UC San Diego
UC San Diego Previously Published Works

Title

Biosynthesis-Assisted Structural Elucidation of the Bartolosides, Chlorinated Aromatic Glycolipids from Cyanobacteria

Permalink

https://escholarship.org/uc/item/4x4377mk

Journal

Angewandte Chemie International Edition, 54(38)

ISSN

1433-7851

Authors

Leão, Pedro N
Nakamura, Hitomi
Costa, Margarida
et al.

Publication Date

2015-09-14

DOI

10.1002/anie.201503186

Peer reviewed

Biosynthesis-Assisted Structural Elucidation of the Bartolosides, Chlorinated Aromatic Glycolipids from Cyanobacteria

Dr. Pedro N. Leão,

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge MA, 02138 (USA). Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093 (USA). Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, Rua dos Bragas, 289, 4050-123 Porto (Portugal)

Hitomi Nakamura,

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge MA, 02138 (USA)

Margarida Costa,

Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, Rua dos Bragas, 289, 4050-123 Porto (Portugal)

Dr. Alban R. Pereira,

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093 (USA)

Prof. Rosário Martins,

Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, Rua dos Bragas, 289, 4050-123 Porto (Portugal). Centre of Health and Environmental Research—CISA, ESTSP, Polytechnic Institute of Porto, Rua Valente Perfeito 322, 4400-330 Vila Nova de Gaia (Portugal)

Prof. Vitor Vasconcelos,

Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, Rua dos Bragas, 289, 4050-123 Porto (Portugal). Department of Biology, Faculty of Sciences, University of Porto, Rua do Campo Alegre, 4169-007 Porto (Portugal)

Prof. William H. Gerwick, and

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093 (USA). Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093 (USA)

Prof. Emily P. Balskus

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge MA, 02138 (USA)

Correspondence to: Vitor Vasconcelos; William H. Gerwick; Emily P. Balskus.

Abstract

The isolation of the bartolosides, unprecedented cyanobacterial glycolipids featuring aliphatic chains with chlorine substituents and C-glycosyl moieties, is reported. Their chlorinated dialkylresorcinol (DAR) core presented a major structural-elucidation challenge. To overcome this, we discovered the bartoloside (brt) biosynthetic gene cluster and linked it to the natural products through in vitro characterization of the DAR-forming ketosynthase and aromatase. Bioinformatic analysis also revealed a novel potential halogenase. Knowledge of the bartoloside biosynthesis constrained the DAR core structure by defining key pathway intermediates, ultimately allowing us to determine the full structures of the bartolosides. This work illustrates the power of genomics to enable the use of biosynthetic information for structure elucidation.

Keywords

biosynthesis; cyanobacteria; dialkylresorcinol; glycolipids; structure elucidation

As ancient photoautotrophs, cyanobacteria have a differentiated ecophysiology among prokaryotes[1] that is abundantly reflected in their secondary metabolite diversity.[2] We have explored the chemical diversity of an in-house culture collection of cyanobacteria[3] and herein report the discovery of four unusual secondary metabolites and the use of biosynthesis in their structure elucidation.

Examination of a fractionated crude organic extract from the filamentous cyanobacterium *Nodosilinea* sp. LEGE 06102 led to the ^{1}H NMR spectroscopy guided isolation of bartoloside A (**1**, 2.0 mg; Figure 1a). Intriguingly, analogous metabolites, namely bartolosides B, C, and D (**2**, 5.6 mg; **3**, 0.7 mg; and **4**, 0.3 mg; Figure 1b), were isolated through bioassay-guided fractionation of a crude extract of *Synechocystis salina* LEGE 06155, a unicellular, free-living cyanobacterium found at the same beach as the bartoloside A producer.[4] These cyanobacterial strains are morphologically and phylogenetically distant (Figure 1; see also the Supporting Information, Figure S1). Together with their geographical co-occurrence, this observation may suggest pathway acquisition by horizontal gene transfer.

Initial structural-elucidation efforts focused on the most abundant metabolites (**1** and **2**). A combination of ^{1}H, ^{13}C, HSQC, HMBC and COSY data revealed a dialkylresorcinol (DAR) scaffold in both compounds (Supporting Information, S7–S39). D-Xylose and L-rhamnose were O-linked to the aromatic rings in **1** and **2**, respectively (Figure 2a, see also Figure S33). The DAR scaffold in **2** was further decorated with a xylose residue C-linked to the C4 position (Figure 2a). To the best of our knowledge, this is the first report of a cyanobacterial C-glycoside.

