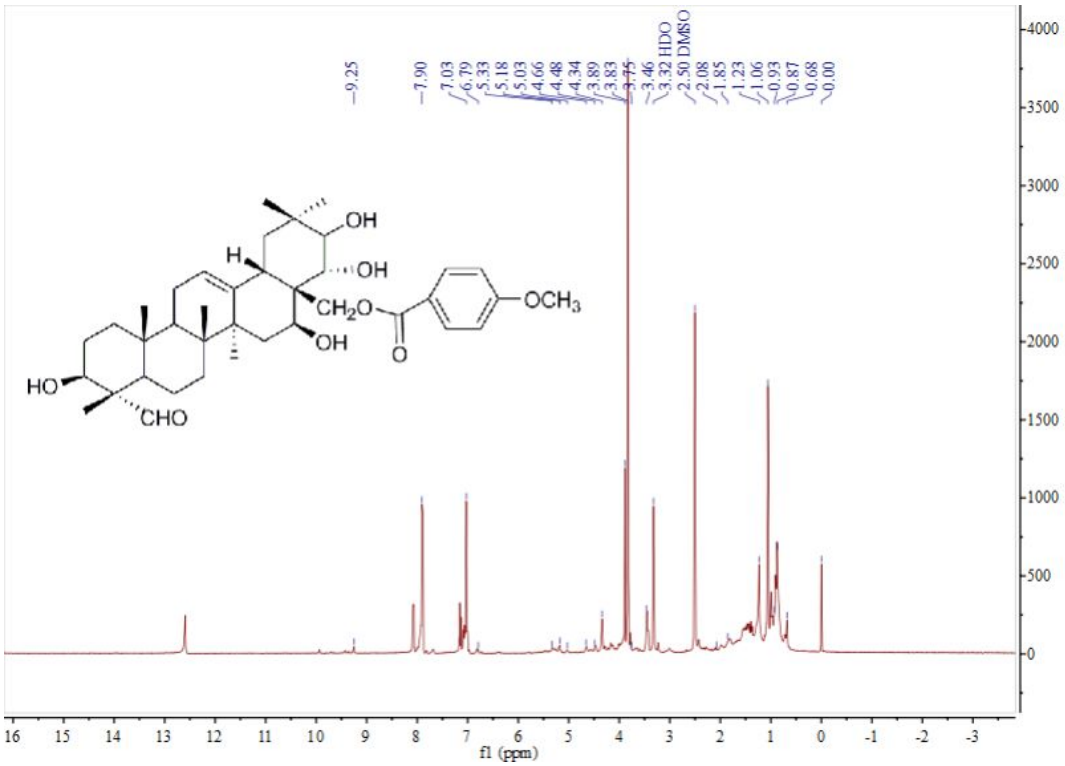


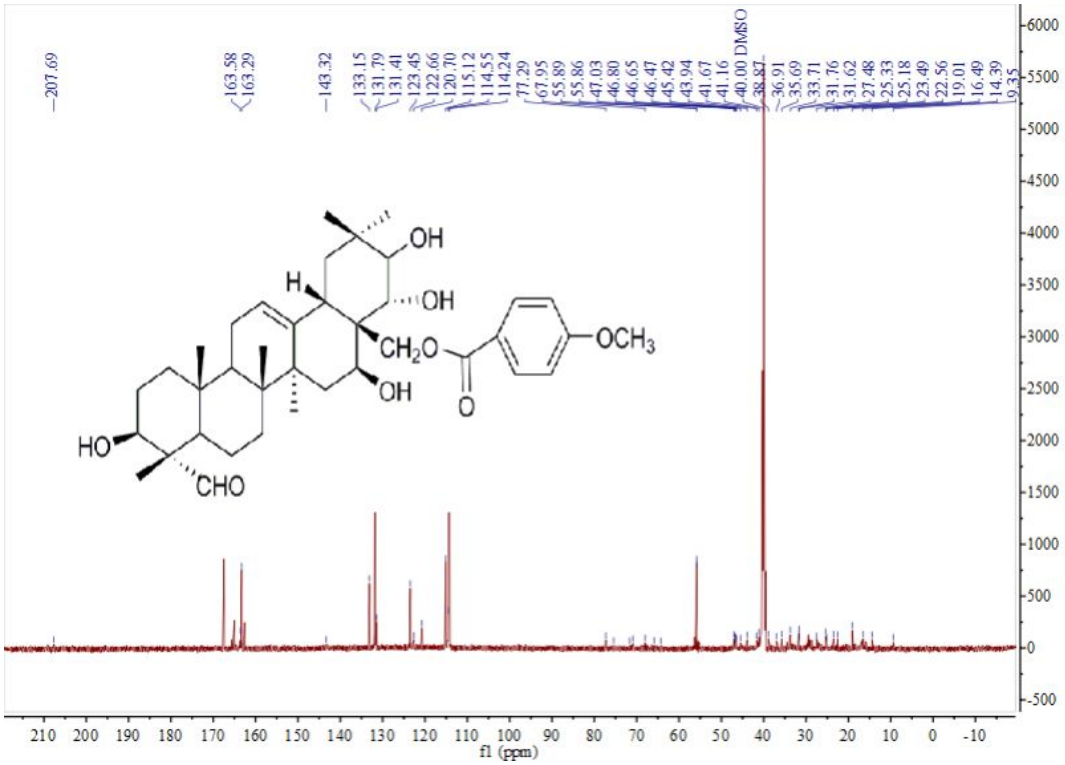
10) ***28-O-p-nitro benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-10)**



