Giant barrel sponges in diverse habitats: a story about the metabolome

Bayona Maldonado, L.M.

Chapter 4

Influence of the geographical location on the metabolic production of giant barrel sponges (*Xestospongia* spp.) revealed by metabolomics tools

Lina M. Bayona [1], Gemma van Leeuwen [1], Özlem Erol [1], Thomas Swierts [2,3], Esther van der Ent [2,3] Nicole J. de Voogd [2,3], Young Hae Choi [1*]

[1]Natural Products Laboratory, Institute of Biology, Leiden University, Leiden, The Netherlands.

[2]Marine Biodiversity, Naturalis Biodiversity Center, Leiden, The Netherlands

[3] Institute of Environmental Sciences, Leiden University, Leiden, The Netherlands

* Corresponding author e-mail: y.choi@chem.leidenuniv.nl

ACS Omega (2020), 5(21), 12398–12408.

https://doi.org/10.1021/acsomega.0c01151

Abstract

Despite their high therapeutic potential, only a limited number of approved drugs originate from marine natural products. A possible reason for this is their broad metabolic variability related to the environment, which can cause reproducibility issues. Consequently, a further understanding of environmental factors influencing the production of metabolites is required. Giant barrel sponges, *Xestospongia* spp.*,* are a source of many new compounds and are found in a broad geographical range. In this study, the relationship between the metabolome and the geographical location of sponges within the genus *Xestospongia* spp. was investigated. One hundred and thirty-nine specimens of giant barrel sponges (*Xestospongia* spp.) collected in four locations, Martinique, Curaçao, Taiwan, and Tanzania, were studied using a multiplatform metabolomics methodology (NMR and LC-MS). A clear grouping of the collected samples according to their location was shown. Metabolomics analysis revealed that sterols and various fatty acids, including polyoxygenated and brominated derivatives, were related to the difference in location. To explore the relationship between observed metabolic changes and their bioactivity, antibacterial activity was assessed against *Escherichia coli* and *Staphylococcus aureus.* The activity was found to correlate with brominated fatty acids. These were isolated and identified as (9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid (**1**), Xestospongic acid (**2**), (7*E*,13*E*,15*Z*)-14,16-dibromohexadeca-7,13,15-trien-5-ynoic acid (**3**) and two previously unreported compounds.

Keywords: Metabolomics, antibacterial, marine sponge, giant barrel sponge, geographic location, brominated fatty acids, *Xestospongia* spp.

1. Introduction

Marine natural products (MNP) have a wide chemical diversity, covering a broader area in the chemical spectrum compared to their terrestrial counterparts (Blunt et al. 2018). The chemical structures of metabolites isolated from marine organisms contain highly characteristic features and many of them have shown diverse bioactivities. In the past decades, the isolation of novel and bioactive molecules from marine organisms has been a hot issue in natural product research resulting, so far, in the development of eight drugs, which have been approved and are currently available for the treatment of cancer, HIV and pain(Altmann 2017; Gerwick and Moore 2012; Newman and Cragg 2004). Despite their potential, the number of approved drugs is low considering the large number of compounds that have been discovered from marine sources. In fact, while more than 1200 new compounds are reported every year, the number of MNP-derived approved drugs has not been increasing at the same rate (Blunt et al. 2016, 2017, 2018).

Although many of the metabolites produced by marine organisms have proved to be active, these compounds are usually produced in very small amounts (Belarbi et al. 2003). During the process of drug development, large quantities of the compound are required to perform all the preclinical and clinical trials that are necessary for a drug to be approved (Gupta 2011). Unfortunately, the large scale harvesting of the organism required for this is not feasible either from an economical or an ecological perspective (Altmann 2017). Moreover, the production of metabolites in marine organisms can change due to environmental factors such as pH, temperature, predation pressure and subsequent changes in symbionts community, making them too unreliable both qualitatively and quantitatively as a natural source of compounds (Viant 2007).

To overcome this, diverse approaches have been suggested, including aqua- and mariculture (Belarbi et al. 2003; Pomponi 1999). Although these techniques have not been used yet for the production of compounds at a commercial scale, it is thought that their implementation could provide sufficient amounts of the compounds to meet the demand for clinical and preclinical trials (Cuevas and Francesch 2009). The successful cultivation of marine organisms, mainly of sponges (de Voogd 2007; Ruiz et al. 2013; Santiago et al. 2019), resulting in the production of higher quantities of active metabolites (Hadas et al. 2005; Page et al. 2005), could guarantee the reliability of the sources, paving the way for their approval for medicinal use. Optimization of growth and production conditions for the cultivation of the organisms requires an understanding of how biotic and environmental factors affect their metabolome. Such a study involving so many variables can benefit from the use of an untargeted approach

that allows the acquisition of the most inclusive picture of the metabolome and then observes how it varies with changing external factors. Metabolomics, defined as comprehensive profiling of all the metabolites produced by an organism, cell, or tissue at a certain point in time can provide the information, which could then be used for guidance on a variety of compounds produced and uncovering the factors associated with their production (Kim et al. 2010).

Among marine organisms, sponges have been considered to be the most prolific in the production of secondary metabolites, most of which have biological activity as proved by their performance in a wide variety of bioassays (Belarbi et al. 2003; Blunt et al. 2018; Mehbub et al. 2014). In particular, giant barrel sponges, which belong to the genus *Xestospongia*, have drawn the attention of the scientific community due to their pharmacological activities and their role in ecosystems (Fiore et al. 2013; Zhou et al. 2010). In ecological systems, their large size allows them to play an essential role in the reef, providing habitat for other organisms and filtering vast amounts of seawater (Diaz and Rützler 2001; Swierts et al. 2018). Therefore, the tight interaction of giant barrel sponges with their environment makes them an interesting model to study the relationship between metabolites and environmental factors. Also, in some locations these sponges have been reported to cover up to 9% of the reef substrate, being more abundant than any other invertebrate (McMurray et al. 2008; Zea 1993). Their chemical composition has been studied, and a wide range of compounds have been isolated including alkaloids, brominated fatty acids, and sterols. Many of these compounds have proved to be bioactive, displaying antibacterial, cytotoxicity, fungicide, and antiretroviral activities (Zhou et al. 2010).

In addition, giant barrel sponges can be found in a wide geographical range: *Xestospongia testudinaria* from the Red Sea to the Indo-Pacific Ocean and Australia, and *Xestospongia muta* in the tropical regions of the Atlantic Ocean. These two species show very similar genetic and morphological markers (Setiawan et al. 2016). Furthermore, recent studies revealed the presence of cryptic species in both ocean basins (Swierts et al. 2017). Interestingly for this study, some of the species present in the Caribbean Sea are genetically much closer to species in the Indo-Pacific than to other species in the same location (Swierts et al. 2013, 2017). These similarities in the cryptic species between locations provides the opportunity to focus on the differences in the metabolome caused by environmental factors.

Geographical location has been identified as one of the most influential factors related to the variation of many sponge metabolites (Page et al. 2005; Rohde et al. 2012; Sacristan-Soriano et al. 2011). However, the results that led to this conclusion were aimed at a few target

metabolites, while the more general effect on the whole metabolome, which requires a holistic approach, has scarcely been studied (Reverter et al. 2018). To study the correlation between the geographical location and metabolic production, 139 specimens of giant barrel sponges (*Xestospongia* spp.), collected in four different geographic locations: Martinique, Curaçao, Taiwan, and Tanzania were studied using a holistic approach. Applying multiplatform metabolomics methodology (nuclear magnetic resonance spectroscopy (NMR) and liquid chromatography-mass spectrometry (LC-MS)), we aimed to investigate the effect of geographical location on the chemical composition of the sponges. Additionally, the correlation between the metabolic changes observed in the samples and their antibacterial activity was evaluated. This proved that the implementation of a metabolomics approach to MNPs can provide relevant information on the conditions required to optimize the production of bioactive compounds. Furthermore, the presence of minor active compounds largely influenced by location-related factors can be revealed using this approach.

2. Results and discussion

The metabolic profile of giant barrel sponge samples collected in four different geographical locations showed clear differences in the chemical composition of the specimens collected in each location. To compare the general metabolic profile of the samples, ^{1}H-NMR and LC-MS were separately applied to the same sample set. These data were further analyzed using an orthogonal partial least-squares discriminant analysis (OPLS-DA) model (Figure 4.1). Both models, ^{1}H-NMR and LC-MS, were validated with a Q^2 value > 0.4 and cross-validation analysis of variance (CV-ANOVA) test $p < 0.05$ (Cai et al. 2012; Zheng et al. 2011).

In fact, for giant barrel sponges *X. muta* and *X. testudinaria*, the composition of sterols (Gauvin et al. 2004) and some brominated fatty acids (Zhou et al. 2010) was previously found to be similar between sponges collected in different oceans. These previous studies showed that despite large geographical separation, giant barrel sponges could share a common metabolic background in qualitative features. In this study, however, a significant separation between the samples collected from different places was observed in the OPLS-DA analysis (Figure 4.1). This result might indicate that the environmental conditions in each location could quantitatively influence the metabolome of the sponges.

The location in which sponges grow involves a number of factors that can influence their development and metabolism, including abiotic factors such as temperature, pH, salinity, or the biotic predatory stress. The effect of the combination of these factors could cause that sponges collected from a specific location produce similar metabolites. Furthermore,

Xestospongia spp. are high microbial abundance (HMA) sponges, and microbial communities have been reported to mainly be affected by geographical location (Swierts et al. 2018). Thus, it is plausible to find differences in the chemical composition of sponges from different locations, as the metabolome corresponds to the holobiont and the metabolites found can either be produced by the sponge, by the microorganisms or they can be the product of the interaction of the sponges with microorganisms (Gerwick and Moore 2012).

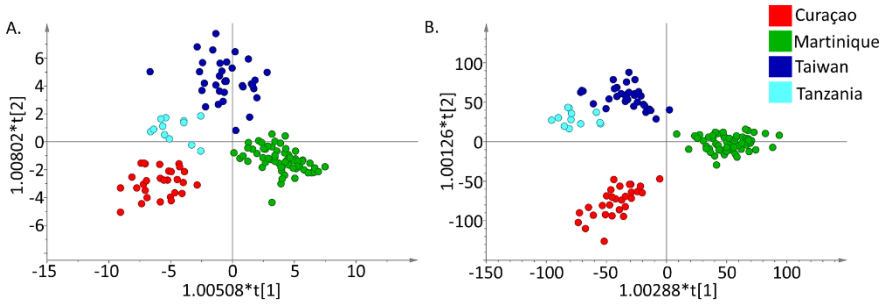

Figure 4.1: First two components of the OPLS-DA analysis based on ^{1}H-NMR (A) and LC-MS (B) of *Xestospongia* spp. samples collected in four locations: Curaçao (Red), Martinique (Green), Taiwan (Dark blue) and Tanzania (Light blue).

The loading plots of the OPLS-DA analysis (NMR and LC-MS data) were analyzed to select the discriminating signals and subsequently identify the corresponding compounds. The characteristic ^{1}H-NMR chemical shifts are shown in a heat map in Figure 4.2, obtained by calculation of the variable importance for the projection (VIP) values. The signals correlated with the samples from Martinique were found mainly in two regions of the spectra. The region between δ_H 0.80 and 1.00 was assigned to methyl groups in sterols. Particularly the singlets in the range of δ_H 0.7-0.8 were assigned to methyls H-18 and H-19 in sterols. Many steroids have been reported in *Xestospongia* spp., including conventional sterols (Kerr et al. 1991), and brominated fatty acids esters (Pham et al. 1999). The aromatic region between δ_H 7.04 and δ_H 7.32 is characteristic of phenolic signals that could correspond to known phenolics of *Xestospongia* such as quinones (Roll et al. 1983), isoquinoline alkaloids (Calcul et al. 2003), and β-carboline alkaloids (Kobayashi et al. 1995). Samples from Curaçao were distinguished by abundant signals in the range of δ_H 2.50-3.80. Signals at downfield of this range (δ_H.3-3.8) correspond to protons attached to oxygen-bearing carbons. These could thus be attributed to hydroxylated polyunsaturated fatty acids, since there are many reports of the isolation of this type of fatty acids from *Xestospongia* spp.(Jiang et al. 2011; Liu et al. 2011; Morinaka et al. 2007). Taiwan samples displayed characteristic signals between δ_H 6.40 and 6.60, which

correspond to double bonds commonly occurring in brominated unsaturated fatty acids. Samples from Tanzania had no distinguishing signals in a specific region of the spectrum, indicating that the changes present in this location do not involve a family of compounds, but rather specific compounds.

The NMR analysis provided a general overview of the metabolic profiles, allowing the detection of families of compounds predominant in each location. However, the congestion of signals in the spectra and the relatively low sensitivity rendered the identification of individual metabolites unfeasible. Thus, LC-MS/quadrupole time of flight (Q-TOF) was used to identify these metabolites, especially the minor ones. As shown in Figure1b, metabolic differences in the samples from each location were as clear as those observed with ^{1}H-NMR. As in the case of ^{1}H-NMR, a VIP plot was also used for the identification of peaks responsible for the separation. However, dereplication of the 50 most relevant peaks obtained from the VIP plot was not successful, because most of the selected MS features could not be identified, or they corresponded to several isomers. Nevertheless, information on a specific metabolites group, brominated fatty acids, was obtained from MS data. Different types of brominated fatty acids were found to be differential features in the samples on each location. Martinique samples showed no bromine-containing signals, while the Curaçao samples were discriminated by their characteristic dibrominated metabolites and the samples from Taiwan and Tanzania by monobrominated ones.

