scientific reports

Check for updates

OPEN

Exploring the chemodiversity of antimicrobial minalemines from *Didemnum granulatum* by neutral loss graph

Vítor F. Freire[1]✉, Jason R. Evans[1], Lucero Martínez-Fructuoso[1], Rohitesh Kumar[2], Rhone K. Akee[2], Svetlana Hogan[2], Christopher C. Thornburg[2], Brian D. Peyser[1], Susan Ensel[2,3], Dongdong Wang[4], Tanja Grkovic[1,4] & Barry R. O'Keefe[1,4]✉

In this study, we report on the identification of the active antimicrobial principles from the tunicate *Didemnum granulatum*. Two new natural products, minalemines G (1) and H (2), were isolated and their structures were elucidated. Both compounds showed potent antibacterial activity, with compound 1 showing sub-microgram inhibition against *Staphylococcus aureus*. Moreover, we also report a new mass spectrometry-based approach, named neutral loss graph, developed to explore minor metabolites from complex mixtures. This approach was applied here to reveal twelve additional minalemine analogues in the organic extract of *D. granulatum*.

Marine organisms are a rich source of bioactive secondary metabolites that have inspired the development of successful drugs based on their natural scaffolds, with a total of 15 compounds approved, and 31 others in various stages of clinical evaluation[1–3]. Three of the approved drugs derived from marine sources were initially identified from tunicates, highlighting the importance of the continue exploration of this subphylum for drug discovery[4]. *Didemnum* is a genus of marine tunicates widely distributed across tropical and subtropical coastal regions of Asia, Oceania, Africa, Europe, and the Americas. Their broad geographical distribution is mirrored by a large chemodiversity of secondary metabolites with over 200 natural products reported, including alkaloids[5], peptides[6], polyketides[7], lipids[8], steroids[9] and nucleosides[9,10]. Many of these compounds possess significant biological activity with potential pharmacological applications, such as the antiproliferative lamellarins[11], the cycle checkpoint inhibitory granulatimides[5], the HIV protease inhibitory didemnaketals[12], and the antimalarial and antitrypanosomal lepadins[13]. The enediynes namenamicin and shishijimicins[14] are potent cytotoxic compounds that closely resemble the microbial natural product calicheamicin, a compound used as an antibody–drug conjugate for treating acute myeloid (Mylotarg®) and lymphocytic leukemias (Besponsa®)[10,15]. Overall, although the chemical diversity of tunicates has been extensively studied, there remains significant potential for new discoveries.

Recently, a large antimicrobial high-throughput screen (HTS) was reported by the National Cancer Institute (NCI) and the National Institute of Allergy and Infectious Diseases (NIAID)[16]. The project involved screening 326,000 natural product fractions against four microbial targets: *Staphylococcus aureus*, *Escherichia coli* (including two strains—wild type and *tol*C efflux mutant), and *Candida albicans* and identified over 3,000 fractions with potent antimicrobial activity. From this, 75 active fractions identified during the discovery campaign were selected for further HPLC fractionation and their chemotypes identified using 1 mg of material according to previously published automated methods[17]. HPLC-subfractions from *D. granulatum* showed activity against all four microbial strains (Fig. S1), and using NMR fingerprints, the major chemical components of those fractions were assigned as lipopeptides, although no complete structure was proposed[16]. Herein, we report additional studies on the antimicrobial fractions from *D. granulatum*, describing the isolation and biological evaluation of minalemines G (**1**) and H (**2**), together with an application of neutral loss graph as a tool to explore minor metabolites in the adjacent chemical space.

[1]Natural Products Branch, Developmental Therapeutic Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Frederick, MD 21702-1201, USA. [2]Natural Products Support Group, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Frederick, MD 21702-1201, USA. [3]Department of Chemistry and Physics, Hood College, Frederick, MD 21701-8599, USA. [4]Molecular Targets Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702-1201, USA. ✉email: vitor.freire@nih.gov; okeefeba@mail.nih.gov

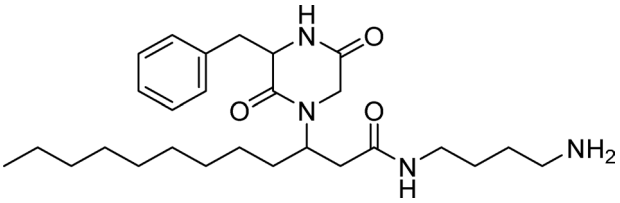

Fig. 1. Structure of rodriguesine A. The compound was identified from *D. granulatum* based on an LC–MS dereplication workflow.

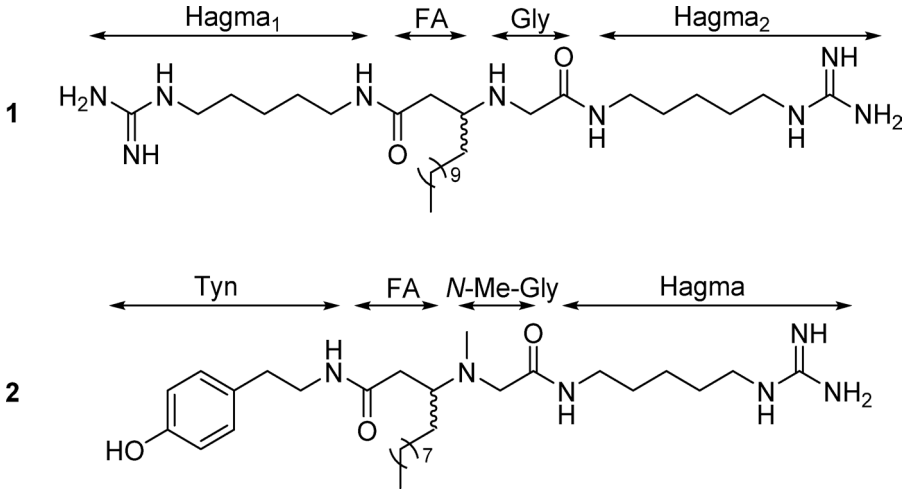

Fig. 2. Structures of minalemines G (**1**) and H (**2**). Compound **1** is composed of two homoagmatine (Hagma) subunits, along with fatty acid (FA) and glycine (Gly), while **2** is composed of tyramine (Tyn), FA, *N*-Me-Gly and Hagma.

Results

Rapid identification of compounds in bioactive subfractions

Despite the potential utility of compounds derived from tunicates, identifying new chemical entities with pharmacological potential in these organisms can be a laborious process, often complicated by the rediscovery of known metabolites[18]. To address this challenge, LC–MS and taxonomic data have been employed in modern dereplication workflows, utilizing analytical datasets and comprehensive databases for metabolite profiling[19]. As the starting point of this study, we performed a mass spectrometry-based annotation of the constituents in the active subfractions from *D. granulatum*. Initially, the accurate mass of *m/z* features present in the active subfractions of *D. granulatum* were searched against a dataset of 182 compounds from the genus *Didemnum* assembled from the *Dictionary of Natural Products* database[20]. Then, a comparison between the MS^2 spectra of annotated features and literature reports was used to support dereplication. This approach led us to identify the *m/z* feature at 459 (R_T = 4.71 min, subfraction 8) as rodriguesine A (Fig. 1)[21], supported by comparison with the previously reported fragmentation profile (Fig. S2, Table S1)[22].

The strategy successfully used in the annotation of rodriguesine A, which was found exclusively in one of the active subfractions, suggesting that other constituents were responsible for the observed activity. LC–MS traces of the active subfractions revealed that the samples were mixtures of multiple components. Subsequent isolation efforts with additional material from the organic extract of *D. granulatum* led to the purification of two new compounds, minalemines G (**1**) and H (**2**) (Fig. 2).

