Supporting Information

Engineering Chemoselectivity in Hemoprotein-Catalyzed Indole Amidation

Oliver F. Brandenberg,[†,‡] *David C. Miller,*[†] *Ulrich Markel,*[†,§] *Anissa Ouald Chaib,*[†,⊥]

Frances H. Arnold [*,†]

[†] Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

[‡] Present address: Department of Human Genetics, University of California Los Angeles, Los Angeles, California 90095, USA

[§] Present address: Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany

[⊥] Present address: Laboratory of Organic Chemistry, ETH Zürich, 8093 Zürich, Switzerland

[*] Corresponding author: frances@cheme.caltech.edu

SI content:

I. Experimental Procedures

General Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar) and used without further purification. Silica gel chromatography was carried out using AMD Silica Gel 60, 230-400 mesh. Synthetic reactions were monitored using thin layer chromatography (Merck 60 gel plates) using a UV-lamp for visualization. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Inova 300 MHz or 500 MHz, or Bruker Prodigy 400 MHz instrument, in $CDCl_3$ and are internally referenced to the residual solvent peak. Data for ^{1}H NMR are reported as follows: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet, dd = doublet of doublets, dt = doublet of triplets, ddd = doublet of doublet of doublets), coupling constant (Hz), integration. High-resolution mass spectra were obtained at the California Institute of Technology Mass Spectrometry Facility.

Chromatography. Analytical reversed-phase high-performance liquid chromatography (HPLC) was carried out using an Agilent 1200 series instrument with water and acetonitrile as mobile phases using an Agilent XDB-C18 column (4.6 x 150 mm, 5 µm). For quantitative HPLC analyses of reaction products, calibration curves using either 1,3,5-trimethoxybenzene or 1-phenyl-3-propanol as internal standards were generated (Figures S11 to S20). Preparative-scale HPLC to purify enzymatic reaction products was performed using an Agilent XDB-C18 column (9.4 x 250 mm, 5 µm). The identity of enzymatic reaction products was confirmed by NMR analysis of products isolated from enzymatic reactions performed on preparative scale (*vide infra*).

Cloning and site-saturation mutagenesis. Plasmid pET22b(+) (Novagen) with the pelB leader sequence removed was used as a cloning and expression vector for all constructs described in this study (see p. S39 for DNA sequences). Site-saturation libraries were generated using the 22c-trick method.[1] Primers were obtained from IDT (primer sequences are available upon request). PCR was performed using Phusion polymerase (NEB) and the resulting PCR products were digested with DpnI (NEB), gel purified (Zymo Research), repaired using the method of Gibson,[2] and used

to directly transform *E. coli* strain BL21(DE3) (Lucigen) by electroporation. Following electroporation, cells were recovered for 45 min at 37º C in Luria-Bertani (LB) medium, aliquots were plated on LB agar plates supplemented with 100 µg/mL ampicillin (LB-Amp plates), and plates were incubated at 37º C overnight.

Site-saturation mutagenesis library screening in 96-well plate format. Single colonies of *E. coli* BL21(DE3) cells transformed with cytochrome P411 site-saturation mutagenesis libraries (or cells transformed with plasmid encoding the corresponding P411 parent variant) were picked with sterile toothpicks and cultured in 96 deep-well plates in LB medium supplemented with 100 µg/mL ampicillin (LB-Amp, 300 µL/well) at 37 ºC, 250 rpm, overnight. In a fresh 96 deep-well plate, Hyperbroth medium (AthenaES, 950 µL/well, supplemented with 100 µg/mL ampicillin (HB-Amp)) was inoculated with the pre-cultures (50 µL/well) and incubated at 37 ºC, 250 rpm, for 2.5 h. The plates were cooled on ice for 20 minutes and then induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 1.0 mM 5-aminolevulinic acid (ALA) (final concentrations). Expression was conducted at 24 ºC, 200 rpm for 16-20 h. The cells were pelleted (3,000 g, 3 min), the supernatant was discarded, and the 96-well plates were transferred to an anaerobic chamber. Cell pellets were resuspended in reaction buffer (M9-N with 20 mM glucose, 370 µL/well; M9-N minimal medium is 47.7 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.6 mM NaCl, 2.0 mM $MgSO_4$, and 0.1 mM $CaCl_2$, adjusted to pH 7.0 at room temperature), and stocks of indole (15 µL/well, in DMSO) and tosyl azide (15 µL/well, in DMSO) were added. Substrate concentration typically varied, depending on the reaction tested, between 10 and 15 mM for both indoles and tosyl azide (see Table S2). Following substrate addition, the plates were sealed with aluminum foil and shaken at 500 rpm, room temperature, overnight. After overnight incubation, the seal was removed and reactions were worked up for HPLC analysis: acetonitrile (typically 400 to 600 µL/well) supplemented with internal standard was added, the plates were resealed, briefly mixed by vortexing, and incubated for 30 min to 1 h at room temperature. The plates were then centrifuged (5,000 g, 10 min), and the cleared supernatant was filtered through an AcroPrep 96-well filter plate (0.2 µm cutoff) into a shallow-well plate for HPLC analysis.

Small-scale analytical amidation reactions using whole cells. Single colonies of *E. coli* BL21(DE3) cells transformed with plasmid encoding P411 variants were grown overnight in 3 mL LB-Amp medium at 37 ºC and 250 rpm. 2 mL of the pre-cultures were used to inoculate 48 mL of HB-Amp medium in 125 mL Erlenmeyer flasks. Cultures were incubated at 37 ºC, 230 rpm for 2 h, typically reaching an OD_{600} = 1.5. Cultures were then cooled on ice (20 min) and induced with 0.5 mM IPTG and 1.0 mM ALA (final concentrations). Expression was conducted at 24 ºC, 130 rpm, for 16-18 h. Cultures were then cooled on ice, centrifuged (3,000 g, 3 min, 4 ºC), and the pellets were resuspended to OD_{600} = 30 in M9-N minimal medium. To determine P411 expression levels, aliquots of the cell suspension were subjected to the carbon monoxide binding spectroscopy assay specified below.

Small-scale biocatalytic reactions were set up in 2 mL crimp vials with 400 µL reaction volume. Typically, 380 µL of P411-expressing cells at OD_{600} = 30 in M9-N buffer were added to the vials and transferred into an anaerobic chamber (no sparging of cell suspensions with argon was conducted; transfer of the cell solution to the anaerobic chamber, including several cycles of applying vacuum and purging with N_2, proved sufficient for degassing). Inside the anaerobic chamber, the vials were put on a shaker (500 rpm), and 10 µL of indole stock solution (in DMSO) were added, followed by 10 µL of tosyl azide stock solution (in DMSO). Final concentrations, depending on the substrates tested, typically ranged from 10 to 25 mM indole and 10 to 25 mM tosyl azide, with 5 % co-solvent. For aerobic reactions, reactions were set up as specified above on the bench under standard atmosphere. For reactions aimed at analyzing indole-azide cycloaddition, a slightly modified protocol was employed: 360 µL cell suspension were used, and 40 µL of a pre-mixed indole and tosyl azide substrate stock (in DMSO) was added. Following addition of substrates, the vials were sealed and shaken at room temperature, 500 rpm for 6–16 h. For sample work-up for HPLC analysis the reactions were quenched by adding acetonitrile (400 µL), samples were mixed thoroughly and incubated for 30 min. Subsequently, 200 µL of this mixture were transferred to fresh 1.5 mL tubes, and 800 µL acetonitrile supplemented with internal standards (1,3,5-trimethoxybenzene or 1-phenyl-3-propanol) was added. This double dilution protocol was employed to assure sample absorption on the HPLC remained in the linear range. Sample solutions in the 1.5 mL tubes were then centrifuged at 13,000 g for 10 minutes, the cleared supernatant was transferred to clean 2 mL vials and analyzed by HPLC.

Determination of whole-cell reaction kinetics and initial rates. Kinetics of 1-methylindole amidation were determined using whole *E. coli* cells expressing P411 variants. Reactions were set up as described above *(Small-scale analytical amidation reactions using whole cells*). Following addition of substrates, reactions were quenched at defined time points by addition of 400 μL acetonitrile, followed by vigorous mixing. The vials were removed from the anaerobic chamber and further sample work-up and HPLC analysis was performed as described above.

Preparative scale reactions. *E. coli* BL21(DE3) cells freshly transformed with plasmid encoding P411 variants were grown overnight in 25 mL LB-Amp (37 ºC, 250 rpm). Next, 500 mL HB-Amp medium in a 2.8 L flask was inoculated with 20 mL of the pre-culture and incubated at 37 ºC, 230 rpm for 2 h (to OD_{600} ca. 1.5). Cultures were then cooled in an ice-water bath for 20 min and induced with 0.5 mM IPTG and 1.0 mM ALA (final concentrations). Expression was conducted at 24 ºC, 140 rpm, for 16–20 h. Cultures were then centrifuged (5,000 g, 5 min, 4 ºC) and the cell pellets were resuspended to OD_{600} = 50 in M9-N buffer supplemented with 20 mM glucose (typically yielding ca. 110 mL of cell suspension). Aliquots of the cell suspension (4 mL) were used to determine the P411 expression level after cell lysis by sonication. Aliquots of the cell suspension (46 mL) were then transferred to 100 mL Erlenmeyer flasks and degassed by sparging with argon for at least 30 minutes. The reaction flasks were then transferred into an anaerobic chamber and indole stock solution (2 mL, in DMSO) and tosyl azide (2 mL, in DMSO) were added. Final concentrations were typically 10 to 20 mM indole and 10 to 20 mM tosyl azide, with 8% co-solvent. The flasks were sealed with parafilm, removed from the anaerobic chamber, and shaken at room temperature, 100 rpm for 18 h. The reactions were quenched by adding acetonitrile (50 mL), transferred to two 50 mL tubes, and then centrifuged (4,000 g, 10 min). The supernatant was concentrated *in vacuo* to remove acetonitrile and extracted with EtOAc or 1:1 EtOAc:cyclohexane (3 x 25 mL). The organic layers were washed with brine (20 mL), dried over $MgSO_4$, filtered, concentrated, and purified by silica gel chromatography on a Biotage Isolera instrument. For several reactions, silica flash chromatography did not succeed to separate indole amidation products from tosyl sulfonamide; in these cases, preparative HPLC on an Agilent XDB-C18 column (*vide supra*) was subsequently used to isolate reaction products.

Determination of P411 concentration. Aliquots (4 mL) of the same *E. coli* cell suspensions used for whole-cell biocatalytic reactions were lysed by sonication (QSonica instrument, 1 min sonication time, 1 s on/off cycles, 40 % output). 1 mL aliquots of the cell lysate were cleared by centrifugation (13,000 g, 10 min, 4°C), and the concentration of P411 enzymes in the lysate was determined from ferrous carbon monoxide binding difference spectra[3] using the previously reported extinction coefficient for serine-ligated enzymes (ε = 103,000 M^{-1} cm^{-1}).[4] The concentration of purified P411 enzymes was determined by quantifying the amount of free hemin using the pyridine/hemochrome assay using the corresponding extinction coefficient (ε = 191,500 M^{-1} cm^{-1}).[5]

Protein purification. *E. coli* BL21(DE3) cells freshly transformed with plasmid encoding P411 variants were grown overnight in 25 mL LB-Amp (30 ºC, 250 rpm). Next, 500 mL HB-Amp medium in a 2.8 L flask was inoculated with 20 mL of the pre-culture and incubated at 37 ºC, 230 rpm for 2 h (to OD_{600} ca. 1.5). Cultures were then cooled in an ice-water bath for 20 min and induced with 0.5 mM IPTG and 1.0 mM ALA (final concentrations). Expression was conducted at 24 ºC, 140 rpm, for 16–20 h. Cultures were then centrifuged (5,000 g, 10 min, 4 ºC) and the cell pellets were frozen at –20 ºC. For protein purification, frozen cells from two such cultures were resuspended in buffer A (25 mM tris, 20 mM imidazole, 100 mM NaCl, pH 7.5, 4 mL buffer per gram of cell wet weight), supplemented with one protease inhibitor tablet (Roche), 1 mg/mL lysozyme, and 0.1 mg/mL DNAse I. The cell suspensions were lysed by sonication (1 min sonication time, 1 s on/off cycles, 40 % output). To pellet insoluble material, cell lysates were centrifuged (20,000 g, 20 min, 4 ºC), followed by filtering the cleared lysate through a 2 µM cut-off syringe-driven filter unit. P411 proteins were purified from the cleared lysate using a nickel NTA column (1 mL HisTrap HP, GE Healthcare) using an AKTAxpress purifier FPLC system (GE Healthcare). P411 enzymes were eluted on a linear gradient from 100% buffer A to 100% buffer B (25 mM tris, 300 mM imidazole, 100 mM NaCl, pH 7.5) over 10 column volumes. Fractions containing eluted protein were pooled and subjected to three rounds of buffer exchange to storage buffer (25 mM Tris-HCl, 100 mM NaCl, pH 7.5) using centrifugal spin filters (10 kDa molecular weight cut-off, Amicon Ultra). Subsequently, the concentrated protein was aliquoted, flash-frozen on powdered dry ice, and stored at –80 ºC. Protein concentrations were determined using the pyridine/hemochrome assay specified above.

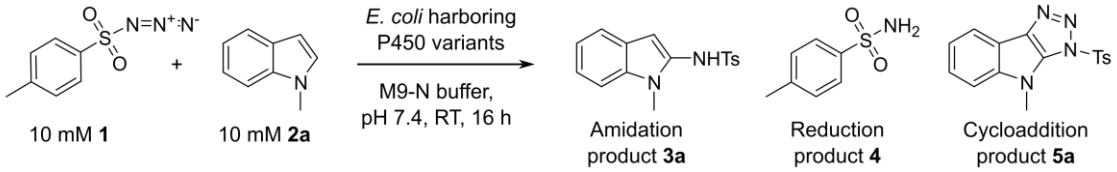
Amidation reactions using purified protein. Portions of M9-N buffer (290 μL) and NADPH (40 μL, 50 mM in M9-N), or multiples thereof, were combined in a 6 mL crimp vial and degassed by sparging with argon for at least 30 minutes. Purified protein solutions were adjusted to 20 μM in M9-N. After degassing was complete, the M9-N/reductant solution and purified protein stocks were brought into the anaerobic chamber. Reactions were set up on a 400 μL scale in 2 mL crimp vials: first, 340 μL of the M9-N/NADPH solution were added per vial, followed by 40 μL of purified P411 solution. Next, 10 μL of indole stock solution (in DMSO) were added, followed by 10 μL of tosyl azide stock solution (in DMSO). Final concentrations were typically 5 mM indole, 5 mM tosyl azide, 5 mM NADPH, and 2 μM P411, with 5% co-solvent. To assess tosyl azide reduction rates, reactions were set up as above but without addition of indole; the volume of M9-N buffer was adjusted accordingly to yield a total reaction volume of 400 μL. The vials were sealed, removed from the anaerobic chamber, and shaken at room temperature, 400 rpm for 1 to 6 h. The reactions were quenched by adding acetonitrile (400 μL) supplemented with internal standard. This mixture was transferred to a 1.5 mL tube and centrifuged at 13,000 g for 10 minutes. The supernatant was transferred to 2 mL vials and analyzed by HPLC.

Assessment of electron transfer rates with the cytochrome *c* reduction assay. Electron transfer rates from NADPH through the reductase domain were assessed as described.[3] First, 920 μL of 0.1 M potassium phosphate buffer (KPi), pH 8.0, were transferred to a 1 cm pathlength spectrophotometric cuvette. Next, 50 μL of a 10 mg/mL solution of horse heart cytochrome *c* (Sigma-Aldrich; in 0.1 M KPi, pH 8.0) were added, followed by 10 μL of purified P411 variant (pre-diluted 20 to 800-fold in 0.1 M KPi, pH 8.0; purified P411 stock solutions are typically 100 to 300 μM). The cuvette was sealed with parafilm, gently inverted a few times to mix, and baseline absorption at 550 nm was recorded for 1 to 2 min in a spectrophotometer in kinetic mode. Next, 20 μL of 5 mM NADPH (in 0.1 M KPi, pH 8.0) were added, the cuvette was inverted a few times to mix, and absorption was recorded immediately for an additional 2 to 5 minutes (as long as the increase in absorption remained linear). Several dilutions were tested for each P411 variant. The resulting data were analyzed as described, using the extinction coefficient of reduced cytochrome *c* (21,000 M^{-1} cm^{-1}), to yield P411 electron transfer activity.[3]

Synthesis of amidation or cycloaddition authentic product standards. All authentic amidation or cycloaddition product standards used in this study (see Figure 4, main text) were synthesized by preparative-scale enzymatic reactions (see *Preparative scale reactions*, p. S5). Authentic standard of tosyl sulfonamide **4** was commercially available (Sigma-Aldrich).

II. Supplementary Tables

Table S1: Hemoprotein variant testing for 1-methylindole amidation.



