Chemistry, bioactivity and biosynthesis of cyanobacterial alkylresorcinols

Teresa P. Martins, a Caroline Rouger, †b Nathaniel R. Glasser, c Sara Freitas, a Nelly B. de Fraissinette, a Emily P. Balskus, c Deniz Tasdemir b and Pedro N. Leão a,* a

Covering: up to 2019

Alkylresorcinols are amphiphilic metabolites, well-known for their diverse biological activities, produced by both prokaryotes and eukaryotes. A few classes of alkylresorcinol scaffolds have been reported from the photoautotrophic cyanobacteria, ranging from the relatively simple hierridins to the more intricate cylindrocyclophanes. Recently, it has emerged that cyanobacteria employ two different biosynthetic pathways to produce unique alkylresorcinol scaffolds. However, these convergent pathways intersect by sharing biosynthetic elements which lead to common structural motifs. To obtain a broader view of the biochemical diversity of these compounds in cyanobacteria, we comprehensively cover the isolation, structure, biological activity and biosynthesis of their mono- and dialkylresorcinols. Moreover, we provide an overview of the diversity and distribution of alkylresorcinol-generating biosynthetic gene clusters in this phylum and highlight opportunities for discovery of novel alkylresorcinol scaffolds. Because some of these molecules have inspired notable syntheses, different approaches used to build these molecules in the laboratory are showcased.

1. Introduction

Alkylresorcinols represent a large structural group of amphiphilic phenolic lipids. These scaffolds are widespread in Nature—alkylresorcinols have been isolated from higher plants, insects, fungi, algae, slime molds, mosses and bacteria.1–3 They play structural roles in the membranes of the producing organisms, and have been associated with many biological activities including antioxidant, antibacterial, cytotoxic, dermatotoxic, genotoxic, and signaling properties.1–3 The chemistry, structural diversity and associated biological activity of alkylresorcinols, mostly from the plant world, has been thoroughly reviewed by Kozubek and co-workers.1,2 It has recently emerged that alkylresorcinols are generated by different pathways that sometimes intersect to create further
chemical diversity. The type III polyketide synthase (PKS), which elongate and cyclize different fatty-acyl precursors give rise to monoalkylresorcinols (MARs) in plants and bacteria and have been studied for the past few decades.\textsuperscript{4,5} However, the biosynthesis of dialkylresorcinols (DARs, so far only encountered in bacteria) was not fully understood until the recent report of two different DAR-generating pathways. In 2013, the Balskus lab showed that the unusual paracyclophane DARs derive from dimerization of type III PKS-derived MAR units.\textsuperscript{6} In 2014, the Bode group revealed that a new type of ketosynthase generated 2,5-DARs from two different fatty acid primary metabolism-derived precursors.\textsuperscript{7}

Cyanobacterial alkylresorcinols, in particular DARs, are structurally distinct from those of heterotrophic bacteria, in that they commonly feature halogenation, glycosylation, long alkyl chains and intra- or inter-molecular cyclization. Our current understanding of alkylresorcinol diversity in cyanobacteria is limited, as many orphan gene clusters that lead to MARs and DARs are found in their genomes.\textsuperscript{8} Despite this small number of cyanobacterial alkylresorcinols, the structural connections among them are evident and reflect their intertwined biosynthetic pathways. This realization has been the main motivation for the present review. We begin by covering the structural diversity of the cyanobacterial alkylresorcinols and their biological activities, before detailing the biosynthetic pathways that generate such structures. We then peer into the genomes of cyanobacteria and look into the distribution of alkylresorcinol-related biosynthetic pathways, including those without any associated natural product. Since these structures have inspired a number of total syntheses that contrast with their natural biosynthesis, some of the strategies that have been used to obtain cyanobacterial alkylresorcinol scaffolds in the laboratory are briefly discussed.

2. Structural variety and biological activity of alkylresorcinols in cyanobacteria

2.1. Monoalkylresorcinols (MARs)

MARs have been reported from a variety of organisms, mainly from plants, bacteria and fungi, but also from protists and...
bryozoans.1–3 Biological activities associated with these metabolites include antitumor, anti-fungal, anti-inflammatory, antibacterial, antiparasitic, antioxidant, and genotoxic, among others.1,2 In higher plants, MARs have been obtained, for example, from rye and wheat bran oils, mango peels as well as from leaves and seeds used in ancient traditional medicines.1,2 Structural variations in plant MARs are usually related to the alkyl substituent; for example, DNA-polymerase inhibiting bis-5-alkylresorcinols isolated from the angiosperm Panapips rubescens8 bear two terminal resorcinol rings on each side of the alkyl chain. Some additional structural features are noteworthy in non-plant MARs: for example, monochasiosols isolated from the slime mold Dictyostelium monochasioides contain a chlorinated phenolic ring, while adeninealkylresorcinol, isolated from the fungus Lasiodiplodia sp.,9 features a nucleobase moiety. O-Methylation is also common in MARs.1

In heterotrophic bacteria, MARs have been reported from Actinobacteria and alpha-, delta- and gamma-Proteobacteria11–13 which also include sulfated, methylated, and glycosylated resorcinol variants (Fig. 1). Interestingly, Azotobacter MARs were found to be the main components of its cyst membranes14,15 and are essential for the encystment process.16 In Streptomyces sp., MARs were found to be associated with the cytoplasmic membrane and to confer β-lactam antibiotic resistance.17

Despite the large number of type III PKS-containing biosynthetic gene clusters (BGCs) found in cyanobacterial genomes (see Section 3), only four cyanobacterial MARs (hierridins A–C18,19 and cylindrofedin A,20 4–7) are known to date.

2.1.1. Hierridins. Hierridins are a group of monoalkylresorcinols isolated in 1992 by González et al.23 for the first time from a lichen, Ramalina hierrensis, that is endemic to the Canary Islands. Hierridin (8) was isolated as a solid compound and the arrangement of the substituents was clarified via semi-

Emily Balskus received her BA from Williams College and her PhD in synthetic organic chemistry from Harvard University, where she worked with Eric Jacobsen. Her postdoctoral studies at Harvard Medical School with Christopher T. Walsh focused on deciphering the biosynthesis of cyanobacterial sunscreens. She began her independent career at Harvard University in 2011 and is now a Professor of Chemistry and Chemical Biology. Her research program explores the intersection of organic chemistry, biochemistry, and microbiology, with a major focus on understanding unusual natural product biosynthetic pathways and metabolic activities of the human microbiota.

Pedro Leão obtained his PhD from the University of Porto in 2010, where he studied allelopathy in cyanobacteria under the supervision of Prof. Vitor Vasconcelos. For his postdoctoral studies at CIIMAR (U. Porto), Pedro focused on marine natural products chemistry and biosynthesis with research stays at the Gerwick lab (Scripps Institution of Oceanography) and Balskus group (Harvard University). In 2015 he became Auxiliary Researcher at CIIMAR and in 2018 he started the Cyanobacterial Natural Products group in the same institution. His research interests lie in the chemistry, biosynthesis and chemical ecology of cyanobacterial secondary metabolites.
showing selective albeit very mild cytotoxicity to a colon adenocarcinoma cell line HT-29, with an IC$_{50}$ value of $\sim$100 µM. This selective bioactivity was further investigated by Freitas $et$ $al.$ Compound 5 seems to interfere with VDAC-1 (voltage-dependent anion-selective channel 1) in mitochondria as well as some cell cycle and structural genes responsible for the cell cycle checkpoints, leading to apoptosis in the HT-29 cells.

During the re-isolation of 5, a halogenated analogue, hierridin C (6) was discovered. The aromatic ring of this metabolite has a chlorine atom adjacent to the alkyl-substituted carbon, confirmed through NMR and MS experiments. Compound 6 showed antimalarial activity against Plasmodium falciparum strains (3D7 and Dd2) with IC$_{50}$ values of 1.5 µM and 2.3 µM, respectively. This is the first bacterial example of a MAR featuring ring halogenation. Curiously, the structurally-related monochasiosins possessing a chlorination in position 2 of the resorcinol ring were recently reported from a Dictyostelium slime mold$^{28}$ and are related to Dictyostelium differentiation-inducing (DIF) polyketides. Similar to hierridin B, the DIF-family compound DIF-3 targets mitochondria in cancer cells.$^{30}$

### 2.1.2. Cylindrofridin A.

The metabolite cylindrofridin A (7) was discovered by Preisitsch $et$ $al.$ and corresponds to a monomeric unit that likely acts as an intermediate in the biosynthesis of the cylindrocyclophanes (see Sections 2.2.1. and 3.2.1.). Compound 7 features halogenation and methylation in the $\omega$-5 and $\omega$-10 positions of its alkyl chain, respectively, and is also acetylated in its benzylic position. The compound was isolated in a screening study of dozens of cyanobacterial strains, by means of an extraction and detection strategy aimed at identifying paracyclophane-related molecules. Metabolite 7 showed mild antibacterial activity towards Gram-positive bacteria, including MRSA.$^{30}$ Interestingly, while investigating the biosynthesis of the cylindrocyclophanes, Nakamura $et$ $al.$ observed potent toxicity effects of a synthetic analogue of 7 towards a cylindrocyclophane-producing strain of Cylindrospermum licheniforme.

