

# Novel Semisynthetic Derivatives of Bile Acids as Effective Tyrosyl-DNA Phosphodiesterase 1 Inhibitors

Oksana V. Salomatina<sup>1</sup>, Irina I. Popadyuk<sup>1</sup>, Alexandra L. Zakharenko<sup>2</sup>, Olga D. Zakharova<sup>2</sup>, Dmitriy S. Fadeev<sup>1</sup>, Nina I. Komarova<sup>1</sup>, Jóhannes Reynisson<sup>3</sup>, H. John Arabshahi<sup>3</sup>, Raina Chand<sup>3</sup>, Konstantin P. Volcho<sup>1,4</sup>, Nariman F. Salakhutdinov<sup>1,4</sup>, Olga I. Lavrik<sup>2,4</sup>

<sup>1</sup> N.N. Vorozhtsov Novosibirsk Institute of Organic Chemistry, SB RAS, 630090, Russia, Novosibirsk, acad. Lavrentjev ave. 9; [ana@nioch.nsc.ru](mailto:ana@nioch.nsc.ru) (O.V.S.), [popadyuk@nioch.nsc.ru](mailto:popadyuk@nioch.nsc.ru) (I.I.P.), [dsf@nioch.nsc.ru](mailto:dsf@nioch.nsc.ru) (D.S.F.), [komar@nioch.nsc.ru](mailto:komar@nioch.nsc.ru) (N.I.K)

<sup>2</sup> Novosibirsk Institute of Chemical Biology and Fundamental Medicine, SB RAS, 630090, Russia, Novosibirsk, acad. Lavrentjev ave. 8; [sashaz@niboch.nsc.ru](mailto:sashaz@niboch.nsc.ru) (A.L.Z), [isar@niboch.nsc.ru](mailto:isar@niboch.nsc.ru) (O.D.Z.)

<sup>3</sup> School of Chemical Sciences, University of Auckland, New Zealand. [j.reynisson@auckland.ac.nz](mailto:j.reynisson@auckland.ac.nz) (J.R.), [j.arabshahi@auckland.ac.nz](mailto:j.arabshahi@auckland.ac.nz) (H.J.A.), [rcha387@aucklanduni.ac.nz](mailto:rcha387@aucklanduni.ac.nz) (R.Ch.)

<sup>4</sup> Novosibirsk State University, 630090, Russia, Novosibirsk, Pirogova str. 2; [anvar@nioch.nsc.ru](mailto:anvar@nioch.nsc.ru) (N.F.S.) [lavrik@niboch.nsc.ru](mailto:lavrik@niboch.nsc.ru) (O.I.L.)

\* Correspondence: [volcho@nioch.nsc.ru](mailto:volcho@nioch.nsc.ru); Tel.: +7 383 3308870

## ELECTRONIC SUPPORTING INFORMATION

|                                                                      |   |
|----------------------------------------------------------------------|---|
| Virtual screening & molecular modeling data.....                     | 2 |
| NMR <sup>1</sup> H and <sup>13</sup> C spectra of new compounds..... | 5 |

Table S1. The scores predicted by the four scoring functions for compounds **1a,b**, **2a,b**, **3a-d**, **4a,b-8a,b** and **9a-11a** docked in Tdp1.

| Compound   | Scoring Function |      |      |      | IC <sub>50</sub> |
|------------|------------------|------|------|------|------------------|
|            | PLP              | GS   | CS   | ASP  |                  |
| <b>1a</b>  | 71.8             | 56.5 | 32.7 | 36.0 | 0.32±0.11        |
| <b>2a</b>  | 70.4             | 58.6 | 26.7 | 31.6 | 0.38±0.12        |
| <b>3a</b>  | 80.8             | 60.8 | 33.9 | 35.8 | 0.65±0.16        |
| <b>3c</b>  | 71.9             | 49.7 | 22.9 | 35.4 | 0.95±0.05        |
| <b>Fur</b> | 45.3             | 47.6 | 27.1 | 33.3 | 1.23±0.33        |
| <b>1b</b>  | 75.2             | 44.6 | 36.3 | 33.6 | 2.65±0.30        |
| <b>2b</b>  | 74.8             | 59.5 | 37.4 | 35.7 | 2.6±0.4          |
| <b>3b</b>  | 80.0             | 63.5 | 33.5 | 34.1 | 2.7±0.2          |
| <b>3d</b>  | 72.7             | 67.9 | 34.1 | 37.4 | 0.48±0.04        |
| <b>4a</b>  | 41.5             | 28.2 | 26.7 | 28.5 | 0.43±0.13        |
| <b>4b</b>  | 65.9             | 41.9 | 31.3 | 28.2 | 6.7±0.7          |
| <b>5a</b>  | 62.1             | 53.9 | 31.2 | 27.3 | 0.42±0.01        |
| <b>5b</b>  | 66.4             | 49.7 | 31.5 | 27.5 | 1.3±0.2          |
| <b>6a</b>  | 64.7             | 40.3 | 29.5 | 27.0 | 1.00±0.05        |
| <b>6b</b>  | 66.5             | 45.8 | 32.4 | 28.2 | 7.6±3.9          |
| <b>7a</b>  | 63.3             | 59.4 | 29.8 | 30.0 | 4.08±0.08        |
| <b>7b</b>  | 63.4             | 50.1 | 29.6 | 32.5 | >15              |
| <b>8a</b>  | 64.5             | 60.2 | 30.8 | 26.8 | 0.47±0.08        |
| <b>8b</b>  | 63.5             | 47.3 | 33.0 | 27.2 | 2.3±0.4          |
| <b>9a</b>  | 69.4             | 59.1 | 31.5 | 35.0 | 0.29±0.12        |
| <b>10a</b> | 56.9             | 56.2 | 25.6 | 26.8 | >15              |
| <b>11a</b> | 56.5             | 61.9 | 22.1 | 28.6 | >15              |

**Table S2.** The calculated molecular descriptors for the bile acid derivatives.

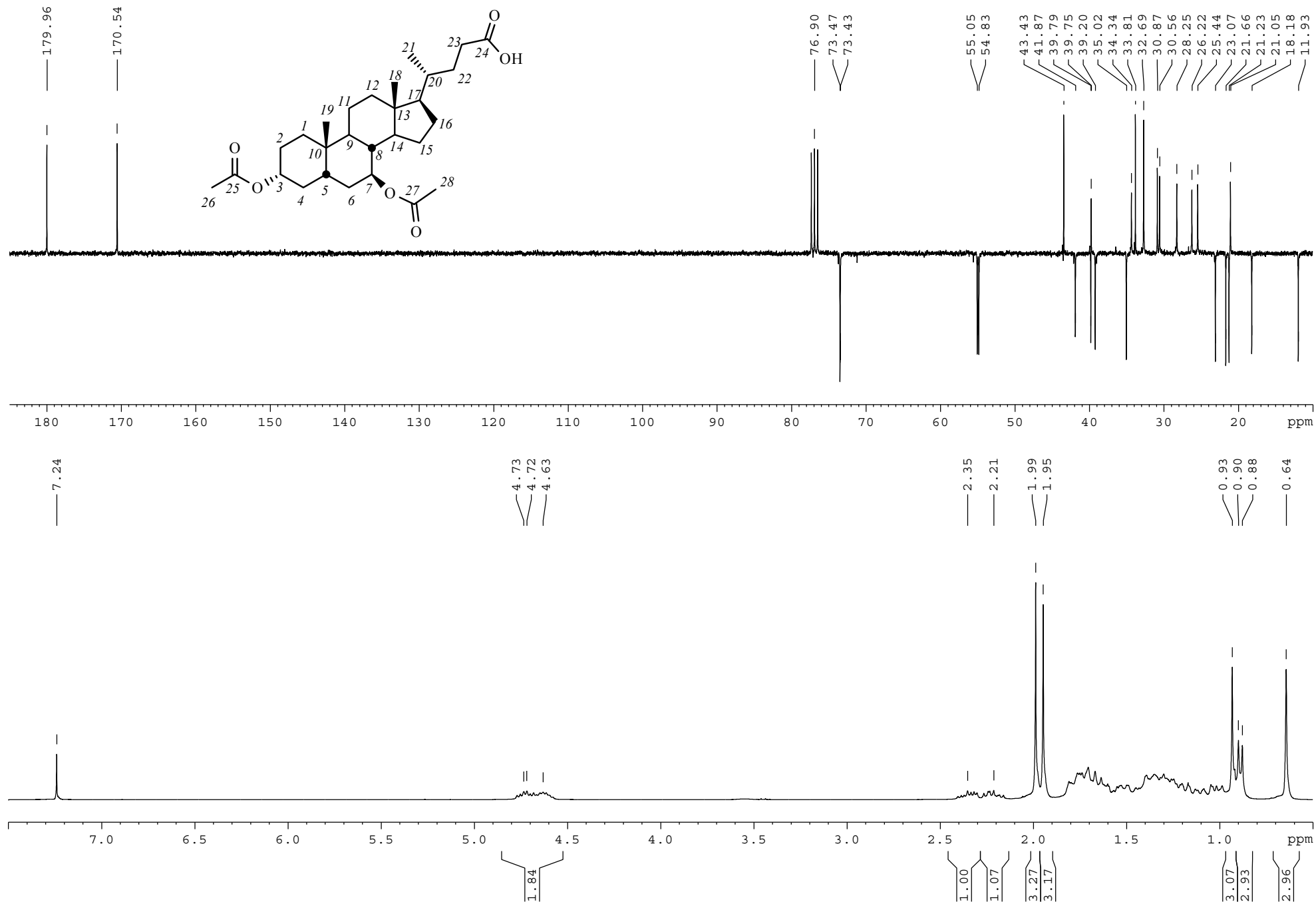
| Compound   | Molecular Descriptor |       |       |              |                 |                |
|------------|----------------------|-------|-------|--------------|-----------------|----------------|
|            | Mol. weight          | Log P | TPSA  | H bond donor | H bond acceptor | Rotatable bond |
| <b>1a</b>  | 618.9                | 5.5   | 97.5  | 2            | 3               | 11             |
| <b>2a</b>  | 618.9                | 5.5   | 97.5  | 2            | 3               | 11             |
| <b>3a</b>  | 618.9                | 5.3   | 97.5  | 2            | 3               | 11             |
| <b>3c</b>  | 576.8                | 5.2   | 91.4  | 3            | 3               | 9              |
| <b>Fur</b> | 304.4                | 2.9   | 112.9 | 4            | 4               | 4              |
| <b>1b</b>  | 534.8                | 5.0   | 85.4  | 4            | 3               | 7              |
| <b>2b</b>  | 534.8                | 5.0   | 85.4  | 4            | 3               | 7              |
| <b>3b</b>  | 534.8                | 5.0   | 85.4  | 4            | 3               | 7              |
| <b>3d</b>  | 576.8                | 5.2   | 91.4  | 3            | 3               | 9              |
| <b>4a</b>  | 551.8                | 5.7   | 81.7  | 1            | 3               | 9              |
| <b>4b</b>  | 467.7                | 5.4   | 69.6  | 3            | 3               | 5              |
| <b>5a</b>  | 630.6                | 6.5   | 81.7  | 1            | 3               | 9              |
| <b>5b</b>  | 546.6                | 6.2   | 69.5  | 3            | 3               | 5              |
| <b>6a</b>  | 565.8                | 6.1   | 81.7  | 1            | 3               | 9              |
| <b>6b</b>  | 481.7                | 5.9   | 69.6  | 3            | 3               | 5              |
| <b>7a</b>  | 552.8                | 4.4   | 94.6  | 1            | 4               | 9              |
| <b>7b</b>  | 468.7                | 4.1   | 82.5  | 3            | 4               | 5              |
| <b>8a</b>  | 609.9                | 5.7   | 81.7  | 1            | 3               | 9              |
| <b>8b</b>  | 525.8                | 5.4   | 69.6  | 3            | 3               | 5              |
| <b>9a</b>  | 708.0                | 9.0   | 101.9 | 2            | 4               | 13             |
| <b>10a</b> | 588.9                | 4.7   | 84.9  | 1            | 4               | 14             |
| <b>11a</b> | 589.8                | 4.0   | 94.2  | 1            | 5               | 12             |

