

Supporting Information

for

Two novel flavin-dependent halogenases from the bacterial consortia of *Botryococcus braunii* catalyze mono- and di-bromination

Pia R. Neubauer¹, Olga Blifernez-Klassen², Lara Pfaff¹, Mohamed Ismail¹, Olaf Kruse² and Norbert Sewald^{1,*}

¹ Organic and Bioorganic Chemistry, Faculty of Chemistry and Center for Biotechnology (CeBiTec), Bielefeld University, D-33615 Bielefeld, Germany

² Algae Biotechnology and Bioenergy, Faculty of Biology and Center for Biotechnology (CeBiTec), Bielefeld University, D-33615 Bielefeld, Germany

* Correspondence: norbert.sewald@uni-bielefeld.de; Tel.: +49 521 106 2051

Table of Contents

General Information	3
Amino acid sequences of the novel, identified F-Hals SpH1 and SpH2	6
SDS-PAGE with heterologously expressed SpH1 and SpH2	6
Enzymatic bromination by SpH1 and SpH2 in presence of catalase	7
Absorption bands of FAD bound in SpH1 and SpH2 during FAD reconstitution	7
HPLC trace of the enzyme reaction with SpH1_K81A and indole	8
Characterization of the compounds brominated by SpH1 and SpH2 by mass spectrometry	8
Determination of regioselective halogenation site <i>via</i> NMR analysis	13
List of detected secondary metabolite biosynthetic gene clusters within the MAG 21	45
Area and volume of the active site pocket for Apo SpH2 and SpH1 and docked with indole and 3-bromoindole before and after MD simulation	48
References	49

General Information

Analytical RP-HPLC

For analytical RP-HPLC analysis a Shimadzu Nexera XR chromatography system was conducted (Luna® C18(2) column (3 μ M, 100 Å, LC column, 100 x 2 mm) from Phenomenex®). Enzymatic conversions from substrate to brominated product were calculated *via* integration of the resulting peaks in the RP-HPLC chromatograms.

Eluent A: H₂O/CH₃CN/TFA= 95:5:0.1

Eluent B: H₂O/CH₃CN/TFA= 5:95:0.1

Flow rate: 650 μ L/min

0 min	95% A	5% B
5.5 min	5% A	95% B
6 min	5% A	95% B
6.1 min	95% A	5% B
9 min	95% A	5% B

LC-ESI-MS

LC-ESI-MS was performed using Agilent Technologies 1200 with C18- Hypersil Gold C18 column (3 μ M, 150 x 2,1 mm; Thermo Fisher Scientific) and a mass spectrometer Agilent Technologies 6220 Accurate Mass TOF/ LC-MS (gas temperature: 325°C, capillary tension: 2500 V, fragmentary tension: 175 V) with Dual-ESI (spray voltage: 2,5 kV).

Eluent A: H₂O/CH₃CN/formic acid = 95:5:0.1

Eluent B: H₂O/CH₃CN/formic acid = 5:95:0.1

Flow rate: 300 μ L/min

0 min	100% A	0% B
10 min	2% A	98% B
11 min	2% A	98% B
11.5 min	100% A	0% B
15 min	100% A	0% B

ESI-MS

Nano ESI-mass spectrometry was recorded on Esquire 3000 Ion trap mass spectrometer from Bruker Daltonik GmbH. Nitrogen serves as atomizer gas as well as dry gas carried out by Bruker—nitrogen generator NGM 11. Helium was employed for cooling of ion trap and collision gas for MS experiments. Measurements were carried out in negative as well as positive mode.

GC-MS

For GC-MS measurements samples were 3 times extracted with ethyl acetate (v/v 1:1) and the organic phase was evaporated for a final volume of 100 μ L. For the analytics a Trace GC Ultra gas chromatography from ThermoScientific with VF-5 column (30 m \times 0.25 mm, 5% diphenylsiloxane, 95% dimethylsiloxane) was utilised. Helium serves as mobile phase and a temperature gradient of 5°C/min from 80°C to 325°C. The mass spectrometer ITQ900 of ThermoFinnigan recorded 20 measurements per min (50-750 *m/z*).

HRMS

HRMS measurements were recorded on Agilent 6200 accurate mass TOF MS. Samples were injected through an Agilent 1200 LC system, Hypersil Gold C18 (50 mm \times 2.1 mm, 1.9 μ m) column. Measurements were performed as described above the LC-ESI-MS measurements. Agilent tuning mix was utilized for external calibration and was performed before measurements.

NMR

NMR spectra were recorded on a Bruker Avance 500 (^1H : 500 MHz, ^{13}C : 126 MHz) or a Bruker DRX-500 spectrometer (^1H : 500 MHz, ^{13}C : 126 MHz). Chemical shifts are reported relative to residual solvent peaks in parts per million (ppm) (DMSO-*d*₆: ^1H : 2.5 ppm; ^{13}C : 39.5 ppm). Coupling constants (J) are reported in Hz with the following abbreviations for indication of the splitting: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad signal. Declared yields were determined at purified, isolated products.

SDS-PAGE

A discontinuous gel system with 4% stacking gel and 15% separating gel described by Laemmli was carried out.^[1] Prestained protein ladder from NEB Inc. (11-245 kDa) was used.

Site-directed mutagenesis of *spH1*

For the site-directed mutagenesis a PCR was carried out to mutate the conserved lysine residue (K85) to an alanine. For the PCR Primers were designed:

SpH1_K85A: ATGTGAAGCCACCGCAGCGCTGGGTATTCGTTTT

SpH1_K85A_rev: AAAACGAATACCCAGCGCTGCGGTGGCTTCACAT

70 ng/ μ L of vector DNA (*spH1_pET28-a*), 125 ng of each primer, 2.5 U/mL Pfu polymerase (Promega, Germany), dNTP mix (200 μ M of each nucleotide) and 1x reaction buffer were added in PCR solution. The PCR program is shown in the following table:

Cycles	Temperature [°C]	Minutes
1	95	0.5
16	95	0.5
	55	1
	68	7
1	68	5

Parental DNA was digested with DpnI restriction enzyme (10 U/ μ L) for 1h at 37°C and transformed in *E. coli* DH5 α . The following protein expression and purification was identical to non-mutated SpH1.

Abbreviations

Br: mono-brominated product

Di-Br: di-brominated product

ESI-MS: electrospray ionisation - mass spectrometry

GC: gas chromatography

HRMS: high resolution mass spectrometry

LC: liquid chromatography

NMR: nuclear magnetic resonance spectroscopy

RP-HPLC: reverse-phase high performance liquid chromatography

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis

Amino acid sequences of the novel, identified F-Hals SpH1 and SpH2

>SpH1

MKHGGMASGAIERVVIVGGGTAGWMAAAALGAYLAGAGTRITLIESEIGTIGVGEATIP
TIRRFYAALGMTDAEVMRACEATAKLGIRFVDWKPGETSFVHPFGRFGQDLRGIDFHQYW
QKARQAGRAAPLEEYSLGAMLAREGHASVPLPNPPSLSVFDWALHFDAARFAAHMRA
YAERAGVARIDARITDVTLHGETGFIEAVTLDSGERVAGDLFVDCSGFRGLLIGEALGTPYR
DWSHWLLCDGAFVQSERVGEPPSCTTARTAGWQWRIPRSREGNGLVFASDFQGGD
EARAELLANIPGAPTMEPRRLRFTPGRRVAVARNVCVSLGLASGFLEPLESTSIALIETGIER
LKQLFPDKGFDPAVIAEYNAQSAEEMERVDFILLHYHLSTRDGPFWQACREMTLPDSL
AKLELWRARGAFLRYRWEMFSPASWLAITYGGFEHLPERLDPGVAAVGTDELAQGLEQMR
AAIARTVADTPHSEFLATVDGAAAEPMRTSA*

>SpH2

MNNPDALRKVCIVGGGTAGWIAAAVMAHHMKGRLFEIELVESDDIGTIGVGESTIPPFME
LIAKLDIDEQDFVRQVQASFKLGIKFDWFRKDESYFHPFGDVGQRVELSDFYQCWRKATL
AGHDFPLQDFAPATVMAHAGRFMLPFKARNTPIAGASYALHVDKRVAQFLRGHAEAR
GVKRTGIVTDVALDARGFVSTLTLKDGRTIDAFFIDCSGFRALLIEKALGSGYTDWSNYL
FCDRAIAAQTEENVGPPRPYTLAEAQDAGWRWRIPLQHRTGNGHVYSSEFMSDEEATRILL
EKVEGEVVAGPMVVPFRTGVREKIWHRNVLAIGLSSGFIEPLESTAIHLIYRGMDFFFRFLP
DRHCDPALADEYNRRMVADYTEIRDFIVLHYCTTARDDTPFWRKCRDMDWPDSLKERVE
LFRSNGTLREGLDELFRSVSWFVLEGMGIRPRTYHPLVDRIDEAMLYAGLDKAREQLAGF
VRQLPTQQQFIDAHCRAEKVDLTRPVMAGV*

SDS-PAGE with heterologously expressed SpH1 and SpH2

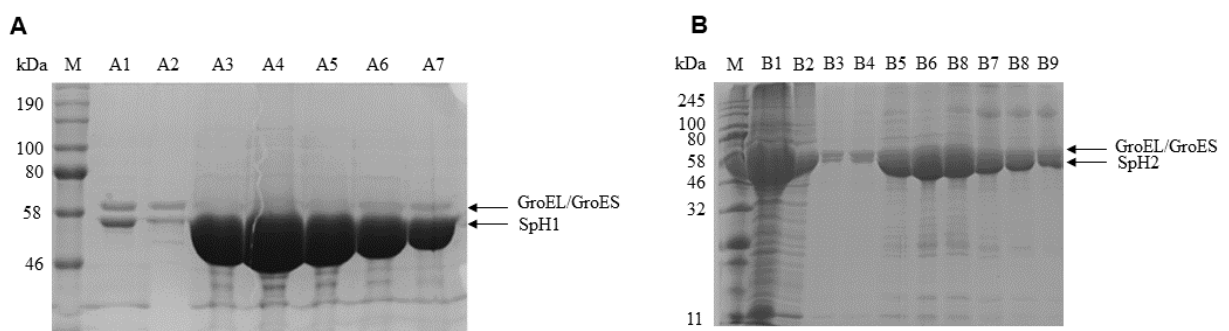


Figure 1: SDS-PAGE of SpH1 (A) and SpH2 (B). A1-A7: Eluted fractions of SpH1 (55.82 kDa) during protein expression; B1-B9: Eluted fractions of SpH2 (58.18 kDa) during protein purification; B1: Flow through, B2-3: Wash steps, B4-B9: Eluted fractions. 1.5 L cultivation resulted in an average rate of 3 mg/mL enzyme.

Enzymatic bromination by SpH1 and SpH2 in presence of catalase

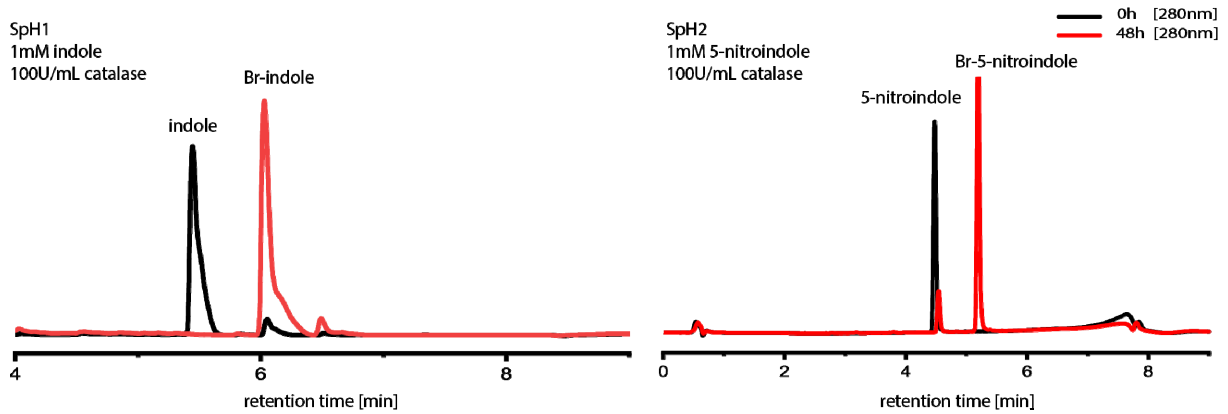


Figure 2: RP-HPLC traces (280nm) of the substrates after 0h and 48h of incubation with the enzymes, SpH1 and SpH2 in presence of catalase. The decomposition of hydrogen peroxide by catalase did not effect the enzymes, SpH1 and

Absorption bands of FAD bound in SpH1 and SpH2 during FAD reconstitution

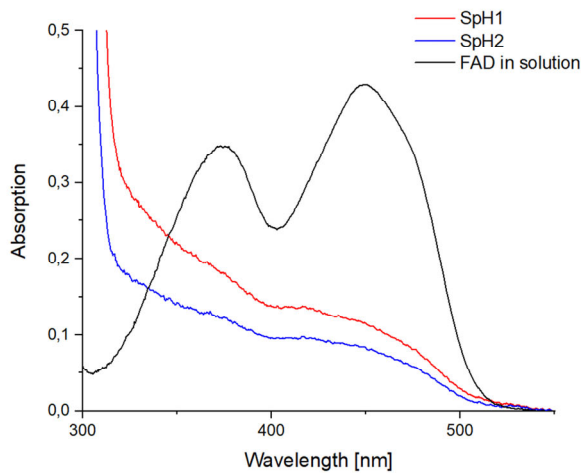


Figure 3: FAD absorption in solution is shifted compared to FAD absorption within the protein, which is caused by the changed environment and can therefore be used to show that FAD is in the protein and not in solution. This was measured for SpH1 and SpH2 and showed a slight shift at 450 nm to around 448 nm and at 373 nm to around 360 nm.

HPLC trace of the enzyme reaction with SpH1_K81A and indole

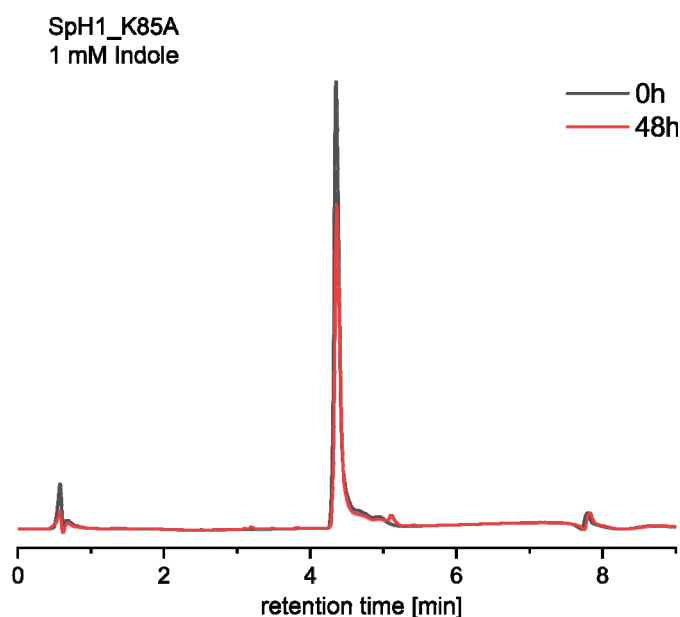


Figure 4: HPLC trace of the enzyme reaction of SpH1_K81A with indole shows that the mutant without the conserved lysine residue is inactivated. This leads to the conclusion that the halogenating mechanism of SpH1 is dependent on the conserved lysine, which is essential for the flavin-dependent halogenating mechanism.

Characterization of the compounds brominated by SpH1 and SpH2 by mass spectrometry

3-Bromoindole 1

theoretical mass [C_8H_6NBr] m/z 194.97, 196.97

Dibromoindole: theoretical mass [$C_8H_5NBr_2$] m/z 274.88, 272.88, 276.87

SpH1: Conversion (RP-HPLC): 100%; ESI-MS [$M+H$]⁺: 195.55, 197.53

SpH2: Conversion (RP-HPLC): Br: 15%; Di-Br: 85%; Br: LC-ESI-MS [$M-H$]⁻: 193.96; 195.95; 35%

Di-Br: LC-ESI-MS [$M-H$]⁻: 271.87; 273.86; 275.86

3-Bromo-2-methylindole 2

theoretical mass [C_9H_8BrN] m/z 208.98, 210.98

Dibromo-2-methylindole: theoretical mass [$C_9H_7Br_2N$] m/z 288.89, 286.89, 290.89

SpH1: Conversion (RP-HPLC): 100%; LC-ESI-MS [$M-H$]⁻: 207.98, 209.97

SpH2: Conversion (RP-HPLC): Br: 10%; Di-Br: 90%; Br: LC-ESI-MS [M-H]⁻: 207.97, 209.97; Di-Br: LC-ESI-MS [M-H]⁻: 285.88, 287.98, 289.88

2-Bromo-3-methylindole 3

theoretical mass [C₉H₈BrN] *m/z* 208.98, 210.98

Dibromo-3-methylindole: theoretical mass [C₉H₅Br₂N] *m/z* 288.89, 286.89, 290.89

SpH1: Conversion (RP-HPLC): 98%; LC-ESI-MS [M-H]⁻: 207.98, 209.88

SpH2: Conversion (RP-HPLC): 35%; LC-ESI-MS [M-H]⁻: 208.02, 210.02

3-Bromo-5-methylindole 4

theoretical mass [C₉H₈BrN] *m/z* 208.98, 210.98

Dibromo-5-methylindole: theoretical mass [C₉H₅Br₂N] *m/z* 288.89, 286.89, 290.89

SpH1: Conversion (RP-HPLC): 98%, LC-ESI-MS [M-H]⁻: 207.98, 209.98

SpH2: Conversion (RP-HPLC): Br: 2%; Di-Br: 98%; Di-Br: LC-ESI-MS [M-H]⁻: 285.88, 287.88, 289.88

3-Bromo-5-nitroindole 5

theoretical mass [C₈H₅BrN₂O₂] *m/z* 239.95, 241.95

SpH1: Conversion (RP-HPLC): 97%; ESI-MS [M-H]⁻: 238.95, 240.95

SpH2: Conversion (RP-HPLC): 98%; LC-ESI-MS [M-H]⁻: 238.95, 240.94

3-Bromo-5-chloroindole 6

theoretical mass [C₈H₅BrClN] *m/z* 228.93, 230.93

SpH1: Conversion (RP-HPLC): 43%; LC-ESI-MS [M-H]⁻: 227.94, 229.95

3-Bromo-5-bromoindole 7

theoretical mass [C₈H₅Br₂N] *m/z* 274.88, 272.88

SpH1: Conversion (RP-HPLC): 55%; LC-ESI-MS [M+H]⁺: 275.88, 273.88

SpH2: Conversion (RP-HPLC): 96%; LC-ESI-MS [M+H]⁺: 273.8, 271.8

3-Bromo-5-fluoroindole 8

theoretical mass [C₈H₅BrFN] *m/z* 212.96, 214.96

Dibromo-5-fluoroindole: theoretical mass [C₈H₄Br₂FN] *m/z* 292.87, 290.87, 294.87

SpH1: Conversion (RP-HPLC): 100%; LC-ESI-MS [M-H]⁻: 211.93, 213.93

SpH2: Conversion (RP-HPLC): Br: 59%; Di-Br: 44%; Br: LC-ESI-MS [M-H]⁻: 212.0, 214.0; Di-Br: LC-ESI-MS [M-H]⁻: 289.93, 291.94, 293.92

3-Bromo-5-cyanoindole 9

theoretical mass [C₉H₅BrN₂] *m/z* 219.96, 221.96

SpH1: Conversion (RP-HPLC): 99%; LC-ESI-MS [M-H]⁻: 218.96, 220.96

SpH2: Conversion (RP-HPLC): 11%; LC-ESI-MS [M-H]⁻: 219.0, 221.0

Bromo-5-hydroxyindole

theoretical mass [C₈H₆BrNO] *m/z* 210.96, 212.96

SpH1: Conversion (RP-HPLC): 94%, LC-ESI-MS [M-H]⁻: 209.96, 211.96

Bromoindole-5-carboxylic acid

theoretical mass [C₉H₆BrNO₂] *m/z* 238.96; 240.96

SpH1: Conversion (RP-HPLC): 20%; LC-ESI-MS [M-H]⁻: 237.95; 239.95

Bromoindole-2-carboxylic acid

theoretical mass [C₉H₆BrNO₂] *m/z* 238.96, 240.96

SpH1: Conversion (RP-HPLC): 9%; LC-ESI-MS [M-H]⁻: 237.96, 239.96

Bromoindole-2-methanol

theoretical mass [C₉H₈BrNO] *m/z* 224.98, 226.98

SpH1: Conversion (RP-HPLC): 99%; LC-ESI-MS [M-H]⁻: 223.99, 225.98

Bromo-2-methyl-5-nitroindole

theoretical mass [C₉H₇BrN₂O₂] *m/z* 253.97, 255.97

SpH1: Conversion (RP-HPLC): 28% + 5 side products; LC-ESI-MS [M-H]⁻: 252.97, 254.96

SpH2: Conversion (RP-HPLC): 74%; LC-ESI-MS [M-H]⁻: 252.96, 254.96

Bromo-7-azaindole 15

theoretical mass [C₇H₅BrN₂] *m/z* 195.96, 197.96

SpH1: Conversion (RP-HPLC): 96%; LC-ESI-MS [M+H]⁺: 196.97, 197.97

SpH2: Conversion (RP-HPLC): Br: 97%; Di-Br: 3%; LC-ESI-MS [M-H]⁻: 195.0, 197.0

Bromo-pyrrolo-2.3-pyrimidine

theoretical mass [C₆H₄BrN₂] *m/z* 196.96, 198.96

SpH1: Conversion (RP-HPLC): 100%; LC-ESI-MS [M-H]⁻: 195.96, 197.95

Bromoazulene

theoretical mass [C₁₀H₇Br] *m/z* 205.97, 207.97

Dibromoazulene: theoretical mass [C₁₀H₆Br₂] *m/z* 285.88, 283.88, 287.88

SpH1: Conversion (RP-HPLC): 98%; GC-MS [M+H]⁺: 206.14, 208.14

SpH2: Conversion (RP-HPLC): Br: 7%; di-Br: 87%, LC-ESI-MS [M-H]⁻: For the mono- and di-brominated products of SpH2 the detection failed due to insufficient ionisation.

Bromo-4-*n*-hexylresorcinol

theoretical mass [C₁₂H₁₇BrNO₂] *m/z* 272.04, 274.04

SpH1: Conversion (RP-HPLC): 21% + 4 side products; LC-ESI-MS [M-H]⁻: 271.03, 273.03

Bromoindole-3-acetic acid

theoretical mass [C₁₀H₈BrNO₂] *m/z* 252.97, 254.97

SpH1: Conversion (RP-HPLC): 0%

SpH2: Conversion (RP-HPLC): Br: 77%; Di-Br: 23%; Br: LC-ESI-MS [M-H]⁻: 251.97, 253.97

Bromotryptophol

theoretical mass [C₁₀H₁₀BrNO] *m/z* 238.99, 240.99

Dibromotryptophol: theoretical mass [C₁₀H₉Br₂NO] *m/z* 318.9, 316.91, 320.9

SpH1: Conversion (RP-HPLC): 5%, LC-ESI-MS [M-H]⁺: 237.99, 239.99

SpH2: Conversion (RP-HPLC): 56%; Di-Br: 44%; Br: LC-ESI-MS [M-H]⁺: 238.0, 240.0; Di-Br: LC-ESI-MS [M-H]⁺: 315.95, 317.95, 319.95

Bromo-indole-3-propionic acid

Bromo-indole-3-propionic acid: theoretical mass [C₁₁H₁₀BrNO₂] *m/z* 266.99, 268.99

Dibromo-indole-3-propionic acid: theoretical mass [C₁₁H₉BrNO₂] *m/z* 346.9, 344.9, 348.9

SpH1: Conversion (RP-HPLC): 2%, LC-ESI-MS [M-H]⁺: 265.99, 267.99

SpH2: Conversion (RP-HPLC): Br: 60%; Di-Br: 33%; Br: LC-ESI-MS [M-H]⁺: 266.06, 268.05

Bromoquinoxaline

theoretical mass [C₈H₅BrN₂] *m/z* 207.96, 209.96

SpH1: Conversion (RP-HPLC): 11%, LC-ESI-MS [M-H]⁺: For the brominated product of SpH1 the mass detection failed due to insufficient ionisation.

