Chemoenzymatic Synthesis of Cryptophycin Anticancer Agents by an Ester Bond Forming NRPS Module

Yousong Ding,[1,2,5,#] Christopher M. Rath,[1,#] Kyle L. Bolduc,[1,2] Kristina Håkansson,[3] and David H. Sherman[1,2,3,4*]

[1]Life Sciences Institute, and Departments of [2]Medicinal Chemistry, [3]Chemistry, [4]Microbiology & Immunology, University of Michigan, Ann Arbor, MI 48109; [5]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.

Supplemental Materials and Methods

1. Protein expression and purification

Proteins were cloned and expressed using standard molecular biology and biochemical techniques. Crp TE was cloned and expressed as previously described.[1,2] CrpD-M2 gene was amplified by PCR with a forward primer with a restriction site of *Bam*HI (CAAGGATCCTTACGTACTACTAATAGCGCA) and a reverse primer with a restriction site of *Xho*I (ATGCTCGAGTAGTTGTTGAATTGGTACTAATGG). The amplicons were purified and digested for cloning into pET28a. The plasmid encoding N-terminal His_6-CrpD-M2 was transformed into *E. coli* BAP1 and grown at 37 °C in TB medium to an OD_{600} of ~0.8 in 2 L flasks. The cultures were cooled to 18 °C, and isopropyl β-D-thiogalactopyranoside was added to a final concentration of 0.2 mM and grown for additional 12-16 h with shaking. The cells were harvested by centrifugation and frozen at -20 °C. Cell pellets were thawed to 4 °C and resuspended in 5X volume of

lysis buffer (20 mM HEPES, pH 7.8, 300 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$, 0.7 mM Tris(2-carboxyethyl) phosphine (TCEP), ~100 mg CelLytic Express (Thermo, Rockfield, IL) before lysis via sonication. Centrifugation at 25,000 x g for 60 min provided clarified lysates. Proteins were purified using Ni-Sepharose affinity chromatography with a gravity column. Briefly, after filtration of the supernatant through a 0.45 μm membrane, the solution was loaded onto a 5 mL HisTrap nickel-nitrilotriacetic acid column. The column was washed with 10 column volumes of buffer A (20 mM HEPES, pH 7.8, 300 mM NaCl, 20 mM imidazole, 1 mM TCEP, 10% glycerol), 10 column volumes of buffer B (20 mM HEPES, pH 7.8, 300 mM NaCl, 50 mM imidazole, 1 mM TCEP, 10% glycerol), and then eluted with buffer C (20 mM HEPES, pH 7.8, 300 mM NaCl, 400 mM imidazole, 1 mM TCEP, 10% glycerol). Protein containing fractions were pooled and desalted with pre-equilibrated PD-10 gravity flow columns in storage buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 1 mM TCEP, 10% glycerol). Fractions were combined, concentrated, frozen, and stored at -80 °C. Purity estimate of 80% was based on SDS-PAGE (**Figure S2**).

2. Analysis of purified CrpD-M2 by FTICR-MS

CrpD-M2 integrity was determined by peptide map fingerprinting. Briefly, CrpD-M2 was reduced (10 mM DTT, 50 ^{o}C 30 min) and digested with trypsin (TPCK modified, Thermo, Rockfield, IL). The sample was desalted with Handee Microspin columns (Thermo, Rockfield, IL) packed with 20 μL of 300 Å polymeric C4 resin (Grace-Vydac, Deerfield, IL). Samples were loaded onto the columns and washed with 30 column volumes of 0.1% formic acid prior to elution with 10 column volumes of 50%

acetonitrile plus 0.1% formic acid. Peptides were then introduced into the FTICR-MS (described below) at a rate of 70 μL/hour with direct infusion. Peaks were identified with the THRASH algorithm as implemented in MIDAS data analysis workstation (National High Magnetic Field Laboratory, Tallahassee, FL). Peaks were matched against a theoretical digest of CrpD-M2 (Protein Prospector, UCSF, San Francisco, CA) with a tolerance of ±20 ppm. Note that the low sequence coverage observed is not surprising due to lack of LC separation, however, sequence coverage is obtained from residue 3-1,922 out of 1,964 suggesting that the purified protein is full-length (**Figure S3**). Identified peptides are presented in **Table S1**.

3. ATP-[^{32}P] PPi exchange assay

The exchange assay for determining A domain substrate specificity was conducted using modified protocol.[3] All acid substrates used in the assay were purchased from Sigma (St. Louis, MO). The reaction mixture (100 μL) contained 75 mM Tris-Cl, pH 7.5, 10 mM Mg Cl_2, 5 mM TCEP, 5 mM ATP, 1 mM tetrasodium pyrophosphate (PP_i), 5 mM free acid substrate, and 0.5 μCi tetrasodium [^{32}P] PP_i (Perkin Elmer, Boston, MA). The ATP-PP_i exchange was initiated by adding 1 μM of CrpD-M2 and allowed to proceed for 10 minutes at room temperature. The reaction was then terminated by the addition of cold charcoal solution (500 μL, 1.6% w/v activated charcoal, 0.1 M tetrasodium pyrophosphate, and 5% perchloric acid in water). Free [^{32}P]-PP_i was removed by centrifugation of the sample, and washing the charcoal pellet twice with wash buffer (0.1 M tetrasodium pyrophosphate, and 5% perchloric acid in water). The charcoal was finally resuspended in water (500 μL) and the bound radioactivity was

determined by scintillation counting on a Beckman LS6500 (Fullerton, CA). All experiments were carried out in duplicate for each substrate with a negative control without enzyme.

4. Biochemical reactions and LC FTICR-MS^2 analysis

Enzymatic reaction conditions were as follows: 100 mM Tris-Cl, pH 7.5, 10 mM $MgCl_2$, 5 mM ATP, 1 mM TCEP, 500 μM free-acid extender unit, 100 μM SNAC-ABC chain elongation intermediate, 1 mM NADH, 1 mM NADPH, 1 μM CrpD-M2, and 1 μM Crp TE. Reaction components were excluded as appropriate (eg. no enzyme for no enzyme control reactions, no chain elongation intermediate for PP_i adenylation domain assays).

For T-domain active site loading experiments samples were incubated for 60 minutes at room temperature. Reactions were raised to pH 8 by the addition of concentrated Tris-base, after 4X dilution in 100 mM ammonium bicarbonate. Trypsin (TPCK-modified, Thermo, Rockfield, IL) at 1 mg/mL was added to a molar ratio of 10:1 (CrpD-M2:trypsin). Samples were incubated at 37 °C for 15 minutes, prior to addition of 10% formic acid to pH 4. Reactions were stored at -80 °C until analysis. The CrpD-M2 T domain active site (QLVEIFQEVLNLPSIGIHDNFFSLGGHSLLAVR) was first identified by accurate mass using LC FTICR-MS (**Figure S3B**). After the tryptic peptide's retention time (80-81 minutes) and the most abundant charge state (4+, 1001.5 m/z) were identified, online MS^2 was performed using external quadruple isolation and IRMPD[1,2] in the FTICR cell for ion activation. The phosphopantetheine ejection ions specific markers for the post-translational modification were observed at 261.1 m/z (Ppt_1)

and 359.1 m/z (Ppt_2).[1,2] In addition, the apo T-domain, charge-reduced, parent containing the residual phosphate (+80) or dehydroalanine (-18) was also observed. A short sequence of b-ions was observed, and added to the confirmation of peptide identity. Identified parent and product ions are shown in **Table S2**, and the parent ion mass spectrum is shown in **Figure S3B**.

Liquid chromatography of trypsin digested CrpD-M2 was performed on an Agilent 1100 with a Jupiter C18 300 Å 1 x 150 mm column (Phenomenex, Torrence, CA) at a rate of 75 µL/min using a column heater at 50 °C. Twenty µL of sample was injected followed by an LC gradient: 2% solvent B 0-20 min, to 60% solvent B at 90 min, to 98% solvent B 105'-108', back to 2% B at 110', and equilibrated at 2% B for 10 minutes. Mobile phase A was water with 0.1% formic acid, and mobile phase B consisted of acetonitrile with 0.1% formic acid.

FTICR-MS (APEX-Q with Apollo II ion source and actively shielded 7 T magnet; Bruker Daltonics, Ballerica, MA) was conducted in positive ion mode from m/z 200–2,000. Electrospray was conducted at 3,000-4,000 V 16-30 scan per spectra utilizing 1 s external ion accumulation in the hexapole prior to analysis in the FTICR using a loop value of 4 for direct inject samples. For online LC FTICR-MS external ion accumulation time was set to 0.33 s with 1 scan per spectra and 128 K signal detected. Collision cell pressure was kept at 5.8e-6 torr and either CID or IRMPD was utilized for MS^2. LC-FTICR MS data was processed in Data Analysis (Bruker Daltonics, Ballerica, MA) and using DECON2LS and VIPER (Pacific Northwest National Labs, Richland, WA) for online THRASH analysis. Protein Prospector (UCSF, San Francisco, CA) was used to

assist with manual assignment of MS^n data. All experiments were performed at least twice to verify the findings.

5. Determination of CrpD-M2 active site loading

The CrpD-M2 T domain active site peptide as identified above, was then loaded with the substrates: L-2HIC, D-2HIC, AKGB, and 2-KIC. Peaks were initially identified through LC FITCR-MS using accurate mass. For online confirmation of peak identity the same LC gradient (described above) was run using a Thermo LTQ linear ion trap (LIT) MS. Product peaks were subjected to MS^2 (phosphopantetheine ejection assay) and MS^3 (apo T-domain, charge-reduced, parent containing the residual dehydroalanine (-18) as precursor) for further confirmation.[1,2] Notably, the phosphopantetheine ejection assay performed well on the low resolution instrument, as previously reported.[4] See **Table S4** for LC FTICR-MS and **Table S5** for LC LIT-MS results. The loading and reaction of substrates on the CrpD-M2 T domain was monitored by looking for mass shifts to the loaded T-domain peptide by LC FITCR-MS. Identified peaks were verified by online MS^2 and MS^3 experiments using a LIT-MS.

Linear ion trap mass spectrometry was performed using an LTQ (Thermo, Rockfield, IL). A number of data independent scans were defined for each sample including: a survey scan, MS^2 fragmentation of the target T-domain peptide in the 3^+ and 4^+ charge state, and MS^3 of the charge reduced Apo - 18 products from the MS^2 scans. An isolation window of 4 m/z was utilized, with normalized collision energy of 40%. Xcaliber (Thermo, Rockfield, IL) was used for initial data processing, followed by Protein Prospector (UCSF, San Francisco, CA) for fragmentation assignments.

6. Cryptophycin MS and MS^2

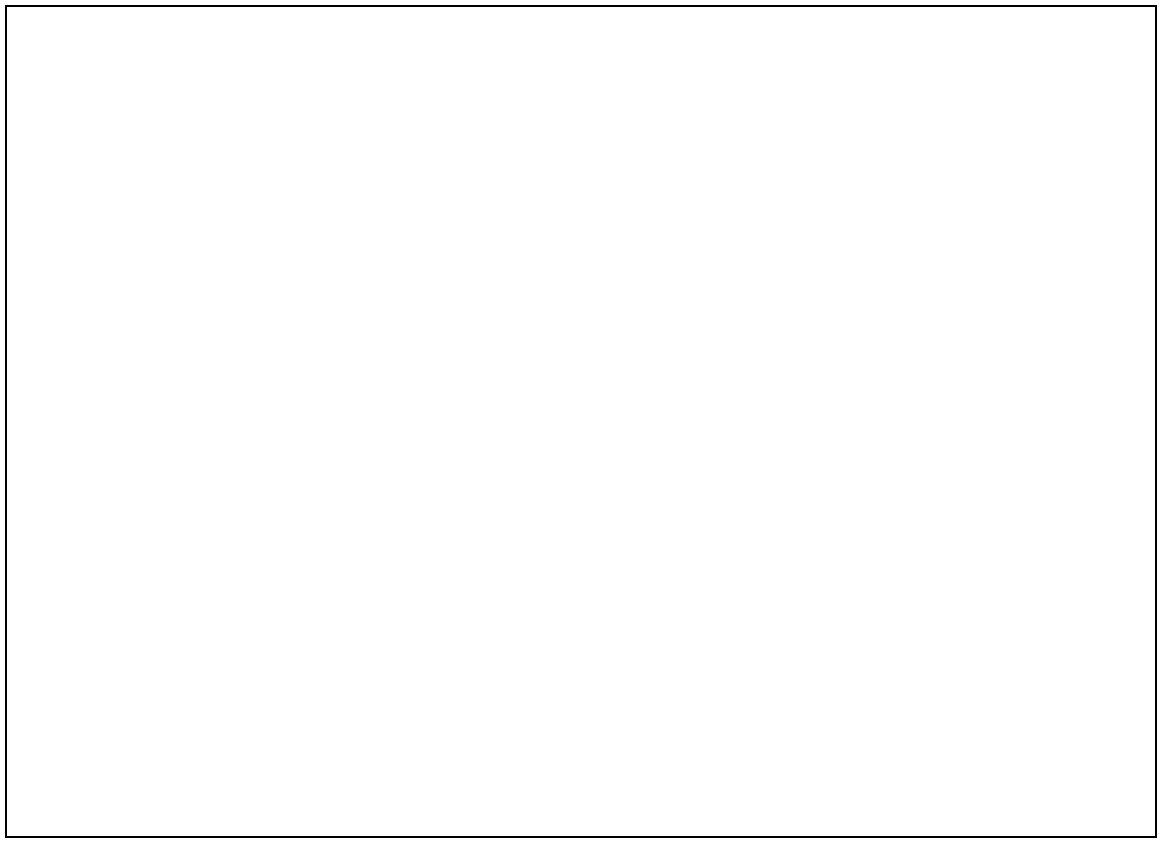
The cryptophycin products were generated by reacting CrpD-M2, SNAC-ABC chain elongation intermediates, unit D extender units, ATP, and Crp TE together as described above. After 60 minutes, the proteins were removed by precipitation with 3 volumes of methanol, pelleted by centrifugation, and the reactions were injected into the LC FTICR-MS (described above). Online MS^2 CID spectra were assigned to cryptophycin 3 and 51 (**Figure S6**). Product spectra were interpreted manually using cryptophycin 3 as a standard based on coelution with an authentic standard (**Figure S7**). Cyclic peptide product ion assignment was aided by tools developed in the Dorrestein and Pevzner laboratories.[5,6] Cryptophycin MS^2 CID fragmentation spectra were assigned with the assumption that fragmentation would occur across the most labile bonds in the gas phase—amide and ester linkages.[5,6] This characteristic breaks cryptophycin into the A, B, C, and D unit constituents with associated water (+/- 18), ammonia (- 17) and CO (- 28) based mass shifts. Two key spectral features are present: the +14 Da shift of the CD + H_2O, ABC - 18, and MH^+ ions between cryptophycin 3 (**Figure S6A**) and 51 (**Figure S6B**),—corresponding exactly to the one methyl group difference due to the different unit C moieties. This diagnostic feature helps to validate both spectral assignments.

7. Cryptophycin 3 authentic standard elution

An authentic standard of cryptophycin 3 was used to verify the identity of chemoenzymatically generated cryptophycin 3 by HPLC co-elution analysis. A SHIMADZU LCMS-2010EV system (Columbia, MD) was used for HPLC separation with a detection wavelength of 218 nm. The product separation was carried out with a

Waters XBridge C_{18} (3.5 μm, 2.1 × 150 mm) column (Milford, MA) at a flow rate of 200 μL/min. Solvent B (acetonitrile with 0.1% formic acid) gradually increased from 50% to 99% and solvent A was water with 0.1% formic acid.

8. Synthesis of SNAC-ABC cryptophycin elongation intermediates (Scheme 1).



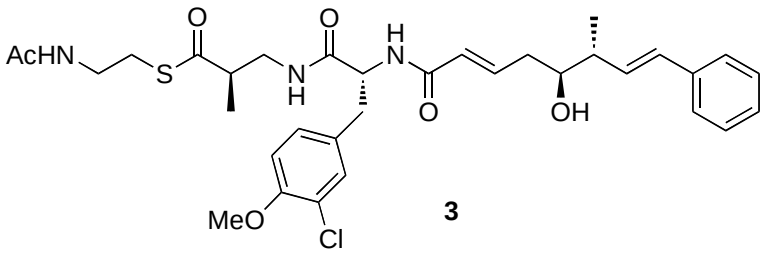
Scheme 1. Synthesis of SNAC-ABC cryptophycin elongation intermediates.

Materials and Methods. Unless otherwise stated, all reactions were performed with no extra precautions taken to exclude moisture or air. Boc-β-Ala-OH was purchased from Advanced ChemTech (Lousville, KY), (R)-3-(Boc-amino)-2-methylpropionic acid was purchased from Sigma (St. Louis, MO), and Boc-3-amino-2,2-dimethyl-propionic acid was purchased from PolyPeptide (San Diego, CA). Solvents were purchased as ACS

Grade (CH_2Cl_2, DMF, and MeOH) from Sigma (St. Louis, MO) or Thermo (Rockfield, IL) and used as received. *N*-acetylcysteamine (SNAC), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), 4-dimethylaminopyridine (DMAP), triethylamine (TEA), benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP), trifluoroacetic acid (TFA), and all other chemicals were obtained from Sigma (St. Louis, MO) or Advanced ChemTech (Lousville, KY) and used directly.

^{1}H and ^{13}C NMR spectra were recorded on a Varian 500 MHz or a Varian Performa IV 600 MHz instrument. Proton chemical shifts are reported in ppm referenced to the centerline of $CDCl_3$ (7.26 ppm) or CD_3OD (3.31 ppm) as the internal standard, carbon chemical shifts are reported in ppm referenced to the centerline of $CDCl_3$ (77.16 ppm) or CD_3OD (49.00 ppm) as the internal standard. Proton chemical data are described as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad), coupling constant (in Hz), and integration. High-resolution mass spectra were obtained on an Agilent Accurate-Mass Q-TOF 6520 (Santa Clara, CA). Semi-preparative HPLC was performed with a Beckman Coulter System Gold 126 HPLC (Brea, CA) utilizing a Waters Prep C18 column (Milford, CA). Analytical thin-layer chromatography (TLC) was performed on silica gel 60 F_{254} TLC glass plates with a fluorescent indicator from EMD Chemicals (Gibbstown, NJ). EMD Silica Gel 60 *Geduran*® (particle size 0.040 – 0.063 mm, Gibbstown, NJ) silica gel was used for flash chromatography. Visualization was accomplished by UV fluorescence quenching (254 nm) and by *p*-anisaldehyde or potassium permanganate staining.

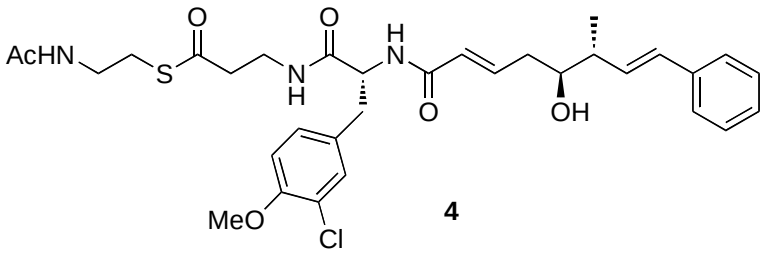
Experimental Procedures



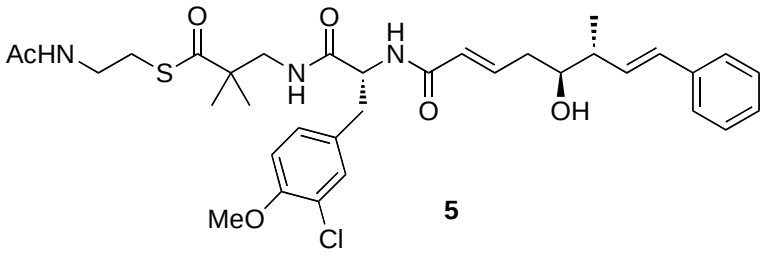
3

(*R*)-*S*-2-acetamidoethyl 3-((*R*)-3-(3-chloro-4-methoxyphenyl)-2-((2*E*,5*S*,6*R*,7*E*)-5-hydroxy-6-methyl-8-phenylocta-2,7-dienamido)propanamido)-2-methylpropanethioate (3). A solution of compound **28** (27 mg, 0.053 mmol) in CH_2Cl_2 (500 μL) was treated with trifluoroacetic acid (250 μL) and stirred at room temperature for 1 hour. Volatiles were then removed under vacuum and residue dissolved in DMF (1 mL). To this solution was added a solution of **30** (10 mg, 0.041 mmol), PyBOP (32 mg, 0.061 mmol), triethylamine (12 μL, 0.081 mmol), and DMAP (1 mg, 0.008 mmol). Reaction stirred for 18 hours at room temperature and then diluted with CH_2Cl_2 (5 mL). Organic layer washed with saturated aqueous NH_4Cl (5 mL), brine (5 mL), and water (5 mL), then dried with Na_2SO_4. Solvent removed under vacuum, residue dissolved in MeOH, and purified by HPLC (XBridge C18 Prep Column, 5 μm, 10 x 250 mm, 3 mL/min, 50:50 to 90:10 MeOH/H_2O over 25 minutes, R_t = 21.1 minutes). Fractions containing the desired compound were concentrated *in vacuo* and dried by lyophilization to recover 4.2 mg of pure **3** as a white solid (13%). ^{1}H NMR (CD_3OD, 600 MHz) δ 7.36 (d, *J* = 7.3 Hz, 2H), 7.27 (m, 3H), 7.18 (m, 1H), 7.13 (dd, *J* = 2.2, 8.4 Hz, 1H), 6.96 (d, *J* = 8.4 Hz, 1H), 6.80 (m, 1H), 6.40 (d, *J* = 15.9 Hz, 1H), 6.24 (dd, *J* = 8.5, 16.0 Hz, 1H), 6.00 (d, *J* = 15.4 Hz, 1H), 4.57 (t, *J* = 8.2 Hz, 1H), 3.83 (s, 3H), 3.64 (m, 1H), 3.41 (m, 1H), 3.33 (m, 1H), 3.16 (m, 1H), 2.99 (m, 4H), 2.82 (m, 2H), 2.36 (m, 4H), 1.91 (s, 3H), 1.15 (d, *J* = 6.8 Hz, 3H), 1.02 (d, *J* = 4.6 Hz, 3H) ppm; ^{13}C NMR (CD_3OD, 150 MHz) δ

202.87, 173.55, 173.47, 170.30, 168.12, 155.39, 143.51, 139.07, 132.55, 132.09, 131.88, 131.58, 129.47, 128.04, 127.13, 126.10, 123.23, 121.75, 113.39, 75.35, 56.59, 56.16, 44.24, 43.28, 39.98, 38.79, 37.95, 29.31, 22.58, 17.49, 15.69 ppm; HRMS (ESI+) *m/z* 666.2386 $[M + Na]^+$ ($C_{33}H_{42}ClN_3NaO_6S$ requires 666.2381).



