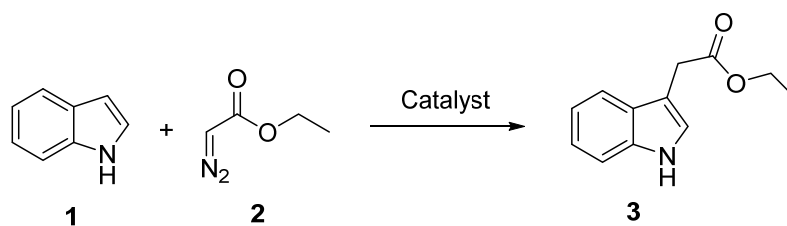


**Table of contents:**

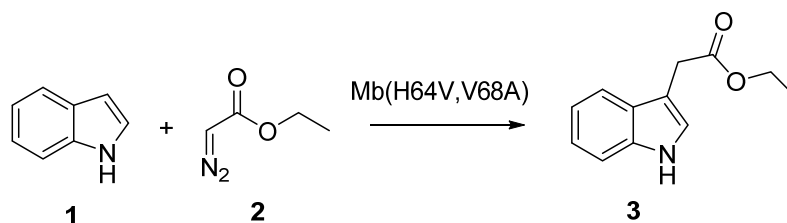
Supplementary Table S1-S3	Pages S2-S4
Supplementary Figure S1-S6	Pages S5-S11
Experimental Procedures	Page S12-S14
Synthetic Procedures	Pages S14-S23
References	Pages S24
NMR Spectra	Pages S25-S57

**Table S1.** Activity and selectivity of wild-type sperm whale myoglobin (Mb), engineered Mb variants, and other hemoproteins in C—H functionalization of indole with ethyl  $\alpha$ -diazooacetate. Reaction conditions: 20  $\mu$ M catalyst, 2.5 mM indole, 2.5 mM EDA, 10 mM sodium dithionite in 50 mM potassium phosphate (pH 7.0), 16 hours, room temperature. Product yield as determined by GC analysis using calibration curve with **3**. N.a. = not determinable due to low activity. Reported values are mean values from  $n \geq 2$  experiments (SE <15%).



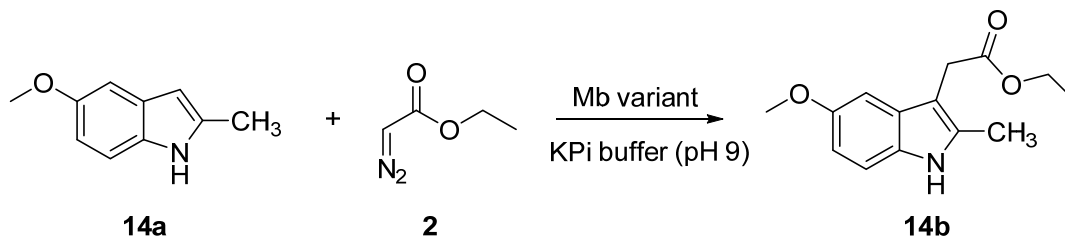
Catalyst	Yield (GC)	TON	% C-H functionalization
-	0%	N/A	-
Hemin	0.9%	1	n.a.
Mb (WT)	0.6%	<1	n.a.
P450 BM3	0.3%	<1	n.a.
Horseradish Peroxidase	0.2%	<1	n.a.
Catalase	0%	0	n.a.
Cytochrome <i>c</i>	0.2%	<1	n.a.
Mb(H64V)	0.2%	0	n.a.
Mb(L29A,H64V)	5%	6	100
Mb(L29F,H64V)	2%	3	100
Mb(F43I,H64V)	0%	0	-
Mb(F43S,H64V)	4%	5	100
Mb(H64V,V68A)	50%	62	100
Mb(H64V,V68F)	0%	0	-
Mb(H64V,I107S)	0%	0	-
Mb(H64V,I107Y)	0%	0	-

**Table S2.** Reaction optimization studies for Mb(H64V,V68A)-catalyzed C—H functionalization of indole with ethyl  $\alpha$ -diazoacetate. Reactions with purified protein also contain 10 mM sodium dithionite. Product yield as determined by GC analysis. Reported values are mean values from  $n \geq 2$  experiments (SE <15%).



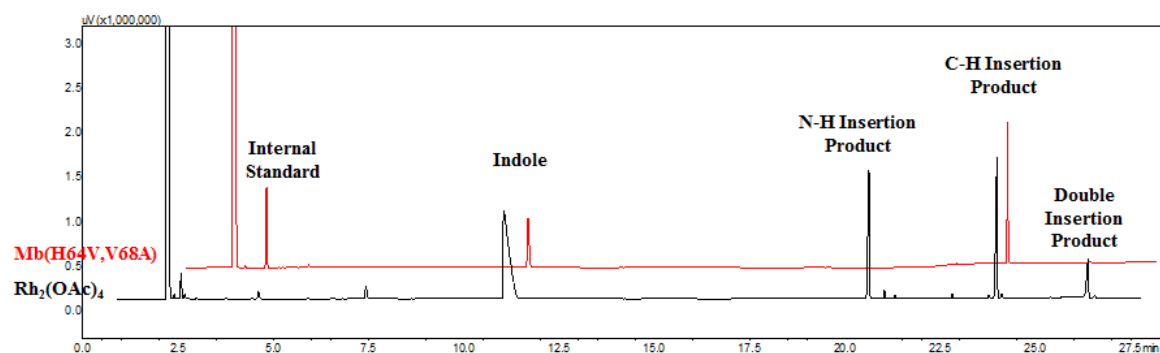
Catalyst	Catalyst loading	[indole] / mM	[EDA] / mM	pH	Yield (GC)	TON
Protein	0.8 mol%	5	5	6	45%	56
Protein	0.8 mol%	5	5	7	51%	64
Protein	0.8 mol%	5	5	8	54%	67
Protein	0.8 mol%	5	5	9	57%	71
Protein	0.8 mol%	2.5	2.5	9	54%	68
Protein	0.8 mol%	2.5	5	9	85%	106
Protein	0.4 mol%	2.5	5	9	51%	128
Protein	0.2 mol%	2.5	5	9	34%	168
Whole cells	OD <sub>600</sub> = 80	2.5	5	7	36%	5
Whole cells	OD <sub>600</sub> = 60	2.5	5	7	31%	5
Whole cells	OD <sub>600</sub> = 40	2.5	5	7	70%	18
Whole cells	OD <sub>600</sub> = 20	2.5	5	7	>99%	82

**Table S3.** Screening of Mb variants for C—H functionalization activity on 2-methyl-5-methoxy indole in the presence of ethyl  $\alpha$ -diazoacetate. Reaction conditions: 20  $\mu$ M protein, 2.5 mM **14a**, 5 mM EDA, 10 mM sodium dithionite in 50 mM potassium phosphate (pH 9.0), 16 hours, room temperature. Product yield as determined by GC analysis using calibration curve with **14b**. n.a. = not active. The triple mutant Mb(L29F,F43S,H64V) was constructed by combining the beneficial mutations at position 29 (Leu $\rightarrow$ Phe) and 43 (Phe $\rightarrow$ Ser) as derived from the screening of the double mutants. This variant shows however comparable activity to Mb(F43S,H64V), indicating a lack of additivity for these mutations. Reported values are mean values from  $n \geq 2$  experiments (SE <10%).

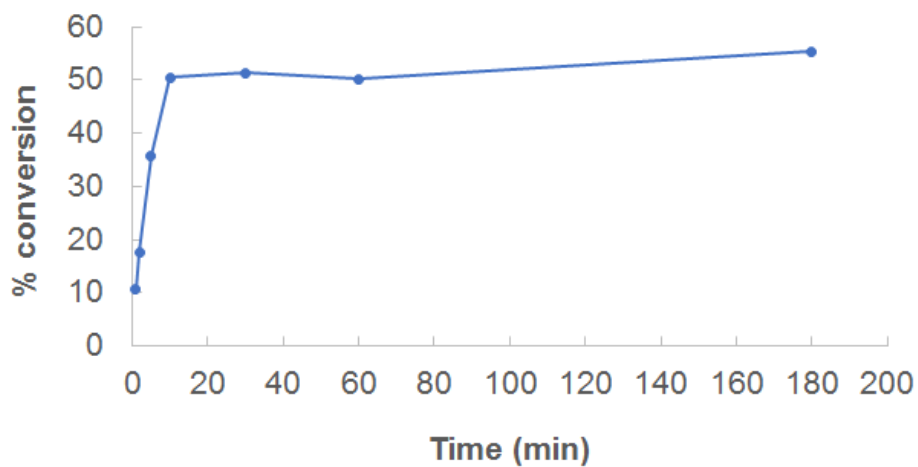


Catalyst	Yield (GC)	TON	% C-H Functionalization
Mb(H64V,V68A)	9%	11	100
Mb(L29A,H64V)	15%	19	100
Mb(L29F,H64V)	41%	51	100
Mb(F43I,H64V)	14%	17	100
Mb(F43S,H64V)	31%	39	100
Mb(H64V,V68F)	n.a.	0	n.a.
Mb(H64V,I107S)	14%	18	100
Mb(H64V,I107Y)	22%	18	100
Mb(L29F,F43S,H64V)	31%	38	100

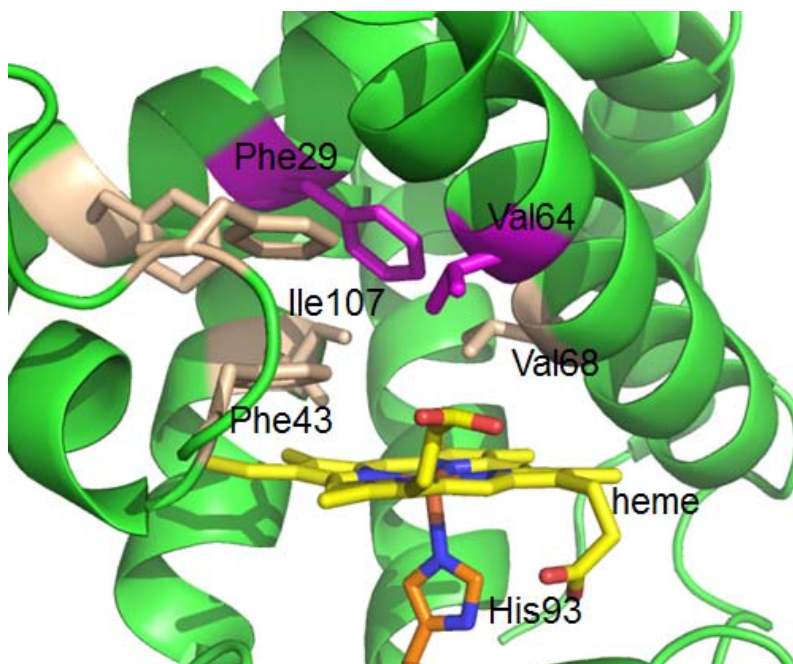
**Figure S1.** Overlay of GC chromatograms corresponding to the reactions with indole and EDA in the presence of either Mb(H64V,V68A) or Rh<sub>2</sub>(OAc)<sub>4</sub> as the catalyst. The Mb(H64V,V68A)-catalyzed reaction produces only the desired C—H insertion product (**3**), whereas the Rh-catalyzed reaction yield a mixture of **3**, the N—H insertion product, and the double insertion product in approximately 3:3:1 ratio based on the GC peak area.



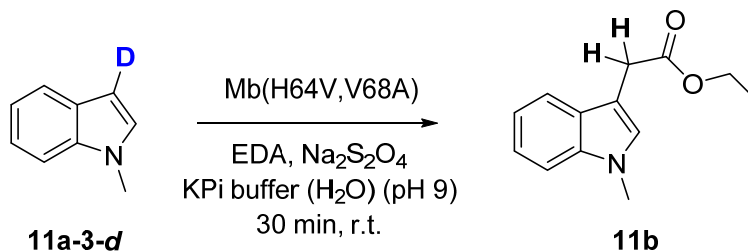
**Figure S2.** Time course experiment for Mb(H64V,V68A)-catalyzed conversion of indole (**1**) and EDA (**2**) to the C—H functionalized product **3** (see **Table 1** of main text). Reaction conditions: 2.5 mM indole, 5 mM EDA, 10 mM sodium dithionite, 20  $\mu$ M Mb variant (0.8 mol%) in 50 mM phosphate buffer (pH 9.0).



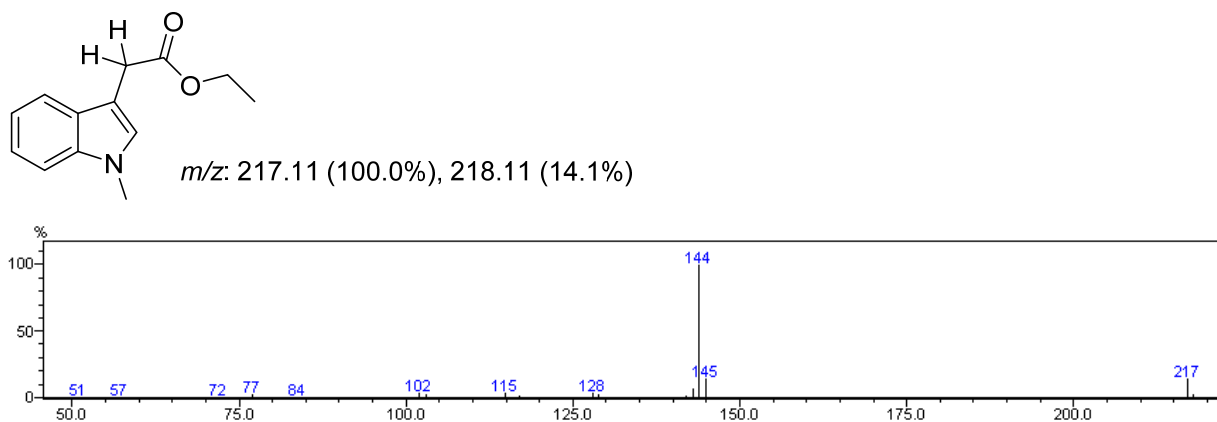
**Figure S3.** View of the distal heme pocket in the model of Mb(L29F,H64V) variant generated using the available crystal structure of a H64V-containing Mb variant (pdb 2MGJ)<sup>[1]</sup>. The heme bound proximal histidine (His93, orange) and the active site residues targeted for mutagenesis are labeled (stick models). The mutated active site residues in Mb(L29F,H64V) compared to wild-type Mb are highlighted in magenta.



**Figure S4. Reactions with 1-methyl-3-deutero-indole (11a-3-d).** Mb(H64V,V68A)-catalyzed conversion of **11a-3-d** with EDA leads to formation of **11b** (see Scheme below) with complete loss of the deuterium label as determined by GC-MS (panel A). Reaction conditions: 2.5 mM **11a-3-d**, 5 mM EDA, 10 mM sodium dithionite, 20  $\mu$ M Mb variant in 50 mM phosphate buffer (pH 9.0). Control experiments confirmed that neither the N-methyl indole substrate (panel B) nor the C—H functionalization product (panel C) undergo H/D exchange with the solvent under the applied reaction conditions (i.e., KPi buffer in H<sub>2</sub>O (pH: 9) or D<sub>2</sub>O (pD: 9), respectively, 30 minutes, room temperature), supporting the occurrence of solvent-mediated D/H exchange during the Mb(H64V,V68A)-catalyzed reaction as described in the main text.



