Design, Synthesis of new Potent and Selective Inhibitors of Matriptase.

Éloïc Colombo,[1] Antoine Désilets,[1] Dominic Duchêne,[2] Félix Chagnon,[1] Rafael Najmanovich,[2] Richard Leduc,[1] and Eric Marsault[1]*
[1]Department of Pharmacology, [2]Departement of Biochemistry, Faculty of Medicine and Health Sciences, Université de Sherbrooke, 3001 12e av Nord, Sherbrooke PQ, Canada J1H5N4.

Supplementary Information

Table of contents

1. **Chemistry general methods** **p. 2**
2. **Synthetic procedures** **p. 2-5**
3. **Biological methods** **p. 6-7**
4. **Modelling procedures** **p. 7-8**
5. **References** **p.9-10**
6. **Chemistry data** **p. 11-34**

1. Chemistry – general methods

Amino acids and coupling reagents were obtained from from ChemImpex International (USA) and used as received. All other reagents and solvents were purchased from Sigma–Aldrich (Canada) or Fisher Scientific (USA). Tetrahydrofuran (THF) was dried on sodium benzophenone ketyl; DCM on P_2O_5; Methanol on magnesium. Analytical HPLC experiments were performed on a Agilent 1100 series instrument equipped with UV detector set at 223 nm and an Agilent Eclipse Plus C18 column (3.0 x 50 mm, 1.8 mm spherical particle size column) with a linear gradient of 2-50% CH_3CN and H_2O containing 0.1% TFA (10 min), 50-100% (4 min), 100% (4 min), 100-2% (1 min) and 2% (3 min). Final products were purified to > 95% purity (HPLC-UV) by preparative HPLC (Beckman 126 instrument) using a Vydac C18, 250 x 22 mm ID, 5 µm particle size column and a linear gradient of acetonitrile containing 0.1% TFA at a flow rate of 7 mL/min. All inhibitors were obtained as TFA salts after lyophilisation. Molecular weights of compounds were confirmed by mass spectrometry (Electrospray micromass ZQ-2000 from Waters). ^1H and ^{13}C NMR spectra were recorded on an ECX-300 (Bruker Inc., USA) at 300 MHz, and are referenced to internal solvent signals.

2. Synthetic procedures

Tert-butyl-N-{1-[methoxy(methyl)carbamoyl]-4-{3-[(3-methoxy-2,5,6-trimethylbenzene)sulfonyl]carbamimidamido}butyl}carbamate (7): To a solution of Boc-Arg(Mtr)-OH (1.36 g, 2.79 mmol) in anhydrous THF (93 mL) were added 1.27 g HATU (3.35 mmol, 1.2 eq), 327 mg HN(Me)OMe • HCl(3.35 mmol, 1 2 eq) and 2.4 mL DIPEA (14 mmol, 5 eq) under nitrogen. The reaction mixture was stirred over 3h at ambient temperature. After completion of the reaction, the mixture was concentrated *in vacuo* and directly purified by flash chromatography (EtOAc/Hexane 80:20 to 100:0) to give Weinreb amide **7** as a white solid (1.4 g, 93%). HPLC: 10.98 min, 90%. ^1H NMR (300 MHz, $CDCl_3$) δ 1.41 (s, 9H), 1.55-1.71 (m, 4H), 2.59 (s, 3H), 2.67 (s. 3H), 2.80 (s,3H), 3.22 (s, 3H), 3.29-3.33 (m, 1H), 3.72 (s, 3H), 3.82 (s, 3H), 4.63 (t, 1H), 5.47 (d, 2H), 6.20-6.31 (m,2H) 6.53 (s, 3H).^{13}C NMR (75.4 MHz, $CDCl_3$) δ 11.94,18.20, 24.12,24.59, 28.34, 32.13, 41.00, 49.31, 55.46, 61.65, 80.38, 111.77, 120.20, 124.91, 136.25, 138.77, 155.74, 156.62, 158.54. MS (ES) m/z 530.2 $(M+H)^+$.

1-[4-amino-5-(1,3-benzothiazol-2-yl)-5-oxopentyl]-3-[(3-methoxy-2,5,5,6-trimethylbenzene)sulfonyl]guanidine (8): A solution of benzothiazole (1.7 mL, 15.5 mmol, and 18 eq) in 52 mL of anhydrous THF was cooled to -78 °C with stirring under nitrogen. N-Butyllithium (1.02 M in hexane, 14.8 mL, 15.1 mmol, 17.5 eq) was added dropwise over 30 min at -78 °C and the mixture was stirred for an additional 30 min at the same temperature. To this solution was added dropwise over 50 min a solution of Weinreb amide **7** (456 mg, 0.86 mmol, 1 eq) in 27 mL anhydrous THF at -78 °C. The reaction mixture was stirred at -78 °C for 2 h, then quenched with 20 mL of saturated aqueous NH_4Cl and extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with water (3 x 50 mL) then brine (2 x 50 mL), dried on anhydrous $MgSO_4$, filtered through fritted glass, and then concentrated *in vacuo*. The residue was dissolved in anhydrous MeOH (17.2 mL), cooled to -20 °C while stirring, and treated with $NaBH_4$ (198 mg, 5.2 mmol, 6 eq). After 1 h, acetone (20 mL) was added and the reaction mixture was warmed to room temperature over 30 min. The reaction mixture was concentrated *in vacuo* and partitioned between water and EtOAc. The aqueous layer was extracted with EtOAc (2 x 20 mL), and the combined organic extracts were washed with brine (2 x 20 mL), dried ($MgSO_4$), filtered through fritted glass and concentrated *in vacuo*. The residue was purified by flash chromatography (EtOAc/Hexane

80:20 to 100:0) to give intermediate **8** as a yellow solid (322 mg, 63%). HPLC: 12.09 min, 100% purity. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.24 (s, 4H), 1.37 (s, 5H), 1.40-1.72 (m, 4H), 2.58 (s, 3H), 2.64 (s,3H), 2.66 (s,3H), 3.08-3.35 (m, 2H), 3.80 (s, 3H), 4.13 (t, 1H), 5.16 (s, 1H), 5.52 (d, 2H), 6.21-6.38 (m, 3H), 6.50 (s, 3H), 7.90 (d, 1H), 7.95 (d, 1H). ^{13}C NMR (75.4 MHz, $CDCl_3$) δ 11.9, 19.1, 23.2,24.1, 26.9, 28.8, 40.9, 55.5, 60.4, 80.2, 112.1, 121.6, 122.5, 124.8, 124.9, 125.0, 126.9, 134.7, 136.7, 138.6, 152.7, 156.6, 157.1, 158.5, 173.0. MS (ES) m/z 606.1$(M+H)^+$.

Benzyl-2-(2-{[(9H-Fluoren-9-ylmethoxy)carbonyl]amino}-4-[(triphenylmethyl)carbamoyl]butanamido)propanoate (9) : To a solution of Fmoc-Gln(Trt)-OH (1.83 g, 3.0 mmol, 1 eq) in 30 mL anhydrous THF were added 1.37 g HATU (3.6 mmol, 1.2 eq), 1.05 g H_2N-Ala-OBn (1.42 mmol, 1.0 eq), and 2.6 mL DIPEA (15.0 mmol, 5.0 eq) under nitrogen. The reaction mixture was stirred over 3h at room temperature. After completion of the reaction, the mixture was concentrated *in vacuo* and purified by flash chromatography, (EtOAc/Hexane 50:50 to 70:30) to give compound **9** as a white solid (2.07 g, 90%). HPLC: 15.19 min, 100% purity. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.29 (d, 1H), 1.95-2.00 (m, 4H), 4.10-4.20 (m, 2H), 4.22-4.37 (m, 1H), 4.47-4.52 (m, 1H), 5.05-5.17 (m, 2H), 7.10-7.41 (m, 26H), 7.76 (d, 2H). ^{13}C NMR (75.4 MHz, $CDCl_3$) δ 17.56, 30.00, 33.29, 47.16, 48.28, 53.35, 66.99, 77.49, 119.97, 125.19, 127.08, 128.22, 128.68, 141.30, 144.47, 144.52, 170.97. MS (ES) m/z 794 $(M+Na)^+$.

2-[2-(2-{[(tert-butoxy)carbonyl]amino}-5-{3-[(4-methoxy-2,3,6-trimethylbenzene)sulfonyl]carbamimidamido}pentanamido)-4-[(triphenylmethyl)carbamoy]butanamido]propanoic acid (10): A solution of Fmoc-Gln(Trt)-Ala-OBn (1.06 mg, 1.37 mmol, 1 eq) in 13.2 ml of CH_2Cl_2/Et_2NH (90:10) was stirred for 3 h at room temperature, then concentrated *in vacuo*. The residue was diluted with CH_2Cl_2 (3 x 20 mL) and concentrated under *in vacuo* to give the intermediate free amine as a colorless oil. To a solution of 798 mg Boc-Arg(Mtr)-OH (1.64 mmol, 1.2 eq) in 14 mL of CH_2Cl_2 was added 314 mg EDC (1.64 mmol, 1.2 eq), 1.37 mmol H_2N-Gln(Trt)-Ala-OBn (1.00 eq) and 222 mg HOBt (1.64 mmol, 1.2 eq). The mixture was stirred at room temperature overnigth. After completion, the reaction mixture was concentrated *in vacuo* and purified by flash chromatography (EtOAc/Hexane 80:20 to 100:0) to yield intermediate **10** as a white solid (813 mg, 77%). HPLC-UV: 14.48 min, 80% purity. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.28 (d, 2H), 1.36 (s, 10H), 1.58-1.66 (m, 1H), 2.42 (s, 3H), 2.59 (s,3H), 2.63 (s,3H), 3.02-3.18 (m, 1H), 3.77 (s, 3H), 4.08-4.15 (m, 1H), 4.34-4.41 (m, 2H), 4.95 (dd, 2H), 5.35 (d, 2H),6.08- 6.30 (s, 3H), 6.47 (s, 3H), 7.14-7.36 (m, 20H). ^{13}CNMR (75.4 MHz, $CDCl_3$) δ 11.0, 17.1, 18.4, 24.2, 25.2, 28.3, 29.8, 48.5, 52.7, 53.6, 55.4, 66.9, 77.5, 79.9, 111.7, 121.1, 126.9, 127.9, 128.9, 131.3, 135.3, 138.7, 144.3, 155.0, 155.6, 158.2, 172.5. MS (ES) m/z 1018.0 $(M+H)^+$.

The intermediate (320.7 mg, 0.31 mmol, 1 eq), 10% Pd on activated carbon (140 mg), and absolute EtOH (3.1 mL) were stirred under an atmospheric pressure of hydrogen for 1 h at ambient temperature. The resulting mixture was filtered on diatomaceous earth through fritted glass and concentrated *in vacuo* to give compound crude compound **10**, which was used as such in the following step.

Tert-butyl N-[1-({1-[(1-{[1-(1,3-benzothiazol-2-yl)-1-hydroxy-5-{3-[(4-methoxy-2,3,6 trimethylbenzene)sulfonyl]carbamimidamido}penta-2-yl]carbamoyl}ethyl)carbamoyl]-3-[(triphenylmethyl)carbamoyl]propyl}carbamoyl)-4-{1-[(4-methoxy-2,3,6-trimethylbenzene)sulfonyl]carbamimidamido}butyl]carbamate (11): To a solution of Boc-Arg(Mtr)-Gln(Trt)-Ala-OH (0.31 mmol, 1 eq) in anhydrous DMF were added 354 mg HATU (0.93 mmol, 3 eq), 0.31 mmol hydroxybenzothiazole, compound **8** (1.0 eq), and 0.3 mL of DIPEA (1.55 mmol, 5 eq). The mixture was stirred overnight, concentrated *in vacuo* and purified by flash chromatography (MeOH/EtOAc 95:5) to give compound **11** (366.2 g, 84%). HPLC: 14.59 min, 83% purity. ^{1}H NMR (300 MHz, $CDCl_3$) δ (s, 9H), 1.39 (d, 3H), 1.47-1.60 (m, 6H), 2.04-2.15 (m, 2H), 2.33-2.58 (m, 10H), 2.84 (s, 6H), 2.97(s, 6H), 3.03-3.11 (m,2H), 3.82 (s, 9H), 4.04-4.10 (m, 1H), 4.18-4.28 (m, 2H), 4.30-4.49 (m,1H), 5.15 (d, 1H), 6.49 (s, 1H), 7.07-7.53 (m, 15H), 7.79-7.92 (m, 2H), 8.32 (d, 1H), 8.65 (d, 1H). ^{13}C NMR (75.4 MHz, $CDCl_3$) δ 11.97, 17.23, 18.59, 24.17, 28.06, 43.24, 55.51, 76.59, 77.43, 127.02, 127.94, 128.94, 128.60, 128.73, 144.10. MS (ES) m/z 1415.0 $(M+H)^+$.

tert-butyl-N-[1-({1-[(1-{[1-(1,3-benzothiazol-2-yl)-5-{1-[(4-methoxy-2,3,6-trimethylbenzene)sulfonyl]carbamimidamido}-1-oxopentan-2-yl]carbamoyl}ethyl)carbamoyl]-3-[(triphenylmethyl)carbamoyl]propyl}carbamoyl)-4-{1-[(4-methoxy-2,3,6-trimethylbenzene)sulfonyl]carbamimidamido}butyl]carbamate (12): IBX (84 mg, 0.3 mmol, 1.2 eq) was added to a solution of compound **11** (355 mg, 0.25 mmol, 1 eq) in 2.5 mL DMSO. The resulting mixture was stirred overnight at ambient temperature. The reaction mixture was washed with water (2 x 10 mL) and brine (2 x 10 mL), dried (MgSO4), filtered through fritted glass, and concentrated *in vacuo*. The residue was purified by flash chromatography (EtOAc/Hexane 80:20 to 100:0) to give compound **12** as a yellow solid (170 mg, 48%).
HPLC: 14.94 min, 95% purity. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.29 (s, 9H), 1.41 (d, 3H), 1.44-1.87 (m, 4H), 2.00-2.15 (m, 4H), 2.30-2.61 (m, 10H), 2.87 (s, 6H), 2.95 (s, 6H), 3.03-3.36 (m, 2H), 3.79 (s, 9H), 4.03-4.15 (m, 2H), 4.30-4.99 (m, 2H), 6.46 (s, 1H), 7.03-7.20 (m, 15H), 7.41-7.52 (m, 2H), 7.86 (d, 1H), 8.16 (d, 1H). ^{13}C NMR (75.4 MHz, $CDCl_3$) δ 11.97, 17.23, 18.59, 24.17, 28.06, 43.24, 55.51, 76.59, 77.43, 127.02, 127.94, 128.60, 128.73, 144.10. MS (ES) m/z 1414.0 $(M+H)^+$.