6-Bromo-2,3-dimethylindole 23

theoretical mass [C₁₀H₁₀BrN] *m/z* 223.00, 225.00

SpH1: Conversion (RP-HPLC): 15%, LC-ESI-MS [M-H]⁺: 222.00, 224.00

SpH2: Conversion (RP-HPLC): Br: 68%; Di-Br: 27%; LC-ESI-MS [M+H]⁺: 222.0, 224.0

Bromophenol

theoretical mass [C₆H₅BrO] *m/z* 171.95, 173.95

SpH1: Conversion (RP-HPLC): 17%; GC-ESI-MS [M+H]⁺: 172.13, 174.12

Bromoanthranilic acid

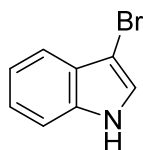
theoretical mass [C₇H₆BrNO₂] *m/z* 214.96, 216.96

SpH1: Conversion (RP-HPLC): 2 product peaks: P1: 71%; P2: 29%; LC-ESI-MS [M-H]⁻: 213.95, 215.95 (P1 *t_R* 7.3 min; P2 *t_R*: 7.6 min)

Determination of regioselective halogenation site *via* NMR analysis

3-Bromoindole 1

HRMS: C₈H₆BrN⁻ *m/z* calculated: 194.9589; observed: 194.9862

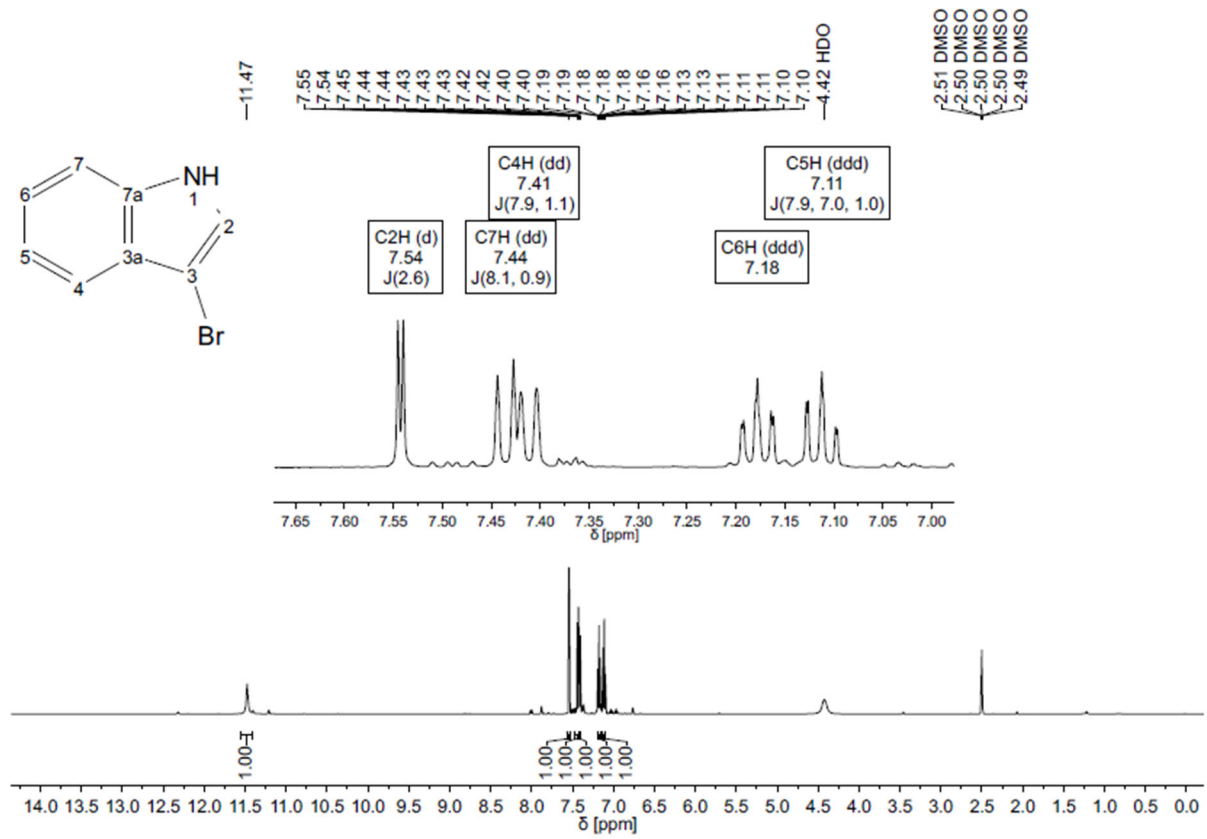


Catalyzed by SpH1

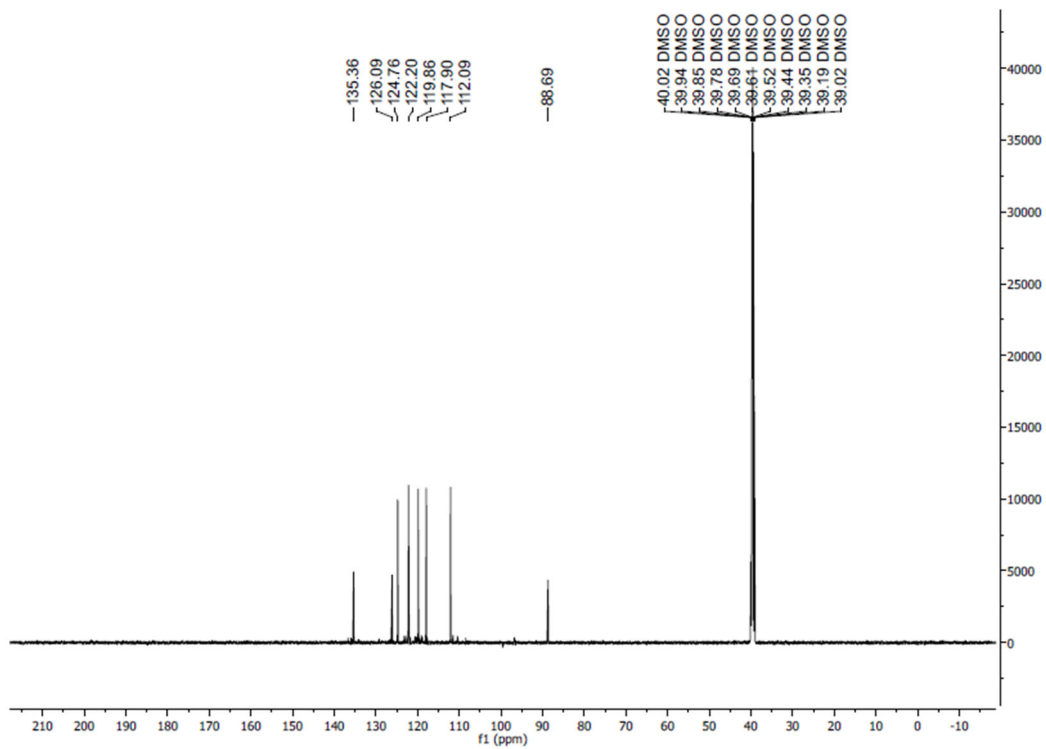
Yield: 35.6%, 0.267 mmol, 52.05 mg

¹H NMR (500 MHz, DMSO-*d*₆) δ 11.47 (s, NH), 7.54 (d, *J* = 2.6 Hz, C²H), 7.44 (dd, *J* = 8.1, 0.9 Hz, C⁷H), 7.41 (dd, *J* = 7.9, 1.1 Hz, C⁴H), 7.18 (ddd, *J* = 8.2, 7.0, 1.3 Hz C⁶H), 7.11 (ddd, *J* = 7.9, 7.0, 1.0 Hz, C⁵H).

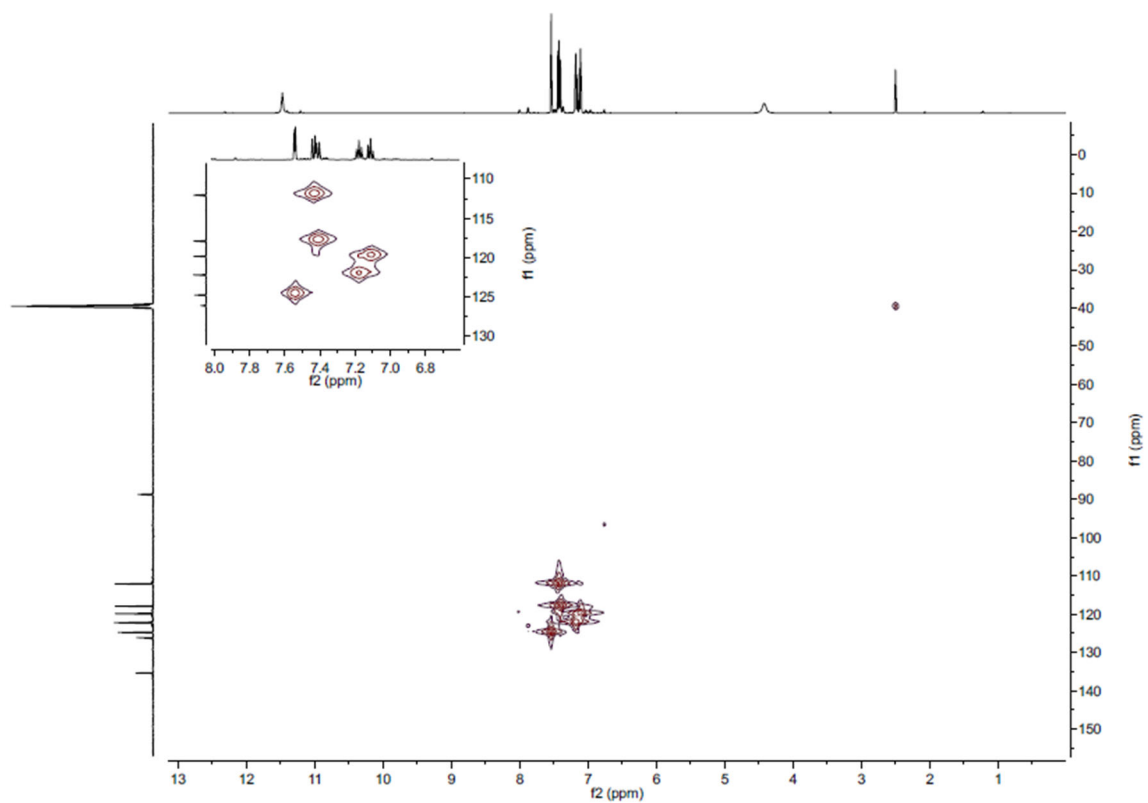
¹³C NMR (126 MHz, DMSO-*d*₆) δ 135.4 (C^{7a}), 126.1 (C^{3a}), 124.8 (C²H), 122.2 (C⁶H), 119.9 (C⁵H), 117.9 (C⁴H), 112.1 (C⁷H), 88.7 (C³H)



1: ¹³C NMR



1: HMQC

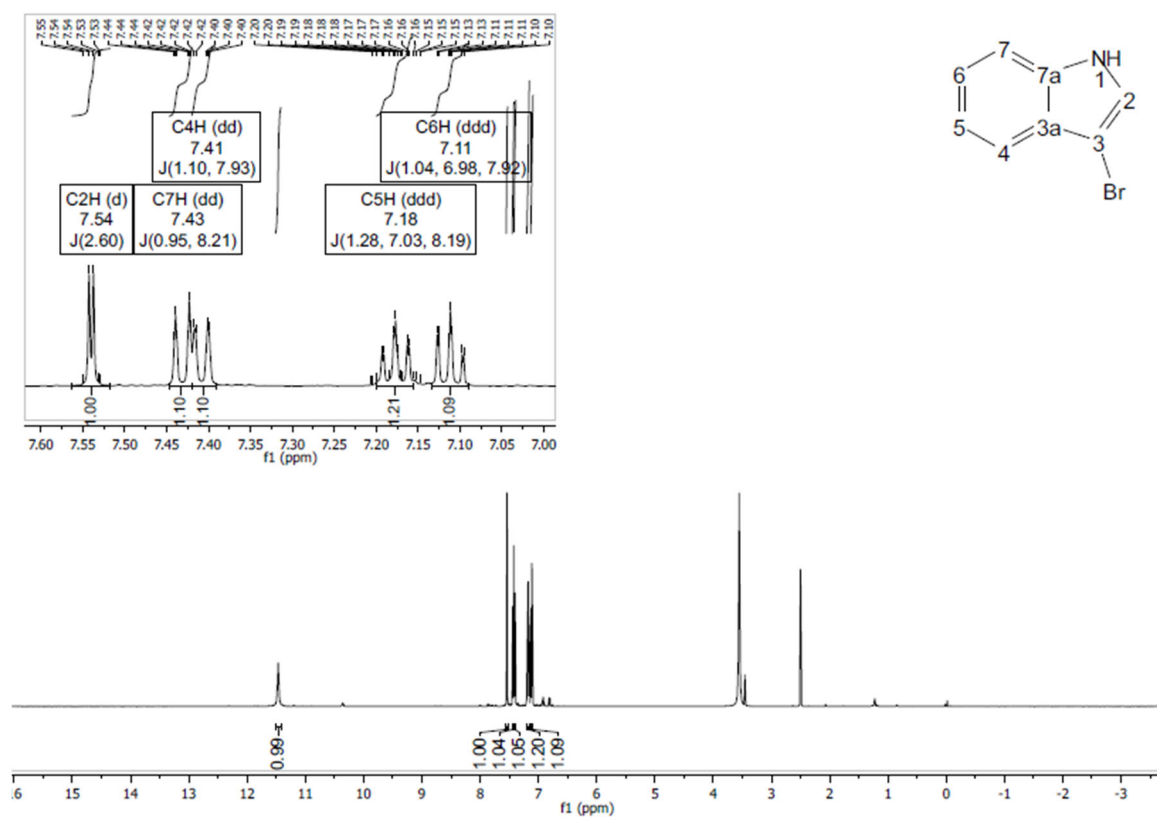


1 Catalyzed by SpH2

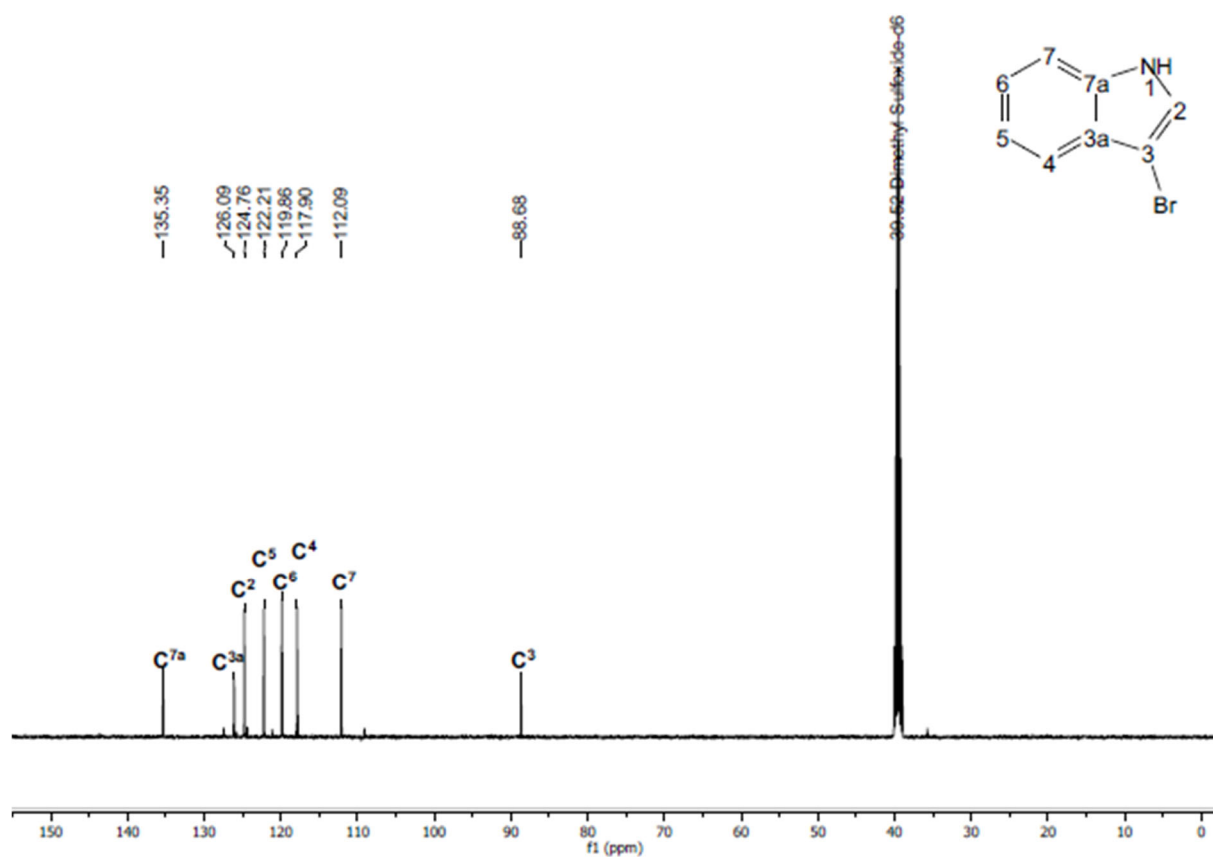
Yield: 10.7%, 0.080 mmol, 15.7 mg

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ [ppm]: 11.47 (s, 1H, N^1H), 7.54 (d, $^3J=2.6$ Hz, 1H, C^2H), 7.43 (dd, $^3J=8.2$ Hz, $^4J=1.0$ Hz, 1H, C^7H), 7.41 (dd, $^3J=7.9$ Hz, $^4J=1.1$ Hz, 1H, C^4H), 7.18 (ddd, $^3J=8.2$ Hz, $^3J=7.0$ Hz, $^4J=1.3$ Hz, 1H, C^6H), 7.11 (ddd, $^3J=7.9$, $^3J=7.0$ Hz, $^4J=1.0$ Hz, 1H, C^5H).

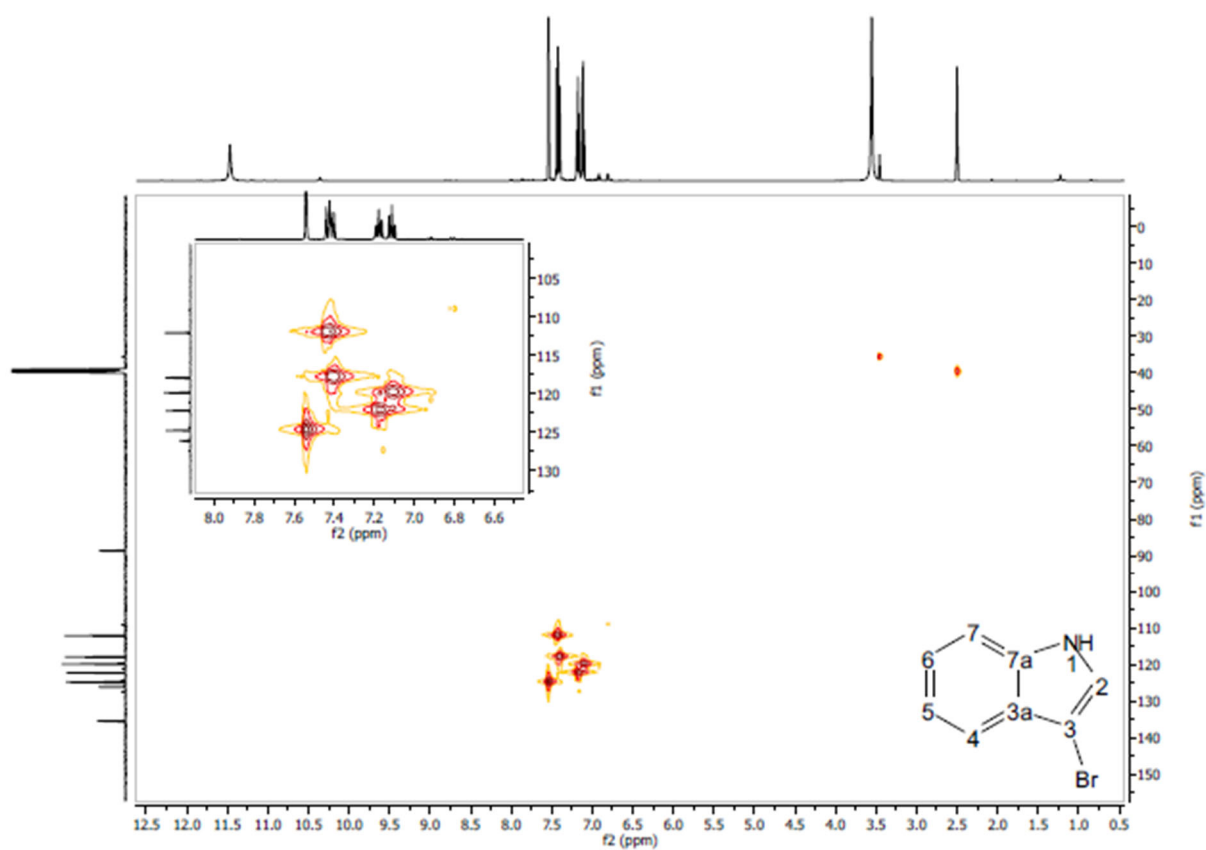
^{13}C NMR (126 MHz, $\text{DMSO-}d_6$) δ : 135.4 (C^7a), 126.1 (C^3a), 124.8 (C^2H), 122.2 (C^5H), 119.9 (C^6H), 117.9 (C^4H), 112.1 (C^7H), 88.7 (C^3Br).