Variant[a]	Description, mutation, references	Yield 3a	Yield 4	Yield 5a
Hemin[b]	Free hemin in M9-N buffer	1.2	52	0.3
E. coli BL21(DE3)	Whole *E. coli* cells, no heterologous protein overexpression	1.7	40	0.4
P450-BM3	Wild-type P450-BM3 from *Bacillus megaterium*	1.6	30	0.5
P450-BM3 T268A C400S	P450-BM3 with mutations T268A and C400S[6]	1.7	30	0.5
P450-CIS	P450-BM3 with mutations V78A F87V P142S T175I A184V S226R H236Q E252G T268A A290V L353V I366V E442K[7]	1.7	28	0.5
P411-CIS	P450-BM3 with mutations V78A F87V P142S T175I A184V S226R H236Q E252G T268A A290V L353V I366V C400S E442K[6]	2.1	43	0.4

Variant[a]	Description, mutation, references	Yield 3a	Yield 4	Yield 5a
P450 HStar H92N H100N	P450-BM3 with mutations V78M H92N H100N L181V T268A C400H L437W[8,9]	1.6	31	0.6
P411-HF	P411-CIS with mutations V87T H92F L181G N201S L215Q I263M V281L T438C K472T N573D F646S Q674*	1.9	34	0.5
CYP119	CYP119 from *Sulfolobus acidocaldarius*[10]	1.8	30	0.6
P450-CAM	P450-CAM from *Pseudomonas putida*[11]	1.7	33	0.5

[a] Reactions were performed on 400 µL scale with whole *E. coli* cells expressing the respective hemoprotein variants, at 10 mM substrate loading (1-methylindole and tosyl azide). Results are the average of duplicate reactions. n.d. = not detected. [b] Reaction with hemin: in the anaerobic chamber, 10 µL of hemin solution (2 mM in DMSO) were added to 350 µL of argon-sparged M9-N buffer, followed by 20 µL of sodium dithionite as reductant (100 mM in M9-N buffer), 1-methylindole (10 µL, 400 mM in DMSO) and tosyl azide (10 µL, 400 mM in DMSO). Final concentrations were 50 µM hemin, 5 mM sodium dithionite, 10 mM 1-methylindole and 10 mM tosyl azide with 5 % co-solvent.

Table S2. Summary of directed evolution for indole amidation.

Round	Parent variant	Libraries evaluated[a]	Screening substrates	Mutations identified
1	P411-CIS (P)	I263X	1-methylindole (10 mM), tosyl azide (10 mM)	I263Y
2	P-I263Y	T438X	1-methylindole (10 mM), tosyl azide (10 mM)	T438S
3	P-I263Y T438S	V87X, L181X, A328X	1-methylindole (10 mM), tosyl azide (10 mM)	A328V
4	P-I263Y T438S A328V	A82X, L181X, E267X, T327X	1-methylindole (10 mM), tosyl azide (10 mM)	A82W
5	P-I263Y T438S A328V A82W	A268X	1-methylindole (15 mM), tosyl azide (15 mM)	A268G
6	P-I263Y T438S A328V A82W A268G (P-YSVWG)	S830X, D1044X, W1046X	1-methylindole (15 mM), tosyl azide (15 mM)	W1046P[a]
7a	P411-IA	W82X, L181X, V328X	indole (10 mM), tosyl azide (10 mM), screen for indole amidation	W82C[b]
7b	P411-IA	W82X, L181X, V328X	indole (10 mM), tosyl azide (10 mM), screen for cycloaddition	W82F[c]
7c	P-I263Y T438S	A82X, V87X, L181X	1-methylindole (10 mM), tosyl azide (10 mM), screen for cycloaddition	A82S[d]

[a]Final variant for 1-methylindole amidation: **P411-IA** = P411-CIS I263Y T438S A328V A82W A268G W1046P

[b]Variant for indole amidation: P411-IA W82C

[c]Variant for indole-azide cycloaddition: P411-IA W82F

[d]Variant for 1-methylindole-azide cycloaddition: P411-CIS I263Y T438S A82S

Table S3: Tabulation of data shown in Figure 2 (main text).

Variant	**3a Yield (%)[a]**	**3a TTN**	**4 Yield (%)**	**4 TTN**	**5a Yield (%)**	**5a TTN**	**Selectivity 3a:4:5a**
E. coli BL21(DE3)	2.9 ± 1.2	n.a.[b]	43.2 ±2.9	n.a.	1.0 ± 0.3	n.a.	3:43:1
P411-CIS (P)	3.9 ± 0.9	440 ± 55	34.5 ± 3.5	4147 ± 1197	1.8 ± 0.1	224 ± 73	2:19:1
P-I263Y	5.4 ± 1.3	584 ± 62	25.9 ± 3.4	2923 ± 749	4.0 ± 0.2	478 ± 172	1:6:1
P-I263Y T438S	10.9 ± 2.9	1336 ± 182	25.8 ± 4.1	3197 ± 227	13.0 ± 0.4	1645 ± 227	1:2:1
P-I263Y T438S A328V	34.9 ± 9.6	2208 ± 44	20.3 ± 5.0	1295 ± 36	11.2 ± 2.0	729 ± 73	3:2:1
P-I263Y T438S A328V A82W	42.5 ± 8.3	2919 ± 160	17.9 ± 2.8	1237 ± 55	1.7 ± 0.2	118 ± 30	25:11:1
P-I263Y T438S A328V A82W A268G "P-YSVWG"	67.9 ± 7.1	5730 ± 554	14.2 ± 2.6	1198 ± 241	0.7 ± 0.1	59 ± 10	97:20:1
P-I263Y T438S A328V A82W A268G W1046P "P411-IA"	77.8 ± 2.4	8413 ± 1024	8.6 ± 2.2	941 ± 331	0.7 ± 0.1	77 ± 16	110:12:1

[a] Assays were performed with 25 mM substrate loading.

[b] n.a., not applicable.

Table S4: Tabulation of data shown in Figure 4 (main text).

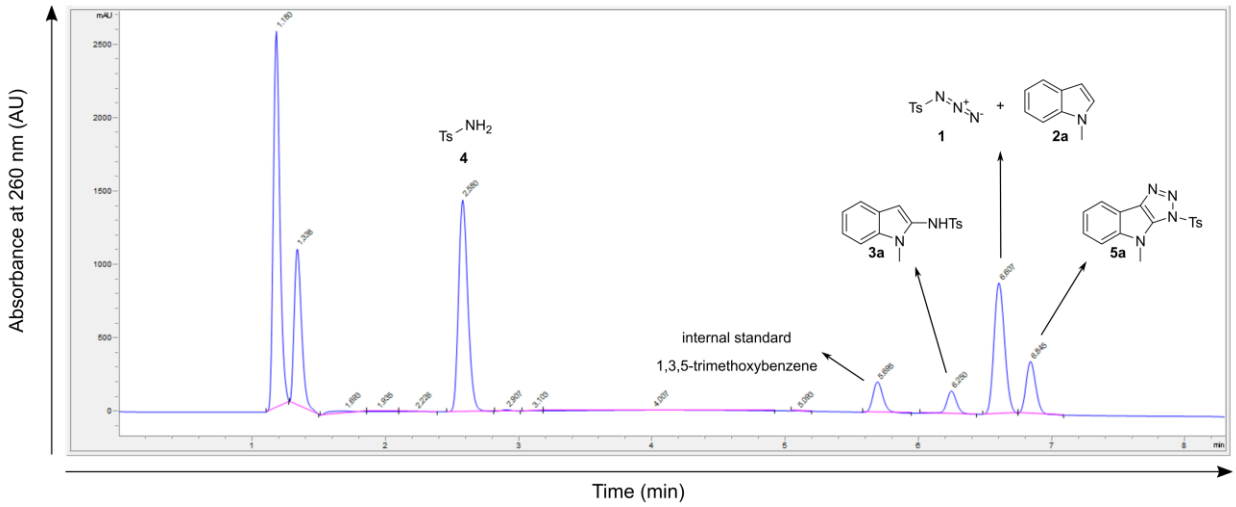
Variant	**3 Yield (%) Amidation**	**4 Yield (%) $TsNH_2$**	**5 Yield (%) Cycloaddition**	**Selectivity 3:4:5**
P411-IA	78 (**3a**)[a]	8.6[a]	0.7 (**5a**)[a]	110:12:1
P411-IA W82C	45 (**3b**)	5.3	12 (**5b**)	9:1:2
P411-IA	81 (**3c**)	18	n.d.[b]	4:1 (**3c**:**4**)
P411-IA	76 (**3d**)	22	n.d.[b]	3:1 (**3d**:**4**)
P411-IA	91 (**3e**)	7	n.d.[b]	12:1 (**3e**:**4**)
P411-IA	64 (**3f**)	30	n.d.[b]	2:1 (**3f**:**4**)
P411-CIS I263Y T438S A82S	8.0 (**3a**)	59	24 (**5a**)	1:7:3
P411-IA W82F	2.0 (**3b**)	5.2	16 (**5b**)	1:3:8

[a] Assays were performed with 25 mM substrate loading; all other reactions were performed at 10 mM substrate loading.

[b] n.d., not determined. Only small amounts of cycloaddition product were produced in these reactions.

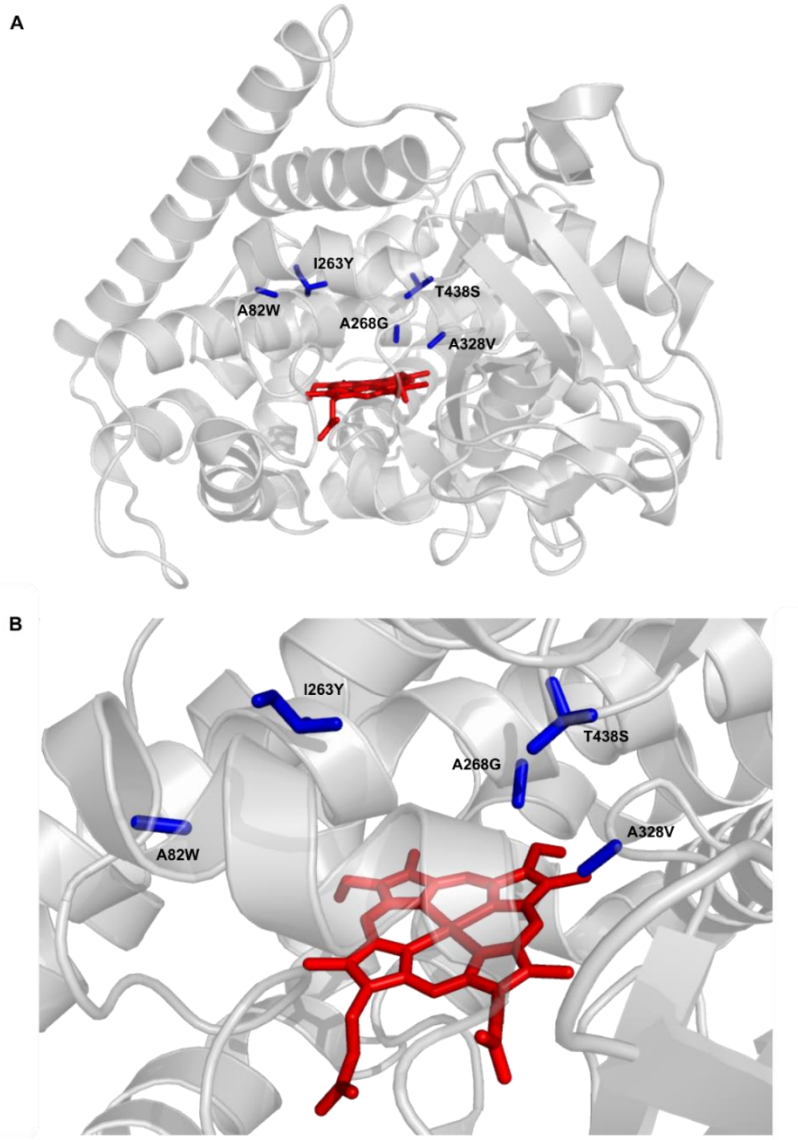
III. Supplementary Figures

Figure S1: Exemplary HPLC trace for enzymatic whole-cell amidation reactions with 1-methylindole and tosyl azide.



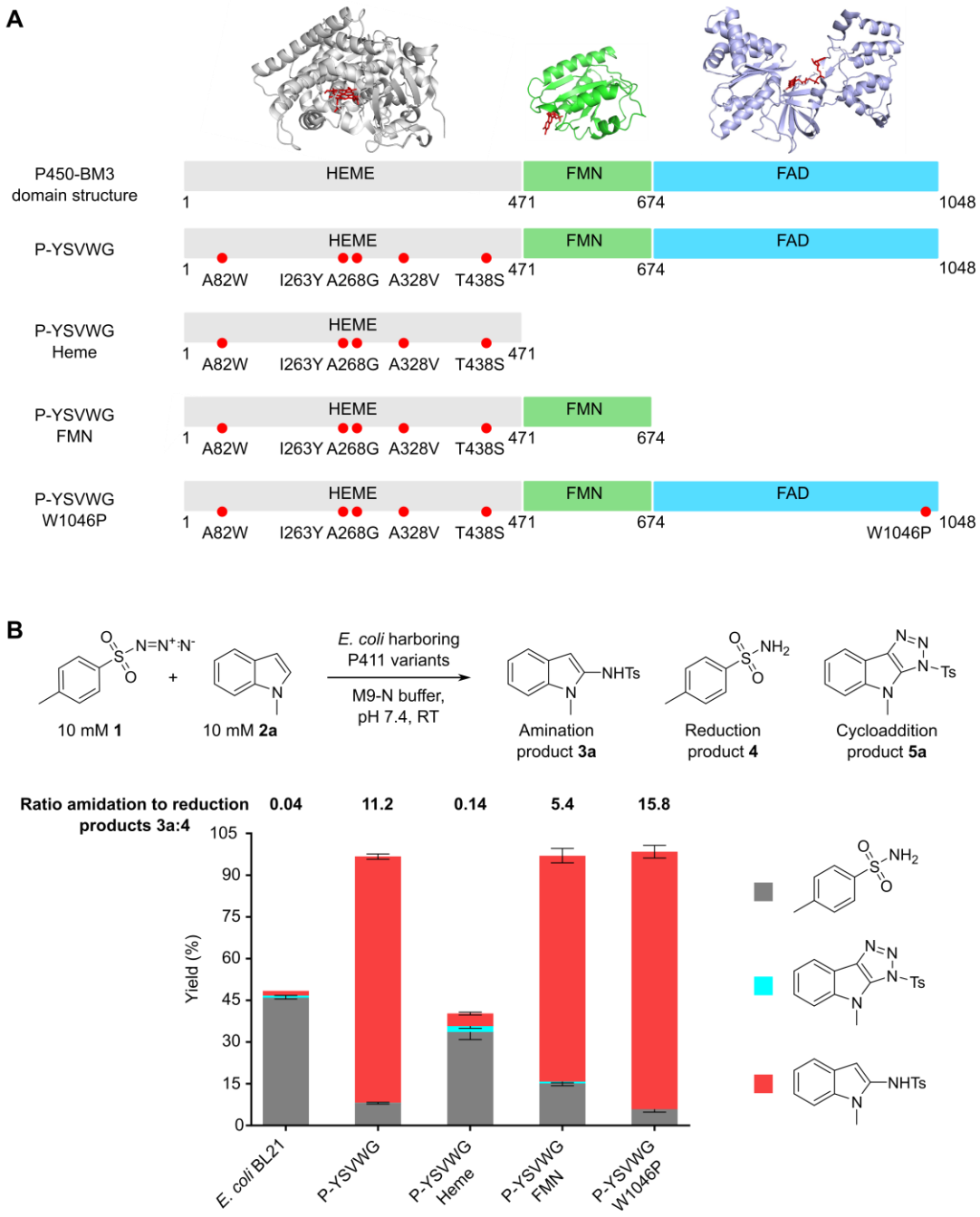
Shown is a HPLC trace obtained from whole-cell reactions performed with variant P411-CIS I263Y T438S (P-YS; see Figure 2 and Table S3). Whole cell reaction mixtures were worked up with acetonitrile and analyzed on an Agilent XDB-C18 column (4.6 x 150 mm, 5 µm), using water and acetonitrile as mobile phase (HPLC method: 50% acetonitrile starting concentration, 50% to 80% acetonitrile over 7 min; 80% to 100% acetonitrile over 0.5 min; 100% acetonitrile for 0.5 min). The identity of all relevant peaks is indicated; 1,3,5-trimethoxybenzene was used as internal standard. Starting materials tosyl azide **1** and 1-methylindole **2a** showed co-elution under these conditions.

Figure S2: Position of heme domain mutations in P411-CIS YSVWG.