The two alkyl substituents on the resorcinol ring accounted for 25 and 27 carbon atoms in **1** and **2**, respectively. Considering the HR-ESI-MS derived molecular formulae ($C_{36}H_{62}Cl_2O_6$ for **1** and $C_{44}H_{76}Cl_2O_{10}$ for **2**) and NMR-based partial structures, these groups had to be linear. Overlapping ^{1}H and ^{13}C resonances for the two chlorinated methine groups and their corresponding ^2J and ^3J HMBC correlations to a total of eight methylene groups placed the Cl atoms on the alkyl chains (Figure 2a) and eliminated the possibility of their vicinal

positioning as observed in several naturally occurring marine lipids.[5] The termini of all partial structures showed HMBC and COSY correlations to the methylene envelope, hindering further assignments. Under ESI conditions, the molecules readily fragmented by loss of HCl and the sugar moieties. Alkyl chain fragmentations at higher energy did not allow for inference of the Cl atom or relative alkyl chain positions, but did confirm that C_{12}/C_{13} chains were present in **1** and C_{12}/C_{15} chains in **2**, **3**, and **4** (Supporting Information, S7–S51). Selective irradiation of the benzylic protons in 1D TOCSY experiments revealed that both alkyl chains were chlorinated in **2** (Figure 2b) as they were in **1** (Supporting Information, S7–S9). Owing to our inability to distinguish the two very similar alkyl groups, an assignment of the chain lengths and the determination of the precise Cl atom positions were not possible using NMR spectroscopy and MS alone.

To gain additional structural insights, we investigated bartoloside biosynthesis. The DAR scaffold is found in just a few cyanobacterial metabolites: the paracyclophanes,[6] microcarbonin A,[7] and nostocyclyne A (Figure S40).[8] In the biosynthesis of the paracyclophanes, DAR formation involves the type III polyketide synthase (PKS) catalyzed formation of a monoalkylresorcinol followed by an intermolecular alkylation reaction.[9] An analogous intramolecular cyclization is conceivable for the biosynthesis of nostocyclyne A (Figure S40). In other bacteria, a ketosynthase (DarB) and a putative aromatase (DarA) have been shown to synthesize DAR scaffolds in a convergent process.[10] DarB catalyzes the head-to-head condensation of a β-ketoacyl acyl carrier protein (ACP) thioester with an α,β-unsaturated acyl ACP thioester, resulting in a dialkylcyclohexanedione carboxylic acid (carboxy-CHD).[11] DarA is thought to be an aromatase based on its in vivo activity in a heterologous host,[11] but has not been characterized in vitro. Homologues of *darA* and *darB* are contiguous in many prokaryotic genomes, including several cyanobacteria.[11] Owing to the multiple cyanobacterial DAR biosynthetic pathways, we could not be sure which biosynthetic machinery would be involved in bartoloside production.

To identify the bartoloside biosynthetic pathway, we sequenced the genome of *S. salina* LEGE 06155 and searched for all of the known DAR biosynthetic pathways. We identified adjacent *darA* and *darB* homologues co-localized with additional genes encoding enzymes with predicted functions consistent with a role in bartoloside biosynthesis, including glycosyltransferases and a putative halogenase (Figure 3, Table S7). We designated this region to be the bartoloside (*brt*) gene cluster (GenBank: KR059027). In contrast to previously characterized *darAB*-containing gene clusters,[10–12] the *brt* cluster lacks a designated ACP, suggesting the use of a trans-acting ACP or ACP thioesters from fatty acid biosynthesis. The participation of acyl-CoA substrates is unlikely as cyanobacteria, including *S. salina* LEGE 06155, lack a canonical β-oxidation pathway.[13]

Based on the *brt* cluster architecture, we envisioned that the biosynthesis of **2** could proceed through the BrtD-catalyzed condensation of the 2-octadecenoyl-ACP thioester with the 3-oxo-hexadecanoyl-ACP thioester. Both precursors are derivatives of abundant (C_{18} and C_{16}) cyanobacterial fatty acids.[14] This reaction would yield carboxy-CHD **5**, which could spontaneously decarboxylate to give dione **6** or undergo BrtC-catalyzed oxidative aromatization to give the DAR backbone of **2–4** (**7**), which contains C_{12} and C_{15} chains at the C2 and C5 positions, respectively (Figure 3b). Alternatively, a DAR scaffold in which

the positions of these substituents are reversed could result from condensation of 2-pentadecenoyl-ACP with 3-oxo-nonadecanoyl-ACP (Figure S36). This scenario is highly unlikely given that C_{15} and C_{19} fatty acids are not produced by cyanobacteria.[14] Therefore, both the logic of bartoloside biosynthesis and the availability of biosynthetic precursors constrain the DAR scaffold of **2**–**4**. A similar analysis revealed that the assembly of **1** should utilize two C_{16} fatty acyl ACP derivatives, leading to a DAR with C_{12} and C_{13} substituents at the C2 and C5 positions.