1: ¹³C NMR

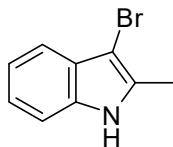


1: HMQC



3-Bromo-2-methylindole 2

HRMS: C₉H₇BrN· *m/z* calculated: 207.9767, observed: 207.9757

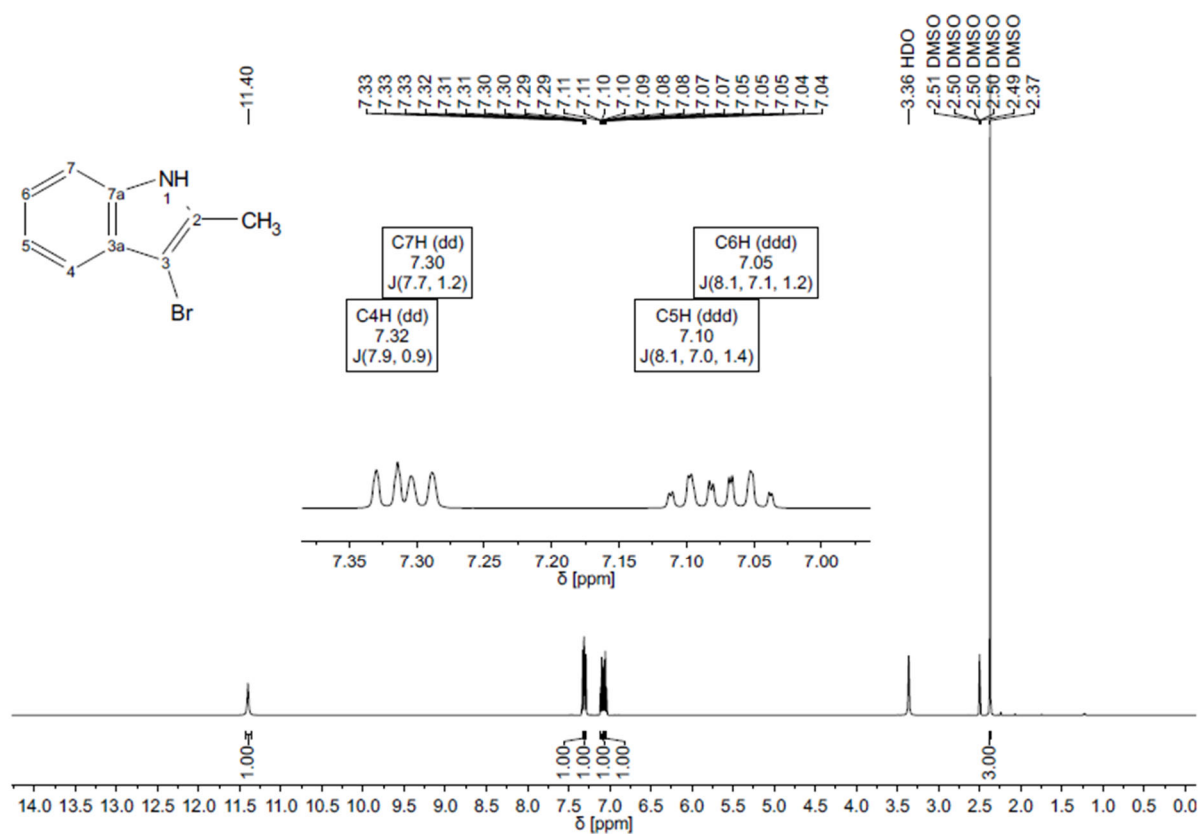


2 Catalyzed by SpH1

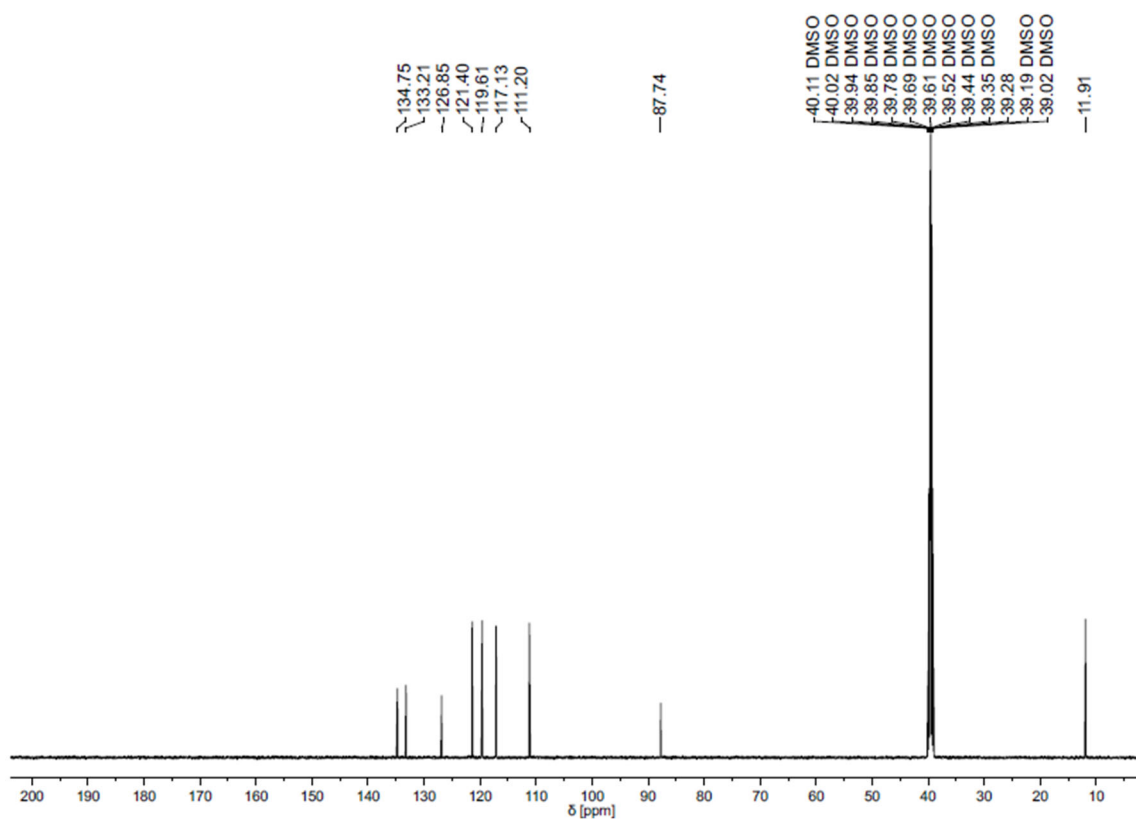
Yield: 18.9%, 0.142 mmol, 29.6 mg

¹H NMR (500 MHz, DMSO-*d*₆) δ 11.40 (s, NH), 7.32 (dd, *J* = 7.9, 0.9 Hz, C⁴H), 7.30 (dd, *J* = 7.7, 1.2 Hz, C⁷H), 7.10 (ddd, *J* = 8.1, 7.0, 1.4 Hz, C⁵H), 7.05 (ddd, *J* = 8.1, 7.1, 1.2 Hz, C⁶H), 2.37 (s, CH₃).

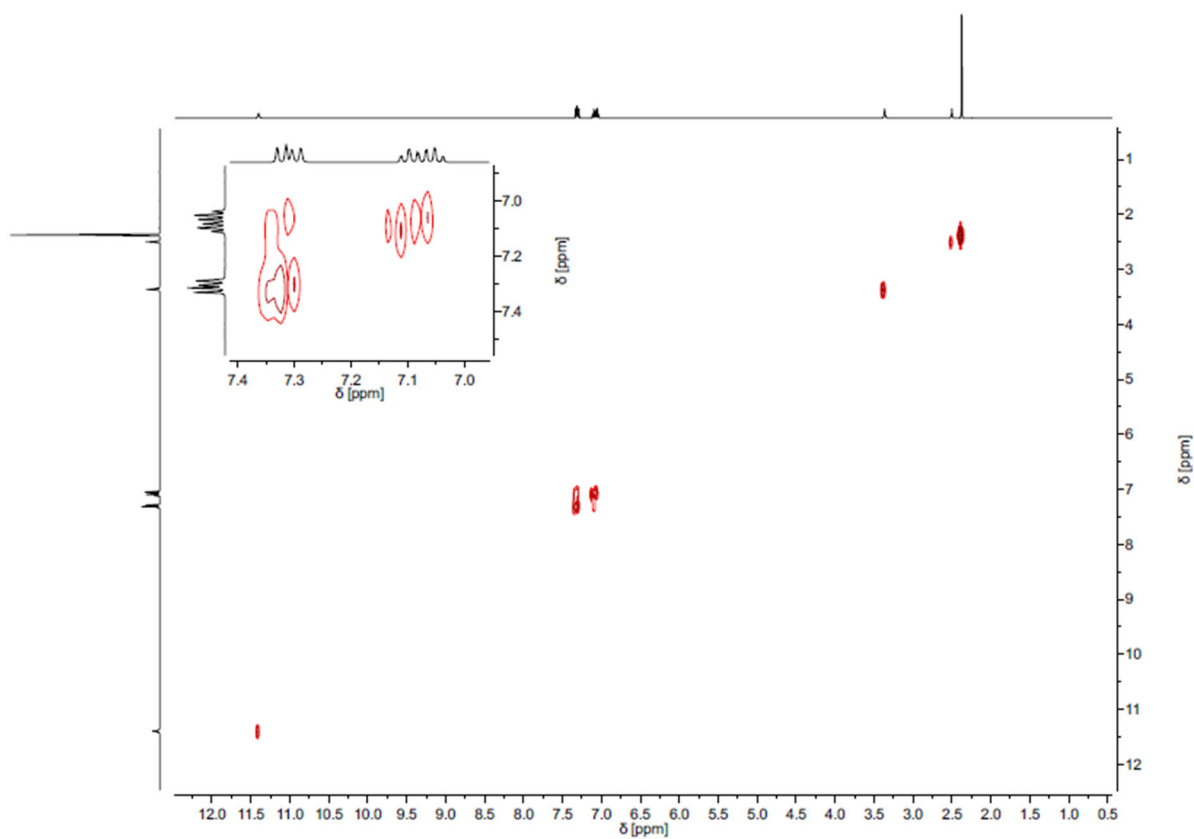
¹³C NMR (126 MHz, DMSO-*d*₆) δ 134.8 (C^{7a}), 133.2 (C²CH₃), 126.9 (C^{3a}), 121.4 (C⁶H), 119.6 (C⁵H), 117.1 (C⁴H), 111.2 (C⁷H), 87.74 (C³Br), 11.9 (C⁸H₃).



2: ^{13}C NMR



2: ^1H , ^1H -COSY

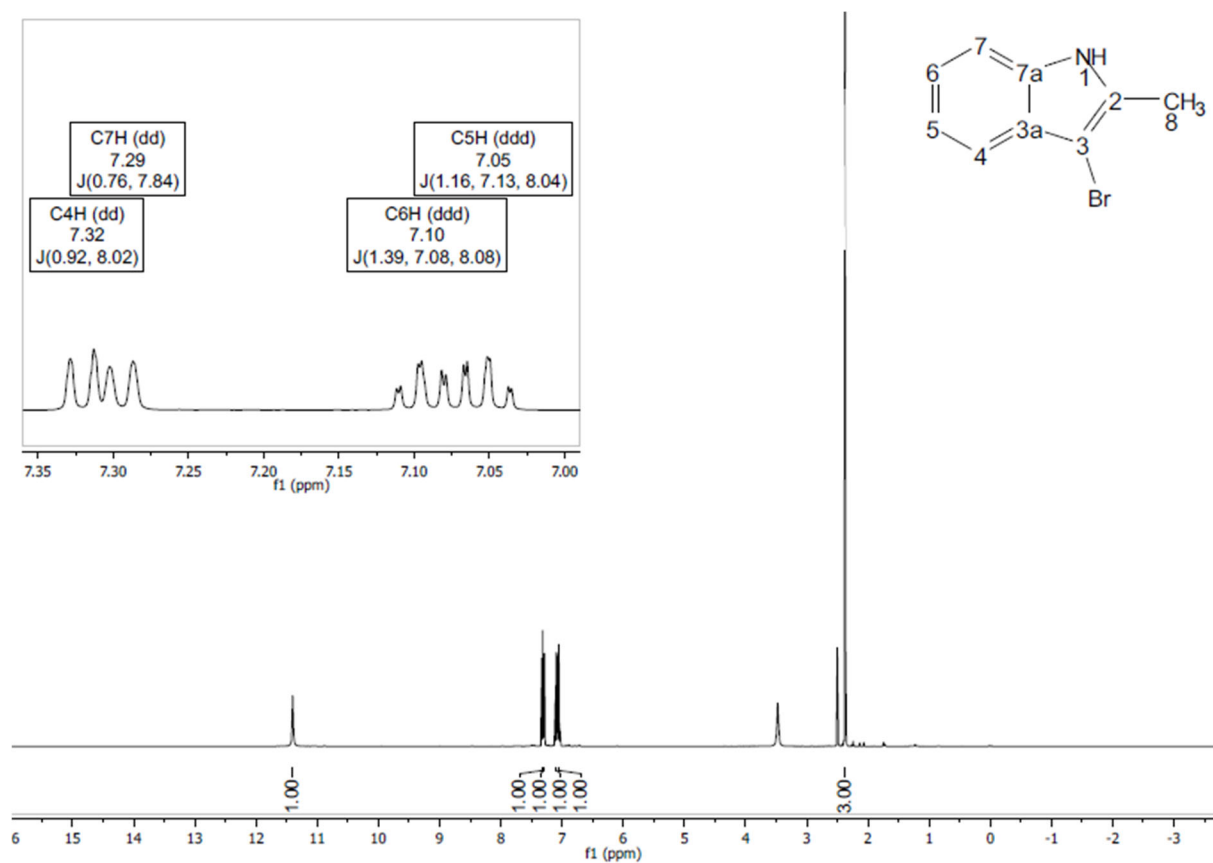


2 Catalyzed by SpH2

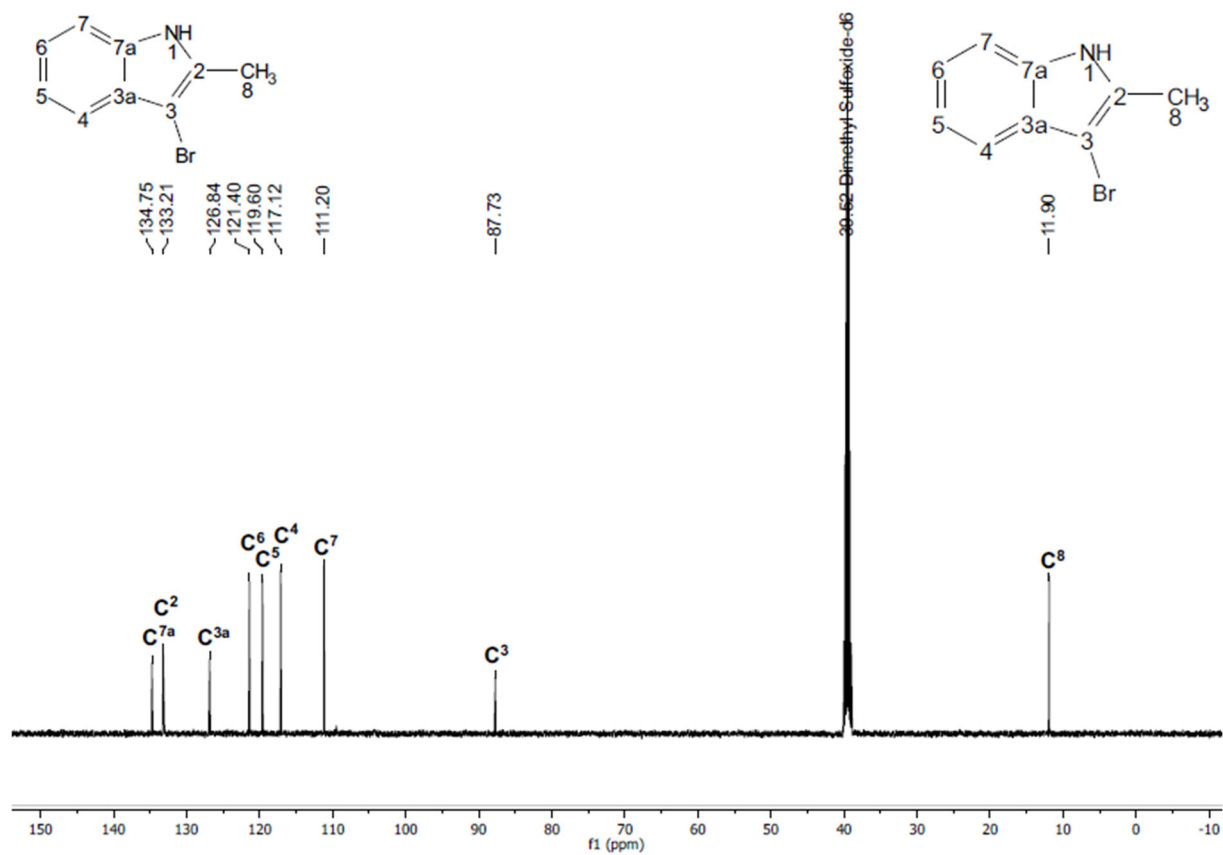
Yield: 24.8%, 0.186 mmol, 39.1 mg

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 11.40 (s, 1H, N^1H), 7.32 (dd, $^3J=8.0$ Hz, $^5J=0.9$ Hz, 1H, C^4H), 7.29 (dd, $^3J=7.8$ Hz, $^5J=0.8$ Hz, 1H, C^7H), 7.10 (ddd, $^3J=8.1$, $^3J=7.1$ Hz, $J=1.4$ Hz, 1H, C^5H), 7.05 (ddd, $^3J=8.0$ Hz, $^3J=7.1$ Hz, $J=1.2$ Hz, 1H, C^6H), 2.37 (s, 3H, C^8H_3).

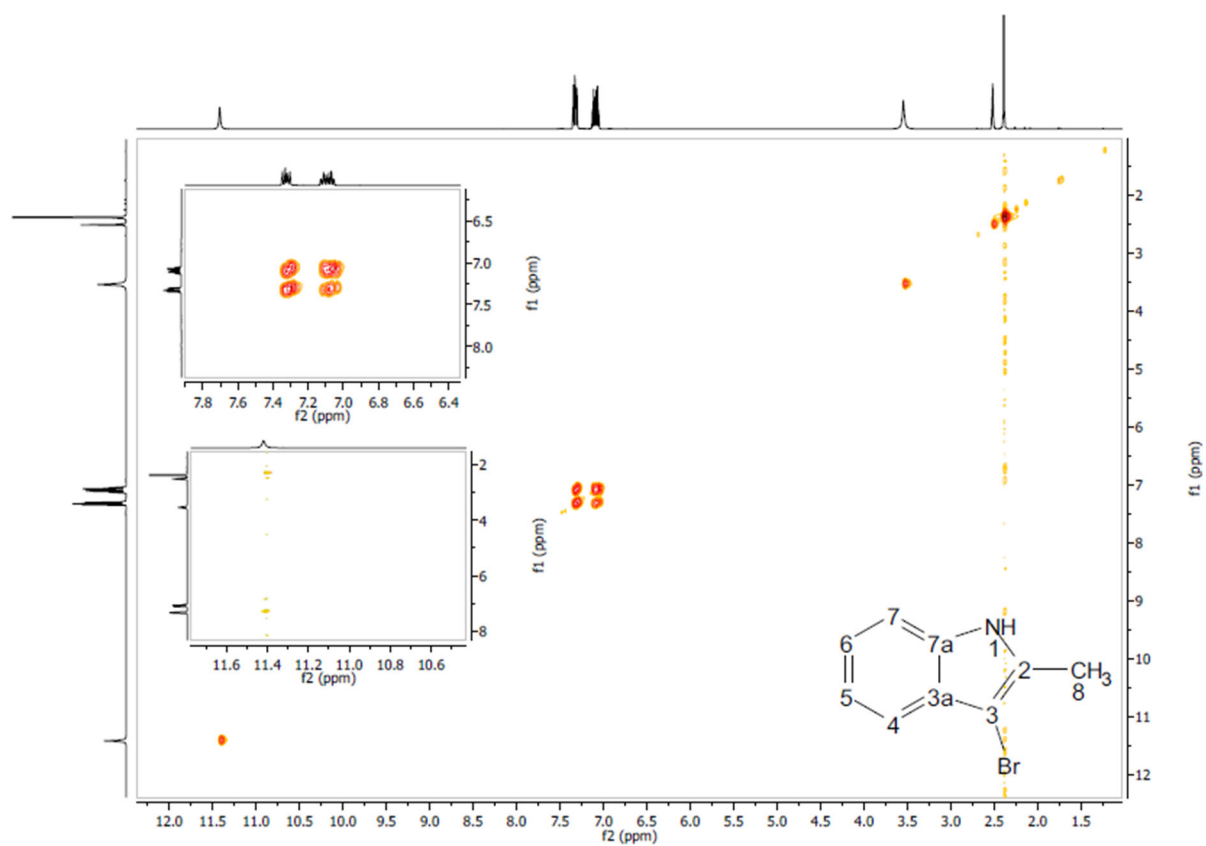
^{13}C NMR (126 MHz, $\text{DMSO-}d_6$) δ 134.8 (C^7a), 133.2 (C^2H), 126.8 (C^3aH), 121.4 (C^6H), 119.6 (C^5H), 117.1 (C^4H), 111.2 (C^7H), 87.7 (C^3Br), 11.9 (C^2CH_3).



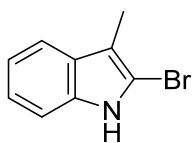
2: ^{13}C NMR



2: ^1H , ^1H -COSY



2-Bromo-3-methylindole 3



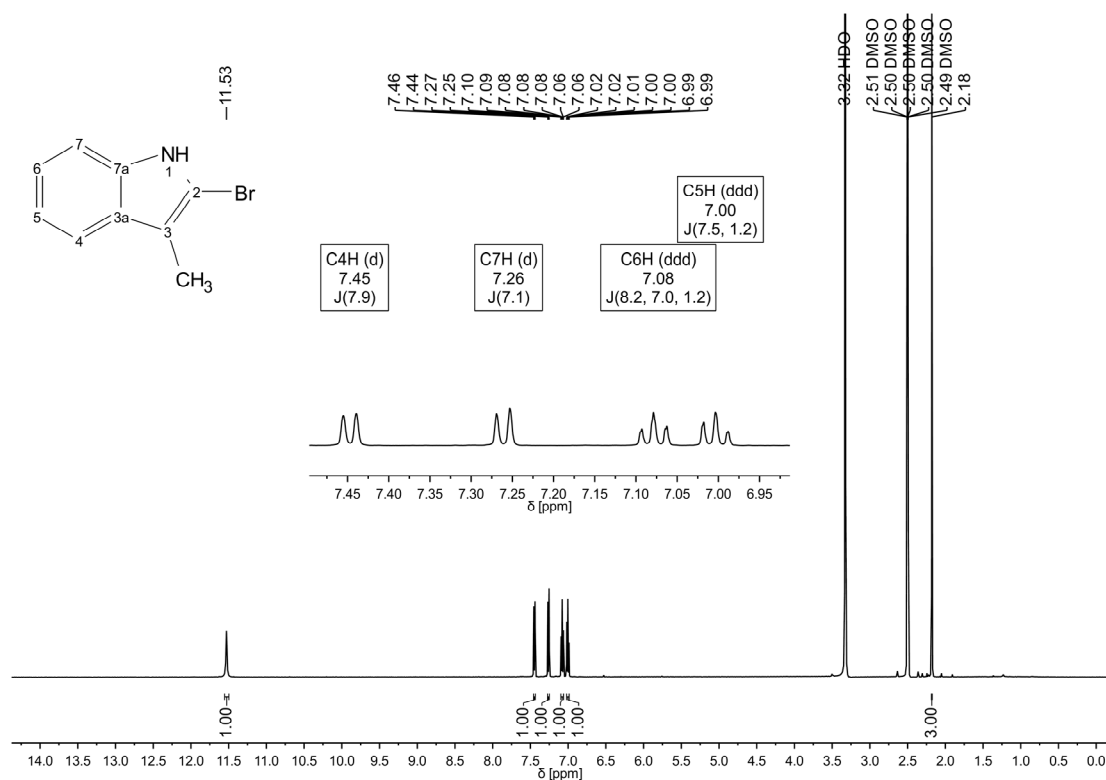
3 Catalyzed by SpH1

Yield: 1.5%, 0.0115 mmol, 2.4 mg

LC-ESI-MS: $[C_9H_7BrN]^+$ m/z : calculated: 207.98 (^{79}Br), 209.98 (^{81}Br); observed: LC-ESI-MS $[M-H]^+$: 207.98, 209.88

1H NMR (500 MHz, DMSO- d_6) δ 11.53 (s, NH), 7.45 (d br., $J = 7.9$ Hz, C⁴H), 7.26 (d br., $J = 7.1$ Hz, C⁷H), 7.08 (ddd, $J = 8.2, 7.0, 1.2$ Hz, C⁶H), 7.00 (ddd, $J = 7.5, 7.1, 1.2$ Hz, C⁵H), 2.18 (s, CH₃).

^{13}C NMR (126 MHz, DMSO- d_6) δ 136.0 (C^{7a}), 127.7 (C^{3a}), 121.5 (C²H), 119.0 (C⁵H), 117.8 (C⁴H), 110.6 (C⁷H), 109.0 (C²Br), 108.3 (C³H), 9.2 (C⁸H₃).