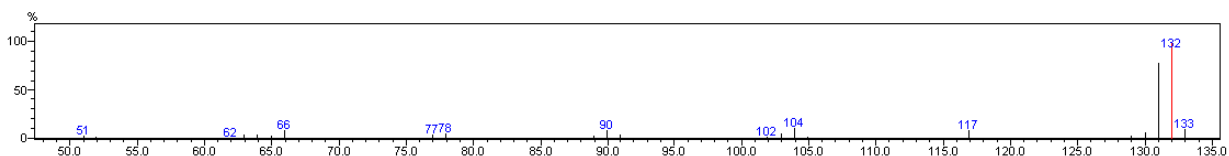
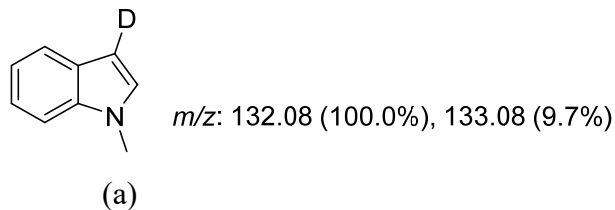
A) MS spectrum of the C-H functionalization product obtained from the Mb(H64V,V68A)-catalyzed reaction with 1-methyl-3-deutero-indole (**11a-3-d**) and EDA.



GC-MS *m/z* (% relative intensity): 218 (2.4), 217 (13.6), 145(13.8), 144(100), 143 (7.2)

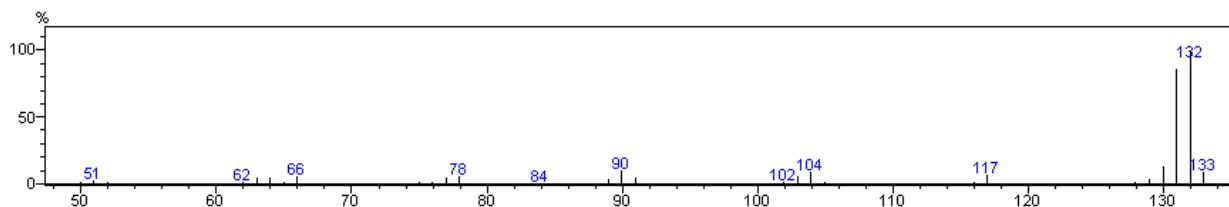


B) GC-MS spectrum of 1-methyl-3-deutero-indole (**11a-3-d**) prior (a) and after (b) incubation in H<sub>2</sub>O (KPi buffer, pH 9) for 30 minutes at room temperature.



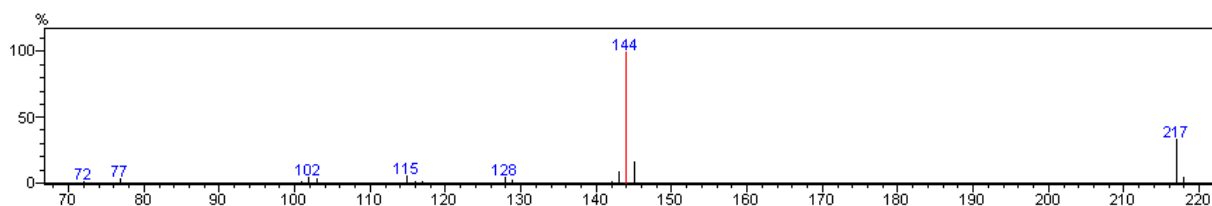
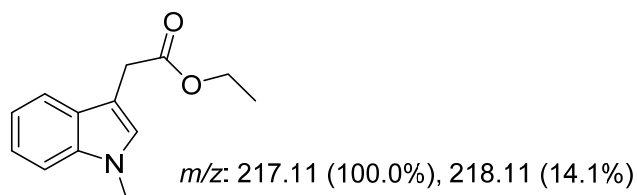
GC-MS *m/z* (% relative intensity): 133 (10.6), 132 (100), 131 (78.9), 130 (6.4), 117 (8.9).

(b)



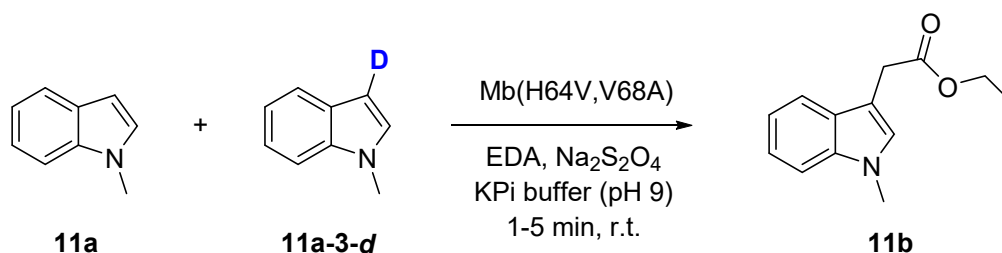
GC-MS *m/z* (% relative intensity): 133 (9.9), 132 (100), 131 (86.6), 130 (13.5), 117 (7.8)

c) GC-MS spectrum of authentic **11b** after (b) incubation in D<sub>2</sub>O (KPi buffer (D<sub>2</sub>O), pD 9) for 30 minutes at room temperature.

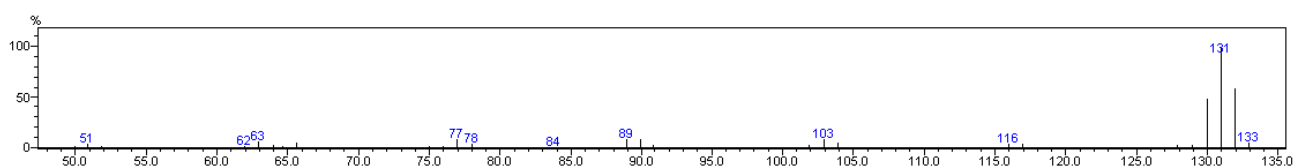


GC-MS *m/z* (% relative intensity): 218 (5.6), 217 (34.1), 145(16.5), 144(100), 143 (9.2)

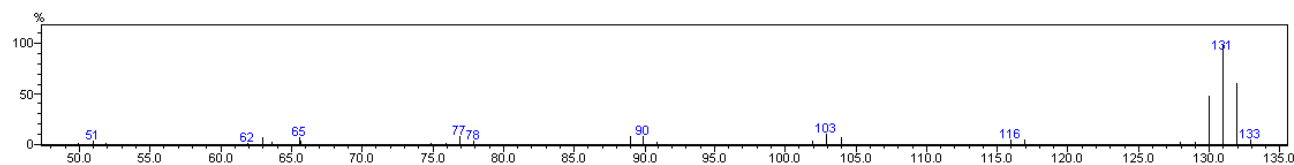
**Table S5.** Competition experiment with N-methylindole (**11a**) and 1-methyl-3-deutero-indole (**11a-3-d**). As shown by the GC-MS spectra and summary table reported below, the Mb(H64V,V68A)-catalyzed conversion of a ~1:1 mixture of N-methylindole and 1-methyl-3-deutero-indole showed no enrichment of the protiated substrate over time, indicating a lack of kinetic isotope effect. Reaction conditions: 2.5 mM mixture of **11a** + **11a-3-d** (~1:0.9 ratio), 5 mM EDA, 10 mM sodium dithionite, 20  $\mu$ M Mb variant in 50 mM phosphate buffer (pH 9.0).



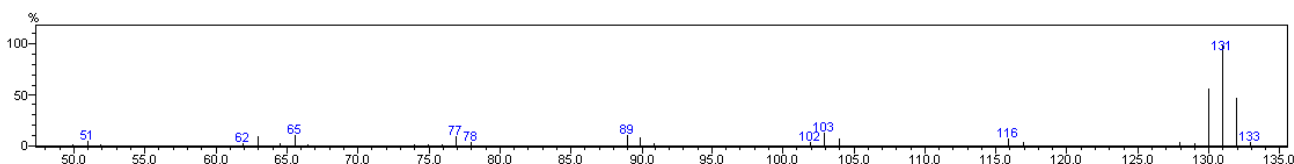
a) GC-MS spectrum of substrate peak prior to addition of the catalyst



b) GC-MS spectrum of substrate peak after 2 min

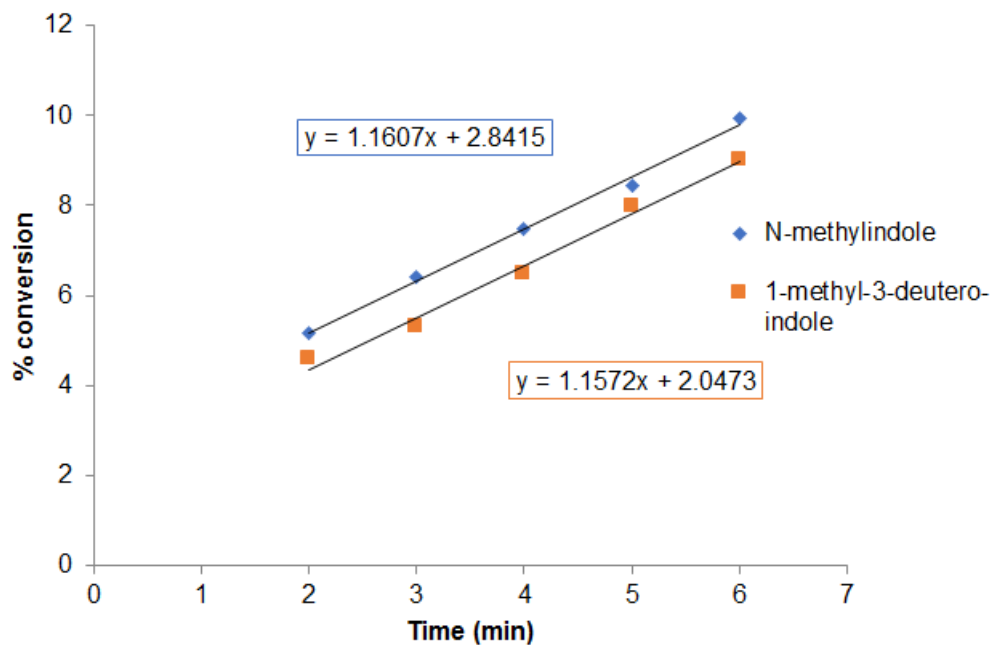


c) GC-MS spectrum of substrate peak after 5 min



Time	Relative Intensity (GC-MS)		
	<i>m/z</i> 130	<i>m/z</i> 131	<i>m/z</i> 132
0 min	49.1	100	56.1
2 min	48.1	100	62.0
5 min	48.5	100	61.6

**Figure S6.** Kinetic isotope effect experiments. The plot reports the initial rates for Mb(H64V,V68A)-catalyzed conversion of 1-methyl-indole (**11a**; blue squares) and of 3d-N-methylindole (**11a-3-d**; orange squares) in the presence of EDA. Comparison of the slopes for the two parallel reactions showed no kinetic isotope effect ( $k_H/k_D = 1.003$ ), indicating that C—H bond cleavage is not part of the rate determining step. Reaction conditions: 2.5 mM **11a** or **11a-3-d**, 5 mM EDA, 10 mM sodium dithionite, 20  $\mu$ M Mb variant in 50 mM phosphate buffer (pH 9.0).



## Experimental Procedures:

**General Information.** All the chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, Alfa Aesar, ACS Scientific, Acros) and used without any further purification, unless otherwise stated. All dry reactions were carried out under argon or nitrogen in oven-dried glassware with magnetic stirring using standard gas-tight syringes, cannula and septa.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were measured on Bruker DPX-500 (operating at 500 MHz for  $^1\text{H}$  and  $^{13}\text{C}$ ).  $\text{CDCl}_3$  was used as the internal standard (77.0 ppm) for  $^{13}\text{C}$  NMR. Silica gel chromatography purifications were carried out using AMD Silica Gel 60 230-400 mesh. Preparative thin layer chromatography was performed on TLC plates (Merck).

**Molecular Cloning.** pET22b(+) (Novagen) was used as the recipient plasmid vector for expression of all of the myoglobin variants. In this construct, the Mb gene is C-terminally fused to a polyhistidine tag and it is under the control of an IPTG-inducible T7 promoter. The cloning of the single-site Mb variants tested in this study was described previously.<sup>[2]</sup> The recombination variants were prepared by combining the desired mutations using a similar cloning procedure and mutagenizing primers reported previously.<sup>[2b]</sup>

**Protein Expression and Purification.** Engineered Mb variants were expressed in *E. coli* C41(DE3) cells as described previously.<sup>[2b]</sup> Briefly, cells were grown in TB medium (ampicillin, 100 mg L<sup>-1</sup>) at 37 °C (200 rpm) until OD<sub>600</sub> reached 1.0-1.2. Cells were then induced with 0.25 mM  $\beta$ -d-1-thiogalactopyranoside (IPTG) and 0.3 mM  $\delta$ -aminolevulinic acid (ALA). After induction, cultures were shaken at 180 rpm and 27 °C and harvested after 18-20 h by centrifugation at 4000 rpm at 4 °C. After cell lysis by sonication, the proteins were purified by Ni-affinity chromatography. The lysate was transfer to a Ni-NTA column equilibrated with Ni-NTA Lysis Buffer. The resin was washed with 50 mL of Ni-NTA Lysis Buffer and then 50 mL of Ni-NTA WashBuffer (50 mM KPi, 250 mM, NaCl, 20 mM imidazole, pH 8.0). Proteins were eluted with Ni-NTA Elution Buffer (50 mM KPi, 250 mM, NaCl, 250 mM histidine, pH 7.0). After elution, the proteins were buffer exchanged against 50 mM KPi buffer (pH 7.0) using 10 kDa Centricon filters. Protein concentration was determined by the CO-binding assay using  $\epsilon_{410} = 156 \text{ mM}^{-1} \text{ cm}^{-1}$  as the extinction coefficient.

**Reactions with purified protein.** Initial reactions (**Table S1**) were carried out at a 400  $\mu\text{L}$  scale using 20  $\mu\text{M}$  myoglobin, 2.5 mM indole, 2.5 mM EDA, and 10 mM sodium dithionite. In a typical procedure, a solution containing sodium dithionite (100 mM stock solution) in potassium phosphate buffer (50 mM, pH 7.0) was degassed by bubbling argon into the mixture for 5 min in a sealed vial. A buffered solution containing myoglobin was carefully degassed in a similar manner in a separate vial. The two solutions were then mixed together via cannulation. Reactions were initiated by addition of 10  $\mu\text{L}$  of indole (from a 100 mM stock solution in ethanol), followed by the addition of 10  $\mu\text{L}$  of EDA (from a 100 mM stock solution in ethanol). The reaction mixture was left under magnetic stirring for 12 hours at room temperature under positive argon pressure. Reactions with hemin were carried out using an identical procedure with the exception that the purified Mb was replaced by hemin (20  $\mu\text{M}$  in THF). For the optimization experiments, reactions were performed according to the general procedure described above, using 20  $\mu\text{M}$  of protein and variable amounts of indole and EDA at different pH (**Table S2**). For the time course experiments (**Figure S2**), reactions were stopped at different time points with 100  $\mu\text{L}$  of 2 N HCl, followed by GC analysis as described below (see **Product Analysis** section for details). For these Mb-catalyzed reactions, higher TON and product conversions were typically observed using the cannulation method described above compared to the same reactions performed in an anaerobic chamber (i.e., Coy Vinyl Anaerobic Chamber).

