Supplementary Material (ESI) for Chemical Communications

Supplementary Information for

Probing the proteasome cavity in three steps: bio-orthogonal photoreactive suicide substrates

Paul P. Geurink,[a] Bogdan I. Florea,[a] Gijs A. van der Marel,[a] Benedikt M. Kessler[b]* and Herman S. Overkleeft[a]*

[a] *Leiden Institute of Chemistry and Netherlands Proteomics Centre, Gorlaeus Laboratories, Einsteinweg 55, 2333 CC Leiden, The Netherlands. Fax: +31 71 527 4307; Tel: +31 71 527 4342; E-mail: h.s.overkleeft@chem.leidenuniv.nl*
[b] *Henry Wellcome Building for Molecular Physiology, Department of Clinical Medicine, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, United Kingdom. E-mail: bmk@ccmp.ox.ac.uk*

Table of contents:

1. General experimental methods ……...……S-2
2. General procedure I: peptide coupling ……...……S-3
3. Synthesis of the compounds ……...……S-3
4. Competition assays ……S-7
5. Photocrosslinking in purified 20S proteasome ……..……S-7
6. LC-MS/MS analysis ……..….S-8
7. Structures of MV151 and biotin-phophane ……..……..S-9
8. References ……S-10
9. NMR and LC-MS spectra of compounds **2-4** ……...……S-11

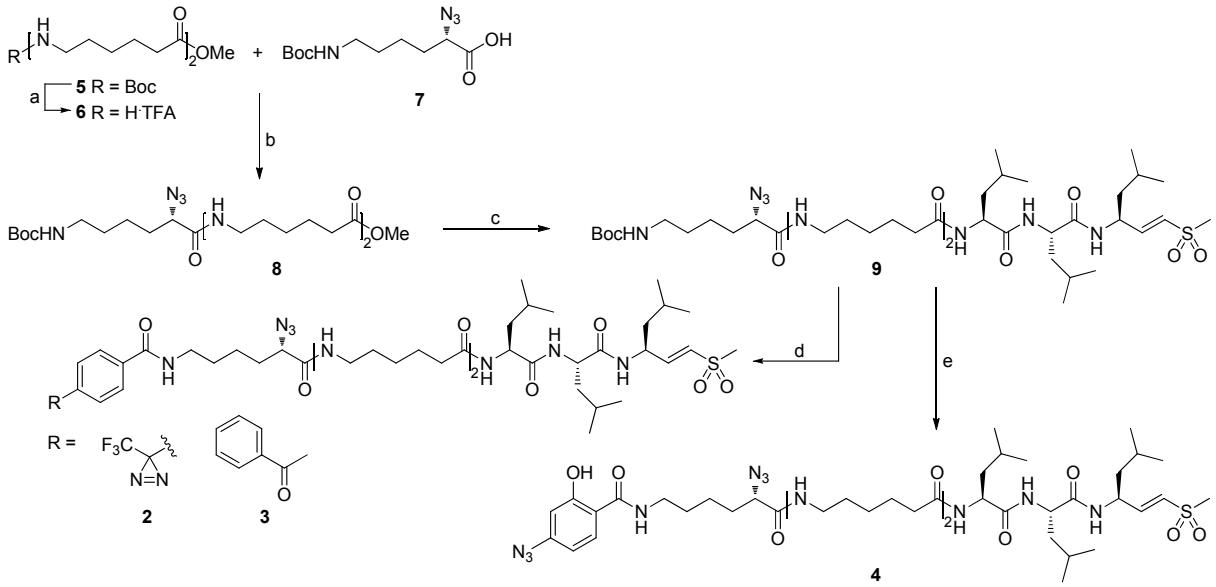
General experimental methods

Acetonitrile (ACN), dichloromethane (DCM), *N*,*N*-dimethylformamide (DMF), methanol (MeOH), diisopropylethylamine (DiPEA) and trifluoroacetic acid (TFA) were of peptide synthesis grade, purchased at Biosolve, and used as received. All general chemicals (Fluka, Acros, Merck, Aldrich, Sigma) were used as received. *O*-(1*H*-6-Chlorobenzotriazolyl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) was purchased at Iris Biotech (Marktrewitz, Germany). Traces of water were removed from reagents used in reactions that require anhydrous conditions by coevaporation with toluene. Solvents that were used in reactions were stored over 4 Å molecular sieves, except methanol and acetonitrile which were stored over 3 Å molecular sieves. Column chromatography was performed on Screening Devices b.v. Silica Gel, with a particle size of 40-63 μm and pore diameter of 60 Å. The eluents toluene, ethyl acetate and petroleum ether (40-60 °C boiling range) were distilled prior to use. TLC analysis was conducted on Merck aluminum sheets (Silica gel 60 F_{254}). Compounds were visualized by UV absorption (254 nm), by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot 4H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot 2H_2O$ (10 g/L) in 10% sulfuric acid, a solution of $KMnO_4$ (20 g/L) and K_2CO_3 (10 g/L) in water, or ninhydrin (0.75 g/L) and acetic acid (12.5 mL/L) in ethanol, where appropriate, followed by charring at ca. 150 °C. ^{1}H- and ^{13}C-NMR spectra were recorded on a Bruker AV-400 (400 MHz) spectrometer. Chemical shifts are given in ppm (δ) relative to CD_3OD as internal standard. High resolution mass spectra were recorded by direct injection (2 μL of a 2 μM solution in water/acetonitrile 50/50 (v/v) and 0.1% formic acid) on a mass spectrometer (Thermo Finnigan LTQ Orbitrap) equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10, capillary temperature 250 °C) with resolution R = 60,000 at *m/z* 400 (mass range *m/z* = 150-2,000) and dioctylpthalate (*m/z* = 391.28428) as a "lock mass". The high resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). Optical rotations $[\alpha]_D^{23}$ were recorded on a Propol automatic polarimeter. LC-MS analysis was performed on a Finnigan Surveyor HPLC system with a Gemini C18 50 × 4.60 mm column (detection at 200-600 nm), coupled to a Finnigan LCQ Advantage Max mass spectrometer with ESI. The applied buffers were H_2O, ACN and 1.0% aq. TFA. HPLC purifications were performed on a Gilson HPLC system coupled to a Phenomenex Gemini 5 μm 250 × 10 mm column and a GX281 fraction collector. The applied buffers was: 0.1% aq. TFA and ACN.

General procedure I: peptide coupling

The Boc-protected amine was treated with a 1:1 (v/v) mixture of DCM/TFA (5 mL/mmol) for 30 min. followed by addition of toluene and concentration of the mixture under reduced pressure. In order to remove excess TFA the mixture was coevaporated with toluene twice. The deprotected amine TFA salt (1 eq.) was dissolved in DCM (5 mL/mmol) and to this were added the carboxylic acid (1 eq.), HCTU (1.1 eq.) and DiPEA (3.5 eq.). The mixture was stirred until TLC analysis revealed a complete conversion (usually after 2 h). The organic layer was washed with 1M aq. HCl (2×), sat. aq. $NaHCO_3$ (2×) and brine, dried over $MgSO_4$ and concentrated under reduced pressure.

Synthesis of the compounds



Reagents and conditions: (a) TFA, DCM, quant.; (b) HCTU, DiPEA, DCM, 81%; (c) i) LiOH, MeOH; ii) TFA·H-Leu$_3$-VS,[1] HCTU, DiPEA, DCM, 71%; (d) i) TFA, DCM; ii) 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid[2] or 4-benzoylbenzoic acid, HCTU, DiPEA, DCM, **2**: 72%, **3**: 77%; (e) i) HCl, 1,4-dioxane; ii) 4-azido-2-hydroxybenzoic acid *N*-hydroxysuccinimide ester,[3] DMF, 53%.

(S)-2-azido-6-((*tert*-butoxycarbonyl)amino)hexanoic acid (7)

This compound was synthesized from H-Lys(Boc)-OH according to the procedure described by R. Stick *et al.*[4] In short: H-Lys(Boc)-OH (1.325 g, 5.38 mmol) was dissolved in MeOH (30 mL) and to this were added K_2CO_3 (3.2 eq., 17.21 mmol, 2.37 g), $CuSO_4 \cdot 5H_2O$ (1 mol%, 54.0 μmol, 13.0 mg) and imidazole-1-sulfonyl azide hydrochloride (1.2 eq., 6.46 mmol, 1.35 g). The mixture was stirred for 15 h after which TLC analysis indicated a complete conversion.

