Supplementary information to:

Selective derivatization of *N*-terminal cysteines using cyclopentenediones

Omar Brun, Jordi Agramunt, Lluis Raich,
Carme Rovira, Enrique Pedroso and Anna Grandas*

1. Materials and methods

2-Methylcyclopentane-1,3-dione, iodomethane, *tert*-butyl acrylate, DIPC, $CuCl_2$, isobutyl chloroformate, TCEP, DTT and 2,2-dimethyl-4-cyclopentene-1,3-dione (**1**) were from Sigma-Aldrich (the latter is no longer available from this supplier, and has been prepared in our laboratory as described below (section 2). HCl in dioxane was from Fluka, and HOBt·H_2O from Iris Biotech. 3-Maleimidopropanoic acid was from Bachem. Fmoc-amino acids (Arg(Pbf), Asn(Trt), Cys(S-*t*Bu), Cys(Trt), Glu(O*s*Bu), Lys(Boc), Ser(*t*Bu), Thr(*t*Bu), Trp(Boc) and Tyr(*t*Bu)) were purchased from Novabiochem. Fmoc-PNA monomers (Bhoc-protected) were from either Link technologies or PolyOrg. CPG supports, nucleoside phosphoramidites (dA^{Pac}, dC^{Ac}, dG^{iPrPac} and dT) and reagents for oligonucleotide synthesis were from Link technologies. *tert*-Butyl (2-aminoethyl)carbamate was synthesized as described in reference 1.

Acid-free DCM was obtained by filtration through basic alumina.

TLC was carried out on silica gel plates 60 F254 from Merck. Samples were lyophilized in either Labconco or FreezeMobile Virtis instruments, and ^1H and ^{13}C NMR spectra were recorded on either a Varian Mercury 400 MHz spectrometer or a Brucker 400 MHz spectrometer. IR spectra were recorded in a Nicolet 6700 FT-IR Thermo Scientific spectrometer.

Reversed-phase HPLC analyses and purifications were performed in Shimadzu systems. HPLC/MS analyses were recorded in an Alliance Waters 2690 separation module with a Waters micromass ZQ4000 MS detector, while a LC/MSD-TOF spectrometer from Agilent Technologies was used for HPLC/HRMS.

MALDI-TOF mass spectra were recorded in a 4800 *Plus* ABSciex instrument. Unless otherwise indicated, reflector was used. ESI (low and high resolution) mass spectra were obtained using an LC/MSD-TOF spectrometer from Agilent Technologies.

1.1 Peptide synthesis, purification and quantification

Peptides were assembled on a NovaSyn TGR resin (MBHA-PS resin derivatized with PEG chains and a Rink amide linker, Novabiochem), which was washed with DCM (3 ×), DMF (3 ×), MeOH (3 ×), DCM (3 ×). The resin was always allowed to swell in DCM for 2 min before Fmoc removal (all synthesis cycles), which was effected by reaction with 20 % piperidine/DMF (1 × 3 min + 1 × 10 min), followed by washing with DMF (5 ×) and DCM (5 ×). Incorporation of the first amino acid onto the resin and peptide assembly were accomplished by using 3 equiv of both Fmoc-amino acid, HOB·H_2O and DIPC, all dissolved in the minimal amount of DCM and a few drops of NMP, for 90 min at rt. Coupling was followed by washing with DCM, DMF and MeOH (3 × each). In case the coupling was not complete, as assessed by the Kaiser test,[2] it was repeated using 2 equiv of the reagents. If necessary, unreacted amines were capped by reaction with Ac_2O/2,6-lutidine/DMF (5:6:89 mixture, 2 × 5 min) for 15 min at rt, which was followed by washing with DMF (5 ×) and DCM (5 ×).

Cysteine-containing peptides were cleaved from the resin and deprotected by using a 94:2.5:1:2.5 TFA/H_2O/TIS/EDT mixture. In the case of peptides not incorporating cysteine residues, cleavage and deprotection were performed by reaction with a 95:2.5:2.5 TFA/H_2O/TIS mixture. Filtrate and washings (TFA) were collected and concentrated by blowing N_2 over the mixture, and diethyl ether was added to the resulting oil or semisolid. After precipitating the peptide, the suspension was centrifuged (7800 rpm, 10 min, 0 ºC) and the diethyl ether supernatant discarded. This procedure was performed three times. The precipitated peptide was dissolved in water and lyophilized before analysis and purification by HPLC.

Pure peptides were quantified by UV/Vis spectroscopy on the basis of their absorbance at 280 nm (Tyr, ε_{280}=1490; Trp, ε_{280}=5500).

1.2 PNA synthesis, purification and quantification

PNAs were assembled on the TentaGel R RAM resin (from Rapp Polymere) derivatized with a Rink amide linker. Prior to chain elongation the resin was washed with DCM (3 ×), DMF (3 ×), MeOH (3 ×), and DCM (3 ×), treated with 20% piperidine/DMF (1 × 3 min + 1 × 15 min), and washed with DCM (3 ×), DMF (3 ×), and DCM (3 ×). The resin was always allowed to swell in DCM for 2 min before Fmoc removal, which was carried out by reaction with 20% piperidine/DMF (1 × 3 min + 1 × 10 min), followed by washing with DCM (3 ×), DMF (3 ×), and DCM (3 ×). Incorporation of the PNA monomers was effected by reaction with 4 equiv of both the Fmoc-protected monomer and COMU and 8 equiv of DIPEA, all dissolved in the minimal amount of NMP. This mixture was preactivated for 1 min before being poured onto the resin and allowed to react for 90 min at rt. Coupling was followed by washing with DCM, DMF, and MeOH (3 × each). In case the coupling was not complete, as assessed by the Kaiser test, it was repeated as before. A capping step was carried out after incorporation of each monomer, as described above, to ensure that no trace of free amines was left. Cleavage and deprotection and purification were performed as for peptides. PNAs were quantified by UV/Vis spectroscopy on the basis of the absorbance of the nucleobases at 260 nm (A, ε_{260}=13700; C, ε_{260}=6600; G, ε_{260}=11700; T, ε_{260}=8600).

In the case of mixed sequences (PNA chains typically incorporate a lysine residue to increase solubility in water, and sometimes a cysteine residue for conjugation), the coupling conditions described for peptide synthesis were used to incorporate amino acids and those described to assemble PNA chains for the coupling of PNA building blocks.

1.3 Oligonucleotide synthesis, purification and quantification

Oligonucleotide chains were assembled in a 3400 ABI automatic synthesizer at the 1 μmol scale. Chain elongation and incorporation of cysteine-containing phosphoramidites were carried out using the phosphite triester approach and standard synthesis cycles, with the only exception of the Cys-containing phosphoramidite (double coupling, 10 min/each). After chain elongation was completed, the Fmoc group was removed by treatment with 20% piperidine in DMF (1 × 10 min), and the resin thoroughly washed with ACN. Treatment with concd. aq. ammonia at rt removed protecting groups from the oligonucleotide (2 h in case of homo-dT oligomers; 3 h when all the bases were present). After filtration and washing, ammonia was evaporated under reduced pressure, and the sample lyophilized.

H-Cys-oligonucleotides were obtained after cysteine deprotection as described below (section 9.4).

Oligonucleotides were quantified by UV spectroscopy (260 nm; A, ε_{260}=15400; C, ε_{260}=7400; G, ε_{260}=11500; T, ε_{260}=8700), and characterized by MALDI-TOF MS.

1.4 Conditions for HPLC analysis and purification

Reversed-phase HPLC analyses and purifications of **polyamides** (peptides and PNAs)**:**

- **Analysis conditions**: Jupiter Proteo column (4 µm, 90 Å, 250 × 4.6 mm) from Phenomenex. Linear gradients of 30 min were always used. Solvent A: 0.045% TFA in water, solvent B: 0.036% TFA in ACN, flow: 1 mL/min, detection wavelength: 200 to 500 nm (diode array detector). Unless otherwise stated, all injections were performed at rt.
- **Purification conditions**: Jupiter Proteo column (10 µm, 90 Å, 250 × 10.0 mm) from Phenomenex. Linear gradients of 30 min were always used. Solvent A: 0.1% TFA in water, solvent B: 0.1% TFA in ACN, flow: 3 mL/min, detection wavelength: 220 and 260 nm. Unless otherwise stated, all injections were performed at rt.

Reversed-phase HPLC analyses and purifications of **oligonucleotides:**

- **Analysis conditions**: Jupiter column (10 µm, 90 Å, 250 × 4.6 mm) from Phenomenex. Linear gradients of 30 min were always used. Solvent A: 0.1 M triethylammonium acetate, solvent B: ACN, flow: 1 mL/min, detection wavelength: 254 nm. Unless otherwise stated, all injections were performed at rt.
- **Purification conditions**: Jupiter column (10 µm, 90 Å, 250 × 10.0 mm) from Phenomenex. Linear gradients of 30 min were always used. Solvent A: 0.1 M triethylammonium acetate, solvent B: ACN, flow: 3 mL/min, detection wavelength: 254 nm. Unless otherwise stated, all injections were performed at rt.

Conditions for HPLC/MS analysis (polyamides)**:**

- Jupiter Proteo column (4 µm, 90 Å, 250 × 4.6 mm) from Phenomenex. Linear gradients of 30 min were always used. Solvent A: 0.1% formic acid in water, solvent B: 0.1% formic acid in ACN, flow: 1 mL/min, detection wavelength: 210 to 600 nm (diode array detector). Unless otherwise stated, all injections were performed at rt.

1.5 Characterization of underivatized peptides and PNAs used in this work

MALDI-TOF mass spectra: All samples were prepared by mixing 1 or 2 µL of sample solution with 1 µL of matrix solution, the latter being chosen depending on the nature of the analyte. The matrix/sample mixture was prepared directly on the plate, and allowed to air-dry. Oligonucleotides were analyzed using the THAP/CA mixture, while DHB was chosen for the analysis of either peptides and PNAs.

The THAP matrix was prepared by dissolving 10 mg of THAP in 1 mL of a 1:1 (v/v) H_2O/ACN mixture. Ammonium citrate solution: 10 mg/mL H_2O. The DHB matrix was prepared by dissolving 10 mg of DHB in 1 mL of a 1:1 (v/v) H_2O/ACN mixture containing 0.1% of TFA.

Amino acids, nucleosides and PNA monomers are designed with the one-letter code. "H-C..." refers to cysteine residues with a free amine in all sequences.[3] Letters A, C, G and T mean alanine, cysteine, glycine and threonine in peptide (and PNA) sequences (where letters different from A, C, G and T appear), but mean 2'-deoxyadenosine, 2'-deoxycytidine, 2'-deoxyguanosine and 2'-deoxythymidine in oligonucleotide sequences (where the only letters appearing are these four). cHex is the abbreviation for the linker between cysteine and the phosphate linkage in the case of Cys-oligonucleotides. Lower case letters (a, c, g and t) refer to PNA monomers.

Product	**HPLC analysis**	**MALDI-TOF MS analysis**
H-WGRGC-NH_2 (**5**)	0-50% B, t_R = 16.5 min	*m/z* 577.5 $[M+H]^+$, M calcd. for $C_{24}H_{36}N_{10}O_5S$ 576.3
H-KYAYCG-NH_2 (**6**)	10-40% B, t_R = 10.5 min	*m/z* 703.5 $[M+H]^+$, M calcd. for $C_{32}H_{46}N_8O_8S$ 702.3
H-CYG-NH_2 (**7**)	0-40% B, t_R = 12.8 min	*m/z* 341.2 $[M+H]^+$, M calcd. for $C_{14}H_{20}N_4O_4S$ 340.1
H-CK-catagctgtttc-NH_2 (**25d**)	5-35% B, 60 ºC, t_R = 12.4 min	*m/z* 3464.9 $[M+H]^+$, M calcd. for $C_{138}H_{181}N_{67}O_{41}S$ 3464,4
H-CSYAKYG-NH_2 (**25e**)	0-50% B, t_R = 16.0 min	*m/z* 790.5 $[M+H]^+$, M calcd. for $C_{35}H_{51}N_9O_{10}S$ 789.9
H-CSYACKYG-NH_2 (**28**)	0-50% B, t_R = 17.4 min	*m/z* 893.5 $[M+H]^+$, M calcd. for $C_{38}H_{56}N_{10}O_{11}S_2$ 892.4
H-CWGRGC-NH_2 (**30**)	10-40% B, t_R = 13.7 min	*m/z* 680.0 $[M+H]^+$, M calcd. for $C_{27}H_{41}N_{11}O_6S_2$ 679.3

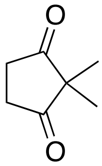
Table S1. Characterization data of polyamides prepared following standard solid-phase procedures (and not described in further sections).

The synthesis and characterization of Mal-gtttc-K-NH_2 **43** is described in reference 4.

2. Synthesis of 2,2-disubstituted cyclopent-4-en-1,3-diones 1, 2 and 3

2,2-Dimethylcyclopentanedione was synthesized following the procedure of Agosta and Smith[5] with minor modifications (see below). Oxidation to cyclopentenedione was carried out essentially as described in ref 6. Purification of **1** by column chromatography must be carried out immediately after work-up, using volatile solvents and taking care during their elimination (otherwise a large amount of **1** is lost upon evaporation under reduced pressure). Curiously, compound **1** was found to undergo degradation upon storage in the freezer, whilst it kept perfectly stable at rt. A solution in 1:3 (v/v) ACN/H_2O was kept in the refrigerator for months, with no sign of product degradation.

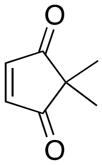
· 2,2-Dimethylcyclopentane-1,3-dione (precursor of 1)



Potassium hydroxide (1.00 g, 17.82 mmol) and MeI (2.53 g, 17.82 mmol) were added to a solution of 2-methylcyclopentane-1,3-dione **I** (see Scheme 1 below, 2.00 g, 17.82 mmol) in dioxane (30 mL) and water (10 mL), and heated to reflux (105 ºC) overnight. The reaction mixture was cooled and extracted with ether (3 × 30 mL), and the ether solution was evaporated under reduced pressure. 10% aq. HCl (50 mL) was added to the resulting crude and the mixture heated to the boiling point. After 5 min boiling, an excess of 10% aq. $NaHCO_3$ was added to the mixture, which was subsequently cooled down. Finally, the mixture was extracted with DCM (3 × 20 mL). The organic layers were dried over $MgSO_4$, filtered and evaporated under reduced pressure to afford a white solid (801 mg, 35% yield). No further purification was required.

mp: 52-53 ºC. TLC (hexanes/AcOEt 4:1): $R_f = 0.53$, dyeing reagent: 100 mg 2,4-dinitrophenyl-hydrazine dissolved in 100 mL EtOH + 1 mL conc. aq. HCl (yellow spot). IR (ATR, solid): ν 2972, 2923, 2865, 1694, 1689, 1454, 1285 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$) δ: 2.80 (s, 4H), 1.15 (s, 6H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ: 216.5, 57.8, 34.7, 20.4 ppm. HRMS (ESI, positive mode): *m/z* 127.0758 $[M+H]^+$, M calcd. for $C_7H_{10}O_2$ 126.0681.

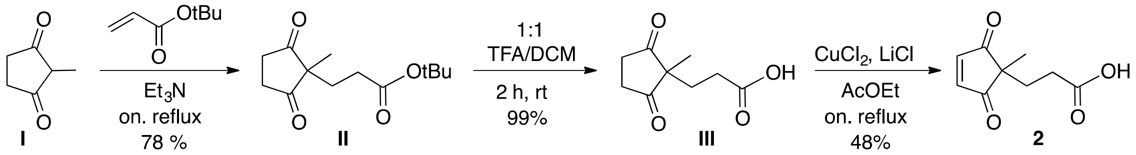
· 2,2-Dimethylcyclopent-4-ene-1,3-dione (1):



Copper(II) chloride (926.73 mg, 6.97 mmol), lithium chloride (292.82 mg, 6.97 mmol) and 2,2-dimethylcyclopentane-1,3-dione (400 mg, 3.17 mmol) were added to EtOAc (25 mL), and the mixture was refluxed for 17 h at 85 ºC. Once the reaction was complete (as assessed by TLC analysis), H_2O (10 mL) was poured onto the round-bottomed flask, and its content fully transferred into a separatory funnel. After separating the two phases, the organic layer was further washed with H_2O until the aqueous phase was colorless. The organic layer was dried over $MgSO_4$ and filtered before solvent was removed under reduced pressure. Purification using silica gel flash column chromatography was immediately carried out, eluting with a 99:1 up to 95:5 (v/v) hexanes/Et_2O mixture. **1** Was obtained as a yellow oil (85.5 mg, 22%).

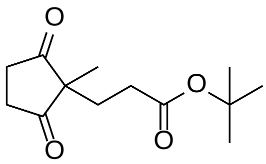
TLC (hexanes/AcOEt 4:1): $R_f = 0.60$. IR (ATR, oil): ν 1695, 1290, 1139, 1045, 854 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$) δ: 7.20 (s, 2H), 1.16 (s, 6H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ: 207.7, 147.1, 46.4, 19.6 ppm. HRMS (ESI): the product does not ionize.

Synthesis of CPD 2.



Scheme S1. Synthesis of CPD **2**. (on. = overnight)

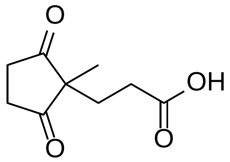
· ***tert*****-Butyl 3-(1-methyl-2,5-dioxocyclopentyl)propanoate** (**II**):



2-Methylcyclopentane-1,3-dione (**I**, 1.22 g, 10.9 mmol) and *tert*-butyl acrylate (5.9 mL, 26.8 mmol, 2.5 equiv) were dissolved in 20.0 mL of triethylamine. The mixture was refluxed overnight with fast stirring. Triethylamine and excess *tert*-butyl acrylate were removed in vacuo to render **II** (2.50 g, 78%) as a dark brown solid that was no further purified.

mp: 63-64 ^{o}C. TLC (DCM/AcOH 98:2): R_f = 0.30. IR (ATR, solid): ν 2980, 1708, 1365, 1294, 1143, 840 cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ: 2.87 – 2.70 (m, 4H), 2.20 (t, *J* = 7.5 Hz, 2H), 1.92 (t, *J* = 7.6 Hz, 2H), 1.40 (s, 9H), 1.12 (s, 3H) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ: 215.9, 172.3, 81.1, 55.4, 35.0, 30.0, 29.3, 28.2, 19.8 ppm. HRMS (ESI, positive mode): *m/z* 241.1446 [M+H]$^+$, M calcd. for $C_{13}H_{20}O_4$ 240.1362.

· **3-(1-Methyl-2,5-dioxocyclopentyl)propanoic acid** (**III**):

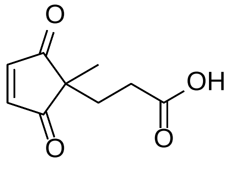


tert-Butyl 3-(1-methyl-2,5-dioxocyclopentyl)propanoate (**II**) (1.0 g, 4.16 mmol) was dissolved in 26.0 mL of a 1:1 TFA/DCM mixture and stirred for 2 h at 25 ºC. Toluene (10-15 mL) was added to the reaction mixture, which was concentrated and dried in vacuo to give **III** as a brown solid (760 mg, 99%). No purification step was necessary.

mp: 121-122 ^{o}C. TLC (DCM/AcOEt/AcOH 78:20:2): R_f = 0.35. IR (ATR, solid): ν 3045, 1697, 1404, 1192 cm^{-1}. ^{1}H NMR (400 MHz, CD$_3$OD) δ: 2.78 (s, 4H), 2.26 (t, *J* = 8.0 Hz, 2H), 1.89 (t, *J* = 8.0 Hz, 2H), 1.10 (s, 3H) ppm. ^{13}C NMR (101 MHz, CD$_3$OD) δ: 217.9, 176.4, 56.5, 35.6, 30.2, 29.5, 19.6 ppm. HRMS (ESI, negative mode): *m/z* 183.0661 [M-H]$^-$, M calcd. for $C_9H_{12}O_4$ 184.0736.

· **3-(1-Methyl-2,5-dioxocyclopent-3-en-1-yl)propanoic acid** (**2**):

3-(1-Methyl-2,5-dioxocyclopentyl)propanoic acid (**III**) (500 mg, 2.72 mmol) was

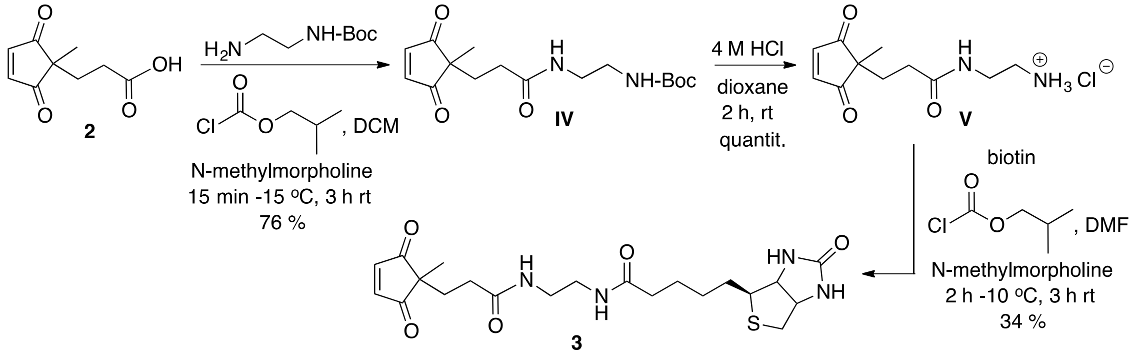


dissolved in 20.0 mL of ethyl acetate, and CuCl$_2$ (803 mg, 5.98 mmol, 2.2 equiv) and LiCl (253 mg, 5.98 mmol, 2.2 equiv) were added to the mixture. After refluxing overnight, *ca.* 20 mL of ethyl acetate were added to the crude, and the resulting mixture was subsequently washed with 5% aq HCl (*ca.* 40 mL × 3). The organic fraction was dried over anhydrous MgSO$_4$, and the solvent removed in vacuo. The resulting crude was purified by silica gel column chromatography (98:2 DCM/AcOH, isocratic), and a yellow powder was obtained (240 mg, 48%).

Note: sometimes it is advisable to reextract the aqueous phase with ethyl acetate (once or twice) to recover small amounts of **2** that may have remained at the aqueous phase.

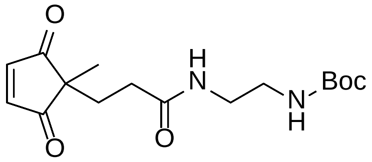
mp: 82-83 °C. TLC (DCM/AcOEt/AcOH 78:20:2): $R_f = 0.55$. IR (ATR, solid): ν 3060, 2916, 1693, 1303 cm^{-1}. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.25 (s, 2H), 2.33 – 2.24 (m, 2H), 2.01 – 1.93 (m, 2H), 1.18 (s, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ: 206.9, 176.9, 148.1, 49.2, 28.8, 28.5, 18.6 ppm. HRMS (ESI, negative mode): *m/z* 181.0500 [M-H]$^-$, M calcd. for $C_9H_{10}O_4$ 182.0579.

Synthesis of CPD 3.



Scheme S2. Synthesis of CPD **3**.

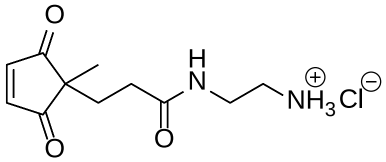
· ***tert*-Butyl (2-(3-(1-methyl-2,5-dioxocyclopent-3-en-1yl)propanamido)ethyl)carbamate (IV)**:



A solution of **2** (50 mg, 0.274 mmol) in DCM (1.0 mL) was cooled to -15 ºC. Subsequently, NMM (30 µL, 0.274 mmol, 1.0 equiv) and isobutyl chloroformate (35 µL, 0.274 mmol, 1.0 equiv) were added. After 10 min with heavy stirring, *tert*-butyl (2-aminoethyl)carbamate (48 mg, 0.301 mmol, 1.1 equiv) was added. The suspension was kept at -15 ºC for another 15 min and afterwards was reacted for 3 h at rt. After removing the solvent under reduced pressure, the crude was dissolved in EtOAc (20 mL) and extracted with water (2 × 10 mL), 5% HCl (2 × 10 mL) and 5% NaHCO$_3$ (2 × 10 mL). The organic layer was dried over MgSO$_4$, filtered and evaporated under reduced pressure to afford the title compound (67 mg, 76% yield) as a yellow solid.

mp: 102-103 ºC. TLC (DCM/AcOEt/AcOH 78:20:2): R$_f$ = 0.17. IR (ATR, solid): ν 3368, 3322, 2962, 1697, 1689, 1237 cm^{-1}. ^{1}H NMR (CDCl$_3$, 400 MHz): δ 7.24 (s, 2H,), 6.31 (bs, 1H), 5.05 (bs, 1H,), 3.32-3.23 (m, 4H), 2.13- 2.09 (m, 2H), 1.99- 1.96 (m, 2H), 1.44 (s, 9H), 1.16 (s, 3H) ppm. ^{13}C NMR (CDCl$_3$, 101 MHz): δ 206.9, 171.8, 156.9, 147.8, 79.8, 49.3, 40.8, 40.1, 31.1, 29.5, 28.4, 18.4 ppm. HRMS (ESI, positive mode): *m/z* 325.1767 [M+H]$^{+}$, 225.1232 [M–Boc+H]$^{+}$, 247.1053 [M–Boc+Na]$^{+}$, 347.1577 [M+Na]$^{+}$, 649.3430 [2M+H]$^{+}$, 671.3249 [2M+Na]$^{+}$; M calcd. for $C_{16}H_{24}N_2O_5$ 324.1685.

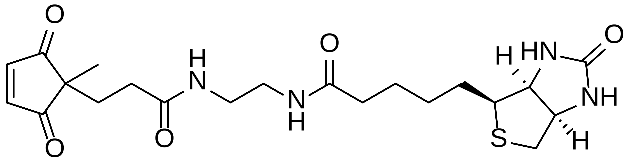
· ***N*-(2-Aminoethyl)-3-(1-methyl-2,5-dioxocyclopent-3-en-1-yl)propanamide (V)**:



4 M Hydrochloric acid in dioxane (0.5 mL, 7.39 mmol) was added to a solution of **IV** (60 mg, 0.184 mmol) in DCM (0.5 mL) and stirred at rt for 2 h. The solvent was removed under reduced pressure to afford the title compound (42 mg, quantitative yield) as a red solid.

mp: 134-136 ºC. TLC (DCM/AcOEt 1:1): R$_f$ = 0.54. IR (ATR, solid): ν 3279, 2922, 1693, 1643, 1444 cm^{-1}. ^{1}H NMR (CD$_3$OD, 400 MHz): δ 7.34 (s, 2H), 3.38 (t, *J* = 8.0 Hz, 2H), 3.02 (t, *J* = 8.0 Hz, 2H), 2.15-2.11 (m, 2H), 1.93-1.89 (m, 2H), 1.12 (s, 3H) ppm. ^{13}C NMR (CD_3OD, 101 MHz): δ 207.3, 174.4, 148.1, 49.0, 39.5, 36.8, 30.2, 29.1, 17.5 ppm. HRMS (ESI, positive mode): *m/z* 225.1234 $[M+H]^+$, 247.1053 $[M+Na]^+$, 471.2220 $[2M+Na]^+$; M calcd for $C_{11}H_{16}N_2O_3$ 224.1161.

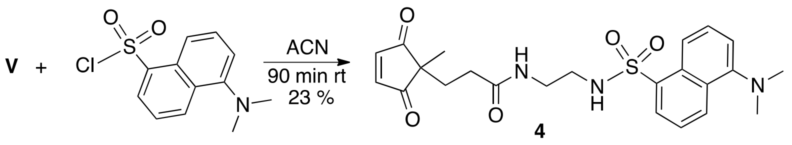
· ***N*-(2-(3-(1-Methyl-2,5-dioxocyclopent-3-en-1-yl)propanamido)ethyl)-5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-d]imidazol-4-yl)pentanamide** (**3**):



D-Biotin (50 mg, 0.204 mmol) was dissolved in anhydrous DMF (3.0 mL) under a nitrogen atmosphere and stirred for 30 min at -10 ºC. Once dissolved, NMM (56 µL, 0.510 mmol, 2.5 equiv) and isobutyl chloroformate (26.5 µL, 0.204 mmol, 1.0 equiv) were added, and the mixture was reacted under heavy stirring at -10 ºC for 1 h. Thereafter, a solution of **V** (46 mg, 0.204 mmol, 1.0 equiv) in anhydrous DMF (2.0 mL) was added, and the mixture was stirred for 1 additional hour at -10 ºC and at rt for 3 h. Afterwards, the solvent was evaporated in vacuo and the crude was purified by reversed-phase chromatography, using a 10 mL-syringe filled with 2.0 mL of C18-derivatized silica (Vydac). Elution was carried out with a gradient of 8.0 mL of H_2O, 8.0 mL of 3:1 (v/v) H_2O/MeOH and 8.0 mL of 1:1 (v/v) H_2O:MeOH. The title compound (28 mg, 34%) was obtained as a yellow solid.

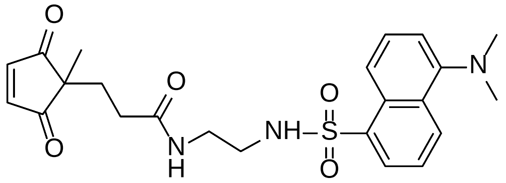
mp: 126-127 ºC. TLC (DCM/MeOH/AcOH 88:10:2): R_f = 0.23. IR (ATR, solid): ν 3296, 3199, 2976, 2923, 2415, 1703, 1628, 1138 cm^{-1}. ^{1}H NMR (CD_3OD, 400 MHz): δ 7.38 (s, 2H), 4.53 (dd, *J* = 4.4 Hz, *J* = 7.6 Hz, 1H), 4.35 (dd, *J* = 4.8 Hz, *J* = 8.0 Hz; 1H), 3.26 (m, 4H); 3.26 (m, 1H), 2.96 (dd, *J* = 4.8 Hz, *J* = 12.4 Hz, 1H), 2.74 (d, *J* = 12.8 Hz, 1H), 2.22 (t, *J* = 6.4 Hz, 2H), 2.09 (m, 2H), 1.91 (m, 2H), 1.81-1.57 (m, 4H), 1.45 (q, *J* = 7.6 Hz, 2H), 1.15 (s, 3H) ppm. ^{13}C NMR (CD_3OD, 101 MHz): δ 207.3, 174.9, 173.4, 164.7, 148.1, 61.9, 60.2, 55.5, 49.0, 39.7, 38.6, 38.5, 35.4, 30.5, 29.6, 28.3, 28.1, 25.4, 17.5 ppm. HRMS (ESI, positive mode): *m/z* 451.2015 $[M+H]^+$, 473.1830 $[M+Na]^+$; 923.3712 $[2M+Na]^+$ M calcd. for $C_{21}H_{30}N_4O_5S$ 450.1937.

Synthesis of CPD 4.



Scheme S3. Synthesis of CPD **4**.

- ***N*-(2-(5-(dimethylamino)naphthalene-1-sulfonamido)ethyl)-3-(1-methyl-2,5-dioxocyclopent-3-en-1-yl)propanamide (4):**



Dansyl chloride (54.67 mg, 0.202 mmol) was added to a solution of **V** (50 mg, 0.223 mmol) in anhydrous ACN (1 mL). Subsequently, TEA (0.5 mL, 3.59 mmol) was added and the mixture left for 3 h under heavy stirring at rt. Afterwards, the solvent was evaporated under reduced pressure, and the resulting residue dissolved in DCM (15 mL) and extracted with water (3 × 15 mL). The organic layer was extracted with 5% HCl (3 × 15 mL). The aqueous layer was basified to pH 6-7 with 5% $NaHCO_3$ and extracted with DCM (3 × 15 mL). The organic layer was dried over $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The title compound (21.10 mg, 23% yield) was obtained as a yellow solid.

mp 128-129 ºC. TLC (DCM/EtOAc 95:5): R_f = 0.35. IR (ATR, Solid): ν 3296, 3074, 2918, 2874, 1698, 1641, 1309, 1140, 788, 727, 622 cm^{-1}. ^1H NMR ($CDCl_3$, 400 MHz): δ 8.54 (d, J = 8.4 Hz, 1H); 8.28 (d, J = 8.4 Hz, 1H); 8.22 (d, J = 7.2 Hz); 7.58-7.50 (m, 2H); 7.20 (s, 2H), 7.19 (d, J = 6.8 Hz); 6.04 (t, J = 2.4 Hz, 1H); 5.82 (t, J = 5.2 Hz, 1H); 3.24-3.20 (m, 2H); 3.02-2.98 (m, 2H); 2.89 (s, 6H); 2.03-1.98 (m, 2H); 1.92-1.89 (m, 2H); 1.13 (s, 3H) ppm. ^{13}C NMR ($CDCl_3$, 101 MHz): δ 207.0, 172.20, 152.1, 147.9, 134.4, 130.6, 129.9, 129.6, 129.4, 128.5, 123.2, 118.6, 115.3, 49.3, 45.4, 42.8, 39.5, 30.9, 29.3, 18.4 ppm. ESI HRMS (positive mode): *m/z* 458.1730 $[M+H]^+$, 480.1571 $[M+Na]^+$, 937.3224 $[2M+Na]^+$; M calcd. for $C_{23}H_{28}N_3O_5S$ 457.1671.

3. Experiments with peptides containing Cys at the *C*-terminal or internal positions, and stability of the adducts

· 2,2-Dimethyl-4-cyclopentene-1,3-dione (1) + H-WGRGC-NH$_2$ (5), and 1 + H-KYAYCG-NH$_2$ (6):

Both peptides (50 nmol) were independently reacted with 5 equiv of 2,2-dimethyl-4-cyclopentene-1,3-dione (**1**) in water (0.1 mM peptide concentration) at 37 ºC. The reaction progress was followed by HPLC (Figure S1), by analyzing aliquots removed at at 15 min, 1 h and 3 h.

Characterization of MTA adducts **8** (from the **1** + **5** reaction):
HPLC: analysis conditions, 10-40% B, t_R = 17.0 and 17.4 min
MALDI-TOF (positive mode, DHB): *m/z* (both peaks) 701.5 $[M+H]^+$, M calcd. for $C_{31}H_{44}N_{10}O_7S$ (**1** + **5** Michael-type adduct) 700.3

Characterization of the MTA adducts from the **1** + **6** reaction:
HPLC: analysis conditions, 10-40% B, t_R = 16.5 min
MALDI-TOF (positive mode, DHB): *m/z* 827.6 $[M+H]^+$, M calcd. for $C_{39}H_{54}N_8O_{10}S$ (**1** + **6** Michael-type adduct) 826.4.

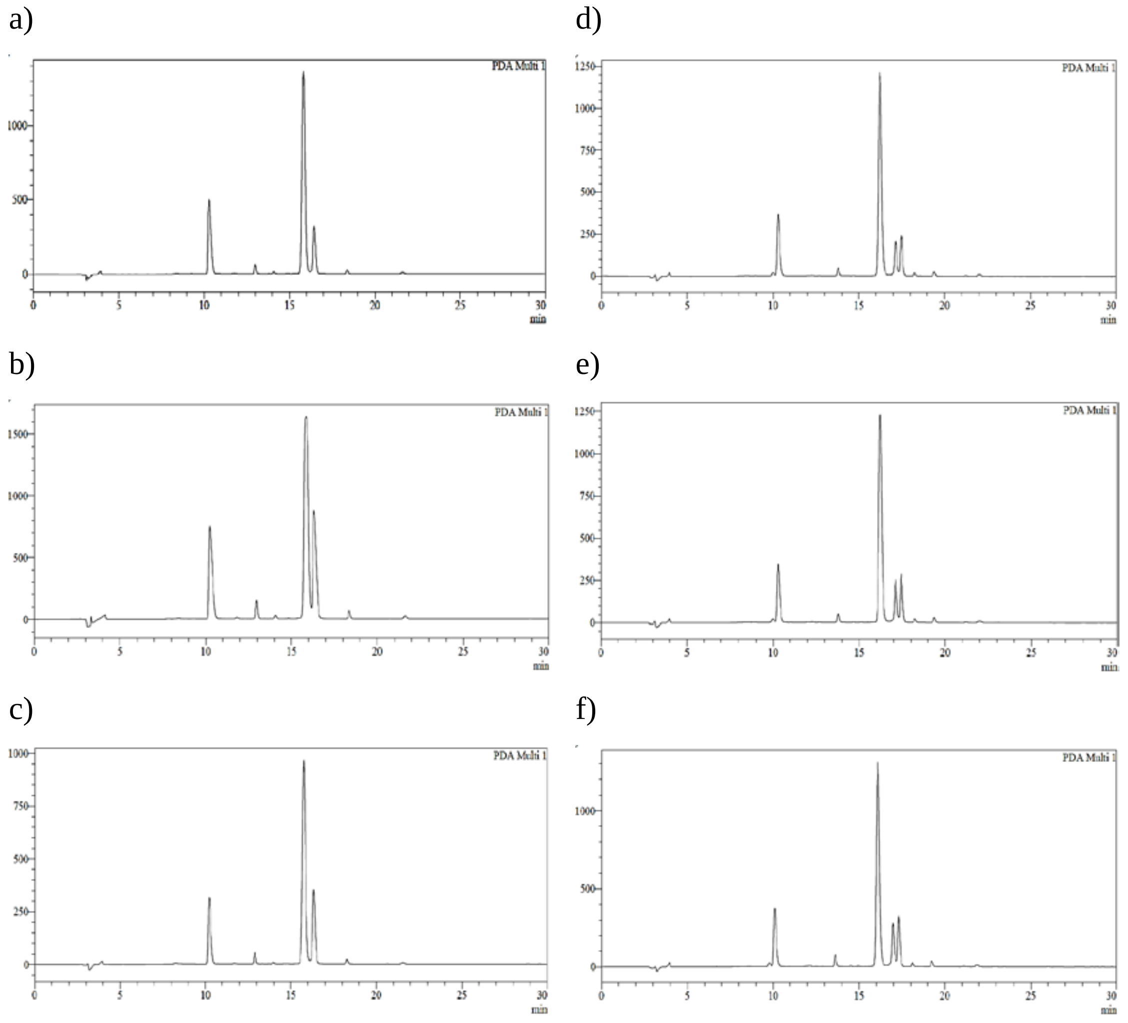


Figure S1: HPLC traces (220 nm) of the **1** + **6** reaction mixture after: a) 15 min, b) 1 h, and c) 3 h; and of the **1** + **5** reaction mixture after: d) 15 min, e) 1 h, and f) 3 h. Peaks at 10.3 and 15.5 min are, in both experiments, the starting reagents (peptides and **1**, respectively).

· 3-(1-Methyl-2,5-dioxocyclopent-3-en-1-yl)propanoic acid (2) + H-KYAYCG-NH$_2$ (6):

H-KYAYCG-NH$_2$ (**6**, 20 nmol) was reacted with 5 equiv of **2** in 0.1 mM phosphate buffer (pH=7.8, 0.1 mM peptide concentration) at rt for 2 h, after which time the mixture was stirred for an additional 4.5 h at 45 ºC. The reaction progress was followed by HPLC, by analyzing aliquots removed at 1 h and at 6.5 h (Figure S2). The new product peaks were collected, lyophilized and analyzed by MALDI-TOF MS.

Product characterization (MTA **9**):
HPLC: analysis conditions, 0-40 % B, t_R = 21.5 min and 22.2 min.
MALDI-TOF (positive mode, DHB): *m/z* (both peaks) 885.7 $[M+H]^+$, M calcd. for $C_{41}H_{56}N_8O_{12}S$ (**2**+**4** Michael-type adduct) 884.4.

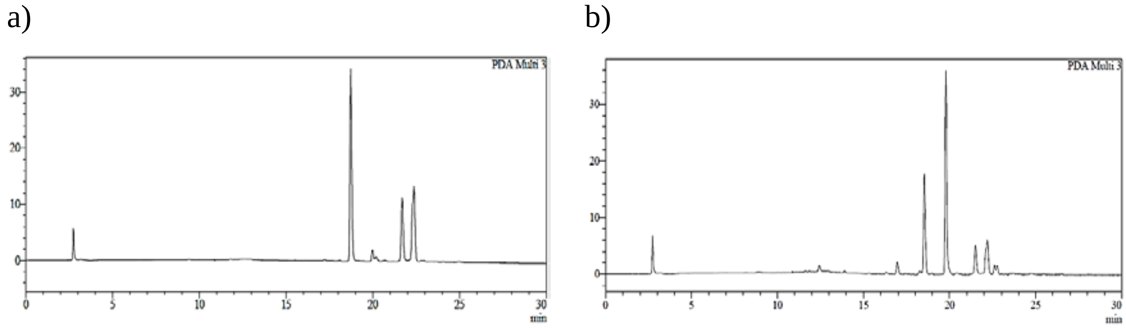


Figure S2: HPLC traces (280 nm) of the **2** + **6** reaction crude (**2** is not detected at this wavelength) at different times: a) 1 h, b) 6.5 h. Peak at 18.5 min is peptide monomer and peak at 20.0 min is peptide dimer.

