Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

The Single-Component Flavin Reductase/Flavin-Dependent Halogenase AetF is a Versatile Catalyst for Selective Bromination and Iodination of Arenes and Olefins

*Y. Jiang, H. M. Snodgrass, Y. S. Zubi, C. V. Roof, Y. Guan, D. Mondal, N. H. Honeycutt, J. W. Lee, R. D. Lewis, C. A. Martinez, J. C. Lewis**

I. General

A) Materials

Greiner Bio-One polypropylene 96-well V-bottom plates (product number 651201) were purchased from Fisher Scientific. Agilent 0.2 µm PVDF 96-well filter plates (product number 203980-100) were purchased from Agilent. NADP, FAD, and antibiotics were purchased from Chem-Impex International Inc. (Wood Dale, IL). Substrates were purchased from Sigma-Aldrich, Chem-Impex, AK Scientific, Enamine, and Santa Cruz Biotechnologies. A gene encoding AetF was purchased as an *E. coli* codon-optimized construct in pET28(a) from Twist Biosciences as described below. GDH-105 (hereafter, GDH; 50 U/mg) was obtained from Codexis, Inc. (Redwood City, CA). Catalase from bovine liver was obtained from Millipore Sigma (2,000-5,000 U/mg; stock solutions were prepared assuming 2,000 U/mg; product number C9322). L-glutathione reduced was obtained from Millipore Sigma. Luria broth (LB) and terrific broth (TB) media were purchased from Research Products International (Mt. Prospect, IL). Qiagen Miniprep Kits were purchased from QIAGEN Inc. (Valencia, CA) and used according to the manufacturer's instructions. Protein ladder (Blue Pre-stained Protein Standard, Broad Range (11-190 kDa); product number P7706) was purchased from New England Biolabs (Ipswitch, MA).

Antibiotics were prepared as 1000x stock solutions: 1000x chloramphenicol was prepared at 25 mg/mL in EtOH, and 1000x kanamycin was prepared at 50 mg/mL. Stock solutions of 10 mM NADP and 10 mM FAD were prepared in 25 mM HEPES, pH 7.4 (reaction buffer) and stored at -20 °C until use. Stock solutions of 1.5 M NaBr, and 1 M glucose were prepared in reaction buffer and stored at 4 °C until use. Stock solutions of substrate were prepared at 100 mM in DMSO or in IPA. GDH was prepared as 180 U/mL stock solution in reaction buffer immediately before reaction setup. Catalase and glutathione stock solutions were prepared immediately before reaction setup.

B) Instruments

Ultra-high pressure liquid chromatography (UHPLC) was performed using one of two systems. Studies of the AetF substrate scope were performed on an Agilent 1200 series system with a 1290 Infinity II high-speed pump, a 1260 Infinity II diode array detector, and a 1290 Infinity II multisampler with single-needle configuration connected to a 6135X single quadrupole mass spectrometer with an Agilent Jet Stream ESI source. All achiral UPLC analysis was performed using an Eclipse Plus C18 2.1x5 mm guard column with a 1.8 µm pore size (part number 821725-901) connected to a ZORBAX rapid resolution C18 column (part number 959757-902) with mobile phase A of water with 0.1% FA and mobile phase B of acetonitrile with 0.1% FA.

The second system was an Acquity Waters UPLC equipped with a Waters BEH C18 column (2.1 × 50 mm, 1.7 µm, part no. 186002350) maintained at 30 °C, using 0.1% trifluoracetic acid in water (mobile phase A) and acetonitrile (mobile phase B) at a flow rate of 0.6 mL min^{-1}. For the iodination experiments, the m/z values of the iodinated products were verified using the same system but equipped with an SQ Detector/Mass spectrometer with the cone voltage set to 30V and the capillary voltage set to 4.5 kV.

Reverse phase purification was performed using either or both: 1.) a Biotage Isolera One with 12 g SNAP-KP-C18-HS columns using 0.1% FA (formic acid) in water as mobile phase A and 0.1% FA in methanol as mobile phase B or 2.) An Agilent 1100 HPLC equipped with a Supelco Discovery C18 semipreparative column (25 cm x 10 mm, 5 µm particle size) and an Agilent 1260 Infinity II fraction collector using 0.1% FA in water as mobile phase A and 0.1% FA in acetonitrile as mobile phase B.

C) Chromatography (UHPLC, UHPLC-MS, SFC)

Biotage method 1: Used to purify quinoline compound **13**:

Column Volumes	%A	%B
1	90	10
15	0	100
20	0	100

Biotage method 2: Used to purify compound **4a**, **7a** and **13a**:

Column Volumes	%A	%B
5	90	10
15	70	30
30	40	60
10	0	100
10	0	100

Biotage method 3: Used to purify compound **1a** and **1b**:

Column Volumes	%A	%B
10	90	10
30	70	30
5	0	100
10	0	100

UHPLC-MS Method 1: Used to analyze products **1a**, **2a**, **4a**, **6a**, **7a**, **13a**, **1b,** and **6b**.
Column: Agilent Eclipse Plus C18 column (3.5 µm particle size; 4.6 x 150 mm)
Mobile phase A: Water with 0.1% TFA
Mobile phase B: ACN with 0.1% TFA
Method:

Time (min)	%A	%B
0	95	5
0.5	95	5
8	20	80
8.01	5	95
8.25	5	95
8.5	95	5

UHPLC-MS Method 2: Used to analyze product **3a**, **5a**, **8a**, **9a**, **10a**, **11a**, **12a**, and **11b**.
Column: Agilent Eclipse Plus C18 column (3.5 µm particle size; 4.6 x 150 mm)
Mobile phase A: Water with 0.1% TFA
Mobile phase B: ACN with 0.1% TFA
Method:

Time (min)	%A	%B
0	80	20
0.5	80	20
8	20	80
8.01	5	95
8.25	5	95
8.5	80	20

UHPLC-MS Method 3: Used to analyze products **14a** and **15a** during substrate scope studies.
Column: Agilent Eclipse Plus C18 column (3.5 µm particle size; 4.6 x 150 mm)
Mobile phase A: Water with 0.1% TFA
Mobile phase B: ACN with 0.1% TFA
Method:

Time (min)	%A	%B
0	60	40
0.5	60	40
4	30	70
4.01	5	95
4.25	5	95
4.5	60	40

Semi-preparative HPLC Method 1: used to purify **2a** from bioconversions using AetF as the catalyst using the HPLC system described in the general methods

Time (min)	%A	%B
0	70	30
13	63	37
14	5	95
22	5	95

Semi-preparative HPLC Method 2: used to purify product **6a**

Time (min)	%A	%B
0	72	28
15	65	35
16	5	95
23	5	95

Semi-preparative HPLC Method 3: used to purify product **9a**

Time (min)	%A	%B
0	65	35
10	57	43
11	5	95
18	5	95

Semi-preparative HPLC Method 4: used to purify product **10a**

Time (min)	%A	%B
0	70	30
11	59	41
12	5	95
18	5	95

Semi-preparative HPLC Method 5: used to purify product **6b**

Time (min)	%A	%B
0	71	29
12	65	35
13	5	95
18	5	95

UHPLC Method 1: Used for AetF optimization studies with 6-fluorotryptamine and L-Tryptophan on UHPLC system #2.
Column: ACQUITY UPLC BEH C18 Column, 130 Å, 1.7 µm, 2.1 x 50 mm (part number 186002350) with a flow rate of 0.6 mL/min.
Mobile phase A: Water with 0.1% TFA
Mobile phase B: Acetonitrile
Gradient:

Time (min)	%A	%B
0	95	5
0.25	95	5
3	50	50
3.01	5	95
3.5	5	95
3.51	95	5

Chiral HPLC method 1: Used to determine the enantioselectivity of the halocyclized product **14a** and **15a** from the crude reaction mixture for e.r. determination.
Column: Phenomenex Lux 3 µm Cellulose-1, LC Column 250 x 4.6 mm
Method: Hexanes/ IPA= 90/10, flow rate: 1 mL/min

SFC-MS Method 1: Used for the chiral separation of **12a** crude reaction mixture for e.r. determination
Column: CHIRALPAK IC-U 3.0 x 50 mm, 1.6 µm particle size (part number 84U82) held to 40 °C with a flow rate of 1.7 mL/min.
Mobile phase A: CO_2
Mobile phase B: Methanol
Gradient:

Time (min)	%A	%B
0	95	5
2.65	55	45
2.7	95	5
3.2	95	5

SFC-MS Method 2: Used for the chiral separation of **13a** from the crude reaction mixture for e.r. determination
Column: CHIRALPAK IG-U 3.0 x 50 mm, 1.6 µm particle size (part number 87U82) held to 40 °C with a flow rate of 1.7 mL/min.
Mobile phase A: CO_2
Mobile phase B: Methanol with 0.1% isopropylamine
Gradient:

Time (min)	%A	%B
0	95	5
2.65	55	45
2.7	95	5
3.2	95	5

II. Enzyme Preparation

A) Molecular Cloning

Standard molecular biology protocols were followed. Oligonucleotides were purchased from either Sigma-Aldrich or Integrated DNA technologies as lyophilized powder and prepared according to the manufacturer's specification. All primer sequence and their use in this project are included in a table below. AetF was purchased as an *E. coli* codon-optimized gene construct in pET28(a) from Twist Biosciences with the 6x-his tag placed at the N-terminus of the protein embedded in a flexible linker.

Full protein sequence for AetF used in this study. The methionine that starts translation of AetF is in bold, and the putative catalytic lysine is highlighted red:

MGSSHHHHHHSSGLVPRGSH**M**LEVCIIGFGFSAIPLVRELARTQTEFQIISAESGSVWDRLSES
GRLDFSLVSSFQTSFYSFDLVRDYEKDYYPTAKQFYEMHERWRSVYEEKIIRDFVTKIENFKDY
SLISTRSGKTYEAKHVVLATGFDRLMNTFLSNFDNHVSNKTFVFDTMGDSANLLIAKLIPNNNKII
LRTNGFTALDQEVQVLGKPFTLDQLESPNFRYVSSELYDRLMMSPVYPRTVNPAVSYNQFPLI
RRDFSWVDSKSSPPNGLIAIKYWPIDQYYYHFNDDLENYISKGYLLNDIAMWLHTGKVILVPSDT
PINFDKKTITYAGIERSFHQYVKGDAEQPRLPTILINGETPFEYLYRDTFMGVIPQRLNNIYFLGYT

RPFTGGLANITEMQSLFIHKLITQPQFHQKIHQNLSKRITAYNQHYYGAAKPRKHDHTVPFGFYT
EDIARLIGIHYQPNECRSVRDLLFYYAFPNNAFKYRLKGEYAVDGVDELIQKVNDKHDHYAQVF
VQALSIRNMNSDEAAEWDHSARRFSFNDMRHKEGYRAFLDTYLKAYRQVENISVDDTVVDEE
WNFMVKEACQVRDKVAPNIEEKTHYSKDEDVNKGIRLILSILDSDISSLPDSNGSRGSGNLKEG
DRLCKFEAQSIEFIRRLLQPKNYELLFIRESTVSPGSHRHGETA

The K258A mutation was introduced using SOE-PCR[1] using the following conditions:

Fragment 1 PCR:

PCR mix:

Component	Volume (uL)	Concentration
Forward primer (HSPF016)	2.5	500 nM
Reverse Primer (HSPF-023)	2.5	500 nM
Template (AetF)	1	50 ng
Water	19	NA
Primestar 2x mix	25	NA

PCR Cycle:

Temperature (°C)	Length of cycle (min:sec)	Repetitions
98	5:00	1
98	0:10	28
60	0:05	
72	0:30	
72	5:00	1
4	hold	1

Fragment 2 PCR:

PCR Mix:

Component	Volume (uL)	Concentration
Forward primer (HSPF022)	2.5	500 nM
Reverse Primer (HSPF-020)	2.5	500 nM
Template (AetF)	1	50 ng
Water	19	NA
Primestar 2x mix	25	NA

PCR cycle:

Temperature (°C)	Length of cycle (min:sec)	Repetitions
98	5:00	1
98	0:10	28
60	0:05	
72	0:30	
72	5:00	1
4	hold	1

Vector PCR:

PCR Mix:

Component	Volume (uL)	Concentration
Forward primer (HSPF021)	2.5	500 nM
Reverse Primer (HSPF-017)	2.5	500 nM
Template (AetF)	1	50 ng
Water	19	NA
Primestar 2x mix	25	NA

PCR cycle:

Temperature (°C)	Length of cycle (min:sec)	Repetitions
98	5:00	1
98	0:10	28
60	0:05	
72	1:30	
72	5:00	1
4	hold	1

After the F1, F2, and vector PCRs, the resulting amplicons were purified via gel extraction using the Qiagen gel extraction kit according to the manufacturers protocol. Samples were eluted off the spin column with 30 µL of molecular-grade water. To each tube, 3 µL of NEB Cutsmart buffer (NEB#B6004S) and 1 µL of DPN1 enzyme (NEB#R0176L) was added to each reaction and allowed to incubate at 37 °C for 1 hour. After digestion of the template DNA, the reaction mixture was purified using the Qiagen PCR purification kit according to the manufacturer's protocol, and

eluted with 10 µL of molecular grade water. After purification of the F1 and F2 fragment, the assembly PCR was performed to generate the AetF(K258A) insert.

Assembly PCR:

PCR mix:

Component	Volume (uL)	Concentration
Forward primer (HSPF016)	2.5	500 nM
Reverse Primer (HSPF-020)	2.5	500 nM
Template (1:1 mixture of F1:F2)	1	50 ng (total)
Water	19	NA
Primestar 2x mix	25	NA

PCR cycles:

Temperature (°C)	Length of cycle (min:sec)	Repetitions
98	5:00	1
98	0:10	28
60	0:05	
72	1:00	
72	5:00	1
4	hold	1

The resulting PCR product was purified via gel extraction according to the manufacturer's protocol. Concentrations of the purified DNA products was measured using a Thermo NanoDrop 8000 by measuring the absorbance at 260 nm. Ligation of the assembled AetF(K258A) fragment into pET28(a) was done using Gibson Assembly. The Gibson assembly was performed as follows:

Component	Volume (uL)	Amount (ng, pmol)
Insert (AetF[K258A])	1.5	150 (0.15)
Vector (pET28a)	1	100 (0.05)
NEB 2x Gibson Mix	2.5	n/a

This reaction mixture was allowed to incubate for thirty minutes at 50 °C. After the reaction was complete, 2 µL of the assembly mixture was used directly to transform *E. coli* BL21(λDE3) cells Agilent, part number 230132) via electroporation using the EC1 program. Electroporated cells were recovered using 750 µL (Expression recovery medium, Biosearch Technologies Item ID 80030-1) and transferred to a 12 mL culture tube, then incubated at 37 °C for one hour with 210 rpm in a Kuhner Shaker. After this recovery period, selection was performed by plating the culture

onto agar containing kanamycin and the plate was incubated at 37 °C overnight. The next day, 5 colonies were picked into 10 mL of LB media supplemented with kanamycin and allowed to grow overnight at 37 °C with 210 rpm. After allowing the cultures to grow overnight, they were removed from the incubator and plasmid was isolated following the Qiagen miniprep protocol. The K258A mutation was verified by Sangar sequencing performed by Functional Biosciences (Madison, WI).

List of primers used in this study:

Primer name	Sequence (5'->3')	Function	Dir.
HSPF-007	TACGGCATTAGATCAGGAAG	Sequencing	F
HSPF-016	CCGCGCGGCAGCCATATGTTGGAAGTCTGTATCATCG	SOE Fragment 1	F
HSPF-017	GATACAGACTTCCAACATATGGCTGCCGCGCGGCAC	SOE Vector	R
HSPF-020	CTCGAGTCATTAAGCGGTTTCTCCATGGC	SOE Fragment 2	R
HSPF-021	GCCATGGAGAAACCGCTTAATGACTCGAG	SOE vector	F
HSPF-022	ATGGACTGATCGCTATTGCGTACTGGCCCATTGACCAG	SOE Fragment 2	F
HSPF-023	CTGGTCAATGGGCCAGTACGCAATAGCGATCAGTCCATTAG	SOE Fragment 1	R

B) Gene Expression

After verification of the correctly assembled sequences, AetF constructs harbored within pET28 were transformed into BL21 Gold (DE3) competent cells (Agilent 230132) with the pGro7 chaperone plasmid (Takara Bio, catalog # 3340) and electroporation and recovery were done as described above. Recovered media was plated onto LB agar plates with kanamycin and chloramphenicol for selection and incubated at 37 °C overnight. The next day, single colonies of *E. coli* were picked into 12 mL culture tubes containing 10 mL of LB media supplemented with the appropriate antibiotics. After incubation for 16 hours at 37 °C in a vertical incubator set to 210 rpm, the cultures were removed from the incubator and 400 µL of each was diluted with 400 µL of 50% glycerol to create a glycerol stock. The remaining culture was used to inoculate 1 L of TB media in a 2.8 L Erlenmeyer flask supplemented with the appropriate antibiotics, at which point the flask was placed into a horizontal incubator and the culture was allowed to incubate at 37 °C for 2.5 hours. At this point, the culture was removed from the incubator and placed into an ice bath for 15 minutes while the incubator was cooled to 18 °C. Gene expression was induced with 100 µL of a 1 M stock of IPTG (final concentration 0.1 mM) and the addition of 2 g of solid L-arabinose directly into the media (final concentration 2 mg/mL). The culture flask was incubated at 18 °C for 24 hours at 210 rpm. After this point in time, the culture flask was removed from the incubator and pelleted in a centrifuge at 15,000 x g for 45 minutes with the centrifuge held at 4 °C. The supernatant was removed from the cell pellet, and the cell pellet was stored at either –20 °C or –80 °C until further use.

For expression in microtiter plates, colonies from the LB agar plate were picked into a 2 mL plate charged with 500 µL of LB media supplemented with the appropriate antibiotics. These cultures were allowed to grow overnight in a Glas-Col incubator set to 37 °C and 900 rpm. The next day, 20 µL of the overnight culture was used to inoculate 800 µL of TB media charged in a 2 mL deep well plate supplemented with the appropriate antibiotics. After inoculation, these cultures were placed into a Glas-Col incubator set to incubate at 37 °C and 900 rpm for 2.5 hours, at which point the culture plate was removed and the incubator was allowed to cool to 18 °C. Once the incubator cooling was complete, the culture plate was set to incubate for 24 hours at 18 °C and 900 rpm. After incubation was complete, the culture plate was placed into a centrifuge and centrifugation was done at 2000 x g and 20 minutes while held at 4 C. After centrifugation was complete, the supernatant was removed from the culture and the cell pellets were stored at –80 °C until use.

C) Protein Preparation

For preparation of clarified cell lysate from microtiter plates, the cell pellets were resuspended in 100 µL of lysis buffer (0.75 mg/mL lysozyme, 1 uL/mL benzonase I in the appropriate buffer for final conditions) in a Glas-col incubator set to 37 °C and 1200 rpm for 15 minutes. After resuspension was complete, the incubation speed was reduced to 900 rpm and the lysis was allowed to continue for an additional 45 minutes. Once cell lysis was complete, the 2-mL deep well plate was placed into the –80 °C freezer until the cell lysate was thoroughly frozen through, at least 1 hour. After the cell lysate was frozen, it was removed from the –80 °C freezer and set in ambient temperature for 10 minutes. To ensure even thawing of the entire plate, it was then placed into lukewarm water for an additional 20 minutes. Once the cell lysate was thawed, the lysate was clarified via centrifugation at 2000 x g for 20 minutes. The supernatent was then immediately aliquoted into a microtiter plate for bioconversions.

For cell pellets prepared from 1 L cultures, cell lysis was performed by either microfluidization (Microfluidics LM10) or sonication. For microfluidization, the cell pellet was resuspended in 100 mL of lysis buffer (0.75 mg/mL lysozyme, 1 mg/mL DNAse1) and incubated at 37 °C in a vertical incubator for 15 minutes to resuspend the pellet. Once resuspension was complete, the resuspended cells were subjected to microfluidization at 20,000 PSI in two passes through the system. The lysed cell was then clarified via centrifugation at 15,000 x g for 45 minutes, and the resulting clarified lysate was poured into a fresh tube for use. When whole cells were used, the cell pellet was resuspended in 100 mL of HEPES buffer (pH 7.4, 25 mM) via incubation in a vertical incubator set to 18C and 140 rpm for 30 minutes.

For sonication, cells were lysed using a QSonica S-4000 with a 0.5” horn at 40 W and a total processing time of 5 minutes with 1 minute on/off cycles. Cell lysate was clarified at 15,000 rpm for 30 minutes, and the resulting clarified lysate was transferred to a fresh 50 mL centrifuge tube and added to pre-equilibrated Ni-NTA (equilibration buffer: 20 mM phosphate, 300 mM NaCl, 10 mM imidazole pH 7.4). Clarified lysate was incubated with resin for approximately 30 minutes at which point it was transferred to uncapped spin columns and the lysate was allowed to flow through. The resin was washed with at least 5 CV wash buffer (20 mM phosphate, 300 mM NaCl, 25 mM imidazole pH 7.4), at which point the spin column was transferred to a new centrifuge tube and the resin was washed with elution buffer (20 mM phosphate, 300 mM NaCl, 250 mM imidazole, pH 7.4). Eluted protein was concentrated via diafiltration using Amicon spin filters Ultra 30K MWCO spin filters and the buffer was exchanged for storage buffer (25 mM HEPES and 10%

glycerol, pH 7.4). For long term storage, proteins were immediately frozen in liquid nitrogen and stored at – 80 °C until use. AetF concentration was determined using the Pierce BCA Protein Assay Kit.

III. Bioconversion Protocol and Analysis

A) Biocatalysis Conditions

Analytical-scale AetF reactions

Purified AetF was used for initial evaluation of the substrate scope and reactions were performed in triplicate. A solution of 125 µM AetF in storage buffer (HEPES pH 7.4, 25 mM with 10% glycerol) was thawed on ice. A stock solution of AetF was prepared at 37.5 uM, and 50 µL of this solution was aliquoted into a microtiter plate. A solution of 500 µM FAD, 500 µM NADP, 100 mM D-glucose, 50 mM NaBr or NaI or 500 mM NaCl, and 5 mM of substrate was prepared in reaction buffer (small molecule stock solution). A solution of 67.5 U/mL GDH and 262.5 U/mL catalase was prepared in reaction buffer (enzyme stock solution). 15 µL of the small molecule stock solution was added to the 96 well microtiter plates containing clarified lysate. 10 µL of the enzyme stock solution was then added to give a total reaction volume of 75 µL and a final AetF concentration of 25 µM. The plate was sealed using an aluminum heat seal and incubated in ambient atmosphere using a plate shaker set to 600 rpm. The next day, reactions were quenched using 1 volume of methanol, and the quenched reactions were further incubated for 15 minutes to facilitate full protein precipitation. Precipitated protein was removed from the solution via centrifugation at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes. From the resulting supernatant, 50 µL was then added to a 0.2 µM filter plate charged with 150 µL of water from a Milli-Q filtration system, and filtration was done gravimetrically at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes into a fresh microtiter plate. The microtiter plate containing filtered reactions was then heat-sealed using an aluminum heat-seal and analyzed via UHPLC-(MS) using the appropriate methods as given in the methods section of this SI. Conversion is defined as % AUC of the product(s) as determined by m/z value compared to the % AUC of the total area of starting material + product(s). Full characterization of all brominated and iodinated products presented in the manuscript is provided below.

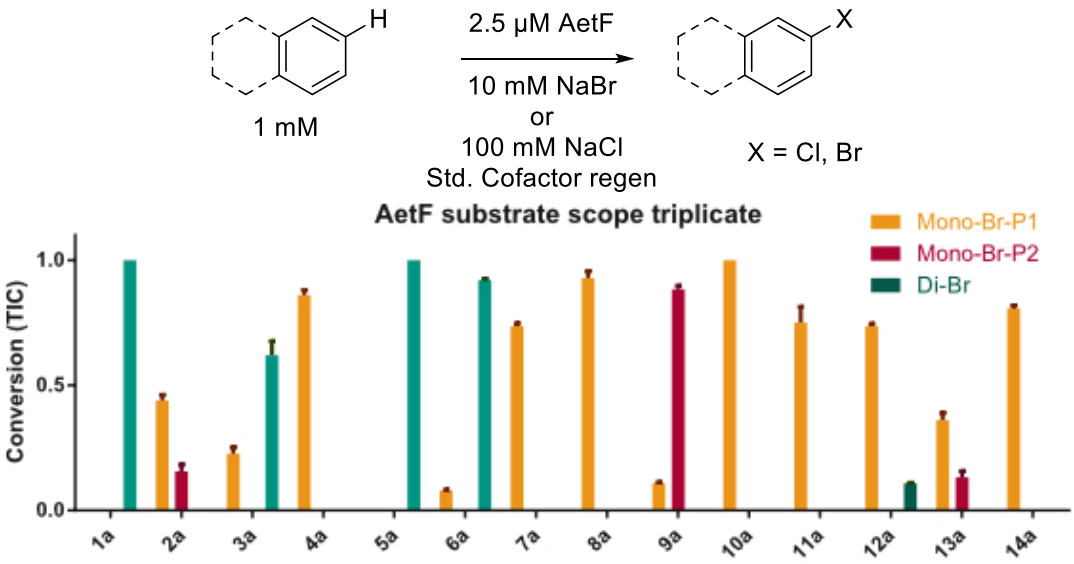


Figure S1. Conversion data used to evaluate AetF substrate scope. Conversion was evaluated as the AUC of the brominated product(s) on the TIC over the total AUC of substrate and all products.

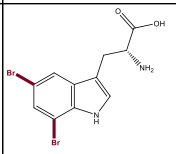
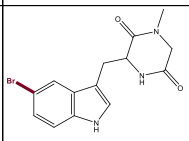
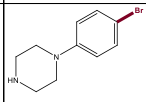
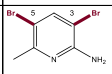
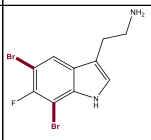
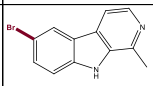
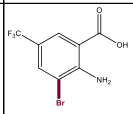
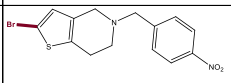
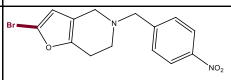
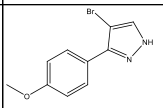
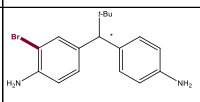
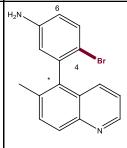
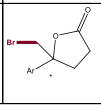
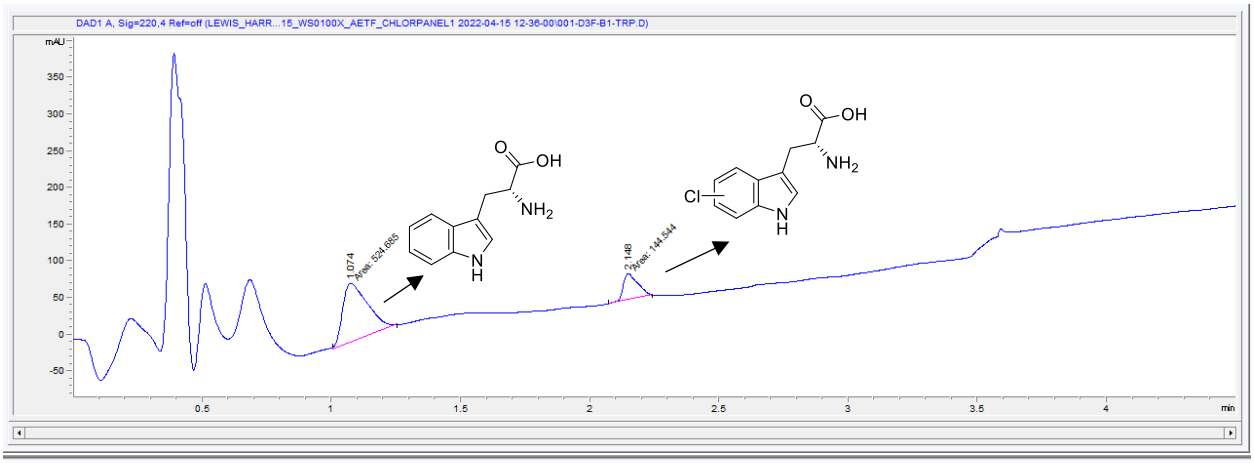
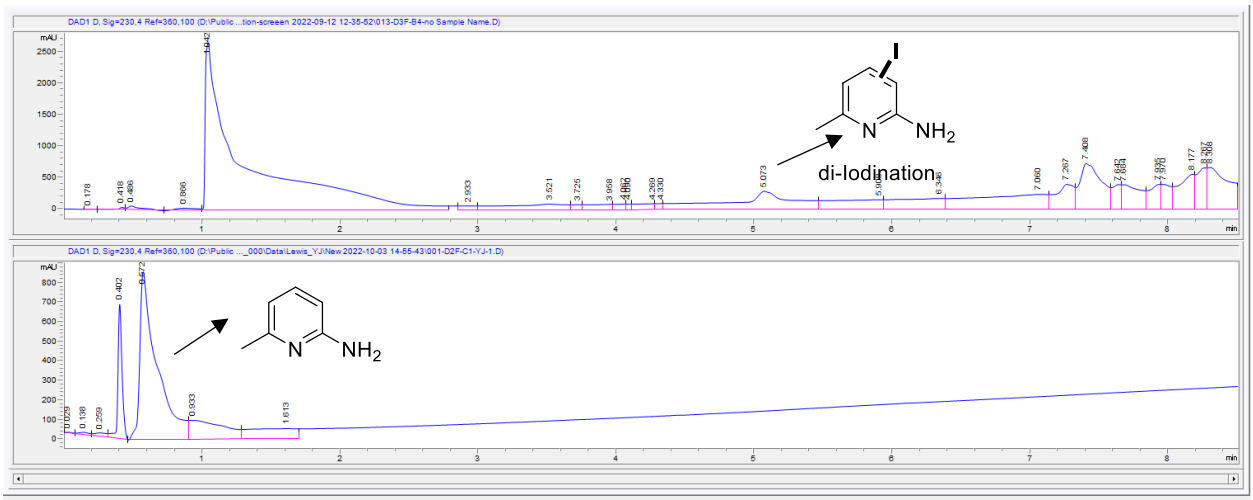
Name	structure	Mono-Br-P1			Mono-Br-P2			Di-Br		
1a		0	0	0	0	0	0	1	1	1
2a		0.463458	0.43486	0.41934	0	0	0	0.147059	0.134282	0.18672
3a		0.256637	0.210526	0.212955	0	0	0	0.556718	0.645559	0.658385
4a		0.845746	0.883801	0.850575	0	0	0	0	0	0
5a		0	0	0	0	0	0	1	1	1
6a		0.077953	0.083724	0.07549	0	0	0	0.922047	0.916276	0.92451
7a		0.723711	0.733716	0.74812	0	0	0	0	0	0
89a		0.93266	0.898734	0.953668	0	0	0	0	0	0
9a		0.113823	0.104082	0.106656	0.886177	0.895918	0.868083	0	0	0
10a		1	1	1	0	0	0	0	0	0
11a		0.821711	0.726323	0.704116	0	0	0	0	0	0
12a		0.737542	0.723246	0.744746	0	0	0	0.108154	0.106396	0.110265
13a		0.332467	0.356144	0.392274	0.158618	0.129219	0.109916	0	0	0
14a		0.818509	0.810115	0.796813	0	0	0	0	0	0

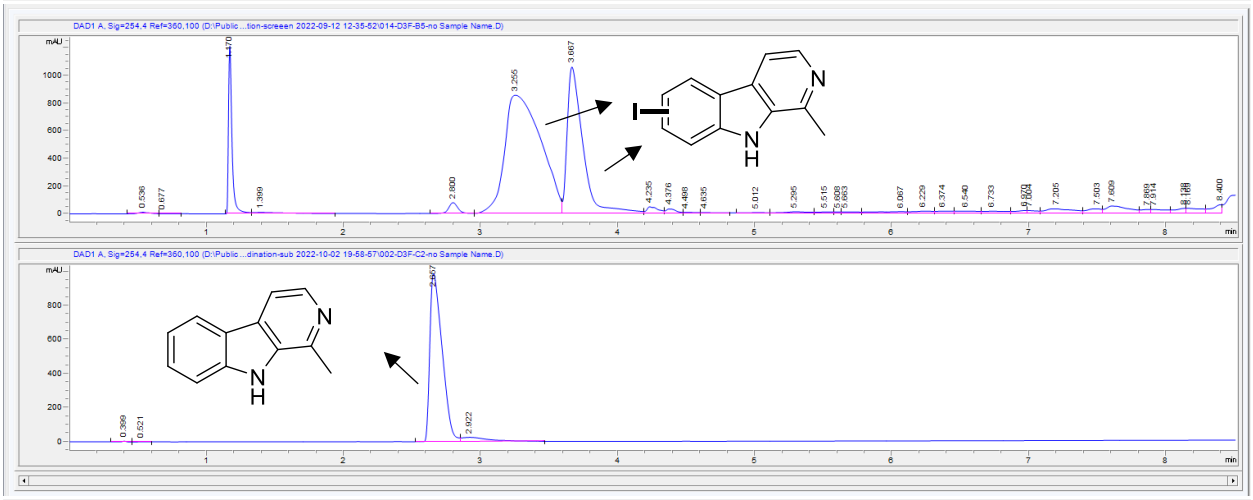
Table S1. Conversion data from Figure S1. Data from each individual replicate is represented.



#	Time	Area	Height	Width	Area%	Symmetry
1	1.074	524.7	80.2	0.109	78.401	0.505
2	2.148	144.5	36.6	0.0659	21.599	0.528

Figure S2. LC/MS chromatogram for chlorination of tryptophan by AetF. Integration was performed at 220 nm and the relative % area is shown. Chlorination was not observed for any other substrates in Figure S1.


di-Iodination



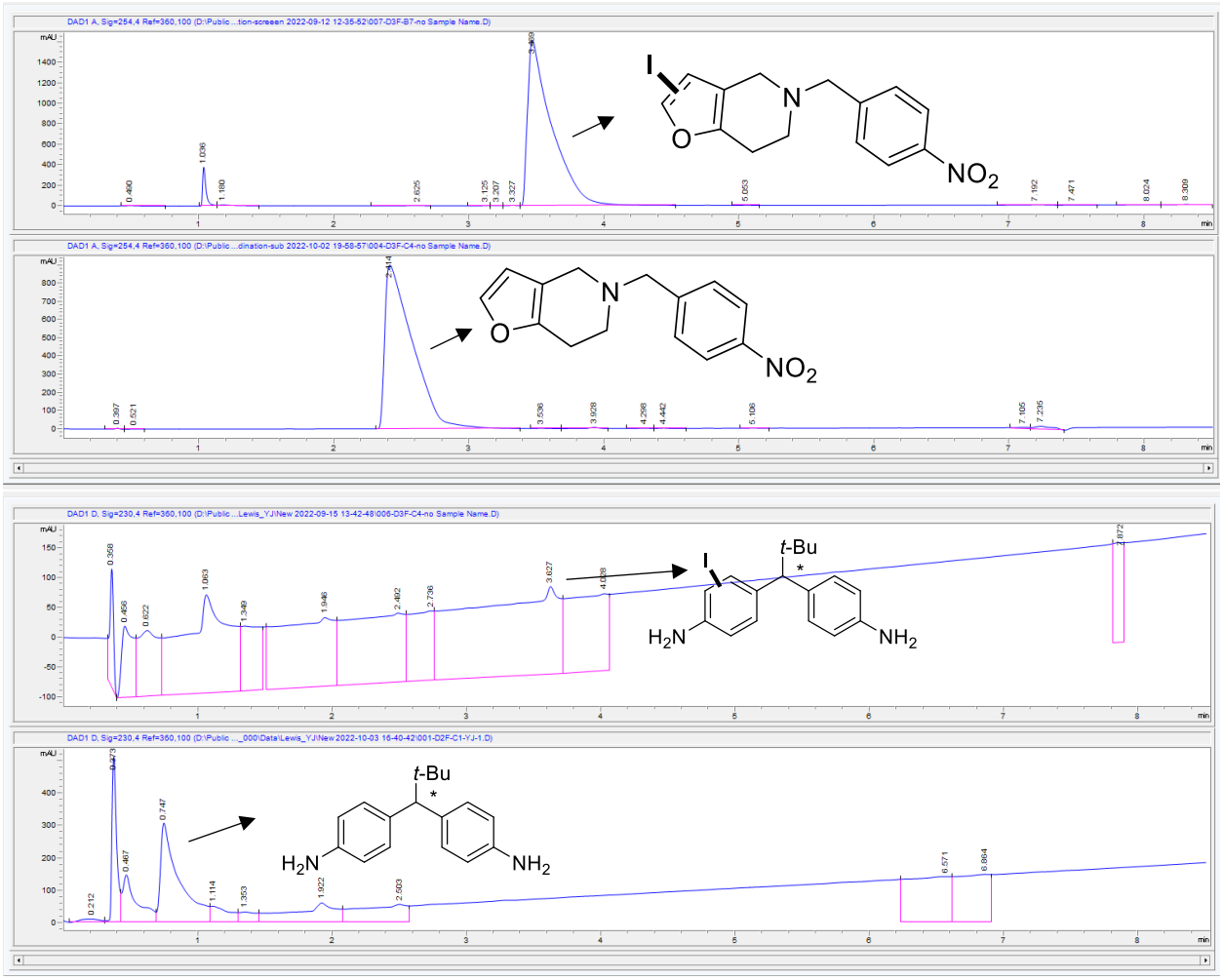


Figure S3. LC/MS chromatograms showing iodination of substrates not discussed in the manuscript by AetF.

Preparative-scale AetF reactions

20 mL and 30 mL bioconversions were conducted in 300 mL plastic beakers (Fisherbrand™ Polypropylene Disposable Beaker, catalog No. FB012915) covered with a breathable plate seal. Substrates were prepared as 30 mM stock solutions in DMSO. The beakers were charged with 25 mM Tris-HCl, pH 9.0, followed by NaBr (10 mM in reaction buffer, 10 equiv.), NADP (100 µM final concentration, in reaction buffer, 0.1 equiv.), FAD (100 µM final concentration in reaction buffer 0.1 equiv.), glucose (40 mM final concentration, prepared in reaction buffer, 40 equiv.), glutathione (1 mM final concentration, prepared in reaction buffer, 1 equiv.), catalase (70 U/mL final concentration), substrate (1 mM final concentration), AetF (25 µM final concentration), and reactions were initiated by the addition of GDH (9 U/mL final concentration). The breathable plate seal was placed over the beaker, which was promptly placed in a Thermo MaxQ 8000 incubator using a 250 mL Erlenmeyer pedestal. The beaker was set to incubate at 150 rpm, 25 °C overnight.

