Stepwise O Atom Transfer in Heme-Based Tryptophan Dioxygenase: Role of Substrate Ammonium in Epoxide Ring Opening

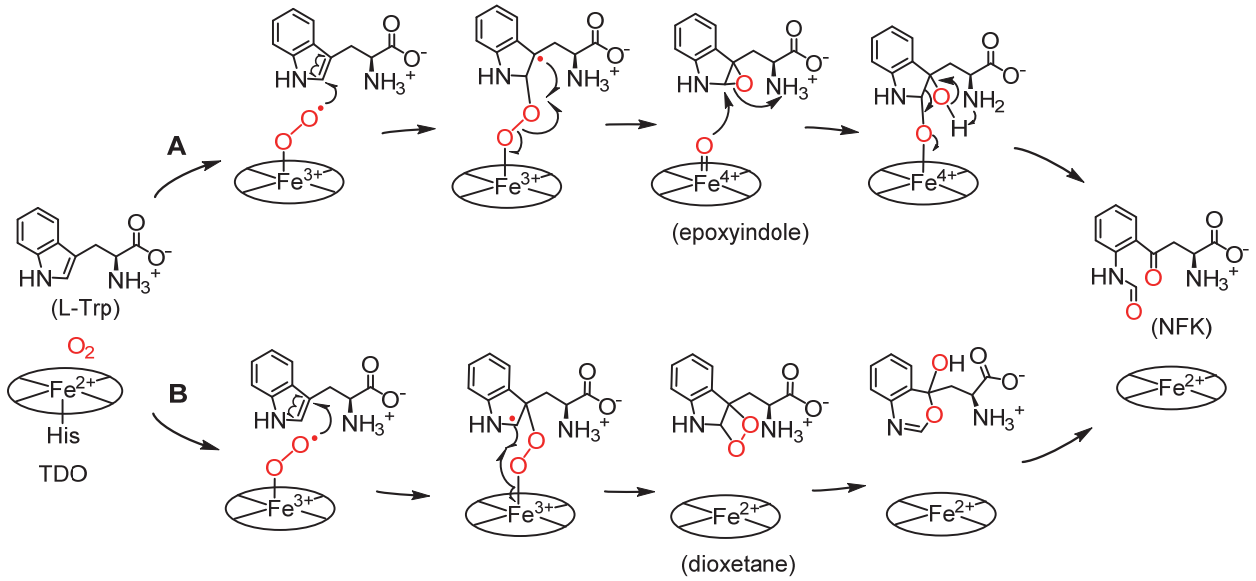
Inchul Shin,[†] Brett R. Ambler,[‡] Daniel Wherritt,[†] Wendell P. Griffith,[†] Amanda Maldonado,[‡] Ryan A. Altman,[‡] and Aimin Liu*[,†]

[†] Department of Chemistry, University of Texas at San Antonio, San Antonio, Texas 78249, United States; [‡] Department of Medicinal Chemistry, The University of Kansas, Lawrence, Kansas 66045, United States

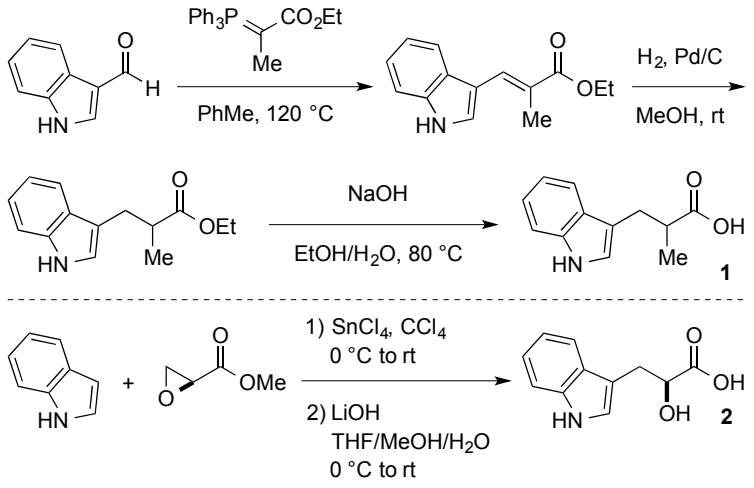
This document contains the following supplementary information: **Scheme S1**, providing the mechanistic models of both stepwise and concerted oxygen insertion; **Scheme S2** and associated text, providing the details of chemical synthesis and characterization of the molecular probes of **1** and **2**; **Figure S1**, **Table S1**, **Scheme S3**, and **Figure S2**, providing the identification of the peak at 7.2 min in **Figure 4A(b)**; **Scheme S4**, providing the chemical structures and masses for the tandem mass spectrometry analysis; **Table S2**, providing the key numbers of NMR shifts; and **Figures S3−S8**, providing the NMR and IR spectroscopic characterization of the TDO reaction products from the synthetic probe **2**.

List of Schemes, Figures and Tables

Scheme S1. Mechanistic Models. (A) Stepwise oxygen insertion model via an epoxyindole intermediate, (B) concerted oxygen insertion mechanisms *via* a dioxetane intermediate



Scheme S2. Chemical Synthesis of 1 and 2

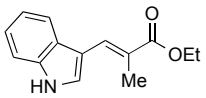


Synthesis and Characterization of Probes 1 and 2

Scheme Methyl-based probe **1** was synthesized from indole-3-carboxaldehyde using a three-step sequence involving Wittig olefination, reduction, and hydrolysis.[1] Hydroxyl-containing tryptophan analog **2** was generated from indole using a $SnCl_4$-mediated epoxide opening followed by saponification.[2]

Reagents were purchased from commercial sources, and used as received. Anhydrous toluene (PhMe) and methanol (MeOH) were dispensed from a solvent purification system, in which the solvent was dried by passage through two columns of activated alumina under argon. A CombiFlash® RF–4x purification system was used for chromatographic purifications. Silica gel was purchased from Sorbent Technologies (cat. #30930M-25, 60 Å, 40–63 μm). Preparative HPLC purification was conducted using an Agilent Preparative System fitted with a Waters T3 C-18 19 mm x 150 mm column. Final compounds were determined to be >95% pure by HPLC analysis using a Waters Analytical System equipped with a Acquity T3 C-18 2.1 mm x 50 mm column.

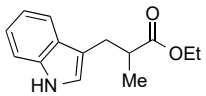
Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on a Bruker 400 AVANCE spectrometer (400 MHz). Chemical shifts (δ) for protons are reported in parts per million (ppm) downfield from tetramethylsilane, and are referenced to the proton resonance of residual protic species in the NMR solvent (e.g. $CHCl_3$ δ = 7.27 ppm). NMR data are represented as follows: chemical shift (ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constant in Hertz (Hz), and integration. Uncorrected melting points were measured on a Thomas Hoover Capillary Melting Point Apparatus.



(*E*)-ethyl 3-(1*H*-indol-3-yl)-2-methylacrylate

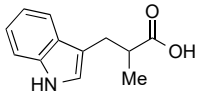
Indole-2-carboxaldehyde (1.45 g, 10.0 mmol) and (carbethoxyethylidene)triphenylphosphorane (5.44 g, 15.0 mmol) were added to a 250 mL round-bottom flask, and the flask was attached to a reflux condenser. The system was evacuated and backfilled with N_2 (3x), PhMe (0.080 L) was injected, and the flask was placed in a 120 °C oil bath. After 4 h, the reaction was cooled to room temperature, and quenched with 1 N HCl (75 mL). The aqueous phase was extracted with EtOAc

(75 mL), and the combined organic layers were washed with H_2O (100 mL) and brine (100 mL). The organic solution was dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. Chromatographic purification (EtOAc/hexanes 1:3) afforded the title compound as a tan solid (2.03 g, 89%). m.p. 133–135 °C (literature[1]: 135–136 °C). The ^{1}H NMR spectrum matched the previously reported data.[1] ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.75 (s, 1 H), 8.11–8.07 (m, 1 H), 7.86–7.81 (m, 1 H), 7.50 (d, J = 2.7 Hz, 1 H), 7.45–7.40 (m, 1 H), 7.32–7.21 (m, 1 H), 4.33 (q, J = 7.1 Hz, 2 H), 2.21 (d, J = 1.3 Hz, 3 H), 1.40 (t, J = 7.1 Hz, 3 H).