After DAR formation, the *brt*-encoded glycosyltransferases should decorate the resorcinol ring with xylose and rhamnose. Chlorine installation presumably requires halogenation at unactivated carbon atoms, a transformation typically performed by α-ketoglutarate-dependent nonheme iron halogenases.[15] However, this type of enzyme was not found in the *S. salina* genome. Instead, we predict that BrtJ is the halogenase based on its homology to CylC[9a] and ColD/ColE[16] (41% identity, 55–59% similarity), enzymes encoded in the biosynthetic gene clusters of the halogenated cyanobacterial natural products cylindrocyclophane and columbamide.[16–17] BrtJ and its homologues do not resemble known halogenases and therefore likely represent a new enzyme class capable of halogenating unactivated carbon centers. Given their weak homology to the N-oxygenase AurF (BrtJ: 15% identity, 26% similarity, HHPred E-value:[18] 2.4×10^{-29}), these enzymes may be diiron halogenases,[19] and this possibility will be the subject of future investigations.

To confirm the involvement of the *brt* gene cluster in bartoloside biosynthesis, BrtD and BrtC were overexpressed in *E. coli* BL21 cells and purified. According to HR-MS analysis, BrtC bound flavin mononucleotide (FMN; Figure S37). *N*-Acetylcysteamine (SNAC) thioester analogues of the predicted BrtD substrates (**8** and **9**) were synthesized and incubated with BrtD; both **5** and **6** were detected in the assay mixture by LC-HR-ESI-MS, which is in agreement with our biosynthetic hypothesis (Figure 4). When BrtC was included in the assay, DAR **7** was observed (Figure 4). These findings provide the first biochemical evidence that DarA homologues are aromatases. Furthermore, the in vivo production of DAR-like metabolites by *E. coli* transformed with *brtCD* was observed (Supporting Information, S63, Figure S38). Overall, these observations confirmed the ability of BrtD and BrtC to utilize C_{16} and C_{18} fatty acyl thioesters for DAR synthesis and provided experimental verification of the biosynthetic hypothesis that enabled structure elucidation.

The only remaining unknown variable in the planar structure of **2** was the position of the Cl atoms. At this point, the acquisition of HSQC-TOCSY data for **2** established the connection of the benzylic moiety in the C_{15} chain to the chlorinated methine group by identification of four bridging methylene groups (Figure 5a–c; Supporting Information, S20–S39). The ω-7 position of the Cl atom in the C_{12} chain was similarly elucidated by the identification of a previously unassigned bridging methylene group in conjunction with a diagnostic HMBC correlation from the deshielded ω-6 methylene protons to the ω-3 carbon atom (Figure 5a, d, e; Supporting Information, S20–S39), thus finalizing the structure of **2**. Our proposed structure for **1** is based on NMR data highly similar to that of **2** (Supporting Information, S7–S19, Figure S2). The overall structures of **3** and **4** were deduced by comparison of their NMR and HRMS data with those of **2** (Supporting Information, S40–S51). A future

challenge lies in establishing the absolute configuration of these glycolipids, which were not amenable to crystallization.

The precise positioning of the Cl atoms in the bartolosides, cylindrocyclophanes, and columbamides suggests that halogenation by BrtJ homologues is highly specific. Aside from the occurrence of this halogenation in cyanobacterial DARs of distinct biosynthetic origin, the similarity of the 5-alkynyl moieties in microcarbonin A and nostocyclyne A is noteworthy (Figure S40). Together, these observations illustrate the functional convergence of two distinct biosynthetic logics and could suggest a highly selected role for DARs in cyanobacteria. Recognizing that microcarbonin A was isolated on the basis of a strong photosynthesis-inhibiting activity,[7] and that DevBCA export machinery is encoded in the *brt* cluster,[2] we envisioned that the bartolosides could act as allelochemicals.[2] Compound **2** was indeed a major organic component in *S. salina* LEGE 06155 culture supernatants (Figure S39). However, we did not observe strong bioactivity in anticyanobacterial, antialgal, or cytotoxicity assays (Supporting Information, S54–S55). Furthermore, although DARs are implicated in quorum-sensing,[20] we were unable to locate homologues of the required components of the signaling system PauR or LuxR in the genome of *S. salina* LEGE 06155.

