

## Supporting Information

# Design of novel RNA ligands that bind stem-bulge HIV-1 TAR RNA

Maria Duca,<sup>1</sup> Vincent Malnuit,<sup>1</sup> Florent Barbault,<sup>2</sup> Rachid Benhida<sup>1\*</sup>

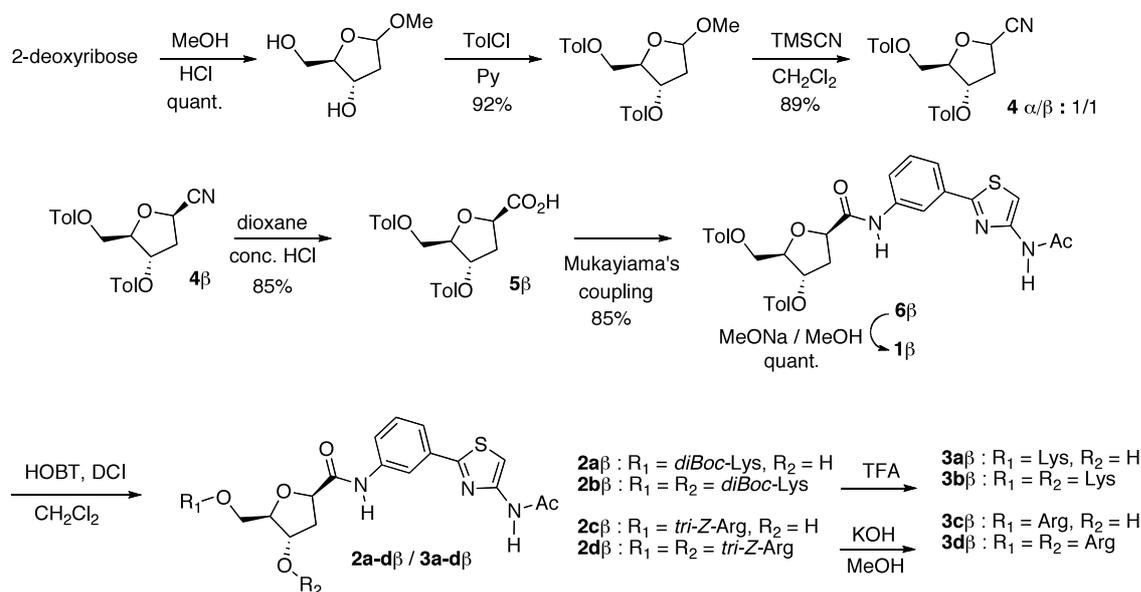
<sup>1</sup>Laboratoire de Chimie des Molécules Bioactives et des Arômes UMR 6001, Université de Nice Sophia Antipolis, CNRS Institut de Chimie de Nice, Parc Valrose, 06108 Nice Cedex 2. <sup>2</sup>Université Paris Diderot, Laboratoire ITODYS 15 rue Jean de Baïf, bâtiment Lavoisier, 75013 Paris France.

**Abbreviations:** Boc (tert-butoxycarbonyl); HOBt (hydroxybenzotriazole); DIC (*N,N'*-diisopropylcarbodiimide), DMAP (dimethylaminopyridine); TFA (trifluoroacetic acid); Z (carbobenzyloxy protecting group).

## Materials and Equipment

**Chemistry.** Solvents and most of the starting material were purchased from Aldrich and Alfa Aesar. High resolution mass spectra (HRMS) were obtained with a LTQ Orbitrap hybrid mass spectrometer with an electrospray ionization probe (ThermoScientific, San Jose, CA) by direct infusion from a pump syringe to confirm correct molar mass and high purity of compounds. The <sup>1</sup>H and <sup>13</sup>C NMR (200 MHz and 500 MHz) spectra were recorded on a Bruker AC 200 spectrometer. Chemical shifts are expressed as parts per million from tetramethylsilane. Splitting patterns have been designated as follows: s (singlet), d (doublet), t (triplet) and m (multiplet), br (broad). Coupling constants (J values) are listed in hertz (Hz). Analytical thin-layer chromatography (TLC) was conducted on Merck (VWR) precoated silica gel 60F254 plates and compounds were visualized with ninhydrin test and/or under ultraviolet light (254 nm). Column chromatographies were carried out on silica gel (Merck, SDS 60A, 40-63 μm, VWR). Final products were analyzed by HPLC using reverse phase analytical column (Thermo, RP C18 250 x 4.6 mm, 5μ) on a Waters<sup>TM</sup> 600E pump and a photodiode array detector Waters<sup>TM</sup>. Data were monitored using Waters Millennium software. All HPLC analysis was run at room temperature. Solvent A and solvent B respectively, 0.1% TFA in water and 0.1% TFA in acetonitrile were used for HPLC studies. For HPLC analysis, a gradient of A/B (100/0 to 40/60 for 30 min) was employed at a flow rate of 1 mL/min.

### Synthesis of ligands 3a-d $\beta$ .



Scheme S1. Synthesis of ligands 3a-d $\beta$ .

**5'-(*N*- $\alpha,\omega$ -di-Boc-L-lysyl)-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (2a $\beta$ ) and 3',5'-di-(*N*- $\alpha,\omega$ -di-Boc-L-lysyl)-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (2b $\beta$ ).** To a solution of commercially available *N*- $\alpha,\omega$ -di-Boc-L-lysine (308 mg, 0.583 mmol) in anhydrous dichloromethane (5 mL) were added successively HOBT (90 mg, 0.663 mmol), DCI (92  $\mu$ L, 0.663 mmol), triethylamine (103  $\mu$ L, 0.663 mmol) and DMAP (6.5 mg, 0.053 mmol). Finally, a solution of 1 $\beta$  (100 mg, 0.265 mmol) in dichloromethane (5 mL) was added under argon atmosphere. The mixture was stirred for 4 h at room temperature then washed twice with water (2 x 15 mL) and dried over MgSO<sub>4</sub>. Concentration under reduced pressure followed by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5) led to desired compounds 2a $\beta$  and 2b $\beta$ . 2a $\beta$ : yield 29% (56.1 mg);  $R_f$  0.54 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (CD<sub>3</sub>OD)  $\delta$  1.15-1.45 (m, 22H), 1.55-1.80 (m, 2H), 2.20 (s, 3H), 2.30-2.40 (m, 2H), 2.80-3.00 (m, 2H), 4.00-4.10 (m, 2H), 4.40-4.50 (m, 1H), 4.67-4.70 (m, 1H), 5.20-5.25 (m, 1H), 7.25-7.40 (m, 2H), 7.55-7.70 (m, 2H), 8.21 (s, 1H); <sup>13</sup>C NMR (CD<sub>3</sub>OD)  $\delta$  23.1, 24.0, 24.6, 29.2, 30.9, 32.5, 40.7, 41.3, 43.1, 55.9, 66.5, 73.5, 80.1, 81.1, 87.6, 109.6, 119.6, 121.4, 123.9, 130.6, 137.1, 139.6, 151.1, 158.6, 158.9, 159.0, 159.8, 171.3, 173.5, 175.3; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup> 706.3115 (C<sub>33</sub>H<sub>47</sub>N<sub>5</sub>O<sub>10</sub>S requires 706.3116). 2b $\beta$ : yield 68% (191.7 mg);  $R_f$  0.75 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (CD<sub>3</sub>OD)  $\delta$  1.20-1.50 (m, 44H), 1.60-1.90 (m, 4H), 2.21 (s, 3H), 2.30-2.60 (m, 2H), 2.80-3.10 (m, 4H), 4.00-4.10 (m, 2H), 4.30-4.50 (m, 3H), 4.60-4.70 (m, 1H), 5.20-5.30 (m, 1H), 7.30-7.40 (m, 2H), 7.60-7.70 (m, 2H), 8.22 (s, 1H); <sup>13</sup>C NMR (CD<sub>3</sub>OD)  $\delta$  23.2, 24.1, 24.6, 29.3, 31.0, 32.5, 38.0, 41.4, 43.1, 55.8, 66.2, 77.8, 80.3, 80.4, 81.1, 85.8, 109.6, 119.7, 121.5, 124.0, 130.6, 137.2, 139.6, 151.1, 158.6, 159.0, 159.4, 159.7, 160.3, 171.2, 172.3, 174.4, 175.1; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup> 1034.5124 (C<sub>49</sub>H<sub>76</sub>N<sub>7</sub>O<sub>15</sub>S requires 1034.5115).

**5'-(*N*- $\alpha,\delta,\omega$ -tri-Z-L-arginyl)-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (2c $\beta$ ) and 3',5'-di-(*N*- $\alpha,\delta,\omega$ -tri-Z-L-arginyl)-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (2d $\beta$ ).** To a solution of *N*- $\alpha,\delta,\omega$ -tri-Z-L-arginine (336 mg, 0.583 mmol) obtained using a previously reported procedure<sup>1</sup> in

anhydrous dichloromethane (5 mL) were added successively HOBt (90 mg, 0.663 mmol), DCI (103  $\mu$ L, 0.663 mmol), triethylamine (92  $\mu$ L, 0.663 mmol) and DMAP (6.5 mg, 0.053 mmol). Finally, a solution of **1 $\beta$**  (100 mg, 0.265 mmol) in dichloromethane (5 mL) was added under argon atmosphere. The mixture was stirred for 6 h at room temperature then washed twice with water (2 x 15 mL) and dried over MgSO<sub>4</sub>. Concentration under reduced pressure followed by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5) led to desired compounds **2c $\beta$**  and **2d $\beta$** . **2c $\beta$** : yield 28% (79 mg); R<sub>f</sub> 0.61 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (DMSO-d<sub>6</sub>)  $\delta$  1.50-1.70 (m, 4H), 2.00-2.10 (m, 2H), 2.20 (s, 3H), 3.50-3.70 (m, 2H), 3.75-3.95 (m, 3H), 4.00-4.05 (m, 1H), 4.10-4.20 (m, 1H), 4.59 (t, J = 7.6 Hz, 1H), 5.45-5.55 (m, 6H), 7.20-7.50 (m, 19H) and 8.33 (s, 1H); <sup>13</sup>C NMR (DMSO-d<sub>6</sub>)  $\delta$  23.5, 24.7, 29.1, 60.7, 61.5, 65.9, 66.2, 73.0, 77.0, 90.2, 105.7, 117.1, 121.3, 123.8, 127.6, 128.2, 129.5, 132.8, 134.3, 140.0, 150.2, 154.7, 156.9, 159.9, 162.3, 168.6, 171.6 and 172.2; mass spectrum (ESI), *m/z* (M+H)<sup>+</sup> 936.3238 (C<sub>47</sub>H<sub>50</sub>N<sub>7</sub>O<sub>12</sub>S requires 936.3233). **2d $\beta$** : yield 64% (287 mg); R<sub>f</sub> 0.87 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  1.50-1.70 (m, 12H), 2.15 (s, 3H), 2.20-2.30 (m, 2H), 3.60-3.70 (m, 3H), 3.80-4.00 (m, 3H), 4.20-4.40 (m, 2H), 5.00-5.20 (m, 12H), 6.21 (t, J = 7.0 Hz, 1H), 7.10-7.40 (m, 34H) and 7.98 (s, 1H); <sup>13</sup>C NMR (CDCl<sub>3</sub>)  $\delta$  23.5, 24.7, 29.1, 42.2, 47.5, 52.3, 67.0, 68.9, 82.6, 108.7, 119.2, 120.0, 128.1, 128.4, 128.8, 130.0, 136.2, 136.8, 136.9, 151.7, 155.6, 155.9, 160.5, 163.8, 172.6 and 172.7; mass spectrum (ESI), *m/z* (M+H)<sup>+</sup> 1494.5348 (C<sub>77</sub>H<sub>80</sub>N<sub>11</sub>O<sub>19</sub>S requires 1494.5347).

**General procedure for the deprotection of Boc-protected compounds.** To a solution of Boc-protected compounds (**2a $\beta$** , **2b $\beta$** , **2a $\alpha$** , **2b $\alpha$** , **2e** or **2f**) respectively, (0.10 mmol) in anhydrous dichloromethane (5 mL) were added 10 eq. of trifluoroacetic acid (1 mmol, 74.3  $\mu$ L) and the reaction mixture was stirred under argon atmosphere overnight at room temperature. The crude product was then evaporated and the residue coevaporated three times with toluene in order to eliminate the remaining trifluoroacetic acid. Crystallization in ethyl ether lead to the desired compounds **3a $\beta$** , **3b $\beta$** , **3a $\alpha$** , **3b $\alpha$** , **3e** and **3f**, respectively, in 74% to 89% yield.

**5'-L-lysyl-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (**3a $\beta$** ).** Yield 87% (44 mg); HPLC retention time = 22.2 min; <sup>1</sup>H NMR (D<sub>2</sub>O)  $\delta$  1.63-1.78 (m, 4H), 2.48 (s, 3H), 2.52-2.58 (m, 1H), 2.62-2.68 (m, 1H), 3.26 (t, 2H, J = 7.5 Hz), 3.90 (t, 2H, J = 6.0 Hz), 3.95-4.00 (m, 2H), 4.05-4.10 (m, 1H), 4.35-4.40 (m, 1H), 4.60-4.65 (m, 1H), 5.00 (t, 1H, J = 8.0 Hz), 7.49 (s, 1H), 7.60-7.68 (m, 2H), 7.65-7.70 (m, 1H) and 8.00 (s, 1H); <sup>13</sup>C NMR (D<sub>2</sub>O)  $\delta$  22.9, 27.2, 31.3, 39.9, 42.7, 55.5, 61.9, 71.9, 78.2, 88.0, 113.6, 115.9, 118.2, 120.6, 128.1, 129.4, 137.2, 160.1, 163.0, 163.3, 163.6 and 176.6; mass spectrum (ESI), *m/z* (M+H)<sup>+</sup> 506.2066 (C<sub>23</sub>H<sub>32</sub>N<sub>5</sub>O<sub>6</sub>S requires 506.2069).

**3',5'-di-L-lysyl-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (**3b $\beta$** ).** Yield 74% (47 mg); HPLC retention time = 18.4 min; <sup>1</sup>H NMR (D<sub>2</sub>O)  $\delta$  1.35-1.50 (m, 4H), 1.60-1.70 (m, 4H), 1.80-1.90 (m, 4H), 2.35 (s, 3H), 2.52-2.60 (m, 1H), 2.70-2.78 (m, 1H), 2.90 (2H, t, J = 8.0 Hz), 3.12 (t, 2H, J = 8.0 Hz), 3.75-3.85 (m, 2H), 4.10-4.20 (m, 1H), 4.33 (t, 1H, J = 6.5 Hz), 4.60-4.65 (m, 1H), 4.68-4.72 (m, 1H), 4.98 (t, 1H, J = 7.0 Hz), 7.49 (s, 1H), 7.52-7.60 (m, 2H), 7.75 (d, 1H, J = 7.5 Hz) and 8.00 (s, 1H); <sup>13</sup>C NMR (D<sub>2</sub>O)  $\delta$  22.7, 26.8, 29.7, 42.5, 66.3, 78.2, 78.7, 83.2, 110.1, 115.7, 118.0, 119.5, 121.9, 123.9, 130.4, 135.1, 137.0, 148.6, 159.6, 163.2, 169.7, 170.1 and 172.3; mass spectrum (ESI), *m/z* (M+H)<sup>+</sup> 634.2999 (C<sub>29</sub>H<sub>44</sub>N<sub>7</sub>O<sub>7</sub>S requires 634.3017).

**General procedure for the deprotection of Z-protected compounds.** To a solution of Z-protected compounds (**2c $\beta$**  or **2d $\beta$** , 0.10 mmol) in MeOH (3 mL) /THF (3 mL) was added PdCl<sub>2</sub> (20 mg) and the mixture was hydrogenated overnight in a Parr hydrogenator apparatus at 50 psi.<sup>2</sup> The catalyst was filtered and the crude product was purified by semipreparative HPLC leading to the desired compounds **3c $\beta$**  and **3d $\beta$**  in 64% and 58% yield, respectively.

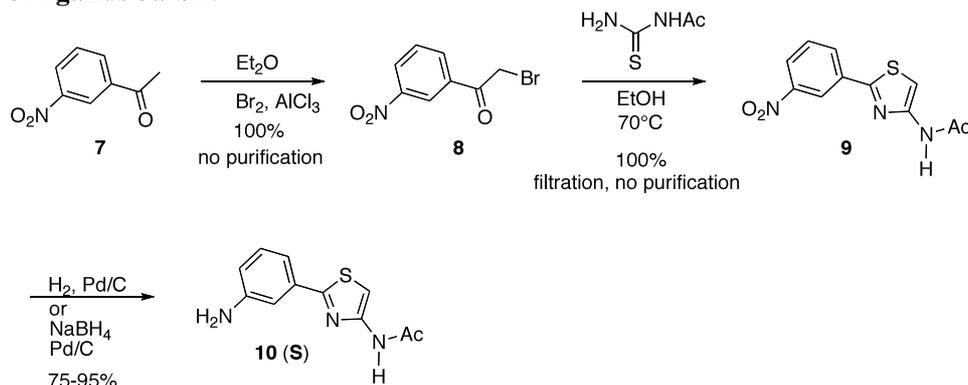
**5'-L-arginyl-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-**

**yl)phenyl]carboxamide (3c $\beta$ ).** Yield 64% (34 mg); HPLC retention time = 20.7 min;  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ )  $\delta$  1.40-1.60 (m, 2H), 1.70-1.90 (m, 2H), 2.20 (s, 3H), 2.30-2.45 (m, 2H), 2.78 (t, 2H,  $J = 6.8$  Hz), 4.01 (t,  $J = 6.6$  Hz, 1H), 4.30-4.50 (m, 3H), 7.28 (s, 1H), 7.40-7.60 (m, 3H) and 7.78 (s, 1H);  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ )  $\delta$  22.4, 23.0, 27.2, 39.9, 42.7, 49.6, 61.9, 71.9, 78.2, 88.0, 105.0, 119.9, 120.5, 121.9, 129.4, 133.2, 137.2, 150.3, 160.1, 163.8, 168.9, 174.2 and 176.6; mass spectrum  $m/z$  556.2 ( $\text{M} + \text{Na}^+$ ) (theoretical 556.1).

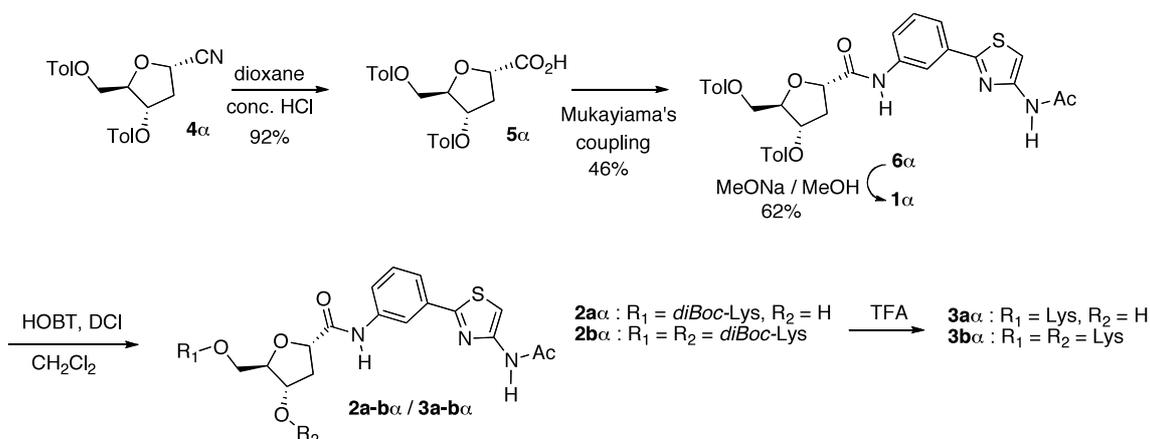
**3',5'-di-L-arginyl-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-**

**yl)phenyl]carboxamide (3d $\beta$ ).** Yield 58% (40 mg); HPLC retention time = 20.9 min;  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ )  $\delta$  1.40-1.60 (m, 4H), 1.60-1.75 (m, 2H), 1.80-1.90 (m, 2H), 2.24 (s, 3H), 2.75 (t,  $J = 6.8$  Hz, 2H), 2.99 (t,  $J = 7.0$  Hz, 2H), 4.00 (t,  $J = 6.2$  Hz, 2H), 4.21 (t,  $J = 6.4$  Hz, 1H), 4.50-4.60 (m, 3H), 5.45-5.50 (m, 1H), 7.35-7.50 (m, 3H), 7.60-7.70 (m, 1H) and 7.91 (s, 1H);  $^{13}\text{C}$  NMR ( $\text{D}_2\text{O}$ )  $\delta$  22.7, 26.7, 29.6, 36.0, 39.5, 53.1, 53.2, 66.3, 78.2, 78.7, 83.2, 110.1, 119.5, 121.9, 124.0, 135.1, 137.0, 148.6, 159.6, 163.5, 169.7, 170.1 and 172.3; mass spectrum  $m/z$  712.3 ( $\text{M} + \text{Na}^+$ ) (theoretical 712.8).

**Synthesis of ligands 3a-b $\alpha$ .**



**Scheme S2a.** Synthesis of S nucleobase



**Scheme S2b.** Synthesis of ligands 3a-b $\alpha$ .

**1-Bromo-3-nitroacetophenone (8).** To a suspension of commercially available 3-nitroacetophenone (4g, 24.2 mmol) in diethyl ether (25 mL) was added 5% of aluminum chloride (160 mg) and the reaction mixture was placed at 0°C. Then, bromine (1 eq., 1.4 mL, 24.2 mmol) was added. After 1h of stirring at room temperature the reaction was complete (verified by NMR), water was added and the mixture was extracted 3 times with diethyl ether. The product was obtained as a yellow solid and employed in the following reaction without further purification: yield 5.83 g (99%); <sup>1</sup>H NMR (DMSO-d<sub>6</sub>) δ: 4.48 (s, 2H), 7.70 (s, 1H), 7.74 (t, 1H, *J* = 8.0 Hz) 8.25-8.35 (m, 1H), 8.45-8.55 (m, 1H) and 8.81 (t, 1H, *J* = 1.8 Hz); <sup>13</sup>C NMR (DMSO-d<sub>6</sub>) δ: 30.2, 123.9, 128.2, 130.3, 134.5, 135.2, 148.5 and 189.5.

**2-*N*-Acetylamino-4-(4-nitrophenyl)thiazole (9).** To a solution of 1-bromo-3-nitroacetophenone (**8**, 5.83 g, 24.0 mmol) in ethanol (130 mL) was added the *N*-acetylurea (1 eq., 2.83 g, 24.0 mmol). The reaction mixture was then stirred at 80 °C 30 min. then left to cool to room temperature. The formed precipitate is filtered and washed with a mixture ethanol/ether leading to compound **9** as a yellow solid: yield 6.2 g (99%); R<sub>f</sub> = 0.5 (cyclohexane/EtOAc 1:1); <sup>1</sup>H NMR (DMSO d<sub>6</sub>) δ: 2.17 (s, 3H), 7.56 (t, *J* = 8.0 Hz, 1H), 7.76 (s, 1H) 7.95-8.05 (m, 1H) 8.18 (d, *J* = 8.0 Hz, 1H) and 8.56 (t, 1H, *J* = 2.0 Hz); <sup>13</sup>C NMR (DMSO d<sub>6</sub>) δ: 22.4, 110.3, 120.0, 122.2, 130.3, 131.6, 135.7, 146.0, 148.3, 158.35 and 168.7.