Structure elucidation of minalemines G (1) and H (2)

Minalemine G (**1**) showed a $[M+H]^+$ at *m/z* 554.4873 by HRESIMS, corresponding to the molecular formula $C_{28}H_{60}N_9O_2^+$. The ^1H NMR of **1** (Table 1) showed one sp^3 methine at δ_H 3.52, multiple methylenes ranging between δ_H 3.88–1.29, and one terminal methyl at δ_H 0.90. A spin system formed by five methylenes ($Hagma_1$-H1–5) was revealed by ^1H–^1H COSY correlations (Fig. 3). Analysis of ^1H–^{15}N and ^1H–^{13}C HMBC showed correlations from $Hagma_1$-H5 (δ_H 3.17) to a terminal guanidine (δ_N 82.9 and δ_C 158.6), forming the first homoagmatine subunit ($Hagma_1$). The methylene at $Hagma_1$-H1(δ_H 3.22) showed correlations to $Hagma_1$-N1 (δ_N 122.7) and to the fatty acid portion (FA) through the carbonyl at FA-C1 (δ_C 172.4). The second homoagmatine subunit ($Hagma_2$) exhibited a similar spin system containing five methylenes with overlapped ^1H, ^{13}C and ^{15}N chemical shifts to $Hagma_1$, differing at CH_2-1 (δ_H 3.27, δ_C 40.4) and N1 (δ_N 116.0). $Hagma_2$ was connected to the C-terminal of a glycine (Gly) residue by a ^1H–^{13}C HMBC correlation from $Hagma_2$-H1 (δ_H 3.27) to Gly-C2

1[a]				2[b]	
Position	**δ_C[c]/δ_N[d], type**	**δ_H[e], mult (*J* in Hz)**	**Position**	**δ_C[c]/δ_N[d], type**	**δ_H[e], mult (*J* in Hz)**
Hagma$_1$			*Tyn*		
1	39.2, CH_2	3.22, t (7.2)	1	42.3, CH_2	3.37, td (2.9, 7.1, 7.3)
2	29.9, CH_2	1.56, m	2	35.6, CH_2	2.71, t (7.1)
3	24.9, CH_2	1.40, m	3	131.0, C	
4	29.4, CH_2	1.61, m	4/8	130.7, CH	7.03, dd (2.2, 8.5)
5	42.3, CH_2	3.17, t (7.1)	5/7	116.3, CH	6.70, dd (2.2, 8.5)
1-NH	122.7, NH	–	6	157.3, C	–
5-NH	82.9, NH	–	*FA*		
C=N	158.6, C	–	1	175.0, C	–
FA			2	37.4, CH_2	2.24, dd (6.0, 14.4) 2.31, dd (8.8, 14.4)
1	172.4, C	–	3	62.4, CH	2.94, m
2	34.9, CH_2	2.57, dd (7.2, 16.4) 2.70, dd (4.6, 16.4)	4	30.8, CH_2	1.27, m 1.55, m
3	57.5, CH	3.52, m	5–11	23.7–33.1, CH_2	1.29–1.30, m
4	24.3, CH_2	1.61, m	12	14.4, CH_3	0.89, t (7.2)
5–13	23.8 – 33.1, CH_2	1.29–1.40, m	*N-Me-Gly*		
14	14.5, CH_3	0.90, t (7.0)	1	58.7, CH_2	3.06, s
Gly			2	174.2, C	-
1	46.5, CH_2	3.82, d (15.5) 3.88, d (15.5)	N-Me	37.4, CH_3	2.21, s
2	166.5, C	–	*Hagma*		
NH-1	45.8, NH_2	–	1	39.7, CH_2	3.23, td (1.8, 7.5)
Hagma$_2$			2	29.4, CH_2	1.59, m
1	40.4, CH_2	3.27, t (7.1)	3	24.9, CH_2	1.39, m
2	29.9, CH_2	1.56, m	4	30.1, CH_2	1.56, m
3	25.0, CH_2	1.40, m	5	42.4, CH_2	3.16, t (1.8, 7.2)
4	29.4, CH_2	1.61, m	C=N	158.6, C	–
5	42.3, CH_2	3.17, t (7.1)			
1-NH	116.0, NH	–			
5-NH	83.0, NH	–			
C=N	158.6, C	–			

Table 1. NMR data for minalemines G (**1**) and H (**2**) in MeOH-d_4. [a]TFA salt, [b]Free base, [c]151 MHz, [d]60 MHz, [e]600 MHz.

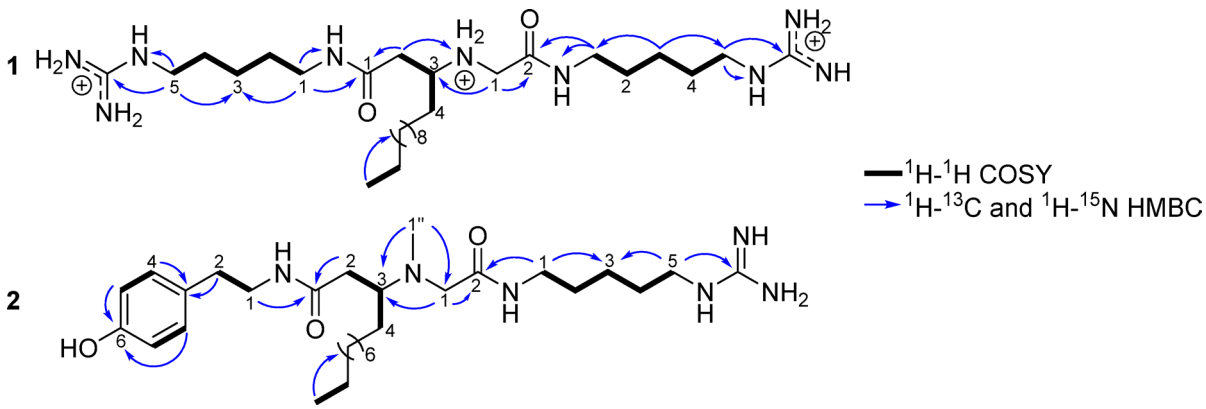

Fig. 3. Key ^{1}H–^{1}H COSY, ^{1}H–^{13}C and ^{1}H–^{15}N HMBC correlations of minalemine G (**1**) of the free-base form of minalemine H (**2**).

(δ_C 166.5). The final planar structure was formed by linking the Gly residue to FA subunit through correlations from Gly-H1 (δ_H 3.82/3.88) to FA-C3 (δ_C 57.5), confirmed by the correlation from FA-H2 (δ_H 2.57/2.70) to Gly-N1 (δ_N 45.8). With the planar structure established, a competing enantioselective conversion method (CEC)[23] was applied to determine the absolute configuration of C-1 in **1**, but the kinetics of the reactions using

R- and *S*-homobenzotetramisole (HBTM) showed no differentiation, suggesting an enantiomeric mixture (Fig. S3). A weak specific rotation $[\alpha]_D^{24}$ of −2.97 suggested **1** to be a scalemic mixture, which was confirmed by chiral chromatography analysis which showed a 6:4 ratio of the two enantiomers (Fig. S4). Compound **1** was determined to be a new natural product and was named minalemine G.

Minalemine H (**2**) showed a $[M+H]^+$ at *m/z* 533.4172 by HRESIMS, corresponding to the molecular formula $C_{29}H_{53}N_6O_3^+$. Analysis of 1D and 2D NMR of the TFA salt form of **2** showed broad signals for the methine and methylene protons of FA and *N*-Me-Gly, resulting in a lack of key correlations for these subunits. Consequently, compound **2** was converted to a free base form by eluting its methanolic solution through an amino SPE cartridge. Figure 4 shows the comparison between the NMR spectra of the TFA salt and free base forms of **2**, highlighting the FA-(*N*-Me-Gly) region. The 1D and 2D NMR data of the free base form of **2** showed clearer signals, multiplicities, and correlations, supporting the structure elucidation outlined below.