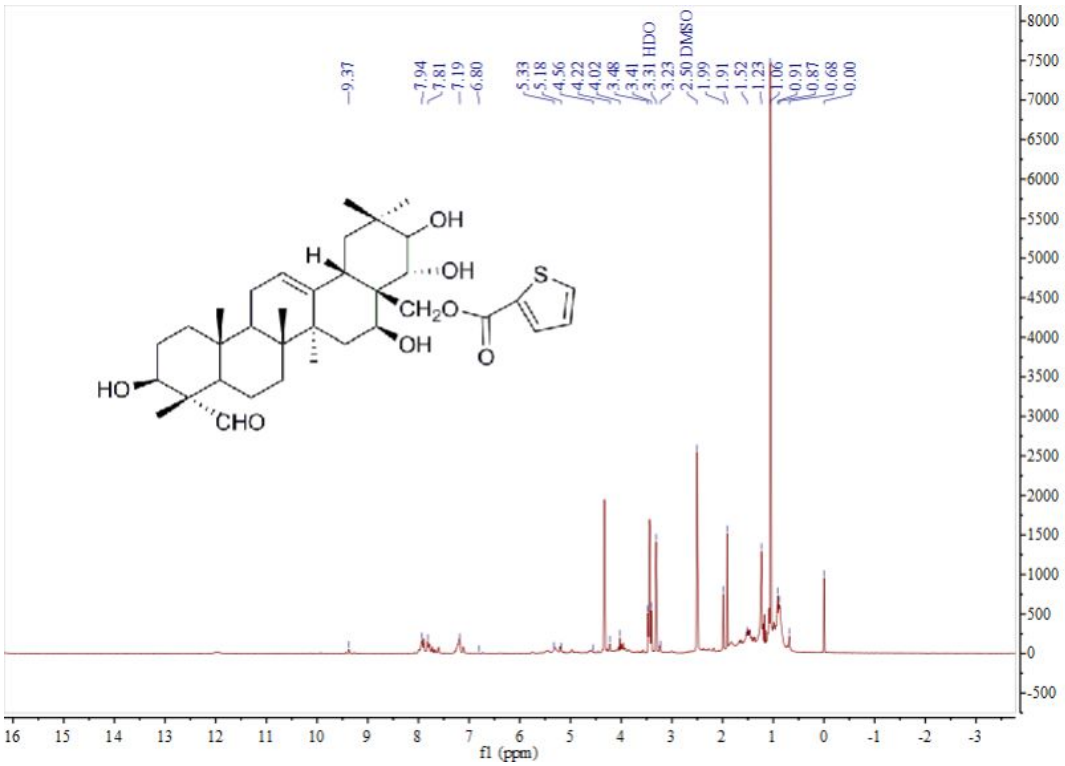


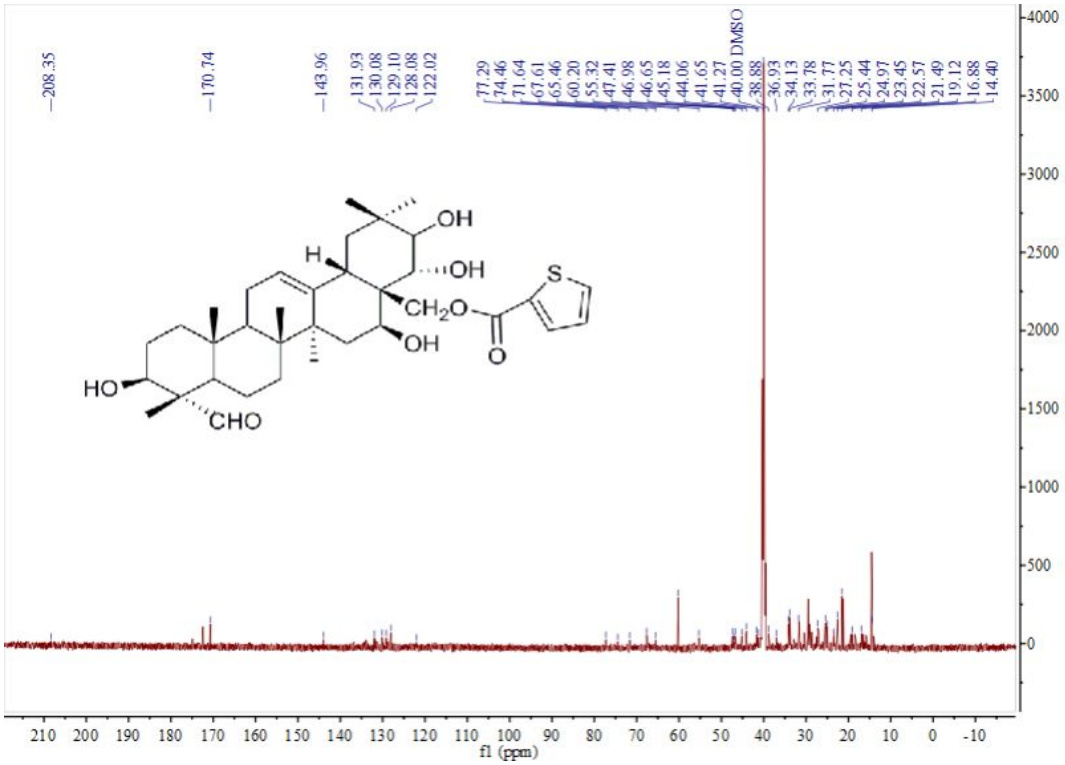
11) ***28-O-p-methoxy benzoyl-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde* (S-11)**