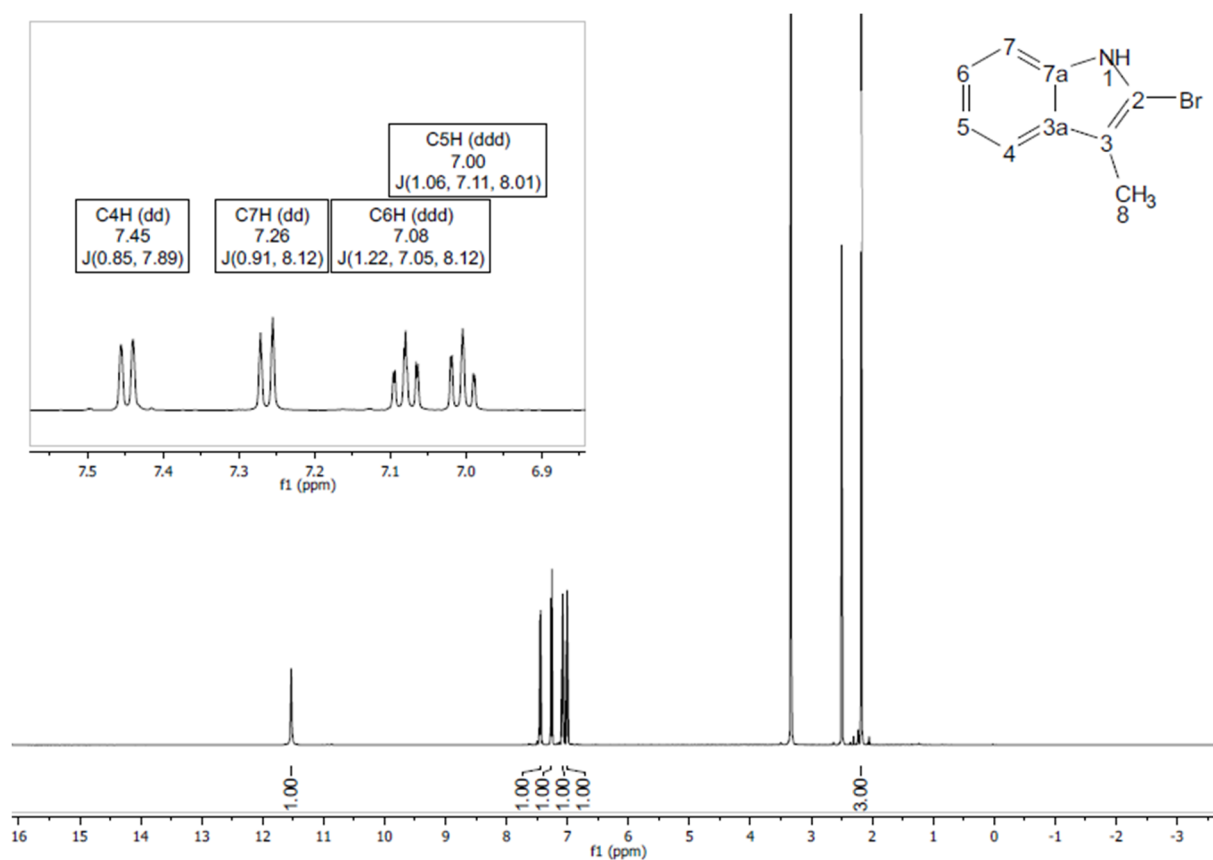
3 Catalyzed by SpH2

Yield: 2%, 0.014 mmol, 3.1 mg

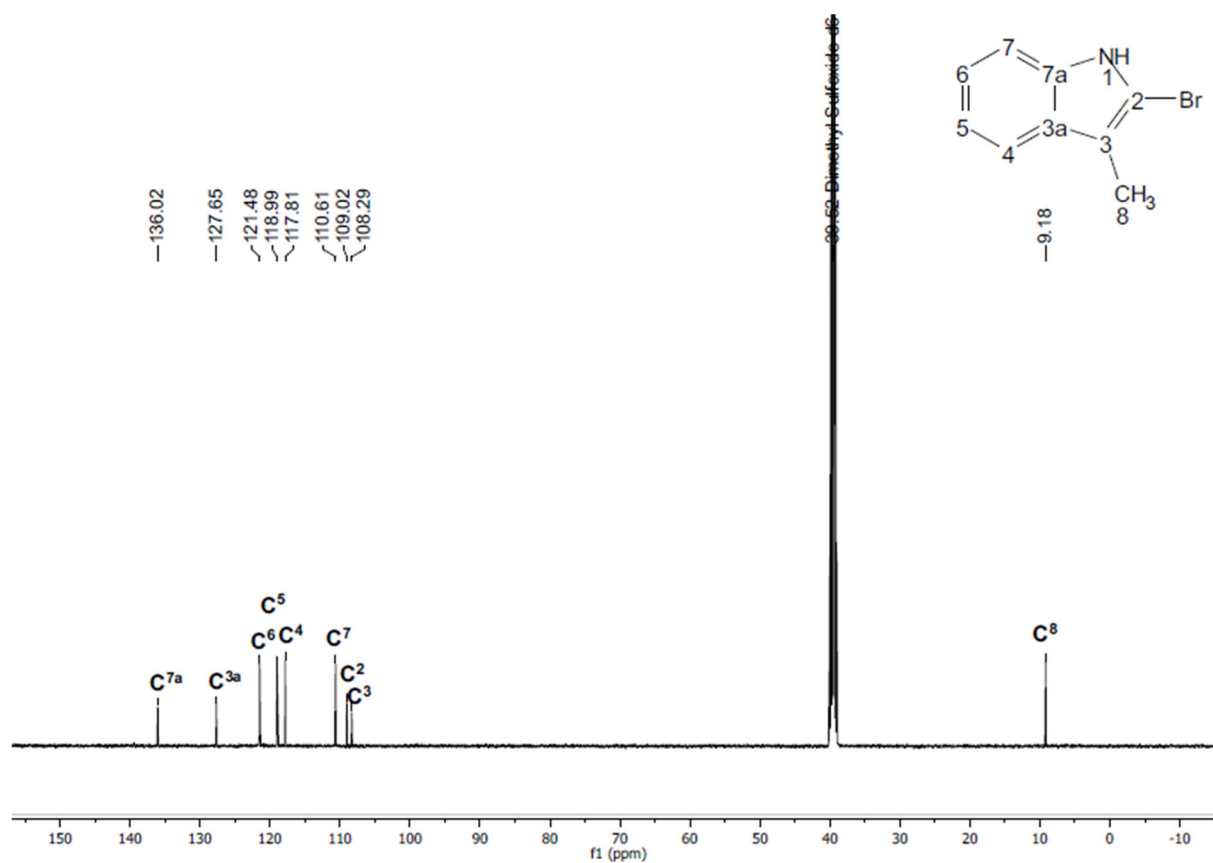
[C₉H₇BrN] *m/z*: calculated: 207.98 (⁷⁹Br), 209.98 (⁸¹Br); observed: LC-ESI-MS [M-H]⁻: 208.0, 210.0

¹H NMR (500 MHz, DMSO-*d*₆) δ 11.53 (s, 1H, N¹H), 7.45 (dd, ³J= 7.9 Hz, ⁵J= 0.9 Hz, 1H, C⁴H), 7.26 (dd, ³J= 8.1 Hz, ⁵J= 0.9 Hz, 1H, C⁷H), 7.08 (ddd, ³J= 8.1 Hz, ³J= 7.1 Hz, ³J= 1.2 Hz, 1H, C⁶H), 7.00 (ddd, ³J= 8.0 Hz, ³J= 7.1 Hz, ³J= 1.1 Hz, 1H, C⁵H), 2.18 (s, 3H, C⁸H₃).

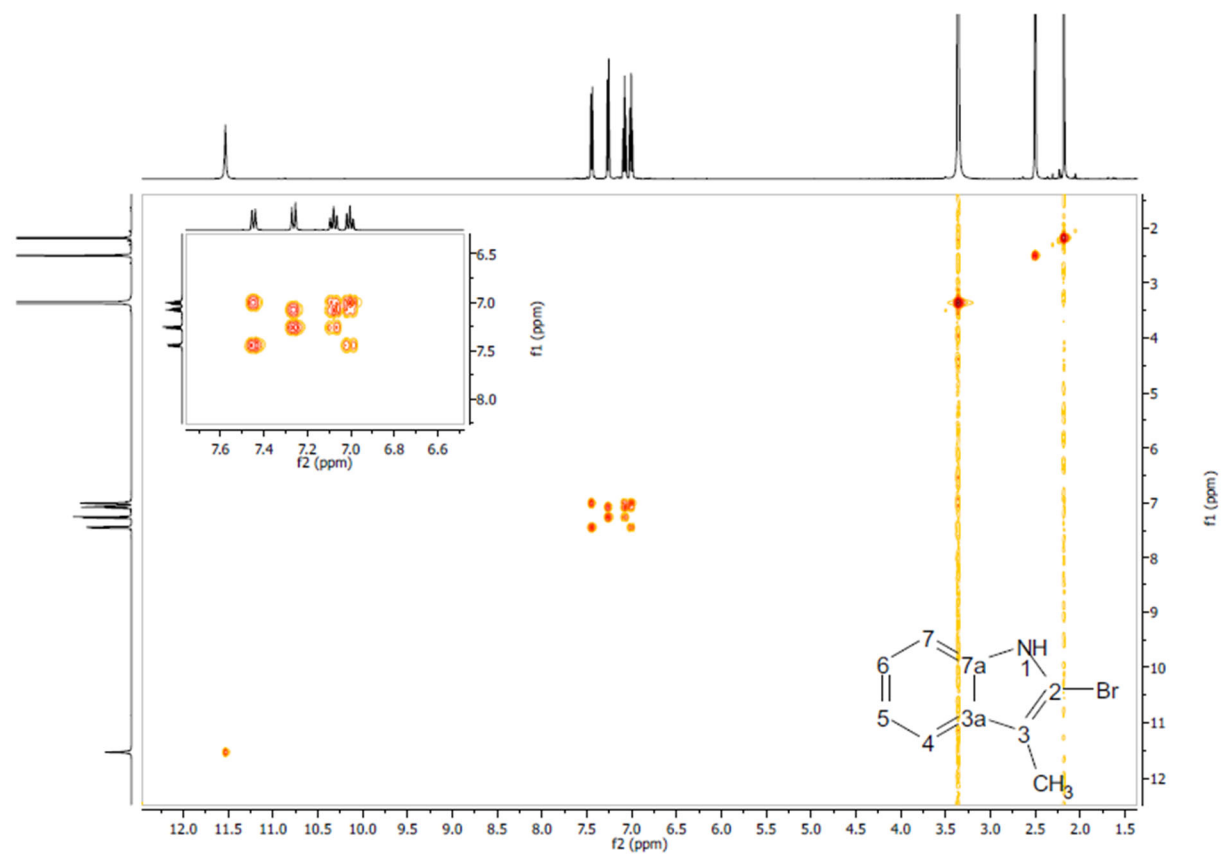
¹³C NMR (126 MHz, DMSO-*d*₆) δ 136.0 (C^{7a}), 127.7 (C^{3a}), 121.5 (C⁶H), 119.0 (C⁵H), 117.8 (C⁴H), 110.6 (C⁷H), 109.0 (C²Br), 108.8 (C³CH₃), 9.2 (C⁸H₃).



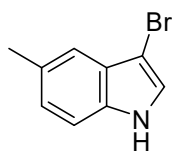
3: ^{13}C NMR



3: $^1\text{H}, ^1\text{H}$ -COSY



3-Bromo-5-methylindole 4



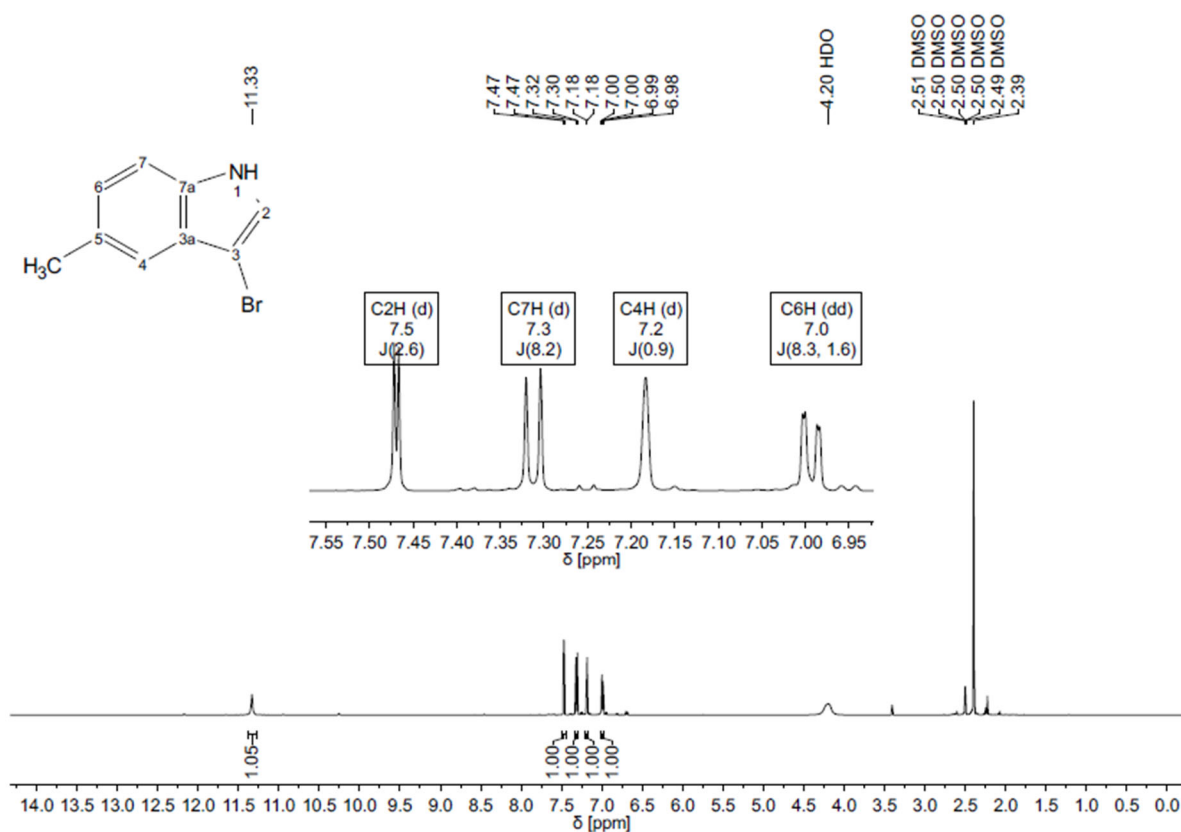
4 Catalyzed by SpH1

Yield: 26.2%, 0.196 mmol, 41 mg

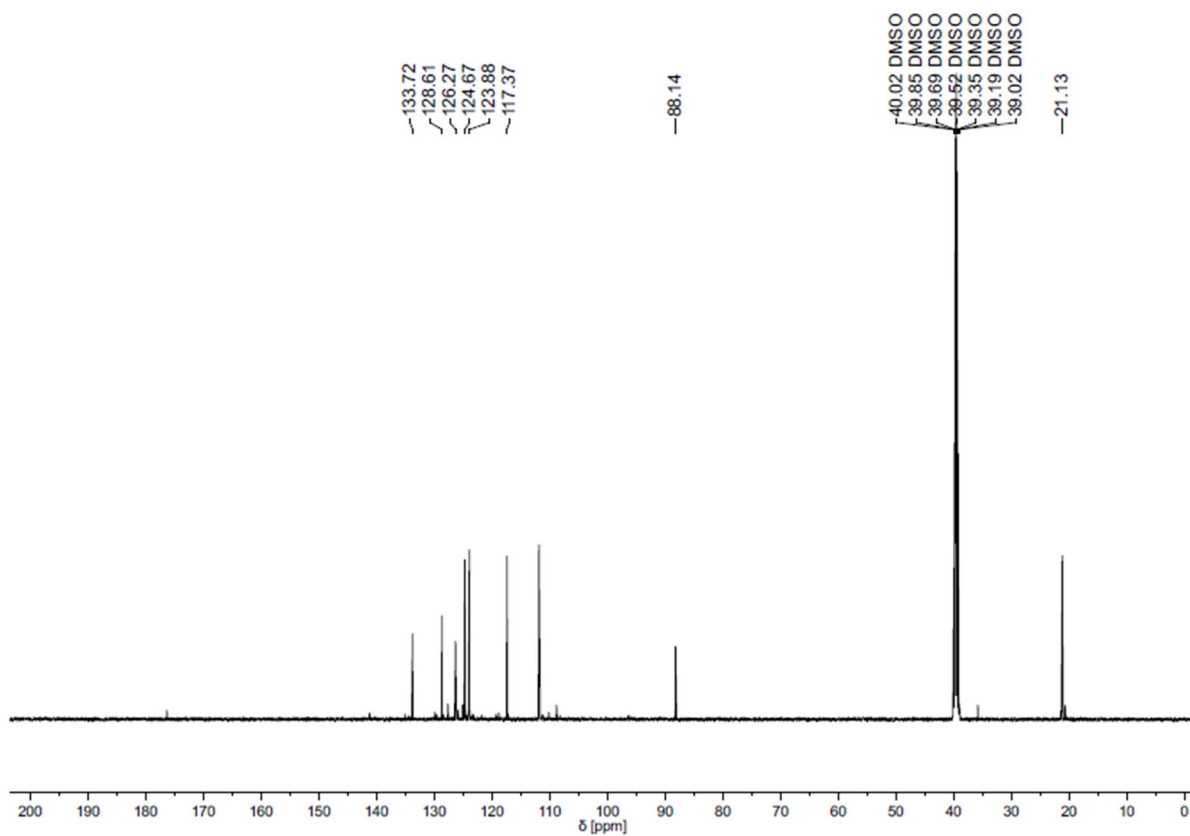
[C₉H₇BrN] m/z: calculated: 207.98 (⁷⁹Br), 209.98 (⁸¹Br); observed: LC-ESI-MS [M-H]⁻: 207.98, 209.98

¹H NMR (500 MHz, DMSO-*d*₆) δ 11.33 (s, NH), 7.47 (d, *J* = 2.6 Hz, C²H), 7.31 (d, *J* = 8.2 Hz, C⁷H), 7.18 (d, *J* = 0.9 Hz, C⁴H), 6.99 (dd, *J* = 8.3, 1.6 Hz, C⁶H), 2.39 (s, CH₃).

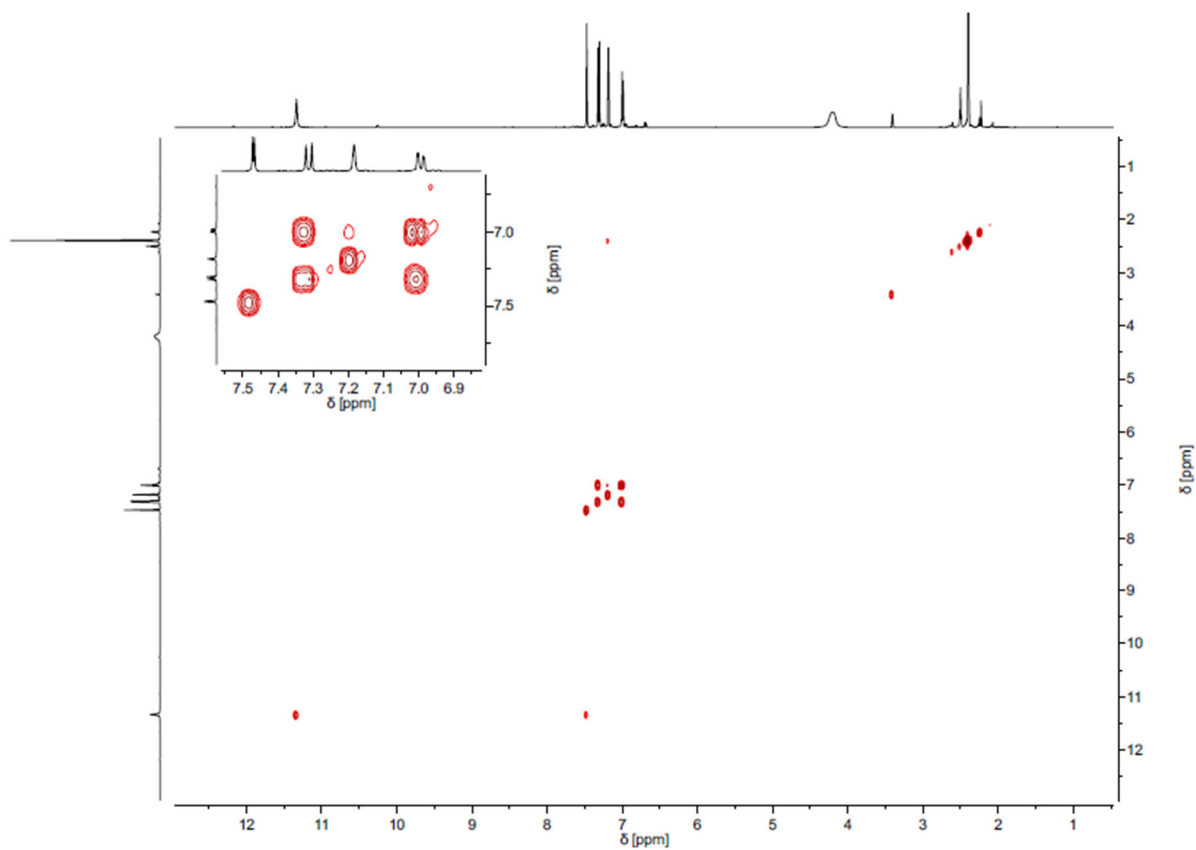
¹³C NMR (126 MHz, DMSO-*d*₆) δ 133.7 (C^{7a}), 128.6 (C^{5a}), 126.3 (C^{3a}), 124.7 (C²H), 123.9 (C⁶H), 117.4 (C⁴H), 111.8 (C⁷H), 88.1 (C³Br), 21.1 (C⁸H₃).