#### 2.2. Dialkylresorcinols

DARs have been known for a few decades, with flexirubin (10)$^{31}$ and 2-n-hexyl-5-n-propylresorcinol (also referred as DB-2073, 11)$^{32}$ being the first described compounds of this family. While flexirubin-type pigments have so far only been reported from Flexibacter, Flavobacteria and Cytophaga strains,$^{33}$ the description of the potent antibiotic DB-2073 marked the discovery of a series of DARs from Pseudomonas spp. Examples include resorstatin (2-hexyl-5-pentylresorcinol, 12), showing a strong inhibitory activity against lipid peroxidation induced by free radicals in rat brain homogenate (IC$_{50}$ = 2.06 µM)$^{34}$ and other short-chain DARs such as 2-butyl-5-propylresorcinol (13) and 2-hexyl-5-methylresorcinol (14) that show antibacterial or antifungal activity.$^{35}$ More recently, arcuflavin (15), a xanthomonadin-DAR hybrid that bears resemblance to the flexirubin pigments was reported from an Azoarcus sp. $\beta$-proteobacterium.$^{36}$ The structures, biosynthesis and biological role of bacterial DARs (mostly from heterotrophic bacteria) have been reviewed by Schöner $et$ $al.$$^{37}$
The first cyanobacterial DARs were discovered in the early 1990s when Richard Moore’s lab described the highly unusual structures of a new class of naturally-occurring cyclic DARs: a series of cyclophanes isolated from Nostocales strains.19,20 Since then, a number of DARs have been reported from cyanobacteria, most of which differ structurally from their non-cyanobacterial counterparts and have a variety of biological activities. Many structural features are shared by different classes of cyanobacterial DARs, including chlorination, alkyne moieties and glycosylations, which reflects common biosynthetic elements (see Section 3.2) but also suggests that these compounds share biological roles.

2.2.1. \([\eta,\eta]\text{Cyclophanes and cylindrofridins.}\) The cyclophane nomenclature was suggested by Cram & Steinberg66a following their effort to synthesize molecules with unique complementary three-dimensional structures, considered as

![Image of cyclophane structures]

### Table 1
Biological activity of cyanobacterial alkylresorcinols

<table>
<thead>
<tr>
<th>Alkylresorcinol class</th>
<th>Congener(s)</th>
<th>Reported biological activity</th>
<th>Potency*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(in µM, unless otherwise stated)</td>
</tr>
<tr>
<td><strong>MARS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hierridins</td>
<td>A (4) and B (5), mixture</td>
<td>Antiplasmodial (Plasmodium falciparum, D6 and W2 strains)18</td>
<td>3.7–5.2 µg mL⁻¹ (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>B (5)</td>
<td>Cytotoxic (HT-29 cells)26</td>
<td>~100 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>B (5) and C (6)</td>
<td>Antiplasmodial (P. falciparum, 3D7 and Dd2 strains)19</td>
<td>1.5–2.3 (IC₅₀)</td>
</tr>
<tr>
<td>Cylindrofridins</td>
<td>A (7)</td>
<td>Antibacterial (Gram-positive: MRSA, Streptococcus pneumoniae)20</td>
<td>8.6–17 (MIC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cytotoxic (HaCaT cells)20</td>
<td>~100 (IC₅₀)</td>
</tr>
<tr>
<td><strong>DARS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartolides</td>
<td>A (58)</td>
<td>Cytotoxic (HT-29 cells)21</td>
<td>21 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A (58), E (62) and I (66)</td>
<td>Cytotoxic (T-47D, MO-63, RKO cells)52</td>
<td>22–60 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>B (59)</td>
<td>Cytotoxic (T-47D, MG-63, HepG2, SH-SYSY, PC-3, HT-29, RKO cells)51</td>
<td>9.5–26 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A-C (35–37)</td>
<td>Antibacterial (Staphylococcus aureus)44</td>
<td>40–100 (MIC)</td>
</tr>
<tr>
<td></td>
<td>A-C (35–37), F (40) and G (41)</td>
<td>Cytotoxic (FI and MCF7 cells)44</td>
<td>0.86–5.1 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A (35), B (36) and F (40)</td>
<td>Antibacterial (S. pneumoniae)45</td>
<td>5.6–8.2 (MIC)</td>
</tr>
<tr>
<td></td>
<td>F (40) and G (41)</td>
<td>Antibacterial (Enterococcus faecalis, S. aureus)45</td>
<td>0.1–1.1 (MIC)</td>
</tr>
<tr>
<td></td>
<td>A-F (35–40) and H-L (42–46)</td>
<td>Antifungal (Candida albicans)45</td>
<td>1.3–5.5 (MIC)</td>
</tr>
<tr>
<td></td>
<td>A-F (16, 21–25)</td>
<td>Antimycobacterial (Mycobacterium tuberculosis)45</td>
<td>0.8–1.8 (MIC)</td>
</tr>
<tr>
<td></td>
<td>A-F (16, 22, 24–26)</td>
<td>20S proteasome inhibition42</td>
<td>2.55–59.3 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>D (30), C (32) and F (34)</td>
<td>Cytotoxic (HT-29 cells)52</td>
<td>0.5–2.8 (EC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A₁₁ A₁, C₁, C₂, C₁₁ and F₁ (32)</td>
<td>Cytotoxic (MDA-MB-435 and MDA-MB-231 cells)53</td>
<td>2.3–5.2 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>F (25)</td>
<td>Cytotoxic (MDA-MB-435, MDA-MB-231 and OVCAR3 cells)53</td>
<td>10–13 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>C (30), C₁₁ and C₂ (32)</td>
<td>Cytotoxic (HT-29 cells)52</td>
<td>1.7–3.3 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A (47) and B (48)</td>
<td>Cytotoxic (MDA-MB-435, MDA-MB-231 and OVCAR3 cells)53</td>
<td>0.9–9.8 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A (47), C (49) and D (50)</td>
<td>Cytotoxic (KB and LoVo cells)53</td>
<td>0.5–2 µg mL⁻¹ (IC₅₀)</td>
</tr>
<tr>
<td>Nostocyclophane</td>
<td>A-D (18–20, 17)</td>
<td>Cytotoxic (MDA-MB-435, MDA-MB-231 and OVCAR3 cells)53</td>
<td>10–12.5 µg/disk (MIC)</td>
</tr>
<tr>
<td></td>
<td>A (70)</td>
<td>Antibacterial (S. aureus and Bacillus subtilis)54</td>
<td>0.6–6.4 (IC₅₀)</td>
</tr>
<tr>
<td></td>
<td>A-D (51–54)</td>
<td>Cytotoxic (MDA-MB-435 and MDA-MB-231 cells)53</td>
<td>0.6–6.4 (IC₅₀)</td>
</tr>
</tbody>
</table>

*As reported in the cited references. **Compounds 35 and 36 only. ***From MABA (microplate alamar blue assay). ****From LORA (low-oxygen recovery assay). Paper value corresponds to compound 40, compound 41 was reported to have a MIC > 10 µM.
interesting vehicles for host–guest chemistry. Many and diverse structural motifs that are found within the natural product space fall into the cyclophane category sensu lato.\(^{41}\) The cyclophanes found in cyanobacteria belong to the [7.7]paracyclophanes, in which two \(p\)-substituted aromatic units are bridged by seven carbon atoms. More precisely, these cyanobacterial metabolites share a unique scaffold consisting of two resorcinol moieties connected via two aliphatic chains and are divided into different structural subclasses, namely the \(p\)-cyclophanes,\(^{38,39,42}\) nostocyclophanes,\(^{38,43}\) carbamidocyclophanes,\(^{44–46}\) merocyclophanes\(^{47,48}\) and ribocyclophanes\(^{49}\) (Fig. 2). The cylindrocyclophanes and nostocyclophanes were discovered on the basis of antitumor activity showed by two Nostocales cyanobacterial extracts.\(^{38}\) The bioassay-guided isolation led to the structure elucidations of cylindrocyclophane \(A\) \((16)\) and halogenated nostocyclophane \(D\) \((17)\), from \(Cylindro\) sp. \(\text{lichenforme}\) ATCC 29204 and \(Nostoc\) \(\text{linckia}\) UTEX B1932, respectively. The absolute stereochemistry of the halogenated metabolite \(17\) was determined using X-ray crystallography.\(^{38}\) In the following years, the same team described nostocyclophanes \(\text{A–C}\) \((18–20)\),\(^{43}\) and cylindrocyclophanes \(\text{B–F}\) \((21–25)\),\(^{39}\) which were also discovered based on their cytotoxicity towards human cancer cell lines. The mid-chain chlorinations are the hallmark of the nostocyclophanes, while all the cylindrocyclophanes feature a methyl substituent adjacent to the benzylic positions. The cylindrocyclophane class was later expanded following the \(20S\) proteasome inhibition activity-guided discovery of congeners \(\text{A}_{4}–\text{A}_{1}\) \((26–29)\), \(\text{C}_{4}–\text{C}_{1}\) \((30–33)\) and \(\text{F}_{4}\) \((34)\) from \(Nostoc\) sp. UIC 10022A that feature chlorination in the terminal alkyl chain positions.\(^{42}\)