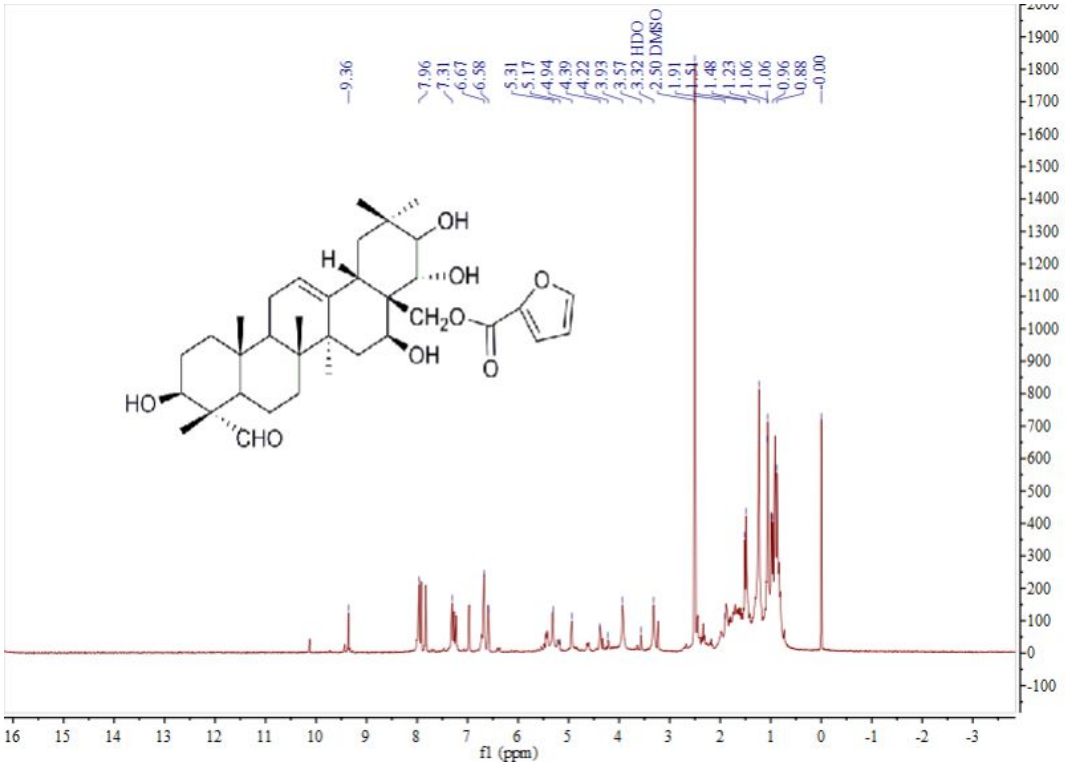


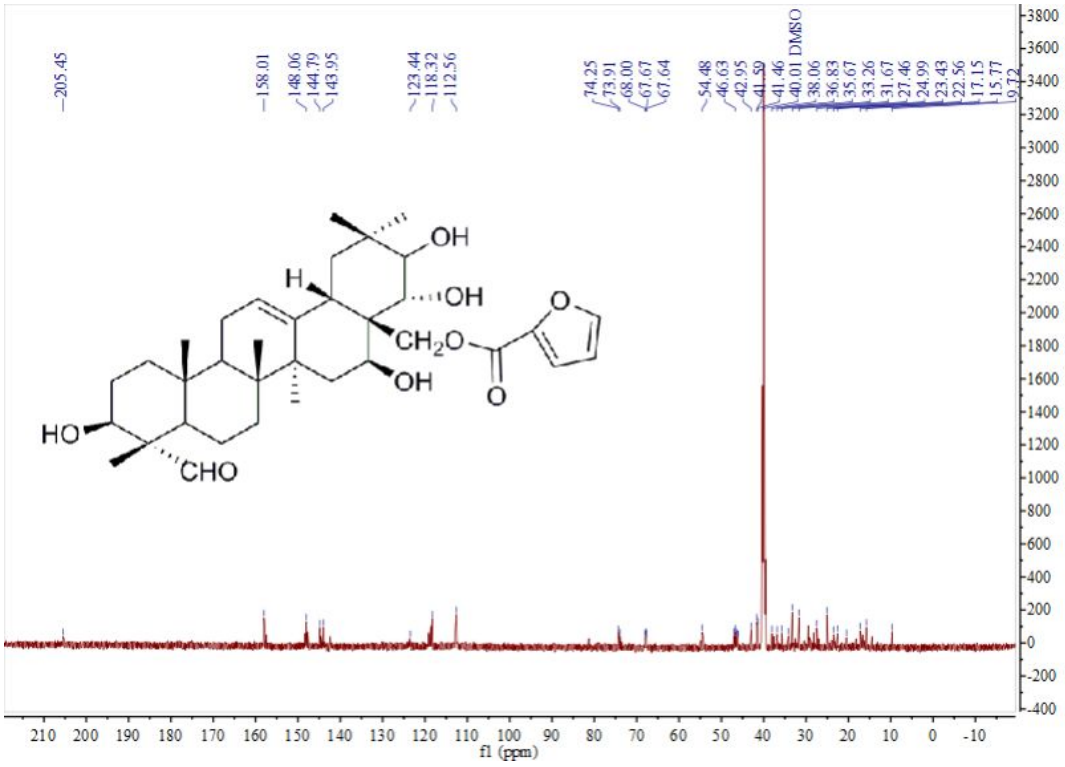
12) ***28-O-(2-thiophenoformyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde* (S-12)**



