CHAPTER 2

MOLECULAR BIOLOGY OF CYANOBACTERIAL TOXINS

Genetic basis of microcystin production

Thomas BÖRNER & Elke DITTMANN

1 Introduction

Cyanobacteria produce different metabolites including alkaloids, lipo-polysaccharides, polyketides and peptides that may act as toxins on other bacteria, lower and higher eukaryotes. Until today, only a few genes involved in the biosynthesis of cyanobacterial toxins have been discovered, all belonging to the group of peptides/polyketides. The way leading to the detection of these genes has been paved by several important discoveries. Microcystins were found to be small circular heptapeptides containing several unusual amino acids (Bishop et al., 1959; Botes et al., 1984; Rinehart et al., 1988). Moore et al. (1991) reported that *Microcystis aeruginosa* PCC 7820 synthesises microcystin-LR from L-methionine, L-phenylalanine, L-glutamic acid, acetate and pyruvate. These precursor studies, together with data from an assay for thio-template activity (Arment and Carmichael, 1996) and with the structure of microcystin, suggested a nonribosomal way of biosynthesis for the peptide backbone and a polyketide pathway for the characteristic Adda ((2*S,* 3*S,* 8*S,* 9*S*)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid) moiety. These two types of biosyntheses were first discovered for the peptide antibiotic gramicidine S and the macrolide antibiotic erythromycin, respectively.

Non-ribosomal peptide synthetases (NRPS) represent a family of large, multifunctional modular enzymes, in which each module is responsible for the activation, modification and condensation of an individual amino acid. These enzymes form large complexes serving as templates for the biosynthesis of peptides as an alternative way to (poly)peptide formation on ribosomes (for review see Marahiel et al., 1997; von Döhren et al., 1997). The large number of bacterial polyketides is synthesised in a similar way through the action of modular polyketide synthases (PKS). The great variety of polyketide structures is achieved by the incorporation of various acyl units that can be further modified by integrated and

J. Huisman, H.C.P. Matthijs and P.M. Visser (eds.), Harmful Cyanobacteria, 25–40.
© 2005 *Springer. Printed in the Netherlands.*

non-integrated domains and enzymes, respectively (for review see Hopwood, 1997; Khosla, 1997). Even though there was a good chance that microcystins are indeed synthesised on multi-enzyme complexes built up by NRPS and PKS, neither the respective genes nor the enzymes had been described for cyanobacteria. Since NRPS of bacteria and fungi are phylogenetically related and share consensus sequences (Turgay and Marahiel, 1994), an obvious approach to the identification of genes for microcystin biosynthesis was to search for consensus sequences of NRPS-encoding genes in cyanobacteria.

NRPS genes could indeed be identified for the first time in cyanobacteria by investigating microcystin-producing and non-producing *Microcystis* cells (Dittmann et al., 1996; Meissner et al., 1996). In parallel, NRPS genes for the synthesis of anabaenopeptilide had been investigated in *Anabaena* (Rouhiainen et al., 2000). Since toxic as well as non-toxic cells harbour NRPS genes, methods for genetic manipulation of toxin-producing cyanobacteria and for directed mutagenesis of NRPS genes had to be established to facilitate identification of those genes that are needed for microcystin production. The goal was finally attained by knock-out mutagenesis of NRPS genes in *M. aeruginosa* PCC 7806 leading to mutant cells no longer capable of synthesising microcystin (Dittmann et al., 1997). This study was completed by sequencing a large gene cluster coding for the essential components of the microcystin synthetase complex in the *M. aeruginosa* strains PCC 7806 (Tillett et al., 2000) and K-139 (Nishizawa et al., 1999; 2000).

Following the same or a similar approach, it will now be possible to detect the genes for all other cyanobacterial peptide toxins that are synthesised non-ribosomally. The microcystin synthetase (*mcy*) genes have been characterised recently for a *Planktothrix* (Christiansen et al., 2003) and an *Anabaena* strain (Rouhiainen et al., 2004). Furthermore, NRPS and PKS genes could be detected that are involved in the biosynthesis of the related hepatotoxin nodularin (M.C. Moffitt and B.A. Neilan, accession no. AY210783). The search for genes with a role in the biosynthesis of cylindrospermopsin, another hepatotoxin, has also revealed NRPS and PKS genes that occur in toxin-producing strains of *Cylindrospermopsis raciborskii* and *Anabaena bergii*, but not in non-toxic cells (Schembri et al., 2001), though direct evidence for their suggested function is lacking. Here we review studies on the microcystin and nodularin synthetase genes and discuss aspects of their function and evolution.

2 Microcystin synthetase: genes and their function in *Microcystis aeruginosa*

The principle organisation of the gene cluster coding for the individual enzymes involved in microcystin biosynthesis is identical in the two strains of *M. aeruginosa* that have been investigated in more detail (Nishizawa et al., 2000; Tillett et al., 2000), but differs between the genera (see below). The *mcy* gene cluster of *M. aeruginosa* spans about 55kb of the chromosomal DNA and comprises 10 genes embedded in two bidirectionally transcribed operons (Tillett et al., 2000; Kaebernick et al., 2002). Downstream of the promoter region, *mcyA-C* encode three NRPS comprising 5 modules, whereas upstream *mcyD-J* encode polyketide synthases

(*mcyD*), hybrid enzymes (*mcyE, G*) and additional tailoring enzymes (*mcyF, I, J*) as well as a component of a putative ABC transporter (*mcyH*) (Fig. 1). Taking into account that the multi-enzyme components of the Mcy complex comprise the minimal set of domains needed for the assembly of seven amino acids and four acetate units plus additional integrated domains with epimerase, methyltransferase and aminotransferase activities, the enzymes encoded by the *mcy* gene cluster are thought to catalyse 45 out of the postulated 48 sequential steps of microcystin biosynthesis (Tillett et al., 2000).

