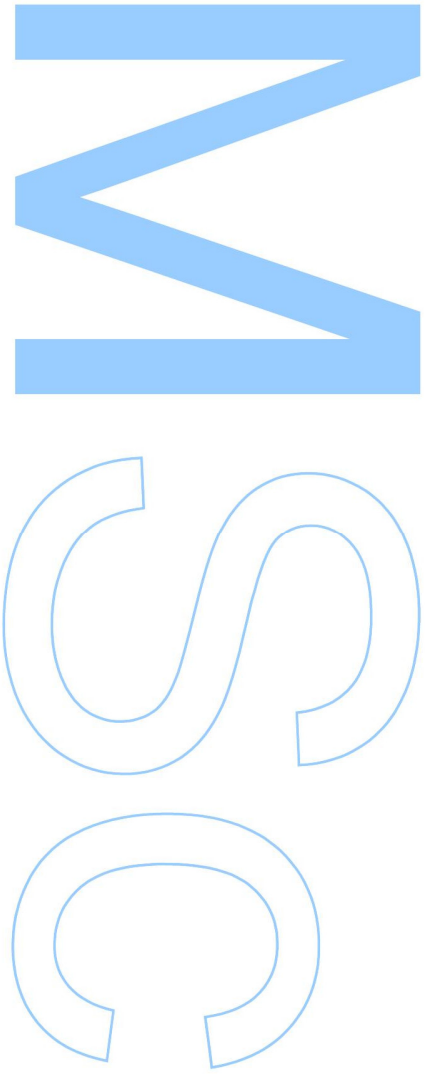
A novel biocatalytic esterification involved in the biosynthesis of branched bartolosides

João Pedro Alves Rodrigues dos Reis
Master's Degree in Cell and Molecular Biology
Department of Biology
2019
Supervisor
Pedro Nuno Leão, Principal Investigator, FCUP, CIIMAR
Cosupervisor
Ralph Urbatzka, CIIMAR

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas
pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____



Acknowledgments

First and foremost, I would like to thank my supervisor Pedro Leão for allowing me to develop my master thesis in his lab. During these two years, he always pushed me to go forward and taught me a lot of things that I will carry with me throughout my scientific and personal life. Your leadership and guidance allowed me to develop a work which makes me proud of.

It has indeed been a pleasure to be part of such a wonderful team that shares so much scientific knowledge and cultivates such a friendly environment. Sara Freitas, Teresa Martins and Adriana Rego I am very glad to have had the opportunity to help assemble our lab, thank you for all your support and motivation. Sandra Figueiredo for helping me achieve all our isolations and teaching me how to work independently with expensive equipment. Ana Vieira for always helping me get my material in time, allowing me faster results. Raquel Castelo Branco for all the support before and after joining CNP. Kathleen Abt and Jorge Antunes for all the kind gestures and friendship. Nádia Eusébio for helping me deal with both my thesis and Halversity project. Thank you all for the shared scientific and social experiences!

I am also thankful for having had the opportunity to work with my co-supervisor Ralph Urbatzka and for his patience when my master thesis' objectives changed. Additionally, I would like to acknowledge all the people from BBE for helping to create a cooperative environment. Individually I want to thank Lígia Sousa for teaching me how to work with mammalian cells and for staying with me until I was able to work independently.

I would also to thank my friends from university/college for all the extracurricular fun and experiences. Filipa Timóteo specially for being such a good and kindhearted friend and Gonçalo Outeiro for all the companionship and gestures throughout our academic and personal journeys. I would like give a special thanks to Eugénia Rocha for always being with me in the good and bad times and for trying to listen to me and help even though you don't understand a word of what I am saying when I talk about science. You are the best in all areas, love you so much. Last but not least, I would like to thank my parents for all the investments and for believing in me and helping me to grow throughout my personal life and for always wanting the best for me. You are the best parents in the universe.

This work was supported by Fundação para a Ciência e a Tecnologia (FCT), through project HALVERSITY (PTDC/BIA-BQM/29710/2017) and by the European Research Council (ERC) through a Starting Grant (759840, FattyCyanos).

Abstract

Cyanobacteria have been extensively studied due to their prolific capacity to biosynthesize natural products with potent biological activities and fascinating chemistry behind their biosynthesis. Bartolosides are a group of halogenated dialkylresorcinols found in certain cyanobacteria that rely on the recruitment of fatty acid derivatives from the primary metabolism. Armed with this knowledge, we envisioned that this could be harnessed to incorporate terminal alkyne moieties into the alkyl chains of bartolosides, generating click chemistry-accessible versions that could be used for probing their biological role. To achieve this, we supplemented cultures of a bartoloside producing cyanobacterium – *Synechocystis salina* LEGE 06099 – with terminal alkyne-containing fatty acids. Instead, our experiments led us to detect and later isolate bartoloside esters of the supplemented 6-heptynoic acid. We then demonstrated that naturally occurring esterified versions of bartolosides could be found in *S. salina* LEGE 06099 in the absence of fatty acid supplementation. Such findings encouraged us to focus our attention on the bartolosides biosynthetic gene cluster, *brt*, particularly the enzyme BrtB which had no ascribed function. Through several *in vitro* assays we established that BrtB is responsible for the esterification of bartolosides. BrtB is a novel enzyme that alkylates free fatty acids using secondary alkyl chlorides (the bartolosides). It is the second enzyme of its family to be characterized following the Friedel-Crafts C-alkylating enzyme, CylK. Phylogenetic analysis of these enzymes revealed a series of cyanobacterial biosynthetic gene clusters that may represent opportunities for the discovery of novel natural products and enzymes.

Resumo

As cianobactérias têm sido extensivamente estudadas devido à sua capacidade prolífica de produzir compostos naturais derivados de processos bioquímicos únicos associados à sua biossíntese. As bartolosidas são um grupo de dialquilresorcinóis halogenados encontrados em certas cianobactérias que dependem do recrutamento de ácidos gordos derivados do metabolismo primário. Com isto em mente, previmos que isso poderia ser aproveitado para incorporar alcinos terminais nas cadeias alifáticas das bartolosidas, gerando versões acessíveis a "click chemistry" que poderiam ser utilizadas para investigar a sua função biológica. Para isso, suplementamos as culturas de uma cianobactéria produtora de bartolosidas - *Synechocystis salina* LEGE 06099 - com ácidos gordos contento alcinos terminais. Contudo, as nossas experiências levaram à deteção e, posterior, isolamento de ésteres de bartolosides com o ácido 6-heptinóico suplementado. Em seguida, demonstramos a ocorrência natural de versões esterificadas das bartolosidas, encontradas na cultura de *S. salina* LEGE 06099 sem suplementação de ácidos gordos. Estes resultados encorajaram-nos a dedicar nossa atenção no "cluster" biossintético das bartolosidas, *brt*, particularmente na enzima BrtB, inicialmente sem função conhecida. Através de vários ensaios *in vitro*, estabelecemos que a BrtB é responsável pela esterificação de bartolosidas. BrtB é uma nova enzima que alquila ácidos gordos livres usando as cadeias alifáticas cloradas presentes nas bartolosidas. Esta é a segunda enzima da sua família a ser caracterizada após a enzima CylK que catalisa a alquilação de Friedel-Crafts. A análise filogenética dessas enzimas revelou uma série de "clusters" biossintéticos cianobacterianos que podem representar oportunidades para a descoberta de novos compostos e enzimas naturais.

Key words

Natural products, secondary metabolites, polyketide, non-ribosomal peptide, terminal alkyne, dialkylresorcinol, esterification, alkylation, structural elucidation, heterologous expression.

Abbreviations list

ATP	adenosine triphosphate
BGC	biosynthetic gene cluster
CAI	codon adaptation index
COSY	COrrelated SpectroscopY
DAR	dialkylresorcinol
DNA	deoxyribonucleic acid
EIC	extracted ion chromatogram
FDA	Food and Drug Administration
HMBC	Heteronuclear Multiple Bond Correlation
HRESIMS	high resolution electron spray ionization mass spectrometry
HSQC	heteronuclear single quantum correlation
IR	infrared
LC	liquid chromatography
MS	mass spectrometry
NMR	nuclear magnetic resonance
NP	Natural product
NRPS	non-ribosomal peptide synthetases
OD	optical density
PCR	polymerase chain reaction
PKS	polyketide synthases
RiPPS	Ribosomally synthesized and post-translationally modified peptides
SAM	S-adenosyl methionine
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
TLC	thin layer chromatography
UV	ultraviolet

Table of contents

Table and Figure List

I. Introduction

1. Impact of natural products in the biological sciences and industry

Natural products have impactful roles in our society. With the serendipitous discovery of penicillin by Alexander Fleming, 1929[1], modern natural products research flourished with potential applications in several fields, namely drug discovery, nutrition, chemical biology and also in the industry.[2-3] However, the understanding of what a natural product is may not be clear and consistent throughout different scientific areas. From the chemical point of view, a natural product can be described as a small molecule derived from a biological source and that is not strictly required for the development and viability of the producing organism.[2, 4]

From a historical point of view, mankind has always entrusted Nature to supply not only food but also ways to improve life quality. Natural products have been used in traditional medicine from ancient times, with records of plant-based substances dating 2600 B.C.[5] However, secondary metabolite chemical diversity and biological importance started to gain more importance after 1817 with the isolation of morphine from *Papaver somniferum* by a German pharmacist Friedrich Sertuerner.[6] Such discovery motivated scientists to pursue plant based compounds capable of helping the treatment of a wide spectrum of diseases, for example, the discovery of quinine, a promising antimalarial drug from the bark of *Cinchona* species.[7]

Plant derived natural products were undoubtedly fundamental to the improvement of life quality throughout the times, but advances in microbiology research stimulated the screening for bioactive secondary metabolites from microbial sources. [8] As mentioned above, the discovery (Alexander Fleming) and isolation (Howard Florey) of penicillin from the fungus *Penicillium notatum* triggered a massive interest in studying microbial organisms' secondary metabolism, with a special focus on bacteria.[9] The huge diversity of bacteria was, by itself, a motivation to seek for bioactive compounds that could increase and expand the scientific horizons in terms of chemical and structural diversity together with the promising biological and industrial role of bacterial metabolites. In fact, even though natural products found in the Plantae kingdom are approximately eight times the ones found in the Bacteria kingdom, the relative amount of bacterial bioactive secondary metabolites is heavily superior.[9-10] Many bioactive compounds obtained from bacterial sources present antibacterial properties[9-10] (Figure 1), representing 51% of all FDA-approved antibacterial natural products obtained from either plants, fungi or bacteria.[11] For instance,

Actinomycetes, a group of filamentous gram positive bacteria, are responsible for the production of almost all antibiotics used in the pharmaceutical industry, with almost 75% of reported bioactive compounds being produced by a single genus, *Streptomyces*.[12-13] One of the most significant antibiotics with bacterial origin is streptomycin, initially isolated by Albert Schatz and Selman Waksman in 1944, which was later introduced in the market as a tuberculosis treatment as result of its strong activity against *Mycobacterium tuberculosis*.[14-15]

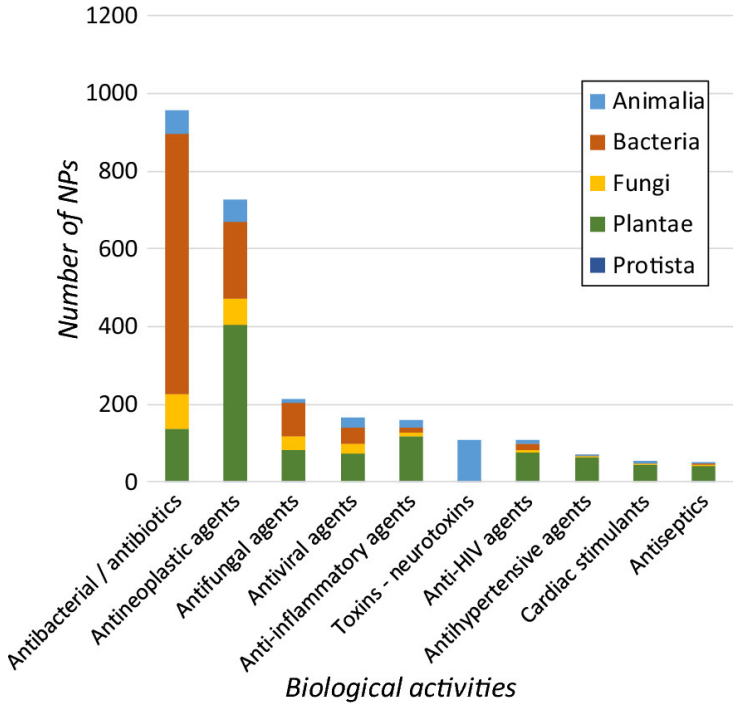


Figure 1. Major bioactivities of natural products and their distribution throughout the different kingdoms of life. Adapted from Chassagne *et al.* 2019.

As discussed above, secondary metabolites were initially studied with the aim to find novel compounds, with special consideration to antibacterial and anticancer compounds. However, the applications of such diverse and complex natural products goes beyond their direct application to the pharmaceutical industry. The fast increase in global population demanded more intense farming practices and with that the need for products capable of both protecting the cultivations and conserving the produced food.[10, 16] Further concern was also related with the handling of environmentally friendly compounds due to their low toxicity when compared to other industrial chemicals.[17] With that in mind, natural products started to gain relevance in these fields and secondary metabolite producers started to be screened for potential applications in agricultural

activities and food industry.[16] For instance, natural products have been actively applied as pesticides and herbicides as well as food additives and preservatives with the financing dedicated to these fields usually rounding 50 billion dollars per year[10, 17-18], once again with a special focus on bacterial secondary metabolites.[19] From the reported and approved natural products used as pesticides, avermectin was one of most recognized compounds due to its potent anthelmintic activity.[20] This secondary metabolite, produced by *Streptomyces avermitilis*, causes the hyperexcitation of invertebrates' nervous system and effectively paralyzes them, protecting plantations from being damaged.[16, 21] Secondary metabolites have also been extensively studied and employed as food ingredients, as in the application of pigments as natural colorants, which can also be derived from microbial sources.[22] In some cases, secondary metabolites may be used both as food additives to confer specific colors but also as bioactive compounds – the cyanobacteria and microalgae-derived scytonemin and astaxantin, respectively, serve as good examples due to their anti-inflammatory properties.[23-24]

Although natural products unarguably have a clear role in our daily lives, their potential applications go beyond their direct use. Advances in interdisciplinary domains, such as chemical sciences, have been driven by attempts to mimic the intricate and complex chemistry and biosynthesis of natural products.[25] The use of enzymes in chemical synthesis (biocatalysis) is currently a very promising field of academic and industrial research. Biocatalysis offers several advantages for the pharmaceutical industry, for example, due to the economic and environmental benefits over traditional "non-green processes".[26] Enzymes allow the synthesis of different products due to their regio- and stereo- selectivity without the need for protection steps as in organic synthesis.[26-27]

The use of biocatalysis over traditional chemical catalysts also allows the expansion of organic synthesis' toolbox with the use of several reactions such as oxidations, reductions, halogenations among other, with a special focus on alkylation reactions. For instance, chemical synthesis reactions that require either O-alkylation, C-alkylation and N-alkylation are time-consuming and are not economical and environmentally viable due to the use of toxic alkylating agents.[27] A great example of enzymatic catalysts are methyltransferases, a widespread group of enzymes involved in the transfer of methyl groups, which can be green alternatives for Friedel-Crafts alkylations (figure 2).[28-29] The catalytic mechanism of enzymatic oxidations are also of great interest to the pharmaceutical industry since they avoid undesired side-products due to their high stereoselectivity.[30] However, enzymatic catalysis is not adapted to industrial processes leading to the necessity for advancements in protein engineering with regard to the reaction thermodynamics

and specificity of the enzymatic reactions.[31] Additionally, innovative methodologies that improve the biocatalysts performance have also been explored, for example in the scale up of Bayer-Villiger oxidations through the use of bioreactors and resin based methodologies (figure 2).[30, 32] This reaction is performed by flavin-dependent monooxygenases, which are usually related with natural product biosynthesis and can be adapted to organic synthesis and in prodrug activation.[33-34] Besides, studying proteins and enzymes allows to better understand how reactions are carried in terms of substrate affinity to active centers as well as substrate interactions and enzymatic folding.

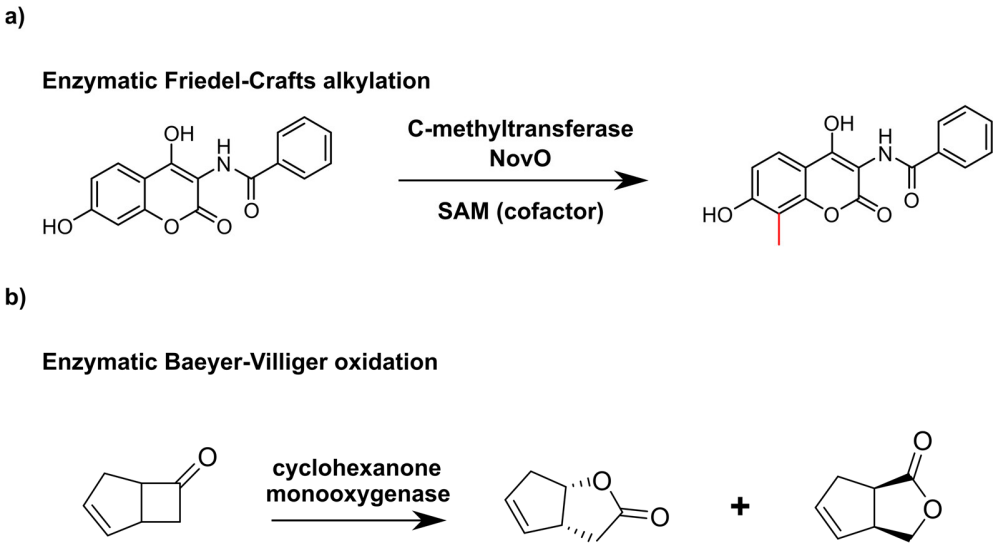


Figure 2. Examples of enzymatic reactions used in natural product biosynthesis. a) Methyltransferase NovO performs an enzymatic Friedel-Crafts alkylation, using S-Adenosyl methionine (SAM) as a cofactor, to introduce a methyl moiety in aromatic substrates. [29] b) Cyclohexanone monooxygenase performs Bayer-Villiger oxidations.[33] Adapted from Tengg *et al.* 2016 and Hilker *et al.* 2005.

2. Cyanobacteria are prolific natural product producers

Terrestrial bacteria are the most well-known source of bioactive metabolites, but extensive exploration programmes led to a decrease in the number of new natural products discovered from these sources each year.[35] On the other hand, the development of efficient strategies and technologies for marine sampling allowed to expand the secondary metabolite discovery to the oceans due to the high biological and chemical diversity associated.[36-37]

Cyanobacteria are widespread photosynthetic gram-negative prokaryotes that can be found in different environments, such as freshwater and marine ecosystems as well as in extreme

habitats.[38-39] They have had an important role in the production of oxygen dating at least 2.3 billion years ago, which make these organisms one of the most ancient and diverse groups of organisms on our planet. [40-41] Cyanobacteria are also unique among bacteria since some species are capable of cell differentiation into structures called heterocysts and akinetes.[41] The ability of compartmentalization allows these microorganisms to fix nitrogen, in the heterocysts, and survive under extreme conditions (e.g. light deprivation) by entering a vegetative-like status through the specialized cells akynetes.[41]

Although they are commonly recognized as oxygen producing microorganisms, cyanobacteria are also a prolific source of secondary metabolites. [38-39, 42] Richard Moore was a pioneer in this field with his group's work in cyanobacterial derived compounds and was responsible for the discovery of dozens of cyanobacterial secondary metabolites, such as majusculamides A and B[43], oscillatoxins or debromoaplysiatoxin.[44-46] There are many applications for cyanobacterial compounds due to their varied biological properties such as UV protection conferred by the scytonemins[47] or anticancer activity, e.g. dolastatin 10, the pharmacophore of the anticancer drug brentuximab vedotin.[48-50] However, the question of what distinguishes cyanobacteria from other secondary metabolite producers arises. The high chemical diversity of cyanobacteria is in fact one of the reasons that make such organisms great targets for natural product discovery.[38-39, 51] The vast and unique mechanisms involved in the biosynthesis of cyanobacterial secondary metabolites allow these microorganisms to generate varied compounds that are of interest to chemists and biologist alike.

The secondary metabolite production relies on a biosynthetic machinery consisting of a contiguous arrangement of a set of genes defined as biosynthetic gene clusters (BGCs).[52-53] Bacterial genes codifying for secondary metabolites are physically clustered while in plants the biosynthesis of natural products does not follow this assembly line, which contributed to the focus on bacterial genomes.[54] The rationale behind BGCs is that the biosynthetical pathway involves the interaction of each enzyme of the cluster with substrates to promote the "brick by brick" construction of the resulting secondary metabolite (figure 3). Additionally, biosynthetic gene clusters can include enzymes capable of performing halogenation reactions (e.g. chlorine and bromine) as well as other mechanistically interesting reactions that enable the production of a broad range of metabolites with different chemical moieties.[55]

Apart from the compounds generally present in other organisms such as terpenes and polysaccharides, these photosynthetic microorganisms produce metabolites that may enclosure

chemically diverse structures that range from lipid- to amino acid-based skeletons., among others.[38, 51, 56] In fact, cyanobacterial metabolites can be classified according to their chemical structural and divided into 10 classes including alkaloids, polyketides and peptides, as well as hybrids thereof.[38] Most of the metabolites produced by cyanobacteria are peptides[38] that are mainly derived from non-ribosomal peptides synthetases (NRPS) or ribossomally-produced and post-translationally modified peptide (RiPP) systems. NRPs are synthesized by the action of several enzymes encoded in their respective BGCs, without the involvement of ribosomal RNA as in the usually seen peptides.[56-57] On the other hand, RiPPs biosynthesis relies on a ribosomally generated peptide containing recognition sequences that will later promote post-translationally modifications, such as cyclization, yielding the final natural product (e.g. sphaerocyclamide).[58-59] Additional secondary metabolites are the fatty acid-based compounds called polyketides (PKs), e.g. oscillatoxins[45]. Their biosynthesis relies on the action of modular systems consisting on the elongation of activated fatty acids and posterior modifications resulting in a broad structural diversity.[60-61]

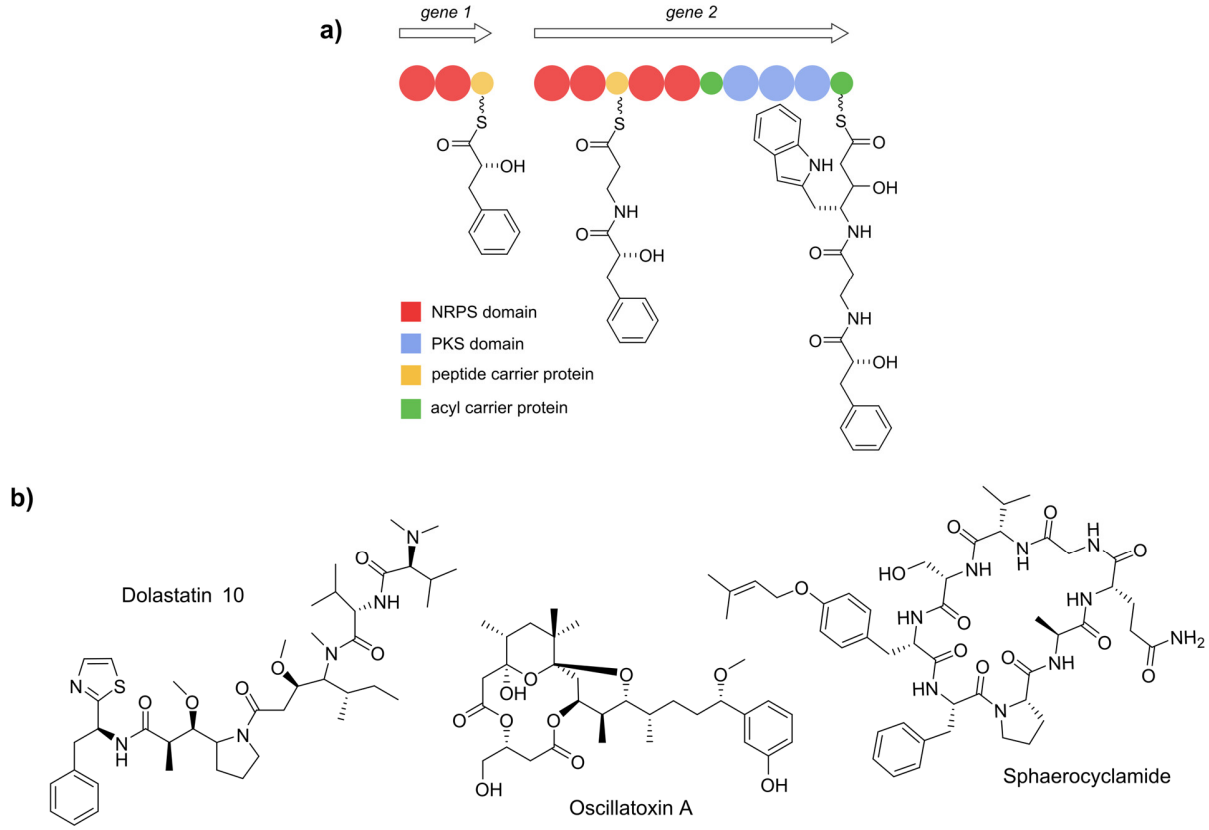


Figure 3. Examples of structural and chemical diversity of cyanobacterial secondary metabolites. a) depicted is scheme of a partial biosynthetic gene cluster. b) Dolastatin 10 is a lipopeptide derived from marine cyanobacterium *Symploca* sp. that has potent anticancer activity.[58]Oscillatoxin A is a toxic metabolite derived from PKS systems[46]. Sphaerocyclamide is a RiPP found in the cyanobacterium *Sphaerospermopsis* sp. LEGE 00249.[60]

In addition, cyanobacteria are capable of producing remarkable compounds such as lipopeptides, which contain lipid chains bound to peptide moieties, e.g. the previously mentioned dolastatin 10.[62]

The cyanobacterial chemical diversity encourages the discovery of natural products not only for the novelty of the enzymatic chemistry behind their biosynthesis but also for the opportunity to discover new potent activities.

3. Approaches for natural product and drug discovery

The discovery of penicillin caused the emergence of the Golden Age of natural products where the main focus was drug discovery. This generated a need for developing efficient approaches capable of providing a high rate of natural product discovery that could be used in the industry. During these years, the application of bio-guided drug discovery was the main and only source of potential candidates due to the simple assays and methodologies that could be applied.[63] Put simply, target cells were exposed to microbial extracts and the phenotype of the assay would dictate the compounds activity and determine the need for further isolation efforts.[63] However, such methodologies soon started to be ineffective due to the time consuming and expensive methodologies applied, which led to the need for high throughput screening approaches.[63-65] When the scientists started to focus on the chemical diversity of natural products techniques such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) were extensively used and allowed faster dereplication processes and secondary metabolite discovery.[66] In fact, the applicability of combinatorial bio-guided and mass spectrometry approaches is still used for the discovery of natural products.[e.g.67]

Nevertheless, with the fast development in DNA sequencing technologies[68] the natural products area witnessed a resurgence of interest relying on genomic and bioinformatic approaches. For instance, metagenomics allowed the discovery of an unprecedented number of compounds from uncultured microorganisms.[69] Additionally, genome mining approaches allowed the fast expansion in knowledge of both biosynthetic gene cluster architechture and natural product chemistry in the beginning of the 21st century.[70] However, predicting the end-product of a gene cluster can be sometimes challenging. Such problem could be addressed, for example, through the use of heterologous systems: entire BGC can be cloned into suitable heterologous hosts (e.g. *Streptomyces* sp.) to access the secondary metabolite derived from the gene cluster inserted.[e.g.71-72] In the case of cyanobacterial secondary metabolites such bioinformatic approaches have been

extensively used for the characterization of PKS and NRPS derived compounds. [73-74] These strategies have allowed the discovery of a huge variety of cyanobacterial natural products but such methodologies are still non-universal, can be time consuming or even discouraging.

One other subject of natural products research is the study of their mode of action, in particular, regarding their ecological or pharmacological relevance. Achieving this, however, is not trivial. Click chemistry has emerged as a powerful set of biorthogonal reactions (i.e. chemical reactions occurring in living systems without affecting their biological processes[75]) that are acknowledged for their fast and simple procedures with a wide-range of applications in chemical biology, e.g. deciphering the mode of action of a natural product.[76-77] One of the most common approaches relies on heterocycle formation between azides and terminal alkyne moieties through a copper-catalyzed Huisgen ,3-dipolar cycloaddition (figure 4).[78] Click chemistry advancements have allowed the creation of studies based in bio-conjugation of functionalities of interest to the azide, later generating a tagged version of the desired compound which is suitable for both *in vivo* and *in vitro* studies.[77, 79-80]

Click Chemistry:

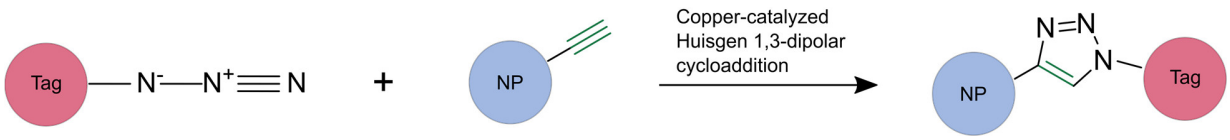


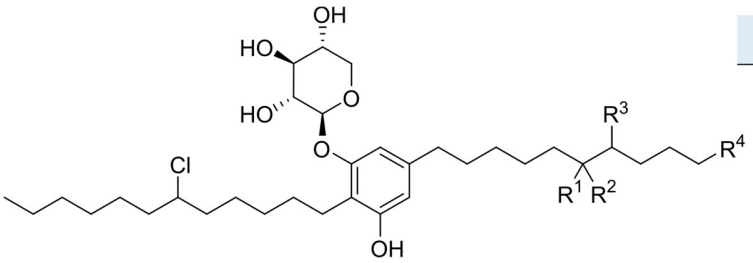
Figure 4. Click chemistry using azides and terminal alkynes. For example, terminal alkyne labelled natural products (NP) react with azides to form a heterocycle through a copper-catalyzed Huisgen 1,3-dipolar cycloaddition. Azides can be tagged with several moieties such as biotin and fluorophores.

4. Bartolosides, a group of cyanobacterial dialkylresorcinols

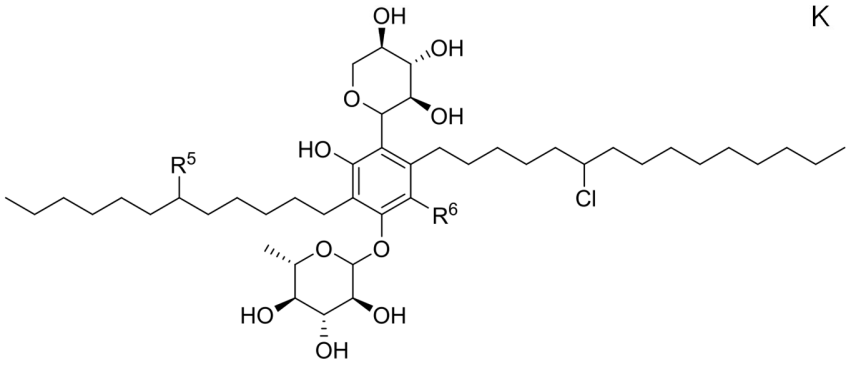
Bartolosides are a recently discovered and unprecedented group of cyanobacterial secondary metabolites present in some cyanobacteria with uncommon features such as a chlorinated dialkylresorcinol (DAR) core or C-glysosyl moieties.[81-82] Non-halogenated DAR scaffolds are somewhat common in other bacteria but only a few DAR compounds have been discovered from cyanobacteria, e.g. paracyclophanes, microcarbonin A and nostocyclyne A. [83-86]

Bartolosides can be divided into two groups according to their structure: i) monoglycosylated versions, containing an O-linked β-D-xylosyl moiety and ii) diglycosylated versions featuring an O-linked α-L-rhamnose and a C-linked xylose while additional differences consist in the alkyl chain length that can range from 11 to 15 carbons.[81-82] Bartolosides A-K have been discovered and

isolated from the coccoid cyanobacterium *Synechocystis salina* LEGE 06155 and *Synechocystis salina* LEGE 06099 as well as from the filamentous *Nodosilinea* sp. LEGE 06102.



Bartoloside	R_1	R_2	R_3	R_4
A (**1**)	Cl	H	H	C_3H_7
E	H	H	Cl	CH_3
F	Cl	H	H	C_4H_9
G (**2**)	H	H	H	C_3H_7
H	Cl	H	H	C_5H_{11}
I	Cl	Cl	H	C_3H_7
J	H	H	H	CH_3
K	Cl	H	H	C_2H_5



Bartoloside	R5	R6
B	Cl	H
C	H	H
D	Cl	H

Figure 5. Bartolosides chemical structures enclose several chlorinations, different alkyl chain lengths and can be mono- or diglycosylated.

The bartolosides biosynthetic gene cluster (*brt*) (Figure 6) encodes two enzymes essential for the dialkylresorcinol formation (DAR). The ketosynthase BrtD performs the condensation between two fatty acyl thioester derivatives leading to the ring formation while the flavin-dependent aromatase BrtC finalizes the resorcinol formation.[82] Additional genes *brtGHI* encoded in the *brt* cluster are involved in the glycolipid export while *brtAEF* encode glycosyltransferases. Halogenation of the alkyl chains in the bartolosides is predicted to be performed by BrtJ since some of its homologs (CylC, ColD/ColE) have been found in cyanobacterial biosynthetic gene cluster encoding for halogenated secondary metabolites, cylindrocyclophanes and columbamides.[87-88] Previous bioinformatic analysis also supports that this enzyme belongs to a class of dimetal halogenases acting on unactivated carbons.[81-82]

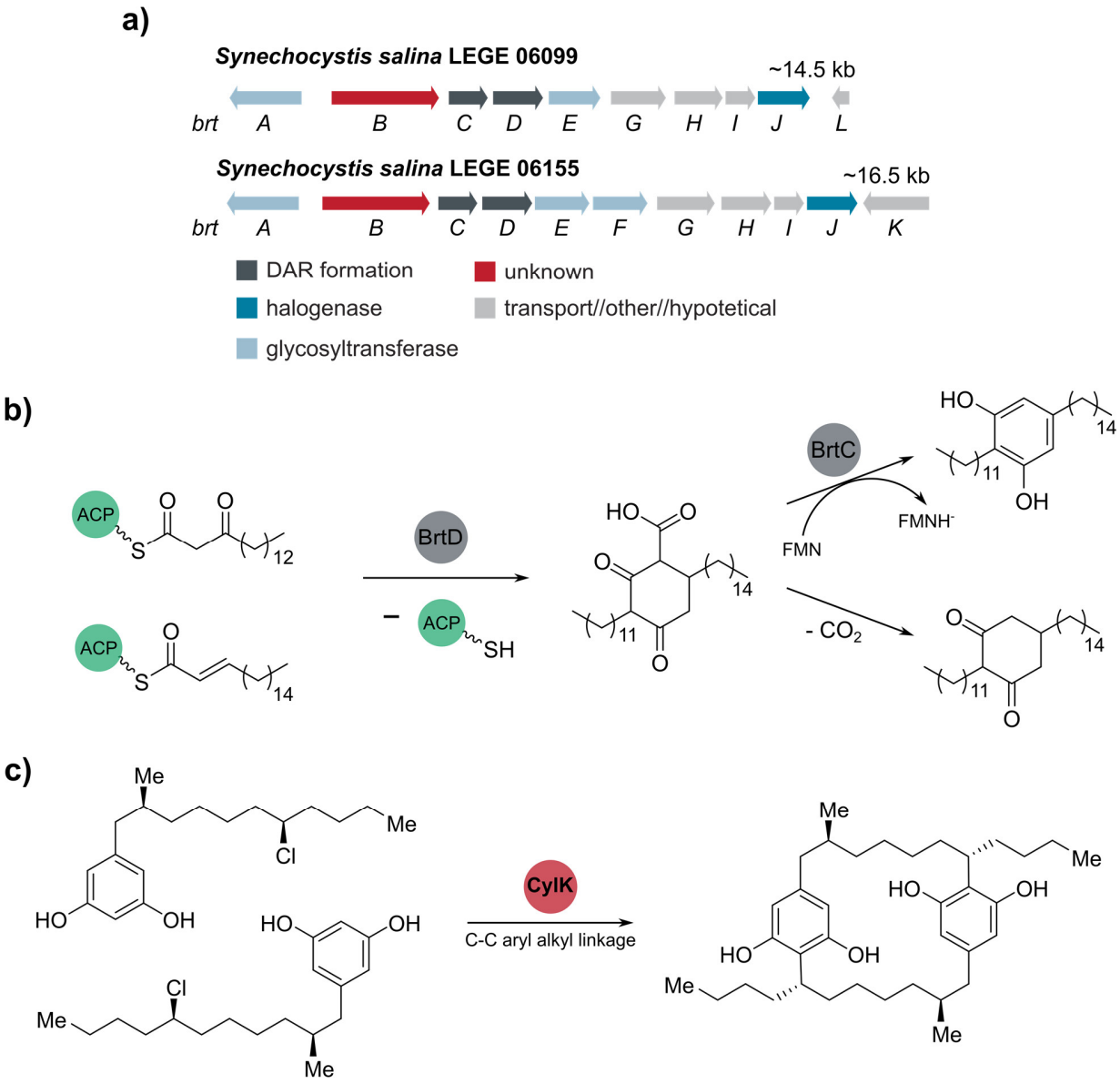


Figure 6. Biosynthesis of bartolosides and cylindrocyclophanes. a) Bartolosides biosynthetic gene cluster, *brt*, from *S. salina* LEGE 06099 and *S. salina* LEGE 061155 with annotated gene functions. b) proposed mechanism of dialkylresorcinol formation in bartolosides B-D. Leão *et al*. 2015.c) CylK catalyzes the paracyclophane formation in the cylindrocyclophanes.