**Whole Cell Reactions.** Whole-cell experiments (Table S2) were carried out at a 400  $\mu\text{L}$ -scale using 370  $\mu\text{L}$  of *E. coli* cells expressing Mb(H64V,V68A), 2.5 mM indole derivative, and 5 mM EDA. In a typical procedure, a sealed vial containing whole cells was degassed with argon for 5 min. The reactions were initiated by addition of 10  $\mu\text{L}$  of indole (from a 0.1 M stock solution in ethanol), followed by the addition of 20  $\mu\text{L}$  of EDA (from a 0.1 M stock solution in ethanol) with a syringe to a suspension of Mb(H64V,V68A)-expressing cells (varying OD<sub>600</sub> in 50 mM KPi buffer, pH 7.2). The reaction mixture was stirred for 16 hours at room temperature under positive argon pressure. The TON for the whole-cell reactions were calculated based on Mb concentration in the reaction mixture as measured via UV-vis spectroscopy ( $\epsilon_{410} = 156 \text{ mM}^{-1} \text{ cm}^{-1}$ ) after cell lysis.

**Product Analysis.** The reactions were analyzed by adding 20  $\mu\text{L}$  of internal standard (50 mM benzodioxole in ethanol) to the reaction mixture, followed by extraction with 400  $\mu\text{L}$  of dichloromethane (DCM) and analyzed by GC-FID using a Shimadzu GC-2010 gas chromatograph equipped with a FID detector, and a Cyclosil-B column (30 m x 0.25 mm x 0.25  $\mu\text{m}$  film). Separation method: 1  $\mu\text{L}$  injection, injector temp.: 200°C, detector temp.: 300°C. Gradient: column temp. set at 140°C for 3 min, then to 160°C at 1.8°C/min, then 165°C at 1°C/min, then to 245°C at 25°C/min for 6 min. Total run time: 28.3 min. Calibration curves for quantification of the different C—H insertion products were constructed with authentic standards prepared using whole cell biotransformations with Mb(H64V, V68A)-expressing cells or using  $\text{Rh}_2(\text{OAc})_4$  as the catalyst as described in **Synthetic Procedures**. All measurements were performed at least in duplicate. For each experiment, negative control samples containing no protein were included.

**H/D Exchange and Kinetic Isotope Effect Experiments.** These reactions were carried out using 20  $\mu\text{M}$  myoglobin, 2.5 mM 1-methylindole and/or 1-methyl-3-deutero-indole, 5 mM EDA, and 10 mM sodium dithionite in 400  $\mu\text{L}$  potassium phosphate buffer (50 mM, pH 9.0). In a typical procedure, a solution containing sodium dithionite (100 mM stock solution) in potassium phosphate buffer (50 mM, pH 9.0) was degassed by bubbling argon into the mixture for 5 min in a sealed vial. A buffered solution containing myoglobin was carefully degassed in a similar manner in a separate vial. The two solutions were then mixed together via cannulation. Reactions were initiated by addition of 10  $\mu\text{L}$  of 100 mM indole in ethanol, followed by the addition of 20  $\mu\text{L}$  of 100 mM EDA in ethanol. The reaction mixture was left under magnetic agitation for the indicated time under positive argon pressure. The reactions were quenched by exposure to air and immediate extraction with  $\text{CH}_2\text{Cl}_2$  (400  $\mu\text{L}$ ), followed by GC-MS analysis. GC-MS analyses were performed on a Shimadzu GCMS-QP2010 equipped with a RTX-XLB column (30 m x 0.25 mm x 0.28  $\mu\text{m}$ ) and a quadrupole mass analyzer. Separation method: 1  $\mu\text{L}$  injection, injector temperature: 250 °C, detector temperature: 225 °C. Gradient: column temperature set to 50 °C for 1 min, then to 280 °C at 30 °C/min, and held at 280 °C for 3.0 min. Total run time was 12.00 min. The reactions for the KIE experiments (Figure S6) were analyzed for product formation as described above under **Product Analysis**. All measurements were performed in duplicate.

## Synthetic Procedures:

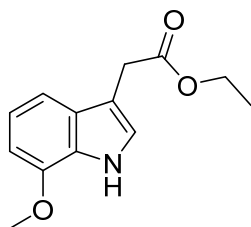
### Chemical synthesis of authentic indole C-H functionalization products (Procedure A):

To generate authentic standards for the different indole functionalized products, Rh<sub>2</sub>(OAc)<sub>4</sub>-catalyzed indole C-H insertion reactions were carried out according to the following general procedure. To a flame dried round bottom flask was added indole (1 equiv.) and Rh<sub>2</sub>(OAc)<sub>4</sub> (5 mol%) in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) under argon. To this solution was added a solution of diazo compound (1 equiv.) in CH<sub>2</sub>Cl<sub>2</sub> (3-5 mL) via slow addition over 3-5 hours using a syringe pump. The resulting mixture was stirred at room temperature overnight. The solvent was removed under vacuum and the crude mixture was purified by flash chromatography using a hexanes:DCM:EtOAc mixture. The identity of the indole products was confirmed by GC-MS, <sup>1</sup>H, and <sup>13</sup>C NMR.

### General procedure for the preparation of N-methyl indoles (Procedure B):

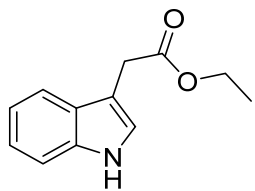
To a stirred solution of indole (1 equiv.) in dry DMF, NaH (1.2 equiv., 60% suspension in mineral oil) was added portion wise under an inert atmosphere at 0 °C. The reaction mixture was then warmed to room temperature and stirred for 30 min. Then, MeI (1.2 equiv.) was added dropwise to the mixture and stirred overnight. Water was added, and the aqueous layer was extracted with ether. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and evaporated under reduced pressure. The residue was purified by column chromatography (hexanes/EtOAc, 20:1). The identity of the N-methylated indole products was confirmed by GC-MS, <sup>1</sup>H, and <sup>13</sup>C NMR.

### Preparative scale procedure for enzymatic synthesis of ethyl 2-(7-methoxy-1h-indol-3-yl)acetate (9b):



To a 500 mL Erlenmeyer flask containing 200mL of a suspension of Mb(H64V,V68A) expressing cells (*E. coli* (C41(DE3), OD<sub>600</sub>=20), 5 mL of 7-methoxyindole (100mM, 0.5 mmol) was added. The flask was sealed with a rubber septum and degassed with argon for 15 min. A solution of EDA (10 mL, 100 mM, 1 mmol) was slowly added using a syringe pump over 2 hours. The reaction mixture was stirred overnight at room temperature under positive argon pressure. The desired product was extracted with ethyl acetate (5 x 10 mL) and washed with water and brine. The organic layers were combined and dried over sodium sulfate, evaporated under reduced pressure, and the residue was purified by column chromatography (hexanes/DCM/EtOAc, 7:2:1) to afford ethyl 2-(7-methoxy-1h-indol-3-yl)acetate **9b** as a clear oil (71.7 mg, 61.0%). GC-MS m/z (% relative intensity): 234(5.6), 233(35.4), 161(15.8), 160(100), 145(24.3); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 8.26 (bs, NH, 1H), 7.21 (d, *J* = 8.0 Hz, 1H), 7.12(s, 1H), 7.02 (t, *J* = 8.0 Hz, 1H), 6.63 (d, *J* = 7.5 Hz, 2H), 4.13 (q, *J* = 7.0 Hz, 2H), 3.92 (s, 3H), 3.72 (s, 2H), 1.23 (t, *J* = 7.0 Hz, 3H) ppm, <sup>13</sup>C NMR (CDCl<sub>3</sub>, 500 MHz): δ 172.0, 146.2, 128.6, 126.7, 122.5, 120.1, 111.7, 109.1, 102.2, 60.8, 55.3, 31.6, 14.2 ppm.

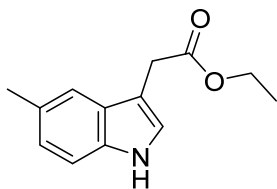
### Ethyl-3-indoleacetate (**3**)



Following the standard procedure **A**, brown oil, yield = 9.7%, GC-MS m/z (% relative intensity): 218(4.2), 217(27.9), 145(12.5), 144(100), 143(10.3); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 8.08 (bs, NH, 1H), 7.62 (d, *J* = 7.5 Hz, 1H), 7.33 (d, *J* = 8.0 Hz, 1H), 7.19 (t, *J* = 7.5 Hz, 1H), 7.12 (d, *J* = 8.0 Hz, 2H), 4.16 (q, *J* = 7.0 Hz, 2H), 3.76 (s, 2H), 1.25 (t, *J* = 7.0 Hz, 3H) ppm, <sup>13</sup>C NMR (CDCl<sub>3</sub>, 500 MHz): δ 172.1, 136.1, 127.3, 123.0, 122.2, 119.6, 118.9, 111.2, 108.6, 60.8, 31.4, 14.3 ppm.

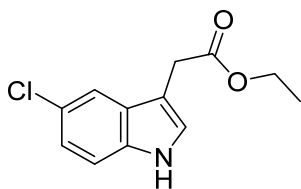


### Ethyl 2-(5-methyl-1h-indol-3-yl)acetate (4b)



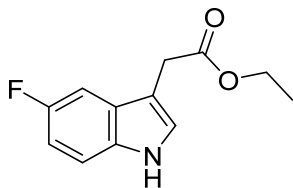
Following the standard procedure A, brown oil, yield = 3.3%, GC-MS  $m/z$  (% relative intensity): 217(22.2), 145(12.9), 144(100), 143(9.3), 115(7.6);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.93 (bs, NH, 1H), 7.37 (s, 1H), 7.21 (d,  $J = 9.0$  Hz, 1H), 7.11 (s, 1H), 6.99 (d,  $J = 8.5$  Hz, 1H), 4.13 (q,  $J = 7.0$  Hz, 2H), 3.71 (s, 2H), 2.42 (s, 3H), 1.22 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  172.2, 135.1, 132.7, 128.5, 121.2, 119.5, 118.1, 110.3, 104.6, 60.7, 30.5, 14.3, 11.7 ppm.

### Ethyl 2-(5-chloro-1h-indol-3-yl)acetate (5b)



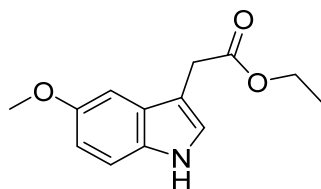
Following the standard procedure A, dark yellow oil, yield = 9.1%, GC-MS  $m/z$  (% relative intensity): 239(6.1), 237(19.6), 166(32.1), 164(100), 128(12.9);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  8.16 (bs, NH, 1H), 7.57 (s, 1H), 7.21 (d,  $J = 8.5$  Hz, 1H), 7.1 (d,  $J = 9.5$  Hz, 1H), 4.16 (q,  $J = 7.0$  Hz, 2H), 3.70 (s, 2H), 1.25 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  171.9, 134.5, 128.4, 125.4, 124.5, 122.5, 118.5, 112.2, 108.4, 61.0, 31.3, 14.2 ppm.

### Ethyl 2-(5-fluoro-1h-indol-3-yl)acetate (6b)



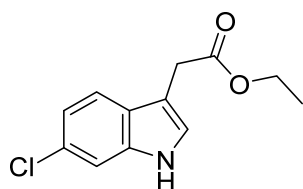
Following the standard procedure **A**, brown oil, yield = 2.4%, GC-MS  $m/z$  (% relative intensity): 221(15.9), 149(10.5), 148(100), 128(9.8), 101(8.1);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  8.03 (bs, NH, 1H), 7.23 (d,  $J = 10.0$  Hz, 1H), 7.19 (s, 1H), 6.91 (t,  $J = 9.0$  Hz, 2H), 4.14 (q,  $J = 7.0$  Hz, 2H), 3.68 (s, 2H), 1.23 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  171.8, 132.6, 124.8, 111.8, 111.7, 110.8, 110.6, 104.1, 103.9, 60.9, 31.4, 14.2 ppm.

**Ethyl 2-(5-methoxy-1h-indol-3-yl)acetate (7b)**



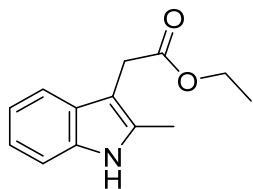
Following the standard procedure **A**, brown oil, yield = 15.7%, GC-MS  $m/z$  (% relative intensity): 233(22.4), 161(12.2), 160(100), 144(13.7), 117(8.8);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.97 (bs, NH, 1H), 7.22 (d,  $J = 9.0$  Hz, 1H), 7.12 (s, 1H), 7.05 (s, 1H), 6.84 (dd,  $J = 8.5, 1.5$  Hz, 1H), 4.16 (q,  $J = 7.0$  Hz, 2H), 3.85 (s, 3H), 3.72 (s, 2H), 1.25 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  172.1, 154.2, 131.3, 127.7, 123.8, 112.6, 111.9, 108.4, 100.7, 60.8, 55.9, 31.6, 14.3 ppm.

**Ethyl 2-(6-chloro-1h-indol-3-yl)acetate (8b)**



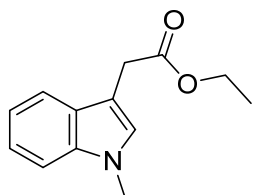
Following the standard procedure **A**, maroon oil, yield = 9.1%, GC-MS  $m/z$  (% relative intensity): 239(11.0), 237(33.7), 166(30.1), 164(100), 128(19.9);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  8.16 (bs, NH, 1H), 7.49 (d,  $J = 8.5$  Hz, 1H), 7.28 (s, 1H), 7.07 (d,  $J = 7.0$  Hz, 2H), 4.16 (q,  $J = 7.0$  Hz, 2H), 3.72 (s, 2H), 1.25 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  172.0, 136.5, 128.1, 125.9, 123.8, 120.4, 119.8, 111.1, 108.8, 60.9, 31.3, 14.2 ppm.

**Ethyl 2-(2-methyl-1h-indol-3-yl)acetate (10b)**



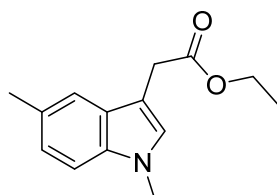
Following the standard procedure A, purple oil, yield = 16.7%, GC-MS  $m/z$  (% relative intensity): 217(30.6), 145(17.5), 144(100), 143(14.7), 115(6.5);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.87 (bs, NH, 1H), 7.53 (d,  $J = 7.5$  Hz, 1H), 7.19 (d,  $J = 7.0$  Hz, 1H), 7.10 (m, 2H), 4.13 (q,  $J = 7.0$  Hz, 2H), 3.68 (s, 2H), 2.34 (s, 3H), 1.23 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  172.2, 135.1, 132.7, 128.5, 121.2, 119.5, 118.1, 110.3, 104.6, 61.0, 30.5, 14.3, 11.7 ppm.

