Abstract: Hepatotoxic microcystins (MCs) are the most widespread class of cyanotoxins and the one that has most often been implicated in cyanobacterial toxicosis. One of the main challenges in studying and monitoring MCs is the great structural diversity within the class. The full chemical structure of the first MC was elucidated in the early 1980s and since then the number of reported structural analogues has grown steadily and continues to do so, thanks largely to advances in analytical methodology. The structures of some of these analogues have been definitively elucidated after chemical isolation using a combination of techniques including nuclear magnetic resonance, amino acid analysis and tandem mass spectrometry (MS/MS). Others have only been tentatively identified using liquid chromatography-MS/MS without chemical isolation. An understanding of the structural diversity of MCs, the genetic and environmental controls for this diversity and the impact of structure on toxicity are all essential to the ongoing study of MCs across several scientific disciplines. However, because of the diversity of MCs and the range of approaches that have been taken for characterizing them, comprehensive information on the state of knowledge in each of these areas can be challenging to gather. We have conducted an in-depth review of the literature surrounding the identification and toxicity of known MCs and present here a concise review of these topics. At present, at least 269 MCs have been reported. Among these, about 20% (54 of 269) appear to be the result of chemical or biochemical transformations of MCs that can occur in the environment or during sample handling and extraction of cyanobacterial, including oxidation products, methyl esters, or post-biosynthetic metabolites. The toxicity of many MCs has also been studied using a range of different approaches and a great deal of variability can be observed between reported toxicities, even for the same congener. This review will help clarify the current state of knowledge on the structural diversity of MCs as a class and the impacts of structure on toxicity, as well as to identify gaps in knowledge that should be addressed in future research.

Keywords: microcystin; cyanobacteria; cyanotoxin; structural elucidation; toxicology
1. Introduction

There are an increasing number of warnings about toxic cyanobacterial blooms observed worldwide and global warming is thought to stimulate their development in eutrophic waters [1-5]. These blooms are often accompanied by production of a variety of cyanotoxins generally classified according to the human organs on which they act: hepatotoxins (liver), neurotoxins (nervous system) and dermatoxins (skin). Among these cyanotoxins, it appears that hepatotoxic and tumor promoting microcystins (MCs) are more commonly found in cyanobacterial blooms and considered to be one of the most hazardous groups throughout the world [6-10]. Despite the significant amount of available information, interest in MCs continues to increase due to their well-known hazards to farm animals, fisheries, aquaculture, human health, and wildlife through exposure via drinking, environmental and recreational waters [7,8,10-12].

Louw [13] was one of the first scientists who isolated and attempted to characterize a cyanotoxin from a cyanobacterial bloom, dominated by Microcystis toxica, which occurred in the Vaal Dam reservoir in South Africa in 1942–3. He concluded that it was an alkaloid of undetermined structure with acute and chronic hepatotoxic properties. Initial attempts to characterize the structures of MCs started in the late 1950s [14]. However, full structural identification of the first microcystin congeners was achieved in 1984, when a combination of amino acid analysis, nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) were used [15,16]. These and subsequent studies showed the chemical structure of these toxins to consist of a cyclic heptapeptide composed of five relatively conserved amino acids plus two variable L-amino acids [14-20].

Two decades later, studies have shown that these toxins are biosynthesized nonribosomally via an MC synthetase gene cluster (mcy), consisting of a combination of polyketide synthases (PKSs), nonribosomal peptide synthetases (NRPSs) and tailoring enzymes [21-24]. Recent studies have reported the bulk of this structural diversity is the result of genetic and/or environmental factors, which have an impact on the functioning of enzymes encoded in the mcy gene cluster [25-32]. This mcy gene cluster is often spontaneously modified through point mutations, deletions and insertions or a series of genetic recombination, which affect the functioning of the microcystin peptide synthetases and result in the chemical diversity observed in nature [33-37].

Because of the large number of literature reports spanning over four decades of research, it can be difficult to glean accurate information on the total number of identified MCs. The phrase ‘more than 100 microcystin congeners’ is still often used in the literature [12,38-41] however estimates as high as 248 known MCs have recently been published [42]. In this review, we update this number to 269 and describe the methods by which this structural elucidation was carried out. In addition, we provide an in-depth review of their toxic potential and a discussion of the structure–activity relationships this information provides.

2. Nomenclature and general chemical structure of microcystins

Hughes et al. [43] first described a hepatotoxic ‘fast death factor’ in an extract of the isolated strain Microcystis aeruginosa NRC-1. This hepatotoxic factor was later renamed microcystin, derived from the genus Microcystis [44-46]. Since, they have also been referred to in the literature as cyanoginosin, with prefix ‘cyano’ from the term cyanobacteria and ‘ginosin’ derived from aeruginosa [19,47]; and cyanoviridin, with the root ‘viridin’ from the species M. viridis [48].
After almost two decades of structural analysis of toxic peptides from the colonial bloom-forming cyanobacterium *M. aeruginosa*, Botes et al. [15-19] and Santikarn et al. [20] provided structural details of these toxins with the general structure of: cyclo-(D-Ala₁-X²-D-Masp³-Z⁴-Adda⁵-D-γ-Glu⁶-Mdha⁷) (Figure 1) in which X and Z are variable L-amino acids, D-Masp³ is D-erythro-β-methyl-isoaspartic acid, and Mdha is N-methyldehydroalanine. The β-amino acid Adda, (2S,3S,4E,6E,8S,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid, is the most unusual substructure, and has not been reported elsewhere in nature apart from in the closely related group of toxic pentapeptides, the nodularins. After a time, leading researchers in the field warned that the continued use of multiple naming systems could generate confusion as the number of publications on these cyclic peptides increased [49]. Therefore, a universal system of nomenclature was proposed based on the original term microcystin-XZ, where X and Z denote the variable amino acid residues in positions 2 and 4 [50].

The number of fully and partially characterized MCs increased significantly throughout the 1990s and structural variations have been reported in all seven amino acids, but most frequently with substitution of the two variable L-amino acids X and Z at positions 2 and 4, respectively, and demethylation of amino acids at position 3 and/or 7. Multiple combinations of the variable L-amino acids (X and Z), and modifications to the other amino acids, lead to the high levels of structural diversity observed in MCs which have been characterized to date from bloom samples and isolated strains of *Microcystis, Oscillatoria, Planktothrix, Anabaena* (syn. *Dolichospermum*), and *Nostoc* spp.. While Carmichael et al. [50] originally proposed abbreviation in the format MCYST-XZ, in recent decades the abbreviation MC-XZ has come in to general use, where X and Z are the one-letter amino acid abbreviations (where these exist) and with any variations at positions 1, 3, and 5–7 (relative to D-Ala, D-Masp, Adda, D-Glu and Mdha, respectively) shown in square brackets, using 3–7-letter amino acid abbreviations [51], in numerical order with the position indicated by a superscript, and separated by commas without spaces in square brackets immediately prior to MC. For example, [D-Leu¹,D-Asp³,Dha⁷]MC-LR contains D-Leu, D-Asp and Dha at positions 1, 3 and 7, and L-Leu and L-Arg at positions 2 and 4, respectively, with Adda and D-Glu assumed by default at positions 5 and 6 (see Figure 1). If an amino acid residue at positions 2 or 4 is not one of the 20 standard amino acids, the three-letter (or more, where necessary) abbreviation is used; e.g., the congener containing Leu at position 2 and homoarginine in position 4 is named MC-LHar. Ring opened MCs are designated with the prefix [seco-a/b], where a and b are the residue numbers between which the amide bond has been hydrolysed. For example, [seco-4/5]MC-LR indicates MC-LR hydrolysed between the Arg⁴- and Adda⁵-residues. A microcystin name-generator is included in version 15 of the microcystin mass calculator tab of a publicly available toxin mass list [52]. To facilitate ongoing efforts to maintain comprehensive lists and databases of toxins and/or cyanobacterial metabolites, it is recommended that researchers use these naming conventions when reporting the identification of new microcystins.
Figure 1. General structure of microcystins and an overview of their observed structural diversity. R$^1$ = H or CH$_3$; R$^2$ = H or CH$_3$; R$^3$ = H, CH$_3$ or C$_3$H$_7$OH; R$^4$ = H, CH$_3$ or COCH$_3$; X and Z = variable L-amino acids.

3. Biosynthesis of microcystins

The nonribosomal biosynthesis of MCs has been described in detail in *Microcystis* [21-23,53], *Planktothrix* [54] and *Anabaena* [55]. This MC synthetase system is encoded by two transcribed operons in *M. aeruginosa*, *mcyABC* and *mcyDEFGHIJ*, from a central bidirectional promoter located between *mcyA* and *mcyD* [23,53]. Insertional gene knockout experiments have demonstrated that all microcystin congeners produced by a strain are synthesized by a single enzyme complex encoded by a 55 kb gene cluster [21,54,56,57]. Therefore, strains that do not produce MCs can result from inactivation either by point mutation, or by partial or complete deletion of *mcy* genes [36,56,58,59]. Recently, Shishido et al. [60] reported that the *mcy* gene cluster did not encode all enzymes for the synthesis of rare variants of microcystin that contain a selection of homo-amino acids by the benthic cyanobacterium *Phormidium* sp. LP904c, but may instead activate homo-amino acids produced during the synthesis of anabaenopeptins.

