Supplementary Material

A genetically encoded photocaged N^{ε}-methyl-L-lysine

Yane-Shih Wang, Bo Wu, Zhiyong Wang, Ying Huang, Wei Wan,[†] Pei-Jing Pai, William K. Russell, Yin N. Moe, David H. Russell, and Wenshe Liu*

Department of Chemistry and [†]Department of Biology, Texas A&M University, College Station, TX 77843
E-mail: wliu@chem.tamu.edu

Table of Contents

1. General Experimental

All reactions involving moisture sensitive reagents were conducted in oven-dried glassware under an argon atmosphere. Anhydrous solvents were obtained through standard laboratory protocols. Analytical thin-layer chromatography (TLC) was preformed on Whatman SiO_2 60 F-254 plates. Visualization was accomplished by UV irradiation at 254 nm or by staining with ninhydrin (0.3% w/v in glacial acetic acid/n-butyl alcohol 3:97). Flash column chromatography was performed with flash silica gel (particle size 32-63 μm) from Dynamic Adsorbents Inc (Atlanta, GA).

Specific rotations of chiral compounds were obtained at the designated concentration and temperature on a Rudolph Research Analytical Autopol II polarimeter using a 0.5 dm cell. Proton and carbon NMR spectra were obtained on Varian 300 and 500 MHz NMR spectrometers. Chemical shifts are reported as δ values in parts per million (ppm) as referenced to the residual solvents: chloroform (7.27 ppm for ^{1}H and 77.23 ppm for ^{13}C) or water (4.80 ppm for ^{1}H). A minimal amount of 1,4-dioxane was added as the reference standard (67.19 ppm for ^{13}C) for carbon NMR spectra in deuterium oxide, and a minimal amount of sodium hydroxide pellet was added to the NMR sample to aid in the solvation of amino acids which have low solubility in deuterium oxide under neutral or acidic conditions. ^{1}H NMR spectra are tabulated as follows: chemical shift, multiplicity (s = singlet, bs = broad singlet, d = doublet, t = triplet, q = quartet, m = multiplet), number of protons, and coupling constant(s). Mass spectra were obtained at the Laboratory for Biological Mass Spectrometry at the Department of Chemistry, Texas A&M University.

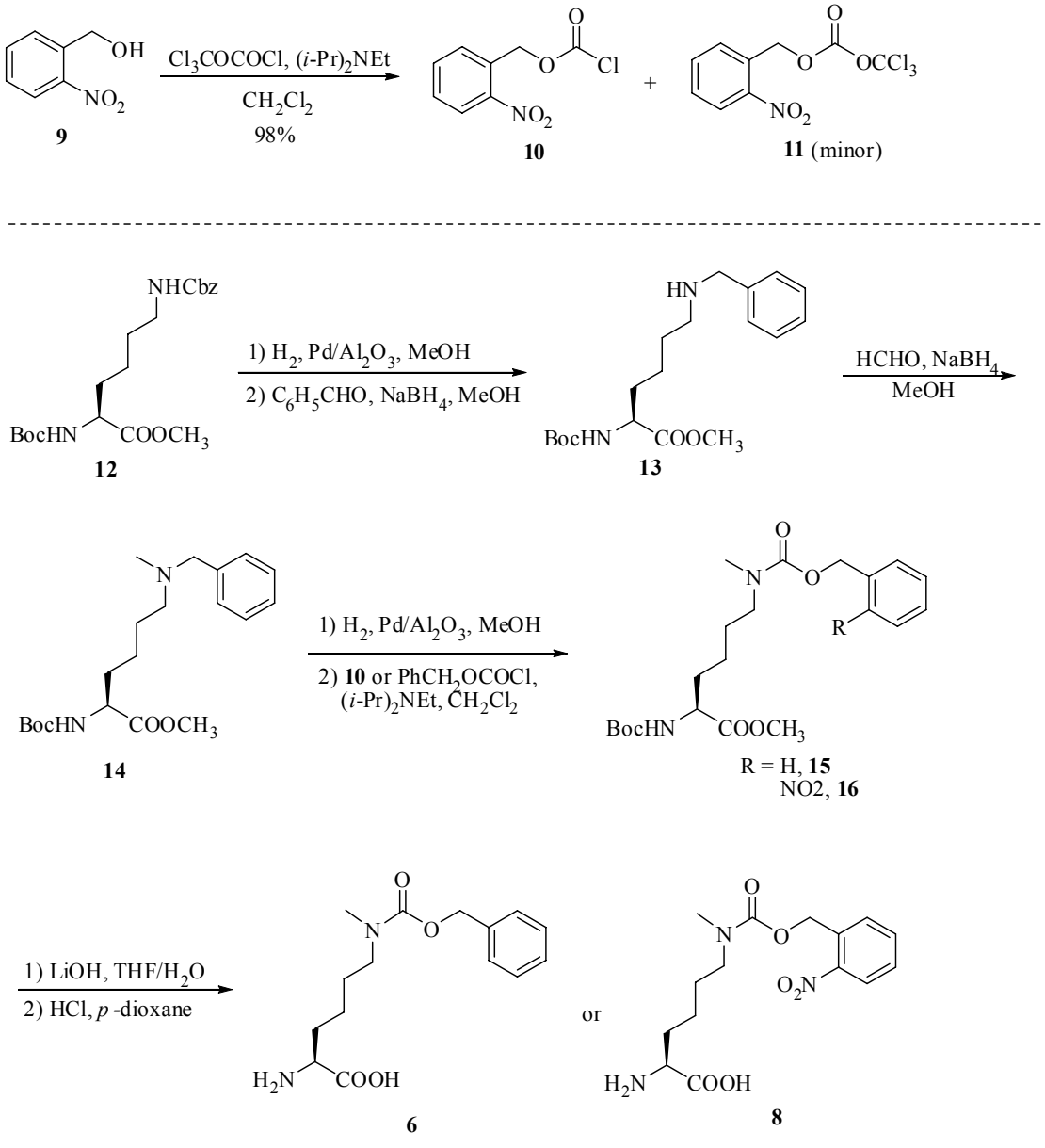
H-Lys(Z)-OH (**5**) and H-Lys(Me)-OH·HCl (**17**) were obtained from Chem-Impex International, Inc. (Wood Dale, IL). Compound **7** was synthesized as reported.[1] All other reagents were obtained from commercial suppliers and used as received.

2. Chemical Synthesis

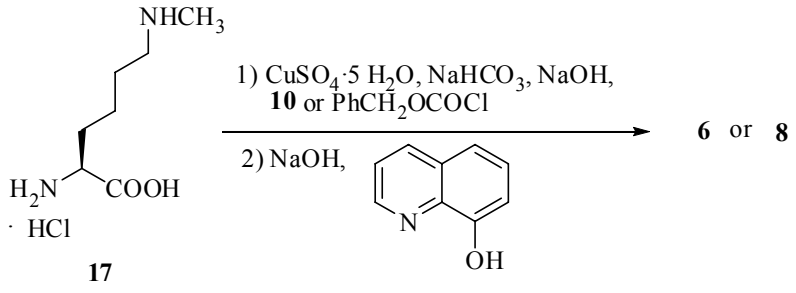
Compounds **6** and **8** were synthesized from **12**[2] in a scalable route (Supplementary Scheme 1). For comparison, a shorter synthesis from the relatively expensive N^{ϵ}-methyl-L-lysine hydrochloride (**17**) following the standard protocol of

copper complexation, reaction with appropriate chloroformate, and decomplexation with 8-hydroxyquinoline[3] was also developed (Supplementary Scheme 2).

Supplementary Scheme 1. Longer synthesis of **6** and **8**.



Supplementary Scheme 2. Shorter synthesis of **6** and **8**.



2-Nitrobenzyl chloroformate (10) and 2-nitrobenzyl trichloromethyl carbonate (11). To a solution of 2-nitrobenzyl alcohol (**9**, 1.97 g, 12.9 mmol) in anhydrous dichloromethane (55 mL) cooled in an ice bath was added diphosgene (1.71 mL, 14.2 mmol) in dichloromethane (10 mL) dropwise over 10 min followed by diisopropylethylamine (2.25 mL, 12.9 mmol) in dichloromethane (10 mL) dropwise over 10 min. The reaction mixture was then stirred at room temperature for 2 h, and sodium hydroxide (1 *N*, 20 mL) was added and stirred further at room temperature for 30 min. The mixture was washed with water (30 mL), saturated sodium bicarbonate (30 mL x 2) and brine (30 mL), dried (Na_2SO_4), and flash chromatographed (EtOAc/hexanes, 1:20) to give a mixture of **10** and **11** (2.71 g, 98%) as a yellow oil. A minor fraction of impurity, presumably **11**, was evident from NMR analysis but did not interfere with the next step reaction. No further purification was performed. For **10**[4]: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.19 (d, 1 H, J = 8.0 Hz), 7.74 (t, 1 H, J = 7.8 Hz), 7.65 (d, 1 H, J = 8.0 Hz), 7.58 (t, 1 H, J = 7.8 Hz), 5.75 (s, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 150.6, 147.7, 134.4, 129.9, 129.8, 129.2, 125.6, 69.7. For **11**: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.20 (d, 1 H, J = 6.5 Hz), 7.74 (t, 1 H, J = 7.7 Hz), 7.68 (d, 1 H, J = 8.0 Hz), 7.58 (t, 1 H, J = 7.7 Hz), 5.75 (s, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 147.3, 147.2, 134.5, 130.2, 129.8, 129.0, 108.0, 68.0.

***N*$^{\alpha}$-Boc-*N*$^{\varepsilon}$-benzyl-L-lysine methyl ester (13).** A solution of **12** (4.20 g, 10.6 mmol) in methanol (100 mL) was hydrogenated under a H_2 balloon in the presence of palladium on alumina (10 wt.% Pd, 0.71 g, 0.67 mmol) at room temperature for 3 h, and TLC analysis showed a complete conversion. The mixture was then filtered over a pad of Celite and the solution was directly used for the next step reaction. The material should be immediately used without purification since prolonged storage at room temperature or flash chromatography would contribute to lactam formation.

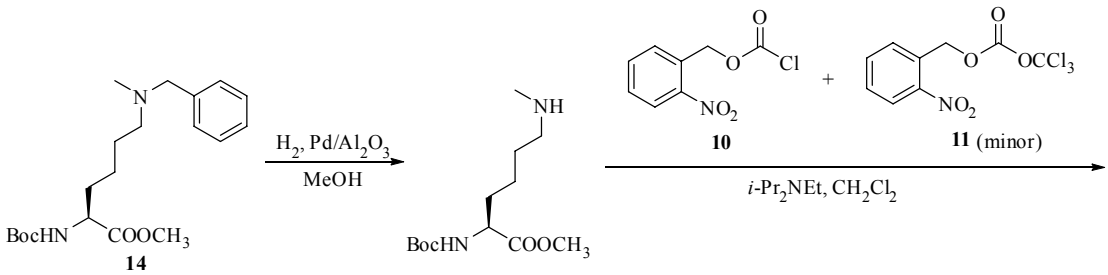
To a solution of the above amine (~10.6 mmol) in methanol was added benzaldehyde (4.00 mL, 39.4 mmol), and the reaction mixture was stirred at room temperature for 30 min. The mixture was then cooled in an ice bath, and sodium borohydride (0.75 g, 19.8 mmol) was added portionwise. The mixture was then stirred at room temperature overnight, and water (10 mL) was added dropwise to quench the reaction. Most of the methanol was evaporated under a reduced pressure, and the residue was dissolved in ethyl acetate (100 mL), washed with water (30 mL), saturated sodium bicarbonate (30 mL) and brine (30 mL), dried (Na_2SO_4), and evaporated to give the crude **13** as a yellow oil, which was used in the next step reaction without further purification. A small fraction of pure **13** was obtained by flash chromatography (10% methanol with 5% triethylamine in dichloromethane) for characterization. $[\alpha]_D^{19}$ +7.4 (*c* 4.85, CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.35-7.30 (m, 4 H), 7.27-7.24 (m, 1 H), 5.05 (d, 1 H, J = 7.0 Hz), 4.32-4.28 (m, 1 H), 3.78 (s, 2 H), 3.74 (s, 3 H), 2.63 (t, 2 H, J =7.0 Hz), 1.83-1.78 (m, 1 H), 1.67-1.60 (m, 1 H), 1.58-1.48 (m, 3 H), 1.44 (s, 9 H), 1.41-1.36 (m, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.5, 155.5, 140.5, 128.5, 128.2, 127.0, 79.9, 54.1, 53.5, 52.3, 49.1, 32.7, 29.7, 28.4, 23.2; HRMS (ESI) calcd for $C_{19}H_{31}N_2O_4$ ($[M + H]^+$) 351.2284, found 351.2282.

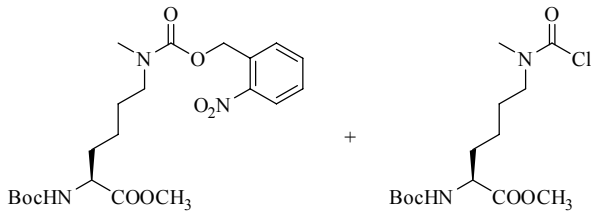
N^{α}-Boc-N^{ε}-benzyl-N^{ε}-methyl-L-lysine methyl ester (14). To a solution of crude **13** (~10.6 mmol) in methanol (100 mL) in methanol was added formaldehyde (37% aqueous solution, 3.00 mL, 40.3 mmol), and the reaction mixture was stirred at room temperature for 30 min. The mixture was then cooled in an ice bath, and sodium borohydride (0.77 g, 20.4 mmol) was added portionwise. The mixture was then stirred further at room temperature for 4 h, and water (30 mL) was added dropwise to quench the reaction. Most of the methanol was evaporated under a reduced pressure, and the residue was dissolved in ethyl acetate (100 mL), washed with water (30 mL), hydrochloric acid (1 *N*, 30 mL), sodium hydroxide (1 *N*, 30 mL) and brine (30 mL), dried (Na_2SO_4), evaporated, and flash chromatographed (EtOAc/hexanes, 1:1 then 5% to 10% methanol in dichloromethane) to give **14** (3.32 g, 86% yield for three steps) as a yellow oil. $[\alpha]_D^{20}$ +7.7 (*c* 2.37, CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.33-7.29 (m, 4 H), 7.26-7.23 (m, 1 H), 5.08 (d, 1 H, *J* = 8.5 Hz), 4.31-4.27 (m, 1 H), 3.73 (s, 3 H), 3.46 (s, 2 H), 2.35 (t, 2 H, *J* =7.2 Hz), 2.17 (s, 3 H), 1.82-1.76 (m, 1 H), 1.66-1.47 (m, 3 H), 1.44 (s, 9 H), 1.40-1.32 (m, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.6, 155.6, 139.3, 129.2, 128.4, 127.1, 80.0, 62.5, 57.1, 53.6, 52.4, 42.4, 32.7, 28.5, 27.0, 23.3; HRMS (ESI) calcd for $C_{20}H_{33}N_2O_4$ ($[M + H]^+$) 365.2440, found 365.2437.

N^{α}-Boc-N^{ε}-Cbz-N^{ε}-methyl-L-lysine methyl ester (15). A solution of **14** (2.58 g, 7.07 mmol) in methanol (50 mL) was hydrogenated under a H_2 balloon in the presence of palladium on alumina (10 wt.% Pd, 0.50 g, 0.47 mmol) at room temperature for 5 h. The mixture was then filtered over a pad of Celite and evaporated to give the crude amine (Boc-Lys(Me)-OMe) as a grey oil. A small fraction of pure amine was obtained by flash chromatography (10% methanol with 5% triethylamine in dichloromethane) for characterization. $[\alpha]_D^{19}$ +6.9 (*c* 1.90, CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 500 MHz) δ 5.38 (d, 1 H, *J* = 7.0 Hz), 4.14-4.12 (m, 1 H), 3.59 (s, 3 H), 2.95 (s, 1 H), 2.46 (t, 2 H, *J* =7.2 Hz), 2.29 (s, 3 H), 1.69-1.64 (m, 1 H), 1.55-1.48 (m, 1 H), 1.43-1.33 (m, 2 H), 1.30 (s, 9 H), 1.27-1.22 (m, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.4, 155.5, 79.6, 53.3, 52.1, 51.2, 35.8, 32.2, 28.8, 28.2, 23.0; HRMS (ESI) calcd for $C_{13}H_{27}N_2O_4$ ($[M + H]^+$) 275.1971, found 275.1968.

To a solution of the above amine (~7.07 mmol) and diisopropylethylamine (2.00 mL, 11.5 mmol) in anhydrous dichloromethane (40 mL) cooled in an ice bath was added benzyl chloroformate (95%, 1.50 mL, 10.5 mmol) dropwise over 10 min, and the mixture was stirred at room temperature for 12 h. The mixture was then diluted in ethyl acetate (100 mL), washed with sodium hydroxide (0.5 *N*, 40 mL) and brine (40 mL), dried (Na_2SO_4), evaporated, and flash chromatographed (EtOAc/hexanes, 1:3) to give **15** (2.59 g, 90% for two steps) as a colorless oil. $[\alpha]_D^{19}$ +1.2 (*c* 1.51, CH_2Cl_2); ^{1}H NMR analysis showed a 1.5:1 mixture of rotamers at room temperature. Major rotamer: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.33-7.32 (m, 4 H), 7.29-7.26 (m, 1 H), 5.26 (d, 1 H, *J* = 6.0 Hz), 5.10 (s, 2 H), 4.24 (m, 1 H), 3.69 (s, 3 H), 3.32-3.19 (m, 2 H), 2.87 (s, 3 H), 1.78-1.66 (m, 2 H), 1.57-1.51 (m, 2 H), 1.41 (s, 9 H), 1.36-1.28 (m, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.4, 156.5, 155.6, 137.0, 128.5, 128.0, 127.9, 79.7, 67.0, 53.4, 52.2, 48.4, 33.9, 32.0, 28.4, 26.9, 22.2; Characteristic peaks of the minor rotamer: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 5.05 (d, 1 H, *J* = 6.0 Hz);

^{13}C NMR ($CDCl_3$, 125 MHz) δ 156.3, 155.4, 79.9, 53.3, 34.6, 32.4, 27.5, 22.5; HRMS (ESI) calcd for $C_{21}H_{33}N_2O_6$ ($[M + H]^+$) 409.2339, found 409.2332.





***N*$^{\alpha}$-Boc-*N*$^{\varepsilon}$-(2-nitrobenzyl)oxycarbonyl-*N*$^{\varepsilon}$-methyl-L-lysine methyl ester (16) and *N*$^{\alpha}$-Boc-*N*$^{\varepsilon}$-chlorocarbonyl-*N*$^{\varepsilon}$-methyl-L-lysine methyl ester (18).** Compound **14** (1.06 g, 2.91 mmol) was converted into the corresponding amine by hydrogenolysis, which was then treated with crude **10** (0.94 g, 4.38 mmol) according to the procedure for **15** to give **16** (0.91g, 69% for two steps) as a yellow oil. A small amount of **18** (yield not determined), the structure of which was assigned based on NMR and MS analysis data, was obtained as a colorless oil. Presumably Boc-Lys(Me)-OMe reacts with **11** to give **16** and generates one molecule of phosgene at the same time, which then acylates the residual Boc-Lys(Me)-OMe to afford **18**.