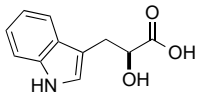
Ethyl 3-(1*H*-indol-3-yl)-2-methylpropanoate

(*E*)-ethyl 3-(1*H*-indol-3-yl)-2-methylacrylate (1.00 g, 4.36 mmol) and Pd/C (10% wt, 185 mg, 0.174 mmol) were added to a 100 mL round-bottom flask, which was sealed with a three-way flushing adaptor. The system was evacuated and backfilled with N_2 (3x), and then MeOH (12 mL) was injected. The flask was evacuated and backfilled with H_2 (3x), and the suspension was stirred under a balloon of H_2 at rt for 16 h. Next, the reaction mixture was passed through celite, and the filter pad was washed with EtOAc (2 x 20 mL). The solvent was removed *in vacuo*, and the resulting oil was filtered through a pad of SiO_2 (EtOAc) to provide the title compound as a brown oil (1.01 g, quantitative). The ^{1}H NMR spectrum matched the previously reported data.[1] ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.01 (s, 1 H), 7.65–7.60 (m, 1 H), 7.36 (dt, J = 8.1, 1.0 Hz, 1 H), 7.20 (ddd, J = 8.2, 7.0, 1.3 Hz, 1 H), 7.14 (ddd, J = 8.0, 7.0, 1.1 Hz, 1 H), 7.01 (d, J = 2.4 Hz, 1 H), 4.12 (q, J = 7.1 Hz, 2 H), 3.25–3.15 (m, 1 H), 2.93–2.80 (m, 2 H), 1.24–1.18 (m, 6 H).



3-(1*H*-indol-3-yl)-2-methylpropanoic acid

Ethyl 3-(1*H*-indol-3-yl)-2-methylpropanoate (0.150 g, 0.649 mmol) and EtOH (2.5 mL) were added to a 25 mL round-bottom flask. 1 N NaOH (2.5 mL) was injected, the flask was attached to a reflux condenser, and the reaction was heated in an 80 °C oil bath for 2 h. The reaction was cooled to room temperature, and the EtOH was removed *in vacuo*. 1 N HCl (20 mL) was added to the mixture, and the aqueous phase was extracted with EtOAc (2 x 20 mL). The combined organic layers were dried over $MgSO_4$, filtered and the solvent was removed *in vacuo* to provide the title compound as a tan solid (105 mg, 80%). m.p. 129–131 °C (literature[1]: 130–131 °C). The ^{1}H NMR spectrum matched the previously reported data.[1] ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.05 (s, 1 H), 10.80 (s, 1 H), 7.53–7.48 (m, 1 H), 7.32 (dt, J = 8.1, 0.9 Hz, 1 H), 7.09 (d, J = 2.4 Hz, 1 H), 7.05 (ddd, J = 8.1, 6.9, 1.2 Hz, 1 H), 6.96 (ddd, J = 8.0, 7.0, 1.1 Hz, 1 H), 3.33 (s, 1 H), 3.07–2.95 (m, 1 H), 2.76–2.63 (m, 2 H), 1.07 (d, J = 6.7 Hz, 3 H).

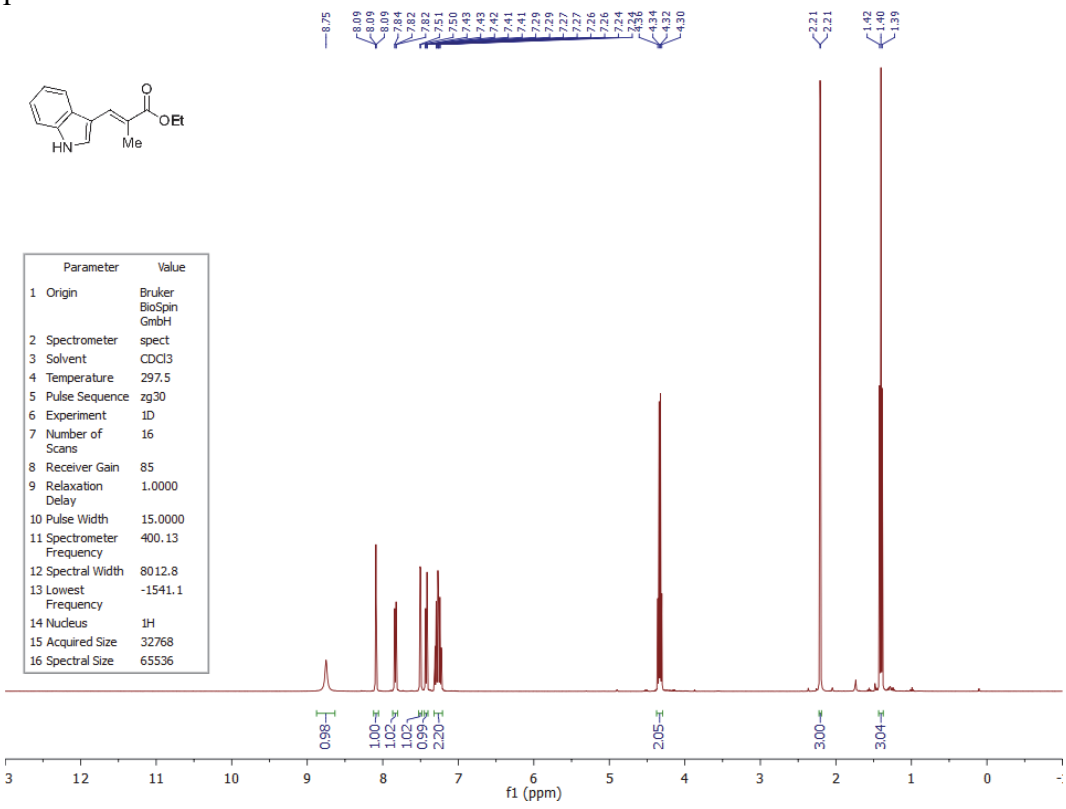


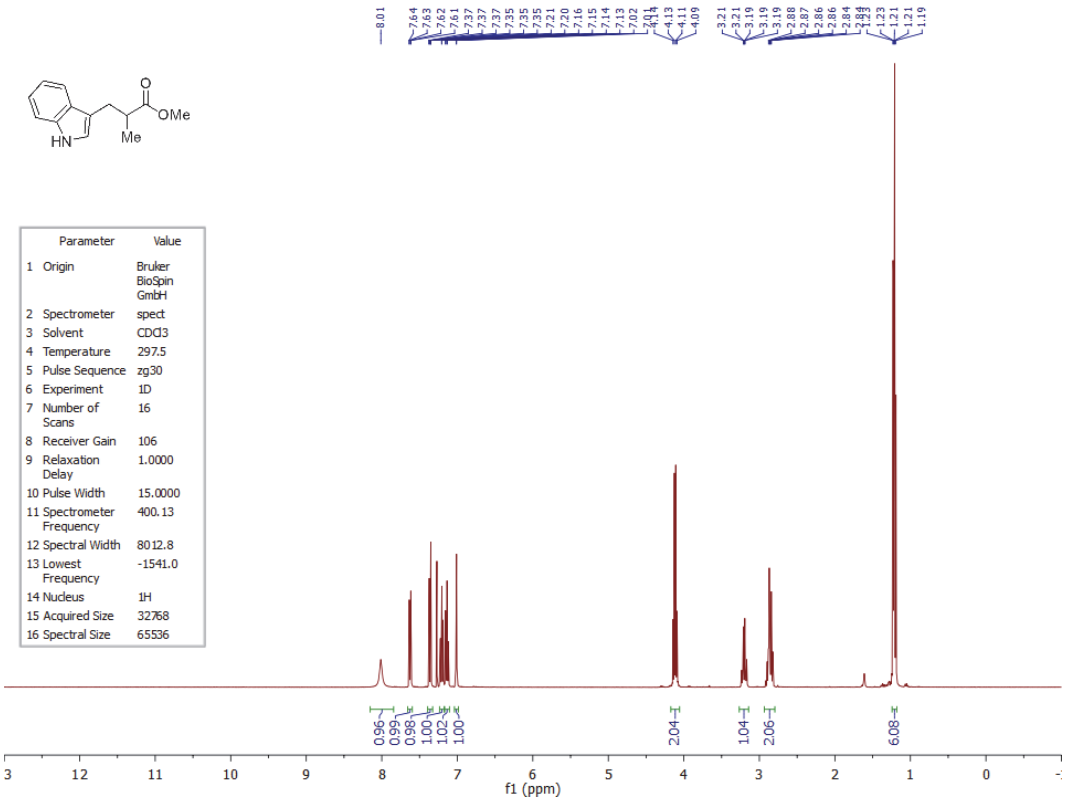
(*S*)-2-hydroxy-3-(1*H*-indol-3-yl)propanoic acid

