Combining OSMAC, metabolomic and genomic methods on termite symbiotic fungi for the production and annotation of halogenated azaphilones and ilicicolins

Téo Hebra[1], Nicolas Pollet[2], David Touboul[1,*], Véronique Eparvier[1,*]

[1] Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, 91198, Gif-sur-Yvette, France.
[2] Université Paris-Saclay, CNRS, IRD, Évolution, Génomes, Comportement & Écologie, 91198 Gif-sur-Yvette, France.
[*]email: veronique.eparvier@cnrs.fr; david.touboul@cnrs.fr

Suppdata

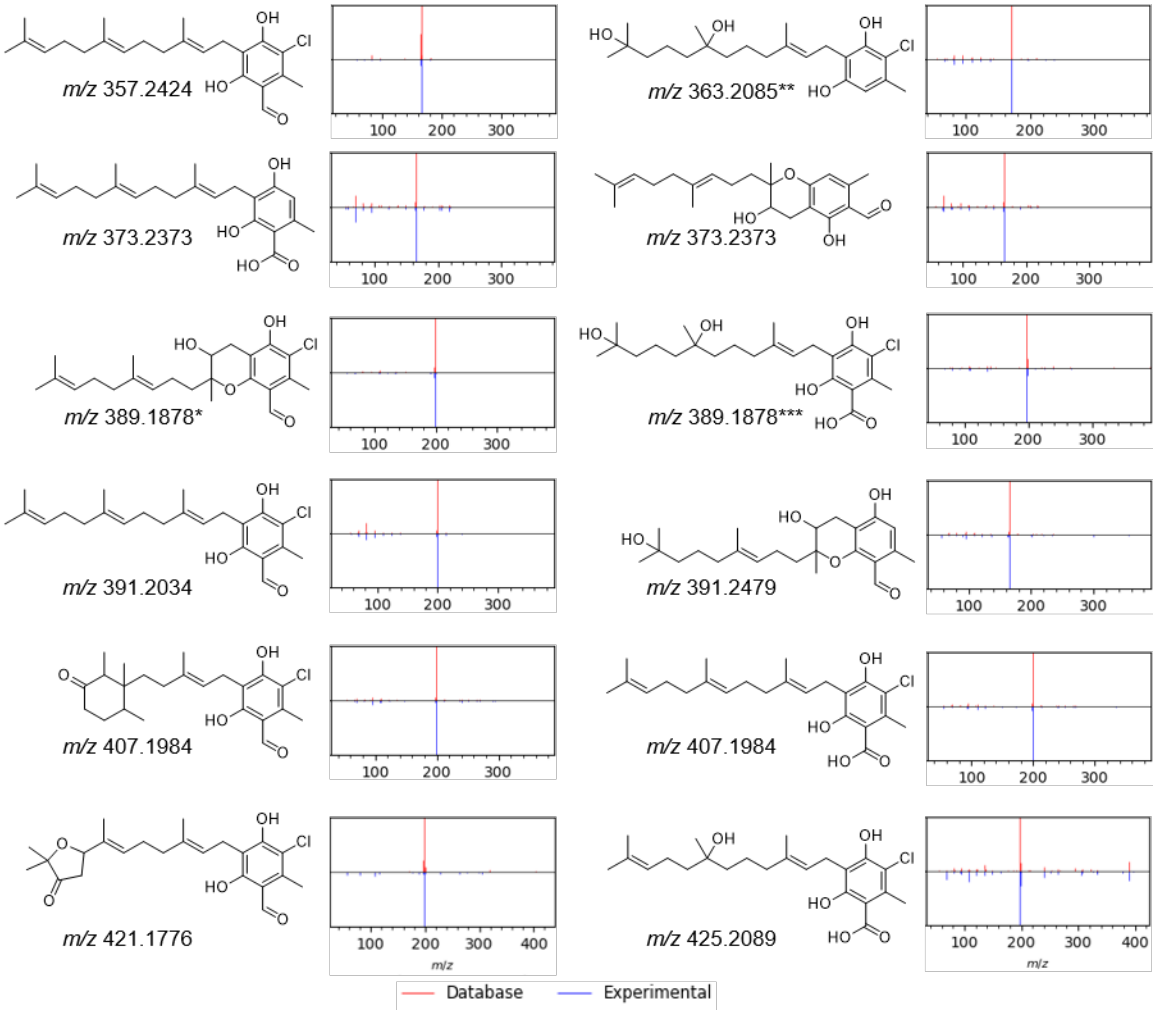
Table S1: Dereplicated ilicicolins from *N. discophora* crude extract.

CN63								
$[M+H]^+$	**Error (ppm)**	**RT**	**Formula**	**Name**	**Database**	**Cosine score**	**Annotation***	**Ref****
357.2422	0.6	14.9	$C_{23}H_{32}O_3$	LL-Z 1272β	GNPS	0.91	Level 2	[1]
363.2080	1.5	8.6	$C_{22}H_{35}ClO_4$	Ilicicolinol-$2H_2O$	In-House	0.92	Level 1	[2]
373.2364	2.5	16.5	$C_{23}H_{32}O_4$	Ilicicolinic acid B	In-House	0.93	Level 1	[3]
373.2373	0.0	13.2	$C_{23}H_{32}O_4$	Ilicicolinal G	In-House	0.85	Level 1	[2]
389.1867	2.8	14.7	$C_{23}H_{31}ClO_4$	Ilicicolinal H-H_2O	In-House	0.92	Level 1	[2]
389.1873	1.3	8.5	$C_{23}H_{35}ClO_6$	Ilicicolinic acid E-$3H_2O$	In-House	0.93	Level 1	[2]
391.2028	1.7	15.9	$C_{23}H_{31}ClO_3$	LL-Z1272α	MS-DIAL	0.89	Level 2	[4]
391.2456	5.9	11.7	$C_{23}H_{34}O_5$	Ilicicolinal I	In-House	0.82	Level 1	[2]
407.1951	8.0	17.0	$C_{23}H_{31}ClO_4$	LL-Z 1272δ	MS-DIAL	0.69	Level 2	[5]
407.1985	-0.4	14.2	$C_{23}H_{31}ClO_4$	Ilicicolinic acid A	In-House	0.93	Level 1	[3]
421.1773	0.8	12.2	$C_{23}H_{29}ClO_5$	Ascofuranone	MS-DIAL	0.74	Level 2	[6]
425.2082	1.7	8.5	$C_{23}H_{33}ClO_5$	Ilicicolinic acid C	In-House	0.82	Level 1	[3]

*Level 0 corresponds annotations from isolated pure compounds; level 1 to annotations by comparison with standards; level 2 to putative annotations (e.g., MS/MS library comparison or tentative structure); level 3 to a chemical class assignment. RT = Retention Time

** The references are indicated p23

Figure S1: Dereplicated ilicicolins structures with their theoretical exact masses and and experimental fragmentation spectrum comparison with databases.



* represents the amount of H_2O losses during ionization.

Table S2: Dereplicated azaphilones from *P. sclerotiorum* crude extract.

CN111							
$[M+H]^+$	Error (ppm)	Formula	Name	Database	Cosine score	Annotation*	Ref**
349.1202	-0.3	$C_{19}H_{21}ClO_4$		In-House	0.65	Level 2	[7]
356.1853	1.0	$C_{21}H_{25}NO_4$		In-House	0.86	Level 2	[7]
357.1700	-1.0	$C_{21}H_{24}O_5$		In-House	0.95	Level 2	[7]
359.1857	-1.1	$C_{21}H_{26}O_5$		In-House	0.74	Level 2	[7]
362.1522	-1.3	$C_{20}H_{24}ClNO_3$		In-House	0.88	Level 2	[7]
363.1353	1.3	$C_{20}H_{24}ClO_4$		In-House	0.85	Level 2	[7]
371.1844	2.4	$C_{22}H_{26}O_5$		In-House	0.84	Level 2	[7]
381.1697	-0.1	$C_{23}H_{24}O_5$	Isorotiorin	In-House	0.96	Level 2	[8]
385.2007	0.7	$C_{23}H_{28}O_5$		In-House	0.92	Level 2	[7]
390.1474	-1.9	$C_{21}H_{24}ClNO_4$	Sclerotioramine	MONA	0.90	Level 1	[7,8]
391.1309	-0.6	$C_{21}H_{23}ClO_5$	Sclerotiorine	MS-DIAL	0.82	Level 1	[7,8]
391.2114	0.3	$C_{22}H_{30}O_6$		In-House	0.86	Level 2	[7]
399.2169	-0.8	$C_{24}H_{30}O_5$		In-House	0.92	Level 2	[7]
404.1627	-1.0	$C_{22}H_{26}ClNO_4$		In-House	0.96	Level 2	[7]
405.1468	-1.2	$C_{22}H_{25}ClO_5$	Sclerketide B	In-House	0.95	Level 2	[7,9]
407.1256	0.0	$C_{21}H_{23}ClO_6$		In-House	0.86	Level 2	[7]
415.1299	1.9	$C_{23}H_{23}ClO_5$	5-Chloroisorotiorin	In-House	0.93	Level 1	[7,8]
417.1902	1.4	$C_{23}H_{28}O_7$	Geumsanol B	In-House	0.86	Level 2	[10]
419.1617	0.7	$C_{23}H_{27}ClO_5$		In-House	0.94	Level 2	[7]
433.1773	0.8	$C_{24}H_{29}ClO_5$		In-House	0.69	Level 2	[7]
434.1731	-0.5	$C_{23}H_{28}ClNO_5$	Isochromophilone VI	MSDIAL	0.83	Level 1	[7,11]
451.1526	1.3	$C_{23}H_{27}ClO_7$	Chlorogeumsanol B	In-House	0.84	Level 1	[7]
500.1839	-0.9	$C_{27}H_{30}ClNO_6$		In-House	0.96	Level 2	[7]

*Level 0 corresponds annotations from isolated pure compounds; level 1 to annotations by comparison with standards; level 2 to putative annotations (e.g., MS/MS library comparison or tentative structure); level 3 to a chemical class assignment.