**Ethyl 2-(1-methylindol-3-yl)acetate (11b)**



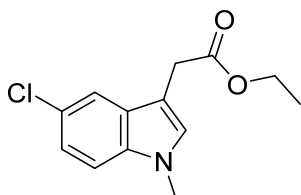
Following the standard procedure A, brown oil, yield = 15.9%, GC-MS  $m/z$  (% relative intensity): 217(18.0), 145(11.9), 144(100), 143(7.1), 127(3.7);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.60 (d,  $J = 8.0$  Hz, 1H), 7.28 (d,  $J = 8.5$  Hz, 1H), 7.21 (t,  $J = 6.5$  Hz, 1H), 7.11 (t,  $J = 7.5$  Hz, 2H), 7.02 (s, 1H), 4.15 (q,  $J = 7.0$  Hz, 2H), 3.74 (s, 5H), 1.25 (t,  $J = 7.0$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  172.2, 136.9, 127.7, 121.7, 122.2, 119.1, 119.0, 109.3, 106.9, 60.8, 32.7, 31.4, 14.3 ppm.

**Ethyl 2-(1,5-dimethylindol-3-yl)acetate (12b)**



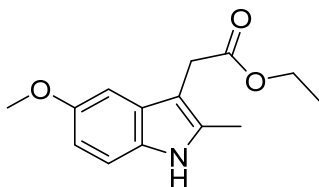
Following the standard procedure **A**, orange oil, yield = 31.1%, GC-MS *m/z* (% relative intensity): 231(44.5), 159(24.2), 158(100), 143(9.9), 115(12.4); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.41 (s, 1H), 7.19 (d, *J* = 8.0 Hz, 1H), 7.07 (d, *J* = 8.0 Hz, 1H), 7.01 (s, 1H), 4.18 (q, *J* = 7.0 Hz, 2H), 3.74 (s, 5H), 2.48 (s, 3H), 1.28 (t, *J* = 7.0 Hz, 3H) ppm, <sup>13</sup>C NMR (CDCl<sub>3</sub>, 500 MHz): δ 172.2, 135.4, 128.3, 127.9, 127.8, 123.4, 118.7, 108.9, 106.3, 60.7, 32.7, 32.7, 31.3, 14.3 ppm.

### Ethyl 2-(5-chloro-1-methylindol-3-yl)acetate (**13b**)



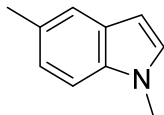
Following the standard procedure **A**, yellow oil, yield = 21.3%, GC-MS *m/z* (% relative intensity): 253(20.1), 251(59.7), 180(59.1), 178(100), 143(24.2); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.57 (s, 1H), 7.20 (d, *J* = 11.0 Hz, 1H), 7.17 (d, *J* = 11.0 Hz, 1H), 7.06 (s, 1H), 4.17 (q, *J* = 9.0 Hz, 2H), 3.74 (m, 5H), 1.28 (t, *J* = 9.0 Hz, 3H) ppm, <sup>13</sup>C NMR (CDCl<sub>3</sub>, 500 MHz): δ 171.8, 135.3, 129.1, 128.7, 125.1, 122.1, 122.0, 118.9, 118.6, 110.4, 110.3, 106.7, 60.9, 32.8, 31.1, 14.2 ppm.

### Ethyl 2-(5-methoxy-2-methyl-1H-indol-3-yl)acetate (**14b**)



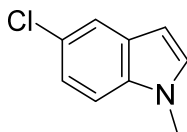
Following the standard procedure **A**, yellow oil, yield = 12.4%, GC-MS *m/z* (% relative intensity): 247(28.0), 175(13.5), 174(100), 159(14.7), 131(7.9); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.91 (bs, NH, 1H), 7.07 (d, *J* = 8.5 Hz, 1H), 6.98 (d, *J* = 2.0 Hz, 1H), 6.73 (dd, *J* = 8.5, 2.0, Hz, 1H), 4.10 (q, *J* = 7.0 Hz, 2H), 3.82 (s, 3H), 3.61 (s, 2H), 2.31 (s, 3H), 1.21 (t, *J* = 7.0 Hz, 3H) ppm, <sup>13</sup>C NMR (CDCl<sub>3</sub>, 500 MHz): δ 172.1, 154.1, 133.6, 130.2, 129.0, 111.0, 110.9, 104.5, 100.5, 60.8, 55.9, 30.6, 14.3, 11.8 ppm.

## 1,5-dimethylindole



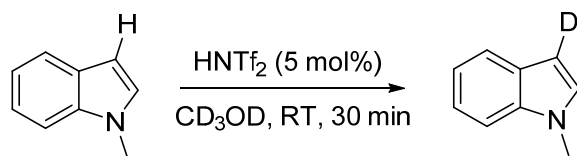
Following the standard procedure **B**, yellow oil, yield = 75.4%, GC-MS  $m/z$  (% relative intensity): 146(12.0), 145(99.9), 144(100), 130(7.6), 115(10.4);  $^1\text{H NMR}$  ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.43 (s, 1H), 7.24 (d,  $J = 8.5$  Hz, 1H), 7.07 (dd,  $J = 8.0, 0.5$  Hz, 1H), 7.02 (d,  $J = 3.0$  Hz, 1H), 6.41 (dd,  $J = 3.0, 0.5$  Hz, 1H), 3.78 (s, 3H), 2.47 (s, 3H), ppm,  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  135.2, 128.8, 128.7, 128.4, 123.1, 120.5, 108.9, 100.3, 32.8, 21.4 ppm.

## 5-chloro-N-methylindole



Following the standard procedure **B**, yellow oil, yield = 66.9%, GC-MS  $m/z$  (% relative intensity): 167(34.5), 166(29.1), 165(100), 164(59.4), 129(7.4);  $^1\text{H NMR}$  ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.59 (d,  $J = 2.0$  Hz, 1H), 7.23 (d,  $J = 10.5$  Hz, 1H), 7.18 (dd,  $J = 10.5, 2.0$  Hz, 1H), 7.08 (d,  $J = 3.5$  Hz, 1H), 6.43 (d,  $J = 3.5$  Hz, 1H), 3.78 (s, 3H) ppm,  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  135.2, 130.1, 129.5, 125.1.0, 121.8, 120.2, 110.2, 100.6, 32.9 ppm.

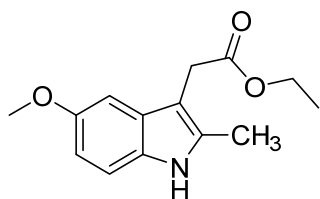
## Synthesis of 1-methyl-3-deutero-indole (**11a-3-d**)



1-Methyl-3-deutero-indole (**11a-3-d**) was prepared according to a modified version of a reported procedure.<sup>[3]</sup> To a flame dried round bottom flask 10 mL of  $\text{CD}_3\text{OD}$  were added via syringe followed by 50 mg (0.38 mmol) of N-methylindole. A solution of bis(trifluoromethanesulfonyl)imide ( $\text{HNTf}_2$ , 5.4 mg, 0.019 mmol) in  $\text{CD}_3\text{OD}$  (1 mL) was added to the reaction mixture and left stirring for 30 minutes under Ar. The product was extracted with dry

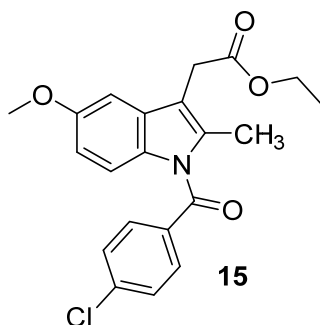
CH<sub>2</sub>Cl<sub>2</sub> and washed with D<sub>2</sub>O (3x) to afford 3-deutero-1-methyl-indole (**11a-3-d**) as a yellow oil (44.6 mg, 89% yield, 100% deuterium incorporation as determined by NMR). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.60 (d, *J* = 7.5 Hz, 1H), 7.30 (d, *J* = 8.0 Hz, 1H), 7.20 (t, *J* = 8.0 Hz, 1H), 7.08 (t, *J* = 7.5 Hz, 1H), 7.02 (s, 1H), 3.77 (s, 3H) ppm.

### Chemo-enzymatic synthesis of indomethacin



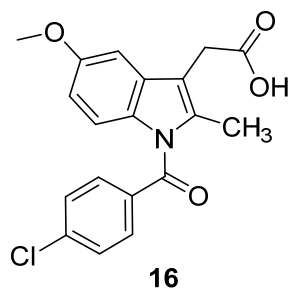
**14b**

**Enzymatic Synthesis of 14b:** To a 500 mL Erlenmeyer flask containing 400mL of a suspension of Mb(L29F,H64V) expressing cells (*E. coli* (C41(DE3), OD<sub>600</sub>=20), 10 mL of 5-methoxy-2-methylindole (100 mM, 1.0 mmol) was added. The flask was sealed with a rubber septum and degassed with argon for 15 min. A solution of EDA (20 mL, 100 mM, 2 mmol) was slowly added using a syringe pump over 2 hours. The reaction mixture was stirred overnight at room temperature under positive argon pressure. The desired products were extracted with ethyl acetate (5 x 10 mL) and washed with water and brine. The organic layers were combined, dried over sodium sulfate, evaporated under reduced pressure, and the residue was purified by column chromatography (hexanes/DCM/EtOAc, 7:2:1) to afford ethyl 2-(5-methoxy-2-methyl-1h-indol-3-yl)acetate **14b** as a yellow oil (99.3 mg, 40%).



**Synthesis of compound 15:** Acylation of **14b** to give indomethacin ethyl ester was performed following a reported procedure.<sup>[4]</sup> Briefly, to a solution of <sup>t</sup>BuOK (25.2 mg, 0.23 mmol) in THF

(2.0 mL) was added a solution of **14b** in THF (44.8 mg, 0.18 mmol) at  $-78^{\circ}\text{C}$ . The resulting solution was left under magnetic agitation for 30 min. To the mixture, a solution of *p*-chlorobenzoyl chloride (28.6  $\mu\text{L}$ , 0.23 mmol) in THF (1.5 mL) was added dropwise and the solution was left stirring for 2 hours. The reaction was quenched with  $\text{NH}_4\text{Cl}$  (10 mL). The organic compounds were extracted with EtOAc, and the combined layers were washed with saturated aqueous  $\text{NaHCO}_3$ ,  $\text{H}_2\text{O}$ , brine, and dried over  $\text{Na}_2\text{SO}_4$ . The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexanes/EtOAc, 8.5:1.5) to afford indomethacin ethyl ester **15** (36.1 mg, 52%) as a white powder. LC-MS  $m/z$  :408.4  $[\text{M}+\text{Na}]^+$ ;  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.67 (d,  $J = 8.0$  Hz, 2H), 7.47 (d,  $J = 8.5$  Hz, 2H), 6.98 (s, 1H), 6.89 (d,  $J = 9.0$  Hz, 1H), 6.68 (dd,  $J = 9.0$  Hz, 1H), 4.19 (q,  $J = 7.5$  Hz, 2H), 3.84 (s, 3H), 3.65 (s, 3H), 2.38 (s, 2H), 1.28 (t,  $J = 7.5$  Hz, 3H) ppm,  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  170.9, 168.3, 156.1, 139.2, 135.9, 133.9, 131.2, 130.8, 130.7, 129.1, 114.9, 112.7, 111.7, 101.3, 77.3, 77.1, 76.8, 61.0, 55.7, 30.4, 14.3, 13.4 ppm.

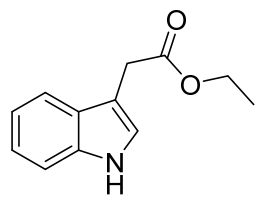


**Synthesis of Indomethacin (16):** To a solution of indomethacin ethyl ester (23.7 mg, 0.061 mmol) in a 9:1 MeOH:H<sub>2</sub>O solution, NaOH (1% vol, 0.426 g, 10.6 mmol) was added and the mixture was left under reflux for 1 hour. After cool down, the solution was acidified with HCl (1M) and the product was extracted in EtOAc (3x). The organic layers were washed with water (2x) and brine (2x) and dried over  $\text{Na}_2\text{SO}_4$ . The solvent was removed under reduced pressure to afford indomethacin **16** (21.7 mg, 99%);  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 500 MHz):  $\delta$  7.94 (d,  $J = 7.0$  Hz, 2H), 7.41 (d,  $J = 8.5$  Hz, 2H), 7.09 (d,  $J = 8.5$  Hz, 1H), 6.91 (s, 1H), 6.64 (dd,  $J = 8.5$  Hz, 1H), 3.76 (s, 3H), 3.57 (s, 2H), 2.31 (s, 3H),  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 500 MHz):  $\delta$  175.1, 167.5, 153.6, 138.9, 133.7, 130.9, 130.8, 129.3, 128.8, 128.3, 110.6, 109.9, 103.5, 99.9, 55.1, 29.6, 10.3 ppm.

## References

- [1] M. L. Quillin, R. M. Arduini, J. S. Olson, G. N. Phillips, *J. Mol. Biol.* **1993**, *234*, 140-155.
- [2] a) M. Bordeaux, V. Tyagi, R. Fasan, *Angew. Chem. Int. Ed.* **2015**, *54*, 1744–1748; b) P. Bajaj, G. Sreenilayam, V. Tyagi, R. Fasan, *Angew. Chem. Int. Ed.* **2016**, *55*, 16110–16114.
- [3] D. V. Patil, S. W. Kim, Q. H. Nguyen, H. Kim, S. Wang, T. Hoang, S. Shin, *Angew. Chem. Int. Ed.* **2017**, *56*, 3670-3674.
- [4] Y. Kasaya, K. Hoshi, Y. Terada, A. Nishida, S. Shuto, M. Arisawa, *Eur. J. Med. Chem.* **2009**, 4606-4613.