2-(2-amino-5-carbamimidamidopentanamido)-N-(1-{[1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-oxopentan-2-yl]carbamoyl}ethyl)pentanediamide (1): A mixture of Boc-Arg(Mtr)-Gln(Trt)-Ala-Arg(Mtr)benzothiazole (206 mg, 0.15 mmol, 1 eq) and anhydrous anisole (200 µL) was placed in a teflon reaction tube of an HF apparatus and cooled to -78 °C. Anhydrous HF (15-20 mL) was condensed into the tube, and the temperature was increased to 0 °C. The reaction mixture was stirred at 0 °C for 3 h, concentrated *in vacuo*, and triturated with Et_2O (3 x 20 mL) to give a yellow solid. This solid was purified by reverse-phase HPLC, (water/MeCN/TFA 70:30:0.1), which allowed the separation of two diastereoisomers (88:12). The major diastereomer **1** was obtained as a white solid after lyophilization (56 mg, 35%). HPLC: 7.31 min, 94% purity. The second diastereomer was also isolated as a pure product (8 mg, 6.87 min, 95% purity). ^{1}H NMR (300 MHz, D_2O) δ 1.38 (d, 3H) 1.49-1.58 (m, 2H), 1.61-1.69 (m, 2H), 1.73-1.92 (m, 2H), 2.19-2.25 (m, 4H), 3.09-3.23 (m, 1H), 3.98 (t, 1H), 4.23-4.32 (m, 2H),7.56-7.68 (m, 2H), 8.11 (d, 1H), 8.19 (d, 1H). ^{13}C NMR (75.4 MHz, D_2O)δ 17.44, 24.12, 24.73, 25.02, 27.94, 33.22. 39.74, 54.66, 55.49, 74.38, 76.60, 121.76, 124.80, 125.94, 144.09, 173.49, 177.48, 177.96. HRMS calcd for $C_{27}H_{42}N_{12}O_5S$: 647.3200; found: 647.3228 $(M+H)^+$.

2-(2-amino-5-carbamimidamidopentanamido)-N-(1-{[1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-hydroxypentan-2-yl]carbamoyl}ethyl)pentanediamide (2): the same

method as compound **1** was used to synthesize compound **2**, which was obtained as a white solid (7.2 mg, 30%). HPLC: 6.08 min, 94%. ^{1}H NMR (300 MHz, D_2O). δ 1.07 (d, 3H), 1.43-1.61 (m, 4H), 1.70-1.99 (m, 4H), 2.15-2.29 (m,4H), 3.00-3.18 (m,1H), 3.93 (t, 1H), 4.04-4.13 (m, 1H), 4.14-4.24 (m, 1H), 5.02 (d, 1H), 7.39-7.55 (m, 2H), 7.89 (d, 1H), 7.97-8.00 (m,1H).^{13}C NMR (75.4 MHz, D_2O) δ 17.62,23.37,24.37,26.86,27.94, 30.82, 40.33, 52.50, 53.07, 73.44, 172.36. HRMS calcd for $C_{27}H_{44}N_{12}O_5S$: 649.3351; found: 649.3366 $(M+H)^+$.

N-(1-{[6-amino-1-(1,3-benzothiazol-2-yl)-1-oxohexan-2-yl]carbamoyl}ethyl)-2-(2-amino-5-carbamimidamidopentanamido)pentanediamide (3): The same method as compound **1** was used to synthesize compound **3**, which was obtained as a white solid (6 mg, 42%). HPLC: 5.11 min, 91%. ^{1}H NMR (300 MHz, D_2O) δ 1.36 (d, 6H), 1.42-1.54 (m,4H), 1.61-1.80 (m, 4H), 1.98-2.12 (m, 4H), 2.21-2.31 (m, 4H), 2.94 (t, 1H), 3.84-3.94 (m, 1H), 4.23-4.32 (m, 1H), 4.34-45 (m, 1H), 7.59-7.70 (m, 2H), 8.11 (m, 1H), 8.19 (d, 1H).^{13}CNMR (75.4 MHz, D_2O) δ 17.64, 22.42, 26.31, 26.80, 30.35, 30.94, 39.06, 40.20, 52.44, 54.98, 122.96, 124.89. HRMS calcd for $C_{27}H_{42}N_{10}O_5S$: 310.1603; found: 310.1603 ($M2H^+/2$).

2-amino-N-(1-{[1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-oxopentan-2-yl]carbamoyl}ethyl)pentanediamide (4): The same method as compound **1** was used to synthesize compound **4**, which was obtained as a white solid (10.0 mg, 30%). The second diastereomer was also isolated as a pure product (2.1 mg, 6.38 min, 93% purity). HPLC: 7.31 min, 94% purity. ^{1}H NMR (300 MHz, D_2O) δ 1.36 (d, 3H) 1.58-1.62 (m, 2H), 1.91-1.96 (m, 2H), 2.11 (d, 2H), 2.20- 2.25 (m, 2H), 3.17 (d, 2H), 3.97(d, 1H), 4.14-4.25 (m, 1H), 4.35 (q, 1H), 7.43-7.57 (m, 2H), 8.09 (d, 1H), 8.14 (d,1H). ^{13}C NMR (75.4 MHz, D_2O) δ 16.47, 24.62, 26.44, 29.89, 30.22, 40.61, 52.13, 122.52, 122.89, 124.85, 127.81, 136.71, 152.61, 156.66, 163.98, 174.58, 176.80, 193.38. HRMS calcd for $C_{27}H_{42}N_{10}O_5S$: 491.2189; found: 491.2161$(M+H)^+$.

N-[1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-oxopentan-2-yl]-2-acetamidopropanamide (5): The same method as compound **1** was used to synthesize compound **5**, which was obtained as a white solid (10.8 mg, 30%). The second diastereomer was also isolated as a pure product (5.1 mg, 8.04 min, 99% purity). HPLC: 8.29 min, 97% purity. ^{1}H NMR (300 MHz, D_2O) δ 1.29 (d, 3H), 1.52-1.63 (m, 2H), 1.77-1.82 (m, 2H), 1.86 (s, 3H), 3.15 (t, 2H), 4.07 (d, 1H), 4.21 (q, 1H), 7.43-7.59 (m, 2H), 8.06 (d, 1H), 8.12 (d,1H).^{13}C NMR (75.4 MHz, D_2O) δ 16.51, 21.49, 24.34, 27.56, 40.36, 49.93, 55.22, 122.85, 124.77, 125.12, 136.66, 152.53, 156.66, 163.92, 173.99, 175.41, 193.26. HRMS calcd for $C_{18}H_{24}N_6O_3S$: 405.1709; found: 405.1683 $(M+H)^+$.

N-[1-(1,3-benzothiazol-2-yl)-5-carbamimidamido-1-oxopentan-2-yl]acetamide (6): The same method as compound **1** was used to isolate compound **6**, which was obtained as a white solid (6.2 mg, 48%). HPLC: 8.22 min, 99% purity. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.52-1.68 (m, 2H), 1.97-2.08 (m, 2H), 2.11(s, 3H), 3.16 (t, 2H), 5-35-5.50 (m, 1H), 7.43-7.59 (m, 2H), 8.08 (d, 1H), 8.17 (d, 1H).^{13}C NMR (75.4 MHz, $CDCl_3$) δ 21.47, 24.26, 27.76, 40.33, 55.36, 122.89, 124.84, 127.83, 136.78, 148.14, 152.66, 163.80, 174.27, 193.91). HRMS calcd for $C_{15}H_{19}N_5O_2S$: 334.1332; found: 334.1340 $(M+H)^+$.

3. Biological methods

Matriptase cDNA was a generous gift from Dr. Chen-Yong Lin (Georgetown University, Washington, USA). Human matriptase-2 cDNA was a generous gift from C. López-Ótin (Universidad de Oviedo, Oviedo, Spain). Human hepsin cDNA was cloned from a human liver cDNA library from Ambion (Foster City, CA, USA). Human TMPRSS11D (HAT) cDNA was from GeneCopoeia (Rockville, MD, USA). Purified human trypsin and thrombin were from Sigma-Aldrich (Oakville, ON, Canada). AMC-coupled fluorogenic peptides and the H-Glu-Gly-Arg-chloromethylketone irreversible inhibitor were from Bachem (Torrance, CA, USA).

Purification of active human serine proteases

Matriptase, Matriptase-2 and Hepsin were expressed and purified as previously described.[1] cDNA corresponding to amino acid 45-418 of TMPRSS11D was amplified by PCR and subcloned in the pMT-BiP/V5-His vector (Life technologies, Burlington, ON, Canada) for expression and purification. Active TMPRSS11D was purified using the same method as for matriptase-2 and Hepsin.[1] Matriptase, matriptase-2, Hepsin, TMPRSS11D, trypsin and thrombin were active-site titrated with the burst titrant 4-methylumbelliferyl-p-guanidino benzoate (MUGB). All enzymatic assays for these enzymes were performed at room temperature in a reaction buffer containing 50 mM Tris-HCl, 15 mM NaCl and 500µg/ml BSA at pH 7, 4. Human soluble furin was expressed, purified, titrated and assayed as described previously.[1] Enzyme activities were monitored by measuring the release of fluorescence from AMC-coupled peptides (excitation, 360 nm; emission, 441 nm) in a FLX-800 TBE microplate reader (Bio-Tek Instruments, Winooski, VT, USA) at room temperature.

Inhibition assays and K_i determination

To evaluate inhibition potency of the benzothiazoles compounds towards proteases, 1 nM of enzymes was added to the reaction buffer containing 0 nM, 10 nM or 10µM of the different ketobenzothiazoles (0,4 µM, 4 µM and 400 µM for the reduced form of RQAR-ketobenzothiazole) and 100µM of Boc-Arg-Arg-Val-Arg-AMC for furin and 100µM of Boc-Gln-Ala-Arg-AMC for the other proteases. If substantial inhibition occurred using a ratio of I/E < 10, compounds were treated as tight-binding inhibitors [2]. For Ki determination of tight-binding inhibitors, enzymes diluted to concentration ranging from 0, 25 to 1 nM were pre-equilibrated with appropriate dilutions of the compounds for 15 minutes at room temperature. Residual enzyme activity was then measured by following the hydrolysis of the fluorogenic substrate Boc-Gln-Ala-Arg-AMC for 30 minutes at room temperature. Data from at least three independent experiments were averaged and residual velocities were plotted as a function of inhibitor concentration. Data were fitted by non-linear regression analysis to the Morrison Ki equation (1):

$$vi/vo = vo*(1-((((Eo + Io + Kiapp) - (((Eo + Io + Kiapp)2) -4*Eo*Io)½)) / (2*Eo))), \quad (1)$$

where vo and vi are the steady-state rates of substrate hydrolysis in the absence and presence of inhibitor, respectively, Eo, the initial concentration of enzyme, Io, the initial concentration of inhibitor and Kiapp the substrate-dependent equilibrium dissociation constant. The substrate-independent constant Ki was calculated using the equation (2):

$$Kiapp = (Ki*(1+(S/Km))), \quad (2)$$

where S is the initial concentration of substrate and Km the Michaelis–Menten constant for the enzyme–substrate interaction.