Indole (1.07 g, 9.10 mmol) was added to an oven-dried 50 mL Schlenk flask, which was sealed with a rubber septum and evacuated and backfilled with N_2 (3x). CCl_4 (8.5 mL) and methyl (2*S*)-glycidate (0.613 mL, 7.00 mmol) were injected, and the flask was placed in a 0 °C ice bath. A

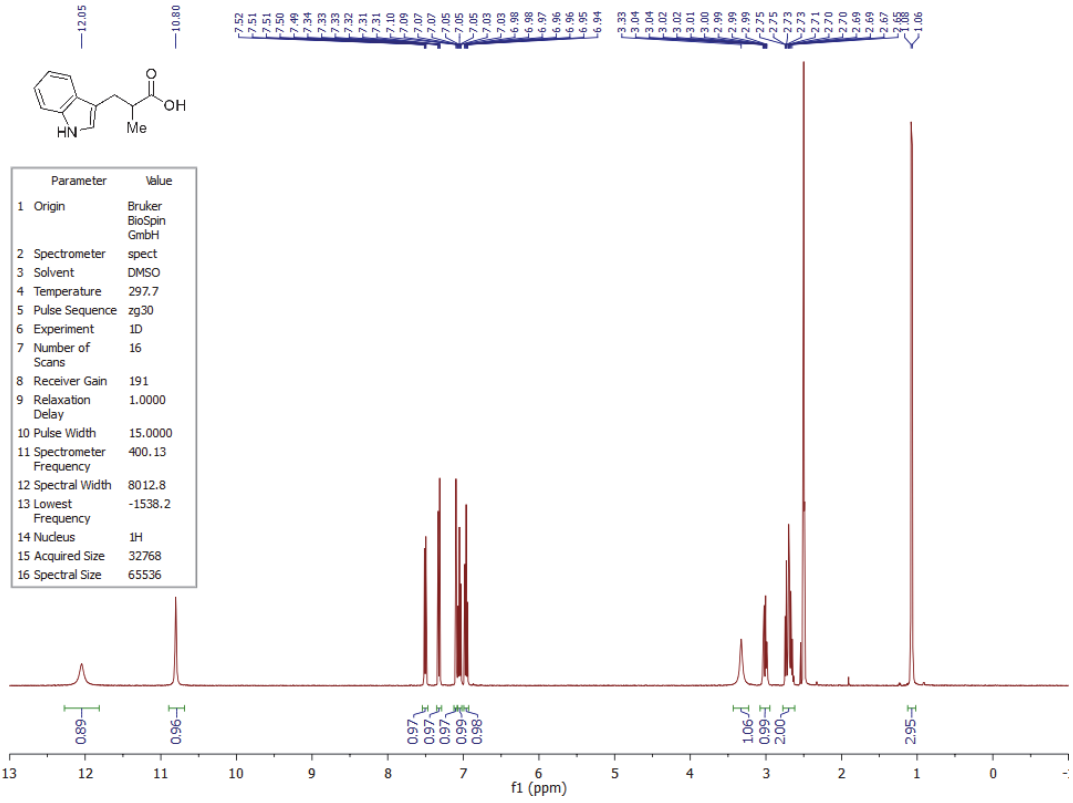
solution of $SnCl_4$ (0.90 mL, 7.7 mmol) in CCl_4 (6.0 mL) was prepared in a separate flask, and transferred to the Schlenk flask *via* syringe over a 5 min period. The resulting orange suspension was allowed to warm to room temperature, and vigorous stirring was continued for 1 h. The reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (40 mL), and extracted with $CHCl_3$ (3 x 40 mL). The organic suspension was washed with brine (100 mL), dried over $MgSO_4$, filtered through a pad of celite, and the solvent was removed *in vacuo*. Chromatographic purification (hexanes : EtOAc, 4 : 1 → 7:3) afforded the methyl ester as a colorless amorphous solid (471 mg). The material was transferred to a 25 mL round-bottom flask, and dissolved in THF:MeOH:H_2O (3 : 1 : 1 mL). The mixture was cooled to 0 °C, and LiOH (0.103 g, 4.30 mmol) was added. The mixture was warmed to room temperature, and stirred for 16 h. The reaction mixture was diluted with Et_2O (50 mL) and acidified with 1 N HCl (50 mL). The aqueous phase was extracted with Et_2O (50 mL), and the combined organic layers were dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. Chromatographic purification (hexanes : EtOAc : AcOH, 74 : 25 : 1), followed by preparative HPLC purification afforded the title compound as a faint red amorphous solid (102 mg, 5% over 2 steps). The ^1H NMR spectrum matched the previously reported data.[2] ^1H NMR (400 MHz, Acetone-d_6) δ 10.03 (s, 1 H), 7.62 (dt, J = 7.9, 1.0 Hz, 1 H), 7.36 (dt, J = 8.2, 0.9 Hz, 1 H), 7.25–7.21 (m, 1 H), 7.08 (ddd, J = 8.2, 7.0, 1.2 Hz, 1 H), 7.00 (ddd, J = 7.9, 7.0, 1.1 Hz, 1 H), 4.45 (dd, J = 7.1, 4.4 Hz, 1 H), 3.27 (ddd, J = 14.7, 4.4, 0.8 Hz, 1 H), 3.10 (ddd, J = 14.7, 7.1, 0.7 Hz, 1 H), 2.54 (s, 1 H).

NMR spectral of **1** and **2** are attached below:

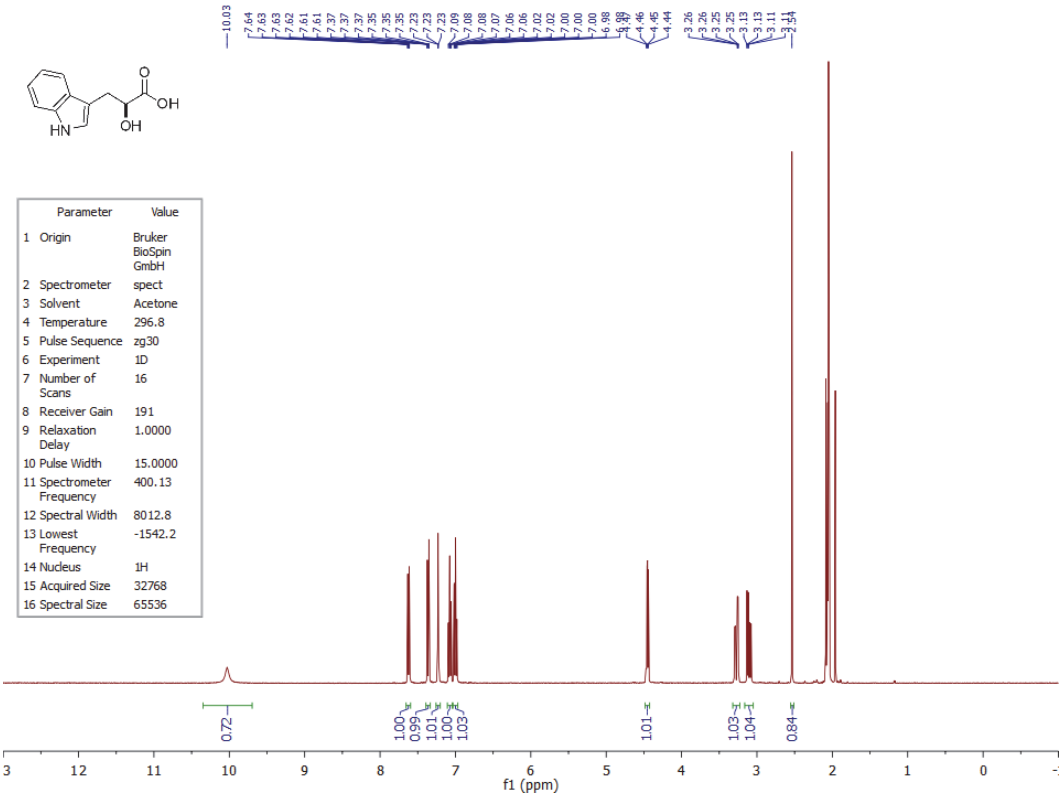




	Parameter	Value
1	Origin	Bruker BioSpin GmbH
2	Spectrometer	spect
3	Solvent	CDCl3
4	Temperature	297.5
5	Pulse Sequence	zg30
6	Experiment	1D
7	Number of Scans	16
8	Receiver Gain	106
9	Relaxation Delay	1.0000
10	Pulse Width	15.0000
11	Spectrometer Frequency	400.13
12	Spectral Width	8012.8
13	Lowest Frequency	-1541.0
14	Nucleus	1H
15	Acquired Size	32768
16	Spectral Size	65536



