NATURAL PRODUCTS

Integrating mass spectrometry and genomics for cyanobacterial metabolite discovery

Nathan A. Moss · Matthew J. Bertin · Karin Kleigrewe · Tiago F. Leão · Lena Gerwick · William H. Gerwick

Received: 23 June 2015 / Accepted: 3 October 2015 / Published online: 17 November 2015
© Society for Industrial Microbiology and Biotechnology 2015

Abstract Filamentous marine cyanobacteria produce bioactive natural products with both potential therapeutic value and capacity to be harmful to human health. Genome sequencing has revealed that cyanobacteria have the capacity to produce many more secondary metabolites than have been characterized. The biosynthetic pathways that encode cyanobacterial natural products are mostly uncharacterized, and lack of cyanobacterial genetic tools has largely prevented their heterologous expression. Hence, a combination of cutting edge and traditional techniques has been required to elucidate their secondary metabolite biosynthetic pathways. Here, we review the discovery and refined biochemical understanding of the olefin synthase and fatty acid ACP reductase/aldehyde deformylating oxygenase pathways to hydrocarbons, and the curacin A, jamaicamide A, lyngbyabellin, columbamide, and a trans-acyltransferase macrolactone pathway encoding phormidolide. We integrate into this discussion the use of genomics, mass spectrometric networking, biochemical characterization, and isolation and structure elucidation techniques.

Keywords Cyanobacteria · Natural products · Biosynthesis · Mass spectrometry · Genomics

Introduction

Over the past three decades, natural products isolated from type III tropical filamentous cyanobacteria have provided numerous bioactive therapeutic lead compounds [51], as well as compounds with deleterious effects to human health [10]. In addition, a variety of methylated alkenes and alkanes are produced by several different clades of the phylum, including type III filamentous cyanobacteria [9]. The environmental role of these secondary metabolites is largely unknown, but potent activity in cytotoxicity assays suggests a potential ecological role as potential antagonistic or defense chemicals [42]. Along with interesting chemical functional groups and structural diversity, the cyanobacterial biosynthetic gene clusters described to date display a variety of novel biochemical features [22, 23]. In some cyanobacterial biosynthetic pathways, inter- and intra-species evolutionary adaptation is suggested by an apparent horizontal gene transfer within biosynthetic pathways, characterized by high gene homology but with a unique gene order and corresponding molecular structure [16].

Traditional isolation and structure elucidation techniques are robust and efficient when a secondary metabolite is produced in sufficient quantities. However, when they are produced in small quantities, several other newer approaches become helpful. For example, heterologous expression has been successfully accomplished to produce
the cyanobacterial natural products O-demethylbarbamide [25] and lyngbyatoxin [41], and further development of a cyanobacterial “toolbox” is likely to facilitate these types of efforts in the future [54]. Online bioinformatics tools such as NCBI DELTA-BLAST, antiSMASH [3, 33], NapDoS [62] and NRPSpredictor [46] enable prediction of biosynthetic enzyme function from DNA sequence information. Advances in mass spectrometric data processing and visualization via molecular networking also enable rapid detection of new compounds that are available in only small quantities [56, 60], and in some cases novel structures can be reliably assigned using innovative algorithmic methods [40]. New cryoprobe designs for high-field NMR coupled with FAST data acquisition techniques are extending the reach of NMR-based structure elucidation to low nanomole quantities of natural products [5, 39].