If substantial inhibition occurred only when using a ratio of I/E > 10, compounds were treated as classical reversible inhibitors. In this case, inhibition assays used for Ki determination were initiated by the addition of 1 nM of protease to a reaction mixture containing the inhibitor and the fluorogenic substrate. Data generated from plots of enzyme velocity versus substrate concentration at several fixed inhibitor concentration were fitted by non-linear regression analysis to equations describing competitive, uncompetitive, noncompetitive and mixed model inhibition[3]. For inhibition of matriptase by R-Benzo and thrombin by RQAR-Benzo, the mixed model inhibition was the preferred model as determined by the goodness of fit (R2) and the Akaike's information criteria (AICc) and therefore used for Ki determination. All non-linear regression and statistical analysis were performed using GraphPad Prisms version 5.04 for Windows (GraphPad Software, San Diego, CA, USA).

Dilution recovery experiments

Dissociation of the enzyme-inhibitor (EI) complex was investigated using dilution recovery experiments. High concentration of matriptase (0,5 μM) with variable concentration of A) RQAR-Benzo or B) EGR-CMK was incubated for 20 minutes at room temperature in reaction buffer to allow for the formation of the EI complex. Complexes were rapidly (1:2000) diluted in a reaction buffer containing 400μM of Boc-Gln-Ala-Arg-AMC (~7 times Km). Activity was continuously recorded for 75 minutes and fluorescence plotted as a function of time.

4. Modelling procedures

Preparation of ligand structures

The 3D coordinates of the ligand with correct chirality were generated with Open Babel 2.30[4] using SMARTS input format.[5] The coordinates were export in PDB format and partial charges were calculated in Open Babel with the Partial Equalization of Orbital Electronegativities (PEOE) algorithm by Gaisteiger & Marsilli. [6] The general AMBER force field (gaff) [7] parameters generated in Antechamber[8] and convert to GROMACS[9-12] formats by ACPYPE[13] were used for the ligand.

Preparation of the protein target

The structure of Matriptase (PDB code 3NCL) was obtained from the RCSB Protein Data Bank (www.pdb.org)[14]. Hydrogens and water molecules were removed from the structure. Flexible residues were located based on the detection of residues seen in different conformation in the binding sites of other Matriptase structures (PDB: 1EAW, 1EAX, 2GV6, 2GV7 and 3BN9). As a result, during the docking, Asp^{705}, Phe^{708}, Tyr^{755}, Gln^{782}, Gln^{783}, Gln^{802}, and Asp^{828} were set as flexible residues during docking simulations.

Docking parameter

The ligand was docked in the protein with our in house developed docking program FlexAID.[15] FlexAID uses a genetic algorithm to optimize the relative position of the ligand with respect to the protein as well as dyhedral angles of the ligand and flexible residues (optimized using an exhaustive search within library of side chain conformations at each step of the simulation). The population size and number of generation were both set at 300 used in the genetic algorithm. These values were seen empirically to offer a good balance between the precision and length of simulations

(Gaudreault *et al.*, personal communication). The ligand was set as semi-flexible. Arg in P1 and the warhead in P1' were set as rigid in a conformation like that found in the thrombin inhibitor (0ZE) in PDB structure 1B5G. All peptide bonds were also set as rigid to diminish the search space. The carbon making the covalent bond with the protein was fixed at a distance of 3.0 ± 0.1Å of the catalytic Ser^{805}. 100 different docking simulations were performed and rescored with a fast MM/GBSA (Molecular Mechanics/Generalised Born Surface Area) method (see below).

MM/GBSA

The MM/GBSA method is used to approximate the binding free energy (ΔG_{bind}) using equation 1:

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} \quad (1)$$
$$\Delta E_{MM} = \Delta E_{internal} + E_{electrostatic} + E_{vdW} \quad (2)$$
$$\Delta G_{solv} = \Delta G_{GB} + \Delta G_{SA} \quad (3)$$

where ΔE_{MM} is the difference in energy between the protein-ligand complex and the protein and ligand alone calculate with basic molecular mechanic equations (2). ΔG_{solv} is the difference in solvation free energy calculated with GB implicit solvation model (ΔG_{GB}) and the nonpolar solvation free energy calculated with solvent accessible surface area (ΔG_{SA}).

All protein-ligand complexes were first energy minimised with GROMACS 4.5.1[9-12] using the conjugate gradient algorithm with AMBER03[16] force field in a TIP3P [17] water box. All 100 different poses were then rescored with MM/GBSA using HCT model[18] of the effective Born radii. The top 5 poses were then visually inspected with PyMOL.[19]

5. References

1. Denault, J. B.; Lazure, C.; Day, R.; Leduc, R. Comparative characterization of two forms of recombinant human SPC1 secreted from Schneider 2 cells. *Protein. Expr. Purif.* **2000**, 19, 113-124.

2. Bieth, J. G. Theoretical and pratical aspects of proteinase inhibition kinetics. *Methods. Enzymol.* **1995**, 248, 59-84.

3. Copeland, R. A Pratical Introduction to Structure,Mechanism and Data Analysis. Wiley-VCH, New-York, 2000, chapter 10, pp 1-397.

4. Open Babel: The Open Source Chemistry Toolbox. http://openbabel.org/ accessed on 07/22/2011.

5. SMARTS- A Language for Describing Molecular Patterns. http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html accessed on 07/2/2011.

6. Gasteiger, J.; M. Marsili, M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*. **1980**, 36, 3219-3228

7. Wang, J.; Wolf, R. M.; Caldwell, J. W.;Kollman, P. A.; Case, D. A. Development and testing of a general AMBER force field. *Journal of Computational Chemistry*. **2004**, 25, 1157-1174.

8. Wang, J.; Wang, W.; Kollman P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*. **2006**, 25, 247-260.

9. Berendsen, H.J.C.; van der Spoel, D.; van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **1995**, 91, 43–56.

10 Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, 7, 306-317.

11. Van der Spoel, D.; Lindahl, E.; Hess, E.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, 26, 1701-1718.

12. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, 4, 435-447.

13. Sousa Da Silva, A. W.; Vranken, W. F.; Laue, E. D. ACPYPE - AnteChamber PYthon Parser interfacE. Manuscript to be submitted.

14. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Research*. **2000**, 28, 235-242.

15. Najmanovich, R. J. Side chain flexibility upon ligand binding: Docking predictions and statistical analysis. PhD Thesis, 2003, Weizmann Institute of Science.

16. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, 24, 1999-2012.

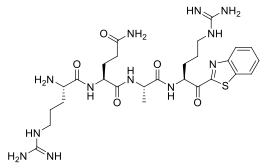
17. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, 79, 926–935.

18. Hawkins, D.; Cramer, C.; Truhlar, D. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* A. **1996**,100, 19824–19839.

19. The PyMOL Molecular Graphics System, Version 1.2r1pre, Schrödinger, LLC.

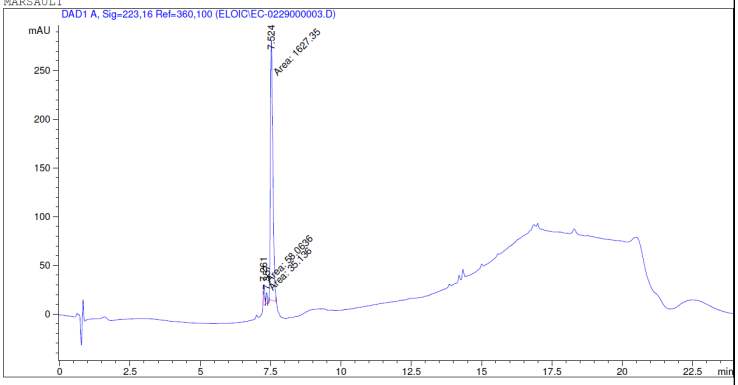
6. Chemical characterization data

Compound 1 :



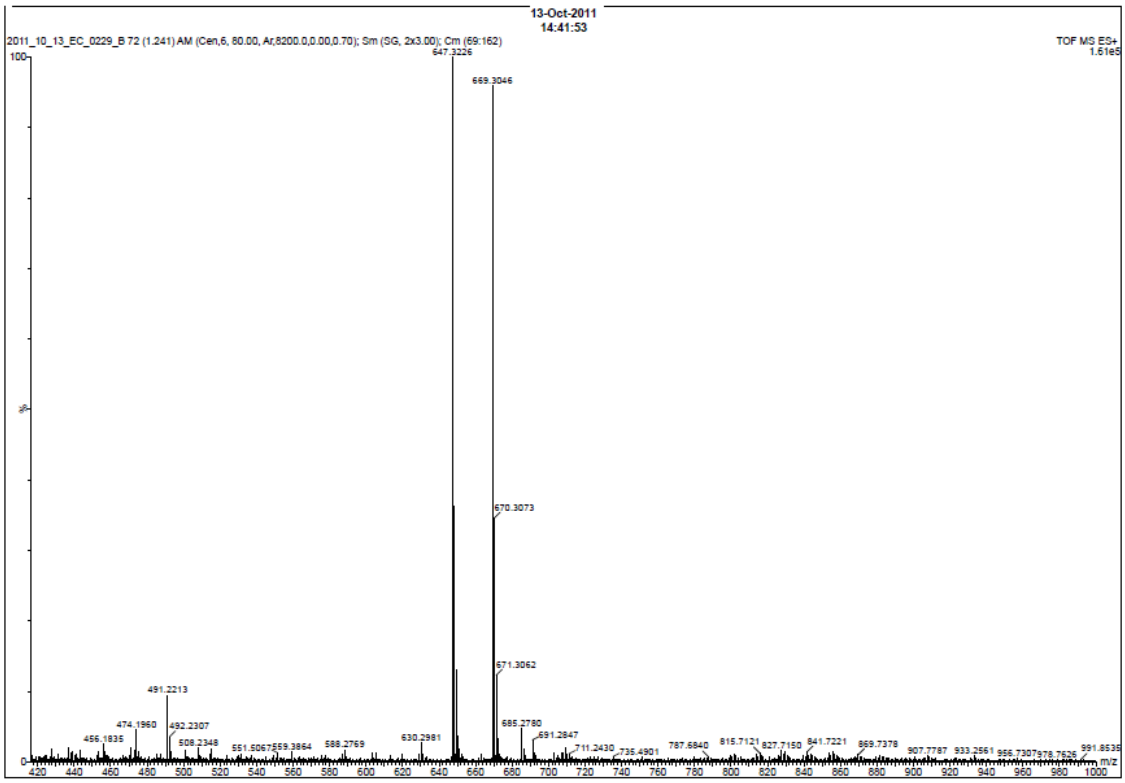
- HPLC

```
=====================================================================
Injection Date  : 5/25/2011 10:31:07 AM          Seq. Line :   3
Sample Name     : EC-0229-01B                     Location : Vial 12
Acq. Operator   : ELO                                Inj :   1
Acq. Instrument : Instrument 1                  Inj Volume : 30 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 20 µl
Acq. Method     : C:\CHEM32\1\METHODS\MARSAULT.M
Last changed    : 5/25/2011 9:38:07 AM by ELO
Analysis Method : C:\CHEM32\1\METHODS\WASHOUT.M
Last changed    : 12/9/2011 5:06:30 PM by ARNAUD
MARSAULT
```

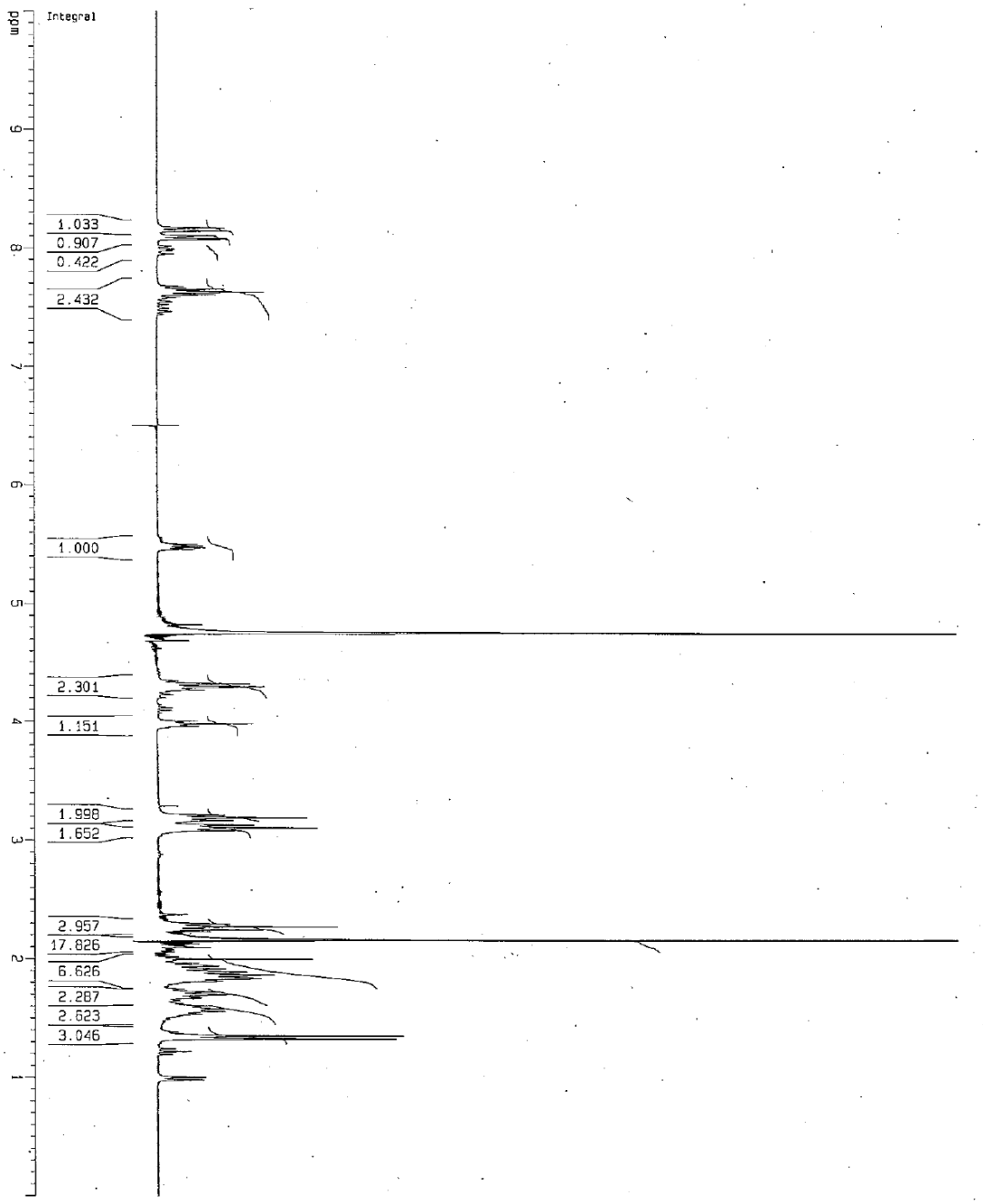



```
=====================================================================
                        Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: DAD1 A, Sig=223,16 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]       [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   7.261 MM    0.0470   58.06359   20.61059   3.3747
   2   7.357 MM    0.0528   35.13601   11.09398   2.0421
   3   7.524 MM    0.1012 1627.35071  268.02475  94.5832

Totals :                  1720.55031  299.72932
```

