The cyanobacterium *Aetokthonos hydrillicola* – more than just an "eagle killer"?

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I

Biowissenschaften

der Martin-Luther-Universität

Halle-Wittenberg,

vorgelegt

von Markus Schwark

Tag der mündlichen Qualifikation: 23.10.2023

Dekan: Prof. Dr. Markus Pietzsch

1. Berichterstatter/in: Prof. Dr. Timo H. J. Niedermeyer

2. Berichterstatter/in: Prof. Dr. Raphael Reher

3. Berichterstatter/in: Prof. Dr. Wolfgang Sippl

Science doesn't always go forwards. It's like doing a Rubik's cube. You sometimes have to make more of a mess with a Rubik's cube before you can get it to go right.

Jocelyn Bell Burnel

Summary

Natural products from cyanobacteria are known for their structural diversity and biological activity. They can be lead molecules for the development of new drug substances, but many of them are far better known as potent toxins. The freshwater cyanobacterium *Aetokthonos hydrillicola* is of special interest in this regard, because it could be identified as the producer of a novel neurotoxin, called aetokthonotoxin (AETX). This toxin, which is transferred through the food chain, could be determined as the reason for the neurological disease "Vacuolar Myelinopathy" (VM). This disease affects bald eagles and other animals in the southeastern part of the United States, where the cyanobacterium occurs. In an initial screening, an extract of *A. hydrillicola* showed a higher cytotoxic activity, than the pure compound AETX. Thus, the main aim of this project was to find an explanation for this observation, and to identify other specialized metabolites produced by *A. hydrillicola.*

An extract of *A. hydrillicola* was microfractionated, and the generated fractions were tested for cytotoxic activity. Indeed, it was found that *A. hydrillicola* produces another cytotoxic compound. This cytotoxic compound was isolated and characterised as a pentapeptide belonging to the dolastatin family. Compounds from this family are well known for their pronounced cytotoxicity, and they were the first substances from cyanobacteria that made their way into the clinic as cytostatic agents. The compound isolated from *A. hydrillicola* differed from the known dolastatin derivatives in a so far unknown C-terminal moiety. Due to the structural relationship to the dolastatins and the name of the producer organism, the compound was named aetokthonostatin (AEST).

Using the tool GNPS (Global Natural Products Social Molecular Networking), AEST derivatives were identified in mass spectrometry data of the *A. hydrillicola* extract. These mainly differed in methylation pattern, but also in the number of amino acids. Three derivatives were isolated and characterized in more detail. All isolated compounds were tested for cytotoxicity and found to be highly active (IC$_{50}$ 1-25 nM). From the cytotoxicity data, first conclusions about structure-activity relationships (SAR) could be drawn. Further biological activity of AEST was tested using the model organism *Caenorhabditis elegans* (*C. elegans*). Surprisingly, in contrast to AETX, AEST showed no acute toxic effect, but decreased the reproduction rate of *C. elegans* (IC$_{50}$ 0.8 µM).

In addition to AEST derivatives, the GNPS analysis also revealed the presence of several compounds chemically related to the neurotoxin AETX. These molecules differed mainly

in the bromine substitution, but also both indole building blocks could be identified. One of these indoles was isolated and structurally characterized. A synthesis of this indole was achieved. In cooperation with partners at the Leibniz Institute of Plant Biochemistry, a total synthesis of AETX could be established.

The development of a PCR protocol, detecting the AETX biosynthetic gene cluster in environmental samples, by collaborators at the Institute of Hydrobiology in Budweis, Czech Republic, was supported by HPLC-MS analyses, to determine the presence/absence of AETX, and to match the analytical results with the results of the PCR assay. It could be shown that the PCR assay works reliable and can detect the AETX-producing cyanobacterium in environmental samples. We were also able to show that plants other than *Hydrilla verticillata* (in this case *Justicia americana*) can be colonised by *A. hydrillicola* and that AETX is also produced under these conditions.

Taken together, the cyanobacterium *A. hydrillicola* was found to produce two highly potent toxins. This is especially important as *A. hydrillicola* is the first known freshwater cyanobacterium producing dolastatin derivatives. In addition, a derivative of AETX could be isolated from the cyanobacterium, and a total synthesis of AETX could be achieved. Finally, the development of a reliable PCR protocol for the detection of AETX production will contribute to the protection of our ecosystems and human health.

Zusammenfassung

Naturstoffe aus Cyanobakterien sind für ihre strukturelle Vielfalt und Bioaktivität bekannt. Sie können als Vorbild für die Entwicklung von neuen Arzneistoffen dienen, aber gleichzeitig sind die Meisten besser als potente Toxine bekannt. Das Cyanobakterium *Aetokthonos hydrillicola* ist hierbei von besonderem Interesse, da es als Produzent eines neuen Neurotoxins, Aetokthonotoxin (AETX), identifiziert wurde. Dieses Toxin, welches über die Nahrungskette weitergegeben wird, wurde als die Ursache für die neurologische Krankheit „Vakuolare Myelinopathie" (VM) identifiziert. Diese Krankheit, die Weißkopfseeadler und andere Tiere beeinträchtigt, wurde an verschiedenen Seen im Südosten der USA, die ein Vorkommen des Cyanobakteriums aufwiesen, lokalisiert.

In einem initialen Screening zeigte ein Extrakt von *A. hydrillicola* eine höhere cytotoxische Aktivität als AETX als Reinsubstanz. Das Hauptziel dieses Projektes war es daher, eine Erklärung für diese Beobachtung zu finden und weitere von *A. hydrillicola* produzierte Verbindungen zu identifizieren.

Ein Extrakt wurde dafür mikrofraktioniert und die erhaltenen Fraktionen auf cytotoxische Aktivität getestet. Tatsächlich zeigte sich, dass eine noch deutlich stärker cytotoxisch aktive Verbindung von *A. hydrillicola* produziert wird. Isolierung und Strukturaufklärung zeigten, dass es sich bei der Verbindung um ein lineares Pentapeptid handelt, welches der Klasse der Dolastatine zuzuordnen ist. Substanzen aus dieser Verbindungsklasse sind für ihre starke cytotoxische Aktivität bekannt und haben es als erste cyanobakterielle Verbindungen in den klinischen Einsatz geschafft. Die isolierte Verbindung unterschied sich strukturell von den bereits bekannten Verbindungen durch eine neue Variation am C-Terminus. Aufgrund der Zugehörigkeit zur genannten Substanzklasse und dem Produktionsorganismus, *A. hydrillicola*, wurde die Verbindung Aetokthonostatin (AEST) genannt.

Mit Hilfe der Anwendung GNPS (Global Natural Products Social Molecular Networking) wurden weitere Derivate des AEST, in den massenspektrometrischen Daten eines Extrakts von *A. hydrillicola*, identifiziert. Diese unterschieden sich hauptsächlich im Methylierungsgrad, aber auch in der Anzahl der Aminosäuren. Drei Derivate wurden so isoliert und weitergehend charakterisiert.

Die isolierten Verbindungen wurden alle auf ihre cytotoxische Aktivität getestet und es konnte gezeigt werden, dass es sich dabei um hochaktive Verbindungen handelt (IC_{50} 1-25 nM). Des Weiteren konnten erste Rückschlüsse über die Struktur-Wirkungsbeziehungen der isolierten Verbindungen getroffen werden. Weitere Daten über die Bioaktivität der Verbindung AEST wurden mit Hilfe des Modellorganismus *Caenorhabditis elegans* gewonnen. Überraschenderweise zeigte AEST, im Gegensatz zu AETX, keine akute toxische Wirkung, verringerte aber die Reproduktionsrate von *C. elegans* (IC_{50} 0.8 µM).

Zusätzlich zu den AEST-Derivaten offenbarte die GNPS-Analyse auch das Vorhandensein

mehrerer Verbindungen, die chemisch mit dem Neurotoxin AETX verwandt sind. Diese unterschieden sich vor allem in der Anzahl der Bromatome, aber auch die beiden Indol-Bausteine des AETX konnten identifiziert werden. Einer der Indol-Bausteine wurde isoliert und strukturell aufgeklärt. Eine Synthese des isolierten Indol-Bausteins wurde etabliert. In Kooperation mit Partnern am Leibniz-Institut für Pflanzenbiochemie wurde die Totalsynthese des AETX realisiert.

Die Entwicklung eines PCR-Protokolls, zum Nachweis des Biosynthesegenclusters von AETX in Umweltproben, durch Kollaborationspartner des Institutes für Hydrobiologie in Budweis, Tschechische Republik, wurde durch HPLC-MS Analysen unterstützt, um das Vorhandensein/Fehlen von AETX zu bestimmen und um die gewonnenen Erkenntnisse mit denen der vorangegangen PCR zu vergleichen. Es konnte gezeigt werden, dass der entwickelte PCR-Assay zuverlässig funktioniert und AETX-produzierende Cyanobakterien in Umweltproben nachweisen kann. Des Weiteren konnte gezeigt werden, dass neben *Hydrilla verticillata* (in diesem Fall *Justicia americana*) auch andere Pflanzen von *A. hydrillicola* besiedelt werden können und dass unter diesen Bedingungen ebenfalls AETX produziert wird.

Zusammengenommen wurde festgestellt, dass das Cyanobakterium *A. hydrillicola* zwei hoch potente Toxine produziert. Dies ist von besonderer Bedeutung, da *A. hydrillicola* das erste bekannte Süßwassercyanobakterium ist, welches Dolastatin derivate produziert. Des Weiteren konnte ein Derivat des AETX aus dem Cyanobakterium isoliert werden und eine Totalsynthese des AETX erreicht werden. Abschließend kann die Entwicklung eines zuverlässigen PCR-Protokolls, für den Nachweis der AETX-Produktion, dazu beitragen unsere Ökosysteme und die menschliche Gesundheit zu schützen.

Contents

1. Introduction

1.1. Natural product research

The broadest definition of a natural product (NP) would describe it as any compound or material, which is produced by a living organism. In the field of NP research, the focus is primarily on novel secondary, bioactive compounds produced by bacteria, fungi, or plants. In the beginning, NPs were described as "substances appearing to have no explicit role in the internal economy of the organism that produces it".[1] This hypothesis was disproven over the years, as NPs play an important role for the respective producing organism (e.g. protection or communication).[2][3] These observations led to the new term "specialized metabolites" in the early 2000s.[4] The effects of NPs on the human organism were also recognised early on, as natural components already have been used to treat ailments for 60,000 years.[5] Nearly 65 % of the world's population consumes natural bioactive compounds for the treatment of many diseases.[6]

The field of NP research was born in 1803, when Friedrich Wilhelm Sertürner isolated a crystalline substance from opium, which he called morphine.[7] About 96 years later, Felix Hoffmann and Arthur Eichengrün synthesised one of the world's most popular NP-based drugs, aspirin.[8] With the discovery of penicillin by Alexander Fleming in 1929, also NPs from other organisms than plants, raised the attention of research scientists.[9] Nowadays, a life without drugs based on NPs would not be imaginable. For example, the statins, one of the largest selling class of drugs worldwide and first-line pharmacologic intervention for dyslipidemia, were isolated from different fungi (e.g. *Aspergillus terreus* or *Monascus ruber*), while paclitaxel (one of the most prescribed anti-cancer drugs) was isolated from the plant *Taxus brevifolia*.[10][11] Indeed, NPs play a key role in drug discovery, especially for cancer and infectious diseases.[12] If one takes a look at all new approved drugs from the FDA between 1981 and 2019, 65 % are NPs, or based on NP-derived pharmacophores (**Fig. 1.1**).[13] In comparison to standard synthetic molecules, NPs offer a higher structural diversity and complexity. Usually they have a higher molecular mass, a larger number of sp^3 carbons, a higher molecular rigidity, a lower amount of rotatable bonds, but more unsaturated bonds, more hydrogen bonds acceptors and donators, and finally different amounts of heteroatoms (less nitrogen, more oxygen and sulfur atoms).[14] These properties can be advantageous or disadvantageous when it comes to the process of drug discovery. Especially, when synthetic chemistry reaches its limits, NPs might offer a

novel toolbox of possible structures and interactions with molecular targets.

In contrast to the advantages of NPs and their key role in drug discovery, several drawbacks led pharmaceutical companies to reduce their efforts in NPs based drug-discovery programs. These problems include the culturing of the organism, the extraction method, dereplication of known NPs, insufficient amounts of bioactive NPs, incompatibilities with High-Throughput-Screening (HTS) approaches, the unknown molecular target of NPs, and the need for benefit sharing with countries of origin of the biological material (Nagoya protocol).[12]
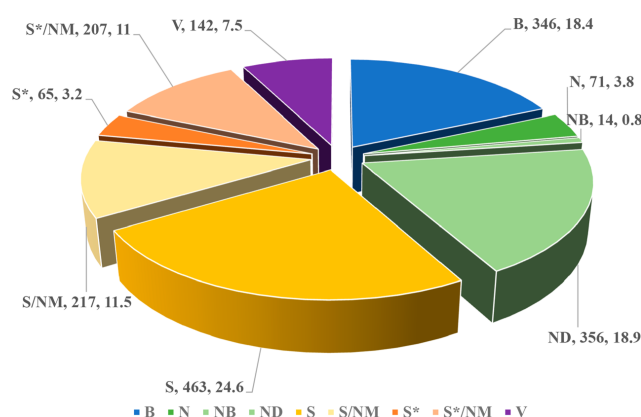
Figure 1.1.: All new approved drugs in the period from 01.01.1980 to 30.09.2019, categorized by their respective source. B – biological macromolecule, N – unaltered natural product, NB – botanical drug (defined mixture), ND - natural product derivative, S – synthetic drug, S* - synthetic drug (NP derived pharmacophore), V – vaccine, /NM –natural product-mimetic. Graphic adopted from the recurring review *Natural products as a source of new drugs* last published by Newman and Cragg.[13]

To overcome these obstacles, several new techniques were implemented in the field of NP research. These techniques slowly replace the traditional bioactivity guided approach. With the help of accurate and fast LC-MS measurements, the field of metabolomics was introduced in NPs research.[15] Crude extracts can now rapidly be screened concerning their metabolomic profile, and new dereplication tools like the GNPS molecular networking platform, or several *in silico* databases (e.g. Dictionary of Natural Products, The Natural Products Atlas, or the LOTUS initiative) can be used to identify and visualize possible new NPs.[16][17][18][19][20][21] Also, changes in the metabolite profile in various physiological states of the producing organism can be detected, allowing to take a deeper look in the complex behavior of NPs producing organisms.[22] Another building block is the incorporation of

genome analysis and genome mining in the process of NPs research.[23] This was driven forward due to two key characteristics, concerning the genome organisms producing NPs. Especially, in the case of bacteria and fungi, the genes of interest are clustered together, and they consist of highly conserved modules.[23] These clusters can be identified, and the structure of possible new NPs can be predicted based on the incorporated enzymes, helping in dereplication and identifying new NPs.[24][25][26] Expression of biosynthetic gene clusters in well-known organisms (e.g. *E. coli*) can improve the yield and purity of NPs, and silent clusters can be addressed as novel sources of NPs.[27][28]

A third major step was the addition of new cultivation techniques to the field of NPs research. The concept "One Strain Many Compounds" was the dominating concept since the early 1960er for cultivation of microorganisms.[29][30] This method has its limits, as it is not a realistic approach to simulate the natural environment of a microorganism. So-called "helper strains" or even insects, which have a symbiotic relationship with the organism of interest, were added to the culture to simulate a more realistic approach, which resulted in the discovery of new NPs.[31][32] Also, techniques where microorganisms can be cultivated in their natural soil (e.g. iChip) could overcome the obstacle of formerly unculturable microorganisms.[33]

In a nutshell NPs were, since their first discovery, a driver of innovation, especially in the field of drug discovery.[34] With the rise of antibiotic resistant pathogens and the rising incidence of cancer, NPs will surely continue to play a key role in future medicine.[35][36] Also, understanding the biological function of NPs, will give us a better understanding of our ecosystems and how we can protect them. Taking all this into account, research in the field of NPs is not only necessary, but also a privilege to discover what "Mother Nature" is still hiding from us.

1.2. Cyanobacteria

Cyanobacteria are photosynthetic prokaryotes, which can be found in a diverse range of environments.[37] The exact origin of cyanobacteria is still unknown, but cell-preserved fossils of cyanobacteria and rock-derived carbon isotope data are consistent with the presence of photosynthetic microorganisms back to ~3.5 billion years ago.[38] Cyanobacteria played a major role in the Great Oxidation Event, and all plant plastids share a common endosymbiotic cyanobacterial ancestor, which demonstrates the key role of cyanobacteria in evolutionary processes.[39] From a quantitative point of view, they are also among the most important organisms, which can be found on our earth. Their estimated wet biomass is around thousand million tons (10^{15} g).[40] So, it is no surprise that cyanobacteria can even be found in extreme environmental systems like deserts, polar ice, or hot springs.[37] The morphology of cyanobacteria is as flexible as their occurrence. It ranges from spherical and unicellular to filamentous and spiral.[41] Some filamentous cyanobacteria have devel-

oped functional cell differentiation into heterocysts (nitrogen fixation), akinetes (resting stage cells), and hormogonia (reproductive, motile filaments).[42] Most cyanobacteria live as independent autotrophs, but they were also found in symbiotic relationships with fungi or marine sponges.[43][44] Nowadays, over 270 cyanobacterial genera with over 3000 species have been described.[45]

In the public, cyanobacteria are mainly known as blue-green algae. This description is misleading, as cyanobacteria are no algae, which are eukaryotic organisms. The name blue-green algae is due to the pigment phycocyanin, which gives the cells a bluish color, when present in sufficient high concentrations.[46] Because of the formation of harmful "algae" blooms, cyanobacteria are often associated with toxic metabolites and environmental damage. These toxins demonstrate the capability of cyanobacteria to produce compounds with an impressing bioactivity, underlining the need for further research.

1.2.1. Cyanobacterial natural products

Considering that cyanobacteria can occur in varied and often extreme habitats, it is no surprise that they are capable of synthesizing a high number of diverse and bioactive specialized metabolites. This includes over 2000 compounds in 260 families, which could be classified into 10 major classes: alkaloids, depsipeptides, lipopeptides, macrolides/lactones, peptides, terpenes, polysaccharides, lipids, polyketides, and others.[47][48] **Fig. 1.2** illustrates the different compounds produced by cyanobacteria. Most of these compounds are peptides (25 %). Adding depsipeptides and lipopeptides, their proportion is rising to 55 %. Regarding their biological activity, cytotoxicity (30 %) and antibacterial activity (12 %) are the prominent observed effects.[47] In addition, recent genome studies have shown that the NPs discovered so far from cyanobacteria appear to be only a small part of what could be discovered in the future.[49] A detailed look into all classes of cyanobacterial metabolites is beyond the scope of this work and can be found in literature reviews.[50][51][52][53][54]

Despite the huge potential of cyanobacterial metabolites, they have been neglected a long time by the scientific community. Until the 1980's, cyanobacteria were only known to produce toxins and not as a source of potential new drug leads.[55][56] Even, if that has changed in the last 20 years, in comparison to other microorganisms (e.g. actinobacteria), cyanobacteria are still an underexplored source of NPs.[57] This is due to their cultivation as phototrophic microorganisms, slow growth, and oftentimes low yields of their NPs.[12] But with the above-mentioned new methods in NPs research, these obstacles will be tackled. Finally, the approval of the drug brentuximab vedotin, by the Food and Drug Administration (FDA) in 2011, demonstrated that cyanobacterial NPs can be the source of new drug candidates, not only on paper, but also in real-life.[58]

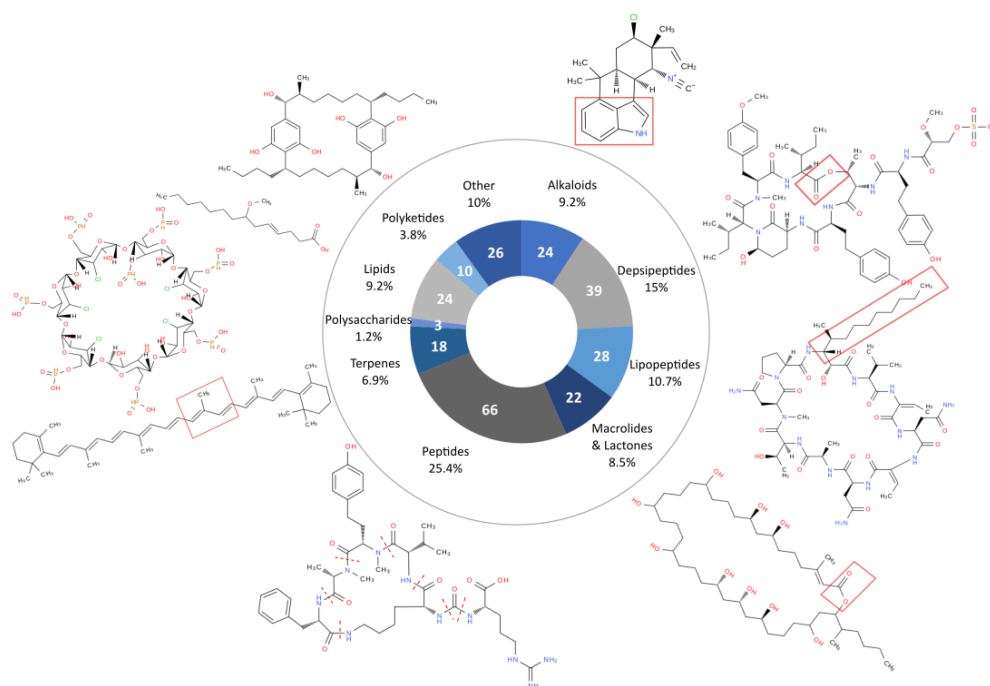**Figure 1.2.:** Classification of cyanobacterial metabolites according to their respective chemical class. The main characteristics of each chemical class are highlighted in red. Graphic adopted from the review *Natural products from cyanobacteria: Focus on beneficial activities published* by Demay et al.[47]
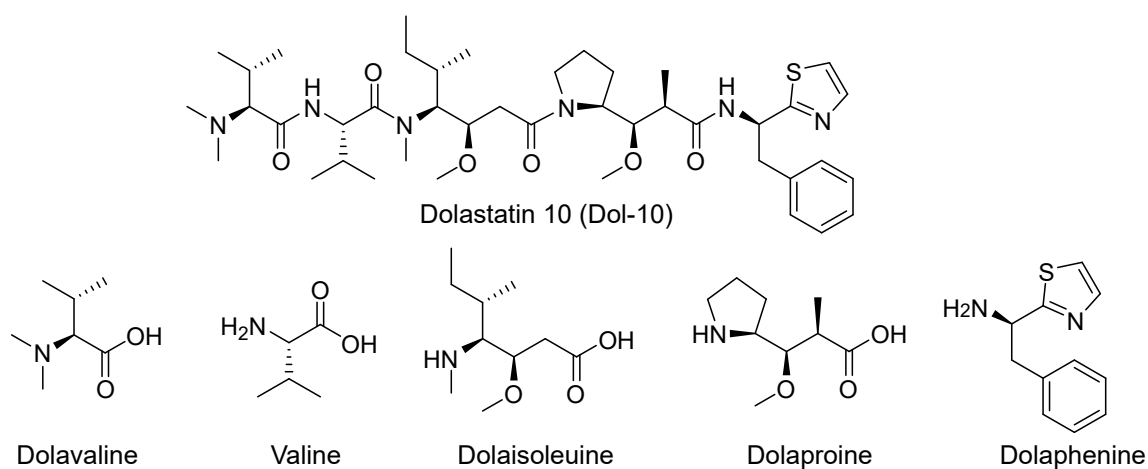

Figure 1.3.: Chemical structures of dolastatin 10 and its amino acid derived monomers.

1.2.2. Peptides from cyanobacteria: the dolastatins

As mentioned above, peptides are the main compound class produced by cyanobacteria. These peptides can be linear or cyclic, and they are often associated with toxic/cytotoxic activities.[51] Most of the peptides discovered so far are non-ribosomal peptides (NRPs), which are synthesised by specialised enzymes that cluster together[59] One of these linear NRPs is dolastatin 10 (Dol-10) (**Fig. 1.3**). Dol-10 was isolated the first time in 1987 from the marine shell-less mollusk *Dolabella auricularia.*[60] In later studies, it was clarified that not the mollusk is producing the compound, but the marine cyanobacterium *Symploca sp. VP642,* which is the main diet of the mollusk.[61] Structurally, Dol-10 is a linear pentapeptide with four amino acid derivatives (dolavaline, dolaisoleuine, dolaproine and dolaphenine), which were unknown so far, and which became characteristic for compounds of this class. From the beginning, Dol-10 raised the interest of the scientific community, as in a first test it was more potent than most of the known anticancer drugs so far (IC$_{50}$ 0.03 nM on L1210 leukemia cells).[62] As the yields of Dol-10 were quite low and the stereochemistry still unknown, a total synthesis was needed and established in a short period of time after the discovery.[63] Stereochemistry and the needed amount for further studies could be solved in one step, and in a continued screening it could be demonstrated that Dol-10 was highly active against different kinds of cancer cells like: L120 leukemia cells, B16 melanoma cells, small cell lung cancer NCI-H69 cells, and human prostate cancer cells.[62][64][65][66] The IC$_{50}$ was always in the low nM region (0.03–0.59 nM). Further studies showed that the toxic effects on different cancer cell lines is due to the ability of Dol-10 to inhibit tubulin polymerization, inducing apoptosis of affected cells. Later, the β-tubulin subunit could be identified as the molecular target of Dol-10 (**Fig. 1.4**).[67][68][69] In the following years, several derivatives from other marine cyanobacteria could be isolated (e.g. symplostatin 1 or malevamide D, **Fig. 1.5**), all with similar activity profile.[70][71][72] Even cyclic derivatives could be detected and isolated.[73][74] Dol-10 was studied in clinical trials, but because of adverse effects (peripheral neuropathy), the clinical trials were stopped in phase II.[75] In the meantime, medicinal chemists synthesized a large library of Dol-10 analogues, resulting in a better understanding of the structure-activity relationships and in two promising candidates for further clinical studies: monomethylauristatin F (MMAF) and E (MMAE)(**Fig. 1.5**).[76][77][78]

With the rise of antibodies in modern medicine, scientists developed the idea to couple highly toxic substances to antibodies, which can bind to a specific cell type, and release the toxic substance only in the addressed type of cell.[79] Thus, antibody-drug conjugates (ADCs) were born. The idea behind this concept is that side effects can be avoided, as the toxic compound, called "warhead", is only released in the targeted cells, e.g. cancer cells (**Fig. 1.6**).[80] MMAE was chosen as a promising candidate for this technique. In 2003, for the first time, a publication was presented that described the successful devel-

opment of an ADC with MMAE as a warhead.[81] Finally, in 2011 the FDA approved the drug, which was called brentuximab vedotin, for the treatment of anaplastic large T-cell systemic malignant lymphoma and hodgkin lymphoma.[82] Encouraged by this success, the efforts in developing ADCs with compounds from the Dol-10 compound family as a "warhead" significantly increased. Four more drugs with MMAE, as a "warhead", and one more drug with MMAF, as the cytotoxic compound, were approved until November 2020.[83] In addition, around 30 clinical trials are ongoing with compounds from the Dol-10 family coupled to an antibody, underlining the success of cyanobacterial NPs in ADCs.[83]

Figure 1.4.: Model of spindle microtubules in the mitosis and binding target of dolastatin 10. Graphic adopted from the review *Marine antitumor peptide Dolastatin 10: Biological activity, structural modification and synthetic chemistry* by Gao et al.[84]

Figure 1.5.: Chemical structures of symplostatin 1, malevamide D, monomethylauristatin E, and monomethylauristatin F.

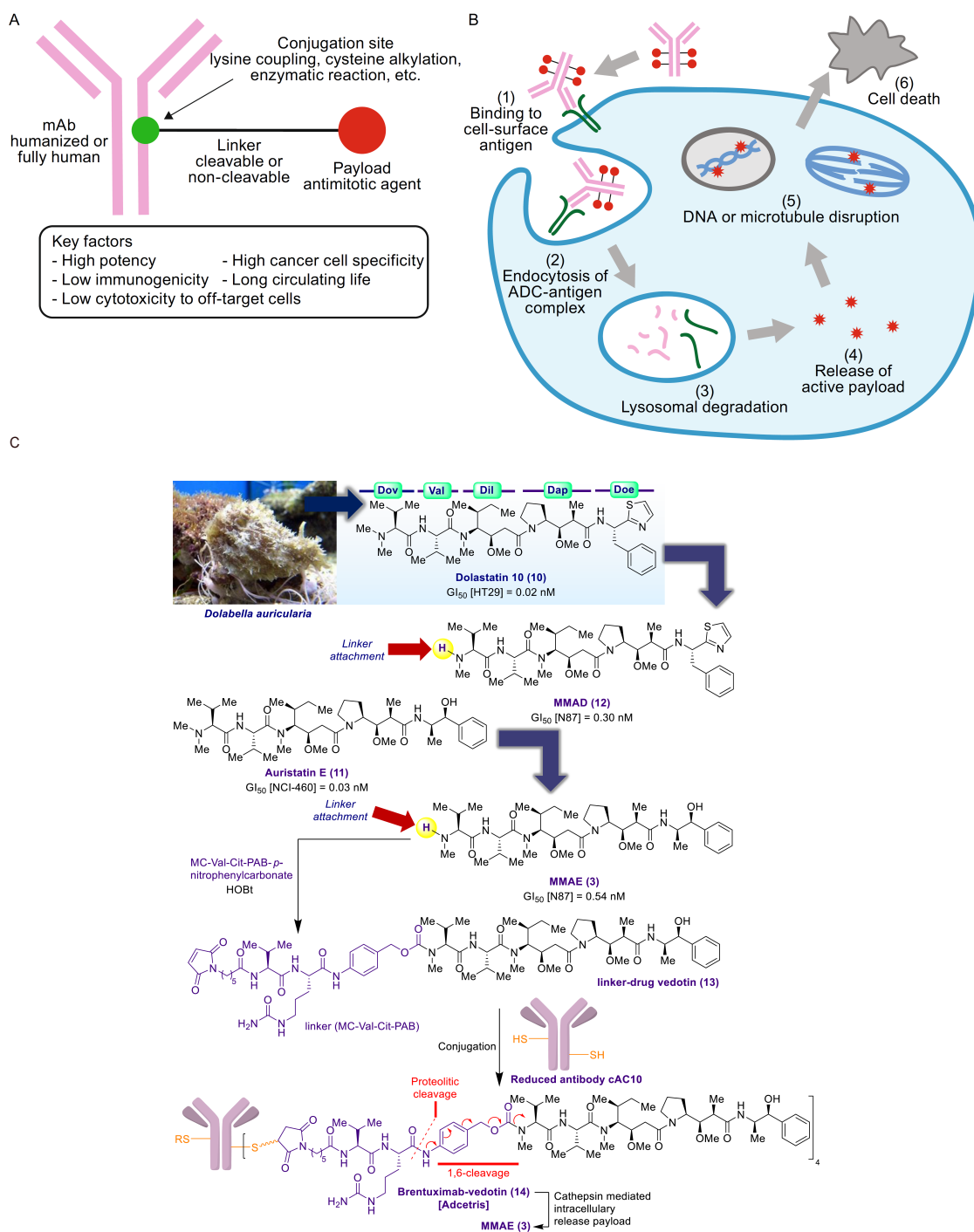# 1. Introduction

Figure 1.6.: (**A, B**) Structure and mechanism of action of ADCs. (**C**) From natural product (Dol-10) to ADC. Graphics adopted from the paper *Antibody-drug conjugates: Recent advances in conjugation and linker chemistries* published by Tsuchikama et al. and the review *Antibody-drug conjugates containing payloads from marine origin* published by Chen-Sánchez et al.[80][85]

1.2.3. *Aetokthonos hydrillicola,* the "eagle-killer"

In winter 1994/1995, at the DeGray Lake (USA), the biggest mass mortality of bald eagles could be observed.[86] As the cause of death, a spongiform myelinopathy could be identified, which was called Vacuolar Myelinopathy (VM).[86][87][88] VM is histologically diagnosed via lesions in the white matter of the brain. Clinical signs of this disease are loss of motor function and neurological impairments.[86] The cause of this disease was unknown, as in VM affected animals no xenobiotics, pathogens, or neuroinflammation could be detected, but it could be revealed that VM is transferred up the food chain (**Fig. 1.7**).[89][90][91][92] The focus shifted to the water bodies where VM occurred, and it could be shown that in all affected lakes a water plant was growing: *Hydrilla verticillata.*[93][94] In the affected water bodies, many of the plant leaves were colonialized by a so far unknown cyanobacterium, raising the suspicion that a cyanobacterial toxin might be the cause of VM.[93][94] Based on 16S rRNA sequencing, the cyanobacterium could not be assigned to any known genus and was described as *Aetokthonos hydrillicola gen. et sp. nov.*[95]

Cultivation of *A. hydrillicola* was difficult. It took two years to generate enough biomass for feeding trials. Unfortunately, birds fed with biomass from cultured *A. hydrillicola* did not develop VM. Mass spectrometry imaging studies of *H. verticillata* leaves, colonized with *A. hydrillicola*, revealed the presence of a hitherto unknown pentabrominated compound.[96] As standard culture media for cyanobacteria does not contain any bromide, it was now obvious why the compound was not produced under standard cultivation conditions. Cultivation of *A. hydrillicola* with bromide supplementation resulted in the production of the novel compound. Structure elucidation revealed this compound to be a pentabrominated bi-indole alkaloid with a rarely seen connection between the indole moieties (**Fig. 1.7**).[96] The isolated compound was fed to chicken, and it could be demonstrated that the compound was the reason for the development of lesions in the white matter of the brain.[96] In addition, the compound could only be detected at lakes where VM occurred. The compound was given the name aetokthonotoxin (AETX), which is based on greek language and means: poison that kills the eagle.[96] AETX was also tested on HeLa cells, the nematode *C. elegans*, and larval zebrafish (*Danio rerio*), where the toxic effect could be confirmed.[96] After 25 years of research, the riddle of these mysterious deaths of bald eagles in the USA could be solved, again demonstrating what an influence cyanobacterial NPs can have on our ecosystems.

Figure 1.7.: From the cyanobacterium to the bald eagle - toxin transmission through the food chain. Graphic adopted from the paper *Hunting the eagle killer: A cyanobacterial neurotoxin causes vacuolar myelinopathy* published by Breinlinger et al.[96]

1.3. Objectives and motivation

As pointed out above, cyanobacterial NPs are a promising source of potential new drug lead candidates, and they also play a significant role in the balance of our ecosystems. Especially, the cyanobacterium *A. hydrillicola* raised the attention of the scientific community, as it could be identified to produce a novel neurotoxin (AETX), which was responsible for the death of bald eagles in the southeastern United States. In a first screening of the crude extract of *A. hydrillicola*, strong cytotoxicity could be observed, which could not be linked to AETX.

The aim of this work was thus, to identify the unknown cytotoxin, and to further characterize the secondary metabolome of *A. hydrillicola* using a combination of the traditional bioactivity guided workflow and new *in silico* tools for dereplication and identification of NPs.

In a first step, an extraction protocol had to be optimized to obtain extracts containing a high yield of the secondary metabolites. For the analysis of the extracts, HPLC/HPLC-MS methods had to be optimised, including the establishment of a microfractionation method, which was used for the bioactivity-guided isolation of the putative cytotoxic compound. Dereplication and identification of derivatives, as well as novel compounds, rested on high-resolution mass spectrometry data, which were analyzed with new computational tools, like the "Global Natural Products Social Molecular Networking" (GNPS). For the isolation of compounds, different chromatographic methods, i.e. flash chromatography, size-exclusion chromatography, and semi-preparative HPLC, were tested. After compound isolation, their structures had to be elucidated. This was based on modern analytical methods, including one- and two-dimensional NMR spectroscopy and high-resolution mass spectrometry. The bioactivity of the fractions and of the obtained pure compounds had to be tested, using a cell-based cytotoxicity assay, as well as assays using the nematode *C. elegans*. Finally, as AETX and AETX-related compounds in the extracts are of interest, it was necessary to establish a synthesis of them.

2. Material and methods

2.1. Chemical work

2.1.1. Chemicals

All chemicals used for extraction, chromatography, synthesis, or any other analyses are listed in **Tab. 2.1**.

Table 2.1.: Chemicals and solvents.

Chemical	Vendor
Acetic Acid	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Acetonitrile HPLC grade (MeCN)	Honeywell Riedel-de Haën, Charlotte, USA
Acetonitrile anhydrous	VWR Chemicals, Darmstadt, DE
Acetone	VWR Chemicals, Darmstadt, DE
Ammonium acetate	VEB Laborchemie GmbH, Apolda, DE
Celite	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Chlorosulfonyl isocyanate (CSI)	Acros Organics, New Jersey, USA
Diammonium hydrogen phosphate	Merck KGaA, Darmstadt, DE
5,7-dibromoindole	Sigma-Aldrich, St.Louis, USA
5,7-dibromoindole-3-carboxaldehyde	Merck KGaA, Darmstadt, DE
Dichloromethane-d_2	Thermo Fisher Scientific, Waltham, USA
Dimethylformamide anhydrous (DMF)	VWR Chemicals, Darmstadt, DE
Dimethyl sulfoxide (DMSO)	AppliChem GmbH, Darmstadt, DE
Dimethyl sulfoxide-d_6	VWR Chemicals, Darmstadt, DE
1-Fluor-2-4-dinitrophenyl-5-L-alaninamid (FDAA; Marfey-reagent)	Thermo Fisher Scientific, Waltham, USA
Formic acid (FA)	VWR Chemicals, Darmstadt, DE
Hydrochloric acid (HCl)	AppliChem GmbH, Darmstadt, DE
Methanol (MeOH)	Honeywell Riedel-de Haën, Charlotte, USA
Monomethyllauristatin E (MMAE)	Hycultec GmbH, Beutelsbach, DE
1-nitropropane	Merck KGaA, Darmstadt, DE
Sodium bicarbonate	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Water (MS purity grade)	AppliChem GmbH, Darmstadt, DE

2.1.2. Cultivation media for cyanobacteria

A. hydrillicola was cultivated in standard BG11 medium with the addition of 0.42 mM potassium bromide (KBr).[96][97] Components are listed in **Tab. 2.2**. All stock solutions were prepared separately by dissolving the compounds in demineralized water at the respective concentrations. Stock solutions 1-8 were autoclaved at 121 °C. Solutions 9-11 were sterile filtered through a membrane filter with 0.22 µm pore size. The stock solutions were added to 944 mL demineralized water, which was autoclaved at 121 °C prior to use, under aseptic conditions, using the volumes specified in **Tab. 2.2**. The ready-to-use BG11 medium and stock solutions 1-8, 10, and 11 were stored at room temperature. Stock solution 7 was protected from light to prevent oxidation of citric acid. Solution 9 was kept at 4 °C and protected from light.

Table 2.2.: Composition of BG11 medium supplemented with 0.042 mM KBr.

