Article

Not peer-reviewed version

Chemical Investigation of the Mediterranean Sponge Crambe crambe by UHPLC-HRMS/MS via manual and computational dereplication approaches

Pinelopi Vlachou , Nikolaos Tsafantakis , Nikola Milic , Alexandros Polyzois , Eirini Baira , Alkaterini Termentzi , Géraldine Le Goff , Jamal Ouazzani , Nikolas Fokialakis *

Posted Date: 28 October 2024

doi: 10.20944/preprints202410.2036.v1

Keywords: Crambe crambe; UHPLC-HRMS/MS; dereplication; computational mass spectrometry; molecular networking; guanidine alkaloids; crambescin analogues

 doi:10.20944/preprints202410.2036.v1

Article

Chemical Investigation of the Mediterranean Sponge *Crambe crambe* by UHPLC-HRMS/MS via Manual and Computational Dereplication Approaches

Pinelopi Vlachou [1], Nikolaos Tsafantakis [1], Nikola Milic [1], Alexandros Polyzois [1], Eirini Baira [2], Aikaterini Termentzi [2], Géraldine Le Goff [3], Jamal Ouazzani [3] and Nikolas Fokialakis [1,*]

[1] Laboratory of Pharmacognosy & Natural Products Chemistry, Department of Pharmacy, National and Kapodistrian University of Athens, 15771 Athens, Greece
[2] Laboratory of Toxicological Control of Pesticides, Department of Pesticides Control and Phytopharmacy, Benaki Phytopathological Institute, 8 St. Delta Street, Kifissia, 14561 Athens, Greece
[3] Institut de Chimie des Substances Naturelles ICSN, Centre National de la Recherche Scientifique, Avenue de la Terrasse, 91198 Gif-sur-Yvette, France
* Correspondence: fokialakis@pharm.uoa.gr

Abstract: The CH2Cl2-MeOH extract of the Mediterranean sponge *Crambe crambe* was investigated via UHPLC-HRMS/MS employing manual dereplication and in silico mass spectrometry tools. A deconvolution approach was implemented toward the extensive metabolic characterization of the sample, resulting in the annotation of 53 compounds. The analysis of data-dependent HRMS/MS scans was conducted to establish fragmentation patterns characteristic of each crambescin A, B and C sub-families. Among the 39 compounds identified from these groups, 22 analogues were reported for the first time including 4 new homologous series that differed by the ratio of methylene units in the upper (n+2) and lower (m+2) alkyl side chains. More specifically, crambesins presenting m= 5 or 6 and n=5 (compounds 7, 11, 22 and 24) as well as m=5 or 6 and n=4 (compounds 5, 6, 8, 9, 12 and 14) were characterized. Additionally, 4 new crambescidin analogues (compounds 13, 15, 35 and 39) were also reported. The identity of the dereplicated features was further validated by studying crambescins spectral similarities through feature-based molecular networking approach. Overall, this study suggests UHPLC-HRMS/MS—through integration of manual and computational dereplication approaches—as a valuable tool for the investigation and high-throughput characterization of *C. crambe* metabolome.

Keywords: *Crambe crambe*; UHPLC-HRMS/MS; dereplication; computational mass spectrometry; molecular networking; guanidine alkaloids; crambescin analogues

1. Introduction

Crambe crambe is a red encrusting sponge of the order *Poecilosclerida* widespread along the sublittoral of the western Mediterranean sea and the Macaronesian archipelagos [1]. Extensive studies have been carried out in this model sponge with regard to its ecology and chemical content [2–4].

The species is well known for the accumulation of a large diversity of guanidine-bearing alkaloids, a group of secondary metabolites that might have a microbial origin [5,6]. Although many years it was believed that *C. crambe* is virtually free of microsymbionts [7], recent studies of the microbiome of the species have reported the presence of a low number of bacteria and fungi, along with a group of Betaproteobacteria which dominates its microbial community and may be involved in biosynthesis of guanidine alkaloids [6,8,9]. However, *C. crambe* is a low microbial abundance sponge harboring much smaller bacterial communities with a lower bacterial diversity comparing to other sponges [10].

 doi:10.20944/preprints202410.2036.v1

The complex guanidine alkaloids derived from the sponge are distributed in two main chemical families, crambescins and crambescidins. The crambescins can be divided into three sub-families crambescins A, B and C. The structure of all crambescins A includes 5,6-fused bicyclic ring system linked to one aliphatic side chain at the C-13 position and a guanidinoalkyl side-chain at the C-7 position. The second type is of a crambescin B skeleton, which is characterized by a [4,5] decane spiro-ring system with similar side chains as crambescin A. Crambescins C represent the third sub-family featured by a monocyclic guanidine ring connected to a linear 3-hydroxypropyl chain and two side chains [11,12] (**Figure 1**). The members of crambescidin family have a pentacyclic guanidine core and are named according to their molecular weight [13–15].

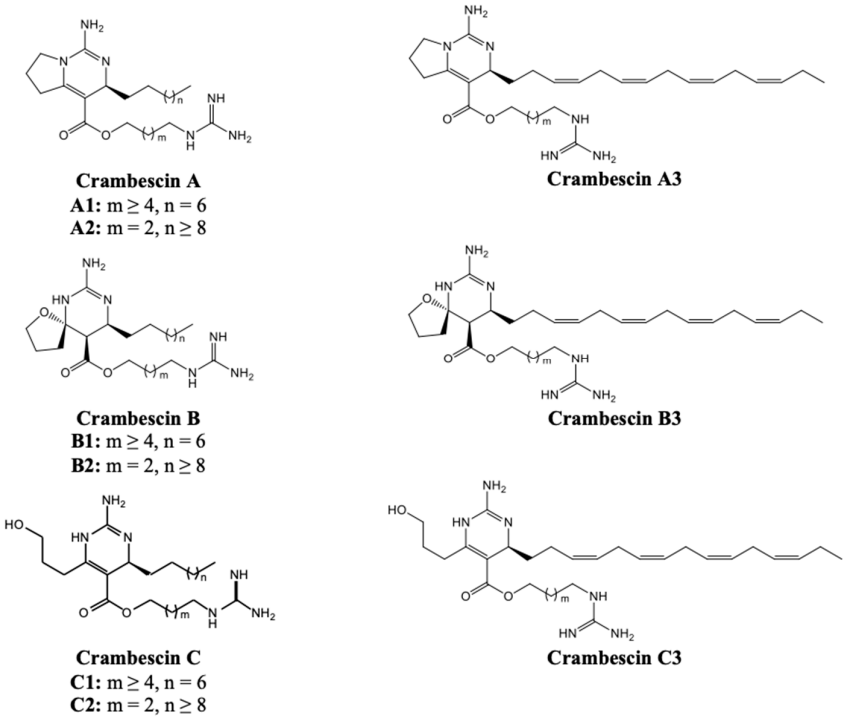

Figure 1. Structures of crambescin sub-families described in the literature.

These specialized metabolites biosynthesized in *C. crambe* are chemotaxonomic markers for *Crambeidae* family and exhibit an outstanding range of biological activities. Several studies have demonstrated the antifungal, antimalarial, antiviral and cytotoxic activity against a diverse panel of human tumor cell lines from members of crambescidin family, while crambescins present a more specific activity on ion channels [14,16–23]. Being one of the strongest organic bases, guanidine might be responsible for the biological activity of these molecules, due to the ionic pocket in the guanidine nucleus, which can interact with biopolymers through hydrogen bonds and/or electrostatic interactions [24]. The unique chemical architectures of polycyclic guanidine alkaloids, along with the broad spectrum of their biological activities make them particularly attractive as starting points for drug development.

The complexity of *C. crambe* metabolic profile lies in the presence of multiple crambescin homologous series constituting a mixture of compounds that display a high degree of structural similarity. In that context, UHPLC-HRMS/MS associated with dereplication approach represents a valuable tool for the high-throughput characterization of the sponge's metabolome. Molecular networking and computational processing contributes on further analysis of MS/MS data. Molecular networking was first introduced as a dereplication strategy over a decade ago, with a key publication in 2013 demonstrating its ability to organize mass spectrometry data and aid in identifying known compounds in complex biological mixtures based on their spectral similarities [25]. Since then, the approach has evolved significantly, with various advancements and applications reported in the

 doi:10.20944/preprints202410.2036.v1

literature [26,27]. Regarding sea sponges, molecular networking as a dereplication strategy has been employed successfully in different species, in some cases leading to the discovery of novel compounds, and even entirely new classes of compounds [28,29].

In the present study, the dichloromethanic-methanolic (CH_2Cl_2-MeOH) extract of *C. crambe* was analyzed by UHPLC-HRMS/MS. The robust mass spectral data were processed and dereplicated employing modern mass spectrometry computational tools that utilize machine learning algorithms and in silico fragmentation to aid in structure elucidation and chemical class prediction [30–34]. Molecular networking visualization was conducted in the comprehensive Cytoscape environment [35] by integrating data from both molecular networking and manual dereplication approaches.

2. Results and Discussion

Figure Our study resulted in the annotation of the guanidine alkaloid class of natural products and, in total, 53 compounds of this class were manually dereplicated, including 22 new analogues of crambescin and 4 new analogues of crambescidin, which are reported here for the first time.

All detected compounds were more abundant both in number and intensity in the ESI(+) mode in comparison to the ESI(-). Consequently, we focused our study only on the ESI(+). The deconvolution process yielded in the detection of 53 compounds within the crude extract as presented in *Error! Reference source not found.*. The two internal standards, yohimbine and reserpine were detected as intense peaks eluted at 9.80 and 12.59 min, organizing into own distinct molecular clusters in the MN analysis. Observed m/z corresponded to [M+H]+ ions and were consistent with their molecular formulas, while characteristic fragments ions were detected with a mass error < 5 ppm. These results demonstrated the suitability of the system regarding mass accuracy at both HRMS and HRMS/MS levels and a mass error of ± 5 ppm was further considered for the annotation of deconvoluted peaks. All 53 annotated metabolites belonged to guanidine alkaloid family, most of them being detected as multiple charged species. Annotation of crambescin analogues was performed according to previous studies suggesting a conventional nomenclature based on sub-family type followed by the molecular weight [11,36]. Accordingly, crambescidin analogues were named according to their molecular weight. Crambescins and guanidine alkaloids in general, are a major component of the overall chemical profile. That can be seen in the sunburst plots visualizing the LC-MS/MS quantification data (**Figure 2**).

