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 Sewald¹,*

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 via 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 μM, 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 μL/min

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Eluent A</th>
<th>Eluent B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>95%</td>
<td>5%</td>
</tr>
<tr>
<td>5.5</td>
<td>5%</td>
<td>95%</td>
</tr>
<tr>
<td>6</td>
<td>5%</td>
<td>95%</td>
</tr>
<tr>
<td>6.1</td>
<td>95%</td>
<td>5%</td>
</tr>
<tr>
<td>9</td>
<td>95%</td>
<td>5%</td>
</tr>
</tbody>
</table>

LC-ESI-MS

LC-ESI-MS was performed using Agilent Technologies 1200 with C18- Hypersil Gold C18 column (3 μM, 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 μL/min

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Eluent A</th>
<th>Eluent B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>10</td>
<td>2%</td>
<td>98%</td>
</tr>
<tr>
<td>11</td>
<td>2%</td>
<td>98%</td>
</tr>
<tr>
<td>11.5</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>15</td>
<td>100%</td>
<td>0%</td>
</tr>
</tbody>
</table>
**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 μL. For the analytics a Trace GC Ultra gas chromatography from ThermoScientific with VF-5 column (30 m x 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 x 2.1 mm, 1.9 μm) 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, 13C: 126 MHz) or a Bruker DRX-500 spectrometer (1H: 500 MHz, 13C: 126 MHz). Chemical shifts are reported relative to residual solvent peaks in parts per million (ppm) (DMSO-d6: 1H: 2.5 ppm; 13C: 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/μL of vector DNA (spH1_pET28-a), 125 ng of each primer, 2.5 U/mL Pfu polymerase (Promega, Germany), dNTP mix (200 μM of each nucleotide) and 1x reaction buffer were added in PCR solution. The PCR program is shown in the following table:

<table>
<thead>
<tr>
<th>Cycles</th>
<th>Temperature [°C]</th>
<th>Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
<td>0.5</td>
</tr>
<tr>
<td>16</td>
<td>95</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>68</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>68</td>
<td>5</td>
</tr>
</tbody>
</table>

Parental DNA was digested with DpnI restriction enzyme (10 U/μL) for 1h at 37°C and transformed in *E. coli* DH5α. 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

MKHGGMAGIERRVIVGAGTAGWAAAAALGAYLAGAGTRITLIESSEIGTIVGGEATIP
TIRRFYAALGMTDAEVMRACEATAKLGIRFVWKPGTSFVHPFGRFQGDLRGLIDFHYW
QKARQAGRAAPEEYSGLAMLAREGHASVLPNPPSQLSVDFWALHFAARFAAHMRA
YAERAGVARIDARITDVTLHTGETFIAETLSDGVRAGDLFVDGSRGLREGIALTGSYR
DWSHWLCDGAFAVQERVGEPPSCTVTARTAGWQWRPLRSREGNGLVFSADSFQGD
EARAELANIPGAPTMEPRRLRTFPGREVAWARNCVSLGLASGFLPLESTSAIETGIER
LKQLFPDKGFDPAVIAEYNAQSAEEMERVDFILLHYHLSTRDGPFWQPACREMTLPDSL
AKLELWRARGAFRLRYWEWFSPAWSLAIYGGFEHLPERLDPGVAAVGTDELAQGLEQMR
AAIARTVADTPHTSEFLATVGDGAAEPMRTSA*

>SpH2

MNNPDARLKVCHRIVGGGTAGWIAAAVMAHMKRLFEIELVESDDITIGVGETISSPME
LIKLIDILDQDFVRQVASYFKCDEYFHPFQGVRVELSDFYQCRWATL
AGHDFFPQLDFAPATVMHAGRFLMPFKARNTPIAGASYALHVDAKRVAQFRLGAHEAR
GVKRTEIGVDVALDARGFVTLTLDKDGRTIDADFFDSCFRALLIEKALGSGYTDWSNL
FCDRAIAAQTVGPPRPTLAEAQADAGWRIPLQHRTGNGHYSFSDEATRILL
EKVEEVEVAGPMVVFRTGVRKIHWRNLAIGLSSGFIEPLESTAIHLYRGMDFRFRLIP
DRHCDPALADEYNRRMVADYTEIRDINLHYCITTARRDTPFWRKCRDMWDPSLKERVE
LFRSNGTLREGDFRSVSWSVLEGMRGPRYHPLVDREAMELYAGLDKAREQLAGF
VRQLPTQQDFIDAHCRAEKEVCLTRPVPMAGV*

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 SpH2.

