SUPPLEMENTARY INFORMATION

From tryptophan to toxin: Nature's convergent biosynthetic strategy to aetokthonotoxin

Sanjoy Adak[1#], April L. Lukowski[1#], Rebecca J. B. Schäfer[1], Bradley S. Moore[1,2*]

[1]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093, United States

[2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, California, 92093, United States

[#]These authors contributed equally to this work

*To whom correspondence should be addressed: bsmoore@ucsd.edu

Table of contents

1. Plasmid construction and cloning

General considerations. *Escherichia coli* cloning strain DH10B was used for DNA propagation. Individual pET-28a(+) plasmids containing *E. coli* codon-optimized *aetA-F* genes between NdeI and XhoI restriction sites were synthesized and subcloned by Twist Bioscience. All PCR and incubation steps were performed in a BioRad MyCycler Thermal Cycler. PCR-grade water was purchased from Teknova.

Codon-optimized DNA sequences

aetA

ATGTATGATGTTGTCATCTGCGGCGGCGGACTCGCCGGGCTGTCACTGGCGCGTCAGCTGAAAATGACCAT
GCCGGAGCTCAGCGTTCTTATTCTGGAAAAAACATCCGGCGAGGTCCCGGTTTCCCCTTATAAGGTGGGAG
AATCCCTGGATGAGACAAGTAACTTCTATCTGTCCAACATTCTGCAGCTGGAAAGCTACTTGCAAGAACAT
CAGTTACATAAACTGGGGCTCCGCTGGTTCTTTGGTGGCGGCAATCTGCCGCTGGAAACACGTCCGGAGAT
CGGTCCGCCTGAGTTCCCGAGCATTCGCGCTCGCCATGTAGACCGCAAAGCTCTTGAGAATCACCTTGTCC
AGCTGAACAAAGAATCGGGGGTGGCGCACTTTACGGGTGTTCTGGTCAAGGATATCTTGCTTAGTGAAGAT
GATAGCCCGCATGAAGTTTTATACACTCAGCAGGGCGGCAATCACTTCACGGTGAAATGCCGTTGGCTTGT
GGATGCCACAGGCCGCCGCCGTTACTTGCAGTCGAAGTTAAACTTGACGCAGCCGAGTGGCCACCGTGCGA
ATGCCGTGTGGTGGCAACTTGAAGGTAAGCACAGCTTAGATAAAATGGCCGACGACGCGGACTGGCTGAGC
CGCAATAACCATCAACGTTGGTTTTCAACTAATCATTTAATGGGTGACGGCTACTGGGTATGGGTCATTCC
GCTCATTACCGGCAATACTTCTATCGGCATCGTAGCCGCGGAAGAAATGCATCCCCTGAGCAAACATAGCA
ATTATCAGTCCGCAATGAAATGGCTGGAAGAACATGAACCCCATTTCGCCAAGTTTATCCAGACCAGCACG
GCGCAAAATTTTTTGGCGGTCAAAGACTTCTGTTACCATTCCCAGCAAGTATTCAGTGAGAAACGCTGGAG
CTGCGTGGGTGATAGCGCCATCTTCACCGATCCATTCTATAGCCCAGGCGTTGTAATGCTGGCGTACAGCA
ACTGCCTGACTGTAAAGATGATTGAGCTGGATCATCAGAACCAATTGACCGCGGAAATTGTTCAGCAATTT
AACAACTTGCTGCTGCGCGACATTGGACACAATTACCTGCAAATTTACCAGGATAATTATCCGGTGTTTGG
CTCTTTTAACGTTATGAGCATCAAACTGTTCTGGGATGCGGTGTATATTTGGTTTTTCCCAGGTGCGCTGT
TTTTTCAGCAGCATTTCACCAATCTCGAAGTGTTAAGCAATTTTTCTCGTATTGCACAGCGCTACTTAGCC
TTGAATCAGCGGGTTCAGAAACTGTTTCAAGATTGGGGCCAAGCAATTAAAAGTCAAAATTTCCCGTGTTT
TCTGGCGTACGATATTTCTTACAAACCGACCTGCCTGGGCAAGCCCCGCGCAAACGGCAAACGTTTCTTTT
TCGATATTTTCAAAGAAGCCCACTTCAGCCTTGTCGAAAAGAAAGACACCCAACAGTATCTGGCGGATCTC
GAAGTGCAAATGCAAGAACTGGAAATTTGGGCGCAGGTGATCTTTCGTCACGCTCTGAACCAGGTGATGCC
GGAACGTCTGAGTCAGTTCCCTGAACCCTTTTGGGTGAACGTCTGGGCCATCGGCCTGAACCCGGAGAAAT
GGGAAGAGGATCGCCTGTTTCAGCCAGAAACCGAACCGCAGGTTGATCTTCGTAACTACATTCTGGCCGAA
TTACCTGATGACCACTATCCTCCTGGCGCGGAACCAGTAACGACATCTAGCGTTGATAACCCGGCCTACGT
CGCAGCCGTAAACTAA

aetB

ATGACGCTTGAACTTCAAAGCTCGATTCAATCCAAGAGTATCATCACGCCGCCTGGCCCGCGCGGCTTGCC
GCTGGTTGGTAGCCTTGCACCATTACTGGGTAAGCACTTACACGTGACCCTGAACCAGCTGGCTAAAGAGT
ACGGCAACATTTTCCAGATCCGTATGGGGATTCGCACGGTCGTGGCACTGAACGGCATCAAAACCATTACC
GAAGCCCTCATTAAACAACAGAAAAATTTCATTAACCGCGCCGATTTCGAATTTTTTAGCATTCCCGCGGA
CGGCGGGCTGTTCGAAGCCAAGTCGGGTCAGGCCTGGGAGAAACATCGCGGTATTGCAGCGAAAGCTATGC
GGACGTTCGTGATTGATAAGTCCGACCGCGTCGAACATCACATTATGGAGGAAGCAGCGAAACTGGCGGAC
TCTTTCCTTAATTATGATGGCCAGCCATTTGATCCGGACATTCCTTCGAGCCTGGCGGTGGGTTCGATCAT
CCACCGGGTGTTATTTGGCACTCGTGGCTCATCAGAAGATCCAGATTTTGTAGCCACGGCTCGCATTCTGC
AGTTATTTCCGCCTGTCACCCAGCGTATTCTCATTAGCGAGTTCCTGCCAAAGATCTTACGTCCGGTGTTC
TTTATTTTTAACCAGAAATCCCTGCAGGACTTTCAGGATGTTCCAGCAACCGTGGATAAGCTGGTCTCTAA
AAACATTGAACAGCATCGCAACAGTTTCGACCCAGATAACATTCGTGACATGACGGATGCGTTTTTGAAAG
CCGCGAATGATCTGTCGGAAAGCGACAAAAAAAATTTTAACTTGTCCGAAAATCACATTCTGCGTGGGACG
CTGATGCAACTGACGGGCGCTGGCTTAGAGCCTGTGCTTATCCAACTGATCTGGGCACTTTTATACATGAT
TGCGTATCCCGATATTCAGGCCAAAATTCAACAAGAATTAGACGAGGTTGTCGGGCTGGGCCAGCAGCCTT
GTTTAAACCACCGGGGCAAACTCCCATTTACCGAAGCCTGTATCAACGAAATCCTGCGCCACTCAAGCATC

AGCCCGATTCCTTTTCTGTACGCAAGCACTACCGACACTACGATCAACGGGTATAGCGTGCATCAGGGCAC
TCCCGTGATCATTAACCTTTATAGTCTGACCCGCGATGAACGTTATTGGCAGAACCCGGACAAATTTGATC
CGGATCGTTTTCTCGATGAAAATTGCAAAATCCGTTCCGATCTGCTGGATAAGTTCTACCCGTTCGGAATT
GGTCAACGGCGTTGCATGGGCGAATATCTGGCGCGCCTGGAAATCTTCATCTTTTTTACGACGCTGATGCA
TAAATGTAAATTGGAGAAAGTACCGCGTGAAAAACTTAGCCTTGAACCAAAACCCGGGTTCTTTATGCGTC
CGCAGAAATATAATATCATCGCCAAATCCCGGTTCTAA

aetC

ATGTCGAAATCGGAGAACCTCAAAGATGTTGAAATTGAGGTGGAAAAGCTGGAGGGTCGGCATCGCCGCAT
TTTTGCCCAGATTCAGATCAGCTACCCTCTCGAACAGGTATGGCAGGTCATCACGGATTATGAAGCCTTTG
CGAAATTCATGCCGAACCTCAAAGAATGTCGGCGTCTGGAACACCCGACCGGAGGCATCCGCTTAGAACAG
GTTCGTACCCTGTCTTTTTTGGGTTTAAACTTTAGCGGCCGCAGTGTGTTCGACATCGCCGAAGAATTCCC
GGACAAGATTCATTATCAGCTGGTGGAAGGTGATTTAAAAGCCTTCAGCGGTGACTGGCGCCTGGCTCCAG
CGAATCTGGGCGAGAAAGCGGGTGTGGTGTTAACGTATAATTTCTCGATTCTGCCAAACCCTCTGCTGCCG
ATTGTCGTCGTTGAACGCGTTTTTTCGCACGACGTTCCGGTCTCTCTGCTTGCTATCCGCCAGCGCGTTGA
GGATCTTTTTGGTTCAGGTAGTAAGTAA

aetD

ATGAAAGCTATCCTGCAACTGATCCTGGAAAAACGCCAAGAATTCGAAAAACTTCCCTGTTTCGAATTCGT
GCGTGACGAAACAATTTCCCCTGAAGAACGCCTGATTCTGTACCCGTGTATCGCTGCGTTTGCCTTAAACT
TCCGCGATCTGAACCGGTACGATTATCGTGACGATAACTCGTCTGATTATTATCAGAAAATCATTAACATC
CATACTCAGGAGGACGCGAAACACTGGGAATGGTTCCTGAATGATCTGGAACTGCTGGGGTTCGATAAAAC
CATGCGCTTTTCGGAGGCCCTGCGCTTTGTCTGGTCTGATGATTTACTGCACACTCGTCGGCTCTGCCACA
ACATTGCCGTGCTGAGTCACGATTTAGAGCCAGTTATGAAGATGGTCGTCATCGAGGCGATGGAAACTGCC
GGGTTAGTCATCTTCCATGCCCTGGCCAAACCGGGGGAGTCAATCGCCAAGGCGACGCGCCGCAAATATCT
GTACGTCGCAGACAGCCATGTCGAAGTGGAAACGGGGCACGCCGTGGGAACTGAAAACATTATCACCATCC
TGGAACAGACTCAGTTGAGCTCTGAGCAGGAGGAAAAAGCTAAAGAGATTGTGAATAAAGTGTTTCAATGG
AGCACAAACTTGATTGGCGAGTTCGAACGGTACGTGAAAGCGCATCGTTCTGAAAAAGCGCAACCAACAGC
GGCCTATTAA

aetE

ATGACACACCATCCGCCGGAGCCGTTTCGCATCAAAATGGTGGAACCTATTAAATTGCTCGACCGCGATGC
GCGCAAAGCAGCGATTCGCCGCTCCGGTTACTCTCTCTTTGGGCTGCGCTCGGAAGAAATCTTTATTGATT
TTTTTACCGATAGCGGCACCACGGCTATGTCACAAACTCAATGGAGCGCGATGTTCGAAGGTGATGAGGCC
TATGCCGGCGGCTCTTCTTATTTTCGCCTGGCCGAGGTGATCCATGACATTTTTGGTTTTGATTATTTTCT
GCCGACGCACCAGGGCCGCGCTGCGGAAAATATTCTGTCCACATGTATGGTTAAGCCAGATCAGTATGTTC
CCTCCAATATGCACTTTGATACTACGTACGCGAACATCCGCGCCCGTGGTGGCCGCCCGATTAATTTAGTC
ATTGACGAAGCTCATCATCCCGGTAGTTATCACCTGTTCAAGGGGAACATGAATATCCAGAAATTACGCGC
TTTTATTGAAGAGGTCGGCCCTAATCAGATTCCGTTCGGTATGATTACTGTCACTAACAACGCAGGTGGAG
GTCAACCCGTAAGCATGGAAAACCTGCGTGCTGTCAGCCAGACGTATAAGGAATTCGGTATTCCATTCTTC
ATTGATGCATGTCGTTTCGCGGAAAACGCGTACCTGATTAAATTACGCGAACCGGATTACGCTGATAAAAC
TCCCCTTGAAATTTCGCGCGAAATGTTCAGTCTGGCAGATGGTATCACTATGTCTGCTAAAAAAGATGGAA
TGGTCAACATTGGAGGGTTTATCGCGATGAACGATGAACTGCTGTTTGAACAGGCTCGCAACGAGCTGTTA
CTGCGCGAAGGTTTCCCGACCTATGGCGGCCTGGCGGGCCGGGATCTTGACGCCATGGCTGTGGGATTCCA
TGAAGTGCTCCAGGAAGACTATCTGGCCTATCGTCTCGCTCAAACCGCGTACCTGGGGGATCGGCTGCGTG
AATTAGAAATCCCGATCGTTGAACCACCGGGCGGTCATGCCATTTACATCGATGCGGGGCGTCTGTTGCCA
CACATCTCACAGAAAGAATTCCCAGCGCATGCGTTAACTGTCGAACTGTACCTGGAGGGCGGCATTCGTAC
AGTGGAGATCGGTTCACTGTCGTTCGCCTATCCGGACCCGAAAACTAACCAAATGGTATATCCGAACCTTG
AGCTGGTTCGTTTAGCACTTCCGCGCCGCGTCTATACCCAGAGCCACCTGGATTACGTCGCCGAAACCCTG
GGCAAGACGGTGACTCGTTGCGCCCAGATTCCGGGCTACCGTATTACCTATGCTCCGAAGGTACTGCGCCA
ATTCACCGCGCAACTGGAACCAATTAAACTGGACGCCCATCCATCTGCGACTCTGTAA

aetF

ATGTTGGAGGTATGCATTATCGGCTTTGGGTTCTCTGCGATTCCACTGGTTCGTGAGCTTGCTCGGACGCA
GACTGAATTCCAGATTATTAGCGCAGAGTCGGGGAGTGTTTGGGACCGCCTGAGCGAGTCCGGTCGTCTGG
ATTTTAGCCTGGTATCCTCATTTCAGACCTCTTTTTATAGCTTCGACCTCGTCCGTGACTACGAGAAAGAT
TATTATCCAACCGCAAAACAATTTTACGAAATGCATGAACGTTGGCGGTCTGTGTATGAAGAAAAAATCAT
CCGCGATTTTGTCACCAAAATCGAAAATTTCAAAGACTACAGCCTGATCTCCACCCGCAGCGGCAAGACAT
ACGAAGCCAAACATGTTGTGCTGGCGACCGGCTTTGATCGCCTGATGAATACCTTTTTGTCTAACTTTGAT
AACCATGTTAGCAATAAAACCTTTGTGTTCGACACCATGGGTGACAGTGCAAACTTGCTGATTGCGAAACT
TATTCCGAACAACAACAAAATCATCTTGCGCACCAATGGCTTTACCGCACTGGACCAGGAAGTGCAGGTGC
TGGGTAAACCATTTACCCTCGATCAGCTGGAAAGCCCAAACTTCCGCTACGTAAGTTCTGAGCTTTATGAT
CGCCTGATGATGTCGCCTGTATACCCCCGTACCGTCAACCCCGCCGTCTCATACAATCAGTTCCCACTGAT
TCGGCGCGACTTCAGTTGGGTAGATTCAAAAAGTTCCCCGCCCAATGGTCTGATCGCCATTAAGTATTGGC
CGATCGACCAGTATTATTACCACTTCAATGACGACCTTGAAAATTATATTAGCAAAGGCTACCTGCTTAAC
GATATCGCCATGTGGCTTCATACGGGCAAAGTCATCCTGGTCCCGTCAGATACTCCGATTAACTTTGATAA
AAAAACCATTACCTACGCGGGCATCGAACGCTCCTTTCATCAATATGTTAAAGGTGACGCTGAGCAGCCTC
GCCTGCCCACCATCTTAATCAATGGCGAGACCCCTTTCGAATACCTTTATCGTGATACTTTCATGGGTGTC
ATCCCCCAGCGCCTGAACAACATTTATTTCCTGGGATACACGCGGCCATTCACCGGTGGTCTGGCAAACAT
TACCGAAATGCAAAGCCTGTTCATCCATAAACTGATTACCCAGCCGCAATTCCACCAGAAAATCCACCAGA
ATCTCAGCAAACGTATCACTGCGTACAACCAACATTACTATGGTGCAGCGAAGCCGCGTAAACATGATCAT
ACGGTTCCCTTCGGCTTCTACACTGAAGATATCGCACGTCTGATTGGCATCCACTATCAGCCGAATGAATG
TCGTTCAGTACGGGATTTGTTGTTCTATTATGCGTTCCCCAACAACGCTTTCAAATATCGTCTGAAGGGGG
AATACGCCGTGGATGGCGTTGACGAGCTCATTCAGAAAGTCAACGACAAACACGATCATTACGCGCAAGTT
TTCGTACAGGCACTGAGCATTCGTAACATGAACAGTGACGAGGCTGCAGAATGGGACCACTCCGCCCGTCG
TTTTAGTTTTAATGATATGCGCCACAAAGAGGGATACCGTGCCTTTTTGGACACATATCTGAAGGCGTATC
GTCAGGTGGAAAACATTTCGGTGGATGATACGGTTGTGGATGAGGAATGGAATTTTATGGTGAAGGAAGCG
TGCCAGGTGCGCGATAAAGTAGCTCCCAACATCGAAGAAAAGACTCATTATAGTAAGGATGAAGACGTAAA
CAAAGGGATTCGTCTGATTCTGTCTATCCTTGACTCAGATATCTCCTCTCTGCCGGATTCTAATGGCTCTC
GCGGTTCTGGTAATCTTAAAGAGGGTGACCGTCTGTGTAAATTTGAAGCGCAAAGTATTGAGTTCATTCGT
CGCCTCCTGCAGCCGAAAAATTATGAACTGCTGTTTATCCGTGAATCAACTGTTAGCCCCGGTTCACATCG
GCATGGGGAAACTGCTTAA

Construction of pET28-MBP-TEV-*aetB* using Gibson assembly. The *aetB* gene was amplified from pET-28a(+)-*aetB* using Phusion HF polymerase, 5x HF reaction buffer, and dNTPs (Thermo Scientific) and primers denoted in Table S1 (Integrated DNA Technologies) following the reaction conditions outlined in Table S2. pET28-MBP-TEV (Addgene) was digested with BamHI and XhoI restriction enzymes (New England Biosciences, NEB) at 37 °C for 2 h. Both the *aetB* PCR product and the digested pET28-MBP-TEV plasmid were purified by gel extraction using the QIAquick Gel Extraction Kit (Qiagen) and EconoSpin spin columns for DNA (Epoch Life Science). *aetB* was inserted into the digested pET-MBP-TEV by incubating 3 µL digested plasmid with 1 µL purified amplified *aetB* and 4 µL 2x Gibson Assembly Mix (NEB) at 50 °C for 1 h. The entire reaction mixture was transformed into 50 µL chemically competent *E. coli* DH10B using standard heat shock procedures and spread onto an LB agar plate supplemented with 50 µg/mL kanamycin (Thermo Scientific). Successful insertion of *aetB* was verified by sequencing (GENEWIZ).

Table S1. PCR primers used in this study.

Primer name	Primer DNA sequence (Gibson overhangs are in lowercase)
AetB_pET28a-MBP-TEV_GA_fwd	5' – gaaaatctatacttccaaggatccATGACGCTTGAACTTCAAAGCTCGATTC – 3'
AetB_pET28a-MBP-TEV_GA_rev	5' – ggtggtggtggtggtgctcgagTTAGAACCGGGATTTGGCGATGATATTATATTTC – 3'

Table S2. PCR reaction conditions for the amplification of *aetB* with Gibson overhangs.