For **16**: $[\alpha]_D^{20}$ +8.9 (*c* 1.70, CH_2Cl_2); R_f = 0.46 (EtOAc/hexanes, 1:1); ^{1}H NMR analysis showed a 1.1:1 mixture of rotamers at room temperature. Major rotamer: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.96 (d, 1 H, *J* = 8.0 Hz), 7.58 (t, 1 H, *J* = 7.5 Hz), 7.50-7.47 (m, 1 H), 7.39 (t, 1 H, *J* = 8.2 Hz), 5.42 (s, 2 H), 5.21 (d, 1 H, *J* = 7.5 Hz), 4.20-4.16 (m, 1 H), 3.63 (s, 3 H), 3.21 (appar. nonet, 2 H, *J* = 7.1 Hz), 2.86 (s, 3 H), 1.74-1.44 (m, 4 H), 1.34 (s, 9 H), 1.30-1.24 (m, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.2, 155.6, 155.4, 147.5, 133.6, 133.1, 128.7, 128.5, 124.8, 79.6, 63.7, 53.2, 48.5, 33.8, 31.9, 28.2, 26.7, 22.2; Characteristic peaks of the minor rotamer: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 5.13 (d, 1 H, *J* = 7.5 Hz), 2.83 (s, 3 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.1, 155.5, 155.4, 147.5, 128.9, 128.5, 79.7, 52.1, 48.4, 34.6, 32.3, 27.4, 22.4; HRMS (ESI) calcd for $C_{21}H_{31}N_3O_8Li$ ($[M + Li]^+$) 460.2271, found 460.2272.

For **18**: $[\alpha]_D^{20}$ +6.6 (*c* 2.20, CH_2Cl_2); R_f = 0.61 (EtOAc/hexanes, 1:1); ^{1}H NMR analysis showed a 1.2:1 mixture of rotamers at room temperature. Major rotamer: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 5.06 (d, 1 H, *J* = 8.0 Hz), 4.30 (m, 1 H), 4.20-3.73 (s, 3 H), 3.44 (t, 1 H, *J* = 7.5 Hz), 3.40-3.36 (m, 1 H), 3.10 (s, 3 H), 1.85-1.56 (m, 4 H), 1.43 (s, 9 H), 1.40-1.31 (m, 2 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 173.3, 155.5, 149.8, 80.1, 53.2, 52.8, 51.2, 38.6, 32.5, 28.5, 26.7, 22.4; Characteristic peaks of the

minor rotamer: ^{1}H NMR ($CDCl_3$, 500 MHz) δ 3.74 (s, 3 H), 3.02 (s, 3 H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 149.3, 80.2, 52.5, 36.8, 32.7, 27.4, 22.4; HRMS (ESI) calcd for $C_{14}H_{25}ClN_2O_5Na$ ([M + Na]$^+$) 361.1320 (^{37}Cl)/359.1350 (^{35}Cl), found 361.1348/ 359.1359; calcd for $C_{14}H_{26}ClN_2O_5$ ([M + H]$^+$) 339.1505/337.1530, found 339.1585/ 337.1557.

***N*$^{\varepsilon}$-Benzyloxycarbonyl-*N*$^{\varepsilon}$-methyl-L-lysine (6).** To a solution of **15** (2.59 g, 6.34 mmol) in THF (20 mL) was added lithium hydroxide solution (0.5 M, 25.0 mL, 12.5 mmol), and the mixture was stirred at room temperature for 3 h. The mixture was diluted in water (20 mL) and extracted with ether (30 mL x 2). The ether extracts were discarded, and the remaining aqueous solution was adjusted to pH 3 with hydrochloric acid (3 *N*), with the concomitant formation of white precipitate. The suspension was extracted with ethyl acetate (50 mL x 2), and the combined organic phases were washed once with brine (30 mL), dried (Na_2SO_4), and evaporated to give the crude carboxylic acid as a colorless oil, which was directly used without further purification.

The above crude acid (~6.34 mmol) was dissolved in 1,4-dioxane (15 mL), and hydrogen chloride in 1,4-dioxane (4.0 M, 5.0 mL, 20.0 mmol) was added. The resulting white suspension was stirred at room temperature for 12 h, evaporated, redissolved in a minimal amount of water, and loaded onto an ion-exchange column made from Dowex 50WX4-400 cation-exchange resin (~14 mL bed volume). The column was washed with excessive water (300 mL) and then eluted with pyridine (1 M, 450 mL) to give **6** (1.51 g, 81% for two steps) as a white powder. $[\alpha]_D^{20}$ +14.1 (*c* 1.07, 3 *N* HCl) (lit.[5] $[\alpha]_D^{25}$ +14.0 (*c* 0.5, acetic acid)); ^{1}H NMR analysis showed a 1:1 mixture of rotamers at room temperature. Major rotamer: ^{1}H NMR (D_2O, 500 MHz) δ 7.46-7.42 (m, 5 H), 5.16 (s, 2 H), 3.68 (m, 1 H), 3.34 (m, 2 H), 2.90 (s, 3 H), 1.83 (m, 2 H), 1.60 (quintet, 2 H, *J* = 7.3 Hz), 1.34 (m, 2 H); ^{13}C NMR (D_2O, 75 MHz) δ 184.1, 158.4, 137.1, 129.4, 129.0, 128.5, 68.0, 56.5, 49.0, 35.1, 34.8, 27.6, 22.8; Characteristic peaks of the minor rotamer: ^{1}H NMR (D_2O, 500 MHz, pH = 14) δ 2.95 (s, 3 H); ^{13}C NMR (D_2O, 75 MHz) δ 128.2, 67.2, 49.2, 34.2, 27.3. HRMS (ESI) calcd for $C_{15}H_{23}N_2O_4$ ([M + H]$^+$) 295.1658, found 295.1656.

***N*$^{\varepsilon}$-(2-Nitrobenzyl)oxycarbonyl-*N*$^{\varepsilon}$-methyl-L-lysine (8).** According to the same procedure for **6**, **16** (0.914 g, 2.02 mmol) afforded **8** (0.514 g, 75% for two steps) as a pale yellow solid. $[\alpha]_D^{20}$ +14.2 (*c* 1.16, 3 *N* HCl); ^{1}H NMR analysis showed a 1.1:1 mixture of rotamers at room temperature. Major rotamer: ^{1}H NMR (D_2O, 500 MHz, pH = 14) δ 8.00 (d, 1 H, *J* = 8.0 Hz), 7.64 (t, 1 H, *J* = 7.7 Hz), 7.50 (d, 1 H, *J* = 8.5 Hz), 7.47 (t, 1 H, *J* = 8.2 Hz), 5.32 (s, 2 H), 3.95 (m, 1 H), 3.20 (m, 2 H), 2.77 (s, 3 H), 1.84 (m, 2 H), 1.49 (m, 2 H), 1.31-1.27 (m, 2 H); ^{13}C NMR (D_2O, 75 MHz) δ 184.2, 158.0, 147.6, 135.1, 132.7, 129.8, 129.4, 125.7, 65.0, 56.5, 49.2, 35.1, 34.9, 27.6, 22.8; Characteristic peaks of the minor rotamer: ^{1}H NMR (D_2O, 500 MHz, pH = 14) δ 2.82 (s, 3 H); ^{13}C NMR (D_2O, 75 MHz) δ 133.0, 129.7, 64.9, 49.3, 34.2, 27.2; HRMS (ESI) calcd for $C_{15}H_{22}N_3O_6$ ([M + H]$^+$) 340.1509, found 340.1513.

Synthesis of 6 from the shorter pathway. To a solution of **17** (1.50 g, 7.63 mmol) in

water (30 mL) was added cupric sulfate pentahydrate (1.00 g, 4.00 mmol), followed by sodium bicarbonate (1.42 g, 16.90 mmol) in small portions to prevent excessive bubble formation. Benzyl chloroformate (2.53 g, 13.36 mmol) in dioxane (5 mL) was then added dropwise in 5 min, followed by sodium hydroxide (0.49 g, 12.25 mmol) in one portion. The reaction mixture was stirred at room temperature for 16 h, filtered, washed with water (100 mL), ethanol (50 mL) and diethyl ether (50 mL), and dried in the open air for 1 h to give the crude copper complex (2.40 g, 97%) as a blue solid.

All the above copper complex (2.40 g, ~7.40 mmol) was suspended in sodium hydroxide solution (0.2 *N*, 100 mL, 20 mmol), and a solution of 8-hydroxylquinoline (1.40 g, 9.64 mmol) in 1,4-dioxane (10 mL). The resulting green suspension was stirred at room temperature overnight and filtered. The filtrate was adjusted to pH 3 with hydrochloric acid (3 *N*) and extracted with ethyl acetate (40 mL x 2). The organic extracts were discarded, and the aqueous phase was concentrated to about 20 mL and loaded onto an ion-exchange column made from Dowex 50WX4-400 cation-exchange resin (~14 mL bed volume). The column was washed with excessive water (300 mL) and then eluted with pyridine (1 M, 450 mL) to give a yellow solid upon evaporation, which was suspended in ethanol, filtered, washed ethyl acetate dried to give **6** (1.46 g, 65% for two steps) as a white solid. $[\alpha]_D^{22}$ +15.3 (*c* 1.02, 3 *N* HCl). All other characterization data were identical to that of **6** from the longer route.

Synthesis of 8 from the shorter pathway. According to the same procedure for **6**, **17** (0.50 g, 2.54 mmol) afforded **8** (0.39 g, 45% yield for two steps). The compound was identical to **8** from the longer route in all aspects.

3. DNA and Protein Sequences

3.1 DNA Sequences

Z Domain:

atgactagtgtagacaactagatcaacaaagaacaacaaaacgccttctatgagatcttacatttacctaacctgaatgagga
gcagcgtgatgccttcatccaaagtttaaaagatgacccaagccaaagcgctaacctttagcagaagctaaaaagctaaat
gatgctcaggcgcctaagggatctgagctccatcaccatcaccatcactaa

GFP_{UV}:

atgagtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaatttttctgt
cagtggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactacctgttccat
ggccaacacttgtcactactttctcttatggtgttcaatgcttttcccgttatccggatcacatgaaacggcatgactttttcaaga
gtgccatgcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggaactacaagacgcgtgctgaagtca
agtttgaaggtgataccttgttaatcgtatcgagttaaaaggtattgattttaaagaagatggaaacattctcggacacaaact
cgaatacaactataactcacacaatgtatacatcacggcagacaaacaaaagaatggaatcaaagctaacttcaaaattcgc
cacaacattgaagatggatccgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttacca
gacaaccattacctgtcgacatagtctgccctttcgaaagatcccaacgaaaagcgtgaccacatggtccttcttgagtttgta
actgctgctgggattacacatggcatggatgaactctacaaagagctccatcaccatcaccatcactaa

pylT:

ggaaacctgatcatgtagatcgaatggactctaaatccgttcagccgggttagattcccggggtttccgcca

Methanosarcina mazei PylRS:

atggataaaaaaccactaaacactctgatatctgcaaccgggctctggatgtccaggaccggaacaattcataaaataaaac
accacgaagtctctcgaagcaaaatctatattgaaatggcatgcggagaccaccttgttgtaaacaactccaggagcagca
ggactgcaagagcgctcaggcaccacaaatacaggaagacctgcaaacgctgcagggtttcggatgaggatctcaataa
gttcctcacaaaggcaaacgaagaccagacaagcgtaaaagtcaaggtcgtttctgcccctaccagaacgaaaaaggca
atgccaaaatccgttgcgagagccccgaaacctcttgagaatacagaagcggcacaggctcaaccttctggatctaaatttt
cacctgcgataccggtttccacccaagagtcagtttctgtcccggcatctgtttcaacatcaatatcaagcatttctacaggag
caactgcatccgcactggtaaaagggaatacgaaccccattacatccatgtctgcccctgttcaggcaagtgcccccgcac
ttacgaagagccagactgacaggcttgaagtcctgttaaacccaaaagatgagatttccctgaattccggcaagcctttcag
ggagcttgagtccgaattgctctctcgcagaaaaaaagacctgcagcagatctacgcggaagaaagggagaattatctgg
ggaaactcgagcgtgaaattaccaggttctttgtggacaggggttttctggaaataaaatccccgatcctgatccctcttgagt
atatcgaaaggatgggcattgataatgataccgaactttcaaaacagatcttcagggttgacaagaacttctgcctgagaccc
atgcttgctccaaacctttacaactacctgcgcaagcttgacagggccctgcctgatccaataaaaatttttgaaataggccca
tgctacagaaaagagtccgacggcaaagaacacctcgaagagtttaccatgctgaacttctgccagatgggatcgggatg
cacacgggaaaatcttgaaagcataattacggacttcctgaaccacctgggaattgatttcaagatcgtaggcgattcctgca
tggtctatggggataccttgatgtaatgcacggagacctggaactttcctctgcagtagtcggacccataccgcttgaccg
ggaatggggtattgataaaccctggataggggcaggtttcgggctcgaacgccttctaaaggttaaacacgactttaaaaat
atcaagagagctgcaaggtccgagtcttactataacgggatttctaccaacctgtaa

3.2 Proteins Sequences

Z Domain:

MTSVDNXINKEQQNAFYEILHLPNLNEEQRDAFIQSLKDDPSQSANLLAEAKK
LNDAQAPKGSELHHHHHH

X represents a noncanonical amino acid.

GFP_{UV}:
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKL
PVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGN
YKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQK
NGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTXSALSKDP
NEKRDHMVLLEFVTAAGITHGMDELYKELHHHHHH

X represents a noncanonical amino acid.

Methanosarcina mazei PylRS:

MDKKPLNTLISATGLWMSRTGTIHKIKHHEVSRSKIYIEMACGDHLVVNNSRS
SRTARALRHHKYRKTCKRCRVSDEDLNKFLTKANEDQTSVKVKVVSAPTRT
KKAMPKSVARAPKPLENTEAAQAQPSGSKFSPAIPVSTQESVSVPASVSTSISSI
STGATASALVKGNTNPITSMSAPVQASAPALTKSQTDRLEVLLNPKDEISLNS
GKPFRELESELLSRRKKDLQQIYAEERENYLGKLEREITRFFVDRGFLEIKSPIL
IPLEYIERMGIDNDTELSKQIFRVDKNFCLRPMLAPNLYNYLRKLDRALPDPIK
IFEIGPCYRKESDGKEHLEEFTMLNFCQMGSGCTRENLESIITDFLNHLGIDFKI
VGDSCMVYGDTLDVMHGDLELSSAVVGPIPLDREWGIDKPWIGAGFGLERL
LKVKHDFKNIKRAARSESYYNGISTNL

4. Construction of plasmids

4.1 Constructions of pY+ and pY-

The plasmid pY+ was derived from the pRep plasmid by replacing the suppressor tRNA in pRep by pylT.[6] The gene of pylT flanked by the lpp promoter at the 5' end and the *rrnC* terminator at the 3' end was amplified from pBK-AcKRS-pylT.[7,8] The plasmid pY- was derived from the pNeg plasmid by replacing the suppressor tRNA with pylT.[6] Similarly, the gene of pylT flanked by the lpp promoter at the 5' end and the *rrnC* terminator at the 3' end was amplified from pBK-AcKRS-pylT.[7,8] pY+ has a tetracycline selection marker and a chloramphenicol acetyltransferase gene with an amber mutation at D112. pY- has an ampicillin selection marker and a barnase gene with two amber mutations at Q2 and D44. The barnase gene is under control of a pBad promoter.

4.2 Construction of pET-pylT-GFP

Plasmid pET-pylT-GFP was derived from the plasmid pAcKRS-pylT-GFP1Amber in which GFP_{UV} has an amber mutation at Q204.[7,8] The restriction enzyme *BglII* was used to cut off the ACKRS gene. The digested pAcKRS-pylT-GFP1Amber plasmid was ligated to form pET-pylT-GFP.

4.3 Construction of pET-pylT-Z

The pET-pylT-Z plasmid was derived from pET-pylT-GFP. This gene was amplified from the pLeiZ plasmid.[9] Two restriction sites, *NdeI* at the 5' end and *SacI* at the 3' end, were introduced in the PCR product which was subsequently digested and used to replace GFP_{UV} in pET-pylT-GFP.

5. Construction of the pRS1 Library

The plasmid pBK-mmPylRS that encodes wild-type *Methanosarcina mazei* PylRS was derived from a pBK plasmid containing *p*-iodophenylalanyl-tRNA synthetase.[10] The pylRS gene is under the control of *E. coli glnS* promoter and terminator. It was amplified from genomic DNA of *Methanosarcina mazei* strain DSM 3647 (ATCC) by flanking primers, pBK-mmPylRS-NdeI-F and pBK-mmPylRS-PstI~NsiI-R. To construct the pRS1 library, NNK (N=A or C or G or T, K=G or T) mutations were introduced at six sites by overlap extension PCR.[11] The following pairs of primers were used to generate a PylRS gene library with randomization at six sites: (1) pBK-mmPylRS-NdeI-F (5'-gaatcccatatggataaaaaaccactaaacactctg-3') and mmPylRS-Mutlib01-R (5'-ggccctgtcaagcttgcgmnngtagttmnnmnngtttggagcaagca tggg-3'); (2) mmPylRS-Mutlib02-F (5'-cgcaagcttgacagggccctgcctgatcc-3') and mmPylRS-Mutlib03-R (5'-gcatcccgatcccatctgmnngaamnncagcatggtaaactcttc-3'); (3) mmPylRS-Mutlib04-F (5'-cagatgggatcgggatgcacacg-3') and mmPylRS-Mutlib05-R (5'-ccgaaacctgcccctatmnngggtttatcaatacccca-3'); (4) mmPylRS-Mutlib06-F (5'-ataggggcaggtttcgggctcgaacgcc-3') and pBK-mmPylRS-PstI~NsiI-R (5'-gtttgaaaatgcatttacaggttggtagaaatccc-3'). The gene library was digested with the restriction enzymes *NdeI* and *NsiI*, gel-purified, and ligated back into the pBK vector

digested by *NdeI* and *PstI* to afford the pRS1 plasmid library. 1 μg of the ligation products were then electroporated into *E. coli* Top10 cells. Electroporated cells were recovered in SOC medium for 60 min at 37 °C, transferred into a 2 L 2YT medium with kanamycin (25 μg/mL) and were incubated at 37 °C to an OD_{600} of 1.0. To calculate the library size, 1 μL of the recovered SOC culture was subjected to serial dilutions in 2YT, plated on LB agar plates with kanamycin (25 μg/mL), and grown overnight in a 37 °C incubator. Based on the colony numbers on these plates, the pRS1 library contains approximately 1.01×10^9 independent transformants. Sequencing pylRS variants in 20 clones did not reveal any significant bias at the randomization sites.

6. Selection Procedure for Evolving Pyrrolysyl-tRNA Synthetase

The selections followed the scheme shown in **Supplementary Scheme 3**. For the positive selection, the pRS1 library was used to transform *E. coli* TOP10 competent cells harboring pY+ to yield a cell library greater than 1×10^9 cfu, ensuring complete coverage of the pRS1 library. Cells were plated on minimal agar plates containing 12 μg/mL tetracycline (Tet), 25 μg/mL kanamycin (Kan), 68 μg/mL chloramphenicol (Cm) and 1 mM **5**. After incubation at 37 °C for 72 h, colonies on the plates were collected. Total plasmids were isolated and separated by 1% agarose gel electrophoresis. pRS1 plasmids were extracted using a gel-extraction kit (QIAGEN). The extracted pRS1 plasmids from the positive selection were used to transform *E. coli* TOP10 cells harboring pY− for the negative selection. After electroporation, the cells were allowed to recover for 1 h at 37 °C in SOC media before being plated on LB agar plates containing 50 μg/mL Kan, 200 μg/mL ampicillin (Amp), and 0.2% arabinose. The plates were incubated for 16 h at 37 °C. Survived cells were then pooled and the pRS1 plasmids were extracted. The selection power to exclude the mutants that also took endogenous amino acids was tested on LB agar plates containing 50 μg/mL Kan, 200 μg/mL Amp, 0.2% arabinose, and 1mM **5**. The plates containing 1 mM **5** showed fewer and fewer colony numbers as the alternative selections were repeated. Five alternative selections (three positive + two negative) finally yielded many colonies. After the final third positive selection 22 single colonies were selected and the plasmids were isolated for sequencing, and another 96 single colonies were chosen to test their ability to grow on plates with 102 μg/mL Cm, 25 μg/mL Kan, 12 μg/mL Tet, and 1 mM of **5**, **6**, **7** or **8**. An LB agar plate without any NAA was used as the control. Images of colonies growing on different plates are shown in **Supplementary Fig. 7**. Sequences of PylRS variants that charge pylT with different NAAs are presented in **Supplementary Table 1**.