NMR data revealed structural resemblance between compounds **1** and **2** by sharing one Hagma and FA subunits. Additionally, the ^{1}H NMR spectrum of **2** showed two singlets at δ_H 3.06 and δ_H 2.21, with reciprocal ^{1}H–^{13}C HMBC correlations, forming an *N*-Me-Gly moiety. An 1,4-disubstituted aromatic containing two doublet of doublets at δ_H 6.70 (Tyn-H5/H7) and δ_H 7.03 (Tyn-H4/H8) with coupling constants of 2.2 and 8.5 Hz, was connected to two methylenes at Tyn-C2 (δ_C 35.6) and Tyn-C1 (δ_C 42.3) through HMBC correlation, establishing a tyramine (Tyn) subunit. The Tyn and FA substructures were connected through correlations between Tyn-H1 (δ_H 3.37) and FA-H2 (δ_H 2.24/2.31) to the carbonyl at FA-C1 (δ_C 175.0). The FA moiety showed HMBC correlations from FA-H3 (δ_H 2.94) to the methyl and methylene of the *N*-Me-Gly subunit. Additionally, *N*-Me-Gly was attached to Hagma, as shown by correlations from *N*-Me-Gly-H1 (δ_H 3.06) and Hagma-H1 (δ_H 3.23) to the carbonyl group at *N*-Me-Gly-C2 (δ_C 174.2). As observed for **1**, compound **2** also showed a low specific rotation value, $[\alpha]_D^{24}$ of −4.44. Chiral chromatographic analysis indicated a 7:3 enantiomeric ratio, revealing it as a scalemic mixture. Compound **2** was identified as a new natural product, and named minalemine H.

Exploring minor metabolites using neutral loss graph method

Due to challenges during isolation and limited availability of organic extract material, only compounds **1** and **2** were isolated in sufficient quantities for complete structural characterization by NMR. However, LC–MS analysis of the fractions revealed several additional peaks with similar ionization profiles, suggesting the presence of analogues of **1** and **2**. To explore the chemodiversity of these minor metabolites, we developed the neutral loss graph (NLG), a mass spectrometry-based approach that uses a combinatorial neutral loss analysis to explore structural similarities of closely related compounds. The NLG uses LC-MS^2 experimental data to generate a list of all existing neutral losses within each *m/z* feature. Then, key neutral losses from the target compounds serve as probes for pair-wise similarity analysis, with the output displayed as a network graph (Fig. 5).

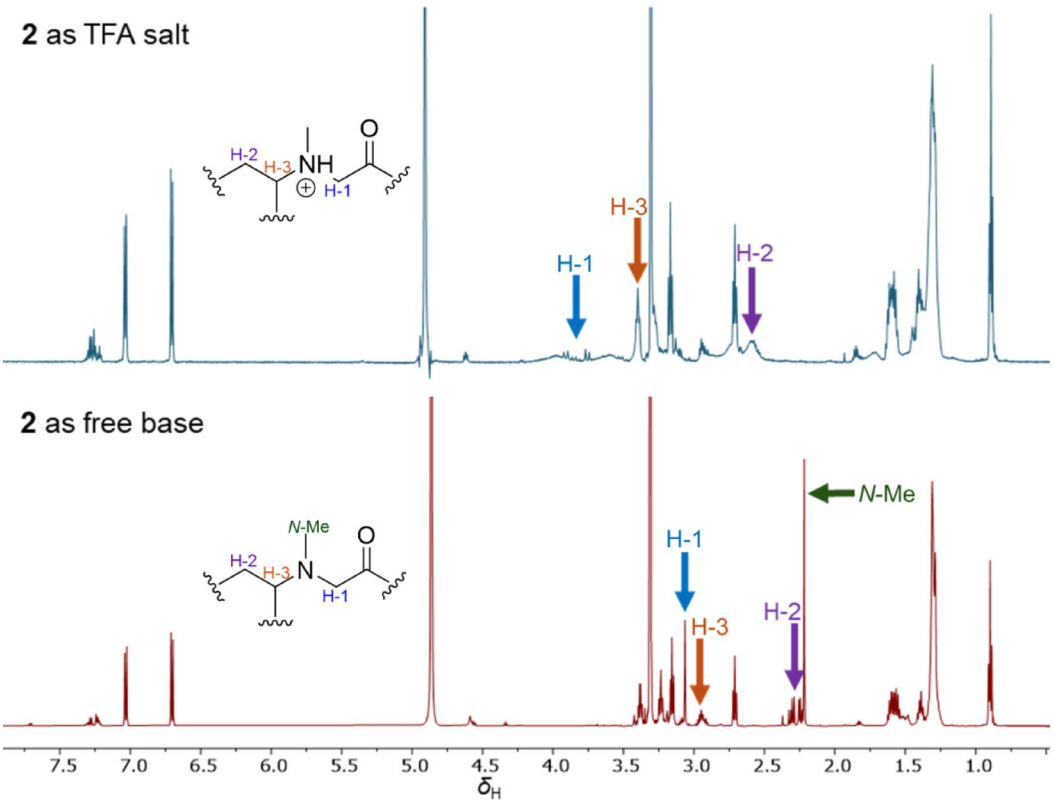

Fig. 4. Comparison of signals on ^{1}H NMR spectra between TFA salt and free base forms of compound **2** in MeOH-d_4.

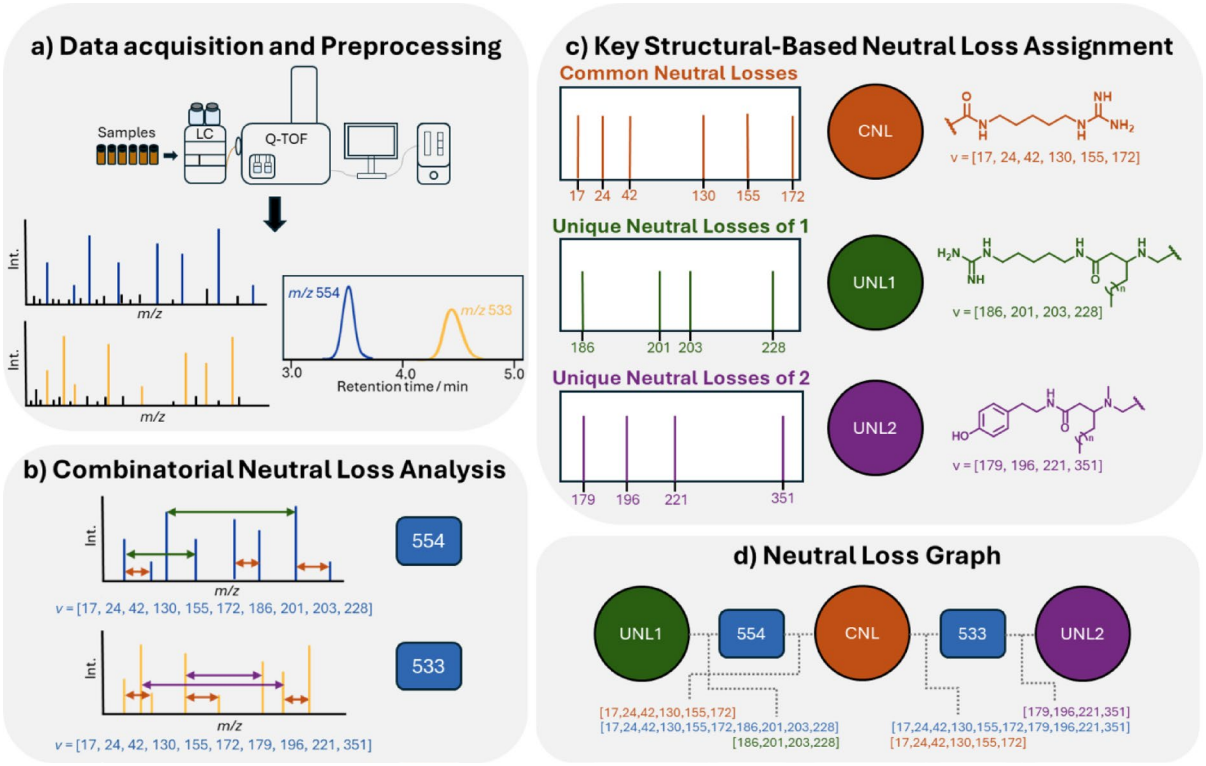

Fig. 5. The concept of neutral loss graph (NLG). **a** LC–MS/MS data acquisition and preprocessing combining feature detection and alignment. **b** Combinatorial analysis of experimental neutral losses between all ions in a *m/z* feature. **c** Determination of key structural neutral losses for each core. **d** Calculation of similarity vectors and network construction results in an NLG.