Component	[stock]	[reaction]	Volume added (50 µL total)
pET-28a(+)-*aetB*	100 ng/µL	2 ng/µL	1 µL
Forward Primer	10 µM	0.2 µM	1 µL
Reverse Primer	10 µM	0.2 µM	1 µL
dNTPs	10 mM	0.2 mM	1 µL
Buffer HF	5x	1x	10 µL
Phusion polymerase	2 U/µL	1 U	0.5 µL
PCR water	-	-	35.5 µL

PCR program

98 °C	30 s	*1 cycle*
98 °C	10 s	*35 cycles*
66 °C	30 s	
72 °C	45 s	
72 °C	10 min	*1 cycle*
4 °C	*hold*	

2. Protein expression and purification

General considerations. All growth media and purification buffers were prepared with water filtered with a Milli-Q water purification system equipped with a Q-POD dispenser. Fast protein liquid chromatography (FPLC) was conducted on a Cytiva ÄKTA Pure 25 L1 system fitted with an F9-C fraction collector and S9 sample pump and controlled by Unicorn v7 software. Buffers were vacuum filtered with 0.2 µm nylon filters (MilliporeSigma). 5 mL HisTrap FF columns (Cytiva) were used for affinity chromatography. Large cultures (>500 mL) were centrifuged with a Beckman Coulter Avanti JXN-26 centrifuge. Cell lysis was performed by sonication with a QSonica Q500 ultrasonic processor. Protein samples were analyzed with Tris SDS-PAGE gels (BioRad supplies) and Quick Coomassie Stain (Anatrace). All purification steps were performed at 4 °C unless otherwise indicated.

Expression protocol for AetA, AetC, AetF, and SsuE.[1] pET-28a(+) plasmids containing *aetA, aetC, aetF*, or *ssuE* genes were transformed into chemically competent *E. coli* BL21(DE3) cells using standard heat shock procedures and plated onto LB agar supplemented with 50 µg/mL kanamycin. Single colonies were used to inoculate 10 mL LB media containing 50 µg/mL kanamycin and these starter cultures were incubated at 37 °C, 200 rpm overnight. The overnight cultures were used to inoculate 1 L TB media supplemented with 0.4% glycerol and 50 µg/mL kanamycin. Cultures were incubated at 37 °C, 180 rpm until the optical density at 600 nm (OD_{600}) reached 0.8-1. The cultures were chilled on ice for approx. 20 min as the incubator cooled to 18 °C, at which point the expression of each protein was induced by the addition of 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG, 100 µL of 1 M stock). The induced 1 L cultures were incubated at 18 °C, 180 rpm overnight (~18 h) prior to harvesting by centrifugation (5000 x *g*, 20 min). Average wet mass of cells varied from 13.6 g (AetC) to 23.9 g (AetF).

Purification protocol for AetA, AetC, AetF, and SsuE. Pellet from 1 L TB culture was resuspended in ~80 mL cold lysis buffer (50 mM Tris-HCl pH 7.4, 300 mM NaCl, 10% glycerol, 10 mM imidazole) and sonicated in a glass beaker on ice, 5 min total "on" time, 3 s on, 7 s off. Lysed cells were centrifuged at 17000 rpm for 30 min at 4 °C. Supernatant was filtered through 0.45 µm syringe filters (CellTreat) prior to loading onto the FPLC for purification on a 5 mL HisTrap FF column. Buffer A was the lysis buffer as previously indicated, and Buffer B was 50 mM Tris-HCl pH 7.4, 300 mM NaCl, 10% glycerol, 250 mM imidazole. The column was equilibrated with 5 column volumes (CV) of 0% Buffer B at 5 mL/min, and the sample was loaded at 2 mL/min via the sample pump followed by a 10 mL wash with Buffer A through the sample pump. The loaded column was washed with 8% Buffer B for 10 CV at 2.5 mL/min and protein was eluted in a linear gradient from 8-100% Buffer B over 15 CV at 2 mL/min. Fractions containing AetA or AetF were analyzed on a 10% Tris SDS-PAGE gel, and AetC was analyzed with a 12.5% gel. Fractions containing the desired protein were concentrated to 2.5 mL using Amicon centrifugal molecular weight cutoff filters (50 kDa for AetA and AetF, 10 kDa for AetC). Concentrated proteins were desalted using PD-10 desalting columns (Cytiva) equilibrated with storage buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, 10% glycerol) following package directions. Yields were determined by standard Bradford assay (~ 5 mg AetA, ~ 2 mg AetC, and ~ 47 mg AetF) and purity assessed by Tris SDS-PAGE (see Figure S1). AetF was bright yellow in color, while AetA and AetC were colorless. SsuE was purified in a similar manner as described above, yielding ~1.5 mg of colorless protein.

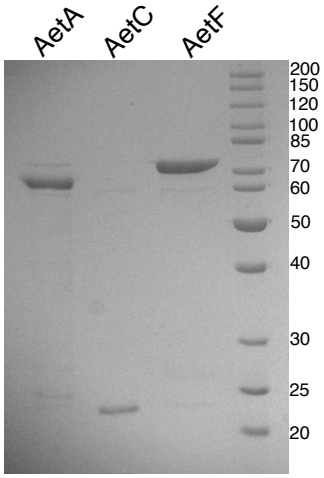


Figure S1. 10% Tris SDS-PAGE gel with samples of purified His-tagged AetA, AetC, and AetF. The purity shown represents the purity of the samples used in enzymatic assays. Sizes estimated by ProtParam are as follows for each His-tagged protein: AetA – 71059.64 Da; AetC – 22151.36 Da; and AetF – 80055.34 Da.

Flavin quantification in AetF. The absorbance spectrum of 1 mL of 10 μM AetF was measured from 300 nm to 700 nm in 2 nm increments in a polystyrene cuvette (BrandTech) using a Cary 60 UV-vis spectrophotometer (Agilent Technologies). The sample was transferred to a microcentrifuge tube, and 20 μL of 10% sodium dodecyl sulfate (SDS) was added. The sample was mixed, incubated at room temperature for 10 min, and the UV absorbance spectrum was measured again from 300 nm to 700 nm in 2 nm increments. The percentage flavin incorporated was determined using the extinction coefficient of free FAD at 450 nm (Figure S2, red trace, e_{450} = 11300 M^{-1} cm^{-1}). The percent FAD incorporation was determined to be approximately 86%, and the extinction coefficient of AetF with FAD bound was found to be approximately 11870 M^{-1} cm^{-1}.

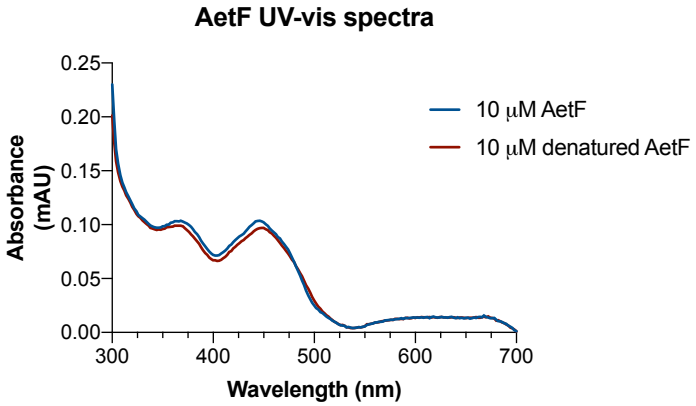


Figure S2. UV-vis absorbance spectra of 10 μM purified AetF (blue) and 10 μM purified and denatured AetF (red).

Expression and purification protocol for AetD

The AetD gene containing overexpression plasmid (pET28a) was transformed into *E. coli* BL21 (DE3). A single colony was used to inoculate 10 mL of LB medium containing 40 µg/mL of kanamycin, and the starter culture was grown overnight at 37 °C with shaking (220 rpm). 1.0 liter of TB medium containing 40 µg/mL of kanamycin was inoculated with this starter culture. The cells were grown at 37 °C with shaking (220 rpm) until the culture reached an OD600 of 0.6. The culture was then incubated at 4°C for ~30 min without shaking. The culture was induced by adding IPTG to a final concentration of 0.5 mM, the temperature was lowered to 18 °C, and the cells were grown with shaking (200 rpm) for a further 20 hours. The cells were then harvested by centrifugation at 10,000 x *g* for 10 min at 4 °C and stored at -80 ˚C.

The cell pellet was re-suspended in 30 mL of lysis buffer (50 mM KH_2PO4, 150 mM NaCl, 10 mM imidazole, pH 8.0). Lysozyme (15 mg) was added, and the cells were lysed by sonication on ice (six cycles of 30 sec durations during which 1.0 sec sonicator pulses at output level 0.6 were followed by 1.0 sec pauses). The resulting suspension was centrifuged (40,000 x *g*, 30 min), and the clarified supernatant was loaded onto a 5 mL Ni-NTA-affinity column pre-equilibrated with lysis buffer kept at 4 °C. The Ni-NTA-affinity column was then washed with 50 mL wash buffer (50 mM KH_2PO_4, 150 mM NaCl, 20 mM imidazole, pH 8.0). The protein was eluted from the column with elution buffer (50 mM KH_2PO_4, 150 mM NaCl, 200 mM imidazole, pH 8.0). The fractions containing protein were collected and concentrated using YM-10 Amicon ultracentrifugal filters at 5000 x *g* to a final volume of 2.5 mL. The concentrated sample was buffer exchanged, using PD-10 desalting column (Cytiva), into 50 mM Tris-HCl pH 8, 100 mM NaCl, 10% glycerol. Yields were determined by standard Bradford assay (a typical yield was 70 mg/liter), and purity assessed by Tris SDS-PAGE (see Figure S3).

Expression and purification protocol for AetE

pET28a overexpression plasmid containing the AetE encoding gene and pGro7 plasmid containing GroEL-GroES encoding genes were transformed into *E. coli* BL21 (DE3). A starter culture was grown overnight in 10 mL of LB medium containing 40 µg/mL of kanamycin and 30 µg/mL of chloramphenicol at 37 °C. 1.0 liter of TB medium containing 40 µg/ml of kanamycin, 30 µg/mL of chloramphenicol and 2 mg/mL L-arabinose was inoculated with this starter culture. The cells were grown at 37 °C with shaking (220 rpm) until the culture reached an OD600 of 0.6. The culture was then incubated at 4°C for ~30 min without shaking. Then the culture was induced by adding IPTG to a final concentration of 0.5 mM. The rest of the overexpression and purification protocol were the same as discussed for AetD. A typical yield was 35 mg/mL.

Expression and purification protocol for AetB

For the overexpression of MBP-AetB fusion protein, pET28-MBP-TEV-*aetB* plasmid was transformed into *E. coli* BL21 (DE3). A starter culture was grown overnight in 10 mL of LB medium containing 40 µg/mL of kanamycin at 37 °C. 1.0 liter of TB medium containing 40 µg/mL of kanamycin was inoculated with this starter culture. The cells were grown at 37 °C with shaking (220 rpm) until the culture reached an OD600 of 0.6. The culture was then incubated at 4 °C for ~30 min without shaking. Then the culture was induced by adding 0.5 mM IPTG and 1 mM 5-aminolevulinic acid. The rest of the overexpression and purification protocol were the same as discussed for AetD. A typical yield was 20 mg/mL.

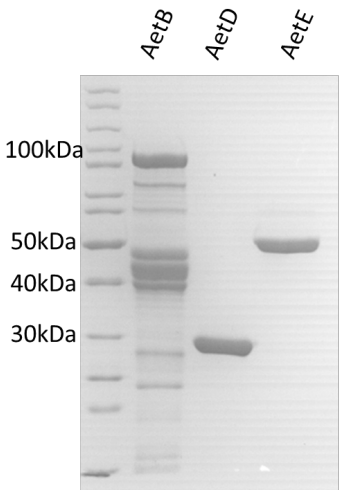


Figure S3. 10% Tris SDS-PAGE gel with samples of purified AetB, AetD, and AetE. The purity shown represents the purity of the samples used in enzymatic assays. Sizes estimated by ProtParam are as follows: His-MBP-AetB – 99290 Da; His-AetD – 29998 Da; and His-AetE – 54602 Da. The two intense lower bands in the AetB sample likely correspond to His-MBP (44718 Da) and cleaved AetB (55096 Da).

3. Enzyme assays and LC-MS analysis

General considerations. HPLC separations were performed using an Agilent Technologies 1200 series HPLC equipped with a degasser, binary pump, autosampler, and diode array detector. Analytical-scale separations were performed with an Agilent Eclipse XDB-C18 5 µm, 4.6 x 150 mm column fitted with a guard cartridge, and semi-preparative separations were performed with a Kinetex 5 µm C18 100 Å, 250 x 10.0 mm column. Solvents used were HPLC-grade. Data were collected and analyzed using ChemStation for LC 3D systems version B.02.01. LC-MS analysis was performed with an Agilent Technologies 1260 Infinity series HPLC equipped with a degasser, binary pump, autosampler, and diode array detector coupled to an Agilent Technologies 6530 Accurate-Mass Q-TOF LC/MS. Separations were performed with a Kinetex 5 µm C18 100 Å, 150 x 4.6 mm column. Solvents used were mass spectrometry-grade. Data was collected and analyzed using MassHunter Workstation Software version B.05.01. For both HPLC and LC-MS systems, solvent A was water + 0.1% formic acid and solvent B was acetonitrile + 0.1% formic acid.

HPLC Method 1 (semi-preparative scale AetF reaction): 2 mL min^{-1} flow rate; 0-5 min 5% B, 5-15 min 5-95% B, 15-18 min 95-100% B, 18-20 min 100-5% B, 20-25 min 5% B.
HPLC Method 2 (analytical-scale AetE reactions): 0.75 mL min^{-1} flow rate; 0-5 min 5% B, 5-15 min 5-95% B, 15-18 min 95-100% B, 18-20 min 100-5% B, 20-25 min 5% B.

LC-MS Method 1 (analytical-scale AetA reactions and flavin analysis): 0.75 mL min^{-1} flow rate; 0-5 min 5% B, 5-15 min 5-95% B, 15-18 min 95-100% B, 18-20 min 100-5% B, 20-25 min 5% B. Dual ESI ion source, negative polarity.
LC-MS Method 2 (analytical-scale AetB and AetD reactions): 0.75 mL min^{-1} flow rate; 0-4 min 10% B, 4-11 min 10-95% B, 11-21 min 95% B, 21-23 min 95%-10% B, 23-27 min 10% B. Dual ESI ion source, negative polarity.

Analytical-scale AetF reactions. 50 µL reactions consisting of 5 µM AetF, 100 µM FAD, 500 µM L-tryptophan (**4**) or 5-bromo-L-tryptophan (**5**), 1 mM NADPH, 10 mM KBr, and 50 mM Tris-HCl pH 8 buffer were prepared in 1.5 mL centrifuge tubes and incubated at 30 °C overnight. Reactions were quenched by the addition of 100 µL methanol and centrifuged at 14,000 rpm for 10 min prior to filtering with 0.2 µM centrifugal filters. 25 µL of the filtered sample was injected on a Kinetex 5 µm C18 (150 x 4.6 mm) column using LC-MS Method 1.

Preparative-scale AetF reaction and purification. A 24.5 mL AetF reaction was prepared with 1 µM AetF, 1 U/mL glucose-6-phosphate dehydrogenase, 5 mM glucose-6-phosphate, 1 mM $NADP^+$, 1 mM L-tryptophan (**4**) (5 mg total), 10 mM KBr, 100 µM FAD, and 50 mM Tris-HCl pH 8 buffer in a 250 mL Erlenmeyer flask, adding AetF last. The reaction was incubated at 30 °C, 120 rpm overnight (19 h total) and quenched by the addition of ~50 mL methanol. The reaction was centrifuged at 10,000 x *g* for 10 min to pellet precipitated protein. The supernatant was transferred to a round bottom flask, and methanol was removed by rotary evaporation. The remaining water mixture was lyophilized overnight and extracted 3 times with 1 mL methanol. The extraction was filtered through an 0.2 µm syringe filter and concentrated to ~500 µL by rotary evaporation. The 5,7-dibromo-L-tryptophan (**6**) product was purified by HPLC using a Kinetex 5 µm C18 (250 x 10.0 mm) column and HPLC Method 1. Acetonitrile was removed by rotary evaporation and the remaining water solution was lyophilized overnight to yield 3.6 mg (40.6% yield) of **6** as a fluffy white powder. **^{1}H-NMR** (500 MHz, d_6-DMSO) δ [ppm] = 11.43 (s, 1H), 8.37 (s, 1H), 7.81 (d, *J* = 1.7 Hz, 1H), 7.43 (d, *J* = 1.7 Hz, 1H), 7.33 (s, 1H), 3.44 (dd, *J* = 7.7, 4.5 Hz, 1H), 3.19 (dd, *J* = 15.1, 4.5 Hz, 1H), 3.02 (dd, *J* = 15.1, 7.7 Hz, 1H). **^{13}C-NMR** (126 MHz, d_6-DMSO) δ [ppm] = 166.2, 133.6, 130.5, 127.2, 124.9, 120.8, 111.2, 110.7, 104.9, 54.7, 27.0.[2]

Analytical-scale AetA reactions. 50 µL reactions containing 30 µM AetA, 1 U/mL glucose-6-phosphate dehydrogenase, 10 mM glucose-6-phosphate, 1 mM $NADP^+$, 1 µM SsuE, 500 µM 5-bromoindole (**7**), 50 mM Tris-HCl pH 8 buffer, 10 mM KBr, and DMSO to a final concentration of 10% were prepared in 1.5 mL centrifuge tubes and incubated at room temperature (approx. 22 °C) overnight. Reactions were quenched by the addition of 100 µL methanol and centrifuged at 14,000 rpm for 10 min prior to filtering with 0.2 µM centrifugal filters. 10 µL of the filtered sample was injected on a Kinetex 5 µm C18 (150 x 4.6 mm) column using LC-MS Method 1. Reactions with 3,5-dibromoindole (**8**), L-tryptophan (**4**), 5-bromo-L-tryptophan (**5**), and desbromo-AETX (**S1**) were conducted under the same conditions. Reactions omitting SsuE were performed with SsuE storage buffer in place of the protein (50 mM Tris-HCl pH 8, 100 mM NaCl, 10% glycerol).

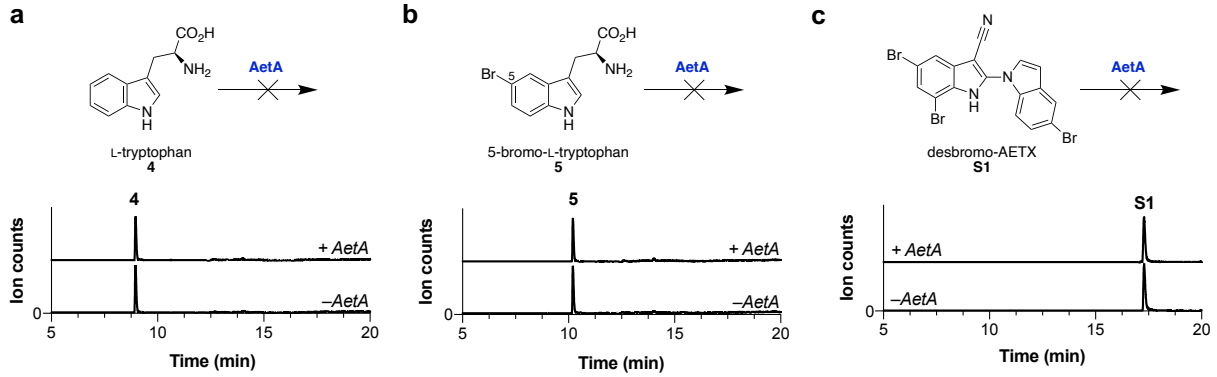


Figure S4. AetA reactions with a) L-tryptophan (**4**), b) 5-bromo-L-tryptophan (**5**), and c) desbromo-AETX (**S1**). LC-MS traces are EICs of the substrates and hypothetical brominated product masses. No brominated products were observed in any of these reactions and no substrate consumption is evident, suggesting AetA is not active on these substrates.