For all compounds besides L-tryptophan, the following workup was used. After overnight incubation, protein denaturation was affected by the addition of 10% v/v of 1 M HCl. Sodium

chloride was added to saturation, and the mixture was stirred vigorously for 15 minutes. The stir bar was removed, and the mixture was filtered through celite. The pH of the filtrate was adjusted to pH >10 with 1 M NaOH, and the mixture was extracted 3x with 50% v/v DCM. The organic layers were combined and dried over magnesium sulfate. Once dry, the solution was filtered and concentrated by rotary evaporation. The resulting solids were then dissolved in DCM, adsorbed onto celite, and purified by reverse phase biotage or semi-prep HPLC using the methods listed in the product characterization section of this SI. Only the major product from each bioconversion was isolated and characterized. Isolated yields and e.r. obtained from chiral chromatography are available in Figure 3 in the main text. Chiral chromatograms are available in the product characterization portion of this SI.

Purification of di-iodinated L-tryptophan from FDH bioconversions followed our previously reported method.[2] After overnight incubation, protein denaturation was affected by the addition of 10% v/v of 1 M HCl. Sodium chloride was added to saturation, and the mixture was stirred vigorously for 15 minutes. The stir bar was removed, and the mixture was filtered through celite. The aqueous filtrate was submitted to strong cation exchange to remove salts. For strong cation exchange,[2] DOWEX™ 50WX8 resin was slurry-packed with methanol in a 250 mL chromatography column. The resin was washed with ~300 mL of methanol and ~300 mL of deionized water. The resin was acidified with HCl (1 M) until the pH of flow through was less than 2. The resin was washed with deionized water until the pH was between 6-7. Acidic filtrate (pH < 2) was added to resin. The resin was washed with ~500 mL of deionized water. The product was eluted with NH_4OH (1 M) until product was no longer eluting from the column, as indicated by HPLC. Product-containing fractions were concentrated to dryness using a rotary evaporator under high vacuum. The residue was further purified by reverse phase Biotage using the methods listed in the product characterization section of this SI.

Analysis of the K258A mutation on bromination activity of AetF and discovery of iodination

During cell lysis, the lysis buffer used for verification of the catalytic lysine residue was HEPES, pH 7.4, 25 mM. 50 µL of the clarified lysate containing either wt-AetF or the K258A variant were aliquoted into a microtiter plate in triplicate. A solution of 500 µM FAD, 500 µM NADP, 100 mM D-glucose, 50 mM of NaBr or NaI and 5 mM of L-tryptophan was prepared in reaction buffer (small molecule stock solution). A solution of 67.5 U/mL GDH and 262.5 U/mL catalase was prepared in reaction buffer (enzyme stock solution). 15 µL of the small molecule stock solution was added to the 96 well microtiter plates containing clarified lysate. 10 µL of the enzyme stock solution was then added to give a total reaction volume of 75 µL. These reactions were incubated at 25 °C in a Thermo Scientific plate mixer (Eppendorf Thermomixer R) at 750 rpm for 20 hours. The next day, reactions were quenched using 1 volume of methanol, and the quenched reactions were further incubated for 15 minutes to facilitate full protein precipitation. 100 µL of this solution was transferred to a 1 mL culture plate (Waters, part # 186002481) and precipitated protein was removed from the solution via centrifugation at 2000 x g rpm in a Sorvall XT centrifuge for 15 minutes. From the resulting supernatant, 50 µL was then added to a 0.2 µM filter plate (OASIS, 186002794) charged with 150 µL of water from a Milli-Q filtration system, and filtration was done gravimetrically at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes into a fresh microtiter plate. The microtiter plate containing filtered reactions was then fitted with a mat (PP cap, part # 186002483) using an aluminum heat-seal and analyzed via UHPLC using method 1. Conversion

is defined as % AUC of the product(s) as determined by m/z value compared to the % AUC of the total area of starting material + product(s).

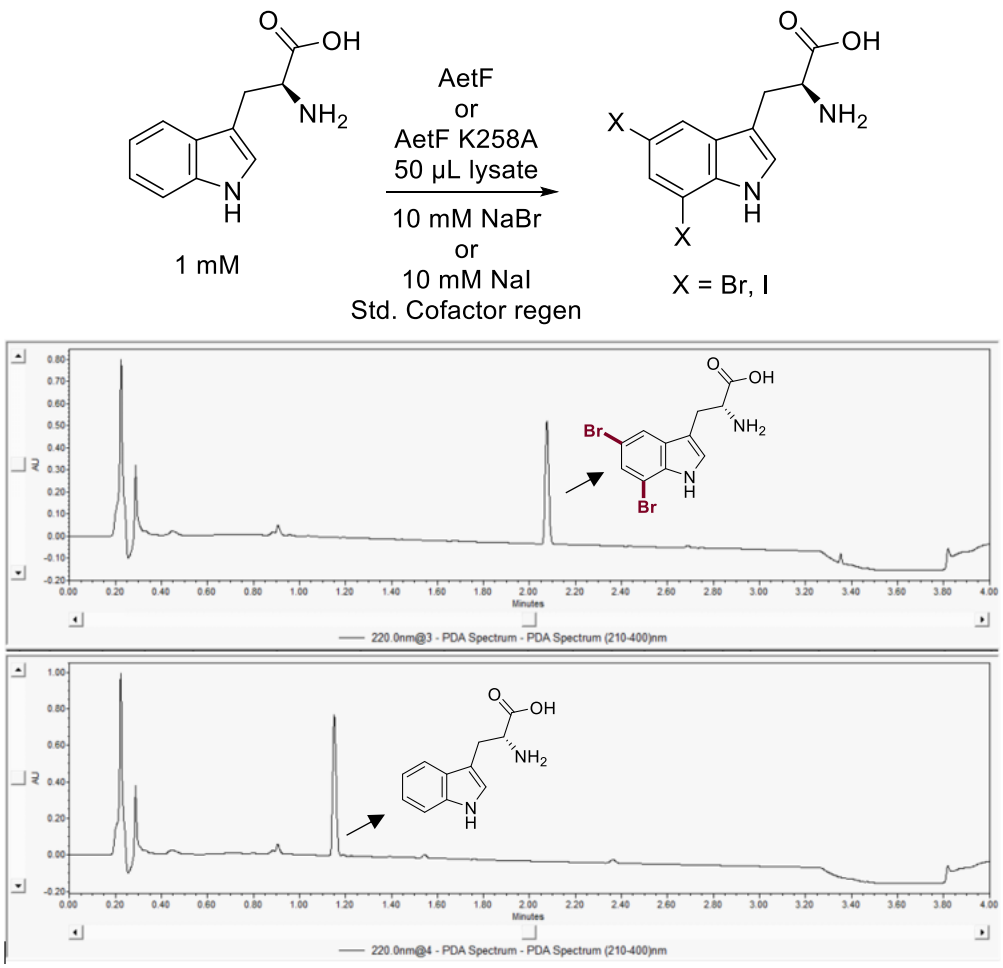


Figure S4. LC/MS chromatogram showing bromination activity using AetF and AetF K258A. Top: Wild-type AetF showing full conversion to dibrominated L-Tryptophan. Bottom: Chromatogram showing no activity with AetF K258A variant.

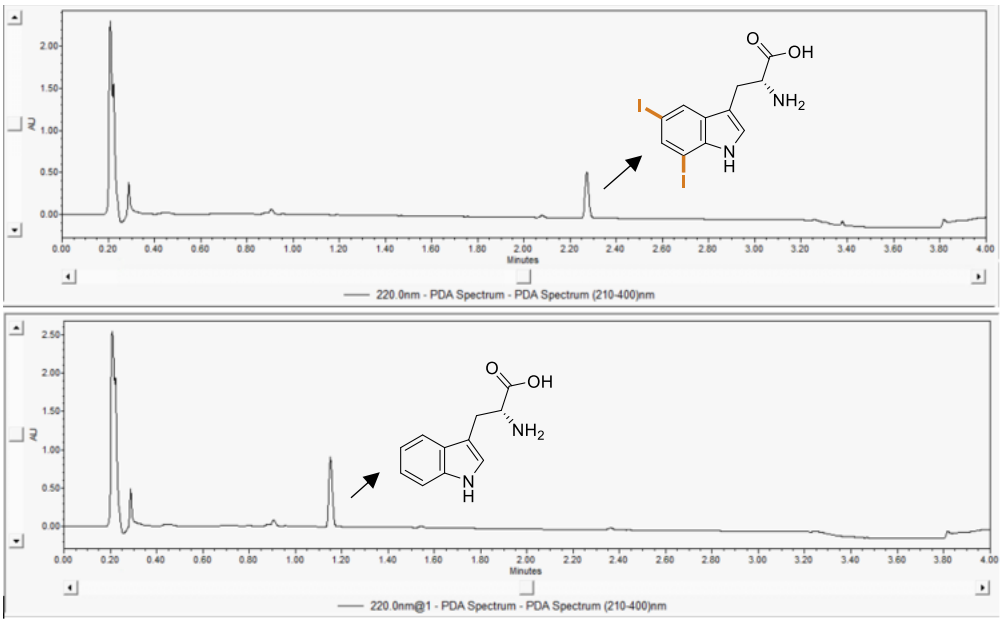


Figure S5. LC/MS chromatogram showing iodination activity using AetF and AetF K258A. Top: Wild-type AetF showing full conversion to diiodinated L-Tryptophan. Bottom: Chromatogram showing no activity with AetF K258A variant.

B) Reaction Optimization

Buffer optimization studies to facilitate preparative halogenation

Study 1:

Scheme SX: Initial optimization studies to identify improved buffer conditions for AetF catalysis.

The layout of the plate was as follows:

Buffer	Concentration	Plate	1	2	3	4	5	6	7	8	9	10	11	12
HEPES	25 mM	A	pH 6.5			pH 7.0			pH 7.5			pH 8.0		
HEPES	100 mM	B	pH 6.5			pH 7.0			pH 7.5			pH 8.0		
Phosphate	25 mM	C	pH 6.5			pH 7.0			pH 7.5			pH 8.0		
Phosphate	100 mM	D	pH 6.5			pH 7.0			pH 7.5			pH 8.0		

During cell lysis, lysis buffer was prepared as a 10x stock solution of lysozyme and benzonase I in HEPES pH 7.4, 5 mM (7.5 mg/mL lysozyme, 10 uL/mL benzonase I). 150 µL of this stock solution was then aliquoted into 1.35 mL of the appropriate pH and concentration of either HEPES or phosphate. From the 100 µL of clarified lysate, 50 µL was aliquoted into a micro titer plate. A solution of 500 µM FAD, 500 µM NADP, 100 mM D-glucose, 50 mM NaBr, and 12.5 mM of 6-fluorotryptamine in DMSO (50 mM stock solution, 5% v/v DMSO final concentration) was prepared in reaction buffer (small molecule stock solution). A solution of 67.5 U/mL GDH and 262.5 U/mL catalase was prepared in reaction buffer (enzyme stock solution). 15 µL of the small molecule stock solution was added to the 96 well microtiter plates containing clarified lysate. 10 µL of the enzyme stock solution was then added to give a total reaction volume of 75 µL and a final AetF concentration of 25 µM. The plate was sealed using an aluminum heat seal and incubated in ambient atmosphere using a plate shaker set (Eppendorf Thermomixer R) to 600 rpm. The next

day, reactions were quenched using 1 volume of methanol, and the quenched reactions were further incubated for 15 minutes to facilitate full protein precipitation. Precipitated protein was removed from the solution via centrifugation at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes. From the resulting supernatant, 50 µL was then added to a 0.2 µm filter plate charged with 150 µL of water from a Milli-Q filtration system, and filtration was done gravimetrically at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes into a fresh microtiter plate. The microtiter plate containing filtered reactions was then heat-sealed using an aluminum heat-seal and analyzed via UHPLC using method 1. Conversion is defined as % AUC of the product(s) as determined by MAX PLOT values compared to the % AUC of the total area of starting material + product(s).

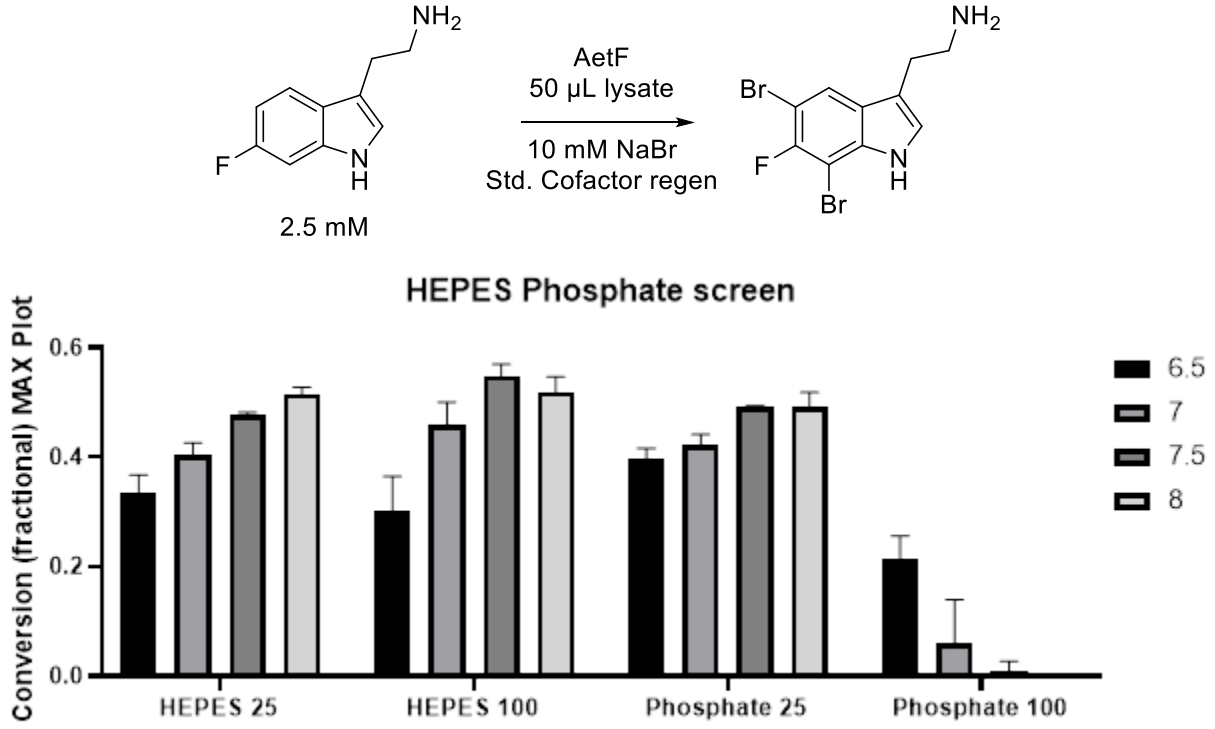


Figure S6. Results of the first optimization study. Interpretation: AetF conversion is highest in HEPES, 100 mM at pH 7.5.

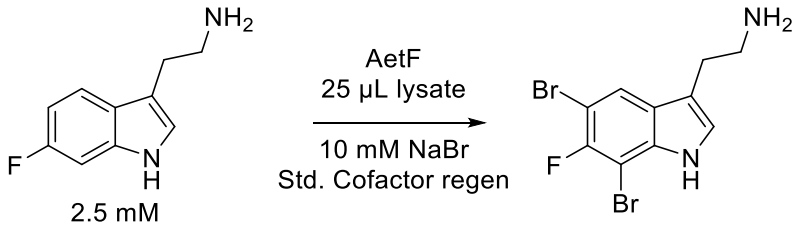
Study 2:

During cell lysis, lysis buffer was prepared as a 10x stock solution of lysozyme and benzonase I in HEPES pH 7.4, 5 mM (7.5 mg/mL lysozyme, 10 uL/mL Benzonase I). 150 µL of this stock solution was then aliquoted into 1.35 mL of the appropriate pH and concentration of either HEPES or phosphate. From the 100 µL of clarified lysate, 25 µL was aliquoted into a micro titer plate followed by 25 µL of the appropriate reaction buffer. Stock solutions of the appropriate composition were comprised as follows:

The final layout of the microtiter plate was as follows. Reactions were conducted in triplicate.

Concentration	Plate	1	2	3	4	5	6	7	8	9	10	11	12
25 mM	A	pH 7.5			pH 8.0			pH 8.5			pH 9.0		
100 mM	B	pH 7.5			pH 8.0			pH 8.5			pH 9.0		
		HEPES, pH 7.5, 100 mM											

15 µL of the small molecule stock solution was added to the 96 well microtiter plates containing clarified lysate. 10 µL of the enzyme stock solution was then added to give a total reaction volume of 75 µL and a final AetF concentration of 25 µM. The plate was sealed using an aluminum heat seal and incubated in ambient atmosphere using an Eppendorf thermomixer R plate shaker set to 600 rpm. The next day, reactions were quenched using 1 volume of methanol, and the quenched reactions were further incubated for 15 minutes to facilitate full protein precipitation. Precipitated protein was removed from the solution via centrifugation at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes. From the resulting supernatent, 50 µL was then added to a 0.2 µM filter plate charged with 150 µL of water from a Milli-Q filtration system, and filtration was done gravimetrically at 3600 x rpm in a Sorvall XT centrifuge for 15 minutes into a fresh microtiter plate. The microtiter plate containing filtered reactions was then heat-sealed using an aluminum heat-seal and analyzed via UHPLC using method 1. Conversion is defined as % AUC of the product(s) as determined by MAX PLOT value compared to the % AUC of the total area of starting material + product(s).



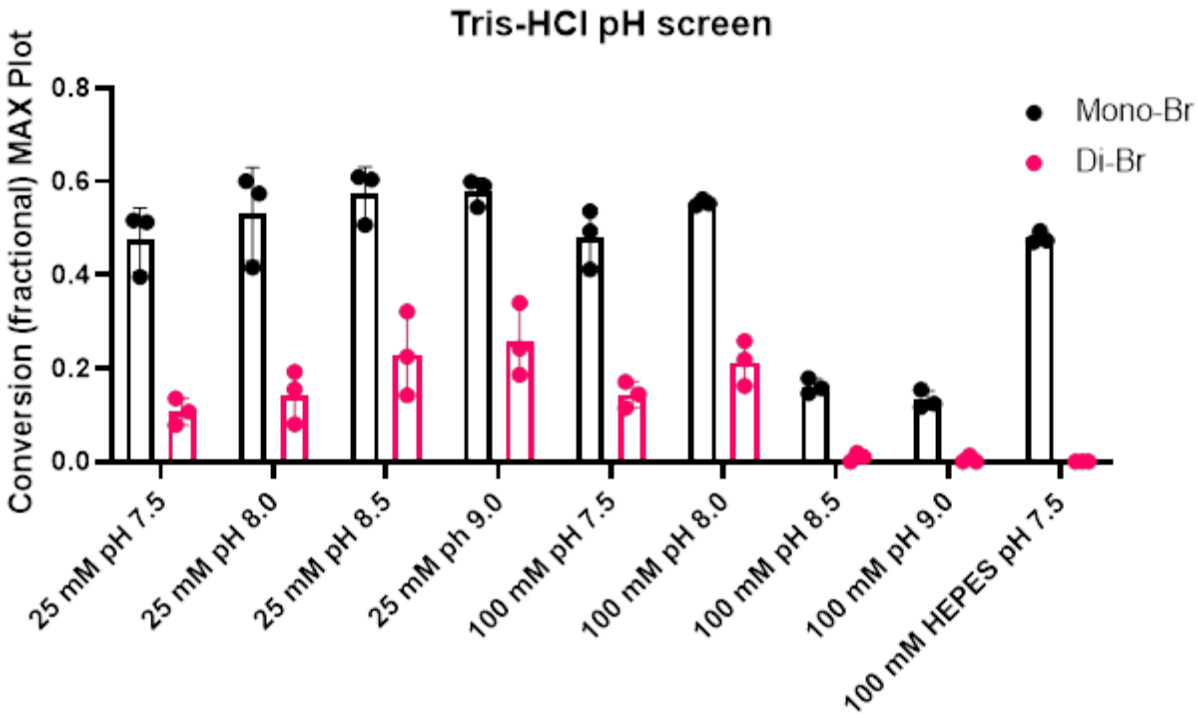
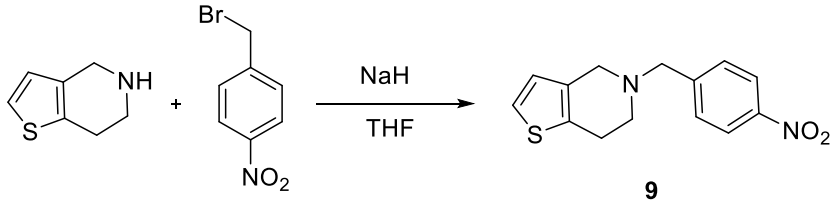


Figure S7. Results of the second optimization study. Interpretation: AetF has the highest activity in Tris pH 9.0, 25 mM.

IV. Synthetic Procedures

*Furan and thiophene compounds **9** and **10**:*



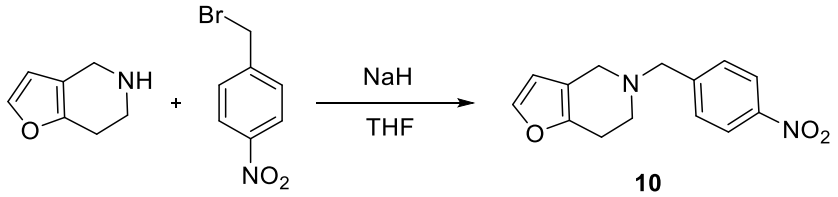
9

To a suspension of NaH (60%, 1.2 mmol, 48 mg, 1.2 equiv.) in THF (4 mL) was added a solution of 4,5,6,7-tetrahydrothieno[3,2-c]pyridine (1 mmol, 139 mg) in THF (1 mL), and the mixture was allowed to stir at room temperature for 30 min. A solution of 1-(bromomethyl)-4-nitrobenzene (1.2 mmol, 260 mg) in THF (1 mL) was then added, and the reaction was stirred overnight. The reaction was quenched by adding MeOH, and then concentrated under reduced pressure. The residue was further purified by flash column chromatography (Hexane/EA= 6/1) to afford the desired product in 84% yield.

^{1}H NMR (500 MHz, Chloroform-d) δ 8.19 (d, J = 8.5 Hz, 2H), 7.60 (d, J = 8.4 Hz, 2H), 7.11 – 7.07 (m, 1H), 6.73 (d, J = 5.1 Hz, 1H), 3.83 (s, 2H), 3.61 (s, 2H), 2.93 (t, J = 5.5 Hz, 2H), 2.84 (t, J = 5.7 Hz, 2H).

^{13}C NMR (126 MHz, Chloroform-d) δ 147.19, 146.73, 133.62, 133.33, 129.45, 125.25, 123.59, 122.90, 61.31, 53.22, 50.79, 25.54.

HRMS (ESI-MS) Calc. for $[C_{14}H_{15}N_2O_2S]^+$ ([M+H])$^+$: 275.0854, found 275.0851.



10

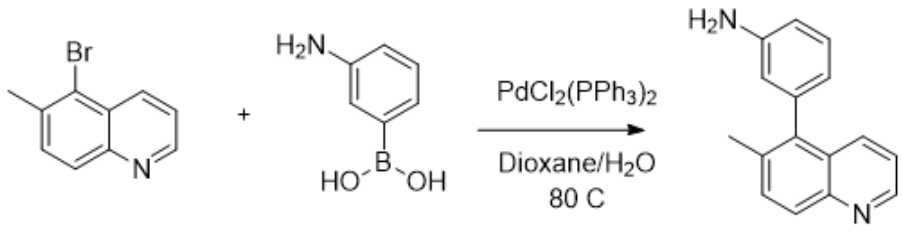
To a suspension of NaH (60%, 1.2 mmol, 48 mg) in THF (4 mL) was added a solution of 4,5,6,7-tetrahydrofuro[3,2-c]pyridine (1 mmol, 123 mg) in THF (1 mL), and the mixture was allowed to stir at room temperature for 30 min. A solution of 1-(bromomethyl)-4-nitrobenzene (1.2 mmol, 260 mg) in THF (1 mL) was then added, and the reaction was stirred overnight. The reaction was quenched by adding MeOH and then concentrated under reduced pressure. The residue was further purified by flash column chromatography (Hexane/EA= 6/1) to afford the desired product in 67% yield.

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.16 (d, *J* = 8.7 Hz, 2H), 7.55 (d, *J* = 8.8 Hz, 2H), 7.24 (d, *J* = 1.8 Hz, 1H), 6.15 (d, *J* = 1.9 Hz, 1H), 3.79 (s, 2H), 3.42 (s, 2H), 2.81 (t, *J* = 5.7 Hz, 2H), 2.71 (t, *J* = 5.7 Hz, 2H).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 148.62, 147.17, 146.73, 141.06, 129.37, 123.56, 115.32, 108.57, 77.41, 77.16, 76.91, 60.89, 50.28, 50.12, 23.88.

HRMS (ESI-MS) Calc. for $[C_{14}H_{15}N_2O_3]^+$ ([M+H])$^+$: 259.1083, found 259.1079.

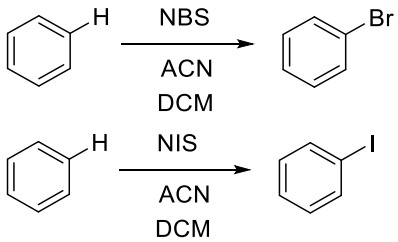
*Quinoline **13**:*



To an oven-dried 50 mL round-bottom flask containing a stir-bar were added 3-aminophenyl boronic acid (HCl salt, 1.73 mmol, 300 mg, 1.3 equiv), 5-bromo-6-methylquinoline (1.35 mmol, 300 mg, limiting reagent), $PdCl_2(PPh_3)_2$ (0.04 mmol, 32 mg, 3 mol % Pd), and cesium carbonate (3.46 mmol, 1.127 g, 2.5 equiv.). The flask was then evacuated and purged with nitrogen 3 times and fitted with a condenser capped with a rubber septum. 15 mL of a 1:4 mixture of degassed dioxane: water was then added to the round bottom flask and a nitrogen balloon was added to the condenser. The solution was heated to 80 °C with vigorous stirring for 14 hours. The next day, the mixture was diluted with 1 volume of methanol, filtered through celite, and the solvent was removed via rotary evaporation under reduced pressure. The resulting residue was taken up in DCM and washed with 0.1 M NaOH, then brine, and the organic layer was dried over magnesium sulfate and the solvent was removed by rotary evaporation.

To purify the desired compound, silica gel was added to the resulting residue and the mixture was taken up in DCM and the compound was adsorbed onto the silica by removal of solvent. This silica gel was then added to a column presaturated with 1:1 Ethyl acetate: Hexanes with 1% TEA additive. NMR of the major component of the reaction showed an impurity from the Buchwald-Hartwig amination product, which was removed via reverse-phase chromatography using Biotage method 1. The desired product was obtained as a brown solid in 27% yield.
^{1}H NMR (500 MHz, Chloroform-d) δ 8.74 (dd, J = 4.2, 1.7 Hz, 1H), 7.94 (d, J = 8.7 Hz, 1H), 7.76 (dt, J = 8.6, 1.2 Hz, 1H), 7.54 (d, J = 8.6 Hz, 1H), 7.22 – 7.19 (m, 1H), 7.16 (dd, J = 8.5, 4.2 Hz, 1H), 6.71 – 6.66 (m, 1H), 6.55 (dt, J = 7.5, 1.3 Hz, 1H), 6.48 (t, J = 1.9 Hz, 1H), 3.69 (s, 2H), 2.22 (s, 3H).
^{13}C NMR (126 MHz, Chloroform-d) δ 149.13, 146.82, 146.49, 139.63, 138.40, 134.63, 133.74, 132.15, 129.46, 128.24, 127.88, 120.77, 120.44, 116.68, 114.13, 20.55.
HRMS (ESI-MS) Calc. for $[C_{16}H_{15}N_2]^+$ ($[M+H]$)$^+$: 235.1230, found 235.1231.

General procedures for preparation of authentic products using NBS or NIS

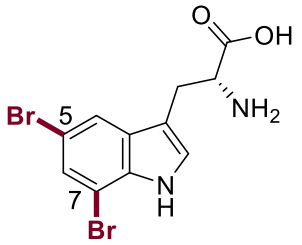


Authentic racemic products were prepared using a modified reported procedure.[3] To a solution of substrate (1 mmol) in DCM (3 mL) and ACN (2 mL) was added *N*-bromosuccinimide or *N*-

iodosuccinimide (1.1 mmol). The resulting mixture was stirred at room temperature and monitored by TLC. Upon completion of reaction, the reaction was concentrated. The residue was purified by flash column chromatography on silica gel or reverse phase biotage to yield the corresponding product. Relevant LC/MS chromatograms and NMR spectra are presented below.

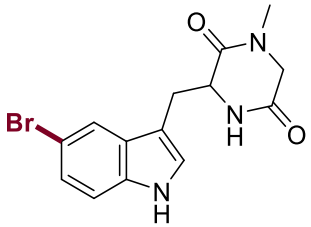
V. Product isolation and characterization

*Product **1a**:*



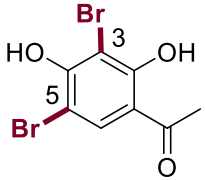
The general procedure for preparative-scale AetF reactions and Biotage method 3 were used to obtain product **1a** as a white solid in 74% yield (10.4 mg). ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.24 (br, 1H), 8.31 (m, NH_2), 7.76 (s, 1H), 7.40 (s, 1H), 7.28 (s, 1H), 3.07 (m, 2H), 2.77 – 2.57 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 133.43, 130.86, 126.40, 124.61, 120.79, 114.36, 110.33, 104.77, 56.78, 25.47. The NMR spectra matched those that were previously reported.[4] HRMS (ESI-MS) Calc. for $[C_{11}H_{11}Br_2N_2O_2]^+$ ($[M+H]$)$^+$: 360.9182, found 360.9182.

*Product **2a**:*



The general procedure for preparative-scale AetF reactions and semi-preparative HPLC Method 1 were used to obtain product **2a** as a white solid in 28% yield (2.2 mg, formic acid salt). ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.63 (d, J = 1.7 Hz, 1H), 7.29 (d, J = 8.6 Hz, 1H), 7.19 (dd, J = 8.6, 1.8 Hz, 1H), 7.06 (s, 1H), 4.57 (s, 2H), 4.23 (t, J = 3.8 Hz, 1H), 3.44 (dd, J = 14.6, 3.1 Hz, 1H), 3.05 (dd, J = 14.7, 4.6 Hz, 1H), 2.57 (s, 3H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 168.73, 168.17, 136.62, 130.24, 128.03, 125.50, 122.42, 114.06, 113.36, 108.69, 57.19, 51.57, 33.79, 31.65. HRMS (ESI-MS) Calc. for $[C_{14}H_{14}BrN_3N_aO_2]^+$ ($[M+Na]$)$^+$: 358.0167, found 358.0163.

*Product **3a**:*



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 4/1) was used to isolated product **3a** was obtained as a white solid. ^{1}H NMR (400 MHz, Chloroform-d) δ 13.32 (s, 1H), 7.88 (s, 1H), 2.60 (s, 3H). The spectra matched those that

were previously reported.[5] This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S8).

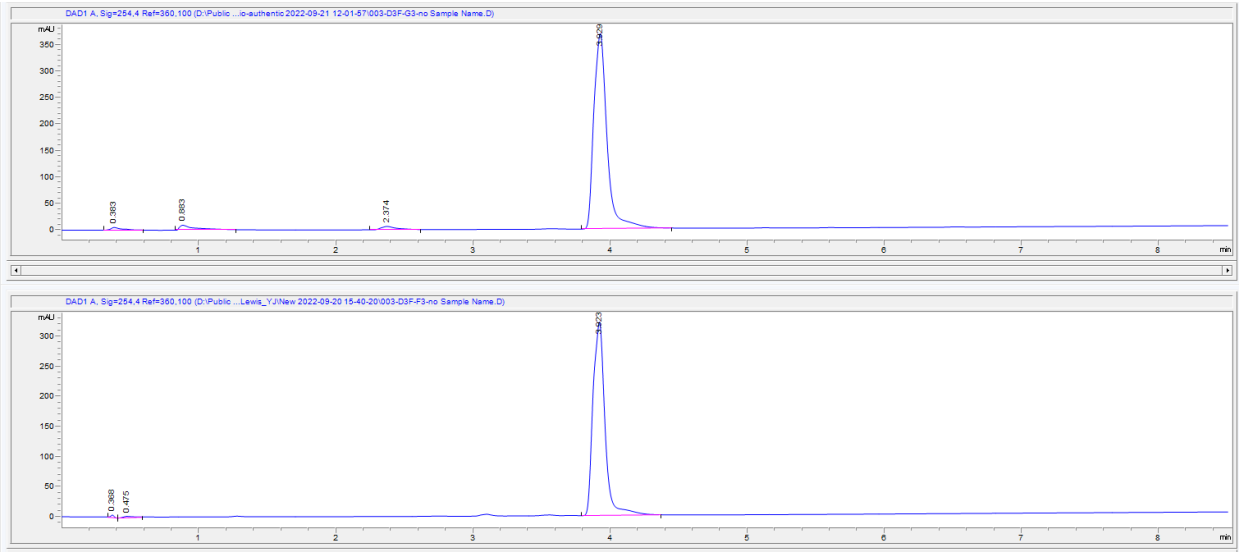
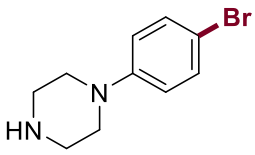


Figure S8. Comparison of chromatograms (254 nm) of **3a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

*Product **4a**:*



The general procedure for bromination using NBS and Biotage method 2 were used to obtain **4a** as a white solid. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.49 (s, 1H), 7.34 (d, J = 8.6 Hz, 2H), 6.89 (d, J = 8.6 Hz, 2H), 3.35 – 3.26 (m, 8H). The spectral data matched those that were previously reported.[6] This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S9).

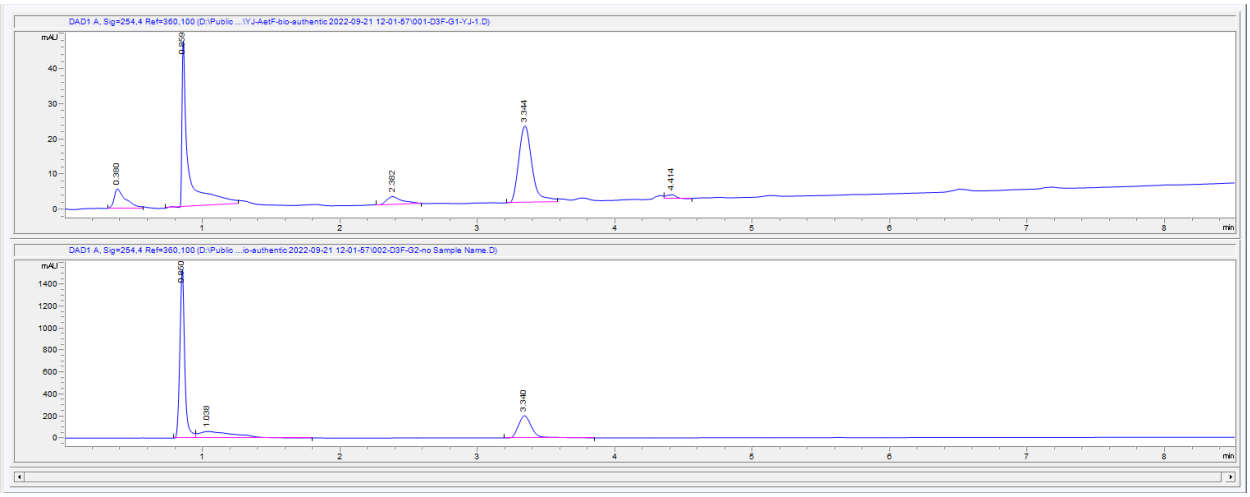
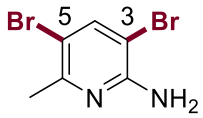


Figure S9. Comparison of chromatograms (254 nm) of **4a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

*Product **5a**:*



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 4/1) were used to obtain product **5a** as a white solid. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.72 (s, 1H), 4.93 (s, 2H), 2.44 (s, 3H). The spectral data matched those that were previously reported.[7] This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S10).