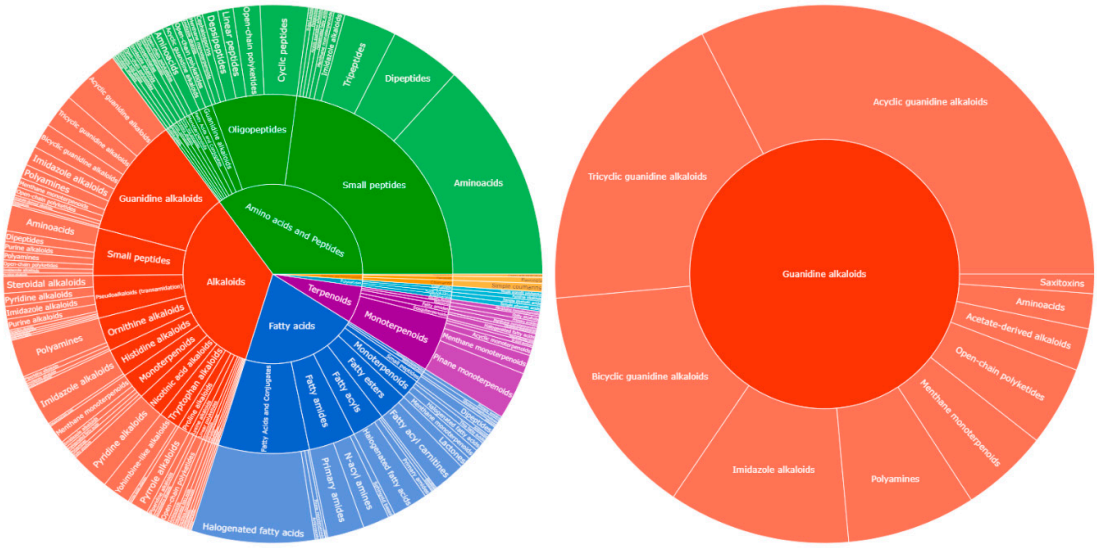

Figure 2. Sunburst plot visualization of relative quantification of Natural Product Classes (NPC), as predicted by CANOPUS ClassyFire, of total chemical profile's NPCs (left chart), and subclasses of guanidine alkaloids (right chart).

 doi:10.20944/preprints202410.2036.v1

Specifically, these 53 annotated metabolites are the dominant features as seen in (***Figure 2***), and when visualized via molecular networking they seed four clusters of crambescins-related compounds (Figure 3). The GNPS2 workflow resulted in the construction of a feature-based molecular network (FBMN) consisting of 1024 nodes (consensus MS2 spectra) and 2629 edges. Following this, we visualized the integrated SIRIUS results onto the MN, concerning the compound classes found in the C. crambe extract, as predicted by SIRIUS built-in tool CANOPUS. The visualized results indicated the presence of guanidine alkaloid compounds in the C. crambe extract, aligning with our previous findings [37] and the existing literature [11,38]. As illustrated in ***Figure 3***, by overlaying the results of the manual dereplication approach onto the constructed FBMN, a compelling clustering pattern emerged. Specifically, we observed several guanidine-containing clusters, collectively consisting of 223 nodes, connected with 537 edges in total.

It is important to emphasize that, based on the constructed molecular network, the interpretation suggests that most of the manually dereplicated A, B and C crambescins form three main clusters depending on their side chain moieties ("1", "2" or "3", rather than their main core moieties (monocyclic, bicyclic, spiro-ring). Thus, we can observe three major clusters: one primarily composed of A1-B1-C1, a second mainly featuring A2-B2-C2, and a third predominantly consisting of A3-B3-C3 crambescins.

One of the clusters (comprising 87 nodes connected by 227 edges), encompassed nodes annotated as CANPOPUS-predicted guanidine alkaloids, alongside those corresponding mainly to crambescins with "1"-type sidechain moieties (crambescins A1, B1 and C1). Another cluster (comprising 55 nodes connected by 149 edges), also contained nodes annotated as guanidine alkaloids, as well as those corresponding mainly to crambescins with "3"-type sidechain moieties (crambescins A3, B3 and C3). The next characteristic guanidine-alkaloids-containing cluster (21 nodes, 29 edges) encompassed nodes corresponding to crambescins with "2"-type sidechain moieties (crambescins A2, B2 and C2). Finally, the last two clusters encompassed features corresponding to crambescidins and guanidine-like compounds, respectively.

The MN inspection confirmed that guanidine-containing clusters, as predicted by CANOPUS, include most nodes linked to the manually dereplicated compounds of interest. This finding opens avenues for exploring adjacent "unknown" nodes. Additionally, visual analysis effectively differentiated clusters of crambescins, crambescidins, and guanidine-related compounds.

 doi:10.20944/preprints202410.2036.v1

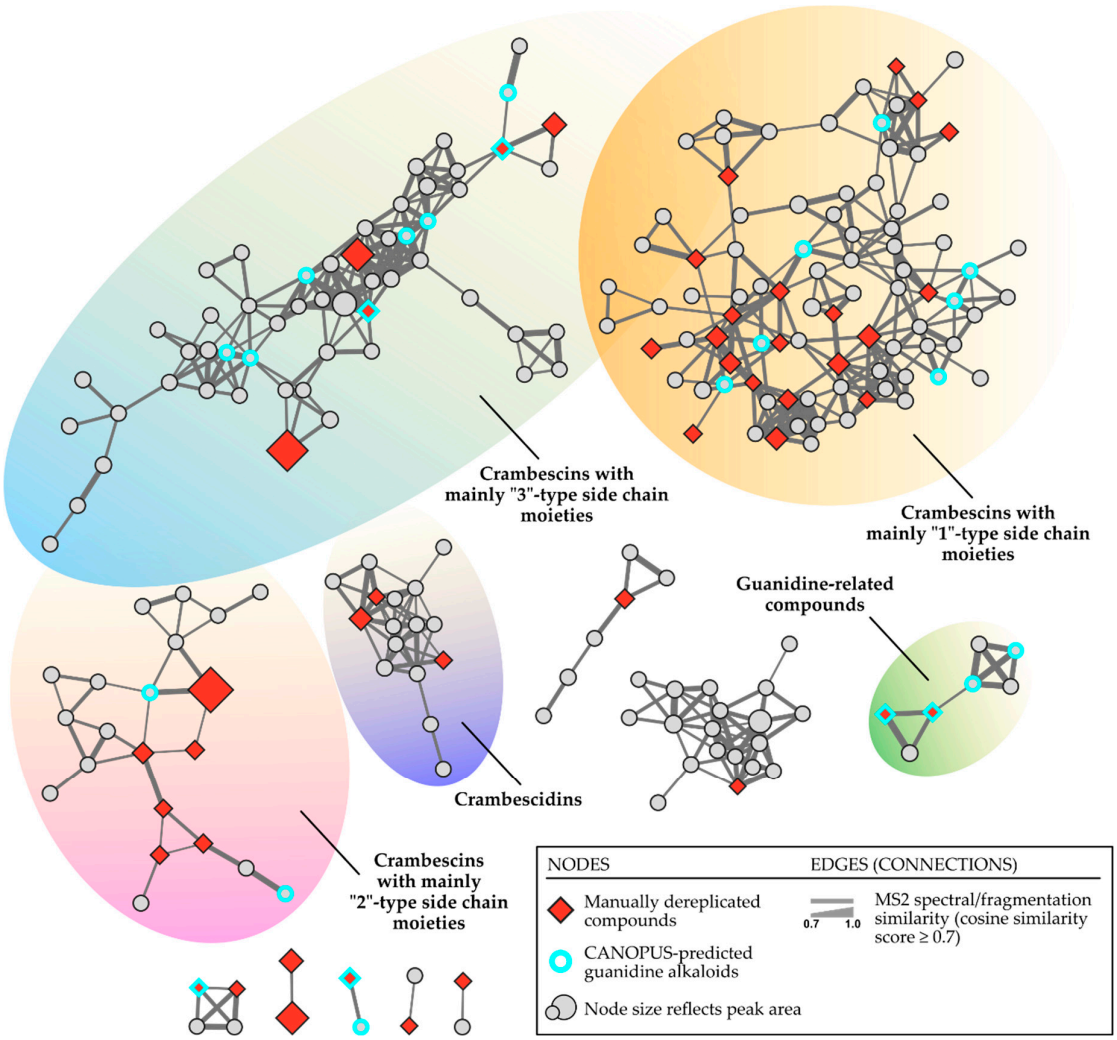

Figure 3. Feature-based molecular networking (FBMN) of the *Crambe crambe* extract visualized in Cytoscape. Refer to the legend for an explanation of the colors and shapes of the nodes, along with details about the clusters. The MN was visually enhanced using the InkScape vector graphics editor [39].

Table 1. Compounds detected in *C. crambe* extract by ESI(+)-UHPLC-HRMS/MS.