4

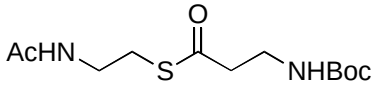


5

(R)-S-2-acetamidoethyl-3-((R)-3-(3-chloro-4-methoxyphenyl)-2-((2E,5S,6R,7E)-5-hydroxy-6-methyl-8-phenylocta-2,7-dienamido)propanamido)-2-methylpropanethioate (4) and ***S*-2-acetamidoethyl 3-((*R*)-3-(3-chloro-4-methoxyphenyl)-2-((2*E*,5*S*,6*R*,7*E*)-5-hydroxy-6-methyl-8-phenylocta-2,7-dienamido)propanamido)-2,2-dimethylpropanethioate (5)** were prepared in a manner analogous to that used to prepare compound **3** and purified by HPLC. Compound **4**: 12%; R_t = 20.5 minutes; 1H NMR (CD_3OD, 600 MHz) δ 7.36 (d, *J* = 7.8 Hz, 2H), 7.25 (m, 3H), 7.17 (dd, *J* = 7.1, 7.4 Hz, 1H), 7.11 (dd, *J* = 2.0, 8.4 Hz, 1H), 6.96 (d, *J* = 8.5 Hz, 1H), 6.80 (m, 1H), 6.41 (d, *J* = 16.0 Hz, 1H), 6.23 (dd, *J* = 8.5, 16.0 Hz, 1H), 5.99 (d, *J* = 15.4 Hz, 1H), 4.55 (t, *J* = 6.0 Hz, 1H), 3.83 (s, 3H), 3.65 (m, 1H), 3.43 (m, 4H), 2.99 (m, 4H), 2.83 (dd, *J* = 8.5, 13.8 Hz, 1H), 2.70 (m, 2H), 2.39 (m, 2H), 2.32 (m, 1H), 1.91 (s, 3H), 1.15 (d, *J* = 6.0 Hz, 3H) ppm; ^{13}C NMR (CD_3OD, 150 MHz) δ 198.43, 173.46,

173.41, 170.29, 168.14, 155.38, 143.53, 139.07, 132.55, 132.09, 131.86, 131.55, 129.47, 128.04, 127.13, 126.08, 123.21, 121.75, 113.36, 75.35, 56.58, 56.12, 44.24, 44.06, 40.00, 38.80, 37.96, 36.53, 29.37, 22.55, 17.50 ppm; HRMS (ESI+) *m/z* 652.2226 $[M + Na]^+$ ($C_{32}H_{40}ClN_3NaO_6S$ requires 652.2224).

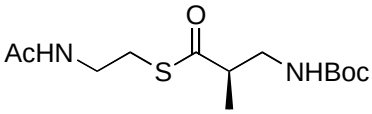
Compound **5**: 8%; R_t = 21.5 minutes; ^1H NMR (CD_3OD, 600 MHz) δ 7.37 (m, 2H), 7.28 (m, 3H), 7.16 (m, 2H), 6.98 (d, *J* = 8.7 Hz, 1H), 6.81, (m, 1H), 6.42 (d, *J* = 15.8 Hz, 1H), 6.24 (dd, *J* = 8.3, 15.9 Hz, 1H), 6.01 (d, *J* = 15.2 Hz, 1H), 4.64 (m, 1H), 3.85 (s, 3H), 3.65 (m, 1H), 3.34 (m, 2H), 3.06-2.92 (m, 4H), 2.82 (m, 1H), 2.36 (m, 4H), 1.94 (s, 3H), 1.31 (s, 6H), 1.17 (m, 3H) ppm; ^{13}C NMR (CD_3OD, 150 MHz) δ 205.97, 173.73, 173.43, 168.18, 155.35, 143.56, 139.07, 132.52, 132.07, 131.85, 131.64, 129.45, 128.01, 127.13, 126.04, 123.25, 121.72, 113.40, 75.35, 56.59, 56.11, 51.98, 48.24, 44.22, 39.80, 38.81, 37.79, 30.72, 29.28, 23.70, 23.55, 22.60, 17.51 ppm; HRMS (ESI+) *m/z* 680.2538 $[M + Na]^+$ ($C_{34}H_{44}ClN_3NaO_6S$ requires 680.2537).



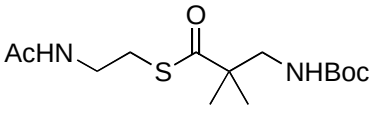
23

2-acetamidoethyl-3-(tert-butoxycarbonylamino)propanethioate (23). A solution of 3-(tert-butoxycarbonylamino)propanoic acid (**20**, 250 mg, 1.32 mmol), EDCI (506 mg, 2.64 mmol), and DMAP (16 mg, 0.13 mmol) in CH_2Cl_2 (4 mL) was prepared and allowed to stir at room temperature for 10 minutes. *N*-acetylcysteamine (169 μL, 1.59 mmol) was then added and reaction stirred at room temperature for 18 hours. Quenched by the addition of saturated aqueous NH_4Cl, the organic layer was removed and aqueous layer extracted with CH_2Cl_2 (2 x 10 mL), washed with saturated aqueous $NaHCO_3$ (10 mL), brine (10 mL), and water (10 mL). The combined organics were then

dried with $MgSO_4$ and solvent removed under vacuum. Flash chromatography (100:0 to 95:5 CH_2Cl_2/MeOH) afforded 230 mg of pure thioester **23** as a clear, colorless oil (60%). TLC R_f = 0.50 (5% MeOH/CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 500 MHz) δ 6.15 (bs, 1H), 4.99 (bs, 1H), 3.41 (q, *J* = 5.0 Hz, 4H), 3.02 (t, *J* = 5.0 Hz, 2H), 2.76 (t, *J* = 5.0 Hz, 2H), 1.96 (s, 3H), 1.41 (s, 9H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 198.61, 170.55, 155.93, 79.69, 44.45, 36.97, 28.94, 28.48, 23.29 ppm; HRMS (ESI+) *m/z* 313.1199 $[M+Na]^+$ ($C_{12}H_{22}N_2NaO_4S$ requires 313.1192).



24

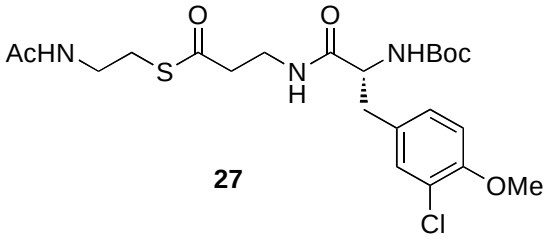


25

(R)-2-acetamidoethyl-3-(tert-butoxycarbonylamino)-2-methylpropanethioate (24) and **2-acetamidoethyl-3-(tert-butoxycarbonylamino)-2,2-dimethylpropanethioate (25)** were prepared in an identical manner to thioester **23**. Compound **24**: 82%; TLC R_f = 0.50 (5% MeOH/CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 500 MHz) δ 6.17 (s, 1H), 4.95 (s, 1H), 3.37 (m, 4H), 2.98 (m, 2H), 2.90 (m, 1H), 1.96 (s, 3H), 1.41 (s, 9H), 1.16 (d, *J* = 5.0 Hz, 3H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 202.85, 170.57, 156.09, 79.68, 49.01, 43.82, 39.10, 28.83, 28.48, 23.29, 15.26 ppm; HRMS (ESI+) *m/z* 327.1364 $[M + Na]^+$ ($C_{13}H_{24}N_2O_4S$ requires 327.1349).

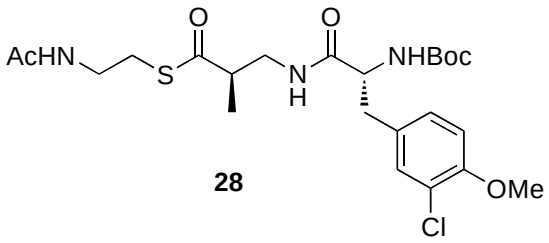
Compound **25**: 78%; TLC R_f = 0.50 (5% MeOH/CH_2Cl_2); ^{1}H NMR ($CDCl_3$, 500 MHz) δ 6.18 (s, 1H), 4.91 (s, 1H), 3.40 (q, *J* = 5.0 Hz, 2H), 3.27 (d, *J* = 5.0 Hz, 2H), 3.00 (t, *J* = 5.0 Hz, 2H), 1.97 (s, 3H), 1.41 (s, 9H), 1.22 (s, 6H) ppm; ^{13}C NMR ($CDCl_3$, 125 MHz) δ 205.81, 170.51, 156.26, 79.60, 51.24, 49.22, 38.92, 28.69, 28.47, 23.32 ppm; HRMS (ESI+) *m/z* 341.1533 $[M + Na]^+$ ($C_{14}H_{26}N_2NaO_4S$ requires 341.1505).

Boc-3-Cl-D-Tyr(Me)-OH (26). Boc-3-Cl-D-Tyr(Me)-OH was used from a preparation for a previous investigation.[7]

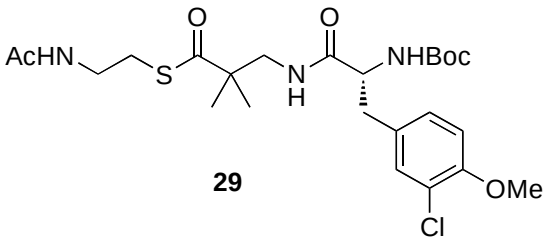


27

(R)-2-acetamidoethyl-3-((R)-2-(tert-butoxycarbonylamino)-3-(3-chloro-4-methoxyphenyl)propanamido)-2-methylpropanethioate (27). Thioester **23** (57 mg, 0.20 mmol) was dissolved in CH_2Cl_2 (1 mL), charged with trifluoroacetic acid (500 μL), and stirred at room temperature for 1 hour. Volatiles removed under vacuum and residue dissolved in DMF (1 mL). To this solution was added a solution of Boc-3-Cl-D-Tyr(Me)-OH (**26**, 79 mg, 0.24 mmol), PyBOP (203 mg, 0.39 mmol), triethylamine (72 μL, 0.52 mmol), and DMAP (2 mg, 0.02 mmol). Reaction stirred for 18 hours at room temperature and then diluted with CH_2Cl_2 (5 mL). Organic layer washed with saturated aqueous NH_4Cl (5 mL), brine (5 mL), and water (5 mL), then dried with Na_2SO_4. Solvent removed under vacuum and residue subjected to flash chromatography (100:0 to 95:5 CH_2Cl_2/MeOH) to afford 63 mg of pure compound **27** as a clear, colorless oil (63%). TLC R_f = 0.45 (5% MeOH/CH_2Cl_2); ^{1}H NMR (CD_3OD, 500 MHz) δ 7.24 (bs, 1H), 7.12 (dd, *J* = 2.2, 8.4 Hz, 1H), 6.97 (d, *J* = 8.4 Hz, 1H), 4.16 (dd, *J* = 6.1, 8.8 Hz, 1H), 3.85 (s, 3H), 3.43 (m, 2H), 3.33 (m, 2H), 3.00 (m, 3H), 2.72 (m, 3H), 1.92 (s, 3H), 1.37 (s, 9H) ppm; ^{13}C NMR (CD_3OD, 125 MHz) δ 198.44, 174.16, 173.46, 157.55, 155.34, 131.97, 131.79, 129.90, 123.12, 113.35, 80.68, 57.42, 56.59, 44.09, 40.00, 38.26, 36.51, 29.37, 28.63, 22.53 ppm; HRMS (ESI+) *m/z* 524.1585 $[M + Na]^+$ ($C_{22}H_{32}ClN_3O_6S$ requires 524.1593).



28

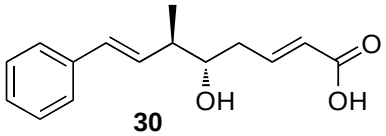


29

(*R*)-2-acetamidoethyl 3-((*R*)-2-(*tert*-butoxycarbonylamino)-3-(3-chloro-4-methoxyphenyl)propanamido)-2-methylpropanethioate (28) and **(*R*)-2-acetamidoethyl 3-(2-(*tert*-butoxycarbonylamino)-3-(3-chloro-4-methoxyphenyl)propanamido)-2,2-dimethylpropanethioate (29)** were prepared in a manner analogous to that used to prepare compound **27**. Compound **28**: 96%; ^{1}H NMR (CD_3OD, 500 MHz) δ 7.25 (bs, 1H), 7.13 (dd, *J* = 2.2, 8.5 Hz, 1H), 6.98 (d, *J* = 8.4 Hz, 1H), 4.18 (dd, *J* = 6.2, 8.6 Hz, 1H), 3.85 (s, 3H), 3.46 (dd, *J* = 7.0, 13.5 Hz, 1H), 3.33 (m, 2H), 3.17 (dd, *J* = 6.5, 13.5 Hz, 1H), 2.98 (m, 3H), 2.84 (q, *J* = 6.9 Hz, 1H), 2.73 (dd, *J* = 8.9, 13.7 Hz, 1H), 1.92 (s, 3H), 1.38 (s, 9H), 1.07 (d, *J* = 7.0 Hz, 3H) ppm; ^{13}C NMR (CD_3OD, 125 MHz) δ 202.84, 174.34, 173.46, 157.52, 155.35, 131.95, 131.81, 129.88, 123.15, 113.37, 80.67, 57.54, 56.60, 43.20, 39.97, 38.27, 29.30, 28.65, 22.56, 15.67 ppm; HRMS (ESI+) *m/z* 538.1748 $[M + Na]^+$ ($C_{23}H_{34}ClN_3O_6S$ requires 538.1749).

Compound **29**: 89%; ^{1}H NMR (CD_3OD, 500 MHz) δ 7.26 (bs, 1H), 7.14 (dd, *J* = 2.2, 8.4 Hz, 1H), 6.97 (d, *J* = 8.4 Hz, 1H), 4.24 (dd, *J* = 5.9, 9.2 Hz, 1H), 3.85 (s, 3H), 3.33 (m, 4H), 2.99 (m, 3H), 2.73 (dd, *J* = 9.3, 13.8 Hz, 1H), 1.93 (s, 3H), 1.37 (s, 9H), 1.19 (s, 3H), 1.16 (s, 3H) ppm; ^{13}C NMR (CD_3OD, 125 MHz) δ 206.00, 174.53, 173.44, 157.56, 155.33, 131.94, 131.88, 129.86, 123.16, 113.38, 80.66, 57.55, 56.60, 52.03, 48.13, 39.79, 38.08, 29.30, 28.65, 23.71, 23.56, 22.59 ppm; HRMS (ESI+) *m/z* 552.1909 $[M + Na]^+$ ($C_{24}H_{36}ClN_3O_6S$ requires 552.1906).

Synthesis of cryptophycin unit A (30)



30

Cryptophycin unit A was synthesized previously according to Eggen et al.[8]

Supplemental Tables and Figures

#	Observed	Expected	dPPM	Start	Stop	Sequence
1	1,493.82	1,493.82	1	3	16	TTNSALSLPPIQPR
2	2,760.52	2,760.48	14	112	136	ESVLHQQAQLAAITPFDLETAPLIR
3	1,410.72	1,410.70	14	242	254	GTTQSFSLNTDLK
4	2,227.18	2,227.15	15	263	282	NSGTTLFMTLHAAFATLLYR
5	2,983.41	2,983.46	-16	335	359	ETTLEAYEHQDVPFEQVVEVLQPQR
6	2,158.00	2,158.03	-13	432	450	MTAHFQNLCSAIVENPQQK
7	2,522.25	2,522.26	-4	578	599	IEQALQTAKGVEDCYVMVRNQK
8	2,784.30	2,784.32	-8	765	789	GITYINSDGSEQVQSYAQLLEDAQR
9	2,549.28	2,549.28	2	872	892	KWSQNNLNDDNFKLETIESLQK
10	3,216.50	3,216.46	11	893	921	FSTDKDKDYYNAQPEDLALFMLTSGSTGMSK
11	1,939.99	1,940.03	-20	1269	1284	SLLKQRFECGEFKSLR
12	2,034.06	2,034.02	19	1323	1339	TLTLIFTDNLGWQQDNR
13	2,571.28	2,571.27	2	1384	1404	QNSQVISQILHLWNYNEQTEK
14	2,399.29	2,399.30	-4	1436	1456	QQAVKLLWIANQSQLVHPTDK
15	1,862.01	1,862.01	0	1441	1456	LLWIANQSQLVHPTDK
16	2,205.17	2,205.12	21	1513	1531	NRERFVSGLEPVDMTAKEK
17	1,392.67	1,392.70	-21	1517	1529	FVSGLEPVDMTAK
18	2,226.15	2,226.14	4	1625	1644	TQLDGVFHMAGIIQETPIEK
19	2,354.28	2,354.28	1	1815	1834	FGIPNQINFVQLEQIPLTQR
20	2,075.06	2,075.07	-1	1840	1858	EQIAAIYGGLNTSEQTKPR
21	1,725.96	1,725.94	9	1908	1922	KNLPLATLFQNPTIER

Table S1. CrpD-M2 peptides identified by accurate mass peptide mass fingerprinting using direct injection FTICR-MS. All masses given are monoisotopic and deconvoluted in Da. Mass error in ppm, start and stop sites, as well as assigned sequence are all provided.

#	m/z mon	z	Obsd.	Theo.	dPPM	ID
1	1,248.3	3	3,741.98	3,741.97	2	Apo + 80
2	1,215.7	3	3,644.04	3,643.97	19	Apo - 18
3	1,335.0	3	4,001.94	4,002.06	-29	Holo 3^+
4	1,001.5	4	4,002.00	4,002.06	-16	Holo 4^+
5	261.1	1	260.12	260.12	9	Ppt_1
6	359.1	1	358.10	358.10	-8	Ppt_2
7	453.2	1	452.23	452.23	0	$b_4 - NH_3$
8	566.3	1	565.32	565.32	4	$b_5 - NH_3$
9	324.2	1	323.19	323.19	-15	$b_3 - NH_3$
10	1,409.8	1	1,408.78	1,408.76	12	$b_{12} -NH_3$
11	788.1	3	2,361.28	2,361.24	18	$b_{14}-NH_3$

Table S2. LC FTICR IRMPD MS^2 verification of CrpD-M2 T domain active site. The observed monoisotopic molecular weight, charge state, deconvoluted monoisotopic mass, and theoretical mass are all provided. Mass errors in ppm and peak identity are also provided.

A Domain	235	236	239	278	299	301	322	330	331	517	Specificity
CrpD-M2-A	V	A	I	F	L	G	S	S	G	K	2KIC/2HIC
PksJ-A1	V	G	W	T	T	A	A	I	C	K	2KIC
BarE-A	V	G	I	L	V	G	G	T	S	K	Trichloro-2KIC
BSLS-A1	G	A	L	M	V	V	G	S	I	K	D-2HIV
BEAS-A1	G	A	L	M	I	V	G	S	I	K	D-2HIV
CseA-A1	V	G	V	W	V	G	T	S	G	K	2-KIC
CseB-A1	V	G	F	W	V	A	V	S	D	K	2-KIV
ENSYN-A1	G	A	L	H	V	V	G	I	C	K	D-2HIV
HctE-A1	V	G	V	W	L	A	L	F	C	K	2KIV
HctF-A1	V	G	V	W	L	A	L	F	C	K	2KIV
KtzG-A	V	T	Y	F	N	G	P	S	G	K	2-KIV
Vlm1-A1	A	A	L	W	I	A	V	S	G	K	2-KIV
Vlm2-A1	V	V	I	W	I	A	E	N	M	K	Pyruvate
BarD	**D**	A	I	L	L	G	G	A	A	K	L-Leucine

Table S3. CrpD-M2 A-domain predicted specificity.