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 [C₈H₆NBr] m/z 194.97, 196.97

Dibromoindole: theoretical mass [C₈H₅NBr₂] 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 [C₉H₈BrN] m/z 208.98, 210.98

Dibromo-2-methylindole: theoretical mass [C₉H₅Br₂N] m/z 288.89, 286.89, 290.89

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

2-Bromo-3-methylindole

theoretical mass [C9H8BrN] m/z 208.98, 210.98

Dibromo-3-methylindole: theoretical mass [C9H8Br2N] 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): 35%; LC-ESI-MS [M-H]: 208.02, 210.02

3-Bromo-5-methylindole

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


3-Bromo-5-nitroindole

theoretical mass [C9H5BrN2O2] 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

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

theoretical mass [C8H5Br2N] m/z 274.88, 272.88


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

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

Bromo-2-methyl-5-nitroindole

theoretical mass \([\text{C}_9\text{H}_7\text{BrN}_2\text{O}_2]\) \(m/z\) 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 \([\text{C}_7\text{H}_5\text{BrN}_2]\) \(m/z\) 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 \([\text{C}_6\text{H}_4\text{BrN}_2]\) \(m/z\) 196.96, 198.96

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

Bromoazulene

theoretical mass \([\text{C}_{10}\text{H}_7\text{Br}]\) \(m/z\) 205.97, 207.97

Dibromoazulene: theoretical mass \([\text{C}_{10}\text{H}_6\text{Br}_2]\) \(m/z\) 285.88, 283.88, 287.88


Bromo-4-n-hexylresorcinol

theoretical mass \([\text{C}_{12}\text{H}_{19}\text{BrNO}_2]\) \(m/z\) 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 \([\text{C}_{10}\text{H}_8\text{BrNO}_2]\) \(m/z\) 252.97, 254.97

SpH1: Conversion (RP-HPLC): 0%

Bromotryptophol

Theoretical mass $[\text{C}_{10}\text{H}_{10}\text{BrNO}] m/z$ 238.99, 240.99

Dibromotryptophol: Theoretical mass $[\text{C}_{10}\text{H}_{9}\text{Br}_{2}\text{NO}] m/z$ 318.9, 316.91, 320.9


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 $[\text{C}_{11}\text{H}_{10}\text{BrNO}_{2}] m/z$ 266.99, 268.99

Dibromo-indole-3-propionic acid: Theoretical mass $[\text{C}_{11}\text{H}_{9}\text{Br}_{2}\text{NO}_{2}] m/z$ 346.9, 344.9, 348.9


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

Bromoquinoxaline

Theoretical mass $[\text{C}_{8}\text{H}_{5}\text{BrN}_{2}] m/z$ 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 $[\text{C}_{10}\text{H}_{10}\text{BrN}] m/z$ 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 $[\text{C}_{4}\text{H}_{4}\text{BrO}] m/z$ 171.95, 173.95


Bromoanthranilic acid

Theoretical mass $[\text{C}_{7}\text{H}_{6}\text{BrNO}_{2}] m/z$ 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 tR 7.3 min; P2 tR: 7.6 min)

**Determination of regioselective halogenation site via NMR analysis**

3-Bromoindole 1

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

\[
\begin{align*}
\text{Br} \\
& \\
\text{H} \\
& \\
\text{N}
\end{align*}
\]

Catalyzed by SpH1

Yield: 35.6%, 0.267 mmol, 52.05 mg

\(^1\text{H} \text{NMR} \ (500 \text{ MHz, DMSO-}d_6) \delta 11.47 \ (s, \text{NH}), \ 7.54 \ (d, \ J = 2.6 \text{ Hz, C}^2\text{H}), \ 7.44 \ (dd, \ J = 8.1, 0.9 \text{ Hz, C}^3\text{H}), \ 7.41 \ (dd, \ J = 7.9, 1.1 \text{ Hz, C}^4\text{H}), \ 7.18 \ (ddd, \ J = 8.2, 7.0, 1.3 \text{ Hz C}^5\text{H}), \ 7.11 \ (ddd, \ J = 7.9, 7.0, 1.0 \text{ Hz, C}^5\text{H}).
\]