Several studies have assessed the impact of environmental factors, including light, iron limitation, temperature and pH on cellular MC content [25,31,32,61-67]. Some of these factors are involved in enhancing or suppressing the expression of the *mcy* gene, leading to an increase or decrease in MC production. For example, Kaebernick et al. [61] reported that light quality has a significant effect on the transcription levels of the genes *mcyB* and *mcyD*. Recently, Yang and Kong [68] reported that the exposure of *M. aeruginosa* cells to UV-B not only remarkably inhibited the growth of *Microcystis*, but also led to a reduction in MC concentration by decreasing transcription of the gene *mcyD*. Sevilla et al. [63] demonstrated that iron starvation causes an increase in *mcyD*...
transcription and MC levels. Moreover, it was also shown that stress conditions caused by nutrient deprivation increased mcyD transcription and MC production [67]. Horst et al. [69] combined laboratory and field experiments and observed a correlation between nitrogen limitation and lower MC cell quotas in a field survey of Lake Erie (situated at the international boundary between Canada and the USA). Since MC is a nitrogen-rich metabolite, Van de Waal et al. [30] have observed an increase in MC cell quotas after excess nitrogen supply, the nitrogen-rich variant MC-RR being the most significantly affected. Similarly, Piddick et al. [32] recently reported that when Microcystis strain CAWBG11 was grown in batch culture until nitrogen-depletion, the relative abundance of the microcystin congeners shifted and the relative abundance of arginine-containing microcystins decreased, whilst the congeners which did not contain arginine increased. These changes coincided with a large decrease in nitrate concentration to 0.35% of initial levels.

Therefore, the biosynthesis of MCs is strain dependent and their diversity is due to variability in the coding of the microcystin synthetase genes among cyanobacterial strains [70] as well as some environmental factors such as nitrogen concentrations [29,30,32,71,72]. The large variety of amino acids that can be incorporated and further modified by tailoring enzymes, having activities such as the epimerization, cyclisation, N-methylation, formylation, and reduction of amino acids, accounts for the production of highly complex peptides [73,74]. In the field, it is important to consider that multiple strains of cyanobacteria co-exist, often leading to more complex profiles of MCs. Typically, cyanobacterial strains in the field and in controlled conditions produce up to twenty microcystin congeners but usually only one or two congeners dominate in any single strain [12,40,75-79]. However, the number of congeners produced in a single strain can exceed 47 when all minor MC analogues are considered [33] but in many other cases MC profiles reportedly contain only a small number of congeners. This highlights the fact that there are still new MCs being constantly identified, with varying degrees of certainty, in the very recent literature [80,81]. However, gaps remain in the understanding of how environmental factors can modulate the relative abundance of MC congeners in a bloom, requiring further research on the topic.

4. Structural elucidation of microcystins

Definitive structural elucidation of new MCs can only be achieved by purifying individual microcystins using preparative isolation techniques followed by comprehensive NMR, amino acid analysis and tandem MS (MS/MS) experiments.

Proper purification, structural elucidation and quantitation are also of critical importance to toxicity studies, where the knowledge of structure, purity and amount of each MC tested for toxicity are required to ensure the validity and comparability of toxicology results. In this section we give a brief overview of the techniques available for structural elucidation of MCs and the structural information that each technique can provide. This will be useful in choosing appropriate methods for structural elucidation as well as in the critical evaluation of published MC identifications and toxicities.

4.1. Mass spectrometry for structural elucidation
In addition to offering sensitive and selective targeted quantitative analysis of known MCs, MS is an invaluable tool for tentative structural elucidation of new microcystins, providing information about their molecular formula (full scan) and structure (MS/MS). The types of MS techniques that have been used for the identification of MCs have evolved as mass spectrometers suitable for intact analysis of peptides have been developed. Early work relied heavily on direct analysis of MCs purified off-line, using fast atom bombardment MS in a variety of low and high resolution as well as tandem MS configurations [15,75,82-85]. Similarly, the usually high-resolution MS information from matrix assisted laser desorption ionization-time of flight MS with and without post-source decay has been invaluable in identifying new purified MCs over the years [86-90]. The widespread availability of electrospray ionization (ESI) has made it commonplace to couple analytical liquid chromatography (LC) separations directly to high resolution and/or tandem mass spectrometers. This has offered the possibility for structural characterization and tentative identification of many MCs in complex mixtures without the need for isolation of milligram amounts of toxin needed for full structural elucidation. Liquid chromatography with sequential multi-stage MS dissociation (LC-MSn), while usually offering only unit mass resolution, provides a great deal of information for the determination of structure and connectivity and has been widely used for MC identification [33,74,76,89,91,92]. Alternatively, liquid chromatography with high resolution MS/MS (LC-HRMS/MS) provides accurate mass data on precursor and product ions suitable for determining their molecular formulae [93-96,80,81].

LC-MS alone is usually not considered definitive proof of chemical structure, but a great deal of structural information can never the less be obtained. LC separation of microcystins is typically carried out in reverse phase using C18 or similar stationary phases and acidic mobile phases, resulting in protonation of carboxyl groups, with acetonitrile as the organic modifier. The hydrophobicity and therefore elution order of MCs in RPLC is governed primarily by the number of arginine residues in their structure, and the relative elution order of new congeners can provide useful information about structure. Microcystins with two arginine or homoarginines (e.g. MC-RR) are the most polar and elute first, followed by MCs with a single Arg at position four (e.g. MC-LR), then those with a single Arg at position 2 (e.g. MC-RY) and finally those with no Arg residues (e.g. MC-LA). Though sometimes difficult to separate chromatographically [97], the relative elution order of desmethylated MCs can also be useful for identification of minor analogues with [DMAdda5]-variants eluting much earlier than [D-Asp3]-variants, which are followed closely by [Dha7]-variants [81].

The nominal molecular weights for known parent MCs range from 882 Da to 1101 Da, although the theoretical range using the known amino acids identified to date in natural MCs is significantly wider than this. These masses are too large for definitive molecular formula elucidation by HRMS when all possible elements are considered (e.g. F, P, Cl, etc.), but when the possible elemental composition and ring or double bond equivalents (RDBE) are limited to a reasonable range similar to known MCs, in most cases a candidate formula can usually be determined. Known parent (unconjugated) MCs have between 44 and 57 and 84 H atoms and 13 N atoms, 12–17 O atoms, and up to one S atom in methionine-containing MCs. It can also be useful to limit possible molecular formulas by the number of RDBEs in the structure, which ranges from 17 to 27. Within these limits there are currently no cases among known microcystin where [M+H]+ or [M+2H]2+ ions are within 5
ppm mass accuracy from one another. However, there are >120 cases reported of isomeric microcystins reported, significantly limiting the utility of accurate mass measurements alone for MC identification. It should be noted that conjugates between MCs and thiol-containing biomolecules (e.g. cysteine, glutathione) have been reported [80,98,99] and that mass ranges and molecular formula tolerances need to be widened accordingly when characterizing these compounds.

The ionization and MS/MS fragmentation of MCs are also generally similar between MCs with the same number of Arg/Har residues. In positive ESI mode, MCs with two Arg/Har residues, or conjugated to Cys or GSH, appear primarily as doubly charged ions after ESI while those with one or no Arg/Har are predominantly singly charged. MCs with no Arg/Har ionize comparatively poorly and therefore tend to form adducts with other cations present in the sample (e.g. [M+Na]+ or [M+NH4]+). In negative ESI all MCs exist predominantly as singly charged [M-H]- ions. The MS/MS dissociation of microcystins has been studied extensively in both positive [100,101], and negative [102,103] ionization modes. In most cases, MCs fragment to produce sensitive class-characteristic fragments at m/z 135.0804 originating from the Adda–residue in positive mode, and m/z 128.0353 originating from the Glu6-residue in negative mode. Since MCs are cyclic peptides, a large body of previous work on the collision induced dissociation (CID) of non-microcystin peptides also exists [104] that can be useful in interpretation of microcystin MS/MS spectra. In addition to the class-characteristic fragments, various “sequence ions” are formed during CID following ring opening that represent cleavages of individual amino acids. These can be used to effectively reconstruct the MC structure one amino acid at a time. For example, in negative mode the B-ion series is useful for determining the mass of the amino acid at position 4 since ions represent successive cleavages from the B1 ion [C5H4O-X4-Adda-(Glu-H2O)-Mdha-Ala]- [103]. Similarly, in positive mode, fragments associated with each individual amino acid can be readily identified from HRMS/MS data. This allows the amino acid at which small structural variations, such as demethylation have occurred, to be determined but not always the definitive structure of amino acid side chains, particularly for low resolution MS data [100].

There are two different types of CID that provide different and often complementary information for structural elucidation; quadrupole and ion trap CID. Quadrupole CID, sometimes termed higher energy collisional dissociation (HCD), is a single-stage process involving a smaller number of more energetic collisions. The main advantage of this approach is that the full m/z range of product ions can be detected simultaneously, allowing for detection of both class-specific and structurally informative product ions in a single experiment. However, it can be more difficult to trace the origin of lower mass product ions. Ion trap CID involves longer time scales and a larger number of lower energy collisions, as well as offering the potential for multiple stages of MS/MS (MSn). Here, structural elucidation is done by constructing a fragmentation tree where the origin of lower mass product ions can be established in higher order experiments (e.g. MS2 or MS3) in relation to higher-mass product ions observed in MS2. The principal limitation of ion trap dissociation is the 1/3 rule, stating that product ions of less than 1/3 the m/z of a selected precursor ion cannot be isolated for detection. This means that class-specific low mass product ions cannot be detected in MS2 regardless of the collision energy used.
Use of established chemical degradation or modification reactions to probe the structure of unknown chemicals is a classical approach that pre-dates the availability of modern analytical instrumentation such as NMR or MS. This approach was an important part of the first structural elucidation studies on MCs, where partial hydrolysis reactions followed by Edman degradation reactions and mass spectrometric methods of characterizing the linearized products were utilised [15,16]. However, when combined with modern LC-MS, similar approaches are even more powerful due to the ease with which multiple products of reactions with multiple reactants can be selectively detected. The reactivity of the double bond in Dha and Mdha of MCs towards thiols has emerged as being particularly useful for characterizing MCs congeners. In particular, mercaptoethanol derivatization has been exploited in such a manner both to identify new MCs from complex samples [76,105] or to quickly differentiate Mdha- from Dhb-containing MCs [106], which is not possible using LC-MS alone unless authentic standards are available. Another useful reaction is methyl esterification that, together with LC-MS, can be used to count the number of reactive carboxylic acid (the CO$_2$H group of Glu is reactive, but not that of Masp$^3$/Asp$^3$) and phenolic groups in MCs [95]. Selective oxidation with mild oxidants can also be combined with LC-MS to identify MCs containing sulfide and sulfoxides, such as methionine (e.g. MC-MR) [80,92]. This approach has also proved useful for identifying sulfide (e.g. Cys, GSH) conjugates of MCs, and their autoxidation products, in natural samples [80]. These conjugates are also readily semi-synthesised to produce authentic standards for confirmation of their presence in samples and cultures by LC-MS (e.g. [80,98]) (see section 5.7). In addition, the identity of S-conjugates at position-7 of MCs can be confirmed through a range of deconjugation reactions [99,107,108].