Stock solution	Ingredients	Concentration in stock solution [g/L]	Volume per L media [ml]	Final concentration [g/L]
1	$NaNO_3$	300	5	1.5
2	K_2HPO_4	6.25	5	0.032
3	$MgSO_4$ x 7 H_2O	15	5	0.750
4	$CaCl_2$ x 2 H_2O	7.2	5	0.036
5	EDTA	0.2	5	1.00×10^{-3}
6	Na_2CO_3	8	5	0.040
7	Citric acid	1.2	5	6.00×10^{-3}
	Ammonium ferric citrate	1.2	5	6.00×10^{-3}
8	KBr	10	5	0.05
9	H_3BO_3	2.86	1	2.86×10^{-3}
	$MnCl_2$ x 4 H_2O	1.81	1	1.81×10^{-3}
	$ZnSO_4$ x 7 H_2O	0.22	1	0.22×10^{-3}
	Na_2MoO_4 x 2 H_2O	0.39	1	0.39×10^{-3}
	$CuSO_4$ x 5 H_2O	0.079	1	0.079×10^{-3}
	$Co(NO_3)_2$ x 6 H_2O	0.049	1	0.049×10^{-3}
10	$NaHCO_3$	84.1	5	0.421
11	TES	229.25	10	2.29

2.1.3. Cultivation conditions

A. hydrillicola was cultivated in 1000 mL glass bottles in 1000 mL of standard BG11 medium supplemented with KBr. Cultures were gently stirred using a magnetic stirrer (100 rpm) and supplemented with 5 % CO_2 (1 L/h). Fluorescent tubes (Sylvania GroLux, F18W/GRO) were used as light source, providing an average light intensity of 35 μE. Cultivation temperature was 28 °C.

2.1.4. Harvesting of cyanobacterial biomass

The cyanobacterial biomass was separated from the culture medium via filtration through a net. The obtained biomass was washed with distilled water, transferred into 50 mL falcon tubes, frozen, dried under reduced pressure, and stored at -18 °C until further use.

2.1.5. Extraction of cyanobacterial biomass

Lyophilized biomass was crushed with a spatula and transferred to 300 mL extraction bottles. In a first step, 50 % (v/v) $MeOH/H_2O$ was added (approx. 1 mL/10 mg biomass), and the mixture was sonicated with a sonotrode (output control: 9; 90 % duty cycle) for 5 min on ice. After sonification, the mixture was shaken on an overhead shaker for 1 h. Subsequently, the mixture was centrifuged (4900 rpm, 10 min, 20 °C), and the supernatant was separated from the biomass. This extraction was repeated twice using 80 % (v/v) $MeOH/H_2O$ in the second and 100 % MeOH in the third step. The supernatants from each extraction step were combined in a flask and dried under reduced pressure using a rotary evaporator. For further use, the dried biomass extract was resuspended in MeOH and filtered to remove undissolved particles.

2.1.6. Analytical HPLC

For analytical HPLC, an UltiMate 3000 system was used (Thermo Fisher Scientific, Waltham, USA). The components of the instrument are summarized in **Tab. 2.3**. All other parameters of the methods are listed in **Tab. 2.4**–**2.6**. Prior to analysis, the samples were filtered through a 0.45 µm filter into HPLC vials. The following HPLC columns were used, as indicated in the methods:

Kinetex C_{18} (2.6 µm, 100 Å, 2.1 mm ID x 100 mm, Phenomenex, Torrance, USA), Luna C_{18} (5 µm, 100 Å, 4.6 mm ID x 250 mm, Phenomenex, Torrance, USA), Luna phenyl-hexyl (5 µm, 100 Å, 4.6 mm ID x 250 mm, Phenomenex, Torrance, USA), and Luna pentafluorophenyl (5 µm, 100 Å, 4.6 mm ID x 250 mm, Phenomenex, Torrance, USA). Data obtained from analytical HPLC was analyzed with Chromeleon 7 (version 7.2.6, Thermo Fisher Scientific).

Table 2.3.: Analytical HPLC setup.

System	Component
Thermo Fisher Scientific	Quaternary pump (LPG-3400SD)
UltiMate 3000	Autosampler (WPS-3000TSL)
	Column oven (TCC-3000SD)
	Diode array detector (DAD-3000)
	Fraction collector (AFC-3000)

Table 2.4.: Analytical HPLC screening method using a Kinetex C_{18} column.

Parameter	Value
Instrument	UltiMate 3000
Column	Kinetex C_{18} (2.6 µm, 100 Å, 2.1 mm ID x 100 mm)
Solvents	A: H_2O + 0.1 % FA
	B: MeCN + 0.1 % FA
Elution	Gradient: 5–100 % B over 25 min; plateau 100 % B for 5 min
Flow rate	0.5 mL/min
Temperature	50 °C
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

Table 2.5.: Analytical HPLC screening method using different Luna columns C_{18}/phenyl-hexyl (PH)/pentafluorophenyl (PFP).

Parameter	Value
Instrument	UltiMate 3000
Column	Luna C_{18}/PH/PFP (5 µm, 100 Å, 4.6 mm ID x 250 mm)
Solvents	A: H_2O + 0.1 % FA
	B: MeCN or MeOH + 0.1 % FA
Elution	Gradient: 5–100 % B over 25 min; plateau 100 % B for 5 min
Flow rate	1.3 mL/min
Temperature	25 °C
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

For method development on Luna columns, the following generic method was applied using solvent modification, gradient, and temperature, as indicated in the results and discussion section.

Table 2.6.: Analytical HPLC generic method using a Luna C_{18} column.

Parameter	Value
Instrument	UltiMate 3000
Column	Luna C_{18} (5 µm, 100 Å, 4.6 mm ID x 250 mm)
Solvents	A: H_2O + 0.1 % FA or 10 mM ammonium acetate
	B: MeCN + 0.1 % FA or 10 mM ammonium acetate
Elution	Gradient: initial gradient as indicated; plateau 100 % B for 5 min
Flow rate	1.3 mL/min
Temperature	as indicated
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

2.1.7. Microfractionation

Microfractionation of the crude extract was performed on an UltiMate 3000 (Thermo Fisher Scientific, Waltham, USA). The components of the instrument are listed in **Tab. 2.3**. Prior to analysis, all samples were filtered through a 0.45 µm filter into HPLC vials. All parameters of the used method are listed in **Tab. 2.7**. Fractions were collected in 96-well plates, air-dried, redissolved in 50 µL 10 % (v/v) DMSO in water, and submitted to cytotoxicity testing.

Table 2.7.: Microfractionation method on an UltiMate 3000 HPLC system.

Parameter	Value
Instrument	UltiMate 3000
Column	Luna C_{18} (5 µm, 100 Å, 4.6 mm ID x 250 mm)
Solvents	A: H_2O + 0.1 % FA
	B: MeCN + 0.1 % FA
Elution	Gradient: 5–100 B % over 25 min; plateau 100 % B for 5 min
Flow rate	1.3 mL/min
Temperature	25 °C
Fraction collection	Time-based: 30 sek per fraction
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

2.1.8. HPLC-MS

For HPLC-MS, an UltiMate 3000 (Thermo Fisher Scientific, Waltham, USA) system coupled to a Q Exactive Plus (Thermo Fisher Scientific, Waltham, USA) mass spectrometer was used. The components of the instruments are summarized in **Tab. 2.8**. Prior to analysis, all samples were filtered through a 0.45 µm filter into HPLC vials. All parameters of the used method are listed in **Tab. 2.9**. Data obtained from HPLC-MS analysis were analyzed with the software FreeStyle 1.6 (Thermo Fisher Scientific, Waltham, USA).

Table 2.8.: HPLC-MS components.

System	Component
Thermo Fisher Scientific UltiMate 3000	Quaternary pump (DGP-3600A)
	Autosampler (WPS-3000TSL)
	Column oven (TCC-3100 1x2P-10P)
	Diode array detector (PDA-3000)
Thermo Fisher Scientific mass spectrometer	Q Exactive Plus

Table 2.9.: HPLC-MS method for an UltiMate 3000 coupled to a Q Exactive Plus.

Parameter	Value
Instrument	UltiMate 3000 and Q Exactive Plus Orbitrap
Column	Kinetex C_{18} (5 µm, 100 Å, 2.1 mm ID x 50 mm)
Solvents	A: H_2O + 0.1 % FA
	B: MeCN + 0.1 % FA
Elution	Gradient: 5–100 % B over 16 min; plateau 100 % B for 4 min
Flow rate	0.4 mL/min
Temperature	50 °C
Injection volume	1 µL
HESI source settings	Spray voltage: 3.5 kV
	Sheath gas flow rate: 50 l/min
	Aux gas flow rate: 13 l/min
	Sweep gas flow rate: 3 l/min
	Capillary temperature: 350 °C
	S-lens RF level: 50 V
	Aux gas heater temp: 425 °C
MS settings	Scan range: $150 - 1950$ m/z
	Polarity: both
	Full MS resolution: 35000
	Stepped NCE: 30, 60, 75

2.1.9. Flash chromatography

Flash chromatography was performed on a Gilson HPLC system (Gilson, Middleton, USA) in the configuration as listed in **Tab. 2.10**. As column, a Chromabond C_{18} 43 g flash cartridge (Macherey-Nagel, Düren, DE) was used. Initially, the cartridge was activated with 5 column volumes of 100 % MeCN, and then equilibrated with another 5 column volumes of the starting conditions. For sample preparation, the extract was resuspended in MeOH and Celite was added in a ratio of 1:4, corresponding to the weight of the dried sample. MeOH was removed under reduced pressure, resulting in a dry powder. The powder was transferred to an empty 12 g flash cartridge, and placed in front of the separation column. First crude fractionation was achieved by a linear gradient from 5-100 % MeCN over 40 min, with a plateau of 8 min MeCN at the end, at a flow rate of 20 mL/min (**Tab. 2.11**). Fractions were collected every 2 min and analyzed via HPLC-MS. Fractions with similar composition were combined.

Table 2.10.: Gilson flash chromatography setup.

System	Component
Gilson	Binary pump (322-H2)
	Diode array detector (171)
	Column oven (TCC-3000SD)
	Liquid handler (GX-271)

Table 2.11.: Flash chromatography method on a Gilson system.

Parameter	Value
Instrument	Gilson
Column	Chromabond C_{18} (43 g)
Solvents	A: H_2O
	B: MeCN
Elution	Gradient: 5–100 % B over 48 min; plateau 100 % B for 12 min
Flow rate	20 mL/min
Temperature	Room temperature
Fraction collection	Time-based: 2 min per fraction
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

2.1.10. Sixe-exclusion chromatography

Sephadex LH-20 dry material was suspended in the mobile phase, let sit to swell, and form a homogeneous slurry for 6 h. Carefully, not to introduce air bubbles into the column bed, the slurry was poured into an empty gravity-flow glass column (25 mm ID x 900 mm), and the stationary phase was let to settle for 4 h by purging with mobile phase. 250 mg of dried extract were redissolved in 5 mL of the initial mobile phase and loaded onto the column. The flow rate was adjusted to about 30 mL/h, and fractions were collected every 20 min over 24 h. The obtained fractions were dried *in vacuo* and resuspended in 1 mL MeOH, before subjecting them to HPLC-MS analysis.

2.1.11. Semi-preparative HPLC

For semi-preparative HPLC, an UltiMate 3000 (Thermo Fisher Scientific, Waltham, USA) system was used. The components of the instrument are summarized in **Tab. 2.3**. All other parameters of the methods are listed in **Tab. 2.12–2.13**. Prior to analysis, all samples were filtered through a 0.45 µm filter into HPLC vials. The following HPLC column was used, as indicated in the methods: Luna C_{18} (5 µm, 100 Å, 10 mm ID x 250 mm). Fractions identical in composition were combined and dried *in vacuo*.

Table 2.12.: Semi-preparative HPLC method on an UltiMate 3000 HPLC system using a gradient of 50-85 % MeCN in water.

Parameter	Value
Instrument	UltiMate 3000
Column	Luna C_{18} (5 µm, 100 Å, 10 mm ID x 250 mm)
Solvents	A: H_2O + 10 mM ammonium acetate
	B: MeCN
Elution	Gradient: 50–85 % B over 17.5 min; plateau 100 % B for 5 min
Flow rate	6.1 mL/min
Temperature	25 °C
Fraction collection	Peak based: threshold 150 mAU, slope 15 mAU/s
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

Table 2.13.: Semi-preparative HPLC method on an UltiMate 3000 HPLC system using isocratic conditions of 55 % MeCN in water.

Parameter	Value
Instrument	UltiMate 3000
Column	Luna C_{18} (5 µm, 100 Å, 10 mm ID x 250 mm)
Solvents	A: H_2O + 0.1 % FA
	B: MeCN + 0.1 % FA
Elution	Isocratic: 55 % B for 18 min; plateau 100 % B for 5 min
Flow rate	6.1 mL/min
Temperature	25 °C
Fraction collection	Peak based: threshold 150 mAU, slope 15 mAU/s
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

2.1.12. Preparative HPLC

For preparative HPLC, the Gilson chromatography system was used. The components of the instrument are summarized in **Tab. 2.10**. All other parameters of the method are listed in **Tab. 2.14**. Prior to analysis, all samples were filtered through a 0.45 µm filter into HPLC vials. The following HPLC column was used as indicated in the method: Kinetex C_{18} (5 µm, 100 Å, 21.2 mm ID x 150 mm). Fractions identical in composition were combined and dried *in vacuo*.

Table 2.14.: Preparative HPLC method on a Gilson system using a Kinetex C_{18} column.

Parameter	Value
Instrument	Gilson
Column	Kinetex C_{18} (5 µm, 100 Å, 21.2 mm ID x 150 mm)
Solvents	A: H_2O
	B: MeCN
Elution	Gradient: 30-70 % B over 12 min; plateau 100 % B for 5 min
Flow rate	20 mL/min
Temperature	Room temperature
Fraction collection	Peak based: threshold 150 mAU, slope 15 mAU/s
DAD settings	Channel 1/2/3/4: 210/238/254/430 nm

2.1.13. Marfey´s analysis

The absolute configuration of dolaproine and valine was established using Marfey's method.[98] AEST (**1**) and MMAE (500 µg each) were hydrolysed in 6 N HCl (1 mL), in a sealed reaction vessel at 110 °C for 24 h. After drying under a stream of nitrogen, FDAA solution [(1-fluoro-2,4- dinitrophenyl)-5-L-alanine amide] in acetone (50 μL, 0.04 M) and NaHCO$_3$ (100 μL, 1 M) were added to each reaction vessel. The solution was kept at 80 °C for 3 min, cooled to room temperature, and acidified with HCl (2 M, 50 μL). After the addition of MeCN/water (200 µL, 1:1), the FDAA-amino acid derivatives from the hydrolyses were analysed and compared by HPLC-MS using the method described in **2.1.8**. The absolute configuration was determined by comparing the retention time of the derivatized monomers from MMAE with the retention time of the derivatized monomers from AEST (**1**).

2.1.14. NMR spectroscopy

1D and 2D NMR data were recorded on an Agilent Varian VNMRS 600 MHz spectrometer, operating at 600 MHz (^{1}H) and 150 MHz (^{13}C), using DMSO-d_6 (referenced to δH 2.51, δC 39.54 for residual solvent) and CD$_2$Cl$_2$ (referenced to δH 5.32, δC 54.00 for residual solvent) as solvents. The instrument was equipped with a 5 mm cryoprobe. The obtained data were analysed using the software Mnova (Version 14.3) and ACD/Spectrus Processor (Version 66513).

2.1.15. Carbonitrile synthesis

2.1.15.1. Method A

A mixture of 5,7-dibromoindole-3-carboxaldehyde (1000 mg, 3.64 mM), diammonium hydrogen phosphate (2.57 g, 19.49 mM), 1-nitropropane (125 mL, 23.35 mM), and 3.68 ml of glacial acetic acid was refluxed for 12.5 hours. The volatile reactants and solvents were removed under reduced pressure, and an excess of water was added to the residue. After a short time, crude indole-3-carbonitrile precipitates rapidly. The precipitate was filtered off, redissolved in DMF, and further purified by preparative HPLC using a linear gradient from 30–70 % MeCN in water (**Tab. 2.14**).

White solid; yield: 60 %; ^{1}H NMR (400 MHz, DMSO-d$_6$) δ: 12.73 (s, 1H, NH), 8.41 (d, 1H, J = 3.1 Hz), 7.85 (d, 1H, J = 1.7 Hz), 7.72 (d, 1H, J = 1.7 Hz).

2.1.15.2. Method B

To a solution of 5,7-dibromoindole (1000 mg, 3.64 mM) in anhydrous MeCN (4.1 mL), chlorosulfonyl isocyanate (CSI) (1.63 mL, 3.89 mM) was added dropwise at 0 °C, and the reaction mixture was stirred at 0 °C for 2 h. Anhydrous DMF (7.0 mL, 91 mM) was added dropwise, and the mixture was stirred at 0 °C for 2 h. The mixture was poured into ice-water, and the obtained precipitate was filtered off, redissolved in DMF, and further purified by preparative HPLC using a linear gradient from 30–70 % MeCN in water (**Tab. 2.14**).

White solid; yield: 85 %; ^{1}H NMR (400 MHz, DMSO-d$_6$) δ: 12.73 (s, 1H, NH), 8.41 (d, 1H, J = 3.0 Hz), 7.85 (d, 1H, J = 1.7), 7.73 (d, 1H, J = 1.7 Hz).

2.2. Biological work

2.2.1. Cell culture

2.2.1.1. Media and reagents

All media, chemicals, and other reagents used in cell culture and cytotoxicity testing are listed in **Tab. 2.15**. 10 x phosphate-buffered saline (PBS) was prepared by dissolving the components in **Tab. 2.16** in Milli-Q water and autoclaving at 121 °C. 1 x PBS was prepared from this stock by 1:10 addition of demineralized water, autoclaved, and stored at 4 °C until further use. Fetal bovine serum (FBS) was heat-inactivated, prior to use, by heating to 60 °C in a temperature-controlled water bath for 30 min and stored at -20 °C until further use. 10 x trypsin-EDTA was diluted 1:10 with sterile PBS to 1 x trypsin–EDTA under aseptic conditions and stored at -20 °C until further use. Working stocks of 1 x trypsin–EDTA were kept at 4 °C to avoid repetitive cycles of thawing and freezing. Dimethyl sulfoxide (DMSO) was sterilized by filtration through a membrane filter with 0.22 µm pore size and stored at 4 °C. Growth medium (Dulbecco's modified Eagle medium low glucose, DMEM) was supplemented with 10 % (v/v) FBS and 2 mM glutamine and kept at 4 °C up to 12 months at maximum.

Table 2.15.: Chemicals and media used in cell culture.

Reagent	Vendor
Acetic acid	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Doxorubicin hydrochloride	Cayman Chemicals, An Arbor, USA
Dulbecco's modified eagle medium low glucose	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Fetal bovine serum (FBS)	Sigma-Aldrich, St.Louis, USA
Roti-Cell glutamine solution (200 mM)	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Sulforhodamine B (SRB)	Sigma-Aldrich, St.Louis, USA
Trichloroacetic acid	Merck KGaA, Darmstadt, DE
Tris-base	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Trypan blue	Sigma-Aldrich, St.Louis, USA
10 x Trypsin-EDTA	Sigma-Aldrich, St.Louis, USA

Table 2.16.: Composition of phosphate-buffered saline (PBS).

Component	Concentration
NaCl	80.0 g/L
KCl	2.0 g/L
Na_2HPO_4 x 2 H_2O	14.4 g/L
KH_2PO_4	2.4 g/L

2.2.1.2. Cell lines

HeLa cells were purchased from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). Cells were cultured in DMEM supplemented with 10 % fetal bovine serum and 2 mM glutamine solution at 37 °C in a humidified incubator with 5 % CO_2. The medium was exchanged every two to three days. The spent culture medium was aspirated, and fresh growth media was let to trickle down the flask wall carefully, so the cells were not accidentally detached due to shear forces. When the cells had reached a confluency of 80–90 % (every three to four days), they were sub-cultured ("passaged") according to the following protocol. Growth media, 1 x PBS, and 1 x trypsin–EDTA were pre-warmed in a temperature-controlled water bath at 37 °C. The culture medium was carefully aspirated, and the cells were washed with 5 mL PBS. After aspiration of the PBS, the cells were treated with 2 mL of trypsin–EDTA and incubated at 37 °C for several minutes to allow for a sufficient cell detachment. Protein digest was stopped by addition of 3 mL growth medium, and the cells were carefully flushed from the flask surface. The cell suspension was transferred into a sterile centrifuge tube and spun down at 150 x g for 7 min at room temperature. In the meantime, 9 mL of growth medium were put into a new T-25 cell culture flask. After centrifugation, the supernatant was aspirated, and the cell pellet was resuspended in growth medium. The cells were split at ratios between 1:3 and 1:5 using between 3 and 5 mL of growth media for resuspending the cell pellet and 1 mL of suspension for inoculation of the new flask. All cells were continuously sub-cultured until a passage number of around 40 was reached.

2.2.1.3. Cell counting

Cells were harvested and trypsinized as described in chapter **2.2.1.2.** After centrifugation, the cell pellet was resuspended in 8 mL growth medium. A 1 mL aliquot is taken out, and the cells were stained with 10 μL of trypan blue stock solution. The cell suspension was applied on an EVE cell counting slide (NanoEnTek Inc., Seoul, KR) and counted with an EVE automated cell counter.

2.2.1.4. Sulforhodamine B (SRB) colorimetric cytotoxicity assay

The SRB assay was based on the protocol published by Vichai and Kirtikara.[99] Exponentially growing HeLa cells (confluency 70–90 %) were harvested by trypsinization and counted as described above. The cell concentration was adjusted in a manner that 100 μL contained the desired number of cells per well (10,000 cells per well). The cells were allowed to adhere overnight. Dilutions of the compounds were prepared in growth medium, and subsequently the culture medium was removed, replaced by the compound-containing dilutions in triplicates, and incubated for 48 h. As a positive control, a sample containing a concentration of 100 μM doxorubicin was used. For cell fixation, 100 μL cold TCA was directly added to the cells, grown in 96-well plates and incubated for 1 h at 4 °C. Subsequently, the medium/TCA mixture was removed, and the fixated cell layer was washed three times with 200 μL demineralized water. The plates were dried completely with pressurized air. The dry cell layer was stained with 100 μL SRB working stock and incubated at 4 °C for 30 min. Having removed the staining solution, the cells were washed three times with 200 μL of 1 % (v/v) acetic acid. The plate was dried with pressurized air. 200 μL of 10 mM tris-base (pH 10.5) was added to solubilize the stain, and the plate was shaken on an orbital shaker until the dye was homogeneously dissolved. The absorbance at 510 nm was measured with a TECAN Infinite M200 Pro plate reader. GraphPad Prism 6 (Dotmatics, Version 6.01) was used for normalization and plotting of the data.

2.2.2. *Caenorhabditis elegans*

2.2.2.1. Media and reagents

All media, chemicals, and other reagents used in *C. elegans* related work are listed in **Tab. 2.17**. 1 M potassium phosphate buffer (pH 6.0) was prepared by dissolving 108.3 g KH_2PO_4 and 35.6 g K_2HPO_4 in 1 L H_2O. Nematode Growth Medium (NGM) was prepared by dissolving NaCl (3 g), agar (17 g), and peptone (2.5 g) in 975 mL H_2O and autoclaving for 50 min. After autoclaving, the mixture was cooled in a water bath at 55 °C for 15 min, and 1 mL $CaCl_2$, 1 mL cholesterol in ethanol (5 mg/mL), 1 mL $MgSO_4$ (1 M), and 25 mL potassium phosphate buffer (1 M) were added.

Table 2.17.: Chemicals used for *C.elegans* cultivation.

Reagent	Vendor
NaCl	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Agar	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Bacto^{TM} Peptone	Thermo Fisher Scientific, Waltham, USA
Cholesterol	VWR Chemicals, Darmstadt, DE
Ethanol	Carl Roth GmbH & Co. KG, Karlsruhe, DE
KH_2PO_4	Carl Roth GmbH & Co. KG, Karlsruhe, DE
K_2HPO_4	Carl Roth GmbH & Co. KG, Karlsruhe, DE
$MgSO_4$	AppliChem GmbH, Darmstadt, DE
NaOH	Chemapol Pharma, Prag, CZ
NaOCl solution 12 % Cl	Carl Roth GmbH & Co. KG, Karlsruhe, DE
Sucrose	Carl Roth GmbH & Co. KG, Karlsruhe, DE

2.2.2.2. Cultivation

C. elegans were cultured at 22 °C in 90 mm petri dishes on NGM agar seeded with *E. coli* OP50 as a food source. Synchronous populations of the worms were obtained by treating gravid worms for 5 min with 5 M NaOH and 5 % NaOCl for lysis and decontamination. The lysate was pelleted by centrifugation (1200 rpm, 2 min). The eggs were separated from the pellet by density gradient centrifugation (300 rpm, 1 min) using a sucrose solution (60 % w/v) and water in the ratio 1:1. The upper layer, in which the eggs float, was collected and transferred to a fresh tube with sterile water (1:3) and subsequently centrifuged (350 rpm, 5 min) to pellet the eggs and wash out the sucrose. The eggs were allowed to hatch in M9 buffer.

2.2.2.3. Reproduction assay

A 24-well plate was prepared with NGM agar, and 20 µL of an *E. coli* OP50 solution was pipetted in every well. The plate was incubated at 37 °C for 24 h and subsequently treated with either 10 % DMSO (negative control) or different concentrations of AEST (**1**) in DMSO. After adding the test compounds, the plate was incubated for another hour. Three age synchronized L4 larvae were placed in every well. After 72 h, the total number of healthy adult worms were counted. IC$_{50}$ was determined by non-linear regression analysis using GraphPad Prism 6 (Dotmatics, Version 6.01).

2.3. Computational methods

2.3.1. Feature-based GNPS molecular networking analysis

The raw data was converted to the mzXML format using the software MSConvert (version 3.0.21026-c1c9c6b09).[100] The converted data was processed with the software MZmine (version 2.53), following the online workflow:

https://ccms-ucsd.github.io/GNPSDocumentationfeature basedmolecularnetworking-with-mzmine2/.[101][102][103][104][105]

Data analysis used default parameters, except for the noise level (1 x E04), group intensity threshold (1 x E05), min. highest intensity (1 x E05), m/z tolerance (0.005 m/z), and RT range for MS/MS scan pairing (0.2 min). The mass spectral network was assembled and visualized using Cytoscape (version 3.8.2).[106]

2.3.2. Classical GNPS molecular networking analysis

The raw data was converted to the mzXML format using the software MSConvert (version 3.0.21026-c1c9c6b09).[100] The converted data was processed, following the online workflow:

https://ccms-ucsd.github.io/GNPSDocumentation/networking/.[101][102][103][104][105]

Data analysis used default parameters, except for precursor ion mass tolerance (0.005 m/z) and fragment ion mass tolerance (0.005 m/z). The mass spectral network was assembled and visualized using Cytoscape (version 3.8.2).[106]

2.3.3. Semiquantitative analysis and statistical evaluation

The peak areas of AETX from colonialized *H. verticillata* and *J. americana* samples were calculated using the software FreeStyle 1.6 (Thermo Fisher Scientific, Waltham, USA). The data were divided by the amount of analyzed sample and submitted to statistical evaluation. Statistical single factorial ANOVA analysis was performed using the software GraphPad Prism 6 (Dotmatics, Version 6.01). The alpha level was set to 0.05, and a Turkey's multiple comparison test was used to compare the reservoirs against each other.

3. Results and discussion

3.1. Discovery of the unknown source of cytotoxicity

Recently, the cyanobacterium *A. hydrillicola* raised the attention of the scientific community, as it could be identified to produce a novel neurotoxin (AETX), which was responsible for the death of bald eagles in the southeastern United States.[96] In a first screening of the crude extract of *A. hydrillicola*, strong cytotoxicity could be observed, which could not be linked to AETX. This raised the question: Is there, beside AETX, another cytotoxic metabolite produced by *A. hydrillicola*?

3.1.1. Testing and characterization of the crude extract of *A. hydrillicola*

In order to confirm that the extract of *A. hydrillicola* is cytotoxic, extracts of the cyanobacterial biomass were generated. Since the polarity of the cytotoxic compound was unknown, the lyophilized biomass was extracted with increasing concentrations of MeOH in water, starting with 50 % MeOH and ending with 100 % MeOH. The obtained extract was a green, viscous solid. For initial bioactivity screening, the extract was dissolved in DMSO to a concentration of 10 mg/mL. The extract was screened via HPLC-DAD, showing few peaks beside the AETX, that was already described from this strain (**Fig. A.2.1**). The SRB colorimetric assay was chosen to assess the cytotoxicity of the extract. As a suitable cell line, HeLa cells were chosen. In an initial screening, the extract was tested, starting from 1 µg/ml as lowest concentration. However, it was found that even at this low concentration, the extract showed high cytotoxicity, demonstrating the potency of the extract. In a second assay, the lowest concentration was reduced to 10 pg/ml. From this assay, an EC_{50} of 0.12 µg/mL could be calculated. The obtained EC_{50} was compared to the corresponding one of pure AETX (0.65 µg/mL), showing that the crude extract was around five times more active than what could be expected, if the activity came from AETX alone (**Fig. 3.1**). After confirmation of the extract's cytotoxicity, identification of the responsible compound(s) was the next step.

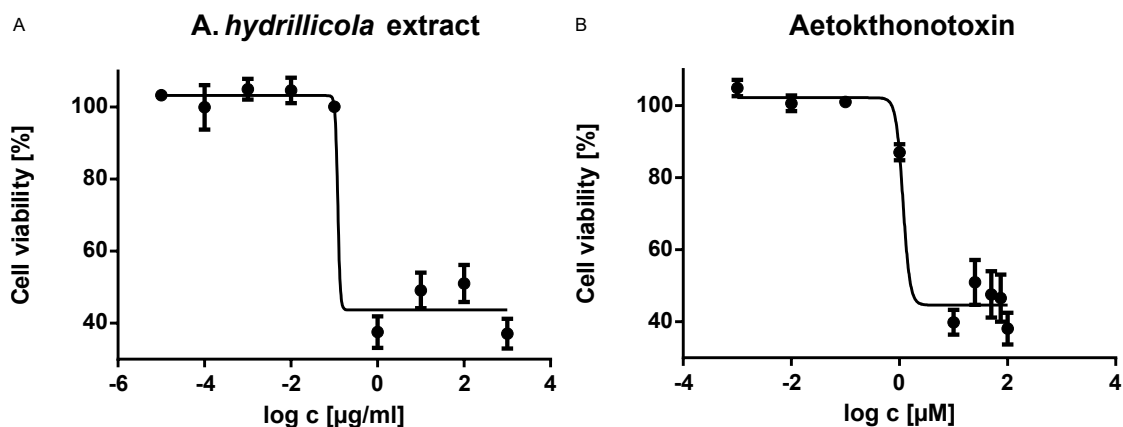**Figure 3.1.:** (**A, B**) Cytotoxicity of an *A. hydrillicola* extract (**A**, EC_{50} 0.12 µg/mL) and pure AETX (**B**, EC_{50} 1 µM; corresponds to about 0.65 µg/mL).
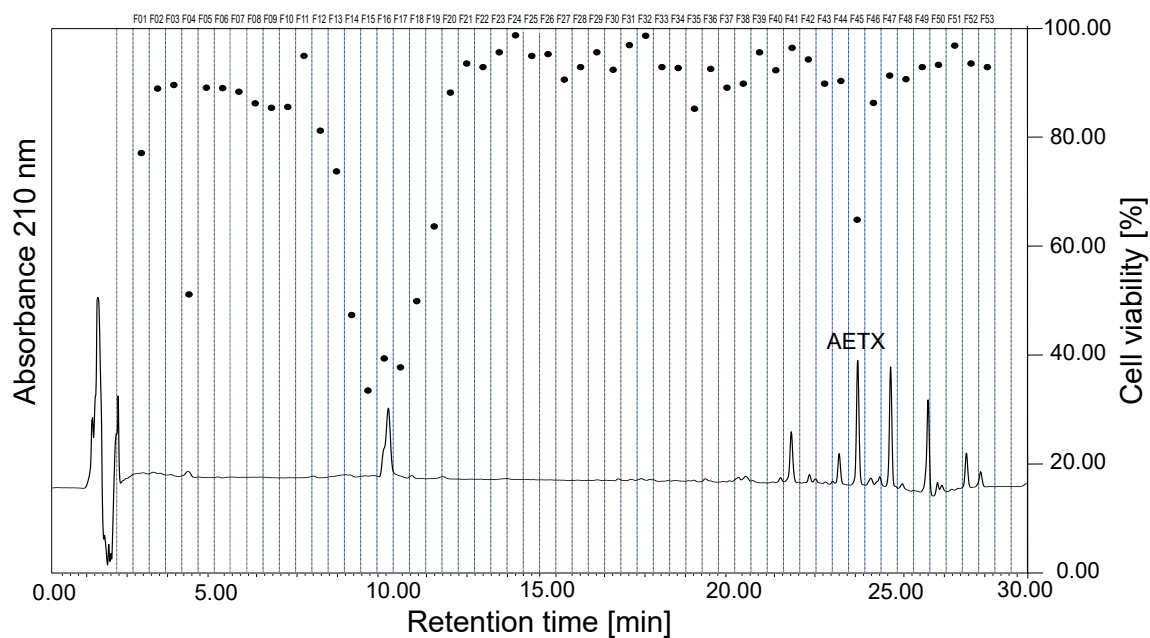

Figure 3.2.: HPLC-UV chromatogram (210 nm) of an *A. hydrillicola* extract (gradient 5-100 % MeCN in water + 0.1 % FA each) on a Luna C_{18} column (4.6 mm ID x 150 mm). Cell viability, in comparison to negative control, of each fraction is represented by a black dot.

3.1.2. Microfractionation and identification of the cytotoxic compound(s)

For rapid identification of the cytotoxic compound(s), a microfractionation approach was chosen. The compounds in the extract were separated by HPLC and collected in a time-based manner, resulting in 53 fractions. All fractions were subsequently tested in the SRB cytotoxicity assay. The results are shown in **Fig. 3.2**. The compounds eluting with a retention time of approx. 10 min, could clearly be identified as the main cause of cytotoxicity in this extract. From the observed peak shape, it already could be concluded that it was not a single molecule, but several compounds eluting closely together. The cell viability was lower compared to AETX, demonstrating once again the potency of the unknown molecules. Also, the minor compound eluting at t_r 4 min was suggested to contribute to the extract's cytotoxicity. The active fractions were analyzed by HPLC-MS, revealing the molecular mass of the main compound ([M+H]+ at *m/z* 774.5726) and its potential sum formula ($C_{43}H_{76}O_7N_5$). As this mass or sum formula did not match to any natural product in the investigated databases (Dictionary of Natural Products, SciFinder) the isolation of the compound was approached.[107][108]

3.1.3. Pre-purification by flash chromatography

A first crude purification of the extract was achieved by flash chromatography. The separation media in flash chromatography are usually based on silica particles, allowing for either normal or reversed-phase separation. For this work, a reversed-phase medium was used for separation, as it is suitable for compounds with a broad range of polarity and allows good sample recovery. The extract was prepared as a dry load (**2.1.9**) and filled into a pre-column. Separation was achieved with a linear gradient from 5 % MeCN in water to 100 % MeCN over 48 min (**Tab. 2.11**). Every 2 min, fractions were collected. The obtained fractions were dried under reduced pressure, redissolved in MeOH, and submitted to HPLC-DAD analysis. Results indicated that the separation of the main cytotoxins from most of the other metabolites could be achieved (**Fig. A.2.2-A.2.3**). Only minor impurities, which seemed to be much more lipophilic, could be detected. After HPLC-DAD analysis, fractions with the same composition were combined and submitted to semi-preparative HPLC.

3.1.4. Compound(s) isolation by semi-preparative HPLC

The combined flash fractions, containing the compounds of interest, were dried, redissolved in MeOH to a concentration of 10 mg/ml, and filtered (0.45 μm pore size). As a starting point, a column with C_{18} column media, and a gradient of 5-100 % MeCN was chosen, which was optimized in several steps afterwards. Finally, a gradient of 20–50 % MeCN in water over 23 min was chosen, which resulted in decent separation (**Fig. 3.3**).
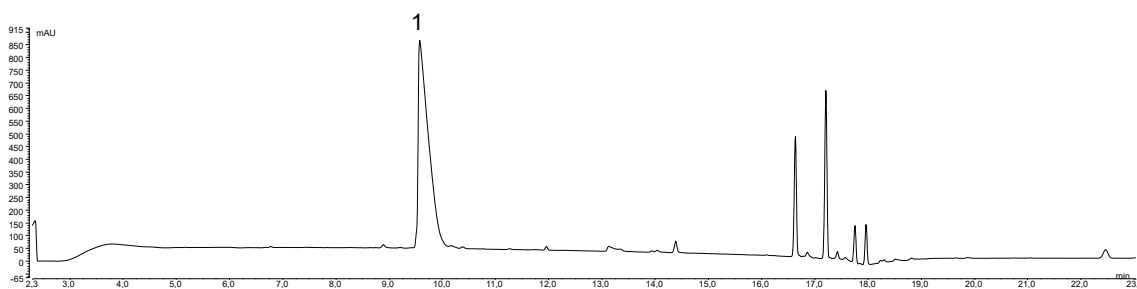
3. Results and discussion

Figure 3.3.: HPLC-UV chromatogram (210 nm) of the pooled flash fractions (gradient 20–50 % MeCN in water + 0.1 % FA each) on a Luna C$_{18}$ column (4.6 mm ID x 250 mm). Compound (**1**) was collected via automatic fraction collection.

The main peak (**1**) was isolated in several runs and submitted to HPLC-MS analysis for a purity check. Unfortunately, the closely eluting compounds could not be separated under these conditions. HPLC-MS analysis revealed that the isolated peak consists of two compounds, which had a mass difference of 14 m/z, indicating that the two compounds only differ in a methyl group, which made the separation challenging (**Fig. A.2.9**).

First, separation via Sephadex LH-20 gravity column chromatography was approached. The sample was eluted with 30 ml/h over 24 h, resulting in 74 fractions. The collected fractions were analyzed in HPLC-MS. Unfortunately, the separation was not successful (**Fig. 3.4**).
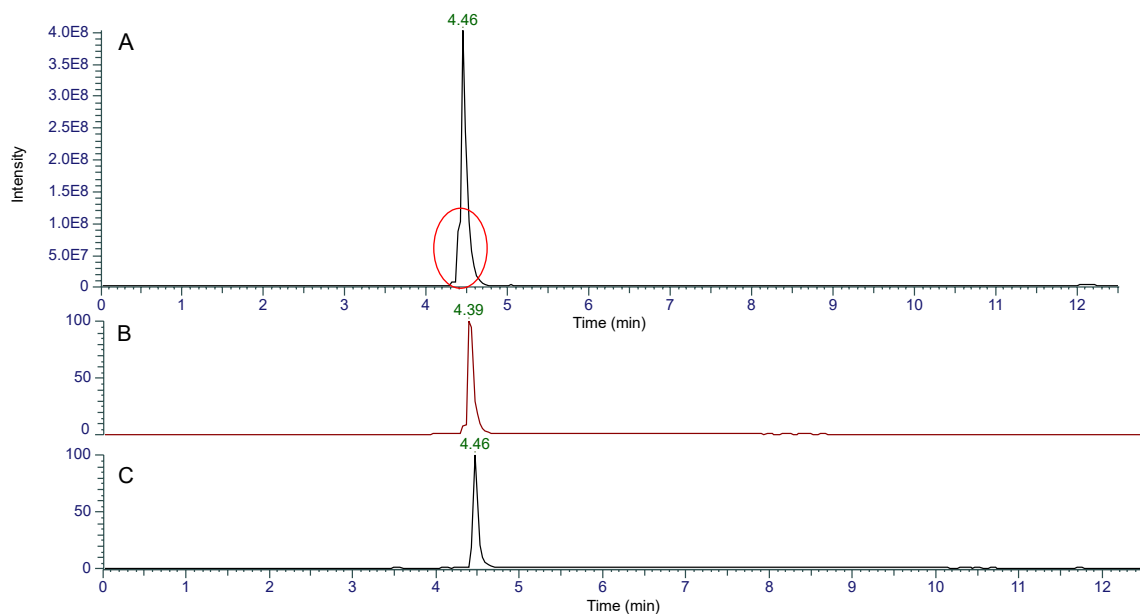
Figure 3.4.: (**A**) TIC of a fraction obtained from Sephadex-LH20 chromatography (**B**) EIC of m/z 774.57 (t$_r$ 4.46 min) (**C**) EIC of m/z 760.55 (t$_r$ 4.39 min).