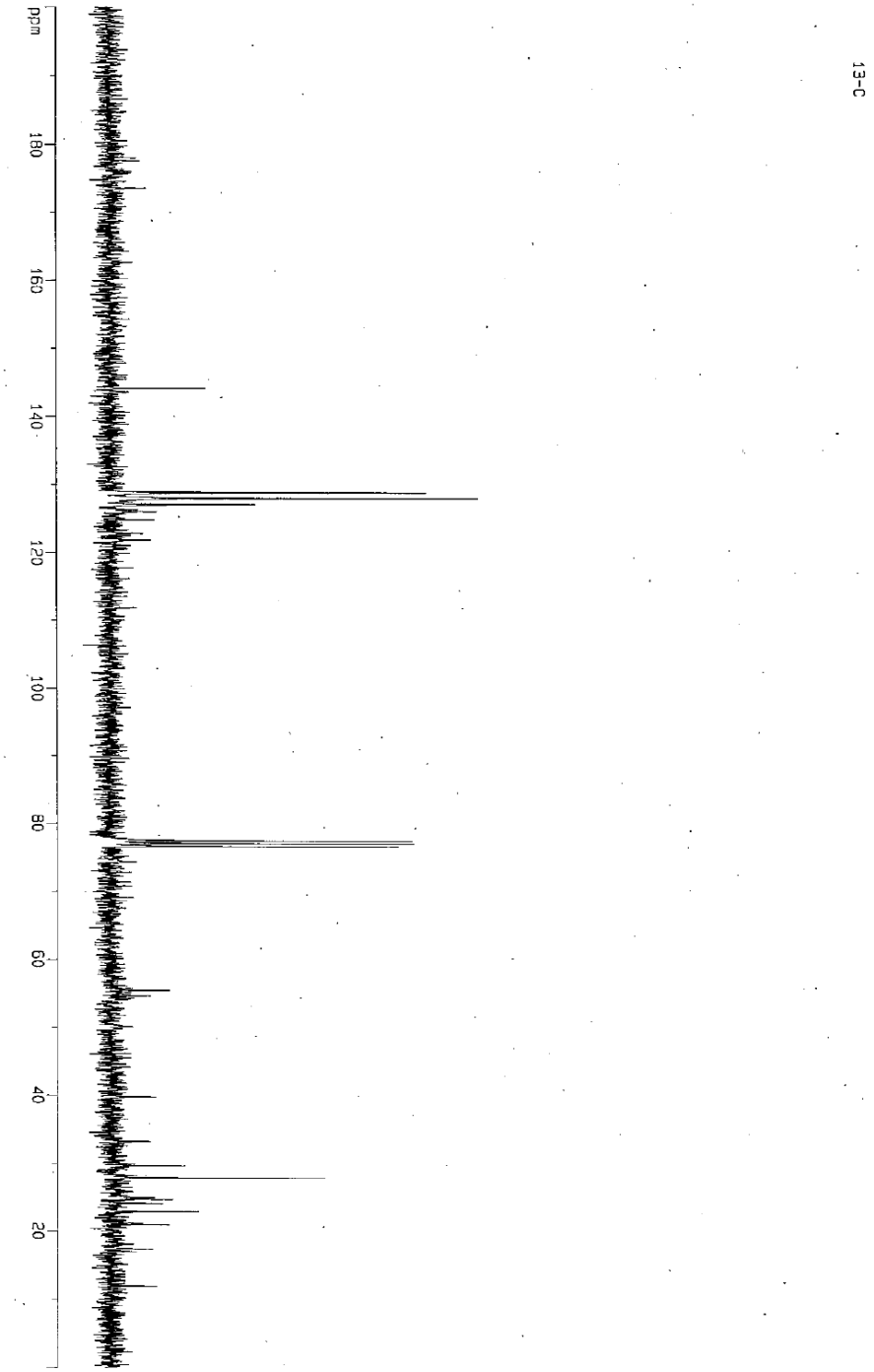
- MS


13-Oct-2011
14:41:53
2011_10_13_EC_0229_B 72 (1.241) AM (Cen,6, 80.00, Ar,8200.0,0.00,0.70); Sm (SG, 2x3.00); Cm (69:162)
TOF MS ES+
1.61e5
100
%
0
647.3226
669.3046
670.3073
671.3062
491.2213
474.1960
492.2307
456.1835
508.2348
551.5067
559.3864
588.2769
630.2981
685.2780
691.2847
711.2430
735.4901
787.6840
815.7121
827.7150
841.7221
869.7378
907.7787
933.2561
956.7307
978.7626
991.8535
420
440
460
480
500
520
540
560
580
600
620
640
660
680
700
720
740
760
780
800
820
840
860
880
900
920
940
960
980
1000
m/z

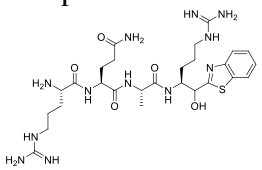
- ^{1}H NMR


Integral
ppm
1.033
0.907
0.422
2.432
1.000
2.301
1.151
1.998
1.652
2.957
17.826
6.626
2.287
2.623
3.046

- ^{13}C NMR

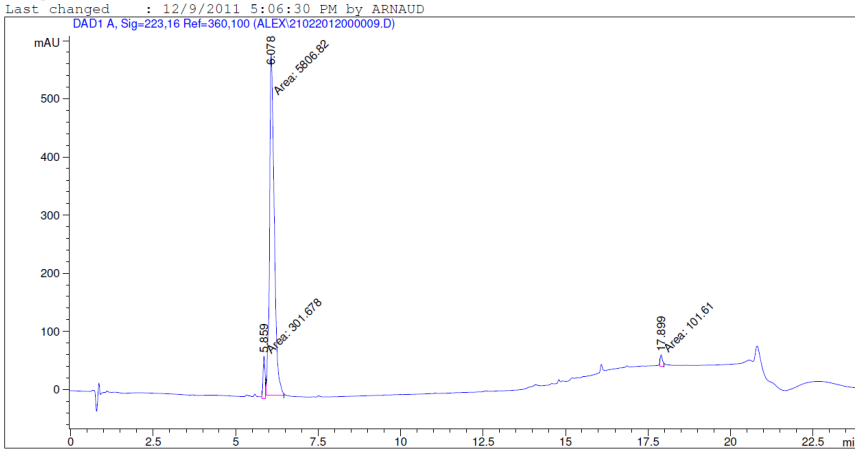


Compound 2:



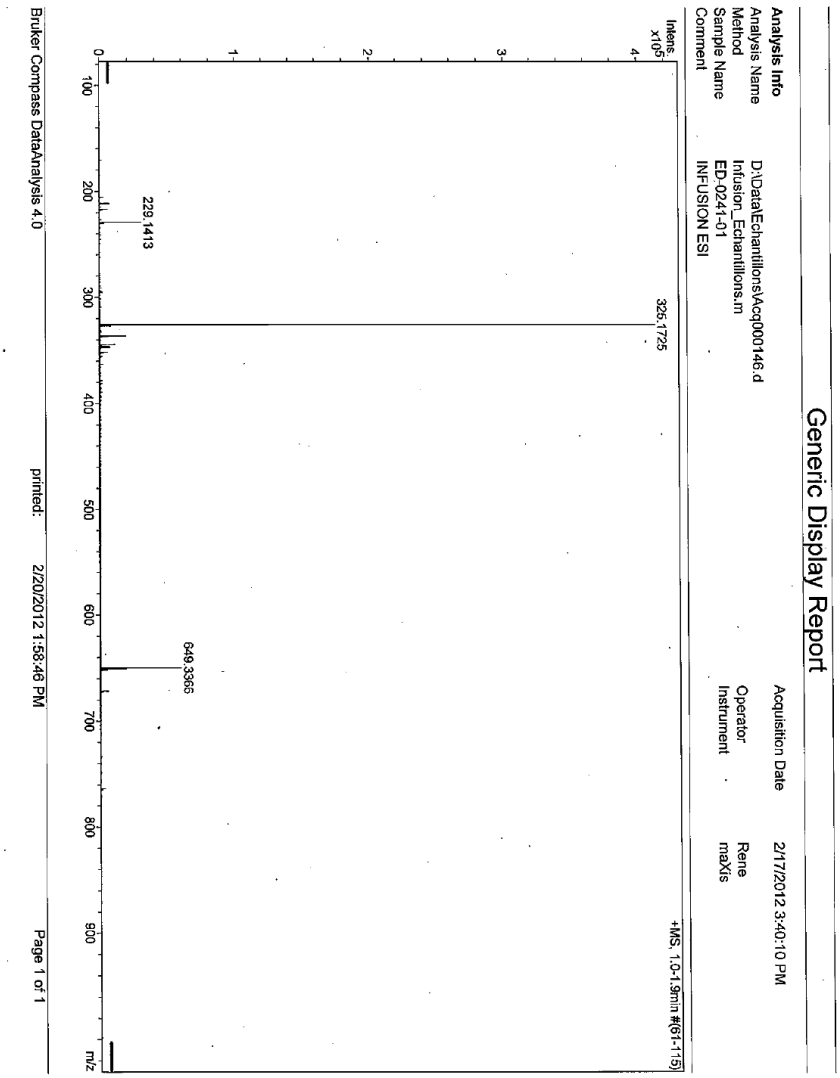
- HPLC

```
======================================================================
Injection Date  : 2/21/2012 1:26:18 PM          Seq. Line :   9
Sample Name     : EC-0241-01                     Location : Vial 27
Acq. Operator   : ALEX                              Inj :   1
Acq. Instrument : Instrument 1                Inj Volume : 30 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 10 µl
Acq. Method     : C:\CHEM32\1\METHODS\MARSAULT.M
Last changed    : 4/26/2012 11:17:12 AM by ALEX
Analysis Method : C:\CHEM32\1\METHODS\WASHOUT.M
Last changed    : 12/9/2011 5:06:30 PM by ARNAUD
```

