Electronic Supplementary Material (ESI) for Chemical Science.

Supporting Information

Promiscuous indolyl vinyl isonitrile synthases in the biogenesis and diversification of hapalindole-type alkaloids

Kuljira Ittiamornkul,[a] Qin Zhu,[a] Danai S. Gkotsi,[b] Duncan R. M. Smith,[b] Matthew L. Hillwig,[a] Nicole Nightingale,[a] Rebecca J. M. Goss[b] and Xinyu Liu*[a]

[a] Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Pittsburgh, PA, USA 15260

[b] School of Chemistry and BSRC, University of St. Andrews, St. Andrews, KY16 9ST, U.K.

* To whom correspondence should be addressed: xinyuliu@pitt.edu

SI Materials and Methods

General methods: All PCRs were carried out on a C1000 thermal cycler (Bio-Rad). DNA sequencing was performed by Elim BioPharm Inc. Analytical reverse-phase HPLC was performed using a Dionex UHPLC with a photo-diode array UV/Vis detector (Thermo Fisher Scientific) and a Luna C18 column (5 μm, 4.6 × 250 mm, Phenomenex). HRMS analysis was conducted using a Q Exactive Benchtop Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a Dionex RSLC (Thermo Fisher Scientific). The NMR spectrum was recorded on a Bruker Avance III 600 MHz spectrometer or a Bruker Avance III 700 MHz spectrometer.

Materials: Synthetic oligonucleotides for gene amplification by PCR were purchased from Life Technologies or Integrated DNA Technology. Kappa HiFi DNA polymerase was obtained from Kappa Biosystems. LB broth and agar used for culturing *E. coli* were obtained from Teknova. M9 minimal media was homemade to contain the final concentration of 2 mM Mg_2SO_4, 0.4% glucose, 0.1 mM $CaCl_2$, 1x M9 salts (12.65 mM $Na_2HPO_4.7H_2O$, 22.04 mM KH_2PO_4, 8.56 mM NaCl, 18.6 mM NH_4Cl)**.** All halo-substituted L-Trp except 6-F L-Trp were obtained via a chemoenzymatic route as previously described.[1] All other reagents including inorganic salts and cofactors were purchased from Sigma-Aldrich or Fisher Scientific unless otherwise stated.

Strains and plasmids: *E. coli* TOP10 cell (Life Technologies) was used for routine cloning, plasmid propagation and heterologous expression of *amb*I1-3 genes in this study. *Saccharomyces cerevisiae* INVSc1 strain (Life Technologies) was used in recombineering-assisted cloning.

General methods for yeast recombineering: Plasmid vector pMQ123i[2] was cut with SmaI restriction enzyme and dephosphorylated by Antarctic phosphatase. Primers were designed to PCR amplify DNA to be joined with the linearized vectors. *S. cerevisiae* (INVSc1) was grown overnight in yeast extract-peptone-dextrose (YPD) at 30°C and transformed using a modified "lazy bones"

transformation protocol where 20 to 200 ng of the dephosphorylated linearized vector and 50 to 500 ng of the amplicon were added. Recombinants were selected on uracil dropout plates. Plasmids were subsequently recovered from yeast using the “smash and grab” method. For screening, the isolated plasmids were transformed into *E. coli*, where gentamicin (10 µg/mL) was used for selection and propagation.

Cloning of pMQ123i_*amb*I1-3: The genes were PCR amplified from UTEX1903 genomic DNA, using the following primers: forward, GAA TTG GAT CCT CTA GAT TCT CCA TAC AGG AGG AAT AAT ATG ATT AGT GAA AAA ATT CTC (*amb*I1-5'-connect123i); reverse, CAT TTA TAT CCT CCT ACG GGT ATG GAG AAC TAA CTC TTG TTG TCA AG (*amb*I1-3'-connectI2); forward, TTC TCC ATA CCC GTA GGA GGA TAT AAA TGA CTC AAA TTA TCA ATA TCA C (*amb*I2-993bp-5'-connectI1); reverse, TCA GAC CGC TTC TGC GTT CTG ATT TAT TAT AAA ATA TGT ACC CGT TGC (*amb*I3-3'-connect123i). The genes were recombined into pMQ123i by yeast homologous recombination approach as described.[2] Plasmid DNA was isolated from selected colonies and digested with appropriate restriction enzymes to confirm positive clones.

Synthesis of 1d and 1f: The procedure for preparing **1d** and **1f** was adopted from that published by Hoppe and Schollkopf.[3] In brief, a solution of diethylisocyanomethyphosphate (0.5 mL, 3 mmol) in THF (2.5 mL) was added dropwise to a stirred solution of sodium bis(trimethylsilyl)amide (3 mmol) in THF (3.5 mL) at -78 °C. The mixture was stirred for 15 min and a solution of 6-fluoro-3-indolecarboxaldehyde or 5-chloro-3-indolecarboxaldehyde (3.0 mmol) in THF (20 mL) was added dropwise. After stirring for ca. 20 h at 0 °C, the reaction was quenched by the addition of acetic acid (198 mg, 3.3 mmol) in THF (1.5 mL). The solvent was removed *in vacuo*. The residue was taken up in ethyl acetate (30 mL), washed with phosphate buffer (15 mL, pH=7) and water **(**15 mL), and dried with Na_2SO_4. The residue was purified by silica gel flash column chromatography to give **1d** (2.8 mg) along with its *trans*-isomer (10.3 mg) and **1f** (6.1 mg) along with its *trans*-isomer (11.7 mg). **1d**:

brown powder, ^{1}H NMR (700 MHz, CD_3OD) (Fig. SI-1) δ 8.0 (brs, 1H), 7.60 (dd, *J* = 8.7, 5.2 Hz, 1H), 7.10 (dd, *J* = 9.6, 2.0 Hz, 1H), 6.90 (ddd, *J* = 9.6, 8.6, 2.3 Hz, 1H), 6.85 (m, 1H), 5.80 (d, *J* = 8.8 Hz, 1H); ^{13}C NMR (175 MHz, CD_3OD) (Fig. SI-2) *δ* 167.66, 159.55 (d, *J* = 234 Hz), 135.56 (d, *J* = 12.3 Hz), 126.83, 124.15, 123.50, 118.52 (d, *J* = 10.5 Hz), 109.38, 108.38 (d, *J* = 24.5 Hz), 103.49 (t, *J* = 10.5 Hz), 97.18 (d, *J* = 26.2 Hz). ***trans*-1d**: brown powder, ^{1}H NMR (600 MHz, CD_3OD) (Fig. SI-1) *δ* 7.68 (dd, *J* = 8.7, 5.2 Hz, 1H), 7.50 (s), 7.21 (d, *J* = 14.4 Hz), 7.13 (dd, *J* = 9.6, 2.2 Hz, 1H), 6.95 (ddd, *J* = 9.2, 9.1, 2.3 Hz, 1H), 6.52 (d, *J* = 14.4 Hz, 1H); ^{13}C NMR (150 MHz, CD_3OD) (Fig. SI-2) *δ* 161.18 (d, *J* = 53.2 Hz), 159.26, 137.70 (d, *J* = 12.4 Hz), 130.55, 128.20 (d, *J* = 2.4 Hz),, 121.30, 120.28 (d, *J* = 10.0 Hz), 109.99, 108.70 (d, *J* = 24.3 Hz), 105.72 (t, *J* = 12.2 Hz), 97.72 (d, *J* = 26.2 Hz). **1f**: brown powder, ^{1}H NMR (600 MHz, CD_3OD) (Fig. SI-3) *δ* 8.09 (s), 7.72 (d, *J* = 1.8 Hz), 7.42 (d, *J* = 8.5 Hz), 7.19 (dd, *J* = 8.5, 1.8 Hz), 6.88 (m), 5.88 (d, *J* = 8.8 Hz); ^{13}C NMR (150 MHz, CD_3OD) (Fig. SI-4) *δ* 167.76, 134.03, 128.03, 127.73, 125.89, 123.92, 122.46, 117.15, 115.97, 112.57, 109.05, 103.63 (t, *J* = 12.0 Hz).

***In vitro* assay of AmbI1-3 with L-Trp and halo-substituted L-Trp:** N-His_7-tagged AmbI1**,** AmbI2, AmbI3 protein were overexpressed and purified as previously described.[4] A standard assay with either a single or double L-Trp was conducted in a 1-mL scale at pH 7.4 with **2** (2.5 mM), Rub5P (2.5 mM), α-KG (2.0 mM), $(NH_4)_2Fe(SO_4)_2$ (0.5 mM) and AmbI1, I2, I3 (20 μM each) for 4 h at 30 °C. All assays were stopped by extraction with ethyl acetate (1 mL x 2). The combined organic layer was dried under a stream of N_2 gas, and redissolved in methanol (100 μL). For product verification, a 20-μL aliquot was used for HPLC and LC-HRMS analysis. The authenticities of all new halogen substituted indole vinyl isonitriles were verified by HRMS (Fig. SI-5 to 13) and UV spectra obtained by a photodiode array detector coupled to HPLC. *cis*-Indole vinyl isonitriles (**1a**–**j**) share common absorptions at 220-230 nm and 280-320 nm (Fig. SI-14). For **1d** and **1f**, synthetic standards were obtained (see above). They were used for co-elution experiment in HPLC analysis. Standard curves

that correlate the mass of **1d** and **1f** with the extent of UV absorbance at 315 nm by HPLC analysis were constructed for the quantification of **2d** and **2f** turnover by AmbI1-3 *in vitro*.