	Parameter	Value
1	Origin	Bruker BioSpin GmbH
2	Spectrometer	spect
3	Solvent	DMSO
4	Temperature	297.7
5	Pulse Sequence	zg30
6	Experiment	1D
7	Number of Scans	16
8	Receiver Gain	191
9	Relaxation Delay	1.0000
10	Pulse Width	15.0000
11	Spectrometer Frequency	400.13
12	Spectral Width	8012.8
13	Lowest Frequency	-1538.2
14	Nucleus	1H
15	Acquired Size	32768
16	Spectral Size	65536


10.03
7.64
7.63
7.63
7.62
7.62
7.61
7.61
7.37
7.37
7.37
7.35
7.35
7.35
7.23
7.23
7.23
7.09
7.08
7.08
7.07
7.06
7.06
7.02
7.02
7.00
7.00
7.00
6.98
4.47
4.46
4.45
4.44
3.26
3.26
3.25
3.25
3.13
3.13
3.11
3.11
2.94
OH
OH
HN
Parameter Value
1 Origin Bruker BioSpin GmbH
2 Spectrometer spect
3 Solvent Acetone
4 Temperature 296.8
5 Pulse Sequence zg30
6 Experiment 1D
7 Number of Scans 16
8 Receiver Gain 191
9 Relaxation Delay 1.0000
10 Pulse Width 15.0000
11 Spectrometer Frequency 400.13
12 Spectral Width 8012.8
13 Lowest Frequency -1542.2
14 Nucleus 1H
15 Acquired Size 32768
16 Spectral Size 65536
0.72
1.00
0.99
1.01
1.00
1.03
1.01
1.03
1.04
0.84
f1 (ppm)

Identification of the peak at 7.2 min in Figure 4Ab.

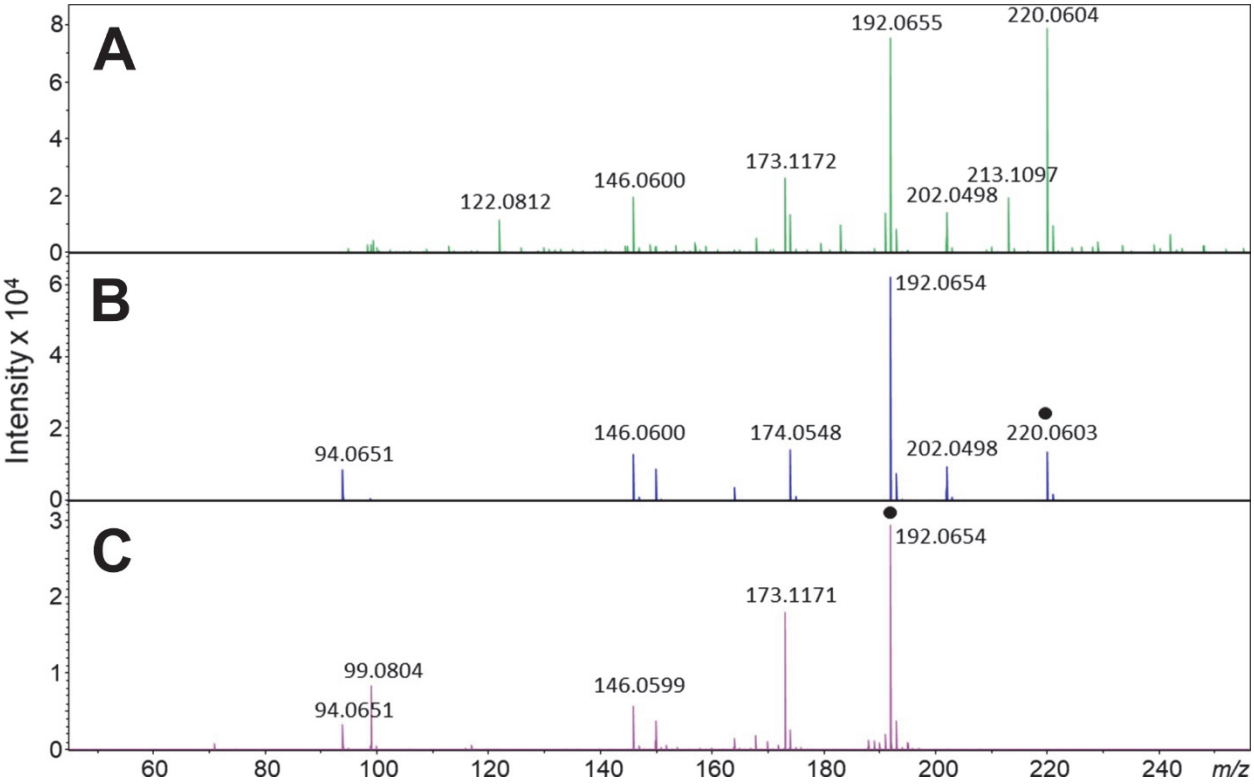
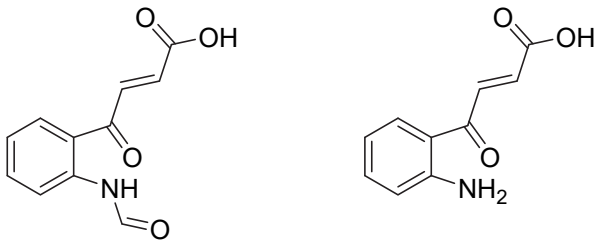


Figure S1. (A) Positive ion ESI mass spectra of the LC fraction of 7.2 min peak in Figure 4Ab. Tandem mass spectra of (B) 220.0604 and (C) 192.0655 as parent ions.

Table S1. Calculated and experimentally measured accurate masses.

Molecular formula	Ion	*m/z* theoretical	*m/z* observed	Mass accuracy (ppm)
$C_{11}H_9NO_4$	$[M+H^+]^+$	220.0604	220.0604	0
$C_{10}H_9NO_3$	$[M+H^+]^+$	192.0655	192.0655	0

Scheme S3. Suggested structure derived from NFK.