An additional enzyme – BrtB – is encoded in the *brt* cluster. Phylogenetic analysis indicates that BrtB and CylK - an enzyme that participates in paracyclophane biosynthesis – are homologs.[89] The paracyclophanes are cyanobacterial natural products that contain aryl-alkyl linkages in their core structure, however, the biocatalytic mechanisms behind paracyclophane assembly were only recently clarified, in particular the function of CylK.[87, 89] Through several *in vitro* assays this enzyme (annotated as hemolysin-type calcium-binding protein) was shown to be responsible for the

paracyclophane formation through a Friedel-Crafts alkylation between the chlorinated chains of the intermediates and the resorcinol ring through a S_N2 nucleophilic substitution (Figure 6c).[89-90] Despite its homology with CylK, the function of BrtB has remained unclear. The reasons behind this are that a CylK-like mechanism is not necessary for the biosynthesis of the known bartolosides and that CylK requires a free C-2 position in the alkylresorcinol, which is unavailable in the bartolosides.

Objective: The biosynthesis of the dialkylresorcinol skeleton in the bartolosides involved recruitment of fatty acids derivatives from primary metabolism. We envisioned that this could be exploited to incorporate terminal alkyne moieties into bartolosides and generate click-chemistry accessible versions of these secondary metabolites for probing their biological role. However, while attempting this we encountered unexpected results that led us to diverge from this initial aim and focus on the:

- Isolation of novel bartolosides from *S. salina* LEGE 06099;
- Biosynthesis of the novel bartolosides in *S. salina* LEGE 06099.

II. Results and Discussion

1. Incorporation of terminal alkyne-containing fatty acids into bartolosides

As previously mentioned, the biosynthesis of bartolosides involves the recruitment of fatty acid derivatives from primary metabolism to assemble the dialkylresorcinol skeleton. We hypothesized that this could be exploited to incorporate terminal alkyne moieties into these natural products, making them accessible for click chemistry experiments to gain insights in their biological role.

In order to understand if cyanobacteria could incorporate the desired terminal alkynes we initially assessed whether they could be toxic to the cyanobacterial cells. To this end, we supplemented the bartoloside producer *Synechocystis salina* LEGE 06099, with both 5-hexynoic and 6-heptynoic acids at three different concentrations (5, 10 and 50 mg L^{-1}). The OD_{750nm} was measured in a 96-well plate in a microplate reader (Synergy HT, BioTek) at day 0 (day of supplementation) and every 2 days until reaching the 30-day period.

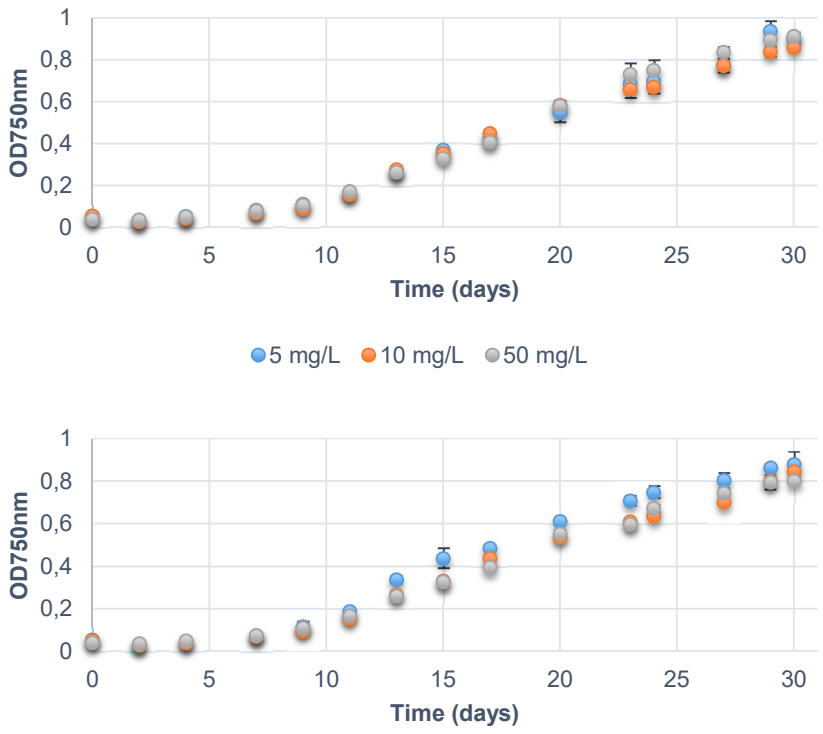


Figure 7. *S. salina* LEGE 06099 growth over 30 days after exposing the culture to both 5-hexynoic (top) and 6-heptynoic acid (bottom).

We observed that supplementation with both acids did not affect the normal growth of *S. salina* LEGE 06099 since they presented the normal growth curve tending to the stationary phase after 20 days, approximately. Additionally, the highest concentration applied (50 mg/L) was still soluble in water and did not precipitate. These allowed us to proceed with the experiments and analyze the incorporation of the terminal alkyne-containing fatty acids into the cells.

Expected results:

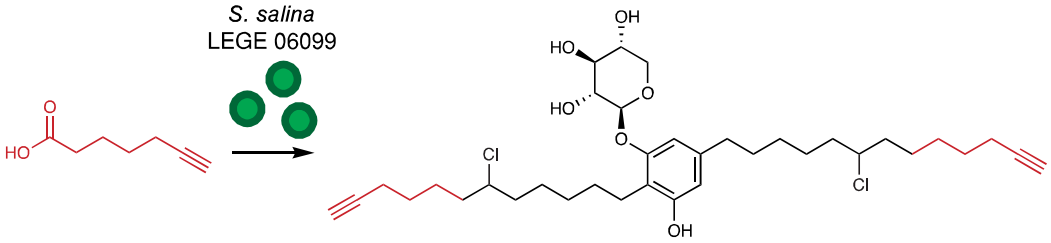


Figure 8. Supplementation of terminal alkyne-containing fatty acids was expected to lead to the incorporation of terminal alkyne moieties into the alkyl chains of bartolosides produced by *S. salina* LEGE 06099.

After harvesting the cells and performing the extraction of organic compounds with a mixture of MeOH/CH_2Cl_2 the samples were analyzed by LC-HRESIMS. Since the target compounds (bartoloside derivatives) have long aliphatic chains we adapted the LC method to be able to observe highly non-polar species (see “Materials and Methods” section). The output analysis revealed a massive depletion of the major metabolite bartoloside A (**1**) levels as well as those of several other analogues (figure 9). However, when searching for the *m/z* values corresponding to terminal-alkyne versions of bartoloside A and other major metabolites we did not detect any ions compatible with these structures. Bartoloside A (**1**) is one of the most abundant secondary metabolites in this strain, representing around 0.6% dry weight[81], which means that the feeding experiment performed could either have affected its biosynthesis or have been incorporated in an unexpected region.

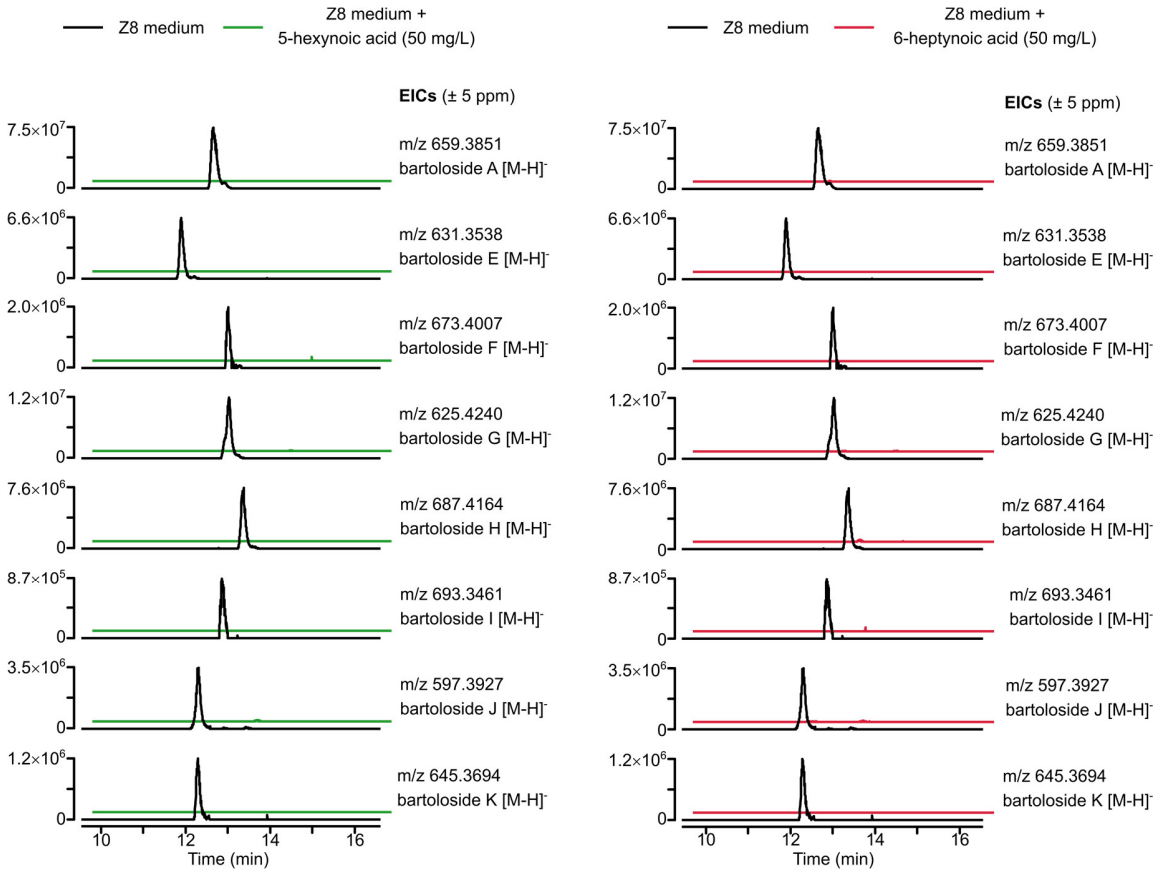


Figure 9. Depletion of bartolosides levels in *S. salina* LEGE 06099 supplemented with fatty acids. Analysis was carried by LC-HRESIMS of the crude extracts obtained and presented are the extracted ion chromatograms (EICs) of the *m/z* correspondent to the $[M-H]^-$ ions of bartolosides.

In order to understand the reason behind bartolosides depletion, we sought to find new masses that were only present in the samples corresponding to supplemented cultures. Further analysis of the obtained chromatograms revealed a very different profile between samples and led to the detection of several peaks that were highly abundant in the cultures exposed to the fatty acids but were absent in the controls (figure 10). We verified that these ions were consistent with the incorporation of one or two units of the supplemented fatty acids with the concomitant loss of one or both Cl atoms in bartolosides A (**1**) and G (**2**) (most abundant metabolites), respectively.

In order to support our hypothesis, we performed fragmentation of some of the new abundant ions through LC-HRESIMS/MS analysis of the crude extracts (figures 11 and 12). The data showed fragments corresponding to the intact alkyne precursors or their neutral loss in the bartolosides core structure. This suggested that the detected peaks could correlate with esterified versions of bartolosides **1** and **2** with 5-hexynoate and 6-heptynoate.

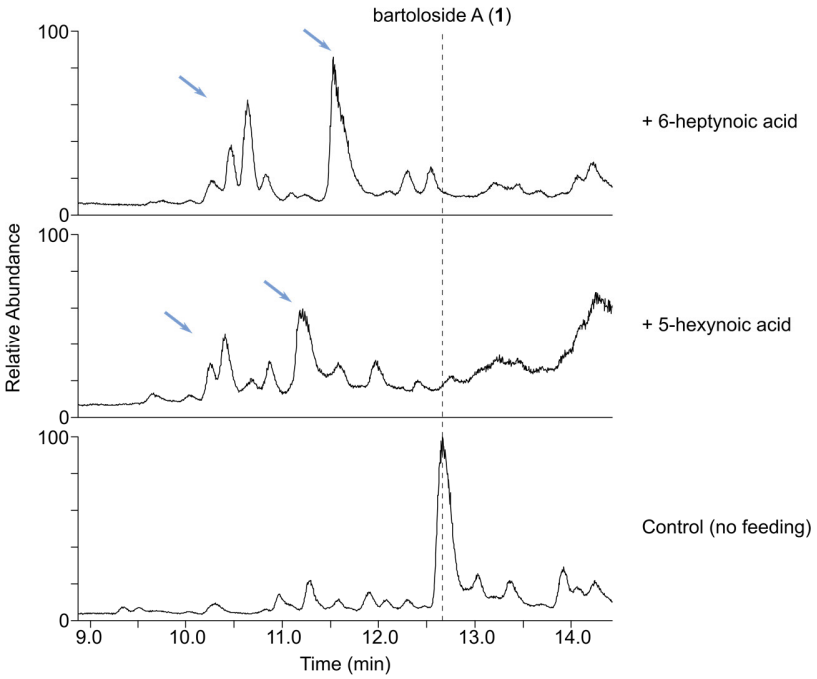


Figure 10. Total Ion Chromatograms obtained from LC-HRESIMS analysis. Supplementation with 5-hexynoic and 6-heptynoic acids led to the formation of new abundant masses absent in the control. Additionally, bartoloside A level was depleted in the exposed cultures.

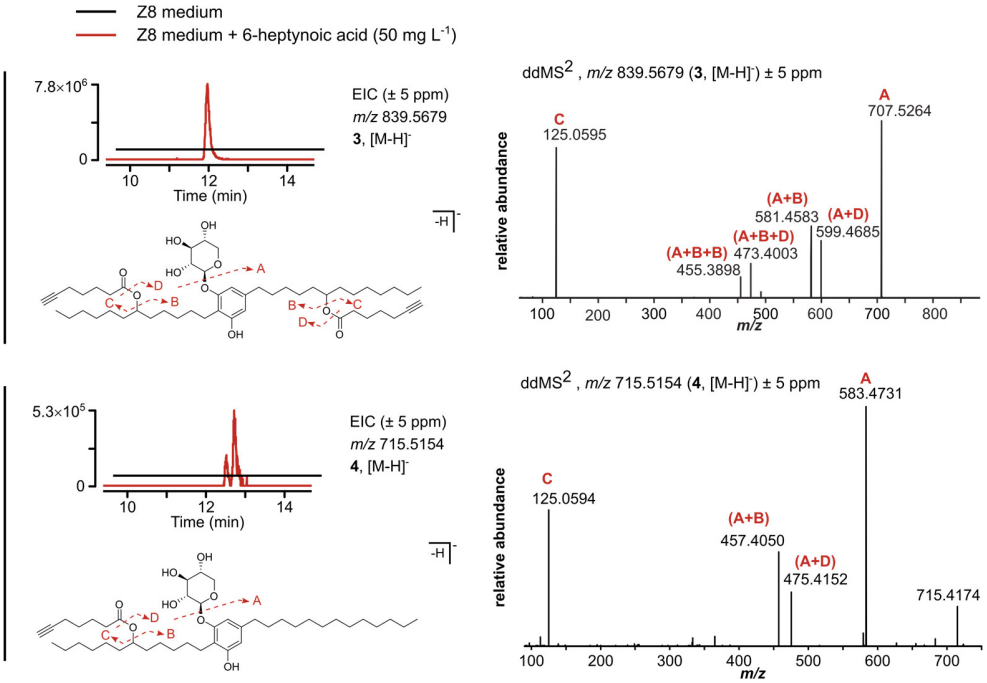


Figure 11. Formation of bartoloside A and G esters upon supplementation of *S. salina* LEGE 06099 with 6-heptynoic acid. Here we present the proposed chemical structures and respective annotated LC-HRESIMS/MS data for compounds **3** and **4**.

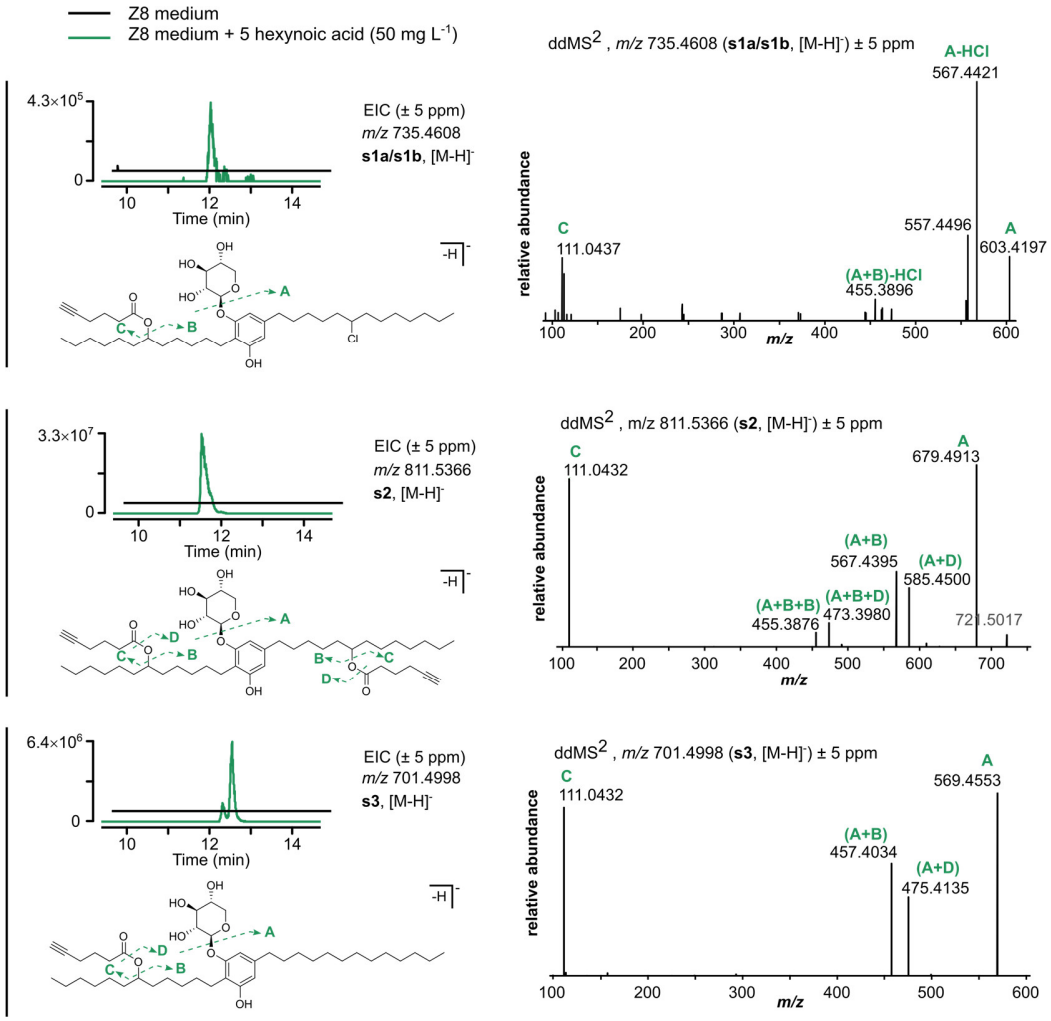


Figure 12. Formation of bartoloside A and G esters upon supplementation of *S. salina* LEGE 06099 with 5-hexynoic acid. Here we present the proposed chemical structures and respective annotated LC-HRESIMS/MS data for compounds **s1a/s1b, s2** and **s3**.

2. Isolation of branched bartolosides formed under supplementation

The previously performed experiments showed that supplementation with 5-hexynoic and 6-heptynoic acids did not generate terminal alkyne moieties in the DAR moiety of the bartolosides but instead led to the formation of new compounds. To unambiguously establish the chemical structure of the newly observed metabolites we sought to isolate and purify the major metabolites found under supplementation with fatty acids.

To fulfill our objective, we performed a 20 L cultivation of *S. salina* LEGE 06099 under supplementation with 6-heptynoic acid at a concentration of 50 mg L^{-1} in order to try to obtain a

sufficient amount of compound for structure elucidation. The harvested freeze-dried biomass (5.6 g, d. w.) was extracted by repeated percolation using a mixture of CH_2Cl_2/MeOH (see "Materials and Methods" section) resulting in a crude extract of 754.6 mg. Subsequent fractionation was achieved by normal phase flash chromatography using a gradient of increased polarity from hexane to ethyl acetate (EtOAc) to methanol (MeOH). The fractionation originated eight fractions (1-8) with similar TLC (thin layer chromatography) profiles. Fractions 4 (13.5 mg), 5 (102.2 mg) and 6 (14.5 mg) eluting with a 2:3 mixture of EtOAc/hexane were analyzed by LC-HRESIMS and indicated the presence of the desired *m/z* corresponding to the hypothetical branched bartolosides A and G (*m/z* 839.5678 $[M-H]^-$ and 715.5154 $[M-H]^-$, respectively). At this stage, depending on the relative abundance of each compound, we could have pooled the fractions together but considering that both compounds were mostly present in fraction 5 (figure 13) we decided to only go forward with this sample for further fractionation.

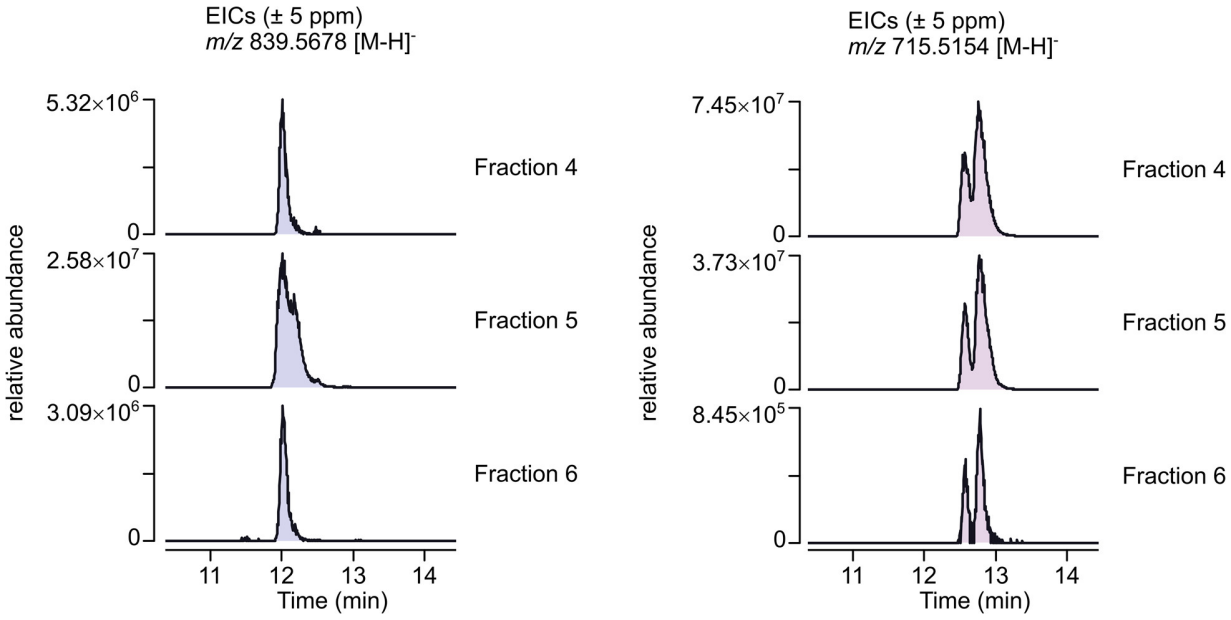


Figure 13. Extracted Ion Chromatograms (EICs) of *m/z* 839.5678 $[M-H]^-$ (compound **3**) and *m/z* 715.5154 $[M-H]^-$ (compound **4**) from the LC-HRESIMS of fractions obtained from the flash chromatography performed for the isolation of **3** and **4**.

The isolation of the desired metabolites involved a new chromatographic step using a reverse-phase semipreparative HPLC which allows a more efficient peak collection since we can track the separation progress in real time with the help of a PDA (photodiode array) detector. The HPLC program was adapted to allow the proper resolution of all peaks. Initially we tested an isocratic gradient of 80% acetonitrile in water during 25 minutes but peak separation was insufficient. After increasing the acetonitrile concentration to 90% we verified a more efficient separation with the

advantage of a faster fractionation since the compounds were eluting faster due to the decreased polarity of the mixture. Finally, we decided to use an isocratic gradient of 92% acetonitrile in water during 25 minutes followed by a washing step with 100% acetonitrile (see "Materials and Methods" section), affording 10 subfractions. All subfractions were submitted to ^{1}H NMR to evaluate their composition and purity. Fraction 5.6 (48.0 mg, figure 14), eluting at 13.5-15.0 minutes, was spectroscopically pure (> 95%, ^{1}H NMR) and showed characteristic signals of bartolosides regarding resorcinol and xylosyl protons.[81-82] Additionally, fraction 5.10 (4.3 mg, figure 14), with a retention time of 24.0-25.0 minutes, was also pure and presented a similar ^{1}H NMR profile to the subfraction 5.6 and consequently to the bartolosides.

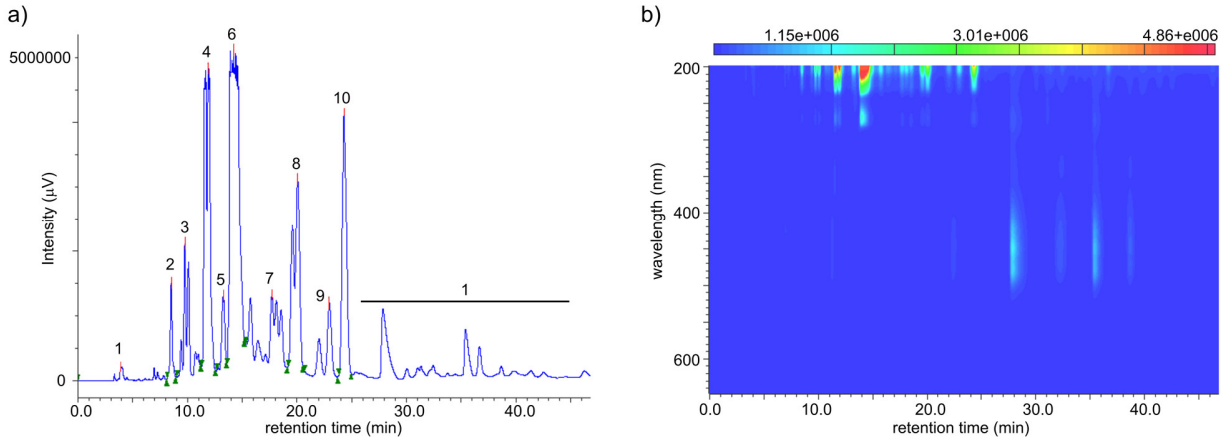


Figure 14. Reverse phase HPLC chromatogram (a) and PDA heat map spectra (b) of derived from the semipreparative chromatography of fraction 5 performed for the isolation of compounds **3** and **4**.

3. Structural elucidation of esterified bartolosides **3** and **4**

The ^{1}H NMR analysis of both 5.6 and 5.10 subfractions indicated that they were suitable for further structural elucidation. Therefore, both samples were submitted to 1D ^{13}C NMR and 2D NMR (HSQC, HMBC and COSY) as well as HRESIMS/MS analysis in order to unequivocally establish their chemical identity.

*bartoloside A-17,29-diyl bis(hept-6-ynoate) (**3**)*

HRESIMS analysis of compound **3** (isolated as an orange oil) showed a $[M-H]^-$ peak at *m/z* 839.5685, compatible with a molecular formula of $C_{50}H_{80}O_{10}$ (calculated for $C_{50}H_{79}O_{10}^-$, *m/z* 839.56867) and 11 degrees of unsaturation. The ^{1}H and ^{13}C NMR spectra showed typical bartoloside resonances, namely for the aromatic ring (δ_C 155.8, 154.6, 141.9, 116.3, 110.4 and

107.8) and xylosyl (δ_C 101.4, 72.6, 74.9, 69.7 and 64.5) (annex table 1). Characteristic resonances for two ester/acid carbonyls (δ_C 174.1 and 173.9) and two alkyne moities (δ_C 68.8, δ_H 1.95) were also observed. The HRESIMS derived molecular formula and isotope pattern of **3** dictated the absence of halogens in the molecule. The molecular formula had an additional $C_{14}H_{18}O_4$ when compared to the major bartoloside **1** produced by *S. salina* LEGE 06099 (bartoloside A) consistent with the incorporation of two 6-heptynoic acid moieties. This initial analysis strongly suggested that compound **3** could correspond to a bartoloside derivative where the chlorinated positions in **1** were substituted by the supplemented 6-heptynoic acid. To obtain further insights regarding the structure of **3**, we performed 2D NMR experiments (HSQC, HMBC, COSY). Correlational data for the glycosylated dialkylresorcinol moiety were highly similar to those of the bartolosides and it was also possible to deduce the structures of two intact 6-heptynoate spin systems from COSY and HMBC (figure 15). Notably, two conspicuous protons resonating at δ_H 4.87 (H-17 and H-29) in **3** are not found in any of the previously reported bartolosides. Each of these protons was positioned in the middle of an alkyl chain (figure 15, annex table 1) and showed HMBC correlations to the carbonyls C-1' (δ_C 174.1) and C-1'' (δ_C 173.9). Strong deshielding indicated that these were oxymethine protons and established the directionality of the ester linkages. HRESIMS/MS data, which show fragments corresponding to 6-heptynoate (*m/z* 125.05 $[M-H]^-$, figure 11) are in accordance with this observation. Additionally, the resonances for the chlorinated carbons in bartoloside A (δ_C 64.9, δ_H 4.01)[81-82] are not present in **3**, further implying ester linkages in positions 17 and 29. While it was not possible to fully connect the oxymethines to the dialkylresorcinol rings, the HRESIMS/MS data of in-source fragments of **3** (figure 16) matches the data for bartoloside A (**1**) and other bartolosides in our previous structural elucidation efforts.[81-82] As such, we propose that the esterified positions in **3** correspond to the chlorinated positions in **1**.

Optical rotation, infrared (IR) and ultraviolet (UV) fingerprint:

$[\alpha]$ $^{25}_D$ -5.0 (c 0.60, MeOH); IR (thin film) υ_{max} 3398, 3308, 2923, 2857, 1726, 1703, 1590, 1428, 1052, 1033, 1017, 627 cm^{-1}. UV (MeOH) λ_{max} (log ε) 217 (3.2), 221 (3.2), 273 (3.0).

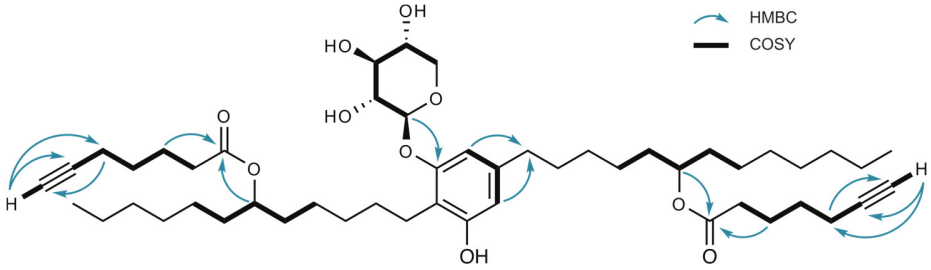


Figure 15. Key HMBC and COSY correlations of bartoloside A-17,29-diyl bis(hept-6-ynoate) (**3**)

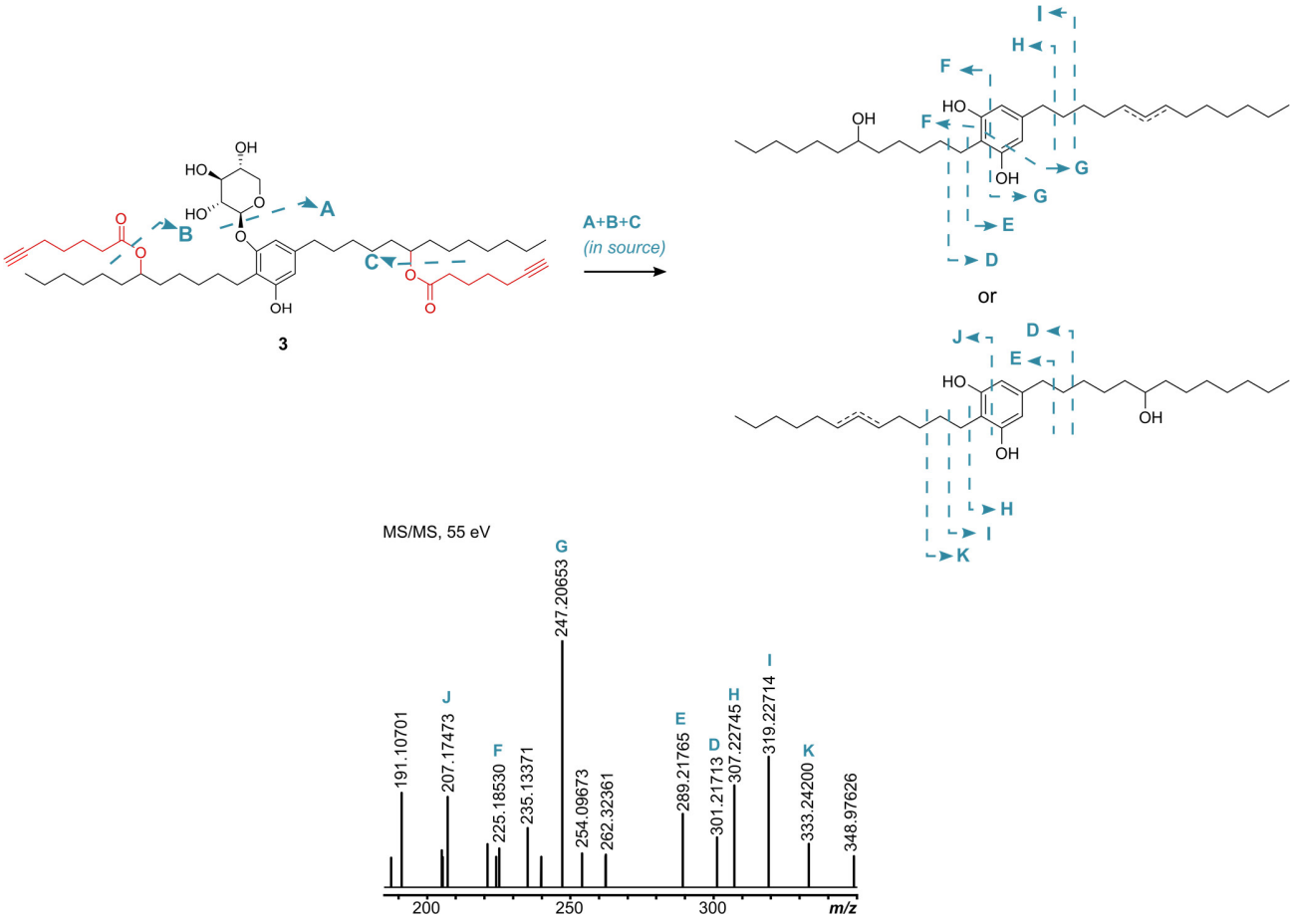


Figure 16. HRESIMS/MS spectra for the in-source (collision energy of 90 eV) generated fragment of **3** providing the required confirmation and validation of the chemical structure of compound **3**.

*bartoloside G-17-yl hept-6-ynoate (**4**)*

HRESIMS data of compound **4** (isolated as a light orange oil) showed a prominent peak at *m/z* 715.5161 $[M-H]^-$ which was compatible with a molecular formula of $C_{43}H_{72}O_8$ (calculated for $C_{43}H_{71}O_8^-$, *m/z* 715.5162) and 8 degrees of unsaturation. The ^1H and ^{13}C NMR data (annex table 2) were highly similar to those from compound **3**, notably featuring a carbonyl resonance (δ_C 174.1) and an alkyne moiety signal (δ_C 68.8, δ_H 1.95). Additionally, the typical bartoloside resonances for the aromatic ring (δ_C 155.7, 154.3, 142.5, 116.1, 110.4, 108.1) and xylosyl group (δ_C 101.4, 72.6, 74.9, 69.7 and 64.5) were present. The HRESIMS-derived molecular formula and isotope pattern of **4** indicated the absence of halogens in the molecule and an additional $C_7H_9O_2$ when compared to another bartoloside produced by *S. salina* LEGE 06099 (bartoloside G, **2**). This would be consistent with **4** resulting from the incorporation of one 6-heptynoic acid moiety into **2**. Two-dimensional NMR experiments (HSQC, HMBC and COSY) showed the presence of an intact 6-heptynoate-derived spin system from COSY and HMBC correlations (figure 17). The HMBC correlations for H-17 (δ_H 4.88) with C-1' (δ_C 174.1) indicated an identical ester linkage to that of **3**, at the C-17 (δ_C 74.7) position (figure 17). In line with these observations, the HRESIMS/MS data for **4** showed a 6-heptynoate-derived fragment (*m/z* 125.05 $[M-H]^-$, figure 11). By isolating and carrying out HRESIMS/MS analysis on an in-source fragment of **4**, we obtained diagnostic fragments indicating a fully aliphatic alkyl chain with 13 carbons (*m/z* 249.22 and *m/z* 291.23, figure 18), as well as fragments consistent with an alcohol group in the C12 alkyl chain (e.g. *m/z* 225.18 and *m/z* 319.23, figure 18). We thus propose that metabolite **4** is a derivative of the naturally occurring **2**, in which the chlorinated position is substituted by a 6-heptynoate moiety.

Optical rotation, infrared (IR) and ultraviolet (UV) fingerprint:

$[\alpha]$ ^{25}D -4.6 (c 0.65, MeOH); IR (thin film) υ_{max} 3410, 3313, 2923, 2855, 1731, 1704, 1590, 1428, 1054, 1033, 1014, 627 cm^{-1}; UV (MeOH) λ_{max} (log ε) 216 (3.1), 221 (3.1), 273 (2.8).

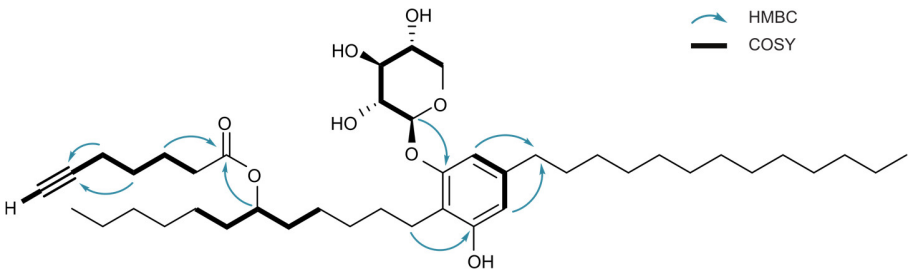


Figure 17. Key HMBC and COSY correlations of bartoloside G-17-yl hept-6-ynoate (**4**).