To explore the chemodiversity of *D. granulatum* using NLG method, fractions 1–7, generated from the SPE fractionation, were analyzed by LC-MS^2 in a data-dependent acquisition mode (DDA). The raw data (.d) was converted to an open-source format (mzML) and preprocessed on MZmine 3.4.27, where the aligned feature list was exported as a mgf file. This file was then used for a combinatorial neutral loss analysis, calculating all-to-all mass differences, from the precursor ion to its fragments, and between fragments within the same *m/z* feature, generating a new mgf file. Parallel to that, three sets of target neutral losses were established: common neutral losses (CNL), unique neutral losses of **1** (UNL1), and **2** (UNL2). The CNL contained six neutral losses (Table S2) related to the Hagma unit found in compounds **1** and **2**, consisting of small mass losses such as NH_3 and CHN, and medium mass losses, such as $C_7H_{16}N_4O$. The UNL1 encompassed four larger neutral losses related to Hagma-FA-Gly, substructures of **1**, while UNL2 was comprised of four large neutral losses related to Tyn-FA-*N*-Me-Gly substructures of **2**. The neutral loss graph was then built by creating a pair-wise similarity matrix, comparing the combinatorial neutral loss of each *m/z* feature and the sets of CNL, UNL1 and UNL2. Figure 6 shows the cluster containing minalemines present in the NLG displayed on Cytoscape, where all features not directly connected to the CNL, and UNL1, or UNL2 were excluded. This cluster showed a total of fourteen *m/z* features with edges connecting them to the CNL and UNL1 or UNL2 (Fig. 6).

Minalemine G (**1**, *m/z* 554) clustered with six other *m/z* features (Fig. 6, Table S3) showing edges to the groups of neutral losses CNL and UNL1, indicating related substructures among them. Analysis of MS^1 spectra for features *m/z* 526, 540, and 568 showed differences of -28, -14, and + 14 Da, respectively, compared to the monoprotonated molecule $[M+H]^+$ of **1**. Characteristic fragments such as $[M+H–C_7H_{16}N_4O]^+$, $[M+H—C_8H_{18}N_4O]^+$, $[M+H–C_8H_{19}N_5O]^+$ and $[M+H–C_9H_{20}N_6O]^+$, suggested these features to be part of a homologous series of **1**, differing only in FA chain length. The MS^2 data of features *m/z* 606, 620, and 634 showed an additional neutral loss of SO_3H (79.956 Da), comprising fragment ions at *m/z* 526, 540 and 554 respectively, suggesting their structures to be sulfamic acid-containing compounds, analogous to previously reported minalemines D-F containing sulfamic acid group at Gly-N1[24].

Minalemine H (**2**, *m/z* 533) shared connection with CNL and UNL2, as well as six other *m/z* features (Fig. 6, Table S3). Analysis of MS^1 spectrum of features *m/z* 477, 505, 519, 547 and 561 showed differences of -56, -28, -14, + 14, and + 28 Da compared to the monoprotonated molecule $[M+H]^+$ of **2**. MS^2 spectra of these features displayed the same fragmentation pattern as **2**, with characteristic fragment ions such as $[M+H–C_{10}H_{13}NO_2]^+$, $[M+H–C_{10}H_{16}N_2O_2]^+$, $[M+H–C_{11}H_5N_3O_2]^+$ and $[M+H—C_{17}H_{29}N_5O_3]^+$, indicating that these *m/z* features form a homologue series of **2**. Inspection of MS^1 spectrum of feature *m/z* 531 implied an additional unsaturation compared to **2**. MS^2 analysis revealed a -2 Da difference for most fragment ions, except for *m/z* 216 [M + H–

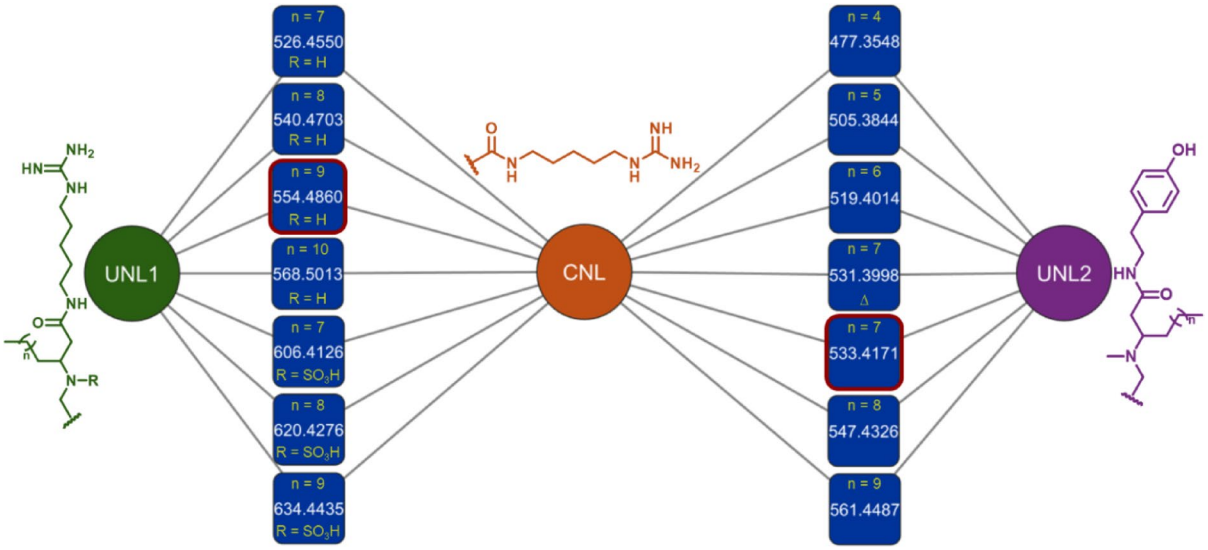

Fig. 6. Minalemine's cluster on the neutral loss graph (NLG). The common neutral loss (CNL) represents the shared subunit of Hagma. The unique neutral loss 1 (UNL1) corresponds to the unique subunits of Hagma-FA-Gly from minalemine G (**1**, node *m/z* 554.4860). While the unique neutral loss 2 (UNL2) is composed by Tyn-FA-*N*-Me-Gly, specific subunits of minalemine H (**2**, node *m/z* 533.4171).

Microbial strains	1	2	Levofloxacin	Amphoteracin B
E. coli (BW25113)	20	> 20	0.03	> 8
E. coli, *tol*C efflux deficient (JW5503-1)	1.25	5	0.08	> 8
P. aeruginosa, efflux mutant (PAM 1626)	5	> 20	0.015	> 8
S. aureus (ATCC 29213)	0.62	2.5	0.12	> 8
E. faecalis (ATCC 29212)	20	20	0.5	> 8
E. faecium, VRE (ATCC 700,221)	10	20	> 4	> 8
C. albicans (ATCC 90,028)	10	> 10	> 2	0.5
A. fumigatus (ATCC MYA-3626)	> 10	> 10	> 2	1

Table 2. Minimum inhibitory concentration (μg/mL) of minalemines G (**1**) and H (**2**) against selected bacterial and fungal strains.

$C_{20}H_{29}NO_2]^+$ and *m/z* 199 $[M + H–C_{20}H_{32}N_2O_2]^+$, which are fragments without the FA chain, strongly suggesting the position of the unsaturation to be at that subunit.