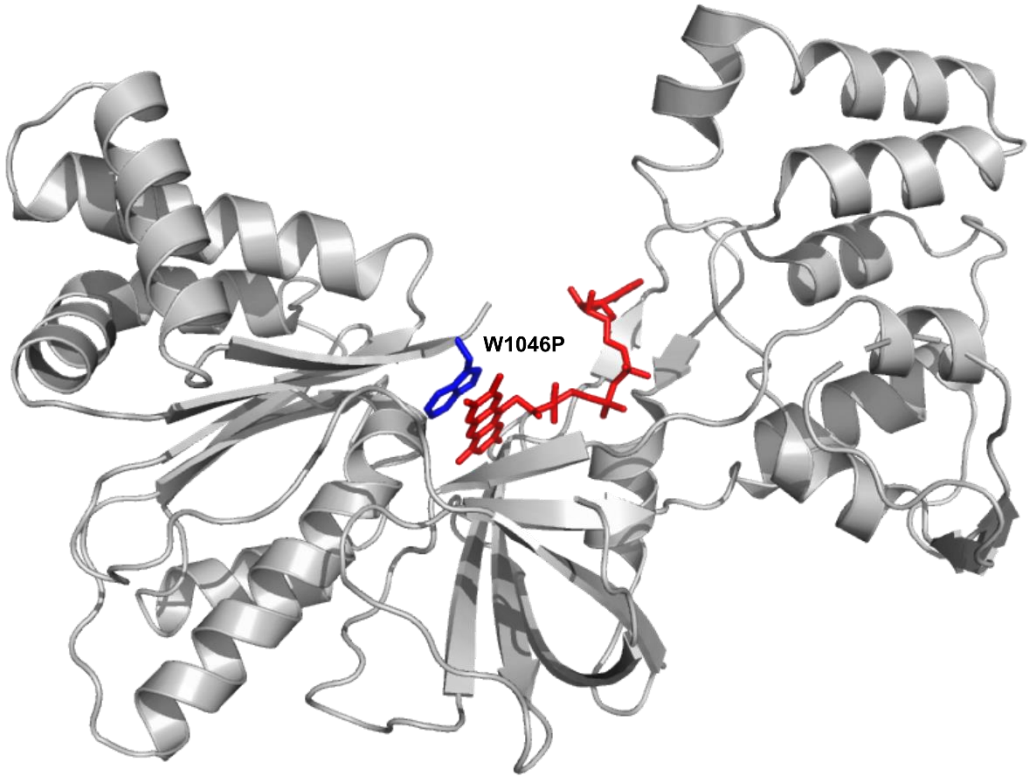
(A) Overall structure of the P411-CIS heme domain (pdb: 4H23),[6] with residues mutated in P411-CIS YSVWG (P-YSVWG) shown as sticks and highlighted in blue; the heme cofactor is shown in red. (B) Close-up view of the active site.

Figure S3: Reductase domain engineering to decrease nitrene reduction.



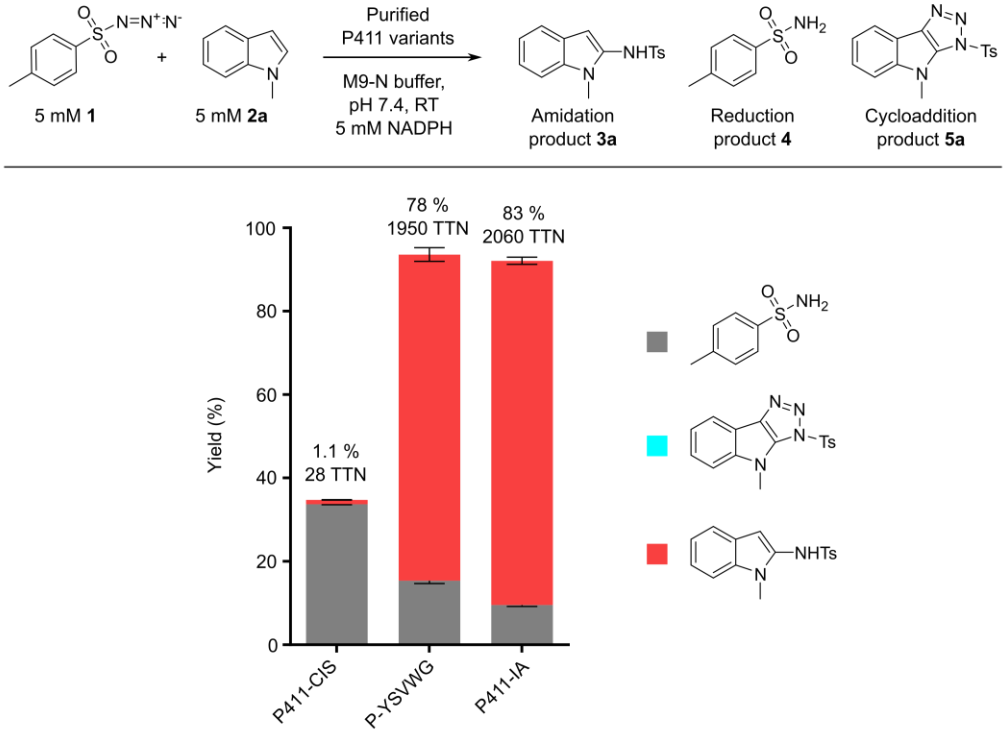
(A) Schematic overview of P411 domain structure and mutants tested to reduce unproductive nitrene reduction. (B) Yield of P-YSVWG and reductase mutants in whole-cell amidation reactions; point mutant W1046P proved superior in reducing formation of sulfonamide **4**, improving the ratio between amidation and nitrene reduction *ca.* 1.4-fold compared to P-YSVWG.

Figure S4: Position of the W1046P mutation in the P450-BM3 FAD domain.



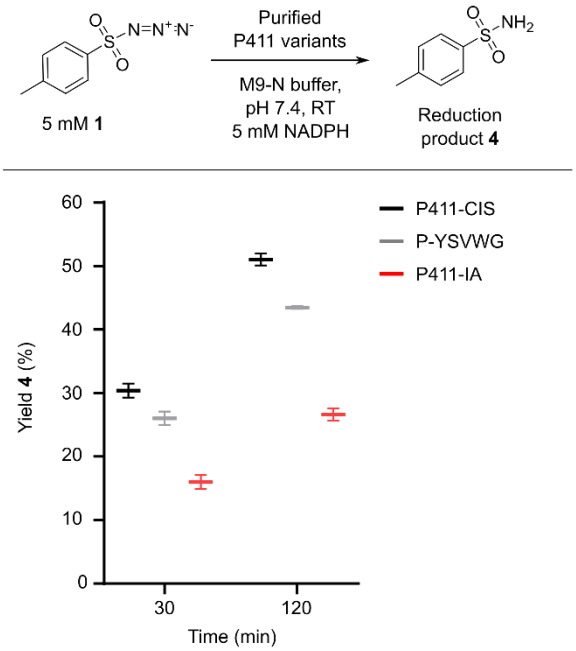
Crystal structure of the P450-BM3 FAD domain (pdb: 4DQK).[12] The FAD cofactor is shown as red sticks; residue W1046, located close to the C-terminus of the protein, is shown as blue sticks.

Figure S5: P411-catalyzed 1-methylindole amidation *in vitro*.



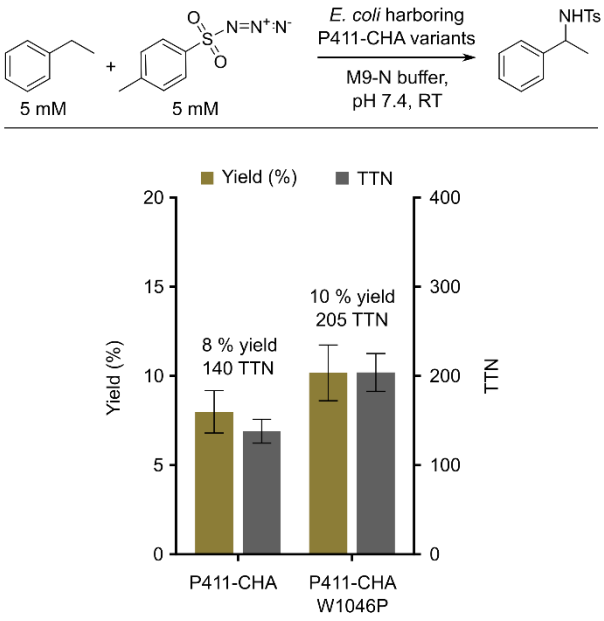
1-methylindole amidation reactions were performed *in vitro* using purified P411 variants. Reactions were set up at 5 mM substrate loading with 5 mM NADPH as reducing agent and with 2 µM P411 variants. Reaction yield and TTN for amidation product **3a** are indicated.

Figure S6: P411-catalyzed tosyl azide reduction *in vitro*.



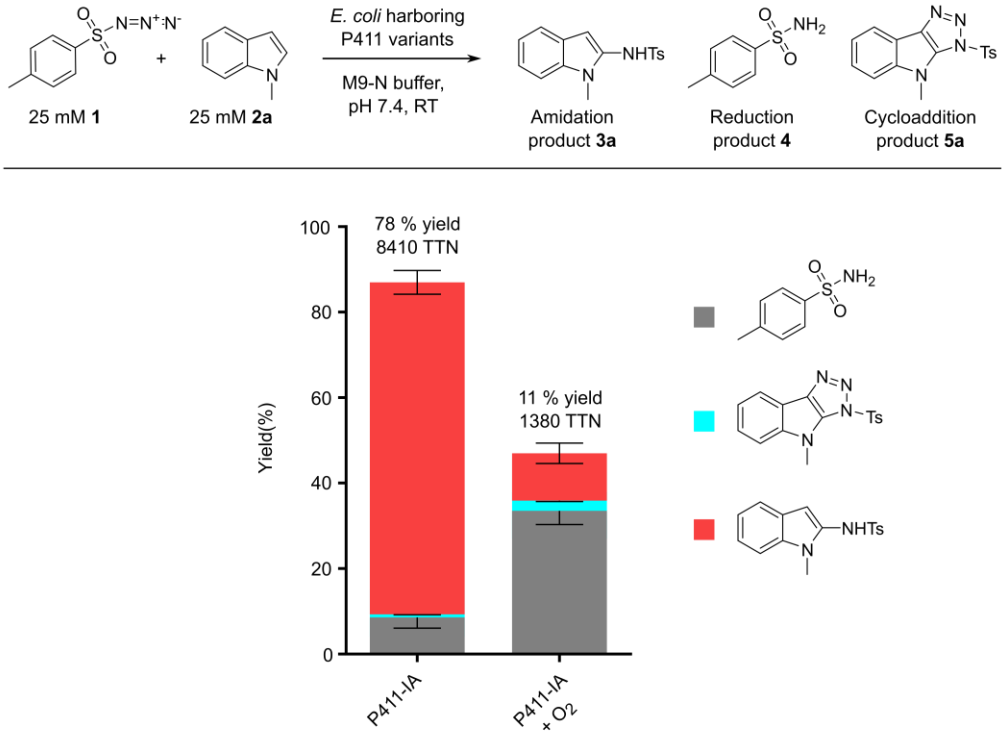
Tosyl azide reduction catalyzed by P411 variants was analyzed by setting up reactions *in vitro* using purified enzyme (2 µM) in absence of a nitrene acceptor substrate. Reactions were analyzed after 30 or 120 minutes; the HPLC yields of sulfonamide **4** are shown. P411-IA showed the lowest rates of nitrene reduction / sulfonamide formation, in agreement with the theory that the W1046P mutation slows down electron transfer and thus nitrene reduction. Of note, variant P-YSVWG, identical with P411-IA but lacking the W1046P mutation, showed slightly reduced sulfonamide formation rates compared to the parent variant P411-CIS, indicating that the accumulated heme domain mutations (I263Y T438S A328V A82W A268G) may have a minor effect on nitrene reduction rate. The 30-minute time point was used to calculate rates of nitrene reduction shown in Figure 3A in the main text.

Figure S7: Effect of W1046P in P411-CHA catalyzed C–H amidation.



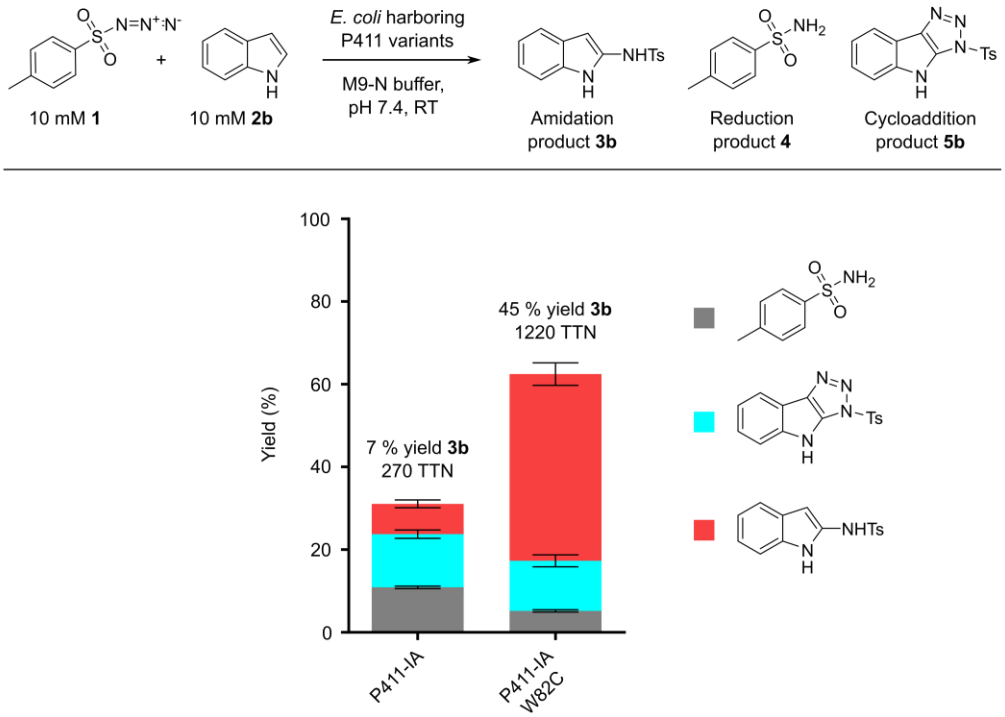
To analyze whether the W1046P mutation represents a generalizable solution to improve activity of P411 nitrene transferases, we cloned the W1046P variant of P411-CHA, an enzyme previously evolved to catalyze amidation of C(*sp3*)-H bonds.[13] We compared the two enzymes in the amidation of ethyl benzene, and found that the W1046P mutation delivered a slight increase in reaction yield (*ca.* 1.25-fold) and TTN (*ca.* 1.5-fold).

Figure S8: P411-IA catalyzed 1-methylindole amidation under aerobic conditions.



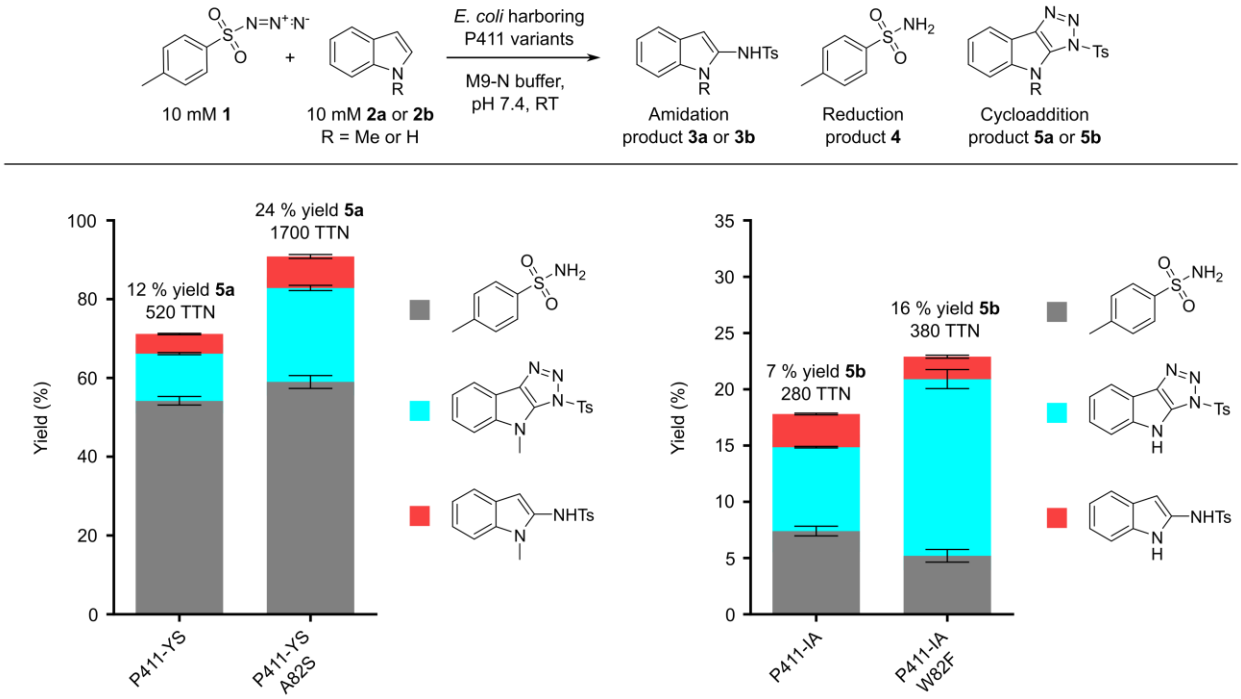
Comparison of P411-IA whole-cell amidation reactions under anaerobic (left column) and aerobic (right column) conditions. The standard protocol for setting anaerobic reactions is provided in SI section I, *Experimental Procedures*, p. S4. Aerobic reactions were set up on the bench under standard atmosphere without prior degassing of cells or substrates. Reaction yields and TTN for amidation product **3a** are given, showing a *ca.* 7-fold reduction in enzymatic amidation activity in presence of oxygen.

Figure S9: Unprotected indole amidation by P411-IA W82C.



