J Ind Microbiol Biotechnol (2007) 34:443–456
DOI 10.1007/s10295-007-0216-6

ORIGINAL PAPER

Genetic analysis of polyketide synthase and peptide synthetase genes in cyanobacteria as a mining tool for secondary metabolites

Martin E. Barrios-Llerena · Adam M. Burja · Phillip C. Wright

Received: 17 October 2006 / Accepted: 6 March 2007 / Published online: 25 April 2007

Abstract Molecular screening using degenerate PCR to determine the presence of secondary metabolite genes in cyanobacteria was performed. This revealed 18 NRPS and 19 PKS genes in the 21 new cyanobacterial strains examined, representing three families of cyanobacteria (*Nostocales*, *Chroococales* and *Oscillatoriales*). A BLAST analysis shows that these genes have similarities to known cyanobacterial natural products. Analysis of the NRPS adenylation domain indicates the presence of novel features previously ascribed to both proteobacteria and cyanobacteria. Furthermore, binding-pocket predictions reveal diversity in the amino acids used during the biosynthesis of compounds. A similar analysis of the PKS ketosynthase domain shows significant structural diversity and their presence in both mixed modules with NRPS domains and individually as part of a PKS module. We have been able to classify the NRPS genes on the basis of their binding-pockets. Further, we show how this data can be used to begin to link structure to function by an analysis of the compounds Scyptolin A and Hofmannolin from *Scytonema* sp. PCC 7110.

Keywords Peptide synthetases · Polyketide synthases · Cyanobacteria · Ketosynthase domain · Adenylation domain · Natural products

Electronic supplementary material The online version of this article (doi:10.1007/s10295-007-0216-6) contains supplementary material, which is available to authorized users.

M. E. Barrios-Llerena · P. C. Wright (✉)
Biological and Environmental Systems Group,
Department of Chemical and Process Engineering,
University of Sheffield, Mappin Street, Sheffield S1 3JD, UK
e-mail: P.C.Wright@Sheffield.ac.uk

A. M. Burja
Metabolic Engineering and Fermentation Group,
Ocean Nutrition Canada, 101 Research Drive,
Dartmouth, NS B2Y 4T6, Canada

Present Address:
M. E. Barrios-Llerena
Marie Curie Excellence Team Pathogen Habitats,
Institute of Immunology and Infection Research,
University of Edinburgh, West Mains Road,
Edinburgh EH9 3JT, UK

Introduction

In recent years, as the number and variety of compounds isolated from cyanobacteria increases, so too has interest in cyanobacteria as potential sources of biologically active secondary metabolites [5, 41, 56, 58]. These compounds demonstrate a diverse range of biological activities and chemical structures, including novel cyclic and linear lipopeptides, fatty acids, alkaloids and other organic chemicals [4, 23, 26, 27, 46, 47, 49, 67]. Many of these have potential pharmaceutical, nutraceutical, agricultural, and other applications [2, 3, 5, 60].

To date, the majority of bioactive metabolites isolated from cyanobacteria have either been polyketides, non-ribosomal peptides, or a hybrid of the two. This is a feature of their biosynthetic complexity, which they share with other bacteria such as actinobacteria [15, 38]. Type I polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) are exclusively involved in the biosynthesis of these natural products [48, 71]. This common attribute is also a characteristic of several important natural products that are currently in pre-clinical pharmaceutical development (e.g. cryptophycin and epothilone), or use (bleomycin, rapamycin and FK506) [41, 58].

Non-ribosomal peptide synthetases possess a modular structure, with each module functioning as a building block responsible for the incorporation and/or modification of a single amino acid unit. The order and number of modules on a NRPS protein dictates the corresponding order and number of amino acids in the peptide product. A typical NRPS module consists minimally of an adenylation (A) domain, responsible for amino acid activation; a thiolation (T) domain (also known as a peptidyl carrier protein), for thioesterification of the activated amino acid; and a condensation (C) domain, for transpeptidation between the aligned peptidyl and amino acyl thioesters resulting in the elongation of a growing peptide chain and the evolution of a water molecule [7, 32, 72]. Furthermore, NRPS and PKS may be found as mixed entities, resulting in molecules containing elements of both, an occurrence well documented in the literature (e.g. curacin A, microcystin, and barbamide). The biosynthesis of these hybrid molecules illustrates how NRPS/PKS-derived systems have evolved to be functionally compatible.

Polyketides on the other hand, can be devided into three classes. Type I PKSs being multifunctional enzymes, organized into modules, responsible for one cycle of chain elongation and are currently representative of all cyanobacterial PKSs isolated to date [62]. Type II PKSs being cyclic multienzyme complexes, typically involved in the biosynthesis of aromatic antibiotics in other bacteria. Finally, type III PKSs, are homodimeric condensation enzymes responsible for variety of unusual flavonoid and chalcone compounds. Typically, type I PKSs minimally contains three domains per module, a β-ketosynthase (KS), a acyltransferase (AT), and a acyl carrier protein (ACP), that select, activate, and catalyze a decarboxylic Claisen condensation between the extender unit and the polyketide chain, generating a β-ketoacyl-*S*-ACP intermediate. Optional domains may also be found between AT and ACP, which carry out a variable set of reductive modifications of the β-keto group before the next round of chain extension. The order of modules in the PKS enzymes dictates the sequence of biosynthetic events, and the variation of domains within the modules affords the structural diversity observed in the resultant natural products [55, 63].

Molecular approaches have been used to successfully identifying gene clusters involved in the biosynthesis of a variety of cyanobacterial secondary metabolites. For example, microcystin from *Microcystis aeruginosa* [16, 48, 71], anabaenopeptilide from *Anabaena flos-aquae* [51], and lyngbyatoxin, curacin A, the jamaicamides and barbamide, from *Lyngbya majuscula* [10, 11, 19, 20]. Alternatively, Christiansen et al. [14] using genetic PCR-based, screening techniques, reported the presence of NRPS genes in 75% of 146 axenic strains from the Pasteur culture collection (PCC) of cyanobacteria [50], yet only five strains were sequenced and no further analysis was carried out. Conversely, the presence of type I KS domains has been reported by Neilan et al. [43] in a range of bacterial strains including cyanobacteria. In that work, a similar genetic PCR-base screening technique was carried out on both culture collection and environmental samples isolated from Shark Bay, Western Australia. Analysis of the resultant protein sequences showed grouping into two functional groups, the first comprised KS domains that use acyl-CoA's as a starter or extender unit, and the second represents those isolates from mixed NRPS/PKS systems [43]. Recently, Ehrenreich et al. [21] using similar PCR screening techniques, conducted a combined NRPS and PKS study whereby 20 marine and freshwater cyanobacteria were probed and compared to cyanobacterial genomes. The presence of NRPS A-domains and PKS KS-domains was subsequently noted in 65 and 90% of the cyanobacteria studied, respectively.

Here, we conduct a combined NRPS and PKS PCR-based screening study, in order to begin to isolate, sequence and analysize the genes responsible for the production of secondary metabolites from 21 cyanobacterial strains. This information will then be used to assess the potential of these strains to produce currently unknown secondary metabolites. Ultimately these data will result in the development of experiments to identify and implement fermentation techniques that will induce the production of these secondary metabolites. (Note: The chemical structures of secondary metabolites mentioned throughout this manuscript are shown in Fig. 1. The number in parentheses references compounds mentioned in the text).

Materials and methods

Cyanobacterial strains and culture conditions

Freshwater and marine cyanobacterial strains were obtained from the PCC, The Culture Collection of Algae at the University of Texas (UTEX), The American Type Culture Collection (ATCC) and The Culture Collection of Algae and Protozoa (CCAP). These include: *Anabaena cylindrica* UTEX 629, *Anabaena* sp. PCC 7120, *Anabaena* sp. PCC 9109, *Anabaena ambigua* CCAP 1403/7, *Anabaena variabilis* ATCC 29413, *Gloeobacter violaceus* PCC 7421, *Gloeothece* sp. ATCC 27152, *Leptolyngbya* sp. PCC 6703, *Leptolyngbya* sp. PCC 7104, *Leptolyngbya* sp. PCC 7410, *Plectonema boryanum* PCC 73110, *Lyngbya aestuarii* PCC 7419, *Lyngbya majuscula* CCAP 1446/4, *Microcoleus* sp. PCC 8701, *Nodularia harveyana* CCAP 1452/1, *Nostoc punctiforme* PCC 73102, *Oscillatoria* sp. CCAP 1459/13, *Oscillatoria* sp. CCAP 1459/26, *Oscillatoria* sp. PCC 6506, *Planktothrix* sp. PCC 7811, *Oscillatoria* sp. PCC 7515, *Plectonema terebrans* CCAP 1463/4, *Scytonema* sp. PCC 7110, and *Tolypothrix* sp. PCC 7601. All strains were grown