Antimicrobial activity

The minimum inhibitory concentration (MIC) of minalemines G (**1**) and H (**2**) were established for various bacterial and fungal strains (Table 2, Fig. S5), which included *E. coli* wild-type (BW25113) and *E. coli tol*C efflux-deficient (JW5503-1), *P. aeruginosa* efflux-mutant (PAM 1626), *S. aureus* wild-type (ATCC 29213), *E. faecalis* wild-type (ATCC29212), vancomycin-resistant *E. faecium* (VRE, ATCC 700221), *C. albicans* (ATCC90028), and *A. fumigatus* (ATCC MYA-3626). Levofloxacin served as positive control for bacterial strains, while amphotericin B was used for the fungal strains. Minalemines G (**1**) and H (**2**) did not show antifungal activity against *C. albicans* or *A. fumigatus*, but both compounds showed activity against *S. aureus*, with MIC values of 0.62 μg/mL for **1** and 2.50 μg/mL for **2**. Furthermore, the Gram-negative *E. coli tol*C efflux-deficient strain was more susceptible to both compounds than the *E. coli* wild type, suggesting these compounds could be substrates of the *tol*C efflux pumps. Minalemine G (**1**) also exhibited a MIC value of 5.00 μg/mL against the efflux mutant strain of *P. aeruginosa*.

Discussion

Antimicrobial activity guided-isolation of the active principles from the organic extract of the tunicate *D. granulatum* resulted in the purification of two new guanidine alkaloids, named minalemines G (**1**) and H (**2**). Compounds **1** and **2** are new representatives of the minalemine class, previously composed of six compounds, minalemines A-F, isolated from a *D. rodriguesi* collected in New Caledonia[24]. The minalemines G (**1**) and H (**2**) described here differ from those previously reported by the absence of a leucine residue, along with the substitution of the agmatine subunit with homoagmatine in compound **1**, and with tyramine in **2**.

The chemical structure of a compound is directly reflected in fragment ions displayed in its MS^2 spectrum. Even analogues with subtle structural modifications can exhibit fragments with distinctive *m/z* values[25]. Visualization tools, such as feature-based networks, can explore similarities in MS/MS fragmentation between known and unknown compounds, expanding the annotation and assisting in the dereplication efforts, leading to new discoveries[26,27]. These innovative approaches have facilitated the identification of new compounds at a faster pace, ushering natural products research into a new era[19,26,28]. However, analogues tend to show similar fragmentation pathways, which are disclosed by their neutral losses. Neutral loss has been used for spectral similarity analysis and metabolite annotation[29], and to enhance molecular similarity analysis in METLIN[25]. Additionally, the use of neutral loss has been shown to outperform the cosine similarity for numerous small molecules[30]. Even though traditional molecular networking approaches combine fragment ions and neutral loss matches using a modified cosine score, it still relies on overlapped fragment ions in order to cluster two compounds together[30]. The minalemines presented no fragment overlap in the MS/MS (Fig. S6), failing to cluster both compounds together. This observation led us to develop the neutral loss graph method, a modified approach that relies on neutral loss analysis to propagate the annotation of target compounds. When applied to *D. granulatum* at the fraction level, NLG resulted in the annotation of twelve additional analogues of compounds **1** and **2**. Six analogues of **2** form a homologous series, differing only in the length of the FA moiety. Moreover, *m/z* features annotated as analogues of **1** include a homologue series of minalemine G, and their respective sulfamic acid derivatives. The NLG demonstrates the applicability of neutral loss analysis as an alternative method for exploring a chemical space, leading to the annotation of minor metabolites and providing insights into their structures. Our results demonstrate that integrating LC–MS/MS data into the dereplication workflow is a reliable strategy to reveal minor compounds of interest, opening possibilities for further research to better explore chemical, biological, and pharmacological space.

Minalemines G (**1**) and H (**2**) demonstrated antibacterial activity, with compound **1** showing notable activity against the Gram-positive *S. aureus*. The more potent activity of **1** suggests that the addition of charged sites in its structure enhances its potency. The amphiphilic nature of the minalemines could lead to cell membrane disruption. However, their lower MIC against specific strains makes it reasonable to suggest that these compounds may possess a different mechanism of action[31]. Maccari and collaborators synthesized linear and cyclic di-guanidine compounds that exhibit potent antibacterial activity. Their structure–activity relationship (SAR) findings showed that modifications to the linker, altering the distance between the guanidine groups, along with derivatization by different substituents, significantly change the activity[31]. One of their synthetic compounds comprising two guanidine subunits linked by a chain of two series of eight methylenes with a secondary amine in the center, and a 1-methylenecyclopropyl group on one guanidine moiety (Fig. S7), showed MIC value below 0.125 µg/mL against a strain of Gram-positive bacteria *Streptococcus pyogenes*[31]. These results indicate that SAR studies could enhance the minalemine class as antibacterial leads for further development.

Methods

General procedures

Optical rotations were measured on a Rudolph Research Analytical AUTOPOL IV automatic polarimeter with a 0.25 dm path length cell in methanol at 25 °C. UV spectra were recorded as methanol solutions on a Varian Cary 50-Bio UV/Vis spectrophotometer. FTIR spectra were recorded as thin films on a Bruker Alpha II spectrometer. NMR spectra were recorded at 25 °C on either a Bruker Avance III HD spectrometer, equipped with a 5 mm TCI Cryo-Probe Prodigy or a Bruker Avance III spectrometer equipped with a 3 mm TCI cryogenic probe, both operating at a frequency of 600 MHz for the ^{1}H nucleus, 151 MHz for the ^{13}C nucleus, and 60 MHz for the ^{15}N nucleus. For the 3 mm TCI cryogenic probe, all 2D NMR experiments were acquired with non-uniform sampling (NUS) set to 25% using the standard Bruker pulse sequences. For the 5 mm TCI cryogenic probe, all 2D NMR experiments were acquired with non-uniform sampling (NUS) set to 40% for ^{1}H–^{1}H detected experiments or 35% for ^{1}H–^{13}C detected experiments using the standard Bruker pulse sequences. Spectra were calibrated to residual solvent signals at δ_H 3.31 and δ_C 49.0 for MeOH-d_4 or δ_H 2.50 and δ_C 39.5 for DMSO-d_6. The δ_N values were not calibrated to an external standard but were referenced to neat NH_3 (δ_N 0.00) using the standard Bruker parameters. NMR FID processing and data interpretation was done using MestReNova software, version 15.0. Semi-preparative scale HPLC purification was performed with a Gilson HPLC purification system equipped with a GX-281 liquid handler, a 322-binary pump, and a 172-photodiode array detector. All solvents used for chromatography and UV were HPLC grade, and H_2O was Millipore Milli-Q PF filtered.

High-resolution mass spectra were recorded on an Agilent 1260 Infinity II UHPLC system coupled to an Agilent 6545 QToF equipped with a dual AJS ESI source. A Kinetex C_{18} column (50.0 × 2.1 mm, 1.7 μm, 100 Å Phenomenex) was used. The mobile phase consisted of H_2O + 0.1% formic acid (A) and MeCN + 0.1% formic acid (B) at a flow rate of 0.7 mL/min. The gradient used was maintained at 95:5 (A:B) for 0.5 min, from 95:5 to 0:100 (A:B) for 8.0 min, maintained at 0:100 (A:B) for 0.5 min, from 0:100 to 95:5 (A:B) for 0.5 min and then equilibrated in a post-run at 95:5 (A:B) during 1.0 min. The positive mode ESI conditions were 3 kV of capillary voltage, 1 kV of nozzle voltage, gas temperature at 300 °C and gas flow of 10 L/min. The data dependent acquisition (DDA) mode was used, under the following parameters: positive ionization; acquisition time from

0.5 to 9 min; *m/z* range 100–3200; scan time of 0.5 s; isolation width ~ 4 amu; precursor selection of 10^5 absolute intensity; 3 MS/MS per MS survey; mass exclusion real time for 0.15 min; and ramp of collision energy for MS/MS 20–30 eV for the lowest mass (*m/z* 200), 40 eV for *m/z* 800, 50–70 eV for *m/z* 1500, and 50–70 eV for the highest mass (*m/z* 3000).