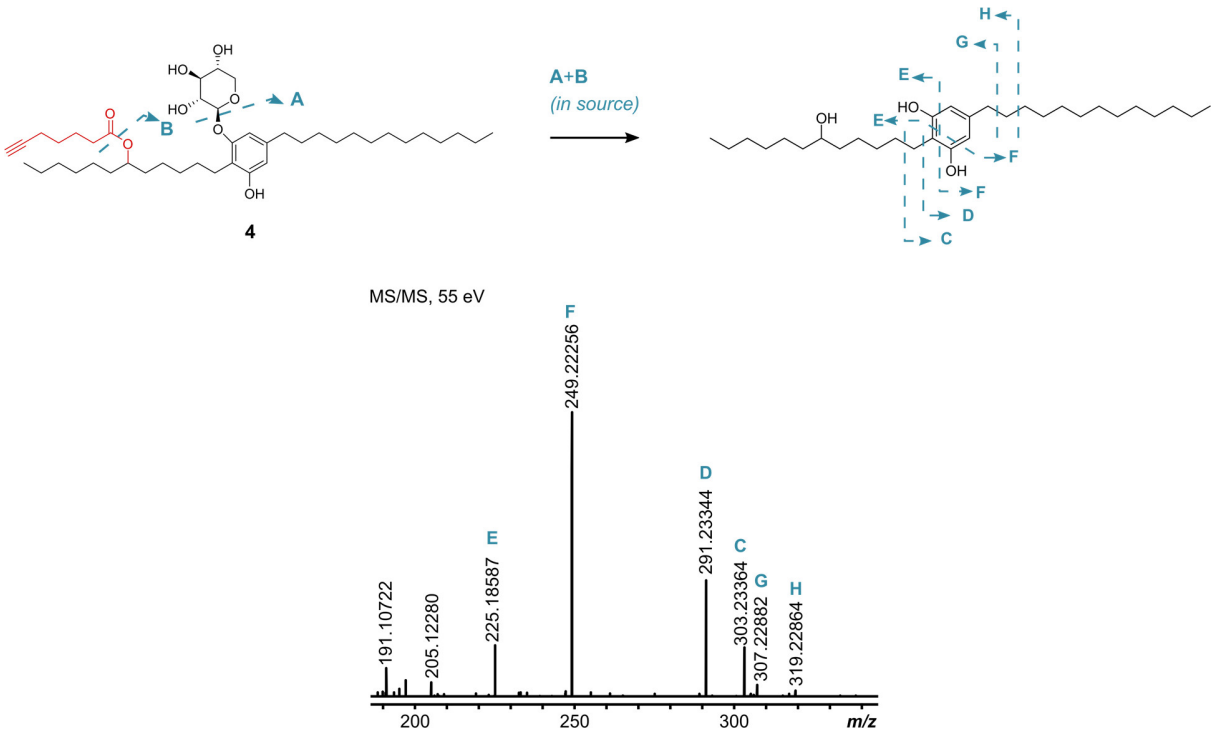


Figure 18. HRESIMS/MS spectra for the in-source (collision energy of 65 eV) generated fragment of **4** providing the required confirmation and validation of the chemical structure of compound **4**.

4. Incorporation of additional substrates into the bartolosides

We hypothesized whether additional substrates could be incorporated in the same fashion as the terminal alkyne-containing fatty acids supplemented. With this in mind, we set out to investigate which type of substrates could be incorporated in the bartolosides alkyl halide chains.

The previously used *S. salina* LEGE 06099 strain was exposed to several substrates containing not only carboxylate moieties (7-bromoheptanoic, trans-3-hexenedioic, butyric, caprylic, lauric and palmitic acids) but also other related nucleophiles commonly used in organic synthesis (hexylalcohol and hexanamide). After supplementing the cultures, the cells were harvested, extracted as carefully described in the “Materials and Methods” section and analyzed by LC-HRESIMS. As expected, the results showed the formation of corresponding mono- and diesters of both bartoloside A (**1**) and G (**2**) with the fatty acids supplemented (figure 19). Since trans-3-hexenedioic acid contains two carboxylate moieties we hypothesized that the branching could occur between two bartolosides or even inside the same molecule through the esterification of both alkyl halide chains to the same substrate. However, we could not detect any ion that could

correlate with branched versions of bartoloside A and G with hexylalcohol, hexanamide nor trans-3-hexenedioic indicating selectivity towards fatty acids. Overall, our results show that exogenously provided fatty acids are converted *in vivo* into esterified versions of bartolosides by *S. salina* LEGE 06099.

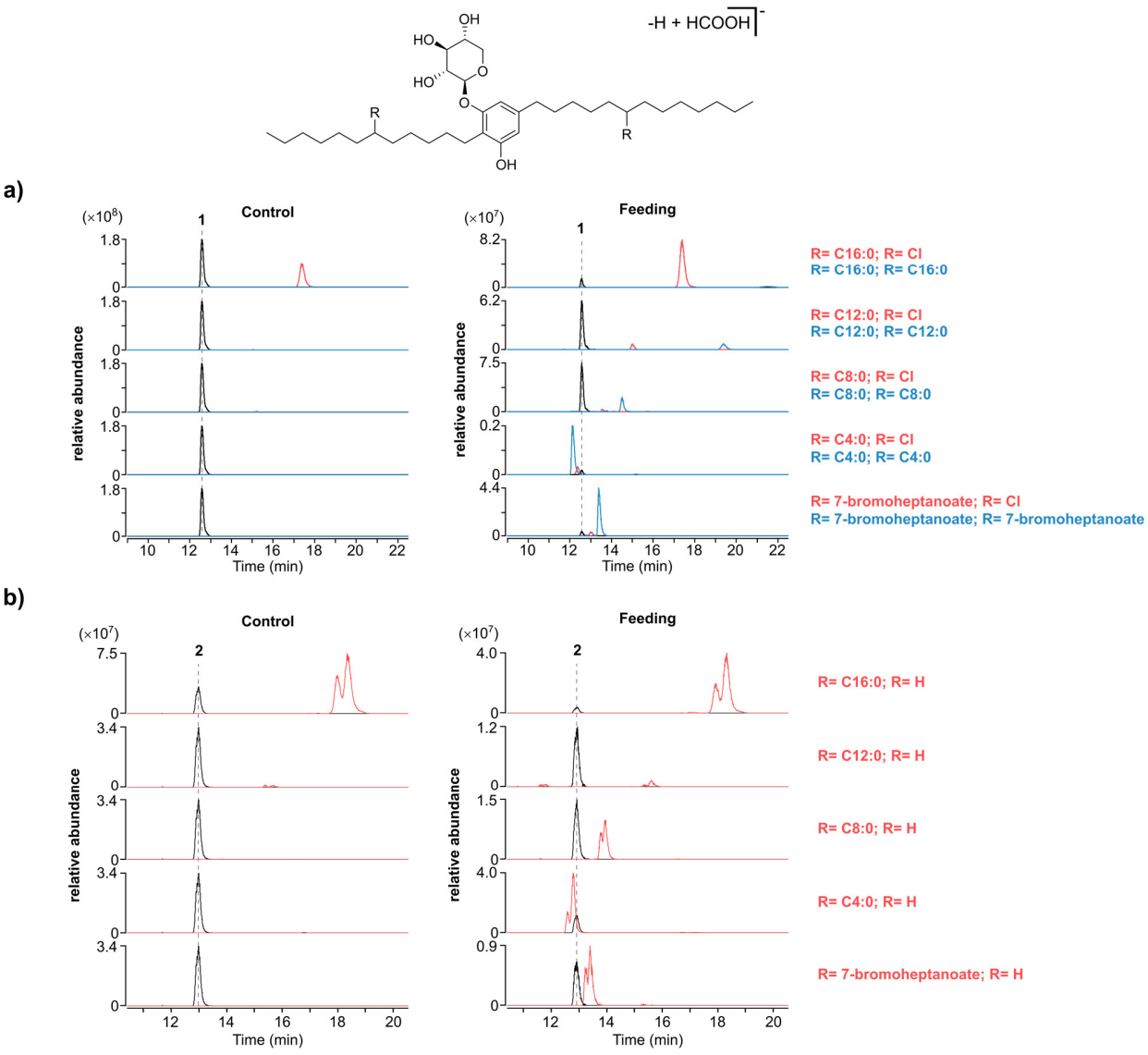


Figure 19. Bartoloside esters formed in *S. salina* LEGE 06099 cells upon supplementation with different fatty acids commonly present in bacteria and with the halogenated fatty acid 7-bromoheptanoate. LC-HRESIMS of extracted ion chromatograms (EICs) of $[M-H+HCOOH]^-$ adducts of the proposed bartoloside A esters (a) and bartoloside G esters (b). Controls without supplementation are depicted in the left panel and supplemented cultures in the right panel.

Interestingly, while analyzing the LC-HRESIMS data of our non-supplemented controls we verified the presence of two abundant compounds with *m/z* values and retention times consistent with both bartoloside A monopalmitate (**5a/5b**) and bartoloside G palmitate (**6**). Such findings led us to hypothesize that esterified bartolosides occur naturally in the cyanobacteria and not only while under supplementation with exogenous fatty acids. Therefore, we set out to investigate whether the detected species was in fact the branched version of bartoloside A monopalmitate (**5a/5b**).

5. Detection of naturally occurring bartoloside esters

The previously obtained LC-HRESIMS results of the control samples presented in figure 18 directed us to search for the naturally occurring bartoloside A monopalmitate (**5a/5b**) ester (*m/z* 879.6486 $[M-H]^-$). Further analysis of the CH_2Cl_2/MeOH (2:1) cellular extract of *S. salina* LEGE 06099 without supplementation used in the chapter revealed the presence of a *m/z* consistent with the diesterified version of bartoloside A with palmitic acid (**7**, *m/z* 1099.9177 $[M-H]^-$). The higher retention time corroborated with the hypothetical structure due to the putative compound's lower polarity owing to the four alkyl chains.

With these data in our hands we decided to perform the mass fragmentation of the detected ions (*m/z* 879.6486 $[M-H]^-$ and 1099.9177 $[M-H]^-$). The LC-HRESIMS/MS data revealed the presence of intact fragments of palmitic acid (*m/z* 255.2330 $[M-H]^-$) as well as fragments corresponding to the previously presented data of the bartoloside A core (figure 20). Fragmentation of the hypothetical bartoloside G palmitate (**6**, *m/z* 845.6876 $[M-H]^-$) also revealed the presence of palmitic acid and bartoloside G fragments, without a chlorine atom, in line with what had been observed for the previously isolated compound **4**.

Knowing this, we extensively examined the LC-HRESIMS data for additional esterified versions of bartolosides A and G with fatty acids known to be present in cyanobacterial cells.[91] This revealed 20 additional peaks matching with *m/z* of branched versions of **1** and **2** with fatty acids with varying chain length and saturation, which were then analyzed by LC-HRESIMS/MS to confirm the presence of corresponding fragments (figure 21, annex figures 1 and 2). Putative bartoloside A and G esters **5a/5b** and **6**, respectively, seem to be the most abundant ones as verified by the integration of the corresponding peak area (figure 22).

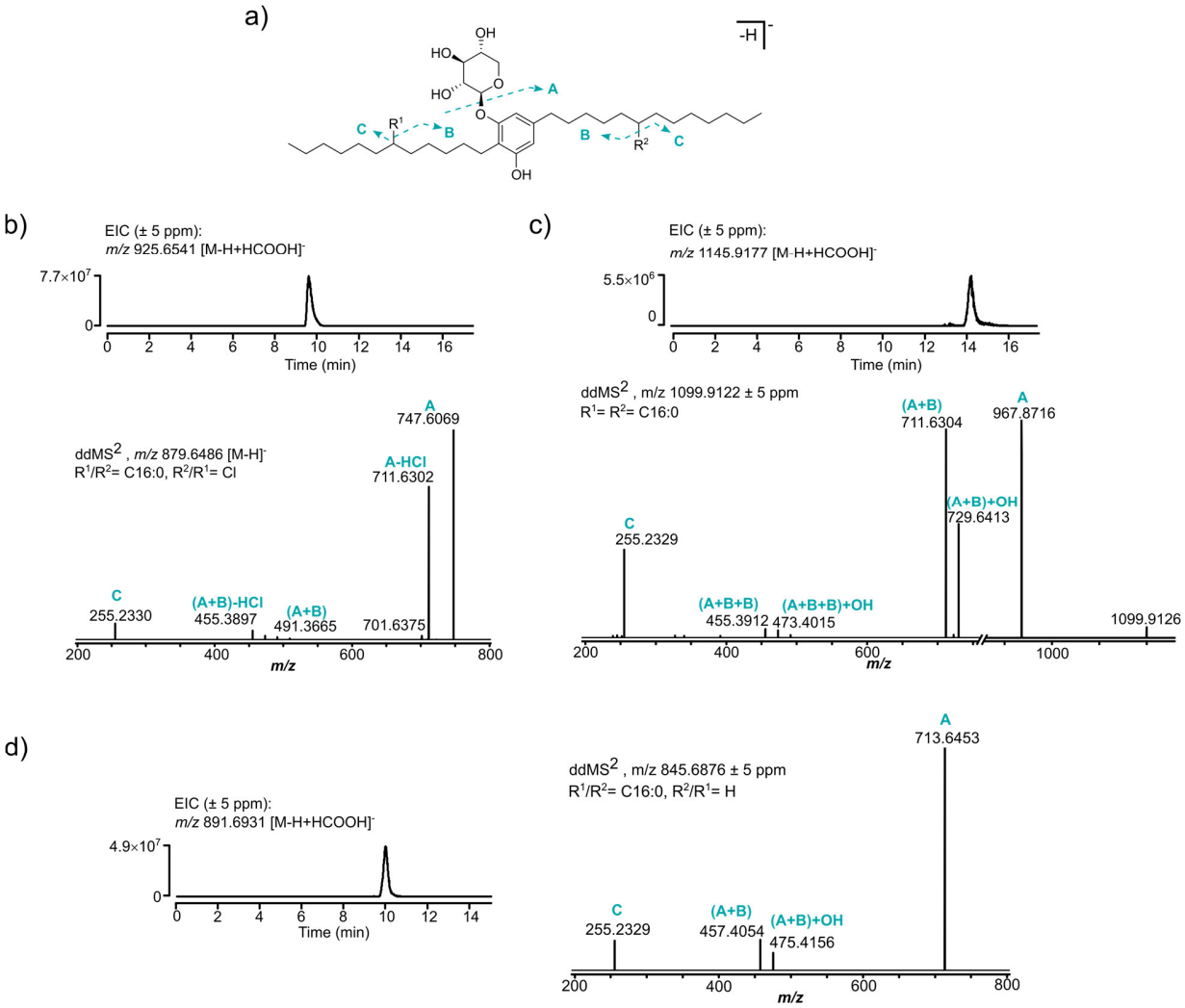


Figure 20. Detection and HRESIMS/MS based structural assignment of naturally occurring bartoloside A and G esters of palmitate in *S. salina* LEGE 06099. a) Annotated MS/MS fragments for the structures of bartoloside A mono- (**5a/5b**) and diester (**7**) and bartoloside G palmitate (**6**). Depicted are the extracted Ion Chromatograms (EICs) for the $[M-H+HCOOH]^-$ ions and MS/MS spectra for the $[M-H]^-$ species of detected bartoloside A esters **5a/5b** and **7** (b and c, respectively) as well as bartoloside G ester **6** (d).

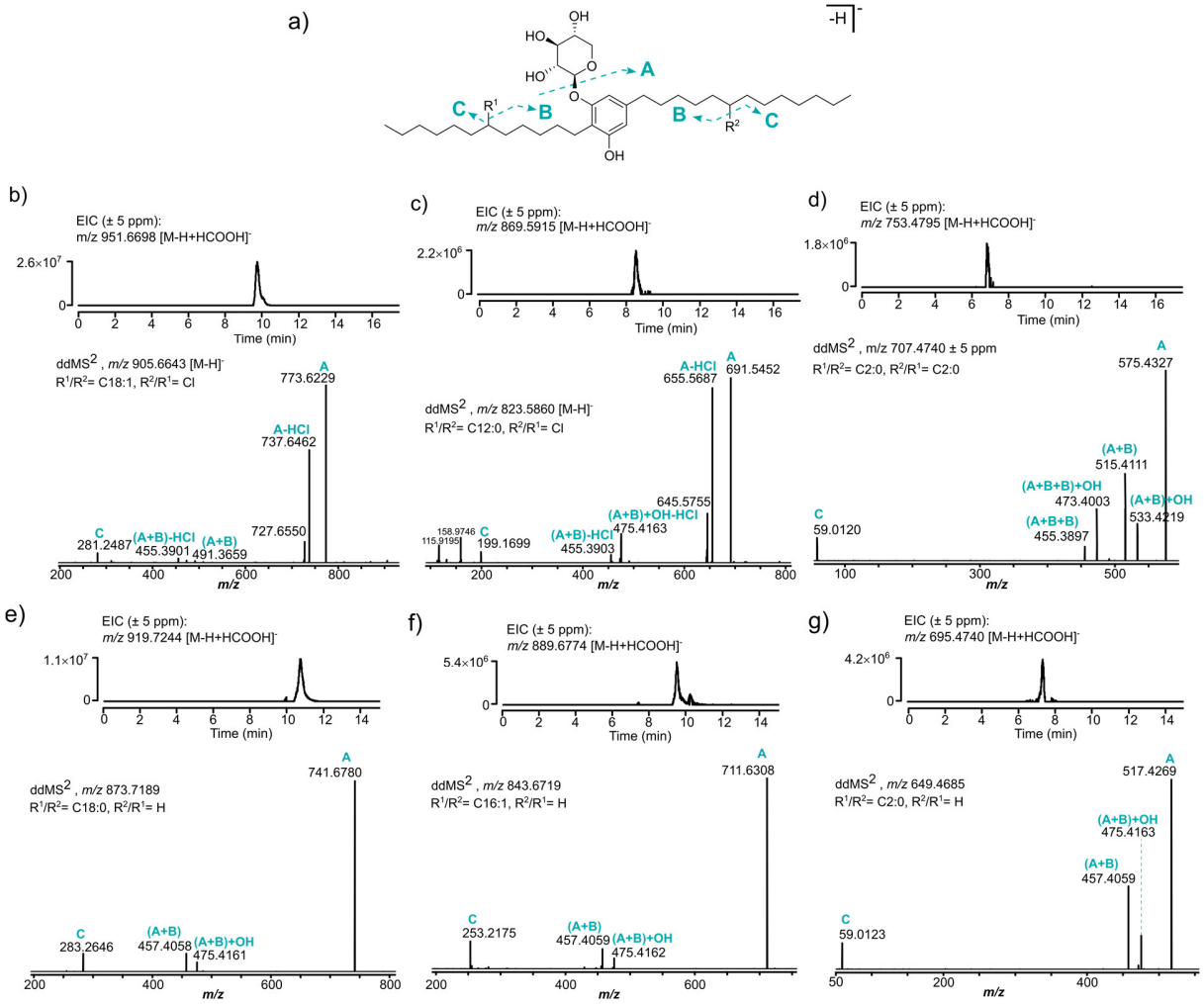


Figure 21. Detection and HRESIMS/MS based structural assignment of naturally occurring bartoloside esters produced by *S. salina* LEGE06099. a) Annotated MS/MS fragments for the structures of bartoloside A and G esters. b-d) Extracted Ion Chromatograms (EICs) for the $[M-H+HCOOH]^-$ ion and MS/MS spectra for the $[M-H]^-$ specie of some of the detected naturally occurring bartoloside A esters. e-g) Extracted Ion Chromatograms (EICs) for the $[M-H+HCOOH]^-$ ion and MS/MS spectra for the $[M-H]^-$ specie of some detected naturally occurring esterified bartoloside G.

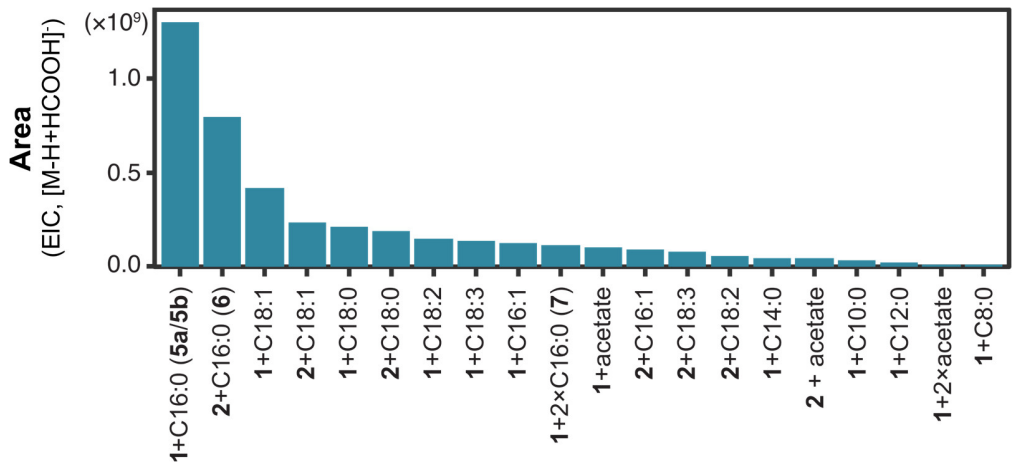


Figure 22. Relative abundance of naturally occurring bartoloside esters found in *S. salina* LEGE06099 cells.

Even though *S. salina* LEGE 06099 produces other bartolosides we could not detect any peaks corresponding to hypothetical esterified versions of such metabolites due to their low abundance in the used strain. [81] On the other hand, proving that the detected peaks were indeed bartoloside esters required purification and structure elucidation.

6. Isolation of bartoloside A monoester **5a**

To undoubtedly establish that the LC-HRESIMS detected compounds were indeed branched versions of bartolosides we cultivated 20 L of *S. salina* LEGE 06099 without the supplementation of exogenous fatty acids. After extracting the organic compounds from the freeze-dried biomass (574.31 mg d.w. crude extract) using a mixture of CH_2/Cl_2:MeOH (2:1) we performed another flash chromatography to fractionate our sample (see “Materials and Methods” section).

At this point, we had to adapt the solvent concentration since our target compounds are more apolar due to their long aliphatic chains. Since we used a normal phase chromatography and started with apolar solvents (hexane) our compounds were predicted to have eluted earlier than compounds **3** and **4**. Therefore, we tried to extend the separation using higher volumes of the more apolar mixtures (until 40% EtOAc, 60% hexane). The derived samples were pooled according to their TLC profiles originating 23 fractions (1-23) and analyzed by LC-HRESIMS. The derived data showed a very weak separation of compounds when compared to the flash chromatography performed in the isolation of **3** and **4** since the target compound (*m/z* 879.6486 $[M-H]^-$) had eluted throughout several fractions. Fraction 15 (18.81 mg) contained the highest relative abundance of the desired peak (putative bartoloside A monopalmitate **5a**/**5b**, *m/z*

879.6486 [M-H]$^-$)and was selected for further fractionation through semipreparative reverse phase HPLC (figure 23).

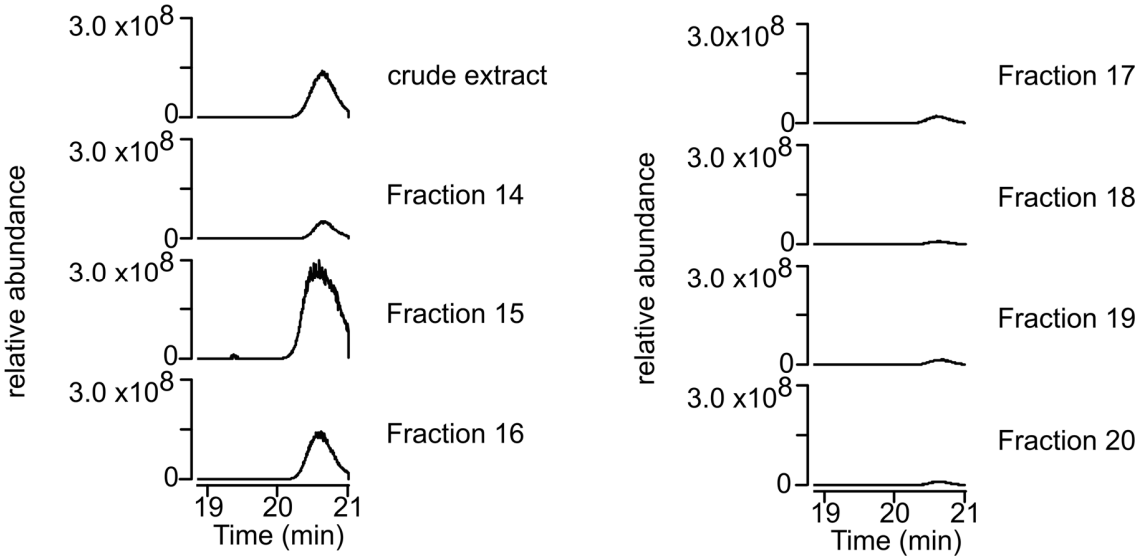


Figure 23. Extracted Ion Chromatograms (EICs) of putative esterified bartoloside A monopalmitate *m/z* 879.6486 [M-H]$^-$. Depicted are fractions 14-20 derived from the flash chromatography performed with the crude extract obtained from 20 L of non-supplemented *S. salina* LEGE 06099 containing the putative esterified bartoloside A monopalmitate *m/z* 879.6486 [M-H]$^-$.

Since we had previously applied a concentration of 92% acetonitrile in water for the isolation of compounds **3** and **4** we decided to try and use a higher concentration of acetonitrile (99%) due to the non-polar nature of the target modified bartoloside **5a/5b**. However, this percentage was not sufficient to cause the elution of all compounds present in fraction 15 under 120 minutes (figure 24a). Therefore, we decided to try methanol for the mobile phase, which, although more polar than acetonitrile, is known to be more efficient in eluting compounds when applied in high proportion in C18 columns.[92] In fact, after adjusting the methanol concentration to an isocratic gradient of 98% MeOH in water we obtained a good separation in under 60 minutes (figure 24b). From the collected 11 peaks, subfraction 15.7 (retention time 32.1-34.5 minutes) contained one of the major peaks and likely the target species with *m/z* 879.6486 [M-H]$^-$. To validate this, we analyzed the eluting peak by HRESIMS/MS using direct injection into the mass spectrometer device. Again, the target *m/z* value was observed and fragments corresponding to palmitic acid as well as its neutral loss were detected along with bartoloside A dialkylresorcinol fragments (annex figure 4).

The obtained subfraction 15.7 (6.52 mg d.w.) was submitted to ^{1}H NMR so that we could verify its purity and forward it to 2D NMR analyses for the structural elucidation. However, the data showed that a minor bartoloside co-eluted (likely the structural isomer of bartoloside A monopalmitate **5b**) since the proton resonances were extremely close and overlapping (annex figure 3), thus further separation was necessary.

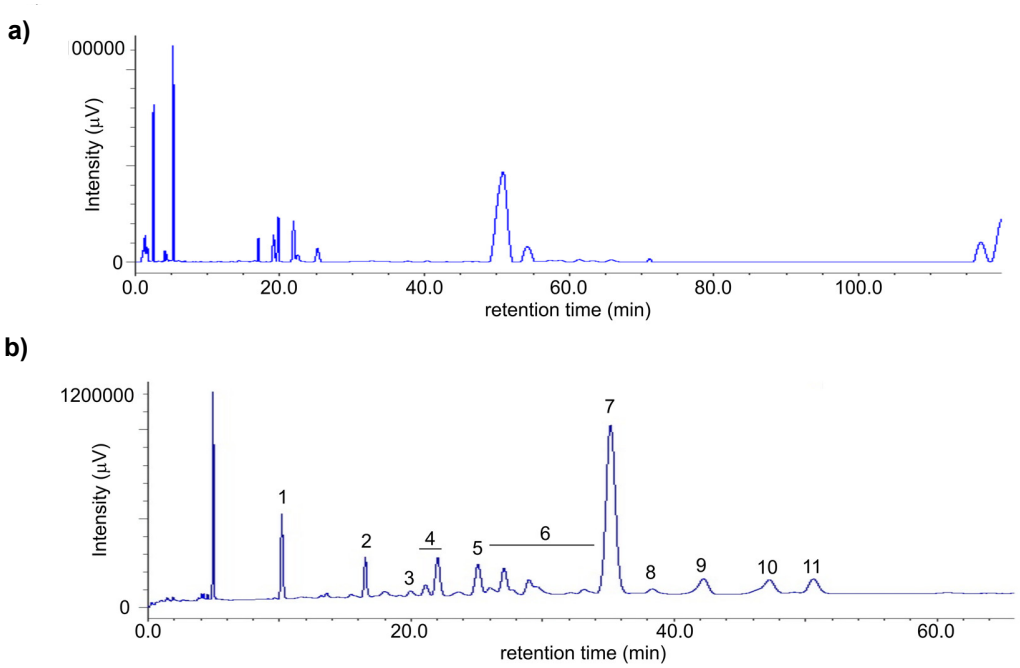


Figure 24. Chromatograms obtained after RP-HPLC of fraction 15 using 99% acetonitrile in water (a) and 98% methanol in water (b). the collected subfractions (15.1-15.11) are evidenced by the numbers on top of the peaks or peak region in b).

Although methanol was efficient for a faster separation of fraction 15 we decided to test acetonitrile since it promoted a much slower peak elution. In fact, after some program adjustments we verified that 98% acetonitrile in water was the ideal to cause the proper separation of subfraction 15.7, even though the elution was occurring only after 120 minutes. Fraction 15.7.2 (3.1 mg, retention time ~122 minutes, figure 25) was then submitted to ^{1}H NMR and since it was spectroscopically pure we proceeded with the 1D and 2D NMR analysis for the structural elucidation.

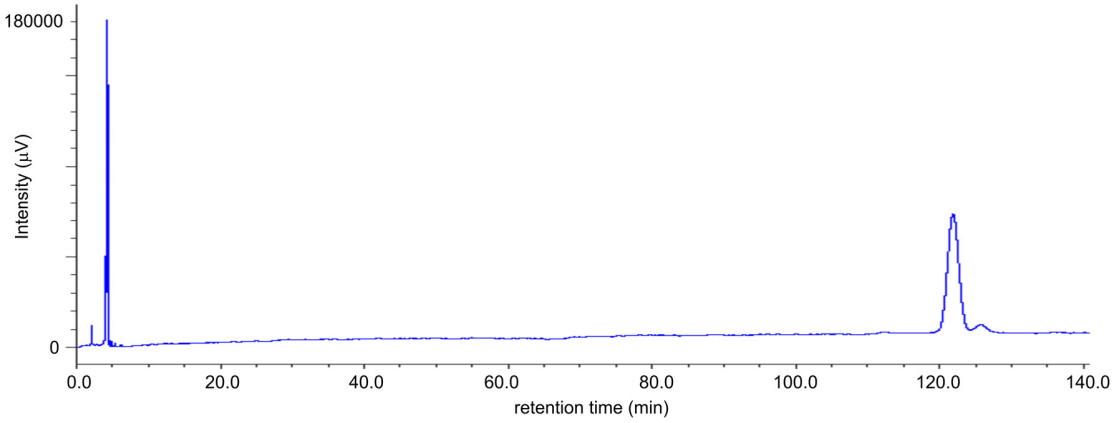


Figure 25. Chromatogram of purified subfraction 15.7.2 after a reverse phase semipreparative HPLC of fraction 15.7 using 98% acetonitrile in water. 15.7 using 98% acetonitrile in water.

7. Structural elucidation of bartoloside A monopalmitate **5a**

The HRESIMS performed for compound **5a** (isolated as a white glassy solid) presented a peak with *m/z* 879.6504 [M-H]- (calculated for $C_{52}H_{92}O_8Cl^-$ *m/z* 879.6499), compatible with a molecular formula of $C_{52}H_{93}O_8Cl$ and 6 degrees of unsaturation. The ^1H and ^{13}CNMR data indicated the presence of the characteristic aromatic resorcinol ring and xylosyl group observed in bartolosides from *S. salina* LEGE 06099 (annex table 3). A carbonyl resonance was also present (δ_C 174.7) as observed for bartoloside esters **3** and **4**. The additional $C_{16}H_{31}O_2$ (compared to **1**) was consistent with the esterification of one palmitic acid moiety with bartoloside A (**1**) and the concomitant loss of one chlorine atom (as seen for **3** and **4**). Support to this hypothesis was gained from HRESIMS/MS analysis of the in-source-formed species with *m/z* 473.40032, which showed diagnostic fragments that are consistent with a backbone similar to that of **1** (figure 26), as well as from 2D NMR experiments: an HSQC correlation between δ_H 3.88 and δ_C 64.5, likely corresponding to a chlorinated methine, was part of a spin system that expanded, on both directions into degenerate CH_2 resonances, indicated that this corresponded to a mid-chain chlorination, characteristic of the bartolosides (figure 27). HRESIMS/MS data was consistent with the presence of a palmitate-derived fragment (*m/z* 255.23 $[M-H]^-$, figure 20). The 1D and 2D NMR data showed an additional methyl group (compared to **1**) at the terminal position of an alkyl chain, as well as a bigger CH_2 envelope, both in line with the presence of a palmitate moiety in compound **5a**. The oxymethine proton resonating at δ_H 4.88 presented HMBC correlations to C-1' (δ_C 174.7) in the palmitate moiety and was itself part of another spin system that degenerated into CH_2 envelopes; hence it was likely positioned in the middle of the non-chlorinated alkyl chain. All these

observations supported our proposal of **5a** being a palmitate ester of **1**. Because **1** has two halogenated methine moieties, to clarify which chain contained the esterification in **5a**, we compared the ^{1}H NMR spectra for **5a** with the previously reported[81-82] data for **1**. Interestingly, we found that the multiplicity of H_2-12 (δ_H 2.58) changed from a complex multiplet in bartoloside A to a triplet in **5a**, while the multiplicity of H_2-24 (δ_H 2.50) remained unaffected (figure 28). As such, we propose that position 17, in the same alkyl chain as the H_2-12 benzylic protons, is esterified in bartoloside A-17-yl palmitate (**5a**). Further HRESIMS/MS analysis is consistent with this assignment (figure 26).

Optical rotation, infrared (IR) and ultraviolet (UV) fingerprint:

[α] $^{25}_{D}$ -9.2 (c 0.42, MeOH); IR (thin film) υ_{max} 3417, 2923, 2854, 2360, 2340, 1730, 1704, 1592, 1463, 1455, 1430, 1159, 1048, 1033 cm^{-1}; UV (MeOH) λ_{max} (log ε) 217 (3.3), 222 (3.3), 273 (2.9).

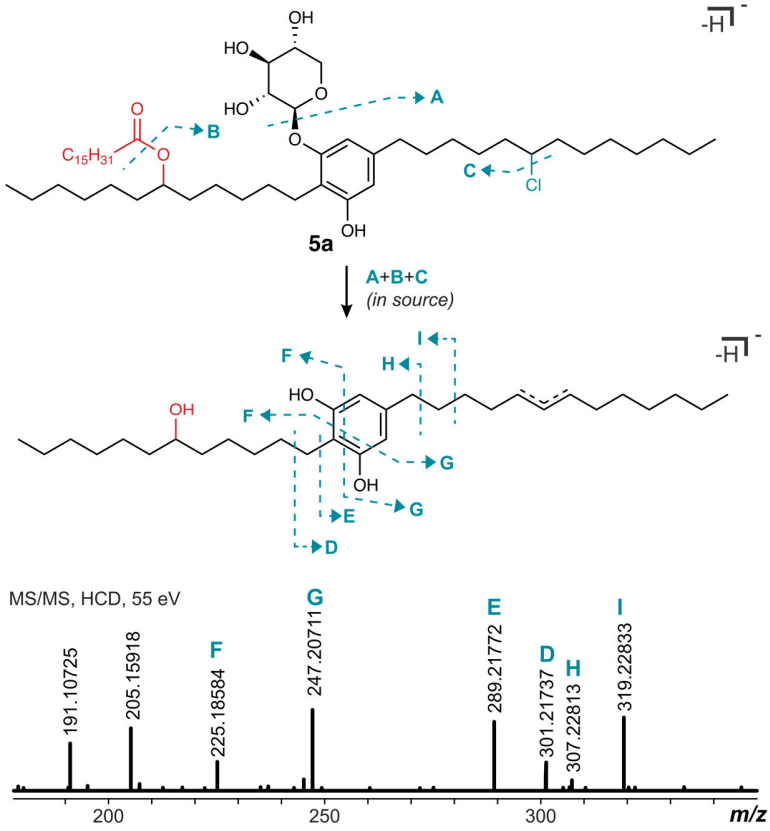


Figure 26. HRESIMS/MS spectra for the in-source (collision energy of 90 eV) generated fragment of **5a** providing the required confirmation and validation of the chemical structure of compound **5a**.

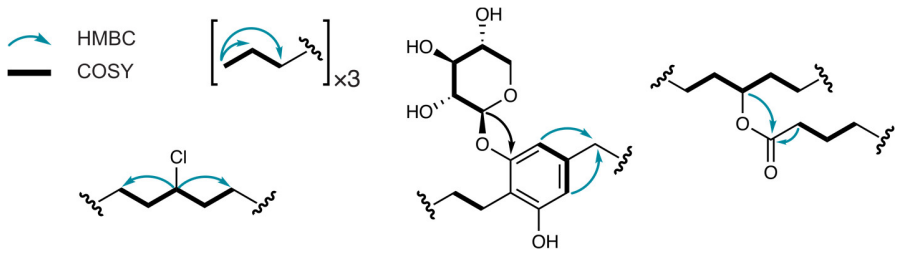


Figure 27. Key HMBC and COSY correlations of bartoloside A-17-yl palmitate (**5a**).