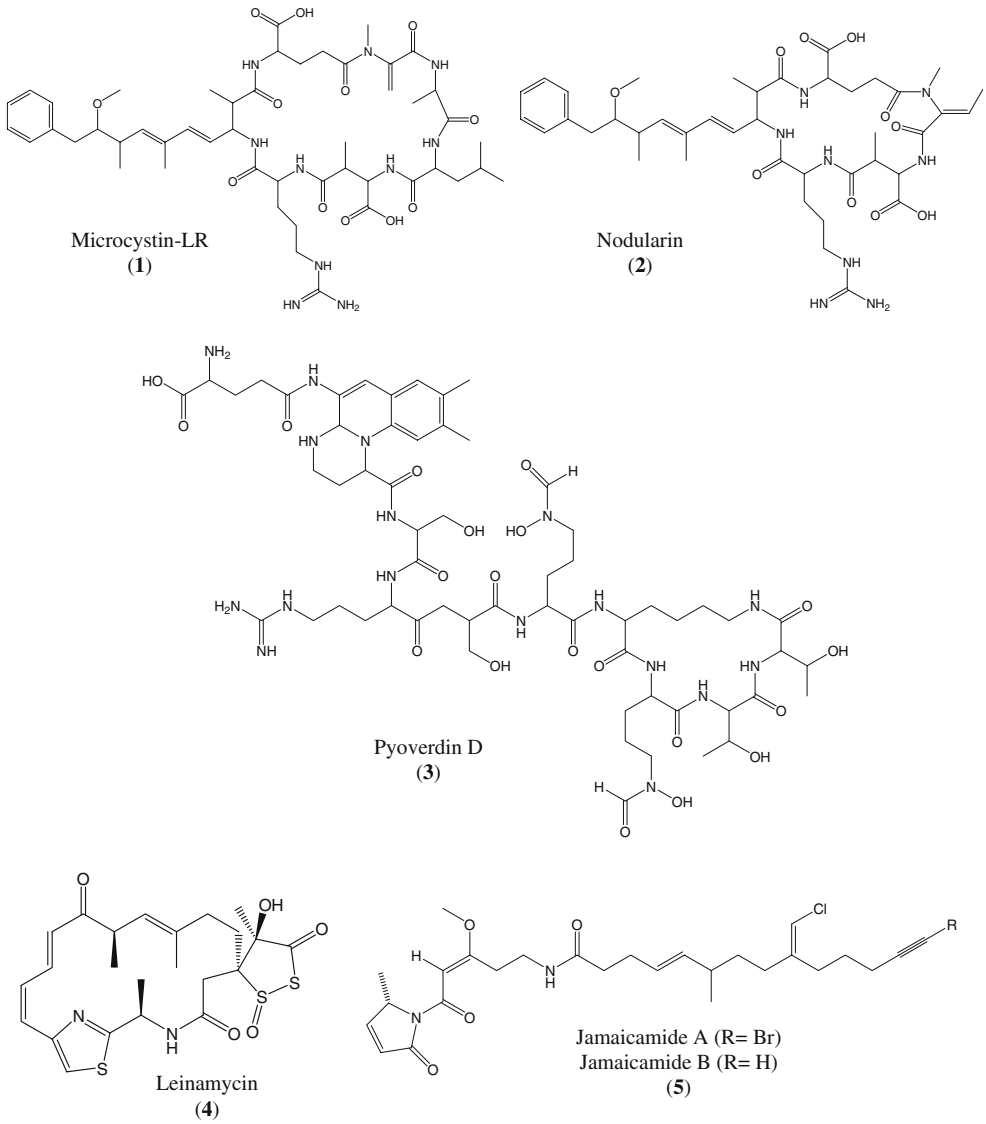


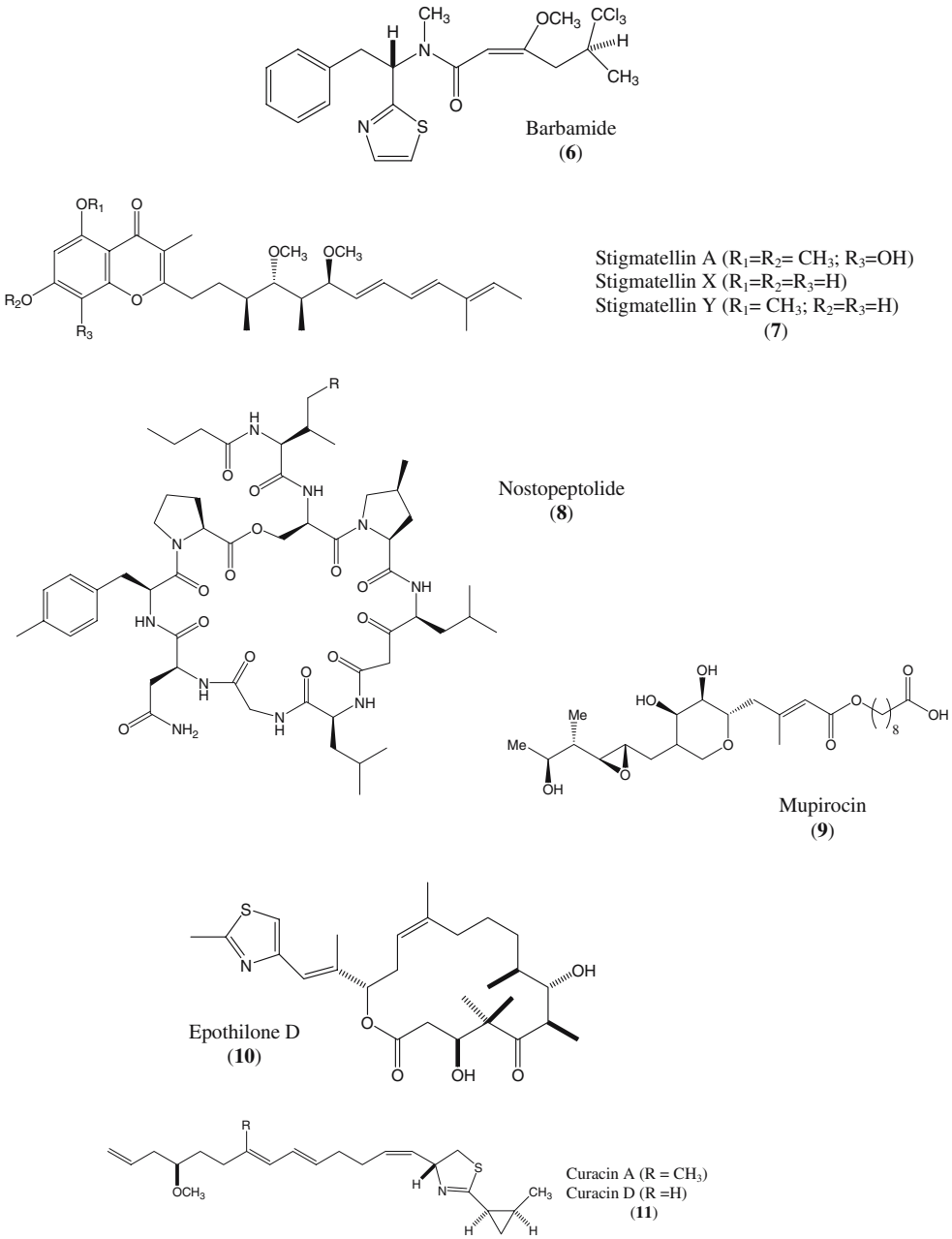
Fig. 1 Structure of selected secondary metabolites produced by cyanobacteria and other bacteria. Number in *parentheses* are used when the compound is mentioned throughout the text

in appropriate growth media as suggested by the relevant culture collections at 25°C under a 1212 hours light:dark cycle. Illumination was via fluorescent, plant growth lamps (Gro-Lux lamps, Osram Sylvania, Danvers, MA, USA) with a photosynthetic photon flux density (PPFD) of 50 µEinstein m^{-2} s^{-1} (determined using a QSL-2000 quantum sensor Biospherical Instruments Inc. San Diego, CA, USA), in a plant growth chamber (MRL-350HT, Sanyo Gallenkamp, Loughborough, UK) set to 60% humidity.

Sample preparation and DNA purification

Genomic DNA was extracted using the method of Tamagnini et al. [66] with slight modifications. Specifically, a pelleted 20 ml aliquot of mid- to late-exponential phase cultures were washed twice and resuspended with TE buffer (50 mM Tris–HCl, pH 8.0, 10 mM EDTA, pH 8.0), containing 800 mM ammonium acetate. Cells were broken by mechanical disruption using glass beads, with the mixture then kept on ice for 10 min, centrifuged, and the aqueous phase extracted with chloroform. The nucleic acids were precipitated with ethanol [53] and redissolved overnight at 4°C in 100 µl of TE buffer (10 mM Tris–HCl, pH 7.4; 0.1 mM EDTA, pH 8.0) or DNA-free sterilized water. Genomic DNA from *Gloeothece* sp. ATCC 27152 and *Oscillatoria* sp. PCC 7515 was isolated as above, with the exception of cellular disruption being performed using liquid nitrogen in a mortar and pestle. DNA extract from tough strains (*Leptolyngbya* sp. PCC 7104, *Lyngbya majuscula* CCAP 1446/4, *Plectonema terebrans* CCAP 1463/4, *Microcoleus* sp. PCC 8701, *Anabaena* sp. PCC 9109, *Oscillatoria* sp CCAP 1459/13 and *Oscillatoria* sp CCAP 1459/26) was extracted with

Fig. 1 continued



the DNeasy Plant Mini Kit (QIAGEN Ltd., UK), following the manufacturer's instructions.