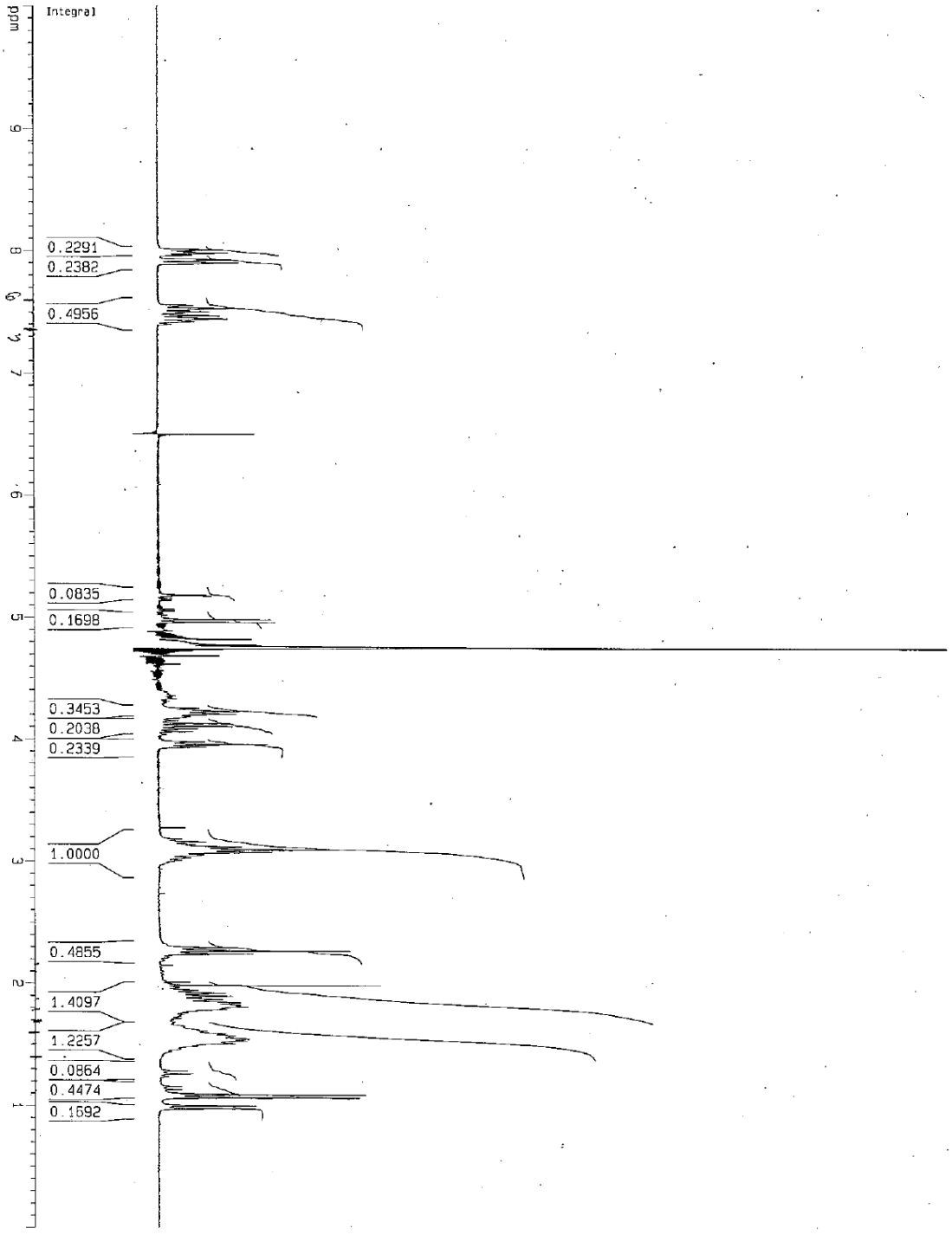



```
======================================================================
                          Area Percent Report
======================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: DAD1 A, Sig=223,16 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
   #   [min]       [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   5.859 MM    0.0687  301.67834   73.23421   4.8579
   2   6.078 MM    0.1645 5806.82227  588.50153  93.5059
   3  17.899 MM    0.0832  101.60999   20.35223   1.6362

Totals :                  6210.11060  682.08797
```

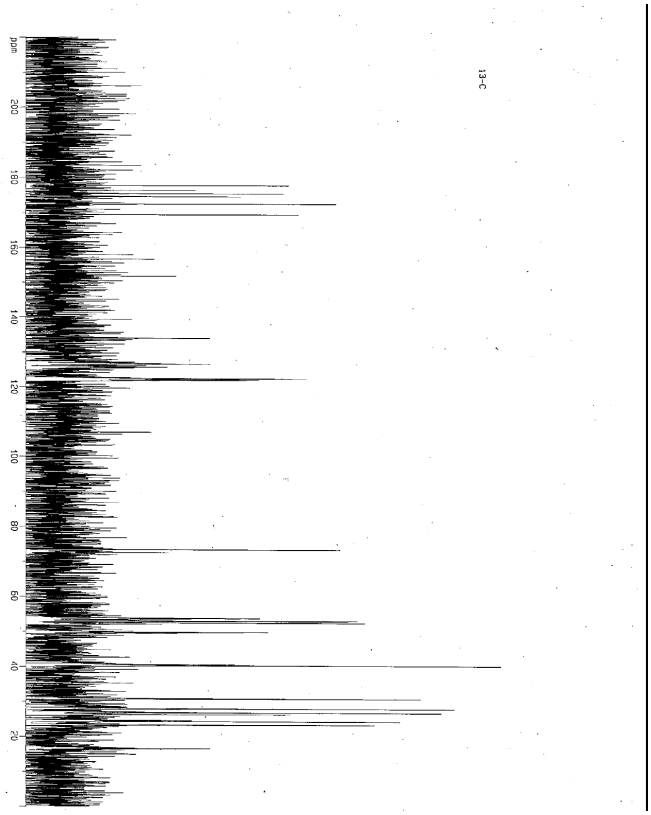
- MS


Generic Display Report
Analysis Info
Analysis Name D:\Data\Echantillons\Acq000146.d
Method Infusion_Echantillons.m
Sample Name ED-0241-01
Comment INFUSION ESI
Acquisition Date 2/17/2012 3:40:10 PM
Operator Rene
Instrument maXis
Intens.
x10^5
4
3
2
1
0
325.1725
229.1413
649.3366
+MS, 1.0-1.9min #(61-115)
100
200
300
400
500
600
700
800
900
m/z
Bruker Compass DataAnalysis 4.0
printed: 2/20/2012 1:58:46 PM
Page 1 of 1

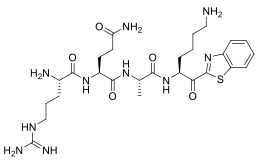
- ^{1}H NMR


Integral
ppm
0.2291
0.2382
0.4956
0.0835
0.1698
0.3453
0.2038
0.2339
1.0000
0.4855
1.4097
1.2257
0.0864
0.4474
0.1692

- ^{13}C NMR

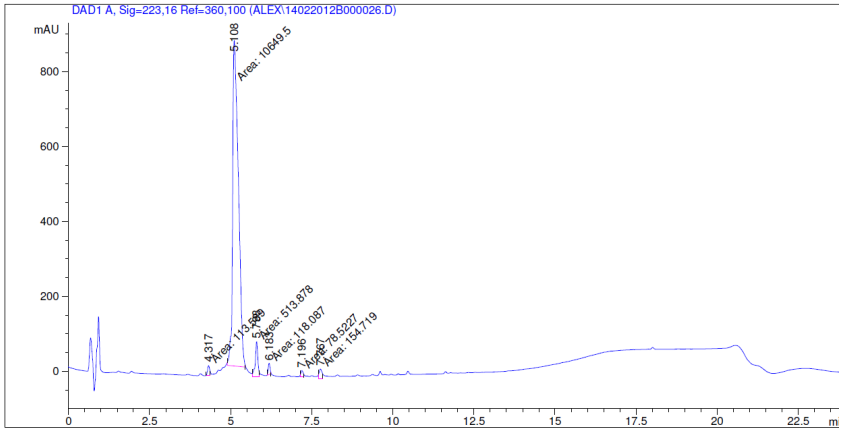


Compound 3:



- HPLC

```
======================================================================
Injection Date  : 2/14/2012 10:19:49 PM          Seq. Line :  26
Sample Name     : EC-0165-02C                     Location : Vial 10
Acq. Operator   : ALEX                               Inj :   1
Acq. Instrument : Instrument 1                 Inj Volume : 30 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 20 µl
Acq. Method     : C:\CHEM32\1\METHODS\MARSAULT.M
Last changed    : 4/26/2012 11:17:12 AM by ALEX
Analysis Method : C:\CHEM32\1\METHODS\1MLSURGITH.M
Last changed    : 2/21/2012 8:02:57 AM by ELO
                  (modified after loading)
MARSAULT
```

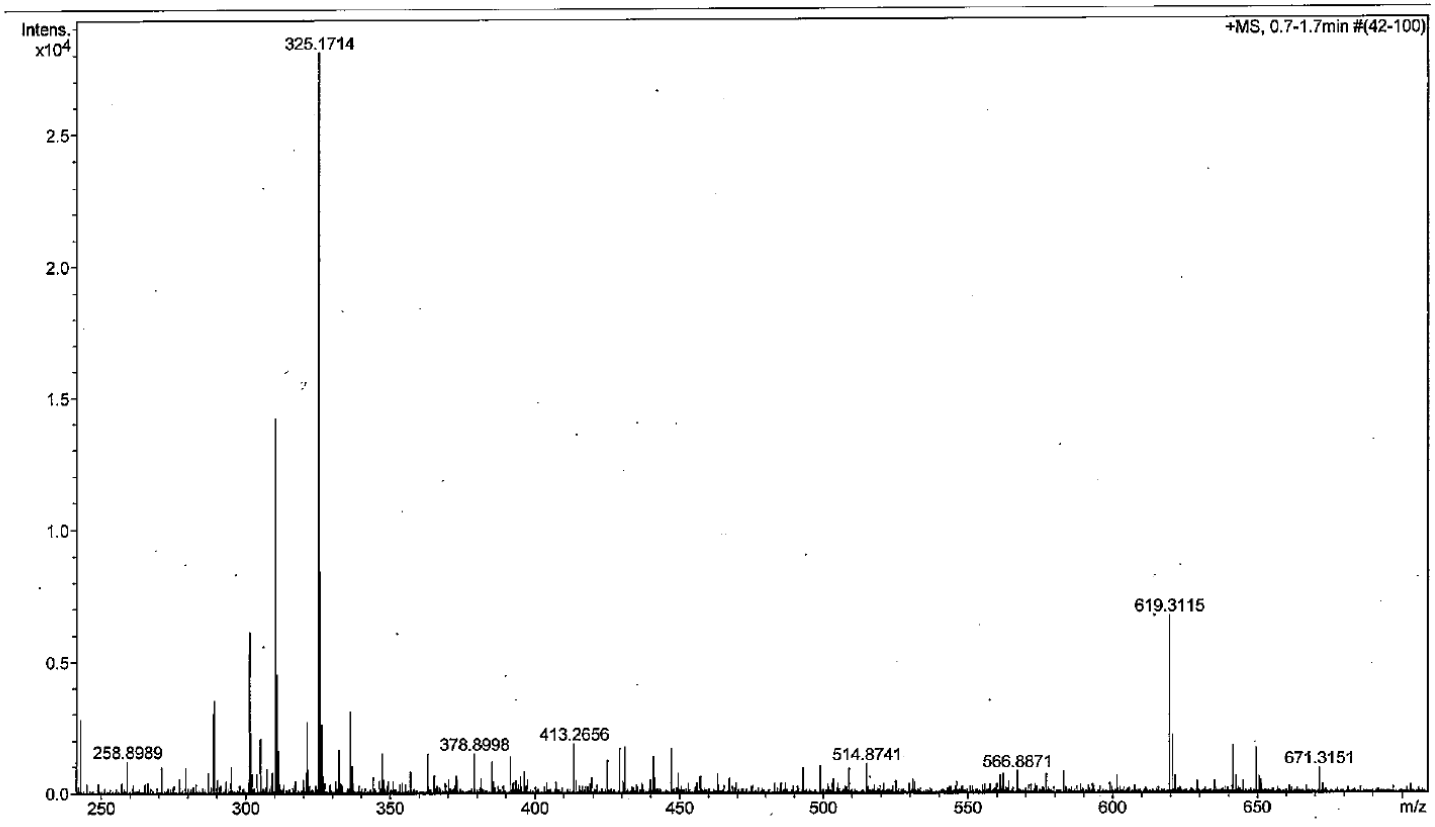



```
======================================================================
                        Area Percent Report
======================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: DAD1 A, Sig=223,16 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]    [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   4.317 MM    0.0727  113.58930   26.05217   0.9768
   2   5.108 MM    0.2042 1.06495e4   869.19843  91.5826
   3   5.798 MM    0.0904  513.87775   94.72636   4.4192
   4   6.183 MM    0.0645  118.08740   30.49323   1.0155
   5   7.196 MM    0.0762   78.52269   17.17111   0.6753
   6   7.767 MM    0.1003  154.71906   25.70735   1.3305

Totals :                  1.16283e4  1063.34865
```