The carbamidocyclophanes are characterized by carbamate substituents at both benzylic positions. Carbamidocyclophanes \(\text{A–E}\) \((35–39)\) were isolated from a Vietnamese \(Nostoc\) sp. strain, guided by antibacterial activity towards Gram-positive bacteria as well as antifungal \((\text{Candida maltosa})\) activity.\(^{44}\) These metabolites showed mild potency towards \(Staphylococcus aureus\) ATCC 6538 but had stronger, low micromolar cytotoxicity towards \(\text{MCF7}\) breast adenocarcinoma cells indicating that cyanobacterial paracyclophanes exhibit generalized cytotoxicity.\(^{44}\) The congeners \(\text{F}\) \((40)\) and \(\text{G}\) \((41)\) are similar to 35 but

![Fig. 2 Cyanobacterial [7.7]paracyclophanes (16–47, 49–54) and related metabolites (48, 55–57).](image-url)
feature hydroxyl and acetyl substituents, respectively, in one of the benzylic positions. Both metabolites showed sub-micromolar cytotoxicity towards human cancer cell lines.\textsuperscript{43} Notably, compound 40 exhibited potent anti-\textit{Mycobacterium} activity and was also reported to strongly inhibit the growth of

\textit{S. aureus} and \textit{Enterococcus faecalis}.\textsuperscript{45} Five new carbamidocyclophanes (42-46) with minor structural differences to known congeners were later reported by the same research group,\textsuperscript{46} and exhibited submicromolar antibacterial (Gram-positive) activity and low-micromolar cytotoxicity in a human keratinocyte cellular assay.

Merocyclophanes have a benzylic methyl substituent and lack the beta methyl group found in the cylindrocyclophanes. Merocyclophane A (47) was isolated from a terrestrial \textit{Nostoc} sp. strain (UIC 10062) together with the hydroxyquinone-containing analogue merocyclophane B (48).\textsuperscript{47} Merocyclophanes C (49) and D (50) were later isolated from the freshwater \textit{Nostoc} sp. UIC 10110 and are analogues of 47 with hydroxylation at one or both terminal alkyl positions.\textsuperscript{48} The merocyclophanes, as with most cyanobacterial paracyclophanes, are strongly cytotoxic (low-to-sub micromolar IC\textsubscript{50}) towards human cancer cell lines.\textsuperscript{49}

Recently, the Orjala group reported members of a new class of paracyclophanes, namely ribocyclophanes A–D (51–54), that contain \(\beta\)-\(\gamma\)-ribopyranose as substituent at one or both benzylic positions.\textsuperscript{50} These sugar moieties were found to be in the rare \(\text{C}_4\) conformation; the authors suggest that this is due to steric hindrance from the cylindrocyclophane core not allowing the more common (and typically more stable) \(\text{C}_1\) conformation. The ribocyclophanes were co-isolated with cylindrocyclophane D (23) suggesting a common biosynthetic origin.\textsuperscript{51}

The isolation of pharacine, a highly symmetrical paracyclophane (a cyclic terephthalate), from the gliding marine Gram-negative bacterial strain \textit{Cytophaga} sp. (AM13.1), suggested that the occurrence of [n.n]cyclophanes is not restricted to the phylum Cyanobacteria.\textsuperscript{20} Still, the presence of traces of this molecule in ubiquitous phthalate plasticizers has raised doubts as to the natural origin of the compound.\textsuperscript{20}

In 2015, Preisitsch and collaborators\textsuperscript{20} isolated two new cylindrocylophane-related DARs, cylindrofridins B (53) and C (56). These compounds are dimers of the co-isolated MAR 7 or its putative deacetylated monomer (that has not been isolated to date) and are reminiscent of the Friedel–Crafts alkylation events that are involved in cyanobacterial paracyclophane biosynthesis\textsuperscript{a} (see Section 3.2.1.). This is evident from the chlorination at the \(\omega-5\) carbon atom in cylindrofridins, the site of \(\text{[7.7]}\)paracyclophane macrocycle formation, and therefore compounds 55 and 56 represent linearized biosynthetic intermediates of the cylindrocyclophanes.\textsuperscript{20} Likewise, the recently described ribocyclophane E (57), isolated from the ribocyclophane-producing strain \textit{Nostoc} sp. UIC 10366,\textsuperscript{49} is a linearized version of compound 54. No significant bioactivity has been reported for compounds 55–57, indicating that the open versions of the cyclophane macrocycles do not conserve their cytotoxic and antibacterial properties.\textsuperscript{20,49}

2.2.2. Bartolosides. While screening cultured cyanobacteria obtained from the Portuguese coast for novel natural products, Leão \textit{et al.}\textsuperscript{41} discovered an unprecedented group of DAR glycolipids, named bartolosides A–D (58–61). These compounds were isolated from the phylogenetically distant cyanobacterial genera \textit{Nodosilinea} and \textit{Synecochystis}. Examination of a fractionated crude organic extract of the filamentous cyanobacterium \textit{Nodosilinea nodulosa} LEGE 06102 led to the \(\text{H}\) NMR spectroscopy-guided isolation of compound 58. Subsequently, an independent bioassay-guided fractionation of the crude extract of \textit{Synecochystis salina} LEGE 06155, a unicellular,
free-living cyanobacterium found at the same beach as the producer of bartoloside A (58), led to the isolation of analogous metabolites, namely bartolosides B, C and D (59, 60 and 61). The authors hypothesized that, taking into account the phylogenetic distance between the two producing strains, coupled with their geographical co-occurrence, horizontal gene transfer of the bartoloside metabolic pathway may have occurred.\(^{31}\) The bartolosides share a core DAR motif, with long alkyl chains with at least one being chlorinated. In addition, the resorcinol ring itself is decorated with an O-linked xylose residue in 58 and C-linked xylose as well as O-linked rhamnose in 59–61. The very similar halogenated alkyl substituents in the DAR scaffold complicated structural elucidation using standard spectroscopy and spectrometric techniques. As the compounds were recalcitrant to crystallization, only by inferring the positioning of each of the two alkyl chains in 59 from biosynthetic studies was it possible to fully elucidate its planar structure by interpretation of NMR and MS\(^{5}\) data. The planar structures of 58, 60 and 61, were then elucidated by comparison of their spectroscopic data with that of 59.\(^{31}\)

In a later study using the cyanobacterium \textit{S. salina} LEGE 06099, closely related to the bartolosides B–D (59–61) producer (\textit{S. salina} LEGE 06155), Afonso \textit{et al.} reported new analogues of metabolite 58, bartolosides E–K (62–68) by following cytotoxic activity towards cancer cell lines. Most of the structural differences of these new analogues related to the length of the alkyl chain and the presence or absence of a chlorine substituent in the 5-alkyl chain. A notable exception is compound 66 that features a rare, mid-chain gem-dichloro moiety.\(^{52}\) Interestingly, this same strain produced bartoloside A in elevated amounts (reaching up to 0.6% of its dry weight), suggesting a potential structural role.\(^{52}\) The bartolosides have been tested for cytotoxicity towards several cancer cell lines but showed only mild activity with compound 59 being the most potent of the series (IC\(_{50}\) value 9 \(\mu\)M against PC-3 human prostate adenocarcinoma).\(^{53,54}\)

\subsection{2.2.3. Microcarbonin A.} In 2006, Beresovsky and collaborators\(^{53}\) reported microcarbonin A (69), a metabolite produced by a \textit{Microcystis} sp. strain (MB-K) collected from the Israeli freshwater Lake Kinneret. During bioassay-guided isolation, fractions containing 69 demonstrated a highly specific activity in inhibiting photosynthesis and causing the death of the dinoflagellate \textit{Peridinium gatunense} by abolishing its internal carbonic anhydrase activity. However, a gradual loss of its allelopathic activity toward the final step of the purification process was observed. This was attributed to the synergistic effect of the fatty acids residing with microcarbonin A in the partially purified fractions.\(^{53}\) Compound 69 features a rare 1-aryl-1,3-diyn in its 5-alkyl substituent and a fully aliphatic 2-alkyl chain. Among the cyanobacterial DARs, it is metabolite 69 that most closely resembles DARs from heterotrophic bacteria.