The mixture was concentrated under reduced pressure, redissolved in EtOAc and washed with 1M aq. HCl. The aqueous layer was extracted twice with EtOAc followed by washing of the combined organic layers with brine, drying over $MgSO_4$ and concentration under reduced pressure. The title compound was obtained after purification by column chromatography (25% → 50% EtOAc/PE) as a colorless oil (yield: 524 mg, 1.92 mmol, 36%). The spectroscopic data corresponded to those reported in literature.[5]

Boc-Ahx(α-N_3)-Ahx_2-OMe (8)

This compound was synthesized according to general procedure I from **7** (1.92 mmol) and Boc-Ahx_2-OMe **5**[6] (2.0 mmol). The title compound was purified by column chromatography (75% EtOAc/PE → 5% MeOH/EtOAc) and obtained as a colorless oil (yield: 0.80 g, 1.55 mmol, 81%). ^{1}H NMR (400 MHz, CD_3OD): δ = 8.14 (t, *J* = 5.58 Hz, 1H), 7.88 (t, *J* = 5.39 Hz, 1H), 3.79 (dd, *J* = 11.98, 4.97 Hz, 1H), 3.64 (s, 3H), 3.21 (dd, *J* = 13.06, 7.28 Hz, 2H), 3.16 (dd, *J* = 12.84, 7.02 Hz, 2H), 3.04 (t, *J* = 6.73 Hz, 2H), 2.32 (t, *J* = 7.40 Hz, 2H), 2.18 (t, *J* = 7.44 Hz, 2H), 1.87-1.70 (m, 2H), 1.62 (td, *J* = 15.35, 7.53 Hz, 4H), 1.57-1.47 (m, 6H), 1.43 (s, 9H), 1.39-1.27 (m, 6H) ppm. ^{13}C NMR (100 MHz, CD_3OD): δ = 175.63, 175.54, 175.44, 172.03, 171.95, 158.22, 79.61, 64.17, 64.12, 51.98, 40.00, 40.31, 40.17, 40.04, 36.94, 34.59, 32.40, 30.45, 30.02, 29.98, 28.86, 27.42, 27.36, 26.57, 25.58, 23.91 ppm. $[\alpha]_D^{23}$ = +9.7° (*c* = 1 in MeOH). LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 7.47. HRMS: calcd. for $C_{24}H_{44}N_6O_6$ $[M + H]^+$ 513.33951; found 513.33949.

Boc-Ahx(α-N_3)-Ahx_2-Leu_3-VS (9)

Compound **8** (0.80 g, 1.55 mmol) was dissolved in MeOH (8 mL) and cooled to 0 °C, after which LiOH (2 mL of a 1M aq. solution) was added slowly. The mixture was slowly warmed to RT and stirred for 15 h, after which TLC analysis indicated complete consumption of starting material. Next, 1M aq, HCl (3.6 mL) was added to neutralize the solution (~pH 5) and the mixture was concentrated under reduced pressure. The product was dissolved in DCM and dried over $MgSO_4$ to remove all traces of water. The resulting carboxylic acid was coupled to Boc-Leu_3-VS[1] (1.61 mmol) according to general procedure I and the title compound was obtained after purification by column chromatography (75% EtOAc/PE → 10% MeOH/EtOAc) as a colorless solid (yield: 0.99 g, 1.10 mmol, 71%). ^{1}H NMR (400 MHz, CD_3OD): δ = 8.15 (t, *J* = 5.61 Hz, 1H), 8.06 (d, *J* = 8.14 Hz, 2H), 7.88 (t, *J* = 5.40 Hz, 2H), 6.80 (dd, *J* = 15.16, 5.09 Hz, 1H), 6.63 (dd, *J* = 15.18, 1.19 Hz, 1H), 4.70-4.64 (m, 1H), 4.42-

4.34 (m, 2H), 3.79 (t, *J* = 6.72 Hz, 1H), 3.21 (t, *J* = 6.98 Hz, 2H), 3.15 (t, *J* = 6.92 Hz, 2H), 3.04 (t, *J* = 6.73 Hz, 2H), 2.98 (s, 3H), 2.26 (dt, *J* = 7.12, 3.26 Hz, 2H), 2.18 (t, *J* = 7.44 Hz, 2H), 1.87-1.73 (m, 2H), 1.71-1.45 (m, 22H), 1.43 (s, 9H), 1.39-1.30 (m, 3H), 0.97-0.89 (m, 18H) ppm. ^{13}C NMR (100 MHz, CD_3OD): δ = 176.06, 175.62, 174.81, 174.04, 172.03, 158.30, 148.36, 130.77, 79.68, 79.23, 64.19, 53.37, 53.26, 49.03, 43.22, 42.90, 41.62, 41.46, 41.00, 40.22, 40.13, 36.91, 36.59, 32.42, 30.45, 30.08, 29.98, 28.86, 27.53, 27.43, 26.59, 26.50, 25.87, 25.82, 25.74, 23.90, 23.49, 23.46, 22.20, 22.11, 22.01 ppm. $[\alpha]_D^{23}$ = –29.5° (*c* = 1 in MeOH). LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 8.50. HRMS: calcd. for $C_{43}H_{79}N_9O_9S$ $[M + H]^+$ 898.57942; found 898.58048.

4-(3-(Trifluoromethyl)-3*H*-diazirin-3-yl)benzamido-Ahx(α-N_3)-Ahx_2-Leu_3-VS (2)

This compound was synthesized according to general procedure I from compound **9** (0.21 mmol) and 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid[2] (0.23 mmol). The title compound was purified by crystallization from MeOH/Et_2O and obtained as a colorless solid (yield: 155 mg, 0.15 mmol, 72%). ^{1}H NMR (400 MHz, CD_3OD): δ = 8.60 (t, *J* = 5.33 Hz, 1H), 8.17 (t, *J* = 5.41 Hz, 2H), 8.06 (t, *J* = 7.59 Hz, 3H), 7.90 (d, *J* = 8.48 Hz, 2H), 7.33 (d, *J* = 8.21 Hz, 2H), 6.79 (dd, *J* = 15.17, 5.08 Hz, 1H), 6.61 (dd, *J* = 15.19, 1.12 Hz, 1H), 4.72-4.62 (m, 1H), 4.42-4.30 (m, 2H), 3.81 (t, *J* = 6.69 Hz, 1H), 3.40 (dd, *J* = 12.09, 6.52 Hz, 2H), 3.23-3.12 (m, 4H), 2.97 (s, 3H), 2.25 (t, *J* = 7.18 Hz, 2H), 2.17 (t, *J* = 7.42 Hz, 2H), 1.93-1.75 (m, 2H), 1.72-1.40 (m, 22H), 1.39-1.27 (m, 3H), 0.93 (ddd, *J* = 9.86, 9.32, 5.12 Hz, 18H) ppm. ^{13}C NMR (100 MHz, CD_3OD): δ = 176.38, 175.88, 175.15, 174.39, 172.30, 168.78, 148.50, 137.35, 133.01, 130.85, 129.05, 127.67, 123.4 (q, *J* = 273.92 Hz), 64.41, 53.56, 49.29, 43.33, 42.86, 41.65, 41.60, 40.91, 40.79, 40.43, 40.31, 40.24, 36.98, 36.66, 32.50, 30.16, 30.07, 29.99, 27.63, 27.51, 26.67, 26.59, 25.98, 25.95, 25.85, 24.13, 23.46, 22.11, 22.02, 21.96 ppm. $[\alpha]_D^{23}$ = –28.2° (*c* = 1 in MeOH). LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 9.16. HRMS: calcd. for $C_{47}H_{74}F_3N_{11}O_8S$ $[M + H]^+$ 1010.54674; found 1010.54811.

4-Benzoylbenzamido-Ahx(α-N_3)-Ahx_2-Leu_3VS (3)

This compound was synthesized according to general procedure I from compound **9** (0.22 mmol) and commercially available 4-benzoylbenzoic acid (0.24 mmol). The title compound was purified by crystallization from MeOH/Et_2O and obtained as a colourless solid (yield: 170 mg, 0.17 mmol, 77%). ^{1}H NMR (400 MHz, CD_3OD): δ = 7.97 (d, *J* = 8.30 Hz, 2H), 7.83

(d, *J* = 8.31 Hz, 2H), 7.78 (d, *J* = 7.17 Hz, 2H), 7.66 (t, *J* = 7.41 Hz, 1H), 7.54 (t, *J* = 7.66 Hz, 2H), 6.80 (dd, *J* = 15.17, 4.97 Hz, 1H), 6.62 (dd, *J* = 15.18, 1.26 Hz, 1H), 4.66 (td, *J* = 8.85, 4.85 Hz, 1H), 4.41-4.29 (m, 2H), 3.83 (t, *J* = 6.73 Hz, 1H), 3.48-3.41 (m, 2H), 3.20 (t, *J* = 7.03 Hz, 2H), 3.14 (t, *J* = 7.02 Hz, 2H), 2.98 (s, 3H), 2.29-2.24 (m, 2H), 2.17 (t, *J* = 7.42 Hz, 2H), 1.94-1.77 (m, 2H), 1.74-1.43 (m, 22H), 1.38-1.28 (m, 3H), 0.97-0.88 (m, 18H) ppm. ^{13}C NMR (100 MHz, CD_3OD): δ = 197.64, 176.49, 175.82, 175.24, 174.37, 172.20, 169.06, 148.50, 141.22, 139.33, 138.34, 134.18, 131.06, 130.98, 129.67, 128.43, 64.30, 53.81, 53.54, 49.90, 49.17, 43.22, 42.88, 41.57, 41.41, 40.80, 40.29, 40.21, 36.97, 36.67, 32.48, 30.13, 30.03, 29.98, 27.62, 27.50, 26.65, 26.56, 25.97, 25.93, 24.13, 23.48, 23.44, 22.08, 21.94 ppm. $[\alpha]_D^{23}$ = –26.8° (*c* = 1 in MeOH). LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 8.67. HRMS: calcd. for $C_{52}H_{79}N_9O_9S$ $[M + H]^+$ 1006.57942; found 1006.58079.