Generic Display Report

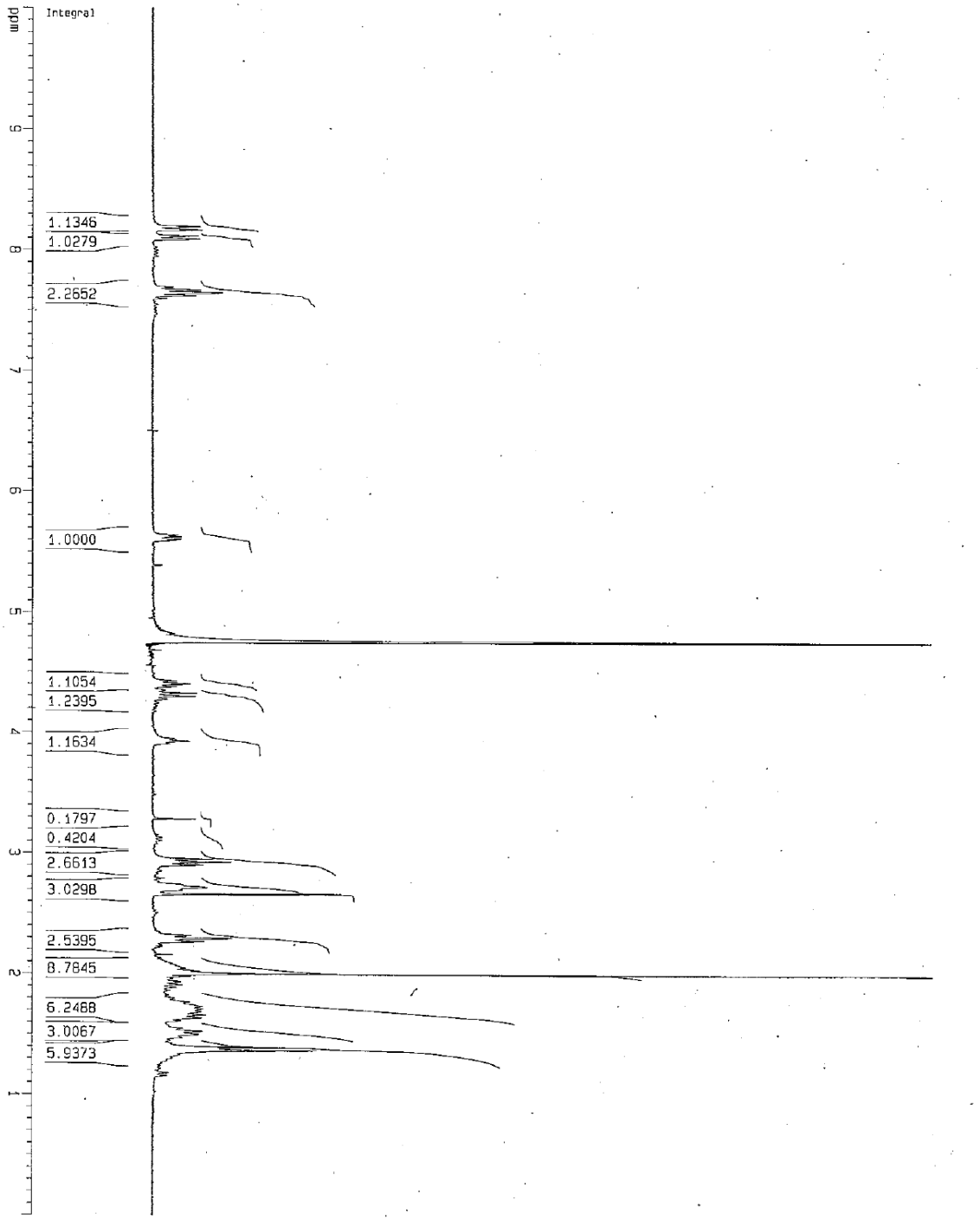
Analysis Info

Analysis Name	D:\Data\Echantillons\Acq000149.d	Acquisition Date	2/21/2012 4:35:44 PM
Method	Infusion_Echantillons.m	Operator	Rene
Sample Name	EC-0165-02	Instrument	maXis
Comment	INFUSION ESI		

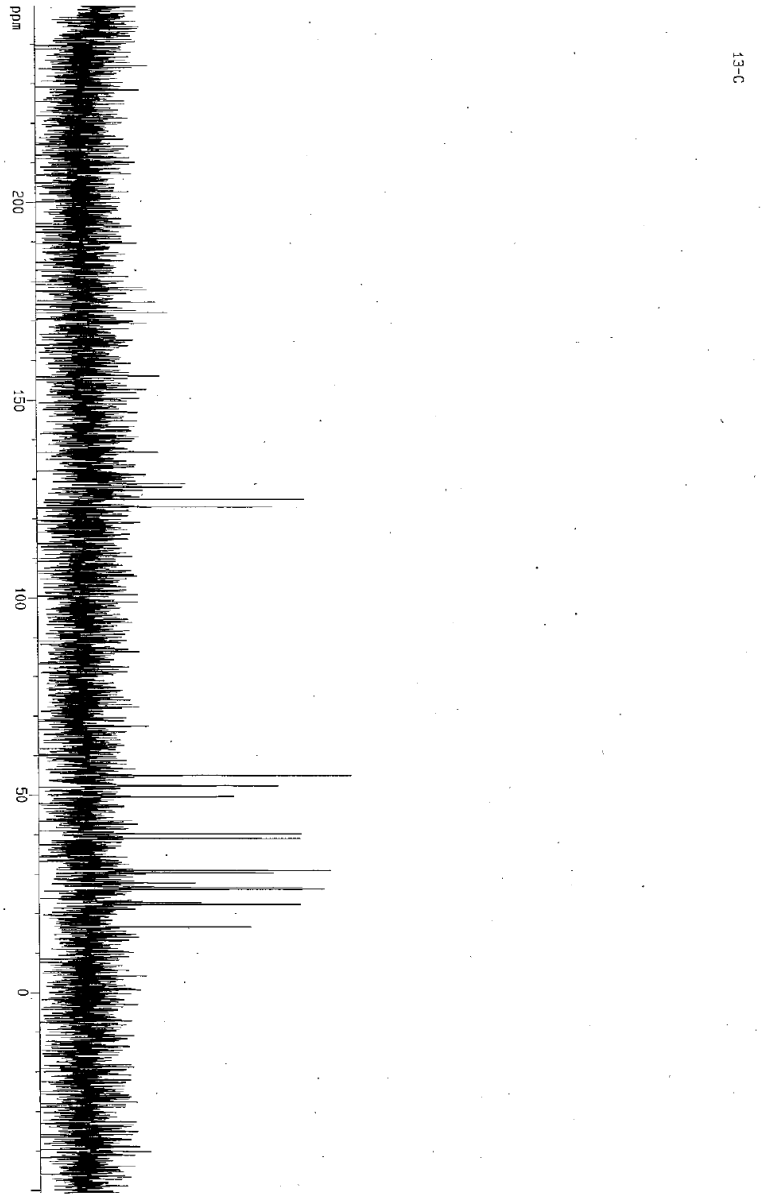


- MS

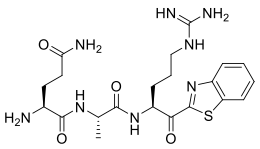
- ^{1}H NMR


ppm
Integral
1.1346
1.0279
2.2652
1.0000
1.1054
1.2395
1.1634
0.1797
0.4204
2.6613
3.0298
2.5395
8.7845
6.2488
3.0067
5.9373

- ^{13}C NMR

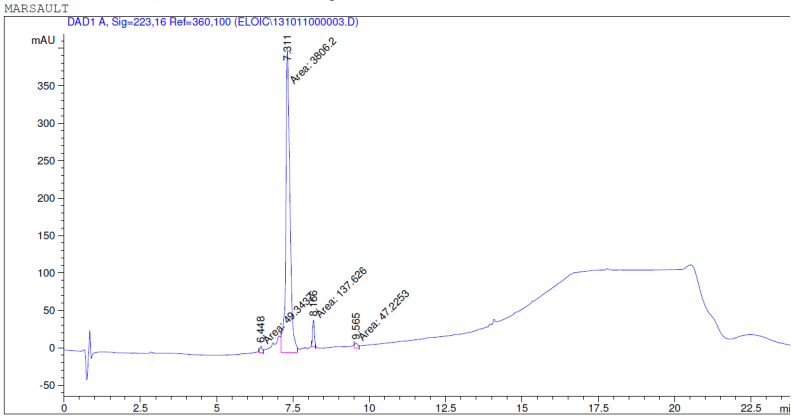


Compound 4:



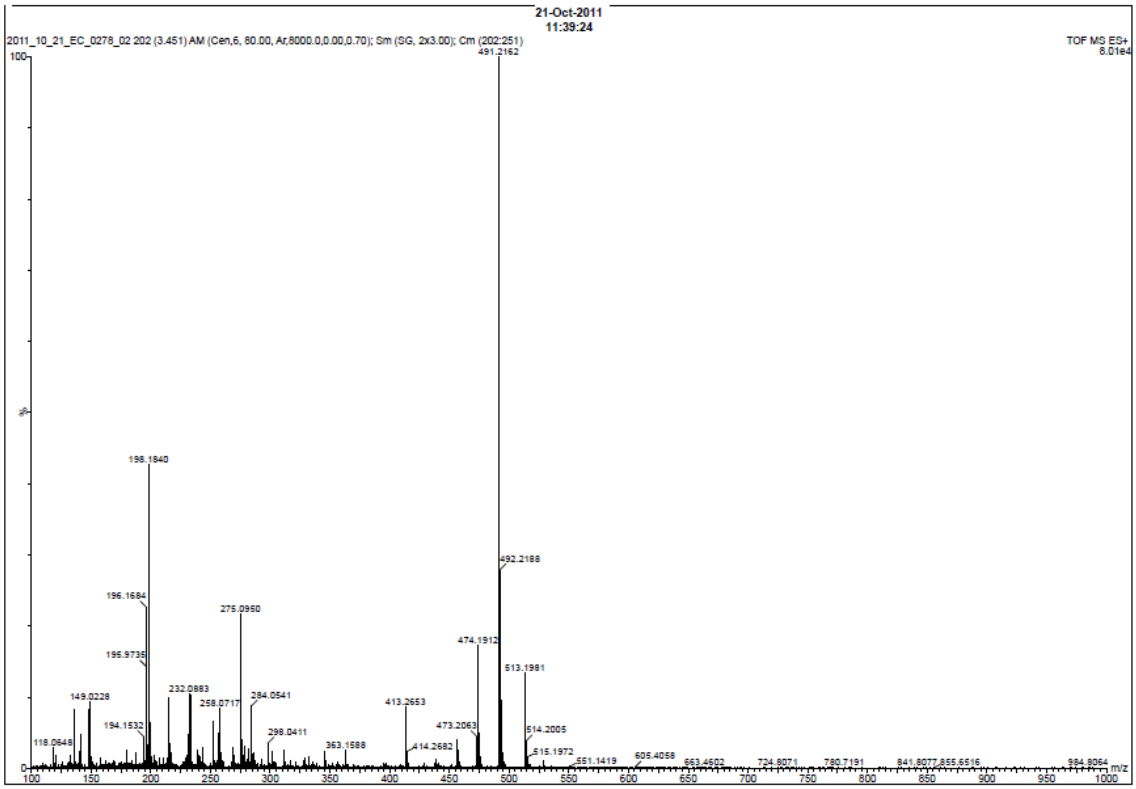
- HPLC

```
=====================================================================
Injection Date  : 10/13/2011 9:58:57 AM          Seq. Line :   3
Sample Name     : EC-0278-01 DIA 2                Location : Vial 62
Acq. Operator   : ELO                                Inj :   1
Acq. Instrument : Instrument 1                 Inj Volume : 30 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 20 µl
Acq. Method     : C:\CHEM32\1\METHODS\MARSAULT.M
Last changed    : 10/5/2011 5:07:28 PM by ELO
Analysis Method : C:\CHEM32\1\METHODS\1MLSURGITH.M
Last changed    : 2/21/2012 8:02:57 AM by ELO
                  (modified after loading)
MARSAULT
```

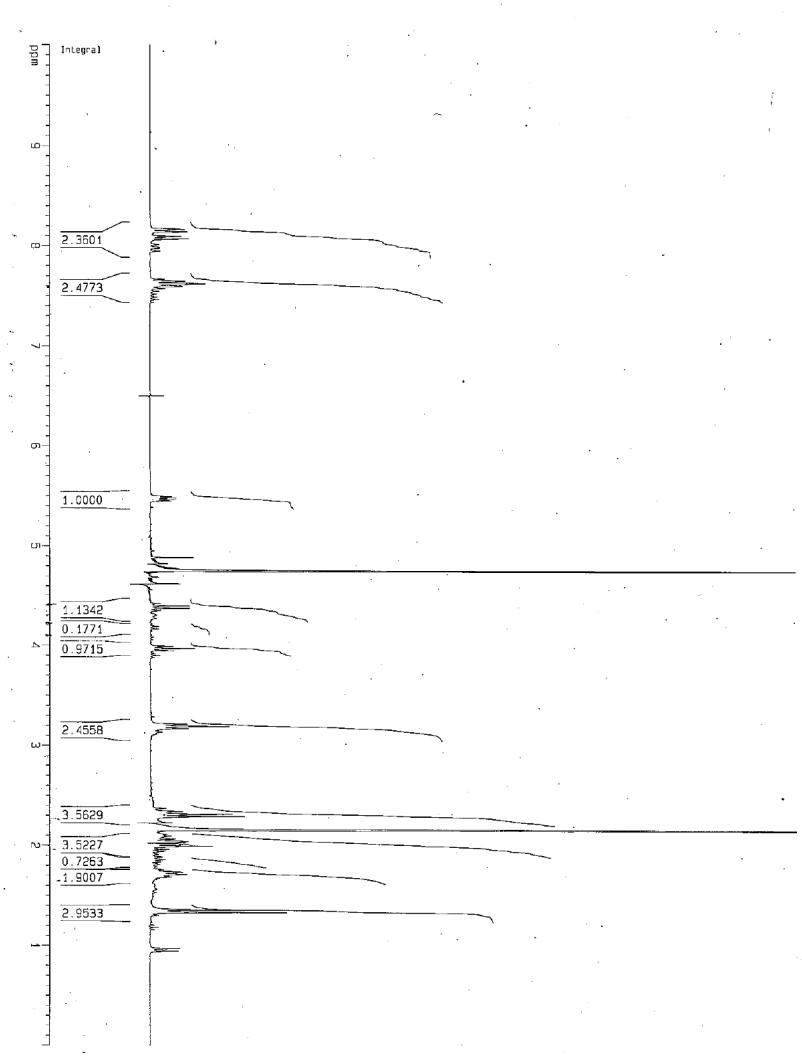



```
=====================================================================
                        Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: DAD1 A, Sig=223,16 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   6.448 MM    0.0961   49.34365    8.55804   1.2213
   2   7.311 MM    0.1570 3806.20068  404.00662  94.2037
   3   8.166 MM    0.0645  137.62650   35.54226   3.4063
   4   9.565 MM    0.1150   47.22528    6.84677   1.1688

Totals :                  4040.39611  454.95369
```