** The references are indicated p23

Figure S2: Dereplicated azaphilone structures with their theoretical exact masses and experimental fragmentation spectrum comparison with databases.

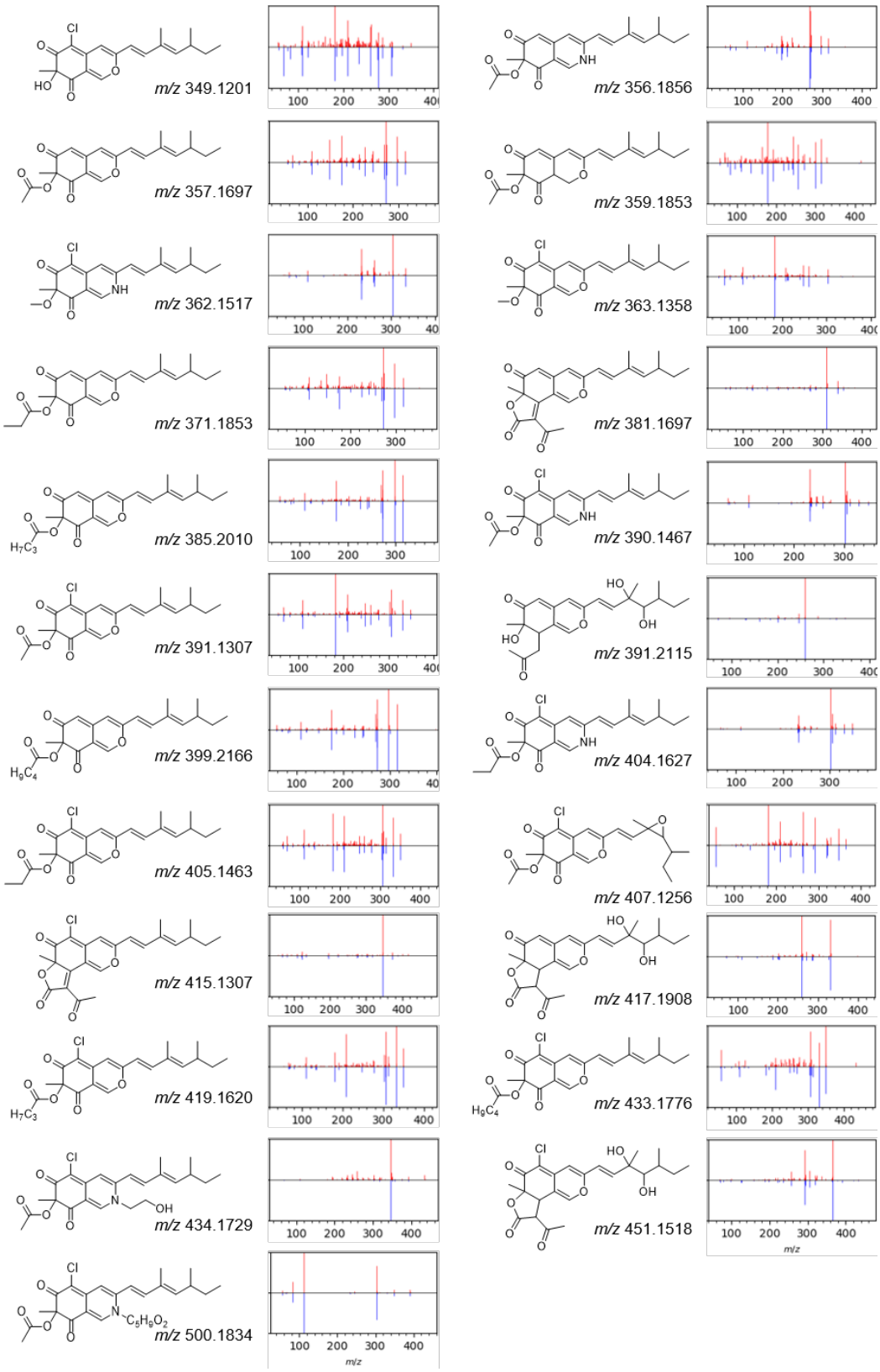


Table S3: Basic genomes statistics.

	Genome size (Mpb)	Contigs	N50 (Mpb)	N75 (Mpb)	L50 (contigs)	L75 (contigs)	%GC
Neonectria discophora	41.6	26	4.03	3.02	3	8	54.2
Penicillium sclerotiorum	34.7	10	4.34	3.95	4	6	48.3

Table S4: BUSCO score.

Gene	Complete		Fragmented (%)	Missing (%)
	Single copy (%)	Duplicated (%)		
Neonectria discophora	99.3	0.0	0.0	0.7
Penicillium sclerotiorum	97.6	0.7	0.0	1.7

Table S5: *Ndi_Ili* biosynthetic gene cluster annotation.

Gene name	Size (AA)	Best match Swissprot	Global alignement (I/S/G) %	Global alignement (I/S/G) AA	Pfam domain	PFAM_ID	Plausible function	Duplicated
Ndi_Ili_A	2014	A0A455R5P9	56/70/8	1201/1506/179	ACP transacylase Beta-ketoacyl synthase Acyl transferase Polyketide synthase dehydratase Phosphopantetheine attachment alpha/beta hydrolase	PF16073.5 PF00109.26 PF00698.21 PF14765.6 PF00550.25 PF07859.13	Non reducing polyketide synthase	N
Ndi_Ili_B	297	A0A455R413	49/59/16	167/203/55	Prenyltransferase	PF01040.18	Prenyltransferase	N
Ndi_Ili_C	1065	A0A455R7E6	60/75/4	662/826/40	AMP-binding, NAD dependent epimerase/dehydratase	PF00501.28 PF01370.21	Non-canonical non-ribosomal peptide synthetase	N
Ndi_Ili_D	553	A0A455R7M0	68/79/5	387/447/31	Tryptophan halogenase	PF04820.14	FAD dependant halogenase	N
Ndi_Ili_E	674	No reviewed protein			Sugar (and other) transporter, Major Facilitator	PF00083.24 PF07690.16	Transporter	N
Ndi_Ili_F	451	No reviewed protein			ATPase family associated with various cellular activities	PF00004.29	ATPase	N
Ndi_Ili_G	532	No reviewed protein						N
Ndi_Ili_H	418	No reviewed protein						N
Ndi_Ili_I	442	No reviewed protein					unknown function	N
Ndi_Ili_J	782	No reviewed protein					unknown function	N

I: identity, S: similarity, G: gap

Table S6: Annotated ilicicolins from *N. discophora* crude extract, synthetized by *Ndi_Ili* biosynthetic pathway.

CN63					
ID	**$[M+H]^+$**	**Error (ppm)**	**Formula**	**Identification**	**Ref***
9	373.2364	2.5	$C_{23}H_{32}O_4$	Dereplication, standard	[3]
2	407.1985	1.1	$C_{23}H_{31}ClO_4$	Dereplication, standard	[3]
8	357.2422	0.6	$C_{23}H_{32}O_3$	Dereplication, *in silico*	[1]
1	391.2028	1.7	$C_{23}H_{31}ClO_3$	Dereplication, *in silico*	[4]

* The references are indicated p23

Table S7: *Psc_Aza* biosynthetic gene cluster annotation.