This review summarizes a number of recent advances in the study of marine cyanobacterial secondary metabolite biosynthesis that have utilized an intriguing diversity of methodologies. For example, the OLS and FAAR/ADO pathways were initially discovered via a comparison of genes and hydrocarbon molecules between cyanobacterial species, and further probed biochemically to determine substrate preferences and mechanisms [32, 34, 47]. Biochemical studies of interacting polyketide synthase (PKS) modules in the curacin A pathway resulted in the characterization of type II docking domains, which mediate module association, thereby facilitating molecular chain elongation [57]. Overexpression of jamaicamide A genes jamA, jamB, and jamC from Moorea producens JHB, followed by in vitro biochemical analyses, has shed light on the mechanism of alkyne formation and has provided a potential mechanism for the creation of natural product derivatives possessing an alkyne for downstream synthetic modification via click chemistry [61]. In a related species, Moorea bouillonii PNG5-198, MS2-based Molecular Networking combined with genome mining was used to uncover the molecule columbamide A and its associated biosynthetic pathway, and revealed possible secondary metabolite pathway regulation features in three closely related Moorea species [26]. Bioinformatic analysis of genome sequences identified the phormidolide biosynthetic pathway in Lepolyngbya sp.; this pathway features trans-acting acyltransferases atypical of cyanobacterial biosynthesis (unpublished data). We discuss these recent biosynthetic findings, update previous reviews on this subject [22, 23], and highlight the potential for new genome comparison technologies in the context of their ability to aid in the discovery of novel cyanobacterial natural products and biosynthetic pathways.

**Genome comparison and heterologous expression characterize cyanobacterial hydrocarbon pathways**

Cyanobacteria have been known to produce odd-chain length hydrocarbons for several decades [18, 59]. The curM gene, which generates the terminal olefin during the biosynthesis of curacin A in Moorea producens 3L, was used as a query sequence for mining the genome of Synechococcus sp. strain PCC 7002, which produces odd-chain length hydrocarbons with a terminal olefin (Fig. 1a) [34]. This research uncovered the olefin synthase (OLS) pathway, which is present in several different clades of cyanobacteria [9]. The biosynthetic portion of the pathway consists of (1) a fatty acyl-ACP ligase (FAAL) that uses ATP to activate a fatty acid of particular length, followed by linkage to the phosphopantetheine prosthetic group of an acyl carrier protein by way of an AMP-bound intermediate, and (2) a modular KS gene which shows significant homology to curM: a canonical KS domain with KS, AT, KR, and ACP domains, followed by a separate sulfotransferase and thioesterase module. The KS domain extends the preceding fatty acyl-ACP via an acetate unit and reduces the β-carbonyl to a hydroxy group. With the exception of Lepolyngbya sp. PCC 7376, all filamentous cyanobacteria which contain the OLS pathway split the FAAL-ACP

---

**Fig. 1** Hydrocarbon producing pathways in cyanobacteria. **a** Organization of the Olefin Synthase (OLS) pathway—FAAL fatty acid-ACP ligase, ACP acyl carrier protein, KS ketosynthase, AT acyltransferase, KR ketoreductase, ST sulfotransferase, TE thioesterase. **b** Organization of the fatty acid ACP reductase (FAAR)/aldehyde deformylating oxygenase (ADO) pathway

[Springer]
and KS-AT-KR-ACP-ST-TE into two open reading frames, while in unicellular and baeocystous cyanobacteria, there is only one [9]. Sulfonation of the hydroxyl group via phosphoadenosine-phosphosulfate (PAPS), followed by concerted decarboxylation and desulfation, creates the terminal double bond [17]. Knockout and upregulation experiments in the native organism alternately eliminated or increased C19-hydrocarbon production, respectively, lending credence to the predicted role of the OLS pathway in terminal olefin production [34]. In later experiments, substrate feeding to both purified curM and the OLS KS/ST/TE gene cassette indicated that the OLS KS was unable to process 3-hydroxy 5-methoxy-dodecanoyl-CoA, in contrast to curM. However, both were able to process 3-hydroxydodecanoyl CoA [32].