\(^{13}\text{C} \text{NMR} \ (126 \text{ MHz, DMSO-}d_6) \delta 135.4 \ (C^7\text{a}), \ 126.1 \ (C^3\text{a}), \ 124.8 \ (C^2\text{H}), \ 122.2 \ (C^6\text{H}), \ 119.9 \ (C^5\text{H}), \ 117.9 \ (C^4\text{H}), \ 112.1 \ (C^7\text{H}), \ 88.7 \ (C^3\text{H})
1: $^{13}$C NMR
1: HMQC
1 Catalyzed by SpH2

Yield: 10.7%, 0.080 mmol, 15.7 mg

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

$^{13}$C NMR (126 MHz, DMSO-$d_6$) δ: 135.4 (C$^7a$), 126.1 (C$^3a$), 124.8 (C$^2$H), 122.2 (C$^5$H), 119.9 (C$^6$H), 117.9 (C$^4$H), 112.1 (C$^7$H), 88.7 (C$^3$Br).
1: $^{13}$C 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, CH$_3$).

$^{13}$C NMR (126 MHz, DMSO-d$_6$) δ 134.8 (C7a), 133.2 (C2CH3), 126.9 (C3a), 121.4 (C6H), 119.6 (C5H), 117.1 (C4H), 111.2 (C7H), 87.74 (C3Br), 11.9 (C8H3).
2: $^{13}\text{C}$ NMR

2: $^1\text{H}/^1\text{H}$-COSY
2 Catalyzed by SpH2

Yield: 24.8%, 0.186 mmol, 39.1 mg

$^1$H NMR (500 MHz, DMSO-$d_6$) $\delta$ 11.40 (s, 1H, N1H), 7.32 (dd, $^3$J= 8.0 Hz, $^5$J= 0.9 Hz, 1H, C4H), 7.29 (dd, $^3$J= 7.8 Hz, $^5$J= 0.8 Hz, 1H, C7H), 7.10 (ddd, $^3$J= 8.1 Hz, $^7$J= 7.1 Hz, $^5$J= 1.4 Hz, 1H, C5H), 7.05 (ddd, $^3$J= 8.0 Hz, $^7$J= 7.1 Hz, $^5$J= 1.2 Hz, 1H, C6H), 2.37 (s, 3H, C8H3).

$^{13}$C NMR (126 MHz, DMSO-$d_6$) $\delta$ 134.8 (C7a), 133.2 (C2H), 126.8 (C3aH), 121.4 (C6H), 119.6 (C5H), 117.1 (C4H), 111.2 (C7H), 87.7 (C3Br), 11.9 (C2CH3).
2: $^{13}$C NMR

2: $^1$H,$^1$H-COSY
**2-Bromo-3-methylindole 3**

3 Catalyzed by SpH1

Yield: 1.5%, 0.0115 mmol, 2.4 mg

LC-ESI-MS: [C₆H⁻BrN]⁻/m/z: calculated: 207.98 (²⁹Br), 209.98 (³¹Br); observed: LC-ESI-MS [M-H]: 207.98, 209.88

¹H NMR (500 MHz, DMSO-d₆) δ 11.53 (s, NH), 7.45 (d br., J = 7.9 Hz, C⁴H), 7.26 (d br., J = 7.1 Hz, C⁵H), 7.08 (ddd, J = 8.2, 7.0, 1.2 Hz, C⁶H), 7.00 (ddd, J = 7.5, 7.1, 1.2 Hz, C⁷H), 2.18 (s, CH₃).

¹³C NMR (126 MHz, DMSO-d₆) δ 136.0 (C₇a), 127.7 (C⁷a), 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-d6) δ 11.53 (s, 1H, N1H), 7.45 (dd, J= 7.9 Hz, J= 0.9 Hz, 1H, C4H), 7.26 (dd, J= 8.1 Hz, J= 0.9 Hz, 1H, C7H), 7.08 (ddd, J= 8.1 Hz, J= 7.1 Hz, J= 1.2 Hz, 1H, C6H), 7.00 (ddd, J= 8.0 Hz, J= 7.1 Hz, J= 1.1 Hz, 1H, C5H ), 2.18 (s, 3H, C8H3).