Gene name	Size (AA)	Best match Swissprot	Global alignement (I/S/G) %	Global alignement (I/S/G) AA	Pfam domain	PFAM_ID	Plausible_function	Duplicated
Psc_Aza_A	2584	Q0CF75	65/77/7	1744/2076/185	ACP transacylase Beta-ketoacyl synthase Acyl transferase Polyketide synthase Dehydratase Methyltransferase Zinc-binding dehydrogenase Ketoreductase	PF16073.5 PF00109.26 PF00698.21 PF14765.6 PF08242.12 PF00107.26 PF08659.10	Highly reducing polyketide synthase	N
Psc_Aza_B	2726	A0A0K0MCJ4	67/78/5	1870/2188/129	ACP transacylase Beta-ketoacyl synthase Acyl transferase Phosphopantetheine attachment site Methyltransferase	PF16073.5 PF00109.26 PF00698.21 PF00550.25 PF13649.6	Non reducing polyketide synthase	N
Psc_Aza_C	352	Q0CF72	44/55/32	208/256/152	FAD-binding domain_3	PF01494.19	Monooxygenase FAD dependant	N
Psc_Aza_D	449	A0A084B9Z3	34/50/8	162/239/40	Transferase	PF02458	Acyltransferase	N
Psc_Aza_E	471	Q0CF74	49/66/4	232/318/18	FAD binding domain	PF01565.23	FAD-linked oxidoreductase	N
Psc_Aza_F	341	No reviewed protein			Zinc-binding dehydrogenase, Alcohol dehydrogenase GroES-like	PF00107.26 PF08240.12	Dehydrogenase	N
Psc_Aza_G	385	No reviewed protein			Alcohol dehydrogenase GroES-like	PF08240.12	Alcohol dehydrogenase	N
Psc_Aza_H	561	Q0CCX4	55/68/8	318/392/43	Tryptophan halogenase	PF04820.14	FAD-dependant halogenase	N
Psc_Aza_I	590	No reviewed protein			No PFAM domain			N
Psc_Aza_J	227	No reviewed protein			Amine oxidoreductase	PF01593.24	Amine oxidase	N
Psc_Aza_K	306	No reviewed protein			G protein-coupled glucose receptor	PF11970.8	G protein-coupled receptor	N
Psc_Aza_L	590	No reviewed protein			Major Facilitator	PF07690.16	Efflux pump	N
Psc_Aza_M	407	No reviewed protein			Unknow function	PF14033.6	Unknow function	N

I: identity, S: similarity, G: gap

Figure S3: MS/MS spectra of [**10**+H]$^+$ and the corresponding fragmentation scheme.

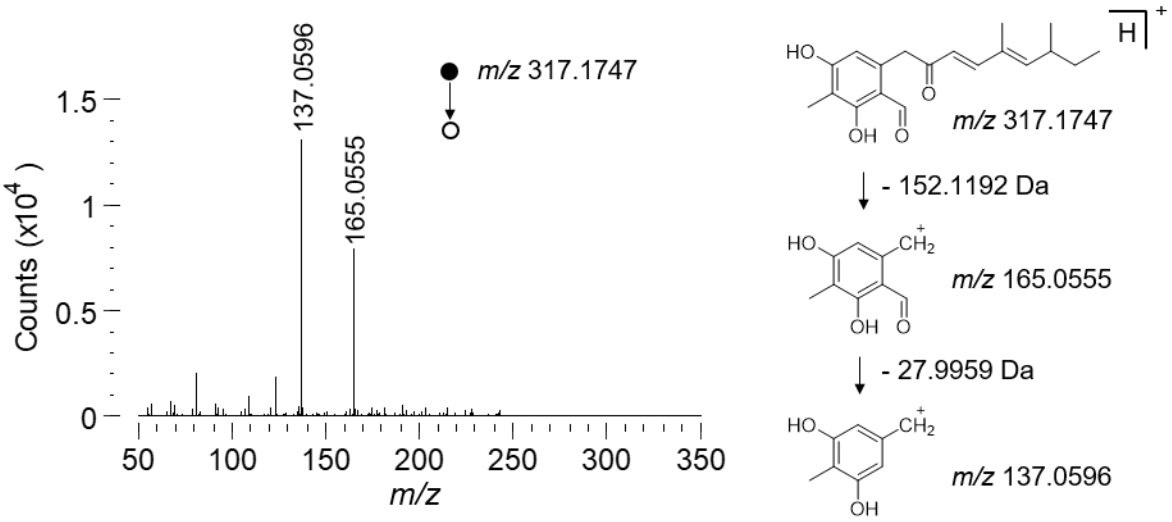


Figure S4: MS/MS spectra of [**11**+H]$^+$ and the corresponding fragmentation scheme.

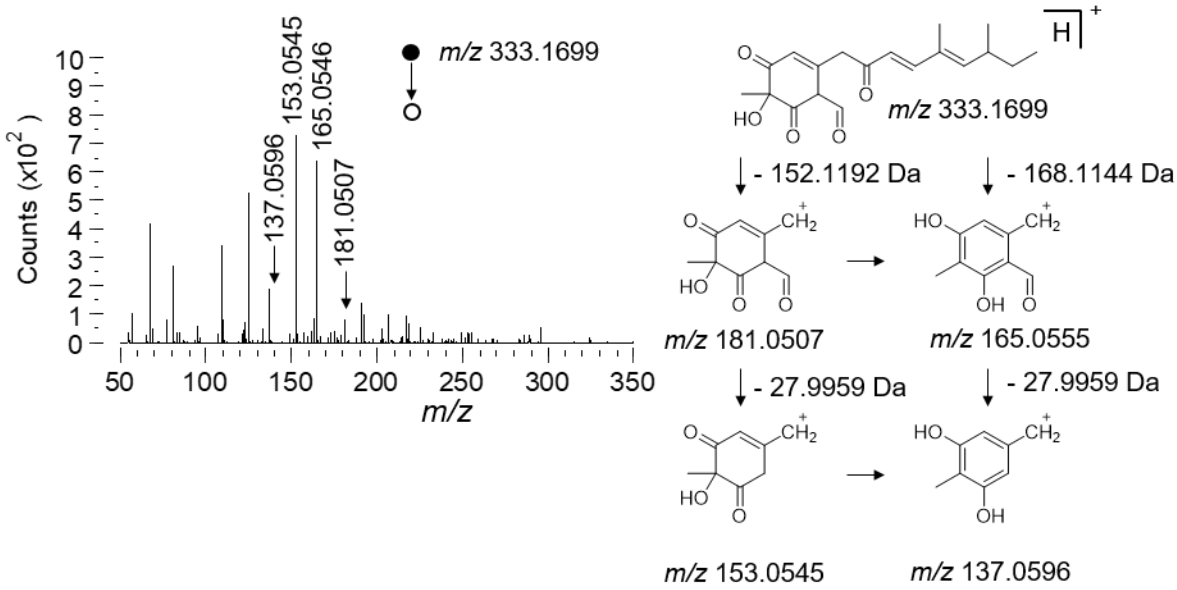


Figure S5: MS/MS spectra of [**12**+H]$^+$ and the corresponding fragmentation scheme.

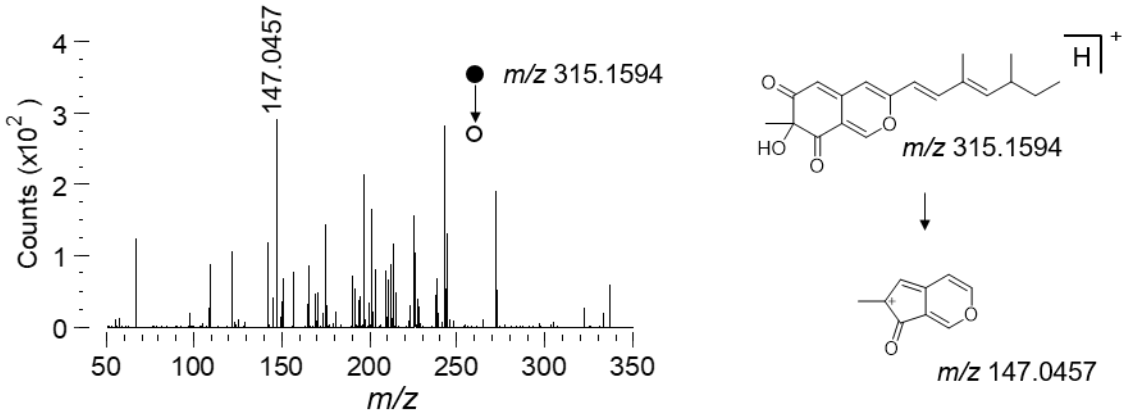


Figure S6: Comparison of intermediates **12** and **21** fragmentation spectra with their common neutral losses corresponding to H/Cl differences.

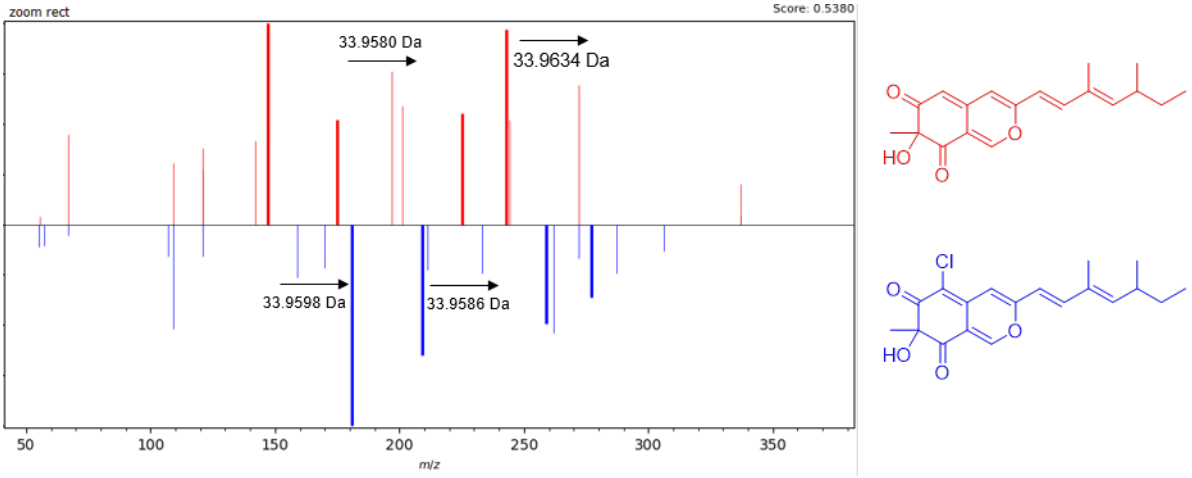


Table S8: Annotated azaphilones from *P. sclerotiorum* crude extract and fractions, synthetized by *Psc_Aza* biosynthetic pathway.