Comparative genomics also helped identify the fatty acid ACP reductase/aldehyde deformylating oxygenase (FAAR/ADO) pathway. By comparing a set of similar genes in cyanobacteria that either did or did not produce odd-chain length alkanes, Schirmer et al. [47] were able to identify two genes in *Synechococcus elongatus* PCC 7942 that were common to all alkane-producing strains. Heterologous expression in *E. coli* identified the genes as forming alkanes, and knock-in and knock-out experiments confirmed their role in alkane production. *In vitro* biochemical experiments determined the first gene to be a fatty acid ACP reductase that catalyzes the reduction of a fatty acid to an aldehyde (Fig. 1b). X-ray crystallography along with expression of a homolog to the second gene lent evidence that it possessed aldehyde decarbonylation activity related to metal atom-catalyzed radical reactions [49]. Further biochemical and structural studies of the ADO enzyme provided evidence for an iron-catalyzed dioxygenase mechanism as the driver for alkane generation [21, 27]. The characterization of these two pathways to hydrocarbons in cyanobacteria was an important discovery that may have consequences to understanding cyanobacterial ecology and evolution [9], and provides an intriguing potential for expression of hydrocarbon-producing enzymes in industrial microbiology applications. In summary, a combination of genome mining and traditional pathway manipulation techniques enabled the identification and activity of these unique biochemical pathways.

**Refinement of pathways via biochemical studies**

A combination of biochemistry, genome mining, and protein expression was employed to achieve a more fundamental understanding of the interaction between KS modules in the curacin A biosynthetic pathway. This pathway is remarkable in that it is in part composed of seven successive and separate KS modules containing the canonical KS-AT-ACP framework and associated tailoring enzymes, encoded by single genes for each KS module from *curG-curM*. Each module must interact consecutively with their downstream partner to ultimately produce curacin A (Fig. 2a) [8]. Docking domains, short 20–40 amino acid residue chains found on the C terminus of the upstream acyl carrier protein and downstream N-terminal ketosynthase enzyme in type-I PKS systems, foster selective and specific interaction between successive pathway modules, allowing the growing molecule to be passed from one module to the next with high fidelity (Fig. 2b) [14].

By comparing the amino acid sequences of several natural product pathways, a new type of docking domain, termed class 2, was identified in myxobacteria and cyanobacteria [57]. In contrast to previously elucidated class 1 docking domains that are found in actinobacteria, class 2 docking domains can employ a flexible structure and non-covalent interaction between ACP α-helix “coiled coils” that is distinct from class 1 docking domains, which employ helices responsible for both dimerization and interaction between ACP and KS, as determined by X-ray crystallography [6, 57]. Additionally, in class 2 docking domains, the ACP and KS can be brought close enough for the ACP to
interact with the catalytic domains of the downstream KS module. This new structural interaction was demonstrated by the expression of chimeric enzymes wherein the class 1 docking domains found on two successive modules from the pikromycin pathway were replaced by cognate class 2 docking pairs identified from the curacin A pathway. This construct enabled the processing of synthetic pikromycin pathway intermediates, as measured by mass spectrometry. The insights gained from this multidisciplinary study revealed a broad potential use of class 2 docking domains in the design of expression systems for natural products.

Alkynes present a unique molecular moiety found in bioactive bacterial and cyanobacterial natural products [37], including jamaicamide B [12] (Fig. 3a), viequeamide A [4], veraguamide [35], and carmabin A [19] (Fig. 3b). However, their biosynthetic mechanism of formation has remained uncharacterized. Comparison of the cyanobacterial biosynthetic gene clusters of carmabin A and jamaicamide A, both of which contain an alkyne in their predicted pathway starter units, pointed toward the protein products of a highly conserved set of three genes at the start of each biosynthetic pathway as being responsible for the formation of this functional group [61]. Further genome mining by Zhu et al. identified more than 80 pathways with a similar conserved set of genes [61]. Subsequent expression and purification of the JamA, B, C enzymes enabled mass spectrometry-based in vitro characterization of substrate specificity and cofactor requirements of the gene cassette. JamA was found to be an acyl-ACP synthetase with preference for 5-hexenoic acid, whereas JamB was characterized as a membrane-bound desaturase which generates the alkyne from hexenoic acid loaded onto the JamC acyl carrier protein (Fig. 3c). Additionally, creation of alkyne-tagged molecules was demonstrated in this study via heterologous expression of JamABC with a plant type III PKS, which could be useful in the creation of synthetic probes when integrated with click chemistry. Again, the combination of cyanobacterial genome mining with biochemical experiments and mass spectrometry was an effective approach to developing an understanding of this biosynthetic process, and has helped to develop a new biosynthetic tool for the creation of alkyne-based cellular probes.