**Table S3.** Criteria of lead-like, drug-like and known drug space (KDS) in terms of molecular descriptors.

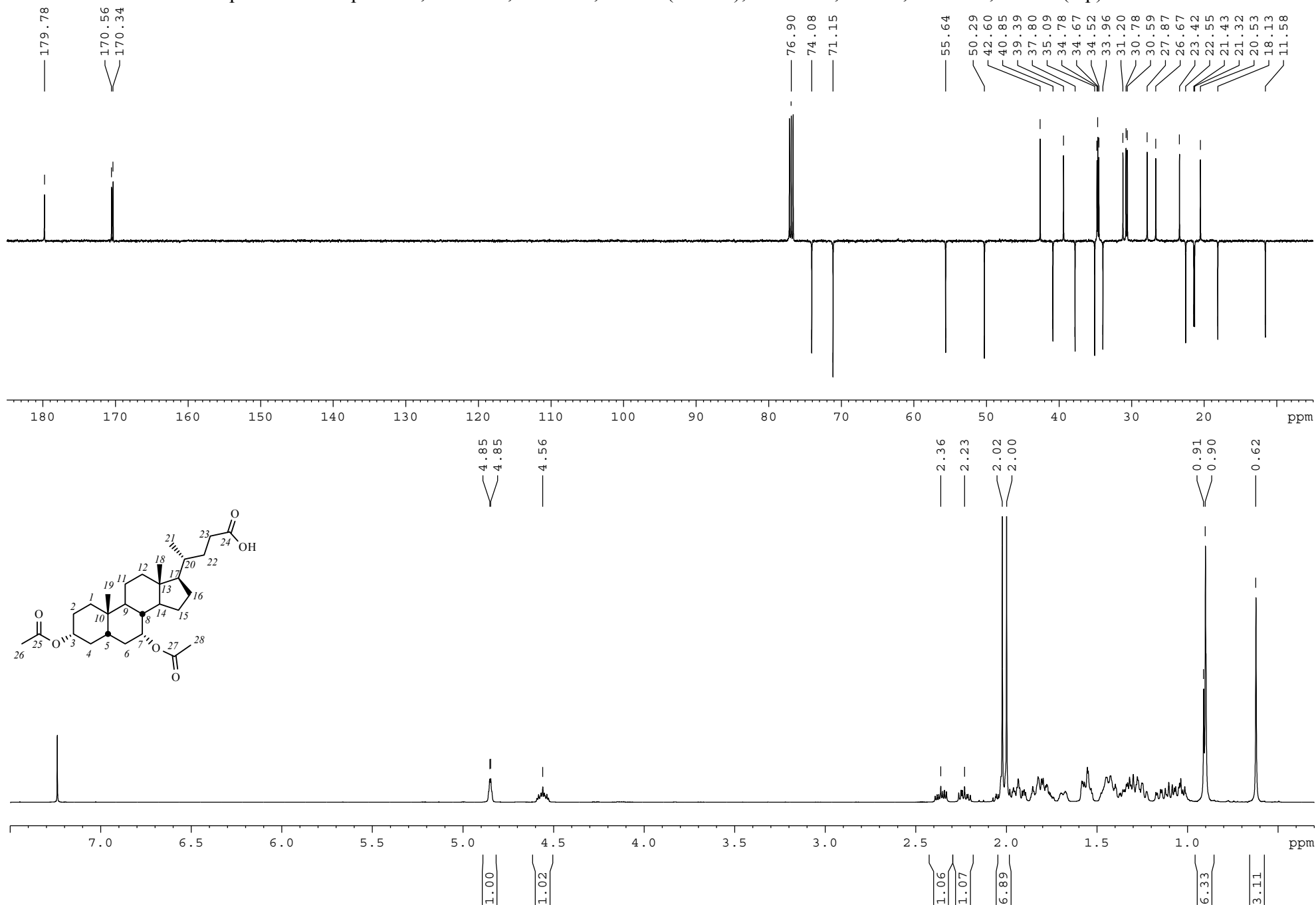
|                                            | Lead-like Space | Drug-like Space | Known Drug Space |
|--------------------------------------------|-----------------|-----------------|------------------|
| Molecular weight (g mol <sup>-1</sup> )    | 300             | 500             | 800              |
| Lipophilicity (Log P)                      | 3               | 5               | 6.5              |
| Hydrogen bond donors (HD)                  | 3               | 5               | 7                |
| Hydrogen bond acceptors (HA)               | 3               | 10              | 15               |
| Polar surface area (Å <sup>2</sup> ) (PSA) | 60              | 140             | 180              |
| Rotatable bonds (RB)                       | 3               | 10              | 17               |

From: Zhu, F; Logan, G.; Reynisson J. Wine Compounds as a Source for HTS Screening Collections. A Feasibility Study. *Mol. Inf.*, **2012**, *31*, 847 – 855, DOI:10.1002/minf.201200103

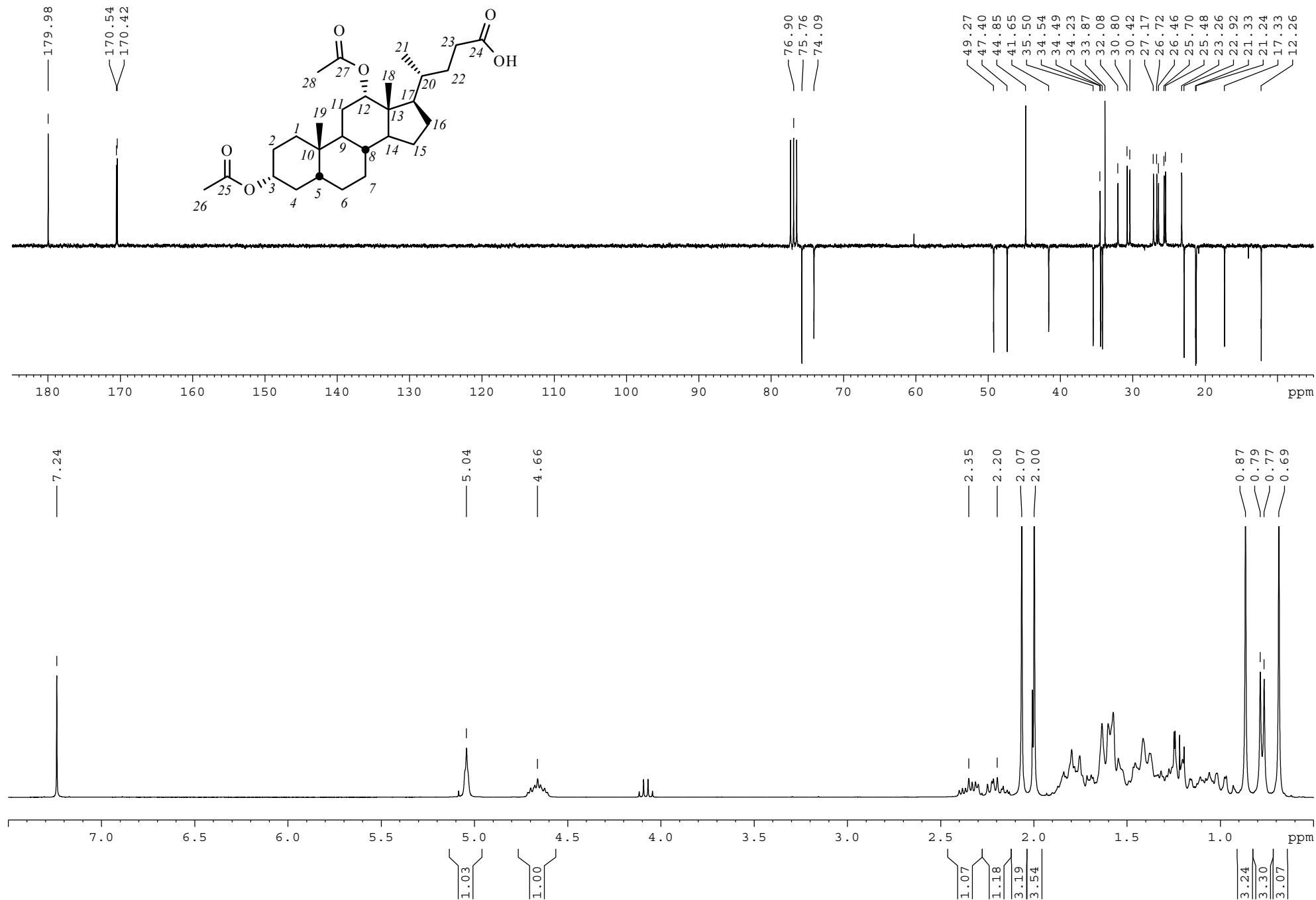
Spectra of Compound **1**,  $^1\text{H}$  NMR, 300MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 75MHz,  $\text{CDCl}_3$  (top)



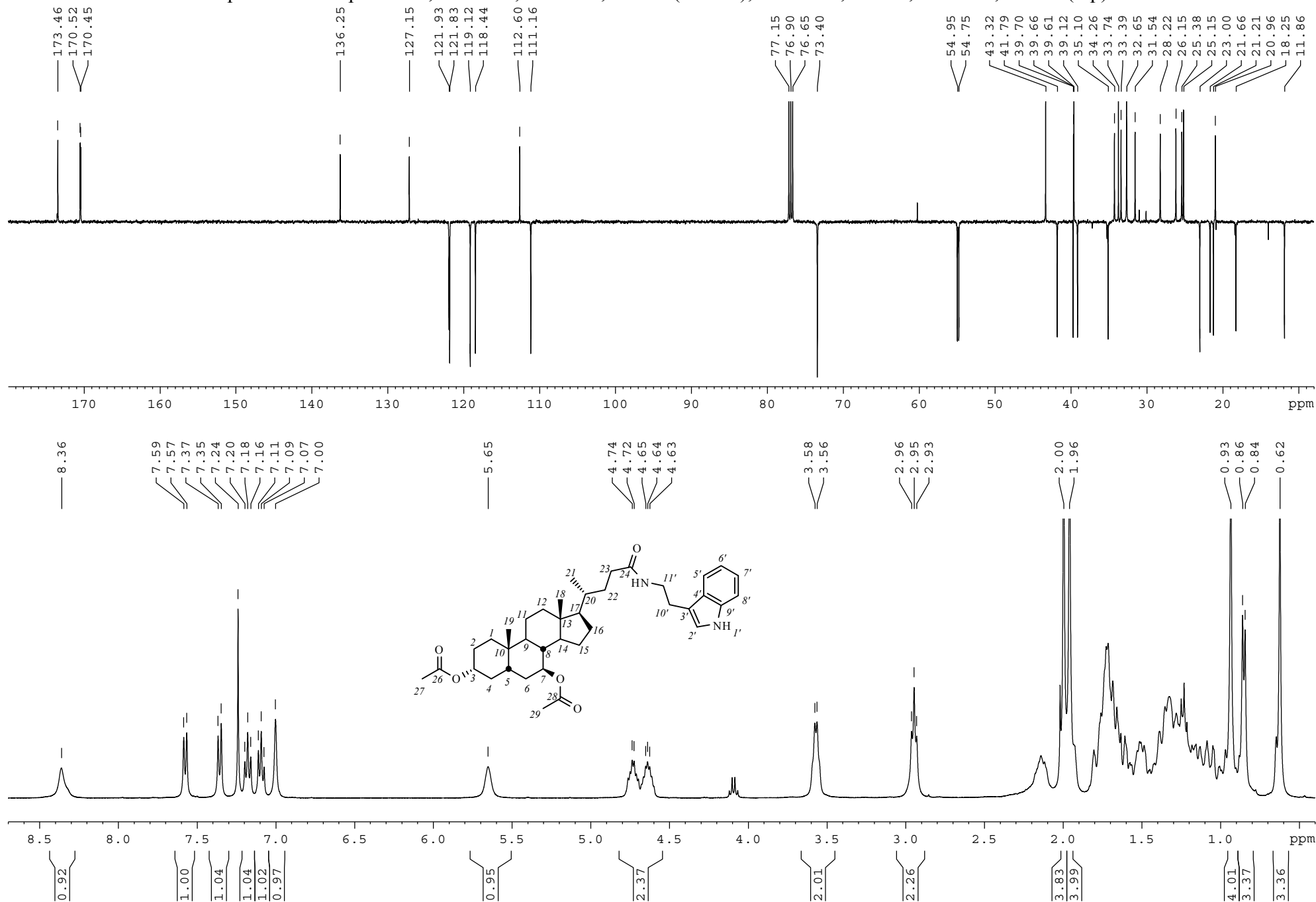
Spectra of Compound **2**,  $^1\text{H}$  NMR, 500MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 125MHz,  $\text{CDCl}_3$  (top)



Spectra of Compound **3**,  $^1\text{H}$  NMR, 300MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 75MHz,  $\text{CDCl}_3$  (top)