4: ^{13}C NMR



4: ^1H , ^1H -COSY



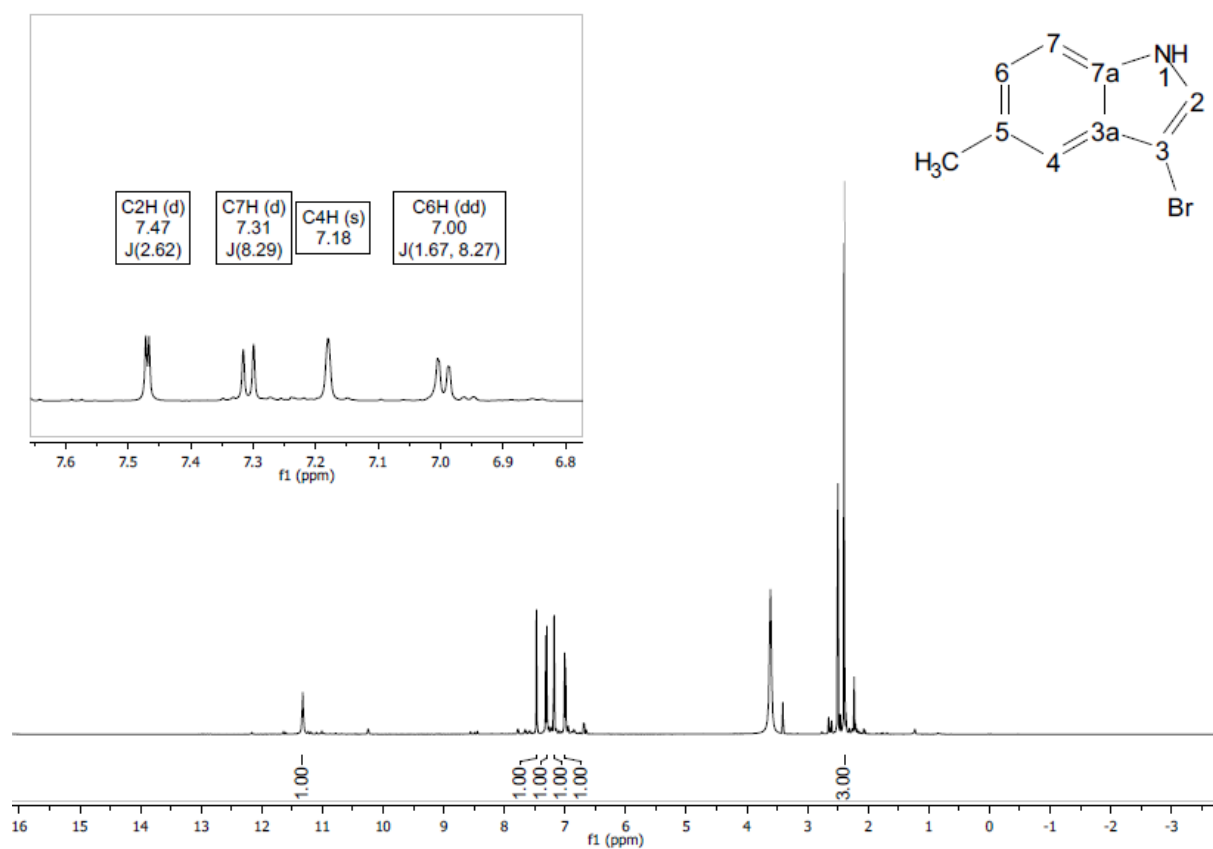
4 Catalyzed by SpH2:

Yield: 4.8%, 0.036 mmol, 7.5 mg

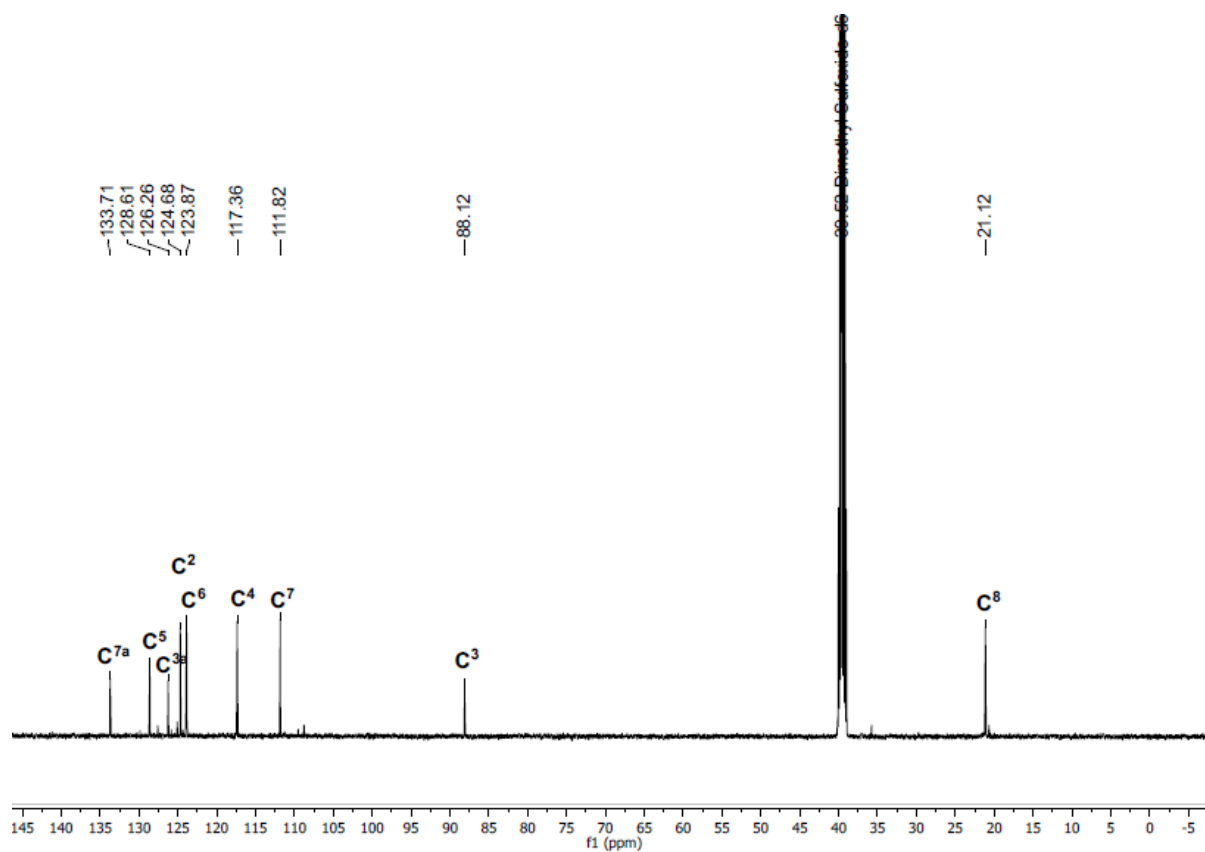
[C₉H₇BrN]⁺ m/z: calculated: 207.98 (⁷⁹Br), 209.98 (⁸¹Br); observed: LC-ESI-MS [M-H]⁺: 208.0, 210.0

¹H NMR (500 MHz, DMSO-*d*₆): δ [ppm]= 11.32 (s, 1H, N¹H), 7.47 (d, ³J= 2.6 Hz, 1H, C²H), 7.31 (d, ³J= 8.3 Hz, 1H, C⁷H), 7.18 (s, 1H, C⁴H), 7.00 (dd, ³J= 8.3 Hz, ⁴J=1.7 Hz, 1H, C⁶H), 2.40 (s, 3H, C⁸H₃).

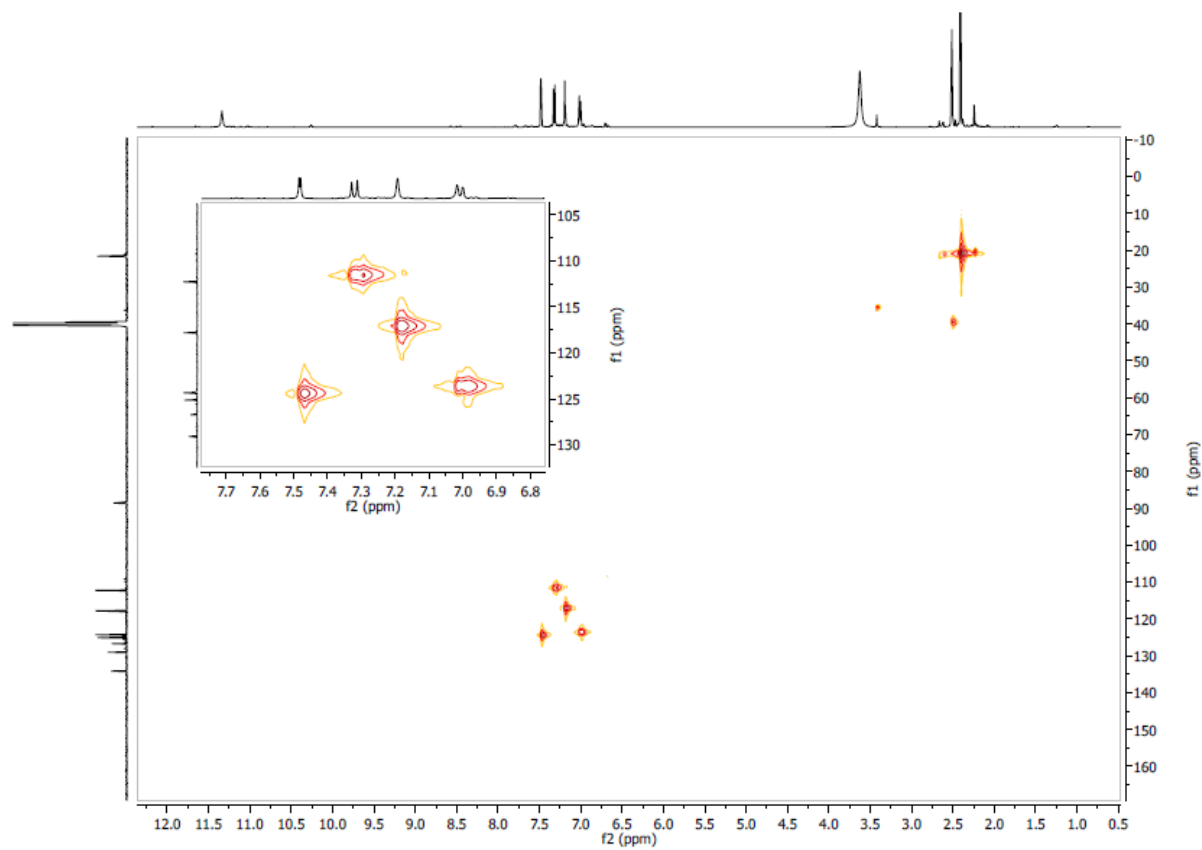
¹³C NMR (126 MHz, DMSO-*d*₆) δ 133.7 (C^{7a}), 128.6 (C⁵H), 126.3 (C^{3a}), 124.7 (C²H), 123.9 (C⁶H), 117.4 (C⁴H), 111.8 (C⁷H), 88.1 (C³Br), 21.1 (C⁸H₃).



4: ^{13}C NMR

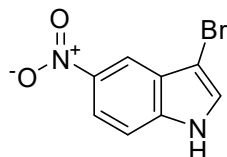


4: HMQC



3-Bromo-5-nitroindole 5

HRMS: C₈H₄BrN₂O₂⁻ *m/z* calculated: 238.462; observed: 238.9448

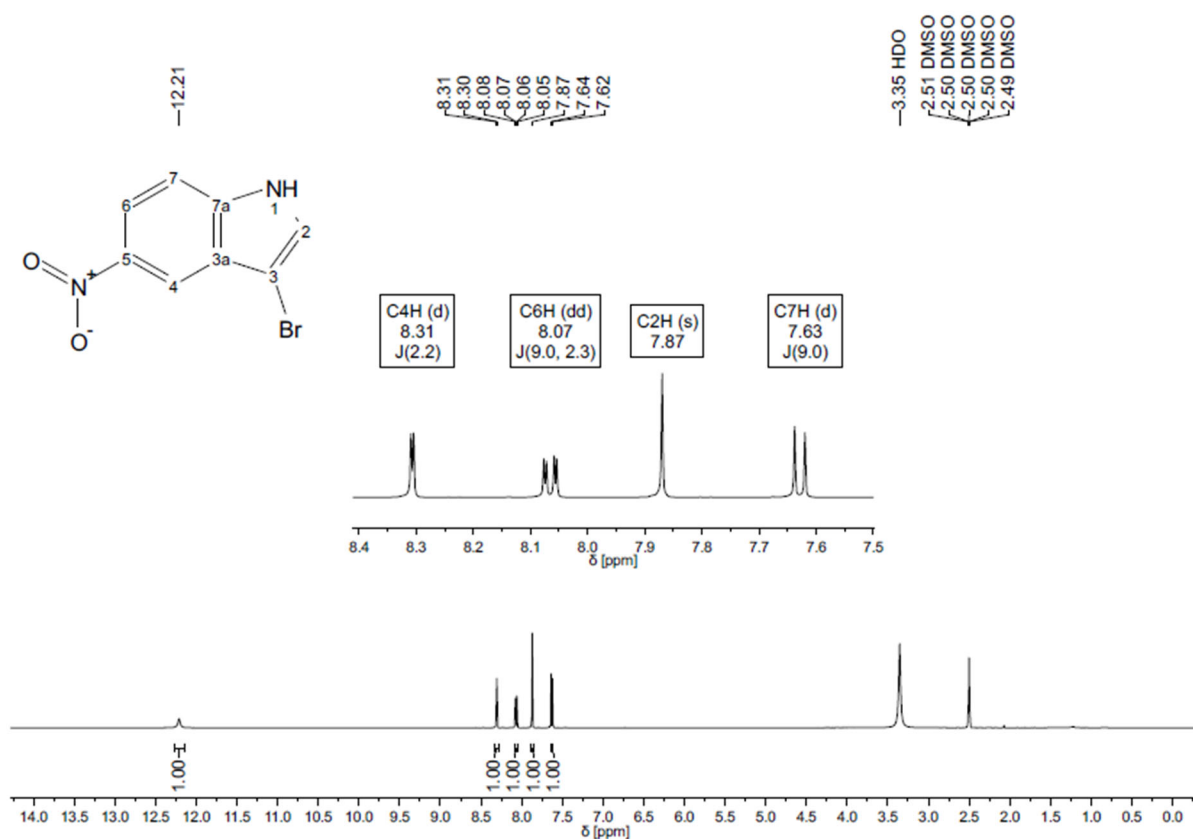


5 Catalyzed by SpH1

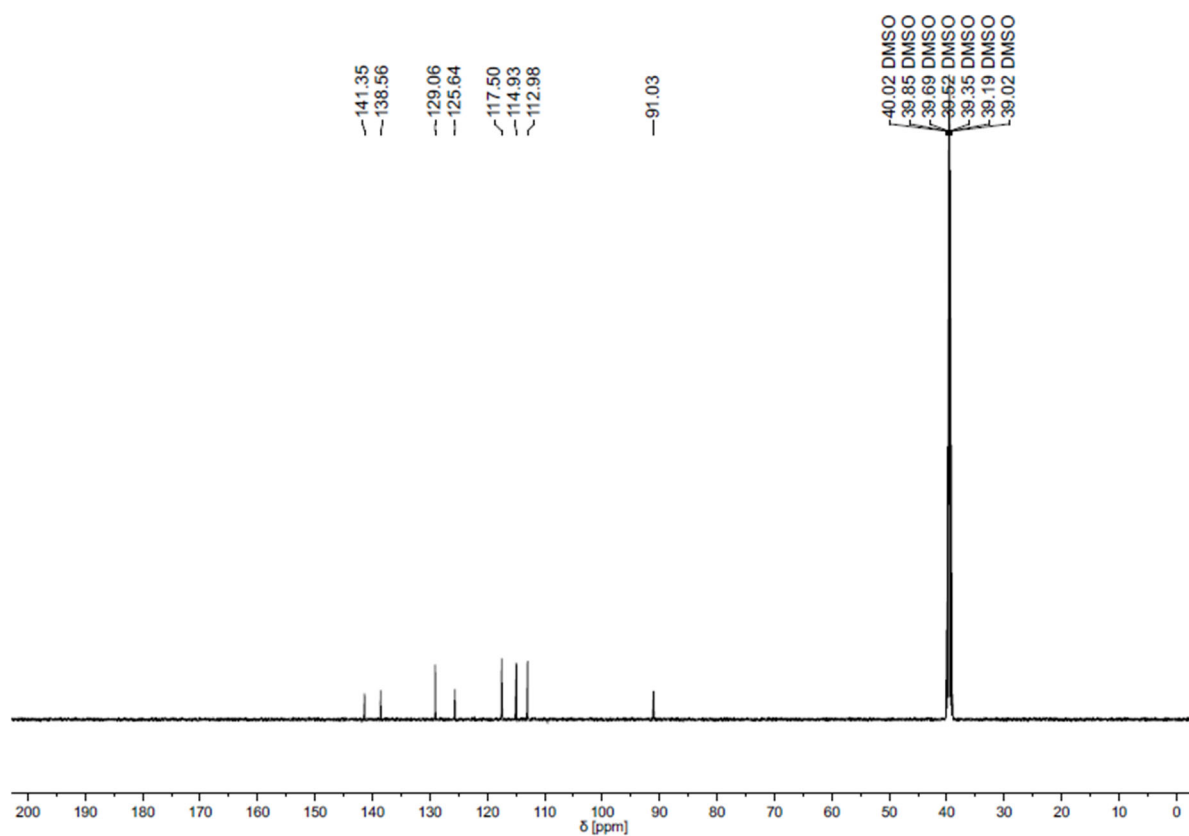
Yield: 12.8%, 0.0954 mmol, 22.9 mg

¹H NMR (500 MHz, DMSO-*d*₆) δ 12.21 (s, NH), 8.31 (d, *J* = 2.2 Hz, C⁴H), 8.07 (dd, *J* = 9.0, 2.3 Hz, C⁶H), 7.87 (s, C²H), 7.63 (d, *J* = 9.0 Hz, C⁷H).

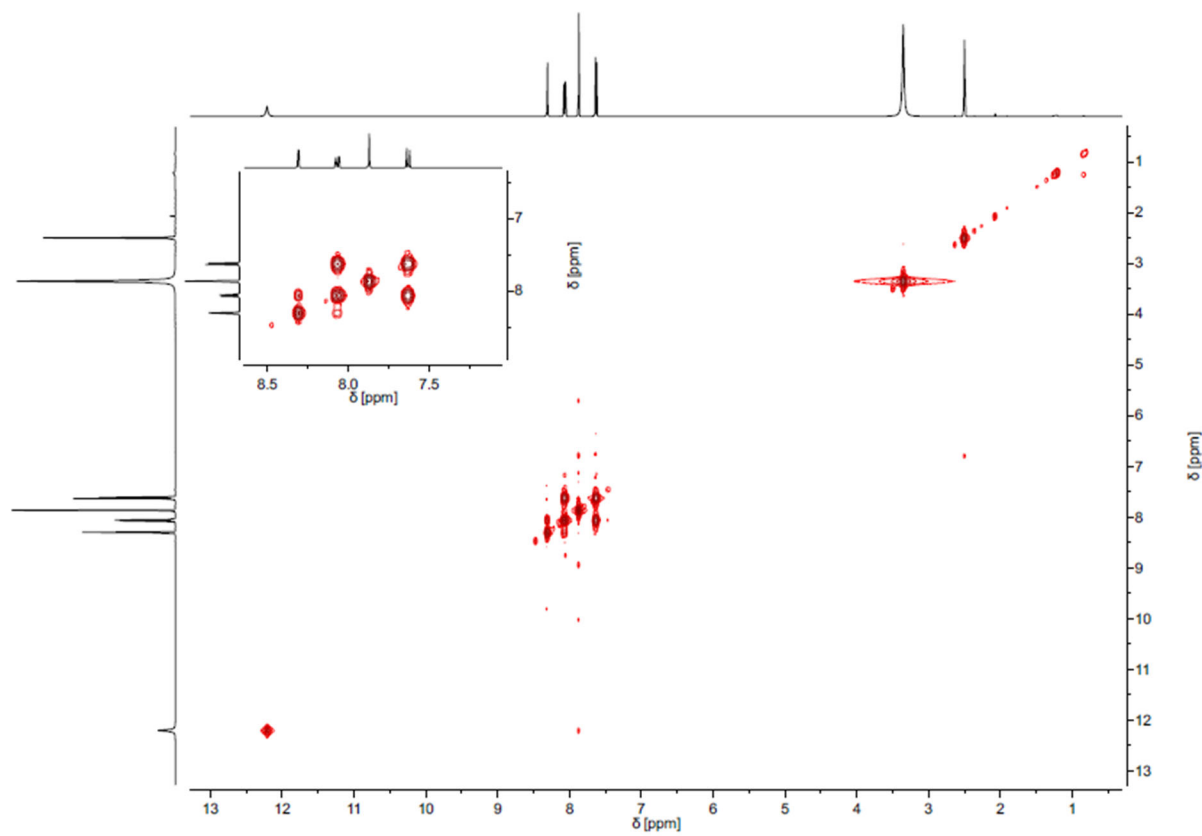
¹³C NMR (126 MHz, DMSO-*d*₆) δ 141.4 (C^{7a}), 138.6 (C⁵NO₂), 129.1 (C^{3a}), 125.6 (C⁴H), 117.5 (C²H), 114.9 (C⁶H), 113.0 (C⁷H), 91.0 (C³Br).



5: ^{13}C NMR



5: $^1\text{H}, ^1\text{H}$ -COSY

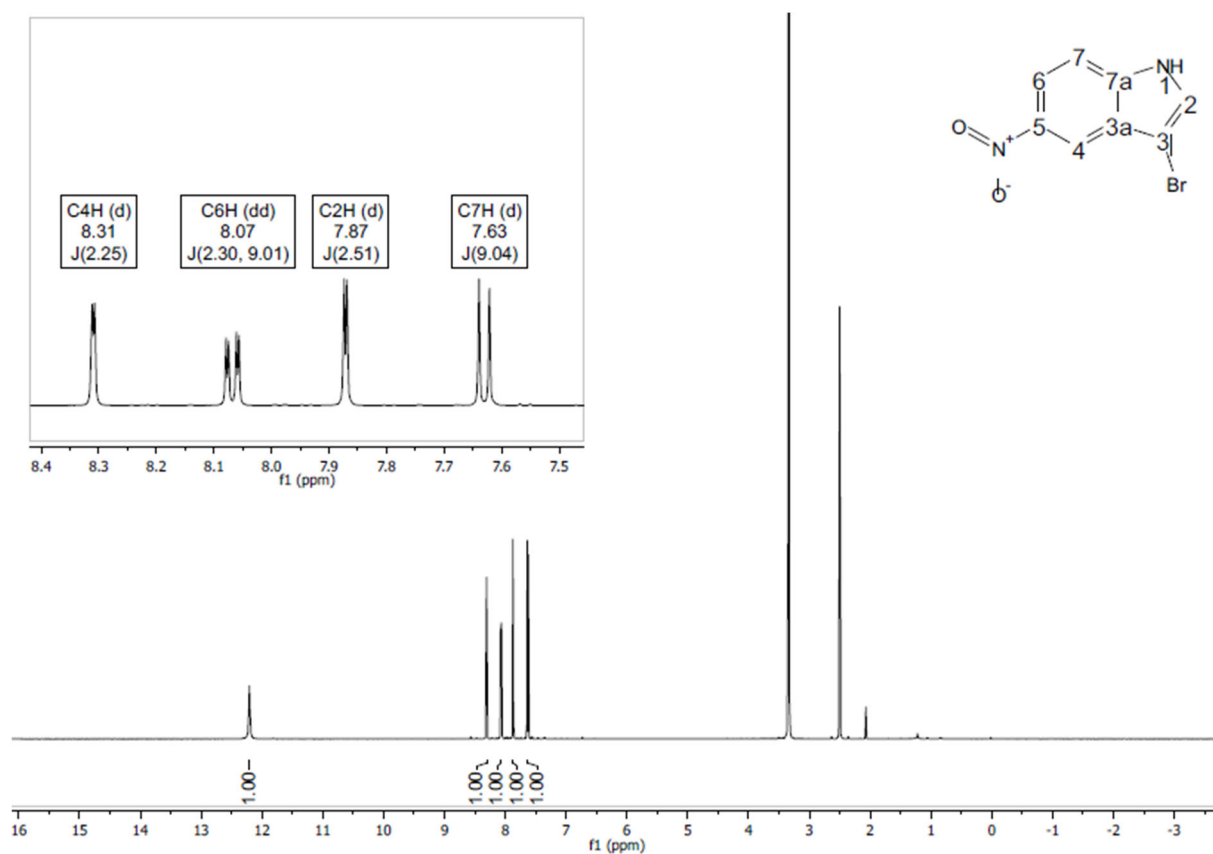


5 Catalyzed by SpH2

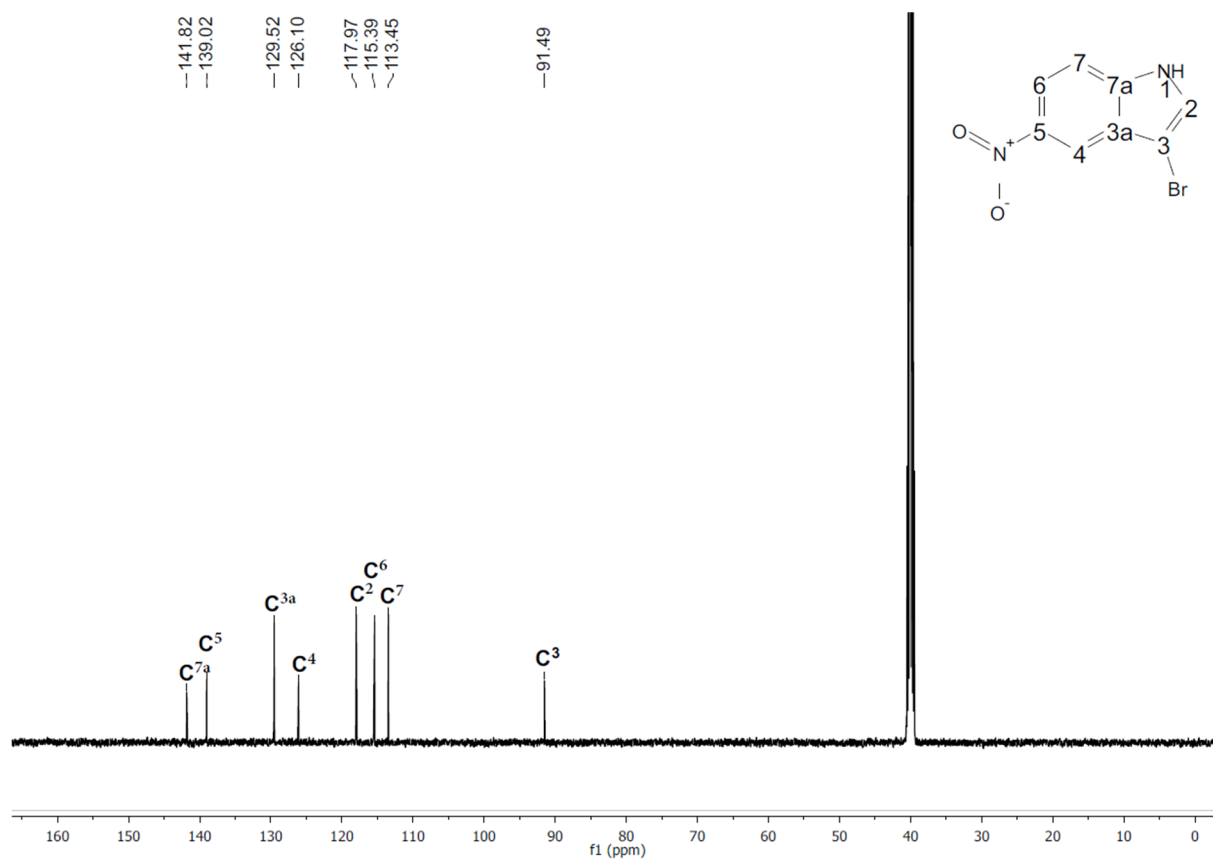
Yield: 11.2%, 0.084 mmol, 20.3 mg

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ [ppm]: 12.21 (s, 1H, N^1H), 8.31 (d, $^4J=2.3$ Hz, 1H, C^4H), 8.07 (dd, $^3J=9.0$ Hz, $^4J=2.3$ Hz, 1H, C^6H), 7.87 (d, $^3J=2.5$ Hz, 1H, C^2H), 7.63 (d, $^3J=9.0$ Hz, 1H, C^7H).