\subsection{2.2.4. Nostocyclyne A.} Following a bacterial sampling campaign aiming at understanding the massive death of the fern \textit{Adiantum capillus veneris} in a greenhouse of Neve Monomicson (Israel), Ploutno & Carmeli\(^{54}\) first described nostocyclyne A (70). The compound was the major active metabolite produced during a bloom of the terrestrial cyanobacterium \textit{Nostoc} sp. (TAU strain IL-220) that was thought to be associated with the observed plant mortality.\(^{24}\) Pure 70 did not show promising bioactivity towards Gram-negative bacteria, and it had only mild activity towards Gram-positives.\(^{24}\) Apart from the rare enediyne,\(^{55}\) compound 70 is, to the best of our knowledge, the only other example of a cyclic alkyne natural product. Curiously, 70 combines features of the cyclophanes (mid-alkyl chain C–C bond with resorcinol ring) with an internal alkyne moiety, which is in turn found in compound 69.

\section{3. Biosynthesis of alkylresorcinols in cyanobacteria}

\subsection{3.1. Monoalkylresorcinols from type III PKS pathways}

MARs are, to date, known only to be biosynthesized via type III PKS enzymes. Homodimeric and iterative, type III PKSs are typically smaller and simpler than type I and II PKSs, but they retain nonetheless elongation and cyclization functionalities.\(^{4}\) Generally, type III PKSs use a thioester as the starter unit and perform elongation through the condensation of extender units, usually malonyl-CoA, to form a polyketide chain that is then intramolecularly cyclized \textit{via} aldol condensation, Claisen condensation or lactonization.\(^{56}\) Despite the simplicity of their reactivity, by diversifying the starter and extender units, number of condensation steps and cyclization mechanisms, these enzymes are responsible for the biosynthesis of a breadth of compounds, not only alkylresorcinols but also alklypyrones, chalcones, stilbenes, quinolones and acylphloroglucinols, among other.\(^{57,58}\) MARs are generated by a specialized group of type III PKS, the alkylresorcinol synthases (ARS), which catalyze the condensation between fatty acyl-CoA (or-ACP) thioesters (starter units) and malonyl-CoA thioesters (or other extender units), followed by an aldol-type cyclization. Most of the currently characterized ARSs are bacterial, and only few plant and fungi-derived examples are known.\(^{5}\) The first ARS enzyme, ArSB, which was characterized in 2006 from \textit{Azotobacter vinelandii}, is involved in the biosynthesis of cyst envelope alkylresorcinols (e.g. compound 1).\(^{6}\) Since then, a few other homologs have been characterized, namely SrsA from \textit{Streptomyces griseus} (that produces 6-methyl-5-alkylresorcinols), MMAR\(_{2470}\) and MMAR\(_{2474}\) from \textit{Mycobacterium marinum} which provide methylated alkylresorcinols,\(^{59}\) the alkylresorcylic acid-generating FtpA from \textit{Myxococcus xanthus},\(^{13}\) BpsA from \textit{Bacillus subtilis} which produces long-chain MARs,\(^{60}\) as well as Dpga\(^{61}\) and Dbps\(^{57}\) which produce the MAR \(S\)-3,5-dihydroxyphenylglycine (a common moiety in actinobacterial glycopeptide antibiotics).

Apart from the MAR-mediated biosynthesis of cyanobacterial cyclophanes\(^{6}\) (see Section 3.2.1.), only very recently the first cyanobacterial biosynthetic pathway leading to MAR end products (hierridins 5 and 6) has been reported.\(^{19}\) The \textit{hid} BGC was found in the genome of the hierridin-producer \textit{Cyanobium} sp. LEGE 06113 by mining it for type III PKSs. This small BGC codes for three proteins with functions consistent with hierridin biosynthesis: a type-III PKS (HidC), a SAM-dependent methyltransferase (HidA) and a FAD-dependent oxidoreductase (HidB) (Fig. 3). Based on \textit{in vivo} assays and the biosynthetic
logic associated with Hid enzymes, hierridin biosynthesis putatively involves four main steps. First, HidC accepts thioester-activated palmitic acid and performs four elongation steps with malonyl-CoA and finally a decarboxylative condensation (C2-to-C7 aldol condensation) forms the alkylresorcinol ring (Fig. 3). Methylation of resorcinol phenol groups by HidA likely ensues, since a recombinant version of HidA was able to methylate 5-pentadecylresorcinol. This observation not only established experimentally the involvement of the *hid* BGC in hierridin biosynthesis, but also suggested that methylation precedes HidB-mediated hydroxylation of the resorcinol ring to yield hierridin B (5). The last biosynthetic step involves regioselective chlorination of 5 generating hierridin C (6). However, as observed for the resorcinol-halogenated bartolosides, the *hid* BGC does not encode a halogenase and no halogenase was found in the genome of *Cyanobium* sp. LEGE 06113, and so the nature of this halogenation is still unknown. 

3.2. Dialkylresorcinols

3.2.1. From monoalkylresorcinol dimerization. The natural paracyclophanes contain an unusual dialkylresorcinol motif and symmetrical structures with an intriguing biosynthetic origin. Nevertheless, the complete biosynthetic pathway for natural paracyclophanes eluded discovery for many years because it involves a cryptic activation step. The final dialkylresorcinol is formed from a head-to-tail dimerization of two monoalkylresorcinol precursors. Paracyclophane biosynthesis has been studied by a combination of isotope labeling experiments, bioinformatics and biochemistry, and the key biosynthetic steps are now generally understood, with the essential enzyme activities having been demonstrated *in vitro*.

The first insights into paracyclophane biosynthesis came from feeding studies with isotopically labeled acetate. Bobzin and Moore fed a mixture of unlabeled and [1,2-13C₂]-labeled acetate to the cylindrocyclophane producer *Cylindrospermum licheniforme* ATCC 29204. The resulting cylindrocyclophane D contained adjacent 13C atoms in specific positions (Fig. 4), indicating that acetate is incorporated into the final product. Feeding of [1-13C,2H₃]acetate and [2-13C,2H₃]acetate further indicated that the C-1 and C-2 positions of acetate are incorporated as expected for an aromatic polyketide (Fig. 4). Similar results were also obtained from the nostocyclophane producer *Nostoc linckia* UTEX B1932. From these data, and the 2-fold symmetry of the paracyclophane scaffold, Bobzin and Moore correctly proposed that natural paracyclophanes are biosynthesized from a polyketide-derived alkylresorcinol precursor that undergoes dimerization and that the β-methyl groups in compound 23 were derived from acetate and not from methionine. This realization was key to the Balskus’ group strategy for the discovery of the biosynthetic gene cluster for

<table>
<thead>
<tr>
<th><img src="https://example.com/diagram.png" alt="Diagram" /></th>
<th><img src="https://example.com/diagram.png" alt="Diagram" /></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fig. 3</strong> Biosynthesis of hierridins. Top – Hierridin biosynthetic gene cluster (<em>hid</em>). Bottom – Proposed biosynthesis of hierridins B (5) and C (6).</td>
<td><img src="https://example.com/diagram.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

**Fig. 4** Key findings that provided insight into the overall biosynthesis of cylindrocyclophanes. (A) Feeding experiments with [13C]-labeled acetate produce cylindrocyclophanes with a distinctive labeling pattern. [1,2-[13C]acetate produces adjacent [13C] atoms in the positions indicated in green, with the 1- and 2-positions of acetate incorporated into specific positions. (B) Discovery of the cylindrofridins, which contain a chlorine atom at the site of carbon–carbon bond formation, matched a retrosynthetic hypothesis for the origin of cylindrocyclophanes. Cylindrofridin A (7) is the monomeric precursor to cylindrocyclophane D (23), while cylindrofridins B (55) and C (56) likely represent the product of partial macrocyclization.
cylindrocyclophanes (cly).\textsuperscript{5,6,8} Because β-methyl groups are introduced into polyketides by a highly-conserved pathway that includes a homolog of hydroxymethylglutaryl-CoA (HMG-CoA) synthase, using degenerate PCR and whole-genome sequencing, Nakamura et al.\textsuperscript{*} identified an HMG-CoA synthase homolog in the cylindrocyclophan producer \textit{C. licheniforme} ATCC 29412. The HMG-CoA synthase homolog gene (clyF) was found to be part of a gene cluster (clyA-\textit{P}) that encodes other enzymes whose predicted activities appeared related to cylindrocyclophane biosynthesis.\textsuperscript{8} The cluster's biosynthetic role was confirmed by heterologous expression and purification of the type III PKS CylII, which was shown to convert a diketide into an alkylresorcinol that resembles the putative precursor of cylindrocyclophane biosynthesis.\textsuperscript{6}