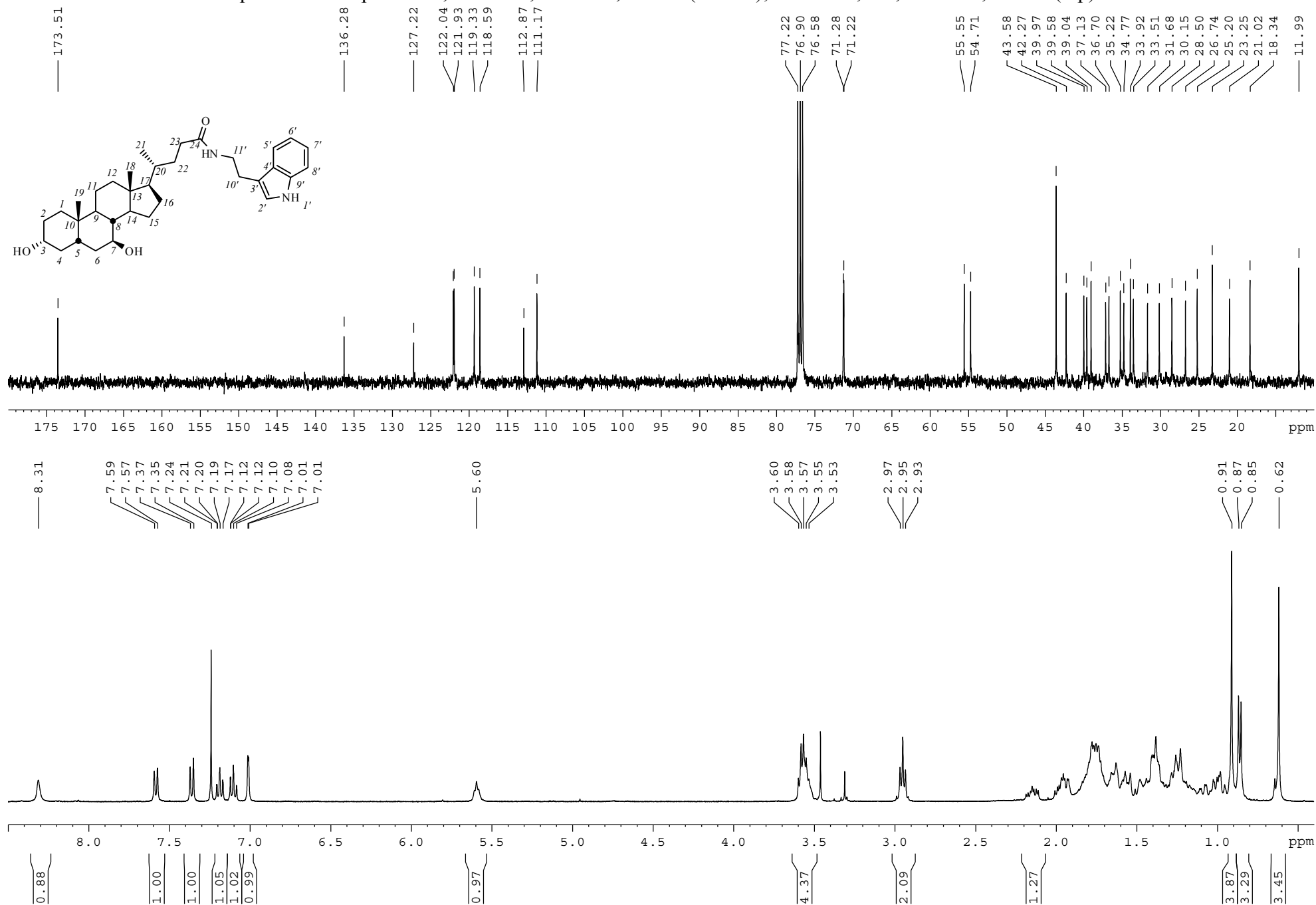


Spectra of Compound **1a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 125MHz,  $\text{CDCl}_3$  (top)

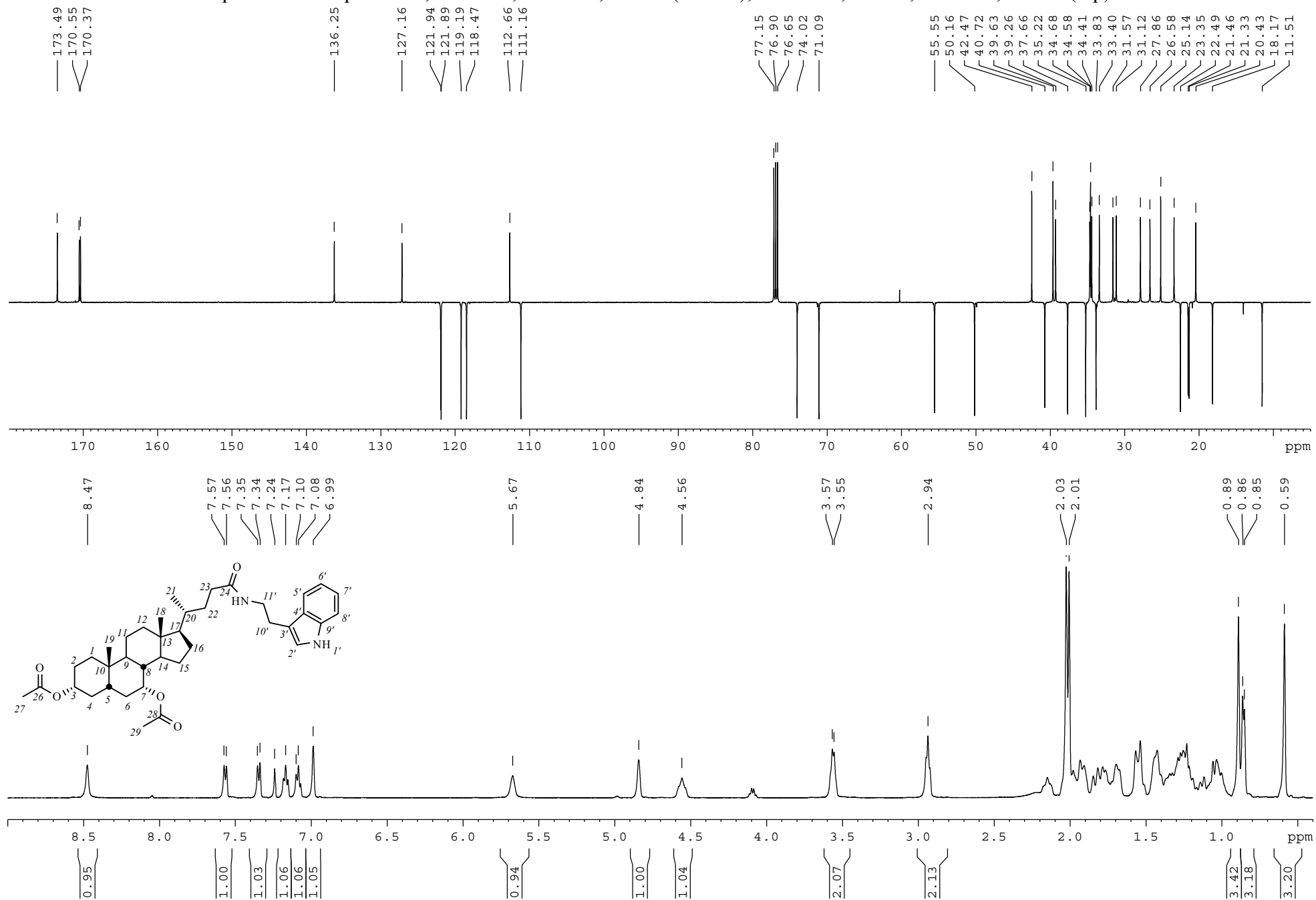




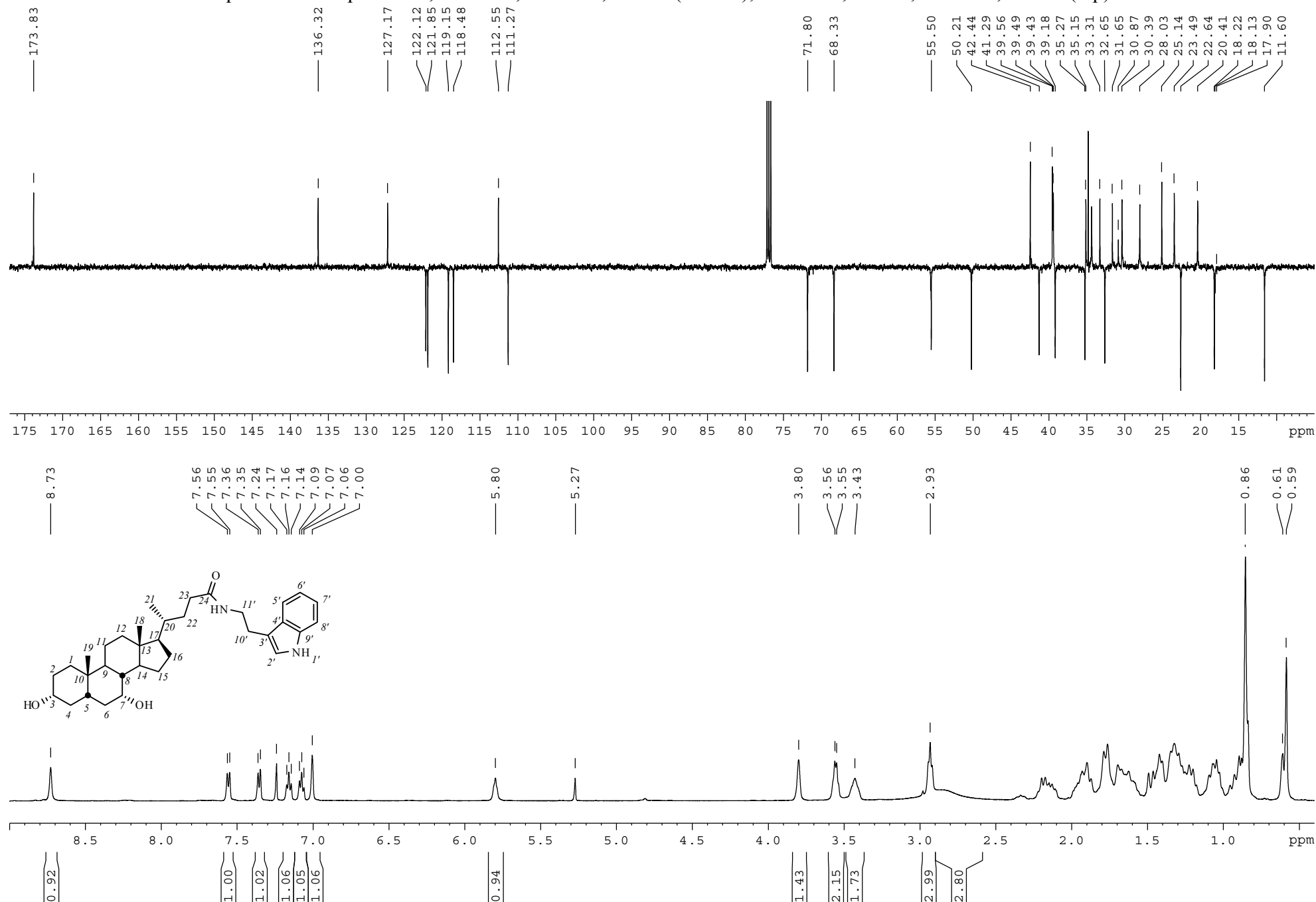
Spectra of Compound **1b**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, BB, 100MHz,  $\text{CDCl}_3$  (top)