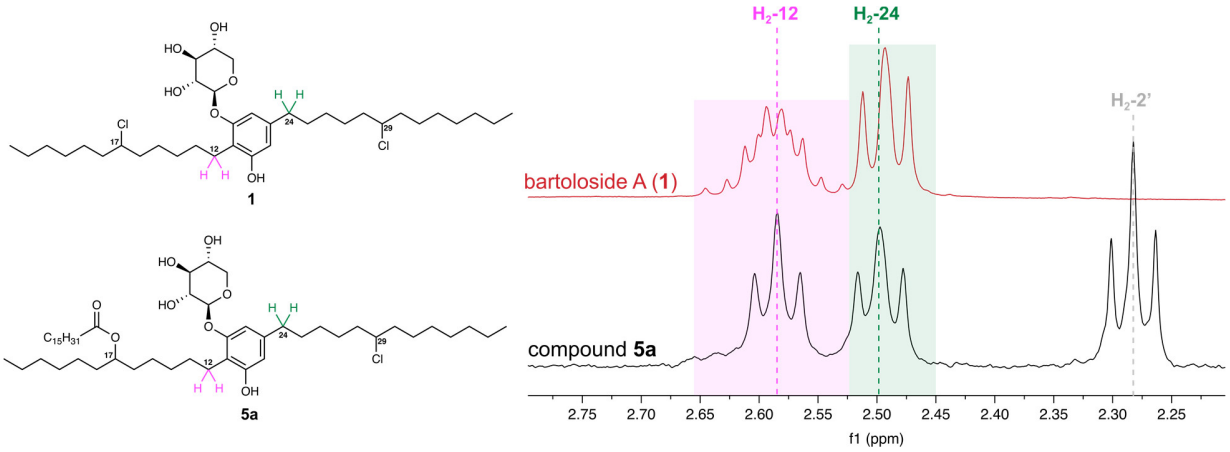


Figure 28. ^{1}H NMR-based assignment of which alkyl chain in **5a** is esterified with palmitic acid. Shown are the structures and regions of the benzylic proton resonances for bartoloside A (**1**, top) and **5a** (bottom). Resonance for protons of H_2-24 (green) were not affected while protons H_2-12 (pink) multiplicity was affected by the new palmitate moiety in **5a**.

8. Bioactivity of compounds **3**, **4** and **5a**

Even though the alkyne moiety was not incorporated in the expected fashion we proceeded with some bioactivity assays using our new compounds **3** and **4**, as well as the novel natural version of bartoloside A palmitate (**5a**). As previously reported, bartoloside A only had activity against the colorectal adenocarcinoma cell line HT-29[82] which was our starting point for the assays using compounds **3**, **4** and **5a** (figure 29). Additionally, since we had isolated decent amounts of compounds **3** and **4** we decided to also screen their bioactivity against other cell lines available (hCMEC/D3 and HCT116, figures 30 and 31, respectively). However, none of the compounds had activity against these cell lines, with the exception of compounds **3** and **4** which presented low activity against hCMEC/D3, a blood-brain barrier cell line that has been immortalized, (but not a cancer-derived cell line). Regarding naturally occurring ester **5a,** it seems that after the alkylation with palmitate bartoloside A loses the activity against HT-29. Furthermore, we exposed several

bacteria along with the fungus *Candida albicans* to our isolated compounds but no antimicrobial activity was detected for any of the isolated natural products (figure 32).

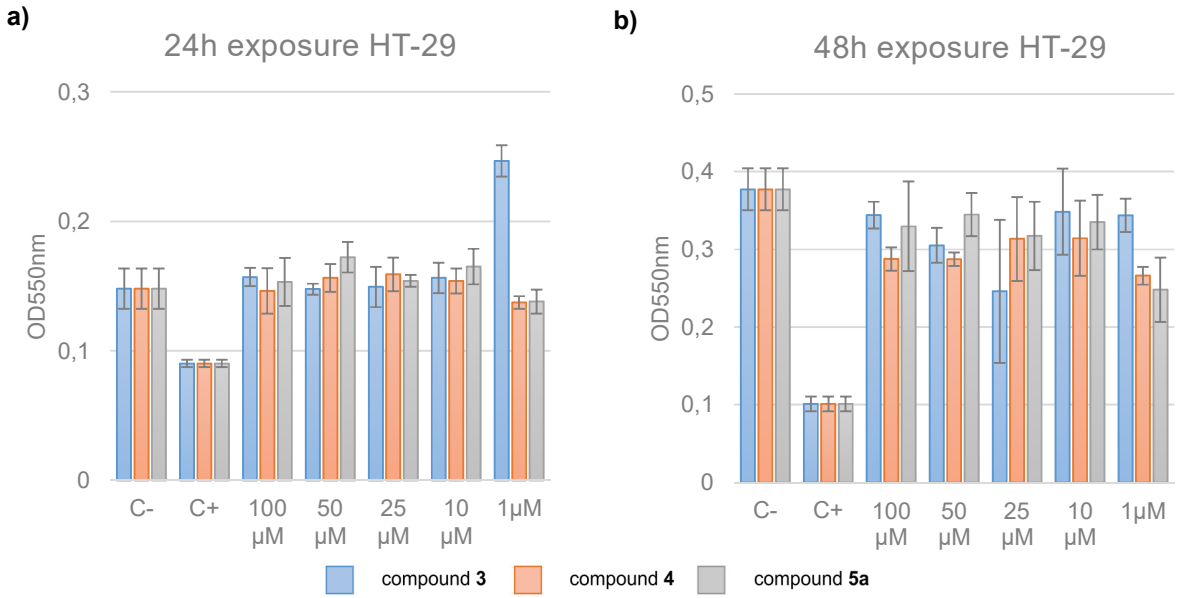


Figure 29. Bioactivity assay results of exposed HT-29 cells to compounds **3**, **4** and **5a** after 24h (a) and 48h (b). C+:1 µM staurosporine.

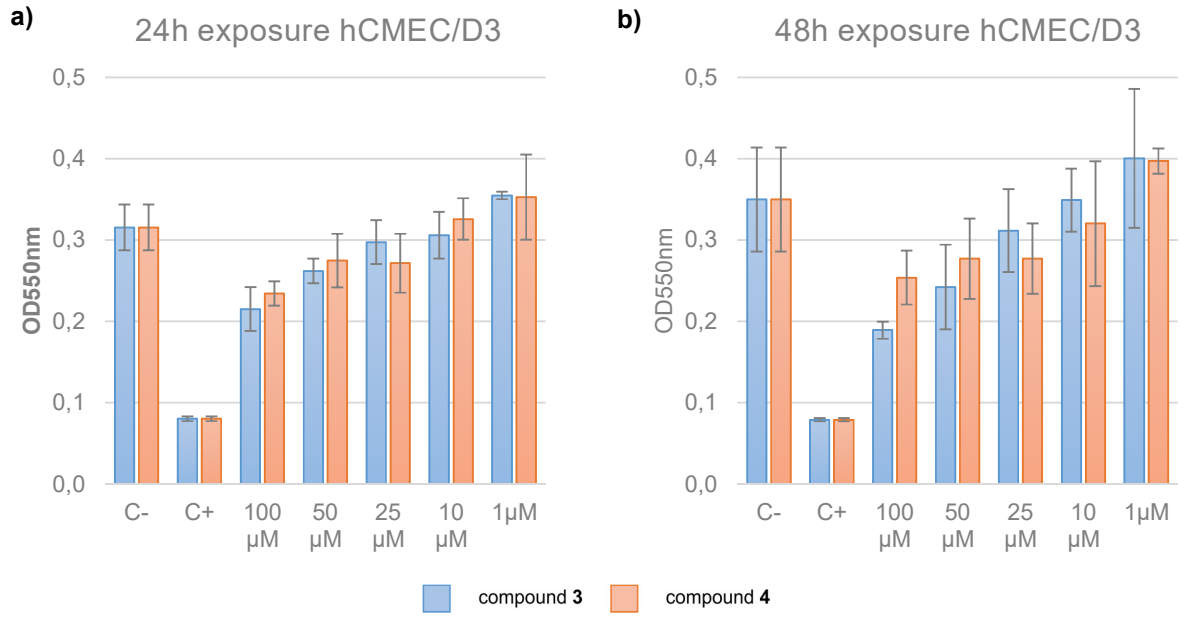


Figure 30. Bioactivity assay results of exposed hCMEC/D3 cells to compounds **3** and **4** after 24h (a) and 48h (b). C+:1 µM staurosporine.

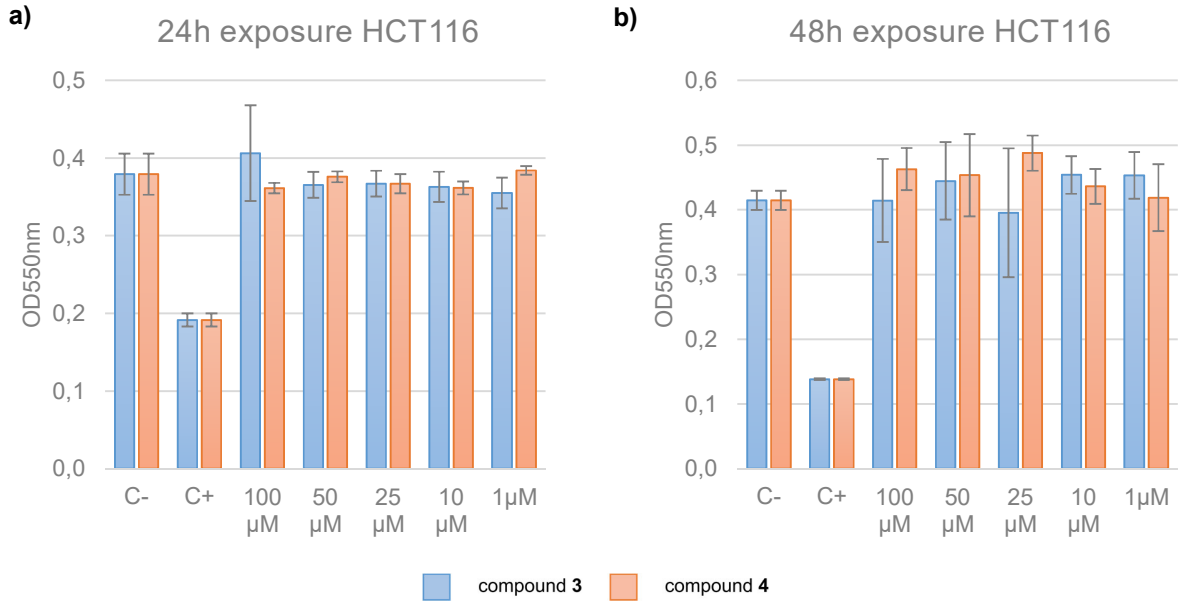


Figure 31. Bioactivity assay results of exposed HTC116 cells to compounds **3** and **4** after 24h (a) and 48h (b). C+:1 µM staurosporine.

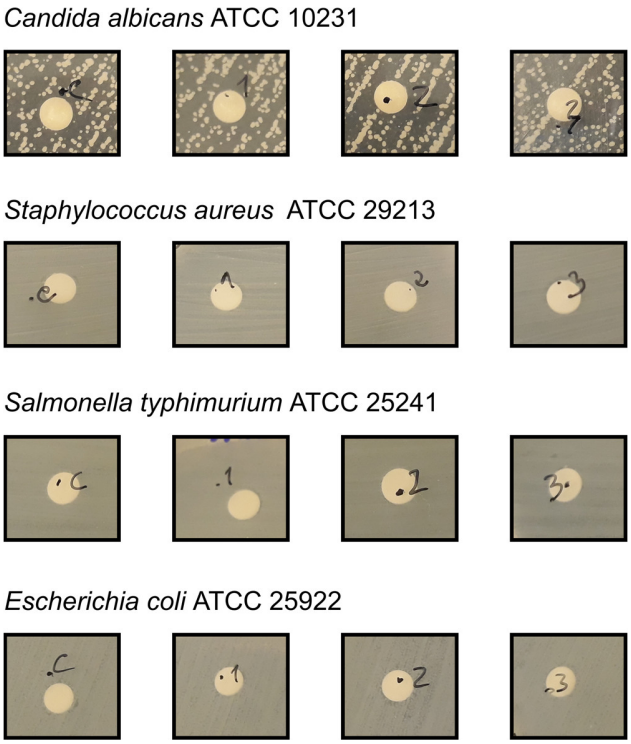


Figure 32. Tested bacterial and fungi strains. C: DMSO; 1- compound **3**, 2- compound **4**; 3- compound **5a**.

Having established that bartoloside esters occur in Nature, we focused on the biochemical aspects of the observed esterification. In organic synthesis, several methods have been described for the formation of carboxyesters, usually exploiting the nucleophilic attack of an alcohol onto the carboxylic carbon, as described by the Fisher-Speier esterification.[93] However, in the present reaction between bartolosides and fatty acids, the carboxylate seems to be acting as the nucleophile performing the attack onto the electrophilic alkyl halide chains of the bartolosides. Recent studies by Balskus and co-workers regarding CylK C-alkylation (see "Introduction" section) encouraged us to consider that BrtB, the only enzyme encoded in the *brt* cluster with no ascribed function, [81-82] could be responsible for the observed esterification.

9. Cloning of NStrep-BrtB

To be able to understand if the ester formation in the bartolosides was promoted by BrtB we first needed to clone it into a suitable model for protein expression, such as *Escherichia coli*. We attempted at using the genomic DNA (gDNA) from *S. salina* LEGE 06099 (isolated using the PureLink™ Genomic DNA Mini Kit, Invitrogen), however, after running a 1% agarose gel electrophoresis we verified that the DNA was degraded. After additional attempts, we faced the same results and decided to perform the gDNA extraction using the cetyl trimethylammonium bromide (CTAB) procedure (figure 33a).

Although the arguably most common procedure for recombinant protein expression and purification relies on the fusion of a polyhistidine tag (HisTag) into the gene sequence of the desired protein, we did not succeed in overexpressing the 6×HisTag version of BrtB. After using both *E. coli* BL21 and *E. coli* BL21(DE3)pLysS competent cells we decided to analyze the protein sequence of the recombinant protein. The codon adaptation index (CAI)[94] of the inserted gene was in fact relatively low (0.63) when crossed against the *E. coli* genome, indicating that we should try and use a different model for protein expression. Therefore, we tested the expression in *E. coli* Rosetta(DE3) since it contains a plasmid encoding for rare codons that could ameliorate transcription. The attempt failed again since we were not able to see any expression. At this point we started to consider protein solubility issues and a recently published study by the Balskus' with a detailed purification of CylK, (BrtB homolog).[90] In the new procedure, Balskus and co-authors described a protocol using a Strep-Tag©-recombinant version of CylK. This tag consists of an aminoacid sequence overhang (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) with binding affinity to a streptavidin analog, Strep-Tactin resin, working as an efficient affinity chromatography system.[95]

We decided to attempt the cloning and expression of Strep-Tag© fused BrtB. The primer design for the amplification of *brtB* was accomplished using both Geneious Prime 2019.0.4 and OligoCalc.[96] This step consisted on the addition of two overhangs complementary to the restriction enzymes SalI (5') and SacI(3'), which were the only enzymes in pET51b (+) (containing Strep-tag) that did not cut the internal sequence of *brtB,* originating a N-terminal Strep-Tag version of BrtB. Since the melting temperatures of the primers was considerably high and because we were using Phusion High-Fidelity DNA Polymerase (ThermoFisher), the PCR program consisted on a 2-step protocol at 72 °C (see "Materials and Methods" section). After the purification and restriction digestion of both PCR product and vector pET51b (+) the ligation was promoted overnight at 16 °C. This ligation mixture was used to transform TOP10 *E. coli* competent cells as described in the "Materials and Methods" section. The extracted plasmid pET51b-NStrep-*brtB* was then used to transform chemically competent *E. coli* Rosetta(DE3) and *E. coli* BL21(DE3)pLysS. Obtained colonies were verified using PCR (figures 33c and d).

Table 1. Primers used for the amplification of NStrep-*brtB*. Underlined are the complementary sequences of the restriction enzymes used.

Primer name	Primer sequence (5' – 3')	T_m (°C)	% GC
BrtB-NStrep-SalI-F	AATGTCGACATGGCCAATCCCTTCG	59.3	52
BrtB-NStrep-SacI-R	ATATGAGCTCCTAGTAGCCGTACCCG	61.1	54

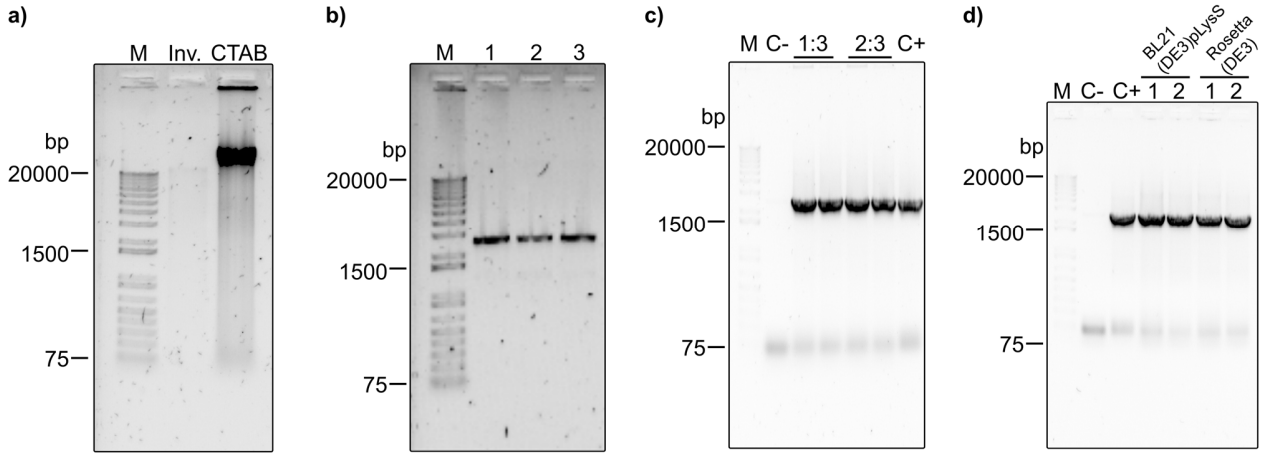


Figure 33. Results from1% agarose gel electrophoresis of cloning procedures of NStrep-*brtB*. a) Genomic DNA of *S. salina* LEGE 06099 was extracted using the Invitrogen Kit (Inv.) and with the cetyl trimethylammonium bromide (CTAB) method. b) PCR products (1-3) of *brtB* amplification after band excision and purification. c) Colony PCR results of TOP 10 *E. coli* colonies transformed with 1:3 and 2:3 ligation mixture. d) Colony PCR results of *E. coli* BL21(DE3)pLysS and Rosetta(DE3) colonies transformed with purified pET51-NStrep-*brtB*. M: GeneRuler 1 kb plus DNA ladder (Invitrogen); C-: PCR mixture without gDNA (only primers); C+: PCR mixture with gDNA from *S. salina* LEGE06099.

10. Expression and purification of NStrep-BrtB

Since we have followed the steps employed by Schultz *et al.* [90] regarding the use of Strep-Tag© technology we decided to replicate their procedures regarding protein expression. Both CylK and BrtB are predicted to be calcium dependent proteins[81-82, 89-90] meaning that supplementing the growing *E. coli* Rosetta(DE3) with $CaCl_2$ while adding isopropyl β-D-1-thiogalactopyranoside (IPTG) could promote the protein's stability. However, after inducing both *E. coli* Rosetta(DE3) and BL21(DE3)pLysS, containing recombinant NStrep-*brtB*, and harvesting the cultures after 4 and 16 hours, we did not obtain overexpression of our protein (91.4 kDa) verified in a SDS-PAGE gel (figure 34).

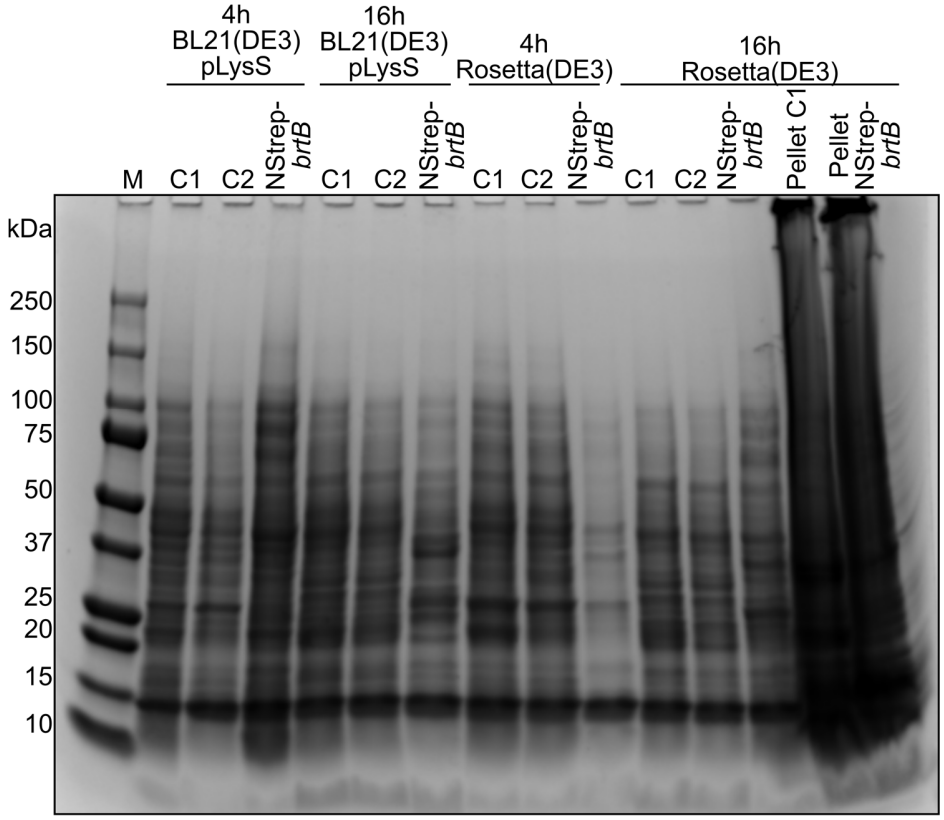


Figure 34. SDS-PAGE results of NStrep-BrtB expression after 4h and 16h induction with IPTG and $CaCl_2$ of both *E. coli* BL21(DE3)pLysS and Rosetta(DE3). Target protein has 91.4 kDa. M: Precision Plus Protein Standard (Biorad)); C1: empty pET51b (+) vector; C2: no induction control (no IPTG).

At this point, we considered that the target protein was insoluble inside the induced cells and prevented the efficient overexpression and detection. Indeed, Schultz *et al.* [90] faced the same issues and published two procedures for the expression and purification of CylK-Strep, one of them requiring protein refolding. However, such a protocol demanded high speed centrifugation technology and additional vectors which were not available in our facilities. Nevertheless, they were able to obtain 0.1 mg of CylK-Strep per liter of culture of a 2 L culture of an *E. coli* strain similar to Rosetta(DE3).[90] For this reason we decided to test if we could obtain soluble NStrep-BrtB from a 2 L culture of our *E. coli* Rosetta(DE3) harboring the pET51b-NStrep-*brtB* plasmid. After performing the affinity chromatography using a pre-packed StrepTactin Superflow column (Novagen) we were able to observe the desired recombinant protein NStrep-BrtB eluting mainly in fractions 3 to 7 (figure 35a). The purified protein was desalted and concentrated (see “Materials and Methods” section) and a yield of 0.3 mg of NStrep-brtB per liter of *E. coli* Rosetta(DE3) culture was obtained (figure 35b).

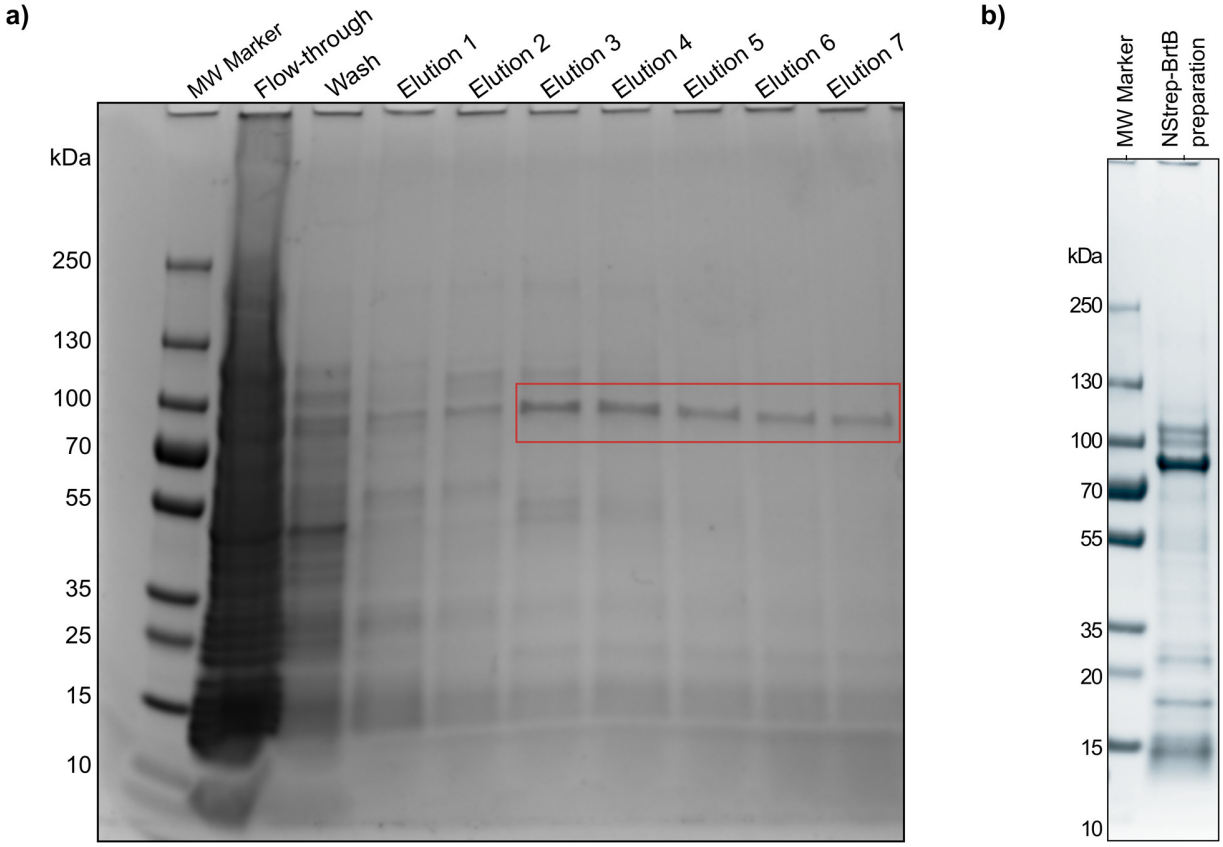


Figure 35. SDS-PAGE of the affinity chromatography performed (a) and the NStrep-BrtB after desalting and concentration (b). MW Marker: Precision Plus Protein Standard (Biorad)). Pooled fractions 3 to 7 are evidenced with a red rectangle.

11. Enzymatic assays using recombinant NStrep-BrtB

To obtain biochemical evidence supporting that BrtB was responsible for the esterification of bartolosides we resorted to *in vitro* enzymatic assays of the protein's activity. Since BrtB is predicted to be calcium dependent we performed the assays in a buffer containing Ca^{2+}, as in the expression and purification. After incubating a mixture of bartoloside A (previously isolated by Leão and co-workers[82]) with both palmitic and 6-heptynoic acids, the reaction was quenched and the collected aliquots analyzed (0, 6 and 24h) by LC-HRESIMS/MS (figures 36 and 37, annex figure 5). The results (figures 36 and 37) showed that BrtB was able to catalyze the reaction between bartoloside A (**1**) and 6-heptynoic acid, generating both the bartoloside A mono- and diester of 6-heptynoate (**8a**/**8b** and **3**, respectively). In line with the *in vivo* data, we were also able

to verify the esterification of bartoloside A (**1**) with palmitic acid to generate bartoloside A palmitate (**5a/5b**) as well as bartoloside A dipalmitate (**7**), though in far less abundance and only after 24h (figure 37). Analysis of the controls without NStrep-BrtB (only **1** and fatty acids) showed the decrease in bartoloside A levels without the formation of any other peaks, which can potentially be explained by the degradation of compound **1** when in the presence of an acidic environment or by solubility issues caused by the presence of the fatty acids.

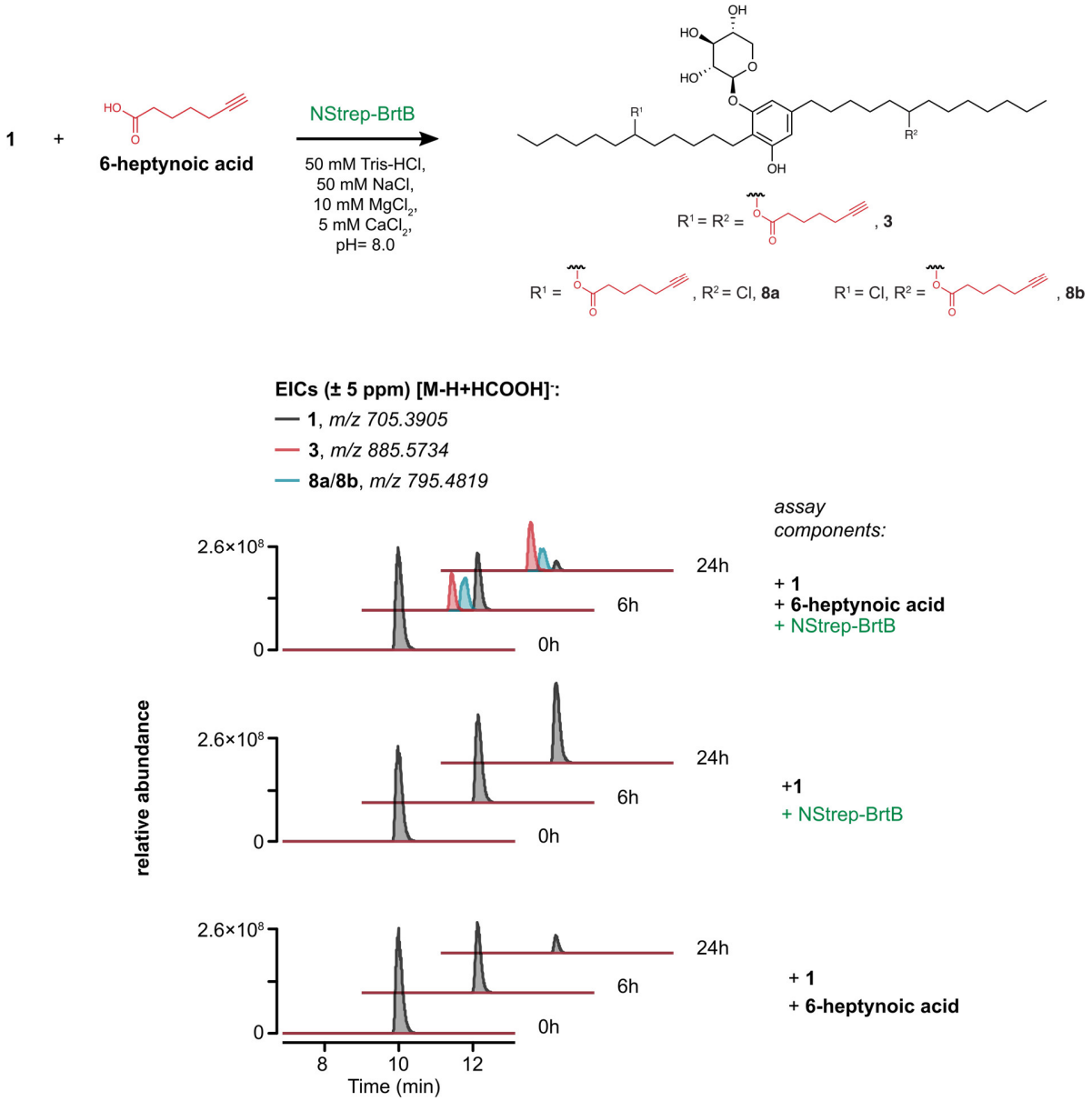


Figure 36. BrtB converts bartoloside A (**1**) and 6-heptynoic acid into bartoloside esters **3** and **8a/8b** *in vitro*. Depicted are the extracted ion chromatograms (EICs) obtained from LC-HRESIMS analysis of the assay containing NStrep-BrtB and control assays with either NStrep-BrtB or 6-heptynoic acid.

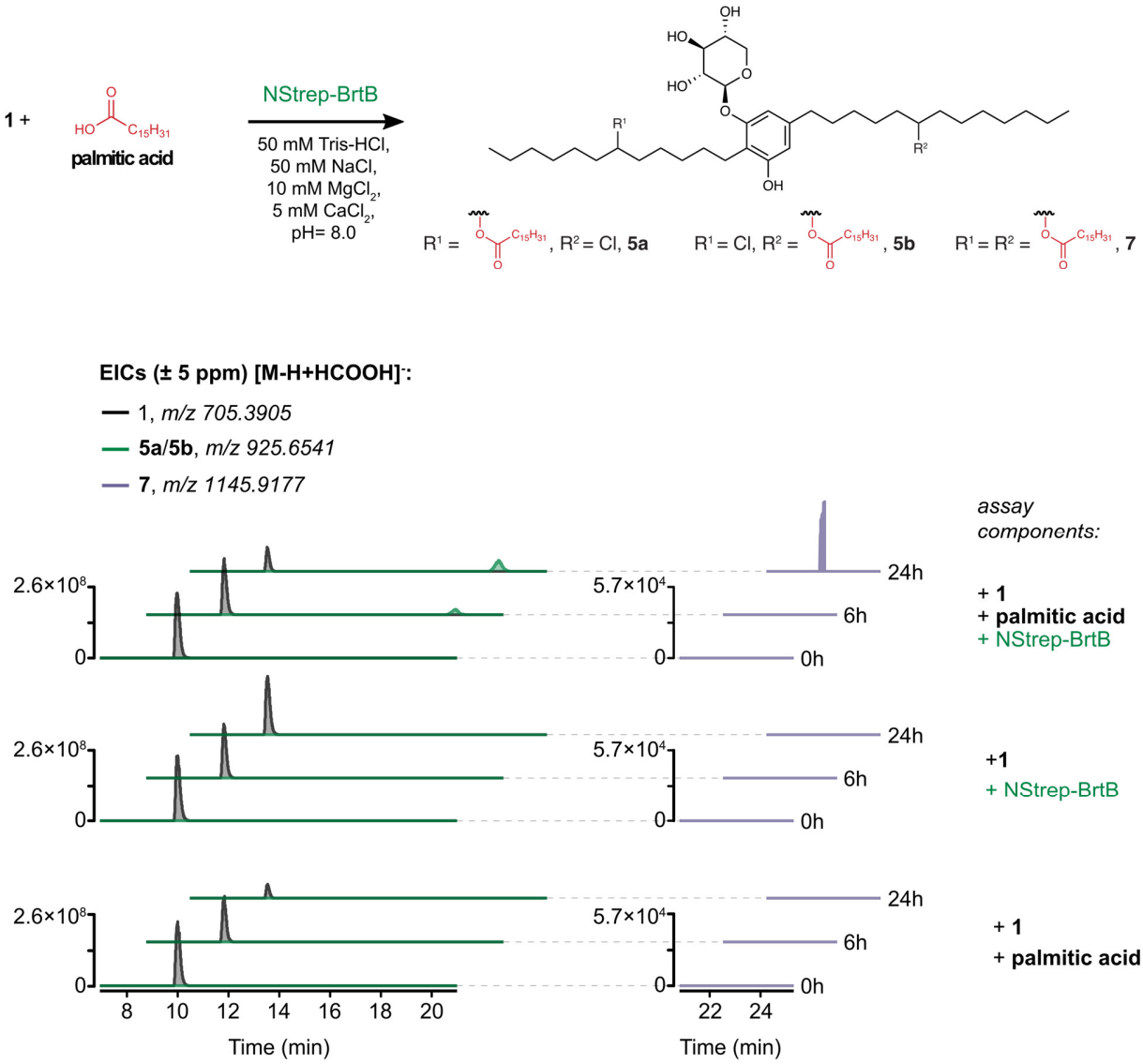


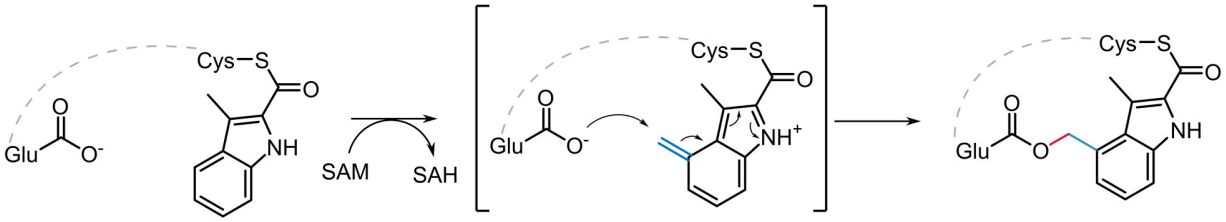
Figure 37. BrtB converts bartoloside A (**1**) and palmitic acid into bartoloside esters **5a/5b** and **7** *in vitro*. Depicted are the extracted ion chromatograms (EICs) obtained from LC-HRESIMS analysis of the assay containing NStrep-BrtB and control assays with either NStrep-BrtB or palmitic acid.

Together, our data show that BrtB catalyzes the C-O alkylation of the secondary alkyl halide moieties from bartolosides. As mentioned above, in organic synthesis, esterification processes usually rely on the nucleophilic attack of an alcohol onto a carboxylic carbon or carbonyl as typified in the Fisher-Speier esterification.[93] Nevertheless, esterification reactions that are independent of alcohols and take advantage of the nucleophilicity of the carboxylate anion have also been described.[97] Such reactivity can be found using primary alkyl halides as the electrophile but the

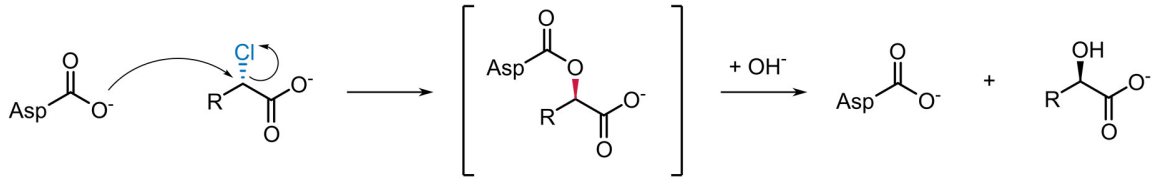
procedures usually rely on the application of expensive or even toxic intermediates, e.g. hexamethylphosphoramide (HMPA).[98]

In biological chemistry, carboxylates acting as nucleophiles can also be found but require substrate activation by adenylation.[99] Examples of such reactions can be found in the polyketide biosynthesis where alpha-keto acids and/or fatty acids are activated by adenosine triphosphate (ATP), through the formation of a bond between the carboxylate and phosphate of ATP, to generate an Acyl-AMP (adenosine monophosphate).[99-100] Nevertheless, to undergo carboxyester formation these reactions rely on the nucleophilic attack onto an activated acyl carbon (through acylation of a suitable alcohol as in the lipase and esterase activities)[101] or through direct oxidation by Bayer-Villiger monooxygenases[34]. To our knowledge, the single example in living organisms in which carboxyester formation is suggested to occur via a carboxylate nucleophile is in nosiheptide biosynthesis.[102] NosN, encoded in the *nos* gene cluster, is reported as the responsible for the catalysis of an intramolecular reaction between a reactive exocyclic electrophilic intermediate and a glutamate residue (Glu6) of the structural peptide, generating an ester linkage and causing the macrolactonization in nosiheptide (figure 38).[102] Furthermore, transient carboxyester formation has also been verified under the action of some haloalkane dehalogenases.[103] For instance, the L-2-Haloacid dehalogenase from *Pseudomonas* sp. YL is proposed to catalyze the nucleophilic attack of carboxylate group from aspartate onto the α-carbon of the L-2-haloacid.[104-105] Such ester intermediate will be later attacked by an water molecule activated by a basic amino acid from the enzyme, causing the hydrolysis of the ester intermediate and originating a hydroxylated version of the haloacid (figure 38).[104-105] Even though these enzymes catalyze esterifications using a carboxylate nucleophile the reactivity and substrates behind their mechanisms are considerably different when compared to BrtB action (figure 38). With reference to nosiheptide biosynthesis, the initial substrate consists on a methylene moiety in the side-ring system of 3,4-dimethylindolic acid which has a clearly different reactivity when compared to a secondary alkyl halide as in the bartolosides. On the other hand, the L-2-haloacid dehalogenase also uses a secondary alkyl moiety, though there is carboxylic carbon withdrawing electron density from the halogenated α-carbon, increasing its electrophilicity and causing this position to be more prone to the nucleophilic attack by the carboxylate than in the halogenated position of bartolosides. As such, BrtB is a novel enzyme responsible for an unprecedented O-alkylation of secondary alkyl halide moieties.