$C_{11}H_9NO_4$ $C_{10}H_9NO_3$

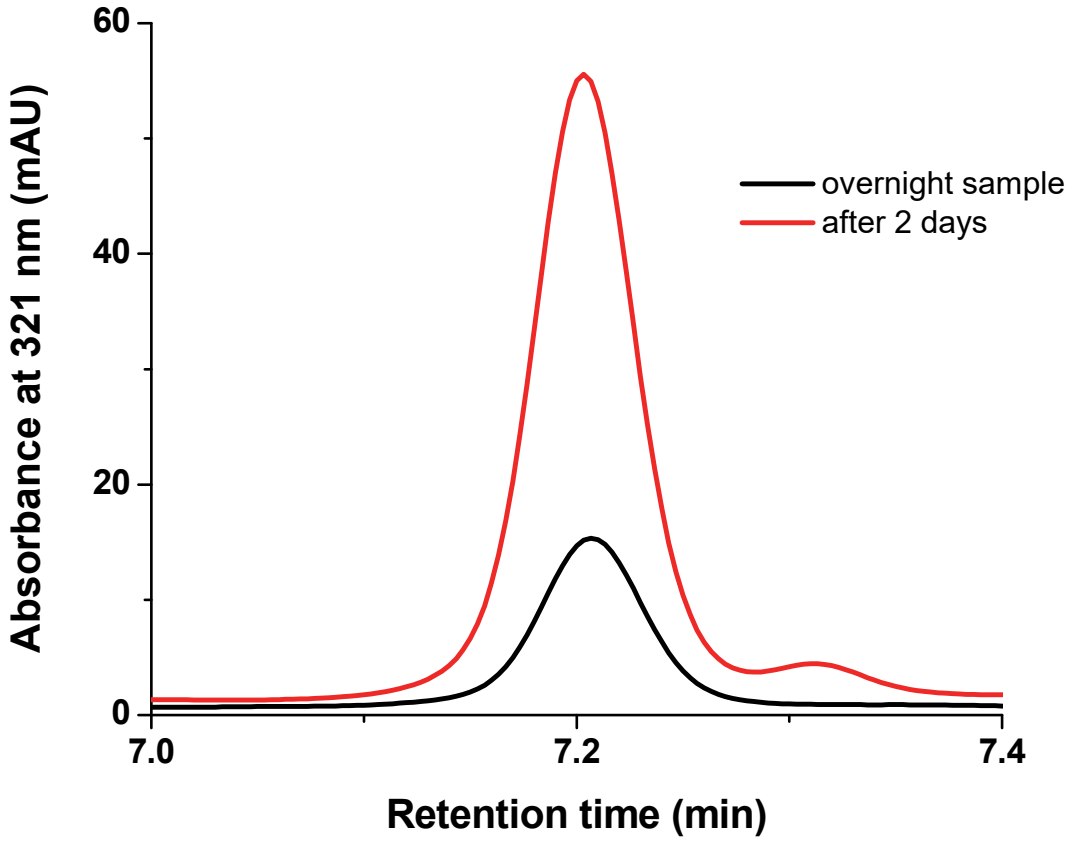
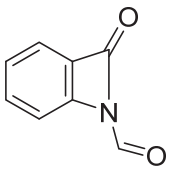


Figure S2. Comparison of peaks at 7.2 min of L-Trp reaction mixtures on HPLC.

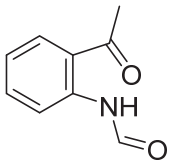
The minor peak at 7.2 min in Figure 4Ab was further analyzed by high resolution mass spectrometry (Figure S1 and Table S1). The ESI mass spectra show 220.0604 and 192.0655 as the major peaks. The 220 amu signal is consistent with NFK less an ammonium group (Scheme S3). Tandem mass on both peaks indicate that *m/z* 192.0655 is a daughter peak from *m/z* 220.0604. Two days after the initial reaction, the same L-Trp reaction mixture filtrate was re-analyzed by HPLC, and the height of the 7.2 min peak increased by ca. 3.8-fold (Figure S2). These observations indicate that the 7.2 min peak is a spontaneous decay product of NFK which accumulates after the TDO-catalyzed reaction.

Scheme S4. Chemical structures for the tandem MS analysis.

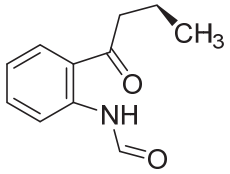
Figure 5 Panel F:



Chemical Formula: $C_8H_5NO_2$
Exact Mass: 147.03

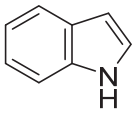


Chemical Formula: $C_9H_9NO_2$
Exact Mass: 163.06

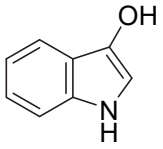


Chemical Formula: $C_{11}H_{13}NO_2$
Exact Mass: 191.09

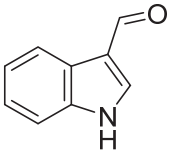
Figure 5 Panel G:



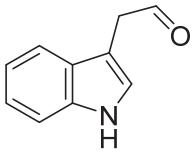
Chemical Formula: C_8H_7N
Exact Mass: 117.06



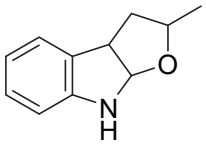
Chemical Formula: C_8H_7NO
Exact Mass: 133.05



Chemical Formula: C_9H_7NO
Exact Mass: 145.05

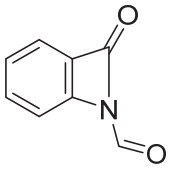


Chemical Formula: $C_{10}H_9NO$
Exact Mass: 159.07

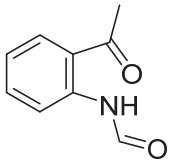


Chemical Formula: $C_{11}H_{13}NO$
Exact Mass: 175.10

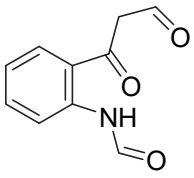
Figure 5 Panel H:



Chemical Formula: $C_8H_5NO_2$
Exact Mass: 147.03

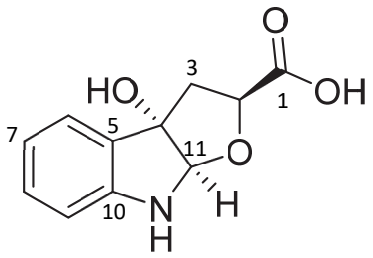


Chemical Formula: $C_9H_9NO_2$
Exact Mass: 163.06



Chemical Formula: $C_{10}H_9NO_3$
Exact Mass: 191.06

Table S2. Summary of the NMR parameters of **2b**.



Atom	^{1}H shift (ppm)	^{1}H mult.	*J* value(s) (Hz)	^{13}C shift (ppm)	^{13}C mult.	HBMC	1D NOESY
1	-	-	-	178.6	s	-	
2	4.15	dd	11.1, 5.6	78.6	s	1, 3	3a, 3b, 11
3a	2.77	dd	12.3, 5.6	43.6	s	4, 5, 11	
3b	2.39	dd	12.3, 11.1	"	s	1, 2, 4, 5	
4	-	-	-	88.2	s	-	
5	-	-	-	130.1	s	-	
6	7.42	dd	7.5, 1.3	124.4	s	4, 5, 8, 10, 11	
7	6.97	t	7.5	120.6	s	5, 6, 8, 9	
8	7.28	td	7.7, 1.4	130.7	s	6, 10	
9	6.81	d	7.9	111.2	s	5, 7, 8	
10	-	-	-	149.4	s	-	
11	5.55	s	-	99.0	s	2, 4, 5, 10	