4-Azido-2-hydroxybenzamido-Ahx(α-N_3)-Ahx_2-Leu_3-VS (4)

Compound **9** (0.23 g, 0.25 mmol) was Boc-deprotected with HCl (3 mL of a 4M solution in 1,4-dioxane), followed by coevaporation with toluene (3×) and the resulting amine HCl salt was dissolved in DMF (2 mL). To this were added DiPEA (1 eq., 0.25 mmol, 43.0 μL) and 4-azido-2-hydroxybenzoic acid *N*-hydroxysuccinimide ester[3] (1.3 eq., 91.0 mg, 0.33 mmol) and the reaction mixture was stirred for 15 h, after which LC-MS analysis indicated complete consumption of the amine. DCM (10 mL) was added and the mixture was washed with 1M aq. HCl (2×), sat. aq. $NaHCO_3$ (2×) and brine, dried over $MgSO_4$ and concentrated under reduced pressure. The title compound was obtained after purification by column chromatography (100% DCM → 6% MeOH/DCM) as a colourless solid (yield: 130 mg, 0.14 mmol, 53%). ^{1}H NMR (400 MHz, CD_3OD): δ = 7.75 (d, *J* = 8.32 Hz, 1H), 6.79 (dd, *J* = 15.17, 5.08 Hz, 1H), 6.61 (dd, *J* = 15.19, 1.27 Hz, 1H), 6.57-6.52 (m, 2H), 4.65 (td, *J* = 9.10, 4.93 Hz, 1H), 4.35 (ddd, *J* = 14.62, 8.85, 6.09 Hz, 2H), 3.80 (t, *J* = 6.69 Hz, 1H), 3.38 (t, *J* = 6.95 Hz, 2H), 3.20-3.12 (m, 4H), 2.97 (s, 3H), 2.24 (dd, *J* = 8.01, 6.18 Hz, 2H), 2.16 (t, *J* = 7.43 Hz, 2H), 1.89-1.74 (m, 2H), 1.71-1.39 (m, 22H), 1.38-1.25 (m, 3H), 0.97-0.86 (m, 18H) ppm. ^{13}C NMR (100 MHz, CD_3OD): δ = 176.27, 175.79, 174.96, 174.18, 172.11, 170.31, 162.90, 148.43, 146.49, 130.83, 130.43, 113.88, 110.89, 108.23, 79.29, 64.31, 53.52, 53.38, 49.15, 43.27, 42.88, 41.62, 41.51, 40.28, 40.21, 40.16, 36.96, 36.64, 32.44, 30.12, 30.00, 27.59, 27.47, 26.61, 26.53, 25.94, 25.90, 25.81, 24.07, 23.45, 22.14, 22.04, 21.98 ppm. $[\alpha]_D^{23}$ = –27.1° (*c* = 1 in MeOH). LC-MS: gradient 10% → 90% ACN/(0.1% TFA/H_2O): R_t (min): 8.95. HRMS: calcd. for $C_{45}H_{74}N_{12}O_9S$ $[M + H]^+$ 959.54952; found 959.55103.

Competition assays

Whole cell lysates of EL4 were made by sonication of cell pellets in 3 volumes of lysis buffer containing 50 mM Tris pH 7.5, 1 mM DTT, 5 mM $MgCl_2$, 250 mM sucrose, 2 mM ATP. Protein concentration was determined by the Bradford assay. Cell lysate (9 μg total protein) was exposed to the inhibitors **2**, **3** or **4** for 1 hour prior to incubation with MV151[7] (0.5 μM) for 1 hour at 37 °C. Reaction mixtures were boiled with Laemmli's buffer containing β-mercaptoethanol for 5 min. before being resolved by 12.5% SDS-PAGE. In-gel detection of residual proteasome activity was performed in the wet gel slabs directly on the Typhoon Variable Mode Imager (Amersham Biosciences) using the Cy3/Tamra settings (λ_{ex} 532nm, λ_{em} 560 nm) to detect MV151.

Photocrosslinking in purified 20S proteasome

Purified 20S proteasome (human erythrocyte, 1 mg/mL, Enzo life sciences) was diluted with lysis buffer (50 mM Tris pH 7.5, 1 mM DTT, 5 mM $MgCl_2$, 250 mM sucrose, 2 mM ATP) to a concentration of 20 ng/μL. From this, 10 μL (200 ng 20S proteasome) was incubated with compound **2**, **3** or **4** (10 μM final concentration) for 2 h at 37 °C in the dark. The samples were irradiated with UV light (λ 365 nm) for the appropiate time at 0 °C, by placing the lamp (Spectroline® ENF-260C/FE, 6W) directly on top of the opened Eppendorff tubes (distance from light source to sample ~4 cm). After irradiation, the samples were stored at 4 °C in the dark. Next, the samples were treated with a Biotin-phospane reagent[8] (200 μM final concentration) for 1 h at 37 °C. After boiling the mixture with Laemmli's buffer containing β-mercaptoethanol for 5 min., the samples were resolved by 12.5% SDS-PAGE and all biotinylated proteins were visualized by Western blotting. The blots were blocked with 1% BSA in TBS-Tween 20 (0.1 % Tween 20) for 30 min. at RT, hybridized for 1 h with Streptavidin-HRP (1:10,000) in blocking buffer, washed and visualized using an ECL+ kit (Amersham Biosciences). Also, in a different gel, the total protein content was visualized by silverstain. The appropriate bands were cut from the gel and an in-gel digestion was performed according to the procedure described in literature.[9]

LC-MS/MS analysis

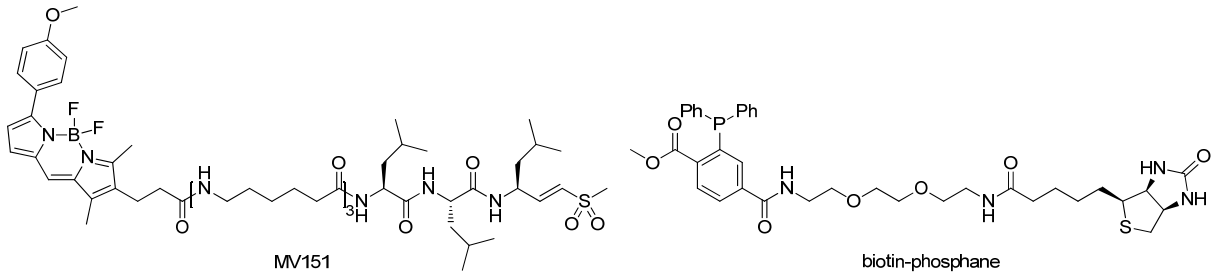
Tryptic peptides were analyzed on a Surveyor nanoLC system (Thermo) hyphenated to a LTQ-Orbitrap mass spectrometer (Thermo). Gold and carbon coated emitters (OD/ID = 360/25 μm tip ID = 5 μm), trap column (OD/ID = 360/100 μm packed with 25 mm robust Poros®10R2/15 mm BioSphere C18 5 μm 120 Å) and analytical columns (OD/ID = 360/75 μm packed with 20 cm BioSphere C18 5 μm 120 Å) were from Nanoseparations (Nieuwkoop, The Netherlands). The mobile phases (A: 0.1% formic acid/H_2O, B: 0.1% formic acid/ACN) were made with ULC/MS grade solvents (Biosolve). The emitter tip was coupled end-to-end with the analytical column via a 15 mm long TFE teflon tubing sleeve (OD/ID 0.3 × 1.58 mm, Supelco, USA) and installed in a stainless steel holder mounted in a nanosource base (Upchurch scientific, Idex, USA). General mass spectrometric conditions were: an electrospray voltage of 1.8 kV was applied to the emitter, no sheath and auxiliary gas flow, ion transfer tube temperature 150 °C, capillary voltage 41 V, tube lens voltage 150 V. Internal mass calibration was performed with air-borne protonated polydimethylcyclosiloxane (*m/z* = 445.12002) and the plasticizer protonated dioctyl phthalate ions (*m/z* = 391.28429) as lock mass.[10] For shotgun proteomics analysis, 10 μL of the samples was pressure loaded on the trap column with a 10 μL/min flow for 5 min. followed by peptide separation with a gradient of 35 min. 5 → 30% B, 15 min. 30 → 60% B, 5 min. A, at a flow of 300 μL/min. split to 250 nL/min. by the LTQ divert valve. For each data dependent cycle, one full MS scan (300-2000 *m/z*) acquired at high mass resolution (60,000 at 400 *m/z*, AGC target 1 × 106, maximum injection time 1,000 ms) in the Orbitrap was followed by 3 MS/MS fragmentations in the LTQ linear ion trap (AGC target 5 × 103, max injection time 120 ms) from the three most abundant ions.[11] MS/MS settings were: collision gas pressure 1.3 mT, normalized collision energy 35%, ion selection threshold of 500 counts, activation q = 0.25 and activation time of 30 ms. Fragmented precursor ions that were measured twice within 10 s were dynamically excluded for 60 s and ions with $z < 2$ or unassigned were not analyzed. Data from MS/MS was validated manually.