In conclusion, the bartolosides are secondary metabolites produced by cyanobacterial genera from which few (*Synechocystis*)[21] or no (*Nodosilinea*) natural products are known. The chlorinated alkyl substituents of these molecules presented a substantial challenge for structure elucidation. By investigating the biosynthesis of the bartolosides prior to completing the structural assignment, we obtained crucial information that ultimately enabled us to solve the complete planar structures and majority of the stereocenters. Given the increasing accessibility of genomic data, we envision that this general strategy will be a powerful adjunct for structure elucidation. Efforts to study the unusual biosynthetic events in this pathway and the natural roles of the bartolosides are underway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Oren, A. Cyanobacteria: An Economic Perspective. Sharma, NK.Rai, AK., Stal, LJ., editors. Wiley; Hoboken: 2014. p. 1-20.

2. Leão PN, Engene N, Antunes A, Gerwick WH, Vasconcelos V. Nat Prod Rep. 2012; 29:372–391. [PubMed: 22237837]

3. a) Leão PN, Ramos V, Goncalves PB, Viana F, Lage OM, Gerwick WH, Vasconcelos VM. Mar Drugs. 2013; 11:1316–1335. [PubMed: 23609580] b) Leão PN, Costa M, Ramos V, Pereira AR, Fernandes VC, Domingues VF, Gerwick WH, Vasconcelos VM, Martins R. PloS ONE. 2013; 8:e69562. [PubMed: 23922738]

4. Brito A, Ramos V, Seabra R, Santos A, Santos CL, Lopo M, Ferreira S, Martins A, Mota R, Frazão B, Martins R, Vasconcelos V. Syst Appl Microbiol. 2012; 35:110–119. [PubMed: 22277323]

5. Dembitsky VM, Srebnik M. Prog Lipid Res. 2002; 41:315–367. [PubMed: 11958814]

6. Moore BS, Chen JL, Patterson GML, Moore RE, Brinen LS, Kato Y, Clardy J. J Am Chem Soc. 1990; 112:4061–4063.

7. Beresovsky D, Hadas O, Livne A, Sukenik A, Kaplan A, Carmeli S. Isr J Chem. 2006; 46:79–87.

8. Ploutno A, Carmeli S. J Nat Prod. 2000; 63:1524–1526. [PubMed: 11087597]

9. a) Nakamura H, Hamer HA, Sirasani G, Balskus EP. J Am Chem Soc. 2012; 134:18518–18521. [PubMed: 23106426] b) Nakamura H, Balskus EP. Synlett. 2013:1464–1470.

10. Nowak-Thompson B, Hammer PE, Hill DS, Stafford J, Torkewitz N, Gaffney TD, Lam ST, Molnár I, Ligon JM. J Bacteriol. 2003; 185:860–869. [PubMed: 12533461]

11. Fuchs SW, Bozhüyük KAJ, Kresovic D, Grundmann F, Dill V, Brachmann AO, Waterfield NR, Bode HB. Angew Chem Int Ed. 2013; 52:4108–4112.Angew Chem. 2013; 125:4202–4206.

12. Schöner TA, Fuchs SW, Reinhold-Hurek B, Bode HB. PLOS ONE. 2014; 9:e90922. [PubMed: 24618669]

13. von Berlepsch S, Kunz H-H, Brodesser S, Fink P, Marin K, Flügge U-I, Gierth M. Plant Physiol. 2012; 159:606–617. [PubMed: 22535424]

14. a) Liu XY, Sheng J, Curtiss R. Proc Natl Acad Sci USA. 2011; 108:6899–6904. [PubMed: 21482809] b) Kenyon CN. J Bacteriol. 1972; 109:827–834. [PubMed: 4621688]

15. Vaillancourt FH, Yeh E, Vosburg DA, Garneau-Tsodikova S, Walsh CT. Chem Rev. 2006; 106:3364–3378. [PubMed: 16895332]

16. Kleigrewe K, Almaliti J, Tian IY, Kinnel RB, Korobeynikov A, Monroe EA, Duggan BM, Di Marzo V, Sherman DH, Dorrestein PC, Gerwick L, Gerwick WH. J Nat Prod. 2015; 78:1671–1682. [PubMed: 26149623]