It should be noted that stereochemical information is not generally accessible from LC-MS/MS analyses. It is therefore generally assumed in the literature that the stereochemistry of the amino acids in MCs whose structures are established using LC-MS methods are the same as those in well-established structures such as MC-LR unless there is evidence to the contrary, and the same assumption is made throughout this review.

4.2. Preparative isolation of microsystins

While a great deal of structural information is available from LC-MS without the need for additional purification of MCs, stereochemistry is not accessible by LC-MS. Definitive structural elucidation by NMR currently requires isolation of 50–100 μg or more of a pure MC. This is most often done by scaling up production of laboratory cultured cyanobacteria, or directly from naturally occurring bloom samples. Dried cyanobacterial biomass is usually first extracted under acidic conditions. Aqueous extracts can then be applied to a reverse phase (e.g. C$_{18}$) open column or large solid phase extraction cartridge, depending on the complexity of the toxin profile. Elution is then carried out in a stepwise fashion, usually with aqueous methanol. In some cases, size exclusion chromatography has been used [82] but is often not required. For more complex toxin mixtures, a final semi-preparative isolation step is required, again using a reverse phase column. Additional selectivity at each step can be achieved by varying the pH of the eluent with various buffers. These
then need to be removed with an additional SPE step. Long-term storage of MCs in acidic methanol should be avoided because of the potential for formation of methyl esters [80,109] (see section 5.7).

### 4.3. Amino acid analysis

One of the classical approaches to structural characterization of MCs after preparative isolation is to hydrolytically digest them to their constituent amino acids, which are then identified chromatographically based on retention time comparison to amino acid standards. Typically, hydrolysis is done in 6 M HCl at 105 °C for 24 hr. It should be noted that under these conditions, some amino acids (e.g. tryptophan, glutamine, serine, threonine, tyrosine, tryptophan and cysteine) are destroyed and their presence must be inferred. Amino acids can then be derivatized with a UV absorbent tag and analyzed by reverse phase LC-UV [82]. Alternatively, they can be derivatized with trifluoroacetic anhydride to make them volatile and then analyzed by gas chromatography [110]. When chiral stationary phases or derivatizing agents are used, it is possible to determine the absolute configuration of the amino acids, which is not possible by MS or NMR alone.

### 4.4. Structural elucidation by nuclear magnetic resonance spectroscopy

Although it is most often used in combination with other approaches mentioned earlier, NMR is usually required for definitive structural identification of MCs including relative stereochemistry. Often CD<sub>3</sub>OD and sometimes D<sub>2</sub>O were used in the past as solvents for NMR spectroscopy of MCs to reduce <sup>1</sup>H signals arising from the solvent signals in 1- and 2-D spectra. The disadvantage is that exchangeable amide protons, and thus their correlations, cannot be detected. It is therefore advantageous to use solvents that do not exchange the amide protons, such as (CD<sub>3</sub>)<sub>2</sub>SO or CD<sub>3</sub>OH, and with modern NMR spectrometers and solvent suppression techniques, the use of partially protonated solvents such as CD<sub>3</sub>OH is not usually problematic. However, NMR spectroscopy is a relatively insensitive technique, and significant amounts of purified compound are generally required. With modern high-field NMR spectrometers, full structural elucidations can be performed on less than 50 μg of an MC in CD<sub>3</sub>OH in 5 mm NMR tubes.

A comprehensive approach to assigning all protons and carbons in MCs by NMR alone was presented nearly 3 decades ago that relied on a combination of three 2D NMR techniques: double quantum filtered correlation spectroscopy (DQF-COSY), <sup>1</sup>H-detected multiple quantum coherence (HMQC), and heteronuclear multiple bond correlation (HMBC) [111]. DQF-COSY was first used to assign H-2 of each constituent amino acid along with several other proton assignments. HMQC was then used to assign C-2 of each constituent amino acid based on the H-2 resonances already assigned. Finally, HMBC NMR spectra, that can detect <sup>2</sup>J- and <sup>3</sup>J-couplings between carbon and hydrogen atoms, were used to assign the carbonyl carbon from each amino acid and ultimately to determine the overall amino acid sequence. Nowadays, TOCSY NMR spectra would typically also be acquired to identify the individual amino acid spin systems, and edited HSQC spectra to identify methyl, methylene and methine <sup>13</sup>C and <sup>1</sup>H resonances. ROESY NMR spectra are generally better suited for
detection of through-space correlations for molecules of this size than are NOESY NMR spectra. In addition to detecting through-space correlations and thus potentially defining the relative stereochemistry of the molecule, ROESY NMR correlations can be useful for establishing the connectivity of the amino acids as a supplement to HMBC NMR correlations. Selective 1D-TOCSY or -DIPSI2 NMR spectra can often be used to obtain 1H signal multiplicities and coupling constants in regions of spectral overlap, which can be useful for chemical shift and stereochemical assignments. The higher resolution afforded in the 13C axis by SHMB and SHSQC spectra is sometimes useful for resolving crowded regions of the HSQC (e.g. methylene region) and HMBC (e.g. amide carbonyl region) spectra, and data acquisition times for 2D spectra can be substantially reduced by use of NUS and NOAH sampling techniques.

NMR spectroscopy can also be used quantitatively to measure the concentrations of algal toxin solutions relative reference solutions [112]. This approach has been applied to produce quantitative certified reference materials (CRMs) for MC-RR, MC-LR, and [Dha−]MC-LR [113] with CRMs for other MCs including MC-LA and [D-Leu1]MC-LY in preparation (P. McCarron, personal communication).

4.5. 3-Dimensional structures of microcystins

Two methods have been used for determining the 3-dimensional structures of microcystins. X-ray crystallography has been used for nearly a century to reveal the 3-dimensional relationships of heavy atoms in crystallized organic molecules. Unfortunately, microcystins have not so far yielded suitable crystals for such studies. However, in recent decades the 3-dimensional structures of microcystins bound to protein phosphatases (PPs) have been determined [114-122] (see Section 6.2). In addition to providing key information about the mode of action and structure-activity relationships of the microcystins, this also provides 3-dimensional structural information about the toxins themselves. The other approach that has been used is NMR spectroscopy, using through-space NOESY and ROESY correlations to estimate the relative intramolecular distances between hydrogen atoms in the microcystins in solution [123-126]. Such measurements can be supplemented with measurements of 1H–1H coupling constants, which provide information about the dihedral angles between spin-coupled hydrogen atoms within the molecule. Such studies have led to several published 3-dimensional NMR-based solution structures, something that is not available from X-ray crystal structures because the structures are based on the solid-state structures that do not necessarily reflect the structure in solution.

Although not in and of itself proof of structure or stereochemistry, observation of high affinity of a purified MC for PPs or MC-specific antibodies can be taken as strong supporting evidence that the absolute stereochemistry of the MC is largely the same as for other MCs. The reason is that recognition of MCs by receptor biomolecules is dependent on the 3-dimensional structure of the ligand (see section 6.2), which is controlled by amino acids present in the MC and their stereochemistries. Biosynthetic reasoning also suggests that most MCs likely share the same stereochemistry, since they are assembled by closely related synthetases produced from genetically similar gene clusters, even in separate MC-producing genera (see section 3).
5. Diversity of characterized microcystin congeners

To date, the identification of at least 269 different MC congeners have been reported in the literature (Table S1) using various combinations of the techniques reviewed in section 4. These congeners include MCs biosynthesized with structural variations at every amino acid position as well as the products of several chemical and biochemical transformations that can occur in the environment or the laboratory, all of which are reviewed in this section.

5.1. Congeners with variable amino acid at position 1

In the first ever report of the structural elucidation of a microcystin, MC-LA, produced by the South African \textit{M. aeruginosa} strain WR70 (UV-0l0), the amino acid residue in position 1 (Figure 1) was found to be \textit{D}-Ala \cite{15,17,18,20}. This amino acid is incorporated by the NRPS module of \textit{mcyA} \cite{23} and was initially considered as relatively conserved in MCs \cite{6,93}. However, subsequent characterization of new congeners has shown that the amino acid residue in position 1 is also variable, although not as diverse as those in positions 2 and 4. \textit{D}-Ala is found in position 1 in 210 of the 269 reported microcystins (Table S1). However, other amino acid residues reported in position 1 include \textit{D}-Ser, \textit{D}-Leu, Gly, \textit{D}-Met, \textit{D}-Phe, \textit{D}-Val, \textit{D}-Hil, and \textit{D}-Met S-oxide. Among these congeners, the first one reported was [{\textit{D}-Ser}$_{1}$,\textit{ADMAdda}]MC-LR isolated from \textit{Nostoc} sp. strain 152 by Sivonen et al. \cite{85}. Subsequently, [{\textit{D}-Leu}$_{1}$]MC-LR was identified as the most abundant microcystin from a laboratory strain RST 9501 of a \textit{Microcystis} sp. isolated from a hepatotoxic \textit{Microcystis} bloom from brackish waters in the Patos Lagoon estuary, southern Brazil \cite{162}, and from a cyanobacterial bloom dominated by \textit{M. aeruginosa} collected in Pakowki Lake, AB, Canada \cite{163}. A decade later, Shishido et al. \cite{74} identified [{\textit{D}-Met}$_{1}$]MC-LR and [{\textit{D}-Met}$_{1}$,\textit{D}-Asp$_{3}$]MC-LR from \textit{M. aeruginosa} strains NPLJ-4 and RST 9501. Furthermore, the authors reported that the \textit{M. aeruginosa} strain NPLJ-4 produced a larger diversity of microcystin variants, but most of them in trace amounts, such as [{\textit{D}-Val}$_{1}$]MC-LR and [{\textit{D}-Phe}$_{1}$]MC-LR. Recently, Puddick et al. \cite{90} described eight new congeners (Table S1) with Gly at position-1 in two hydro-terrestrial cyanobacterial mat samples from the McMurdo Dry Valleys of Eastern Antarctica. Recently, Foss et al. \cite{80} identified a new congener by LC-MS/MS with homoisoleucine (Hil) at position 1, from a \textit{Microcystis} bloom at Poplar Island, MD, USA.