1.581  
1.279  
1.267  
1.253  
1.239

4.181  
4.167  
4.153  
4.139  
3.761

8.083  
7.624  
7.609  
7.336  
7.320  
7.240  
7.202  
7.187  
7.172  
7.126  
7.110

4.71

2.74

2.37

2.07

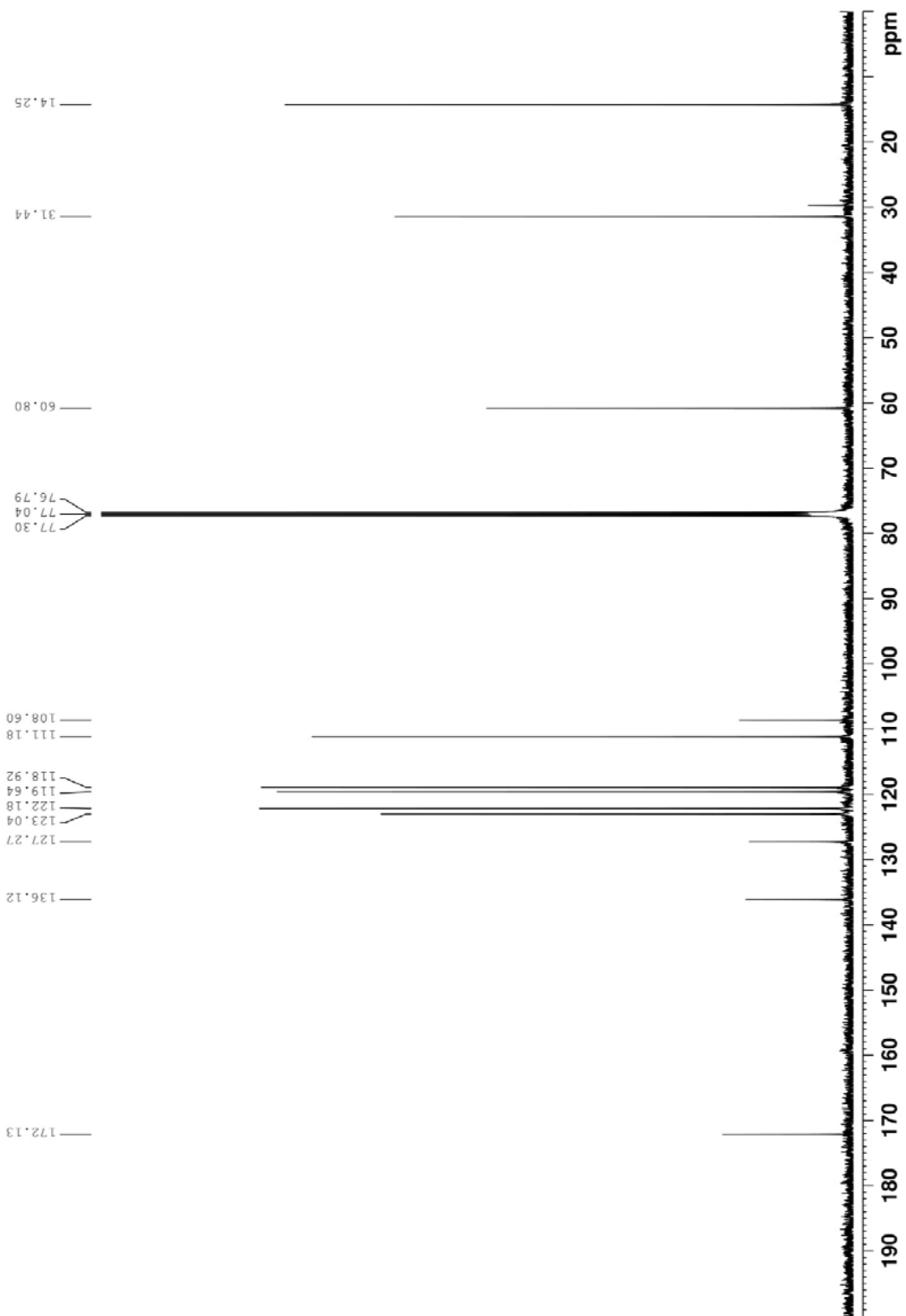
1.14

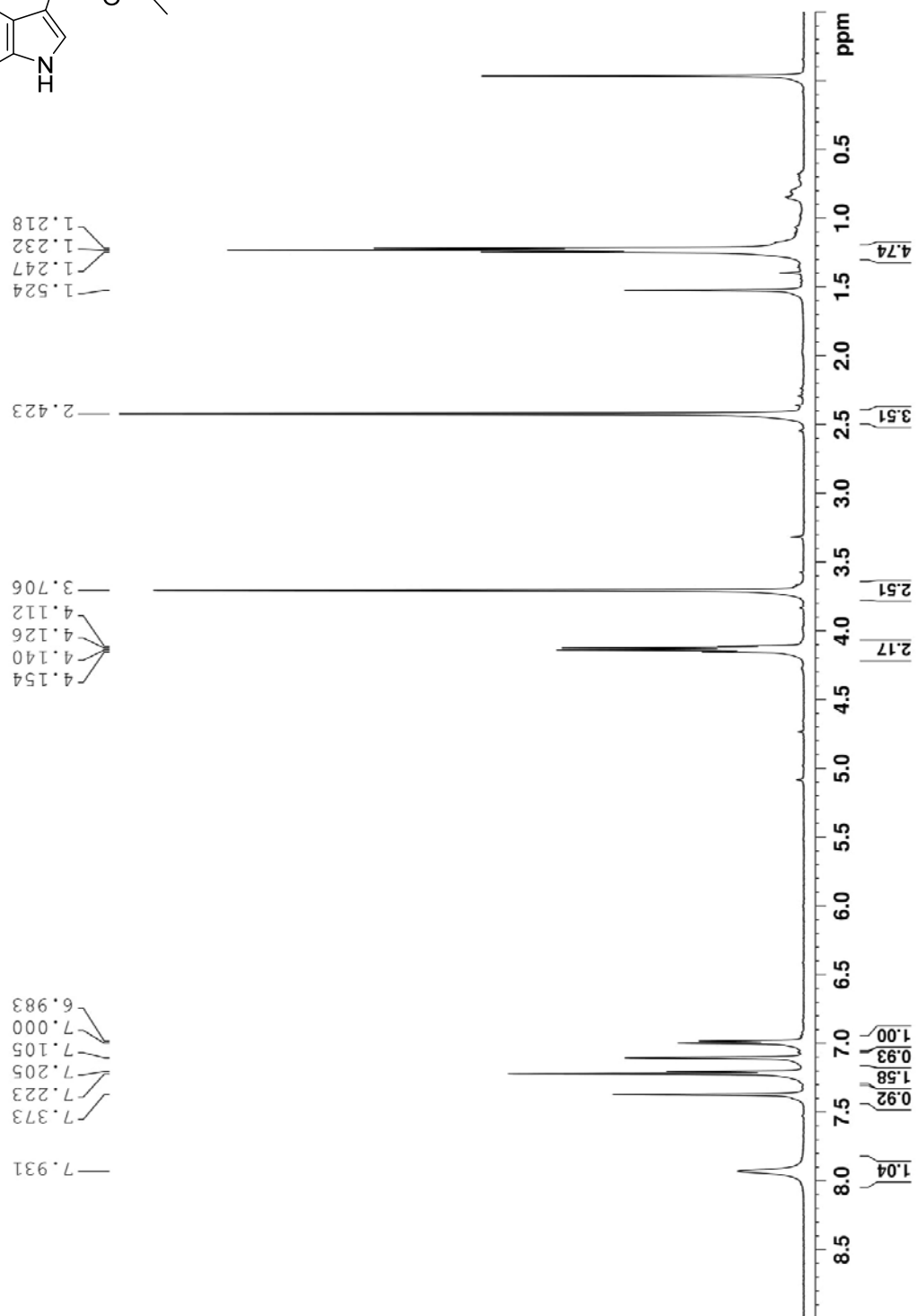
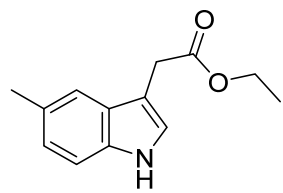
1.00

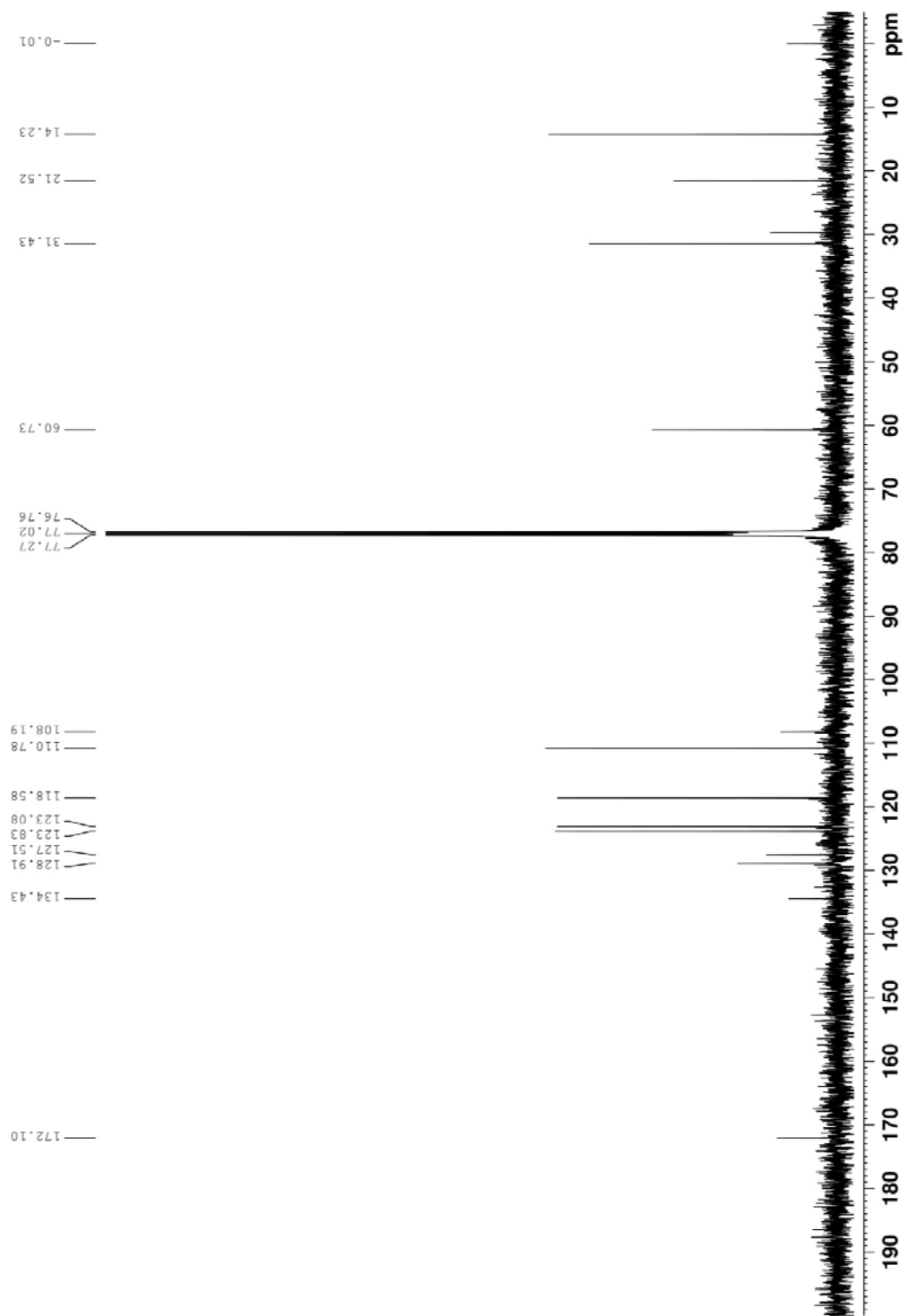
1.00

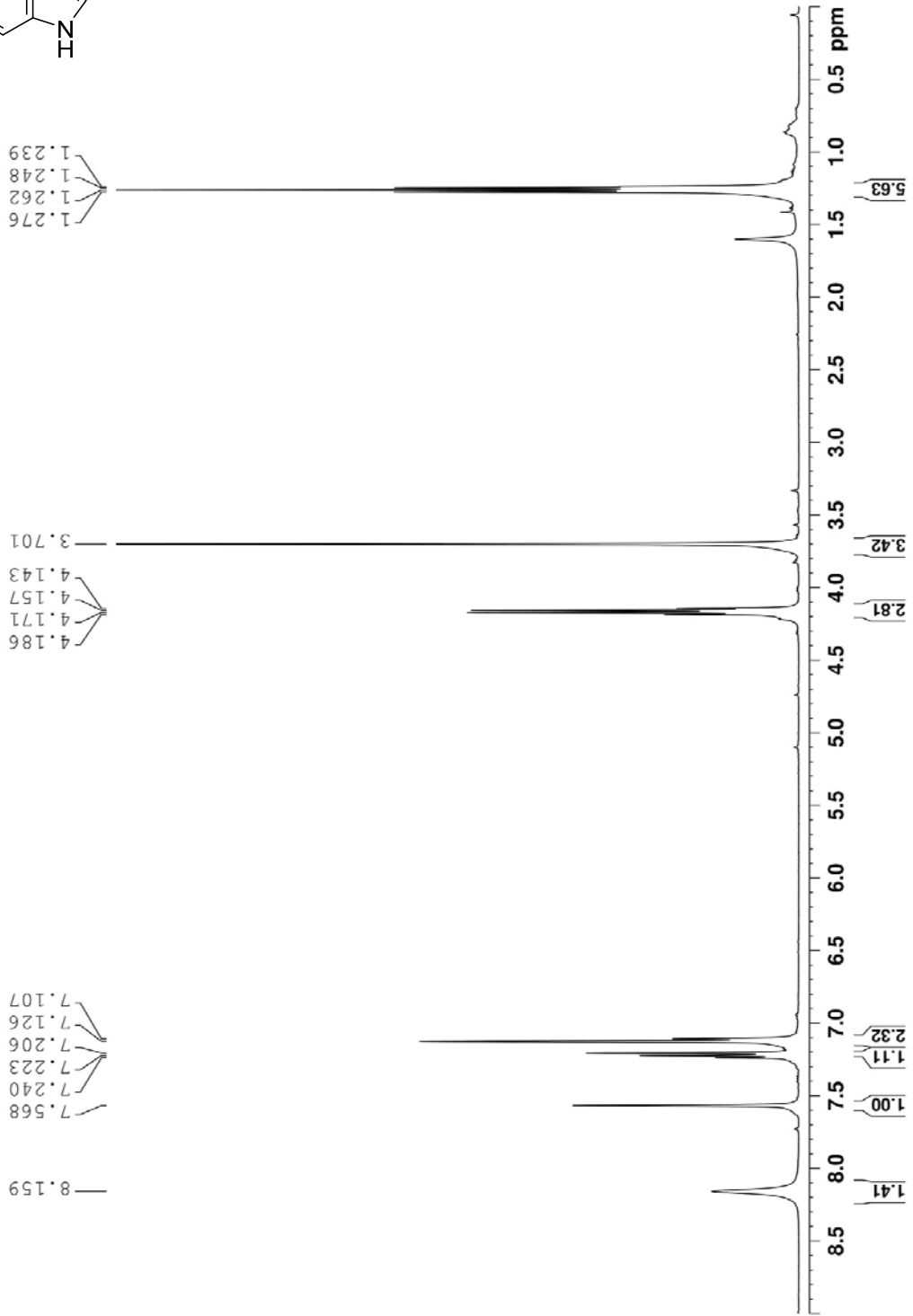
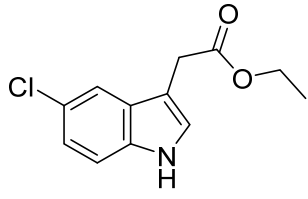
1.10