Directed biosynthesis of halogen-substituted indole vinyl isonitriles in *E. coli*:

Individual colony of *E. coli* TOP10 cell containing *amb*I1-I3-pMQ123i plasmid was picked and inoculated into a 3 mL LB culture overnight which was further inoculated to a 100 mL M9 minimal media (2 mM Mg_2SO_4, 0.4% glucose, 0.1 mM $CaCl_2$, 1x M9 salts containing 12.65 mM $Na_2HPO_4.7H_2O$, 22.04 mM KH_2PO_4, 8.56 mM NaCl, 18.6 mM NH_4Cl)) initially supplemented with 50 µM halo-substituted L-Trp at 37°C overnight until OD_{600} reached 0.7-0.9. Then it was supplemented with additional 200 µM halo-substituted L-Trp and induced for 2 days by the addition of isopropylthio-β-galactoside (IPTG) to a final concentration of 1 mM. To extract, the culture was centrifuged and the supernatant was transferred into a separating funnel, mixed with an equal volume of ethyl acetate, and shaken vigorously. The mixture was then allowed to stand for the layers to be separated. The extraction was repeated twice. The organic layer was dried with Na_2SO_4, filtered and evaporated to dryness. The crude extract was dissolved in methanol (100 µL), centrifuged to remove particles, and analyzed by HPLC and LC-HRMS. A gradient from 5-70% acetonitrile in 35 min was used for the mobile phase. For quantification of **1d**, a standard curve that correlates the mass of **1d** with the extent of UV absorbance at 315 nm by HPLC analysis was constructed using synthetic **1d**.

Antibacterial activity assay of 1a, 1e and 1h: Compounds **1e** and **1h** were isolated to homogeneity by semi-preparative HPLC from 4-L cultures of *E. coli* TOP10 cells containing *amb*I1-I3-pMQ123i plasmid in M9 medium and supplemented with **2e** and **2h** respectively. Compound **1a** was obtained synthetically.[3] Antimicrobial susceptibility of *Vibrio cholerae* N16961, *Escherichia coli* Imp, *Bacillus subtilis* 168 towards **1a, 1e** and **1h** were determined using disc diffusion method. Blank discs were impregnated with 250 nmol of **1a** or **1e** or **1h** (5 µL of 50 mM stock solution of **1a** or **1e** or **1h** in DMSO) and allowed to dry. Bacterial isolates were grown overnight on LB medium and diluted into

OD_{600} = 0.01. These were used to inoculate LB agar plates by streaking swabs over the entire agar surface followed by the application of the respective isonitrile containing discs. Plates were then incubated for 16 h at 30 °C. Testing was done in duplicate. Zone diameters were determined and averaged. Pure DMSO (5 µL) was used as a negative control for each bacterial strain, in which all bacteria grew normally with no inhibition zone observed.

Figure SI-1. ^{1}H NMR spectra of **1d** (CD_3OD, 700 MHz) and ***trans*-1d** (CD_3OD, 600 MHz)

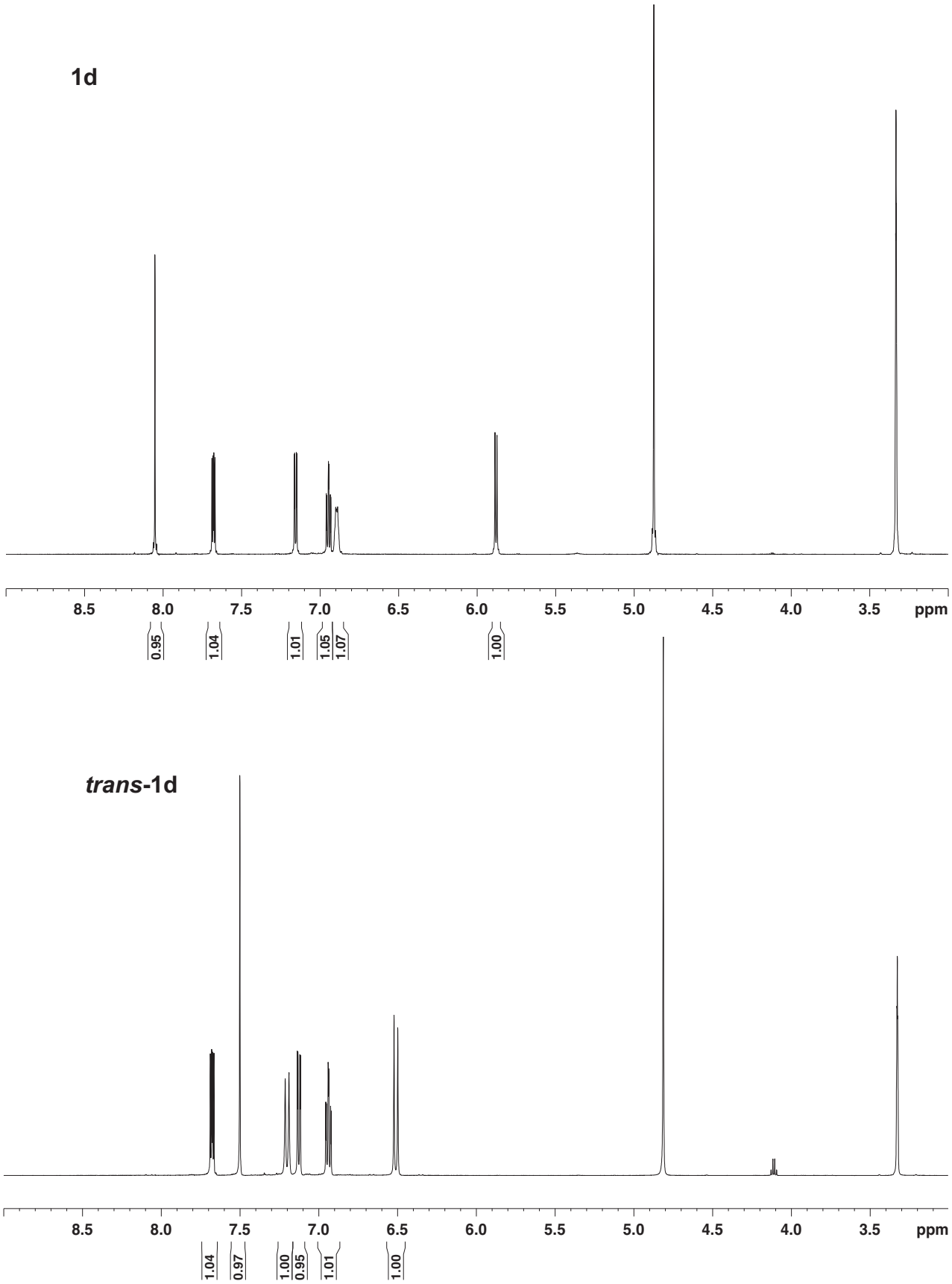
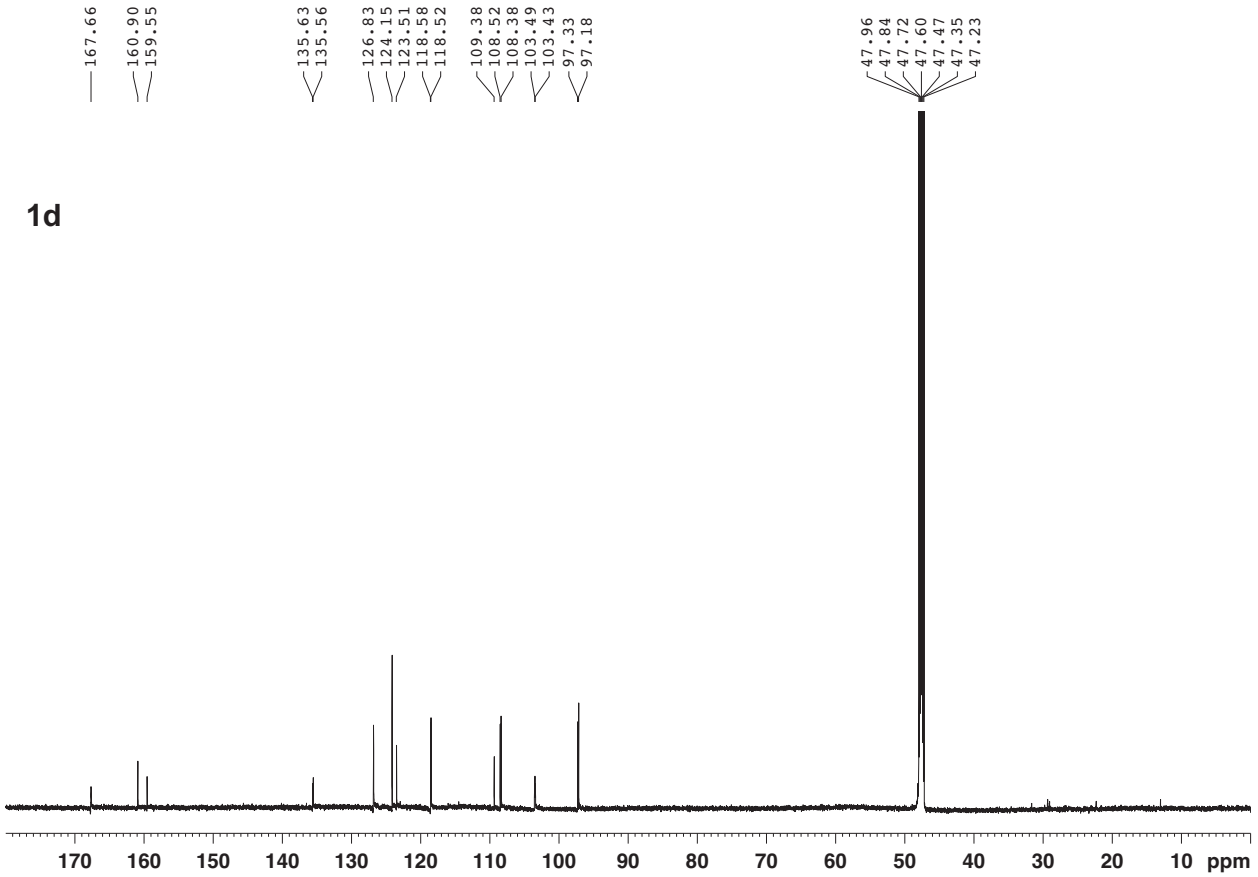