5.2. Congeners with variable amino acid at positions 2 and 4

The two \textit{L}-amino acids in positions 2 and 4 are the most variable of the seven amino acids in microcystins (Table S1). The adenylation domains of \textit{mcyB} and \textit{mcyC} are responsible for the activation of the amino acids that appear in positions 2 and 4, respectively \cite{23,53-55}. Therefore, the variable \textit{L}-amino acids incorporated into microcystins are dependent upon the substrate specificity of the adenylation domain responsible for loading the amino acid onto the microcystin synthetase module \cite{32,172}. To date, 20 different amino acids have been identified at position 2 and 40 at position 4 (Table S1). Leu is present in position 2 in about one third of reported microcystins (94 of the 269) and Arg is present in position 4 in more than half of the reported microcystins (168 of the 269). For example,
Puddick et al. [12], when conducting an analysis of the relevant literature on 49 isolated strains of cyanobacteria, reported that most of the strains assessed were only able to incorporate one or two amino acids into position 4, usually including Arg. Consequently, the congener MC-LR is the most commonly observed MC in most cyanobacterial blooms worldwide [6-9,173]. However, in many other cases, different congeners with differing residues at positions 2 and 4 can dominate a bloom and methods targeting MC-LR exclusively are known to significantly underestimate total MC content in many cases.

### 5.3. Congeners with variable amino acid at position 3

Position 3 is highly conserved in microcystins, as only D-erythro-β-methyl isoaspartic acid (D-Masp) and D-erythro-isoaspartic acid (D-Asp) have been reported in MCs. D-Masp is most commonly reported (155 of 269 MCs), with D-Asp reported in all remaining microcystins (Table S1). The amino acid at position-3 is incorporated by the activity of the NRPS module of mcyB [23]. Furthermore, Pearson et al., [174] and Sielaff et al. [175] reported that one of the moieties, D-erythro-β-methyl-aspartic acid, required the activity of two additional enzymes prior to incorporation in the peptide chain: the 2-hydroxy acid dehydrogenase mcyI and the aspartate racemase mcyF. Both dicarboxylic acids, D-Masp/Asp at position 3 and D-Glu in position 6, are inserted into the cyclic peptide structure in the iso-configuration [15,16].

### 5.4. Congeners with variable amino acid at position 5

The most characteristic twenty-carbon β-amino acid in position 5 of microcystins, and which plays an important role in their biological activity, is Adda (reported in 211 of 269 MCs). Four PKS modules, mcyG, J, D, and E are responsible for the synthesis of the Adda moiety of microcystins [24]. The natural structural diversity of the Adda moiety is largely limited to the group on the C-9 oxygen, that can be present as 9-O-desmethylAdda (DMAdda), the 9-hydroxy derivative of Adda, or as 9-acetoxy-9-O-desmethylAdda (ADMAdda). The O-methylation on the C-9 oxygen of microcystins is carried out by the mcyJ O-methyltransferase [54]. Recently, Fewer et al. [176] reported that the production of O-acylated variants of microcystin in Nostoc is the result of specific enzymatic O-acetylation of the microcystin catalyzed by the mcyL O-acetyltransferase. In fact, several earlier studies reported that O-acylated microcystins are produced almost exclusively by strains of the genus Nostoc isolated from a range of aquatic and terrestrial habitats in free-living and lichen-associated growth states [78,85,133,149]. However, O-acylated microcystins have also been reported in a single strain of Planktothrix agardhii PHI123 [171].

The first DMAdda microcystin congener, [DMAdda₅]MC-LR, was identified from a cyanobacterial bloom containing M. aeruginosa (dominant), M. viridis and M. wesenbergii collected from Homer Lake, IL, USA [75,79]. Furthermore, 15 different DMAdda MC congeners have been described (Table S1), that could originate from either incomplete methylation by the mcyJ gene during biosynthesis or from hydrolysis of the 9-methoxy or 9-acetoxy group in the [Adda₅] and [ADMAdda₅] microcystin variants, respectively. However, since most reports of DMAdda₅-microcystins have occurred in strains or blooms where the corresponding ADMAdda-congeners were not present, and ethers are relatively resistant to hydrolysis, the origin of DMAdda₅-microcystins is likely to be due...
primarily to incomplete methylation during biosynthesis. LC-MS studies recently indicated the existence of small amounts of incompletely characterized MC congeners containing oxidized Adda (hydroxyAdda and ketoAdda) and Adda demethylated (dmAdda) at two positions other than at the 9-O-methyl group [80]. Furthermore, [epoxyAdda]MC-LR was tentatively identified in a field sample [81]. These oxidized congeners are likely to arise through metabolism or autoxidation, while the two dmAdda variants presumably arise through biosynthesis.

5.5. Congeners with variable amino acid at position 6

D-Glu is incorporated in position 6 by the activity of the NRPS module adjacent to the PKS module of mcyE [23] and is highly conserved in this position. The esterification of the α-carboxyl group of the D-Glu residue is rare, and only 11 such MC congeners have been characterized in the literature (Table S1). Namikoshi et al. [75,79] characterized the only non-methyl (-C\(_3\)H\(_6\)OH) monoester of the α-carboxyl on the D-Glu\(^{6}\) unit of MC-LR ([D-Glu(OC\(_2\)H\(_3\)(CH\(_3\)))OH] MC-LR) from a toxic bloom containing M. aeruginosa (dominant), M. viridis, and M. wesenbergii collected from Homer Lake (Illinois) in 1988. However, the mono-methyl ester derivatives at the D-Glu\(^{6}\) unit [D-Glu(OMe)] were more frequent and 11 different congeners have been characterized (Table S1). These methyl esters appear to be artefacts produced by exposure to methanol (section 5.7).

5.6. Congeners with variable amino acid at position 7

The amino acid residue in position 7 is also variable; however, it’s not as diverse as those in positions 2 and 4. N-methyldehydroalanine (Mdha) is found in position 7 in 185 of the 269 reported MCs (Table S1). However, other amino acids have been reported in this position, including dehydroalanine (Dha) [33,82,105,106,134,156], dehydrobutyrine (Dhb) [74,76,78,90,138], N-methyldehydrobutyrine (Mdhb) [78], Ser [33,83,100,134,148,156], N-methylserine (Mser) [79,81,85,92,105,106], Thr [165], and conjugates formed by addition of Cys, γ-GluCys and glutathione (GSH) with the reactive double bond of Mdha\(^{7}\) [79,80]. Chemical and metabolic transformations of MCs are discussed in section 5.7.

The N-terminal adenylation domain of mcyA catalyzes the N-methylation of microcystins [21,23,54,55]. The N-methyltransferase (NMT) domain recognizes and activates Ser as an AMP-adenylate, causing the Ser residue to undergo N-methylation and dehydration to form Mdha [21,23,34,54,55]. Thus, the NMT domain catalyzes the transfer of the methyl group from S-adenosylmethionine (SAM) to the α-amino group of the thioesterified amino acid [177]. However, Fewer et al. [34] reported that the order and timing of the N-methylation, dehydration and condensation reactions have not yet been determined. For example, deletion of the entire NMT domain of mcyA in Anabaena [34] and Planktothrix [178] or point mutations in this gene in Microcystis [37], were associated with the production of Dha\(^{7}\)-microcystins by these genera. Therefore, simultaneous production of Mdha- and Dha-containing microcystins in Anabaena could be the result of the inactivation of the NMT domain during the assembly of Dha-containing microcystins [34]. However, Marahiel et al. [179] suggested that the concomitant biosynthesis of Dha- and Mdha-containing microcystin variants could also be the result of the limitation of SAM, the source of the N-methyl group. Such mixtures of microcystins containing both Mdha and Dha are common in
Planktothrix [110] and Microcystis [135], and partial inactivation of the NMT domain may be a widespread feature of microcystin biosynthesis.

MCs-containing Dhb instead of Dha in Planktothrix are thought to be the result of a homologous recombination event in the first module of mcyA that replaces the adenylation domain, leading to a change in substrate specificity (from Ser to Thr) and presumably the concomitant deletion of the NMT domain [178]. For the biosynthesis of nodularin, a microcystin-like pentapeptide containing Mdhb instead of Mdha, the nodularin synthetase gene cluster is proposed to have arisen by deletion of two NRPS modules and a change in substrate specificity from Ser to Thr in the C-terminal adenylation domain of NdaA [180]. Kurmayer et al. [178] reported a significant correlation between mcyAAd1 genotypes with NMT and mcyAAd1 genotypes without NMT and the occurrence of either Mdha or Dhb in position 7 of the MC structures in Planktothrix strains. Therefore, strains of cyanobacteria producing microcystin variants containing either Dha or Dhb, exclusively, are known from all major MC producing genera (Microcystis, Nostoc, Anabaena, and Planktothrix) [6,33,90,106,134,140]. Furthermore, Sano and Kaya [138] reported that the configuration of the Dhb7 unit of Dhb-microcystins isolated from the CCAP strains of O. agardhii was E, while the Dhb of nodularins from Nodularia spumigena has been determined as Z [181], suggesting that the Dhb unit in Dhb-microcystins is biosynthesized via a dehydration of Thr or allo-threonine.