7. Protein Expression and Purification

To express GFP_{UV} incorporated with a NAA, we cotransformed *E. Coli* BL21(DE3) cells with pBK-mKRS1 and pET-pylT-GFP. Cells were recovered in 1 mL of LB medium for 1 h at 37 ºC before being plated on LB agar plate containing

Kan (25 μg/mL) and Amp (100 μg/mL). A single colony was then selected and grown overnight in a 10 mL culture, which was then used to inoculate 100 mL of M9 minimal media supplemented with 1% glycerol, 300 μM leucine, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.2% NaCl, 25 μg/mL Kan, and 100 μg/mL Amp. Cells were grown at 37 ºC in an incubator (300 r.p.m.) and protein expression was induced when the OD_{600} reached 0.7 by adding both IPTG and **5** to a final concentration of 1 mM. After 6 h induction, cells were harvested, resuspended in a lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0), and sonicated. The cell lysate was clarified by centrifugation (60 min, 11,000 r.p.m., 4 ºC). The supernatant was injected onto a Ni^{2+}-NTA column (30 mL) equipped on a GE Healthcare ÄKTApurifier FPLC system and washed with 45 mL of lysis buffer and 45 mL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 40 mM imidazole, pH 8.0). Proteins were then eluted out by running an imidazole gradient from 40 mM to 250 mM in lysis buffer. Pure fractions were collected and concentrated. The buffer was later changed to 1 mM ammonium bicarbonate using an Amicon Ultra-15 Centrifugal Filter Device (Millopore, 10,000 MWCO cut). The purified proteins were analyzed by 15% SDS-PAGE. GFP_{UV} proteins incorporated with other NAAs were expressed and purified similarly except that the supplemented NAA was changed. For all NAAs, 1 mM final concentration was used.

Z domain proteins incorporated with different NAAs were expressed and purified in the same way as for GFP_{UV} proteins except that pET-pylT-Z was used to cotransform *E. coli* BL21(DE3) together with pBK-mKRS1.

8. Photolysis to Form the Monomethylated Protein

Z-7, **Z-8**, and **GFP-8** (1 mg/mL) in 1 mM ammonium bicarbonate buffer were treated with 365 nm UV light from a hand-held UV lamp for one hour.

9. Protein LC-ESI-MS Analysis

An Agilent (Santa Clara, CA) 1200 capillary HPLC system was interfaced to an API QSTAR Pulsar Hybrid QTOF mass spectrometer (Applied Biosystems/MDS Sciex, Framingham, MA) equipped with an electrospray ionization (ESI) source. Liquid chromatography (LC) separation was achieved using a Phenomenex Jupiter C4 microbore column (150 × 0.50 mm, 300 Å) (Torrance, CA) at a flow rate of 10 μL per min. The proteins were eluted using a gradient of (A) 0.1% formic acid versus (B) 0.1% formic acid in acetonitrile. The gradient timetable was as follows: 2% B for 5 min, 2-30% in 3 min, 30-60% in 44 min, 60-95% in 8 min, followed by holding the gradient at 95% for 5 min, for a total run time of 65 min. The MS data were acquired in positive ion mode (500-1800 Da) using spray voltage of +5000 V. BioAnalyst software (Applied Biosystems) was used for spectral deconvolution. For the GFPuv protein analysis, a mass range of m/z 500-1800 was used for deconvolution and the output range was 10000-50000 Da using a step mass of 0.1 Da and a S/N threshold of

20. For the Z-Domain protein analysis, a mass range of m/z 500-2000 was used for deconvolution and the output range was 5000-15000 Da for Z-domain-His6X using a step mass of 0.1 Da and a S/N threshold of 20.

10. References

1 Y. Tatsu, Y. Shigeri, A. Ishida, I. Kameshita, H. Fujisawa and N. Yumoto, *Bioorg. Med. Chem. Lett*., 1999, **9**, 1093-1096.

2 W. Wan, Y. Huang, Z. Wang, W. K. Russell, P. J. Pai, D. H. Russell and W. Liu, *Angew. Chem. Int. Ed.,* 2010, **in press**.

3 S. Wiejek, E. Masiukiewiez, B. Rzeszotarska, *Chem. Pharm. Bull*., 2001, **49**, 1189-1191.

4 R. G. Dyer and K. D. Turnbull, *J. Org. Chem*., 1999, **64**, 7988-7995.

5 H. Yamamoto and J. T. Yang, *Biopolymers*, 1974, **13**, 1093-1107.

6 S. W. Santoro, L. Wang, B. Herberich, D. S. King and P. G. Schultz, *Nat. Biotechnol.*, 2002, **20**, 1044-1048.

7 Y. Huang, W. K. Russell, W. Wan, P. J. Pai, D. H. Russell and W. Liu, *Mol. Biosyst.*, 2010, **DOI**: 10.1039/b920120c.

8 Y. Huang, W. Wan, W. K. Russell, P. J. Pai, Z. Wang, D. H. Russell and W. Liu, *Bioorg. Med. Chem.*, 2010, **20**, 878-880.

9 L. Wang, Z. Zhang, A. Brock and P. G. Schultz, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 56-61.

10 J. Xie, L. Wang, N. Wu, A. Brock, G. Spraggon and P. G. Schultz, *Nat. Biotechnol.*, 2004, **22**, 1297-1301.

11 J. P. Tam, Q. Yu and Z. Miao, *Biopolymers*, 1999, **51**, 311-332.

11. Supplementary Tables, Schemes, and Figures

Supplementary Table 1. Evolved PylRS variants that charge pylT with different NAAs.

Position	305	306	309	346	348	477	Remark[1]
WT	L	Y	L	N	C	W	
mKRS1[2]	L	M	A	N	T	W	**5, 7, 6, 8**
mKRS2	L	V	A	N	A	W	**5, 7, 6**
mKRS3[3]	L	M	A	N	C	W	**5, 7, 6, 8**
mKRS4	L	A	A	H	L	W	**5, 6**
mKRS5	L	M	P	N	C	W	**5, 6, 7, 8**
mKRS6	L	M	A	N	S	W	**5**
mKRS7	L	Y	A	N	A	W	**5**
mKRS8[4]	L	M	T	N	A	W	**5**
mKRS9	L	A	A	N	A	W	**5**
mKRS10	L	A	L	N	A	W	**5, 7**
mKRS11[5]	L	A	L	N	C	W	**5, 7**

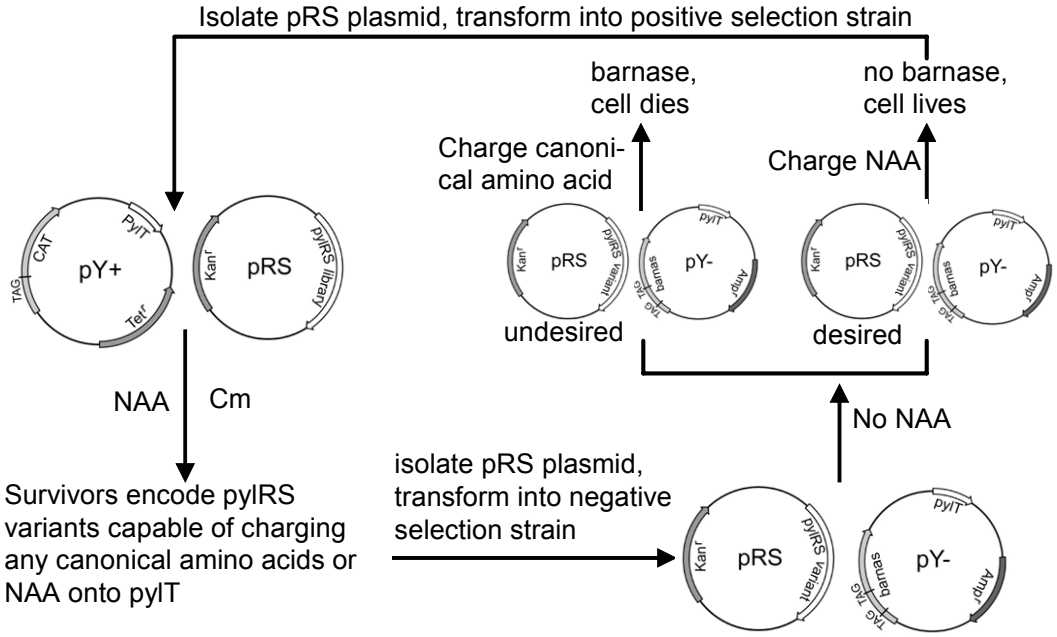
[1] This column represents the NAAs that can be taken by the mutant PylRS variants. The order of compounds also indicates decreasing encoding efficiency based on the screening results.

[2] The mutant was found seven times out of 22 sequenced mutants and has an extra mutation T364K.

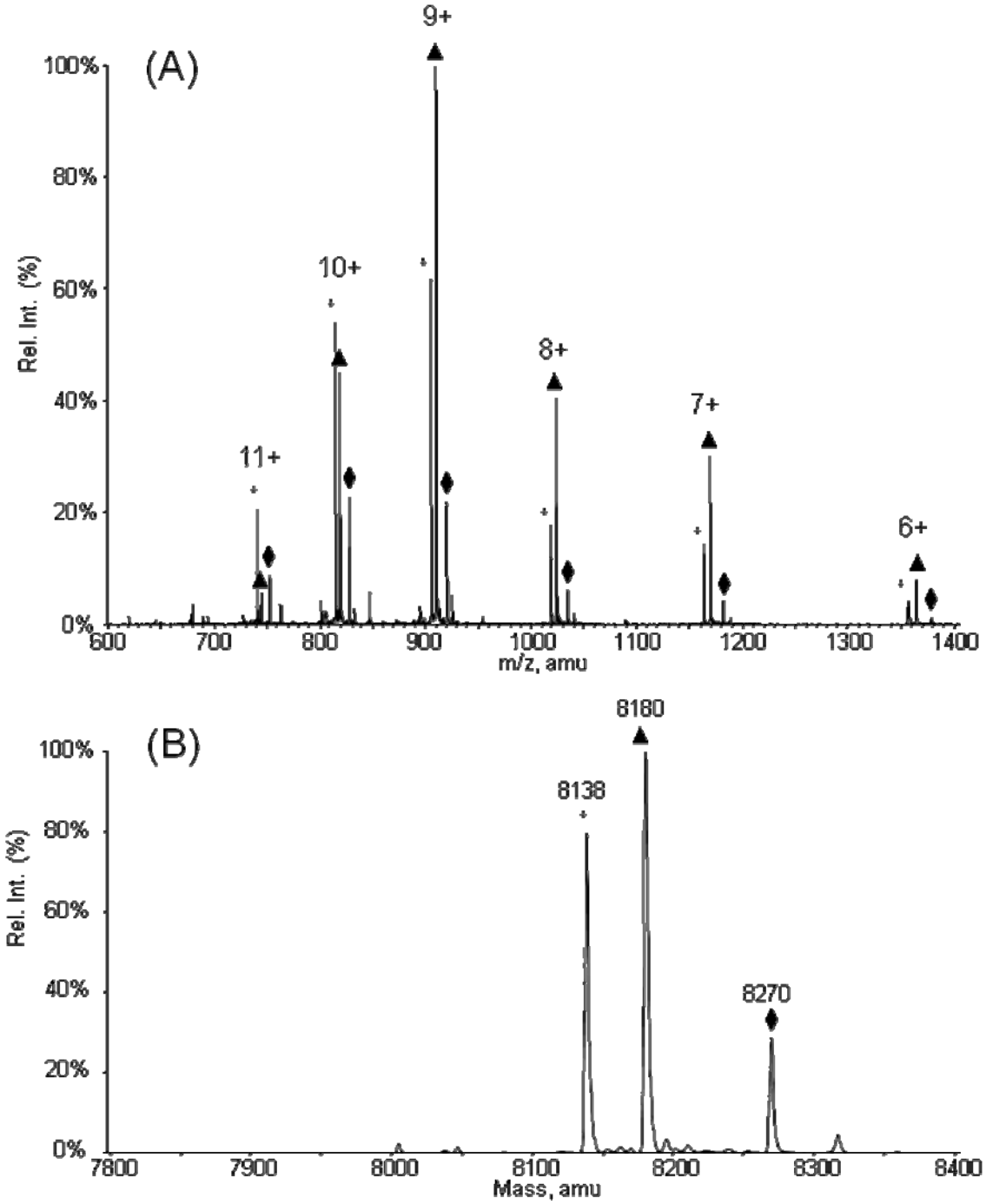
[3] The mutant was found six times out of 22 sequenced mutants.

[4]The mutant has extra mutation P297S.

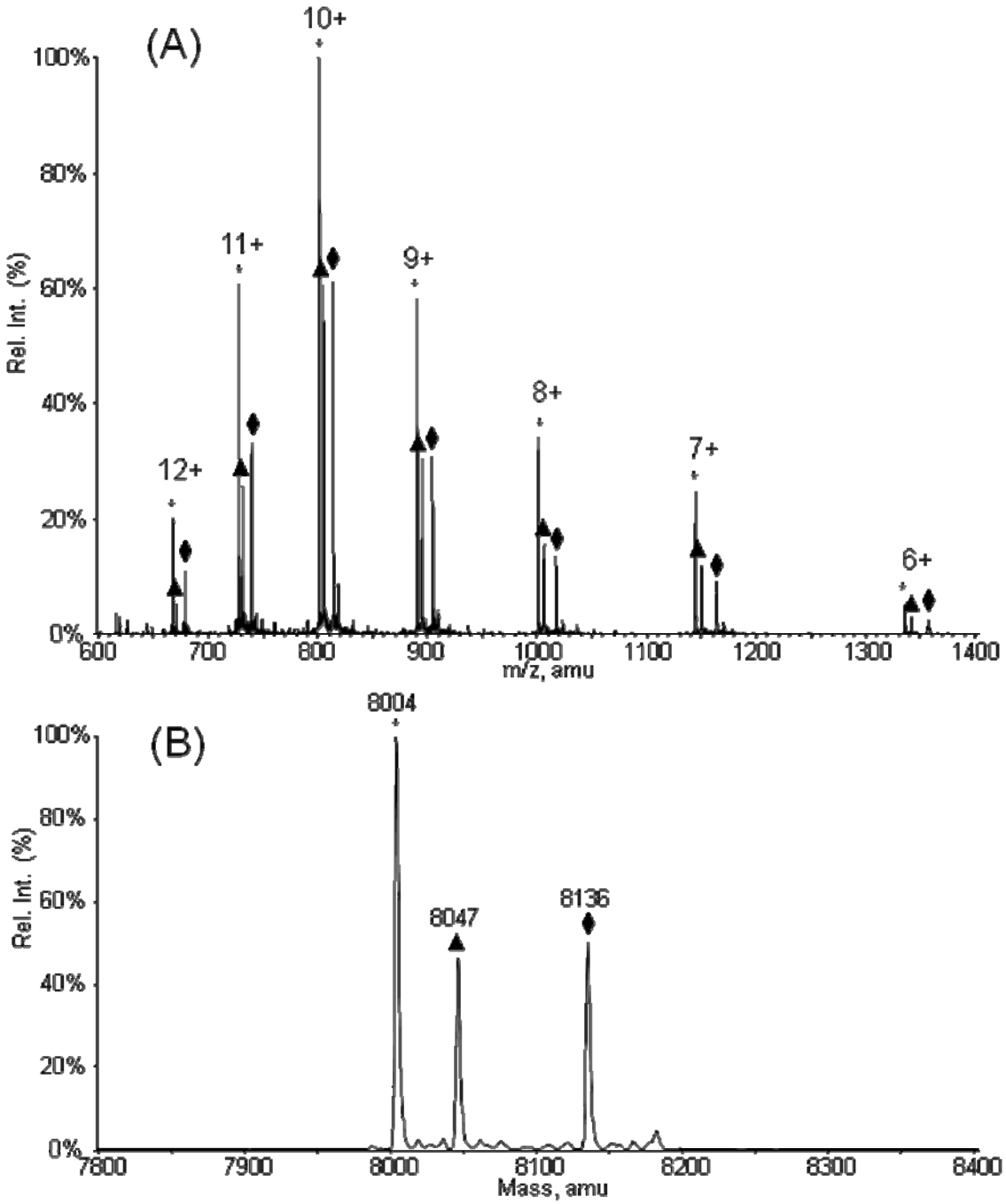
[5] The mutant was found twice out of 22 sequenced mutants.



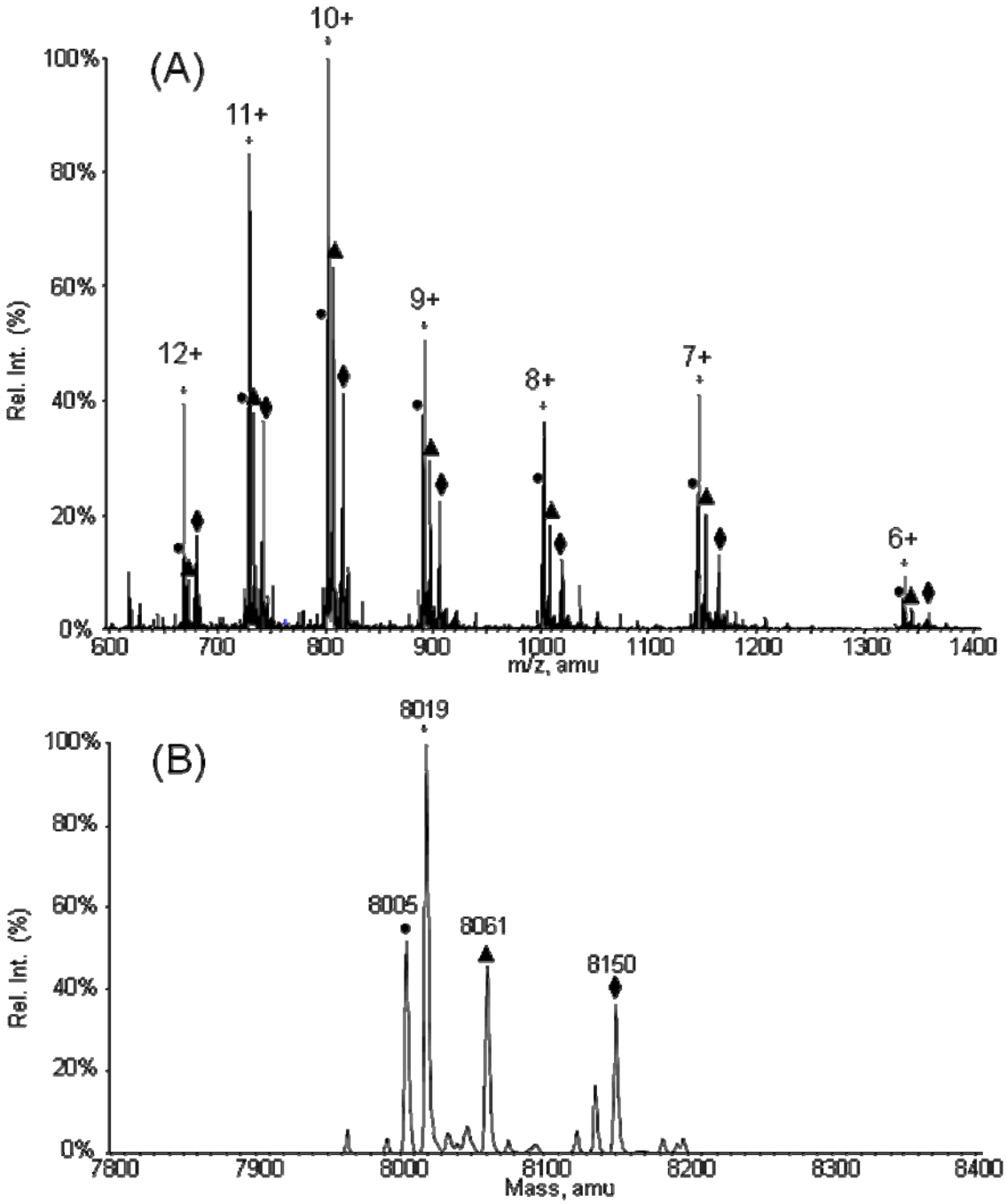
Supplementary Scheme 3. The selection scheme to identify PylRS variants specific for a noncanonical amino acid.