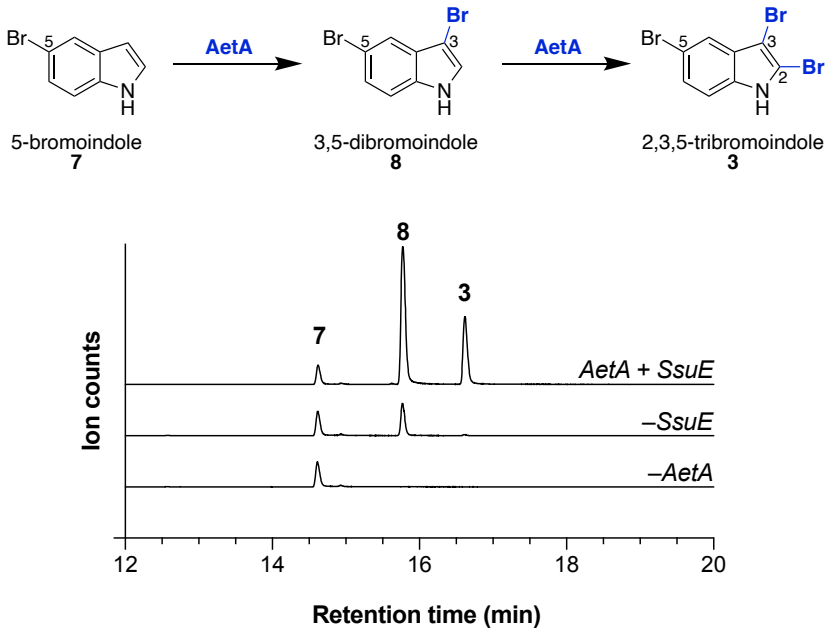


Figure S5. Comparison of AetA reaction with and without SsuE added to reduce FAD in reaction with 5-bromoindole (**7**).

Flavin cofactor identification in AetF. 50 µL of 50 µM purified AetF was precipitated with 50 µL methanol, vortexed vigorously, and centrifuged at 14,000 rpm for 10 min to pellet precipitated protein. The supernatant was filtered through an 0.2 µM centrifugal filter, and 2 µL of the sample was injected on a Kinetex 5 µm C18 (150 x 4.6 mm) column using LC-MS Method 1 and compared to an equimolar concentration of pure FAD standard.

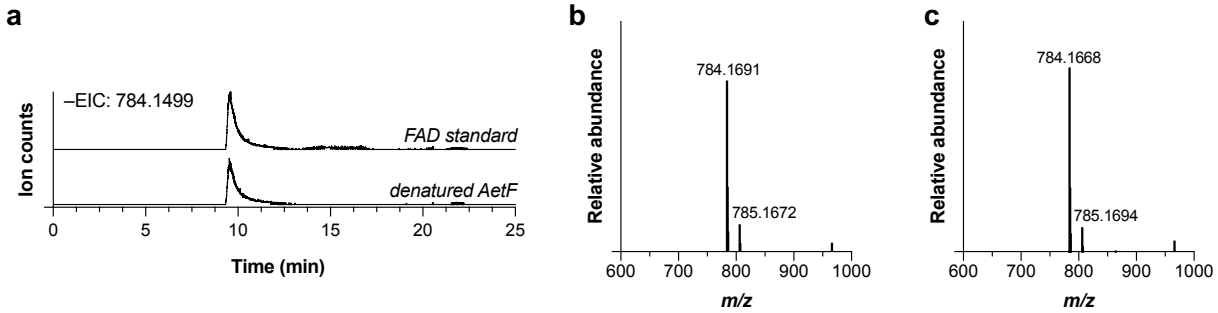


Figure S6. LC-MS comparison of denatured AetF to an FAD standard. a) HPLC separation showing the EIC of FAD in negative ionization mode (*m/z* 784.1499). b) MS extraction of denatured AetF. c) MS extraction of FAD standard.

AetF NAD(P)H preference. 1 mL reactions containing 1 µM AetF, 100 µM NADH or NADPH, 250 µM L-tryptophan (**4**), 10 mM KBr, and 50 mM Tris-HCl pH 8 buffer were prepared, adding AetF last. In no substrate controls, equal volume of DMSO relative to the volume of substrate added was used. The consumption of NAD(P)H was monitored at 340 nm for 5 min using polystyrene cuvettes (BrandTech) and a Cary 60 UV-vis spectrophotometer (Agilent Technologies). The extinction coefficients of NADH and NADPH were used to calculate the concentration of NAD(P)H being consumed. The rapid consumption of NADPH compared to NADH in Figure S7 suggests that NADPH is the preferred reducing cofactor of AetF.

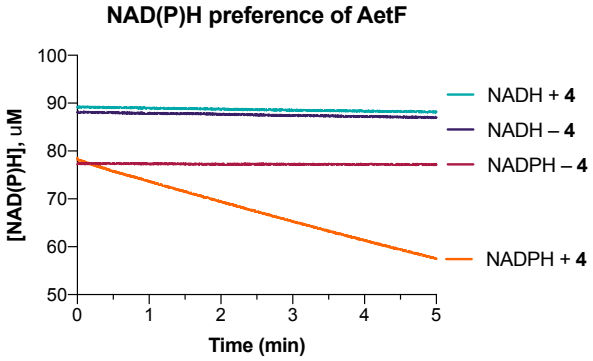


Figure S7. Consumption of NAD(P)H by AetF over time in the presence and absence of L-tryptophan (**4**).

Sequence similarity network generation. The amino acid sequences of representative flavin-dependent monooxygenases (FMOs) from each categorical group are listed in Table S3.[3] These sequences were used to generate a limited sequence similarity network (SSN) using the EFI-EST server at an E-value of 5 and an alignment score of 15. The network was visualized and annotated in Cytoscape v3.6.1.[4] An expanded sequence similarity network consisting of a broader array of flavin-dependent halogenase sequences was constructed using the published dataset and analysis parameters provided in Fisher et al.[5]

Table S3. Enzymes included in the representative FMO sequence similarity network presented in this work. UniProt ID numbers are shown for each sequence, with the exception of SatH and PCMO sequences that are indicated by UniParc IDs.

Enzyme abbreviation	Type	Uniprot
AetF	halogenase (single component)	A0A861B9Z9
AetA	halogenase (two component)	A0A861B8S3
Bmp5	halogenase (single component)	A0A166V3F8
hs_Bmp5	halogenase (single component)	A0A1S6M1M2
VemK	halogenase (single component)	A0A410P811
PrnA	halogenase (two component)	G0ZGJ1
HalB	halogenase (two component)	Q71ME2
RebH	halogenase (two component)	Q8KHZ8
PyrH	halogenase (two component)	K7QVV7
CmdE	halogenase (two component)	Q0VZ69
ThdH	halogenase (two component)	A1E280
KtzQ	halogenase (two component)	A8CF75
KtzR	halogenase (two component)	A8CF74
SttH	halogenase (two component)	E9P162
MibH	halogenase (two component)	E2IHC5
SpmH	halogenase (two component)	A0A1W7D4T7
MalA	halogenase (two component)	L0E155
BrvH	halogenase (two component)	B4WBL8
Tar14	halogenase (two component)	W5VG40
VirX1	halogenase (two component)	M4SKV1
BorH	halogenase (two component)	M9QSI0
SatH	halogenase (two component)	UPI000996F5C4
PltA	halogenase (two component)	B3G2A5
CmlS	halogenase (two component)	Q9AL91
Pyr29	halogenase (two component)	A3R4S0
CndH	halogenase (two component)	B9ZUJ5
Bmp2	halogenase (two component)	U6BGC3
Clz5	halogenase (two component)	U6A3L4
Mpy16	halogenase (two component)	J7H1A1
CtoA	halogenase (two component)	A0A0M4KXQ6
PHBH	Group A	P00438
MICAL	Group A	Q8VDP3
HPCDO	Group A	A0A7W6SQT2
PdxH	Group A	P0AFI7
NahG	Group A	P23262
PheA	Group A	P31020
KMO	Group A	P38169
DhpH	Group A	Q93NG3
HPAO	Group A	A0A4Y1MVS5
TfdB	Group A	Q8KN28
MobA	Group A	Q6SSJ6
XlnD	Group A	Q9F131
ABMO	Group A	Q92402
NphA	Group A	A0A285B4J2
OnpA	Group A	A7KS54
TetX	Group A	B1MM05
NamA	Group A	A0A109YL79
HpbA	Group A	O06647
PcpB	Group A	P42535
BdxA	Group A	A0A375IET4
PhacB	Group A	Q078T0
HpxO	Group A	A6T923
NicC	Group A	P86491
MhpA	Group A	P77397
NicB	Group A	B1N1A2
DntB	Group A	Q2PWU9
OpdA	Group A	A7YJX0
NPMO	Group A	A0A1E1F8I8
UbiH	Group A	P25534
UbiF	Group A	P75728
UbiI	Group A	P25535
TmuM	Group A	I7ASS7
SdgC	Group A	Q7X281
PhzS	Group A	Q9HWG9
RIFMO	Group A	F2R776
VioD	Group A	Q9S3U8
VioC	Group A	Q9S3U9
RebC	Group A	Q8KI25
PgaE	Group A	Q93LY7
RdmE	Group A	Q54530
SibG	Group A	C0LTM1
AuaG	Group A	H1ZZA4
FmqB	Group A	Q4WLW7
AspB	Group A	P0DOW1
AsuE1	Group A	D7P5V0
XiaO	Group A	I3VFC5
CHMO	Group B	P12015
FMO3	Group B	P31513
PvdA	Group B	Q51548
YUC2	Group B	Q9SVQ1
PAMO	Group B	Q47PU3
CpnB	Group B	Q8GAW0
CzcO	Group B	O50641
MbtG	Group B	Q1B6B3
HapE	Group B	Q93TJ5
Sno1	Group B	Q8MP06
TMO	Group B	Q83XK4
OTEMO	Group B	H3JQW0
FMO1	Group B	P38866
SCO7468	Group B	Q8CJJ9
CddA	Group B	Q938F6
CpdB	Group B	T2HVF7
AcmA	Group B	A1IHE6
EthA	Group B	P9WNF8
MtmOIV	Group B	Q194P4
PenE	Group B	E3VWK3
StcW	Group B	Q00730
IucD	Group B	B0R9Q2
PubA	Group B	Q8ECU3
SsDesB	Group B	B5HNG5
LuxA	Group C	P19907
CamE36	Group C	D7UER1
SsuD	Group C	P80645
LadA	Group C	O50325
HPAH	Group C	Q6Q272
DmoA	Group C	E9JFX9
CamP	Group C	Q6STM1
NtaA	Group C	P54989
DszC	Group C	Q64F43
BdsA	Group C	Q8GRC7
EmoA	Group C	Q9F9T3
RutA	Group C	P75898
KijD3	Group D	B3TMR1
PCMO	Group D	UPI00197ED330
NpcB	Group D	Q6F4M9
TftD	Group D	O87009
Reut_A1585	Group D	Q471I2
VlmH	Group D	P96072
NcnH	Group D	D8IPV8
StyA	Group E	O50214
ZEP	Group E	Q9FGC7
LimB	Group E	Q9EUT9
SQLE	Group E	Q14534
SpeA	Group G	B7H6U9
DavB	Group G	B3IVI6
PAO	Group G	Q5W9R9
IaaM	Group G	P06617
LMO	Group H	P21795
NMO	Group H	Q12723

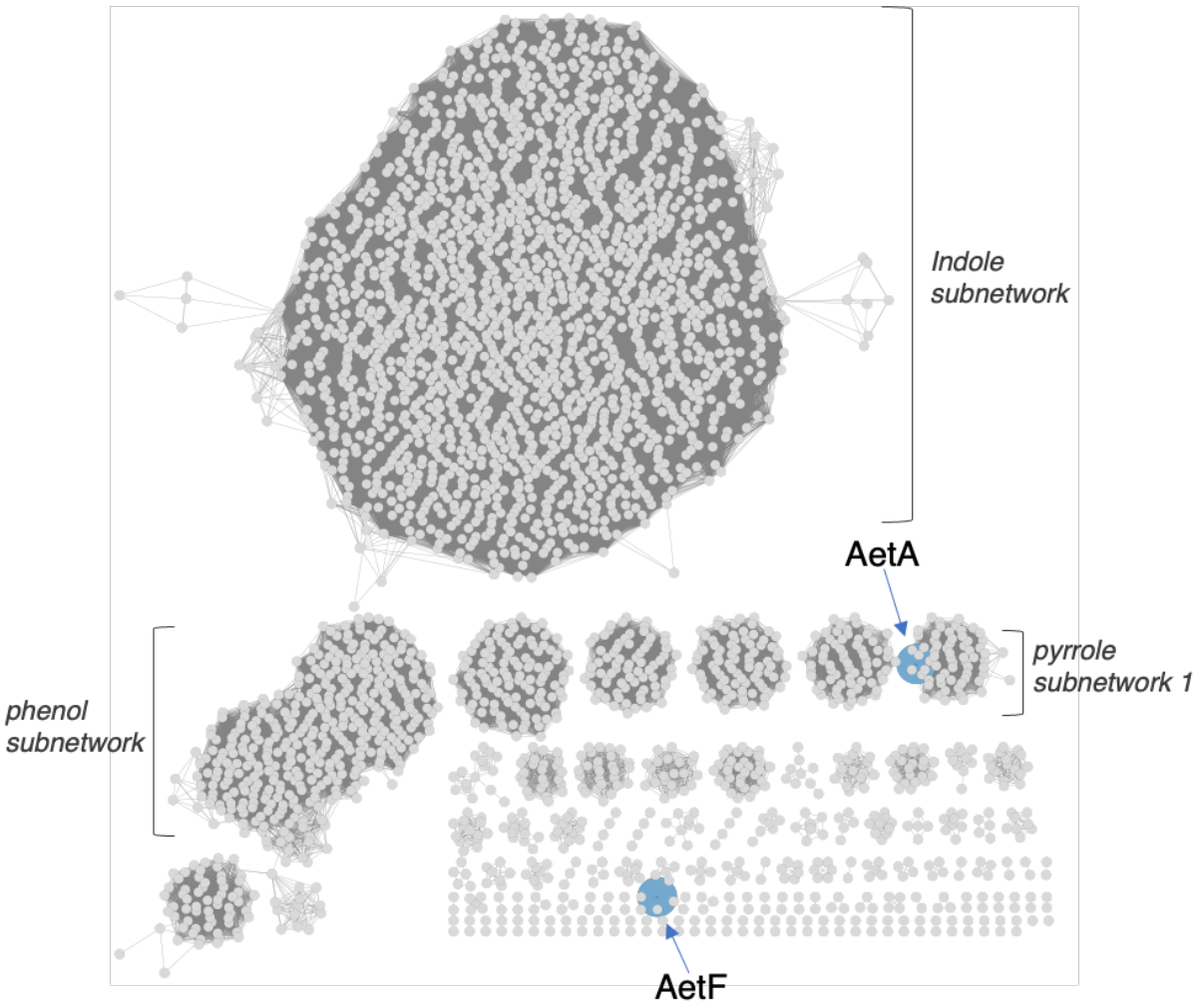


Figure S8. Expanded sequence similarity network of flavin-dependent halogenases. The curated flavin-dependent halogenase sequence dataset and selected subnetwork labels are derived from Fisher et al.[5] Protein sequences for AetA and AetF were seeded into the dataset and analyzed following the published data analysis parameters. AetA clusters with a subset of pyrrole halogenases while AetF is a singleton.

Analytical-scale AetB reaction. A typical 100 µL reaction mixture contained 500 µM 2,3,5-tribromoindole (**3**), 500 µM 5,7-dibromoindole-3-carbonitrile (**2**), 3 µM spinach ferredoxin (Fdx), 1 µM spinach ferredoxin reductase (FdR), 75 µg AetB in 100 mM phosphate buffer, pH 7.5. The reaction was initiated by adding 2 mM NADPH and incubated at RT overnight. The reaction was quenched by the addition of 100 µL acetonitrile, and the resulting solution was filtered using 0.2 µM centrifugal filters. The samples were analyzed by LC-MS Method 2.

Analytical-scale AetB reaction using different indoles. The reaction was set up using the similar condition discussed above. The only difference being 500 µM 5-bromoindole (**7**) and 500 µM 5,7-dibromoindole (**S2**) were also added in the reaction mixture. The EICs of the corresponding coupled products were looked for during LC-MS analysis.

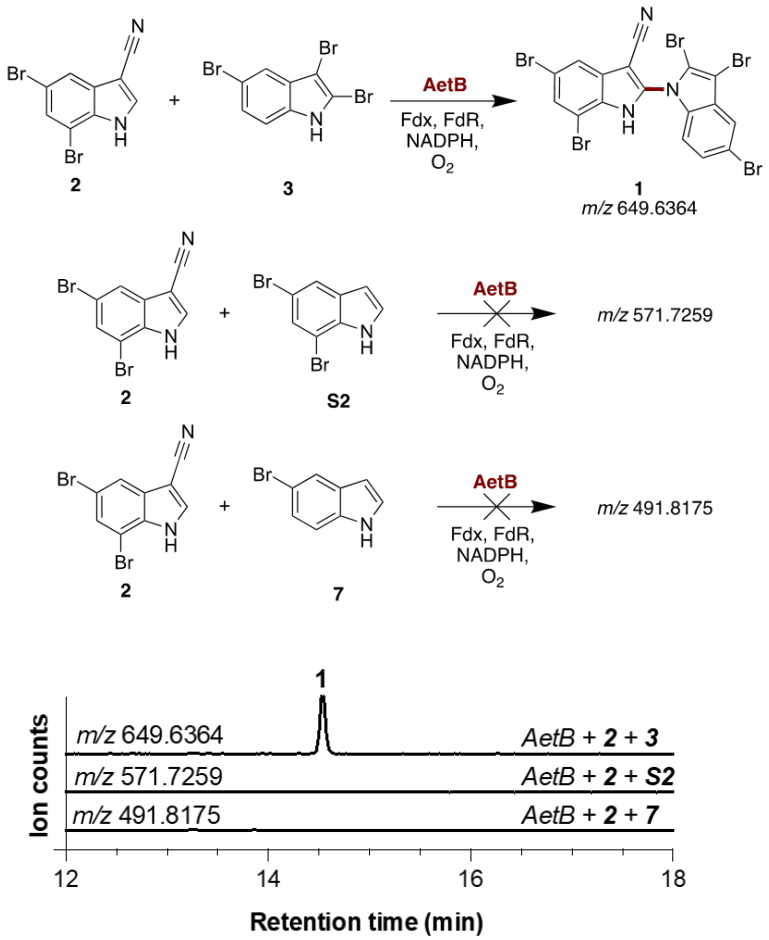


Figure S9. AetB reactions with **2** and 5-bromoindole (**7**) and 5,7-dibromoindole (**S2**). The AetB reaction with **2** and **3** is shown as a positive control.

Analytical-scale AetE reaction: The AetE catalyzed reaction was performed in 100 mM phosphate buffer, pH 7.5. A 100 µL aliquot of the reaction mixture containing 500 µM 5-bromo-L-tryptophan (**5**) and 50 µM AetE was incubated at RT overnight. The reaction was quenched by the addition of 100 µL acetonitrile, and the resulting solution was filtered using 0.2 µM centrifugal filters. The samples were analyzed by HPLC Method 2.

Detection of pyruvate in AetE reaction: The AetE catalyzed reaction was performed as discussed above. After the overnight incubation 1 mM *o*-(2,3,4,5,6-pentafluorobenzyl)-hydroxylamine (PFBHA) was added to the reaction mixture and incubated for 30 min at 50 ºC. Upon denaturing the protein using 50% acetonitrile, the sample was filtered using 0.2 µM centrifugal filters. The samples were analyzed by LC-MS Method 2.