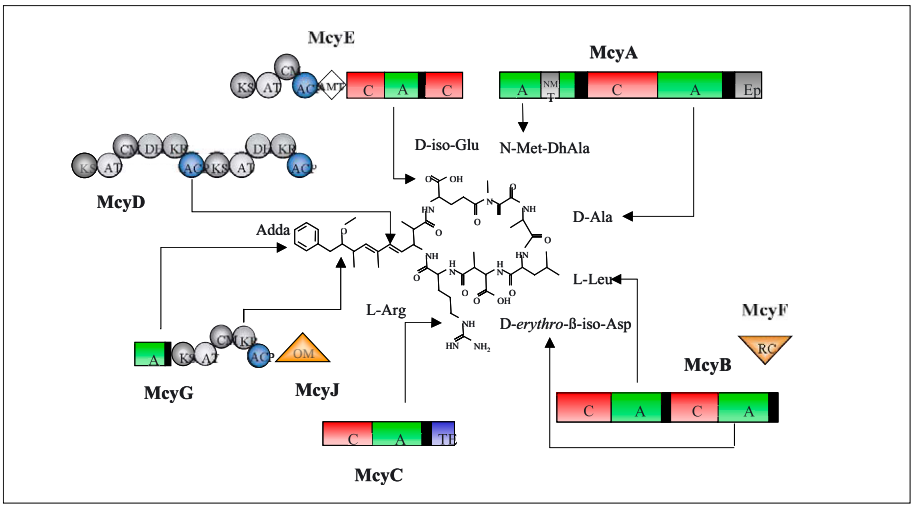

Figure 1. The role of components of the microcystin synthetase (Mcy proteins) deduced from the nucleotide sequence of their genes (Tillett et al., 2000) and functional analyses (Dittmann et al., 1997; Christiansen et al., 2003; Sielaff et al., 2003). PKS domains: AT: acyltransferase; ACP: acyl carrier protein; KS: ß-ketoacyl synthase, KR: ketoacyl reductase; DH: dehydratase; CM: C-methyltransferase; AMT: aminotransferase; NRPS domains: A: aminoacyl adenylation; C: condensation; NMT: N-methyltransferase; Ep: epimerase, TE: thioesterase; McyF: racemase, OM (McyJ): O-methyltransferase. Black bars represent the thiolation motif of NRPS modules. Arrows indicate the assignment of individual proteins to steps of microcystin biosynthesis.

The Adda-D-Glu precursor is synthesised by the activities of at least four enzymes, McyG, J, D and E. McyG is a hybrid enzyme that combines an unusual NRPS adenylation domain with similarity to Acyl-CoA ligases with a typical PKS elongation module. McyG presumably starts microcystin synthesis by activation of phenylacetate followed by transfer to the 4-phosphopantetheine of the first carrier domain. Subsequently, Adda is formed by the four PKS modules of McyG, D and E. McyJ is needed for an O-methylation step. McyE is a hybrid protein composed of PKS and NRPS modules. An integrated domain with similarity to glutamate semialdehyde aminotransferases most likely provides the β-amino group of the

Adda moiety. Finally, the first condensation domain of McyE condenses Adda with the activated glutamate thereby linking the PKS with the NRPS part of microcystin biosynthesis. The NRPS modules of McyA, B and C activate the remaining five amino acids and incorporate them into the growing peptide structure. It has been proposed that the TE domain of McyC (most probably together with a separate enzyme; Christiansen et al., 2003) is responsible for cyclisation and release of microcystin from the synthetase complex. McyF is a racemase thought to provide D-aspartate and most likely also D-methyl-aspartate (Sielaff et al., 2003). The role of McyI remains to be determined (Tillett et al., 2000). Since McyH shows significant similarity to exporters of the ABC transporter family it might play a role in microcystin export (Pearson et al., 2004).

More than 65 structural variants of microcystin have been described. Usually, the individual strains produce more than one microcystin variant (Sivonen and Jones, 1999). This observation raises the question as to whether the biosynthesis of each variant needs its own set of *mcy* genes (then a strain that produces e.g. three variants of microcystin should possess three *mcy* genes clusters) or if one and the same cluster codes for the enzymes needed to synthesise all microcystin variants of a strain. Mutation of single *mcy* genes in different *Microcystis* strains has led in each case to mutant clones that were no longer capable of producing any kind of microcystin. This clearly demonstrates that one *mcy* gene cluster is responsible for the synthesis of all microcystin variants in a given strain (Dittmann et al., 1997; Nishizawa et al., 1999; 2000; Tillett et al., 2000). The substrate specificity of non-ribosomal peptide synthetases is determined by the amino acid sequence of the adenylation domain. Therefore, the ability of strains to produce more than one isoform of microcystin is likely caused by relaxed substrate specificity of the adenylation domains of the involved peptide synthetases (see below 6.). Mikalsen et al. (2003) could identify two groups of sequence variants of an adenylation domain of the *mcyB* gene. Presence of one or the other variant in *Microcystis* strains correlated with the formation of microcystin-LR isoforms and microcystin-RR (in a subgroup in combination with microcystin-LR), respectively. Rouhiainen et al. (2004) observed a higher variability in the specifying amino acids of adenylation domains that are responsible for the incorporation of amino acids at variable positions of microcystins as compared with the other adenylation domains in *mcy* genes (see also 6. below).

3 Distribution of *mcy* genes in cyanobacterial communities

The detection of NRPS genes in cyanobacteria (Meissner et al., 1996) and of *mcy* genes in *Microcystis* (Dittmann et al., 1997) led to investigations into the distribution of these genes among cyanobacteria. Primers deduced from NRPS sequences were used in PCR studies and revealed the presence of NRPS genes in about 75% of 146 axenic strains from the Pasteur Culture Collection representing all traditional botanical orders of cyanobacteria, the *Chroococcales, Pleurocapsales, Oscillatoriales, Nostocales,* and *Stigonematales* (Christiansen et al., 2001). With the help of *mcy*-specific primers, the existence of genes for microcystin biosynthesis

was detected in *Microcystis, Planktothrix (Oscillatoria), Anabaena* and *Nostoc,* and of related genes in *Nodularia* (Neilan et al., 1999). At the same time it became evident that some closely related strains (genotypes) did not contain *mcy* genes and that there was a good correlation between the presence of *mcy* genes and the ability to produce microcystin(s) and also between the lack of these genes and the inability to synthesise microcystin(s). Thus, it was recognised that PCR with *mcy*-specific primers provided a way of discriminating between potentially hepatotoxic and non-hepatotoxic strains (genotypes) and a method for detecting potentially harmful cyanobacteria in water bodies (see Ouellette and Wilhelm, 2003 for review). A combination of *mcy*-specific PCR with RFLP analysis readily identifies potential microcystin producers in environmental samples (Fig. 2).