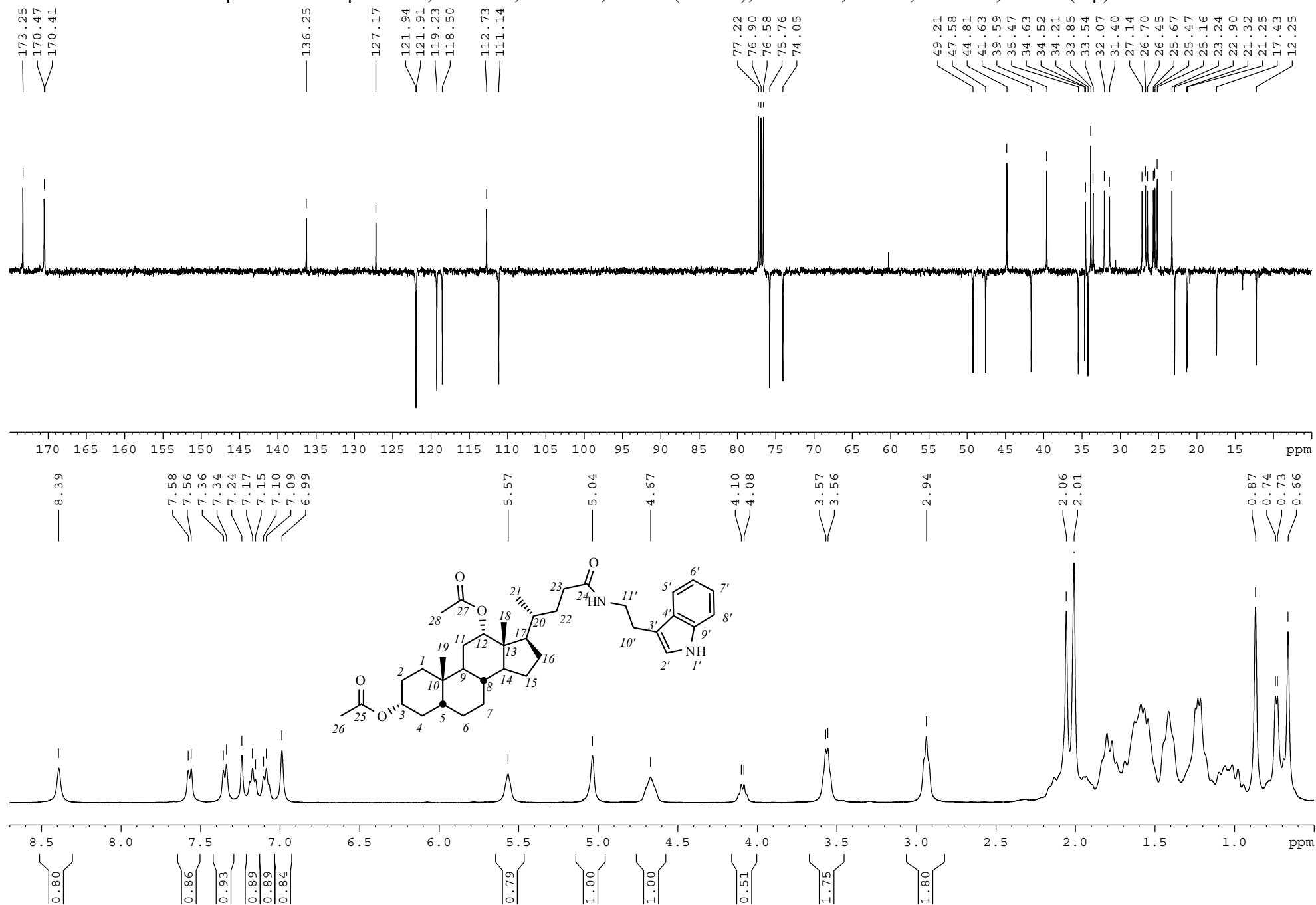
Spectra of Compound **2a**,  $^1\text{H}$  NMR, 500MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 125MHz,  $\text{CDCl}_3$  (top)



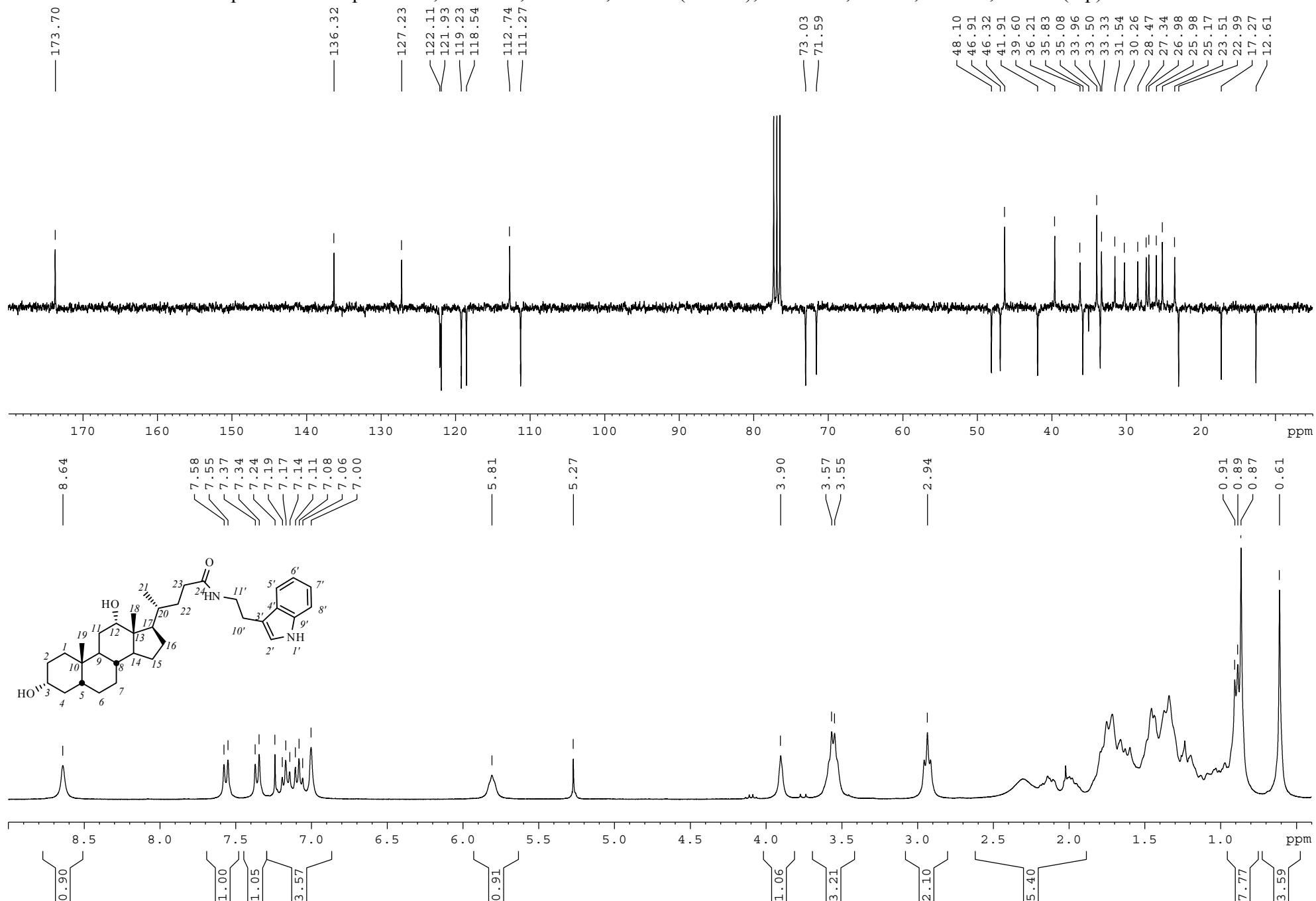
Spectra of Compound **2b**,  $^1\text{H}$  NMR, 500MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 125MHz,  $\text{CDCl}_3$  (top)



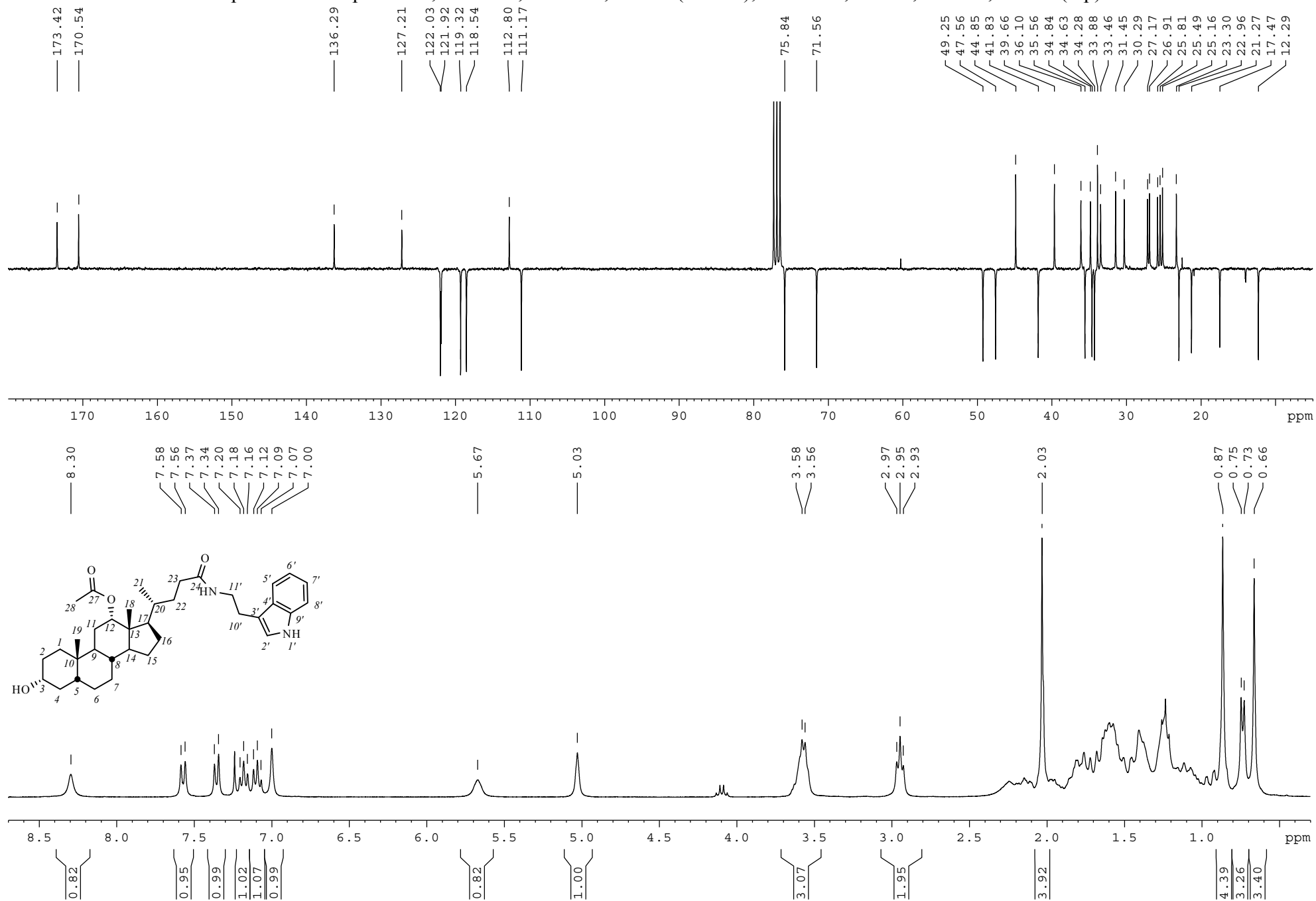
Spectra of Compound **3a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)