#	*m/z*	Rt	Charge state, *z*	M_w exp.[a]	Proposed Formula	Δ(ppm)[b]	Major MS/MS fragments: *m/z* (charge state, *z*)	Proposed Identification	Ref.
1	282.1806	8.16	1	281.1728	$C_{14}H_{23}N_3O_3$	-2.28	246.1587 (1); 264.1706 (1); 114.9612 (1); 60.0562 (1)	guanidine related compound ($C_{14}H_{23}N_3O_3$)	
2	282.1806	9.34	1	281.1728	$C_{14}H_{23}N_3O_3$	-2.13	114.9611 (1); 60.0562 (1); 223.1323 (1)	guanidine related compound ($C_{14}H_{23}N_3O_3$)	
3	264.1704	9.66	1	263.1626	$C_{14}H_{21}N_3O_2$	-0.99	246.1604 (1); 60.0562 (1)	guanidine related compound ($C_{14}H_{21}N_3O_2$)	
IS	355.2012	9.80	1	354.1934	$C_{21}H_{26}N_2O_3$	-0.75	144.0807 (1); 212.1274 (1)	Yohimbine	
4	296.1964	10.44	1	295.1886	$C_{15}H_{25}N_3O_3$	-1.52	237.1476 (1); 205.1217 (1); 60.0562 (1); 159.1160 (1)	guanidine related compound ($C_{15}H_{25}N_3O_3$)	
5	227.1803	10.46	2	452.3450	$C_{23}H_{44}N_6O_3$	-3.31	174.1600 (1); 148.6021 (2)	crambescin C 452	

 doi:10.20944/preprints202410.2036.v1

6	234.1884	10.77	2	466.3612	$C_{24}H_{46}N_6O_3$	-1.93	188.1756 (1); 155.6100 (2)	crambescin C 466 homologue (m=6, n=4)	
7							174.1600 (1); 148.6021 (2)	crambescin C 466 homologue (m=5, n=5)	
8	218.1757	10.87	2	434.3357	$C_{23}H_{42}N_6O_2$	-0.65	197.1646 (2); 174.1600 (1); 148.1095 (2); 220.1689 (1)	crambescin A 434 homologue (m=5, n=4)	
9	227.1803	10.88	2	452.3450	$C_{23}H_{44}N_6O_3$	-3.32	128.1431 (1); 174.1599 (1); 111.0442 (1); 284.1956 (1)	crambescin B 452 homologue (m=5, n=4)	
10	234.1884	10.98	2	466.3612	$C_{24}H_{46}N_6O_3$	-1.91	160.1441 (1); 141.5942 (2)	crambescin C1 466 (m=4, n=6)	[11]
11	241.1961	11.14	2	480.3765	$C_{25}H_{48}N_6O_3$	-2.42	188.1756 (1); 155.6096 (2)	crambescin C 480 homologue (m=6, n=5)	
12	225.1835	11.16	2	448.3514	$C_{24}H_{44}N_6O_2$	-0.45	204.1722 (2); 188.1756 (1); 155.1178 (2); 220.1689 (1)	crambescin A 448 homologue (m=6, n=4)	
13	292.8887	11.18	3	875.6426	$C_{44}H_{87}N_6O_{11}$	1.73	246.1587 (1); 139.0751 (1); 162.1598 (1); 381.3460 (1)	crambescidin 875	
14	234.1884	11.18	2	466.3612	$C_{24}H_{46}N_6O_3$	-1.93	128.1432 (1); 111.0443 (1); 188.1757 (1); 298.2129 (1)	crambescin B 466 homologue (m=6, n=4)	
15	279.2133	11.19	3	834.6166	$C_{45}H_{82}N_6O_8$	-1.01	246.1586 (1); 264.1707 (1); 70.0657 (1); 139.0750 (1) [d]	crambescidin 834	
16	241.1961	11.24	2	480.3765	$C_{25}H_{48}N_6O_3$	-2.46	174.1600 (1); 148.6021 (2)	crambescin C1 480 (m=5, n=6)	[11]
17	218.1757	11.27	2	434.3357	$C_{23}H_{42}N_6O_2$	-0.58	127.0864 (2); 197.1646 (2); 132.1130 (1); 262.2150 (1)	crambescin A2 434 (m=2, n=7)	
18	249.1829	11.33	2	496.3502	$C_{28}H_{44}N_6O_2$	-2.79	114.1028 (1); 384.2650 (1); 132.1131 (1); 228.1724 (2)	crambescin A3 496 (m=2) (cis)	
19	225.1834	11.36	2	448.3513	$C_{24}H_{44}N_6O_2$	-0.78	204.1722 (2); 141.1018 (2); 160.1441 (1); 248.2002 (1)	crambescin A1 448 (m=4, n=6)	
20	234.1884	11.43	2	466.3612	$C_{24}H_{46}N_6O_3$	-1.93	132.1131 (1); 127.5783 (2)	crambescin C2 466 (m=2, n=8)	
21							156.1748 (1); 111.0443 (1); 160.1441 (1)	crambescin B1 466 (m=4, n=6)	[11]
22	248.2038	11.49	2	494.3921	$C_{26}H_{50}N_6O_3$	-2.46	188.1756 (1); 155.6091 (2)	crambescin C1 494 (m=6, n=6)	[11]
23	238.1909	11.51	2	474.3662	$C_{26}H_{46}N_6O_2$	-2.02	188.1756 (1); 217.1806 (2); 155.1178 (2); 170.1650 (1)	didehydrocrambescin A1 474 (m=6, n=6)	
24	241.1960	11.55	2	480.3765	$C_{25}H_{48}N_6O_3$	-2.58	142.1585 (1); 111.0443 (1); 188.1756 (1); 298.2106 (1)	crambescin B 480 homologue (m=6, n=5)	
25	256.1910	11.55	2	510.3664	$C_{29}H_{46}N_6O_2$	-1.57	128.1178 (1); 384.2652 (1); 235.1809 (2); 146.1287 (1)	crambescin A3 510 (m=3) (cis)	
26	249.1829	11.62	2	496.3502	$C_{28}H_{44}N_6O_2$	-2.73	132.1131 (1); 114.1028 (1); 127.0864 (2); 384.2638 (1)	crambescin A3 496 (m=2) (trans)	

 doi:10.20944/preprints202410.2036.v1

27	232.1909	11.62	2	462.3663	$C_{25}H_{46}N_6O_2$	-1.99	211.1798 (2); 174.1599 (1); 148.1093 (1); 248.2003 (1)	crambescin A1 462 (m=5, n=6)	[11]
28	241.1961	11.68	2	480.3765	$C_{25}H_{48}N_6O_3$	-2.46	156.1749 (1); 111.0443 (1); 174.1600 (1); 284.1957 (1)	crambescin B1 480 (m=5, n=6)	[11]
29	248.2039	11.68	2	494.3921	$C_{26}H_{50}N_6O_3$	-2.40	132.1131 (1); 127.0865 (2); 114.1026 (1)	crambescin C2 494 (m=2, n=10)	
30	225.1835	11.70	2	448.3513	$C_{24}H_{44}N_6O_2$	-0.64	127.0863 (2); 204.1721 (2); 132.1130 (1); 276.2321 (1)	crambescin A2 448 (m=2, n=8)	[11]
31	272.2038	11.71	2	542.3920	$C_{30}H_{50}N_6O_3$	-2.40	160.1441 (1); 142.1335 (1)	crambescin C3 542 (m=4)	[12]
32	267.8778	11.73	3	800.6100	$C_{45}H_{80}N_6O_6$	-2.50	70.0657 (1); 349.2648 (2); 206.1536 (1); 392.3078 (2) [d]	crambescidin 800 or isocrambescidin 800	[13,14]
33	263.1990	11.79	2	524.3824	$C_{30}H_{48}N_6O_2$	-0.76	142.1334 (1); 384.2649 (1); 160.1440 (1); 242.1875 (2)	crambescin A3 524 (m=4) (cis)	[12]
34	256.1910	11.81	2	510.3664	$C_{29}H_{46}N_6O_2$	-1.65	146.1287 (1); 134.0943 (2); 128.1178 (1); 235.1805 (2)	crambescin A3 510 (m=3) (trans)	
35	286.8851	11.83	3	857.6318	$C_{44}H_{85}N_6O_{10}$	1.41	264.1703 (1); 139.0749 (1); 246.1602 (1); 381.3456 (1)	crambescidin 857	
36	273.2094	11.84	3	816.6048	$C_{45}H_{80}N_6O_7$	-2.53	264.1707 (1); 246.1587 (1); 139.0751 (1); 70.0657 (1) [d]	crambescidin 816	[14]
37	234.1884	11.88	2	466.3613	$C_{24}H_{46}N_6O_3$	-1.77	184.2054 (1); 132.1131 (1); 114.1028 (1); 242.1484 (1)	crambescin B2 466 (m=2, n=8)	
38	239.1989	11.90	2	476.3821	$C_{26}H_{48}N_6O_2$	-1.39	188.1756 (1); 156.174 (1); 111.0443 (1); 218.1873 (2)	crambescin A1 476 (m=6, n=6)	
39	281.5530	11.94	3	841.6355	$C_{44}H_{85}N_6O_9$	-0.20	263.1982 (1); 70.0657 (1); 116.1071 (1); 139.0751 (1)	crambescidin 841	
40	267.8778	11.94	3	800.6099	$C_{45}H_{80}N_6O_6$	-2.58	70.0657 (1); 349.2647 (2); 392.3078 (2); 206.1537 (1) [d]	crambescidin 800 or isocrambescidin 800	[13,14]
41	248.2039	11.94	2	494.3921	$C_{26}H_{50}N_6O_3$	-2.34	156.1749 (1); 111.0444 (1); 188.1756 (1); 298.2130 (1)	crambescin B1 494 (m=6, n=6)	[11]
42	232.1910	11.98	2	462.3664	$C_{25}H_{46}N_6O_2$	-1.73	211.1799 (2); 134.0942 (2); 146.1287 (1); 276.2325 (1)	crambescin A 462 homologue (m=3, n=8)	
43	238.1908	11.98	2	474.3659	$C_{26}H_{46}N_6O_2$	-2.74	127.0864 (2); 132.1131 (1); 217.1807 (2); 114.1028 (1)	didehydrocrambescin A2 474 (m=2, n=10)	
44	270.2066	12.03	2	538.3977	$C_{31}H_{50}N_6O_2$	-1.36	497.3828 (1); 384.2652 (1); 522.3766 (1); 174.1602 (1) [c]	crambescin A3 538 (m=5) (cis)	[12]
45	263.1990	12.03	2	524.3825	$C_{30}H_{48}N_6O_2$	-0.63	160.1441 (1); 141.1018 (2); 242.1877 (2); 384.2650 (1)	crambescin A3 524 (m=4) (trans)	[12]
46	232.1910	12.11	2	462.3664	$C_{25}H_{46}N_6O_2$	-1.82	127.0864 (2); 211.1799 (2); 132.113 (1); 290.2473 (1)	crambescin A2 462 (m=2, n=9)	[11]