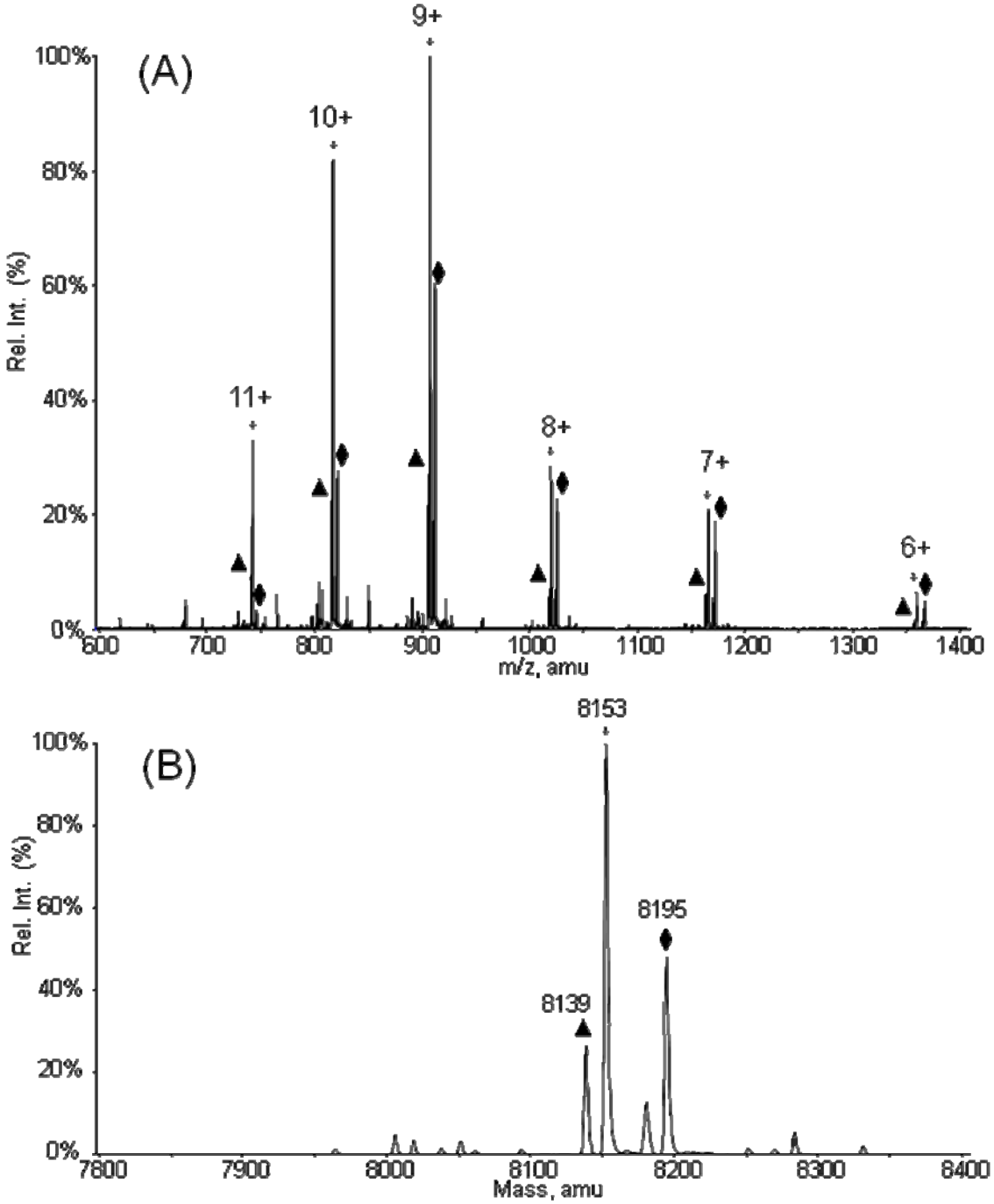
Supplementary Fig. 1. Mass determination of Protein **Z-5** (A) ESI-MS spectrum of **Z-5** (B) The deconvoluted ESI-MS spectrum of **Z-5**.



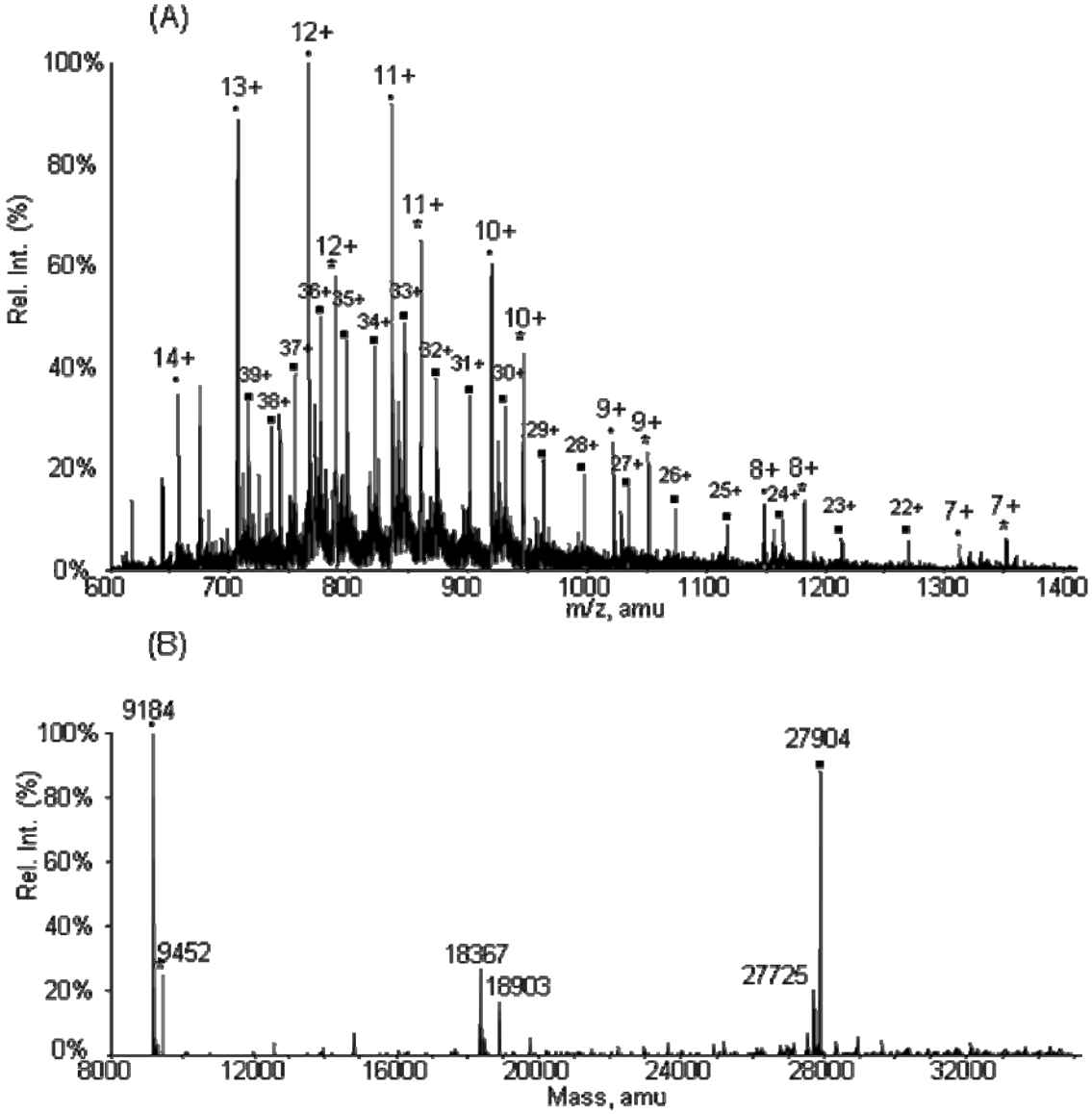
Supplementary Fig. 2. Mass determination of Protein **Z-6** (A) ESI-MS spectrum of **Z-6** (B) The deconvoluted ESI-MS spectrum of **Z-6**.



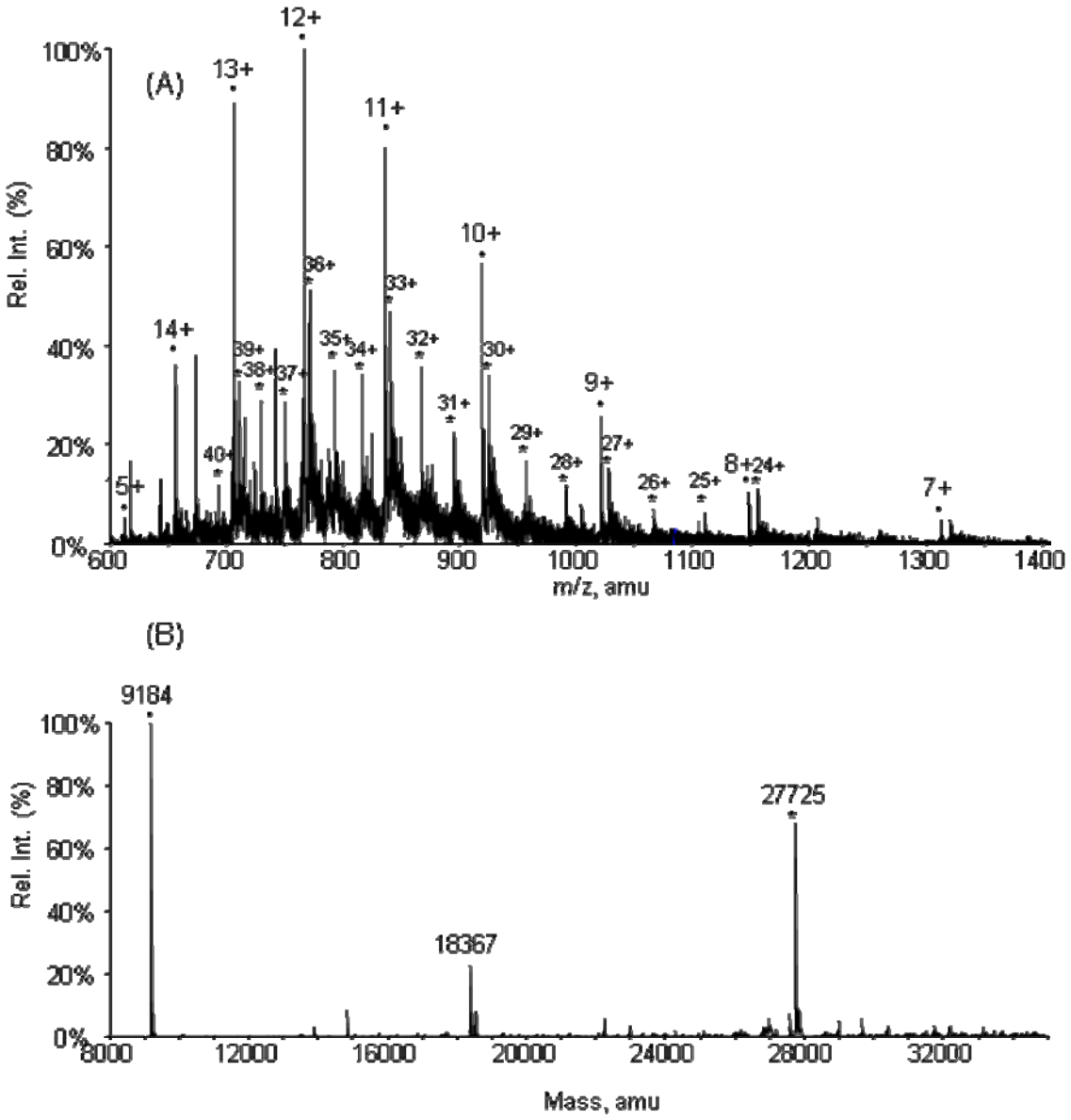
Supplementary Fig. 3. Mass determination of Protein **Z-7** (A) ESI-MS spectrum of **Z-7** (B) The deconvoluted ESI-MS spectrum of **Z-7**.



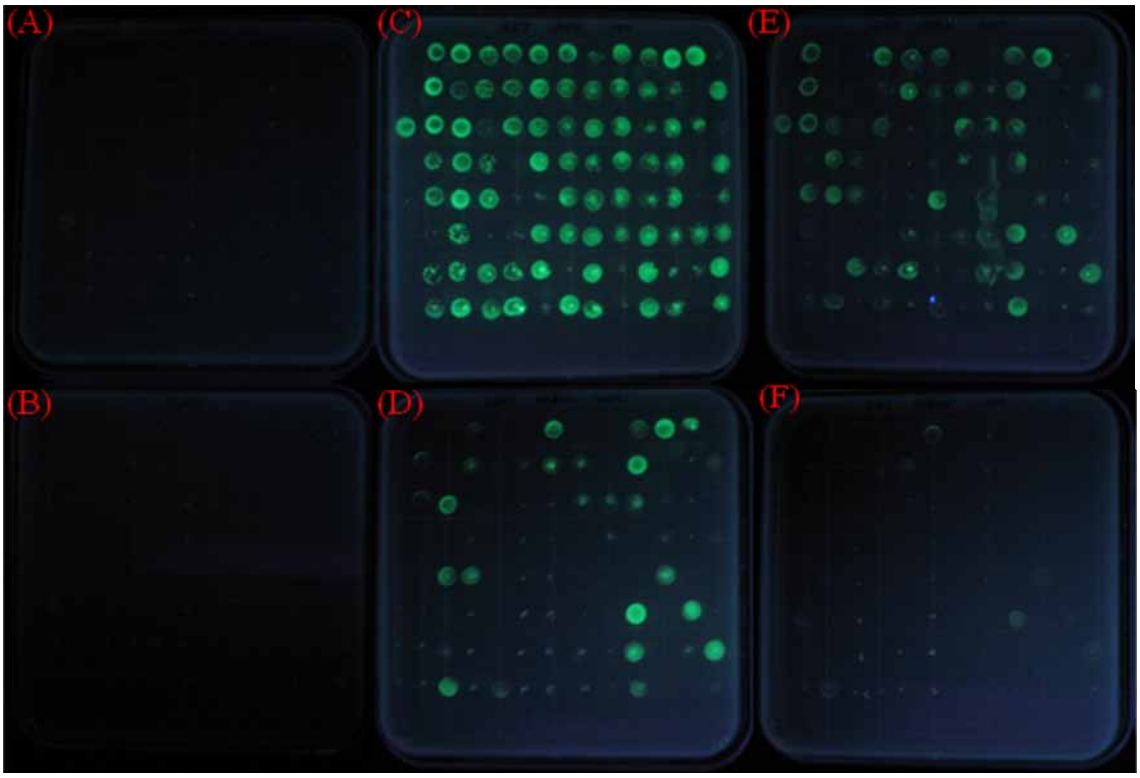
Supplementary Fig. 4. Mass determination of Protein **Z-8** (A) ESI-MS spectrum of **Z-8** (B) The deconvoluted ESI-MS spectrum of **Z-8**.



Supplementary Fig. 5. Mass determination of Protein **GFP-8** (A) ESI-MS spectrum of **GFP-8** (B) The deconvoluted ESI-MS spectrum of **GFP-8**.

