CAWTHRON

Exploring the production of microcystins by *Dolichospermum lemmermannii* from Aotearoa New Zealand

Cawthron Report 4033

World-class science
for a better future

REVIEWED BY: Laura Kelly

APPROVED FOR RELEASE BY: Kirsty Smith

PROJECT NUMBER: 18782

ISSUE DATE: 30 July 2024

RECOMMENDED CITATION: Wood S, Pearman J, Thompson L, Challenger S, Puddick J. 2024. Exploring the production of microcystins by *Dolichospermum lemmermannii* from Aotearoa New Zealand. Nelson: Cawthron Institute. Cawthron Report 4033. Prepared for Ministry for the Environment.

Document no. F-PC21 | Issue no. 2

Exploring the production of microcystins by *Dolichospermum lemmermannii* from Aotearoa New Zealand

Susie Wood, John Pearman, Lucy Thompson, Sarah Challenger, Jonathan Puddick

Prepared for Ministry for the Environment

CAWTHRON

Contents

Executive summary

Cyanobacterial blooms are becoming increasingly common in waterbodies around Aotearoa New Zealand. Some bloom-forming species can produce natural toxins, known as cyanotoxins, which pose a health risk to humans and animals. To protect human health, guidelines for recreational cyanobacteria exposure were developed for use in Aotearoa New Zealand in 2009; these guidelines have recently been updated and are due for release in the second half of 2024. In the revised guidelines, the alert-level thresholds for planktonic cyanobacteria are now based on taxa-specific cell concentrations. This relies on accurate knowledge of the cyanobacterial taxa that produce cyanotoxins in Aotearoa New Zealand. Currently, only four planktonic cyanobacterial taxa are known cyanotoxin producers in Aotearoa New Zealand: *Cuspidothrix issatschenkoi* (anatoxins), *Microcystis* spp. (microcystins), *Nodularia spumigena* (nodularins) and *Raphidiopsis raciborskii* (cylindrospermopsins).

In May 2023, the presence of a gene involved in the production of the hepatotoxic cyanotoxin microcystin (*mcyE*) was detected in a sample from Lake Opuha (Canterbury), despite no known toxin producers being present in the sample. Further work identified *Dolichospermum lemmermannii* as the species containing the cyanotoxin production gene, and chemical analysis indicated that it might produce microcystins. *Dolichospermum lemmermannii* is currently categorised as non-toxic in Aotearoa New Zealand. If *Dolichospermum lemmermannii* does produce microcystins, then implementation of the updated cyanobacteria guidelines may result in an underestimation of the human health risk for recreational water users.

To further investigate the possible microcystin production by *Dolichospermum lemmermannii* in Aotearoa New Zealand, Cawthron Institute was commissioned by the Ministry for the Environment (MfE) to:

- evaluate *Dolichospermum lemmermannii* cultures for the presence of the *mcyE* gene
- evaluate *mcyE*-positive cultures for microcystins using liquid chromatography–tandem mass spectrometry (LC–MS/MS)
- undertake genome sequencing of *Dolichospermum lemmermannii* to determine whether all the genes involved in microcystin production were present.

Twelve cultures of *Dolichospermum lemmermannii* sourced from the Cawthron Institute Culture Collection of Microalgae were screened for the *mcyE* gene using polymerase chain reaction (PCR). Only three cultures were positive for the gene.

One of the *mcyE*-positive cultures (CAWBG680), a *mcyE*-negative culture, and the original water sample from Lake Opuha were 'cleaned-up' and concentrated using solid-phase extraction cartridges and analysed using LC–MS/MS to determine if microcystins were present. No microcystins were identified in the water sample or cultures; instead, multiple anabaenopeptilides were identified. Investigations determined that the anabaenopeptilides were the cause of the interference that led to the tentative LC–MS/MS identification of microcystins in the Lake Opuha sample. Other possible microcystin signals were also investigated using LC–MS/MS techniques, but these were confirmed to be unrelated compounds.

Genomic sequencing was undertaken on one of the *mcyE*-positive cultures (CAWBG680) and the microcystin synthase (*mcy*) gene operon annotated. The exact number and arrangement of microcystin genes within a genome can vary among species and strains, although it typically includes the core genes (*mcyA*, *mcyB*, *mcyC*, *mcyD* and *mcyE*), genes encoding tailoring enzymes (*mcyG*, *mcyI* and *mcyJ*), a racemase (*mcyF*) and a possible transporter gene (*mcyH*). In the *mcyE*-positive culture that was sequenced, the *mcyA–G* and *mcyJ* genes were present; however, the taxonomic classification of *mcyF* suggested that it was present only in bacteria associated with the culture and was not from *Dolichospermum lemmermannii*. The *mcyH*, and *mcyI* genes were also absent from the *Dolichospermum lemmermannii* genome.

The combined evidence from the LC–MS/MS analysis and *mcy* gene cluster sequencing demonstrates that *Dolichospermum lemmermannii* CAWBG680 cannot produce microcystins. The identification of anabaenopeptilides as the cause for the tentative microcystin identification in the original Lake Opuha sample also alleviates initial concerns. These results support the adoption of the revised alert-level framework for planktonic cyanobacteria in the cyanobacteria guidelines for recreational fresh waters, where *Dolichospermum* spp. are not included as a toxin producer. There is a small possibility that other *Dolichospermum* strains / species can produce microcystins or other cyanotoxins in Aotearoa New Zealand, or could be introduced here. We therefore recommend that when *Dolichospermum* is dominant, continued testing of environmental samples using PCR and LC–MS/MS is undertaken. The results from this study also highlight that while PCR approaches are very valuable for detecting potential cyanotoxin producers, positive results need to be confirmed using chemical analyses.

1. Introduction

Cyanobacterial blooms are becoming increasingly prevalent in lakes across Aotearoa New Zealand, particularly in waterbodies that are eutrophic. These blooms can cause a suite of environmental issues such as reduced water quality and low dissolved oxygen levels in bottom waters (e.g. Smucker et al. 2021). Additionally, because some bloom-forming species produce natural toxins, which are known as cyanotoxins, they also pose a health risk to humans and animals who come in contact with or consume cyanotoxin-contaminated water (Wood et al. 2006). Because of the human health risk posed by cyanobacteria, guidelines for cyanobacteria in recreational fresh waters were developed for use in Aotearoa New Zealand in 2009 (MfE and MoH 2009). The guidelines include three-tier alert-level frameworks for monitoring and managing human health risks associated with planktonic and benthic cyanobacteria. Originally, the thresholds used in the alert-level framework for planktonic cyanobacteria included all known cyanotoxin-producing taxa (including those identified as toxin producers overseas). The guidelines have recently been revised (due for release in the second half of 2024) and the revised alert-level framework for planktonic cyanobacteria uses thresholds based on taxa-specific cell concentrations. To be effective, this risk management approach relies on accurate knowledge of the cyanobacterial species that produce cyanotoxins in Aotearoa New Zealand. Currently, only four planktonic genera / species are known cyanotoxin producers in Aotearoa New Zealand: *Cuspidothrix issatschenkoi* (anatoxins), *Microcystis* spp. (microcystins), *Nodularia spumigena* (nodularins) and *Raphidiopsis raciborskii* (cylindrospermopsins). The only benthic cyanobacteria included in the guidelines is *Microcoleus*. This is not included in the present report because the thresholds for *Microcoleus* genera are based on percent cover, not cell counts. To effectively implement the updated version of recreational cyanobacteria guidelines for planktonic cyanobacteria, it is essential to have robust knowledge on which genera / species are toxin producers in Aotearoa New Zealand, and regional councils and other water managers have requested that research is undertaken to explore events where the presence of a new toxin production is suspected.

In May 2023, Environment Canterbury sent four water samples to the Cawthron Institute (Cawthron; Nelson). These samples contained cyanobacteria and were evaluated using targeted quantitative polymerase chain reaction (PCR) assays for genes involved in cyanotoxin production. Of most interest was the presence of the microcystin / nodularin synthase gene *mcyE / ndaF* (a gene involved in the production of the cyanotoxin microcystin) in the sample from Lake Opuha, despite the presence of no known producers (Wood and Puddick 2023). DNA sequencing of a portion of the *mcyE* gene can sometimes identify the toxin-producing species due to variations in the genetic code between different cyanobacterial taxa. Comparing the sequence from the sample with genetic databases indicated that the gene sequence from Lake Opuha matched with 99.5% similarity to the *mcyE* gene from *Dolichospermum* sp. (previously *Anabaena* sp.; Genbank accession number KF219514).

Microscopic analysis of the Lake Opuha sample confirmed the presence of *Dolichospermum lemmermannii*. In previous unpublished work, we have identified the *mcyE* gene in *Dolichospermum lemmermannii* cultures from the Cawthron Institute Culture Collection of Microalgae (CICCM; Wood, unpublished data). *Dolichospermum* spp. are known to produce microcystins (Li et al. 2016) in Europe, Canada, and most recently they have been found in the USA (e.g. Harada et al. 1991; Kobos et al. 2013;

Dreher et al. 2019). To the best of our knowledge, microcystin-producing *Dolichospermum* spp. have not been documented in the Southern Hemisphere.

The observation of the *mcyE* gene from *Dolichospermum lemmermannii* in Lake Opuha (Wood and Puddick 2023) prompted further analysis for microcystins using liquid chromatography–tandem mass spectrometry (LC–MS/MS). While no commonly observed microcystins congeners were detected,[1] further MS/MS analysis tentatively indicated the possibility of new microcystin congeners in the Lake Opuha sample (Wood and Puddick 2023). However, it was not possible to confirm this finding with the resources available and without a more concentrated sample.

Currently, *Dolichospermum lemmermannii* is categorised as non-toxic in Aotearoa New Zealand. If *Dolichospermum lemmermannii* does produce microcystins, then implementation of the updated guidelines for cyanobacteria in recreational fresh waters may result in an underestimation of the human health risk for water users. Analysis of *Dolichospermum lemmermannii* cultures from the CICCM provides an efficient way to explore whether this species is likely to be a toxin producer, as the work can be carried out without identifying and collecting suitable samples from cyanobacterial blooms.

To further investigate the possible presence of microcystin production by *Dolichospermum lemmermannii* in Aotearoa New Zealand, Cawthron was commissioned by the Ministry for the Environment (MfE) to:

1) evaluate cultures of *Dolichospermum lemmermannii* from the CICCM for the presence of the *mcyE* gene, a ubiquitous gene involved in microcystin production
2) re-examine the Lake Opuha sample and *Dolichospermum lemmermannii* cultures that had *mcyE* detections in Milestone 1 and assess for novel microcystins using LC–MS/MS
3) undertake genome sequencing on a *Dolichospermum lemmermannii* culture that had a *mcyE* detection in Milestone 1 and annotate the genomes to a level where the genes involved in microcystin production can be explored.

The next three sections of this report sequentially describe the work undertaken in Milestones 1–3 (Sections 2–4), and the final section provides overall conclusions and recommendations (Section 5).

[1] There are over 300 known microcystin congeners: see Spoof and Catherine (2017).

2. The presence of the *mcyE* gene in *Dolichospermum lemmermannii* cultures

2.1 Introduction

The microcystin synthase (*mcy*) gene cluster in cyanobacteria typically consists of multiple genes organised into several modules responsible for different steps in microcystin biosynthesis. While the exact number and arrangement of genes can vary among species and strains, a typical *mcy* gene cluster includes core genes involved in the synthesis of microcystin, such as *mcyA*, *mcyB*, *mcyC*, *mcyD* and *mcyE* (Yancey et al. 2022). Additionally, there are often genes encoding tailoring enzymes responsible for modifying the microcystin structure, such as *mcyG*, *mcyI* and *mcyJ*, a racemase gene (*mcyF*) and a putative transporter gene, *mcyH* (Tillett et al. 2000; Sielaff et al. 2003; Pearson et al. 2004; Pearson et al. 2007). The variation in the *mcy* gene cluster among species and strains primarily arises from differences in the presence, absence, arrangement and sequence of the genes encoding these biosynthetic enzymes. This variability can lead to the production of structurally diverse microcystin variants with varying toxicity levels and environmental persistence (Rantala et al. 2006).