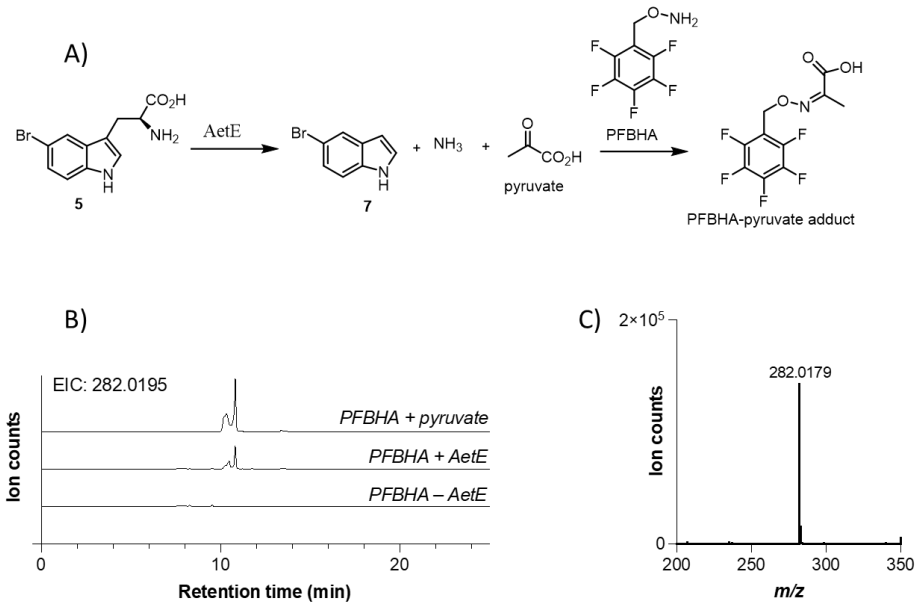


Figure S10. A) PFBHA was used to trap pyruvate generated in AetE reaction. B) Chromatograms shown are EICs of PFBHA-pyruvate adduct (*m/z* 282.0195) analyzed in negative ionization mode. C) Exact mass of the PFBHA-pyruvate adduct in negative ionization mode.

Analytical-scale AetC and AetD reaction: The reaction was carried out in 100 mM phosphate buffer, pH 7.5. A 100 µL reaction mixture containing 500 µM 5,7-dibromotryptophan (**6**), 2 µM AetC, and 25 µM AetD was incubated at RT overnight. The reaction was quenched by the addition of 100 µL acetonitrile, and the resulting solution was filtered using 0.2 µM centrifugal filters. The samples were analyzed by LC-MS Method 2.

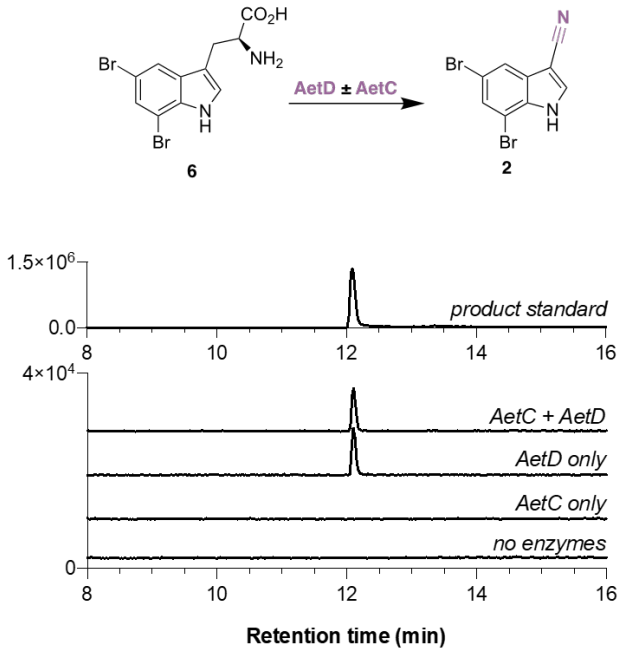


Figure S11. AetD reactions with **6** in the presence and absence of AetC. Chromatograms shown are EICs of **2** (*m/z* 298.8648) analyzed in negative ionization mode.

Screening of first row transition metals for improved AetD activity: Here is the list of metal salts screened for improved AetD activity: $MnCl_2$, $(NH_4)_2Fe(SO_4)_2$, $FeCl_3$, $CoCl_2$, $NiCl_2$, $CuCl_2$, $ZnSO_4$.The reaction was carried out in 100 mM phosphate buffer, pH 7.5. A 100 µL reaction mixture containing 500 µM 5,7-dibromotryptophan (**6**), 50 µM AetD, and 1 mM of the individual metal ion was incubated at RT overnight. The reaction was quenched by the addition of 100 µL acetonitrile, and the resulting solution was filtered using 0.2 µM centrifugal filters. The samples were analyzed by LC-MS Method 2.

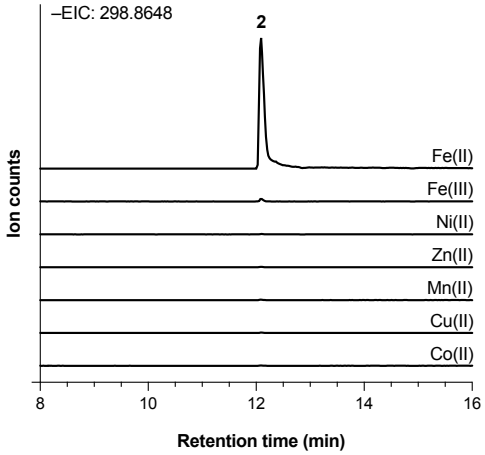
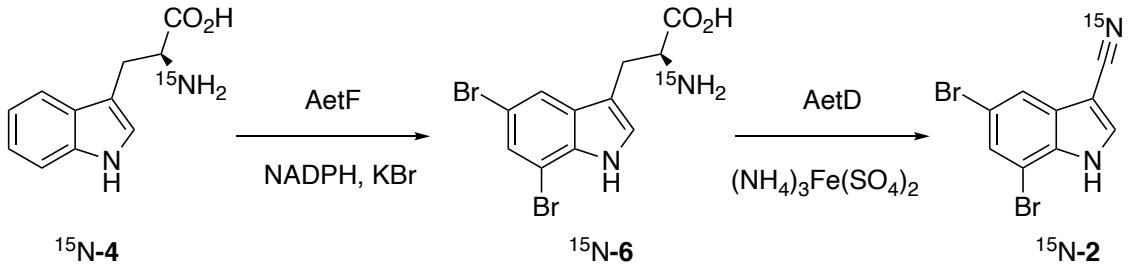


Figure S12. AetD reactions with **6** in the presence of different transition metals. Chromatograms shown are EICs of **2** (*m/z* 298.8648) analyzed in negative ionization mode.

AetD reaction in the presence of $^{15}NH_4Cl$: A 100 µL reaction mixture containing 500 µM 5,7-dibromotryptophan (**6**), 50 µM AetD, 1mM $(NH_4)_2Fe(SO_4)_2$, 10mM $^{15}NH_4Cl$ in 100 mM phosphate buffer, pH 7.5 was incubated at RT overnight. The reaction was quenched by the addition of 100 µL acetonitrile, and the resulting solution was filtered using 0.2 µM centrifugal filters. The samples were analyzed by LC-MS method 2.

AetD reaction in the presence of enzymatically prepared 5,7-dibromotryptophan-(amino-15N): To prepare 5,7-dibromotryptophan-(amino-15N) (^{15}N-**6**), a reaction mixture containing 500 µM L-tryptophan-(amino-15N) (^{15}N-**4**), 5 µM AetF, 10 mM KBr, 2 mM NADPH in 100 mM phosphate buffer, pH 7.5 was incubated in RT for 5 hrs. Afterwards, 50 µM AetD and 1 mM $(NH_4)_2Fe(SO_4)_2$ were added to the reaction mixture and incubated overnight. The reaction was quenched by the addition of 100 µL acetonitrile, and the resulting solution was filtered using 0.2 µM centrifugal filters. The samples were analyzed by LC-MS method 2.



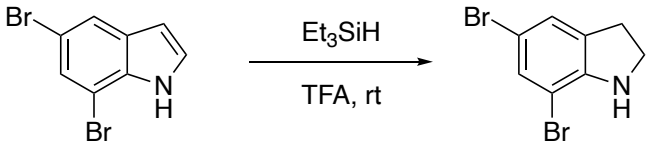
Scheme S1. Schematic showing AetD catalyzed reaction in the presence of enzymatically generated 5,7-dibromotryptophan-(amino-15N) (^{15}N-**6**).

4. Synthetic procedure of substrates and synthetic standards

General considerations. All chemicals were purchased from Fisher Scientific, Alfa Aesar, or MilliporeSigma without further purification. All solvents were of HPLC grade or higher. Preparative flash column chromatography was carried out on a Teledyne ISCO CombiFlash® Rf+ Lumen™ system using using silica gel 60 (EMD, 40-63 μm) for the stationary phase. NMR spectroscopic data were obtained on a 500 MHz JEOL NMR spectrometer with either a 3.0 mm probe or a 5.0 mm probe. The values of the chemical shifts are described in ppm and coupling constants are reported in Hz. NMR chemical sifts were referenced to the residual solvent peaks (δH 2.50 and δC 39.5 for DMSO-d_6; δH 7.26 and δC 77.16 for $CDCl_3$; δH 3.58 δC 67.57 for THF-d_8. NMR data analysis was performed using MestreNova® 14.21-27684, 2021.

4.1 Synthesis of aetokthonotoxin (1)

*5,7-Dibromoindoline (**S3**)*



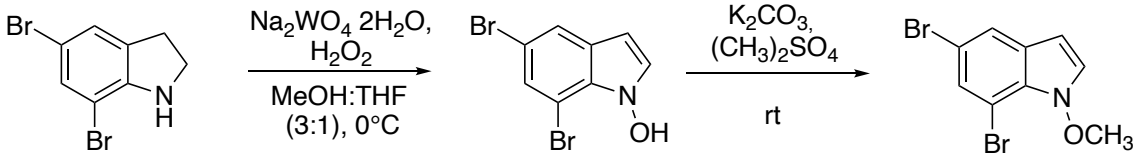
S2 **S3**

Adapted from Tomakinian *et al.*[7]
5,7-Dibromoindole **S2** (100 mg, 0.724 mmol, 1.0 equiv.) was weighed into a 10 mL flask. TFA (2.0 mL) was added dropwise, and the reaction mixture was stirred until 5,7-dibromoindole **S2** was completely dissolved. Then Et_3SiH (116 μL, 1.452 mmol, 2.0 equiv.) was added dropwise, and the reaction mixture was heated to 60 °C overnight. After removal of TFA through evaporation, sat. $NaHCO_{3\,(aq)}$ was added under ice cooling. The aqueous phase was extracted twice with CH_2Cl_2, and the combined organic layers were washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by normal phase column chromatography with a gradient of 0–100% EtOAc in hexanes. 5,7-dibromoindoline **S3** was obtained as a yellow oil in 85% yield (178 mg, 0.615 mmol). TLC (Hex:EtOAc (2:1) R_F = 0.8).

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 7.28 (dd, J = 1.7 Hz, 1H), 7.11 (d, J = 1.5 Hz, 1H), 3.97 (bs, 1H*, 3.62 (t, J = 8.5, 2H), 3.23–3.03 (m, 2H).**^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] =149.50, 132.21, 131.64, 126.51, 109.53, 102.98, 46.90, 30.76.

The following synthetic route was adapted from Ricardo et al.[6]

*5,7-Dibromo-1-methoxy-1H-indole (**S5**)*

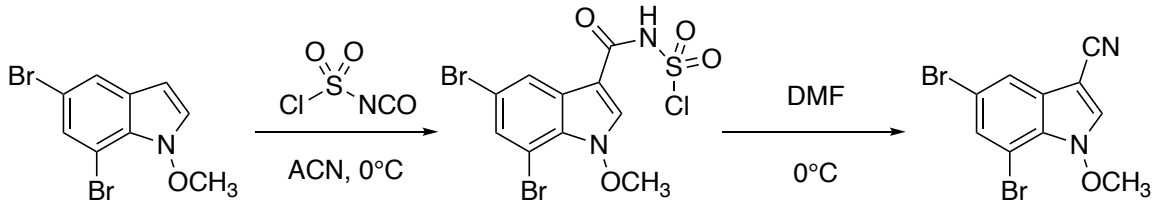


S3 **S4** **S5**

5,7-dibromoindoline **S3** (178 mg, 0.642 mmol, 1.0 equiv.) was dissolved in 4 mL MeOH:THF (3:1) at 0 °C. Na_2WO_5 (38 mg, 0.129 mmol, 0.2 equiv.) was dissolved in 500 µL H_2O and added to the stirred solution of **S3** followed by dropwise addition of 30% H_2O_2 (656 µL, 10 equiv.). The resulting solution was stirred at 0 °C. After 1h additional Na_2WO_5 (29 mg, 0.098 mmol, 0.15 equiv.) dissolved in 400 µL H_2O and 30% H_2O_2 (600 µL, 9 equiv.) was added, and the reaction was stirred until complete consumption of starting material was observed by TLC (R_f (toluene) = 0.29). Dimethyl sulfate (340 µL, 1.8 mmol, 5.6 equiv.) was added to the crude reaction mixture at room temperature, followed by K_2CO_3 (390 mg, 2.88 mmol, 4.5 equiv.). The reaction mixture was vigorously stirred for 1 h until full consumption of starting material **S4** was observed by TLC (Hex:EtOAc (9:1) R_F = 0.55). Brine and $CHCl_3$ were added in equal ratios, phases were separated, and the aqueous phase was extracted twice with $CHCl_3$. The combined organic phases were dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by automated column chromatography (gradient: 5 min 100% hexanes, 5-15 min 0-10% EtOAc). 5,7-dibromo-1-methoxy-1*H*-indole (**S5**) was obtained as a white solid in 26% yield (50 mg, 0.166 mmol).

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 7.64 (d, J = 1.7 Hz, 1H), 7.51 (d, J = 1.7 Hz, 1H), 7.29 (d, J = 3.4 Hz, 1H), 6.31 (d, J = 3.5 Hz, 1H), 4.09 (s, 3H). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] = 129.08, 128.31, 128.15, 126.34, 122.90, 113.06, 102.60, 98.25, 67.84.

*5,7-Dibromo-1-methoxy-1H-indole-3-carbonitrile (**S7**)*



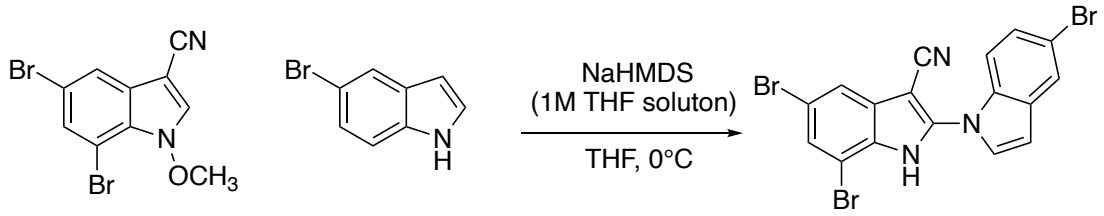
S5 **S6** **S7**

S5 (50 mg, 0.164 mmol, 1.0 equiv.) was dissolved in CH_3CN (1 mL), cooled to 0 °C, and CSI (50 µL, 0.575 mmol, 3.4 equiv.) was added. The reaction mixture turned slightly pink and after 30 min a white precipitate formed. To convert intermediate **S6** to **S7**, DMF (0.5 mL) was added at 0 °C and the precipitate redissolved, rendering a slightly yellow solution. TLC analysis showed full consumption of starting material and formation of a single new product (Hex:EtOAc (4:1) R_F = 0.3). The reaction mixture was concentrated under reduced pressure and resuspended in equal volumes of H_2O and EtOAc. Phases were separated, and the organic phase was washed three times with $LiCl_{aq}$ (5%) and once with brine. The organic layer was dried over $MgSO_4$, filtered, and

concentrated under reduced pressure. 5,7-dibromo-1-methoxy-1*H*-indole-3-carbonitrile (**S7**) was obtained in 58% yield (32 mg, 54 µmol) as a white powder.

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 7.84 (d, J = 1.6 Hz, 1H), 7.80 (s, 1H), 7.67 (d, J = 1.7 Hz, 1H), 4.18 (s, 3H). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] = 132.05, 131.77, 127.72, 127.51, 121.85, 116.28, 113.72, 103.57, 82.90, 69.16.

*Desbromo-aetokthonotoxin (**S1**)*

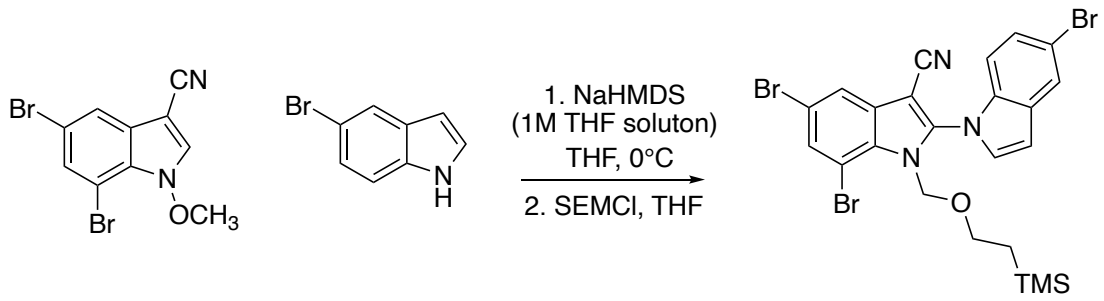


S7 **7** desbromo-AETX **(S1)**

5-bromoindole (**7**) (13 mg, 65 µmol, 1.2 equiv.) was dissolved in anhydrous THF (400 µL), cooled to 0 °C, and flushed with argon. NaHMDS (1 M THF solution) (110 µL, 109 µmol, 2.0 equiv.) was added dropwise under inert gas atmosphere. The reaction mixture was stirred for 30 min at 0 °C. 5,7-dibromo-1-methoxy-1*H*-indole-3-carbonitrile (**S7**) (18 mg, 55 µmol, 1.0 equiv.) was dissolved in anhydrous THF (400 µL), purged with argon, and then added dropwise to the reaction mixture at 0 °C, which immediately turned dark red. The reaction was allowed to slowly warm up to room temperature. The reaction was quenched with sat NH_4Cl_{aq}, followed by addition of equal volume EtOAc. The phases were separated, and the organic phase was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by reversed phase HPLC (80-100% CH_3CN, 20 min), and the desired product desbromo-AETX (**S1**) was obtained as a white solid in 78% yield (21 mg, 42 µmol). TLC: Hex:DCM (2:1) R_F = 0.7

^{1}H-NMR (500 MHz, d_8-THF) δ [ppm] = 10.87 (s, 1H, NH), 7.92 (d, J = 1.8 Hz, 1H), 7.86 (d, J = 2.0 Hz, 1H), 7.73 (d, J = 1.7 Hz, 1H), 7.70 (d, J = 3.4 Hz, 1H), 7.54 (d, 8.7 Hz, 1H), 7.42 (dd, J = 8.8, 2.0 Hz, 1H), 6.81 (d, J= 3.4 Hz, 1H). **^{13}C-NMR** (126 MHz, d_8-THF) δ [ppm] = 142.61, 136.09, 132.57, 132.34, 130.55, 130.22, 127.20, 124.78, 121.93, 116.46, 116.06, 113.97, 113.57, 107.09, 106.73, 82.21.

*Protected desbromo-aetokthonotoxin (**S8**)*

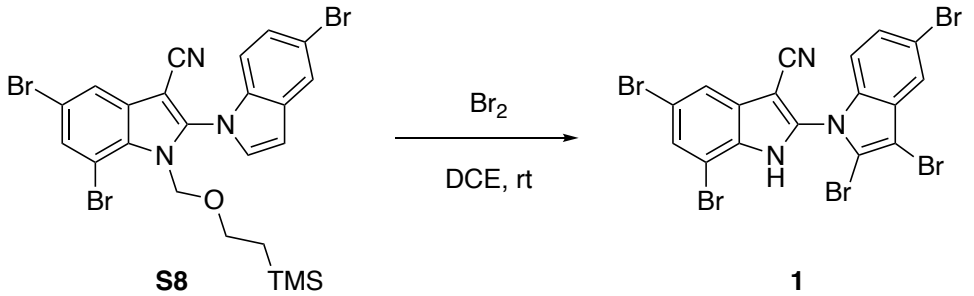


S7 **7** **S8**

5-bromoindole **7** (22 mg, 112 µmol, 1.2 equiv.) was dissolved in anhydrous THF (0.8 mL), cooled to 0 °C, and flushed with argon. NaHMDS (1 M THF solution) (224 µL, 224 µmol, 2.0 equiv.) was added dropwise, and the reaction was stirred for 15 min at 0 °C. Dibrominated nitrile indole **S7** (31 mg, 94 µmol, 1.0 equiv.) was dissolved in anhydrous THF (0.8 mL) and added dropwise at 0 °C to the reaction mixture. The reaction mixture was allowed to slowly warm up to room temperature. After 2 h, SEMCl (40 µL, 225 µmol, 2.4 equiv.) was added and stirred for 5 min. The reaction mixture was quenched through addition of NH_4Cl and diluted with EtOAc. Phases were separated, and the aqueous phase was extracted three times with equal volumes EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by normal phase column chromatography (0-20% EtOAc in hexanes). The desired product **S8** was obtained in 57% yield (33 mg, 53 µmol) as a white powder. TLC (Hex:EtOAc) R_F = 0.8.