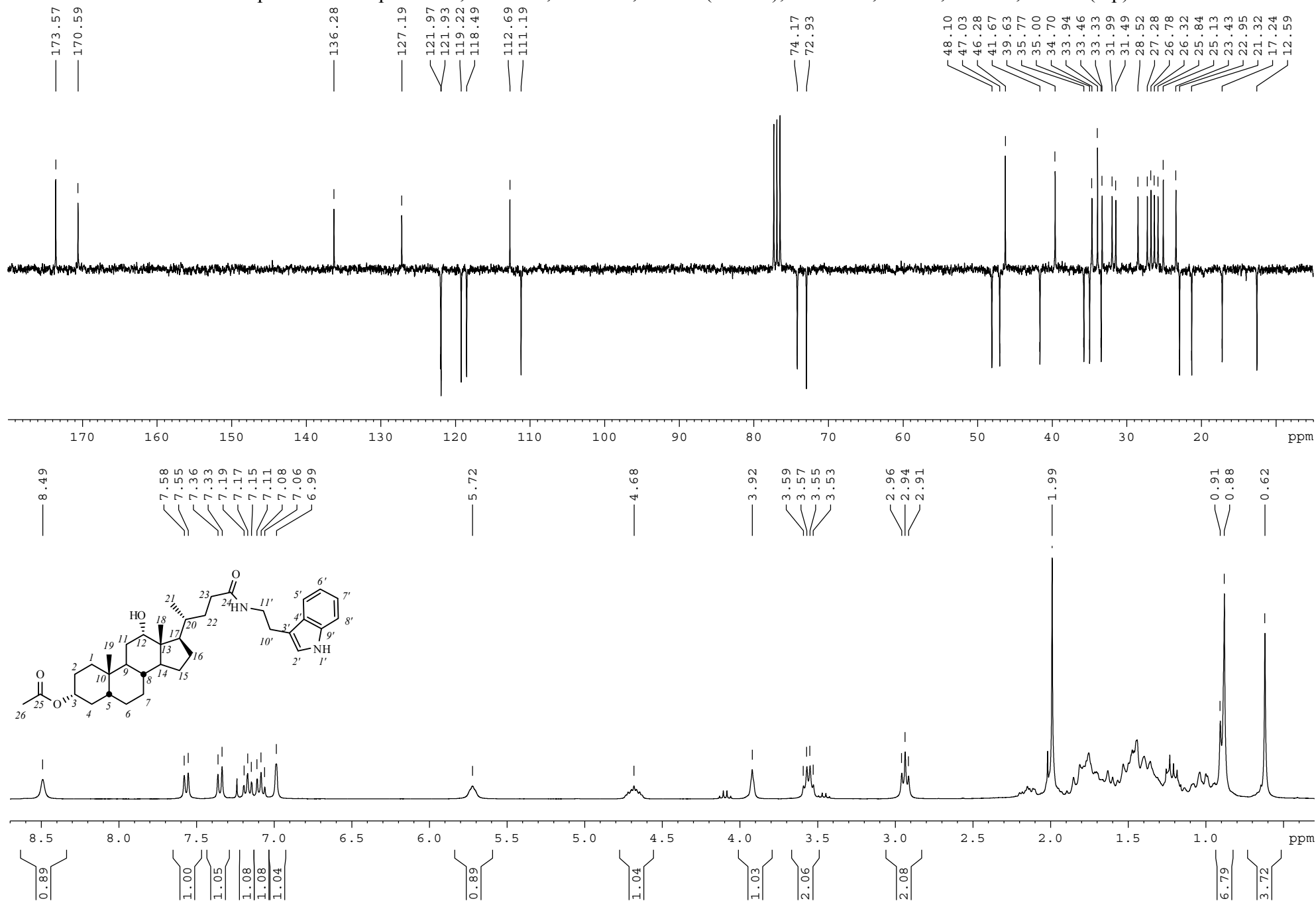
Spectra of Compound **3b**,  $^1\text{H}$  NMR, 300MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 75MHz,  $\text{CDCl}_3$  (top)



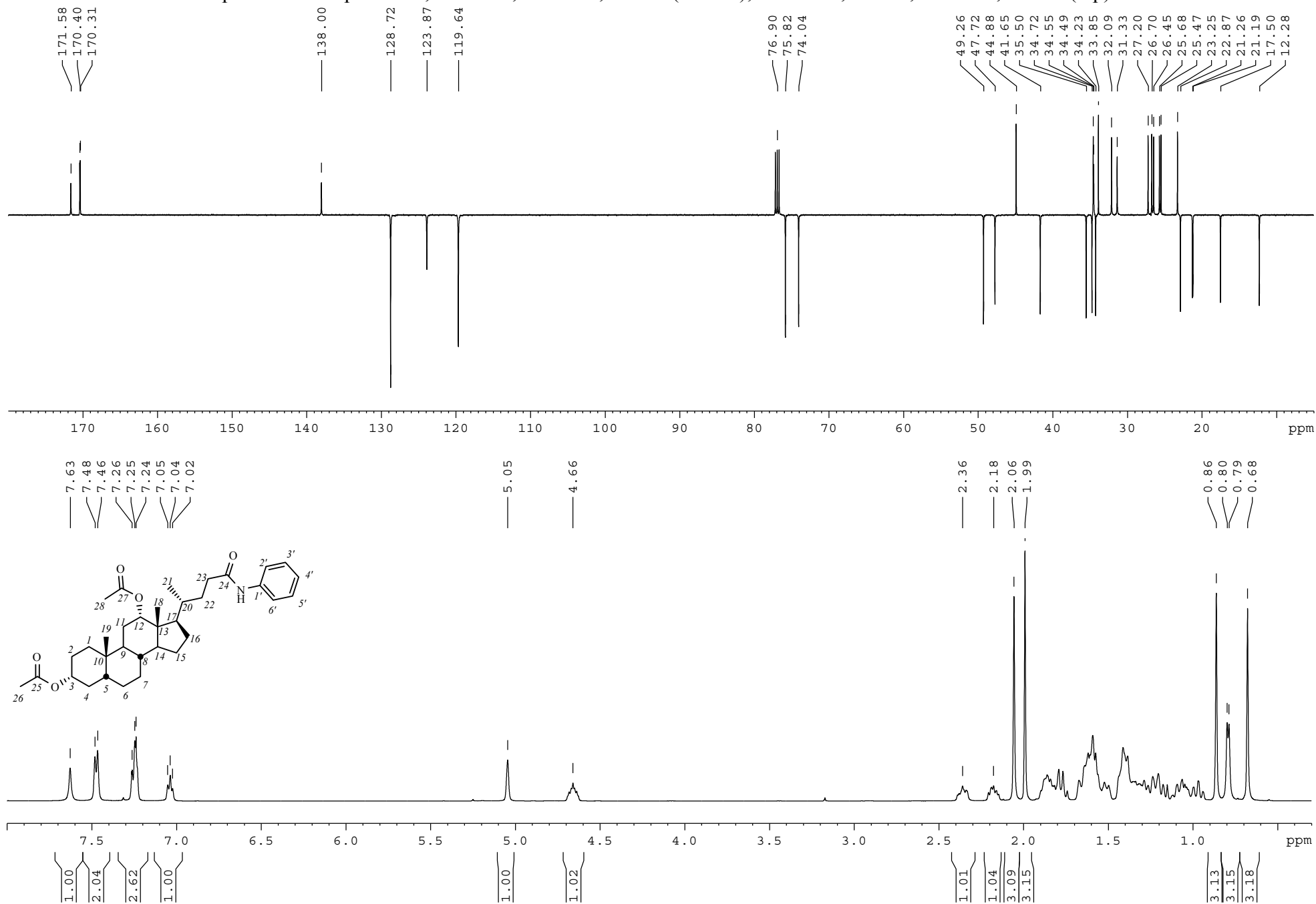
Spectra of Compound **3c**,  $^1\text{H}$  NMR, 300MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 75MHz,  $\text{CDCl}_3$  (top)



Spectra of Compound **3d**,  $^1\text{H}$  NMR, 300MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 75MHz,  $\text{CDCl}_3$  (top)

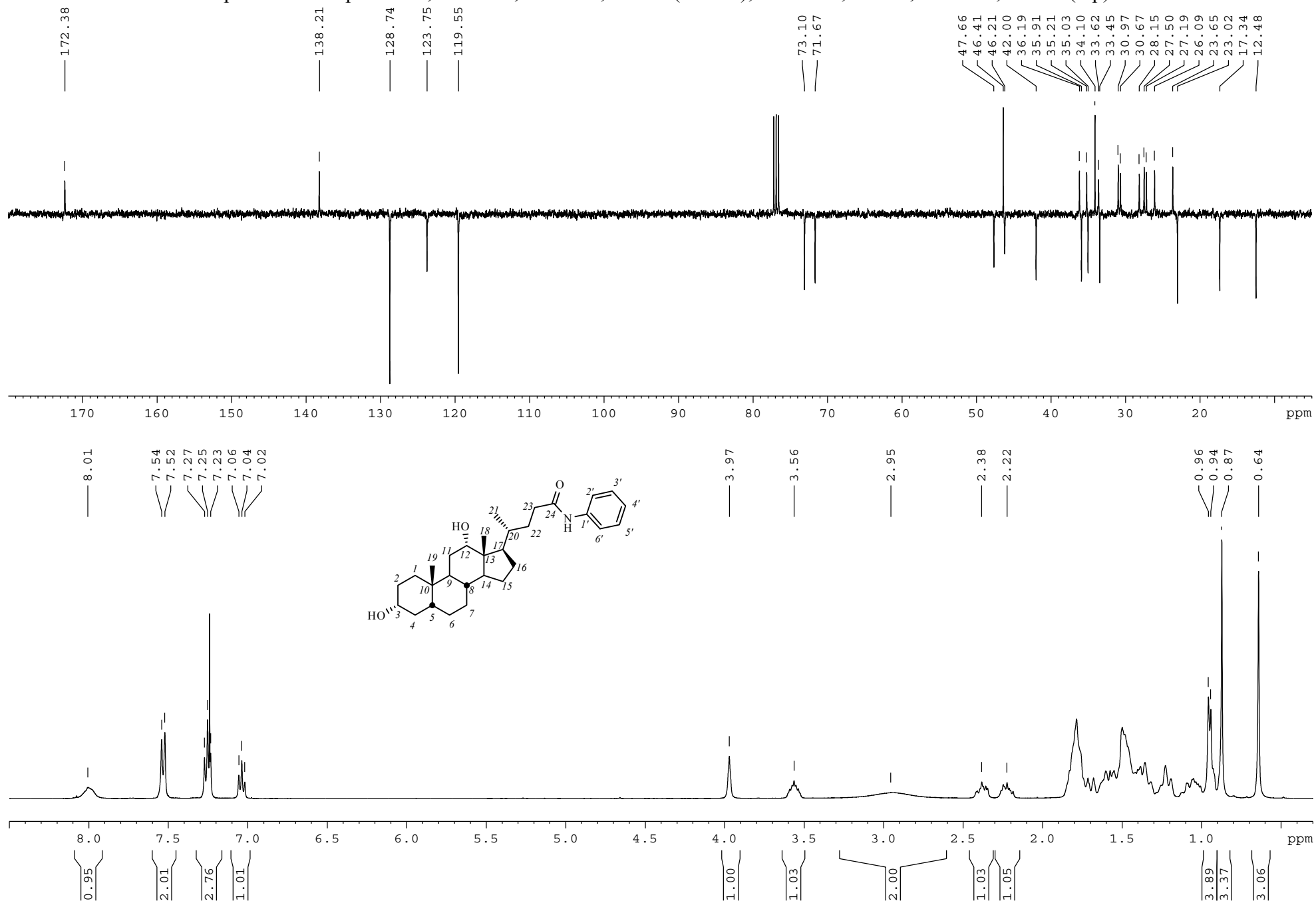


Spectra of Compound **4a**,  $^1\text{H}$  NMR, 500MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 125MHz,  $\text{CDCl}_3$  (top)