CN111					
ID	**$[M+H]^+$**	**Error (ppm)**	**Formula**	**Identification**	**Ref***
10	317.1747	0.1	$C_{19}H_{24}O_4$	Manual annotation, *in silico*	This study
11	333.1699	-0.8	$C_{19}H_{24}O_5$	Manual annotation, *in silico*	This study
12	315.1594	-1	$C_{19}H_{22}O_4$	Manual annotation, *in silico*	[12]
21	349.1205	-1.1	$C_{19}H_{21}ClO_4$	Dereplication, *in silico*	[7]
13	357.1705	-2.4	$C_{21}H_{24}O_5$	Dereplication, *in silico*	[7]
4	391.1309	-0.6	$C_{21}H_{23}ClO_5$	Dereplication, standard	[7,8]
19	361.2021	-3.2	$C_{21}H_{28}O_5$	Manual annotation, *in silico*	This study
20	395.1626	-1.6	$C_{21}H_{27}Cl_{O5}$	Dereplication, *in silico*	[7,8]
15	381.1691	1.5	$C_{23}H_{24}O_5$	Dereplication, *in silico*	[7,8]
6	415.1312	-1.3	$C_{23}H_{23}ClO_5$	Dereplication, *standard*	[7,8]
16	383.1859	-1.6	$C_{23}H_{26}O_5$	Dereplication, *in silico*	[7,8]
17	417.1467	-0.9	$C_{23}H_{25}ClO_5$	Dereplication, *in silico*	[7,8]

* The references are indicated p24

Figure S7: Extracted ion chromatogram of halogenated ilicicolins from scaffold A (H, Cl and Br) with their isotopic patterns.

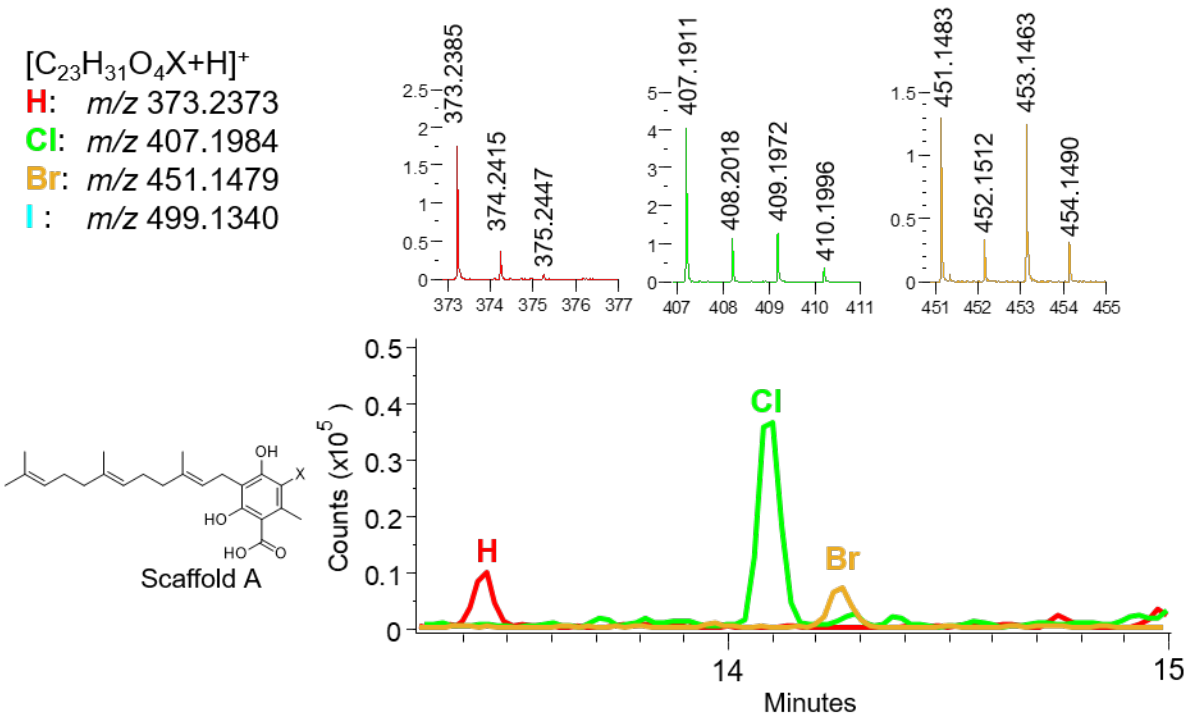


Figure S8: Extracted ion chromatogram of halogenated ilicicolins from scaffold B (H, Cl and Br) with their isotopic patterns.

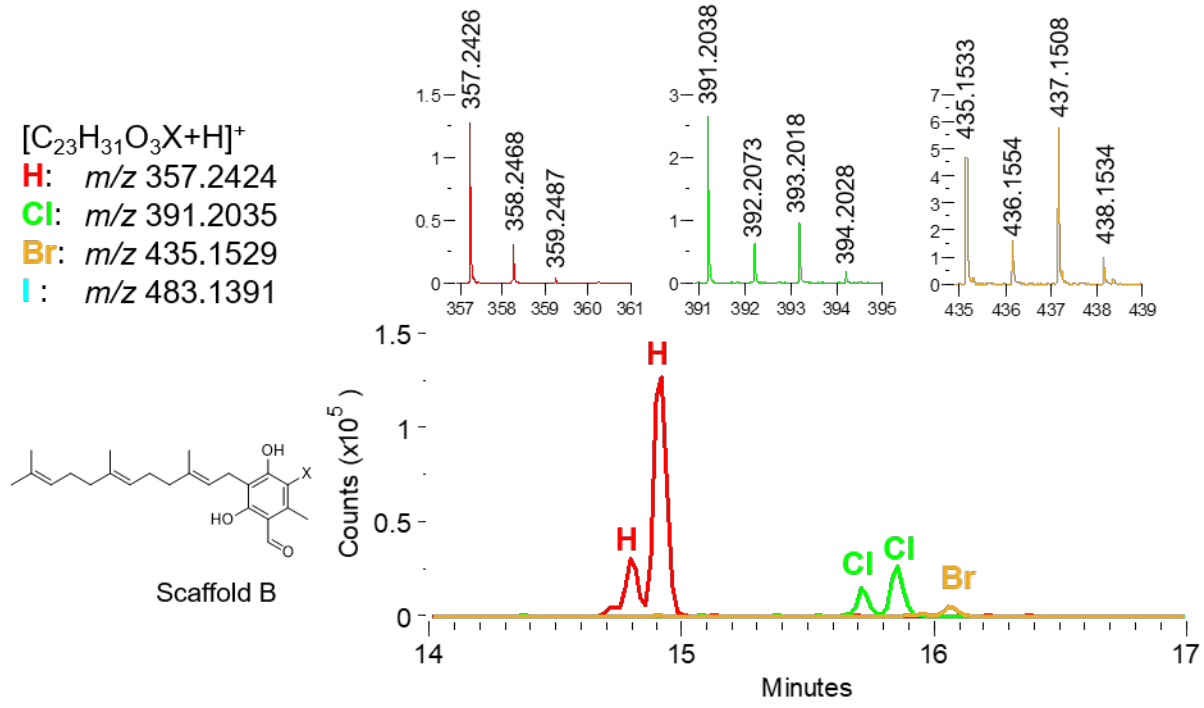


Figure S9: Extracted ion chromatogram of halogenated azaphilones from scaffold B' (H, Cl, Br and I) with their isotopic patterns.