Table S1 Proteasome subunits identified from the indicated bands in Fig. 4b after in-gel tryptic digestion and LC-MS/MS analysis[a]

Accession #		Mass (Da)	Sequence coverage	*z*	ppm	Peptide score	Peptide sequence
IPI00479306	(β5)	22444	36.3%	2	1.42	33.3	AIYQATYR
				2	0.80	74.6	LLANMVYQYK
				2	1.48	60.0	ATAGAYIASQTVK
				2	1.47	59.7	GPGLYYVDSEGNR
				3	1.05	25.5	DAYSGGAVNLYHVR
				2	0.00	83.7	GYSYDLEVEQAYDLAR
IPI00025019	(β6)	26472	25.7%	3	1.13	20.8	LSEGFSIHTR
				2	0.06	49.9	NMQNVEHVPLSLDR
				3	0.42	36.6	AMTTGAIAAMLSTILYSR
				3	2.52	43.7	AMTTGAIAAMLSTILYSR
				2	5.84	83.0	AGGSASAMLQPLLDNQVGFK

[a] Protein name, mass of the β subunit, % coverage of the protein by amino acids identified by LC-MS, charge of the peptide (*z*), measurement error (ppm) and Mascot peptide scores. Mascot identifications were manually validated. M: oxidized methionine.

Structures of MV151[7] and biotin-phosphane[8]



References

1 M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg and H. Ploegh, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 6629-6634.

2 M. Nassal, *Liebigs Ann. Chem.*, 1983, 1510-1523.

3 M. H. Beale, R. Hooley, S. J. Smith and R. P. Walker, *Phytochemistry*, 1992, **31**, 1459-1464.

4 E. D. Goddard-Borger and R. V. Stick, *Org. Lett.*, 2007, **9**, 3797-3800.

5 J. T. Lundquist and J. C. Pelletier, *Org. Lett.*, 2001, **3**, 781-783.

6 J. Folwarczna, W. Janiec, L. Sliwinski and R. Paruszewski, *Pol. J. Pharmacol. Pharm.*, 1989, **41**, 585-589.

7 M. Verdoes, B. I. Florea, V. Menendez-Benito, C. J. Maynard, M. D. Witte, W. A. van der Linden, A. M. C. H. van den Nieuwendijk, T. Hofmann, C. R. Berkers, F. W. B. van Leeuwen, T. A. Groothuis, M. A. Leeuwenburgh, H. Ovaa, J. J. Neefjes, D. V. Filippov, G. A. van der Marel, N. P. Dantuma and H. S. Overkleeft, *Chem. Biol.*, 2006, **13**, 1217-1226.

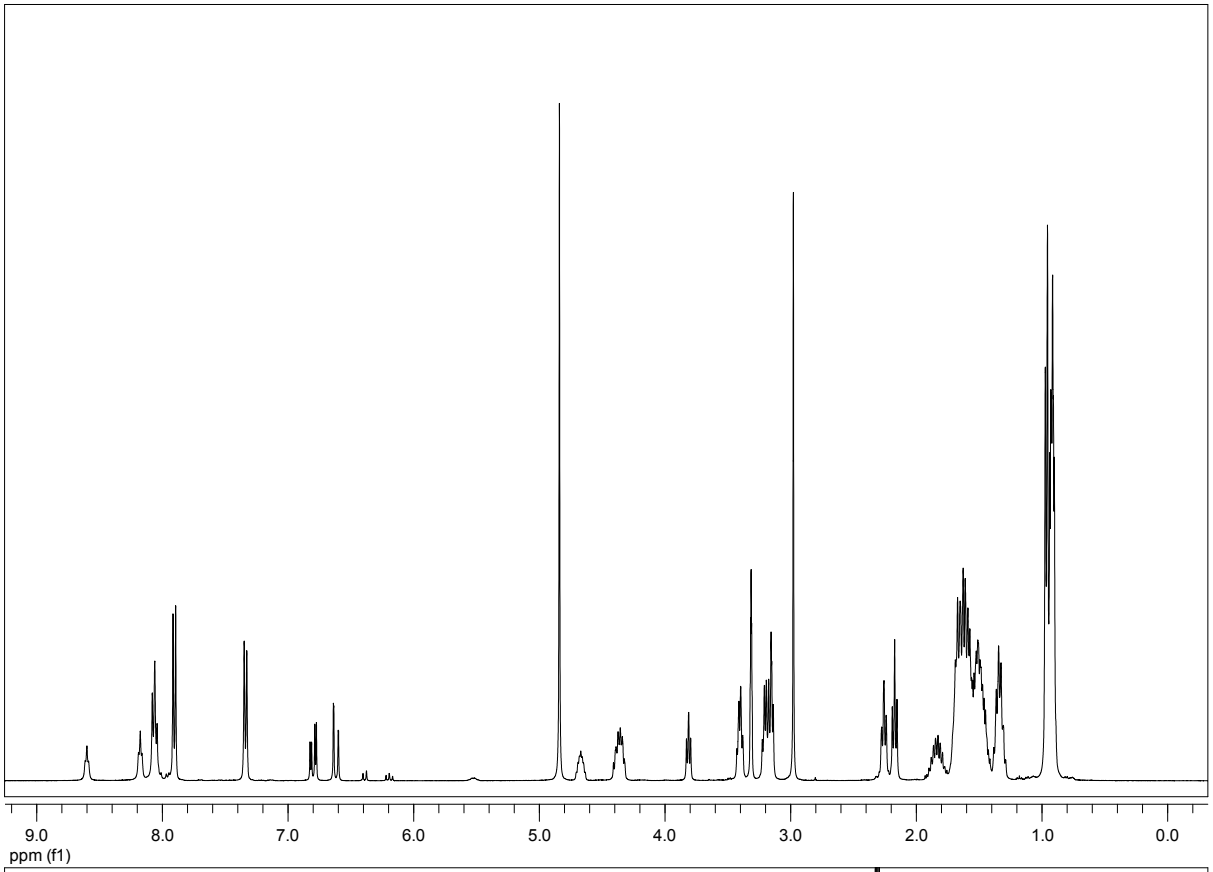
8 M. Verdoes, B. I. Florea, U. Hillaert, L. I. Willems, W. A. van der Linden, M. Sae-Heng, D. V. Filippov, A. F. Kisselev, G. A. van der Marel and H. S. Overkleeft, *ChemBioChem*, 2008, **9**, 1735-1738.