The variation in the chemical composition of the samples observed in this study proves the plasticity of *Xestospongia* spp. in terms of their biosynthesis processes. This could partly explain the great diversity in compounds isolated from this same sponge genus all over the world. Considering that these compounds exhibit a wide range of biological activities, it could be presumed that this metabolic differentiation observed in samples from different locations could be reflected in their bioactivity (Page et al. 2005). To investigate this potential correlation, the antimicrobial activity of *Xestospongia* spp. extracts against a Gram positive (*Staphylococcus aureus*) and a Gram negative (*Escherichia coli*) bacteria was assayed. This particular bioactivity was chosen due to numerous reports of antimicrobial compounds in *Xestospongia* spp. collected throughout the world (Bourguet-Kondracki et al. 1992; He et al. 2015).

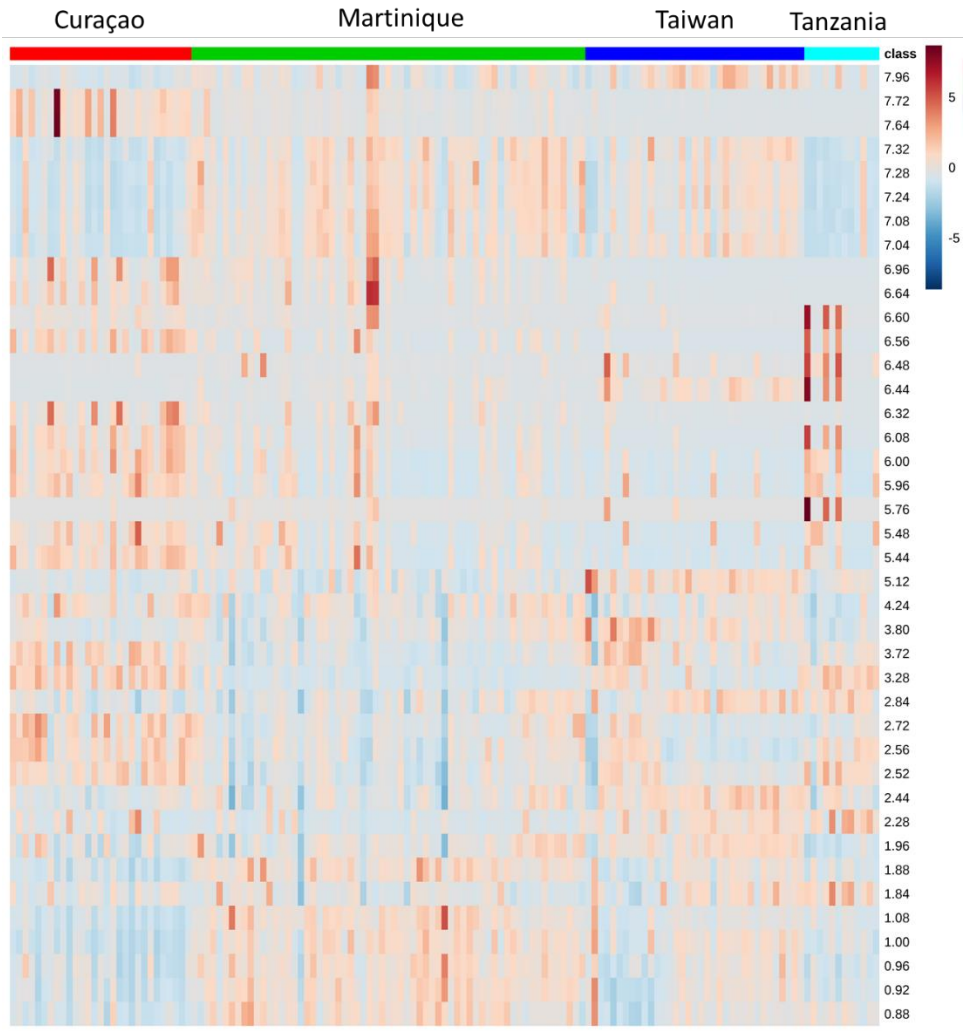

Figure 4.2: Heat map of characteristic signals from the ^{1}H-NMR data obtained from the variable importance for the projection (VIP) plot of orthogonal partial least square discriminant analysis (OPLS-DA).

The result of the activity test showed that some sponge extracts were active against *S. aureus* at a concentration of 512 µg/mL. From the whole sample set, 11.5% of the collected samples displayed activity, although there was a large variation in the activity according to the location. For example, while 20% of the samples collected in Taiwan had antimicrobial activity, none of the samples from Tanzania displayed activity. Although differences in the activity between collection places were observed, the ratio of active and nonactive samples was not significantly related to the collection places (χ^2 (2) = 2.72, p = 0.256). The lack of relation

between these two factors suggests that the production of antibacterial compounds is triggered by a factor occurring within a smaller spatial scale or is driven by genetic variation.

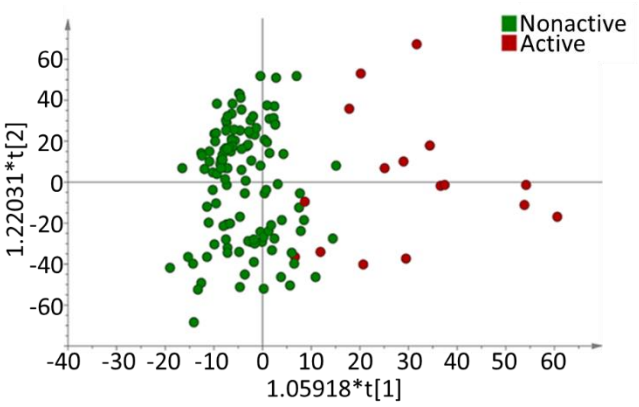

Figure 4.3: orthogonal partial least square discriminant analysis (OPLS-DA) model for the 139 *Xestospongia* spp. samples categorized by their activity against *Staphylococcus aureus* using LC-MS data.

On the other hand, none of the samples showed activity against *E. coli*, a proteobacteria, when tested at a concentration of 512 µg/mL. The lack of activity against *E.coli* can be explained by the fact that proteobacteria are one of the most predominant phyla among the bacterial communities of *Xestospongia* spp.(Fiore et al. 2013; Swierts et al. 2018). Therefore, it is a natural result that they do not produce compounds that could inhibit the growth of these types of bacteria.

To identify the compounds specifically involved in the antibacterial activity against *S. aureus*, an OPLS-DA model was built, grouping the samples as active (showed activity at 512 µg/mL) and nonactive (no activity shown at concentrations of 512 µg/mL) and using both NMR and LC-MS data. The model based on NMR data was not validated and did not reveal differences between the two groups. In this case, overlapping of signals belonging to compounds of the same family or low sensitivity could explain the lack of validation, as the activity must be related to specific compounds. On the other hand, with the LC-MS data, it was possible to separate the samples that displayed activity from the non-active samples as shown in Figure 4.3. Although variation in the chemical composition among the active samples was observed, a list of the masses of potentially active compounds was made using an S-plot, (appendix 1 Table S2). These features, together with the list obtained previously from the OPLS-DA analysis using location as a factor, were used to target the compounds of interest from samples collected in Martinique, Curaçao and Taiwan. The great dispersion observed between the

active samples suggested that although all samples exhibited activity, it was not necessarily due to the same compounds or alternatively, that there was a significant variation in the amount of active compounds present in the samples depending on the location. To clarify this, some of the most active compounds were isolated and tested, and their resulting activity was compared with their occurrence in different locations.

2.1 Isolation and structural elucidation

Ethanolic extracts of samples from Martinique, Curaçao, and Taiwan that were active against *S. aureus* were prepared to isolate potentially active antimicrobial compounds. These extracts were subjected to fractionation with liquid chromatography, using the list of potential active features as a criterion for fraction selection. This led to the isolation of five brominated fatty acid analogues (Figure 4.4): two from Martinique extracts (**1**,**2**), two from Curaçao extracts (**3**,**4**), and one from Taiwan extracts (**5**).

Compound **1** was isolated from a Martinique sample as a white powder. The (+)-HRESIMS spectrum of **1** showed the proton adduct $[M+H]^+$ ions at *m/z* 347.0646 and 349.0631, with relative intensities of 1:1, suggesting the presence of one bromine atom in the molecule. The molecular formula was deduced to be $C_{18}H_{19}BrO_2$. The ^{1}H-NMR (CH_3OH-d_4, 600 MHz) spectrum of the compound showed the presence of two double bonds and the ^{13}C-NMR (CH_3OH-d_4, 150 MHz) showed one carboxylic acid carbon, and the presence of three triple bonds. The molecular formula together with the characteristic NMR signals were dereplicated using the Dictionary of Natural Products. The compound was identified as (9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid, which had been previously isolated from *X. muta* collected in Columbus Island, Bahamas, and reported to inhibit the HIV-1 protease with an IC_{50} of 8 µM (Patil et al. 1992).

Compound **2** was also isolated from a Martinique sample as a white powder. The (+)-HRESIMS spectrum of **2** showed the proton adduct $[M+H]^+$ ions at *m/z* 351.0956 and 353.0939, with relative intensities of 1:1. This isotopic pattern suggested the presence of a bromine atom in the molecule. The molecular formula was deduced to be $C_{18}H_{23}BrO_2$. The ^{1}H-NMR (CH_3OH-d_4, 600 MHz) spectroscopic data, showed the presence of two double bonds and ^{13}C-NMR (CH_3OH-d_4, 150 MHz) showed the presence of two triple bonds. A search using the molecular formula and the characteristic NMR signals in the Dictionary of Natural Products, yielded the compound (9*E*,17*E*)-18-bromooctadeca-9,17-dien-7,15-diynoic acid also known as xestospongic acid. This compound had been originally isolated from *Xestospongia sp* samples

collected in Australia as one of the most abundant compounds in the sample accounting for 0.1% of the dry weight material (Quinn and Tucker 1985).

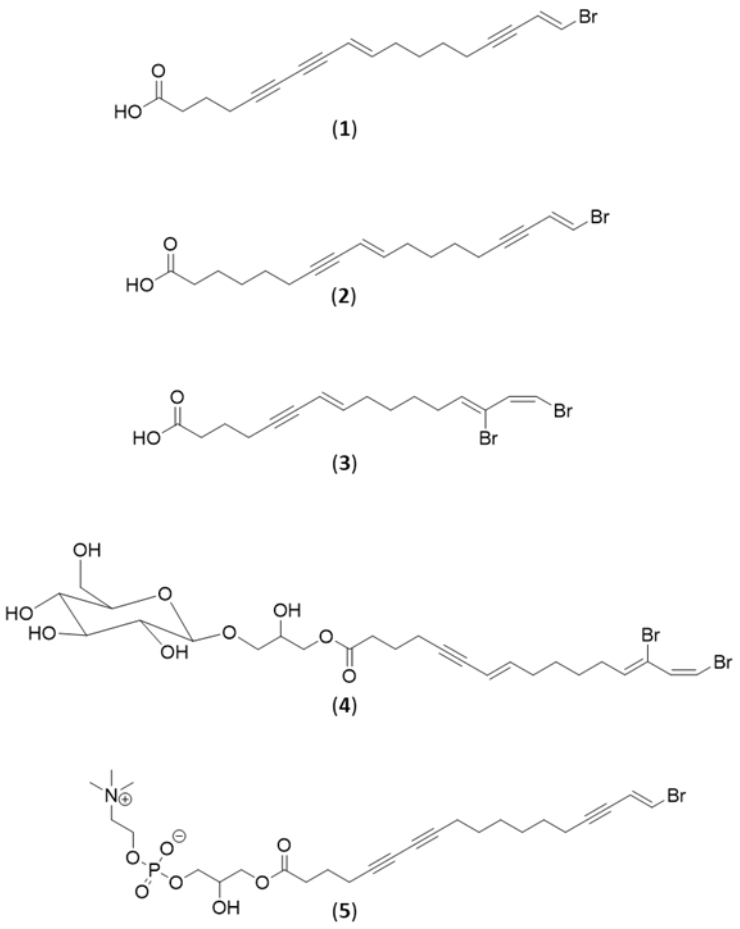

Figure 4.4: Structure of the compounds isolated from the giant barrel sponge (*Xestospongia* spp.). (9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid (**1**), Xestospongic acid (**2**), (7*E*,13*E*,15*Z*)-14,16-dibromohexadeca-7,13,15-trien-5-ynoic acid (**3**), compound (**4**) and compound (**5**)

Compound **3** was isolated from a Curaçao sample as a white powder. Its (+)-HRESIMS spectrum showed the proton and sodium adduct $[M+H]^+$ and $[M+Na]^+$ ions at *m/z* 402.9904, 404.9884, and 406.9867, and 424.9727, 426.9706, and 428.9685, respectively, both having a relative intensity of 1:2:1. This isotopic pattern indicates the presence of two bromine atoms in the molecule. The molecular formula was deduced to be $C_{16}H_{20}Br_2O_2$. The ^{1}H-NMR (CH_3OH-d_4, 600

MHz) spectroscopic data, showed the presence of three double bonds in the molecule and the ^{13}C-NMR (CH_3OH-d_4, 150 MHz) showed one carboxylic acid carbon and one triple bond. The molecular weight together with the NMR signal in the Dictionary of Natural Products led to the identification of the compound as (7*E*,13*E*,15*Z*)-14,16-dibromohexadeca-7,13,15-trien-5-ynoic acid. This compound had been previously reported from *X. muta* collected in Summerland Key, Florida, USA and in Portobelo Bay, Panama (Schmitz and Gopichand 1978; Villegas-Plazas et al. 2019).