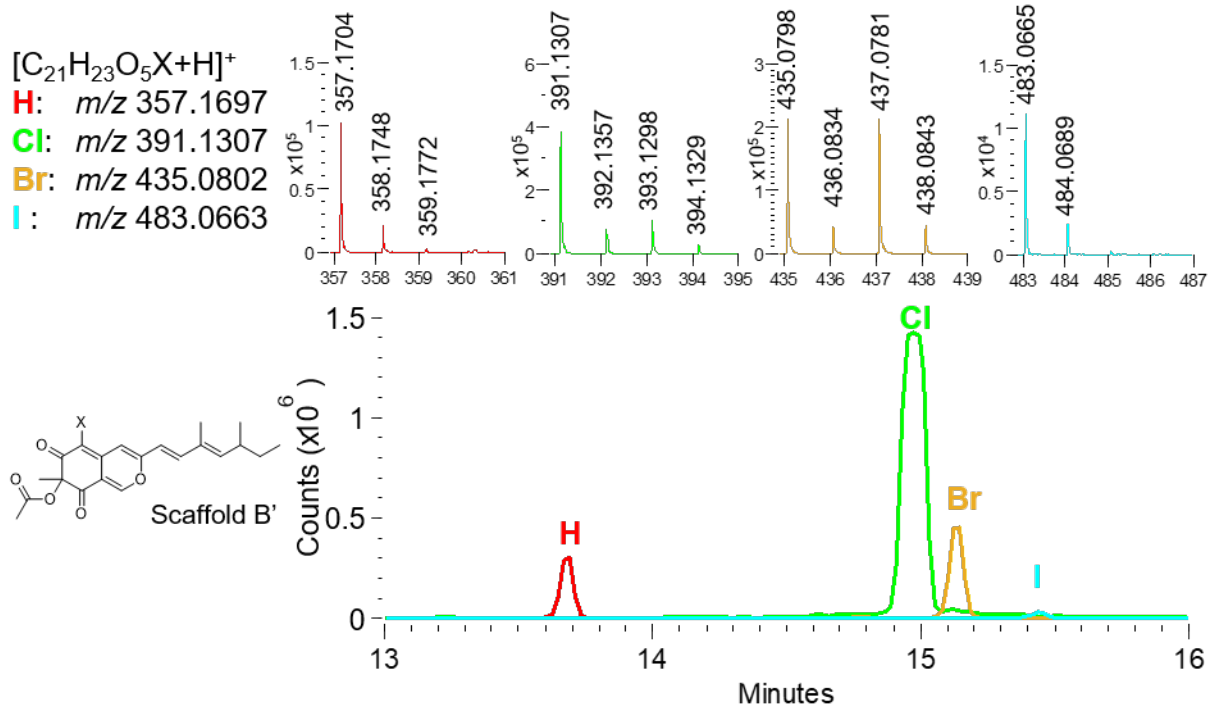


Figure S10: Extracted ion chromatogram of halogenated azaphilones from scaffold C' (H, Cl, Br and I) with their isotopic patterns.

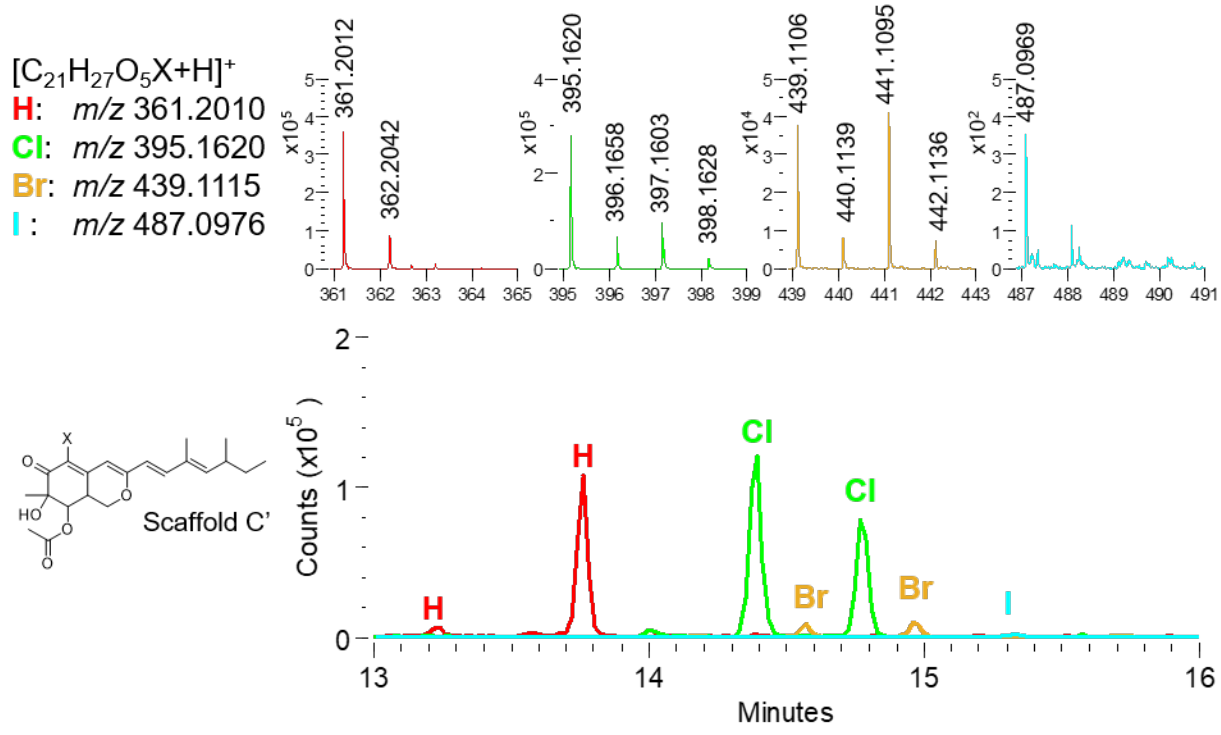


Figure S11: Extracted ion chromatogram of halogenated azaphilones from scaffold D' (H, Cl, Br and I) with their isotopic patterns.

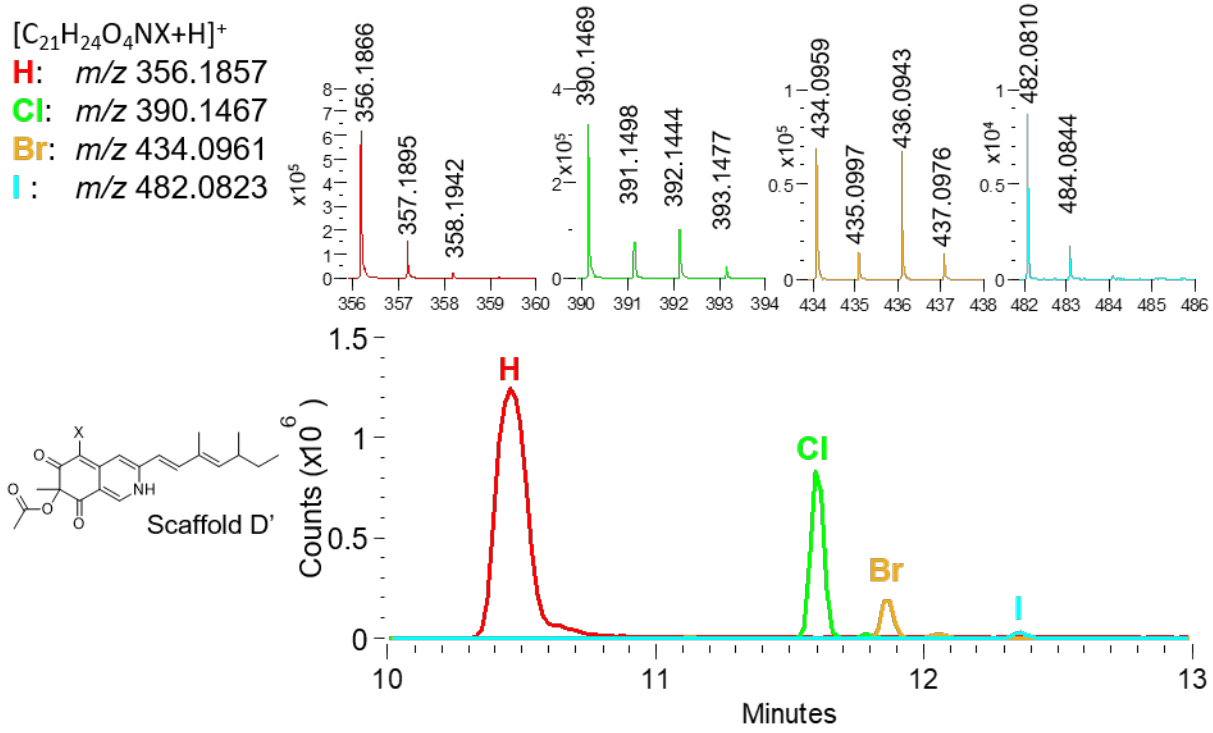


Figure S12: Extracted ion chromatogram of halogenated azaphilones from scaffold E' (H, Cl, Br and I) with their isotopic patterns.