	Rxn	m obs	m theo	dppm	ID	Max Int.	FTICR-t
1	L-2HIC	4,116.21	4,116.12	22	2HIC	2.1E+05	82
2	D-2HIC	4,116.18	4,116.12	15	2HIC	4.3E+05	82
3	KIC	4,114.15	4,114.12	9	KIC	8.8E+04	81
4	KIC+NADPH	4,116.14	4,116.12	5	2HIC	1.5E+05	82
5	KIC+NADH	4,116.20	4,116.12	20	2HIC	7.6E+05	82

Table S4. CrpD-M2 active site bound intermediates identified by accurate mass using LC FTICR-MS. All masses given are monoisotopic, in Da. Mass error is provided in ppm. Product identification is provided as ID. Intensity is the average of peak intensity of the most abundant charge state averaged over the three most intense scans. Retention times are given in minutes, note that due to instrument configuration (dead-volume, actual performance, scan times) LC FTICR and LC LIT–MS retention times differ by approximately 8 minutes, although the order of eluted peaks and relative elution time in relation to total ion chromatogram is maintained. See **Figures S26-30** for associated chromatograms and spectra supporting these assignments.

#	Rxn	Parent Ion	IT-t	ID	Fragment Ions Identified
1	L-2HIC	1030^{4+}	74	2HIC	2HIC-Ppt_2, Apo - 18 (1216^{3+}), Apo + 80 (1249^{3+})
		1373^{3+}	74		Apo - 18 (1824^{2+}), Apo + 80 (1873^{2+})
		1030^{4+}→			Apo - 18 MS^2 product (1215^{3+}) → MS^3 products: b_6,
		1216^{3+}	74		b_7, b_9, b_{10}, b_{11}, b_{12}, b_{13}, y_9, y_{12}, y_{13}, y_{14}
		1373^{3+}→			
		1824^{2+}	74		Apo - 18 MS^2 product (1842^{2+}) → MS^3 products: y_{16}
2	D-2HIC	1030^{4+}	74	2HIC	2HIC-Ppt_1, 2HIC-Ppt_2, Apo - 18 (1216^{3+}), Apo + 80 (1249^{3+})
		1373^{3+}	74		Apo - 18 (1824^{2+}), Apo + 80 (1873^{2+})
		1030^{4+}→			Apo - 18 MS^2 product (1215^{3+}) → MS3 products: b_4, b_6, b_7, b_9, b_{10}, b_{11}, b_{12}, b_{15}, b_{16}, y_4, y_5, y_7, y_8, y_9, y_{10}, y_{11},
		1216^{3+}	74		y_{12}, y_{13}, y_{14}, y_{15}, y_{16}, y_{17}, y_{18}
		1373^{3+}→			Apo - 18 MS^2 product (1823^{2+}) → MS3 products: y_{10},
		1824^{2+}	74		y_{14}, y_{15}, y_{16}, y_{18}
3	KIC	1030^{4+}	74	KIC	KIC-Ppt_1, Apo - 18 (1216^{3+}), Apo + 80 (1249^{3+})
		1373^{3+}	74		
					Apo - 18 MS^2 product (1215^{3+}) → MS3 products: b5,
		1030^{4+}→			b_6, b_7, b_9, b_{10}, b_{11}, b_{12}, b_{13}, b_{14}, y_3, y_4, y_5, y_7, y_8, y_9, y_{11},
		1216^{3+}	74		y_{12}, y_{13}, y_{14}, y_{18}
		1373^{3+}→			Apo - 18 MS^2 product (1823^{2+}) → MS3 products: b_{15},
		1824^{2+}	74		y_9, y_{14}, y_{15}, y_{16}, y_{18}
4	KIC+NADPH	1030^{4+}	74	2HIC	2HIC-Ppt_1, 2HIC-Ppt_2, Apo - 18 (1216^{3+}), Apo + 80 (1249^{3+})
		1373^{3+}	74		2HIC-Ppt_2, Apo - 18 (1824^{2+}), Apo + 80 (1873^{2+})
		1030^{4+}→			Apo - 18 MS^2 product (1215^{3+}) → MS3 products: b_8,
		1216^{3+}	74		b_9, b_{10}, b_{11}, b_{12}, b_{13}, b_{15}, y_3, y_7, y_9, y_{10}, y_{11}, y_{13}, y_{14}, y_{16}
		1373^{3+}→			Apo - 18 MS^2 product (1823^{2+}) → MS3 products: b_{15},
		1824^{2+}	74		y_{14}, y_{15}, y_{16}
5	KIC+NADH	1030^{4+}	74	2HIC	2HIC-Ppt_1, 2HIC-Ppt_2, Apo + 80 (1249^{3+})
		1373^{3+}	4		Apo - 18 (1824^{2+}), Apo + 80 (1873^{2+})
					Apo - 18 MS^2 product (1215^{3+}) → MS3 products: b_4,
		1030^{4+}→			b_5, b_6, b_9, b_{10}, b_{11}, b_{12}, b_{13}, b_{15}, y_4, y_6, y_8, y_9, y_{10}, y_{11},
		1216^{3+}	74		y_{12}, y_{13}, y_{14}, y_{15}, y_{17}, y_{18}
		1373^{3+}→			Apo - 18 MS^2 product (1823^{2+}) → MS3 products: b_7,
		1824^{2+}	74		y_9, y_{11}, y_{14}, y_{16}

Table S5. CrpD-M2 peptides identified by MS^2 and MS^3 LC LIT-MS. All masses given are average, in Da. Mass errors of +/- 1 m/z unit were used for peak assignment. Only Ppt, apo – 18, and apo + 80 ions are shown for MS^2 and only main sequence 1^+ b and y ions are shown as results from MS^3 for clarity. Retention times are given in minutes, note that due to instrument configuration (dead-volume, actual performance, scan times) LC-FTICR and LC-LIT –MS retention times differ by approximately 8 minutes, although the order of eluted peaks and relative elution time in relation to total ion chromatogram is maintained. Product identification is provided as ID. The Ppt_1 ion is the cyclic fragment formed after cleavage at the distal end of the phosphate. Ppt_2 ion is formed when cleavage occurs at the proximal end of the arm, resulting in an ejected ion containing the phosphate. Note that MS^2 of the 3+ or 4+

parent predominantly give phosphopantetheine ejection ions, apo - 18, and apo + 18 charge reduced species. MS^3 fragmentation of the Apo -18 peaks result in a, b and y ion series sufficient for peak identification. See **Figures 31-35** for further information.

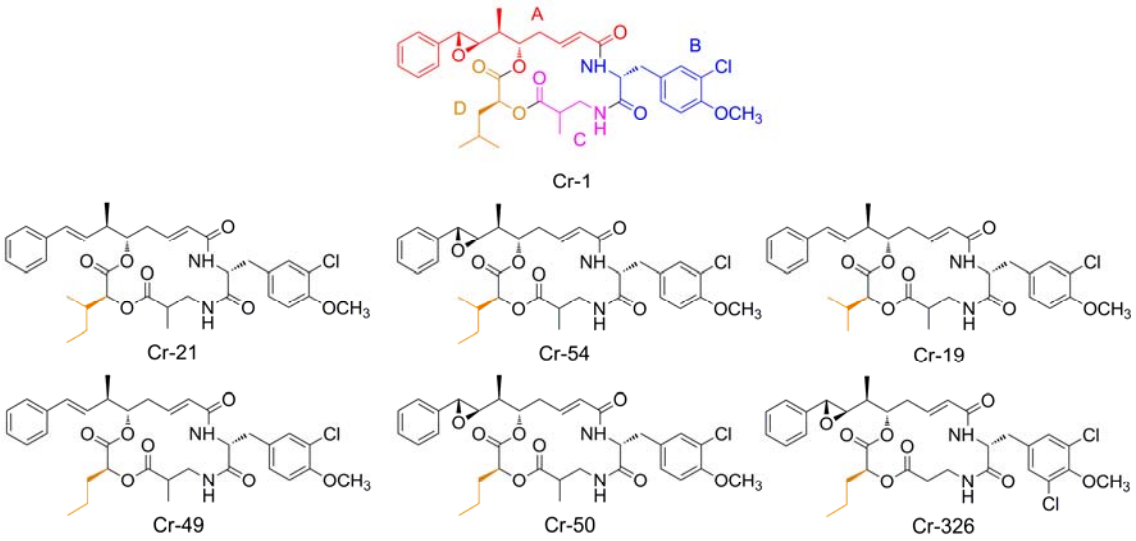


Figure S1. Chemical structures of natural cryptophycin analogs. Cryptopycin-1 is comprised of four units linked in clockwise order of δ-hydroxyoctenoic acid (unit A), 3-chloro *O*-methyl-D-tyrosine (unit B), methyl-β-alanine (unit C), and L-leucic acid (unit D). Several cryptophycin analogs carrying variable unit D moieties are also shown.

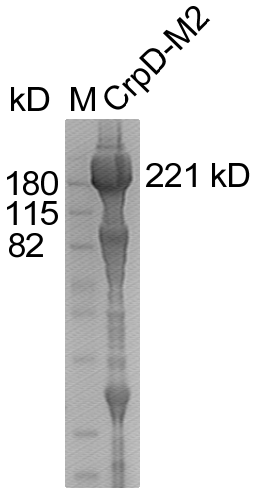


Figure S2. 4-12 % SDS-PAGE analysis of N-terminally His-tagged CrpD-M2 after Ni-NTA resin.

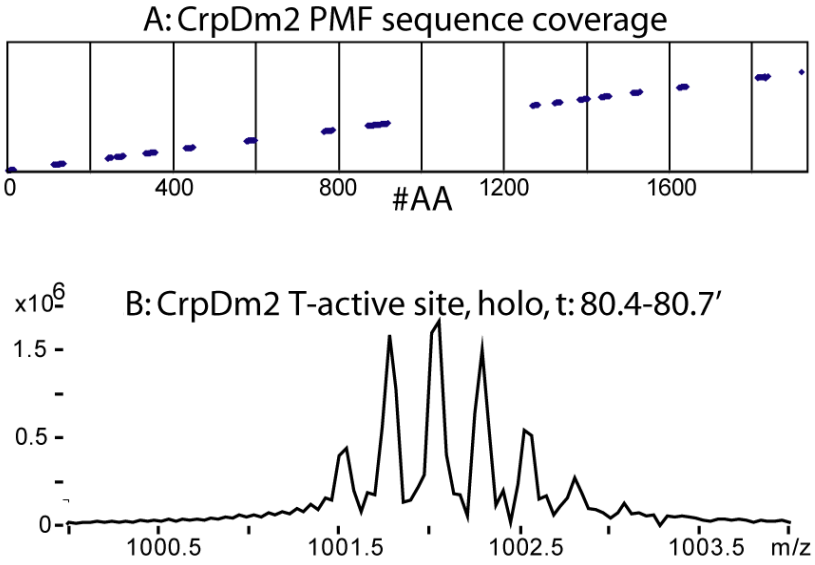
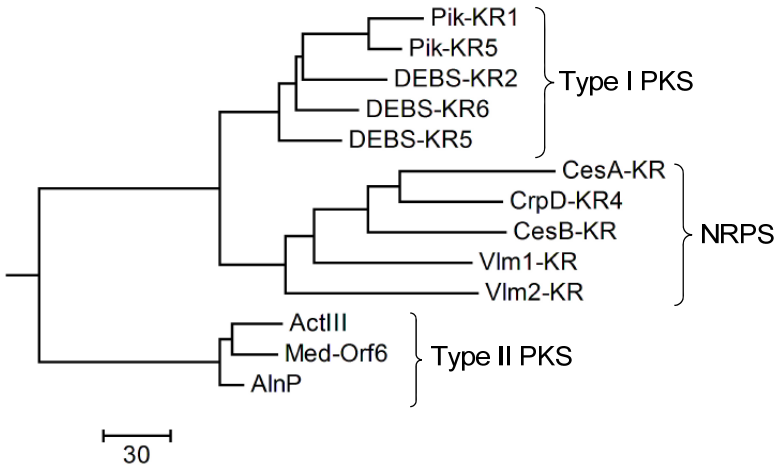


Figure S3. CrpD-M2 characterization. (**A**) Sequence coverage of purified CrpD-M2 by amino acid number, as determined by direct inject FTICR-MS peptide fingerprint mapping. (**B**) Identification of CrpD-M2 T domain active site with LC FTICR-MS as m/z versus absolute intensity. The deconvoluted 4^{+} ion at a monoisotopic mass of 1,001.51 m/z with has an observed molecular weight of 4,002.00 Da as compared to an expected molecular weight of 4,002.06 Da.

A.



B.

```
Region 88-103      Region 134-149
HMAGIIQETPIEKETP...FCSVNGFFGGTNVAAY CrpDm2_KR → ?-type
HLAGIIQEILIEKQTE...TSSARTLLPGMTISAY CesA_KR   → ?-type
HLAGAGNLEEHWSEID...FSSVNSYFGGASFTAY CesB_KR   → ?-type
EDAGYAARQIPQTSVF...CSSSLIGLHSAYKSLY PksJ_KR   → ?-type
HAAATLDDGTVDTLTG...FSSFASAFGAPGLGGY DEBSKR1   → B-type
HTAGALDDGIVDTLTA...FSSVSSTLGIPGQGNY PikKR1    → B-type
HAAGLPQQVAINDMDE...FSSGAGVWGSARQGAY DEBSKR2   → A-type
HAAGVSTSTPLDDLTE...FSSNAGVWGSPGLASY DEBSKR5   → A-type
HTAGVPESRPLHEIGE...FSSGAGVWGSANLGAY DEBSKR6   → A-type
```

Figure S4. (**A)** Phylogenetic analysis of CrpD-M2 KR domain. CrpD-M2 KR domain was grouped with other NRPS KR domains, which grouped into a distinct category compared to type I module PKS and type II discrete PKS β-KR domains. The tree was constructed using the Neighbor-Joining method and was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The scale bar indicates the number of amino acid substitutions per site. Alignment of sequences and the phylogenetic tree were constructed using MEGA 4.0.[9] (**B)** Multiple alignments showing specificity determining regions for PKS β-KR domains. Based on sequence analysis of multiple biosynthetic pathways and product analysis, guidelines were established for determining PKS-KR domain stereospecificity.[10] This findings were later verified with in vitro biochemistry for a series of free PKS enzyme domains with model substrates.[11] When this analysis is applied to CrpD-M2, we can determine that it is neither an A-type nor B-type PKS KR based upon its inability to meet the criteria for either. For example in region 88-103 CrpD-M2 KR does not contain the LDD motif typical of group B, with a strictly conserved L93, D94 or E94 and an invariant D95. In region 134-149 it does not contain the B-group P144 and N148, nor does it contain the A-group W141. Thus this NRPS KR domain appears to fall into a separate KR class based upon primary amino acid sequence as well as analysis of known product structures (**Figure S5**).

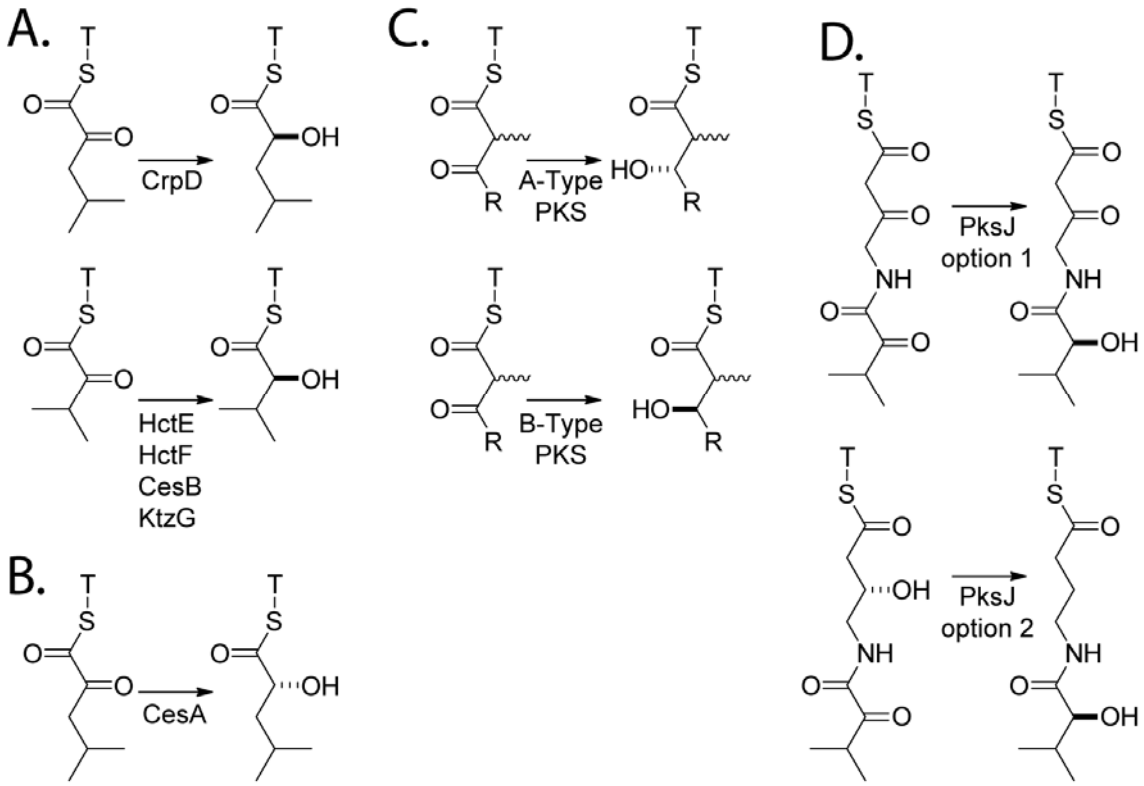


Figure S5. Known ketoreductase catalyzed reactions for T-domain bound substrates in NRP natural products.

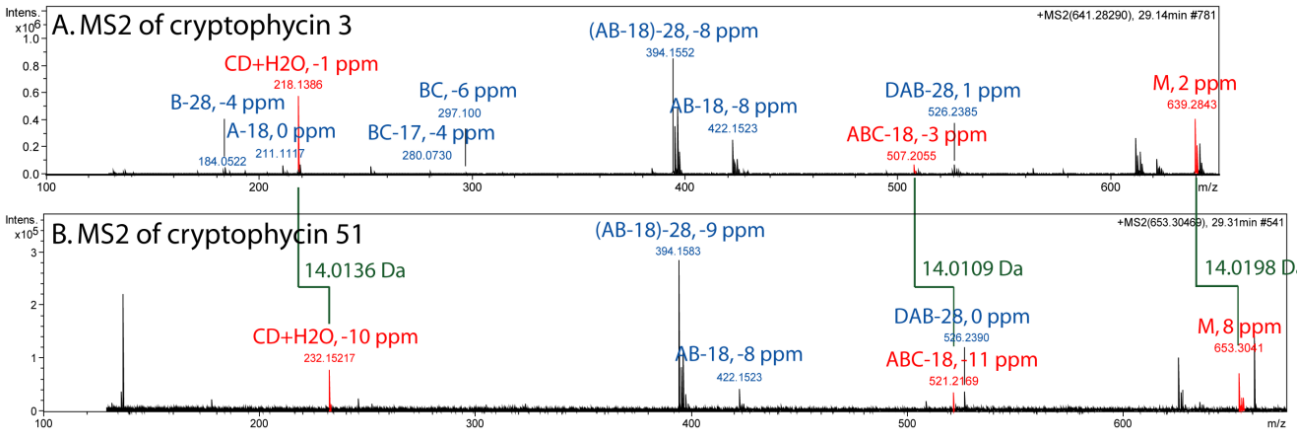


Figure S6. LC FTICR-MS^2 spectra of cryptophycins. Data are presented as m/z versus absolute signal. (**A**) cryptophycin 3 CID MS^2 fragmentation spectra. (**B**) cryptophycin 51 CID MS^2 fragmentation spectra. Assigned ions are denoted in blue. Key ions are noted in red, with the mass shift in green. Mass errors and observed MH$^+$ values are also provided. All assigned peaks were within 15 ppm of the theoretical value. All unit B containing peaks had the predicted ^{37}Cl +2 isotope profile.