Table 4.1: NMR spectroscopic data for compounds **4** and **5**

	Compound 4		Compound 5	
Position	^{13}C-NMR δ, type	^{1}H-NMR δ, (*J* in Hz)	^{13}C-NMR δ, type	^{1}H-NMR δ, (*J* in Hz)
1’	174.9, C	---	174.6, C	---
2’	33.8, CH_2	2.48 t (7.3)	33.7, CH_2	2.49 t (7.6)
3’	25.2, CH_2	1.82 m	24.8, CH_2	1.83 quint (7.2)
4’	19.4, CH_2	2.35 td (7.0,1.8)	19.2, CH_2	2.35 t (7.0)
5’	88.1, C	---	76.8, C	---
6’	80.9, C	---	66.3, C	---
7’	111.3, CH	5.45 dm (15.8)	66.3, C	---
8’	144.0, CH	5.99 dt (15.8, 7.1)	78.2, C[b]	---
9’	33.6, CH_2	2.08 m	19.6, CH_2	2.28 m
10’	29.5, CH_2	1.40m	29.3, CH_2	1.43 m
11’	29.1, CH_2	1.44 m	29.3, CH_2	1.54 m
12’	32.0, CH_2	2.04 m	29.3, CH_2	1.54 m
13’	137.4, CH	6.07 td (7.7,1.5)	29.4, CH_2	1.43 m
14’	114.8, CH	----	19.9, CH_2	2.30 m
15’	132.3, CH	6.78 dm (7.6)	78.3, C[b]	---
16’	113.4, CH	6.56 d (7.6)	93.8, C	---
17’	---	---	119.2, CH	6.24 dt (14.0, 2.3)
18’	---	---	117.9, CH	6.70 d (14.0)
1	66.7, CH_2	4.17 m	67.8, CH_2	3.92 m
2	69.6, CH	4.00 m	69.8, CH	3.99 m
3	71.9, CH_2	3.92 dd (10.5, 5.2), 3.66 m	66.3, CH_2	4.21 dd (11.4, 4.5), 4.14 dd (11.4, 6.2)
1’’	104.7, CH	4.28 d (7.8)	60.4, CH_2	4.31 m
2’’	75.1, CH	3.21 m	67.0, CH_2	3.66 m
3’’	77.9, CH	3.36 bs	---	---
4’’	71.6, CH	3.29 m	---	---
5’’	78.0, CH	3.28 bs	---	---
6’’	62.7, CH_2	3.87 dd (12.1, 1.8) 3.67 m	---	---
N-Me	---	---	54.7, CH_3	3.24 s

[a] NMR spectra were recorded in CH_3OH-d_4, ^{1}H 600 MHz, ^{13}C 150 MHz. [b] these carbons are interchangeable

Compound **4** was also isolated from a Curaçao sample as a white powder. Its (+)-HRESIMS spectrum showed the proton and sodium adducts $[M+H]^+$ and $[M+Na]^+$ ions at *m/z* 639.0777, 641.0761, and 643.0749 and 661.0580, 663.0581, and 665.0563, respectively, both sets of ions with a relative intensity of 1:2:1. This isotopic pattern indicates the presence of two bromine atoms in the molecule. The molecular formula was deduced to be $C_{25}H_{36}Br_2O_9$, which requires 7 degrees of unsaturation. The ^{1}H-NMR and ^{13}C-NMR spectroscopic data (Table 4.1) and heteronuclear single quantum correlation (HSQC) spectrum revealed 10 methylene (δ_H/δ_C 1.40/29.5, 1.44/29.1, 1.82/25.2, 2.04/32.0, 2.08/33.6, 2.35/19.4, 2.48/33.8, 3.67-3.87/62.7, 3.92-3.66/71.9 4.17/66.7), six methine (δ_H/δ_C 3.21/75.1, 3.28/78.0, 3.29/71.6, 3.36/77.9, 4.00/69.6, 4.28/104.7) and five olefinic protons (δ_H/δ_C 5.45/111.3, 5.99/144.0, 6.07/137.4, 6.56/113.4, 6.78/132.3). The signal at 104.7 ppm is very characteristic for a carbon atom joined to two oxygen atoms, which indicates the presence of a sugar moiety in the molecule. In addition, the ^{13}C-NMR spectrum showed four nonprotonated carbons, consisting of one ester carbonyl (δ_C 174.9), two *sp* carbons (δ_C 80.9, 88.1) and one olefinic carbon (δ_C 114.8). The presence of aliphatic signals together with a carbonyl and *sp* and sp^2 carbons indicates that the structure contains an unsaturated fatty acid moiety. Two of the olefinic carbons are shifted upfield, indicating the presence of a substituent that increases the protection over those carbons. This is in agreement with the presence of two bromine atoms observed in the mass spectra and with the lack of any terminal methyl or methylene groups. It was thus possible to establish the attachment of bromine atoms to terminal olefinic carbons at δ_c 114.8 and δ_c 113.4.

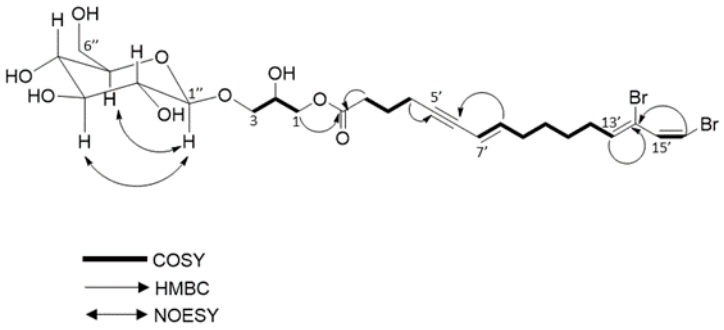

Figure 4.5: Important COSY, HMBC and NOE correlations of compound **4**

Further examination of HMBC and COSY correlations allowed us to establish the full structure of compound **4** (Figure 4.5) as consisting of three moieties: a dibrominated unsaturated fatty acid, a glycerol molecule and a sugar moiety. The brominated fatty acid and the sugar are attached to C1 and C3 of the glycerol molecule respectively. The chemical shift of δ_c 104.7 was

assigned to the anomeric carbon of the sugar, which is attached to C3 of the glycerol moiety through a glycosidic bond. Additionally, NOESY showed correlations between the anomeric proton and those in positions 3'' and 5''. This correlation together with the coupling constants of the anomeric proton (*J* = 7.8 Hz) and protons 3''and 4'' (*J* > 8 Hz) obtained from *J*-Resolved spectra allowed the identification of the sugar moiety as β-glucose. This was also supported by reported ^{13}C-NMR chemical shifts of β-glucose moiety in similar analogues (Fan 1996; Wicke et al. 2000). The identical chemical shift and coupling constants of H-13' indicated that the double bond in position 13'would have the same configuration as that of compound 3 Lastly, the double bond in position 7' was confirmed to have an E configuration with its characteristic coupling constant (*J* = 15.8 Hz), while the terminal double bond was found to have a Z configuration with the coupling constant (*J* = 7.6 Hz) (Schmitz and Gopichand 1978; Villegas-Plazas et al. 2019).

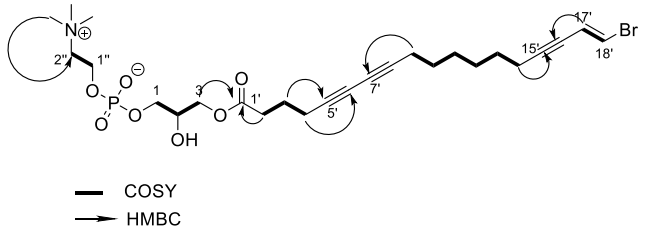

Figure 4.6: COSY and HMBC important correlations of compound **5**

Compound **5** was isolated from a sample from Taiwan as a white powder. The (+)-HRESIMS spectrum of **5** showed the proton adduct $[M+H]^+$ ions at *m/z* 588.1718 and 590.1702. The ions have a relative intensity of 1:1, indicating the presence of a bromine atom in the molecule. The molecular formula was deduced to be $C_{26}H_{39}BrNO_7P$. The ^{1}H-NMR and APT^{13}C-NMR spectroscopic data (Table 4.1) and HSQC correlation revealed the presence of three overlapping methyl groups joined to a nitrogen atom (δ_H/δ_C 3.24/54.7 x 3), 13 methylene (δ_H/δ_C 1.43/29.4, 1.43/29.3, 1.54/29.3 x 2, 1.83/24.8, 2.28/19.6, 2.30/19.9, 2.35/19.2, 2.49/33.7, 3.66/67.0, 3.92/67.8, 4.14-4.21/66.3, 4.31/60.4), one methyne (δ_H/δ_C 3.99/69.8), two olefinic protons (δ_H/δ_C 6.24/119.2, 6.70/117.9), and seven carbons with no protons attached, consisting of one carbonyl ester (δ_C 174.6) and six *sp* carbons (δ_C 66.3 x 2, 76.8, 78.2, 78.3, 93.8). The *sp* carbons indicate the presence of three triple bonds in the molecule. However, some of the δ_C are shifted to downfield, suggesting that two of the triple bonds are conjugated. As for compound **5,** the lack of a terminal methyl or methylene along with the low chemical shift of the olefinic carbon indicates the presence of a terminal olefinic bond attached to a bromine atom. Further examination of HMBC and COSY correlations established

the structure of **5** (Figure 4.6) as consisting of three moieties: a brominated fatty acid, a molecule of glycerol and a molecule of phosphatidylcholine. The presence of a phosphate group can be deduced from analysis of the exact mass of the molecule.

All of the isolated compounds contained one or more triple bonds in their structures, thus they are classified as polyacetylenes. This kind of compound has been reported in a wide range of marine organisms such as algae, corals, mollusks and sponges. In the case of sponges, the genera *Petrosia*, *Callyspongia* and *Xestospongia* are the main sources of polyacetylene compounds, and in some cases they have even been considered to be a chemotaxonomic marker of these genera (Zhou et al. 2015). Although the biosynthetic pathway and ecological function of this kind of compound are still unclear, they have shown a wide range of biological activities. In this study, all the isolated compounds exhibited mild activity against *S. aureus* (**1**: 64 µg/mL, **2** 256 µg/mL, **3** 64 µg/mL, **4** 64 µg/mL, **5** 128 µg/mL). Thus, the inconsistency in the relationship between the activity and location in which the sponges were collected can be explained by the fact that the compounds responsible for the activity might differ in their concentration or their structure in each location.

A comparison of the occurrence of the isolated compounds between the locations showed different patterns for each compound (Figure 4.7). Interestingly, compound **2**, isolated from a sample collected in Martinique, was more abundant in samples from the other three locations. This compound has been previously isolated from *Xestospongia* spp. samples collected in Australia(Quinn and Tucker 1985), the Red Sea(Hirsh et al. 1987), and Mayotte in the coast of Africa (Bourguet-Kondracki et al. 1992). The occurrence of **2** in *Xestospongia* spp. samples collected all over the world indicates that although it is a constitutive metabolite of sponges of the genus *Xestospongia,* the environmental factors prevalent in each location may affect the amount in which this metabolite is produced.

