Tetrahedron Letters 59 (2018) 934–937

Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

Identification of the new chymotrypsin inhibitor micropeptin 996 by metabolomics-guided analysis

Wendy K. Strangman [a], Allison K. Stewart [a], Megan C. Herring [b], Jeffrey L.C. Wright [a,b,*]

[a] MARBIONC, University of North Carolina Wilmington, 5600 Marvin Moss Lane Wilmington, NC 28409, USA
[b] Department of Chemistry and Biochemistry, University of North Carolina Wilmington, 601 S. College Road Wilmington, NC 28403, USA

ARTICLE INFO

Article history:
Received 15 December 2017
Revised 24 January 2018
Accepted 28 January 2018
Available online 31 January 2018

Keywords:
Cyanobacteria
Micropeptin
Metabolomics
Chymotrypsin inhibitor

ABSTRACT

An untargeted metabolomics approach was used to investigate a cultured strain of *Microcystis aeruginosa* (UTEX LB2386) known to be a prolific producer of a diverse class of cyanopeptides. Identification of a putative new compound with a molecular weight of 996 led to the purification and structure elucidation of this new member of the micropeptin class of cyanopeptides. Micropeptin 996 displayed potent inhibition of the serine protease enzyme chymotrypisin relative to structurally related members of this class.

Introduction

Cyanobacteria are extremely prolific producers of biologically active small molecules. Some freshwater genera such as *Microcystis* are of enhanced interest due to their propensity for forming harmful algal blooms (HABs) with high concentrations of toxins. In addition to their acute toxins such as microcystins, cylindrospermopsins, and anatoxin, many species also produce a wide variety of other peptide natural products that act as potent and selective inhibitors of proteases.[1] In the environment, this inhibitory activity is thought to protect the cyanobacteria by interfering with the digestive enzymes of crustacean grazers.[2] These compounds are also of great interest in the search for new small molecule therapeutics and for chemical probes of biological systems as proteases are known to play key roles in infections as well as in cardiovascular diseases and cancer.[3–6]

Micropeptins are members of the cyanopeptolin family of compounds and are produced by species of *Microcystis*.[7] This family is characterized by a 6-amino acid-membered cyclic core cyclized through a depsi-linkage between the side chain hydroxyl of the *N*-terminus amino acid (almost always threonine) with the carboxy terminus of the adjacent residue. They are further defined by the presence of the unique 3-amino-6-hydroxy piperadone (Ahp) moiety and a side chain containing 1–4 amino, hydroxyl, or fatty acids connected to the ring through the amide nitrogen of the lactone forming threonine. Structurally similar compounds have been described from other freshwater cyanobacterial genera including oscillapeptins (*Oscillatoria*), planktopeptins (*Planktothrix*), nostopeptins (*Nostoc*), and aeruginopeptins (*M. aeruginosa*).[7–13]Members have also been isolated from marine cyanobacteria (eg. Lyngbyastatins[14] and somamides[15] from *Lyngbya;* symplocamide from *Symploca*.[6] Partially due to their modular non-ribosomal peptide synthase (NRPS) biosynthetic origin, there is a high degree of variability within this structural class with over 150 congeners isolated to date.[16]

In our continuing search for novel chemistry and molecules with potential therapeutic value from cultured microorganisms, we employed the emerging technique of non-targeted metabolomic screening to detect new compounds present in our cyanobacterial culture collection. Metabolomics is a method for comprehensively assessing the small-molecule content of a complex biological system.[17] We applied a UPLC-QTOF MS metabolomics method towards laboratory cultures of the *M. aeruginosa* strain UTEX LB2386 (Supplemental). Originally isolated from Little Rideau Lake, Ontario, Canada, this strain is not reported as acutely toxic and does not produce microcystins. Previous large-scale extractions of this strain followed by traditional chromatographic separation, compound purification, and structure confirmation by NMR indicated that this strain is a prolific producer of several classes of bioactive cyanopeptides (Fig. 1). These include 3 new chlorinated microginins with the rare 3-amino-2-hydroxy- octanoic acid (Ahoa) side chain[18] as well

* Corresponding author.
E-mail address: wrightj@uncw.edu (J.L.C. Wright).

https://doi.org/10.1016/j.tetlet.2018.01.087
0040-4039/

as the previously described aeruginosamides B and C,[19] ferintoic acid A,[20] and several large molecular weight compounds assumed to be members of the microviridin class of cyanopeptides.[21] Given our knowledge of the metabolites produced by this culture using a traditional extraction and purification procedure, we employed a non-targeted metabolomics approach to identify any additional metabolites that remained undetected.

Results and discussion

To perform the metabolomics analysis, replicate cultures were grown, chemically extracted, and processed through SPE cartridges prior to UPLC-HRQTOF spectral acquisition (see Supplemental for experimental details). The chromatographic and mass spectral data were then imported into Progenesis QI software for spectral processing which resulted in a data set of 8079 ions. This set was then filtered to remove compounds with anova p-values > 0.05 and those with a highest mean value in the solvent blank, bringing the total down to 5528. These compound ions were then exported to the program EZinfo and subjected to orthogonal partial least square discriminant statistical analysis (OPLS-DA) to construct an S-plot comparing the cellular extracts to the media components. A loadings value of >0.07 was chosen and 20 compound ions with values greater than this were tagged and imported back into Progenesis QI. Manual annotation of these ions allowed us to identify ions falsely identified as separate compounds by the software but were typically high intensity fragment ions from the most abundant compounds produced by the organism. As others have noted, there is a quite large dynamic range in the production of natural products[16] and samples loaded on to the QToF at concentrations high enough to see the more minor compounds often result in false positives around the major metabolites. Detailed analysis of the untargeted metabolomics screen of cells vs the media blank rapidly highlighted the compounds previously described from this strain (Fig. 1). Additionally, a compound with an $[M+H]^+$ m/z of 997.5 was identified and yielded no matches in our search of published (Progenesis, Scifinder, MarinLit and Antibase) and internal databases of compounds and was targeted for purification and structure determination.