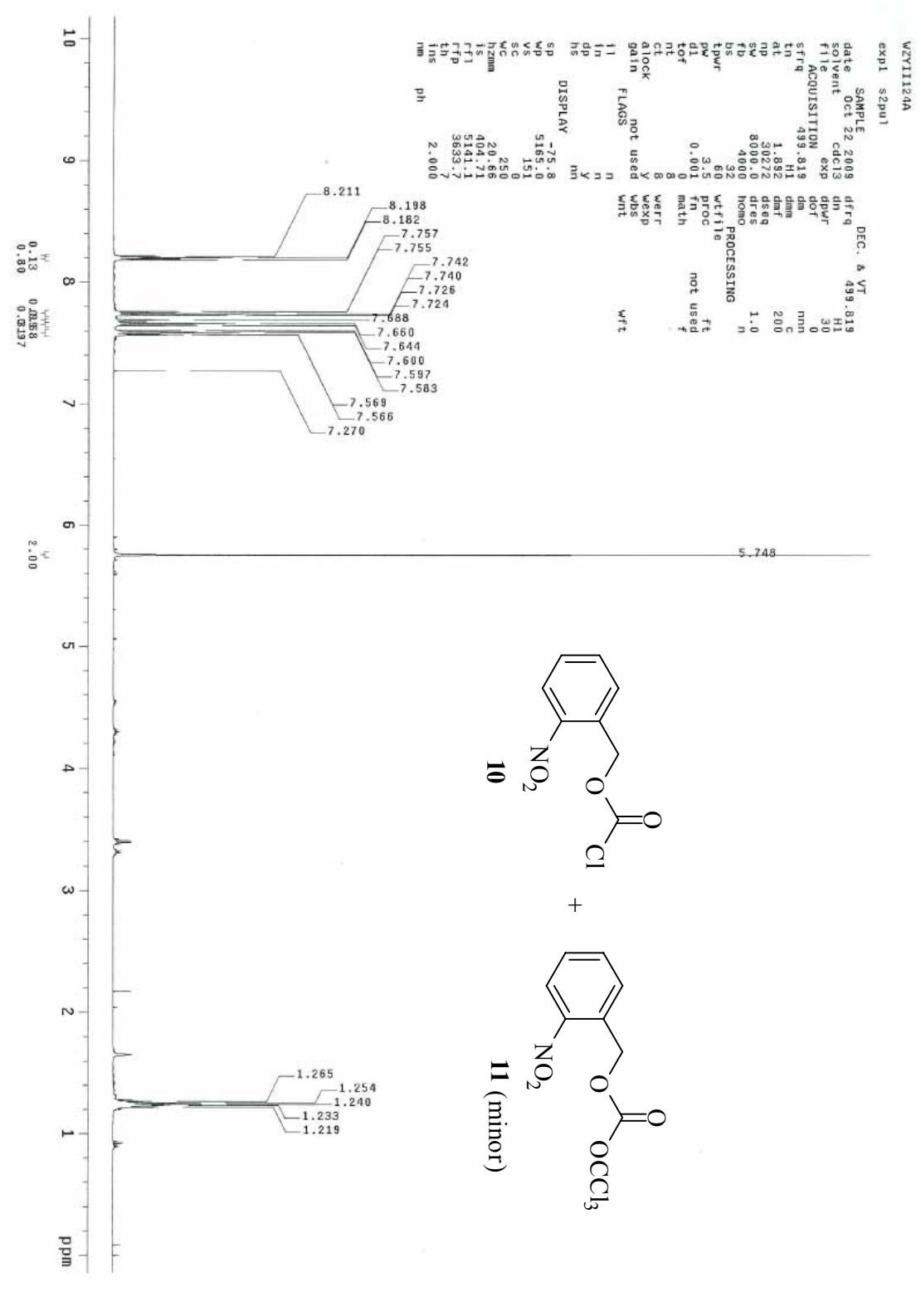


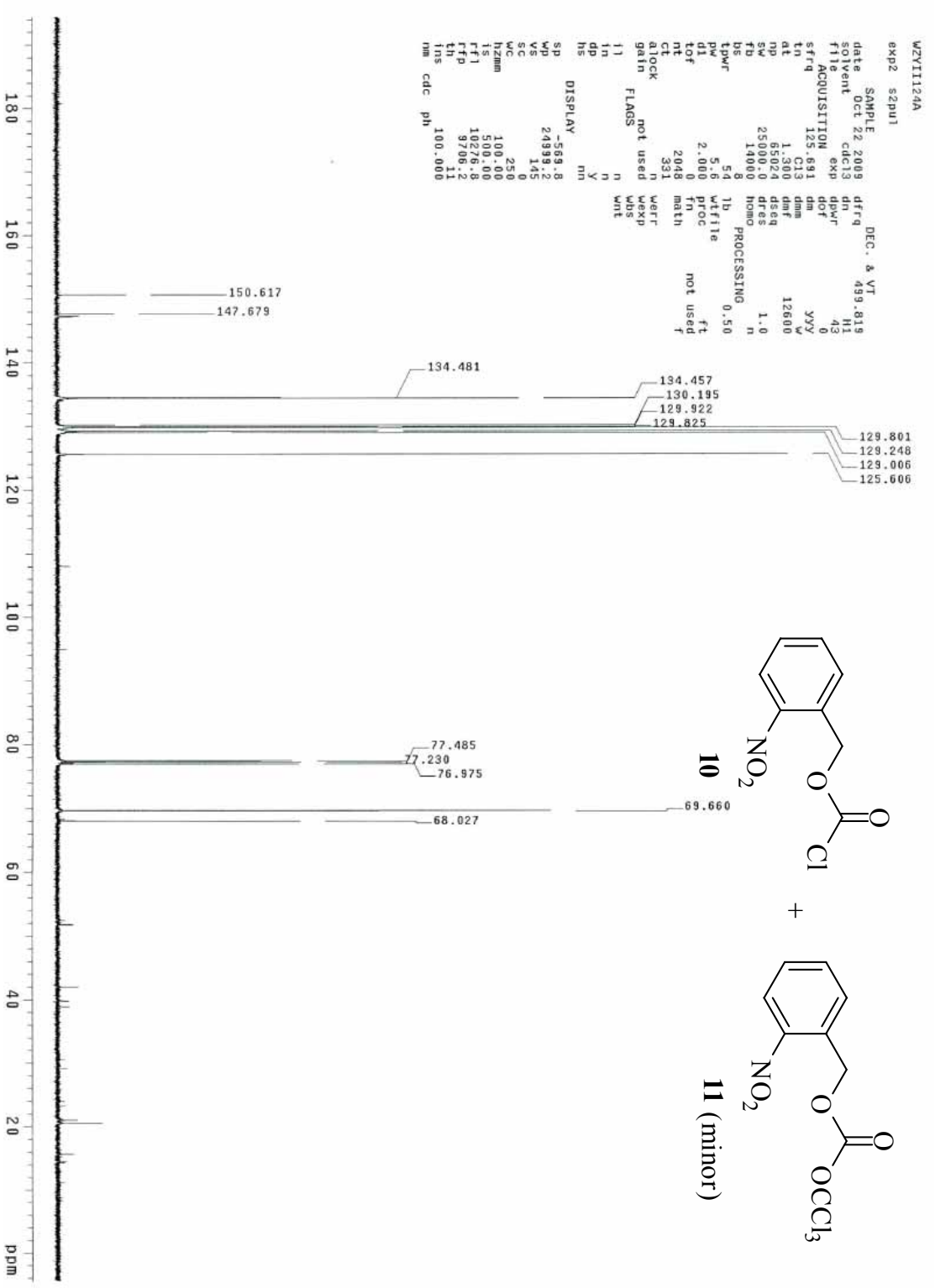
Supplementary Fig. 6. Mass determination of Protein **GFP-8** after photolysis. (A) ESI-MS spectrum of **GFP-8** (B) The deconvoluted ESI-MS spectrum of **GFP-8**.

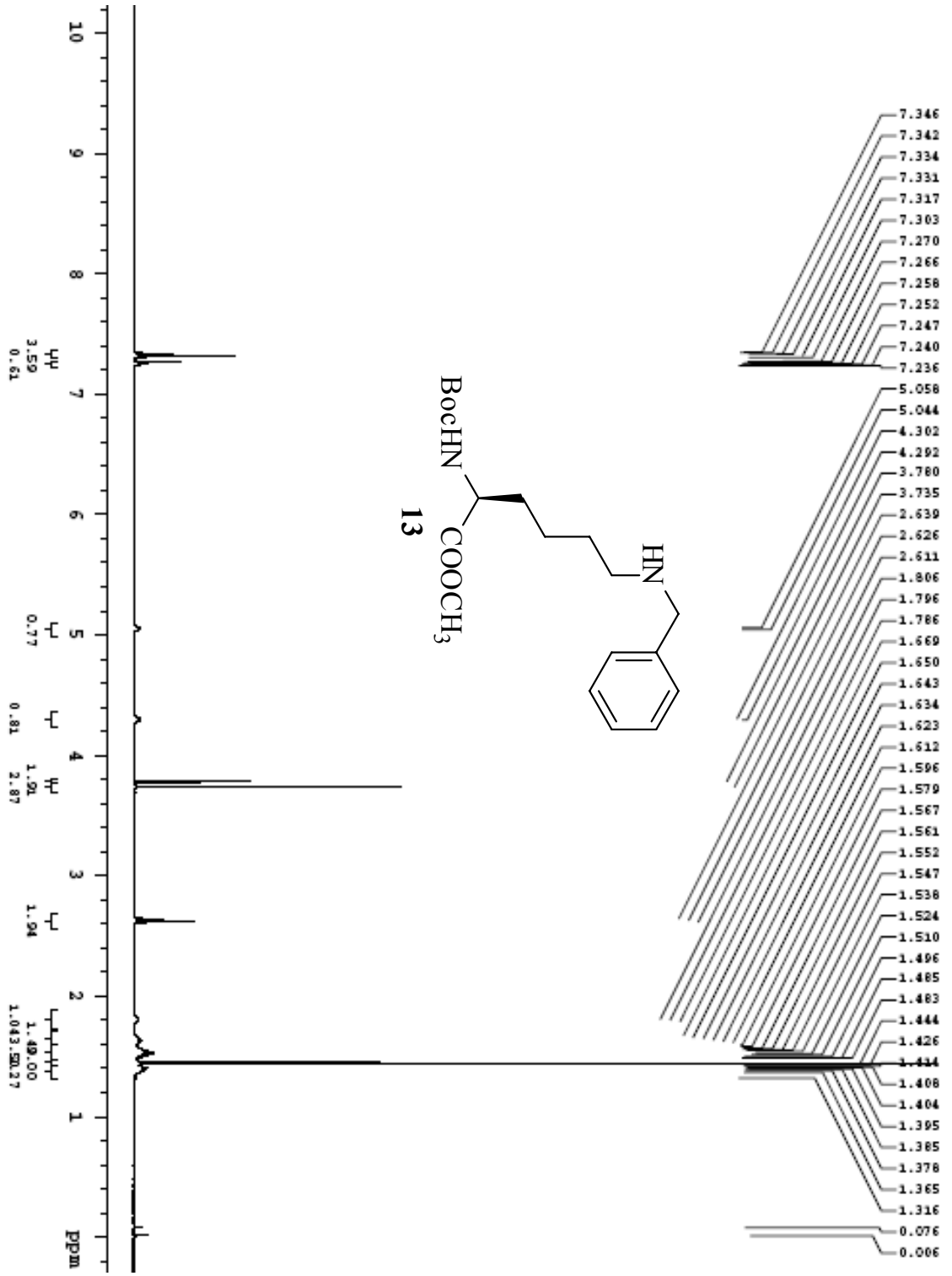


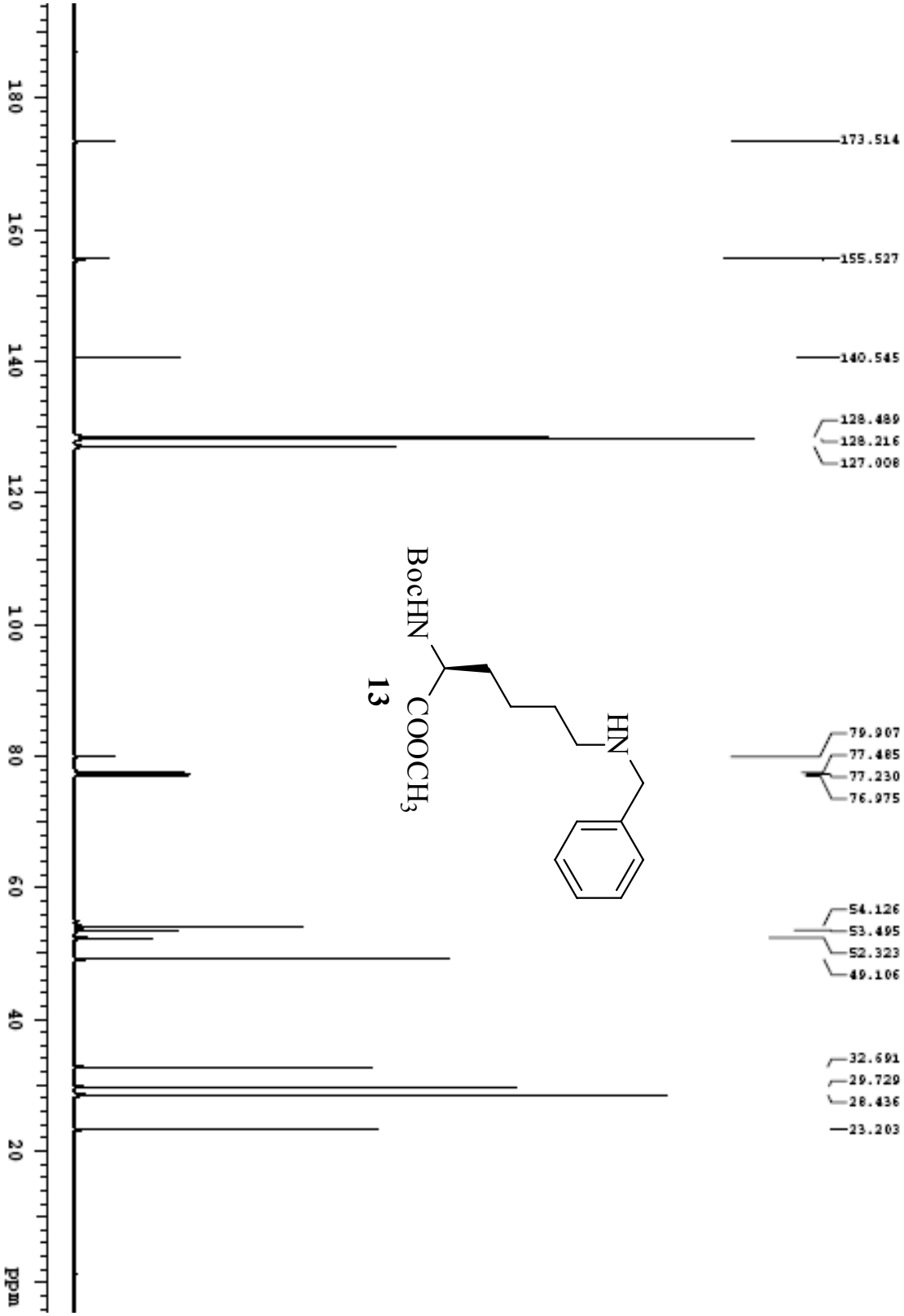
Supplementary Fig. 7. Growth of 96 single colonies from the final third positive selection of **5** on LB plates with different supplements: (**A**) 68 μg/mL Cm, 25 μg/mL Kan and 12 μg/mL Tet; (**B**) 102 μg/mL Cm, 25 μg/mL Kan and 12 μg/mL Tet; (**C**) 1 mM **5**, 102 μg/mL Cm, 25 μg/mL Kan and 12 μg/mL Tet; (**D**) 1 mM **6**, 102 μg/mL Cm, 25 μg/mL Kan and 12 μg/mL Tet; (**E**) 1 mM **7**, 102 μg/mL Cm, 25 μg/mL Kan and 12 μg/mL Tet; (**F**) 1 mM **8**, 102 μg/mL Cm, 25 μg/mL Kan and 12 μg/mL Tet. The pY+ plasmid has a GFP_{UV} gene under control of a T7 promoter. Its expression is promoted by the suppression of two amber mutations at positions 1 and 107 of a T7 RNA polymerase gene in pREP. The fluorescent intensity of the expressed full-length GFP_{UV} roughly represents the suppression efficiency at amber codons.

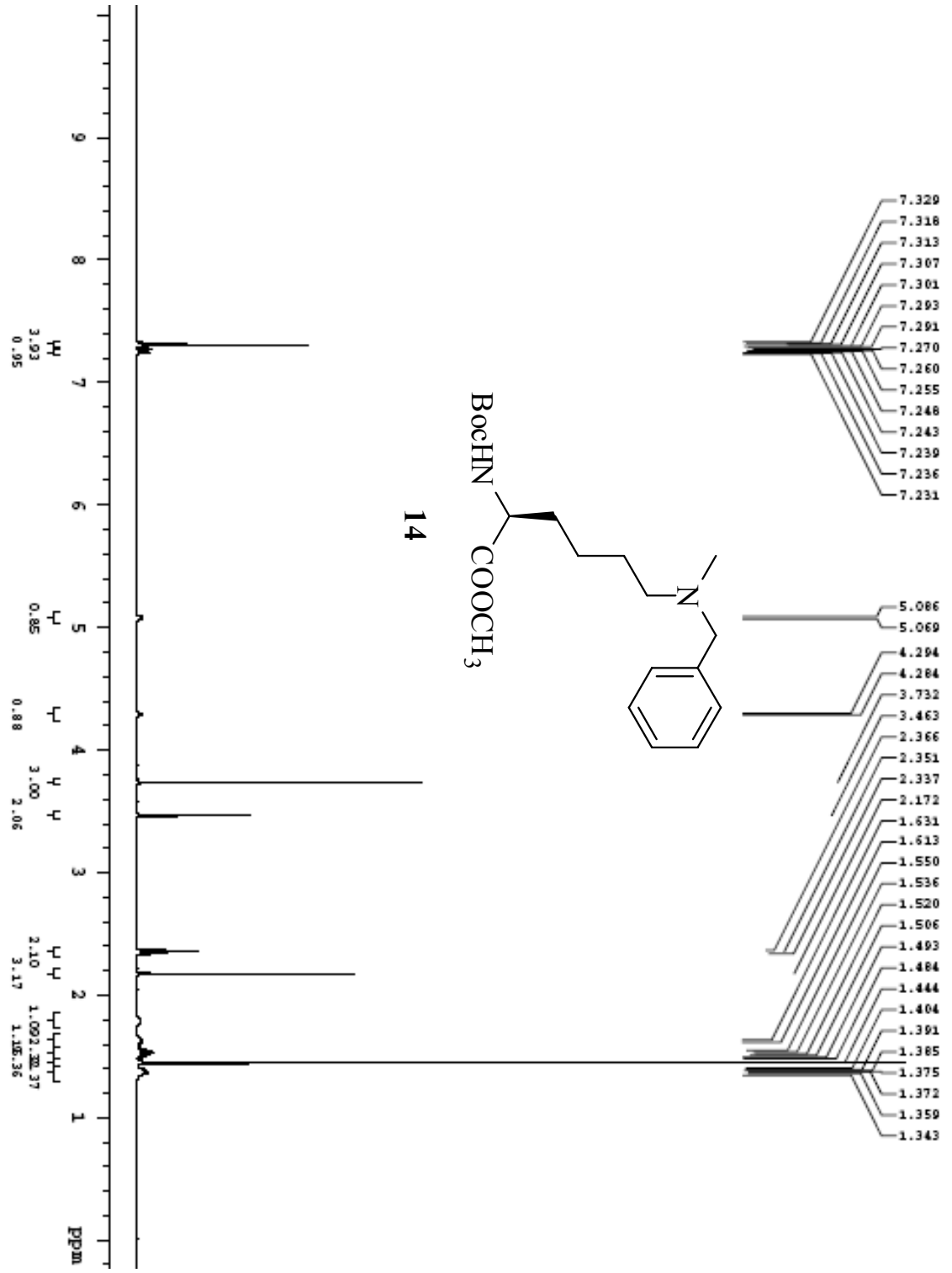
13. NMR and MS Spectra of Synthesized Compounds