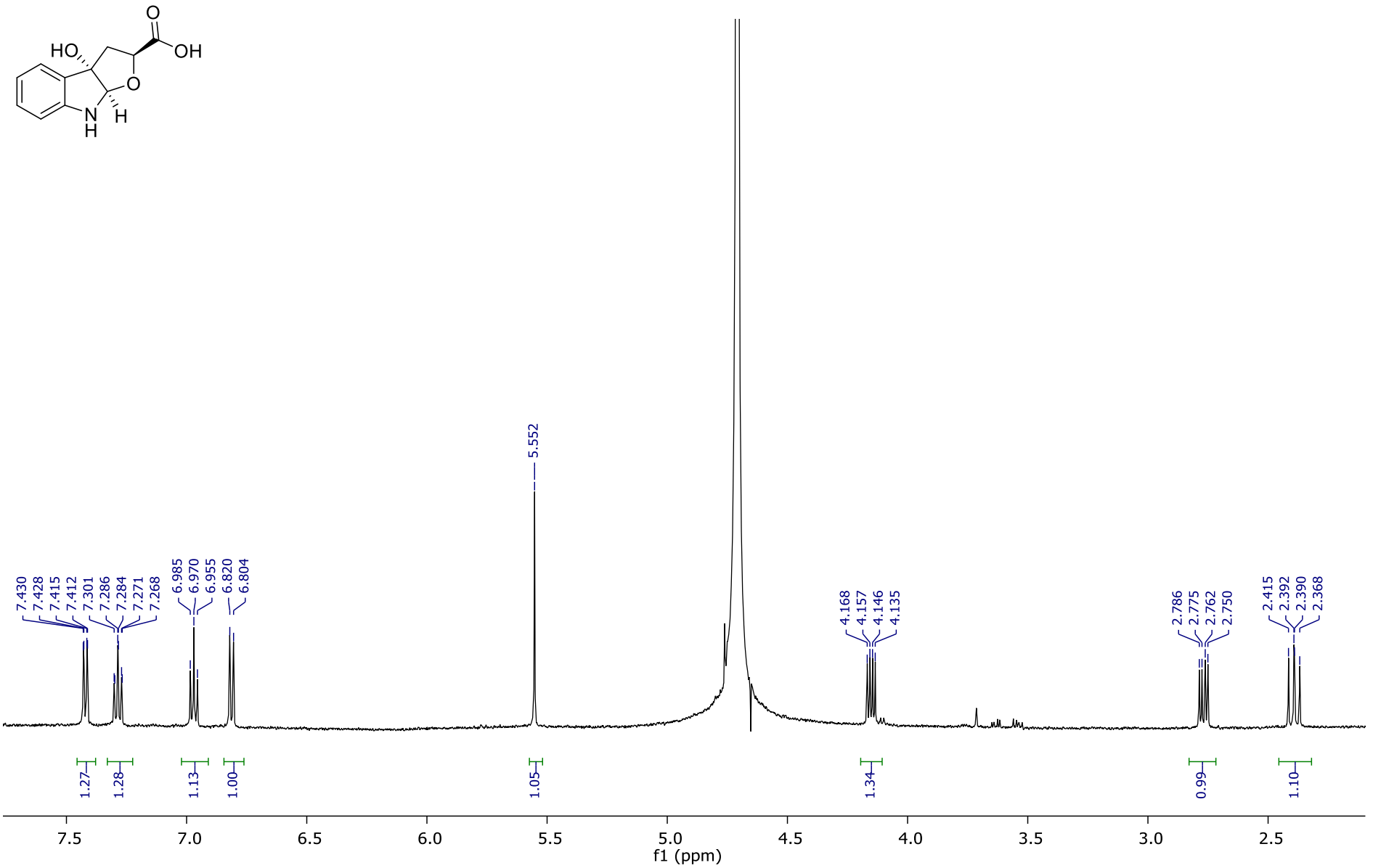


Figure S3. 500 MHz ^{1}H NMR spectrum of monooxygenated product **2b** from the TDO reaction of probe **2** in D_2O.

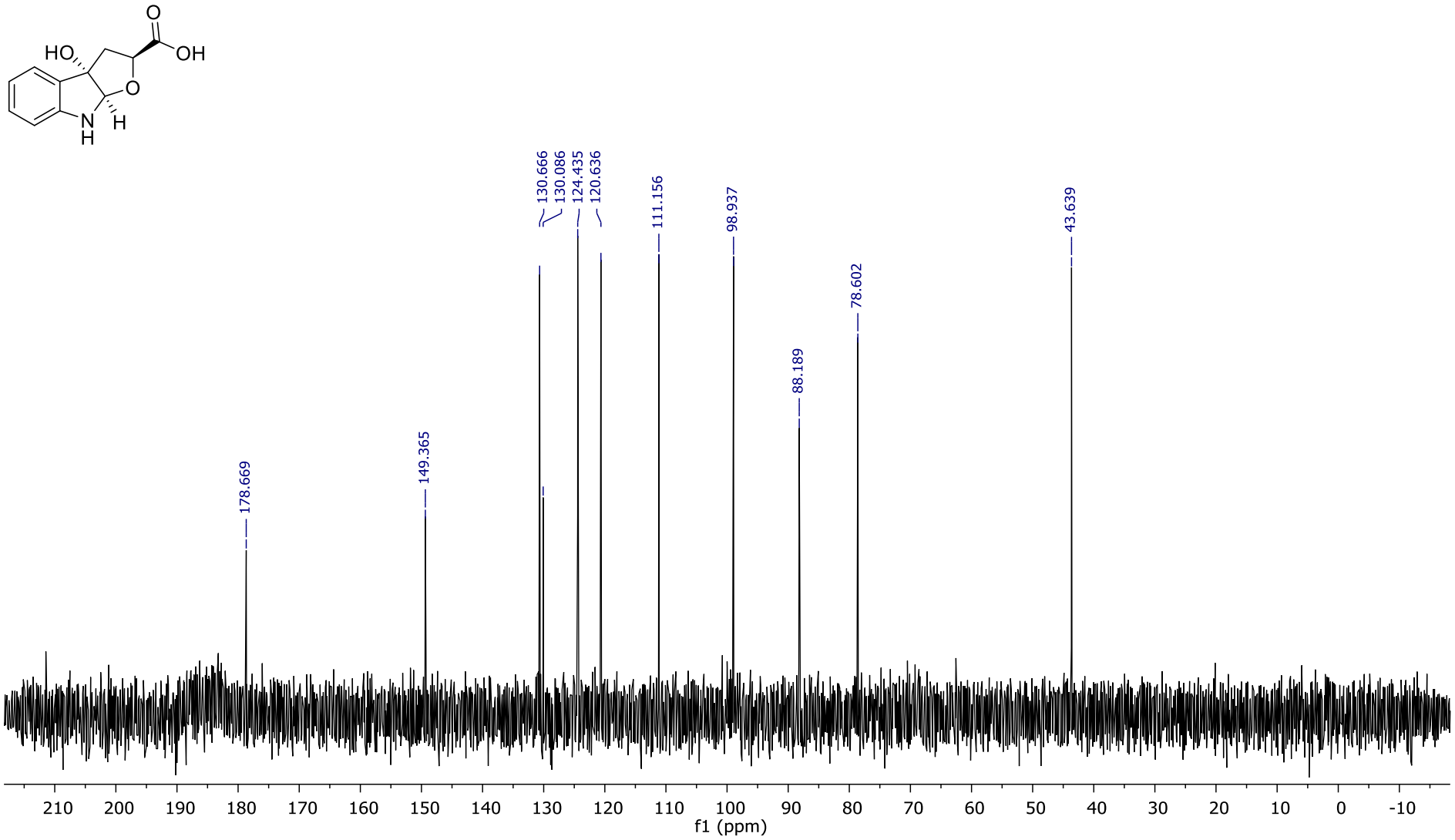


Figure S4. 125 MHz ^{13}C NMR spectrum of monooxygenated product **2b** in D_2O.