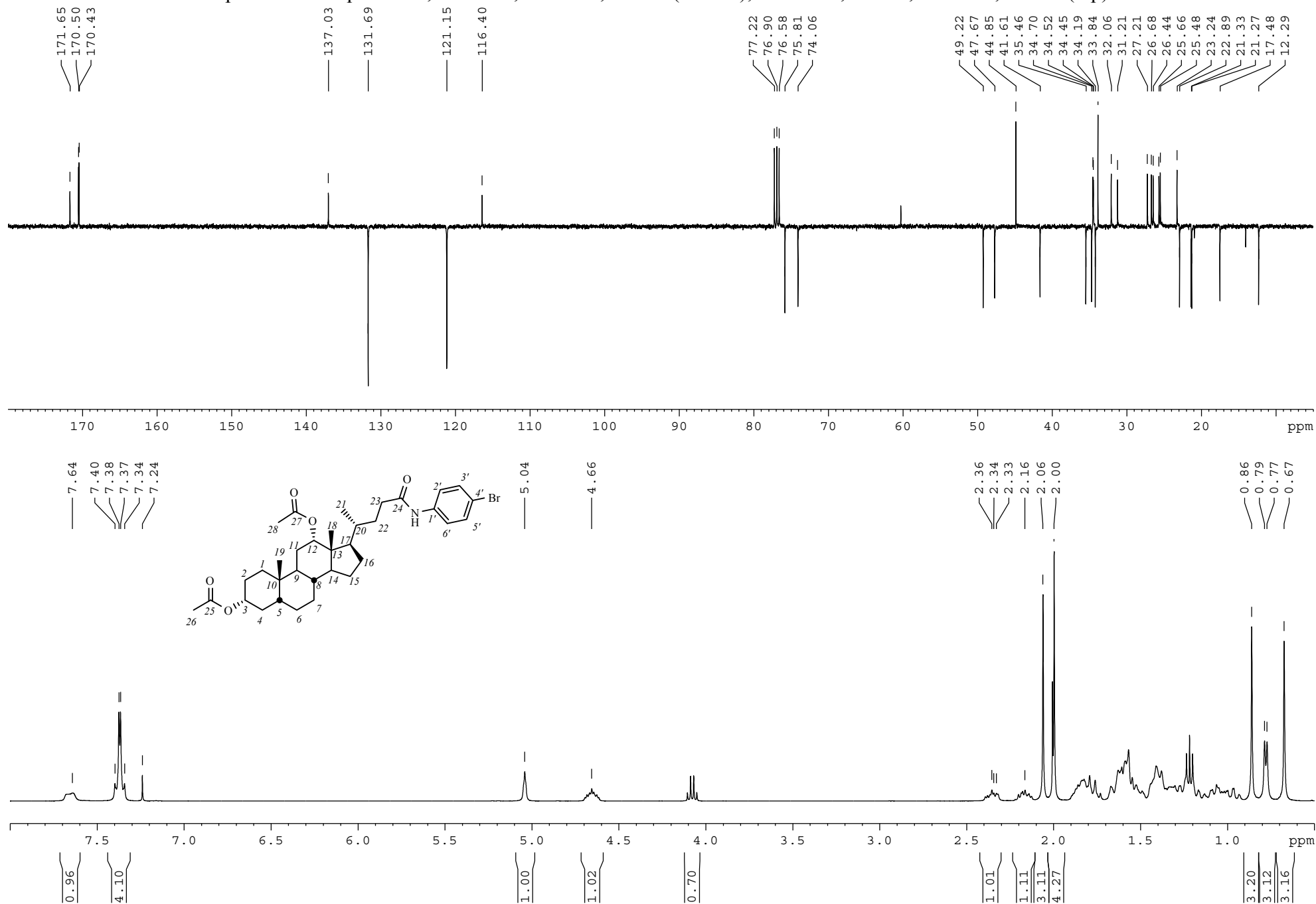




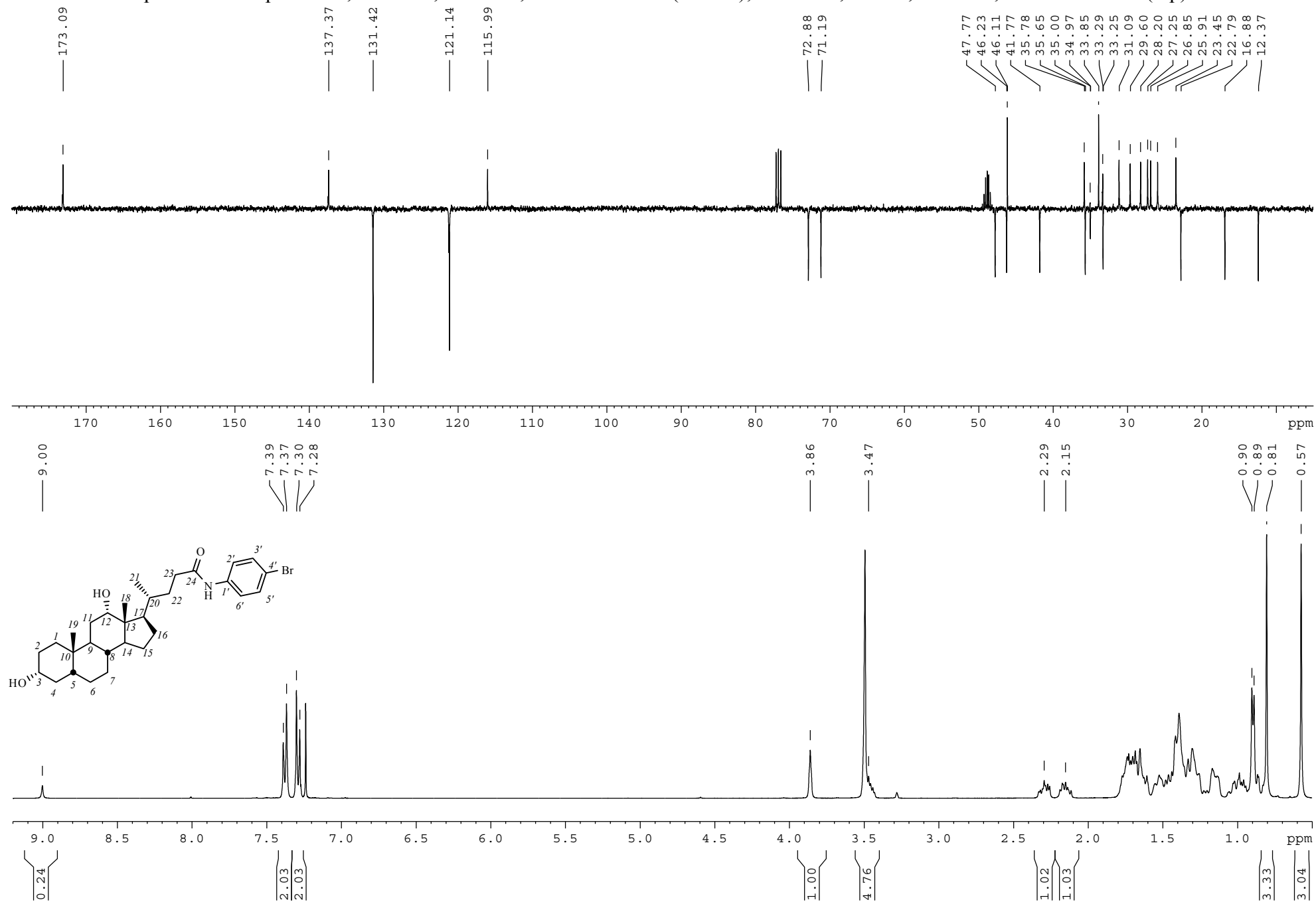
Spectra of Compound **4b**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)



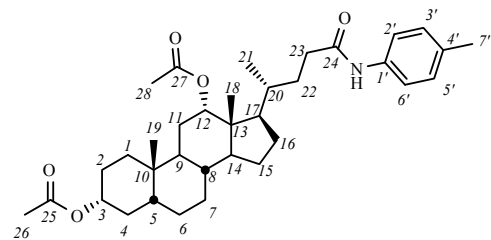
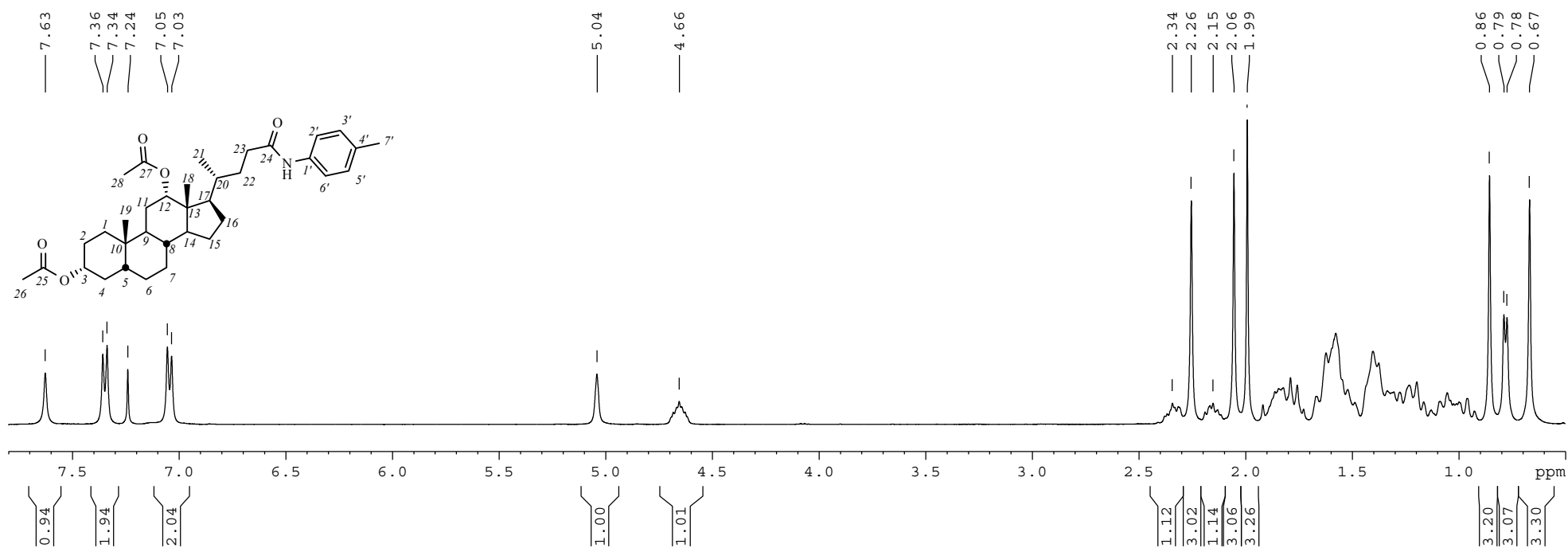
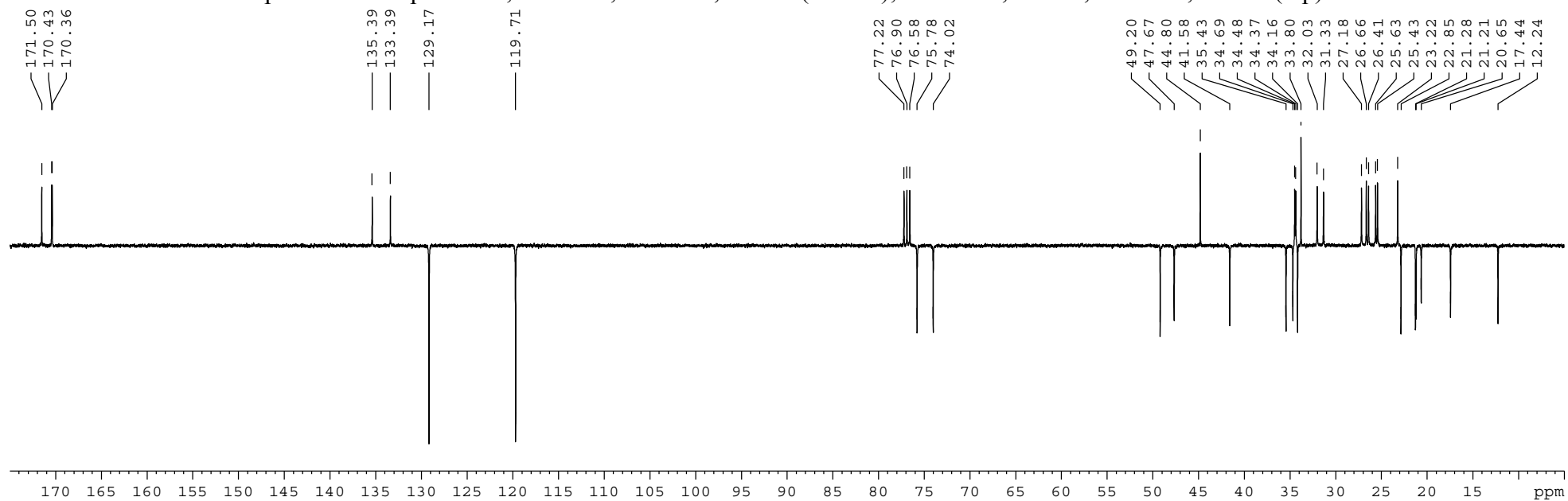
Spectra of Compound **5a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)



Spectra of Compound **5b**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3+\text{CD}_3\text{OD}$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3+\text{CD}_3\text{OD}$  (top)



Spectra of Compound **6a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)