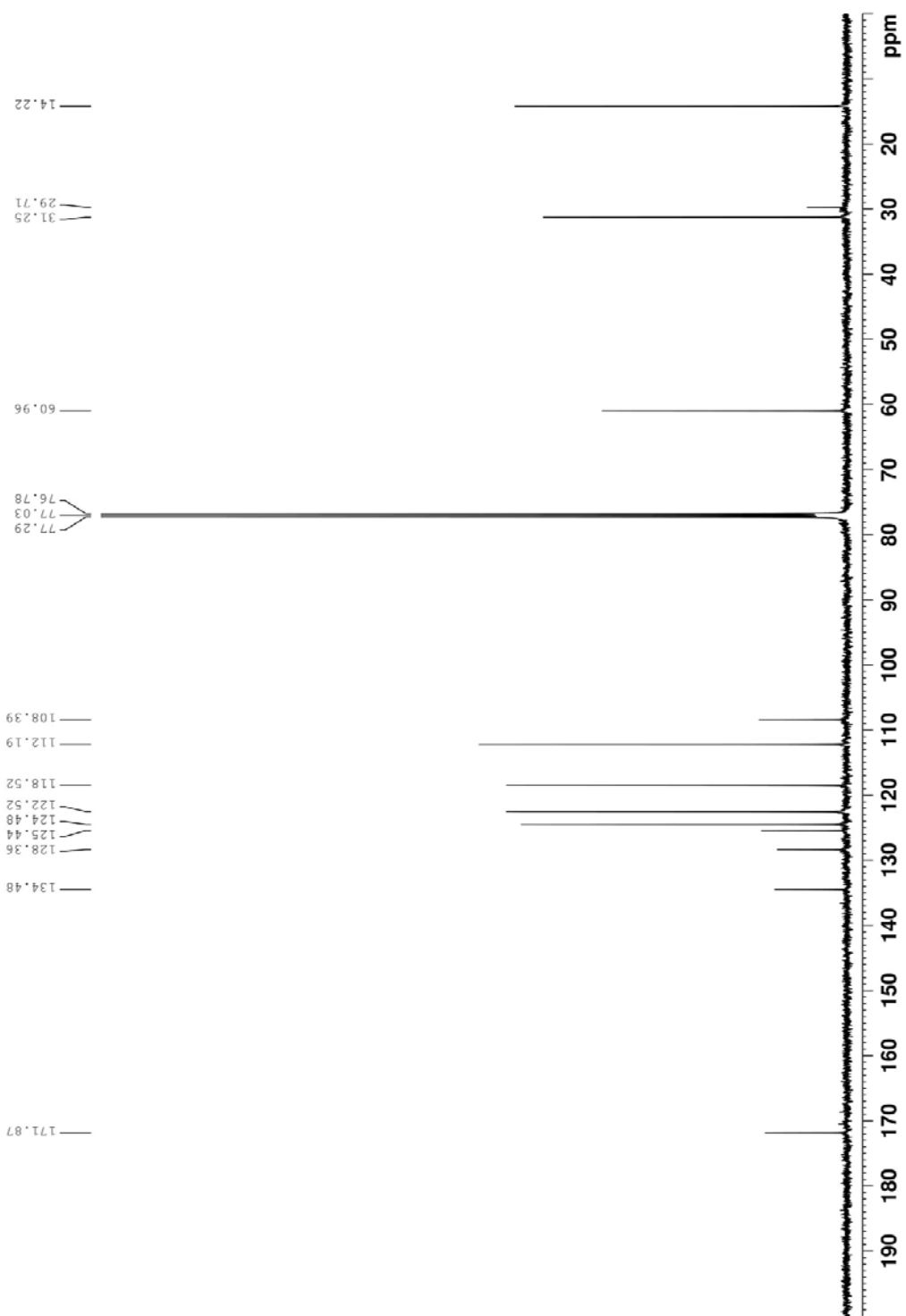
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 ppm

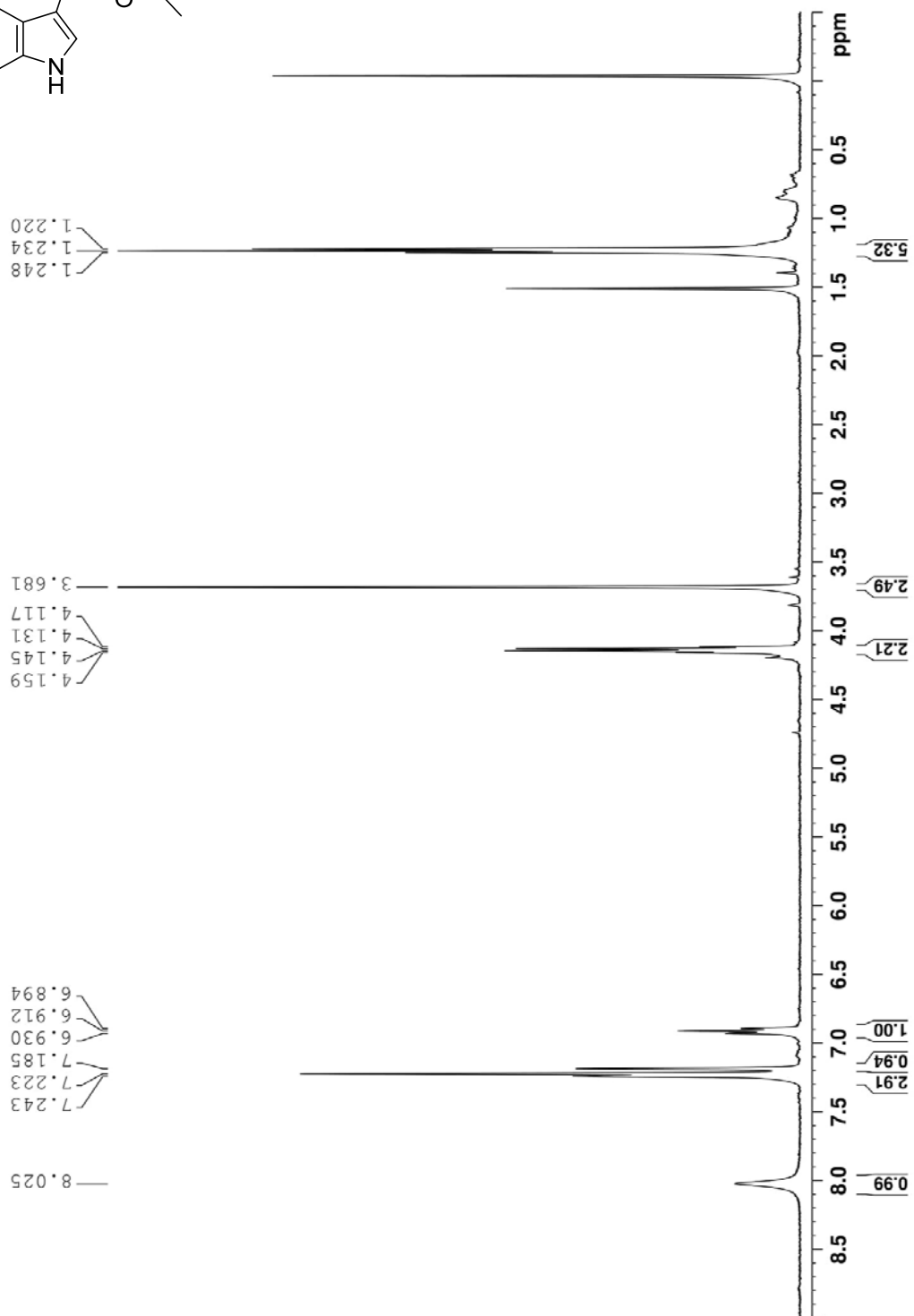
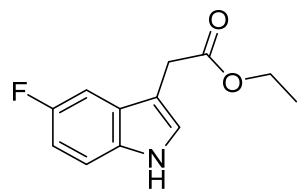


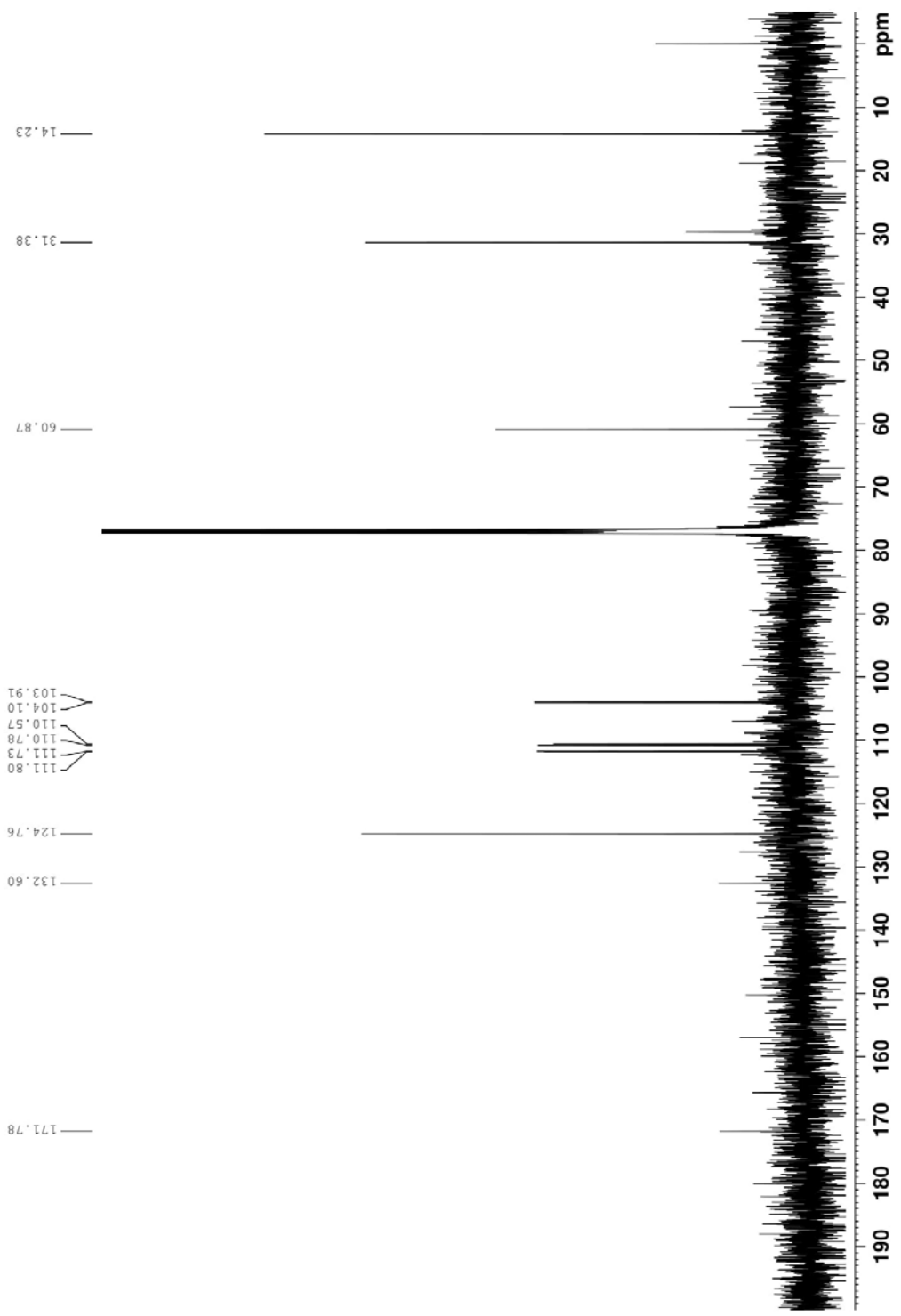




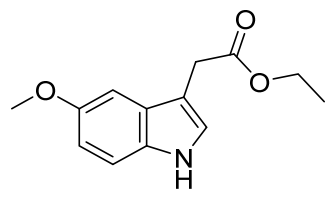










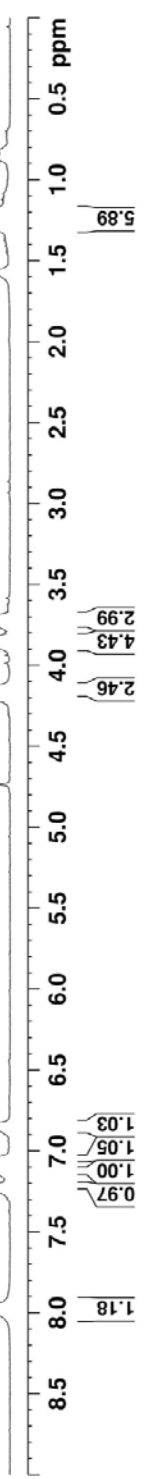


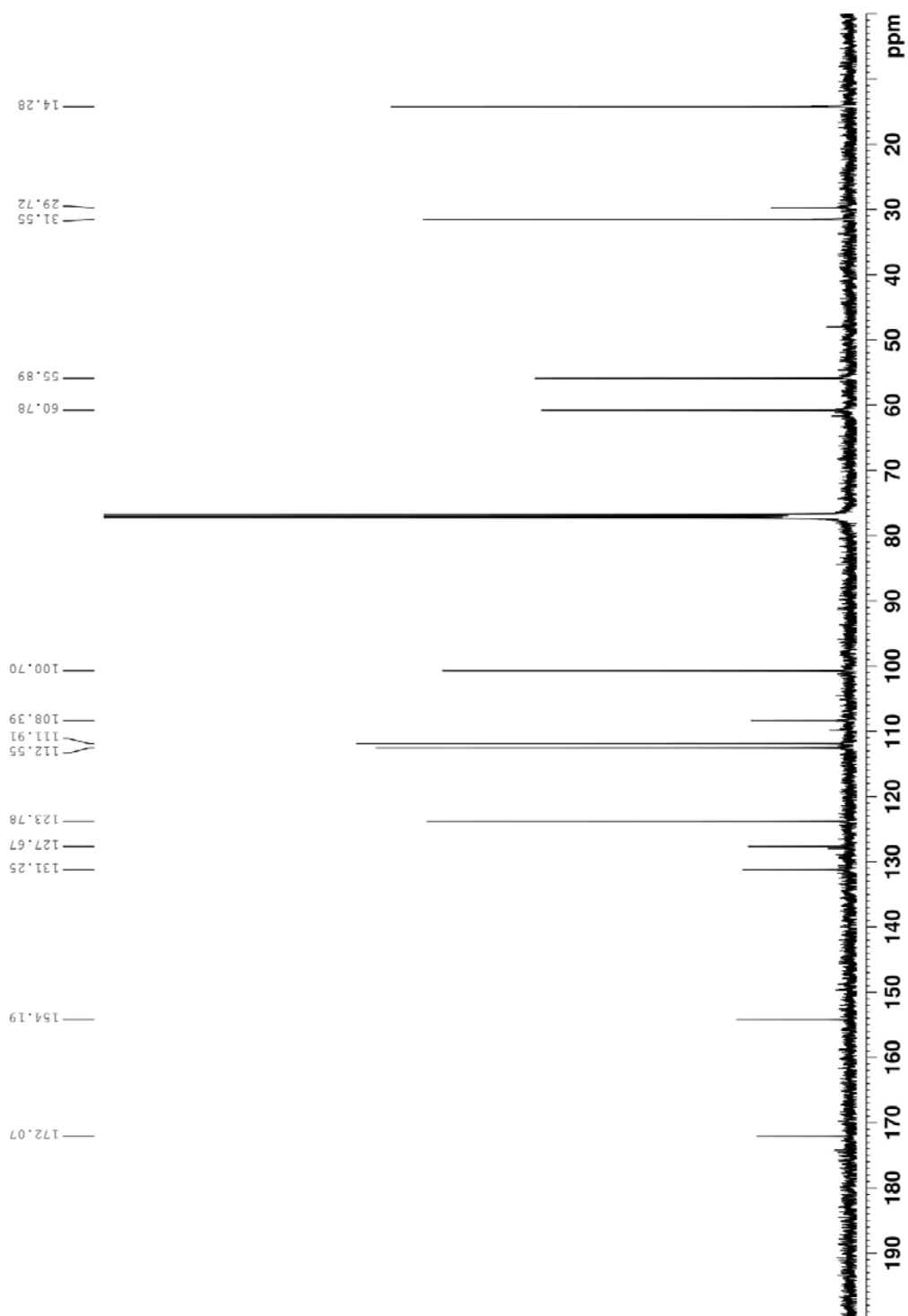
1.265  
1.251  
1.237

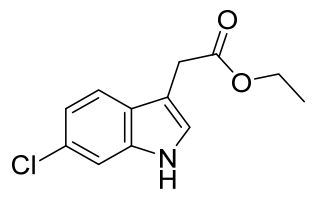
3.716  
3.846  
4.132  
4.146  
4.160  
4.174