13C NMR (126 MHz, DMSO-d6) δ 136.0 (C7a), 127.7 (C3a), 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: $^1$H, $^1$H-COSY
3-Bromo-5-methylindole 4

4 Catalyzed by SpH1

Yield: 26.2%, 0.196 mmol, 41 mg

\([\text{C}_9\text{H}_7\text{BrN}]\) \(m/z\): calculated: 207.98 (\(^{79}\text{Br}\)), 209.98 (\(^{81}\text{Br}\)); observed: LC-ESI-MS \([\text{M-H}]\): 207.98, 209.98

\(^1\text{H}\) NMR (500 MHz, DMSO-\(d_6\)) \(\delta\) 11.33 (s, NH), 7.47 (d, \(J = 2.6\) Hz, C\(^2\)H), 7.31 (d, \(J = 8.2\) Hz, C\(^3\)H), 7.18 (d, \(J = 0.9\) Hz, C\(^4\)H), 6.99 (dd, \(J = 8.3, 1.6\) Hz, C\(^5\)H), 2.39 (s, CH\(_3\)).

\(^{13}\text{C}\) NMR (126 MHz, DMSO-\(d_6\)) \(\delta\) 133.7 (C\(^7a\)), 128.6 (C\(^5a\)), 126.3 (C\(^3a\)), 124.7 (C\(^2\)H), 123.9 (C\(^6\)H), 117.4 (C\(^4\)H), 111.8 (C\(^7\)H), 88.1 (C\(^3\)Br), 21.1 (C\(^8\)H\(_3\)).
4: $^{13}$C NMR

4: $^1$H, $^1$H-COSY
4 Catalyzed by SpH2:

Yield: 4.8%, 0.036 mmol, 7.5 mg

\[ \text{[C}_9\text{H}_7\text{BrN]} \text{ } m/z: \text{ calculated: } 207.98 \text{ (}^{79}\text{Br}), 209.98 \text{ (}^{81}\text{Br)}; \text{ observed: LC-ESI-MS [M-H]}: 208.0, 210.0 \]

\(^1\text{H} \text{ NMR (500 MHz, DMSO-d}_6\): } \delta \text{ [ppm]}= 11.32 \text{ (s, } 1\text{H, N}_1\text{H}), 7.47 \text{ (d, } ^3\text{J}= 2.6 \text{ Hz, } 1\text{H, } C_2\text{H}), 7.31 \text{ (d, } ^3\text{J}= 8.3 \text{ Hz, } 1\text{H, } C_7\text{H}), 7.18 \text{ (s, } 1\text{H, } C^4\text{H}), 7.00 \text{ (dd, } ^3\text{J}= 8.3 \text{ Hz, } ^4\text{J}=1.7 \text{ Hz, } 1\text{H, } C^6\text{H}), 2.40 \text{ (s, } 3\text{H, } C^8\text{H}).

\(^{13}\text{C} \text{ NMR (126 MHz, DMSO-d}_6\): } \delta \text{ 133.7 (C}_7\text{a)}, 128.6 (C^5\text{H}), 126.3 (C^3\text{a}), 124.7 (C^2\text{H}), 123.9 (C^6\text{H}), 117.4 (C^4\text{H}), 111.8 (C^7\text{H}), 88.1 (C^3\text{Br}), 21.1 (C^8\text{H}).
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-d₆) δ 12.21 (s, NH), 8.31 (d, J = 2.2 Hz, C₄H), 8.07 (dd, J = 9.0, 2.3 Hz, C₅H), 7.87 (s, C₂H), 7.63 (d, J = 9.0 Hz, C₅H).