It is anticipated that the knowledge of genes for microcystin production and the biosynthesis of other cyanobacterial toxins will be important for the molecular ecology of harmful cyanobacteria. The sensitivity of PCR will permit the analysis of even single colonies or filaments. Real-time PCR with *mcy*-specific primers will provide data on the quantitative distribution of toxic and non-toxic genotypes in populations (Kurmayer and Kutzenberger, 2003; Vaitomaa et al., 2003). The development of DNA chips has been initiated containing genus-, species-, and also *mcy*-specific sequences. The DNA chips are expected to provide information, by simple hybridisation experiments, on the composition of cyanobacterial

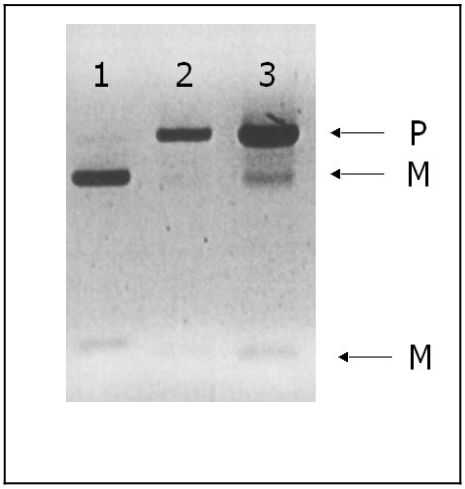

Figure 2. Identification of Microcystis *and* Planktothrix *as microcystin-producers in a water sample from lake Wannsee (Berlin). DNA was isolated from axenic* M. aeruginosa *(lane 1) and* P. agardhii *strains (lane 2) and from the environmental sample (lane 3);* mcy*-specific fragments were amplified by PCR, digested by a restriction endonuclease (*Eco*RV) and electrophoretically separated on an agarose gel. M and P mark the position of* Microcystis*- and* Planktothrix*-specific fragments, respectively. The DNA-based technique revealed the presence of* mcy*-genes from* Microcystis *and* Planktothrix *in the water sample (Hisbergues et al., 2003).*

communities or blooms and their potential toxicity (EU project MidiChip; http://www.ulg.ac.be/ cingprot/midichip/index.htm).

4 Upregulation of *mcy* gene expression by light

Knowledge of the regulation of cyanotoxin biosynthesis may help prevent the appearance of environmental conditions that support toxin production and provide clues on the cellular function of these substances. It has been suggested that several environmental factors influence the biosynthesis of cyanotoxins. For discussion of environmental effects on toxin production see chapter 3 of this volume.

Studies on transcript accumulation of genes encoding the peptide synthetases involved in microcystin biosynthesis in two unrelated strains of *M. aeruginosa* clearly demonstrated a light-dependent increase in the steady state level of those mRNAs (Nishizawa et al., 1999; Kaebernick et al., 2000) (Fig. 3). No data on the effects of other environmental factors on the expression of *mcy* genes have been reported so far. Interestingly, the positive effect of light was only detectable with cells harvested from the early (OD_{750} 0.4 – 0.7) and middle growth phases (OD_{750} 0.84 – 0.87) of batch cultures, but not from the late phase (OD_{750} 1.6 – 2.9)

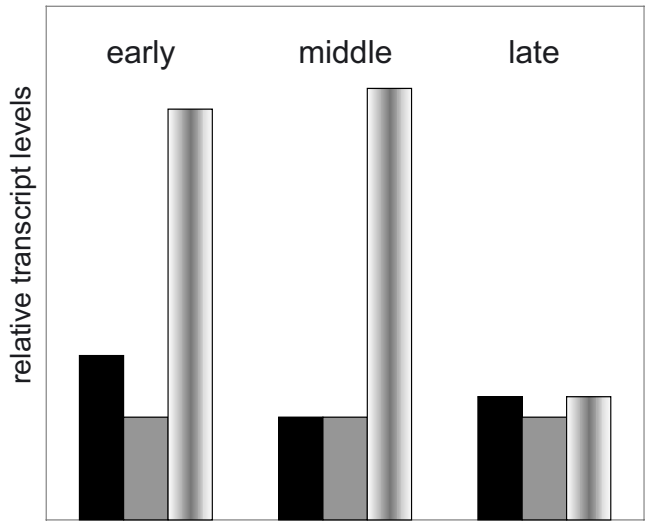

Figure 3. Relative mcyB *transcript levels in* M. aeruginosa *PCC 7806 cells harvested in the early, middle or late growth phase of batch culture. Cultures were illuminated with a photon irradiance of 16 µmol m^{-2} s^{-1} (black column), 31 µmol m^{-2} s^{-1} (grey column), and 68 µmol m^{-2} s^{-1} (shaded column), respectively (Kaebernick et al., 2000).*

(Fig. 3; Kaebernick *et al*. 2000). Shading effects by high cell densities could explain this striking result (Kaebernick *et al.,* 2000). This observation might also indicate an effect of cell density on the transcription of *mcy* genes as known from many genes that are regulated by a quorum sensing mechanism in other bacteria (cf. Dittmann et al., 2001). Studies are under way to discriminate between these alternatives. Since the stimulating effect of light on *mcy* gene transcription could be observed already after a few minutes of illumination of *M. aeruginosa* PCC 7806 cells, light may act directly on transcription and/or transcript stability rather than indirectly, e.g. via a general stimulation of growth (Kaebernick et al., 2000 and unpubl. data). The connection between light, growth and microcystin production is complex and has been the subject of several studies (see chapter 3 of this volume).

The microcystin content could be controlled at the level of transcription, transcript stability, translation, enzyme activities and/or microcystin turnover. Thus, an enhanced *mcy* transcript accumulation after illumination by certain light intensities and qualities (Kaebernick et al., 2000) may not necessarily lead to an enhanced amount of microcystins under the same conditions. However, light has also a marked positive influence on the accumulation of Mcy proteins (shown for McyF in Fig. 4, but also found for other Mcy proteins; M. Hisbergues et al., unpubl. data), which suggests that the enhanced transcript level leads to increased translation. Hence, regulation at the transcriptional level seems to be important for controlling the synthesis of the Mcy proteins and, most probably, also of the toxin.

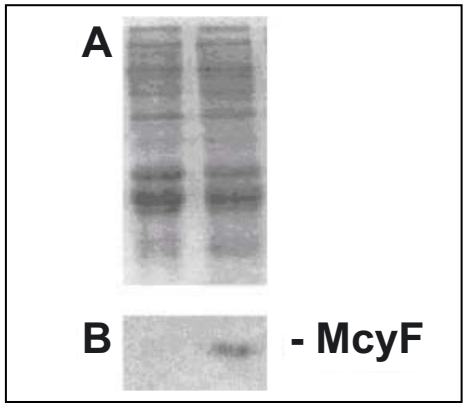

Figure 4. (A) Proteins were isolated from M. aeruginosa *PCC 7806 grown under low light conditions (left lane) or medium light conditions (right lane), electrophoretically separated on SDS polyacrylamide gels and stained by Commassie blue. (B) Western blot of the same gel. McyF was detected by reaction with specific antibodies only in cells grown under medium light conditions (data of M. Hisbergues).*

Studies on the regulation of microcystin production, not to mention the synthesis of other cyanobacterial toxins, are still at their beginning. The presence of a gene encoding a component of a putative ABC transporter in *mcy* gene clusters, and also in the *nda* gene cluster (coding for nodularin biosynthesis), suggests an export of

these hepatotoxins from the producing cells. Future analyses of the effects of environmental factors on toxin production will have to consider both intra- and extracellular microcystins.