Collection, extraction, and isolation

The tunicate *D. granulatum* was collected in a depth range of 8–13 m in Northern Australia in September 2002 by the Museum & Art Gallery of the Northern Territory under contract to the Natural Products Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, NCI. The specimen was taxonomically identified by Dr. Patricia Kott Mather, and a voucher (0M9H2221) was deposited at the Smithsonian Institution. The ascidian (wet weight 1110 g) was extracted in water, followed by a MeOH/DCM overnight soak according to the Natural Products Branch's standard marine extraction procedure[32], to provide 6.17 g of the organic extract (NSC #C23983). A portion of the organic extract C23983 (3.5 g) was prefractionated on a C_8 SPE column (20 g), generating seven fractions[33]: 95:5 H_2O:MeOH (C23983_1), 80:20 H_2O:MeOH (C23983_2), 60:40 H_2O:MeOH (C23983_3), 40:60 H_2O:MeOH (C23983_4), 20:80 H_2O:MeOH (C23983_5), 0:100 H_2O:MeOH (C23983_6), 50:50 MeCN:MeOH (C23983_7).

Fractions C23983_3–7 were combined (1.7 g, named C23983_3) and subjected to five HPLC separations using mobile phase of H_2O + 0.5% TFA (solvent A) and MeCN + 0.5% TFA (solvent B). (*I*) HPLC separation using a Kinetex C_8 column (150 × 21.2 mm, 5 μm, 100 Å, Phenomenex) on a gradient of 95:5 to 0:100 (A:B) for 23 min at 9.0 mL/min, and fraction collection was performed in 0.33 min increments. HPLC fractions 30–49 (159 mg, named C23983_3_30) were combined and subjected to (*II*) separation using Luna C_4(2) column (250 × 10 mm, 5 μm, 100 Å, Phenomenex) on a gradient of 68:32 to 50:50 (A:B) for 34 min at 3.8 mL/min, and fraction collection was performed in 0.50 min increments. HPLC fractions 25 – 50 (53 mg, named C23983_3_30_25) were combined and subjected to (*III*) separation using Luna C_4(2) column (250 × 10 mm, 5 μm, 100 Å, Phenomenex) on a gradient of 67:33 to 55:45 (A:B) for 32 min at 3.8 mL/min, and fraction collection was performed in 0.33 min increments. HPLC fractions 45–49 (5.5 mg, named C23983_3_30_25_45) were combined and subjected to (*IV*) separation using Luna C_4(2) column (250 × 10 mm, 5 μm, 100 Å, Phenomenex) on a gradient of 65:35 to 62:38 (A:B) for 31 min at 3.8 mL/min, and fraction collection was performed in 0.33 min increments yielding minalemine G (3.4 mg, 0.10% of organic extract yield; **1**). HPLC fractions from separation (*III*) 55–59 (4.6 mg, named C23983_3_30_25_55) were combined and subjected to (*V*) separation using Luna C_4(2) column (250 × 10 mm, 5 μm, 100 Å, Phenomenex) on a gradient of 65:35 to 60:40 (A:B) for 29 min at 3.8 mL/min, and fraction collection was performed in 0.33 min increments yielding minalemine H (3.9 mg, 0.11% of organic extract yield; **2**).

Minalemine G (***1***): colorless oil, $[\alpha]_D^{24}$ −2.97 (*c* 0.31, MeOH); UV (MeOH) λ_{max} (log ε) 220 (5.67) nm; IR (film) ν_{max} 3268, 3182, 2926, 2857, 1648, 1549, 1524, 1454, 1368, 1235, 1191, 828, 709 cm^{-1} (Fig. S8); ^1H and ^{13}C data in MeOH-d_4 and DMSO-d_6 are in Tables 1 and S4 and in Figs. S9–S28; HRESIMS *m/z* 554.4873 $[M+H]^+$ (calcd. for $C_{28}H_{60}N_9O_2^+$ *m/z* 554.4865, Δ = 1.44 ppm, Figs. S29–S30).

Minalemine H (***2***): colorless oil, $[\alpha]_D^{24}$ −4.44 (*c* 0.45, MeOH); UV (MeOH) λ_{max} (log ε) 220 (5.87), 278 (5.12) nm; IR (film) ν_{max} 3339, 3292, 3193, 3112, 2928, 2857, 1669, 1576, 1435, 1202, 1184, 1136, 839, 802, 723 cm^{-1} (Fig. S31); ^1H and ^{13}C data in MeOH-d_4 and DMSO-d_6 are Tables 1 and S4 and in Figs. S32–S47; HRESIMS *m/z* 533.4172 $[M+H]^+$ (calcd. for $C_{29}H_{53}N_6O_3^+$ *m/z* 533.4174, Δ = −0.37 ppm, Figs. S48–S49).

Compound identification of active subfraction using LC–MS/MS data

The UHPLC-MS/MS data from active subfractions C23983_4_1 to 9 was acquired using the method described above in the general procedures section. The raw data files (.d) were converted to mzML files by MSConvert 3.0.22 (ProteoWizard), with binary encoding precision 64-bit, write index and applied filter of peak picking MS levels 1–2. The mzML files were uploaded to MZmine 3.4.27[34]. Mass detection for MS^1 and MS^2 was performed using the mass detector centroid and noise level of 1.0E2. Chromatograms were reconstructed using the function ADAP Chromatogram Builder with retention time between 1.0 and 8.5 min, MS level 1, min group size in # of scans 2, group highest intensity 1.0E5, minimum height 1.0E5 and *m/z* tolerance of 0.005 *m/z* or 20.0 ppm. Deconvolution was performed using the function "group MS^2 scans with features" with MS^1 to MS^2 precursor tolerance of 0.005 *m/z* or 20.0 ppm and retention time filter of 0.50 min. The isotopic peak grouper function was applied with *m/z* tolerance of 0.005 *m/z* or 20.0 ppm, retention time tolerance of 0.50 min, maximum charge of 3, and most intense peak as the representative isotope. Data were aligned using the function join aligner, with *m/z* tolerance of 0.005 m/z or 20.0 ppm, weight for *m/z* of 75, R_T tolerance of 0.5 min and weight for R_T of 25. The feature list generated after the described steps was exported using the function Export/Submit to GNPS, generating one mgf and one csv file.

The generated csv file was used to search against an *in-silico* library comprising 182 compounds from the genus *Didemnum* assembled from the *Dictionary of Natural Products* database. First, the accurate mass of *m/z* features present in the active subfractions of *D. granulatum* was searched against an *in-silico* library comprising protonated molecules and sodium adducts with exact mass (± 20 ppm). A comparison between the MS^2 spectrum of annotated features and literature reports was then used to support the annotation.

Competing enantioselective conversion (CEC)

The CEC reaction followed by LC–MS analysis was based on a previously described procedure[23,35]. Compound **1** (0.20 mg, 0.36 μmol) was transferred to two different vials, and dimethylformamide (90 μL) was added as an organic solvent. The catalysts *S*- and *R*-HBTM (5.0 μL, 4.3 μmol) were added to their respective vials, along with *N*,*N*-diisopropylethylamine (0.7 μL, 4.3 μmol). Propionic anhydride (0.5 μL, 4.3 μmol) was added to initiate the

reaction. Aliquots of 2 μL were taken every 5 min and quenched with 100 μL of MeOH prior to LC–MS analysis, for a total reaction time of 60 min.