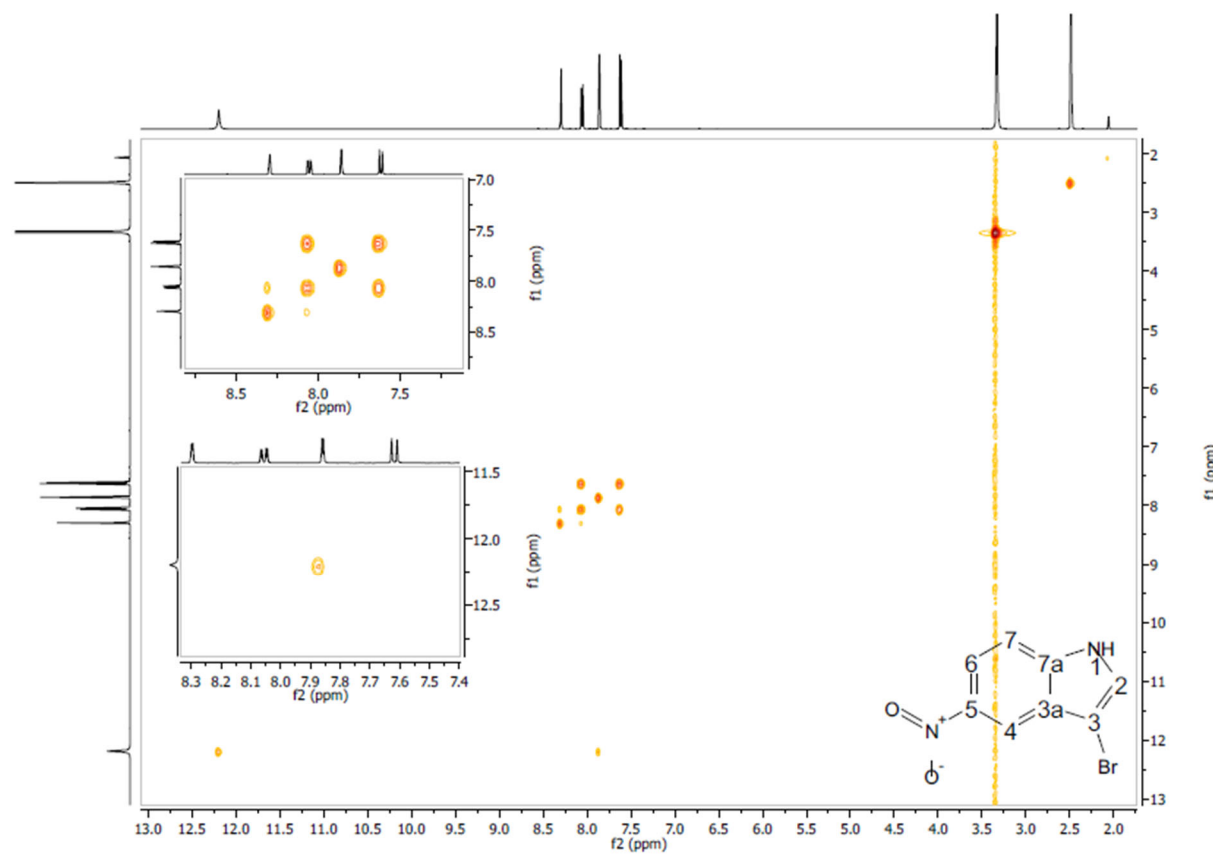
^{13}C NMR (126 MHz, $\text{DMSO-}d_6$) δ 141.8 (C^7a), 139.0 (C^5NO_2), 129.5 (C^3a), 126.1 (C^4H), 118.0 (C^2H), 115.4 (C^6H), 113.5 (C^7H), 91.5 (C^3Br).



5: ¹³C NMR

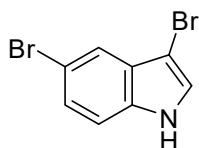


5: ¹H, ¹H-COSY



3, 5-Dibromoindole 7

HRMS: $C_8H_4Br_2N$ m/z calculated: 271.8716; observed: 271.8711

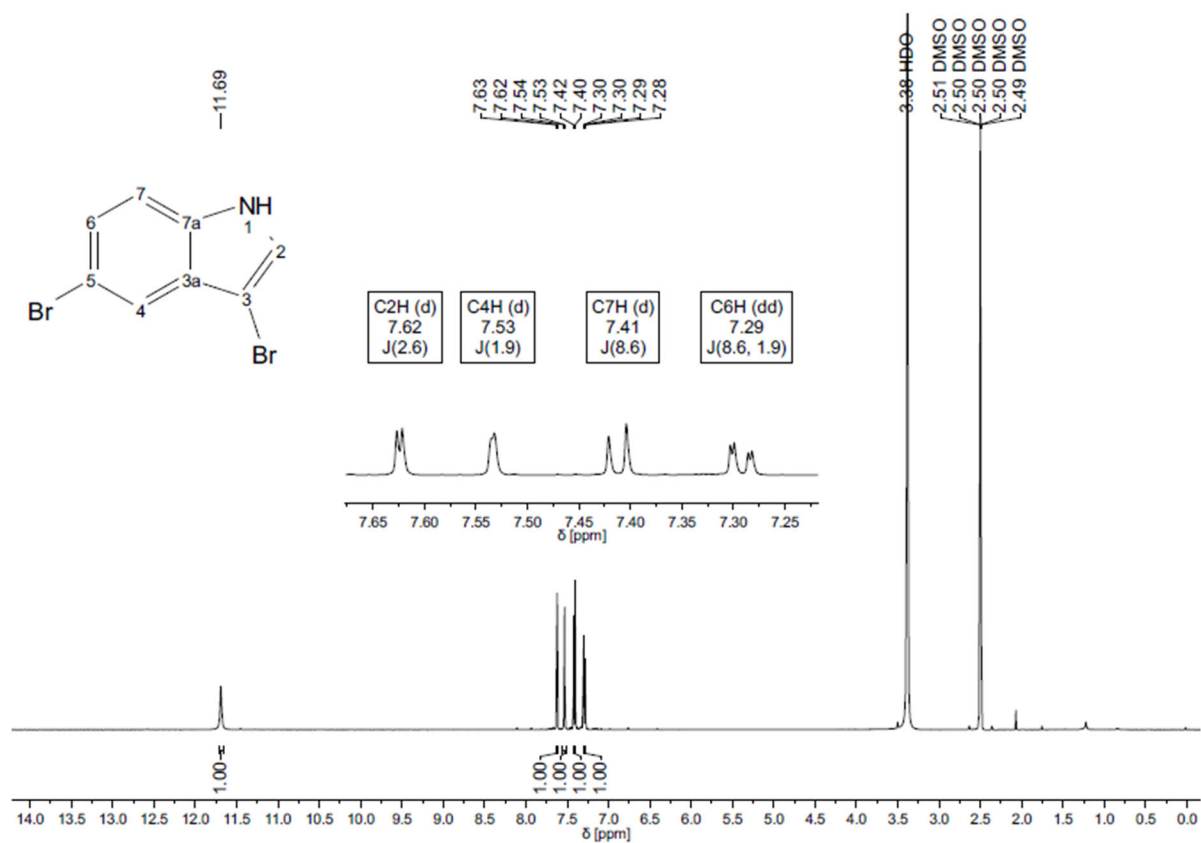


7 Catalyzed by SpH1

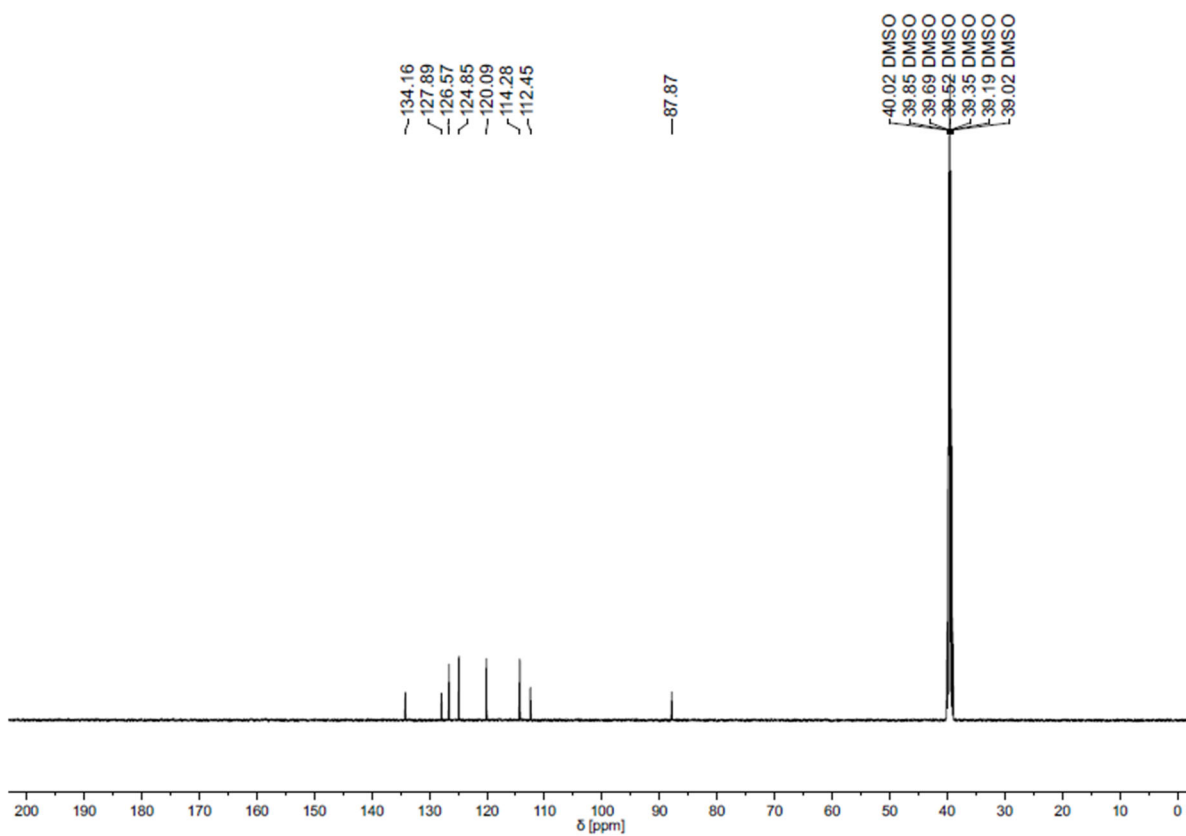
Yield: 8.8%, 0.066 mmol, 18 mg

1H NMR (500 MHz, $DMSO-d_6$) δ 11.69 (s, NH), 7.62 (d, $J = 2.6$ Hz, C²H), 7.53 (d, $J = 1.9$ Hz, C⁴H), 7.41 (d, $J = 8.6$ Hz, C⁷H), 7.29 (dd, $J = 8.6, 1.9$ Hz, C⁶H).

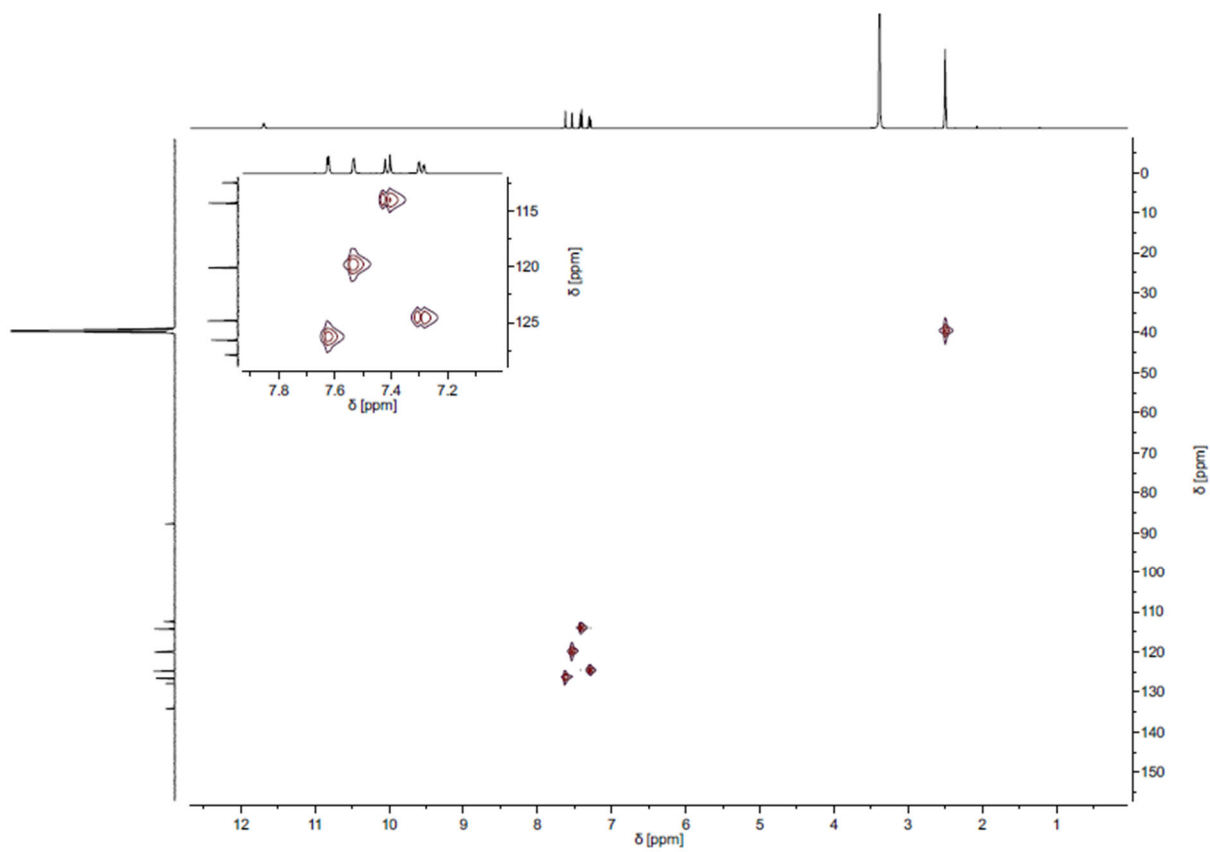
^{13}C NMR (126 MHz, $DMSO-d_6$) δ 134.2 (C^{7a}), 127.9 (C^{3a}), 126.6 (C⁶H), 124.9 (C²H), 120.1 (C⁴H), 114.3 (C⁷H), 112.5 (C⁵Br), 87.9 (C³Br).



7: ¹³C NMR



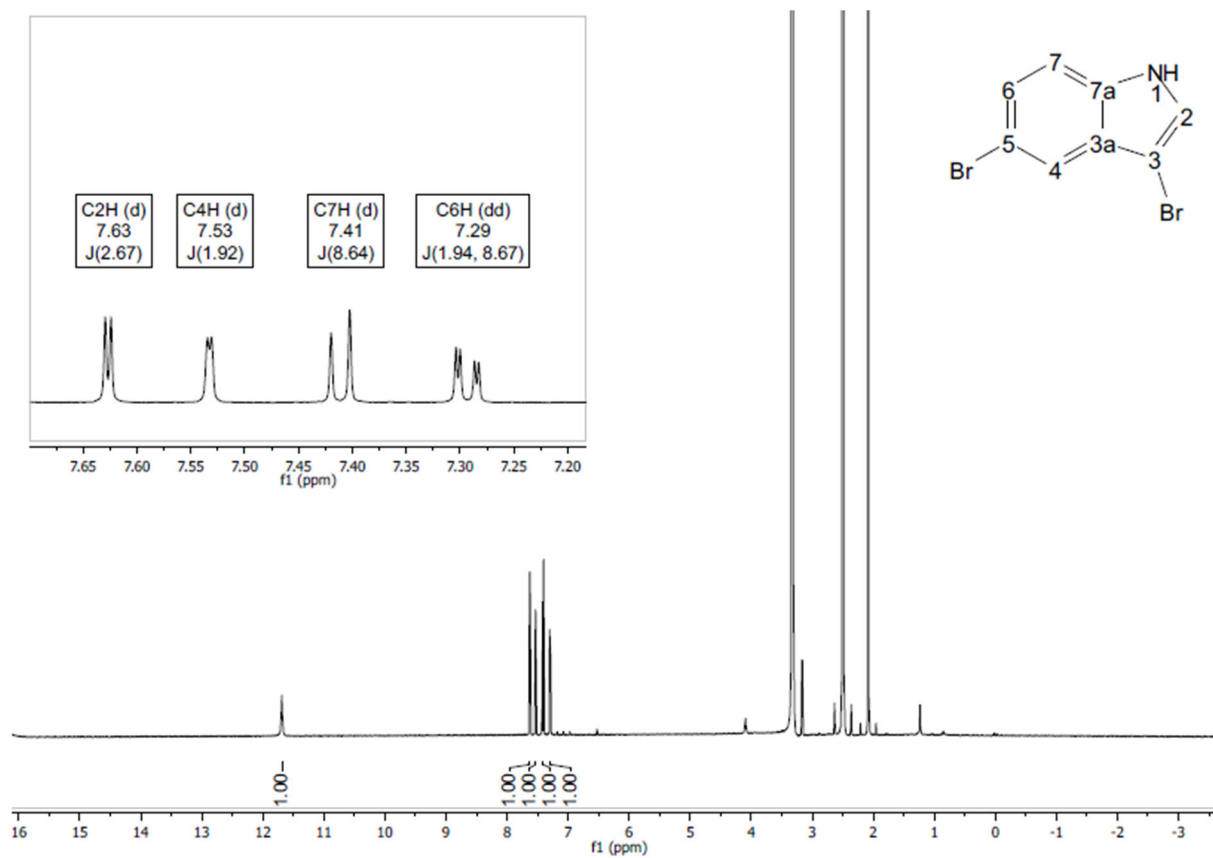
7: HMQC



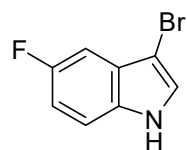
7 Catalyzed by SpH2

Yield: 0.6%, 0.004 mmol, 1.2 mg

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ [ppm]: 11.69 (s, 1H, N^1H), 7.63 (d, $^3J=2.7$ Hz, 1H, C^2H), 7.53 (d, $^4J=1.9$ Hz, 1H, C^4H), 7.41 (d, $^3J=8.6$ Hz, 1H, C^7H), 7.29 (dd, $^3J=8.7$ Hz, $^4J=1.9$ Hz, 1H, C^6H).



3-Bromo-5-fluoroindole 8

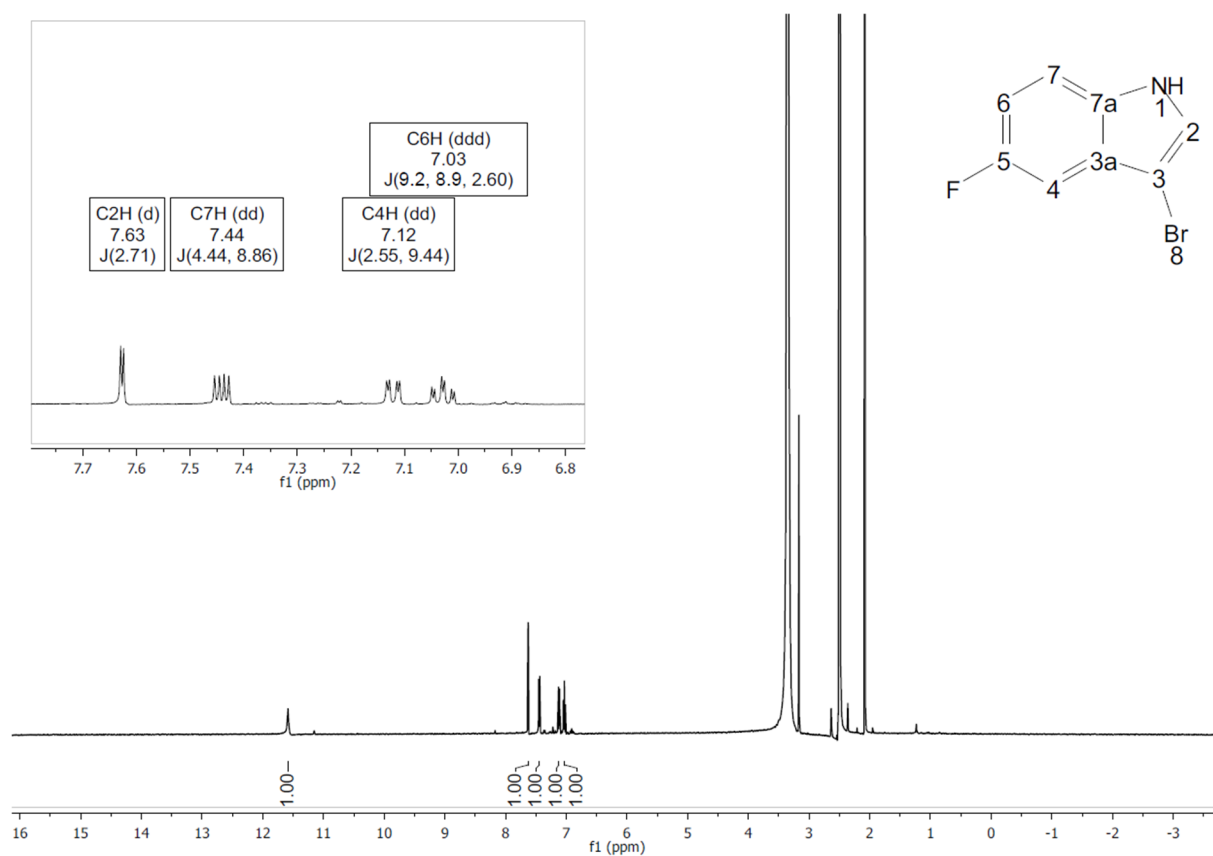


8 Catalyzed by SpH2

Yield: 0.7%, 0.005 mmol, 1.1 mg

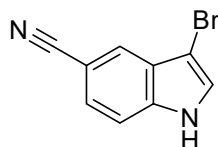
[C₈H₄BrFN] *m/z*: calculated: 211.96 (⁷⁹Br), 213.96 (⁸¹Br), 276.87; observed: LC-ESI-MS [M-H]: 212.0, 214.0

¹H NMR (500 MHz, DMSO-*d*₆) δ [ppm]: 11.58 (s, 1H, N¹H), 7.63 (d, ³J= 2.7 Hz, 1H, C²H), 7.44 (dd, ³J= 8.9, ⁴J_{HF}= 4.4 Hz, 1H, C⁷H), 7.12 (dd, ³J_{HF}= 9.4, ⁴J= 2.6 Hz, 1H, C⁴H), 7.03 (ddd, ³J_{HF}= 9.2, ³J= 8.9, ⁴J= 2.6 Hz, 1H, C⁶H).