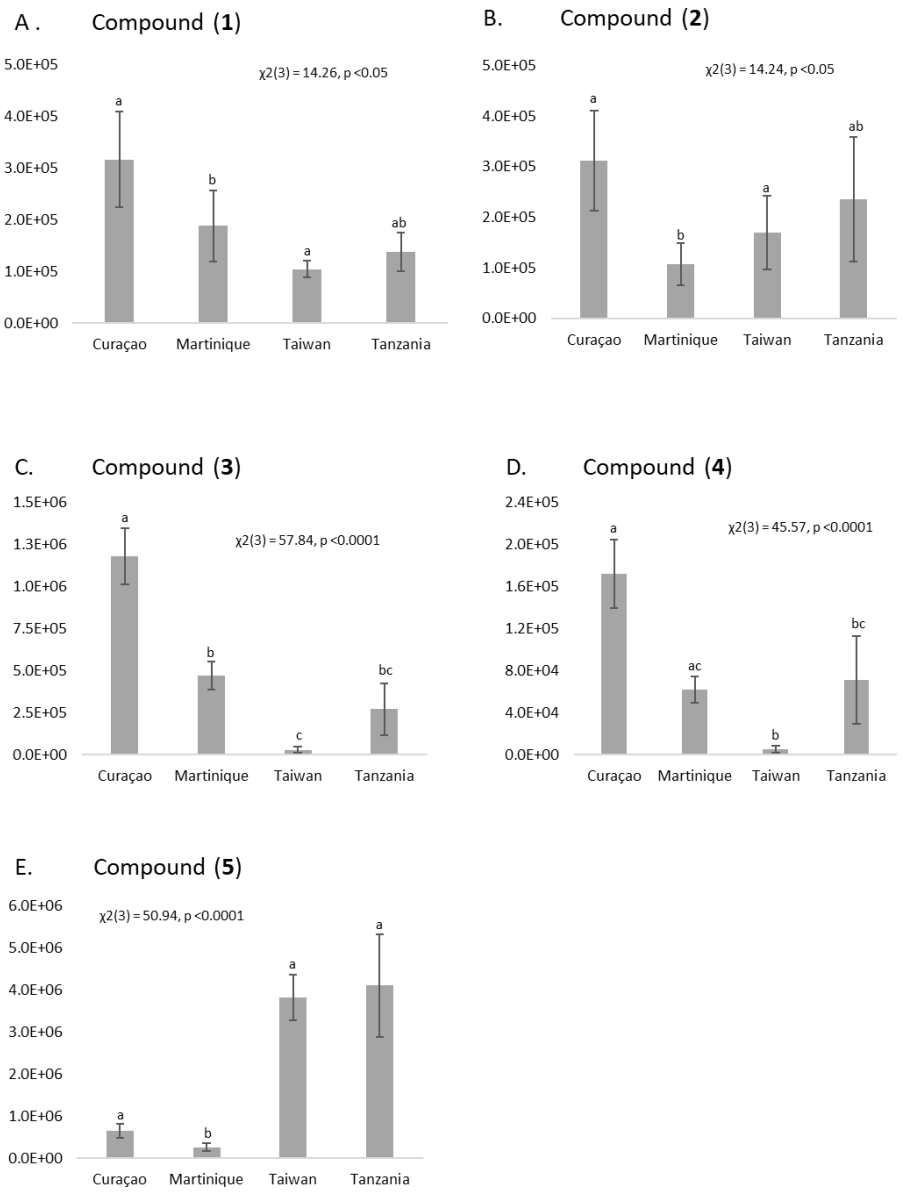

Figure 4.7: Intensity of the buckets of the most intense peak of the mass spectra for compound **1**-**5** in each location. Error bars indicate the standard error. Results of a Kruskal-Wallis Test are shown in each graph. Different letters indicate significant differences in the Post-Hoc test.

Compounds **3** and **4** were more abundant in samples from the Caribbean region, mainly Curaçao. Both compounds have two atoms of bromine in their structures that distinguish them from the other compounds isolated in this study. Compound **3** has been previously isolated

from a sample collected in Florida, USA, and although compound **4** as such, has not been previously reported, but its fatty acid moiety corresponds to compound **3**. Moreover, compound has been previously isolated from a sample collected in Florida and Panamá (Schmitz and Gopichand 1978; Villegas-Plazas et al. 2019). This suggests that these compounds occur higher quantities in sponges located in the Caribbean region and this fact might be used to distinguish the sponges from this region. Lastly, compound **5** is a phospholipid from the phosphatidylcholine group. These compounds are known to be part of the cellular membrane in animals, having not only structural functions but also playing a role in the signaling of metabolic pathways (D'Arrigo and Servi 2010). The variability observed in the amount of **5,** which is more abundant in samples from Taiwan and Tanzania than samples from the Caribbean, suggests that, similarly to what occurs in animal cell membranes, this compound also has more than just a structural role in *Xestospongia* spp. and its production is thus conditioned by the environmental factors related to each location.

3. Experimental section

3.1 Sample collection and extraction

Xestospongia spp. samples were collected in Martinique, Curaçao, Tanzania, and Taiwan and stored in ethanol at -20°C (appendix 1 Table S1). Samples were transported to the Institute of Biology of Leiden University for further analysis. The *Xestospongia* samples were ground and extracted with ethanol and sonicated for 20 min. The extraction was done in triplicate. An aliquot of 1 mL of each extract was dried and used for ^{1}H-NMR analysis. The remaining extracts were dried. The salt from the extracts was removed using C-18 SPE Supelco Supelclean LC-18, (Merck, Darmstadt, Germany) cartridges. For each extract, 50 mg were loaded into the cartridge and eluted with solvents of decreasing polarity, i.e., H_2O (F1), MeOH (F2), and MeOH/DCM (1:1) (F3). The methanol fraction (F2) was used for further LC-MS analysis.

3.2 ^{1}H-NMR Analysis and data preprocessing

The dry extract was resuspended in 1mL of deuterated methanol (CH_3OH-d_4) with hexamethyl disiloxane (HMDSO) as the internal standard. The ^{1}H-NMR spectra were measured at 25 °C in an AV-600 MHz NMR spectrometer (Bruker, Karlsruhe, Germany), operating at the ^{1}H-NMR frequency of 600.13 MHz, and equipped with a TCI cryoprobe and Z gradient system. For internal locking, CH_3OH-d_4 was used. A presaturation sequence was used to suppress the residual water signal, using low power selective irradiation at the H_2O frequency during the recycle delay.

The resulting spectra were phased, baseline corrected, and calibrated to HMDSO at 0.07 ppm using TOPSPIN V. 3.0 (Bruker, Karlsruhe, Germany). The NMR spectra were bucketed using AMIX 3.9.12 (Bruker BioSpin GmbH, Rheinstetten, Germany). Bucket data were obtained by spectra integration at 0.04 ppm intervals from 0.20 to 10.02 ppm. The peak intensity of individual peaks was scaled to the total intensity of the buckets. The regions between 3.32 and 3.28, 4.9 and 4.8, 3.62 and 3.57, and 1.15 and 1.19 ppm were excluded from the analysis because they correspond to solvent residual signals.

3.3 LC-MS analysis and data processing

The methanol fractions obtained from the SPE were dried, and 1 mg was dissolved in ACN/H_2O 1:1 to obtain solutions with a final concentration of 1mg/mL. The fractions were analyzed using an UHPLC-DAD-MS, Thermo Scientific (Dreieich, Germany) UltiMate 3000 system coupled to a Bruker (Bremen, Germany) OTOF-Q II spectrometer with electrospray ionization (ESI). The UHPLC separation was performed on a Phenomenex (Utrecht, The Netherlands), Kinetex, C18 (2.1 x 150 mm, 2.6 µm) using a two-step gradient of 0.1% formic acid in H_2O (A) and 0.1% formic acid in ACN (B), starting at 45% B to 60% in 15 minutes, 60% to 90% in 12.5 min and 90% to 98% B in 2.5 minutes. The flow rate was 0.300 mL/min, and the column temperature was maintained at 40°C. The injection volume was set at 1 µL. The mass spectrometer parameters were set as follow: nebulizer gas 2.0 bar, drying gas 10.0 mL/min, temperature 250°C, capillary voltage 3500 V. The mass spectrometer was operated in positive mode with a scan range of 100 - 1650 m/z and sodium formate was used as a calibrant.

The resulting chromatogram was processed to obtain a matrix for further analysis using Brucker Daltonics Profile Analysis (version 2.1, Bremen, Germany). The spectra were divided into buckets of 1 minute between 1 and 30 minutes and 1 m/z between 100 and 1450 m/z. The buckets were organized in a matrix, and data were filtered to remove those buckets that presented a %CV above 20% in the quality control samples.

3.4 Statistical analysis

The matrixes obtained from the NMR and LC-MS were used to perform multivariate data analysis using SIMCA-P software (v.15.0.2, Umetrics, Umeå, Sweden). Principal component analysis PCA, discriminant analysis of partial least square PLS-DA, and orthogonal partial least square OPLS-DA were performed. For the analysis, data were scaled using united variance scaling (NMR) and pareto scaling (LC-MS), and the models were tested using a permutation test and a cross-validation ANOVA (CV-ANOVA) test. The model was considered valid if CV-

ANOVA showed $p < 0.05$. For the prediction power of the model, Q^2 values above 0.4 were required; otherwise, the model was considered valid but with no prediction power.

A heatmap was created using a data matrix with the top 40 signals of the VIP plot. This matrix was uploaded on the Metaboanalyst R2.0 Web site (http://www.metaboanalyst.ca) (Chong et al. 2019). The dendrogram was obtained by hierarchical cluster analysis using the Euclidean distance and the "Ward" algorithm.

To test if the concentration of each compound differed among locations, the intensity of buckets corresponding to the most intense ion observed in its mass spectra was used. Data were analyzed with IBM SPSS Statistics Version 22 (Armonk, NY, USA) using a Kruskal-Wallis Test. Location was used as a factor, and the number of samples was 139. For the compounds that appeared to be significantly different between locations, a Post-Hoc test was done, and the P values were Bonferroni-corrected.

3.5 Isolation and elucidation

For the isolation of active compounds, extracts of samples from Martinique, Curaçao and Taiwan were prepared as mentioned in section 1. The crude extracts were fractionated using an SPE 20 mL LC-18 Supelco Supelclean (Merck, Darmstadt, Germany) cartridge and eluted using two different methods according to the sample location. The sample from Martinique (1.0 g) was eluted with 100 mL of H_2O, MeOH, and MeOH/DCM (1:1), yielding three fractions: FM1, FM2, and FM3, respectively. The samples from Curaçao (1.9 g) and Taiwan (2.5 g) were eluted using 50 mL of each of the following solvents: 100% H_2O; H_2O/MeOH 8:2, H_2O/MeOH 6:4, H_2O/MeOH 4:6, H_2O/MeOH 2:8, 100% MeOH and MeOH/DCM 1:1. This resulted in fractions of each extract FC1–FC7 for Curaçao samples and FT1–FT7 for Taiwan samples, respectively.

Fraction FM2 (212 mg) was submitted to a size-based separation. The fraction was resuspended in 10 mL of MeOH and injected into a Sepacore flash system (Büchi, Hendrik-Ido-Ambacht, The Netherlands) with a Sephadex LH-20 (Merck KGaA, Darmstadt, Germany) column and a sample loop of 20 mL. Samples were eluted at a flow rate of 2.5 mL/min with MeOH. Fractions were collected automatically every minute and combined into nine FM2.1-FM2.9 fractions based on their TLC profiles. The purification of fractions FM2.7, FM2.9, FC4 and FT4 was performed using an Agilent (Santa Clara, CA, USA) 1200 series system on a Phenomenex (Utrecht, The Netherlands) Luna 5 µm, C-18, 250 mm x 10 mm column and eluted at a flow rate of 3.50 mL/min with different gradients of 0.1% formic acid in H_2O (A) and 0.1% formic acid in MeOH (B). The fractions FM.2.7-FM2.9 (49.87mg) were eluted with a

gradient of 75% B to 80% B in 25 min, 15 min of 80% B, 80% to 100% B in 2 min and 100% B for 5 min. This yielded 2.26 mg of **1** and 1.58 mg of **2.** The fraction FC4 (70.6 mg) was eluted using the following gradient: 72% to 85% B in 52 min, 85% to 100% B in 2 min and 100% B for 10 min. This allowed the isolation of **3** (7.16 mg) and **4** (2.14 mg). The fraction FT5 (94.08 mg) was eluted using the following gradient: 72% to 80% B in 34 min, 80% to 85% B in 16 min, 85% to 100% B in 3 min and 100% B for 3 min. This led to compounds **5** (1.51 mg) and **1**(1.00 mg).

(9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid (**1**)

White amorphous powder;^{1}H-NMR (CH_3OH-d_4, 600 MHz) δ_H 1.52 m, 1.82 quint (*J* = 7.3 Hz), 2.18 m, 2.29 m, 2.31 m, 2.38 t (*J* = 7.0 Hz), 5.55 dm (*J* = 15.9 Hz), 6.23 m, 6.27 m, 6.70 d (*J* = 14.0 Hz). ^{13}C-NMR (CH_3OH-d_4, 150 MHz) δ_c 19.5, 19.6, 25.8, 28.6, 28.6, 33.2, 36.9, 65.9, 73.3, 73.9, 80.2, 83.3, 92.8, 109.8, 117.7, 118.9 and 148.3. HRESIMS *m/z* $[M+H]^+$ 347.0646 and 349.0631 (Calcd for $C_{18}H_{20}BrO_2^+$, 347.0647 and 349.0626).

Xestospongic acid (**2**)

White amorphous powder;^{1}H-NMR (CH_3OH-d_4, 600 MHz) δ_H 1.43 m, 1.49 m, 1.51 m, 1.63 quint (*J* = 7.5 Hz), 2.09 m, 2.24 m, 2.26 m, 2.28 m, 5.45 dm (*J* = 15.8 Hz),5.96 dt (*J* = 15.8, 7.2 Hz), 6.22 dt (*J* = 14.0, 2.3 Hz), 6.68 d (*J* = 14.0 Hz). ^{13}C-NMR (CH_3OH-d_4, 150 MHz) δ_C 19.8 x 2, 26.3, 28.9, 29.2, 29.7, 29.8, 33.3, 36.6, 78.3, 80.2, 89.2, 93.6, 111.6, 117.9, 119.2, and 143.3. HRESIMS *m/z* $[M+H]^+$ 351.0956 and 353.0939 (Calcd for $C_{18}H_{24}BrO_2^+$, 351.0960 and 353.0939).