**Genome mining: identifying trans-AT systems with intriguing biochemistry in cyanobacteria**

Advances in sequencing technology have led to a rapid increase in the number of sequenced cyanobacterial genomes, which has in turn increased the number of identified modular biosynthetic gene clusters for secondary metabolites [23]. Our investigation of the secondary metabolite potential of a _Leptolyngbya_ sp. collected from Sulawesi, Indonesia, led to the identification of a PKS pathway that was ultimately determined to encode for the toxic macrocyclic polyketide, phormidolide [2, 58]. This pathway possessed several unusual features from canonical PKS biosynthetic pathways and was annotated as a trans-AT system [2].

Trans-AT biosynthetic systems represent an emerging class of the biosynthetic pathway in heterotrophic bacteria with examples utilizing PKS [30] and hybrid PKS/NRPS modules [52]. In these systems, a separately encoded AT domain, typically upstream of the PKS megasynthase, loads substrates onto the modular ACPs. In some cases where a trans-AT is present, the modular AT domains, also known as “cis-”AT domains, may be absent [20] or in the case of leinamycin [52], be reduced (~160 AA) when compared to full-length cis-AT proteins. In the latter case, these KS-AT adaptor regions (AT^d) lack the conserved malonate-specific motif GHS[LVIFAM]G [11] which is present in the discrete trans-ATs, and show very low amino acid sequence identity.
to full-length cis- as well as trans-AT domains. It is possible
that the KS-AT adaptor regions are non-functional and
evolutionary artifacts [52]. In this latter case, exemplified
by leinamycin, the discrete ATs act in trans and possess the
acyltransferase activity. Trans-AT systems are hallmarked
by PKS modules that are split between megasynthases [20,
30] and the possession of KS domains that are predicted
to be inactive and are termed non-elongating KS0 domains
[20, 30, 38]. These inactive or non-elongating KSs lack the
catalytic His residue in the conserved motif HGTGT
[13, 50] that is always present in elongating KS domains.

Many trans-AT pathways encode hydroxymethylglutaryl
coenzyme A (HMG-CoA) synthase (HCS) cassettes [13,
43, 50], which introduce methyl branches in growing pol-
yketide chains at the C-3 position of the extended acetate
unit. Enoyl-CoA hydratase pairs (ECH) are components
of the aforementioned HCS cassette and have been shown
to catalyze the successive dehydration and decarboxyla-
tion of an HMG intermediate during β-branch-type meth-
ylation events. An interesting feature of the phormidolide
pathway is the presence of ECH pairs embedded within the
PKS megasynthases of the modular pathway in addition to
the ECH pair in the HCS cassette of the tailoring enzyme
suite. To date, we do not know the functional role of these
“embedded” modular ECH pairs and a full biochemical
analysis is needed to determine their role. Duplicated or
tandem ACP domains are often observed at β-branch points
within modules in the PKS megasynthetic architecture [13,
20, 31]. These tandem domains appear to be effective in
biosynthetic processes that require multienzyme events
such as β-branching methylation [15].