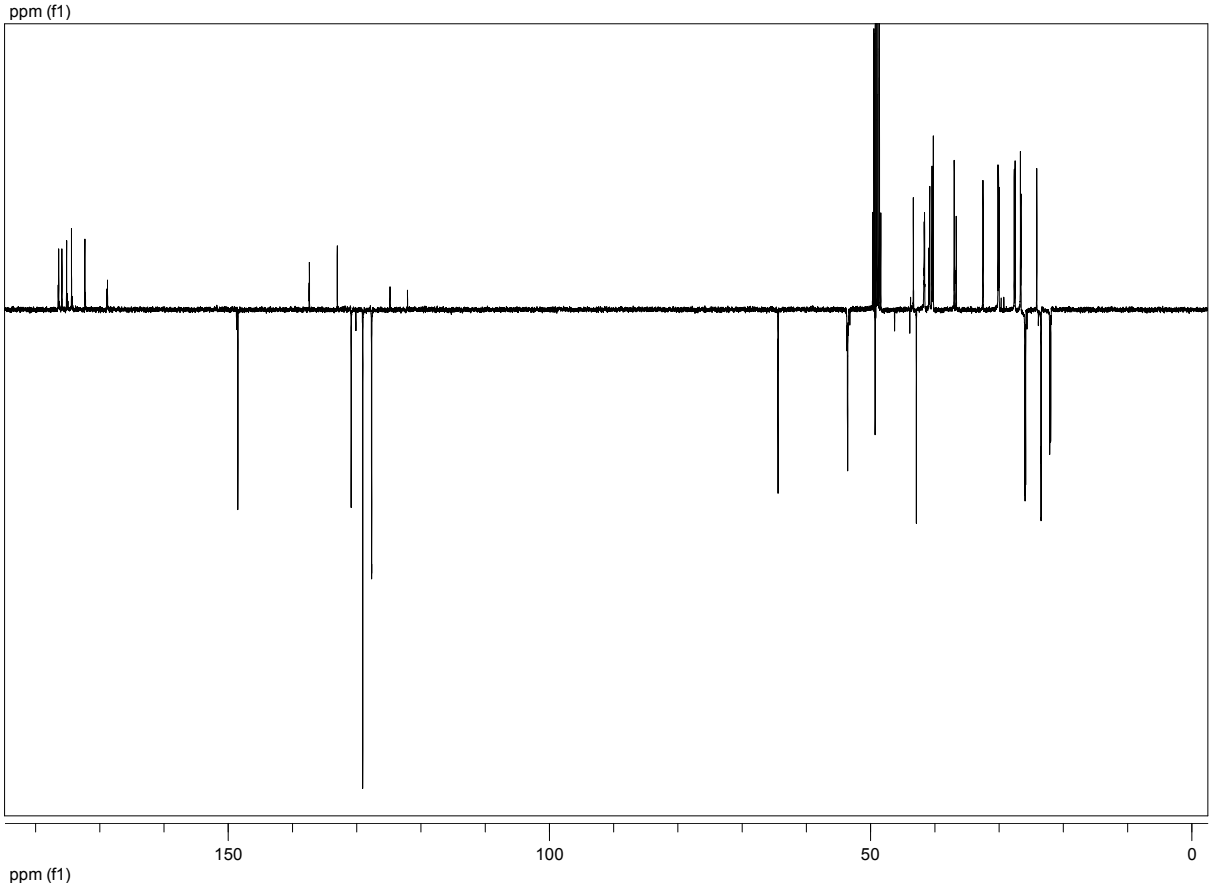
9 A. Shevchenko, H. Tomas, J. Havlis, J. V. Olsen and M. Mann, *Nat. Protoc.*, 2006, **1**, 2856-2860.

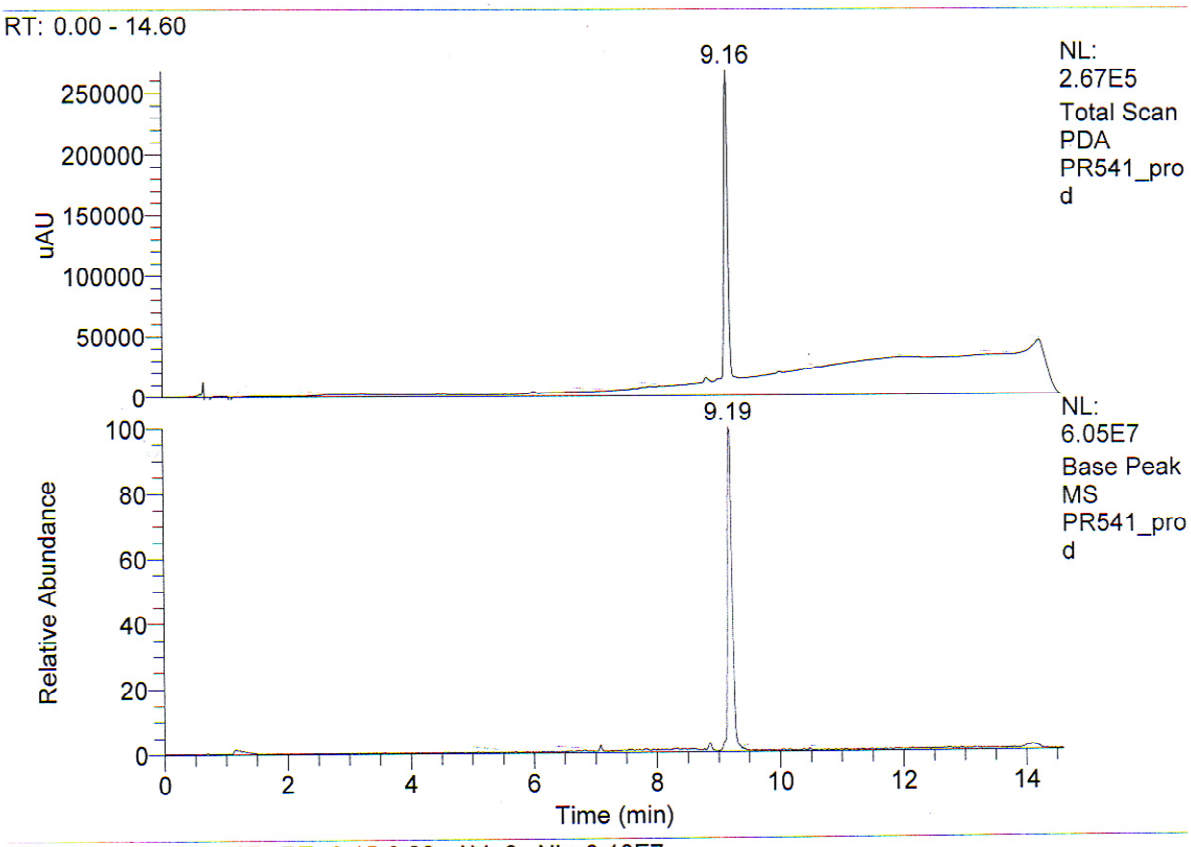
10 J. V. Olsen, L. M. F. de Godoy, G. Q. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning and M. Mann, *Mol. Cell. Proteomics*, 2005, **4**, 2010-2021.

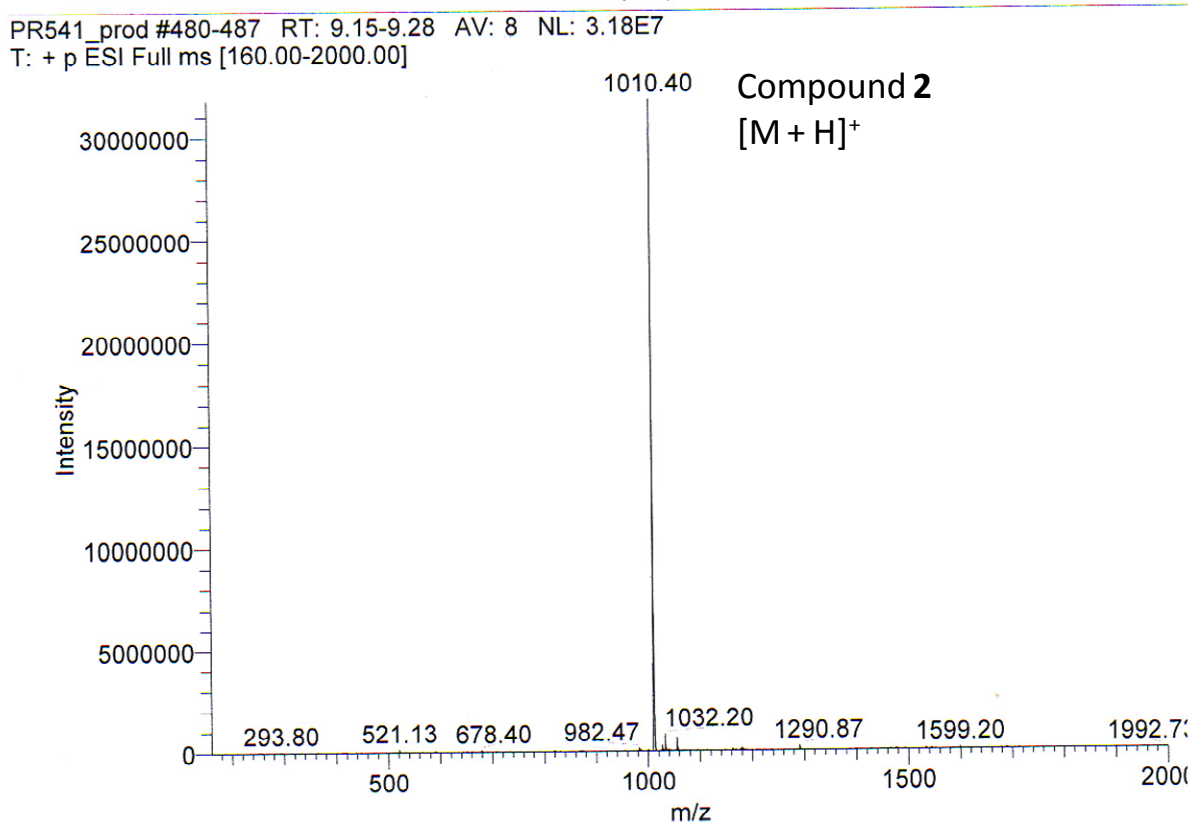
11 D. Baek, J. Villen, C. Shin, F. D. Camargo, S. P. Gygi and D. P. Bartel, *Nature*, 2008, **455**, 64-71.