NosN radical SAM:



L-2-Haloacid dehalogenase:



BrtB (O-alkylation):

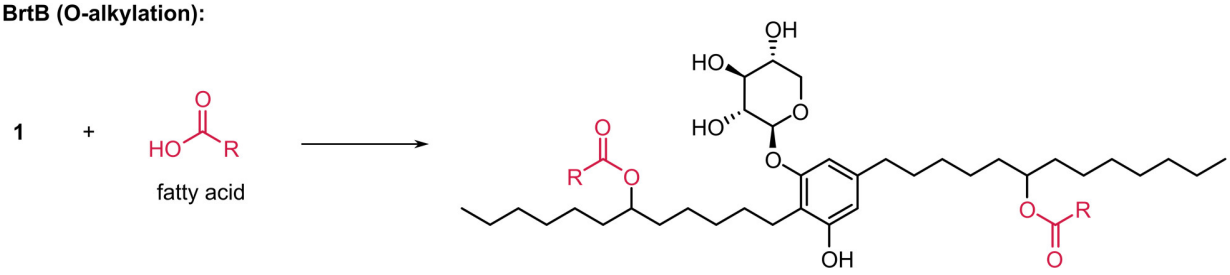


Figure 38. Selected O-alkylating enzymes. A exomethylene intermediate is generated by NosN, a SAM depedent enzyme, which will then drive an intramolecular ester bond formation. Dehalogenation of alkyl halides catalyzed by L-2-haloacid dehalogenase involves the formation of an ester species, which results from the nucleophilic attack of a side chain carboxylate onto the chlorinated position of the haloacid. BrtB catalyzes a C-O alkylation between non-activated fatty acids and bartolosides.

12. Phylogenetic Analysis of BrtB homologs

After establishing that BrtB is the enzyme responsible for the novel biocatalytic O-alkylation reaction behind the formation of bartoloside esters, we decided to look more closely to biosynthetic gene clusters containing BrtB homologs. To accomplish this, we performed a phylogenetic analysis and verified the presence of two major, well-supported clades of BrtB/CylK homologs within the dataset. One of the clades is composed by cyanobacterial homologs that are within biosynthetic gene clusters encoding for halogenases, while the other clade was formed by essentially non-cyanobacterial sequences. Interestingly, within this larger clade, BrtB homologs diverge into an independent sub-clade from CylK homologs, involved in the biosynthesis of

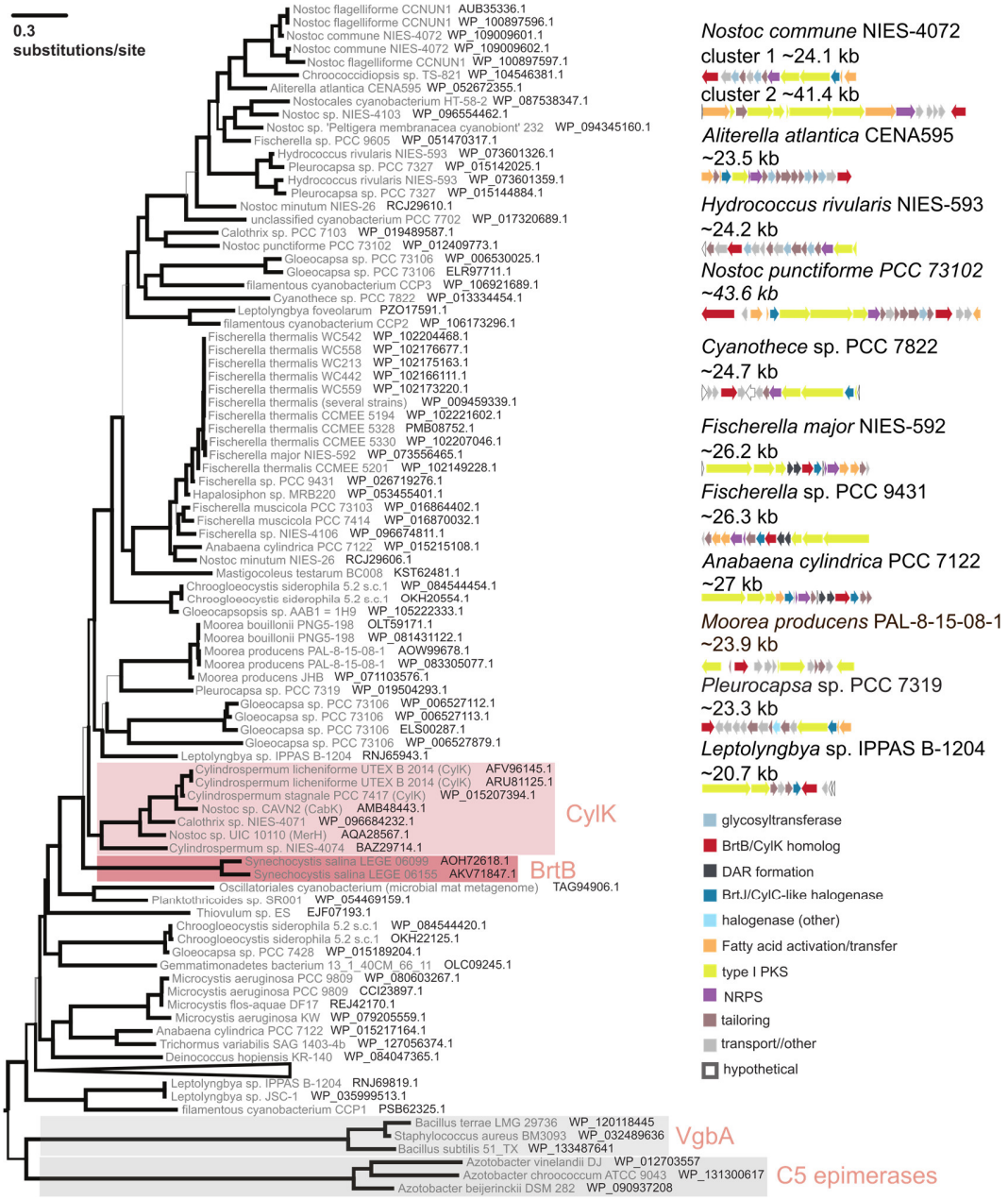


Figure 39. Phylogenetic analysis of BrtB homologs (left) and exemples of putative biosynthetic gene cluster containing BrtB/CylK homologs.

different cyclophanes. Annotation of the putative BGC within both these BrtB/CylK clades indicated that these gene clusters conspicuously encode for type I polyketide synthases, fatty acid activating enzymes, non-ribosomal peptide synthetases and varied tailoring functionalities. For instance, *Anabaena* cylindrical PCC 7122 contains a putative BGC interestingly similar to *brt*, with the presence of a BrtJ (halogenase) homolog, possibly encoding for bartoloside-like secondary

metabolites. Knowing this, our phylogenetic analysis along with previous reported data[89] strongly suggest that BGCs encoding CylK/BrtB-like enzymes from cyanobacteria represent prospects for the discovery of novel enzymatic reactivity and natural product chemistry.

III. Final Remarks

Taking into consideration that the biosynthesis of bartolosides require the recruitment of fatty acid derivatives from primary metabolism, we anticipated that by supplying terminal alkyne fatty acids to *S. salina* LEGE 06099 (a major bartoloside producer) cultures, we would obtain click chemistry-accessible versions of the bartolosides. However, the attempt to incorporate terminal alkyne moieties into previously known bartolosides failed and instead led to the discovery and later isolation of two compounds (**3** and **4**), formed upon supplementation, that were elucidated and confirmed to be bartoloside esters of the supplemented fatty acid 6-heptynoate. We eventually detected putative natural occurring bartoloside A palmitate, suggesting that the esterification in bartolosides could happen without supplementation of exogenous fatty acids. After successfully isolating and structurally elucidating bartoloside A palmitate (**5a**) we concluded that the previously reported bartolosides were not, in all likelihood, the end-product of the *brt* gene cluster. These findings drove us to perform *in vitro* experiments to determine the function of the only enzyme in the bartolosides gene cluster without ascribed function. Overall, we show that BrtB is a novel enzyme responsible for the catalysis of free (non-activated) fatty acids with the chlorinated positions found in the bartolosides, generating a significant diversity of branched bartolosides. Such reactivity has only been transiently observed in intermediaries formed by haloalkane dehalogenases action. BrtB is the second enzyme of its family to be characterized, along with the Friedel-Crafts alkylating enzyme CylK.[89] Phylogenetic analysis of CylK/BrtB-like enzymes revealed a series of putative cyanobacterial biosynthetic gene clusters that paved the way to the discovery of novel biocatalytic reactivity and natural product architecture. Furthermore, the relatively high abundance of these bartolosides indicates that there must exist an important biological role for such secondary metabolites. Bearing in mind the conspicuous depletion of bartolosides in the presence of exogenously supplied free fatty acids, we speculate that the formation of bartoloside esters, by BrtB, might represent a fatty acid storage or scavenging mechanism.

IV. Materials and Methods

1. Cyanobacterial culture conditions

The cyanobacterium *Synechocystis salina* LEGE 06099 was obtained from the LEGEcc.[106] Cultures were grown in Z8 medium supplemented with 25 g L^{-1} sea salts (Tropic Marin), at 25 °C, under a 14:10 h light/dark cycle and constant aeration. For the feeding studies with 5-hexynoic and 6-heptynoic acids, small-scale cultures were inoculated to a final OD_{750nm} of 0.04~0.1 while large scale cultures (20 L) were inoculated using 1.5 L of stationary phase cultures. Feeding experiments with other substrates used a 3:50 inoculum of stationary phase cultures.

2. Feeding experiments

Supplementation with 5-hexynoic and 6-heptynoic acids:

Feeding experiments of *S. salina* LEGE 06099 cultures (in Z8 medium, see above) were performed through the supplementation of 5-hexynoic acid or 6-heptynoic acid to a final concentration of 50 mg L^{-1} from a 1000× concentrated solution of each acid in DMSO. The experiments were carried out in 100 mL cultures in Erlenmeyer flasks and cells were harvested after 30 days by centrifugation at 4500 ×*g* for 10 min at 4 °C, rinsed with deionized H_2O and centrifuged again. The resulting biomass pellets were lyophilized prior organic extraction (see section “Organic extraction”). Crude extracts were analyzed by LC-HRESIMS and LC-HRESIMS/MS.

Supplementation with other substrates:

A similar procedure was used for supplementation with 7-bromoheptanoic, trans-3-hexenedioic, hexylphosphoric, caprylic, butyric, lauric and palmitic acids as well as hexanamide and hexylalcohol but a final concentration of each substrate of 0.4 mM was used. Cells were harvested after three days of exposure and organic extraction (see section “Organic extraction”) was performed without prior lyophilization. The resulting crude extracts were analyzed by LC-HRESIMS.

3. Large-scale *S. salina* LEGE 06099 cultivation

In order to obtain sufficient biomass to isolate sufficient amounts of bartoloside esters formed upon supplementation for structure elucidation, a 20 L culture of *S. salina* LEGE 06099 was prepared in Z8 medium supplemented with 6-heptynoic acid (50 mg L^{-1}), as described above. After a 45-day growth period, the biomass was harvested by centrifugation (4500 ×*g*, 10 min, 4 °C), washed with deionized water and centrifuged again. The corresponding pellet was freeze-dried and stored at -20 °C until further use.

For the isolation of bartoloside esters formed without any supplementation (instead, 0.5% DMSO was added to the culture), a 20 L culture of *S. salina* LEGE 06099 was grown in Z8 medium during 30 days and the biomass was collected as previously described.

4. Organic extraction

For the small scale cultures (100 mL) the organic extraction was carried through repeated percolation using a mixture of CH_2Cl_2/MeOH (2:1, v/v) at room temperature while for the large scale cultures (20 L) the extraction was carried both at room temperature and at 40 °C. Solvent volume depended on the sample size. The solvent was removed *in vacuo* and the extract stored at – 20 ºC until further use.

5. LC-HRESIMS analysis

For LC-HRESIMS analyses, separation was performed in a Luna C18 column (100 × 3 mm, 3 µm, 100 Å, Phenomenex). Mixtures of MeOH/H_2O 1:1 (v/v) with 0.1% formic acid (eluent A) and isopropanol (IPA) with 0.1% formic acid (eluent B) were used as mobile phase, with a flow rate of 0.4 mL min^{-1}. Depending on the sample to be analyzed, different elution programs and injection volumes/concentrations were used, as detailed below.

<u>LC-HRESIMS of samples derived from feeding with 5-hexynoic and 6-heptynoic acids:</u>

Crude extracts and flash chromatography fractions obtained from the initial feeding experiments were separated (10 µL of a 1 mg mL^{-1} solution were injected) using a gradient from 9:1 to 3:7 eluent A/eluent B in 10 min and held for 7 min before returning to the initial conditions.

LC-HRESIMS of samples derived from feeding with other substrates:

Crude extracts obtained from feeding experiments with 7-bromoheptanoic, trans-3-hexenedioic, hexylphosphoric, caprylic, butyric, lauric and palmitic acids, hexanamide and hexylalcohol were separated (10 µL of a 1 mg mL^{-1} solution were injected) using a gradient from 9:1 to 3:7 eluent A/eluent B in 10 min and held for 12 min before returning to the initial conditions.

LC-HRESIMS of samples derived from 20 L culture of *S. salina* LEGE 06099 without feeding:

Crude extracts of biomass from cultures that were not supplemented with fatty acids were analyzed by LC-HRESIMS using a gradient from at 9:1 to 1:4 eluent A/eluent B in 3 min and held for 28 min before returning to initial conditions. Fractions obtained from semipreparative HPLC during purification of natural bartoloside esters were analyzed (10 µL at 1 mg mL^{-1} per injection) using a similar program, but with an isocratic step of only 24 min. Semipreparative HPLC subfractions that were analyzed by LC- HRESIMS/MS were separated (10 µL at 0.5 mg mL^{-1} injected) using a gradient from 9:1 to 7:13 eluent A/eluent B in 5 min, increasing to 3:7 eluent A/eluent B over 15 min and holding the isocratic for additional 5 min before returning to initial conditions.

LC-HRESIMS of enzymatic reaction samples:

Analysis of enzymatic reaction samples was carried out by injecting 10 µL of supernatant from the methanol/acetonitrile-quenched reaction mixture (see Enzymatic assays). Separation involved an isocratic step of 9:1 eluent A/eluent B over 2 min, followed by a linear gradient to 7:13 eluent A/eluent B over 3 min and held for 10 min before returning to the initial conditions.

6. HRESIMS/MS analysis

MS/MS parameters for the LC-HRESIMS/MS analysis of crude extracts, HPLC fractions and enzymatic assays were: resolution of 35000, with a 1 m/z isolation window, a loop count of 3, AGC target of 5×10^4 and collision energy of 35 eV.

HRESIMS/MS analysis of purified compounds **3** and **4** was performed by direct injection (0.1 mg mL^{-1} solutions) into the spectrometer, using a resolution of 35000, a 1 *m/z* isolation window, a loop count of 3 and an AGC target of 5×10^4. Stepped collision energies of 35, 40 and 45 eV were applied. To obtain structural information regarding the dialkylresorcinol moiety, isolation of

insource-formed species were selected for ddMS^2 events. For purified compounds **3** and **4**, the in-source collision induced dissociation (CID) energy was set to 90 and 65 eV, respectively (to isolate species corresponding to loss of xylosyl, one C_7H_9O and $C_7H_9O_2$ for **3** as well as loss of xylosyl and C_7H_9O for **4**), resolution of 35000, with 1 *m/z* isolation window, loop count of 1 and AGC target of 5×10^4. CID energy was set to 55 eV. For purified compound **5a** the in-source collision energy was set to 90 eV (to isolate a species corresponding to the loss of xylosyl and HCl).

7. NMR, spectrometry and spectroscopy analysis

One- and two-dimensional NMR spectra were obtained on a Bruker Avance III, 400 MHz, in the Materials Center of the University of Porto (CEMUP). Chemical shifts are reported in parts per million (ppm) and using the residual solvent resonances as reference for ^{1}H ($CDCl_3$, H= 7.26 ppm) and ^{13}C ($CDCl_3$, C= 77.16 ppm). Optical rotation was measured in a P-2000 polarimeter (JAS.CO). Infrared spectra were acquired in a Nicolet iS5 FTIR spectrometer (ThermoScientific). UV spectra were obtained from a UV-1600PC spectrometer (VWR).

8. Fractionation and isolation of branched bartolosides

Isolation of bartoloside esters **3** and **4**:

The crude extract (754.6 mg) obtained from a 20 L culture of *S. salina* LEGE 06099 in Z8 medium supplemented with 6-heptynoic acid (details in “Feeding Experiments” section) was fractionated by normal phase flash chromatography (Si gel 60, 0.040-0.063 mm, Macherey Nagel) using a gradient of increasing polarity gradient from hexane to EtOAc to MeOH, yielding eight fractions (1-8). Fraction 5 (102.2 mg), eluting with a 2:3 mixture of EtOAc/hexane was further separated by RP-semipreparative HPLC with a ACE C18 column (100 Å pore size, 250 × 10 mm, 5 µm) (ACE). Mobile phase A was ddH2O and mobile phase B was MeCN (aq). The LC method used a constant flow rate of 3 mL min^{-1} with an isocratic step 8% A, 92% B over 25 min followed by a gradient to 100% B over 2 min and held for 15 minutes before returning to the initial conditions. This procedure afforded 10 subfractions. Fractions 5.6 (compound **3**, 48.0 mg, t_R = 13.5 -15.0 min) and 5.10 (compound **4**, 4.3 mg, t_R = 24.0 – 25.0 min) were spectroscopically pure (>95%, ^{1}H NMR), showing signals characteristic of bartolosides.[81-82]

Isolation of bartoloside ester **5a**:

The resulting organic extract (574.3 mg) from 20 L culture of *S. salina* LEGE 06099 was fractionated by normal phase flash chromatography (Si gel 60, 0.040-0.063 mm, Macherey-Nagel). An increasing polarity gradient from hexane to EtOAc to MeOH was used for elution. The collected fractions were pooled according to their TLC profiles, affording 23 sub-fractions (1-23). Sub-fraction 15, eluting with 2:3 and 1:1 EtOAc/hexane, contained *m/z* 879.648 (**5a**/**5b**, $[M-H]^-$), as verified by LCHRESIMS analysis, and was further fractionated by RP-semipreparative HPLC (ACE C18, 250 × 10 mm, 5 µm, 100 Å, ACE), using 98% MeOH (aq) isocratically over 65 minutes and with a flow rate of 3 mL min^{-1}. ^1H NMR analysis showed that a minor bartoloside co-eluted (likely compound **5b**), and thus further separation was necessary. Fraction 15.7 (t_R= 32.1 – 34.5 min) was reinjected into the RP-semipreparative system using the above-mentioned column and further separated with 98% MeCN (aq) to yield metabolite **5a** (3.1 mg, t_R ~122 min). ^1H NMR analysis indicated that the compound was pure (>95%) and showed characteristic resonances for bartolosides.[81-82]

9. Cloning of *brtB*

Genomic DNA from S. salina LEGE 06099 was extracted using the PureLink™ Genomic DNA MiniKit (Invitrogen) and CTAB method but only the second was used for the PCR amplification of brtB using primers BrtB-NStrep-SalI-F and BrtB-NStrep-SacI-R, which contained restriction sites for SalI and SacI, respectively. The reaction was prepared in a volume of 20 µL, containing 1× Phusion HF Buffer, 400 µM dNTPs, 0.5 µM of each primer, 0.4U Phusion polymerase (ThermoScientific) and 2 µL of template gDNA. The PCR cycling program consisted in a two-step protocol with extension at 72 °C for 1.5 min. PCR products were gel purified (Illustra GFX PCR DNA and Gel Band Purification kit, GE Healthcare), restriction digested using SalI and SacI (Thermo Scientific) for 1.5 h at 37 °C and gel purified again. Plasmid purifications were performed using NZYMiniprep (NZYTech). Digests were ligated into linearized StrepTag-containing pET-51b (+), using T4 DNA Ligase at 16 °C overnight.

Competent *E. coli* TOP10 cells were transformed with a 5 µL ligation sample by incubating the mixture on ice for 30 min, heat shocking at 42 °C for 45 s and recovering on ice for 2 min. Then, 250 µL of warm S.O.C. medium was added, the mixture was incubated at 37 °C for 1 h and 100

µL of cells was plated on LB agar supplemented with 100 µg mL^{-1} ampicillin. The plate was incubated overnight at 37 °C. Individual colonies were inoculated into 5 mL of LB supplemented with 100 µg mL^{-1} ampicillin and grown overnight at 37 °C with 175 rpm shaking. The purified plasmids were sequenced to verify construct identity and amplification fidelity. pET-51b-NStrep-brtB was transformed into chemically competent *E. coli* BL21(DE3)pLysS and Rosetta (DE3) by incubating the plasmid and cells on ice for 30 min and 5 min, respectively, applying a heat-shock at 42 °C for 45 s and recovering for 2 min on ice. Then, 80 µL of warm S.O.C. medium was added, the mixture was incubated at 37 °C for 1 h and 70 µL of cells was plated on LB agar supplemented with 100 µg mL^{-1} ampicillin and 34 µg mL^{-1} chloramphenicol. Individual colonies were inoculated into 5 mL of LB supplemented with 100 µg mL^{-1} ampicillin and 34 µg mL^{-1} chloramphenicol and grown overnight at 37 °C with 175 rpm shaking. The resulting transformants were frozen in LB/glycerol 1:1 (v/v) stocks and stored at -80 °C.

10. Small-scale NStrep-BrtB expression tests

Glycerol stocks of *E. coli* BL21(DE3)pLysS and Rosetta (DE3) harboring the pET-51b-NStrep-*brtB* plasmid were used to prepare 50 mL cultures in LB medium supplemented with 100 µg mL^{-1} ampicillin and 34 µg mL^{-1} chloramphenicol. Cultures were incubated at 37 °C with 190 rpm and transferred to 15 °C after reaching OD_{600nm}=0.4. Induction was carried at OD_{600nm}=0.6-0.7 with 0.5 mM IPTG and supplemented with 5 mM $CaCl_2$.

After 4h and 16h of induction 20 mL of each culture were centrifuged (3000 ×*g*, 10 min, 4 °C (Gyrozen 2236R, Rotor GRF-L-85-6), the pellets washed with Phosphate-buffered saline (PBS), centrifuged again and stored at -20 ºC. Pellets were resuspended in 500 µL of lysis buffer (50 mM Tris-HCl, 200 mM NaCl, 5% glycerol, 1 mM DTT and 20 mM $CaCl_2$, pH 8.0) supplemented with 1× Halt Protease Inhibitor Cocktail, lysozyme was added to a final concentration of 300 µg/mL The mixtures were incubated at 4 °C for 4h and 10 mM MgCl2 was added to each tube and incubated again for additional 30 min. the samples were centrifuged (17000 ×*g*, 30 minutes, 4 °C). Collected samples analyzed by SDS-PAGE (4-20% Mini-PROTEAN TGX precast gel, BIO-RAD). The gels were washed with ddH_2O for 15 min, stained with Coomassie Blue G250 BioSafe (BioRad) for 1h and washed again.

11. Expression and purification of NStrep-BrtB in *E. coli* Rosetta(DE3)

A glycerol stock of *E. coli* Rosetta (DE3) harboring the pET-51b-NStrep-*brtB* plasmid was used to prepare a 50 mL starter culture in LB medium supplemented with 100 μg mL^{-1} ampicillin and 34 μg mL^{-1} chloramphenicol. The starter culture was incubated overnight at 37 °C with 175 rpm shaking. The overnight culture was used to inoculate (1:100 dilution) 2 L of LB medium supplemented with 100 μg mL^{-1} ampicillin and 34 μg mL^{-1} chloramphenicol. The culture was incubated at 37 °C with 190 rpm shaking. At OD_{600nm}=0.4-0.5 the culture was transferred to 15 °C and orbital shaking decreased to 175 rpm. Once the culture reached OD_{600nm}=0.8-0.9, protein expression was induced with 500 μM IPTG and the culture supplemented with 5 mM $CaCl_2$. After four additional hours of incubation at 15 °C with 175 rpm shaking, the cells were harvested by centrifugation (3000 ×g, 10 min, 4 °C), washed with PBS and re-centrifuged. The pellet was flash frozen in liquid nitrogen and stored at -80 °C until it was resuspended in lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 20 mM $CaCl_2$, 10 mM $MgCl_2$, 1 mM EDTA, pH 8.0) to which Pierce Protease Inhibitor Mini Tablets EDTA-free were added according to the manufacturer's instructions (Thermo Scientific). The cells were lysed in a cell disruptor (CF Range, Constant Sytems Ltd) by a single cycle at 20000 psi and the lysate was clarified by centrifugation (19500 ×g, 30 min, 4 °C). A pre-packed 1 mL StrepTactin Superflow column (Novagen) was equilibrated with 2 mL of lysis buffer at 4 °C for 45 min and the clarified lysate was transferred to the column, discarding the flow-through. The column was washed with 5 mL lysis buffer and the protein was eluted in multiple 500 μL fractions using elution buffer (100 mM Tris-HCl, 500 mM NaCl, 20 mM $CaCl_2$, 10 mM $MgCl_2$, 1 mM EDTA, 2.5 mM desthiobiotin, pH 8.0). The collected fractions were analyzed by SDS-PAGE (4-20% Mini-PROTEAN TGX precast gel, BIO-RAD). Fractions containing the desired protein (~91.4 kDa) were pooled and desalted using PD-10 desalting columns (GE Healthcare). The columns were equilibrated with storage buffer (50 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$ and 10% glycerol, pH 8.0) and the protein sample was loaded into the columns (up to 2.5 mL per column) and eluted with 3.5 mL storage buffer. The sample was concentrated to 34.6 μM using Pierce™ Protein Concentrator tubes, 10K MWCO (Thermo Scientific) through two consecutive 15 min centrifugations (8000 ×g, 4 °C). The concentrated protein was frozen in liquid nitrogen as pellets and stored at –80 °C. This procedure yielded 0.3 mg of NStrep-BrtB per L of *E. coli* culture.

12. Enzymatic assays

In vitro enzymatic assays were carried out in Eppendorf tubes to a final volume of 100 µL in reaction buffer (50 mM Tris-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH = 8.0). Final concentrations of each substrate were 1 µM for NStrep-BrtB, 100 µM for bartoloside A (**1**) and 200 µM for both 6-heptynoic and palmitic acids and the reactions were incubated at 37 °C with 180 rpm shaking. Aliquots (30 µL) were collected at 0, 6 and 24h and quenched with 60 µL of a cold mixture of MeOH/MeCN (1:1, v/v), vortexed immediately and incubated in ice for 10 min. Quenched reactions were then centrifuged at 17 000 ×g, 4 °C) and analyzed by LC-HRESIMS/MS.

13. Phylogenetic analysis

Amino acid sequences of the BrtB homologs with lowest e-value (BlastP) were retrieved from the NCBI database (232 homologs retrieved). Their amino acidic sequences were aligned, together with those of a set of the distantly related virginiamycin lyases and alginate C5-epimerases and with the sequence of BrtB, using MUSCLE, from within Geneious R11 (Biomatters). The resulting alignment (with a total of 239 sequences) was trimmed to its core region and contained 1558 positions. FastTree 2.1.528 (from within Geneious R11) was used to compute an approximately-maximum-likelihood phylogenetic tree, using pseudocounts and 1000 rate categories of sites.

14. Bioactivity assays

Animal cell line assays:

All cell lines were sub-cultivated and grown in supplemented medium as recommended by the providers. All HCT116 colon carcinoma cell lines were maintained in McCoy's 5A modified medium while HT-29 and hCMEC/D3 cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM). All media were also supplemented with 10% fetal bovine serum (Biochrom), 1% of penicillin/streptomycin (Biochrom), and 0.1% of Amphotericin B (GE Healthcare), and all cells lines incubated at 37 °C in 5% CO_2.

For the bioactivity assays cells were seeded at 3.3×10^4 cells mL^{-1} in 96 well-plates, exposed to compounds **3**, **4** and **5a** after 24h and kept in the incubator for 24 and 48 h. All compounds were prepared in ×100 concentrated stock solutions dissolved in dimethylsulphoxide (DMSO). A volume

of 1 µL of each stock solution was added to each well to a final concentration of 0.1% DMSO; negative control wells received solvent only; positive control used a solution of staurosporine dissolved in DMSO and was added to a final concentration of 1 µM. After the exposure, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) (MTT) was added at a final concentration of 0.05 mg ml^{-1} per well for 3 h. The formation of formazan crystals was visually evaluated by microscopy, dissolved in DMSO and absorbance measured at 550 nm on a Synergy HT Microplate Reader (Biotek).

Antimicrobial susceptibility assay:

Solutions of 0.5 mg mL^{-1} of compounds **3**, **4** and **5a** prepared in DMSO were tested against Gram-positive bacterial strains (*Staphylococcus aureus* ATCC 29213), Gram-negative bacterial strains (*Escherichia coli* ATCC 25922 and *Salmonella typhimurium* ATCC 25241) and against a yeast strain (*Candida albicans* ATCC 10231). The bacteria were grown in Mueller-Hinton agar (MH – BioKar diagnostics) from stock cultures and incubated at 37 °C. The yeast was grown in Sabouraud Dextrose Agar (BioKar diagnostics). For the antibacterial screening, a method of disc diffusion was carried out. Bacterial pure colonies were picked from overnight cultures in MH and suspended in LB liquid medium, the turbidity adjusted to OD_{600nm}= 0.090-0.110 (0.5 McFarland standard) and the MH plates seeded with the resulting inoculum. Blank discs (6 mm in diameter, Oxoid) were placed in the inoculated plates and impregnated with 15 µL of a 0.5 mg mL^{-1} solution of each compound. The plates were left for 30 min at room temperature and then incubated overnight at 37 °C. After 24h, the plates were checked for inhibition halos, indicative of antimicrobial activity. The diameter of inhibition halos was recorded. A disc with 15 µL DMSO was used as negative control.

V. References

1. Fleming, A., On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. 1929. *B World Health Organ* **2001,** *79* (8), 780-90.
2. Pye, C. R.; Bertin, M. J.; Lokey, R. S.; Gerwick, W. H.; Linington, R. G., Retrospective analysis of natural products provides insights for future discovery trends. *Proc Natl Acad Sci U S A* **2017,** *114* (22), 5601-5606.
3. Vaishnav, P.; Demain, A. L., Unexpected applications of secondary metabolites. *Biotechnol Adv* **2011,** *29* (2), 223-9.
4. Craney, A.; Ahmed, S.; Nodwell, J., Towards a new science of secondary metabolism. *J Antibiot (Tokyo)* **2013,** *66* (7), 387-400.
5. Cragg, G. M.; Newman, D. J., Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta* **2013,** *1830* (6), 3670-3695.
6. Sertuerner, Ueber das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandtheile des Opiums. *Annalen der Physik* **1817,** *55* (1), 56-89.
7. Buss, A. D.; Cox, B.; Waigh, R. D., Natural Products as Leads for New Pharmaceuticals. *Burger's Medicinal Chemistry and Drug Discovery* **2003**, 847-900.
8. Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. *Metabolites* **2012,** *2* (2), 303-36.
9. Chassagne, F.; Cabanac, G.; Hubert, G.; David, B.; Marti, G., The landscape of natural product diversity and their pharmacological relevance from a focus on the Dictionary of Natural Products®. *Phytochem Rev* **2019**.
10. Bérdy, J., Bioactive Microbial Metabolites. *J Antibiot (Tokyo)* **2005,** *58*, 1.
11. Patridge, E.; Gareiss, P.; Kinch, M. S.; Hoyer, D., An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov Today* **2016,** *21* (2), 204-7.
12. Baltz, R., Antimicrobials from actinomycetes: Back to the future. *Microbe* **2007,** *2*, 125-131.
13. Demain, A. L., Importance of microbial natural products and the need to revitalize their discovery. *J Ind Microbiol Biot* **2014,** *41* (2), 185-201.
14. Hinshaw, H. C.; Pyle, M. M.; Feldman, W. H., Streptomycin in tuberculosis. *Am J Med* **1947,** *2* (5), 429-435.
15. Schatz, A.; Bugle, E.; Waksman, S. A., Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.∗†. *P Soc Exp Biol Med* **1944,** *55* (1), 66-69.
16. Yan, Y.; Liu, Q.; Jacobsen, S. E.; Tang, Y., The impact and prospect of natural product discovery in agriculture: New technologies to explore the diversity of secondary metabolites in plants and microorganisms for applications in agriculture. *EMBO Rep* **2018,** *19* (11).
17. Demain, A. L.; Sanchez, S., Microbial drug discovery: 80 years of progress. *J Antibiot (Tokyo)* **2009,** *62* (1), 5-16.
18. Cantrell, C. L.; Dayan, F. E.; Duke, S. O., Natural Products As Sources for New Pesticides. *J Nat Prod* **2012,** *75* (6), 1231-1242.
19. Singh, R.; Kumar, M.; Mittal, A.; Mehta, P. K., Microbial metabolites in nutrition, healthcare and agriculture. *3 Biotech* **2017,** *7* (1), 15.
20. Burg, R. W.; Miller, B. M.; Baker, E. E.; Birnbaum, J.; Currie, S. A.; Hartman, R.; Kong, Y. L.; Monaghan, R. L.; Olson, G.; Putter, I.; Tunac, J. B.; Wallick, H.; Stapley, E. O.; Oiwa, R.; Omura, S., Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. *Antimicrob Agents Chemother* **1979,** *15* (3), 361-367.

21. Wolstenholme, A. J.; Rogers, A. T., Glutamate-gated chloride channels and the mode of action of the avermectin/milbemycin anthelmintics. *Parasitology* **2005,** *131* (S1), S85-S95.
22. Sen, T.; Barrow, C. J.; Deshmukh, S. K., Microbial Pigments in the Food Industry-Challenges and the Way Forward. *Front Nutr* **2019,** *6*, 7-7.
23. Stevenson, C. S.; Capper, E. A.; Roshak, A. K.; Marquez, B.; Grace, K.; Gerwick, W. H.; Jacobs, R. S.; Marshall, L. A., Scytonemin--a marine natural product inhibitor of kinases key in hyperproliferative inflammatory diseases. *Inflamm Res* **2002,** *51* (2), 112-4.
24. Guerin, M.; Huntley, M. E.; Olaizola, M., *Haematococcus* astaxanthin: applications for human health and nutrition. *Trends Biotechnol* **2003,** *21* (5), 210-6.
25. Khan, R. A., Natural products chemistry: The emerging trends and prospective goals. *Saudi Pharm J* **2018,** *26* (5), 739-753.
26. Truppo, M. D., Biocatalysis in the Pharmaceutical Industry: The Need for Speed. *ACS Medl Chem Lett* **2017,** *8* (5), 476-480.
27. Turner, N. J.; Truppo, M. D., Biocatalysis enters a new era. *Curr Opin Chem Biol* **2013,** *17* (2), 212-4.
28. Dalhoff, C.; Hüben, M.; Lenz, T.; Poot, P.; Nordhoff, E.; Köster, H.; Weinhold, E., Synthesis of S-Adenosyl-L-homocysteine Capture Compounds for Selective Photoinduced Isolation of Methyltransferases. *ChemBioChem* **2010,** *11* (2), 256-265.
29. Tengg, M.; Stecher, H.; Offner, L.; Plasch, K.; Anderl, F.; Weber, H.; Schwab, H.; Gruber-Khadjawi, M., Methyltransferases: Green Catalysts for Friedel–Crafts Alkylations. *ChemCatChem* **2016,** *8* (7), 1354-1360.
30. Holtmann, D.; Fraaije, M. W.; Arends, I. W.; Opperman, D. J.; Hollmann, F., The taming of oxygen: biocatalytic oxyfunctionalisations. *Chem Commun (Camb)* **2014,** *50* (87), 13180-200.
31. Woodley, J. M., Protein engineering of enzymes for process applications. *Curr Opin Chem Biol* **2013,** *17* (2), 310-316.
32. Hilker, I.; Wohlgemuth, R.; Alphand, V.; Furstoss, R., Microbial transformations 59: first kilogram scale asymmetric microbial Baeyer-Villiger oxidation with optimized productivity using a resin-based in situ SFPR strategy. *Biotechnol Bioeng* **2005,** *92* (6), 702-710.
33. Mihovilovic, Marko D.; Müller, B.; Stanetty, P., Monooxygenase-Mediated Baeyer−Villiger Oxidations. *Eur J Org Chem* **2002,** *2002* (22), 3711-3730.
34. Leisch, H.; Morley, K.; Lau, P. C., Baeyer-Villiger monooxygenases: more than just green chemistry. *Chem Rev* **2011,** *111* (7), 4165-222.
35. Bernan, V. S.; Greenstein, M.; Maiese, W. M., Marine Microorganisms as a Source of New Natural Products. *Adv Appl Microbiol* **1997** (43), 57-90.
36. Fenical, W.; Jensen, P. R., Developing a new resource for drug discovery: marine actinomycete bacteria. *Nat Chem Biol* **2006,** *2* (12), 666-73.
37. Montaser, R.; Luesch, H., Marine natural products: a new wave of drugs? *Future Med Chem* **2011,** *3* (12), 1475-89.
38. Demay, J.; Bernard, C.; Reinhardt, A.; Marie, B., Natural Products from Cyanobacteria: Focus on Beneficial Activities. *Mar Drugs* **2019,** *17* (6).
39. Leao, P. N.; Engene, N.; Antunes, A.; Gerwick, W. H.; Vasconcelos, V., The chemical ecology of cyanobacteria. *Nat Prod Rep* **2012,** *29* (3), 372-91.
40. Kasting, J. F.; Siefert, J. L., Life and the Evolution of Earth's Atmosphere. *Science (New York)* **2002,** *296* (5570), 1066.
41. Flores, E.; Herrero, A., Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol* **2009,** *8*, 39.
42. Wase, N. V.; Wright, P. C., Systems biology of cyanobacterial secondary metabolite production and its role in drug discovery. *Expert Opin Drug Dis* **2008,** *3* (8), 903-29.
43. Marner, F.-J.; Moore, R. E.; Hirotsu, K.; Clardy, J., Majusculamides A and B, two epimeric lipodipeptides from *Lyngbya majuscula*. *J Org Chem* **1977,** *42* (17), 2815-2819.