PCR amplification and cloning

Amplification of NRPS A-domain and PKS KS-domain regions was performed using the degenerate oligonucleotide primer pairs MTF2/MTR [16] and DKF/DKR [44], respectively. The PCR thermal cycle included an initial denaturation at 94°C for 4 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing for 30 s at 50 and 60°C respectively for primer pairs MTF2/MTR and DKF/DKR, followed by extension at 72°C for 1 min, with a final extension at 72°C for 7 min. The amplification reaction was performed using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA). PCR products were visualised on agarose gels using standard electrophoresis protocols using 1×TAE buffer [53], purified with the QIAquickspin PCR purification kit (Qiagen, Hilden, Germany) and plasmids cloned using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. Sequencing was subsequently carried out by MWG Biotechnology.

Sequence analysis

Nucleotide sequences were analysed using CLUSTAL W [70] for multiple sequence alignments. Translated protein

sequences were compared with those in databases through the National Center for Biotechnology Information (US National Institutes of Health, Bethesda, MD) using the BLASTx program [1]. Cyanobacterial NRPS and PKS protein sequences were aligned using the PAM matrix and the multiple sequence alignment tool from CLUSTAL W and MEGA package [34]. Identification of the predicted amino acid activated by a specific unknown NRPS A module was performed using software located at http://www.tigr.org/jravel/nrps [9]. The activation of amino acids by the identified A-domain motif will indicate their presence, unmodified or modified in the final natural product structure. These translated protein sequences were subsequently compared to the NRPS of known microbial and plant natural products using software located at http://www.npbiogene.com/.

Nucleotide sequences

Nucleotide sequences of the putative non-ribosomal peptide synthetase A-domain and polyketide synthase KS genes reported in this study were deposited in GenBank under accession numbers AY768427–AY768458, AY768460–AY768463, AY768474–AY768516 and DQ439637–DQ439644.

Results

Identification of adenylation domains (NRPS)

Putative peptide synthetase gene sequences, confirmed via sequencing and BLASTx analysis, were obtained (~1 kbp size) from 18 cyanobacterial strains (75%) screened (Table 1). *Gloeobacter violaceus* PCC 7421 (Chroococcales) and *Leptolyngbya* sp. PCC 6703 and 7104; *Lyngbya* sp. PCC 7419; *Oscillatoria* sp. PCC 7515 and *Plectonema terebrans* CCAP 1463/4 (Oscillatoriales), gave no PCR product under the conditions tested, and potentially lack the ability, as far as we can tell at this point, to non-ribosomally, synthesize peptides. From these 18 cyanobacterial strains, 33 unique A-domains were subsequently isolated and sequenced. BLASTx sequence analysis of these clones showed varying similarities (43 and 97% identity) to other known cyanobacterial NRPS sequences (Table 2).

BLASTp analysis of these 33 clones showed that 25 were homologous to genes that synthesize compounds through cyanobacterial NRPSs, with the remaining derived from bacterial NRPS mechanics. Interestingly, 6 of the 18 strains screened possessed only 1 A-domain, while the remaining possessed 2 or more. The most prolific strains being *Anabaena ambigua* CCAP 1403/7, *Planktothrix* sp. PCC 7811 and *Microcoleus* sp. PCC 8701 harbouring five and four A-domains, respectively (Table 2). These A-domains can result in the production of one compound, or multiple independent compounds.

Finally, clustering of these NRPS sequences using Neighbour-joining techniques, illustrates a lack of taxonomic affiliations between cyanobacteria and A-domains (Fig. 2). For example, only clone three from *Planktothrix* sp. PCC 7811 showed homology (55% of identity with JamO from *Lyngbya majuscula*) to a peptide synthetase from a member of the same cyanobacterial family. Rather, as seen elsewhere, clustering was possible according to substrate-binding pocket specificity, with main branches diverging according to specific amino acid activation [38] (Fig. 2). These include valine (Val), glycine (Gly), cysteine (Cys), serine (Ser), threonine (Thr), proline (Pro), glutamic

Table 1 Cyanobacterial strains analyzed in this study by degenerate peptide synthetase PCR (NRPS) and degenerate ketosynthase PCR (PKS)

Strain	NRPS	PKS	Strain	NRPS	PKS
Anabaena cylindrica UTEX 629	+	–	*Leptolyngbya* sp. PCC 7410	+	+
Anabaena sp. PCC 9109	+	+	*Lyngbya* sp. CCAP 1446/5	+	+
Anabaena ambigua CCAP 1403/7	+	+	*Lyngbya aestuarii* PCC 7419	–	–
***Anabaena variabilis* ATCC 29413**[a]	**+**	**+**	*Lyngbya majuscula* CCAP 1446/4	+	+
***Nostoc punctiforme* PCC 73102**[a]	**+**	**+**	*Microcoleus* sp. PCC 8701	+	+
Nodularia harveyana CCAP 1452/1	+	+	*Planktothrix* sp. PCC 7811	+	+
Scytonema sp. PCC 7110	+	+	*Oscillatoria* sp. PCC 7515	–	+
Tolypothrix sp. PCC 7601	+	+	*Oscillatoria* sp. PCC 6506	+	+
***Gloeobacter violaceus* PCC 7421**[a]	–	**+**	*Oscillatoria* sp. CCAP 1459/13	+	+
Gloeothece sp. ATCC 27152	+	+	*Oscillatoria* sp. CCAP 1459/26	+	+
Leptolyngbya sp. PCC 6703	–	–	*Plectonema boryanum* PCC 73110	+	+
Leptolyngbya sp. PCC 7104	–	–	*Plectonema terebrans* CCAP 1463/4	–	–

Note: (+) indicates a positive PCR result (i.e. production of a PCR-amplified DNA product); (–) indicates a negative PCR result

[a] Cyanobacterial strains highlighted in bold have their genomes completely sequenced and were used as positive controls

Table 2 Cyanobacterial NRPS sequences analyzed using the BLASTp tool for natural product biosynthesis

Strain	Clone	Accession number[a]	Domain	Compound[e]	*E* value	Signature sequence[c]	Predicted amino acid[d]
Anabaena cylindrica UTEX 629	Clone-8	AF484556	A2	Leinamycin	2×10^{-96}	DLYSFSLV	BacA-M2-Cys
	Clone-9	AY212249	A6	Microcystin	1×10^{-73}	DIWHVSLI	PvdD-M1-Ser
***Anabaena* sp. PCC 7120**[b]	**17131676**	**AY212249**	**A7**	**Microcystin**	$\mathbf{1 \times 10^{-112}}$	**DLFNNALT**	**SafB-M2-Gly**
Anabaena sp. PCC 9109	Clone-4	AY210783	NdaB-A8	Nodularin	8×10^{-88}	DLKNFGVG	FxbC-M1-5hOrn
	Clone-7	CAC01606	A7	anabaenopeptolide	1×10^{-121}	DAFFLGVT	AdpD-M1-Ile
Anabaena ambigua CCAP 1403/7	Clone-2	AF007865	A3	Bacitracin	2×10^{-77}	DAWYCGNV	McyB-M1-Leu
	Clone-3	AF516145	barG-A1	Barbamide	1×10^{-112}	DAFTIAAV	BacC-M2-Phe
	Clone-4	AB050629	A7	Iturin	2×10^{-94}	DVWSFSLV	EntF-M1-Ser
	Clone-5	AF007865	A3	Bacitracin	2×10^{-72}	DALYIVNV	CepA-M1-Leu
	Clone-8	AF204805	nosC-A6	Nostopeptolide	1×10^{-110}	DILALGMI	Cda2-M2-Gly
***Anabaena variabilis* ATCC 29413**[b]	**Clone-4**	**AF204805**	**nosA-A2**	**Nostopeptolide**	$\mathbf{1 \times 10^{-138}}$	**DVWHISLI**	**NosA-M2-Ser**
***Nostoc punctiforme* PCC 73102**[b]	**30581770**	**AJ566197**	**lgrB-A5**	**pentadecapeptide**	$\mathbf{1 \times 10^{-132}}$	**DLFNNALT**	**SafB-M2-Gly**
	30581848	**AY522504**	**JamO-A7**	**Jamaicamide**	$\mathbf{1 \times 10^{-131}}$	**DLFNNALT**	**SafB-M2-Gly**
Nodularia harveyana CCAP 1452/1	Clone-2	AJ566197	lgrB-A6	pentadecapeptide	1×10^{-105}	DALWIGGT	GrsB-M2-Val
Scytonema sp. PCC 7110	Clone-2	AJ441056	A10	Microcystin	2×10^{-64}	VDWITSL-	β-Ala
	Clone-6	AJ566197	lgrB-A6	pentadecapeptide	1×10^{-106}	DAFWLGGT	GrsB-M2-Val
Tolypothrix sp. PCC 7601	Clone-3	AF516145	barG-A1	Barbamide	1×10^{-83}	DVEAIGSI	BacB-M1-Lys
	Clone-5	AY210783	NdaB-A8	Nodularin	1×10^{-74}	DSASIAEV	BacC-M3-His
Gloeothece sp. ATCC 27152	Clone-1	AY212249	A7	Microcystin	1×10^{-110}	DLFN-ALT	PchE-M1-Cys
Leptolyngbya sp. PCC 7410	Clone-3	AF204805	nosC-A7	Nostopeptolide	1×10^{-88}	DATQVGEV	NosC-M3-Asp/Asn
	Clone-7	AF204805	nosC-A6	Nostopeptolide	1×10^{-89}	DVFNLGLI	CepB-M1-HPG[e]
Lyngbya sp. CCAP 1446/5	Clone-1	AF204805	nosC-A6	Nostopeptolide	1×10^{-66}	DILFIGVV	BacA-M1-Ile
	Clone-3	NP_251092	A5	pyoverdin D	1×10^{-126}	DAWHVSLI	PvdD-M1-Ser
Lyngbya majuscula CCAP 1446/4	Clone-4	AF183408	A6	Microcystin	1×10^{-125}	DVWHFSLI	McyA-M1-Ser
	Clone-6	AY212249	A7	Microcystin	1×10^{-115}	DLFNNALT	SafB-M2-Gly
Microcoleus sp. PCC 8701	Clone-1	AF204805	nosC-A7	Nostopeptolide	1×10^{-99}	DATKIGEV	NosC-M3-Asp/Asn
	Clone-4	CAC01606	A7	anabaenopeptolide	2×10^{-48}	DAMHVGGF	FenC-M1-Glu
	Clone-11.3	CAC01603	A2	anabaenopeptolide	1×10^{-140}	DFWNIGMV	FxbC-M2-Thr
	Clone-11.4	AJ269505	A4	anabaenopeptolide	1×10^{-188}	DIENIGGV	FxbC-M1-5hOrn
Planktothrix sp. PCC 7811	Clone-1	AF204805	nosA-A1	Nostopeptolide	1×10^{-148}	DAFFLGVT	AdpD-M1-Ile
	Clone-2	AJ269505	A4	anabaenopeptolide	1×10^{-143}	DVENAGVV	AdpB-M2-Ahp[f]
	Clone-3	AY522504	JamO-A7	Jamaicamide	1×10^{-100}	DLFNNAL-	SafB-M2-Gly
	Clone-6	AJ566197	lgrC-A8	pentadecapeptide	9×10^{-84}	DALWIGGV	GrsB-M2-Val
Oscillatoria sp. PCC 6506	Clone-1	AF204805	nosA-A3	Nostopeptolide	3×10^{-91}	DFHFITHD	Pps4-M1-Pro
Oscillatoria sp. CCAP 1459/13	Clone-11.4	AL939115	cdaPS1-A1	Antibiotic CDA	1.7	unknown	unknown
Oscillatoria sp. CCAP 1459/26	Clone-4	NP_251092	A5	pyoverdin D	1×10^{-134}	DVWHVSLI	PvdD-M1-Ser
Plectonema boryanum PCC 73110	Clone-1	AF204805	nosC-A7	Nostopeptolide	1×10^{-135}	DATQVGEV	NosC-M3-Asp/Asn