Figure SI-2. ^{13}C NMR spectra of of **1d** (CD_3OD, 700 MHz) and ***trans*-1d** (CD_3OD, 600 MHz)



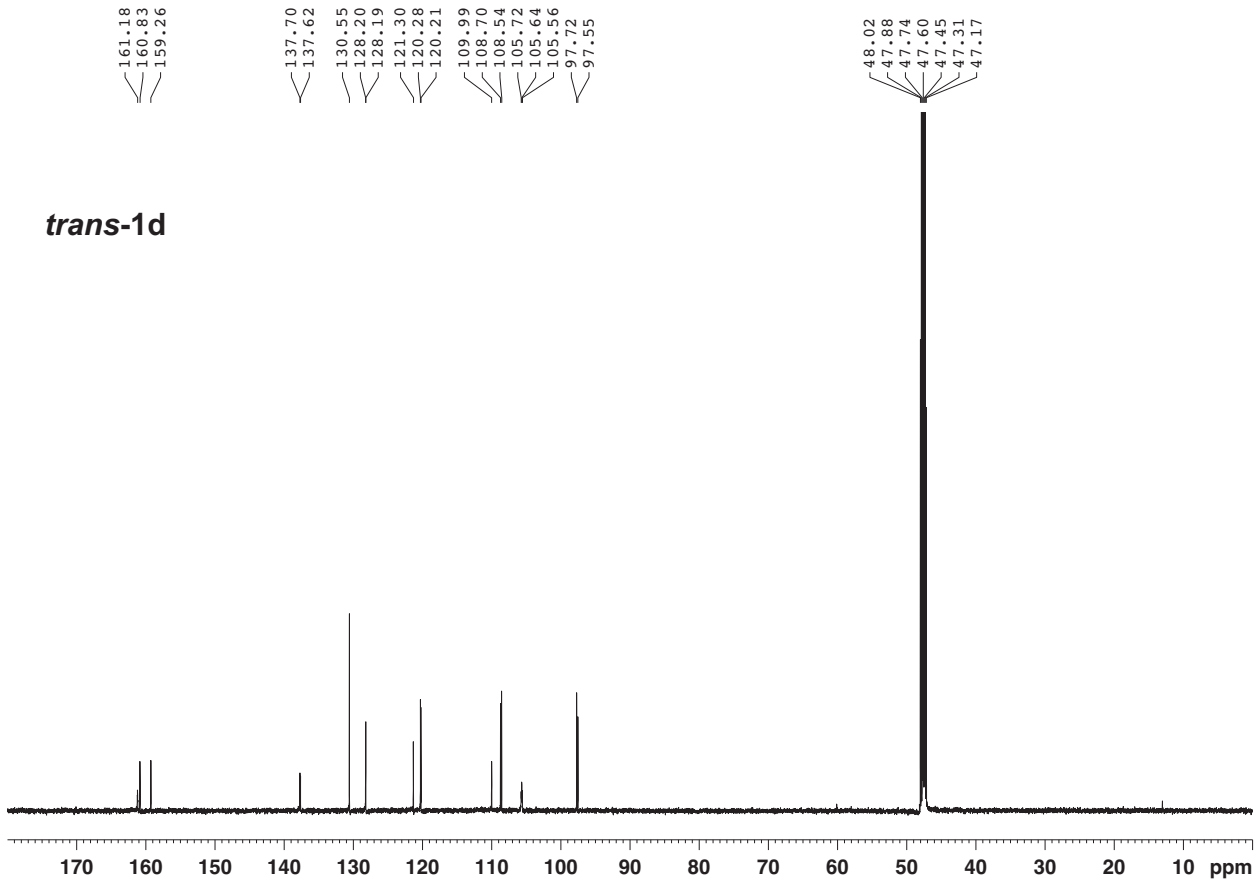


Figure SI-3. ^{1}H NMR (CD_3OD, 600 MHz) spectrum of **1f**.

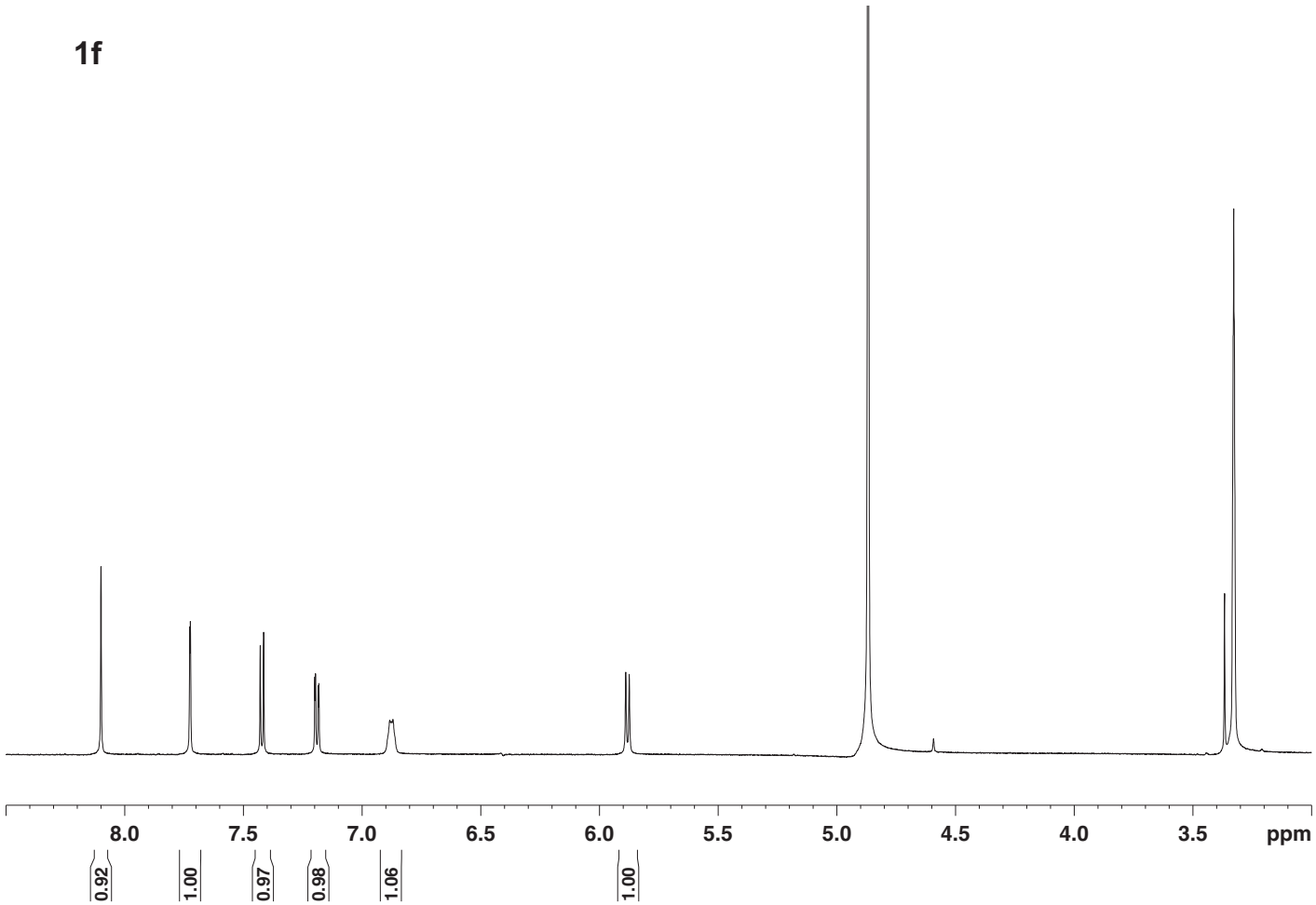


Figure SI-4. ^{13}C NMR (CD_3OD, 150 MHz) spectrum of **1f**.