The microcystin congener [D-Asp3, Thr7]MC-RR was identified from an environmental bloom of P. rubescens collected from Austrian alpine lakes [165]. Additionally, 9 microcystin analogues containing Ser and 13 containing Mser instead of Mdha in position 7 have been reported (Table S1).

According to Tillett et al. [23], the module mcyA is responsible for the incorporation of Ser into the growing molecule. Ser is further N-methylated by the corresponding NMT domain and transformed by dehydration into Mdha prior to or following the condensation reaction. Analogous to Ser, Thr is transformed by dehydration into Dhb [182]. Therefore, the production of the unique congener with Thr residue and the congener with Ser or Mser residues in position 7 could be the result of the partial inactivation of the NMT domain during the dehydration and N-methylation reactions of Thr- and Ser-containing microcystins. In fact, Fewer et al. [34] reported that there is a 10–500-fold increase in the amount of Ser incorporated into the MCs in strains of Anabaena lacking a functional NMT domain, suggesting that there may be a functional coupling of the dehydration and N-methylation reactions.

5.7. Chemical and biochemical transformations of microcystins

Several chemical and biochemical reactions have been reported to occur in the cell, the environment or during sample extraction and storage in the laboratory that lead to the formation of new MC analogues beyond those that are biosynthetic products of cyanobacteria (Figure 2). Even though these are not of biosynthetic origin, their occurrence can still be important to measure for studies of MC occurrence and toxicity. It is recommended that authors describing the identification of new microcystin variants try to clearly differentiate between those that are metabolites of cyanobacteria, those that form through conjugation or metabolism after biosynthesis, and those that form artefactually during sample handling or storage.
Eight MCs containing sulfur oxidation in their Met moiety have been reported (Table S1) that are likely artefacts produced by autoxidation. This is consistent with the findings of Miles et al. [92] who reported that a fresh culture of Dolichospermum flos-aquae contained only traces of sulfoxides [D-Asp]3MC-M(O)R and MC-M(O)R, but that these increased markedly during storage or sample extraction and preparation. This suggests that microcystins containing methionine sulfoxide are primarily post-extraction oxidation artefacts, rather than being produced by biosynthesis in cyanobacteria. These can then be further oxidized to the corresponding sulfones (e.g. MC-M(O2)R) though an additional step of autoxidation [92]. In addition, Puddick et al. [89] reported microcystin congeners containing the oxidized tryptophan moieties kynurenine (Kyn), oxindolylalanine (Oia), and N-formylkynurenine (Nfk) that are also considered as autooxidation artefacts.

Isomerization of the C-6 double bond in the Adda group is rarely reported and is thought to be due to ultraviolet (UV) irradiation rather than arising through biosynthesis [183-185]. In early studies, Harada et al. [111] and Namikoshi et al. [75] isolated geometrical isomers at C-6 of [(6Z)-Adda]MC-LR and -RR, as minor components with along with the corresponding normal (6E)-Adda MCs (MC-LR and -RR) from Microcystis spp. The DMAdda derivative at the Adda unit could be also an artefact produced during extraction of cyanobacterial samples dominated by ADMAdda derivatives. In fact, Beattie et al. [146] observed that [ADMAdda] microcystins are unstable and hydrolyze at room temperature in aqueous solution and/or in methanol, and that [D-Asp,ADMAdda,Dhb]MC-RR isolated from a Nostoc strain DUN901 was converted to [D-Asp,DMAdda,Dhb]MC-RR within one week. Similarly, Ballot et al. [127] reported the half-life for hydrolysis of several [ADMAdda]MCs (to form the corresponding [DMAdda]MCs) to be ca 30 h at pH 9.7 and 30 °C.
In addition, MCs containing methyl esters in their Glu\(^{6}\) moiety are almost certainly artefacts produced by esterification in methanolic solutions, possibly due to traces of residual acid from extraction or isolation steps. For example, [D-Glu(OMe)\(^{6}\)]MC-LR and [D-Glu(OMe)\(^{6}\)]MC-FR characterized recently by Bouhaddada et al. [40] were identified in the presence of their much more abundant non-esterified congeners MC-LR and MC-FR, respectively. In addition, the same results were reported for the eight other described congeners with methyl esters on the α-carboxyl group of the D-Glu residue in position 6 [75,79,84,94,148,164]. Harada et al. [109] demonstrated formation of Glu\(^{6}\) methyl esters of MCs in acidic methanolic solutions, and Foss et al. [80] observed substantial esterification of [Leu\(^{1}\)]MC-LR that was purified with an acidic eluent then evaporated to dryness before storage in MeOH at −20 °C. In fact, del Campo and Ouahid [38] identified the presence of [D-Asp\(^{3}\),D-Glu(OMe)\(^{6}\)]MC-Laba from a cultured Microcystis strain UTEX-2666, but not its non-esterified residue [D-Asp\(^{3}\)]MC-Laba, suggesting that near-complete esterification is possible under some conditions. Therefore, the technique described by Fastner et al. [186] consisting of analyzing of a single colony of Microcystis directly, without methanol extraction may circumvent this problem. As a precaution, use of both acids and methanol during sample preparation or storage should be avoided where possible. Esters of MCs containing L-Glu at position-2 and -4 have also been reported [134] and given the facile esterification of D-Glu\(^{6}\) it seems probable that these two are esterification artefacts as well.

MCs with either Dha or Mdha at position 7 possess an α,β-unsaturated carbonyl group that is reactive toward thiols (e.g., cysteine), and to a lesser extent amines and hydroxyl groups, through the Michael addition reaction. Consequently, the unusual microcystin congener [Mlan\(^{7}\)]MC-LR that was characterized from a bloom of Microcystis spp. collected from Homer Lake (IL, USA) can be considered the result of a chemical reaction between the major constituent (MC-LR) biosynthesized in the cyanobacterium and cysteine [79]. On the other hand, the biosynthesis of lanthionine has been proposed to involve addition of the thiol group of a Cys residue to the double bond of Dha, generated by dehydration from Ser, forming a sulfide bridge [79]. It has also been proposed that glutathione-S-transferase catalyzes the addition of glutathione (GSH) to microcystins, but also that the reaction proceeds at a similar rate with or without the enzyme under physiological conditions [99,187], suggesting that the reaction may be primarily chemical rather than purely biochemical. Recently, 3 new microcystins containing cysteine (Cys), glutathione (GSH), and γ-glutamylcysteine (γ-GluCys) conjugated to the Mdha\(^{7}\) olefinic bond of [D-Leu\(^{1}\)]MC-LR were identified in extracts of freeze-dried Microcystis bloom material from Poplar Island (MD, USA) [80]. Small amounts of the corresponding sulfoxides of the conjugates (Cys(O), cysteine sulfoxide; GS(O), glutathione sulfoxide; and γ-GluCys(O), γ-glutamylcysteine sulfoxide), probably produced by autoxidation, were also detected. Conjugated MCs present in the bloom material were characterized for the first time, but they could nonetheless be metabolites produced during detoxification of microcystin congeners in aquatic organisms. The principal pathway of the metabolism and the excretion of MCs in mammals and aquatic organisms has been found to be as MC–GSH and MC–Cys conjugates [8,188-190]. However, cyanobacteria are known to produce both cysteine and glutathione [191], and low levels of GSH conjugates of microcystins have been detected in several cyanobacterial cultures using untargeted LC-HRMS methods (D. G. Beach and C. O. Miles, unpublished observations). GSH conjugates can be further metabolized to γ-GluCys and Cys conjugates [192], so it is possible that the γ-GluCys and Cys conjugates observed from cyanobacteria are produced via this route or by direct conjugation of.
MCs to these smaller peptides. Nucleophilic addition of water (to produce \( \text{D,L-Ser}^7 \) and \(-\text{Mser}^7\) congeners) and methanol (to produce the corresponding \( \text{O-methoxyserine} \) congeners), and intramolecular addition of the 2-arginyl group, have all also been observed at the reactive double bond of \( \text{Dha}^7 \) and \( \text{Mdha}^7 \) of microcystins under basic conditions during studies of their derivatization with thiols [99], as shown in Fig. 2.

Eight linear MCs (Table S1) have been reported to date that could well be the products of bacterial degradation of MCs, since many studies have shown that the first step of biodegradation of microcystins by multiple different species of bacteria is through the cleavage and opening of the macrocycle (reviewed by Li et al. [193]).

MC congeners containing oxidized Adda (an epoxide, ketone and hydroxylated derivative) have also been reported recently in small amounts in a bloom, and partially characterized by LC-HRMS/MS and chemical derivatizations [80]. At this stage it is unclear whether these are the result of enzyme-catalyzed oxidation or are products of autoxidation.

Researchers should consider the detection in a sample of a potentially artefactual MC congener, such as an oxidized, esterified or conjugated variant, as possible evidence for the presence in the original sample of its unmodified precursor microcystin.

6. Toxicity of microcystin congeners

6.1. Toxicity in vivo

It should be borne in mind when comparing LD\(_{50}\) values (the dose of toxin that kills 50% of exposed animals) that toxicological measurements are inherently subject to uncertainty caused by biological variability, and the possibility that in some cases isolated MCs may not have been adequately characterized with respect to purity and amount of toxin prior to toxicity assessment. For this reason, only large differences in toxicity should be considered significant when comparing toxicities from different experiments and different MC congeners.

The toxic effects of MCs on animals were first studied with unpurified cyanobacterial extracts. Several studies reported enlargement and congestion of the liver, with necrosis of the hepatic cells in laboratory and domestic animals after oral administration or parenteral injection of toxic algal extracts in which \( \text{M. aeruginosa} \) was the predominant species [44,194,195]. A decade later, Elleman et al. [161] confirmed these observations and reported that the lethal dose (LD\(_{100}\)) in mice by intraperitoneal (i.p.) injection of an \( \text{M. aeruginosa} \) (then \( \text{Anacystis cyan} \)) bloom extract was 15–30 mg of lyophilized algal bloom per kilogram body weight (b.w.). The LD\(_{50}\) of the purified toxin from this extract was estimated to be 56 \( \mu \)g kg\(^{-1}\) in mice, and the approximate LD\(_{100}\) was 70 \( \mu \)g kg\(^{-1}\). Acute toxicity by parenteral administration of the purified toxin to mice produced extensive liver lobular hemorrhage and death within 1–3 h. However, repeated administration of sublethal doses daily over some weeks produced progressive hepatocyte degeneration and necrosis and the development of fine hepatic fibrosis.