The \textit{cly} cluster contains an abbreviated set of polyketide synthases that perform only four elongation reactions. This activity is insufficient to explain the biosynthesis of the paracyclophane alkyl chain. Feeding studies with \textit{d}_{13}C-decanoic acid led to the production of \textit{d}_{13}C-cylindrocyclophane F\textsubscript{6},\textsuperscript{6} indicating that the alkyl tail originates from fatty acids instead of from a polyketide, which is nonetheless consistent with the original feeding studies of Bobzin and Moore.\textsuperscript{64} Purification and characterization of CylA and CylB, a fatty acid adenylating enzyme and acyl carrier protein (ACP), respectively, further confirmed that cylindrocyclophane biosynthesis begins with the recruitment of free decanoic acid.\textsuperscript{8} Because the alkyl chain of decanoic acid is incorporated into cylindrocyclophanes \textit{in toto}, the dimerizing macrocyclization event necessarily requires functionalization of an unactivated carbon center.\textsuperscript{6} However, the true identity of the activated precursor remained elusive because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}

The dimerization mechanism was clarified with the discovery of the cylindrofridins from other cylindrocyclophane producers.\textsuperscript{65} Cylindrofridins 55 and 56 are nearly chemically identical to the cylindrocyclophanes, except that one of the dialkylresorcinol motifs is disconnected. Instead, a chlorine atom occupies the position of C–C bond formation in one alkyl tail (Fig. 4). Chlorination had been proposed as an activating dialkylresorcinol motifs is disconnected. Instead, a chlorine identical to the cylindrocyclophanes, except that one of the alkylation,\textsuperscript{65} chlorination has been proposed as an activating dialkylresorcinol motifs is disconnected. Instead, a chlorine because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}

Because the alkyl chain of decanoic acid is incorporated into cylindrocyclophanes \textit{in toto}, the dimerizing macrocyclization event necessarily requires functionalization of an unactivated carbon center.\textsuperscript{6} However, the true identity of the activated precursor remained elusive because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}

The dimerization mechanism was clarified with the discovery of the cylindrofridins from other cylindrocyclophane producers.\textsuperscript{65} Cylindrofridins 55 and 56 are nearly chemically identical to the cylindrocyclophanes, except that one of the dialkylresorcinol motifs is disconnected. Instead, a chlorine atom occupies the position of C–C bond formation in one alkyl tail (Fig. 4). Chlorination had been proposed as an activating dialkylresorcinol motifs is disconnected. Instead, a chlorine because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}

The dimerization mechanism was clarified with the discovery of the cylindrofridins from other cylindrocyclophane producers.\textsuperscript{65} Cylindrofridins 55 and 56 are nearly chemically identical to the cylindrocyclophanes, except that one of the dialkylresorcinol motifs is disconnected. Instead, a chlorine atom occupies the position of C–C bond formation in one alkyl tail (Fig. 4). Chlorination had been proposed as an activating dialkylresorcinol motifs is disconnected. Instead, a chlorine because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}

The dimerization mechanism was clarified with the discovery of the cylindrofridins from other cylindrocyclophane producers.\textsuperscript{65} Cylindrofridins 55 and 56 are nearly chemically identical to the cylindrocyclophanes, except that one of the dialkylresorcinol motifs is disconnected. Instead, a chlorine atom occupies the position of C–C bond formation in one alkyl tail (Fig. 4). Chlorination had been proposed as an activating dialkylresorcinol motifs is disconnected. Instead, a chlorine because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}

The dimerization mechanism was clarified with the discovery of the cylindrofridins from other cylindrocyclophane producers.\textsuperscript{65} Cylindrofridins 55 and 56 are nearly chemically identical to the cylindrocyclophanes, except that one of the dialkylresorcinol motifs is disconnected. Instead, a chlorine atom occupies the position of C–C bond formation in one alkyl tail (Fig. 4). Chlorination had been proposed as an activating dialkylresorcinol motifs is disconnected. Instead, a chlorine because it was not initially detected in culture extracts.\textsuperscript{5,6,8,63}
(CylJ, CylL, CylM, and CylN) may provide other as-yet undiscovered decorations to the paracyclophane scaffold. For example, the ribocyclophanes (51–54, 57)* might originate from the action of CylN, annotated as a glycosyltransferase.*

3.2.2. From monoalkylresorcinol intramolecular cyclization. Echoing the original proposal for paracyclophane biosynthesis, Ploutno and Carmeli hypothesized that the entirety of nostocyclene A (70) originates from a polyketide precursor* (Fig. 6). However, in light of the known steps for cylindrocyclophane biosynthesis, the structure of 70 suggests it might have a similar biosynthetic origin to the cyanobacterial paracyclophanes, where a fatty acid precursor is functionalized and intramolecularly cyclized using homologs of CylC and CylK, with the monoalkylresorcinol being synthesized by a homolog of CylI (Fig. 6). The presence of a single internal alkyne bond is especially interesting, as the enzymes that give rise to internal alkynes are poorly documented. Unfortunately, Ploutno and Carmeli noted that the *Nostoc* strain isolated from the bloom material (from which 70 was isolated) did not produce the metabolite under laboratory conditions,* suggesting the strain is either not the native producer or that 70 is the degradation product of a parent molecule. An investigation of cyanobacteria

---

**Fig. 5** Biosynthesis of cyanobacterial cyclophanes. Overview of the BGCs for different cyclophane classes (top) and biosynthetic pathway for the cylindrocyclophanes (bottom), as illustrated by cylindrocyclophane D (23). Enzymes that have been purified and characterized in vitro* are shown in bold.
that contain homologs of CylC, CylK, and CylI may allow the identification of a nostocyclyne-producing strain.

3.2.3. From fatty acid condensation. The biosynthesis of bacterial dialkylresorcinols (DARs) was first studied in the 70s and 80s, with Sankawa and colleagues proving the involvement of two polyketide chains in the generation of the DAR scaffold. In 2003, Nowak-Thompson et al. provided the first experimental evidence suggesting that DARs are synthetized by a head-to-head condensation of two fatty-acid-derived substrates through the activity of a ketosynthase, DarB, followed by a putative aromatase DarA. The two enzymes and an ACP were found to be encoded by adjacent genes in the genome of Pseudomonas aurantiaca strain BL915 and gave rise to heterologous production of the DAR 2-hexyl-5-propylresorcinol in Escherichia coli (Fig. 7). More recently, Fuchs et al. experimentally showed, through in vivo and in vitro assays, that β-keto- and α,β-unsaturated-fatty acyl thioesters acted as DarB substrates and gave rise to carboxy-dialkylcyclohexanediones (carboxy-CHDs) and CHDs that could potentially act as substrates for the putative aromatase DarA. The authors found that DarB was a new phylogenetically distinct class of ketosynthase enzyme, which had lost the elongation ability in favor of the head-to-head fatty acid condensation ability and that darAB genes are widely spread among bacteria, including in cyanobacterial strains belonging to the genera Fischerella and Microcystis.

The first characterized cyanobacterial DarAB-like pathway was the bartoloside BGC (brt) from the cyanobacterium Synechocystis salina LEGE 06155. This gene cluster (Fig. 7) is involved in the biosynthesis of diglycosylated 59–61 and encodes, apart from DarA and DarB, three glycosyltransferases, a glycolipid export operon, a CylC homolog and a CylK homolog, as well as a hypothetical protein. An additional brt BGC was later reported by Afonso et al. from the closely-related cyanobacterium S. salina LEGE 06099 and is involved in the biosynthesis of the monoglycosylated 58 and 62–68. Just as in the previously characterized DarAB system, the biosynthesis of bartolosides is initiated by BrtD (DarB homolog) that connects the two fatty acid derivative thioesters via Claisen condensation and simultaneous Michael addition, yielding a carboxy-CHD (Fig. 7). All diglycosylated bartolosides (B–D) are synthetized by condensation of C16 and C18 fatty acyl-derivatives, namely 2-octadecenoyl-ACP and 3-oxo-hexadecanoyl-ACP thioesters. Regarding monoglycosylated bartolosides, the initial substrates likely have different chain lengths (C14–C18), including odd-chain C15 and C17 α,β-unsaturated fatty acyl-ACP thioesters. The origin of these odd-chain precursors is still unclear. Although brt lacks a dedicated ACP, unlike previous characterized DAR clusters, it is foreseen that the ACP-bound fatty acyl derivatives are recruited from primary metabolism, especially considering the unlikely participation of acyl-CoA substrates due to the apparent absence of a canonical β-oxidation pathway.
in cyanobacteria. By using heterologously-expressed BrtC (DarA homolog) and BrtD in a coupled assay, Leao et al. provided the first biochemical evidence of DarA enzymes acting as aromatases, as conversion of the fatty acyl derivative surrogates to the DAR backbone of was observed only in the presence of BrtC. Backed by crystallographic data, Mori et al. showed that in bacterial stilbene biosynthesis by a DarAB system (StlC/StlD from Photorhabdus luminescens), decarboxylation occurs during aromatization, clarifying that carboxy-CHDs are the substrates for DarA aromatases. Following DAR backbone generation, the biosynthesis of the bartolosides should proceed with the decoration of the resorcinol with sugar moieties, namely xylose and rhamnose by glycosyltransferases BrtA, E and F (diglycosylated bartolosides) or Brt A and E (monoglycosylated bartolosides), although it is unclear whether all these enzymes are involved in the glycosylations. Halogenation (including gem-dichlorination) in the alkyl chains is predicted to be carried out by the CylC homolog BrtJ. The timing of halogenation in bartoloside biosynthesis is unclear, but it is foreseeable that BrtJ—like CylC—acts on ACP tethered fatty acyl derivatives, i.e., before DAR formation. As for hierridin C (6), bartoloside D (61) is halogenated in the resorcinol ring but no additional halogenase has been found in the genome of S. salina that could explain this functionality. Intriguingly, BrtB enzymes, found in both brt clusters, are CylK homologs that could alkylate nucleophilic substrates using the mid-chain halogenated positions in the bartolosides as electrophiles. Still, such reactivity is not required to explain the biosynthesis of currently known bartolosides and bartolosides do not have a free C2 position as do known BrtK nucleophiles. Apart from the biosynthetic genes, the brt cluster also encodes a DevBCA-like glycolipid export system, BrtGHI, which could be responsible for the presence of bartolosides in spent culture medium or for potential incorporation into the outer membrane.