A number of different PCR assays that screen for a region of core *mcy* genes have been designed (e.g. Vaitomaa et al. 2003). Screening for one or more of the core *mcy* genes has become a routine tool for assessing the toxin production potential of a culture or environmental sample. However, research has shown that some strains can have a deletion of one or more genes in the *mcy* cluster or a loss-of-function mutation, and thus are not able to produce microcystins (Christiansen et al. 2008). Screening for a single *mcy* gene can therefore result in false positives (i.e. there is a gene detection, but there is no toxin present in the sample). Current best practice is to follow up positive PCR results with toxin analysis.

For this study, we screened for the *mcyE* gene in *Dolichospermum lemmermannii* cultures from the CICCM. The *mcyE* gene encodes an enzyme involved in the formation of the peptide bond that links key amino acids together to form the microcystin structure. It is considered essential for the production of microcystins, and we have used this assay widely in the past (e.g. Wood et al. 2017).

2.2 Methods and results

A total of 12 *Dolichospermum lemmermannii* cultures were sourced from the CICCM (Table 1). Approximately 50 mL from each culture was centrifuged (3,000 × *g*, 10 min) and the supernatant was discarded. The resulting pellet was extracted using the DNeasy PowerSoil™ DNA isolation kit (QIAGEN, Germany) and the automated QIAcube DNA extraction robot (QIAGEN), following the manufacturer's protocol.

Table 1. List of *Dolichospermum lemmermannii* cultures sourced from the Cawthron Institute Culture Collection of Microalgae (CICCM) and screened for the *mcyE* gene using polymerase chain reaction (PCR) in this study.

CICCM code	Original location strain was isolated from	*mcyE* (+/−)
CAWBG564	Lake Pauri (Wellington)	−
CAWBG565	Lake Pauri (Wellington)	−
CAWBG567	Lake Tūtira (Hawkes Bay)	+
CAWBG568	Lake Tūtira (Hawkes Bay)	+
CAWBG572	Lake Rotokare (Taranaki)	−
CAWBG573	Lake Rotokare (Taranaki)	−
CAWBG574	Lake Rotokauri (Waikato)	−
CAWBG576	Lake Dudding (Manawatū)	−
CAWBG577	Lake Dudding (Manawatū)	−
CAWBG592	Lake Waihola (Otago)	−
CAWBG593	Lake Waihola (Otago)	−
CAWBG680	Lake Taupō (Waikato)	+

PCR amplification was undertaken on the DNA extracts using primers targeting a portion of the *mcyE* gene, as described in Wood et al. (2017). The PCR products were run on a 1.5% agarose gel, and bands of the expected size (400 bp) were observed for the *mcyE* gene in three of the cultures: CAWBG567, CAWBG568 and CAWBG680 (Table 1).

2.3 Discussion

The presence of the *mcyE* gene was identified in three cultures using targeted PCR analysis. This suggests that these cultures could have the genetic potential to produce microcystins. To further investigate this, genome analysis was used to explore the entire microcystin synthetase gene cluster of *Dolichospermum lemmermannii* CAWBG680 (Section 4). Because of the high biomass required to isolate sufficient DNA for genome sequencing, only one of the three *mcyE*-positive cultures was evaluated. Sufficient biomass of CAWBG567 and CAWBG568 could not be grown within the time frame of the project.

3. Screening for potential microcystins using LC–MS/MS

3.1 Introduction

Microcystins are the most common cyanotoxins globally (Harke et al. 2016), including in Aotearoa New Zealand (Wood et al. 2017). Current research indicates that microcystin production in Aotearoa New Zealand is limited to the cyanobacterial genera *Microcystis* (several species or strains that grow as planktonic colonies or single cells), *Planktothrix* (an unclassified species that grows as a benthic film in shallow ponds) and *Nostoc* (several species that grow as benthic colonies; Puddick et al. 2019). However, new toxin-producing cyanobacterial species are always being identified, and there is the possibility that overseas microcystin-producing species could be introduced to freshwater habitats in Aotearoa New Zealand.

Currently, microcystin analysis is primarily undertaken by LC–MS/MS. This technique separates compounds in a sample extract using LC (generally based on the hydrophobicity of the compound) and analyses the mass-to-charge ratio (*m/z*) of the compound using MS. This is a powerful means for identification, as all compounds have a certain mass, allowing for broad differentiation between different microcystin congeners. Additional specificity is added by fragmenting the ion of interest, measuring the *m/z* of the resulting fragments and analysing for fragments specific to the compound of interest. Furthermore, the MS signals generated through LC–MS/MS analyses are generally proportional to the concentration of the compound, allowing for simultaneous quantitation alongside highly specific identification.

Beyond the quantification and rapid identification of known compounds, LC–MS/MS is also useful for the structural characterisation of compounds. This process can be used to confirm the identity of a compound to a higher degree of certainty, or to determine the structure of an unknown compound.

In this study, we used LC–MS/MS to investigate whether microcystins were present in a *mcyE*-positive water sample and *Dolichospermum lemmermannii* culture.

3.2 Methods and results

One *mcyE*-positive *Dolichospermum lemmermannii* culture (CAWBG680) and a *mcyE*-negative culture (CAWBG564) were grown at 18 °C using a 12 h:12 h light / dark regime (20 µmol photon/m^2/s). Potential microcystins were extracted from the cultures (40 mL each) by supplementing with 40 µL of concentrated formic acid (0.1% final concentration), mixing and freezing at −20 °C. The frozen culture was thawed and sonicated for 30 min. This freezing, thawing and sonicating process was repeated once more before the extract was clarified by centrifugation (3,200 × *g* for 10 min). The Lake Opuha water

sample was similarly extracted, but as only 10 mL of sample was available, only 10 µL of concentrated formic acid was used.

The aqueous extracts were cleaned-up and concentrated using Strata-X solid-phase extraction (SPE) cartridges (33 µm polymeric reversed phase, 100 mg/3 mL; Phenomenex). Each SPE cartridge was conditioned with 100% methanol (10 mL) and equilibrated with 10% methanol + 0.1% formic acid (10 mL), before the entire sample extract was loaded. The SPE cartridge was washed with 10% methanol + 0.1% formic acid (10 mL) to remove compounds with low affinity to the Strata-X, before the compounds of interest were eluted using 80% methanol + 0.1% formic acid (10 mL). The SPE eluent was dried down at 40 °C under a stream of nitrogen gas and resuspended in 80% methanol + 0.1% formic acid (1 mL). Samples were stored at −20 °C between LC–MS/MS analyses.

Cleaned-up and concentrated samples were analysed using an Acquity I-Class ultra-performance liquid chromatography system (Waters Co.) using a C_{18} column (Acquity BEH-C_{18}, 1.7 µm, 50 × 2.1 mm; Waters Co.) maintained at 40 °C in a column oven. Sample components were eluted using a flow rate of 0.4 mL/min and a gradient of 10% acetonitrile (mobile phase A) to 90% acetonitrile (mobile phase B), each containing 100 mM formic acid and 4 mM ammonia. The samples were injected at 5% mobile phase B and held for 12 s before a linear gradient up to 35% mobile phase B over 24 s, to 50% mobile phase B over a further 72 s and to 65% mobile phase B over a final 42 s, before flushing with 100% mobile phase B (at 0.6 mL/min) and returning to the initial column conditions. Sample components were analysed on a Xevo-TQS mass spectrometer (Waters Co.) operated in positive-ion electrospray ionisation mode (source temperature 150 °C; capillary voltage 1.5 kV; nitrogen desolvation gas 1,000 L/h at 500 °C; cone gas 150 L/h).

Three LC–MS/MS analyses were used to investigate whether microcystins were present in the samples. Microcystin analysis using multiple-reaction monitoring (MRM) was undertaken as per the microcystin quantification described in Puddick et al. (2016). MS/MS characterisation of selected ions was undertaken similar to that described for candidate microcystins in Kleinteich et al. (2018). Microcystin precursor ion screening was undertaken as described in Kleinteich et al. (2018).

Microcystin MRM analysis coupled with LC–MS scans in positive and negative ion modes were used to identify most target ions for further investigation using MS/MS characterisation techniques (*m/z* 930, 958, 977, 984, 1,032 and 1,048). An additional target ion (*m/z* 533) was identified using the microcystin precursor ion screening method (which detects the parent ion responsible for fragment ions that are diagnostic of microcystins; Kleinteich et al. 2018). The microcystin MRM method (Puddick et al. 2016) analyses for 21 microcystin congeners frequently observed in microcystin-producing cyanobacteria from Aotearoa New Zealand (Puddick et al. 2019) and nodularin-R. Analysis of the cleaned-up and concentrated Lake Opuha and CAWBG680 samples using the microcystin MRM method had signals in the channels for MC-YR, MC-LR, dmMC-LR, MC-FR, MC-RA and MC-RAba (Table 2). No interfering signals were observed in the cleaned-up and concentrated CAWBG564 sample when it was analysed.

Table 2. Information on the interfering signals observed in the microcystin multiple-reaction monitoring (MRM) analysis for the cleaned-up and concentrated Lake Opuha and *Dolichospermum lemmermannii* CAWBG680 samples.

Microcystin (MC)	Expected RT (min)	Expected $[M+H]^+$	Lake Opuha		CAWBG680		Actual target $[M+H]^+$	Explanation
			Detected by MRM	Observed RT (min)	Detected by MRM	Observed RT (min)		
MC-YR	1.33	1,045	+	1.40	+	1.40	1,048	Interference from closely related mass (M_r 1,047 Da)
MC-LR	1.38	995	+	1.50	–	NA	977	Sodium ion of M_r 976 Da compound
dmMC-LR	1.39	981	+	1.27	+	1.63	958	Sodium ion of M_r 957 Da compound
MC-FR	1.44	1,029	+	1.61	+	1.41	1,032	Interference from closely related mass (M_r 1,032 Da)
MC-RA	1.60	953	–	NA	+	1.45	930	Sodium ion of M_r 930 Da compound
MC-RAba	1.77	967	+	1.54	+	1.50	984	Water-loss ion of M_r 983 Da compound

m/z = Mass-to-charge ratio

M_r = Molecular mass

$[M+H]^+$ = Singly protonated ion

RT = Retention time using the chromatography described in Puddick et al. (2016)

Although the signals had similar mass spectrometric properties to each intended microcystin congener (similar parent mass and the formation of a *m/z* 135 fragment ion), the retention time of each compound was different from that expected for the microcystin (Table 2). This suggested that the compounds were not the intended microcystins and could be a different microcystin structural analogue or an entirely different compound. Further analysis of the Lake Opuha and CAWBG680 samples using mass spectrometry scans (to determine the *m/z* of the positive and negative ions formed from the samples) suggested that all the tentative microcystin signals were due to different parent masses from the expected microcystins (Table 2). For example, the compound causing the signal in the MC-YR MRM channel (*m/z* 1,045.7 > 135.1) was due to a compound with a closely related mass to MC-YR (M_r 1,047 Da; 3 Da higher than MC-YR). When a non-target compound is present in very high concentrations, or ionises very effectively, this can lead to some of the signal bleeding into the mass window for the target compound – causing interference in an MRM channel. Other interfering signals were due to the formation of sodium ions and water-loss products that had a closely related *m/z* to the target microcystin. In these cases, the parent mass for each compound was different from the *m/z* anticipated by the microcystin MRM method.