5 Comparison of *mcy* genes in different genera – evolutionary aspects

Most microcystin-producing cyanobacteria are strains of the genera *Anabaena, Microcystis* and *Planktothrix*, while the only known nodularin producers are strains of the brackish water species *Nodularia spumigena.* Cyanobacteria of these genera possess gas vesicles and show a tendency to form surface water-blooms. According to their positions in phylogenetic trees based on 16S rDNA sequences, the microcystin producing genera are only distantly related (e.g. Lyra et al., 2001; Tillett et al., 2001; Rantala et al., 2004) and differ also from each other in their morphology. *Microcystis* is a unicellular cyanobacterium that forms characteristic colonies in its natural environment (e.g. Otsuka et al., 2000). *Planktothrix* is a filamentous cyanobacterium; *P. rubescens* usually forms distinct layers while *P. agardhii* is more often evenly distributed in the water column, especially in shallow lakes (e.g. Van Liere and Mur, 1980; Kurmayer et al., 2004). Planktic species of *Anabaena* and *Nodularia* are filamentous cyanobacteria that differentiate heterocysts, cells specialised for nitrogen fixation (e.g. Lehtimäki et al., 2000; Lyra et al., 2001; Golden and Yoon, 2003).

The ability to produce microcystin or nodularin shows a patchy distribution among the cyanobacteria. All four genera have members that contain the toxin biosynthesis genes (in most strains correlated with the ability to produce the toxins), and other members that do not. Toxic and non-toxic strains (genotypes) usually coexist in field populations and usually differ with respect to presence *vs*. absence of the genes needed for toxin biosynthesis in their genomes (see section 3 above). Comparisons of the known gene clusters for microcystin and nodularin biosynthesis from the four investigated genera, *Anabaena, Microcystis, Nodularia,* and *Planktothrix*, allow conclusions to be drawn on how this patchy distribution of toxin biosynthesis genes could have evolved among cyanobacteria. In principle, the *mcy* gene clusters might have been distributed by lateral (horizontal) gene transfer between the genera, species and genotypes within a species. If this occurred, genotypes that were non-producers would be expected never to have possessed the genes for the production of peptide toxins. Alternatively, gene clusters might have evolved from common ancestral genes. Had this occurred, many genera, species and strains that harbour no genes for microcystin or nodularin biosynthesis, although phylogenetically related to the producing taxa, must have once had those genes and subsequently lost them. A combination of both alternatives is conceivable. Complete nucleotide sequences of the *mcy* gene clusters are available from *Anabaena* 90*, Microcystis aeruginosa* PCC 7806 and K-139*,* and *Planktothrix agardhii* CYA 126 in databases; the *nda* gene cluster for nodularin biosynthesis has been sequenced in *Nodularia spumigena* NSOR10 (Fig. 5). Comparison of the structural organisation of the clusters and the sequences of individual genes indicates how microcystin biosynthesis might have evolved.

The multi-enzyme components encoded by the individual gene clusters are similar, but the content and arrangement of the modules and consecutive domains are not identical in the four genera. The general organisation of the clusters and the arrangement of certain genes within the clusters are different, as shown in Figure 5. In *Microcystis*, *Anabaena* and *Nodularia* the genes are transcribed from a central bi-directional promoter region (cf. Kaebernick et al., 2002), whereas in *Planktothrix* all *mcy* genes, except *mcyT,* seem to be transcribed unidirectionally from a promoter located upstream of gene *mcyD*.

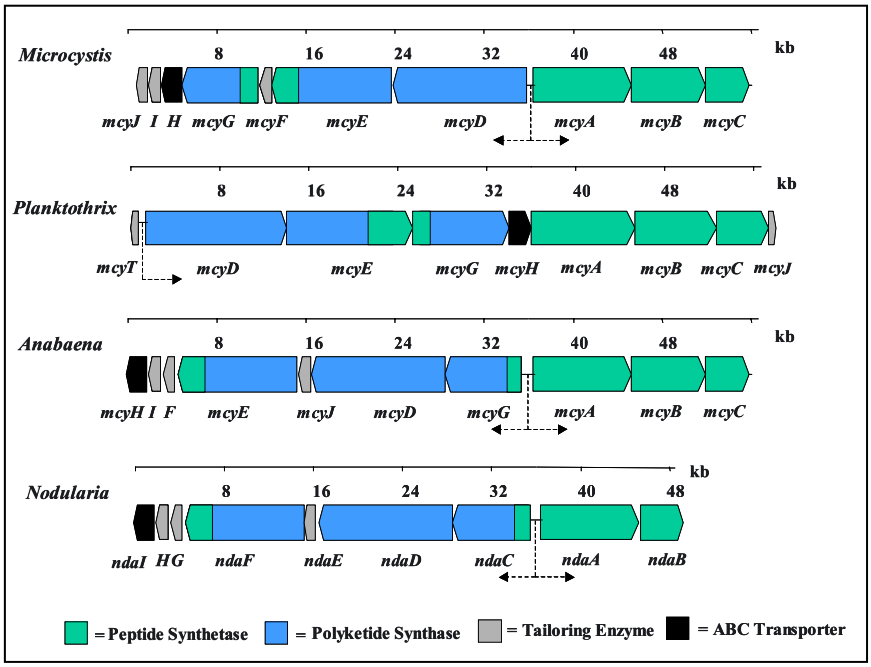

Figure 5. Gene clusters coding for the biosynthesis of microcystin in Microcystis *(Tillett et al., 2000; Nishizawa et al., 2000),* Planktothrix *(Christiansen et al., 2003), and* Anabaena *(Rouhiainen et al., 2004) and of nodularin in* Nodularia *(M.C. Moffitt and B.A. Neilan; accession no. AY210783). Arrows indicate the transcriptional start sites from the putative promoter regions.*

There is a further difference between *Microcystis/Planktothrix* and *Anabaena*/*Nodularia.* NRPS and PKS genes are usually found to follow a "colinearity rule", i.e. the order of genes is the same as the order of the single enzymatic steps of the respective biosynthesis (Marahiel et al., 1997; von Döhren et al., 1997). As described above, microcystin synthesis starts with the step catalysed by McyG, followed by the steps involving McyD, E, A, B and C. An identical arrangement of genes is found in *Anabaena* (Fig. 5). The order of *nda* genes in

Nodularia follows also the "colinearity rule" (Fig. 5). In *Microcystis* and *Planktothrix,* however, the order of the first three genes is *mcyD*, *E* and *G,* i.e. there is a clear deviation from the "colinearity rule". Similarly, the position of tailoring genes and the gene encoding a component of a putative ABC transporter within the cluster is different in the individual genera, except in *Anabaena* and *Nodularia* where, remarkably, it is identical. *Microcystis* still exhibits the same complement of tailoring genes as *Anabaena* and *Nodularia*, but arranged in a different order. *Planktothrix* is most distinctive: two of the tailoring genes (*mcyF, I*) are completely missing, whereas an additional gene, encoding a thioesterase (*mcyT*), was found in close proximity to the biosynthesis gene cluster (Christiansen et al., 2003).