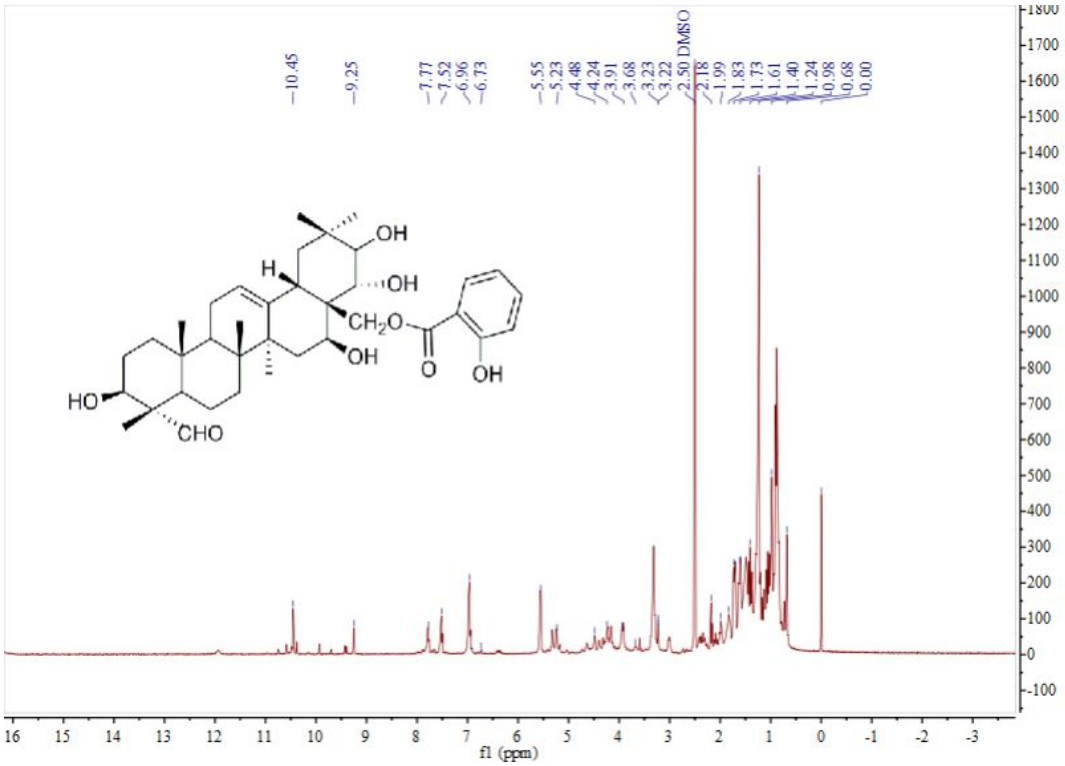


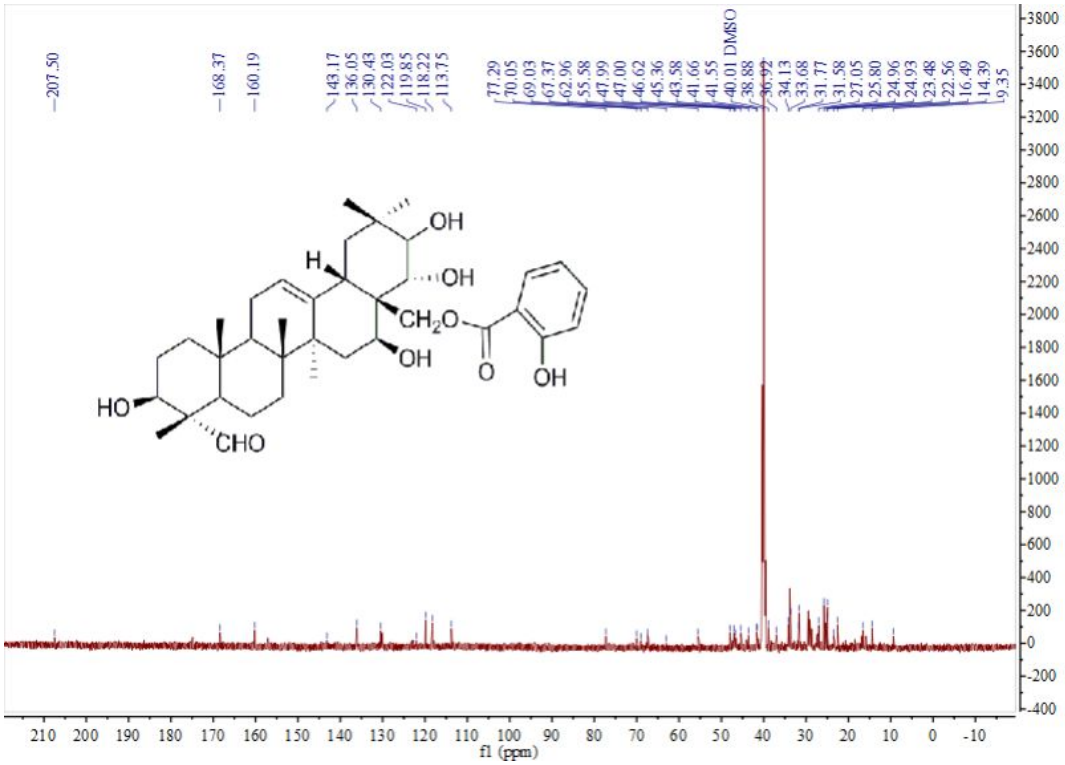
13) ***28-O-(2-furfuroyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-13)**