To determine the identity of the interfering compounds, MS/MS investigations were undertaken. Here, a first mass analyser is used to separate the ion of interest (in this case the singly protonated ion for each interfering compound: $[M+H]^+$) from other ions present, and the ion is fragmented by collision-induced dissociation (CID; where the parent ion is accelerated into a neutral gas, in this case argon, to break it apart). The *m/z* of the resulting fragments are then measured using a second mass analyser, and the inferred masses of these fragments provide structural clues about the compound. For example, microcystin congeners will commonly form a *m/z* 135 and 163 fragments diagnostic of the Adda sidechain found in microcystins, and differences in the *m/z* of other fragments relate to the masses of the amino acids that comprise the cyclic peptide (e.g. Puddick et al. 2013; Puddick et al. 2014).

Based on the microcystin MRM evaluation, we investigated the identity of the *m/z* 930, 958, 977, 984, 1,032 and 1,048 ions in the Lake Opuha and / or CAWBG680 samples using MS/MS. All of the MS/MS spectra for the compounds of interest contained a *m/z* 164 fragment ion (Figure 1), suggesting that the compounds might be structurally related. The *m/z* 958, 984, 1,032 and 1,048 compounds all contained fragment ions at *m/z* 86, 209, 309 and 400 (Figure 1), suggesting a higher degree of structural relatedness. Literature evaluation indicated that several of the observed fragment ions (*m/z* 86, 164, 181, 209) are present in the MS/MS spectra of anabaenopeptilides (Sanz et al. 2015).

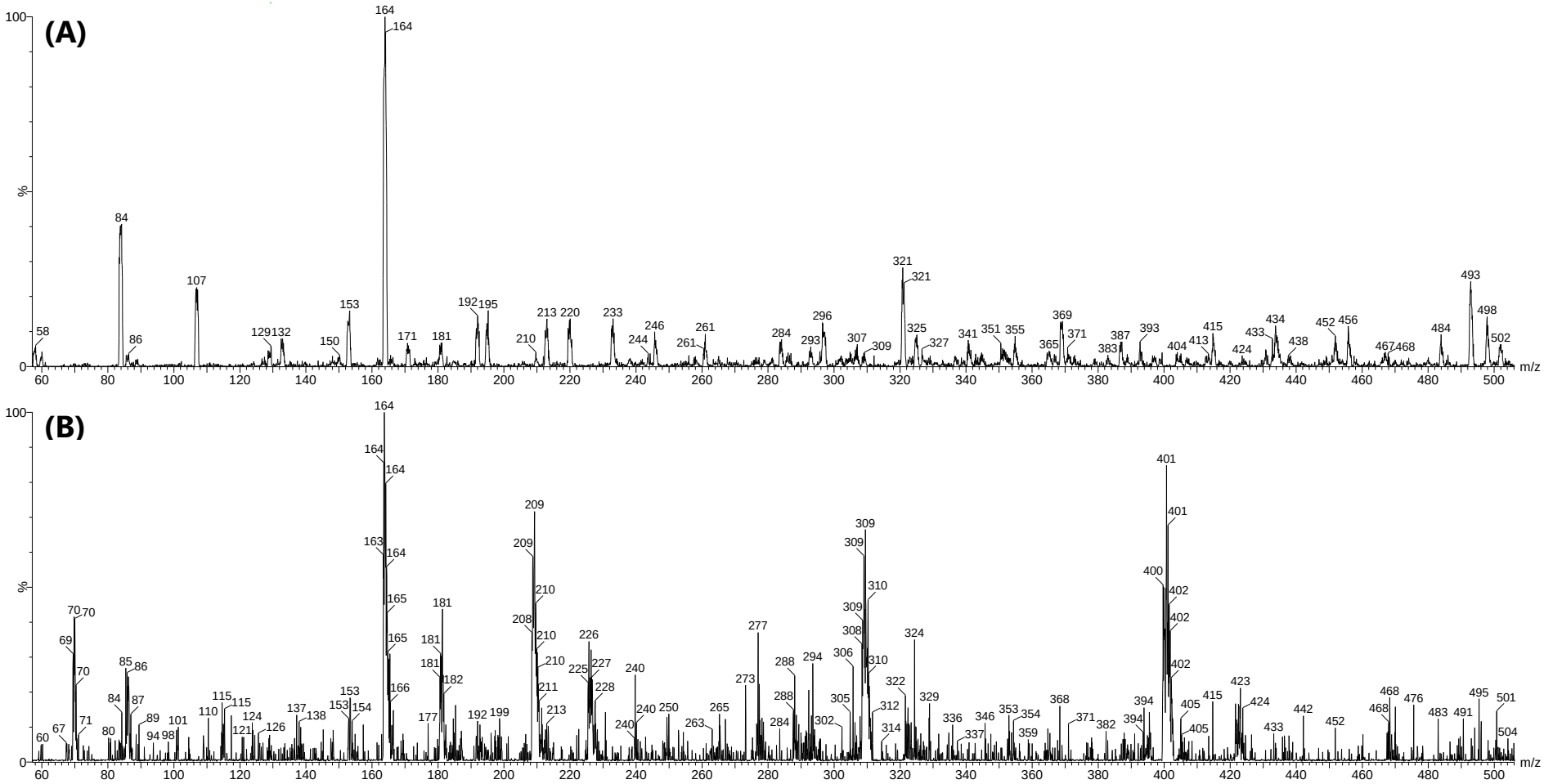

Figure 1. Tandem mass spectra for the *m/z* 930 (A), 958 (B), 977 (C), 984 (D), 1,032 (E) and 1,048 (F) ions in the Lake Opuha and / or *Dolichospermum lemmermannii* CAWBG680 samples.

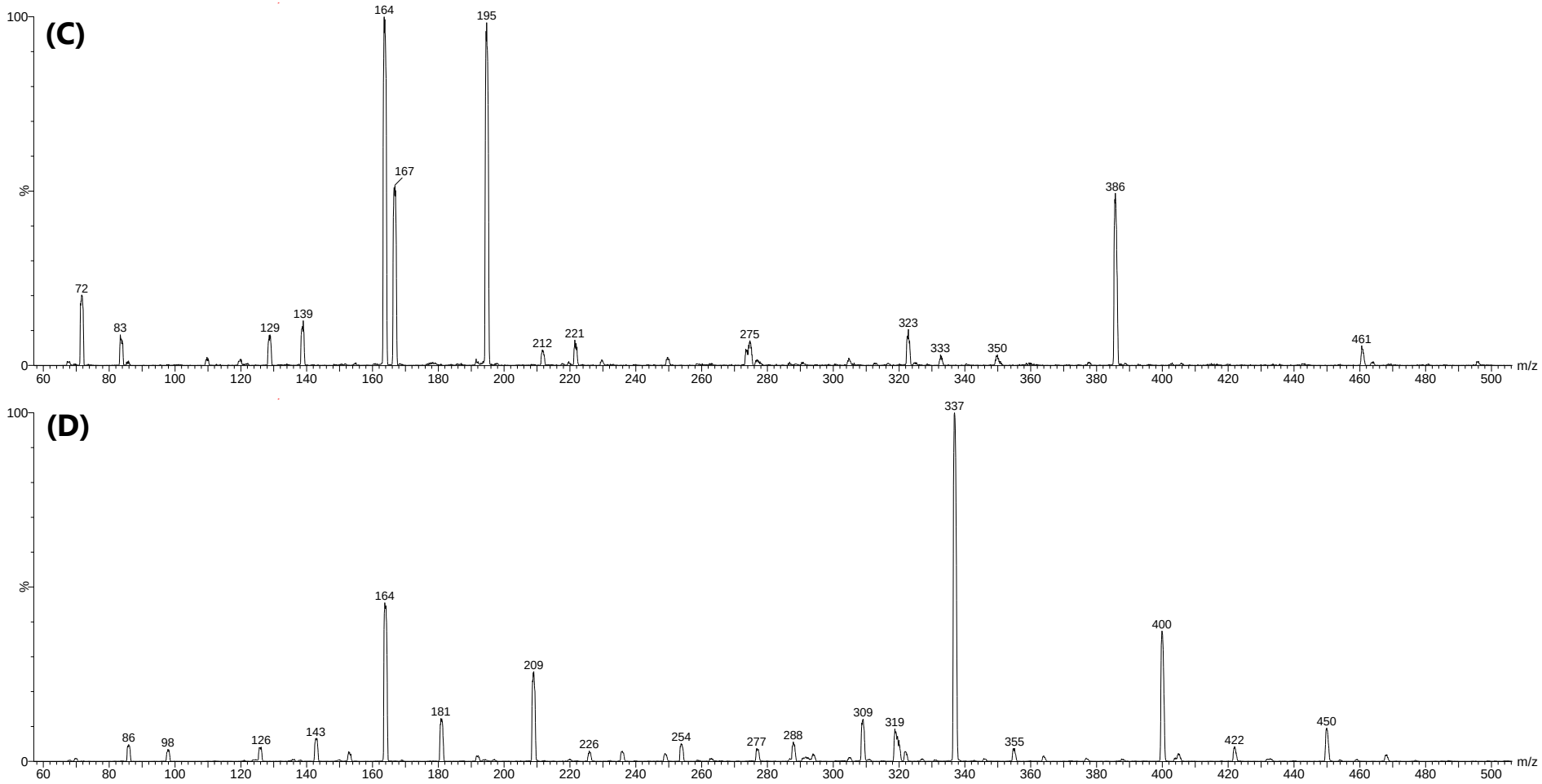

Figure 1 (continued). Tandem mass spectra for the *m/z* 930 (A), 958 (B), 977 (C), 984 (D), 1,032 (E) and 1,048 (F) ions in the Lake Opuha and / or *Dolichospermum lemmermannii* CAWBG680 samples.

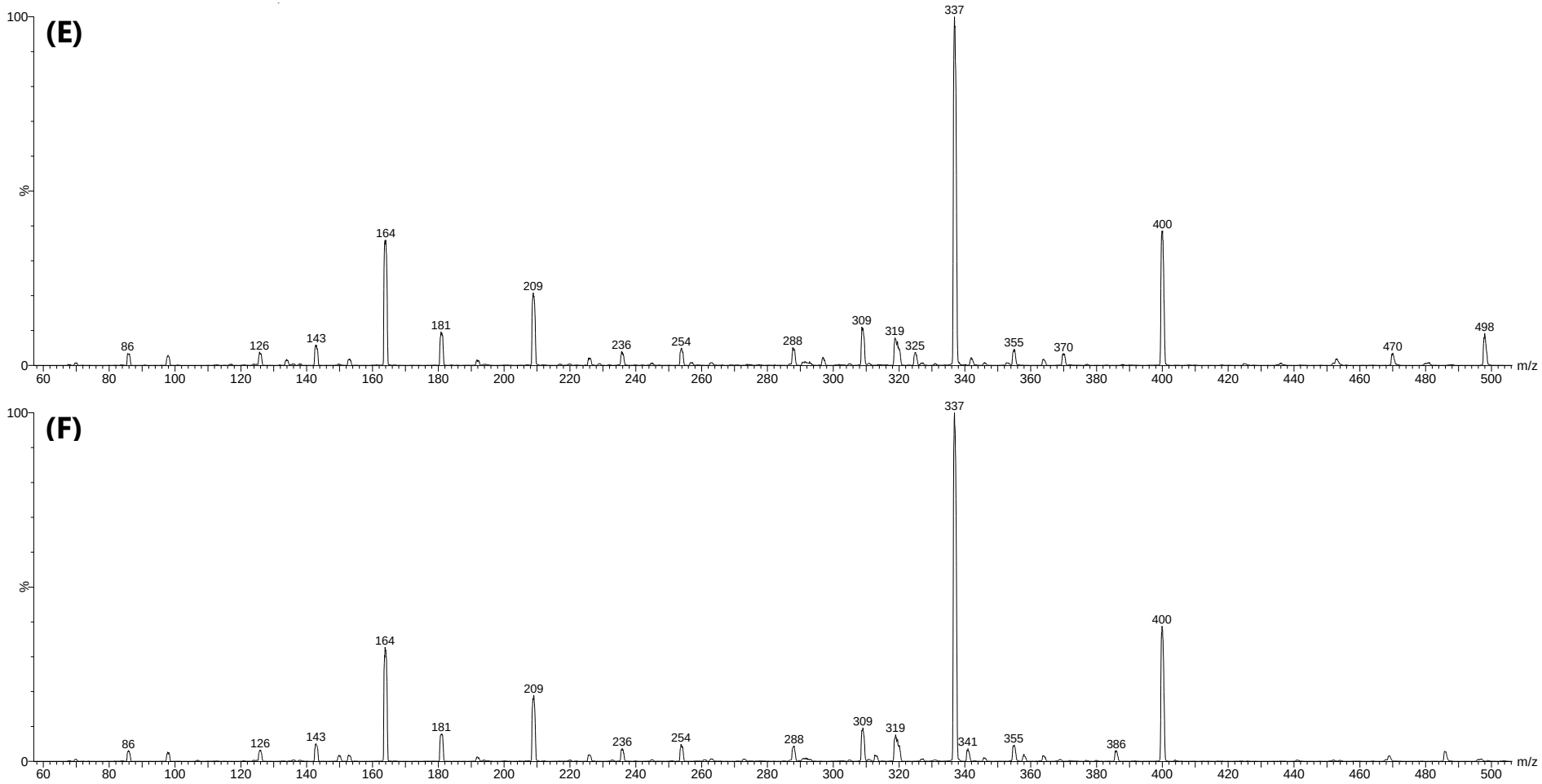

Figure 1 (continued). Tandem mass spectra for the *m/z* 930 (A), 958 (B), 977 (C), 984 (D), 1,032 (E) and 1,048 (F) ions in the Lake Opuha and / or *Dolichospermum lemmermannii* CAWBG680 samples.