HPG p-hydroxy phenyl glycine, *Ahp* 3-amino-6-hydroxy-2-piperidone

[a] Accession numbers corresponding to the NCBI website for amino acid sequences

[b] Sequences highlighted in bold were obtained from GenBank (genomes)

[c] Eight variable amino acids of the signature sequences determined as described by Stachelhaus et al. [61]

[d] Nomenclature of the reference compounds as described by Challis et al. [9]

[e] The compound here is for illustrative purposes, in that the domain is similar. This does not imply the strain makes this exact compound

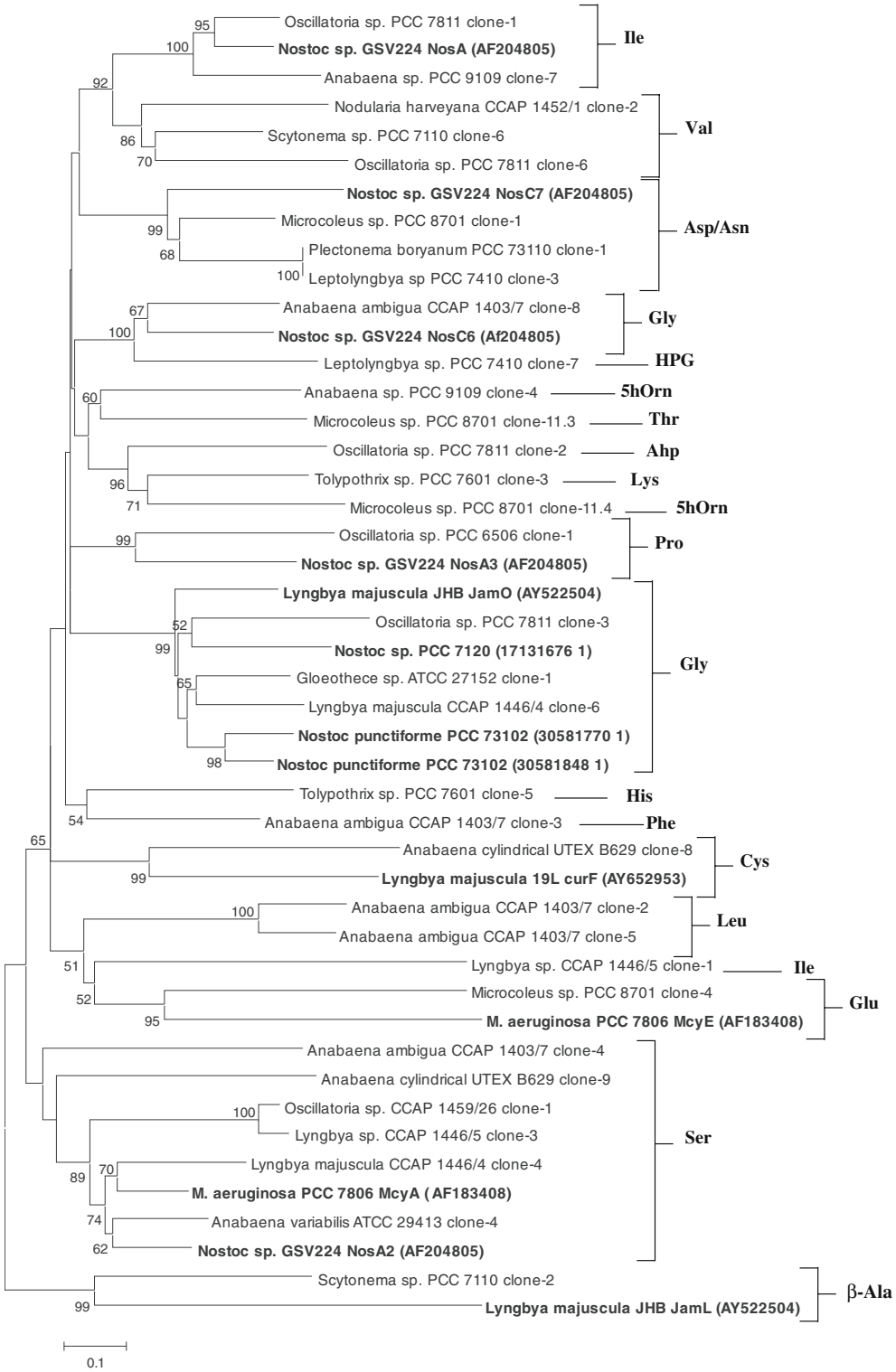


Fig. 2 Phylogenetic analysis of A-domains from non-ribosomal peptide synthetase genes. Sequences obtained from GenBank are given in *boldface* and their accession numbers in *parentheses*. Sequences were aligned using the Clustal W program. Divergence between amino acid sequences was calculated using a PAM matrix and the tree constructed using the neighbor-joining method in the MEGA3 software package (version 3.0). Bootstrap values were calculated from 1,000 bootstrap resamplings

acid (Glu), aspartic acid (Asp), asparagine (Asn), leucine (Leu), isoleucine (Ile), alanine (Ala), phenylalanine (Phe), histidine (His) and lysine (Lys) and non-protein amino acids such as 5-hydroxy-*L*-Ornithine (5 hOrn), 3-amino-6-hydroxy-2-piperidone (Ahp) and 4-hydroxy-*L*-phenylglycine (HPG).

Analysis of the translated NRPS A-domain proteins, using the tool for natural product biosynthesis (http://www.npbiogene.com/) was also carried out. Results obtained showed high homology to sequences of modules and domains present in metabolites from both cyanobacteria and others such as proteobacteria and actinobacteria (Table 2). In some cases, these were common cyanobacterial toxins such as microcystin (**1**), a potent hepatotoxic heptapeptide; or nodularin (**2**), a cyclic pentapeptide with comparable activity to microcystin, both produced by a diverse range of cyanobacteria, including species of the genera *Microcystis*, *Anabaena*, *Nostoc*, *Nodularia* and *Oscillatoria* [59]. Here, we demonstrate the presence of genes with homology to microcystin synthetase A (McyA) from *M. aeruginosa* PCC 7806, in *A. cylindrica* UTEX 629, *Scytonema* sp. PCC 7110, *Gloeothece* sp. ATCC 27152 and *L. majuscula* CCAP 1446/4. Additionally, a *Scytonema* sp. PCC 7110 PCR product showed homology to McyC (*E* value 2.0×10^{-64}). While PCR products from *Anabaena* sp. PCC 9109 and *Tolypothrix* sp. PCC 7601 (Table 2) showed homology to the A-domain in nodularin synthetase B (NdaB) from *Nodularia spumigena* strain NSOR10 [42]. MycC and NdaB are the terminal activation domains in microcystin and nodularin biosynthesis, and in both cases the substrate-binding pocket was found to be specific for the activation of *L*-Arg [42].