44. Nunnery, J. K.; Mevers, E.; Gerwick, W. H., Biologically active secondary metabolites from marine cyanobacteria. *Curr Opin Biotechnol* **2010,** *21* (6), 787-93.
45. Mynderse, J. S.; Moore, R. E., Toxins from blue-green algae: structures of oscillatoxin A and three related bromine-containing toxins. *J Org Chem* **1978,** *43* (11), 2301-2303.
46. Mynderse, J. S.; Moore, R. E.; Kashiwagi, M.; Norton, T. R., Antileukemia activity in the Osillatoriaceae: isolation of Debromoaplysiatoxin from *Lyngbya*. *Science (New York)* **1977,** *196* (4289), 538-40.
47. Garcia-Pichel, F.; Sherry, N. D.; Castenholz, R. W., Evidence for an ultraviolet sunscreen role of the extracellular pigment scytonemin in the terrestrial cyanobacterium *Chlorogloeopsis* sp. *Photochem Photobiol* **1992,** *56* (1), 17-23.
48. Luesch, H.; Moore, R. E.; Paul, V. J.; Mooberry, S. L.; Corbett, T. H., Isolation of dolastatin 10 from the marine cyanobacterium *Symploca* species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1. *J Nat Prod* **2001,** *64* (7), 907-10.
49. Patel, S.; Keohan, M. L.; Saif, M. W.; Rushing, D.; Baez, L.; Feit, K.; DeJager, R.; Anderson, S., Phase II study of intravenous TZT-1027 in patients with advanced or metastatic soft-tissue sarcomas with prior exposure to anthracycline-based chemotherapy. *Cancer* **2006,** *107* (12), 2881-7.
50. Newman, D. J.; Cragg, G. M., Current Status of Marine-Derived Compounds as Warheads in Anti-Tumor Drug Candidates. *Mar Drugs* **2017,** *15* (4), 99.
51. Shah, S. A. A.; Akhter, N.; Auckloo, B. N.; Khan, I.; Lu, Y.; Wang, K.; Wu, B.; Guo, Y. W., Structural Diversity, Biological Properties and Applications of Natural Products from Cyanobacteria. A Review. *Mar Drugs* **2017,** *15* (11).
52. Pfannenstiel, B. T.; Keller, N. P., On top of biosynthetic gene clusters: How epigenetic machinery influences secondary metabolism in fungi. *Biotechnol Adv* **2019,** *37* (6), 107345.
53. Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; de Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; Cruz-Morales, P.; Duddela, S.; Düsterhus, S.; Edwards, D. J.; Fewer, D. P.; Garg, N.; Geiger, C.; Gomez-Escribano, J. P.; Greule, A.; Hadjithomas, M.; Haines, A. S.; Helfrich, E. J. N.; Hillwig, M. L.; Ishida, K.; Jones, A. C.; Jones, C. S.; Jungmann, K.; Kegler, C.; Kim, H. U.; Kötter, P.; Krug, D.; Masschelein, J.; Melnik, A. V.; Mantovani, S. M.; Monroe, E. A.; Moore, M.; Moss, N.; Nützmann, H.-W.; Pan, G.; Pati, A.; Petras, D.; Reen, F. J.; Rosconi, F.; Rui, Z.; Tian, Z.; Tobias, N. J.; Tsunematsu, Y.; Wiemann, P.; Wyckoff, E.; Yan, X.; Yim, G.; Yu, F.; Xie, Y.; Aigle, B.; Apel, A. K.; Balibar, C. J.; Balskus, E. P.; Barona-Gómez, F.; Bechthold, A.; Bode, H. B.; Borriss, R.; Brady, S. F.; Brakhage, A. A.; Caffrey, P.; Cheng, Y.-Q.; Clardy, J.; Cox, R. J.; De Mot, R.; Donadio, S.; Donia, M. S.; van der Donk, W. A.; Dorrestein, P. C.; Doyle, S.; Driessen, A. J. M.; Ehling-Schulz, M.; Entian, K.-D.; Fischbach, M. A.; Gerwick, L.; Gerwick, W. H.; Gross, H.; Gust, B.; Hertweck, C.; Höfte, M.; Jensen, S. E.; Ju, J.; Katz, L.; Kaysser, L.; Klassen, J. L.; Keller, N. P.; Kormanec, J.; Kuipers, O. P.; Kuzuyama, T.; Kyrpides, N. C.; Kwon, H.-J.; Lautru, S.; Lavigne, R.; Lee, C. Y.; Linquan, B.; Liu, X.; Liu, W.; Luzhetskyy, A.; Mahmud, T.; Mast, Y.; Méndez, C.; Metsä-Ketelä, M.; Micklefield, J.; Mitchell, D. A.; Moore, B. S.; Moreira, L. M.; Müller, R.; Neilan, B. A.; Nett, M.; Nielsen, J.; O'Gara, F.; Oikawa, H.; Osbourn, A.; Osburne, M. S.; Ostash, B.; Payne, S. M.; Pernodet, J.-L.; Petricek, M.; Piel, J.; Ploux, O.; Raaijmakers, J. M.; Salas, J. A.; Schmitt, E. K.; Scott, B.; Seipke, R. F.; Shen, B.; Sherman, D. H.; Sivonen, K.; Smanski, M. J.; Sosio, M.; Stegmann, E.; Süssmuth, R. D.; Tahlan, K.; Thomas, C. M.; Tang, Y.; Truman, A. W.; Viaud, M.; Walton, J. D.; Walsh, C. T.; Weber, T.; van Wezel, G. P.; Wilkinson, B.; Willey, J. M.; Wohlleben, W.; Wright, G. D.; Ziemert, N.; Zhang, C.; Zotchev, S. B.; Breitling, R.; Takano, E.; Glöckner, F. O., Minimum Information about a Biosynthetic Gene cluster. *Nat Chem Biol* **2015,** *11* (9), 625-631.
54. Walsh, C. T.; Fischbach, M. A., Natural products version 2.0: connecting genes to molecules. *J Am Chem Soc* **2010,** *132* (8), 2469-2493.

55. Jones, A. C.; Monroe, E. A.; Eisman, E. B.; Gerwick, L.; Sherman, D. H.; Gerwick, W. H., The unique mechanistic transformations involved in the biosynthesis of modular natural products from marine cyanobacteria. *Nat Prod Rep* **2010,** *27* (7), 1048-65.
56. Tidgewell, K.; Clark, B.; Gerwick, W., 2.06. The Natural Products Chemistry of Cyanobacteria. *Comprehensive Natural Products II* **2010,** *2*.
57. Finking, R.; Marahiel, M. A., Biosynthesis of Nonribosomal Peptides. *Annu Rev Microbiol* **2004,** *58* (1), 453-488.
58. Yang, X.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: new insights into the role of leader and core peptides during biosynthesis. *Chemistry* **2013,** *19* (24), 7662-7677.
59. Martins, J.; Leikoski, N.; Wahlsten, M.; Azevedo, J.; Antunes, J.; Jokela, J.; Sivonen, K.; Vasconcelos, V.; Fewer, D. P.; Leão, P. N., Sphaerocyclamide, a prenylated cyanobactin from the cyanobacterium *Sphaerospermopsis* sp. LEGE 00249. *Sci Rep-UK* **2018,** *8* (1), 14537.
60. Shen, B., Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* **2003,** *7* (2), 285-95.
61. Staunton, J.; Weissman, K. J., Polyketide biosynthesis: a millennium review. *Nat Prod Rep* **2001,** *18* (4), 380-416.
62. Hamley, I. W., Lipopeptides: from self-assembly to bioactivity. *Cheml Commun* **2015,** *51* (41), 8574-8583.
63. Katz, L.; Baltz, R. H., Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol* **2016,** *43* (2-3), 155-76.
64. Chaguturu, R.; Patwardhan, B., Chapter 1 - Drug Discovery Impasse: Pharmacognosy Holds the Key. In *Innovative Approaches in Drug Discovery* Eds. Academic Press: Boston, **2017**; pp 1-22.
65. Cronk, D., Chapter 8 - High-throughput screening. In *Drug Discovery and Development (Second Edition)* Eds. Churchill Livingstone: **2013**; pp 95-117.
66. Griffiths, J., A Brief History of Mass Spectrometry. *Anal Chem* **2008,** *80* (15), 5678-5683.
67. Kildgaard, S.; Subko, K.; Phillips, E.; Goidts, V.; de la Cruz, M.; Diaz, C.; Gotfredsen, C. H.; Andersen, B.; Frisvad, J. C.; Nielsen, K. F.; Larsen, T. O., A Dereplication and Bioguided Discovery Approach to Reveal New Compounds from a Marine-Derived Fungus *Stilbella fimetaria*. *Mar Drugs* **2017,** *15* (8).
68. Heather, J. M.; Chain, B., The sequence of sequencers: The history of sequencing DNA. *Genomics* **2016,** *107* (1), 1-8.
69. Niu, G., Genomics-Driven Natural Product Discovery in Actinomycetes. *Trends Biotechnol* **2018,** *36* (3), 238-241.
70. Challis, G. L., Genome Mining for Novel Natural Product Discovery. *J Med Chem* **2008,** *51* (9), 2618-2628.
71. Chen, R.; Wong, H. L.; Burns, B. P., New Approaches to Detect Biosynthetic Gene Clusters in the Environment. *Medicines (Basel)* **2019,** *6* (1), 32.
72. Rodríguez Estévez, M.; Myronovskyi, M.; Gummerlich, N.; Nadmid, S.; Luzhetskyy, A., Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain *Streptomyces albus* subsp. chlorinus NRRL B-24108. *Mar Drugs* **2018,** *16* (11), 435.
73. Micallef, M. L.; D'Agostino, P. M.; Al-Sinawi, B.; Neilan, B. A.; Moffitt, M. C., Exploring cyanobacterial genomes for natural product biosynthesis pathways. *Mar Genom* **2015,** *21*, 1-12.
74. Dittmann, E.; Gugger, M.; Sivonen, K.; Fewer, D. P., Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends Microbiol* **2015,** *23* (10), 642-652.
75. Sletten, E. M.; Bertozzi, C. R., Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew Chem Int Ed Engl* **2009,** *48* (38), 6974-6998.
76. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew Chem Int Ed Engl* **2001,** *40* (11), 2004-2021.

77. Kolb, H. C.; Sharpless, K. B., The growing impact of click chemistry on drug discovery. *Drug Discov Today* **2003,** *8* (24), 1128-37.
78. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew Chem Int Ed Engl* **2002,** *41* (14), 2596-2599.
79. Speers, A. E.; Adam, G. C.; Cravatt, B. F., Activity-Based Protein Profiling in Vivo Using a Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J Am Chem Soc* **2003,** *125* (16), 4686-4687.
80. Presolski, S. I.; Hong, V. P.; Finn, M. G., Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation. *Curr Protoc Chem Biol* **2011,** *3* (4), 153-162.
81. Afonso, T. B.; Costa, M. S.; Rezende de Castro, R.; Freitas, S.; Silva, A.; Schneider, M. P.; Martins, R.; Leao, P. N., Bartolosides E-K from a Marine Coccoid Cyanobacterium. *J Nat Prod* **2016,** *79* (10), 2504-2513.
82. Leao, P. N.; Nakamura, H.; Costa, M.; Pereira, A. R.; Martins, R.; Vasconcelos, V.; Gerwick, W. H.; Balskus, E. P., Biosynthesis-assisted structural elucidation of the bartolosides, chlorinated aromatic glycolipids from cyanobacteria. *Angew Chem Int Ed Engl* **2015,** *54* (38), 11063-7.
83. Martins, T. P.; Rouger, C.; Glasser, N. R.; Freitas, S.; de Fraissinette, N. B.; Balskus, E. P.; Tasdemir, D.; Leao, P. N., Chemistry, bioactivity and biosynthesis of cyanobacterial alkylresorcinols. *Nat Prod Rep* **2019**.
84. Moore, B. S.; Chen, J. L.; Patterson, G. M. L.; Moore, R. E.; Brinen, L. S.; Kato, Y.; Clardy, J., [7.7]Paracyclophanes from blue-green algae. *J Am Chem Soc* **1990,** *112* (10), 4061-4063.
85. Beresovsky, D.; Hadas, O.; Livne, A.; Sukenik, A.; Kaplan, A.; Carmeli, S., Toxins and Biologically Active Secondary Metabolites of *Microcystis* sp. isolated from Lake Kinneret. *Israel J Chem* **2006,** *46* (1), 79-87.
86. Ploutno, A.; Carmeli, S., Nostocyclyne A, a Novel Antimicrobial Cyclophane from the Cyanobacterium *Nostoc* sp. *J Nat Prod* **2000,** *63* (11), 1524-1526.
87. Nakamura, H.; Hamer, H. A.; Sirasani, G.; Balskus, E. P., Cylindrocyclophane biosynthesis involves functionalization of an unactivated carbon center. *J Am Chem Soc* **2012,** *134* (45), 18518-21.
88. Kleigrewe, K.; Almaliti, J.; Tian, I. Y.; Kinnel, R. B.; Korobeynikov, A.; Monroe, E. A.; Duggan, B. M.; Di Marzo, V.; Sherman, D. H.; Dorrestein, P. C.; Gerwick, L.; Gerwick, W. H., Combining Mass Spectrometric Metabolic Profiling with Genomic Analysis: A Powerful Approach for Discovering Natural Products from Cyanobacteria. *J Nat Prod* **2015,** *78* (7), 1671-82.
89. Nakamura, H.; Schultz, E. E.; Balskus, E. P., A new strategy for aromatic ring alkylation in cylindrocyclophane biosynthesis. *Nat Chem Biol* **2017,** *13* (8), 916-921.
90. Schultz, E. E.; Braffman, N. R.; Luescher, M. U.; Hager, H. H.; Balskus, E. P., Biocatalytic Friedel-Crafts Alkylation Using a Promiscuous Biosynthetic Enzyme. *Angew Chem Int Ed Engl* **2019,** *58* (10), 3151-3155.
91. Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustaquio, A. S.; El Gamal, A. A.; Moore, B. S., Enzymatic Halogenation and Dehalogenation Reactions: Pervasive and Mechanistically Diverse. *Chem Rev* **2017,** *117* (8), 5619-5674.
92. Stalcup, A. M.; Martire, D. E.; Wise, S. A., Thermodynamic comparison of monomeric and polymeric C18 bonded phases using aqueous methanol and acetonitrile mobile phases. *J Chromatogr A* **1988,** *442*, 1-14.
93. Fischer, E.; Speier, A., Darstellung der Ester. *Berichte der deutschen chemischen Gesellschaft* **1895,** *28* (3), 3252-3258.
94. Sharp, P. M.; Li, W. H., The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* **1987,** *15* (3), 1281-1295.

95. Schmidt, T. G. M.; Skerra, A., The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* **2007,** *2*, 1528.
96. Kibbe, W. A., OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* **2007,** *35* (Web Server issue), W43-W46.
97. Otera, J.; Nishikido, J., Conversion of Carboxylic Acids into Esters without Use of Alcohols. *Esterification: Methods, Reactions and Applications (second edition)* **2010**; pp 173-192.
98. Haslam, E., Recent developments in methods for the esterification and protection of the carboxyl group. *Tetrahedron* **1980,** *36* (17), 2409-2433.
99. Fischbach, M. A.; Walsh, C. T., Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* **2006,** *106* (8), 3468-96.
100. Arora, P.; Goyal, A.; Natarajan, V. T.; Rajakumara, E.; Verma, P.; Gupta, R.; Yousuf, M.; Trivedi, O. A.; Mohanty, D.; Tyagi, A.; Sankaranarayanan, R.; Gokhale, R. S., Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat Chem Biol* **2009,** *5* (3), 166-173.
101. Lortie, R., Esterification - Biological. In *Encyclopedia of Catalysis* I Horváth (Ed.) **2010**.
102. LaMattina, J. W.; Wang, B.; Badding, E. D.; Gadsby, L. K.; Grove, T. L.; Booker, S. J., NosN, a Radical S-Adenosylmethionine Methylase, Catalyzes Both C1 Transfer and Formation of the Ester Linkage of the Side-Ring System during the Biosynthesis of Nosiheptide. *J Am Chem Soc* **2017,** *139* (48), 17438-17445.
103. O'Hagan, D.; Schmidberger, J. W., Enzymes that catalyse SN2 reaction mechanisms. *Nat Prod Rep* **2010,** *27* (6), 900-918.
104. Liu, J. Q.; Kurihara, T.; Hasan, A. K.; Nardi-Dei, V.; Koshikawa, H.; Esaki, N.; Soda, K., Purification and characterization of thermostable and nonthermostable 2-haloacid dehalogenases with different stereospecificities from *Pseudomonas* sp. strain YL. *Appl Environ Microb* **1994,** *60* (7), 2389-93.
105. Liu, J. Q.; Kurihara, T.; Miyagi, M.; Esaki, N.; Soda, K., Reaction mechanism of L-2-haloacid dehalogenase of *Pseudomonas* sp. YL. Identification of Asp10 as the active site nucleophile by 18O incorporation experiments. *J Biol Chem* **1995,** *270* (31), 18309-12.
106. Ramos, V.; Morais, J.; Castelo-Branco, R.; Pinheiro, Â.; Martins, J.; Regueiras, A.; Pereira, A. L.; Lopes, V. R.; Frazão, B.; Gomes, D.; Moreira, C.; Costa, M. S.; Brûle, S.; Faustino, S.; Martins, R.; Saker, M.; Osswald, J.; Leão, P. N.; Vasconcelos, V. M., Cyanobacterial diversity held in microbial biological resource centers as a biotechnological asset: the case study of the newly established LEGE culture collection. *J Appl Phycol* **2018,** *30* (3), 1437-1451.

VI. Annexes

Annex Text 1 – Protein and DNA sequences

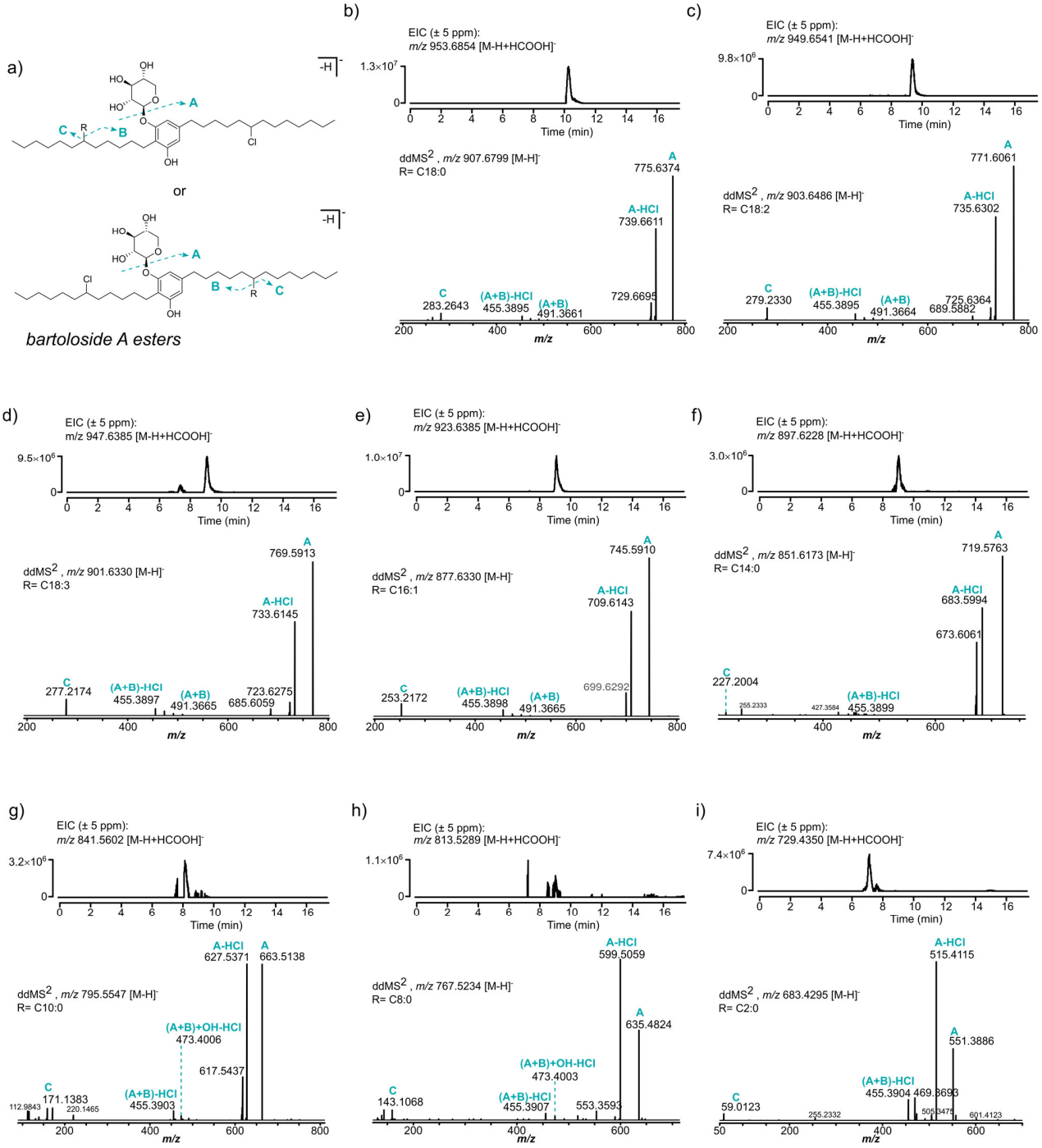
Protein Sequence (NStrep-BrtB, 91.4 kDa)

MASWSHPQFEKGADDDDKVPDPTSVDMANPFDVALTNLNSLYSGIPDVPVLQTYNSLNPYDISLTNLNNLSNFVIEPA
VLPTAAGTYGPQISEIAFALGGGIITGALADLLSTELGIEIENEALITKFLSFWTPGSDYAFGQDYVEFIQGREGFDT
IIGYDPGFNSALAPVQIDFVLGGPRQDFNFFAPLDEPARYVLGDWRTPYYVDNDIQSEGLNEFMYVSLFTLNPDEFSP
DNAENHRIQLYGSLDNYYFVPTTVVIDGSALDPSLPSIPVPGTEIYYKTGQVFADLIAIVPFQDPNDPNLISAWIDFE
GYLPPEQADLSGQTANPLLSGSPLLQLGYEGIDMGTATALAPDGDIYVAGSTSSLLGPFPRGGSDNFIARYNDDGSLE
WIVQFGSTEFDIITDIVSDSAGNVYATGWTRGNVETGFGLAPGVQDNWVAKFDANGSQLWVEQFTIDGYLDRSMGIAL
DEVNGRIHLTGHSNTDGSVADTNDPAIVPNGNVVTWIAGFDANTGAEDYRNVFEVAQTSNSRGRIDEGFGIDVDDAGN
VFSTGWAGENAYDVYLVKSDTDGAVLWTKSFGTGSNATQSYAWDVASDGTNAYILGWTQGELNTFRDQIPPTDEFAGN
ATAPLTPQNAYQGGTFDAFVAAYDTDGTELWTWHVGGTGDDGTFFGKIVAAGDFVYATGYTDGFIGTLGGSNAGDYDA
WIGKFDKLTGQTAWIQQIGSTKLDYATGISVDGDDIFVTGFTEGSLGGLNGGASDAWVAKLNQDGALEAFNVSAPIAP
PAPFAPPAPLTPPAPPAPDLSSLQSLGSTPQPTYPALGDSYGTNPFGTQPTQPTYDLSSMYQNAYYGGFGGYSGYGY

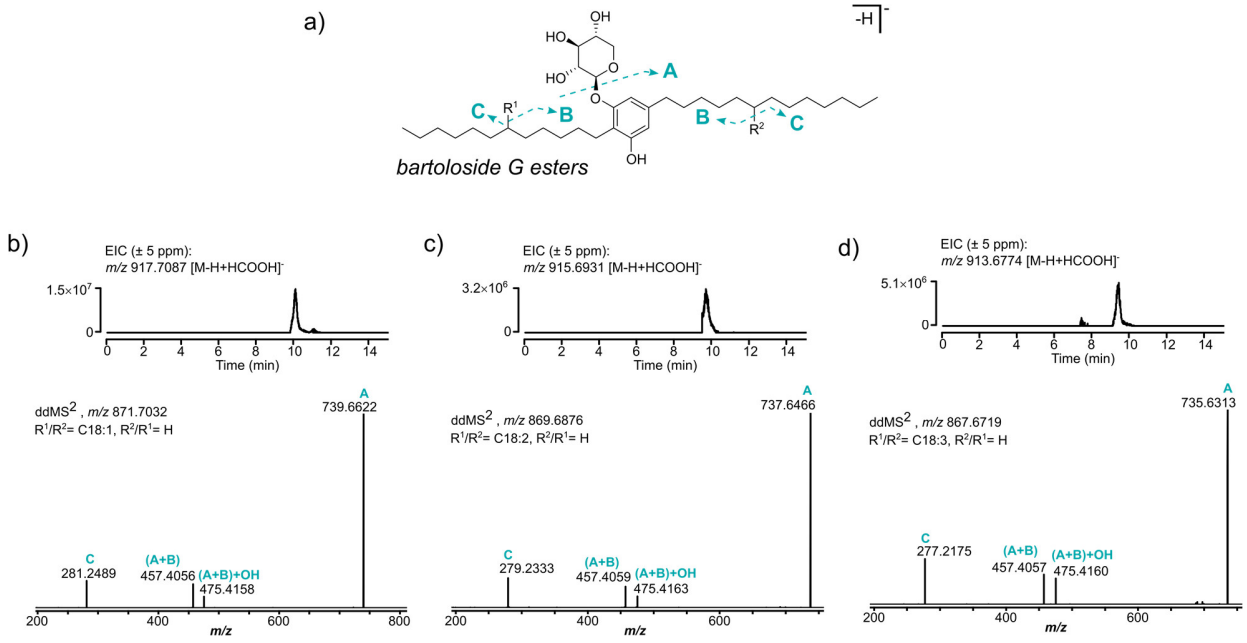
DNA sequence (NStrep-*brtB*)

ATGGCAAGCTGGAGCCACCCGCAGTTCGAAAAGGGTGCAGATGACGACGACAAGGTACCGGATCCGACGTCGGTCGAC
ATGGCCAATCCCTTCGACGTCGCTTTAACTAATCTTAACAGTCTTTACAGCGGTATCCCTGATGTTCCCGTTCTTCAG
ACGTACAACTCTCTTAACCCCTACGACATTTCTTTAACTAATCTTAACAATCTTAGCAACTTCGTCATTGAGCCAGCC
GTGCTCCCTACAGCGGCAGGTACCTACGGCCCTCAAATCAGTGAAATTGCTTTTGCGTTGGGGGGTGGGATCATAACA
GGGGCACTAGCTGATCTGCTATCCACTGAGTTGGGGATCGAGATCGAGAACGAAGCTCTGATAACTAAGTTCCTTTCC
TTCTGGACTCCCGGTTCCGATTACGCCTTCGGTCAGGATTATGTTGAATTTATTCAAGGTCGGGAAGGCTTTGACACC
ATCATCGGTTACGACCCCGGATTCAACAGTGCATTAGCTCCCGTACAAATTGACTTTGTCCTTGGCGGCCCCAGGCAG
GACTTTAACTTCTTTGCCCCCTTAGATGAGCCAGCACGCTACGTTTTGGGTGATTGGCGGACCCCCTATTATGTGGAT
AACGATATCCAGAGTGAGGGGCTGAACGAGTTTATGTATGTTTCGCTCTTCACACTTAACCCGGATGAATTTTCGCCA
GACAATGCAGAAAATCATCGTATCCAGCTCTACGGCAGTCTGGACAATTACTACTTTGTACCAACGACAGTTGTAATT
GATGGAAGCGCCCTAGATCCAAGTCTTCCATCAATTCCCGTCCCAGGCACAGAAATTTACTATAAAACCGGGCAGGTT
TTTGCAGATCTGATTGCCATTGTGCCTTTCCAGGATCCTAACGACCCAAATCTTATTTCCGCCTGGATTGATTTTGAG
GGCTATCTGCCCCCGGAACAGGCCGACCTGAGTGGCCAGACCGCGAACCCCTTGCTCTCCGGTTCTCCCCTCCTTCAA
TTGGGCTACGAAGGCATCGATATGGGTACGGCAACGGCCCTCGCCCCGGATGGTGATATCTACGTCGCTGGCTCCACT
AGCAGTCTTTTGGGACCTTTTCCCCGTGGAGGTTCAGACAACTTCATCGCCCGCTACAACGACGATGGTAGCTTGGAA
TGGATTGTTCAATTTGGTAGCACTGAGTTCGACATAATCACCGACATTGTCAGCGACAGCGCTGGTAACGTTTATGCC
ACCGGTTGGACCCGGGGCAATGTGGAGACGGGTTTTGGCCTGGCCCCAGGTGTTCAAGACAATTGGGTAGCCAAATTT
GATGCCAATGGCAGCCAACTCTGGGTAGAGCAATTTACCATTGATGGCTACCTCGACCGTTCCATGGGCATTGCCCTC
GATGAAGTCAATGGGAGAATTCACCTGACAGGGCACTCCAACACGGATGGCTCCGTTGCGGATACAAACGACCCCGCC

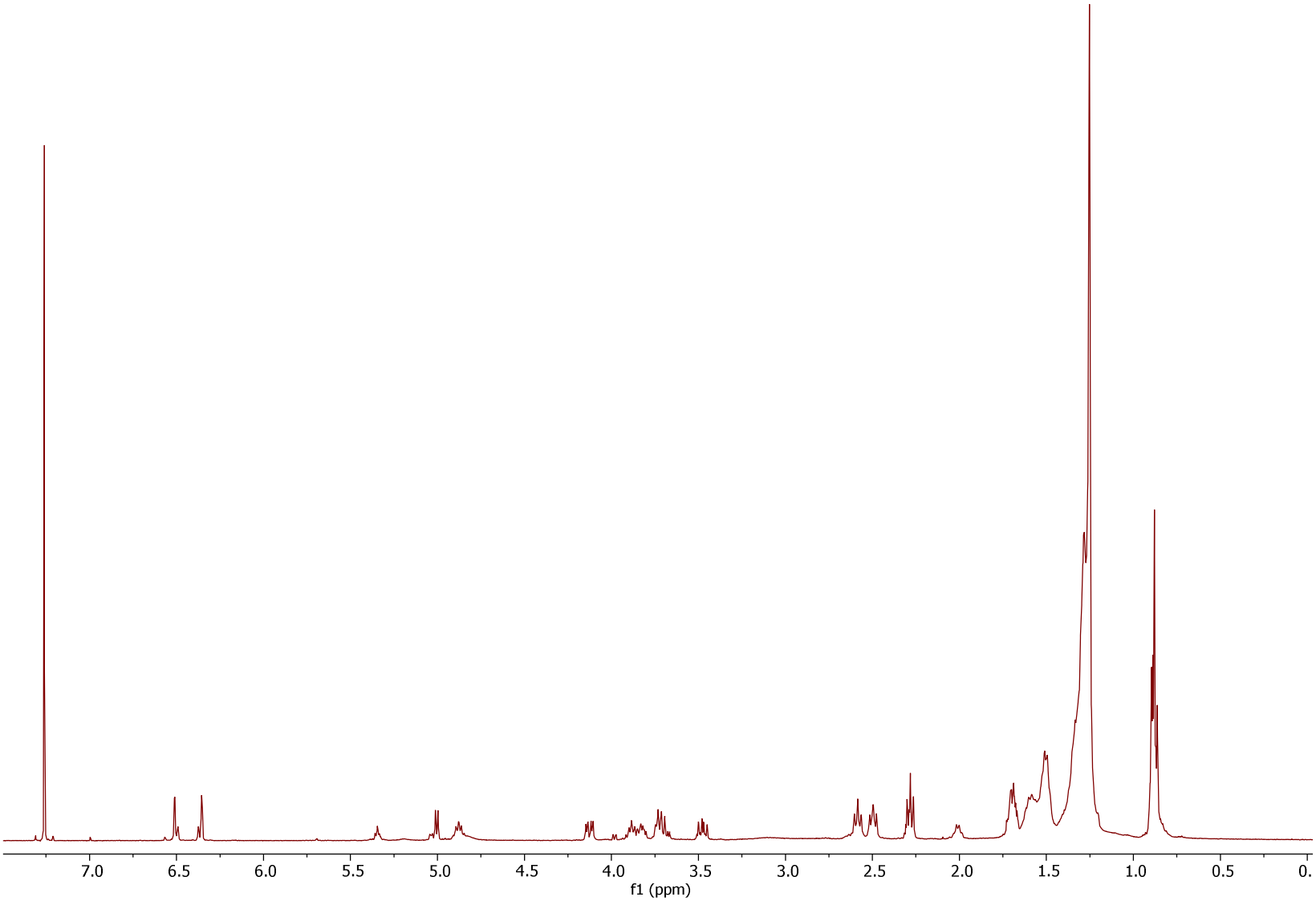
ATTGTCCCCAATGGTAATGTGGTCACCTGGATTGCTGGTTTTGATGCTAACACTGGAGCAGAAGACTATCGCAATGTC
TTTGAAGTCGCCCAAACCTCTAACAGCCGGGGACGCATTGATGAAGGTTTTGGCATAGATGTTGACGATGCTGGCAAT
GTGTTCTCCACCGGTTGGGCCGGAGAAAACGCCTACGACGTCTATCTGGTTAAATCCGATACGGATGGCGCTGTACTC
TGGACCAAGAGCTTCGGAACTGGGTCTAATGCCACCCAAAGTTATGCCTGGGATGTTGCTAGTGATGGCACCAATGCC
TATATCCTAGGTTGGACCCAGGGCGAACTCAACACATTCCGCGATCAAATTCCTCCCACGGATGAATTTGCGGGTAAT
GCGACGGCTCCACTTACTCCGCAAAATGCTTACCAAGGTGGTACCTTCGATGCCTTTGTTGCCGCCTATGACACCGAT
GGTACTGAACTCTGGACTTGGCATGTAGGTGGAACAGGGGATGATGGAACCTTCTTCGGTAAAATTGTCGCTGCTGGG
GACTTTGTTTACGCCACAGGCTATACCGATGGCTTTATTGGCACCCTAGGTGGCAGTAACGCTGGTGACTACGATGCC
TGGATCGGTAAGTTTGATAAGTTAACCGGTCAGACTGCATGGATCCAGCAGATTGGTAGCACCAAGCTAGACTACGCC
ACTGGTATTTCCGTTGATGGCGATGACATTTTCGTAACAGGCTTTACGGAAGGTTCCTTGGGCGGCTTGAACGGTGGT
GCTTCCGACGCCTGGGTTGCTAAGTTAAATCAAGATGGCGCTTTAGAAGCTTTTAATGTTTCTGCCCCTATTGCTCCC
CCAGCTCCTTTTGCTCCCCCAGCCCCTCTTACTCCTCCAGCTCCCCCGGCTCCTGATCTCAGTAGCTTGCAGTCATTG
GGTAGCACCCCTCAGCCGACCTACCCGGCCCTTGGGGACAGCTATGGCACTAACCCCTTTGGTACTCAGCCTACTCAA
CCTACCTATGACCTCAGCAGCATGTACCAAAATGCTTACTATGGCGGCTTCGGAGGTTATTCCGGGTACGGCTACTAG



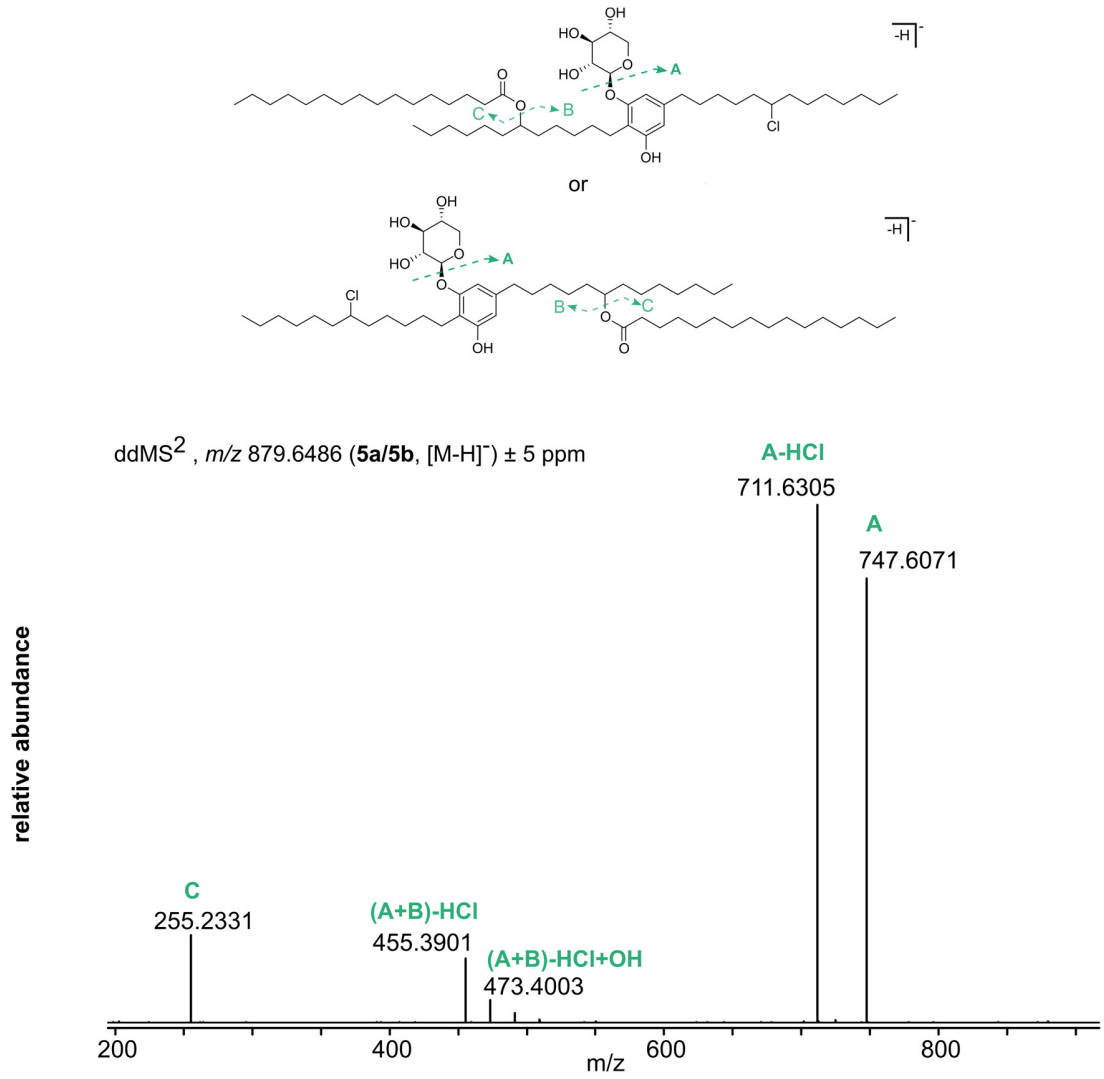
Annex Figure 1. Detection and HRESIMS/MS based structural assignment of naturally occurring bartoloside A esters produced by *S. salina* LEGE06099. a) Annotated MS/MS fragments for the structures of bartoloside A and G esters. b-i) Extracted Ion Chromatograms (EICs) for the $[M-H+HCOOH]^-$ ion and MS/MS spectra for the $[M-H]^-$ specie of some of the detected naturally occurring bartoloside A esters.