**2-*N*-Acetylamino-4-(3-aminophenyl)thiazole (10).** To a solution of compound **9** (1.90 g, 7.22 mmol) in a mixture of dichloromethane and methanol (1:1 v/v, 30 mL) was added NaBH<sub>4</sub> (5 eq. 1.37 g, 36.1 mmol) and the reaction mixture was stirred in the presence of palladium on activated carbon (10%) for 3 h. After removal of the catalyst by filtration through a pad of Celite, the filtrate was concentrated under reduced pressure and the product purified on a silica gel column (eluents CH<sub>2</sub>Cl<sub>2</sub>/MeOH 90:10) to give pure compound **10** as a white solid: yield 1.26 g (75%); R<sub>f</sub> = 0.55 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (DMSO d<sub>6</sub>) δ: 2.15 (s, 3H), 5.12 (br s, 2H), 5.26-5.28 (br s, 2H), 6.65-6.75 (m, 2H), 7.20-7.30 (m, 3H) and 7.37 (s, 1H); <sup>13</sup>C NMR (DMSO d<sub>6</sub>) δ: 22.9, 107.3, 111.6, 113.9, 129.5, 135.3, 149.2, 149.9, 157.5, 157.9 and 168.9.

**Synthesis of C-(2'-Deoxy-3',5'-di-*O*-toluoyl- $\alpha$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (6 $\alpha$ ).** To a solution of 3,5-di-*O*-toluoyl- $\alpha$ -1-cyano-2-deoxy-D-ribofuranose (**4 $\alpha$** , 828 mg, 2.18 mmol) in dioxane (5 mL) were added 5 mL of HCl 37% and the reaction was stirred at reflux for 2h. The mixture was cold, neutralized with NaOH 1N (1x 10mL) and extracted with ethyl acetate. The organic phases were dried over MgSO<sub>4</sub> and the solvent evaporated to give compound **5 $\alpha$**  in 92% yield (800 mg). This latter was employed in the following reaction without further purification. Compound **5 $\alpha$**  (700 mg, 1.76 mmol) was dissolved in dry methylene chloride (30 mL) and then triethylamine (2 equiv, 0.491 mL), chloromethylpyridinium iodide (2 equiv, 899 mg) and **S** (1.2 equiv, 492 mg) were added successively. After refluxing for 1 h, the solvent was evaporated and the crude product was extracted with methylene chloride, dried over MgSO<sub>4</sub>, filtered and evaporated. The obtained oil was purified under silica gel chromatography using cyclohexane/ethyl acetate (9/1) as solvents giving 810 mg (46%) of **6 $\alpha$**  as a slightly yellow solid: R<sub>f</sub> 0.30 (1:1 cyclohexane/ethyl acetate); <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 2.18 (s, 3H), 2.21 (s, 3H), 2.43 (s, 3H), 2.75-2.85 (m, 2H), 4.40-4.50 (m, 2H), 4.60-4.70 (m, 1H), 4.80-4.90 (m, 1H), 5.59 (t, 1H, *J* = 3.6 Hz), 6.88 (d, *J* = 8.0 Hz, 2H), 7.08 (s, 1H), 7.20-7.40 (m, 2H), 7.50-7.60 (m, 2H), 7.69 (d, *J* = 8.2 Hz, 2H), 7.90-8.05 (m, 3H) and 8.65 (s, 1H); <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 22.0, 22.1, 23.6, 30.1, 36.1, 64.5, 75.8, 79.4, 84.4, 108.7, 110.0, 118.0, 119.8, 122.5, 127.1, 127.5, 135.5, 138.0, 144.6, 144.7, 149.4, 158.1, 158.4, 166.1, 166.7, 168.4 and 171.1.

**Synthesis of C-(2'-Deoxy- $\alpha$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (1 $\alpha$ ).** To a stirred solution of compound **6 $\alpha$**  (460 mg, 0.750 mmol) in dry MeOH (15 mL) was added MeONa (3 equiv, 122 mg). The reaction mixture was stirred at room temperature for 16 h, and then evaporated under reduced pressure to give a crude oil. Silica gel column chromatography purification using 5% of methanol in methylene chloride afforded **1 $\alpha$**  as a white solid

(175 mg, 62%).  $R_f$  0.25 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (DMSO-d<sub>6</sub>) δ 2.00-2.10 (m, 1H), 2.16 (s, 3H), 2.30-2.40 (m, 1H), 4.00-4.15 (m, 2H), 4.40-4.50 (m, 1H), 4.78 (t, J = 5.2 Hz, 1H), 5.03 (d, J = 3.6 Hz, 1H), 7.30-7.40 (m, 1H), 7.50-7.60 (m, 3H), 8.31 (s, 1H) and 9.60 (s, 1H); <sup>13</sup>C NMR (DMSO-d<sub>6</sub>) δ 22.4, 61.6, 70.9, 77.5, 87.8, 107.9, 117.4, 119.3, 120.9, 128.8, 134.6, 138.8, 148.6, 157.9, 168.6 and 171.8.

**Synthesis of 5'-(N- $\alpha$ , $\omega$ -di-Boc-L-lysyl)-C-(2'-Deoxy- $\alpha$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (2a $\alpha$ ) and 3',5'-di-(N- $\alpha$ , $\omega$ -di-Boc-L-lysyl)-C-(2'-Deoxy- $\alpha$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (2b $\alpha$ ).** By the same procedure as described above for the synthesis of 2a $\beta$  and 2b $\beta$ , 1 $\alpha$  (175 mg) was converted into compounds 2a $\alpha$  and 2b $\alpha$ . 2a $\alpha$ : yield 27% (50 mg);  $R_f$  0.62 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ 1.43 (s, 18H), 1.60-1.85 (m, 2H), 2.20 (s, 3H), 2.60-2.80 (m, 2H), 2.95-3.10 (m, 2H), 3.25-3.30 (m, 2H), 4.00-4.15 (m, 2H), 4.20-4.30 (m, 1H), 4.45-4.55 (m, 1H), 4.85-4.90 (m, 1H), 5.22 (t, 1H, J = 4.8 Hz), 7.15-7.20 (m, 2H), 7.25-7.35 (m, 2H) and 8.24 (s, 1H); <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ 24.0, 24.6, 29.3, 30.6, 31.0, 31.2, 32.6, 32.7, 41.4, 43.1, 55.4, 55.7, 66.1, 78.0, 80.5, 81.0, 81.2, 85.5, 109.6, 120.0, 121.7, 123.9, 124.0, 130.7, 137.2, 139.8, 151.1, 158.6, 158.8, 160.5, 171.2, 171.3, 171.4 and 174.6; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup> 706.55 (calculated 706.31). 2b $\alpha$ : yield 73% (200 mg);  $R_f$  0.73 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ 1.40 (s, 36H), 1.60-1.80 (m, 4H), 2.19 (s, 3H), 2.50-2.80 (m, 4H), 2.95-3.10 (m, 4H), 3.25-3.30 (m, 4H), 4.00-4.15 (m, 2H), 4.20-4.30 (m, 1H), 4.45-4.55 (m, 1H), 4.70-4.80 (m, 1H), 5.15-5.30 (m, 1H), 7.30-7.40 (m, 2H), 7.50-7.70 (m, 2H) and 8.24 (s, 1H); <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ 24.0, 24.6, 29.4, 31.0, 31.2, 32.2, 32.6, 41.4, 55.4, 55.7, 66.0, 78.0, 80.3, 80.5, 81.0, 81.2, 85.5, 109.6, 120.0, 121.6, 124.0, 130.7, 137.2, 139.8, 158.5, 159.0, 160.3, 171.2, 174.3 and 174.6; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup> 1034.30 (calculated 1034.51).

**5'-L-lysyl-C-(2'-Deoxy- $\alpha$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (3a $\alpha$ ).** Yield 87% (35 mg); retention time 19.8 min; <sup>1</sup>H NMR (D<sub>2</sub>O) δ 1.55-1.65 (m, 2H), 1.75-1.85 (m, 2H), 2.05-2.20 (m, 2H), 2.38 (s, 3H), 3.05-3.15 (m, 2H), 3.75-3.85 (m, 2H), 4.05-4.10 (m, 1H), 4.49 (dd, 1H, J = 5.5 and 12 Hz), 5.05-5.10 (m, 1H), 5.55 (d, 1H, J = 5.5 Hz), 7.57 (s, 1H), 7.60-7.65 (m, 2H), and 8.10 (s, 1H); <sup>13</sup>C NMR (D<sub>2</sub>O) δ 22.9, 26.3, 26.8, 42.5, 47.2, 49.4, 53.0, 66.0, 78.3, 79.0, 83.5, 110.2, 115.7, 118.0, 130.5, 135.2, 137.3, 148.7, 159.8, 159.9, 170.2, 172.5 and 173.8; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup> 506.2066 (C<sub>23</sub>H<sub>32</sub>N<sub>5</sub>O<sub>6</sub>S requires 506.2069).

**3',5'-di-L-lysyl-C-(2'-Deoxy- $\beta$ -D-ribofuranosyl-1')-N-[3-(2-acetylaminothiazole-4-yl)phenyl]carboxamide (3b $\alpha$ ).** Yield 85% (89 mg); retention time 16.3 min; <sup>1</sup>H NMR (D<sub>2</sub>O) δ 1.35 (t, 4H, J = 7.5 Hz), 1.55-1.65 (m, 4H), 1.75-1.85 (m, 4H), 2.05-2.20 (m, 2H), 2.38 (s, 3H), 3.05-3.15 (m, 2H), 3.28 (q, 4H, J = 7 Hz), 3.75-3.85 (m, 2H), 4.05-4.10 (m, 1H), 4.36 (t, 1H, J = 6.5 Hz), 4.49 (dd, 1H, J = 5.5 and 12 Hz), 5.05-5.10 (m, 1H), 5.55 (d, 1H, J = 5.5 Hz), 7.57 (s, 1H), 7.60-7.65 (m, 2H), and 8.10 (s, 1H); <sup>13</sup>C NMR (D<sub>2</sub>O) δ 22.7, 26.3, 26.8, 35.7, 39.5, 42.5, 47.2, 49.4, 53.0, 53.1, 66.1, 78.3, 79.0, 83.5, 110.2, 113.3, 115.7, 118.0, 119.4, 120.3, 121.9, 123.8, 130.5, 135.2, 137.3, 148.7, 159.8, 159.9, 163.2, 163.5, 169.8, 170.2, 172.5 and 173.8; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup> 634.3016 (C<sub>29</sub>H<sub>44</sub>N<sub>7</sub>O<sub>7</sub>S requires 634.3017).

**Synthesis of 5'-(N- $\alpha$ , $\omega$ -di-Boc-L-lysyl)-2'-Deoxyuridine (2e) and 3',5'-di-(N- $\alpha$ , $\omega$ -di-Boc-L-lysyl)-2'-Deoxyuridine (2f).** To a solution of commercially available N- $\alpha$ , $\omega$ -di-Boc-L-lysine (1.62 g, 3.07 mmol) in anhydrous dichloromethane (15 mL) were added successively HOBt (415 mg, 3.07 mmol), DIC (475  $\mu$ L, 3.07 mmol), triethylamine (428  $\mu$ L, 3.07 mmol) and DMAP (15 mg, 0.123 mmol). Finally, a solution of 2'-deoxyuridine (300 mg, 1.23 mmol) in dichloromethane (5 mL) was added under argon atmosphere. The mixture was stirred for 4 h at room temperature then washed twice with water (2 x 15 mL) and dried over MgSO<sub>4</sub>. Concentration under reduced pressure followed by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 95:5) led to the desired compounds 2e and 2f. 2e: yield 27% (185 mg); <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.42 (s, 18H), 1.60-1.80 (m, 4H), 2.35-2.50 (m, 2H), 3.00-3.50 (m, 2H), 3.85-3.95 (m, 1H), 4.30-4.50 (m, 3H), 4.70-4.80 (m, 1H), 5.83 (d, J = 8.0 Hz, 1H), 6.24 (t, J = 7.2 Hz, 1H) and 7.82 (d, J = 8.0 Hz, 1H); <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 22.9, 28.8, 30.1, 30.8, 40.1, 40.5, 54.0, 64.6, 70.7, 79.7, 85.7, 93.4, 103.8, 140.0, 150.6, 156.0, 163.5 and 173.1; mass spectrum (ESI),  $m/z$  (M+H)<sup>+</sup>

557.2816 ( $C_{25}H_{41}N_8O_{10}S$  requires 557.2817). **2f**: yield 70% (762 mg);  $^1H$  NMR ( $CDCl_3$ )  $\delta$  1.42 (m, 36H), 1.60-1.85 (m, 4H), 2.20-2.40 (m, 1H), 2.50-2.60 (m, 1H), 3.00-3.20 (m, 4H), 3.70-3.90 (m, 1H), 4.10-4.20 (m, 4H), 4.30-4.50 (m, 2H), 4.70-4.80 (m, 1H), 5.85 (d,  $J = 8.0$  Hz, 1H), 6.23 (t,  $J = 5.8$  Hz, 1H), 7.54 (d,  $J = 8.0$  Hz, 1H);  $^{13}C$  NMR ( $CDCl_3$ )  $\delta$  22.8, 23.0, 23.9, 28.8, 30.1, 31.8, 37.7, 40.1, 42.5, 54.1, 64.7, 79.7, 80.6, 82.6, 85.7, 103.8, 140.0, 150.6, 156.0, 156.7, 157.4, 158.1, 163.5 and 173.1; mass spectrum (ESI),  $m/z$  ( $M+H$ ) $^+$  885.4812 ( $C_{41}H_{69}N_6O_{15}S$  requires 885.4815).

**5'-L-lysyl-2'-deoxyuridine (3e)**. Yield 89%; retention time 21.4 min;  $^1H$  NMR ( $D_2O$ )  $\delta$  1.45-1.60 (m, 2H), 1.70-1.80 (m, 2H), 1.95-2.10 (m, 2H), 2.40-2.50 (m, 2H), 2.95-3.10 (m, 2H), 3.80-3.90 (m, 2H), 4.09 (t, 1H,  $J = 6.2$  Hz), 4.20-4.30 (m, 1H), 6.10 (d,  $J = 8.0$  Hz, 1H), 6.25 (t,  $J = 6.8$  Hz, 1H) and 7.70 (d,  $J = 8.0$  Hz, 1H);  $^{13}C$  NMR ( $D_2O$ )  $\delta$  22.6, 26.7, 29.5, 29.7, 39.3, 39.4, 58.0, 61.6, 70.8, 77.3, 85.9, 86.8, 102.6, 142.1, 151.8, 155.9, 166.6 and 172.4; mass spectrum (ESI),  $m/z$  ( $M+H$ ) $^+$  357.1768 ( $C_{15}H_{25}N_4O_6$  requires 357.1774).

**3',5'-di-L-lysyl-2'-deoxyuridine (3f)**. Yield 74%; retention time 18.2 min;  $^1H$  NMR ( $D_2O$ )  $\delta$  1.55-1.70 (m, 4H), 1.75-1.90 (m, 4H), 2.00-2.20 (m, 4H), 2.70-2.80 (m, 2H), 3.05-3.15 (m, 4H), 4.30-4.40 (m, 3H), 4.60-4.70 (m, 3H), 5.60 (br s, 1H), 5.99 (d,  $J = 8.0$  Hz, 1H), 6.32 (t, 1H,  $J = 7.0$  Hz) and 7.80 (d,  $J = 8.0$  Hz, 1H);  $^{13}C$  NMR ( $D_2O$ )  $\delta$  21.9, 22.1, 24.4, 26.7, 29.6, 35.8, 39.5, 53.1, 53.8, 65.8, 76.3, 81.6, 87.2, 87.3, 102.9, 142.7, 151.8, 166.6, 169.8 and 170.1; mass spectrum (ESI),  $m/z$  ( $M+H$ ) $^+$  485.2716 ( $C_{21}H_{37}N_6O_7$  requires 485.2718).

### Binding studies

Unless otherwise stated, all reagents and solvents were of analytical grade and from Sigma (St Louis, U.S.A.). HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] and all inorganic salts for buffers were purchased from Calbiochem (molecular biology grade). RNA (figure S1) and DNA oligonucleotides were purchased from IBA GmbH and used without further purification. A mixture of pre- and mature yeast tRNAs (containing over 30 different species) was purchased from Sigma (type X-SA). Stocks of tRNAmix can be quantified in its native form (without base hydrolysis) using an extinction coefficient of  $9640\text{ cm}^{-1}\text{ M}^{-1}$  per base.<sup>3</sup>

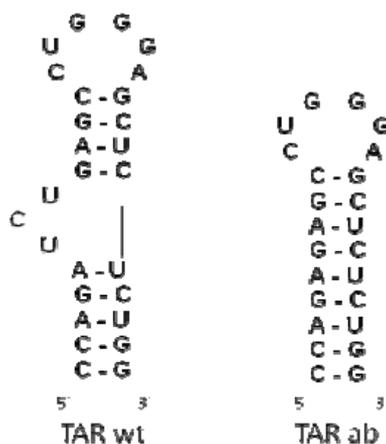


Figure S1. RNA structures

### Buffers

All buffers were filtered through 0.22- $\mu\text{m}$  Millipore filters (GP ExpressPLUS membrane). A small aliquot (50–100 ml) was first filtered and then discarded to avoid any contaminants that might be leached from the filter. The solutions to be used in the fluorescence experiments were prepared by diluting the concentrated stocks in Milli-Q water and filtered again as described above. All standard fluorescence measurements were performed in buffer A (20 mM HEPES (pH 7.4 at 25  $^{\circ}\text{C}$ ), 20 mM NaCl, 140 mM KCl and 3 mM  $\text{MgCl}_2$ ). For competitive experiments in the presence of a hairpin DNA, a 27-mer sequence (5'-CCAGATCTGAGCCTGGGAGCTCTCTGG-3'), annealed beforehand,

were added to buffer A to obtain a 10-fold nucleotide excess regarding TAR RNA (50 nM duplex; 5 nM RNA). For competitive experiments in the presence of a *t*RNA, a mixture of pre- and mature yeast *t*RNAs (containing over 30 different species from baker's yeast (*S. cerevisiae*, Sigma, type X-SA)) was added to buffer A to obtain a 100-fold nucleotide excess regarding TAR RNA. Stock solutions of *t*RNA were prepared in water and quantified using an extinction coefficient of  $9640 \text{ cm}^{-1} \text{ M}^{-1}$  per base.<sup>3</sup>

#### Fluorescence binding assay.<sup>4</sup>

Ligand solutions were prepared as serial dilutions in buffer A at a concentration four times higher than the desired final concentration to allow for the subsequent dilution during the addition of the RNA solution. Pipetting automated system Eppendorf epmotion® 5075 was used in order to perform these analysis. The appropriate ligand solution (50  $\mu\text{L}$ ) was then added to a well of a non-treated black 96-well plate (Nunc 237105), in duplicate. Refolding of the RNA was performed using a thermocycler (ThermoStat Plus Eppendorf) as follows: the RNA, diluted in 1mL of buffer A, was first denatured by heating to  $90^\circ\text{C}$  for 2 min; then cooled to  $4^\circ\text{C}$  for 10 min followed by incubation at  $20^\circ\text{C}$  for 15 min. After refolding, the RNA was diluted to a working concentration of 10 nM through addition of the appropriate amount of buffer A. The tube was mixed and 50  $\mu\text{L}$  of the RNA solution was added to each well containing ligand. This subsequent dilution lowered the final RNA concentration to 5 nM. The fluorescence was measured on a GeniosPro (Tecan) with an excitation filter of  $485 \pm 10 \text{ nm}$  and an emission filter of  $535 \pm 15 \text{ nm}$ . Each point was measured 10 times with a 500  $\mu\text{s}$  integration time and averaged. Binding was allowed to proceed overnight at  $5^\circ\text{C}$  to achieve equilibrium. To study the temperature dependence, the plates were incubated after overnight equilibrium at  $5^\circ\text{C}$  and  $25^\circ\text{C}$  (figure S3).

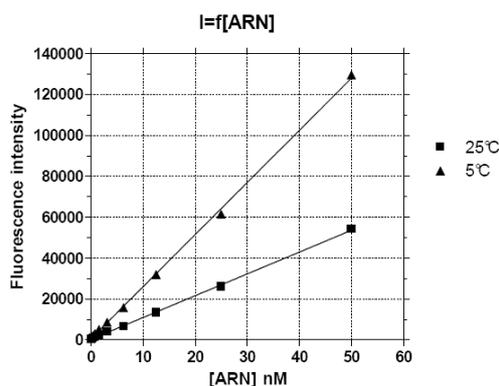


Figure S3. Fluorescence Intensity of TAR RNA

Neomycin was taken as a control as its binding to TAR has already been studied using several methods. Its  $K_d$  value ( $21 \pm 2 \mu\text{M}$ ) is in good agreement with previously reported values.<sup>5</sup> To check that the fluorescence quenching was not a result of nonspecific interactions between our ligands and the fluorophore, titrations of a 5 nM solution of A488 hexyl-ester derivative with **3b $\beta$**  ranging from 0.01 to 1500  $\mu\text{M}$  were carried out. No significant changes of the fluorescence intensity was observed (less than 10 %, data not shown), thus confirming that the observed fluorescence decreases very likely reflect the binding of our ligands (**3a-d $\beta$** , **3a-b $\alpha$** ) to the labeled TAR fragment rather than nonspecific interactions with the fluorophore.

#### Data analysis

Binding data were analyzed using the non-linear least-squares numerical solver-based binding data global analysis program BIOEQS, in which the calculated binding surface is obtained using a numerical constrained optimization chemical equilibrium solver.<sup>6</sup> Unless otherwise stated, binding profiles were well modeled using a simple model assuming the one to one stoichiometry. A higher

initial fluorescence value is observed in the presence of *dsDNA* and *tRNA*, which is consistent with the modification of the polarity of the solvent and a small fluorescence of the *tRNA* mixture. Curves for all compounds measured at 5°C are shown in figures S4, S5, S6 and S7.