The "shoulder" in front of the main peak was still present, and the mass of both compounds could be found in the analyzed fractions. As Sephadex gel chromatography failed in separating the two molecules, HPLC was tried next. As the structure of the desired compound was still unknown, in a first step different separation media and eluents were screened. Nowadays, several column media for reversed-phase HPLC are available, offering different interaction characteristics, and thus more options in separation. As a starting point, C_{18} material was chosen and compared against pentafluorophenyl (PFP) and phenylhexyl (PH) column material. In addition, MeOH and MeCN were tested as organic solvents. The results are illustrated in **Fig. A.2.4-A.2.8**. No column media was initially successful in separation, but in the chromatogram of the C_{18} column with MeCN, as well as in the PFP column, a small fronting shoulder could be detected. The switch from MeCN to MeOH brought no advantage, as only an increase in retention time was observed. Thus, the C_{18} column in combination with MeCN was chosen, as it provided the best peak shape and a decent retention time. In a next step the gradient was optimized, but attempts to achieve baseline separation were futile. The best separation can be seen in **Fig. 3.5** (**A**). As baseline separation could not be achieved by gradient optimization, the temperature was varied next. The temperature was increased from 30 °C up to 60 °C. Indeed, from 50 °C on, a baseline separation could be achieved (**Fig. 3.5**).
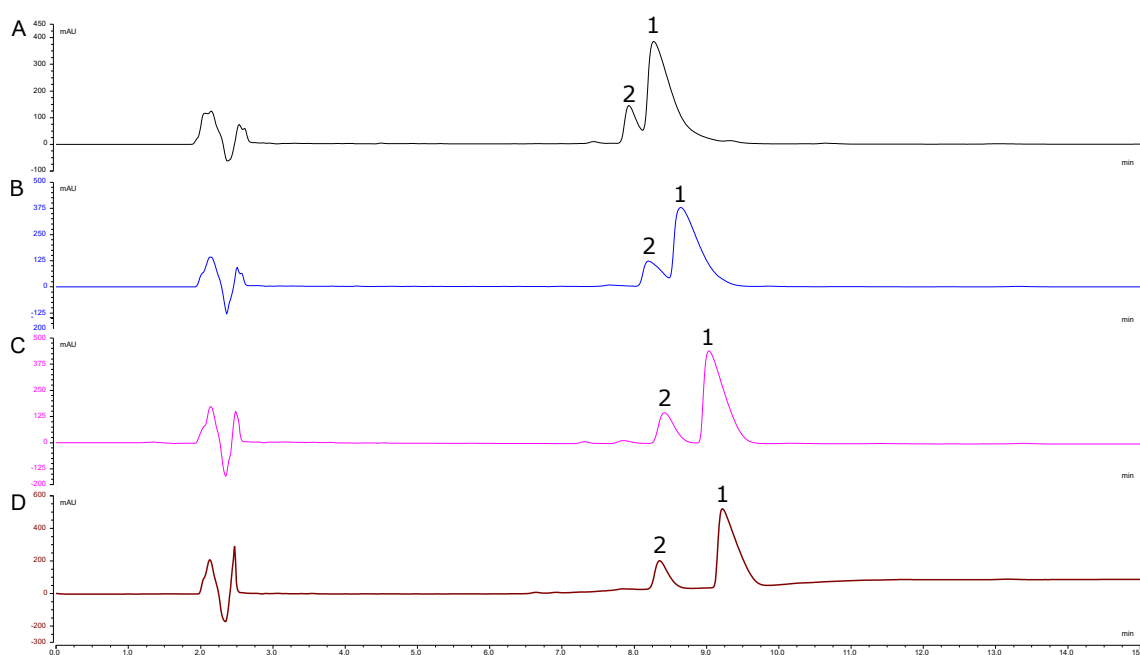

Figure 3.5.: HPLC-UV chromatograms (210 nm) of pooled fractions (gradient 20-50 % MeCN in water + 0.1 % FA each) on a Luna C_{18} column (4.6 mm ID x 250 mm). (**A**) Temperature: 30 °C (**B**) Temperature: 40 °C (**C**) Temperature: 50 °C (**D**) Temperature: 60 °C. Cytotoxic compound (**1**) and derivative (**2**).

3. Results and discussion

At 60 °C, the best separation could be seen, but the two compounds still eluted quite closely. In addition, a degradation of the compound at this temperature could not be excluded, thus the search for a method with less harsh conditions continued. As a final parameter, the pH of the eluents was modified. So far, all separations were carried out under acidic pH conditions, as the eluents were modified with 0.1 % formic acid (FA). Thus, the pH was increased using an ammonium acetate buffer system. First, the pH was increased to ~ 5 and finally to ~ 9. Interestingly, the retention changed dramatically. The compounds eluted at 80 % instead of 35 % of organic solvent content in the mobile phase, and also a better separation with a better peak shape could be achieved (**Fig. A.2.11**). The gradient was further optimized under these conditions, which resulted in a satisfying separation of the two compounds (**Fig. 3.6**). Now, with a suitable method in hand, the two compounds were isolated. Purity was checked via HPLC-MS and HPLC-DAD. The purity was satisfying, so that the solvents were removed using a rotary evaporator. For final drying, and to remove the ammonium acetate buffer completely, the samples were frozen and lyophilized for 24 h.

Finally, from 30.8 g of cyanobacterial biomass, 46 mg (yield: 0.15 %) of the main compound (**1**) and 9 mg (yield: 0.03 %) of the second compound (**2**), carrying a methyl group less, could be isolated.
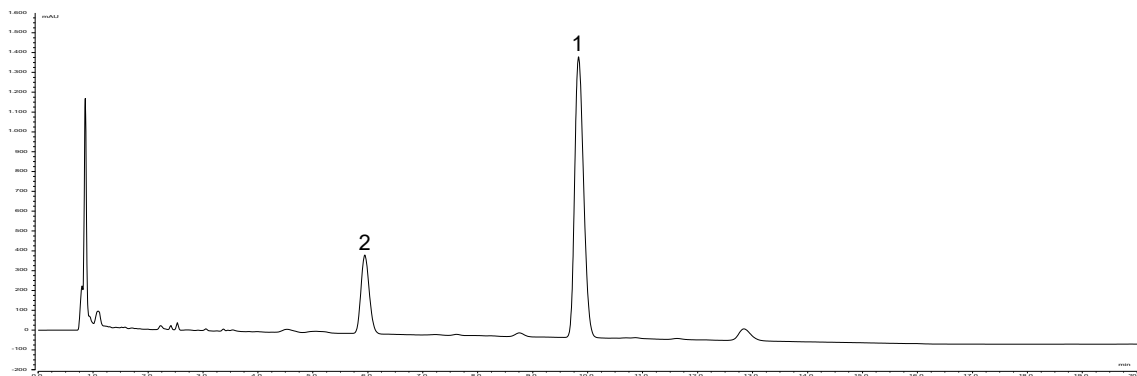

Figure 3.6.: HPLC-UV chromatogram (210 nm) of the optimized separation (gradient 50–85 % MeCN in water at pH 9 using an ammonium acetate buffer system) on a Luna C_{18} column (4.6 mm ID x 250 mm). Cytotoxic compound (**1**) and derivative (**2**).

3.1.5. NMR and MS based structure elucidation

For NMR analysis, 5 mg of each compound were prepared, and analyzed as described in materials and methods section (**2.1.14**). Significant signal overlap in the ^{1}H-NMR spectrum complicated the NMR analysis of **1** (**Fig. A.2.13**). A preliminary analysis of the ^{13}C and ^{1}H-NMR data suggested the presence of four amide or ester carbonyls (173.4, 168.8, 172.9, and 169.7 ppm) and two NH protons (7.90 and 8.02 ppm; NMR data **Tab. A.2.1**). The presence of amide bonds was confirmed by four signals consistent with α-carbonyl protons of amino acids (2.25, 2.35, 4.64 and 2.79 ppm). This gave confidence that the compound was a peptide. Signals from six N- or O-methyl groups were also observed. Evaluation of the HSQC-DEPT, TOCSY and COSY spectra revealed the typical pattern of valine and isoleucine, further supporting the hypothesis of a peptidic structure (**Fig. A.2.15-A.2.18**). The HMBC data confirmed valine and isoleucine as the two N-terminal amino acids. The third amino acid had a side chain like isoleucine, but three extra protons in its spin system (3.98 and 2.35/2.17 ppm), attached to two carbons. Evaluation of the HMBC and COSY spectra revealed correlations that indicated a CH_2 group (2.35/2.17 ppm) adjacent to the carbonyl carbon (168.8 ppm), and an O-methyl group attached to a CH (3.20 ppm), which was bonded to the carbon carrying the isoleucine side chain. This amino acid has therefore been identified as dolaisoleuine (Dil), suggesting that the compound may belong to the dolastatin compound class. With this information, the NMR data of the fourth amino acid was easily assignable, as it was consistent with dolaproine (Dap), another amino acid typical for dolastatins. However, the NMR-data of the C-terminus did not match to any known C-terminus in the dolastatin compound family. The presence of an aromatic phenyl (7.12 – 7.20 ppm) and an O-methyl group (3.32 ppm) was revealed by further analysis of ^{1}H, HMBC and HSQC spectra. Using HMBC correlations from CH (3.38 ppm) to the phenyl group and the carbonyl carbon, the structure was elucidated as a hitherto unknown monomer at the C-terminus. This C-terminal monomer was given the name aetophine (Aph). The *N,N*-dimethyl-Ile-Val-Dil-Dap-Aph monomer sequence of the compound was confirmed by ROESY data. The observed MS/MS fragmentation pattern supported this amino acid sequence (**Fig. A.2.36**). After structure elucidation, the observed signal overlap could also be explained: cis/trans isomerism between dolaproine and dolaisoleuine has been described for dolastatins before, resulting in a doubled number of signals.[109] As the compound was produced by *A. hydrillicola*, and belongs to the dolastatin compound family the name aetokthonostatin (AEST) was chosen for **1** (**Fig. 3.7**). After solving the structure of **1**, structure elucidation of **2**, lacking a methyl group, was straight forward. Comparing the MS/MS data of the two compounds already revealed a difference of - 14 *m/z* at the N-terminal *N,N*-dimethyl-Ile moiety, indicating that the nitrogen might be carrying one methyl group less (**Fig. A.2.37**). Unfortunately, the NH signal could not be observed in the NMR spectra, but all other signals

of the N-terminal amino acid were present, and the methyl group attached to the nitrogen shifted from 41.6 ppm to 34.5 ppm, which would fit to a single methylated amine (**Tab. A.2.3**). In addition, the integral of this signal in the ^{1}H-spectrum was lower than in the original spectrum of **1**. The rest of the NMR data was in consistence with the one obtained for AEST (**1**). For (**2**) the name monomethylaetokthonostatin (MMAEST) was chosen. (**Fig. 3.7**).
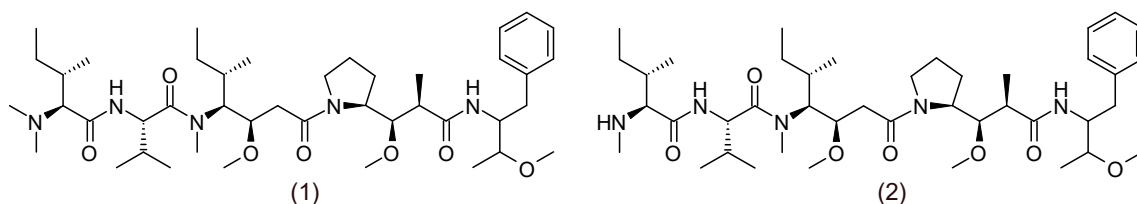

Figure 3.7.: Chemical structures of AEST (**1**) and MMAEST (**2**).

3.1.6. Stereochemistry of AEST

The elucidation of stereochemistry is still one of the most challenging processes in natural product chemistry. The gold standard method relies on X-ray crystallography, but not every molecule provides crystals, which are suitable for X-ray measurements. In this work, the stereochemistry of four monomers (*N,N*-dimethyl-Ile, Val, Dil, and Dap) of AEST (**1**) was determined using Marfey's and NMR analysis. Unfortunately, the stereochemistry of the new C-terminal moiety (Aph) could not be elucidated, as suitable crystals for X-ray measurements could not be obtained.

3.1.6.1. Marfey's analysis

Initially, Marfey's analysis was used to determine the stereochemistry of Dap, Dil and Val. For the C-terminus, no suitable standard was available, and the N-terminal *N,N*-dimethyl-Ile would not react with the derivatization agent, as tertiary amines are not suitable for Marfey's analysis. MMAE was chosen as the standard, because three of the five monomers (Val, Dil, Dap) are identical to those found in AEST (**1**). Unfortunately, Dil degraded during the hydrolysis process, as the ether moiety was not stable under the applied acidic conditions. Even shorter hydrolysis time was not able to solve this problem. Finally, the retention times of the derivatized Dap and Val were compared with those obtained from MMAE, revealing that the stereochemistry has to be identical (**Fig. 3.8-3.9**). For Dap even the two possible cis/trans isomers were detected, explaining the observation of two peaks in both spectra.
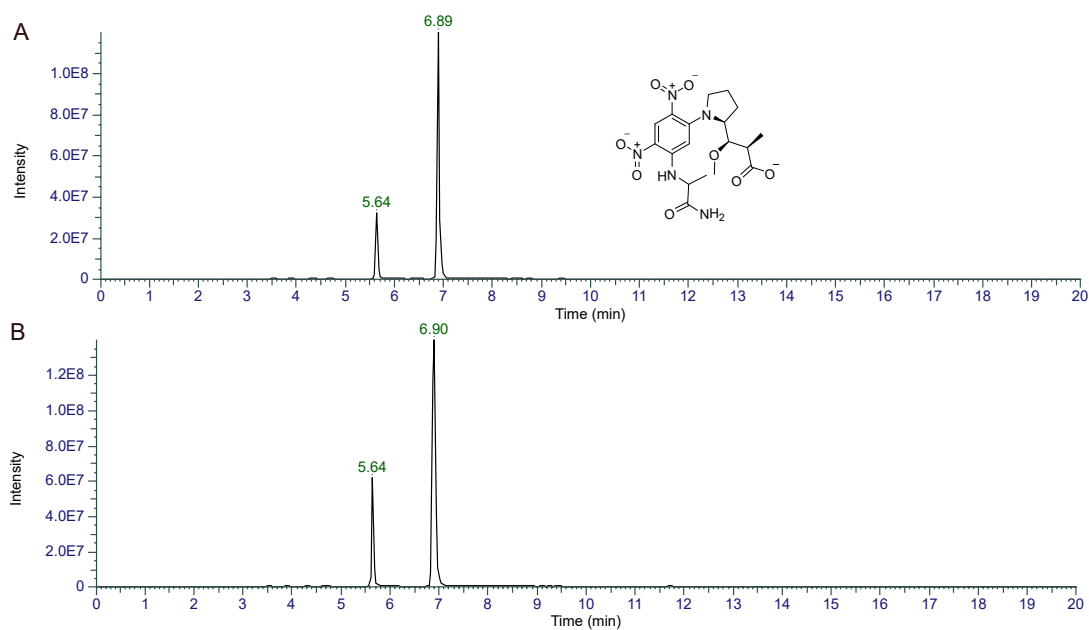
Figure 3.8.: Marfey's analysis of dolaproine. EIC of dolaproine derivatized with Marfey's reagent (m/z 438.1630, structure shown in the chromatogram). (**A**) AEST (**1**) and (**B**) MMAE.
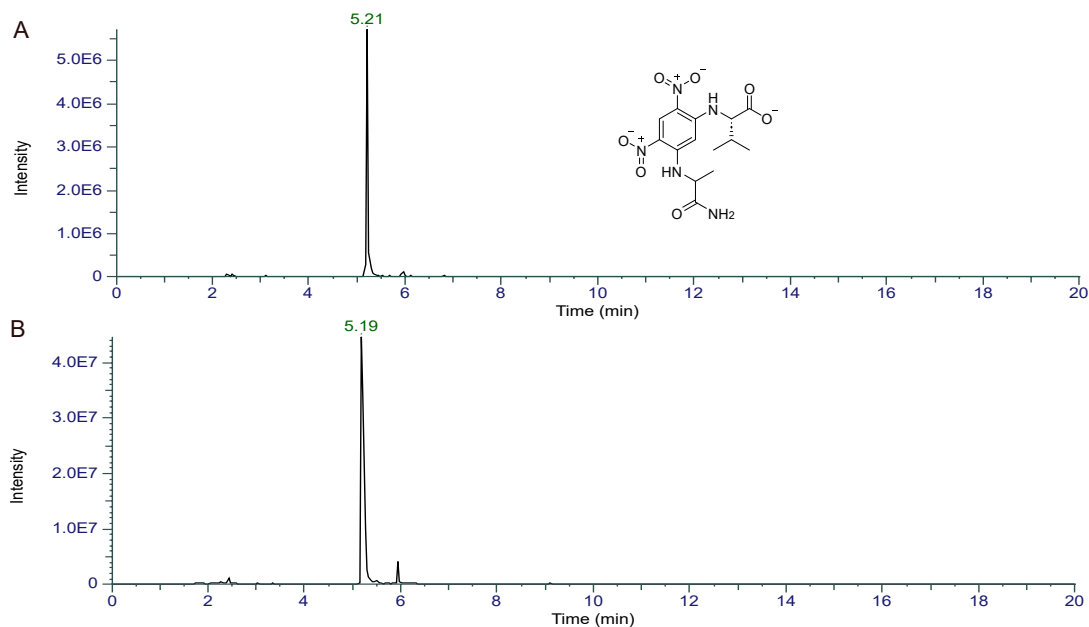

Figure 3.9.: Marfey's analysis of valine. EIC of valine derivatized with Marfey's reagent (m/z 368.1211, structure shown in the chromatogram). (**A**) AEST (**1**) and (**B**) MMAE.

3.1.6.2. NMR analysis

To elucidate the stereochemistry of the Dil and *N,N*-dimethyl-Ile moiety, ^{13}C-NMR data of AEST (**1**) and symplostatin 1, which has the same structure as AEST, except for the C-terminus, were compared. To compare the data, a ^{13}C-NMR of AEST in CD_2Cl_2 was recorded, as NMR-data of symplostatin 1 was only available in the mentioned solvent (**Fig. A.2.19**-**A.2.20**).[70] Comparing the obtained NMR-data revealed that the stereochemistry of the four monomers (Dap, Dil, Val, *N,N*-dimethyl-Ile) of AEST (**1**) and symplostatin 1 has to be identical, as only minor differences in chemical shifts were observed (**Tab. A.2.2**). These minor differences are most likely due to the different C-Termini of both molecules.

3.1.7. Feature based molecular networking

As mentioned before, in a first manual analysis of the HPLC-MS data of the *A. hydrillicola* extract, several masses of compounds potentially related to the main compound AEST (**1**), could be detected. To verify this relatedness and to discover more AEST derivates, a chemoinformatic analysis of the extract via the "Global Natural Products Social Molecular Networking" (GNPS) platform was performed.

GNPS was released in 2016 as a new data-driven platform for MS/MS data storage, analysis, and knowledge dissemination, allowing the user to build molecular networks based on the alignment of MS/MS spectra to each other.[103] Similarities of the compared MS/MS spectra are defined by a modified cosine scoring scheme, that determines the similarity of two MS/MS spectra with scores ranging from 0 (totally dissimilar) to 1 (completely identical).[103] Depending on the set threshold, (e.g. 0.7) MS/MS spectra with a cosine score above the set limit are clustered together in a molecular network. Assuming that similarities in MS/MS spectra reflecting similarities in chemical structures, this can help in identification of structurally related molecules in the analysed data. In addition, the MS/MS spectra are also searched against GNPS spectral libraries, helping in dereplication of already known molecules.[103] The resulting molecular networks can be visualized online in a browser or can be exported for use in third-party visualization software, such as Cytoscape.[106]

This general approach can be refined by a so-called "Feature based molecular networking analysis" (FBMN). For FBMN, not only the MS/MS data of compounds are evaluated, but also the retention time of a compound in the HPLC-MS analysis is considered.[104] This is especially useful if two or more compounds with the same mass are present in a sample. In classical GNPS data analysis, all MS/MS spectra of this parent mass would be combined into one spectrum, obfuscating differences in the MS/MS spectra of the compounds eluting at different times. With the addition of the retention time as discriminator

in FBMN, the individual MS/MS spectra are handled separately, allowing the discrimination of compounds with the same parent mass but different molecular structure.[104] This approach was chosen for the present study, because in the manual data analysis, peaks with the same mass but a different retention time were observed. The detailed procedure and the chosen parameters are described in the materials and methods section (**2.3.1**). FBMN analysis revealed the presence of AEST related compounds in the *A. hydrillicola* extract (**Fig. 3.10**). Most of the detected AEST derivatives were also pentapeptides, but tetrapeptides could be observed as well. The second most abundant compound was the already isolated MMAEST (**2**). In addition to that, three more derivatives with one methyl group less were revealed. The structure of these three compounds and most of the other observed derivatives could be deduced based on their MS/MS data (**Fig. A.2.41-A.2.56**). Adding to the four monomethyl-derivatives, four di-desmethyl-derivatives (m/z 746.5413) were detected, and one derivative with a methyl group more (m/z 788.5893). Among the revealed tetrapeptides, the one at m/z 613.4524 was the most abounded one. Evaluation of its MS/MS data showed that this compound is a tetrapeptide lacking the C-terminus. The other tetrapeptides could be identified as mono and di-desmethylated versions of the major tetrapeptide. Based on the abundance of the observed ions, it was decided to isolate the tetrapeptide (m/z 613.4524) and the derivative at m/z 790.5676.
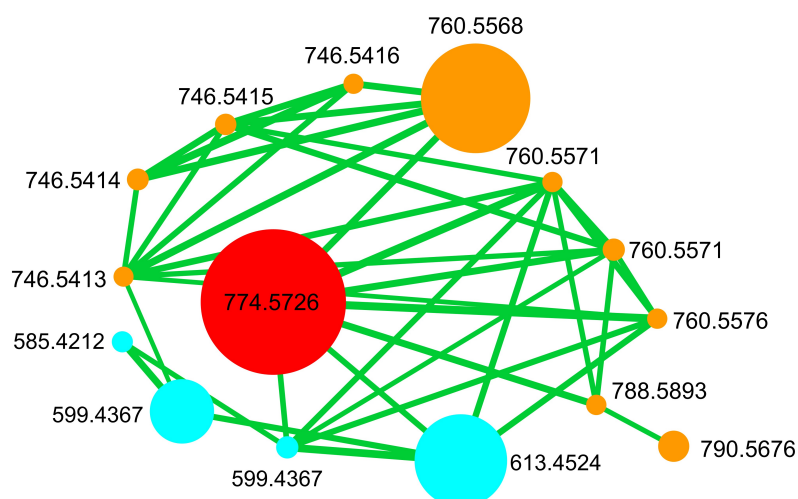

Figure 3.10.: Feature-based molecular network of AEST related compounds. Size of the nodes correlates with the abundance of the detected molecular ion. Red–AEST (**1**); orange–pentapeptides; cyan–tetrapeptides.

3.1.8. Isolation and structure elucidation of AEST-derivatives

For the isolation of the tetrapeptide, the flash chromatography fractions containing the compound were used. The tetrapeptide was isolated by the diploma student Kristin Röhrborn under my supervision.[110] For isolation, a linear gradient from 10–50 % MeCN in water was used (**Fig. A.2.10**). In contrast to AEST (**1**), sufficient separation was achieved without basic pH modification of the mobile phase. The purity of the isolated compound was checked via HPLC-MS, and the compound was analysed by NMR spectroscopy. Structure elucidation was straight forward, as the observed signals matched with the structure that was predicted based on the MS/MS data (**Fig. A.2.38**). It was obvious from the NMR data that, in comparison to AEST (**1**), the C-terminus was missing, as no proton signals were observable in the aromatic region. The remaining signals were in consistence with the NMR-data of AEST (**1**) (**Tab. A.2.4** and **Fig. A.2.27-A.2.30**). As the compound, compared to AEST (**1**), lacks the C-terminal monomer Aph, it was named des-Aph-aetokthonostatin (DAAEST (**3**), **Fig. 3.11**).

Next, the derivative with the molecular weight of 789.5616 Da was isolated from the respective flash fractions. For this compound, the pH of the mobile phase was modified again with ammonium acetate to a pH of ~ 8, and a linear gradient from 50–85 % MeCN was used (**Fig. A.2.12**). As the retention time of this compound, compared to AEST (**1**), was slightly higher, the compound had to be more lipophilic. After isolation, purity was confirmed via HPLC-MS, and the compound was submitted to NMR spectroscopy. Based on the obtained MS/MS data, no structure could be postulated. The NMR data of the four C-terminal amino acids was identical with those observed for AEST (**1**) (**Tab. A.2.5**), indicating that the difference must be in the N-terminal amino acid. Several signals characteristic for *N,N*-dimethyl-Ile were missing, and a new signal at 5.85 ppm was observed in the ^{1}H-NMR spectrum. In addition, two new carbons signals with a relatively high chemical shift of 117.8 and 153.9 ppm appeared, suggesting a double bond to be present in the N-terminal monomer of this compound. This agreed with the new signal in the ^{1}H-spectrum, which, according to HSQC data, was found to be attached to the carbon at δ_C 117.3 ppm. After analysis of the NMR data, the structure shown in **Fig. 3.11** could be proposed. Unfortunately, this proposed structure was not agreeing with the obtained MS/MS data of the molecule, because the calculated molecular mass of this compound was only 728.5088 Da, 61.0528 Da less than it should have, according to the initial MS analysis. Interestingly, a neutral loss with exactly this mass difference could be seen in the MS/MS data of the parent ion at 790.5676 *m/z*, indicating that the molecule might be not stable and could lose this moiety also during storage of the compound. From the accurate mass, a molecular formula of C_2H_7NO (Δ 1.39 ppm) could be calculated for this neutral loss. To confirm the hypothesis that the compound had degraded, the compound was recovered from the NMR tube, and again analysed via HPLC-MS. In fact, the molec-

ular mass of the molecule was found to be 728.5070 Da, which now matched well with the structure proposed after evaluation of the NMR data (**Fig. A.2.40**). This proved that the molecule was decomposed between the first HPLC-MS analysis and the NMR measurement. As it was not possible to deduce the complete structure from the MS/MS data, and the compound decomposed before NMR analysis, the structure could not be solved unambiguously. Due to time constraints, the compound was not reisolated for new NMR measurements. The clarification of this structure has to be addressed in future studies.
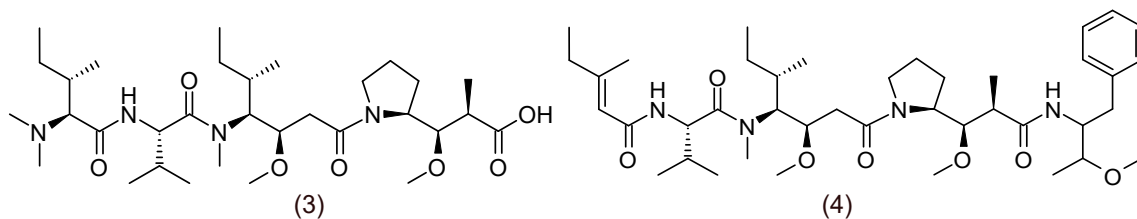
Figure 3.11.: Chemical structures of DAAEST (**3**) and the degraded-AEST derivative (**4**).

3.1.9. Cytotoxicity testing

As the main cytotoxic compound in the *A. hydrillicola* extract was found to be structurally closely related to the dolastatins, the extract's cytotoxicity could be explained. To experimentally confirm the cytotoxicity of the isolated compounds, the same SRB cytotoxicity assay, that has initially been used for testing the crude extract and the fractions, has been utilized.

First, AEST (**1**) was tested in a concentration from 1 µM to 1 pM. This range of concentration was chosen based on the IC_{50} of Dol-10 and MMAE, which are among the most cytotoxic molecules from this compound class, with an IC_{50} around 1–10 nM.[64][111] Indeed, it could be demonstrated that AEST (**1**) is equally toxic, having an IC_{50} of 1 nM (**Fig. 3.12**). With that result in hand, it was clear that AEST (**1**) is the source of the highly cytotoxic activity, which was found in the initial screening of the crude extract.

Next, the cytotoxicity of AEST (**1**) was compared with the cytotoxicity of MMAE and MMAF. As discussed in the introduction, MMAE and MMAF are already on the market as payloads in ADCs used in cancer therapy. AEST (**1**) was found to be equally toxic as the clinically used MMAE, which had an IC_{50} of 3 nM in my assay, demonstrating the potency of the isolated compound. In contrast, MMAF ($IC_{50} = 17$ nM) was around 17x times less active, compared to AEST (**1**). This is due to the fact that MMAF contains a free carboxylic group at the C-terminus, which is charged under physiological pH. With this negatively charged carboxylic acid, the penetration of the compound into the cell is decreased, explaining the lower potency. Finally, the two isolated AEST derivatives MMAEST (**2**) and DAAEST (**3**) were also tested. It could be demonstrated that a methyl group less at the N-terminal nitrogen has little influence on the cytotoxicity, as the IC_{50}

3. Results and discussion

of MMAEST (4 nM) was close to the one observed for AEST (**1**). This was not surprising, as structure-activity-relationships (SAR) of Dolastatin 10 derivatives were extensively studied in the past, indicating the absence of one methyl group at the N-terminus did not have an impact on the cytotoxicity.[77]

To test if the C-terminal moiety is necessary for activity, DDAEST (**3**) was tested. With an IC_{50} of 25 nM, DAAEST (**3**) was less active than AEST (**1**), but still on a similar level as MMAF. A key factor is that DAAEST (**3**) also has a free carboxylic acid that adds a negative charge to the molecule, resulting in reduced cell penetration. Interestingly, this result is not in total agreement with the literature, because previous SAR studies were indicating that the benzene moiety at the C-terminus is necessary for cytotoxic activity.[77] However, a following study by Pettit *et al.* demonstrated that a methylester of an amino-terminal tetrapeptide of dolastatin 10 was around 30-fold times less cytotoxic than dolastatin 10 itself, even though both compounds were interacting comparable with the molecular target.[112] These differences in potency are matching well with the observed ones for AEST (**1**) and DDAEST (**3**). These results suggest that the lower activity of DAAEST (**3**) might not only based on the reduced uptake.

Taking the observed activity of DAAEST (**3**) and the results from Pettit *et. al* into account, the C-terminal benzene moiety seems not to be necessary for cytotoxic activity.
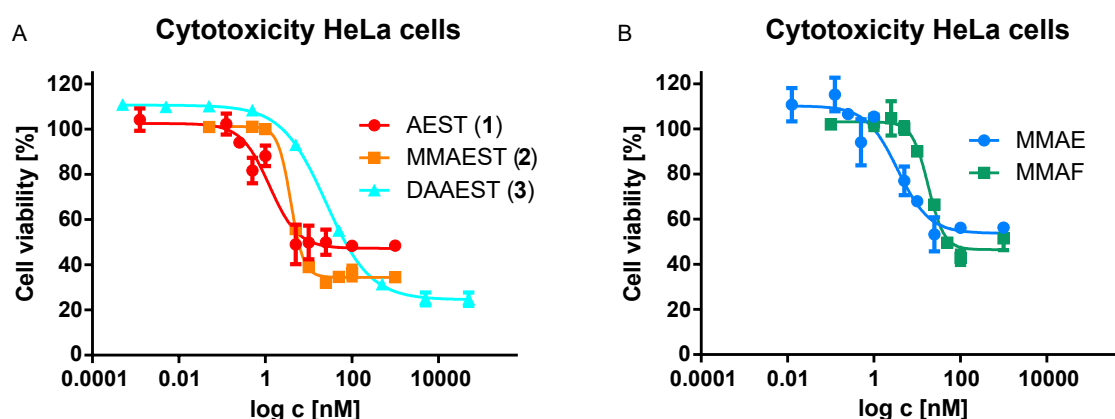

Figure 3.12.: (**A**) Cytotoxicity of AEST (**1**, EC_{50} 1 $\pm$ 0.2 nM), MMAEST (**2**, EC_{50} 4 $\pm$ 0.01 nM), DAAEST (**3**, EC_{50} 25 $\pm$ 1.4 nM) on HeLa cells (n = 3) (**B**) Cytotoxicity of MMAE (EC_{50} 3 $\pm$ 0.4 nM), and MMAF (EC_{50} 170 $\pm$ 30 nM) on HeLa cells (n = 3). Data are represented as average $\pm$ SEM.

3.1.10. *Caenorhabditis elegans*

The discovery of a second highly toxic compound in the *A. hydrillicola* extract raised the question whether the two toxins AETX and AEST (**1**) might act additive or synergistically. To answer this question, *Caenorhabditis elegans* was chosen as a model organism. *C. elegans* is a small, free-living nematode, which occurs worldwide. The adult worms are about 1 mm long, which make it easy to observe them via a standard microscope. Furthermore, *C. elegans* is a self-fertilizing hermaphrodite with a rapid life cycle (3 days at 25 °C from egg-laying adult), which consists of four different larvae stages (L1–L4). In addition, *C. elegans* offers a handful of advantages as a model organism: it has a small size, large brood size, is easy to cultivate and to maintain, long-term cryopreservable, has a quick generation time, and transparency.[113] Also, the similarities between the cellular and molecular processes in *C. elegans* and other animals across evolutionary time have made *C. elegans* an excellent model organism. Metabolism, organelle structure and function, gene regulation, and protein biology are research topics that can be addressed by studying *C. elegans*. 60–80 % of human genes have an ortholog in the genome of *C. elegans*, and 40 % of genes known to relate to human diseases have orthologs in the *C. elegans* genome, making the nematode also interesting for studying human health and diseases.[113]

AETX has already been tested on *C. elegans* before, showing acute toxicity with an LD_{50} of 38 nM.[96] The first questions to be answered was, whether AEST (**1**) is also toxic to *C. elegans*. A defined number of worms were treated with different concentrations of AEST (**1**) (56 µM–5.6 nM), and after 24 h the number of healthy worms were counted. In contrast to AETX, no acute toxicity could be observed, which was consistent with previous reports on the toxicity of Dol-10 to *C. elegans*.[114] To study if the treatment with AEST (**1**) had any long-term effects on the worms, they were cultivated two days more. Indeed, a strong decrease in the reproduction rate could be observed. As mentioned above, *C. elegans* has a rapid life cycle (3 days). In the control group (treated with 10 % DMSO in water), numerous new worms had developed after three days of cultivation. In the AEST (**1**) treated group, none or only a few worms, which were also smaller than normal, had developed over the same time. To verify these results, an assay was designed, where three L4 larvae were treated with different concentrations of AEST (**1**) and AETX. After 72 h of incubation, the number of healthy-looking worms were counted. The results of the treatment with AEST (**1**) can be seen in **Fig. 3.13**. A strong reduction in reproduction can be seen up to a concentration of 1 µM. In the AETX-treated culture, its acute toxicity, as reported before, could be verified. However, it was observed that in sub-lethal concentrations, the surviving worms were developing similar to those in the negative control, resulting in a healthy colony in the longer term.

Summarizing these results, it could be demonstrated that *A. hydrillicola* is producing two harmful compounds (AETX and AEST), where one compound (AETX) shows acute

toxicity, while the other (AEST) shows a more chronic form of toxicity (decrease in reproduction rate) to the nematode *C. elegans*. These results point out that the occurrence of *A. hydrillicola* in water bodies, used for recreational purposes, or as drinking water, should be monitored closely in the future, as the production of two toxic compounds, with different types of toxicities, could be confirmed.
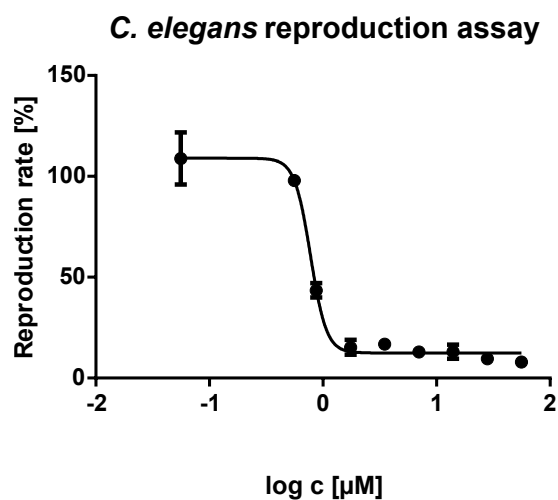

Figure 3.13.: Effect on the reproduction rate (EC$_{50}$ 0.8 ± 0.2 µM) of *C. elegans* treated with different concentration of AEST (**1**).

3.2. Discovery of AETX related compounds

While manually analyzing the *A. hydrillicola* extract for AEST derivatives, the presence of derivatives of the neurotoxin AETX was also noted. Thus, I became also interested in their structures. As most of these metabolites had lower molecular masses than AETX, I hypothesized that they might be biosynthesis intermediates of AETX. An analysis using the software tool haloseeker, which allows the user to detect halogenated molecules in mass spectrometry data, by S. Breinlinger, already had revealed the presence of several single indoles and AETX derivatives with different halogenation pattern, also including iodine instead of bromine.[115][116] As the amounts of these derivatives in the extract were quite low, their isolation has not been achieved, yet. In this work, a GNPS based approach, which resulted in the discovery of additional derivatives, was used. Based on these findings, the isolation and synthesis of one of the AETX indole building blocks was achieved, which finally supported the total synthesis of AETX in cooperation with our partners at the Leibniz Institute for Plant Biochemistry.

3.2.1. GNPS analysis

In order to identify derivatives of AETX, a crude extract of *A. hydrillicola* was analyzed via the platform GNPS (**Fig. 3.14**). This time, a classic networking analysis was chosen, as it was not necessary to include further parameters, as for example the retention time. The analysis revealed eight compounds clustering with AETX. Three of these were already discovered by S. Breinlinger, including both single indole moieties (at m/z 298.864 and m/z 351.779) and the single debrominated AETX (m/z 571.725).[115] The other four detected metabolites (m/z 495.804, 628.636, 667.629, and 704.549) were so far unknown. Possible structures could be hypothesized based on their MS/MS data (**Fig. A.3.1**). These derivatives mainly differ in the number of bromine atoms and the absence of the nitrile moiety.

3.2.2. Isolation

To draw conclusions about the possible biosynthesis of AETX, the compound with the m/z 298.864 was chosen for isolation, as it represents a single indole moiety, which could be one of the key building blocks of AETX biosynthesis. For isolation, the fractions, which were acquired by flash chromatography, were screened via HPLC-MS. The compound could be located in fractions 50-53. These fractions were combined and subjected to semi-preparative HPLC. The compound was isolated in cooperation with my diploma student K. Röhrborn.[110] Isolation was convenient, as a satisfactory separation could be achieved under isocratic conditions of 55 % MeCN in water (**Fig. A.3.9**). After evaporation of the solvents and lyophilization, the yield was, with 2.8 mg, suitable for NMR analysis.