 doi:10.20944/preprints202410.2036.v1

47	416.3195	12.16	2	830.6233	$C_{46}H_{82}N_6O_7$	0.98	264.1707 (1); 246.1587 (1); 70.0657 (1); 139.0751 (1)	crambescidin 830	[14]
48	272.2037	12.18	2	542.3917	$C_{30}H_{50}N_6O_3$	-2.92	160.1441 (1); 111.0443 (1); 232.2046 (1); 274.2275 (1)	crambescin B3 542 (m=4)	[12]
49	270.2066	12.27	2	538.3976	$C_{31}H_{50}N_6O_2$	-1.49	174.1597 (1); 148.1098 (2); 249.1955 (2); 156.1493 (1)	crambescin A3 538 (m=5) (trans)	[12]
50	404.2534	12.32	1	403.2463	$C_{22}H_{33}O_4N_3$	-1.74	360.2640 (1); 206.1536 (1); 342.2542 (1); 60.0562 (1)	crambescidin acid	[37]
51	239.1988	12.51	2	476.3820	$C_{26}H_{48}N_6O_2$	-1.65	127.0864 (2); 218.1873 (2); 132.113 (1); 304.2617 (1)	crambescin A2 476 (m=2, n=10)	[11]
52	254.2221	12.51	1	253.2148	$C_{14}H_{27}N_3O$	-1.97	195.1740 (1); 97.0651 (1); 60.0562 (1); 111.0442 (1)	crambescin 253	[37]
IS	609.2800	12.59	1	608.2722	$C_{33}H_{40}N_2O_9$	-0.22	195.0642 (1); 174.0913; 397.2098 (1); 448.1952 (1)	reserpine	
53	282.2534	13.77	1	281.2460	$C_{16}H_{31}N_3O$	-1.42	114.9612 (1); 223.2051 (1); 97.0651 (1); 60.0562 (1)	crambescin 281	[37]

[a] Experimental molecular weight was calculated from the observed m/z as follows: $M_w calc. = (m/z) \times z - (z \times M_H)$ *where* M_H*= 1.0078 Da.* [b] Mass error between the molecular weight calculated from the experimental m/z and from the theoretical formula as follows: $\Delta(ppm) = 10^6 \times \frac{(M_w\ exp. - M_w\ theor.)}{M_w\ theor.}$. For determination of the theoretical molecular weight, the following monoisotopic mass values were considered for each element H, C, N and O: M_H*= 1.0078 Da,* M_C*=12 Da,* M_N*=14.0031 Da,* M_O*=15.9949 Da.* [40]. [c] MS/MS data dependent scans were not triggered on the $[M+2H]^{2}$ ion, data from the MS/MS fragmentation of the $[M+H]^+$ ion is reported. [d] data from the MS/MS fragmentation of the $[M+2H]^{2+}$ ion is reported.

2.1. Crambescins A

The most abundant metabolite of the extract (compound **30**) was detected at 11.71 min with *m/z* 225.1835 corresponding to a $[M+2H]^{2+}$ consistent with a molecular formula of $C_{24}H_{44}N_6O_2$. In the MN analysis, the corresponding feature node is a part of the crambescins "2"-type side chain moiety cluster, connected with another A2 crambescin (compound 51) and an annotated node congruent with CANOPUS-derived predictions for guanidine compounds. The single charged ion $[M+H]^+$ was also observed at *m/z* 449.3600 (Δ 1.34 ppm) with an intensity 50 times lower than this of the doubled charged ion. The MS/MS spectrum of the $[M+2H]^{2+}$ ion displayed characteristic fragments which provided insights on both alkyl and guanidinoalkyl side chains (Figure 4). More specifically, the loss of methylenediamine (-42 Da, CH_2N_2) yielded in the *m/z* 204.1721 fragment ion ($[C_{23}H_{44}N_4O_2]^{2+}$, Δ -1.47 ppm) and further dissociation of the aliphatic side chain (-154 Da, $C_{11}H_{22}$) resulted in the *m/z* 127.0863 ($[C_{12}H_{22}N_4O_2]^{2+}$, Δ -2.36 ppm). Therefore, doubled charged fragments can serve to determine the length of the aliphatic side chain at C-13. On the other hand, singly charged fragment *m/z* 132.1130 ($[C_5H_{14}N_3O]^+$, Δ -3.03 ppm) provided a direct characterization of the guanidinoalkyl side chain. Finally, the fragment ion *m/z* 276.2321 ($[C_{18}H_{30}NO]^+$, Δ 0.36 ppm) could be related to the cleavage of both guanidine cyclic core and guanidine ester. Consequently compound **30** was identified as crambescin A2 448.

 doi:10.20944/preprints202410.2036.v1

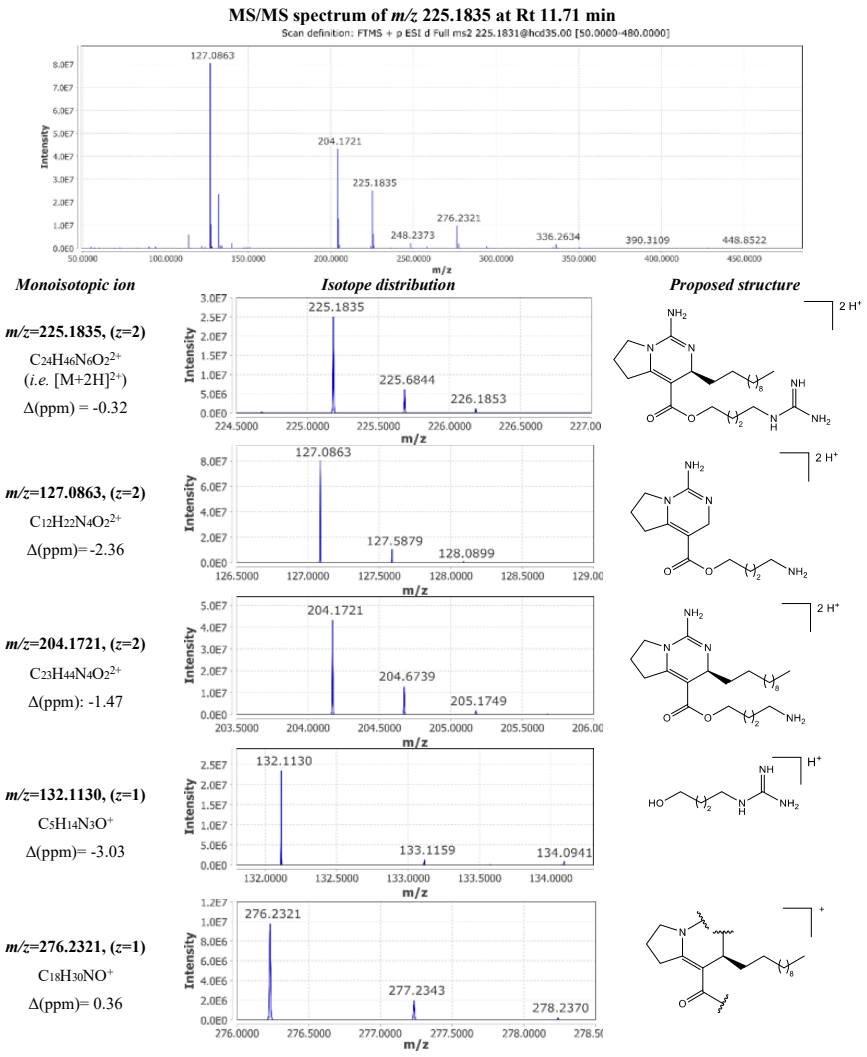

Figure 4. MS/MS fragmentation of Crambescin A2 448 (compound **30**).

Interestingly, this fragmentation pattern led to the characterization of new homologues from the crambescins A family, which displayed different ratios regarding the number of methylene units composing the alkyl side chains. MS/MS spectra of compound **8** was consistent with a crambescin A 434 homologue with m=5 and n=4. Double charged fragments with *m*/z 197.1646 and 148.1095 corresponded to $[C_{22}H_{42}N_4O_2]^{2+}$ and $[C_{15}H_{28}N_4O_2]^{2+}$ respectively (Δ, -0.76 and -2.70 ppm). As mentioned previously, neutral loss between these two fragments corresponded to the dissociation of the C-13 alkyl side chain i.e. C_7H_{14}, while guanidinoalkyl chain was characterized by the fragment *m*/z 174.1600 ($[C_8H_{20}N_3O]^+$, Δ -1.15 ppm).

By analogy to this fragmentation pattern, 10 compounds of crambescin A group were annotated. Compounds **17** and **19** were identified as crambescin A2 434 and crambescin A1 448. Crambescin A1 462 and crambescin A2 462 were assigned to compounds **27** and **46** while compounds **38** and **51** were respectively corresponded to crambescin A1 476 and crambescin A2 476. The oxidized forms of **38** and **51**, which were consistent with a formula of $C_{24}H_{46}N_6O_2$, were detected at 11.51 min (didehydrocrambescin A1 474, compound **23**) and at 11.98 min (didehydrocrambescin A2 474, compound **43**).

Additionally, another two new crambescin A homologues were characterized, compounds **12** and **42**, which were identified as crambescin A 448 (m=6, n=4) and crambescin A 462 (m=3, n=8), respectively.

 doi:10.20944/preprints202410.2036.v1

A series of compounds presenting empirical molecular formulae of the type $C_nH_{2n-12}N_6O_2$ were found to belong to the crambescin A3 sub-family, which were first introduced by Genta-Jouve et al. [41]. These metabolites exhibited a Ring Double Bond Equivalents (RDBE) of 10 which indicates an extra four degrees of unsaturation comparing to other crambescins A ($C_nH_{2n-4}N_6O_2$, RDBE 6). Following a similar fragmentation pathway, neutral loss of 202 Da (-$C_{15}H_{22}$) between double charged fragment ions enabled to confirm that these additional unsaturated bonds were located on the upper aliphatic side chain. Consequently, compounds **18** and **26** were determined to be crambescin A3 496 isomers, compounds **25** and **34** were found to be crambescin A3 510 isomers, compounds **44** and **49** were identified as crambescin A3 538 isomers and compounds **33** and **35** were assigned to crambescin A3 524 isomers. To the authors' knowledge, it is the first time that crambescin A3 isomerism has been reported. The same characteristic fragment ions were observed for each pair of isomers, differing only by their intensity as shown in Figure 5. The annotation of *cis-* and *trans-* isomers was based on their expected relative polarity.