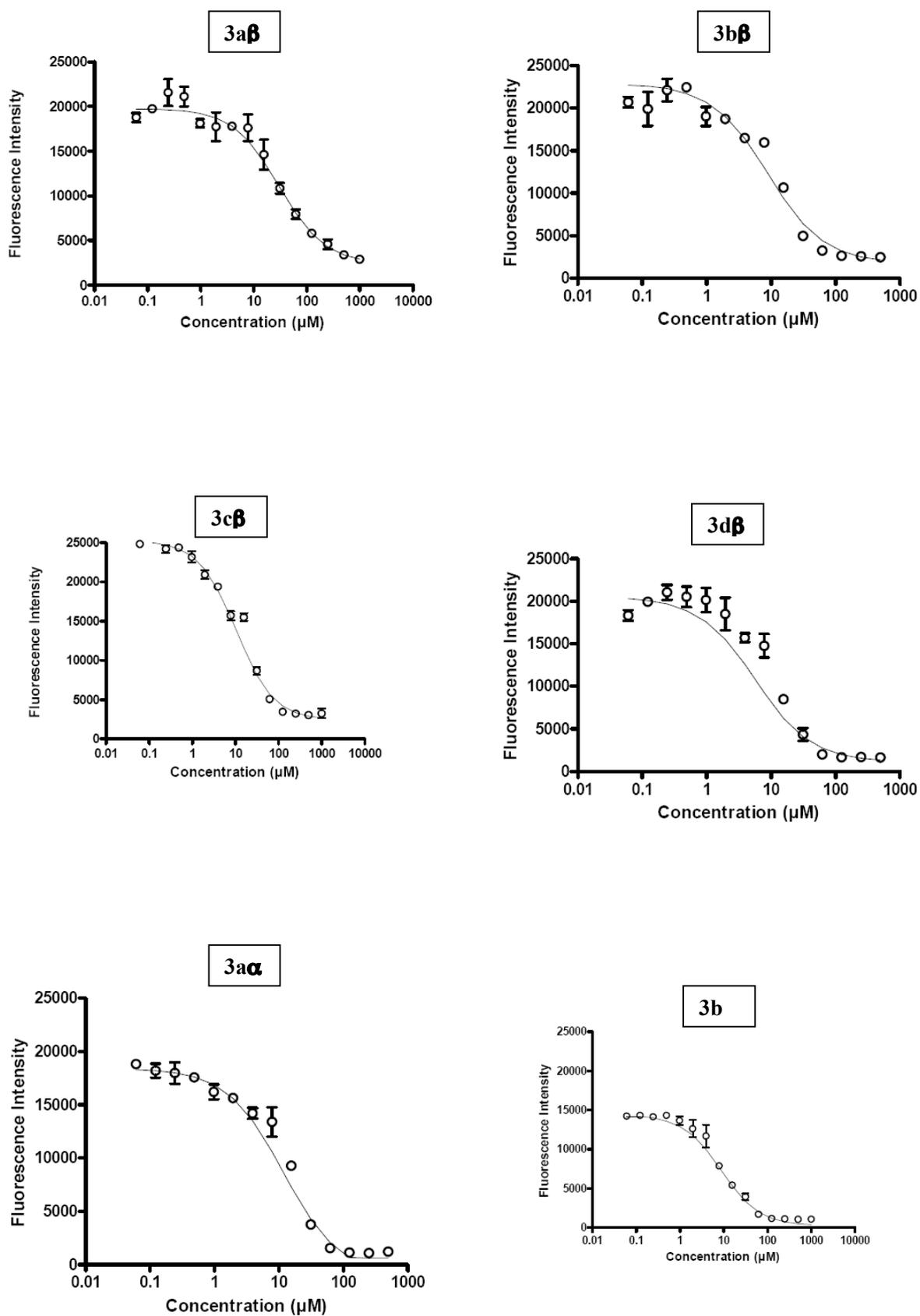


Figure S4. Binding curves of compounds **3aβ-3dβ** and **3a-3bα** with TAR RNA without competitor.

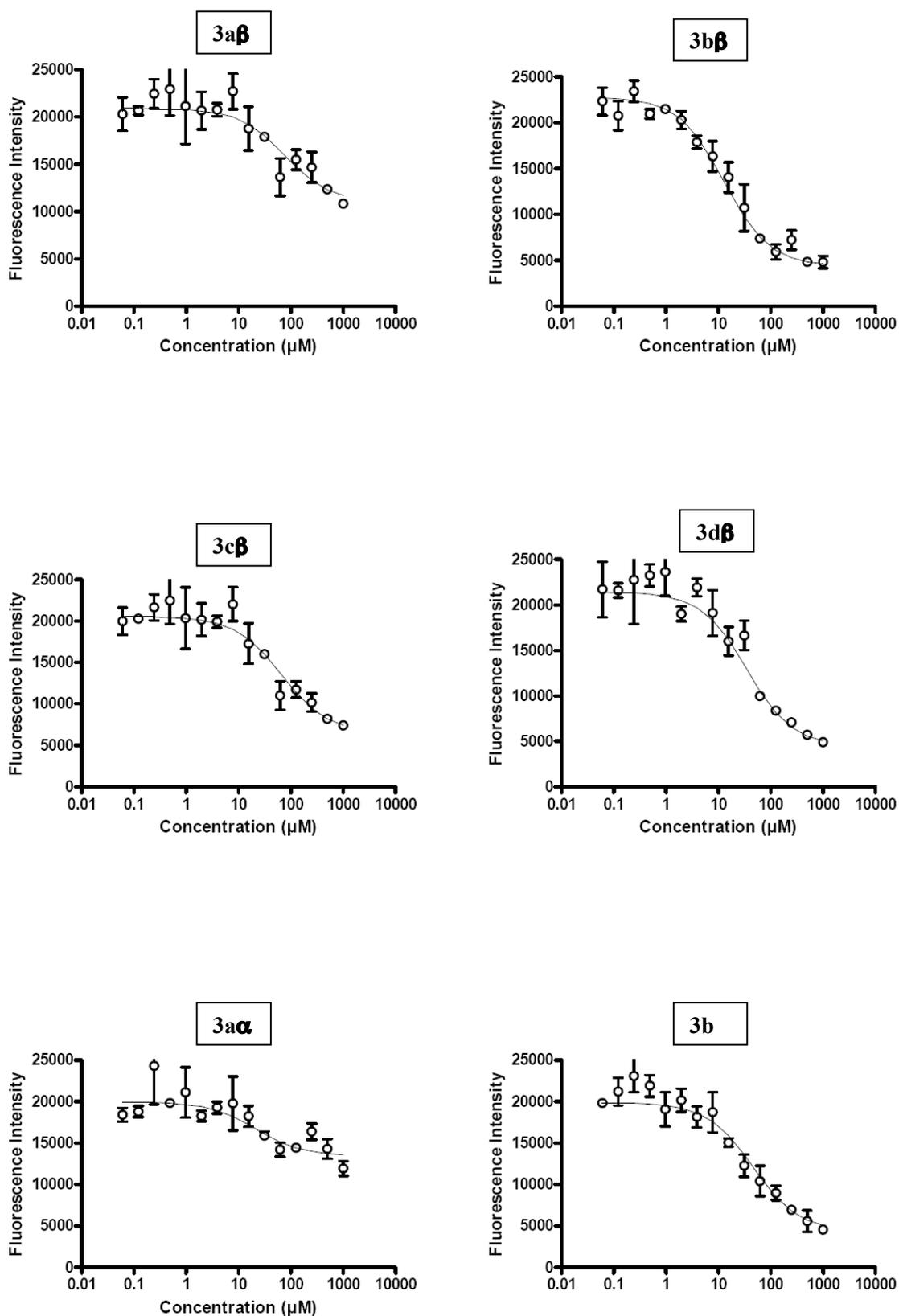


Figure S5. Binding curves of compounds **3a $\beta$ -3d $\beta$**  and **3a $\alpha$ -3b $\alpha$**  with TAR RNA in the presence of 100 eq. of tRNA.

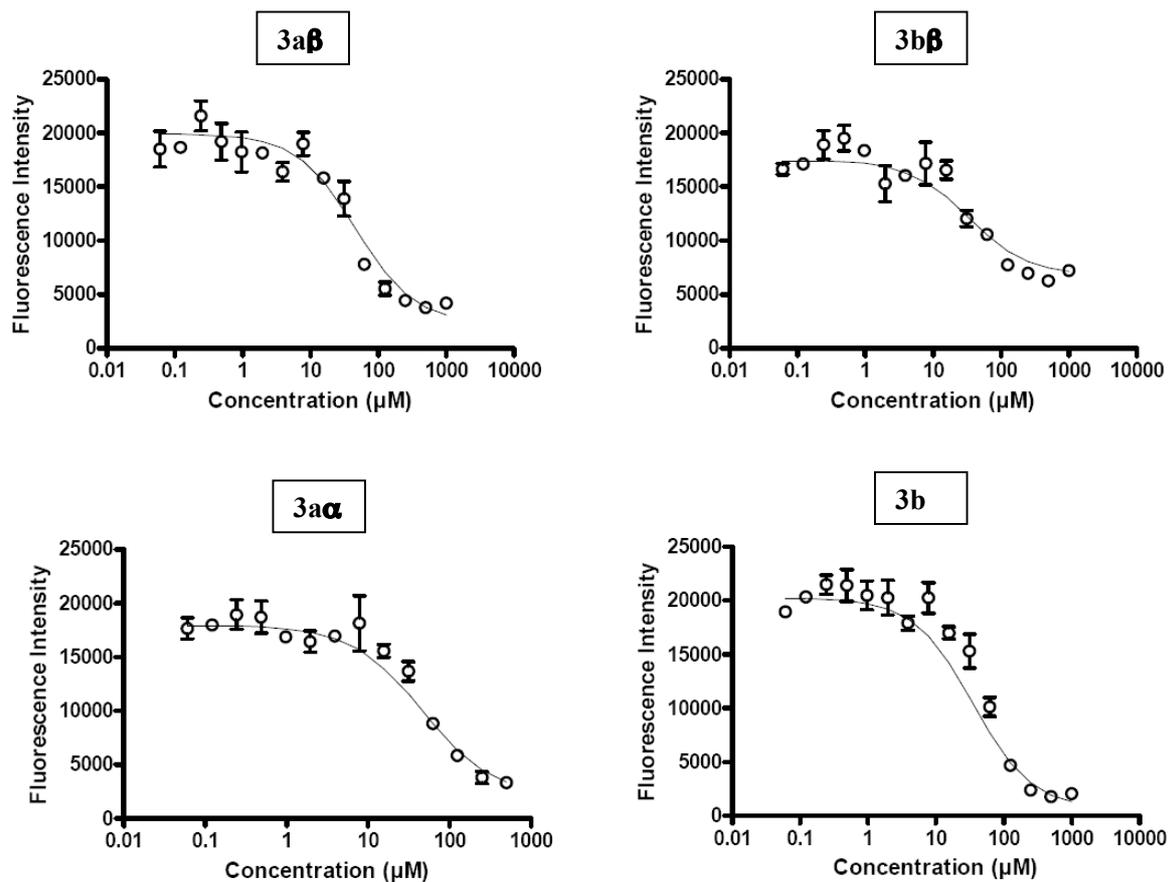


Figure S6. Binding curves of compounds **3aβ-3bβ** and **3aα-3bα** with TAR RNA in the presence of 10 eq. of DNA.

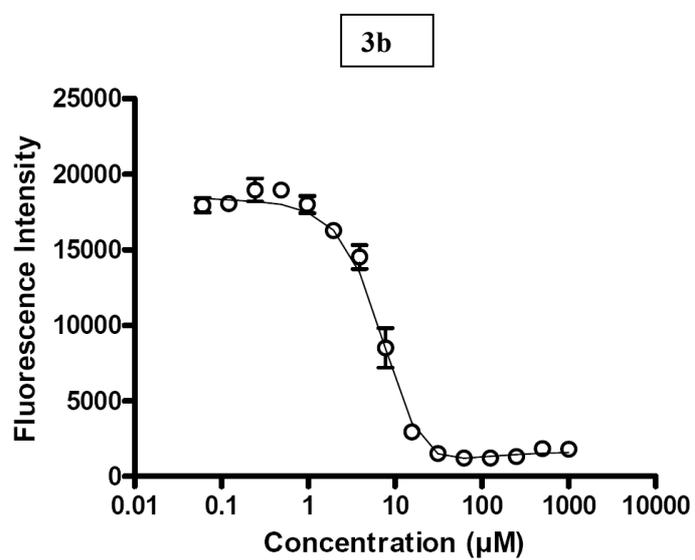
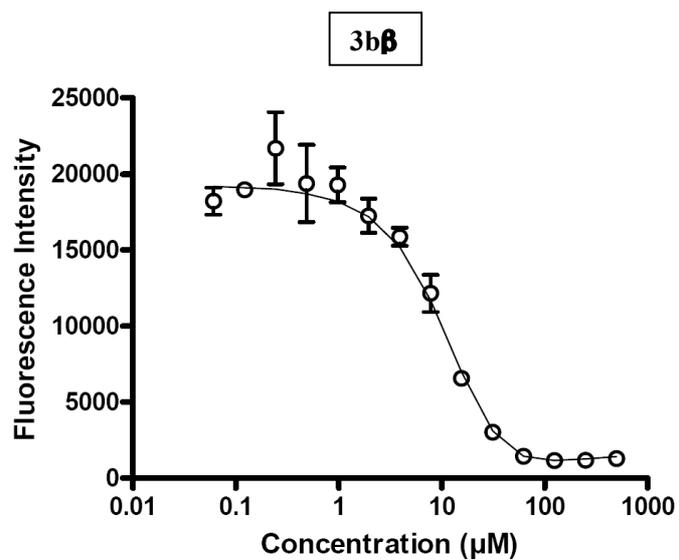
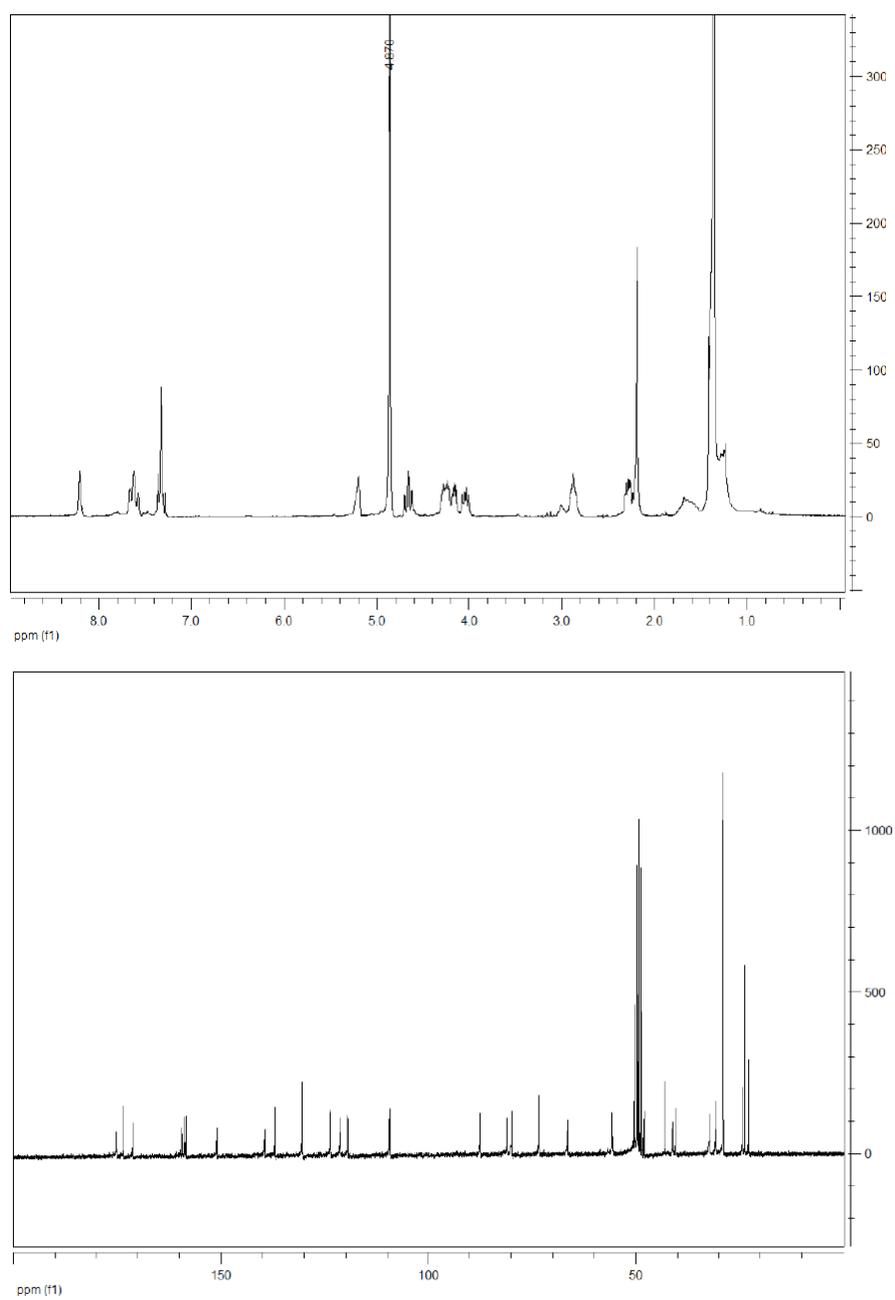


Figure S7. Binding curves of compounds **3b $\beta$**  and **3b $\alpha$**  with TARab RNA.

Figure S8.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (200 MHz) and HRMS spectra for compound **2a $\beta$** .



DUCA\_100211144704 #3 RT: 0.08 AV: 1 NL: 1.35E5  
T: FTMS + p ESI SIM ms [701.00-711.00]

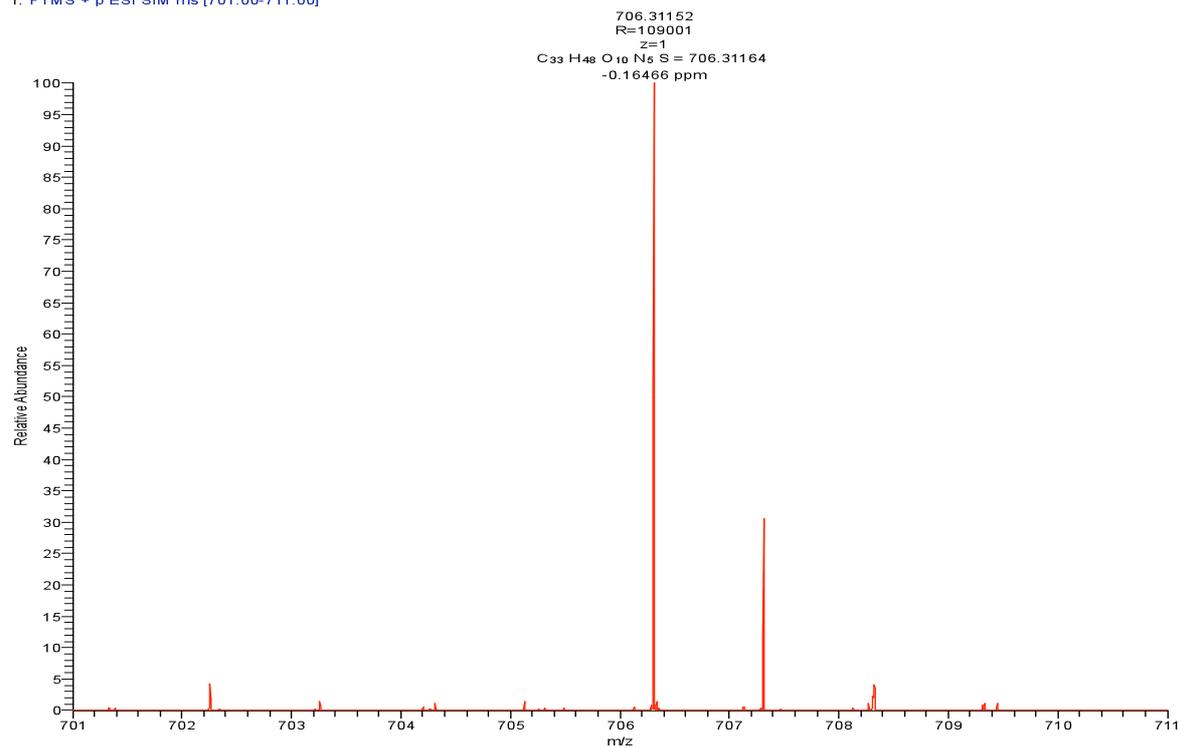
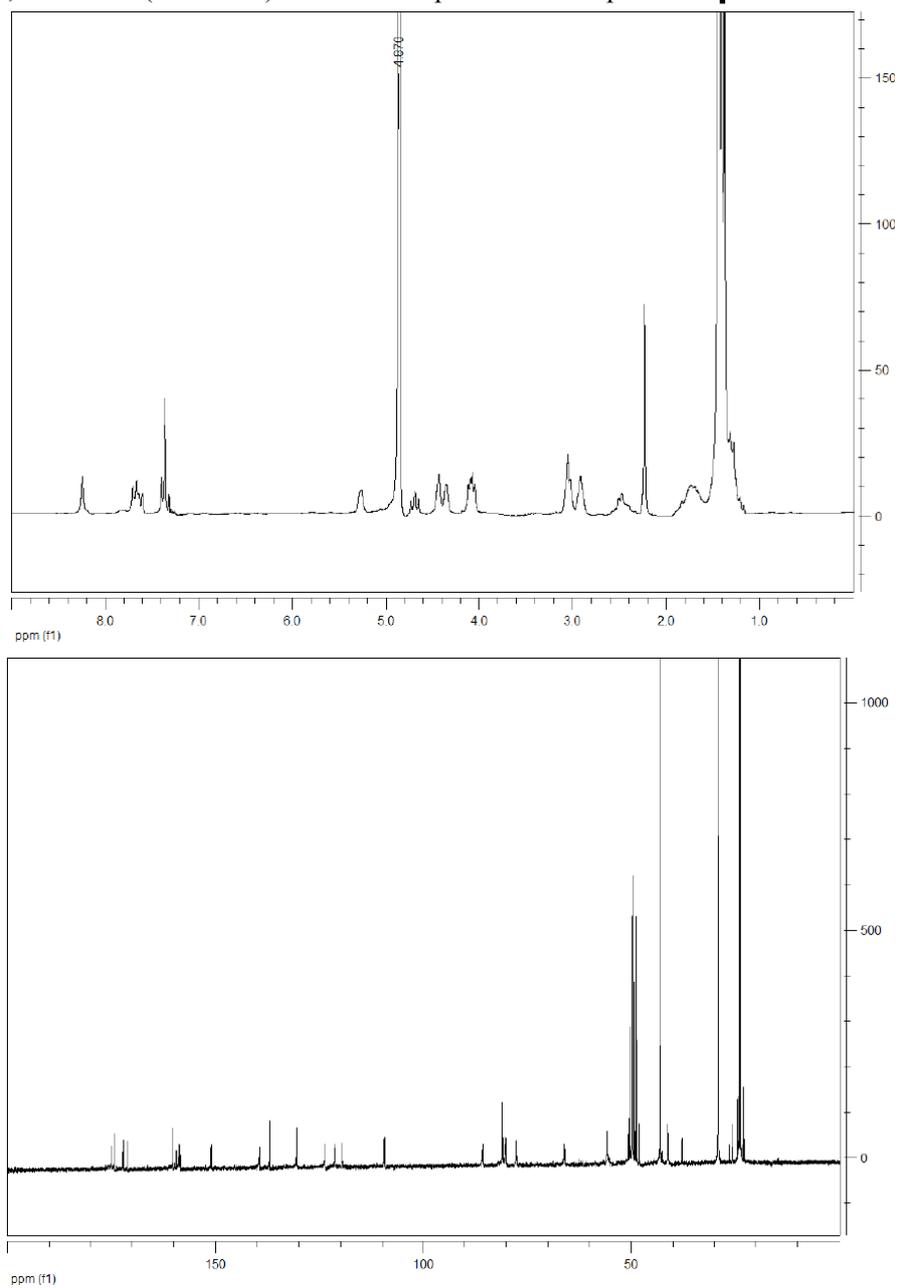


Figure S9.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (200 MHz) and HRMS spectra for compound **2b $\beta$** .