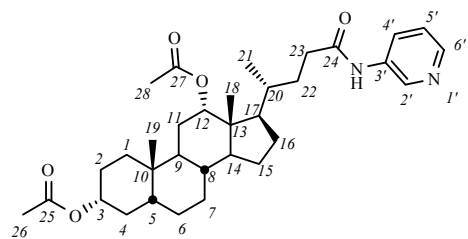
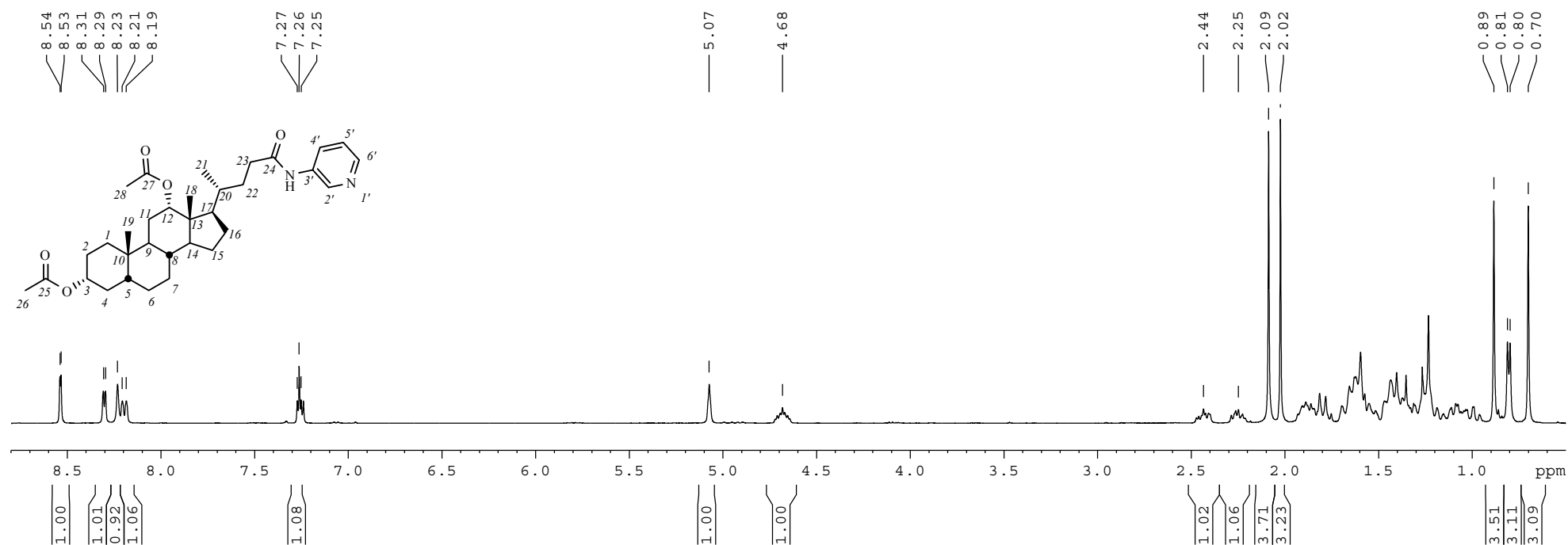
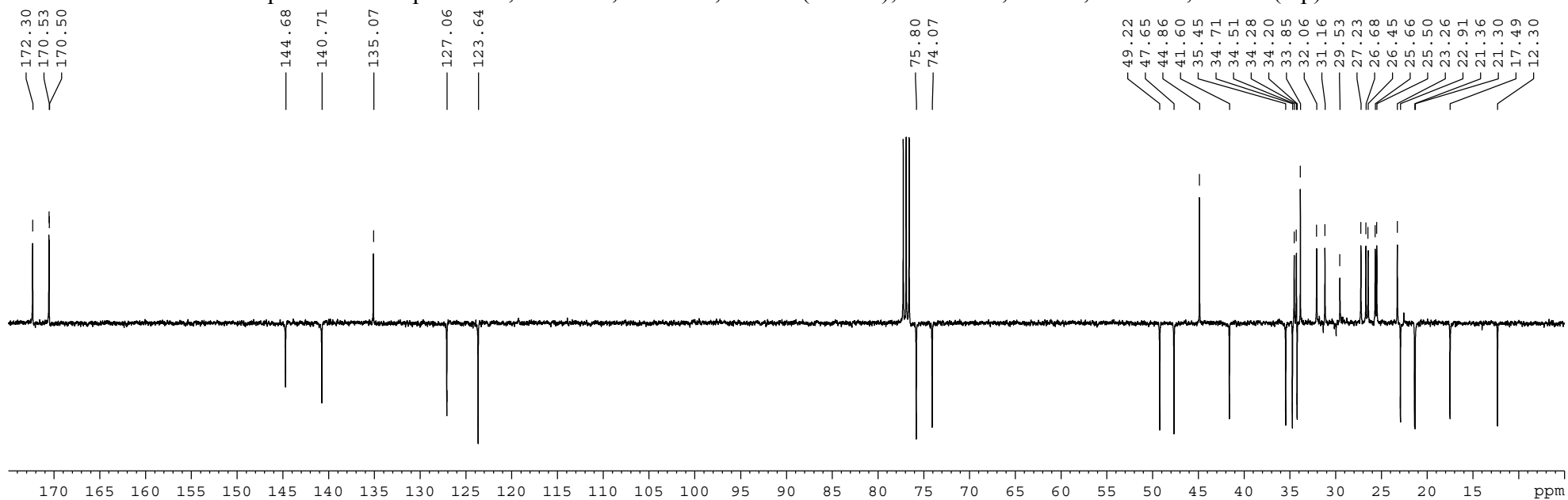


**Chemical structure of compound 1:** A steroid derivative with a 3-hydroxyl group, a 17-phenylacetate group, and a 13-methyl group. The structure is labeled with carbons 1 through 24 and protons 1' through 7'.

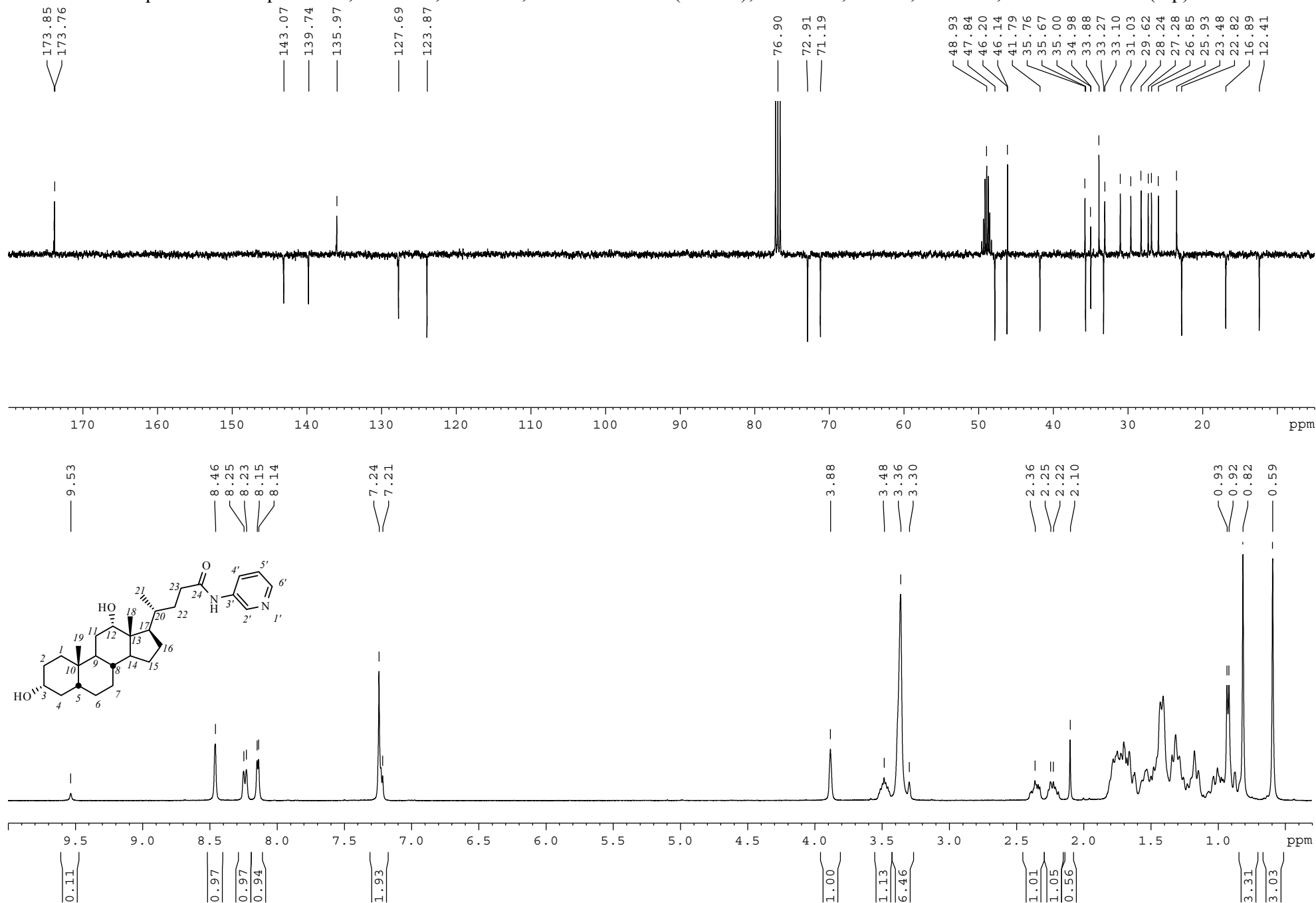
**<sup>13</sup>C NMR spectrum (101 MHz, CDCl<sub>3</sub>):** The spectrum shows peaks at 172.18, 135.61, 133.34, 129.23, 119.69, 73.13, 71.65, 47.80, 46.40, 46.37, 41.96, 36.19, 35.88, 35.18, 35.02, 34.05, 33.67, 33.43, 31.11, 30.43, 28.25, 27.46, 27.11, 26.07, 23.63, 22.99, 20.71, 17.34, and 12.53 ppm.

**<sup>1</sup>H NMR spectrum (400 MHz, CDCl<sub>3</sub>):** The spectrum shows peaks at 7.89, 7.41, 7.38, 7.24, 7.07, 7.04, 3.96, 3.57, 2.90, 2.26, 0.97, 0.95, 0.87, and 0.64 ppm. Integration values are provided for several peaks: 0.91, 2.04, 2.11, 1.00, 1.00, 2.85, 5.33, 4.00, 3.65, and 3.06.

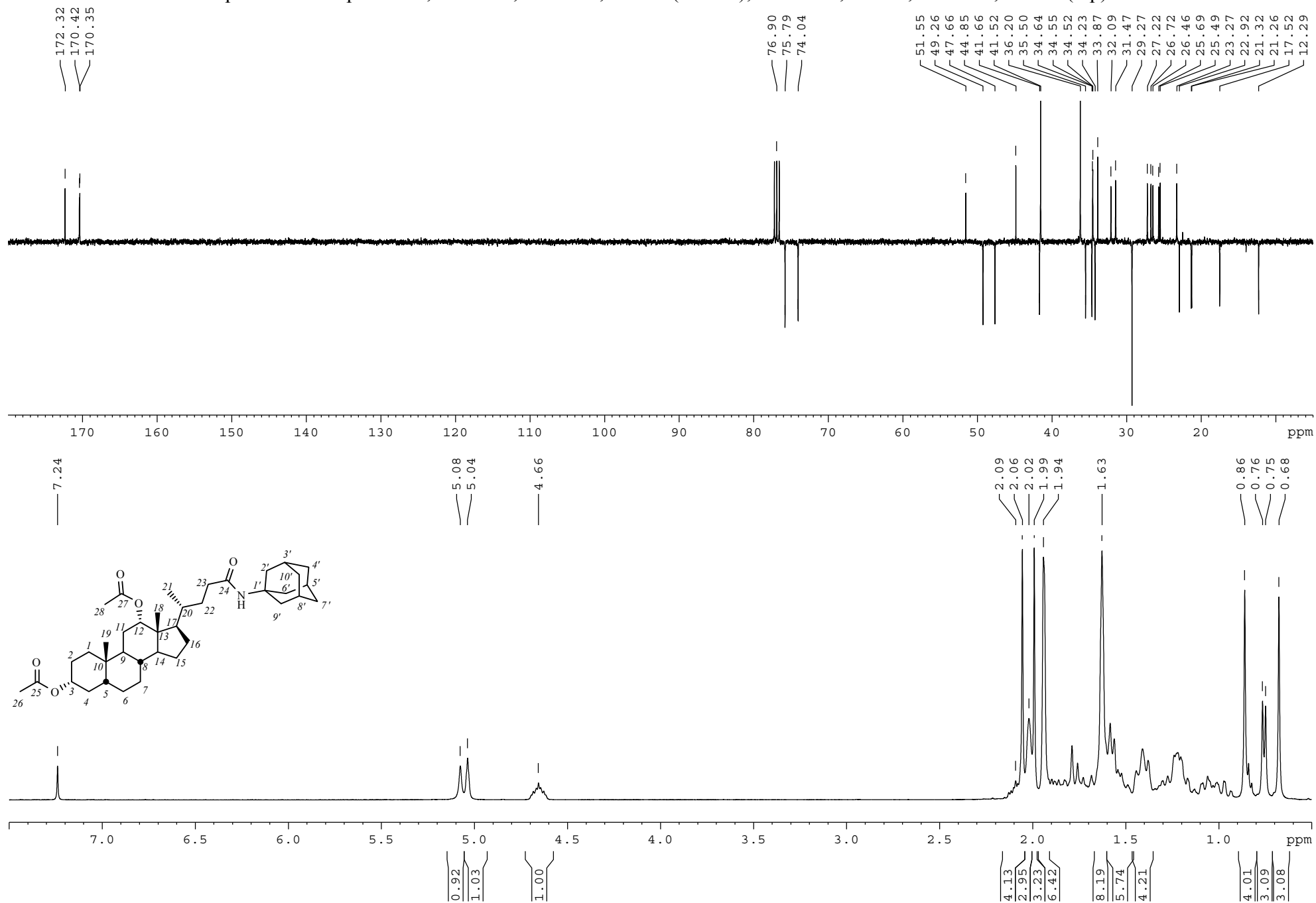
Spectra of Compound **7a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)