Trans-AT pathways have been described in a diverse
array of heterotrophic bacteria, but thus far have been rarely
found in cyanobacteria. The first cyanobacterial trans-AT
pathway was discovered from a lichen-associated Nostoc
sp. encoding the polyketide natural product nosperin [24].
Interestingly, many of the reported trans-AT pathways have
been described from bacteria that engage in symbiotic rela-
tionships with multicellular eukaryotes [43, 44], and this
appears to be the case with this Nostoc sp. Intriguingly,
an examination of genomic data from several filamen-
tous marine cyanobacteria has led to the identification of
the first trans-AT pathway described from this sub-group
of cyanobacteria, and raises interesting questions about
the distribution of these pathways as well as trends in the
construction of stereochemically complex macrocyclic pol-
yketides in these taxa. The pathway contains several of the
deviations from canonical PKS systems described above,
including discrete ATs, split modules, non-elongating KS0
domains, tandem ACP domains at β-branch points, and an
HCS cassette [2] (Fig. 4). The organism in which the path-
way was found, Leptolyngbya sp., is not known to engage
in symbiotic relationships with eukaryotes, but rather has a
strong tendency to form a biofilm-like substance. As more
cyanobacterial genome sequence information becomes
available, it will be interesting to observe if trans-AT path-
ways are of a more common occurrence and to further
understand the evolution of these pathways and the biologi-
cal role of the products they encode.

Discovery via genomic comparison and MS2-based
molecular networking

Combining new techniques like mass spectrometric profil-
ing and bioinformatics can help identify interesting targets
for chemical isolation and provide orthogonal means for
natural product discovery. Mass spectrometric molecular
networking [56, 60] is a useful tool for the identification of
structural analogs within a chemical extract. Related families
of natural products can be detected with algorithms that con-
vert MS-profiled metabolomes to maps of structurally linked
molecules on the basis of common fragment ions [55]. In
addition, bioinformatic analyses of the biosynthetic gene
clusters encoded in DNA sequence data can help to evalu-
ate the uniqueness of compounds present within a genome
[3, 33]. In a recent study, the mass spectrometric metabolic
profiles of extracts from three marine cyanobacterial strains
M. producens 3L, M. producens JHB and M. bouillonii
PNG51905-8 were compared with respect to the biosyn-
thetic pathways predicted from their sequenced genomes
[26]. These same three extracts were subject to MS2-based
molecular networking (Fig. 5). In addition to the well-stud-
ied cyanobacterial products curacin A, jamaicamide A and
apratoxin A, a new class of chlorinated compounds, the
columbamides, was discovered in relatively high yields in
the extract of M. bouillonii PNG. Comparison of the known
biosynthetic gene clusters from the Moorea strains led to
the discovery of a putative regulatory gene, encoding a ser-
ine–histidine kinase, next to a biosynthetic gene cluster of
unknown function in the M. bouillonii PNG genome. This
same gene is adjacent to the curacin A and jamaicamide A
biosynthetic gene clusters in M. producens 3L and JHB and,
therefore, may be involved in the observed robust constitu-
tive expression of these compounds. Subsequent isolation
and structure elucidation primarily by NMR led to the dis-
covery of the columbamides, alkyl amides with cannabino-
mimetic activity. Based on the bioinformatics analysis, nine
enzymes are involved in the assembly of the columbamides.
First, two novel halogenases are proposed to insert chlor-
ine atoms at the terminal carbon and ω-7 positions of an
acyl carrier protein-bound dodecanoic acid. This intermedi-
ate is then extended twice via the addition of acetate with
two polyketide synthase modules, and then serine is added
via an NRPS module. The molecule is ultimately released from the megasynthase through a reductase mechanism, and the resulting primary hydroxy group is acetylated to yield columbamide A. Figure 6 shows the proposed biosynthetic pathway and the chemical structures of the columbamides. The combined use of molecular networking and genomic comparison enabled the discovery of these unique molecules; integration of these techniques provides a powerful framework for novel bioactive metabolite discovery [26, 56].
Development of a cyanobacterial reference genome to foster more accurate genome assembly