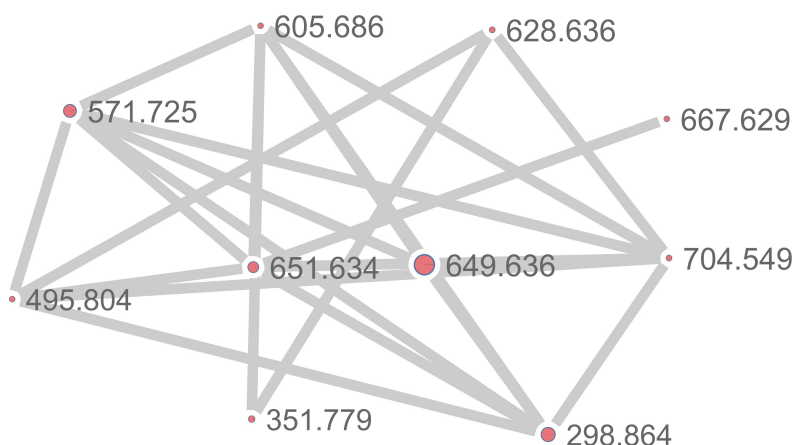**Figure 3.14.:** Classical molecular network of AETX and related compounds. Size of the node correlates with the abundance of the detected ions.

3.2.3. Structure elucidation

A potential structure of the isolated compound (**5**) could already be assumed by analyzing the MS data, as the sum formula ($C_9H_4N_2Br_2$) and the observed fragments in the MS/MS spectrum matched with the carbonitrile indole moiety of AETX (**Fig. A.3.10**). To confirm the postulated structure, the compound was analyzed by NMR spectroscopy. The ^{13}C-NMR spectrum confirmed the number of carbon atoms (9), and the chemical shifts were in accordance with the published NMR data of the western indole of AETX (**Fig. A.3.2-A.3.6** and **Tab. A.3.1**). Also, the chemical shifts in the ^{1}H-NMR spectrum were consistent with the data of the respective indole in AETX. One new signal was observed in the proton spectrum, which could be assigned to a hydrogen connected to the carbon in position 2, where in AETX the eastern indole moiety is attached. Two-dimensional NMR spectra supported this hypothesis, as a correlation between the new signal and the indole NH hydrogen could be observed in the COSY spectrum, indicating that these two hydrogen atoms were adjacent. The ^{1}H signals belonging to the aromatic hydrogen atoms at 7.85 and 7.72 ppm were splitting up in doublets, which was surprising at first, as they should give just a singular signal. Measuring the coupling constants (1.71 Hz) revealed that this coupling was due to a long-range aromatic $^4J_{HH}$ coupling between the two aromatic protons at 7.85 and 7.72 ppm. In case of a geminal or $^3J_{HH}$ coupling, the coupling constant should be much higher (6-15 Hz). Further analysis of the two-dimensional NMR spectra (HSQC and HMBC) supported the proposed structure, as the data were consistent with the previously published data of AETX. Taking all data into account, (**5**) was identified as 5,7-dibromo-3-carbonitrile indole – the western building block of AETX (**Fig. 3.15**).

Indeed, this metabolite has been shown to play a key role in AETX biosynthesis, as the coupling to the eastern building block (2,3,5-tribromo indole), catalyzed by the cytochrome P450 enzyme AetB, represents the final reaction sequence in AETX biosynthesis.[117]
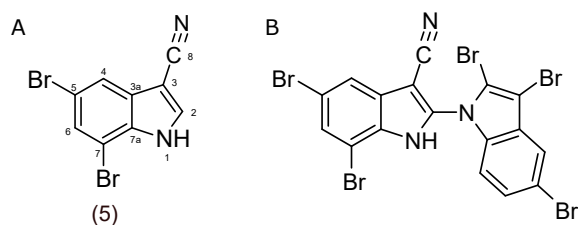

Figure 3.15.: (**A**) Chemical structure of the isolated and synthesized 5,7-dibromo-3-carbonitrile indole (**5**). (**B**) Chemical structure of AETX.

3.2.4. Synthesis

One of the main problems in natural product research is the limited yield of isolable natural products. For more detailed bioactivity and biosynthesis studies, more material was necessary, thus a total synthesis of AETX and of the two indole building blocks was needed. The successful isolation of the indole-carbonitrile moiety suggested synthesizing this moiety first. Literature offered two ways to introduce a carbonitrile at position 3 to an indole core structure. (**Fig. 3.16**).

The first option (method A) was via the conversion of a respective aldehyde in position 3 into a carbonitrile function.[118] The reaction mechanism includes the *in situ* generation of hydroxylamine from 1-nitropropane, and the following formation of an aldoxime, which rapidly dehydrates to the corresponding nitrile.

The second option (method B) was based on the use of chlorosulfonyl isocyanate (CSI), which reacts with nitrogen containing heterocycles to N-chloro-sulfonyl amides as intermediates, which rapidly decompose to the corresponding nitriles.[119]

Both methods were tested, and both were found suitable to synthesize the desired compound. The NMR spectra of the isolated natural product were identical to the spectra of the synthetic compounds, demonstrating that the synthesis was successful (**Fig. A.3.7-A.3.8**). As the yield of the compound synthesized via method B was higher compared to method A, and the presence of an aldehyde in position 3 was not necessary for the reaction, method B was chosen as the way of choice to introduce a carbonitrile in position 3. Based on these findings, a total synthesis of AETX could be established in cooperation with chemists at the Leibniz Institute for Plant Biochemistry. The synthesis has been published and can be found in the appendix (**A.5.2**).
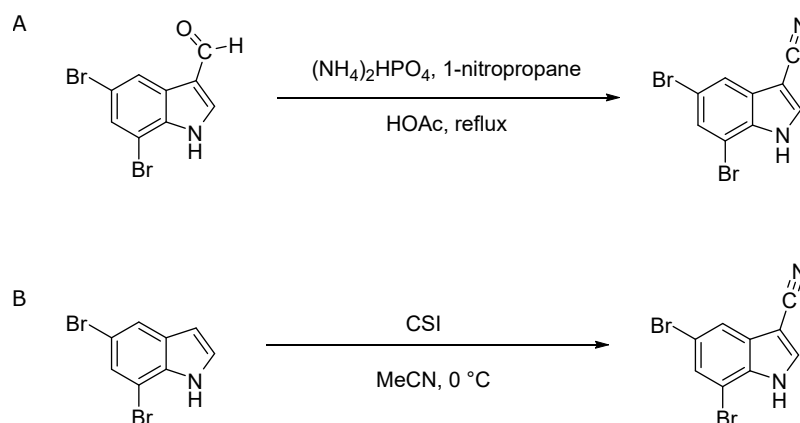
Figure 3.16.: Synthesis of 5,7-dibromo-3-carbonitrile via method A (**A**) and method B (**B**).

3.3. Analysis of AETX in environmental samples

Tracking the spreading of *A. hydrillicola*, a producer of two potent toxins, will be crucial in the future prevention and intervention, controlling its impact on environmental and human health. A convenient means to detect the cyanobacterium in environmental samples would be a PCR protocol for the detection of the AETX biosynthesis gene cluster. Thus, such a protocol was designed by our cooperation partners at the Czech Academy of Science in Budweis. In order to establish a correlation between the PCR test and the presence or absence of AETX in the samples, all samples were analyzed with PCR and HPLC-MS. The HPLC-MS data was on one hand used to make a qualitative decision (AETX present/ not present), but on the other hand it was also used for semi-quantitative assessment of AETX concentrations in the samples.

3.3.1. HPLC-MS analysis

For HPLC-MS analysis, 60 samples of colonialized *H. verticillata* leaves from 4 different samples locations in the southeastern USA, and 3 samples from *Justicia americana*, from a fifth location, were extracted following the previous mentioned protocol (**2.1.5**). After extraction, samples were analyzed via HPLC-MS using the analytical screening method (**Tab. 2.9**). To confirm presence or absence of AETX, the samples were compared with a standard containing AETX. As AETX is highly lipophilic, and tends to stick to the injection needle or other parts of the HPLC, blanks were run between the samples, excluding cross contamination. The results of all samples are summarized in **Tab. A.4.1**. Except for one sample (N 14), the results of the PCR analysis were confirmed by HPLC-MS. In addition, samples from Lake Sinclair, which were taken as a negative control as no *A. hydrillicola* was detected there, were all negative, demonstrating the reliability of the PCR and HPLC-MS experiments. Furthermore, samples from J. Strom Thurmond

reservoir were also AETX positive. However, in these samples, *A. hydrillicola* did not grow on the invasive plant *H. verticillata*, which in the meanwhile has been eradicated from this lake, but on the native water plant *J. americana*. This demonstrates the capability of *A. hydrillicola* to grow on different hosts, and not losing the ability to synthesize AETX. The different sampling locations were compared with each other, investigating, if there is a difference in the AETX quantity present in the samples. For this purpose, the areas were determined and related to the mass of the analyzed sample. Statistical analysis (one factorial ANOVA) showed that the observed differences in AETX content in the sample locations and hosts was statistically significant (**Fig. 3.17**). The amount of AETX was significant higher in samples from Long Branch Reservoir than in samples from all the other sampling locations. The lowest amounts were found in J. Strom Thurmond reservoir. This is interesting, because it demonstrates that the capability of synthesizing AETX seems to be reduced if *A. hydrillicola* is not growing on *H. verticillata*. These findings strengthen the hypothesis that *H. verticillata* has an impact on the biosynthesis of AETX. One possibility is that *H. verticillata* can accumulate bromine from the water, which in turn could be used for the synthesis of AETX, but this has to be evaluated in further studies.[96]
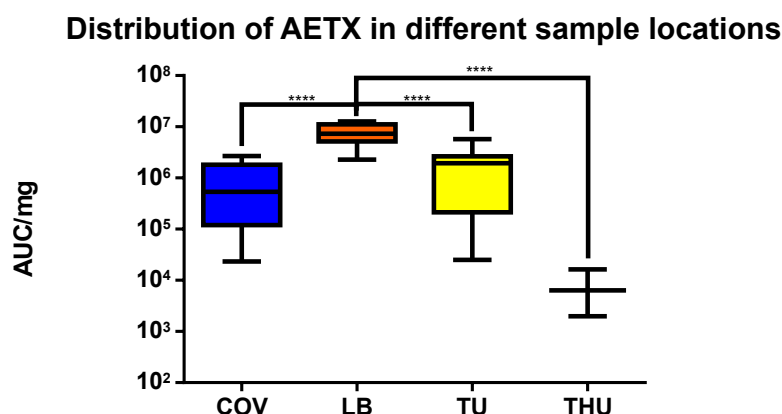

Figure 3.17.: Boxplot diagram, based on semiquantitative measurements, of AETX distribution in the respective sample location (****p< 0.0001). Covington Reservoir (COV), Long Branch Reservoir (LB), Jackson Lake on Tussahaw Creek (TUS), Lake Sinclair (SIN), and J. Strom Thurmond reservoir (THU).

4. Conclusion and outlook

In the presented work, I took a closer look at the secondary metabolome of the cyanobacterium *A. hydrillicola.* Initial screening of an *A. hydrillicola* extract showed a high cytotoxicity, which was not only due to AETX, the only known metabolite produced by this bacterium so far. Thus, the major aim of this work was to identify the main cytotoxin in this extract. Furthermore, a preliminary analysis by S. Breinlinger revealed the presence of several AETX related compounds in the crude extract.[115] Thus, the second aim of this project was to identify these metabolites, to reach a better understanding of the AETX biosynthesis and maybe also of its mode of action. Finally, the implementation of a reliable PCR protocol for the detection of the AETX biosynthetic gene cluster, which could play a key role in the future tracking and prevention of AETX related environmental issues, was supported.

4.1. Aetokthonostatins: dolastatin derivatives produced by *A. hydrillicola*

A traditional bioactivity-guided isolation approach could solve the question, why extracts of *A. hydrillicola* were highly cytotoxic, although the already known metabolite AETX was only moderately cytotoxic. Microfractionation and subsequent cytotoxicity testing suggested the main metabolite of this strain to be responsible for the observed toxicity. Isolation of the metabolite was achieved by semi-preparative HPLC under basic conditions. The isolated compound was a pentapeptide belonging to the class of dolastatins, with a so far unknown C-terminus. The main compound was named aetokthonostatin (AEST). The dolastatins are among the most interesting secondary metabolites from cyanobacteria, as they already made their way into the clinic.

In addition to the classical bioactivity-guided process, the extract was analyzed by the platform GNPS, which led to the identification of several AEST derivatives, and finally to the isolation of three of these derivatives. Comparative bioactivity testing of the derivatives, including one compound missing the C-terminal monomer, showed that an aromatic moiety at the C-terminus is not necessary for activity of this compound class, as it has been suggested in the literature.[77]

Furthermore, treatment of the nematode *C. elegans* with AEST, revealed an inhibition of nematode reproduction, as the worms were unable to produce offspring.

4. Conclusion and outlook

A. hydrillicola was identified as the first freshwater cyanobacterium producing dolastatin related compounds – all other known members of this compound family were produced by marine cyanobacteria of the genus *Symploca/Caldora*.[44][70][71][120][121] This raises one question, which is unanswered so far, and should be addressed by further studies: are there relevant concentrations of AEST in affected lakes, which could have a negative influence on the ecosystem and human health? Also, the question why cyanobacteria are producing these highly toxic metabolites is still unanswered. It could be an act of self-defense against possible predators, but this has to be evaluated in future studies. The disease Vacuolar Myelinopathy (VM), which is responsible for the deaths of bald eagles in the southeastern United States and which could be linked to the metabolite AETX, has already demonstrated the potential of toxic cyanobacterial metabolites and their impact on our ecosystems. Taking these new findings into account, a monitoring program of the further spreading of *A. hydrillicola*, as the producer of two potent toxins, should be considered as necessary in the future.

4.2. Identification and synthesis of AETX related compounds

Mass spectrometry based molecular networking with the software platform GNPS was used to identify AETX derivatives in the *A. hydrillicola* extract. A carbonitrile indole monomer could be isolated in quantities sufficient for NMR analysis. The relative high concentration in the crude extract suggested that this indole moiety is one of the major intermediates of AETX biosynthesis. The biosynthesis of AETX could be solved in the meantime, and indeed the isolated 5,7-dibromo-3-carbonitrile indole is one of the key building blocks of the biosynthesis.[117] To supply larger amounts of this compound for further biosynthetic studies, as well as activity studies, a synthesis of the 5,7-dibromo-3-carbonitrile indole was established. Two different approaches were tested, and the one with CSI as the main reagent turned out as the most efficient one. Based on this reaction the first total synthesis of AETX, in collaboration with partners at the Leibniz Institute for Plant Biochemistry, could be achieved. The total synthesis of AETX is a major step forwards, as now enough material for mode of actions studies can be provided, and several derivatives, which can be used to understand structure-activity relationships, can be synthesized.

The analyses furthermore revealed the presence of iodinated AETX derivatives. Indeed, in the meanwhile it has been demonstrated that the flavin reductase/flavin-dependent halogenase AetF is able to iodinate compounds, in addition to the already known bromination.[122] Future studies should show if these iodinated derivatives are produced in quantities sufficient for isolation, NMR analysis, and bioactivity testing, to see whether differences in the halogenation have an influence on the bioactivity of AETX.

With this work, the basis for further studies with AETX, regarding e.g. the elucidation of its mode of action, as well as structure-activity relationships, were set. This will help to get a better understanding of this neurotoxin.

4.3. Analysis of AETX in environmental samples

PCR primers for the detection of biosynthesis gene clusters have been developed for many natural products.[123] As AETX was shown to be of high ecosystem relevance, the need to track the spreading of the producing cyanobacterium has become mandatory. The development of PCR primers for AETX biosynthesis genes, by collaborators at the Czech Academy of Sciences, was a major step forward in this regard. The reliability of this protocol was confirmed by HPLC-MS analysis – AETX has only been detected in PCR-positive samples. The semiquantitative analysis of the generated data could show that the amount of AETX in the samples varies between different samples locations. This could be due to a different availability of bromide, as this is necessary for AETX biosynthesis. In addition, it could be shown that AETX biosynthesis is not solely depending on the interaction between *A. hydrillicola* and *H. verticillata*, as AETX could also be detected in J. Strom Thurmond reservoir, which is free of *H. verticillata*. The native aquatic plant *J. americana* was also found to be colonializable by *A. hydrillicola*, but the amount of AETX detected in relationship with *J. americana* samples was much lower, compared to the sample locations where *H. verticillata* was present. This could relate to previous findings that *H. verticillata* is capable of accumulating bromide, which could enhance the biosynthesis of AETX. To prove this hypothesis, it would be important to study the bromide concentrations in the different sample locations. Nevertheless, a fast protocol for the detection of the AETX biosynthetic gene cluster in environmental samples could be established, which could help in monitoring the AETX producing cyanobacterium.

Bibliography

[1] Williams, D.H.; Stone, M.J.; Hauck, P.R.; Rahman, S.K., Why are secondary metabolites (natural products) biosynthesized?, *Journal of Natural Products*, **1989**, *52*, 1189–1208.

[2] Wink, M., Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores, *Theoretical and Applied Genetics*, **1988**, *75*, 225–233.

[3] Teasdale, M.E.; Liu, J.; Wallace, J.; Akhlaghi, F.; Rowley, D.C., Secondary metabolites produced by the marine bacterium Halobacillus salinus that inhibit quorum sensing-controlled phenotypes in gram-negative bacteria, *Applied and Environmental Microbiology*, **2009**, *75*, 567–572.

[4] Oladipo, A.; Enwemiwe, V.; Ejeromedoghene, O.; Adebayo, A.; Ogunyemi, O.; et al., Production and Functionalities of Specialized Metabolites from Different Organic Sources, *Metabolites*, **2022**, *12*, 534–553.

[5] Solecki, R.S., Shanidar IV, a Neanderthal Flower Burial in Northern Iraq, *Science*, **1975**, *190*, 880–881.

[6] Farnsworth, N.R.; Olayiwola, A.; Bingel, A.S.; Soejarto, D.D.; Guo, Z., Medicinal plants in therapy, *Bulletin of the World Health Organization*, **1985**, *63*, 965–981.

[7] Schmitz, R., Friedrich Wilhelm Sertürner and the discovery of morphine, *Pharmacy in History*, **1985**, *27*, 61–74.

[8] Montinari, M.R.; Minelli, S.; de Caterina, R., The first 3500 years of aspirin history from its roots - A concise summary, *Vascular pharmacology*, **2019**, *113*, 1–8.

[9] Ligon, B.L., Penicillin: Its discovery and early development, *Seminars in Pediatric Infectious Diseases*, **2004**, *15*, 52–57.

[10] Waltenberger, B.; Mocan, A.; Šmejkal, K.; Heiss, E.H.; Atanasov, A.G., Natural products to counteract the epidemic of cardiovascular and metabolic disorders, *Molecules*, **2016**, *21*, 807–831.

Bibliography

[11] Wani, M.C.; Taylor, H.L.; Wall, M.E.; Coggon, P.; McPhail, A.T., The isolation and structure of taxol, a novel antileukemic and antitumor agent from Taxus brevifolia, *Journal of the American Chemical Society*, **1971**, *93*, 2325–2327.

[12] Atanasov, A.G.; Zotchev, S.B.; Dirsch, V.M.; Supuran, C.T., Natural products in drug discovery: advances and opportunities, *Nature Reviews Drug Discovery*, **2021**, *20*, 200–216.

[13] Newman, D.J.; Cragg, G.M., Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019, *Journal of Natural Products*, **2020**, *83*, 770–803.

[14] Feher, M.; Schmidt, J.M., Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry, *Journal of Chemical Information and Computer Sciences*, **2003**, *43*, 218–227.

[15] Wolfender, J.L.; Marti, G.; Thomas, A.; Bertrand, S., Current approaches and challenges for the metabolite profiling of complex natural extracts, *Journal of chromatography. A*, **2015**, *1382*, 136–164.

[16] Allard, P.M.; Genta-Jouve, G.; Wolfender, J.L., Deep metabolome annotation in natural products research: Towards a virtuous cycle in metabolite identification, *Current Opinion in Chemical Biology*, **2017**, *36*, 40–49.

[17] Stuart, K.A.; Welsh, K.; Walker, M.C.; Edrada-Ebel, R., Metabolomic tools used in marine natural product drug discovery, *Expert opinion on drug discovery*, **2020**, *15*, 499–522.

[18] Wolfender, J.L.; Nuzillard, J.M.; van der Hooft, J.J.J.; Renault, J.H.; Bertrand, S., Accelerating metabolite identification in natural product research: toward an ideal combination of LC-HRMS/MS and NMR profiling, in silico databases and chemometrics, *Analytical Chemistry*, **2019**, 704–732.

[19] Wang, M.; Carver, J.J.; Phelan, V.V.; et al, Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, *Nature Biotechnology*, **2016**, *34*, 828–837.

[20] Rutz, A.; Sorokina, M.; Galgonek, J.; Mietchen, D.; Willighagen, E.; et al., The LOTUS initiative for open knowledge management in natural products research, *eLife*, **2022**, *11*, 70780–70821.

[21] van Santen, J.A.; Poynton, E.F.; Iskakova, D.; McMann, E.; Alsup, T.A.; et al., The Natural Products Atlas 2.0: a database of microbially-derived natural products, *Nucleic Acids Research*, **2022**, *50*, 1317–1323.

[22] Liu, X.; Locasale, J.W., Metabolomics: a Primer, *Trends in Biochemical Sciences*, **2017**, *42*, 274–284.

[23] Ziemert, N.; Alanjary, M.; Weber, T., The evolution of genome mining in microbes - a review, *Natural Product Reports*, **2016**, *33*, 988–1005.

[24] Kayrouz, C.M.; Zhang, Y.; Pham, T.M.; Ju, K.S., Genome mining reveals the phosphonoalamide natural products and a new route in phosphonic acid biosynthesis, *ACS Chemical Biology*, **2020**, *15*, 1921–1929.

[25] Viehrig, K.; Surup, F.; Volz, C.; Herrmann, J.; Abou Fayad, A.; et al., Structure and biosynthesis of crocagins: polycyclic posttranslationally modified ribosomal peptides from Chondromyces crocatus, *Angewandte Chemie - International Edition*, **2017**, *56*, 7407–7410.

[26] Surup, F.; Viehrig, K.; Rachid, S.; Plaza, A.; Maurer, C.K.; et al., Crocadepsins-depsipeptides from the myxobacterium Chondromyces crocatus found by a genome mining approach, *ACS Chemical Biology*, **2018**, *13*, 267–272.

[27] Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; et al., Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens, *Proceedings of the National Academy of Sciences*, **2011**, *108*, 6258–6263.

[28] Zhang, H.; Boghigian, B.A.; Armando, J.; Pfeifer, B.A., Methods and options for the heterologous production of complex natural products, *Natural Product Reports*, **2011**, *28*, 125–151.

[29] Hemphill, C.F.P.; Sureechatchaiyan, P.; Kassack, M.U.; Orfali, R.S.; Lin, W.; et al., OSMAC approach leads to new fusarielin metabolites from Fusarium tricinctum, *The Journal of Antibiotics*, **2017**, *70*, 726–732.

[30] Hussain, A.; Rather, M.A.; Dar, M.S.; Aga, M.A.; Ahmad, N.; et al., Novel bioactive molecules from Lentzea violacea strain AS 08 using one strain-many compounds (OSMAC) approach, *Bioorganic & Medicinal Chemistry Letters*, **2017**, *27*, 2579–2582.

[31] Abdel-Razek, A.S.; Hamed, A.; Frese, M.; Sewald, N.; Shaaban, M., Penicisteroid C: New polyoxygenated steroid produced by co-culturing of Streptomyces piomogenus with Aspergillus niger, *Steroids*, **2018**, *138*, 21–25.

[32] Moussa, M.; Ebrahim, W.; Bonus, M.; Gohlke, H.; Mándi, A.; et al., Co-culture of the fungus Fusarium tricinctum with Streptomyces lividans induces production of cryptic naphthoquinone dimers, *RSC Advances*, **2019**, *9*, 1491–1500.

Bibliography

[33] Nichols, D.; Cahoon, N.; Trakhtenberg, E.M.; Pham, L.; Mehta, A.; et al., Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species, *Applied and Environmental Microbiology*, **2010**, *76*, 2445–2450.

[34] Dias, D.A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery, *Metabolites*, **2012**, *2*, 303–336.

[35] Torre, L.A.; Siegel, R.L.; Ward, E.M.; Jemal, A., Global cancer incidence and mortality rates and trends–an update, *Cancer Epidemiology and Prevention Biomarkers*, **2016**, *25*, 16–27.

[36] Wohlleben, W.; Mast, Y.; Stegmann, E.; Ziemert, N., Antibiotic drug discovery, *Microbial biotechnology*, **2016**, *9*, 541–548.

[37] Welker, M.; Dittmann, E.; von Döhren, H., Cyanobacteria as a source of natural products, *Methods in Enzymology*, **2012**, *517*, 23–46.

[38] Schopf, J.W.; Packer, B.M., Early archean (3.3-billion to 3.5-billion-year-old) microfossils from warrawoona group, australia, *Science*, **1987**, *237*, 70–73.

[39] Martin, W.F.; Garg, S.; Zimorski, V., Endosymbiotic theories for eukaryote origin, *Philosophical Transactions of the Royal Society of London, B: Biological Sciences*, **2015**, *370*, 20140330–20140348.

[40] Hachicha, R.; Elleuch, F.; Ben Hlima, H.; Dubessay, P.; de Baynast, H.; et al., Biomolecules from microalgae and cyanobacteria: applications and market survey, *Applied Sciences*, **2022**, *12*, 1924–1950.

[41] Mehdizadeh Allaf, M.; Peerhossaini, H., Cyanobacteria: Model Microorganisms and Beyond, *Microorganisms*, **2022**, *10*, 696–718.

[42] Adams, D.G.; Duggan, P.S., Heterocyst and akiniete differentiation in cyanobacteria, *New Phytologist*, **1999**, *144*, 3–33.

[43] Jiang, L.; Li, T.; Jenkins, J.; Hu, Y.; Brueck, C.L.; et al., Evidence for a mutualistic relationship between the cyanobacteria Nostoc and fungi Aspergilli in different environments, *Applied Microbiology and Biotechnology*, **2020**, *104*, 6413–6426.

[44] Luesch, H.; Moore, R.E.; Paul, V.J.; Mooberry, S.L.; Corbett, T.H., Isolation of dolastatin 10 from the marine cyanobacterium Symploca species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1, *Journal of Natural Products*, **2001**, *64*, 907–910.

[45] Huang, I.S.; Zimba, P.V., Cyanobacterial bioactive metabolites-A review of their chemistry and biology, *Harmful Algae*, **2019**, *86*, 139–209.

[46] Bastien, C.; Cardin, R.; Veilleux, E.; Deblois, C.; Warren, A.; et al., Performance evaluation of phycocyanin probes for the monitoring of cyanobacteria, *Journal of Environmental Monitoring*, **2011**, *13*, 110–118.

[47] Demay, J.; Bernard, C.; Reinhardt, A.; Marie, B., Natural products from cyanobacteria: focus on beneficial activities, *Marine Drugs*, **2019**, *17*, 320–369.

[48] Jones, M.R.; Pinto, E.; Torres, M.A.; Dörr, F.; Mazur-marzec, H.; et al., CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria, *Water Research*, **2021**, *196*, 117017–117029.

[49] Dittmann, E.; Gugger, M.; Sivonen, K.; Fewer, D.P., Natural product biosynthetic diversity and comparative genomics of the cyanobacteria, *Trends in Microbiology*, **2015**, *23*, 642–652.

[50] Chlipala, G.; Mo, S.; Carcache de Blanco, E.J.; Ito, A.; Bazarek, S.; et al., Investigation of antimicrobial and protease-inhibitory activity from cultured cyanobacteria, *Pharmaceutical Biologyiology*, **2009**, *47*, 53–60.

[51] Mi, Y.; Zhang, J.; He, S.; Yan, X., New peptides isolated from marine cyanobacteria, an overview over the past decade, *Marine Drugs*, **2017**, *15*, 132–159.

[52] Sivonen, K.; Leikoski, N.; Fewer, D.P.; Jokela, J., Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria, *Applied Microbiology and Biotechnology*, **2010**, *86*, 1213–1225.

[53] Wang, M.; Zhang, J.; He, S.; Yan, X., A review study on macrolides isolated from cyanobacteria, *Marine Drugs*, **2017**, *15*, 126–145.

[54] Shah, S.A.A.; Akhter, N.; Auckloo, B.N.; Khan, I.; Lu, Y.; et al., Structural diversity, biological properties and applications of natural products from cyanobacteria. A review, *Marine Drugs*, **2017**, *15*, 354–384.

[55] Pearson, L.; Mihali, T.; Moffitt, M.; Kellmann, R.; Neilan, B., On the chemistry, toxicology and genetics of the cyanobacterial toxins, microcystin, nodularin, saxitoxin and cylindrospermopsin, *Marine Drugs*, **2010**, *8*, 1650–1680.

[56] van Apeldoorn, M.E.; van Egmond, H.P.; Speijers, G.J.A.; Bakker, G.J.I., Toxins of cyanobacteria, *Molecular Nutrition & Food Research*, **2007**, *51*, 7–60.

[57] Niedermeyer, T.H., Anti-infective natural products from cyanobacteria, *Planta Medica*, **2015**, *81*, 1309–1325.

[58] Katz, J.; Janik, J.E.; Younes, A., Brentuximab Vedotin (SGN-35), *Clinical Cancer Research*, **2011**, *17*, 6428–6436.

[59] Welker, M.; von Dohren, H., Cyanobacterial peptides - nature's own combinatorial biosynthesis, *FEMS Microbiology Reviews*, **2006**, *30*, 530–563.

[60] Pettit, G.R.; Kamano, Y.; Herald, C.L.; Tuinman, A.A.; Boettner, F.E.; et al., The isolation and structure of a remarkable marine animal antineoplastic constituent: dolastatin 10, *Journal of the American Chemical Society*, **1987**, *109*, 6883–6885.

[61] Luesch, H.; Moore, R.E.; Paul, V.J.; Mooberry, S.L.; Corbett, T.H., Isolation of dolastatin 10 from the marine cyanobacterium Symploca species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1, *Journal of Natural Products*, **2001**, *64*, 907–910.

[62] Bai, R.; Pettit, G.R.; Hamel, E., Dolastatin 10, a powerful cytostatic peptide derived from a marine animal. Inhibition of tubulin polymerization mediated through the vinca alkaloid binding domain, *Biochemical Pharmacology*, **1990**, *39*, 1941–1949.

[63] Pettit, G.R.; Singh, S.B.; Hogan, F.; Lloyd-Williams, P.; Herald, D.L.; et al., The absolute configuration and synthesis of natural (-)-dolastatin 10, *Journal of the American Chemical Society*, **1989**, *111*, 5463–5465.

[64] Kalemkerian, G.P.; Ou, X.; Adil, M.R.; Rosati, R.; Khoulani, M.M.; et al., Activity of dolastatin 10 against small-cell lung cancer in vitro and in vivo: induction of apoptosis and bcl-2 modification, *Cancer chemotherapy and pharmacology*, **1999**, *43*, 507–515.

[65] Maki, A.; Mohammad, R.; Raza, S.; Saleh, M.; Govindaraju, K.D.; et al., Effect of dolastatin 10 on human non-Hodgkin's lymphoma cell lines, *Anti-cancer drugs*, **1996**, *7*, 344–350.

[66] Turner, T.; Jackson, W.H.; Pettit, G.R.; Wells, A.; Kraft, A.S., Treatment of human prostate cancer cells with dolastatin 10, a peptide isolated from a marine shell-less mollusc, *The Prostate*, **1998**, *34*, 175–181.

[67] Bai, R.L.; Pettit, G.R.; Hamel, E., Binding of dolastatin 10 to tubulin at a distinct site for peptide antimitotic agents near the exchangeable nucleotide and vinca alkaloid sites, *Journal of Biological Chemistry*, **1990**, *265*, 17141–17149.

[68] Li, Y.; Kobayashi, H.; Hashimoto, Y.; Shirai, R.; Hirata, A.; et al., Interaction of marine toxin dolastatin 10 with porcine brain tubulin: competitive inhibition of rhizoxin and phomopsin A binding, *Chemico-Biological Interactions*, **1994**, *93*, 175–183.

[69] Ludueña, R.F.; Roach, M.C.; Prasad, V.; Pettit, G.R., Interaction of dolastatin 10 with bovine brain tubulin, *Biochemical Pharmacology*, **1992**, *43*, 539–543.

[70] Harrigan, G.G.; Luesch, H.; Yoshida, W.Y.; Moore, R.E.; Nagle, D.G.; et al., Symplostatin 1: A dolastatin 10 analogue from the marine cyanobacterium Symploca hydnoides, *Journal of Natural Products*, **1998**, *61*, 1075–1077.

[71] Horgen, F.D.; Kazmierski, E.B.; Westenburg, H.E.; Yoshida, W.Y.; Scheuer, P.J., Malevamide D: isolation and structure determination of an isodolastatin H analogue from the marine cyanobacterium Symploca hydnoides, *Journal of Natural Products*, **2002**, *65*, 487–491.

[72] Simmons, T.L.; McPhail, K.L.; Ortega-Barr\'ıa, E.; Mooberry, S.L.; Gerwick, W.H., Belamide A, a new antimitotic tetrapeptide from a Panamanian marine cyanobacterium, *Tetrahedron Letters*, **2006**, *47*, 3387–3390.

[73] Pettit, G.R.; Kamano, Y.; Herald, C.L.; Dufresne, C.; Cerny, R.L.; et al., Isolation and structure of the cytostatic depsipeptide dolastatin 13 from the sea hare Dolabella auricularia, *Journal of the American Chemical Society*, **1989**, 5015–5017.

[74] Harrigan, G.G.; Luesch, H.; Yoshida, W.Y.; Moore, R.E.; Nagle, D.G.; et al., Symplostatin 2: a dolastatin 13 analogue from the marine cyanobacterium Symploca hydnoides, *Journal of Natural Products*, **1999**, *62*, 655–658.

[75] Simmons, T.L.; Andrianasolo, E.; McPhail, K.; Flatt, P.; Gerwick, W.H., Marine natural products as anticancer drugs, *Molecular Cancer Therapeutics*, **2005**, *4*, 333–342.

[76] Maderna, A.; Doroski, M.; Subramanyam, C.; Porte, A.; Leverett, C.A.; et al., Discovery of cytotoxic dolastatin 10 analogues with N-terminal modifications, *Journal of Medicinal Chemistry*, **2014**, *57*, 10527–10543.

[77] Miyazaki, K.; Kobayashi, M.; Natsume, T.; Gondo, M.; Mikami, T.; et al., Synthesis and antitumor activity of novel dolastatin 10 analogs, *Chemical and pharmaceutical bulletin*, **1995**, *43*, 1706–1718.

[78] Dugal-Tessier, J.; Barnscher, S.D.; Kanai, A.; Mendelsohn, B.A., Synthesis and evaluation of dolastatin 10 analogues containing heteroatoms on the amino acid side chains, *Journal of Natural Products*, **2017**, *80*, 2484–2491.

[79] de Cecco, M.; Galbraith, D.N.; McDermott, L.L., What makes a good antibody-drug conjugate?, *Expert Opinion on Biological Therapy*, **2021**, *21*, 841–847.

[80] Tsuchikama, K.; An, Z., Antibody-drug conjugates: recent advances in conjugation and linker chemistries, *Protein & cell*, **2018**, *9*, 33–46.

Bibliography

[81] Doronina, S.O.; Toki, B.E.; Torgov, M.Y.; Mendelsohn, B.A.; Cerveny, C.G.; et al., Development of potent monoclonal antibody auristatin conjugates for cancer therapy, *Nature Biotechnology*, **2003**, *21*, 778–784.

[82] Senter, P.D.; Sievers, E.L., The discovery and development of brentuximab vedotin for use in relapsed Hodgkin lymphoma and systemic anaplastic large cell lymphoma, *Nature Biotechnology*, **2012**, *30*, 631–637.

[83] Newman, D.J., Natural product based antibody drug conjugates: clinical status as of november 9, 2020, *Journal of Natural Products*, **2021**, *84*, 917–931.

[84] Gao, G.; Wang, Y.; Hua, H.; Li, D.; Tang, C., Marine antitumor peptide dolastatin 10: biological activity, structural modification and synthetic chemistry, *Marine Drugs*, **2021**, *19*, 363–393.

[85] Cheng-Sánchez, I.; Moya-Utrera, F.; Porras-Alcalá, C.; López-Romero, J.M.; Sarabia, F., Antibody-Drug Conjugates Containing Payloads from Marine Origin, *Marine Drugs*, **2022**, *20*, 494–531.

[86] Thomas, N.J.; Meteyer, C.U.; Sileo, L., Epizootic vacuolar myelinopathy of the central nervous system of bald eagles (Haliaeetus leucocephalus) and American coots (Fulica americana), *Veterinary Pathology*, **1998**, *35*, 479–487.

[87] Fischer, J.R.; Lewis-Weis, L.A.; Tate, C.M.; Gaydos, J.K.; Gerhold, R.W.; et al., Avian vacuolar myelinopathy outbreaks at a southeastern reservoir, *Journal of Wildlife Diseases*, **2006**, *42*, 501–510.

[88] Larsen, R.S.; Nutter, F.B.; Augspurger, T.; Rocke, T.E.; Tomlinson, L.; et al., Clinical features of avian vacuolar myelinopathy in american coots, *Journal of the American Veterinary Medical Association*, **2002**, *221*, 80–85.

[89] Dodder, N.G.; Strandberg, B.; Augspurger, T.; Hites, R.A., Lipophilic organic compounds in lake sediment and american coot (Fulica americana) tissues, both affected and unaffected by avian vacuolar myelinopathy, *Science of The Total Environment*, **2003**, *311*, 81–89.

[90] Larsen, R.S.; Nutter, F.B.; Augspurger, T.; Rocke, T.E.; Thomas, N.J.; et al., Failure to transmit avian vacuolar myelinopathy to mallard ducks, *Journal of Wildlife Diseases*, **2003**, *39*, 707–711.

[91] Rocke, T.E.; Thomas, N.J.; Meteyer, C.U.; Quist, C.F.; Fischer, J.R.; et al., Attempts to identify the source of avian vacuolar myelinopathy for waterbirds, *Journal of Wildlife Diseases*, **2005**, *41*, 163–170.

[92] Birrenkott, A.H.; Wilde, S.B.; Hains, J.J.; Fischer, J.R.; Murphy, T.M.; et al., Establishing a food-chain link between aquatic plant material and avian vacuolar myelinopathy in mallards (Anas platyrhynchos), *Journal of Wildlife Diseases*, **2004**, *40*, 485–492.

[93] Wilde, S.B.; Murphy, T.M.; Hope, C.P.; Habrun, S.K.; Kempton, J.; et al., Avian vacuolar myelinopathy linked to exotic aquatic plants and a novel cyanobacterial species, *Environmental Toxicology*, **2005**, *20*, 348–353.

[94] Williams, S.K.; Kempton, J.; Wilde, S.B.; Lewitus, A., A novel epiphytic cyanobacterium associated with reservoirs affected by avian vacuolar myelinopathy, *Harmful Algae*, **2007**, *6*, 343–353.