In others, PCR products with homology to the genes encoding the A domains of the siderophore pyoverdin D (**3**), involved in iron-gathering and virulence and synthesized by *Pseudomonas aeruginosa* were identified in *Lyngbya* sp. CCAP 1446/5 and *Oscillatoria* sp. CCAP 1459/26 (Table 2). Furthermore, a probable gene encoding the module 2 A-domain of leinamycin (**4**), a novel thiazole-containing antimicrobial and antitumor compound [28], produced by several *Streptomyces atroolivaceus* species, was discovered in *A. cylindrica* UTEX 629.

Identification of ketosynthase domains (PKS)

Putative polyketide synthetase KS-domain gene sequences were obtained (~700 bp size) in 19 of the 21 cyanobacterial strains screened (Table 1). Strains that showed no PCR product under the conditions tested were *A. cylindrica* UTEX 629 (Nostocales), *Leptolyngbya* sp. PCC 6703 and PCC 7104, *L. aestuarii* PCC 7419 and *P. terebrans* CCAP 1463/4 (Oscillatoriales). As seen in the NRPS analysis, multiple products were obtained from a number of these strains. In this case, 29 unique KS-domains were present in 18 cyanobacteria (Table 3). These sequences are similar to known cyanobacterial PKS sequences, with homologies of between 49 and 95% (Table 3). BLASTp analysis of the PKS sequences reveal that twenty two clones are homologous to modular KS-domains present in a variety cyanobacterial secondary metabolites (Table 3), with the remainder homologous to other bacterial PKS systems.

Twelve KS sequences were homologous to the Type I KS-domain of jamaicamide (**5**). This compound shows a spectrum of biological activities, including sodium channel blocking and cytotoxicity [20]. Specifically, products from *Anabaena* sp. PCC 9109, *Scytonema* sp. PCC 7110, *G. violaceus* PCC 7421 and *Oscillatoria* sp. CCAP 1459/13 (Table 3) show homology to the module 1 KS-domain from JamE. JamE is highly unusual, in that it contains three consecutive ACPs downstream of the KS-domain [20]. Interestingly, clone 5 of *Gloeothece* sp. ATCC 27152 shows homology to JamM, a type I PKS with homology to the mixed NRPS/PKS BarE (56% identity) involved in barbamide biosynthesis. Lastly, *Oscillatoria* sp. CCAP 1459/26 clone 1 and *Lyngbya* sp. CCAP 1446/5 clone 6 showed homology to JamP, the penultimate module in the Jam pathway, flanked by a terminal thioesterase (TE) domain of putative NRPS origin [20]. Additionally, *Planktothrix* sp. PCC 7811 clone 1 (Table 3) also showed homology to barbamide (**6**) (BarE). This compound, a chlorinated lipopeptide produced by *L. majuscula* strain 19L, is composed of a NRPS followed by a PKS [10]. It is thought that within this module, the trichloroisoleucine is loaded from the A-domain to the KS-domain for further extention.

Five other sequences were homologous to KS domains in stigmatellin (Table 3). Stigmatellin (**7**) is a natural product that possesses an aromatic moiety produced through a PKS system, from the mixobacterium *Stigmatella aurantiaca* [24]. The biosynthesis of this compound is a special case, where an aromatic compound is produced by a bacterial type I PKS, instead of a type II PKS [24]. Furthermore, products from *Tolypothrix* sp. PCC 7601, *Leptolyngbya* sp. PCC 7410 and *P. boryanum* PCC 73110 show high homology to the KS-domain in NosB, in the cluster responsible for the biosysnthesis of nostopeptolide (**8**), a cyclic heptapeptide produced by the terrestrial cyanobacterium *Nostoc* sp. GSV224 [29]. Although NosB is a type I PKS located between two NRPSs (NosA and NosC), it is also similar to a KS-domain in a mixed NRPS/PKS [18].

The PCR products of *A. ambigua* CCAP 1403/7 and *Tolypothrix* sp. PCC 7601 show homology to the KS-domain in mupirocin (**9**) (Table 3). This compound is produced by *Pseudomonas fluorescens* NCIMB 10586, and it is proposed that a separate PKS and fatty acid synthase (FAS) are involved in its assembly from monic acid (MA) and 9-hydroxynonanoic acid via esterification [22]. The sequences obtained in this work are similar to the KS-domains in modules 1, 2 and 4 from MmpD involved in the biosynthesis of MA.

The most interesting finding revolved around the KS-domain PCR product isolated from *Oscillatoria* sp. PCC 7515, which has homology with the EposC module 5 in

Table 3 Cyanobacterial PKS sequences analyzed using the BLASTp tool for natural product biosynthesis

Strain	Clone	Accession number[a]	Domain	Compound[d]	*E* value
***Anabaena sp.* PCC 7120[b]**	**17135468**	**AJ421825**	**KS3**	**Stigmatellin**	$\mathbf{5 \times 10^{-89}}$
Anabaena sp. PCC 9109	Clone-2	AY522504	JamE-KS1	Jamaicamide	1×10^{-67}
	Clone-3	AY522504	JamE-KS1	Jamaicamide	2×10^{-90}
Anabaena ambigua CCAP 1403/7	Clone-3	AF318063	KS2	Mupirocin	1×10^{-61}
	Clone-4	AF319998	KS4	Myxalamid	3×10^{-47}
	Clone-5	AF318063	KS4	Mupirocin	4×10^{-69}
***Anabaena variabilis* ATCC 29413[c]**	**Clone-7.1**	**AJ421825**	**KS3**	**Stigmatellin**	$\mathbf{5 \times 10^{-88}}$
	Clone-7.2	**AJ421825**	**KS6**	**Stigmatellin**	$\mathbf{1 \times 10^{-82}}$
***Nostoc punctiforme* PCC 73102[b]**	**23125974**	**AAT70096**	**curL-KS8**	**Curacin**	$\mathbf{8 \times 10^{-77}}$
	23126921	**AJ421825**	**KS5**	**Stigmatellin**	$\mathbf{1 \times 10^{-78}}$
Nodularia harveyana CCAP 1452/1	Clone-4	AY522504	JamK-KS3	Jamaicamide	1×10^{-79}
	Clone-5	AY522504	JamK-KS4	Jamaicamide	2×10^{-90}
Scytonema sp. PCC 7110	Clone-2	AY522504	JamE-KS1	Jamaicamide	4×10^{-91}
Tolypothrix sp. PCC 7601	Clone-2	AF204805	nosB-KS	Nostopeptolide	4×10^{-72}
	Clone-3	AF318063	KS1	Mupirocin	1×10^{-69}
***Gloeobacter violaceus* PCC 7421[c]**	**Clone-3**	**AY522504**	**JamE-KS1**	**Jamaicamide**	$\mathbf{8 \times 10^{-88}}$
Gloeothece sp. ATCC 27152	Clone-1	AY522504	JamK-KS3	Jamaicamide	3×10^{-81}
	Clone-5	AY522504	JamM-KS6	Jamaicamide	5×10^{-80}
	Clone-7	AAT70096	curA-KS1	Curacin	2×10^{-76}
Leptolyngbya sp. PCC 7410	Clone-1	AF204805	nosB-KS	Nostopeptolide	3×10^{-81}
Lyngbya sp. CCAP 1446/5	Clone-6	AY522504	JamP-KS8	Jamaicamide	5×10^{-75}
Lyngbya majuscula CCAP 1446/4	Clone-2	AF484556	KS8	Leinamycin	5×10^{-69}
	Clone-6	AB032549	KS2	Microcystin	4×10^{-59}
	Clone-7	AJ421825	KS7	Stigmatellin	3×10^{-78}
Microcoleus sp. PCC 8701	Clone-2	AJ421825	KS2	Stigmatellin	4×10^{-81}
Planktothrix sp. PCC 7811	Clone-1	AF516145	barE-KS	Barbamide	4×10^{-98}
Oscillatoria sp. PCC 7515	Clone-1	AF217189	KS5	Epothilone	1×10^{-80}
Oscillatoria sp. PCC 6506	Clone-5	AY522504	JamK-KS3	Jamaicamide	2×10^{-75}
	Clone-7	AJ421825	KS7	Stigmatellin	1×10^{-80}
Oscillatoria sp. CCAP 1459/13	Clone-4	AY522504	JamE-KS1	Jamaicamide	3×10^{-68}
Oscillatoria sp. CCAP 1459/26	Clone-1	AY522504	JamP-KS8	Jamaicamide	6×10^{-68}
Plectonema boryanum PCC 73110	Clone-1	AF204805	nosB-KS	Nostopeptolide	8×10^{-81}

[a] Accession numbers corresponding to the NCBI website for amino acid sequences

[b] Sequences highlighted in Bold were obtained from GenBank (genomes)

[c] Sequence available from GenBank used as a positive control

[d] The compound here is for illustrative purposes, in that the domain identified is similar. This does not imply the strain makes this exact compound

epothilone (**10**) biosynthesis. This compound is produced by the myxobacterium *Sorangium cellulosum* So ce90, and, like paclitaxel (Taxol®), inhibits microtubule depolymerisation and arrests the cell cycle at the G2-M phase [33, 65]. According to Molnar et al. [45], this KS-domain performs an elongation step with the incorporation of a propionate unit into the growing natural product structure.