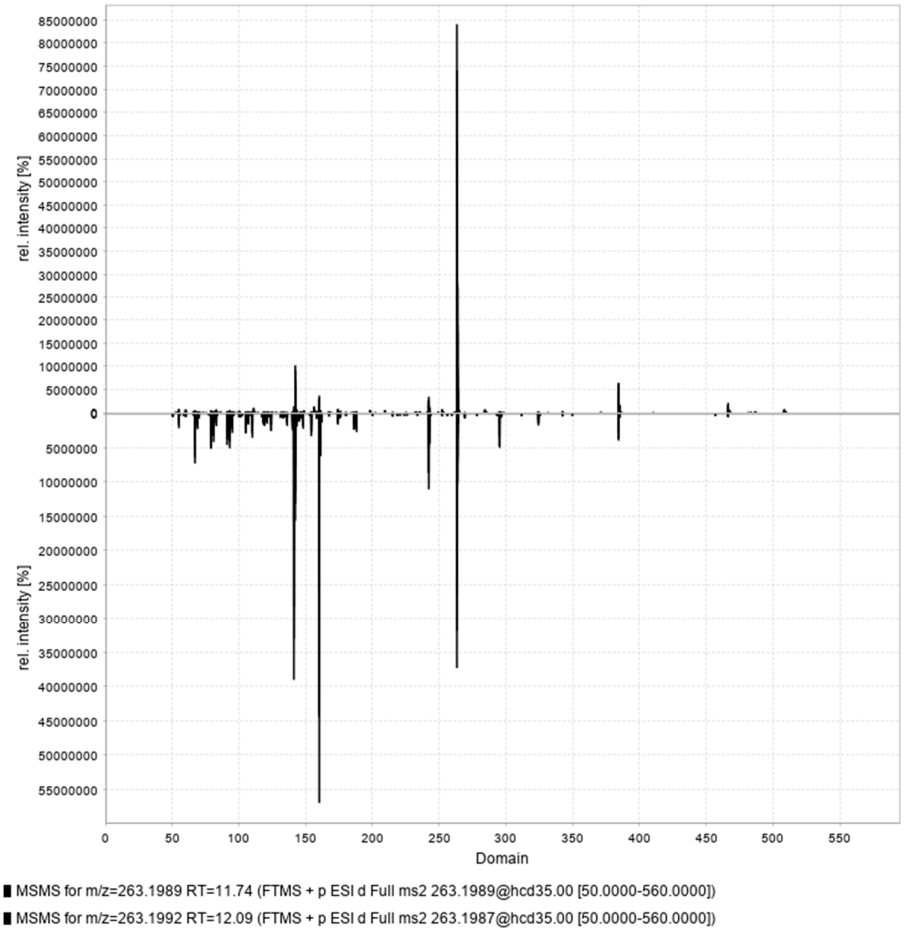

Figure 5. Comparison of MS/MS spectra of cis- and trans- isomers of crambescin A3 524 (compound **33** and **45**, respectively).

2.2. Crambescins B and C

In addition to crambescins A, the investigation of the *C. crambe* metabolic profile revealed the presence of 16 compounds belonging to crambescin B and C families, which exhibited a different fragmentation pattern. These analogues display empirical molecular formulae of the type $C_nH_{2n-2}N_6O_3$ and their discrimination was based on the intensity of specific fragment ions. Due to the presence of two conjugated double bonds in the guanidine ring, members of crambescin C family may undergo a Retro-Diels-Alder reaction less easily comparing to crambescins B, which possess only one intracyclic bond [42].

 doi:10.20944/preprints202410.2036.v1

A pair of isomers consistent with a molecular formula of $C_{25}H_{48}N_6O_3$ were detected at 11.24 min and 11.68 min (compounds **16** and **28**). Following the analysis of the MS/MS spectrum presented in **Figure 6**, the double charged ion $[M+2H]^{2+}$ with *m/z* 241.1961 observed at 11.68 min was identified as the known metabolite crambescin B1 480 (compound **28**). The guanidine cyclic core underwent a pericyclic reaction via a Retro-Diels-Alder mechanism yielding into two fragment ions *m/z* 198.1965 ($[C_{11}H_{24}N_3]^+$, Δ 0.00 ppm) and *m/z* 284.1957 ($[C_{14}H_{26}N_3O_3]^+$, Δ -3.87 ppm). Consequently, loss of methylenediamine from the first fragment resulted in *m/z* 156.1749 ($[C_{10}H_{22}N]^+$, Δ 1.28 ppm), while dissociation of the guanidine ester from the second fragment produced the *m/z* 111.0443 ($[C_6H_7O_2]^+$, Δ -0.90 ppm) and *m/z* 174.1600 ($[C_8H_{20}N_3O]^+$ Δ -1.15 ppm) assigned to the spiroaminal ring and the lower guanidoalkyl chain, respectively. The constructed FBMN clustered crambescin B1 480 with crambescins B1 494, C1 466, and C1 480, creating a larger constellation. The manual dereplication method indicated that the majority of these compounds feature "1"-type moieties or side chains, regardless of the central cyclic guanidine core. Additionally, all these nodes are positioned within the previously mentioned subnetwork that includes CANOPUS-predicted guanidine compounds.

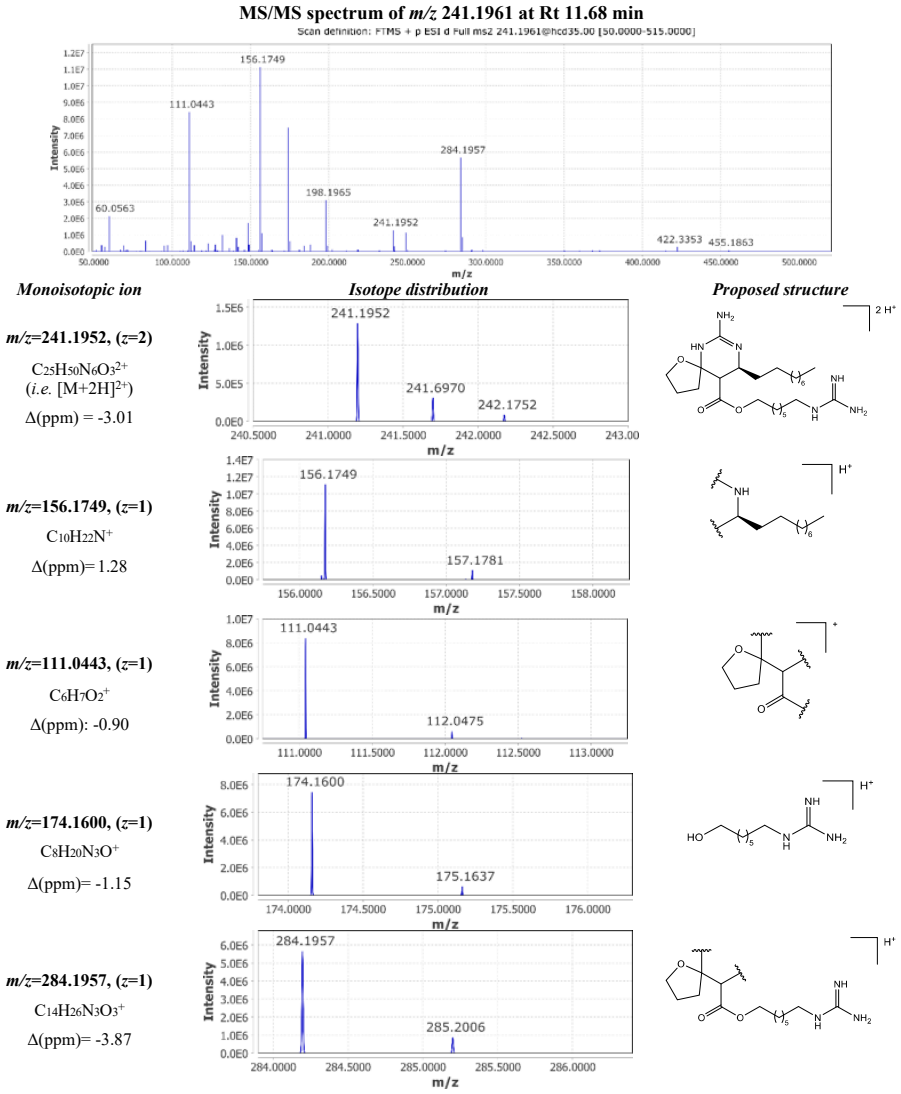

Figure 6. MS/MS fragmentation of Crambescin B1 480 (compound 28).

The fragmentation pattern observed for the metabolite corresponded to the double charged ion $[M+2H]^{2+}$ with *m/z* 241.1961 at 11.24 min was very similar with this of crambescin B1 480. MS/MS spectrum inspection revealed the presence of the same fragments in significantly lower intensity. However, the difference lied in the presence of an intense double charged fragment ion *m/z* 148.6021

 doi:10.20944/preprints202410.2036.v1

($[C_{15}H_{27}N_3O_3]^{2+}$ Δ -0.67 ppm), implying the losses of the upper aliphatic chain and the guanidine moiety of the lower chain. Moreover, the other major fragment with *m/z* 174.1600 ($[C_8H_{20}N_3O]^+$ Δ -1.15 ppm) allowed the determination of the length of the lower aliphatic chain (Figure 7). Based on these observations, compound **16** was identified as crambescin C1 480.