DUCA\_100211143120 #1 RT: 0.01 AV: 1 NL: 3.24E6  
T: FTMS + p ESI SIM ms [1029.00-1039.00]

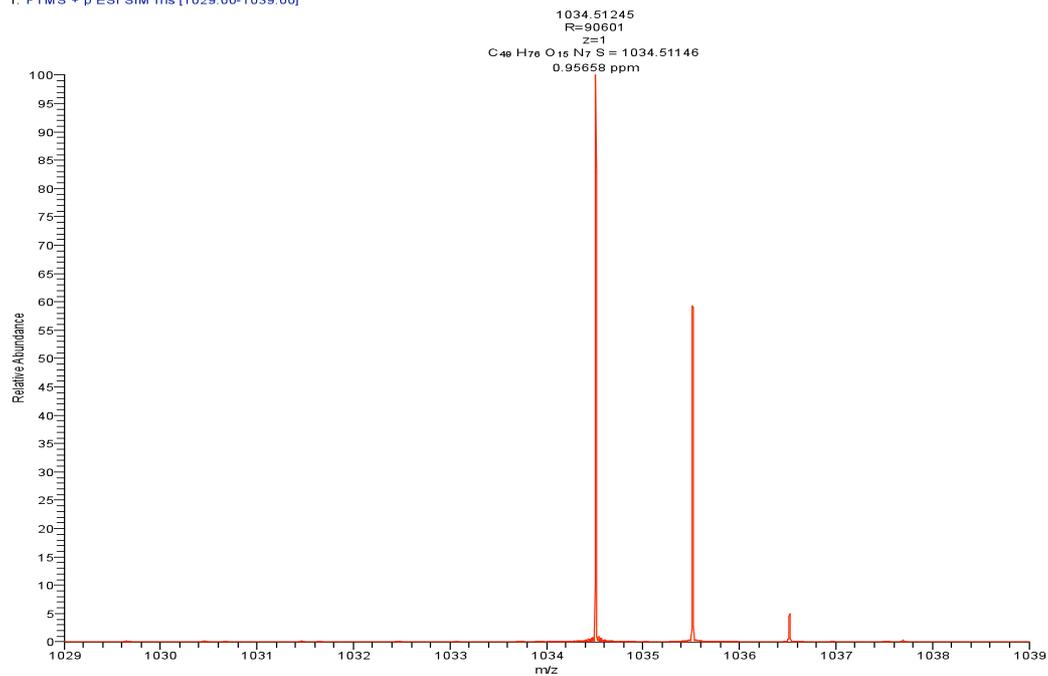
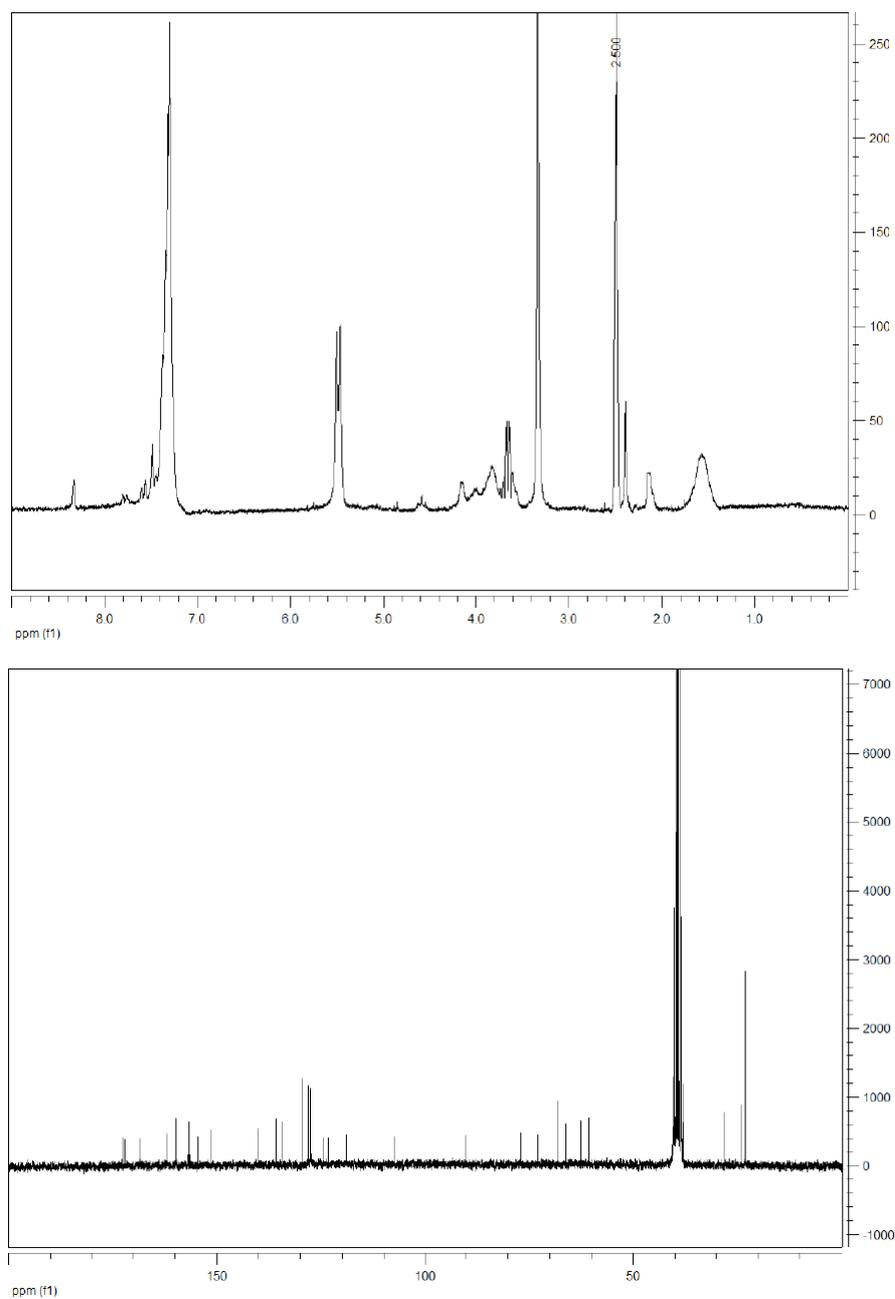


Figure S10.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (200 MHz) and HRMS spectra for compound **2c $\beta$** .



DUCA\_100211150036 #2 RT: 0.04 AV: 1 NL: 3.85E5  
T: FTMS - p ESI SIM ms [931.00-941.00]

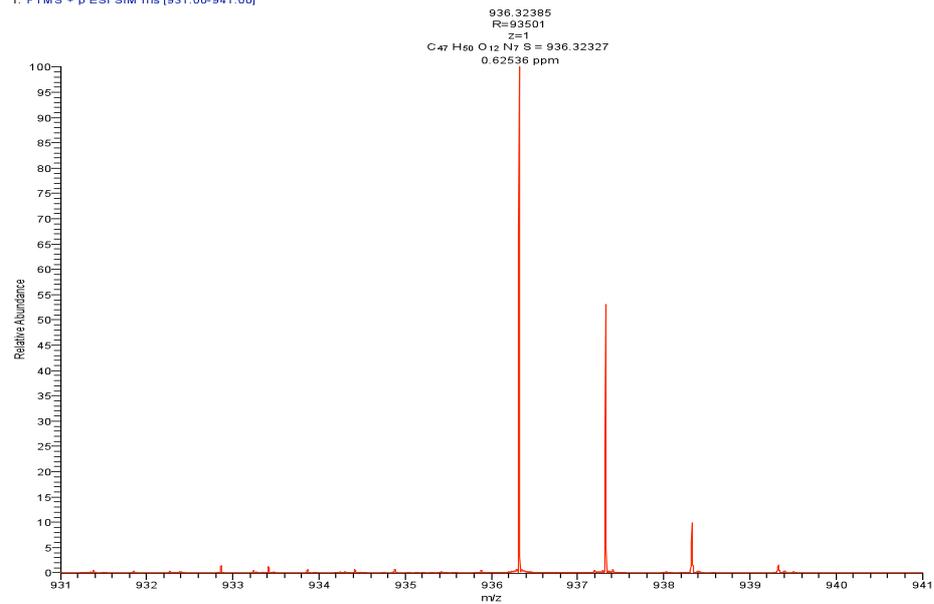
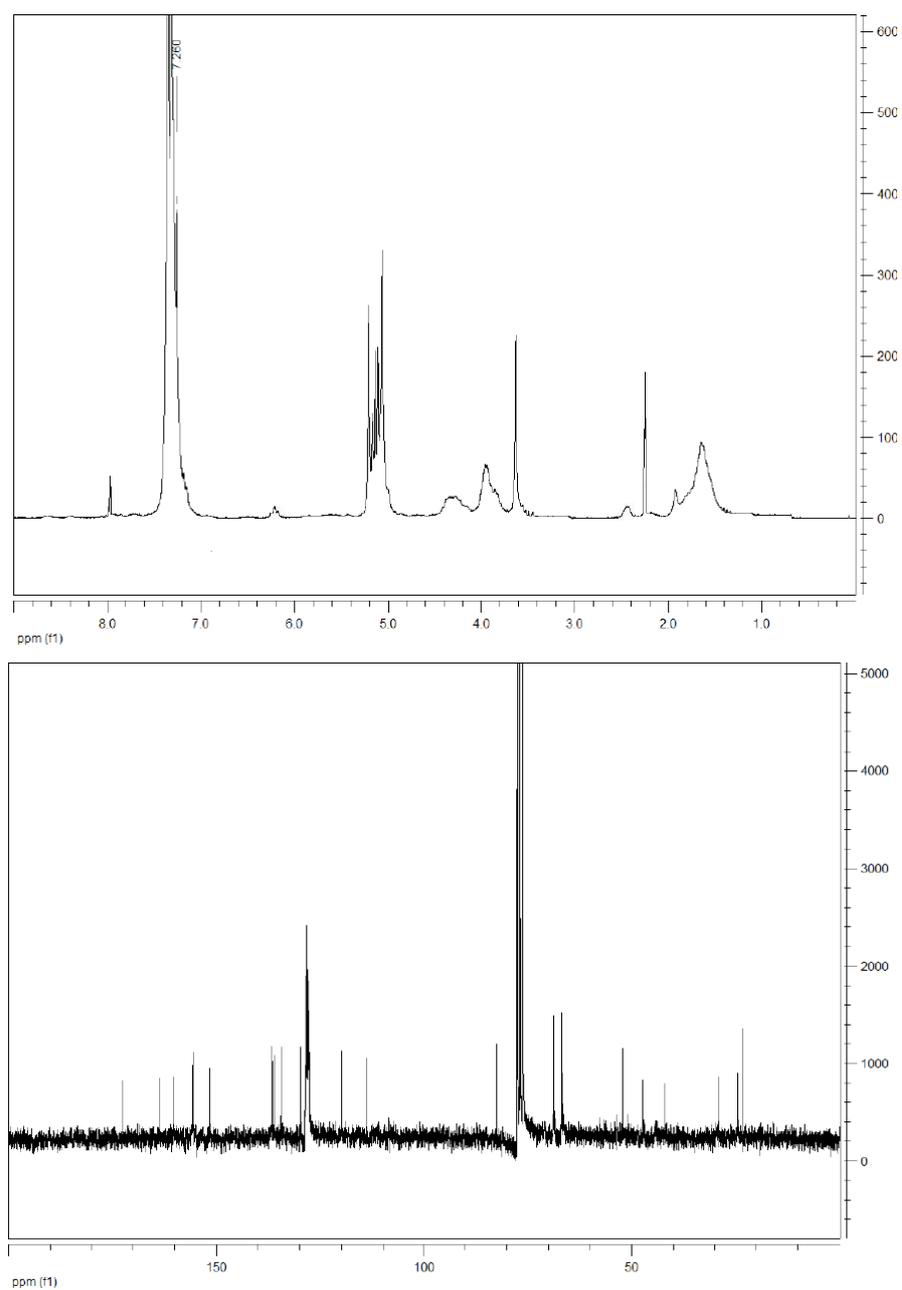


Figure S11.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (200 MHz) and HRMS spectra for compound **2d $\beta$** .



DUCA\_100211164952 #1 RT: 0.02 AV: 1 SM: 76 NL: 1.01E4  
T: FTMS - p ESI sid=35.00 SIM ms [1489.00-1499.00]  
1494.53479  
R=52000  
z=1

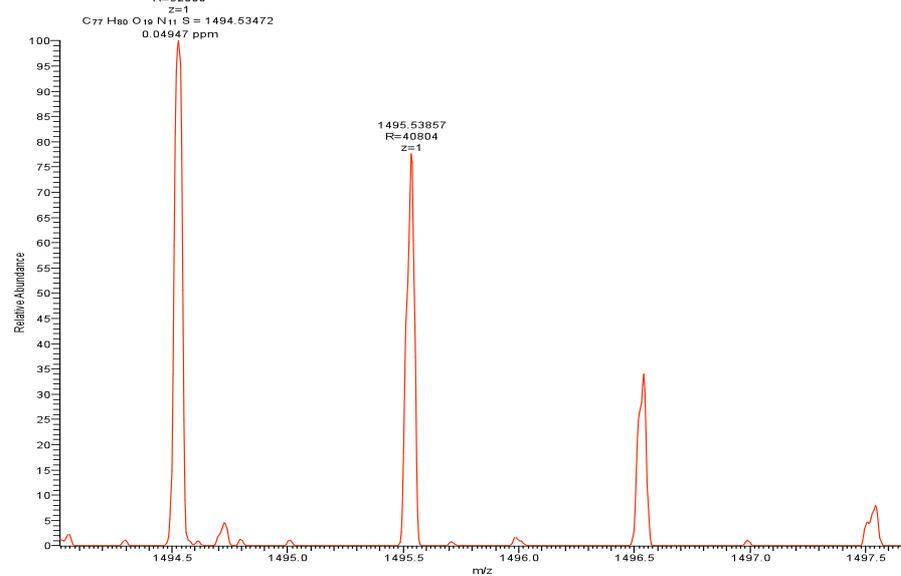
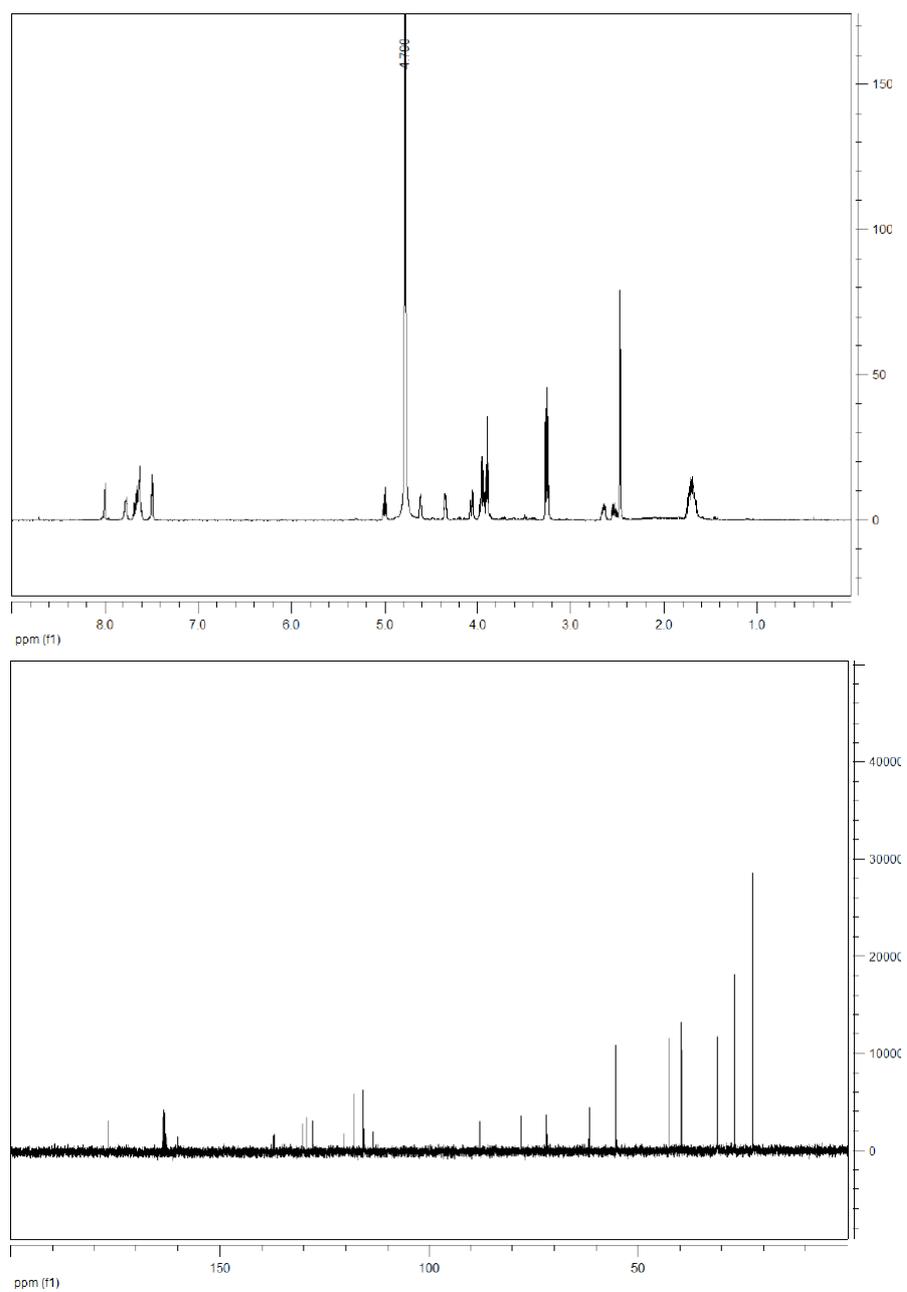


Figure S12.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (500 MHz) and HRMS spectra for compound **3a $\beta$** .



DUCA\_100211141836 #1 RT: 0.02 AV: 1 NL: 3.73E7  
T: FTMS + p ESI SIM ms [501.00-511.00]

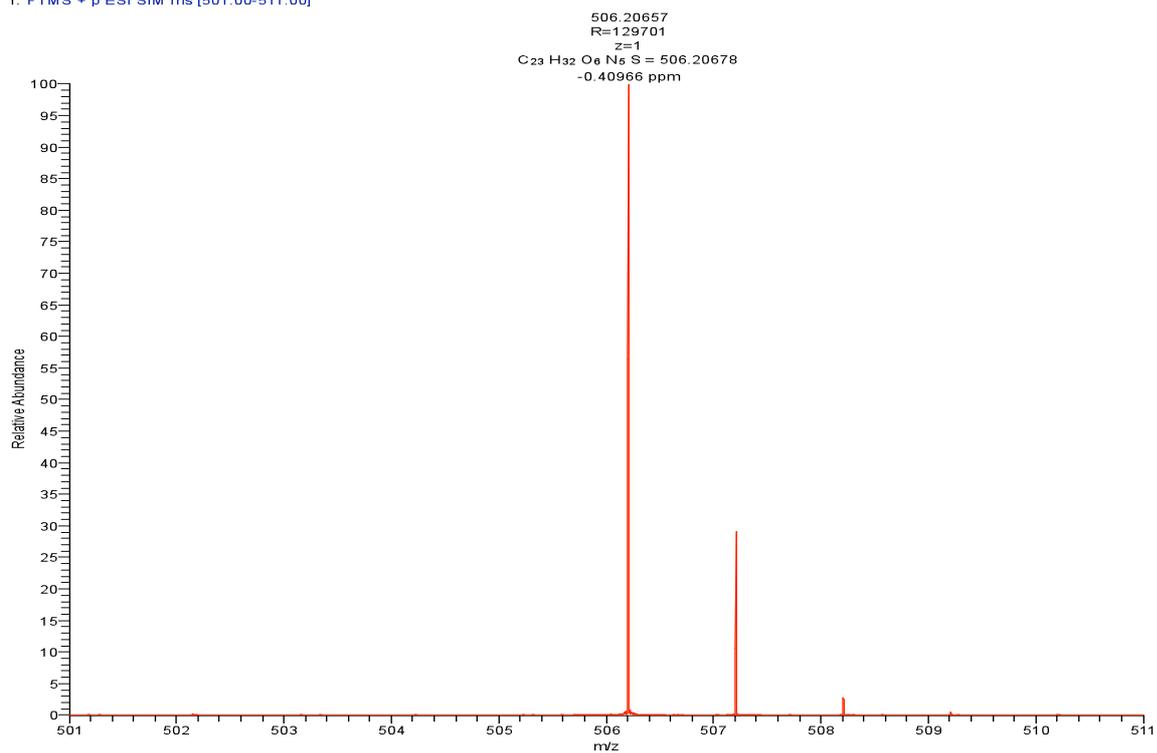
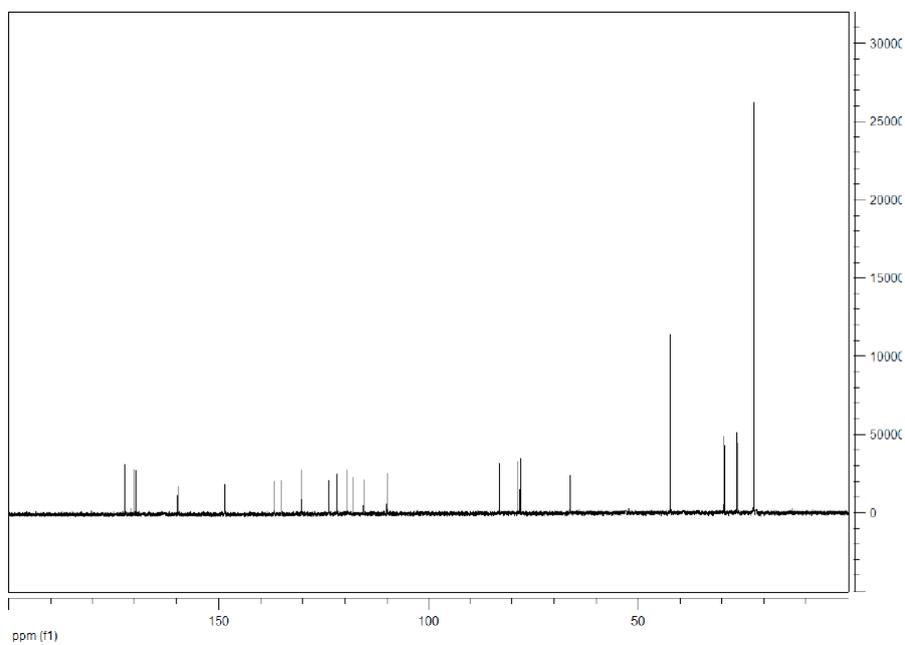
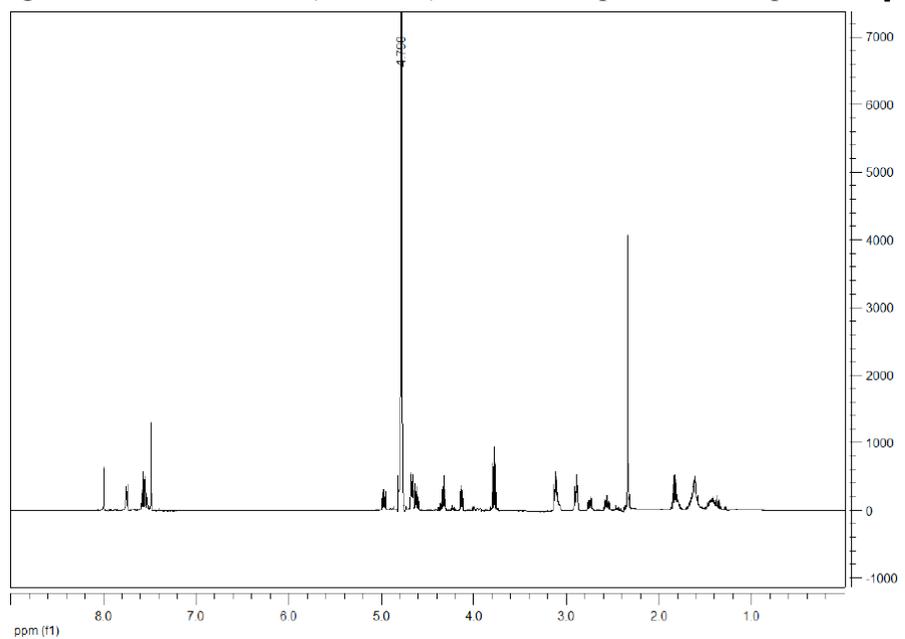


Figure S13.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (500 MHz) and HRMS spectra for compound **3b $\beta$** .