P411-IA showed only low amidation activity on indole **2b** to yield product **3b** (7% yield, 270 TTN); curiously, the enzymatic reaction predominantly delivered the indole-azide cycloaddition product **5b**. However, screening of site-saturation mutagenesis libraries at positions W82, L181, and V328 delivered P411-IA W82C, showing ca. 6-fold higher yields of the desired indole amidation product **3b**. This variant may provide a starting point for further evolution of unprotected indole C_2 amidation activity.

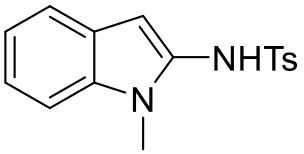
Figure S10: Evolving P411 variants for indole-azide cycloaddition activity.



To test whether the observed indole-azide cycloaddition activity was indeed susceptible to P411 active site mutations and could thus potentially be targeted by directed evolution, we performed one additional round of site-saturation mutagenesis and screening. Variant P411-YS, providing the highest yields of cycloaddition product **5a** within the P411-IA lineage (Figure 2), was randomized at positions W82, V87, and L181 (Table S2), and libraries were screened with **1** and **2a** for enhanced cycloaddition activity. This delivered P-YS A82S, showing higher yield and TTN in the formation of **5a**. Similarly, P411-IA site-saturation libraries of positions W82, L181, and V328 were screened with **1** and **2b** for formation of **5b**, which delivered P411-IA W82F, showing enhanced cycloaddition activity. While this indicates that P411 indole-azide cycloaddition activity is evolvable and that the reaction takes place within or close to the P411 active site, further work is required to investigate the P411 cycloaddition reaction mechanism.

Figure S11: Calibration curve for **3a**

Analyte:



Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

Wavelength: 230 nm

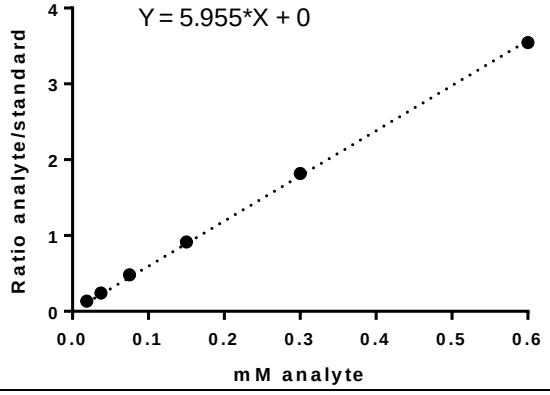
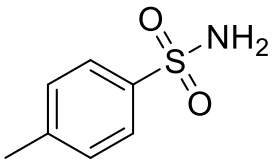


Figure S12: Calibration curve for **4**

Analyte:



Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

Wavelength: 230 nm

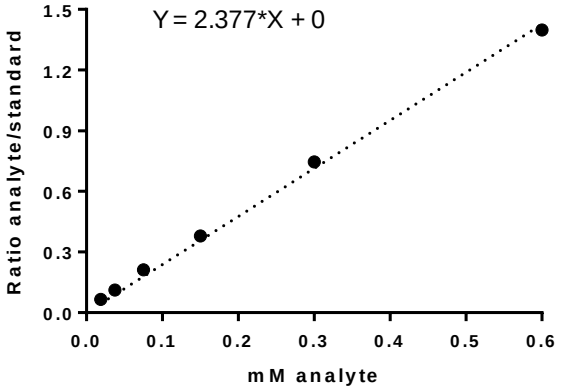
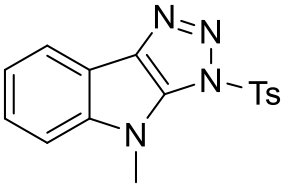


Figure S13: Calibration curve for **5a**

Analyte:



Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

Wavelength: 230 nm

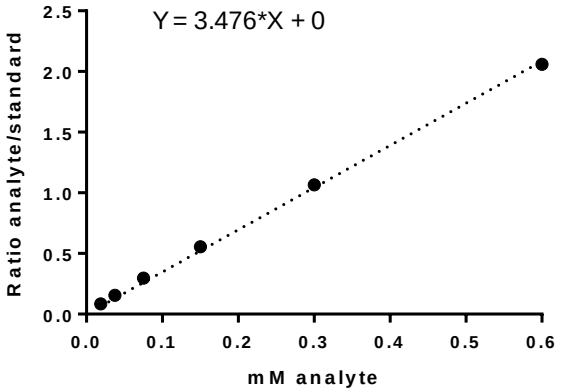
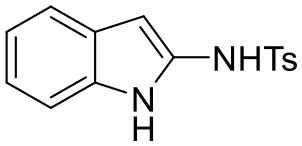


Figure S14: Calibration curve for **3b**

Analyte:



Column: Agilent XDB-C18

Standard: 1-phenyl-3-propanol

Wavelength: 260 nm

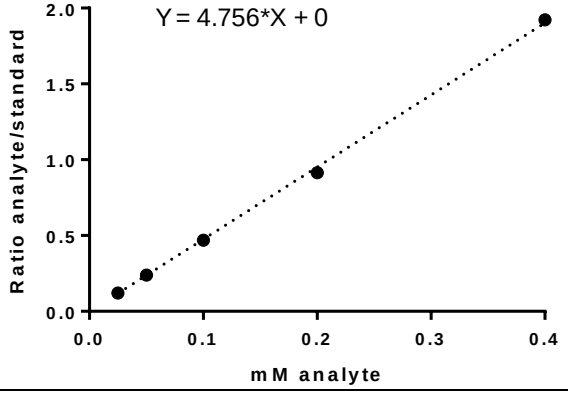
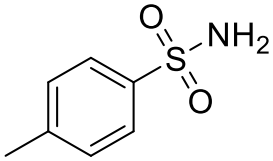


Figure S15: Calibration curve for **4**

Analyte:



Column: Agilent XDB-C18

Standard: 1-phenyl-3-propanol

Wavelength: 260 nm

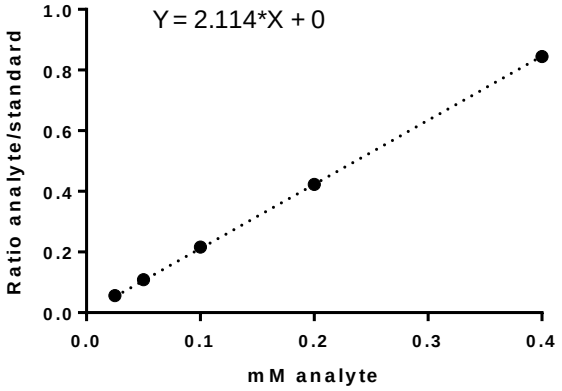
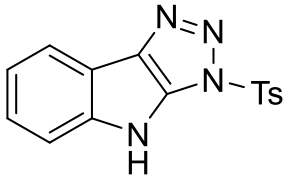


Figure S16: Calibration curve for **5b**

Analyte:



Column: Agilent XDB-C18

Standard: 1-phenyl-3-propanol

Wavelength: 260 nm

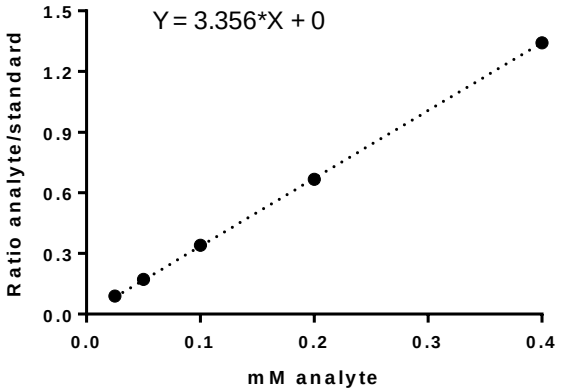
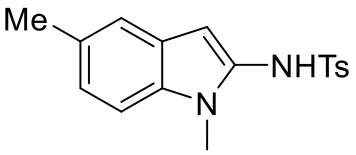


Figure S17: Calibration curve for **3c**

Analyte:



Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

Wavelength: 230 nm

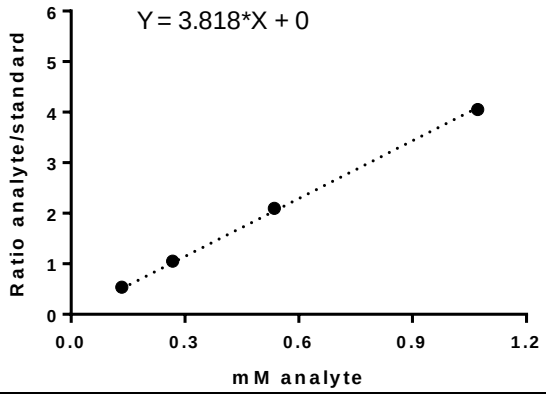
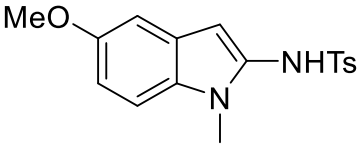


Figure S18: Calibration curve for **3d**

Analyte:



Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

Wavelength: 230 nm

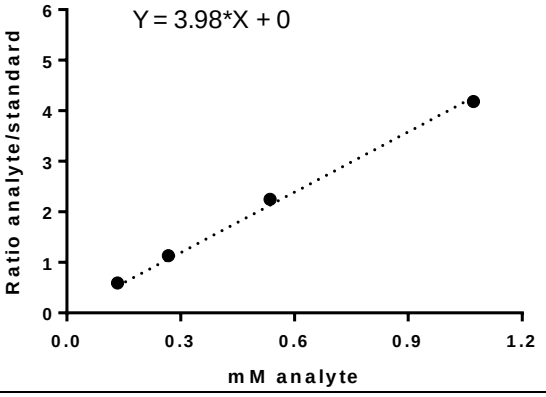
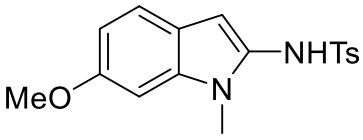


Figure S19: Calibration curve for **3e**

Analyte:



Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

Wavelength: 230 nm

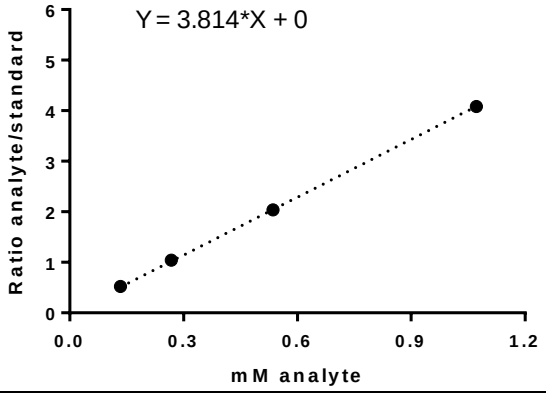
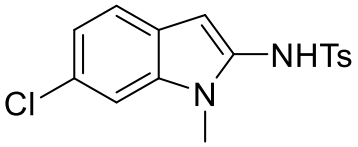


Figure S20: Calibration curve for **3f**

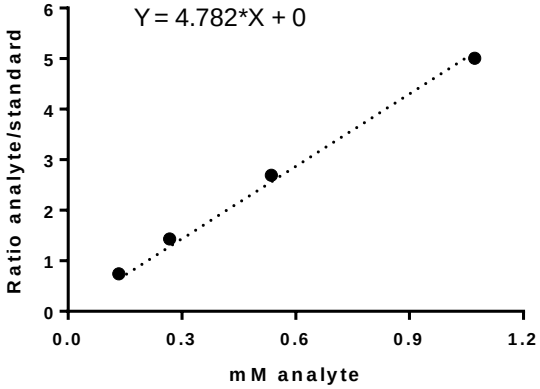
Analyte:



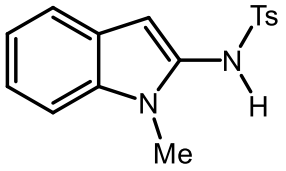
Column: Agilent XDB-C18

Standard: 1,3,5-trimethoxybenzene

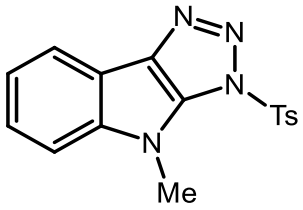
Wavelength: 230 nm



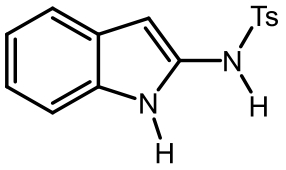
IV. Characterization of Reaction Products



4-methyl-N-(1-methyl-1H-indol-2-yl)benzenesulfonamide 3a. Obtained as a light tan solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.68 – 7.58 (m, 2H), 7.25 – 7.22 (m, 1H), 7.20 – 7.13 (m, 3H), 7.06 – 7.01 (m, 2H), 6.96 (br s, 1H), 3.74 (s, 3H), 2.37 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.43, 136.80, 135.22, 129.41, 127.37, 126.53, 124.45, 122.16, 119.77, 117.44, 110.44, 109.42, 31.96, 21.51. HRMS (FAB+) exact mass calculated for $C_{16}H_{16}N_2O_2S^+$ requires *m/z* 300.0933, found 300.0923.

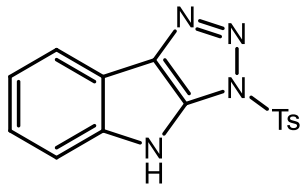


4-methyl-3-tosyl-3,4-dihydro-[1,2,3]triazolo[4,5-b]indole 5a. Obtained as a light pink solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.92 – 7.85 (m, 2H), 7.32 – 7.26 (m, 3H), 7.25 – 7.17 (m, 2H), 7.09 (dt, *J* = 7.9, 0.9 Hz, 1H), 3.46 (s, 3H), 2.42 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.32, 142.45, 140.12, 134.26, 129.35, 126.33, 125.89, 123.02, 118.97, 116.88, 109.70, 29.06, 21.51. HRMS (FAB+) exact mass calculated for $C_{16}H_{15}N_4O_2S^+$ requires *m/z* 327.0916, found 327.0937. Characterization data is consistent with prior reported data and supports the assignment of the regiochemistry of the cycloaddition product.[14]

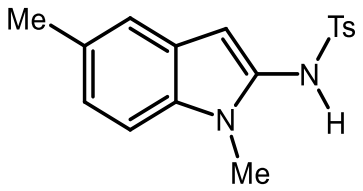


***N*-(1H-indol-2-yl)-4-methylbenzenesulfonamide 3b.** Obtained as a colorless solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 (br s, 1H), 7.67 – 7.58 (m, 2H), 7.32 – 7.28 (m, 1H), 7.19 – 7.11 (m, 5H), 7.03 – 6.94 (m, 1H), 6.20 (s, 1H), 2.36 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.50, 136.59,

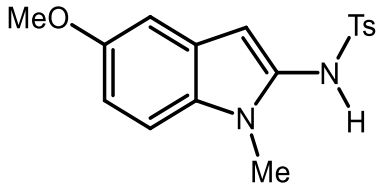
134.38, 129.42, 127.39, 124.01, 122.73, 121.66, 120.31, 117.45, 112.41, 111.38. HRMS (FAB+) exact mass calculated for $C_{15}H_{15}N_2O_2S^+$ requires *m/z* 287.00854, found 287.0842.



3-tosyl-3,4-dihydro-[1,2,3]triazolo[4,5-b]indole 5b. Obtained as a colorless solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 – 7.84 (m, 2H), 7.32 – 7.20 (m, 4H), 7.20 – 7.10 (m, 2H), 2.41 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.70, 142.95, 139.48, 138.93, 132.37, 129.43, 126.34, 126.31, 123.48, 118.01, 116.57, 111.66, 21.52. HRMS (FAB+) exact mass calculated for $C_{15}H_{13}N_4O_2S^+$ requires *m/z* 313.0759, found 313.0766.

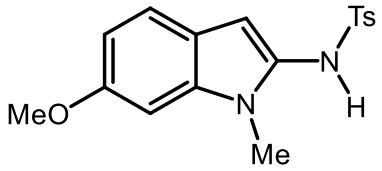


***N*-(1,5-dimethyl-1H-indol-2-yl)-4-methylbenzenesulfonamide 3c.** Obtained as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.68 – 7.60 (m, 2H), 7.21 – 7.15 (m, 2H), 7.12 (d, *J* = 8.4 Hz, 1H), 6.99 – 6.94 (m, 2H), 6.69 (br s, 1H), 6.09 (s, 1H), 3.70 (s, 3H), 2.38 (s, 3H), 2.27 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.38, 136.92, 133.73, 129.37, 129.08, 127.53, 126.80, 124.77, 123.80, 117.09, 109.78, 109.09, 32.97, 21.47, 21.19. HRMS (FAB+) exact mass calculated for $C_{17}H_{19}N_2O_2S^+$ requires *m/z* 315.1167, found 315.1168.