¹³C NMR (126 MHz, DMSO-d₆) δ 141.4 (C₇a), 138.6 (C₅NO₂), 129.1 (C₃a), 125.6 (C₄H), 117.5 (C₂H), 114.9 (C₆H), 113.0 (C₇H), 91.0 (C₃Br).
5: $^{13}$C NMR

5: $^1$H,$^1$H-COSY
Catalyzed by SpH2

Yield: 11.2%, 0.084 mmol, 20.3 mg

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

$^{13}$C NMR (126 MHz, DMSO-$d_6$) δ 141.8 (C7a), 139.0 (C5NO2), 129.5 (C3a), 126.1 (C4H), 118.0 (C2H), 115.4 (C6H), 113.5 (C7H), 91.5 (C3Br).
5: $^{13}$C NMR

5: $^1$H-$^1$H-COSY
3, 5-Dibromoindole 7

HRMS: C₈H₄Br₂N⁺ 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-d₆) 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).

¹³C NMR (126 MHz, DMSO-d₆) δ 134.2 (C₇a), 127.9 (C₃a), 126.6 (C₆H), 124.9 (C₂H), 120.1 (C₄H), 114.3 (C₇H), 112.5 (C₅Br), 87.9 (C₃Br).
7: $^{13}$C NMR

7: HMQC
7 Catalyzed by SpH2

Yield: 0.6%, 0.004 mmol, 1.2 mg

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

![NMR Spectra Image]
3-Bromo-5-fluoroinole 8

8 Catalyzed by SpH2

Yield: 0.7%, 0.005 mmol, 1.1 mg

[C₈H₄BrFN] m/z: calculated: 211.96 (⁸⁷Br), 213.96 (⁸⁵Br), 276.87; observed: LC-ESI-MS [M-H]: 212.0, 214.0

¹H NMR (500 MHz, DMSO-d₆) δ [ppm]: 11.58 (s, 1H, N₁H), 7.63 (d, J = 2.7 Hz, 1H, C₂H), 7.44 (dd, J = 8.9, J_HF = 4.4 Hz, 1H, C⁷H), 7.12 (dd, J_HH = 9.4, J = 2.6 Hz, 1H, C⁵H), 7.03 (ddd, J_HH = 9.2, J = 8.9, J = 2.6 Hz, 1H, C⁶H).
3-Bromo-5-cyanoindole 9

HRMS: C₉H₄BrN₂ 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-d₆) δ 12.04 (s, NH), 7.91 (d, J = 1.6 Hz, C⁴H), 7.79 (s, C³H), 7.61 (d, J = 8.4 Hz, C⁷H), 7.53 (dd, J = 8.5, 1.6 Hz, C⁶H).

¹³C NMR (126 MHz, DMSO-d₆) δ 137.2 (C⁷a), 127.8 (C³a), 126.1 (C⁴H), 124.9 (C⁶H), 123.7 (C²H), 120.1 (C⁸N), 113.6 (C⁷H), 102.2 (C⁵CN), 89.5 (C³Br).
9: $^{13}$C NMR

9: HMQC
Catalyzed by SpH2

Yield: 4.0%, 0.030 mmol, 6.6 mg

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

$^{13}$C NMR (126 MHz, DMSO-$d_6$) δ 137.19 (C$^7a$), 127.8 (C$^3a$), 126.1 (C$^4$H), 124.9 (C$^6$H), 123.7 (C$^2$H), 120.1 (C$^8$N), 113.6 (C$^7$H), 102.2 (C$^5$CN), 89.5 (C$^3$Br).
9: $^{13}$C NMR

9: HMQC
3-Bromo-7-azaindole 15

15 Catalyzed by SpH1

Yield: 7.4%, 0.056 mmol, 11 mg

$[\text{C}_7\text{H}_4\text{BrN}_2]^+$ $m/z$: calculated: 196.96 ($^{79}\text{Br}$), 198.96 ($^{81}\text{Br}$); observed: LC-ESI-MS $[\text{M+H}]^+$: 196.97, 198.97

$^1\text{H}$ NMR (500 MHz, DMSO-$d_6$) $\delta$: 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).