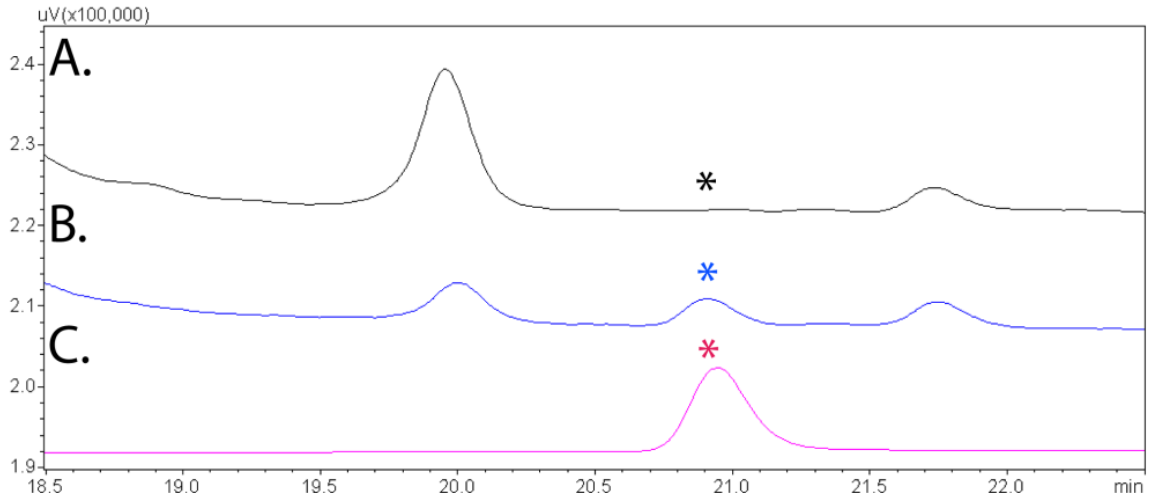


Figure S7. Crp 3 HPLC co-elution analysis with authentic standard. (**A**) A negative control reaction (ATP, L-2HIC extender unit, SNAC-ABC chain elongation intermediate (3-amino-2(*R*)-methylpropionyl, **3**), and boiled enzymes (CrpD-M2 and Crp TE). (**B**) A full reaction (ATP, L-2HIC extender unit, SNAC-ABC chain elongation intermediate (3-amino-2(*R*)-methylpropionyl, **3**), CrpD-M2, and Crp TE). (**C**) Authentic standard of cryptophycin 3.

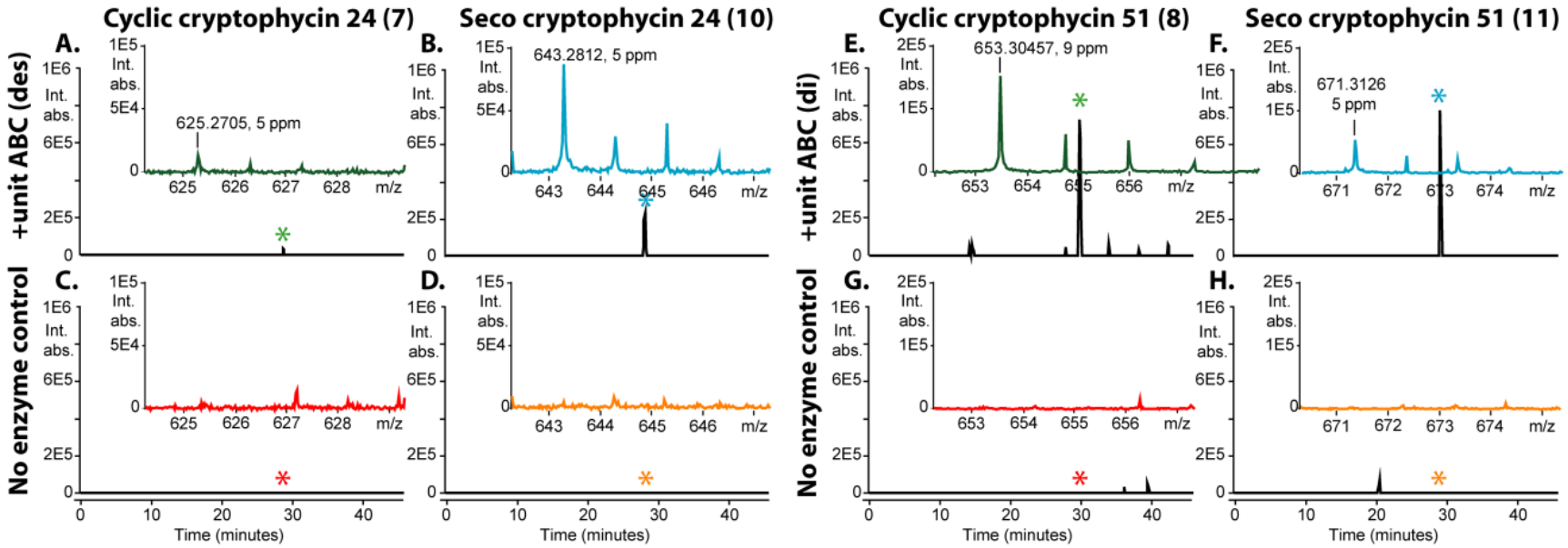


Figure S8 A-D. FTICR MS analysis of cryptophycin products from the reaction of unit C 3-amino-propionyl chain elongation intermediate (**4**) with L-2HIC, ATP, CrpD-M2 and Crp TE. Extracted ion chromatograms are presented at +/- 15 ppm as time versus absolute intensity. Inset mass spectra are time averaged over the one minute elution window corresponding to the asterisk in the extracted ion chromatogram. Inset mass spectra are presented as m/z versus absolute intensity. Monoisotopic $[MH]^+$ mass and the experimental mass error in ppm are also reported. SNAC-ABC intermediate (5) reactions monitoring cyclic (**A**) and linear product formation (**B**) are provided. No enzyme control reactions monitoring cyclic (**C**) and linear product formation (**D**) are provided. **Figure S8 E-H.** FTICR MS analysis of cryptophycin products from the reaction of unit C dimethyl chain elongation unit (**5**) with L-2HIC, ATP, CrpD-M2 and Crp TE. Extracted ion chromatograms are presented at +/- 15 ppm as time versus absolute intensity. Inset mass spectra are time averaged over the one minute elution window corresponding to the asterisk in the extracted ion chromatogram. Inset mass spectra are presented as m/z versus absolute intensity. Monoisotopic mass and the experimental mass error in ppm are also reported. +L-2HIC reactions monitoring cyclic (**E**) and linear product formation (**F**) are provided. No enzyme control reactions monitoring cyclic (**G**) and linear product formation (**H**) are provided.

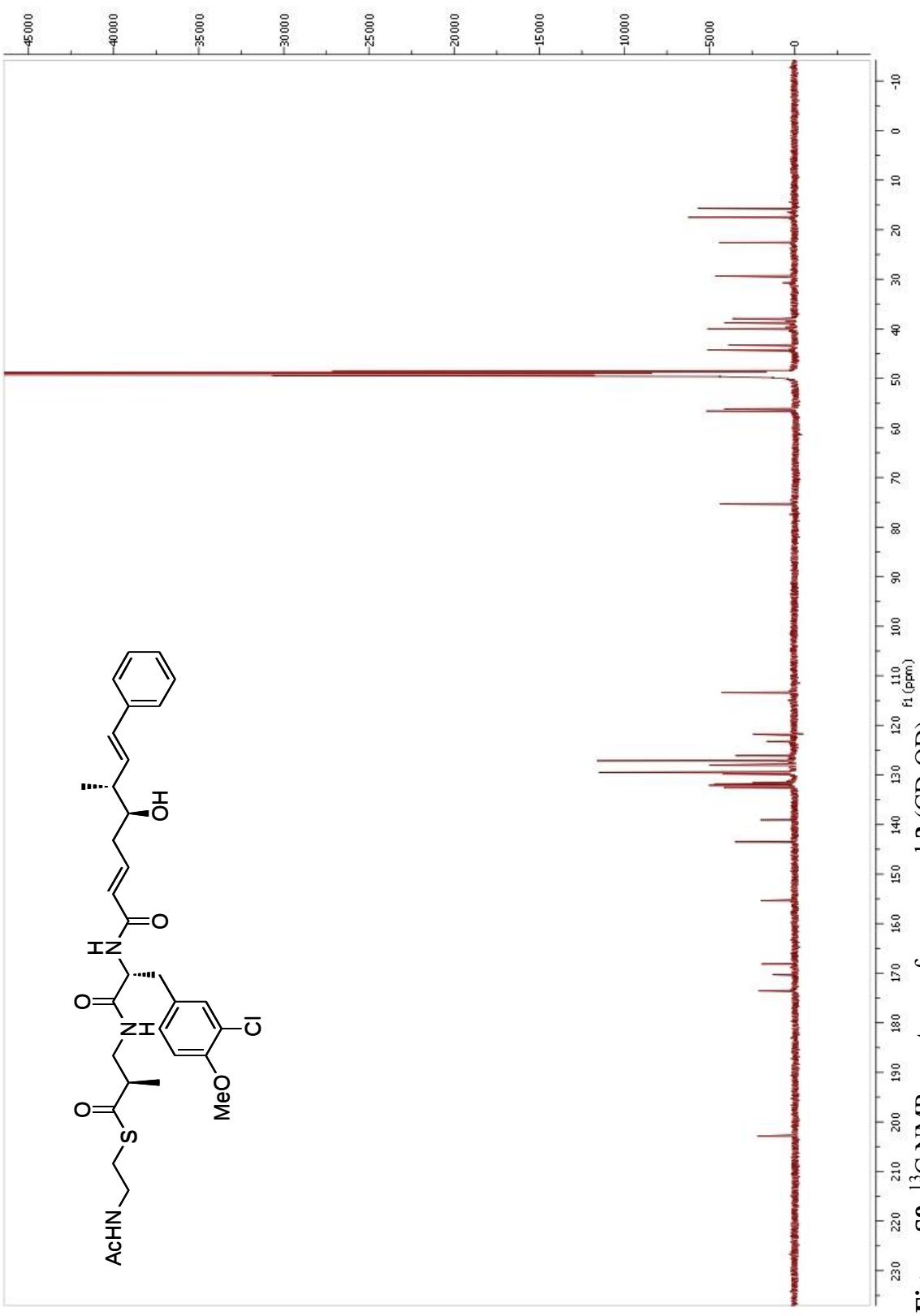


Figure S9. ^{13}C NMR spectrum of compound **3** (CD_3OD).

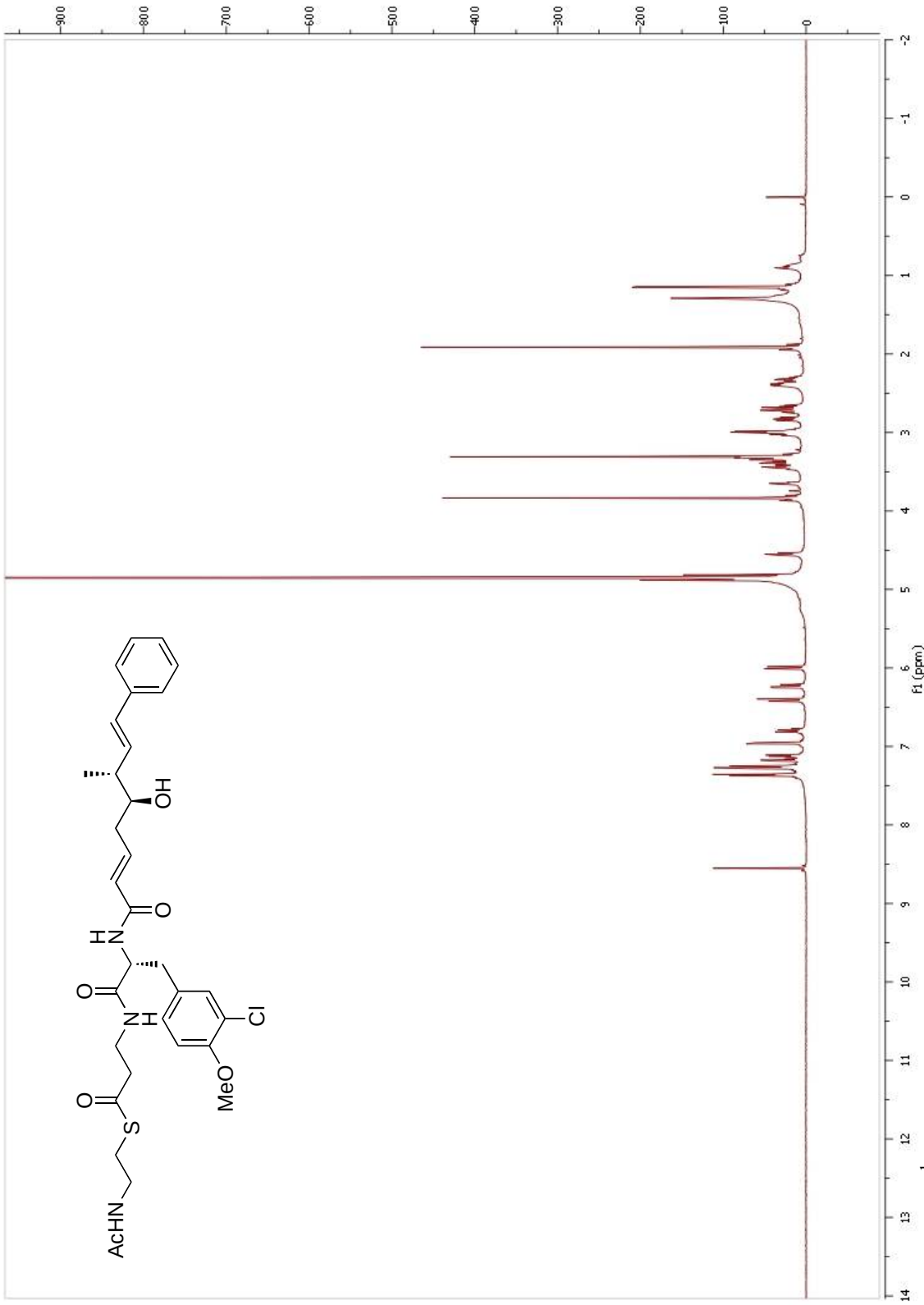


Figure S10. ^{1}H NMR spectrum of compound **4** (CD_3OD).

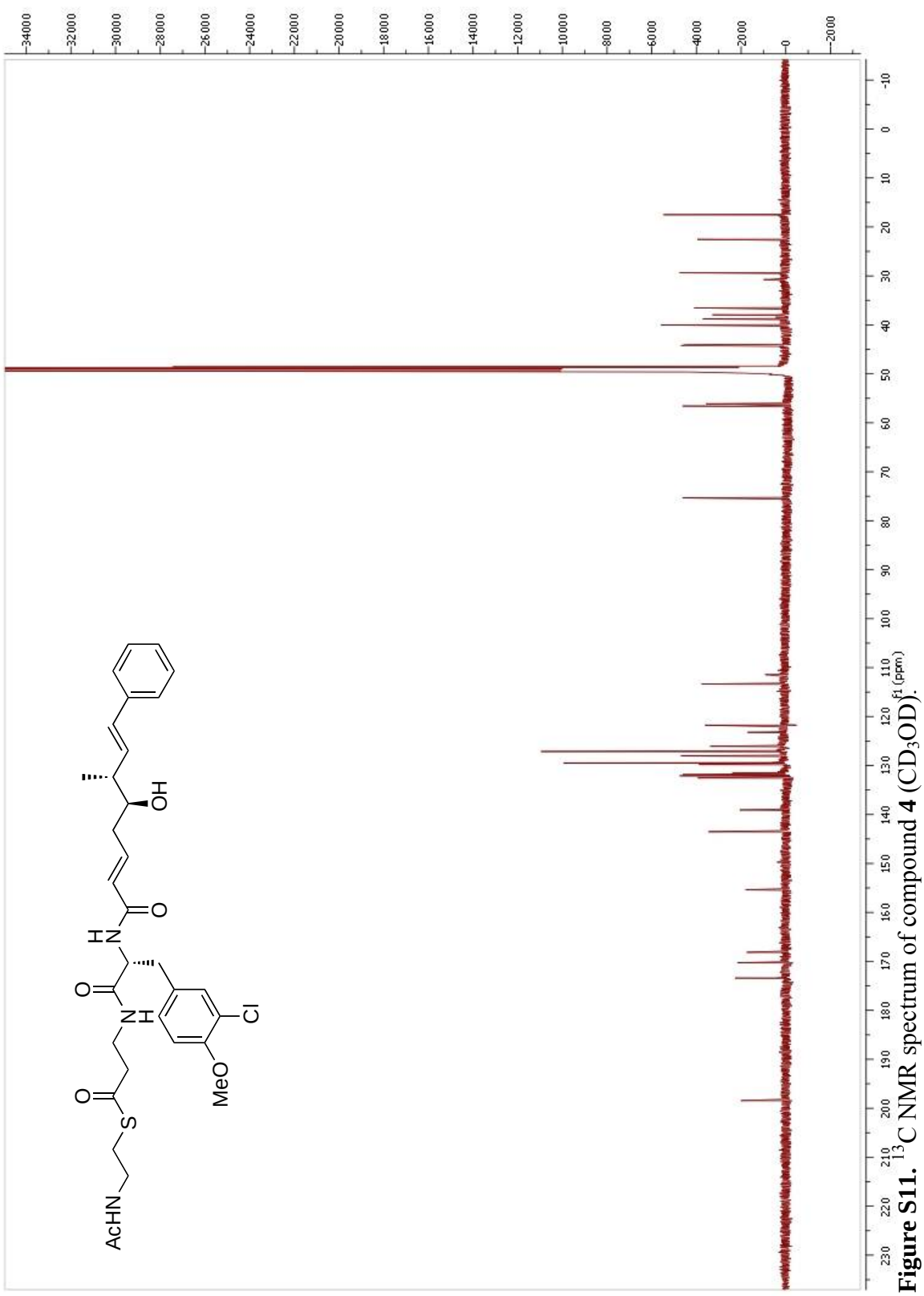


Figure S11. ^{13}C NMR spectrum of compound **4** (CD_3OD).

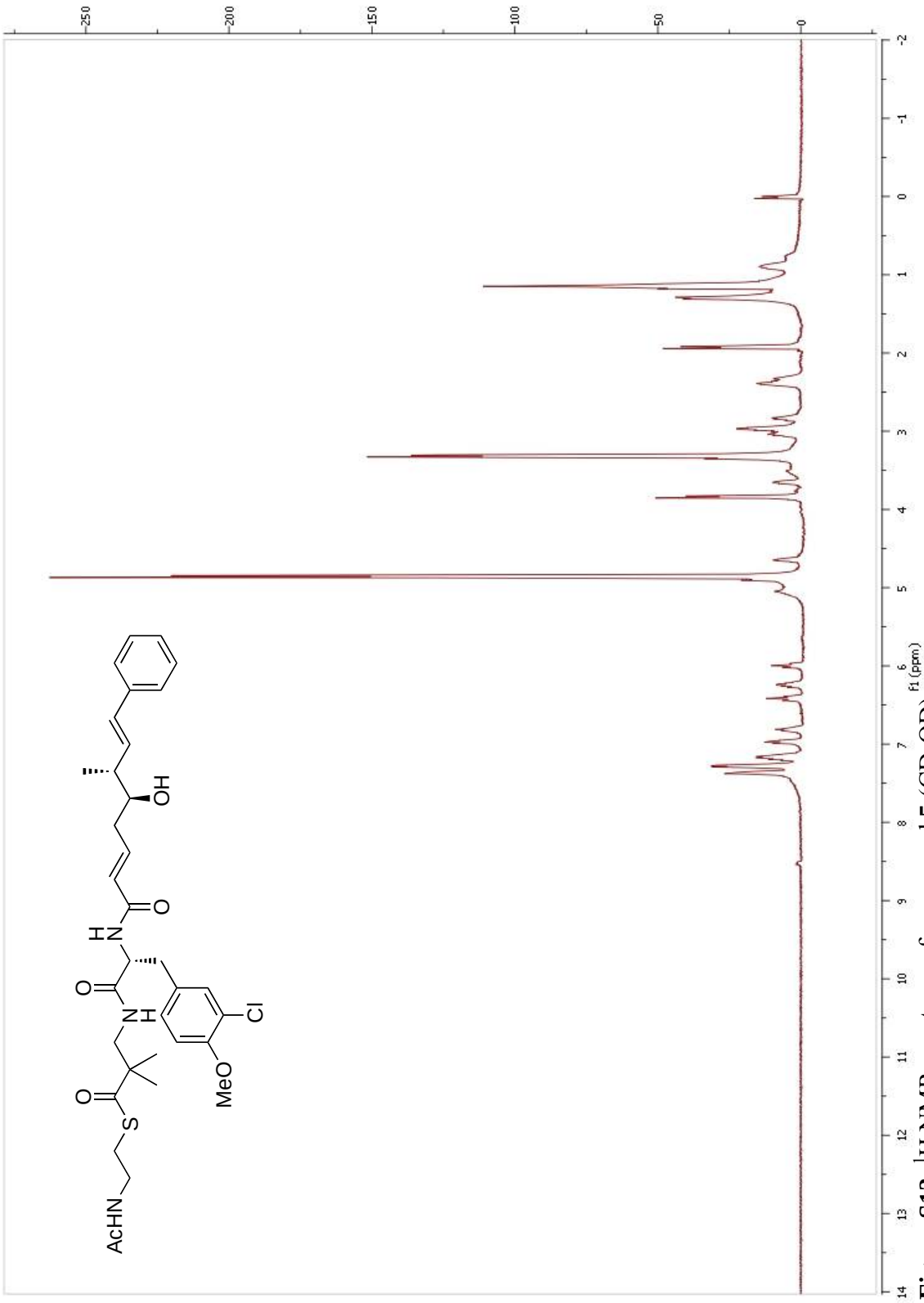


Figure S12. ^{1}H NMR spectrum of compound **5** (CD_3OD).