Compound **2**

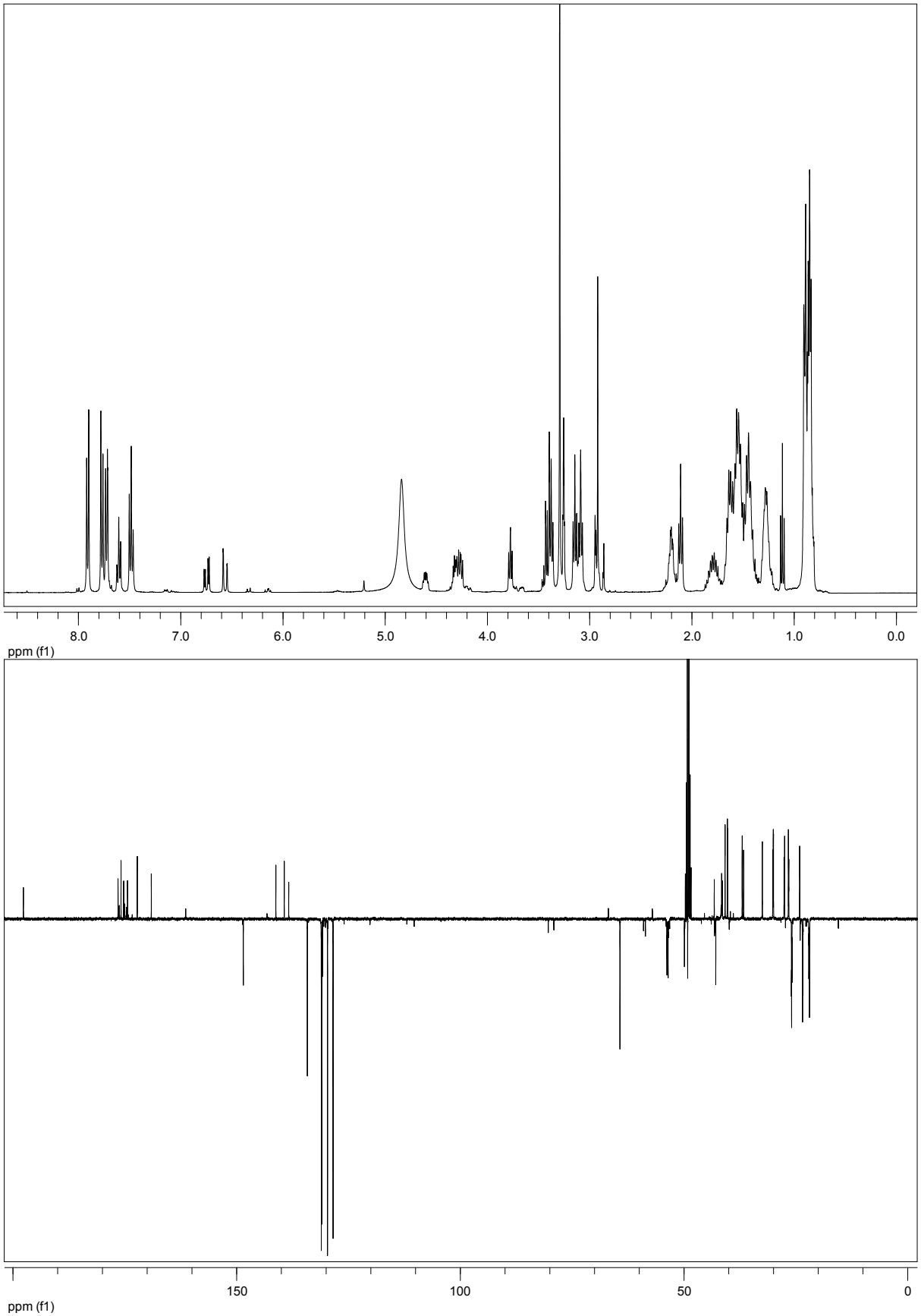


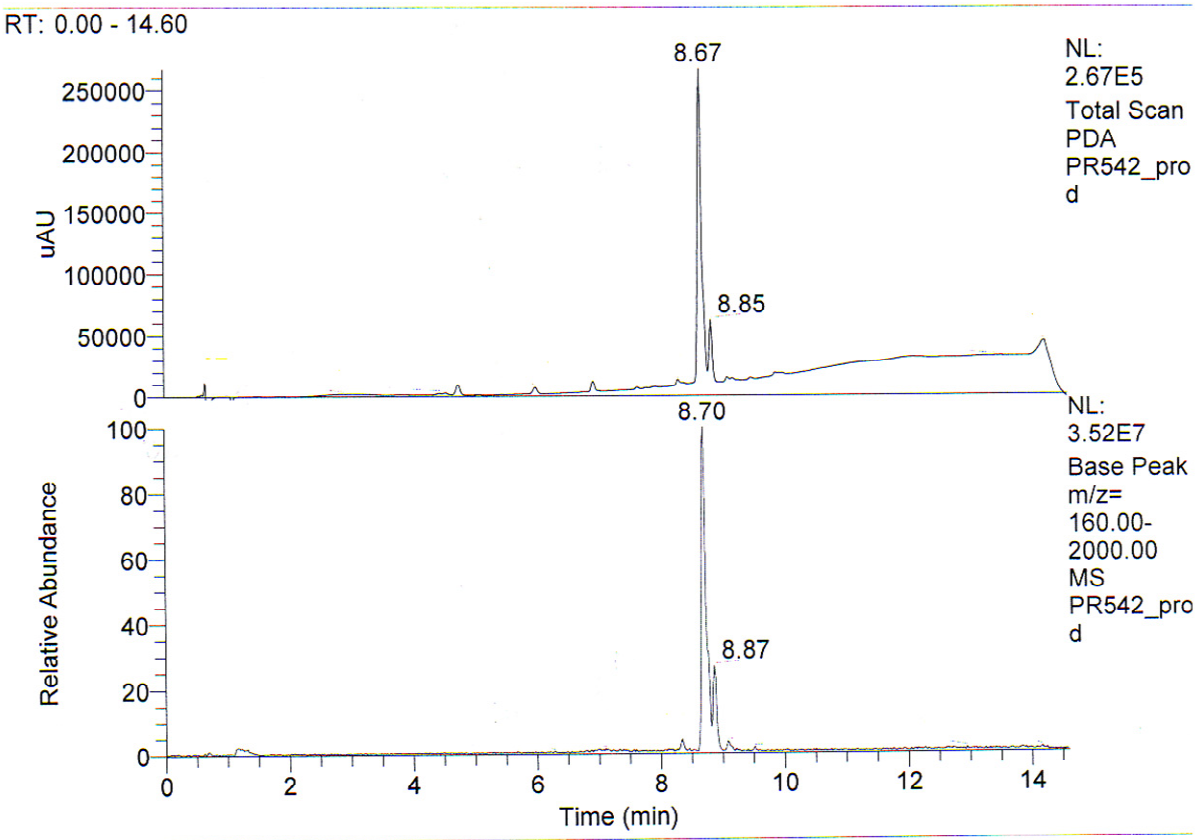


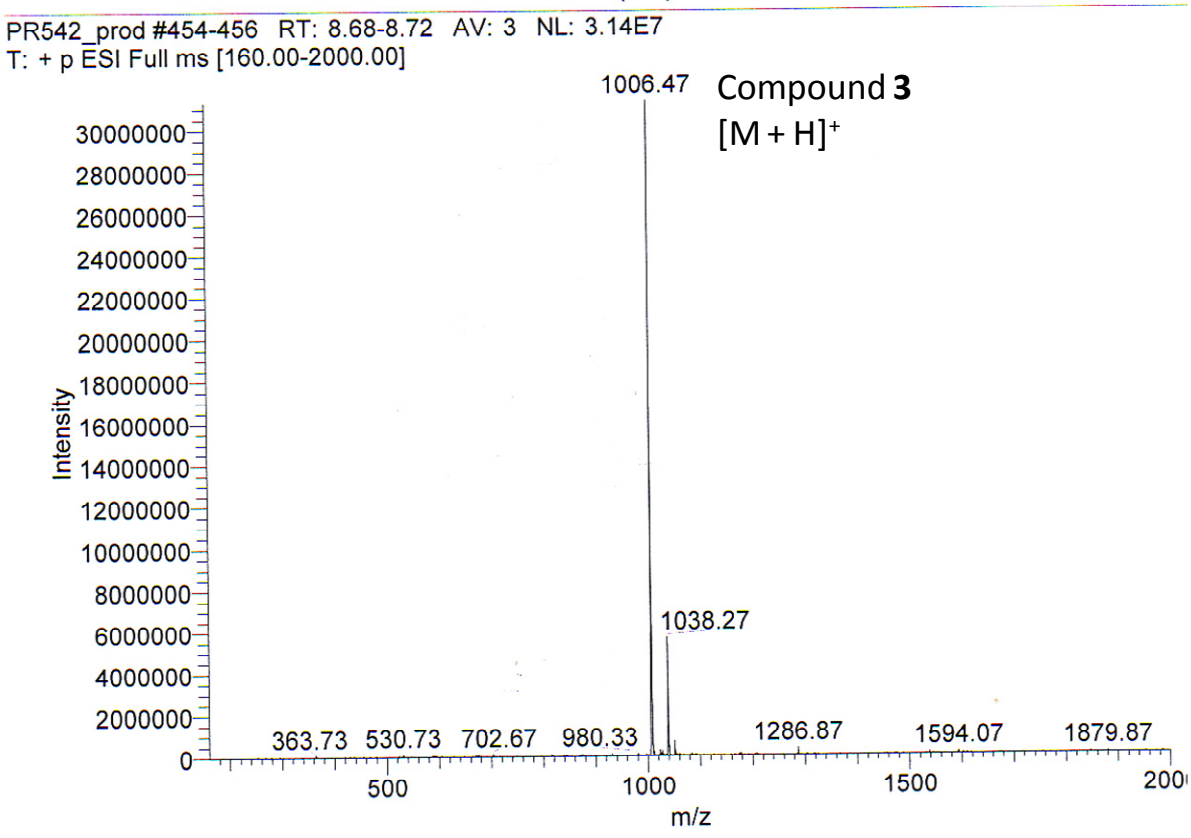

RT: 0.00 - 14.60
9.16
NL:
2.67E5
Total Scan
PDA
PR541_pro
d
uAU
250000
200000
150000
100000
50000
0
9.19
NL:
6.05E7
Base Peak
MS
PR541_pro
d
Relative Abundance
100
80
60
40
20
0
0
2
4
6
8
10
12
14
Time (min)


PR541_prod #480-487 RT: 9.15-9.28 AV: 8 NL: 3.18E7
T: + p ESI Full ms [160.00-2000.00]
1010.40
Compound 2
[M + H]+
30000000
25000000
20000000
15000000
10000000
5000000
0
Intensity
293.80
521.13
678.40
982.47
1032.20
1290.87
1599.20
1992.7
500
1000
1500
m/z