Aliquots (5 μL) of the samples collected at different time intervals were injected on an Agilent 1260 Infinity II HPLC system coupled to an Agilent 6230 ToF equipped with a dual AJS ESI source. A Kinetex C_{18} column (50.0 × 2.1 mm, 5 μm, Phenomenex) was used for the analysis. The mobile phase consisted of H_2O + 0.1% formic acid (A) and MeCN + 0.1% formic acid (B) at a flow rate of 1 mL/min. The gradient used was maintained at 95:5 (A:B) for 1.0 min, from 95:5 to 0:100 (A:B) for 8.0 min, maintained at 0:100 (A:B) for 1.0 min, from 0:100 to 95:5 (A:B) for 0.5 min and then equilibrated in a post-run at 95:5 (A:B) during 1.5 min. The positive mode ESI conditions were 3.5 kV of capillary voltage, 1.5 kV of nozzle voltage, gas temperature at 325 °C and gas flow of 10 L/min. MS spectra were acquired using positive mode, mass range of *m/z* 100–3000, and scan rate of 1 spectra/s. The reaction rates were determined by measuring the peak areas of the fully acylated derivative of compound **1**, using the extracted ion chromatogram (EIC) for the sodiated molecule $[M + Na]^+$ at *m/z* 722.5651 (± 20 ppm) on MassHunter Qualitative Analysis 10.0 software.

Chiral chromatography

The chiral chromatography analysis were performed on a Shimadzu system equipped with a CBM-40 controller, an SPD-M40 PDA detector, and two LC-20AR pumps. A Lux amylose-3 column (250.0 × 4.6 mm, 5 μm, Phenomenex) was used for the analysis. The mobile phase consisted of H_2O + 0.1% TFA (A) and MeCN + 0.1% TFA (B) using an isocratic mode at 80% of B at a flow rate of 1 mL/min. The injection volume was 5 µL.

LC–MS/MS analysis and data processing for creating the neutral loss graph

UHPLC-MS/MS data on the scale-up fractions after C_8 SPE (C23983_1–7) was acquired, converted to mzML, and processed using MZmine using the same parameters as described in the compound identification section. The mgf file was used for a combinatorial neutral loss analysis within each MS^2 spectrum, where the mass difference was calculated for the precursor ion and its fragments, and between fragments within the same feature. The combinatorial analysis generated a mgf file containing neutral losses for each *m/z* feature.

Three sets of target neutral losses were established, namely common neutral losses (CNL), and unique neutral losses of **1** (UNL1) and **2** (UNL2). The CNL contained six neutral losses from the Hagma unit in compounds **1** and **2** (Table S2). The UNL1 is constituted by four larger neutral losses from Hagma-FA-Gly, substructures of **1** (Table S2), while UNL2 is composed of four neutral losses from Tyn-FA-*N*-Me-Gly, substructures of **2** (Table S2).

Neutral loss graph package

A simplified version of the molecular networking algorithm[36] was used to generate a neutral loss graph. The similarity of two features was based on the intersection of their respective neutral loss vectors. Neutral losses for a given feature were derived from the all-to-all differences of the MS2 peak. Given two features with neutral loss vectors v_1 and v_2, the similarity between features was calculated as:

$$0 \leq \max\left(\frac{\sum_{x\in(v_1\cap v_2)}^{x}}{\sum_{x\in v_1}^{x}}, \frac{|v_1 \cap v_2|}{|v_1|}\right) \leq 1$$

A pair wise matrix of all similarities (1 representing a complete or proper subset) was used as input for generating an undirected adjacency graph. The resultant graph object was exported as a graphml file to be imported and visualized using Cytoscape 3.9.1[37] (Fig. S50). The algorithm implementation was written using the R statistical computing software (version 4.3.0)[38]. The MSnbase package[39,40] was used to process mgf files and the igraph package[41] to build the network and generate the graphml file. Code is available on GitHub at https://github.com/NCI-DCTD/neutralLossGraph.

Antimicrobial assay

Microbes, growth, and testing conditions were identical to those reported previously for the primary screen[16].

Data availability

The data that support the findings of the present study are available in the SI of this article. In addition, the raw HRMS and NMR data have been deposited in the Harvard Dataverse (dataverse.harvard.edu) and can be found at [https://doi.org/10.7910/DVN/W3VBPS]).

Received: 17 September 2025; Accepted: 8 December 2025
Published online: 07 January 2026

References

1. Mayer, A. M. S. et al. Marine pharmacology in 2019–2021: marine compounds with antibacterial, antidiabetic, antifungal, anti-inflammatory, antiprotozoal, antituberculosis and antiviral activities; affecting the immune and nervous systems, and other miscellaneous mechanisms of action. *Mar. Drugs* **22**, 309. https://doi.org/10.3390/md22070309 (2024).
2. Zhang, Q. et al. Recent advances of bioactive marine natural products in drug discovery. *J. Oc. Uni. China* **23**, 1297–1318. https://doi.org/10.1007/s11802-024-5975-4 (2024).
3. Mayer, A. M. S. & Pierce, M. L. *Marine pharmacology*, (accessed on Jul 2025) https://www.marinepharmacology.org/approved (2021).
4. Cooreman, K. et al. Emerging pharmaceutical therapies of ascidian-derived natural products and derivatives. *Environ. Toxicol. Pharmacol.* **102**, 104254. https://doi.org/10.1016/j.etap.2023.104254 (2023).