Anabaenopeptilides (also called aeruginopeptins, cyanopeptolins, dolostatins, hofmannolins, microcystilides, micropeptins, nostocyclins, nostopeptins, oscillapeptilides, oscillapeptins, planktopeptins, scyptolins, somamides, symplostatins) are cyclic peptides produced by a range of cyanobacteria (Welker and von Döhren 2006), including *Dolichospermum*. They are characterised by the presence of Ahp (3-amino-6-hydroxy-2-piperidone) in Position 3 and a peptide ring generally cyclised through an ester bond between the β-hydroxy group of the Position 1 threonine and the carboxyl group of the Position 6 amino acid. Anabaenopeptilides also contain a sidechain of variable lengths attached to the Position 1 amino acid (which is commonly threonine). Besides the Position 3 Ahp, a range of amino acids have been observed in the ring and sidechain – resulting in many different anabaenopeptilide structural congeners. See Figure 2 for the structure of cyanopeptolin-A and Welker and von Döhren (2006) for more information on the structural diversity of anabaenopeptilides.

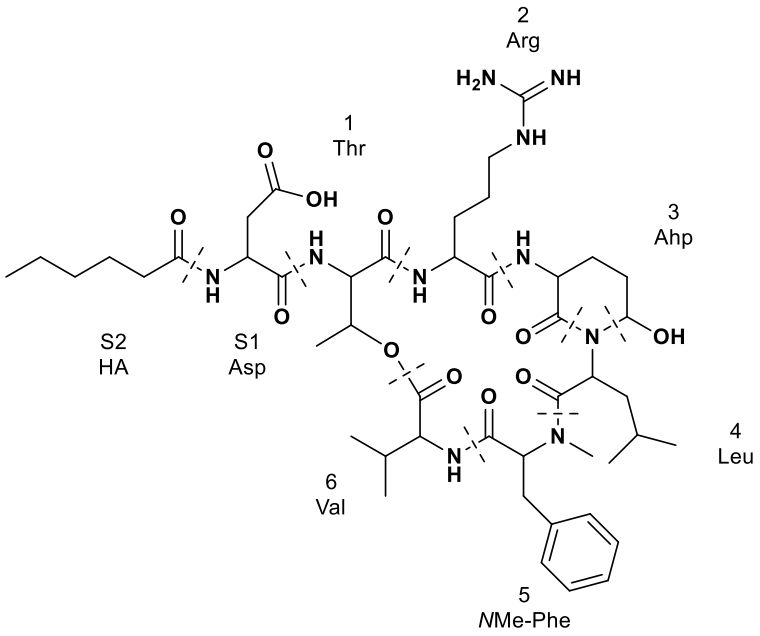

Figure 2. Chemical structure of cyanopeptolin-a (cyanopeptolins are also referred to as anabaenopeptilides; Ahp = 3-amino-6-hydroxy-2-piperidone, Arg = arginine, Asp = aspartic acid, HA = hexanoic acid, Leu = leucine, *N*Me-Phe = *N*-methylphenylalanine, Thr = threonine, Val = valine).

The MS/MS spectra for the *m/z* 984, 1,032 and 1,048 ions in the *Dolichospermum lemmermannii* CAWBG680 sample (Figure 3) were investigated in greater detail to determine the putative structures of the compounds and confirm that they were anabaenopeptilides. The *m/z* 126 and 154 fragments in the MS/MS spectra for each compound suggested that the sidechain terminated in a *N*-propyl-proline formed from propanoic acid (PA; S3) and proline (Pro; S2). The *m/z* 254 and 337 fragment ions indicated that valine (Val) was the sidechain amino acid (S1) and that threonine (Thr) was present in Position 1 (forming the β-lactone ring characteristic of anabaenopeptilides through an ester bond). The presence of *N*-methyl-*O*-methyltyrosine (*N*Me-*O*Me-Tyr) was inferred by a *m/z* 164 immonium ion and the *m/z* 400 fragment ion where it was joined to Ahp–Leu (minus water).

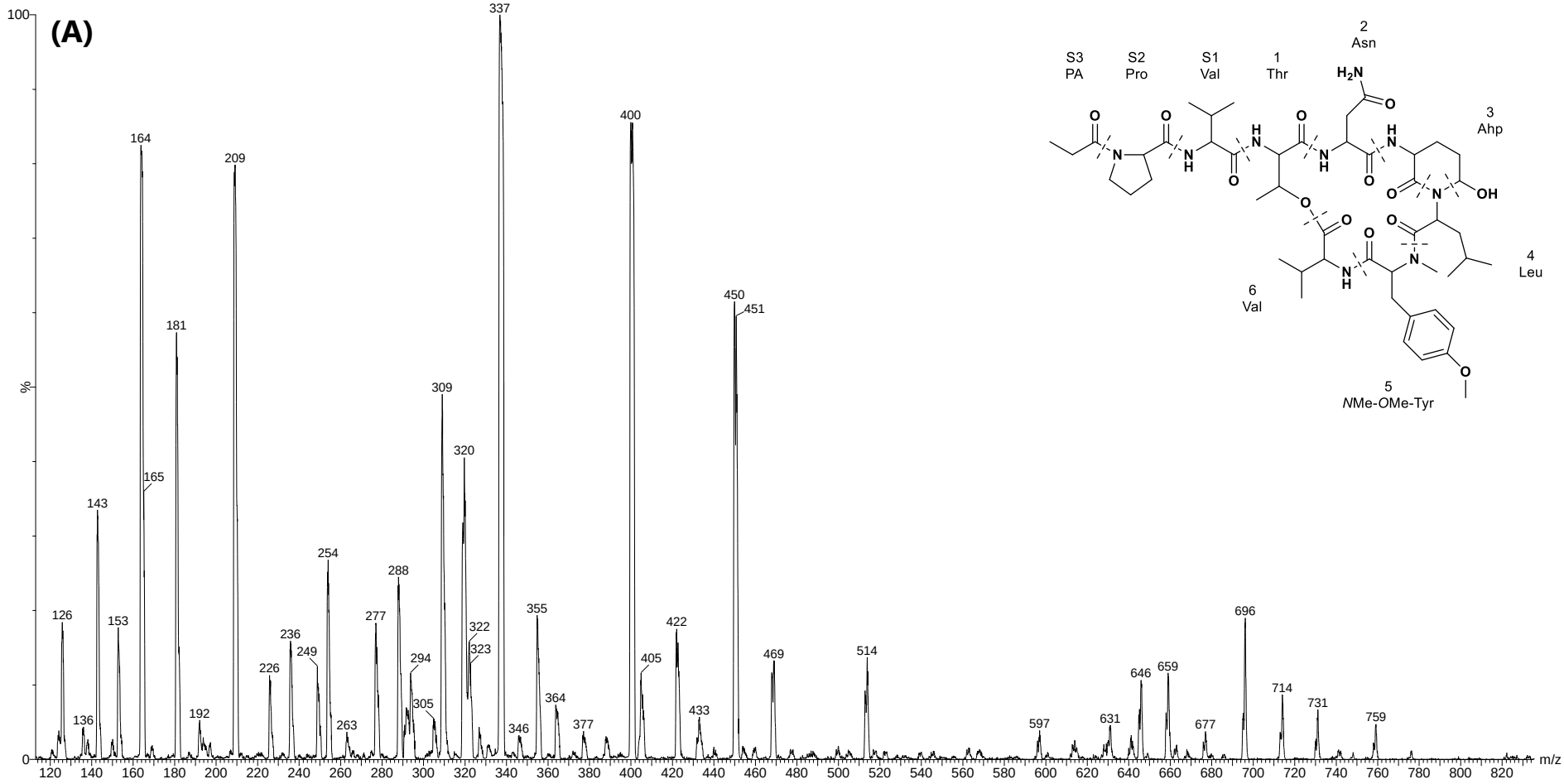

Figure 3. Tandem mass spectra for the *m/z* 984 (A), 1,032 (B) and 1,048 (C) ions from the *Dolichospermum lemmermannii* CAWBG680 sample, including putative structures of the anabaenopeptilides determined using the masses and pattern of the fragment ions.