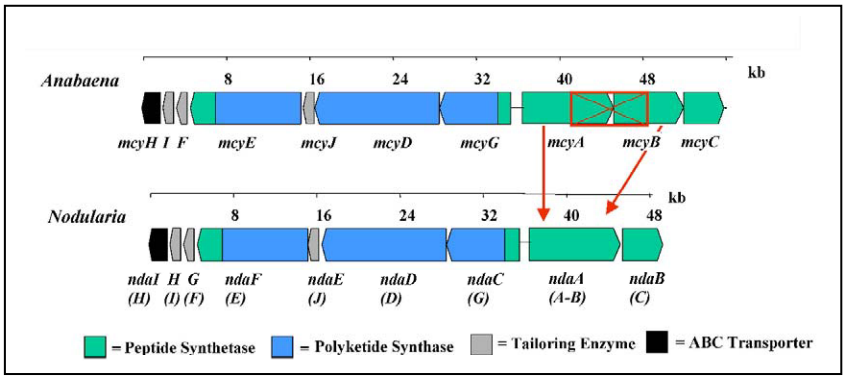

*Figure 6. The microcystin biosynthesis gene (*mcy*) cluster of* Anabaena *90 and the nodularin biosynthesis gene cluster (*nda*) of* Nodularia spumigena *NSOR10 show the same order of genes. The* nda *cluster is proposed to have evolved from an* Anabaena*-type* mcy *cluster by deletion of two NRPS modules from* mcyA *and* B *and fusion of the remaining sequences of* mcyA *and* B *resulting in gene* ndaA. *Letters in brackets indicate the* mcy *gene with homology to the respective* nda *gene.*

Although the overall similarity of the Mcy and Nda multi-enzymes is very high, there is a striking difference between the *mcy* clusters and the *nda* cluster: the *nda* cluster lacks two NRPS modules that are present in all *mcy* clusters. In *Nodularia,* the genes *mcyA* and *B* found separately in the other genera, have been fused together with a simultaneous loss of two NRPS modules resulting in the *ndaA* gene (Fig. 6). The loss of two NRPS modules is consistent with the fewer amino acid residues (five *vs.* seven) in nodularin compared to microcystin (Fig. 7; Annila et al., 1996). Despite the loss of two modules, microcystin/nodularin biosynthesis gene clusters in *Anabaena* and *Nodularia* are the most similar, as expected from the position of these genera in phylogenetic trees (Iteman et al., 2002; Janson and Graneli, 2002; Rantala et al., 2004). It is, therefore, proposed that the nodularin gene cluster evolved from a microcystin biosynthesis gene cluster of the *Anabaena*-type by one major step, the deletion of two NRPS modules (cf. Rantala et al., 2004).

In the case of *Microcystis*, *mcy* clusters of two strains have been sequenced and found to have the same gene order and a high similarity in their nucleotide sequences (Dittmann et al., 1997; Nishizawa et al., 1999 and 2000; Tillett et al., 2000). In contrast, several recombination and many mutation events must have occurred to lead to the differences in structural organisation and nucleotide sequence between the *mcy* gene clusters in *Anabaena, Microcystis* and *Planktothrix* (Fig. 6; Nishizawa et al., 2000; Tillett et al., 2000; Christiansen et al., 2003; Rouhiainen et al., 2004). These differences make a lateral transfer of *mcy* gene clusters between genera unlikely and rather suggest an independent evolution over a long time span. Important data concerning the evolution of *mcy* genes were obtained by analysing the phylogenetic relationship of conserved *mcy* sequences and of sequences of housekeeping genes. The results of this study are not in agreement with lateral gene transfer between the genera, but strongly support the idea of a common ancestor of the *mcy* genes in *Anabaena, Microcystis, Nostoc* and *Planktothrix* (Rantala et al., 2004). Microcystin biosynthesis, therefore, appears to be a very old pathway of secondary metabolism. The lack of *mcy* (or *nda*) genes in many cyanobacterial taxa must consequently be due to losses of these genes during evolution (Rantala et al., 2004). The loss of *mcy* genes seem to be an ongoing process that might start with a loss of function by mutation. There are reports of *Microcystis* and *Planktothrix* genotypes (strains) that contain *mcy* genes, but do not produce microcystin. Whereas such genotypes seem to be rare in *Microcystis* (Mikalsen et al., 2003; Nishizawa et al., 1999; Tillett et al., 2001), they have been more frequently observed in some Austrian (Kurmayer et al., 2004) and German lakes (S. Mbedi and C. Wiedner, pers. commun.). Preliminary data suggest that their inability to produce microcystin is due to mutation of *mcy* gene(s) (Kurmayer et al., 2004).

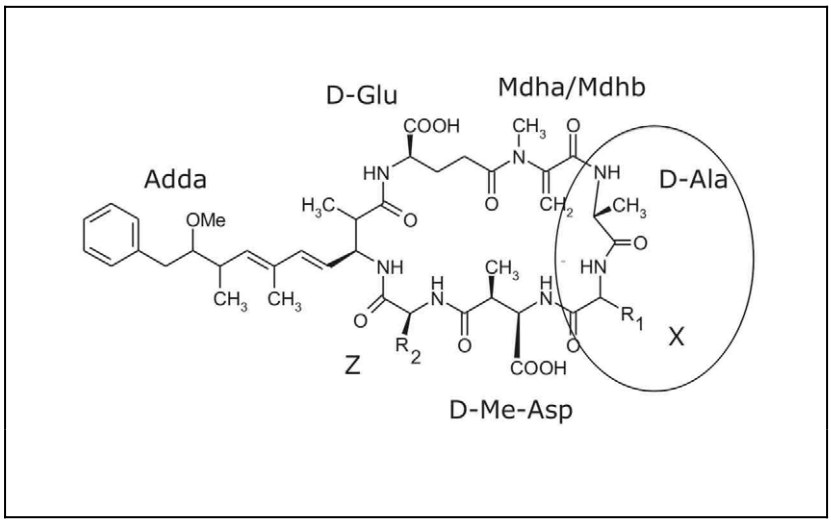

Figure 7. The molecular structure of microcystins and nodularins is similar. Nodularins lack the two amino acids shown in the oval and have Mdhb instead of Mdha.