Similarly, cluster analysis was performed on the translated KS-domain protein sequences, with the resulting sequence alignments clustering into two functional groups [43] (Fig. 3). The first group represents type I KS-domains. In this group, our sequences clustered together with KS-domains from microcystin synthetase (MycD and MycE), curacin (**11**) (CurA and CurI), jamaicamide (JamE) and mixobacterial KS-domains from epothilone (EposA and EposD). The second cluster contains sequences representative of hybrid or mixed NRPS/PKS systems. Examples include KS-domains from barbamide (BarE), curacin (CurF) and nostopeptolide (NosB) synthetase from known cyanobaceria, together with others such as mixobacterial KS-domains from Bleomycin (BlmVIII) and Mixothiazol (MtaD).

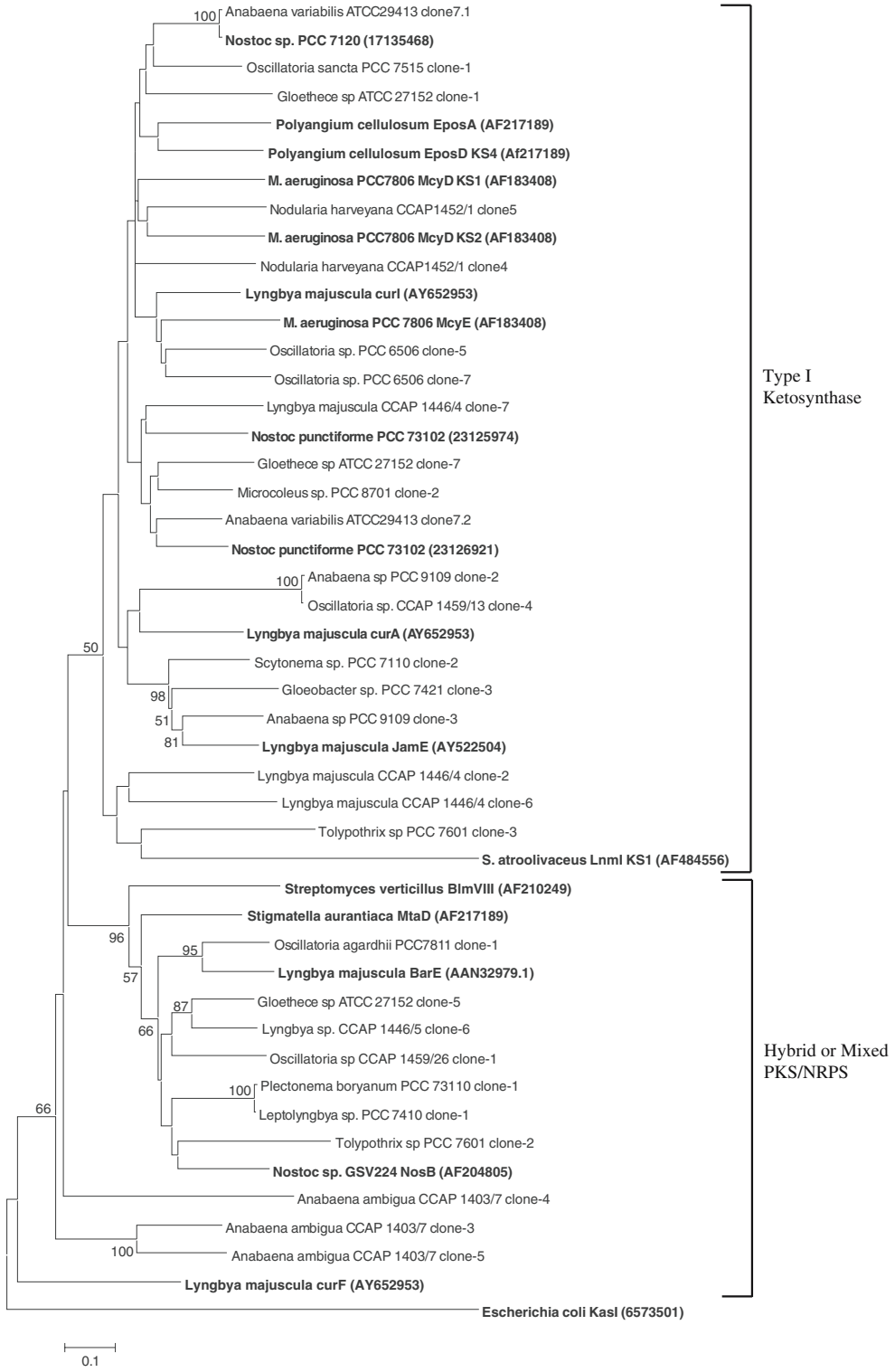


Fig. 3 Phylogenetic analysis of type I ketosynthase domains. Sequences obtained from GenBank are given in *boldface* and their accession numbers in *parentheses*. The fatty acid synthase KS-domain from *Escherichia coli* (Accession no. 6573501), was used as an out-group. Sequences were aligned using the Clustal W program, and tree reconstruction performed using the neighbor-joining method supplied by MEGA3 package (version 3.0). Significant bootstrap values were calculated from 1000 bootstrap resamplings

Five members of the Oscillatoriales family were found to contain KS-domains exclusive to mixed NRPS/PKS gene clusters. Specific examples include *P. boryanum* PCC 73110, *Leptolyngbya* PCC7410, *P. agardhii* PCC 7811, *Oscillatoria* sp. CCAP 1459/26 and *Lyngbya* sp. CCAP 1456/5. Conversely, only a single Chroococcales strain,

Gloeothece sp. ATCC 27152, has a single KS-domain of mixed NRPS/PKS type. Furthermore, analysis of the KS-domain active-site sequence motif (Supplementary Fig. 1a, b) revealed a relatively highly conserved dtaCSSSL motif (where lower case letters denote unconserved amino acids) around the active-site in all strains. This observation is in agreement to that of Moffitt and Neilan [43].

Discussion

It has been emphasised that cyanobacterial secondary metabolites are produced through large multienzyme complexes, constituted by NRPS and PKS modules responsible for the addition of an amino acid or chain-elongation step, respectively. Several authors have performed similar screening of these two genes in cyanobacteria as relates to a specific compound, for example microcystin in *Microcystis aeruginosa* PCC 7806 or nodularin in *Nodularia spumigena* strain NSOR10 [42, 71], or alternatively in a wide range of cyanobacterial families [6, 14, 21]. In this study, we screened a broad range of cyanobacterial families for both NRPS and PKS genes, and analysed these results in order to infer some characteristics of the putative natural products (secondary metabolites) produced.

To our knowledge, only a single study attempted to cluster NRPS protein sequences in cyanobacteria [21]. In that work, the authors found no clear phylogenetic correlation, and in fact high variability within the A-domains was observed. In contrast, our analysis showed that clustering of the full A-domain is possible when substrate conferring tendencies are taken into account (Fig. 2). These results confirm those previously described by Marahiel et al. [61]. Here we can discern branches specific to Ile, Val, Asp/Asn, Gly, HPG, Thr, Ahp, Lys, Pro, His, Phe, Cys, Leu, Glu, Ser and β-alanine (β-Ala) substrates.

In clones 3 and 7, from *Leptolyngbya* sp. PCC 7410, we found two A-domains that are homologous to those in the NosC module responsible for nostopeptolide biosynthesis in *Nostoc* sp. GSV224. These A-domains activate the non-protein amino acids HPG, and Asp/Asn. HPG is synthesized through an unusual biosynthetic pathway [30] through the stoichiometric conversion of *L*-tyrosine. Three enzymes have been identified in HPG biosynthesis and are common to the secondary metabolites teicoplanin [37], complestatin [13] and nocardicin A [25]. These are 4-hydroxymandelic acid synthase, 4-hydroxymandelic acid oxidase, and 4-hydroxyphenylglycine aminotransferase [36]. This information should be useful in the design of specific molecular probes targeting this pathway.

Analysis of the specificity of A-domains towards particular amino acids (Table 2) shows the presence of motifs that recognize three amino acids, Cys, Ser and Thr in six strains. These amino acids are associated with the formation of five-membered heterocyclic rings, which form either thiazole or oxazole, or their reduced structures [73], in bacterial secondary metabolites. These compounds, being signature pharmacophores, are common to many clinically important natural products [52], such as bacitracin [31], leinamycin [12] (antibacterial), bleomycin [17], and epothilone [45, 69] (anticancer). The incorporation of these amino acids is done through a cyclization (Cy) domain that catalyzes amide bond formation, cyclization of the side chain nucleophile with the newly formed amides, and a final dehydration [54].