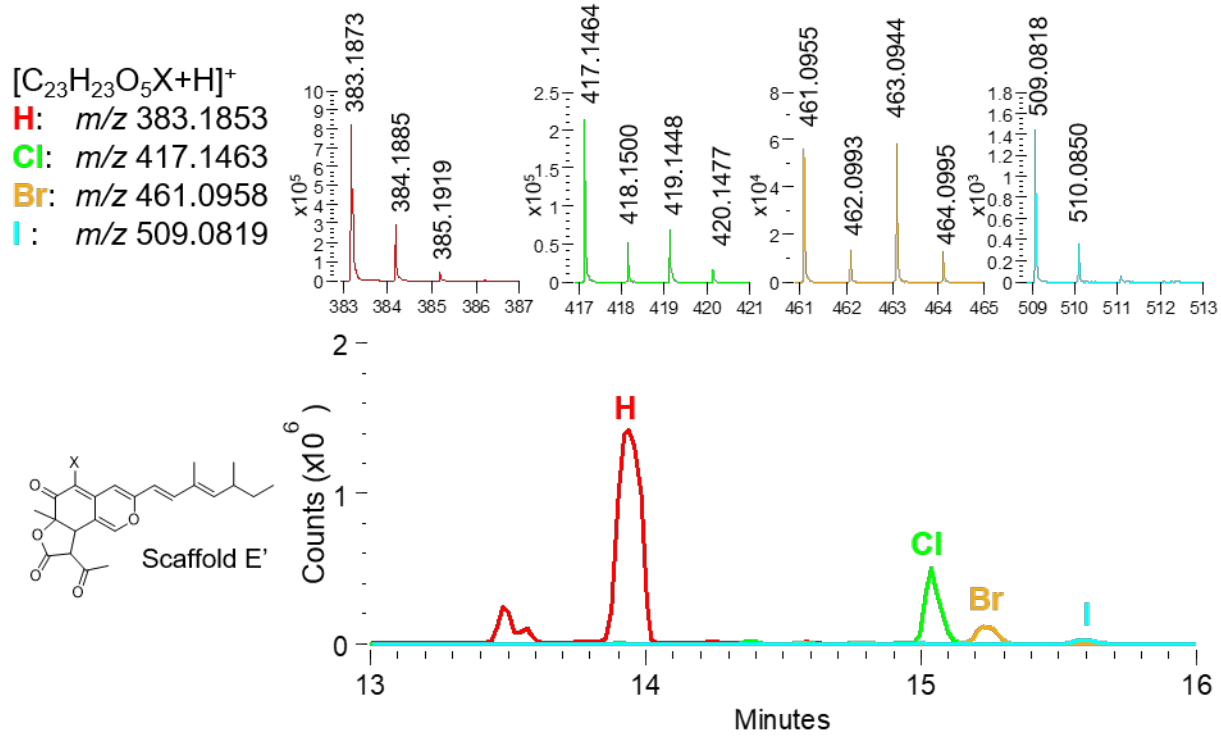


Figure S13: HRMS of compound **22.**

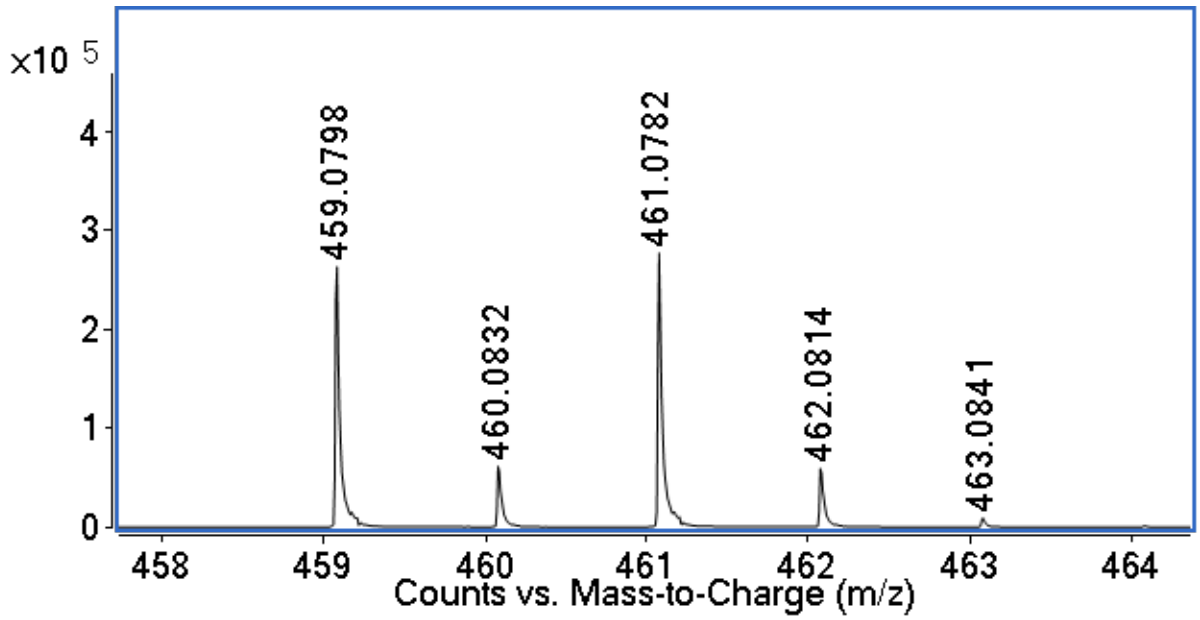


Figure S14: ^{1}H NMR spectrum (DMF-*d7*) of compound **22.**

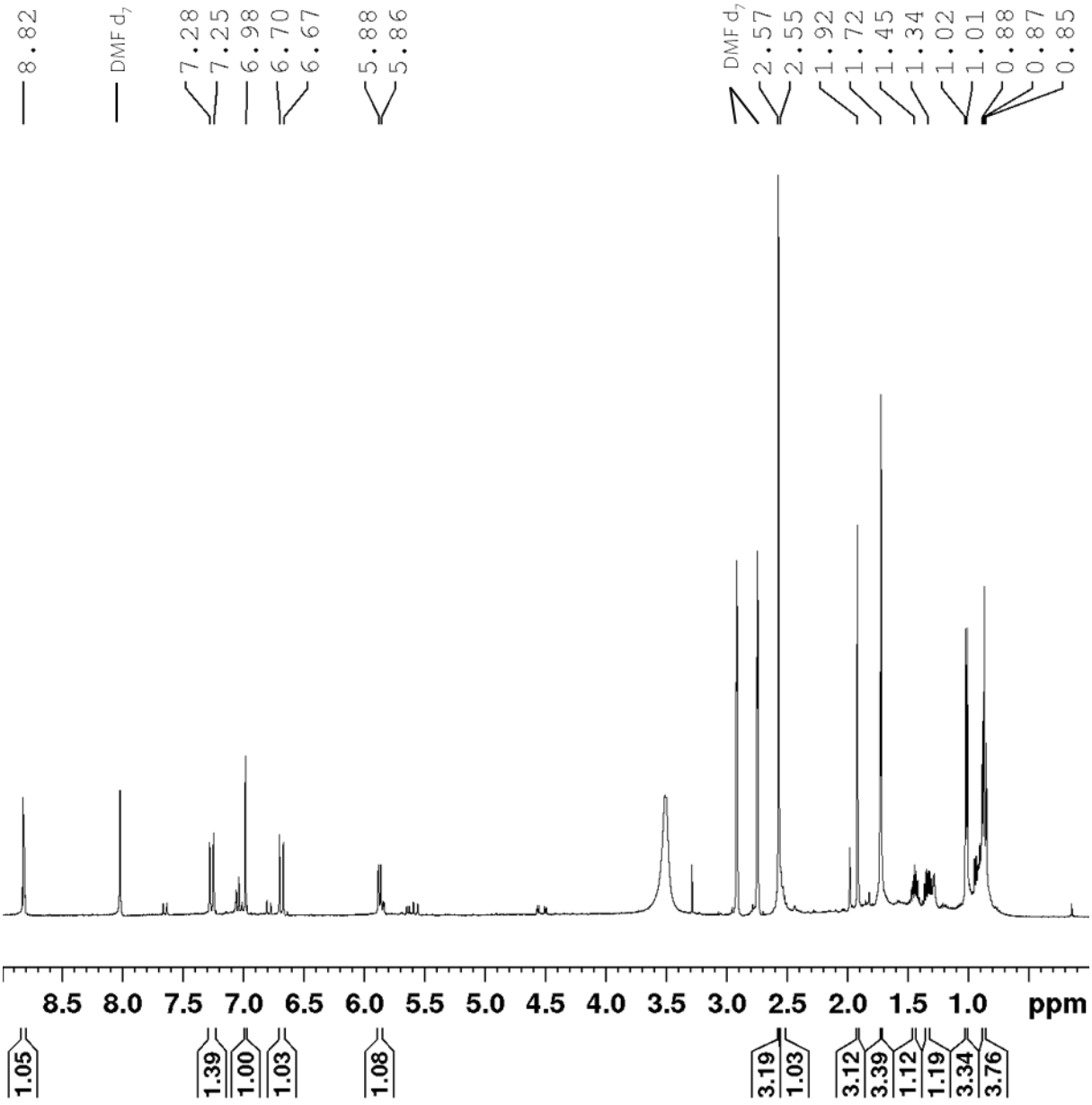


Figure S15: ^{13}C NMR spectrum (DMF-*d7*) of compound **22.**

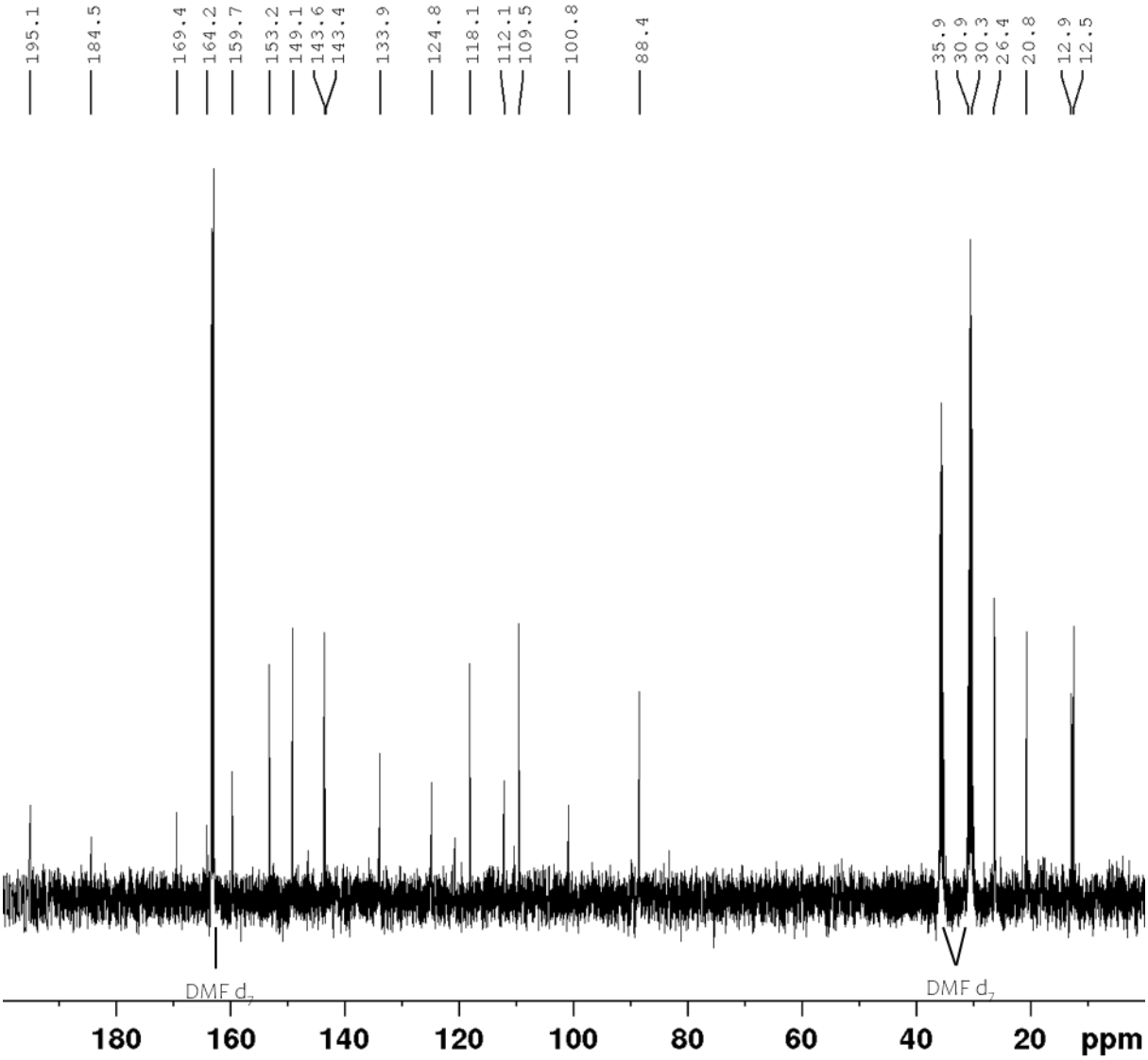


Figure S16: COSY NMR spectrum (DMF-*d7*) of compound **22.**

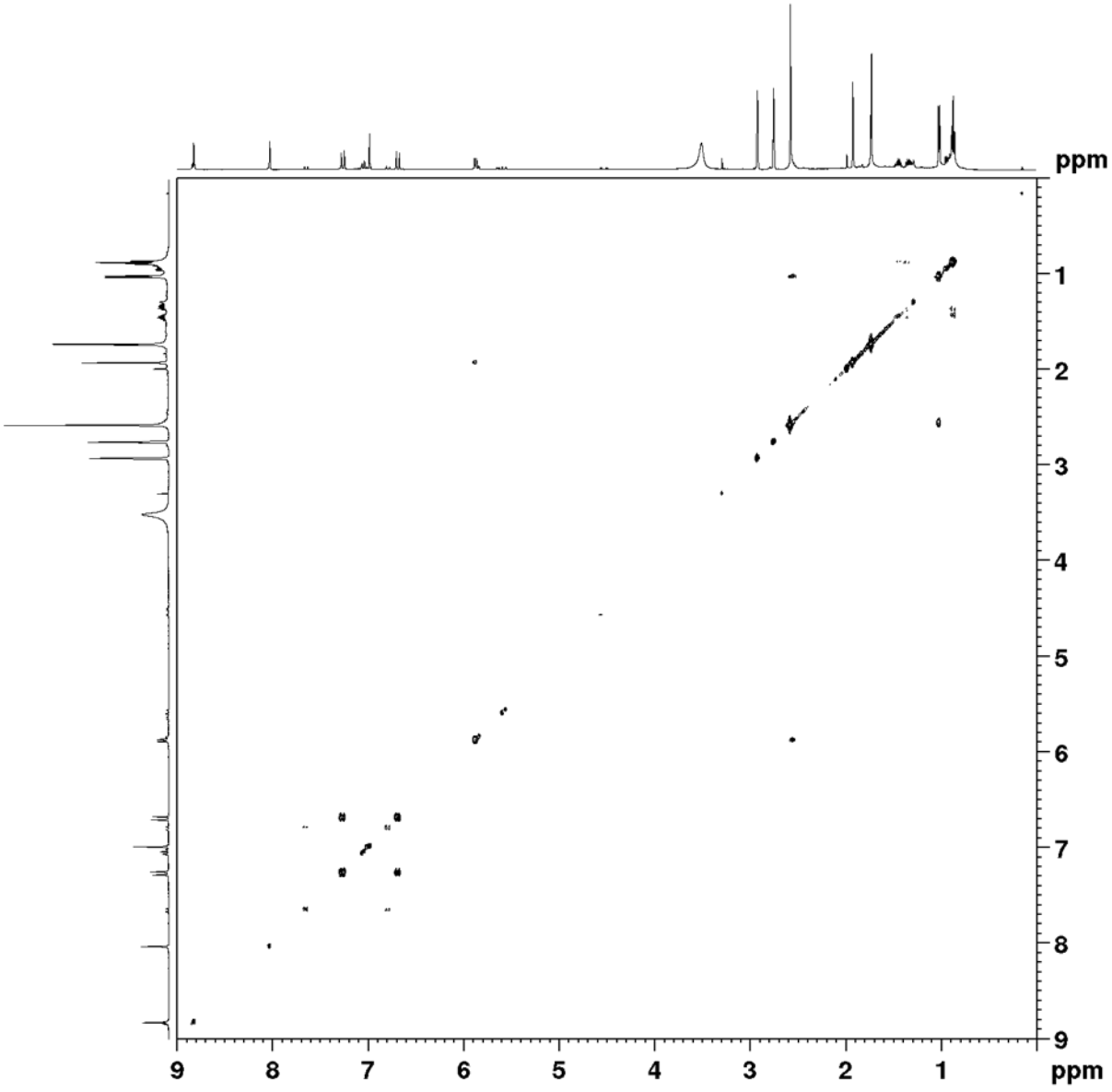


Figure S17: HSQC NMR spectrum (DMF-*d7*) of compound **22.**

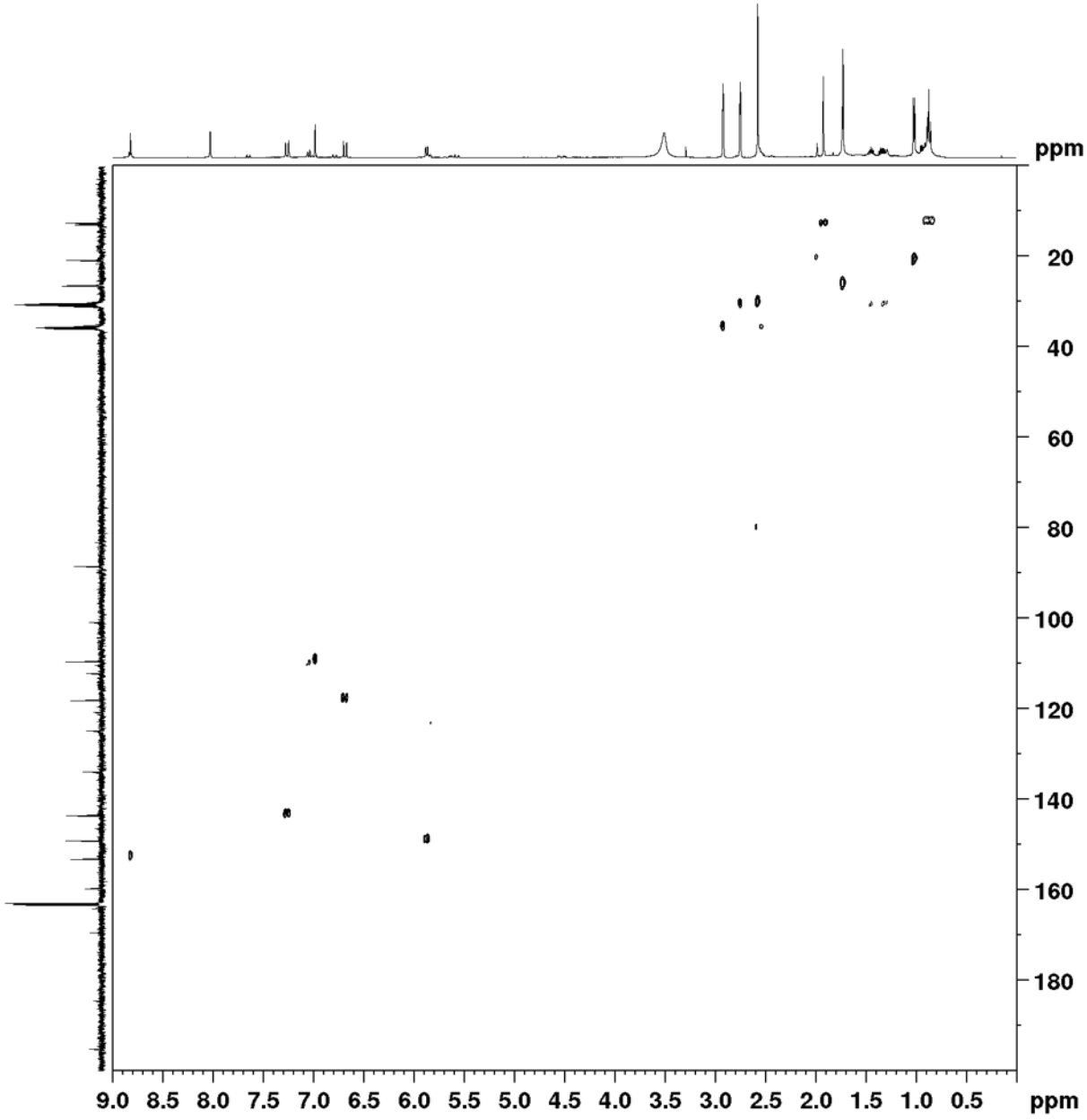
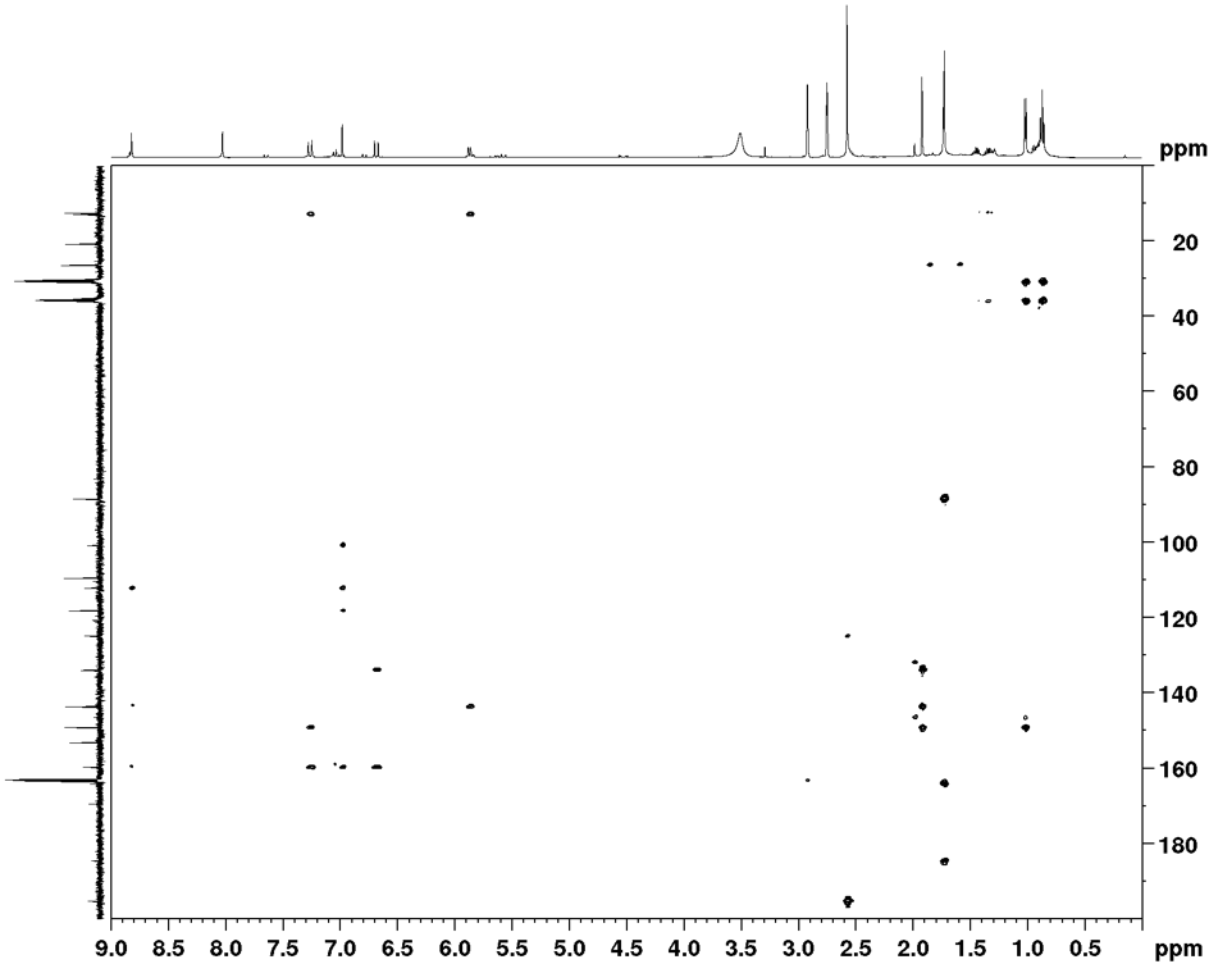


Figure S18: HMBC NMR spectrum (DMF-*d7*) of compound **22.**



References

1. Aldridge, D. C. *et al.* Metabolites of *Nectria coccinea*. *J. Chem. Soc., 1* 2136 (1972) doi:10.1039/p19720002136.
2. Sorres, J., Sabri, A., Brel, O., Stien, D. & Eparvier, V. Ilicicolinic acids and ilicicolinal derivatives from the fungus *Neonectria discophora* SNB-CN63 isolated from the nest of the termite *Nasutitermes corniger* found in French Guiana show antimicrobial activity. *Phytochemistry,* **151**, 69–77 (2018).