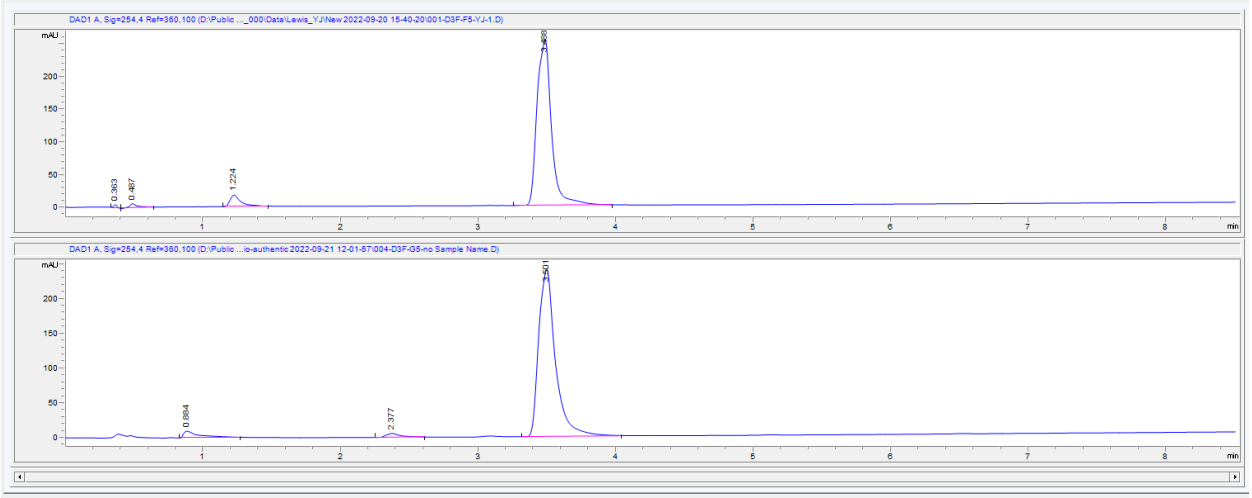
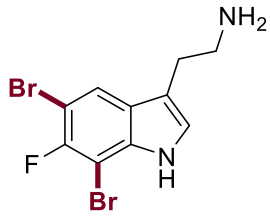


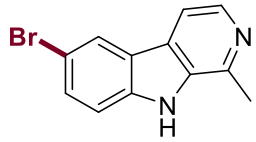
Figure S10. Comparison of chromatograms (254 nm) of **5a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

*Product **6a**:*



The general procedure for preparative-scale AetF reactions and semi-preparative HPLC Method 2 were used to obtain product **6a** as a white solid in 51% yield (4.1 mg, formic acid salt).
^{1}H NMR (600 MHz, 25:75 Deuterium Oxide: Acetonitrile-d_3) δ 8.40 (s, 1H), 7.82 (d, J = 6.1 Hz, 1H), 7.31 (s, 1H), 3.16 (t, J = 7.5 Hz, 2H), 3.01 (t, J = 7.5 Hz, 2H).
^{13}C NMR (126 MHz, 25:75 Deuterium Oxide: Acetonitrile-d_3) δ 151.88 (d, J = 235.2 Hz), 135.27, 126.72, 125.57, 122.02, 111.17, 100.28 (d, J = 27.1 Hz), 92.24 (d, J = 26.9 Hz), 40.02, 23.01.
HRMS (ESI-MS) Calc. for $[C_{10}H_{10}Br_2FN_2]^+$ ([M+H])$^+$: 334.9195, found 334.9190 and Calc. for (M-NH_3)$^+$: 317.8929, found 317.8925

*Product **7a**:*



The general procedure for bromination using NBS and Biotage method 2 were used to obtain **7a** as a white solid. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.26 (d, J = 1.8 Hz, 1H), 8.15 (d, J = 5.5 Hz, 1H), 7.90 (d, J = 5.5 Hz, 1H), 7.62 (dd, J = 8.7, 1.9 Hz, 1H), 7.48 (d, J = 8.7 Hz, 1H), 2.79 (s, 3H). The spectral data matched those that were previously reported.[8] This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S11).

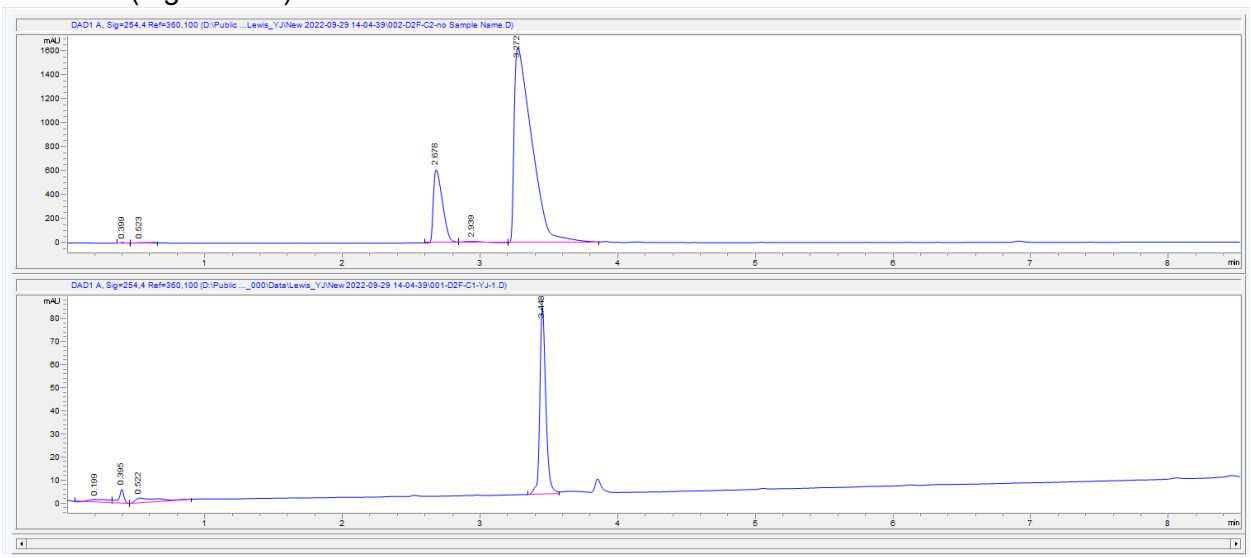
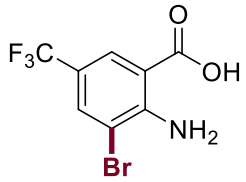


Figure S11. Comparison of chromatograms (254 nm) of **7a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

*Product **8a**:*



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 4/1) were used to obtain product **8a** as a white solid. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.12 – 8.10 (d, J = 1.9 Hz, 1H), 7.80 (d, J = 1.9 Hz, 1H). ^{13}C NMR (126 MHz, Methanol-d_4) δ 169.75, 151.84, 134.29, 129.65, 125.14 (q), 118.35 (q), 112.36, 110.73. HRMS (APCI-MS) Calc. for $[C_8H_4BrF_3NO_2]^-$ ($[M-H]^-$: 281.9378, found 281.9379. This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S12).

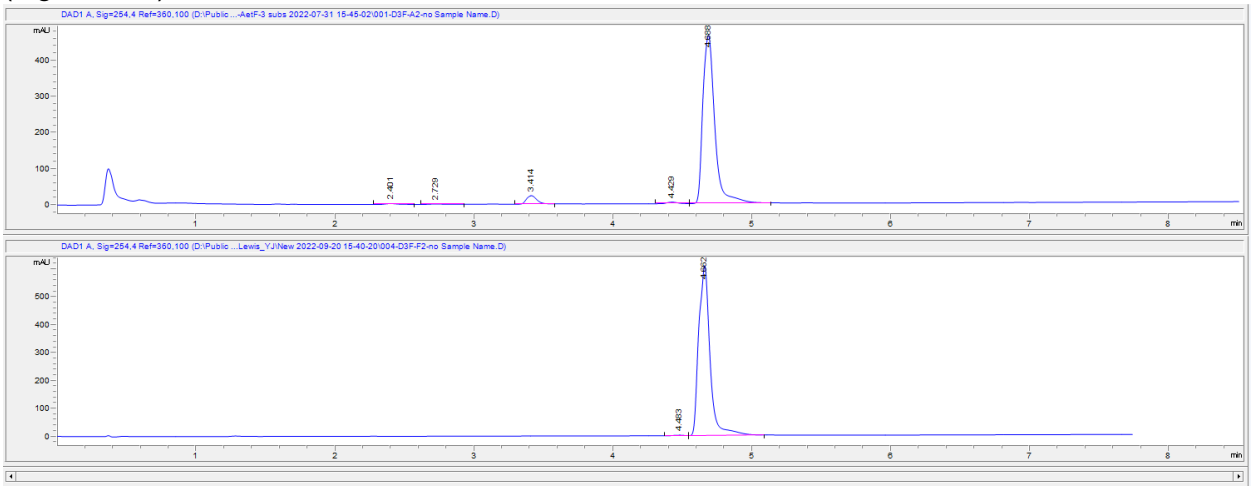
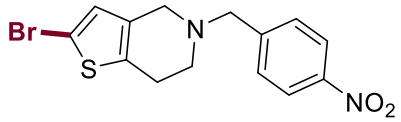


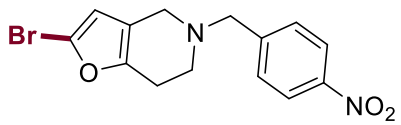
Figure S12. Comparison of chromatograms (254 nm) of **8a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

*Product **9a**:*



The general procedure for preparative-scale AetF reactions and semi-preparative HPLC Method 3 were used to obtain product **9a** as a white solid in 48% yield (3.7 mg). ^{1}H NMR (400 MHz, Chloroform-d) δ 8.20 (d, J = 8.6 Hz, 2H), 7.57 (d, J = 8.5 Hz, 2H), 6.65 (s, 1H), 3.80 (s, 2H), 3.51 (s, 2H), 2.82 (s, 4H). ^{13}C NMR (126 MHz, Chloroform-d) δ 147.44, 146.32, 135.08, 134.12, 129.49, 128.00, 123.80, 109.79, 61.27, 52.71, 50.56, 25.46. HRMS (ESI-MS) Calc. for $[C_{14}H_{14}BrN_2O_2S]^+$ ($[M+H]^+$: 352.9959, found 352.9957.

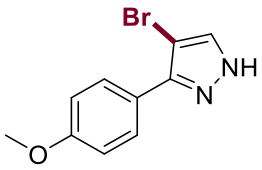
*Product **10a**:*



The general procedure for preparative-scale AetF reactions and semi-preparative HPLC Method 4 were used to obtain product **10a** as a white solid in 60% yield (4.7 mg). ^{1}H NMR (500 MHz,

Chloroform-*d*) δ 8.19 (d, *J* = 8.7 Hz, 2H), 7.55 (d, *J* = 8.6 Hz, 2H), 6.08 (s, 1H), 3.80 (s, 2H), 3.39 (s, 2H), 2.83 (t, *J* = 5.7 Hz, 2H), 2.72 (d, *J* = 5.7 Hz, 2H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 150.59, 147.49, 146.35, 129.52, 123.84, 120.31, 118.29, 110.24, 60.82, 50.18, 49.72, 23.89. HRMS (ESI-MS) Calc. for $[C_{14}H_{14}BrN_2O_3]^+$ $([M+H])^+$: 337.0188, found 337.0183.

*Product **11a**:*



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 5/1) were used to obtain product **11a** as a white solid. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 11.85 (s, 1H), 7.63 – 7.57 (d, 2H), 7.49 (s, 1H), 6.94 – 6.85 (d, 2H), 3.78 (s, 3H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 160.16, 144.18, 137.40, 129.23, 122.10, 114.27, 91.98, 55.47. HRMS (ESI-MS) Calc. for $[C_{10}H_{10}BrN_2O]^+$ $([M+H])^+$: 252.9977, found 252.9972. This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S13).

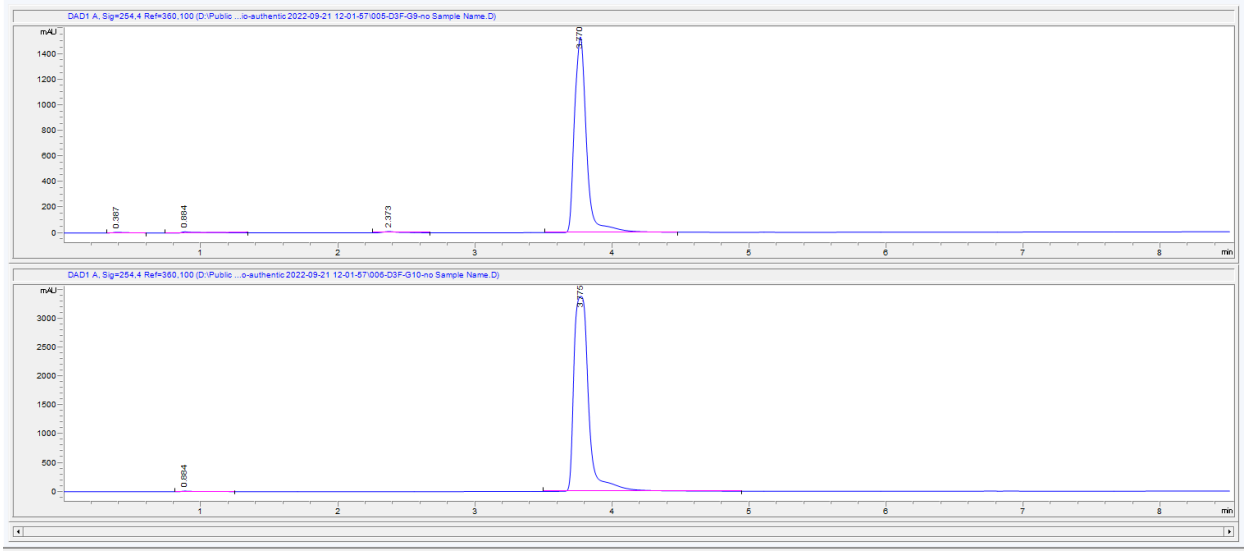
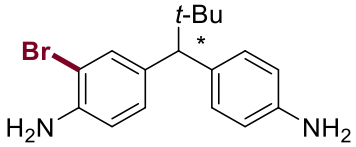


Figure S13. Comparison of chromatograms (254 nm) of **11a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

*Product **12a**:*



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 2/1) were used to obtain product **12a** as a brown solid. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.42 (s, 1H), 7.14 (d, *J* = 8.4 Hz, 3H), 6.66 (d, *J* = 8.2 Hz, 1H), 6.60 (d, *J* = 8.5 Hz, 2H), 3.92 (s, 2H), 3.54 (s, 2H), 3.45 (s, 1H), 0.97 (s, 9H). The spectral data matched those that were previously reported.[9] This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S14/15).

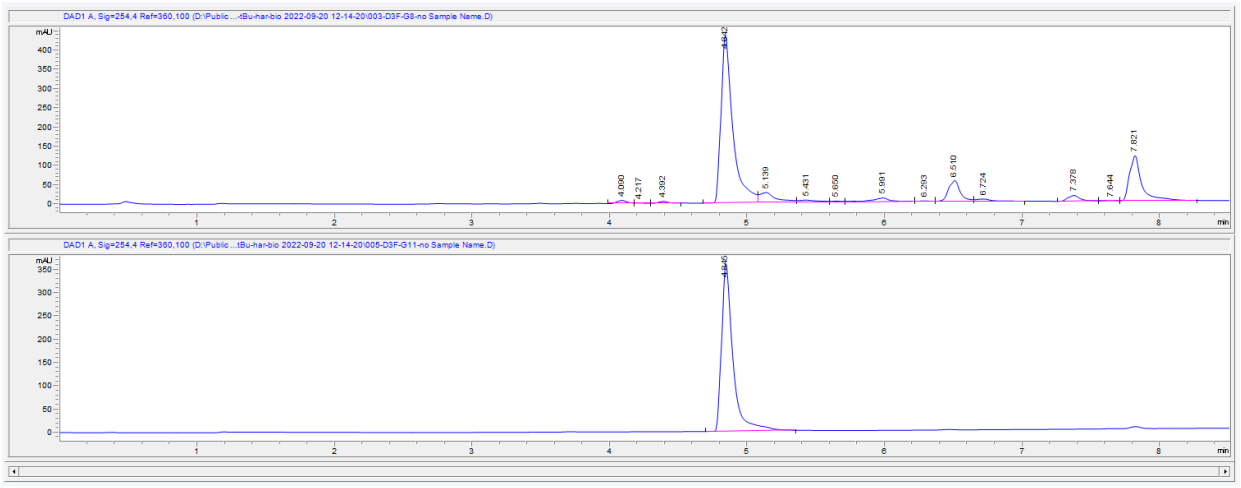
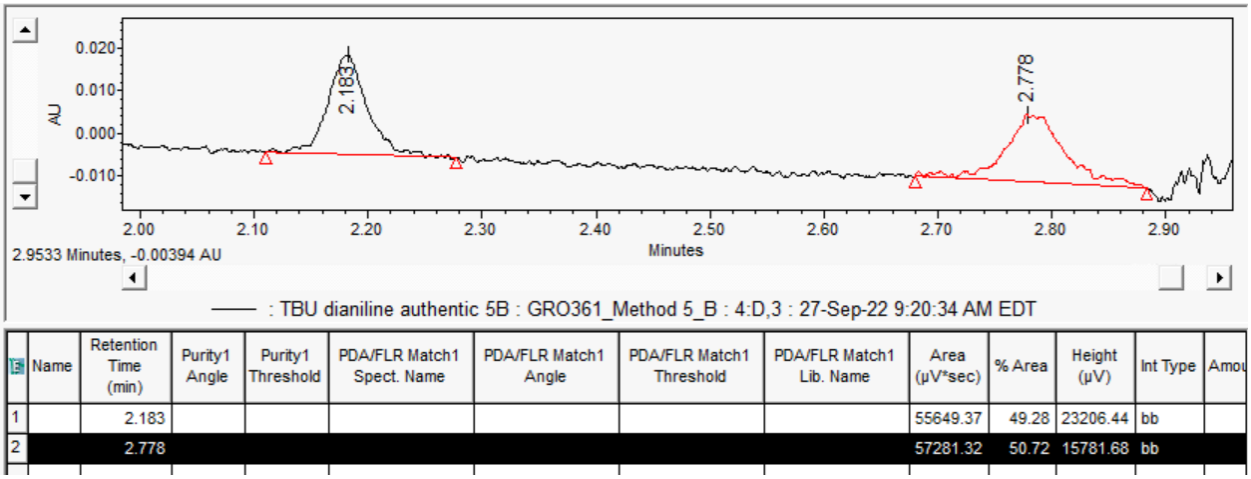
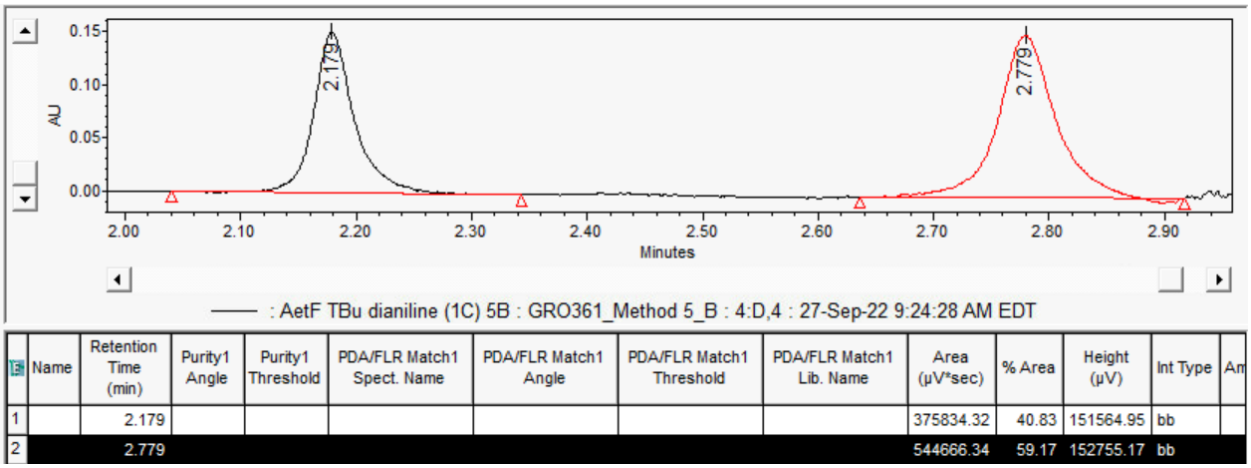


Figure S14. Comparison of chromatograms (254 nm) of **12a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).



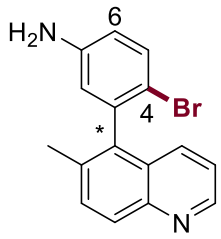
	Name	Retention Time (min)	Purity1 Angle	Purity1 Threshold	PDA/FLR Match1 Spect. Name	PDA/FLR Match1 Angle	PDA/FLR Match1 Threshold	PDA/FLR Match1 Lib. Name	Area (µV*sec)	% Area	Height (µV)	Int Type	Amou
1		2.183							55649.37	49.28	23206.44	bb	
2		2.778							57281.32	50.72	15781.68	bb	



	Name	Retention Time (min)	Purity1 Angle	Purity1 Threshold	PDA/FLR Match1 Spect. Name	PDA/FLR Match1 Angle	PDA/FLR Match1 Threshold	PDA/FLR Match1 Lib. Name	Area (µV*sec)	% Area	Height (µV)	Int Type	Am
1		2.179							375834.32	40.83	151564.95	bb	
2		2.779							544666.34	59.17	152755.17	bb	

Figure S15. Comparison of chiral chromatograms (230 nm) of **12a** produced using general procedures for NBS reactions (top) and analytical-scale AetF reactions (bottom). Based on these, the e.r. value of enzymatic material is 40:60.

*Product **13a**:*



The general procedure for bromination using NBS and Biotage method 2 were used to obtain **13a** as a brown solid. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.84 (dd, *J* = 4.1, 1.4 Hz, 1H), 8.06 (d, *J* = 8.7 Hz, 1H), 7.67 (d, *J* = 8.0 Hz, 1H), 7.64 (d, *J* = 8.7 Hz, 1H), 7.48 (d, *J* = 8.6 Hz, 1H), 7.30 – 7.26 (m, 1H), 6.67 (dd, *J* = 8.6, 2.8 Hz, 1H), 6.55 (d, *J* = 2.8 Hz, 1H), 3.76 (s, 2H), 2.26 (s, 3H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 149.41, 146.90, 146.06, 139.90, 137.28, 134.38, 134.05, 133.53, 132.18, 129.02, 127.29, 121.20, 117.95, 116.42, 112.47, 29.82. HRMS (ESI-MS) Calc. for $[C_{16}H_{14}BrN_2]^+$ ($[M+H]$)$^+$: 313.0340, found 313.0336. This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S16/17).

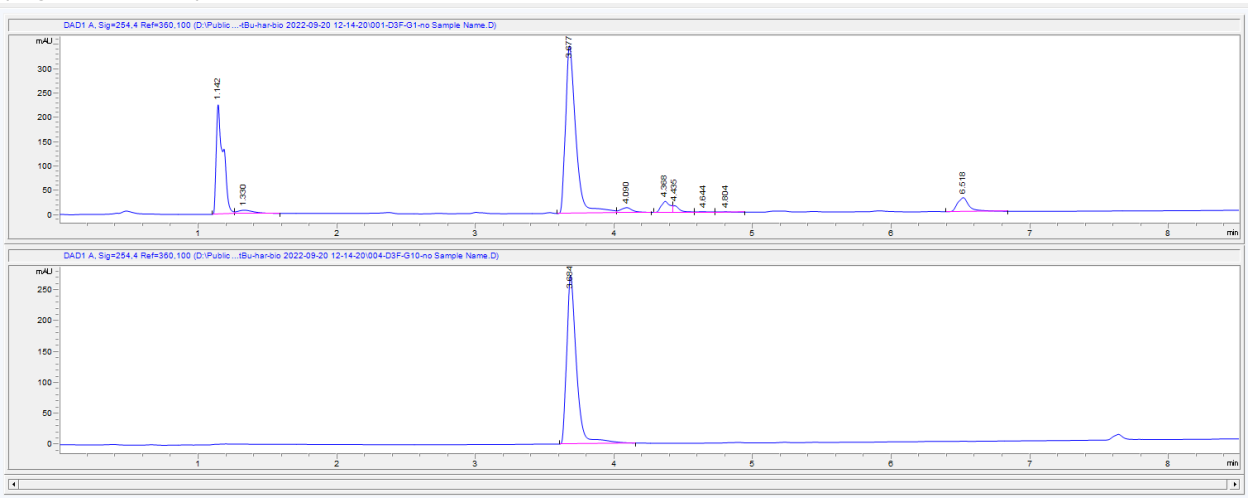
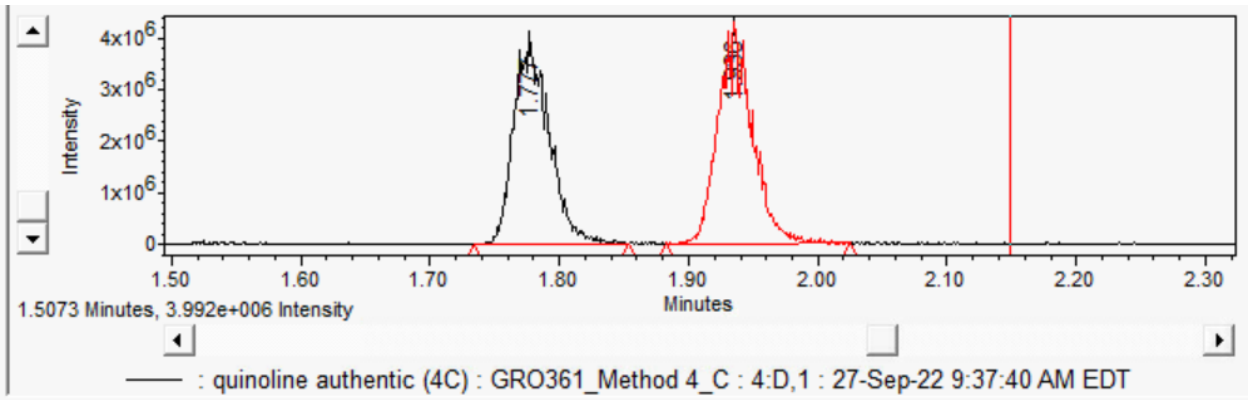
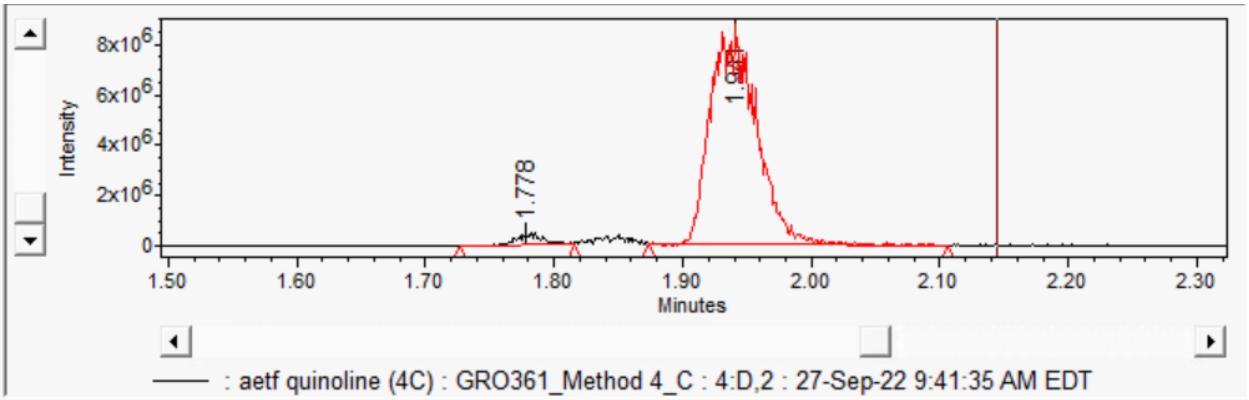


Figure S16. Comparison of chromatograms (254 nm) of **13a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).



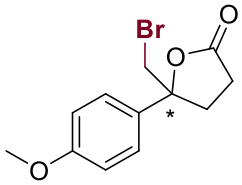
	Name	Retention Time (min)	Area (μV*sec)	% Area	Height (μV)	Int Type	Amount	Units	Peak Type	Peak Codes	A_Pfizer_Conc
1		1.777	7545017.98	49.61	3989106.93	bb			Unknown	I21 I19 C01 C06	
2		1.936	7664662.63	50.39	4294481.25	bb			Unknown	C01 C06	



	Name	Retention Time (min)	Area (μV*sec)	% Area	Height (μV)	Int Type	Amount	Units	Peak Type	Peak Codes	A_Pfizer_Conc	A
1		1.778	607486.48	2.76	423778.03	bb			Unknown	C01 C06		6
2		1.941	21398760.24	97.24	8549455.16	bb			Unknown	C01 C06		213

Figure S17. Comparison of chiral chromatograms (230 nm) of **13a** produced using general procedures for NBS reactions (top) and analytical-scale AetF reactions (bottom). Based on these, the e.r. value of enzymatic material is 97:3.

*Product **14a**:*



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 4/1) were used to obtain product **14a** as a white solid. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.38 – 7.31 (m, 2H), 6.98 – 6.92 (m, 2H), 3.89 (d, J = 11.3 Hz, 1H), 3.80 (s, 3H), 3.79 (d, J = 11.3 Hz, 1H), 2.83 – 2.72 (m, 2H), 2.60 – 2.50 (m, 2H). The spectrum data matched those that

were previously reported.[10] This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S18/19). Yield of AetF-catalyzed reaction: 72%. Product (P) assay yields (AY) were determined relative to internal standard (IS) with appropriate response factors (RF) using the formula AY = RF•(integral P/integral IS)+C.

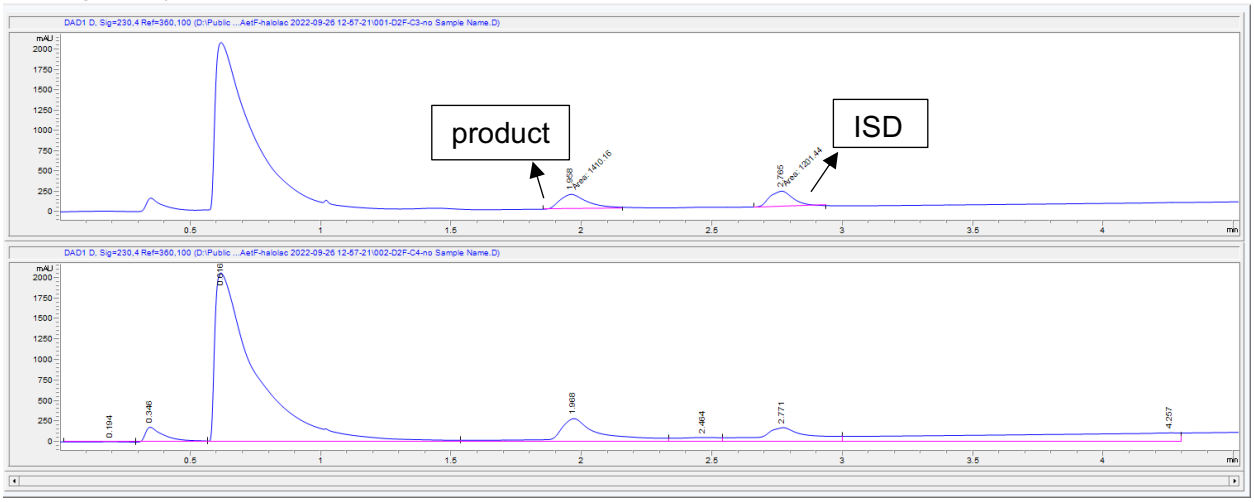


Figure S18. Comparison of chromatograms (230 nm) of **14a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

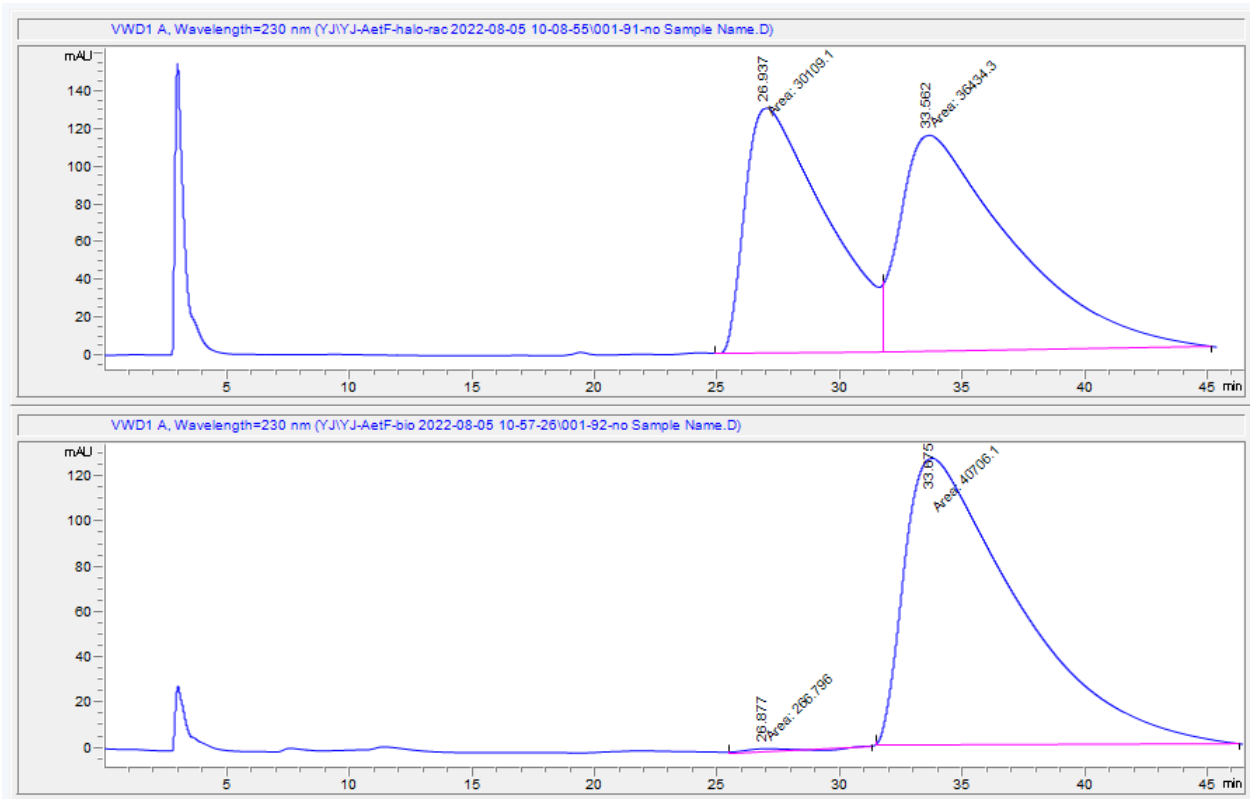
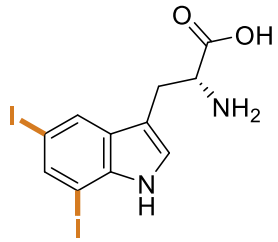


Figure S19. Comparison of chromatograms (230 nm) of **14a** produced using general procedures for NBS reactions (top) and analytical-scale AetF reactions (bottom). Based on these, the e.r. value of enzymatic material is 99:1.

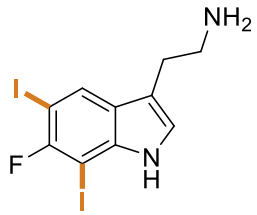
*Product **1b**:*



The general procedure for preparative-scale AetF reactions and Biotage method 3 were used to obtain product **1b** as a white solid in 71% yield (10.1 mg).
^{1}H NMR (600 MHz, Deuterium Oxide with TFA) δ 7.66 (s, 1H), 7.59 (s, 1H), 7.07 (s, 1H), 4.14 (dd, *J* = 7.5, 4.8 Hz, 1H), 3.19 (dd, *J* = 15.4, 4.8 Hz, 1H), 3.09 (dd, *J* = 15.3, 7.5 Hz, 1H).

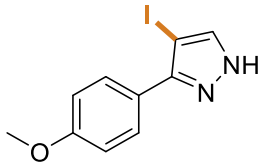
^{13}C NMR (126 MHz, Deuterium Oxide with TFA) δ 171.30, 137.32, 137.30, 128.37, 126.84, 126.15, 107.05, 82.63, 77.29, 52.91, 25.97.
HRMS (ESI-MS) Calc. for $[C_{11}H_{11}I_2N_2O_2]^+$: $([M+H])^+$ 456.8904, found 456.8905.

*Product **6b**:*



The general procedure for preparative-scale AetF reactions and semi-preparative HPLC Method 4 were used to obtain product **6b** as a white solid in 47% yield (5.2 mg, formic acid salt).
^{1}H NMR (500 MHz, 25:75 Deuterium Oxide: Acetonitrile-d_3) δ 8.43 (s, 1H), 8.01 (d, *J* = 5.9 Hz, 1H), 7.30 (s, 1H), 3.18 (t, *J* = 7.5 Hz, 2H), 3.02 (t, *J* = 7.5 Hz, 2H).
^{13}C NMR (126 MHz, 25:75 Deuterium Oxide: Acetonitrile-d_3) δ 156.53 (d, *J* = 230.6 Hz), 139.67 (d, *J* = 5.9 Hz), 129.24 (d, *J* = 2.7 Hz), 126.36, 126.33, 111.14, 69.73 (d, *J* = 30.9 Hz), 62.96 (d, *J* = 34.8 Hz), 40.17, 23.33.
HRMS (ESI-MS) Calc. for $[C_{10}H_{10}I_2FN_2]^+$ $([M+H])^+$: 430.8912, found 430.8914 and Calc. for $(M-NH_2)^+$: 413.8646, found 413.8650

<u>Product **11b**:</u>



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 10/1) were used to obtain product **11b** as a white solid.
^{1}H NMR (400 MHz, Chloroform-*d*) δ 12.62 (s, 1H), 7.62 (d, *J* = 8.6 Hz, 2H), 7.56 (s, 1H), 6.94 (d, *J* = 8.6 Hz, 2H), 3.86 (s, 3H).
^{13}C NMR (126 MHz, Chloroform-*d*) δ 160.18, 147.34, 142.33, 129.47, 122.71, 114.15, 56.95, 55.46.
HRMS (APCI-MS) Calc. for $[C_{10}H_{10}IN_2O]^+$ $([M+H])^+$: 300.9832, found 300.9833.

This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S20).