The binding pocket amino acid signature codes found within *A. cylindrica* clone 8 and *Gloethece* sp. clone 1 are homologous to those found in bacitracin and pyochelin, both natural products known to have a thiazoline heterocycle. While, *Lyngbya* sp. CCAP 1446/5 clone 1 showed specificity for Ile with homology to the module involved in the formation of the heterocycle in bacitracin.

Moreover, clone 3 of this strain and *Oscillatoria* sp. CCAP 1459/26 clone 1 contain a domain that appears to use Ser as a building block, similar to the type of domain found in *Pseudomonas aeruginosa* for the the formation of pyoverdin. A similar homology is seen in the sequence from *A. variabilis* clone 4, with repect to Ser incorporation in nostopeptolide in *Nostoc* sp. GSV224.

In order to elucidate the specificity of the *Scytonema* sp. clone 2 amino acid binding pocket it was necessary to compare its signature sequence with similar organisms. These results showed an unusual start amino acid (**VDWITSL-G**), with Val instead of Asp as the starting amino acid. Previously, this feature was identified in the A-domains of bleomycin NRPS-2, exochelin FxB-2 and jamaicamide JamL which showed specificity to β-Ala. Comparison of these sequences with our data showed 56, 44 and 33% identity. Here, we propose β-Ala as the amino acid activated by the binding pocket of the A-domain found in *Scytonema* sp. PCC 7110 clone 2. This strain is known to produce scytonemin (an antiproliferative pharmacophore) [64], and the cyclic depsipeptides scyptolin A and B (inhibits porcine elastase in vitro) [40] and hofmannolin [39]. Analysis of the amino acid binding pocket in the A-domain identified in *Scytonema* sp. PCC 7110 shows specificity towards the amino acids Val and β-Ala, and proves the presence of genetic machinery associated with the production of these depsipeptides (Fig. 4 grey area), and is the subject of ongoing work to confirm the assignment of these modules. This is especially important in the case of the A-domain that recognises β-Ala, as the amino acid present in scyptolin A is *L*-Ala rather than the former. Modifications to scyptolin could thus either occur post compound formation, or alternatively be involved in the formation of a hitherto unknown compound.

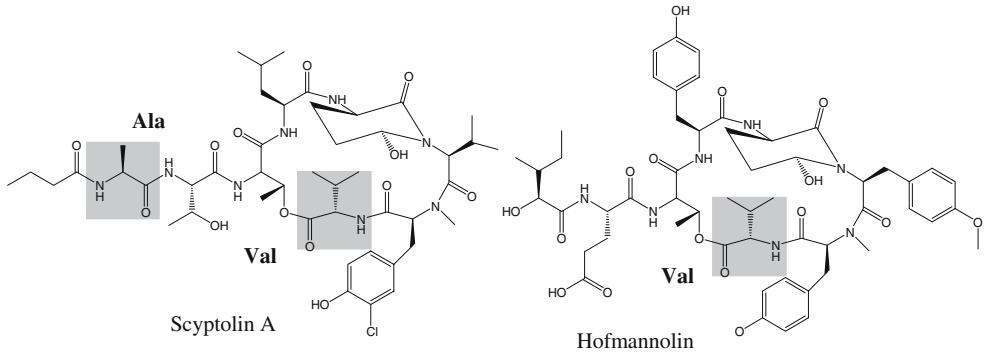


Fig. 4 Illustration of the structure of Scyptolin A and Hofmannolin produced by *Scytonema* sp. PCC 7110. Amino acids encoded by the A-domains identified here in clones 2 and 6 are highlighted in *grey*. Substrate amino acids are abbreviated with a three-letter code

Cluster analysis of polyketide synthase sequences, on the other hand, identified two distinct KS types. These being, those that are exclusively type I and those from hybrid or mixed NRPS/PKS modules, and is in agreement with results from Neilan et al. [43]. The exceptions in this case being the KS-domain of LmnI (associated with leinamycin synthesis) [68] and that of MtaD (associated with myxothiaxol synthesis) [57], both mixed NRPS/PKS that cluster with type I PKS sequences. LmnI was found to cluster within the same sub-group, together with clone 3 from *Tolypothrix* sp. PCC 7601 and clone 2 and 6 from *L. majuscula* CCAP 1446/4. Furthermore, five of our clone sequences were found to be homologous to the JamE KS-domain (Table 3). The three consecutives ACPs present in JamE are believed to be involved in channeling intermediates and providing enzyme docking stations for vinyl or vinyl chloride formation [20]. A similar mechanism may be present here.

Conclusions

Our results reinforce the concept of exploiting cyanobacteria as viable producers of secondary metabolites, by demonstrating their potential to produce a wide range of natural products. Specifically, insight was obtained by screening the adenylation domain and ketosynthase domain of NRPS and PKS genes from both marine and freshwater cyanobacteria. Analysis of these genes reveals the presence of moieties from putative natural products commonly observed in pharmacologically active secondary metabolites. This methodology has previously been demonstrated in the elucidation of cryptic-gene clusters responsible for a variety of antibiotics in actinomycetes [74], and in the structural prediction of a novel peptide encoded in the genome of *Streptomyces coelicolor* [8, 35]. Moreover, analysis of *Scytonema* sp. PCC 7110 NRS A-domain sequences has identified motifs possessing specificities towards the amino acids, valine and alanine, present in the natural products Scyptolin A and Hofmannolin previously isolated and characterised in this strain. The genetic results presented here provide the foundation to guide future research towards specific metabolites, or classes of metabolites within these organisms.

Acknowledgments We would like to thank Helia Radianingtyas for the supply of most of the genomic DNA used here. We also thank the Overseas Research Student's Award Scheme and The University of Sheffield's PhD Scholarship Programme for provision of a PhD scholarship to MBL. This work was supported by the UK Engineering and Physical Science Research Council (EPSRC), and The Carbon Trust. PCW wishes to thank the EPSRC for an advanced research fellowship.

References

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
2. Becher PG, Beuchat J, Gademann K, Juttner F (2005) Nostocarboline: isolation and synthesis of a new cholinesterase inhibitor from nostoc 78-12A. J Nat Prod 68:1793–1795
3. Biondi N, Piccardi R, Margheri MC, Rodolfi L, Smith GD, Tredici MR (2004) Evaluation of Nostoc strain ATCC 53789 as a potential source of natural pesticides. Appl Environ Microbiol 70:3313–3320
4. Borowitzka MA (1995) Microalgae as a source of pharmaceuticals and other biologically active compounds. J Appl Phycol 7:13–15
5. Burja AM, Banaigs B, Abou-Mansour E, Burgess JG, Wright PC (2001) Marine cyanobacteria—a prolific source of natural products. Tetrahedron 57:9347–9377
6. Burns BP, Goh F, Allen M, Neilan BA (2004) Microbial diversity of extant stromatolites in the hypersaline marine environment of Shark Bay, Australia. Environ Microbiol 6:1096–1101
7. Cane DE, Walsh CT (1999) The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem Biol 6:R319–R325
8. Challis GL, Ravel J (2000) Coelichelin, a new peptide siderophore encoded by the Streptomyces coelicolor genome: structure prediction from the sequence of its non-ribosomal peptide synthetase. FEMS Microbiol Lett 187:111–114
9. Challis GL, Ravel J, Townsend CA (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol 7:211–224
10. Chang Z, Flatt P, Gerwick WH, Nguyen VA, Willis CL, Sherman DH (2002) The barbamide biosynthetic gene cluster: a novel marine