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 7.89 (d, J = 1.8 Hz, 1H), 7.87 (d, J = 1.9 Hz, 1H), 7.78 (d, J = 1.8 Hz, 1H), 7.42 (dd, J = 8.7, 1.9 Hz, 1H), 7.35 (d, J = 3.4 Hz, 1H), 7.15 (d, J = 8.7 Hz, 1H), 6.86-6.78 (m, 1H), 5.74 (d, J = 10.6 Hz, 1H), 5.24 (d, J = 10.6 Hz, 1H), 3.33-3.03 (m, 2H), 0.75-0.70 (m, 2H), –0.12 (s, 9H). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] = 142.3, 135.8, 133.1, 130.8, 130.3, 130.2, 129.4, 127.2, 124.5, 121.83, 117.1, 115.8, 112.5, 112.1, 106.8, 106.4, 84.7, 72.9, 66.6, 17.8, -1.5.

*Aetokthonotoxin (**1**)*



S8 **1**

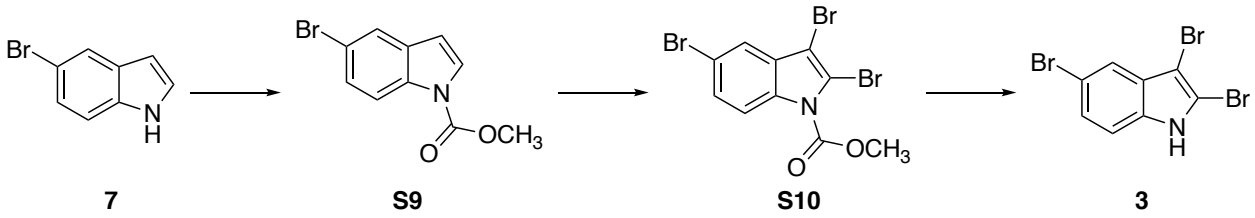
Protected biindole **S8** (20 mg, 32 µmol, 1.0 equiv) was dissolved in dichloroethane (0.5 mL). To the solution was then added 400 µL of a freshly prepared bromine stock solution (50 µL bromine in 2.5 mL dichloroethane; 160.2 µmol, 5.0 equiv.) at room temperature. The reaction mixture was stirred at room temperature for 1 h until TLC showed full consumption of starting material (Hex:EtOAc (9:1), R_F = 0.5). The mixture was quenched with sat. $Na_2S_2O_4$ until decolorization of the reaction mixture was observed. Equal volumes of EtOAc were added, and phases were

separated. The organic phase was washed three times with saturated $NaHCO_3$ solution and once with brine. The combined organic fraction was dried over $MgSO_4$, filtered, and concentrated in vacuo. The crude product was dissolved in small amounts of CH_2Cl_2. After 5 min a white precipitate formed. The supernatant was removed, and the white solid was taken up in toluene and concentrated under reduced pressure to dryness. The final natural product AETX (**1**) was obtained as a crystalline white solid in 72% yield (15 mg, 23 µmol). The ^{1}H and ^{13}C NMR signals agree with the reported literature value.[6]

^{1}H-NMR (500 MHz, d_8-THF) δ [ppm] = 12.50 (s, 1H), 8.04 (d, J = 1.7 Hz, 1H), 7.83 (d, J = 1.6 Hz, 1H), 7.77 (d, J= 1.8 Hz, 1H), 7.47 (dd, J = 8.8, 1.9 Hz, 1H), 7.26 (d, J = 8.8 Hz, 1H). **^{13}C-NMR** (126 MHz, d_8-THF) δ [ppm] = 138.13, 137.79, 132.91,131.26, 130.15, 129.68, 129.01, 122.72, 122.66, 117.37, 117.30, 116.81, 113.72, 112.60, 107.56, 98.22, 88.23.

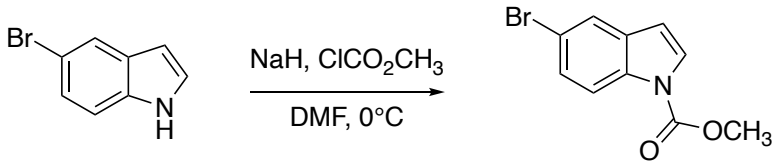
4.2 Synthesis of 2,3,5-tribromoindole (3)

This synthesis was carried out following the protocol reported in Ricardo et al.[6]



7 **S9** **S10** **3**

*Methyl 5-bromo-1H-indole-1-carboxylate (**S9**)*

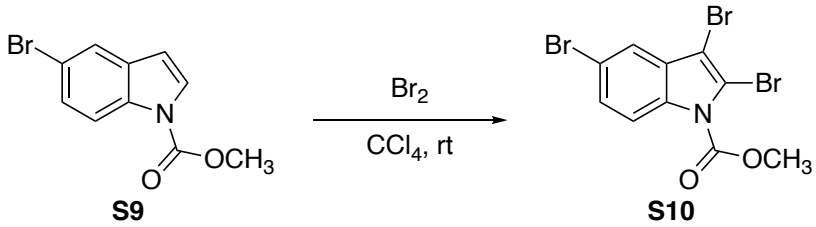


7 **S9**

5 mL DMF was added to a mixture containing 5-bromoindole (**7**) (1 g, 5.1 mmol) and 60% NaH in mineral oil (0.24 g, 6.1 mmol) at 0 ºC. After 30 min of stirring, methyl chloroformate (0.47 mL, 6.1 mmol) was added, and the reaction was stirred until total consumption of the starting material. The reaction was quenched with saturated NH_4Cl, diluted with EtOAc, washed with brine, and dried over $MgSO_4$. Following the removal of the solvent by evaporation at reduced pressure and purification by flash column chromatography on SiO_2 (EtOAc/hexane), **S9** was obtained in 50% yield.

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 8.05 (d, *J* = 8.9 Hz, 1H), 7.70 (d, *J* = 2.0 Hz, 1H), 7.60 (d, *J* = 3.7 Hz, 1H), 7.42 (dd, *J* = 8.7, 2.0 Hz, 1H), 6.54 (d, *J* = 3.7 Hz, 1H), 4.04 (s, 3H). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] = 151.35, 134.08, 132.29, 127.48, 126.77, 123.79, 116.65, 116.49, 107.48, 54.14.[8]

*Methyl 2,3,5-tribromo-1H-indole-1-carboxylate (**S10**)*

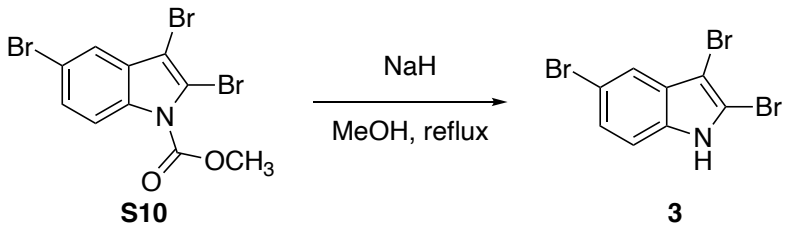


S9 **S10**

Bromine (0.10 mL, 1.9 mmol) dissolved in CCl_4 (1 mL) was added dropwise to a solution containing compound **S9** (60 mg, 0.24 mmol) in CCl_4 (10 mL). The reaction was stirred at room temperature for 12 h. The reaction was quenched with saturated $Na_2S_2O_3$, washed with saturated $NaHCO_3$, brine, and dried over $MgSO_4$. After evaporation of the solvent at reduced pressure, the crude product was purified by flash column chromatography on SiO_2 (EtOAc/hexane) to yield compound **S10** in 40% yield**.**

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 7.92 (d, *J* = 8.9 Hz, 1H), 7.60 (d, *J* = 2.0 Hz, 1H), 7.41 (dd, *J* = 8.9, 2.0 Hz, 1H), 4.09 (s, 3H). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] = 150.35, 134.45, 130.05, 128.69, 121.98, 117.65, 117.13, 112.61, 104.36, 54.45.

*2,3,5-Tribromoindole (**3**)*



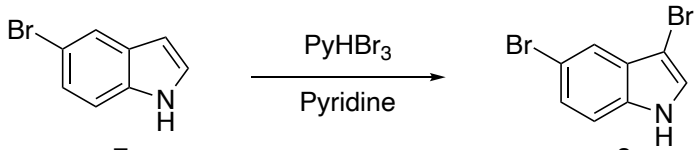
S10 **3**

Compound **S10** (50 mg, 0.12 mmol) was suspended in MeOH (3 mL). 60% NaH in mineral oil (8 mg, 0.18 mmol) was added, and the reaction mixture was refluxed for 2 h. The reaction was quenched with saturated NH_4Cl, diluted with EtOAc, washed with brine, and dried over $MgSO_4$. After removing the solvent by evaporation at reduced pressure, the resulting crude product was purified by flash column chromatography on SiO_2 (EtOAc /hexane) to afford **3** in 50% yield.

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 8.40 (s, 1H), 7.64 (d, *J* = 1.7 Hz, 1H), 7.31 (dd, *J* = 8.6, 1.9 Hz, 1H), 7.17 (d, *J* = 8.7 Hz, 1H). **^{13}C-NMR** (126 MHz, $CDCl_3$**) δ [ppm]** = 134.30, 129.14, 126.50, 121.55, 114.67, 112.30, 111.38, 93.79, 77.40, 77.15, 76.90

4.3 Synthesis of 3,5-dibromoindole (8)

This synthesis was carried out following the protocol reported by Parrick *et al*.[9]



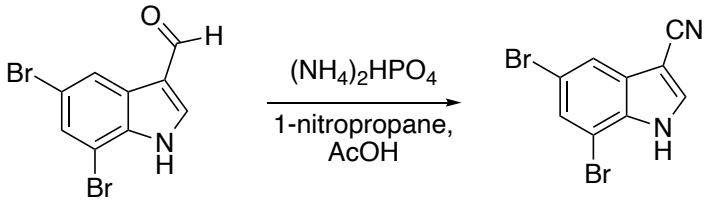
7 **8**

A cold solution of pyridinium hydrobromide perbromide (360 mg, 1.12 mmol) in pyridine (2 mL) was added dropwise to a stirred solution of 5-bromoindole (**7**) (200 mg**,** 1.02 mmol) in pyridine (2 mL) kept on ice. After 0.5 h, the mixture was poured into cold diethyl ether. The precipitated material was filtered off, and the filtrate was washed with dilute hydrochloric acid, dilute aqueous sodium hydroxide and water, dried over $MgSO_4$, and evaporated to yield 3,5-dibromoindole (**8**) in 80% yield.

^{1}H-NMR (500 MHz, $CDCl_3$) δ [ppm] = 8.24 (s, 1H), 7.73 (d, *J* = 1.9 Hz), 7.33 (dd, *J* = 8.6, 1.9 Hz, 1H), 7.24 (d, *J* = 8.7 Hz, 1H), 7.23 (d, *J* = 2.6 Hz, 1H). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ [ppm] = 134.07, 128.68, 126.29, 124.64, 122.01, 114.09, 113.01, 91.09.

4.4 Synthesis of 5,7-dibromoindole-3-carbonitrile (2)

This synthesis was adapted from the experimental protocol reported by Degnan *et al*.[10]



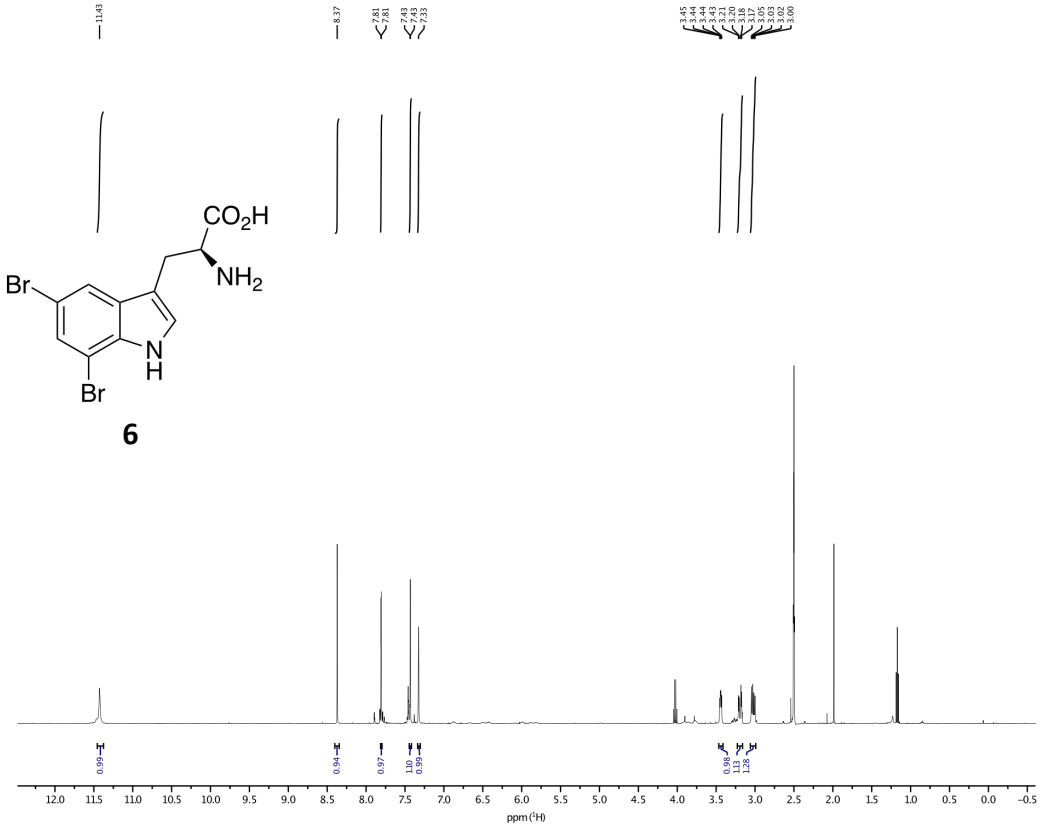
S11 **2**

A mixture containing compound **S11** (100 mg, 0.33 mmol), ammonium phosphate, dibasic (217 mg, 1.65 mmol), and 1-nitropropane (0.056 mL, 0.66 mmol) in acetic acid (3 mL) was heated overnight under reflux condition. The solvent was evaporated, and the residue was taken up in aqueous sodium bicarbonate and extracted with EtOAc. The organic layer was dried over $MgSO_4$ and concentrated under reduced pressure. The crude product was purified by flash column chromatography on SiO_2 (EtOAc /hexane) to give **2** in 30% yield.

^{1}H-NMR (500 MHz, d_6-DMSO) δ [ppm] = 8.40 (s, 1H), 7.84 (d, J = 1.7 Hz, 1H), 7.71 (d, J = 1.7 Hz, 1H). **^{13}C-NMR** (126 MHz, d_6-DMSO) δ [ppm] = 137.08, 133.02, 129.30, 128.00, 120.46, 115.12, 114.36, 106.51, 85.65. **HRMS** (ESI)*: m/z: calcd for $C_9H_4Br_2N_2$* $[M+H]^+$: 297.8741; found: 296.8660

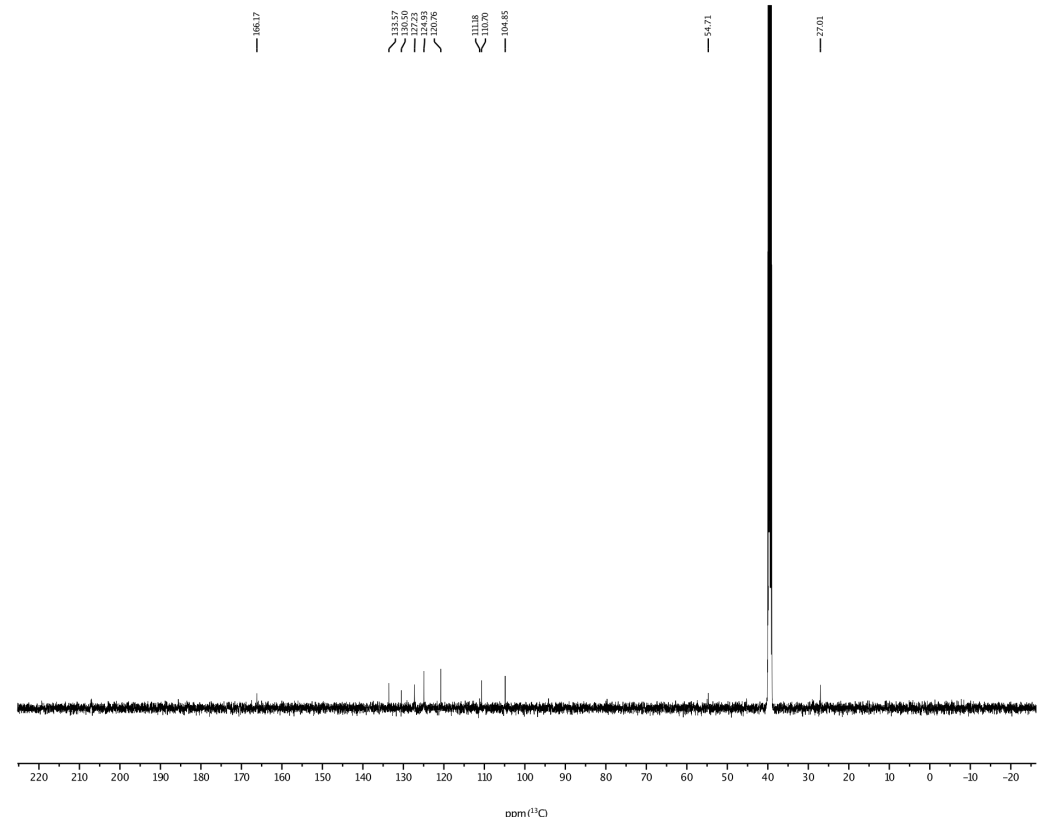
5. NMR spectra

^{1}H NMR (500 MHz, d_6-DMSO)