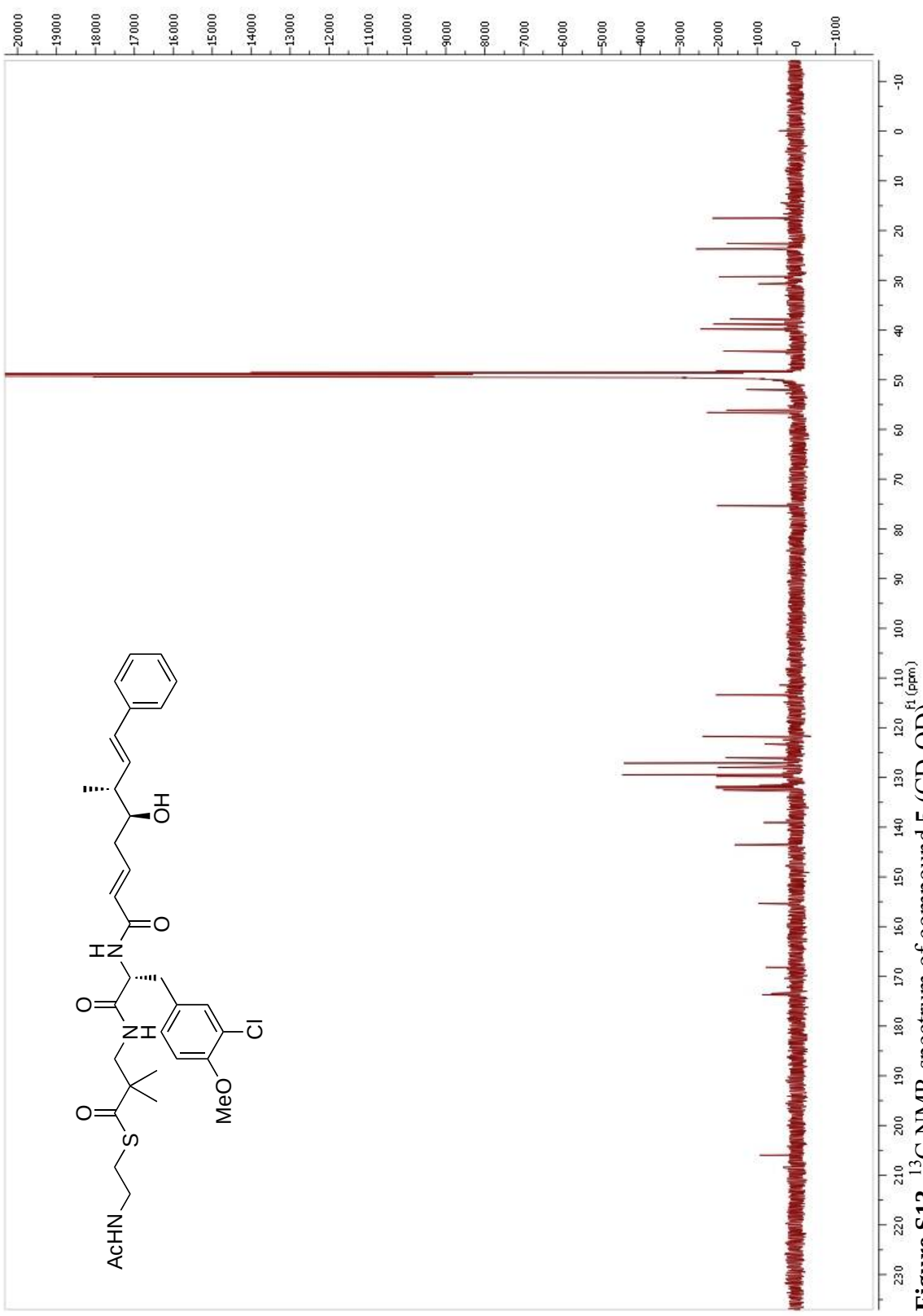


Figure S13. ^{13}C NMR spectrum of compound **5** (CD_3OD).

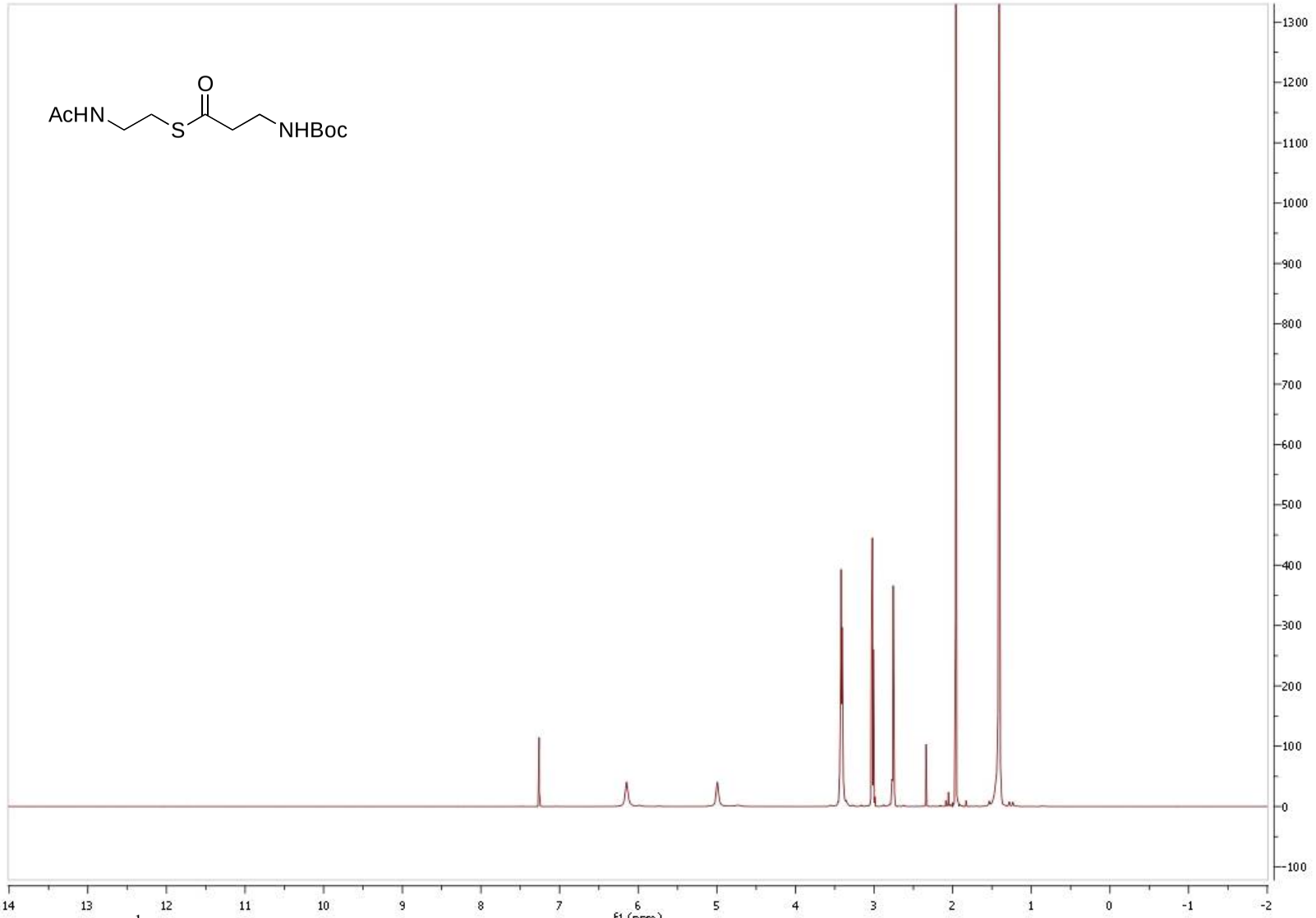


Figure S14. ^{1}H NMR spectrum of compound **23** ($CDCl_3$).

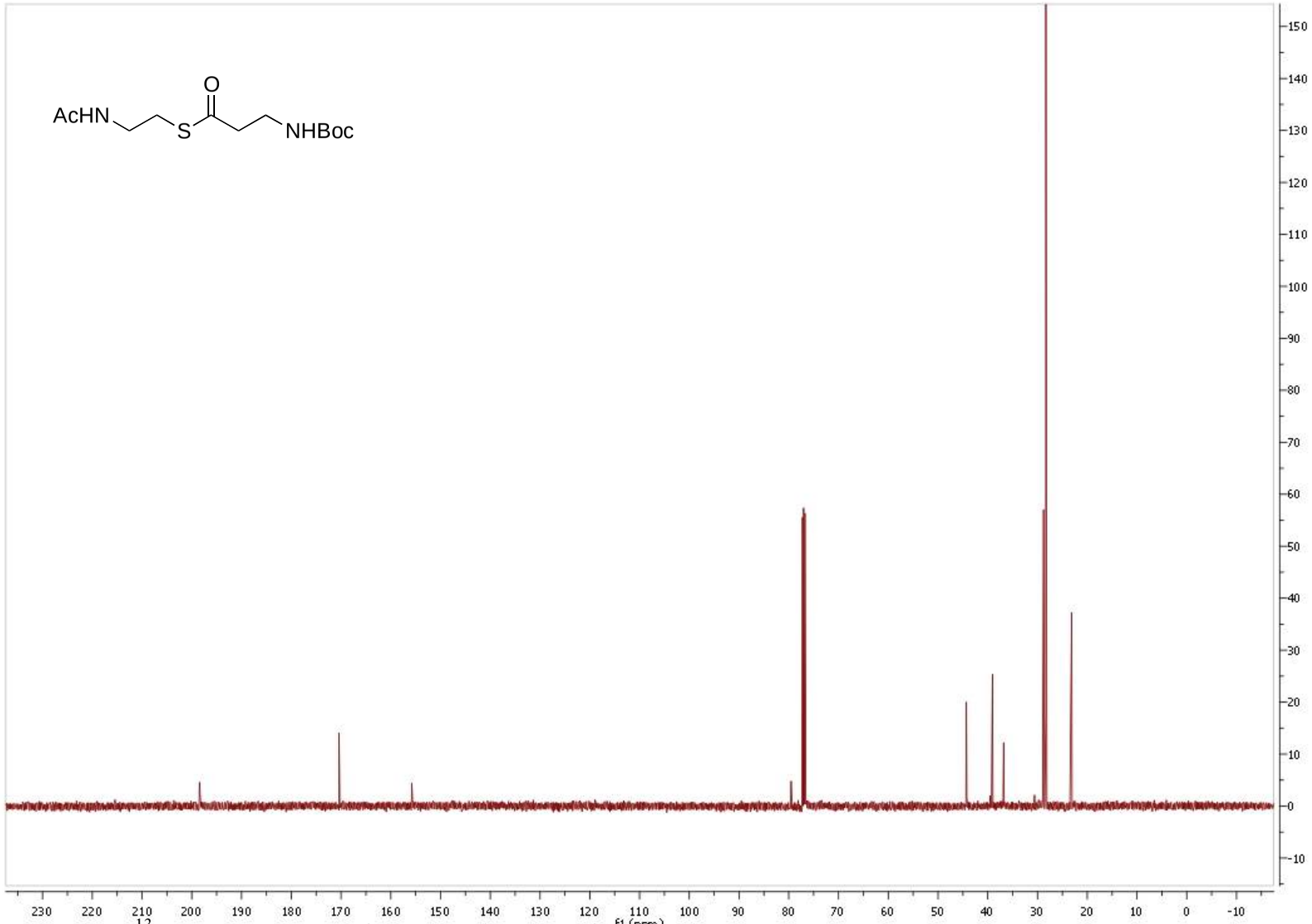


Figure S15. ^{13}C NMR spectrum of compound **23** ($CDCl_3$).

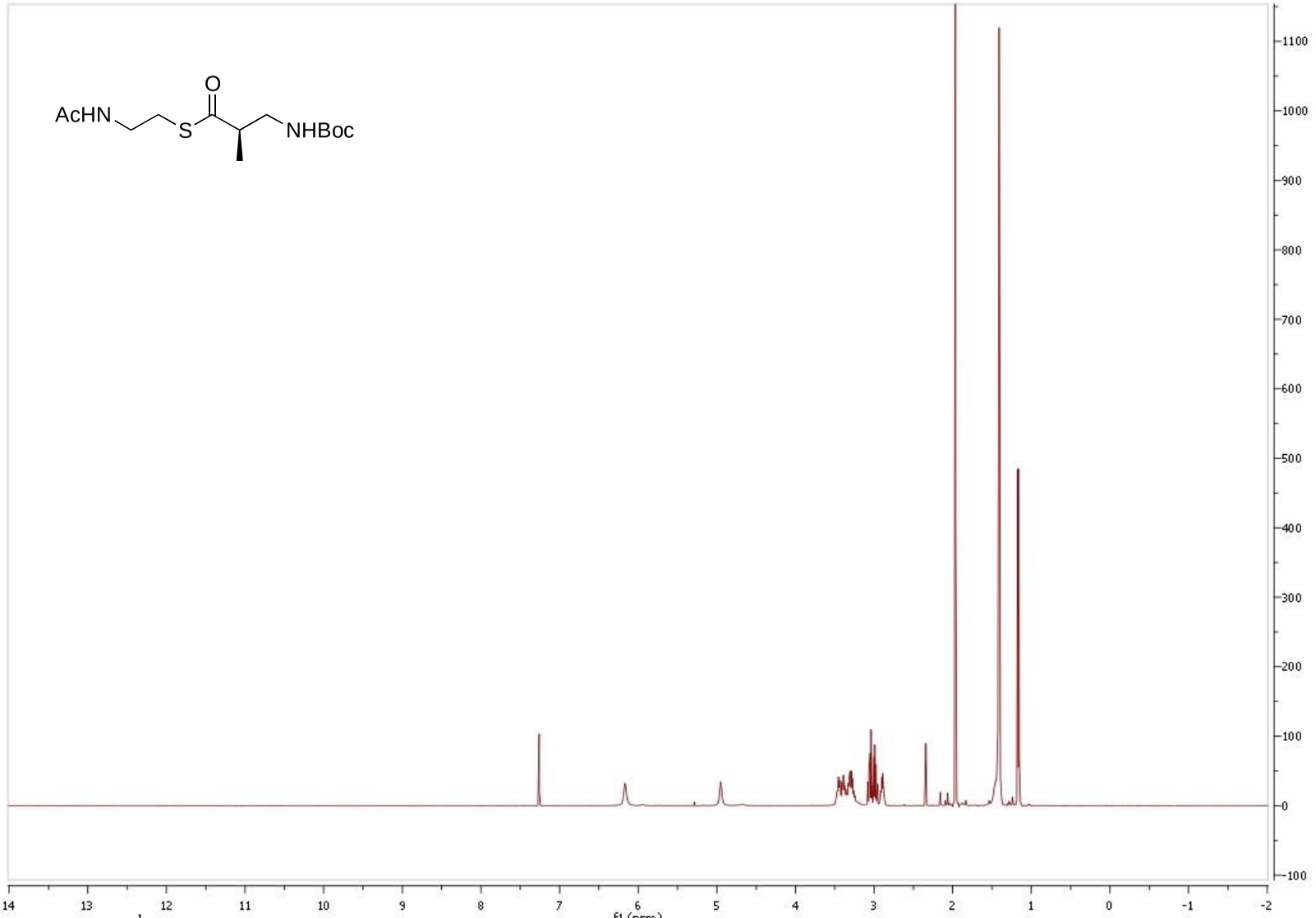


Figure S16. ^{1}H NMR spectrum of compound **24** ($CDCl_3$).

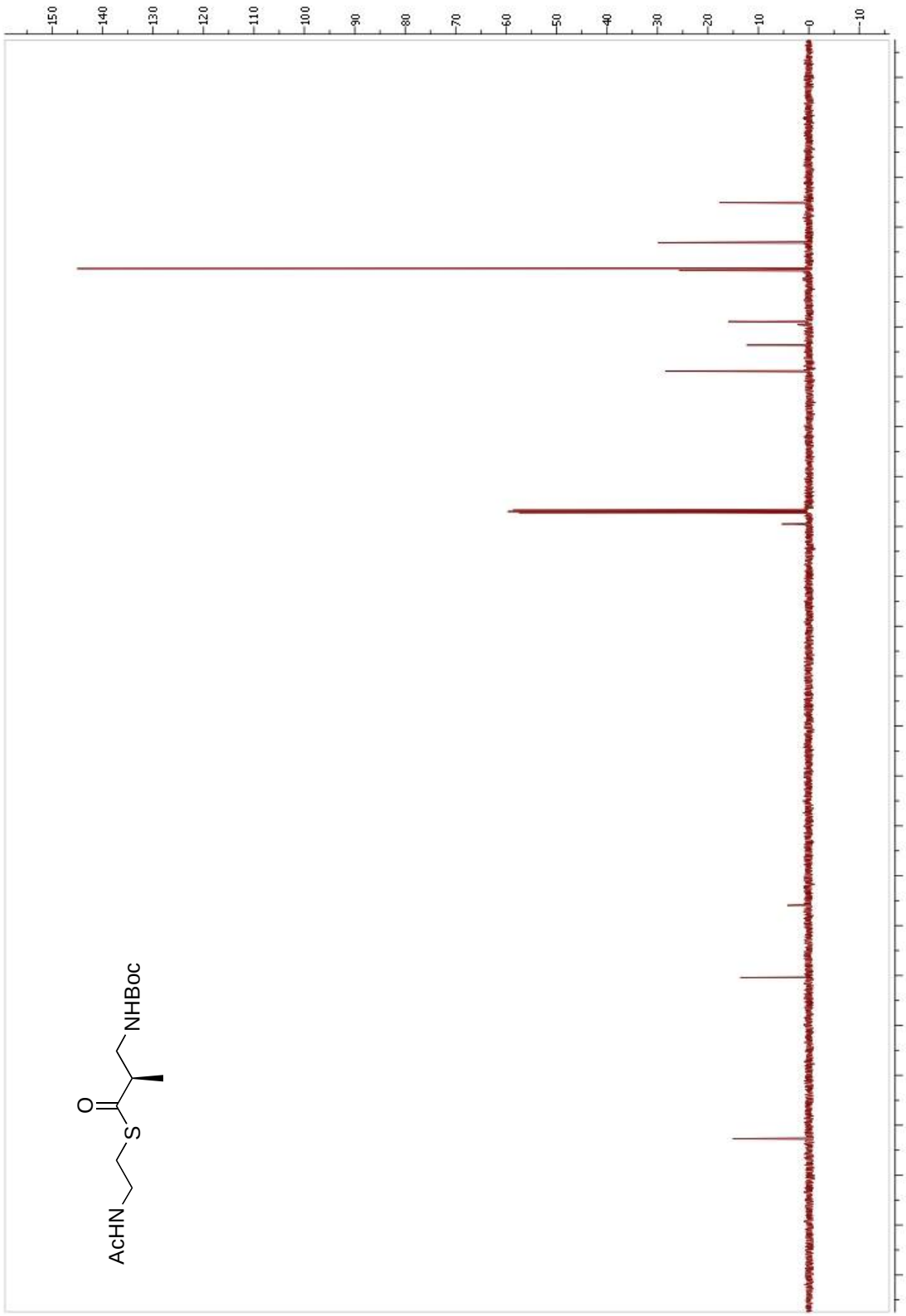


Figure S17. ^{13}C NMR spectrum of compound **24** ($CDCl_3$).