· **Stability of the adducts in water**:

Michael-type adducts from the **1** + **6** and **1** + **5** reactions were collected and lyophilized, redissolved in water (400 µL) and stirred at 37 ºC. Aliquots of the aqueous solutions were analyzed by HPLC (analysis conditions, 10-40% B) after 1 h, 3 h, 6 h and 24 h (Figure S3):

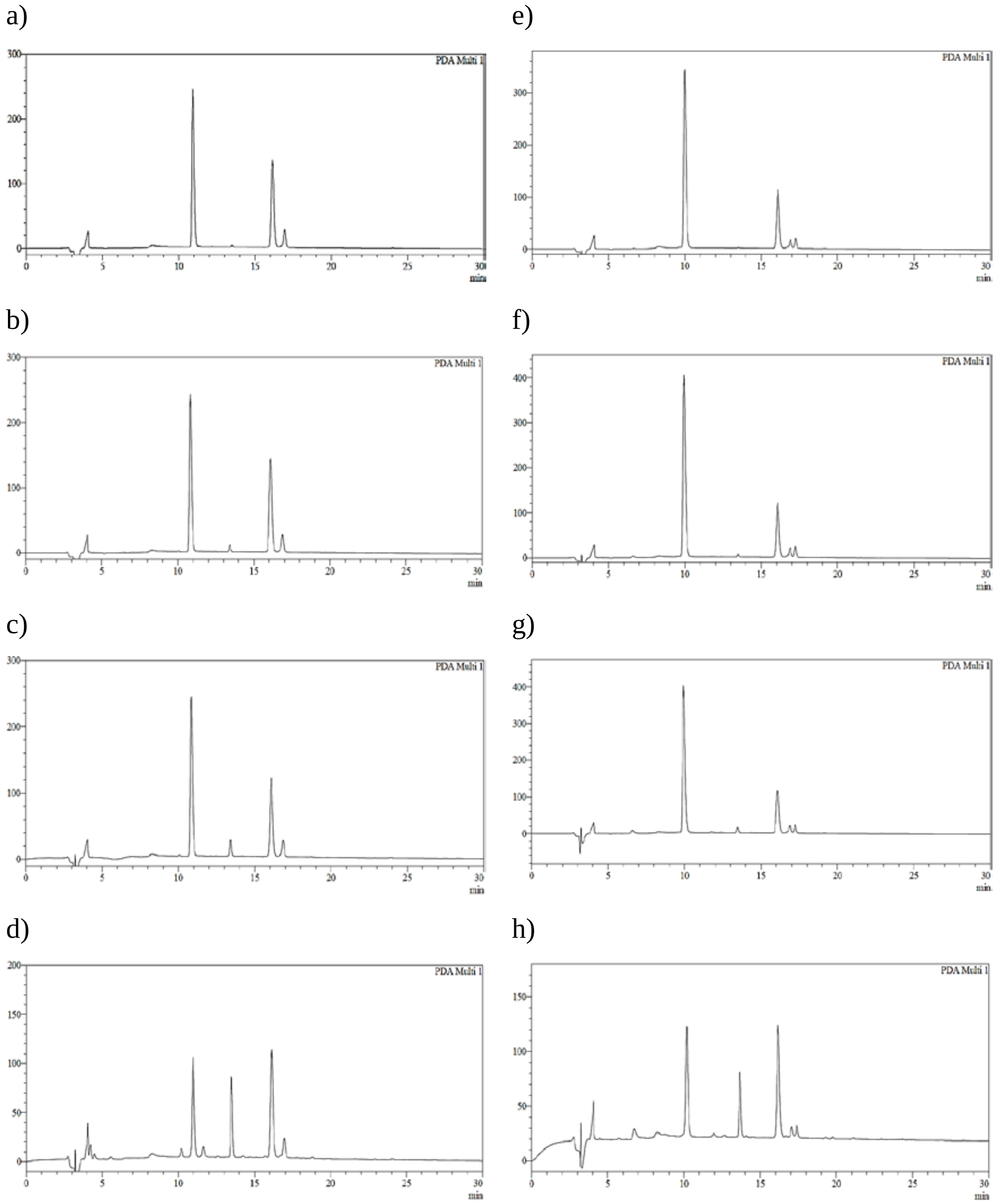


Figure S3: HPLC traces (220 nm) of the aqueous solutions of the MTAs isolated from the **1** + **6** reaction after: a) 1 h, b) 3 h, c) 6 h, d) 24 h, and of the aqueous solutions of the MTAs isolated from the **1** + **5** reaction after e) 1 h, f) 3 h, g) 6 h, h) 24 h. Peaks at 10.3 and 15.5 min are, in both experiments, the starting reagents (peptides and **1**, respectively).

4. Experiments with a peptide containing Cys at the *N*-terminal position

· Reaction of 2,2-dimethyl-4-cyclopentene-1,3-dione (**1**) **with H-CYG-NH$_2$** (**7**):
Peptide H-CYG-NH$_2$ (**7**, 100 nmol) was reacted with 5 equiv of 2,2-dimethyl-4-cyclopentene-1,3-dione (**1**) in water (0.2 mM peptide concentration) at 37 ºC, and the crude analyzed by HPLC at 15 min, 1 h, 3 h, 6 h and at 24 h reaction time (analysis conditions 10-40 % B, Figure S4). Peaks corresponding to the two newly generated products were independently collected, lyophilized, and dissolved in water. A small aliquot of these solutions was immediately analyzed by MALDI-TOF MS, while the rest was left with stirring at 37 ºC for further HPLC analysis (see Figure S5 below).

Product characterization:
HPLC: analysis conditions, 10-40% B, t_R = 22.4 and 23.6 min.
MALDI-TOF (positive mode, DHB): *m/z* (both products) 467.2 $[M+Na]^+$, M calcd. for $C_{21}H_{24}N_4O_5S$ 444.2.

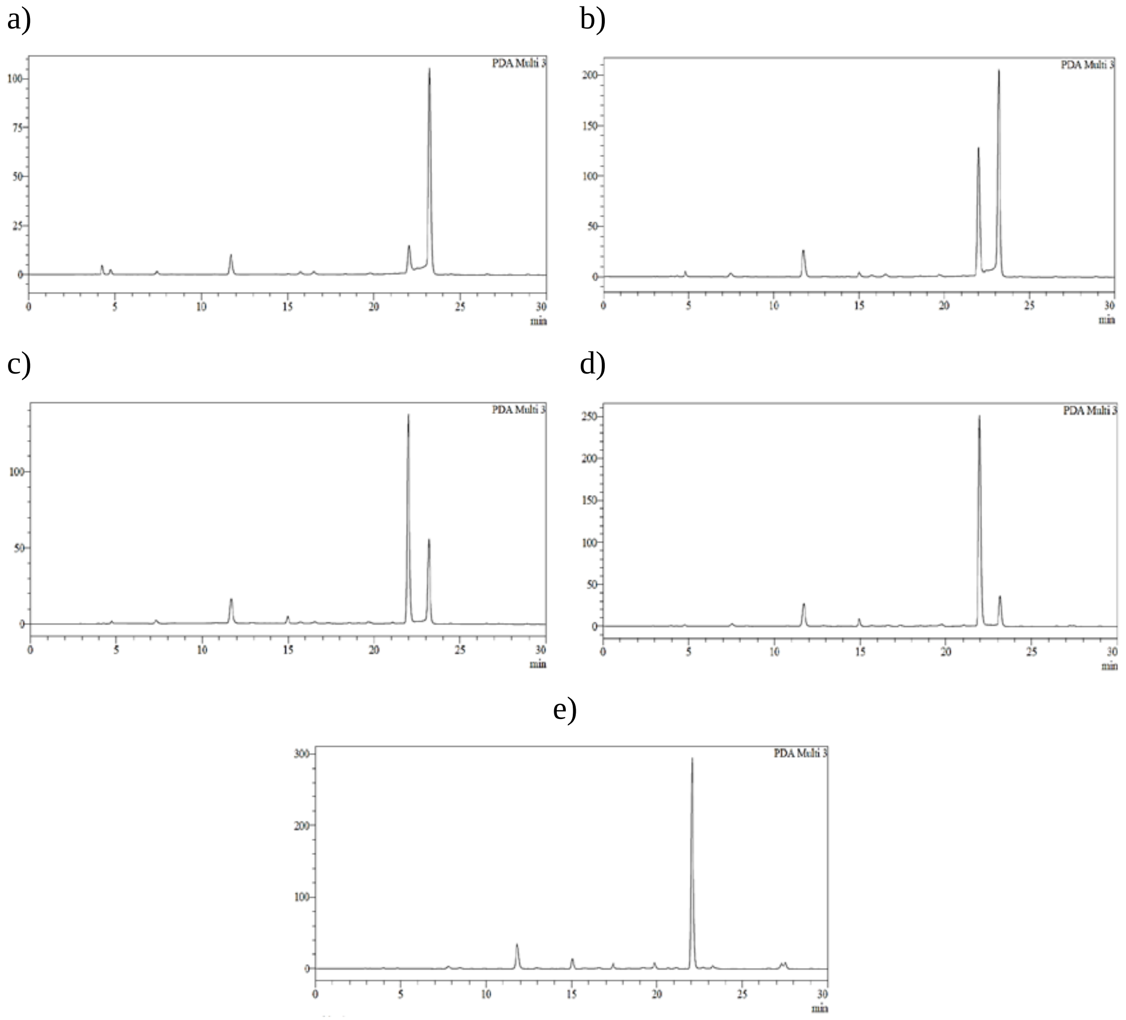


Figure S4: HPLC traces (280 nm) of the **1** + **7** reaction crude (**1** is not detected at this wavelength) at different reaction times: a) 15 min, b) 1 h, c) 3 h, d) 6 h, e) 24 h.

· **Stability of the isolated new products:**

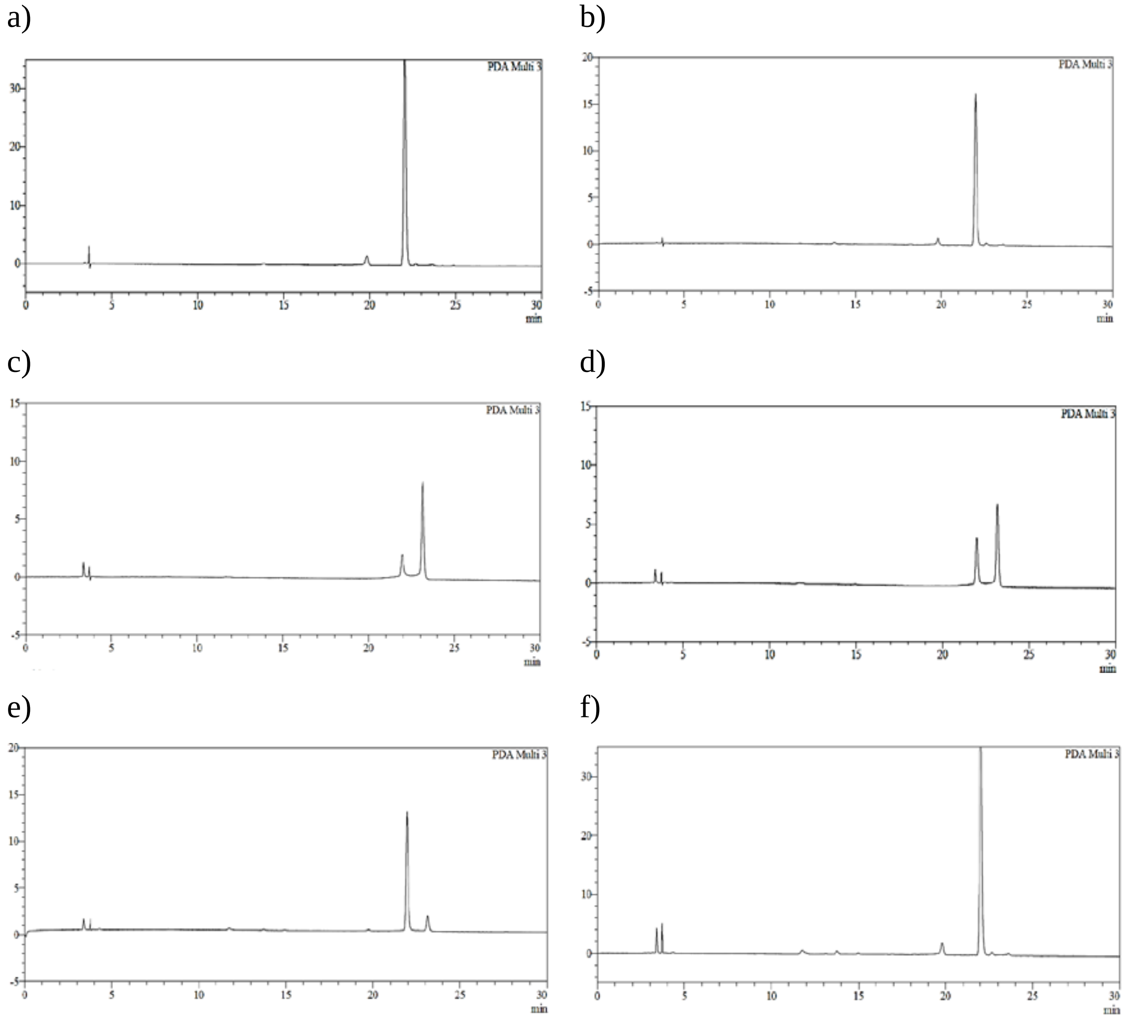


Figure S5: HPLC traces (280 nm) of the peak eluting at 22.4 min after being collected, lyophilized, dissolved in water and stirred at 37 ºC for: a) 15 h, b) 24 h, and of the peak eluting at 23.6 min after the same procedure: c) 15 min, d) 1 h, e) 6 h, f) 24 h.

These experiments showed that the compound eluting at 22.4 min was stable, whilst the adduct eluting at 23.6 min was transformed into this stable compound.

· Reaction of CPD 1 with H-CYG-NH$_2$ (7) - HPLC/MS monitoring:

The reaction **1** + **7** was repeated under the same conditions as above, and the reaction mixture analyzed after 2 h by HPLC/MS (ESI) (Figure S6) instead of carrying out HPLC analyses of the mixture and characterizing the different components by MALDI-TOF MS. It was then discovered that the two newly formed adducts had different masses (which showed that they were indeed different compounds), as indicated below.

Product characterization (HPLC-MS):
HPLC/MS: 10-40% B, t_R = 18.5 min and 19.8 min.
ESI (positive mode) of the peak eluting at 18.5 min: *m/z* 444.6 $[M+H]^+$, M calcd. for $C_{21}H_{24}N_4O_5S$ 444.2 (**M-**20 Da adduct **11**, where **M** = mass of the Michel-type adduct).
ESI (positive mode) of the peak eluting at 19.8 min: *m/z* 446.6 $[M+H]^+$, M calcd. for $C_{21}H_{26}N_4O_5S$ 446.2 (**M-**18 Da adduct **10**, where **M** = mass of the Michel-type adduct).

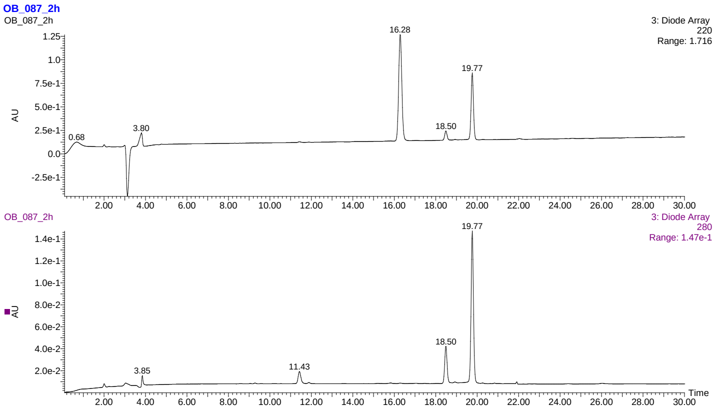


Figure S6: HPLC/MS traces of the **1** + **7** reaction crude after 2 h incubation, monitored at 220 nm (top) and at 280 nm (bottom). Peak at 16.3 min (top) is excess CPD **1** (which can be detected at 220 nm).

5. Reaction of methyl cysteinate with CPDs 1, 2, 3 and 4, and characterization of the final adduct 13

5.1 Reaction of H-Cys-OMe with 1 and characterization of adduct 13:

2,2-Dimethyl-4-cyclopentene-1,3-dione (**1**) (8.0 µL, 0.071 mmol) was dissolved in 1.64 mL of ACN and mixed with an aqueous solution of methyl cysteinate hydrochloride (24.2 mg in 4.92 mL of water, 0.141 mmol, 2 equiv). After 1 h stirring at 25 °C (the reaction had gone to completion, as assessed by HPLC/MS, Figure S7), the mixture was diluted with 60 mL of water, taken to pH=1 with 10 % aq HCl, and extracted with DCM (3 × 60 mL). The organic phase was dried over anhydrous $MgSO_4$, and the solvent removed under reduced pressure. The resulting yellow-to-orange solid was purified by silica gel column chromatography eluting with DCM and increasing amounts of ethyl acetate (→ 40%) to render **13** as a pale yellow solid (7.2 mg, 42%).

mp: 100-101 °C. TLC (DCM/AcOE 1:1): $R_f = 0.61$. IR (ATR, solid): ν 2978, 1734, 1682, 1645 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$) δ: 6.37 (s, 1H), 4.81 (dd, $J = 6.9, 4.6$ Hz, 1H), 3.83 (s, 3H), 3.13 – 3.04 (m, 2H), 1.25 (s, 3H), 1.24 ppm (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ: 205.0, 171.9, 170.2, 149.9, 129.7, 58.2, 53.1, 46.8, 24.2, 22.2, 21.9 ppm. HRMS (ESI, positive mode): m/z 240.0694 $[M+H]^+$, M calcd. for $C_{11}H_{13}NO_3S$ 239.0616. HPLC/MS: 10-50 %B, $t_R = 17.9$ min.

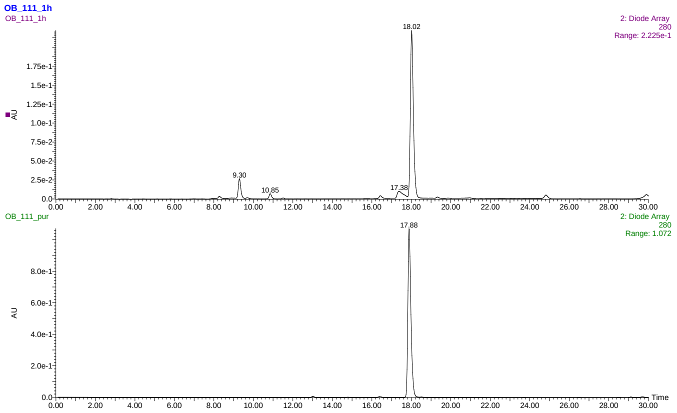


Figure S7: HPLC/MS traces (280 nm) of the crude after 1h reaction (top) and the pure compound (**13**) (bottom).

In addition to the ^1H NMR and ^{13}C NMR spectra, the DEPT at 135 degrees spectrum (101 MHz, $CDCl_3$), the DEPT at 90 degrees spectrum (101 MHz, $CDCl_3$) and the gHSQC spectrum (400 MHz, $CDCl_3$) are shown at the end of this Supporting Information file.

5.2 Reaction of H-Cys-OMe with 2:

H-Cys-OMe (500 nmol) was incubated with 5 equiv of CPD **2** (2500 nmol) in water (1 mM H-Cys-OMe concentration) at 37 ºC, and the crude analysed after 15, 60 and 120 min (Figure S8). New peaks were collected and characterized by ESI-MS.

Product characterization:
HPLC (**M** - 18 Da adducts **14**): analysis conditions, 20-50% B in 20 min, t_R = 16.2 min and 17.7 min.
HPLC (**M** - 20 Da adducts **15**): analysis conditions, 20-50% B in 20 min, t_R = 11.3 min and 11.6 min.
ESI (positive mode) of both peaks: *m/z* 298.2 $[M+H]^+$, M calcd. for $C_{13}H_{15}NO_5S$ 297.0 (**M**-20 adducts **15**, where **M** = mass of the Michel-type adduct).

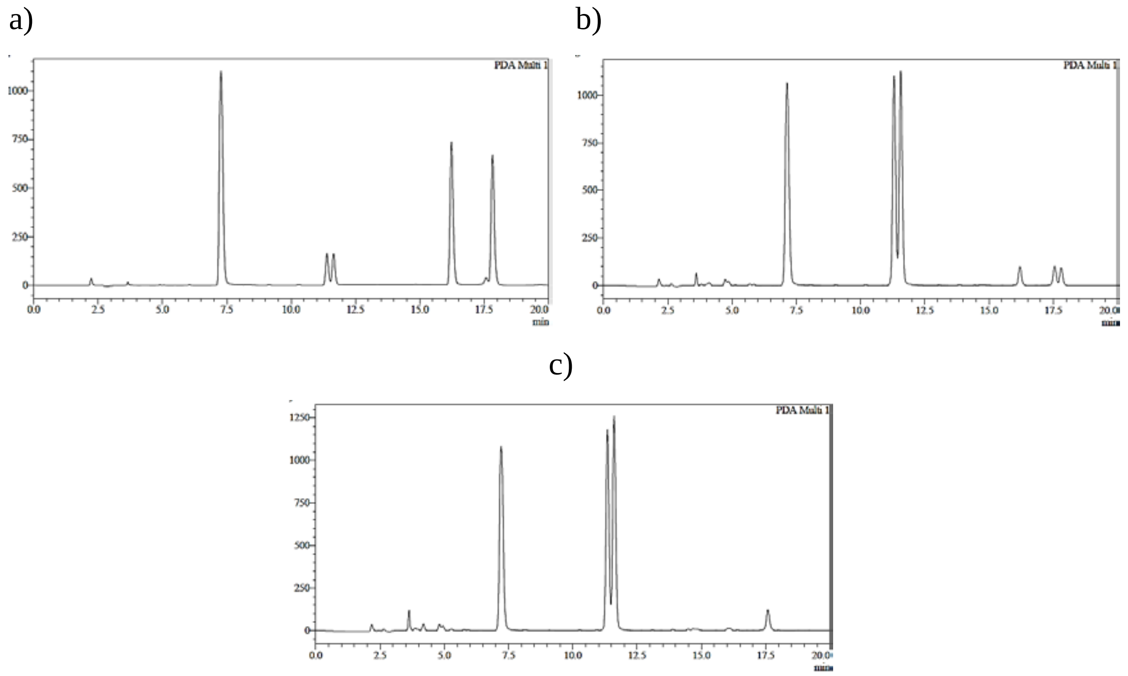


Figure S8: HPLC traces (240 nm) of the reaction of H-Cys-OMe + **2** at a) 15 min, b) 60 min and c) 120 min. Peak at 7.4 is excess **2,** and peaks at 16.2 + 17.7 min are the M-18 Da adducts.

5.3 Reaction of H-Cys-OMe with 3:

CPD-biotin (**3**, 100 nmol) was incubated with 5 equiv of methyl cysteinate (500 nmol) in water (0.2 mM CPD concentration) at 37 ºC, and the crude was analyzed by HPLC/MS after 15 min, 1 h and 24 h reaction times (0-50 %B, Figure S9).

Product characterization:

HPLC/MS: 0-50% B, t_R = 20.1 min, 22.4 min and 23.4 min.

ESI (positive mode) of the products eluting at 20.1 min: *m/z* 565.9 $[M+H]^+$, M calcd. for $C_{25}H_{35}N_5O_6S_2$ 565.20 (**M**-20 adducts **17**, where **M**=mass of the Michel-type adduct).

ESI (positive mode) of the products eluting at 22.4 and 23.4 min: *m/z* 567.9 $[M+H]^+$, M calcd. for $C_{25}H_{37}N_5O_6S_2$ 567.2 (**M**-18 adducts **16**, where **M**=mass of the Michel-type adduct).

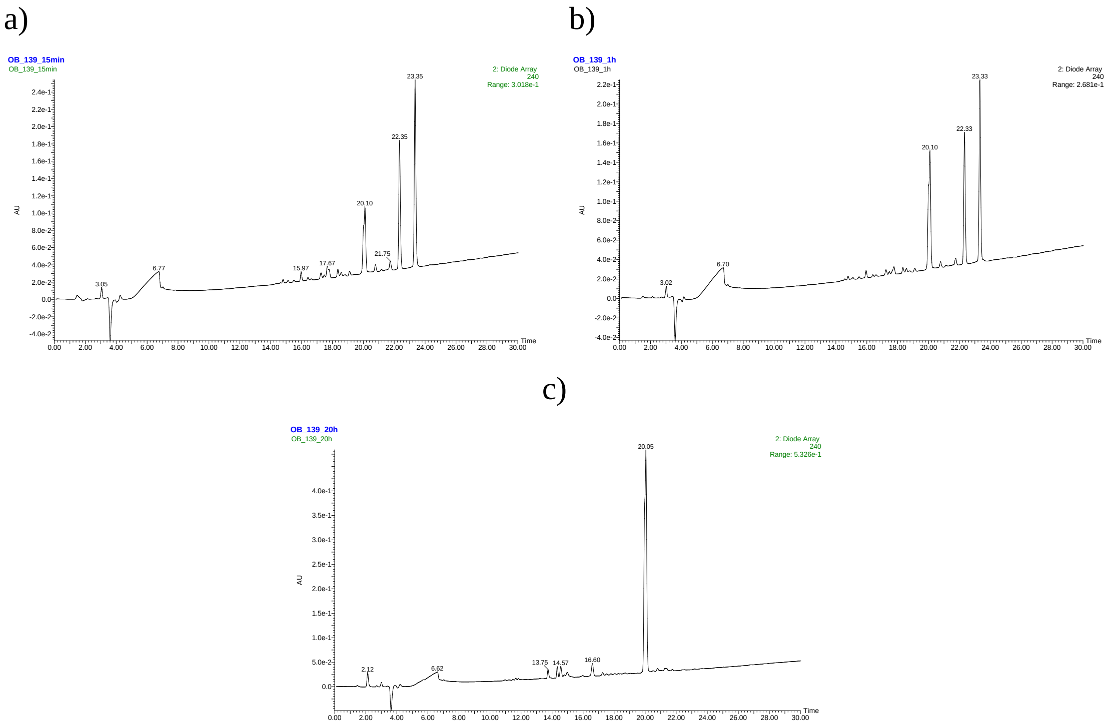


Figure S9: HPLC profiles (240 nm) of the H-Cys-OMe + **3** reaction crude after a) 15 min, b) 1 h and c) 24 h.

5.4 Reaction of H-Cys-OMe with 4:

CPD-Dansyl (**4**, 100 nmol) was incubated with 5 equiv of methyl cysteinate (500 nmol) in water (0.2 mM CPD concentration) at 60 ºC, and the crude was analyzed by HPLC/MS after 1 h and 2.5 h (20-70% B, Figure S10).

Product characterization:

HPLC/MS: 20-70 % B, t_R = 18.8 min, 19.1 min, 22.6 min and 24.1 min.

ESI (positive mode) of the products eluting at 18.7 and 19.1 min: *m/z* 572.8 $[M+H]^+$, M calcd. for $C_{27}H_{32}N_4O_6S_2$ 572.2 (**M**-20 adducts **19**, where **M**=mass of the Michel-type adduct).

ESI (positive mode) of the products eluting at 22.6 and 24.1 min: *m/z* 574.8 $[M+H]^+$, M calcd. for $C_{27}H_{34}N_4O_6S_2$ 574.2 (**M**-18 adducts **18**, where **M**=mass of the Michel-type adduct).

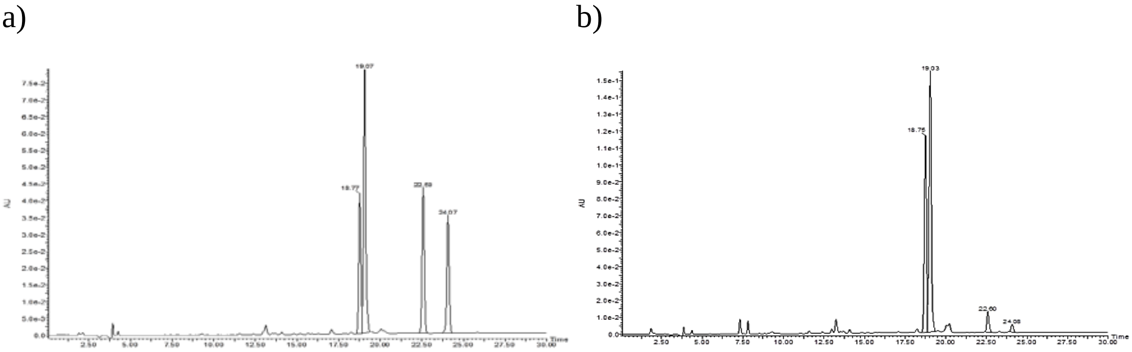


Figure S10: HPLC traces (280 nm) of the reaction of H-Cys-OMe + **4** at a) 60 min and b) 150 min. Peaks at 22.6 + 24.1 min are the M-18 Da adducts.

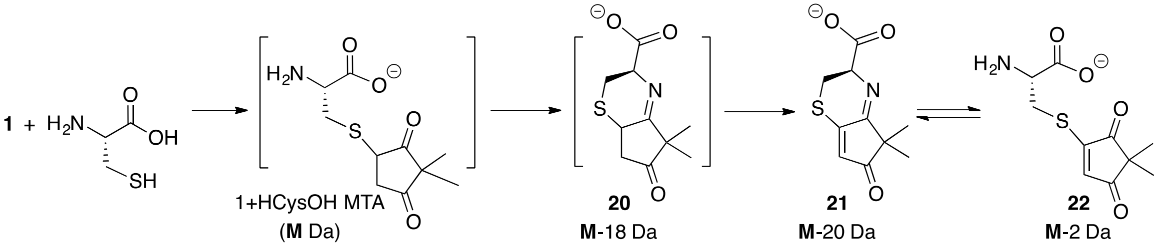
6. Screening of reaction conditions

The next table (Table S2) shows all the conditions tested in this screening. For the preparation of the reaction mixtures, aqueous solutions of methyl cysteinate and **2** were preincubated at the reaction temperature for 10 min. Afterwards, the required amounts of both reagents were mixed. Aliquots of the reaction mixtures (10% of the total volume each) were withdrawn after 15, 60, 120, 240 and 360 min and stored in liquid nitrogen until HPLC-analyzed. Those samples containing an organic co-solvent were diluted with water until the organic solvent content was about 15% (in volume). The **M**-18:**M**-20 relative ratios were determined by HPLC (measuring absorbance at 240 nm). **2** had completely reacted after 1 h in all cases except when the reaction was performed with organic co-solvents or in pure water at 5 ºC, where 2 h were necessary for complete reaction.

				M-20 to M-18 relative ratio				
				Time (min)				
[H-Cys-OMe] (mM)	[**2**] (mM)	Temperature (ºC)	Solvent	15	60	120	240	360
1	1	37	H_2O	4:96	25:75	52:48	87:13	100:0
0.5	0.5	37	H_2O	13:87	36:64	65:35	96:4	-
0.25	0.25	37	H_2O	19:81	49:51	82:18	98:2	-
1	1	60	H_2O	100:0	100:0	-	-	-
1	1	5	H_2O	18:82	21:79	37:63	69:31	83:17
1	1	37	H_2O/ACN 1:1	0:100	1:99	3:97	15:85	48:52
1	1	37	H_2O/MeOH 1:1	1:99	3:97	10:90	36:64	69:31
1	1	37	H_2O/DMSO 1:1	1:99	2:98	6:94	55:45	83:17
1	5	37	H_2O	17:83	92:8	100:0	-	-
1	3	37	H_2O	11:89	81:19	100:0	-	-
5	1	37	H_2O	5:95	20:80	54:46	95:5	-
3	1	37	H_2O	5:95	20:80	45:55	82:18	-

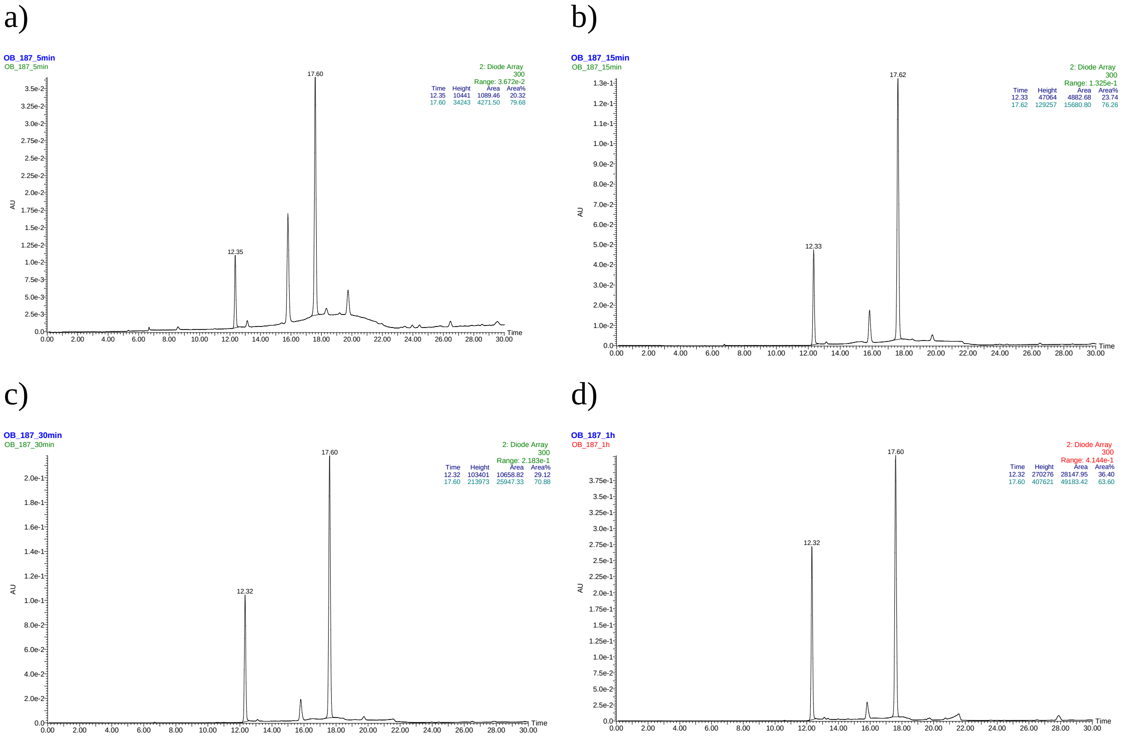
Table S2: **M**-20 Da to **M**-18 Da adduct relative ratios, at different times and under different reaction conditions.

7. Reaction of H-Cys-OH with 1



Scheme S4. Outcome of the reaction **1** + H-Cys-OH.

H-Cys-OH (300 mmol) was reacted with 5 equiv of **1** in water (1 mM cysteine concentration) at 37 ºC. After 5, 15, 30, 60, 120 and 240 min aliquots of the reaction mixture were removed and analyzed by HPLC/MS (0-50% B, Figure S11 and Table S3). Isolated new peaks were incubated in water and reanalyzed to assess their stability (Figure S12). In a separate experiment under the same conditions, aliquots of the reaction mixture were removed after 30, 60 and 120 min and analyzed by ESI-MS (direct perfusion, Figure S13). In this case, HPLC-MS analysis of the mixture showed, from the initial stages, the presence of two adducts exhibiting *m/z* ratios corresponding to **M**-20 Da (**21**) and **M**-2 Da (**22**), and the relative area of the latter increased with time until equilibrium was reached. The **M**-18 Da adduct (**20**) could only be detected by direct perfusion MS analysis, which, for the first and only time, also allowed the expected MTA to be detected.



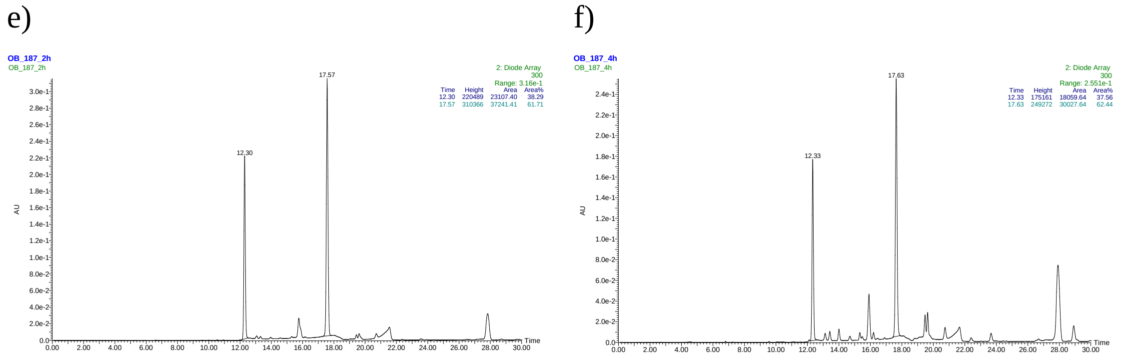


Figure S11: HPLC traces (300 nm) of the reaction between H-Cys-OH and **1** after a) 5 min, b) 15 min, c) 30 min, d) 1 h, e) 2 h and f) 4 h at 37 ºC.

Product characterization:

HPLC/MS: 0-50% B, t_R = 12.3 min and 17.6 min.

ESI (positive mode) of the peak eluting at 12.3 min: 243.7 $[M+H]^+$, M calcd. for $C_{10}H_{13}NO_4S$ 243.0 (**M**-2 Da adduct **22**, where **M** = mass of the Michel-type adduct).

ESI (positive mode) of the peak eluting at 23.0 min: *m/z* 225.7 $[M+H]^+$, M calcd. for $C_{10}H_{11}NO_3S$ 225.0 (**M**-20 Da adduct **21**, where **M** = mass of the Michel-type adduct).

Time (min)	**M**-2 Da adduct **22** (%)	**M**-20 Da adduct **21** (%)
5	20	80
15	24	76
30	29	71
60	36	64
120	38	62
240	38	62

Table S3: Relative ratios of the **M**-2 and **M**-20 adducts at different times, as determined by HPLC (300 nm).

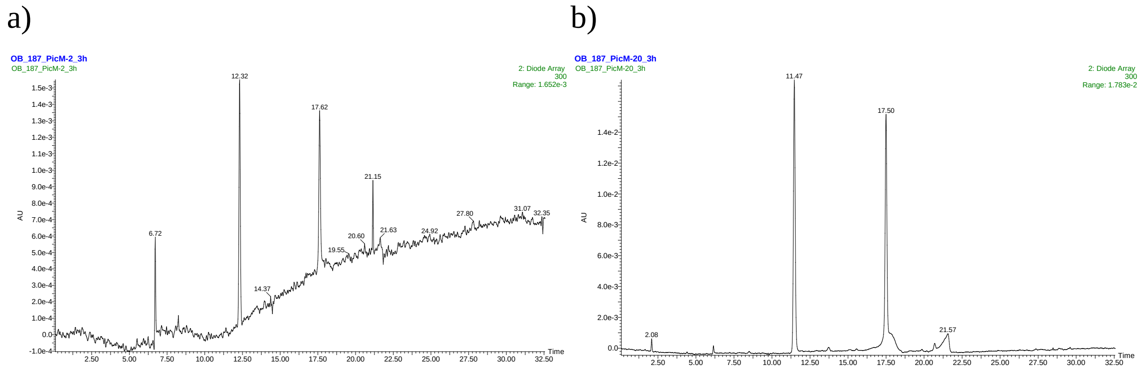


Figure S12: HPLC traces (300 nm) of the isolated **M**-2 Da (a) and **M**-20 Da (b) adducts after incubation in water for 3 h at 37 ºC.

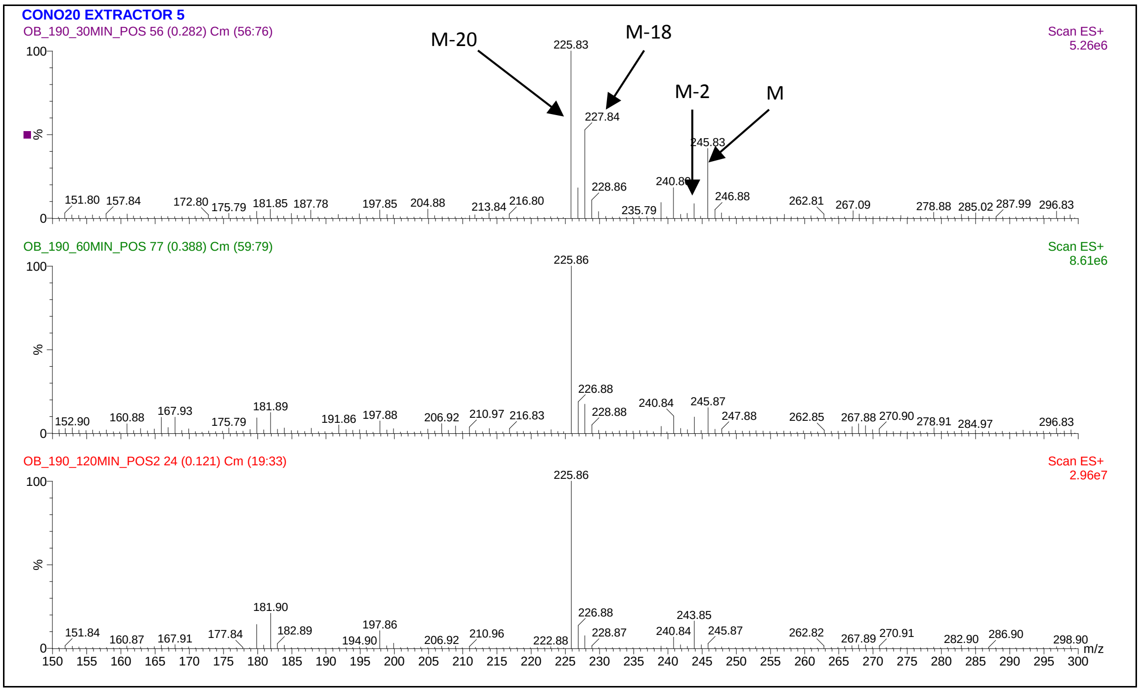
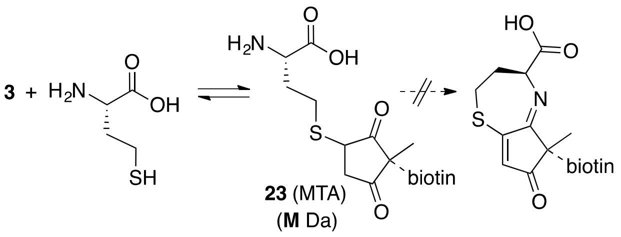


Figure S13: MS analysis (direct perfusion, 20 V, positive mode) of the crude after 30, 60 and 120 min, from top to bottom.