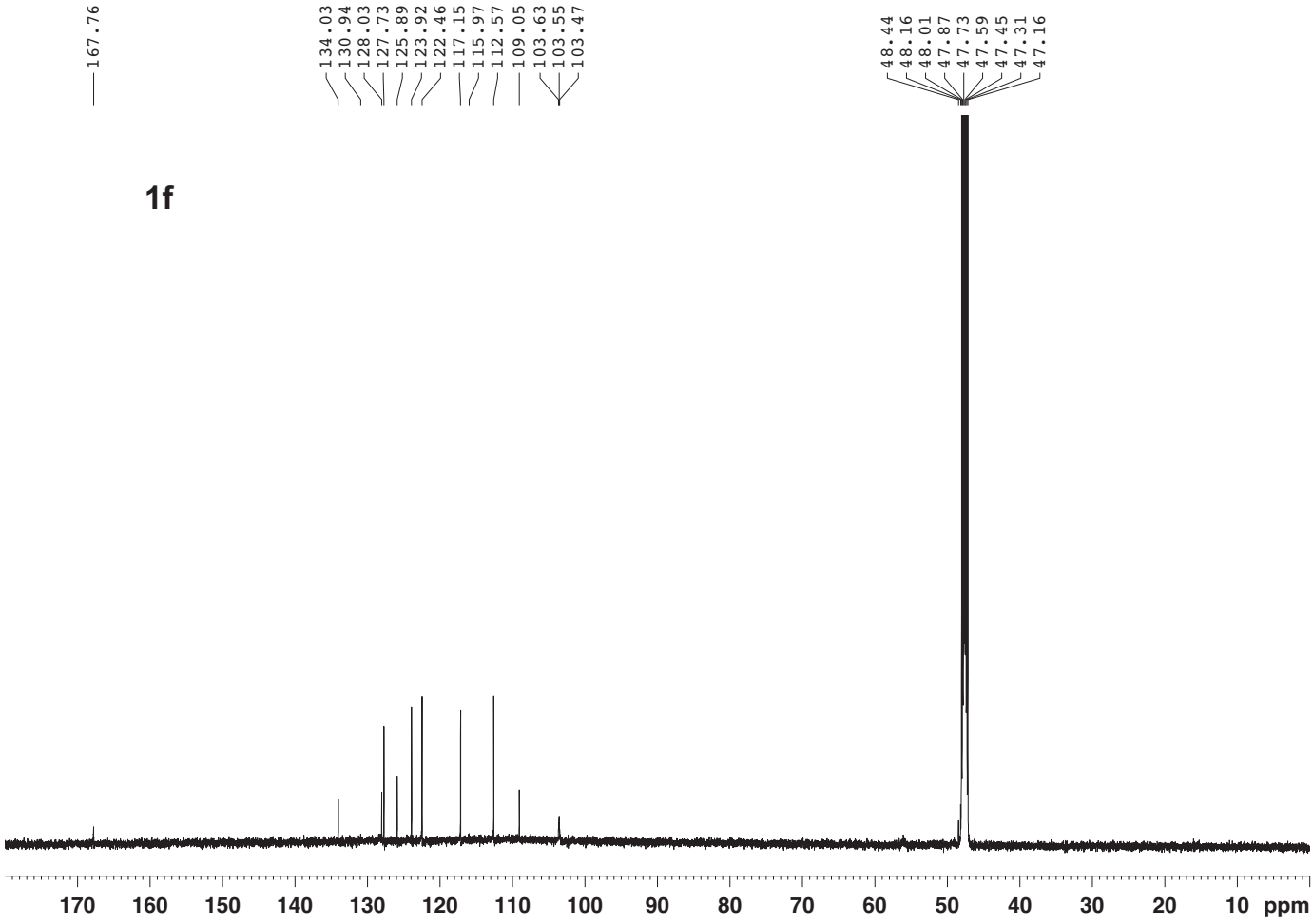


Figure SI-5. HRMS of **1b**

c:\xcalibur\data\xliu\72729lcmsp 10/15/2014 12:42:59 PM Kuljira-4F-Liu

72729lcmsp#4787-4897 RT: 25.86-26.46 AV: 111
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 187.00000-188.00000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
187.06611	189632240.0	100.00	187.06660	-0.50	$C_{11}H_8N_2F$

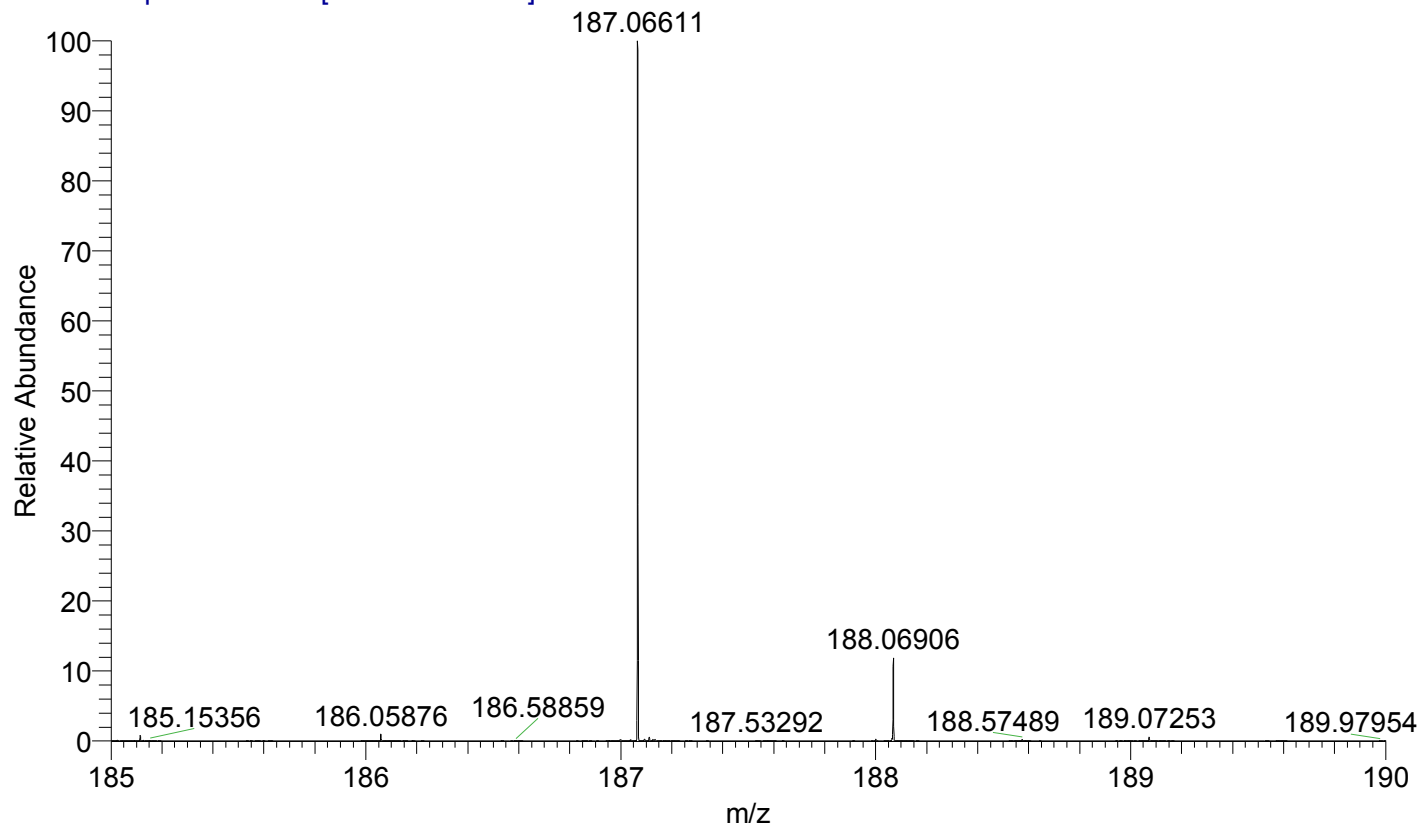


Figure SI-6. HRMS of **1c**

c:\xcalibur\data\xliu\72665lcmsp2 10/8/2014 4:46:17 PM 5F

72665lcmsp2#4775-4849 RT: 26.09-26.49 AV: 75
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 187.00000-188.00000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
187.06677	7089549.0	100.00	187.06660	0.91	$C_{11}H_8N_2F$

72665lcmsp2 #4775-4849 RT: 26.09-26.49 AV: 75 NL: 7.03E6
T: FTMS + p ESI Full ms [100.00-1000.00]