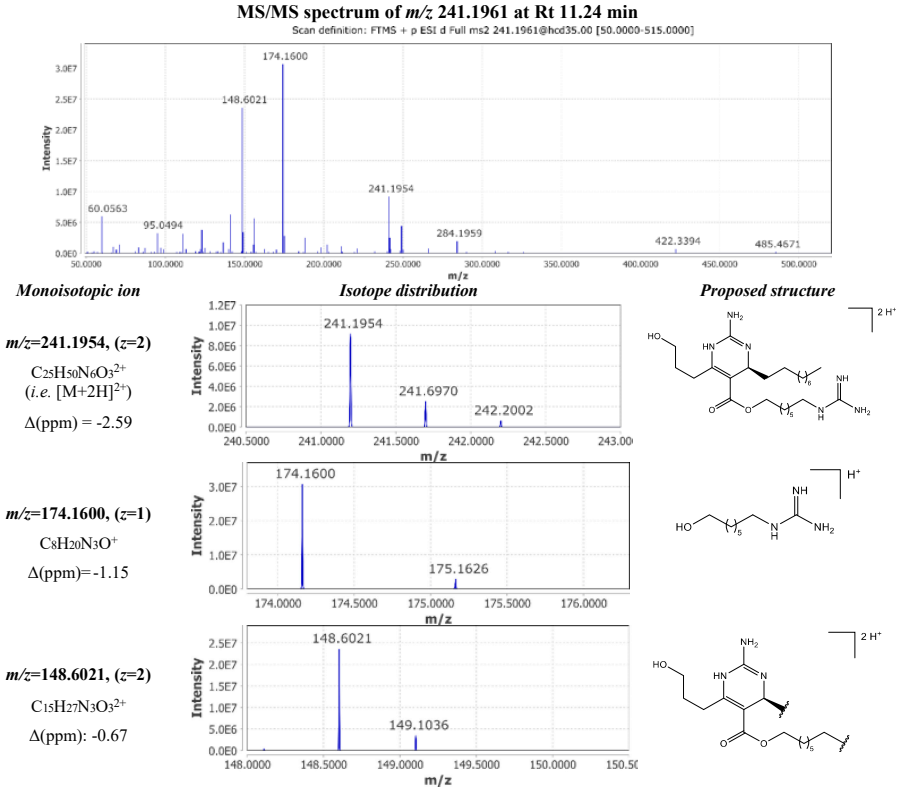

Figure 7. MS/MS fragmentation of Crambescin C1 480 (compound 16).

Another pair of isomers eluting at 11.14 min (compound **11**) and 11.55 min (compound **24**) presented the same molecular formula with crambescins B1 and C1 480. MS/MS spectra of these metabolites displayed the same fragmentation motives, varying only in the length of the side chains. As a result, two new analogues were identified belonging to crambescins B and C group, which bore the same number of methylene units in the lower and upper aliphatic chain (m=6, n=5).

Based on the characteristic fragmentation scheme of crambescins B and C, other analogues were annotated accordingly. Compounds **22**, **29** and **41** were identified as crambescins C1, C2 and B1 494 respectively, compounds **10** and **21** were assigned to crambescins C1 and B1 466 and compounds **20** and **37** corresponded to crambescins C2 and B2 466. Moreover, 5 new analogues were reported for the first time including 2 crambescin C 466 homologues (m=6, n=4; compound **6** and m=5, n=5; compound **7**), a crambescin B 466 with m=6, n=4 (compound **14**), as well as a crambescin C 452 homologue (m=5, n=4; compound **5**) and a crambescin B 452 with m=5, n=4 (compound **9**). Mass spectra and detailed fragmentation patterns for these new analogs are presented in supplementary information.

Finally, 2 compounds presenting a molecular formula of $C_{30}H_{50}N_6O_3$ indicating 9 degrees of unsaturation were found to be the known crambescins C3 and B3 542 (compounds **31** and **48**, respectively) [41] along with 2 minor metabolites, crambescins B 253 and 281 (compounds **52** and **53**) [37] .

2.3. Crambescindins and Other Guanidine Alkaloids

 doi:10.20944/preprints202410.2036.v1

A triple charged ion *m/z* 273.2094 eluting at 11.84 min accounted among the major metabolites detected in the extract. This feature was consistent with a molecular formula of $C_{45}H_{80}N_6O_7$ and was assigned to be crambescidin 816 (compound **37**), a member of crambescidin family which was introduced by Jares-Erijman et al. [14]. Other crambescidin metabolites present at lower concentration were also identified including isomeric forms of crambescindin 800 (compounds **32** and **40**) and crambescidin 830 (compound **47**). Moreover, 4 new crambescindin analogues were reported, compound **13** (crambescidin 875, molecular formula $C_{44}H_{87}N_6O_{11}$), compound **15** (crambescidin 834, molecular formula $C_{45}H_{82}N_6O_8$), compound **35** (crambescidin 857, molecular formula $C_{44}H_{85}N_6O_{10}$), and compound **39** (crambescidin 841, molecular formula $C_{44}H_{85}N_6O_9$).

At MS level, the base peak ion of all crambescidin analogues typically corresponded to the $[M+3H]^{3+}$ form, with the $[M+2H]^{2+}$ ion being present at lower intensity. At MS/MS level, characteristic fragment ions were observed for all the compounds mentioned here above. More specifically, the fragment ion *m/z* 264.17 ($[C_{14}H_{22}N_3O_2]^+$, Δ -1.51 to 0.00 ppm) corresponded to the guanidine core after the loss of the ethyl oxepine moiety and further dehydratation yielded in the *m/z* 246.16. Another common fragment ion, *m/z* 70.0657 ($[C_4H_8N]^+$, Δ 2.85 ppm), was produced by the fragmentation of the hydroxyspermidine unit and subsequent cleavage of the hydroxyl group.

Additionally, 4 guanidine related compounds of lower molecular weight compounds were observed as single charged ions potentially belonging to crambescidin family (compounds **1-4**) along with the known crambescidin acid (compound **50**).

In the context of molecular networking, the manually dereplicated crambescidins, as well as guanidine-related compounds form their own distinct clusters.

3. Materials and Methods

3.1. Chemicals

Analytical grade dichloromethane and methanol that were used for sample extraction were purchased from Sigma-Aldrich (Steinheim, Germany). For sample reconstitution of the crude extract and further UHPLC-HRMS analysis, LC-MS grade solvents were used. More specifically, Acetonitrile (LC-MS grade) was purchased from Fluka/Riedel-de Haën (Seelze, Germany), Methanol (LC-MS grade) was acquired from Sigma-Aldrich (Steinheim, Germany) and Formic acid (LC-MS grade) was obtained from Thermo Fisher Scientific (Geel, Belgium). High purity water was provided by a Millipore Milli-Q Plus water purification system (Millipore, Milford, MA, USA). Reserpine and yohimbine were purchased from Sigma-Aldrich (Steinheim, Germany).

3.2. Sample Preparation

A sample (wet weight 277g) from *C. crambe* species (Class: *Demospongiae*; Order: *Poecilosclerida*) [43] collected from the Mediterranean sea, La Herradura, Spain, 36°43'23.149"N 3°43'35.91"W. It was then lyophilized and further extracted using the Accelerated Solvent Extraction (ASE) method with a mixture of dichloromethane and methanol (50:50, v/v). Crude extract was obtained after evaporation of the solvents under reduced pressure and reconstituted in MeOH in order to obtain a concentration 1 $mg.mL^{-1}$. Reconstituted sample was vortexed thoroughly for 10 min and further subjected to centrifugation at 12 000 rpm for 10 minutes at 4 °C. Supernatant was diluted and spiked with a mixture of reserpine and yohimbine, used as internal standards. Final concentration was of 0.1 µg/mL crude *C. crambe* extract and of 270 µg/mL for each internal standard.

3.3. UHPLC-HRMS/MS Analysis

Liquid chromatography analysis was performed on a Dionex Ultimate 3000 UHPLC system and detection was carried out on an Q-Exactive Orbitrap mass spectrometer equipped with an HESI source (Thermo Scientific, Germany). More specifically, separation was achieved on a Hypersil Gold UPLC C18 (2.1 × 100 mm, 1.9 µm) (Thermo Scientific, Germany) column heated at 40 °C. Solvent A was Milli-Q water with 0.1% (v/v) formic acid and solvent B was acetonitrile. Gradient elution started with 95% A, decreasing to 5% A in 24 min. These conditions were kept for 4 min before getting back

 doi:10.20944/preprints202410.2036.v1

to initial conditions for a 2 min re-equilibration. Auto-sampler tray temperature was maintained at 8 °C and injection volume was of 10 µL. HRMS data were acquired in both negative and positive modes, in the full scan m/z range of 110–1000, with a resolution of 70 000 FWHM in profile mode. Data dependent acquisition was simultaneously performed using a HCD-fragmentation with a normalized collision energy of 35% and a mass resolution of 35 000. Capillary temperature was set at 320 °C in both polarities, whereas source voltage was 3.6 kV in ESI+ and 3.2 kV in ESI-. Probe Heater Temperature was of 200 °C and the S-Lens RF Level was set at 55. Finally, nitrogen was used as the sheath gas (40 arbitrary units) and auxiliary gas (8 arbitrary units).

3.4. Molecular Networking/Computational Chemical Dereplication

UHPLC-HRMS/MS chromatograms were acquired using Xcalibur version 2.1 (Thermo Fisher Scientific).