(7*E*,13*E*,15*Z*)-14,16-dibromohexadeca-7,13,15-trien-5-ynoic acid (**3**)

White amorphous powder;^{1}H-NMR (CH_3OH-d_4, 600 MHz) δ_H,1.40 m, 1.44 m, 1.78 quint (*J* = 7.2 Hz), 2.04 m, 2.08 m, 2.34 dt (*J* = 2.0, 7.0 Hz), 2.40 t (*J* = 7.4 Hz), 5.45 dm (*J* = 15.8 Hz), 5.99 dt (*J* = 15.8, 7.1 Hz), 6.07 td (*J* = 7.7, 1.5 Hz), 6.55 d (*J* = 7.6 Hz), 6.78 dm (*J* = 7.6 Hz). ^{13}C-NMR (CH_3OH-d_4,150 MHz) δ_c 19.4, 25.4, 29.1, 29.5, 32.0, 33.6, 33.9, 80.8, 88.1, 111.3, 113.4, 114.8, 132.3, 137.4, 143.9 and 177.3. HRESIMS *m/z* $[M+H]^+$ 402.9904, 404.9884, 406.9867 (Calcd for $C_{16}H_{21}Br_2O_2^+$, 402.9903, 404.9883, 406.9862) and $[M+Na]^+$ 424.9727, 426.9706, 428.9685 (Calcd for $C_{16}H_{20}Br_2NaO_2^+$, 424.9723, 426.9702, 428.9682).

Compound (**4**)

White amorphous powder;^{1}H-NMR (CH_3OH-d_4, 600 MHz) δ_H in Table 4.1. ^{13}C-NMR (CH_3OH-d_4, 150 MHz) δ_C in Table 4.1. HRESIMS *m/z* $[M+H]^+$ 639.0777, 641.0761, 643.0749 (Calcd for $C_{25}H_{37}Br_2O_9^+$, 639.0804, 641.0784, 643.0763) and , $[M+Na]^+$ 661.0580, 663.0581, 665.0563 (Calcd for $C_{25}H_{36}Br_2NaO_9^+$, 661.0624, 663.0603, 665.0583)

Compound (**5**)

White amorphous powder;^{1}H-NMR (CH_3OH-d_4,600MHz) δ_H in Table 4.1. ^{13}C-NMR (CH_3OH-d_4, 150 MHz) δ_C in Table 4.1. HRESIMS *m/z* $[M+H]^+$ 588.1718, 590.1702 (Calcd for $C_{26}H_{40}BrNO_7P^+$ m/z: 588.1726, 590.1705)

3.6 Antibacterial activity test

Strains used in this study were the Gram-positive bacteria *S. aureus* (CECT976) and Gram-negative bacteria *E. coli* (DH5α,Promega). The strains had been kept at – 80 °C (in 100% glycerol). For their use, the strains were transferred onto Mueller-Hinton agar plates (MHA) (Sigma-Aldrich, Zwijndrecht,The Netherlands) and incubated overnight at 37 °C.

A broth microdilution method was used to determine the minimum inhibitory concentration (MIC) according to the CLSI (Clinical Laboratory Standards Institute) guidelines using 96-wells microtiter plates. The MIC is defined as the lowest concentration of each extract, which completely inhibits bacterial growth. For antimicrobial testing, the extracts were dissolved in 100% DMSO in a concentration of 10 mg/mL. All experiments were performed in triplicate. Ninety microliters of Mueller-Hinton broth (MHB) and 10 μL of the tested extract were added into the first well. Then two-fold serial dilutions of the extracts were prepared by dilution with MHB to achieve a decreasing range of concentrations from 512–16 μg/mL in the microtiter plates. The highest concentration of DMSO after dilution was < 5 %, to avoid affecting the growth of the bacterial strains. From the overnight cultures of the bacterial strains, a single colony was used to inoculate the MHB at 37 °C with agitation (150 rpm). The cultures were then further diluted in MHB and adjusted to a turbidity level of 0.5 McFarland standard solution (approximately 10^6 CFU/mL). Each well was then inoculated with 50 μL of the bacterial solution at a density of 10^6 CFU/mL. Spectinomycin (100 mg/mL) (Sigma-Aldrich) was used as a positive control and 5% dimethyl sulfoxide (DMSO) as a negative control. The inoculated microtiter plates were incubated at 30 °C for 24 h. Bacterial growth was detected by optical density.

Acknowledgment

The authors greatly appreciate the contribution of Dr. Erika G. Wilson in the scientific discussion of this paper. The specimens from Martinique were collected during the Madibenthos expeditions (PI Philippe Bouchet) organised by the Museum National d'Histoire Naturelle (MNHN) and the Marine Protected Areas Agency (AAMP), the Regional Directorate for the Environment (DEAL) and the Martinique Water Bureau (ODE) with funding from the

European Regional Development Fund (ERDF), the Territorial Collectivity of Martinique (CTM) and Saint-James Plantations and BRED Banque populaire. Yusheng Huang of the National Penghu University of Science and Technology, Penghu, Taiwan and Christian Vaterlaus of Marine Cultures, Jambiana Tanzania are thanked for their assistance in the field. This work was supported by the COLCIENCIAS (science technology and innovation ministry, Colombian government) and NWO-VIDI (#16.161.301) and NWO-Aspasia (#105-010.030).

Appendix 1

Table S1: Collection places of *Xestospongia* samples and their geographical information.

Code	Location	Latitude	Longitude
LMB29902	Curaçao	12.10771	-68.94976
LMB30102	Curaçao	12.10771	-68.94976
LMB30302	Curaçao	12.10771	-68.94976
LMB29502	Curaçao	12.10771	-68.94976
LMB31302	Curaçao	12.10771	-68.94976
LMB30902	Curaçao	12.10771	-68.94976
LMB30702	Curaçao	12.10771	-68.94976
LMB30502	Curaçao	12.06500	-68.86027
LMB30402	Curaçao	12.06500	-68.86027
LMB30602	Curaçao	12.06500	-68.86027
LMB30802	Curaçao	12.06500	-68.86027
LMB31002	Curaçao	12.06500	-68.86027
LMB29402	Curaçao	12.12206	-68.96925
LMB31102	Curaçao	12.12206	-68.96925
LMB31202	Curaçao	12.12206	-68.96925
LMB30002	Curaçao	12.12206	-68.96925
LMB29802	Curaçao	12.12206	-68.96925
LMB29702	Curaçao	12.12206	-68.96925
LMB31402	Curaçao	12.32977	-69.15191
LMB29602	Curaçao	12.32977	-69.15191
LMB30202	Curaçao	12.32977	-69.15191
LMB59202	Curaçao	12.39386	-69.15723
LMB59302	Curaçao	12.39386	-69.15723
LMB59402	Curaçao	12.39386	-69.15723
LMB59502	Curaçao	12.39386	-69.15723
LMB59602	Curaçao	11.98617	-68.64665
LMB59702	Curaçao	11.98617	-68.64665
LMB59802	Curaçao	11.98617	-68.64665
LMB59902	Curaçao	11.98617	-68.64665
LMB7202	Martinique	14.53330	-61.08789
LMB7602	Martinique	14.63160	-61.12880
LMB7802	Martinique	14.63160	-61.12880
LMB7402	Martinique	14.44440	-61.03769

Code	Location	Latitude	Longitude
LMB19102	Martinique	14.44440	-61.03769
LMB23702	Martinique	14.44380	-61.04059
LMB7702	Martinique	14.44380	-61.04059
LMB19202	Martinique	14.44380	-61.04059
LMB7102	Martinique	14.44460	-60.89990
LMB7902	Martinique	14.44460	-60.89990
LMB7302	Martinique	14.45510	-60.92529
LMB7502	Martinique	14.45510	-60.92529
LMB25302	Martinique	14.45510	-60.92529
LMB20302	Martinique	14.44800	-60.89970
LMB20602	Martinique	14.44800	-60.89970
LMB18402	Martinique	14.44800	-60.89970
LMB17602	Martinique	14.49630	-60.77170
LMB17402	Martinique	14.44210	-61.03960
LMB19402	Martinique	14.44210	-61.03960
LMB26102	Martinique	14.44210	-61.03960
LMB18902	Martinique	14.44070	-61.02909
LMB24302	Martinique	14.44070	-61.02909
LMB8002	Martinique	14.44070	-61.02909
LMB8102	Martinique	14.44070	-61.02909
LMB24602	Martinique	14.46490	-61.01940
LMB25002	Martinique	14.46490	-61.01940
LMB20002	Martinique	14.46490	-61.01940
LMB19302	Martinique	14.46490	-61.01940
LMB20102	Martinique	14.46490	-61.01940
LMB22602	Martinique	14.46490	-61.01940
LMB23602	Martinique	14.86850	-60.89429
LMB24402	Martinique	14.64620	-60.85079
LMB24002	Martinique	14.91740	-61.14710
LMB22302	Martinique	14.91440	-61.14900
LMB24102	Martinique	14.47600	-61.08590
LMB20202	Martinique	14.47600	-61.08590
LMB17302	Martinique	14.51850	-61.09770
LMB20702	Martinique	14.51850	-61.09770
LMB17702	Martinique	14.84160	-61.22770
LMB18202	Martinique	14.84160	-61.22770

Code	Location	Latitude	Longitude
LMB19802	Martinique	14.84160	-61.22770
LMB24702	Martinique	14.84160	-61.22770
LMB19502	Martinique	14.66950	-61.17509
LMB20802	Martinique	14.65710	-61.15750
LMB17902	Martinique	14.65710	-61.15750
LMB23002	Martinique	14.65710	-61.15750
LMB22402	Martinique	14.65710	-61.15750
LMB24202	Martinique	14.65710	-61.15750
LMB19602	Martinique	14.74070	-61.18029
LMB19902	Martinique	14.86250	-61.20749
LMB20402	Martinique	14.86250	-61.20749
LMB25102	Martinique	14.86250	-61.20749
LMB8902	Martinique	14.57630	-61.05480
LMB26002	Martinique	14.57630	-61.05480
LMB19002	Martinique	14.57630	-61.05480
LMB18702	Martinique	14.45850	-60.96949
LMB25702	Martinique	14.45850	-60.96949
LMB18002	Martinique	14.45850	-60.96949
LMB17502	Martinique	14.46700	-61.03440
LMB20502	Martinique	14.46700	-61.03440
LMB23902	Martinique	14.46700	-61.03440
LMB24802	Martinique	14.46700	-61.03440
LMB19702	Martinique	14.51850	-61.09770
LMB24902	Taiwan	23.5517	119.6412
LMB25202	Taiwan	23.3986	119.3231
LMB4902	Taiwan	23.5725	119.4931
LMB5002	Taiwan	23.5725	119.4931
LMB5102	Taiwan	23.5725	119.4931
LMB5202	Taiwan	23.5517	119.6412
LMB5302	Taiwan	23.5517	119.6412
LMB5402	Taiwan	23.5725	119.4931
LMB5502	Taiwan	23.5517	119.6412
LMB5602	Taiwan	23.3986	119.3231
LMB56102	Taiwan	23.2504	119.6743
LMB56202	Taiwan	23.2504	119.6743
LMB56302	Taiwan	23.5371	119.5443

Code	Location	Latitude	Longitude
LMB56402	Taiwan	23.2442	119.6199
LMB56502	Taiwan	23.2504	119.6743
LMB56602	Taiwan	23.5371	119.5443
LMB56702	Taiwan	23.5371	119.5443
LMB56802	Taiwan	23.5371	119.5443
LMB56902	Taiwan	23.5371	119.5443
LMB57002	Taiwan	23.5371	119.5443
LMB5702	Taiwan	23.3986	119.3231
LMB57102	Taiwan	23.2504	119.6743
LMB5802	Taiwan	23.3986	119.3231
LMB5902	Taiwan	23.5725	119.4931
LMB6002	Taiwan	23.5725	119.4931
LMB6102	Taiwan	23.3986	119.3231
LMB6202	Taiwan	23.5517	119.6412
LMB6302	Taiwan	23.5725	119.4931
LMB6402	Taiwan	23.5517	119.6412
LMB6502	Taiwan	23.5725	119.4931
LMB6602	Taiwan	23.2575	119.6791
LMB6702	Taiwan	23.3986	119.3231
LMB6802	Taiwan	23.5517	119.6412
LMB6902	Taiwan	23.5725	119.4931
LMB7002	Taiwan	23.5517	119.6412
LMB61002	Tanzania	-6.31011	39.58258
LMB61102	Tanzania	-6.31011	39.58258
LMB60302	Tanzania	-6.70895	39.28272
LMB60202	Tanzania	-6.70895	39.28272
LMB60002	Tanzania	-6.70895	39.28272
LMB60602	Tanzania	-6.70895	39.28272
LMB60402	Tanzania	-6.70895	39.28272
LMB60902	Tanzania	-6.70895	39.28272
LMB60102	Tanzania	-6.31011	39.58258
LMB60802	Tanzania	-6.31011	39.58258
LMB60702	Tanzania	-6.31011	39.58258
LMB60502	Tanzania	-6.31011	39.58258

Table S2: List of mass features selected from S-plot of Orthogonal partial least square – discriminant analysis (OPSD-DA) with two classes (active and non-active samples).

m/z	P1 value
749.5	0.0570212
405.5	0.0562273
404.5	0.0546859
425.5	0.0531678
597.5	0.0530375
385.5	0.0521159
642.5	0.0504513
403.5	0.049805
323.5	0.0489471
238.5	0.0483625
370.5	0.0475088
347.5	0.0474848
378.5	0.0472359
296.5	0.0471683
324.5	0.0470832
231.5	0.0470479
377.5	0.0468182
375.5	0.0467828
295.5	0.0467675
348.5	0.0461731
369.5	0.045865
218.5	0.0457138
366.5	0.0456151
351.5	0.0454663
240.5	0.045461
184.5	0.0452214
226.5	0.0445743