Despite the enormous biosynthetic potential of tropical filamentous cyanobacteria, extensive genome mining has remained elusive due to a lack of sequences, which themselves are sparse due to difficulties in culturing these organisms. Currently available molecular biology tools and heterologous expression systems are typically designed for unicellular, heterotrophic bacteria. However, the few available genome sequences for tropical filamentous cyanobacteria are rich with biosynthetic gene clusters. For example, in a comparison of 126 different cyanobacterial genomes, *Moorea producens* 3L contained 14 PKS, NRPS, or PKS-NRPS biosynthetic gene clusters, the third highest total in the study [7, 48]. However, only four of these clusters encode known products, one cluster is shared between several different genomes but is not elucidated, and nine cannot be connected by bioinformatics to any known metabolite. The genome of *M. producens* 3L currently contains 161 scaffolds and cannot be completed with the available data due to the lack of a closely related reference sequence; this situation is mirrored for several unpublished genome sequences as well. The closest completed genome to serve as a potential reference to *Moorea* species is from *Microcoleus* sp. PCC 7113. This latter genome contains one chromosome and 8 plasmids comprising 7.95 megabases (unpublished data, GenBank access number CP003630.1).

*In silico* predictions of phylogenetic homology using the Genome-to-Genome Distance Calculator [1] indicate that *Microcoleus* sp. PCC 7113 has only a 12.7–20.7 % similarity index to *M. producens* 3L. This value is very low compared to the similarity of *M. producens* 3L to *Moorea* sp. JHB (58.5–66.8 %), and *Moorea* sp. PAL 15AUG08-1 (48–57.7 %).

Development of a completed reference genome of a single *Moorea* species is crucial to future genome comparisons, genome-driven discovery of novel natural products, and a more complete understanding of the evolutionary aspects of this interesting genus. In this regard, the genome of a *Moorea* sp. PAL15AUG08-1 was sequenced with both Pacific Biosciences (PacBio) and Illumina MiSeq platforms, and is thus the most suitable candidate to be a reference as it is the most complete marine filamentous cyanobacterial genome to date (unpublished). Additionally, *Moorea* sp. PAL15AUG08-1 is considered a “super-producer”, as it is estimated that 19.8 % of its genome is devoted to natural products biosynthesis; this is nearly four times the average cyanobacterial genome of 5 % encoding for natural products [48]. A total of 43 biosynthetic gene clusters have been identified, including 9 NRPS, 5 PKS and 12 hybrid NRPS/PKS clusters; this is 5 times the average number of NRPS/PKS clusters in cyanobacterial genomes (Fig. 7) [7]. However, only two known compounds have been isolated from this strain to date, palmyramide A [53] and curacin D [29], indicating that there is considerable potential for discovery of new compounds from this strain.

One example of a 50 Kb hybrid NRPS/PKS cluster that is under current investigation encodes for a natural product with significant homology to the biosynthetic pathway for the antifungal lipopeptide hectochlorin A [45] and cytotoxic metabolite lyngbyabellin A [28] (Fig. 8a, b). The first predicted biosynthetic gene of this cluster has 97 % identity with *hctA* and encodes for a fatty acyl-ACP ligase. However, it is translated in the opposite direction of the remainder of the cluster, possibly indicating an inversion in this ORF relative to the other clusters. The homologous gene in the lyngbyabellin pathway is smaller, which explains the lower identity with the first gene in the PAL cluster, but still shows significant conservation in the regions that overlap. The second gene in this new cluster is highly homologous to the second genes in the hectochlorin and lyngbyabellin gene clusters, and is presumably responsible for the formation of the geminal-dichloro group in these molecules. One disparity is that *hctC*, a transposase of unknown function, is not present in the new cluster; however, this gene is most likely not related to compound biosynthesis. However, a similar putative transposase gene is found in the lyngbyabellin A gene cluster, and suggests that the new cluster is more distant in its evolutionary history. The third