[95] Wilde, S.B.; Johansen, J.R.; Wilde, H.D.; Jiang, P.; Bartelme, B.; et al., Aetokthonos hydrillicola gen. et sp. nov.: Epiphytic cyanobacteria on invasive aquatic plants implicated in Avian Vacuolar Myelinopathy, *Phytotaxa*, **2014**, *181*, 243–260.

[96] Breinlinger, S.; Phillips, T.J.; Haram, B.N.; Mareš, J.; Martínez Yerena, J.A.; et al., Hunting the eagle killer: a cyanobacterial neurotoxin causes vacuolar myelinopathy, *Science*, **2021**, *371*, 1335–1343.

[97] Rippka, R.; Deruelles, J.; Waterbury, J.B.; Herdman, M.; Stanier, R.Y., Generic assignments, strain histories and properties of pure cultures of cyanobacteria, *Journal of General Microbiology*, **1979**, *111*, 1–61.

[98] Ishida, K.; Nakagawa, H.; Murakami, M., Microcyclamide, a cytotoxic cyclic hexapeptide from the cyanobacterium Microcystis aeruginosa, *Journal of Natural Products*, **2000**, *63*, 1315–1317.

[99] Vichai, V.; Kirtikara, K., Sulforhodamine B colorimetric assay for cytotoxicity screening, *Nature Protocols*, **2006**, *1*, 1112–1116.

[100] Chambers, M.C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D.L.; et al., A cross-platform toolkit for mass spectrometry and proteomics, *Nature Biotechnology*, **2012**, *30*, 918.

[101] Katajamaa, M.; Miettinen, J.; Orešič, M., MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data, *Bioinformatics*, **2006**, *22*, 634–636.

[102] Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics*, **2010**, *11*, 395.

Bibliography

[103] Mingxun, W.; Jeremy, J.C.; Vanessa, V.P.; Laura, M.S.; Neha, G.; et al., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, *Nature Biotechnology*, **2016**, *34*, 828–837.

[104] Nothias, L.F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; et al., Feature-based molecular networking in the GNPS analysis environment, *Nature Methods*, **2020**, *17*, 905–908.

[105] Myers, O.D.; Sumner, S.J.; Li, S.; Barnes, S.; Du, X., One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: new algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks, *Analytical Chemistry*, **2017**, *89*, 8696–8703.

[106] Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome research*, **2003**, *13*, 2498–2504.

[107] CRC Press, Taylor & Francis Group, Dictionary of Natural Products, **2022**.

[108] Gabrielson, S.W., SciFinder, *Journal of the Medical Library Association*, **2018**, *106*, 588–590.

[109] Johansson, M.P.; Maaheimo, H.; Ekholm, F.S., New insight on the structural features of the cytotoxic auristatins MMAE and MMAF revealed by combined NMR spectroscopy and quantum chemical modelling, *Scientific Reports*, **2017**, *7*, 15920.

[110] Röhrborn, K., Isolierung und Charakterisierung neuer Naturstoffe des Cyanobakteriums Aetokthonos hydrillicola., Diploma thesis, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), **2022**.

[111] Moquist, P.N.; Bovee, T.D.; Waight, A.B.; Mitchell, J.A.; Miyamoto, J.B.; et al., Novel Auristatins with High Bystander and Cytotoxic Activities in Drug Efflux-positive Tumor Models, *Molecular Cancer Therapeutics*, **2021**, *20*, 320–328.

[112] Pettit, G.R.; Srirangam, J.K.; Barkoczy, J.; Williams, M.D.; Boyd, M.R.; et al., Antineoplastic agent 365. Dolastatin 10 SAR probes, *Anti-Cancer Drug Design*, **1998**, *13*, 243–277.

[113] Corsi, A.K.; Wightman, B.; Chalfie, M., A Transparent window into biology: a primer on Caenorhabditis elegans, *WormBook, the C. elegans research community*, **2015**.

[114] Zubovych, I.; Doundoulakis, T.; Harran, P.G.; Roth, M.G., A missense mutation in Caenorhabditis elegans prohibitin 2 confers an atypical multidrug resistance, *Proceedings of the National Academy of Sciences*, **2006**, *103*, 15523–15528.

[115] Breinlinger, S., Investigations into bioactive natural products from cyanobacteria - a search for drug leads and the discovery of a novel cyanotoxin., Dissertation, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), **2021**.

[116] Léon, A.; Cariou, R.; Hutinet, S.; Hurel, J.; Guitton, Y.; et al., HaloSeeker 1.0: A user-friendly software to highlight halogenated chemicals in nontargeted high-resolution mass spectrometry data sets, *Analytical Chemistry*, **2019**, *91*, 3500–3507.

[117] Adak, S.; Lukowski, A.L.; Schäfer, R.J.B.; Moore, B.S., From tryptophan to toxin: nature's convergent biosynthetic strategy to aetokthonotoxin, *Journal of the American Chemical Society*, **2022**, *144*, 2861–2866.

[118] Jiang, B.; Gu, X.H., Syntheses and cytotoxicity evaluation of bis(indolyl)thiazole, bis(indolyl)pyrazinone and bis(indolyl)pyrazine: analogues of cytotoxic marine bis(indole) alkaloid, *Bioorganic & Medicinal Chemistry*, **2000**, *8*, 363–371.

[119] Carbone, A.; Parrino, B.; Cusimano, M.G.; Spanò, V.; Montalbano, A.; et al., New thiazole nortopsentin analogues inhibit bacterial biofilm formation, *Marine Drugs*, **2018**, *16*, 274–289.

[120] Engene, N.; Tronholm, A.; Salvador-Reyes, L.A.; Luesch, H.; Paul, V.J., Caldora penicillata gen. nov., comb. nov. (cyanobacteria), a pantropical marine species with biomedical relevance, *Journal of Phycology*, **2015**, *51*, 670–681.

[121] Luesch, H.; Yoshida, W.Y.; Moore, R.E.; Paul, V.J.; Mooberry, S.L.; et al., Symplostatin 3, a new dolastatin 10 analogue from the marine cyanobacterium Symploca sp. VP452, *Journal of Natural Products*, **2002**, *65*, 16–20.

[122] Jiang, Y.; Snodgrass, H.M.; Zubi, Y.S.; Roof, C.V.; Guan, Y.; et al., The single-component flavin reductase/flavin-dependent halogenase AetF is a versatile catalyst for selective bromination and iodination of arenes and olefins, *Angewandte Chemie - International Edition*, **2022**, *61*, 14610–14615.

[123] Ayuso-Sacido, A.; Genilloud, O., New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups, *Microbial Ecology*, **2005**, *49*, 10–24.

A. Appendix

A.1. List of abbreviations

| Abbreviation	Meaning	Abbreviation	Meaning
ADC	Antibody-drug-conjugate	HSQC	Heteronuclear single quantum correlation
AEST	Aetokthonostatin	LC	Liquid chromatography
AETX	Aetokthonotoxin	mAB	Monoclonal antibody
COSY	Correlation spectroscopy	MeCN	Acetonitrile
Dap	Dolaproine	MeOH	Methanol
DAD	Diode array detector	MMAE	Monomethylauristatin E
DAAEST	Des-aph-aetokthonostatin	MMAEST	Monomethylaetokthonostatin
Dil	Dolaisoleuine	MMAF	Monomethylauristatin F
DMF	Dimethylformamide	MS	Mass spectrometry
DMSO	Dimethyl sulfoxide	MS/MS	Tandem-mass spectrometry
Dol-10	Dolastatin 10	NP	Natural product
EC_{50}	50 % efficacy concentration	NRP	Nonribosomal peptide
EDTA	Ethylenediaminetetraacetic acid	PCR	Polymerase chain reaction
ESI	Electrospray ionization	SAR	Structure-activity relationship
FA	Formic acid	SEC	Size-exclusion chromatography
FBS	Fetal bovine serum	SRB	Sulforhodamine B
FDA	Food and drug administration	TCA	Trichloroacetic acid
GNPS	Global natural products social molecular networking	TOCSY	Total correlation spectroscopy
HMBC	Heteronuclear multiple quantum correlation	t_r	Retention time
HPLC	High-performance liquid chromatography	VM	Vacuolar myelinopathy
HTS	High-throughput-screening		

A.2. Supplementary information to chapter 3.1
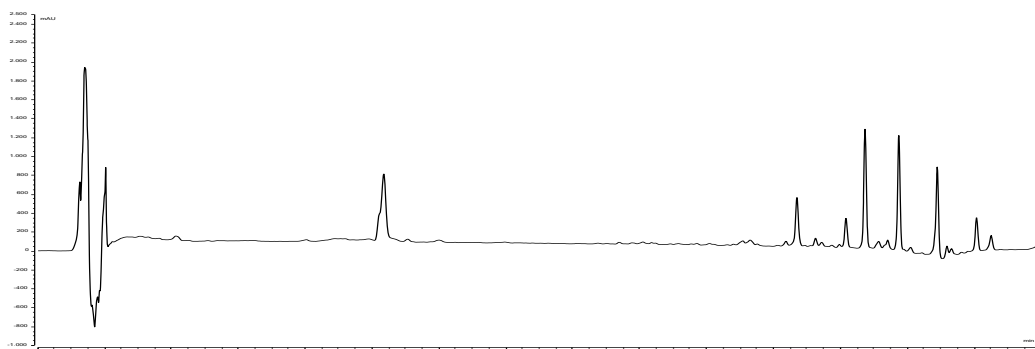
Figure A.2.1.: HPLC-UV chromatogram (210 nm) of an *A. hydrillicola* extract (gradient 5–100 % MeCN in water + 0.1 % FA each) on a Kinetex C_{18} column (2.1 mm ID x 100 mm).
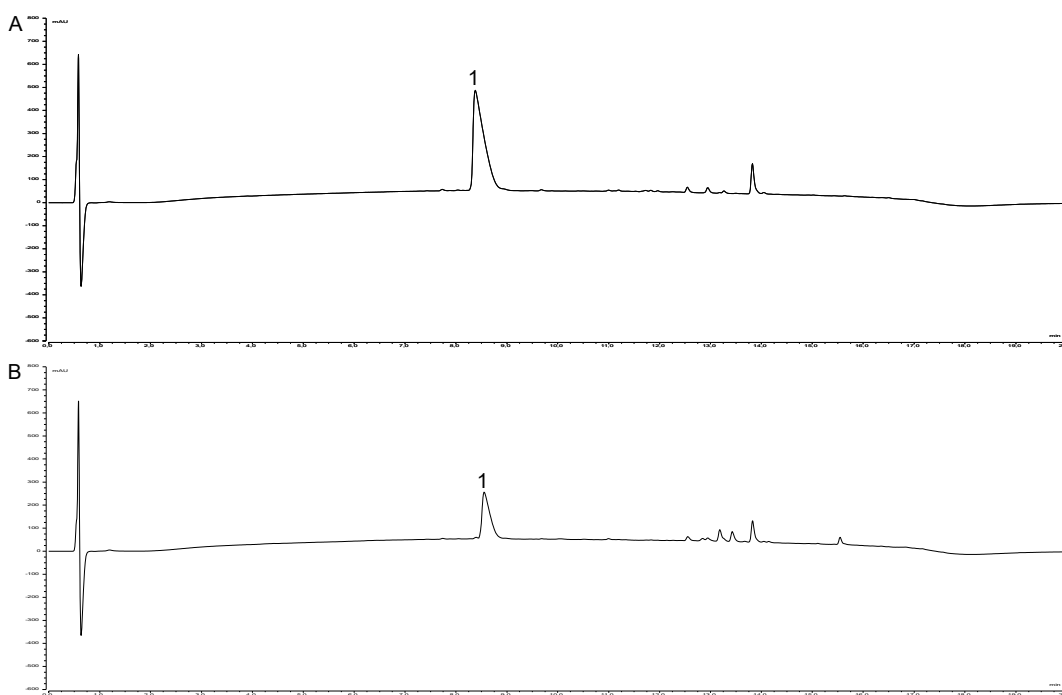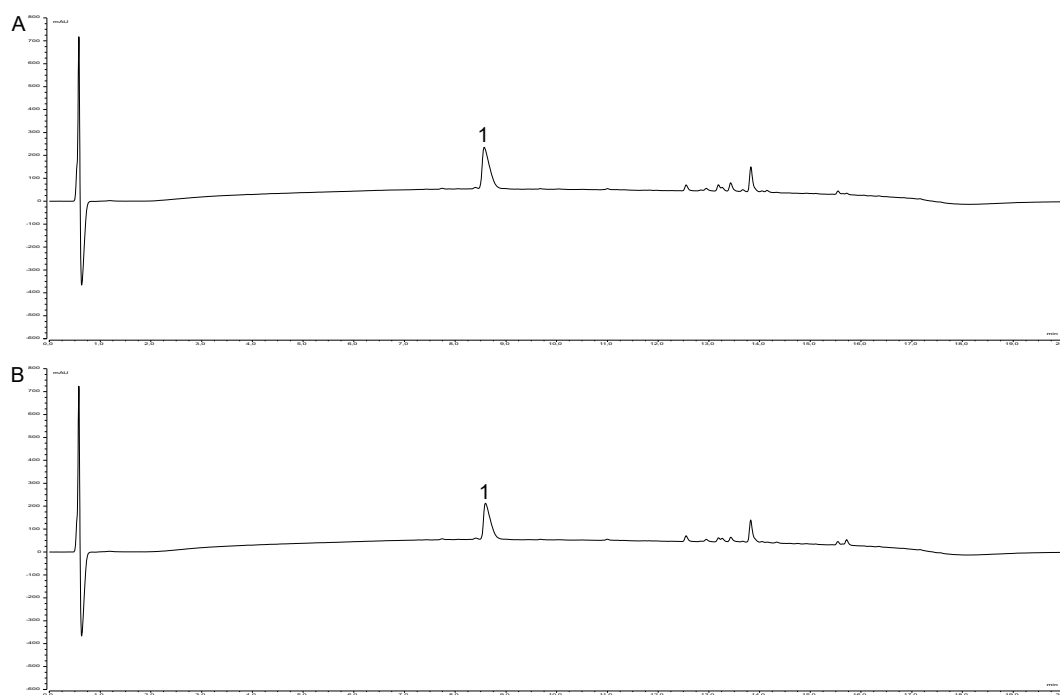

Figure A.2.2.: HPLC-UV chromatograms (210 nm) of flash fraction 12 (**A**) and 13 (**B**) (gradient 5–100 % MeCN in water + 0.1 % FA each) on a Kinetex C_{18} column (2.1 mm ID x 100 mm). Cytotoxic compounds (**1**).
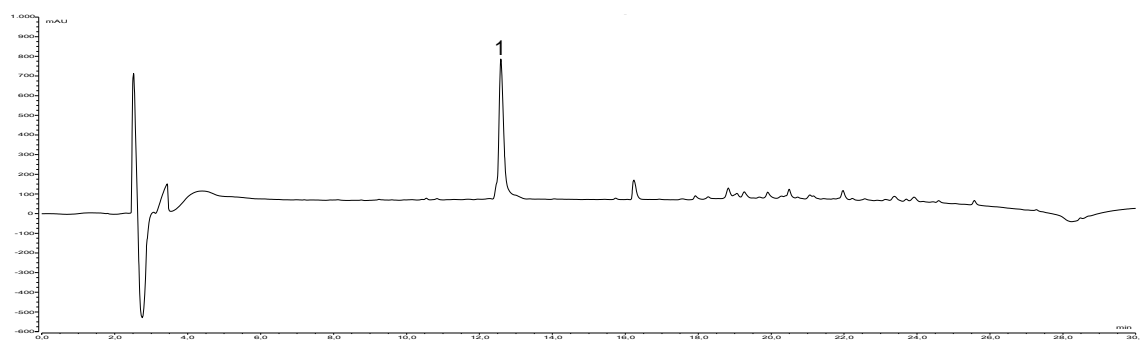
Figure A.2.3.: HPLC-UV chromatograms (210 nm) of flash fraction 14 (**A**) and 15 (**B**) (gradient 5–100 % MeCN in water + 0.1 % FA each) on a Kinetex C_{18} column (2.1 mm ID x 100 mm). Cytotoxic compounds (**1**).
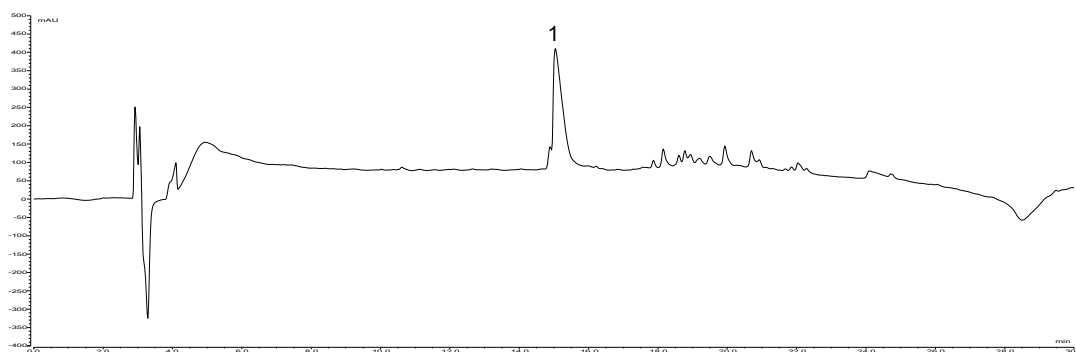

Figure A.2.4.: HPLC-UV chromatogram (210 nm) of combined fractions (gradient 5-100 % MeCN in water + 0.1 % FA each) on a Luna C_{18} column (4.6 mm ID x 250 mm). Cytotoxic compound (**1**).
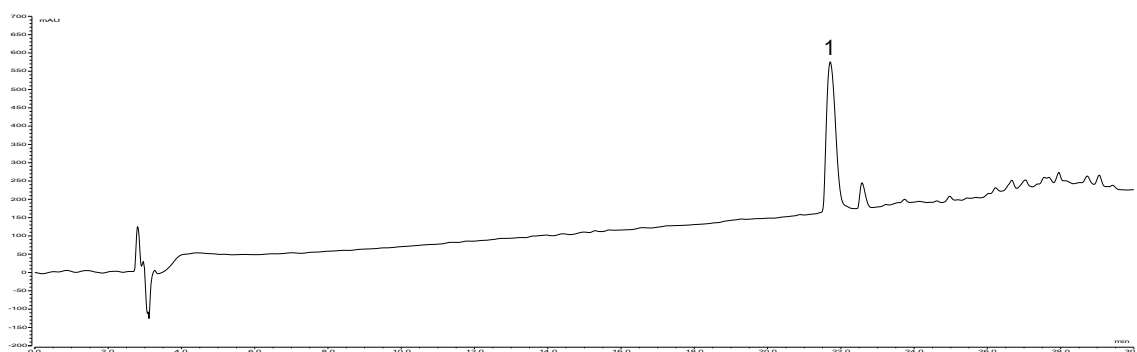
Figure A.2.5.: HPLC-UV chromatogram (210 nm) of combined fractions (gradient 5-100 % MeCN in water + 0.1 % FA each) on a Luna PFP column (4.6 mm ID x 250 mm). Cytotoxic compound (**1**).
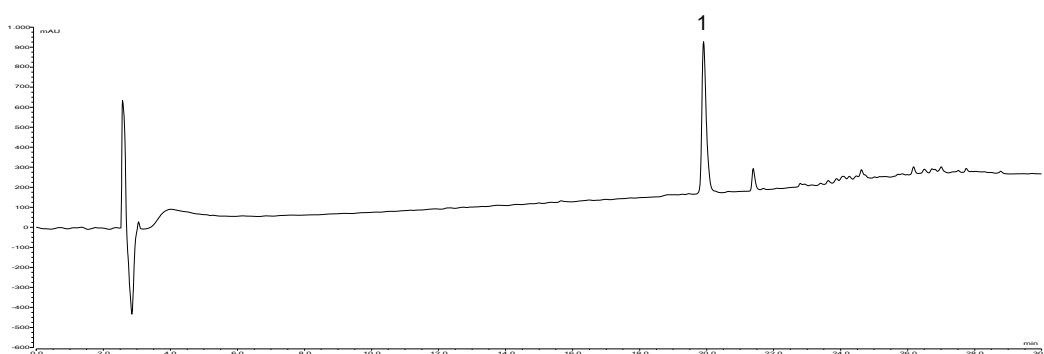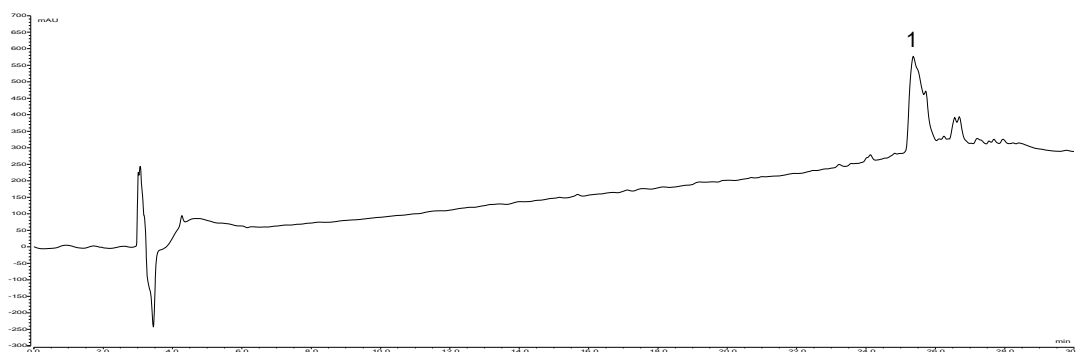

Figure A.2.6.: HPLC-UV chromatogram (210 nm) of combined fractions (gradient 5-100 % MeCN in water + 0.1 % FA each) on a Luna PH column (4.6 mm ID x 250 mm). Cytotoxic compound (**1**).

Figure A.2.7.: HPLC-UV chromatogram (210 nm) of combined fractions (gradient 5-100 % MeOH in water + 0.1 % FA each) on a Luna C_{18} column (4.6 mm ID x 250 mm). Cytotoxic compound (**1**).

Figure A.2.8.: HPLC-UV chromatogram (210 nm) of combined fractions (gradient 5-100 % MeOH in water + 0.1 % FA each) on a Luna PFP column (4.6 mm ID x 250 mm). Cytotoxic compound (**1**).
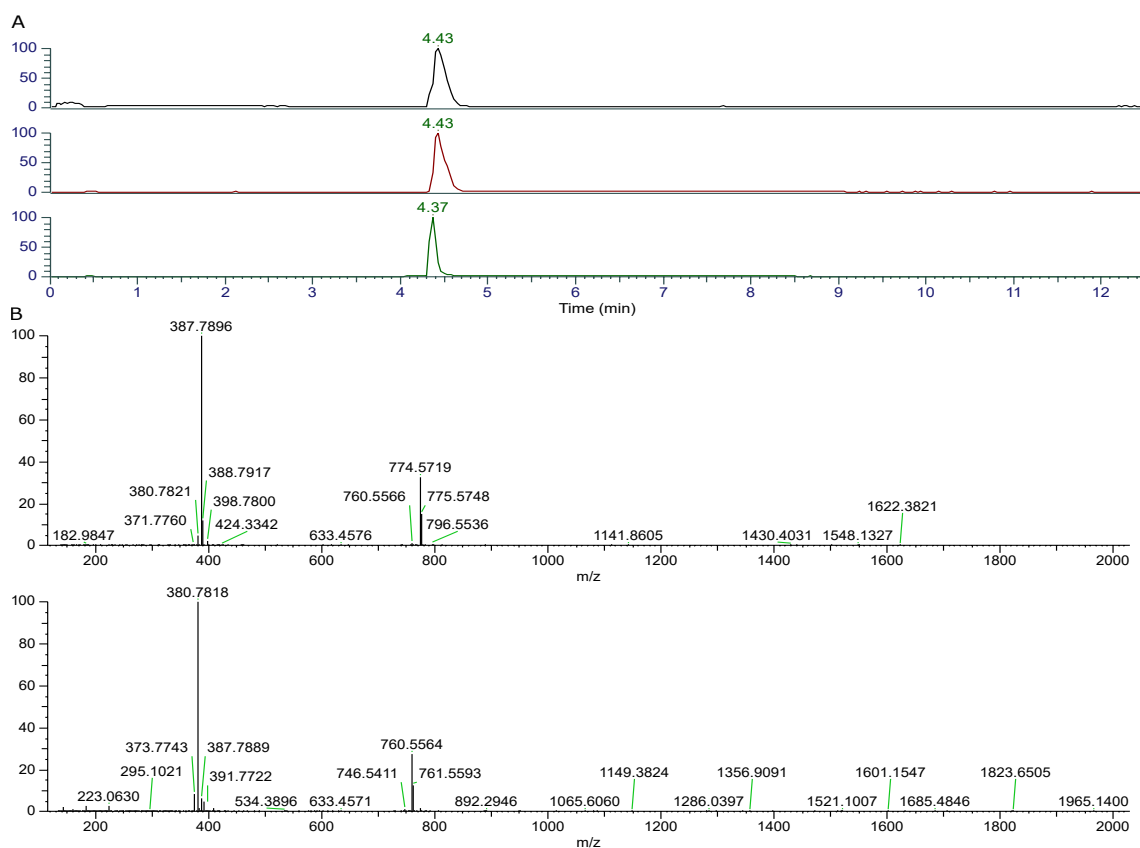

Figure A.2.9.: (**A**) HPLC-MS TIC (positive mode) of the isolated peak (black, t_r 4.43 min), EIC of m/z 774.57 (red, t_r 4.43 min) and EIC of m/z 760.56 (green, t_r 4.37 min). (**B**) Mass spectrum from the compound with m/z 774.57 (top) and 760.56 (bottom).
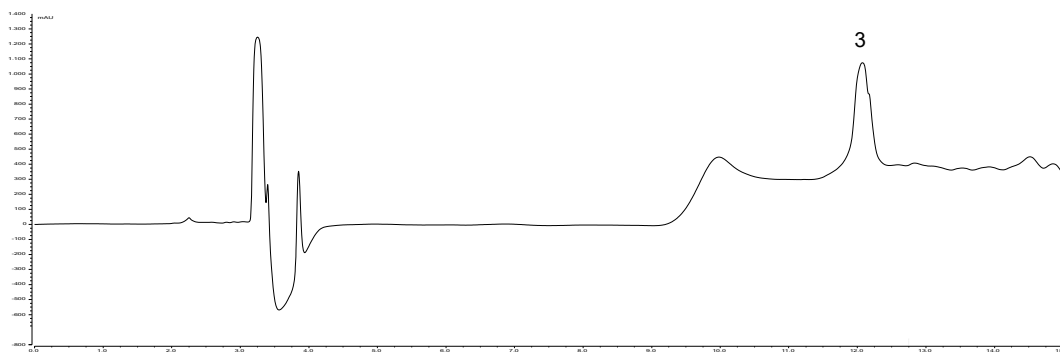
Figure A.2.10.: HPLC-UV chromatogram (210 nm) of the isolation of DAAEST (**3**) (gradient 10-50 % MeCN in water) on a Luna C_{18} column (4.6 mm ID x 250 mm).
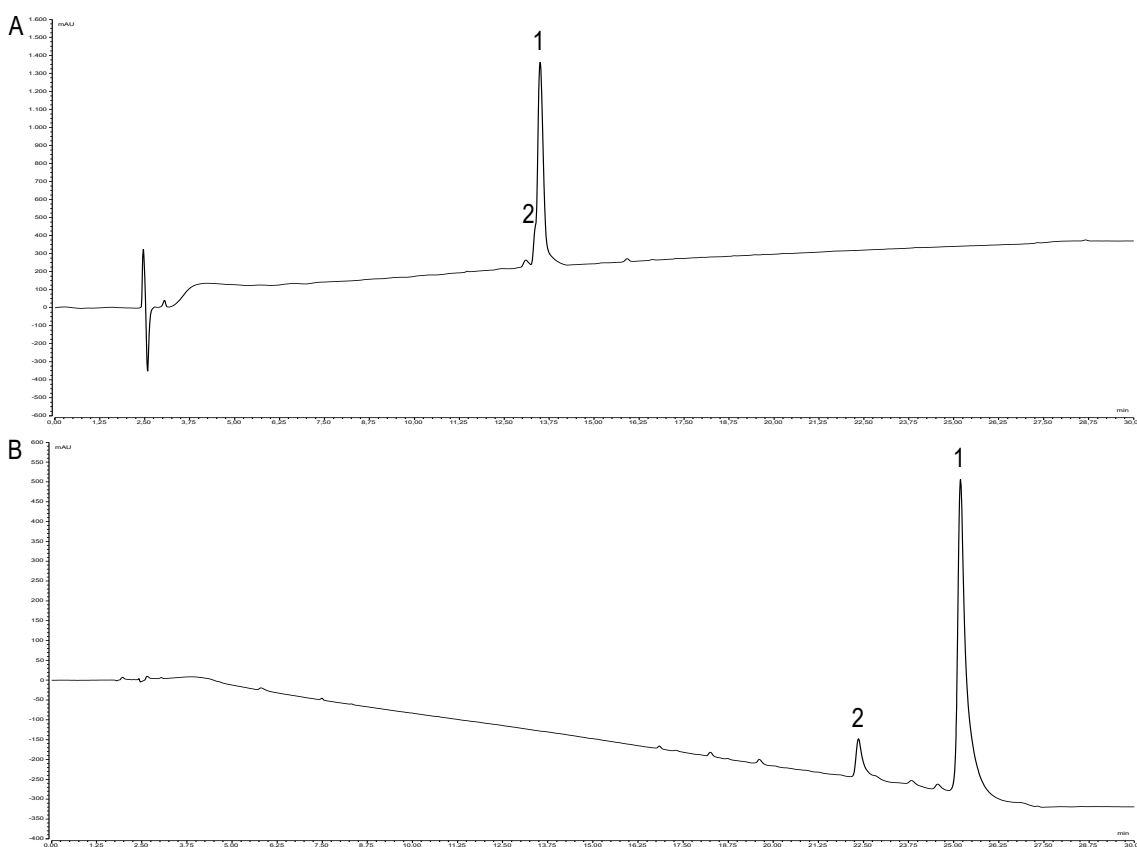

Figure A.2.11.: HPLC-UV chromatograms (210 nm) of combined fractions (gradient 5-100 % MeCN in water) on a Luna C_{18} column (4.6 mm ID x 250 mm). (**A**) pH of the mobile phase adjusted to 2-3 using formic acid. (**B**) pH of the mobile phase adjusted to 8-9 using ammonium acetate. Cytotoxic compound (**1**) and derivative (**2**).
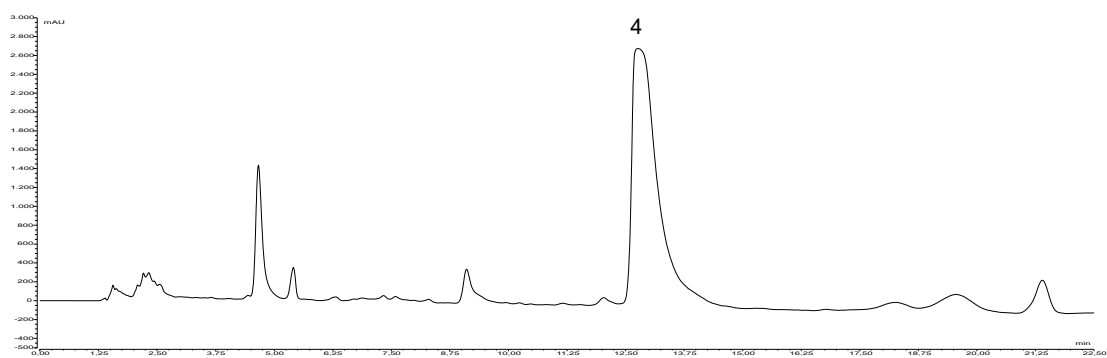
Figure A.2.12.: HPLC-UV chromatogram (210 nm) of the isolation of the degraded AEST derivative (**4**) (gradient 50–85 % MeCN in water + 10 mM ammonium acetate buffer) on a Luna C_{18} column (4.6 mm ID x 250 mm).

Table A.2.1.: One- and two-dimensional NMR data of both conformers of AEST (**1**) in DMSO-d$_6$. n.o = not observed.

unit	C/H no.	δH (J in Hz)	δC, mult	COSY	^{1}H-^{13}C	δH (J in Hz)	δC, mult
		Conformer 1				Conformer 2	
Aph	1	4.23, m	51.5, CH	H-2a, H-2b, H-7, NH	1 (Dap), 2, 3, 7, 8	4.13, m	52.5, CH
	2a	2.84, m	35.0, CH$_2$	H-1, H-2b	1, 3, 4, 7	2.80, m	34.8, CH$_2$
	2b	2.69, m	35.0, CH$_2$	H-1, H-2a	1, 3, 4, 7	2.66, m	34.8, CH$_2$
	3		139.0, qC				139.3, qC
	4a	7.20, m	128.6, CH	H-5a	2, 3, 5, 6	7.19, m	128.8, CH
	4b	7.20, m	128.6, CH	H-5b	2, 3, 5, 6	7.19, m	128.8, CH
	5a	7.19, m	127.9, CH	H-4a, H6	3, 4, 6	7.19, m	128.0, CH
	5b	7.19, m	127.9, CH	H-4b, H6	3, 4, 6	7.19, m	128.0, CH
	6	7.12, m	125.8, CH	H-5a, H-5b	4, 5	7.11, m	125.7, CH
	7	3.38, m	77.1, CH	H-1, H-8	1, 2, 8	3.32, m	76.9, CH
	8	1.07, m	14.9, CH$_3$	H-7	1, 7	1.07, m	14.6, CH$_3$
	9 (O-Me)	3.32, s	56.0, CH$_3$		7	3.30, s	55.9, CH$_3$
	NH	7.90, d (9.1)		H-1	1 (Dap), 1, 2	7.64, d (8.8)	
Dap	1		173.4, qC				173.0, qC
	2	2.25, m	42.9, CH	H-3, H-8	3, 4, 8	2.21, m	43.5, CH
	3	1.07, m	15.9, CH$_3$	H-2	1, 2, 3	1.06, m	15.9, CH$_3$
	4	3.30, m	85.3, CH	H-2, H-4	2, 4, 5	3.79, m	81.1, CH
	5	3.15, m	58.1, CH	H-6a, H-6b		n.o.	n.o.
	6a	1.69, m	25.1, CH$_2$	H-5, H-7a	3, 4, 6, 7	1.66, m	25.2, CH$_2$
	6b	1.22, m	25.1, CH$_2$	H-5, H-7a, H-7b	4, 5, 7	1.28, m	25.2, CH$_2$
	7a	1.79, m	24.2, CH$_2$	H-6a, H-6b, H-7b		1.72, m	23.1, CH$_2$
	7b	1.48, m	24.2, CH$_2$	H-7a		1.45, m	23.1, CH$_2$
	8a	3.54, m	46.1, CH$_2$	H-7a, H-7b	1 (Dil) 4, 5, 6	3.45, m	47.1, CH$_2$
	8b	3.02, m	46.1, CH$_2$	H-7a, H-7b		3.20, m	47.1, CH$_2$
	9 (O-Me)	3.24, s	60.7, CH$_3$		3	3.21, s	60.0, CH$_3$
Dil	1		168.8, qC				168.6, qC
	2a	2.35, m	34.8, CH$_2$	H-2b, H-3	1, 3, 4	2.43, m	37.2, CH$_2$
	2b	2.17, m	34.8, CH$_2$	H-2a, H-3	1, 3	2.25, m	37.2, CH$_2$
	3	3.98, m	77.0, CH	H-2a, H-2b, H-4	1, 2, 4	n.o.	n.o.
	4	4.73, m	54.7, CH	H-3, H-5	1 (Val), 2, 3, 5	4.64, m	55.5, CH
	5	1.83, m	31.3, CH	H-4, H-6, H-7b		1.76, m	32.3, CH
	6	0.87, m	15.1, CH$_3$	H-5	4, 5, 7	0.89, m	15.3, CH$_3$
	7a	1.30, m	25.0, CH$_2$	H-7b, H-5	6	n.o.	n.o.
	7b	0.91, m	25.0, CH$_2$	H-7a, H-8		n.o.	n.o.
	8	0.76, m	10.2, CH$_3$	H-7a, H-7b	5, 7	n.o.	n.o.
	9 (O-Me)	3.20, s	56.9, CH$_3$		3	3.17, s	56.8, CH$_3$
	10 (N-Me)	3.09, s	31.0, CH$_3$		1 (Val), 4	3.00, s	31.3, CH$_3$
Val	1		172.9, qC				173.0, qC
	2	4.64, m	53.5, CH	H-3, NH	1 (N-diMe-Ile), 1, 3, 4, 5	4.54, m	53.7, CH
	3	2.01, m	29.6, CH	H-2, H-4, H-5	2, 4, 5	1.95, m	29.6, CH
	4	0.96, m	18.8, CH$_3$	H-3	2, 3, 5	0.92, m	18.7, CH$_3$
	5	0.96, m	18.8, CH$_3$	H-3	2, 3, 4	0.92, m	18.7, CH$_3$
	NH	8.02, br		H-2		8.03, br	
N-diMe-Ile	1		169.7, qC				169.8, qC
	2	2.79, m	70.4, CH	H-3		n.o.	n.o.
	3	1.76, m	32.3, CH	H-1, H-4, H-5b		n.o.	n.o.
	4	0.71, m	15.1, CH$_3$	H-3	2, 3, 5	n.o.	n.o.
	5a	1.57, m	24.6, CH$_2$	H-3, H-5b, H-6	2, 3, 4	n.o.	n.o.
	5b	1.08, m	24.6, CH$_2$	H-3, H-5a, H-6		n.o.	n.o.
	6	0.83, m	10.2, CH$_3$	H-5a, H-5b	3, 5	n.o.	n.o.
	7 (N-Me)	2.21, m	41.6, CH$_3$			n.o.	n.o.
	8 (N-Me)	2.21, m	41.6, CH$_3$			n.o.	n.o.