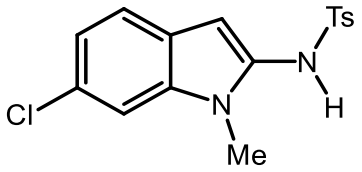


***N*-(5-methoxy-1-methyl-1H-indol-2-yl)-4-methylbenzenesulfonamide 3d.** Obtained as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72 – 7.61 (m, 2H), 7.22 – 7.15 (m, 2H), 7.12 (d, *J* = 8.9 Hz, 1H), 6.97 (s, 1H), 6.79 (dd, *J* = 8.8, 2.4 Hz, 1H), 6.39 (d, *J* = 2.4 Hz, 1H), 6.22 (br s, 1H), 3.70 (s, 3H), 3.63 (s, 3H), 2.38 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.31, 143.38, 137.07, 130.54,

129.43, 127.50 (two overlapping signals, confirmed by HSQC), 125.09, 112.99, 110.39, 109.82, 98.53, 55.40, 33.10, 21.47. HRMS (FAB+) exact mass calculated for $C_{17}H_{19}N_2O_3S^+$ requires *m/z* 331.1116, found 331.1137.

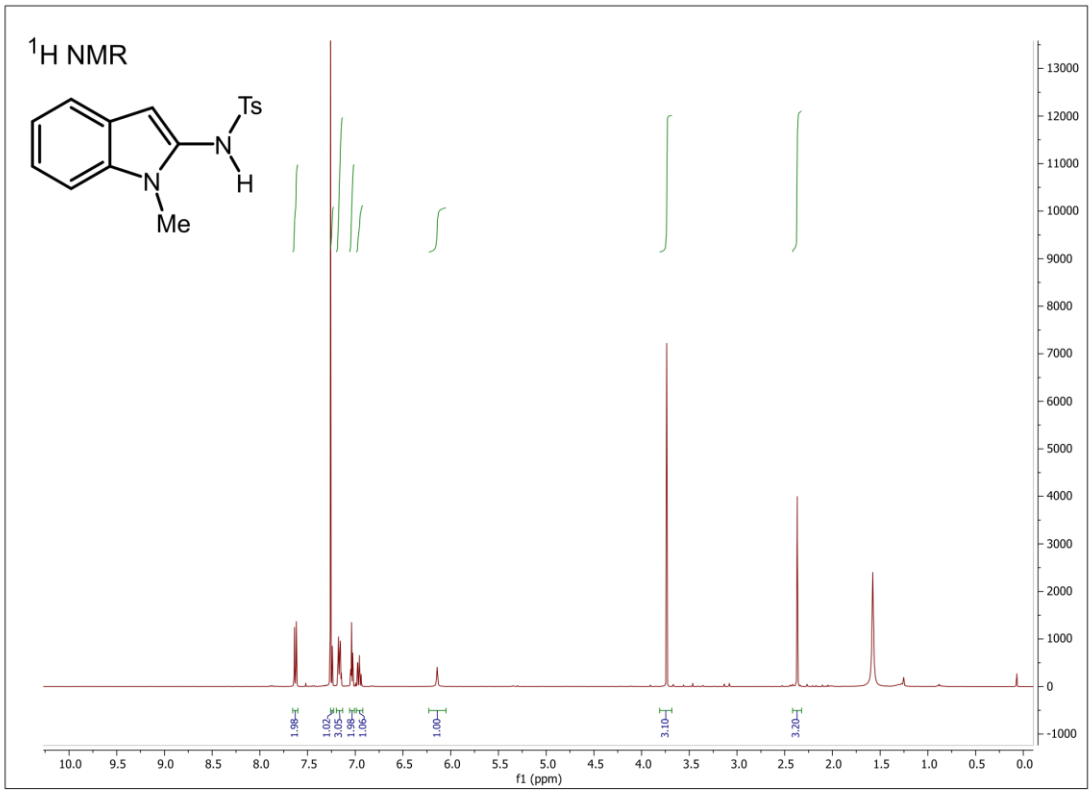


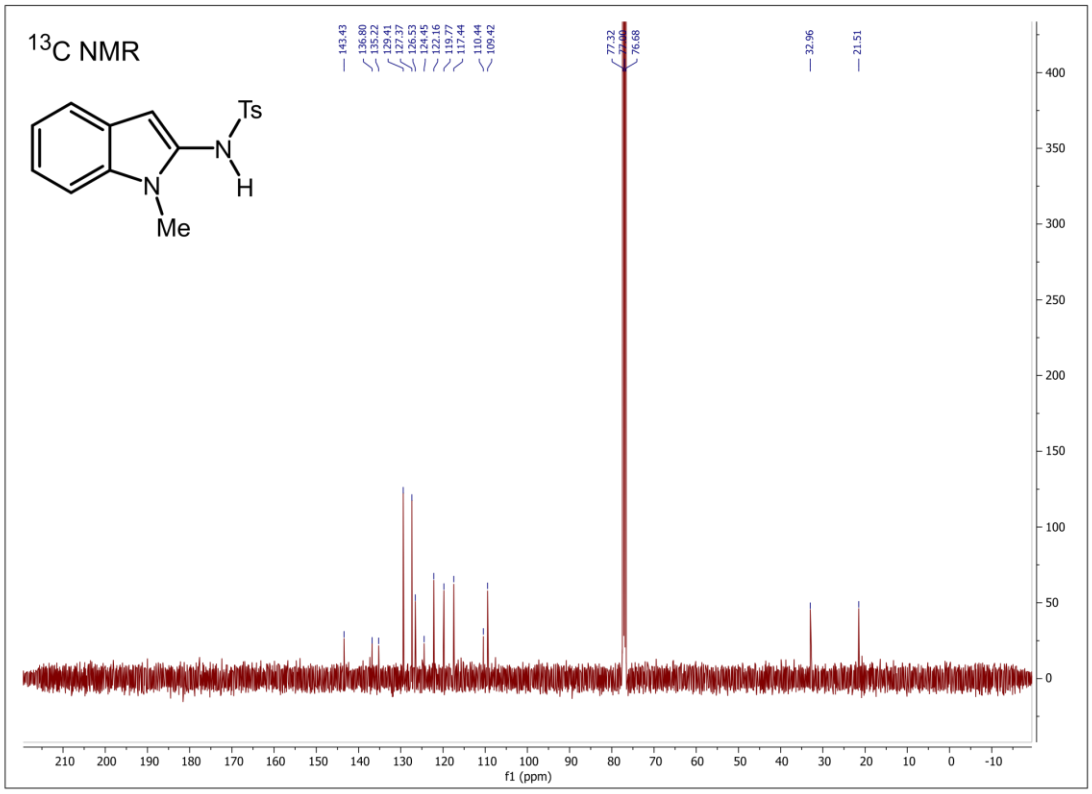
***N*-(6-methoxy-1-methyl-1H-indol-2-yl)-4-methylbenzenesulfonamide 3e.** Obtained as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.69 – 7.60 (m, 2H), 7.21 – 7.11 (m, 2H), 6.97 – 6.85 (m, 2H), 6.67 (d, J = 2.1 Hz, 1H), 6.62 (dd, J = 8.7, 2.2 Hz, 1H), 6.31 (br s, 1H), 3.83 (s, 3H), 3.66 (s, 3H), 2.37 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 156.64, 143.42, 136.79, 136.00, 129.44, 127.41, 125.07, 118.83, 118.42, 110.63, 109.97, 92.65, 55.65, 32.97, 21.55. HRMS (FAB+) exact mass calculated for $C_{17}H_{19}N_2O_3S^+$ requires *m/z* 331.1116, found 331.1131.

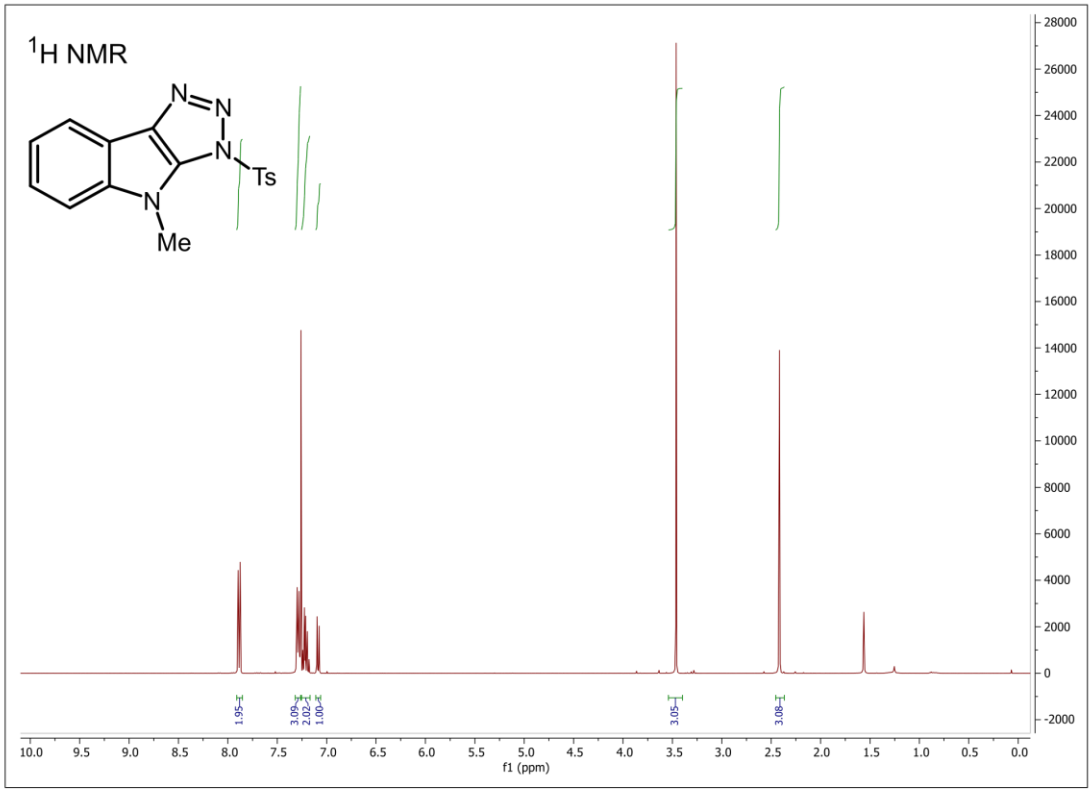


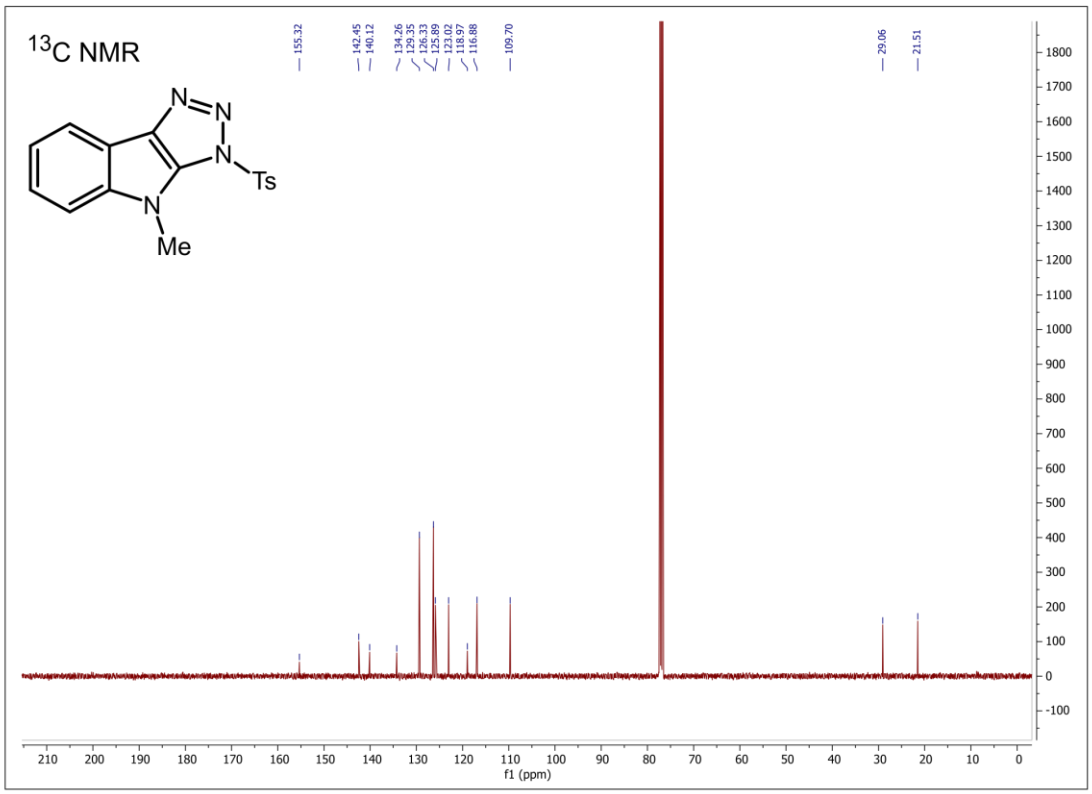
***N*-(6-chloro-1-methyl-1H-indol-2-yl)-4-methylbenzenesulfonamide 3f.** Obtained as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.64 – 7.58 (m, 2H), 7.24 (dd, *J* = 1.6, 0.7 Hz, 1H), 7.18 (m, 1H), 7.01 (d, *J* = 0.9 Hz, 1H), 6.93 (m, 2H), 6.14 (br s, 1 H), 3.70 (s, 3H), 2.38 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.63, 136.60, 135.59, 129.49, 128.33, 127.34, 127.25, 123.11, 120.59, 118.55, 110.76, 109.56, 33.08, 21.54. HRMS (FAB+) exact mass calculated for $C_{16}H_{16}N_2O_2SCl^+$ requires *m/z* 335.0621, found 335.0614.