6.834  
6.837  
6.851  
6.854  
7.050  
7.124  
7.210  
7.228  
7.239

7.974





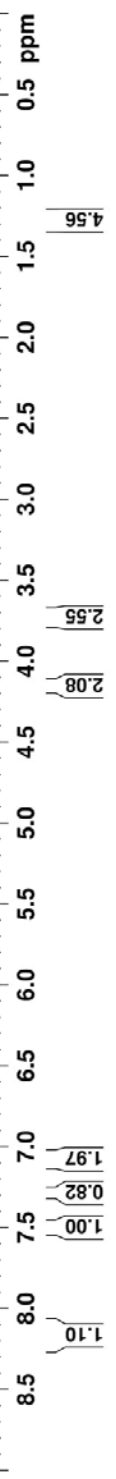


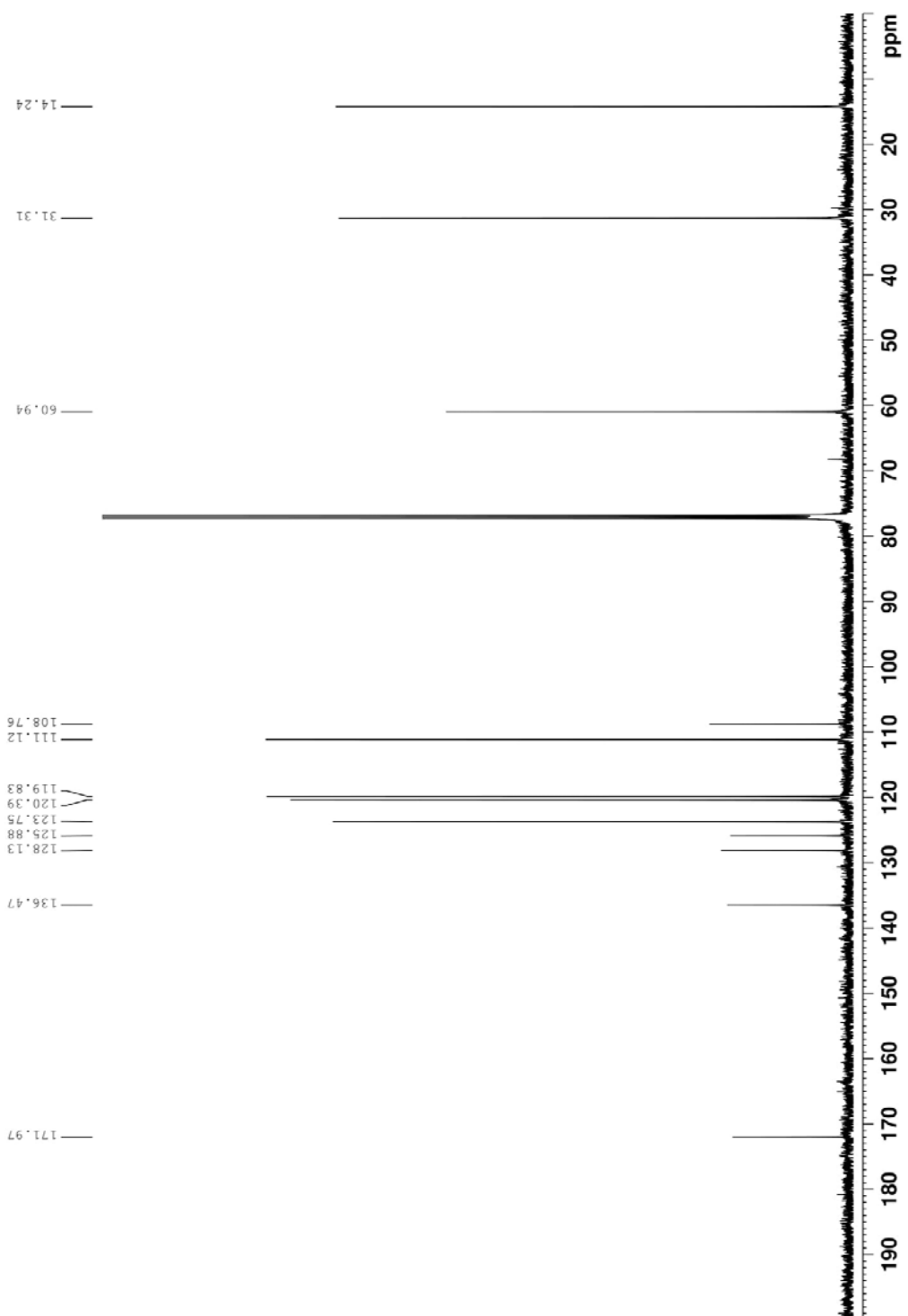
1.263  
1.249  
1.235

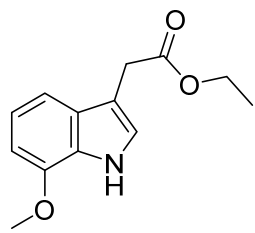
4.218  
4.180  
4.166  
4.152  
4.138  
3.716

7.500  
7.483  
7.272  
7.239  
7.081  
7.067

8.155





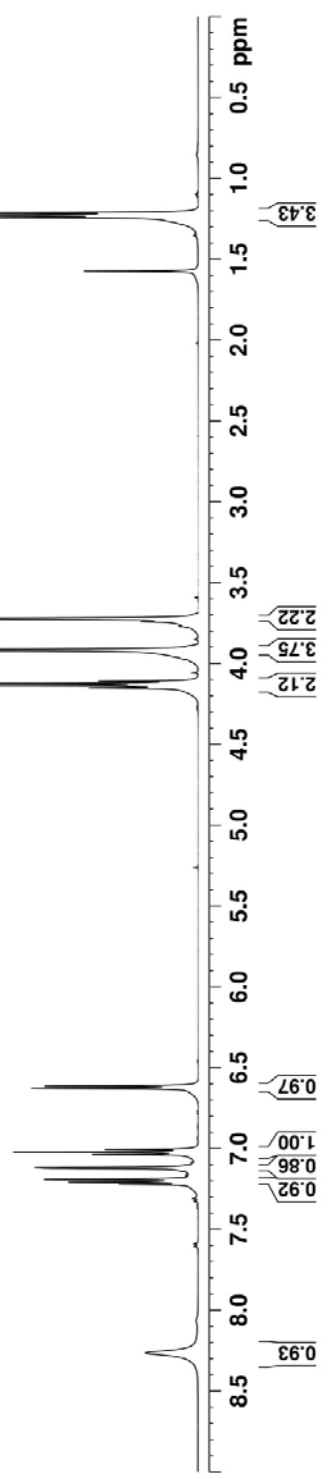


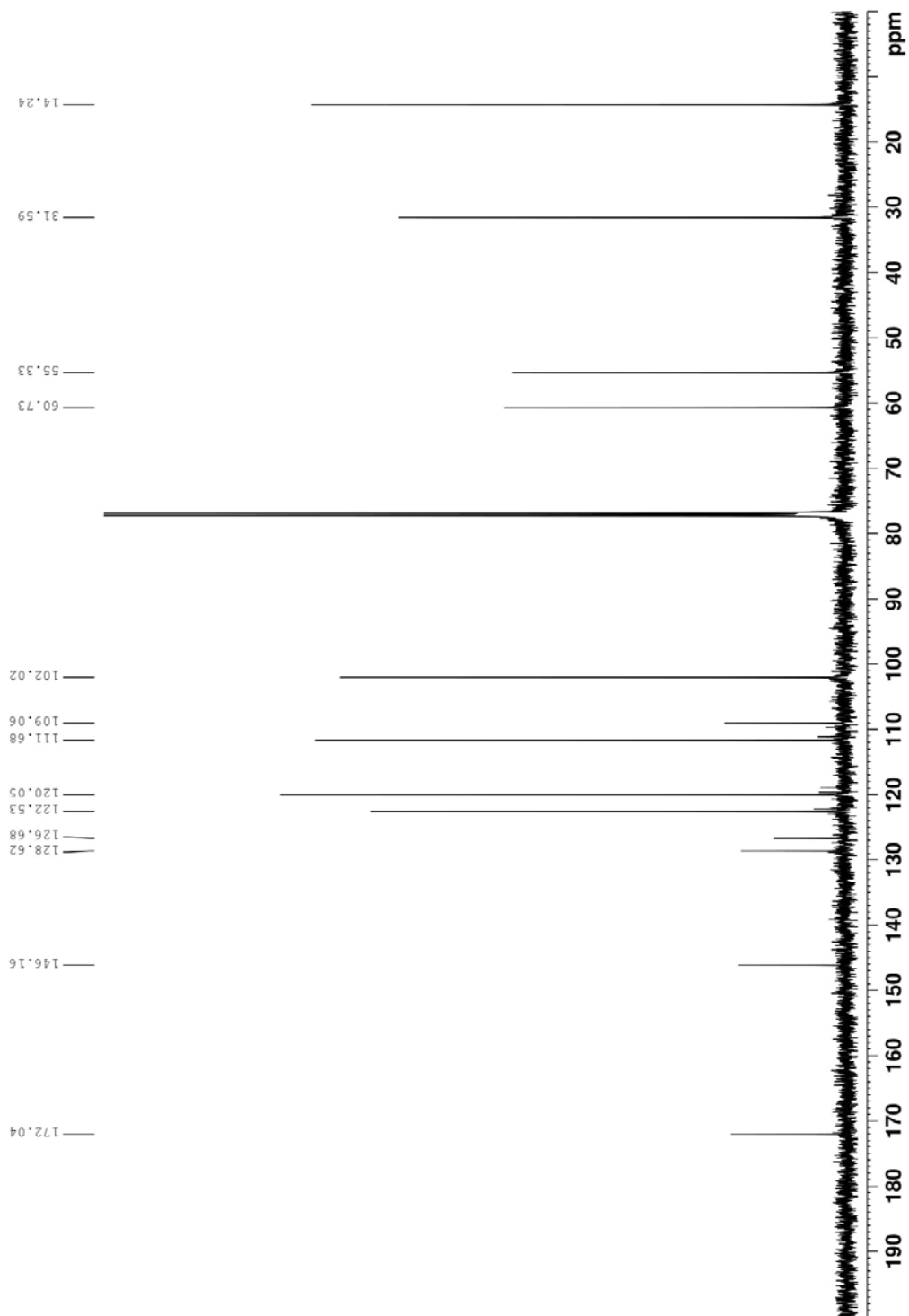
1.242  
1.228  
1.213

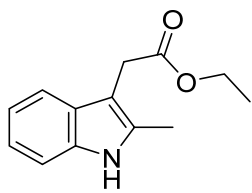
4.153  
4.139  
4.124  
4.110  
3.918  
3.723

7.224  
7.208  
7.192  
7.119  
7.038  
7.022  
7.006  
6.629  
6.614

8.264





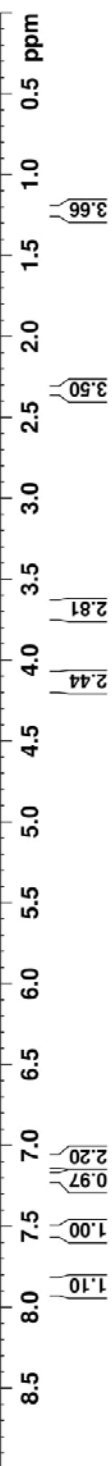


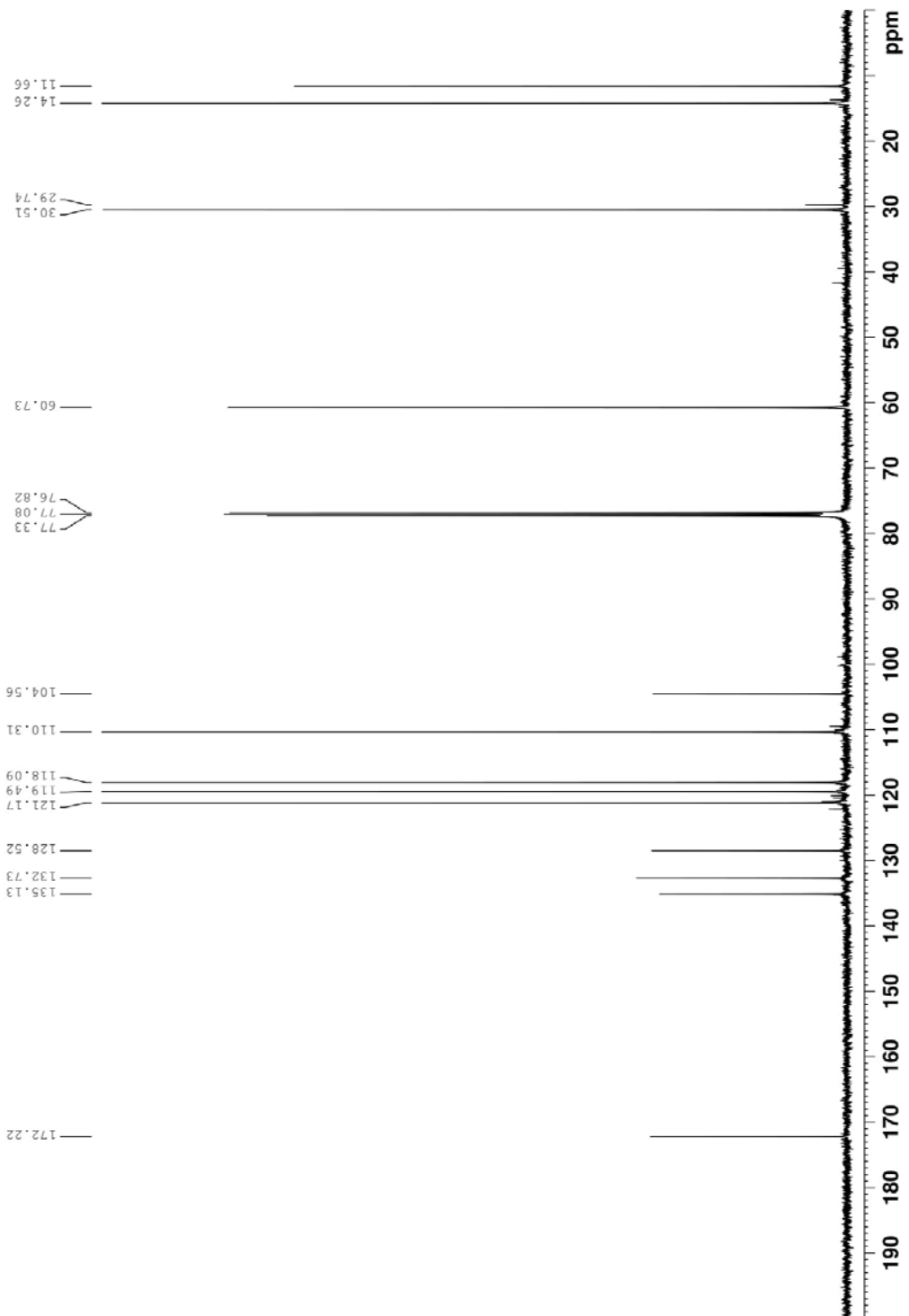
1.266  
1.249  
1.235  
1.221

2.372  
2.339

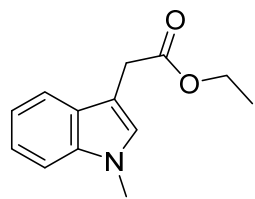
3.676  
4.105  
4.119  
4.133  
4.147

7.874  
7.541  
7.526  
7.240  
7.203  
7.189  
7.121  
7.104  
7.098  
7.089  
7.075