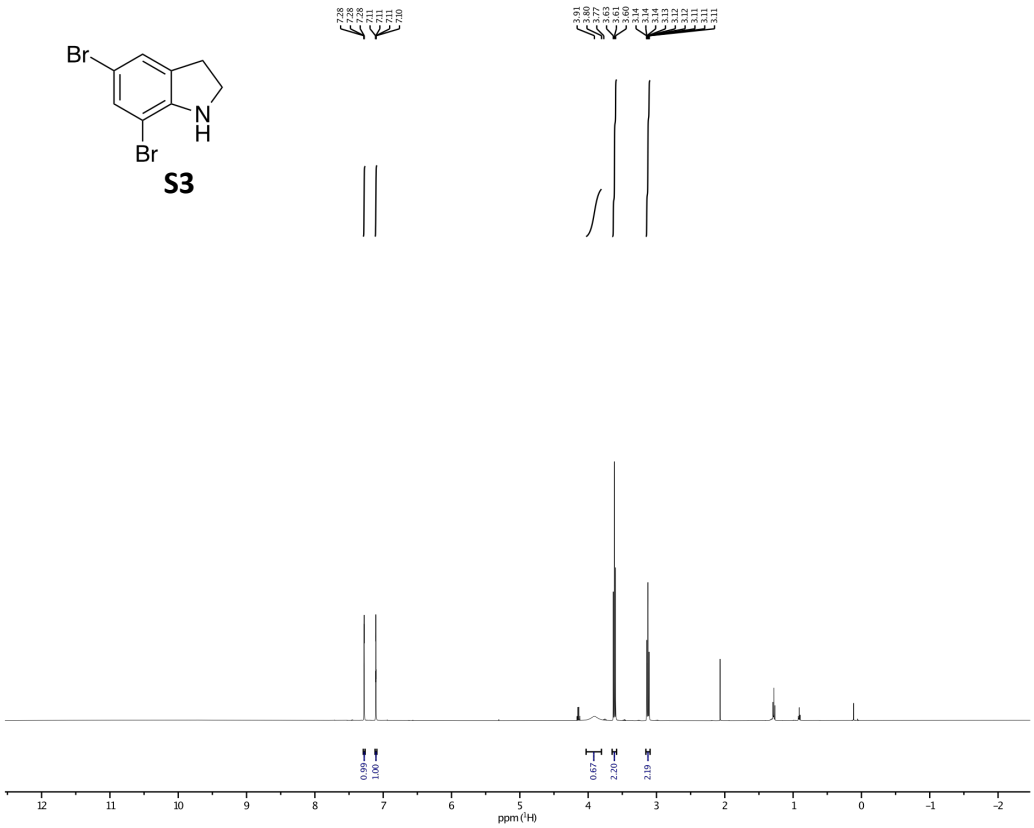


6

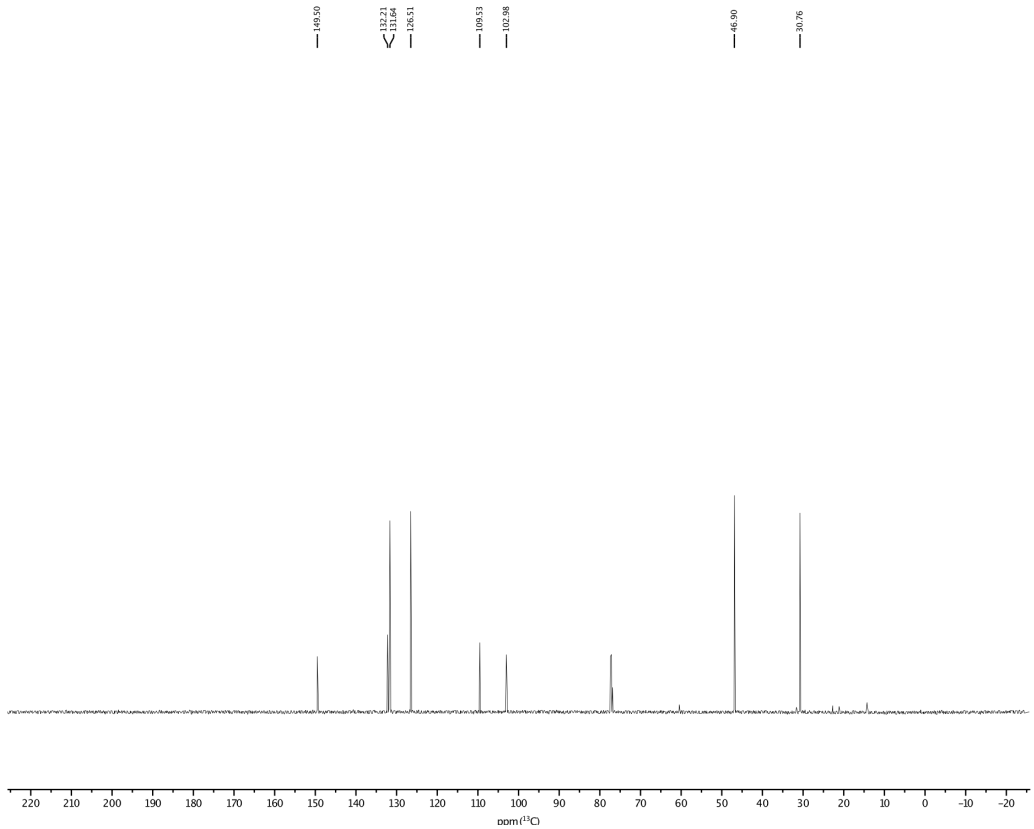
^{13}C NMR (126 MHz, d_6-DMSO)



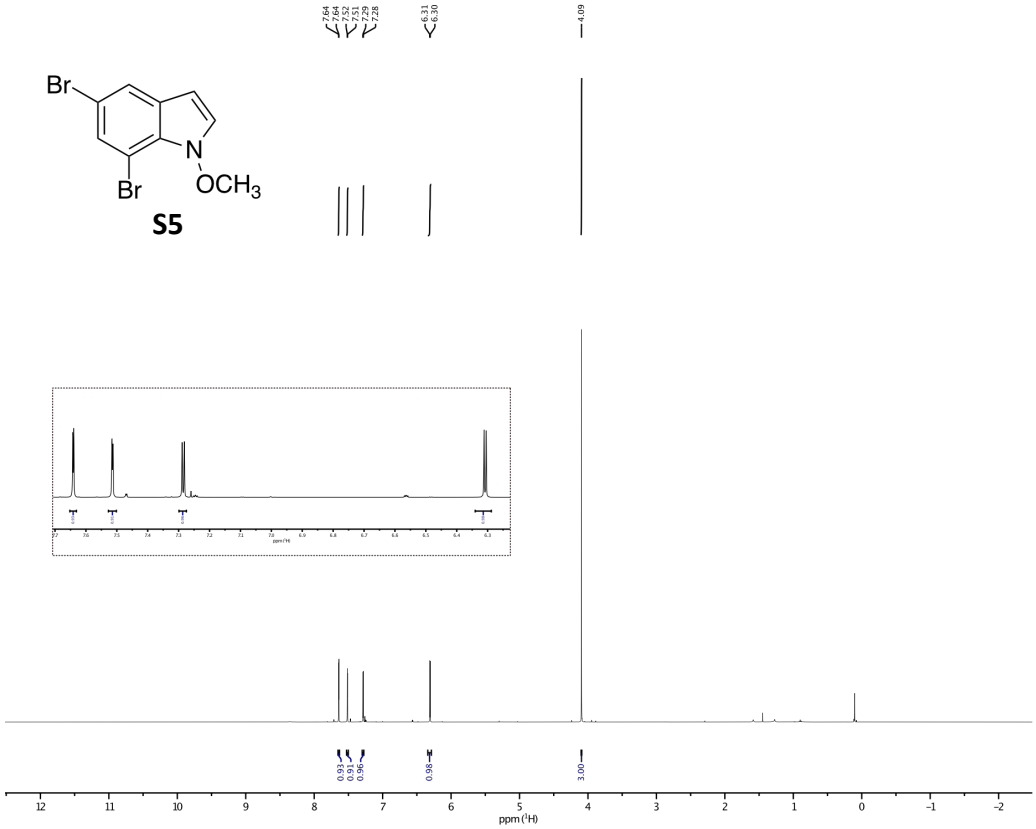
^{1}H NMR (500 MHz, $CDCl_3$)



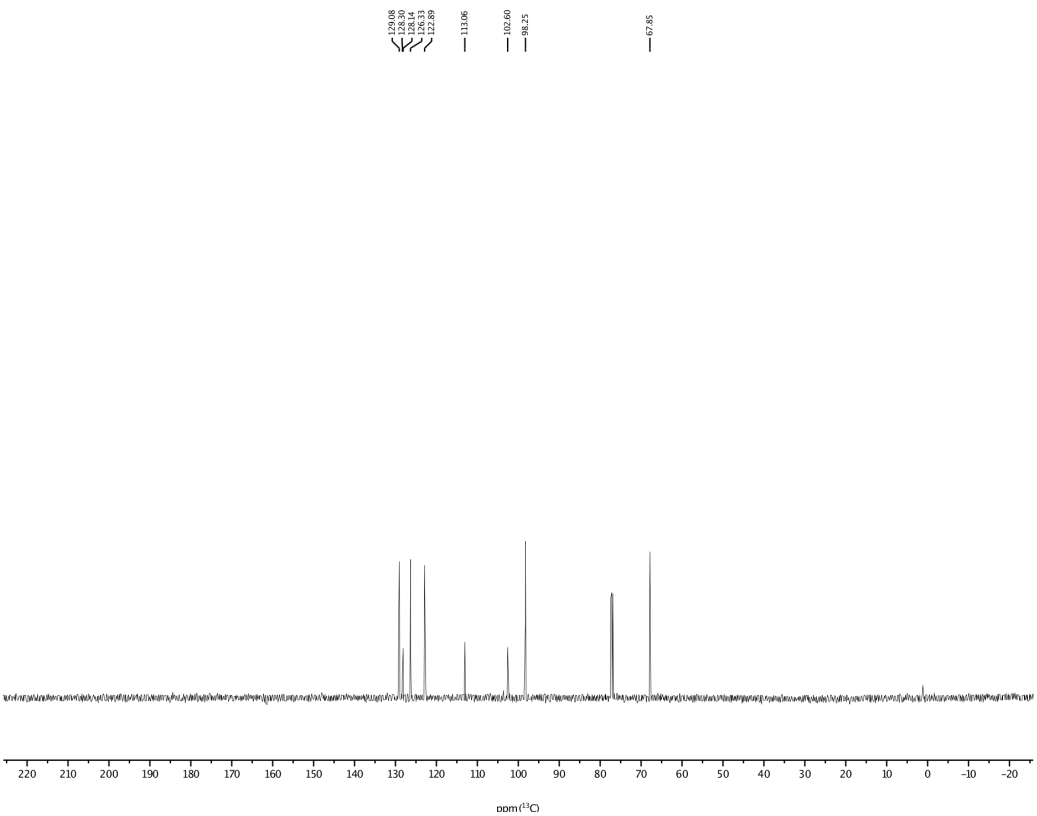
^{13}C NMR (126 MHz, $CDCl_3$)



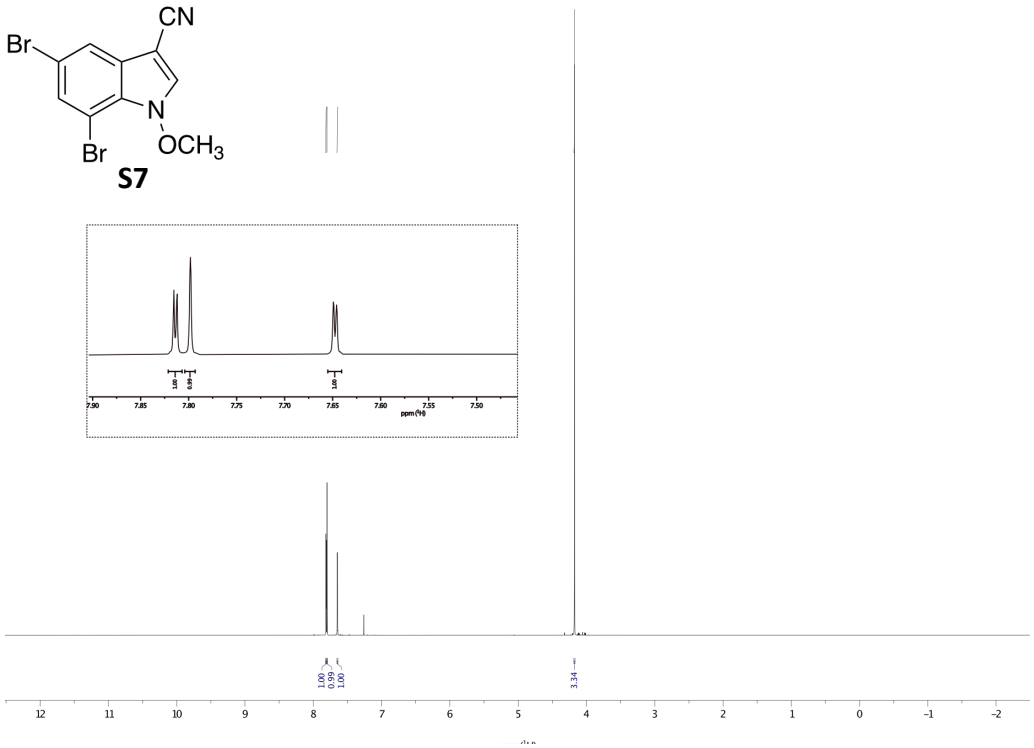
^{1}H NMR (500 MHz, $CDCl_3$)



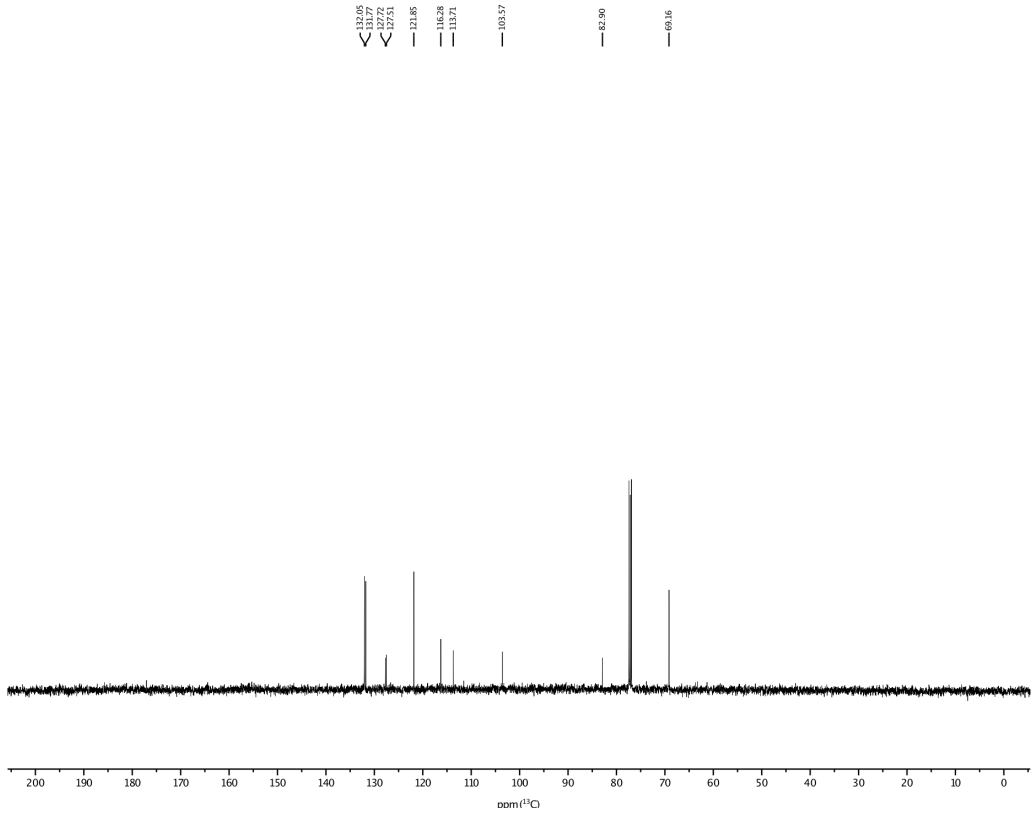
^{13}C NMR (126 MHz, $CDCl_3$)



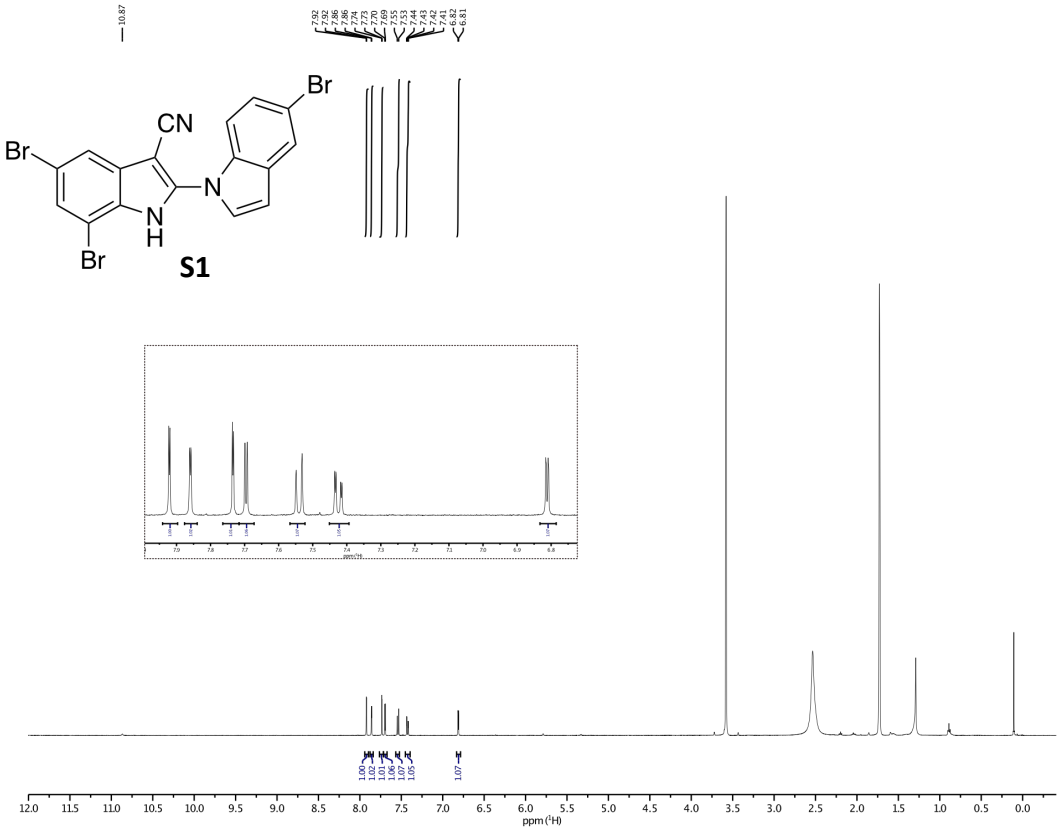
^{1}H NMR (500 MHz, $CDCl_3$)



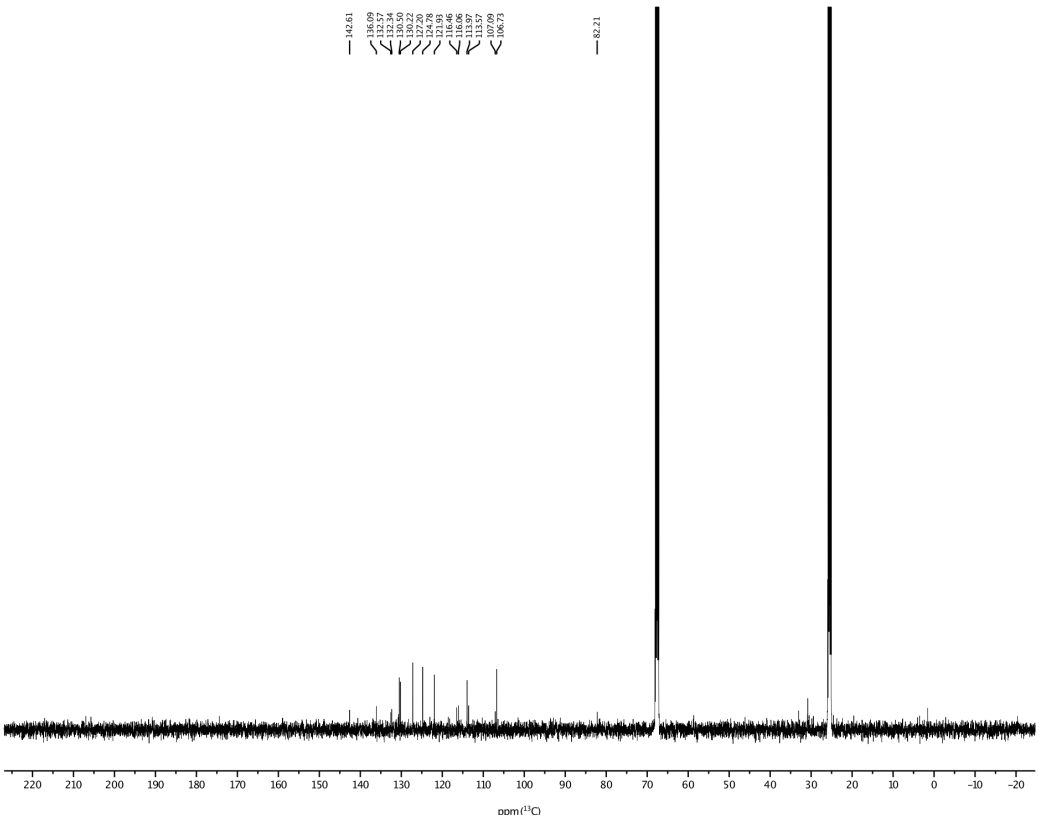
^{13}C NMR (126 MHz, $CDCl_3$)