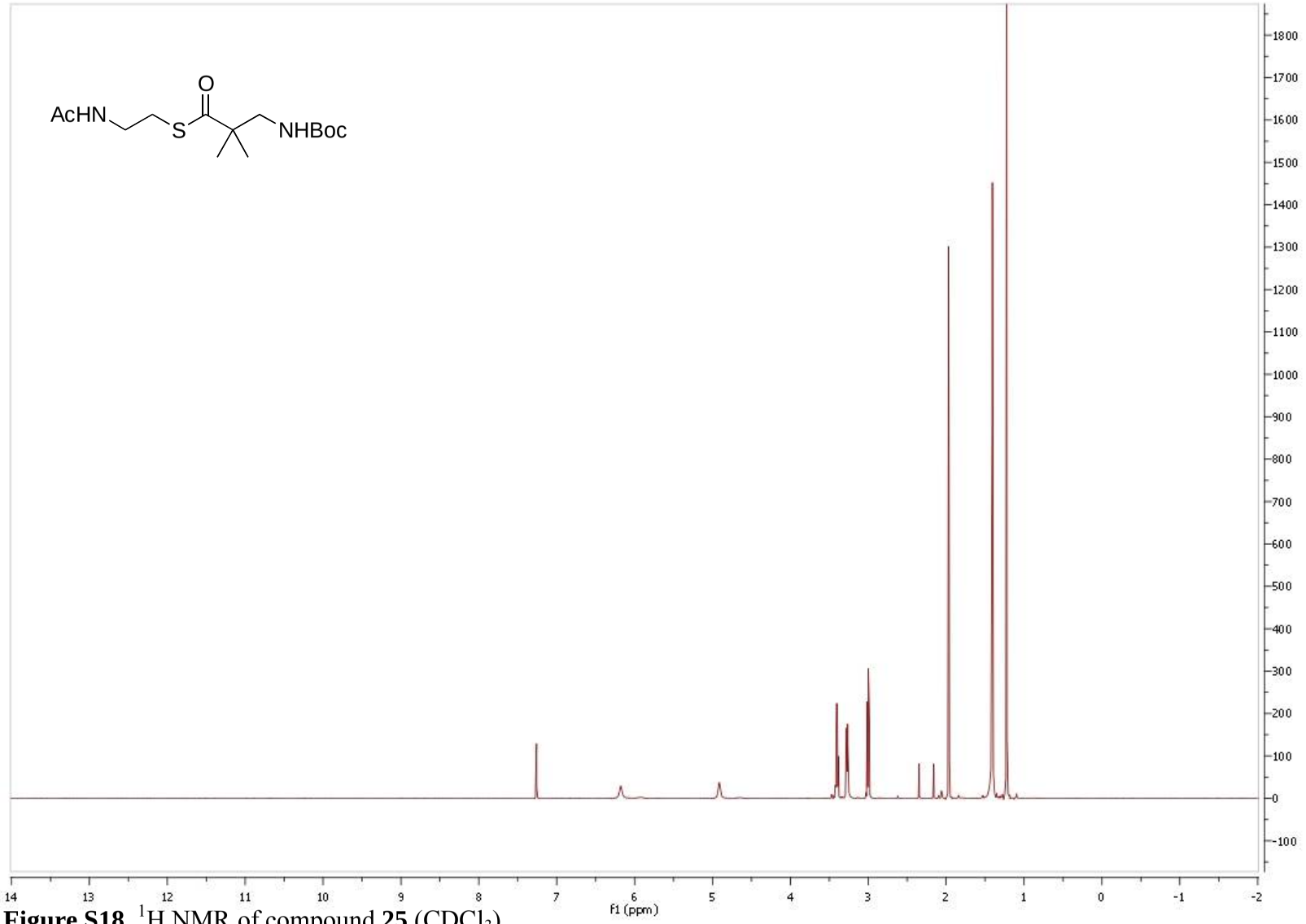


Figure S18. ^{1}H NMR of compound **25** ($CDCl_3$).

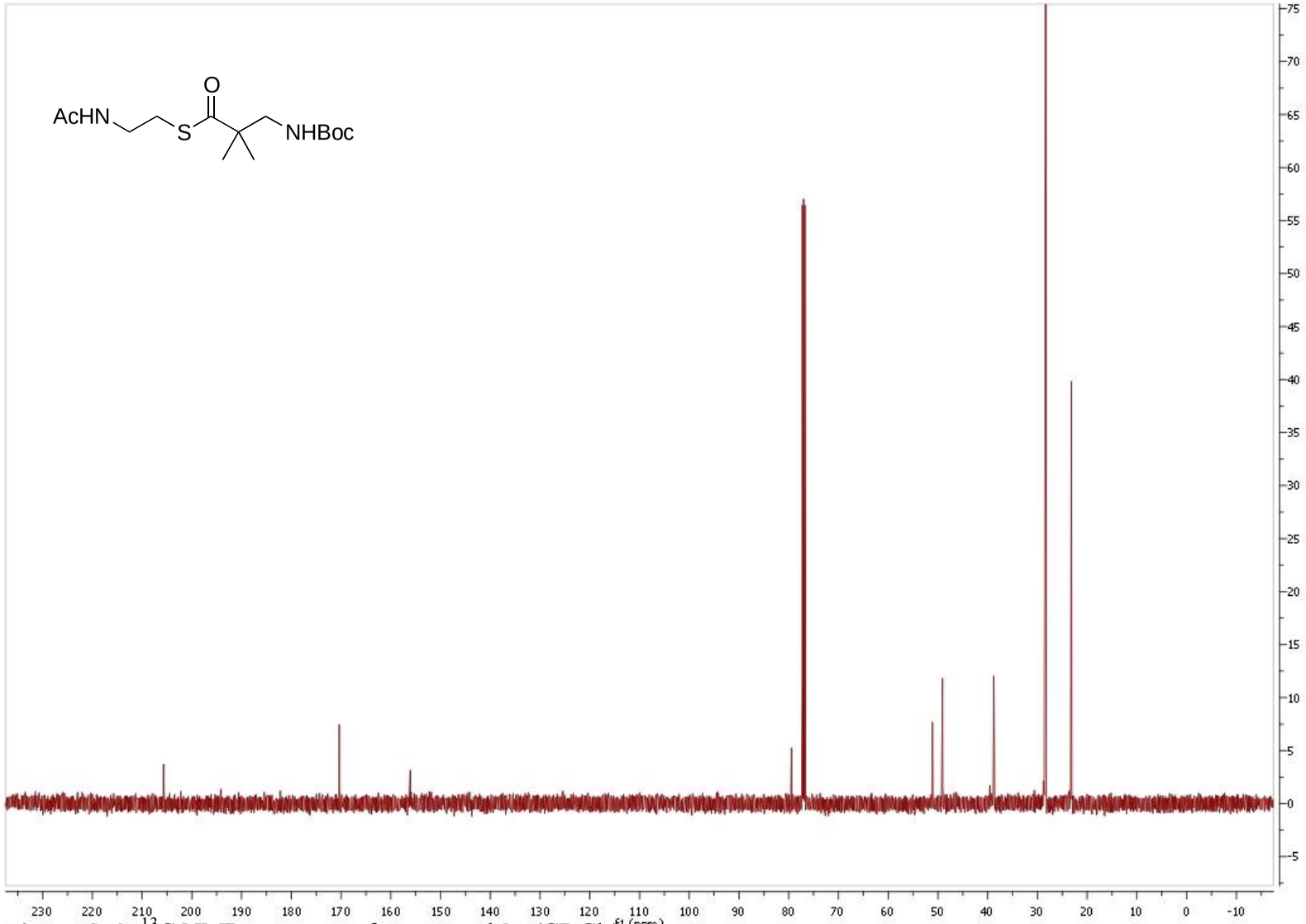


Figure S19. ^{13}C NMR spectrum of compound **25** ($CDCl_3$).

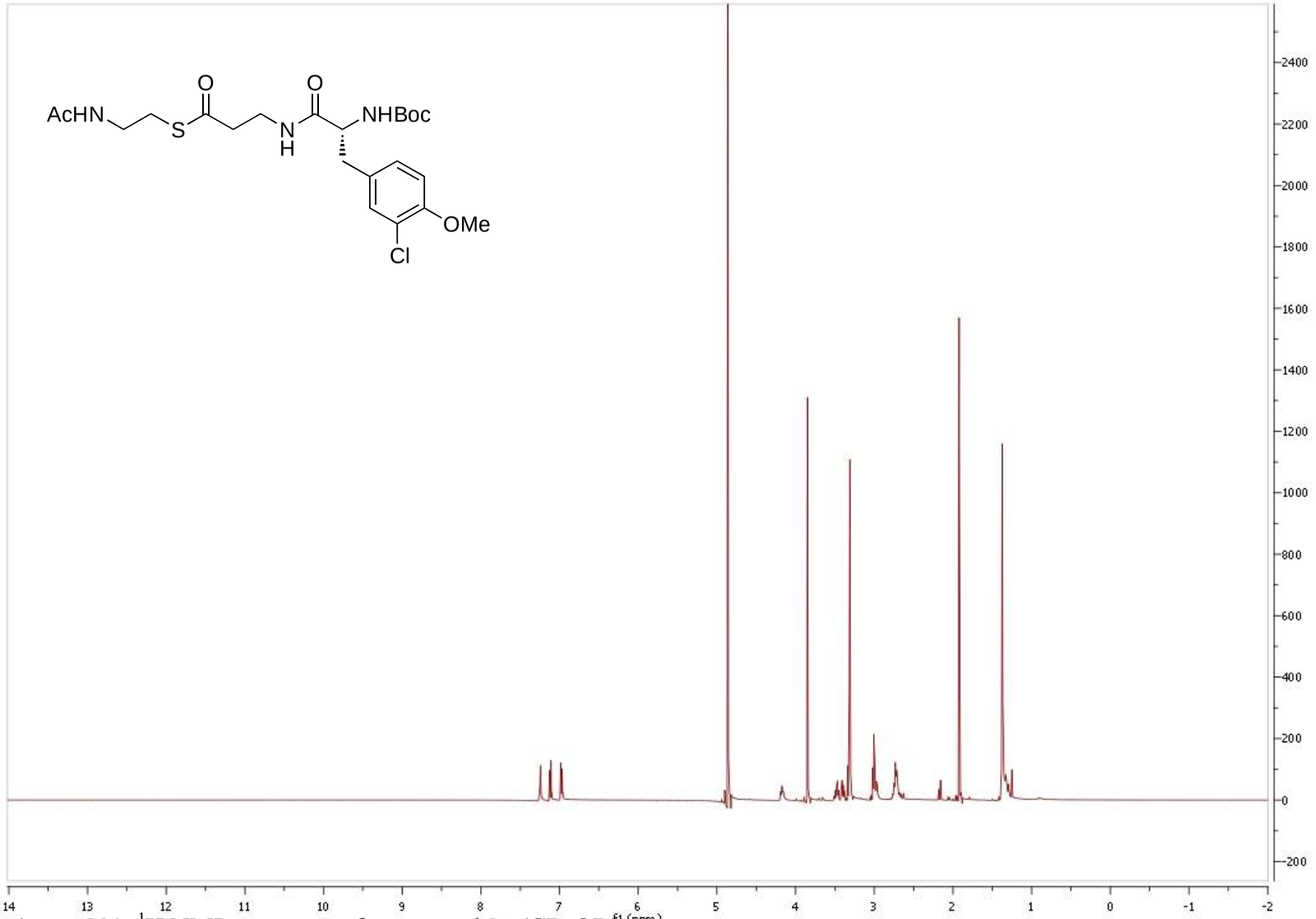


Figure S20. ^{1}H NMR spectrum of compound **27** (CD_3OD).

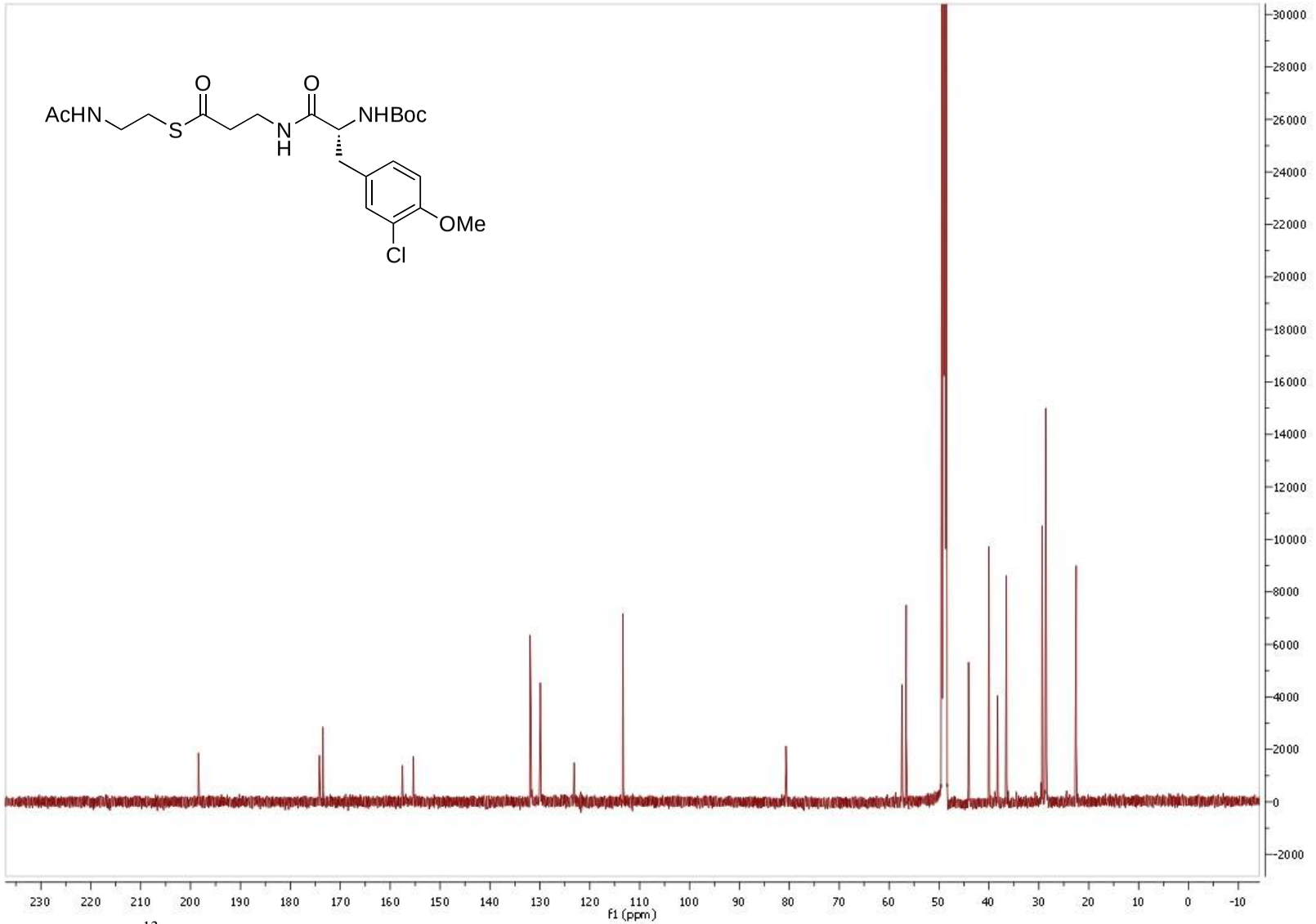


Figure S21. ^{13}C NMR spectrum of compound **27** (CD_3OD).

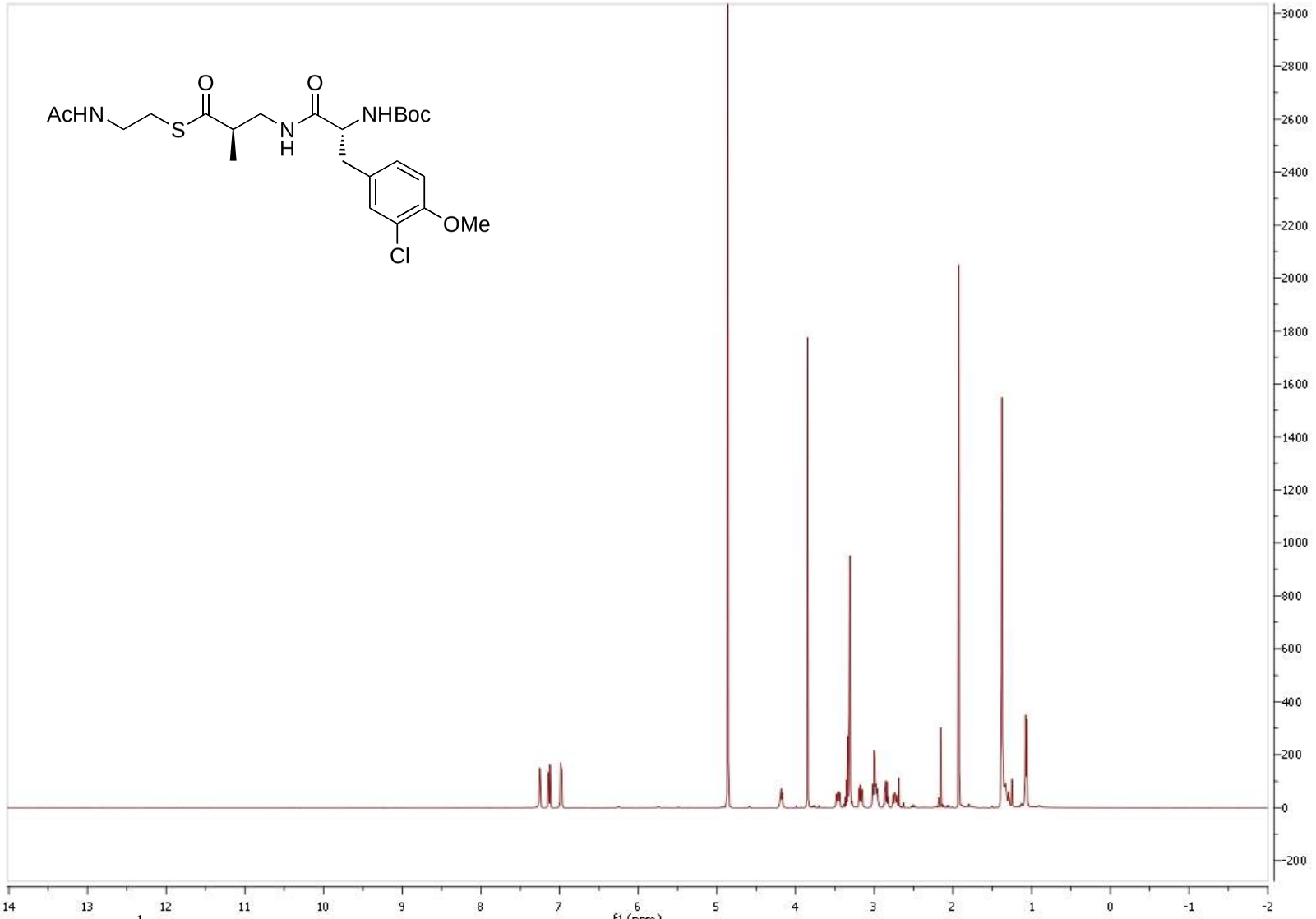


Figure S22. ^{1}H NMR spectrum of compound **28** (CD_3OD).

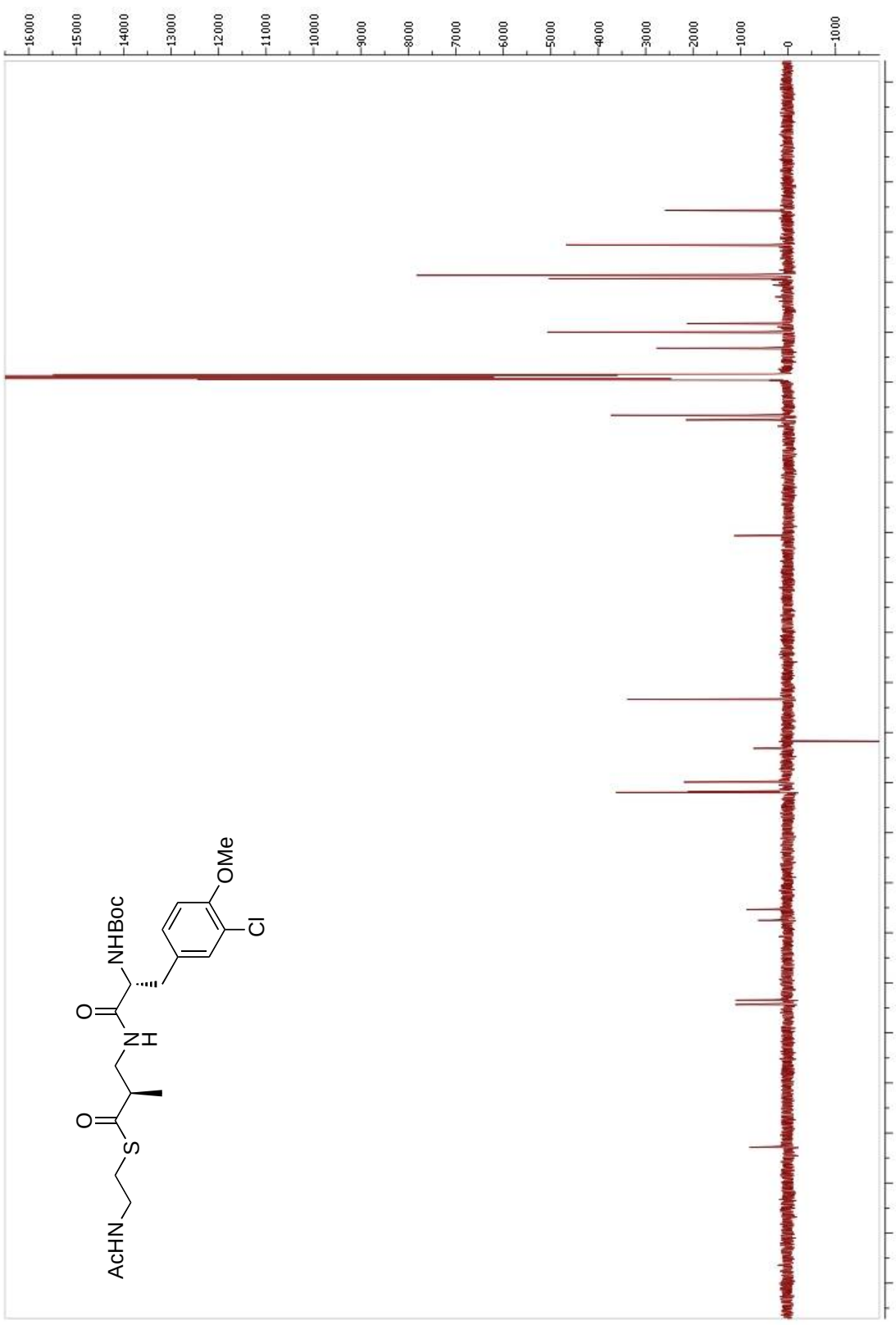


Figure S23. ^{13}C NMR spectrum of compound **28** (CD_3OD).