3. Nirma, C., Eparvier, V. & Stien, D. Antibacterial ilicicolinic acids C and D and ilicicolinal from *Neonectria discophora* SNB-CN63 isolated from a termite nest. *J. Nat. Prod.,* **78**, 159–162 (2015).
4. Li, Y., Liu, D., Cen, S., Proksch, P. & Lin, W. Isoindolinone-type alkaloids from the sponge-derived fungus Stachybotrys chartarum. *Tetrahedron,* **70**, 7010–7015 (2014).
5. Takamatsu, S. *et al.* A Novel Testosterone 5.ALPHA.-Reductase Inhibitor, 8',9'-Dehydroascochlorin Produced by *Verticillium* sp. FO-2787. *Chem. Pharm. Bull.,* **42**, 953–956 (1994).
6. Mori, K. & Takechi, S. Synthesis of the natural enantiomers of ascochlorin, ascofuranone and ascofuranol. *Tetrahedron,* **41**, 3049–3062 (1985).
7. Hebra, T. *et al.* Dereplication, annotation, and characterization of 74 potential antimicrobial metabolites from *Penicillium sclerotiorum* using t-SNE molecular networks. *Metabolites,* **11**, 444 (2021).
8. Gao, J.-M., Yang, S.-X. & Qin, J.-C. Azaphilones: chemistry and biology. *Chem. Rev.,* **113**, 4755–4811 (2013).
9. Liu, Z. *et al.* Identification of anti-inflammatory polyketides from the coral-derived fungus *Penicillium sclerotiorun*: In vitro approaches and molecular-modeling. *Bioorg. Chem.,* **88**, 102973 (2019).
10. Son, S. *et al.* Structures and biological activities of azaphilones produced by *Penicillium* sp. KCB11A109 from a ginseng field. *Phytochemistry,* **122**, 154–164 (2016).
11. Arai, N. *et al.* Isochromophilones III -VI, inhibitors of acyl-CoA: cholesterol acyltransferase produced by *Penicillium multicolor* FO-3216. *J. Antibiot.,* **48**, 696–702 (1995).
12. Huang, X., Zhang, W., Tang, S., Wei, S. & Lu, X. Collaborative biosynthesis of a class of bioactive azaphilones by two separate gene clusters containing four PKS/NRPSs with transcriptional crosstalk in fungi. *Angew. Chem. Int. Ed.,* **59**, 4349–4353 (2020).