DUCA\_100211122020 #1 RT: 0.01 AV: 1 NL: 2.71E8  
T: FTMS + p ESI SIM ms [629.00-639.00]

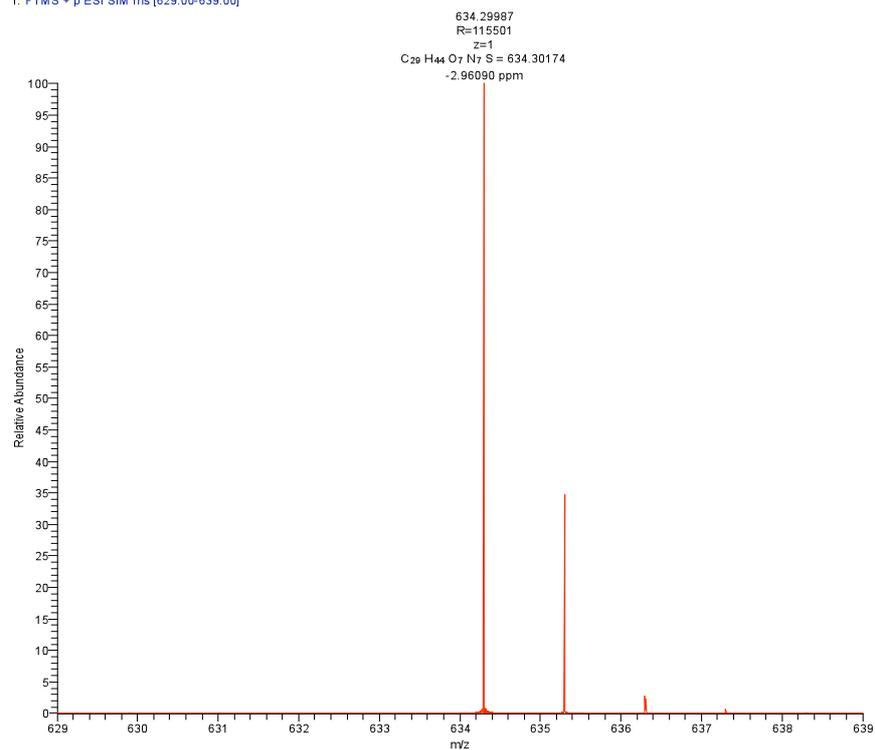
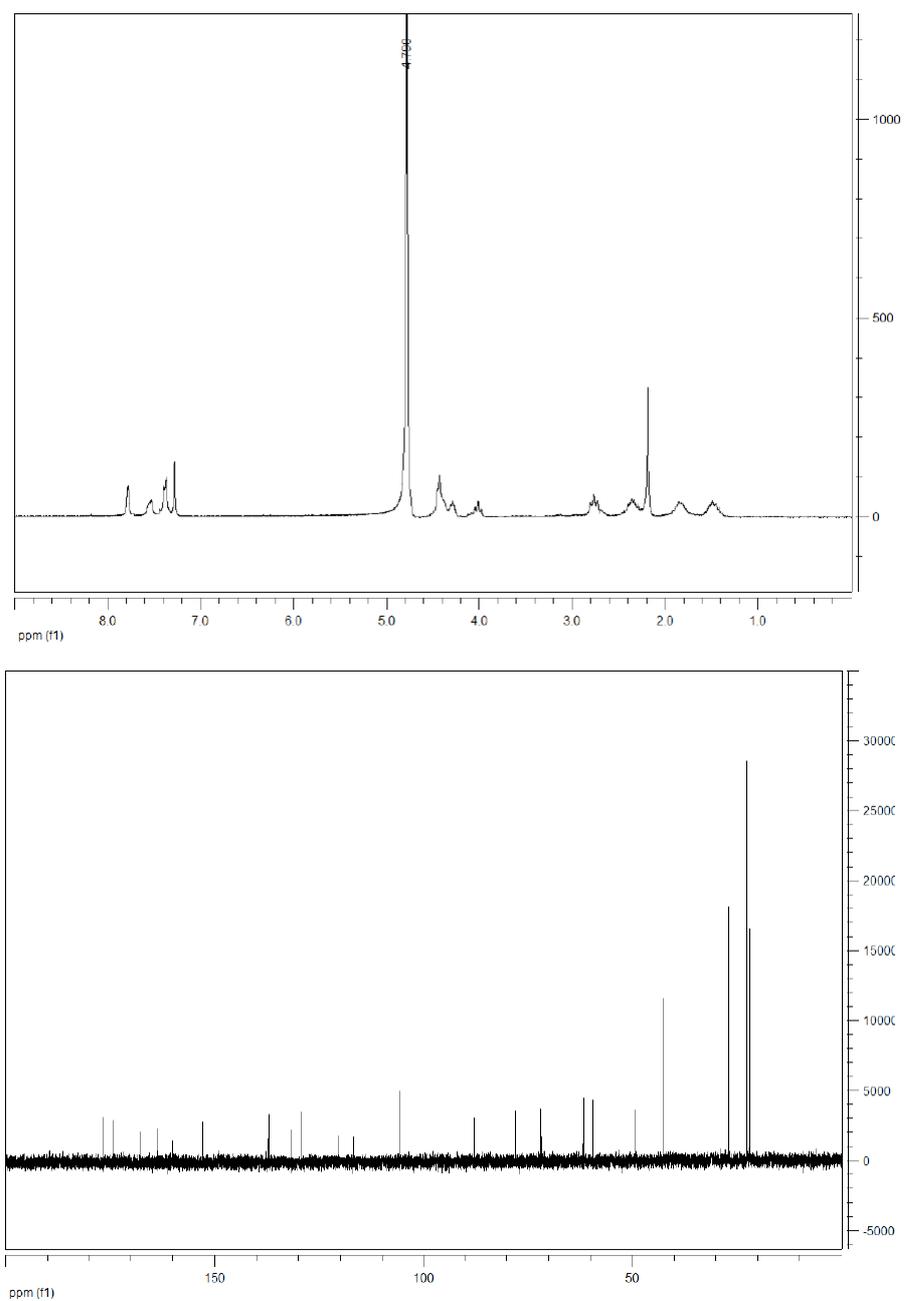


Figure S14.  $^1\text{H}$  and  $^{13}\text{C}$  NMR (500 MHz) and MS spectra for compound **3c $\beta$** .



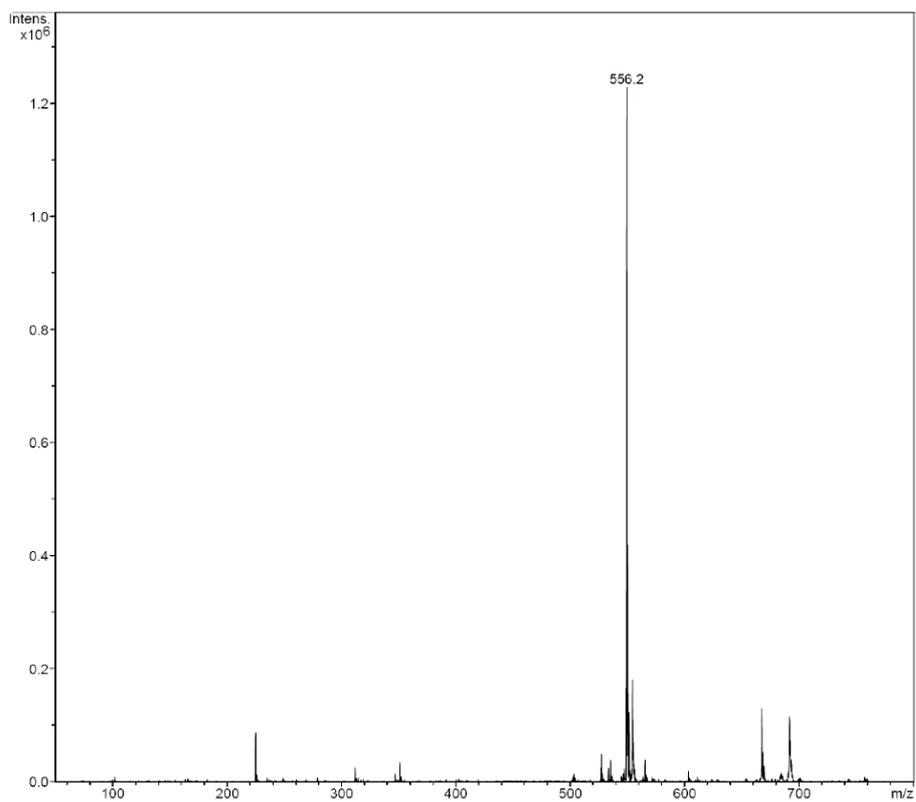
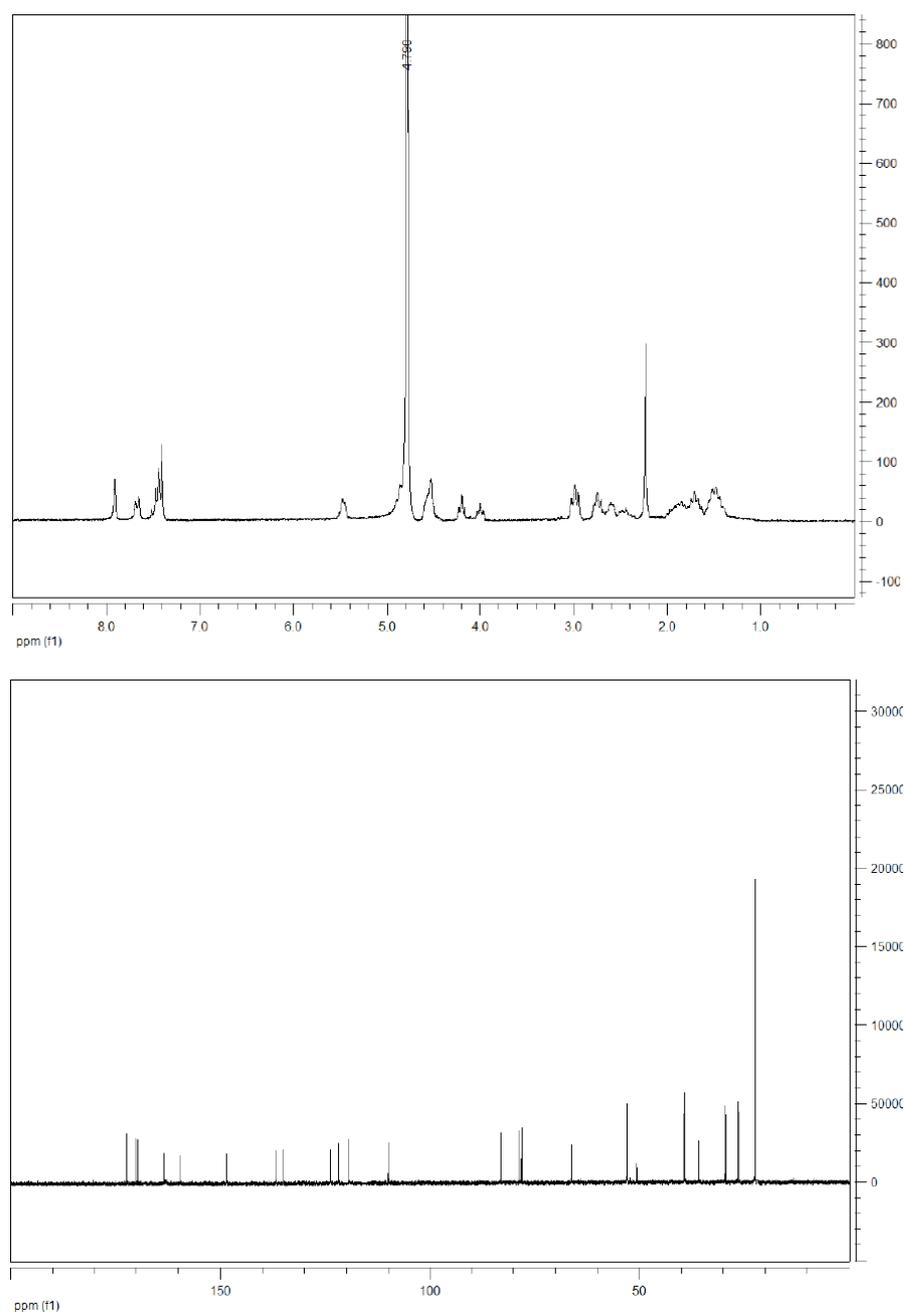


Figure S15.  $^1\text{H}$  and  $^{13}\text{C}$  NMR (500 MHz) spectra for compound **3d $\beta$** .



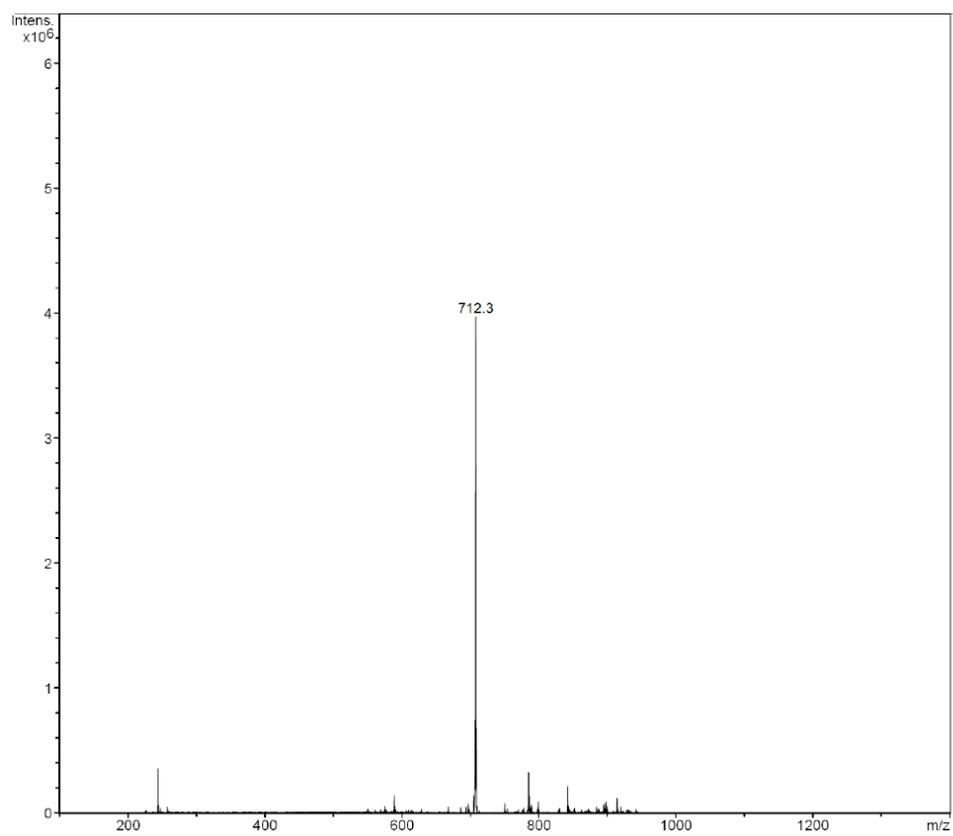


Figure S16.  $^1\text{H}$  and  $^{13}\text{C}$  NMR (200 MHz) spectra for compound **6a**.

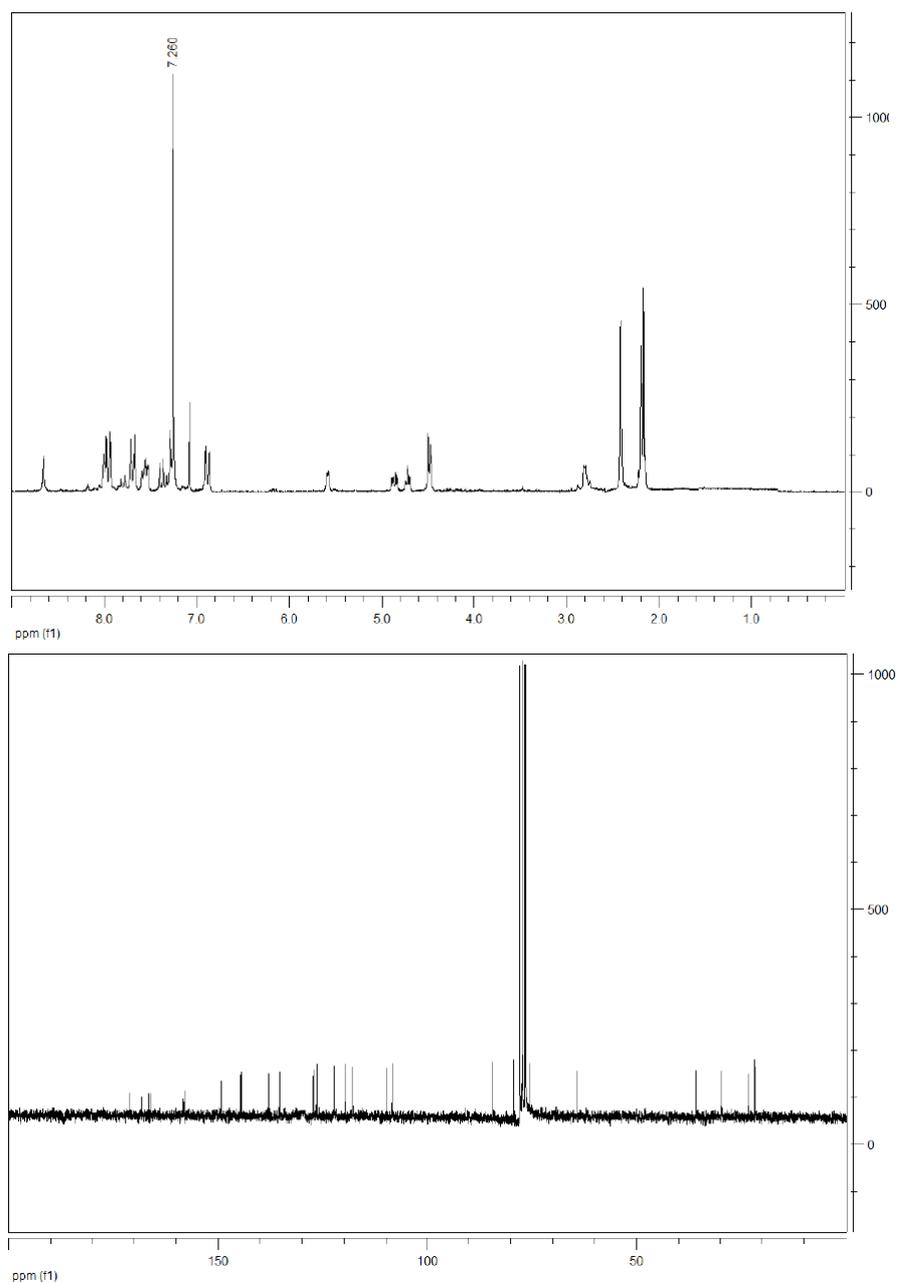


Figure S17.  $^1\text{H}$  and  $^{13}\text{C}$  NMR (200 MHz) spectra for compound **1a**.

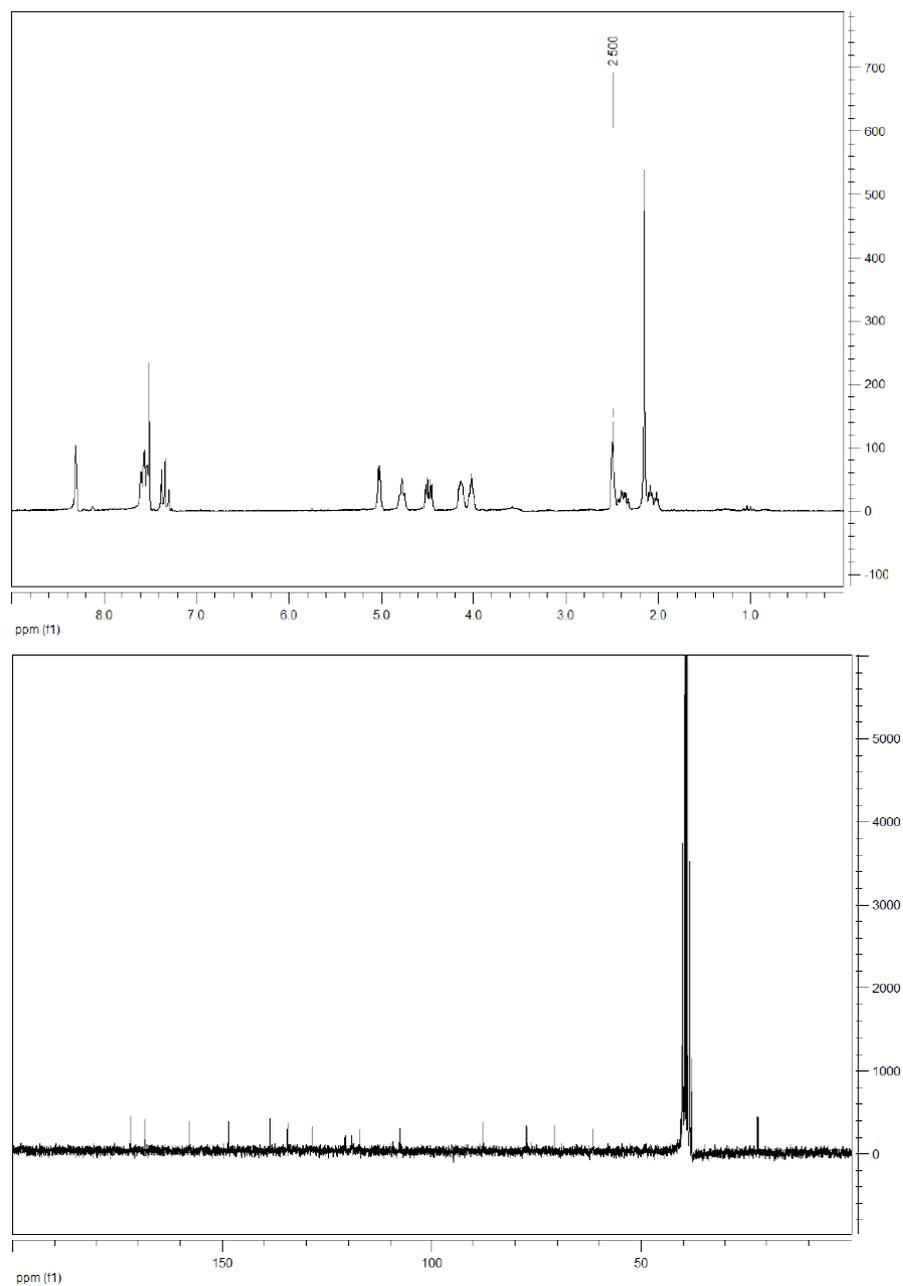


Figure S18.  $^1\text{H}$  and  $^{13}\text{C}$  NMR (200 MHz) spectra for compound **2a $\alpha$** .