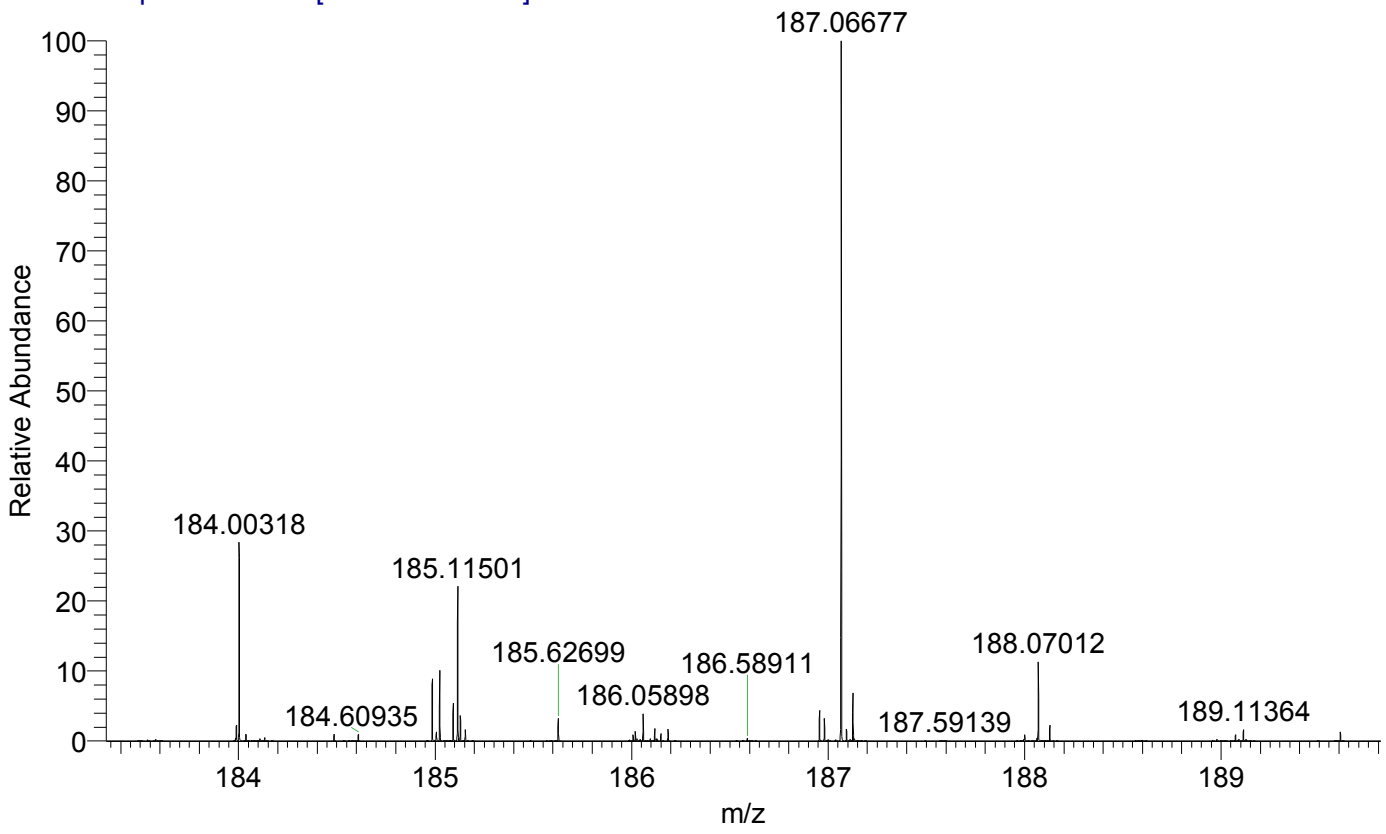


Figure SI-7. HRMS of **1d**

C:\Xcalibur\data\XLiu\72675LCMSP 10/8/2014 10:58:06 PM KULJIRA-6FV-LIU

72675LCMSP#4845-4900 RT: 26.19-26.49 AV: 56
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 187.00000-188.00000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
187.06712	17003176.0	100.00	187.06660	0.51	$C_{11}H_8N_2F$

72675LCMSP #4845-4901 RT: 26.19-26.49 AV: 57 SB: 518 24.72-25.99 , 26.73-28.24 NL: 1.65E7
T: FTMS + p ESI Full ms [100.00-1000.00]

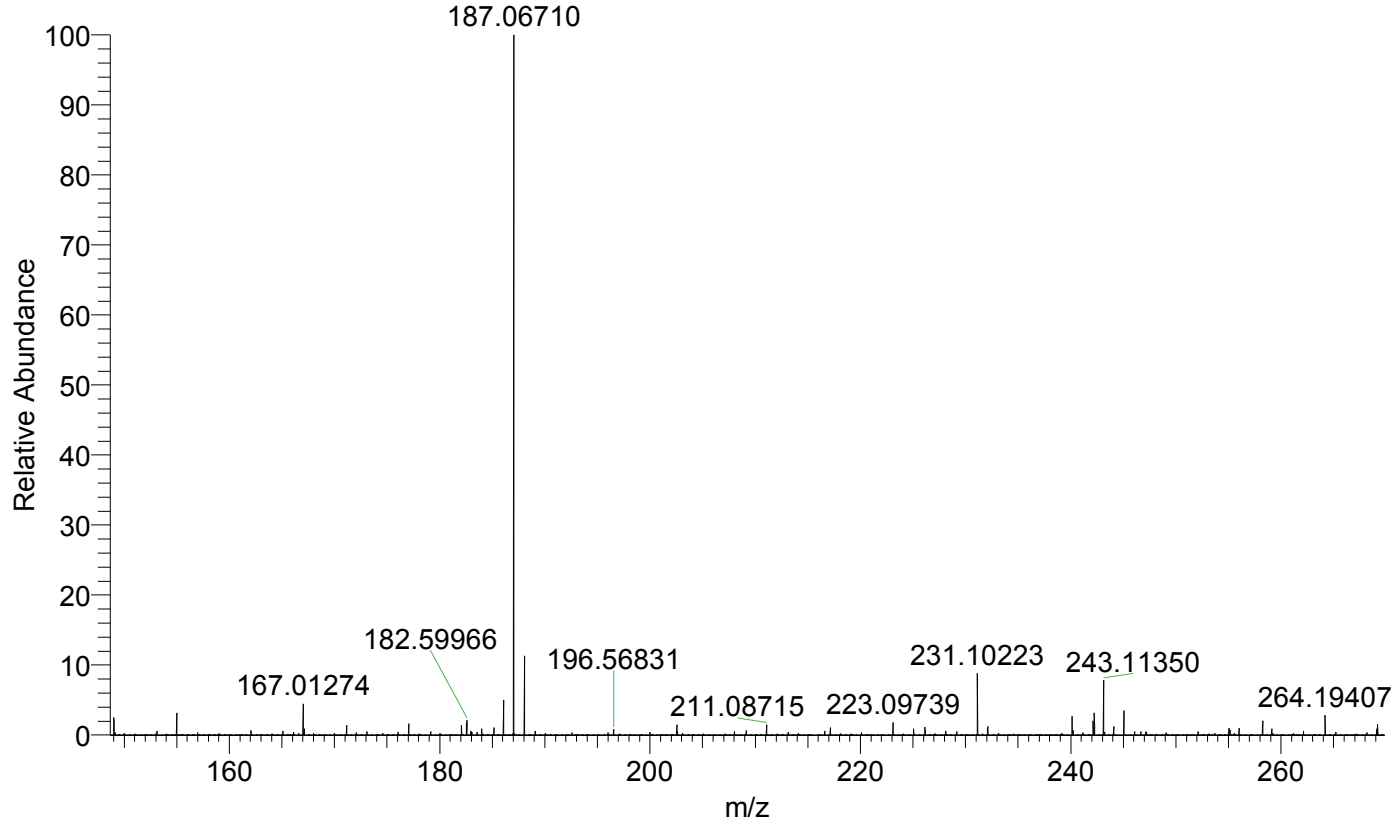


Figure SI-8. HRMS of **1e**

c:\xcalibur\data\xliu\72668lcmsp2 10/8/2014 6:50:12 PM 7F

72668lcmsp2#4826-4908 RT: 26.36-26.80 AV: 83
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 187.00000-187.10000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
187.06702	3516525.8	100.00	187.06660	2.23	$C_{11}H_8N_2F$

72668lcmsp2 #4826-4907 RT: 26.36-26.80 AV: 82 SB: 462 25.08-25.99 , 27.43-29.01 NL: 3.45E6
T: FTMS + p ESI Full ms [100.00-1000.00]