For the MN-based dereplication, the obtained Thermo Fisher *.raw data were first converted to *.mzML format with the use of MSConvert, a tool from the ProteoWizard suite [44,45], in order to be processed in the MZmine2 environment [30]. During the conversion the data is centroided (transforming profile data into a format that represents peak intensities at discrete m/z values) by applying a peak picking filter. Data processing involved the following steps and parametrizations: a) mass detection based on a defined threshold of 100 for MS1 and 10 for MS2; b) chromatogram building was achieved implementing the GridMass - 2D module [46]. More specifically algorithm was set to generate a grid of equally spaced probes of 0.05 min and 0.05 Da in the entire time-range (1st dimension) and the m/z-range (2nd dimension) of the chromatogram. c) chromatogram deconvolution through local minimum search module, with chromatographic threshold 1%, minimum RT 0.05, minimum relative height 1%, minimum absolute height 200, and minimum ratio of peak to edge 1. d) Lastly, we applied feature filtering to keep only the ones with MS2-containing scans. Following MZmine processing, data were forwarded to GNPS2 online platform for molecular network construction and feature prediction through library matching [34]. FBMN workflow parameters in the GNPS2 environment included: general parameters—precursor ion tolerance 0.02 Da; fragment ion tolerance 0.02; networking parameters—minimum cosine score 0.7; minimum matched peaks 6; library search parameters—minimum cosine score: 0.7; minimum matched peaks 6; search for analogs enabled. Furthermore, the data were imported into SIRIUS [31] environment to predict the identity and chemical classes of the compounds based on the implemented CSI:FingerID [32] and the Support Vector Machine algorithms of CANOPUS and ClassyFire [33,47]. The NPC ClassyFire results were visualized in sunburst plots through *Plotly* Python [48], and were also integrated into the GNPS2-constructed FBMN [27,34] by utilizing the robust Cytoscape environment [35]. The Cytoscape platform has been utilized for the visualization of FBMN with a significant advantage being its ability to display any tabular data as attributes of the network, applicable to both nodes and edges. For the MN layout, *yfiles* [49] organic style was chosen since it reveals inherent symmetric and clustered structures within undirected graphs, while ensuring a balanced distribution of nodes and minimizing edge crossings. The acquired data and workflows are available under Findability, Accessibility, Interoperability, and Reuse (FAIR) principles. Specifically, MS/MS data are stored in the MassIVE repository (massive.ucsd.edu) with the identifier MSV000096117. The MN output and its parametrization can be found on GNPS repository (gnps2.org/status?task=2e9e2fe9d05743118d93cbb64f30acb0).

4. Conclusions

An in-depth qualitative investigation of Mediterranean sponge *Crambe crambe* was performed using UHPLC-HRMS/MS. The study of the CH_2Cl_2-MeOH extract, led to the annotation of 53 compounds belonging to crambescin and crambescidin families. A computational mass spectrometry approach, based on molecular networking studies and Natural Product Classes predictions (NPC), enabled the further study of the crambescins guanidine alkaloids family and the prioritization of compounds for future isolation. A detailed fragmentation pattern is proposed for the characterization of crambescin compounds, which enabled the identification of 22 analogues reported for the first

 doi:10.20944/preprints202410.2036.v1

time. Interestingly, MS/MS data analysis highlighted the presence of 4 new homologous series of crambescins, regarding the ratio of methylene units that compose the upper (n+2) and lower (m+2) alkyl side chains. These new homologues included crambescin B/C 452 (compounds **9**/**5**) and crambescin A 434 (compound **8**) with m=5, n=4, crambescin C 466 (compound 7) with m=5, n=5, crambescin B/C 466 (compounds **14**/**6**) and crambescin A 448 (compound 12) with m=6, n=4, as well as crambescin B/C 480 (compound **24**/**11**) with m=6, n=5. Moreover, the presence of 8 crambescidin compounds is reported among which 4 of them constitute new analogues. Due to its sensitivity, the high mass accuracy at both MS and MS/MS levels as well as the extensive structural information provided through HCD-fragmentation, the implemented UHPLC-HRMS/MS methodology demonstrated its fitness for purpose regarding the characterization of both major and minor guanidine alkaloids from *C. crambe* sponge.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, MS and MS/MS spectra and identified fragments for compounds **9, 14, 5, 6** and **7.**

Author Contributions: Conceptualization, Data Analysis & Interpretation, Writing - Original Draft, P.V.; Review & Editing, N.T.; Sample preparation, UHPLC-HRMS/MS Data Acquisition, E.B. & A.T.; Sample preparation, Writing - Review & Editing, G.L.G.; Computational Data analysis, Writing - Review & Editing N.M. & A.P.; Writing - Review & Editing, J.O.; Review & Editing, Coordination & Supervision, N.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported and conducted in the frame of the H2020 TASCMAR project, which was funded by the European Union under grant agreement number 634674 (www.tascmar.eu). This research was co-financed by Greece and the European Union (European Social Fund) through the Operational Programme " Human Resources Development, Education and Lifelong Learning" in the context of the Project "Strengthening Human Resources Research Potential via Doctorate Research" (Project No. MIS- 5000432), implemented by the State Scholarships Foundation (ΙΚΥ).

Data Availability Statement: The acquired data and workflows are stored as mentioned above in the MassIVE repository (massive.ucsd.edu) with the identifier MSV000096117, whereas the MN output can be found on GNPS repository (gnps2.org/status?task=2e9e2fe9d05743118d93cbb64f30acb0).

Acknowledgements: iMare Nature S.L. is acknowledged for its contribution and assistance in the sponge collection process.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Duran, S.; Pascual, M.; Estoup, A.; Turon, X. Strong Population Structure in the Marine Sponge Crambe Crambe (Poecilosclerida) as Revealed by Microsatellite Markers. *Molecular Ecology* **2004**, *13*, 511–522, doi:https://doi.org/10.1046/j.1365-294X.2004.2080.x.
2. Sfecci, E.; Lacour, T.; Amade, P.; Mehiri, M. Polycyclic Guanidine Alkaloids from Poecilosclerida Marine Sponges. *Marine Drugs* **2016**, *14*, 77, doi:10.3390/md14040077.
3. Silva, S.B.L.; Oberhänsli, F.; Tribalat, M.-A.; Genta-Jouve, G.; Teyssié, J.-L.; Dechraoui-Bottein, M.-Y.; Gallard, J.-F.; Evanno, L.; Poupon, E.; Thomas, O.P. Insights into the Biosynthesis of Cyclic Guanidine Alkaloids from Crambeidae Marine Sponges. *Angewandte Chemie* **2019**, *131*, 530–535, doi:10.1002/ange.201809539.
4. Ternon, E.; Perino, E.; Manconi, R.; Pronzato, R.; Thomas, O.P. How Environmental Factors Affect the Production of Guanidine Alkaloids by the Mediterranean Sponge Crambe Crambe. *Mar Drugs* **2017**, *15*, E181, doi:10.3390/md15060181.
5. Makarieva, T.; Shubina, L.; Kurilenko, V.; Isaeva, M.; Chernysheva, N.; Popov, R.; Bystritskaya, E.; Dmitrenok, P.; Stonik, V. Marine Bacterium Vibrio Sp. CB1-14 Produces Guanidine Alkaloid 6-Epi-Monanchorin, Previously Isolated from Marine Polychaete and Sponges. *Mar Drugs* **2019**, *17*, doi:10.3390/md17040213.
6. Öztürk, B.; de Jaeger, L.; Smidt, H.; Sipkema, D. Culture-Dependent and Independent Approaches for Identifying Novel Halogenases Encoded by Crambe Crambe (Marine Sponge) Microbiota. *Scientific Reports* **2013**, *3*, 2780, doi:10.1038/srep02780.
7. Becerro, M.A.; Lopez, N.I.; Turon, X.; Uriz, M.J. Antimicrobial Activity and Surface Bacterial Film in Marine Sponges. *Journal of Experimental Marine Biology and Ecology* **1994**, *179*, 195–205, doi:10.1016/0022-0981(94)90114-7.