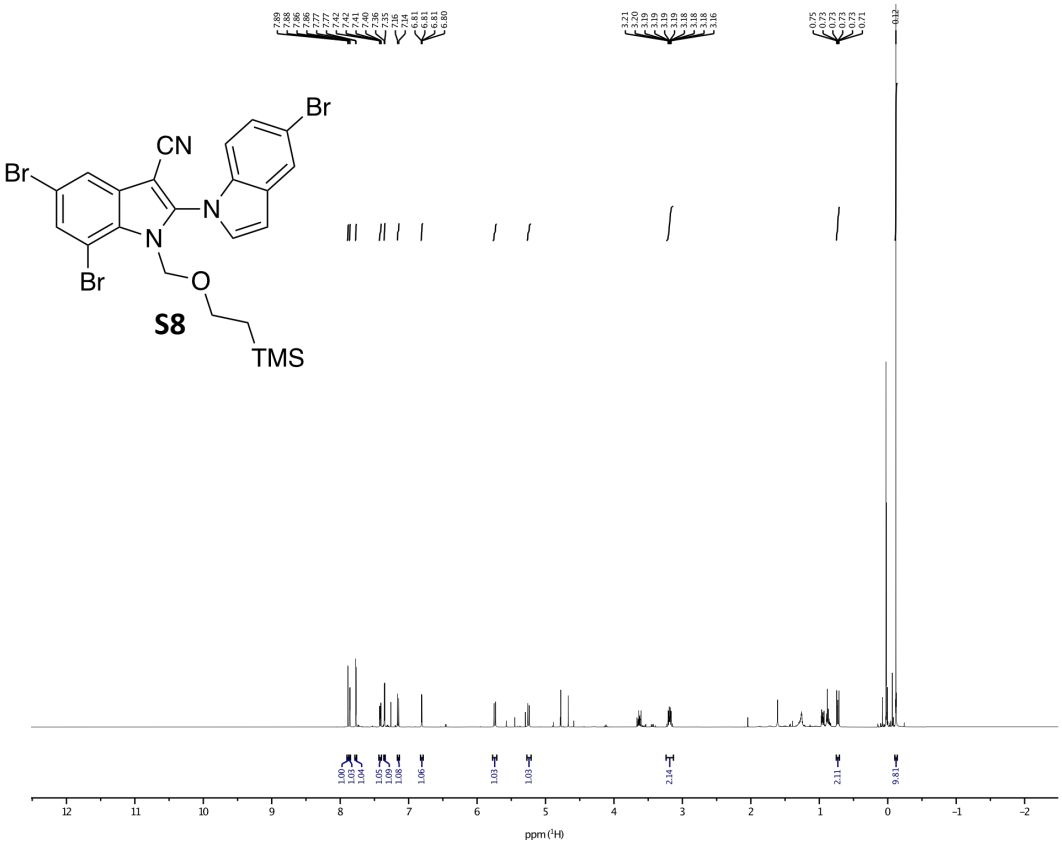
^{1}H NMR (500 MHz, d_8-THF)



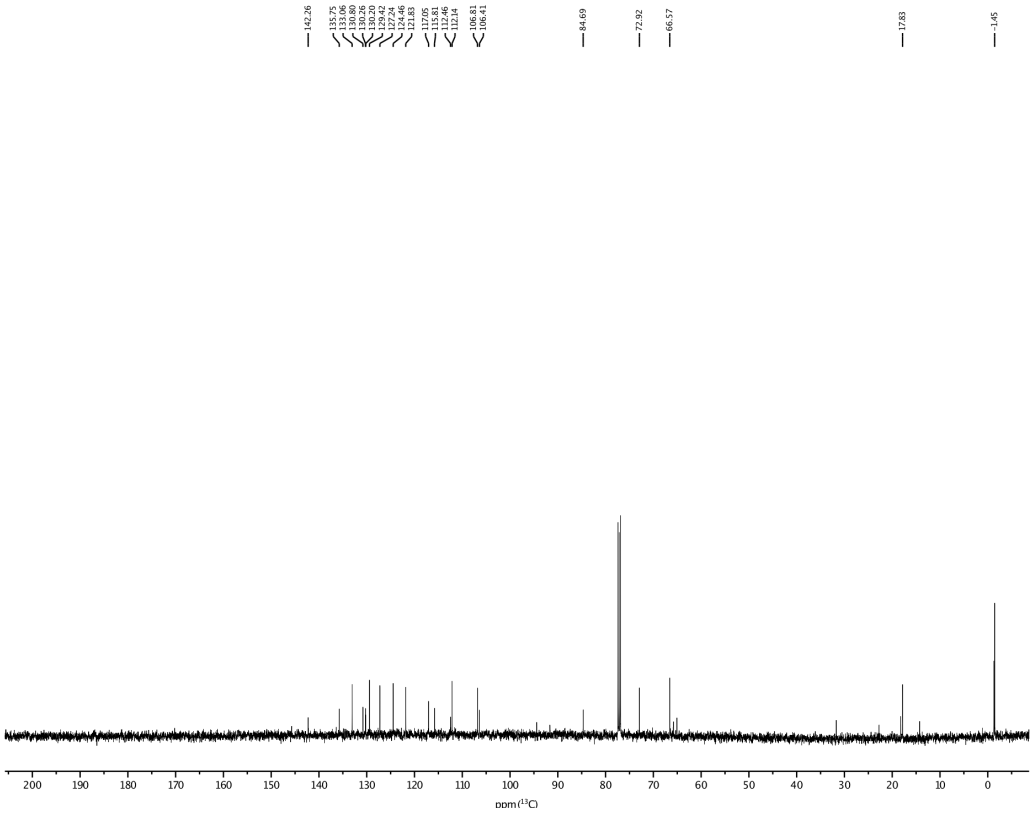
^{13}C NMR (126 MHz, d_8-THF)



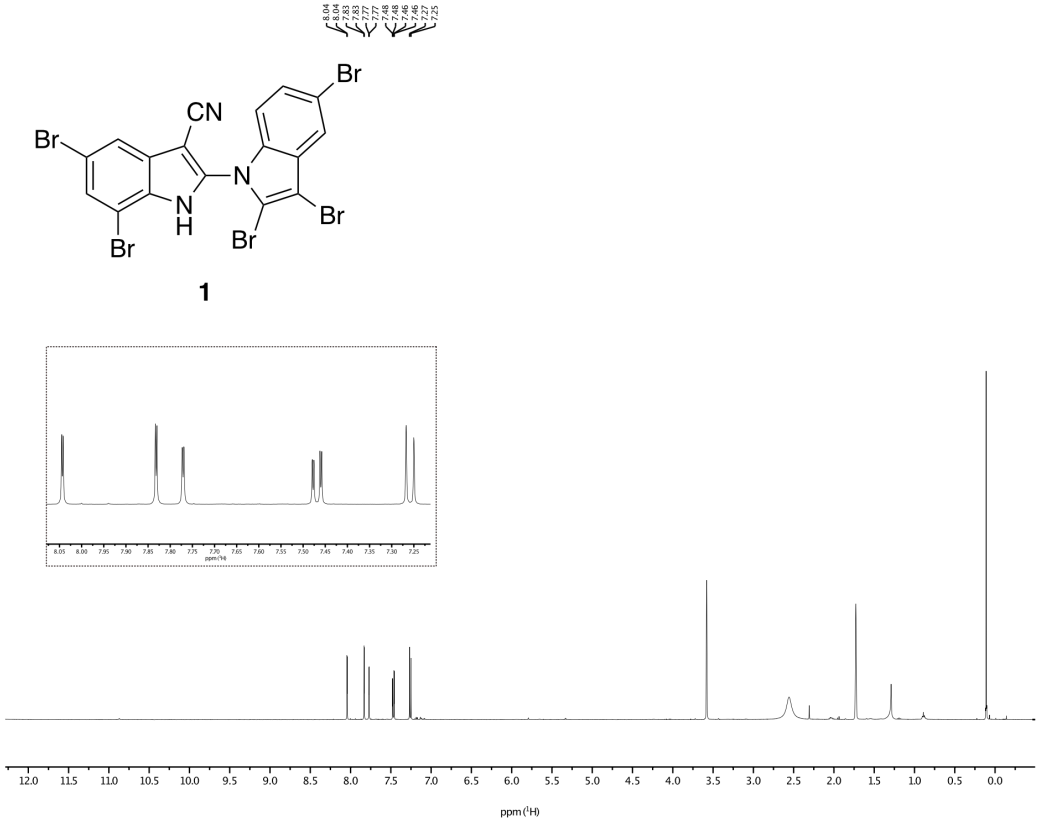
^{1}H NMR (500 MHz, $CDCl_3$)



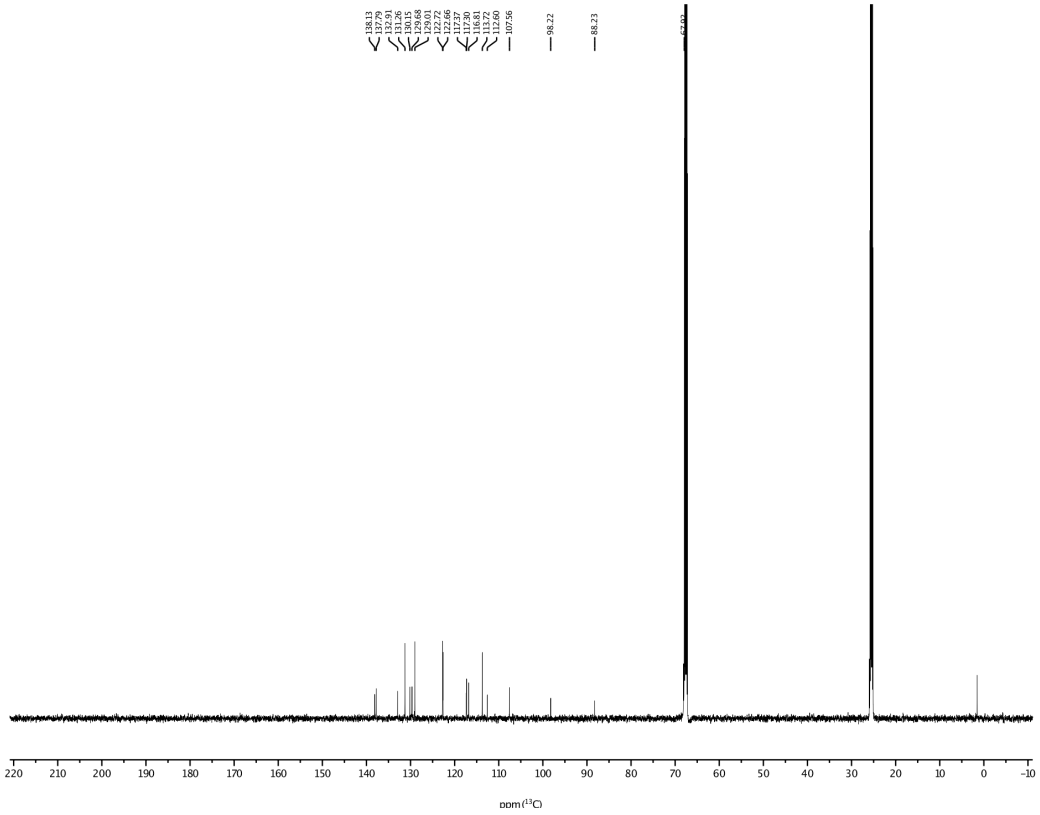
^{13}C NMR (126 MHz, $CDCl_3$)



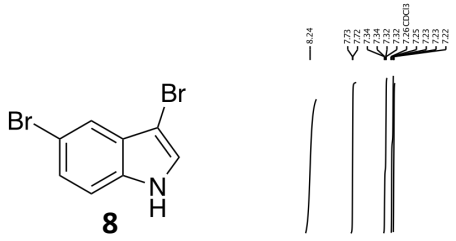
^{1}H NMR (500 MHz, d_8-THF)



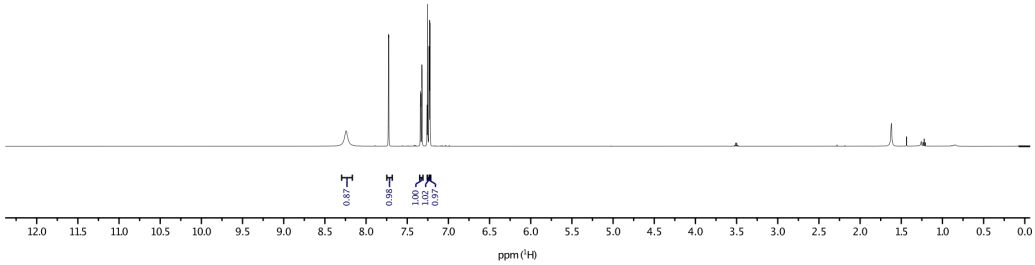
^{13}C NMR (126 MHz, d_8-THF)



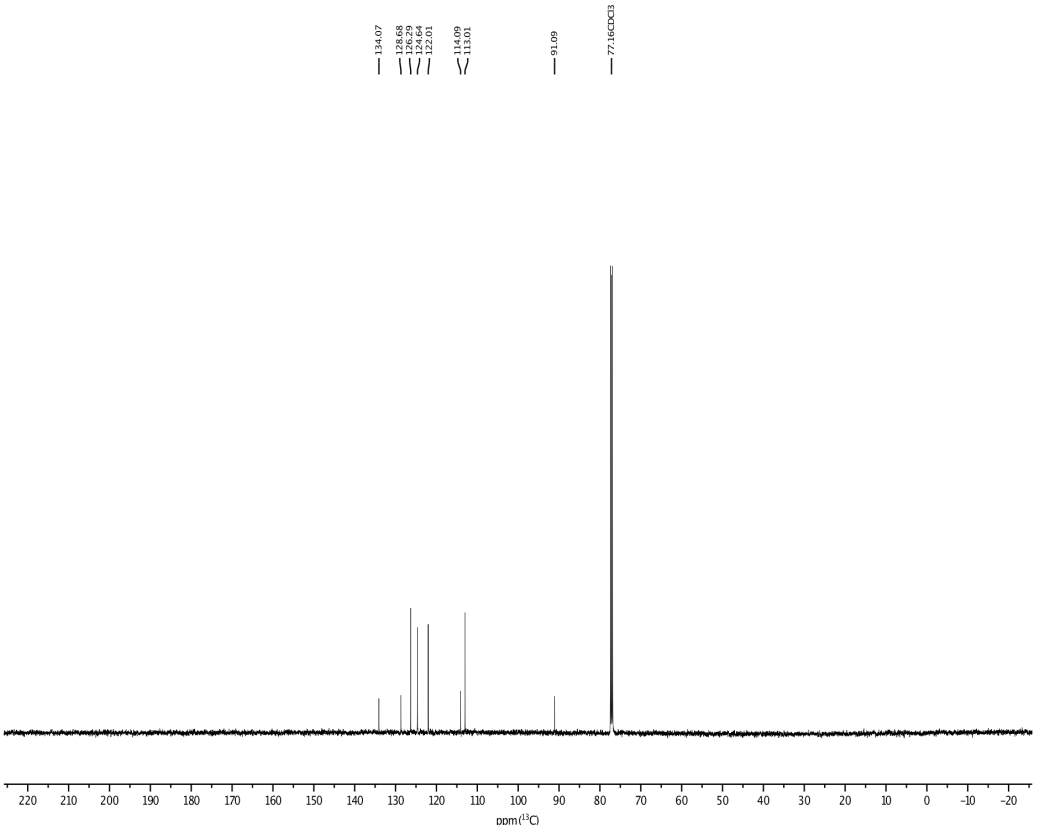
^{1}H NMR (500 MHz, $CDCl_3$)



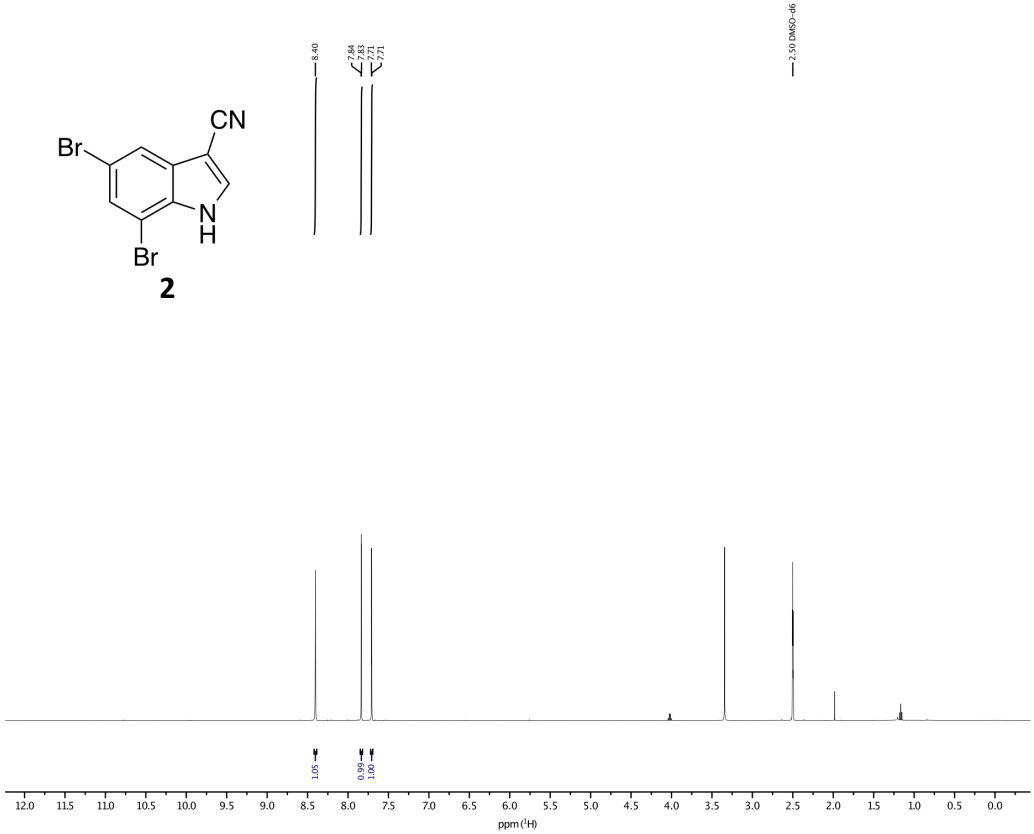
8



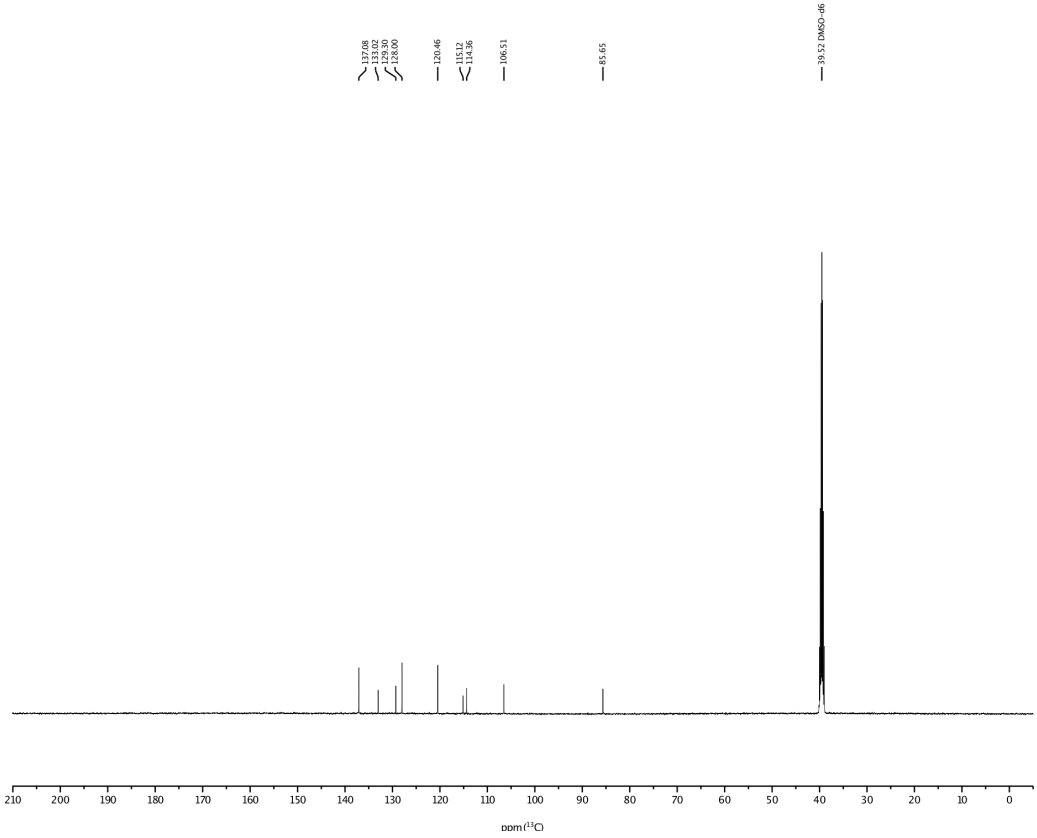
^{13}C NMR (126 MHz, $CDCl_3$)