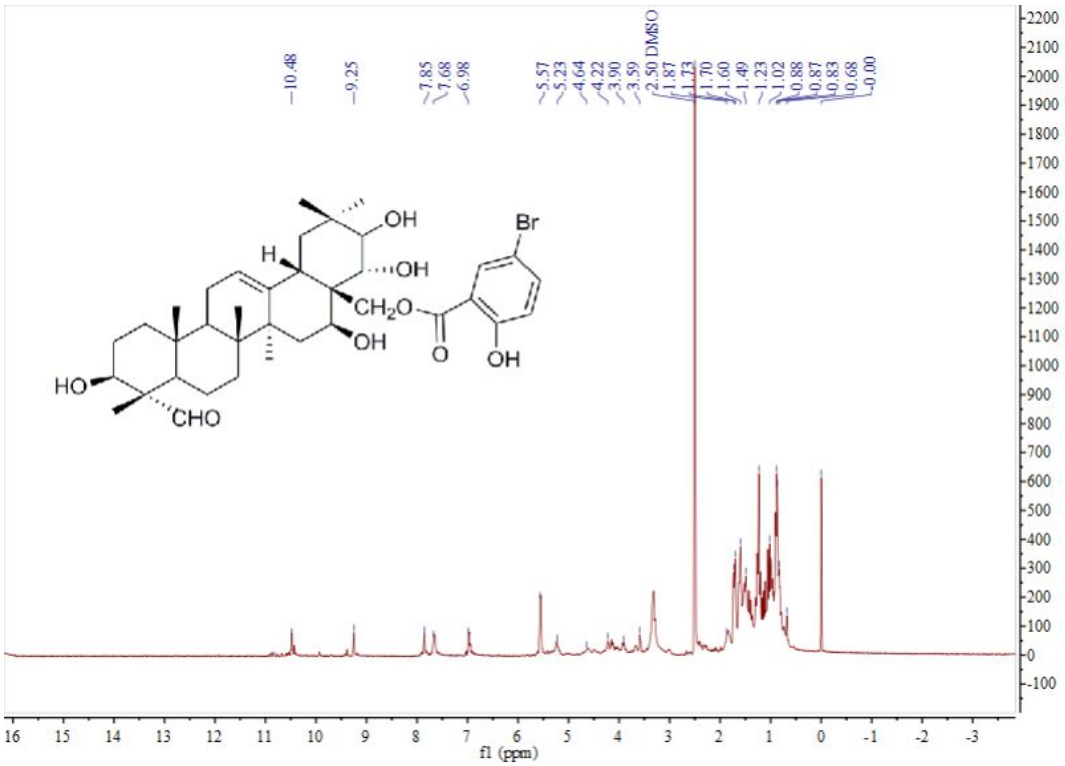


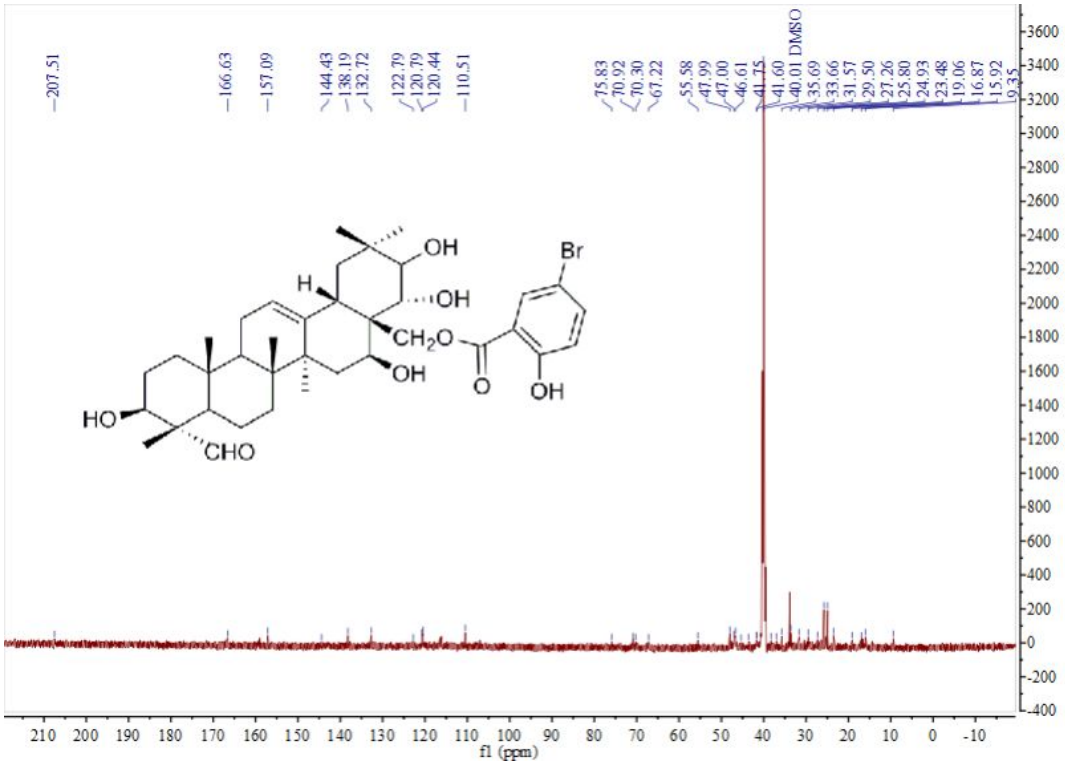
14) ***28-O-(2-hydroxybenzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-14)**