Since then and following the identification of the toxin responsible, toxicological research has focused on purified MCs. The most extensive toxicological information is available for MC-LR. The acute toxicity of MC-LR has been tested in several studies on mice, with i.p. LD\(_{50}\) values ranging from 25 to 150 \( \mu \)g kg\(^{-1}\) b.w. [196,197]. An LD\(_{50}\) value of 50 \( \mu \)g kg\(^{-1}\) b.w. is generally accepted (Table 1) (for a
review, see [198]). The inhalation toxicity (intratracheal and intranasal) in mice is similar, with an 
LD50 of 36 to 122 μg kg⁻¹ b.w. [199-202]. However, MC-LR is 100 times less toxic by oral ingestion than 
it is by intra-peritoneal injection, with an LD50 of 5000 μg kg⁻¹ b.w. [196,203]. Although the toxicity of 
all characterized microcystin congeners has not been determined, there are some variations within 
the microcystin family with i.p. LD50 values ranging from 50 to >1200 μg kg⁻¹ b.w. in mice (see Table 
1). Firstly, the cyclic structure of microcystins is very important for the toxicity and opening of the 
macrocycle at any position makes the molecule much less toxic. For example, the two congeners 
[seco-1/2]MC-LR and [seco-4/5]MC-LR, corresponding to the opening of the cyclic structure of the 
congener MC-LR between the amino acid residues at positions 1/2 and 4/5, respectively, showed no 
toxicity to mice at 22500 and 1100 μg kg⁻¹ b.w., respectively [79]. Secondly, two of the most conserved 
residues, at position 5 (Adda, present in 211 of 269 identified microcystin congeners) and position 6 
(d-Glu, present in 257 of 269 identified microcystin congeners) are also very important for the 
toxicological potential. Adda group stereochemistry and its double bond configuration (6E) are 
crucial for the toxicity since [(6Z)-Adda]MC-LR showed an LD50 in mice >1200 μg kg⁻¹ b.w. [111,143]. 
However, the toxicity appears to be only slightly affected by 9-O-demethylation (DMAAdda) or 9-O- 
acetylation (ADMAdda) of the Adda group. [DMAAdda]MC-LR and [ADMAdda]MC-LR were 
found to have LD50 values of 90–100 and 60 μg kg⁻¹ b.w., respectively, by i.p. injection in mice. The 
free carboxylic acid on the D-Glu⁴ residue seems to be essential for toxicity because the esters ([D- 
Glu(OC₂H₃O)⁴]MC-LR [75] and [D-Glu(OMe)⁴]MC-LR) [84,136,182] of the MC-LR showed no toxicity 
in the i.p. mouse bioassay at 1000 μg kg⁻¹ b.w. These results are consistent with the results of studies 
with synthetic fragments containing elements of the microcystin structure, where the presence of 
Adda and a suitably placed carboxylic acid group mimicking the carboxyl group of Glu⁴ was enough 
to elicit PP inhibition, albeit at potencies far below that of MC-LR [204,205]. However, substitution of 
amino acids in positions 1, 3 and 7 only had a small impact on the toxic potential of MCs in vivo. For 
example, MC toxicity seems to be barely affected by substitution of the moderately conserved D-Ala 
in position 1 by other D-amino acids. Matthiensen et al. [162] and Park et al. [163] reported that the 
LD50 of [D-Leu]MC-LR by i.p injection in mice was 100 μg kg⁻¹ b.w. Likewise, MC toxicity was only 
slightly affected by demethylation of the amino acid at position 3 (D-Asp³ instead of D-Masp³) or by 
demethylation or amino acid substitution at position 7. For example, demethylated forms of MCs 
containing D-Asp³ or Dha⁷ residues are two to five times less or more toxic, respectively, compared 
to their methylated forms [82,135,206,207]. Similarly, Sano and Kaya [138] reported that the toxicity 
slightly affected by the addition of a methyl group to the double bond of the Dha⁷ unit to form the 
Dhb⁷, whereas, conjugation of the Mdha⁷ residue with Cys residue (to give [L-Mlan⁷] instead of 
Mdha⁷) decreased the toxicity 20-fold compared to MC-LR by i.p injection in mice [79]. Nevertheless, 
the acute toxicity is significantly affected by substitution at the two most variable amino acid residues 
in positions 2 and 4. Most of the structural microcystin congeners containing a hydrophobic L-amino 
acid in position 2 or 4, such as MC-LA, MC-YR, MC-YA and MC-YM(O), have similar LD50 values to 
MC-LR (Table 1), whereas the LD50 of the hydrophilic MC-RR is about ten times higher (Table 1). 

Most of the differences in in vivo toxicity between microcystin congeners can be attributed to a 
combination of differences in their uptake into the liver via a carrier-mediated transport system 
(organic anion transporting polypeptides (OATPs)) and/or their inhibitory potency toward hepatic 
PPs that are widely recognized as the principal targets of MCs [94,197,208-214]. MC toxicity requires 
active transport by OATP1B1 and OATP1B3, members of the OATP family of transporters, for uptake
into cells and the high expression of these transporters in the liver accounts for their selective hepatotoxicity [94,215]. While some studies suggest that differences in *in vivo* and *in vitro* toxicity are related to PP inhibition rather than transport [207,216], other reports show that *in vivo* toxicity depends mainly on OATP transport kinetics rather than on differences in PP inhibition [206,215]. Therefore, future study is required to determine the relative importance/contribution of OATP versus PP inhibition to overall toxicity of MCs.

### Table 1. LD<sub>50</sub> values (μg kg<sup>-1</sup> b.w., i.p. mouse) of microcystin congeners

<table>
<thead>
<tr>
<th>Entry&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Microcystin</th>
<th>LD&lt;sub&gt;50&lt;/sub&gt; (μg kg&lt;sup&gt;-1&lt;/sup&gt; b.w., i.p. mouse)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>MC-LA</td>
<td>50</td>
<td>[217]</td>
</tr>
<tr>
<td>18</td>
<td>MC-AR</td>
<td>250</td>
<td>[75]</td>
</tr>
<tr>
<td>22</td>
<td>MC-YA</td>
<td>60–70</td>
<td>[218]</td>
</tr>
<tr>
<td>25</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;,Dha]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>160–300</td>
<td>[82]</td>
</tr>
<tr>
<td>38</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>200–500</td>
<td>[135,206]</td>
</tr>
<tr>
<td>39</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;,(E)-Dhb]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>70</td>
<td>[138]</td>
</tr>
<tr>
<td>41</td>
<td>[Dha]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>250</td>
<td>[82]</td>
</tr>
<tr>
<td>42</td>
<td>[2DMAAdda]&lt;sup&gt;5&lt;/sup&gt;MC-LR</td>
<td>90 – 100</td>
<td>[75,145]</td>
</tr>
<tr>
<td>65</td>
<td>MC-LR</td>
<td>50</td>
<td>[198]</td>
</tr>
<tr>
<td>66</td>
<td>[(6Z)-Adda]&lt;sup&gt;3&lt;/sup&gt;MC-LR</td>
<td>&gt;1200</td>
<td>[111,143]</td>
</tr>
<tr>
<td>81</td>
<td>MC-LY</td>
<td>90</td>
<td>[199]</td>
</tr>
<tr>
<td>87</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;,ADMAdda]&lt;sup&gt;5&lt;/sup&gt;MC-LR</td>
<td>160</td>
<td>[77,78]</td>
</tr>
<tr>
<td>90</td>
<td>MC-HilR</td>
<td>100</td>
<td>[79]</td>
</tr>
<tr>
<td>92</td>
<td>[D-Glu(OCH&lt;sub&gt;3&lt;/sub&gt;)&lt;sub&gt;6&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>&gt;1000</td>
<td>[84,136,182]</td>
</tr>
<tr>
<td>103</td>
<td>[Mser]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>150</td>
<td>[79]</td>
</tr>
<tr>
<td>109</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-FR</td>
<td>90±10</td>
<td>[207]</td>
</tr>
<tr>
<td>119</td>
<td>[2DMAAdda]&lt;sup&gt;5&lt;/sup&gt;MC-LR</td>
<td>60</td>
<td>[77,78]</td>
</tr>
<tr>
<td>127</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-RR</td>
<td>350±10</td>
<td>[207]</td>
</tr>
<tr>
<td>128</td>
<td>[Dha]&lt;sup&gt;7&lt;/sup&gt;MC-RR</td>
<td>180</td>
<td>[153]</td>
</tr>
<tr>
<td>129</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;,(E)-Dhb]&lt;sup&gt;7&lt;/sup&gt;MC-RR</td>
<td>250</td>
<td>[138]</td>
</tr>
<tr>
<td>136</td>
<td>MC-M(O)R</td>
<td>700–800</td>
<td>[75]</td>
</tr>
<tr>
<td>137</td>
<td>MC-FR</td>
<td>250</td>
<td>[75]</td>
</tr>
<tr>
<td>155</td>
<td>MC-YM(O)</td>
<td>56</td>
<td>[161]</td>
</tr>
<tr>
<td>157</td>
<td>[2DMAAdda]&lt;sup&gt;5&lt;/sup&gt;MC-LHar</td>
<td>60</td>
<td>[77,78]</td>
</tr>
<tr>
<td>159</td>
<td>[D-Leu]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>100</td>
<td>[162,163]</td>
</tr>
<tr>
<td>162</td>
<td>MC-RR</td>
<td>500–800</td>
<td>[48,75]</td>
</tr>
<tr>
<td>163</td>
<td>[(6Z)-Adda]&lt;sup&gt;3&lt;/sup&gt;MC-RR</td>
<td>&gt;1200</td>
<td>[111,143]</td>
</tr>
<tr>
<td>177</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-HtyR</td>
<td>80–100</td>
<td>[143]</td>
</tr>
<tr>
<td>179</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;,(E)-Dhb]&lt;sup&gt;7&lt;/sup&gt;MC-HtyR</td>
<td>70</td>
<td>[138]</td>
</tr>
<tr>
<td>182</td>
<td>MC-YR</td>
<td>70</td>
<td>[16]</td>
</tr>
<tr>
<td>202</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;,ADMAdda]&lt;sup&gt;5&lt;/sup&gt;Dhb&lt;sup&gt;7&lt;/sup&gt;MC-RR</td>
<td>200</td>
<td>[146]</td>
</tr>
<tr>
<td>203</td>
<td>[D-Glu(OCH&lt;sub&gt;3&lt;/sub&gt;)&lt;sub&gt;6&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-LR</td>
<td>&gt;1000</td>
<td>[75]</td>
</tr>
<tr>
<td>205</td>
<td>[D-Asp&lt;sub&gt;3&lt;/sub&gt;]&lt;sup&gt;7&lt;/sup&gt;MC-WR</td>
<td>95±10</td>
<td>[207]</td>
</tr>
<tr>
<td>214</td>
<td>MC-HtyR</td>
<td>160–300</td>
<td>[143]</td>
</tr>
</tbody>
</table>
6.2. In vitro toxicity: inhibition of serine/threonine PPs