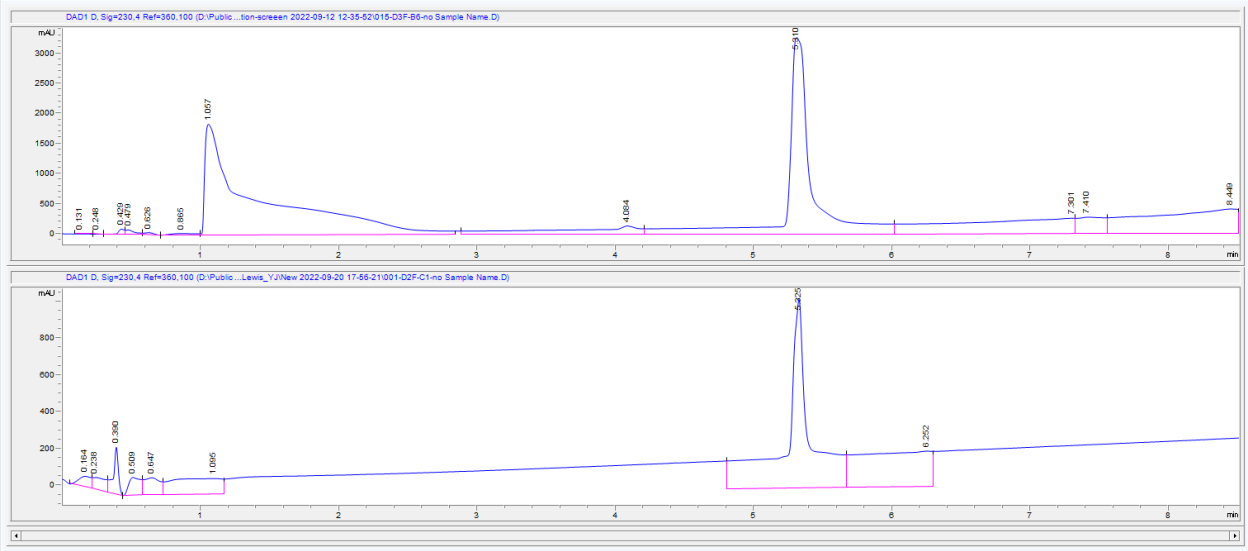
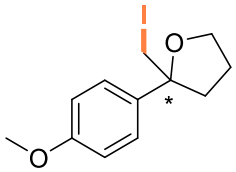


Figure S20. Comparison of chromatograms (254 nm) of **11b** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

Product **15a**:



The general procedure for bromination using NBS and flash column chromatography (silica gel, Hex/EA = 20/1) were used to obtain product **15a** as a white solid. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.33 – 7.29 (m, 2H), 6.89 – 6.85 (m, 2H), 4.06 (q, *J* = 7.1 Hz, 1H), 3.91 (td, *J* = 8.0, 5.7 Hz, 1H), 3.81 (s, 3H), 3.56 – 3.50 (m, 2H), 2.36 – 2.26 (m, 2H), 2.06 (tdd, *J* = 13.1, 7.7, 5.5 Hz, 1H), 1.88 – 1.79 (m, 1H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 158.92, 135.98, 126.74, 113.73, 84.47, 68.52, 55.39, 37.59, 26.39, 19.68. HRMS (EI-MS) Calc. for $[C_{12}H_{15}IO_2]^+$ M^+ : 318.0117, found 318.0110. This material was used to validate the selectivity of material produced using the general procedure for analytical-scale AetF reactions (Figure S21/22). Yield of AetF-catalyzed reaction: 81%. Product (P) assay yields (AY) were determined relative to internal standard (IS) with appropriate response factors (RF) using the formula AY = RF•(integral P/integral IS)+C.

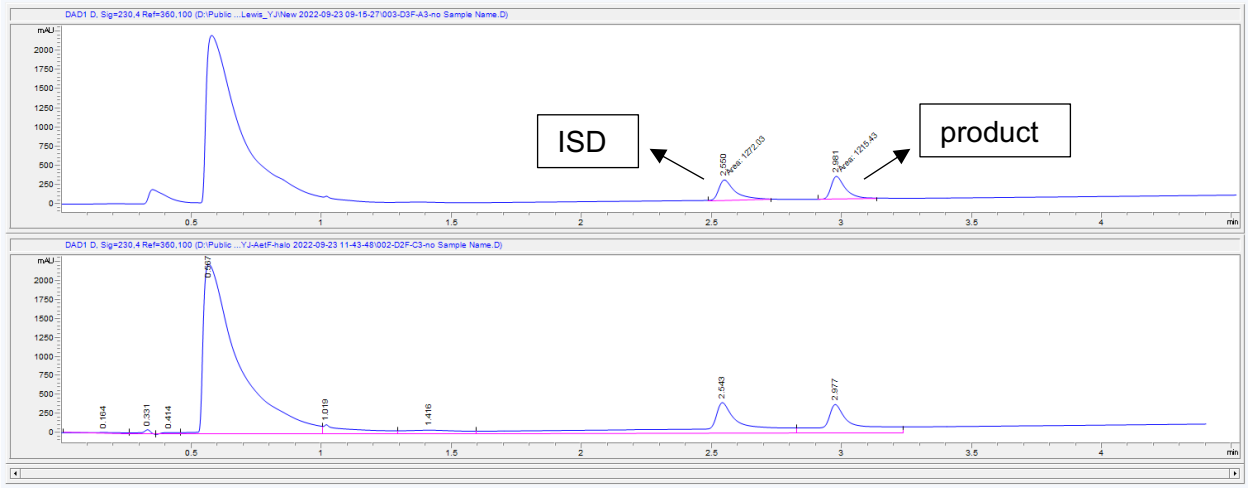


Figure S21. Comparison of chromatograms (230 nm) of **15a** produced using general procedures for analytical-scale AetF reactions (top) and NBS reactions (bottom).

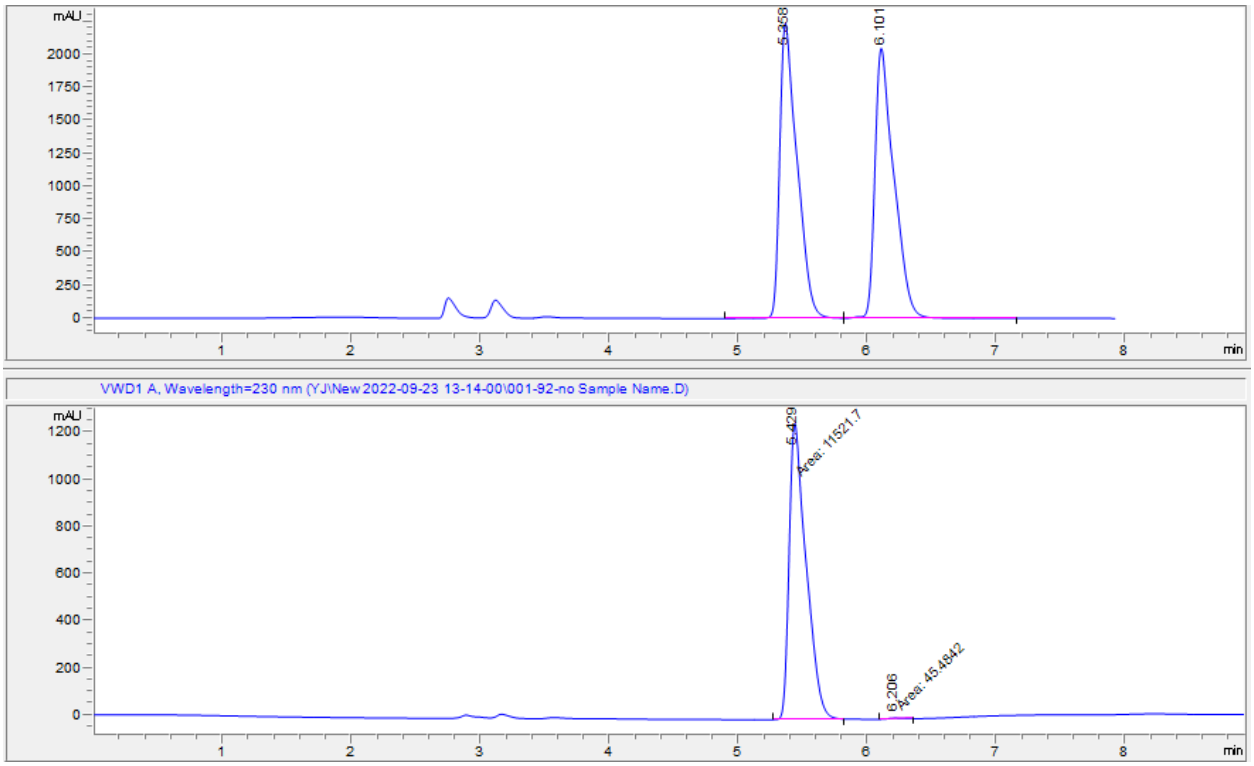


Figure S22. Comparison of chromatograms (254 nm) of **15a** produced using general procedures for NBS reactions (top) and analytical-scale AetF reactions (bottom). Based on these, the e.r. value of enzymatic material is 99:1.

VI. HalA Calculations

Helenium affinity (HalA) of a given compound **(A)** at site **(a)** as a quantitative descriptor of the bond strengths to halenium ions are defined as reaction enthalpy of the following reaction:

$$A + Cl^{+} \rightarrow A \cdot Cl^{+} \qquad HalA = \Delta H \qquad (eq1)$$

Workflow for HalA calculation was derived from Borhan et al.[11] and summarized here:

1. 3D structure generation for molecule **A** from SMILES strings using distance geometry method through RDKit
2. Conformer searching with extensive meta-dynamic sampling (MTD), using GFN2-xtb level of semi-empirical methods[12]
3. Structure optimization and frequency calculation for the lowest lying conformer from step2 with B3LYP/6-31G* level of theory in the gas phase. Imaginary frequencies were checked and fixed to make sure the optimized structure converges into the minimum point.
4. Taking the optimized structure from step 3, detect all aromatic SP2 carbons, and attach a triplet Cl^{+} ion to the SP2 carbon. This will generate a $\boldsymbol{A \cdot Cl^{+}}$ complex for each aromatic SP2 carbon site. For aromatic molecules with chiral faces, two complexes will be generated with each face being attacked.
5. Structure optimization and frequency calculation for all $\boldsymbol{A \cdot Cl^{+}}$ complexes generated in step 4 with B3LYP/6-31G* level of theory in the gas phase. Imaginary frequencies were checked and fixed
6. HaLA for each SP2 carbon were calculated through eq1

A python workflow was implemented to automate above steps and run molecules in batch.

VII. Computational Modeling

The online AlphaFold Colab notebook[13] was used to construct a model of the full-length AetF sequence noted above. The top-ranked structure (see coordinates below) was used to query the DALI protein structure comparison server[14]. The top 5 hits (unique PDBs) were structures of ancestral flavin-containing monooxygenases (FMOs) that were predicted based on the sequences of extant mammalian FMOs (Table S2).

Table S2. DALI results obtained using the AlphaFold-generated structure of AetF

Index	PDB(Chain)	Z	rmsd	lali	nres	%id	Description
1	6sem(A)	26.5	3.7	440	519	13	Ancestral FMO 2
2	6sf0(B)	25.8	3.8	434	530	12	Ancestral FMO 2
3	7al4(A)	25.0	3.1	432	517	15	Ancestral FMO 1
4	6se3(F)	24.7	3.1	437	528	12	Ancestral FMO 3-6
5	6sek(A)	23.9	3.0	426	504	15	Ancestral FMO 5

Structures of AetF were also prepared using trRosetta[15] and compared to the AlphaFold structures (Figure S23). Comparison to the AlphaFold structure reveals good agreement from residues 1-501 (alignment of truncated proteins in PyMOL: 4.347 Å RMSD after refinement;

10.642 Å RMSD before refinement;), however there is significant discrepancy between the C-terminal regions of the proteins (alignment of full-length proteins in PyMOL: 7.269 Å RMSD after refinement; 29.150 Å RMSD before refinement).

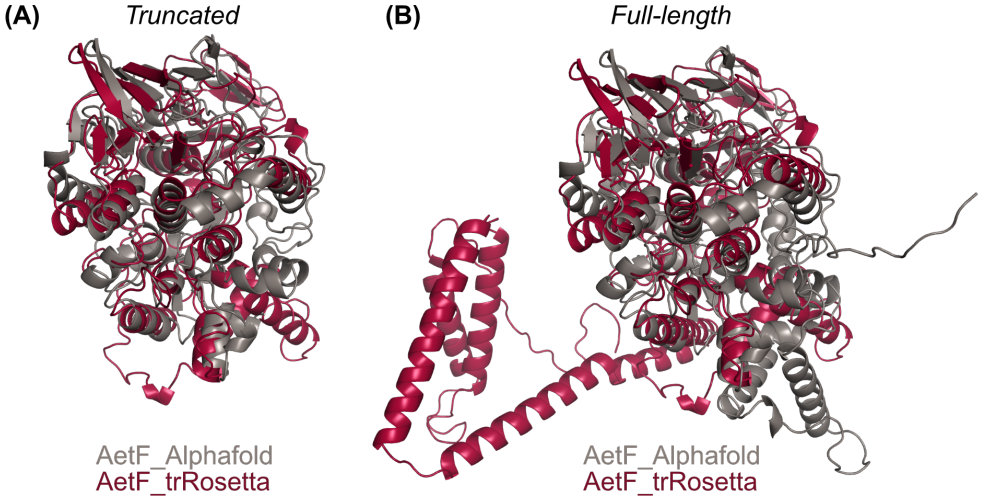


Figure S23. Alignment of **(A)** truncated or **(B)** full-length protein structures of AetF generated from AlphaFold (grey) and trRosetta (crimson). Alignment was performed using the cmd.align in PyMOL[16] version 2.5.0.

The C-terminal region (residues 502 – 668) of AetF not only varied between the generated structures but also substantially deviated from the ancestral FMOs found during the DALI search. Furthermore, it is known that C-terminal helices of the structurally homologous proteins 6sf0 and 6se3 are known to anchor the proteins to membranes.[17] Given this, we predicted that truncation of the C-terminus of AetF could simplify further computations and improve soluble heterologous expression. Using the AlphaFold model of AetF, the structures of the protein truncated after different residues (490, 494, 495, and 501) were relaxed with the Rosetta forcefield[18].

All Rosetta jobs were performed using on Indiana University's Cray XC40 supercomputer, Big Red 3, and data was stored within Slate, a centralized, high-performance Lustre file system. PDB files of starting structures for the different truncated proteins were first prepared in Notepad by deleting the appropriate C-terminus residues. The protein structures were then formatted for use in Rosetta using the script clean_pdb.py found within the rosetta_tools/protein_tools/scripts/ directory. The following command was used to format the PDB file (chain A of AetF_AlphaFold), generating both PDB and FASTA files:

>python
~/rosetta/gnu/mpi/rosetta_bin_linux_2020.08.61146_bundle/main/tools/protein_tools/scripts/clean_pdb.py AetF_AlphaFold.pdb A

A general_relax_flags file was generated with the following options:

-nstruct 5
-default_repeats 5

A Job file specifying the executable (directed to ~/relax.static.linuxgccrelease), general_relax_flags file, and details required for submitting a batch job to Big Red 3 using the Slurm workload manager was prepared. The following command was used to start a job that generated 5 output models for each input structure:

/geode2/soft/hps/cle7/rosetta/gnu/mpi/rosetta_bin_linux_2020.08.61146_bundle/main/source/bin/relax.static.linuxgccrelease -s /N/slate/yszubi/Rosetta/2022/YSZ_05312022/Truncated/491/AetF_Trunc.pdb @general_relax_flags -overwrite -database /geode2/soft/hps/cle7/rosetta/gnu/mpi/rosetta_bin_linux_2020.08.61146_bundle/main/database >& output

Models were sorted by total score, available in the score.sc output file. The top model (truncated after residue 501) was selected based on lowest overall total score (Table S3).

Table S3. Top structures from relaxation of truncated AetF_AlphaFold models in Rosetta

Index	Truncation location	Trajectory #	Total score (REU)
1	490	2	-1488
2	494	3	-1516
3	495	3	-1514
4	501	4	-1520

Molecular docking simulations of cofactors FAD and NADP (Figure S24) were then sequentially performed within Rosetta using RosettaLigand[19,20]. The locations of bound FAD and $NADP^+$ in the ancestral FMO 6sf0 were used as starting points. After aligning AetF to 6sf0 in PyMOL, the cofactors were saved as individual PDB files. The PDB files were converted to SDF files in Avogadro (1.2.0)[21]. Next, the SDF file of the cofactor was converted to a PARAMS file using the script molfile_to_params.py located within the python/apps/public/ directory. As an example, the command used to generate PARAMS and PDB files of the FAD ligand for docking simulations is shown below:

>python ~/rosetta/gnu/mpi/rosetta_bin_linux_2020.08.61146_bundle/main/source/scripts/python/apps/public/molfile_to_params.py -n FAD FAD.sdf

The protein and the FAD ligand PDB files were then concatenated using the following command:

>cat AetF.pdb FAD.pdb > AetF_FAD.pdb

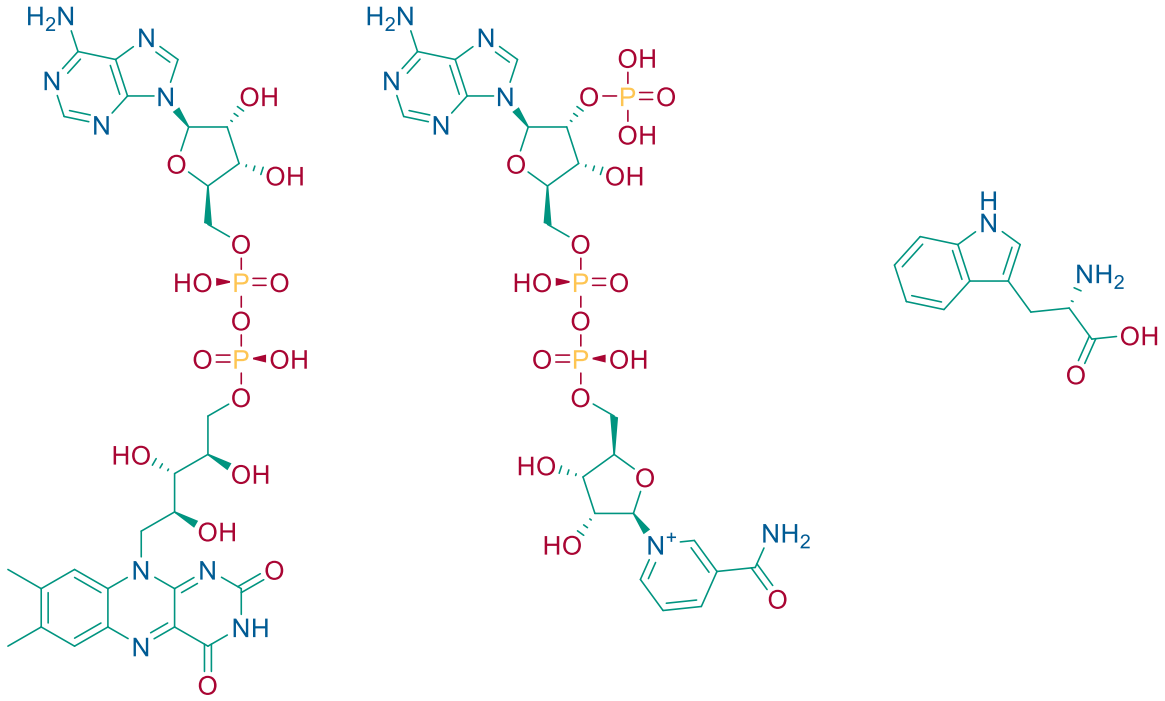


Figure 24. Structures of ligands docked into AetF.

An XML file (adapted from Combs et al.[20]) was then prepared with appropriate variables defined. The contents of the XML file were as follows:

```
<ROSETTASCRIPTS>

		<SCOREFXNS>
			<ScoreFunction name="ligand_soft_rep" weights="ligand_soft_rep">
			</ScoreFunction>
			<ScoreFunction name="hard_rep" weights="ligand">
			</ScoreFunction>
		</SCOREFXNS>

		<LIGAND_AREAS>
			<LigandArea name="inhibitor_dock_sc" chain="X" cutoff="6.0"
add_nbr_radius="true" all_atom_mode="false"/>
			<LigandArea name="inhibitor_final_sc" chain="X" cutoff="6.0"
add_nbr_radius="true" all_atom_mode="false"/>
			<LigandArea name="inhibitor_final_bb" chain="X" cutoff="7.0"
add_nbr_radius="false" all_atom_mode="true" Calpha_restraints="0.3"/>
		</LIGAND_AREAS>

		<INTERFACE_BUILDERS>
```

```
                <InterfaceBuilder name="side_chain_for_docking"
ligand_areas="inhibitor_dock_sc"/>
                <InterfaceBuilder name="side_chain_for_final"
ligand_areas="inhibitor_final_sc"/>
                <InterfaceBuilder name="backbone" ligand_areas="inhibitor_final_bb"
extension_window="3"/>
        </INTERFACE_BUILDERS>

        <MOVEMAP_BUILDERS>
                <MoveMapBuilder name="docking"
sc_interface="side_chain_for_docking" minimize_water="false"/>
                <MoveMapBuilder name="final" sc_interface="side_chain_for_final"
bb_interface="backbone" minimize_water="false"/>
        </MOVEMAP_BUILDERS>

        <SCORINGGRIDS ligand_chain="X" width="30">
                <ClassicGrid grid_name="classic" weight="1.0"/>
        </SCORINGGRIDS>

        <MOVERS>
                <Transform name="transform" chain="X" box_size="20.0"
move_distance="0.2" angle="45" cycles="500" repeats="1" temperature="5"/>
                <HighResDocker name="high_res_docker" cycles="6"
repack_every_Nth="3" scorefxn="ligand_soft_rep" movemap_builder="docking"/>
                <FinalMinimizer name="final" scorefxn="hard_rep"
movemap_builder="final"/>
                <InterfaceScoreCalculator name="add_scores" chains="X"
scorefxn="hard_rep"/>
        </MOVERS>

        <PROTOCOLS>
                <Add mover_name="transform"/>
                <Add mover_name="high_res_docker"/>
                <Add mover_name="final"/>
                <Add mover_name="add_scores"/>
        </PROTOCOLS>

</ROSETTASCRIPTS>
```

An Options file was generated next:

```
-in
        -file
                -s AetF.pdb'
                -extra_res_fa FAD.params
-packing
        -ex1
```

```
        -ex2
        -no_optH false
        -flip_HNQ true
        -ignore_ligand_chi true
-parser
        -protocol dock.xml
-overwrite
-mistakes
        -restore_pre_talaris_2013_behavior true
```

A Job file specifying the executable (directed to ~/rosetta_scripts.static.linuxgccrelease), Options file, and details required for submitting a batch job to Big Red 3 using the Slurm workload manager was prepared. The following command was used to start a job that generated 500 output structures with FAD docked into AetF:

```
>/geode2/soft/hps/cle7/rosetta/gnu/mpi/rosetta_bin_linux_2020.08.61146_bundle/main/source/bin/rosetta_scripts.static.linuxgccrelease @ options.txt -nstruct 500 >& output
```

Models were sorted by total score, available in the score.sc output file. A top model was selected based on overall total score (reflective of the overall quality of the protein-ligand model) and the interface delta X score (difference between ligand-bound and unbound ligand complex).

After obtaining a protein model of AetF with the bound FAD, $NADP^+$ was then docked. Molecular docking simulations were performed within Rosetta using RosettaLigand[19,20] as described above (*vide supra*) except with the changes described here. The receptor was the AetF_FAD model obtained from docking rather than apo AetF. The PDB of $NADP^+$ was concatenated together with both the PDB files of AetF and the FAD cofactor so that AetF was chain A, FAD was chain B, and the cofactor $NADP^+$ was chain X. The Options file was modified so that it included parameters for both FAD and $NADP^+$ (nicknamed here 'NDD'):

```
-in
        -file
                -s AetF.pdb'
                -extra_res_fa FAD.params
                -extra_res_fa NDD.params
```

Models again were sorted by total score, available in the score.sc output file. A top model was selected based on overall total score and the interface delta X score. Finally, after obtaining a protein model of AetF with the bound FAD and $NADP^+$, the native substrate tryptophan was then docked. Molecular docking simulations were performed within Rosetta using RosettaLigand[19,20] as described above (*vide supra*) except with the changes described here. The receptor was the AetF_FAD model obtained from docking both cofactors rather than apo AetF. However, only FAD was left in the simulation since $NADP^+$ is predicted to dissociate and allow substrate access to the active site (there is not reasonable space for both $NADP^+$ and tryptophan to occupy the active site simultaneously). After removing $NADP^+$ from the PDB, tryptophan was manually placed in the suspected binding site of $NADP^+$ in PyMOL. The PDB of tryptophan was obtained from a crystal structure of a protein (PDB: 1csm) with which it co-crystallized. After preparing the PARAMS file

for tryptophan, the PDB of the ligand was concatenated together with both the PDB files of AetF and the FAD cofactor so that AetF was chain A, FAD was chain B, and the substrate tryptophan was chain X. The Options file was modified so that it included parameters for both FAD and tryptophan (nicknamed here 'TYP'):

```
-in
        -file
                -s AetF.pdb'
                -extra_res_fa FAD.params
                -extra_res_fa TYP.params
```

Models were sorted by total score, available in the score.sc output file. Plots of the interface delta scores vs. total scores of poses for each of the ligands are shown in Figure S25.

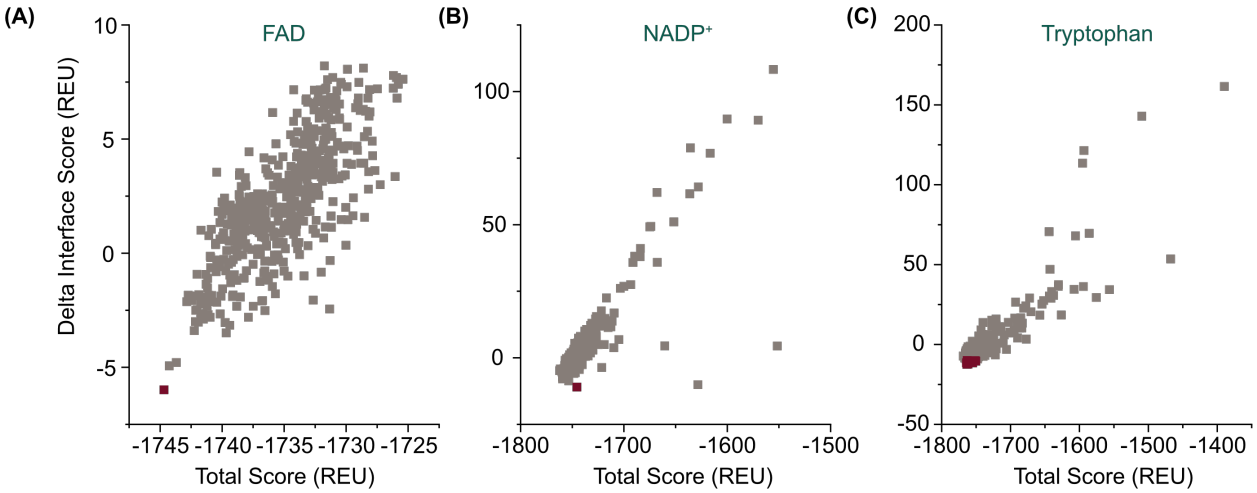


Figure S25. Scores of docked poses of **(A)** FAD, **(B)** $NADP^+$, and **(C)** Tryptophan in AetF. The top poses used for subsequent docking are highlighted in **(A)** and **(B)** as crimson points. The top 10 poses from docking of the substrate tryptophan are also highlighted in **(C)** as crimson points.

The top 10 poses of tryptophan (interface delta score) are also visualized (Figure S26).

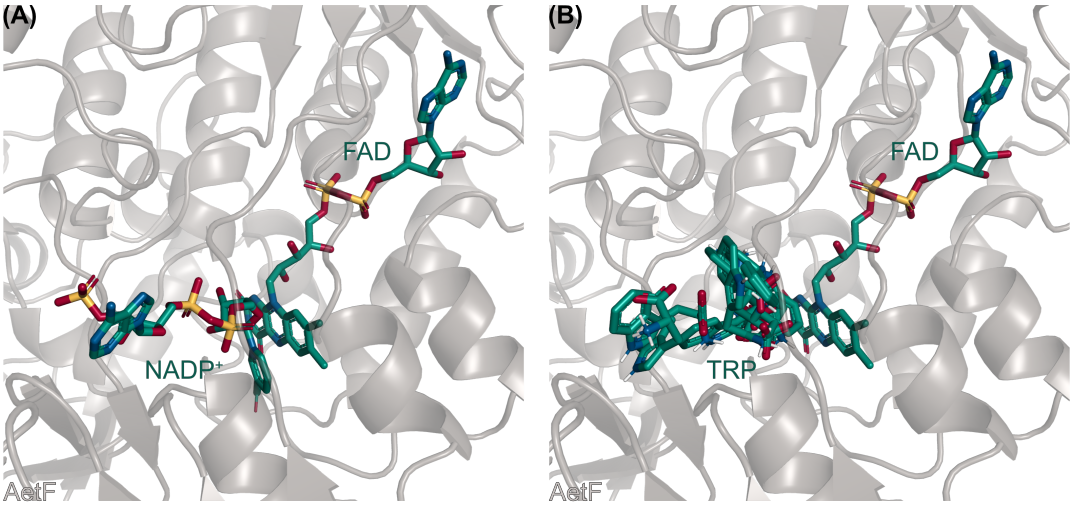


Figure S26. **(A)** Top model of AetF with FAD and $NADP^+$ docked. **(B)** Top 10 poses of tryptophan (TRP) docked into AetF_FAD.

Notably, the model of AetF contains a lysine residue (K158) that projects into the active site and is 5.4 Å from C_5 of tryptophan (Figure S27). This is similar to the relative positioning of the catalytic lysine residue (K79) compared to tryptophan in the well-characterized flavin-dependent halogenase, RebH[22], in which the lysine is 4.1 Å from C_7 of the substrate.

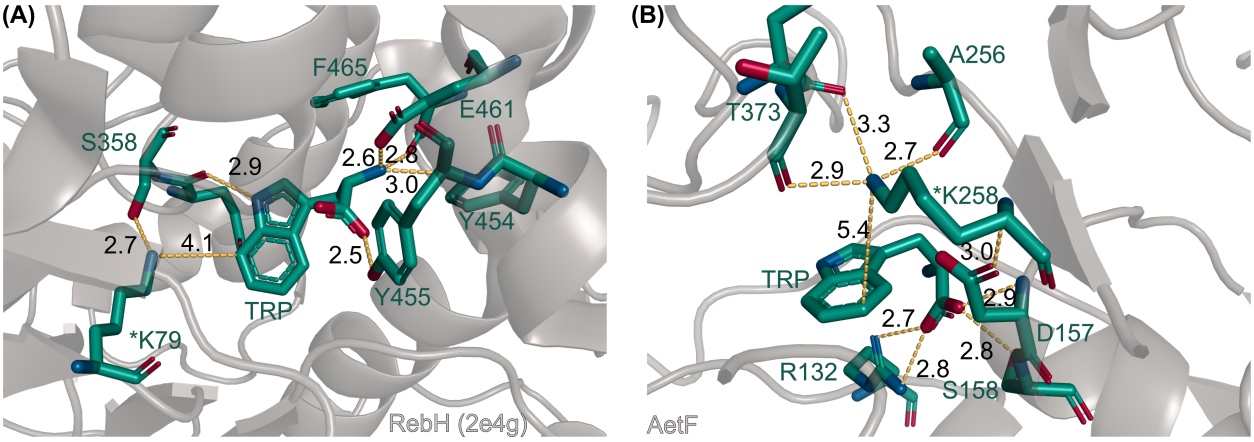
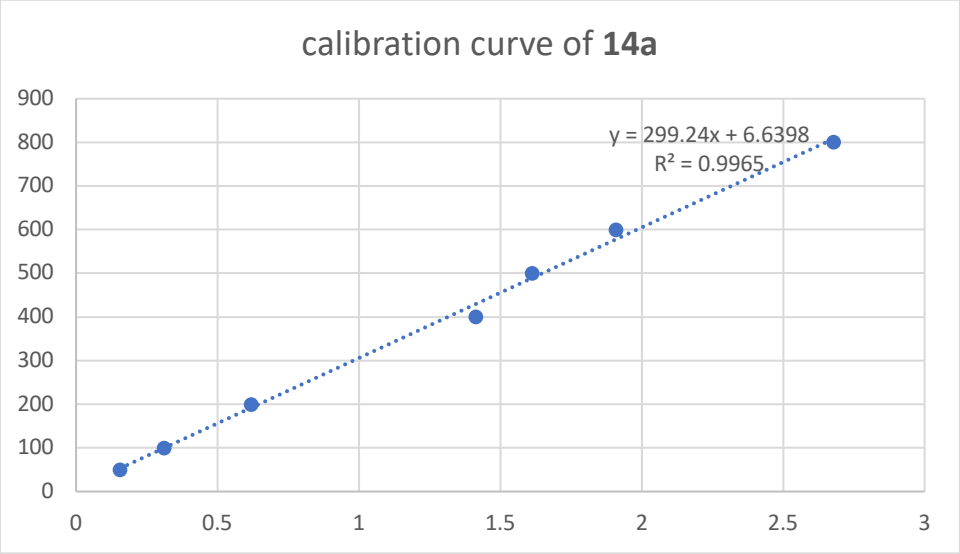


Figure 27. Binding of tryptophan (TRP) to the **(A)** flavin-dependent halogenase RebH (PDB: 2e4g)[22] and **(B)** AetF (from docking). Close contacts to the catalytic lysine residue (*K79 and *K158 in RebH and AetF, respectively) and to the tryptophan substrate are highlighted.

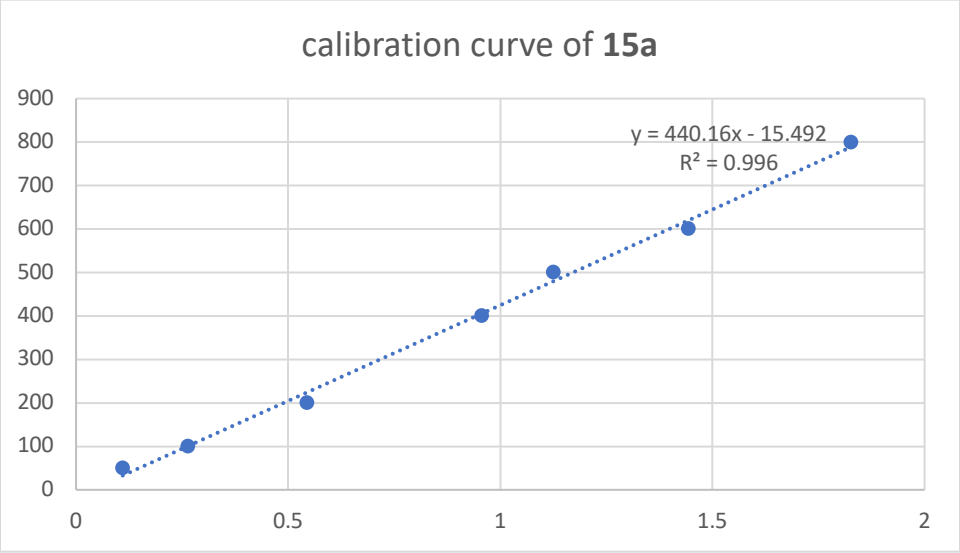
VII. LCMS Calibration Curves

Calibration curves were used to obtain conversions for **14a** and **15a** since these are expected to differ significantly from the starting materials due to loss of the styrenyl olefin.