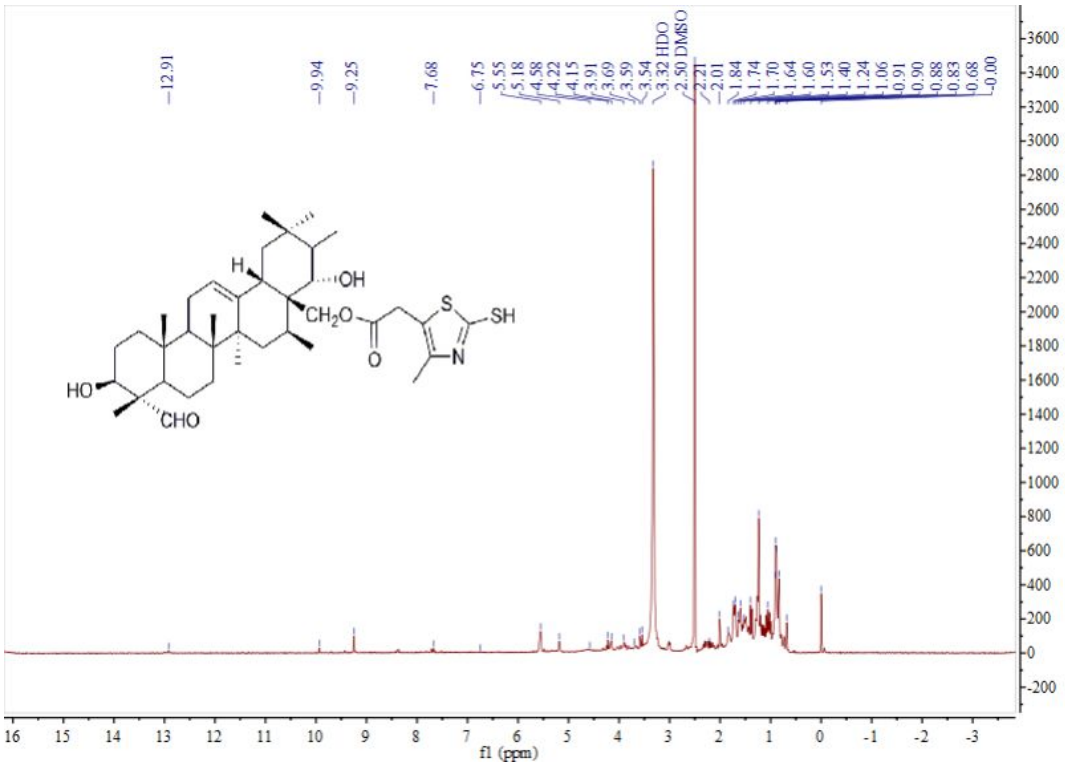


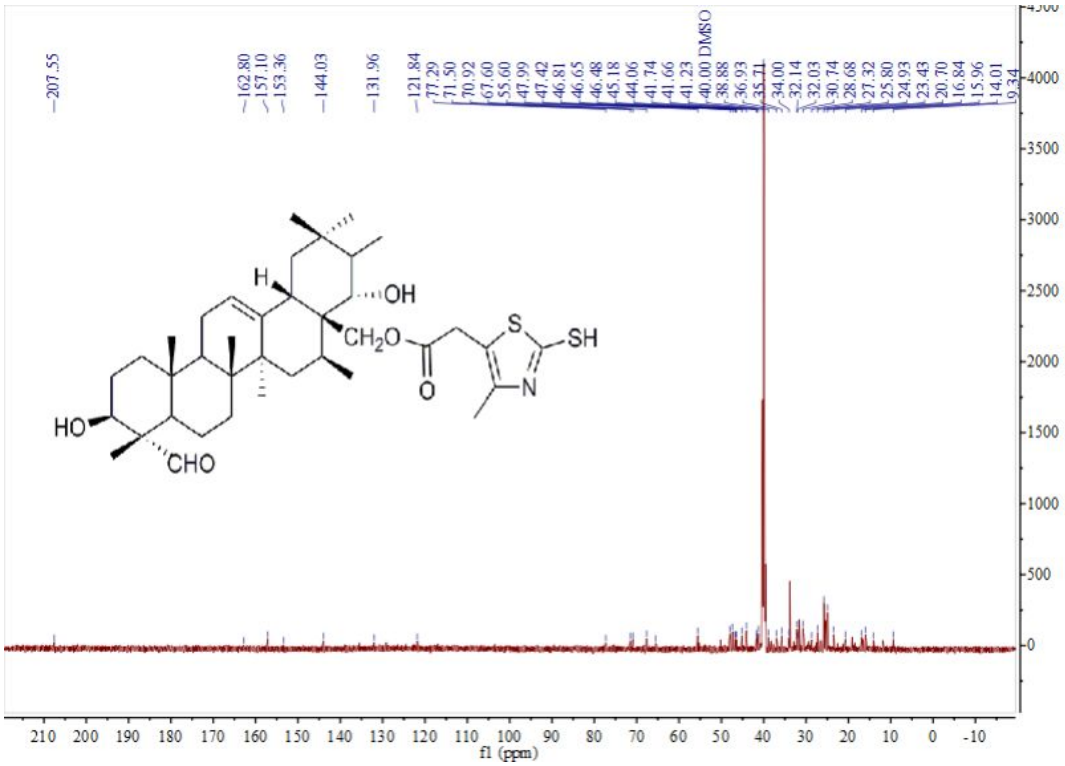
15) 28-O-(5-bromo-2-hydroxybenzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde **(S-15)**