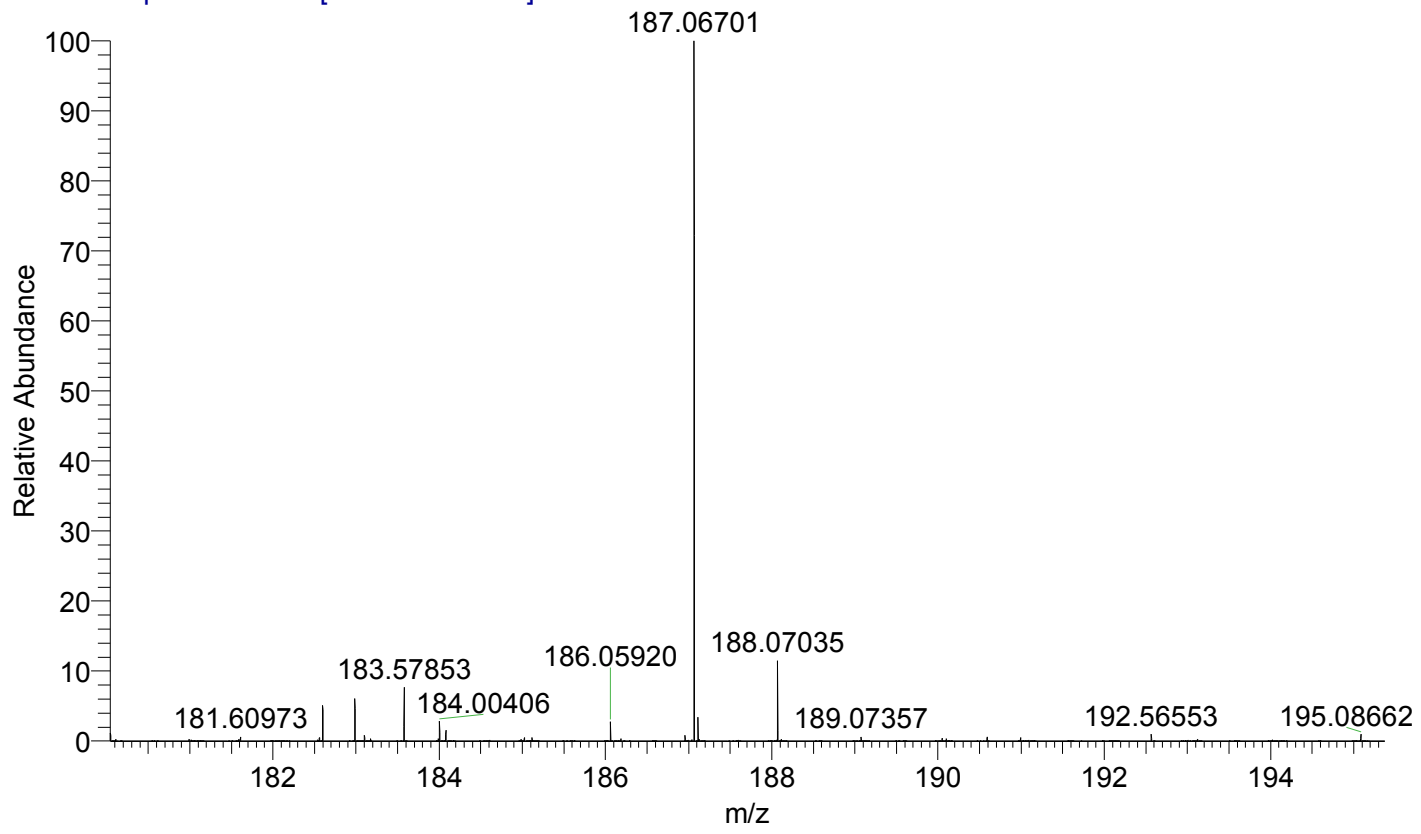


Figure SI-9. HRMS of **1f**

c:\xcalibur\data\xliu\72676lcmsp 10/8/2014 11:39:23 PM KULJIRA-5CIV-LIU

72676lcmsp#5324-5370 RT: 28.57-28.81 AV: 47
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 202.60497-203.60497

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
203.03738	29854188.0	100.00	203.03705	0.33	$C_{11}H_8N_2Cl$

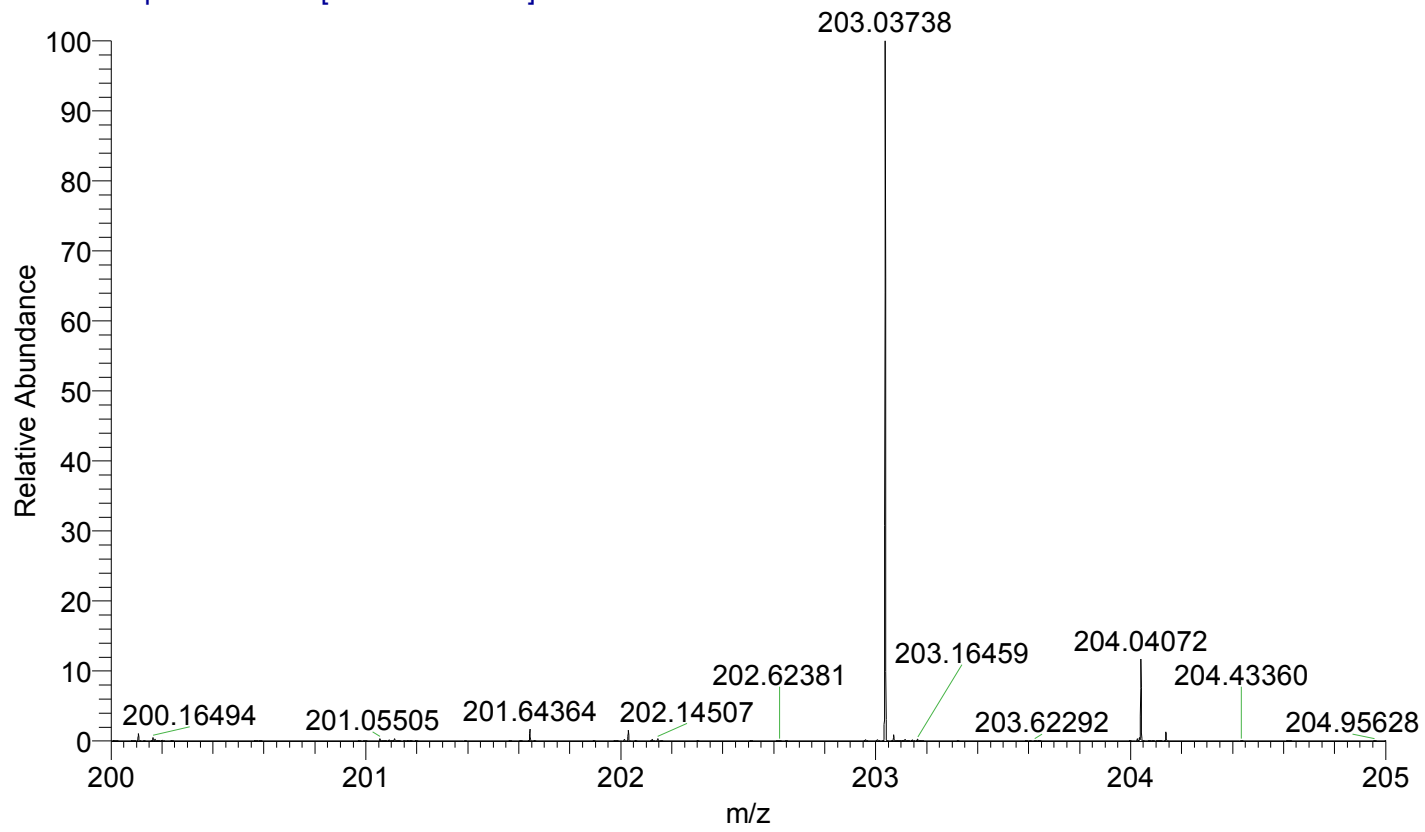


Figure SI-10. HRMS of **1g**

C:\Xcalibur\data\XLiu\72667LCMSP2 10/8/2014 6:08:55 PM

72667LCMSP2#5362-5397 RT: 29.27-29.46 AV: 36
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 246.90000-247.00000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
246.98691	2038514.5	100.00	246.98654	0.38	$C_{11}H_8N_2Br$

72667LCMSP2 #5363-5397 RT: 29.27-29.46 AV: 35 SB: 100 28.89-29.11 , 29.65-29.95 NL: 1.96E6
T: FTMS + p ESI Full ms [100.00-1000.00]

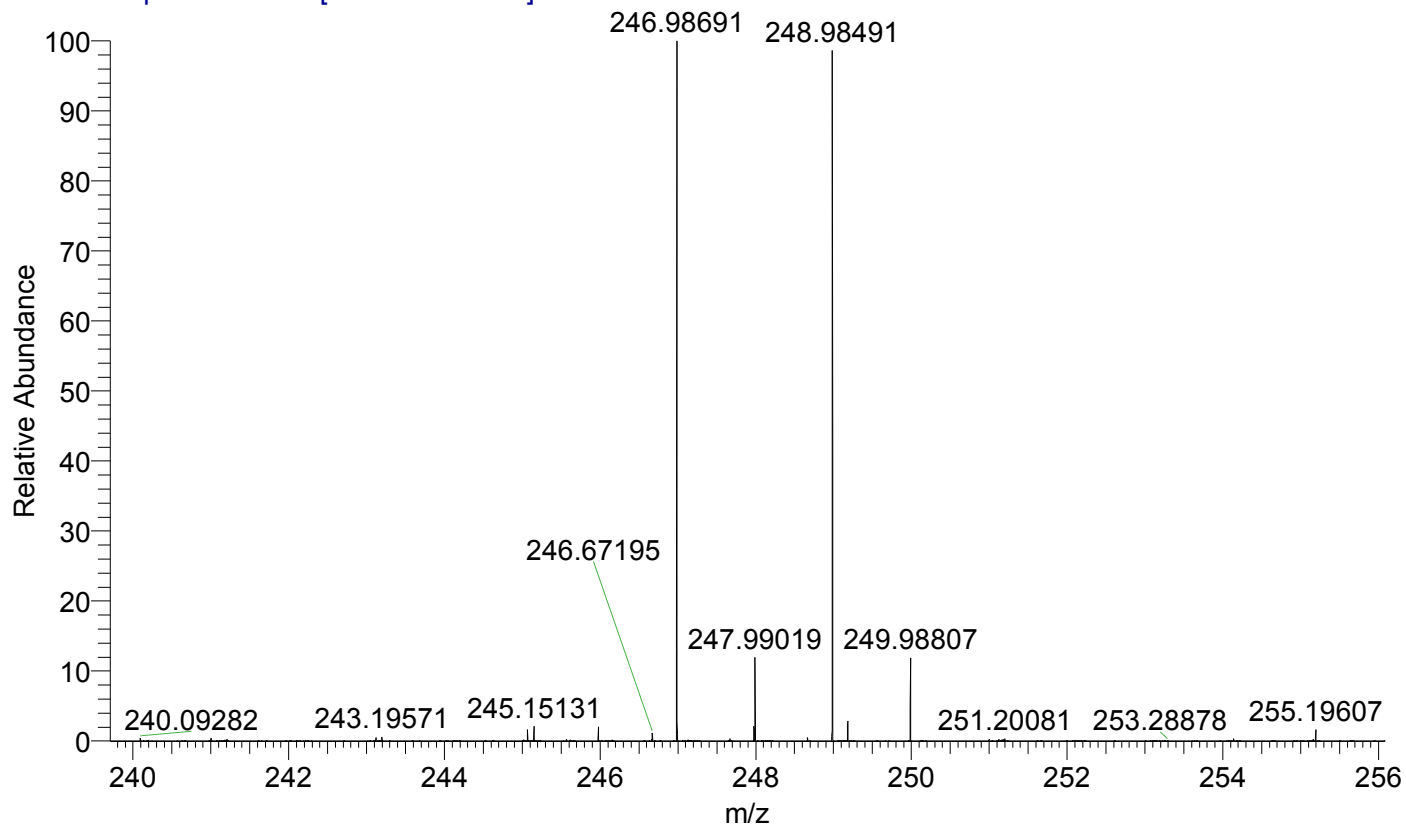


Figure SI-11. HRMS of **1h**

C:\Xcalibur\data\XLiu\72669LCMSP2 10/8/2014 7:31:32 PM 7CL

72669LCMSP2#5197-5284 RT: 28.37-28.84 AV: 88
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 203.00000-204.00000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
203.03760	1885703.3	100.00	203.03707	0.52	$H_7O_7N_6$