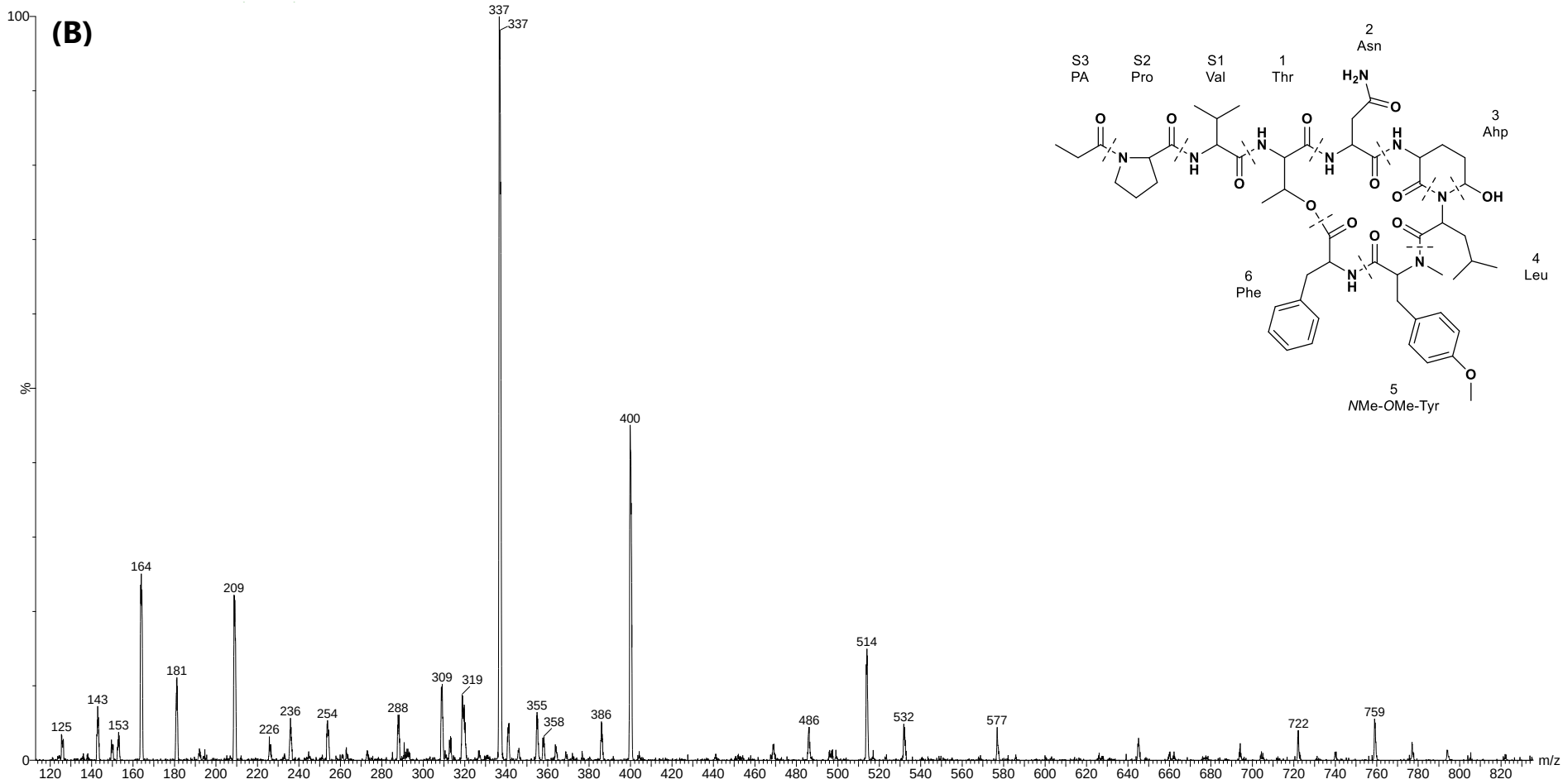

Figure 3 (continued). Tandem mass spectra for the *m/z* 984 (A), 1,032 (B) and 1,048 (C) ions from the *Dolichospermum lemmermannii* CAWBG680 sample, including putative structures of the anabaenopeptilides determined using the masses and pattern of the fragment ions.

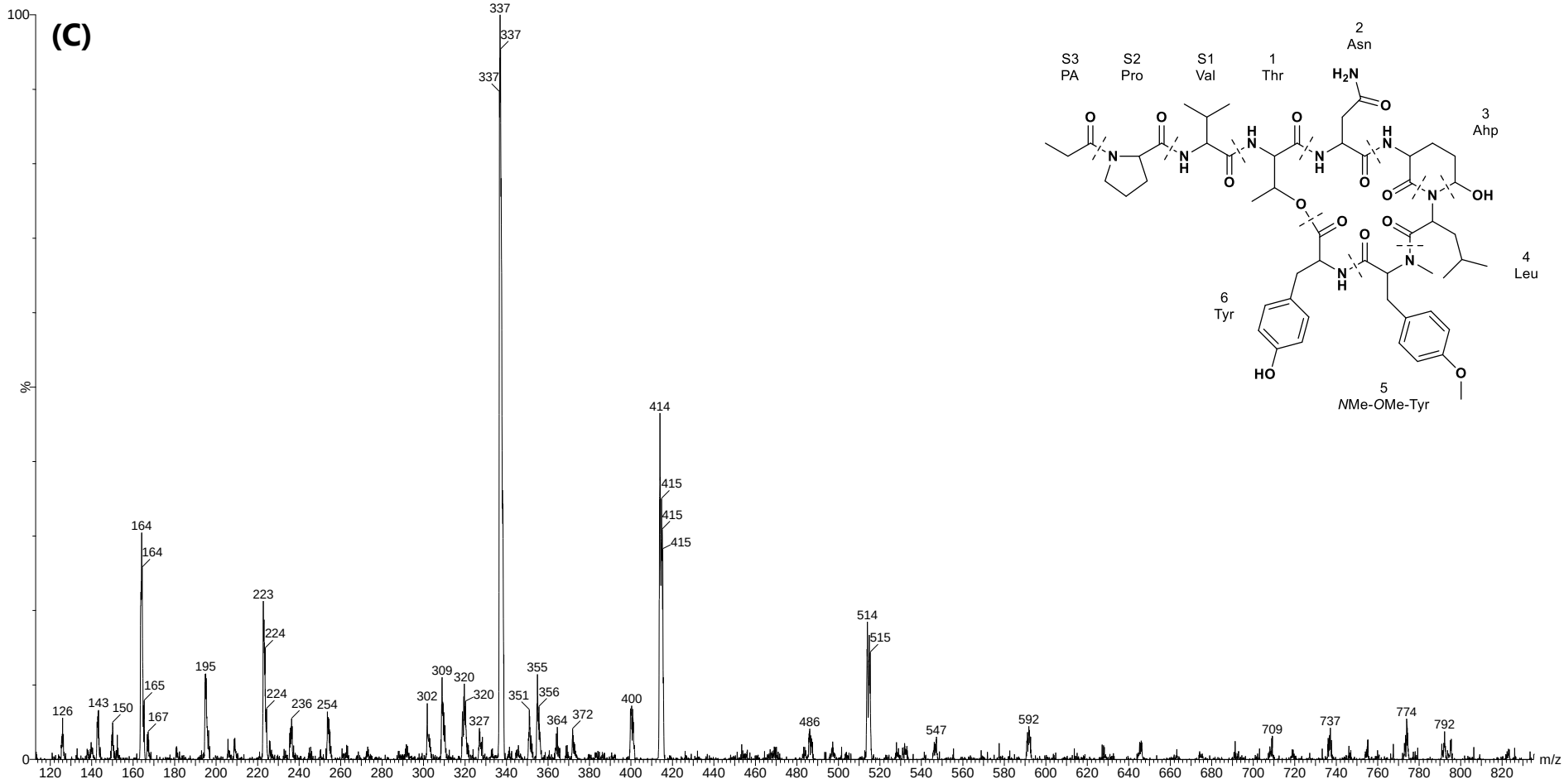

Figure 3 (continued). Tandem mass spectra for the *m/z* 984 (A), 1,032 (B) and 1,048 (C) ions from the *Dolichospermum lemmermannii* CAWBG680 sample, including putative structures of the anabaenopeptilides determined using the masses and pattern of the fragment ions.

Mass differences between the parent ions (*m/z* 984 vs *m/z* 1,032 = +48 Da; *m/z* 1,032 vs *m/z* 1,048 = +16 Da) indicated that the variable amino acids between the three anabaenopeptilides were likely to be valine (Val; 117 Da), phenylalanine (Phe; 165 Da) and tyrosine (Tyr; 181 Da). The residual mass was that of an asparagine (Asn) joined in a peptide chain.

The high frequency of methylated aromatic amino acids in Position 5 (Welker and von Döhren 2006) suggested that the *N*Me-*O*Me-Tyr residue was located here. The previous observation glutamine in Position 2 and valine in Position 6 were used to make the tentative placement of these two amino acids, with the Phe and Tyr also presumed to be in Position 6.

Microcystin precursor ion screening mostly detected the compounds already identified in Table 2. The only additional potential microcystin signal was a *m/z* 533 ion observed in the *Dolichospermum lemmermannii* CAWBG680 sample at a retention time of 1.07 min. The approximate *m/z* and retention time suggested that it could be a doubly charged ion of a MC-RR congener. MS/MS analysis of the *m/z* 533 ion (Figure 4) indicated that this was not the case, as no fragment ions with *m/z* > 533 were observed (which would have been the case for a doubly charged ion of a compound with a mass around 1,060 Da). This suggested that the ion was a singly charged ion of a compound with a mass around 530 Da, which is much lower than the mass of a microcystin.

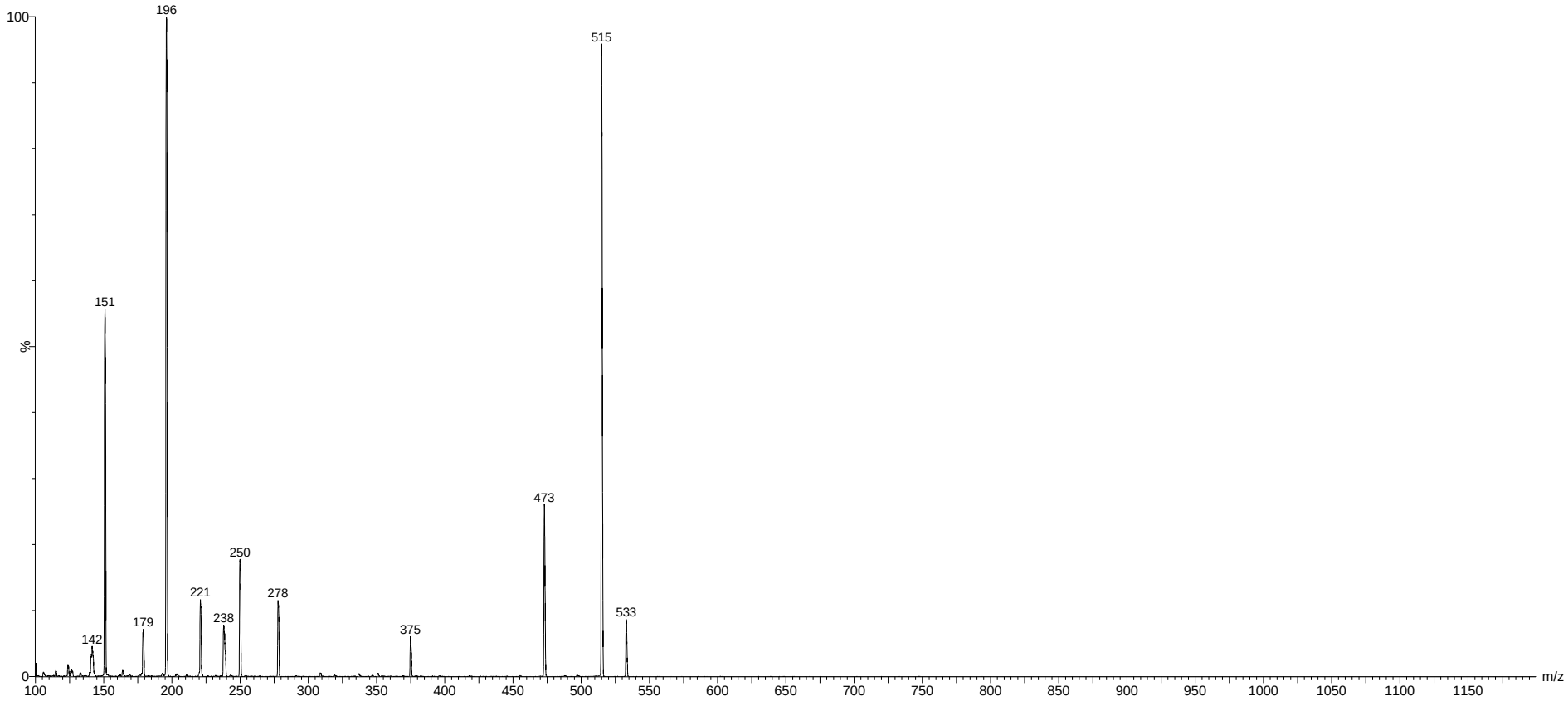

Figure 4. Tandem mass spectrum for the *m/z* 533 ion from the *Dolichospermum lemmermannii* CAWBG680 sample.

3.3 Discussion

LC–MS/MS investigations on samples from Lake Opuha and *Dolichospermum lemmermannii* CAWBG680 did not indicate the production of microcystins. However, multiple compounds present in the two samples caused interferences with the microcystin MRM analysis. The interfering compounds were identified as anabaenopeptilides, which are cyclic peptides produced by an array of cyanobacteria, including *Dolichospermum* (Welker and von Döhren 2006). Due to the widespread production of anabaenopeptilides amongst cyanobacteria, it is likely that they will be observed again in environmental samples from Aotearoa New Zealand and potentially cause interferences. In the current study, the interfering signals from anabaenopeptilides were all at different retention times than those expected for the respective microcystin being analysed. The incorporation of 'qualitative' MRM channels to LC–MS/MS methods for microcystin analysis would add further assurance that the signals were due to interferences rather than novel microcystin congeners. Here, the ratio of the 'qualitative' MRM channel and the MRM channel used for microcystin quantitation can be diagnostic for the compound of interest. Comparison of the ratios (between the signals from the 'quantitative' and 'qualitative' MRM channels) for microcystin standards and the detected analytes can be used to determine if the signal was due to a microcystin or a non-target compound.