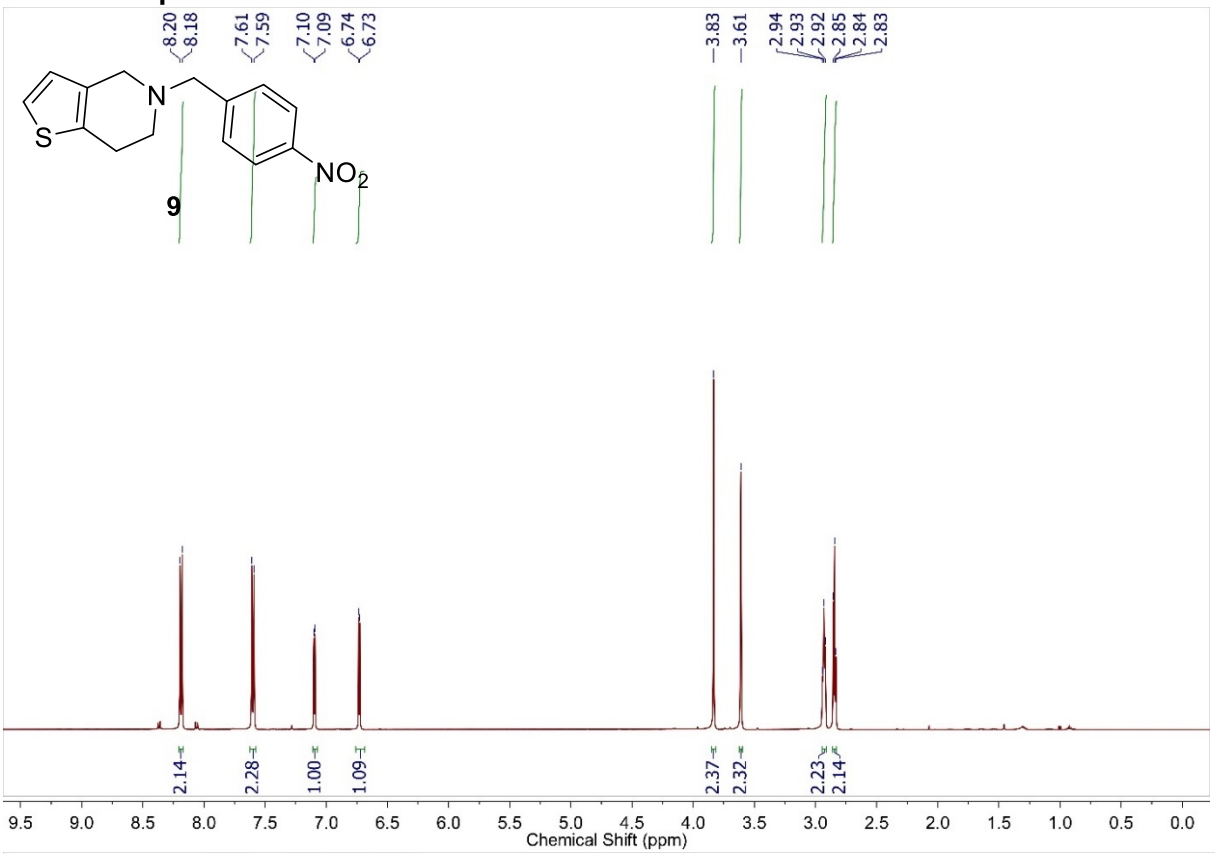
Calibration curve of product **14a** using 4-methoxybenzoate as internal standard:

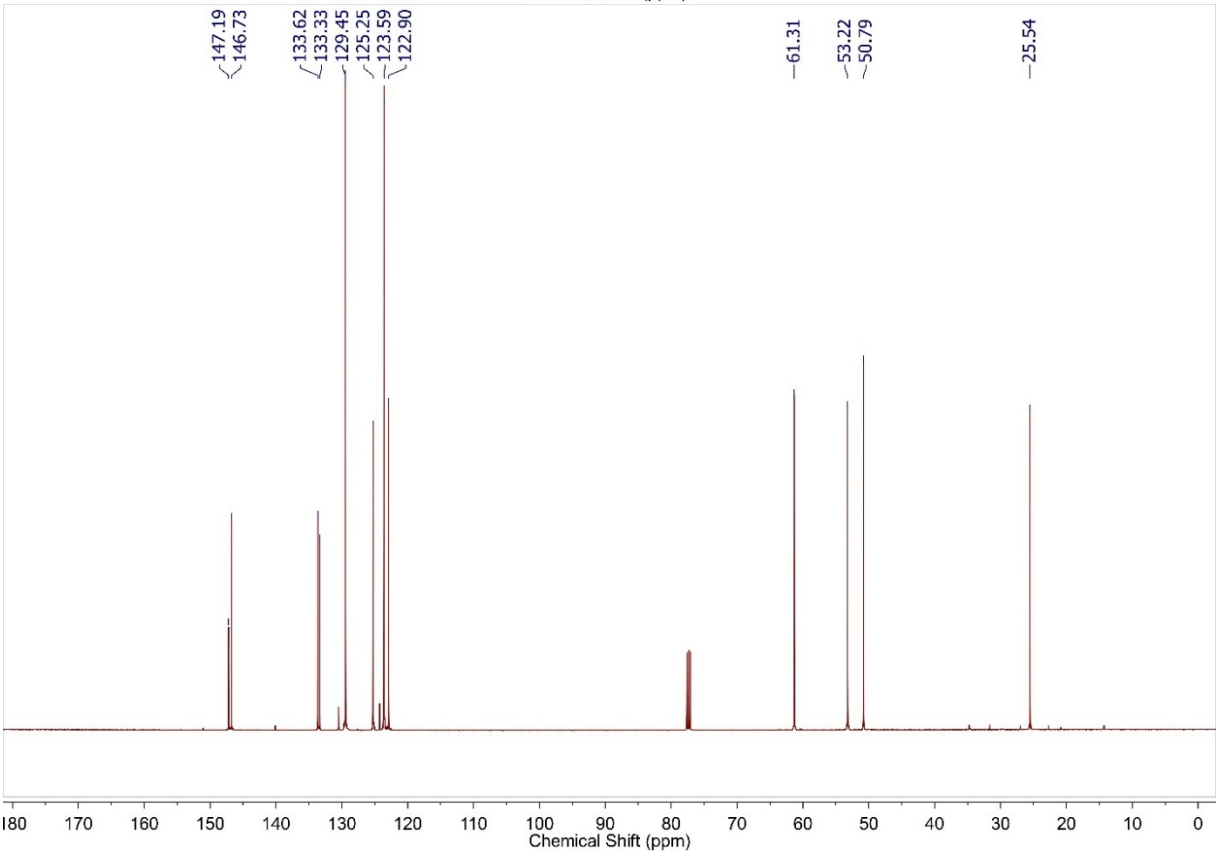


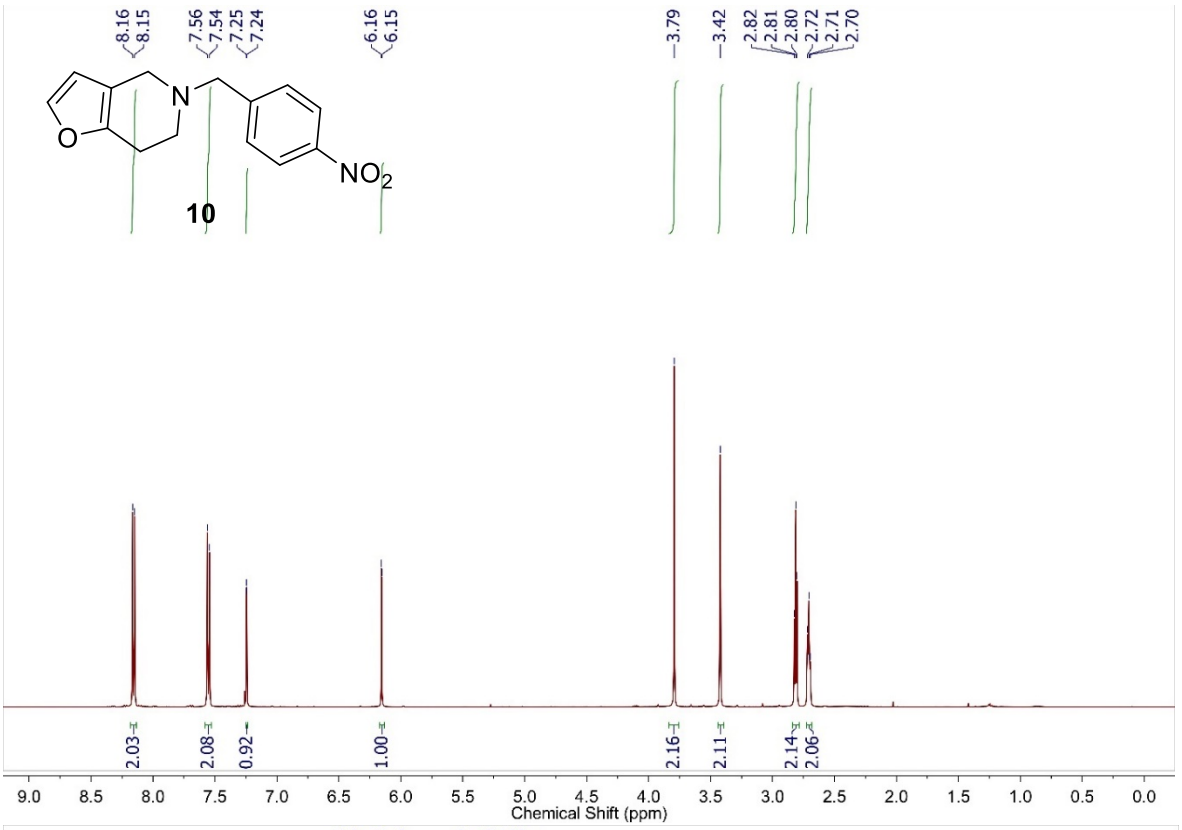
Calibration curve of product **15a** using 4-methoxybenzoate as internal standard:

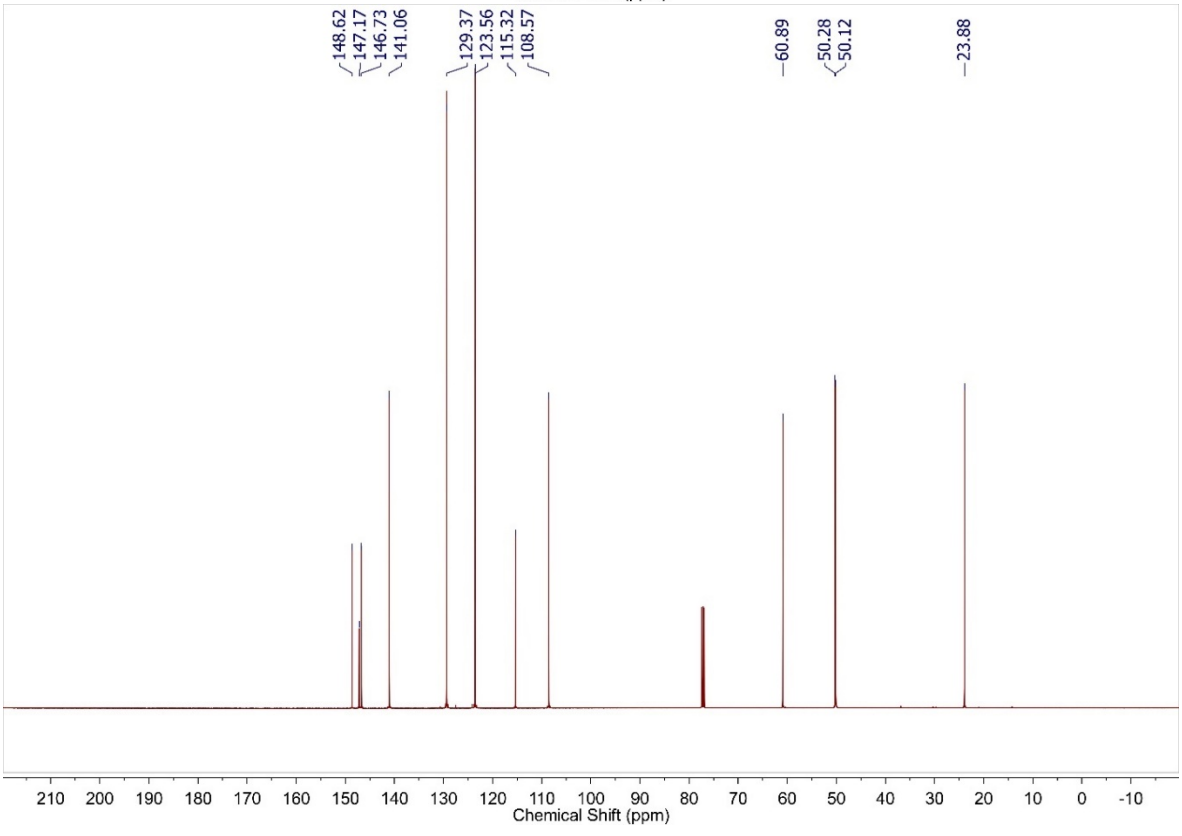


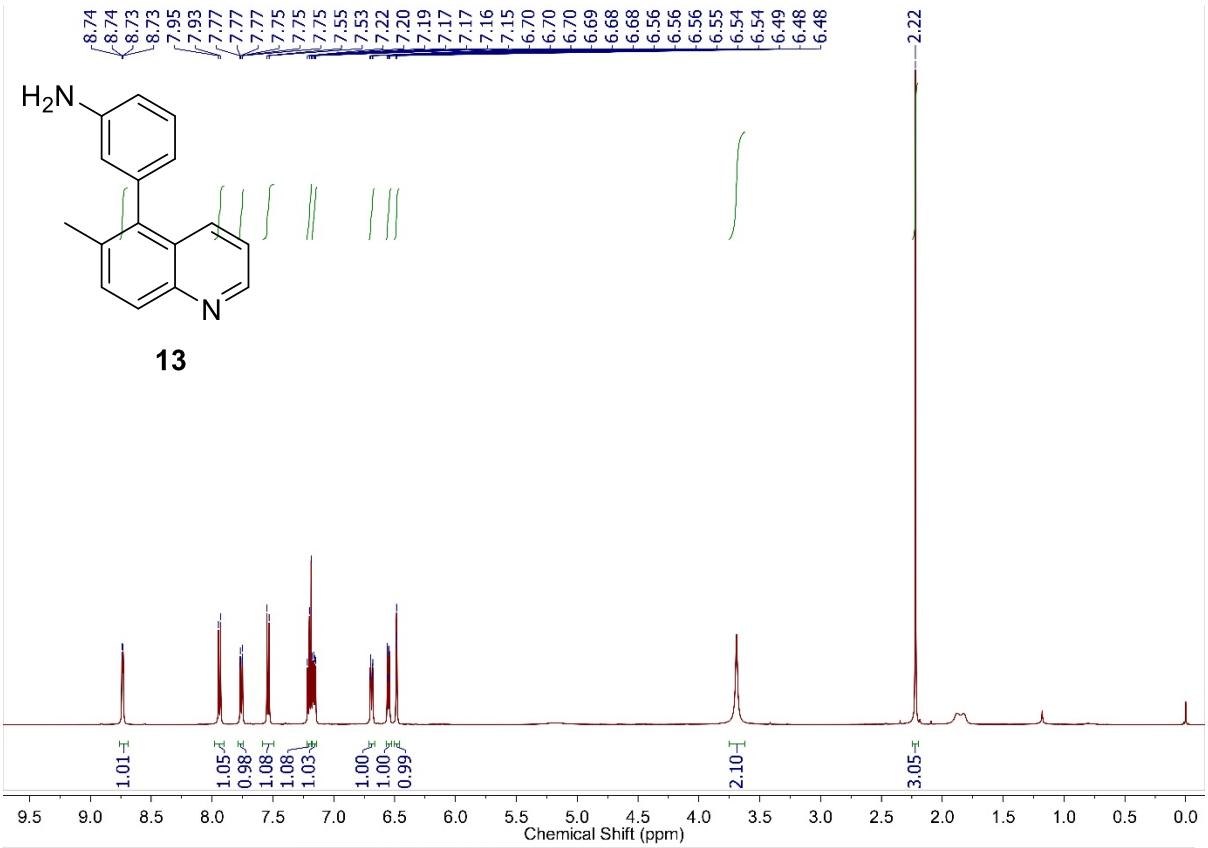
VIII. NMR Spectra



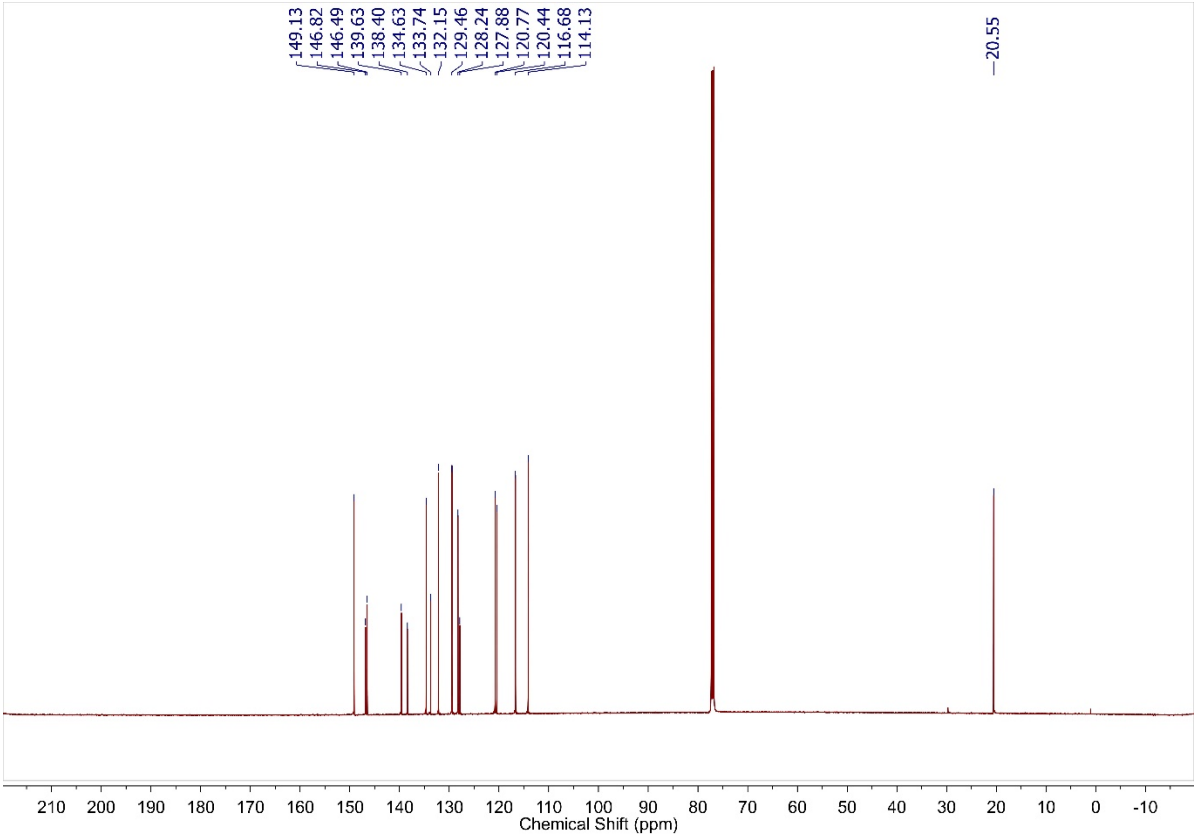


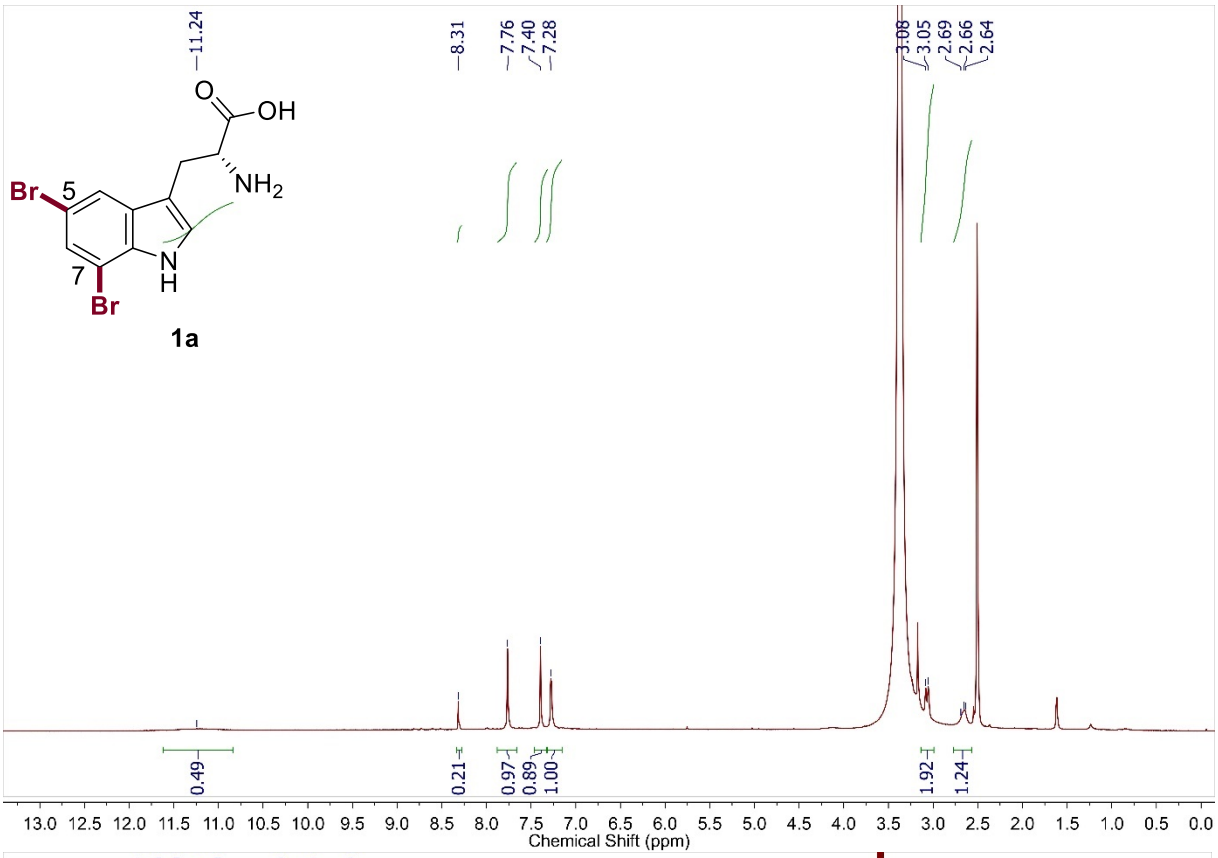

8.16
8.15
7.56
7.54
7.25
7.24
6.16
6.15
3.79
3.42
2.82
2.81
2.80
2.72
2.71
2.70
N
O
NO2
10
2.03
2.08
0.92
1.00
2.16
2.11
2.14
2.06
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
Chemical Shift (ppm)

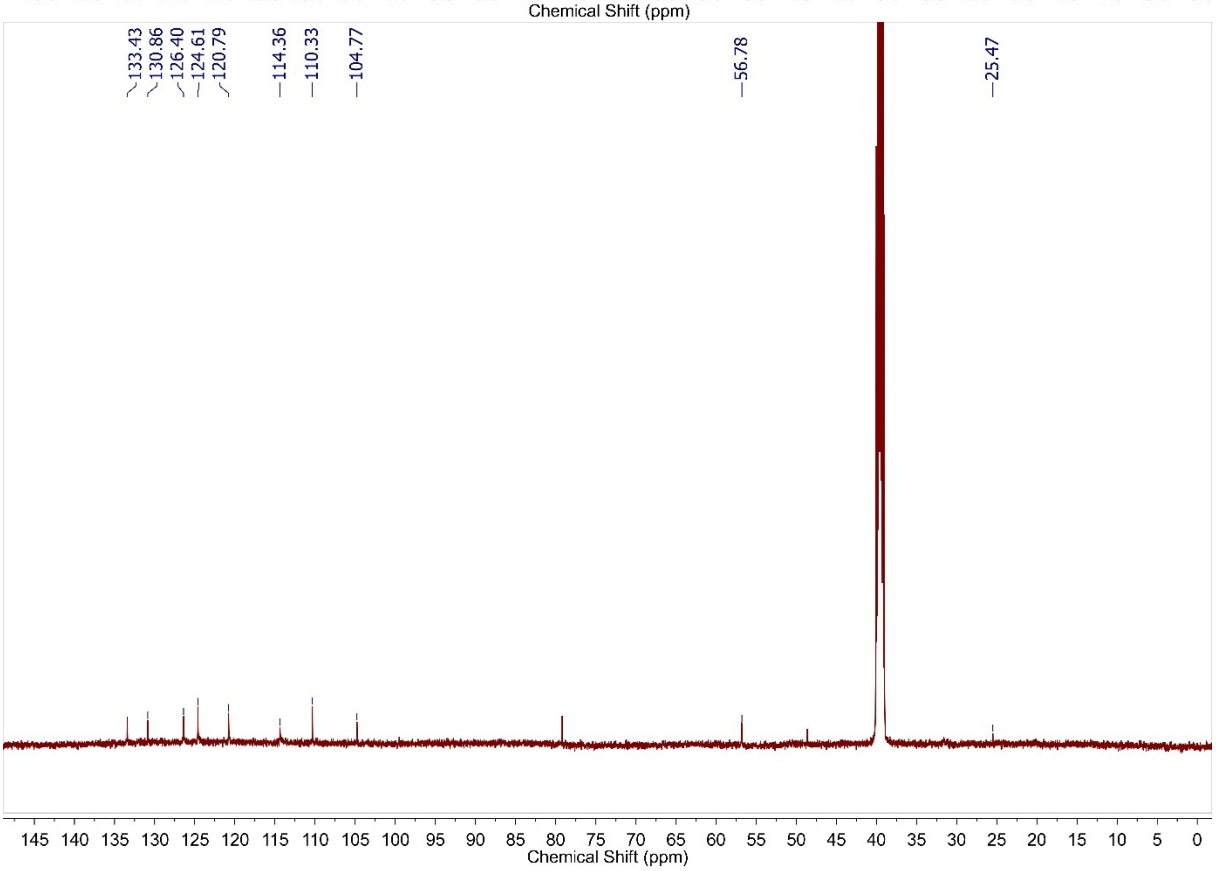

148.62
147.17
146.73
141.06
129.37
123.56
115.32
108.57
60.89
50.28
50.12
23.88
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
Chemical Shift (ppm)

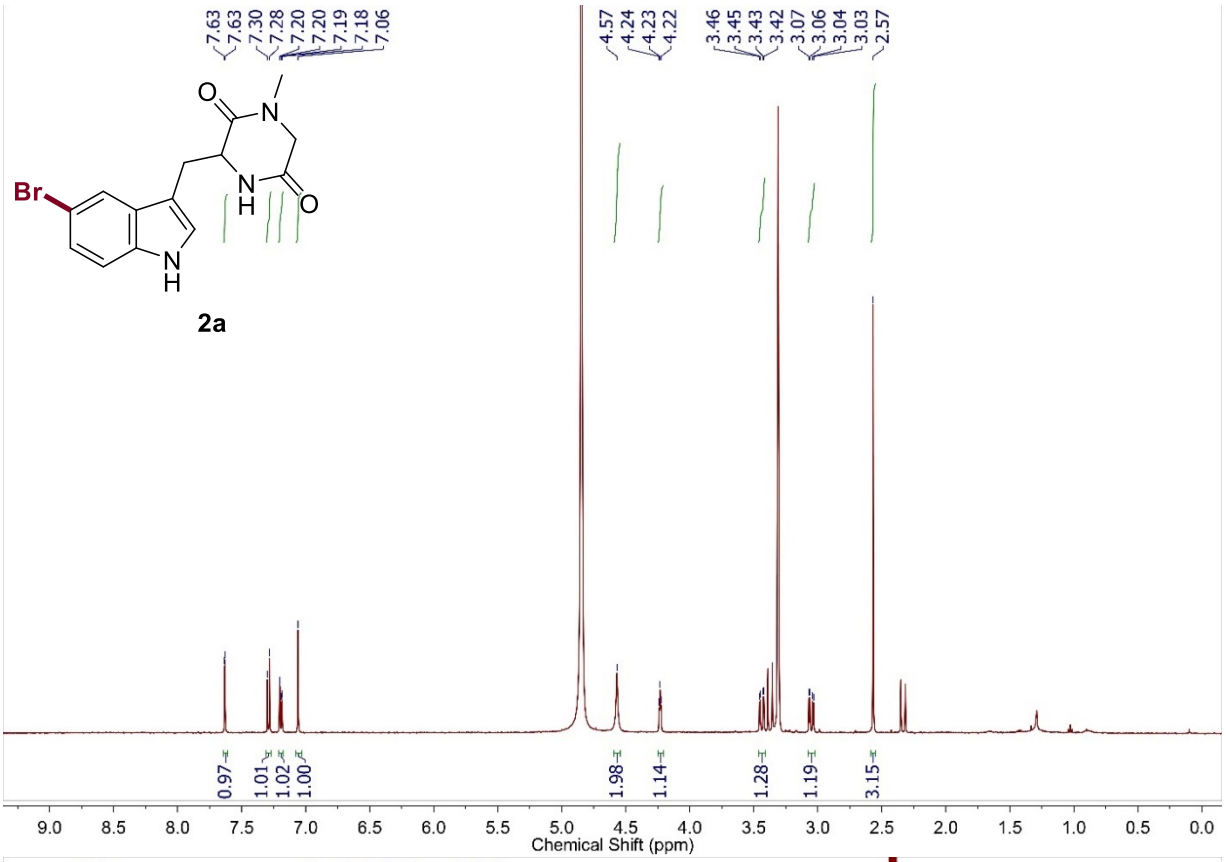

8.74
8.74
8.73
8.73
7.95
7.93
7.77
7.77
7.77
7.75
7.75
7.75
7.55
7.53
7.22
7.20
7.19
7.17
7.17
7.16
7.15
6.70
6.70
6.69
6.68
6.68
6.56
6.56
6.56
6.55
6.54
6.54
6.49
6.48
6.48
2.22
H2N
N
1.01
1.05
0.98
1.08
1.08
1.03
1.00
1.00
0.99
2.10
3.05
Chemical Shift (ppm)

13

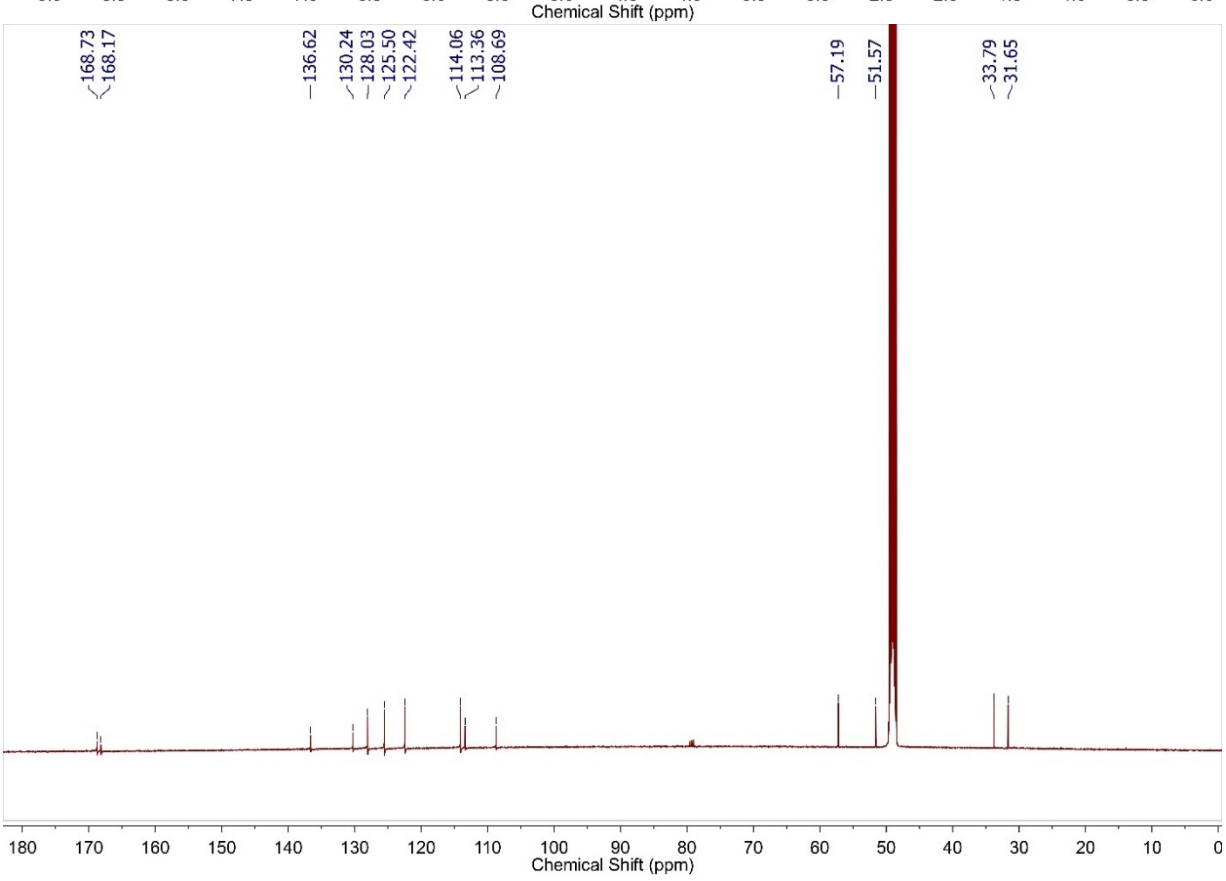

149.13
146.82
146.49
139.63
138.40
134.63
133.74
132.15
129.46
128.24
127.88
120.77
120.44
116.68
114.13
20.55
Chemical Shift (ppm)

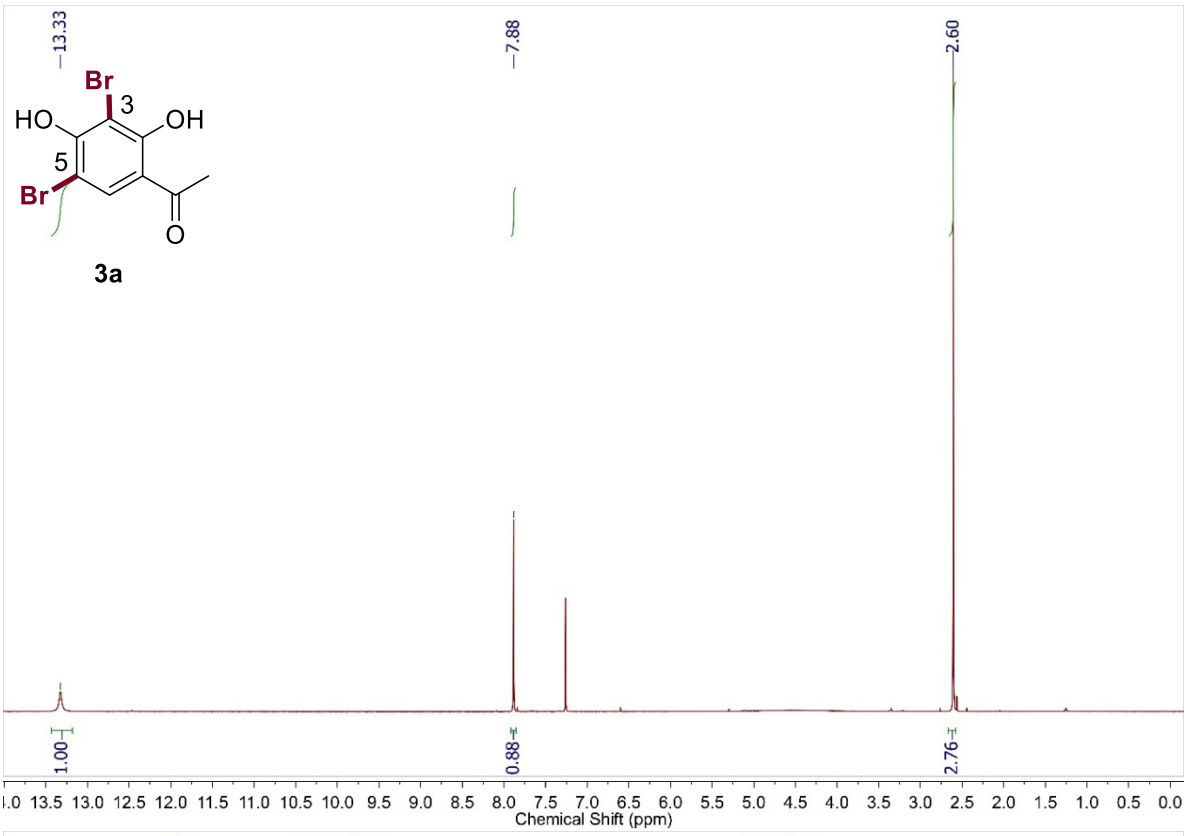

11.24
8.31
7.76
7.40
7.28
3.08
3.05
2.69
2.66
2.64
O
OH
NH2
Br
5
7
N
H
Br
1a
0.49
0.21
0.97
0.89
1.00
1.92
1.24
13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
Chemical Shift (ppm)

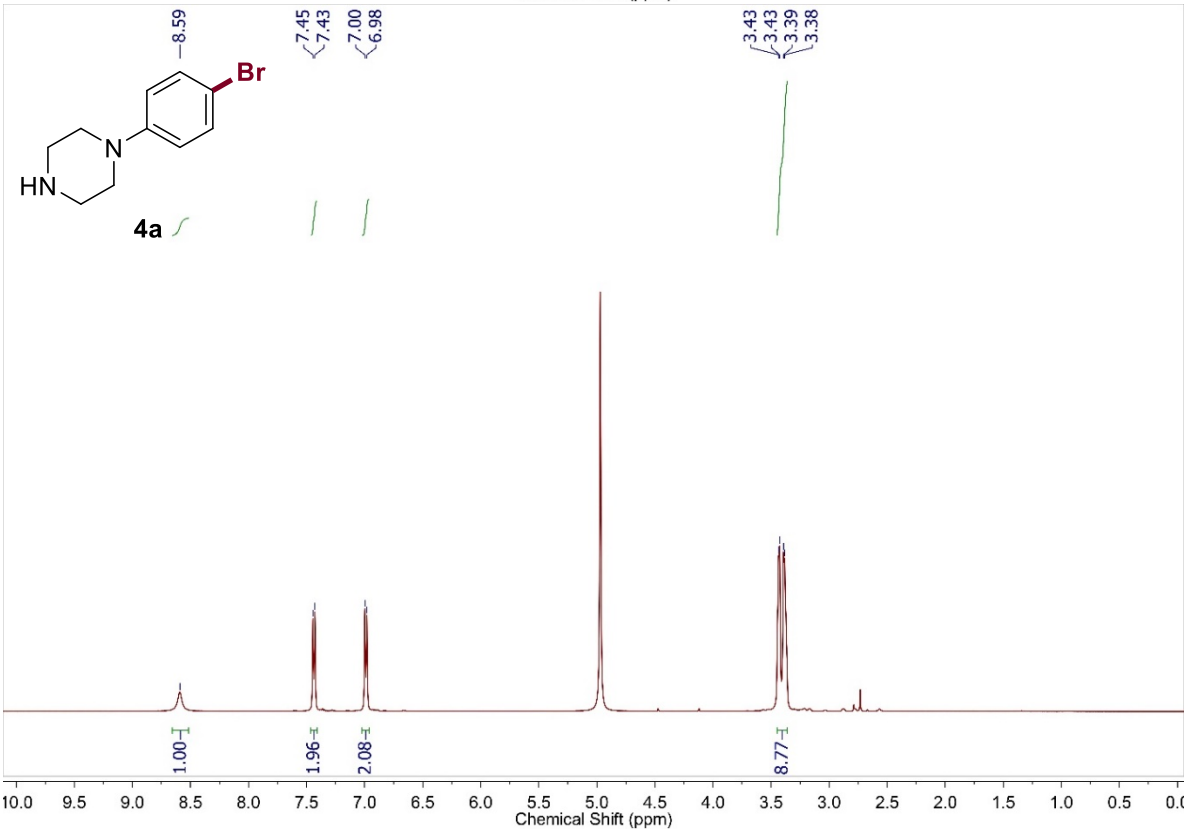

133.43
130.86
126.40
124.61
120.79
114.36
110.33
104.77
56.78
25.47
145 140 135 130 125 120 115 110 105 100 95 90 85 80 75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0
Chemical Shift (ppm)