6 Structure and function of adenylation domains

Further insight into the evolution of microcystin biosynthesis can be expected from analyses of the sequences of individual *mcy* genes and their modules and domains (Christiansen et al., 2003; Kurmayer et al., 2001; Mikalsen et al., 2003; Rouhiainen et al., 2004). A fast evolution should occur of those enzymes and domains that contribute to the observed large variation in microcystin structure. Most of this variation is due to differences in individual amino acid positions, whereas the polyketide side chain of the Adda moiety is highly conserved. The specificity of the so-called adenylation (A) domain determines which amino acid(s) are activated by an individual NRPS module.

A-domains specifically recognise their substrate amino acids and form an amino-acyl-AMP intermediate. Analysis of the crystal structure of the phenylalanine activating domain of the tyrocidine synthetase from *Bacillus subtilis* (PheAT), as well as comparison of the growing number of NRPS adenylation domains, have led to the development of a "specificity-conferring code of the amino acids" (Stachelhaus et al., 1999; Challis et al., 2000). Eight single residues of the amino acid binding pocket were proposed to discriminate between the different amino acid substrates. This code can be used for the prediction of the specificity of NRPSs that have not been characterised.

The four individual strains that have been used for sequencing microcystin biosynthesis gene clusters show different complements of microcystin variants. Another two strains have been investigated that produce different variants of nodularin. Differences occur at four amino acid positions, namely the Mdha/Mdhb position, the L-Leu/L-Arg position, the D-Asp/D-MeAsp position and the L-Arg/L-Har position (Table 1). The different specificities can at least in part be recognised by comparing the "amino acid specificity codes" of the corresponding A-domains. The first A-domain in McyA is supposed to activate serine as the precursor of the Mdha moiety. All microcystin producing strains conform to the code, i.e. show high similarity to serine-activating domains in other bacteria. The analogous NdaAA1 domain, however, shows a clear threonine code, which is expected to be the precursor of the Mdhb moiety. The McyBA1 domain seems to either recognise only L-leucine (as seen for the two *Microcystis* examples) or to be multispecific for L-leucine and L-arginine. In the first case a "leucine code" can be observed, whereas the *Planktothrix* CYA 126 and *Anabaena* 90 McyBA1 domains are not homologous to other NRPS A-domains, but even show some variation between each other. McyBA2 activates D-aspartate and D-methyl-aspartate in all strains investigated, except *Planktothrix* CYA 126, which seems to accept only D-aspartate as a substrate. It is not yet clear, whether this is related to the single amino acid difference in the corresponding codes (Table 1). Alternatively, the enzyme providing D-methyl-aspartate as a substrate might be missing in *Planktothrix*. Strikingly, McyCA1 activates L-arginine in all microcystin-producing strains, even though in all four cases no clear "arginine code" can be observed. The *Nodularia* strains investigated incorporate L-Arg (*Nodularia spumigena* NSOR10) or L-homoarginine (Har, *Nodularia harveyana* PCC7804) in the nodularin structure. There is a single amino acid difference in the specificity code between the two *Nodularia* strains that

could account for the deviating substrate specificity (Table 1), but biochemical evidence is needed to verify this assumption.

Differences in the specificities of A-domains can be deduced not only from the few amino acids determining the 'specificity codes', but also by comparing complete sequences of binding pockets. In contrast to the major part of the Mcy complex, these domains are highly variable in their sequence. Recombination

Table 1. Adenylation domains of microcystin synthetase and nodularin synthetase gene clusters: Comparison of specificity conferring amino acids in the substrate binding pocket, predicted substrates and real substrate. (Org.: organism; Ma = Microcystis aeruginosa *PCC7806, Pa* = Planktothrix agardhii*, A90* = Anabaena *90, Ns* = Nodularia spumigena *NSOR10, Nh* = N. harveyana *PCC7804. Ref.: reference; 1 = Tillett et al., 2000, 2 = Christiansen et al., 2003, 3 = Rouhiainen et al., 2004, 4 = accession no. AY210783)*

A domain	*Org*	*Binding pocket*	*Predicted substrate*	*Substrate*	*Ref*
McyAA1	*Ma*	DVWHFSLI	Ser	L-Ser	1
McyAA1	*Pa*	DVWHISLI	Ser	L-Ser	2
McyAA1	*A*90	DVWHISLI	Ser	L-Ser	3
NdaAA1	*Ns*	DFWNIGMV	Thr	L-Thr	4
McyAA2	*Ma*	DLFNNALT	Gly	L-Ala	1
McyAA2	*Pa*	DLFNNALS	Gly	L-Ala	2
McyAA2	*A*90	DLFNNALT	Gly	L-Ala	3
---	*Ns*	— —	--	--	
McyBA1	*Ma*	DVWFLGNV	Leu	L-Leu	1
McyBA1	*Pa*	DALFFGLV	--	L-Arg/L-Leu	2
McyBA1	*A*90	DVWFFGLV	--	L-Arg/L-Leu	3
---	*Ns*	— —	--	--	
McyBA2	*Ma*	DARHVGIV	--	D-MeAsp/D-Asp	1
McyBA2	*Pa*	DPRHVGIF	--	D-Asp	2
McyBA2	*A*90	DARHVGIF	--	D-MeAsp/D-Asp	3
NdaAA2	*Ns*	DARHVGIF	--	D-MeAsp	4
McyCA1	*Ma*	DVWTIGAV	--	L-Arg	1
McyCA1	*Pa*	DPWVFGLV	--	L-Arg	2
McyCA1	*A*90	DVWVFGLV	--	L-Arg	3
NdaBA1	*Ns*	DVWNFGFV	--	L-Arg	4
NdaBA1	*Nh*	DVWSFGFV	--	L-Har	4
McyEA1	*Ma*	DPRHSGVV	--	D-Glu	1
McyEA1	*Pa*	DPRHSGVV	--	D-Glu	2
McyEA1	*A*90	DPRHSGVV	--	D-Glu	3
NdaFA1	*Ns*	DPRHSGVV	--	D-Glu	4
McyGA1	*Ma*	VGIWVAAS	--	Phenylacetate	1
McyGA1	*Pa*	VGIWVAAS	--	Phenylacetate	2
McyGA1	*A*90	VGIWVAAS	--	Phenylacetate	3
NdaCA1	*Ns*	VGIWVAAS	--	Phenylacetate	4

between different modules of the *mcy* gene cluster seems to contribute to the observed sequence variability (Christiansen et al., 2003; Mikalsen et al., 2003; Rouhiainen et al., 2004). Since individual strains usually produce several different peptides and polyketides, they should also harbour several NRPS and PKS gene clusters. It remains to be determined to what extent an exchange of DNA sequences between the different clusters has contributed to their evolution.