cyanobacterial system of mixed polyketide synthase (PKS)-nonribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. Gene 296:235–247
11. Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, Sherman DH, Gerwick WH (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. J Nat Prod 67:1356–1367
12. Cheng YQ, Tang GL, Shen B (2002) Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in *Streptomyces atroolivaceus* S-140. J Bacteriol 184:7013–7024
13. Chiu HT, Hubbard BK, Shah AN, Eide J, Fredenburg RA, Walsh CT, Khosla C (2001) Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. Proc Natl Acad Sci USA 98:8548–8553
14. Christiansen G, Dittmann E, Via Ordorika L, Rippka R, Herdman M, Borner T (2001) Nonribosomal peptide synthetase genes occur in most cyanobacterial genera as evidenced by their distribution in axenic strains of the PCC. Arch Microbiol 176:452–458
15. Dittmann E, Neilan BA, Borner T (2001) Molecular biology of peptide and polyketide biosynthesis in cyanobacteria. Appl Microbiol Biotechnol 57:467–473
16. Dittmann E, Neilan BA, Erhard M, von Dohren H, Borner T (1997) Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Mol Microbiol 26:779–787
17. Du L, Sanchez C, Chen M, Edwards DJ, Shen B (2000) The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. Chem Biol 7:623–642
18. Du L, Sanchez C, Shen B (2001) Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules. Metab Eng 3:78–95
19. Edwards DJ, Gerwick WH (2004) Lyngbyatoxin biosynthesis: sequence of biosynthetic gene cluster and identification of a novel aromatic prenyltransferase. J Am Chem Soc 126:11432–11433
20. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, Gerwick WH (2004) Structure and biosynthesis of the jamaicamides, new mixed polyketide–peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. Chem Biol 11:817–833
21. Ehrenreich IM, Waterbury JB, Webb EA (2005) Distribution and diversity of natural product genes in marine and freshwater cyanobacterial cultures and genomes. Appl Environ Microbiol 71:7401–7413
22. El-Sayed AK, Hothersall J, Cooper SM, Stephens E, Simpson TJ, Thomas CM (2003) Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. Chem Biol 10:419–430
23. Falch BS, Konig GM, Wright AD, Sticher O, Angerhofer CK, Pezzuto JM, Bachmann H (1995) Biological activities of cyanobacteria: evaluation of extracts and pure compounds. Planta Med 61:321–328
24. Gaitatzis N, Silakowski B, Kunze B, Nordsiek G, Blocker H, Hofle G, Muller R (2002) The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. J Biol Chem 277:13082–13090
25. Gunsior M, Breazeale SD, Lind AJ, Ravel J, Janc JW, Townsend CA (2004) The biosynthetic gene cluster for a monocyclic beta-lactam antibiotic, nocardicin A. Chem Biol 11:927–938
26. Han B, Goeger D, Maier CS, Gerwick WH (2005) The wewakpeptins, cyclic depsipeptides from a Papua New Guinea collection of the marine cyanobacterium *Lyngbya semiplena*. J Org Chem 70:3133–3139
27. Han B, McPhail KL, Ligresti A, Di Marzo V, Gerwick WH (2003) Semiplenamides A-G, fatty acid amides from a Papua New Guinea collection of the marine cyanobacterium *Lyngbya semiplena*. J Nat Prod 66:1364–1368
28. Hara M, Asano K, Kawamoto I, Takiguchi T, Katsumata S, Takahashi K, Nakano H (1989) Leinamycin, a new antitumor antibiotic from Streptomyces: producing organism, fermentation and isolation. J Antibiot (Tokyo) 42:1768–1774
29. Hoffmann D, Hevel JM, Moore RE, Moore BS (2003) Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from Nostoc sp. GSV224. Gene 311:171–180
30. Hubbard BK, Thomas MG, Walsh CT (2000) Biosynthesis of L-*p*-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. Chem Biol 7:931–942
31. Konz D, Klens A, Schorgendorfer K, Marahiel MA (1997) The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. Chem Biol 4:927–937
32. Konz D, Marahiel MA (1999) How do peptide synthetases generate structural diversity? Chem Biol 6:R39–R48
33. Kowalski RJ, Giannakakou P, Hamel E (1997) Activities of the microtubule-stabilizing agents epothilones A and B with purified tubulin and in cells resistant to paclitaxel (Taxol®). J Biol Chem 272:2534–2541
34. Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163
35. Lautru S, Deeth RJ, Bailey LM, Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. Nat Chem Biol 1:265–269
36. Li TL, Choroba OW, Charles EH, Sandercock AM, Williams DH, Spencer JB (2001) Characterisation of a hydroxymandelate oxidase involved in the biosynthesis of two unusual amino acids occurring in the vancomycin group of antibiotics. Chem Commun (Camb) 1752–1753
37. Li TL, Huang F, Haydock SF, Mironenko T, Leadlay PF, Spencer JB (2004) Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. Chem Biol 11:107–119
38. Marahiel MA, Stachelhaus T, Mootz HD (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem Rev 97:2651–2674
39. Matern U, Oberer L, Erhard M, Herdman M, Weckesser J (2003) Hofmannolin, a cyanopeptolin from *Scytonema hofmanni* PCC 7110. Phytochemistry 64:1061–1067
40. Matern U, Oberer L, Falchetto RA, Erhard M, Konig WA, Herdman M, Weckesser J (2001) Scyptolin A and B, cyclic depsipeptides from axenic cultures of *Scytonema hofmanni* PCC 7110. Phytochemistry 58:1087–1095
41. Mayer AM, Gustafson KR (2004) Marine pharmacology in 2001–2: antitumour and cytotoxic compounds. Eur J Cancer 40:2676–2704
42. Moffitt MC, Neilan BA (2004) Characterization of the nodularin synthetase gene cluster and proposed theory of the evolution of cyanobacterial hepatotoxins. Appl Environ Microbiol 70:6353–6362
43. Moffitt MC, Neilan BA (2003) Evolutionary affiliations within the superfamily of ketosynthases reflect complex pathway associations. J Mol Evol 56:446–457
44. Moffitt MC, Neilan BA (2001) On the presence of peptide synthetase and polyketide synthase genes in the cyanobacterial genus *Nodularia*. FEMS Microbiol Lett 196:207–214
45. Molnar I, Schupp T, Ono M, Zirkle R, Milnamow M, Nowak-Thompson B, Engel N, Toupet C, Stratmann A, Cyr DD, Gorlach

J, Mayo JM, Hu A, Goff S, Schmid J, Ligon JM (2000) The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. Chem Biol 7:97–109
46. Moore RE (1996) Cyclic peptides and depsipeptides from cyanobacteria: a review. J Ind Microbiol 16:134–143
47. Namikoshi M, Rinehart KL (1996) Bioactive compounds produced by cyanobacteria. J Ind Microbiol 17:373–384
48. Nishizawa T, Ueda A, Asayama M, Fujii K, Harada K, Ochi K, Shirai M (2000) Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin. J Biochem (Tokyo) 127:779–789
49. Patterson GM, Bolis CM (1994) Scytophycin production by axenic cultures of the cyanobacterium *Scytonema ocellatum*. Nat Toxins 2:280–285
50. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J Gen Microbiol 111:1–61
51. Rouhiainen L, Paulin L, Suomalainen S, Hyytiainen H, Buikema W, Haselkorn R, Sivonen K (2000) Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in Anabaena strain 90. Mol Microbiol 37:156–167
52. Roy RS, Gehring AM, Milne JC, Belshaw PJ, Walsh CT (1999) Thiazole and oxazole peptides: biosynthesis and molecular machinery. Nat Prod Rep 16:249–263
53. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbour Laboratory Press, Cold Spring Harbour
54. Schwarzer D, Marahiel MA (2001) Multimodular biocatalysts for natural product assembly. Naturwissenschaften 88:93–101
55. Shen B, Du L, Sanchez C, Edwards DJ, Chen M, Murrell JM (2001) The biosynthetic gene cluster for the anticancer drug bleomycin from *Streptomyces verticillus* ATCC15003 as a model for hybrid peptide–polyketide natural product biosynthesis. J Ind Microbiol Biotechnol 27:378–385
56. Shimizu Y (1996) Microalgal metabolites: a new perspective. Annu Rev Microbiol 50:431–465
57. Silakowski B, Schairer HU, Ehret H, Kunze B, Weinig S, Nordsiek G, Brandt P, Blocker H, Hofle G, Beyer S, Muller R (1999) New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. J Biol Chem 274:37391–37399
58. Simmons TL, Andrianasolo E, McPhail K, Flatt P, Gerwick WH (2005) Marine natural products as anticancer drugs. Mol Cancer Ther 4:333–342
59. Sivonen K (1996) Cyanobacterial toxins and toxin production. Phycologia 35:12–24
60. Smith GD, Thanh Doan N (1999) Cyanobacterial metabolites with bioactivity against photosynthesis in cyanobacteria, algae and higher plants. J Appl Phycol 11:337–344
61. Stachelhaus T, Mootz HD, Marahiel MA (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem Biol 6:493–505
62. Staunton J, Weissman KJ (2001) Polyketide biosynthesis: a millennium review. Nat Prod Rep 18:380–416
63. Staunton J, Wilkinson B (2001) Combinatorial biosynthesis of polyketides and nonribosomal peptides. Curr Opin Chem Biol 5:159–164
64. Stevenson CS, Capper EA, Roshak AK, Marquez B, Eichman C, Jackson JR, Mattern M, Gerwick WH, Jacobs RS, Marshall LA (2002) The identification and characterization of the marine natural product scytonemin as a novel antiproliferative pharmacophore. J Pharmacol Exp Ther 303:858–866
65. Su D-S, Balog A, Meng D, Bertinato P, Danishefsky SJ, Zheng Y-H, Chou T-C, He L, Horwitz SB (1997) Structure-activity relationship of the Epothilones and the first in vivo comparison with Paclitaxel. Angew Chem Int Ed Engl 36:2093–2096
66. Tamagnini P, Troshina O, Oxelfelt F, Salema R, Lindblad P (1997) Hydrogenases in Nostoc sp. strain PCC 73102, a strain lacking a bidirectional enzyme. Appl Environ Microbiol 63:1801–1807
67. Tan LT, Sitachitta N, Gerwick WH (2003) The guineamides, novel cyclic depsipeptides from a Papua New Guinea collection of the marine cyanobacterium *Lyngbya majuscula*. J Nat Prod 66:764–771
68. Tang GL, Cheng YQ, Shen B (2004) Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. Chem Biol 11:33–45
69. Tang L, Shah S, Chung L, Carney J, Katz L, Khosla C, Julien B (2000) Cloning and heterologous expression of the epothilone gene cluster. Science 287:640–642
70. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
71. Tillett D, Dittmann E, Erhard M, von Dohren H, Borner T, Neilan BA (2000) Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. Chem Biol 7:753–764
72. von Dohren H, Keller U, Vater J, Zocher R (1997) Multifunctional peptide synthetases. Chem Rev 97:2675–2706
73. Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, Patel HM (2001) Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. Curr Opin Chem Biol 5:525–534
74. Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways. Nat Biotechnol 21:187–190