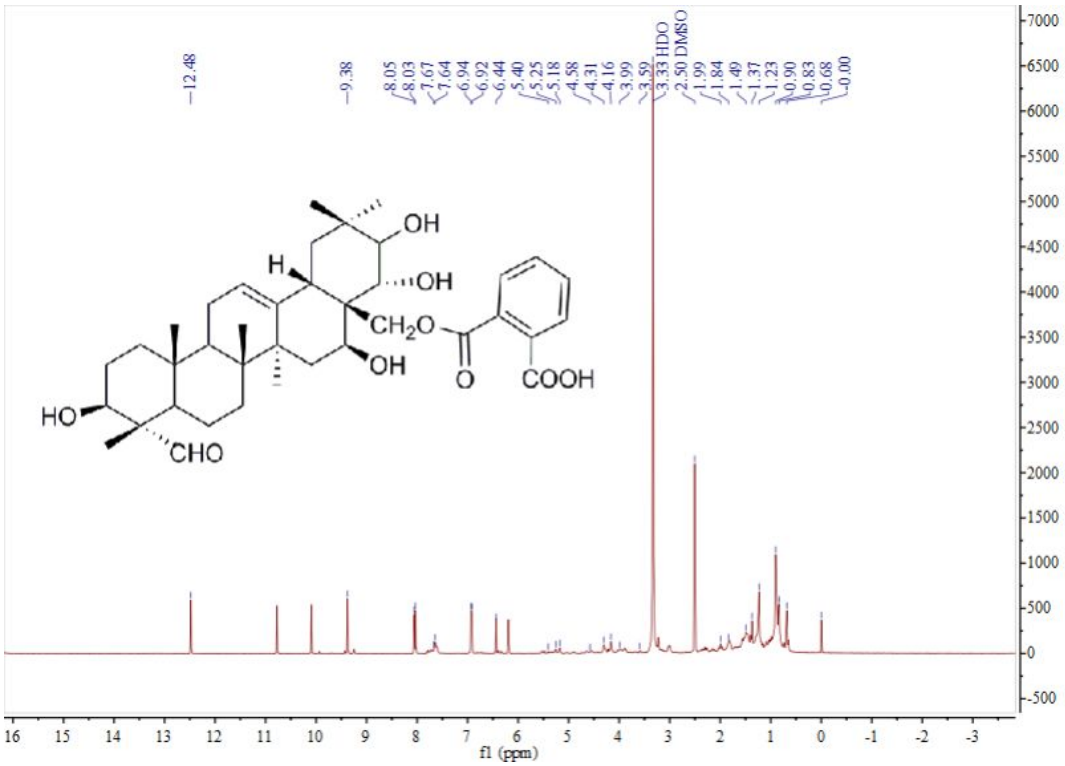


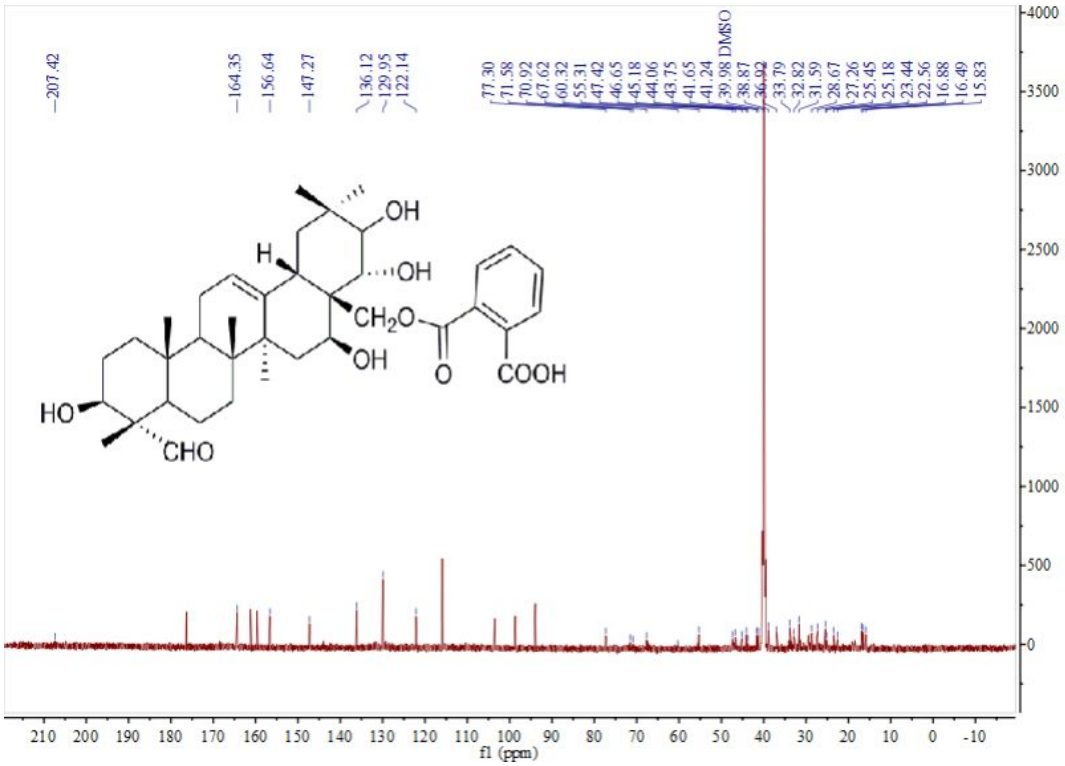
16) ***28-O-(2-mercapto-4-methyl-5-thiazolyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-16)**