5. Berlinck, R. G. S. et al. Granulatimide and isogranulatimide, aromatic alkaloids with g2 checkpoint inhibition activity isolated from the brazilian ascidian *Didemnum granulatum*: structure elucidation and synthesis. *J. Org. Chem.* **63**, 9850–9856. https://doi.org/10.1021/jo981607p (1998).
6. Fernandez, A. M. et al. Structural studies of marine peptides. *Pure Appl. Chem.* **70**, 2130 (1999).
7. Teruya, T., Suenaga, K., Maruyama, S., Kurotaki, M. & Kigoshi, H. Biselides A-E: Novel polyketides from the Okinawan ascidian *Didemnidae* sp. *Tetrahedron* **61**, 6561–6567. https://doi.org/10.1016/j.tet.2005.04.052 (2005).
8. Gonzalez, N., Rodriguez, J. & Jimenez, C. Didemniserinolipids A-C, unprecedented serinolipids from the tunicate *Didemnum* sp. *J. Org. Chem.* **64**, 5705–5707. https://doi.org/10.1021/jo9903914 (1999).
9. Takeara, R. et al. Antileukemic effects of *Didemnum psammatodes* (Tunicata: Ascidiacea) constituents. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **151**, 363–369. https://doi.org/10.1016/j.cbpa.2007.02.011 (2008).
10. Youssef, D. T. A., Almagthali, H., Shaala, L. A. & Schmidt, E. W. Secondary metabolites of the genus *Didemnum*: a comprehensive review of chemical diversity and pharmacological properties. *Mar. Drugs* **18**, 307. https://doi.org/10.3390/md18060307 (2020).
11. Wei, M. et al. Progress on synthesis and structure-activity relationships of lamellarins over the past decade. *Eur. J. Med. Chem.* **269**, 116294. https://doi.org/10.1016/j.ejmech.2024.116294 (2024).
12. Potts, B. C. M. et al. Didemnaketals A and B, HIV-1 protease inhibitors from the ascidian *Didemnum* sp. *J. Am. Chem. Soc.* **113**, 6321–6322. https://doi.org/10.1021/ja00016a087 (2002).
13. Wright, A. D., Goclik, E., Konig, G. M. & Kaminsky, R. Lepadins D-F: antiplasmodial and antitrypanosomal decahydroquinoline derivatives from the tropical marine tunicate *Didemnum* sp. *J. Med. Chem.* **45**, 3067–3072. https://doi.org/10.1021/jm0110892 (2002).
14. Oku, N., Matsunaga, S. & Fusetani, N. Shishijimicins A-C, novel enediyne antitumor antibiotics from the ascidian *Didemnum proliferum*. *J. Am. Chem. Soc.* **125**, 2044–2045. https://doi.org/10.1021/ja0296780 (2003).
15. Vollmar, B. S. et al. Calicheamicin antibody–drug conjugates with improved properties. *Mol. Can. Ther.* **20**, 1112–1120. https://doi.org/10.1158/1535-7163.Mct-20-0035 (2021).
16. Martinez-Fructuoso, L. et al. Screen for new antimicrobial natural products from the nci program for natural product discovery prefractionated extract library. *ACS Infect. Dis.* **9**, 1245–1256. https://doi.org/10.1021/acsinfecdis.3c00067 (2023).
17. Grkovic, T. et al. National Cancer Institute (NCI) program for natural products discovery: rapid isolation and identification of biologically active natural products from the NCI prefractionated library. *ACS Chem. Biol.* **15**, 1104–1114. https://doi.org/10.1021/acschembio.0c00139 (2020).
18. Gaudencio, S. P. et al. Advanced methods for natural products discovery: bioactivity screening, dereplication, metabolomics profiling, genomic sequencing, databases and informatic tools, and structure elucidation. *Mar. Drugs* **21**, 308. https://doi.org/10.3390/md21050308 (2023).
19. FoxRamos, A. E., Evanno, L., Poupon, E., Champy, P. & Beniddir, M. A. Natural products targeting strategies involving molecular networking: different manners, one goal. *Nat. Prod. Rep.* **36**, 960–980. https://doi.org/10.1039/c9np00006b (2019).
20. Taylor & Francis Group. *Dictionary of Natural Products 33.2* (acessed Feb 2025) https://dnp.chemnetbase.com/chemical/ChemicalSearch.xhtml?dswid=-408z (2024).
21. Kossuga, M. H. et al. Antibacterial modified diketopiperazines from two ascidians of the genus *Didemnum*. *J. Braz. Chem. Soc.* **20**, 704–711. https://doi.org/10.1590/s0103-50532009000400014 (2009).
22. Kouchaksaraee, R. M., Li, F., Nazemi, M., Farimani, M. M. & Tasdemir, D. Molecular networking-guided isolation of new etzionin-type diketopiperazine hydroxamates from the persian gulf sponge *Cliona celata*. *Mar. Drugs* **19**, 439. https://doi.org/10.3390/md19080439 (2021).
23. Freire, V. F. et al. Antimicrobial bianthrones from the crinoid *Heterometra* sp. *RSC Adv.* **14**, 38200–38207. https://doi.org/10.1039/d4ra05594b (2024).
24. Expósito, M. A. et al. Minalemines A-F: Sulfamic acid peptide guanidine derivatives isolated from the marine tunicate *Didemnun rodriguesi*. *Tetrahedron* **54**, 7539–7550. https://doi.org/10.1016/s0040-4020(98)00388-3 (1998).
25. Aisporna, A. et al. Neutral loss mass spectral data enhances molecular similarity analysis in METLIN. *J. Am. Soc. Mass. Spectrom.* **33**, 530–534. https://doi.org/10.1021/jasms.1c00343 (2022).
26. Vitale, G. A. et al. Connecting metabolome and phenotype: recent advances in functional metabolomics tools for the identification of bioactive natural products. *Nat. Prod. Rep.* **41**, 885–904. https://doi.org/10.1039/d3np00050h (2024).
27. Freire, V. F. et al. Feature-based molecular networking discovery of bromopyrrole alkaloids from the marine sponge *Agelas dispar*. *J. Nat. Prod.* **85**, 1340–1350. https://doi.org/10.1021/acs.jnatprod.2c00094 (2022).
28. Meunier, M., Schinkovitz, A. & Derbre, S. Current and emerging tools and strategies for the identification of bioactive natural products in complex mixtures. *Nat. Prod. Rep.* **41**, 1766–1786. https://doi.org/10.1039/d4np00006d (2024).
29. Xing, S. et al. Retrieving and utilizing hypothetical neutral losses from tandem mass spectra for spectral similarity analysis and unknown metabolite annotation. *Anal. Chem.* **92**, 14476–14483. https://doi.org/10.1021/acs.analchem.0c02521 (2020).
30. Bittremieux, W. et al. Comparison of cosine, modified cosine, and neutral loss based spectrum alignment for discovery of structurally related molecules. *J. Am. Soc. Mass Spectrom.* **33**, 1733–1744. https://doi.org/10.1021/jasms.2c00153 (2022).
31. Maccari, G. et al. Synthesis of linear and cyclic guazatine derivatives endowed with antibacterial activity. *Bioorg. Med. Chem. Lett.* **24**, 5525–5529. https://doi.org/10.1016/j.bmcl.2014.09.081 (2014).
32. McCloud, T. G. High throughput extraction of plant, marine and fungal specimens for preservation of biologically active molecules. *Molecules* **15**, 4526–4563. https://doi.org/10.3390/molecules15074526 (2010).
33. Thornburg, C. C. et al. NCI program for natural product discovery: a publicly-accessible library of natural product fractions for high-throughput screening. *ACS Chem. Biol.* **13**, 2484–2497. https://doi.org/10.1021/acschembio.8b00389 (2018).
34. Schmid, R. et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **41**, 447–449. https://doi.org/10.1038/s41587-023-01690-2 (2023).
35. Lee, S. R., Park, H. B. & Kim, K. H. Highly sensitive, simple, and cost- and time-effective method to determine the absolute configuration of a secondary alcohol using competing enantioselective acylation coupled with LC/MS. *Anal. Chem.* **90**, 13212–13216. https://doi.org/10.1021/acs.analchem.8b03385 (2018).
36. Yang, J. Y. et al. Molecular networking as a dereplication strategy. *J. Nat. Prod.* **76**, 1686–1699. https://doi.org/10.1021/np400413s (2013).
37. Shannon, P. et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504. https://doi.org/10.1101/gr.1239303 (2003).
38. Team, R. C. *R: A Language and Environment for Statistical Computing*. https://www.r-project.org (2021).
39. Gatto, L., Gibb, S. & Rainer, J. MSnbase, efficient and elegant R-based processing and visualization of raw mass spectrometry data. *J. Proteome Res.* **20**, 1063–1069. https://doi.org/10.1021/acs.jproteome.0c00313 (2021).
40. Gatto, L. & Lilley, K. S. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics* **28**, 288–289. https://doi.org/10.1093/bioinformatics/btr645 (2012).
41. Csárdi, G., Nepusz, T., Müller, K., Horvát, S., Traag, V., Zanini, F., Noom, D. igraph for R: R interface of the igraph library for graph theory and network analysis (v2.1.4). *Zenodo*. https://doi.org/10.5281/zenodo.14736815 (2025).

Acknowledgements

The authors thank Drs. Erin Zeituni and Joseph Campbell from the National Institute of Allergy and Infectious

Diseases for helpful discussions, Lin Du from the Molecular Target Program for assistance in the chiral chromatography analysis, and Samantha C. Waterworth from the Molecular Target Program for assistance with coding evaluation. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health (NIH), under contract HHSN261200800001E and by the National Cancer Institute's NCI Program for Natural Products Discovery-Cure (ZIA BC 011854), as well as the Extramural and the Intramural Research Programs of the NIH. The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Author contributions

V.F.F.: conceptualization, investigation, methodology, data curation, visualization, formal analysis and writing—original draft; J.E.: investigation, software, data curation, formal analysis and writing—original draft; L.M.F. formal analysis and review; R.K.: formal analysis and review; R.K.A.: formal analysis; S.H.: formal analysis; C.C.T.: formal analysis, resources and review; B.D.P.: formal analysis and review; S.E.: formal analysis and review; D.W.: formal analysis and review; T.G.: project administration, supervision, methodology, review; B.R.O.: project administration, supervision, funding acquisition and review.

Funding

Open access funding provided by the National Institutes of Health

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-32070-2.

Correspondence and requests for materials should be addressed to V.F.F. or B.R.O.