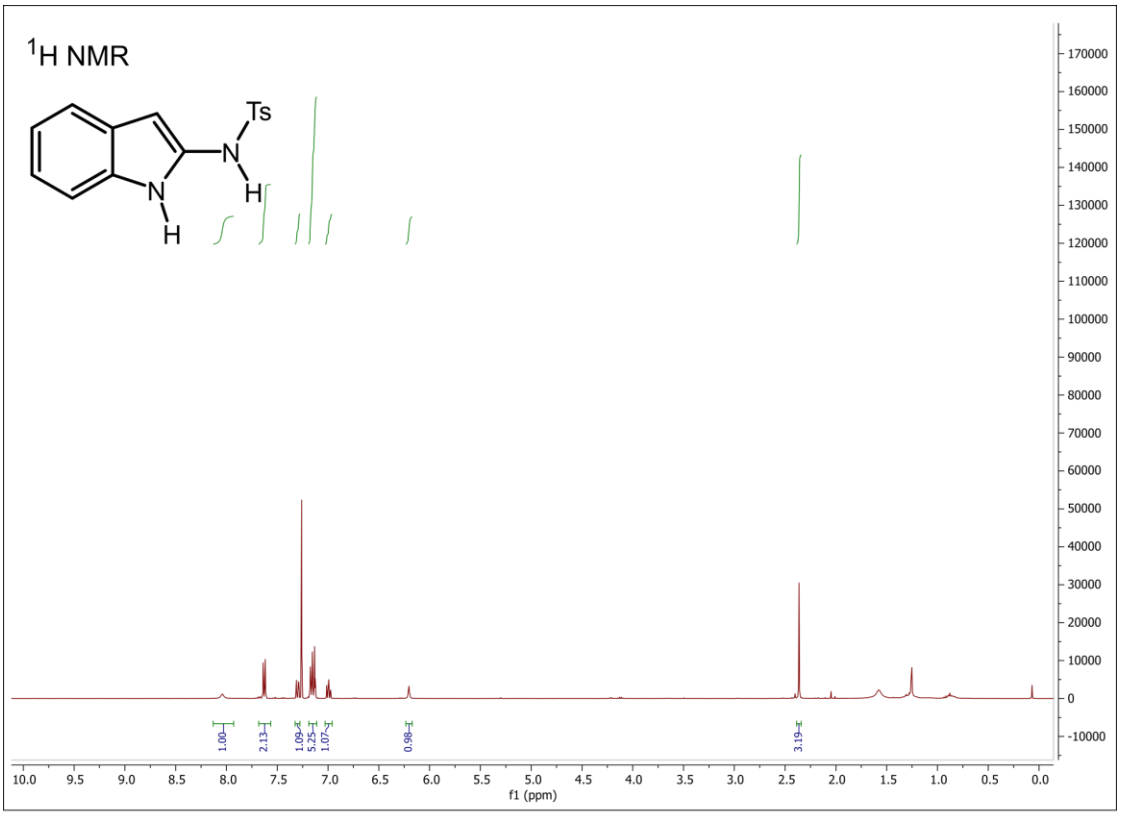
V. NMR Spectra

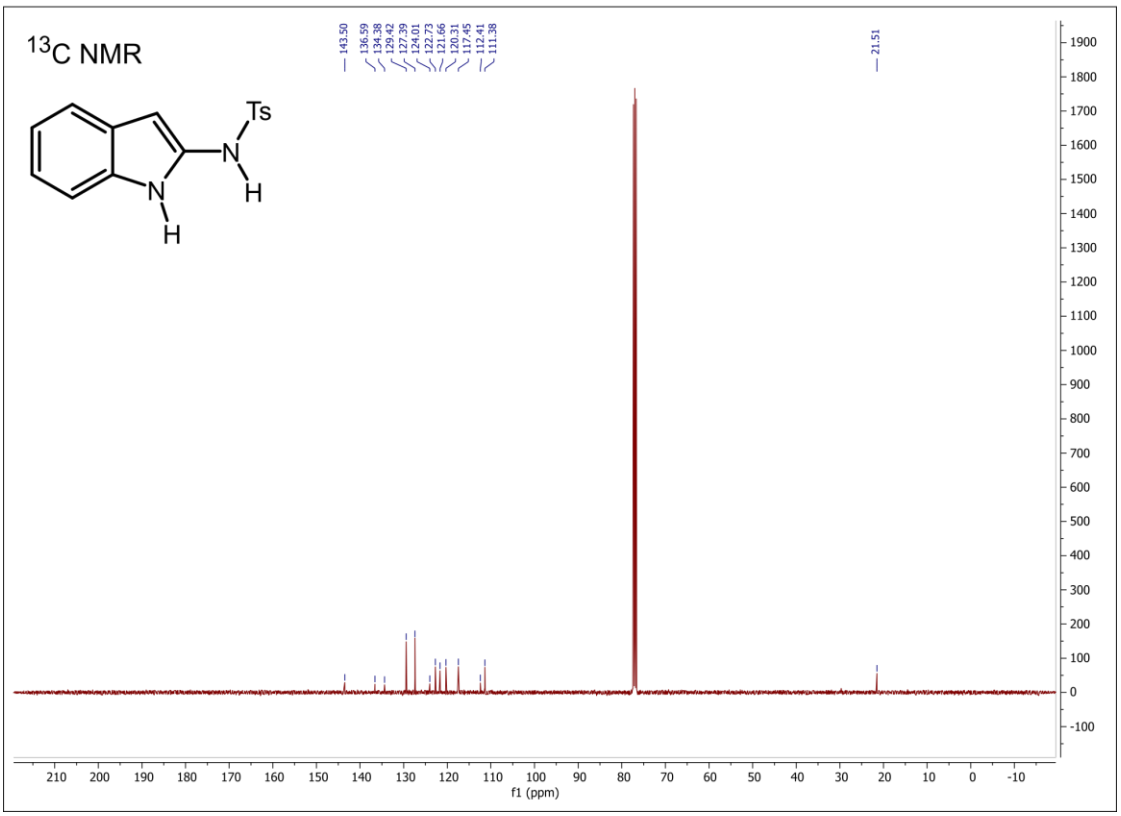


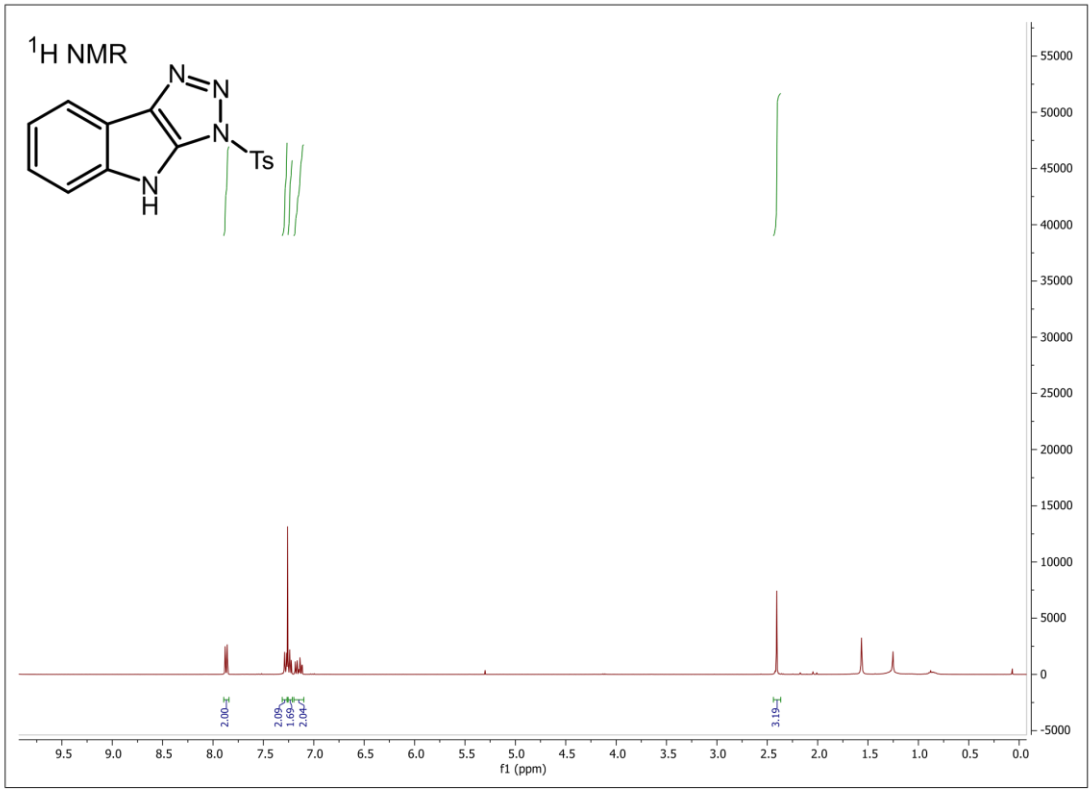


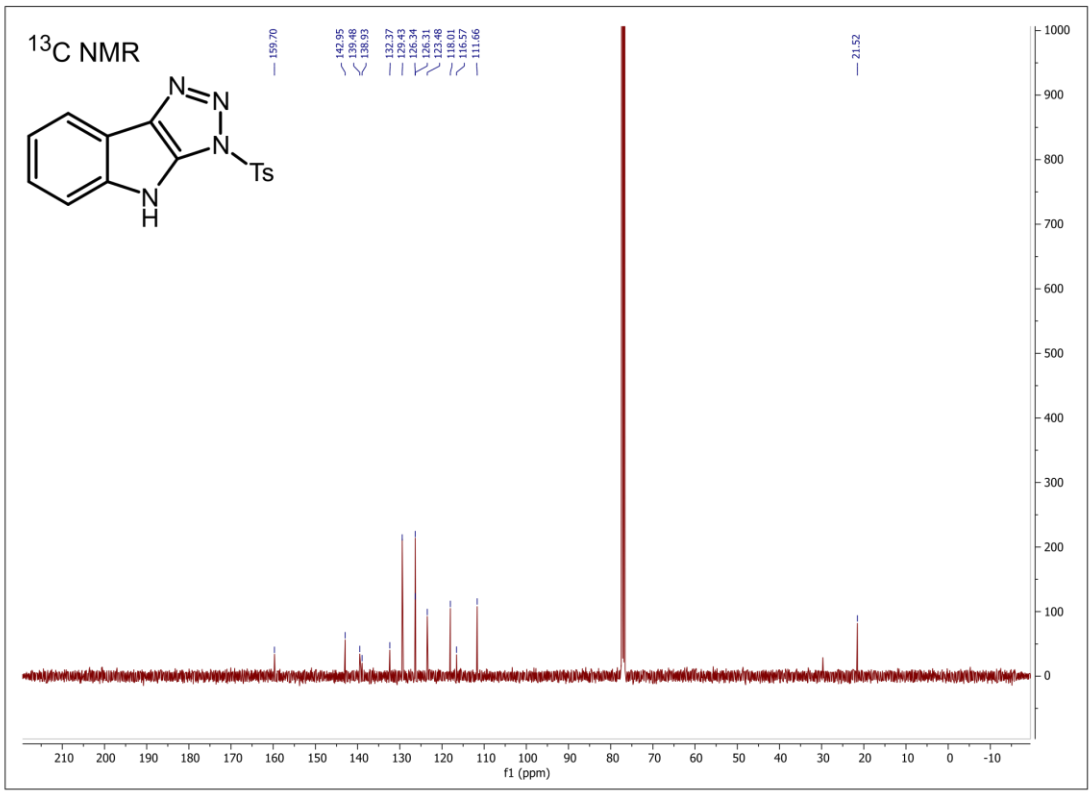

^{1}H NMR
N
N
N
Ts
N
Me
1.95
3.09
2.02
1.00
3.05
3.08
f1 (ppm)

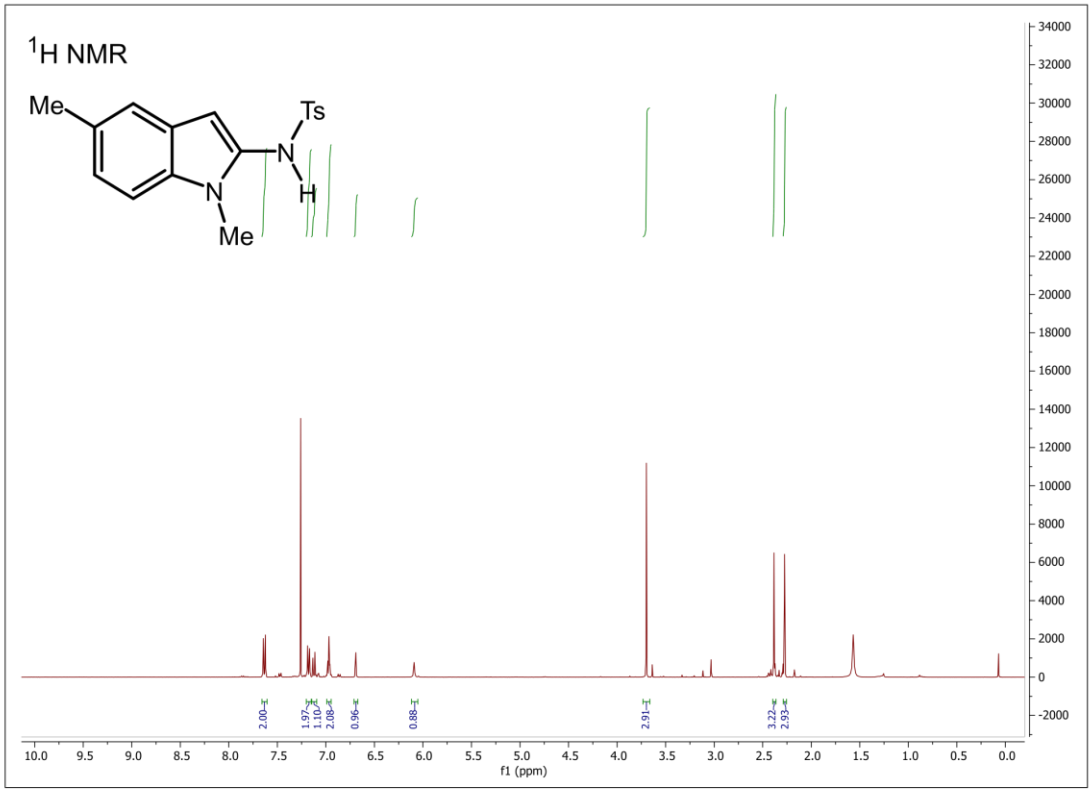

^{13}C NMR
N
N
N
Ts
N
Me
155.32
142.45
140.12
134.26
129.35
126.33
125.89
123.02
118.97
116.88
109.70
29.06
21.51
f1 (ppm)

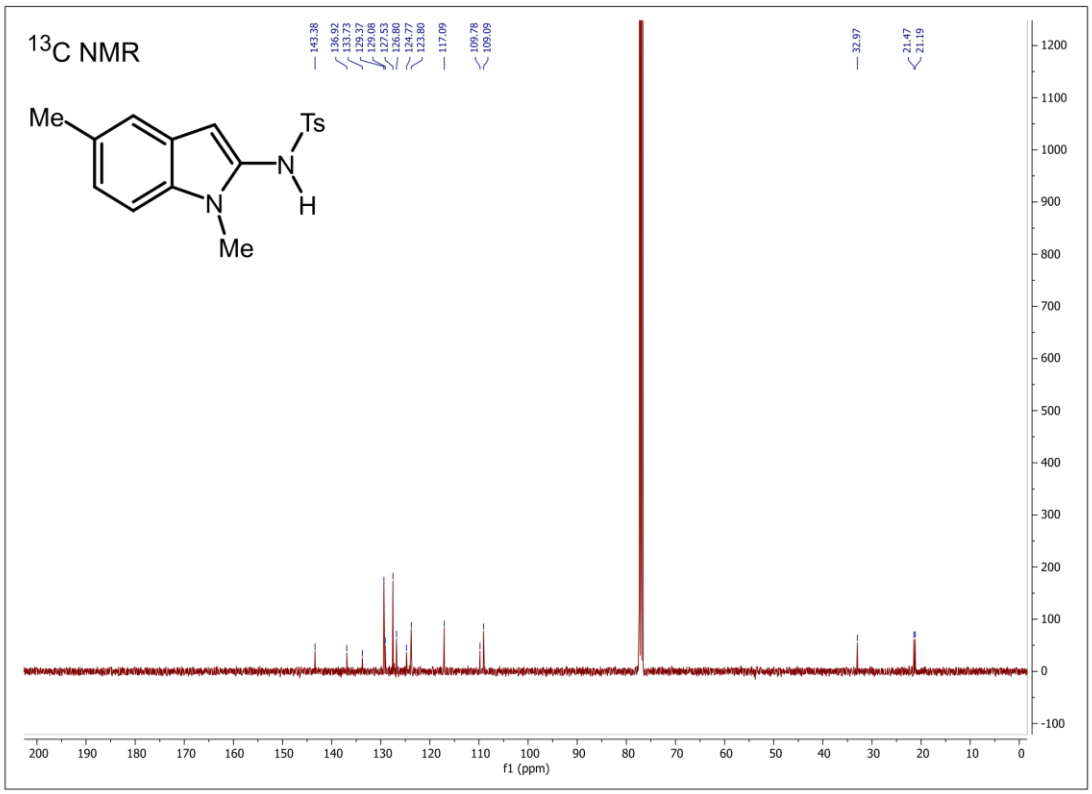

1H NMR
Ts
N
H
N
H
1.00
2.13
1.09
5.25
1.07
0.98
3.19
f1 (ppm)

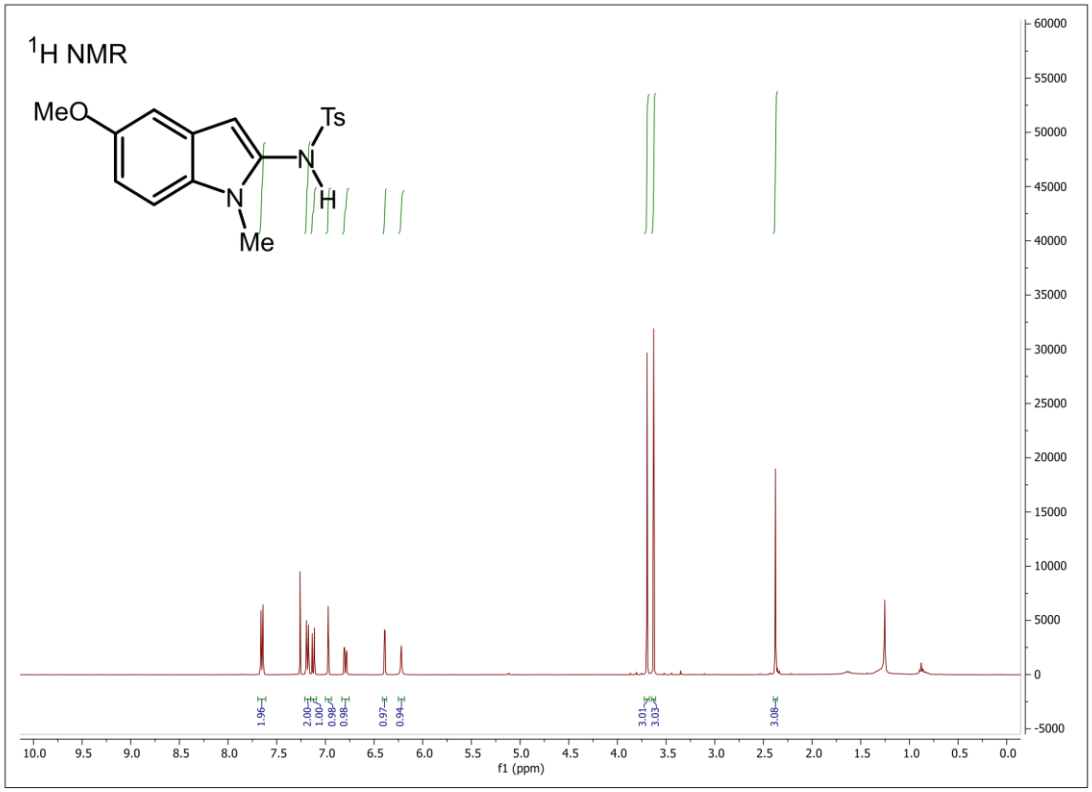

13C NMR
143.50
136.59
134.38
129.42
127.39
124.01
122.73
121.66
120.31
117.45
112.41
111.38
21.51
Ts
N
H
N
H
f1 (ppm)