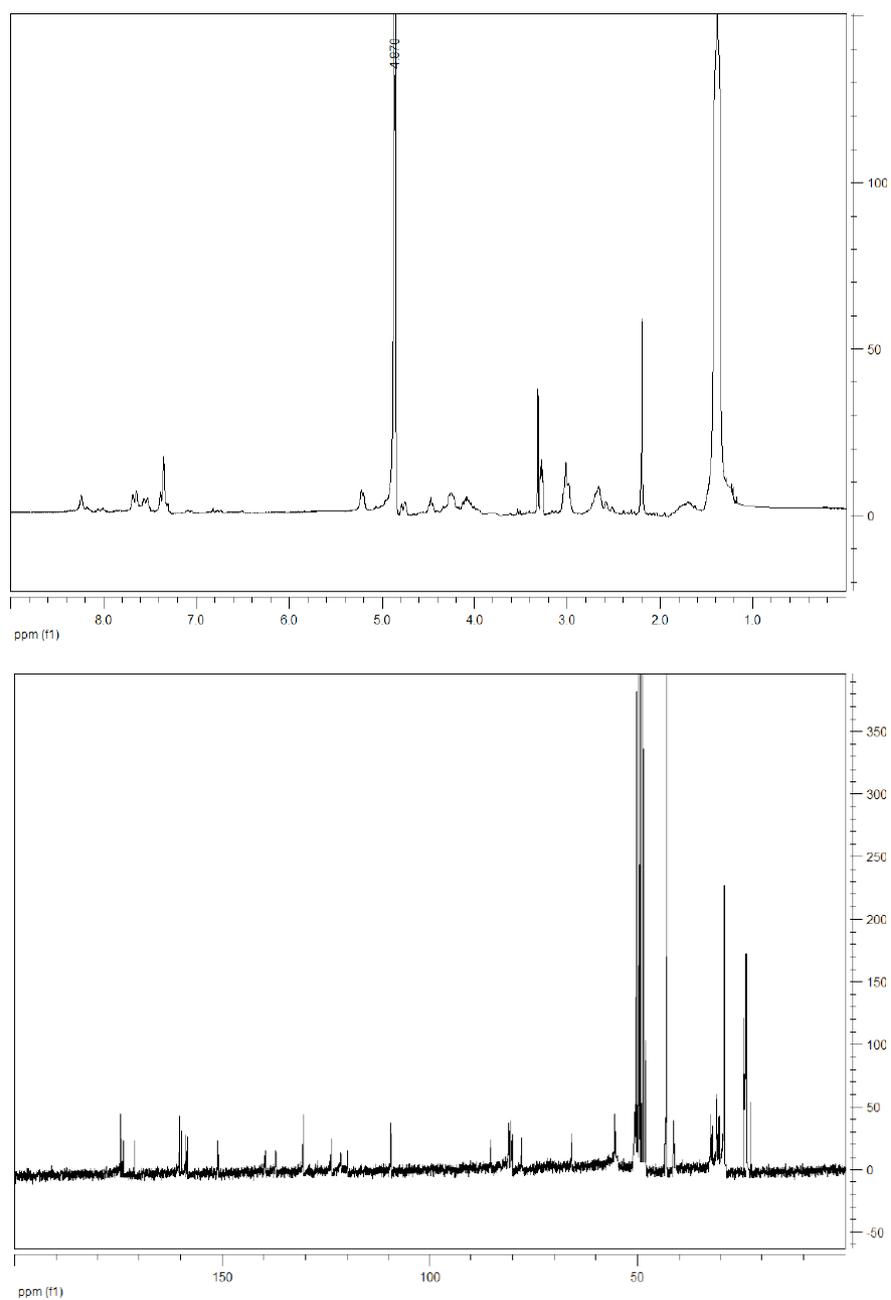


Figure S19.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra (200 MHz) for compound **2ba**.

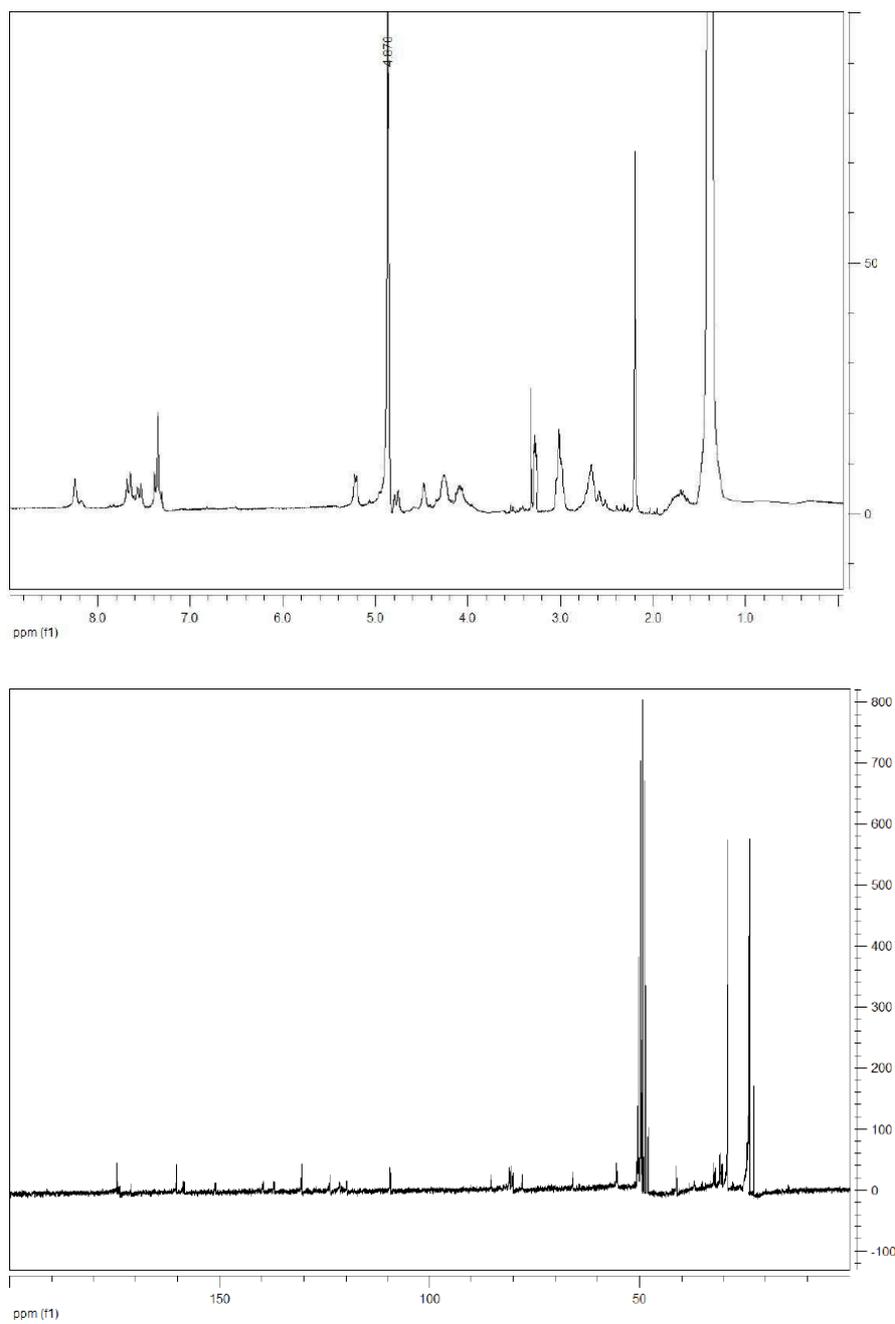
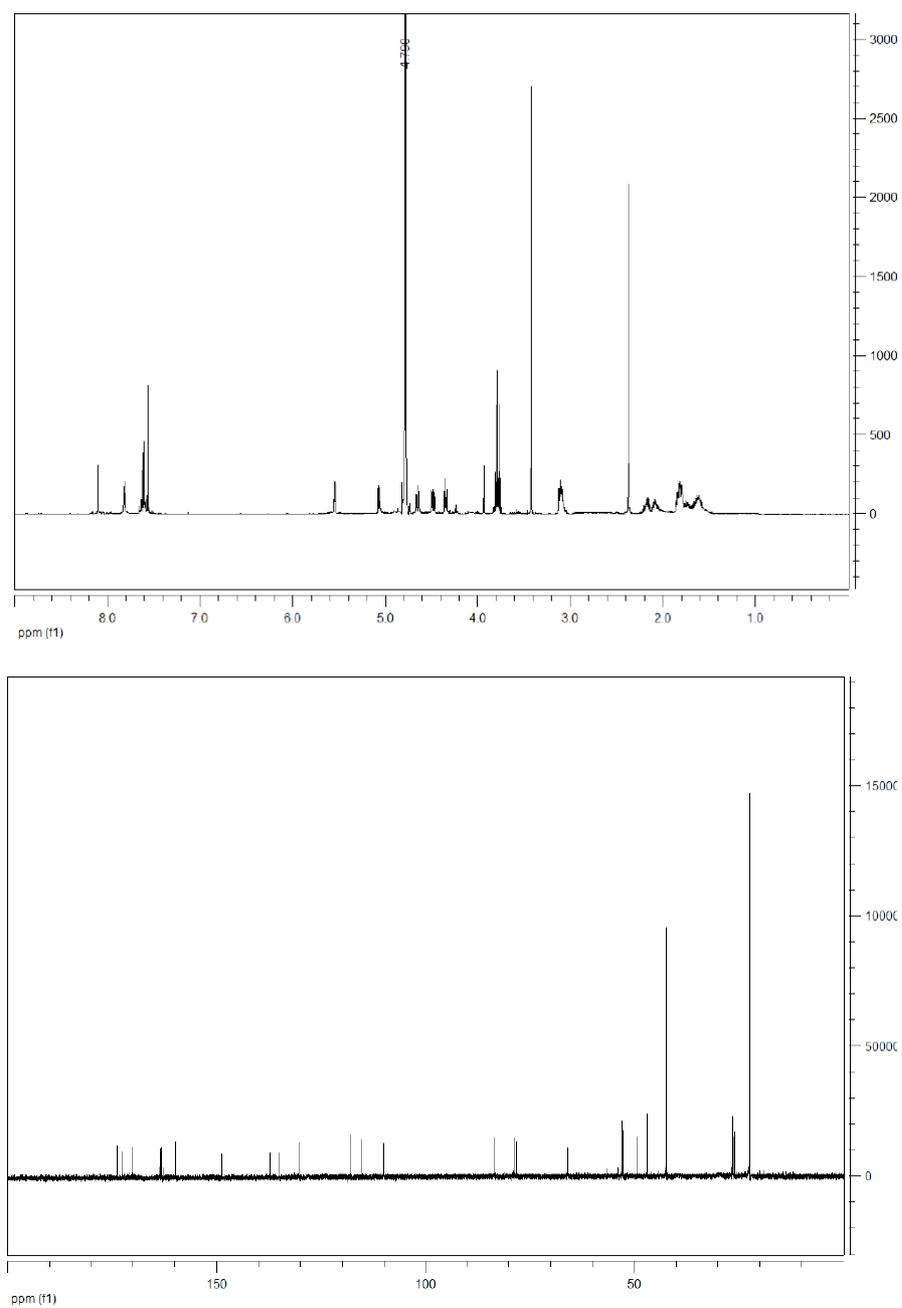


Figure S20.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (500 MHz) and HRMS spectra for compound **3a $\alpha$** .



DUCA\_100211160017 #1 RT: 0.02 AV: 1 NL: 2.99E6  
T: FTMS - p ESI SIM ms [501.00-511.00]

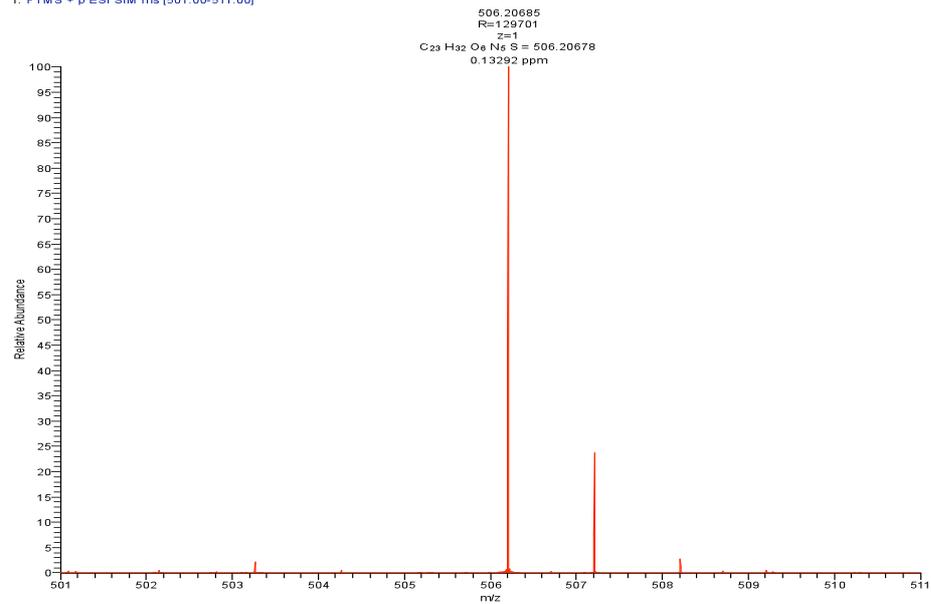
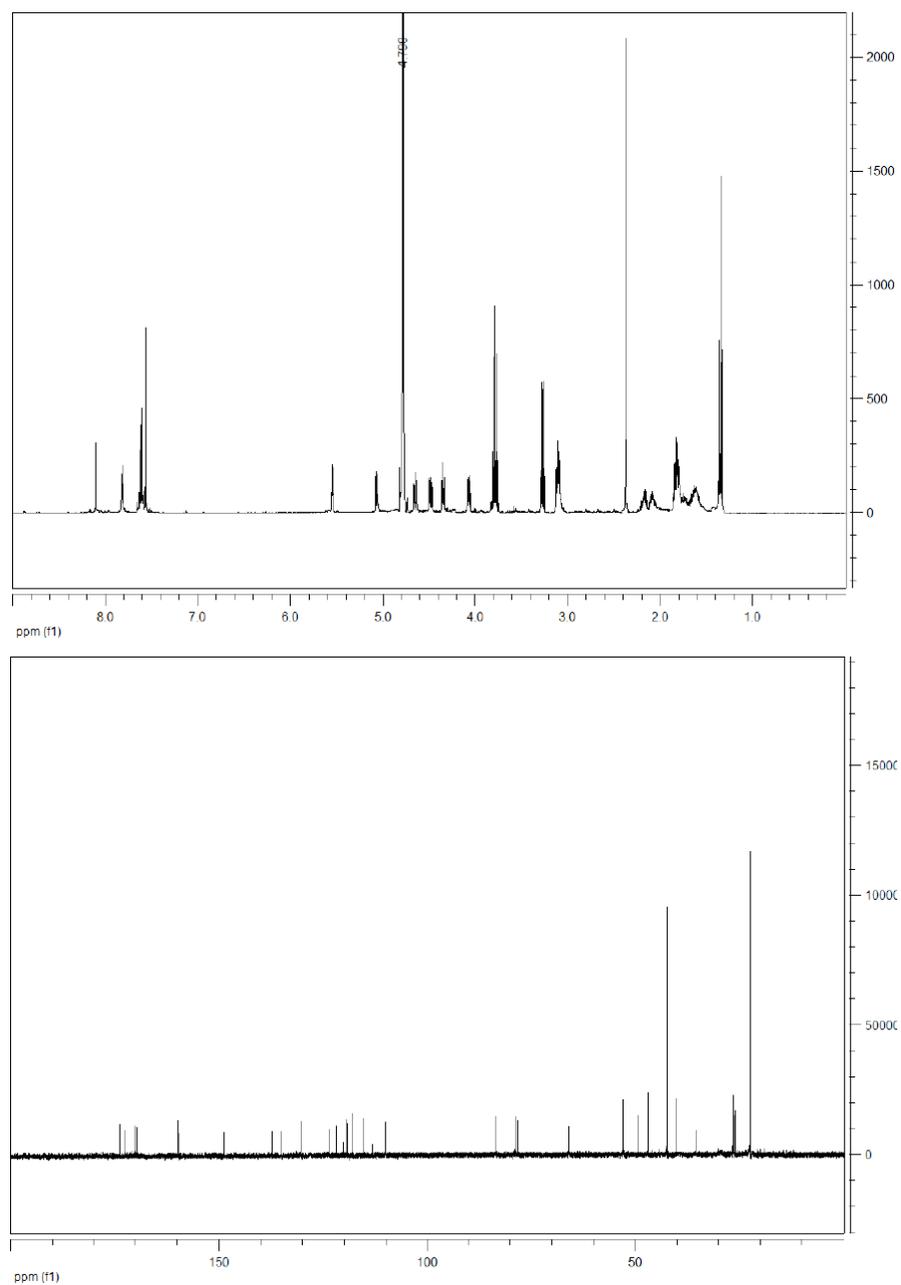


Figure S21.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (500 MHz) and HRMS spectra for compound **3ba**.



DUCA\_100211153915 #2 RT: 0.03 AV: 1 NL: 3.14E7  
T: FTMS + p ESI SIM ms [629.00-639.00]

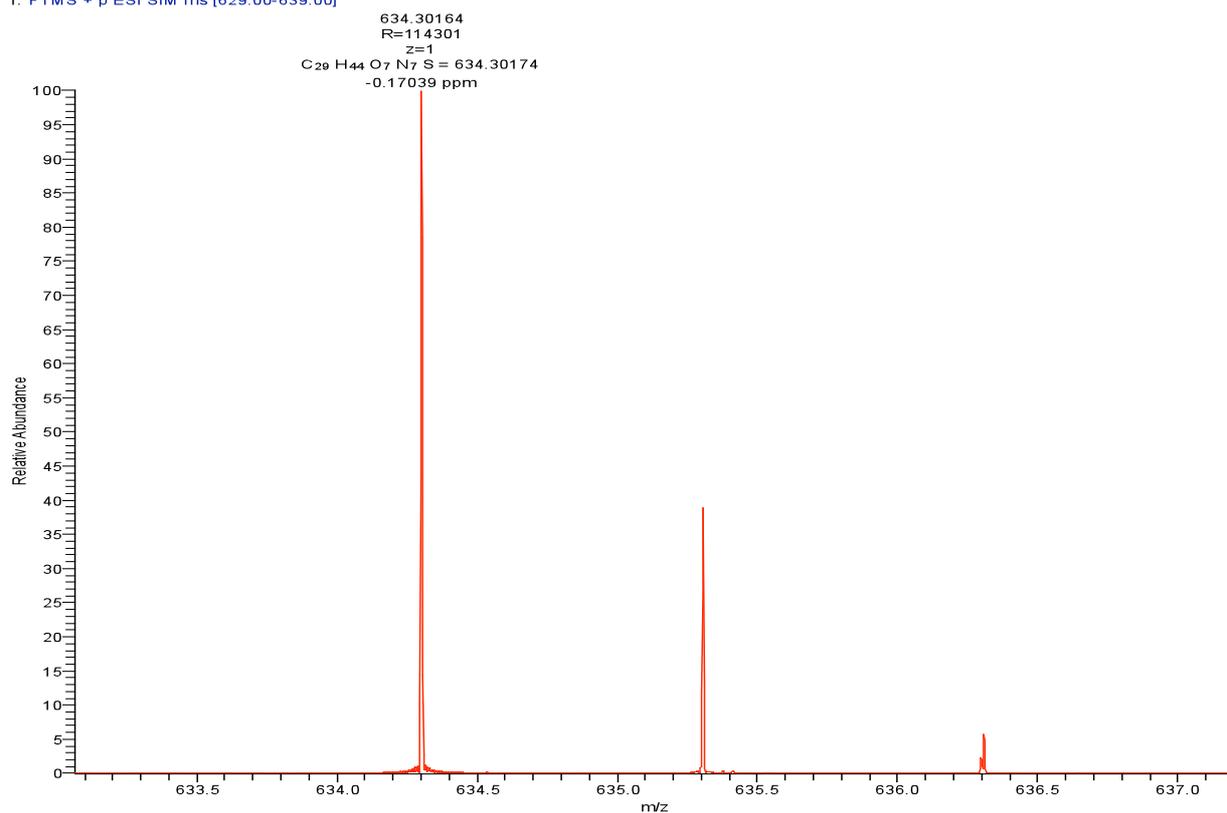
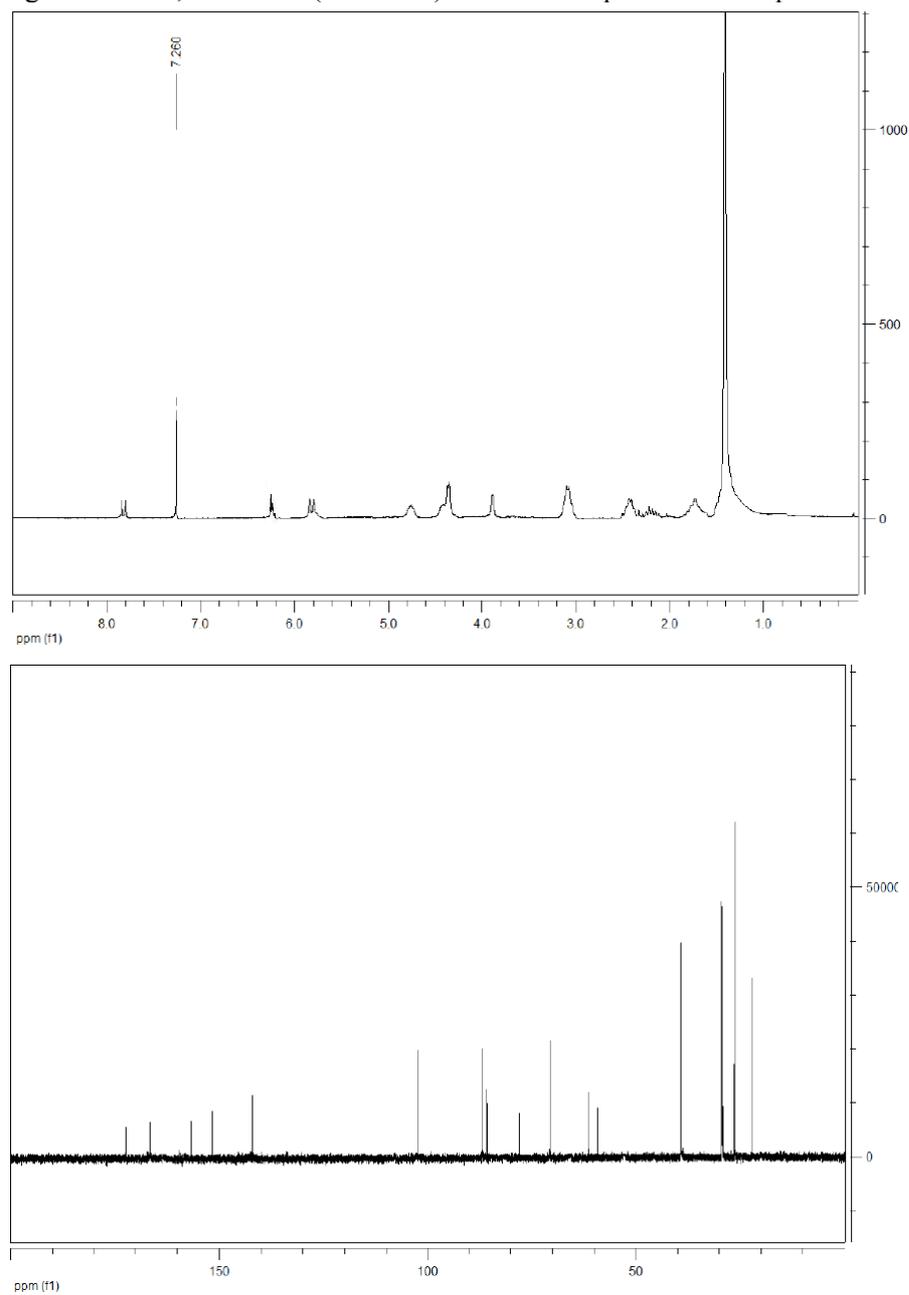


Figure S22.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (200 MHz) and HRMS spectra for compound **2e**.