Table A.2.2.: Comparison of the ^{13}C-NMR spectroscopic data of AEST (**1**) (spectra recorded in CD$_2$Cl$_2$, 150 MHz) and Symplostatin 1. n.o. = not observable due to solvent signal.

unit	C no.	AEST δ_C	Symplostatin 1 δ_C
Aph	1	n.o.	
	2	38.7	
	3	139.4	
	4	129.9	
	5	128.8	
	6	126.6	
	7	76.3	
	8	16.1	
	9	62.0	
Dap	1	174.1	174.0
	2	45.4	44.7
	3	15.0	14.5
	4	82.4	81.9
	5	59.9	59.7
	6	25.3	24.9
	7	25.5	25.4
	8	48.2	48.0
	9 (O-Me)	61.1	60.9
Dil	1	170.1	170.4
	2	38.1	37.9
	3	78.7	78.7
	4	57.1	57.2
	5	33.8	33.4
	6	15.8	15.9
	7	26.2	26.1
	8	10.8	10.8
	9 (O-Me)	58.3	58.1
	10 (N-Me)	32.2	32.3
Val	1	173.7	173.4
	2	54.7	54.5
	3	31.5	31.2
	4	18.2	18.3
	5	20.0	19.6
N-diMe-Ile	1	174.2	174.0
	2	75.4	74.5
	3	34.9	34.5
	4	15.1	15.0
	5	26.4	26.8
	6	12.1	11.8
	7 (N-Me)	43.3	42.6

Table A.2.3.: One- and two-dimensional NMR data of both conformers of MMAEST (**2**) in DMSO-d_6. n.o = not observed.

| unit | C/H no. | Conformer 1 | | | | Conformer 2 | |
		δH (J in Hz)	δC, mult	COSY	^{1}H-^{13}C	δH (J in Hz)	δC, mult
Aph	1	4.22, m	51.5, CH	H-2a, H-2b, H-7, NH	1 (Dap), 2, 3, 7, 8	4.11, m	52.3, CH
	2a	2.84, m	35.0, CH$_2$	H-1, H-2b	1, 3, 4, 7	2.80, m	34.7, CH$_2$
	2b	2.70, m	35.0, CH$_2$	H-1, H-2a	1, 3, 4, 7	2.65, m	34.7, CH$_2$
	3		139.0, qC				139.3, qC
	4a	7.20, m	128.6, CH	H-5a	2, 3, 5, 6	7.19, m	128.9, CH
	4b	7.20, m	128.6, CH	H-5b	2, 3, 5, 6	7.19, m	128.9, CH
	5a	7.19, m	127.8, CH	H-4a, H-6	3, 4, 6	7.20, m	128.0, CH
	5b	7.19, m	127.8, CH	H-4b, H-6	3, 4, 6	7.20, m	128.0, CH
	6	7.11, m	125.8, CH	H-5a, H-5b	4, 5	7.11, m	125.7, CH
	7	3.37, m	77.5, CH	H-1, H-8	1, 2, 8	3.32, m	76.9, CH
	8	1.06, m	15.0, CH$_3$	H-7	1, 7	1.06, m	14.8, CH$_3$
	9 (O-Me)	3.32, s	56.0, CH$_3$		7	3.30, s	56.0, CH$_3$
	NH	7.88, d (9.2)		H-1	1 (Dap), 1, 2	7.64, d (8.8)	
Dap	1		173.6, qC				173.2, qC
	2	2.24, m	43.0, CH	H-3, H-8	3, 4, 8	2.22, m	43.6, CH
	3	1.06, m	15.9, CH$_3$	H-2	1, 2	1.06, m	15.5, CH$_3$
	4	3.30, m	85.3, CH	H-2, H-5	2, 5	3.79, m	81.3, CH
	5	3.15, m	58.2, CH	H-6a, H-6b		3.20, m	55.0, CH
	6a	1.68, m	25.1, CH$_2$	H-6b	4, 5, 7, 8	1.66, m	25.0, CH$_2$
	6b	1.22, m	25.1, CH$_2$	H-5, H-7a, H-7b	4, 5, 7	1.28, m	25.0, CH$_2$
	7a	1.78, m	24.0, CH$_2$	H-6a, H-6b, H-8a, H-8b	5, 6, 8	1.72, m	22.9, CH$_2$
	7b	1.48, m	24.0, CH$_2$	H-6a, H-6b, H-8a, H-8b		1.41, m	22.9, CH$_2$
	8a	3.53, m	46.2, CH$_2$	H-7a, H-7b	1 (Dil), 5, 6, 7	3.44, m	47.2, CH$_2$
	8b	3.02, m	46.2, CH$_2$	H-7a, H-7b		3.20, m	47.2, CH$_2$
	9 (O-Me)	3.23, s	60.9, CH$_3$		4	3.20, s	60.3, CH$_3$
Dil	1		168.9, qC				168.6, qC
	2a	2.35, m	34.8, CH$_2$	H-2b, H-3	1, 3	2.43, m	36.9, CH$_2$
	2b	2.17, m	34.8, CH$_2$	H-2a, H-3	1, 3	2.25, m	36.9, CH$_2$
	3	3.98, m	77.0, CH	H-2a, H-2b, H-4	1, 4	n.o.	n.o.
	4	4.73, m	54.7, CH	H-3, H-5	5	4.62, m	55.5, CH
	5	1.83, m	31.3, CH	H-4, H-6, H-7b		1.77, m	31.8, CH
	6	0.86, m	15.2, CH$_3$	H-5	4, 5, 7	0.88, m	15.3, CH$_3$
	7a	1.29, m	25.1, CH$_2$	H-7b, H-5	6	n.o.	n.o.
	7b	0.91, m	25.1, CH$_2$	H-7a, H-5		n.o.	n.o.
	8	0.77, m	10.0, CH$_3$	H-7a, H-7b	5, 7	n.o.	n.o.
	9 (O-Me)	3.20, s	57.0, CH$_3$		3	3.17, s	56.8, CH$_3$
	10 (N-Me)	3.05, s	31.1, CH$_3$		1 (Val), 4	2.97, s	31.3, CH$_3$
Val	1		172.6, qC				172.8, qC
	2	4.67, m	53.3, CH	H-3, NH	1, 3, 4, 5	4.54, m	53.5, CH
	3	2.04, m	29.9, CH	H-2, H-4, H-5		1.95, m	29.9, CH
	4	0.94, m	18.9, CH$_3$	H-3	2, 3, 5	n.o.	n.o.
	5	0.94, m	18.9, CH$_3$	H-3	2, 3, 4	n.o.	n.o.
	NH	8.03, br		H-2		8.02, br	
N-diMe-Ile	1		169.8, qC				169.8, qC
	2	2.77, m	68.1, CH	H-3		n.o.	n.o.
	3	1.50, m	37.4, CH	H-1, H-4, H-5b		n.o.	n.o.
	4	0.80, m	15.2, CH$_3$	H-3	2, 3, 5	n.o.	n.o.
	5a	1.54, m	24.8, CH$_2$	H-3, H-5b, H-6	2, 3, 4	n.o.	n.o.
	5b	1.10, m	24.8, CH$_2$	H-3, H-5a, H-6	3, 4, 6	n.o.	n.o.
	6	0.82, m	11.1, CH$_3$	H-5a, H-5b	3, 5	n.o.	n.o.
	7 (N-Me)	2.17, s	34.5, CH$_3$			2.15, s	34.4 CH$_3$

Table A.2.4.: One- and two-dimensional NMR data of both conformers of DAAEST (**3**) in DMSO-d_6. n.o = not observed.

| unit | C/H no. | Conformer 1 | | | | Conformer 2 | |
		δH (J in Hz)	δC, mult	COSY	^{1}H-^{13}C	δH (J in Hz)	δC, mult
Dap	1		175.8, qC				n.o.
	2	2.29, m	42.6, CH	H-3, H-4	1, 3, 4	n.o.	n.o.
	3	1.09, m	13.0, CH$_3$	H-2	1, 2, 4	1.12, m	14.6, CH$_3$
	4	3.83, m	81.1, CH	H-2, H-5	1, 2, 3, 5, 6, 9	n.o.	n.o.
	5	4.00, m	58.6, CH	H-4		4.05, m	58.8, CH
	6a	1.91, m	24.5, CH$_2$	H-5, H-6b	4, 5, 7, 8	n.o.	n.o.
	6b	1.77, m	24.5, CH$_2$	H-5, H-6a H-7a,	4, 7	n.o.	n.o.
	7a	1.93, m	23.9, CH$_2$	H-6a, H-6b, H-8a, H-8b	4, 5, 8	n.o.	n.o.
	7b	1.74, m	23.9, CH$_2$	H-6a, H-6b, H-8a, H-8b	4	n.o.	n.o.
	8a	3.52, m	46.7, CH$_2$	H-7a, H-7b	7	3.61, m	45.8, CH$_2$
	8b	3.35, m	46.7, CH$_2$	H-7a, H-7b	7	3.15, m	45.8, CH$_2$
	9 (O-Me)	3.28, s	59.7, CH$_3$		4	3.30, s	60.4, CH$_3$
Dil	1		169.2, qC				n.o.
	2a	2.45, m	36.8, CH$_2$	H-2b, H-3	1, 3	2.43, m	36.7, CH$_2$
	2b	2.34, m	36.8, CH$_2$	H-2a, H-3	1	2.31, m	36.7, CH$_2$
	3	3.98, m	77.5, CH	H-2a, H-2b, H-4	1, 7	n.o.	n.o.
	4	4.63, m	55.5, CH	H-3, H-5		4.67, m	55.2, CH
	5	1.76, m	31.9, CH	H-4, H-6, H-7b	7	1.81, m	31.7, CH
	6	0.90, m	14.9, CH$_3$	H-5	4, 5, 7	n.o.	n.o.
	7a	1.30, m	25.0, CH$_2$	H-7b, H-5, H-8		n.o.	n.o.
	7b	0.90, m	25.0, CH$_2$	H-7a, H-8		n.o.	n.o.
	8	0.76, m	10.1, CH$_3$	H-7a, H-7b	5, 7	n.o.	n.o.
	9 (O-Me)	3.19, s	56.9, CH$_3$		3	3.18, s	56.9, CH$_3$
	10 (N-Me)	3.00, s	31.3, CH$_3$		1 (Val), 3, 4	3.05, s	31.3, CH$_3$
Val	1		173.0, qC				172.9, qC
	2	4.54, m	53.5, CH	H-3, NH	1(N-diMe-Ile), 1, 3, 4, 5	n.o.	n.o.
	3	1.95, m	29.9, CH	H-2, H-4, H-5	2, 4, 5	n.o.	n.o.
	4	0.92, m	18.6, CH$_3$	H-3	2, 3, 5	n.o.	n.o.
	5	0.92, m	18.6, CH$_3$	H-3	2, 3, 4	n.o.	n.o.
	NH	8.02, br		H-2		n.o.	
N-diMe-Ile	1		169.8, qC				169.8, qC
	2	2.77, m	70.4, CH	H-3	1, 3, 4, 5, 7, 8	n.o.	n.o.
	3	1.74, m	32.2, CH	H-1, H-4, H-5b	5	n.o.	n.o.
	4	0.69, m	14.9, CH$_3$	H-3	2, 3, 5	n.o.	n.o.
	5a	1.57, m	24.5, CH$_2$	H-3, H-5b, H-6	2, 3, 4, 6	n.o.	n.o.
	5b	1.07, m	24.5, CH$_2$	H-3, H-5a, H-6	3, 4, 6	n.o.	n.o.
	6	0.83, m	10.1, CH$_3$	H-5a, H-5b	3, 5	n.o.	n.o.
	7 (N-Me)	2.21, s	41.3, CH$_3$		2, 8	n.o.	n.o.
	8 (N-Me)	2.21, s	41.3, CH$_3$		2, 7	n.o.	n.o.

Table A.2.5.: One- and two-dimensional NMR data of the degraded AEST derivative (**4**) in DMSO-d_6.

unit	C/H no.	δH (J in Hz)	δC, mult	COSY	^{1}H-^{13}C
Aph	1	4.24, m	51.4, CH	H-2a, H-2b, H-7, NH	
	2a	2.85, m	35.0, CH$_2$	H-1, H-2b	1, 3, 4, 7
	2b	2.70, m	35.0, CH$_2$	H-1, H-2a	1, 3, 4, 7
	3		139.0, qC		
	4a	7.20, m	128.4, CH	H-5a	2, 3, 5, 6
	4b	7.20, m	128.4, CH	H-5b	2, 3, 5, 6
	5a	7.20, m	127.8, CH	H-4a, H-6	2, 3, 4, 6
	5b	7.20, m	127.8, CH	H-4b, H-6	2, 3, 4, 6
	6	7.11, m	126.3, CH	H-5a, H-5b	4, 5
	7	3.37, m	77.2, CH	H-1, H-8	2, 8, 9
	8	1.06, m	14.5, CH$_3$	H-7	7
	9 (O-Me)	3.32, s	55.8, CH$_3$		7
	NH	7.86, d (9.2)		H-1	1 (Dap)
Dap	1		173.4, qC		
	2	2.25, m	42.7, CH	H-3, H-4	1, 3, 4
	3	1.07, m	15.6, CH$_3$	H-2	1, 2, 4
	4	3.31, m	85.0, CH	H-2, H-5	2, 3, 5
	5	3.16, m	57.9, CH	H-6a, H-6b	
	6a	1.68, m	24.9, CH$_2$	H-5, H-7a, H-7b	
	6b	1.22, m	24.9, CH$_2$	H-5, H-7a, H-7b	
	7a	1.74, m	22.8, CH$_2$	H-6b, H-8a, H-8b	
	7b	1.42, m	22.8, CH$_2$	H-6b, H-8a, H-8b	
	8a	3.54, m	45.9, CH$_2$	H-7a, H-7b	6, 7
	8b	3.02, m	45.9, CH$_2$	H-7a, H-7b	1(Dil), 6, 7
	9 (O-Me)	3.24, s	60.6, CH$_3$		4
Dil	1		168.7, qC		
	2a	2.35, m	34.7, CH$_2$	H-3	1, 3
	2b	2.19, m	34.7, CH$_2$	H-3	1, 3
	3	3.98, m	77.0, CH	H-2a, H-2b, H-4	
	4	4.72, m	54.7, CH	H-3, H-5	
	5	1.81, m	31.2, CH	H-4, H-6, H-7b	
	6	0.86, m	15.1, CH$_3$	H-5	4, 5, 7
	7a	1.32, m	25.0, CH$_2$	H-7b, H-5	
	7b	0.94, m	25.0, CH$_2$	H-7a, H-8	
	8	0.78, m	10.0, CH$_3$	H-7a, H-7b	5, 7
	9 (O-Me)	3.20, s	56.7, CH$_3$		3
	10 (N-Me)	3.08, s	31.0, CH$_3$		1 (Val), 4
Val	1		172.7, qC		
	2	4.62, m	53.4, CH	H-3, NH	1(N-diMe-Ile), 1, 3, 4, 5
	3	2.06, m	29.6, CH	H-2, H-4, H-5	2, 4, 5
	4	0.88, m	18.1, CH$_3$	H-3	2, 3, 5
	5	0.88, m	18.1, CH$_3$	H-3	2, 3, 4
	NH	7.86, br		H-2	
N-diMe-Ile	1		165.9, qC		
	2		154.4, qC		
	3	5.84, m	117.3, CH	H-4	4, 5
	4	2.06, m	29.6, CH$_3$		1, 2, 5
	5	2.06, m	32.5, CH$_2$	H-6	2, 3, 4, 6
	6	0.99, m	19.1, CH$_3$	H-5	2, 5
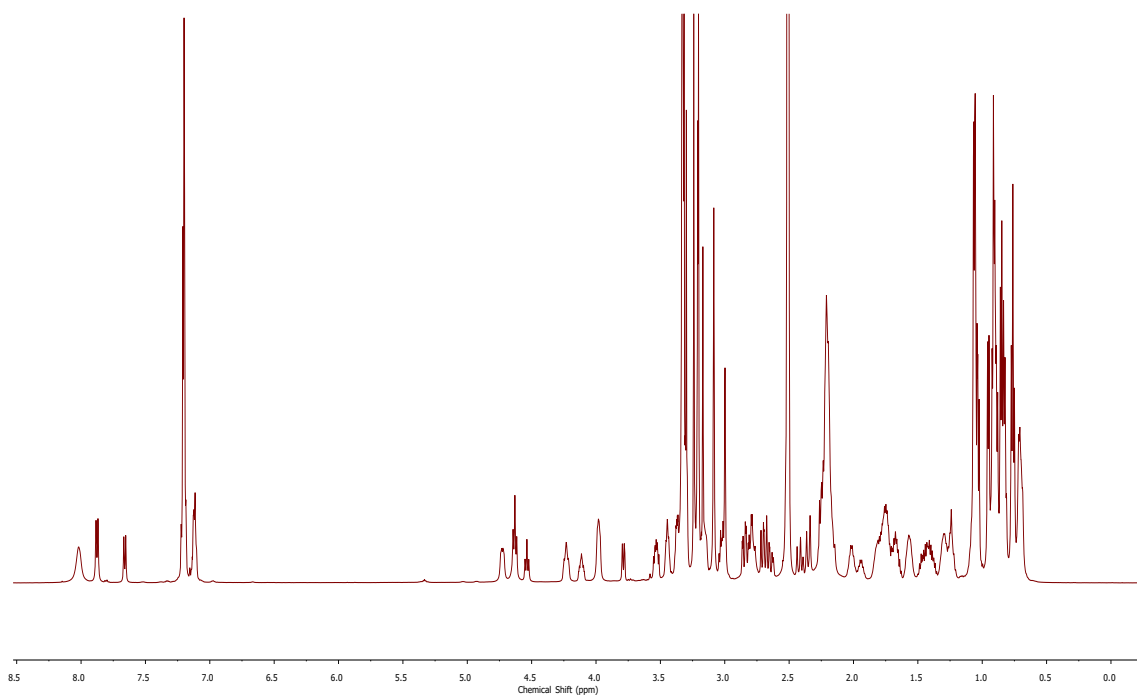
Figure A.2.13.: ^{1}H-NMR spectrum of AEST (**1**) in DMSO-d_6 (600 MHz). See **Tab. A1** for a list of the observed chemical shifts.
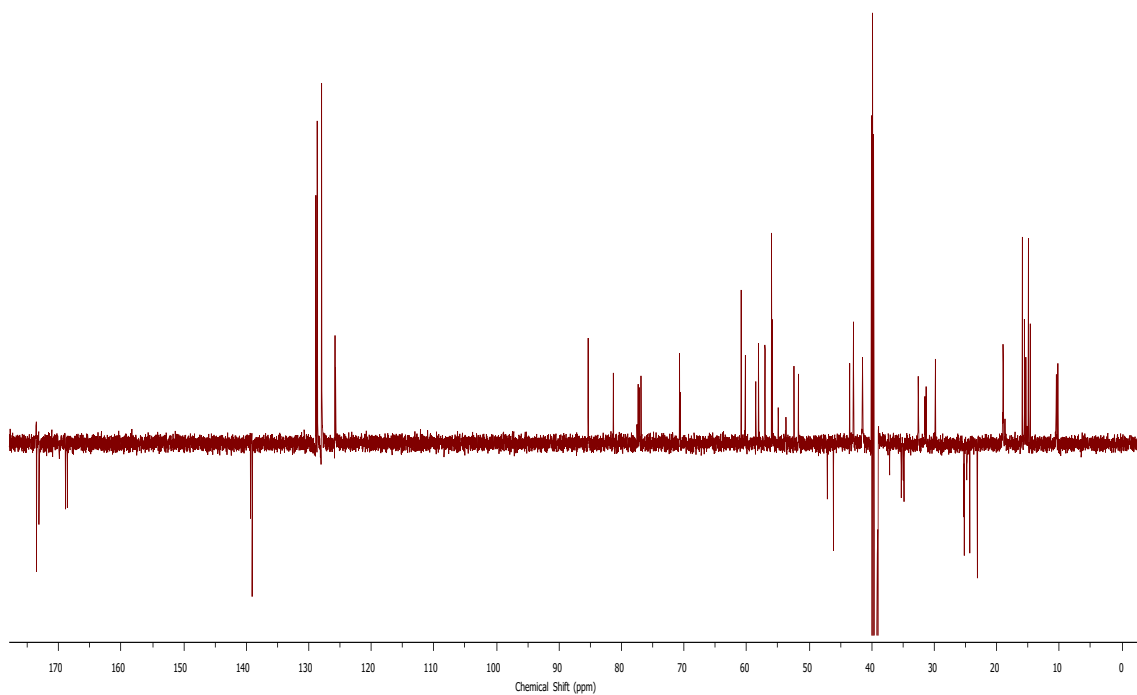
Figure A.2.14.: ^{13}C-APT NMR spectrum of AEST (**1**) in DMSO-d_6 (150 MHz). See **Tab. A1** for a list of the observed chemical shifts.
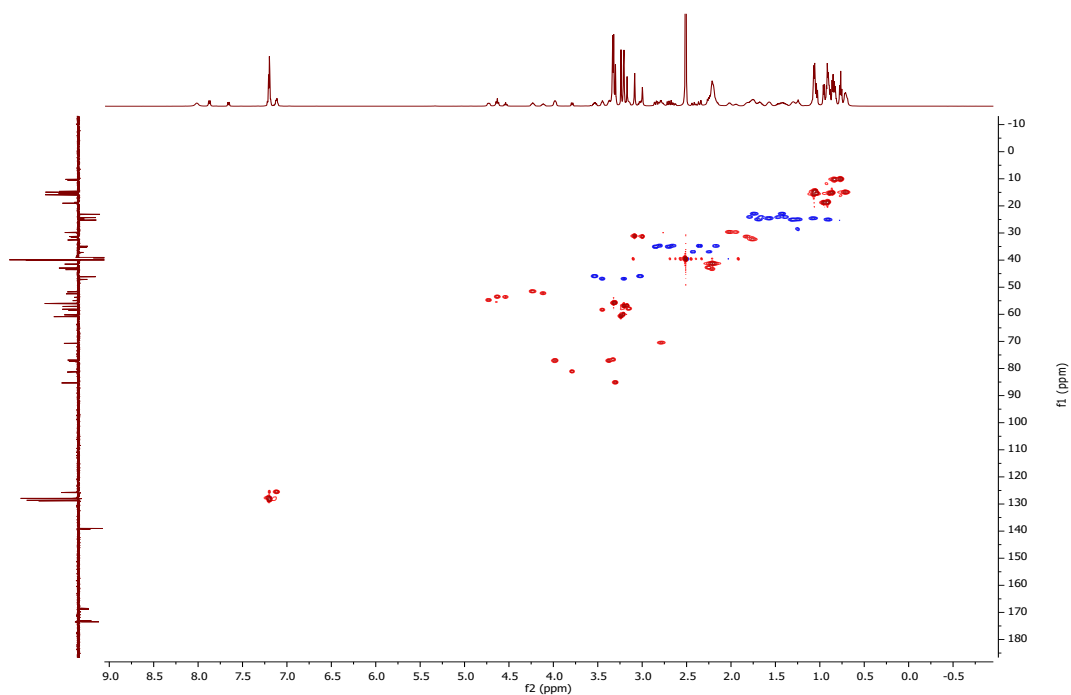
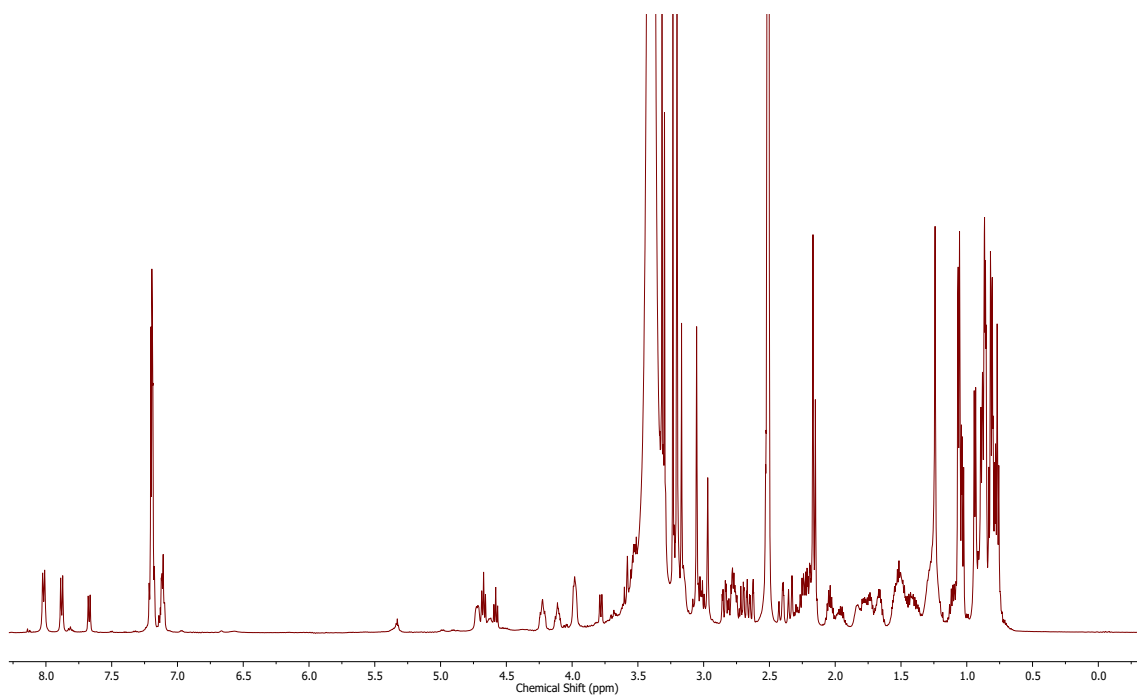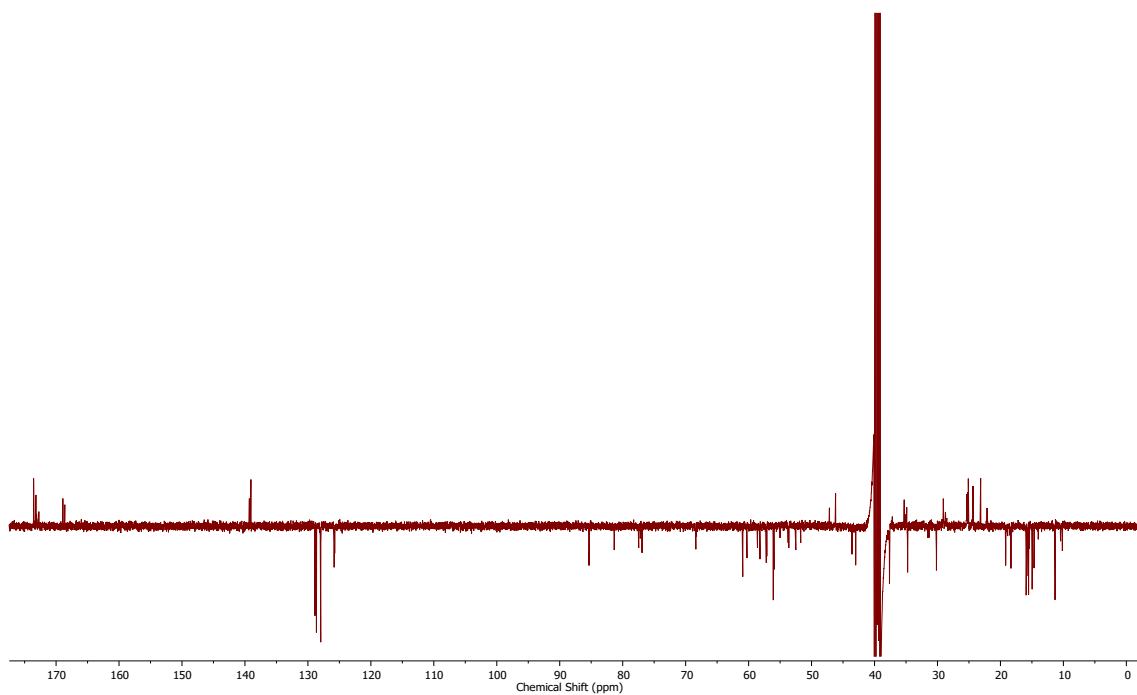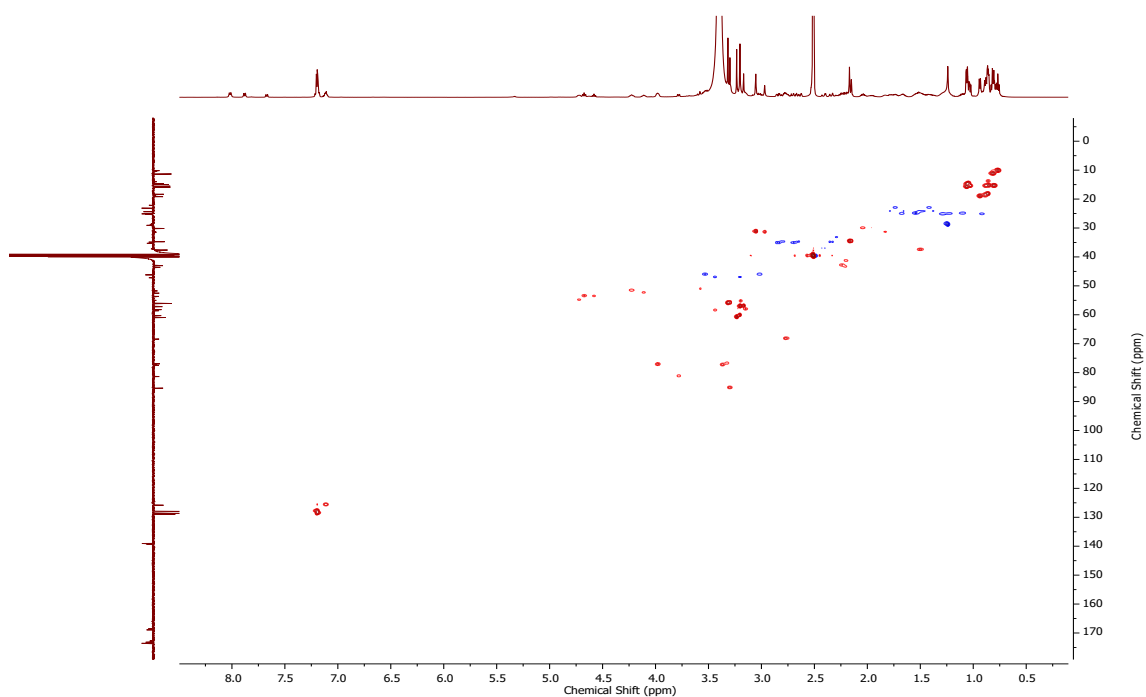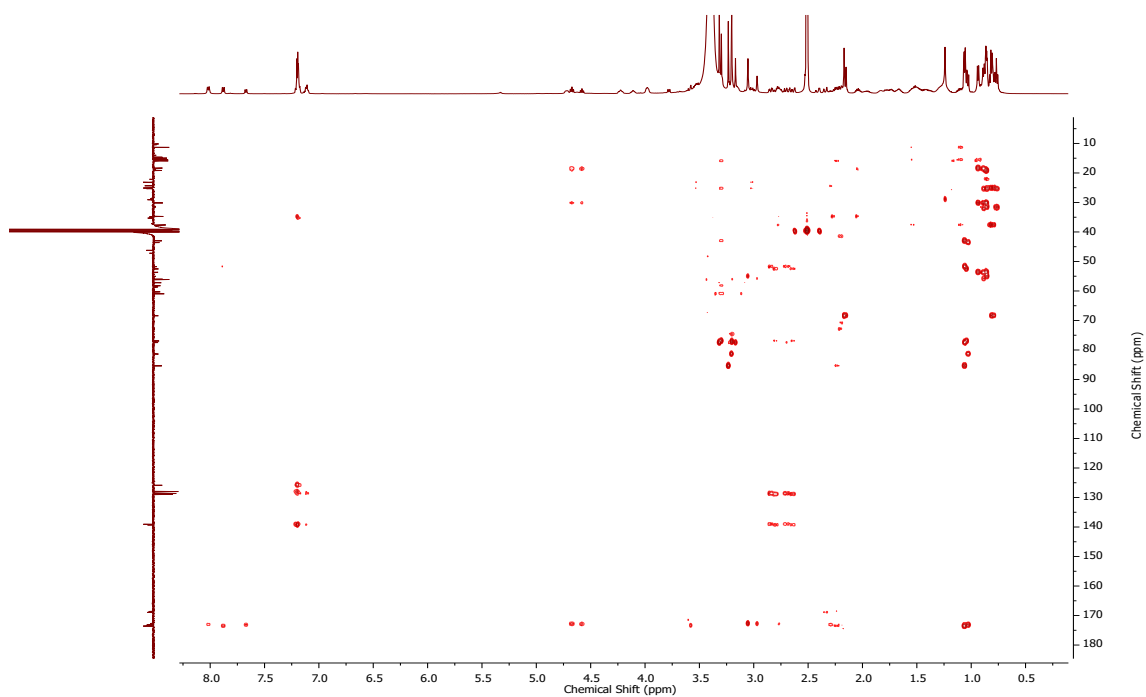**Figure A.2.15.:** ^{13}C-HSQC NMR spectrum of AEST (**1**) in DMSO-d_6 (600 MHz). See **Tab. A1** for a list of the observed chemical shifts.
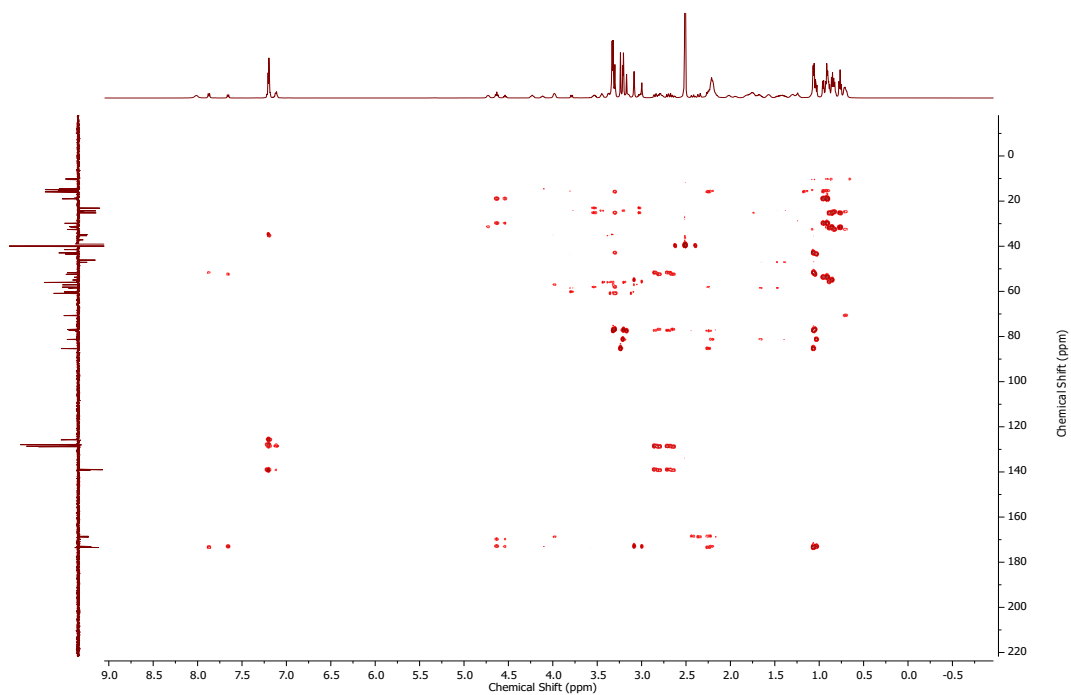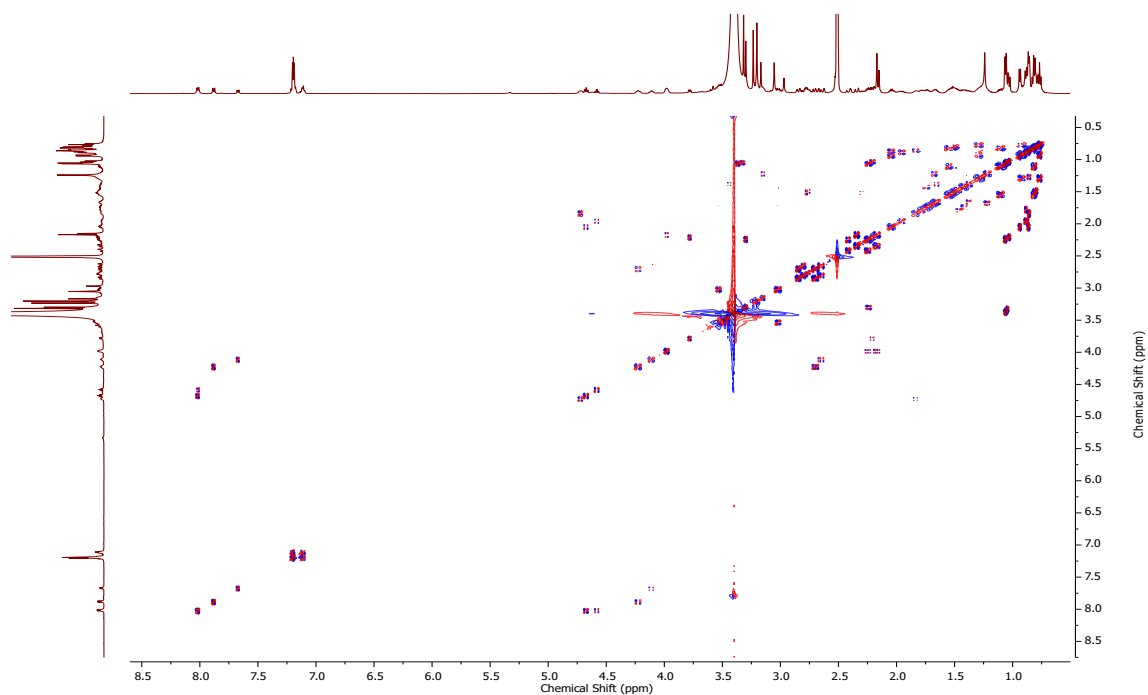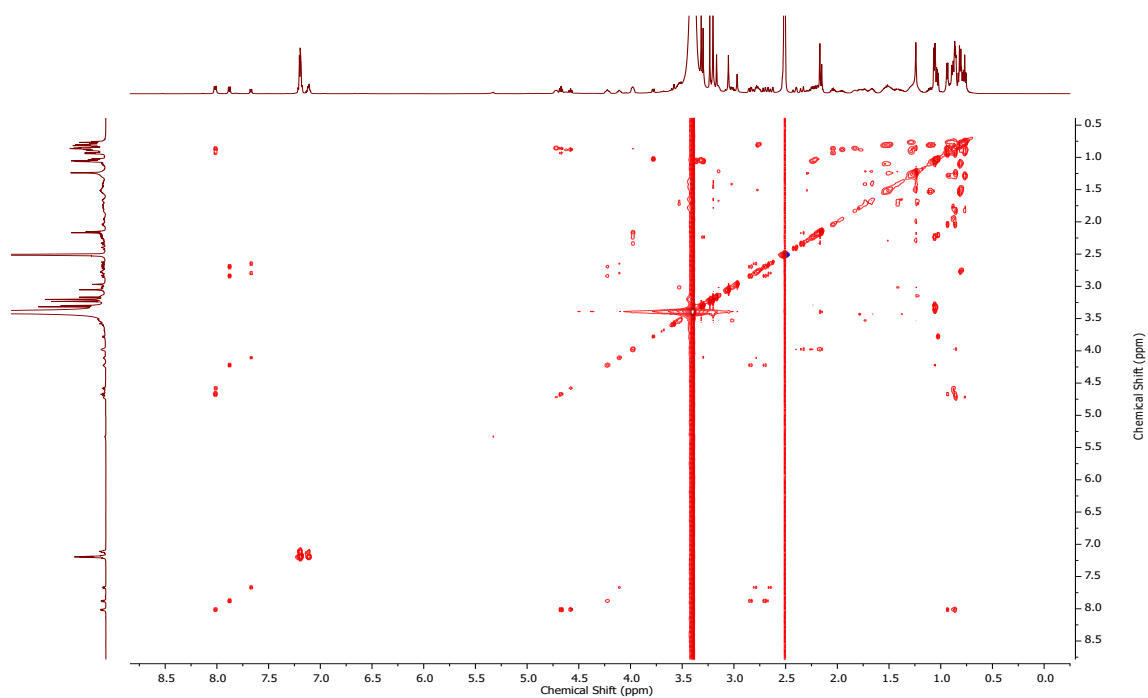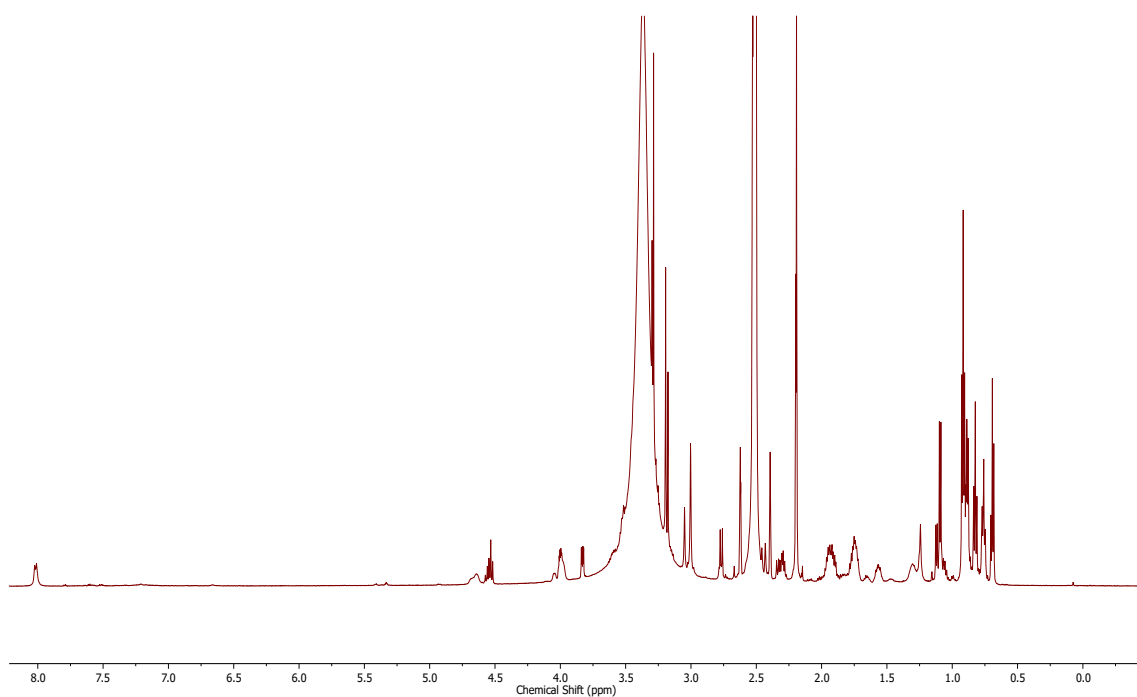