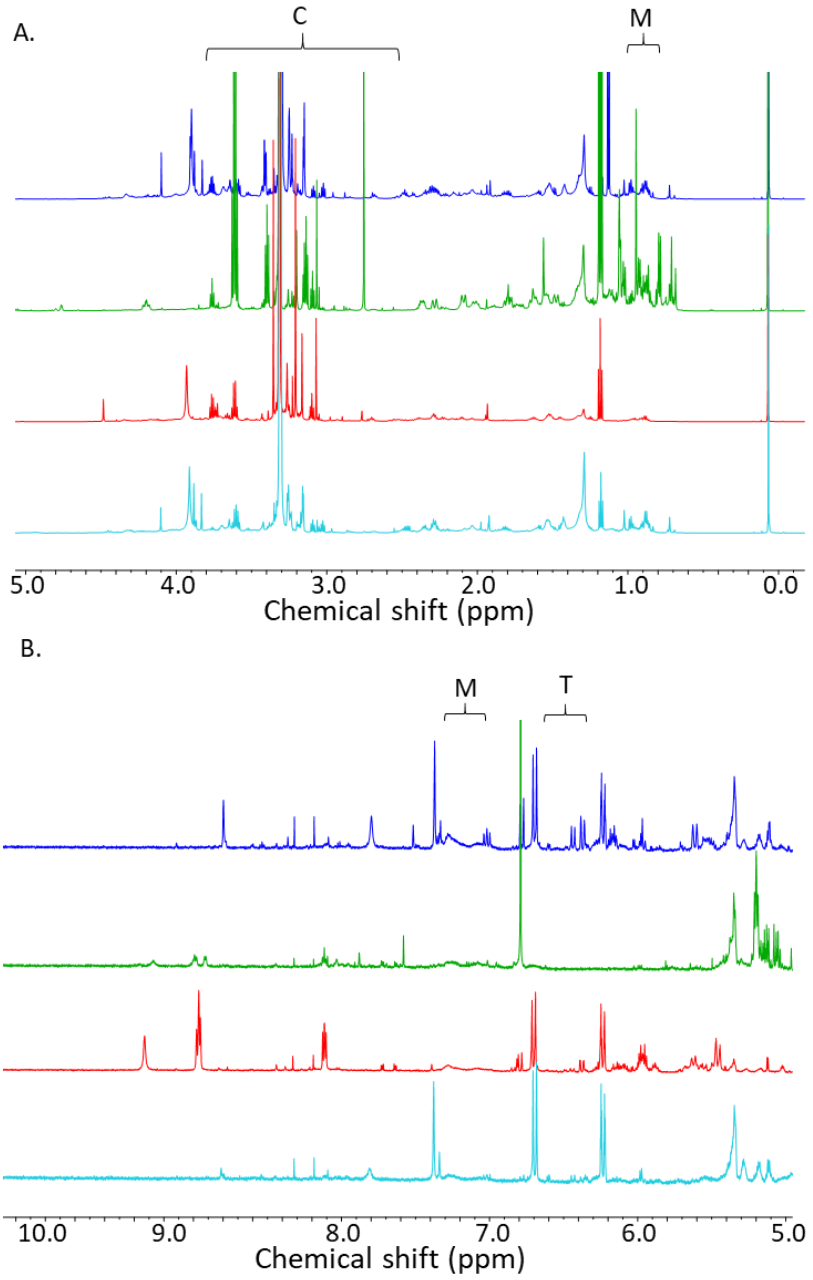

Figure S1: Typical ^{1}H-NMR spectra profile (CH_3OH-d_4, 600 MHz) of *Xestospongia* samples from four locations: Martinique (Green), Curaçao (Red), Taiwan (Dark blue), and Tanzania (Light blue) divided in two regions A: from δ_H 0 - 5 and B: from δ_H 5 - 10. The characteristic ^{1}H-NMR signal ranges of Martinique (M), Curaçao (C), and Taiwan (T) are shown in each spectrum.

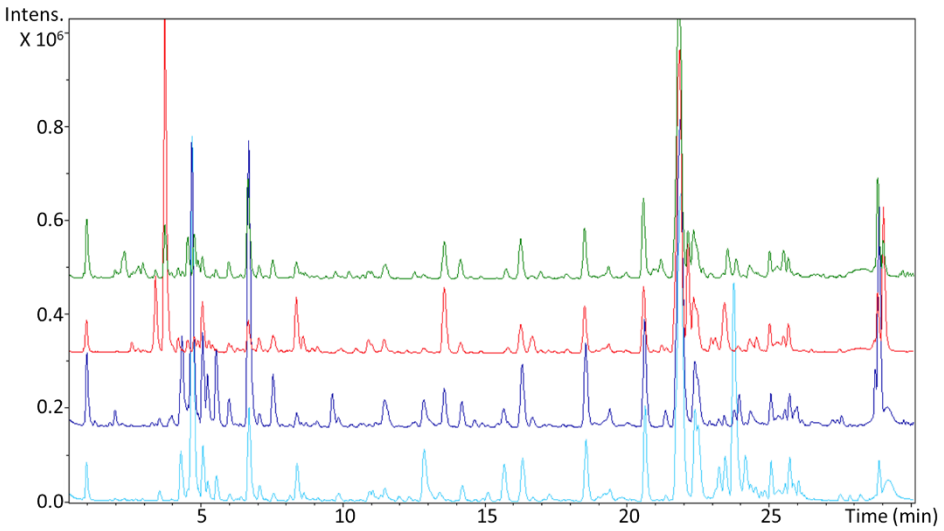

Figure S2: Typical LC-MS chromatographic profiles of *Xestospongia* samples from the four locations: Martinique (Green), Curaçao (Red), Taiwan (Dark blue), and Tanzania (Light blue).

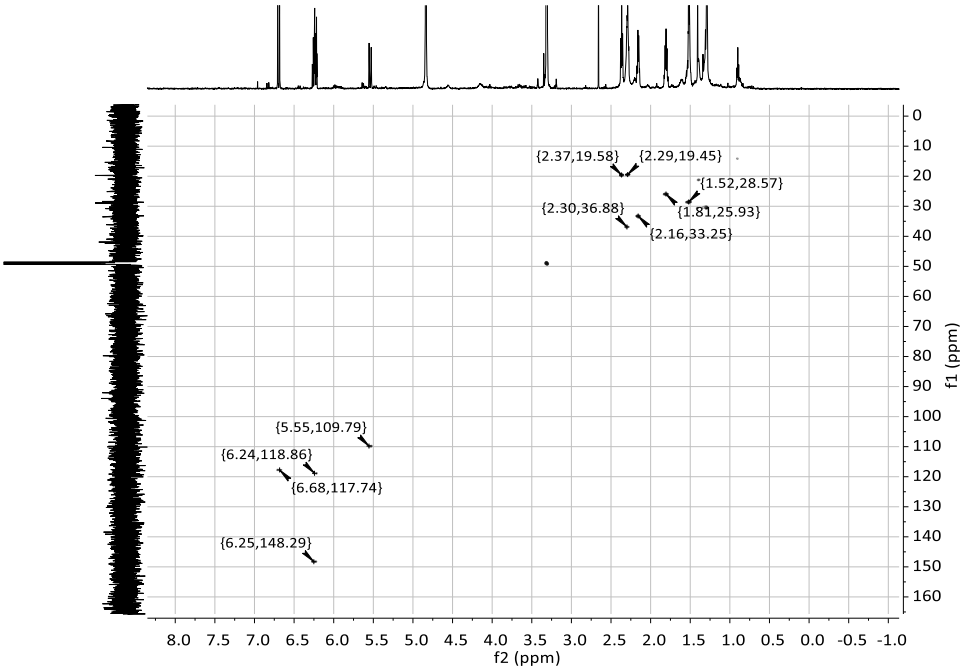

Figure S3: Heteronuclear single quantum coherence (HSQC) spectrum (CH_3OH-d_4, 600 MHz) of (9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid (**1**).

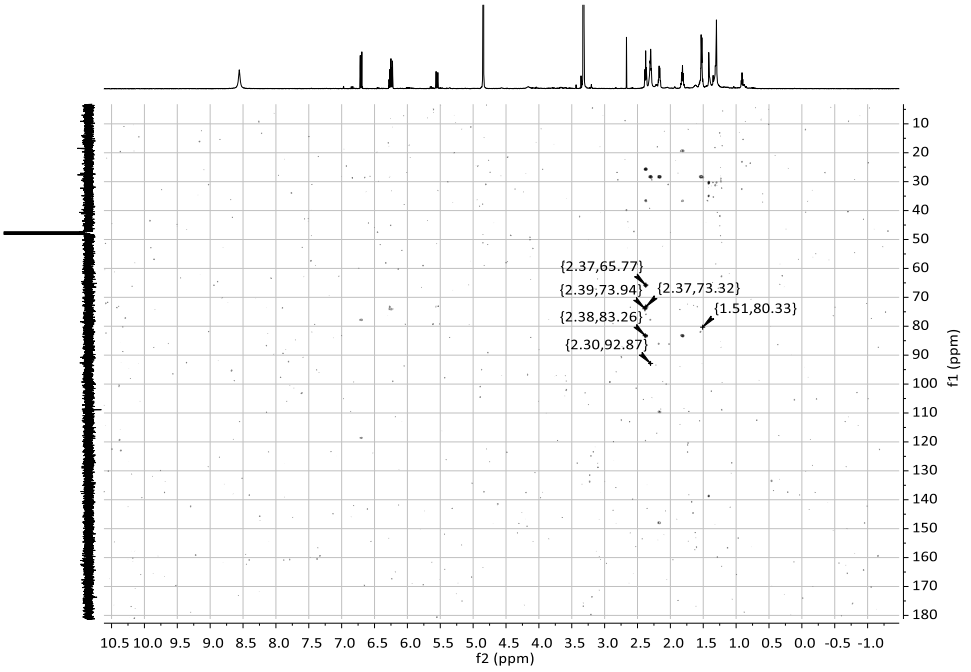

Figure S4: Heteronuclear multiple bond correlation (HMBC) spectrum (CH_3OH-d_4, 600 MHz) of (9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid (**1**).

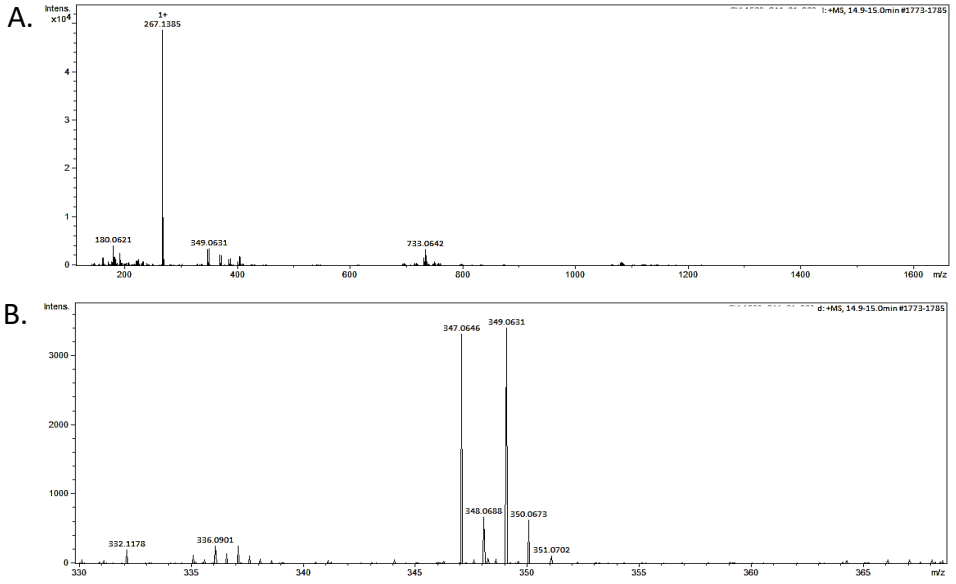

Figure S5: Electrospray ionization-quadrupole-time of flight (ESI-qTOF) MS spectrum in full range (A) and the expanded region around the $[M+H]^+$ ion (B) of (9*E*,17*E*)-18-bromooctadeca-9,17-dien-5,7,15-triynoic acid (**1**).

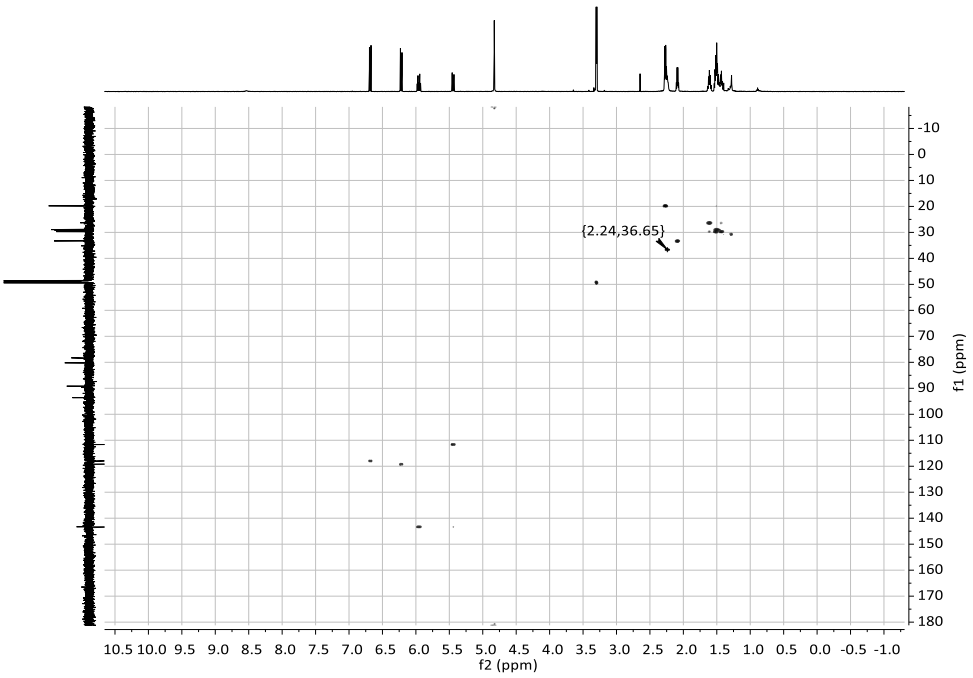

Figure S6: Heteronuclear single quantum coherence (HSQC) spectrum (CH_3OH-d_4, 600 MHz) of Xestospongic acid (**2**).