100210MD08 #380 RT: 14.63 AV: 1 NL: 4.82E4  
T: FTMS - p ESI SIM ms [551.00-561.00]

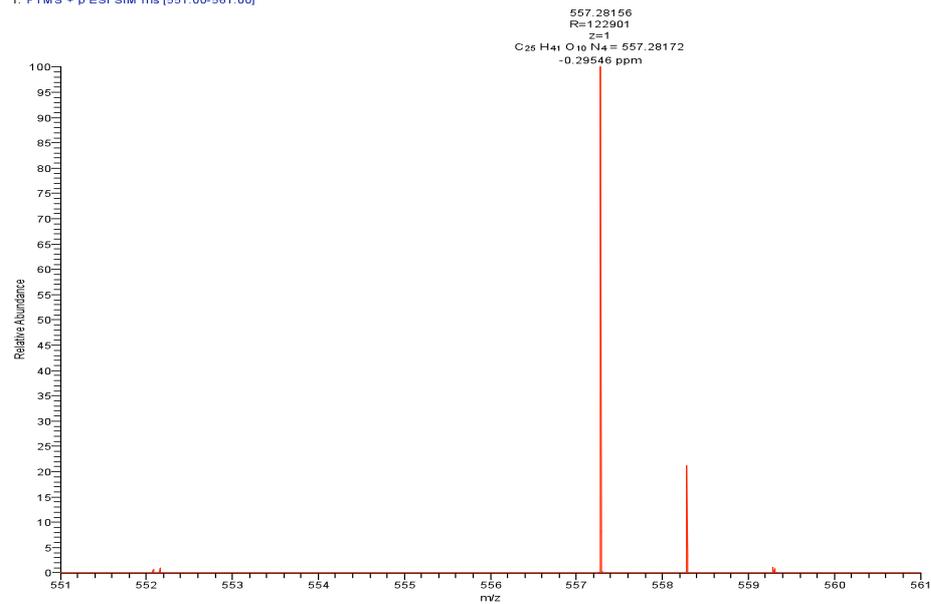
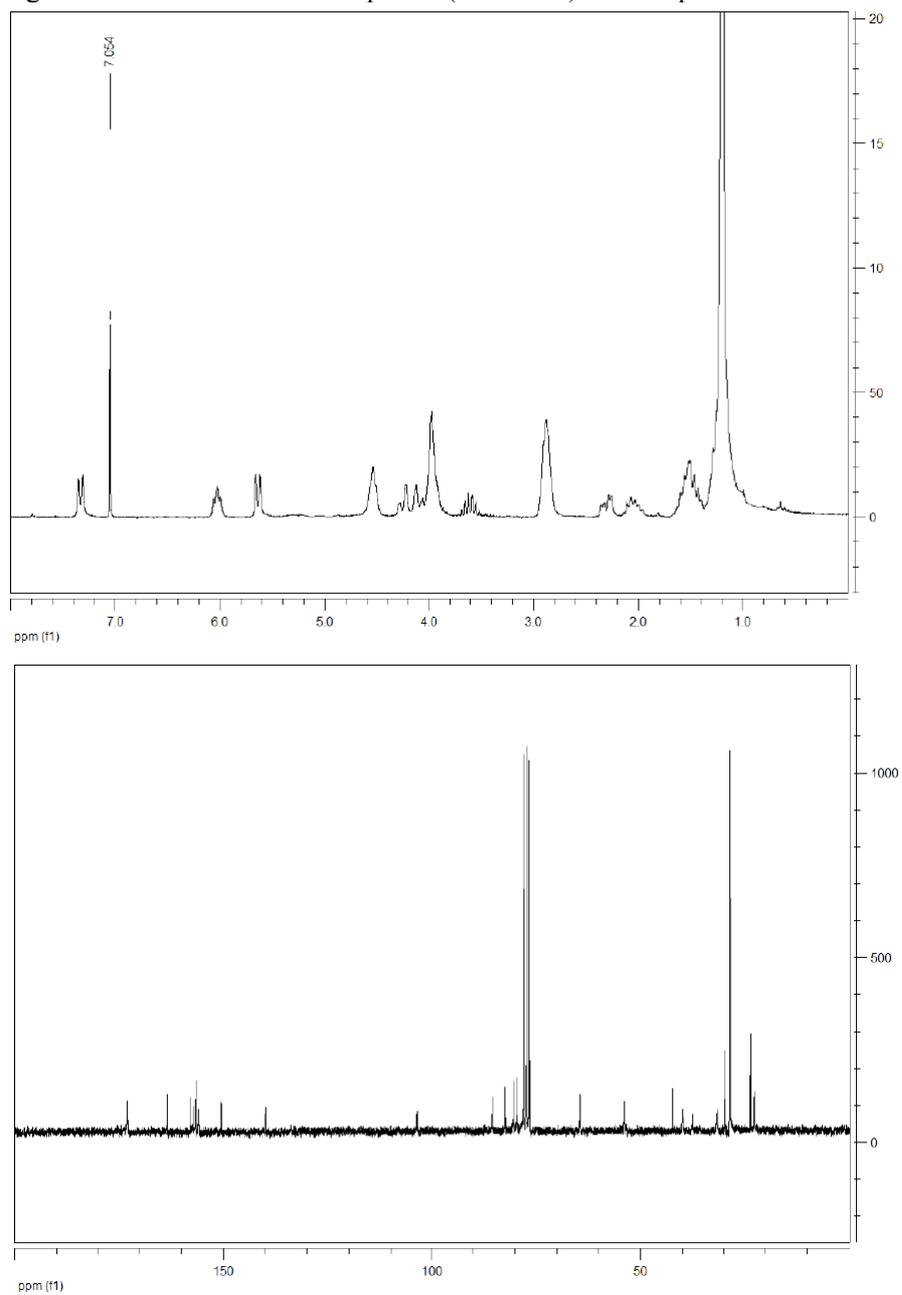


Figure S23.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra (200 MHz) for compound **2f**.



100210MD07 #546 RT: 20.99 AV: 1 NL: 2.06E7  
T: FTMS + p ESI SIM ms [879.00-889.00]

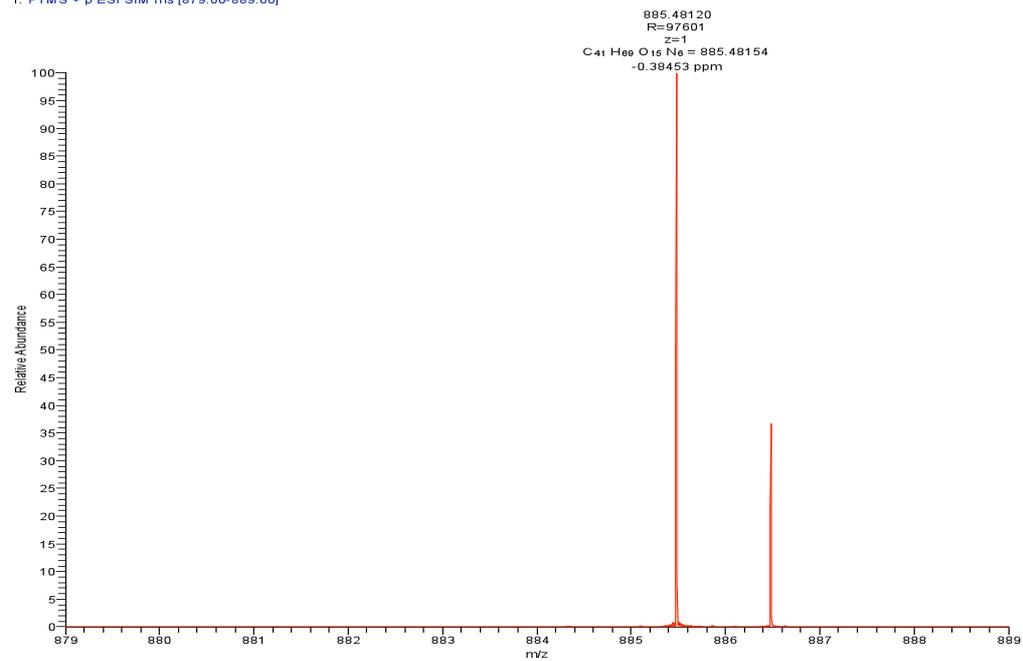
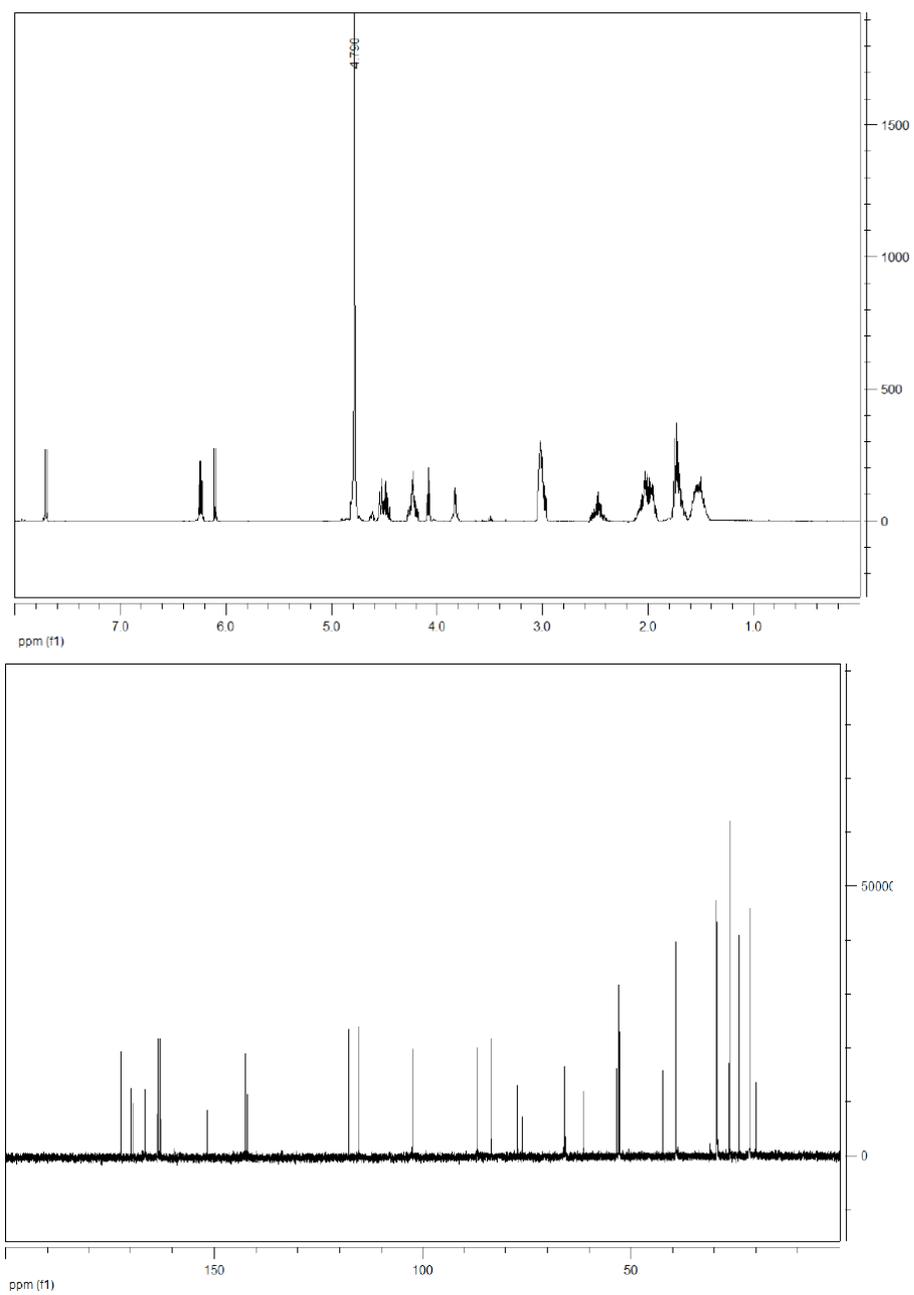


Figure S24.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR and HRMS spectra (500 MHz) for compound **3e**.



100210MD10 #36 RT: 1.29 AV: 1 NL: 8.97E6  
T: FTMS - p ESI SIM ms [351.00-361.00]

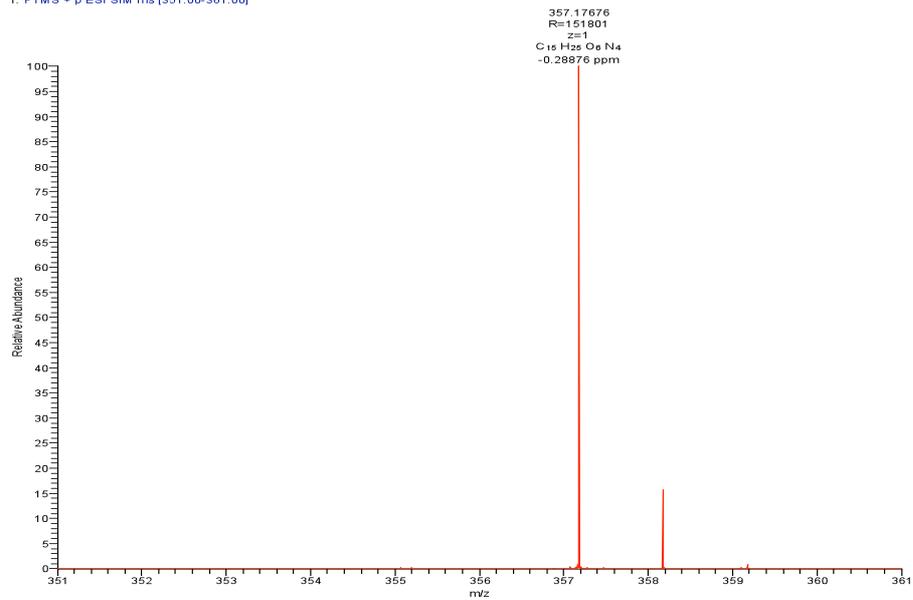
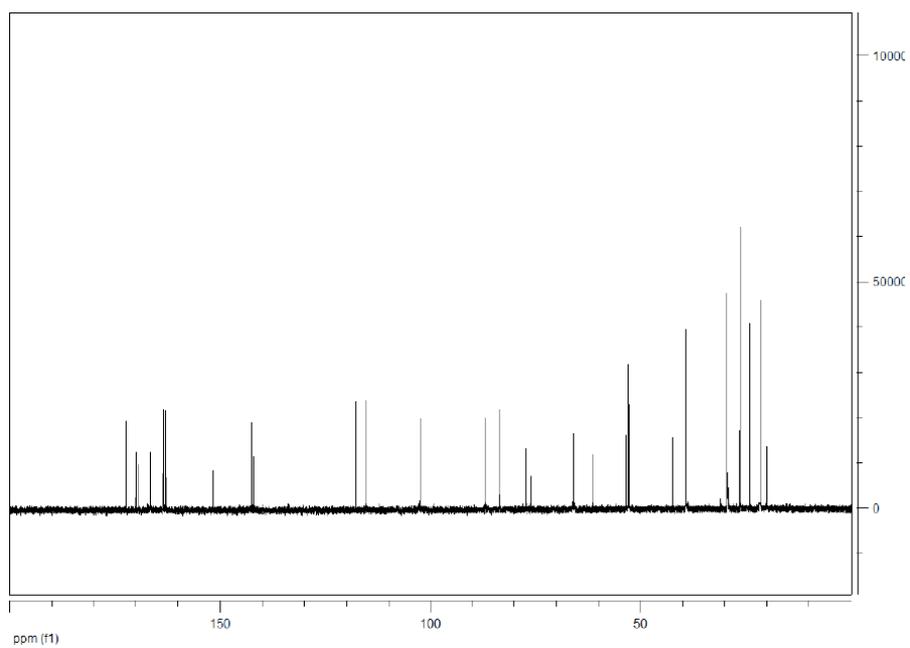
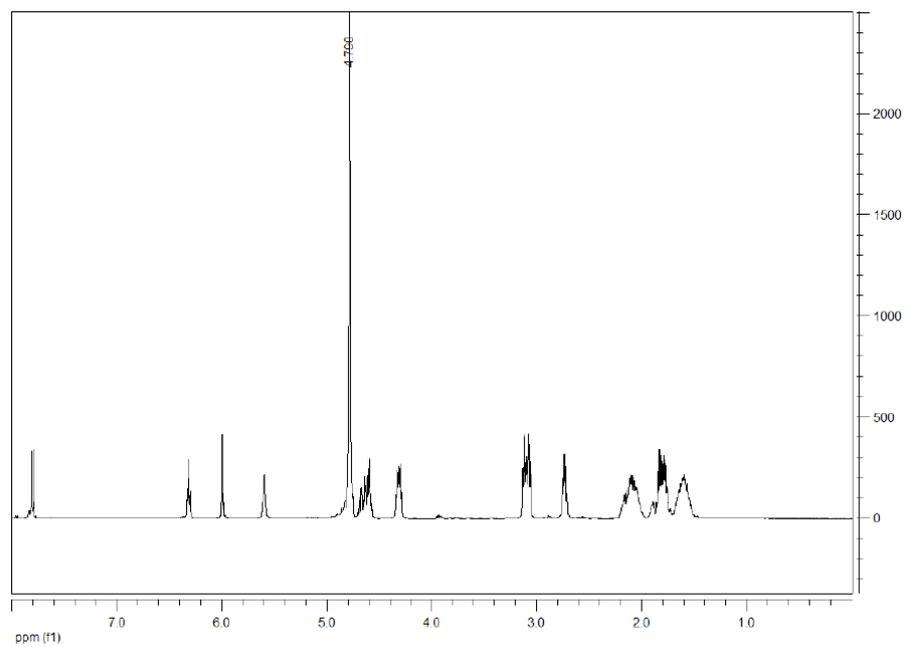
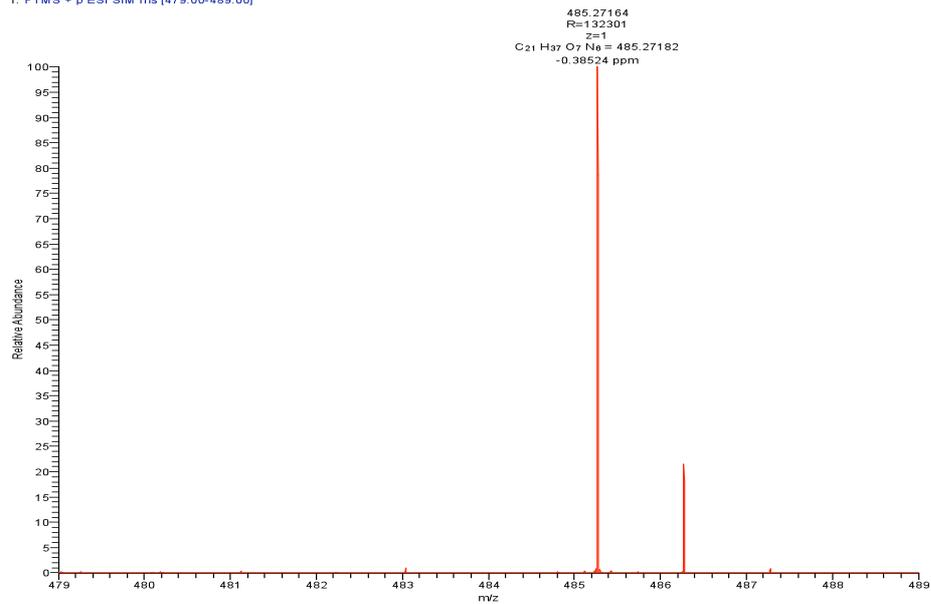


Figure S25.  $^1\text{H}$ ,  $^{13}\text{C}$  NMR (500 MHz) and HRMS spectra for compound **3f**.



100210MD09 #38 RT: 1.41 AV: 1 NL: 4.28E5  
T: FTMS - p ESI SIM ms [479.00-489.00]



## Computational details

Molecular modeling studies were addressed to understand position of ligand inside the TAR binding pocket for compounds **3a $\beta$**  and **3b $\beta$** . Such simulations are challenging since RNA are known to be flexible targets. Consequently, we employed a combination of molecular docking calculations and small molecular dynamic trajectories in implicit solvent to mimic the RNA induced fit effect. This simulation protocol has been already designed and validated for aminoglycoside ligands complexed with RNA-16S and has proved its reliability.<sup>7,8</sup> All calculation details can be found in these two last references.

The experimental structure of RNA TAR complexed with ligand rbt158<sup>9</sup> (pdb code 1UUI) was used as input for our calculations. For molecules **3a $\beta$**  and **3b $\beta$** , the best molecular docking conformation were selected as the one which presented the lowest-docking free energy of binding in the most populated cluster positions. The resulting complexes were then subjected to 500 ps of molecular dynamics simulations with an implicit water solvent method. This molecular dynamics protocol was also applied for all species, alone, as the TAR RNA fragment, compounds **3a $\beta$**  and **3b $\beta$** . Structural analyses were performed with the Visual Molecular Dynamics software.<sup>10</sup> We recorded a hydrogen bond (HB) when the donor-acceptor atomic distance is lower than 3.5 Å and the angle H-donor-acceptor is smaller than 45°.

## Results & Discussions

The molecular docking step made in this work was used to generate as correct as possible binding positions for compounds **3a $\beta$**  and **3b $\beta$**  prior to the molecular dynamics relaxations. However, in order to validate our docking calculations, the ligand rbt158 was extracted from the experimental TAR RNA structure and its position inside the target was predicted and compared to the experimental one. The good reproduction of the molecular functional groups positions, usually called pharmacophoric points, onto the targeted RNA structure definitely validates our calculations (figure S26).

For each compound, the selected conformation was the one, which presented the lowest-docking free energy of binding in the most populated cluster. The stability of the potentials and kinetics energies, which were monitored along the molecular dynamics trajectories, evidences the stability of the systems. Root Mean Square Deviation (RMSD) calculations were performed to analyze the movement of Tar RNA and ligand along the 500 ps of molecular dynamics. Such graph is represented as an example on figure S27.

The interpretations of the RMSD plots indicate that the systems are equilibrated quickly, generally after only 50 ps of trajectory. The TAR RNA complexed with compounds **3a $\beta$**  or **3b $\beta$**  are more twisted than without ligand, demonstrating this way the induced fit effect of RNA. After this structural adaptation of the ligand to the target and the target to the ligand, the global molecular movements are less loose than without compound **3a $\beta$**  or **3b $\beta$** , thereby evidencing the restraint effect of the ligands.