Hydrolysis of adduct 21 to yield 22. The first step of the hydrolysis reaction involves protonation of the N atom of the cycloadduct, and in the case of **21** the resulting intermediate is a zwitterionic neutral species. This points to the formation of **22** as resulting from imine hydrolysis of the **M**-20 Da adduct rather than from oxidation of the MTA (which could be the other alternative). In contrast, protonation of the **M**-20 Da adducts derived from the reaction between CPDs and both peptides and H-Cys-OMe gives a positively charged species and not a zwitterion. Therefore, these **M**-20 Da adducts are more stable and do not undergo imine hydrolysis.

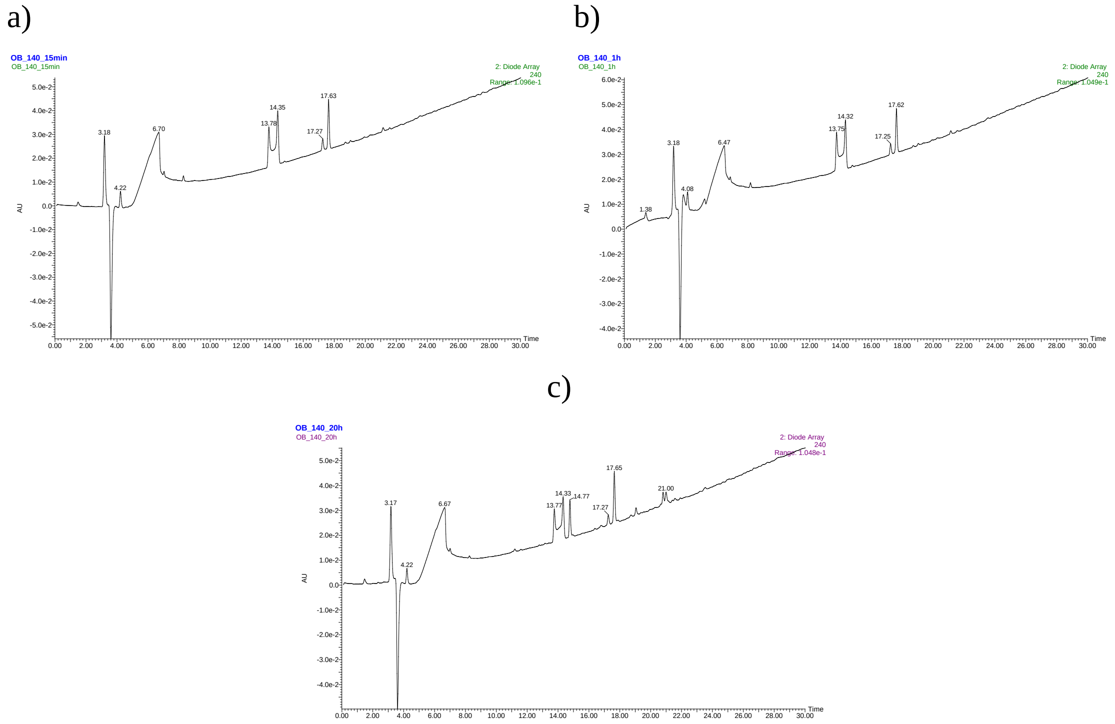
8. Reaction of H-Hcy-OH with 3



Scheme S5. Outcome of the reaction **3** + H-Hcy-OH.

CPD-biotin (**3**, 100 nmol) was incubated with 5 equiv of homocysteine (500 nmol) in water (0.2 mM CPD concentration) at 37 ºC, and the crude was analyzed by HPLC/MS after 15 min, 1 h and 24 h reaction times (0-50% B, Figure S14). New peaks (MTA adducts **23**) formed in the reaction between CPD-biotin **3** and Hcy were collected (together) and lyophilized. Afterwards, they were dissolved in water (400 µL) and reanalyzed after 1, 3 and 6 h of incubation at 37 ºC (Figure S15).

Characterization of the adducts from the Hcy + CPD-biotin (**3**) reaction:
HPLC/MS: 0-50% B, t_R = 13.8 min and 14.4 min.
ESI (positive mode): *m/z* (both products) 586.0 $[M+H]^+$, M calcd. for $C_{25}H_{39}N_5O_7S_2$ (Michael-type adducts **23**) 585.2.


a)

b)

c)

Figure S14: HPLC profiles (240 nm) of the Hcy + **3** reaction after: d) 15 min, e) 1 h and f) 24 h. Peak at 17.6 min is unreacted **3**.

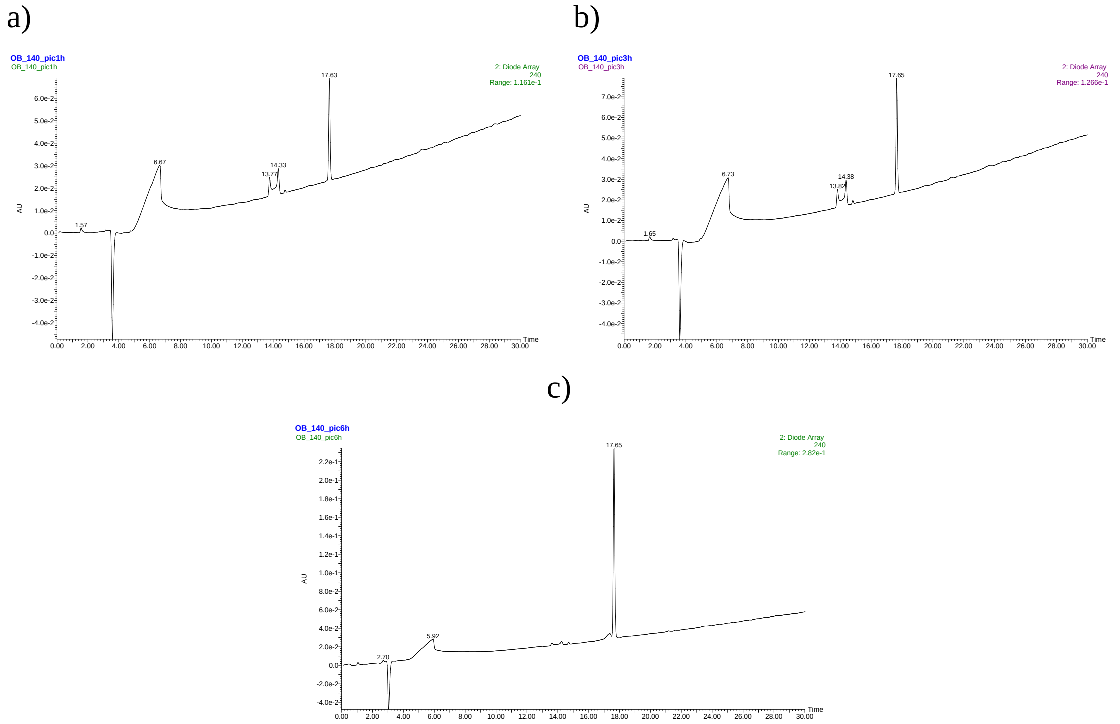


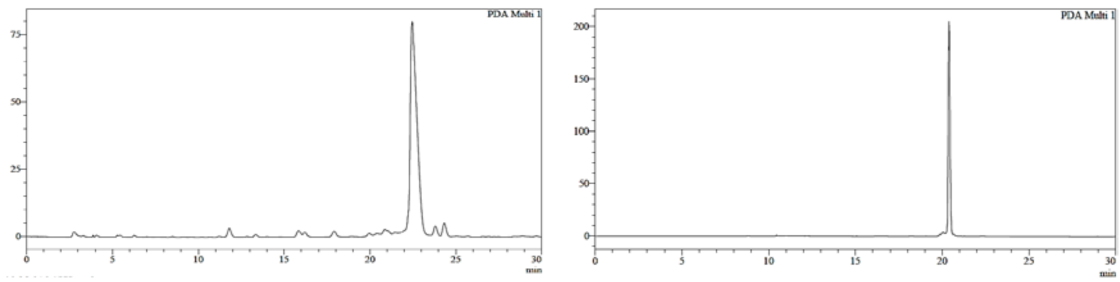
Figure S15: HPLC traces (240 nm) of the aqueous solutions of the products (**23**) isolated from the Hcy + **3** reaction, after: a) 1 h, b) 3 h and c) 6 h incubation in water at 37 ºC. Peak at 17.4 min is biotin-cyclopentenedione **3**.

9. Synthesis of conjugates using CPDs

9.1 CPD-polyamides:

After chain elongation using the standard conditions described above (sections 1.1 & 1.2 of this SI), **2** was incorporated to the protected polyamide-resin by reacting **2** and DIPC with the polyamide-resin for 90 min (5 equiv for incorporation to peptide-resins and 10 equiv for incorporation to PNA-resins). Cleavage and deprotection of the CPD-derivatized polyamide was carried out by treatment with a 95:2.5:2.5 TFA/H_2O/TIS mixture for 2 h at rt. The CPD-polyamide was purified by reversed-phase HPLC (yields given below account for incorporation of **2**, cleavage and deprotection, and purification).

· **CPD-GRGSYEAYK-NH_2 (24a):** 21% yield.
HPLC: purification gradient: 0-40% B; analysis conditions, 0-40% B, t_R = 21.0 min (Figure S16).
MALDI-TOF (positive mode, DHB): *m/z* 1193.5 $[M+H]^+$, M calcd. for $C_{54}H_{76}N_{14}O_{17}$ 1192.6.



a) b)

Figure S16: HPLC traces (280 nm) of crude (a) and purified (b) CPD-derivatized peptide **24a**.

· **CPD-LVPRGSYKTSKLNYL-NH_2 (24b):** 17% yield
HPLC: purification gradient: 20-40% B; analysis conditions, 25-40% B, t_R = 14.3 min (Figure S17).
MALDI-TOF (positive mode, DHB): *m/z* 1959.8 $[M+H]^+$, M calcd. for $C_{91}H_{143}N_{23}O_{25}$ 1958.0.

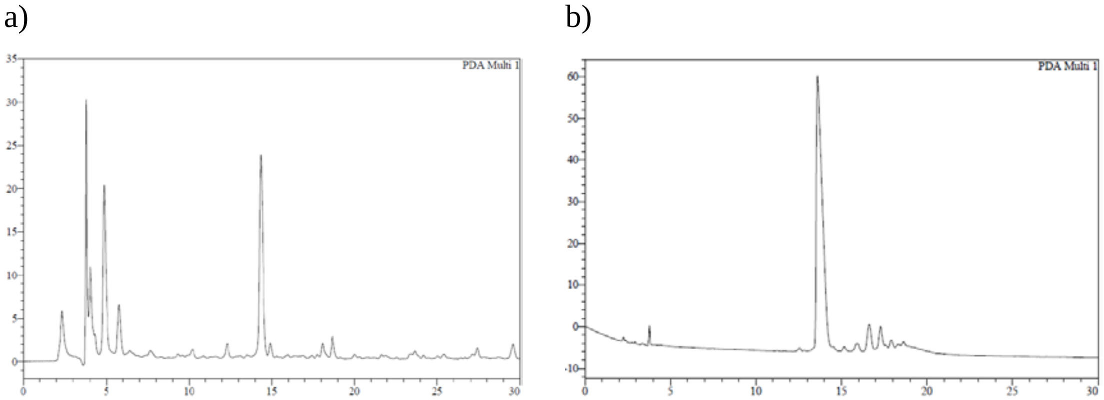


Figure S17: HPLC traces (280 nm) of crude (a) and purified (b) CPD-derivatized peptide **24b**.

· **CPD-catgtcat-K-NH$_2$ (24c):** 11% yield
HPLC: purification gradient: 0-40% B; analysis conditions, 0-40% B, t_R = 19.1 min (Figure S18).
MALDI-TOF (positive mode, DHB): *m/z* 2452.5 [M+H]$^+$, M calcd. for $C_{101}H_{130}N_{46}O_{29}$ 2451.0.

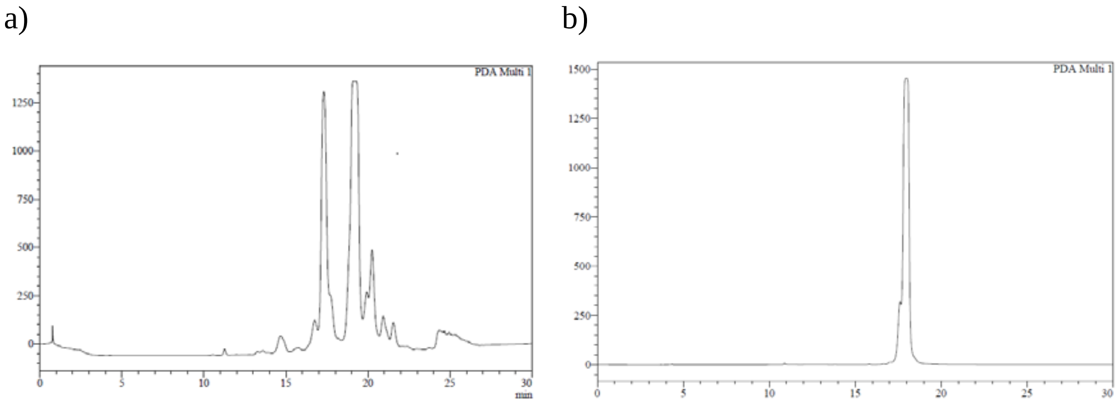


Figure S18: HPLC traces (260 nm) of crude (a) and purified (b) CPD-derivatized PNA **24c**.

9.2 Peptide-PNA conjugates

· H-CK-catagctgtttc-NH_2 + CPD-GRGSYEAYK-NH_2 → conjugate 26ad:

PNA H-CK-catagctgtttc-NH_2 (**25d**, 130 nmol) was reacted with 1.2 equiv of CPD-GRGSYEAYK-NH_2 (**24a**) in water at 37 ºC (0.85 mM PNA concentration). Progress of the reaction was monitored by HPLC (analysis conditions, 5-35% B, 60 ºC). Although the reaction seemed to have finished after 3 h, it was left stirring for 5 h to make sure that all the **M**-18 Da adducts had evolved into the **M**-20 Da ones (one of the **M**-18 Da adducts coeluted with one of the **M**-20 Da adducts).

The final conjugate (**26ad**) was purified (purification conditions, 10-30% B, 60 ºC) and isolated in 51% yield.

Product characterization:

HPLC: analysis conditions, 5-35% B, 60 ºC, t_R = 15.1 min and 15.4 min (Figure S19).

MALDI-TOF (positive mode, DHB): *m/z* (both products) 4639.7 $[M+H]^+$, M calcd. for $C_{192}H_{253}N_{81}O_{57}S$ (conjugate) 4636.9.

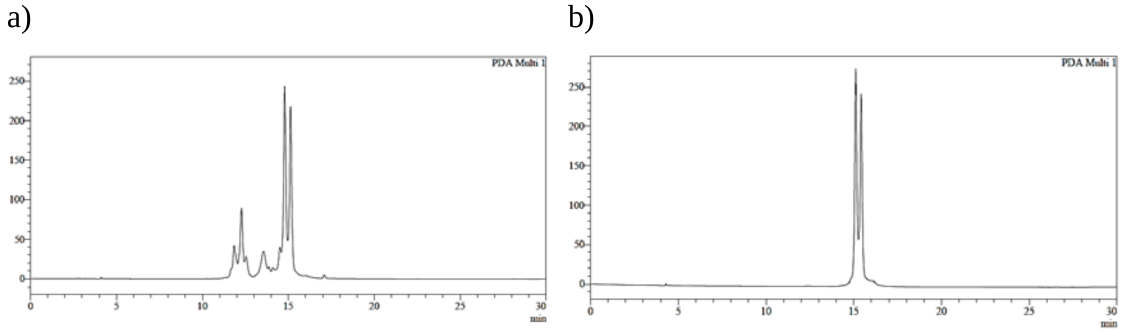


Figure S19: HPLC profiles (260 nm) of crude (a) and purified (b) conjugate **26ad**. Peaks around 12 min in (a) are impurities present in the PNA used for conjugation.

· H-CSYAKYG-NH_2 + CPD-catgtcat-K-NH_2 → conjugate 26ce:

Peptide H-CSYAKYG-NH_2 (**25e**, 50 nmol) was reacted with 1.2 equiv of CPD-catgtcat-K-NH_2 (**24c**) in water at 37 ºC (0.5 mM peptide concentration). Progress of the reaction was monitored by HPLC-MS (analysis conditions, 0-15% B, 60 ºC). This reaction seemed to have finished after 2 h (90% conjugation yield, HPLC-based).

The final conjugate (**26ce**) was purified by HPLC (0-30% B, 60 ºC) and isolated in 40% yield.

Product characterization:

HPLC: analysis conditions, 0-15% B, t_R = 26.7 and 27.7 min (Figure S20).

MALDI-TOF (positive mode, DHB): *m/z* 3222.4 $[M+H]^+$, M calcd. for $C_{136}H_{177}N_{55}O_{36}$ 3220.3.

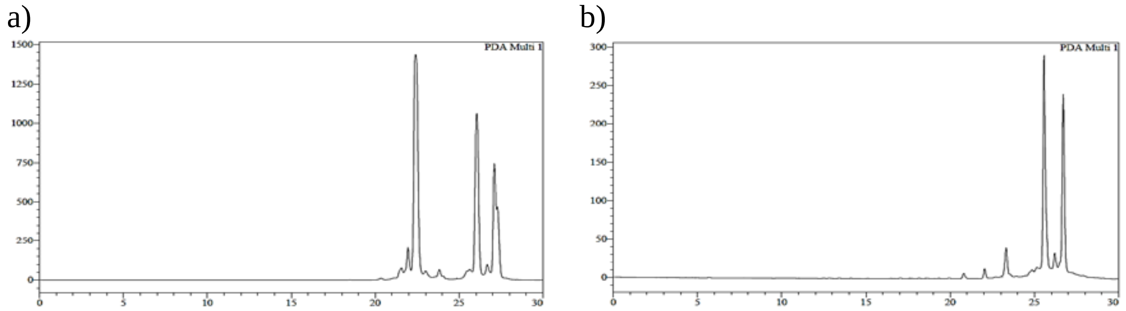


Figure S20: HPLC traces (260 nm) of crude (a) and purified (b) conjugate **26ce**. Peaks around 22 min in (a) is PNA excess used for conjugation.

9.3 Assessment of conjugate stability

· **Stability at different pHs**:

The stability of the 1,2-aminothiol-CPD adducts was studied with PNA-peptide conjugate **26ad** at pH 6.0, 7.0 and 8.0. 5 nmol-samples of the conjugate were dissolved with 500 µL of the appropriate buffer (0.2 M phosphates). Each sample was incubated at 37 ºC and analyzed after 2, 6 and 24 h by HPLC (analysis conditions, 5-35% B, 60 ºC, 260 nm, Table S4).

Time (h)	pH = 6.0	pH = 7.0	pH = 8.0
2	100	100	100
6	99	98	87
24	97	86	65

Table S4: Remaining percentage of conjugate **26ad** after incubation at different times and different pHs.

· **Stability to other thiols:**

5 nmol-samples of the conjugate **26ad** were dissolved with 500 µL of pH = 7.0 buffer (0.2 M phosphates) and mixed with either 100 equiv of methyl cysteinate or 100 equiv of glutathione. Each sample was incubated at 37 ºC and analyzed after 2, 6 and 24 h by HPLC (analysis conditions, 5-35% B, 60 ºC, 260 nm, Table S5).

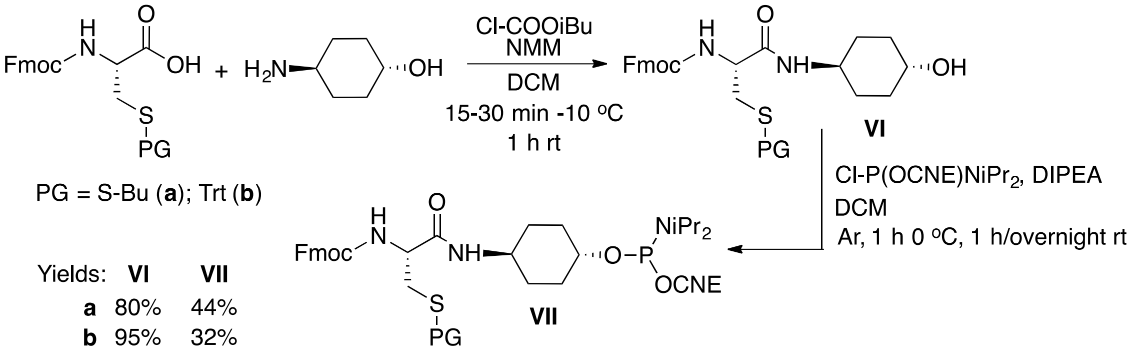
Time (h)	+ Glutathione	+ H-Cys-OMe
2	100	79
6	93	62
24	86	53

Table S5: Remaining percentage of conjugate **26ad** after incubation at different times, with different thiols, at pH = 7.0

9.4 Cys-oligonucleotides

9.4.1 Cysteine phosphoramidites

Cysteine phosphoramidites were prepared following described procedures[7,8] with minor modifications (see below).



Scheme S6. Synthesis of cysteine-containing phosphoramidites **VII**.

Fmoc-Cys(PG)-OH (0.696 mmol) was dissolved in anhydrous DCM (10 mL) at low temperature (–10 ºC). Subsequently, NMM (168 μL, 1.531 mmol) and isobutyl chloroformate (197 μL, 1.531 mmol) were added under an argon atmosphere. After 30 min of preactivaction, *trans*-4-aminocyclohexanol (160 mg, 1.392 mmol), dissolved in anhydrous DCM (8 mL), was added. The reaction mixture was stirred for 30 min at –10ºC and then left 1 hour at rt. Finally, the crude was poured onto 100 mL of hexanes, and the precipitate filtered in a büchner funnel and washed with H_2O (3 × 30 mL) and hexanes (3 × 30 mL) successively. If insoluble product was observed in the mother liquor the filtering process was repeated. Finally, the solid was dissolved in DCM (10 mL), dried over $MgSO_4$ and filtered, and the solvent was removed under reduced pressure to yield the title product (**VI**) as a white solid. **VIa**: 498 mg, 80% yield; **VIb**: 648 mg, 95% yield.
DIPEA (369 μL, 2.118 mmol) was added to a solution of **VI** (0.706 mmol) in anhydrous DCM (4 mL for **VIa** and 10 mL for **VIb**) under an argon atmosphere, and the mixture cooled to 0 ºC. Subsequently, a solution of Cl-P(OCNE)NiPr$_2$ (250 mg, 1.059 mmol) in anhydrous DCM (1 mL) was added dropwise. The mixture was reacted at 0 ºC for 1 h, and then at rt (1 h in the case of **VIa** and overnight in the case of **VIb**). Once the reaction was finished, as assessed by TLC (hexanes/EtOAc/Et_3N 58:40:2, Rf = 0.4) 0.1 mL of MeOH was added and the solvent removed under low pressure. The resulting crude was dissolved in EtOAc (20 mL) and washed with a saturated aq. solution of $NaHCO_3$ (2 × 20 mL) and brine (2 × 20 mL). The organic layer was dried over anhydrous $MgSO_4$, and the residue obtained after solvent removal was purified by silica gel column chromatography eluting with a hexanes/EtOAc/TEA mixture of increasing polarity (from 83:25:2 to 40:58:2), which afforded the title product (**VII**)as a white foam. **VIIa**: 240 mg, 44% yield; **VIIb**: 200 mg, 32% yield.

Fmoc-Cys(S-tBu)-NH-cHex-O-P(OCNE)N^{i}Pr$_2$ VIIa: ^{1}H NMR (400 MHz, CD_3CN): δ 7.85 (d, *J* = 8.0 Hz, 2H), 7.70 (d, *J* = 7.5 Hz, 2H), 7.44 (t, J = 8 Hz, 2H), 7.36 (td, *J* = 7.5, 1.2 Hz, 2H), 6.60 (d, *J* = 8.1 Hz, 1H), 6.08 (d, *J* = 8.3 Hz, 1H), 4.44 – 4.22 (m, 4H), 3.84 – 3.55 (m, 6H), 2.65 (t, *J* = 6.0 Hz, 2H), 2.02 – 1.85 (m, 4H), 1.53 – 1.38 (m, 2H), 1.34

(s, 9H), 1.30 – 1.25 (m, 2H), 1.19 (d, *J* = 1.9 Hz, 6H), 1.18 (d, *J* = 1.9 Hz, 6H). ^{13}C NMR (101 MHz, CD$_3$CN): δ 169.28, 156.13, 144.36, 141.43, 138.21, 129.21, 128.51, 128.02, 127.45, 127.42, 125.55, 125.50, 120.30, 118.93, 72.11, 71.94, 66.75, 58.54, 58.35, 54.89, 32.66, 30.07, 29.37, 24.26, 24.19, 24.12, 24.05, 20.37, 20.30. ^{31}P NMR (162 MHz, CD$_3$CN): δ 145.42 ppm.

· **Fmoc-Cys(Trt)-NH-cHex-O-P(OCNE)N^{i}Pr$_2$ VIIb:** ^{1}H NMR (400 MHz, CDCl$_3$): δ 7.73 (t, *J* = 7.0 Hz, 2H), 7.54 (d, *J* = 7.5 Hz, 2H), 7.44 – 7.32 (m, 8H), 7.29 – 7.15 (m, 11H), 5.72 (d, *J* = 7.9 Hz, 1H), 5.18 (d, *J* = 7.5 Hz, 1H), 4.35 (d, *J* = 6.6 Hz, 2H), 4.17 (t, *J* = 7.0 Hz, 1H), 3.83 – 3.52 (m, 6H), 2.73 – 2.52 (m, 4H), 2.00 – 1.83 (m, 3H), 1.54 – 1.39 (m, 2H), 1.25 – 1.11 (m, 16H). ^{13}C NMR (101 MHz, CDCl$_3$): δ 169.15, 144.34, 143.63, 141.28, 129.56, 128.07, 127.75, 127.06, 126.90, 124.98, 124.94, 120.00, 117.64, 71.86, 71.68, 67.33, 66.92, 58.20, 58.02, 54.14, 47.57, 47.08, 43.11, 42.99, 34.00, 32.37, 30.36, 24.69, 24.61, 24.50, 24.42, 22.51, 22.39, 20.39, 20.32. ^{31}P NMR (162 MHz, CDCl$_3$): δ 145.91 ppm.

9.4.2 Cysteine(PG) oligonucleotides

H-Cys(S-tBu)-NH-cHex-O-P(O)(O^{-})P-O-$^{5'}$dT$_{10}$:

No purification was required (see Figure S21). Yield: 80%.

HPLC: analysis conditions, 0-50 % B, t$_R$ = 15.0 min (Figure S21).

MALDI-TOF (negative mode, THAP/CA): *m/z* 3345.1 [M-H]$^{-}$, M calcd. for $C_{113}H_{156}N_{22}O_{72}P_{10}S_2$ 3346.6.

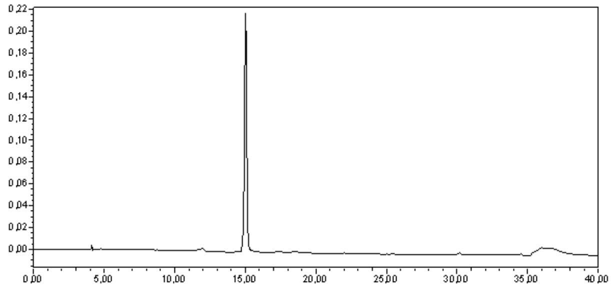


Figure S21. HPLC profile (254 nm) of crude oligonucleotide H-Cys(S-tBu)-NH-cHex-O-P(O)(O^{-})P-O-$^{5'}$dT$_{10}$.

· **H-Cys(Trt)-NH-cHex-O-P(O)(O^{-})P-O-$^{5'}$dT$_{10}$:**

No purification was required (see Figure S22). Yield: 70%.

HPLC: analysis conditions, 10-50%B, t$_R$ = 22.8 min (Figure S22).

MALDI-TOF (negative mode, THAP/CA): *m/z* 3498.9 [M-H]$^{-}$, M calcd. for $C_{128}H_{162}N_{22}O_{72}P_{10}S$ 3500.6.

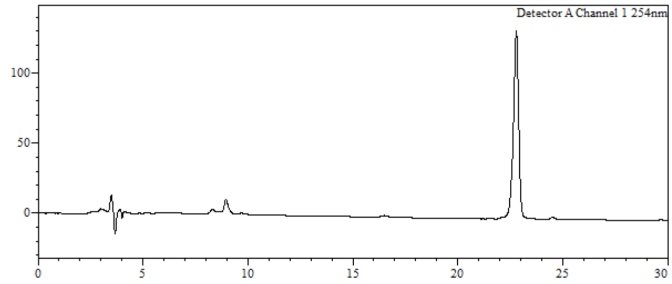


Figure S22. HPLC trace (254 nm) of crude oligonucleotide H-Cys(Trt)-NH-cHex-O-P(O)(O$^-$)P-O-$^{5'}$dT$_{10}$.

· H-Cys(S-tBu)-NH-cHex-O-P(O)(O$^-$)-O-$^{5'}$dCAGATGTCAC:

No purification was required (see Figure S23). Yield: 71%

HPLC: analysis conditions, 10-40% B, t_R = 11.0 min (Figure S23).

MALDI-TOF (negative mode, THAP/CA): *m/z* 3375.1 [M-H]$^-$, M calcd. for $C_{110}H_{148}N_{40}O_{61}P_{10}S_2$ 3378.6.

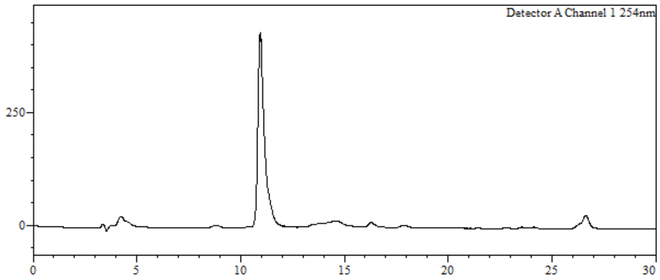


Figure S23. HPLC profile (254 nm) of crude oligonucleotide H-Cys(S-tBu)-NH-cHex-O-P(O)(O$^-$)P-O-$^{5'}$dCAGATGTCAC.

· H-Cys(Trt)-NH-cHex-O-P(O)(O$^-$)-O-$^{5'}$dCAGATGTCAC:

No purification was required (see Figure S24). Yield: 72%

HPLC: analysis conditions, 10-40% B, t_R = 23.4 min (Figure S24).

MALDI-TOF (negative mode, THAP/CA): *m/z* 3531.1 [M-H]$^-$, M calcd. for $C_{125}H_{154}N_{40}O_{61}P_{10}S$ 3532.7.

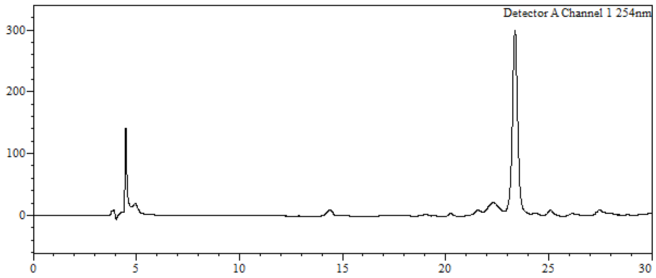


Figure S24. HPLC profile (254 nm) of crude oligonucleotide H-Cys(Trt)-NH-cHex-O-P(O)(O$^-$)P-O-$^{5'}$dCAGATGTCAC.

· **H-Cys(S-tBu)-NH-cHex-O-P(O)(O$^-$)-O-$^{5'}$dCAGCAGCAGAGTCTTCATCAT:**
Purification gradient: 10-40% B . Yield: 28%
HPLC: analysis conditions, 10-40% B, t_R = 11.8 min (Figure S25).
MALDI-TOF (negative mode, THAP/CA): *m/z* 6753.0 [M-H]$^-$, M calcd. for $C_{217}H_{283}N_{80}O_{127}P_{21}S_2$ 6755.2.

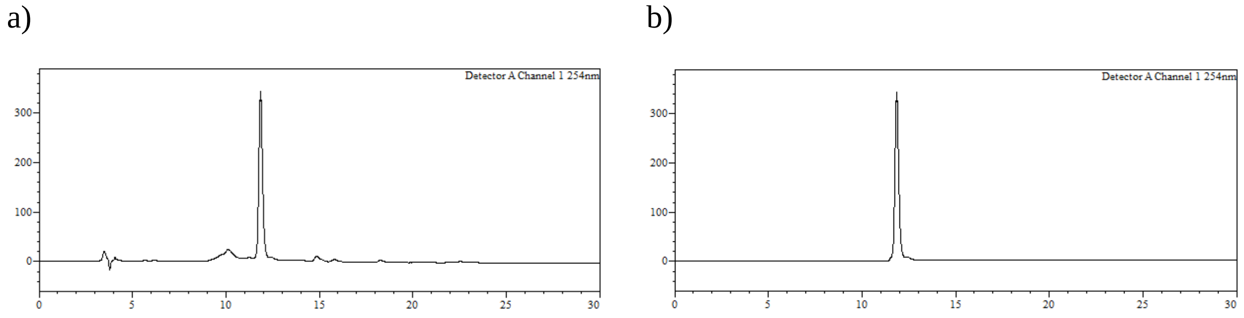


Figure S25. HPLC profiles (254 nm) of oligonucleotide H-Cys(S-tBu)-NH-cHex-O-P(O)(O$^-$)P-O-$^{5'}$dCAGCAGCAGAGTCTTCATCAT, crude (a) and purified (b).

· **H-Cys(Trt)-NH-cHex-O-P(O)(O$^-$)-O-$^{5'}$dCAGCAGCAGAGTCTTCATCAT:**
Purification gradient: 10-40% B . Yield: 32%
HPLC: analysis conditions, 10-40% B, t_R = 24.5 min (Figure S26).
MALDI-TOF (negative mode, THAP/CA): *m/z* 6905.7 [M-H]$^-$, M calcd. for $C_{232}H_{289}N_{80}O_{127}P_{21}S$ 6909.3.

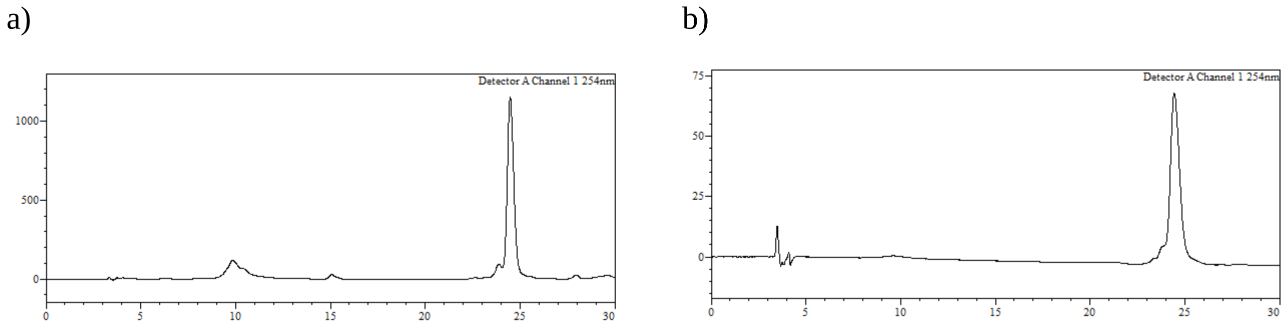


Figure S26. HPLC profiles (254 nm) of oligonucleotide H-Cys(Trt)-NH-cHex-O-P(O)(O$^-$)P-O-$^{5'}$dCAGCAGCAGAGTCTTCATCAT, crude (a) and purified (b).

9.4.3 Cysteine deprotection and Cys-oligonucleotides characterization

· **Deprotection of Cys(S-tBu)-oligonucleotides:**
Cysteine deprotection was carried out by adding TCEP (500 equiv) and water (400 μL) to the lyophilized oligonucleotide (200 nmol; crude or purified, depending on the quality of the crude). The pH was adjusted to 5 by adding 5% aq NaOH, and the mixture was

stirred overnight at 37 °C. Gel filtration of the resulting crude (Sephadex G-25, elution with H_2O, 3 mL/min) allowed most of the excess TCEP to be removed, and the Cys-oligonucleotide was finally purified by reversed-phase HPLC. It is very important to fully remove TCEP, otherwise the remaining phosphine reacts with the CPD and conjugation does not take place (or very dirty crudes are obtained).

· **Deprotection of Cys(Trt)-oligonucleotides:**

$AgNO_3$ (10 μL of a 10 mM solution in water, 5 equiv) was added to a solution of the oligonucleotide (20 nmol) in water (30 μL), and the mixture was reacted for 45 min at 37 °C. DTT was then added (14 μL of a 10 mM solution, 7 equiv), and the mixture left to react for an additional 30 min (37 °C). After centrifugation and decantation, the supernatant was cleaned by gel filtration, which was followed by HPLC purification.

· **H-Cys-NH-cHex-O-P(O)(O^-)P-O-$^{5'}$dT$_{10}$ (25f):**

Purification gradient: 0-50% B. Yield: 30%.

HPLC: analysis conditions, 0-50% B, t_R = 16.0 min (Figure S27).

MALDI-TOF (negative mode, THAP/CA): *m/z* 3257.2 $[M-2H+Na]^-$, M calcd. for $C_{109}H_{148}N_{22}O_{72}P_{10}S$ 3258.6

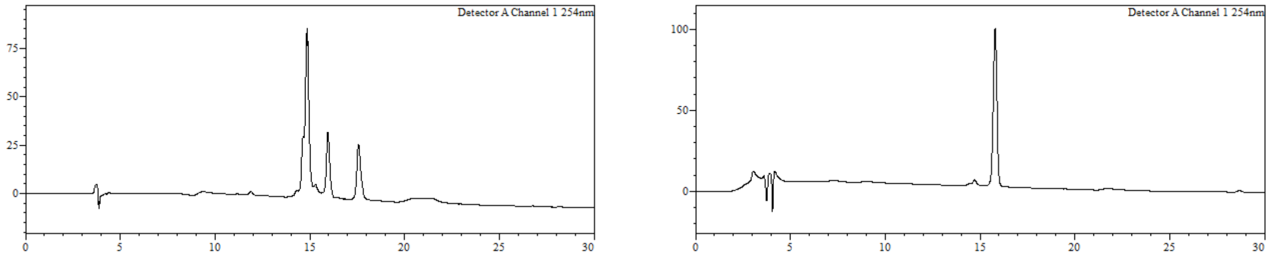

a) b)

Figure S27. HPLC profiles (254 nm) of oligonucleotide **25f**, crude (a) and purified (b).

· **H-Cys-NH-cHex-O-P(O)(O^-)-O-$^{5'}$dCAGATGTCAC (25g):**

Purification gradient: 0-25% B. Yield: 58%.

HPLC: analysis conditions, 0-25 % B, 60 °C t_R = 19.8 min (Figure S28).

MALDI-TOF (negative mode, THAP/CA): *m/z* 3290.1 $[M-H]^-$, M calcd. for $C_{106}H_{140}N_{40}O_{61}P_{10}S$ 3290.6

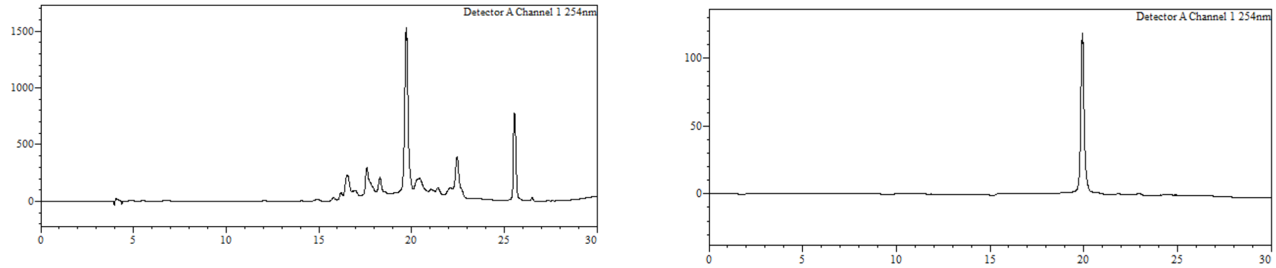

a) b)

Figure S28. HPLC profiles (254 nm) of oligonucleotide **25g**, crude (a) and purified (b).

· **H-Cys-NH-cHex-O-P(O)(O$^-$)-O-$^{5'}$dCAGCAGCAGAGTCTTCATCAT (25h):**
Purification: 0-30% B, 60 ºC . Yield: 50%
HPLC: analysis conditions, 0-30% B, 60 ºC t_R = 22.5 min (Figure S29).
MALDI-TOF (negative mode, THAP/CA): *m/z* 6662.3 [M-H]$^-$, M calcd. for $C_{213}H_{275}N_{80}O_{127}P_{21}S$ 6667.2.