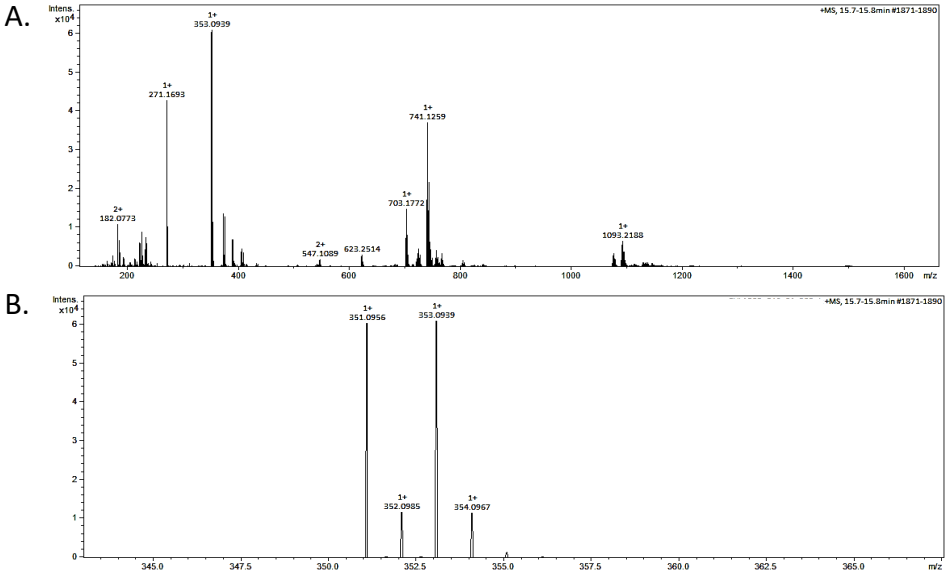

Figure S7: Electrospray ionization-quadrupole-time of flight (ESI-qTOF) MS spectrum in full range (A) and the expanded region around the $[M+H]^+$ ion (B) of Xestospongic acid (**2**).

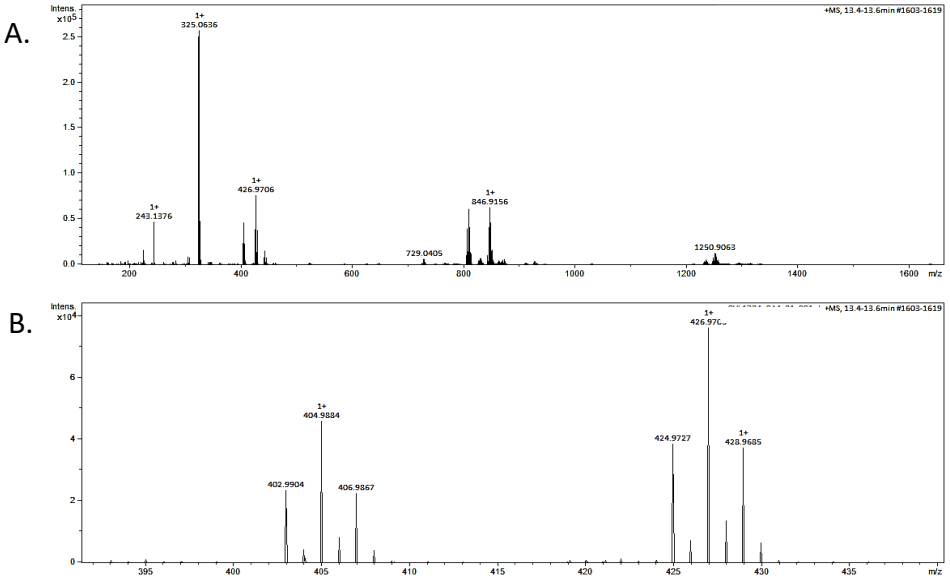

Figure S8: Electrospray ionization-quadrupole-time of flight (ESI-qTOF) MS spectrum in full range (A) and the expanded region around the $[M+H]^+$ and $[M+Na]^+$ ions (B) of (7*E*,13*E*,15*Z*)-14,16-dibromohexadeca-7,13,15-trien-5-ynoic acid (**3**).

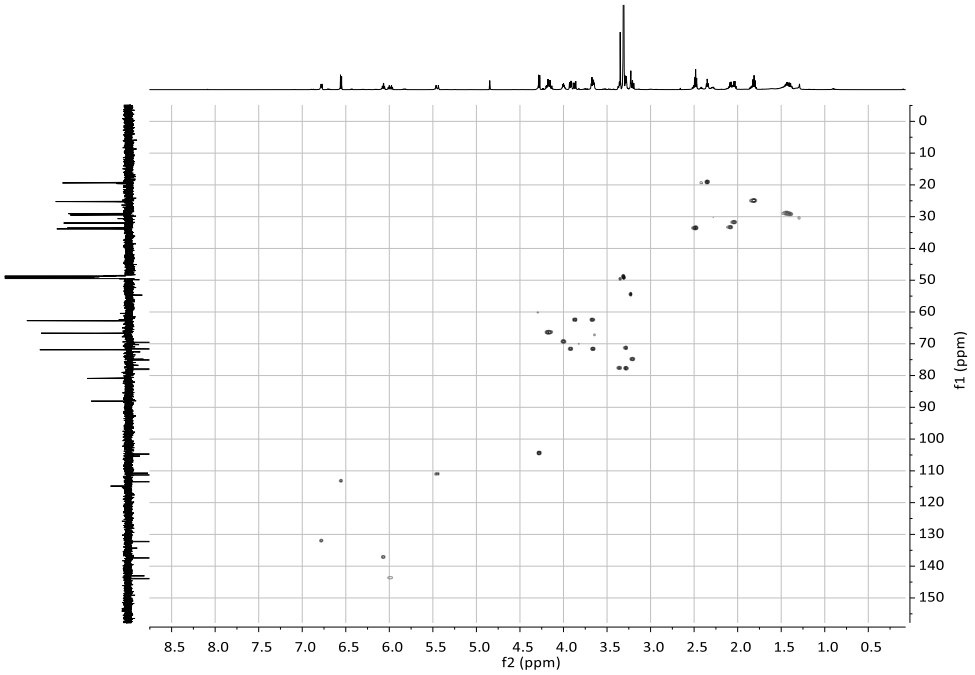

Figure S9: Heteronuclear single quantum coherence (HSQC) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**4**).

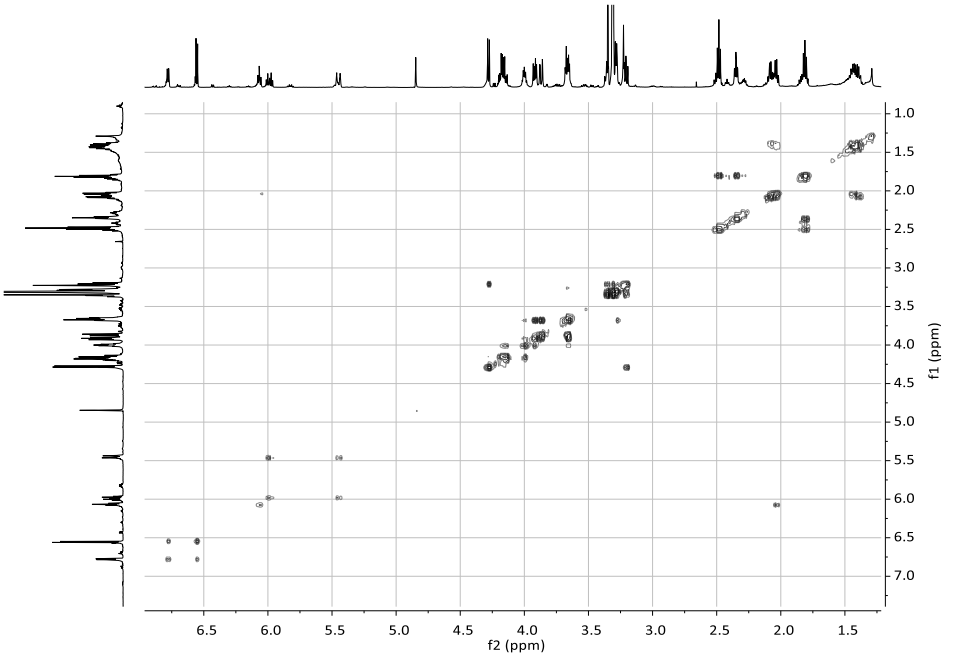

Figure S10: ^{1}H-^{1}H Correlation spectroscopy (COSY) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**4**).

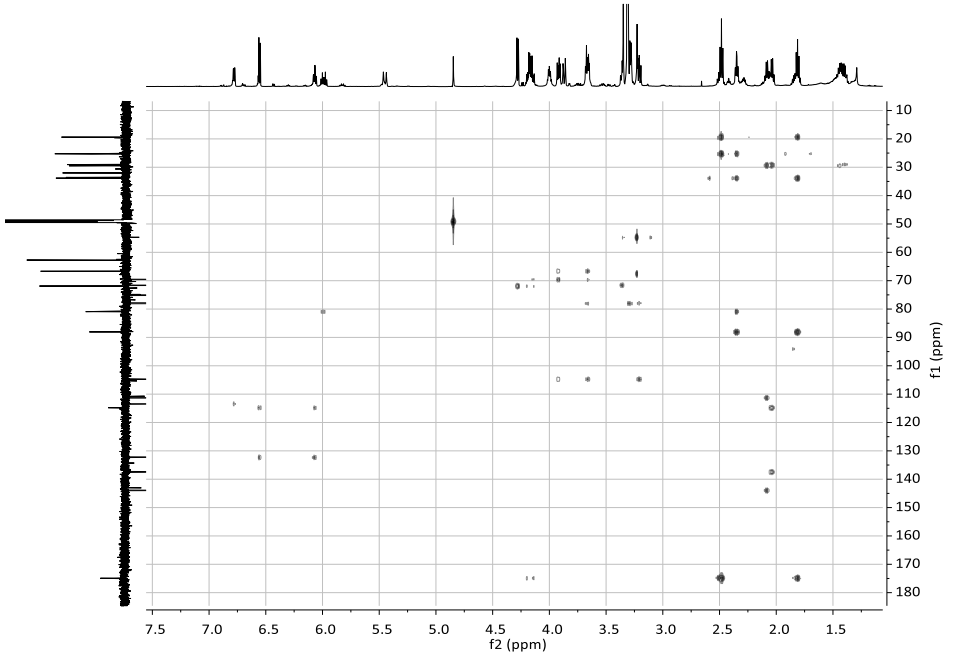

Figure S11: Heteronuclear multiple bond correlation (HMBC) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**4**).

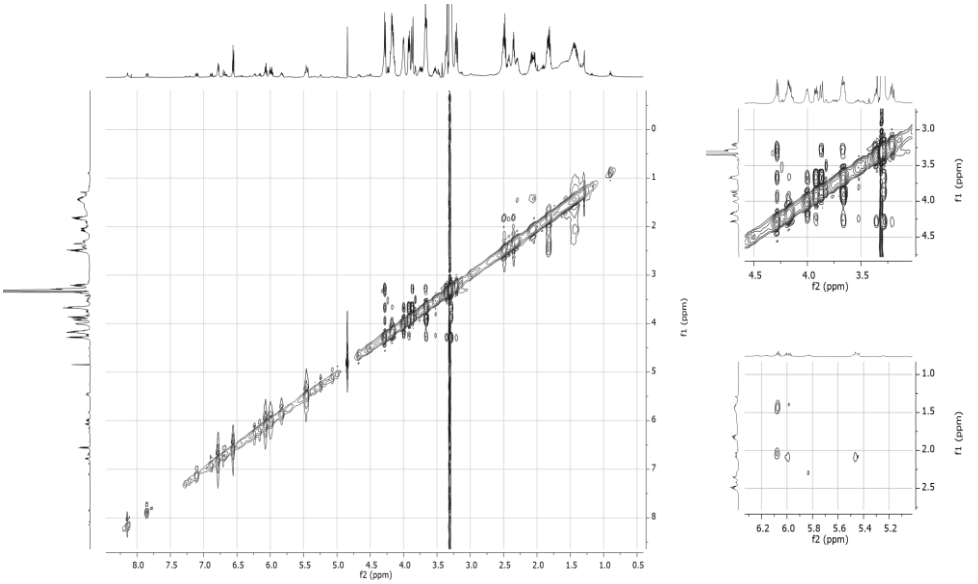

Figure S12: Nuclear Overhauser effect spectroscopy (NOESY) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**4**).