The evolution of *mcy* gene clusters is a dynamic and continuing process. Differences in the general structure of the *mcy* clusters and co-evolution of conserved *mcy* sequences with house-keeping genes do not favour lateral gene transfer between the genera and clearly point to the early appearance of microcystin biosynthesis in the evolution of cyanobacteria (see above). There has been rapid evolution of certain regions of the *mcy* cluster that bear the genetic information for the variability of microcystins. This evolution has occurred 'conventionally' by mutation and by recombination between modules and domains, though a contribution of intraspecific lateral transfer to the evolution of these sequences has not been ruled out.

7 Acknowledgements

We thank Michael Hisbergues for providing the data shown in Figure 4 and Tony Walsby for critically reading the manuscript. The work of the authors was supported by grants from the European Union (CYANOTOX-ENV4-CT98-0802 and TOPIC-FMRX-CT98-0246), the Deutsche Forschungsgemeinschaft (Bo 1045/13-3) and the Fond der Chemischen Industrie.

8 References

Annila, A., Lehtimäki, J., Mattila, K., Eriksson, J.E., Sivonen, K., Rantala, T.T. and Drakenberg, T. (1996) Solution structure of nodularin: an inhibitor of serine/threonine-specific protein phosphatases, *Journal of Biological Chemistry* **27**, 16695-16702.

Arment, A.R. and Carmichael, W.W. (1996) Evidence that microcystin is a thio-template product, *Journal of Phycology* **32**, 591-597.

Bishop, C.T., Anet, E.F.L.J. and Gorham, P.R. (1959) Isolation and identification of the fast-dead factor in *Microcystis aeruginosa* NRC-1, *Canadian Journal of Biochemistry and Physiology* **37**, 453-458.

Botes, D.P., Tuiman, A.A., Wessels, P.L., Viljoen, C.C., Kruger, H., Williams, D.H., Santikarn, S., Smith, R.J. and Hammond, S.J. (1984) The structure of cyanoginosin-LA, a cyclic peptide from the cyanobacterium *Microcystis aeruginosa*, *Journal of the Chemical Society - Perkin Transactions* **I**, 2311-2319.

Challis, G.L., Ravel, J. and Townsend, C.A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, *Chemistry and Biology* **7**, 211-224.

Christiansen, G., Dittmann, E., Ordorika, L.V., Rippka, R., Herdman, M. and Börner, T. (2001) Nonribosomal peptide synthetase genes occur in most cyanobacterial genera as evidenced by their distribution in axenic strains of the PCC, *Archives of Microbiology* **76**, 452-458.

Christiansen, G., Fastner, J., Erhard, M., Börner, T. and Dittmann, E. (2003) Microcystin biosynthesis in *Planktothrix*: genes, evolution. and manipulation, *Journal of Bacteriology* **185**, 564-572.

Dittmann, E., Meissner, K. and Börner, T. (1996) Conserved sequences of peptide synthetase genes in the cyanobacterium *Microcystis aeruginosa, Phycologia* **35**, 62-67.

Dittmann, E., Neilan, B.A., Erhard, M., von Döhren, H. and Börner, T. (1997) Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806, *Molecular Microbiology* **26**, 779-787.

Dittmann, E., Erhard, E., Kaebernick, M., Scheler, C., Neilan, B.A., von Döhren, H. and Börner, T. (2001) Altered expression of two light-dependent genes in a microcystin-lacking mutant of *Microcystis aeruginosa* PCC 7806, *Microbiology* **147**, 3113-3119.

Golden, J.W. and Yoon, H.S. (2003) Heterocyst development in *Anabaena*, *Current Opinion in Microbiology* **185**, 6995-7000.

Hisbergues, M., Christiansen, G., Rouhiainen, L., Sivonen, K. and Börner, T. (2003) PCR-based identification of microcystin-producing genotypes of different cyanobacterial genera, *Archive for Microbiol*ogy **180,** 402-410.

Hopwood, D.A. (1997) Genetic contributions to understanding polyketide synthases, *Chemical Reviews* **97**, 2465-2497.

Iteman, I., Rippka, R., Tandeau de Marsac, N. and Herdman, M. (2002) rDNA analyses of planktonic heterocystous cyanobacteria, including members of the genera *Anabaenopsis* and *Cyanospira, Microbiology* **148**, 481-496.

Janson, S. and Graneli, E. (2002) Phylogenetic analyses of nitrogen-fixing cyanobacteria from the Baltic Sea reveal sequence anomalies in the phycocyanin operon, *International Journal of Systematic and Evolutionary Microbiology* **52**, 1397-1404.

Kaebernick, M., Neilan, B.A., Börner, T. and Dittmann, E. (2000) Light and the transcriptional response of the microcystin gene cluster, *Applied and Environmental Microbiology* **66**, 3387-3392.

Kaebernick, M., Dittmann, E., Börner, T. and Neilan, B.A. (2002) Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide, *Applied and Environmental Microbiology* **68**, 449-455.

Khosla, C. (1997) Harnessing the biosynthetic potential of modular polyketide synthases, *Chemical Reviews* **97**, 2577-2590.

Kurmayer, R., Dittmann, E., Fastner, J. and Chorus, I. (2001) Diversity of microcystin genes in lake Wannsee (Germany), *Microbial Ecology* **43**, 107-118.

Kurmayer, R. and Kutzenberger, T. (2003) Application of real-time PCR for quantitation of microcystin genotypes in a population of the toxic cyanobacterium *Microcystis sp.*, *Applied and Environmental Microbiology* **69**, 6723-6730.

Kurmayer, R., Christiansen, G., Fastner, J. and Börner, T. (2004) Abundance of active and inactive microcystin genotypes in populations of the toxic cyanobacterium *Planktothrix* spp., *Environmental Microbiology*, in press.

Lehtimäki, J., Lyra, C., Suomalainen, S., Sundman, P., Rouhiainen, L., Paulin, L., Salkinoja-Salonen, M. and Sivonen, K. (2000) Characterization of *Nodularia* strains, cyanobacteria from brackish waters, by genotypic and phenotypic methods, *International Journal of Systematic and Evolutionary Microbiology* **50***,* 1043-1053.