72669LCMSP2 #5197-5284 RT: 28.37-28.84 AV: 88 SB: 858 25.99-28.07 , 29.25-31.79 NL: 1.83E6
T: FTMS + p ESI Full ms [100.00-1000.00]

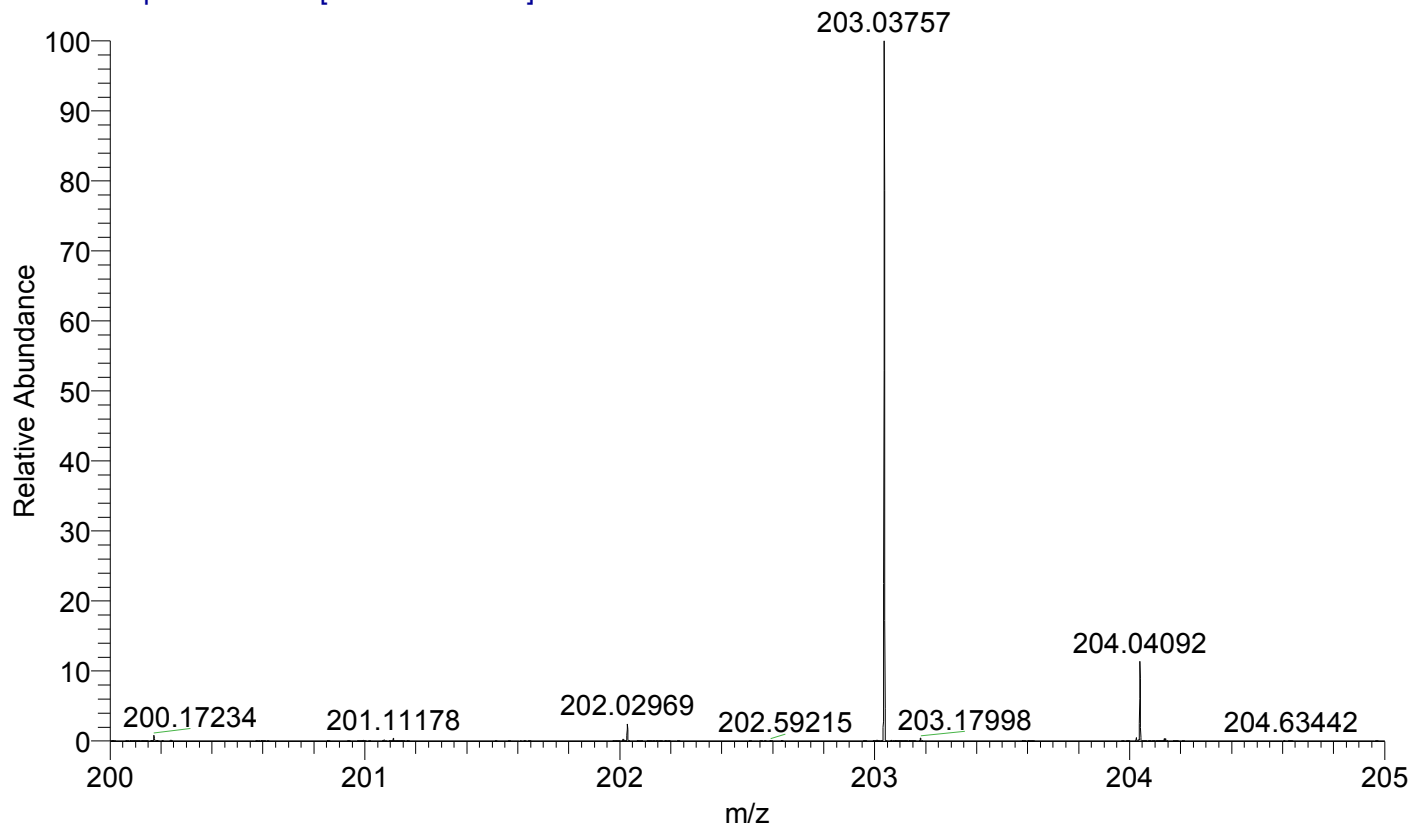


Figure SI-12. HRMS of **1i**

C:\Xcalibur\data\XLiu\72728LCMSP 10/15/2014 12:01:38 PM Kuljira-7Br-Liu

72728LCMSP#5302-5408 RT: 28.81-29.38 AV: 107
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 246.98500-246.98600

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
246.98598	28958042.0	100.00	246.98654	-0.56	$C_{11}H_8N_2Br$

72728LCMSP #5302-5408 RT: 28.81-29.38 AV: 107 NL: 2.85E7
T: FTMS + p ESI Full ms [100.00-1000.00]

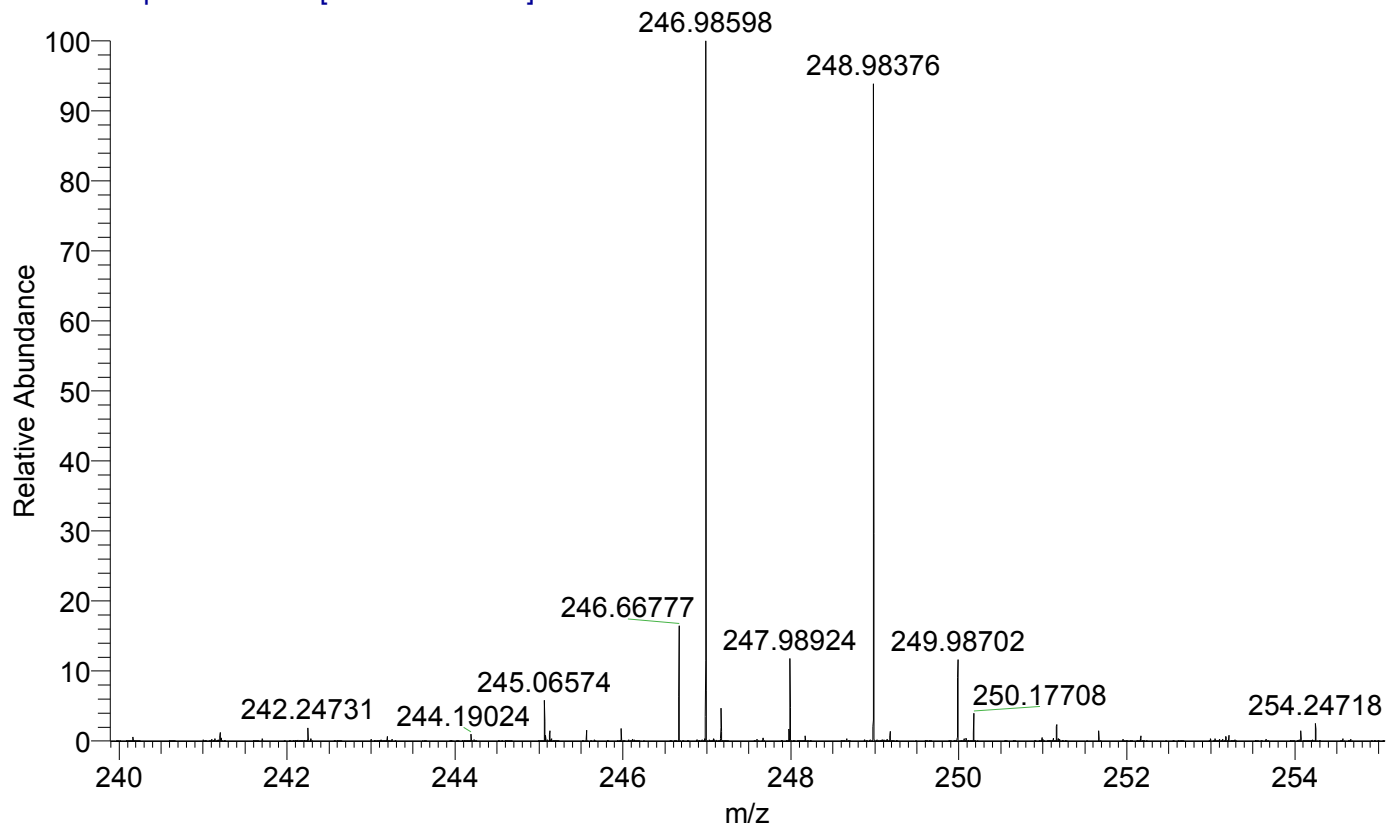


Figure SI-13. HRMS of **1j**

c:\xcalibur\data\xliu\72671lcmsp2 10/8/2014 8:54:08 PM 7I

72671lcmsp2#2252-2284 RT: 12.41-12.58 AV: 33
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 294.00000-295.00000

m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	Composition
294.97274	1681964.4	100.00	294.97267	0.08	$C_{11}H_8N_2I$