![Fig. 7 Comparison between PKS/NRPS clusters of 4 *Moorea* strains with 7 (out of 126) additional cyanobacteria reviewed by Calteau et al., 2014 [7]. The 7 genomes were selected to compare two low secondary metabolite producers (*Synechococcus* genomes), two average secondary metabolite producers (*Microcoleus* sp. and *Gloeocapsa* sp.) and three high secondary metabolite producers (*Nostoc punctiforme*, *C. stagnela* and *Fischerella* sp.). With the exception of *Fischerella* sp. PCC9339, only final scaffolds were considered, as incomplete genomes tend to present fragmented biosynthetic gene clusters](image-url)
The biosynthetic gene in the new cluster is 97.2% identical to the lynD gene, a PKS module responsible for a single acetate extension, along with a C-methylation from S-adenosyl methionine (sAM). Curiously, in the formation of lyngbyabellin A, a double methylation occurs at this juncture instead of the single methylation observed in hectochlorin. The biochemical basis for this double methylation from a single cMT domain is currently not understood. However, due to the significant conservation between amino acid residues in the lyngbyabellin A cMT domain and the PAL cluster cMT domain, it is reasonable to predict that a double methylation occurs at this step in the new cluster. The fourth gene in the cluster is 94% identical to the hctE gene and is composed of one NRPS/PKS module and one NRPS module (C-A-KR-PCP and Heterocyclization(Cy)-A-Oxidase(Ox)-ACP). In the first amino acid module, online A-domain prediction software (NRPSPredictor2) is unable to discern which amino acid (or keto acid) is activated; however, given its 93% identity with the first A domain in the hctE gene, it is reasonable to predict that it may also activate 2-hydroxyisovaleric acid. In this regard, a KR domain is observed between the A and PCP domains, similar to those found in the hctE and lynE genes, suggesting the initial activation of an α-keto acid with subsequent in situ reduction to an α-hydroxy acid. The second module in this ORF has the same structure of the second module of the hctE gene, and is thus predicted to activate and then cyclize a cysteine residue after its incorporation. The fifth gene in this cluster does not share significant homology to any part of the hectochlorin or lyngbyabellin A gene clusters. This gene has 3 NRPS modules that are predicted to incorporate glycine, valine and cysteine; the latter cysteine is further predicted to be heterocyclized. This is followed by a thioesterase that may catalyze the hydrolysis from the biosynthetic assembly line. Lastly, the sixth gene encodes for a putative P450 monooxygenase that may be responsible for oxidation events during the biosynthesis of this compound, analogous to similar motifs involved in the biosynthesis of hectochlorin and lyngbyabellin A. The final structure of this metabolite is predicted to be similar to lyngbyabellin B (Fig. 9) which is produced by M. producens; however, the biosynthesis of lyngbyabellin B remains to be elucidated [36]. Analysis of LCMS data and molecular
networks generated from *M. producens* PAL 15AUG08-1 chemical extracts has yet to uncover any lyngbyabellin analogs and, hence, the compound appears not to be produced under its current conditions of laboratory culture.

**Conclusion**

Recent progress in understanding marine cyanobacterial natural product pathways has required a broad range of methods to explore questions concerning their biosynthesis. While the development of more robust heterologous expression tools would usher a transformative approach to studying cyanobacterial natural products, it is clear that the combination of genome analysis, mass spectrometry, and biochemical studies has provided effective tools for understanding their fundamental pathways. This combination of techniques will most certainly be used, along with improvements in expression tools, to probe the molecules and biosynthetic pathways in cultured cyanobacterial specimens as well as metagenomic and environmental samples. Advances in mass spectrometry visualization have enabled efficient identification of molecular analogs, in particular natural product families. Similarly, development of a complete reference genome for a filamentous marine cyanobacterium will facilitate the more rapid mapping of genomic data and hence the discovery of new biosynthetic gene clusters.

**Acknowledgments**

This work was supported by NIH Grants CA108874 and GM107550-01. Karin Kleigrewe was supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD). Tiago Ferreira Leão is supported by the Ministry of Education of Brazil, D.F., CAPES fellowship, no. 13425137.

**References**

genomics unravel the diversity of secondary metabolism in Cyanobacteria. BMC Genom 15:977