17. Chlipala GE, Sturdy M, Krunic A, Lantvit DD, Shen Q, Porter K, Swanson SM, Orjala J. J Nat Prod. 2010; 73:1529–1537. [PubMed: 20825206]

18. Söding J. Bioinformatics. 2005; 21:951–960. [PubMed: 15531603]

19. Li N, Korboukh VK, Krebs C, Bollinger JM. Proc Natl Acad Sci USA. 2010; 107:15722–15727. [PubMed: 20798054]

20. Brameyer S, Kresovic D, Bode HB, Heermann R. Proc Natl Acad Sci USA. 2015; 112:572–577. [PubMed: 25550519]

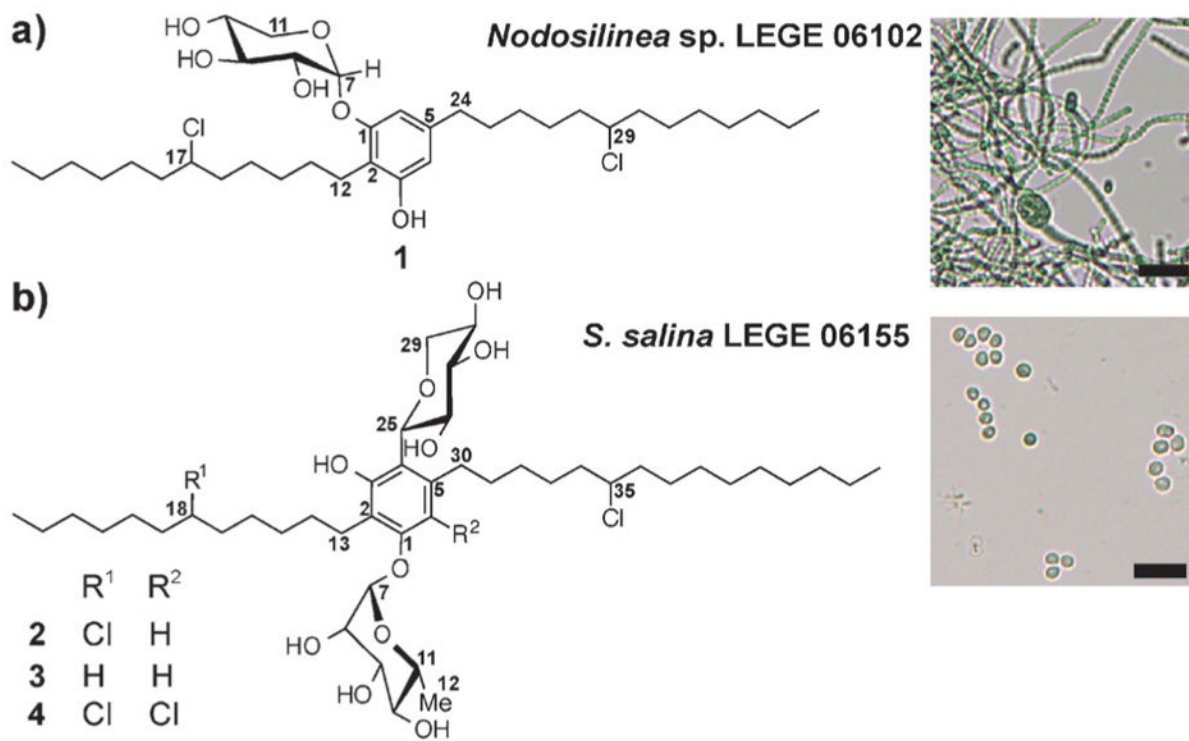21. Nagle DG, Gerwick WH. Tetrahedron Lett. 1995; 36:849–852.