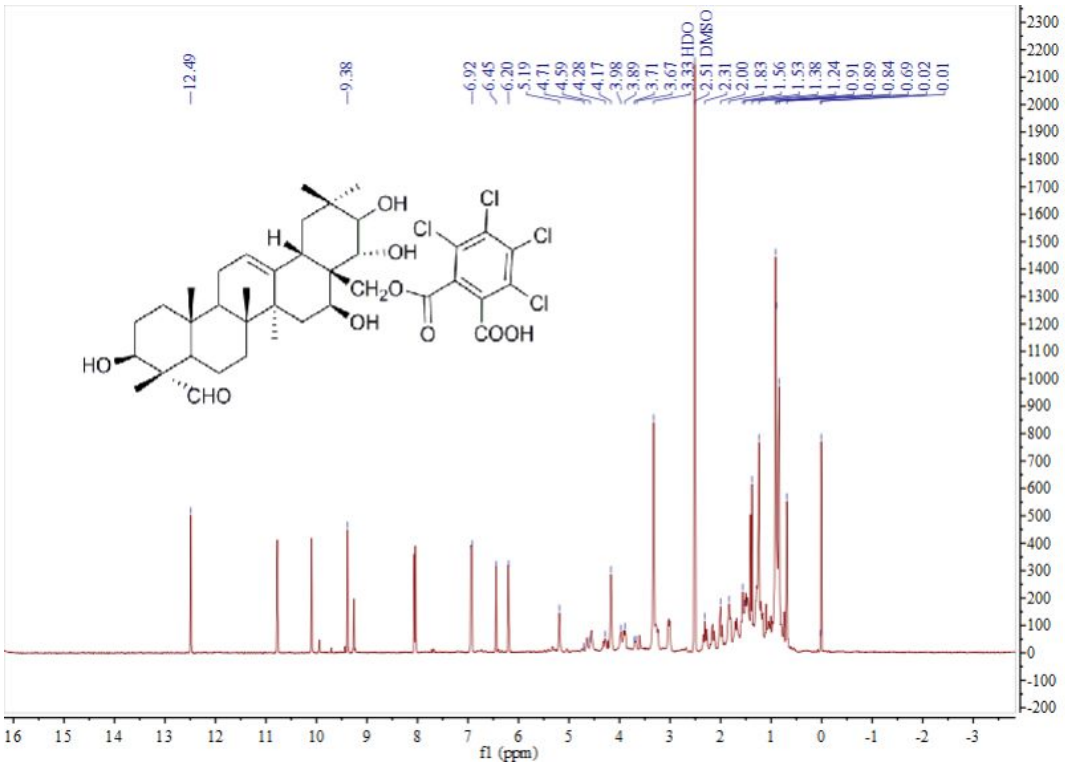


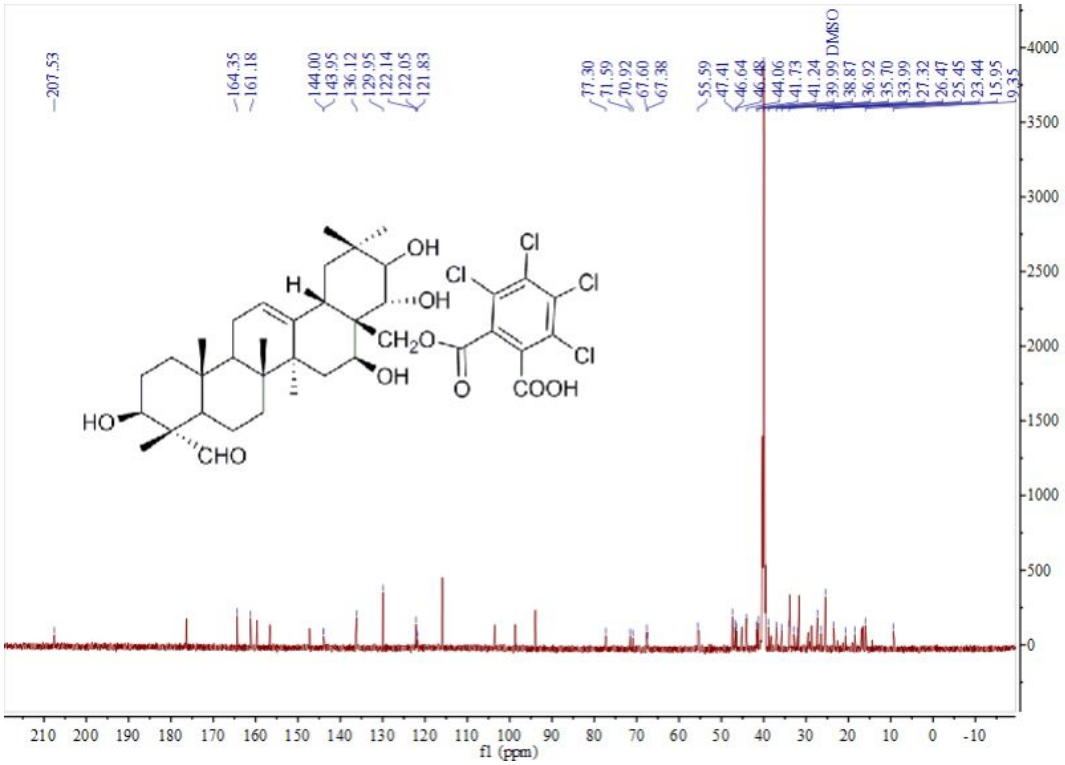
17) ***28-O-(2-formicacid-benzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-17)**





18) ***28-O-(3,4,5,6-tetrachloro-2-formicacid-benzoyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde*** **(S-18)**





19) **28-O-(4-acetylamino-benzenesulfonyl)-3β,16α,21β,22α-O-tetrahydroxy-oleantel-12-Ene-23-aldehyde (S-19)**