3-Bromo-5-cyanoindole **9**

HRMS: $C_9H_4BrN_2$ m/z calculated: 218.9563; observed: 218.9551

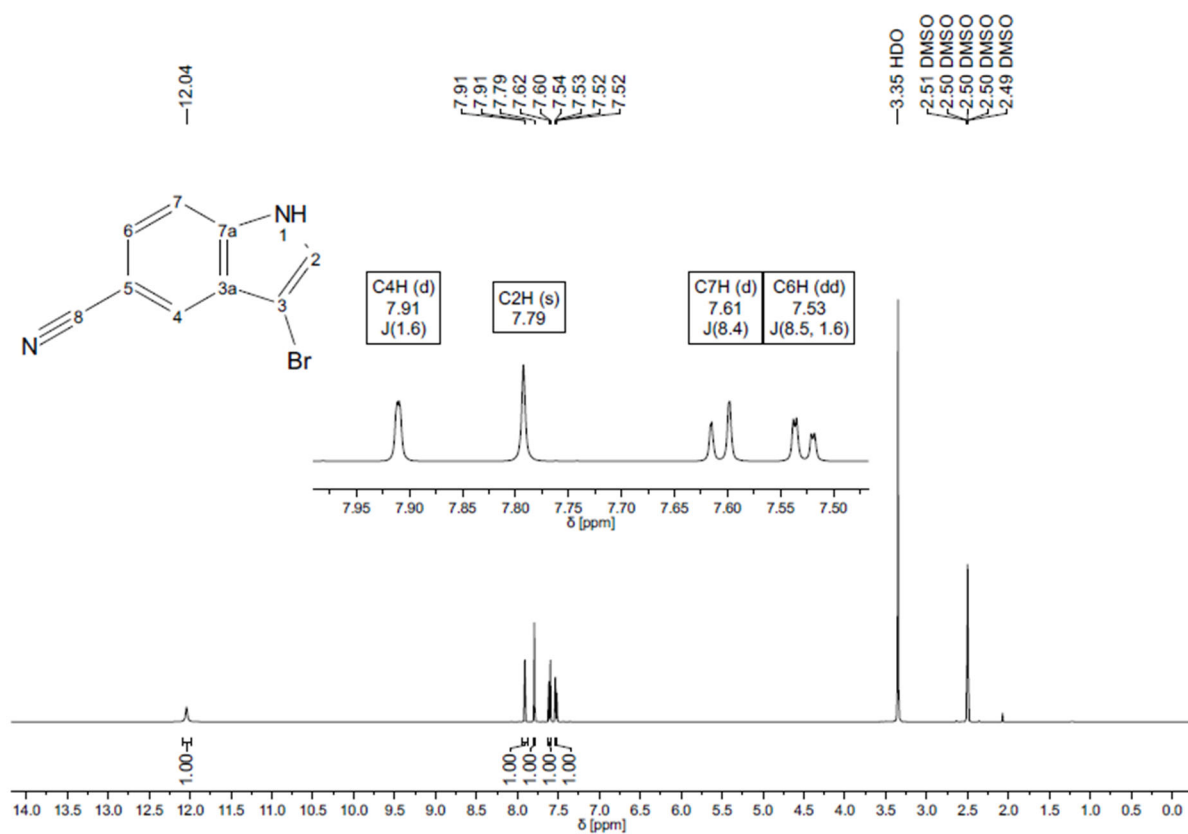


9 Catalyzed by SpH1

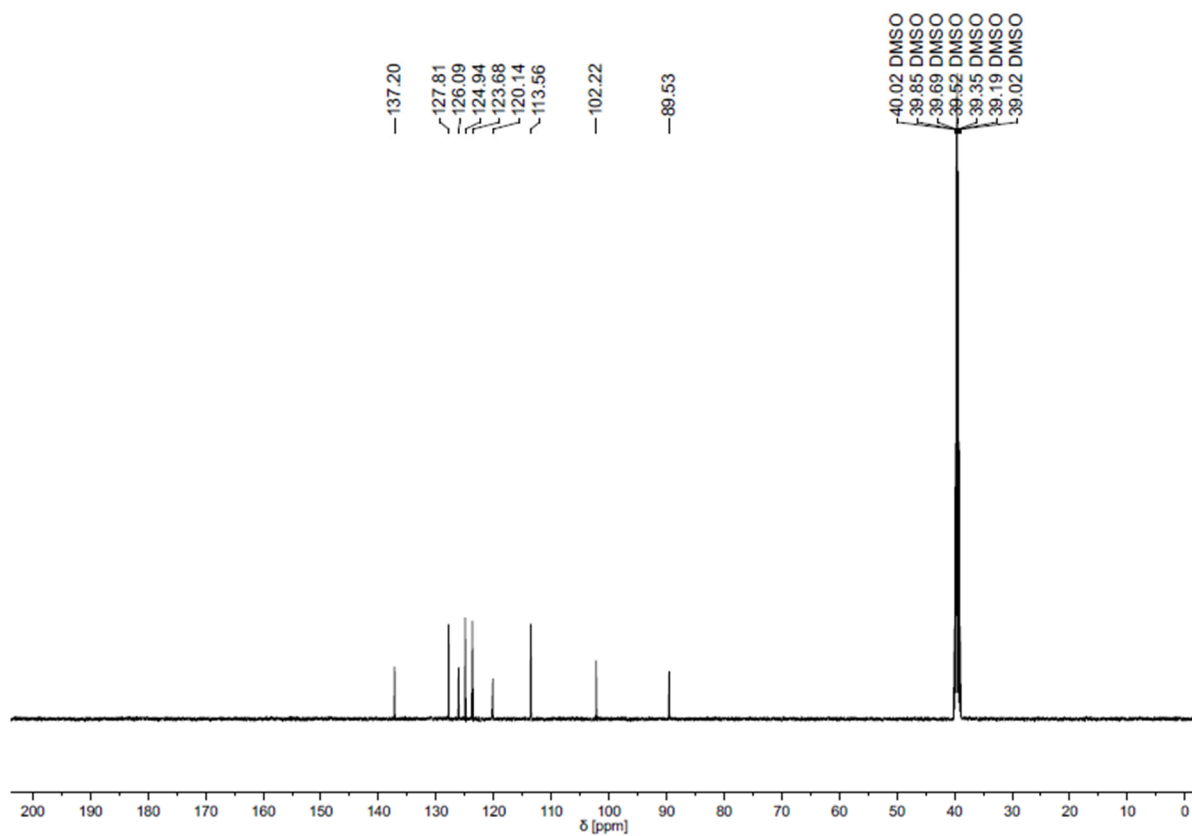
Yield: 37.8%, 0.282 mmol, 62 mg

1H NMR (500 MHz, $DMSO-d_6$) δ 12.04 (s, NH), 7.91 (d, $J = 1.6$ Hz, C⁴H), 7.79 (s, C²H), 7.61 (d, $J = 8.4$ Hz, C⁷H), 7.53 (dd, $J = 8.5, 1.6$ Hz, C⁶H).

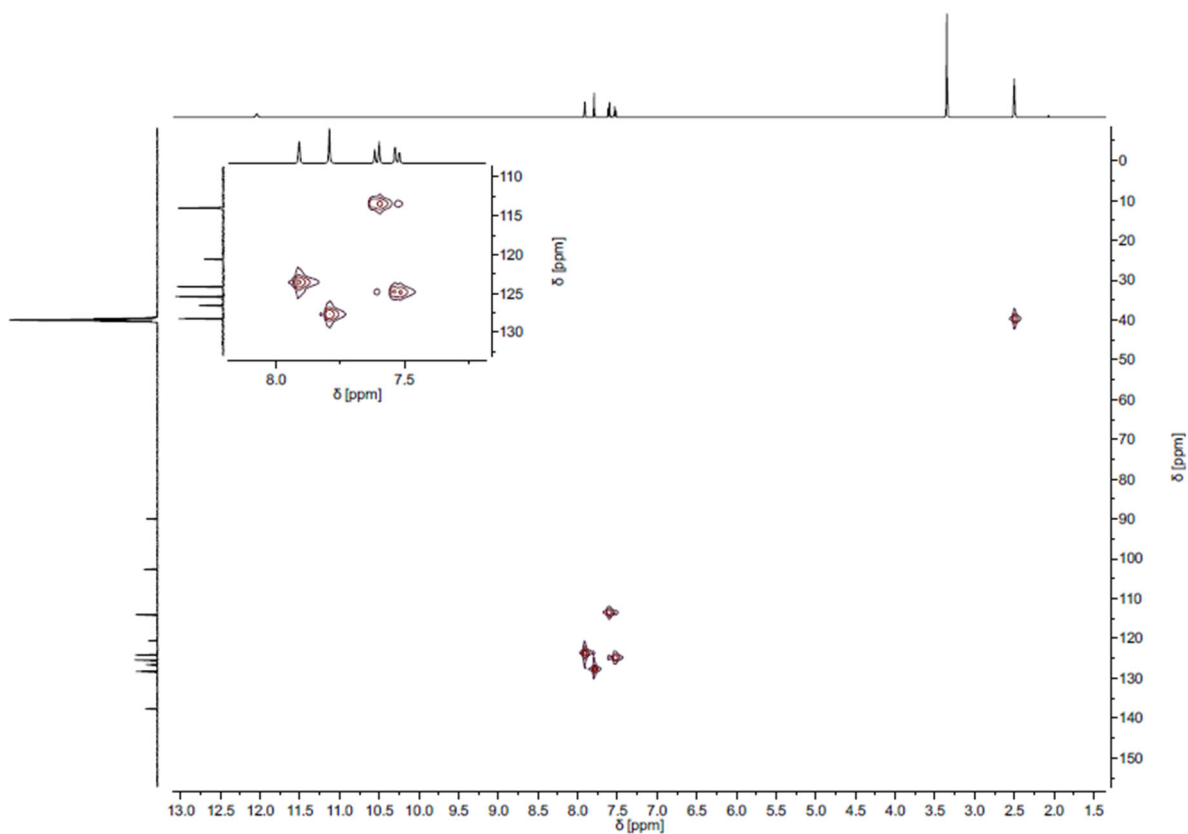
^{13}C NMR (126 MHz, $DMSO-d_6$) δ 137.2 (C^{7a}), 127.8 (C^{3a}), 126.1 (C⁴H), 124.9 (C⁶H), 123.7 (C²H), 120.1 (C^{8N}), 113.6 (C⁷H), 102.2 (C⁵CN), 89.5 (C³Br).



9: ¹³C NMR



9: HMQC

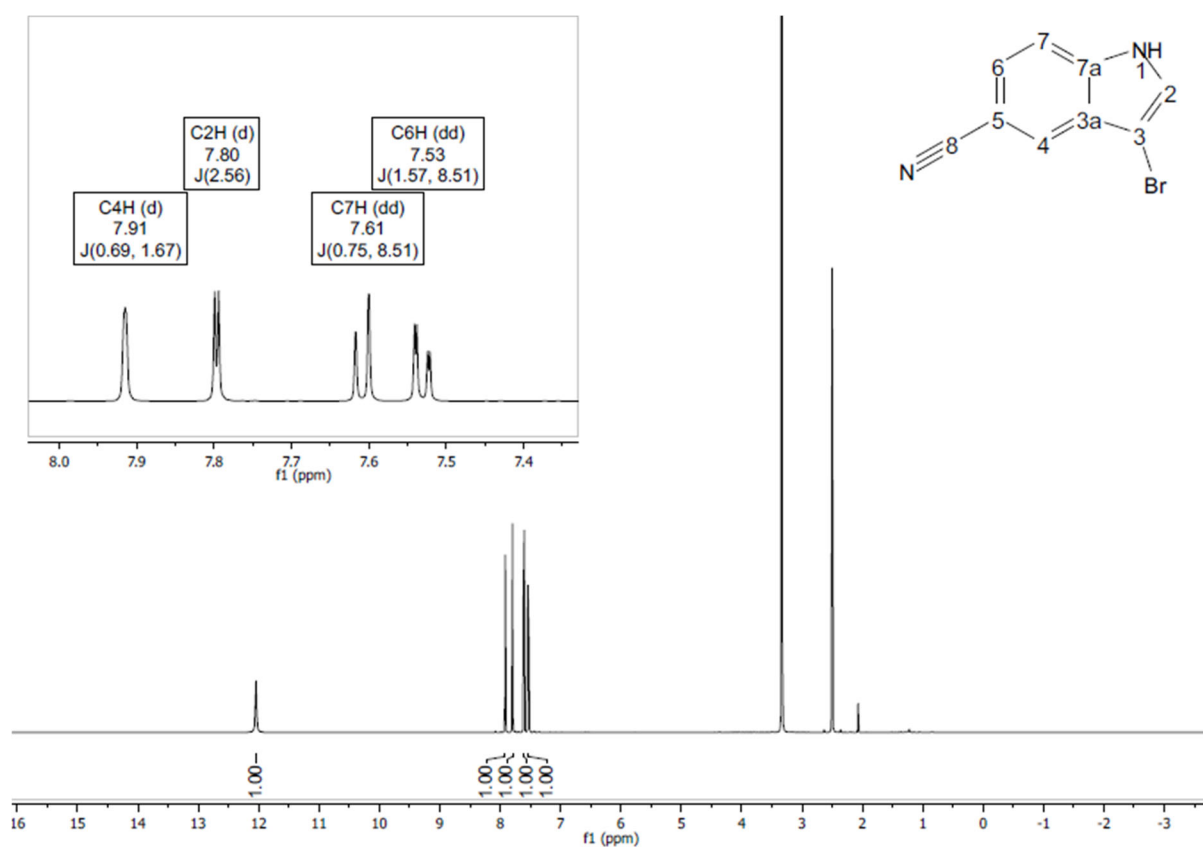


9 Catalyzed by SpH2

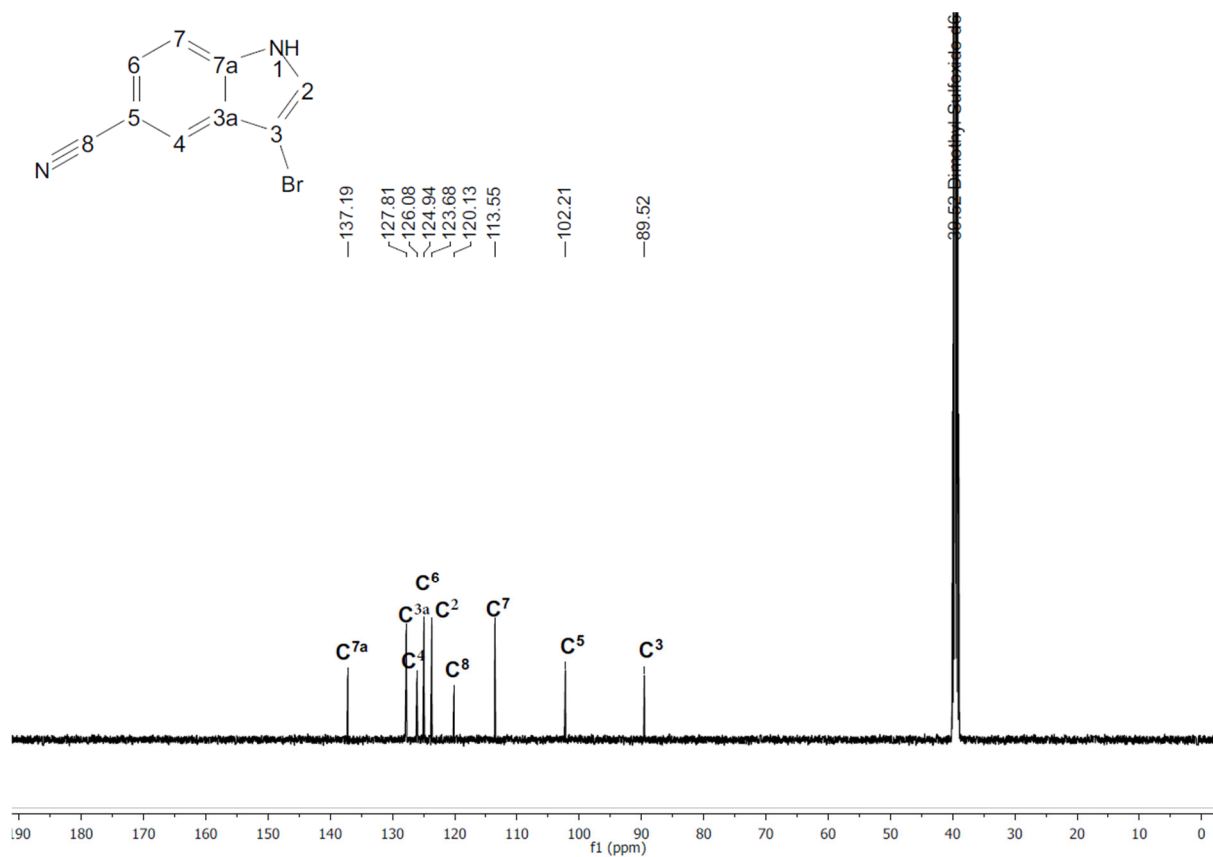
Yield: 4.0%, 0.030 mmol, 6.6 mg

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ [ppm]: 12.04 (s, 1H, N^1H), 7.91 (d, $^4J=1.7$ Hz, $^5J=0.7$ Hz, 1H, C^4H), 7.80 (d, $^3J=2.6$ Hz, 1H, C^2H), 7.61 (dd, $^3J=8.5$ Hz, $^5J=0.8$ Hz, 1H, C^7H), 7.53 (dd, $^3J=8.5$ Hz, $^4J=1.6$ Hz, 1H, C^6H).

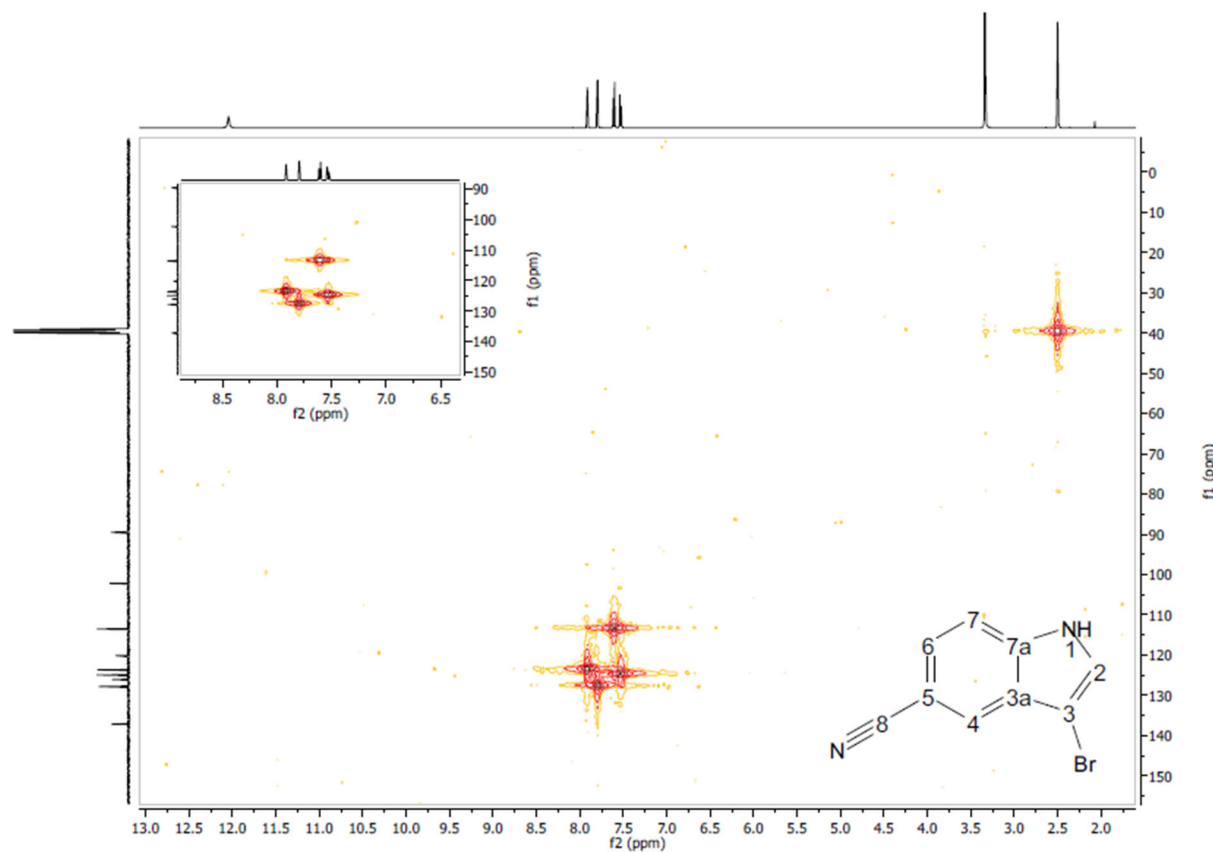
^{13}C NMR (126 MHz, $\text{DMSO-}d_6$) δ 137.19 (C^7a), 127.8 (C^3a), 126.1 (C^4H), 124.9 (C^6H), 123.7 (C^2H), 120.1 (C^8N), 113.6 (C^7H), 102.2 (C^5CN), 89.5 (C^3Br).