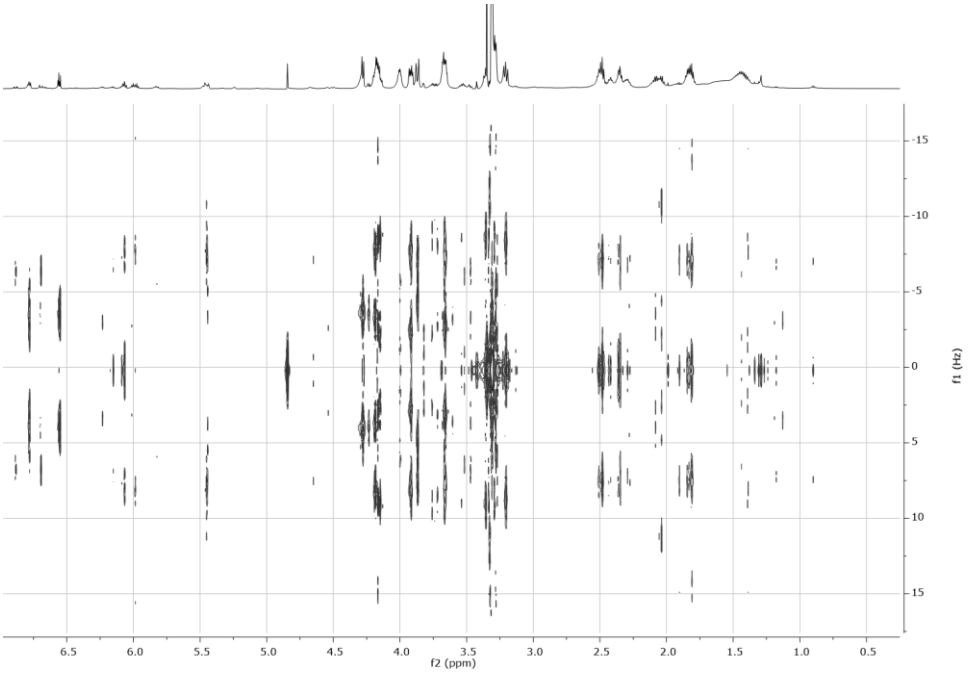

Figure S13: Two-dimensional ^1H-^1H *J*-Resolved spectrum (CH_3OH-d_4, 600 MHz) of Compound (**4**).

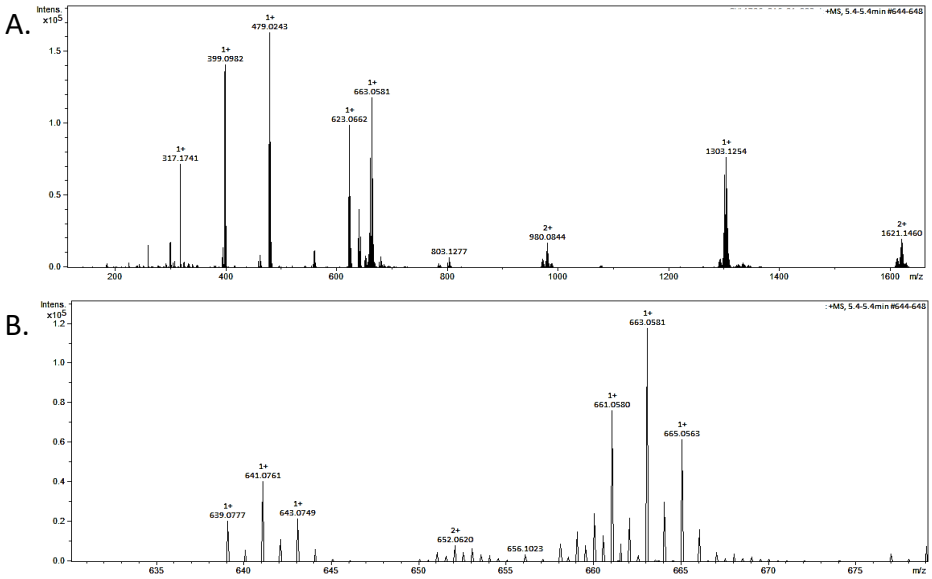

Figure S14: Electrospray ionization-quadrupole-time of flight (ESI-qTOF) MS spectrum in full range (A). and the expanded region around the $[M+H]^+$ and $[M+Na]^+$ ions (B) of Compound (**4**).

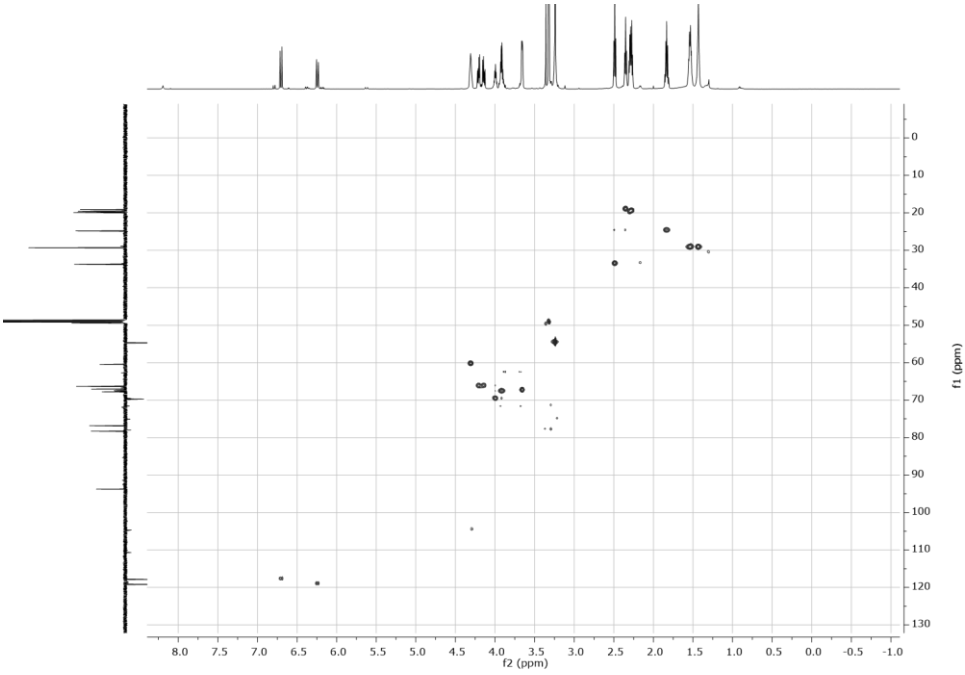

Figure S15: Heteronuclear single quantum coherence (HSQC) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**5**).

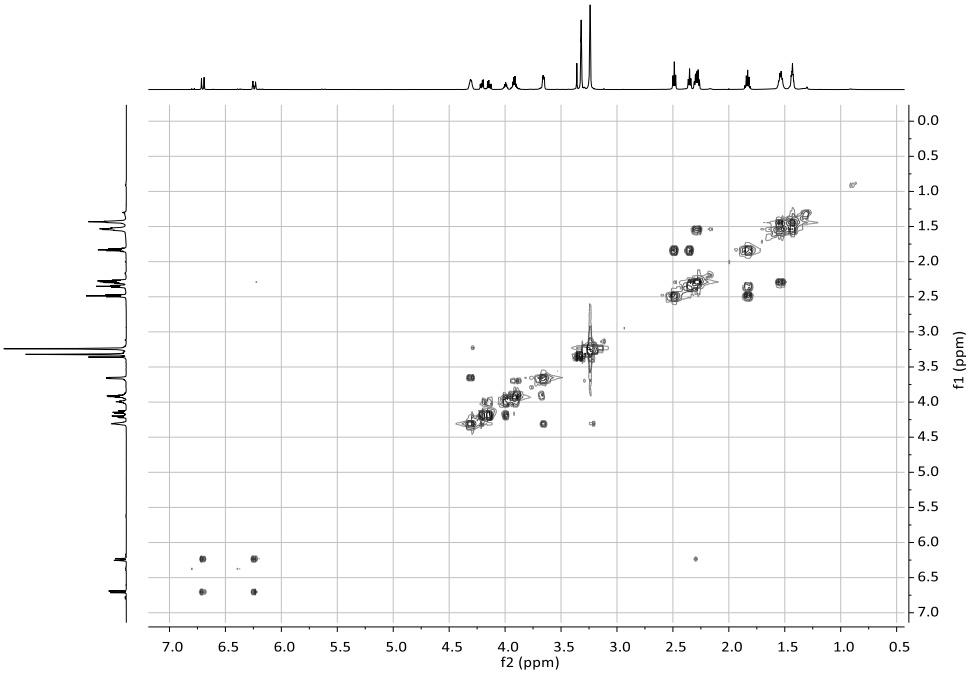

Figure S16: ^{1}H-^{1}H Correlation spectroscopy (COSY) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**5**).

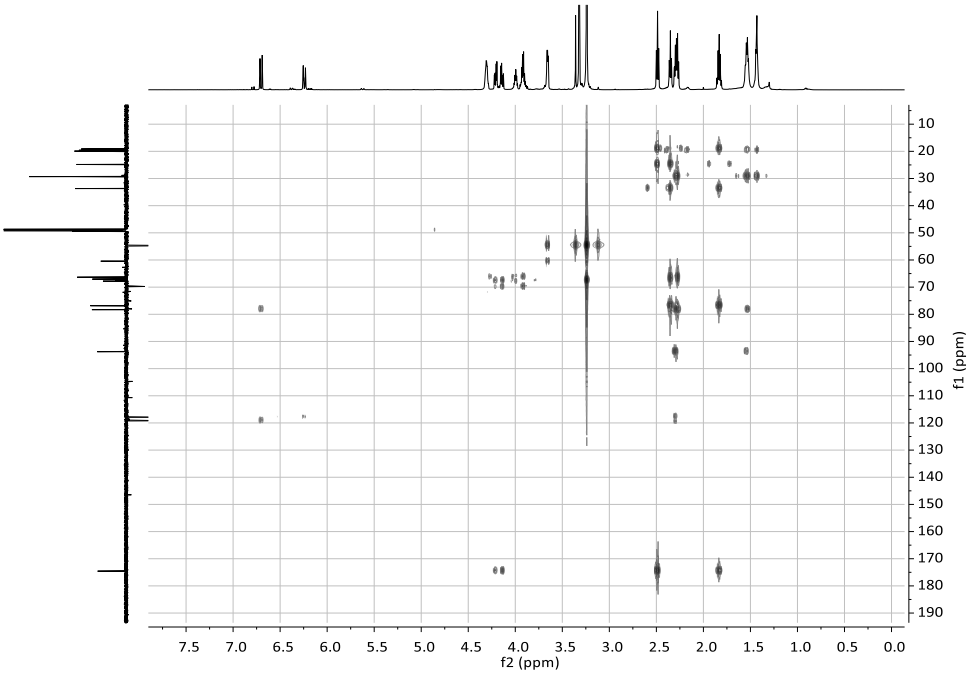

Figure S17: Heteronuclear multiple bond correlation (HMBC) spectrum (CH_3OH-d_4, 600 MHz) of Compound (**5**).

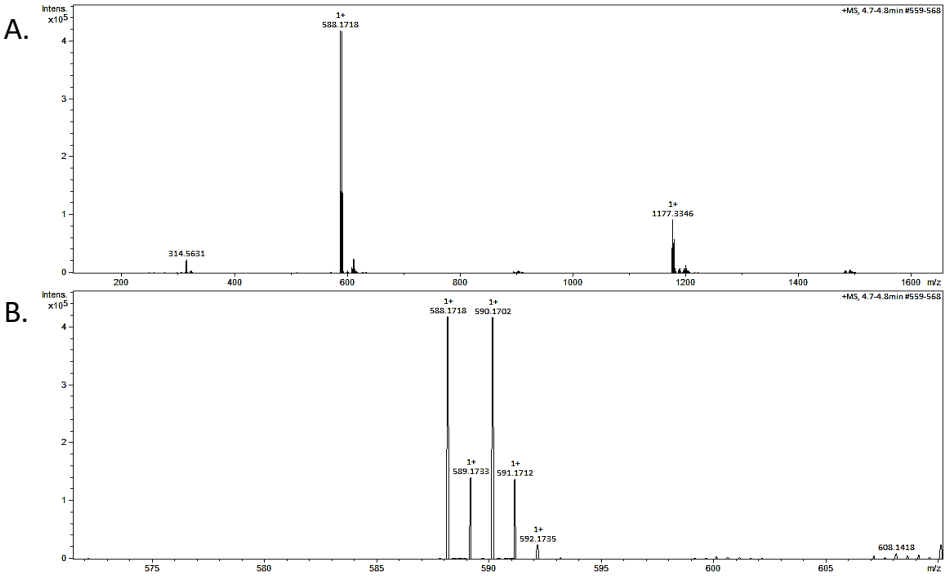

Figure S18: Electrospray ionization-quadrupole-time of flight (ESI-qTOF) MS spectrum in full range (A) and the expanded region around the $[M+H]^+$ ion (B) of Compound (**5**).