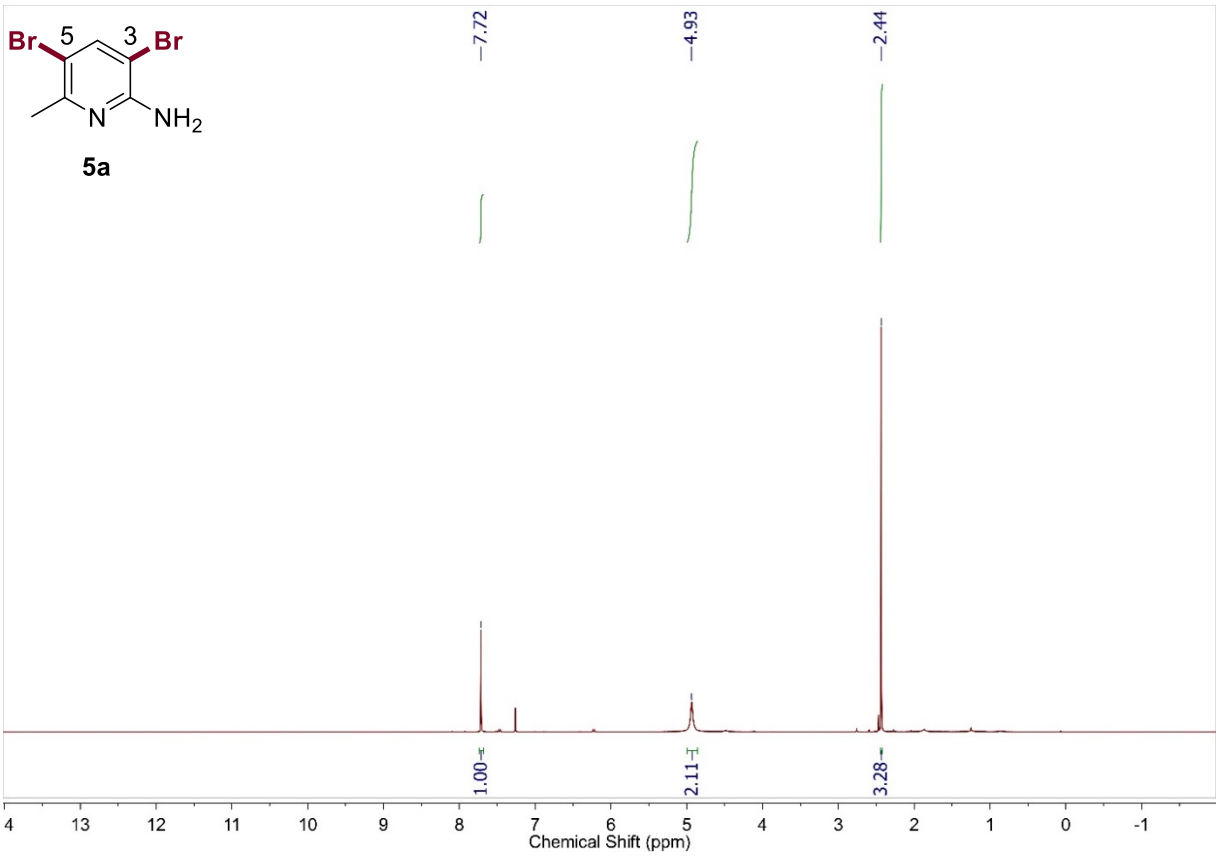

7.63
7.63
7.30
7.28
7.20
7.20
7.19
7.18
7.06
4.57
4.24
4.23
4.22
3.46
3.45
3.43
3.42
3.07
3.06
3.04
3.03
2.57
Br
O
N
N
H
O
N
H
0.97
1.01
1.02
1.00
1.98
1.14
1.28
1.19
3.15
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
Chemical Shift (ppm)

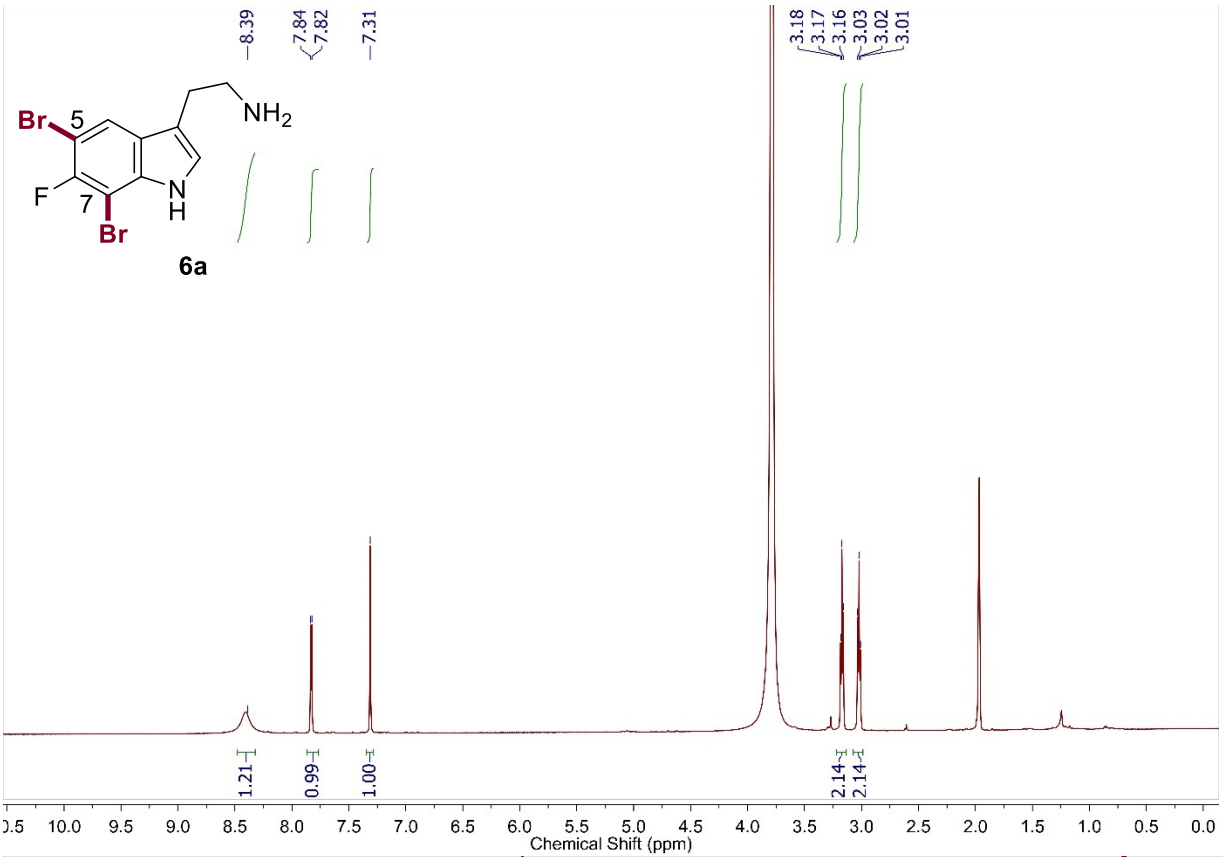
2a

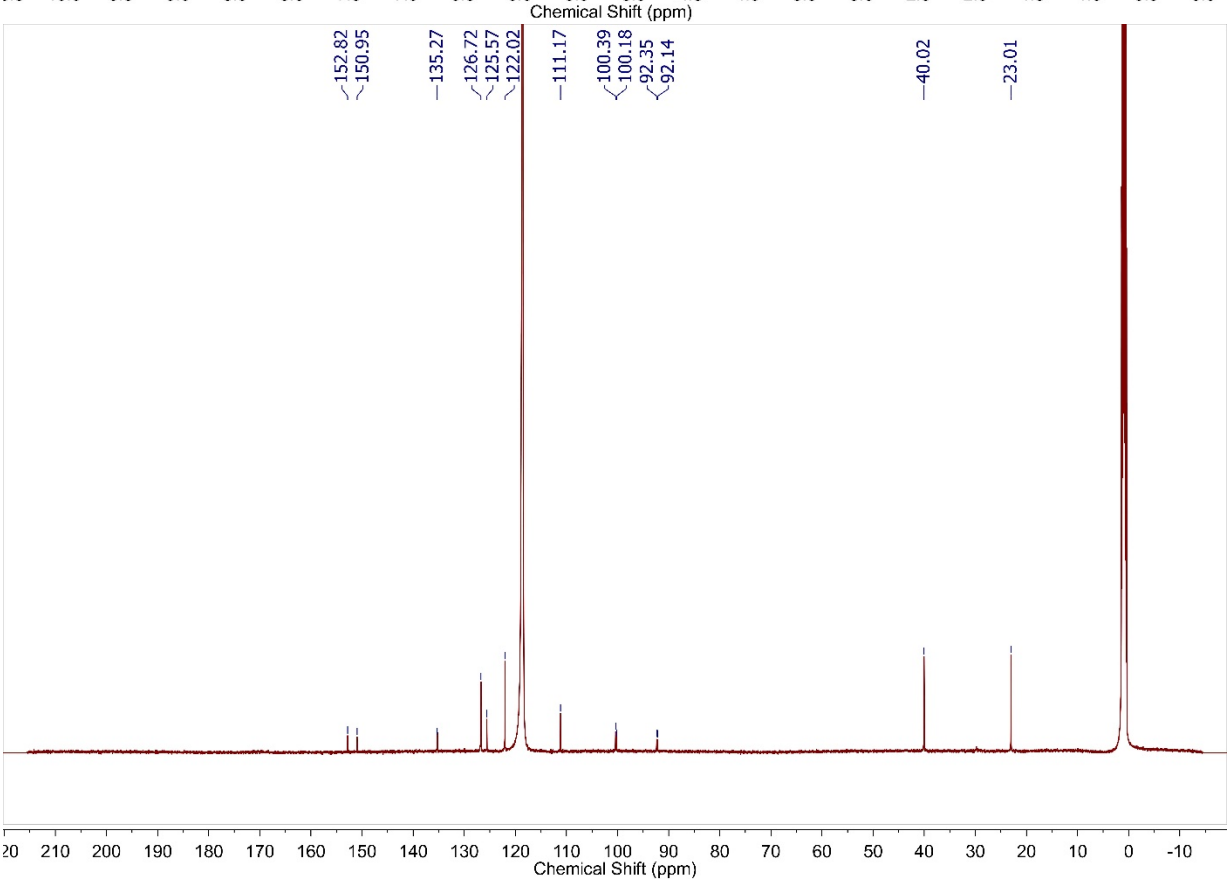

168.73
168.17
136.62
130.24
128.03
125.50
122.42
114.06
113.36
108.69
57.19
51.57
33.79
31.65
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
Chemical Shift (ppm)

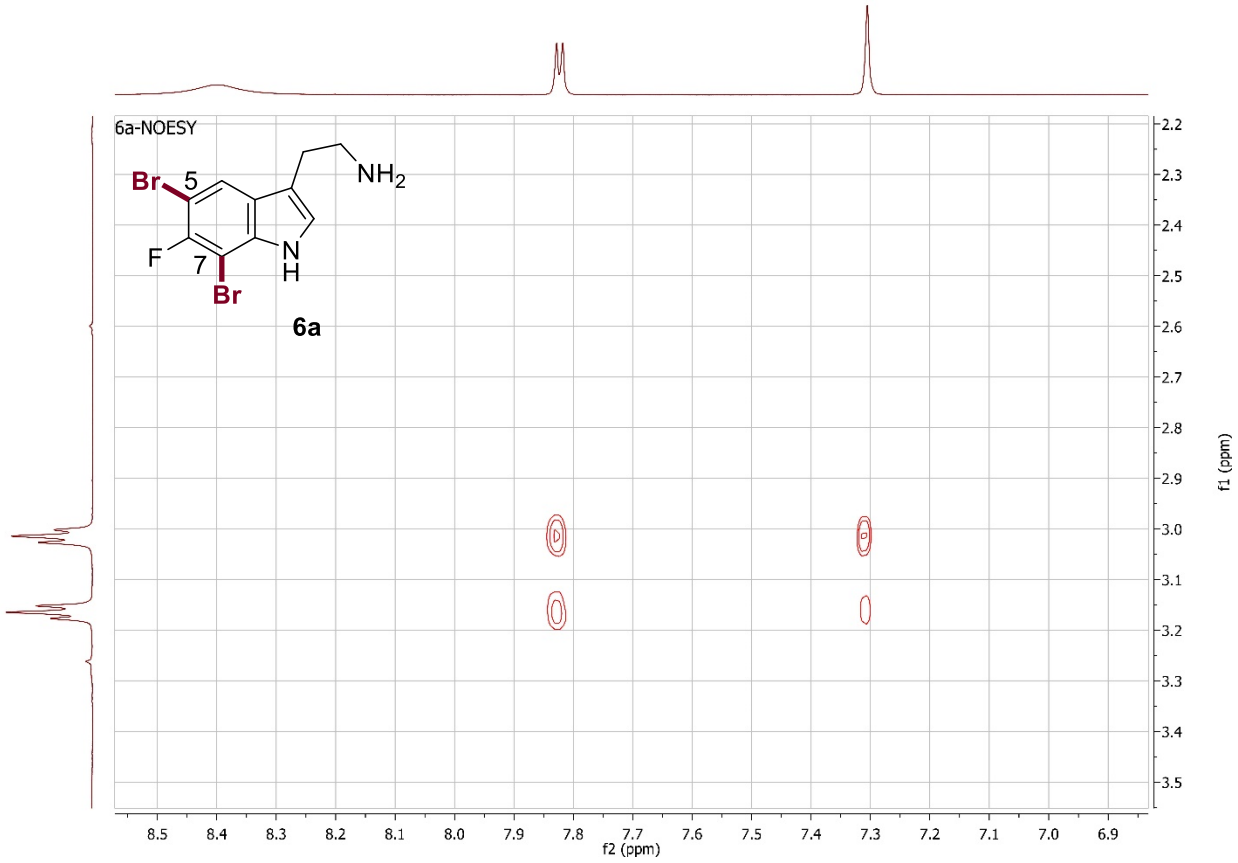

13.33
7.88
2.60
Br
3
HO
OH
5
Br
O
3a
1.00
0.88
2.76
Chemical Shift (ppm)

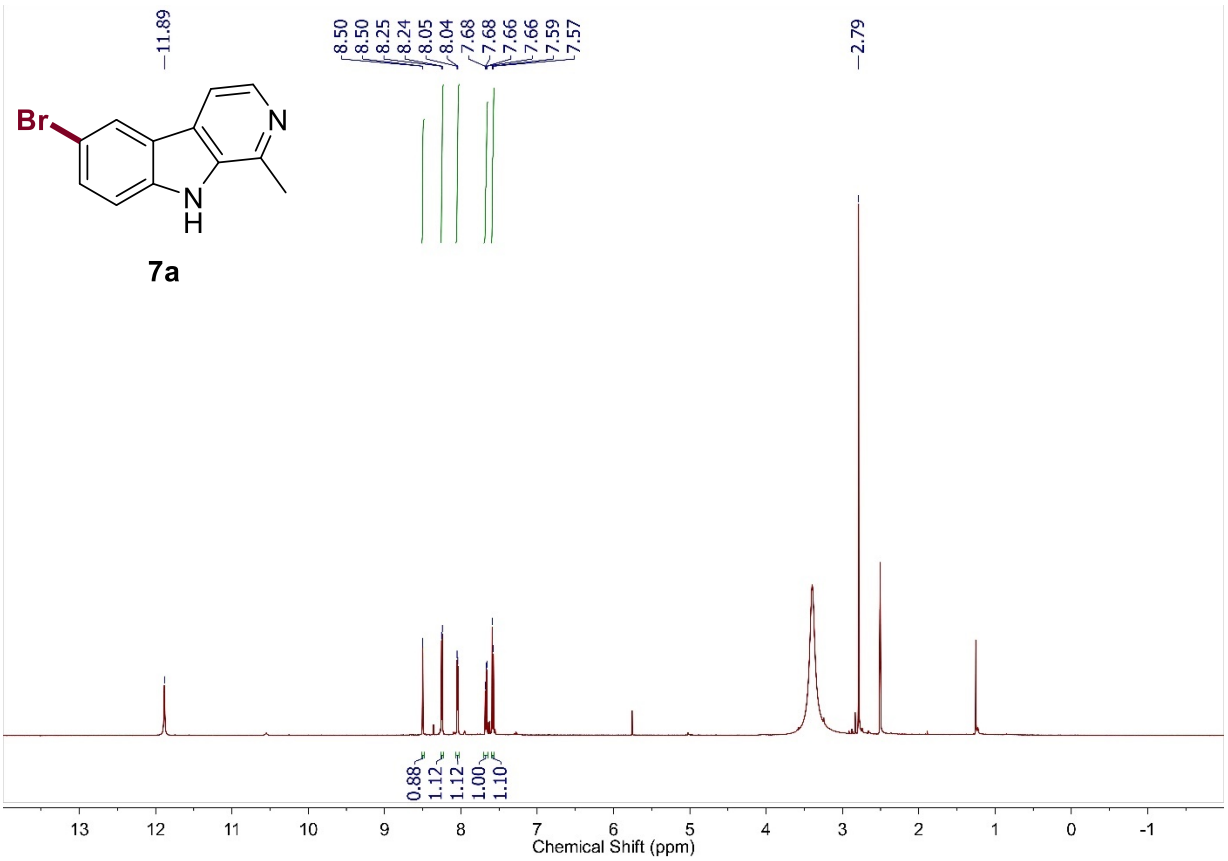

8.59
7.45
7.43
7.00
6.98
3.43
3.43
3.39
3.38
Br
N
HN
4a
1.00
1.96
2.08
8.77
Chemical Shift (ppm)

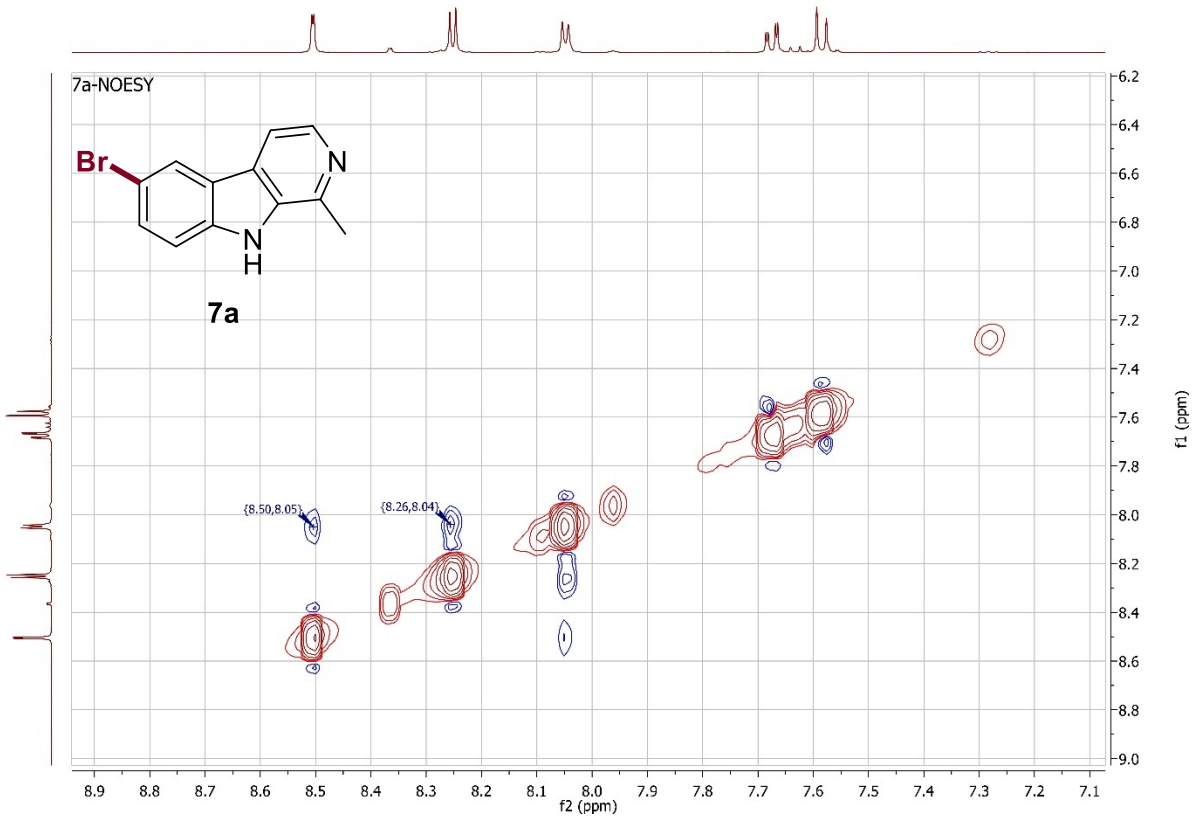

Br
5
3
Br
N
NH2
5a
7.72
4.93
2.44
1.00
2.11
3.28
4
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
Chemical Shift (ppm)

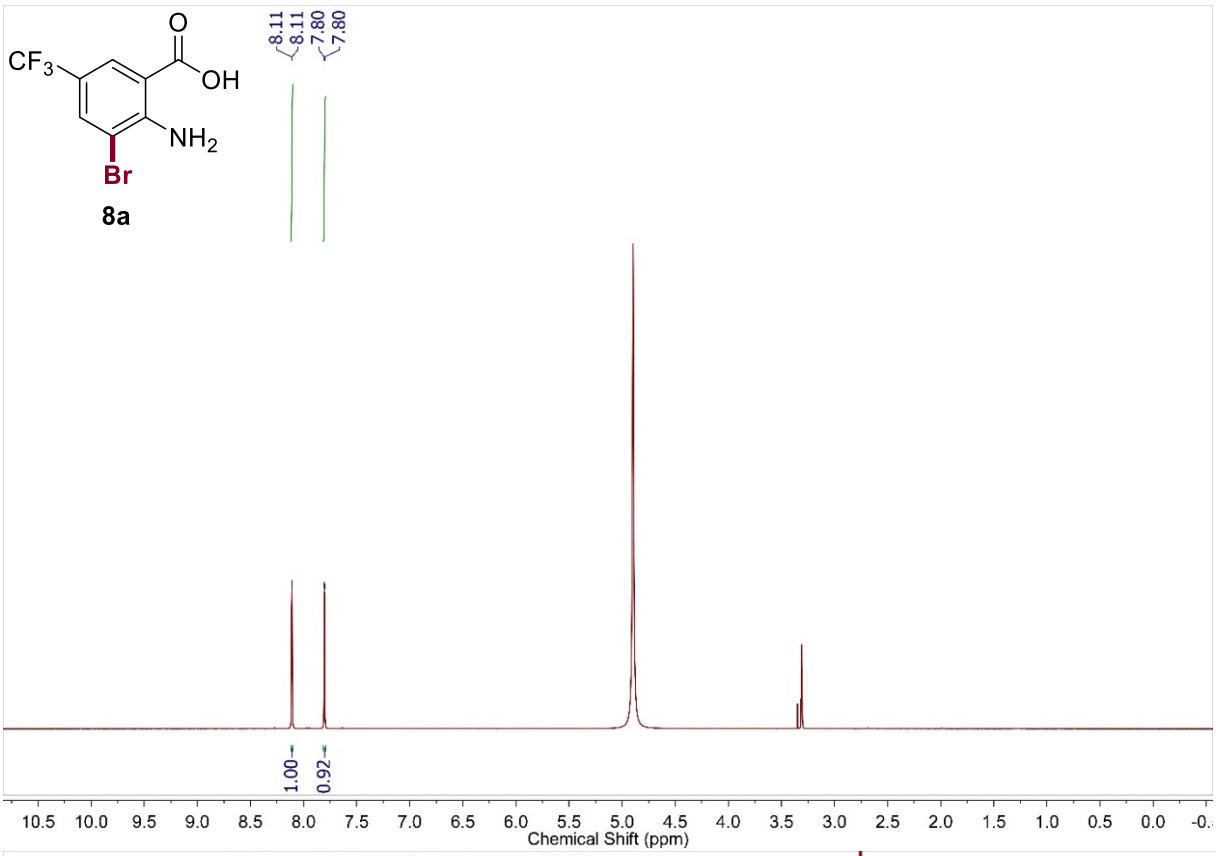

8.39
7.84
7.82
7.31
3.18
3.17
3.16
3.03
3.02
3.01
Br
5
NH2
F
7
N
H
Br
6a
1.21
0.99
1.00
2.14
2.14
Chemical Shift (ppm)

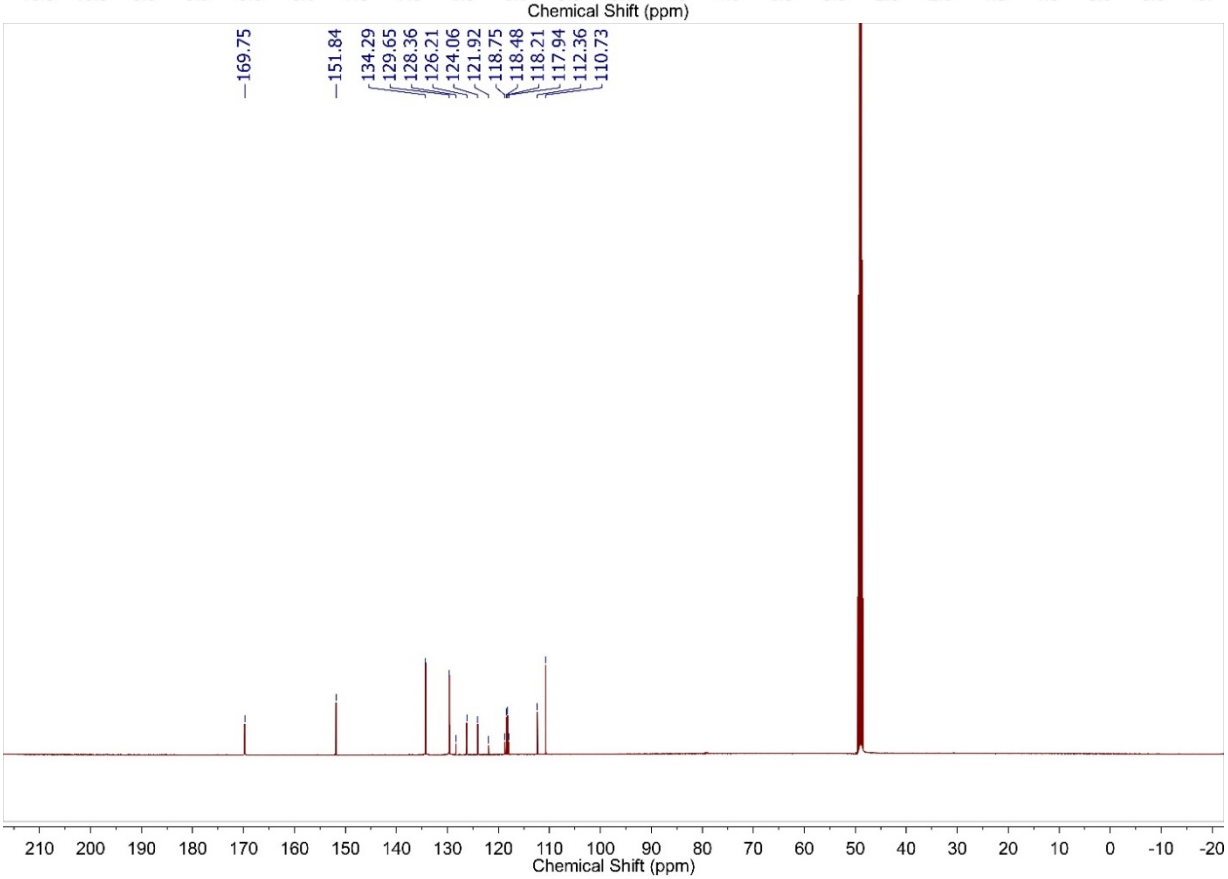

152.82
150.95
135.27
126.72
125.57
122.02
111.17
100.39
100.18
92.35
92.14
40.02
23.01
Chemical Shift (ppm)

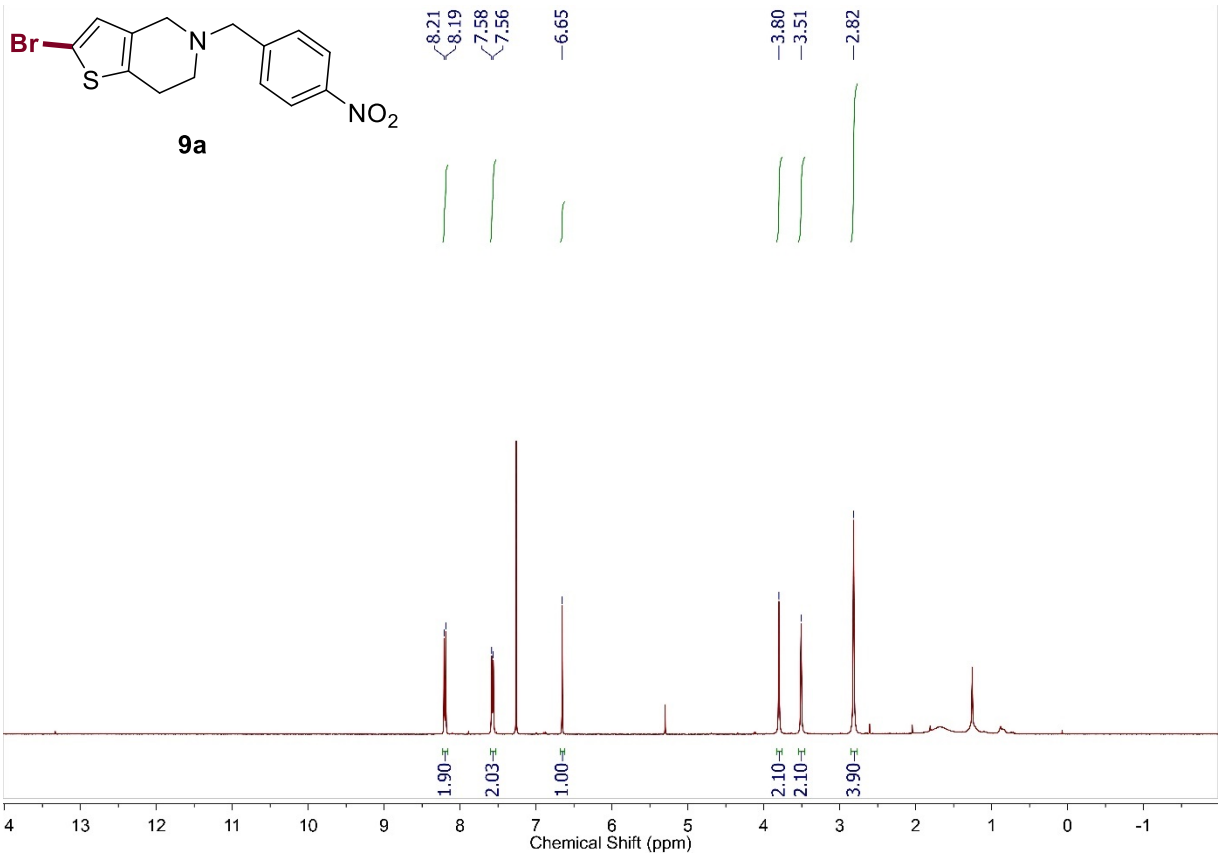

6a-NOESY
Br
5
NH2
F
7
N
H
Br
6a
f1 (ppm)
f2 (ppm)

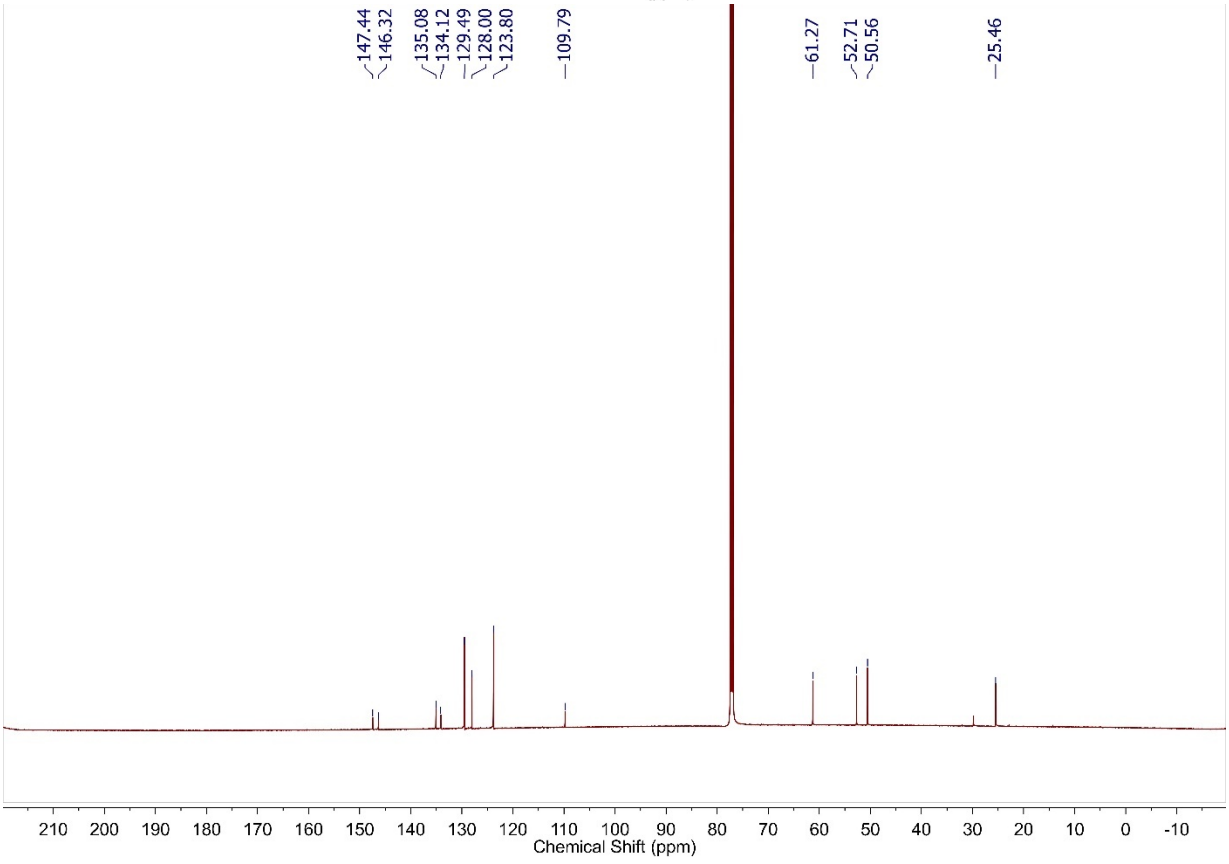

11.89
8.50
8.50
8.25
8.24
8.05
8.04
7.68
7.68
7.66
7.66
7.59
7.57
2.79
Br
N
N
H
7a
0.88
1.12
1.12
1.00
1.10
13 12 11 10 9 8 7 6 5 4 3 2 1 0 -1
Chemical Shift (ppm)

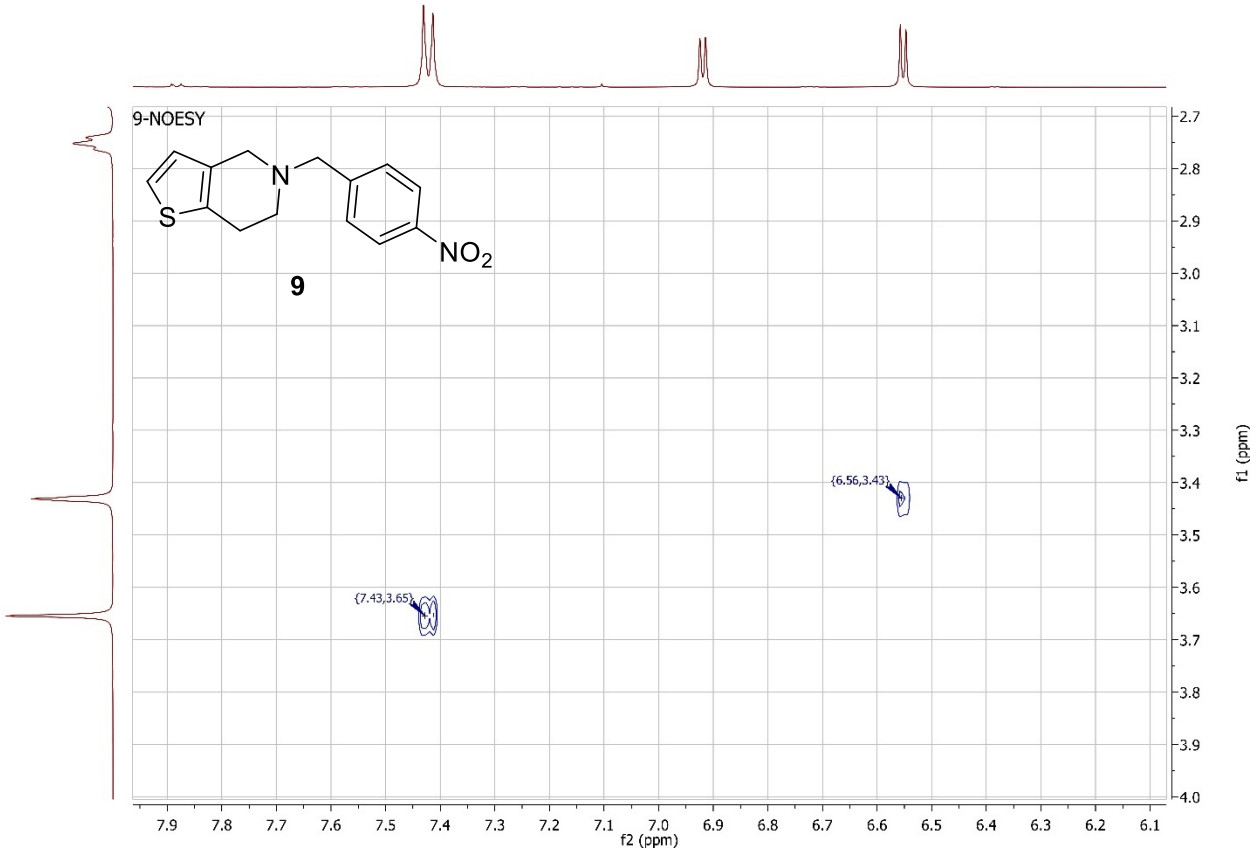

7a-NOESY
Br
N
N
H
7a
{8.50,8.05}
{8.26,8.04}
6.2 6.4 6.6 6.8 7.0 7.2 7.4 7.6 7.8 8.0 8.2 8.4 8.6 8.8 9.0
f1 (ppm)
8.9 8.8 8.7 8.6 8.5 8.4 8.3 8.2 8.1 8.0 7.9 7.8 7.7 7.6 7.5 7.4 7.3 7.2 7.1
f2 (ppm)