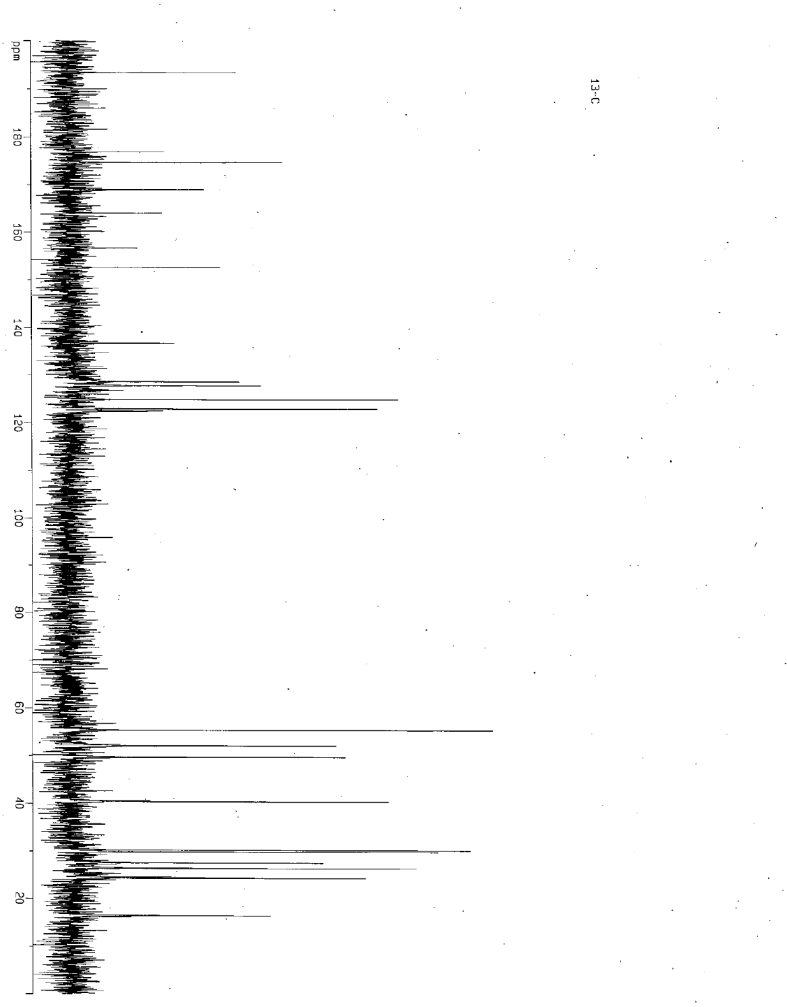
- MS


21-Oct-2011
11:39:24
2011_10_21_EC_0278_02 202 (3.451) AM (Cen,6, 80.00, Ar,8000.0,0.00,0.70); Sm (SG, 2x3.00); Cm (202:251)
TOF MS ES+
8.01e4
100
%
0
491.2162
198.1840
492.2188
196.1684
275.0950
474.1912
195.9735
513.1981
232.0883
149.0228
284.0541
258.0717
413.2653
194.1532
473.2063
514.2005
298.0411
118.0648
363.1588
414.2682
515.1972
551.1419
605.4058
663.4603
724.8071
780.7191
841.8077
855.5516
984.8064
m/z
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000

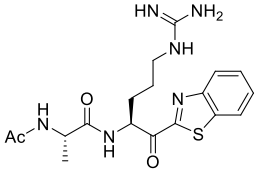
- ^{1}H NMR


Integral
2.3601
2.4773
1.0000
1.1342
0.1771
0.9715
2.4558
3.5629
3.5227
0.7263
1.9007
2.9533

- ^{13}C NMR

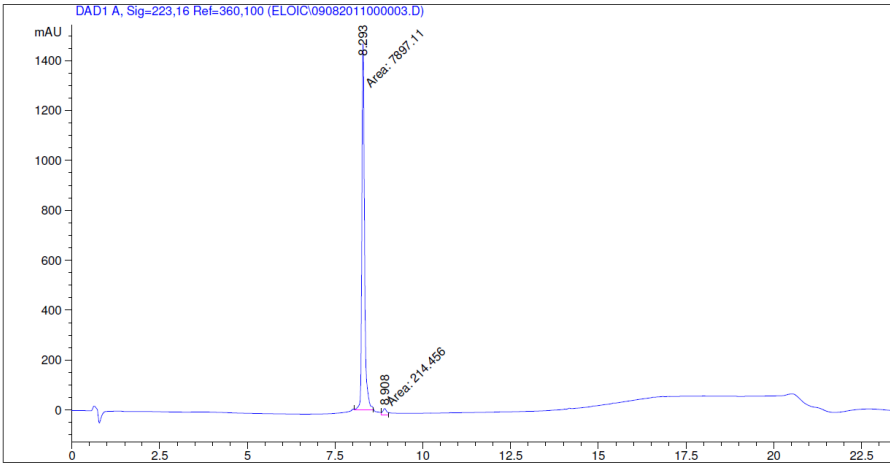


Compound 5:



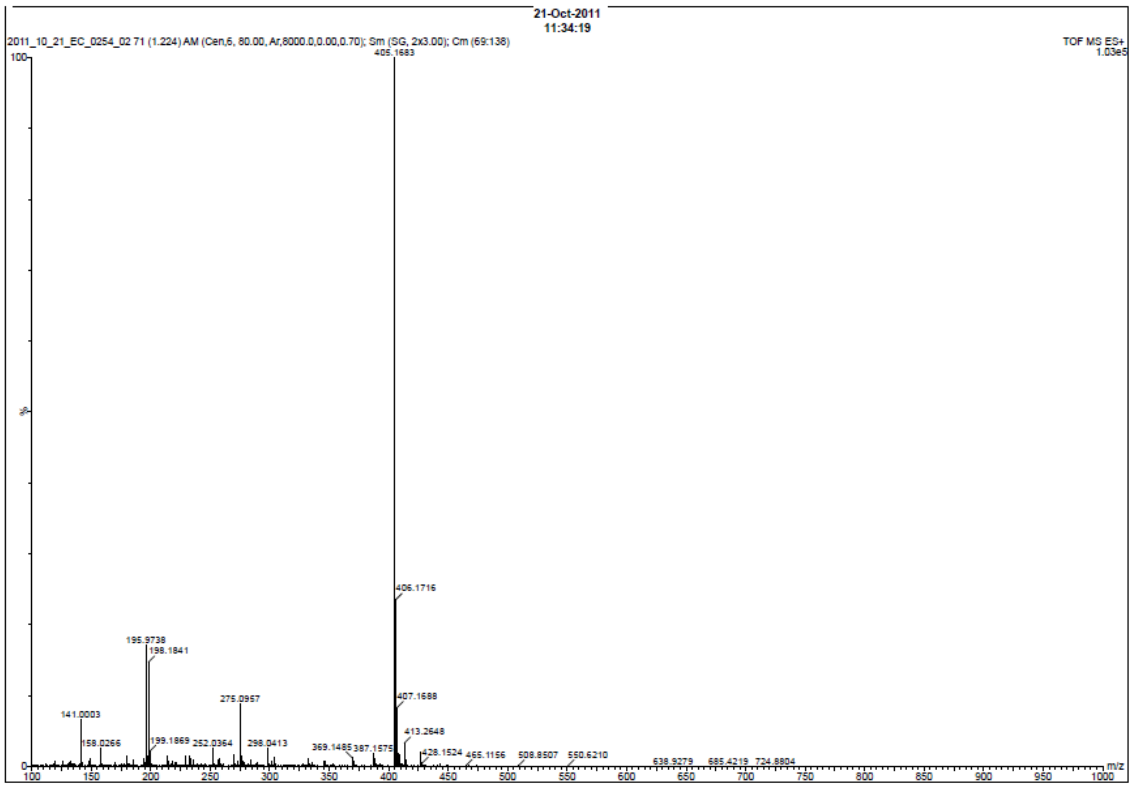
- HPLC

```
=====================================================================
Injection Date  : 8/9/2011 11:31:33 AM          Seq. Line :   3
Sample Name     : EC-0254-01B                    Location : Vial 22
Acq. Operator   : ELO                               Inj :   1
Acq. Instrument : Instrument 1                 Inj Volume : 30 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 20 µl
Acq. Method     : C:\CHEM32\1\METHODS\MARSAULT.M
Last changed    : 6/22/2011 10:51:16 AM by Nadon
Analysis Method : C:\CHEM32\1\METHODS\1MLSURGITH.M
Last changed    : 2/21/2012 8:02:57 AM by ELO
                  (modified after loading)
MARSAULT
```

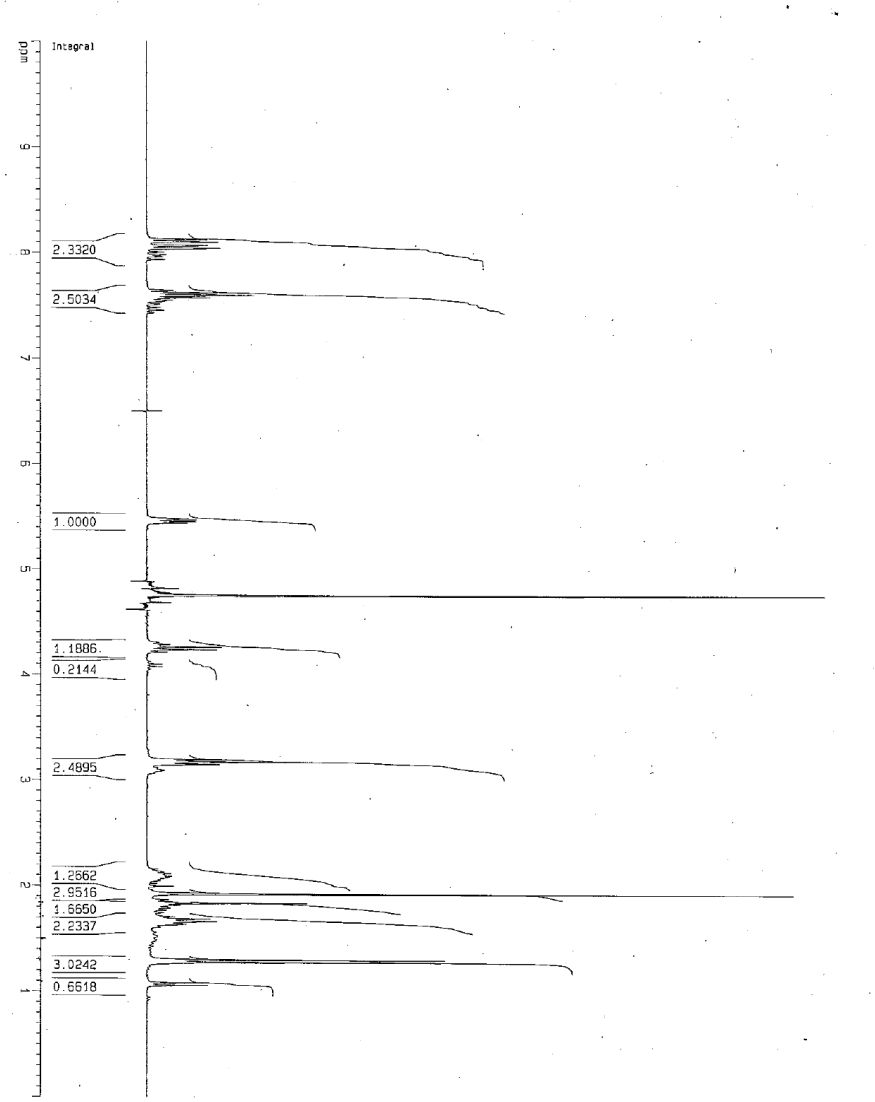



```
=====================================================================
                        Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: DAD1 A, Sig=223,16 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   8.293 MM    0.0896 7897.10742 1468.61414  97.3562
   2   8.908 MM    0.1373  214.45638   26.02627   2.6438

Totals :                  8111.56380 1494.64040
```