72671lcmsp2 #2252-2284 RT: 12.41-12.58 AV: 33 SB: 859 25.99-28.07 , 29.24-31.80 NL: 1.65E6
T: FTMS + p ESI Full ms [100.00-1000.00]

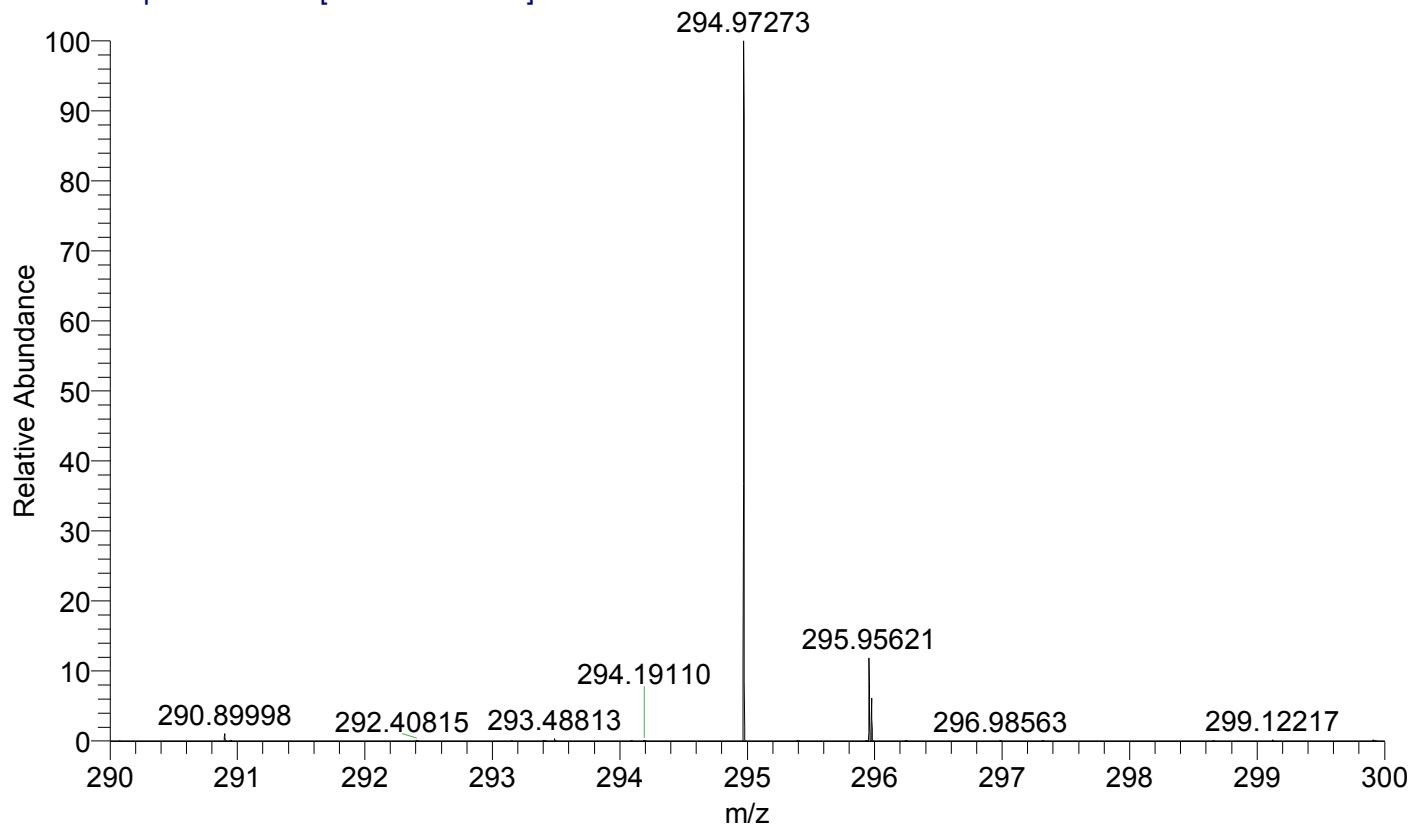
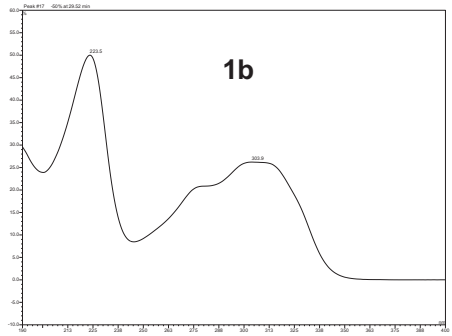
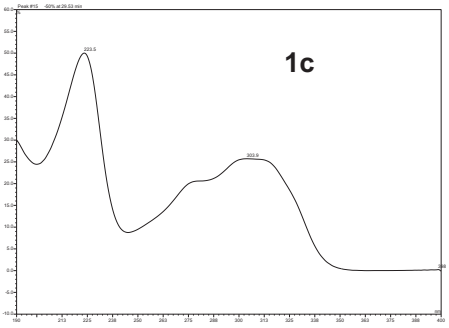
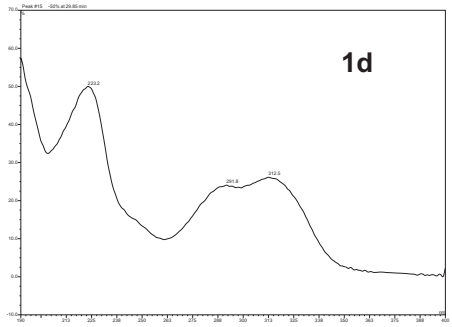
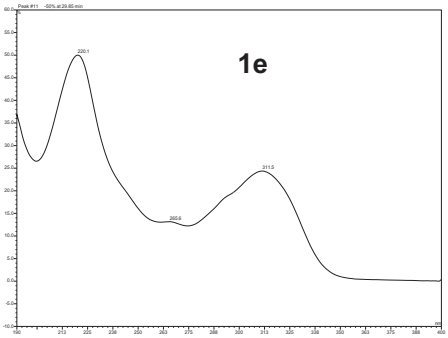


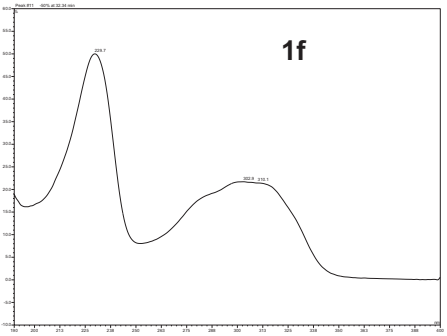
Figure SI-14. UV absorption spectra of **1b**–**1j**

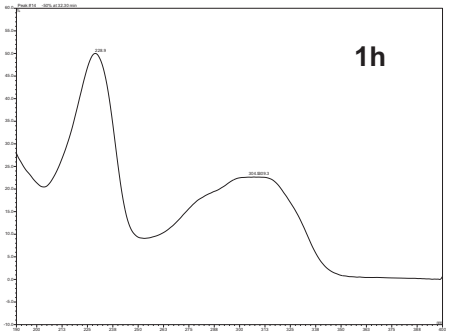


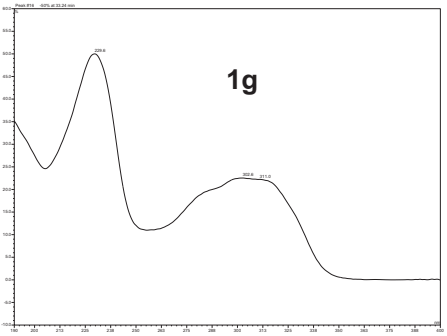


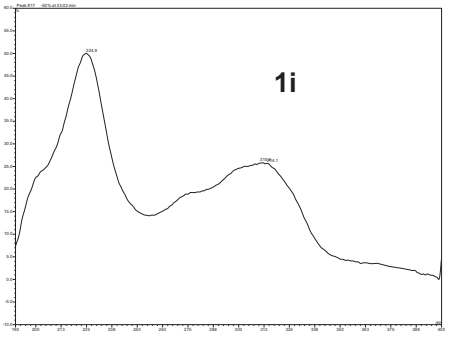


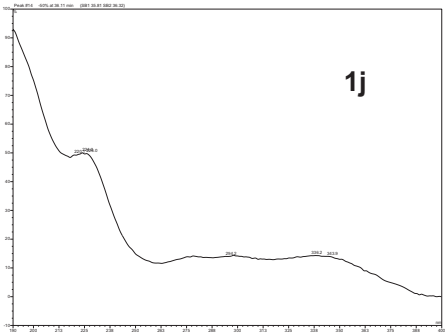












References:

1. D. R. Smith, T. Willemse, D. S. Gkotsi, W. Schepens, B. U. Maes, S. Ballet and R. J. Goss, *Org. Lett.*, 2014, **16**, 2622-2625.
2. R. M. Shanks, D. E. Kadouri, D. P. MacEachran and G. A. O'Toole, *Plasmid*, 2009, **62**, 88-97.
3. I. Hoppe and U. Schöllkopf, *Liebigs Ann. Chem.*, 1984, **1984**, 600-607.
4. M. L. Hillwig, Q. Zhu and X. Liu, *ACS Chem. Biol.*, 2014, **9**, 372-377.