Figure 1.
The bartolosides are chlorinated aromatic glycolipids found in cyanobacteria belonging to different genera. Optical photomicrographs of the producing organisms and structures of: a) bartoloside A (**1**) and b) bartolosides B, C, and D (**2–4**). Scale bar: 10 μm.
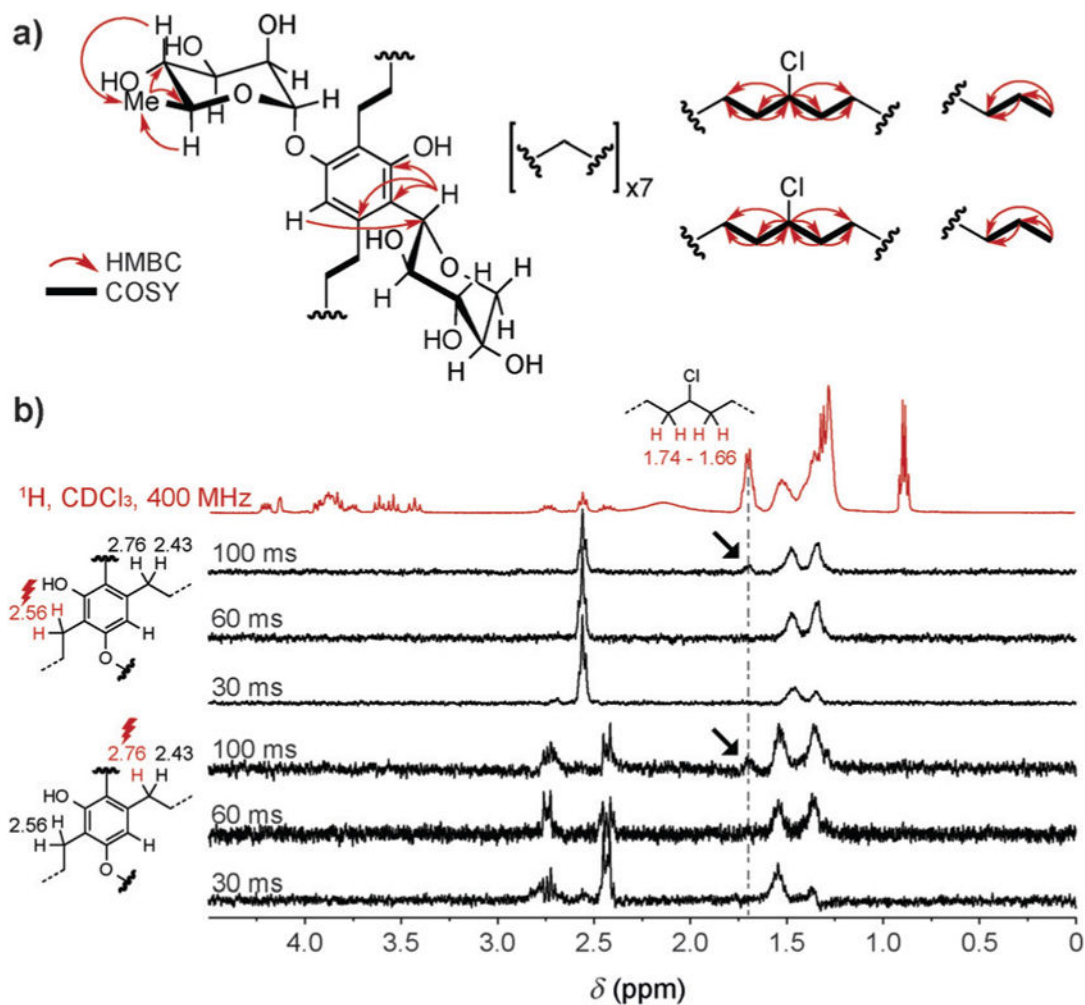
Figure 2.
NMR-derived structural information for **2**. a) Fragments obtained from 1D and 2D NMR spectroscopy, with key HMBC correlations highlighted. b) 1D TOCSY experiments with