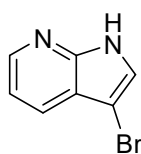
9: ¹³C NMR



9: HMQC



3-Bromo-7-azaindole **15**



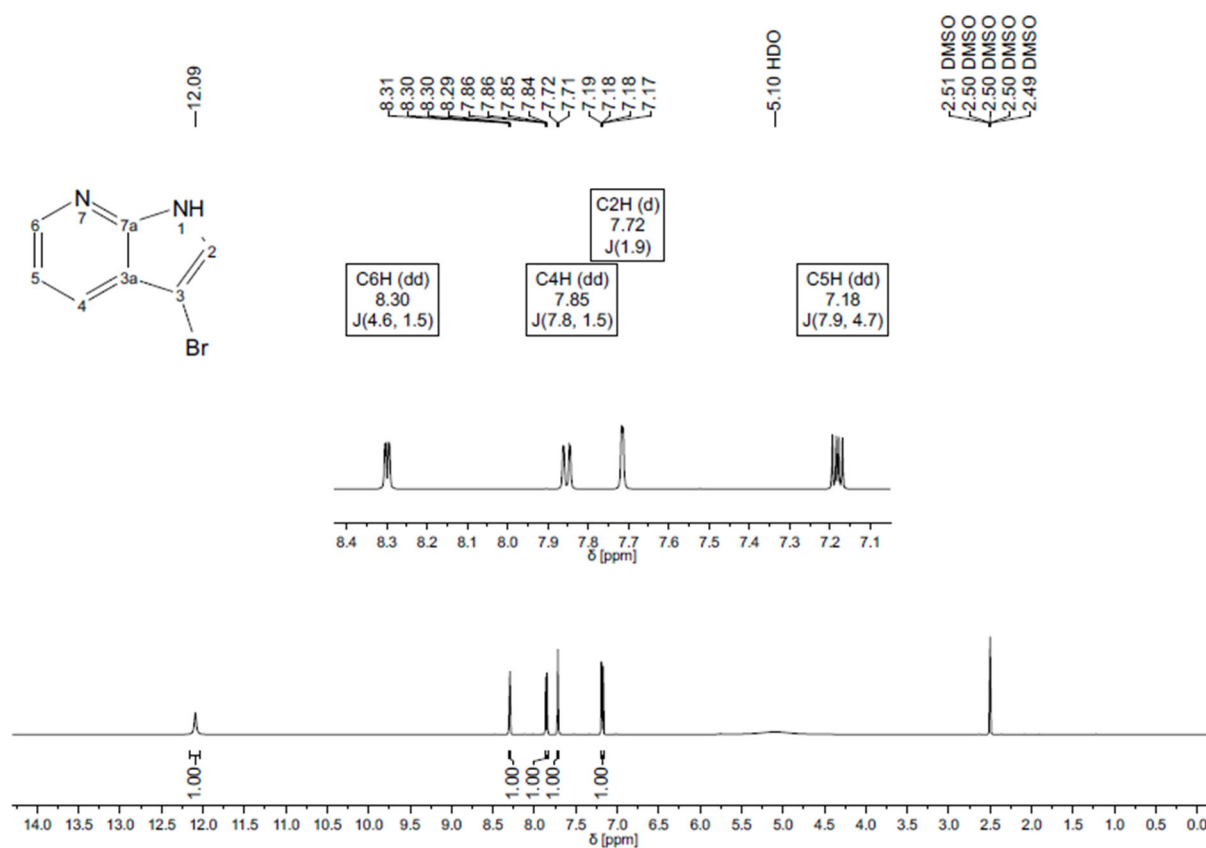
15 Catalyzed by SpH1

Yield: 7.4%, 0.056 mmol, 11 mg

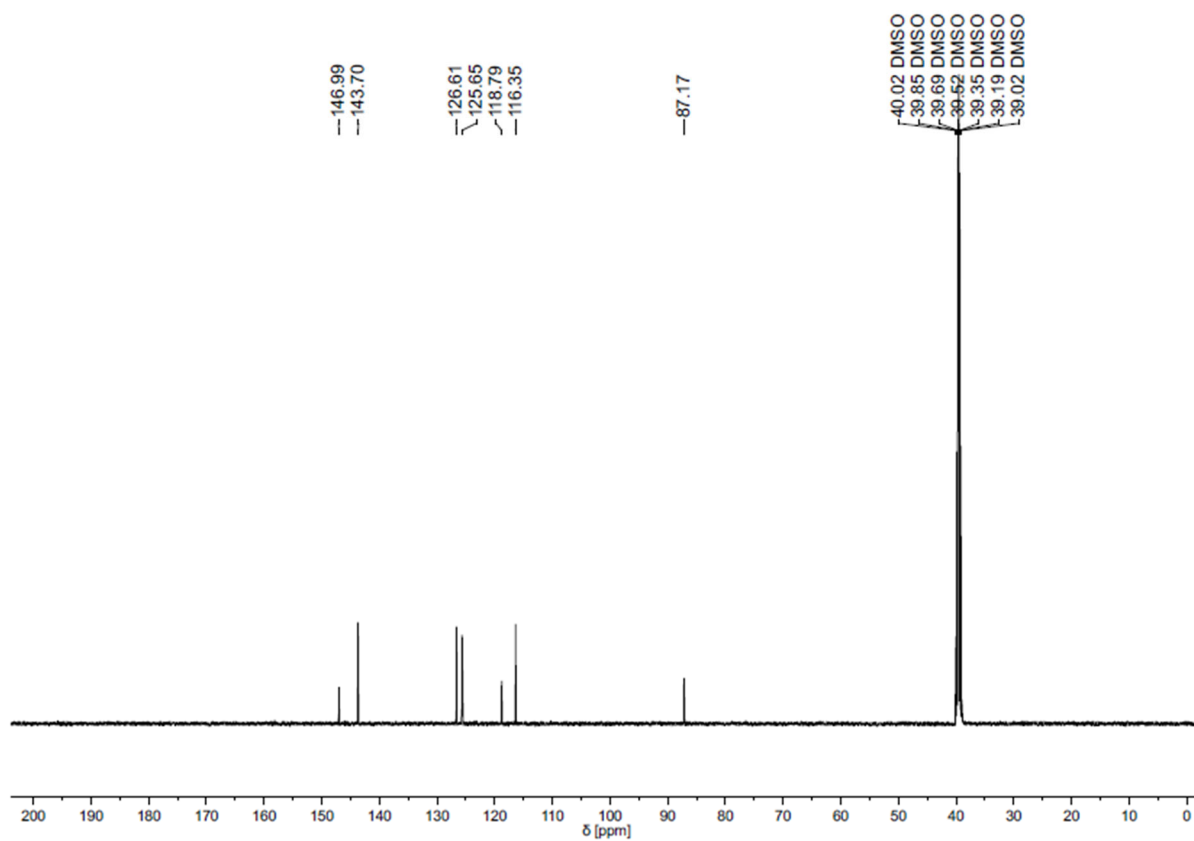
$[C_7H_4BrN_2]^+$ m/z : calculated: 196.96 (^{79}Br), 198.96 (^{81}Br); observed: LC-ESI-MS $[M+H]^+$: 196.97, 198.97

1H NMR (500 MHz, DMSO- d_6) δ 12.09 (s, NH), 8.30 (dd, $J = 4.6, 1.5$ Hz, C 6 H), 7.85 (dd, $J = 7.8, 1.5$ Hz, C 4 H), 7.72 (d, $J = 1.9$ Hz, C 2 H), 7.18 (dd, $J = 7.9, 4.7$ Hz, C 5 H).

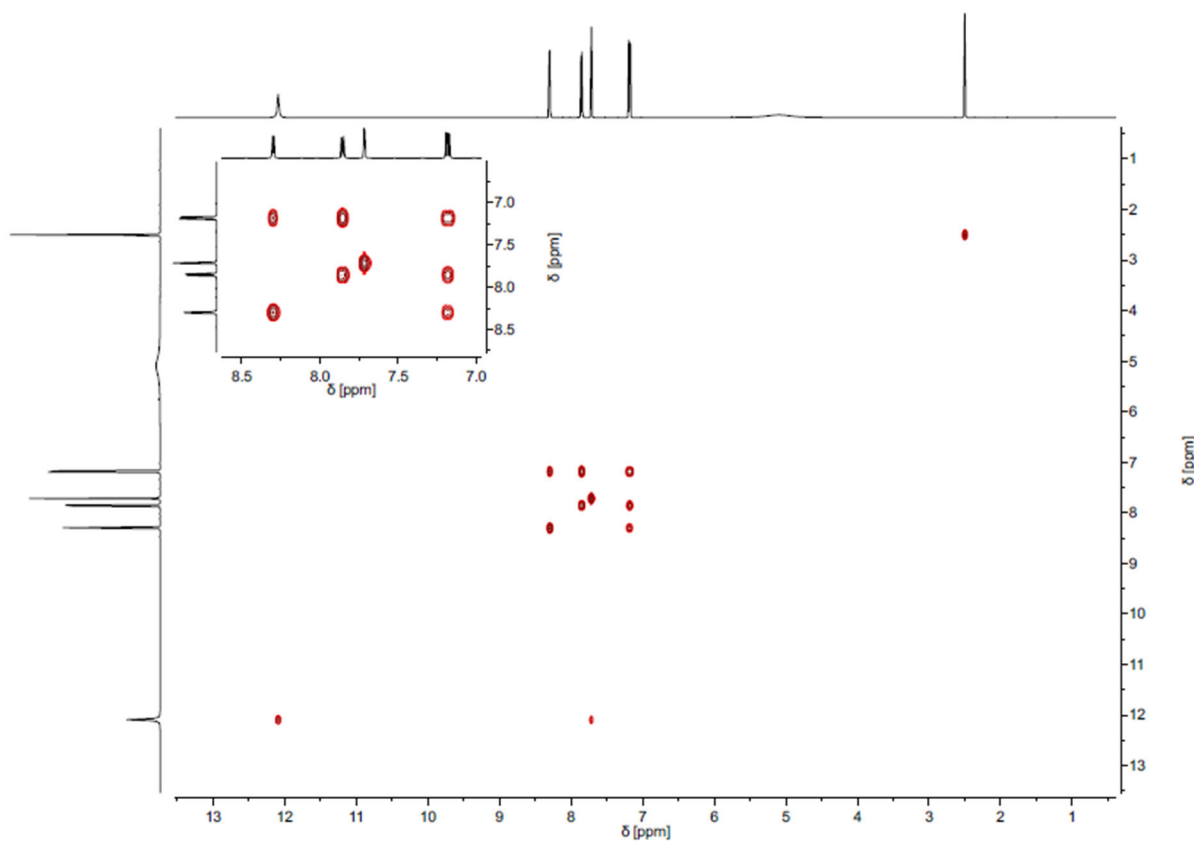
^{13}C NMR (126 MHz, DMSO- d_6) δ 147.0 (C 7a), 143.7 (C 6H), 126.6 (C 4H), 125.7 (C 2H), 118.8 (C 3a), 116.4 (C 5H), 87.2 (C 3Br).



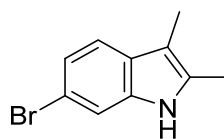
15: ¹³C NMR



15: ¹H, ¹H-COSY



6-Bromo-2,3-dimethylindole **23**



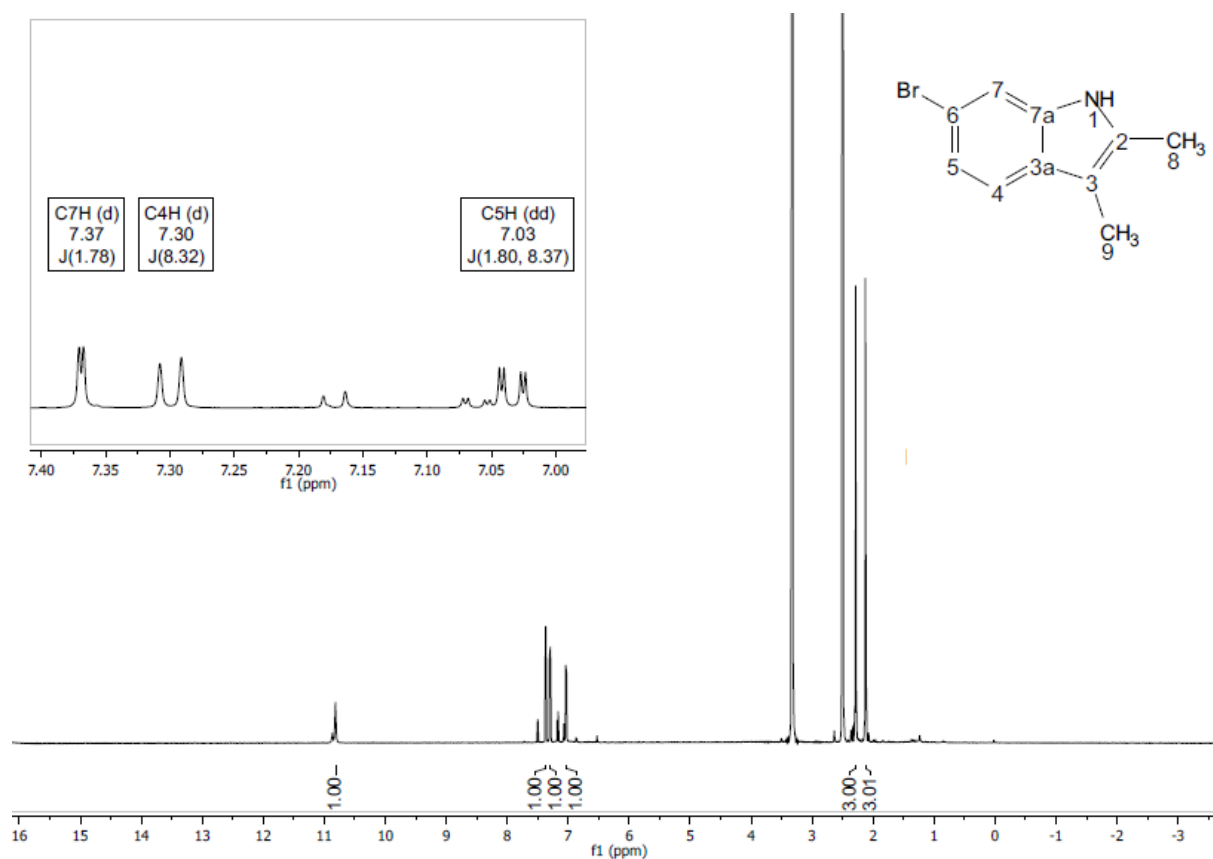
23 Catalyzed by SpH2

Yield: 0.8%, 0.006 mmol, 1.4 mg

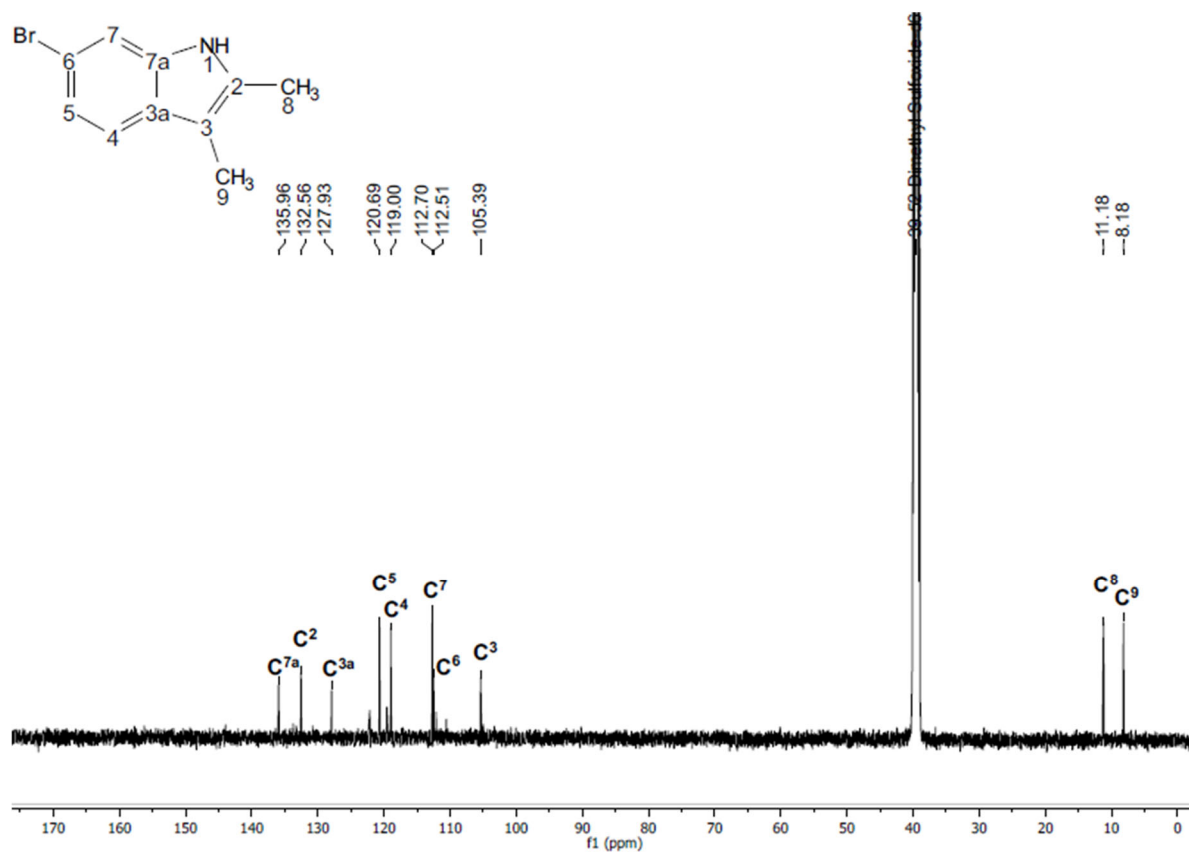
[C₁₀H₉BrN]⁺ *m/z*: calculated: 222.0 (⁷⁹Br), 224.0 (⁸¹Br); observed: LC-ESI-MS [M-H]⁺: 222.0, 224.0

¹H NMR (500 MHz, DMSO-*d*₆) δ [ppm]: 11.69 (s, 1H, N¹H), 7.37 (d, ⁴J = 1.8 Hz, 1H, C⁷H), 7.30 (d, ³J = 8.3 Hz, 1H, C⁴H), 7.03 (dd, ³J = 8.4 Hz, ⁴J = 1.8 Hz, 1H, C⁵H), 2.28 (s br, 3H, CH₃); 2.13 (s br, 3H, CH₃)

¹³C NMR (126 MHz, DMSO-*d*₆) δ 136.0 (C^{7a}), 132.6 (C²CH₃), 127.9 (C^{3a}), 120.7 (C⁵H), 119.0 (C⁴H), 112.7 (C⁷H), 112.5 (C⁶Br), 105.4 (C³CH₃), 11.2 (C⁸H₃), 8.2 (C⁹H₃).



23: ¹³C NMR



List of detected secondary metabolite biosynthetic gene clusters within the MAG 21

The dataset includes the identified secondary metabolite biosynthetic gene clusters encoded within the metagenome assembled genome 21 (MAG21) of the *B. braunii* metagenome sequence data, resulting from the combined metagenome *de novo* assembly and differential coverage and tetra-nucleotide signature based binning.^[2] The analysis was accomplished using the antibiotics & Secondary Metabolite Analysis Shell (antiSMASH, Version 4.1.0).^[3]

Cluster	Input	Type	From	To	Most similar known cluster
Cluster 1	contig-13151000046	Cf_putative	3	7095	-
Cluster 2	contig-46000040	Nrps	1	3078	-
Cluster 3	contig-12860000038	Hserlactone	1	3272	-
Cluster 4	contig-15000000	Cf_putative	3	6624	-
Cluster 5	contig-54000025	Cf_putative	1513	8377	-
Cluster 6	contig-45000027	Cf_fatty_acid	1	4853	-
Cluster 7	contig-529000008	Cf_saccharide	1	4279	-
Cluster 8	contig-446000012	Cf_putative	3	4189	-
Cluster 9	contig-354000029	Cf_putative	3	8222	-
Cluster 10	contig-29000042	Terpene	1	11734	-
Cluster 11	contig-12868000026	Cf_putative	11232	15395	-
Cluster 12	contig-13030000033	Cf_putative	31	13103	-
Cluster 13	contig-3639000000	Cf_fatty_acid	1	5305	-
Cluster 14	contig-29000047	Terpene	1	6928	Astaxanthin_dideoxyglycoside_biosynthetic_gene_cluster (75% of genes show similarity; BGC0001086_c1)
Cluster 15	contig-12916000015	Hserlactone	1	7196	-
Cluster 16	contig-76000017	Cf_putative	2	2990	-
Cluster 17	contig-2211000019	Cf_putative	3	4297	-
Cluster 18	contig-15000041	Siderophore	1	6752	-

Cluster	Input	Type	From	To	Most similar known cluster
Cluster 19	contig-14000013	T1pks	1	10855	-
Cluster 20	contig-620000022	Cf_fatty_acid	1	5844	-
Cluster 21	contig-158000018	Cf_putative	3	6408	-
Cluster 22	contig-12842000026	Cf_putative	196	7594	-
Cluster 23	contig-270000028	Cf_putative	106	7141	-
Cluster 24	contig-268000031	Cf_putative	2	3593	Galactoglucan_biosynthetic_gene_cluster (12% of genes show similarity; BGC0000801_c1)
Cluster 25	contig-1091000038	Cf_fatty_acid	1	5482	-
Cluster 26	contig-380000002	Cf_putative	2	12531	Sphingan_polysaccharide_biosynthetic_gene_cluster (17% of genes show similarity; BGC0000797_c1)
Cluster 27	contig-12844000026	Cf_saccharide	1	16114	-
Cluster 28	contig-20000044	Cf_putative	6400	18603	-
Cluster 29	contig-112000000	Cf_putative	285	6877	-
Cluster 30	contig-370000027	Cf_fatty_acid	1	8972	-
Cluster 31	contig-750000041	Cf_fatty_acid	1	12158	Asukamycin_biosynthetic_gene_cluster (4% of genes show similarity; BGC0000187_c1)
Cluster 32	contig-979000022	Cf_fatty_acid	4189	17707	-
Cluster 33	contig-213000021	Bacteriocin	1	5155	-
Cluster 34	contig-280000007	Cf_putative	1733	7143	-
Cluster 35	contig-12660000040	Cf_putative	6339	13208	-
Cluster 36	contig-12649000039	Cf_putative	4224	19885	Emulsan_biosynthetic_gene_cluster (9% of genes show similarity; BGC0000760_c1)
Cluster 37	contig-13059000033	Cf_fatty_acid	1	14860	-
Cluster 38	contig-13194000046	Cf_putative	2	9476	-
Cluster 39	contig-250000042	Cf_putative	13360	25949	Azinomycin_B_biosynthetic_gene_cluster (4% of genes show similarity; BGC0000960_c1)
Cluster 40	contig-624000005	Cf_putative	328	4559	-

Cluster	Input	Type	From	To	Most similar known cluster
Cluster 41	contig-90000034	Cf_saccharide	1	3955	-
Cluster 42	contig-1491000034	Cf_putative	3	8288	-
Cluster 43	contig-12892000015	Cf_putative	14216	24750	-
Cluster 44	contig-674000020	Cf_putative	3	4752	-
Cluster 45	contig-311000004	Other	1	4111	-
Cluster 46	contig-170000011	Cf_putative	335	8704	-
Cluster 47	contig-121000046	Cf_fatty_acid	1	9103	-
Cluster 48	contig-8000031	Lassoptide	1	8016	-
Cluster 49	contig-254000002	Bacteriocin	1	4268	-
Cluster 50	contig-13168000046	Bacteriocin	5160	13484	-
Cluster 51	contig-472000023	Cf_putative	2	6320	Lipopolysaccharide_biosynthetic_gene_cluster (8% of genes show similarity; BGC0000774_c1)
Cluster 52	contig-156000028	Cf_putative	2	3814	-
Cluster 53	contig-290000000	Cf_putative	1	6658	-
Cluster 54	contig-71000021	Cf_fatty_acid	1	11491	-
Cluster 55	contig-81000021	Cf_putative	609	13474	-
Cluster 56	contig-137000046	Cf_putative	3	7962	-
Cluster 57	contig-1344000023	Cf_putative	437	7335	Lipopolysaccharide_biosynthetic_gene_cluster (5% of genes show similarity; BGC0000774_c1)
Cluster 58	contig-37000040	Cf_putative	2	4429	-
Cluster 59	contig-1262000037	Cf_putative	1	4232	-
Cluster 60	contig-12559000022	Cf_fatty_acid	1	8107	-
Cluster 61	contig-453000026	Cf_putative	1	4304	O&K-antigen_biosynthetic_gene_cluster (3% of genes show similarity; BGC0000780_c1)
Cluster 62	contig-13252000030	Ectoine	15769	26170	Ectoine_biosynthetic_gene_cluster (80% of genes show similarity; BGC0000857_c1)

Cluster	Input	Type	From	To	Most similar known cluster
Cluster 63	contig-18000018	Cf_putative	3188	6860	-
Cluster 64	contig-1317600000	Cf_putative	3	7349	Xenoclyoins_biosynthetic_gene_cluster (25% of genes show similarity; BGC0000189_c1)
Cluster 65	contig-688000009	Cf_putative	5027	9594	-
Cluster 66	contig-200000007	Cf_putative	423	15643	-
Cluster 67	contig-261000005	Cf_putative	9556	20797	-
Cluster 68	contig-50000033	Cf_putative	2	13098	-
Cluster 69	contig-100000032	Bacteriocin	1218	6759	-
Cluster 70	contig-12725000036	Cf_putative	933	15577	-
Cluster 71	contig-12782000009	Other	1	5882	-

Area and volume of the active site pocket for Apo SpH2 and SpH1 and docked with indole and 3-bromoindole before and after MD simulation

Apoenzyme		
	Cavity/Pocket Area (Å ²)	
SpH2	136.80	
SpH1	77.39	
Enzyme / docked indole		
	Area (Å ²)	Volume (Å ³)
SpH2	45.45	23.68
SpH1	77.39	15.60
Enzyme / docked indole / 1ns MD simulation		
	Area (Å ²)	Volume (Å ³)
SpH2	197.12	97.62
SpH1	115.23	43.54
Enzyme/ docked 3-Br-indole / 1ns MD simulation		
	Area (Å ²)	Volume (Å ³)
SpH2	342.65	173.61
SpH1	11.28	2.72

References

- [1] U. K. Laemmli, *Nature* **1970**, 227, 680–685.
- [2] O. Blifernez-Klassen, V. Klassen, D. Wibberg, E. Cebeci, C. Henke, C. Rückert, S. Chaudhari, O. Rupp, J. Blom, A. Winkler, *Sci Rep.* **2021**, 11, 1–13.
- [3] K. Blin, T. Wolf, M. G. Chevrette, X. Lu, C. J. Schwalen, S. A. Kautsar, H. G. Suarez Duran, de Los Santos, Emmanuel LC, H. U. Kim, M. Nave, *Nucleic Acids Res.* **2017**, 45, W36-W41.