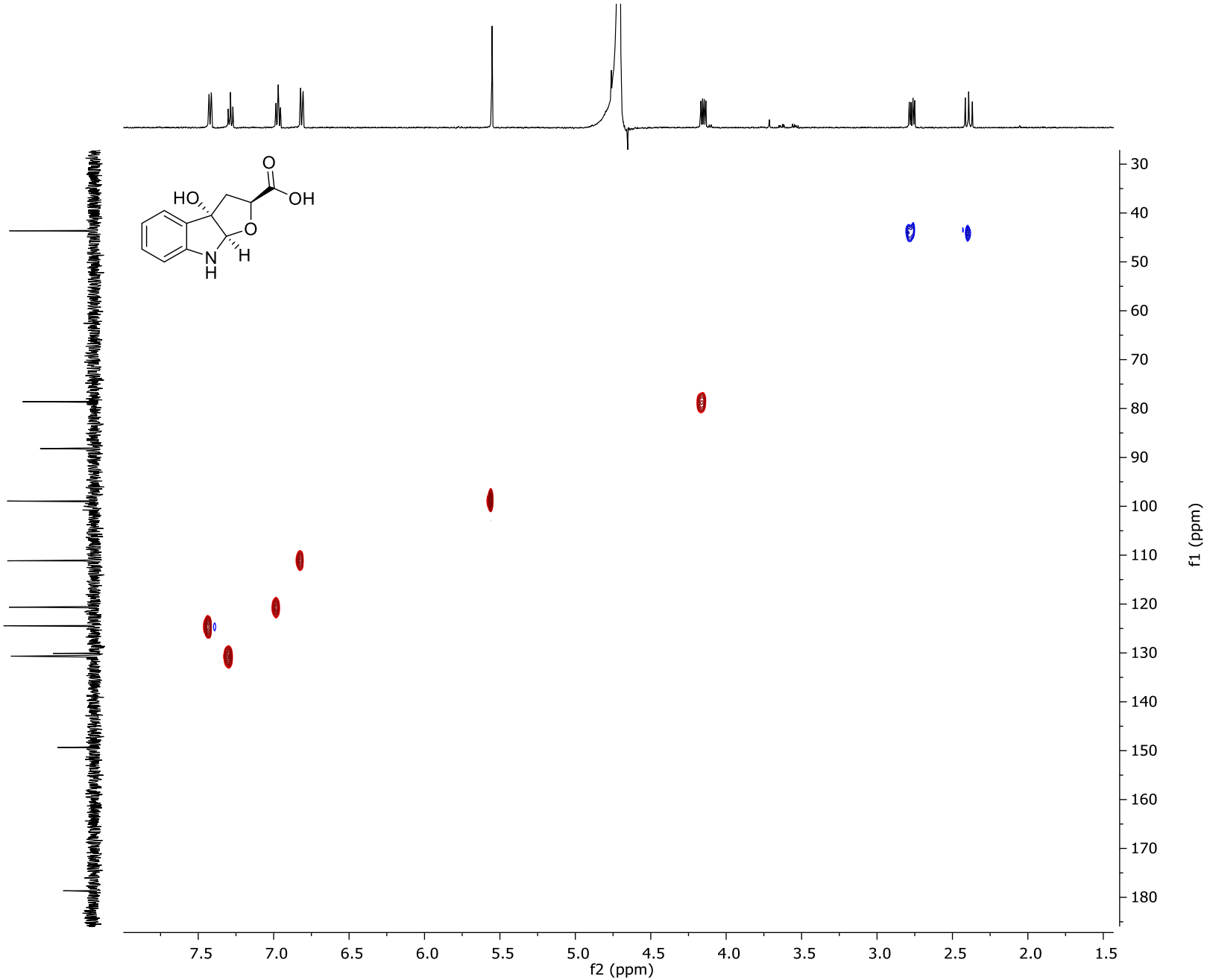


Figure S5. Edited HSQC spectrum of monooxygenated product **2b** in D_2O. Blue signals are CH_2's and red signals are CH's or CH_3's.

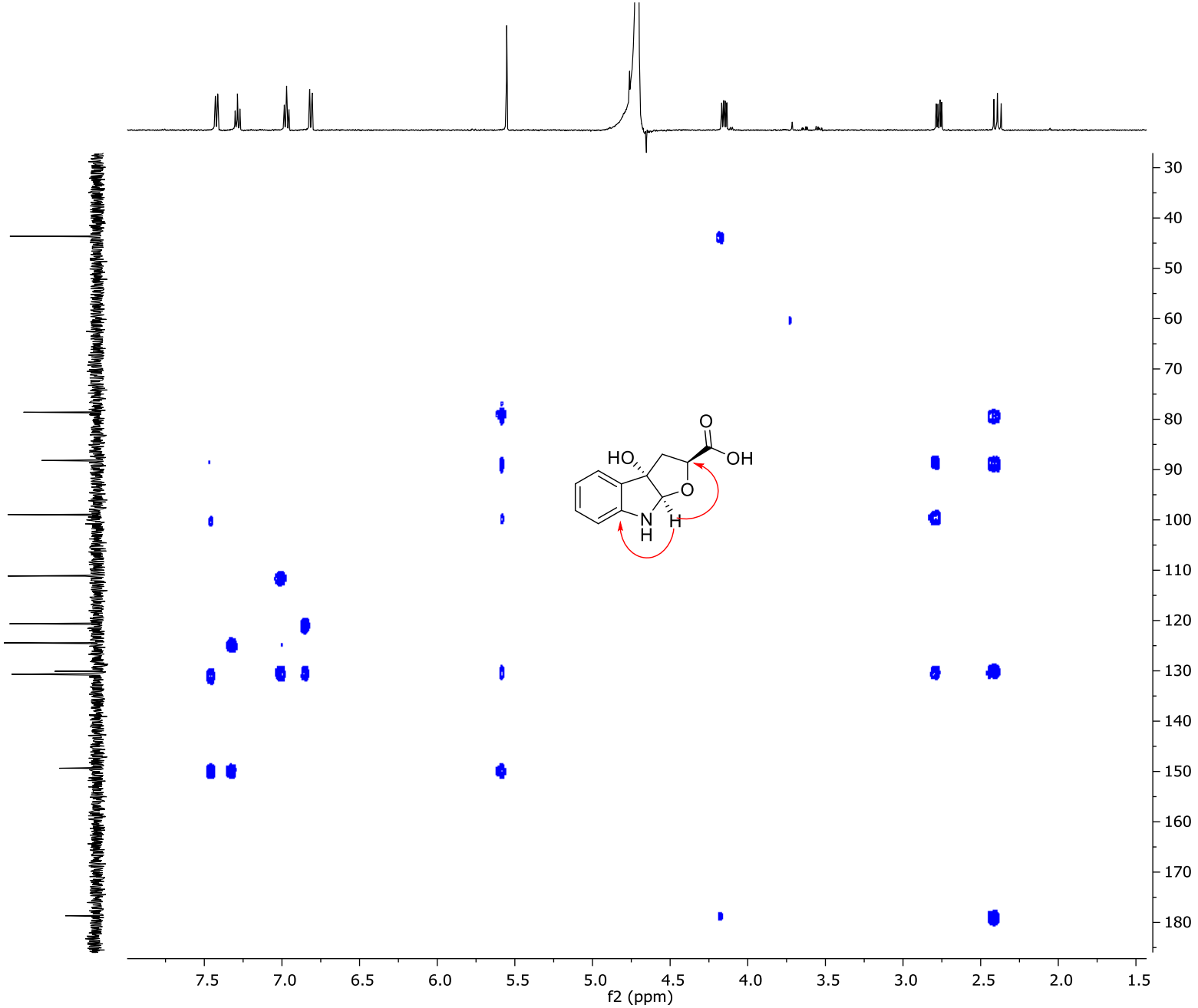


Figure S6. HMBC spectrum of monooxygenated product **2b** in D_2O. Red arrows show key HMBC signals used in the determination of the product structure.