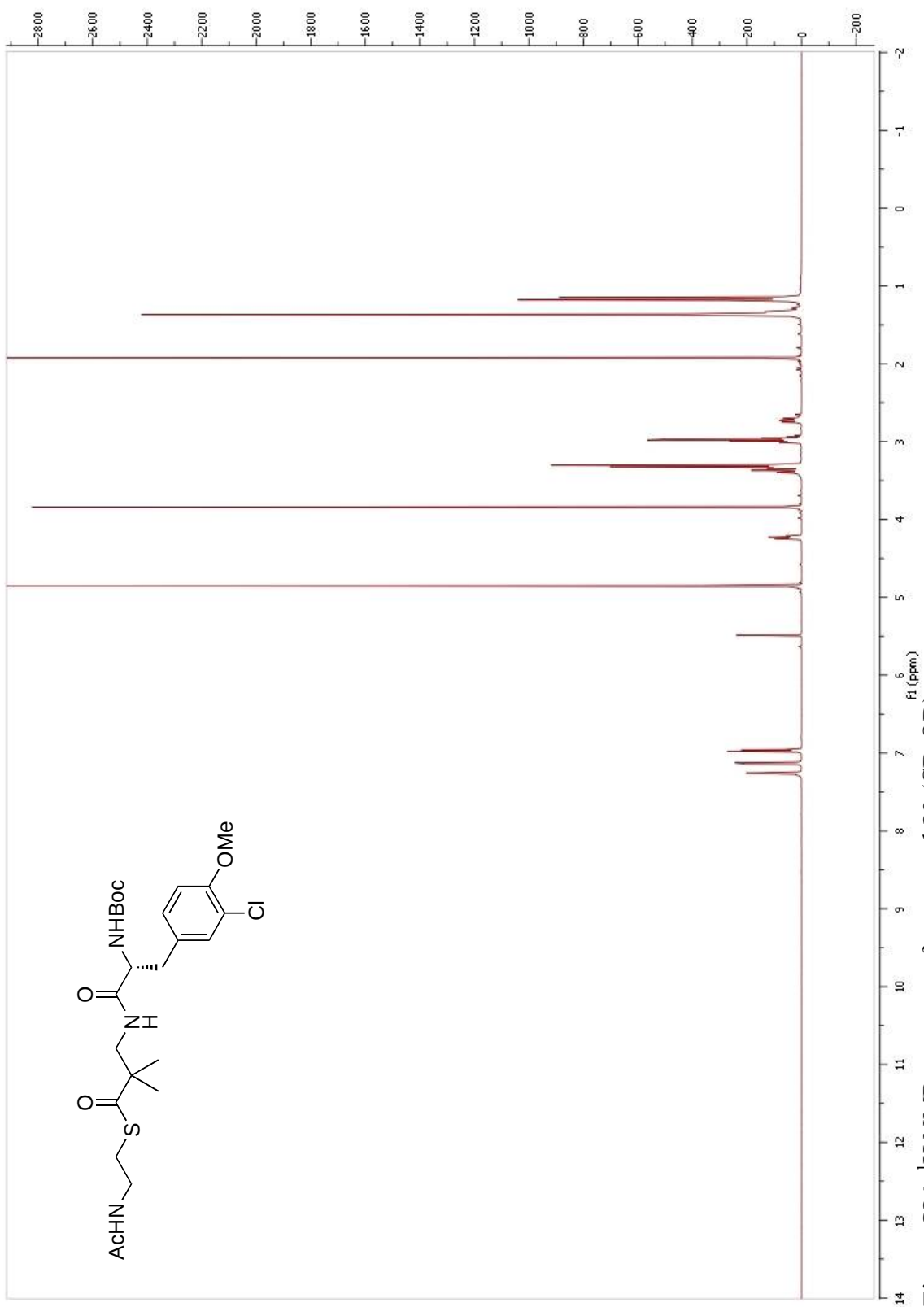


Figure S24. ^{1}H NMR spectrum of compound **29** (CD_3OD).

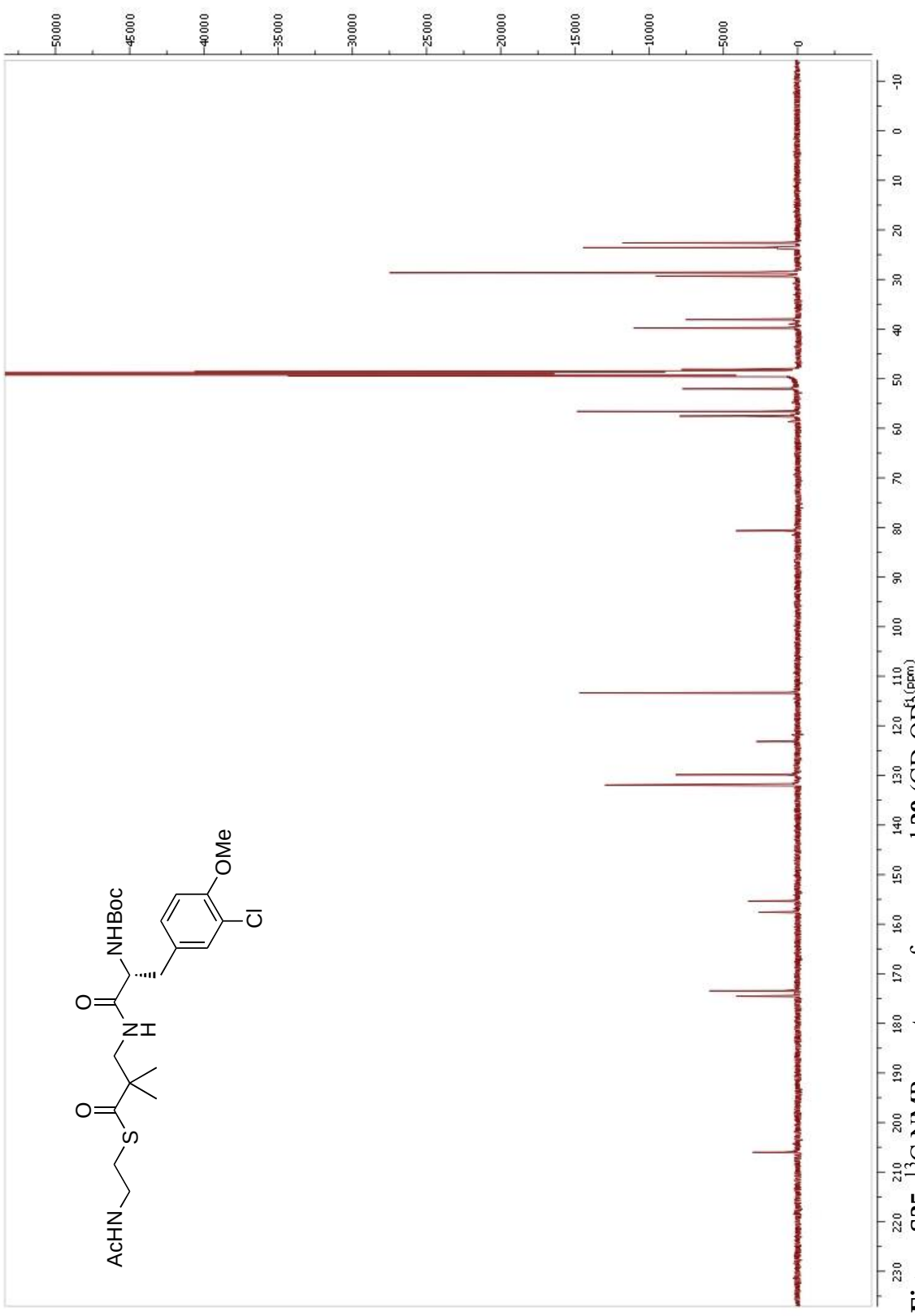


Figure S25. ^{13}C NMR spectrum of compound **29** (CD_3OD).

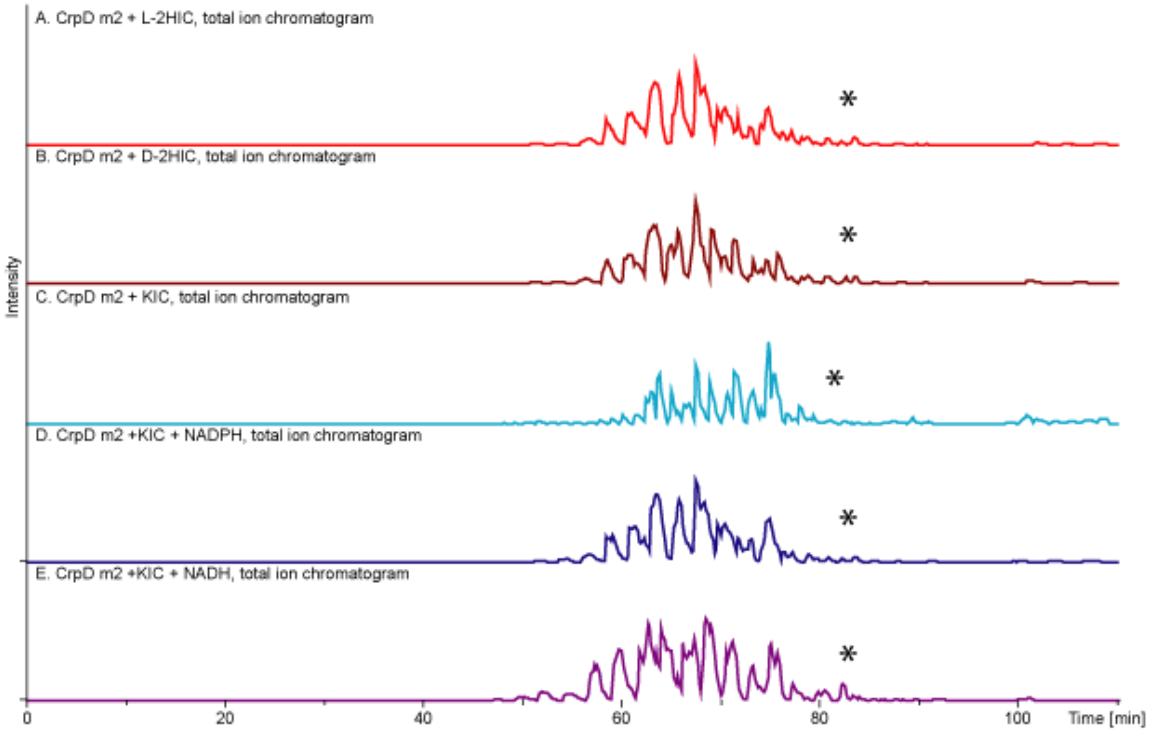


Figure S26. LC-FTICR MS total ion chromatograms for CrpD-M2 samples reacted with ATP and compounds as noted above (**A-E**). Data are presented as normalized intensity versus time, with the elution window for the identified species noted (*).

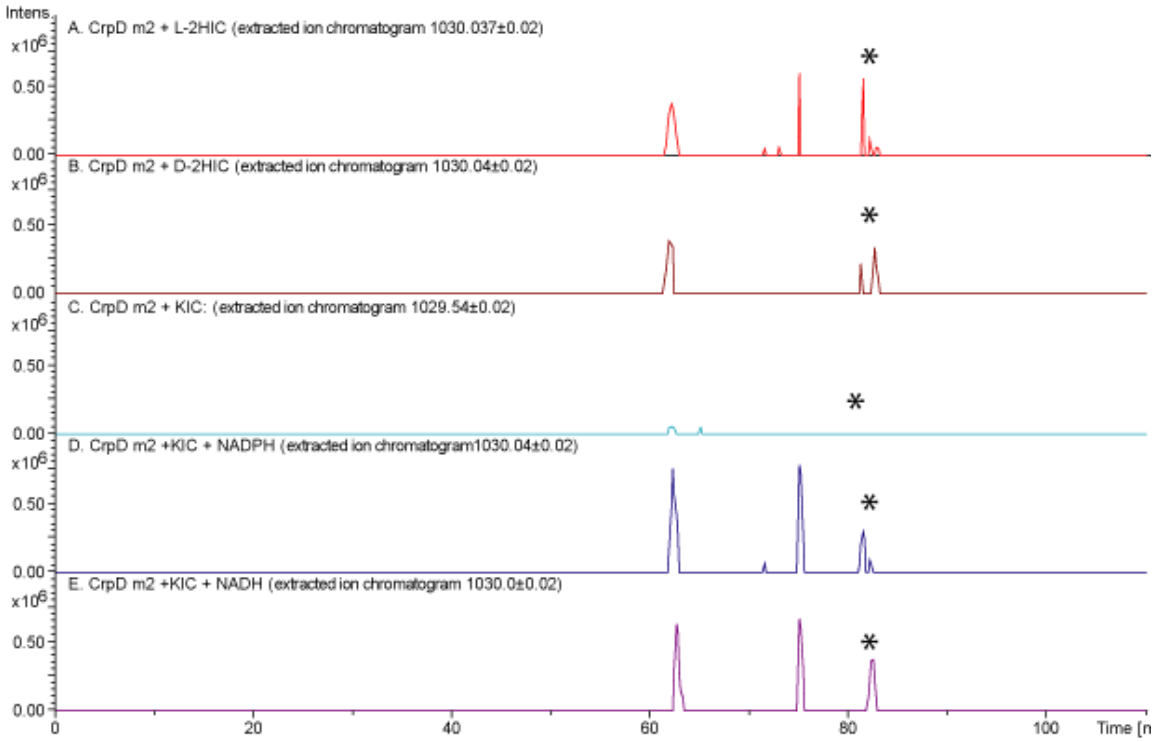


Figure S27. LC-FTICR MS extracted ion chromatograms for CrpD-M2 samples reacted with ATP and compounds as noted above with extracted m/z value and tolerance displayed (**A-E**). The targeted ions in the 4^+ charge state as presented in **Figure 3** are illustrated. Other peaks at a similar m/z value were manually inspected and found to not match the targeted peptide (e.g. incorrect charge, mass value is not the monoisotopic peak, present in negative control reaction, not confirmed by MS^2 and MS^3). Data are presented as normalized intensity versus time, with the elution window for the identified species noted (*). The signal intensity in **Figure S27 C** is lower as compared to other species, although present (**Figure S28-S29 C**).

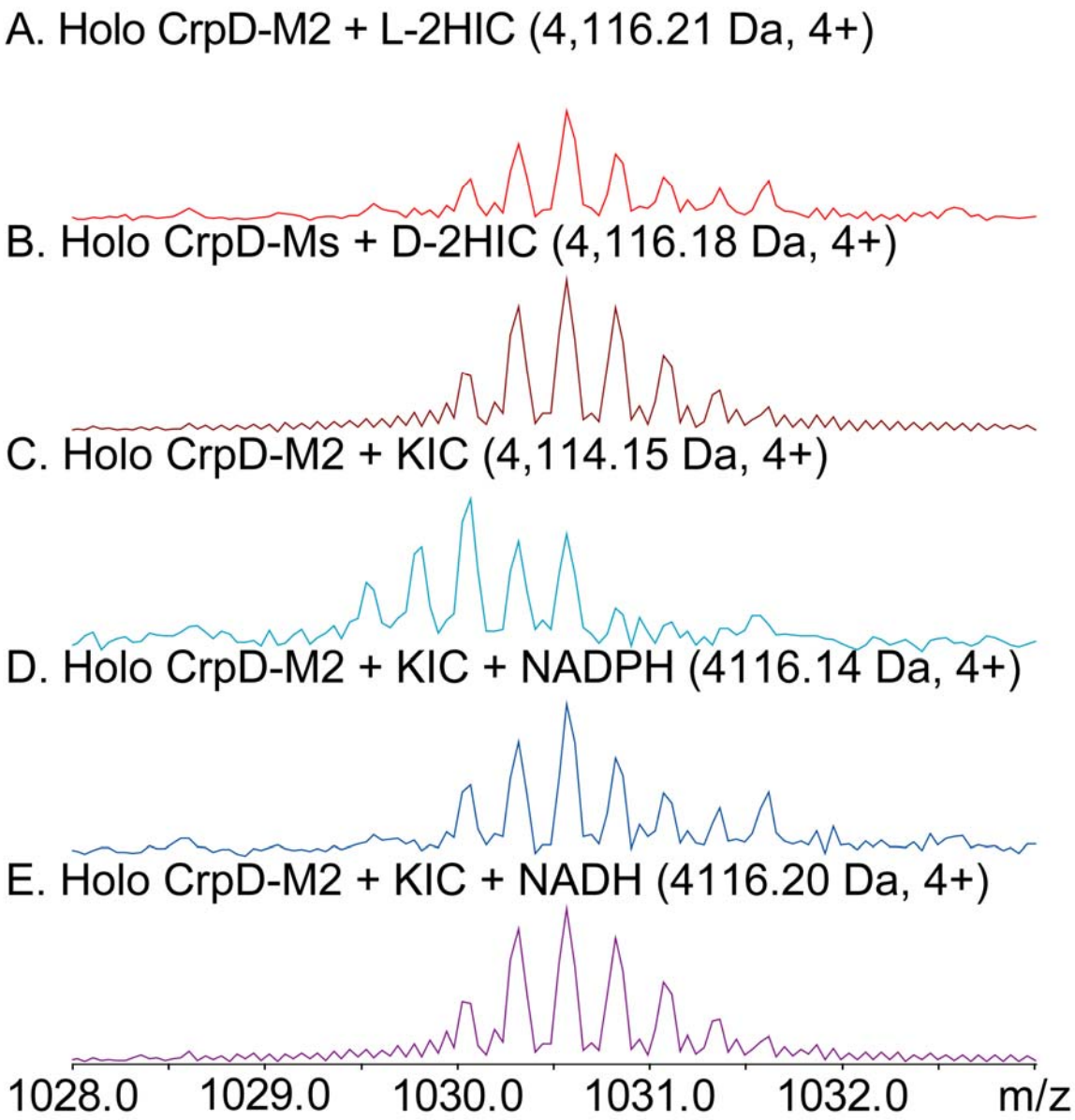


Figure S28. CrpD-M2 T domain active-site bound with extender unit intermediates monitored by LC FTICR-MS. A zoomed mass spectrum averaged over the elution window is presented. The deconvoluted monoisotopic mass, and observed charge state are shown. Reactions of CrpD-M2 with ATP and: (**A**) L-2HIC (2), (**B**) D-2HIC (20), (**C**) 2KIC (1), (**D**) 2KIC +NADPH, (E) 2KIC +NADH. Further data are provided such as LC IT-MS/MS^4 (**Table S4-5**).

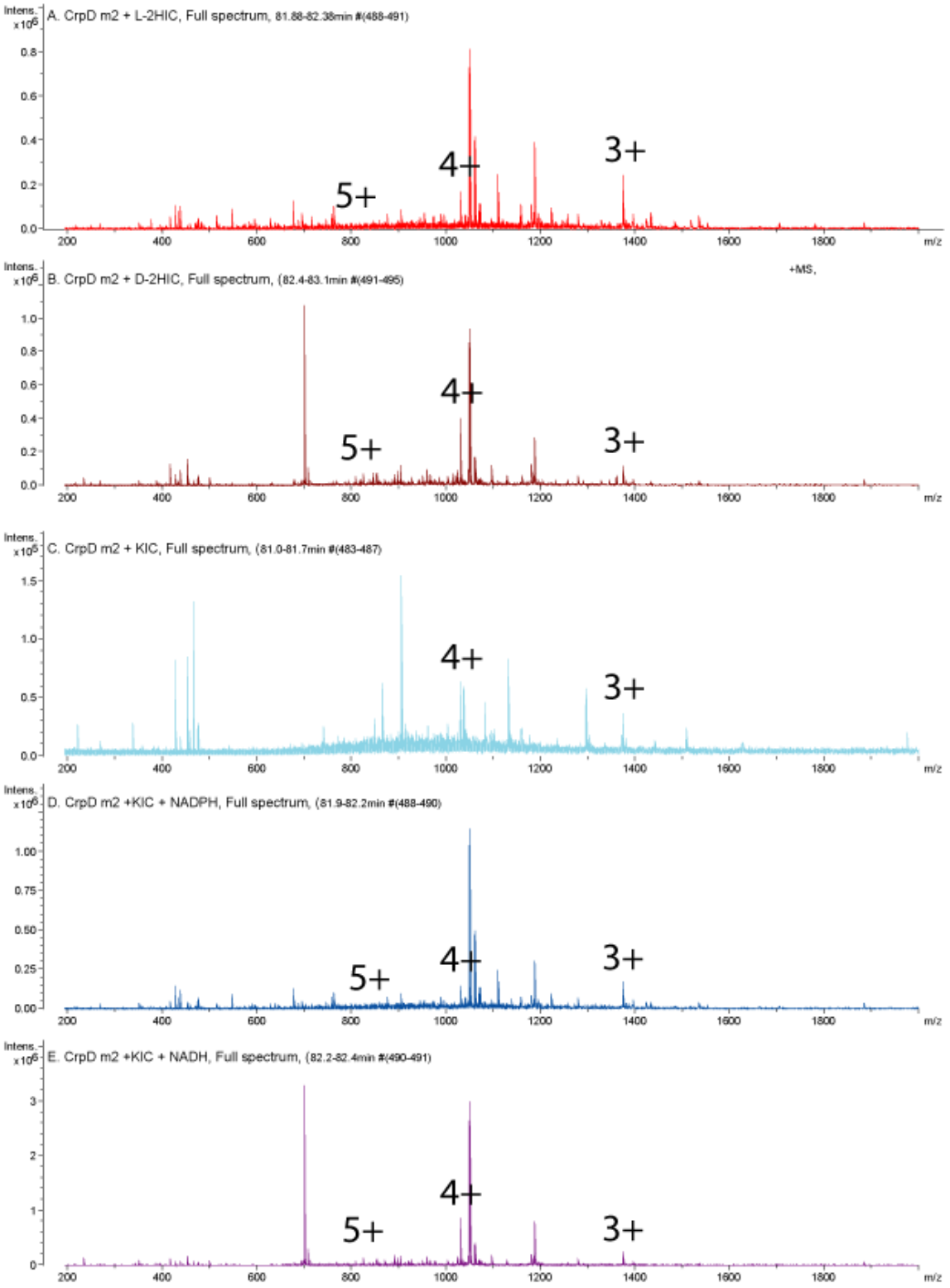


Figure S29. LC-FTICR MS full spectra averaged over the elution window for CrpD-M2 rections (**A-E**). Data are presented as absolute intensity versus m/z values. The 4^{+}, and 3^{+} charge states can be observed for all samples, indicating that indicating that the signals are real and greater than the noise.