Structure elucidation

Micropeptin 996 was isolated as a clear glassy material. High resolution ToF mass spectrometry of the purified compound provided a molecular mass of $[M+H]^+$ m/z 997.5032 indicating a molecular formula of $C_{52}H_{68}N_8O_{12}$ for **1**. Characteristic signals in the 1D NMR spectrum including a singlet hydroxyl resonance at δ_H 6.06, a quartet ester methine shifted downfield at δ_H 5.43, a singlet *N*-methyl resonance (δ_H 2.78) and a methyl doublet at δ_H 1.21 suggested that this compound was a member of the cyanopeptolin family of cyanopeptides. As it was isolated from a strain of *Microcystis aeruginosa*, **1** is more precisely called a micropeptin. Further 2D NMR analysis including COSY, TOCSY, HSQC and HMBC allowed us to assign the component amino acids of this compound in sequence as valine, *N*-methyl phenyl alanine, phenylalanine, amino hydroxyl piperidone (Ahp), homotyrosine, threonine, glutamine, and butyric acid (Fig. 2 a and Table 1, please see Supplemental for full correlation data table and NMR spectra). Further 2D NMR analysis of the HMBC and ROESY spectra allowed the full planar assignment of the cyclic depsipeptide including the side chain extending from the threonine amine.

These assignments were further confirmed through analysis of the high-resolution MS-MS fragmentation data. Fragment ions corresponding to reported values for Ahp-Phe-MePhe- H_2O (m/z 404.1978), Ahp-Phe-H_2O (m/z 243.1127), and Phe-MePhe-H_2O (m/z 292.1136) confirm the connectivity of residues 2–4[13], (Supplemental). Additional ions at m/z 282.1446 and 199.1099 can be ascribed to Htyr-Thr-H_2O and BTA-Gln-H_2O respectively.

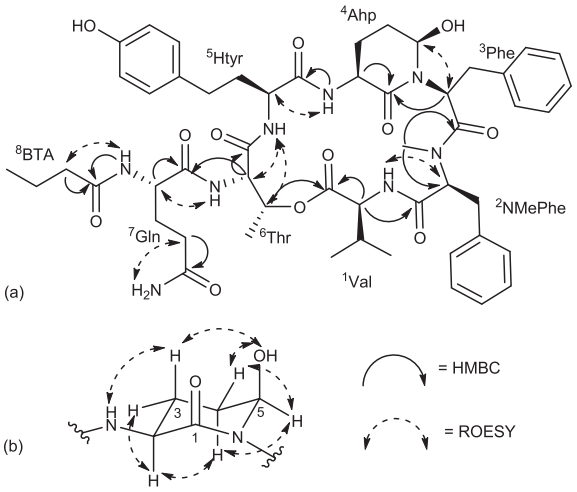

Fig. 2. Micropeptin 996 (**1**) structure with (a) key ROESY and HMBC correlations (b) ROESY correlations for Ahp stereochemistry. BTA refers to butyric acid.

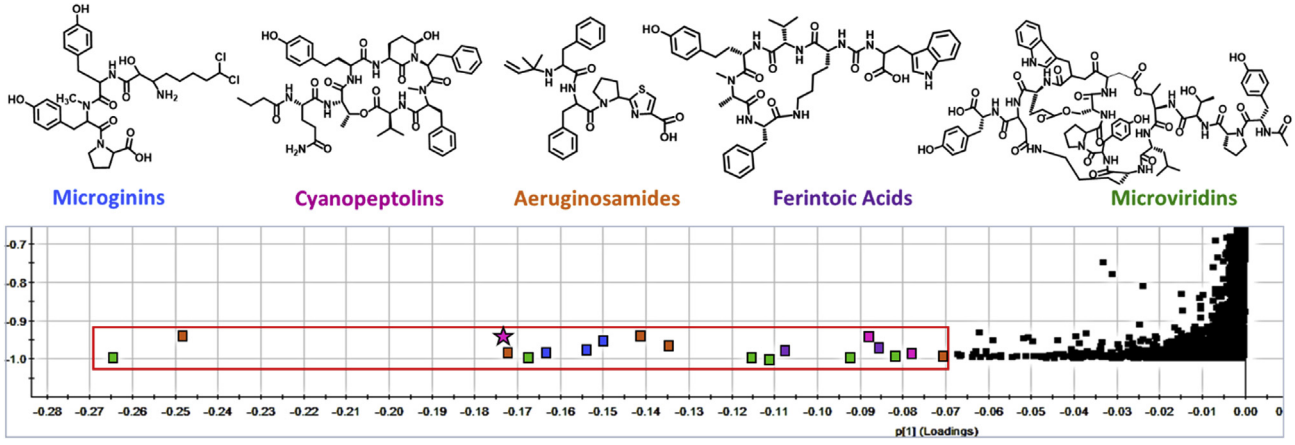

Fig. 1. S-Plot analysis expansion and representative classes of top 20 cyanopeptide mass features identified (red-box, color coded - micropeptin 996 is denoted by a pink star). Each square corresponds to a high-resolution mass feature enabling accurate database searching.