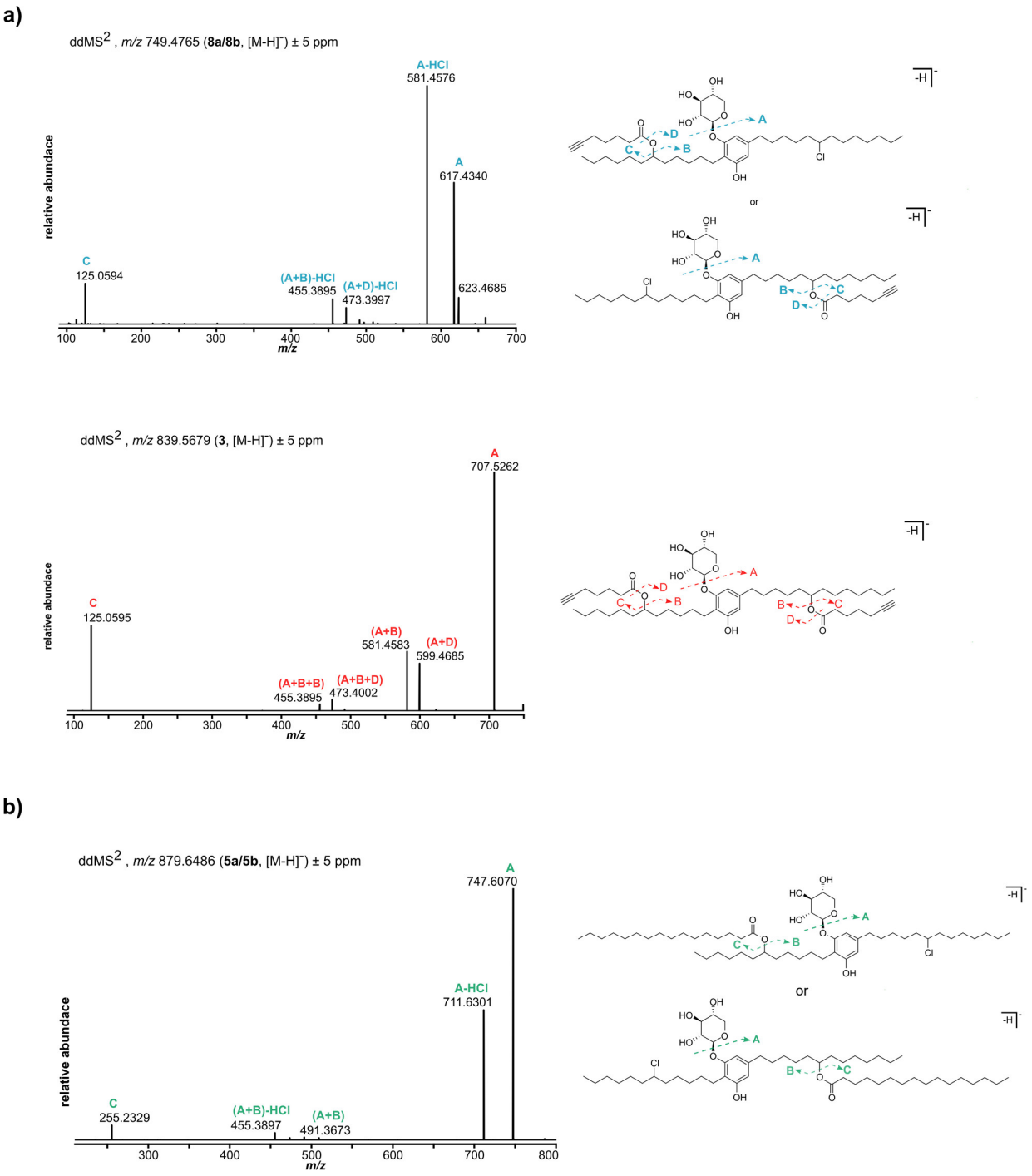
Annex Figure 2. Detection and HRESIMS/MS based structural assignment of naturally occurring bartoloside G esters produced by *S. salina* LEGE06099. a) Annotated MS/MS fragments for the structures of bartoloside A and G esters. b-d) Extracted Ion Chromatograms (EICs) for the $[M-H+HCOOH]^-$ ion and MS/MS spectra for the $[M-H]^-$ specie of some of the detected naturally occurring bartoloside G esters.



Annex Figure 3. ^{1}H NMR spectra for subfraction 15.7 obtained during the isolation of bartoloside A monopalmitate.

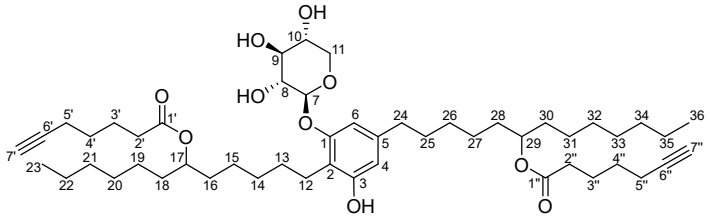


Annex Figure 4. HRESIMS/MS of fraction 15.7.2 obtained during the isolation of compound **5a** with respective structural assignment (top).



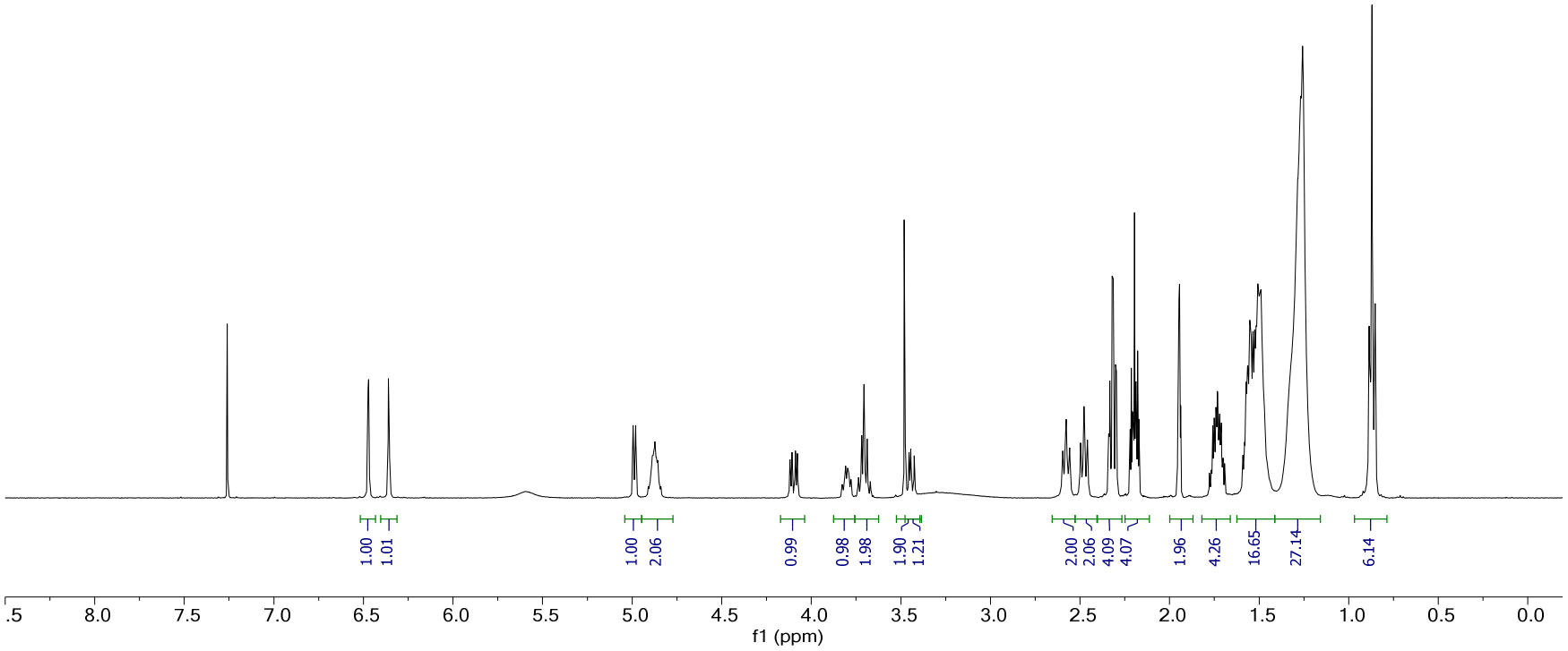
Annex Figure 5. LC-HRESIMS/MS analysis of the bartoloside esters **3** and **8a/8b** (a) and **5a** (b) formed in the full assay. It was not possible to fragment the diester **7**due to low abundance of the precursor ion.

Annex Table 1. NMR Spectroscopic Data (^{1}H 400 MHz, ^{13}C 100 MHz, $CDCl_3$) for bartoloside A-17,29-diyl bis(hept-6yonate) (**3**).

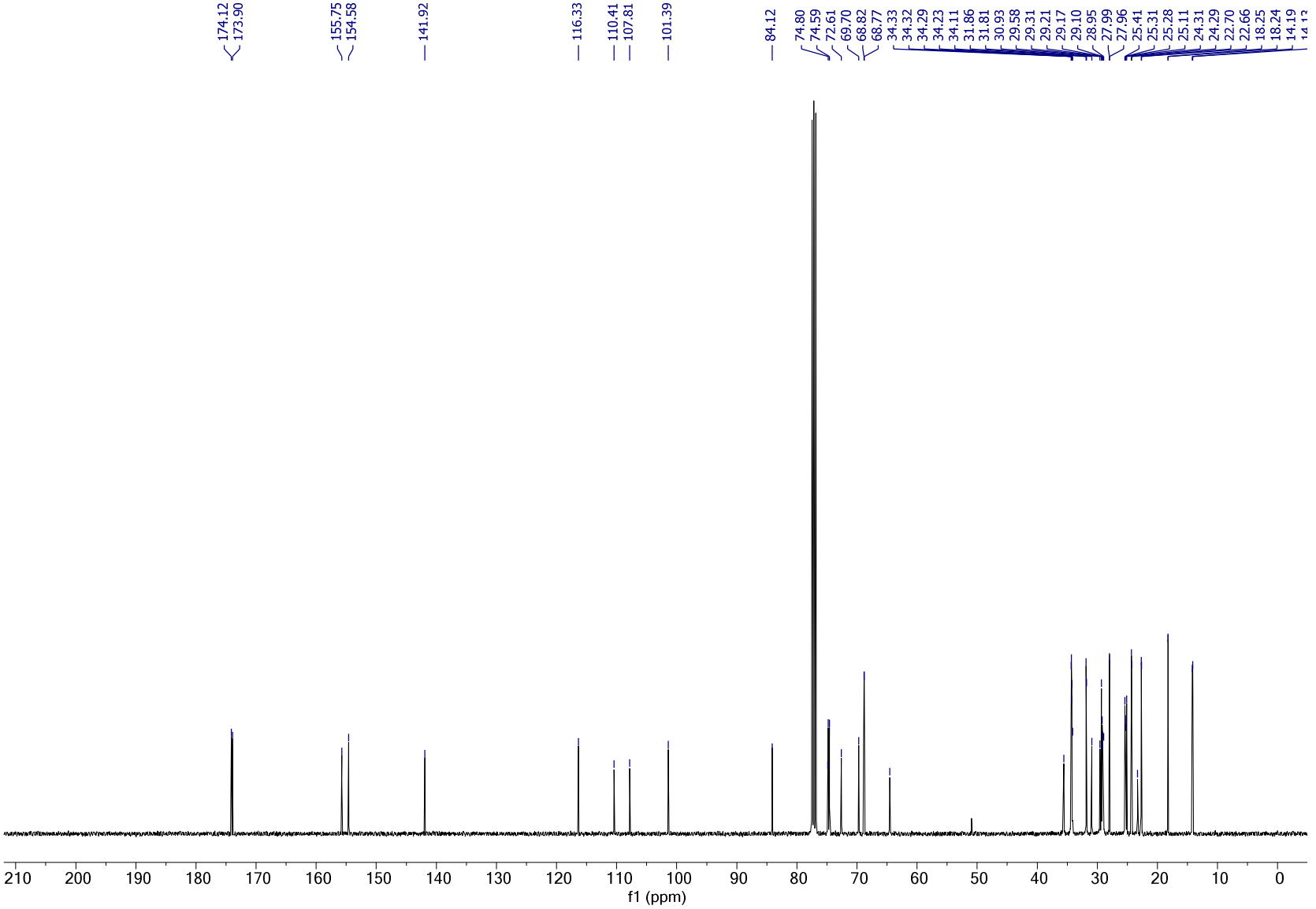


position	δc	type	δH[a]	mult, *J* (Hz)	HMBC[b]	COSY
1	155.8	C				
2	116.3	C				
3	154.6	C-OH				
4	110.4	CH	6.36	d, 1.4	154.6, 116.3, 107.8, 35.6	6.47, 2.48
5	141.9	C				
6	107.8	CH	6.47	d, 1.4	155.8, 116.3, 110.4, 35.6	6.36, 2.48
7	101.4	CH	4.99	d, 5.6	155.8, 74.9, 64.5	3.71, 3.70
8	72.6	CH-OH	3.71	m	101.4, 74.9	4.99, 3.80
9	74.9	CH-OH	3.70	m	72.6, 69.7	4.99, 3.80
10	69.7	CH-OH	3.80	m	74.9	4.10, 3.71, 3.70, 3.45
11	64.5	CH_2	4.10/3.45	dd, 11.9, 4.4/12.0, 8.2	101.4, 74.9, 69.7	3.80, 3.45/4.10, 3.80
12	23.3	CH_2	2.58	t, 7.8	155.8, 154.6, 116.3, 29.3	1.51-1.49, 1.50
13	29.3	CH_2	1.50	m		2.58
14, 20, 26, 32, 33	29.6-29.0	5 × CH_2	1.32-1.28	m	31.9, 31.8, 29.6-29.0, 14.2, 14.1	1.51-1.49, 0.87
15a, 19a, 27a, 31a	25.4-25.1	4 × CH_2	1.51-1.49	m	74.8, 74-6, 31.9, 31.8, 29.6-29.0	4.87, 2.58, 1.32-1.28, 1.27-1.24
15b, 19b, 27b, 31b			1.27-1.24	m	31.9, 31.8, 29.6-29.0	1.51-1.49
16a, 18a, 28a, 30a	34.3-34.1	4 × CH_2	1.51-1.49	m	74.8, 74.6, 31.9, 31.8, 29.6-29.0, 25.4-25.1	4.87, 1.32-1.28, 1.27-1.24
16b, 18b, 28b, 30b			1.27-1.24	m	29.6-29.0, 25.4-25.1	1.51-1.49
17/29	74.8/74.6	2 × CH	4.87	m	174.1, 173.9, 34.3-34.1, 25.4-25.1	1.51-1.49
21/34	31.9/31.8	2 × CH_2	1.27-1.24	m	29.6-29.0, 25.4-25.1, 14.2, 14.1	1.51-1.49, 0.87
22/35	22.7	2 × CH_2	1.27-1.24	m	31.9, 31.8, 29.6-29.0, 25.4-25.1, 14.2, 14.1	1.51-1.49, 0.87
23/36	14.2/14.1	2 × CH_3	0.87	m	31.9, 31.8, 22.7	1.32-1.28, 1.27-1.24
24	35.6	CH_2	2.48	t, 7.5	141.9, 110.4, 107.8, 34.3-34.1, 30.9, 29.6-29.0	6.47, 6.36, 1.55
25	30.9	CH_2	1.55	m	141.9, 29.6-29.0	2.48
1'/ 1''	174.1/173.9	2 × C				
2'/ 2''	34.3/34.3	2 × CH_2	2.32	dt, 7.4, 2.2	174.1, 173.9, 28.0, 24.3	1.73
3'/3''	24.3	2 × CH_2	1.73	m	174.1, 173.9, 34.3, 28.0, 18.3, 18.2	2.32, 1.55
4'/ 4''	28.0	2 × CH_2	1.55	m	84.1, 34.3, 24.3, 18.3, 18.2	2.20, 1.73
5'/ 5''	18.3/18.2	2 × CH_2	2.20	tt, 7.0, 2.7	84.1, 68.8, 28.0, 24.3	1.95, 1.55
6'/ 6''	84.1	2 × C				
7'/ 7''	68.8	2 × CH	1.95	m	84.1, 18.2, 18.3	2.20

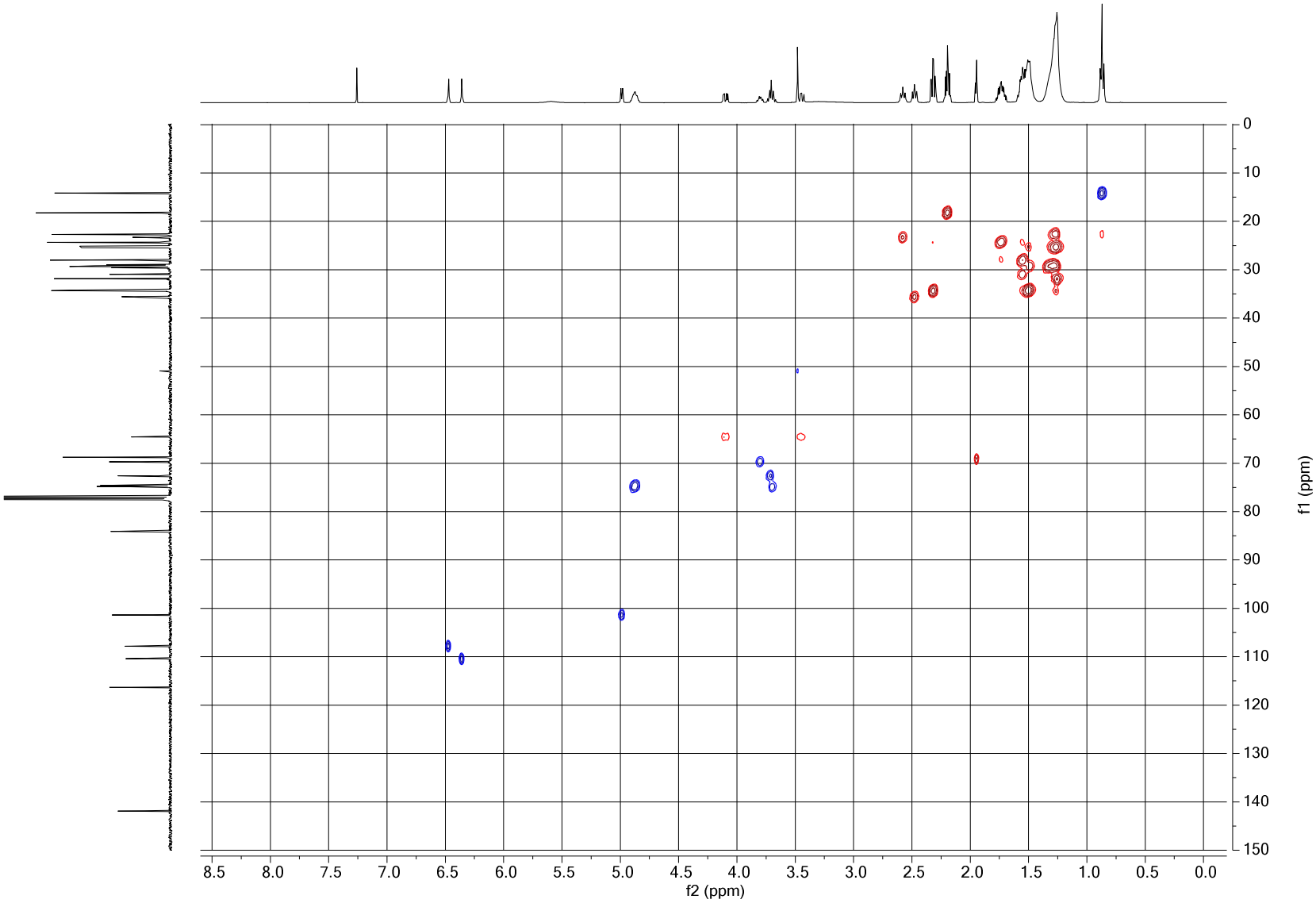
[a]from HSQC; [b]from proton to indicated carbon.



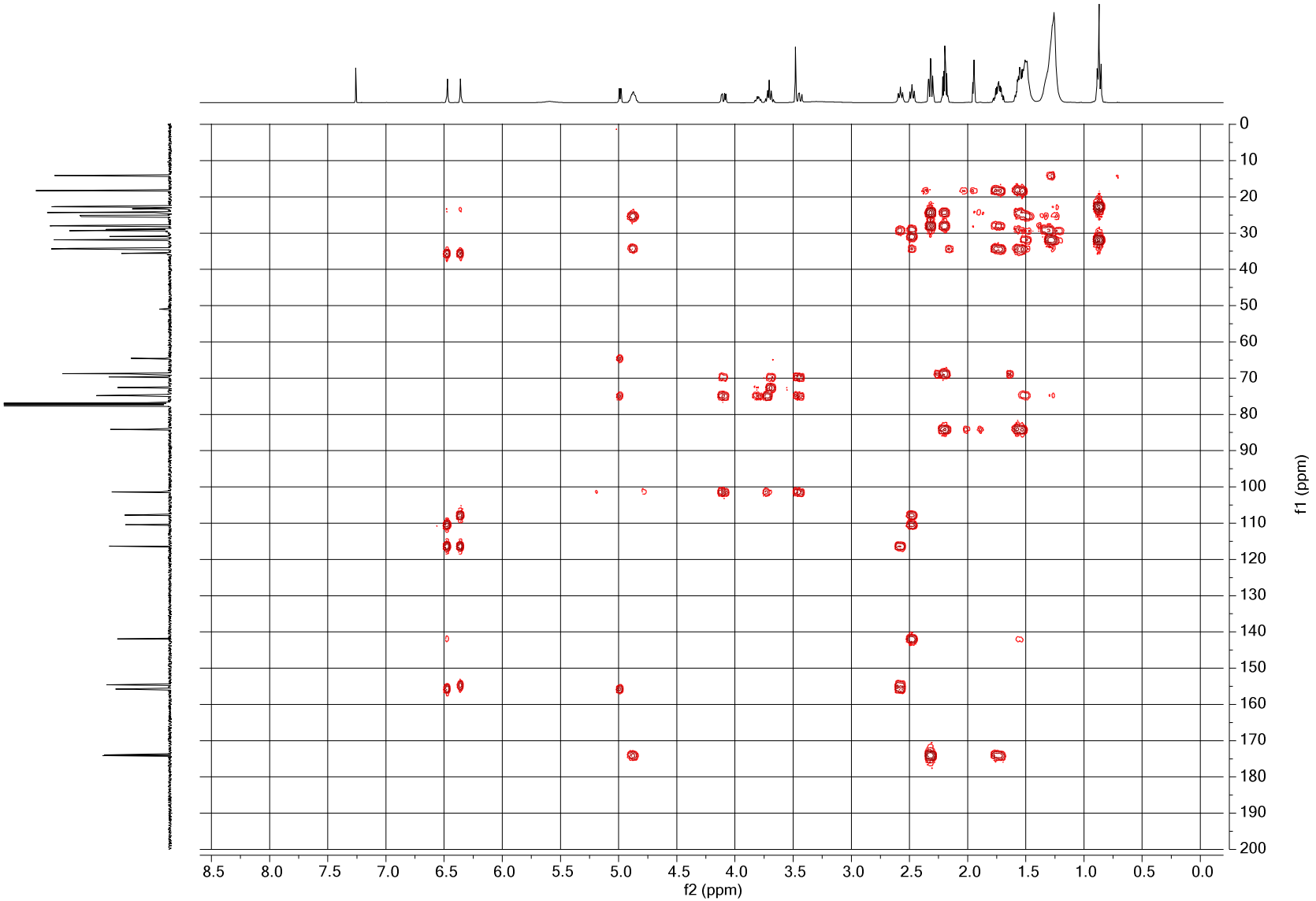
Annex Figure 6. ^{1}H NMR ($CDCl_3$, 400 MHz) spectrum of compound **3**.



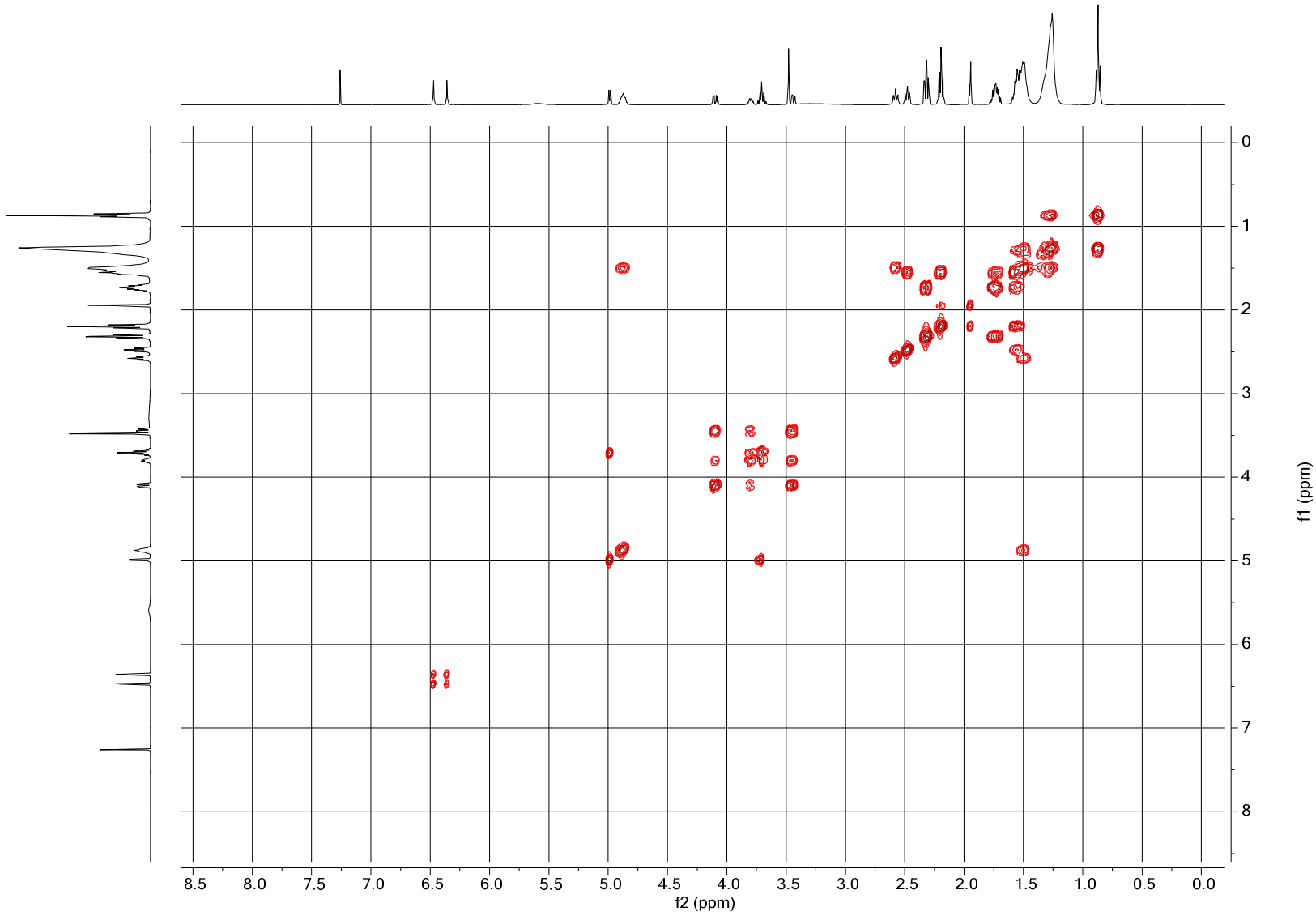
Annex Figure 7. ^{13}C NMR ($CDCl_3$, 100 MHz) spectrum of compound **3**.



Annex Figure 8. HSQC ($CDCl_3$, 400 MHz) spectrum of compound **3**.

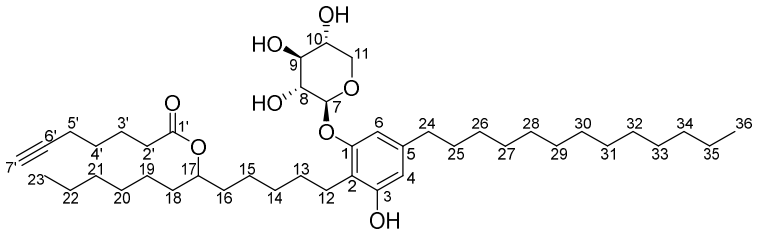


Annex Figure 9. HMBC ($CDCl_3$, 400 MHz) spectrum of compound **3**.



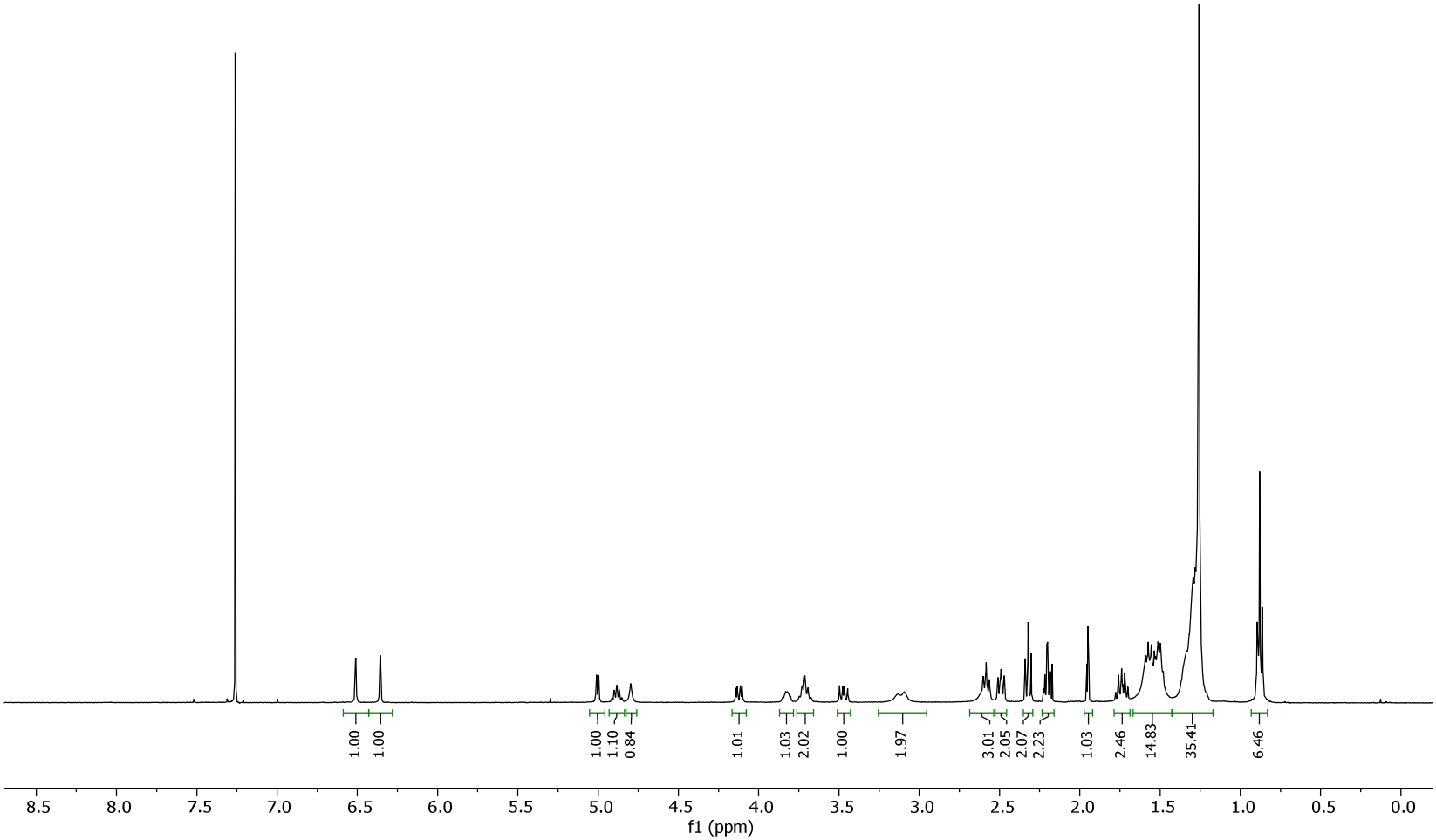
Annex Figure 10. COSY ($CDCl_3$, 400 MHz) spectrum of compound **3.**

Annex Table 2. NMR Spectroscopic Data (^{1}H 400 MHz, ^{13}C 100 MHz, $CDCl_3$) for bartoloside G-17-yl hept-6-ynoate (**4**)