The nature of the interaction of MCs with serine/threonine PPs has been intensively studied [114,208,212,219-222]. The concentrations required to inhibit 50% of the enzyme activity (IC$_{50}$) of serine/threonine phosphatases (PP1–6) by microcystin congeners are shown in Table 2. MC-LR, the most toxic congener, has similar potency in the inhibition of PP1, PP2, and PP4–6, presenting an IC$_{50}$ of around 0.1–1 nM [223-226]. However, it is 1000-fold less potent towards PP3 (IC$_{50}$ ca. 1 μM) and seems to have no effect on the activity of PP7 [226,227].

### Table 2. IC$_{50}$ (nM) values for inhibition of serine/threonine protein phosphatases (PPs) by microcystin congeners

<table>
<thead>
<tr>
<th>Entry</th>
<th>Microcystin</th>
<th>Type of PP</th>
<th>Origin of PP</th>
<th>IC$_{50}$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>MC-LA</td>
<td>PP1</td>
<td>Rabbit muscle</td>
<td>2.3</td>
<td>[228]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP2A</td>
<td>Human hepatocytes</td>
<td>0.56</td>
<td>[229]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP2A</td>
<td>Rabbit muscle</td>
<td>0.05</td>
<td>[228]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.161 ± 0.002</td>
<td>[230]</td>
</tr>
<tr>
<td>14</td>
<td>MC-LV</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>0.06–0.45</td>
<td>[129]</td>
</tr>
<tr>
<td>17</td>
<td>MC-LL</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>0.06–0.45</td>
<td>[129]</td>
</tr>
<tr>
<td>25</td>
<td>[d-Asp$_3$,Dha$_7$]MC-LR</td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.254 ± 0.004</td>
<td>[230]</td>
</tr>
<tr>
<td>33</td>
<td>MC-LM</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>0.06–0.45</td>
<td>[129]</td>
</tr>
<tr>
<td>38</td>
<td>[d-Asp$_3$]MC-LR</td>
<td>PP2A</td>
<td>Rabbit skeletal muscle</td>
<td>0.09</td>
<td>[206]</td>
</tr>
<tr>
<td>39</td>
<td>[d-Asp$_3$,(E)-Dhb$_7$]MC-LR</td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.201 ± 0.003</td>
<td>[230]</td>
</tr>
<tr>
<td>40</td>
<td>[d-Asp$_3$,(Z)-Dhb$_7$]MC-LR</td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.16 ± 0.01</td>
<td>[230]</td>
</tr>
<tr>
<td>41</td>
<td>[Dha$_7$]MC-LR</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>0.54–5</td>
<td>[231,232]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP2A</td>
<td>Bovine kidney</td>
<td>0.11 ± 0.04</td>
<td>[232]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.167 ± 0.003</td>
<td>[230]</td>
</tr>
<tr>
<td>42</td>
<td>[DMAdda$_5$]MC-LR</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>1.5</td>
<td>[231]</td>
</tr>
<tr>
<td>51</td>
<td>MC-LF</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>1.8</td>
<td>[215]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>0.06–0.45</td>
<td>[129]</td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP1</td>
<td>Rabbit muscle</td>
<td>0.1 – 1.9</td>
<td>[206, 208, 228, 231-236]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP1</td>
<td>Chicken gizzard myosin B</td>
<td>6</td>
<td>[237]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP1</td>
<td>Liver of grass carp</td>
<td>0.90</td>
<td>[238]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>1.2</td>
<td>[215]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Rabbit skeletal muscle</td>
<td>0.04–0.5</td>
<td>[206, 208, 228, 234, 239]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Human hepatocytes</td>
<td>0.46</td>
<td>[229]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Human erythrocytes</td>
<td>0.03–2.2</td>
<td>[232, 233, 240, 241]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Bovine heart</td>
<td>0.05–2</td>
<td>[232, 237, 242]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Bovine kidney</td>
<td>0.2 ± 0.1</td>
<td>[232]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Mouse brain</td>
<td>0.28–3.15</td>
<td>[209, 238, 243, 244]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>PP2A</td>
<td>Liver of grass carp</td>
<td>0.28</td>
<td>[238]</td>
<td></td>
</tr>
<tr>
<td>MC-LR</td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.032 ± 0.004</td>
<td>[230]</td>
<td></td>
</tr>
<tr>
<td>[(6Z)-Adda]MC-LR</td>
<td>PP2A</td>
<td>Mouse brain</td>
<td>80</td>
<td>[243]</td>
<td></td>
</tr>
<tr>
<td>[(6Z)-Adda]MC-LR</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>&gt; 100</td>
<td>[231]</td>
<td></td>
</tr>
<tr>
<td>MC-LY</td>
<td>PP2A</td>
<td>Human hepatocytes</td>
<td>0.34</td>
<td>[229]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Glu(OMe)a]MC-LR</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>&gt; 100</td>
<td>[231]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Asp,Dha7]MC-RR</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>0.22 ± 0.01</td>
<td>[230]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Asp]MC-RR</td>
<td>PP2A</td>
<td>Human red blood cells</td>
<td>0.45–11.5</td>
<td>[245]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Asp]MC-RR</td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.30 ± 0.01</td>
<td>[230]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Dha]MC-RR</td>
<td>PP1</td>
<td>Rabbit muscle</td>
<td>8.3 ± 0.8</td>
<td>[232]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Dha]MC-RR</td>
<td>PP2A</td>
<td>Bovine kidney</td>
<td>4 ± 1</td>
<td>[232]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Dha]MC-RR</td>
<td>rPP2Ac</td>
<td>Recombinant human PP2A catalytic subunit</td>
<td>0.29 ± 0.01</td>
<td>[230]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Dha]MC-RR</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>2.6–5.7</td>
<td>[245]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Asp,(E)-Dhb7]MC-RR</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>1.8–56.4</td>
<td>[206, 245]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Asp,(E)-Dhb7]MC-RR</td>
<td>PP2A</td>
<td>Rabbit skeletal muscle</td>
<td>2.4</td>
<td>[206]</td>
<td></td>
</tr>
<tr>
<td>[(6S)-Asp,(E)-Dhb7]MC-RR</td>
<td>PP2A</td>
<td>Human red blood cells</td>
<td>17.9– 49.4</td>
<td>[245]</td>
<td></td>
</tr>
<tr>
<td>MC-LW</td>
<td>rPP1c</td>
<td>Recombinant rabbit skeletal muscle PP1</td>
<td>1.9</td>
<td>[215]</td>
<td></td>
</tr>
<tr>
<td>MC-LW</td>
<td>PP2A</td>
<td>Human hepatocytes</td>
<td>0.29</td>
<td>[229]</td>
<td></td>
</tr>
<tr>
<td>PP2A Human red blood cells</td>
<td>1.1</td>
<td>[215]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rPP2Ac PP2A catalytic subunit</td>
<td>0.114 ± 0.003</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>137 MC-FR rPP2Ac PP2A catalytic subunit</td>
<td>0.069 ± 0.003</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>144 [Dha⁺]MC-YR rPP2Ac PP2A catalytic subunit</td>
<td>0.379 ± 0.003</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>159 [π-Leu¹]MC-LR PP1 Recombinant rabbit skeletal muscle</td>
<td>0.5– 4.43</td>
<td>[162,163]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>163 MC-RR PP1 Rabbit skeletal muscle</td>
<td>0.68</td>
<td>[206]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP1 Chicken gizzard myosin B</td>
<td>3</td>
<td>[237]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP1 Liver of grass carp</td>
<td>3.60</td>
<td>[238]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rPP1c Recombinant rabbit skeletal muscle</td>
<td>1.5</td>
<td>[215]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Human red blood cells</td>
<td>0.241–175</td>
<td>[232,240,241]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Human hepatocytes</td>
<td>0.60</td>
<td>[229]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Mouse brain</td>
<td>0.72–1.4</td>
<td>[209,215,238,243]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Bovine cardiac muscle</td>
<td>1</td>
<td>[237]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Bovine Kidney</td>
<td>10 ± 2</td>
<td>[232]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Rabbit skeletal muscle</td>
<td>0.1</td>
<td>[206]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Liver of grass carp</td>
<td>0.64</td>
<td>[238]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rPP2Ac Recombinant human PP2A catalytic subunit</td>
<td>0.056 ± 0.002</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>163 [(6Z)-Adda⁵]MC-RR rPP2Ac PP2A catalytic subunit</td>
<td>10.1 ± 0.3</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>177 [π-Asp³]MC-HtyR rPP2Ac PP2A catalytic subunit</td>
<td>0.098 ± 0.006</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>179 [π-Asp³,(E)-Dhb⁷]MC-HtyR rPP2Ac PP2A catalytic subunit</td>
<td>0.122 ± 0.005</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180 [π-Asp³,(Z)-Dhb⁷]MC-HtyR rPP2Ac PP2A catalytic subunit</td>
<td>0.110 ± 0.008</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>182 MC-YR PP1 Rabbit skeletal muscle</td>
<td>1.0</td>
<td>[206]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP1 Liver of grass carp</td>
<td>0.90</td>
<td>[238]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Human red blood cells</td>
<td>0.26–9.0</td>
<td>[232,240,241]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Human hepatocytes</td>
<td>0.84</td>
<td>[229]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Bovine kidney</td>
<td>0.09 ± 0.02</td>
<td>[232]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Rabbit skeletal muscle</td>
<td>0.26</td>
<td>[206]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Mouse brain</td>
<td>0.39–1.3</td>
<td>[209,238]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP2A Liver of grass carp</td>
<td>0.40</td>
<td>[238]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rPP2Ac Recombinant human PP2A catalytic subunit</td>
<td>0.125 ± 0.005</td>
<td>[230]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry</td>
<td>Microcystin</td>
<td>Source</td>
<td>IC50 (µM)</td>
<td>Ref.</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>--------</td>
<td>-----------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>234</td>
<td>MC-WR</td>
<td>rPP2Ac</td>
<td>0.18 ± 0.01</td>
<td>[230]</td>
<td></td>
</tr>
<tr>
<td>243</td>
<td>[(-\text{Asp}^3,\text{ADMA}^5,\text{Dhb}^7)]MC-HtyR</td>
<td>PP1</td>
<td>Rabbit skeletal muscle</td>
<td>0.15–0.24</td>
<td>[245]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP2A</td>
<td>Human red blood cells</td>
<td>0.06</td>
<td>[245]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP2A</td>
<td>Bovine heart</td>
<td>0.06</td>
<td>[242]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP4</td>
<td>Porcine testis</td>
<td>0.04</td>
<td>[242]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PP5</td>
<td>Recombinant human PP5</td>
<td>0.5</td>
<td>[242]</td>
</tr>
</tbody>
</table>