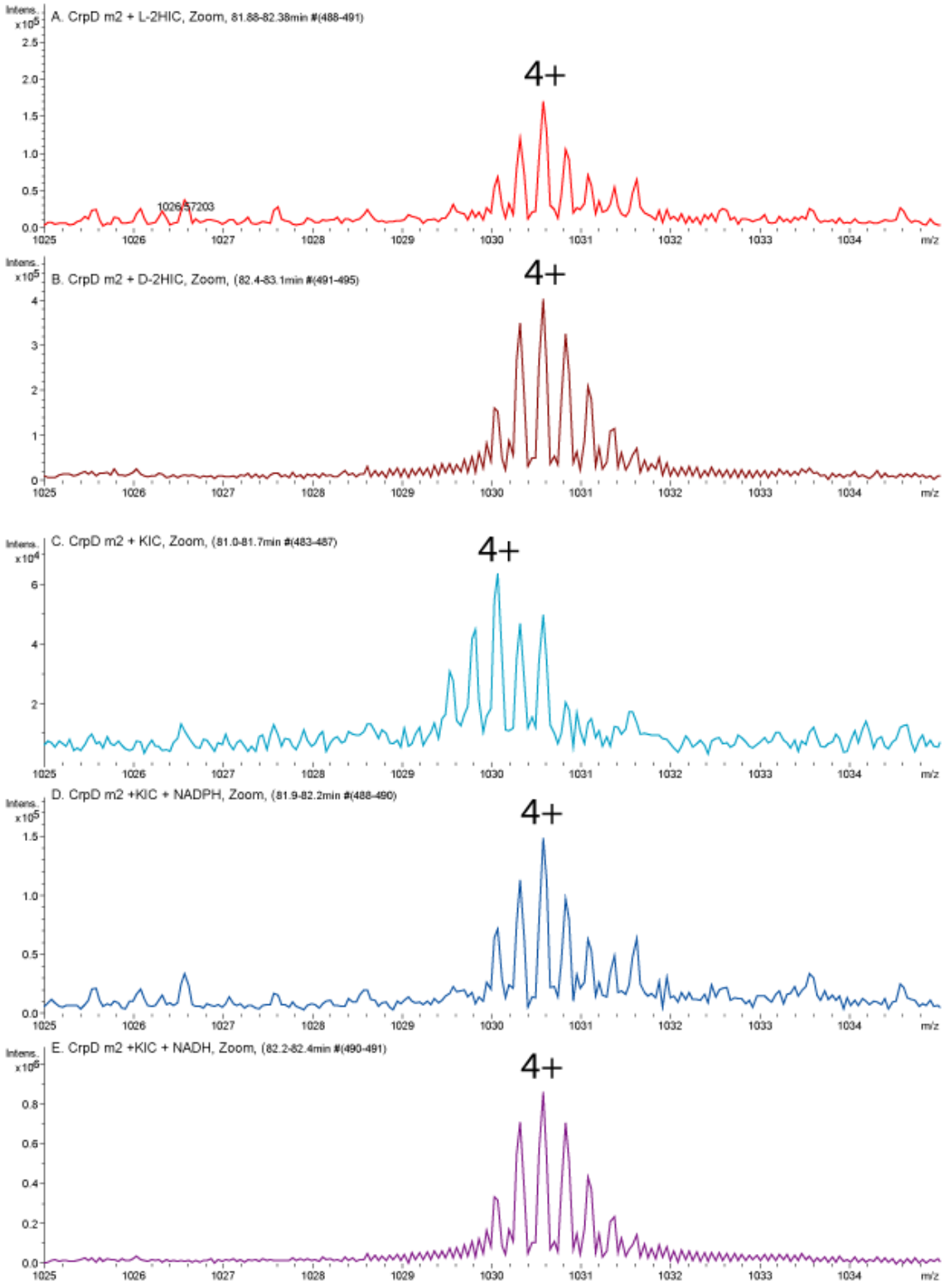


Figure S30. LC-FTICR MS intermediate zoomed spectra averaged over the elution window for CrpD-M2 rections (**A-E**) in the 4^{+} charge state. Data are presented as absolute intensity versus m/z values. This data illustrates that the identified peaks are greater than the noise, as well as the fine isotopic splitting of the ions.

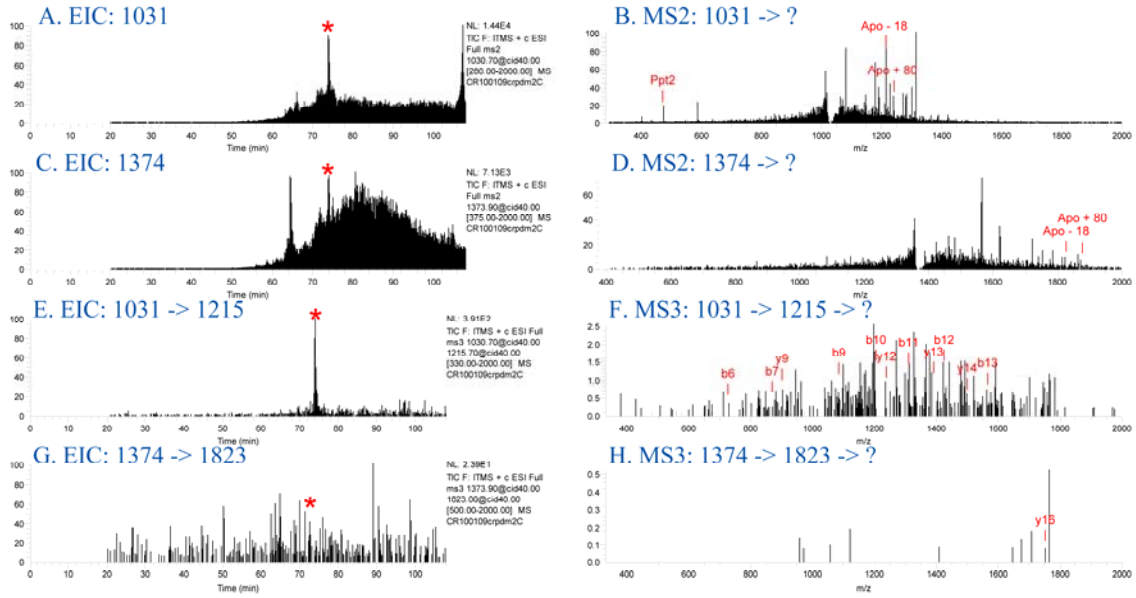


Figure S31. LC-IT MS2 and MS3 data for CrpDm2, ATP, and L-2HIC. Extracted ion chromatograms are presented for the 4^+ loaded active site peptide (**A**) and the 3+ loaded active site loaded peptide (**C**) as well as the 4^+ apo – 18 active site peptide (**E**) and the 3+ apo – 18 active site loaded peptide (**G**) as time versus normalized intensity. Assigned MS^2 (**B**, **D**) and MS^3 (**F**, **H**) spectra are shown as m/z values versus absolute intensity.

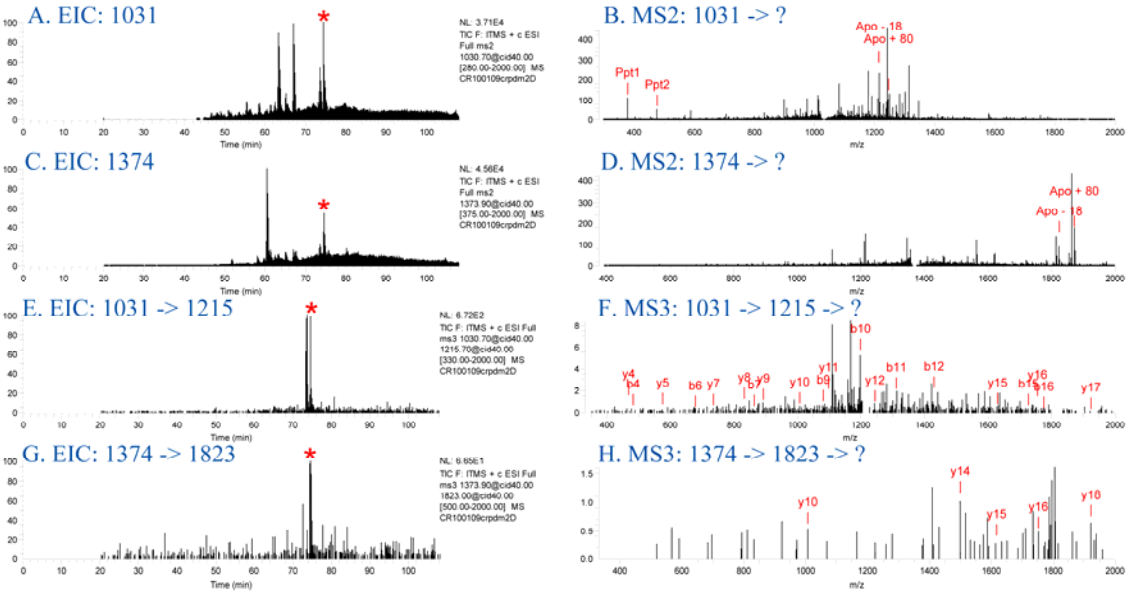


Figure S32. LC-IT MS2 and MS3 data for CrpDm2, ATP, and D-2HIC. Extracted ion chromatograms are presented for the 4^+ loaded active site peptide (**A**) and the 3+ loaded active site loaded peptide (**C**) as well as the 4^+ apo – 18 active site peptide (**E**) and the 3+ apo – 18 active site loaded peptide (**G**) as time versus normalized intensity. Assigned MS^2 (**B**, **D**) and MS^3 (**F**, **H**) spectra are shown as m/z values versus absolute intensity.

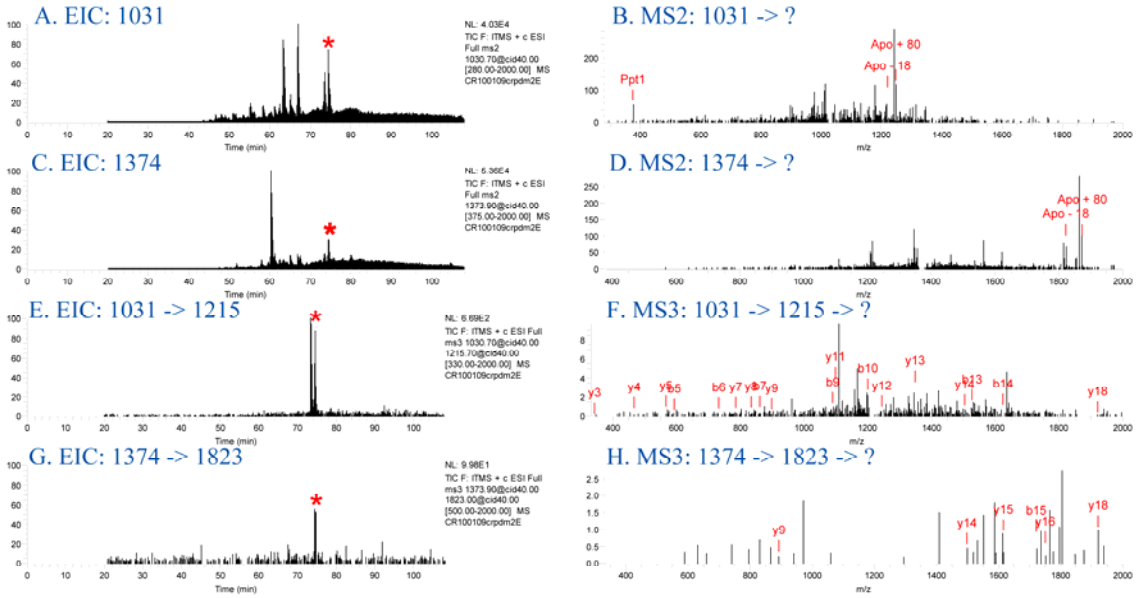


Figure S33. LC-IT MS2 and MS3 data for CrpDm2, ATP, and 2KIC. Extracted ion chromatograms are presented for the 4^+ loaded active site peptide (**A**) and the 3+ loaded active site loaded peptide (**C**) as well as the 4^+ apo – 18 active site peptide (**E**) and the 3+ apo – 18 active site loaded peptide (**G**) as time versus normalized intensity. Assigned MS^2 (**B**, **D**) and MS^3 (**F**, **H**) spectra are shown as m/z values versus absolute intensity.

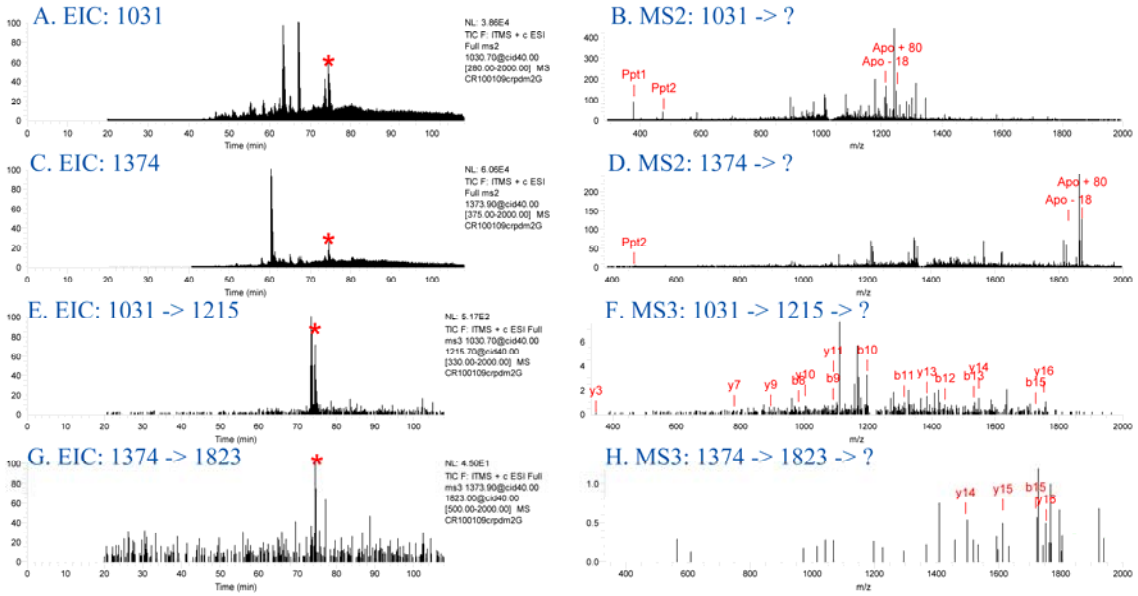


Figure S34. LC-IT MS2 and MS3 data for CrpDm2, ATP, and 2KIC +NADPH. Extracted ion chromatograms are presented for the 4^+ loaded active site peptide (**A**) and the 3+ loaded active site loaded peptide (**C**) as well as the 4^+ apo – 18 active site peptide (**E**) and the 3+ apo – 18 active site loaded peptide (**G**) as time versus normalized intensity. Assigned MS^2 (**B**, **D**) and MS^3 (**F**, **H**) spectra are shown as m/z values versus absolute intensity.

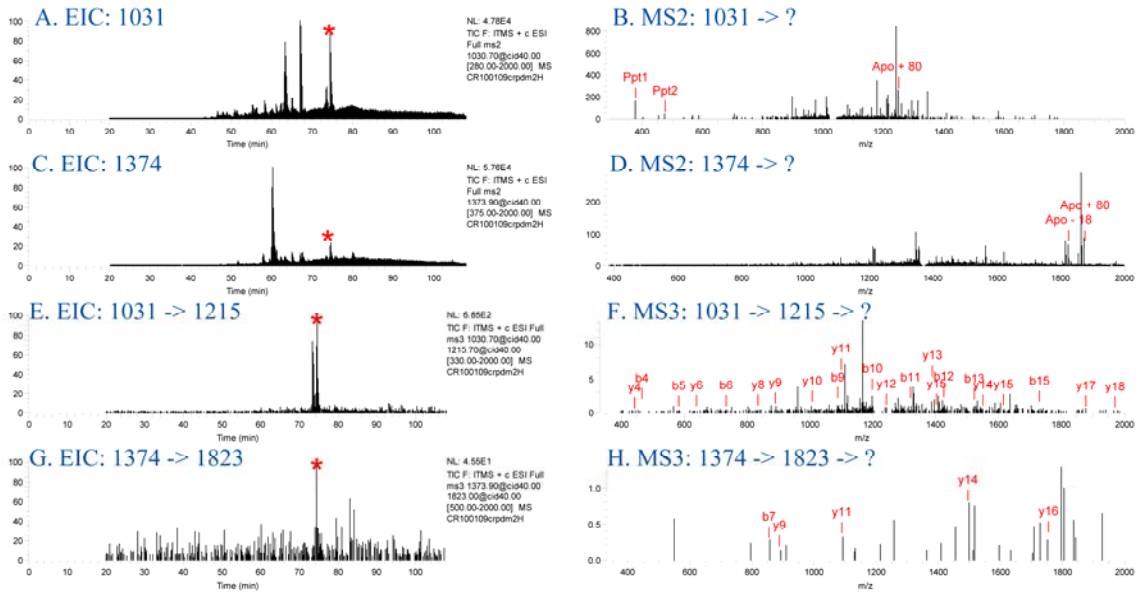


Figure S35. LC-IT MS2 and MS3 data for CrpDm2, ATP, and and 2KIC +NADH. Extracted ion chromatograms are presented for the 4^+ loaded active site peptide (**A**) and the 3+ loaded active site loaded peptide (**C**) as well as the 4^+ apo – 18 active site peptide (**E**) and the 3+ apo – 18 active site loaded peptide (**G**) as time versus normalized intensity. Assigned MS^2 (**B**, **D**) and MS^3 (**F**, **H**) spectra are shown as m/z values versus absolute intensity.

References:

(1) Beck, Z. Q.; Aldrich, C. A.; Magarvey, N. A.; Georg, G. L., Sherman, D. H. *Biochemistry* **2005**, *44*, 13457-13466.

(2) Seufert, W.; Beck, Z. Q.; Sherman, D. H. *Angewandte Chemie. Int. Ed.* **2007**, *48*, 9458-9460.

(3) Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V.-A.; Willis, C. L.; Sherman, D. H. *Gene* **2002**, *296*, 235-247.

(4) Meluzzi, D.; Zheng, W. H.; Hensler, M.; Nizet, V.; Dorrestein, P. C. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3107-3111.

(5) Liu, L.; Greger, J.; Shi, H.; Liu, Y.; Greshock, J.; Annan, R.; Halsey, W.; Sathe, G. M.; Martin, A. M.; Gilmer, T. M. *Cancer Res.* **2009**, *69*, 6871-6878.

(6) Ng, J.; Bandeira, N.; Liu, W.-T.; Ghassemian, M.; Simmons, T. L.; Gerwick, W. H.; Linington, R.; Dorrestein, P. C.; Pevzner, P. A. *Nat. Methods* **2009**, *6*, 596-601.

(7) Seufert, W.; Beck, Z. Q.; Sherman, D. H. *Angew Chem Int Ed* **2007**, *46*, 9298-9300.

(8) Eggen, M.; Mossman, C. J.; Buck, S. B.; Nair, S. K.; Bhat, L.; Ali, S. M.; Reiff, E. A.; Boge, T. C.; Georg, G. I. *J. Org. Chem.* **2000**, *65*, 7792-7798.

(9) Tamura, K.; Dudley, J.; Nei, M.; Kumar, S. *Mol. Biol. Evol.* **2007**, *24*, 1596-1599.

(10) Caffrey, P. *Chembiochem* **2003**, *4*, 654-657.

(11) Weissman, K. J.; Timoney, M.; Bycroft, M.; Grice, P.; Hanefeld, U.; Staunton, J.; Leadlay, P. F. *Biochemistry* **1997**, *36*, 13849-13855.