While anabaenopeptilides do not appear to have high toxicity to mammals, studies have shown cytotoxicity (towards cell lines; Weckesser et al. 1996), and toxicity towards *Caenorhabditis elegans* (nematodes; Lenz et al. 2019) and *Thamnocephalus platyurus* (freshwater crustacean; Gademann et al. 2010). Anabaenopeptilides can also be potent protease inhibitors (Weckesser et al. 1996; Gademann et al. 2010), which is likely to be linked to their observed toxicity.

A comprehensive assessment of the anabaenopeptilides present in CAWBG680 was not undertaken in the current study, and it is likely that other anabaenopeptilide analogues are present. While three of the anabaenopeptilide structures were tentatively assigned using MS/MS data, the positions of several amino acids were based on previous studies rather that structural evidence. More in-depth analysis of the LC–MS/MS data would be required to complete this work; however, this was beyond the scope of the current study. Ultimately, confirming the structure of the compounds would require purification of the compounds, amino acid analysis and nuclear magnetic resonance spectroscopy (or X-ray crystallography, if the compound was able to be crystalised). We did not investigate whether the putative anabaenopeptilide structures were new or had been previously reported. The wide diversity of naming systems for this cyanobacterial peptide makes this difficult to undertake without access to databases that allow for searches based on structure.

4. *Dolichospermum lemmermannii* microcystin and anabaenopeptilide synthase gene clusters

4.1 Introduction

Microcystins are biosynthesised by the *mcy* gene cluster, an approximately 50 kbp region of DNA containing non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) genes (Dittmann et al. 2013). As noted in Section 2.1, a functional *mcy* gene cluster includes the core genes (*mcyA*, *mcyB*, *mcyC*, *mcyD* and *mcyE*), genes encoding the tailoring enzymes, (*mcyG*, *mcyI* and *mcyJ*), a racemase gene (*mcyF*) and a putative transporter gene (*mcyH*).

The exact length, number and arrangement of genes can vary among cyanobacterial species and strains. The genome of *Anabaena* sp. 90 (NCBI accession number GCA000312705) includes a 56 kbp region containing the *mcyHIFEJDGABC* genes for microcystin synthesis (Wang et al. 2012). Given that *Anabaena* sp. 90 is a closely related strain to *Dolichospermum lemmermannii* and has a genome sequence available, its genome was used as a reference for this study.

Anabaenopeptilides are biosynthesised by the anabaenopeptilide synthase (*apd*) gene cluster, an approx. 29 kbp region of DNA containing NRPS and tailoring genes (coding for enzymes such as methyltransferases, a formyl-transferase and a halogenase; Rouhiainen et al. 2000; Welker and von Döhren 2006). Although not part of the original scope for this study, the identification of anabaenopeptilides in *Dolichospermum lemmermannii* CAWBG680 led us to investigate the presence of *apd* genes using the genome data.

In this study, genome sequencing was undertaken on *Dolichospermum lemmermannii* CAWBG680 to determine which genes of the *mcy* gene cluster were present in the strain, and whether any *mcy* genes were missing and could render the strain incapable of producing microcystins.

4.2 Methods and results

A dense *Dolichospermum lemmermannii* CAWBG680 culture (200 mL), grown as described in Section 3, was centrifuged (3,000 × *g*, 10 min) in two batches. The two resulting pellets were extracted using the DNeasy PowerSoil™ DNA Isolation Kit (QIAGEN) and the automated QIAcube (QIAGEN), following the manufacturer's protocol. The DNA was combined and shipped to Azenta Life Science (China) for sequencing using Illumina PE150 and PacBio sequencing.

A total of 21,510,324 reads (length 151 bp each) were obtained from the Illumina sequencing, while 78,171 reads were obtained from the PacBio sequencing, with an average length of 5,832 bp (min 485; max 31,638).

The raw reads from the sequencing were mapped to the reference genome and the microcystin cluster of *Anabaena* sp. 90 obtained from the National Centre for Biotechnology Information (NCBI; CP003284.1 and AY212249.1, respectively) using minimap2 (Li 2018). The presence of anabaenopeptilide synthase genes was also assessed in the same manner using sequences from *Anabaena* sp. 90.

Raw sequences were assembled using the hybrid metagenome assembler OPERA-MS (Bertrand et al. 2019), which assembles the short Illumina reads into contigs using Spades (Bankevich et al. 2012) with default parameters. The PacBio and Illumina reads were then mapped onto the contigs to identify connectivity between the contigs. This information was then used to cluster the contigs into genomes. Coding sequences on the contigs were identified using Prodigal (Hyatt et al. 2010) and annotated against the Kyoto Encyclopaedia of Genes and Genomes (KEGG; Kanehisa and Goto 2000) database using Blastx in DIAMOND (Buchfink et al. 2021). KEGG annotations were subsequently checked against the NCBI nr database.

Because the *Dolichospermum lemmermannii* culture used for sequencing was not axenic, taxonomic analysis revealed sequences from other bacterial species, including *Caulobacter mirabilis*, *Silanimonas lenta* and *Lacibacter* sp.

Mapping of the Illumina reads against the *Anabaena* sp. 90 genome indicated a coverage of 73% at an average mapping depth (number of times raw reads map against the position) of 47.6 (Figure 5). Similarly, mapping of the PacBio reads covered 76% of base pairs from the genome with an average sequencing read depth of 4.3. Against the microcystin cluster, the Illumina raw reads mapped against 61% of the *Anabaena* sp. 90 cluster with an average mapping depth of 51.8. The PacBio reads covered 75% of the *Anabaena* sp. 90 microcystin cluster with an average sequencing read depth of 3.5 (Figure 5).

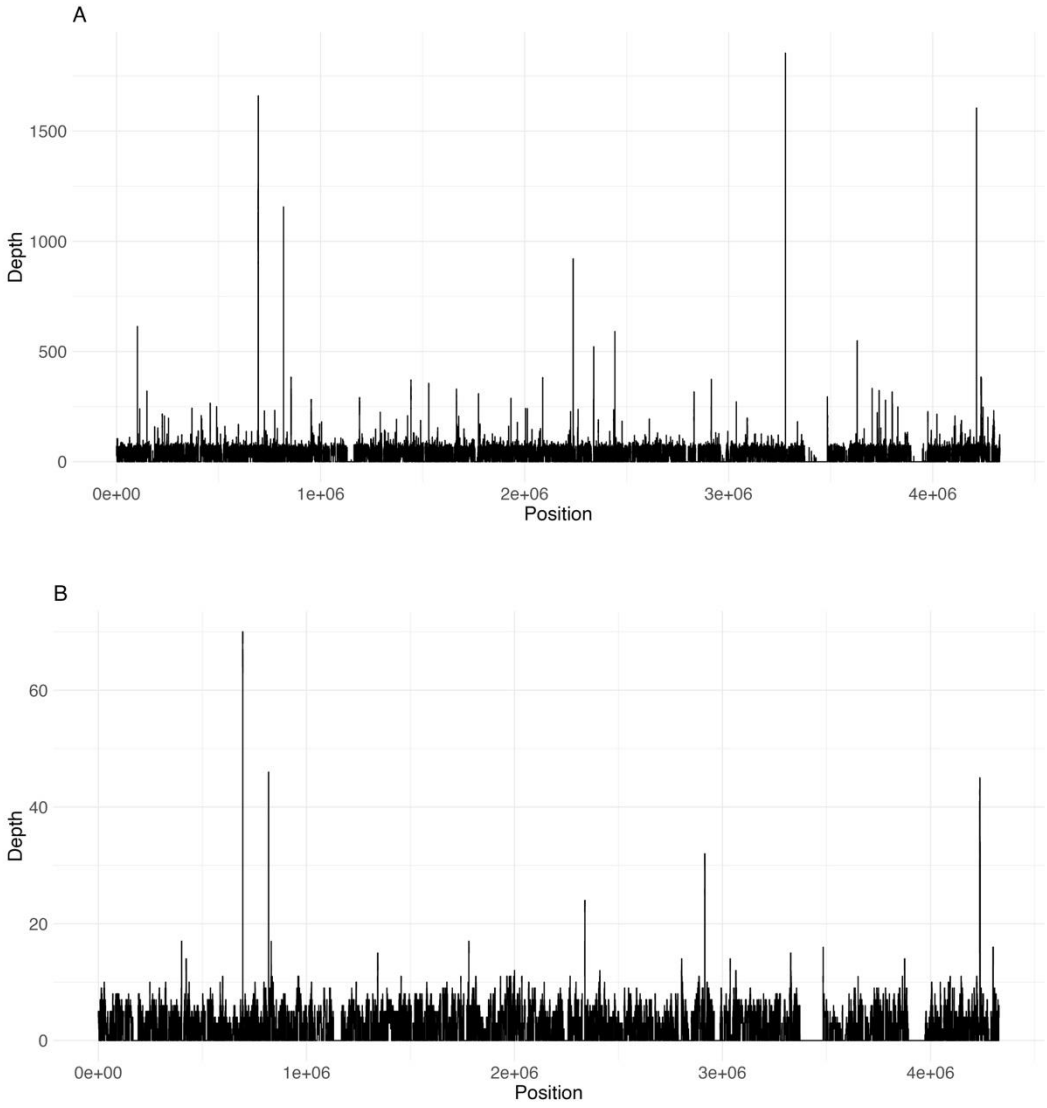

Figure 5. Mapping of the Illumina (A) and PacBio (B) sequencing reads from *Dolichospermum lemmermannii* CAWBG680 against the *Anabaena* sp. 90 genome. The depth (y-axis) shows the number of sequences (depth) at each position on the genome. Note the different y-axis scales between panels A and B.

Using the KEGG annotations, sequences matching the microcystin genes *mcyA–G* were found (Figure 6). However, the taxonomic classification of *mcyF* showed that it was not of cyanobacterial origin, which was in agreement with the raw read mapping against the *Anabaena* microcystin cluster. The annotated *mcyF* sequences obtained matched DNA sequences from *Silanimonas* sp., *Caulobacter* sp., *Porphyrobacter* sp., *Elioraea* sp. and *Rubrivivax* sp. No matches were observed for the *mcyH* and *mcyI* genes. The order of the genes in the cluster could not be established, as the observed genes were not obtained on a single contig, with the genes spread over three contigs in this dataset. One contig of the assembly contained a partial *mcyG*, *mcyA* and *mcyB* while another contig contained a partial *mcyE*, *mcyJ* and *mcyD*. The gene *mcyC* was observed on a contig of its own.