$^{13}\text{C}$ NMR (126 MHz, DMSO-$d_6$) $\delta$: 147.0 (C$^7$a), 143.7 (C$^6$H), 126.6 (C$^4$H), 125.7 (C$^2$H), 118.8 (C$^3$a), 116.4 (C$^5$H), 87.2 (C$^3$Br).
15: $^{13}$C NMR

15: $^1$H,$^1$H-COSY
6-Bromo-2,3-dimethylindole 23

Catalyzed by SpH2

Yield: 0.8%, 0.006 mmol, 1.4 mg

[C_{10}H_{9}BrN] m/z: calculated: 222.0 (^{79}\text{Br}), 224.0 (^{81}\text{Br}); observed: LC-ESI-MS [M-H]: 222.0, 224.0

$^1$H NMR (500 MHz, DMSO-$d_6$) δ [ppm]: 11.69 (s, 1H, N\text{H}), 7.37 (d, $^1J$= 1.8 Hz, 1H, C\text{H}), 7.30 (d, $^3J$= 8.3 Hz, 1H, C\text{H}), 7.03 (dd, $^3J$= 8.4 Hz, $^4J$= 1.8 Hz, 1H, C\text{H}), 2.28 (s br, 3H, CH$_3$); 2.13 (s br, 3H, CH$_3$)

$^{13}$C NMR (126 MHz, DMSO-$d_6$) δ: 136.0 (C\text{7a}), 132.6 (C\text{2CH$_3$}), 127.9 (C\text{3a}), 120.7 (C\text{5H}), 119.0 (C\text{4H}), 112.7 (C\text{7H}), 112.5 (C\text{6Br}), 105.4 (C\text{3CH$_3$}), 11.2 (C\text{8H$_3$}), 8.2 (C\text{9H$_3$}).
23: $^{13}$C 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. The analysis was accomplished using the antibiotics & Secondary Metabolite Analysis Shell (antiSMASH, Version 4.1.0).