Lyra, C., Suomalainen, L.C., Gugger, M., Vezie, C., Sundman, P., Paulin, L. and Sivonen, K. (2001) Molecular characterization of planktic cyanobacteria of *Anabaena*, *Aphanizomenon*, *Microcystis* and *Planktothrix* genera*, International Journal of Systematic and Evolutionary Microbiology* **51**, 513-526.

Marahiel, M.A., Stachelhaus, T. and Mootz, H.D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis, *Chemical Reviews* **97**, 2651-2673.

Meissner, K., Dittmann, E. and Börner, T. (1996) Toxic and non-toxic strains of the cyanobacterium *Microcystis aeruginosa* contain sequences homologous to peptide synthetase genes, *FEMS Microbiology Letters* **135**, 295-303.

Mikalsen, B., Boison, G., Skulberg, O.M., Fastner, J., Davies, W., Gabrielsen, T.M., Rudi, K. and Jacobsen, K.S. (2003) Natural variation in the microcystin synthetase operon *mcyABC* and impact on microcystin production in *Microcystis* strains, *Journal of Bacteriology* **185**, 2774-2785.

Moore, R.E., Chen, J.L., Moore, B.S., Patterson, G.M.L. and Carmichael, W.W. (1991) Biosynthesis of microcystin-LR: origin of carbons in the Adda and Masp units, *Journal of the American Chemical Society* **113**, 5083-5084.

Neilan, B.A., Dittmann, E., Rouhiainen, L., Bass, R.A., Schaub, V., Sivonen, K. and Börner, T. (1999) Nonribosomal peptide synthesis and toxigenicity of cyanobacteria*, Journal of Bacteriology* **181**, 4089-4097.

Nishizawa, T., Asayama, M., Fujii, K., Harada, K. and Shirai, M. (1999) Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in *Microcystis* ssp., *Journal of Biochemistry* **126**, 520-526.

Nishizawa, T., Ueda, A., Asayama, M., Fujii, K., Harada, K.-I., Ochi, K. and Shirai, M. (2000) Polyketide synthase gene coupled to the peptide synthetase module involved in the biosynthesis of the cyclic heptapeptide microcystin, *Journal of Biochemistry* **127**, 779-789.

Otsuka, S., Suda, S., Matsumoto, S. and Watanabe, W.W. (2000) Morphological variability of colonies of *Microcystis* morphospecies in culture, *Journal of General and Applied Microbiology* **46**, 39-50.

Ouellette A.J.A. and Wilhelm, S.W. (2003) Toxic cyanobacteria: the evolving molecular toolbox, *Frontiers for Ecology and Environment* **1**, 359-366.

Pearson, L., Hisbergues, M., Dittmann, E., Börner, T. and Neilan, B.A. (2004) Inactivation of an ABC transporter, McyH, results in loss of microcystin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806, *Applied and Environmental Microbiology*, in press.

Rantala, A., Fewer, D.P., Hisbergues, M., Rouhiainen, L., Vaitomaa, J., Börner, T. and Sivonen, K. (2004) Phylogenetic evidence for the early evolution of microcystin synthesis, *Proceedings of the National Academy of Sciences (USA)* **101**, 568-573.

Rinehart, K.L., Harada, K.I., Namikoshi, M., Chen, C., Harvis, C.A., Munroe, M.H.G., Blunt, J.W., Mulligan, P.E., Beasley, V.R., Dahlem, A.M. and Carmichael, W.W. (1988) Nodularin, microcystin, and the configuration of Adda, *Journal of the American Chemical Society* **110**, 8557-8558.

Rouhiainen, L., Paulin, L., Suomalainen, S., Hyytiainen, H., Buikema, W., Haselkorn, R. and Sivonen, K. (2000) Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90, *Molecular Microbiology* **37**, 156-167.

Rouhiainen, L., Vakkilainen, T., Lumbye Siemer, B., Buikema, W., Haselkorn, R. and Sivonen, K. (2004) Genes coding for the synthesis of hepatotoxic heptapeptides (microcystins) in the cyanobacterium *Anabaena* strain 90, *Applied and Environmental Microbiology* **70**, 686-692.

Schembri, M.A., Neilan, B.A. and Saint, C.P. (2001) Identification of genes implicated in the toxin production in the cyanobacterium *Cylindrospermopsis raciborskii*, *Environmental Toxicology* **16**, 413-421.

Sielaff, H., Dittmann, E., Tandeau de Marsac, N., Bouchier, C., von Döhren, H., Börner, T. and Schwecke, T. (2003) The *mcyF* gene of the microcystin biosynthetic gene cluster from *Microcystis aeruginosa* encodes an aspartate racemase, *Biochemical Journal* **373**, 903-916.

Sivonen, K. and Jones, G.J. (1999) Cyanobacterial toxins, in I. Chorus and J. Bartram (eds.), *Toxic Cyanobacteria in Water: A Guide to their Public Health Consequences, Monitoring and Management*, Spon, London, pp. 41-111.

Stachelhaus, T., Mootz, H.D. and Marahiel, M.A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases, *Chemistry and Biology* **6**, 493-505.

Tillett, D., Dittmann, E., Erhard, M., von Döhren, H., Börner, T. and Neilan, B.A. (2000) Structural organisation of microcystin biosynthesis in *Microcystis aeruginosa* PCC 7806: an integrated peptide-polyketide synthetase system, *Chemistry and Biology* **7**, 753-764.

Tillett, D., Parker, D.L. and Neilan, B.A. (2001) Detection of toxigenicity by a probe for the microcystin synthetase A gene (*mcyA*) of the cyanobacterial genus *Microcystis*: comparison of toxicities with 16S rRNA and phycocyanin operon (phycocyanin intergenic spacer), *Applied and Environmental Microbiology* **67**, 2810-2818.

Turgay, K. and Marahiel, M.A. (1994) A general approach for identifying and cloning peptide synthetase genes, *Peptide Research* **7**, 753-764.

Vaitomaa, J., Rantala, A., Halinen, K., Rouhiainen, L., Tallberg, P., Mokelke, L. and Sivonen, K. (2003) Quantitative real-time PCR for determination of microcystin synthetase E copy numbers for *Microcystis* and *Anabaena* in lakes, *Applied and Environmental Microbiology* **69**, 7289-7297.

Van Liere, L. and Mur, L.R. (1980) Occurrence of *Oscillatoria agardhii* and some related species, a survey, in J. Barica and L.R. Mur (eds.), *Hypertrophic Ecosystems*; *Developments in Hydrobiology* **2**, 67-77.

Von Döhren, H., Keller, U., Vater, J. and Zocher, R. (1997) Multifunctional peptide synthetases, *Chemical Reviews* **97**, 2675-2705.