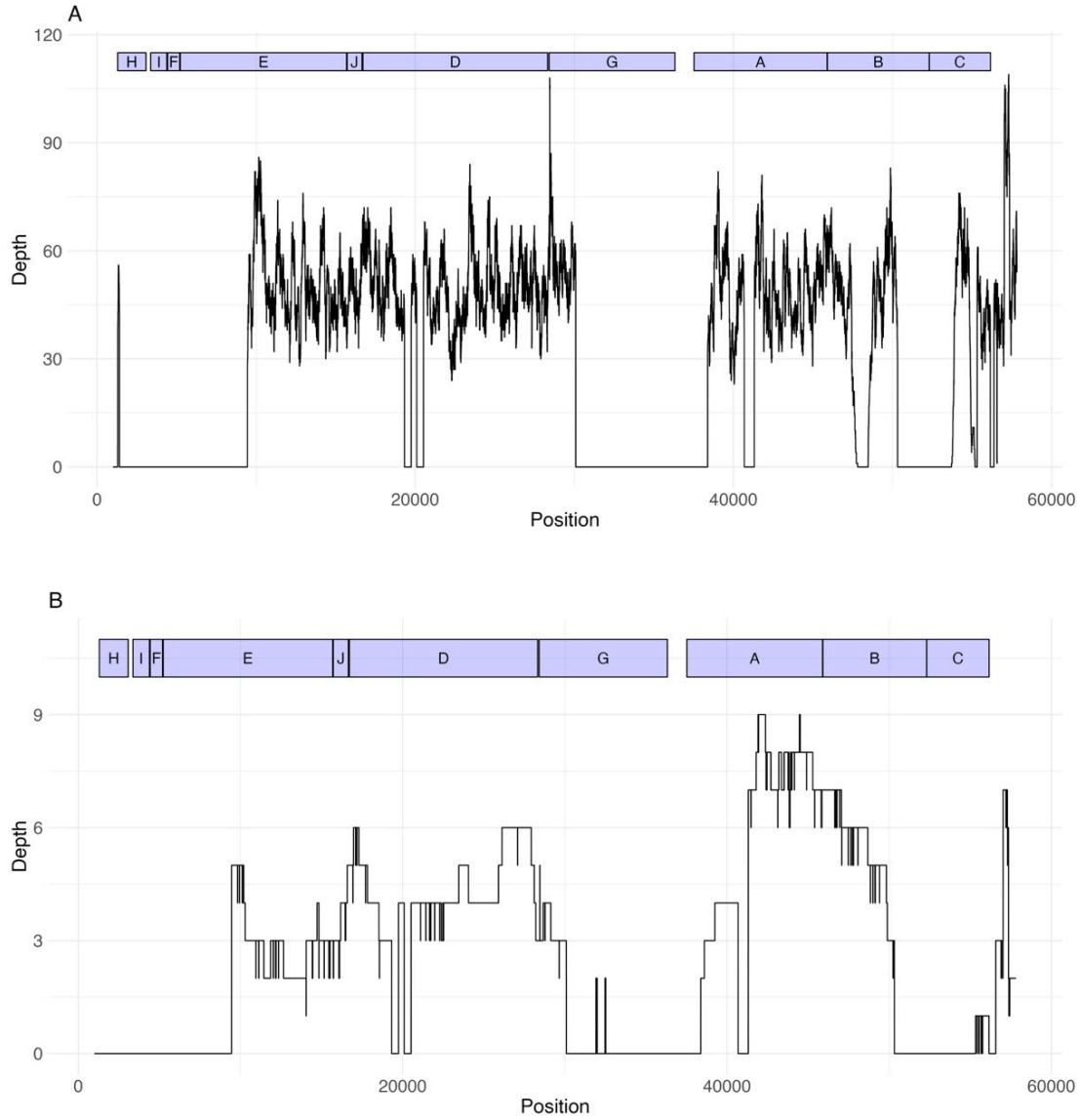

Figure 6. Mapping of the Illumina (A) and PacBio (B) sequencing reads from *Dolichospermum lemmermannii* CAWBG680 against the microcystin cluster from the *Anabaena* sp. 90 genome. The depth (y-axis) shows the number of sequences (depth) at each position on the cluster. Note the different y-axis scales between panels A and B. Boxes highlight the position of the *mcy* genes (denoted by the letters) on the *Anabaena* sp. 90 cluster.

The identification of anabaenopeptilides in *Dolichospermum lemmermannii* CAWBG680 (Section 3.2) led us to investigate whether anabaenopeptilide synthase (*apd*) genes were present in the cyanobacterial strain. Mapping against the *apd* gene cluster of *Anabaena* sp. 90 indicated almost complete presence of *apdB*, *apdD*, *apdE* and *apdF*, and a portion of *apdA* (Figure 7). The *apdC* gene was absent in *Dolichospermum lemmermannii* CAWBG680.

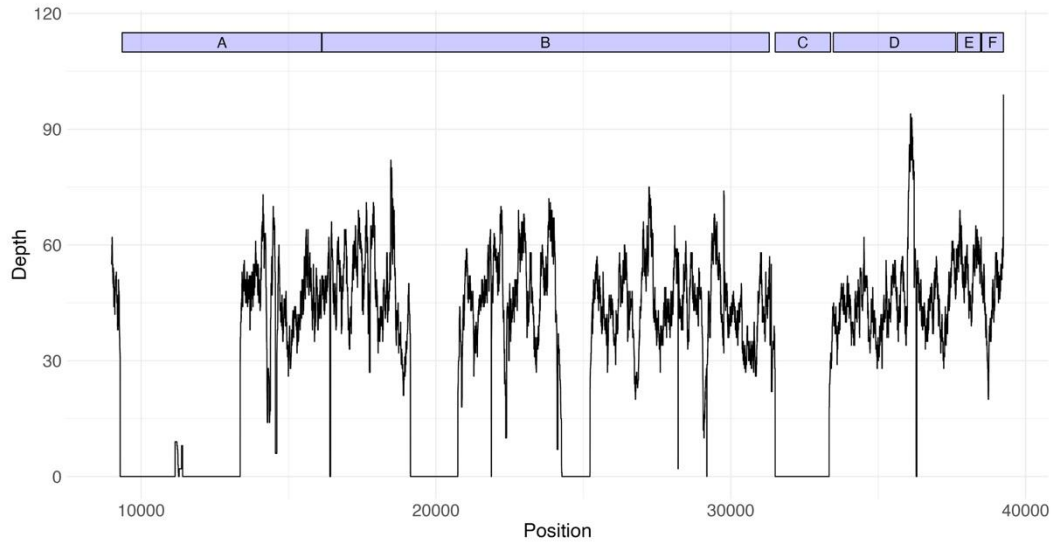

Figure 7. Mapping of the Illumina sequencing reads from *Dolichospermum lemmermannii* CAWBG680 against the anabaenopeptilide synthase (*apd*) cluster from the *Anabaena* sp. 90 genome. The depth (y-axis) shows the number of sequences (depth) at each position on the cluster. Boxes depict the position of the different *apd* genes on the cluster.

4.3 Discussion

Mapping against the *Anabaena* sp. 90 genome indicated that *Dolichospermum lemmermannii* shared a large proportion of the genome. However, approximately 25% of the genome from CAWBG680 could not be mapped to *Anabaena* sp. 90. While this could be due to insufficient sequencing effort, it could also indicate that the *Dolichospermum lemmermannii* CAWBG680 genome is smaller and that there has been a series of gene deletions.

Genomic analysis of the *mcy* gene cluster indicated that the *mcyF*, *mcyH* and *mcyI* genes were absent in *Dolichospermum lemmermannii* CAWBG680. There is a small possibility that this is because we did not have enough sequencing depth of the genome; however, the fact that both the Illumina and PacBio sequencing efforts had the same areas of the gene cluster present and missing suggests that the genes are absent or partly absent. The genes *mcyF* and *mcyI* likely code for enzymes involved with the biosynthesis of the D-methylaspartic acid found in Position 3 of microcystins (Rouhiainen et al. 2004; Pearson et al. 2007). The effect of the loss of the *mcyF* and *mcyI* genes is not well studied, but microcystin-producing *Planktothrix agardhii* CYA126/8 does not contain these two genes – suggesting that microcystin production is possible in their absence (Christiansen et al. 2003; Rouhiainen et al. 2004). While the function of the *mcyH* gene has not been firmly established, its deletion has resulted in the knockout of microcystin production in *Microcystis aeruginosa* (Pearson et al. 2004). The absence of the *mcyH* gene in *Dolichospermum lemmermannii* CAWBG680 suggests that the strain is unlikely to be able to produce microcystin, unless a homologue of the gene was fulfilling its role.

Recently, an additional enzyme thought to be involved in microcystin production has been characterised, *mcyK* (Ouyang et al. 2024). This enzyme was identified in *Fischerella* sp. (a cyanobacteria species from the same order as *Dolichospermum*) and is hypothesised to be involved in the production of microcystin variants containing homoarginine. Because this gene was not annotated in the *Anabaena* sp. 90 genome, even if it was present in *Dolichospermum lemmermannii* CAWBG680, we would not have detected it in our analysis. Given that multiple other microcystin production genes are absent in *Dolichospermum lemmermannii* CAWBG680, and because there are no annotated sequences of this gene in NCBI, we did not pursue investigating it presence further.

We could not resolve the microcystin cluster in a single contig through the current work and were thus unable to completely determine the order of the microcystin genes in the cluster. Due to the presence of other organisms in the culture, a proportion of the sequencing effort was taken up sequencing these other organisms and not *Dolichospermum lemmermannii*. With lower levels of sequencing, assembling contigs across the whole of the *mcy* gene cluster was not possible with the current genome data. Further genome sequencing, especially using PacBio long-read sequencing, would enable the order of the *mcy* genes to be determined in the future.

Because of the detection of anabaenopeptilides in *Dolichospermum lemmermannii* CAWBG680, the *apd* gene cluster was also investigated using the genome data. This indicated that there was good coverage of the *apd* gene cluster, but a portion of *apdA* was missing and *apdC* was absent. The absence of *apdC* did not impact anabaenopeptilide production in *Dolichospermum lemmermannii* CAWBG680, as it codes for a tailoring enzyme involved in the chlorination of tyrosine residues (Rouhiainen et al. 2000) – a structural moiety that was not observed in the three anabaenopeptilide structures putatively characterised during this study (see Figure 3 in Section 3.2). At this stage, we do not know if the missing portion of *apdA* was due to a lack of sequencing data or an actual reduction in the size of the *apd* gene in *Dolichospermum lemmermannii* CAWBG680. As with the *mcy* gene cluster, further sequencing data would be required to address this.

5. Conclusions and recommendations

The three research components of this project provided new insights into potential microcystin production in *Dolichospermum lemmermannii* from Aotearoa New Zealand. Twelve *Dolichospermum lemmermannii* strains from different parts of the country (Hawkes Bay, Manawatū, Otago, Taranaki, Waikato and Wellington) were assessed for the presence of a ubiquitous portion of the *mcyE* gene and it was detected in only three strains (Section 2). This demonstrates that not all *Dolichospermum lemmermannii* strains in Aotearoa New Zealand contain the *mcyE* gene indicative of microcystin production potential. LC–MS/MS investigations did not detect microcystins in one of the *mcyE*-positive cultures (CAWBG680; Section 3), and genome analysis demonstrated that portions of the *mcy* gene cluster were absent in *Dolichospermum lemmermannii* CAWBG680 (Section 4). In combination, these two results (described in Sections 3 and 4) indicate that even though a portion of the *mcyE* gene might be detected in an environmental sample or cyanobacterial culture, it does not confirm that microcystins will be present.

In contrast to the observations made in the current study, Capelli et al. (2017) investigated microcystin production in *Dolichospermum lemmermannii* strains from Europe. The researchers tested 12 strains from different parts of Europe for the *mcyE* gene and microcystin production using a microcystin-specific ELISA (enzyme-linked immunosorbent assay). Five of the cultures contained the *mcyE* gene and were also positive for microcystins by ELISA. This study, and other studies reporting microcystin production in *Dolichospermum* spp. from overseas (e.g. Harada et al. 1991; Kobos et al. 2013; Li et al. 2016; Dreher et al. 2019), demonstrate why continued vigilance is required regarding the management of *Dolichospermum* in Aotearoa New Zealand. While this study clarified that microcystin production was not present in a sample from Lake Opuha collected in 2023 and a culture of *Dolichospermum lemmermannii* (CAWBG680), continued testing of environmental samples where *Dolichospermum* is abundant should be undertaken. Because *Dolichospermum* spp. have also been reported to produce anatoxins, cylindrospermopsins and saxitoxins (other prominent cyanotoxin classes), this testing could be carried out using the cyanotoxin gene screen, which tests for *mcyE* / *ndaF* as well genes responsible for anatoxin, cylindrospermopsin and saxitoxin production. Any positive results could then be followed up by chemical analyses, such as LC–MS/MS. The revised cyanobacteria guidelines for recreational fresh waters encourages water managers to follow this approach.