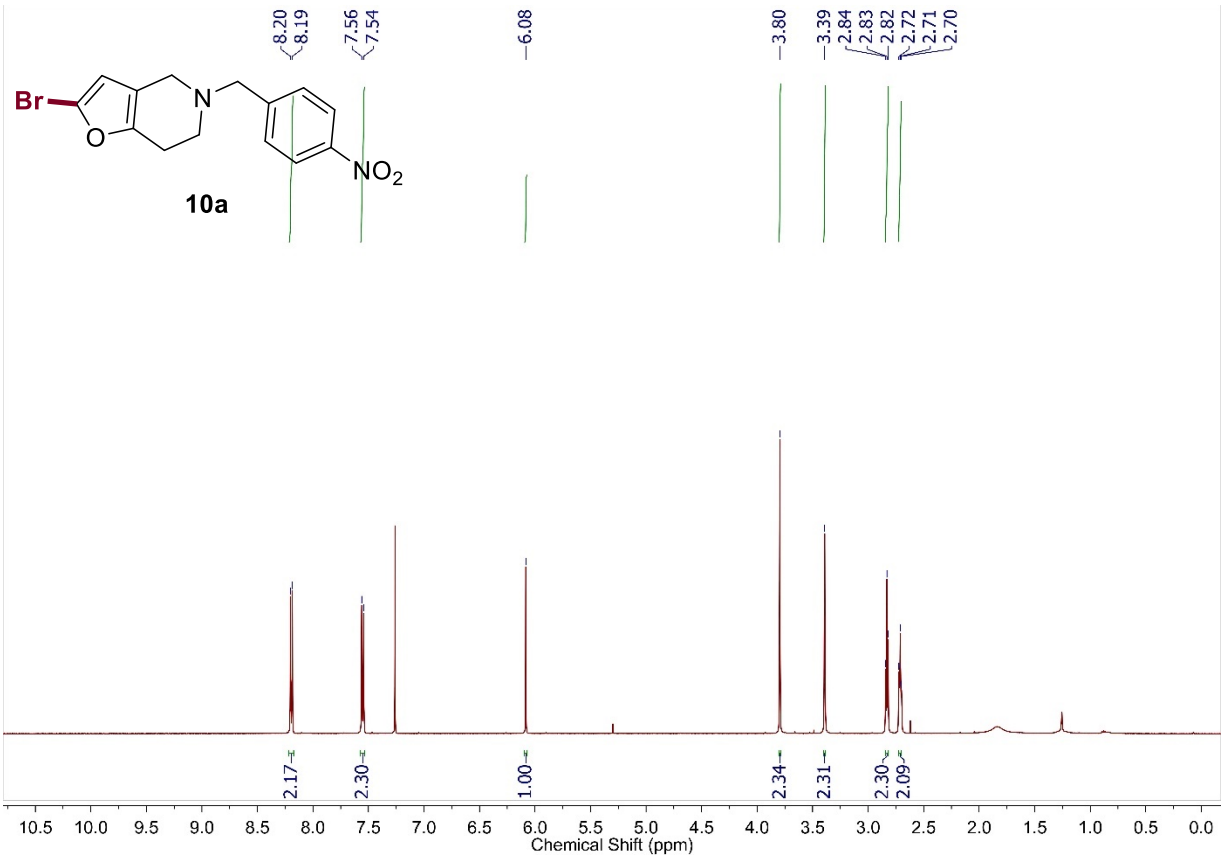

O
OH
CF3
NH2
Br
8a
8.11
8.11
7.80
7.80
1.00
0.92
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
Chemical Shift (ppm)

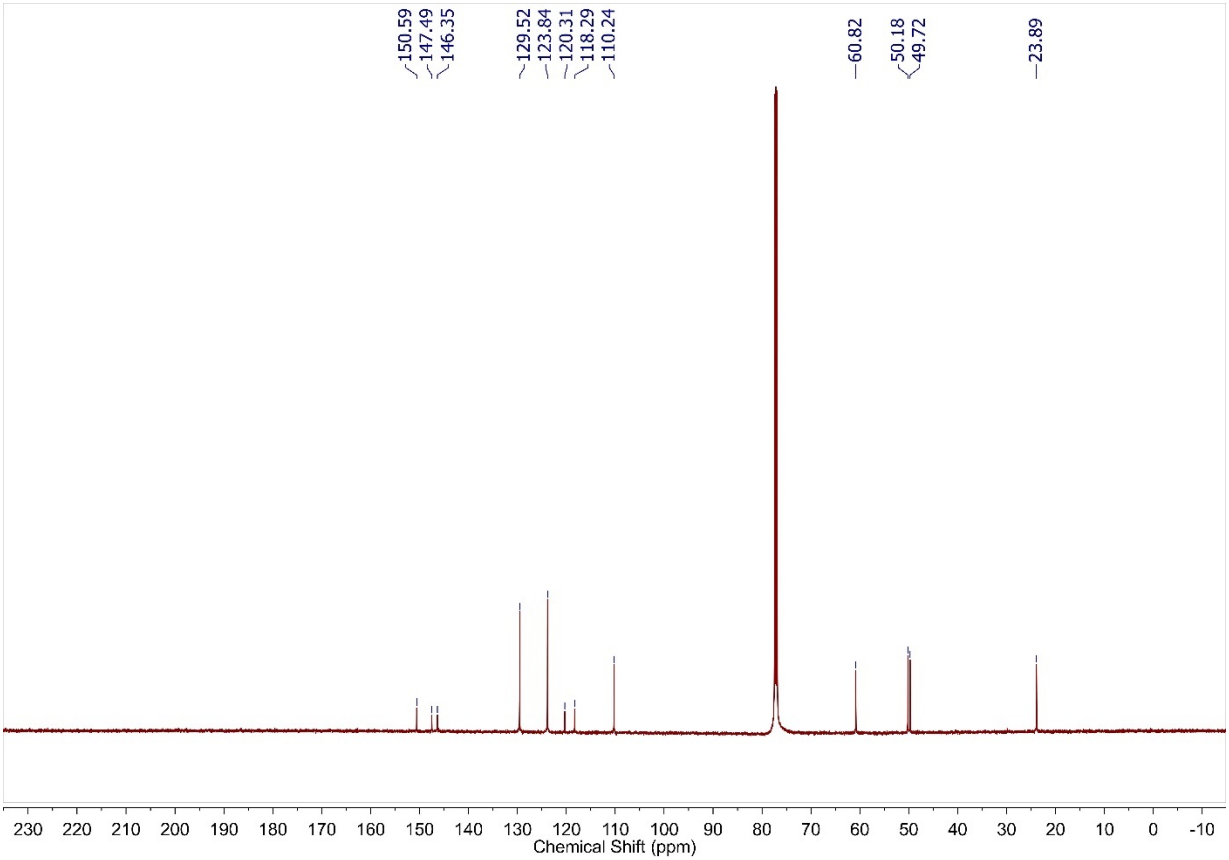

169.75
151.84
134.29
129.65
128.36
126.21
124.06
121.92
118.75
118.48
118.21
117.94
112.36
110.73
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
Chemical Shift (ppm)

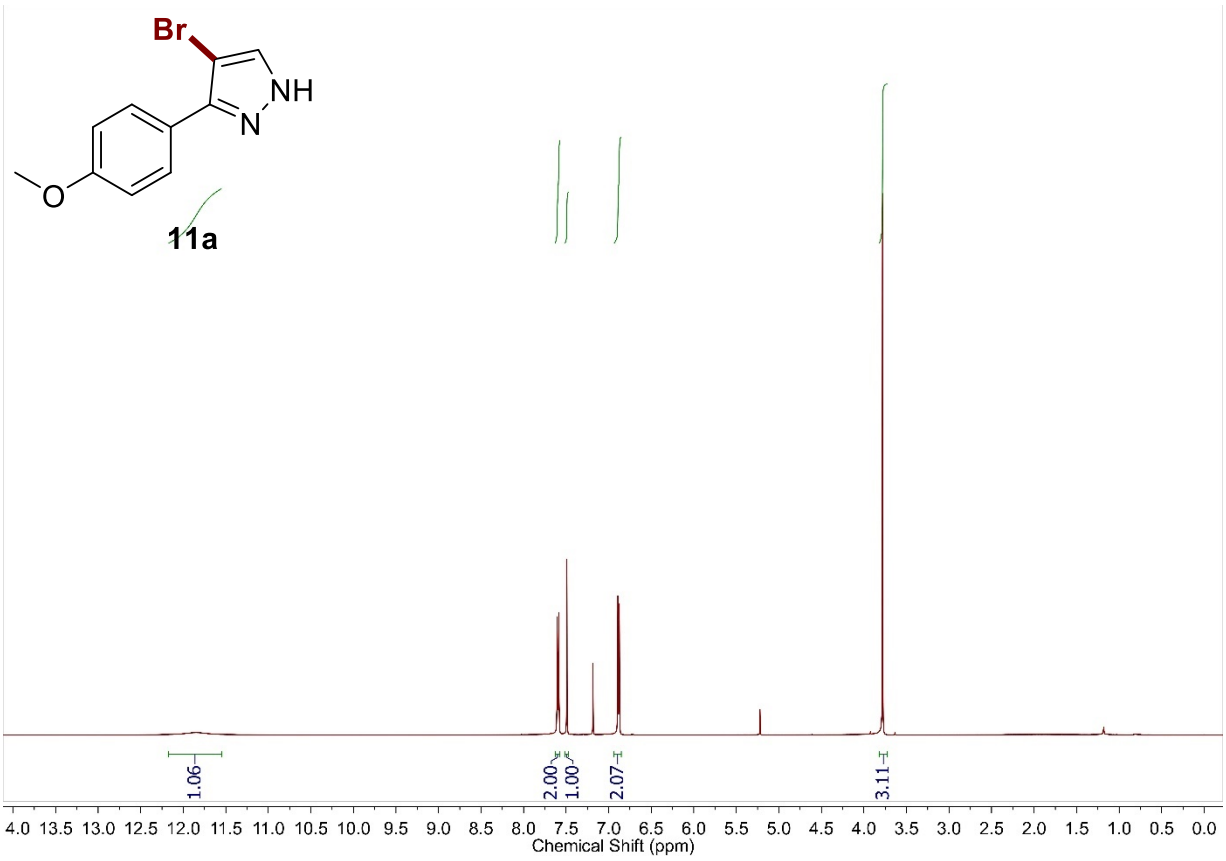

Br
S
N
NO2
9a
8.21
8.19
7.58
7.56
6.65
3.80
3.51
2.82
1.90
2.03
1.00
2.10
2.10
3.90
Chemical Shift (ppm)

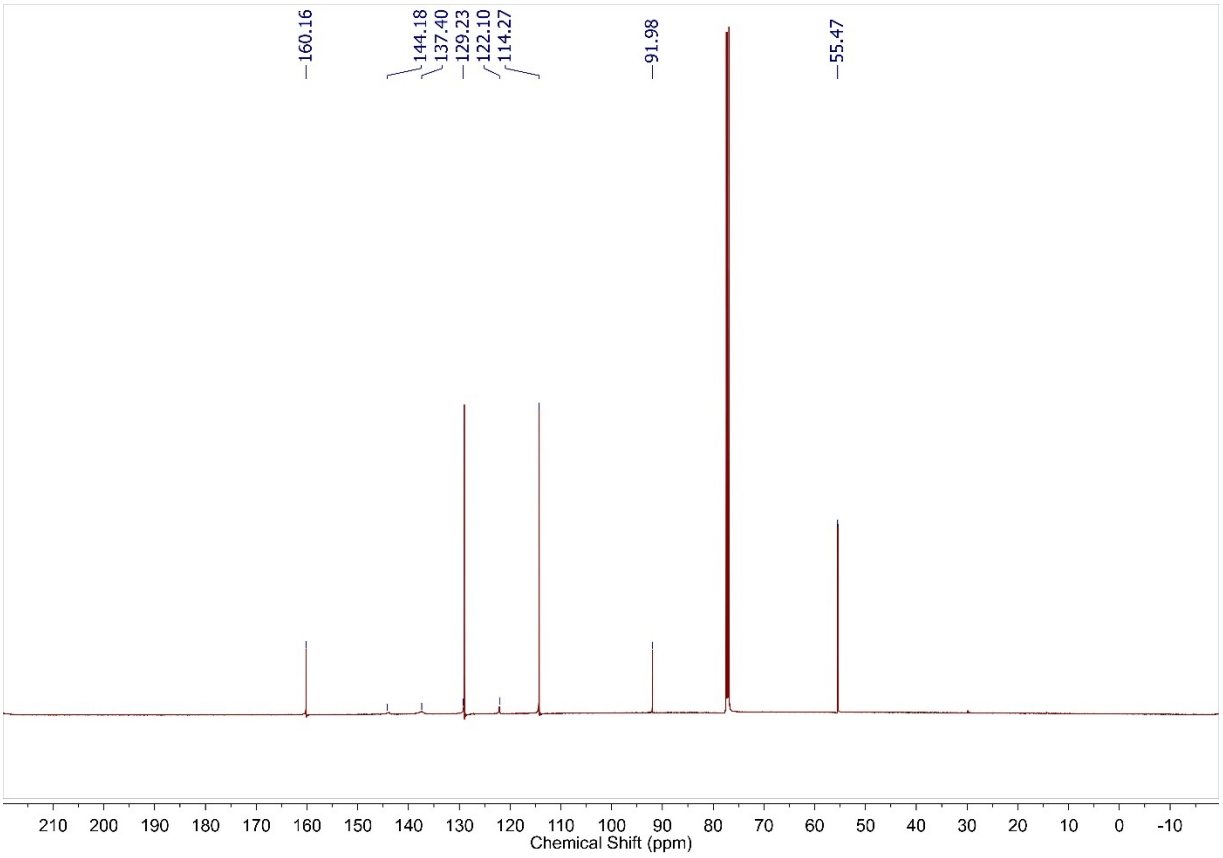

147.44
146.32
135.08
134.12
129.49
128.00
123.80
109.79
61.27
52.71
50.56
25.46
Chemical Shift (ppm)

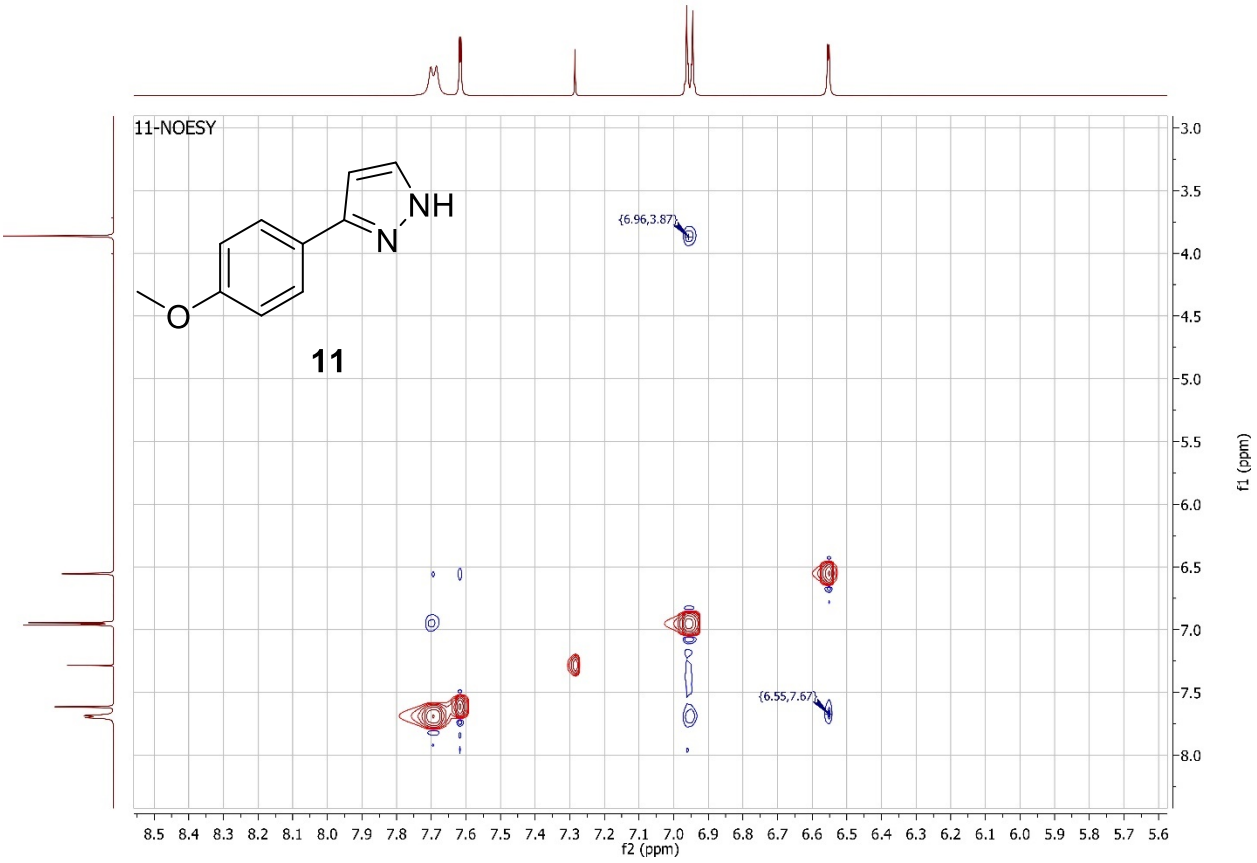

9-NOESY
N
S
NO2
9
{6.56,3.43}
{7.43,3.65}
f1 (ppm)
f2 (ppm)

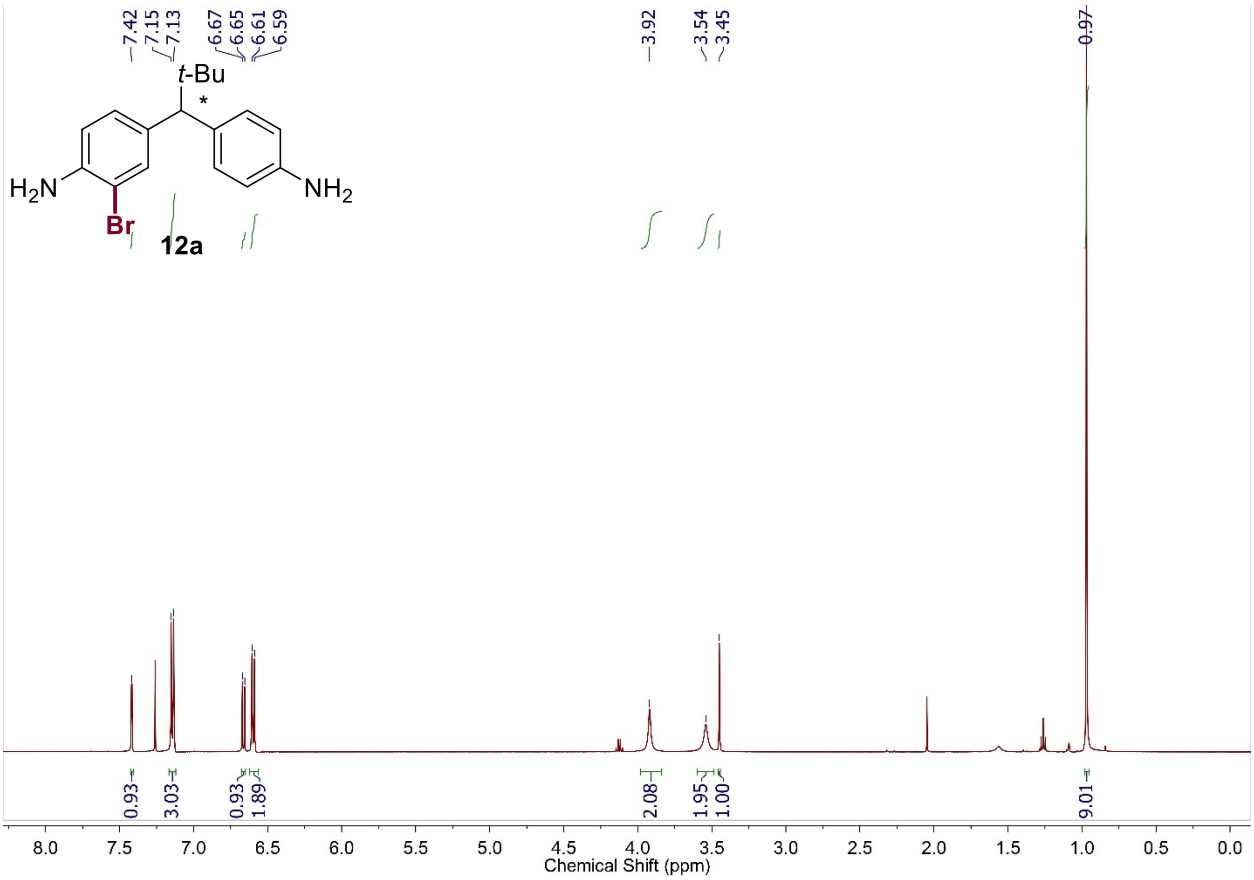

8.20
8.19
7.56
7.54
6.08
3.80
3.39
2.84
2.83
2.82
2.72
2.71
2.70
Br
O
N
NO2
10a
2.17
2.30
1.00
2.34
2.31
2.30
2.09
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
Chemical Shift (ppm)

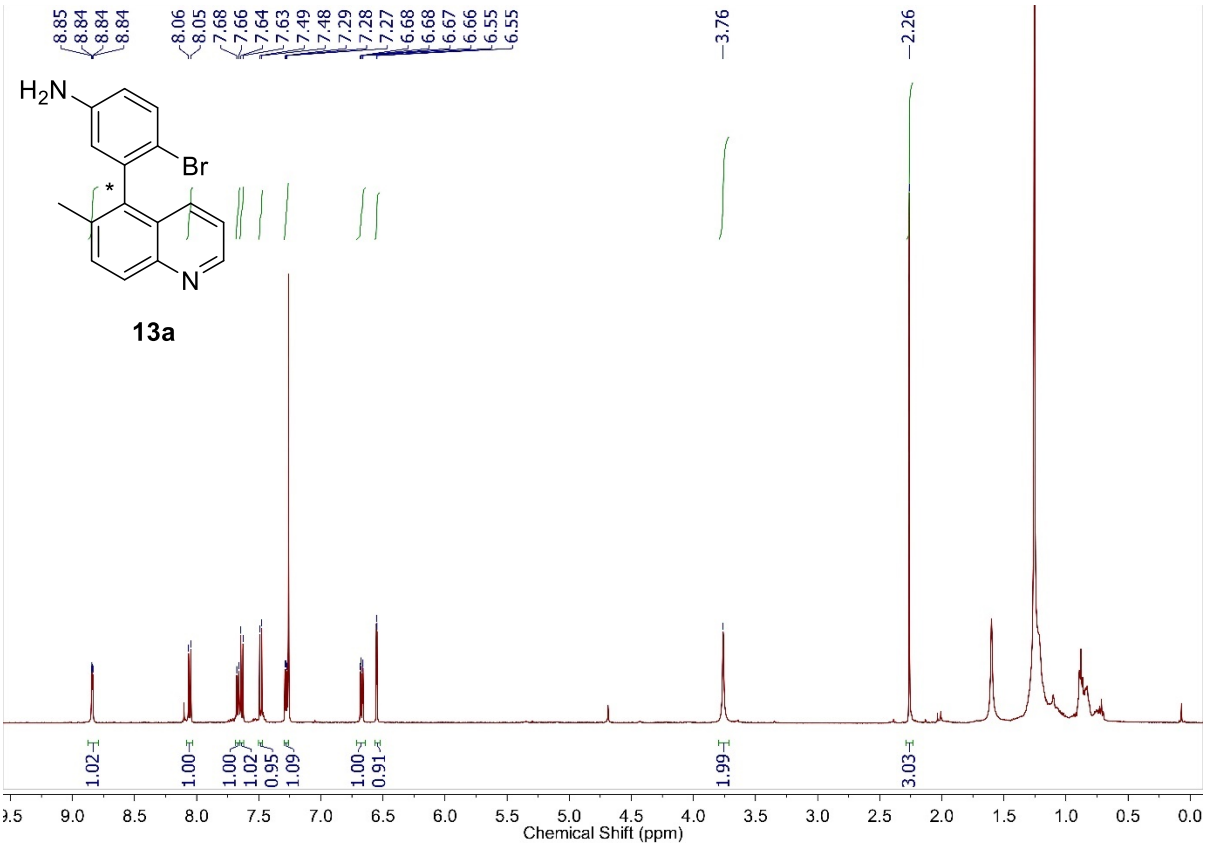

150.59
147.49
146.35
129.52
123.84
120.31
118.29
110.24
60.82
50.18
49.72
23.89
230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
Chemical Shift (ppm)

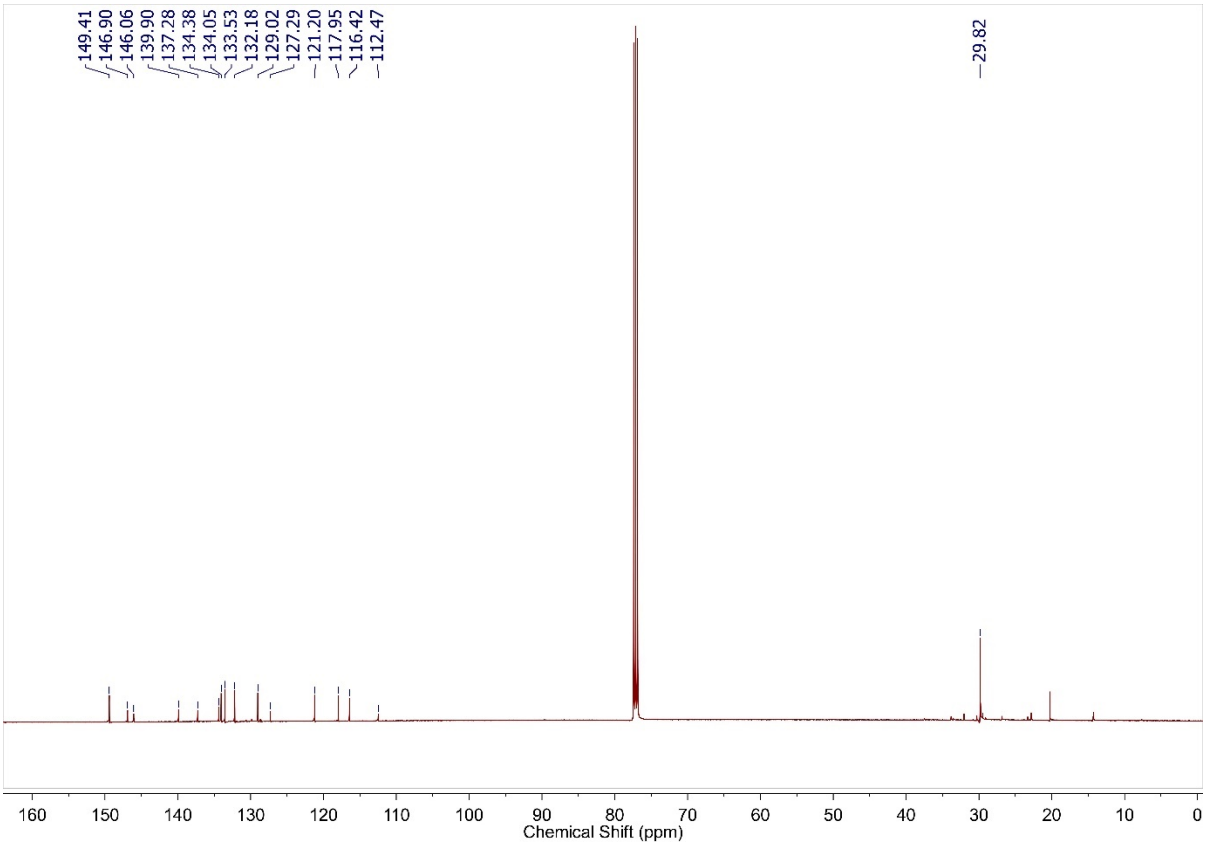

Br
NH
N
O
11a
1.06
2.00
1.00
2.07
3.11
Chemical Shift (ppm)

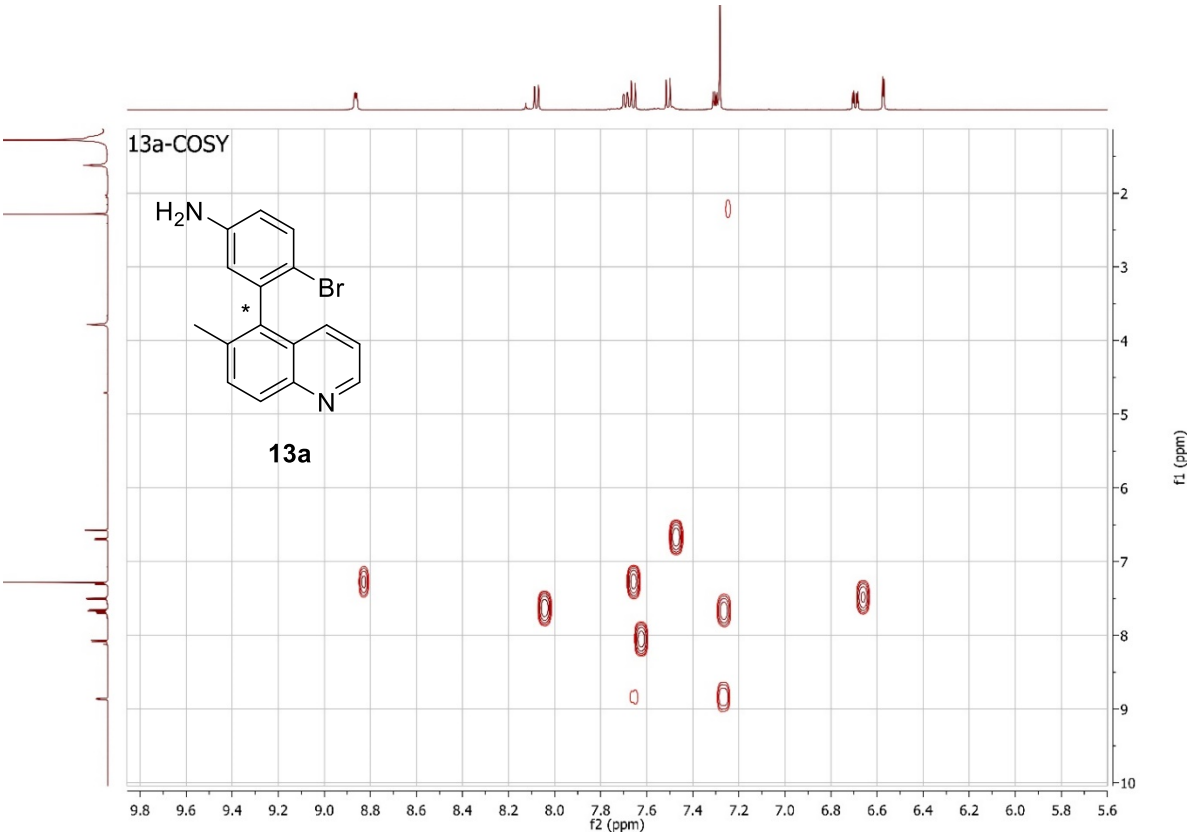

160.16
144.18
137.40
129.23
122.10
114.27
91.98
55.47
Chemical Shift (ppm)

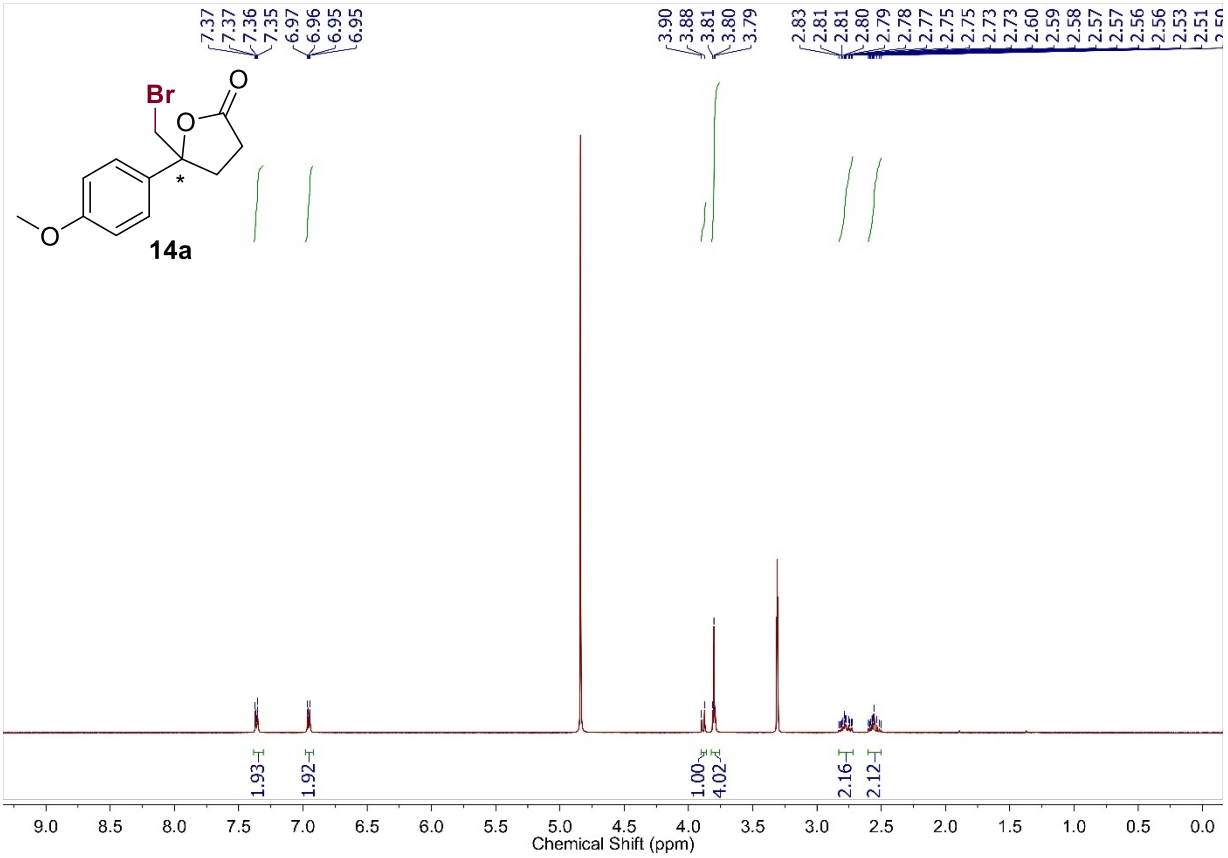

11-NOESY
NH
N
O
11
{6.96,3.87}
{6.55,7.67}
f1 (ppm)
f2 (ppm)

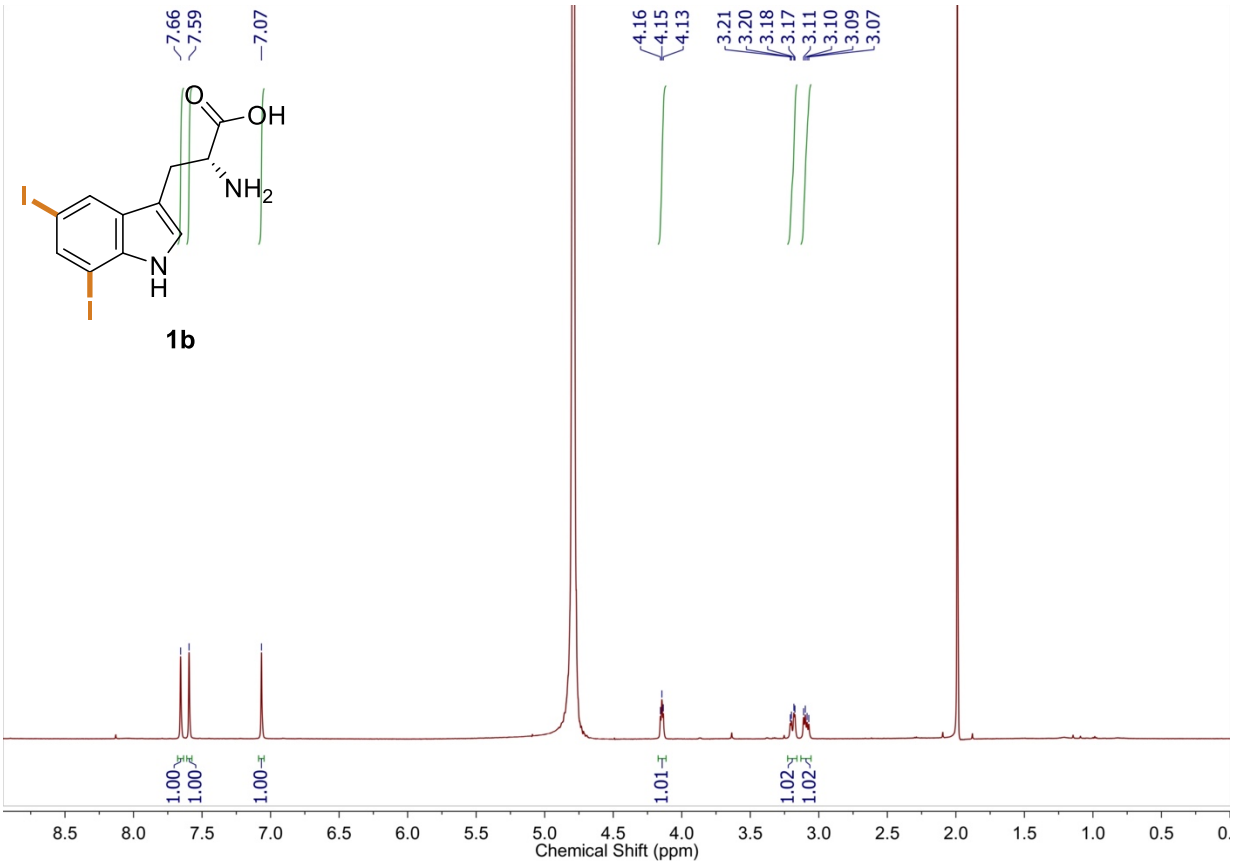

7.42
7.15
7.13
6.67
6.65
6.61
6.59
3.92
3.54
3.45
0.97
t-Bu
*
H2N
NH2
Br
12a
0.93
3.03
0.93
1.89
2.08
1.95
1.00
9.01
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Chemical Shift (ppm)