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Input</th>
<th>Type</th>
<th>From</th>
<th>To</th>
<th>Most similar known cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>contig-131510000046</td>
<td>Cf_putative</td>
<td>3</td>
<td>7095</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>contig-46000040</td>
<td>Nrps</td>
<td>1</td>
<td>3078</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>contig-12860000038</td>
<td>Hserlactone</td>
<td>1</td>
<td>3272</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>contig-150000000</td>
<td>Cf_putative</td>
<td>3</td>
<td>6624</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>contig-54000025</td>
<td>Cf_putative</td>
<td>1513</td>
<td>8377</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 6</td>
<td>contig-45000027</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>4853</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 7</td>
<td>contig-529000008</td>
<td>Cf_saccharide</td>
<td>1</td>
<td>4279</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 8</td>
<td>contig-446000012</td>
<td>Cf_putative</td>
<td>3</td>
<td>4189</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 9</td>
<td>contig-354000029</td>
<td>Cf_putative</td>
<td>3</td>
<td>8222</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 10</td>
<td>contig-29000042</td>
<td>Terpene</td>
<td>1</td>
<td>11734</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 11</td>
<td>contig-12868000026</td>
<td>Cf_putative</td>
<td>11232</td>
<td>15395</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 12</td>
<td>contig-13030000033</td>
<td>Cf_putative</td>
<td>31</td>
<td>13103</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 13</td>
<td>contig-369000000</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>5305</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 14</td>
<td>contig-290000047</td>
<td>Terpene</td>
<td>1</td>
<td>6928</td>
<td>Astaxanthin_dideoxyglycoside_biosynthetic_gene_cluster (75% of genes show similarity; BGC0001086_c1)</td>
</tr>
<tr>
<td>Cluster 15</td>
<td>contig-12916000015</td>
<td>Hserlactone</td>
<td>1</td>
<td>7196</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 16</td>
<td>contig-760000017</td>
<td>Cf_putative</td>
<td>2</td>
<td>2990</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 17</td>
<td>contig-2211000019</td>
<td>Cf_putative</td>
<td>3</td>
<td>4297</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 18</td>
<td>contig-15000041</td>
<td>Siderophore</td>
<td>1</td>
<td>6752</td>
<td>-</td>
</tr>
<tr>
<td>Cluster</td>
<td>Input</td>
<td>Type</td>
<td>From</td>
<td>To</td>
<td>Most similar known cluster</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cluster 19</td>
<td>contig-14000013</td>
<td>T1pks</td>
<td>1</td>
<td>10855</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 20</td>
<td>contig-620000022</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>5844</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 21</td>
<td>contig-15800018</td>
<td>Cf_putative</td>
<td>3</td>
<td>6408</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 22</td>
<td>contig-12842000026</td>
<td>Cf_putative</td>
<td>196</td>
<td>7594</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 23</td>
<td>contig-27000028</td>
<td>Cf_putative</td>
<td>106</td>
<td>7141</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 24</td>
<td>contig-268000031</td>
<td>Cf_putative</td>
<td>2</td>
<td>3593</td>
<td>Galactogluca_n_biosynthetic_gene_cluster (12% of genes show similarity; BGC0000801_c1)</td>
</tr>
<tr>
<td>Cluster 25</td>
<td>contig-1091000038</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>5482</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 26</td>
<td>contig-38000002</td>
<td>Cf_putative</td>
<td>2</td>
<td>12531</td>
<td>Sphingan_polysaccharide_biosynthetic_gene_cluster (17% of genes show similarity; BGC0000797_c1)</td>
</tr>
<tr>
<td>Cluster 27</td>
<td>contig-12844000026</td>
<td>Cf_saccharide</td>
<td>1</td>
<td>16114</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 28</td>
<td>contig-2000044</td>
<td>Cf_putative</td>
<td>6400</td>
<td>18603</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 29</td>
<td>contig-1120000000</td>
<td>Cf_putative</td>
<td>285</td>
<td>6877</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 30</td>
<td>contig-370000027</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>8972</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 31</td>
<td>contig-75000041</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>12158</td>
<td>Asukamycin_biosynthetic_gene_cluster (4% of genes show similarity; BGC0000187_c1)</td>
</tr>
<tr>
<td>Cluster 32</td>
<td>contig-979000022</td>
<td>Cf_fatty_acid</td>
<td>4189</td>
<td>17707</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 33</td>
<td>contig-213000021</td>
<td>Bacteriocin</td>
<td>1</td>
<td>5155</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 34</td>
<td>contig-28000007</td>
<td>Cf_putative</td>
<td>1733</td>
<td>7143</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 35</td>
<td>contig-1266000040</td>
<td>Cf_putative</td>
<td>6339</td>
<td>13208</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 36</td>
<td>contig-12649000039</td>
<td>Cf_putative</td>
<td>4224</td>
<td>19885</td>
<td>Emulsan_biosynthetic_gene_cluster (9% of genes show similarity; BGC0000760_c1)</td>
</tr>
<tr>
<td>Cluster 37</td>
<td>contig-13059000033</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>14860</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 38</td>
<td>contig-13194000046</td>
<td>Cf_putative</td>
<td>2</td>
<td>9476</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 39</td>
<td>contig-25000042</td>
<td>Cf_putative</td>
<td>13360</td>
<td>25949</td>