Spectra of Compound **7b**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3+\text{CD}_3\text{OD}$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3+\text{CD}_3\text{OD}$  (top)

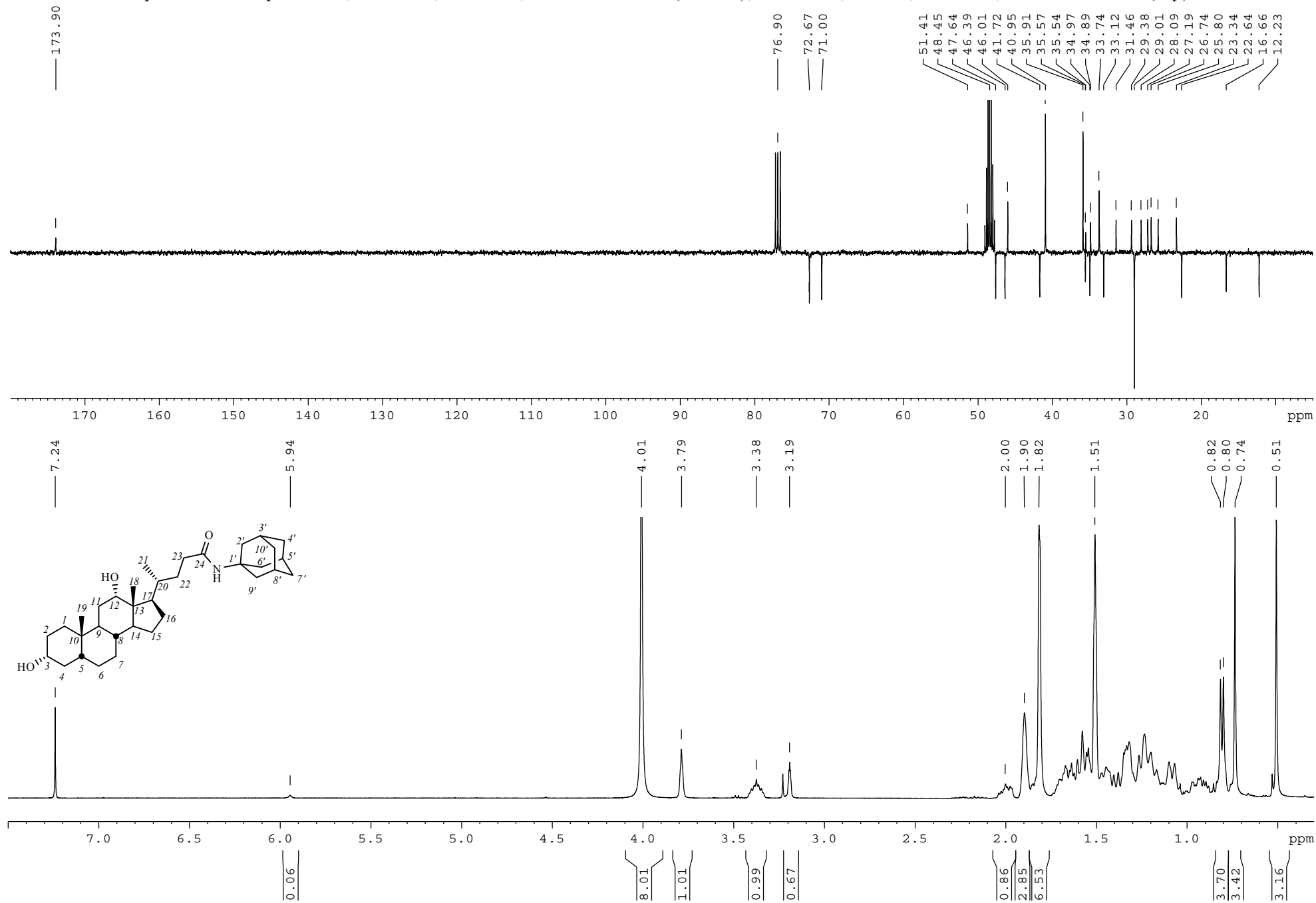


Spectra of Compound **8a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)

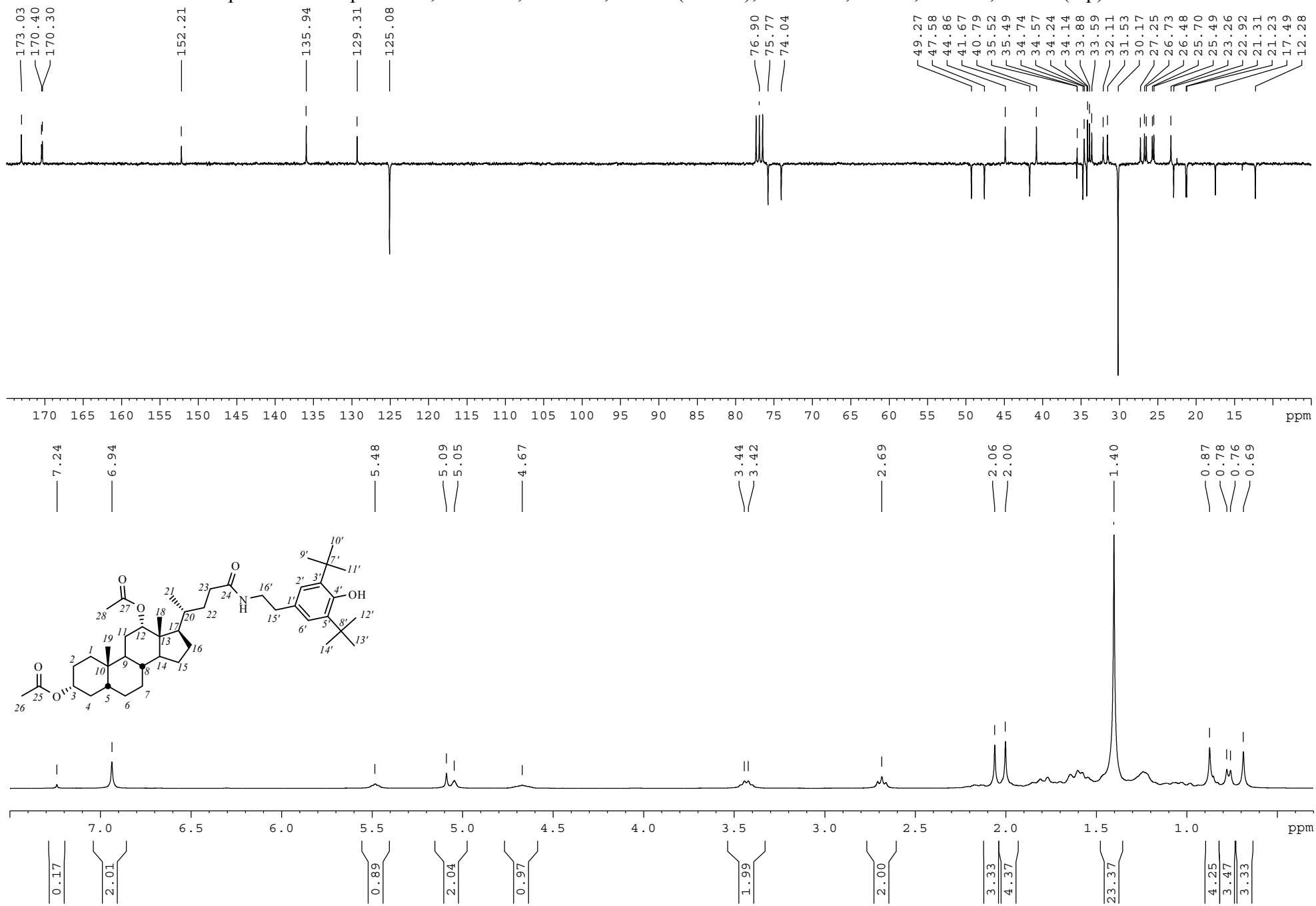




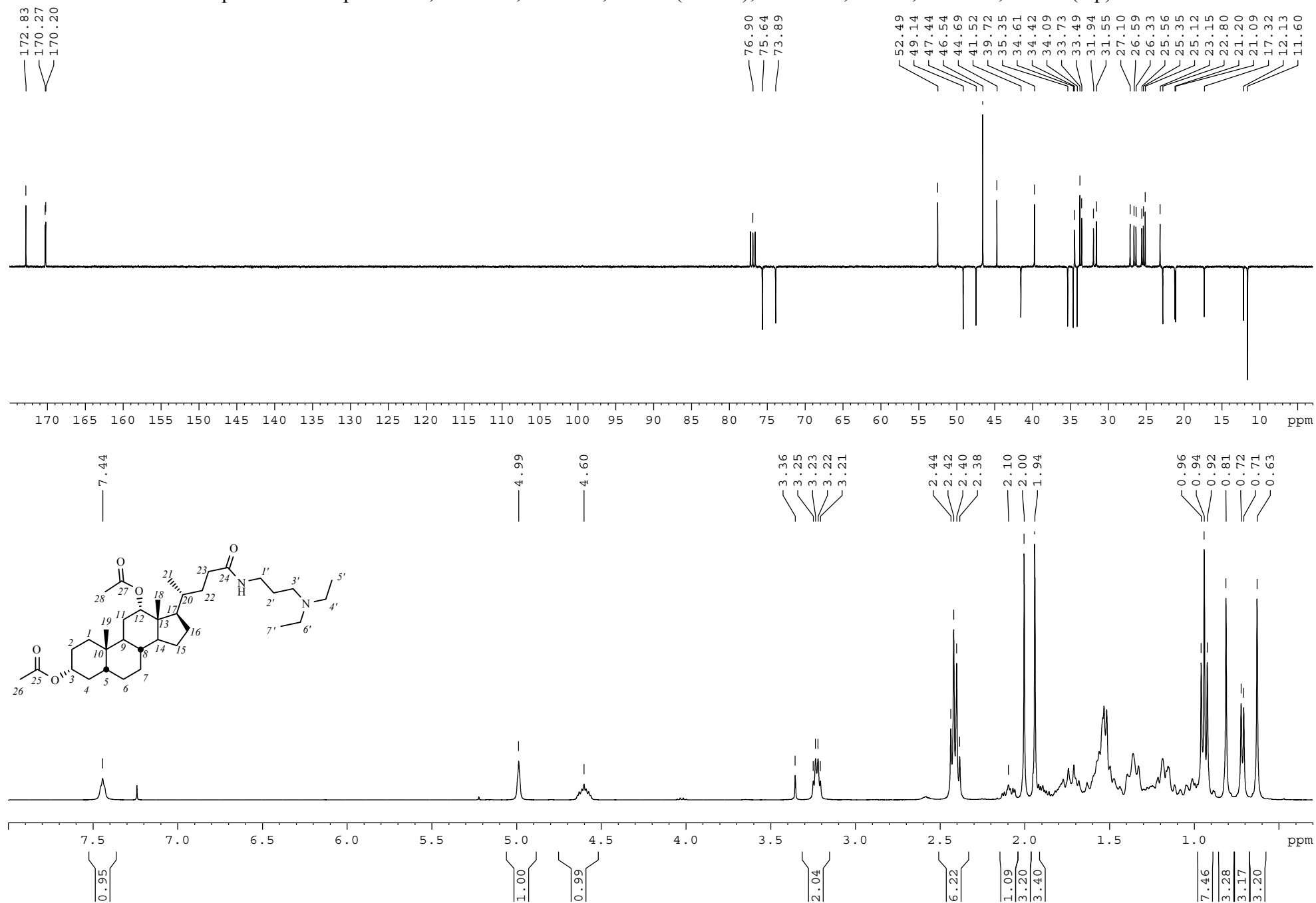
Spectra of Compound **8b**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3+\text{CD}_3\text{OD}$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3+\text{CD}_3\text{OD}$  (top)



Spectra of Compound **9a**,  $^1\text{H}$  NMR, 300MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 75MHz,  $\text{CDCl}_3$  (top)



Spectra of Compound **10a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)



Spectra of Compound **11a**,  $^1\text{H}$  NMR, 400MHz,  $\text{CDCl}_3$  (bottom);  $^{13}\text{C}$  NMR, JMOD, 100MHz,  $\text{CDCl}_3$  (top)