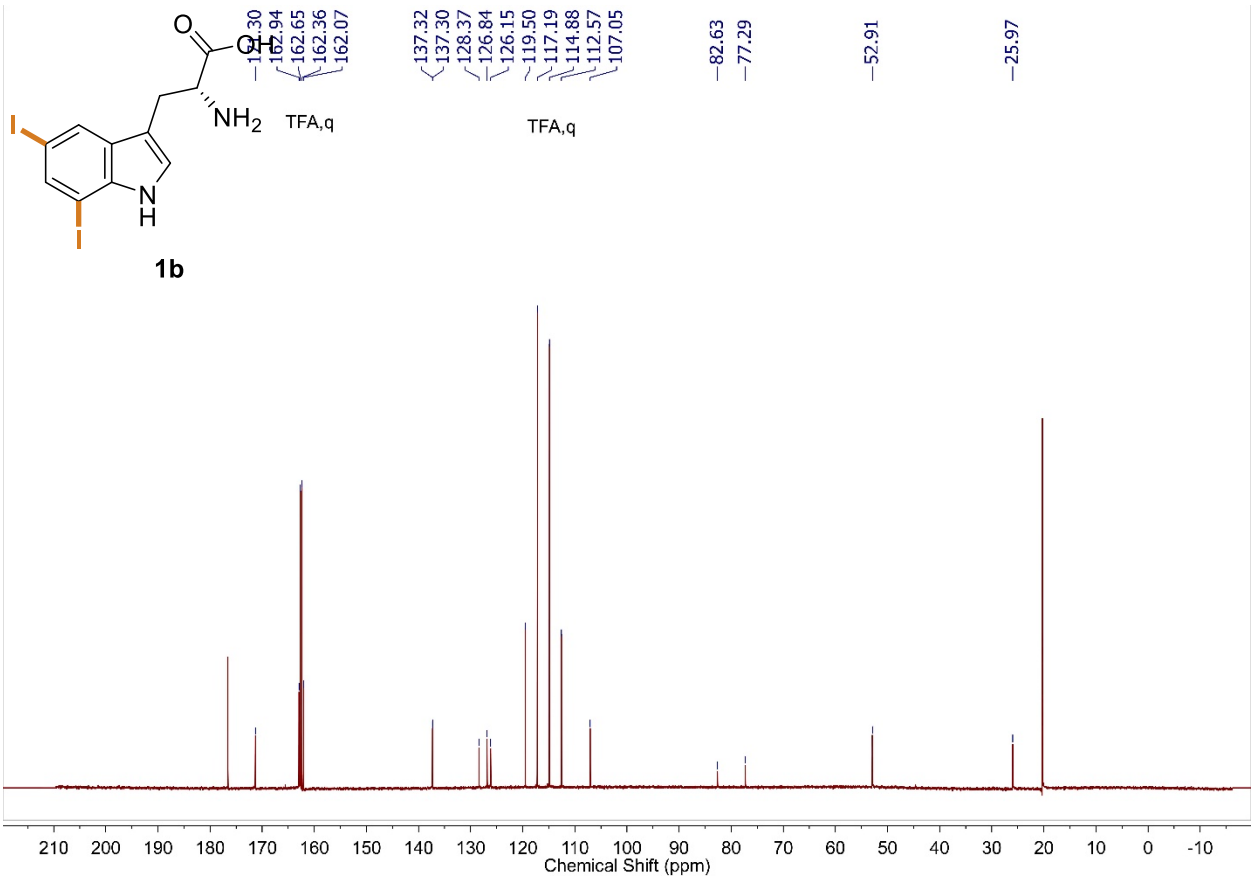

8.85
8.84
8.84
8.84
8.06
8.05
7.68
7.66
7.64
7.63
7.49
7.48
7.29
7.28
7.27
6.68
6.68
6.67
6.66
6.55
6.55
3.76
2.26
H_2N
Br
*
N
13a
1.02
1.00
1.00
1.02
0.95
1.09
1.00
0.91
1.99
3.03
Chemical Shift (ppm)

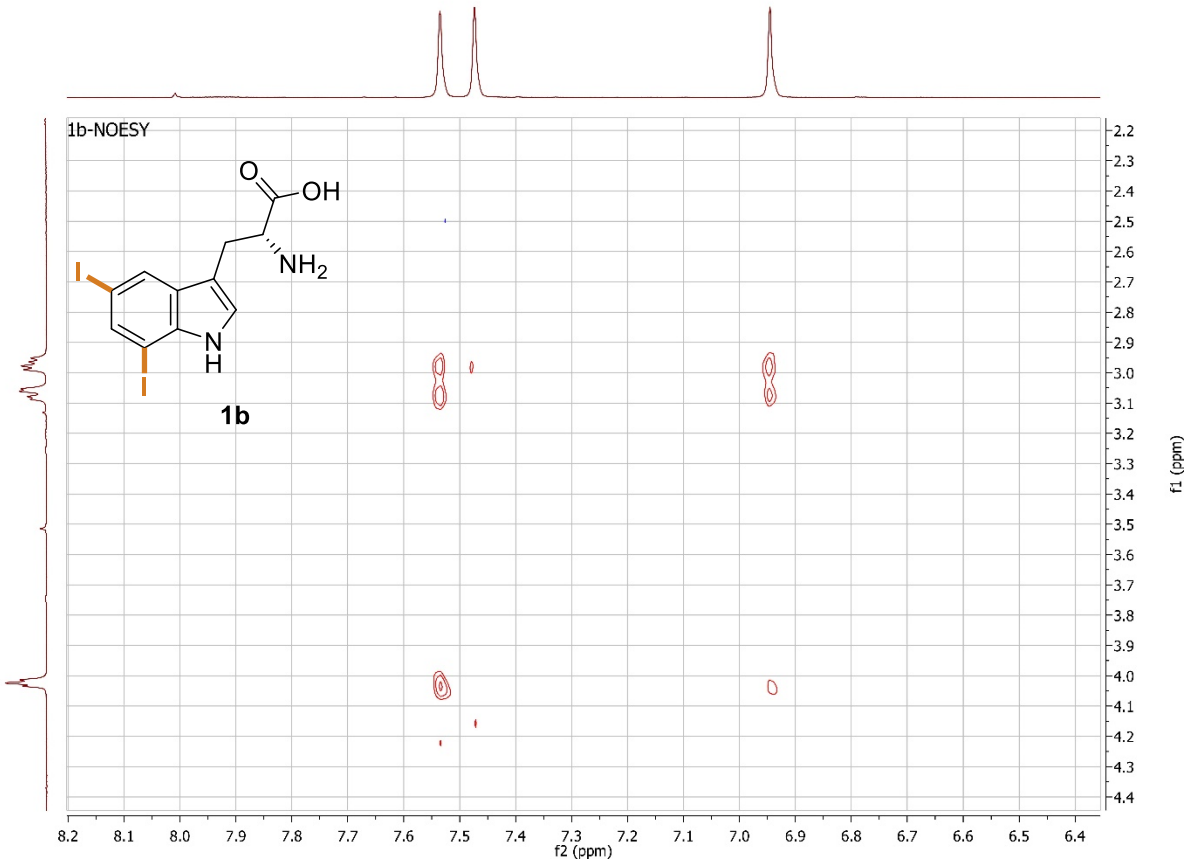

149.41
146.90
146.06
139.90
137.28
134.38
134.05
133.53
132.18
129.02
127.29
121.20
117.95
116.42
112.47
29.82
Chemical Shift (ppm)

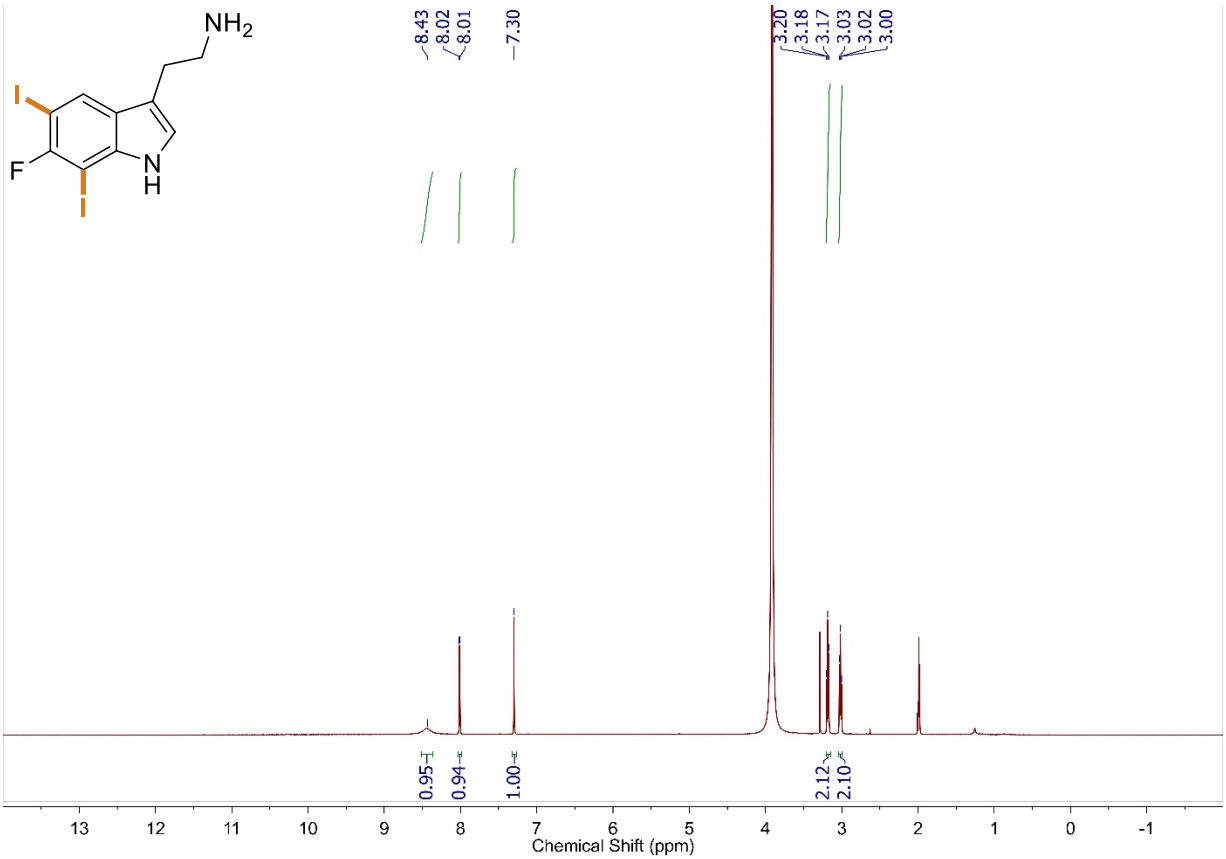

13a-COSY
H_2N
Br
*
N
13a
f1 (ppm)
f2 (ppm)

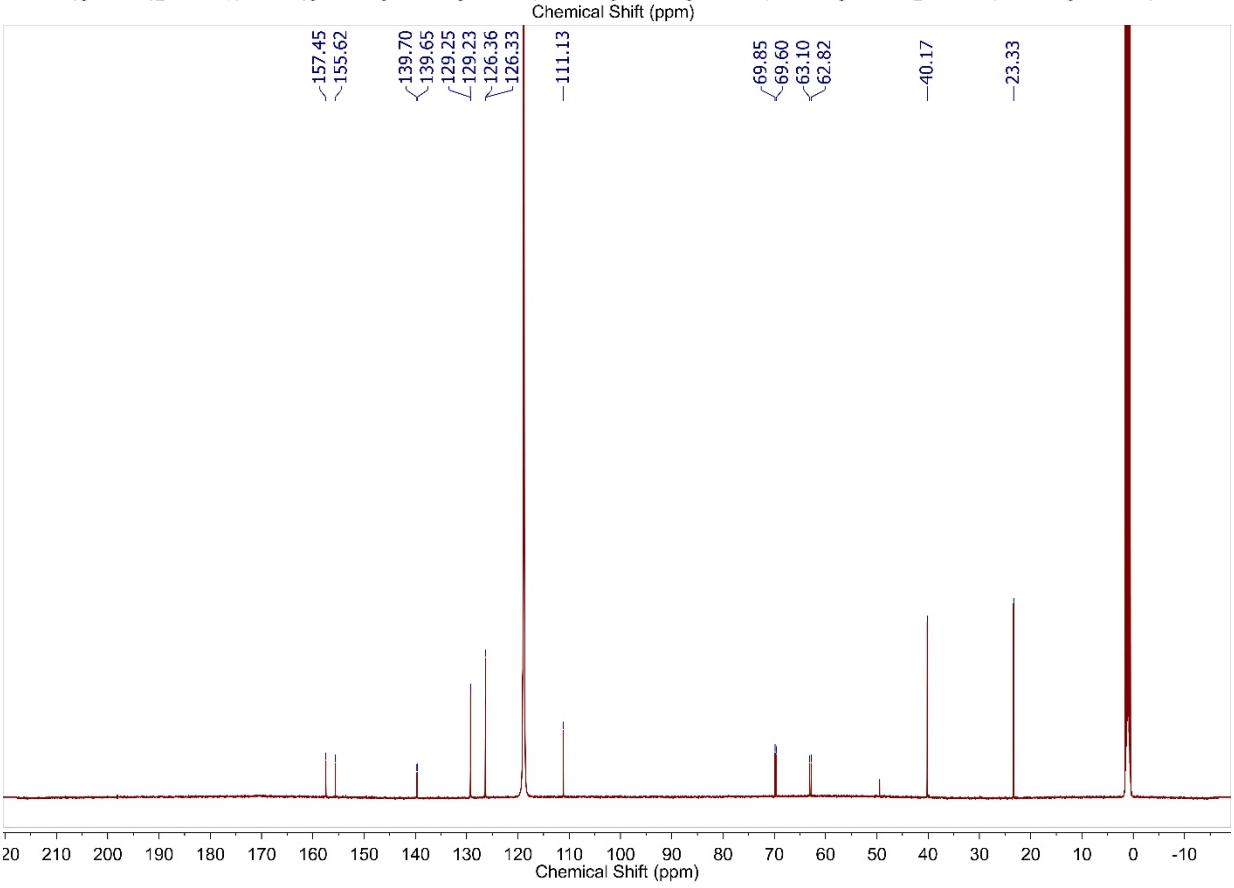

Br
O
O
O
*
O
14a
Chemical Shift (ppm)

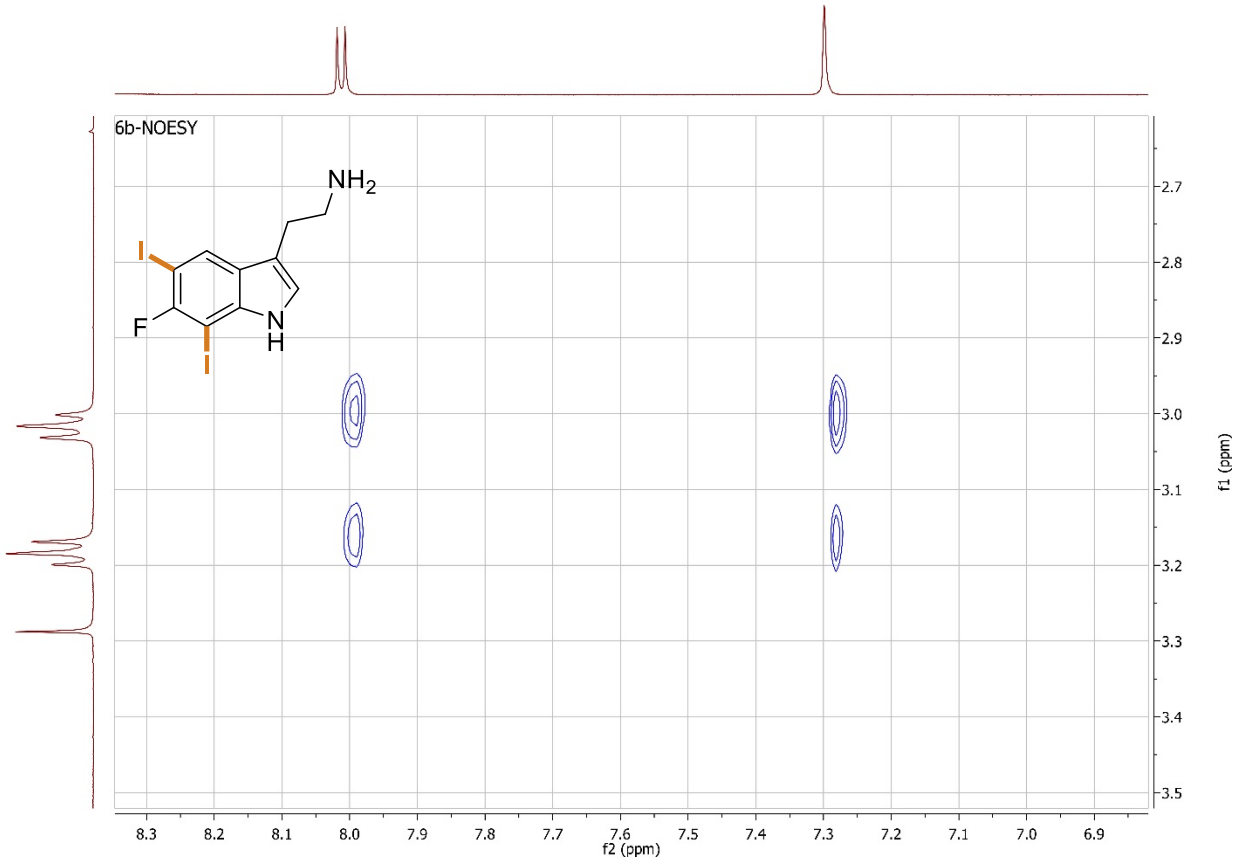

7.66
7.59
7.07
4.16
4.15
4.13
3.21
3.20
3.18
3.17
3.11
3.10
3.09
3.07
O
OH
NH2
N
H
I
I
1b
1.00
1.00
1.00
1.01
1.02
1.02
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
Chemical Shift (ppm)

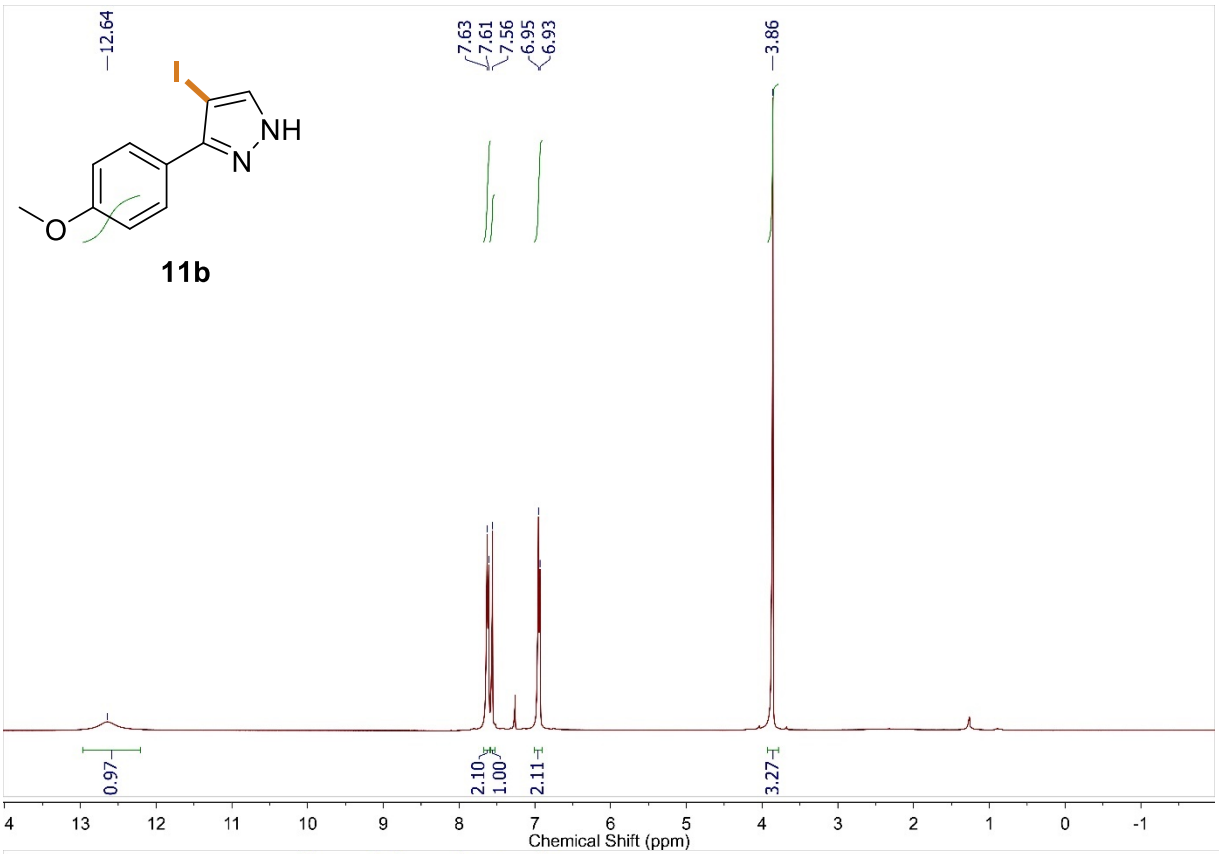

O
OH
NH2
TFA,q
N
H
1b
137.32
137.30
128.37
126.84
126.15
119.50
117.19
114.88
112.57
107.05
82.63
77.29
52.91
25.97
TFA,q
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
Chemical Shift (ppm)

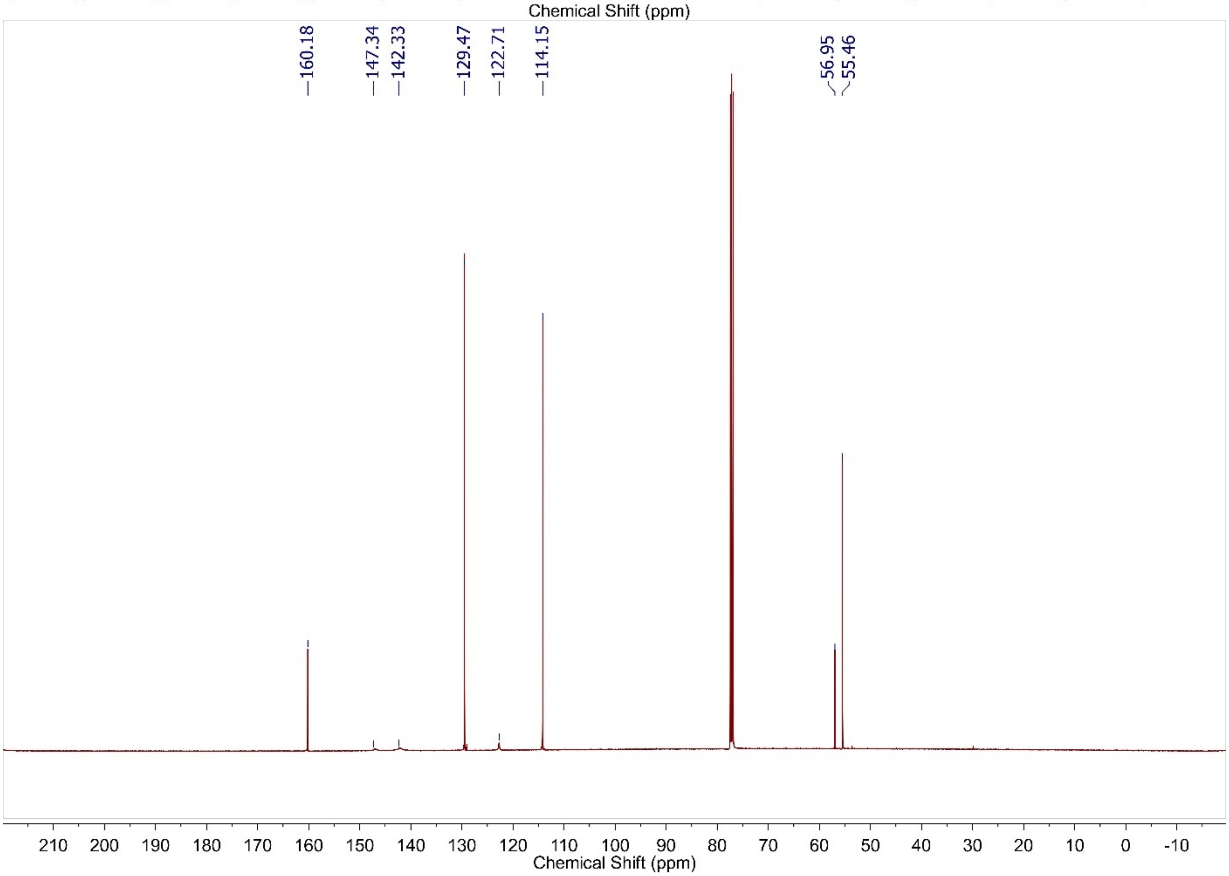

1b-NOESY
O
OH
NH2
I
I
N
H
1b
f1 (ppm)
f2 (ppm)

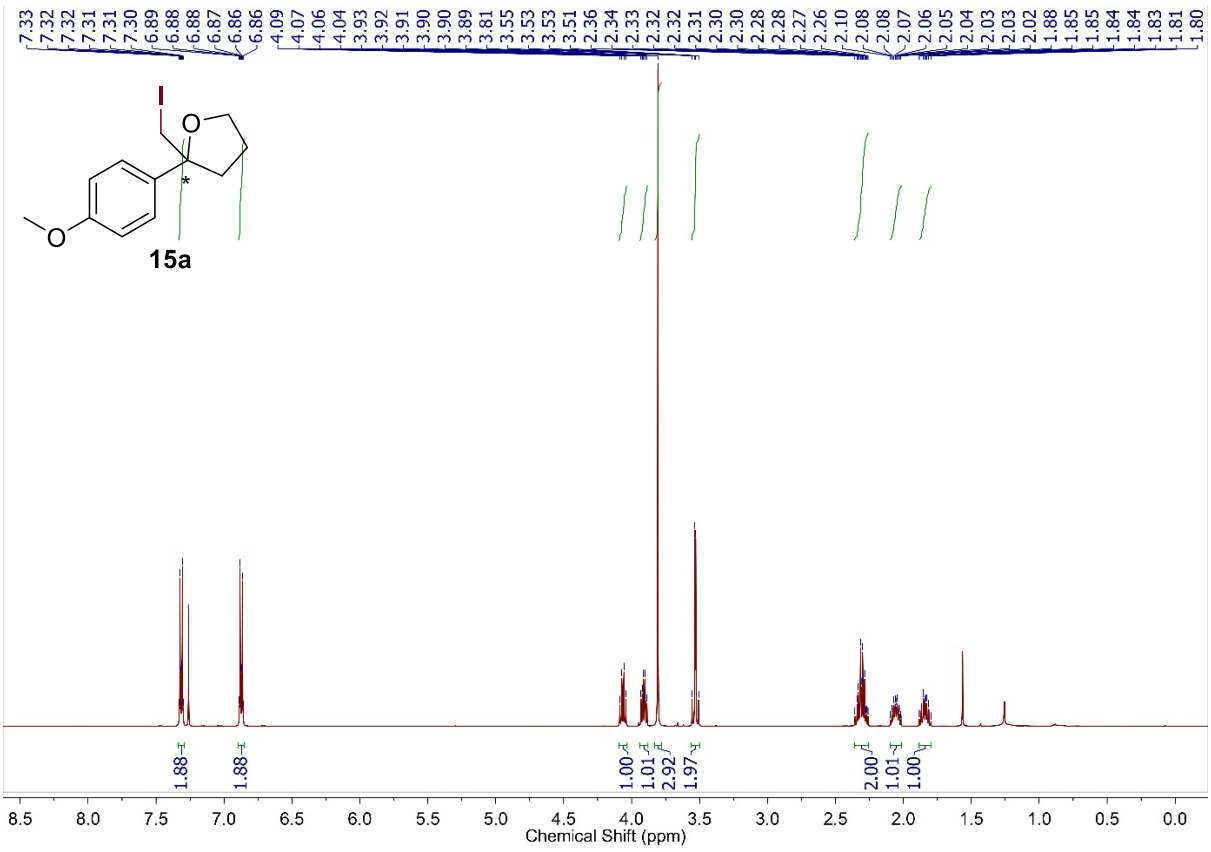

NH2
I
F
I
N
H
8.43
8.02
8.01
7.30
3.20
3.18
3.17
3.03
3.02
3.00
0.95
0.94
1.00
2.12
2.10
13 12 11 10 9 8 7 6 5 4 3 2 1 0 -1
Chemical Shift (ppm)

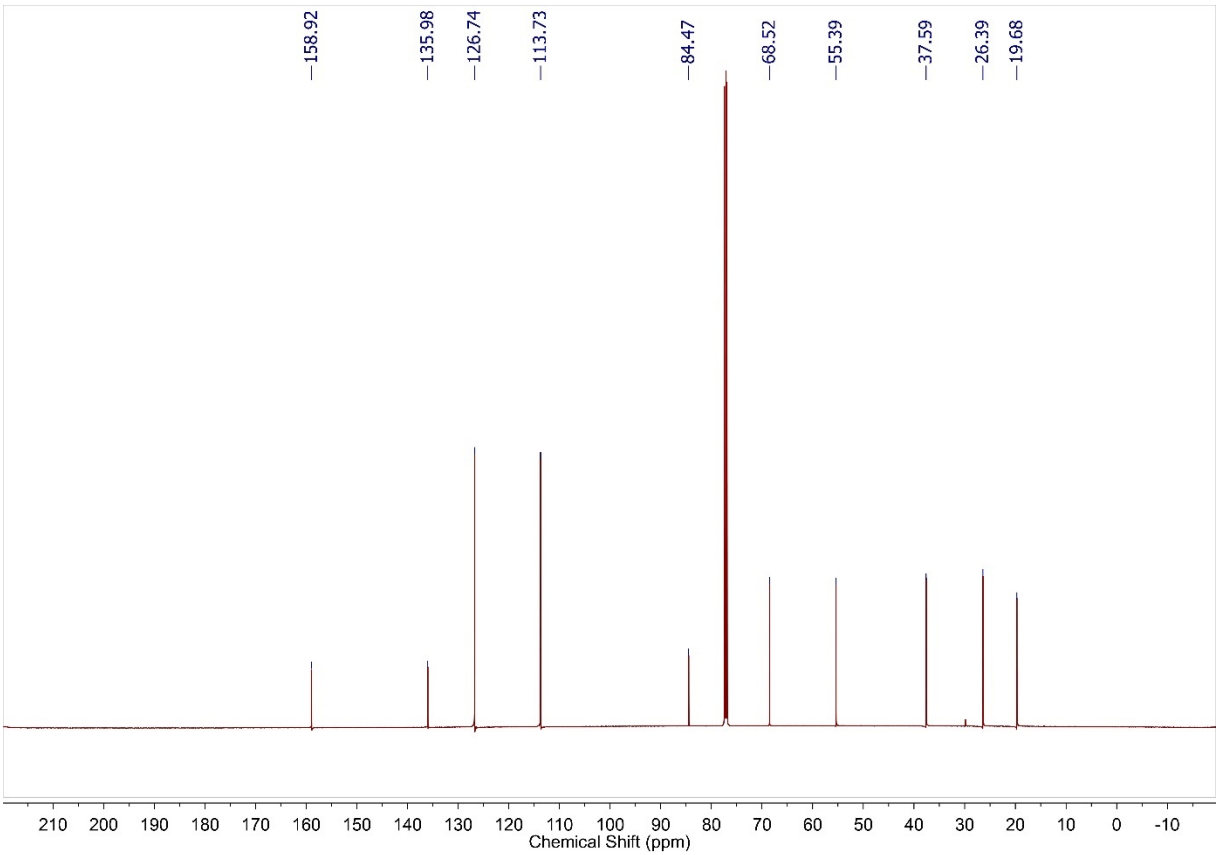

157.45
155.62
139.70
139.65
129.25
129.23
126.36
126.33
111.13
69.85
69.60
63.10
62.82
40.17
23.33
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
Chemical Shift (ppm)


6b-NOESY
NH2
I
F
I
N
H
2.7
2.8
2.9
3.0
3.1
3.2
3.3
3.4
3.5
f1 (ppm)
8.3
8.2
8.1
8.0
7.9
7.8
7.7
7.6
7.5
7.4
7.3
7.2
7.1
7.0
6.9
f2 (ppm)


12.64
7.63
7.61
7.56
6.95
6.93
3.86
NH
N
O
11b
0.97
2.10
1.00
2.11
3.27
Chemical Shift (ppm)


160.18
147.34
142.33
129.47
122.71
114.15
56.95
55.46
Chemical Shift (ppm)


7.33 7.32 7.32 7.31 7.31 7.30 6.89 6.88 6.88 6.87 6.86 6.86
4.09 4.07 4.06 4.04 3.93 3.92 3.91 3.90 3.90 3.89 3.81 3.55 3.53 3.51 2.36 2.34 2.33 2.32 2.32 2.31 2.30 2.30 2.28 2.28 2.27 2.26 2.10 2.08 2.08 2.07 2.06 2.05 2.04 2.03 2.03 2.02 1.88 1.85 1.85 1.84 1.84 1.83 1.81 1.80
I
O
*
O
15a
1.88 1.88 1.00 1.01 2.92 1.97 2.00 1.01 1.00
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
Chemical Shift (ppm)


158.92 135.98 126.74 113.73 84.47 68.52 55.39 37.59 26.39 19.68
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
Chemical Shift (ppm)

IX. References

1. Heckman, K.L., and Pease, L.R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. Nature Protocols *2*, 924–932. 10.1038/nprot.2007.132.

2. Payne, J.T., Andorfer, M.C., and Lewis, J.C. (2013). Regioselective Arene Halogenation Using the FAD-Dependent Halogenase RebH. Angewandte Chemie International Edition in English *52*, 5271–5274. 10.1002/anie.201300762.

3. Andorfer, M.C., Grob, J.E., Hajdin, C.E., Chael, J.R., Siuti, P., Lilly, J., Tan, K.L., and Lewis, J.C. (2017). Understanding Flavin-Dependent Halogenase Reactivity via Substrate Activity Profiling. ACS Catalysis *7*, 1897–1904. 10.1021/acscatal.6b02707.

4. Breinlinger, S., Phillips, T.J., Haram, B.N., Mareš, J., Yerena, J.A.M., Hrouzek, P., Sobotka, R., Henderson, W.M., Schmieder, P., Williams, S.M., et al. (2021). Hunting the eagle killer: A cyanobacterial neurotoxin causes vacuolar myelinopathy. Science *371*. 10.1126/science.aax9050.

5. Dan, W.-J., Zhang, Q., Zhang, F., Wang, W.-W., and Gao, J.-M. (2019). Benzonate derivatives of acetophenone as potent α-glucosidase inhibitors: synthesis, structure–activity relationship and mechanism. J Enzym Inhib Med Ch *34*, 937–945. 10.1080/14756366.2019.1604519.

6. Scattolin, T., Gharbaoui, T., and Chen, C. (2022). A Nucleophilic Deprotection of Carbamate Mediated by 2-Mercaptoethanol. Org Lett *24*, 3736–3740. 10.1021/acs.orglett.2c01410.

7. Joshi, G., and Adimurthy, S. (2011). Environment-Friendly Bromination of Aromatic Heterocycles Using a Bromide–Bromate Couple in an Aqueous Medium. Ind Eng Chem Res *50*, 12271–12275. 10.1021/ie2004863.

8. Ponce, M.A., and Erra-Balsells, R. (2001). Synthesis and isolation of bromo-β-carbolines obtained by bromination of β-carboline alkaloids. J Heterocyclic Chem *38*, 1087–1095. 10.1002/jhet.5570380512.

9. Kim, B., Chinn, A.J., Fandrick, D.R., Senanayake, C.H., Singer, R.A., and Miller, S.J. (2016). Distal Stereocontrol Using Guanidinylated Peptides as Multifunctional Ligands: Desymmetrization of Diarylmethanes via Ullman Cross-Coupling. J Am Chem Soc *138*, 7939–7945. 10.1021/jacs.6b03444.

10. Mondal, D., Fisher, B.F., Jiang, Y., and Lewis, J.C. (2021). Flavin-dependent halogenases catalyze enantioselective olefin halocyclization. Nat Commun *12*, 3268. 10.1038/s41467-021-23503-3.

11. Ashtekar, K.D., Marzijarani, N.S., Jaganathan, A., Holmes, D., Jackson, J.E., and Borhan, B. (2014). A New Tool To Guide Halofunctionalization Reactions: The Halenium Affinity (HalA) Scale. Journal of the American Chemical Society *136*, 13355–13362. 10.1021/ja506889c.

12. Pracht, P., Bohle, F., and Grimme, S. (2020). Automated exploration of the low-energy chemical space with fast quantum chemical methods. Phys Chem Chem Phys *22*, 7169–7192. 10.1039/c9cp06869d.

13. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. 10.1038/s41586-021-03819-2.

14. Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Research *38*, W545–W549. 10.1093/nar/gkq366.

15. Du, Z., Su, H., Wang, W., Ye, L., Wei, H., Peng, Z., Anishchenko, I., Baker, D., and Yang, J. (2021). The trRosetta server for fast and accurate protein structure prediction. Nat Protoc *16*, 5634–5651. 10.1038/s41596-021-00628-9.

16. The PyMOL Molecular Graphics System, Version 2.3., Schrodinger, LLC. (2010).

17. Nicoll, C.R., Bailleul, G., Fiorentini, F., Mascotti, M.L., Fraaije, M.W., and Mattevi, A. (2020). Ancestral-sequence reconstruction unveils the structural basis of function in mammalian FMOs. Nat Struct Mol Biol *27*, 14–24. 10.1038/s41594-019-0347-2.

18. Nivón, L.G., Moretti, R., and Baker, D. (2013). A Pareto-Optimal Refinement Method for Protein Design Scaffolds. Plos One *8*, e59004. 10.1371/journal.pone.0059004.

19. Davis, I.W., and Baker, D. (2009). RosettaLigand Docking with Full Ligand and Receptor Flexibility. J Mol Biol *385*, 381–392. 10.1016/j.jmb.2008.11.010.

20. Combs, S.A., DeLuca, S.L., DeLuca, S.H., Lemmon, G.H., Nannemann, D.P., Nguyen, E.D., Willis, J.R., Sheehan, J.H., and Meiler, J. (2013). Small-molecule ligand docking into comparative models with Rosetta. Nat Protoc *8*, 1277–1298. 10.1038/nprot.2013.074.

21. Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E., and Hutchison, G.R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminformatics *4*, 17. 10.1186/1758-2946-4-17.

22. Yeh, E., Blasiak, L.C., Koglin, A., Drennan, C.L., and Walsh, C.T. (2007). Chlorination by a Long-Lived Intermediate in the Mechanism of Flavin-Dependent Halogenases. Biochemistry *46*, 1284–1292. 10.1021/bi0621213.