<td>Azinomycin_B_biosynthetic_gene_cluster (4% of genes show similarity; BGC0000960_c1)</td>
</tr>
<tr>
<td>Cluster 40</td>
<td>contig-624000005</td>
<td>Cf_putative</td>
<td>328</td>
<td>4559</td>
<td>-</td>
</tr>
<tr>
<td>Cluster</td>
<td>Input</td>
<td>Type</td>
<td>From</td>
<td>To</td>
<td>Most similar known cluster</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>------</td>
<td>-----</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cluster 41</td>
<td>contig-90000034</td>
<td>Cf_saccharide</td>
<td>1</td>
<td>3955</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 42</td>
<td>contig-149100034</td>
<td>Cf_putative</td>
<td>3</td>
<td>8288</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 43</td>
<td>contig-1289200015</td>
<td>Cf_putative</td>
<td>14216</td>
<td>24750</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 44</td>
<td>contig-674000020</td>
<td>Cf_putative</td>
<td>3</td>
<td>4752</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 45</td>
<td>contig-31100004</td>
<td>Other</td>
<td>1</td>
<td>4111</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 46</td>
<td>contig-17000011</td>
<td>Cf_putative</td>
<td>335</td>
<td>8704</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 47</td>
<td>contig-121000046</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>9103</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 48</td>
<td>contig-8000031</td>
<td>Lassopeptide</td>
<td>1</td>
<td>8016</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 49</td>
<td>contig-25400002</td>
<td>Bacteriocin</td>
<td>1</td>
<td>4268</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 50</td>
<td>contig-1316800046</td>
<td>Bacteriocin</td>
<td>5160</td>
<td>13484</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 51</td>
<td>contig-47200023</td>
<td>Cf_putative</td>
<td>2</td>
<td>6320</td>
<td>Lipopolysaccharide_biosynthetic_gen_e_cluster (8% of genes show similarity; BGC0000774_c1)</td>
</tr>
<tr>
<td>Cluster 52</td>
<td>contig-15600028</td>
<td>Cf_putative</td>
<td>2</td>
<td>3814</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 53</td>
<td>contig-29000000</td>
<td>Cf_putative</td>
<td>1</td>
<td>6658</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 54</td>
<td>contig-71000021</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>11491</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 55</td>
<td>contig-81000021</td>
<td>Cf_putative</td>
<td>609</td>
<td>13474</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 56</td>
<td>contig-137000046</td>
<td>Cf_putative</td>
<td>3</td>
<td>7962</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 57</td>
<td>contig-134400023</td>
<td>Cf_putative</td>
<td>437</td>
<td>7335</td>
<td>Lipopolysaccharide_biosynthetic_gen_e_cluster (5% of genes show similarity; BGC0000774_c1)</td>
</tr>
<tr>
<td>Cluster 58</td>
<td>contig-37000040</td>
<td>Cf_putative</td>
<td>2</td>
<td>4429</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 59</td>
<td>contig-1262000037</td>
<td>Cf_putative</td>
<td>1</td>
<td>4232</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 60</td>
<td>contig-1255900022</td>
<td>Cf_fatty_acid</td>
<td>1</td>
<td>8107</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 61</td>
<td>contig-45300026</td>
<td>Cf_putative</td>
<td>1</td>
<td>4304</td>
<td>O&amp;K-antigen_biosynthetic_gene_cluster (3% of genes show similarity; BGC0000780_c1)</td>
</tr>
<tr>
<td>Cluster 62</td>
<td>contig-1325200030</td>
<td>Ectoine</td>
<td>15769</td>
<td>26170</td>
<td>Ectoine_biosynthetic_gene_cluster (80% of genes show similarity; BGC0000857_c1)</td>
</tr>
<tr>
<td>Cluster</td>
<td>Input</td>
<td>Type</td>
<td>From</td>
<td>To</td>
<td>Most similar known cluster</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-----------</td>
<td>------</td>
<td>------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cluster 63</td>
<td>contig-18000018</td>
<td>Cf_putative</td>
<td>3188</td>
<td>6860</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 64</td>
<td>contig-13176000000</td>
<td>Cf_putative</td>
<td>3</td>
<td>7349</td>
<td>Xenocyloins_biosynthetic_gene_cluster (25% of genes show similarity; BGC0000189_c1)</td>
</tr>
<tr>
<td>Cluster 65</td>
<td>contig-688000009</td>
<td>Cf_putative</td>
<td>5027</td>
<td>9594</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 66</td>
<td>contig-20000007</td>
<td>Cf_putative</td>
<td>423</td>
<td>15643</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 67</td>
<td>contig-261000005</td>
<td>Cf_putative</td>
<td>9556</td>
<td>20797</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 68</td>
<td>contig-5000033</td>
<td>Cf_putative</td>
<td>2</td>
<td>13098</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 69</td>
<td>contig-10000032</td>
<td>Bacteriocin</td>
<td>1218</td>
<td>6759</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 70</td>
<td>contig-1272500036</td>
<td>Cf_putative</td>
<td>933</td>
<td>15577</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 71</td>
<td>contig-12782000009</td>
<td>Other</td>
<td>1</td>
<td>5882</td>
<td>-</td>
</tr>
</tbody>
</table>

**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**

<table>
<thead>
<tr>
<th>Apoenzyme</th>
<th>Cavity/Pocket Area (Å²)</th>
<th>Enzyme / docked indole</th>
<th>Enzyme / docked indole / 1ns MD simulation</th>
<th>Enzyme/ docked 3-Br-indole / 1ns MD simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Area (Å²)</td>
<td>Volume (Å³)</td>
<td>Area (Å²)</td>
</tr>
<tr>
<td>SpH2</td>
<td>136.80</td>
<td>45.45</td>
<td>23.68</td>
<td>197.12</td>
</tr>
<tr>
<td>SpH1</td>
<td>77.39</td>
<td>77.39</td>
<td>15.60</td>
<td>115.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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
References