Fuchs et al. reported that in Microcystis sp. PCC 9808, a pair of DarA and DarB homologs were part of a small BGC that also encoded two desaturases, a DevBCA exporter system and an oxygenase (Fig. 8). Taking into consideration such predicted reactivity, the cluster was proposed to encode microcarbonin A (69), produced by a Microcystis sp. strain, although there is no direct experimental evidence supporting this proposal. Hypothetically, the carboxy-CHD resulting from the DarB-mediated condensation of C14 and C12 fatty acyl derivatives is aromatized by the DarA homolog (Fig. 8). Like bartolosides, microcarbonin A BGC also lacks an ACP, suggesting that the fatty-acyl precursors are obtained from primary metabolism. Potentially after DAR formation, the two desaturases likely catalyze the dehydrogenation of the fatty-acid chains, creating the two...
internal alkyne moieties. A similar alkyne system is found in 70 and these are of considerable biosynthetic interest. The DevBCA transporter, as in the case of bartolosides, suggests export to the outer membrane or to the extracellular medium. The oxygenase has no corresponding homolog in characterized DAR clusters and it is unclear what its role might be in the biosynthesis of 69.

### 3.3. Diversity of alkylresorcinol pathways in cyanobacterial genomes

Genomic data for the phylum Cyanobacteria have been increasing over the past years, mostly as a result of a few major sequencing efforts,71–73 but most genera have either none or a single assembly available publicly. As of August 2018, over a thousand cyanobacterial genomes were publicly available in the GenBank database, more than half of which corresponding to the smaller-genome genera Prochlorococcus and Synechococcus picocyanobacteria. This wealth of genomic data can be used to obtain a predictive overview of MAR and DAR diversity (according to currently understood biosynthesis) in the phylum Cyanobacteria. In fact, Nakamura et al. have done this exercise for CyI/CyI/K containing cyanobacterial genomes and found a series of different orphan BGC archetypes that likely correspond to unprecedented natural product classes. Likewise, Costa et al. who looked into the distribution of hid-like BGCs in cyanobacteria observed that these are widely distributed in picocyanobacteria. Within the scope of this review, we carried out a distribution analysis of MAR- and DAR-encoding pathways among currently available cyanobacterial amino acid sequence data (mostly from genome-sequencing studies) available in the GenBank. We performed phylogenetic analysis of type II PKSs (Fig. 9) and DarB homologues (Fig. 10) which overall suggest that most scaffolds of MAR and DAR types are still to be uncovered in this phylum.

Type II PKSs in cyanobacteria are distributed among eight well-supported clades (Fig. 9). The majority of these enzymes are HidC homologs and are found among picocyanobacteria (hidABC clade) and other Synechococcales (hidCA/hidCBA clade). These are part of short clusters that include always a methyltransferase homolog of HidA and are likely to correspond exclusively to MAR end-products. The paracyclophane-associated type III PKSs clade together, as expected, and are, in turn, closely-related to three additional clades (clades “1”, “2” and “3”, Fig. 9). One such clade is composed of Microcystis and Nostoc enzymes, and they are part of clusters that include fatty acid activation enzymes and type I PKSs. Three Moorea spp. type II PKSs compose “clade 2” and are part of BGCs encoding type I PKSs and methyltransferases. “Clade 3” is composed of type III PKS homologs from Cylindrospermum, Cynothece and Pleurocapsa strains. These are encoded within heterogeneous BGCs that share fatty acid activation and type I PKSs. Interestingly, from all the BGCs extracted for the three clades closest to the paracyclophanes clade, only in Cylindrospermum sp. NIES-4074 (“clade 3”) a pair of CyIC/CyIK homologs was detected, suggesting that all other pathways encode MARs. An additional cyanobacterial type III PKS clade observed in the phylogenetic tree (Fig. 9) is composed only of a Moorea and a Roseofilum strain. These homologs are closely related to the actinobacterial DpgA, that is involved in the biosynthesis of the non-proteinogenic amino acid 3,5-hydroxyphenylglycine, and are found in small (~5 kb) clusters that also encode DpgB (dehydratase that is associated with DpgA). The remaining well-supported clade of cyanobacterial type III PKSs includes homologs of MtbPKS18, that takes part in alkylpyrone biosynthesis in Mycobacterium tuberculosis. The potential alkylpyrone-generating cyanobacterial homologs are part of relatively small gene clusters (~15 kb), most of which contain a putative cytochrome b_{5} (desaturase) adjacent to the type III PKS.

The DarB phylogeny shows a less diverse topology for cyanobacteria, with three related clades with strong bootstrap support that are closely-related and likely involved in DAR biosynthesis since they have adjacent DarA homologs (Fig. 10). One of these clades includes only the two BrtD homologs from S. salina. Another small clade is composed of two DarB homologs from Microcystis aeruginosa that correspond to potential
microcarbon-generating BGCs\(^7\) (see also Fig. 8). Considerably more diversity of BGC archetypes is found within the third cluster, which comprises homologs from heterocystous Nostoccales and Stigonematales strains. The associated BGCs are relatively large (>25 kb) and encode diverse functionalities including type I PKSs, DevBCA transporters, acetyltransferases and oxidoreductases. Most also include CylC/CylK pairs.\(^8\)

Interestingly, other cyanobacterial DarB homologs have no DarA homolog in an adjacent position (nor in their immediate genomic context). These DarB homologs are nevertheless evolutionarily more distant from their DarA-associated cyanobacterial counterparts than are those from other bacteria (including the biochemically characterized DarB and StlD), which suggests that their reactivity might be different from DarB and, consequently, that their reaction products may not be carboxy-CHDs. These “DarA-less” DarB enzymes are grouped in two clades (one with homologs from two Acaryochloris sp. strains and a larger that encompasses homologs from a wider variety of genera), as well as a loner Nostoc punctiforme homolog.

The latter (Nostoc) homolog is encoded in a large BGC (~60 kb) that encodes multiple type I PKSs and different fatty acid activating- or modifying functionalities. The BGCs associated with
the two clades are more simple and do not contain type I PKSs (Fig. 10).

Overall, the phylogenetic analysis and annotation of the genomic context of cyanobacterial type III PKS and DarB homologs shows that MARs and DARs are produced by a wide range of cyanobacteria, from the simpler picocyanobacteria to the cell-differentiating Nostocales and Stigonematales. For the overwhelming majority of BGCs that encode these MAR and DAR-generating enzymes, we cannot currently assign a natural product. Moreover, these orphan BGCs are diverse in size and encoded functionalities. Still, and as expected, several fatty acid activating and modifying enzymes are present in the BGCs. Interestingly, the presence of CylC/CylK homolog pairs seems punctuated among the phylogenetic clades and among BGCs of similar architecture. Taking into account that these enzyme pairs are involved in dimerization or eventually polymerization, their incorporation in BGCs might result in gain of new functions for the encoded MARs or DARs. Also of considerable interest is the relatively widespread occurrence of DarB homologs in BGCs without an associated aromatase, whose end-products remain to be discovered.