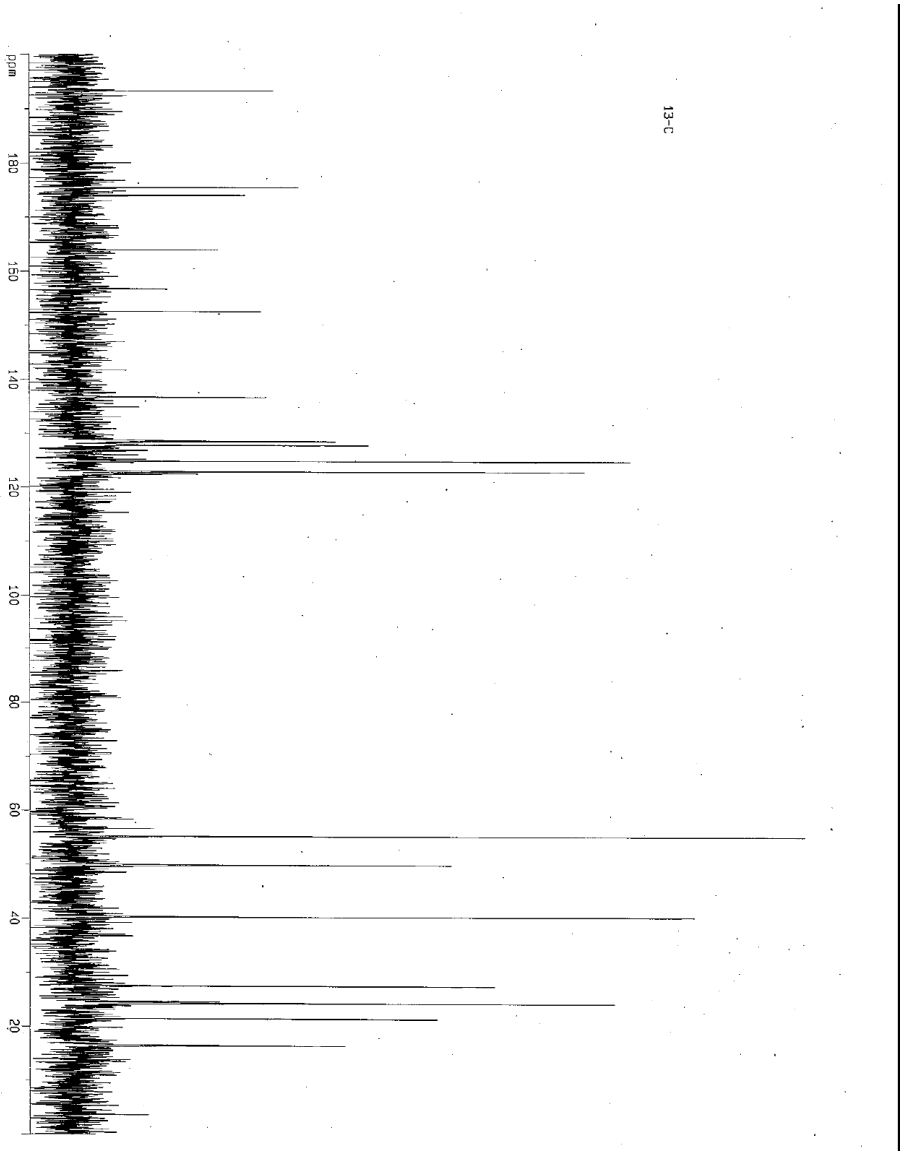
- MS


21-Oct-2011
11:34:19
2011_10_21_EC_0254_02 71 (1.224) AM (Cen,6, 80.00, Ar,8000.0,0.00,0.70); Sm (SG, 2x3.00); Cm (69:138)
TOF MS ES+
1.03e5
405.1683
406.1716
195.9738
198.1841
275.0957
407.1688
141.0003
413.2648
158.0266
199.1869
252.0364
298.0413
369.1485
387.1575
428.1524
465.1156
508.8507
550.6210
636.9279
685.4219
724.8804
%
m/z

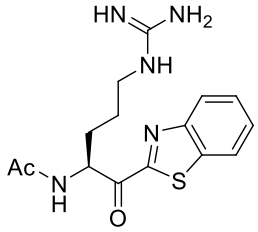
- ^{1}H NMR


Integral
2.3320
2.5034
1.0000
1.1886
0.2144
2.4895
1.2562
2.9516
1.6650
2.2337
3.0242
0.6618

- 13C NMR

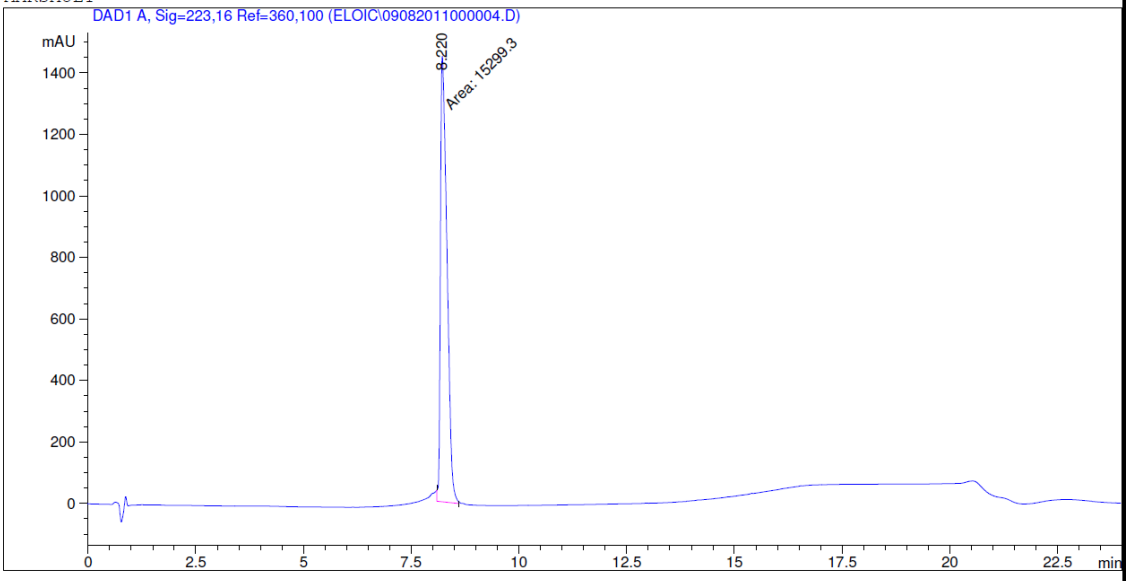


Compound 6



- HPLC

```
=====================================================================
Injection Date  : 8/9/2011 11:57:01 AM          Seq. Line :   4
Sample Name     : EC-0253                        Location : Vial 23
Acq. Operator   : ELO                              Inj :   1
Acq. Instrument : Instrument 1                 Inj Volume : 30 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 20 µl
Acq. Method     : C:\CHEM32\1\METHODS\MARSAULT.M
Last changed    : 6/22/2011 10:51:16 AM by Nadon
Analysis Method : C:\CHEM32\1\METHODS\1MLSURGITH.M
Last changed    : 2/21/2012 8:02:57 AM by ELO
                  (modified after loading)
MARSAULT
```



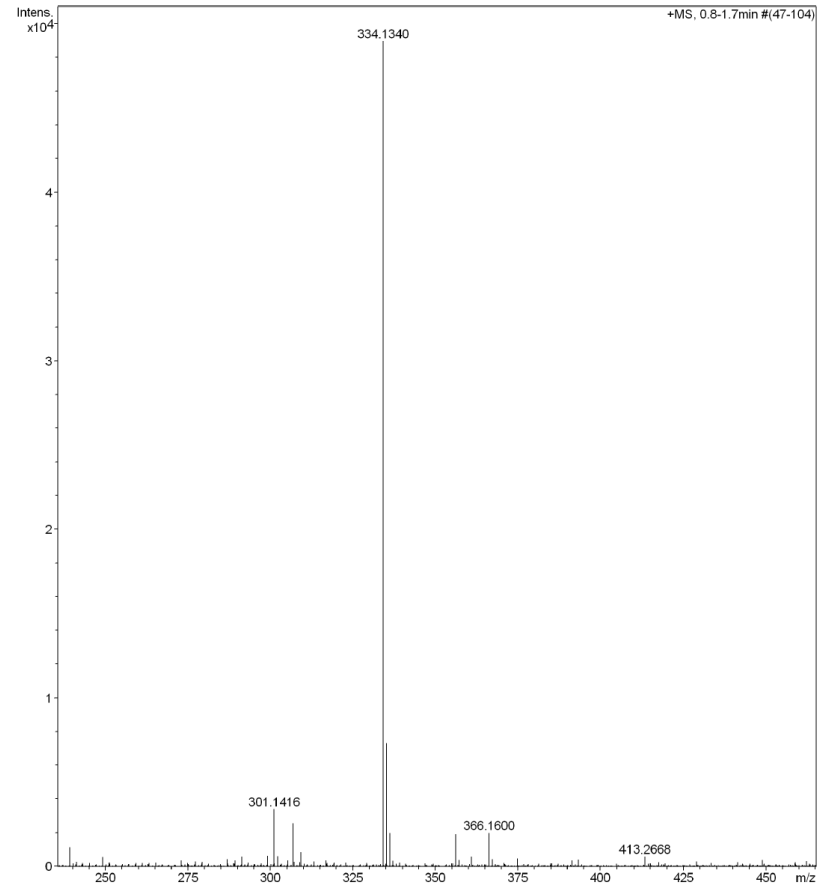

```
=====================================================================
                        Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: DAD1 A, Sig=223,16 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   8.220 MM    0.1759 1.52993e4  1449.96448 100.0000

Totals :                  1.52993e4  1449.96448
```

- MS

Generic Display Report

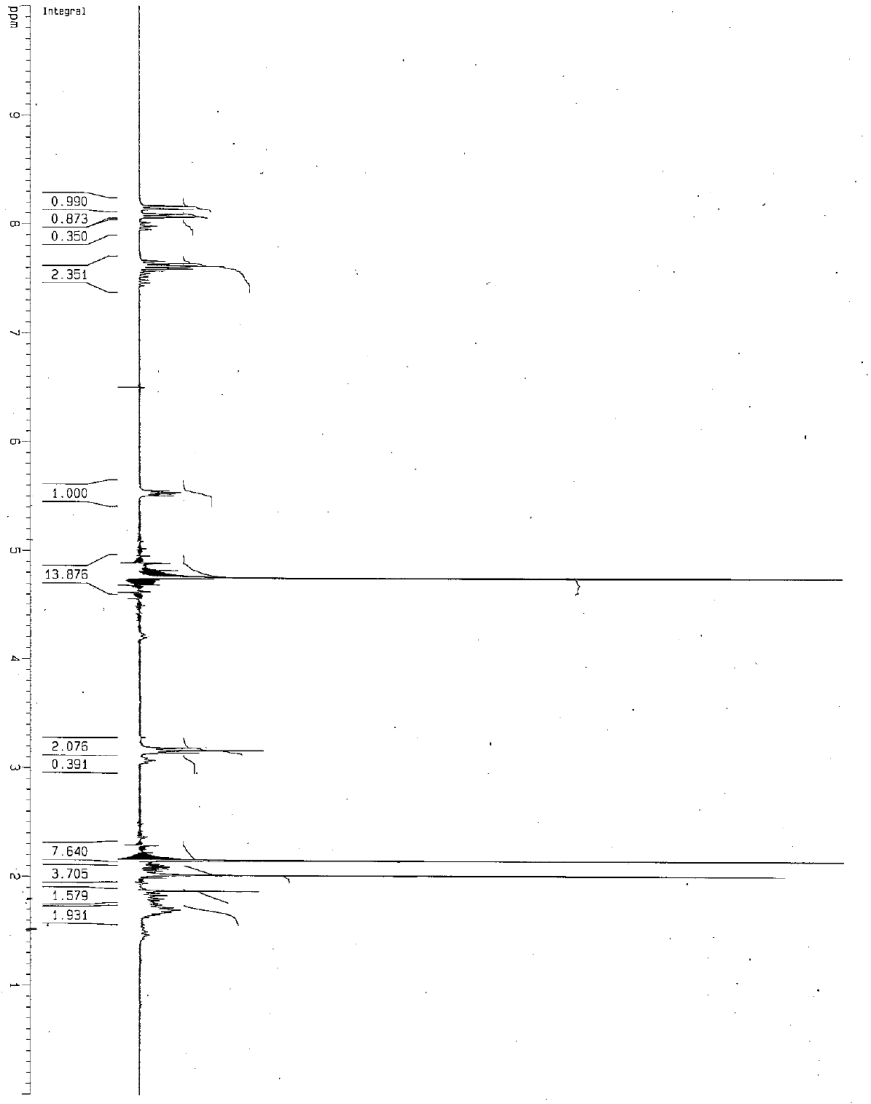
Analysis Info

Analysis Name	D:\Data\Echantillons\Acq000135.d	Acquisition Date	2/10/2012 3:55:11 PM
Method	Infusion_Echantillons.m	Operator	Rene
Sample Name	EC-0253-01	Instrument	maXis
Comment	INFUSION ESI		



Bruker Compass DataAnalysis 4.0 printed: 2/10/2012 3:59:55 PM Page 1 of 1

- 1H NMR


Integral
ppm
0.990
0.873
0.350
2.351
1.000
13.876
2.076
0.391
7.640
3.705
1.579
1.931

- ^{13}C NMR