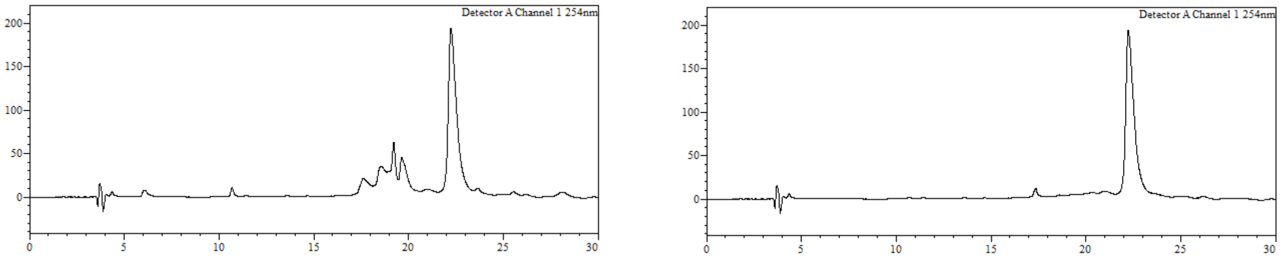

a)
b)

Figure S29. HPLC profiles (254 nm) of oligonucleotide **25h**, crude (a) and purified (b).

9.5 Synthesis of peptide-oligonucleotide conjugates

A mixture containing Cys-oligonucleotide and CPD-peptide (1.2 equiv) was prepared (either by mixing solutions of each of the reagents, or by dissolving one reagent and pouring this solution into an eppendorf containing a lyophilized of the other reagent) making the oligonucleotide concentration 0.5 mM. This mixture was reacted at 37 °C, and the reaction progress was monitored by HPLC. Formation of the conjugate (2-5 h) was assessed by the appearance of two new peaks absorbing at 330 nm (absorption maximum of the stable CPD-1,2-aminothiol adduct). The target conjugate was purified by HPLC.

No differences were observed when the conjugation reaction was carried out with Cys-oligonucleotides of different origin, that is, obtained from either Cys(S-tBu)-oligonucleotides or Cys(Trt)-oligonucleotides.

· H-C-dT$_{10}$ (25f) + CPD-GRGSYEAYK-NH$_2$ (24a) → conjugate 26af

Purification gradient: 10-60% B. Yield: 50%.

HPLC: analysis conditions, 10-60% B, t_R = 11.1 and 11.5 min (Figure S30).

MALDI-TOF (negative mode, THAP/CA): *m/z* 4433.2 [M-H]$^-$, M calcd. for $C_{163}H_{220}N_{36}O_{88}P_{10}S$ 4431.1.

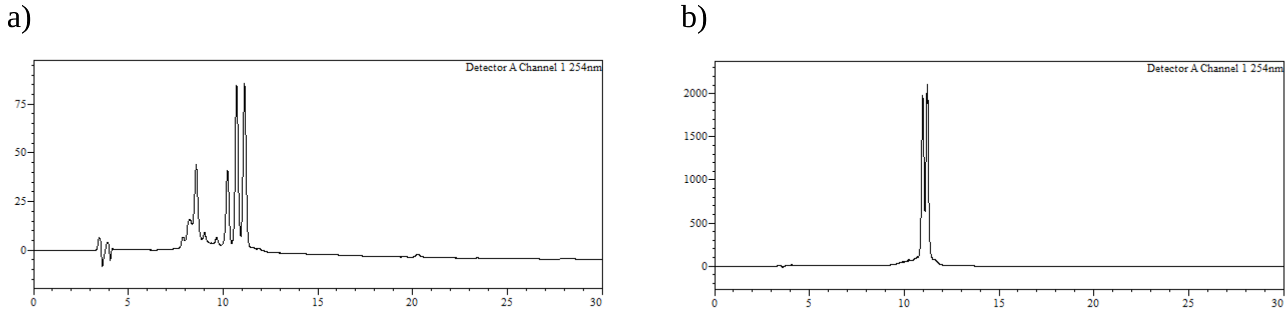


Figure S30. HPLC profiles (254 nm) of conjugate **26af**, crude (a) and purified (b).

· H-C-$^{5'}$dCAGATGTCAC (25g) + CPD-GRGSYEAYK-NH$_2$ (24a) → conjugate 26ag

Purification gradient: 5-35% B. Yield: 60%.

HPLC: analysis conditions, 5-35% B, 60 °C t_R = 23.4 and 24.3 min (Figure S31).

MALDI-TOF (negative mode, THAP/CA): *m/z* 4466.4 [M-H]$^-$, M calcd. for $C_{160}H_{212}N_{54}O_{77}P_{10}S$ 4463.4.

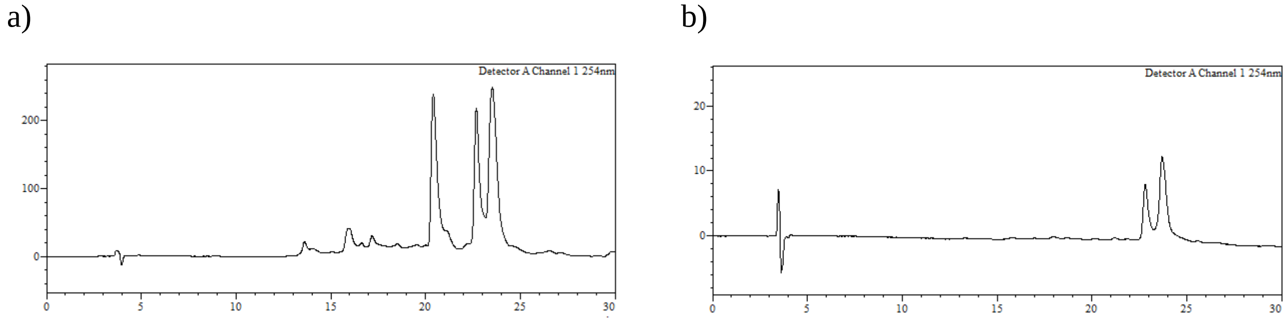
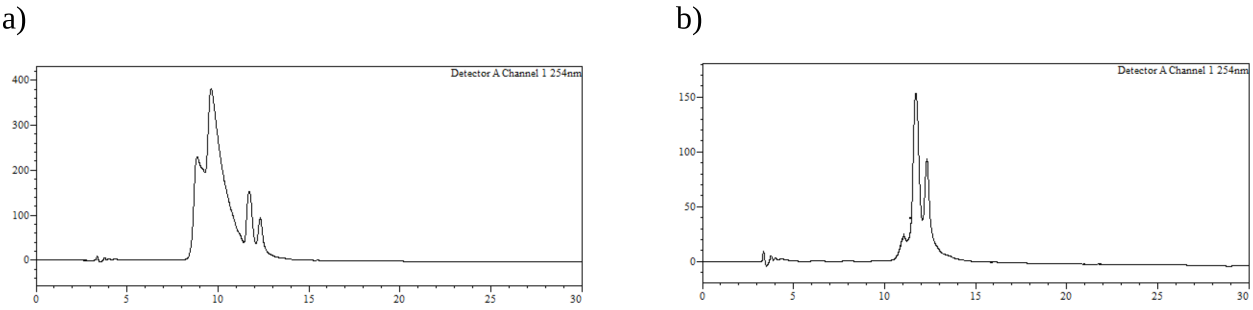


Figure S31. HPLC profiles (254 nm) of conjugate **26ag**, crude (a) and purified (b).

· H-C-$^{5'}$dCAGCAGCAGAGTCTTCATCAT (25h) + CPD-GRGSYEAYK-NH$_2$ (24a) → conjugate 26ah

Purification gradient: 10-40% B. Yield: 44%.
HPLC: analysis conditions, 10-40% B, 60 ºC t_R = 11.9 and 12.3 min (Figure S32).
MALDI-TOF (negative mode, THAP/CA): *m/z* 7832.4 $[M-H]^-$, M calcd. for $C_{267}H_{347}N_{94}O_{143}P_{21}S$ 7839.7.



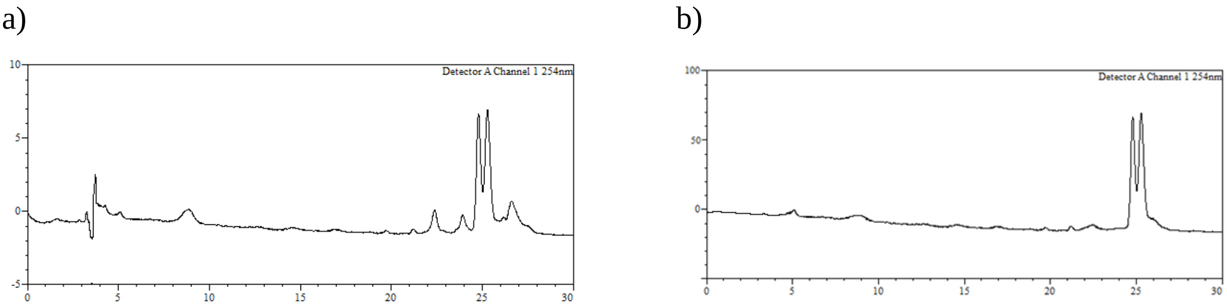
a)

b)

Figure S32. HPLC profiles (254 nm) of conjugate **26ah**, crude (a) and purified (b).

· H-C-$^{5'}$dCAGATGTCAC (25g) + CPD-LVPRGSYKTSKLNYL-NH$_2$ (24b) → conjugate 26bg

Purification gradient: 30-70% B, 60 ºC. Yield: 20%
HPLC: analysis conditions, 20-50% B, 60 ºC t_R = 24.8 and 25.3 min (Figure S33).
MALDI-TOF (negative mode, THAP/CA): *m/z* 5225.1 $[M-H]^-$, M calcd. for $C_{197}H_{279}N_{63}O_{85}P_{10}S$ 5228.6.



a)

b)

Figure S33. C HPLC profiles (254 nm) of conjugate **26bg**, crude (a) and purified (b).

· H-C-$^{5'}$dCAGCAGCAGAGTCTTCATCAT (25h) + CPD-LVPRGSYKTSKLNYL-NH$_2$ (24b) → conjugate 26bh

Purification gradient: 30-70% B. Yield: 33%.
HPLC: analysis conditions, 10-40% B, 60 ºC t_R = 20.0 min (Figure S34).
MALDI-TOF (negative mode, THAP/CA): *m/z* 8601.3 $[M-H]^-$, M calcd. for $C_{304}H_{414}N_{103}O_{151}P_{21}S$ 8605.2.

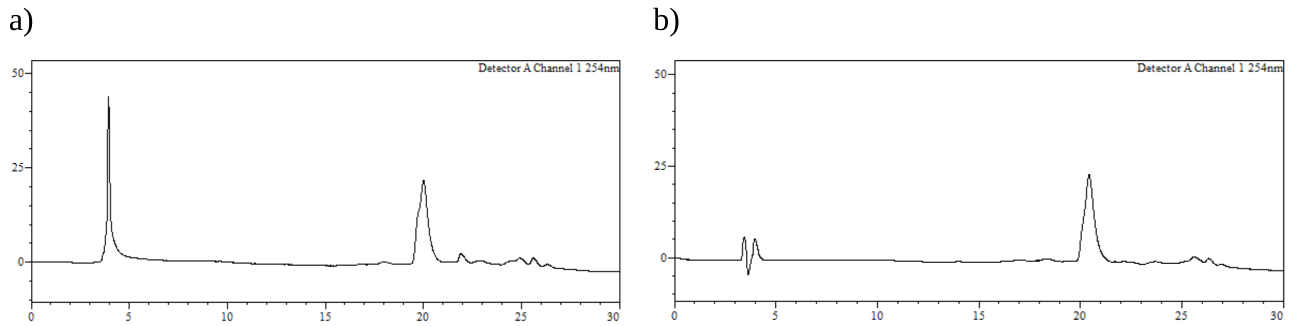


Figure S34. HPLC profiles (254 nm) of conjugate **26bh**, crude (a) and purified (b).

10. Competition experiments with mixtures of peptides

· Experiment 1: peptides 25e and 6 + excess CPD (**2**)

Peptides H-CSYAKYG-NH$_2$ (**25e**) and H-KYAYCG-NH$_2$ (**6**) (100 nmol each) were dissolved in 150 µL of water and CPD **2** was added (3.5 equiv, 50 µL of a 7.0 mM solution). The mixture was analyzed by HPLC (analysis conditions, 0-50% B, Figure S35) before adding **2** (the two peptide peaks exhibited almost identical areas, which confirmed that their relative ratio was 1:1, see Figure S35a), and after its addition, at 30 and 90 min (at 90 min formation of the final **M**-20 Da adduct **27** between H-CSYAKYG and **2** was almost complete). 120 min after the beginning of the incubation process, H-CSYAKYG-NH$_2$ (400 nmol) was added, and the mixtured analyzed 60 min later (the **2** + **25e** reaction had gone to completion). After HPLC purification (purification conditions, 20-60% B) the final **2** + **25e** **M**-20 Da adducts (**27**) were isolated (45% global yield).

Characterization of **M**-20 Da adducts **27**:
HPLC: analysis conditions, 0-50% B, t_R = 20.9 min and 21.7 min (Figure S35).
ESI (positive mode): *m/z* (both products) 952.7 $[M+H]^+$, M calcd. for $C_{44}H_{57}N_9O_{13}S$ 951.4.

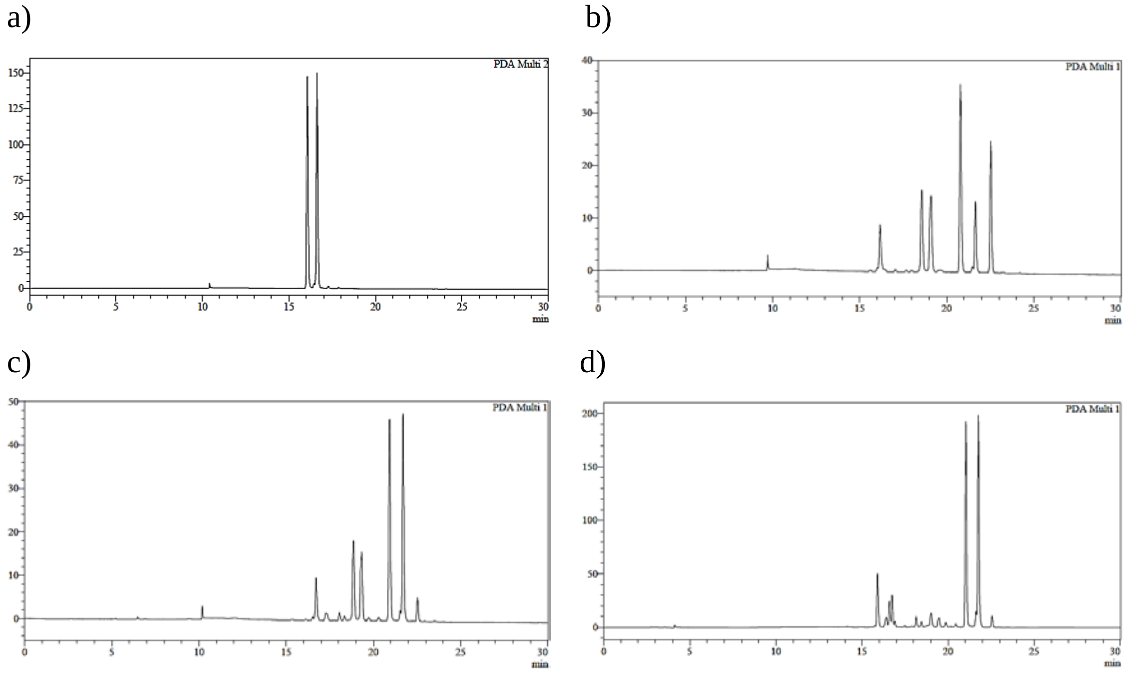


Figure S35: HPLC traces (280 nm) of the **25e** + **6** + **2** reaction crude: before adding **2** (a), 30 and 90 min after the addition of **2** (b and c, respectively), and 60 min after the addition of 400 nmol of peptide **25e** H-CSYAKYG-NH$_2$ (d). Peaks at 18.8 and 19.3 min correspond to the reversible Michael-type adducts formed by reaction between **2** and H-KYAYCG-NH$_2$ (**6**) (see also Table S6 below).

· Experiment 2: excess peptides 25e and 6 + CPD (**2**)

Peptides H-CSYAKYG-NH$_2$ (**25e**) and H-KYAYCG-NH$_2$ (**6**) (100 nmol each) were dissolved in 100 µL of water and CPD **2** was added (0.7 equiv, 100 µL of a 0.7 mM solution). The reaction was monitored by HPLC (analysis conditions, 0-50% B, Figure S36) before adding **2** (the two peptides exhibited almost identical areas, see Figure S36a), and 30 min after its addition. At that point, an additional 0.3 equiv of **2** was added. Further HPLC analyses were carried out 120 and 180 min after the first (0.7 equiv) addition of **2** (reaction was completely finished at 180 min, as assessed by the complete disappearance of peptide **25e** and complete evolution of the intermediate to the final **M**-20 Da adduct **27**). After HPLC purification (purification conditions, 20-60% B) the two final adducts **27** were independently isolated (41% global yield) and characterized (for characterization data of adducts **27** see above).

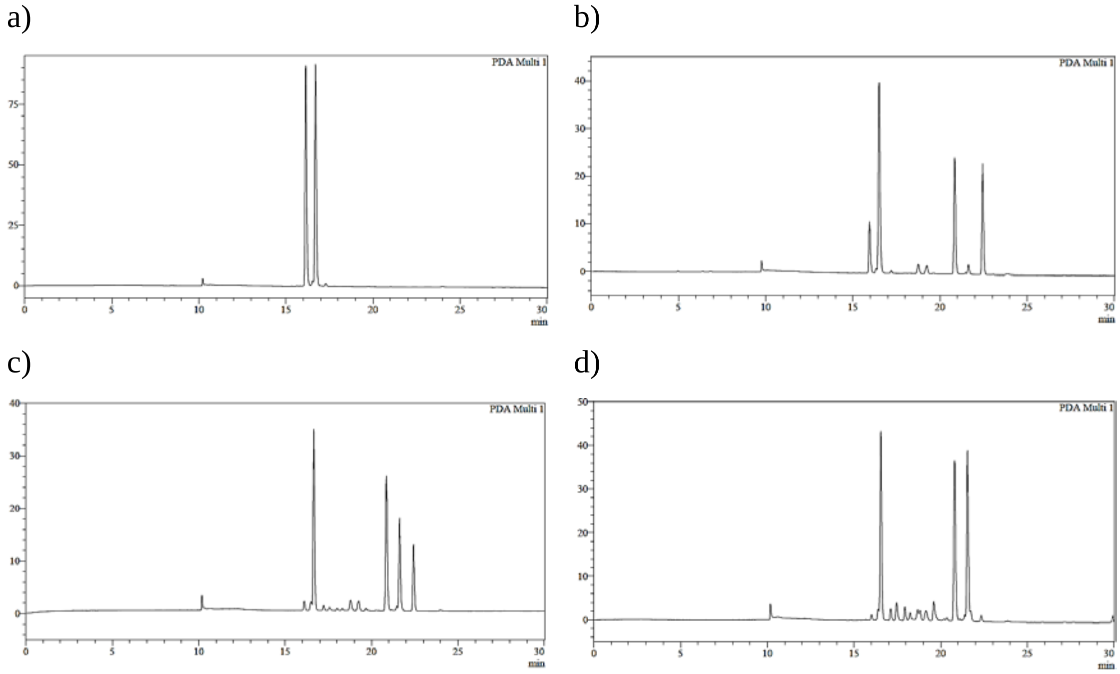


Figure S36: HPLC traces (280 nm) of the reaction crude before adding **2** (a), 30 min after the addition of 0.7 equiv of **2**, and 120 and 180 min total reaction times (c and d, respectively) (see also Table S6 below).

· **Experiment 3: equimolar amounts of peptides (25e and 6) and CPD** (**2**)

Peptides H-CSYAKYG-NH_2 (**25e**) and H-KYAYCG-NH_2 (**6**) (100 nmol of both) were dissolved in 58 µL of water, and CPD **2** was added (1 equiv, 142 µL of a 0.7 mM solution). The mixture was analyzed by HPLC (analysis conditions, 0-50% B, Figure S37) before adding **2**, and after its addition, at 30 min 90 and 180 min (at this point the reaction was finished, as the complete conversion to the **M**-20 Da adduct indicated). After HPLC purification (purification conditions, 20-60% B) the two final adducts **27** were independently isolated (47% global yield) and characterized (for characterization data of adducts **27** see above).

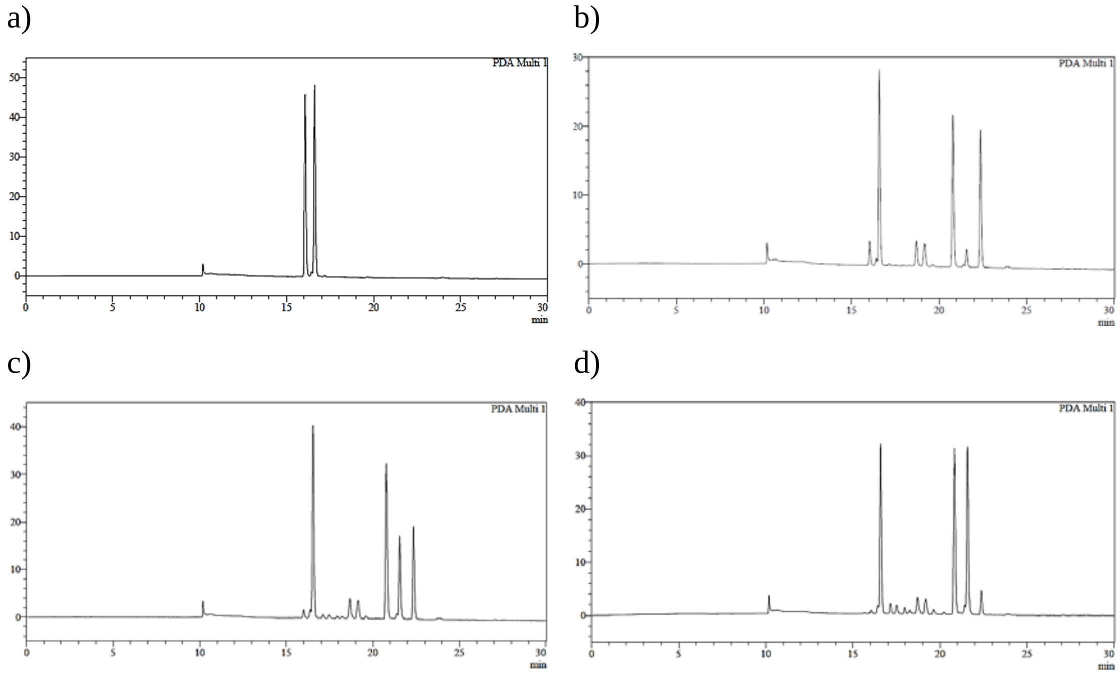


Figure S37: HPLC traces (280 nm) of the reaction crude before adding **2** (a), and 30 min (b), 90 min (c) and 180 min (d) after its addition (see also Table S6 below).

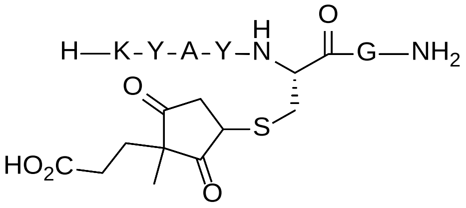
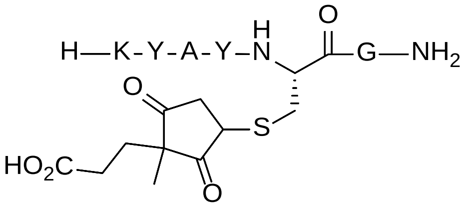
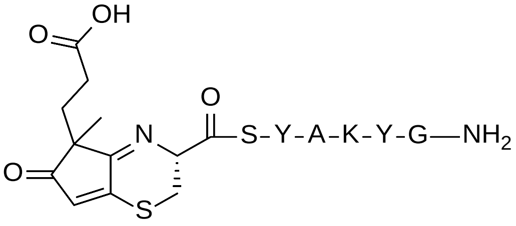
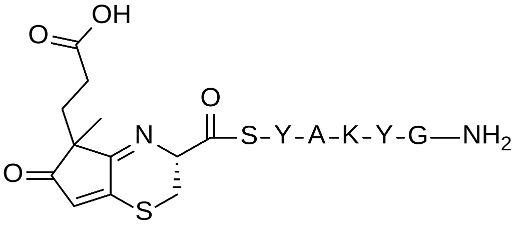
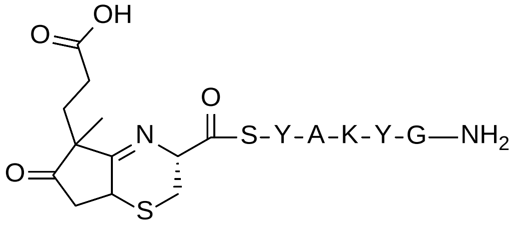
Retention time (min)	Structure
16.1	H—C-S-Y-A-K-Y-G—NH_2 (**25e**)
16.7	H—K-Y-A-Y-C-G—NH_2 (**6**)
18.8	 MTA (**M** Da)
19.3	 MTA (**M** Da)
20.9	 **M**-20 Da, **27**
21.7	 **M**-20 Da, **27**
22.5	 **M**-18 Da

Table S6: Structures and retention times of the products formed in the previous three experiments.

11. Double derivatization of peptides with two cysteines

11.1 Proof of principle experiment:

Peptide H-CSYACKYG-NH_2 (**28**) (100 nmol) was incubated with 1.1 equiv of **2** in water, at 37 ºC and for 90 min (the final peptide concentration was 0.5 mM). 3-Maleimidopropanoic acid (5 equiv, 40 µL of a 12.5 mM solution) was then added, and the mixture left stirring for an additional 1 h. After purification (purification conditions, 20-60% B), the final product (doubly conjugated peptide **29a**) was obtained in 44% yield.

Characterization of conjugate **29a**:

HPLC/MS: analysis conditions, 0-40% B, t_R = 18.0 min and 18.7 min (Figure S38).

HRMS (ESI, positive mode): *m/z* 1224.4310 $[M+H]^+$, M calcd. for $C_{54}H_{69}N_{11}O_{18}S_2$ 1223.4263.

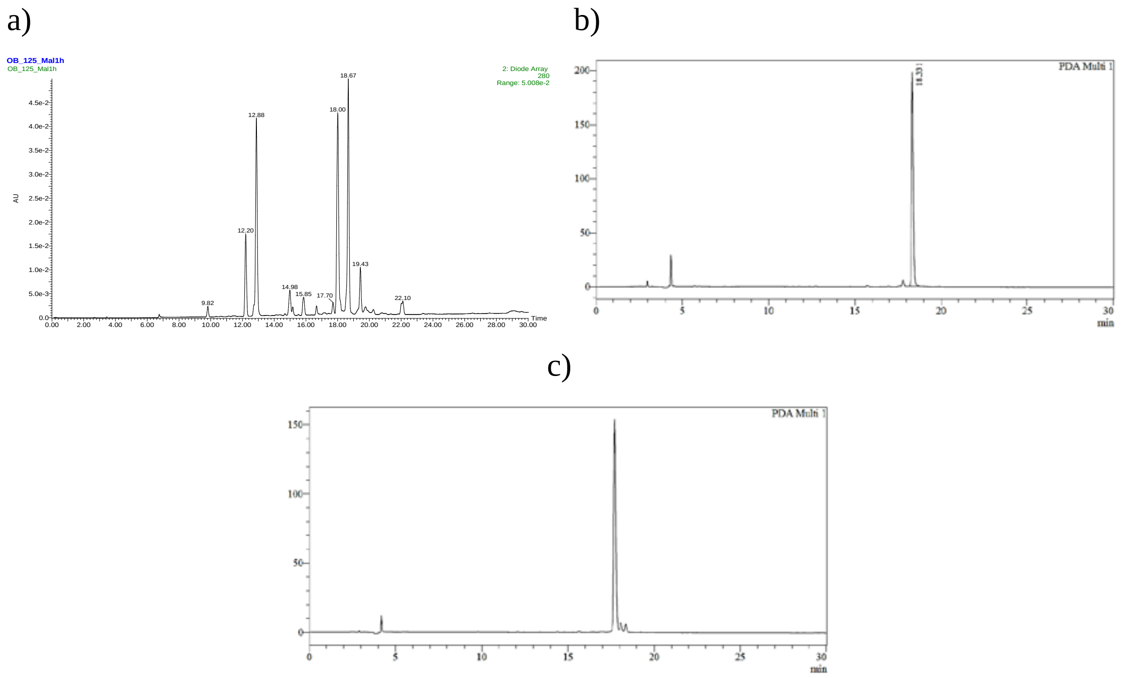


Figure S38: HPLC traces (280 nm) of the crude 1 h after the addition of 3-maleimidopropanoic acid (a, 18.0 and 18.7 min are the peaks corresponding to peptide+CPD), and purified double conjugate **29a** (b and c, each peak was independently collected, 280 nm). The peak at 13 min in the crude (a) is excess 3-maleimidopropanoic acid.

11.2 Synthesis of conjugates incorporating three different moieties

· Conjugation of H-CWGRGC-NH_2 (30) with CPD-biotin (3) and Mal-gtttcK-NH_2 (31) → conjugate 29b:

Peptide H-CWGRGC-NH_2 (**30**, 50 nmol) was dissolved in water (95 µL) and reacted with CPD-biotin **3** (4.46 µL of a 13.45 mM solution, 60 nmol) at 37 ºC (0.5 mM peptide concentration). The reaction was monitored after 2 h by HPLC-MS (analysis conditions, 0-50% B). Subsequently, Mal-gtttc-NH_2 (**31**, 50 µL of a 3 mM solution, 150 nmol) was added and the resulting mixture left reacting for an additional hour. The final conjugate (**29b**) was purified by HPLC (purification conditions, 0-50% B) and isolated in 29% yield.

Product characterization:
HPLC: analysis conditions, 0-50% B, t_R = 13.4 min and 13.9 min (Figure S39)
MALDI-TOF (positive mode, DHB): *m/z* (both products) 2747.6 $[M+H]^+$, M calcd. for $C_{115}H_{155}N_{43}O_{32}S_3$ (conjugate **29b**) 2746.1.

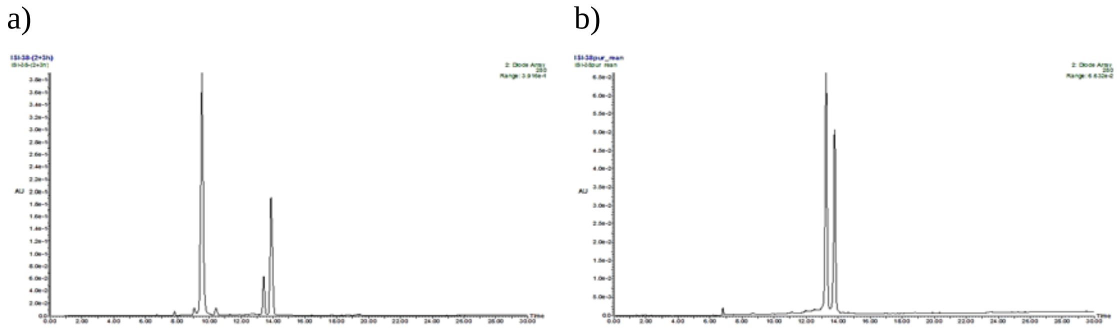


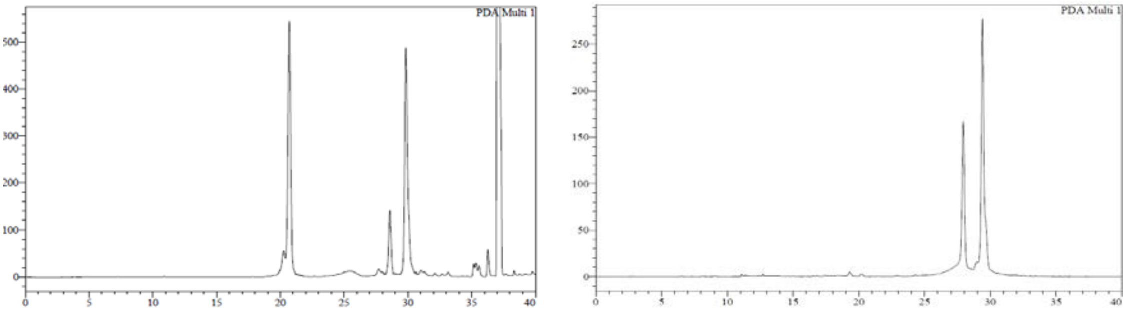
Figure S39. HPLC profiles (254 nm) of conjugate **29b**, crude (a) and purified (b).

· Conjugation of H-CWGRGC-NH_2 (30) with CPD-catgtcatK-NH_2 (24c) and 1-[2-(maleimido)ethyl]-4-[2-(3,4-dihydro-2H-1-benzopyran-6-yl)-5-oxazolyl]pyridinium triflate → conjugate 29c:

Peptide H-CWGRGC-NH2 (**30**, 50 nmol) was reacted with 1.2 equiv of CPD-catgtcatK-NH_2 (**24c**, 60 nmol) in water (100 µL) at 37 ºC (0.5 mM peptide concentration). The reaction was monitored after 2 h by HPLC-MS (analysis conditions, 0-30% B, 60 ºC). Subsequently, the maleimido-fluorophore (15 µL of a 10 mM solution, 150 nmol) was added and the mixture left reacting for an additional h.

The final conjugate (**29c**) was purified (purification conditions, 0-30% B) and isolated in 34% yield.

Product characterization:
HPLC: analysis conditions, 0-30% B, 60 ºC t_R = 28.5 min and 29.8 min (Figure S40)
ESI-HRMS (positive mode): *m/z* (both products) 3513.31 $[M+H]^+$, M calcd. for $C_{151}H_{187}N_{60}O_{38}S_2$ (conjugate **29c**) 3512.40.



a) b)

Figure S40. HPLC profiles (254 nm) of conjugate **29c**, crude (a) and purified (b).

12. Quantum chemical calculations

12.1 Calculations

The Gaussian09 package[9] was used for all calculations. The electronic structure was computed within Density Functional Theory (DFT),[10] using M06-2X functional[11] and an the 6-311++G** basis set to expand Kohn-Sham orbitals[12] . This functional has been proved to provide accurate kinetics and thermodynamics for thiol addition reactions to α,β-unsaturated ketones[13]. Furthermore, single point CCSD(T)/6-311++G** calculations on the M06-2X/6-311++G** optimized structures of the reactants and MTAs for both compound **1** and NMMal show that errors in the electronic energy are below 1 kcal/mol (0.2 kcal/mol for **1** and 0.6 kcal/mol for NMMal).

A SMD solvation model,[14] parametrized for M06 functionals, was used in all calculations (energy evaluations, optimizations and frequencies), considering water as solvent. An additional (explicit) water molecule was included in the calculations to stabilize the nascent charges along the reaction. Structure optimizations and transition state (TS) searches were performed with the Berny algorithm. Minimum and TS structures were characterized by exhibiting zero and one imaginary frequency, respectively. The connections between the minimum structures and the transition states were checked by analyzing the corresponding imaginary frequency modes.

12.2 Results: CPD+amino acid reactions

· H-Cys-OMe + CPD/Mal

To explain the different reactivity of CPDs vs. maleimides when forming MTAs, we compared the free energies (ΔG^{o}) of the two Michael-type reactions (ΔG^{o}) by means of DFT methods. The reaction of methylcysteinate with *N*-methylmaleimide (NMMal) turned out to be highly exothermic (ΔG^{o} = -10.6 kcal/mol), whereas the exothermicity for the reaction with CPDs was much lower (ΔG^{o} = -2.4 kcal/mol, Figure S41). Therefore, the relative stabilities of reagents and adducts in each case is in agreement with the greater experimental stability of thiol-maleimide MTAs with respect to thiol-CPD MTAs.

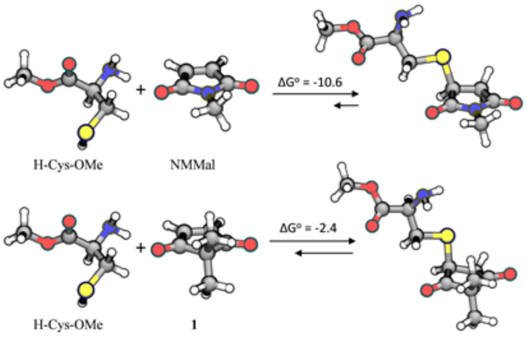


Figure S41. Optimized structures (M06-2X/6-311++G**, considering solvent effects) of the molecules investigated and free energy change, in kcal/mol, of the corresponding Michael-type reaction.

Analysis of the structures of NMMal and CPD **1** revealed that the C=C bond is slightly longer for **1** than for NMMal (1.34 Å vs. 1.33 Å, respectively) and the CO-CH bond is shorter in **1** compared to NMMal (1.49 Å vs. 1.50 Å), indicating that the electrons in the π system are more delocalized in the CPD than in NMMal. In fact, the values for NMMal are typical of a purely localized π-system such as cyclopentene. Consistently, in CPD **1** we found a significant charge transfer from the $\pi_{C=C}$ orbital to the $\pi^*_{C=O}$ (see orbital analysis below). Thus we concluded that the different π-electron delocalization in the Michael acceptor accounts for the higher ΔG^o value in the thiol-NMMal reaction.

· H-Cys-OH + CPD

Formation of the M-20 Da adduct. To further assess the mechanism proposed for the MTA to M-20 Da adduct transformation, we computed the free energy profile of the first two reaction steps for both systems (H-Cys-OH and H-Cys-OMe). Calculations showed (Figure S42, orange and blue profiles, respectively) that formation of the intermediate resulting from nucleophilic attack (first step of the reaction) turned out to be more favoured H-Cys-OH+**1** MTA (orange profile) than for H-Cys-OMe+**1** MTA (blue profile), clearly supporting that the **M**-18 Da adduct is the precursor of the **M**-20 Da one in both cases.

· H-Hcy-OH + CPD

DFT calculations confirm that the first energy barrier (that corresponding to the reaction between the amine and one of the CPD carbonyl groups) is higher for Hcy than for H-Cys-OH or H-Cys-OMe (Figure S42, green profile). Interestingly, formation of the second intermediate (that resulting from nucleophilic attack and proton transfer) is also less favored for Hcy. This likely explains why cyclization does not occur.

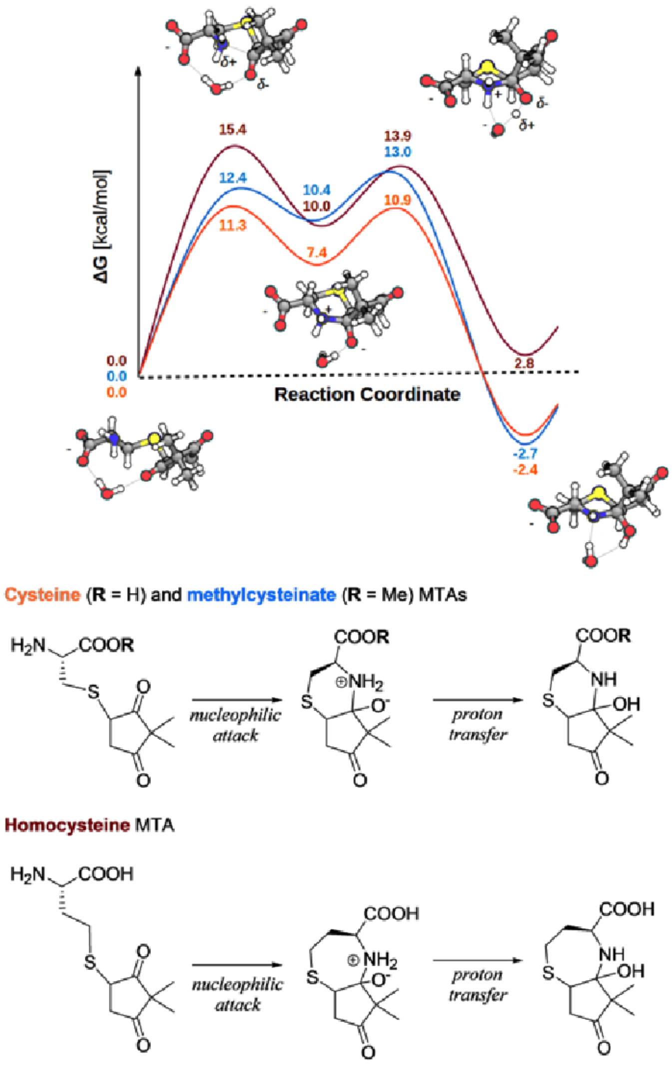


Figure S42. Free energy profiles (M06-2X/6-311++G**/SMD) along the first steps of the cyclization reaction between the amine and the closest carbonyl group of CPD **1**. Cysteine (orange), methyl cysteinate (blue) and homocysteine (green) were studied. The cysteine and homocysteine carboxylic groups were taken as deprotonated. One water molecule was explicitely considered to mediate the proton transfer in the second step.

13. Orbital analysis

A natural bond order (NBO) analysis[15] was performed to rationalize the greater stability of NMMal MTAs with respect to compound **1** (representative of CPD MTAs). The energetic stabilization associated with electron density delocalization between Lewis-type and non-Lewis NBOs can be taken into account by 2nd-order perturbation theory.[16] The results (Table S1) show that the delocalization of the electronic density between the donor π(C=C) orbital and the acceptor π*(C=O) orbital, measured as the difference of the orbital energies, is more pronounced (by ~ 4 kcal/mol) in compound **1** (24.6 kcal/mol) than NMMal (20.8 kcal/mol). The lower delocalization of electrons in NMMal in comparison to **1** arises from the higher contribution of the donor nitrogen lone pair (LP) to the acceptor π*(C=O) orbitals (65.7 kcal/mol) – note NMMal can be regarded as two independent π systems, the localized π(C=C) and the delocalized LP(N)-π*(C=O) –. In other words, the π(C=C) orbital of **1** is delocalized through the two C=O bonds, making it less reactive, while the π(C=C) orbital of NMMal is substantially more localized, thus more reactive. It is worth noting that the π(C=C) orbital of NMMal cannot delocalize throughout the whole ring (according to Hückel rules,[17] this is a 4n-π antiaromatic system).