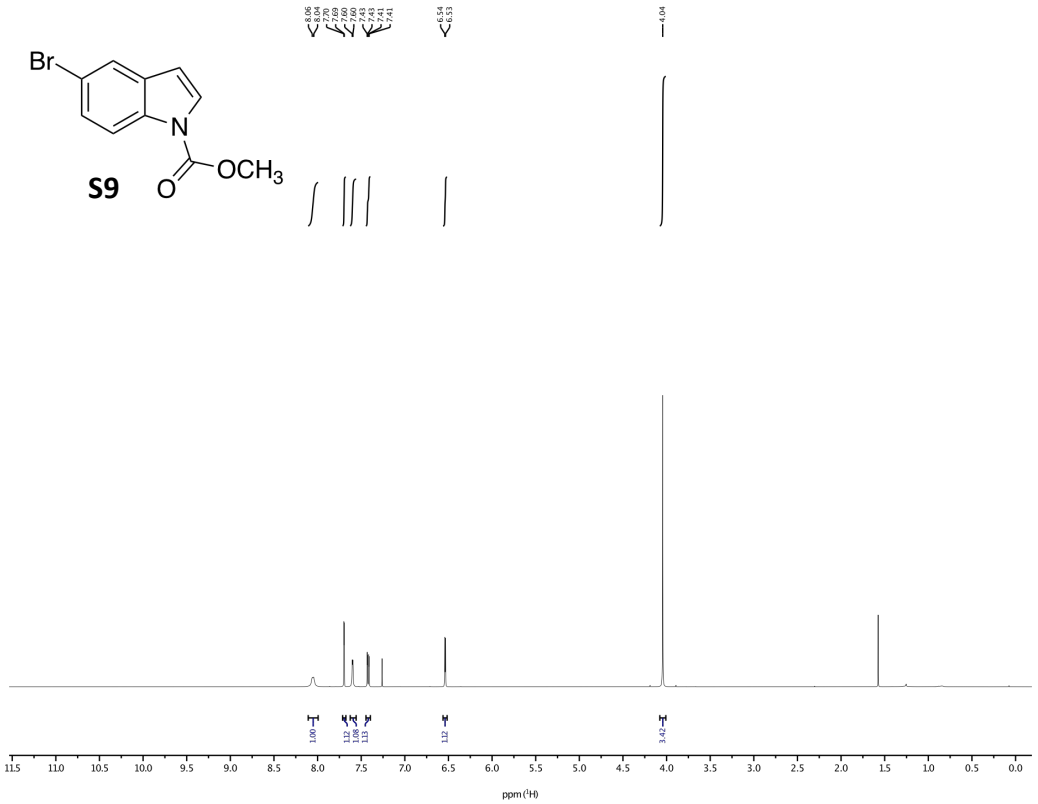
^{1}H NMR (500 MHz, d_6-DMSO)



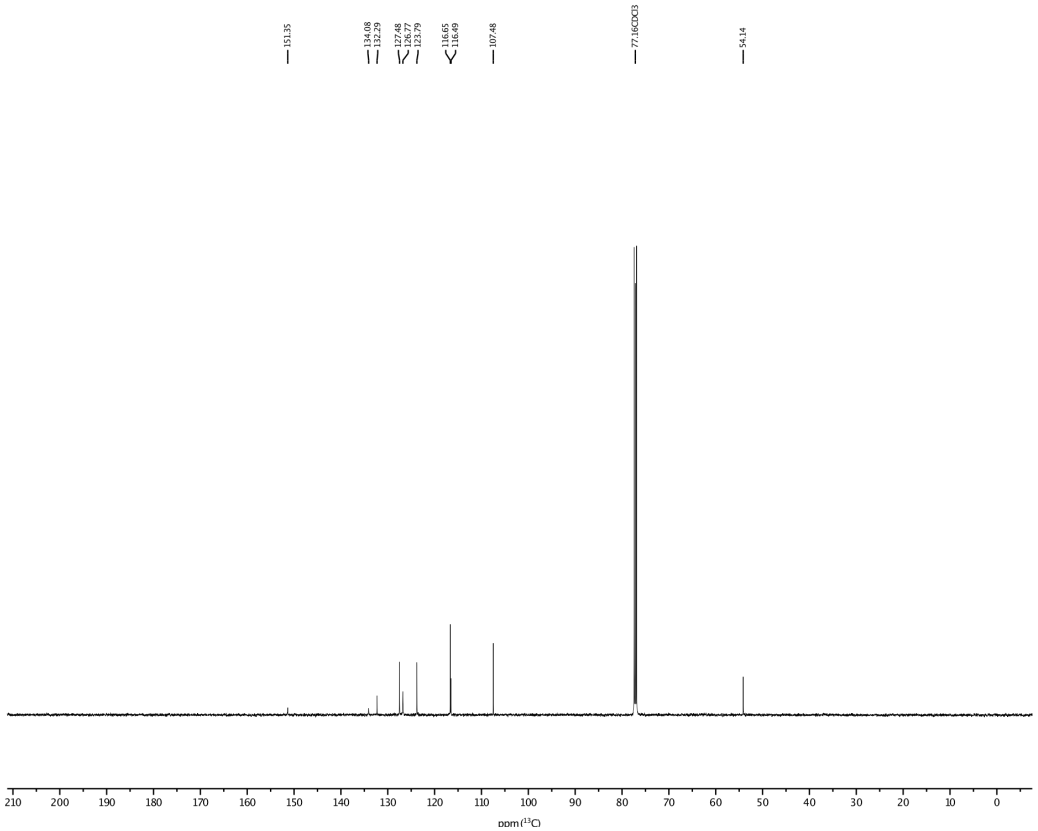
^{13}C NMR (126 MHz, d_6-DMSO)



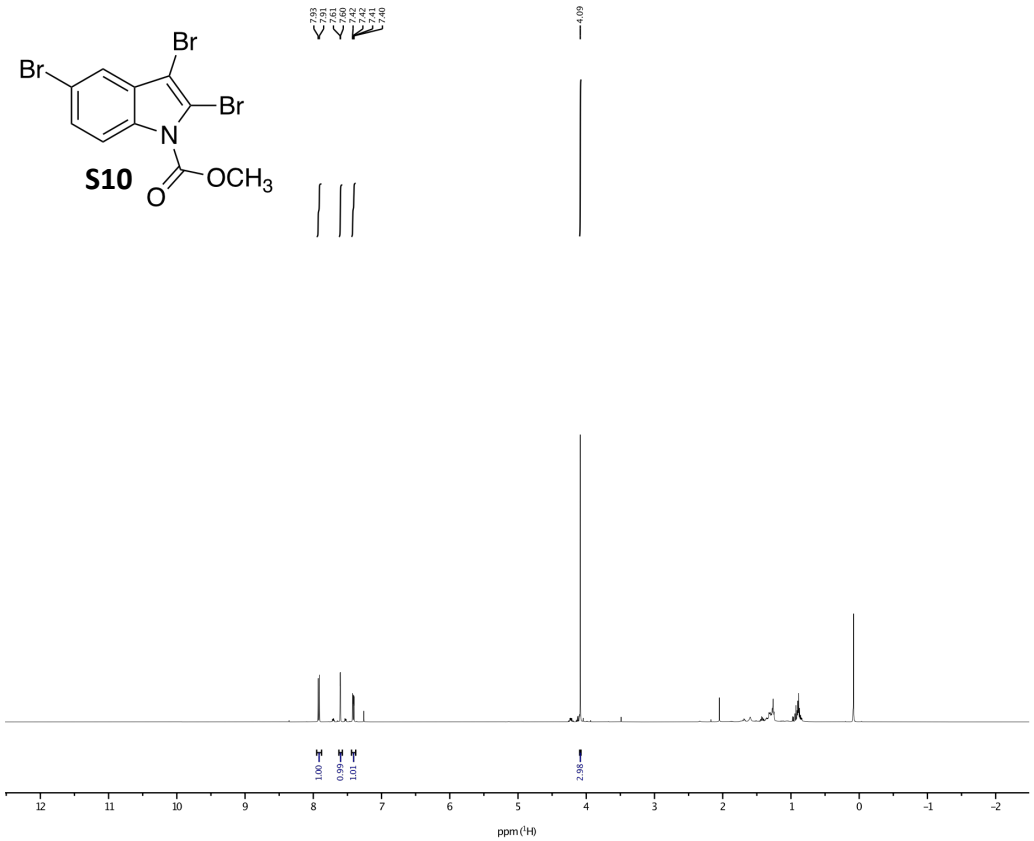
^{1}H NMR (500 MHz, $CDCl_3$)



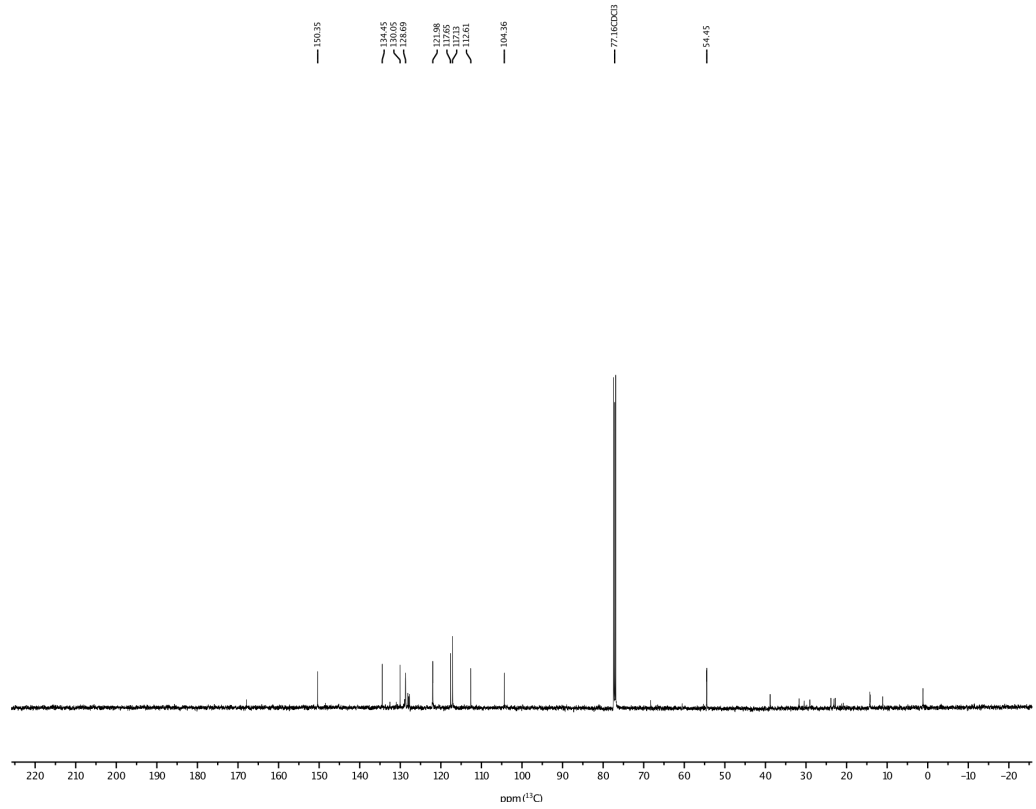
^{13}C NMR (126 MHz, $CDCl_3$)



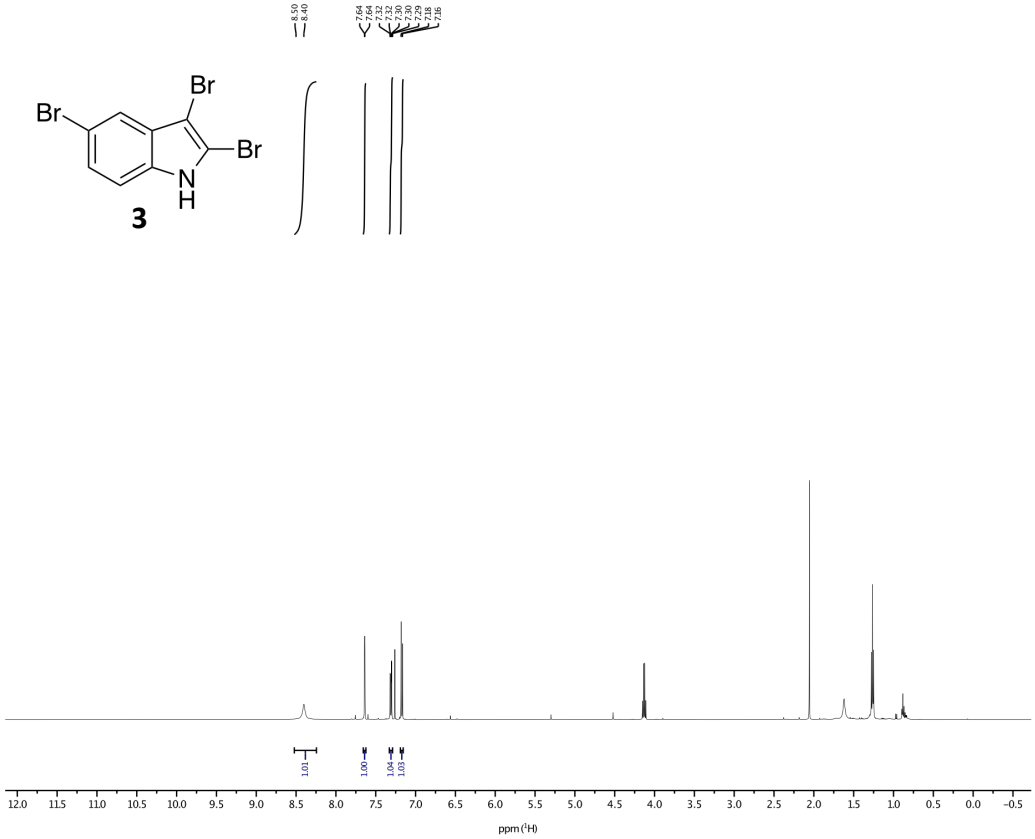
^{1}H NMR (500 MHz, $CDCl_3$)



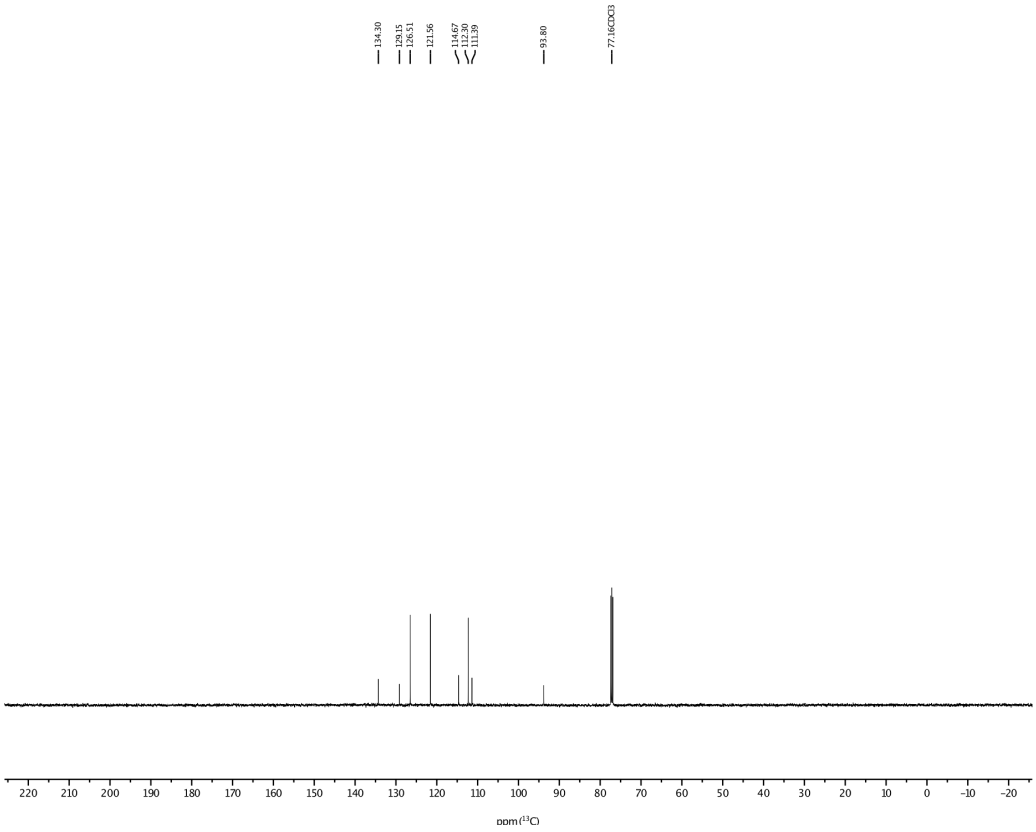
^{13}C NMR (126 MHz, $CDCl_3$)



^{1}H NMR (500 MHz, $CDCl_3$)



^{13}C NMR (126 MHz, $CDCl_3$)



References

(1) Eichhorn, E.; Van der Ploeg, J. R.; Leisinger, T. *J. Biol. Chem.* **1999**, *274*, 26639-26646
(2) Breinlinger, S.; Philipps, T. J.; Haram, B. N.; Mares, J.; Martinez Yerena, J. A.; Hrouzek, P.; Sobotka, R.; Henderson, W. M.; Schmieder, P.; Williams, S. M.; Lauderdale, J. D.; Wilde, H. D.; Gerrin, W.; Kust, A.; Washington, J. W.; Wagner, C.; Geier, B.; Liebeke, M.; Enke, H.; Niedermeyer, T. H. J., Wilde, S. B. *Science* **2021**, *371*, eaax9050
(3) Huijbers, M. M. E.; Montersino, S.; Westphal, A. H.; Tischler, D.; Van Berkel, W. J. H. *Arch. Biochem. Biophys.* **2014**, *544*, 2-17;
(4) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. *Biochim. Biophys. Acta.* **2015**, *1854*, 1019-1037.
(5) Fisher, B. F.; Snodgrass, H. M.; Jones, K. A.; Andorfer, M. C.; Lewis, J. C. *ACS Cent. Sci.* **2019**, *5*, 1844-1856.
(6) Ricardo, M. G.; Schwark, M.; Llanes, D.; Niedermeyer, T. H. J.; Westermann, B. *Chem. Eur. J.* **2021***, 27*, 12032-12035.
(7) Tomakinian, T.; Kouklovsky, C.; Vincent, G. *Synlett* **2015**, *26***,** 1269**.**
(8) Ueno, A.; Kayaki, Y.; Ikariya, T. *Organometallics* **2014**, *33*, 4479–4485
(9) Parrick, J.; Yahya, A.;Ijaz, A. S., Yizun, J. *J. Chem. Soc., Perkin Trans 1*, **1989**, 2009-2015.
(10) Degnan, A. P.; Tora, G. O.; Denhart, D. J.; Vrudhula, V. M.; Macor, J. E., Bronson, J. J. PCT Int. Appl. 2009009411, 15 Jan 2009.