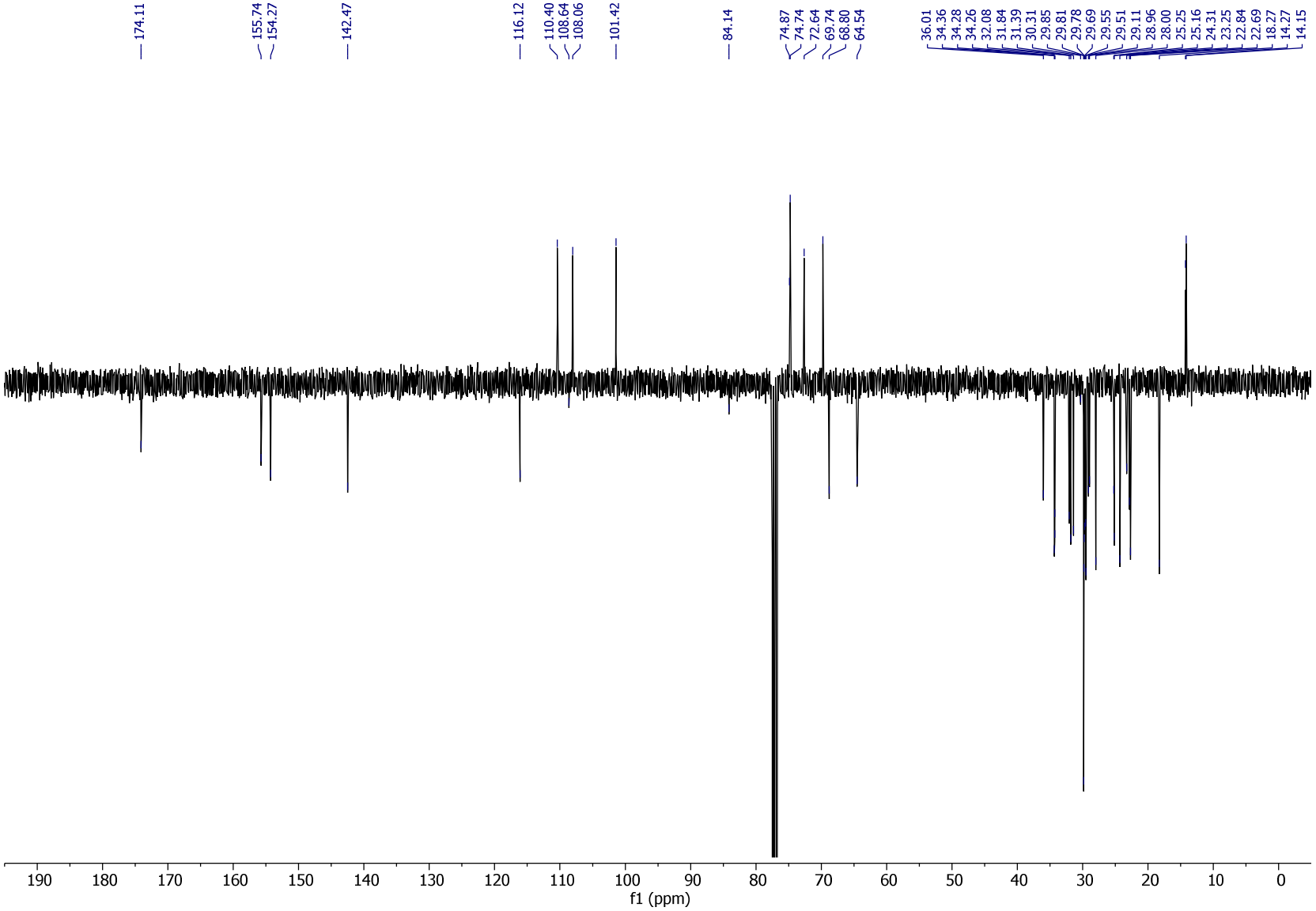


position	δc	type	δH[a]	mult, *J* (Hz)	HMBC[b]	COSY
1	155.7	C				
2	116.1	C				
3	154.3	C-OH				
4	110.4	CH	6.36	d, 1.5	154.3, 116.1, 108.1, 36.0	6.51, 2,49
5	142.5	C				
6	108.1	CH	6.51	d, 1.5	155.7, 116.1, 110.4, 36.0	6.36, 2.49
7	101.4	CH	5.00	d, 5.7	155.7, 74.9, 64.5	3.72, 3.70
8	72.6	CH-OH	3.72	m		5.00, 3.83
9	74.9	CH-OH	3.70	m		5.00, 3.83
10	69.7	CH-OH	3.83	m		4.12, 3.72, 3.70, 3.47
11	64.5	CH_2	4.12/3.47	dd, 11.9, 4.4/11.9, 8.1	101.4, 74.9/101.4, 74.9, 69.7	3.83, 3.47/4.12, 3.83
12	23.3	CH_2	2.58	t, 7.8	155.7, 154.3, 116.1, 29.1	1.51-1.50, 1.50
13	29.1	CH_2	1.50	m		2.58
14/20, 26-33	29.9-29.0	10 × CH_2	1.27-1.24	m	32.1, 31.8, 29.9-29.0, 22.8, 22.7, 14.3, 14.2	1.51-1.50, 0.87
15a, 19a	25.3/25.2	2 × CH_2	1.51-1.50	m	32.1, 31.8	4.88, 2.58, 1.30-1.28
15b, 19b			1.30-1.28	m	32.1, 31.8, 29.9-29.5	1.51-1.50
16a, 18a	34.3	2 × CH_2	1.51-1.50	m	32.1, 31.8, 25.3, 25.2	4.88, 1.30-1.28, 1.27-1.24
16b, 18b			1.27-1.24	m	29.9-29.0	1.51-1.50
17	74.7	CH	4.88	p, 6.2	174.1, 34.3, 25.3, 25.2	1.51-1.50
21/34	32.1/31.8	2 × CH_2	1.27-1.24	m	29.9-29.0, 22.8, 22.7, 14.3, 14.2	1.51-1.50, 0.87
22/35	22.8/22.7	2 × CH_2	1.27-1.24	m	32.1, 31.8, 29.85-29.51, 14.3, 14.2	1.51-1.50, 0.87
23/36	14.3/14.2	2 × CH_3	0.87	m	32.1, 31.8, 22.8, 22.7	1.30-1.28, 1.27-1.24
24	36.0	CH_2	2.49	t, 7.7	142.5, 110.4, 108.1, 31.4, 29.9-29.0	6.51, 6.36, 1.56
25	31.4	CH_2	1.56	m	29.9-29.0	2.49
1'	174.1	C				
2'	34.4	CH_2	2.32	t, 7.4	174.1, 28.0, 24.3	1.74
3'	24.3	CH_2	1.74	m	174.1, 34.4, 28.0, 18.3	2.32, 1.55
4'	28.0	CH_2	1.55	m	84.1, 34.4, 24.3, 18.3	2.20, 1.74
5'	18.3	CH_2	2.20	td, 7.0, 2.6	84.1, 68.8, 28.0, 24.3	1.95, 1.55
6'	84.1	C				
7'	68.8	CH	1.95	t, 2.7		2.20

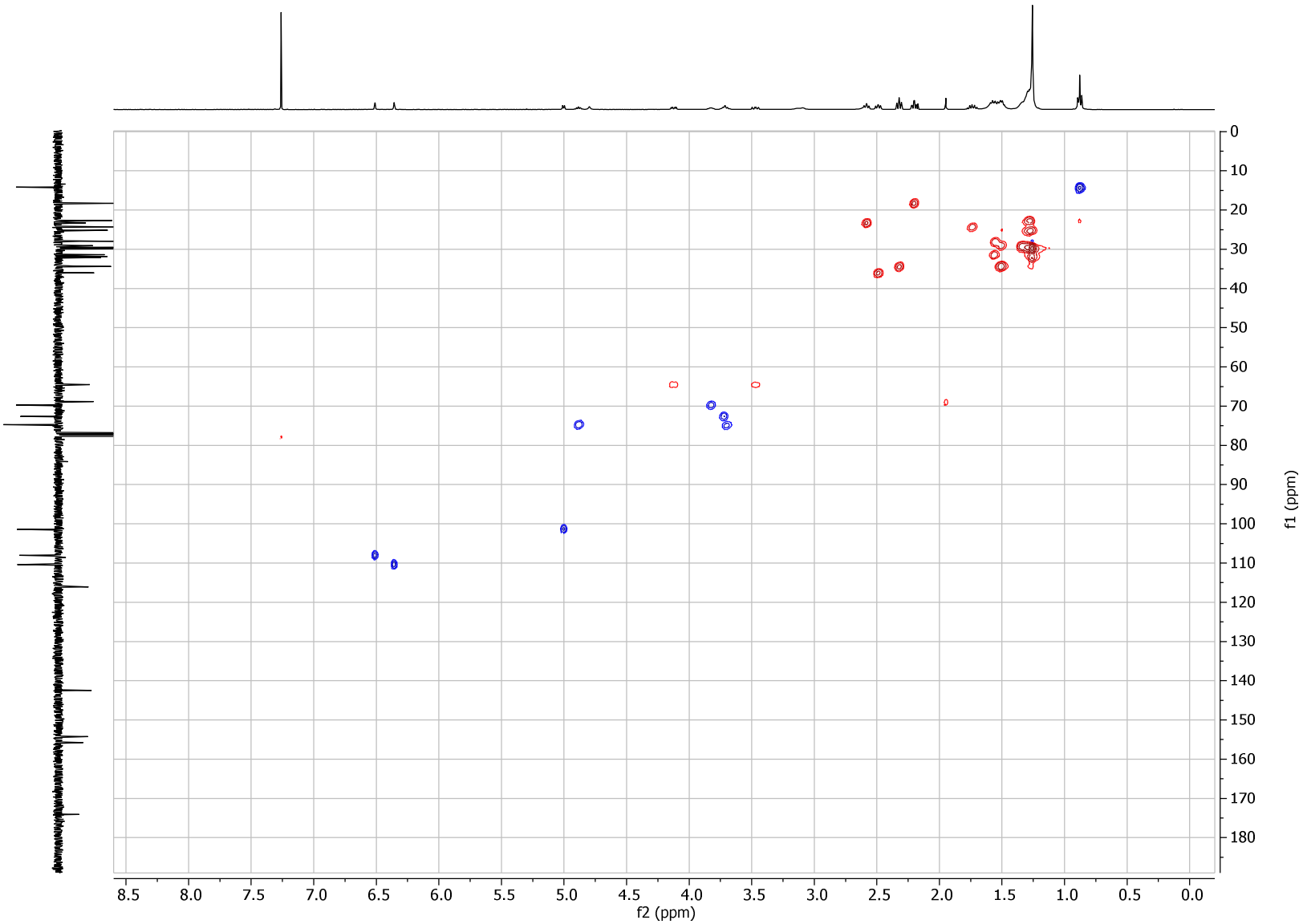
[a]from HSQC; [b]from proton to indicated carbon.



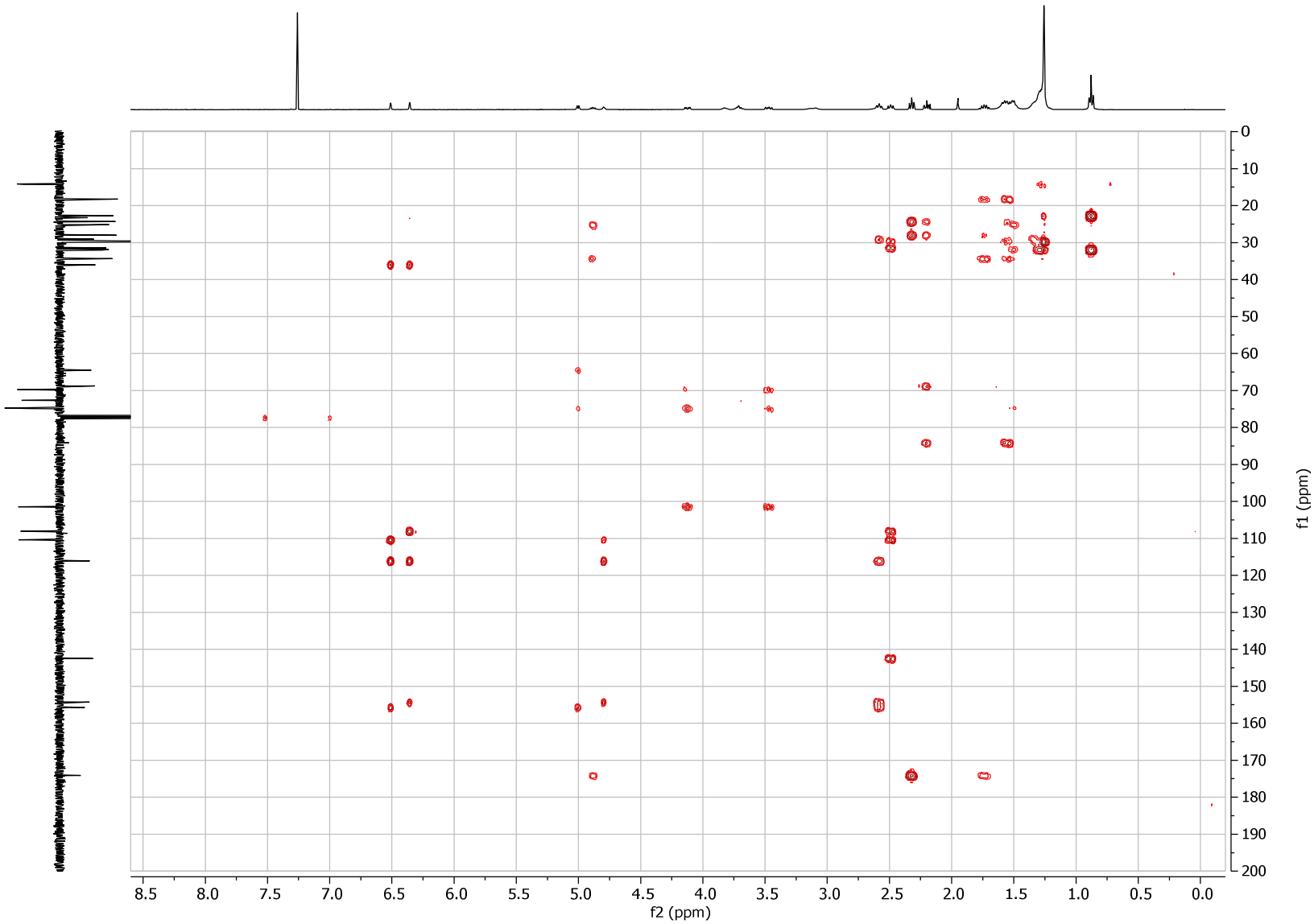
Annex Figure 11. ^{1}H NMR ($CDCl_3$, 400 MHz) spectrum of compound **4**.



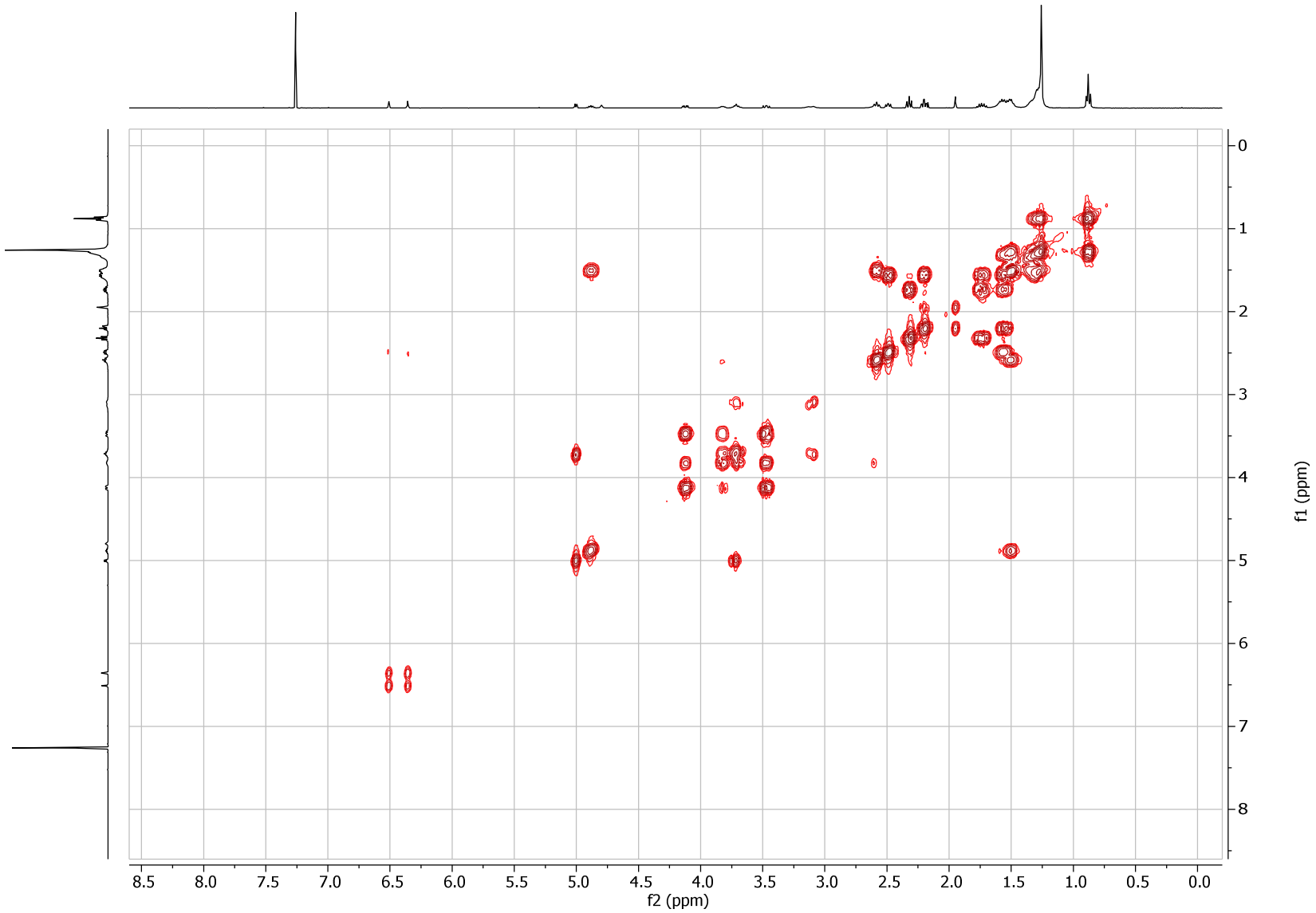
Annex Figure 12. ^{13}C NMR (APT, $CDCl_3$, 100 MHz) spectrum of compound **4**.



Annex Figure 13. HSQC ($CDCl_3$, 400 MHz) spectrum of compound **4**.

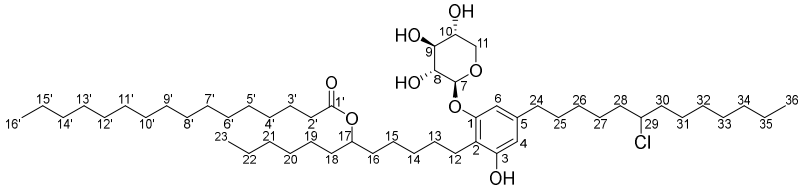


Annex Figure 14. HMBC ($CDCl_3$, 400 MHz) spectrum of compound **4**.



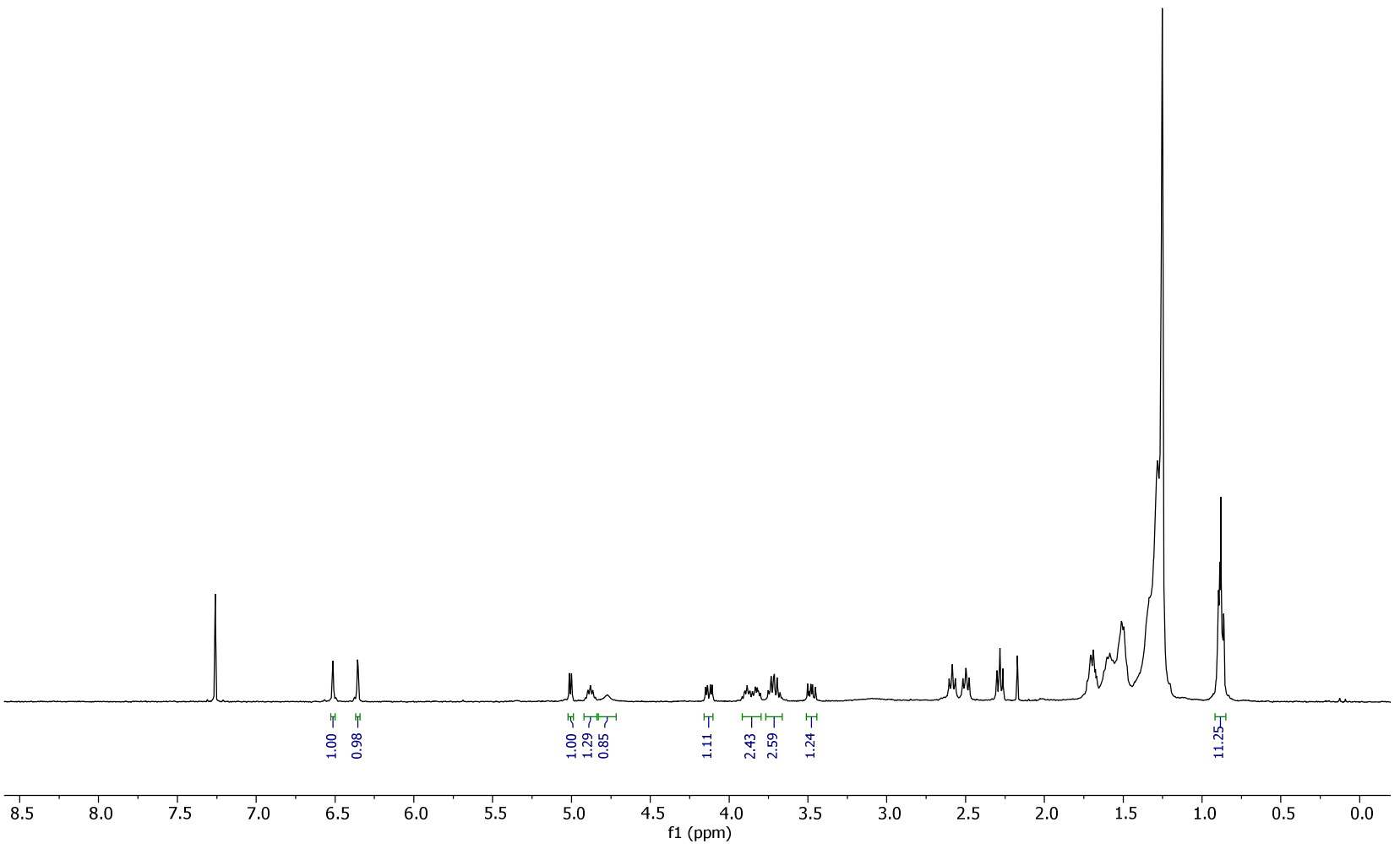
Annex Figure 15. COSY ($CDCl_3$, 400 MHz) spectrum of compound **4**.

Annex Table 3. NMR Spectroscopic Data (^{1}H 400 MHz, ^{13}C 100 MHz, $CDCl_3$) for bartoloside A-17-yl palmitate (**5a**)

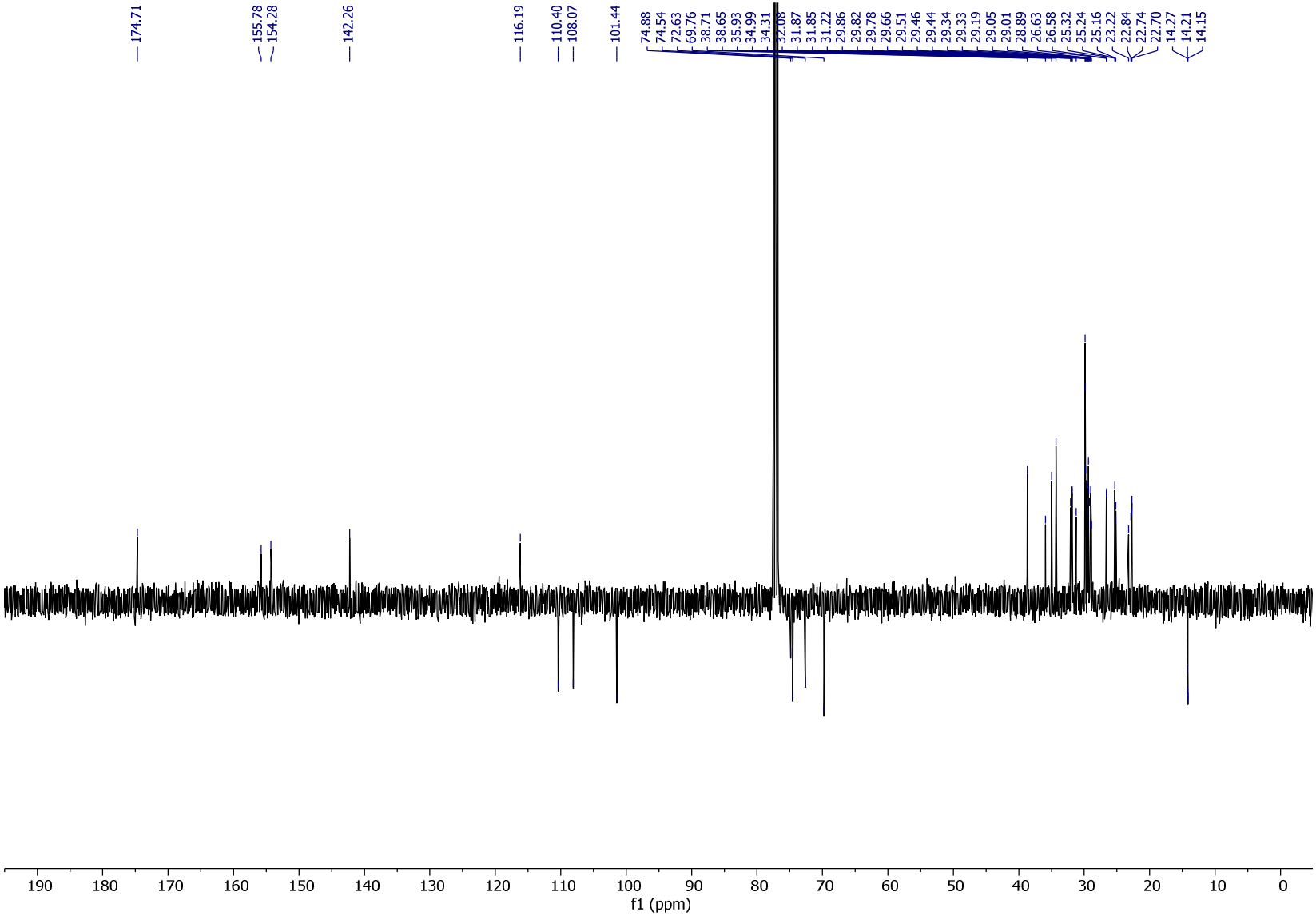


position	δc	type	δH[a]	mult, *J* (Hz)	HMBC[b]	COSY
1	155.8	C				
2	116.2	C				
3	154.3	C-OH				
4	110.4	CH	6.36	s	154.3, 116.2, 108.1, 35.9	6.51
5	142.3	C				
6	108.1	CH	6.51	s	155.8, 116.2, 110.4, 35.9	6.36
7	101.4	CH	5.01	d, 5.7	155.8	3.72, 3.70
8	72.6	CH-OH	3.72	m	101.4, 74.9	5.01, 3.83
9	74.9	CH-OH	3.70	m	72.6, 69.8	5.01, 3.83
10	69.8	CH-OH	3.83	m		4.13, 3.72, 3.70, 3.48
11	64.6	CH_2	4.13/3.48	dd, 11.9, 4.4/dd, 11.9, 8.0	101.4, 74.9/101.4, 69.8	3.83, 3.48/4.13, 3.83
12	23.2	CH_2	2.58	t, 7.7	155.8, 154.3, 116.2, 29.2	1.51, 1.51-1.50
13	29.2	CH_2	1.51	m		2.58
14, 20, 26, 32, 33, 4'-13'	29.9-28.9	15 × CH_2	1.27-1.24	m	29.9-28.9, 22.8-22.7	1.51-1.50, 0.88
15a, 19a	25.2	2 × CH_2	1.61-1.60		29.9-28.9	2.28, 1.30-1.28
15b, 19b			1.30-1.28		32.1-31.9, 29.9-28.9, 22.8-22.7	1.61-1.60, 1.51-1.50
16a, 18a	34.3	2 × CH_2	1.51-1.50	m		4.88, 1.30-1.28, 1.41-1.40
16b, 18b			1.27-1.24	m	29.9-28.9	1.51-1.50
17	74.5	CH	4.88	p, 6.4	174.7, 25.2	1.51-1.50
21, 34, 14'	32.1-31.9	3 × CH_2	1.27-1.24	m	29.9-28.9, 22.8-22.7	1.51-1.50, 0.88
22, 35, 15'	22.8-22.7	3 × CH_2	1.30-1.28	m	32.1-31.9, 29.9-28.9, 22.8-22.7	1.61-1.60, 0.88
23, 36, 16'	14.3-14.2	3 × CH_3	0.88	m	32.1-31.9, 22.8-22.7	1.30-1.28
24	35.9	CH_2	2.5	t, 7.7	142.3, 110.4, 108.1, 31.2	1.58
25	31.2	CH_2	1.58	m		2.5
27a, 31a	26.6	2 × CH_2	1.51-1.50	m		1.7, 1.41-1.40
27b, 31b			1.41-1.40	m		1.51-1.50
28/30	38.7	2 × CH_2	1.70	m	64.5	3.88, 1.51-1.50
29	64.5	CH-Cl	3.88	m	26.6	1.70
1'	174.7	C				
2'	35.0	CH_2	2.28	t, 7.5	174.7, 29.9-28.9, 25.3	1.61-1.60
3'	25.3	CH_2	1.61-1.60	m	29.9-28.9	2.28

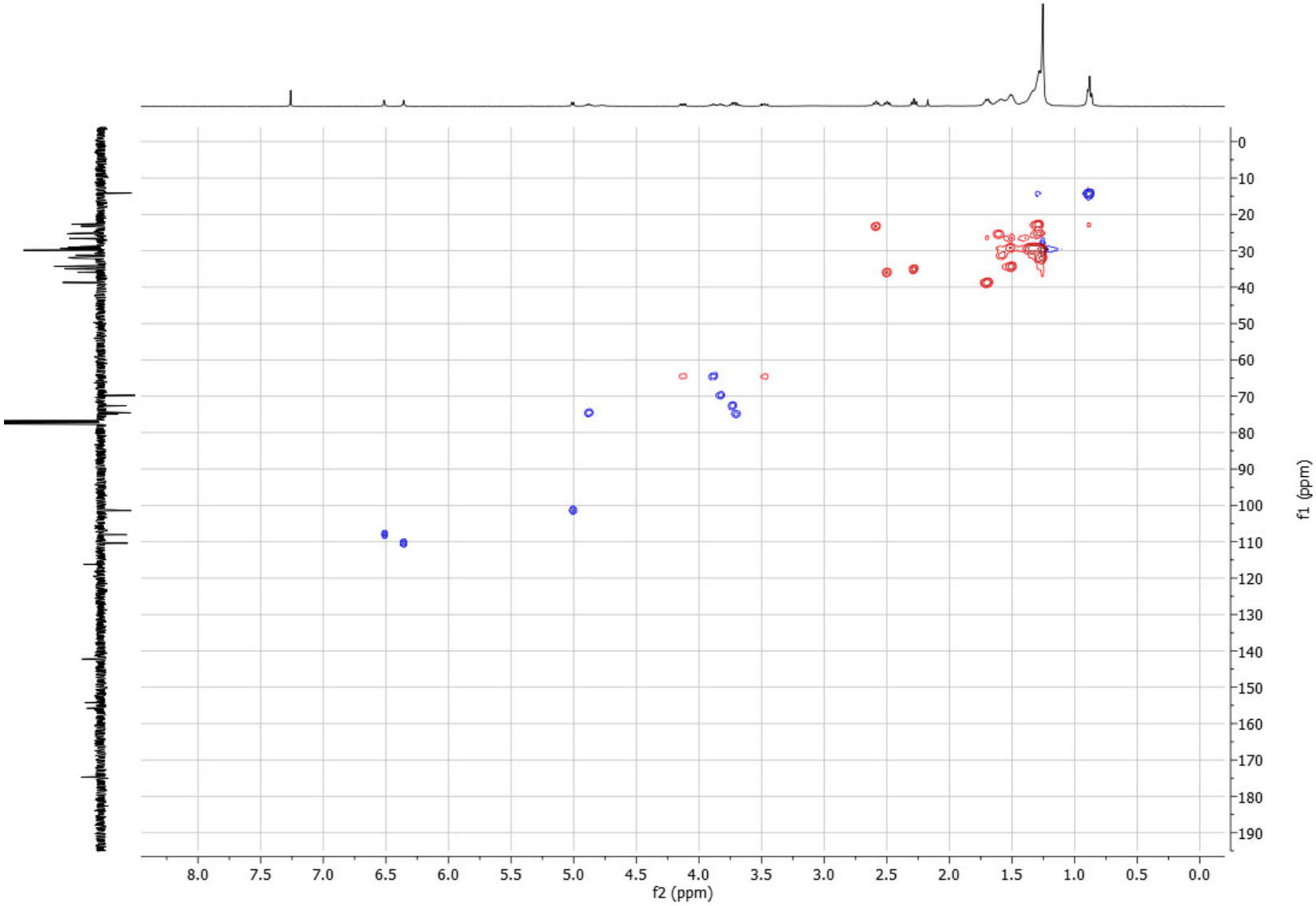
[a]from HSQC; [b]from proton to indicated carbon.



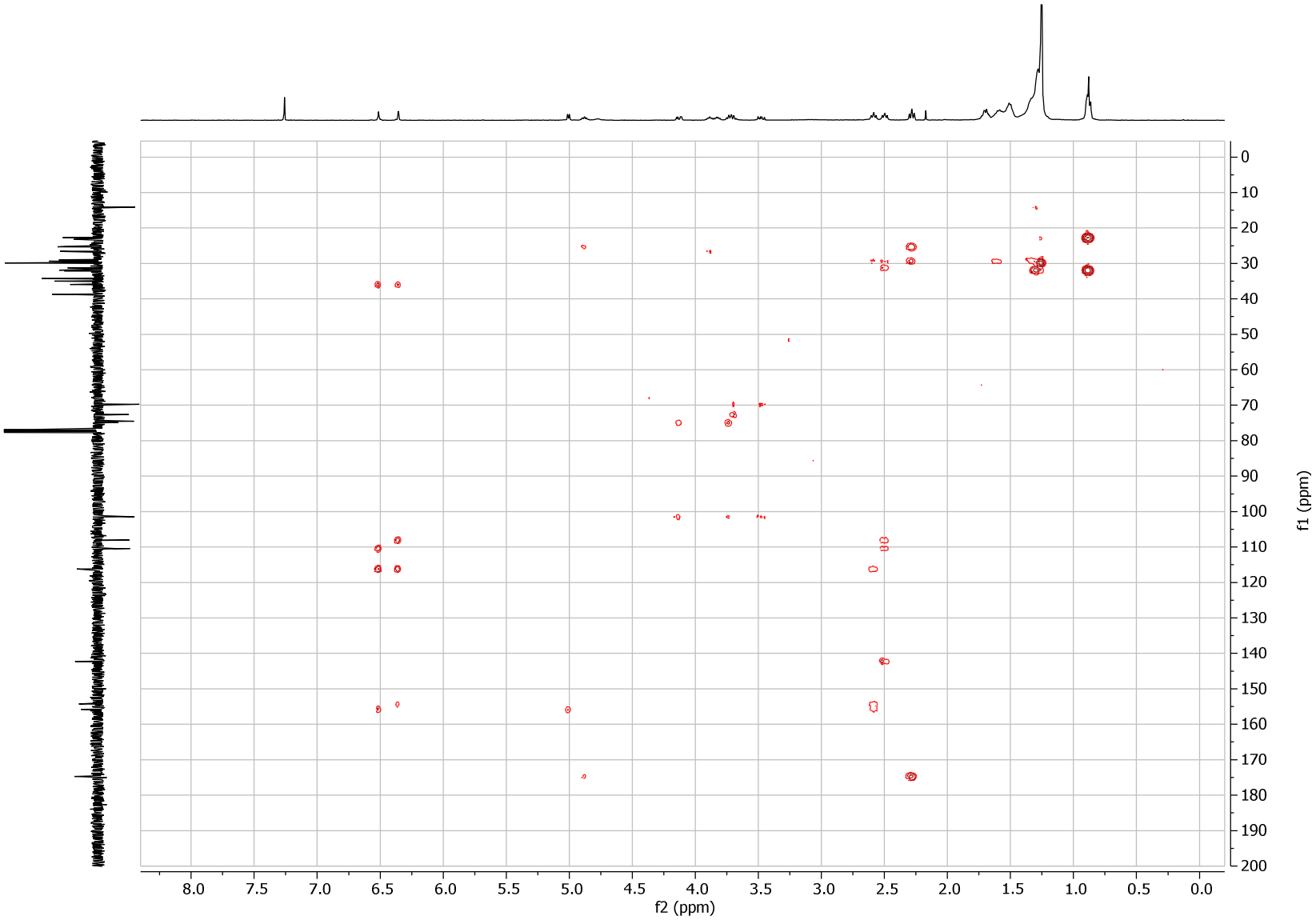
Annex Figure 16. ^{1}H NMR ($CDCl_3$, 400 MHz) spectrum of compound **5a**.



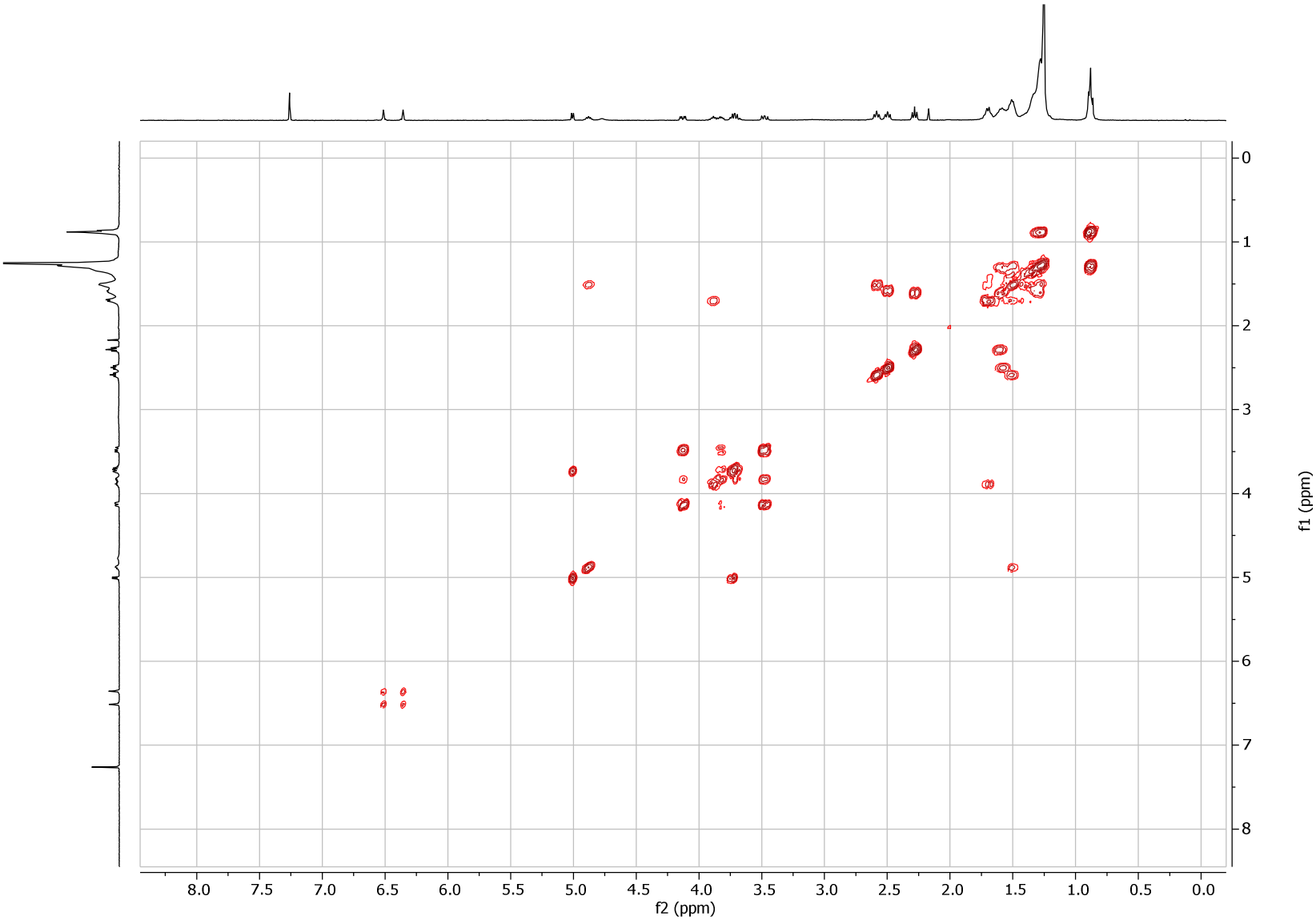
Annex Figure 17. ^{13}C NMR (APT, $CDCl_3$, 100 MHz) spectrum of compound **5a**.



Annex Figure 18. HSQC ($CDCl_3$, 400 MHz) spectrum of compound **5a**.



Annex Figure 19. HMBC ($CDCl_3$, 400 MHz) spectrum of compound **5a**



Annex Figure 20. COSY ($CDCl_3$, 400 MHz) spectrum of compound **5a**