irradiation of the benzylic protons on each chain reveal propagation up to methylene groups vicinal to the chlorinated methine moieties.
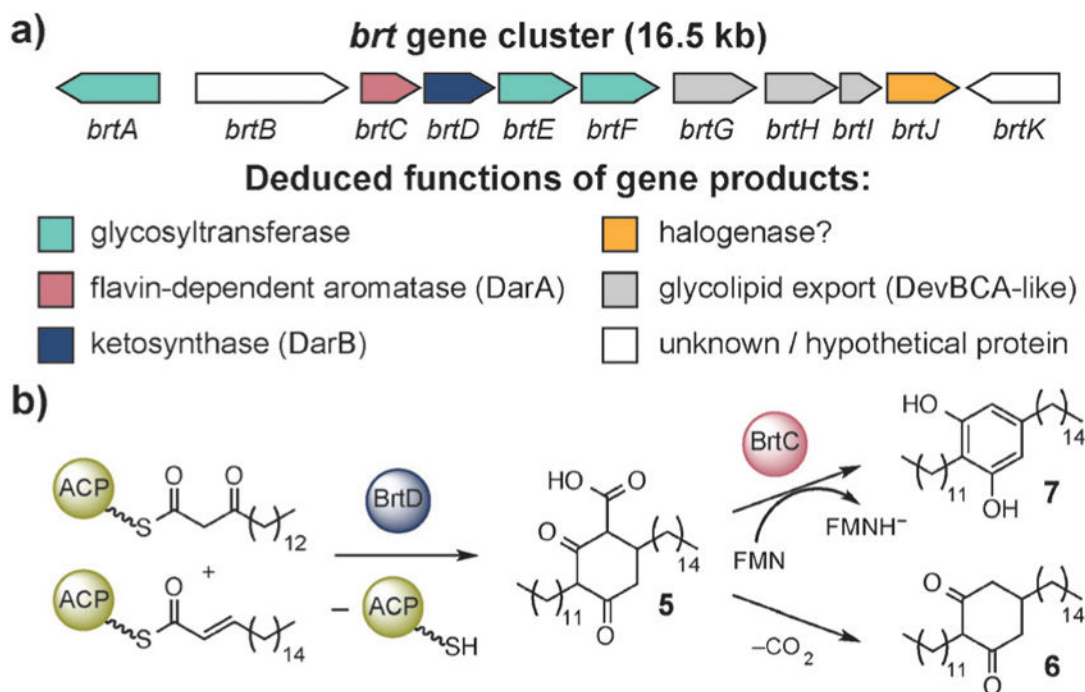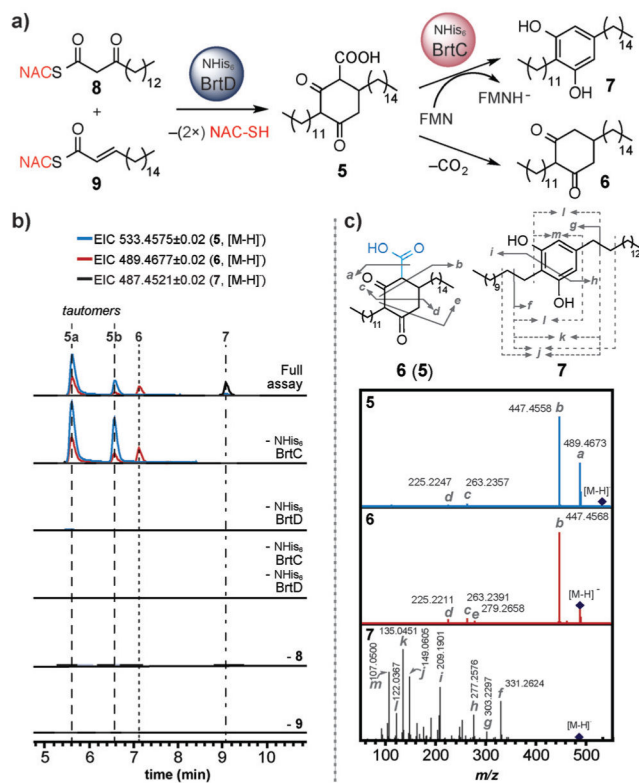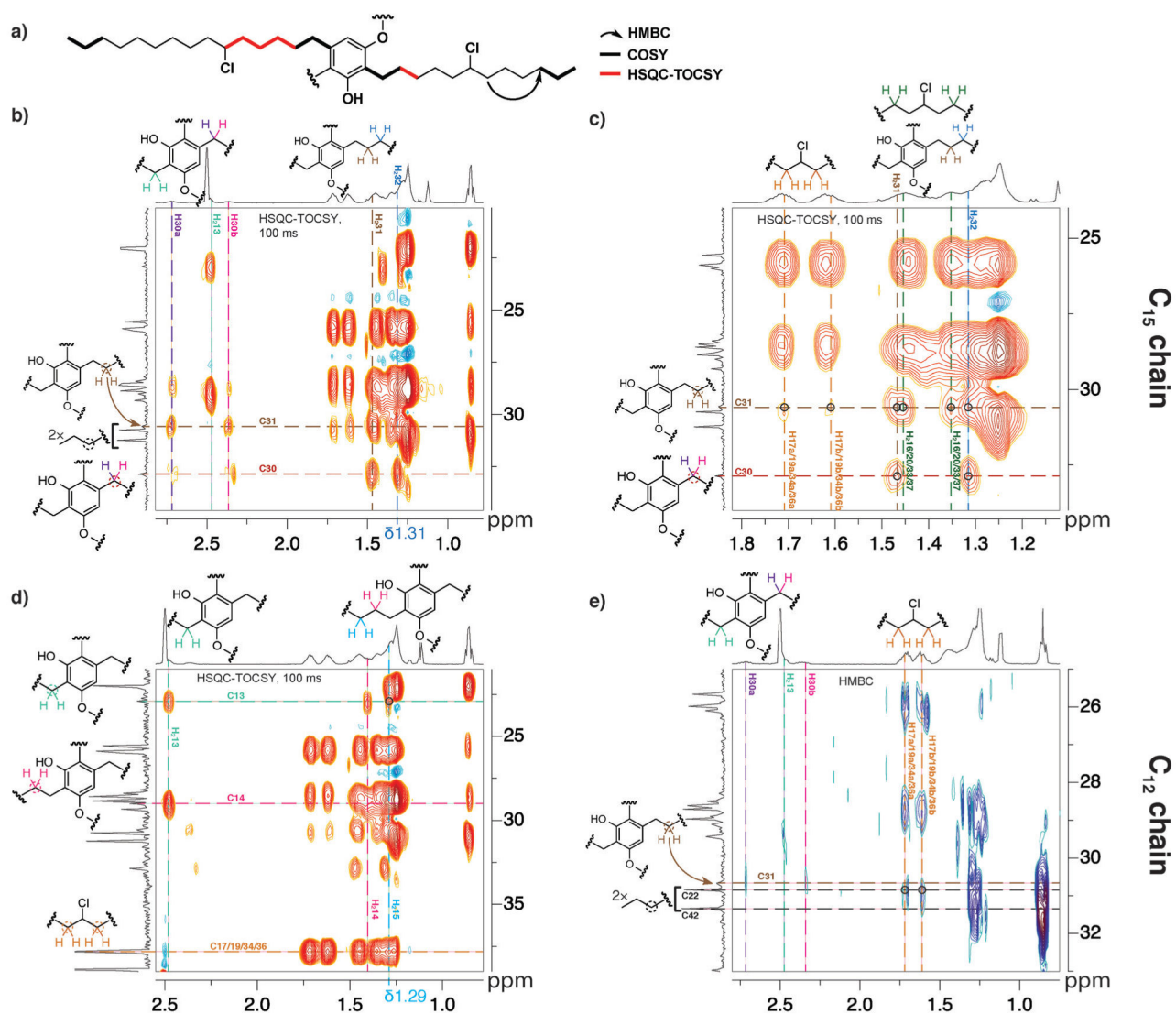
Figure 3.
Biosynthesis of the bartolosides. a) Organization of the *brt* gene cluster from *S. salina* LEGE 06155 and annotated gene functions. b) Proposal for the biosynthesis of the dialkylresorcinol scaffold found in **2**–**4**.