4. Synthetic approaches to cyanobacterial alkylresorcinols

The time-consuming growth culture experiments and low yields obtained for a number of bioactive alkylresorcinols have raised the need for alternative and more effective means to access and
study these secondary metabolites. Due to their simple and planar structure, relatively short and simple organic syntheses of the monoalkyl derivatives hierridins have been achieved already in the 1990s. The uniqueness of the [7.7]paracyclophane scaffold has further drawn the attention of the synthetic community. This has culminated in the development of admirable series of reactions towards the preparation of the dimeric macrocycles cylindrocyclophanes A (16) and F (25).

4.1. Monoalkylresorcinols (MARs)

Many studies have described the synthesis of natural MARs, employing several approaches to accomplish the critical step of carbon–carbon bond formation between the resorcinol ring and the alkyl side chain. In some cases, additions using a Gilman reagent,75 a Suzuki76 or a Sonogashira3 coupling were adopted; however, Wittig and Grignard reactions are the most typically used.77,78 The few procedures dedicated to the synthesis of hierridins relied on one of the two latter alkylation strategies, using a readily available brominated or methoxylated salicylaldehyde derivative as starting material.

In 1998, Jinno and Okita79 reported the total synthesis of the yeast-derived alkylphenol phaffiol (also known as hierridin A,4 see Section 2.1.1.). To this end, 3,5-dibromosalicylaldehyde (71) was subjected to a Wittig olefination to produce the key intermediate 72 as a 1:1 mixture of E/Z isomers (57% yield). Reduction of the alkyl phenol 72 was then carried out by hydrogenation, providing the alkyl phenol 73. Finally, the nucleophilic aromatic substitution of the dibromide 73 with sodium methoxide afforded the expected MAR 4 (67% yield from 73) (Scheme 1).

The extended chain analogue hierridin (8), isolated from the lichen Ramalina hierrensis, was synthesized in five steps by González and co-workers.80 The phenolic aldehyde precursor o-vanillin (74) was first alkylated in the presence of an n-eicosanyl Grignard reagent to generate compound 75 (yield 65%). Subsequent hydrogenation followed by oxidation in the presence of salcomine under a stream of oxygen provided the quinone intermediate 76, which was further subjected to sodium dithionite reduction to afford the corresponding hydroquinone. The methylation of the sterically most accessible

Scheme 1  Synthesis of hierridins 4–6 and 8.19,79,80
hydroxy group by methyliodide and K$_2$CO$_3$ under reflux for 24 h provided hierridin (8) (Scheme 1).

Costa and co-workers$^{19}$ recently achieved the total synthesis of hierridin B (5), which contains a 15-carbon length alkyl side chain moiety. In contrast to the synthetic routes described above for hierridins 4 and 8, their method used 2-hydroxy-3,5-dimethoxybenzaldehyde (77) as the initial building block, avoiding a subsequent methoxylation step in position 5. The synthesis was thus achieved in only two steps, via Grignard reaction of 77 to produce the intermediate 78, which was then hydrogenated to afford 5 in moderate yield (38% from 78) (Scheme 1). The authors suggested that the methoxy group in position 5 negatively influences the efficiency of the reduction step, since the product yield obtained was significantly lower than that reported in the synthesis of compound 8 (38% versus >90%).$^{18}$

The chlorinated derivative hierridin C (6) has been synthesized concurrently to hierridin B (5).$^{19}$ Since direct chlorination of 5 in the appropriate position failed, a procedure consisting in monohalogenation of 77 prior to the Grignard reaction was undertaken, but the outcome of the subsequent hydrogenation step was unsuccessful. An alternate synthetic route for the carbon–carbon bond formation was applied, using a similar approach employed for compound 4. This included the regioselective chlorination of 77 to afford the compound 79. Wittig olefination of 79 followed by palladium-catalyzed hydrogenation led to monoalkylresorcinol 6 (Scheme 1).

These strategies provided 14% to 79% overall recovery of hierridins. Despite the efficiency of the Grignard addition, the subsequent reactions leading to hierridins 5 or 8 led to substantial decrease of the global yield. Furthermore, using the precursor 2-hydroxy-3,5-dimethoxybenzaldehyde (77) as starting material enabled shorter synthesis of 5 and 6. Thus it may be worthwhile to revisit the synthesis of hierridins 4, 5 and 8 by involving a Wittig reaction directly on compound 77 to improve the overall yields.

**4.2. Dialkylresorcinols (DARs)**

Effective synthetic routes to simple natural DARs have been designed.$^{81,82}$ In these studies, alkylated α,β-unsaturated ketones and diethylmalonate were subjected to
Michael addition/Dieckmann condensation followed by saponification/decarboxylation to generate alkylcyclohexane-1,3-diones. Different strategies were then employed to form 2,5-dialkylresorcinols, involving first a 2-alkylation, if necessary, prior to aromatization. However, to the best of our knowledge, the total synthesis of the functionalized cyanobacterial DARS bartolosides and microcarbomin A (69) has not been reported yet.

Cyclophanes possess a more complex architecture, with a macrocyclic system representing a rigid cavity of great interest in host–guest chemistry. For this reason, their synthesis has been extensively investigated since the 1950s. From a synthetic perspective, closure of the ring core accounts for the most challenging step, which has led to the consideration of multiple strategies for the formation of such key carbon–carbon bonds in cyclophanes. A few of these strategies have been applied to the total synthesis of compounds from the [7.7]paracyclophane series (Scheme 2). Mascal and co-workers proceeded to the synthesis of a diester precursor prior to cyclization in [7.7]paracyclophane via intramolecular acyloin condensation. In an attempt to prepare benzocyclononanone, Schubert et al. used high dilution Friedel–Crafts reaction between monoalkylated monomers, which resulted in the fortuitous synthesis of a [7.7]paracyclophanedione (5% yield). Another study by Staab and co-workers used bis(dithiane) alkylation for preparation of a tetramethoxy[7.7]paracyclophane-4,17-dione. As for the synthesis of the C$_2$-symmetric cylindrocyclophanes A (16) and F (25), head-to-tail cyclodimerizations were performed via more advanced procedures, furnishing diene macrocycles as intermediates. Smith and co-workers as well as Yamakoshi et al. employed a tandem cross metathesis/ring-closing metathesis (RCM). Macro cyclic dimerization by a double Horner–Wadsworth–Emmons olefination was reported by Hoye et al. while Nicolaou and co-workers envisioned the Ramberg–Bäcklund olefination reaction to successfully obtain the [7.7]paracyclophane scaffold leading to the formation of 16 and 25.

The natural compounds (−)-cylindrocyclophane A (16) and (−)-cylindrocyclophane F (25) possess six and four stereocenters, respectively. Two strategies were used to obtain the desired configurations of (−)-16 and (−)-25. While two groups, Smith et al. and Yamakoshi and co-workers, introduced all stereocenters before the dimerization step, two other groups, Hoye et al. and Nicolaou and co-workers, preferred establishing only the stereogenic center bearing the n-butyl unit prior to macrocyclization. Asymmetric syntheses were therein

**Scheme 3** Synthesis of cylindrocyclophane F (25) by Smith and co-workers.
required at the end of the process to settle the desired stereochemistry for hydroxy- and methyl-bearing carbons.

Smith and co-workers reported the first total synthesis of cylindrocyclophane F (25) in 1999, through a stepwise procedure to prepare the dimer structure before closing the macro-ring (Scheme 3). The imide 81 was used as starter unit to construct intermediate 82, ensuring stereoccontrol in the initial step introducing the allyl chain. A Danheiser benzannulation between 82 and cyclobutenone 83 and subsequent methylation furnished the resorcinol intermediate 84, which was then converted to iodide 85 and tosyl hydrazone 86. The assembly of 85 and 86 using lithium–halogen exchange followed by Myers reductive coupling led to compound 87 in 73% yield. Conversion of the latter to diene 88 followed by olefin metathesis using a Grubbs ruthenium catalyst provided the paracyclophane 89, which after hydrogenation and removal of the methyl ethers afforded cylindrocyclophane F (25).

The same authors published a second-generation synthesis of compound 25 revisiting the macrocycle assembling strategy through the one-pot cross metathesis/RCM (Scheme 3).

To this end, instead of fragment 83, cyclobutenone 90 was synthesized and united to siloxyalkyne 82 to form the paracyclophane monomer 91. For dimerization of 91, substantial improvements in reactivity were found in the presence of Schrock molybdenum catalyst in comparison to Grubbs ruthenium catalysts. Hydrogenation of the diene macrocycle 92 and cleavage of the methyl ethers afforded cylindrocyclophane F (25). While the first-generation synthesis with 20 steps gave 8.3% overall yield of targeted compound 25, only 11 steps were required with the second method to result in 22% global yield.