8. Bovio, E.; Sfecci, E.; Poli, A.; Gnavi, G.; Prigione, V.; Lacour, T.; Mehiri, M.; Varese, G.C. The Culturable Mycobiota Associated with the Mediterranean Sponges Aplysina Cavernicola, Crambe Crambe and Phorbas Tenacior. *FEMS Microbiol Lett* **2019**, *366*, doi:10.1093/femsle/fnaa014.
9. Croué, J.; West, N.J.; Escande, M.-L.; Intertaglia, L.; Lebaron, P.; Suzuki, M.T. A Single Betaproteobacterium Dominates the Microbial Community of the Crambescidine-Containing Sponge Crambe Crambe. *Scientific Reports* **2013**, *3*, 2583, doi:10.1038/srep02583.
10. Sipkema, D.; Caralt, S. de; Morillo, J.A.; Al-Soud, W.A.; Sørensen, S.J.; Smidt, H.; Uriz, M.J. Similar Sponge-Associated Bacteria Can Be Acquired via Both Vertical and Horizontal Transmission. *Environmental Microbiology* **2015**, *17*, 3807–3821, doi:https://doi.org/10.1111/1462-2920.12827.
11. Bondu, S.; Genta-Jouve, G.; Leirós, M.; Vale, C.; Guigonis, J.-M.; Botana, L.M.; Thomas, O.P. Additional Bioactive Guanidine Alkaloids from the Mediterranean Sponge Crambe Crambe. *RSC Adv.* **2012**, *2*, 2828–2835, doi:10.1039/C2RA00045H.
12. Ternon, E.; Zarate, L.; Chenesseau, S.; Croué, J.; Dumollard, R.; Suzuki, M.T.; Thomas, O.P. Spherulization as a Process for the Exudation of Chemical Cues by the Encrusting Sponge C. Crambe. *Scientific Reports* **2016**, *6*, 29474, doi:10.1038/srep29474.
13. Jares-Erijman, E.; Ingrum, A.L.; Carney, J.; Rinehart, K.; Sakai, R. Polycyclic Guanidine-Containing Compounds from the Mediterranean Sponge Crambe Crambe: The Structure of 13,14,15-Isocrambescidin 800 and the Absolute Stereochemistry of the Pentacyclic Guanidine Moieties of the Crambescidins. **1993**, doi:10.1021/JO00070A012.
14. Jares-Erijman, E.A.; Sakai, R.; Rinehart, K.L. Crambescidins: New Antiviral and Cytotoxic Compounds from the Sponge Crambe Crambe Available online: https://pubs.acs.org/doi/pdf/10.1021/jo00019a049 (accessed on 28 April 2021).
15. Rinehart, K.L.; Jares-Erijman, E.A. Antiviral and Cytotoxic Compounds from the Sponge Crambe Crambe 1999.
16. Aoki, S.; Kong, D.; Matsui, K.; Kobayashi, M. Erythroid Differentiation in K562 Chronic Myelogenous Cells Induced by Crambescidin 800, a Pentacyclic Guanidine Alkaloid. *Anticancer Research* **2004**, *24*, 2325–2330.
17. El-Demerdash, A.; Metwaly, A.M.; Hassan, A.; Abd El-Aziz, T.M.; Elkaeed, E.B.; Eissa, I.H.; Arafa, R.K.; Stockand, J.D. Comprehensive Virtual Screening of the Antiviral Potentialities of Marine Polycyclic Guanidine Alkaloids against SARS-CoV-2 (COVID-19). *Biomolecules* **2021**, *11*, 460, doi:10.3390/biom11030460.
18. El-Demerdash, A.; Petek, S.; Debitus, C.; Al-Mourabit, A. Crambescidin Acid from the French Polynesian Monanchora n. Sp. Marine Sponge. *Chem Nat Compd* **2020**, *56*, 1180–1182, doi:10.1007/s10600-020-03262-1.
19. Kasmiati, K.; Yoshioka, Y.; Okamoto, T.; Ojika, M. New Crambescidin-Type Alkaloids from the Indonesian Marine Sponge Clathria Bulbotoxa. *Marine Drugs* **2018**, *16*, 84, doi:10.3390/md16030084.
20. Martín, V.; Vale, C.; Bondu, S.; Thomas, O.P.; Vieytes, M.R.; Botana, L.M. Differential Effects of Crambescins and Crambescidin 816 in Voltage-Gated Sodium, Potassium and Calcium Channels in Neurons. *Chem Res Toxicol* **2013**, *26*, 169–178, doi:10.1021/tx3004483.
21. Mendez, A.G.; Juncal, A.B.; Silva, S.B.L.; Thomas, O.P.; Martín Vázquez, V.; Alfonso, A.; Vieytes, M.R.; Vale, C.; Botana, L.M. The Marine Guanidine Alkaloid Crambescidin 816 Induces Calcium Influx and Cytotoxicity in Primary Cultures of Cortical Neurons through Glutamate Receptors. *ACS Chem Neurosci* **2017**, *8*, 1609–1617, doi:10.1021/acschemneuro.7b00096.
22. Roel, M.; Rubiolo, J.A.; Guerra-Varela, J.; Silva, S.B.L.; Thomas, O.P.; Cabezas-Sainz, P.; Sánchez, L.; López, R.; Botana, L.M. Marine Guanidine Alkaloids Crambescidins Inhibit Tumor Growth and Activate Intrinsic Apoptotic Signaling Inducing Tumor Regression in a Colorectal Carcinoma Zebrafish Xenograft Model. *Oncotarget* **2016**, *7*, 83071–83087, doi:10.18632/oncotarget.13068.
23. Shrestha, S.; Sorolla, A.; Fromont, J.; Blancafort, P.; Flematti, G.R. Crambescidin 800, Isolated from the Marine Sponge Monanchora Viridis, Induces Cell Cycle Arrest and Apoptosis in Triple-Negative Breast Cancer Cells. *Marine Drugs* **2018**, *16*, 53, doi:10.3390/md16020053.
24. Lazaro, J.E.H.; Nitcheu, J.; Mahmoudi, N.; Ibana, J.A.; Mangalindan, G.C.; Black, G.P.; Howard-Jones, A.G.; Moore, C.G.; Thomas, D.A.; Mazier, D.; et al. Antimalarial Activity of Crambescidin 800 and Synthetic Analogues against Liver and Blood Stage of Plasmodium Sp. *The Journal of Antibiotics* **2006**, *59*, 583–590, doi:10.1038/ja.2006.78.
25. Yang, J.Y.; Sanchez, L.M.; Rath, C.M.; Liu, X.; Boudreau, P.D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; et al. Molecular Networking as a Dereplication Strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699, doi:10.1021/np400413s.
26. Aron, A.T.; Gentry, E.C.; McPhail, K.L.; Nothias, L.-F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J.M.; Sikora, N.; Vargas, F.; et al. Reproducible Molecular Networking of Untargeted Mass Spectrometry Data Using GNPS. *Nat Protoc* **2020**, *15*, 1954–1991, doi:10.1038/s41596-020-0317-5.
27. Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; et al. Feature-Based Molecular Networking in the GNPS Analysis Environment. *Nat Methods* **2020**, *17*, 905–908, doi:10.1038/s41592-020-0933-6.

28. Freire, V.F.; Gubiani, J.R.; Spencer, T.M.; Hajdu, E.; Ferreira, A.G.; Ferreira, D.A.S.; De Castro Levatti, E.V.; Burdette, J.E.; Camargo, C.H.; Tempone, A.G.; et al. Feature-Based Molecular Networking Discovery of Bromopyrrole Alkaloids from the Marine Sponge *Agelas Dispar*. *J. Nat. Prod.* **2022**, *85*, 1340–1350, doi:10.1021/acs.jnatprod.2c00094.
29. Khushi, S.; Salim, A.A.; Capon, R.J. Case Studies in Molecular Network-Guided Marine Biodiscovery. *Marine Drugs* **2023**, *21*, 413, doi:10.3390/md21070413.
30. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010**, *11*, 395, doi:10.1186/1471-2105-11-395.
31. Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A.A.; Melnik, A.V.; Meusel, M.; Dorrestein, P.C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat Methods* **2019**, *16*, 299–302, doi:10.1038/s41592-019-0344-8.
32. Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching Molecular Structure Databases with Tandem Mass Spectra Using CSI:FingerID. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 12580–12585, doi:10.1073/pnas.1509788112.
33. Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M.A.; Petras, D.; Gerwick, W.H.; Rousu, J.; Dorrestein, P.C.; et al. Systematic Classification of Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nat Biotechnol* **2021**, *39*, 462–471, doi:10.1038/s41587-020-0740-8.
34. Wang, M.; Carver, J.J.; Phelan, V.V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kapono, C.A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **2016**, *34*, 828–837, doi:10.1038/nbt.3597.
35. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504, doi:10.1101/gr.1239303.
36. Jamison, M.T.; Molinski, T.F. Antipodal Crambescin A2 Homologues from the Marine Sponge *Pseudaxinella Reticulata* . Antifungal Structure–Activity Relationships. *J. Nat. Prod.* **2015**, *78*, 557–561, doi:10.1021/np501052a.
37. Vlachou, P.; Le Goff, G.; Alonso, C.; Álvarez, P.A.; Gallard, J.-F.; Fokialakis, N.; Ouazzani, J. Innovative Approach to Sustainable Marine Invertebrate Chemistry and a Scale-Up Technology for Open Marine Ecosystems. *Marine Drugs* **2018**, *16*, 152, doi:10.3390/md16050152.
38. Coffey, D.S.; McDonald, A.I.; Overman, L.E.; Rabinowitz, M.H.; Renhowe, P.A. A Practical Entry to the Crambescidin Family of Guanidine Alkaloids. Enantioselective Total Syntheses of Ptilomycalin A, Crambescidin 657 and Its Methyl Ester (Neofolitispates 2), and Crambescidin 800. *J. Am. Chem. Soc.* **2000**, *122*, 4893–4903, doi:10.1021/ja000234i.
39. Inkscape Project 2021.
40. de Laeter, J.R.; Böhlke, J.K.; De Bièvre, P.; Hidaka, H.; Peiser, H.S.; Rosman, K.J.R.; Taylor, P.D.P. Atomic Weights of the Elements. Review 2000 (IUPAC Technical Report). *Pure Appl. Chem.* **2003**, *75*, 683–800, doi:10.1351/pac200375060683.
41. Genta-Jouve, G.; Croué, J.; Weinberg, L.; Cocandeau, V.; Holderith, S.; Bontemps, N.; Suzuki, M.; Thomas, O.P. Two-Dimensional Ultra High Pressure Liquid Chromatography Quadrupole/Time-of-Flight Mass Spectrometry for Semi-Targeted Natural Compounds Identification. *Phytochemistry Letters* **2014**, *10*, 318–323, doi:10.1016/j.phytol.2014.10.029.
42. Smith, R.M. *Understanding Mass Spectra: A Basic Approach*; 2nd ed.; Wiley Interscience: Hoboken, N.J, 2004; ISBN 978-0-471-42949-4.
43. WoRMS Editorial Board World Register of Marine Species. Available from Https://Www.Marinespecies.Org at VLIZ. Accessed Yyyy-Mm-Dd. 2024.
44. Chambers, M.C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D.L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; et al. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat Biotechnol* **2012**, *30*, 918–920, doi:10.1038/nbt.2377.
45. Adusumilli, R.; Mallick, P. Data Conversion with ProteoWizard msConvert. In *Proteomics*; Comai, L., Katz, J.E., Mallick, P., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2017; Vol. 1550, pp. 339–368 ISBN 978-1-4939-6745-2.
46. Treviño, V.; Yañez-Garza, I.-L.; Rodriguez-López, C.E.; Urrea-López, R.; Garza-Rodriguez, M.-L.; Barrera-Saldaña, H.-A.; Tamez-Peña, J.G.; Winkler, R.; Díaz de-la-Garza, R.-I. GridMass: A Fast Two-Dimensional Feature Detection Method for LC/MS: GridMass: 2D Feature Detection for LC/MS. *J. Mass Spectrom.* **2015**, *50*, 165–174, doi:10.1002/jms.3512.
47. Kim, H.W.; Wang, M.; Leber, C.A.; Nothias, L.-F.; Reher, R.; Kang, K.B.; Van Der Hooft, J.J.J.; Dorrestein, P.C.; Gerwick, W.H.; Cottrell, G.W. NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. *J. Nat. Prod.* **2021**, *84*, 2795–2807, doi:10.1021/acs.jnatprod.1c00399.
48. Plotly Technologies Inc. Collaborative Data Science 2015.
49. yWorks yFiles Layouts in Cytoscape 3.6.0. 2024.

 doi:10.20944/preprints202410.2036.v1