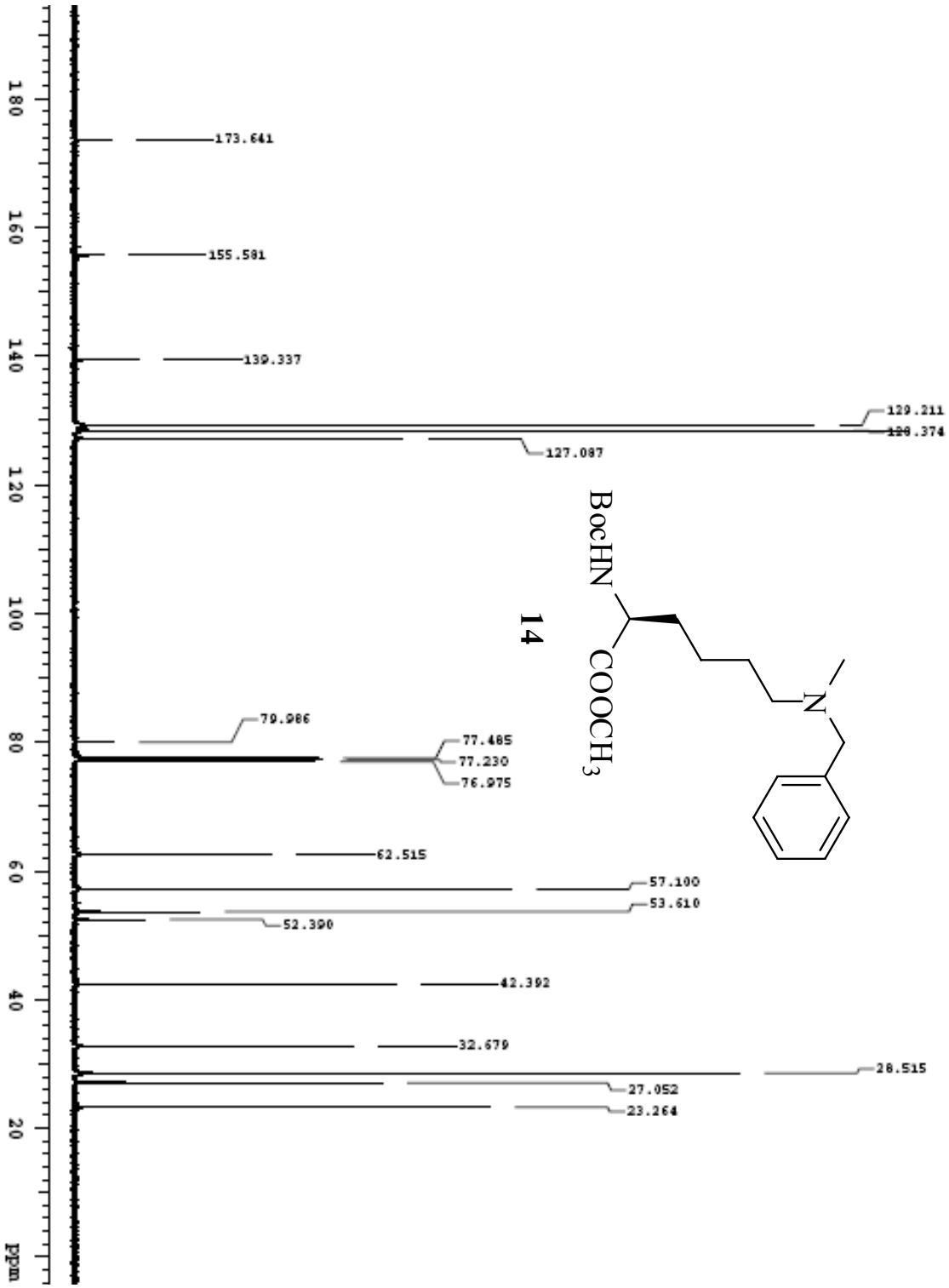


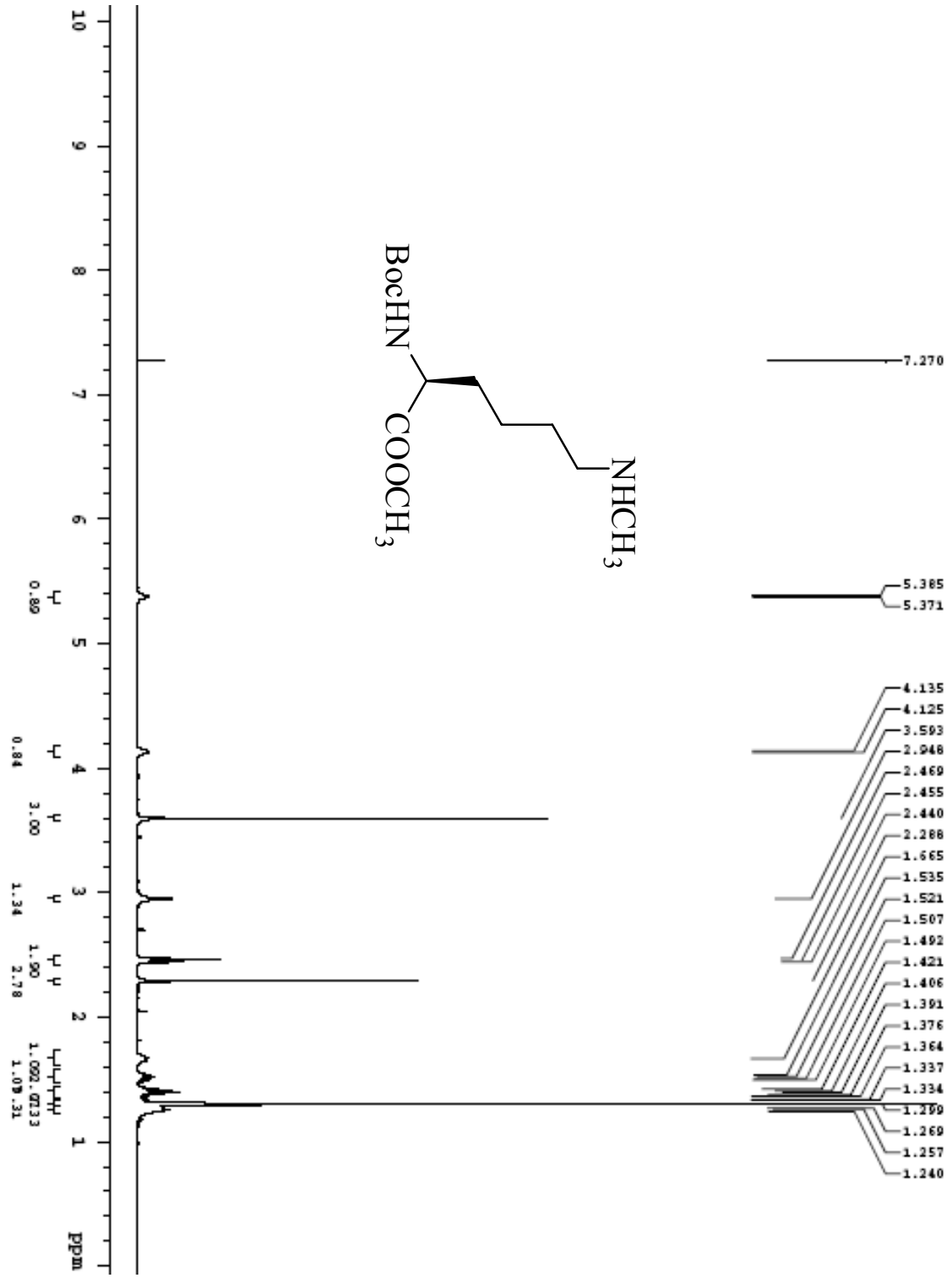

WZYII124A
150.617
147.679
134.481
134.457
130.195
129.922
129.825
129.801
129.248
129.006
125.606
77.485
77.230
76.975
69.660
68.027
180
160
140
120
100
80
60
40
20
ppm
10
11 (minor)
+

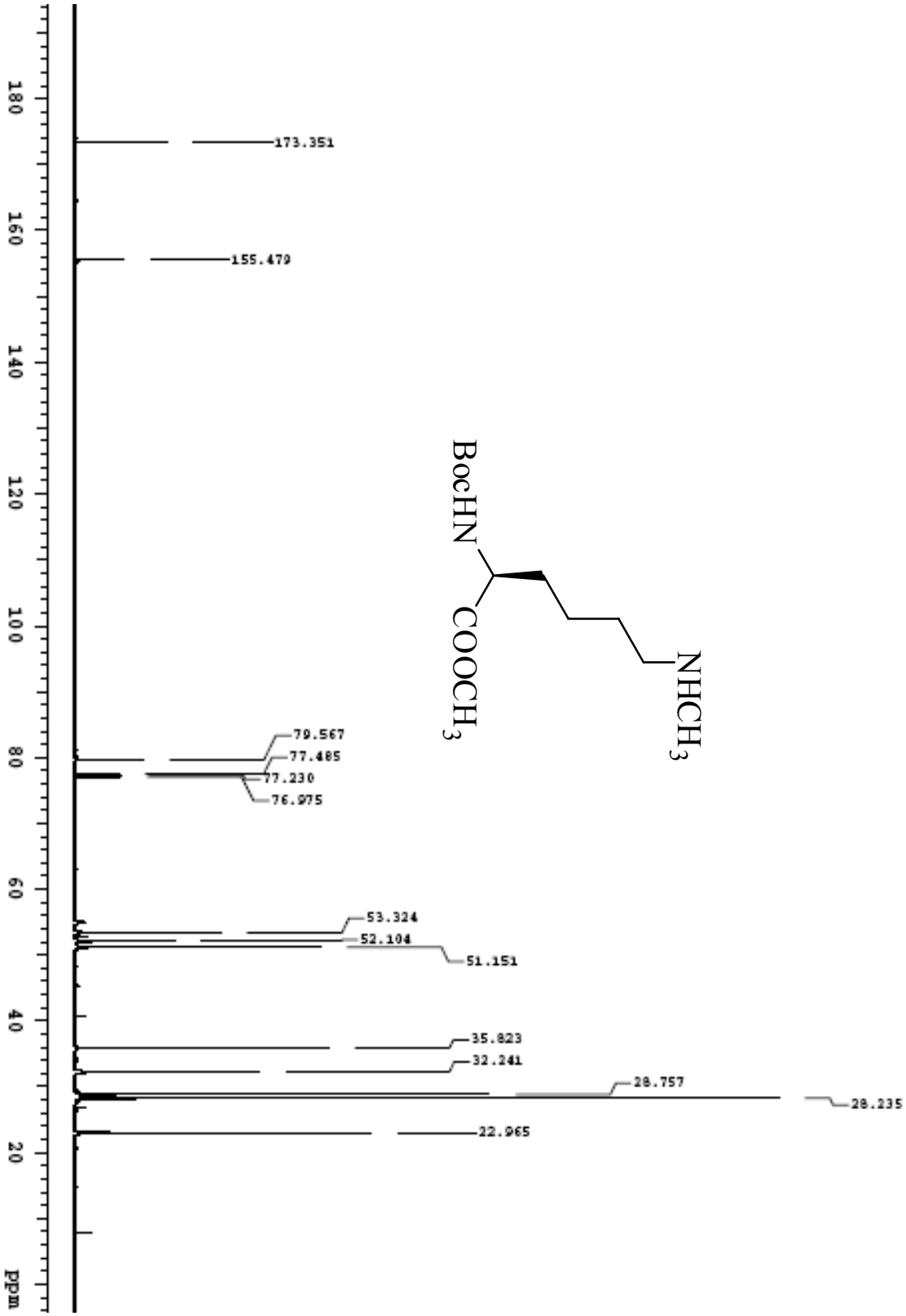

BocHN
COOCH3
HN
13
7.346
7.342
7.334
7.331
7.317
7.303
7.270
7.266
7.258
7.252
7.247
7.240
7.236
5.058
5.044
4.302
4.292
3.780
3.735
2.639
2.626
2.611
1.806
1.796
1.786
1.669
1.650
1.643
1.634
1.623
1.612
1.596
1.579
1.567
1.561
1.552
1.547
1.538
1.524
1.510
1.496
1.485
1.483
1.444
1.426
1.414
1.408
1.404
1.395
1.385
1.378
1.365
1.316
0.076
0.006
ppm

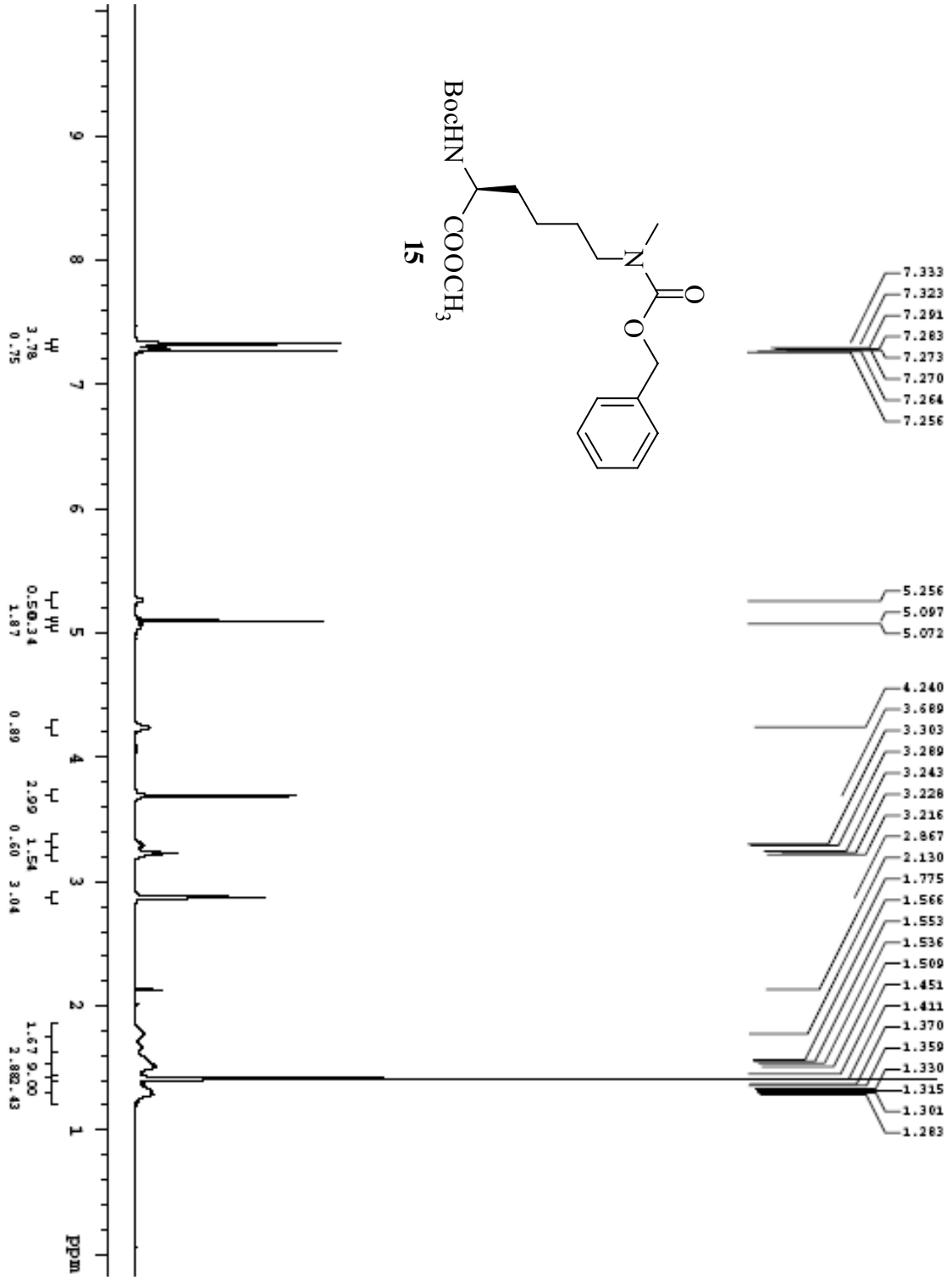

173.514
155.527
140.545
128.489
128.216
127.008
79.907
77.485
77.230
76.975
54.126
53.495
52.323
49.106
32.691
29.729
28.436
23.203
180
160
140
120
100
80
60
40
20
ppm
BocHN
COOCH3
HN
13

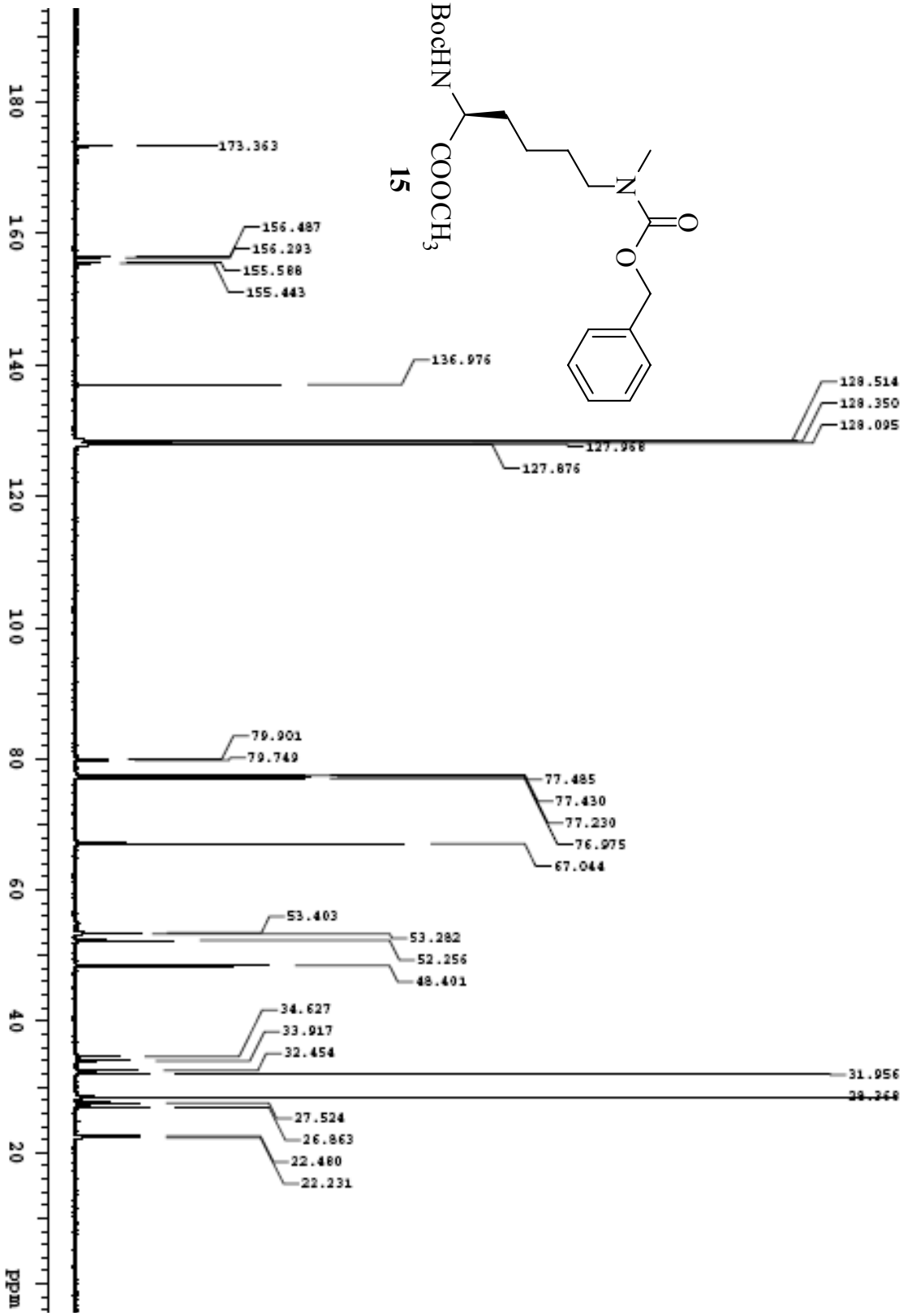

7.329
7.318
7.313
7.307
7.301
7.293
7.291
7.270
7.260
7.255
7.248
7.243
7.239
7.236
7.231
5.086
5.069
4.294
4.284
3.732
3.463
2.366
2.351
2.337
2.172
1.631
1.613
1.550
1.536
1.520
1.506
1.493
1.484
1.444
1.404
1.391
1.385
1.375
1.372
1.359
1.343
BocHN
COOCH3
N
14
ppm

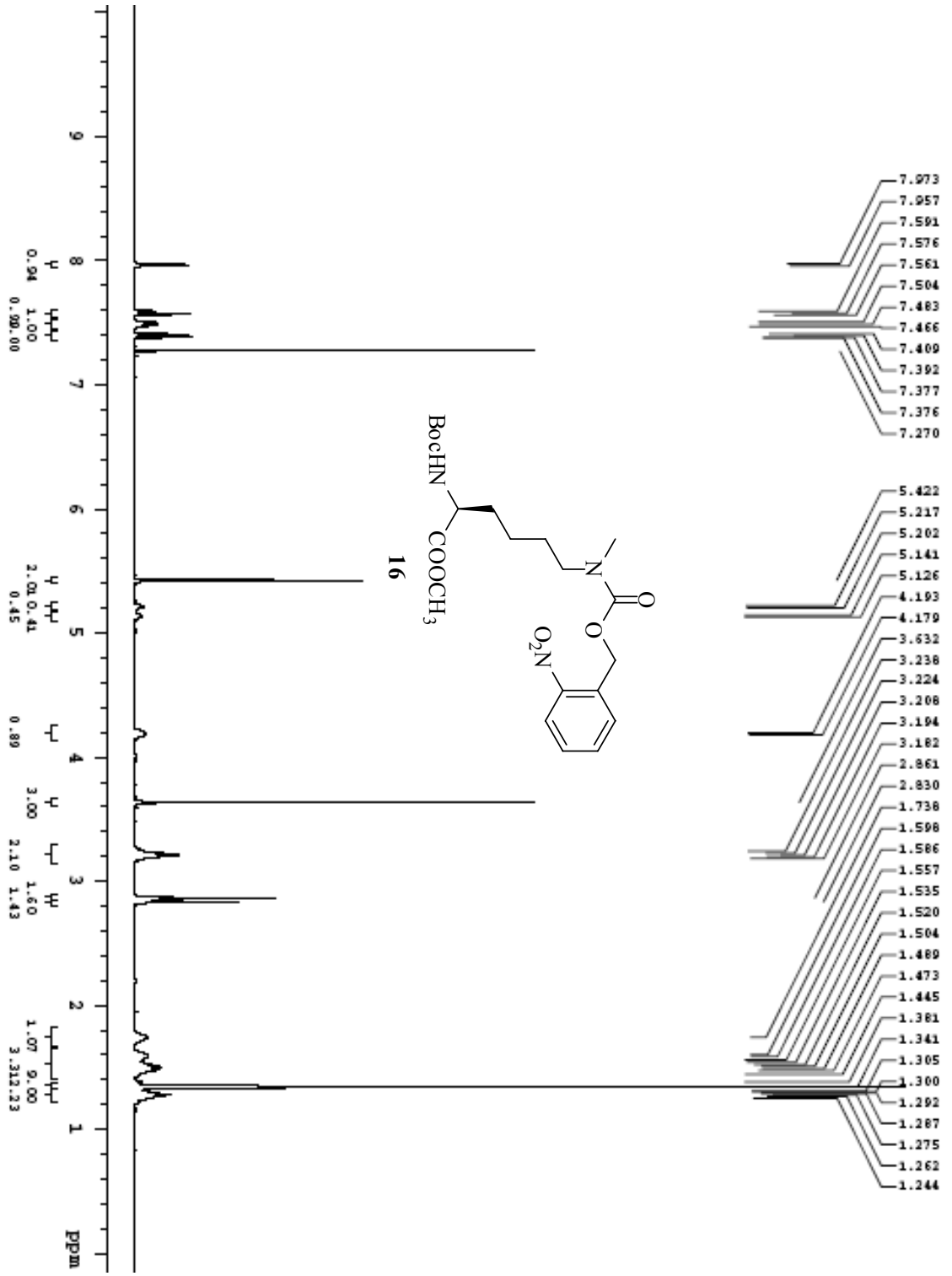

173.641
155.581
139.337
129.211
128.374
127.087
79.986
77.485
77.230
76.975
62.515
57.100
53.610
52.390
42.392
32.679
28.515
27.052
23.264
ppm
BocHN
COOCH3
N
14