Since the cross olefin metathesis provides the thermodynamically most favored isomer possessing double bonds with trans geometry and seven-carbon atom tethers, this reaction was also exploited for the synthesis of the natural hydroxylated analogue 16. A disconnection study revealed a monomer hydroxylated on the side alkyl chain, which was synthesized by Smith and co-workers as follows. Stannyl cyclobutenone 93 underwent a Danheiser annulation with compound 82 to provide aryl iodide 94. Addition of the amide 95 via lithium–halogen exchange and conversion of the ketone intermediate

---

**Scheme 4** Total synthesis of cylindrocyclophane A (16) by Smith and co-workers (A and B) and Yamakoshi and co-workers (C and B).
obtained to the corresponding alcohol in a stereoselective manner provided compound 96 (Scheme 4A). Finally, dimerization of monomer 96 using Schrock catalyst gave the bis(olefin) 97, which afforded cylindrocyclophane A (16) after alcohol deprotection, hydrogenation and ether demethylation (Scheme 4B). An overall yield of 8.1% was recovered from a total of 16 steps.

The tandem dimerization-macrocyclization scheme developed by Smith and co-workers in 2000 (ref. 89) was also implemented by Yamakoshi and co-workers in 2009, who synthesized both cylindrocyclophane A (16) and its non-natural dextrorotatory enantiomer (Scheme 4C). A Sonogashira coupling of syringaldehyde (98) and alkyne 99 prepared from l-diethyl tartrate resulted in the tetrasubstituted resorcinol 100. In the synthetic route leading to 16, a diastereoselective reduction of the alkyne led to an intermediary E-alkene, which was subjected to a Johnson–Claisen rearrangement prior to a Wittig reaction to establish the n-butyl group. This afforded benzaldehyde 101 in ten steps with 33% yield. Within eight subsequent steps, compound 101 underwent a diastereoselective aldolisation in the presence of oxazolidinone 102, followed by extension of the side alkyl chain via Horner-

**Scheme 5** Synthesis of cylindrocyclophane A (16) by Hoye and co-workers and Nicolaou’s approach to synthesise cylindrocyclophanes A (16) and F (25).
Wadsworth–Emmons olefination to furnish the paracyclophane monomer 96. The final steps were common to the procedure adopted by Smith and co-workers, apart from the choice of ring-closing metathesis catalyst, being a Grubbs second-generation complex (Scheme 4B). Overall, 3.5% yield of compound 16 was obtained in a 20-step synthesis.

Hoye and co-workers91 also used a resorcinol derivative, 4-bromo-3,5-dimethoxybenzyl alcohol (103), as starting material for the total synthesis of cylindrocyclophane A (16) (Scheme 5). The synthetic route targeted the paracyclophane monomer 106, which possesses a seven-carbon length side chain bearing the butyl moiety. To this end, alkene 104 was synthesized in four steps, prior to enzyme-assisted conversion of the allylic alcohol to (R)-acetate 105. A [3,3]-sigmatropic rearrangement followed by reduction of the alkene provided the n-butyl group, while the ester side chain formed was elongated via Horner–Wadsworth–Emmons reaction. Transformation to a phosphono ester and oxidation of the benzylic alcohol furnished the bifunctional monomeric unit 106, which yielded E,E-diene 107 through Horner–Wadsworth–Emmons macrocyclic dimerization. In order to introduce the four stereocenters holding the methyl and the hydroxy groups, the ester groups were converted to methyls and stereoselective hydroboration followed by H2O2 oxidation led to the requisite enantiomer (53% yield over four steps). The final step entailed an efficient demethylation to release cylindrocyclophane A (16). This alternate synthetic route revealed an overall yield of 1.6% obtained in 22 steps. Furthermore, the olefination reaction that used the saturated monomer 106 was found to be stereoselective, in contrast to its (E)-4,5-alkene homologue that led to the formation of a mixture of stereoisomers.

Nicolaou and co-workers employed a similar building block to construct the monomer 111 from bromide 109 (Scheme 5).92 Reaction with pentanal and nucleophilic addition of a vinyl lithium followed by allylic transposition-oxidation of the resulting tertiary alcohol provided enone 110. This product 110 bears an n-butyl residue and the five-atom side chain intended to connect the two resorcinol units. Subsequent enantioselective ketone reduction and catalytic olefin hydrogenation established the stereochemistry of the bifunctional monomer precursor 111. The macrocyclization occurred via displacement of the mesylate substituent to generate a (bis)thioether, which was further oxidized to bis(sulfone) 112. With the macrocyclization completed, a Ramberg–Bäcklund reaction of 112 with Chan’s modification provided a bis(olefin), which after asymmetric dihydroxylation afforded the macrocyclic diol 113, the common precursor for the synthesis of both cylindrocyclophanes A (16) and F (25). Cylindrocyclophane F (25) was obtained with good yield (71%) via a one-pot sequential reaction including methylation of the side chains and demethylation of the methoxy groups. Concomitantly, the route to cylindrocyclophane A (16) was envisaged through oxidation of alcohol to ketone, enol triflate formation and cross coupling with a Grignard reagent to furnish bis(olefin) 114. Compound 114 was then subjected to hydroboration, H2O2 oxidation and ether demethylation according to Hoye’s procedure. Nicolaou’s approach gave 1.7% (19 steps) and 0.6% (21 steps) overall yields of paracyclophanes 25 and 16, respectively.

In summary, the syntheses of MARs 4–6 and 8 were based on alkylation of a salicylaldehyde precursor via Wittig or Grignard reaction. The synthetic approaches leading to cylindrocyclophanes 16 and 25 first consisted in the preparation of a tetrasubstituted resorcinol monomer through Danheiser annulation, Sonogashira coupling followed by Johnson–Claissen rearrangement, or asymmetric functionalization. Head-to-tail cyclodimerization then occurred via olefin metathesis, double Horner–Emmons olefination or Ramberg–Bäcklund reaction. Final minor synthetic steps enabled to release or install functional groups. Thus, conventional organic synthesis methods have been employed to prepare the aforementioned compounds (16, 25). In contrast to the assumed biosynthetic pathways, the resorcinol ring was either constructed by reaction of a cyclobutene and an acetylene, or it was introduced through a commercially available functionalized aromatic derivative. Moreover, the macrocycle assembling step required p-substituted monomers, in contrast to the suggested natural dimerization of the resorcinol ring and the non-functionalized alkyl chain. As an alternative to conventional syntheses of natural products, biomimetic approaches have become popular in the last twenty years, revealing new powerful synthetic routes and enabling to confirm or invalidate biosynthetic hypotheses.96 Such a strategy has, for instance, been applied to the naturally occurring paracyclophane dimer longithorone A.97 Despite the challenging linkage formation at C(7) and C(20) and the configuration of the stereocenters, similar approaches mimicking natural biosynthetic pathways may open up new prospects towards efficient organic synthesis and stereochemical elucidation of cyanobacterial DAR series.

5. Conclusions

Only a small fraction of cyanobacterial alkylresorcinol diversity has been documented, but it is now evident that these structural motifs are present throughout the phylum. In particular, the DAR motif seems to be so relevant for cyanobacterial ecophysiology that two fundamentally different biosynthetic pathways have converged to create it in the same bacterial phylum. This is remarkable since there are limited known examples of mechanistically-distinct natural product biosynthetic pathways that have converged to create the same scaffold, and even then this usually occurs in separate kingdoms.98,99,100 Despite the dissimilar biochemistry involved in creating the DAR moiety through MAR biosynthesis followed by Friedel–Crafts alkylation or by head-to-head condensation of fatty acyl precursors, a number of decorations to the DAR scaffold are installed for both pathways, including internal alkynes, glycosylation and halogenation. Illustrative of this are the pairs of CylC and CylK homologs that are found in a considerable diverse set of cyanobacterial gene clusters of DarA and DarB but also type III PKS biosynthetic origin. These shared features likely reflect evolutionary pressure associated with the DAR skeleton in cyanobacteria and are probably reminiscent of related ecological
roles. Still, and contrary to what is known for MARs and DARs produced by other bacteria, the current understanding of the functions that cyanobacterial alkylresorcinols may play in the natural setting is minimal, although many have potent biological activity in pharmacologically-relevant assays. The uncovered diversity of alkylresorcinols in cyanobacteria represents an opportunity to perform genome-guided isolation efforts that have the potential to reveal new compound classes and important bioactivities; the study of their ecological roles will nevertheless remain a challenge.

6. Conflicts of interest

There are no conflicts to declare.

7. Acknowledgements

This work was funded by the BLUEandGREEN project (H2020 TWINNING, Grant Agreement no. 692419) and was further supported by an European Research Council (ERC) Starting Grant (759840) to PNL and National Science Foundation (NSF) through grant CHE-1454007 to EPB. We acknowledge funding from Fundação para a Ciência e a Tecnologia (FCT) through projects PTDC/BIA-BQM/29710/2017 (co-financed by COMPETE 2020, Portugal 2020, ERDF) UID/Multi/04423/2013, as well as through scholarship SFRH/BD116009/2016 to SP and grant IF/01358/2014 to PNL.

8. Notes and references