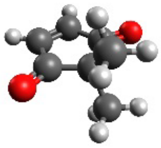
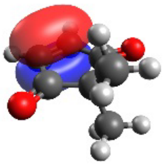
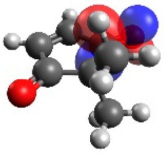
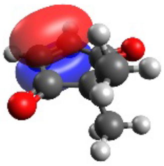
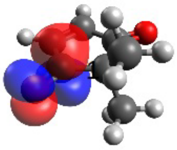
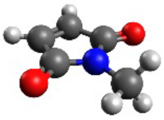
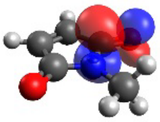
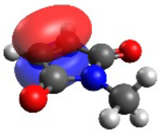
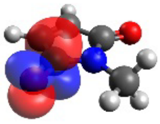
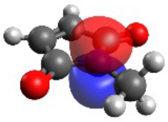
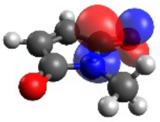
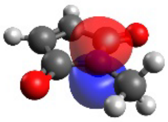
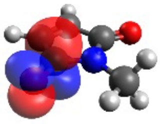
Molecule	Donor orbital	Acceptor orbital	Stabilization Energy (kcal/mol)
 **1**	 BD(C_2-C_3)	 BD*(C_1-O_1)	24.6
	 BD(C_2-C_3)	 BD*(C_4-O_2)	24.6
 **NMMal**	 BD(C_2-C_3)	 BD*(C_1-O_1)	20.8
	 BD(C_2-C_3)	 BD*(C_4-O_2)	20.5
	 LP(N)	 BD*(C_1-O_1)	65.6
	 LP(N)	 BD*(C_4-O_2)	65.7

Table S7. Second order perturbation theory analysis of the Fock matrix for compound **1** and NMMal. BD stands for bonding orbital, BD* for antibonding orbital and LP for a lone pair orbital.

Interestingly, the C=C bond distance in NMMal is identical to the C=C bond of cyclopentene (1.33 Å, Figure S43), a molecule with a fully localized π(C=C) orbital. In contrast, the C=C bond in compound **1** is 0.01 Å longer (1.34 Å), being more similar to the C=C bond in cyclopentadiene (1.45 Å, Figure S43), a molecule with a delocalized π(C=C) orbital. That the distances of the two vicinal C-C bonds of NMMal and compound **1** follow the same trend as cyclopentene and cyclopentadiene is another indication of the larger delocalization of the π(C=C) orbital in compound **1** with respect to NMMal

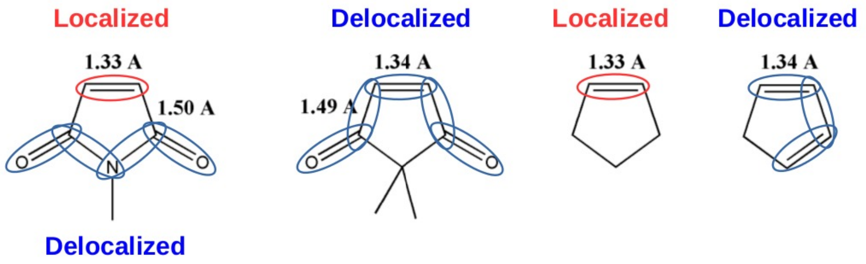


Figure S43. Bond distances of the optimized structures (M06-2X/6-311++G**) of NMMal, compound **1**, cyclopentene and cyclopentadiene.

14. Abbreviations

ACN = acetonitrile
Boc = *tert*-butoxycarbonyl
Bhoc = benzhydryloxycarbonyl
CNE = 2-cyanoethyl
COMU = 1-[(1-(cyano-2-ethoxy-2-oxoethylideneaminooxy)-dimethylamino-morpholino-methylene)] methanaminium hexafluorophosphate
CPD = 2,2-disubstituted cyclopent-4-ene-1,3-dione
DCM = dichloromethane
DHB = 2,5-dihydroxybenzoic acid
DIPC = *N*,*N'*-diisopropylcarbodiimide
DIPEA = *N*,*N*-diisopropylethylamine
DMF = *N*,*N*.dimethylformamide
DTT = dithiothreitol
EDT = 1,2-ethanedithiol
ESI = electrospray ionization
Fmoc = 9-Fluorenylmethoxycarbonyl
Hcy = homocysteine
HOBt = 1-hydroxybenzotriazole
HPLC = high performance liquid chromatography
HRMS = high resolution mass spectrometry
MALDI-TOF = matrix-assisted laser desorption ionization - time of flight
MBHA = *p*-methylbenzhydrylamine
MTA = Michael-type adduct
NMM = *N*-methylmorpholine
NMMal = *N*-methylmaleimide
NMP = *N*-methylpyrrolidone
PNA = peptide nucleic acid
TCEP = tris(2-carboxyethyl)phosphine
TFA = trifluoroacetic acid
THAP = 2,4,6-trihydroxyacetophenone
TIS = triisopropylsilane

15. References

[1] X. Elduque, E. Pedroso and A. Grandas, *J. Org. Chem.* **2014**, *79*, 2843-2853.

[2] E. Kaiser, R. L. Colescot, C. D. Bossinge and P. I. Cook *Anal. Biochem.* **1970**, *34*, 595-598.

[3] IUPAC-IUB Joint Commission on Biochemical Nomenclature for Symbolism for Amino Acids and Peptides, *Eur. J. Biochem.*, 1984, **138**, 9–37.

[4] O. Brun, X. Elduque, E. Pedroso and A. Grandas, *J. Org. Chem.* **2015**, *80*, 6093-6101

[5] W. C. Agosta and A. B. Smith, III, *J. Org. Chem.* **1970**, 35, 3856-3860.

[6] W. Kreiser, A. Wiggermann, A. Krief and D. Swinnen *Tetrahedron Lett.* **1996**, 37, 7119-7122.

[7] D. A. Stetsenko, M. J. Gait, *J. Org. Chem.* **2000**, 65, 4900-4908.

[8] D. A. Stetsenko, M. J. Gait, *Nucleosides, Nucleotides & Nucleic Acids*, **2000**, 19, 1751-1764.

[9] Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, **2009**.

[10] P. Hohenberg and W. Kohn, *Phys. Rev.*, **1964**, *136*, B864-B71.

[11] Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, **2008**, *120*, 215-41.

[12] W. Kohn and L. J. Sham, *Phys. Rev.*, **1965**, *140*, A1133-A1138.

[13] E. H. Krenske, R. C. Petter, Z. Zhu and K. N. Houk, *J. Org. Chem.* **2011**, *76*, 5074-5081.

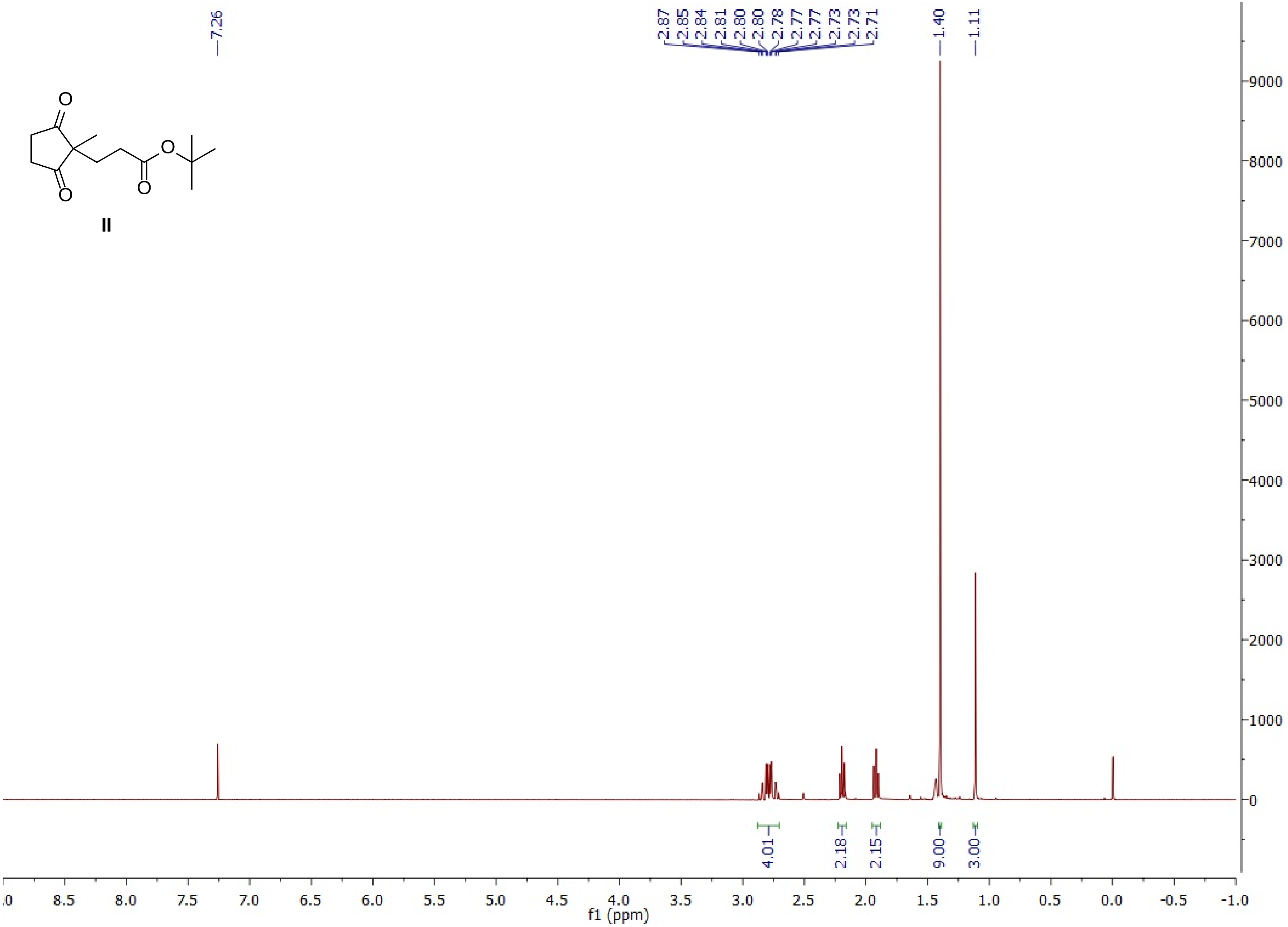
[14] A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B,* **2009**, *113*, 6378-96.

[15] A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.,* **1985**, *83*,735-46.

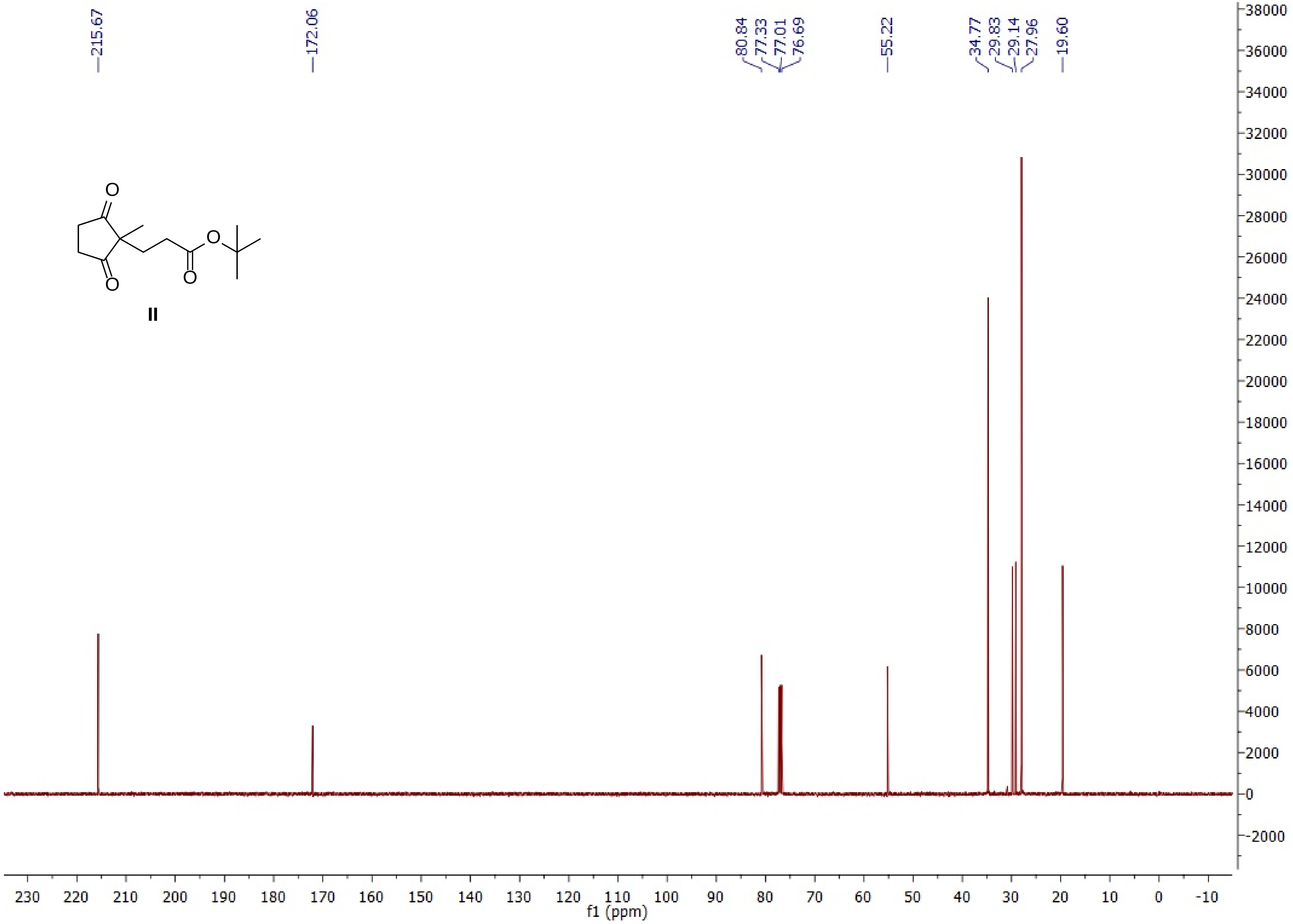
[16] A. E. Reed, L. A. Curtiss, and F. Weinhold, *Chem. Rev.*, **1988**, *88*, 899.

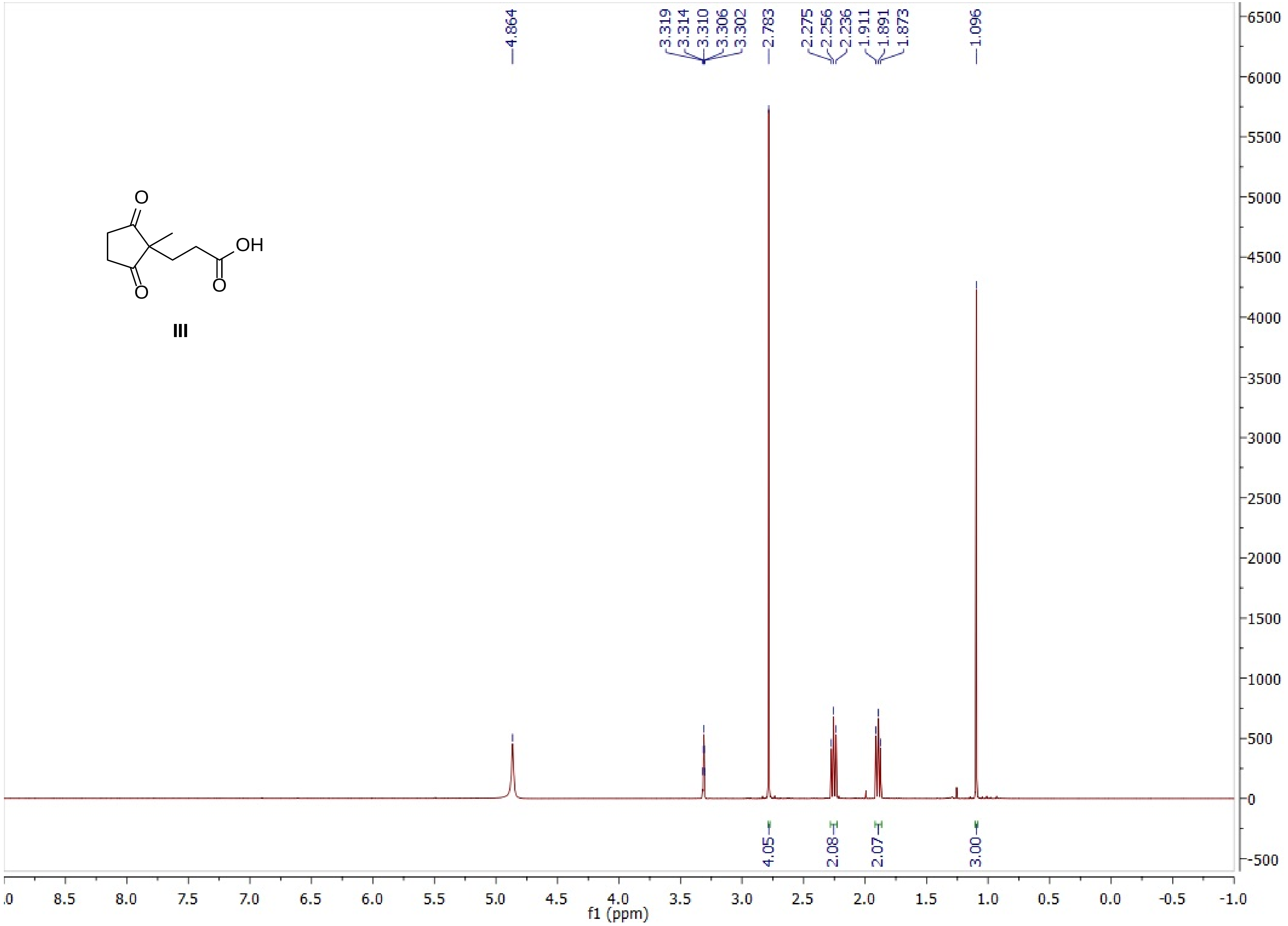
[17] IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: http://goldbook.iupac.org (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8

16. Spectra

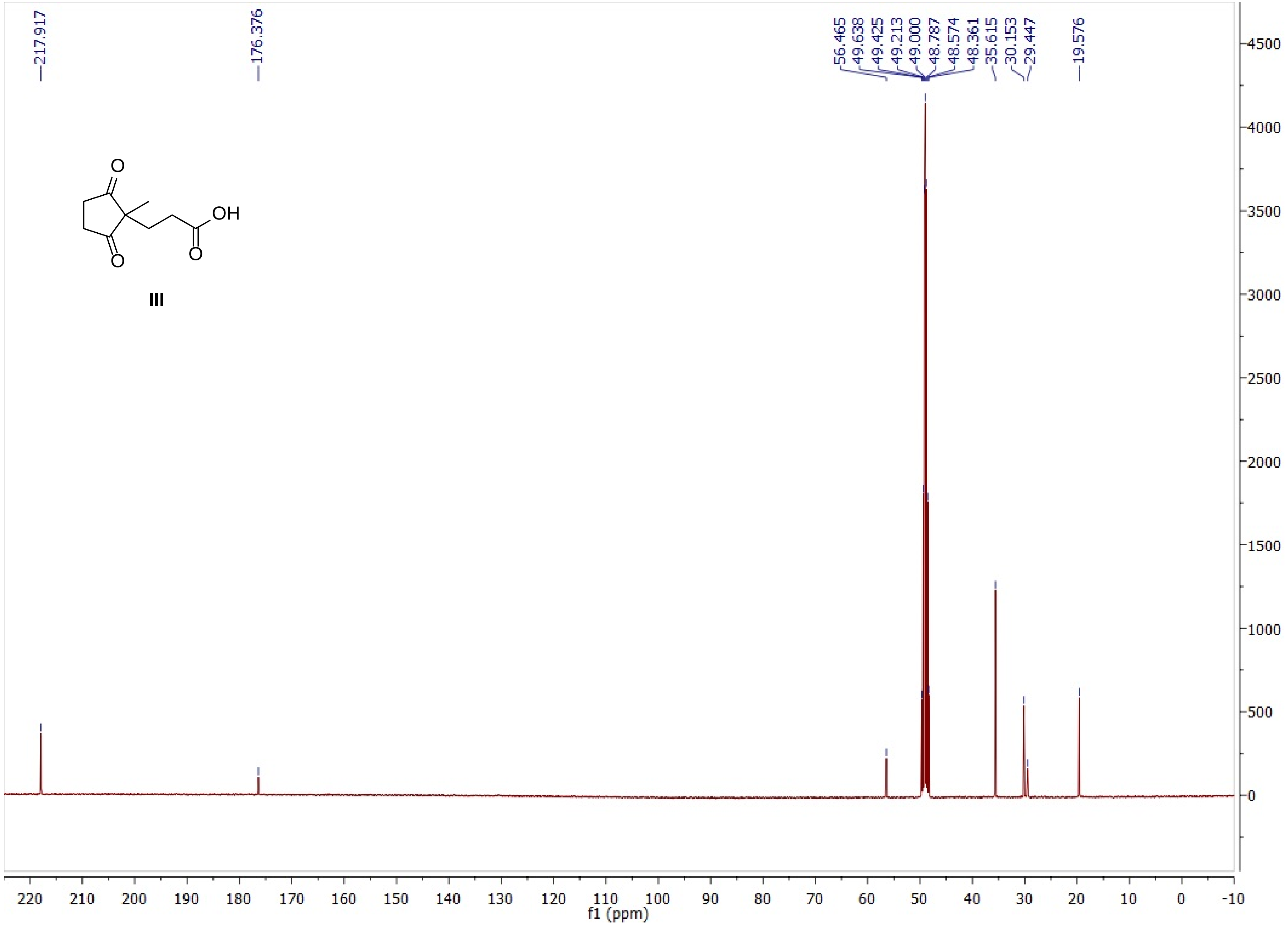


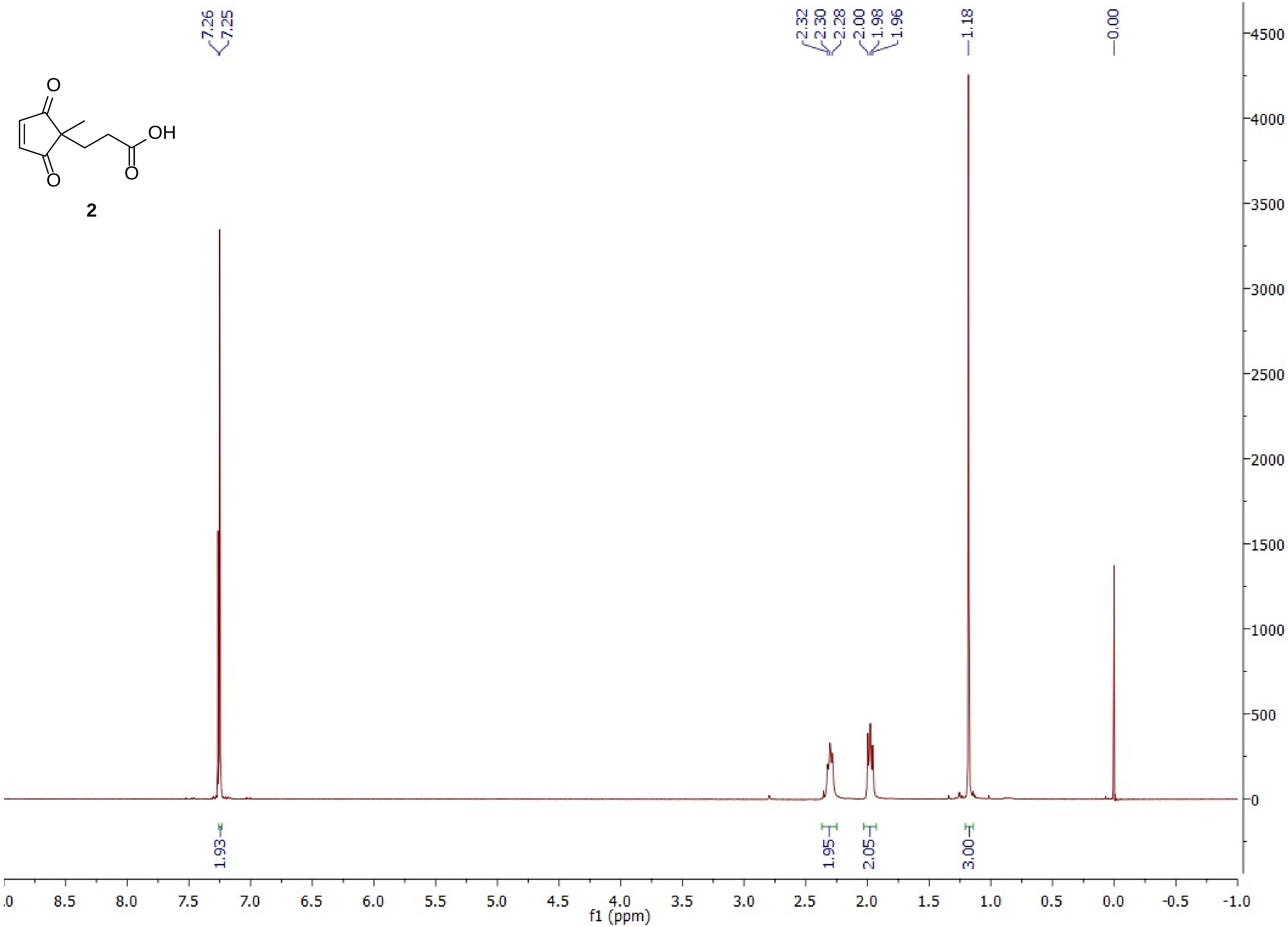
II

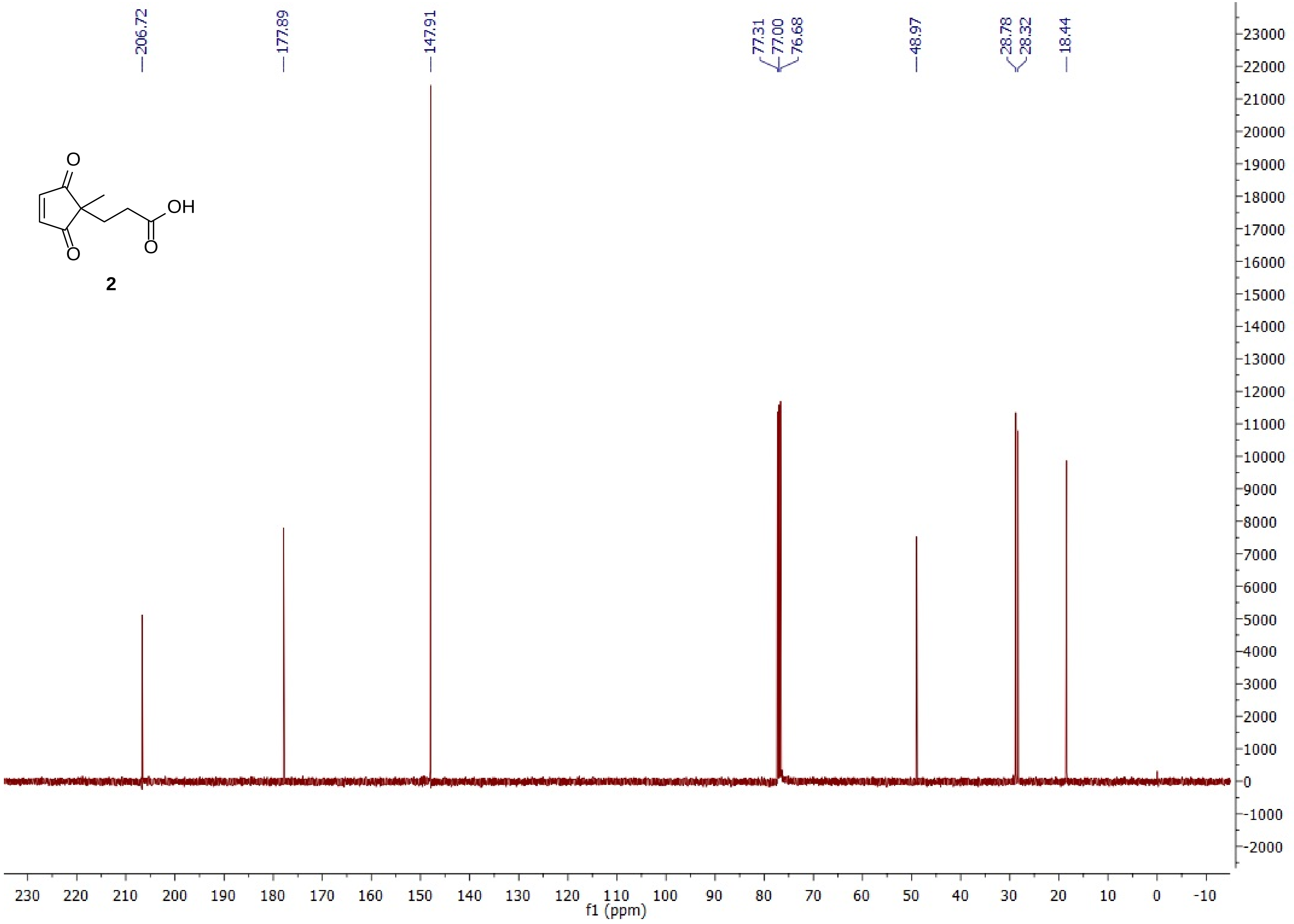

215.67
172.06
80.84
77.33
77.01
76.69
55.22
34.77
29.83
29.14
27.96
19.60
38000
36000
34000
32000
30000
28000
26000
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
-2000
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)
O
O
O
O
II

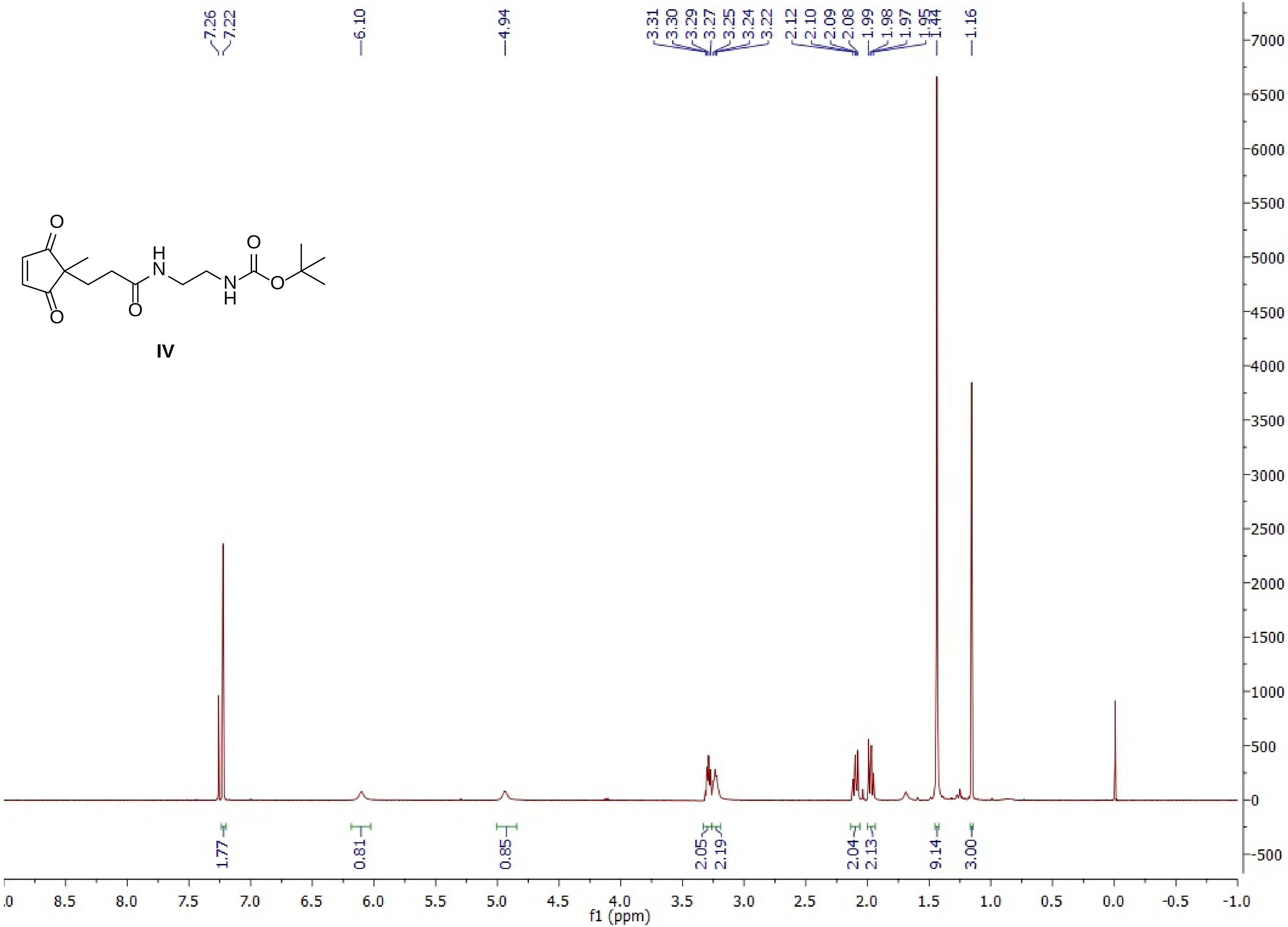

4.864
3.319
3.314
3.310
3.306
3.302
2.783
2.275
2.256
2.236
1.911
1.891
1.873
1.096
O
OH
O
O
4.05
2.08
2.07
3.00
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)

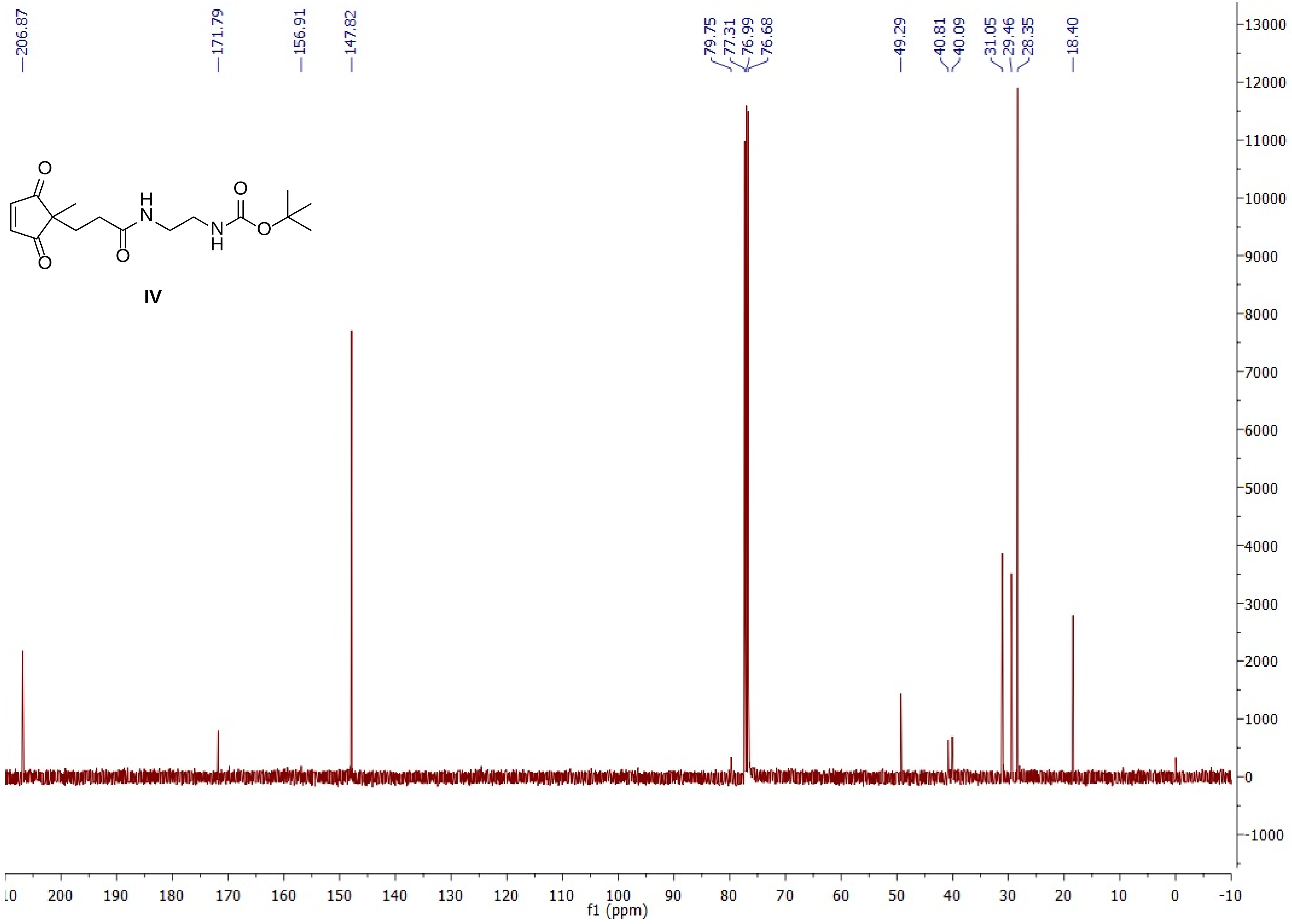
III

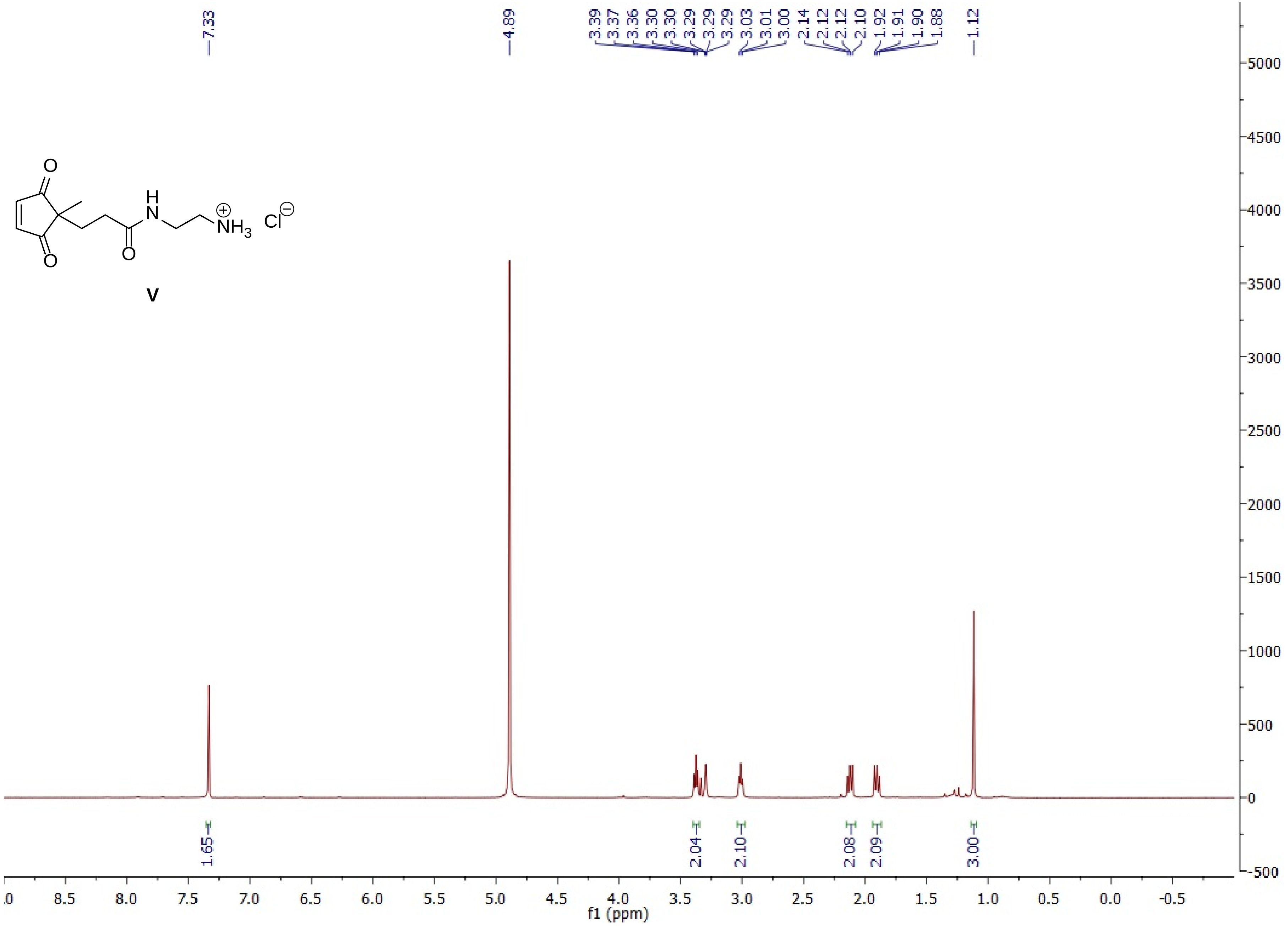

217.917
176.376
56.465
49.638
49.425
49.213
49.000
48.787
48.574
48.361
35.615
30.153
29.447
19.576
O
OH
O
O
III
4500
4000
3500
3000
2500
2000
1500
1000
500
0
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