Figure A.2.16.: ^{13}C-HMBC NMR spectrum of AEST (**1**) in DMSO-d_6 (600 MHz). See **Tab. A1** for a list of the observed chemical shifts.
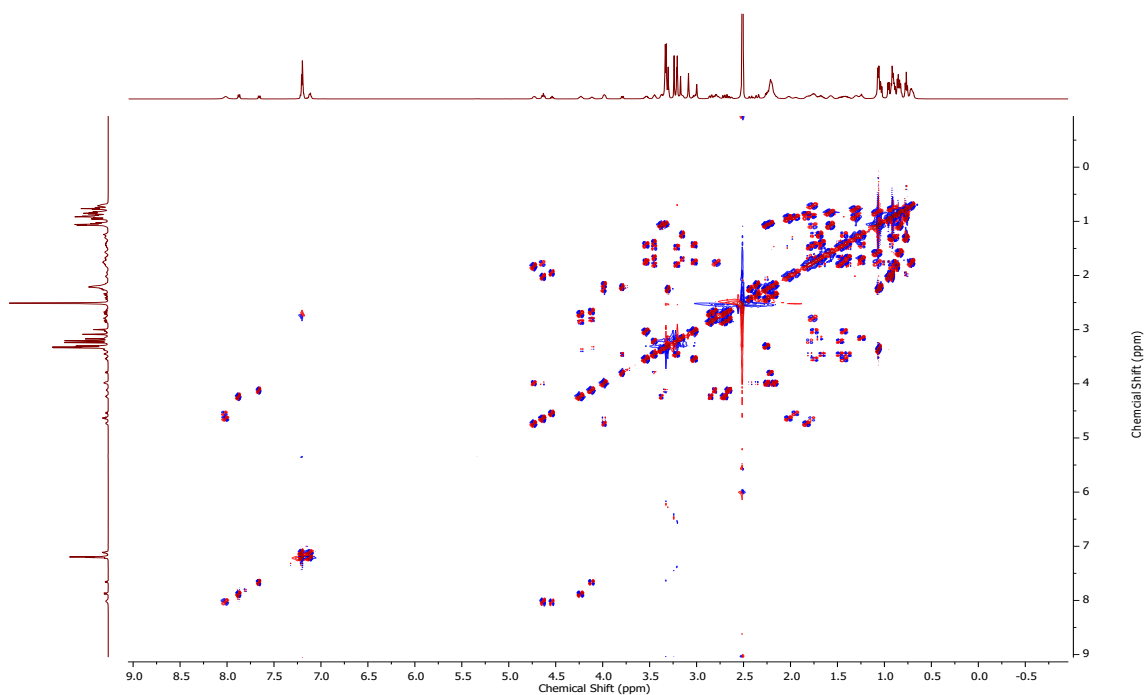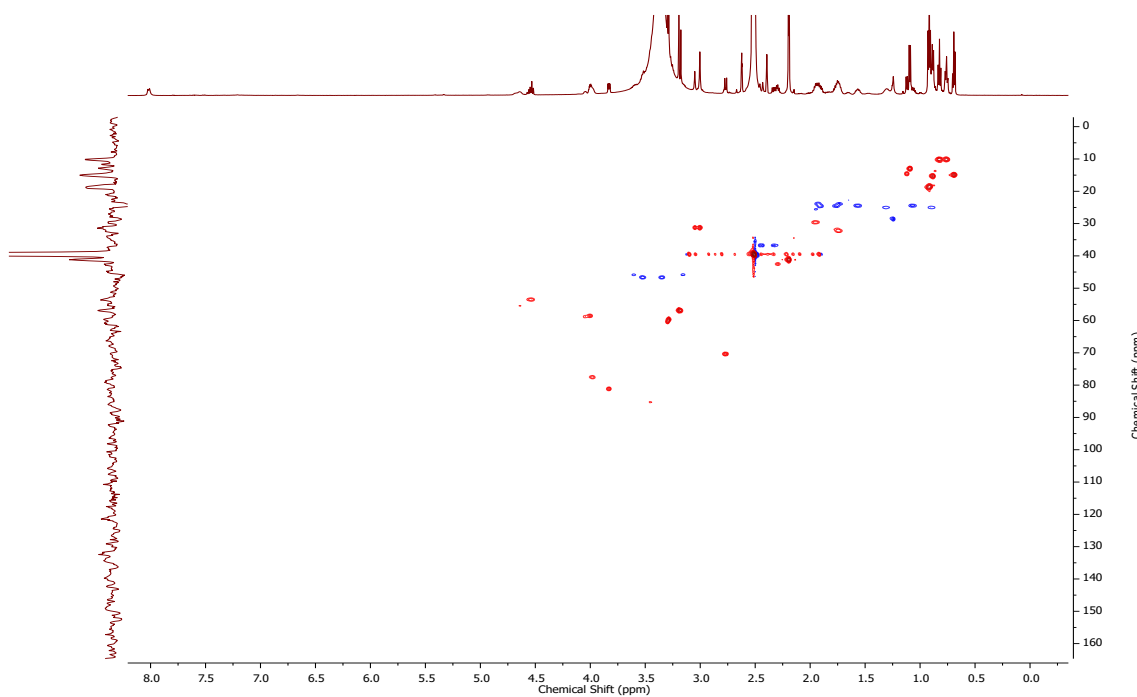
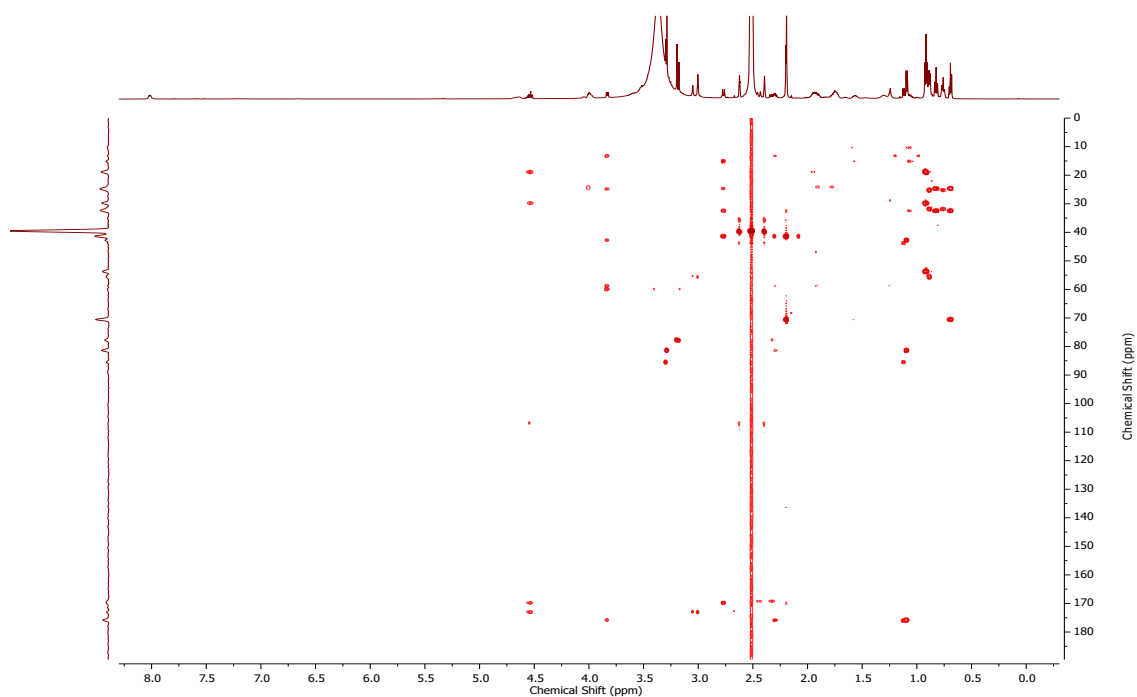**Figure A.2.17.:** DQF-COSY NMR spectrum of AEST (**1**) in DMSO-d_6 (600 MHz). See **Tab. A1** for a list of the observed chemical shifts.
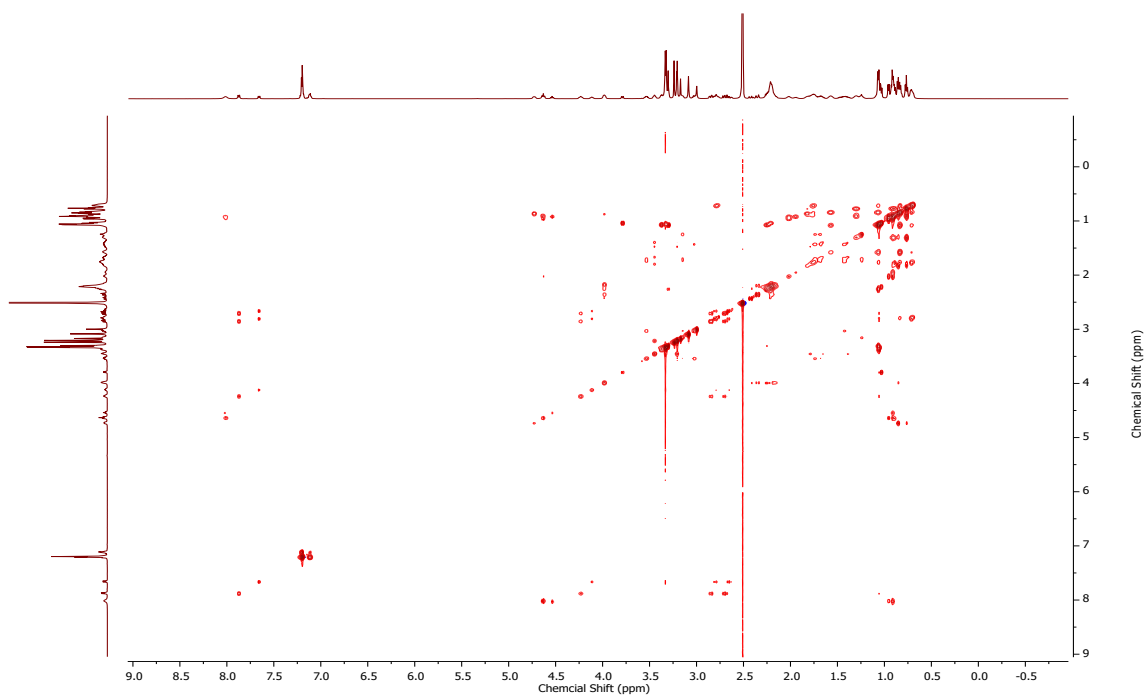
Figure A.2.18.: TOCSY NMR spectrum of AEST (**1**) in DMSO-d_6 (600 MHz).
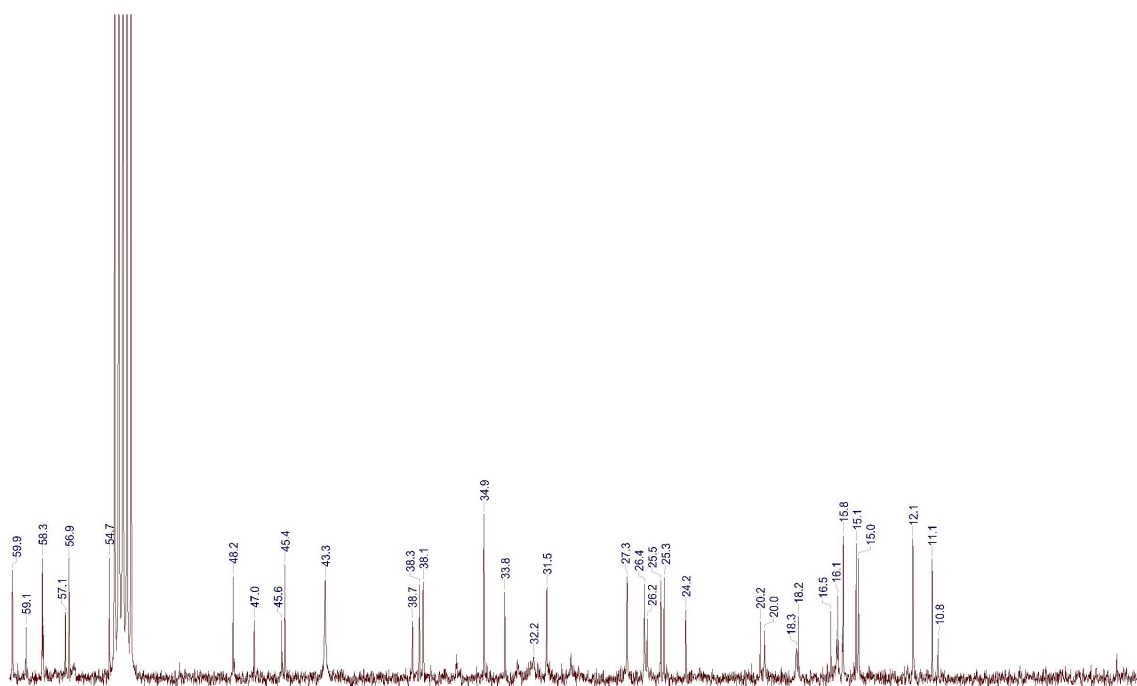
Figure A.2.19.: ¹³C-NMR spectrum of AEST (**1**) in CD₂Cl₂ (150 MHz) from 0 to 60 ppm.
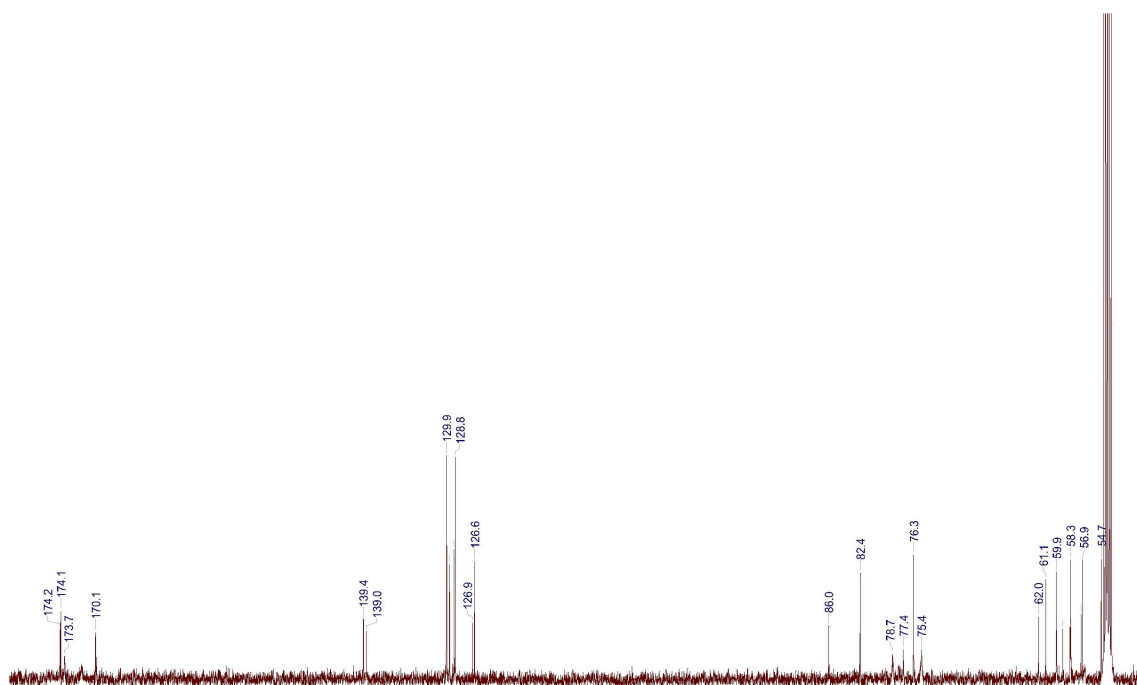
Figure A.2.20.: ¹³C-NMR spectrum of AEST (**1**) in CD₂Cl₂ (150 MHz) from 50 to 180 ppm.

Figure A.2.21.: ^{1}H-NMR spectrum of MMAEST (**2**) in DMSO-d_6 (600 MHz). See **Tab. A3** for a list of the observed chemical shifts.


Figure A.2.22.: ^{13}C-APT NMR spectrum of MMAEST (**2**) in DMSO-d_6 (150 MHz). See **Tab. A3** for a list of the observed chemical shifts.

Figure A.2.23.: ^{13}C-HSQC NMR spectrum of MMAEST (**2**) in DMSO-d_6 (600 MHz). See **Tab. A3** for a list of the observed chemical shifts.

Figure A.2.24.: ^{13}C-HMBC NMR spectrum of MMAEST (**2**) in DMSO-d_6 (600 MHz). See **Tab. A3** for a list of the observed chemical shifts.

Figure A.2.25.: DQF-COSY NMR spectrum of MMAEST (**2**) in DMSO-d_6 (600 MHz). See **Tab. A3** for a list of the observed chemical shifts.

**Figure A.2.26.:** TOCSY NMR spectrum of MMAEST (**2**) in DMSO-d_6 (600 MHz).

Figure A.2.27.: ^{1}H-NMR spectrum of DAAEST (**3**) in DMSO-d_6 (600 MHz). See **Tab. A4** for a list of the observed chemical shifts.

Figure A.2.28.: ^{13}C-HSQC NMR spectrum of DAAEST (**3**) in DMSO-d_6 (600 MHz). See **Tab. A4** for a list of the observed chemical shifts.

Figure A.2.29.: ^{13}C-HMBC NMR spectrum of DAAEST (**3**) in DMSO-d_6 (600 MHz). See **Tab. A4** for a list of the observed chemical shifts.
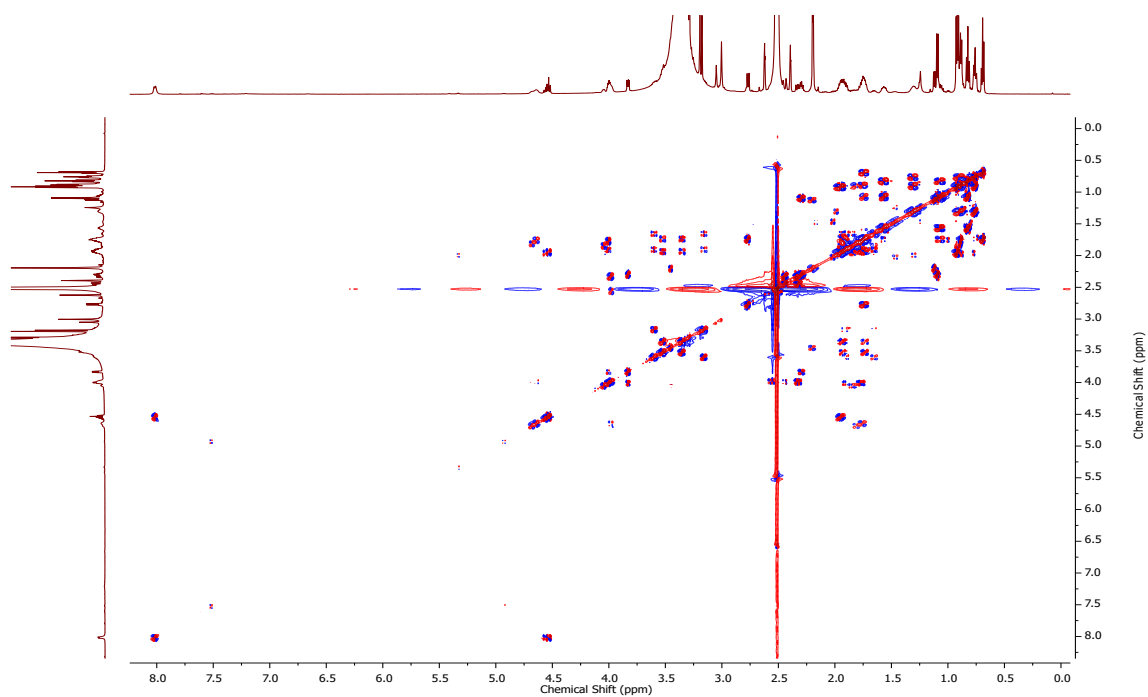
Figure A.2.30.: DQF-COSY NMR spectrum of DAAEST (**3**) in DMSO-d_6 (600 MHz). See **Tab. A4** for a list of the observed chemical shifts.
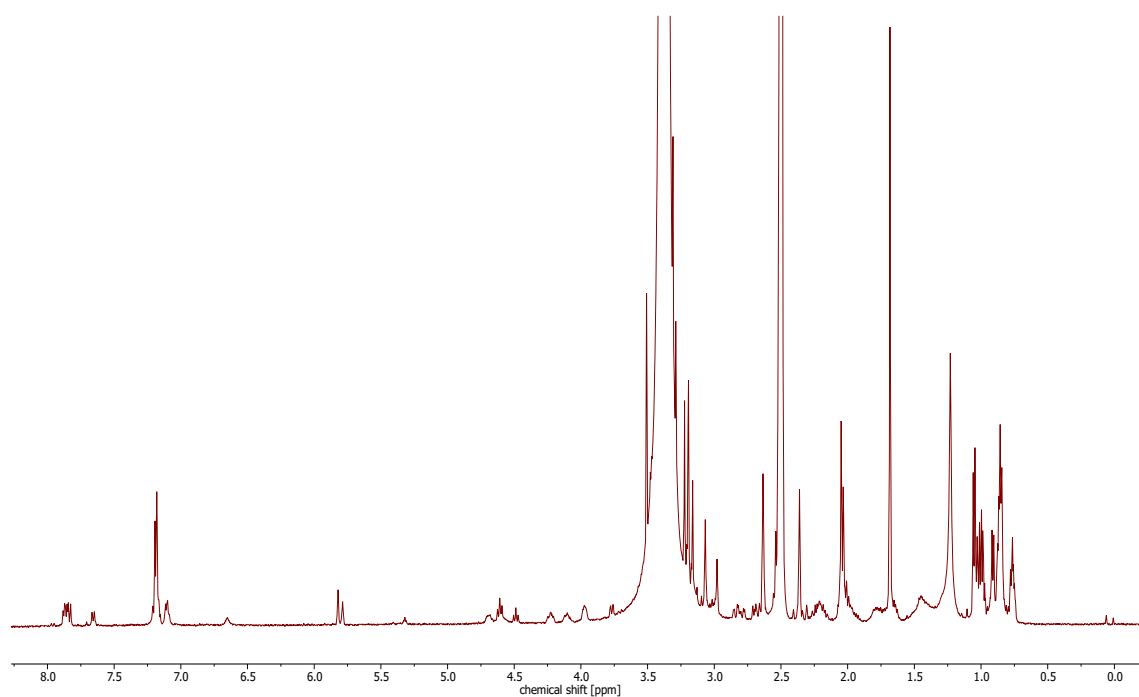
Figure A.2.31.: ¹H-NMR spectrum of the degraded AEST derivative (**4**) in DMSO-d_6 (600 MHz). See **Tab. A5** for a list of the observed chemical shifts.
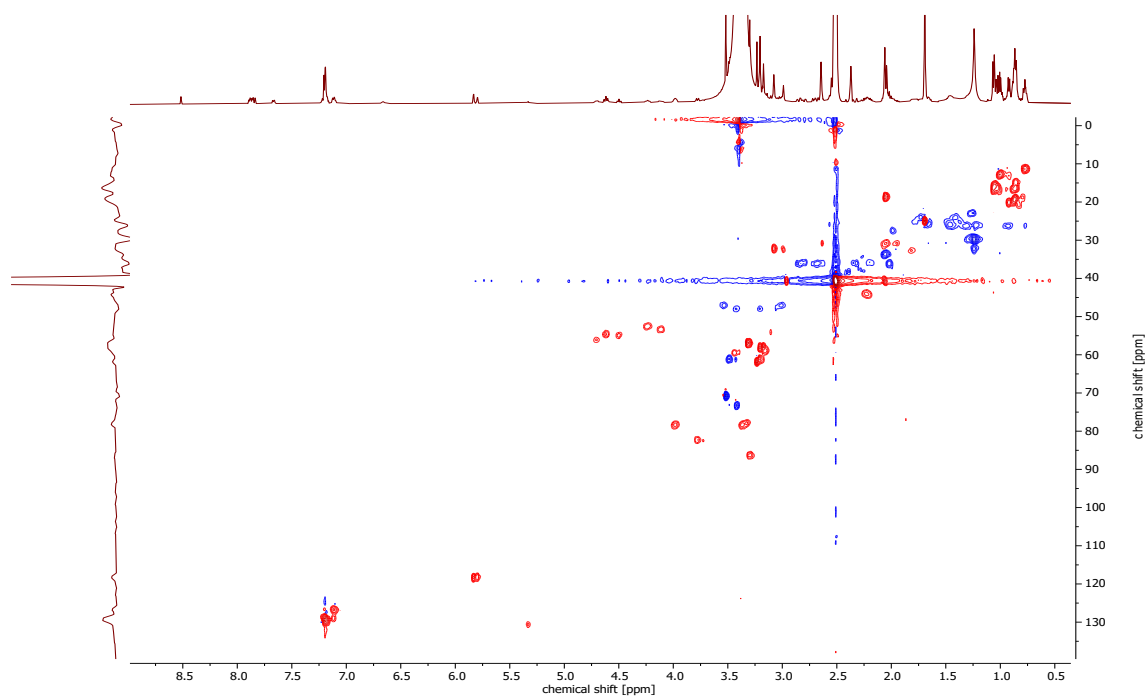
Figure A.2.32.: ¹³C-HSQC NMR spectrum of the degraded AEST derivative (**4**) in DMSO-d_6 (600 MHz). See **Tab. A5** for a list of the observed chemical shifts.
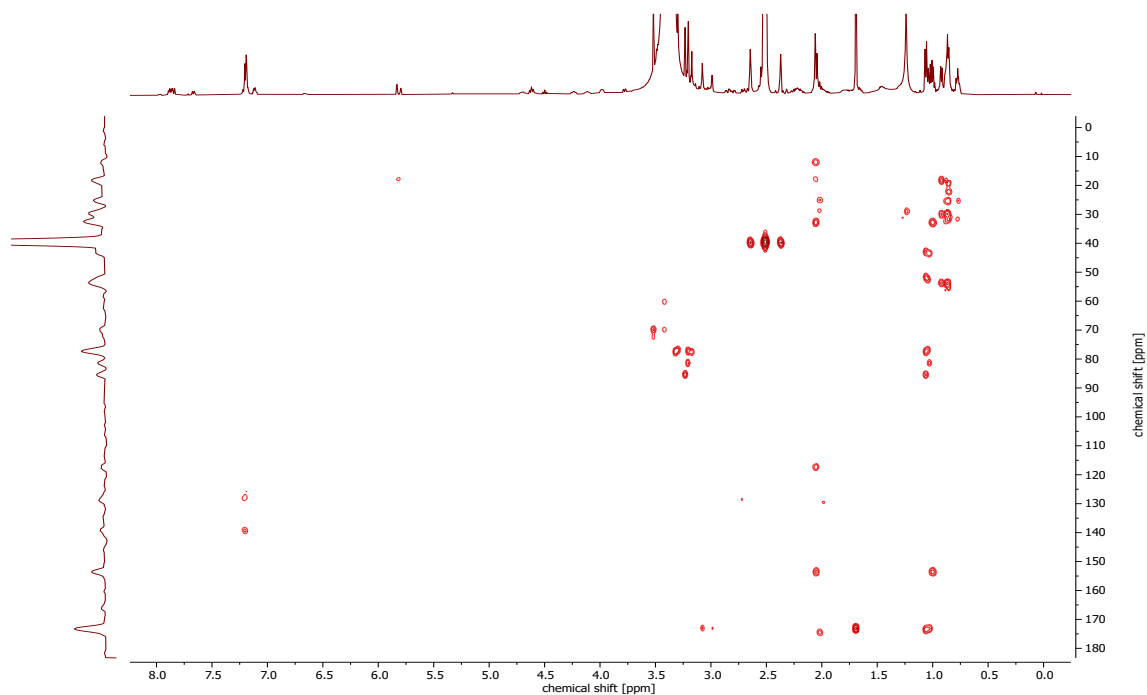
Figure A.2.33.: ^{13}C-HMBC NMR spectrum of the degraded AEST derivative (**4**) in DMSO-d_6 (600 MHz). See **Tab. A5** for a list of the observed chemical shifts.
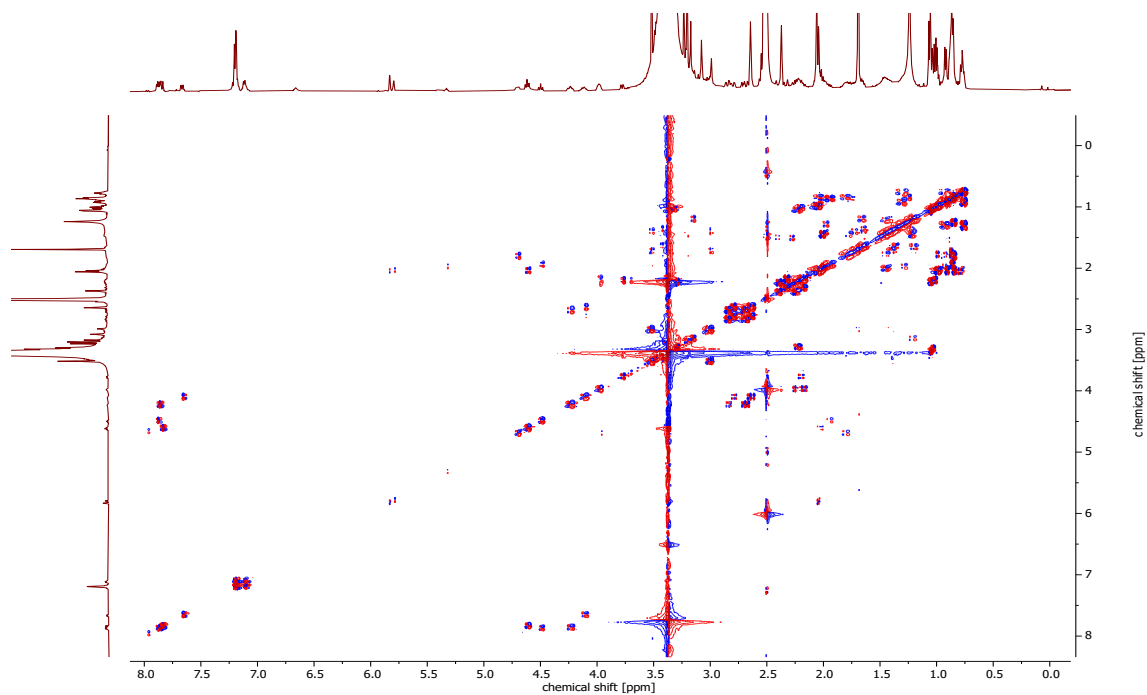

Figure A.2.34.: DQF-COSY NMR spectrum of the degraded AEST derivative (**4**) in DMSO-d_6 (600 MHz). See **Tab. A5** for a list of the observed chemical shifts.
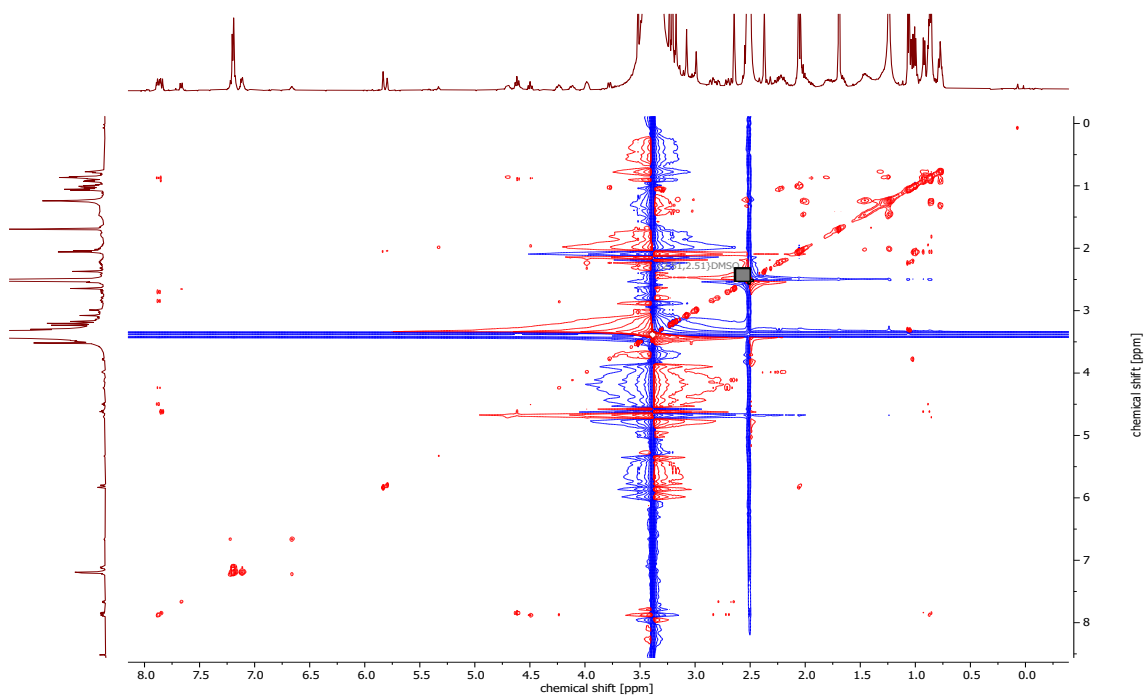
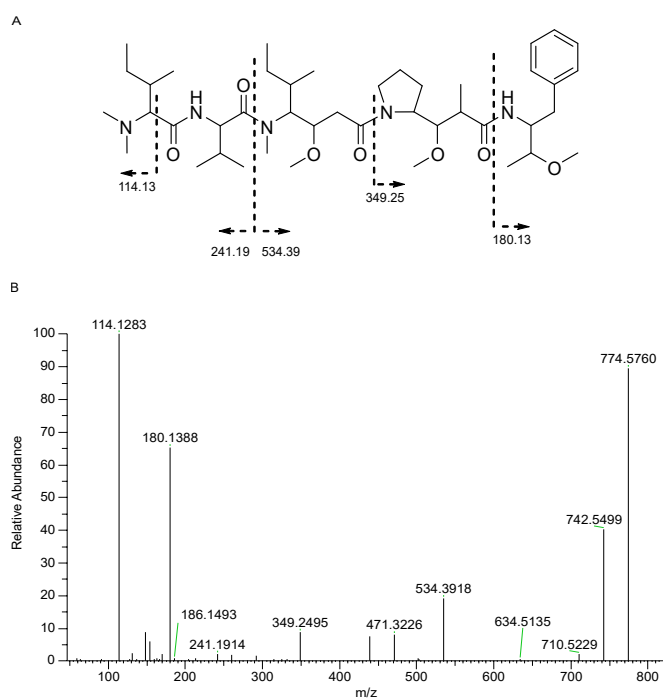**Figure A.2.35.:** TOCSY NMR spectrum of the degraded AEST derivative (**4**) in DMSO-d_6 (600 MHz).

Figure A.2.36.: (**A**) Key MS/MS fragments and (**B**) MS/MS spectrum of AEST (**1**).
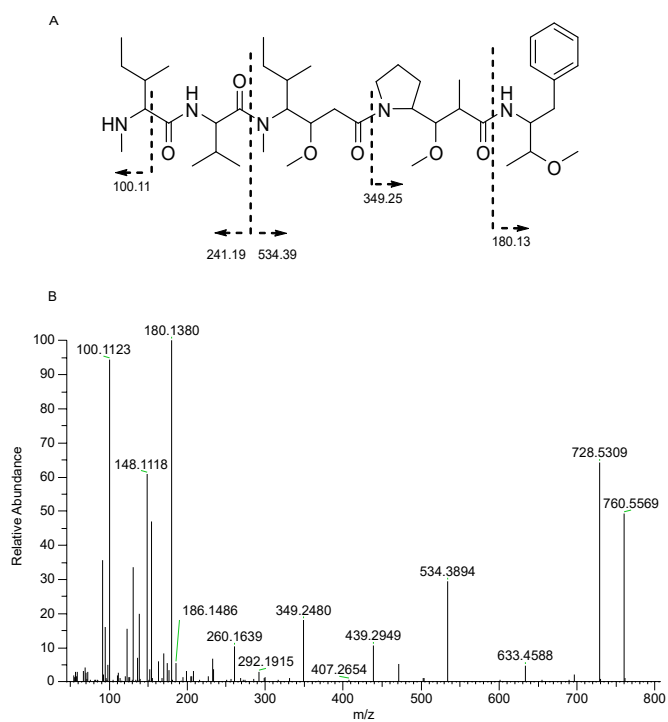
Figure A.2.37.: (**A**) Key MS/MS fragments and (**B**) MS/MS spectrum of MMAEST (**2**).
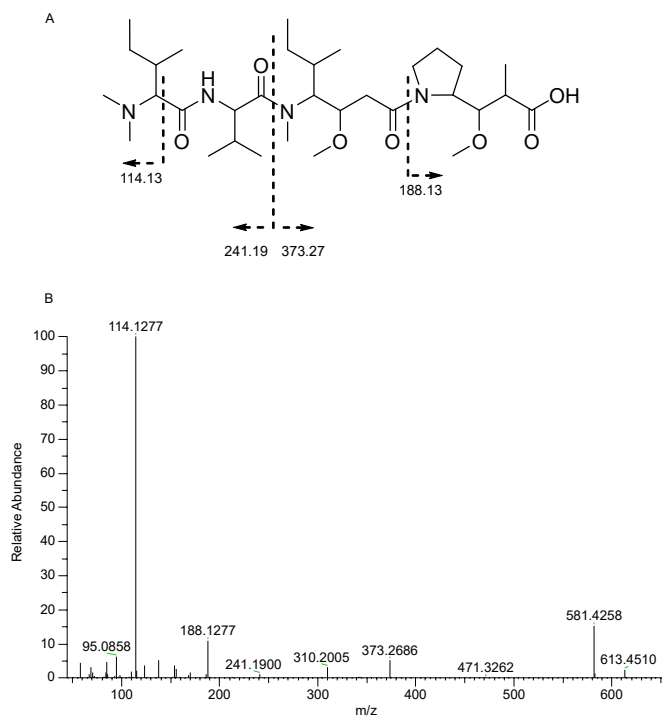
Figure A.2.38.: (**A**) Key MS/MS fragments and (**B**) MS/MS spectrum of DAAEST (**3**).