While enough genome sequencing data was acquired to demonstrate that several genes were absent from the *mcy* cluster of *Dolichospermum lemmermannii* CAWBG680, we did not acquire sufficient sequencing data to fully understand the arrangement of the various *mcy* genes in this strain. This limited our ability to compare the *mcy* cluster of *Dolichospermum lemmermannii* CAWBG680 with the *mcy* clusters of other microcystin producers. Further genome sequencing could allow for this in the future and possibly provide further insights into how the *mcyF*, *mcyH* and *mcyI* genes were lost from the *Dolichospermum lemmermannii* CAWBG680 *mcy* cluster. The *mcyE* gene was detected in two other *Dolichospermum lemmermannii* strains (CAWBG567 and CAWBG568), but the cyanobacterial biomass grown was insufficient for genome sequencing within the time frame of the project. Future work undertaking genome sequencing on these strains, and potentially *mcyE*-negative strains, would allow for better understanding on the inter-strain diversity of *Dolichospermum lemmermannii* in Aotearoa New Zealand.

The combined evidence from the LC–MS/MS analysis and *mcy* gene cluster sequencing indicates that *Dolichospermum lemmermannii* CAWBG680 cannot produce microcystins. The identification of anabaenopeptilides as the cause for the tentative microcystin identification in the original Lake Opuha sample also alleviates initial concerns. These results support the adoption of the revised alert-level framework for planktonic cyanobacteria in the updated recreational cyanobacteria guidelines, where *Dolichospermum* spp. is not included as a toxin producer. There is still a possibility that there are other *Dolichospermum* strains / species in Aotearoa New Zealand that can produce microcystins or other cyanotoxins, or that other strains / species will be introduced, and we recommend continued testing of environmental samples, using PCR and LC–MS/MS, when *Dolichospermum* is dominant. These results also highlight that while PCR approaches are very valuable for detecting potential cyanotoxin producers, positive results need to be confirmed using chemical analyses.

6. Acknowledgements

We thank Laura Kelly (Cawthron) and and Kirsty Smith (Cawthron) for their reviews of this report, and Louisa Fisher for scientific editing. We acknowledge Graeme Clarke (MfE) for his support throughout this project.

7. References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology. 19(5):455–477.

Bertrand D, Shaw J, Kalathiyappan M, Ng AHQ, Kumar MS, Li C, Dvornicic M, Soldo JP, Koh JY, Tong C. 2019. Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. Nature Biotechnology 37(8):937-944.

Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nature Methods. 18(4):366–368.

Capelli C, Ballot A, Cerasino L, Papini A, Salmaso N. 2017. Biogeography of bloom-forming microcystin producing and non-toxigenic populations of *Dolichospermum lemmermannii* (Cyanobacteria). Harmful Algae. 67:1–12.

Christiansen G, Fastner J, Erhard M, Börner T, Dittmann E. 2003. Microcystin biosynthesis in *Planktothrix*: Genes, evolution, and manipulation. Journal of Bacteriology. 185(2):564–572.

Christiansen G, Molitor C, Philmus B, Kurmayer R. 2008. Nontoxic strains of cyanobacteria are the result of major gene deletion events induced by a transposable element. Molecular Biology and Evolution. 25(8):1695–1704.

Dittmann E, Fewer DP, Neilan BA. 2013. Cyanobacterial toxins: biosynthetic routes and evolutionary roots. FEMS Microbiology Reviews. 37(1):23–43.

Dreher TW, Collart LP, Mueller RS, Halsey KH, Bildfell RJ, Schreder P, Sobhakumari A, Ferry R. 2019. *Anabaena/Dolichospermum* as the source of lethal microcystin levels responsible for a large cattle toxicosis event. Toxicon: X. 1:100003.

Gademann K, Portmann C, Blom JF, Zeder M, Jüttner F. 2010. Multiple toxin production in the cyanobacterium *Microcystis*: isolation of the toxic protease inhibitor cyanopeptolin 1020. Journal of Natural Products. 73(5):980–984.

Harada K-I, Ogawa K, Kimura Y, Murata H, Suzuki M, Thorn PM, Evans WR, Carmichael WW. 1991. Microcystins from *Anabaena flos-aquae* NRC 525-17. Chemical Research in Toxicology. 4(5):535–540.

Harke M, Steffen M, Gobler C, Otten T, Wilhelm S, Wood SA, Pearl H. 2016. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis*. Harmful Algae. 54:4–20.

Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 11:1–11.

Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research. 28(1):27–30.

Kleinteich J, Puddick J, Wood SA, Hildebrand F, Laughinghouse IV HD, Pearce D, Dietrich D, Wilmotte A. 2018. Toxic cyanobacteria in Svalbard: Chemical diversity of microcystins detected using a liquid chromatography mass spectrometry precursor ion screening method. Toxins. 10(4):147.

Kobos J, Błaszczyk A, Hohlfeld N, Toruńska-Sitarz A, Krakowiak A, Hebel A, Sutryk K, Grabowska M, Toporowska M, Kokociński M. 2013. Cyanobacteria and cyanotoxins in Polish freshwater bodies. Oceanological and Hydrobiological Studies. 42:358–378.

Lenz KA, Miller TR, Ma H. 2019. Anabaenopeptins and cyanopeptolins induce systemic toxicity effects in a model organism the nematode *Caenorhabditis elegans*. Chemosphere. 214:60–69.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34(18):3094–3100.

Li X, Dreher TW, Li R. 2016. An overview of diversity, occurrence, genetics and toxin production of bloom-forming *Dolichospermum* (*Anabaena*) species. Harmful Algae. 54:54–68.

[MfE and MoH] Ministry for the Environment, Ministry of Health. 2009. New Zealand guidelines for cyanobacteria in recreational fresh waters: interim guidelines. Wellington: Ministry for the Environment.

Ouyang X, Wahlsten M, Pollari M, Delbaje E, Jokela J, Fewer DP. 2024. Identification of a homoarginine biosynthetic gene from a microcystin biosynthetic pathway in *Fischerella* sp. PCC 9339. Toxicon. 243:107733.

Pearson LA, Barrow KD, Neilan BA. 2007. Characterization of the 2-hydroxy-acid dehydrogenase McyI, encoded within the microcystin biosynthesis gene cluster of *Microcystis aeruginosa* PCC7806. Journal of Biological Chemistry. 282(7):4681–4692.

Pearson LA, Hisbergues M, Borner T, Dittmann E, Neilan BA. 2004. Inactivation of an ABC transporter gene, *mcyH*, results in loss of microcystin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Applied and Environmental Microbiology. 70(11):6370–6378.

Puddick J, Prinsep MR, Wood SA, Cary SC, Hamilton DP, Wilkins AL. 2013. Isolation and structure determination of two new hydrophobic microcystins from *Microcystis* sp. (CAWBG11). Phytochemistry Letters. 6:575–581.

Puddick J, Prinsep MR, Wood SA, Kaufononga SAF, Cary SC, Hamilton DP. 2014. High levels of structural diversity observed in microcystins from *Microcystis* CAWBG11 and characterization of six new microcystin congeners. Marine Drugs. 12:5372–5395.

Puddick J, Thomson-Laing G, Wood SA. 2019. Microcystins in New Zealand: a review of occurrence, congener diversity and cell quotas. New Zealand Journal of Botany. 57(2):93–111.

Puddick J, Wood SA, Hawes I, Hamilton DP. 2016. Fine-scale cryogenic sampling of planktonic microbial communities: Application to toxic cyanobacterial blooms. Limnology and Oceanography: Methods. 14:600–609.

Rantala A, Rajaniemi-Wacklin P, Lyra C, Lepistö L, Rintala J, Mankiewicz-Boczek J, Sivonen K. 2006. Detection of microcystin-producing cyanobacteria in Finnish lakes with genus-specific microcystin synthetase gene E (*mcyE*) PCR and asssociations with environmental factors. Applied and Environmental Microbiology. 72(9):6101–6110.

Rouhiainen L, Paulin L, Suomalainen S, Hyytiäinen H, Buikema W, Haselkorn R, Sivonen K. 2000. Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90. Molecular Microbiology. 37(1):156–167.

Rouhiainen L, Vakkilainen T, Siemer BL, Buikema W, Haselkorn R, Sivonen K. 2004. Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium *Anabaena* strain 90. Applied and Environmental Microbiology. 70(2):686–692.

Sanz M, Andreote APD, Fiore MF, Dörr FA, Pinto E. 2015. Structural characterization of new peptide variants produced by cyanobacteria from the Brazilian Atlantic coastal forest using liquid chromatography coupled to quadrupole time-of-flight tandem mass spectrometry. Marine Drugs. 13(6):3892–3919.

Sielaff H, Dittmann E, De Marsac NT, Bouchier C, von Döhren H, Börner T, Schwecke T. 2003. The *mcyF* gene of the microcystin biosynthetic gene cluster from *Microcystis aeruginosa* encodes an aspartate racemase. Biochemical Journal. 373:903–916.

Smucker NJ, Beaulieu JJ, Nietch CT, Young JL. 2021. Increasingly severe cyanobacterial blooms and deep water hypoxia coincide with warming water temperatures in reservoirs. Global Change Biology. 27(11):2507–2519.

Spoof L, Catherine A. 2017. Appendix 3: Tables of microcystins and nodularins. In: Meriluoto J, Spoof L, Codd J, editors. Handbook of cyanobacterial monitoring and cyanotoxin analysis. Hoboken (NJ): Wiley. p. 526–538.

Tillett D, Dittmann E, Erhard M, von Döhren H, Börner T, Neilan BA. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: An integrated peptide-polyketide synthetase system. Chemistry & Biology. 7(10):753–764.

Vaitomaa J, Rantala A, Halinen K, Rouhiainen L, Tallberg P, Mokelke L, 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(12):7289–7297.

Wang H, Sivonen K, Rouhiainen L, Fewer DP, Lyra C, Rantala-Ylinen A, Vestola J, Jokela J, Rantasärkkä K, Li Z. 2012. Genome-derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90. BMC Genomics. 13:1–17.

Weckesser J, Martin C, Jakobi C. 1996. Cyanopeptolins, depsipeptides from cyanobacteria. Systematic and Applied Microbiology. 19(2):133–138.

Welker M, von Döhren H. 2006. Cyanobacterial peptides: nature's own combinatorial biosynthesis. FEMS Microbiology Reviews. 30(4):530–563.

Wood SA, Holland PT, Stirling DJ, Briggs LR, Sprosen J, Ruck JG, Wear RG. 2006. Survey of cyanotoxins in New Zealand water bodies between 2001 and 2004. New Zealand Journal of Marine and Freshwater Research. 40(4):585–597.

Wood SA, Maier MY, Puddick J, Pochon X, Zaiko A, Dietrich DR, Hamilton DP. 2017. Trophic state and geographic gradients influence planktonic cyanobacterial diversity and distribution in New Zealand lakes. FEMS Microbiology Ecology. 93(2):fiw234.

Wood SA, Puddick J. 2023. Cyanotoxin gene analysis in selected Canterbury lakes. Nelson: Cawthron Institute. Cawthron Advice Letter 2355. Prepared for Prepared for the National Public Health Service of Te Whatu Ora (Waitaha Canterbury District).

Yancey CE, Smith DJ, Den Uyl PA, Mohamed OG, Yu F, Ruberg SA, Chaffin JD, Goodwin KD, Tripathi A, Sherman DH. 2022. Metagenomic and metatranscriptomic insights into population diversity of *Microcystis* blooms: Spatial and temporal dynamics of *mcy* genotypes, including a partial operon that can be abundant and expressed. Applied and Environmental Microbiology. 88(9):e02464–02421.

World-class science for a better future.

CAWTHRON

98 Halifax Street East, Nelson 7010 | Private Bag 2, Nelson 7042 | New Zealand | Ph. +64 3 548 2319

cawthron.org.nz