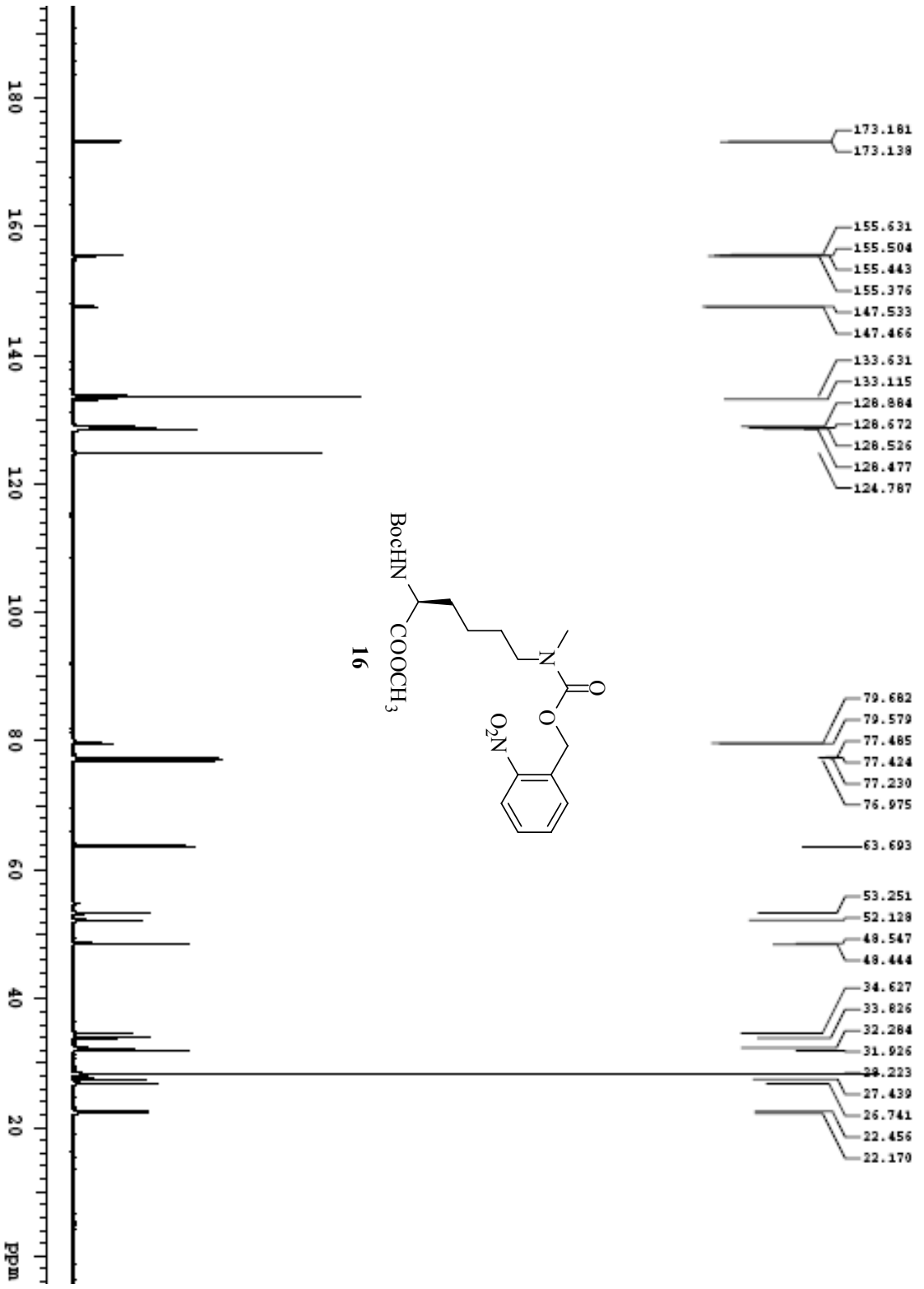

BocHN
COOCH3
NHCH3
7.270
5.385
5.371
4.135
4.125
3.593
2.948
2.469
2.455
2.440
2.288
1.665
1.535
1.521
1.507
1.492
1.421
1.406
1.391
1.376
1.364
1.337
1.334
1.299
1.269
1.257
1.240
0.89
0.84
3.00
1.34
1.00
2.78
1.00
1.09
31
ppm

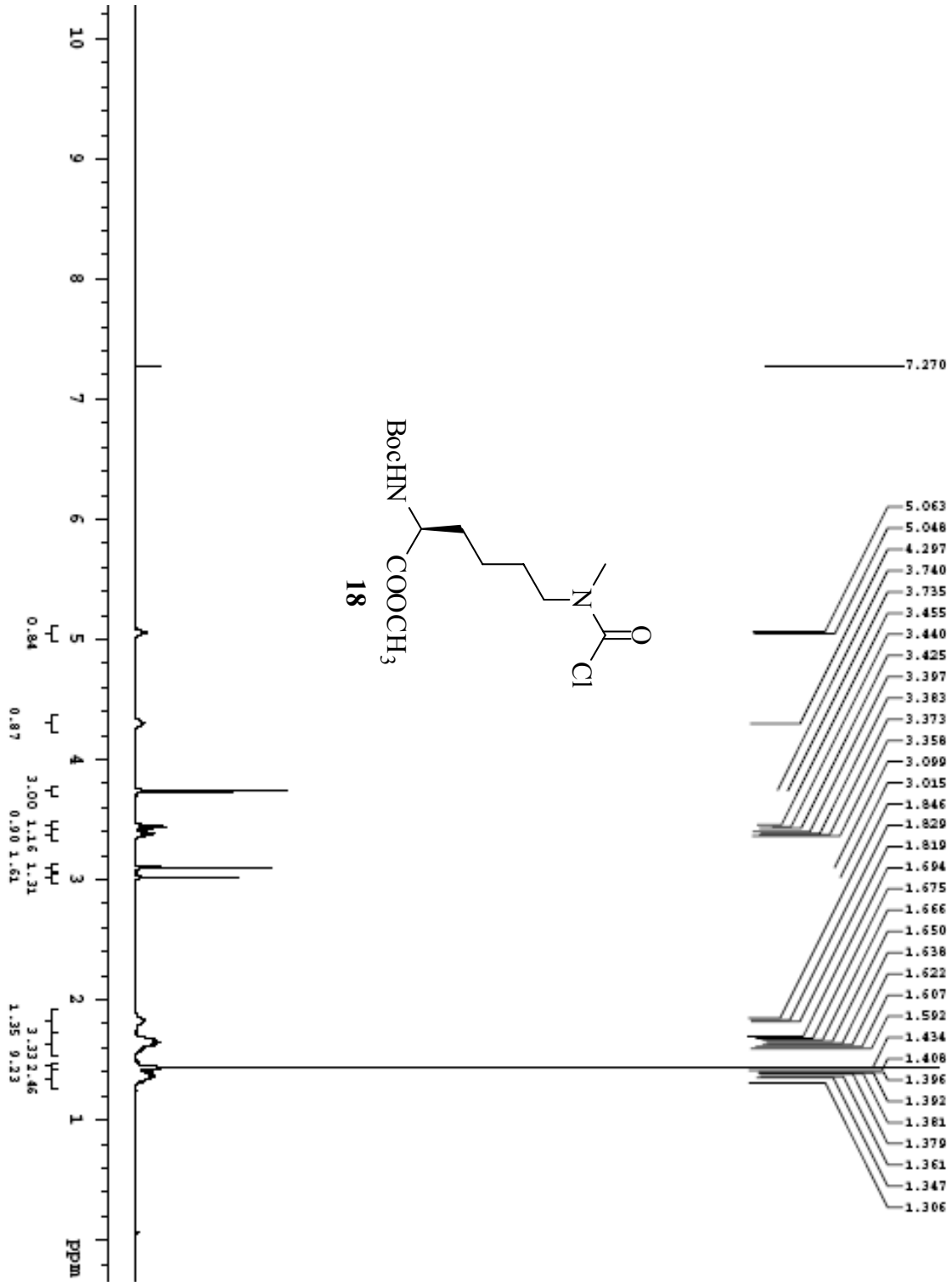

180
160
140
120
100
80
60
40
20
ppm
173.351
155.479
79.567
77.485
77.230
76.975
53.324
52.104
51.151
35.823
32.241
28.757
28.235
22.965
BocHN
COOCH3
NHCH3

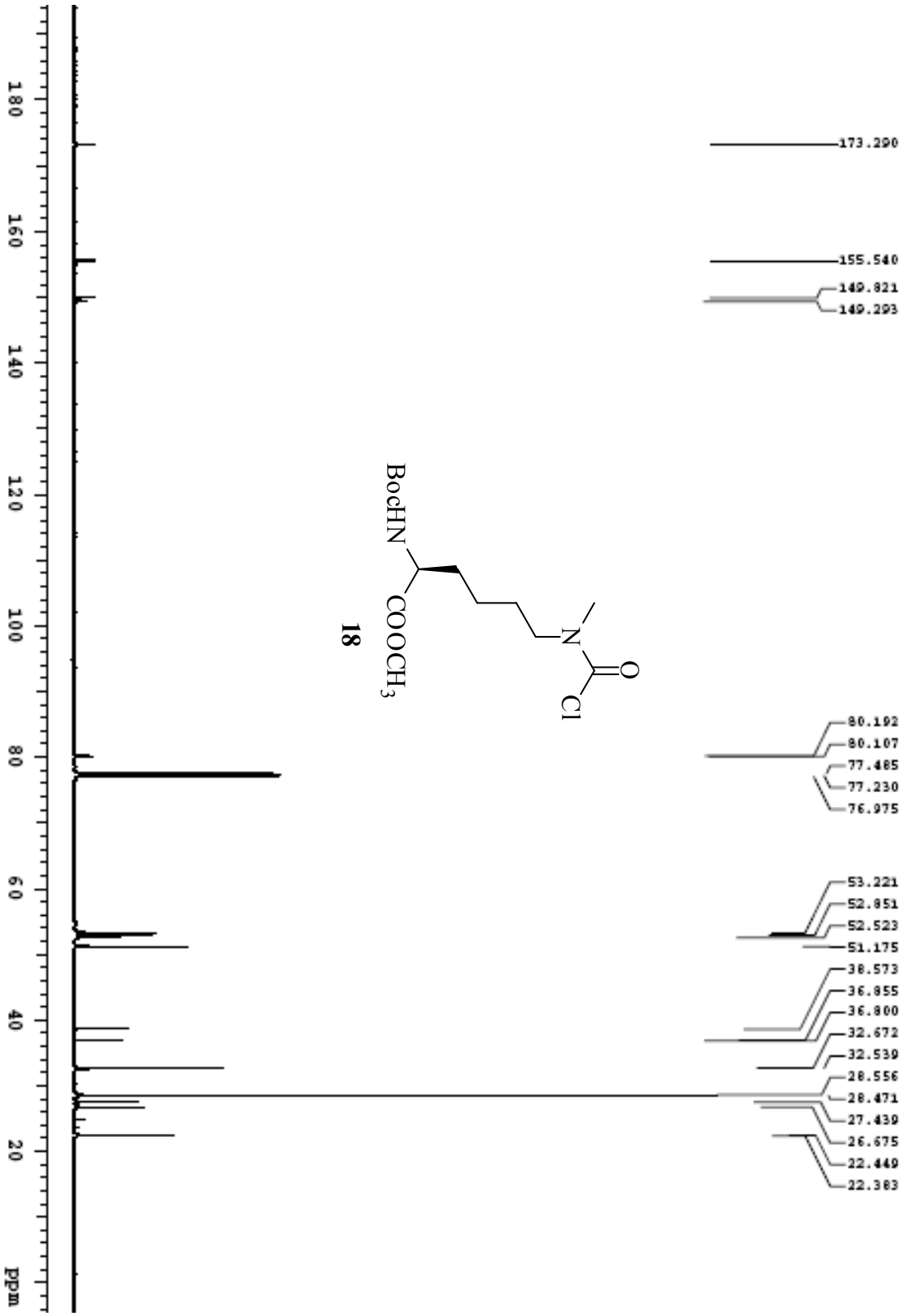

BocHN
COOCH3
15
N
O
O
7.333
7.323
7.291
7.283
7.273
7.270
7.264
7.256
5.256
5.097
5.072
4.240
3.689
3.303
3.289
3.243
3.228
3.216
2.867
2.130
1.775
1.566
1.553
1.536
1.509
1.451
1.411
1.370
1.359
1.330
1.315
1.301
1.283
ppm

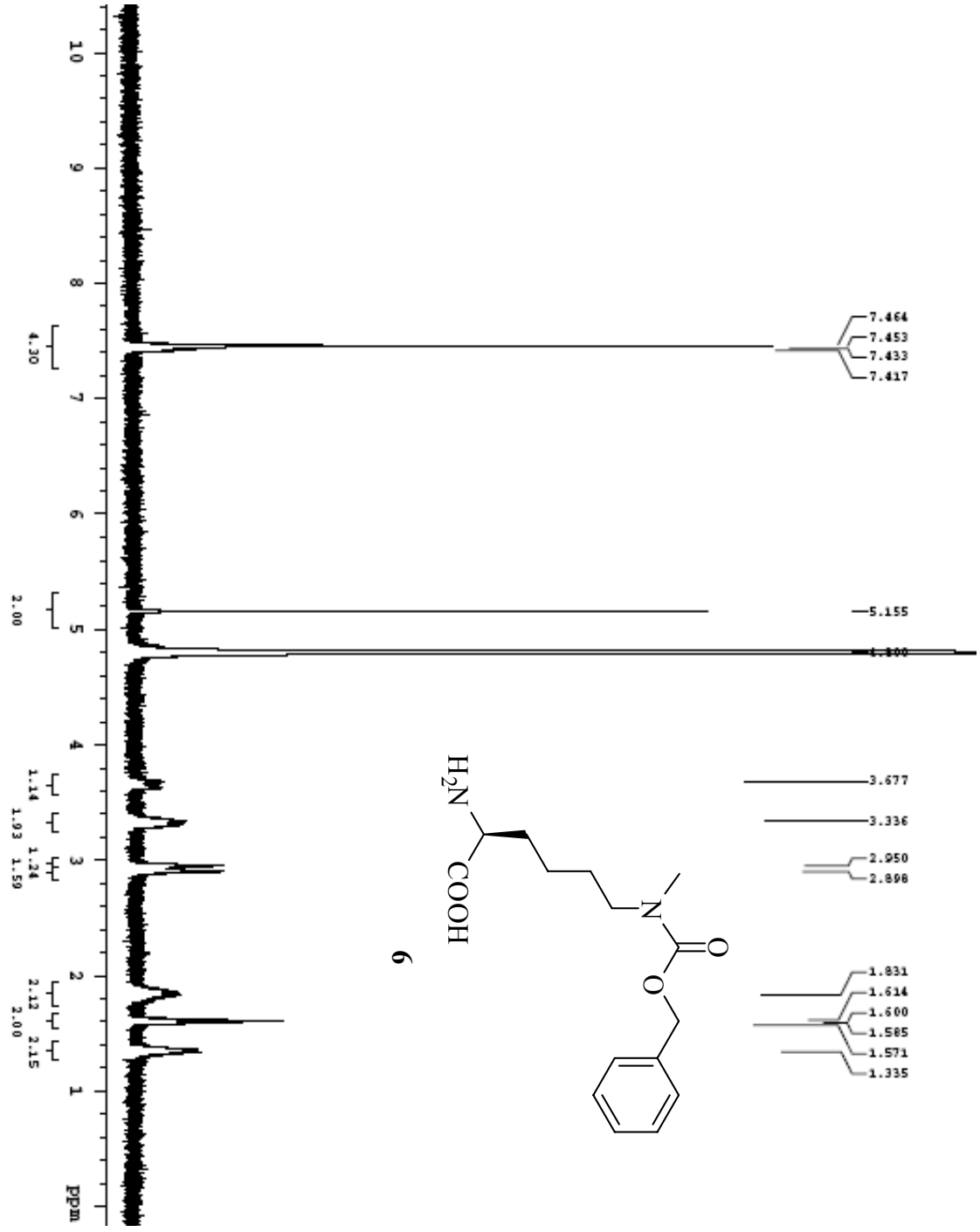

BocHN
COOCH3
15
N
O
O
173.363
156.487
156.293
155.588
155.443
136.976
128.514
128.350
128.095
127.968
127.676
79.901
79.749
77.485
77.430
77.230
76.975
67.044
53.403
53.282
52.256
48.401
34.627
33.917
32.454
31.956
28.368
27.524
26.863
22.480
22.231
180
160
140
120
100
80
60
40
20
ppm


7.973
7.957
7.591
7.576
7.561
7.504
7.483
7.466
7.409
7.392
7.377
7.376
7.270
5.422
5.217
5.202
5.141
5.126
4.193
4.179
3.632
3.238
3.224
3.208
3.194
3.182
2.861
2.830
1.738
1.598
1.586
1.557
1.535
1.520
1.504
1.489
1.473
1.445
1.381
1.341
1.305
1.300
1.292
1.287
1.275
1.262
1.244
BocHN
COOCH3
16
O2N
ppm