* Entry number of each microcystin as indicated in Tables 1 and S1

b Not specified

The inhibition of the PP activity by MCs results primarily from non-covalent interactions largely mediated by the hydrophobic Adda side-chain in position 5, and the Glu\(^6\) moiety. The Adda side-chain is closely accommodated in the hydrophobic groove of the catalytic subunit of the enzyme (Figure 3) and accounts for a significant portion of the toxin’s binding potential. Xing et al. [115] reported that in the catalytic subunit of PP2A, four amino acids Gln122, Ile123, His191, and Trp200 form a hydrophobic cage that accommodates the long hydrophobic Adda side-chain in MC-LR (Figure 3). The second important interactions consist of those of the carboxylic acid of the Glu\(^6\) residue and its adjacent carbonyl group (Adda-amide), which make hydrogen bonds to metal-bound water molecules in the active site of the catalytic subunits of PPs (Figure 3). These observations from X-ray crystallography agree well with in vitro and in vivo toxicity data that found Glu and Adda residues to be essential for the toxicity of MCs, and that modifications to the configuration of the Adda side-chain and esterification of the free carboxylic acid of Glu\(^6\) lead to loss of toxic potential, both [75,246,247].
**Figure 3.** X-ray crystallographic structure showing the interaction of MC-LR with the PP2A catalytic subunit and its adjacent amino acid side chains (based on data provided in [248]). Important interactions are described in the text (Section 6.2). The Mdha<sup>7</sup>, D-Glu<sup>6</sup> and Adda<sup>5</sup> moieties, and selected atoms in these amino acids, are labelled in magenta, red and pale blue, respectively. The large blue spheres are the catalytic metal ions, and the dashed yellow line shows the location of the covalent bond formed over time between the thiol of Cys269 and the Mdha<sup>7</sup> moiety of MC-LR.

A third interaction also occurs between MCs and the catalytic sub-units of PPs via covalent linkage of the electrophilic α,β-unsaturated carbonyl of the Mdha residue in position 7 to the thiol of the cysteine residue in the catalytic subunits of six serine/threonine-specific phosphatases (PP1, PP2, PP3, PP4, PP5 and PP6) by Michael addition (reviewed in [247]). Interactions in the binding pocket between PP1 [114,219] and PP2A [115] and MC-LR are strengthened by a covalent linkage between the Sγ atom of Cys-273 and Cys-269, respectively and the terminal carbon atom (C-3) of the Mdha side chain. However, formation of the covalent bond occurs slowly, and is not essential for the inactivation of the PPs by MCs [98,219,220,242]. For example, Miles et al. [105,106] have shown that Dhb<sup>7</sup>-containing MCs react with thiols about 200 times more slowly than Mdha<sup>7</sup> and Dha<sup>7</sup>-containing MCs [105]. Nodularin-R, which is structurally related to MC-LR but is a pentapeptide containing N-methyldehydrobutyryl (Mdhb) at position 5, did not react at a detectable rate with mercaptoethanol [105]. However, nodularin-R exhibits an IC<sub>50</sub> (1.8 nM), similar to that of MC-LR (2.2 nM) for human red blood cells PP2A [241] and nodularins do not form a covalent linkage to thiol group of Cys-273 in PP1 [116,249]. Similarly, MC congeners containing Dha<sup>7</sup> (e.g. [D-Asp<sup>5</sup>,Dha<sup>7</sup>]MC-LR) and its analogue containing Dhb<sup>7</sup> ([D-Asp<sup>5</sup>,Dhb<sup>7</sup>]MC-LR) have similar IC<sub>50</sub> values for PP2A of 0.254 ± 0.004 and 0.201 ± 0.003 nM, respectively [230].

Two other interactions that also play an important role in the inactivation of PPs have been recently described [247]. The first consists of a hydrogen-bonding interaction between the carboxyl group of the Masp<sup>3</sup> residue to the conserved Arg and Tyr residues in PPs' sub-catalytic units (Arg96 and Tyr134 of PP1; and Arg89 and Tyr127 of PP2). The second consists of a hydrophobic interaction between the aromatic ring of a conserved tyrosine (Tyr272 of PP1; and Tyr265 of PP2) in the PP's β12–β13 loop to the l-Leu<sup>2</sup> residue of MC-LZ congeners [247]. The presence of l-leucine in position 2 can, therefore, explain the higher PP-inhibition potential of MC-LR and its analogues with hydrophobic amino acids in the same position (Table 2).

### 7. Conclusion and future perspectives

Many MCs have been isolated and fully characterized over the last three decades using a combination of analytical techniques and advances in LC-MS/MS instrumentation have recently provided a rapid influx of new MC congeners by tentative identification of trace variants in complex samples. As described in this review, multiple combinations of the variable l-amino acids at positions 2 and 4 and modifications to the other amino acids lead to the structural diversity of 269 MC variants characterized to date. These are formed because of genetic diversity in the mcy gene cluster as well as subsequent chemical and biochemical transformations that can occur after biosynthesis. Although the discovery of mcy gene clusters has led to new avenues for investigating MC regulation at the molecular level, gaps remain in the understanding of how environmental factors can modulate the...
relative abundance of MC congeners in a bloom. Consequently, because the toxicity of a
cyanobacterial bloom depends not only on the total MC accumulation, but also on the structure and
the abundance of each congener, further research focusing on the effects of environmental factors on
modulating the MC profile during bloom episodes is therefore recommended. This is particularly
important in the context of new public health guidelines emerging around the world based on ‘total
microcytins’, which remains a poorly defined concept in the absence of comprehensive data on the
structural diversity and toxic potential of the broader class.

Variations in \textit{in vivo} toxicity between microcystin congeners can be attributed to the relative
importance/contribution of OATP transport versus PP inhibition to overall toxicity of MCs. However,
the large range of values reported in the literature for the LD_{50} by the same route of exposure within
each congener can in part be explained by a combination of biological variability as well as
uncertainty in toxin purity and quantitation. It is, therefore, very important for researchers to
carefully assess and report on the purity and amount of toxins used for toxicity assessments to ensure
comparability of results between different studies and different MC variants.

**Supplementary Materials:** Table S1: Microcystin congeners reported in the literature as identified
from cyanobacterial cultures and field samples.

**Author Contributions:** N.B. provided conceptualization of the need for a review on this topic, led
writing of this article and additional editing. C.O.M. analysis of data including extensive literature
investigation and additional input on literature sources for microcystin congeners, Figure 2, drafting
and editing of manuscript. D.G.B. analysis of data including extensive literature investigation on
microcystin congeners, Figure 3, drafting and editing of manuscript. Z.L. recompilation of data and
input on literature sources for microcystin congeners, Table S1. A.D. recompilation of data and input
on literature sources of toxicity of microcystin congeners, Tables 1 and 2. N.Y.B. recompilation of data,
Figure 1 and analysis and verification of bibliographic references. T.N.Q. provided conceptualization
and additional editing of manuscript.

**Funding:** This work received no external funding.

**Acknowledgments:** We thank Alistair L. Wilkins and Patricia LeBlanc for helpful discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


2. Rigosi, A.; Carey, C.C.; Ibelings, B.W.; Brookes, J.D. The interaction between climate warming and
eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. \textit{Limnol.


75. Namikoshi, M.; Rinehart, K.L.; Sakai, R.; Stotts, R.R.; Dahlem, A.M.; Beasley, V.R.; Carmichael, W.W.; Evans, W.R. Identification of 12 hepatotoxins from a Homer lake bloom of the cyanobacteria...


102. Ferranti, P.; Fabbrocino, S.; Nasi, A.; Caira, S.; Bruno, M.; Serpe, L.; Gallo P. Liquid chromatography coupled to quadrupole time of flight tandem mass spectrometry for microcystin analysis in