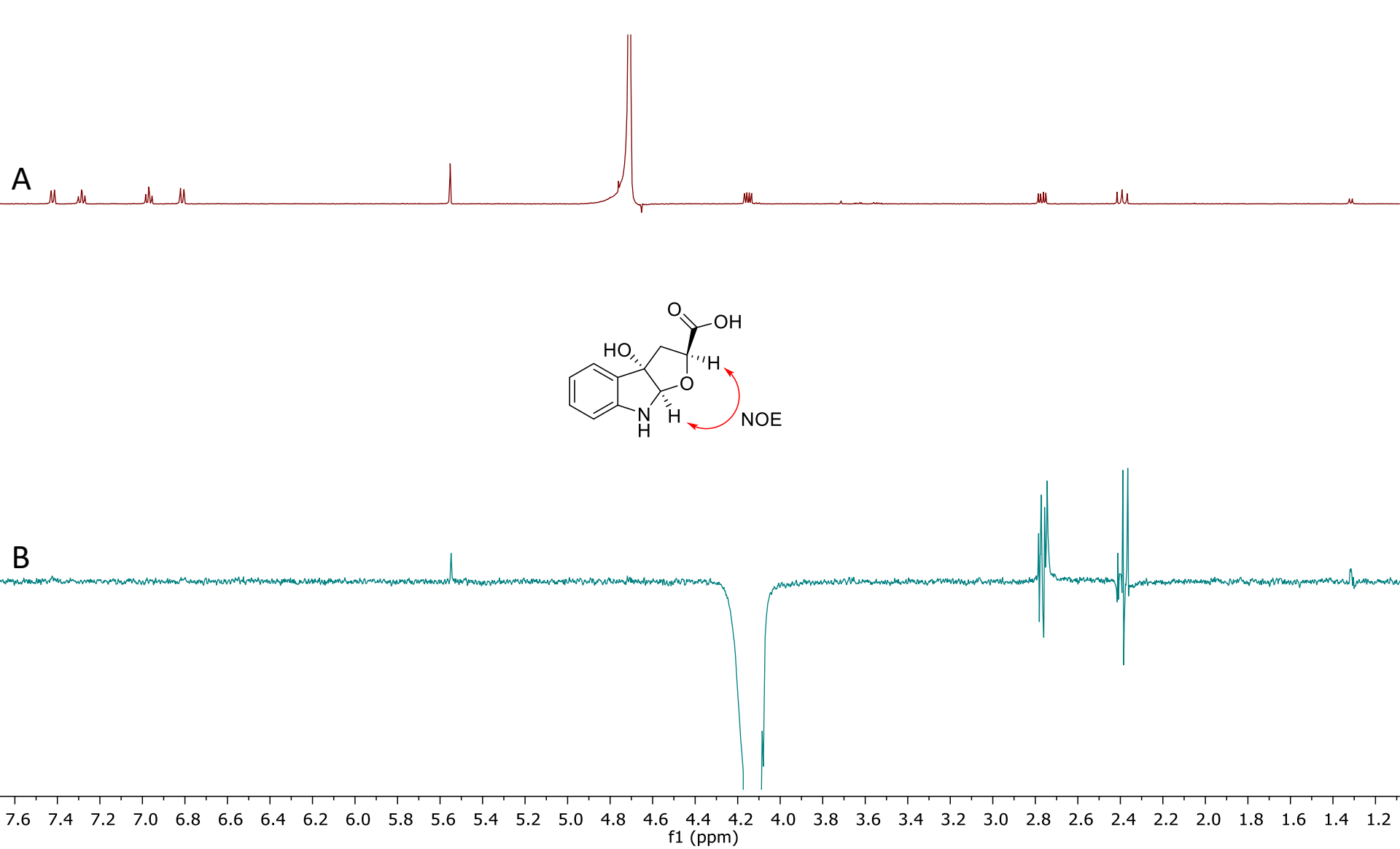
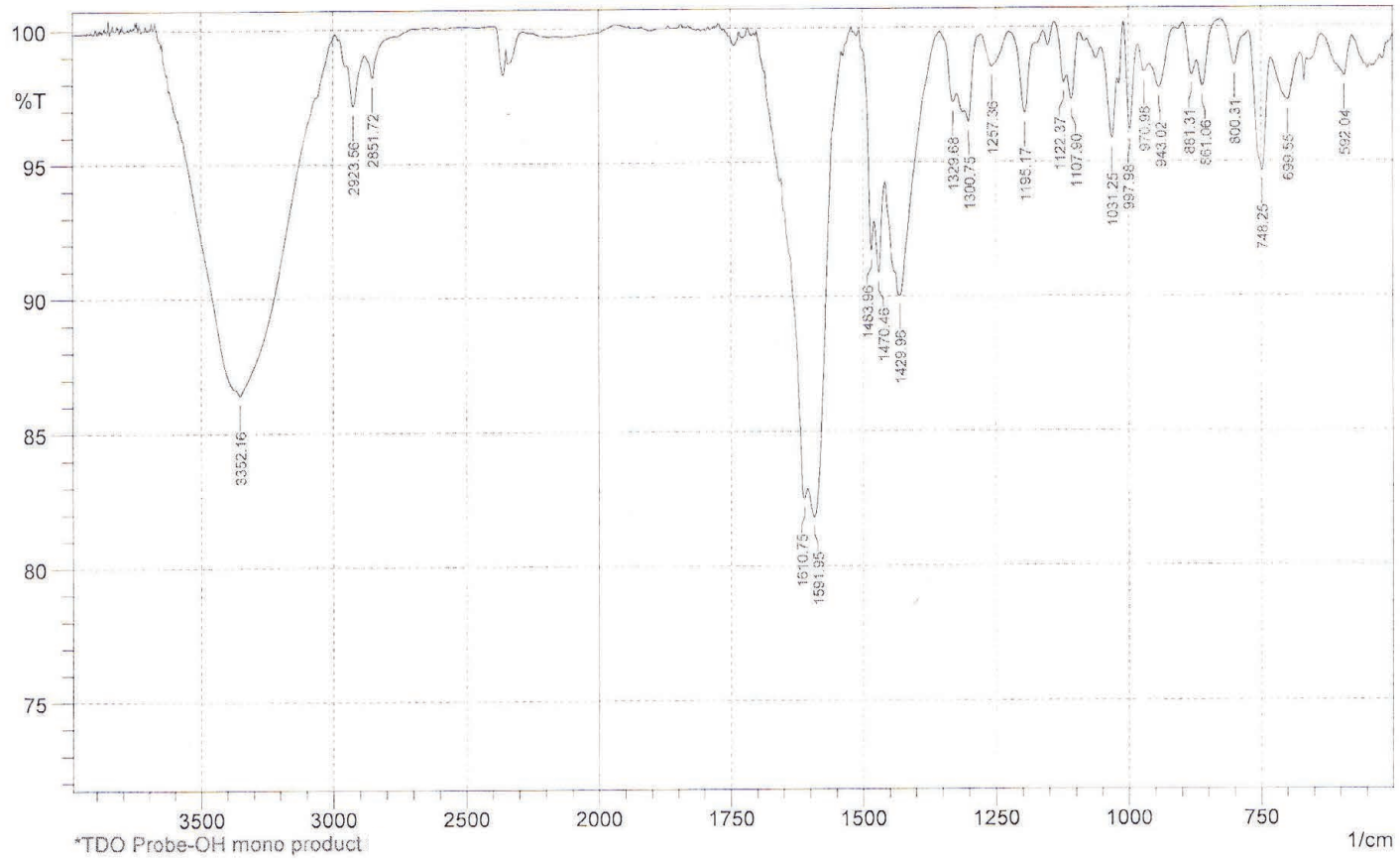


Figure S7. (A) 500 MHz ^{1}H NMR spectrum of monooxygenated product **2b** in D_2O and (B) 1D selective NOESY spectrum in D_2O. Double ended red arrows show key NOE signals used to assign stereochemistry of the newly formed stereocenter.

Figure S8. IR spectrum of the monooxygenated product **2b** of probe **2**.



No.	Peak	Intensity	Corr. Intensity	Base (H)	Base (L)	Area	Corr. Area
1	592.04	98.162	1.494	634.95	577.58	0.272	0.189
2	699.55	97.26	1.65	730.89	675.93	0.471	0.207
3	748.25	94.669	4.551	773.8	731.37	0.565	0.439
4	800.31	98.582	1.457	827.79	774.28	0.114	0.125
5	861.06	97.8	1.294	871.67	827.79	0.157	0.059
6	881.31	98.217	1.019	898.18	871.67	0.12	0.053
7	943.02	97.74	1.353	960.38	913.13	0.269	0.122
8	970.98	98.329	0.613	982.55	960.86	0.128	0.03
9	997.98	96.163	3.65	1009.55	982.55	0.23	0.2
10	1031.25	95.858	2.592	1051.98	1021.61	0.335	0.157
11	1107.9	97.31	1.51	1117.06	1093.92	0.176	0.073
12	1122.37	97.938	0.707	1140.21	1117.06	0.099	0.016
13	1195.17	96.792	2.778	1222.17	1175.4	0.291	0.209
14	1257.36	98.545	1.213	1277.61	1225.06	0.202	0.149
15	1300.75	96.487	1.097	1308.46	1277.61	0.244	0.012
16	1329.68	97.211	0.81	1354.27	1322.93	0.204	0.023
17	1429.96	90.02	1.751	1438.15	1354.75	1.745	0.157
18	1470.46	90.865	2.536	1478.65	1458.4	0.693	0.108
19	1483.96	91.71	2.327	1507.1	1479.13	0.47	0.07
20	1591.95	81.806	4.237	1604.48	1540.36	3.171	0.565
21	1610.75	82.479	1.757	1648.36	1604.97	2.471	0.106
22	2851.72	98.195	1.036	2879.68	2769.28	0.373	0.103
23	2923.56	97.168	1.646	2946.7	2880.17	0.537	0.209
24	3352.16	86.401	0.832	3369.03	3064.33	12.561	1.424

References

1. Matsuda, K.; Toyoda, H.; Nishio, H.; Nishida, T.; Dohgo, M.; Bingo, M.; Matsuda, Y.; Yoshida, S.; Harada, S.; Tanaka, H.; Komai, K.; Ouchi, S. *J. Agr. Food Chem.* **1998**, *46*, 4416-4419.
2. Deechongkit, S.; You, S. L.; Kelly, J. W. *Org. Lett.* **2004**, *6*, 497-500.