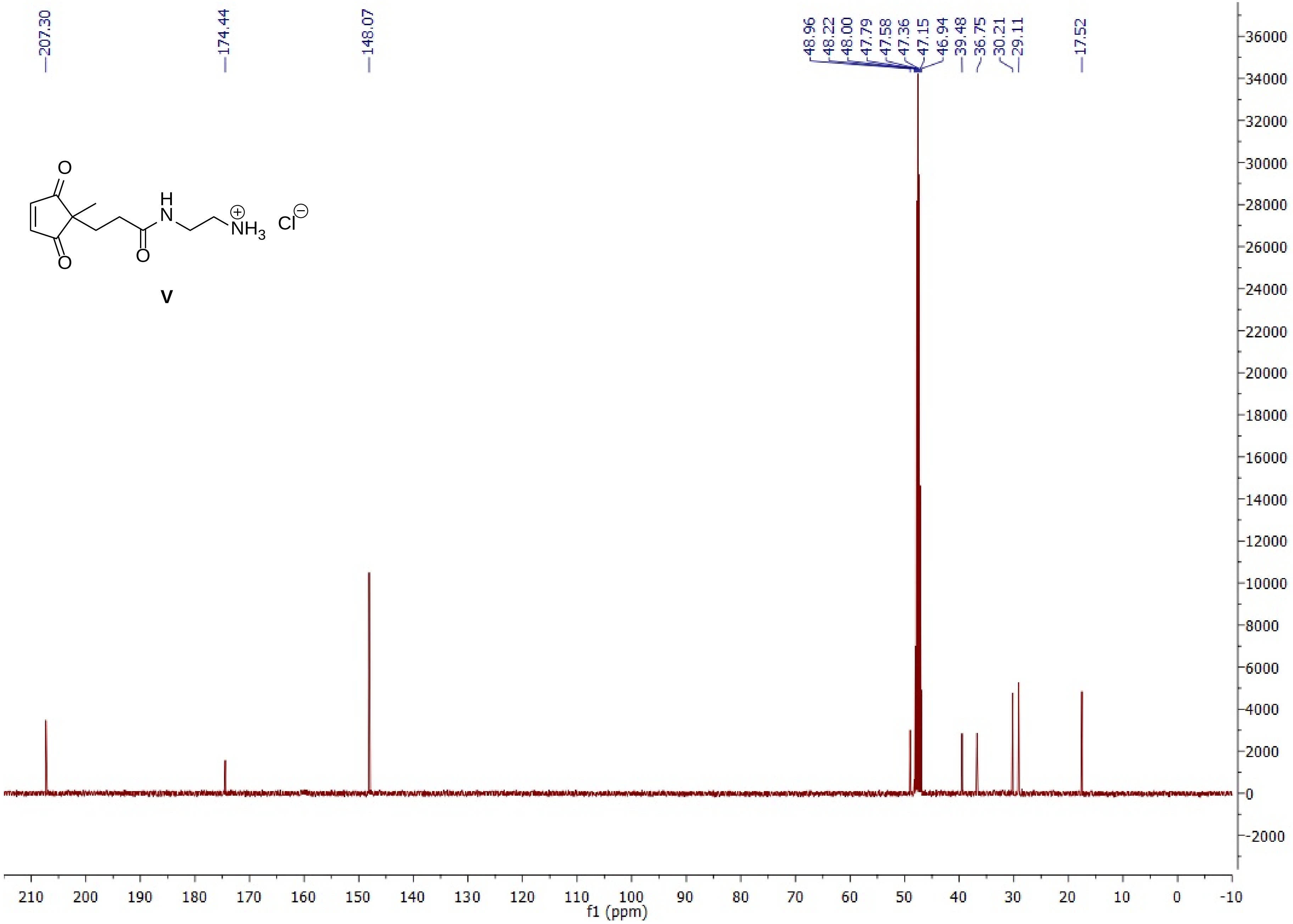

7.26
7.25
2.32
2.30
2.28
2.00
1.98
1.96
1.18
0.00
O
OH
O
O
2
1.93
1.95
2.05
3.00
4500
4000
3500
3000
2500
2000
1500
1000
500
0
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)

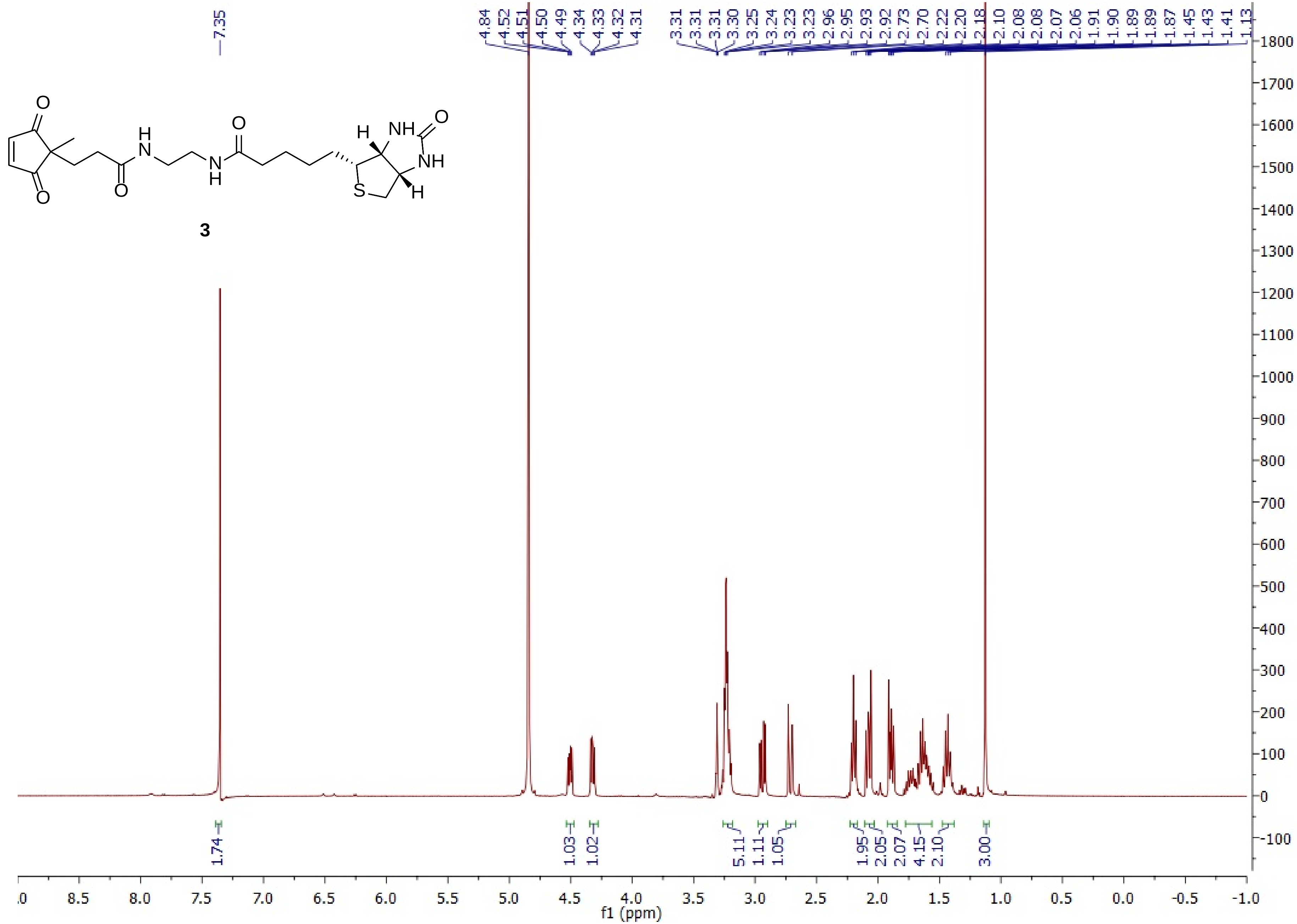

206.72
177.89
147.91
77.31
77.00
76.68
48.97
28.78
28.32
18.44
O
OH
O
O
2
23000
22000
21000
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
-2000
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

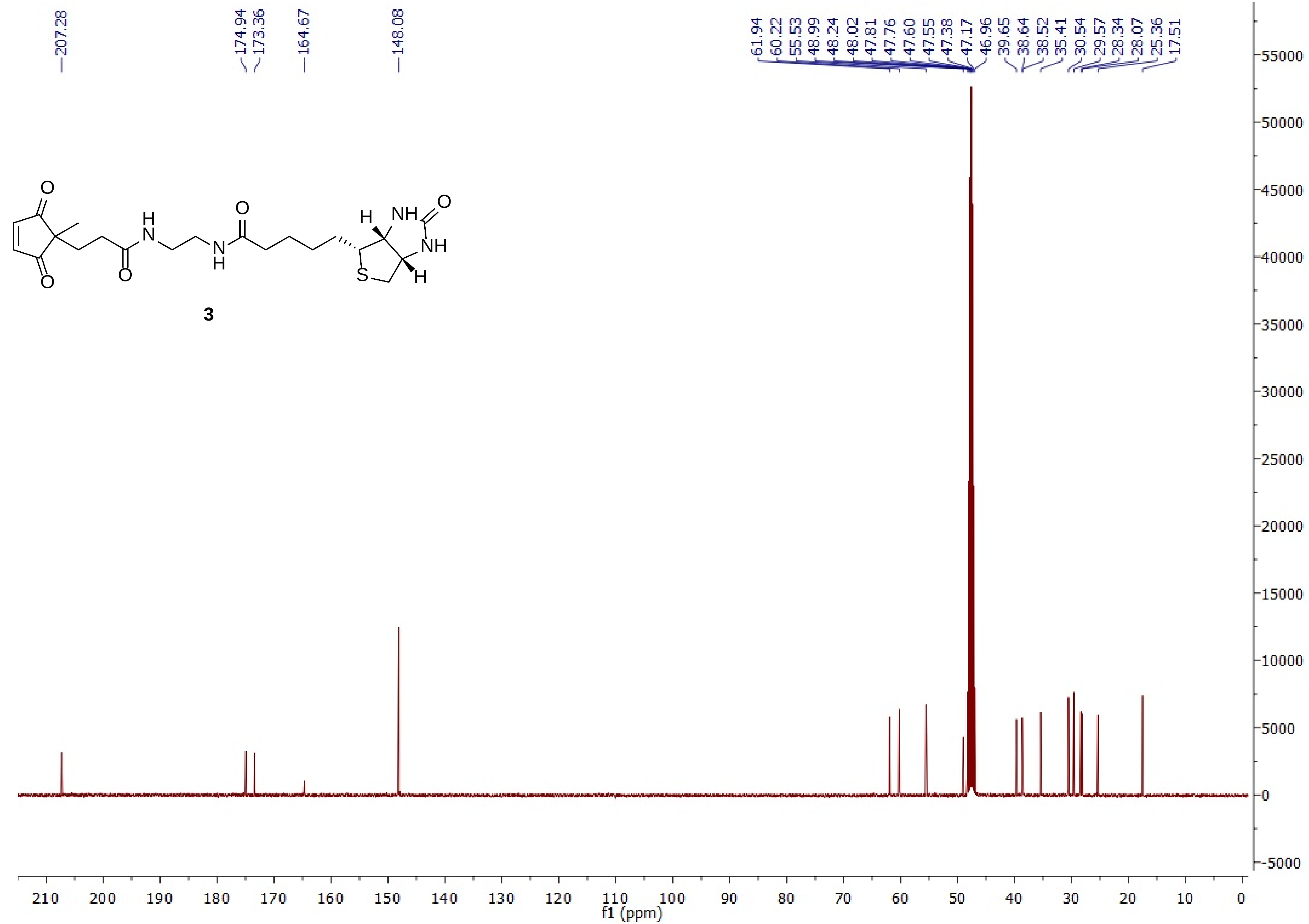

7.26
7.22
6.10
4.94
3.31
3.30
3.29
3.27
3.25
3.24
3.22
2.12
2.10
2.09
2.08
1.99
1.98
1.97
1.95
1.44
1.16
IV
1.77
0.81
0.85
2.05
2.19
2.04
2.13
9.14
3.00
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500

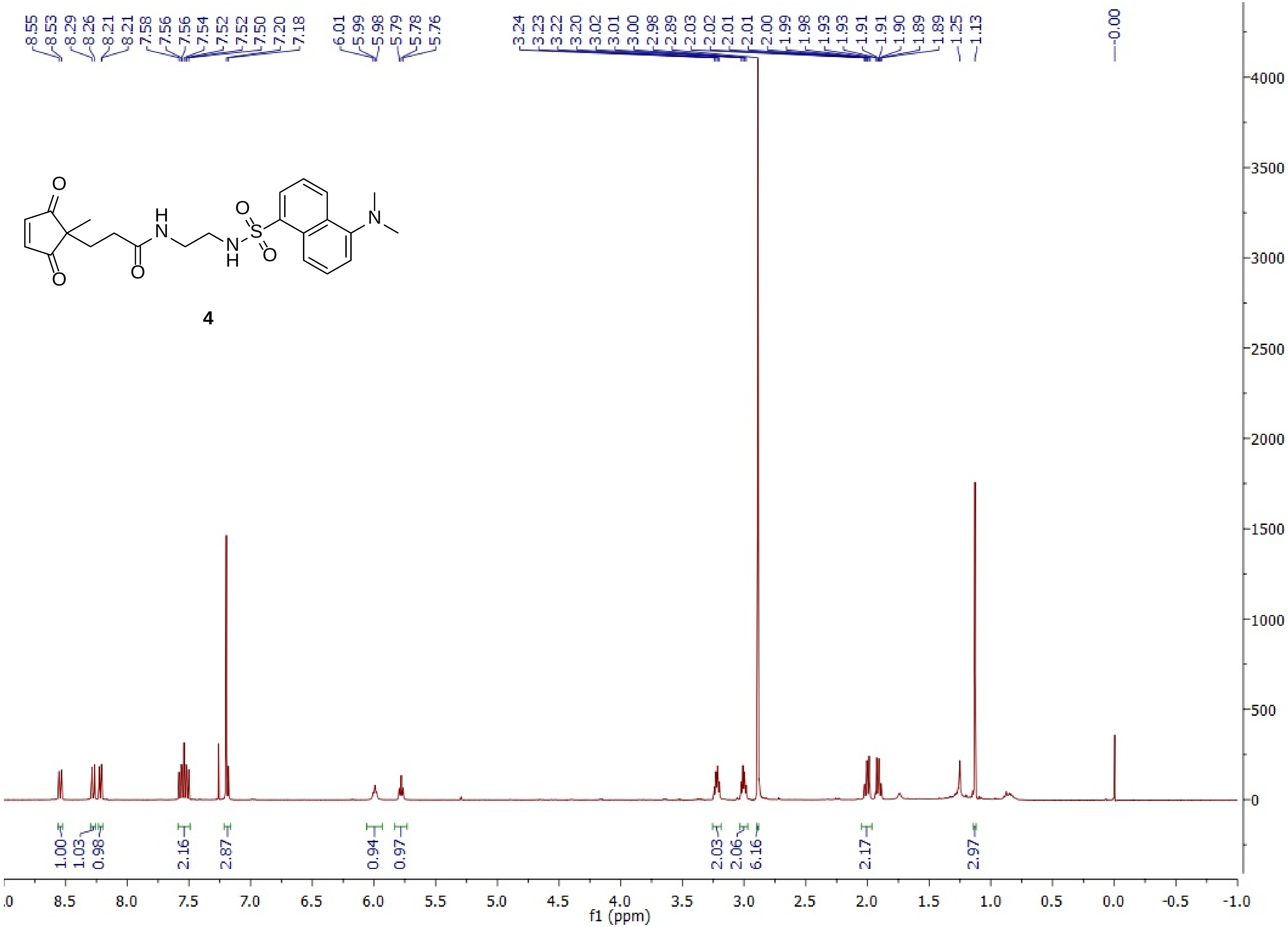

206.87
171.79
156.91
147.82
79.75
77.31
76.99
76.68
49.29
40.81
40.09
31.05
29.46
28.35
18.40
O
O
H
N
O
N
H
O
O
IV
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
10
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

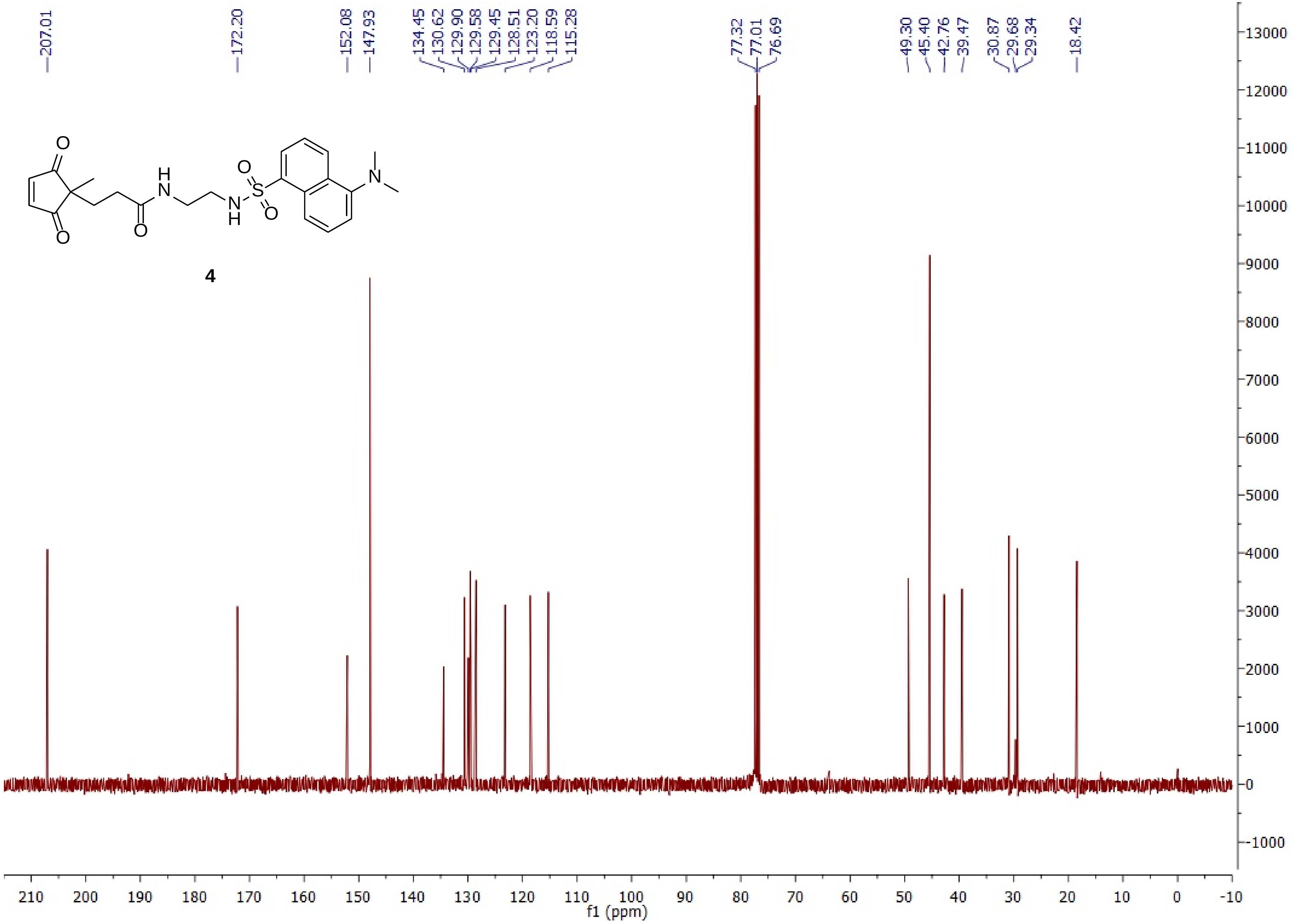

7.33
4.89
3.39
3.37
3.36
3.30
3.30
3.29
3.29
3.29
3.03
3.01
3.00
2.14
2.12
2.12
2.10
1.92
1.91
1.90
1.88
1.12
O
H
N
NH3
Cl
O
O
V
1.65
2.04
2.10
2.08
2.09
3.00
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500

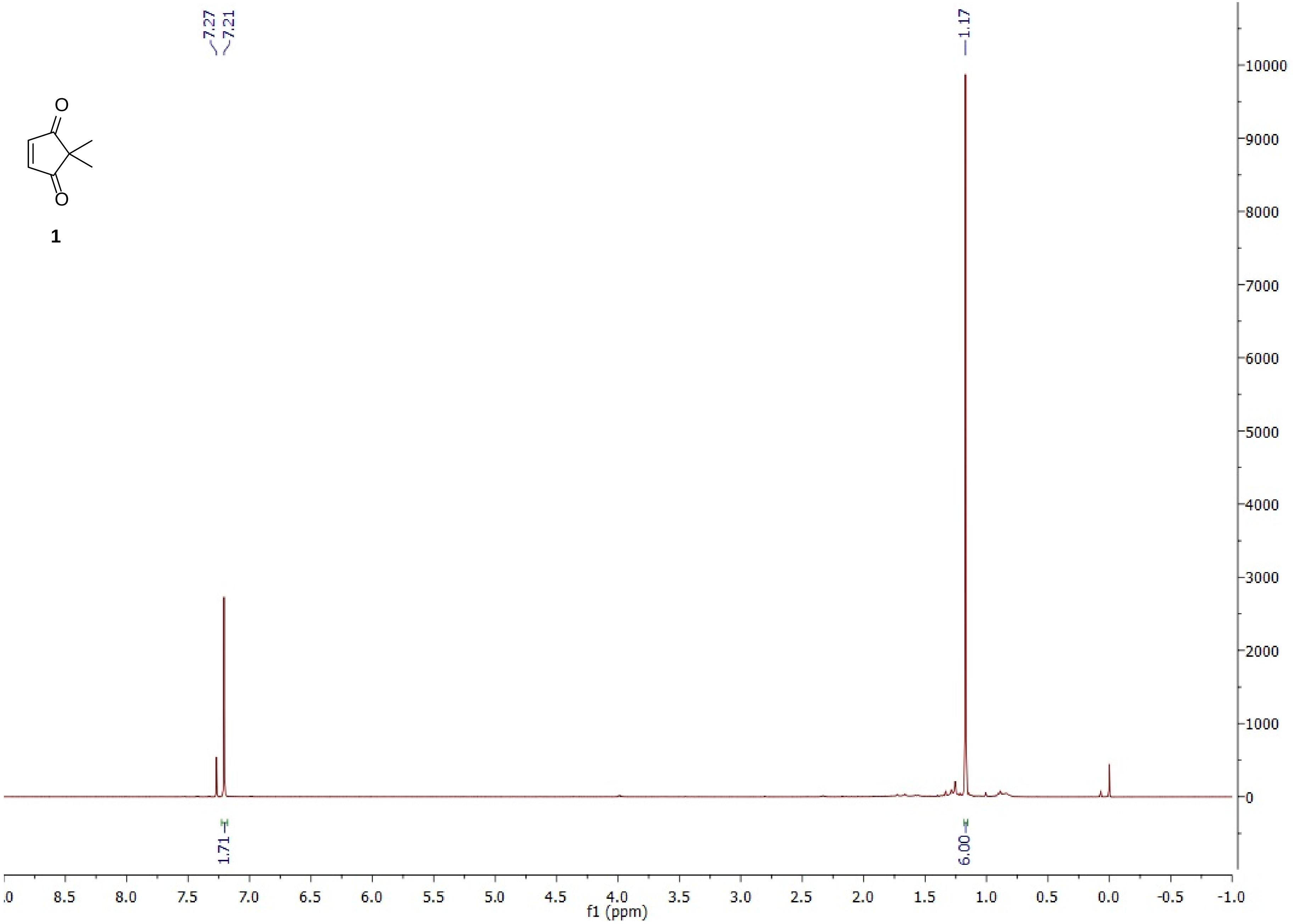

207.30
174.44
148.07
48.96
48.22
48.00
47.79
47.58
47.36
47.15
46.94
39.48
36.75
30.21
29.11
17.52
O
O
O
H
N
NH3
Cl
V
36000
34000
32000
30000
28000
26000
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
-2000
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

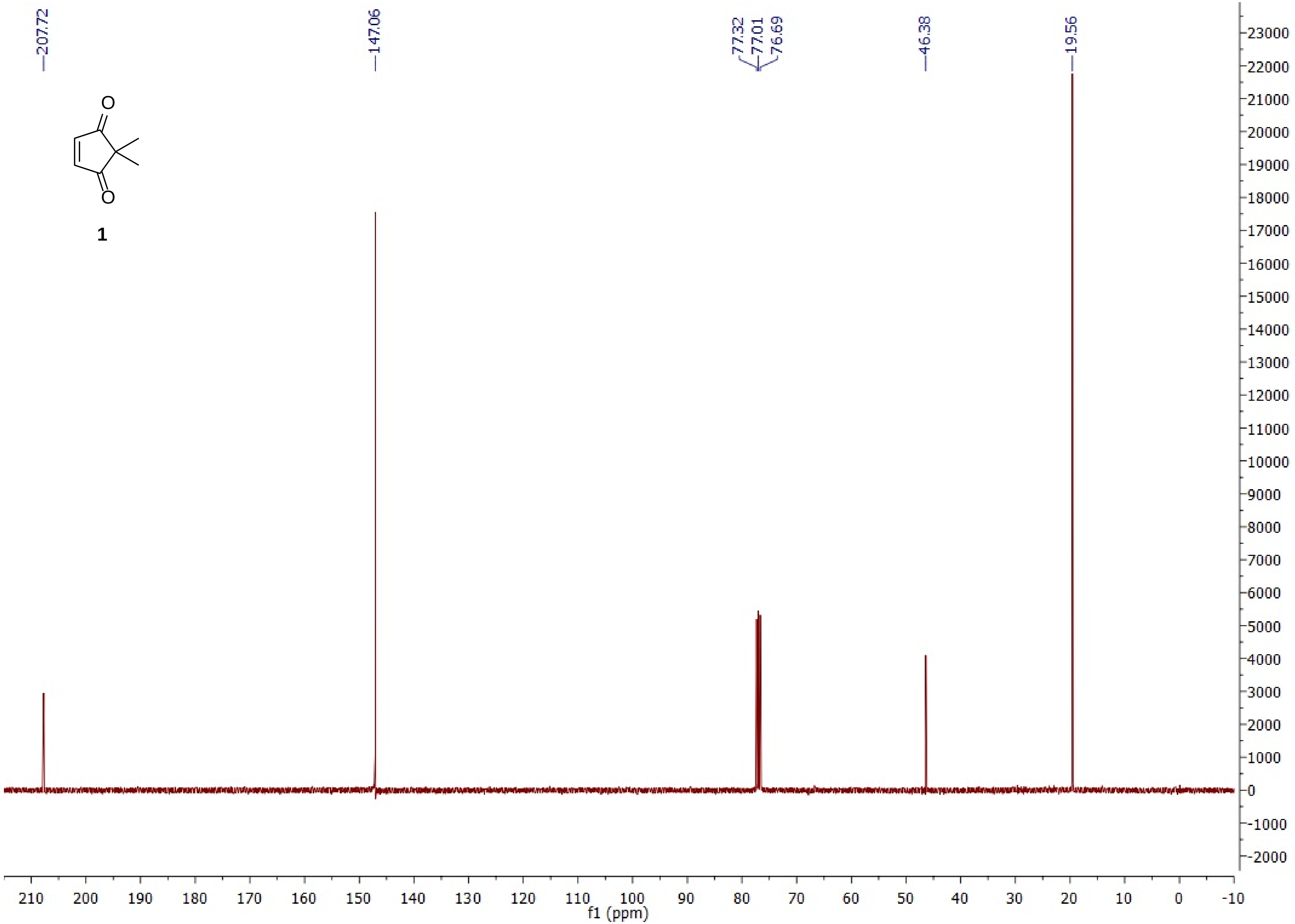

3
7.35
4.84
4.52
4.51
4.50
4.49
4.34
4.33
4.32
4.31
3.31
3.31
3.31
3.30
3.25
3.24
3.23
3.23
2.96
2.95
2.93
2.92
2.73
2.70
2.22
2.20
2.18
2.10
2.08
2.08
2.07
2.06
1.91
1.90
1.89
1.89
1.87
1.45
1.43
1.41
1.13
1.74
1.03
1.02
5.11
1.11
1.05
1.95
2.05
2.07
4.15
2.10
3.00
f1 (ppm)
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100


207.28
174.94
173.36
164.67
148.08
61.94
60.22
55.53
48.99
48.24
48.02
47.81
47.76
47.60
47.55
47.38
47.17
46.96
39.65
38.64
38.52
35.41
30.54
29.57
28.34
28.07
25.36
17.51
O
O
O
O
O
H
N
N
H
H
NH
NH
H
S
3
55000
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
-5000
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

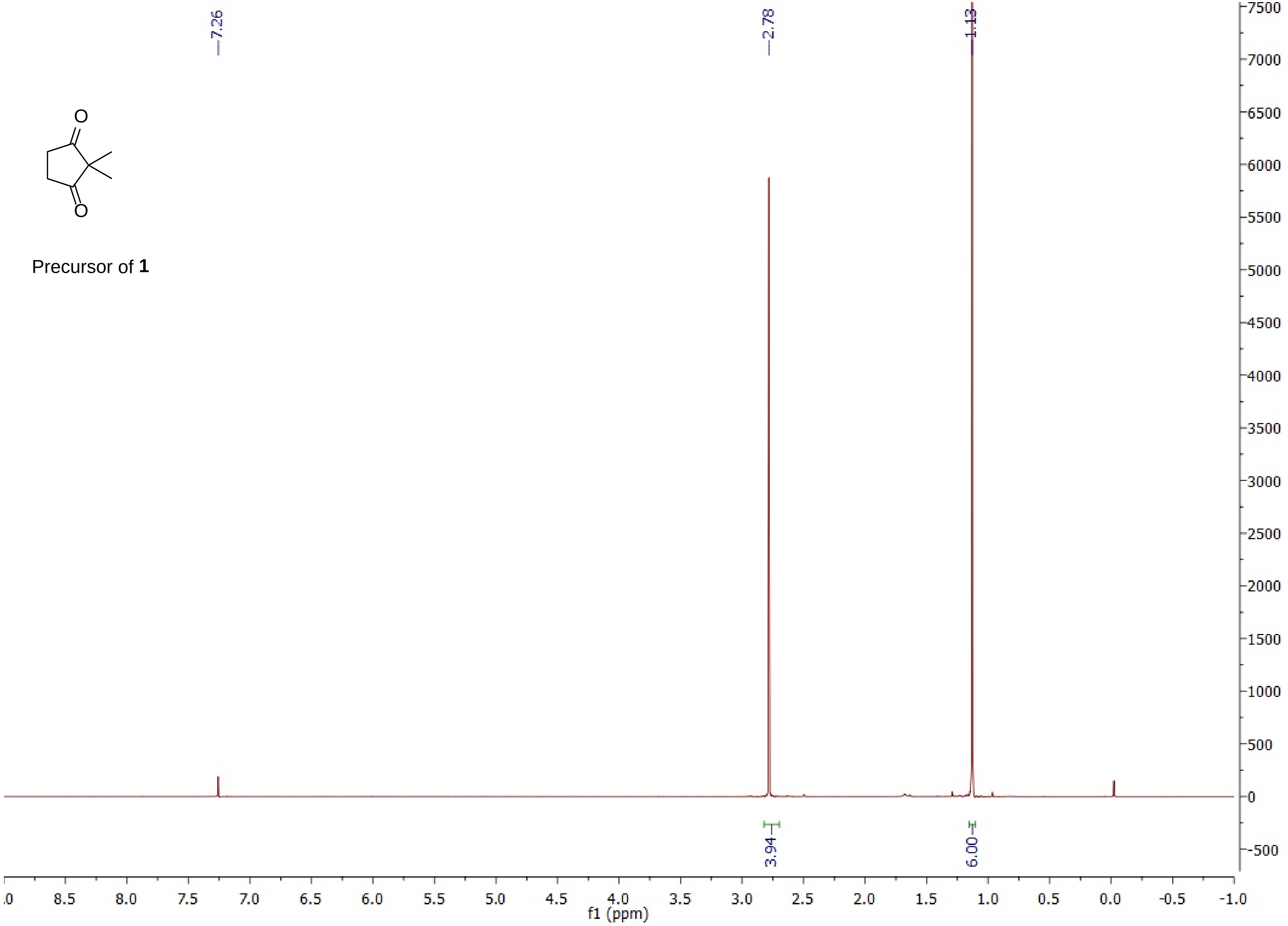

8.55
8.53
8.29
8.26
8.21
8.21
7.58
7.56
7.56
7.54
7.52
7.52
7.50
7.20
7.18
6.01
5.99
5.98
5.79
5.78
5.76
3.24
3.23
3.22
3.20
3.02
3.01
3.00
2.98
2.89
2.03
2.02
2.01
2.01
2.00
1.99
1.98
1.93
1.93
1.91
1.91
1.90
1.89
1.89
1.25
1.13
0.00
4
1.00
1.03
0.98
2.16
2.87
0.94
0.97
2.03
2.06
6.16
2.17
2.97
4000
3500
3000
2500
2000
1500
1000
500
0
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)


207.01
172.20
152.08
147.93
134.45
130.62
129.90
129.58
129.45
128.51
123.20
118.59
115.28
77.32
77.01
76.69
49.30
45.40
42.76
39.47
30.87
29.68
29.34
18.42
4
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

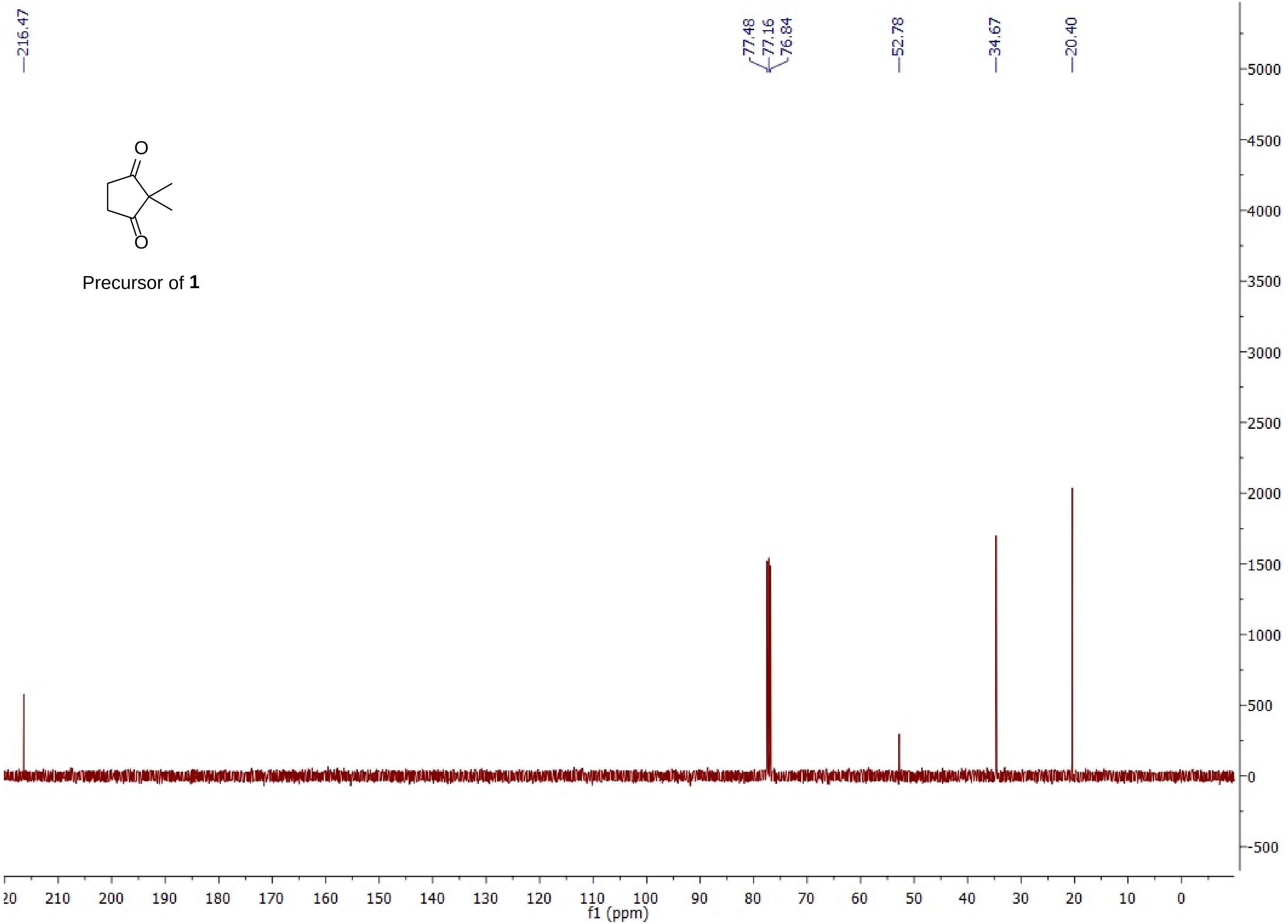

O
O
1
7.27
7.21
1.17
1.71
6.00
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)


1

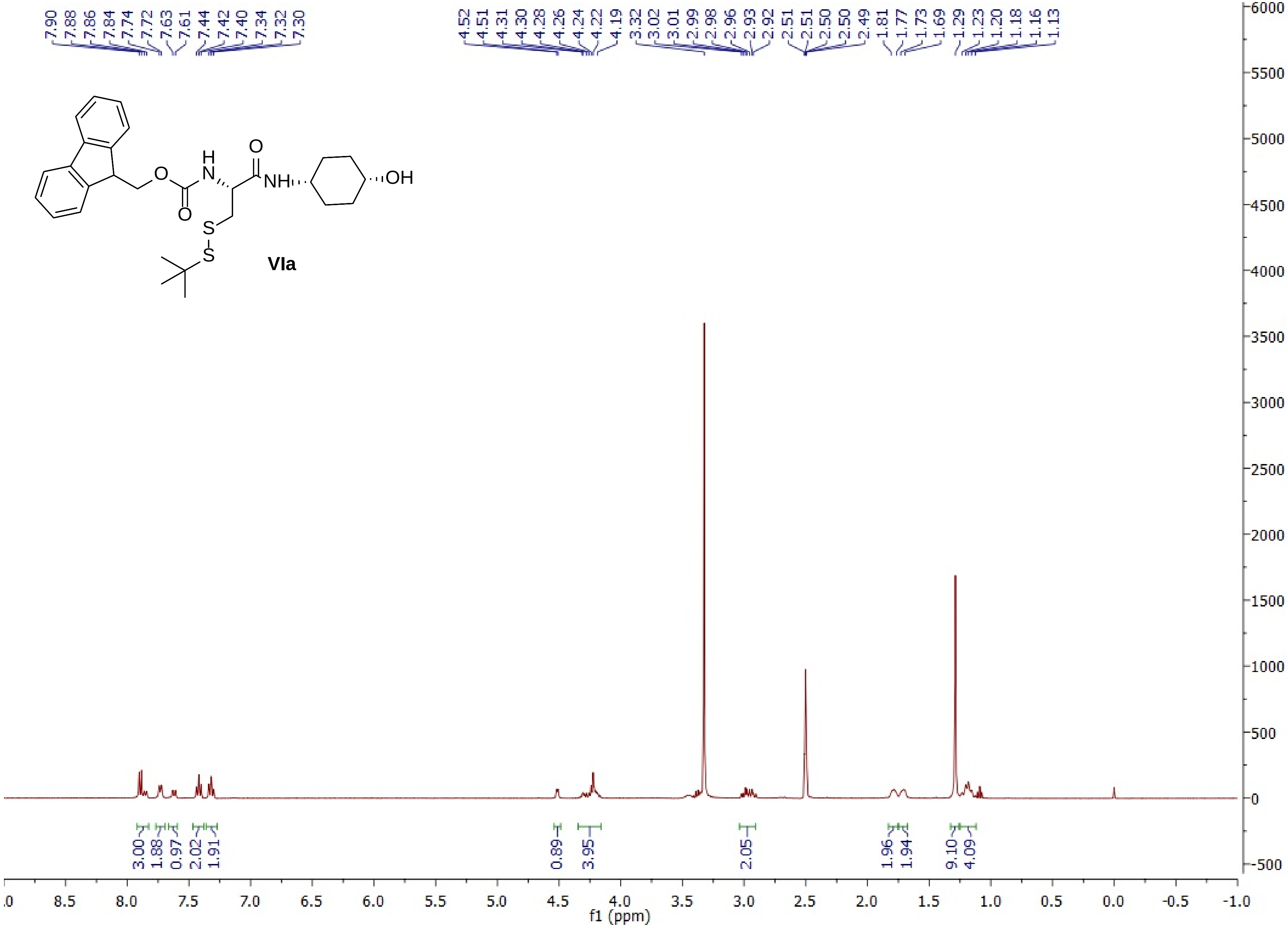
207.72
147.06
77.32
77.01
76.69
46.38
19.56
O
O
23000
22000
21000
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
-2000
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)