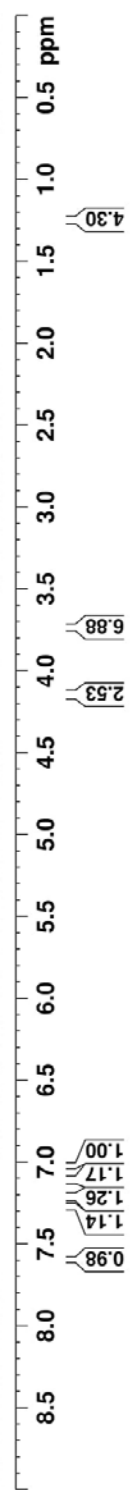


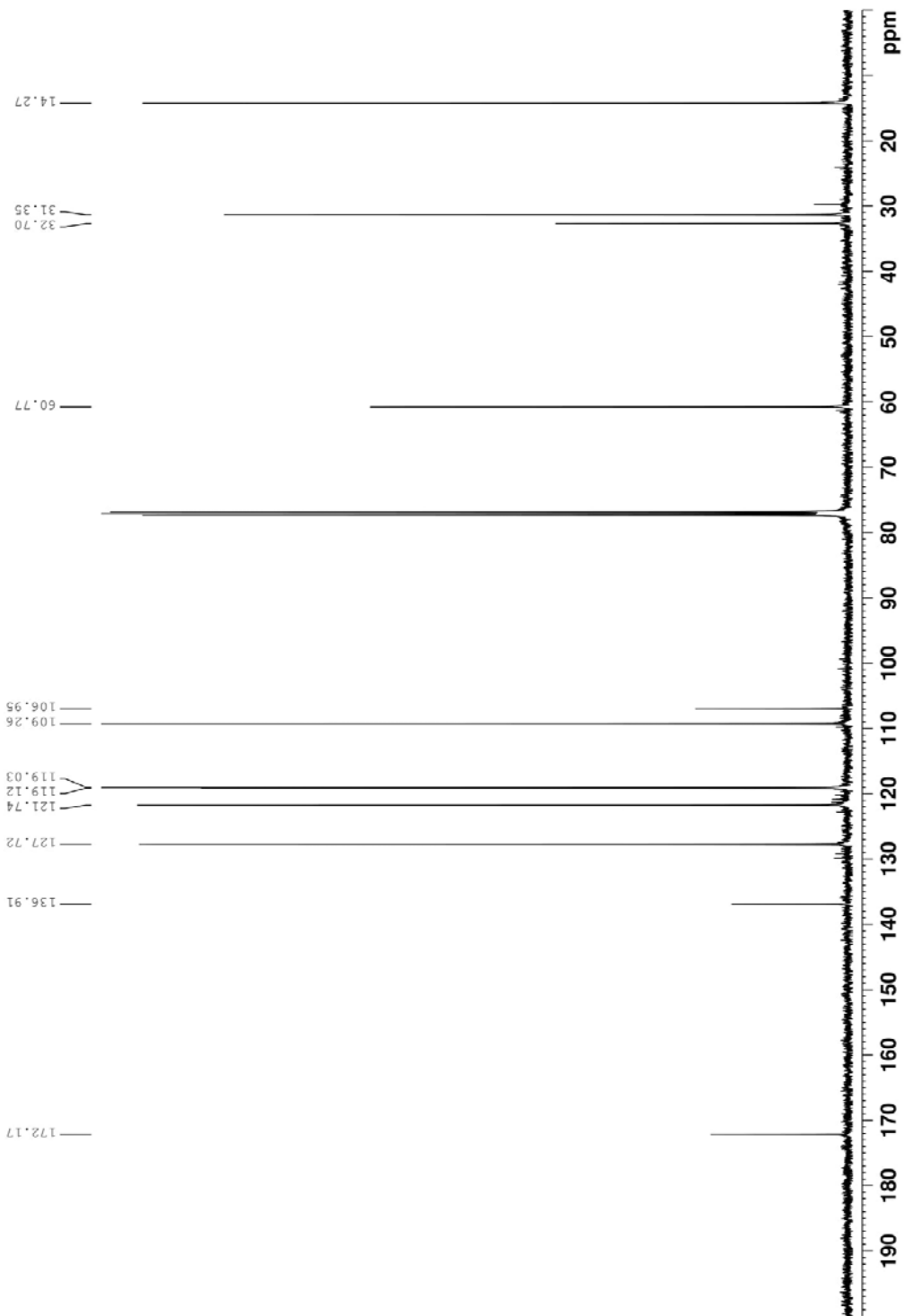


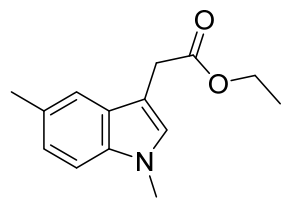
1.577  
1.259  
1.245  
1.231

4.166  
4.152  
4.138  
4.124  
3.737  
3.731

7.602  
7.586  
7.282  
7.265  
7.225  
7.212  
7.196  
7.122  
7.107  
7.092  
7.022





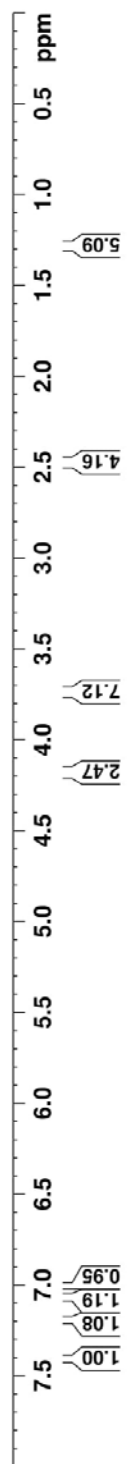


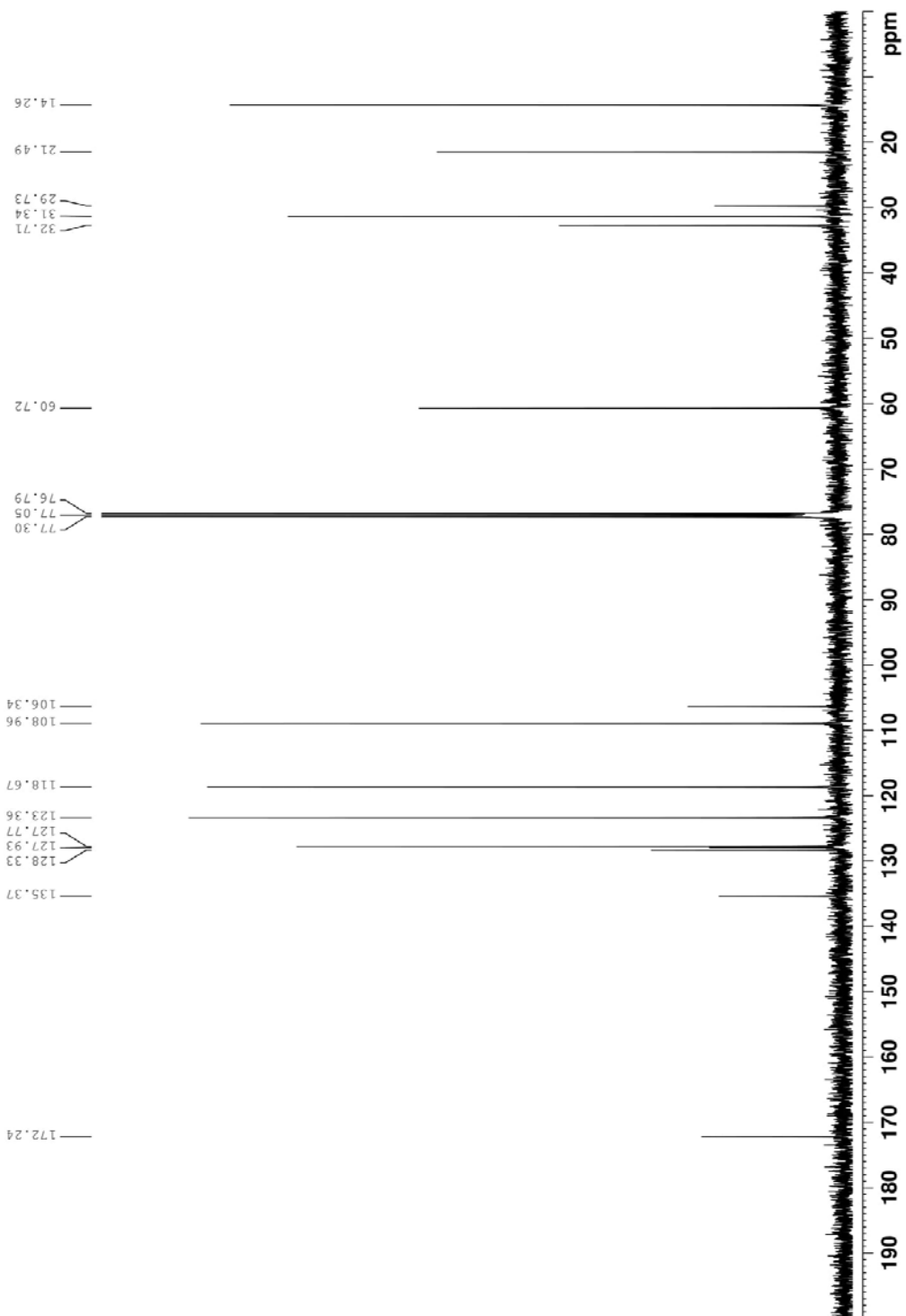
1.299  
1.284  
1.270

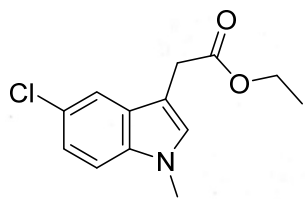
2.478

3.737  
4.159  
4.173  
4.187  
4.202

7.405  
7.260  
7.200  
7.184  
7.074  
7.058  
7.006



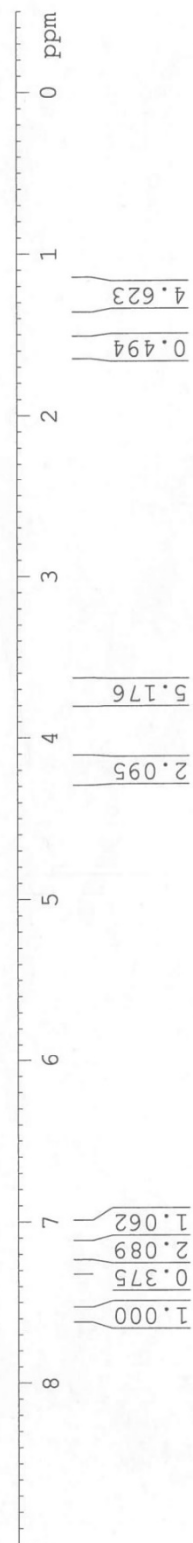


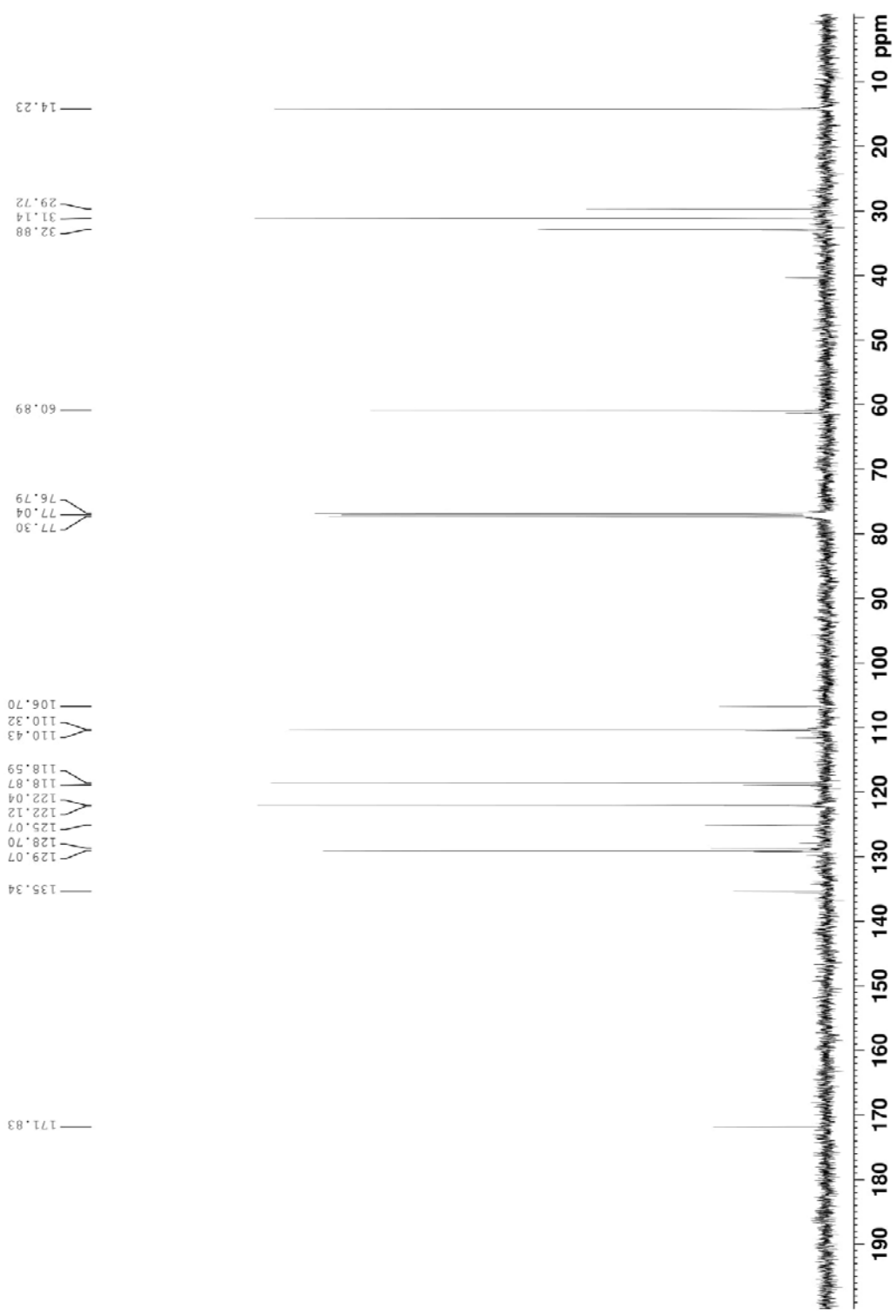


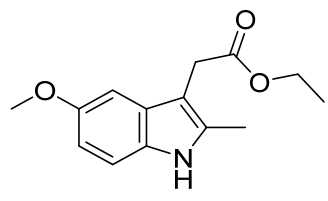
-1.581  
-1.579  
1.295  
1.277  
-1.259

4.199  
4.181  
4.163  
-4.145  
3.743  
3.718  
-3.701

-7.573  
7.260  
-7.212  
7.190  
-7.176  
7.154  
7.060







1.228  
1.214  
1.200

2.340  
2.312

3.613  
3.823  
4.078  
4.093  
4.107  
4.121

7.791  
7.223  
7.075  
7.058  
6.977  
6.973  
6.744  
6.740  
6.727  
6.723