Figure 4.
BrtD and BrtC synthesize the DAR core of bartoloside B. a) Schematic representation of the N-His$_6$-tagged BrtD and BrtC coupled assay. b) LC-HR-ESI-MS derived extracted ion chromatograms (EICs) showing diketone production by N-His$_6$-BrtD and DAR production by N-His$_6$-BrtC. All chromatograms have the same scale. c) HR-MS^2 characterization of the reaction products **5**, **6**, and **7** (dashed lines represent ions generated by a fragmentation on an alkyl chain).

Figure 5.
NMR-based (HSQC-TOCSY, 600 MHz and HMBC, 400 MHz) deduction of the position of the Cl atoms in each chain in **2**. a) Proposed positioning and key HMBC and HSQC-TOCSY correlations leading to the structural assignment. b) HSQC-TOCSY spectrum of **2**, highlighting the spectral deconvolution leading to the assignment of the methylene group resonating at δ=1.31 ppm in the C_{15} chain. c) Key HSQC-TOCSY correlations (circles) connecting the benzylic position to the chlorinated position in the C_{15} chain. d) Assignment of the bridging methylene group in the C_{12} chain by HSQC-TOCSY analysis (key correlation circled). e) HMBC data highlighting the correlation (circled) from a pair of diastereotopic protons (vicinal to the chlorinated position) to the ω-3 methylene carbon atom.