The energies of TAR RNA, ligands and complexes are averaged along the last 100 ps of their respective molecular dynamics trajectory. This way, the interaction energy can be easily obtained from the following equation:

$$\text{RNA} + \text{ligand} \rightarrow \text{Complex} \quad \text{gives} \quad E_{\text{int}} = \langle E_{\text{complex}} \rangle - \langle E_{\text{RNA}} \rangle - \langle E_{\text{ligand}} \rangle$$

Even if the interaction energy does not represent a "real" free energy of binding, these values can be compared to the experimental values of dissociation constant. Indeed, interaction energies of RNA TAR fragment with compounds **3a $\beta$**  and **3b $\beta$**  are, respectively, of -50.4 kcal/mol and -86.0 kcal/mol whereas the dissociation constant is of 36.0  $\mu$ M for ligand **3a $\beta$**  and 1.6  $\mu$ M for ligand **3b $\beta$** . Interestingly, we then observed the coherence between experimental results and molecular modeling since the complex associated to the lowest energy is associated to the lower dissociation constant.

The structure of compound **3a $\beta$**  on the TAR RNA cavity is displayed on figure S27. This compound is bearing a global net charge of +2 owing to the ammonium groups of the N-terminal and side chain parts of the lysine residue. Obviously, these positive hot spots strongly interact with the phosphates negative charges. To be more precise, the "N-terminal" ammonium group is localized between the phosphates of the residues Guanine 36 and Uracil 23. The second ammonium group (lysine side chain moiety) is interacting with the phosphate group of Adenine 22. These strong electrostatic interactions fix ligand **3a $\beta$**  in the major groove of TAR RNA structure. In addition to these electrostatic contacts, three hydrogen bonds are recorded. These are qualified as strong hydrogen bonds since they are observed more than 85% of time of the trajectory. The first of these three H-bonds involves the hydroxy group of the sugar moiety, which is making such hydrogen bond with the O2 atom of the Cytosine 39 phosphodiester group. This H-bond induces the orientation of molecule **3a $\beta$**  and limits the molecular movement of the aromatic part. Interestingly, this part performs a "head to tail" structure with the neighbouring bases A20 and C19 on the major groove. Indeed, the aromatic part of **3a $\beta$**  accepts two hydrogen bonds, one with the amino group of Adenine 20 and the sulfur atom and the other one with the amino group of Cytosine 19 and the oxygen of the terminal amido group. All interactions described above appear on figure S28.

A view of the molecular surface of TAR RNA is displayed on figure S29. This figure exhibits a supplemental cavity at the hydroxy vicinity of the sugar moiety. Ligand **3b $\beta$**  is bearing a supplementary lysine residue at this hydroxy position, which therefore fills this cavity. The structure of compound **3b $\beta$**  on the TAR RNA cavity is displayed on figure S30. As expected, the electrostatic interactions of the **3b $\beta$**  ammonium moieties with the phosphates of residues G36, U23 and A22 are identical for compound **3b $\beta$** . According to the fact that there is a second lysine residue at the 3'-position, the previously observed hydrogen bond is obviously missing for compound **3b $\beta$** . However, the two new positive charges of this residue generate new "hot-spots" which makes electrostatic interaction with the phosphate of residues U38, C39 and U40. This way, the second lysine of compound **3b $\beta$**  follows the RNA strand direction and orientation. However, the supplementary lysine is making steric contacts with the aromatic part and slightly pushes this group from the conformation found for compound **3a $\beta$** . In the molecular complex involving **3b $\beta$**  and TAR RNA, the aromatic part is making two strong hydrogen bonds with the O2 atoms of phosphates G18 and A20.

Figure S26: The predicted structure of ligand rbt158 (green) is compared to its experimental complexed structure (blue). The overall good reproduction of the position of this ligand validates the molecular modeling calculations.

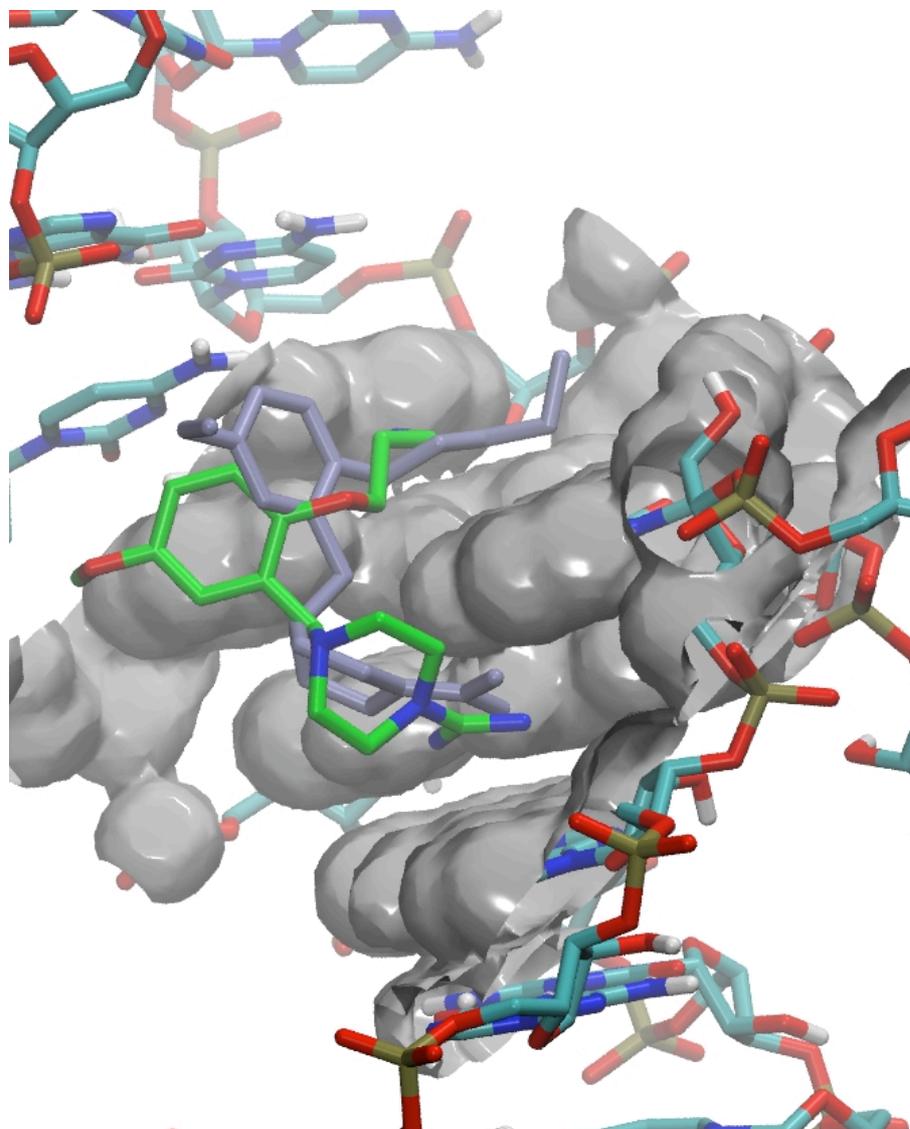


Figure S27: RMSD versus time plots for TAR RNA complexed with molecule  $3a\beta$  = molecule 6.

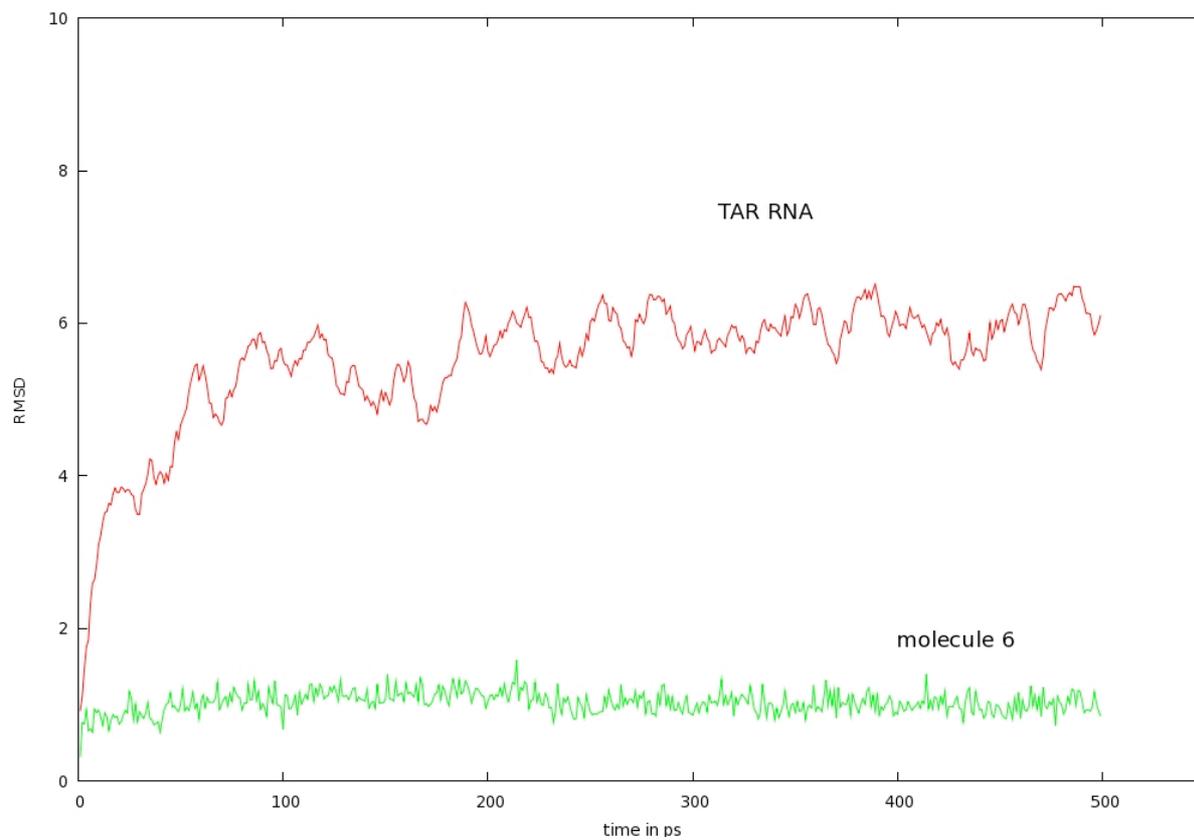


Figure S28: Molecular structure of compound **3a $\beta$**  in complex with TAR RNA fragment.

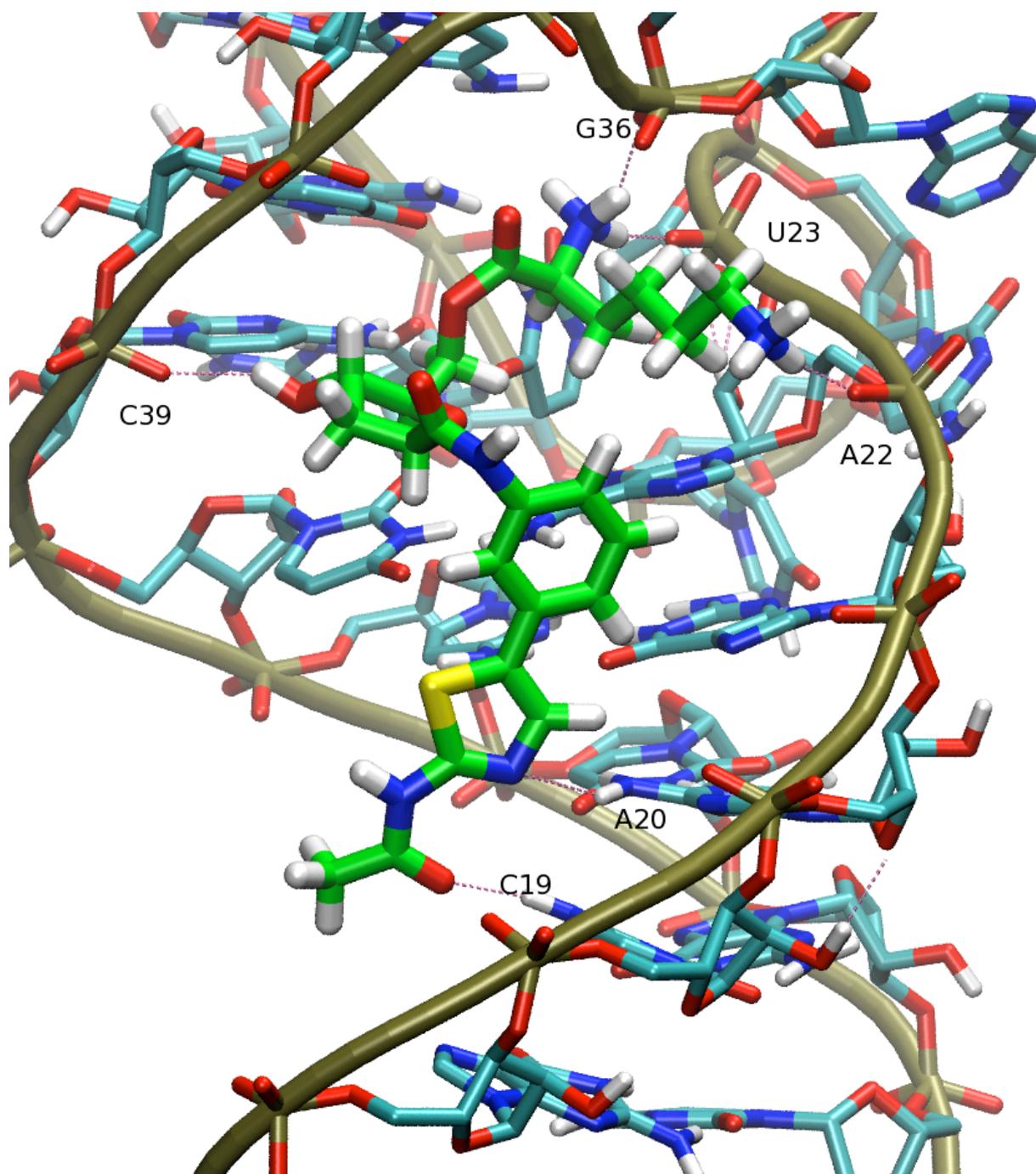


Figure S29: Molecule **3a $\beta$**  in complex with TAR RNA fragment. The molecular surface of RNA exhibits a supplementary cavity at the hydroxy vicinity.

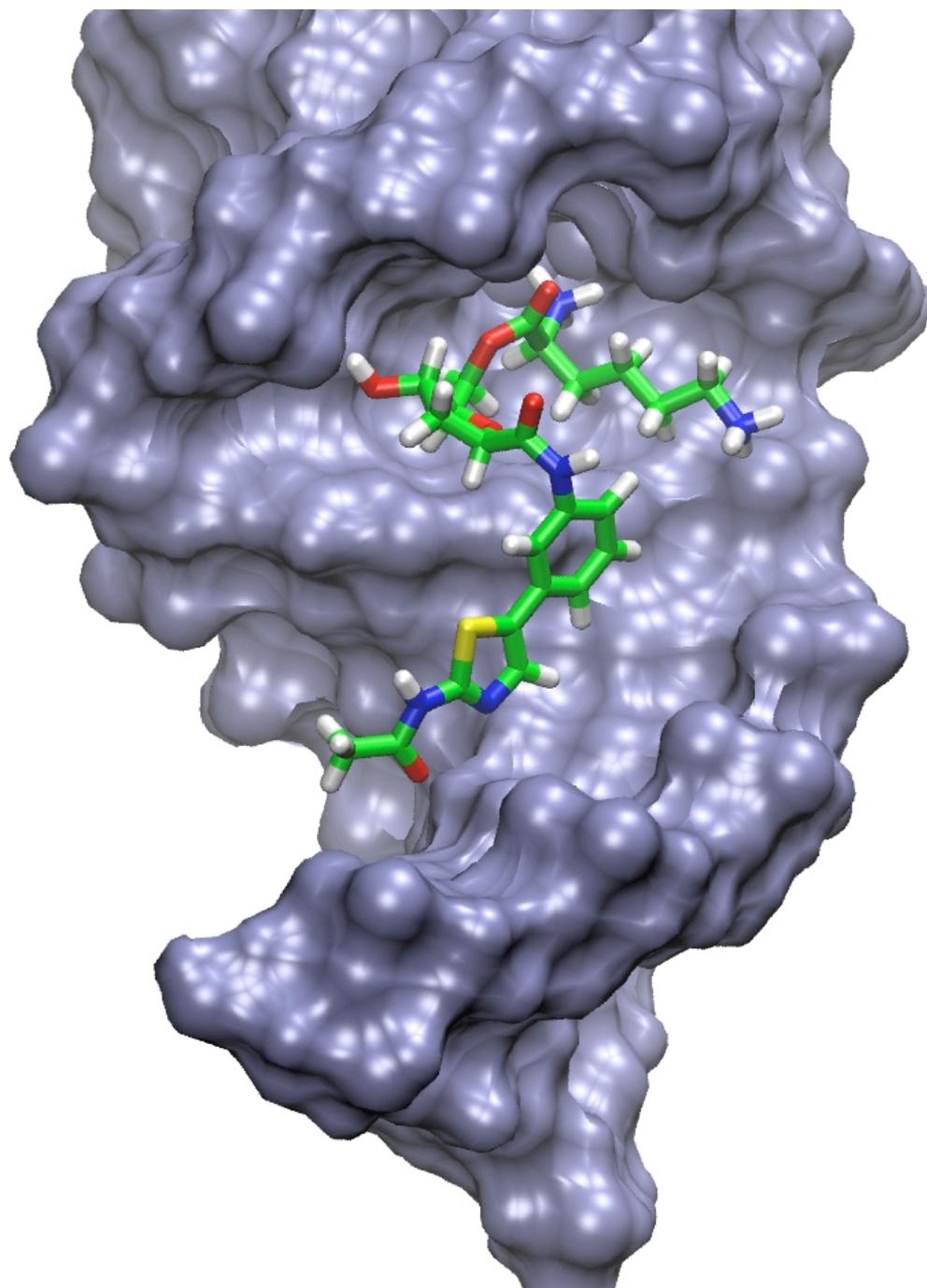
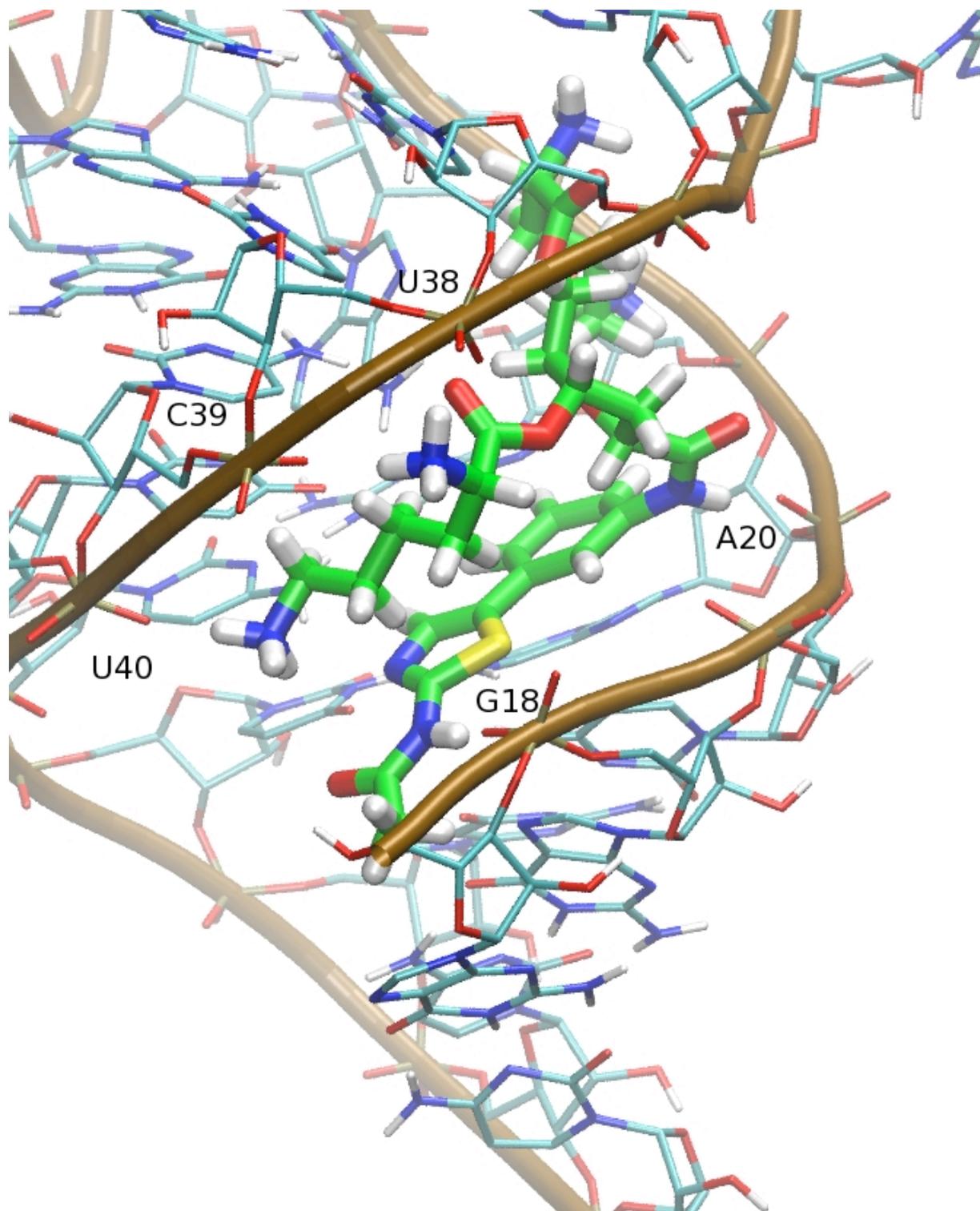


Figure S30: Molecular structure of compound **3b $\beta$**  in complex with TAR RNA fragment.



## References

1. M. Jetten, A. M. Peters, W. F. M. Van Nispen, W, C. J. Ottenheijm, *Tetrahedron Lett.* 1991, **32**, 6025.
2. C. Youan, R. M. Williams *J. Am. Chem. Soc.* 1997, **119**, 11777.
3. N. W. Luedtke, Q. Liu, Y. Tor, *Biochemistry* 2003, **42**, 11391.
4. V. Bonnard, S. Azoulay, A. Di Giorgio, N. Patino, *Chem. Commun.* 2009, 2302.
5. N. Tassew, M; Thompson, *Biophysical Chemistry*, 2003, 241.
6. (a) C. A. Royer, *Anal. Biochem.*, 1993, **210**, 91 ; (b) C. A. Royer, J. M. Beechem, *Methods Enzymol.*, 1992, **210**, 481.
7. F. Barbault, L. Zhang, L. Zhang, B. T. Fan, *Chemometrics and Intelligent Laboratory System*, 2006, **82**, 269.
8. F. Barbault, B. Ren, J. Rebehmed, C. Teixeira, Y. Luo, O. Smila-Castro, F. Maurel, B. T. Fan, L. Zhang, L. Zhang, *Eur. J. Med. Chem.*, 2008, **43**, 1648.
9. B. Davis, M. Afshar, G. Varani, Al Murchie, J. Karn, G. Lentzen, M. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. Swarbick, F. About-Ela, *J. Mol. Biol*, 2004, **336**, 343.
10. W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33.