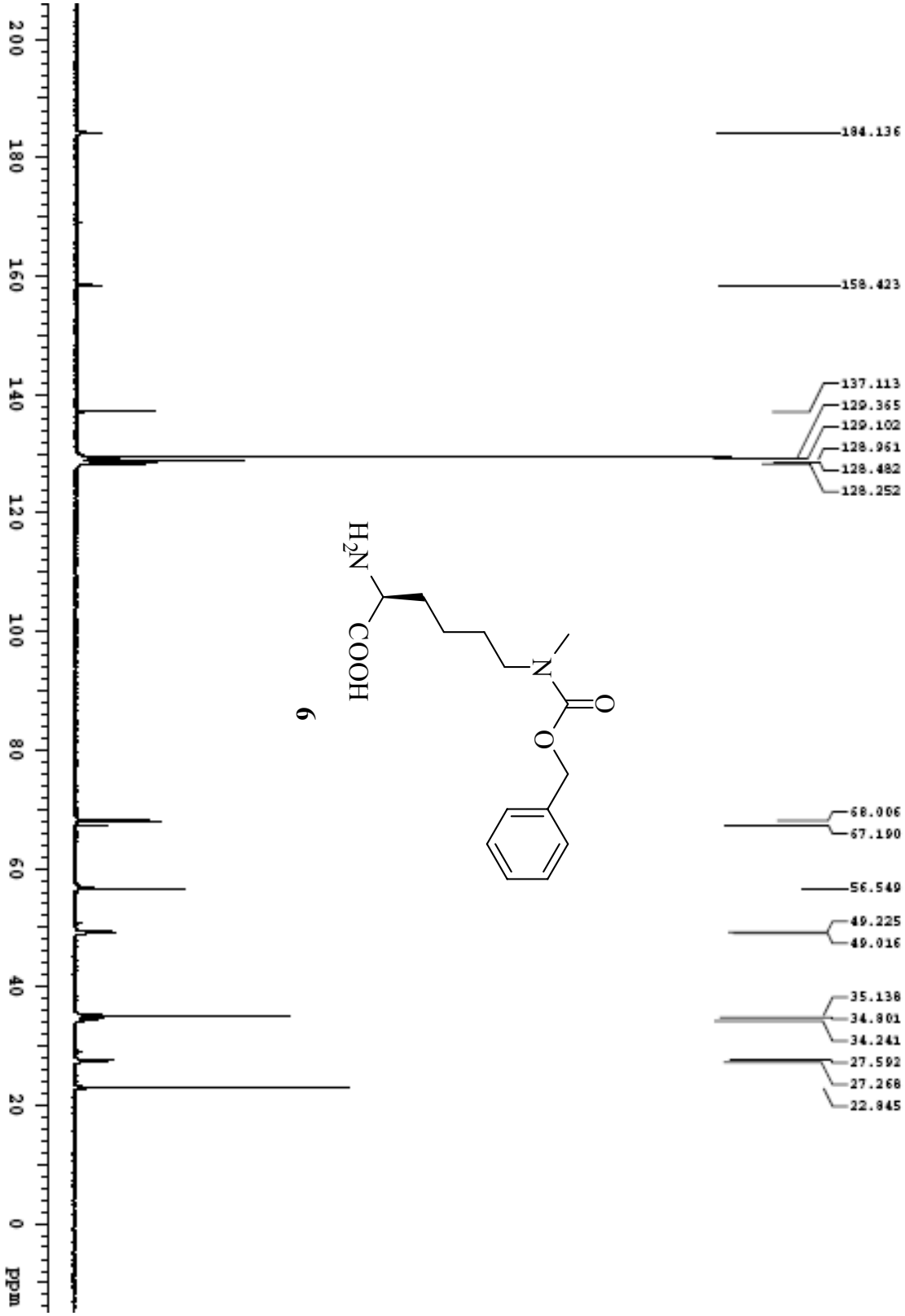

173.181
173.138
155.631
155.504
155.443
155.376
147.533
147.466
133.631
133.115
128.884
128.672
128.526
128.477
124.787
79.682
79.579
77.485
77.424
77.230
76.975
63.693
53.251
52.128
48.547
48.444
34.627
33.826
32.284
31.926
28.223
27.439
26.741
22.456
22.170
180
160
140
120
100
80
60
40
20
ppm
BocHN
COOCH3
16
N
O
O
O2N

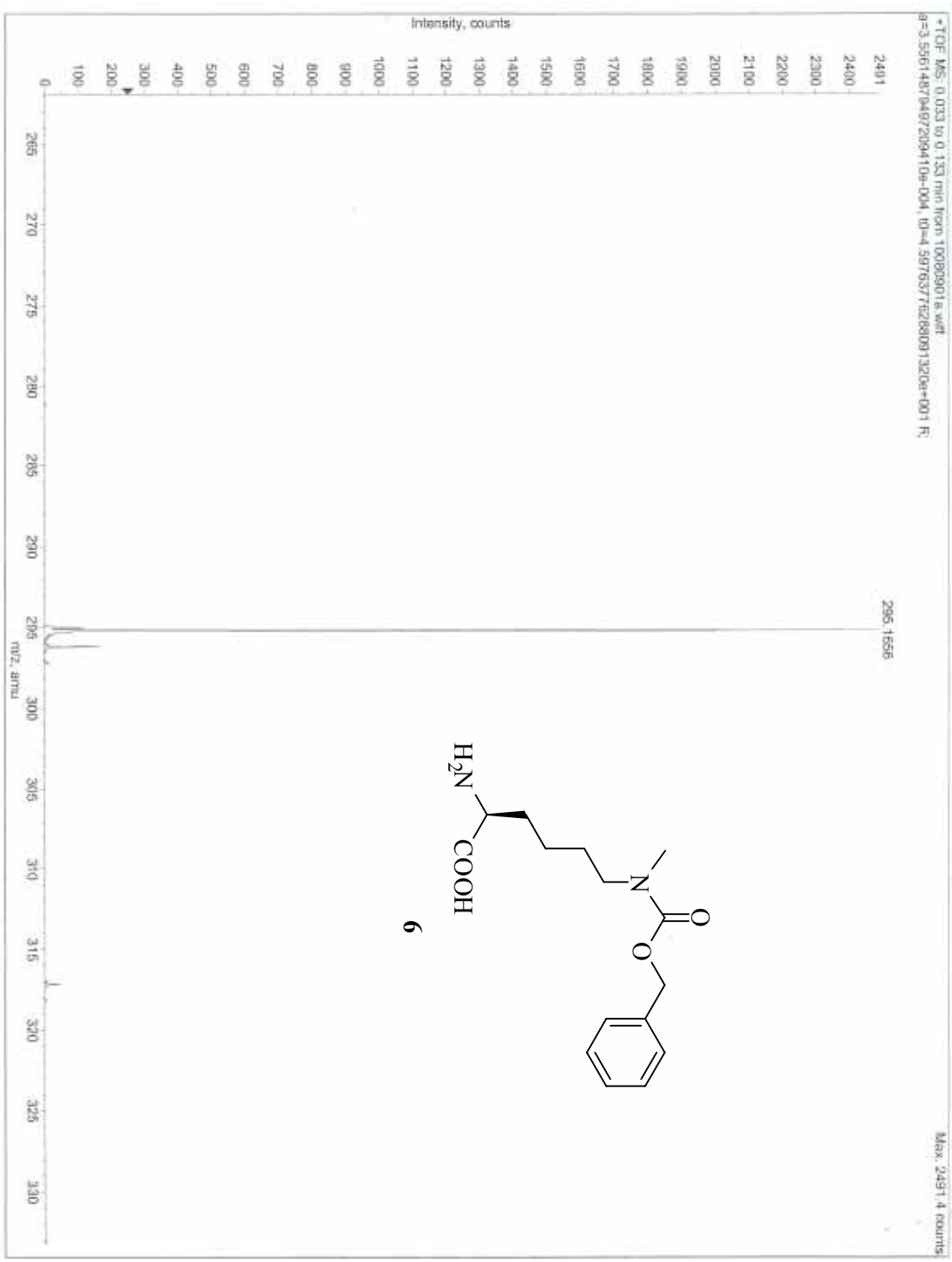

BocHN
COOCH3
18
N
O
Cl
7.270
5.063
5.048
4.297
3.740
3.735
3.455
3.440
3.425
3.397
3.383
3.373
3.358
3.099
3.015
1.846
1.829
1.819
1.694
1.675
1.666
1.650
1.638
1.622
1.607
1.592
1.434
1.408
1.396
1.392
1.381
1.379
1.361
1.347
1.306
0.84
0.87
3.00
1.16
1.31
0.90
1.61
3.33
2.46
1.35
9.23
ppm


BocHN
COOCH3
N
O
Cl
18
173.290
155.540
149.821
149.293
80.192
80.107
77.485
77.230
76.975
53.221
52.851
52.523
51.175
38.573
36.855
36.800
32.672
32.539
28.556
28.471
27.439
26.675
22.449
22.383
180
160
140
120
100
80
60
40
20
ppm

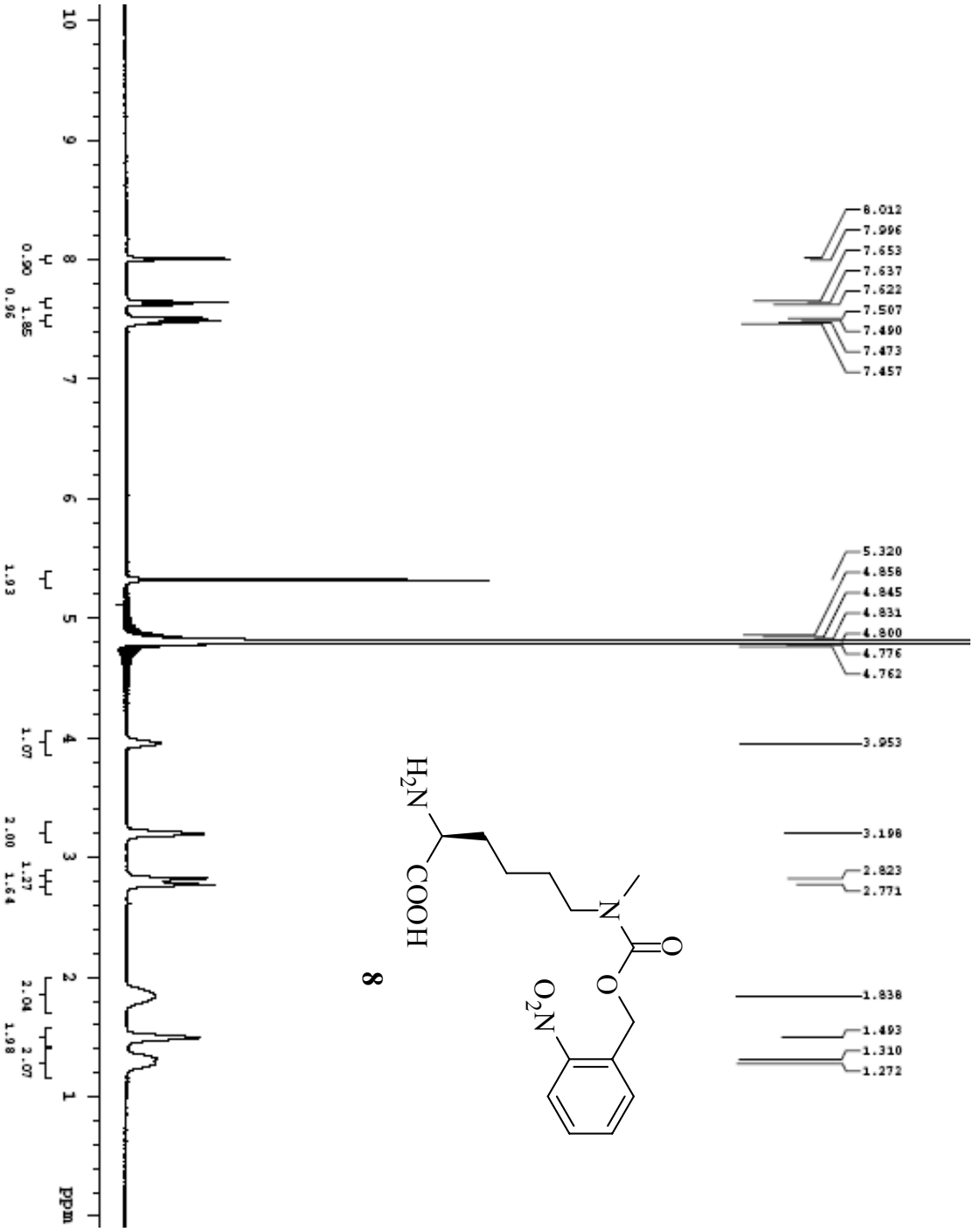

7.464
7.453
7.433
7.417
5.155
3.677
3.336
2.950
2.898
1.831
1.614
1.600
1.585
1.571
1.335
4.30
2.00
1.14
1.24
1.93
1.59
2.12
2.15
2.00
ppm
H_2N
COOH
N
O
O

6

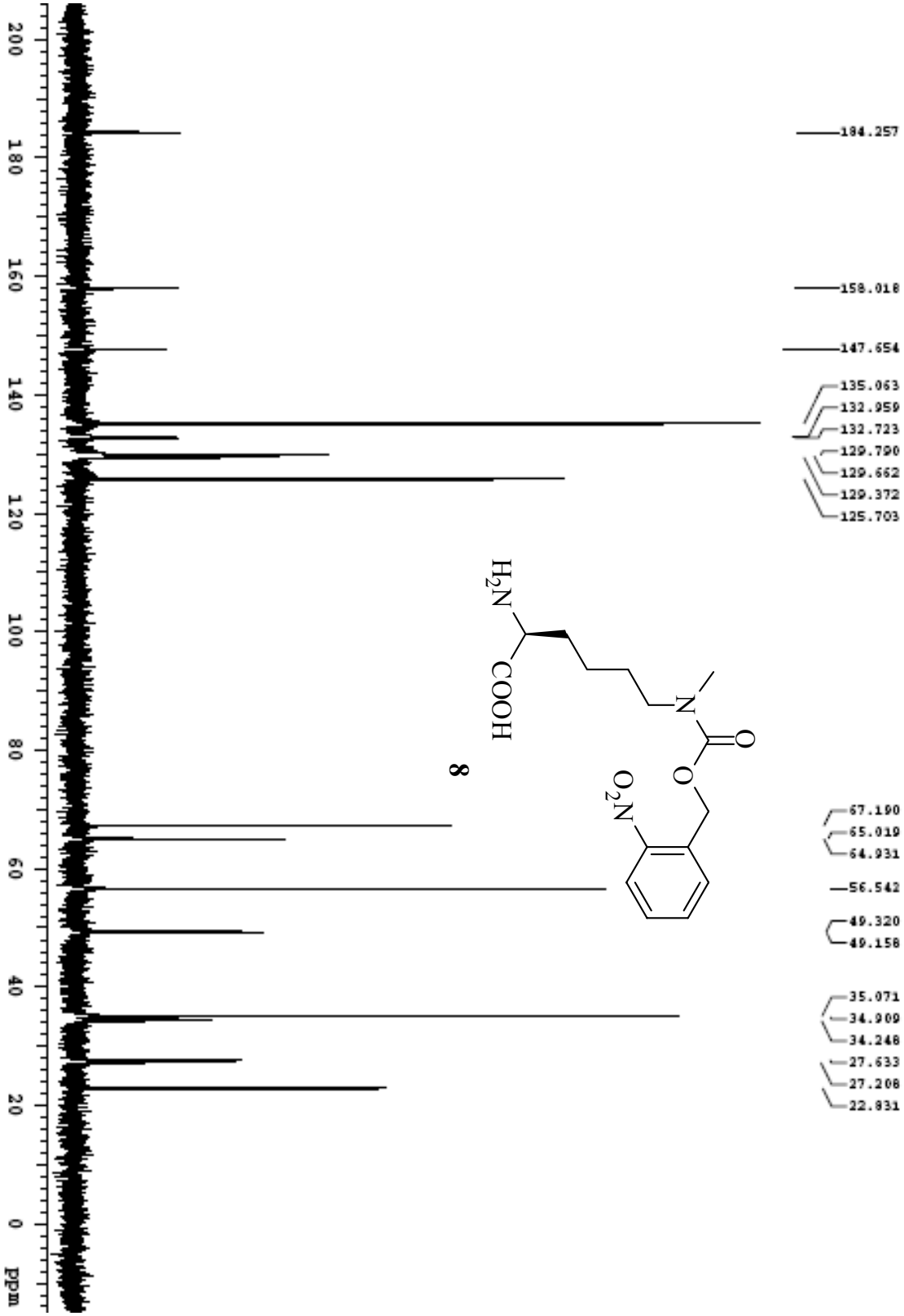

184.136
158.423
137.113
129.365
129.102
128.961
128.482
128.252
68.006
67.190
56.549
49.225
49.016
35.138
34.801
34.241
27.592
27.268
22.845
200
180
160
140
120
100
80
60
40
20
0
ppm
H2N
COOH
N
O
O
6

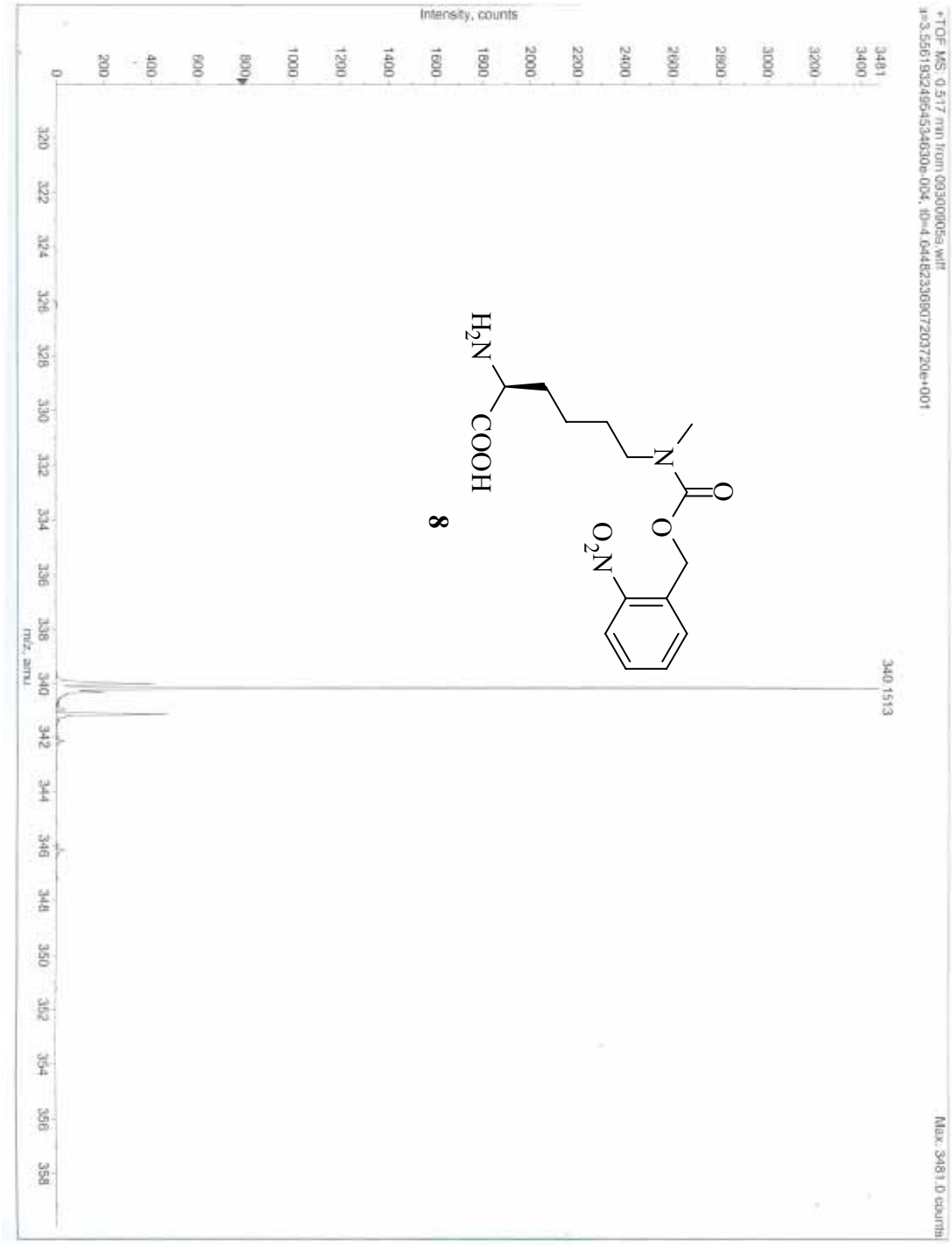

+TOF MS: 0.033 to 0.133 min from 100809013s.wiff
a=3.55614879497209410e-004, t0=4.59763771628809132 0e+001 R;
Intensity, counts
2491
2400
2300
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
295.1556
265
270
275
280
285
290
295
300
305
310
315
320
325
330
m/z, amu
Max. 2491.4 counts

6


8.012
7.996
7.653
7.637
7.622
7.507
7.490
7.473
7.457
5.320
4.858
4.845
4.831
4.800
4.776
4.762
3.953
3.198
2.823
2.771
1.838
1.493
1.310
1.272
0.90
1.85
0.96
1.93
1.07
2.00
1.27
1.64
2.04
2.07
1.98
ppm
H_2N
COOH
N
O
O
O_2N

8


184.257
158.018
147.654
135.063
132.959
132.723
129.790
129.662
129.372
125.703
67.190
65.019
64.931
56.542
49.320
49.158
35.071
34.909
34.248
27.633
27.208
22.831
200
180
160
140
120
100
80
60
40
20
0
ppm
8


+TOF MS: 0.517 min from 093006805a.wiff
a=3.55619324954534630e-004, t0=4.64482336807203720e+001
Intensity, counts
m/z, amu
340.1513
Max. 3481.0 counts
H_2N
COOH
N
O
O
O_2N

8