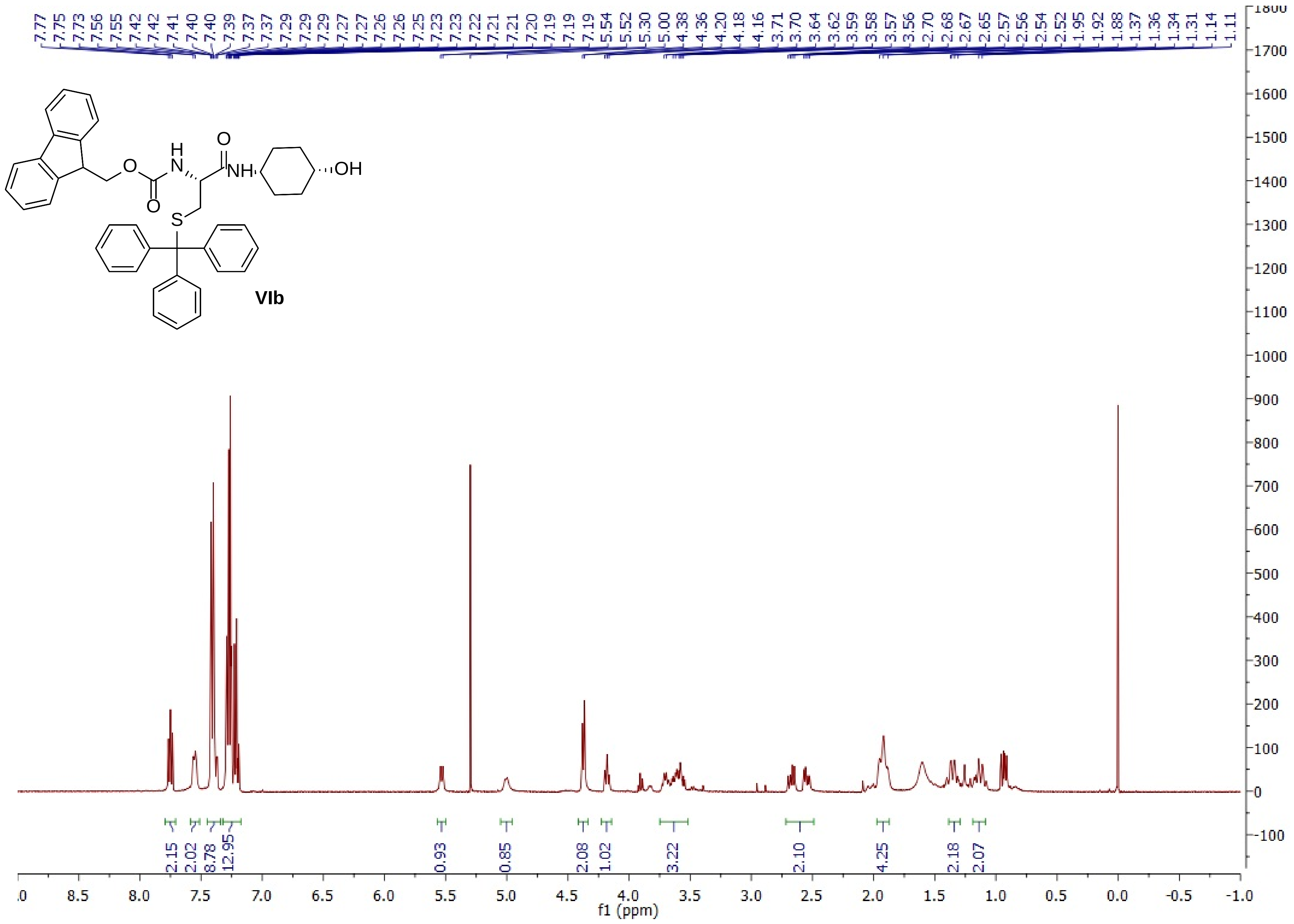
Precursor of **1**

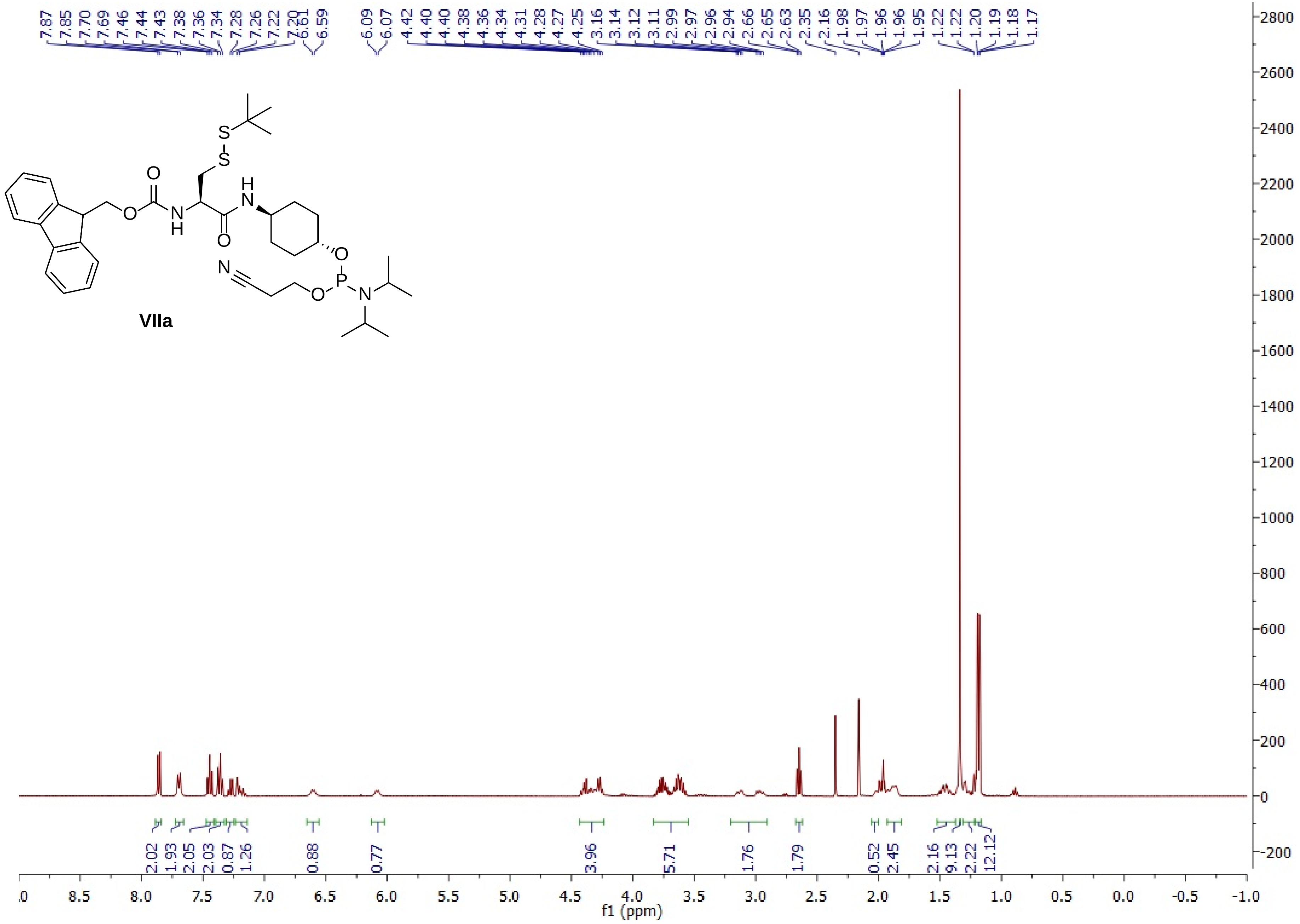


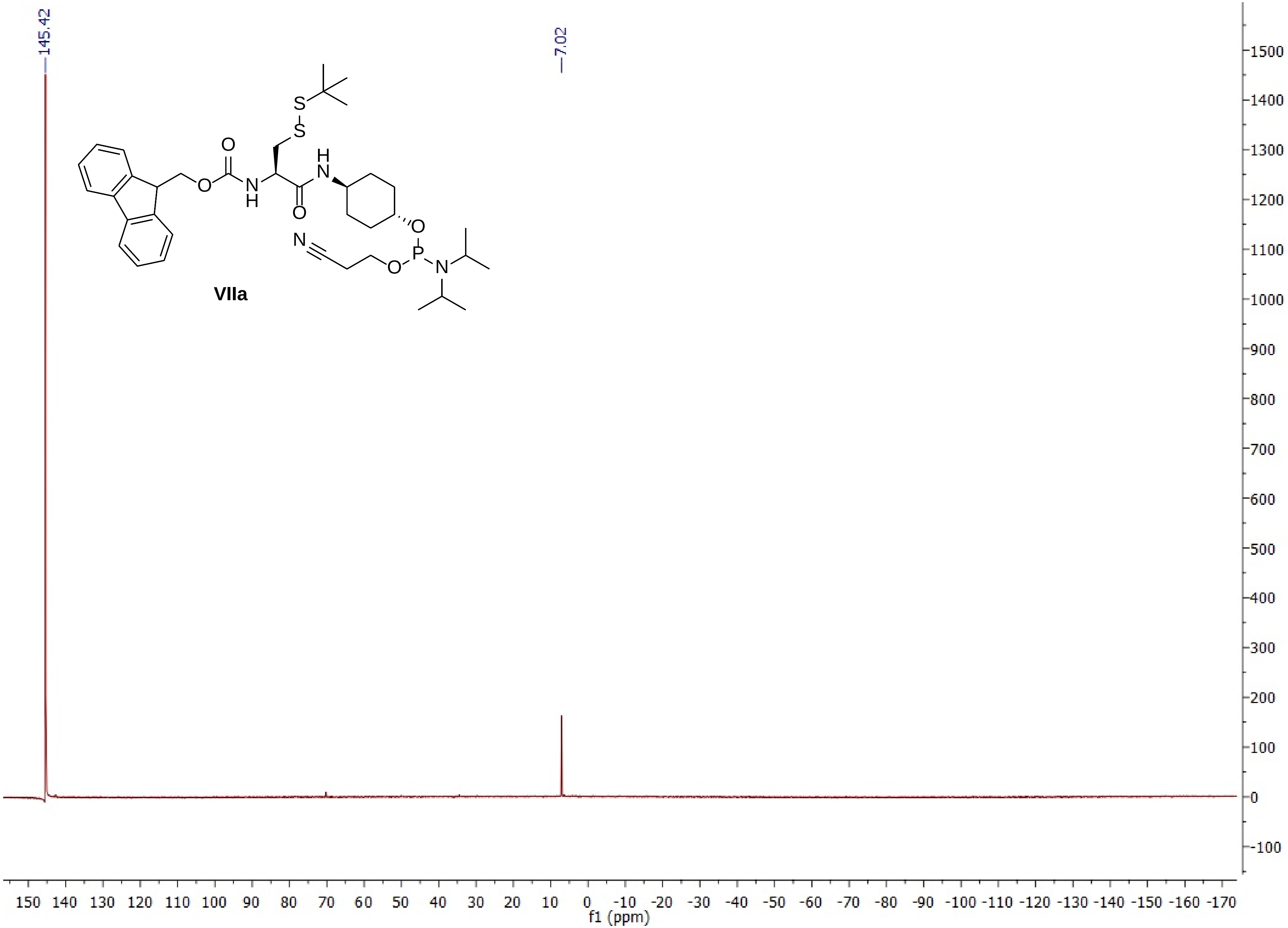
Precursor of **1**

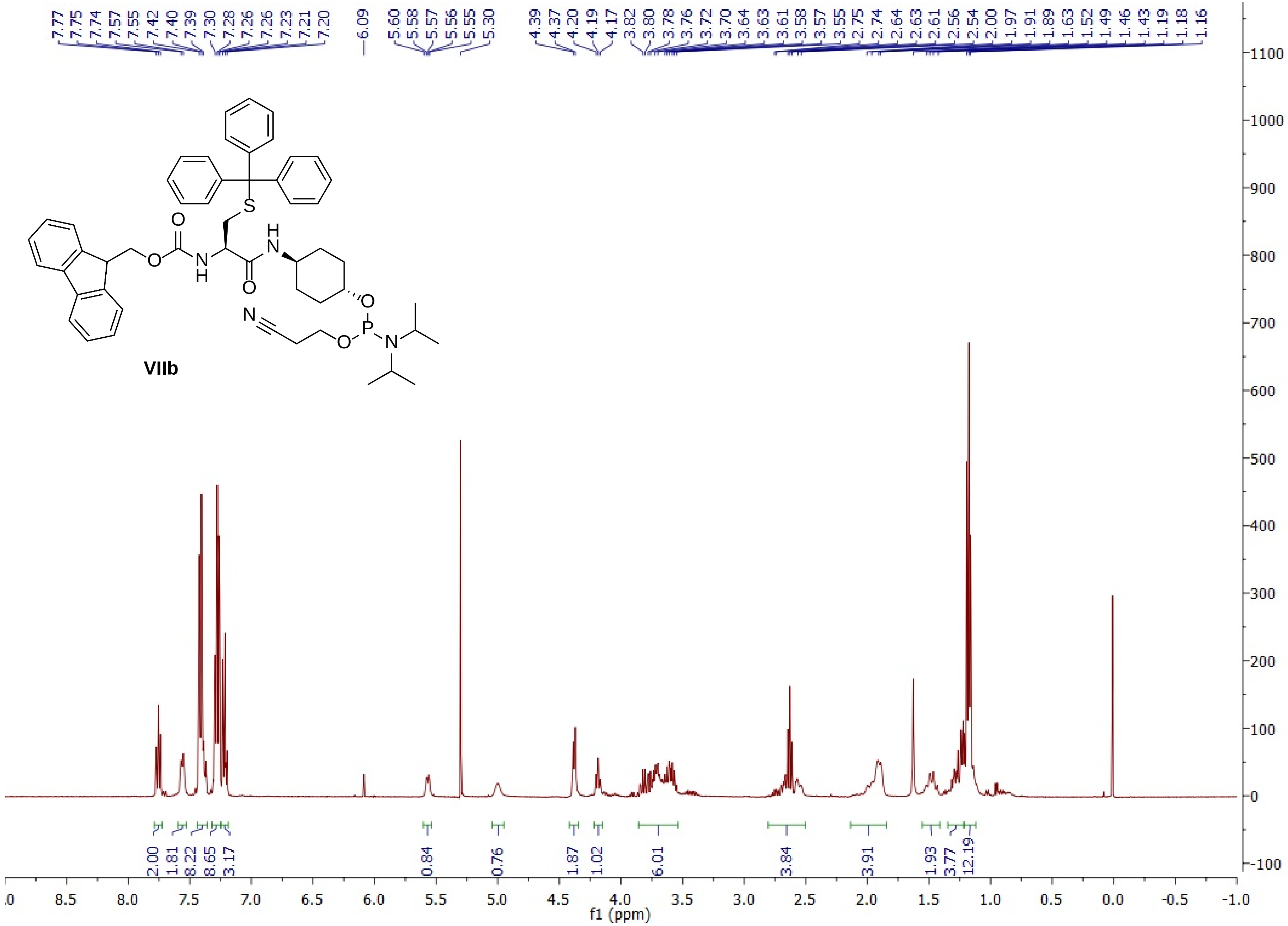

7.90
7.88
7.86
7.84
7.74
7.72
7.63
7.61
7.44
7.42
7.40
7.34
7.32
7.30
4.52
4.51
4.31
4.30
4.28
4.26
4.24
4.22
4.19
3.32
3.02
3.01
2.99
2.98
2.96
2.93
2.92
2.51
2.51
2.50
2.50
2.49
1.81
1.77
1.73
1.69
1.29
1.23
1.20
1.18
1.16
1.13
O
H
N
O
NH
OH
O
S
S
3.00
1.88
0.97
2.02
1.91
0.89
3.95
2.05
1.96
1.94
9.10
4.09
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500

VIa

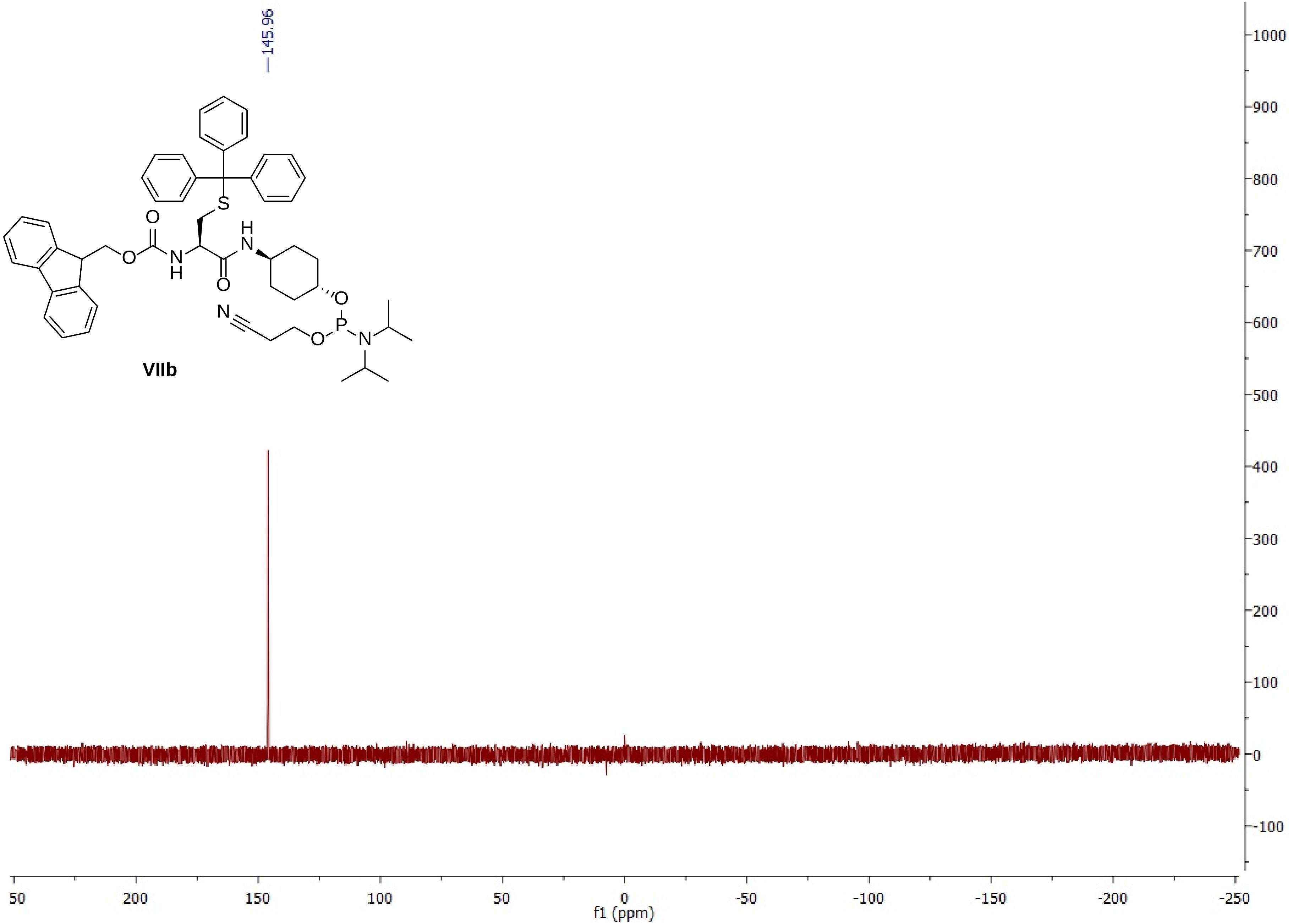

7.77
7.75
7.73
7.56
7.55
7.42
7.42
7.41
7.40
7.40
7.39
7.37
7.37
7.29
7.29
7.29
7.27
7.27
7.26
7.26
7.25
7.23
7.23
7.22
7.21
7.21
7.20
7.19
7.19
7.19
5.54
5.52
5.30
5.00
4.38
4.36
4.20
4.18
4.16
3.71
3.70
3.64
3.62
3.59
3.58
3.57
3.56
2.70
2.68
2.67
2.65
2.57
2.56
2.54
2.52
1.95
1.92
1.88
1.37
1.36
1.34
1.31
1.14
1.11
VIb
O
H
N
NH
OH
O
O
S
2.15
2.02
8.78
12.95
0.93
0.85
2.08
1.02
3.22
2.10
4.25
2.18
2.07
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100


7.87
7.85
7.70
7.69
7.46
7.44
7.43
7.38
7.36
7.34
7.28
7.26
7.22
7.20
6.61
6.59
6.09
6.07
4.42
4.40
4.40
4.38
4.36
4.34
4.31
4.28
4.27
4.25
3.16
3.14
3.12
3.11
2.99
2.97
2.96
2.94
2.66
2.65
2.63
2.35
2.16
1.98
1.97
1.96
1.96
1.95
1.22
1.22
1.20
1.19
1.18
1.17
VIIa
2.02
1.93
2.05
2.03
0.87
1.26
0.88
0.77
3.96
5.71
1.76
1.79
0.52
2.45
2.16
9.13
2.22
12.12
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
-200

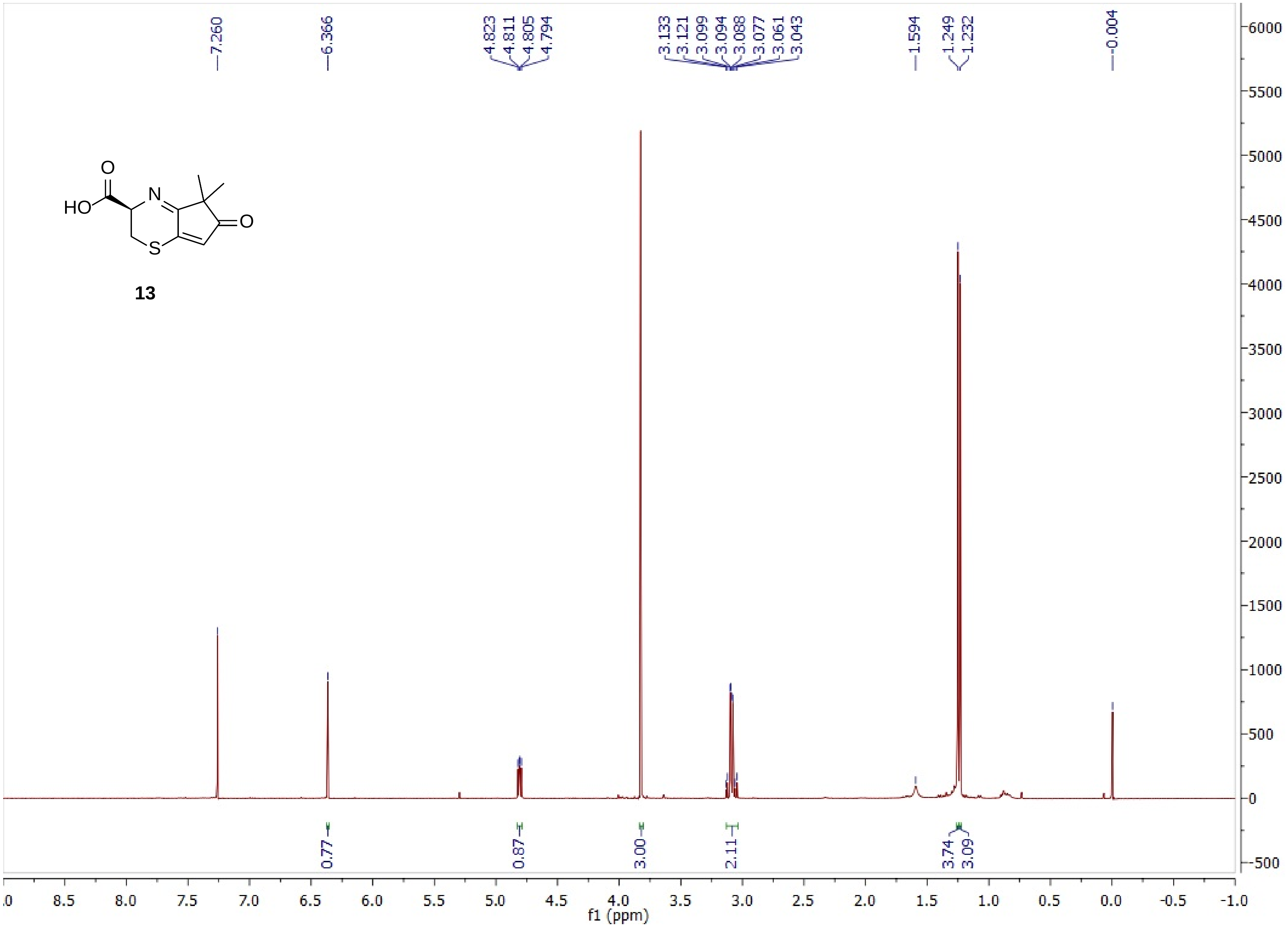

VIIa
145.42
7.02
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -110 -120 -130 -140 -150 -160 -170
f1 (ppm)

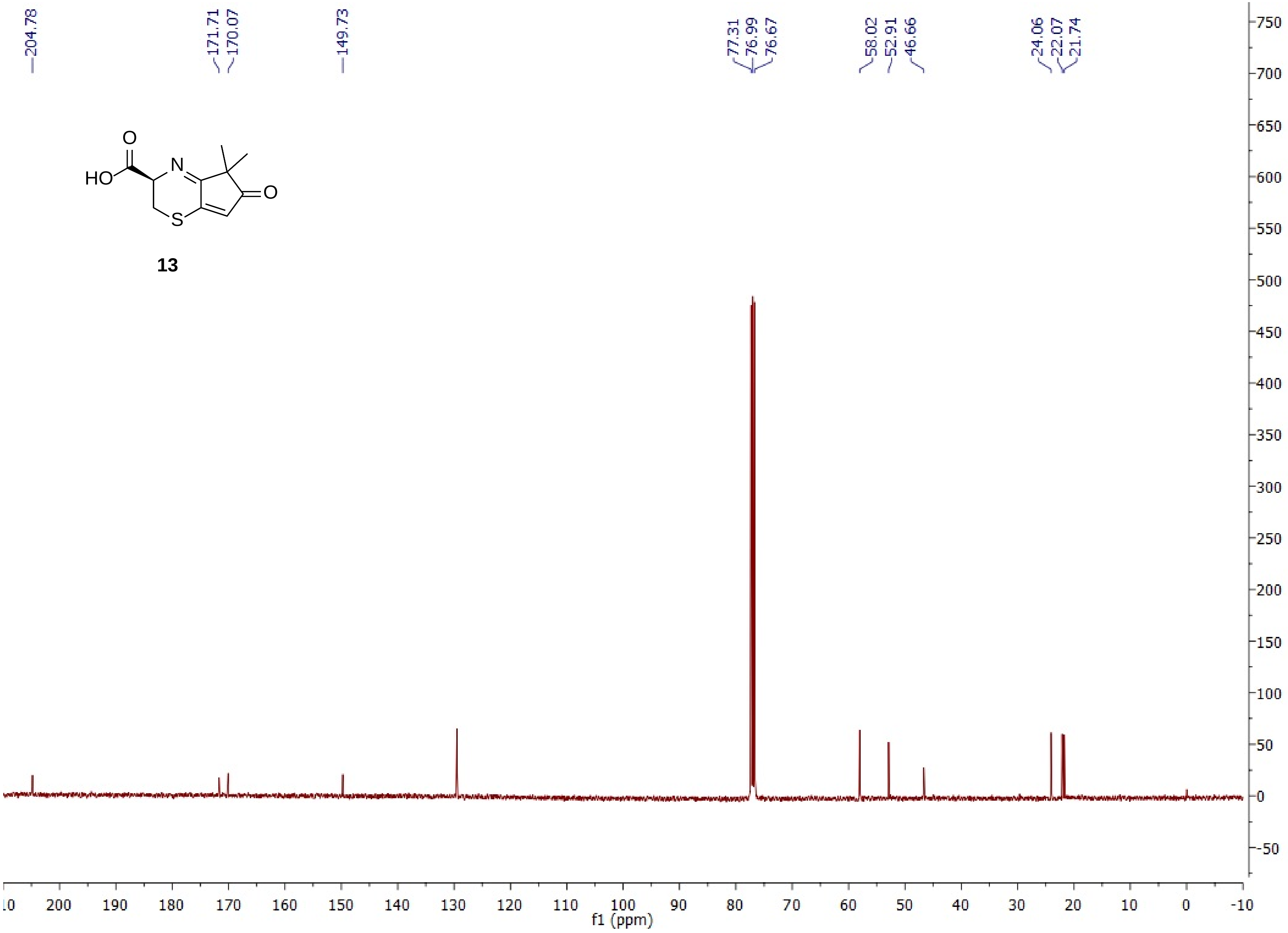

7.77
7.75
7.74
7.57
7.55
7.42
7.40
7.39
7.30
7.28
7.26
7.26
7.23
7.21
7.20
6.09
5.60
5.58
5.57
5.56
5.55
5.30
4.39
4.37
4.20
4.19
4.17
3.82
3.80
3.78
3.76
3.72
3.70
3.64
3.63
3.61
3.58
3.57
3.55
2.75
2.74
2.64
2.63
2.61
2.56
2.54
2.00
1.97
1.91
1.89
1.63
1.52
1.49
1.46
1.43
1.19
1.18
1.16
O
S
H
N
N
H
O
O
O
N
P
O
N
2.00
1.81
8.22
8.65
3.17
0.84
0.76
1.87
1.02
6.01
3.84
3.91
1.93
3.77
12.19
.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
f1 (ppm)
1100
1000
900
800
700
600
500
400
300
200
100
0
-100

VIIb

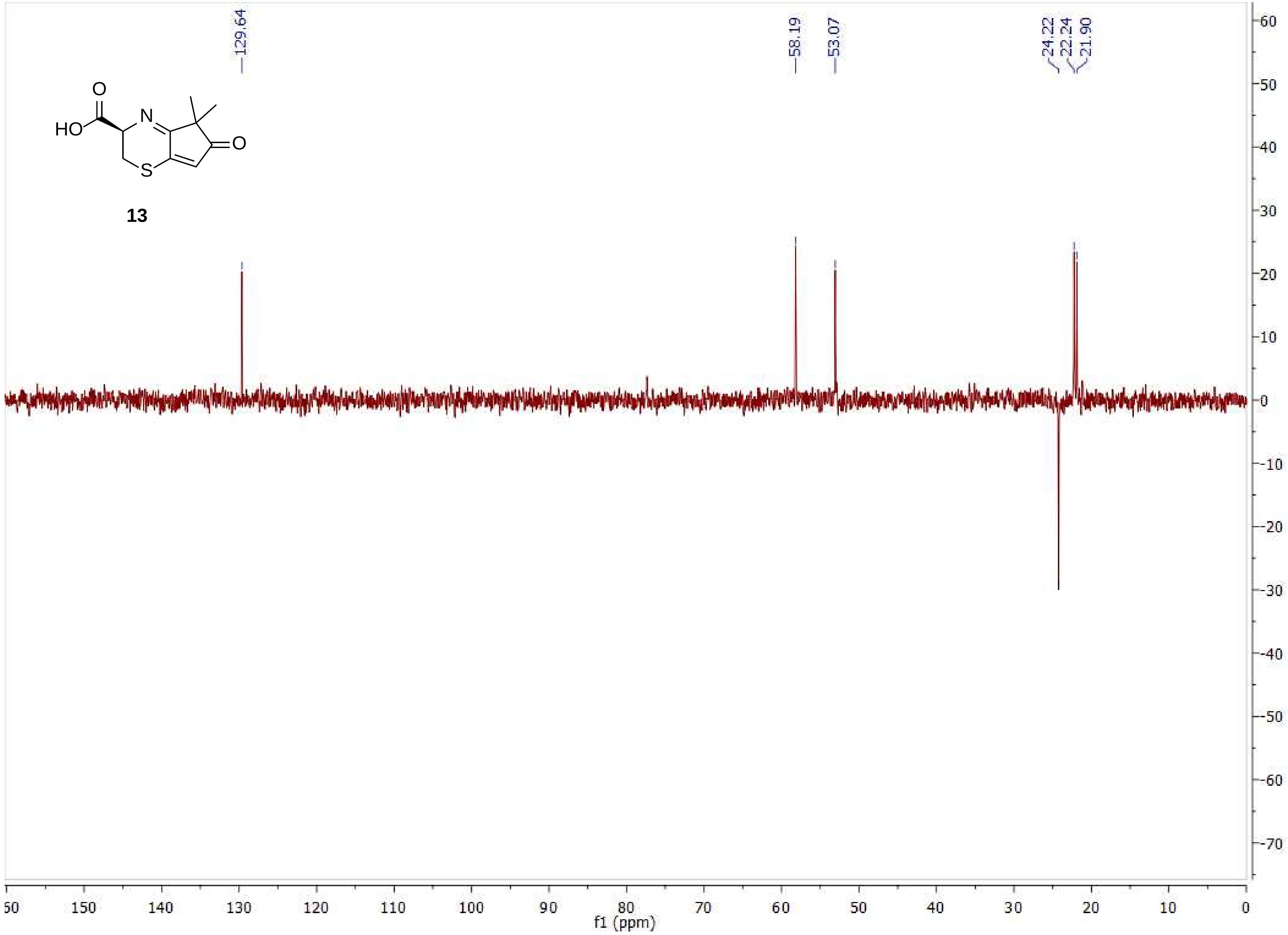

145.96
1000
900
800
700
600
500
400
300
200
100
0
-100
50
200
150
100
50
0
-50
-100
-150
-200
-250
f1 (ppm)

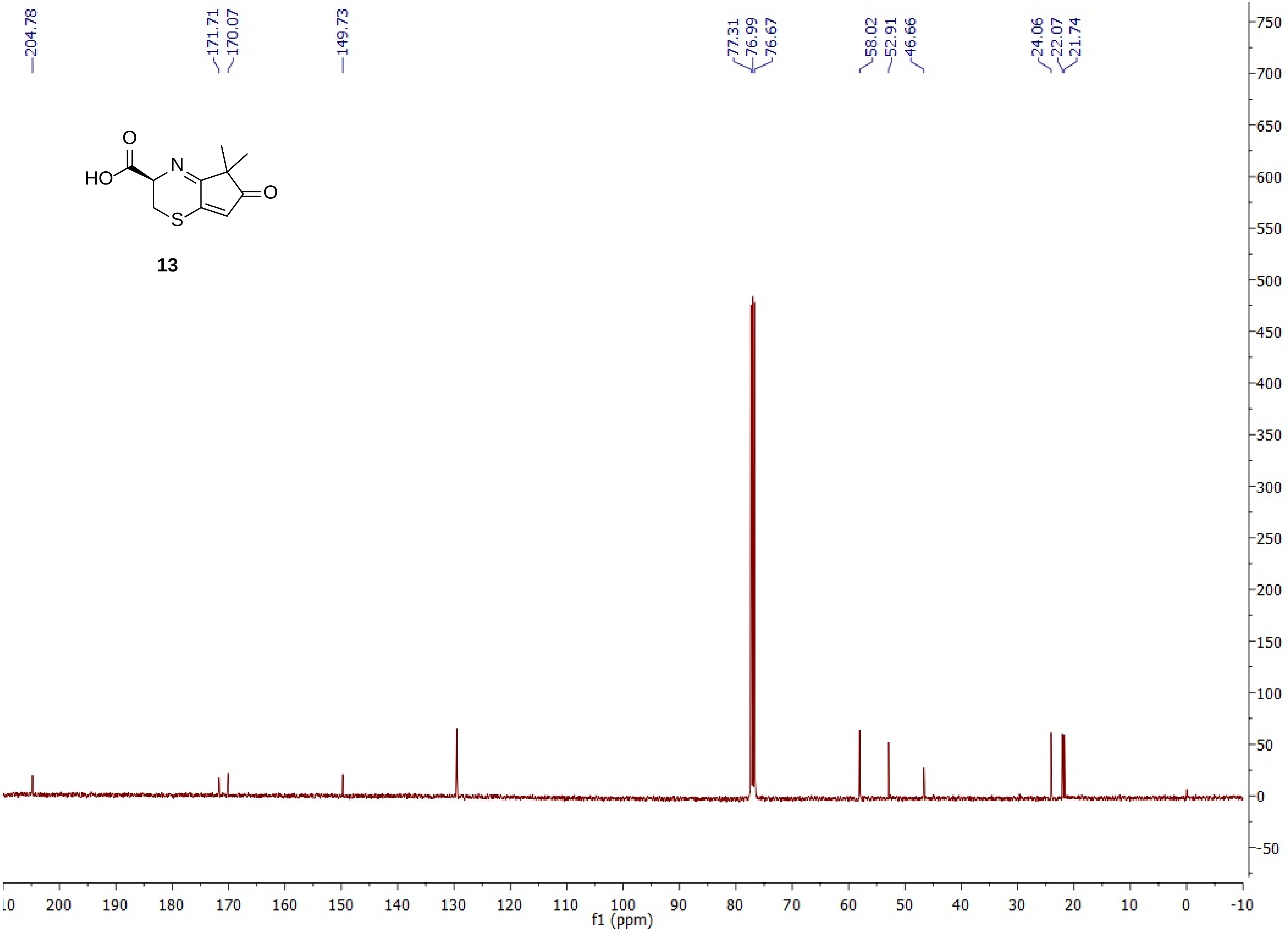
VIIb


13
HO
O
N
S
O
7.260
6.366
4.823
4.811
4.805
4.794
3.133
3.121
3.099
3.094
3.088
3.077
3.061
3.043
1.594
1.249
1.232
-0.004
0.77
0.87
3.00
2.11
3.74
3.09
f1 (ppm)
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500

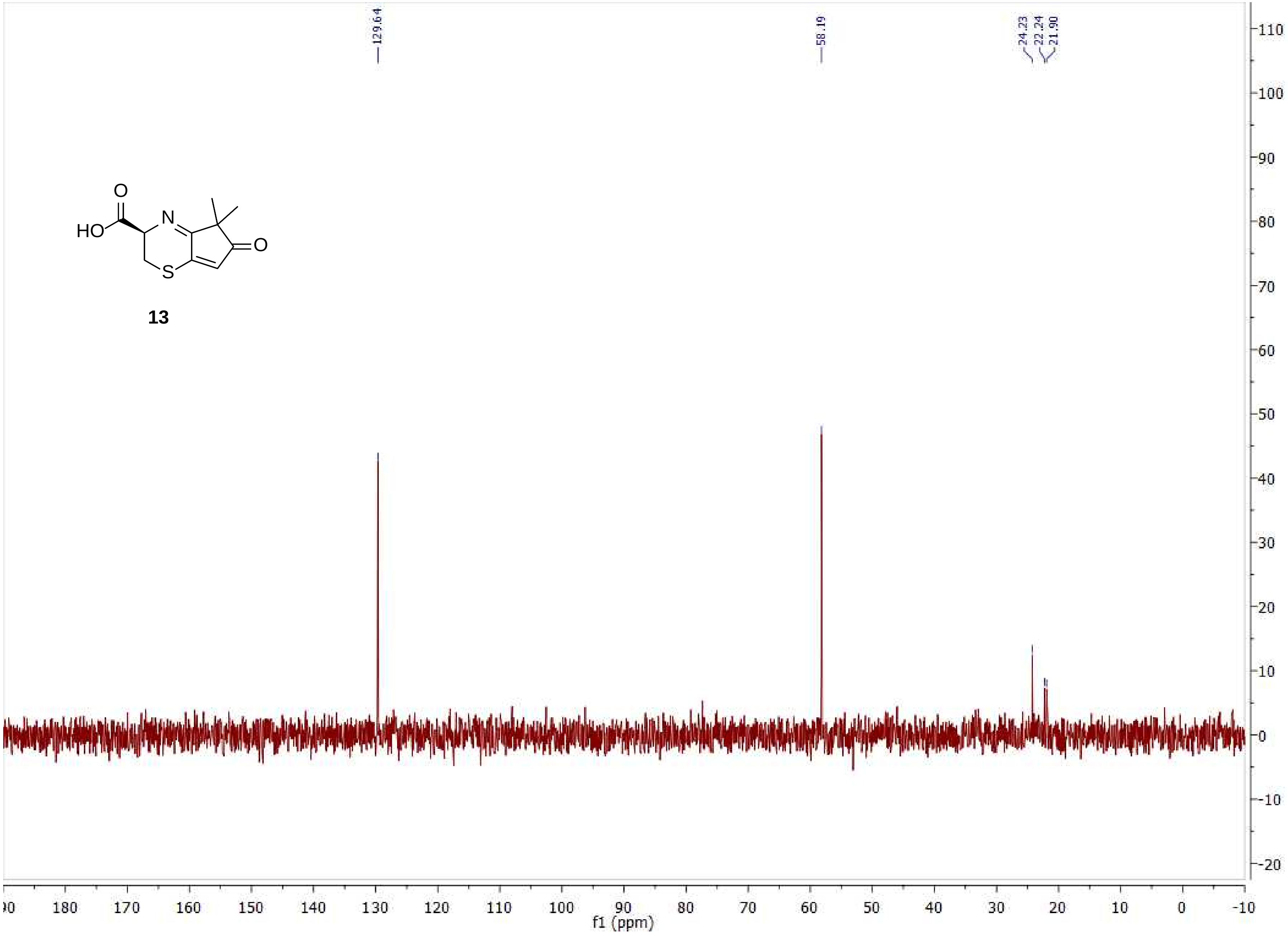

204.78
171.71
170.07
149.73
77.31
76.99
76.67
58.02
52.91
46.66
24.06
22.07
21.74
HO
O
N
S
O
750
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
-50
10
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

13

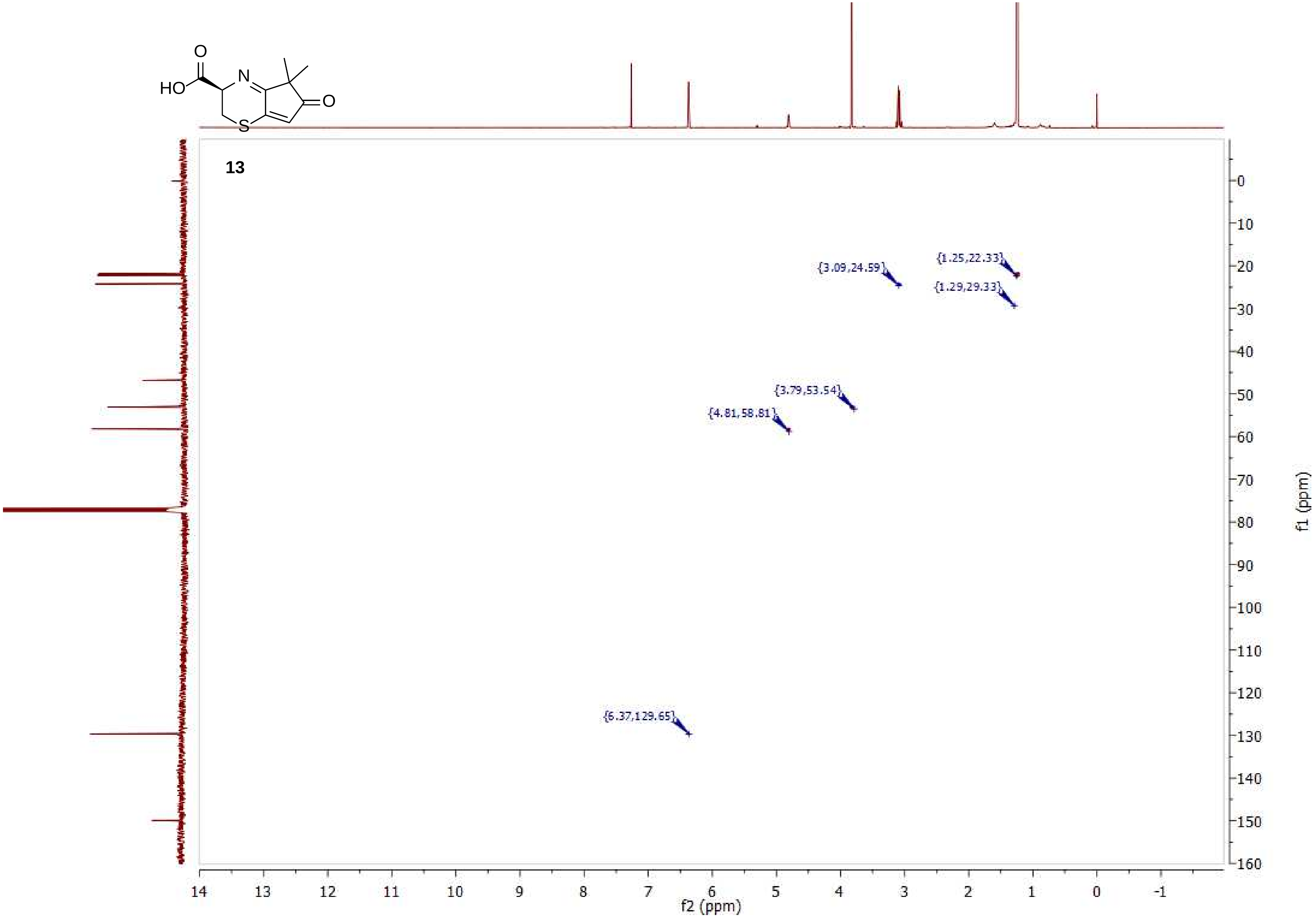

13
HO
O
N
S
O
129.64
58.19
53.07
24.22
22.24
21.90
f1 (ppm)
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70


204.78
171.71
170.07
149.73
77.31
76.99
76.67
58.02
52.91
46.66
24.06
22.07
21.74
O
HO
N
S
O
750
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
-50
10
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

13


129.64
58.19
24.23
22.24
21.90
O
HO
N
S
O
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
)0
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

13



13

HO
O
N
S
O
{3.09,24.59}
{1.25,22.33}
{1.29,29.33}
{3.79,53.54}
{4.81,58.81}
{6.37,129.65}
f2 (ppm)
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160

17. Cartesian coordinates (Å), energy and imaginary frequency (transition state structures) for all the stationary points.