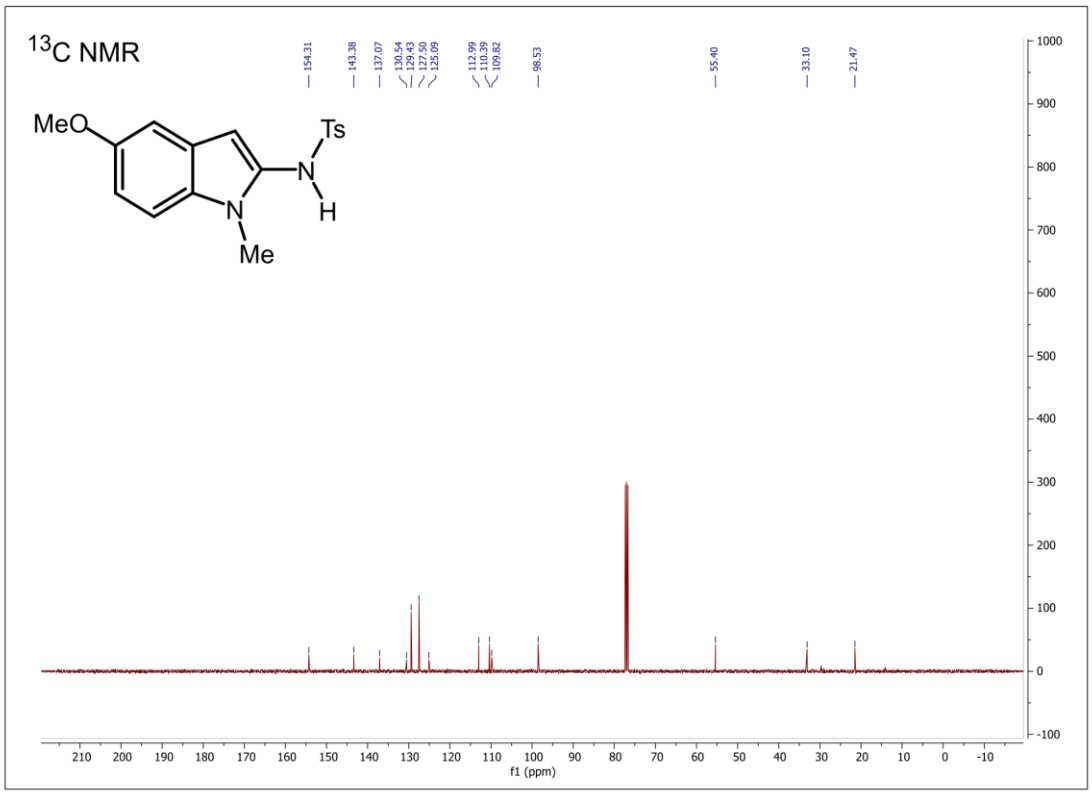

^{1}H NMR
Ts
2.00
2.09
1.68
2.04
3.19
f1 (ppm)

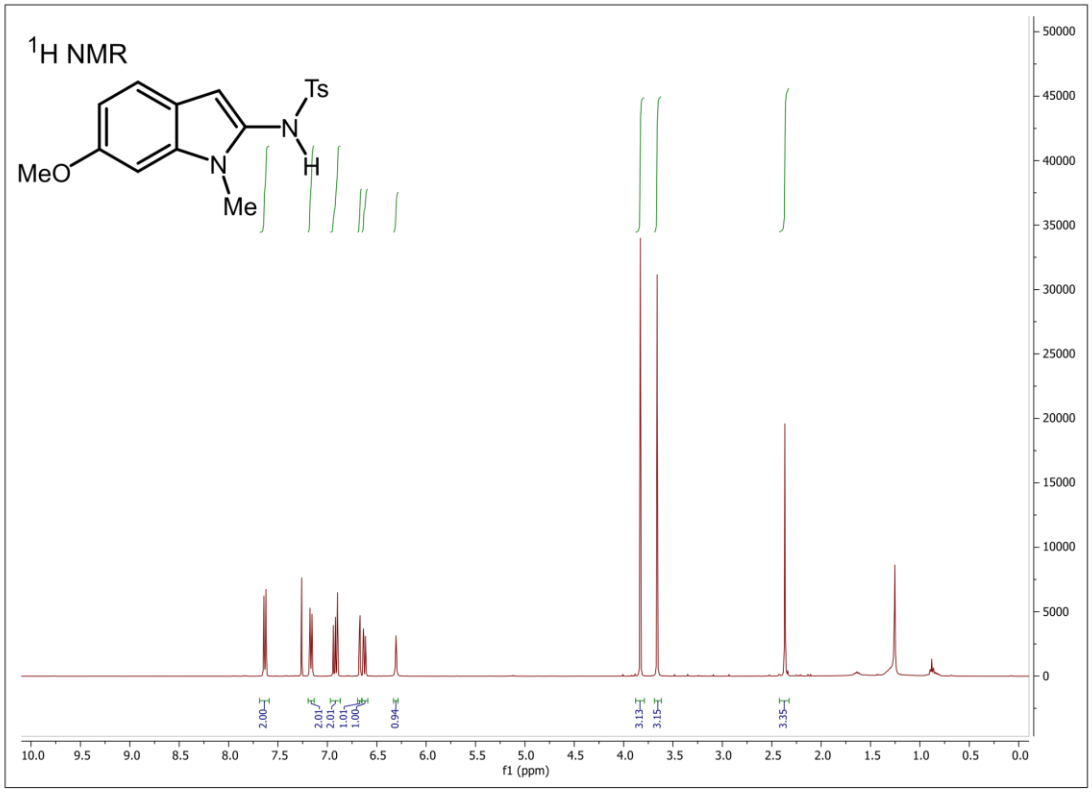

^{13}C NMR
Ts
159.70
142.95
139.48
138.93
132.37
129.43
126.34
126.31
123.48
118.01
116.57
111.66
21.52
f1 (ppm)

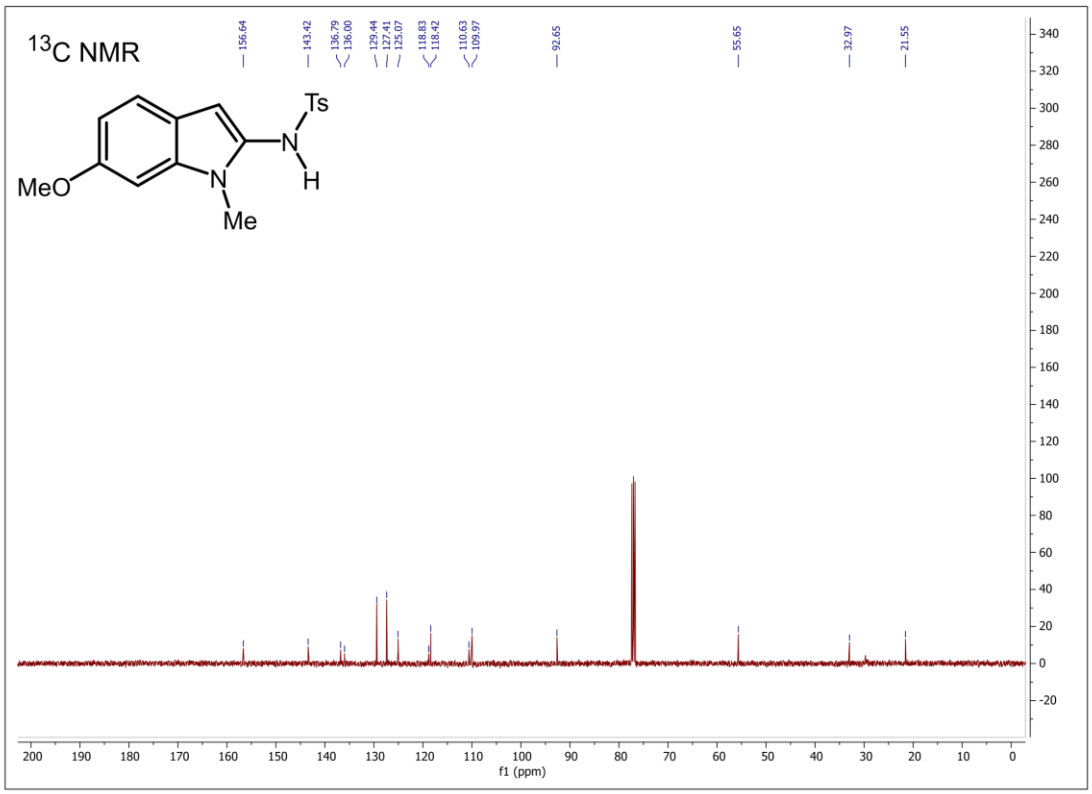

^{1}H NMR
Me
Ts
N
H
N
Me
2.00
1.97
1.10
2.08
0.96
0.88
2.91
3.22
2.93
f1 (ppm)

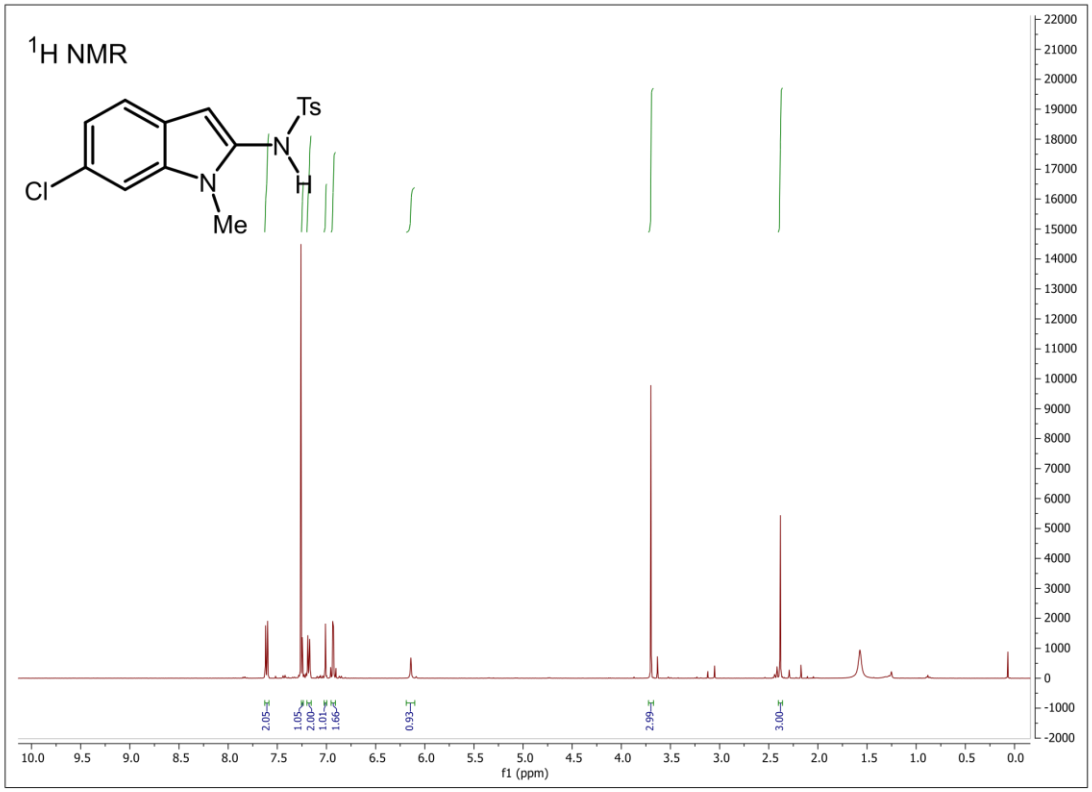

^{13}C NMR
Me
Ts
N
H
N
Me
143.38
136.92
133.73
129.37
129.08
127.53
126.80
124.77
123.80
117.09
109.78
109.09
32.97
21.47
21.19
f1 (ppm)

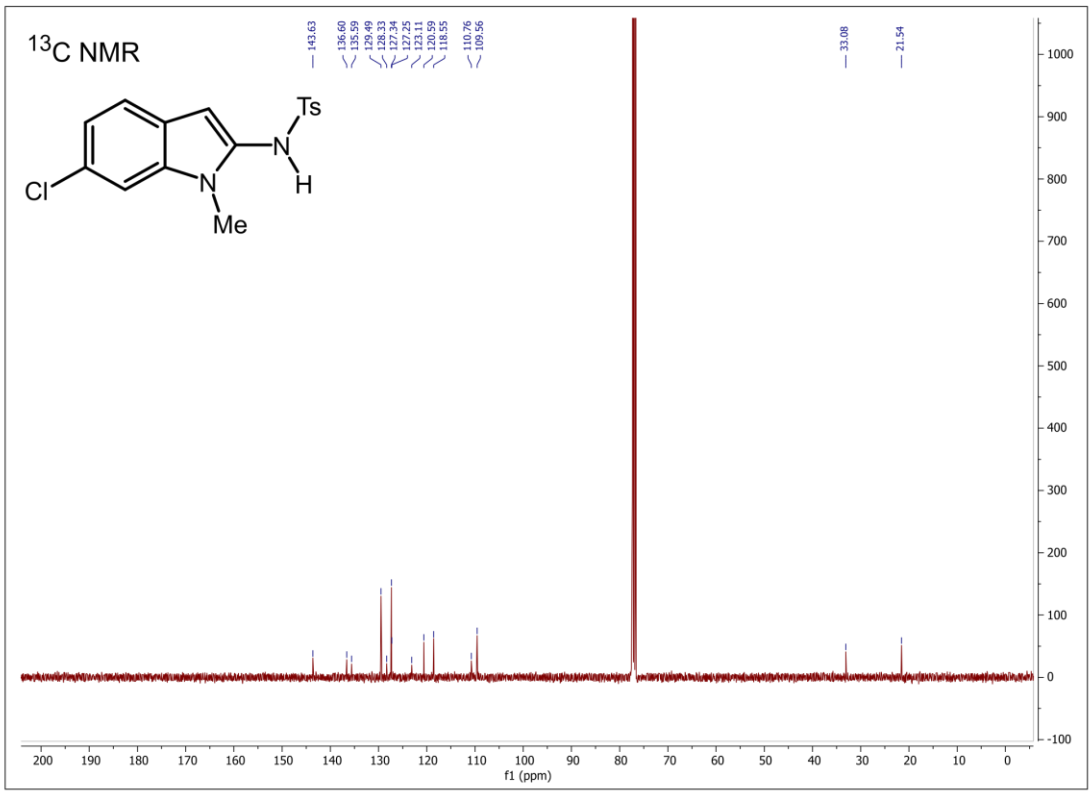

1H NMR
MeO
Ts
N
H
N
Me
1.96
2.00
1.07
1.08
0.98
0.97
0.94
3.01
3.03
3.08
f1 (ppm)


13C NMR
154.31
143.38
137.07
130.54
129.43
127.50
125.09
112.99
110.39
109.82
98.53
55.40
33.10
21.47
MeO
Ts
N
H
N
Me
f1 (ppm)


^{1}H NMR
Ts
N
H
MeO
N
Me
2.00
2.01
2.01
1.01
1.00
0.94
3.13
3.15
3.35
f1 (ppm)


^{13}C NMR
156.64
143.42
136.79
136.00
129.44
127.41
125.07
118.83
118.42
110.63
109.97
92.65
55.65
32.97
21.55
Ts
N
H
MeO
N
Me
f1 (ppm)


^{1}H NMR
Ts
N
H
Cl
N
Me
2.05
1.05
2.00
1.01
1.66
0.93
2.99
3.00
f1 (ppm)


^{13}C NMR
143.63
136.60
135.59
129.49
128.33
127.34
127.25
123.11
120.59
118.55
110.76
109.56
33.08
21.54
Ts
N
H
Cl
N
Me
f1 (ppm)

VI. Nucleotide and amino acid sequence of P411-IA constructs

P411-IA DNA sequence

ATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTAAAAATTTACC
GTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGGATGAATTAGG
AGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTATCAAGTCAGCG
TCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAACTTAAGTCAAGCGCT
GAAATTTGCACGTGATTTTTGGGGAGACGGGTTAGTCACAAGCTGGACGCATGAAA
AAAATTGGAAAAAAGCGCATAATATCTTACTTCCAAGCTTTAGTCAGCAGGCAATG
AAAGGCTATCATGCGATGATGGTCGATATCGCCGTGCAGCTTGTTCAAAAGTGGGA
GCGTCTAAATGCAGATGAGCATATTGAAGTATCGGAAGACATGACACGTTTAACGC
TTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTTTTACCGAGATCA
GCCTCATCCATTTATTATAAGTATGGTCCGTGCACTGGATGAAGTAATGAACAAGCT
GCAGCGAGCAAATCCAGACGACCCAGCTTATGATGAAAACAAGCGCCAGTTTCAAG
AAGATATCAAGGTGATGAACGACCTAGTAGATAAAATTATTGCAGATCGCAAAGCA
AGGGGTGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAAAGATCCAGA
AACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTACATTCTTATA
TGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTTCTTAGTGAA
AAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTCTAGTAGATC
CTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGGTCTTAAACG
AAGCGCTGCGCTTATGGCCAACTGTGCCTGCGTTTTCCCTATATGCAAAAGAAGATA
CGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATGGTTCTGATT
CCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAGTTCCGTCCA
GAGCGTTTTGAAAATCCAAGTGCGATTCCGCAGCATGCGTTTAAACCGTTTGGAAAC
GGTCAGCGTGCGTCTATCGGTCAGCAGTTCGCTCTTCATGAAGCAACGCTGGTACTT
GGTATGATGCTAAAACACTTTGACTTTGAAGATCATACAAACTACGAGCTCGATATT
AAAGAAACTTTAAGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAATCGAAAAA
AATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAAAAAGTACG
CAAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGGTTCAAATAT
GGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGAGCAAAGGAT
TTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGCGCGAAGGA
GCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACGCAAAGCAA
TTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTCGCTACTCC
GTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGCCTGCTTTT
ATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGACCGCGGTGAAGC
AGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGGCGTGAACATATGTGGA
GTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAGATAATAAATCTA
CTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAAAATGCACG
GTGCGTTTTCAACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAGGCAGTGCA
CGAAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATCAAGAAGG
AGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGTGTAACAGC
AAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGAAGAAAAAT
TAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCAATACGTGG

AGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGCAATGGCTGCTAAAACGGTCT
GCCCGCCGCATAAAGTAGAGCTTGAAGCCTTGCTTGAAAAGCAAGCCTACAAAGAA
CAAGTGCTGGCAAAACGTTTAACAATGCTTGAACTGCTTGAAAAATACCCGGCGTGT
GAAATGAAATTCAGCGAATTTATCGCCCTTCTGCCAAGCATACGCCCGCGCTATTAC
TCGATTTCTTCATCACCTCGTGTCGATGAAAAACAAGCAAGCATCACGGTCAGCGTT
GTCTCAGGAGAAGCGTGGAGCGGATATGGAGAATATAAAGGAATTGCGTCGAACTA
TCTTGCCGAGCTGCAAGAAGGAGATACGATTACGTGCTTTATTTCCACACCGCAGTC
AGAATTTACGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTCGGACCGGGAA
CAGGCGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTAAAAGAACAA
GGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCTCATGAAGAC
TATCTGTATCAAGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATCATTACGCTTCAT
ACCGCTTTTTCTCGCATGCCAAATCAGCCGAAAACATACGTTCAGCACGTAATGGAA
CAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGAGCGCACTTCTATATTTGC
GGAGACGGAAGCCAAATGGCACCTGCCGTTGAAGCAACGCTTATGAAAAGCTATGC
TGACGTTCACCAAGTGAGTGAAGCAGACGCTCGCTTATGGCTGCAGCAGCTAGAAG
AAAAAGGCCGATACGCAAAAGACGTGCCGGCTGGGCTCGAGCACCACCACCACCAC
CACTGA

P411-IA amino acid sequence

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIK
EACDESRFDKNLSQALKFARDFWGDGLVTSWTHEKNWKKAHNILLPSFSQQAMKGYH
AMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFII
SMVRALDEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKIIADRKARGEQSDD
LLTQMLNGKDPETGEPLDDGNIRYQIITFLYAGHEGTSGLLSFALYFLVKNPHVLQKVAE
EAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTVPAFSLYAKEDTVLGGEYPLEK
GDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFGNGQRASIGQQFALH
EATLVLGMMLKHFDFEDHTNYELDIKETLSLKPKGFVVKAKSKKIPLGGIPSPSTEQSAK
KVRKKAENAHNTPLLVLYGSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPR
EGAVLIVTASYNGHPPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVP
AFIDETLAAKGAENIADRGEADASDDFEGTYEEWREHMWSDVAAYFNLDIENSEDNKS
TLSLQFVDSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDH
LGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEKLAHLPLAKTVSVEELLQYVELQDPV
TRTQLRAMAAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKFSEF
IALLPSIRPRYYSISSSPRVDEKQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITC
FISTPQSEFTLPKDPETPLIMVGPGTGVAPFRGFVQARKQLKEQGQSLGEAHLYFGCRSP
HEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTYVQHVMEQDGKKLIELLDQGAHFYI
CGDGSQMAPAVEATLMKSYADVHQVSEADARLWLQQLEEKGRYAKDVPAGLEHHHH
HH*

DNA sequence of P411-IA in expression plasmid pET22b+

The P411-IA open reading frame is capitalized (position 5205-8378)

tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccc
tagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttcc
gatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccc
tttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggat
tttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaatttcaggtggca
cttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgctt
caataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcaccca
gaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttg
agagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaaga
gcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaag
agaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgctt
ttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccac
gatgcctgcagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatgga
ggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctc
gcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacg
aaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaactt
cattttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccc
cgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggttt
gtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgta
gttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtc
gtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagctt
ggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacag
gtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtt
tcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggtt
cctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccg
ctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggtattttctccttacgcatctgt
gcggtatttcacaccgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgtga
ctgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagc
tgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggcagctgcggtaaagctcatcagcgtgg
tcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagtttctccagaagcgttaatgtctggcttctgataaagcgggc
catgttaagggcggttttttcctgtttggtcactgatgcctccgtgtaagggggatttctgttcatgggggtaatgataccgatgaaacgagaga
ggatgctcacgatacgggttactgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcggtatggatgcggcggga
ccagagaaaaatcactcagggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgcag
atccggaacataatggtgcagggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaagaccattcatgttgttgctcaggtc
gcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctgctaaccagtaaggcaaccccgccagcctagccg
ggtcctcaacgacaggagcacgatcatgcgcacccgtggggccgccatgccggcgataatggcctgcttctcgccgaaacgtttggtggc
gggaccagtgacgaaggcttgagcgagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaag
cggtcctcgccgaaaatgacccagagcgctgccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgata
gtcatgccccgcgcccaccggaaggagctgactgggttgaaggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaa
cttacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggag

aggcggtttgcgtattgggcgccagggtggtttttcttttcaccagtgagacgggcaacagctgattgcccttcaccgcctggccctgagaga
gttgcagcaagcggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcg
gtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgt
tggcaaccagcatcgcagtgggaacgatgccctcattcagcatttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgtt
ccgctatcggctgaatttgattgcgagtgagatatttatgccagccagccagacgcagacgcgccgagacagaacttaatgggcccgctaa
cagcgcgatttgctggtgacccaatgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtg
tctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcatccagcggatagttaatg
atcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgct
ggcacccagttgatcggcgcgagatttaatcgccgcgacaatttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatca
gcaacgactgtttgcccgccagttgttgtgccacgcggttgggaatgtaattcagctccgccatcgccgcttccacttttcccgcgttttcgca
gaaacgtggctggcctggttcaccacgcgggaaacggtctgataagagacaccggcatactctgcgacatcgtataacgttactggtttcac
attcaccaccctgaattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctc
tcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaaggaatggtgcatgcaagga
gatggcgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccgaagtggcgagcccg
atcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagag
gatcgagatctcgatcccgcgaaattaatacgactcactataggggaattgtgagcggataacaattcccctctagaaataattttgtttaacttt
aagaaggagatatacatATGACAATTAAAGAAATGCCTCAGCCAAAAACGTTTGGAGAGCTTA
AAAATTTACCGTTATTAAACACAGATAAACCGGTTCAAGCTTTGATGAAAATTGCGG
ATGAATTAGGAGAAATCTTTAAATTCGAGGCGCCTGGTCGTGTAACGCGCTACTTAT
CAAGTCAGCGTCTAATTAAAGAAGCATGCGATGAATCACGCTTTGATAAAAACTTA
AGTCAAGCGCTGAAATTTGCACGTGATTTTTGGGGAGACGGGTTAGTCACAAGCTGG
ACGCATGAAAAAAATTGGAAAAAAGCGCATAATATCTTACTTCCAAGCTTTAGTCA
GCAGGCAATGAAAGGCTATCATGCGATGATGGTCGATATCGCCGTGCAGCTTGTTCA
AAAGTGGGAGCGTCTAAATGCAGATGAGCATATTGAAGTATCGGAAGACATGACAC
GTTTAACGCTTGATACAATTGGTCTTTGCGGCTTTAACTATCGCTTTAACAGCTTTTA
CCGAGATCAGCCTCATCCATTTATTATAAGTATGGTCCGTGCACTGGATGAAGTAAT
GAACAAGCTGCAGCGAGCAAATCCAGACGACCCAGCTTATGATGAAAACAAGCGCC
AGTTTCAAGAAGATATCAAGGTGATGAACGACCTAGTAGATAAAATTATTGCAGAT
CGCAAAGCAAGGGGTGAACAAAGCGATGATTTATTAACGCAGATGCTAAACGGAAA
AGATCCAGAAACGGGTGAGCCGCTTGATGACGGGAACATTCGCTATCAAATTATTA
CATTCTTATATGCGGGACACGAAGGTACAAGTGGTCTTTTATCATTTGCGCTGTATTT
CTTAGTGAAAAATCCACATGTATTACAAAAAGTAGCAGAAGAAGCAGCACGAGTTC
TAGTAGATCCTGTTCCAAGCTACAAACAAGTCAAACAGCTTAAATATGTCGGCATGG
TCTTAAACGAAGCGCTGCGCTTATGGCCAACTGTGCCTGCGTTTTCCCTATATGCAA
AAGAAGATACGGTGCTTGGAGGAGAATATCCTTTAGAAAAAGGCGACGAAGTAATG
GTTCTGATTCCTCAGCTTCACCGTGATAAAACAGTTTGGGGAGACGATGTGGAGGAG
TTCCGTCCAGAGCGTTTTGAAAATCCAAGTGCGATTCCGCAGCATGCGTTTAAACCG
TTTGGAAACGGTCAGCGTGCGTCTATCGGTCAGCAGTTCGCTCTTCATGAAGCAACG
CTGGTACTTGGTATGATGCTAAAACACTTTGACTTTGAAGATCATACAAACTACGAG
CTCGATATTAAAGAAACTTTAAGTTTAAAACCTAAAGGCTTTGTGGTAAAAGCAAAA
TCGAAAAAAATTCCGCTTGGCGGTATTCCTTCACCTAGCACTGAACAGTCTGCTAAA
AAAGTACGCAAAAAGGCAGAAAACGCTCATAATACGCCGCTGCTTGTGCTATACGG
TTCAAATATGGGTACCGCTGAAGGAACGGCGCGTGATTTAGCAGATATTGCAATGA

GCAAAGGATTTGCACCGCAGGTCGCAACGCTTGATTCACACGCCGGAAATCTTCCGC
GCGAAGGAGCTGTATTAATTGTAACGGCGTCTTATAACGGTCATCCGCCTGATAACG
CAAAGCAATTTGTCGACTGGTTAGACCAAGCGTCTGCTGATGAAGTAAAAGGCGTTC
GCTACTCCGTATTTGGATGCGGCGATAAAAACTGGGCTACTACGTATCAAAAAGTGC
CTGCTTTTATCGATGAAACGCTTGCCGCTAAAGGGGCAGAAAACATCGCTGACCGCG
GTGAAGCAGATGCAAGCGACGACTTTGAAGGCACATATGAAGAATGGCGTGAACAT
ATGTGGAGTGACGTAGCAGCCTACTTTAACCTCGACATTGAAAACAGTGAAGATAA
TAAATCTACTCTTTCACTTCAATTTGTCGACAGCGCCGCGGATATGCCGCTTGCGAA
AATGCACGGTGCGTTTTCAACGAACGTCGTAGCAAGCAAAGAACTTCAACAGCCAG
GCAGTGCACGAAGCACGCGACATCTTGAAATTGAACTTCCAAAAGAAGCTTCTTATC
AAGAAGGAGATCATTTAGGTGTTATTCCTCGCAACTATGAAGGAATAGTAAACCGT
GTAACAGCAAGGTTCGGCCTAGATGCATCACAGCAAATCCGTCTGGAAGCAGAAGA
AGAAAAATTAGCTCATTTGCCACTCGCTAAAACAGTATCCGTAGAAGAGCTTCTGCA
ATACGTGGAGCTTCAAGATCCTGTTACGCGCACGCAGCTTCGCGCAATGGCTGCTAA
AACGGTCTGCCCGCCGCATAAAGTAGAGCTTGAAGCCTTGCTTGAAAAGCAAGCCT
ACAAAGAACAAGTGCTGGCAAAACGTTTAACAATGCTTGAACTGCTTGAAAAATAC
CCGGCGTGTGAAATGAAATTCAGCGAATTTATCGCCCTTCTGCCAAGCATACGCCCG
CGCTATTACTCGATTTCTTCATCACCTCGTGTCGATGAAAAACAAGCAAGCATCACG
GTCAGCGTTGTCTCAGGAGAAGCGTGGAGCGGATATGGAGAATATAAAGGAATTGC
GTCGAACTATCTTGCCGAGCTGCAAGAAGGAGATACGATTACGTGCTTTATTTCCAC
ACCGCAGTCAGAATTTACGCTGCCAAAAGACCCTGAAACGCCGCTTATCATGGTCG
GACCGGGAACAGGCGTCGCGCCGTTTAGAGGCTTTGTGCAGGCGCGCAAACAGCTA
AAAGAACAAGGACAGTCACTTGGAGAAGCACATTTATACTTCGGCTGCCGTTCACCT
CATGAAGACTATCTGTATCAAGAAGAGCTTGAAAACGCCCAAAGCGAAGGCATCAT
TACGCTTCATACCGCTTTTTCTCGCATGCCAAATCAGCCGAAAACATACGTTCAGCA
CGTAATGGAACAAGACGGCAAGAAATTGATTGAACTTCTTGATCAAGGAGCGCACT
TCTATATTTGCGGAGACGGAAGCCAAATGGCACCTGCCGTTGAAGCAACGCTTATGA
AAAGCTATGCTGACGTTCACCAAGTGAGTGAAGCAGACGCTCGCTTATGGCTGCAG
CAGCTAGAAGAAAAAGGCCGATACGCAAAAGACGTGCCGGCTGGGCTCGAGCACC
ACCACCACCACCACTGAgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataa
ctagcataaccccttggggcctctaaacgggtcttgaggggtttttttgctgaaaggaggaactatatccggat

VII. Supplementary References

(1) Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92. https://doi.org/10.1021/sb300037w.

(2) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison Iii, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6*, 343–345. https://doi.org/10.1038/nmeth.1318.

(3) Guengerich, F. P.; Martin, M. V.; Sohl, C. D.; Cheng, Q. Measurement of Cytochrome P450 and NADPH–Cytochrome P450 Reductase. *Nat. Protoc.* **2009**, *4*, 1245–1251. https://doi.org/10.1038/nprot.2009.121.

(4) Vatsis, K. P.; Peng, H.-M.; Coon, M. J. Replacement of Active-Site Cysteine-436 by Serine Converts Cytochrome P450 2B4 into an NADPH Oxidase with Negligible Monooxygenase Activity. *J. Inorg. Biochem.* **2002**, *91*, 542–553. https://doi.org/10.1016/S0162-0134(02)00438-5.

(5) Fuhrhop, J.-H.; Smith, K. M. *Laboratory Methods in Porphyrin and Metalloporphyrin Research*; Elsevier Scientific Pub. Co.: Amsterdam; New York, 1975.

(6) Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M. A Serine-Substituted P450 Catalyzes Highly Efficient Carbene Transfer to Olefins *in Vivo*. *Nat. Chem. Biol.* **2013**, *9*, 485–487. https://doi.org/10.1038/nchembio.1278.

(7) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin Cyclopropanation via Carbene Transfer Catalyzed by Engineered Cytochrome P450 Enzymes. *Science* **2013**, *339*, 307–310. https://doi.org/10.1126/science.1231434.

(8) Wang, Z. J.; Renata, H.; Peck, N. E.; Farwell, C. C.; Coelho, P. S.; Arnold, F. H. Improved Cyclopropanation Activity of Histidine-Ligated Cytochrome P450 Enables the Enantioselective Formal Synthesis of Levomilnacipran. *Angew. Chem. 126*, 6928–6931. https://doi.org/10.1002/ange.201402809.

(9) Renata, H.; Lewis, R. D.; Sweredoski, M. J.; Moradian, A.; Hess, S.; Wang, Z. J.; Arnold, F. H. Identification of Mechanism-Based Inactivation in P450-Catalyzed Cyclopropanation Facilitates Engineering of Improved Enzymes. *J. Am. Chem. Soc.* **2016**, *138*, 12527–12533. https://doi.org/10.1021/jacs.6b06823.

(10) Rabe, K. S.; Kiko, K.; Niemeyer, C. M. Characterization of the Peroxidase Activity of CYP119, a Thermostable P450 from Sulfolobus Acidocaldarius. *Chembiochem* **2008**, *9*, 420–425. https://doi.org/10.1002/cbic.200700450.

(11) Poulos, T. L.; Finzel, B. C.; Howard, A. J. High-Resolution Crystal Structure of Cytochrome P450cam. *J. Mol. Biol.* **1987**, *195*, 687–700.

(12) Joyce, M. G.; Ekanem, I. S.; Roitel, O.; Dunford, A. J.; Neeli, R.; Girvan, H. M.; Baker, G. J.; Curtis, R. A.; Munro, A. W.; Leys, D. The Crystal Structure of the FAD/NADPH-Binding Domain of Flavocytochrome P450 BM3. *FEBS J.* **2012**, *279*, 1694–1706. https://doi.org/10.1111/j.1742-4658.2012.08544.x.

(13) Prier, C. K.; Zhang, R. K.; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. Enantioselective, Intermolecular Benzylic C–H Amination Catalysed by an Engineered Iron-Haem Enzyme. *Nat. Chem.* **2017**, *9*, 629–634. https://doi.org/10.1038/nchem.2783.

(14) Xing, Y.; Sheng, G.; Wang, J.; Lu, P.; Wang, Y. Preparation of Triazoloindoles via Tandem Copper Catalysis and Their Utility as α-Imino Rhodium Carbene Precursors. *Org. Lett.* **2014**, *16*, 1244–1247. https://doi.org/10.1021/ol5002347.