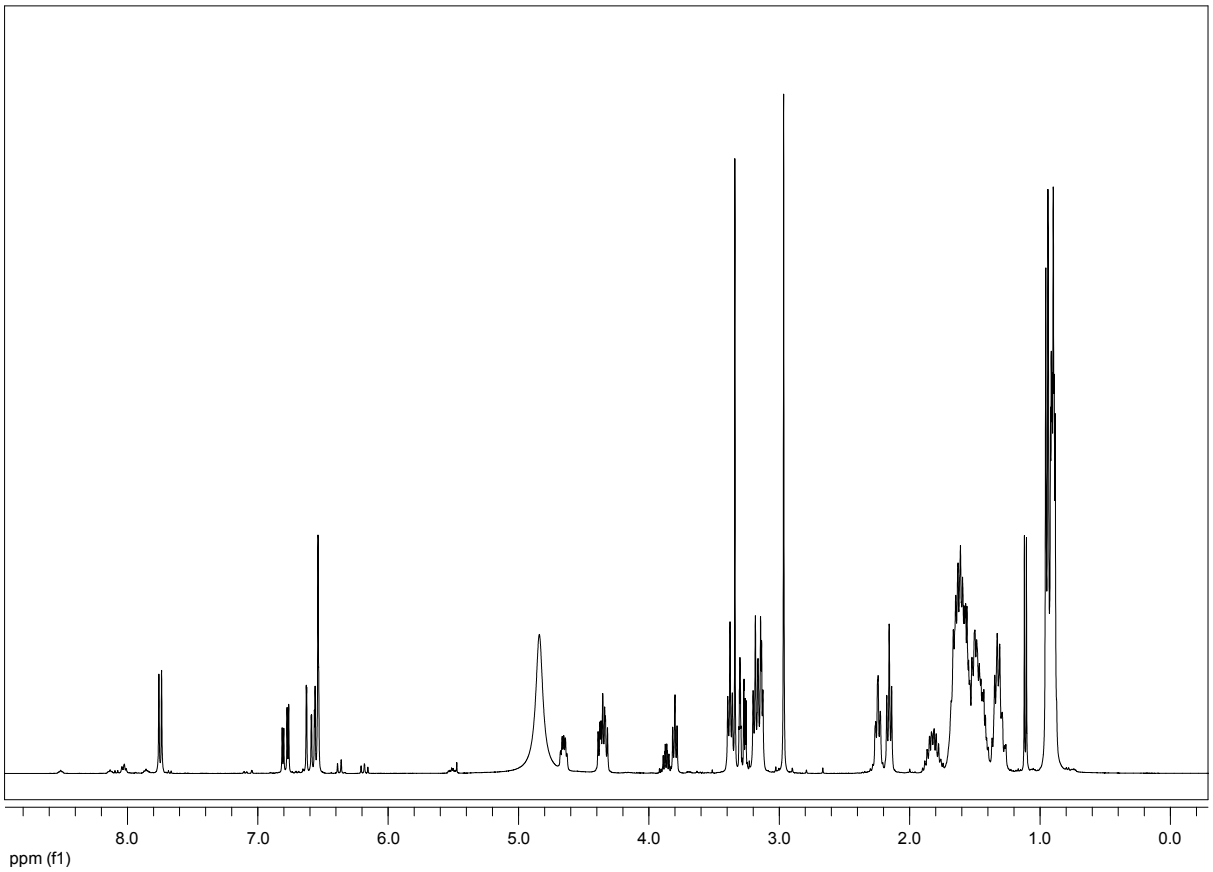
Compound **3**

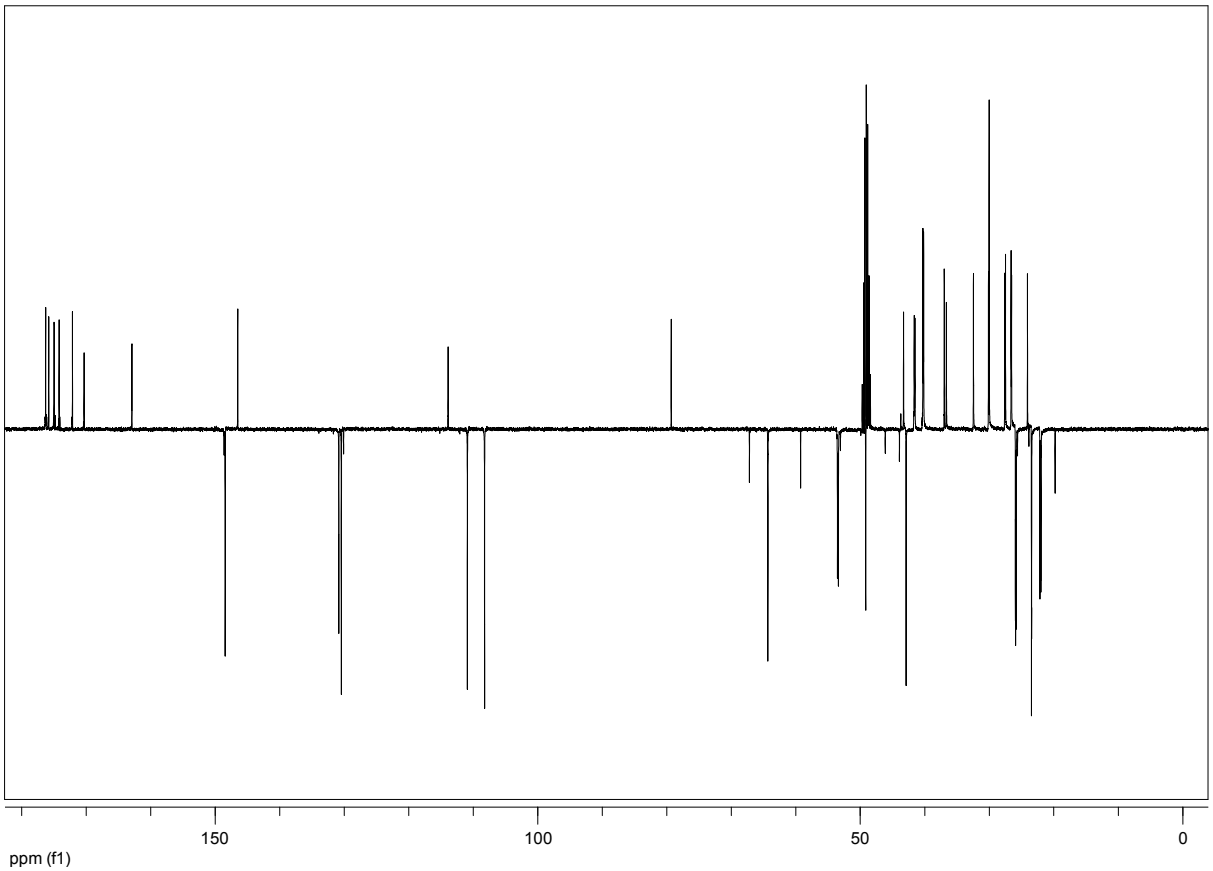


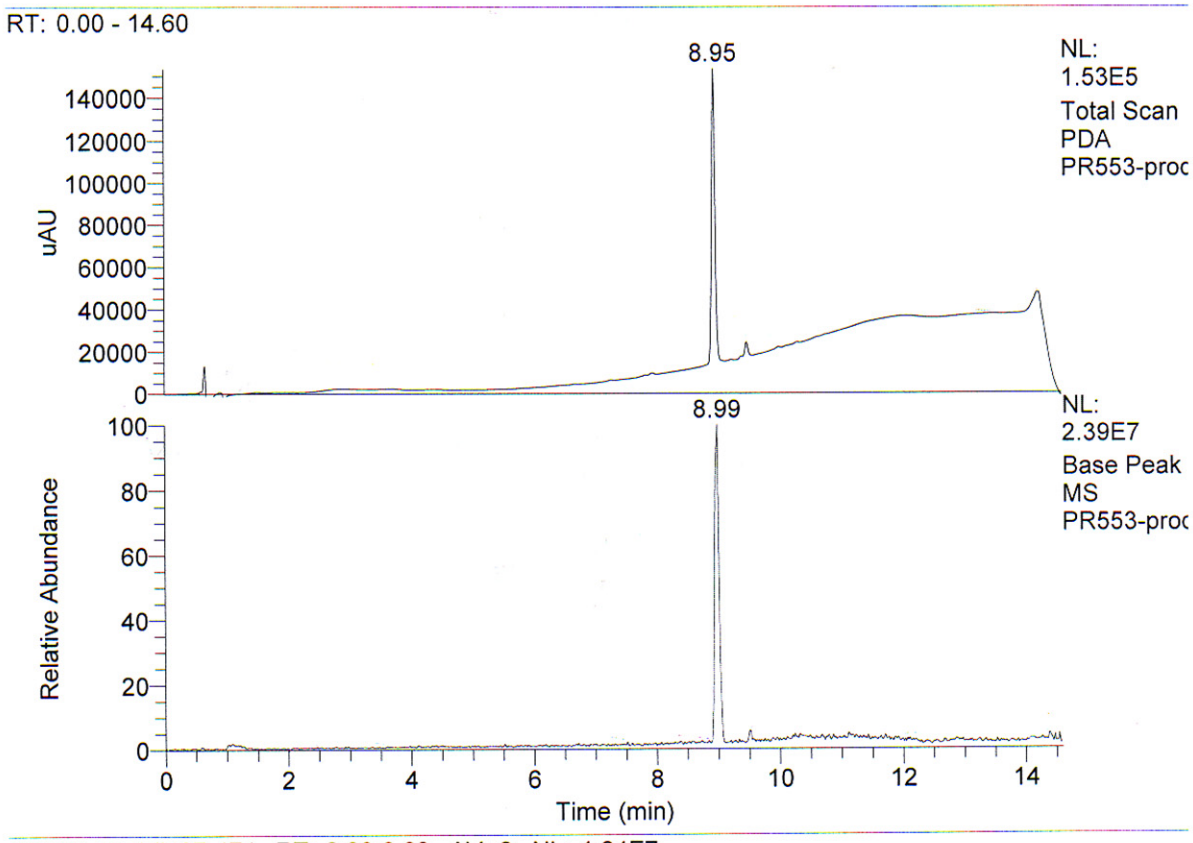

RT: 0.00 - 14.60
8.67
8.85
NL:
2.67E5
Total Scan
PDA
PR542_pro
d
uAU
250000
200000
150000
100000
50000
0
8.70
8.87
NL:
3.52E7
Base Peak
m/z=
160.00-
2000.00
MS
PR542_pro
d
Relative Abundance
100
80
60
40
20
0
0
2
4
6
8
10
12
14
Time (min)

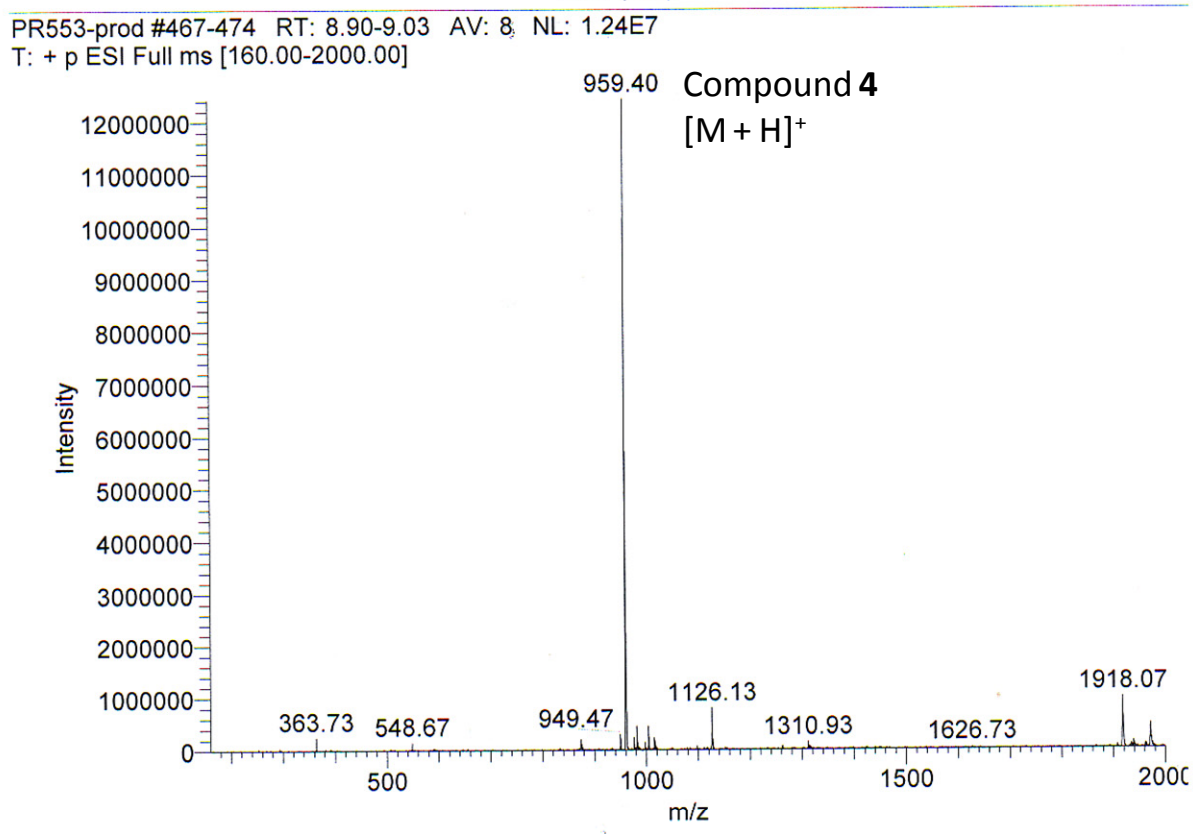

PR542_prod #454-456 RT: 8.68-8.72 AV: 3 NL: 3.14E7
T: + p ESI Full ms [160.00-2000.00]
1006.47
Compound 3
[M + H]+
30000000
28000000
26000000
24000000
22000000
20000000
18000000
16000000
14000000
12000000
10000000
8000000
6000000
4000000
2000000
0
Intensity
1038.27
363.73
530.73
702.67
980.33
1286.87
1594.07
1879.87
500
1000
1500
m/z

Compound **4**






RT: 0.00 - 14.60
8.95
NL:
1.53E5
Total Scan
PDA
PR553-proc
uAU
140000
120000
100000
80000
60000
40000
20000
0
8.99
NL:
2.39E7
Base Peak
MS
PR553-proc
Relative Abundance
100
80
60
40
20
0
0
2
4
6
8
10
12
14
Time (min)


PR553-prod #467-474 RT: 8.90-9.03 AV: 8 NL: 1.24E7
T: + p ESI Full ms [160.00-2000.00]
959.40
Compound 4
[M + H]+
Intensity
12000000
11000000
10000000
9000000
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
363.73
548.67
949.47
1126.13
1310.93
1626.73
1918.07
500
1000
1500
2000
m/z