Compound 1

Sum of electronic and thermal Free Energies (a.u)= -421.842956; CCSD(T) (a.u)= -421.0537717

C	1.17874900	0.36009000	0.00013500
C	0.66859800	1.75700500	-0.00003800
C	-0.66880900	1.75693600	-0.00034300
H	1.32774700	2.61597600	-0.00011600
H	-1.32800600	2.61586900	-0.00093700
C	-1.17885300	0.35992300	-0.00006500
C	0.00001600	-0.60106000	-0.00024300
O	2.34644800	0.03719000	0.00054100
O	-2.34648800	0.03691500	0.00061100
C	0.00020800	-1.46175400	-1.26603900
H	0.00039800	-0.84524100	-2.16796000
H	-0.88750900	-2.09722100	-1.27378600
H	0.88797300	-2.09722600	-1.27341900
C	0.00003100	-1.46193000	1.26552800
H	0.88799100	-2.09711500	1.27309600
H	-0.88753800	-2.09763300	1.27273400
H	-0.00038700	-0.84551200	2.16754900

Compound 1 · H-Cys-OMe MTA

Sum of electronic and thermal Free Energies (a.u)= -1182.946881; CCSD(T) (a.u)= -1181.038184

C	-3.38827400	-0.74665600	0.08881000
C	-2.77143200	-1.20447600	-1.21827500
C	-1.48216000	-0.38994000	-1.31698600
H	-3.45379900	-0.91041000	-2.02357200
H	-1.11042400	-0.21147300	-2.32423200
C	-1.79432800	0.89384000	-0.56723000
C	-2.96366900	0.68405300	0.38325600
O	-4.12551300	-1.40327300	0.78710800
O	-1.19563700	1.93562400	-0.70389500
C	3.58046200	0.28009700	-0.02032100
C	2.43653600	-0.71682200	0.09005700

O	3.44894000	1.48297300	-0.04232600
O	4.76237500	-0.32610900	-0.06386100
H	2.68061100	-1.55186000	-0.56898100
C	1.14613800	-0.04422500	-0.36760900
N	2.42442200	-1.21660800	1.46649800
H	1.24499400	0.31601700	-1.39388600
H	0.91318700	0.79681400	0.28765700
S	-0.21513200	-1.25760600	-0.31375700
H	1.89784100	-2.08373400	1.51033700
H	1.94071500	-0.54671900	2.06043000
C	5.91552400	0.53246100	-0.08721000
H	5.90136500	1.15267800	-0.98307300
H	6.77469600	-0.13227300	-0.10043700
H	5.93181700	1.15811900	0.80486900
H	-2.64087300	-2.28613900	-1.25082400
C	-2.58092700	0.91542400	1.84309400
C	-4.11215700	1.61939800	-0.04113500
H	-3.44130500	0.70709200	2.48230700
H	-2.28652100	1.95755100	1.98195300
H	-1.75221200	0.27063100	2.14334000
H	-3.78374700	2.65471800	0.06641200
H	-4.97336200	1.44507700	0.60653600
H	-4.40741900	1.44921900	-1.07912900

Cyclopentadiene

Sum of electronic and thermal Free Energies (a.u)= -193.994835

C	1.17605400	-0.27872400	-0.00015600
C	0.73218400	0.99162800	-0.00017000
C	-0.73788300	0.98738400	0.00006000
C	-1.17446200	-0.28555900	0.00008500
C	0.00350400	-1.21382400	0.00011000
H	0.00565100	-1.87340800	0.87661700
H	0.00548900	-1.87393600	-0.87597600
H	-2.20443600	-0.61712000	0.00020600
H	-1.35416500	1.87775800	0.00013100
H	1.34319100	1.88563800	-0.00028000
H	2.20788800	-0.60436300	-0.00027500

Cyclopentene

Sum of electronic and thermal Free Energies (a.u)= -195.188562

C	0.66610500	1.07144600	-0.05265500
C	-0.66604900	1.07148100	-0.05263700
C	-1.22677400	-0.32130100	0.11567800
C	-0.00003000	-1.21897900	-0.15309300
C	1.22675500	-0.32136200	0.11568800
H	1.60847400	-0.44834300	1.13568500
H	2.05756000	-0.53080700	-0.56192100
H	-0.00003600	-1.51408400	-1.20544200
H	-0.00005500	-2.12875600	0.44810700
H	-2.05758200	-0.53069800	-0.56194200
H	-1.60851000	-0.44827100	1.13567000
H	-1.28975900	1.95666200	-0.11400400
H	1.28986200	1.95659300	-0.11403700

Cysteine minimum 1

Sum of electronic and thermal Free Energies (a.u)= -1219.671037

C	-3.14409300	-0.67932800	0.00874600
C	-2.52912700	-1.14548900	-1.29362500
C	-1.11448400	-0.56720500	-1.26380500
H	-3.09766500	-0.67585100	-2.10424900
H	-0.67095700	-0.37571700	-2.23878100
C	-1.23900500	0.68399800	-0.41457600
C	-2.50061200	0.62866300	0.43651200
O	-4.02845200	-1.24109000	0.61342400
O	-0.46718500	1.61735200	-0.42541800
C	3.77467400	-0.29238800	0.33104600
C	2.33923800	-0.66947100	0.76220300
O	4.62952400	-1.20961300	0.31148000
O	3.98457000	0.90511700	-0.00507600
H	2.41205500	-1.57779700	1.36408200
C	1.59815000	-0.98381900	-0.53694500
N	1.64891400	0.35123000	1.53908800
H	2.18438800	-1.68710300	-1.12973800
H	1.47101600	-0.07124400	-1.12287000
S	-0.05585000	-1.73054300	-0.31417500

H	1.55060300	1.18868500	0.97001900
H	2.23590000	0.61281900	2.32672800
H	-2.58380700	-2.22749900	-1.41027600
C	-2.23476000	0.70935400	1.93627500
C	-3.42028200	1.77993400	-0.02116200
H	-3.18606100	0.73959400	2.47127300
H	-1.67941400	1.62197300	2.16253500
H	-1.66279000	-0.15195000	2.28575100
H	-2.93849800	2.73114100	0.21290100
H	-4.36778100	1.71447000	0.51669700
H	-3.61654000	1.73966300	-1.09520800
O	2.14552300	2.51979100	-1.32904600
H	2.75158600	1.92612600	-0.84330600
H	1.25913700	2.20481400	-1.10330600

Cysteine minimum 2

Sum of electronic and thermal Free Energies (a.u)= -1219.659245; Note: there is a very low lying imaginary frequency that does not disappear even with tight criterion: iFreq (cm-1)= -29.7074

C	-2.76256100	-0.34669500	0.05664300
C	-2.32143300	-0.31226400	-1.39084200
C	-0.80485800	-0.12201400	-1.34501800
H	-2.79890200	0.56055200	-1.84730400
H	-0.43084700	0.46449300	-2.18458200
C	-0.62719700	0.72007300	-0.04283300
C	-1.65930500	0.14728900	0.98654200
O	-3.85587300	-0.71058100	0.43346000
O	-0.82299000	2.01583000	-0.31224200
C	3.14715000	-0.20979900	0.82515700
C	1.69510200	-0.59938400	0.45378600
O	3.92881400	-1.17226200	0.98681600
O	3.41363400	1.00919200	0.91017700
H	1.33489500	-1.32209100	1.18216400
C	1.66491100	-1.19001900	-0.94570200
N	0.83916900	0.62194200	0.51709700
H	2.32060900	-2.05933500	-0.97730800
H	2.01666600	-0.45871900	-1.68034300
S	0.00618100	-1.75455200	-1.40520900

H 1.34876600 1.35945300 0.01242100
H 0.80056500 0.95124100 1.48782300
H -2.66216100 -1.19783400 -1.92979900
C -1.19979700 -1.04522600 1.84520500
C -2.19535800 1.23052600 1.92262500
H -2.05328800 -1.39374000 2.43086100
H -0.42548300 -0.72691100 2.54909600
H -0.82839600 -1.88169700 1.25533600
H -1.37289600 1.65327600 2.50641900
H -2.91516400 0.78883200 2.61449300
H -2.68201000 2.03654800 1.37559500
O 1.22497500 2.70297700 -1.80482900
H 1.27079500 2.05215100 -2.51315800
H 0.37537000 2.49682700 -1.31848500

Cysteine minimum 3

Sum of electronic and thermal Free Energies (a.u)= -1219.674799

C -2.80043500 0.17622400 -0.19145200
C -2.33856900 1.16405600 0.85850900
C -0.81954000 1.00634700 0.92755100
H -2.79966500 0.85887100 1.80317500
H -0.43414300 1.12395700 1.94209500
C -0.58533500 -0.46262700 0.48126300
C -1.67908600 -0.77671000 -0.58437700
O -3.91987500 0.13943100 -0.65534100
O -0.92705200 -1.27173800 1.60431100
C 3.04725600 -0.44193700 -0.80732800
C 1.60955900 0.11942200 -0.68139800
O 3.86074500 0.29142200 -1.42441400
O 3.29389400 -1.57244800 -0.32889600
H 1.23917700 0.18539000 -1.71028900
C 1.64334500 1.52538000 -0.10348900
N 0.80126700 -0.76938400 0.15898800
H 2.27174700 2.16497700 -0.72107200
H 2.03479500 1.52059900 0.91817300
S -0.01281100 2.27065900 -0.10642900
H 1.63289900 -0.71253800 1.81952700

H	0.85372500	-1.71784300	-0.20676500
H	-2.68801300	2.17280900	0.63202000
C	-1.26245700	-0.44687500	-2.02799100
C	-2.16066300	-2.22969900	-0.55374600
H	-2.12130500	-0.61458200	-2.68140400
H	-0.46215900	-1.12291800	-2.33947400
H	-0.92879700	0.58218800	-2.15219500
H	-1.31177400	-2.90191100	-0.70616600
H	-2.87049700	-2.38955600	-1.36769500
H	-2.64699300	-2.48792600	0.38574900
O	1.83518600	-0.66242400	2.78216100
H	2.05854700	0.26238400	2.93421100
H	-0.29905200	-1.06418400	2.31225500

Cysteine transition state 1

Sum of electronic and thermal Free Energies (a.u)= -1219.652987; iFreq (cm-1)= -151.1932

C	-2.87838200	-0.41828100	-0.11564800
C	-2.29566400	-0.52586900	-1.50634800
C	-0.79454700	-0.29447800	-1.29550700
H	-2.71753800	0.30150200	-2.08955800
H	-0.28464000	0.09227800	-2.17811700
C	-0.84299800	0.80145100	-0.21013500
C	-1.99395000	0.45706600	0.77930500
O	-3.91515800	-0.92673900	0.24757000
O	-0.63971400	1.97729200	-0.56166700
C	3.32657000	-0.11581500	0.72084600
C	1.86898200	-0.62491400	0.73308300
O	4.20276400	-0.88381600	1.17520300
O	3.52597800	1.01172300	0.19509200
H	1.73853800	-1.33474500	1.54973900
C	1.65595400	-1.35065300	-0.59667300
N	0.90279700	0.46923900	0.89082600
H	2.27916000	-2.24547700	-0.63015400
H	1.92625800	-0.70860900	-1.43981100
S	-0.05554300	-1.89284000	-0.83639000
H	1.33925600	1.33574100	0.58092600
H	0.68348400	0.59365400	1.87551600

H	-2.55803300	-1.46507200	-1.99428700
C	-1.64200700	-0.32018200	2.05298600
C	-2.78356000	1.71613100	1.15223000
H	-2.56670600	-0.69821800	2.49534000
H	-1.17315200	0.34417800	2.78211200
H	-0.98015700	-1.16277500	1.85202700
H	-2.14431800	2.39678300	1.71795700
H	-3.63408200	1.43590700	1.77703200
H	-3.15297600	2.23934800	0.26847100
O	1.90447700	2.08892400	-1.92194800
H	2.46226400	1.70554000	-1.22181000
H	1.00258900	2.08493600	-1.55776800

Cysteine transition state 2

Sum of electronic and thermal Free Energies (a.u)= -1219.653678; iFreq (cm-1)= -181.5284

C	-2.75683700	0.27337600	-0.21342300
C	-2.34160200	0.91113200	1.09398600
C	-0.82324600	0.74419200	1.15849700
H	-2.82012500	0.33453500	1.89221300
H	-0.45708400	0.60195500	2.17604000
C	-0.60299300	-0.57414400	0.36700100
C	-1.63307400	-0.58287800	-0.79301600
O	-3.84068700	0.40594600	-0.73658700
O	-0.86290000	-1.63444900	1.22820900
C	3.13276300	-0.18308300	-0.77156800
C	1.68517100	0.33845500	-0.59214300
O	3.94546000	0.65987000	-1.21148200
O	3.36050400	-1.37631900	-0.47774500
H	1.32173000	0.63588900	-1.57455400
C	1.65111700	1.52105900	0.35962000
N	0.84124700	-0.77317700	-0.06603400
H	2.30461000	2.30346600	-0.02270600
H	1.99603500	1.23020100	1.35640100
S	-0.01355900	2.22948500	0.47917500
H	1.30255000	-1.13290400	0.83704500
H	0.86382400	-1.55310000	-0.73106800
H	-2.69647800	1.94004300	1.16766100

C	-1.16279700	0.07344800	-2.10226700
C	-2.14338200	-1.98807600	-1.11933600
H	-2.00406100	0.07945100	-2.79800100
H	-0.36182800	-0.52031700	-2.55099500
H	-0.82314500	1.09942700	-1.97000700
H	-1.30640500	-2.62140100	-1.42601200
H	-2.85034200	-1.92539700	-1.94842400
H	-2.64007600	-2.45392200	-0.26976300
O	1.28195800	-1.66705100	2.38384300
H	1.40053600	-0.90091100	2.95460100
H	-0.00692100	-1.68139800	1.88900100

Homocysteine minimum 1

Sum of electronic and thermal Free Energies (a.u)= -1258.948820

C	-2.73564200	-1.35009400	-0.60070600
C	-2.74594300	-0.21222700	-1.60231400
C	-1.72352800	0.77526200	-1.03863400
H	-3.74538400	0.23633700	-1.57521900
H	-1.85791100	1.81558600	-1.32949200
C	-1.78696800	0.55474600	0.45776500
C	-2.36016800	-0.81799200	0.77353000
O	-3.00558200	-2.50427600	-0.84107900
O	-1.43969900	1.35914800	1.29709000
C	-3.65838000	-0.62025700	1.57994400
H	-4.36318800	0.03126000	1.05799200
H	-3.41160900	-0.17181200	2.54412200
H	-4.12728300	-1.59194700	1.74516100
C	-1.38230300	-1.71434100	1.52974500
H	-1.87616100	-2.65084400	1.79609400
H	-1.06745900	-1.21274500	2.44737100
H	-0.50006100	-1.93794100	0.92631800
C	4.06260600	-0.91055500	0.07554700
C	2.64306300	-0.33764100	0.25839100
O	4.27496500	-1.54132200	-0.99525500
O	4.91426800	-0.70634200	0.97594700
H	1.94907700	-1.13607500	-0.01828100
C	1.04972100	1.38671000	-0.76425500

N	2.31704400	0.09683100	1.61376500
H	0.69690600	1.59275800	0.24538900
S	-0.06526100	0.16259500	-1.53541700
H	3.00912100	0.78935900	1.89340500
H	2.45591800	-0.68765400	2.24474600
H	-2.54384700	-0.55646300	-2.61633300
C	2.46960500	0.83096000	-0.71856200
H	2.78971800	0.52311700	-1.71702200
H	3.13411000	1.64438300	-0.40409100
H	1.01232100	2.30185400	-1.35704200
O	-0.83579800	4.03159900	0.43732300
H	0.00843100	3.99793200	-0.02450400
H	-0.99544500	3.11566200	0.71679100

Homocysteine minimum 2

Sum of electronic and thermal Free Energies (a.u)= -1258.932886

C	2.72652700	0.46740900	-0.47748100
C	2.48247000	-1.01017200	-0.70880700
C	1.06507000	-1.30595000	-0.19053700
H	3.24614400	-1.58131200	-0.17857000
H	1.07837300	-2.12877300	0.52416000
C	0.73322900	-0.01719200	0.64270400
C	1.43979400	1.17114900	-0.08070000
O	3.79977400	1.01255800	-0.61820100
O	1.22391600	-0.17559900	1.88377900
C	1.68536000	2.36782000	0.82941900
H	2.26877200	2.09867400	1.70851500
H	0.73136100	2.79283700	1.15666900
H	2.22189300	3.14290900	0.27804600
C	0.83399300	1.67643100	-1.40701800
H	1.55403300	2.35432600	-1.87166400
H	-0.07581700	2.24899200	-1.21081800
H	0.61523700	0.87704400	-2.11727200
C	-2.74594500	1.45234200	-0.05916000
C	-1.78519400	0.26711200	-0.29249200
O	-3.71587900	1.53309500	-0.84272300
O	-2.45499200	2.24849800	0.86629400

H	-1.24493700	0.49139500	-1.20730800
C	-1.51466900	-2.24289700	-0.58328300
N	-0.79426200	0.21867800	0.82607700
H	-1.17765400	-2.57947600	0.40067200
S	-0.01357500	-1.87885000	-1.55545900
H	-1.07732500	-0.49643800	1.50835800
H	-0.88779500	1.11867800	1.32259300
H	2.60938900	-1.21454900	-1.77654500
C	-2.49027000	-1.06853600	-0.47090100
H	-3.09119600	-0.99641000	-1.37787700
H	-3.17140100	-1.25076800	0.36693500
H	-2.00541300	-3.08963800	-1.06321500
O	-0.19683300	-2.06363400	3.00474100
H	-0.07178400	-2.85638300	2.47261800
H	0.42507200	-1.38468500	2.60792000

Homocysteine minimum 3

Sum of electronic and thermal Free Energies (a.u)= -1258.944322

C	-2.81734600	-0.24952600	-0.32659300
C	-2.48381200	1.22562000	-0.41738700
C	-1.01314900	1.36442400	0.00888500
H	-3.16197800	1.78866400	0.22539900
H	-0.90890700	2.08686700	0.82042000
C	-0.66169600	-0.03427300	0.60881400
C	-1.55230800	-1.06709100	-0.14162700
O	-3.93687300	-0.70672700	-0.41058800
O	-1.22513900	-0.02286300	1.92118400
C	-1.78994800	-2.35528900	0.63793900
H	-2.31190200	-2.17865100	1.57719900
H	-0.83330400	-2.84167900	0.85147600
H	-2.38754700	-3.04128200	0.03402800
C	-1.07758500	-1.43105800	-1.56197900
H	-1.86904800	-2.00549100	-2.04940900
H	-0.19045900	-2.06592800	-1.50738000
H	-0.85830000	-0.55971200	-2.18115900
C	2.61453000	-1.52052700	-0.32956100
C	1.66758300	-0.30304600	-0.37551100

O	3.49916900	-1.56729900	-1.22100300
O	2.43264100	-2.38489100	0.56292000
H	1.12275800	-0.40931400	-1.31357500
C	1.56879500	2.24660800	-0.46875800
N	0.75693700	-0.31541100	0.77906300
H	1.30861500	2.55872500	0.54398600
S	-0.01032000	2.04501500	-1.35956200
H	1.22155700	0.88084000	2.19087400
H	0.83353300	-1.23571800	1.21237700
H	-2.66744200	1.55173800	-1.44595800
C	2.45872700	1.00194400	-0.46845700
H	3.04581000	0.97472200	-1.38778900
H	3.15773900	1.07468000	0.37234500
H	2.08490700	3.08206000	-0.94317900
O	1.12447200	1.38640700	3.02810000
H	1.38284100	2.28965100	2.81581400
H	-0.73920600	0.63374000	2.44458000

H-Cysteine-OMe

Sum of electronic and thermal Free Energies (a.u)= -761.100143; CCSD(T) (a.u)= -759.9560672

C	-0.80789000	0.29913700	0.15579600
C	0.41465800	0.93465400	-0.47568000
O	-1.07773600	0.41015000	1.33147800
O	-1.55649100	-0.37632500	-0.70747500
H	0.41997400	0.69284600	-1.53871300
C	1.67379400	0.37619500	0.18559400
N	0.28015500	2.38241500	-0.32089200
H	1.66129000	0.60329800	1.25363000
H	2.55362500	0.84165500	-0.25828800
S	1.80796400	-1.44144100	0.06845200
H	1.82773800	-1.50217400	-1.27556700
H	1.08176200	2.82741700	-0.76053800
H	0.33465600	2.60786500	0.67073700
C	-2.72541300	-1.01539700	-0.16423000
H	-2.43056500	-1.73448200	0.59984500
H	-3.19910600	-1.52072100	-1.00090800
H	-3.39496700	-0.26767200	0.25992200

Homocysteine transition state 1

Sum of electronic and thermal Free Energies (a.u)= -1258.924321; iFreq (cm-1)= -96.4092

C	-2.85059000	-0.61113900	-0.61519300
C	-2.73248100	0.88989000	-0.76420000
C	-1.32111500	1.25349900	-0.27247400
H	-3.47980300	1.32233000	-0.08878900
H	-1.29694400	2.21792900	0.23607300
C	-1.11601900	0.14375000	0.78222200
C	-1.73278000	-1.17364500	0.26596800
O	-3.72226700	-1.28613700	-1.11469400
O	-1.15975300	0.43897900	1.97926400
C	-2.30093500	-2.01631800	1.40414400
H	-3.01025600	-1.45148500	2.01092100
H	-1.48669700	-2.35708900	2.04794100
H	-2.80730700	-2.89065900	0.99113000
C	-0.85197900	-2.04509000	-0.64296600
H	-1.47422800	-2.84416200	-1.05256700
H	-0.04588100	-2.50004300	-0.06533500
H	-0.42634000	-1.47711700	-1.46906600
C	3.29588700	-1.22374800	0.03361500
C	2.05312600	-0.39030200	-0.33484100
O	4.16360200	-1.34770800	-0.87041400
O	3.35987100	-1.73327600	1.17863900
H	1.57976800	-0.91828200	-1.16713300
C	1.31114900	1.99223200	-0.91293200
N	1.06467000	-0.27269400	0.73577700
H	1.00977600	2.33661400	0.07726200
S	-0.20598800	1.35310200	-1.71159600
H	1.33402000	0.46277500	1.38897100
H	1.04366700	-1.13321300	1.27995100
H	-2.96783700	1.21612900	-1.77831200
C	2.46810100	0.99868200	-0.82476900
H	2.94972300	0.89410500	-1.79777000
H	3.21050300	1.42329300	-0.13851300
H	1.62157700	2.87037200	-1.47998800
O	0.22997600	2.78007200	2.61056000

H	0.07573000	3.37938500	1.87272500
H	-0.29197400	1.98602700	2.38930500

Homocysteine transition state 2

Sum of electronic and thermal Free Energies (a.u)= -1258.926604; iFreq (cm-1)= -136.1740

C	-2.77690300	-0.27119500	-0.44626600
C	-2.49019800	1.21473400	-0.39773500
C	-1.03618900	1.36668600	0.07745900
H	-3.19655300	1.68250700	0.29042500
H	-0.97099900	2.05216200	0.92269200
C	-0.70117000	-0.05510800	0.63203800
C	-1.50827000	-1.07686500	-0.20196000
O	-3.86357900	-0.75499500	-0.67010800
O	-1.19441600	-0.09890100	1.93946700
C	-1.78461700	-2.38475800	0.53181500
H	-2.34642100	-2.22635900	1.45076800
H	-0.84338900	-2.88664700	0.77519400
H	-2.35952400	-3.04648100	-0.11875100
C	-0.94633200	-1.39955600	-1.60151400
H	-1.69741700	-1.98323800	-2.13789200
H	-0.04779300	-2.01493200	-1.51248000
H	-0.72252000	-0.51047600	-2.19366700
C	2.68377400	-1.43383300	-0.34831700
C	1.73818400	-0.21217300	-0.35464900
O	3.62592600	-1.39817800	-1.16908600
O	2.41046800	-2.36657700	0.44491000
H	1.18617200	-0.26433400	-1.28907700
C	1.53631800	2.32758400	-0.28000900
N	0.77423900	-0.34851900	0.77573800
H	1.23875200	2.57238000	0.74171500
S	-0.00546400	2.11698200	-1.23295200
H	1.04562100	0.26756000	1.62579700
H	0.85944600	-1.32562400	1.09573800
H	-2.67300000	1.63996300	-1.38875800
C	2.47560800	1.11762000	-0.31478700
H	3.10070300	1.15687600	-1.20725000
H	3.13656300	1.15556500	0.55716900

H	2.03686800	3.20401700	-0.69152300
O	0.71923100	1.05094300	2.98533800
H	0.70936800	2.00341700	2.84827800
H	-0.51039300	0.47441400	2.50807800

Methylcysteinate minimum 1

Sum of electronic and thermal Free Energies (a.u)= -1259.368036

C	3.50336800	-0.57313900	0.03540400
C	2.90190300	-1.00894100	1.35515000
C	1.46420200	-0.49327500	1.29620300
H	3.44598900	-0.48395700	2.14824200
H	1.00608800	-0.27988500	2.25968600
C	1.53741000	0.72166800	0.39017500
C	2.80157000	0.68144800	-0.45727100
O	4.41671700	-1.12066200	-0.53804100
O	0.72431100	1.61856800	0.35685900
C	-3.41281400	-0.68205900	-0.41971100
C	-1.94281800	-0.89070000	-0.77060600
O	-4.22396600	-1.57734400	-0.34933700
O	-3.70807500	0.58998400	-0.15785300
H	-1.91403500	-1.82906900	-1.32829600
C	-1.21898200	-1.07402500	0.56669100
N	-1.34275000	0.15911800	-1.57106400
H	-1.77727500	-1.76990800	1.19373700
H	-1.15356100	-0.11901900	1.09190800
S	0.46917200	-1.74620200	0.39265400
H	-1.33006300	1.03040000	-1.04662100
H	-1.90556000	0.31772200	-2.40187300
C	-5.06672300	0.87311900	0.22577400
H	-5.73927500	0.61785100	-0.59231400
H	-5.09590100	1.94113900	0.42369200
H	-5.32645000	0.30879700	1.12071300
H	3.00237200	-2.08205400	1.51704000
C	2.52951700	0.67177500	-1.95857600
C	3.66786500	1.89632500	-0.06477800
H	3.47747400	0.71709500	-2.49832600
H	1.93385300	1.54640300	-2.22850100

H	1.99520200	-0.23067700	-2.26055900
H	3.14191600	2.81059200	-0.34585300
H	4.61547600	1.84633300	-0.60394800
H	3.86920900	1.92022600	1.00869500
O	-1.89233300	2.46739900	1.22463400
H	-2.47423400	1.83422100	0.78271100
H	-0.99894200	2.17245100	0.99402900

Methylcysteinate minimum 2

Sum of electronic and thermal Free Energies (a.u)= -1259.351441

C	3.04979800	0.04825100	-0.36518100
C	2.80005200	-0.83030300	0.84197800
C	1.28013900	-0.86025200	1.00661300
H	3.26328900	-0.32993300	1.69845200
H	0.96724200	-0.94660700	2.04749200
C	0.88138200	0.55670400	0.49081500
C	1.80738700	0.85560900	-0.73541000
O	4.10358900	0.11708700	-0.95912000
O	0.99412200	1.44958700	1.46799500
C	-2.85789700	-0.25751500	-0.70367100
C	-1.36167500	-0.52941600	-0.60493300
O	-3.61951600	-1.06698700	-1.17472200
O	-3.21242600	0.92583800	-0.23629800
H	-1.00037600	-0.61007200	-1.62917500
C	-1.12332100	-1.83289100	0.14339200
N	-0.65970600	0.59543600	0.07175700
H	-1.68132400	-2.62961000	-0.34708600
H	-1.46406800	-1.74904200	1.17941500
S	0.62290000	-2.30032300	0.09985300
H	-1.12443200	0.75857200	0.97617500
H	-0.78221000	1.46005600	-0.46764500
C	-4.61649100	1.24713900	-0.31123700
H	-4.93778400	1.24176800	-1.35180400
H	-4.70945700	2.24121100	0.11560400
H	-5.18841300	0.52360900	0.26815900
H	3.27463700	-1.80645300	0.72849100
C	1.32582800	0.37673000	-2.11623100

C	2.15561700	2.34104800	-0.83905800
H	2.11555800	0.58001200	-2.84272700
H	0.44204900	0.94265700	-2.42464100
H	1.10556600	-0.68940700	-2.14824600
H	1.24509900	2.92379000	-1.00379200
H	2.82263700	2.49577900	-1.68942400
H	2.64520400	2.70763700	0.06206300
O	-0.94958000	0.80740700	3.13041900
H	-0.79268100	-0.09690700	3.42229700
H	-0.12343500	1.06597000	2.63778700

Methylcysteinate minimum 3

Sum of electronic and thermal Free Energies (a.u)= -1259.372393

C	-3.08614600	-0.17363100	-0.21518400
C	-2.78867400	0.98234200	0.71671000
C	-1.26388700	1.05782900	0.80108700
H	-3.21383400	0.72358500	1.69130200
H	-0.91327300	1.34445400	1.79458700
C	-0.81387100	-0.40088500	0.52073100
C	-1.82921500	-0.98393100	-0.50719800
O	-4.18124000	-0.42932400	-0.66632500
O	-1.04950100	-1.12653000	1.72166600
C	2.78655600	0.06789800	-0.72155200
C	1.29254100	0.36724200	-0.68906900
O	3.57297700	0.77209400	-1.31512800
O	3.13235900	-1.03602600	-0.07841500
H	0.97220900	0.25978700	-1.73264100
C	1.10531000	1.82305900	-0.27702800
N	0.60925100	-0.54452000	0.22852500
H	1.64986900	2.47223900	-0.96132800
H	1.47486000	1.98637800	0.73941200
S	-0.63869400	2.29982700	-0.37489000
H	1.39247300	-0.14459300	1.93562600
H	0.78176300	-1.50868600	-0.05068600
C	4.52751000	-1.38826900	-0.12137400
H	4.83928200	-1.54113200	-1.15402500
H	4.61006200	-2.31185300	0.44427100

H	5.12184100	-0.60153200	0.34172900
H	-3.27652200	1.89751800	0.37668800
C	-1.44115900	-0.74302800	-1.97593800
C	-2.08506900	-2.48253900	-0.32760000
H	-2.25665500	-1.09694700	-2.61009700
H	-0.54695800	-1.32349600	-2.21782400
H	-1.25891700	0.30611300	-2.20370200
H	-1.13959600	-3.02751100	-0.40027300
H	-2.74003500	-2.83220400	-1.12769800
H	-2.55010000	-2.71399800	0.62939600
O	1.49448700	0.08303200	2.88265400
H	1.29897700	1.02639000	2.92065000
H	-0.46712300	-0.76260100	2.40605500

Methylcysteinate transition state 1

Sum of electronic and thermal Free Energies (a.u)= -1259.348291; iFreq (cm-1)= -102.6715

C	3.13490700	-0.34200300	-0.22384200
C	2.80414900	-0.48607100	1.24656300
C	1.28227500	-0.29316800	1.29875300
H	3.29881300	0.34112700	1.76880200
H	0.92012300	0.05526100	2.26587500
C	1.13136300	0.83102800	0.24890000
C	2.08687800	0.51983700	-0.93668300
O	4.10533000	-0.81703900	-0.76878800
O	1.01950900	1.99992700	0.67079700
C	-3.01487400	-0.54258800	-0.60882600
C	-1.50469400	-0.70992000	-0.48466400
O	-3.76056700	-1.48378100	-0.75998400
O	-3.42064900	0.71424200	-0.51518900
H	-1.19582500	-1.29446500	-1.35366700
C	-1.20602700	-1.49620200	0.79078700
N	-0.75899400	0.54413400	-0.45448900
H	-1.74918100	-2.44123800	0.78757800
H	-1.50574500	-0.92303000	1.67332900
S	0.54969200	-1.91899600	0.92082300
H	-1.18652700	1.17399200	0.22510600
H	-0.78913300	1.01222900	-1.35890900

C	-4.84038000	0.93634100	-0.60942500
H	-5.20605600	0.58757900	-1.57454200
H	-4.97314700	2.01026900	-0.51563700
H	-5.35048500	0.41445300	0.19942800
H	3.16942600	-1.42621900	1.66146100
C	1.51106500	-0.25270400	-2.12692200
C	2.76570300	1.79724300	-1.43860900
H	2.33401900	-0.53589400	-2.78733700
H	0.82998700	0.38400400	-2.69547500
H	0.99012300	-1.15786100	-1.81857800
H	2.01694500	2.47009000	-1.86183600
H	3.48458300	1.53875300	-2.21839300
H	3.28791000	2.32045700	-0.63585700
O	-1.11429700	2.01196300	2.48933100
H	-1.10664500	1.11739900	2.84739600
H	-0.28810000	2.07482900	1.97399000

Methylcysteinate transition state 2

Sum of electronic and thermal Free Energies (a.u)= -1259.347375; iFreq (cm-1)= -596.3485

C	-3.04898400	-0.02571100	-0.31332500
C	-2.78790800	0.73918100	0.96572300
C	-1.26574600	0.78784900	1.09472000
H	-3.21576100	0.14741400	1.78155100
H	-0.92593300	0.78594900	2.13089800
C	-0.84185400	-0.55027600	0.43121200
C	-1.80365700	-0.78664600	-0.76720300
O	-4.10544800	-0.05044400	-0.90323800
O	-0.97979400	-1.55778000	1.36348300
C	2.83141600	0.33266000	-0.65758600
C	1.33542900	0.59858600	-0.54341600
O	3.58885700	1.16158900	-1.10099500
O	3.19082100	-0.86458400	-0.23268600
H	0.97656700	0.75568000	-1.55976600
C	1.10353300	1.84796100	0.29404200
N	0.64874000	-0.57564200	0.05817900
H	1.65276000	2.67726600	-0.15002000
H	1.45574200	1.69657200	1.31832900

S	-0.64614600	2.30565600	0.29903300
H	1.10314700	-0.78857400	1.01435500
H	0.79368600	-1.40204500	-0.53300600
C	4.59636700	-1.17711400	-0.31564400
H	4.91781300	-1.14470900	-1.35572500
H	4.69350700	-2.18099800	0.08649600
H	5.16488800	-0.46566300	0.28173200
H	-3.28345300	1.71097800	0.96145800
C	-1.35102700	-0.20683900	-2.11804500
C	-2.13484700	-2.26739400	-0.97083100
H	-2.15266000	-0.36800200	-2.84124300
H	-0.46801100	-0.74309500	-2.47678300
H	-1.14084300	0.86119200	-2.07966000
H	-1.21879200	-2.82454600	-1.18446800
H	-2.80498100	-2.36469900	-1.82691600
H	-2.61467700	-2.70524500	-0.09703800
O	1.07266200	-1.23565400	2.56714900
H	1.03395600	-0.44724400	3.11890900
H	-0.10291200	-1.43977800	2.07152400

NMMal

Sum of electronic and thermal Free Energies (a.u)= -398.645133; CCSD(T) (a.u)= -397.8802513

C	-1.11800300	-0.25041100	-0.00000200
C	-0.59418400	-1.65483500	0.00001900
C	0.73247400	-1.59963100	0.00001400
H	-1.24868400	-2.51430500	0.00001600
H	1.45633000	-2.40175800	0.00003400
C	1.14485400	-0.15659800	0.00000100
N	-0.02074300	0.59176600	-0.00002200
O	-2.27095100	0.11824400	-0.00002400
O	2.26801000	0.29182600	-0.00003000
C	-0.10972000	2.04213600	0.00001400
H	0.90176400	2.44256300	0.00074400
H	-0.63711600	2.38321000	0.89074900
H	-0.63609000	2.38340100	-0.89124500

NMMal · H-Cys-OMe MTA

Sum of electronic and thermal Free Energies (a.u)= -1159.762115; CCSD(T) (a.u)= -1157.8777308

C	-3.84760000	-0.53008100	-0.14223900
C	-2.83843100	-1.31079700	-0.94578800
C	-1.58173600	-0.43994500	-0.93779600
H	-3.23412400	-1.42874500	-1.95645200
H	-1.14655700	-0.28564300	-1.92421600
C	-2.04988000	0.88351000	-0.36230600
N	-3.32126000	0.71459300	0.14596900
O	-4.95674700	-0.88575100	0.19344000
O	-1.42544800	1.92127300	-0.32402600
C	3.51536700	0.29709700	-0.02606900
C	2.37333400	-0.67458000	0.22633700
O	3.42953100	1.50098400	0.06850300
O	4.64422400	-0.33319300	-0.33357100
H	2.51243200	-1.51811900	-0.45117300
C	1.04859600	0.01921900	-0.07350400
N	2.52744500	-1.16539500	1.59828800
H	1.02376900	0.35667600	-1.11203200
H	0.91355400	0.87678600	0.58584200
S	-0.32015000	-1.16713500	0.16589900
H	1.98246200	-2.01353200	1.72011000
H	2.15420800	-0.47402200	2.24522800
C	5.80697000	0.49334000	-0.51477200
H	5.64300400	1.19276500	-1.33393700
H	6.61644100	-0.19034700	-0.75437300
H	6.02614000	1.03458300	0.40541600
H	-2.71029900	-2.30428600	-0.51417500
C	-4.07220600	1.75196600	0.83564300
H	-4.59957800	1.30436200	1.67640500
H	-4.78988600	2.21461700	0.15761000
H	-3.37140000	2.49979300	1.19975500