Figure A.2.39.: COSY (bold bonds) and key HMBC correlations (arrows) of AEST (**1**).

Figure A.2.40.: (**A**) MS and MS/MS spectrum of the degraded AEST derivative (**4**) after isolation. (**B**) MS and MS/MS spectrum after NMR spectroscopy and proposed structure.
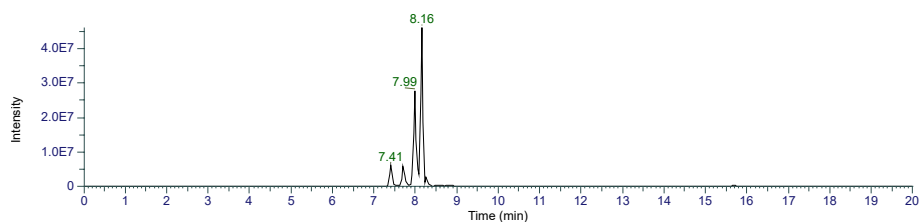
Figure A.2.41.: Extracted ion chromatogram (EIC) of monomethylaetokthonostatin (MMAEST) derivatives (m/z 760.5) in the crude extract.

Figure A.2.42.: Predicted structure based on MS/MS data of MMAEST derivative at retention time 7.41 min.

Figure A.2.43.: Predicted structure based on MS/MS data of MMAEST derivative at retention time 7.70 min.
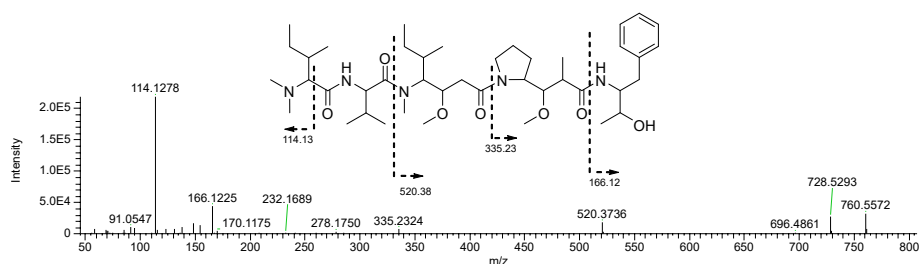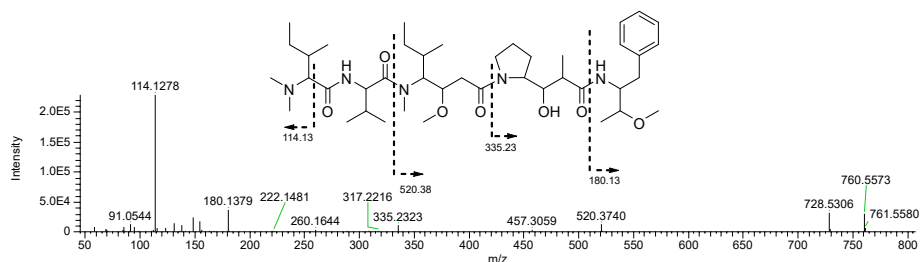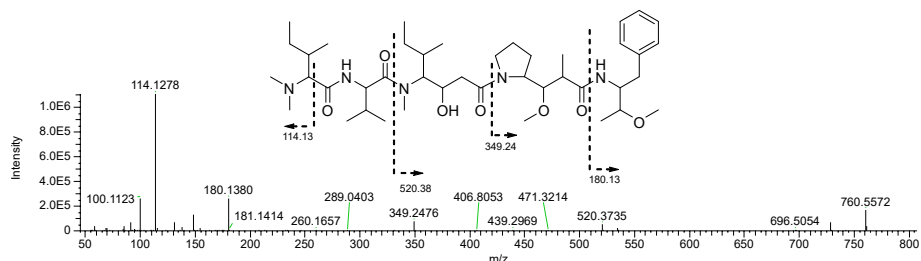

Figure A.2.44.: Predicted structure based on MS/MS data of MMAEST derivative at retention time 7.99 min.

Figure A.2.45.: Predicted structure based on MS/MS data of MMAEST derivative at retention time 8.16 min.
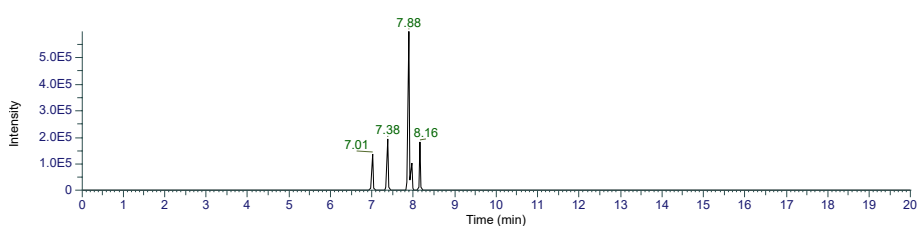
Figure A.2.46.: Extracted ion chromatogram (EIC) of di-desmethylaetokthonostatin (DDAEST) derivatives (m/z 746.5).

Figure A.2.47.: Predicted structure based on MS/MS data of DDAEST derivative at retention time 7.01 min.

Figure A.2.48.: Predicted structure based on MS/MS data of DDAEST derivative at retention time 7.38 min.
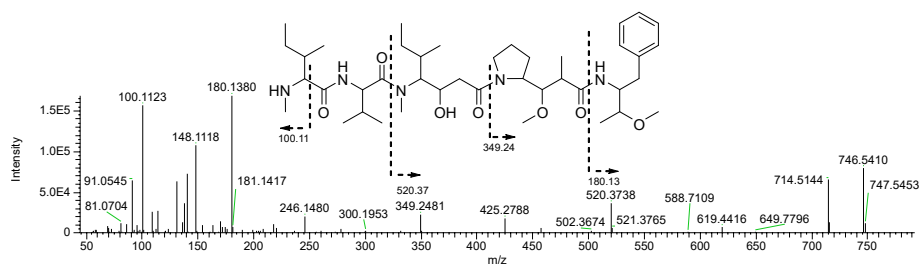
Figure A.2.49.: Predicted structure based on MS/MS data of DDAEST derivative at retention time 7.88 min.
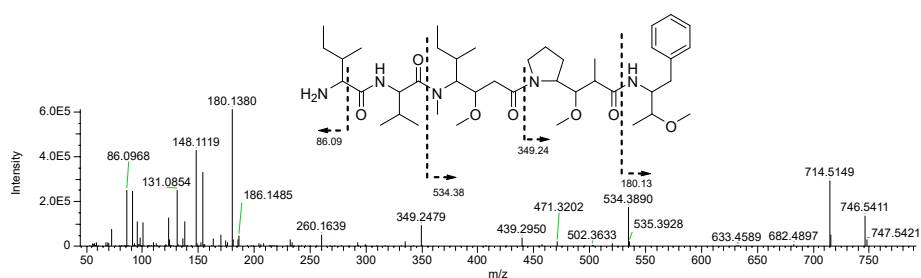
Figure A.2.50.: Predicted structure based on MS/MS data of DDAEST derivative at retention time 8.16 min.
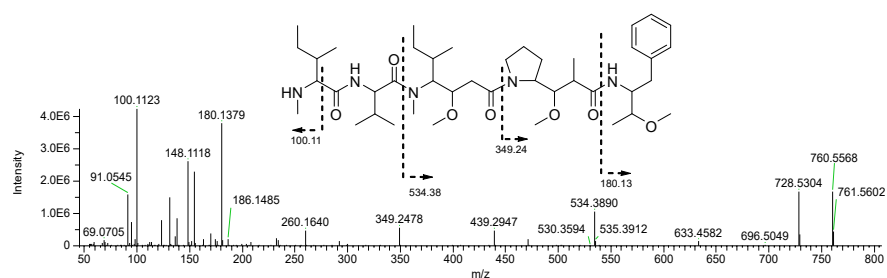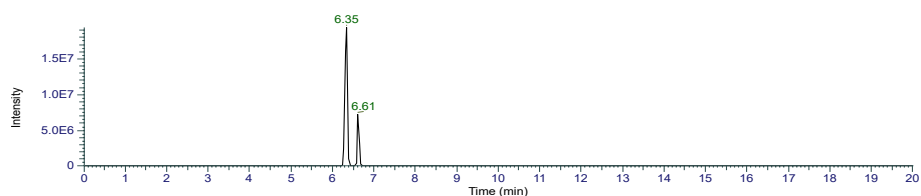

Figure A.2.51.: Extracted ion chromatogram (EIC) of monomethyltetrapeptide (MMtetrapeptide) derivatives (*m/z* 599.4) in the crude extract.
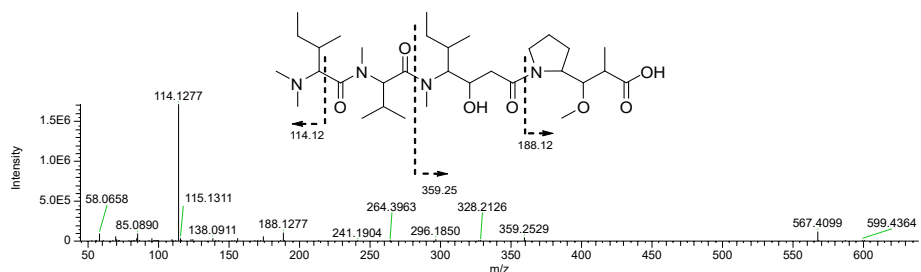
Figure A.2.52.: Predicted structure based on MS/MS data of the MMtetrapeptide derivative at retention time 6.35 min.
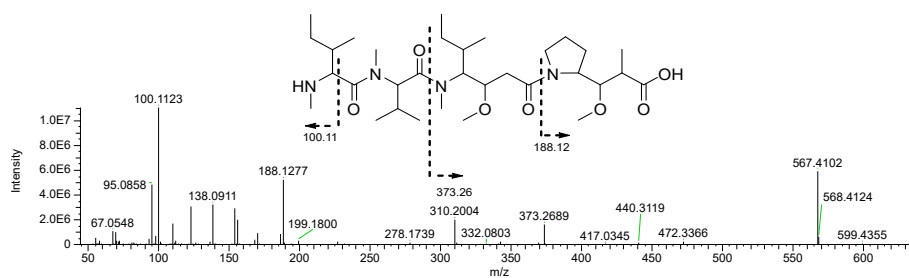
Figure A.2.53.: Predicted structure based on MS/MS data of the MMtetrapeptide derivative at retention time 6.60 min.
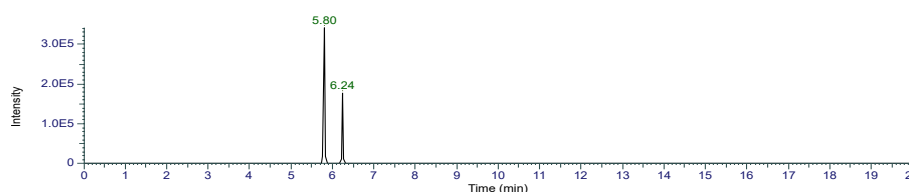
Figure A.2.54.: Extracted ion chromatogram (EIC) of di-desmethyltetrapeptide (DDtetrapeptide) derivatives (*m/z* 585.4) in the crude extract.
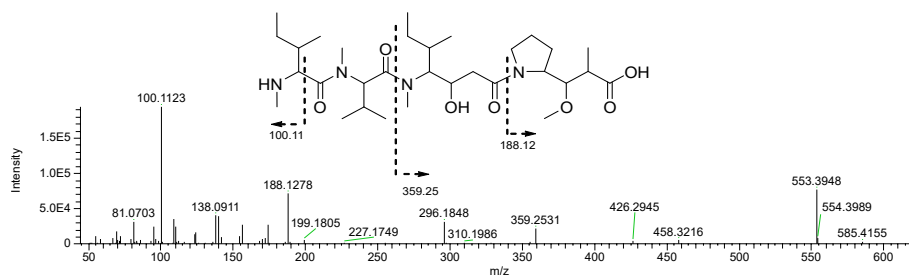
Figure A.2.55.: Predicted structure based on MS/MS data of DDtetrapeptide derivative at retention time 5.80 min.
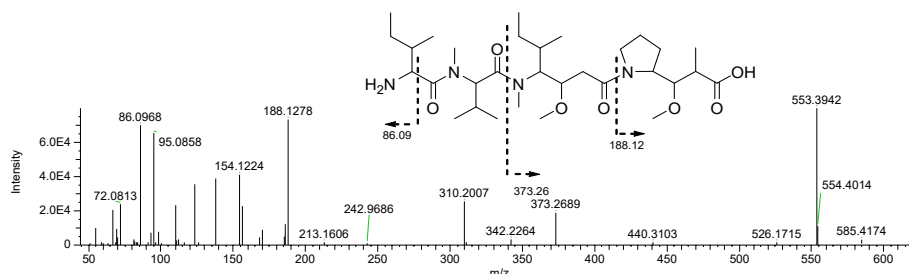
Figure A.2.56.: Predicted structure based on MS/MS data of DDtetrapeptide derivative at retention time 6.24 min.
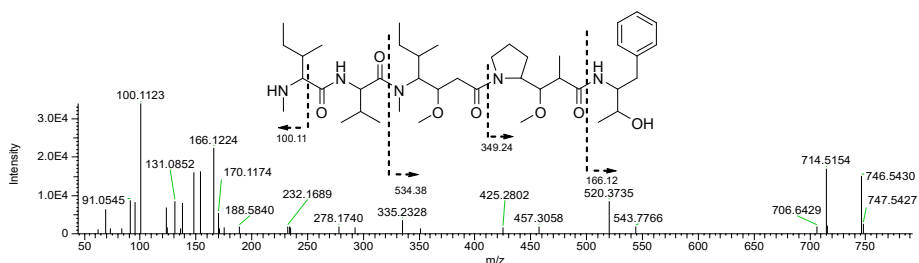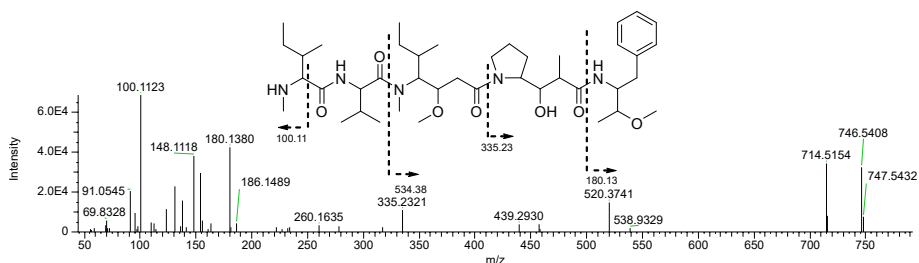
A.3. Supplementary information to chapter 3.2
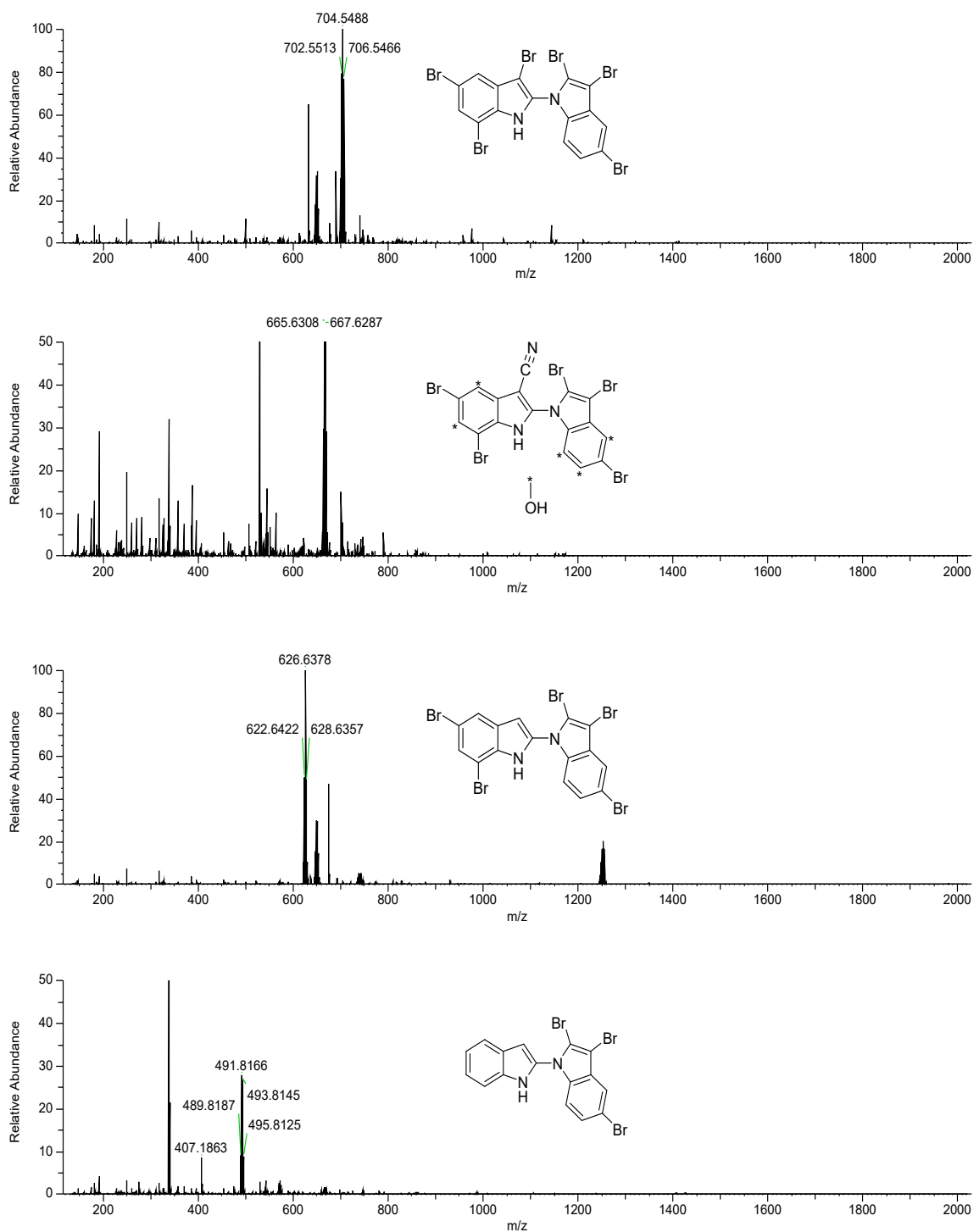
Figure A.3.1.: MS spectra and possible structures of AETX derivatives identified via classical GNPS networking analysis.
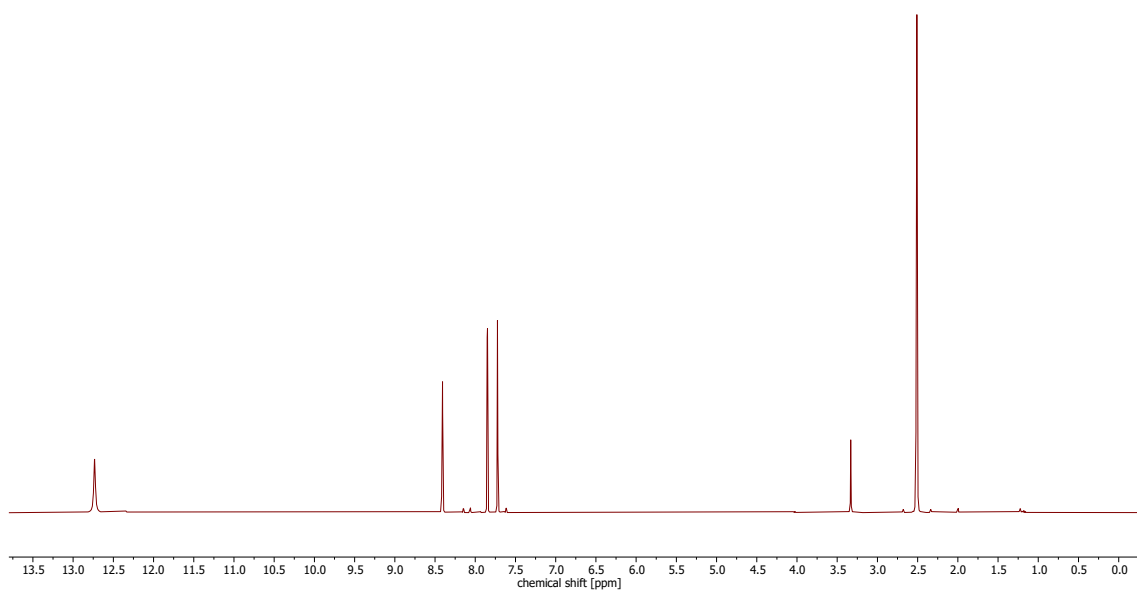
Figure A.3.2.: ^{1}H-NMR spectrum of the isolated 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (400 MHz). See **Tab. A6** for a list of the observed chemical shifts.
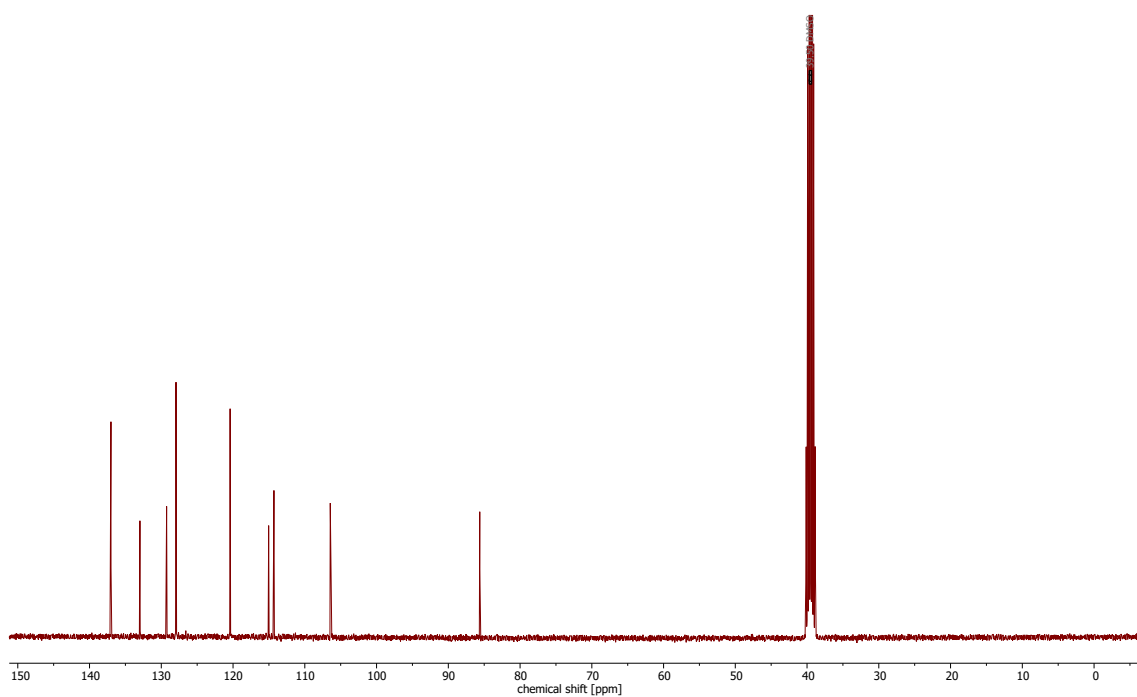
Figure A.3.3.: ^{13}C-NMR spectrum of the isolated 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (150 MHz). See **Tab. A6** for a list of the observed chemical shifts.
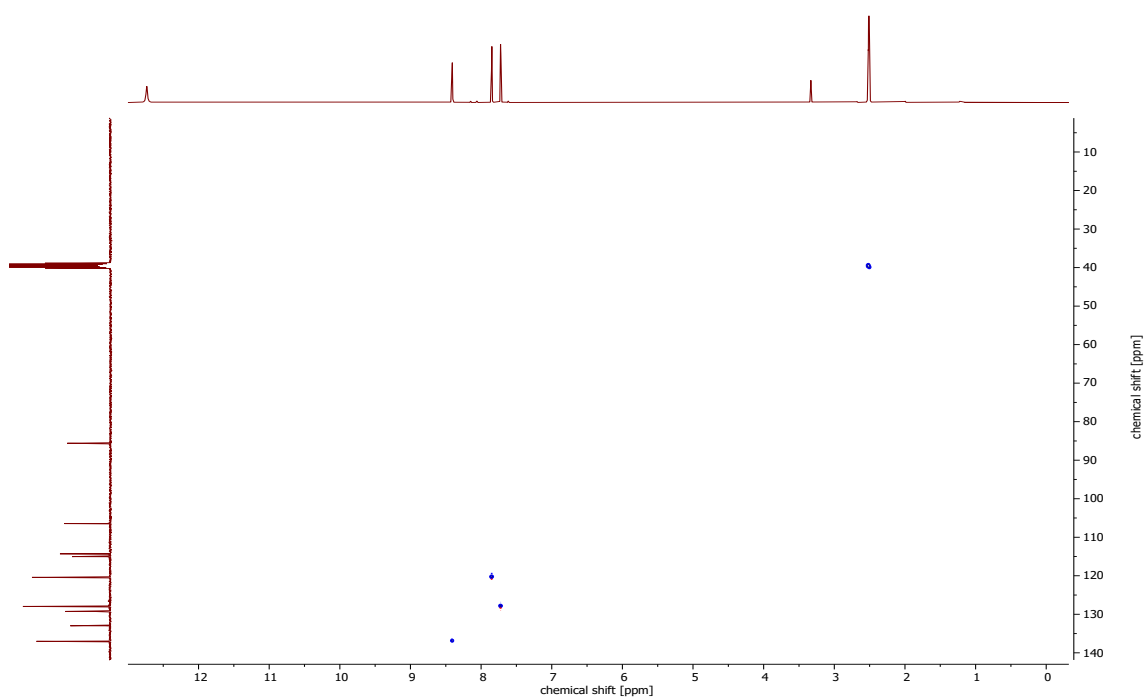
Figure A.3.4.: ^{13}C-HSQC NMR spectrum of the isolated 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (400 MHz).
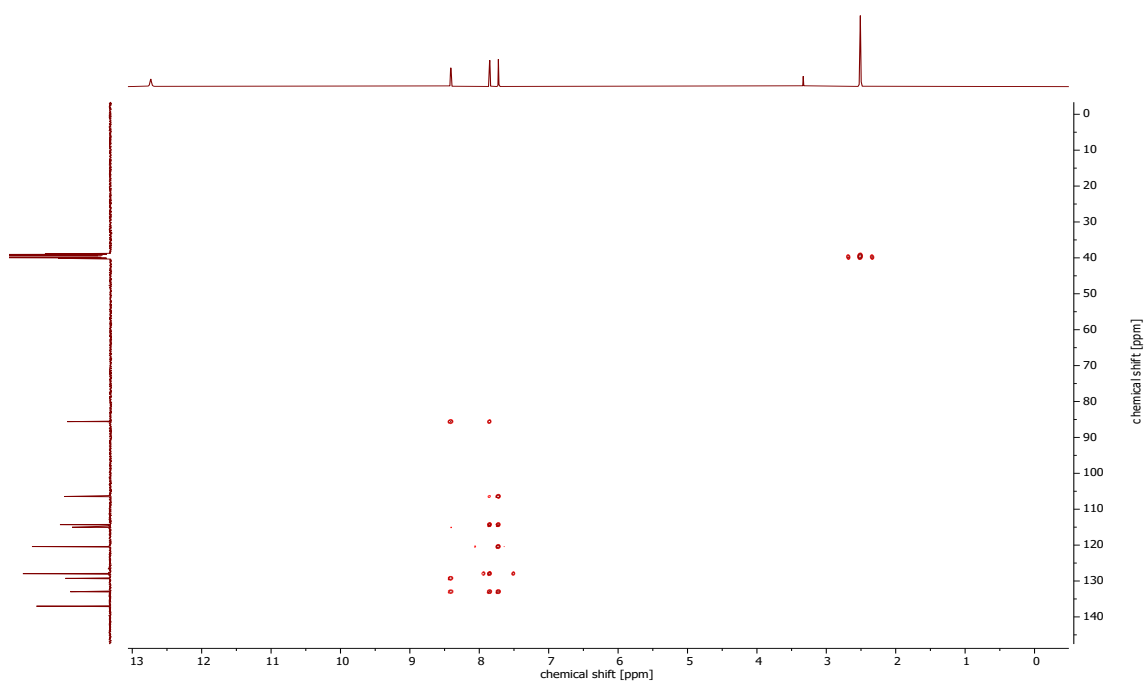
Figure A.3.5.: ^{13}C-HMBC NMR spectrum of the isolated 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (400 MHz).
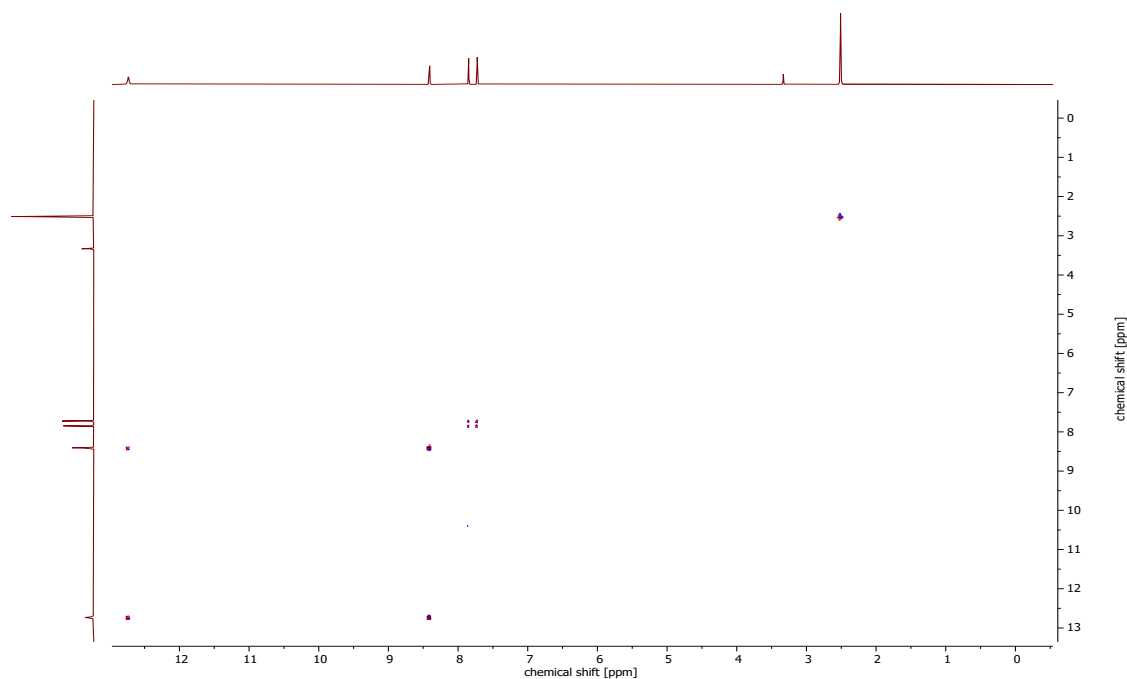
Figure A.3.6.: DQF-COSY NMR spectrum of the isolated 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (400 MHz).

Table A.3.1.: NMR data of 5,7-dibromo-3-carbonitrile indole (**5**) compared to the NMR data of the corresponding part of AETX in DMSO-d_6.

| C/H no. | 5,7-dibromo-3-carbonitrile | | AETX | |
	δH (J in Hz)	δC, mult	δH (J in Hz)	δC, mult
1	12.73, s		13.76, br, s	
2	8.41, d (2.69)	137.0, CH		137.7, qC
3		85.6, qC		85.2, qC
3a		133.0, qC		131.5, qC
4	7.85, d (1.71)	120.4, CH	8.04, s	121.1, CH
5		114.3, qC		115.3, qC
6	7.72, d (1.71)	128.0, CH	7.90, s	129.7, CH
7		106.5, qC		107.0, qC
7a		129.2, qC		127.9, qC
8		115.0, qC		112.7, qC
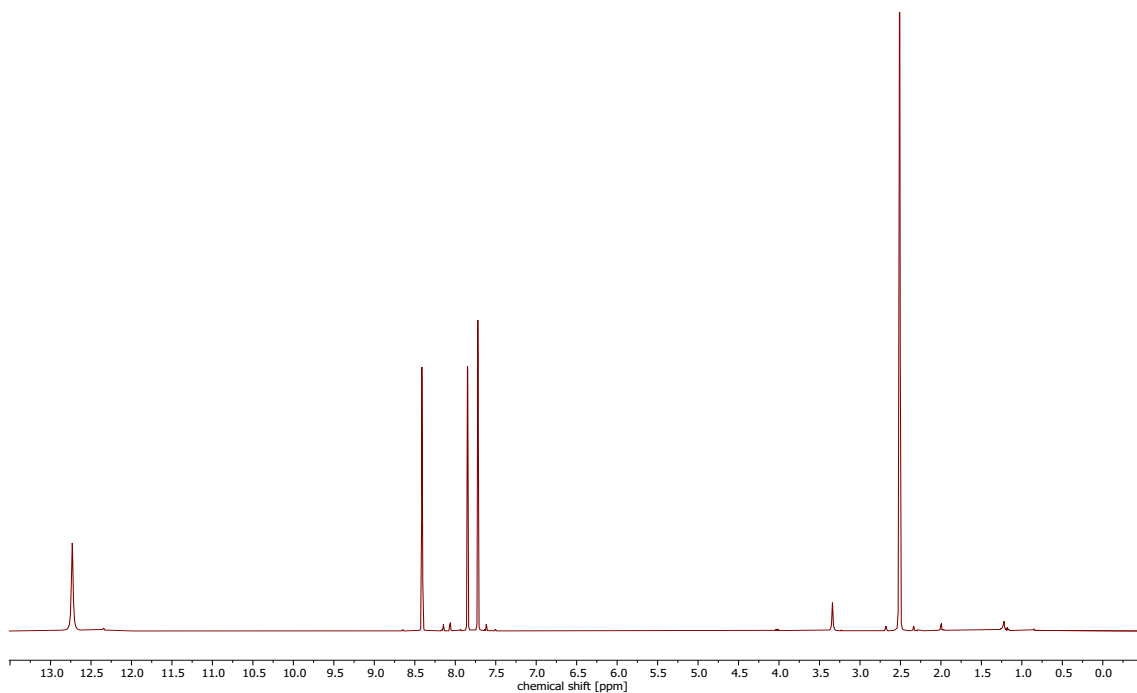
Figure A.3.7.: ^{1}H-NMR spectrum of the via method A synthesized 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (400 MHz).
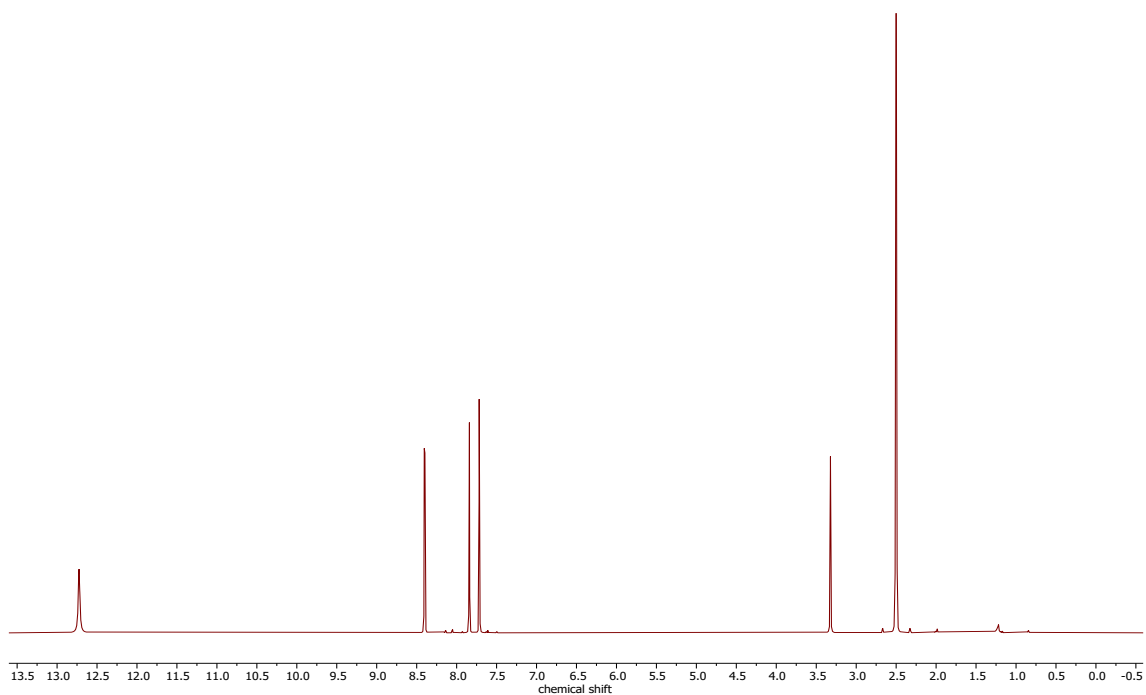
Figure A.3.8.: ^{1}H-NMR spectrum of the via method B synthesized 5,7-dibromo-3-carbonitrile indole (**5**) in DMSO-d_6 (400 MHz).
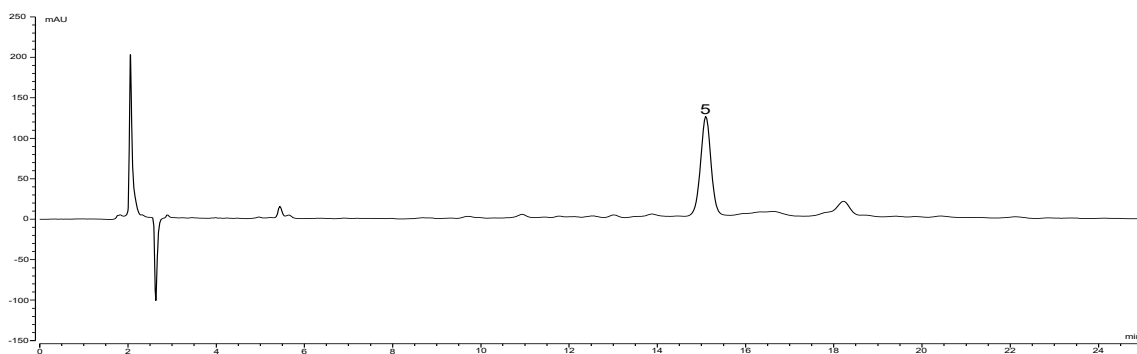
Figure A.3.9.: HPLC-UV chromatogram (210 nm) of the isolation of 5,7-dibromo-3-carbonitrile indole (**5**) (isocratic conditions of 55 % MeCN in water + 0.1 % FA each) on a Luna C_{18} column (4.6 mm ID x 250 mm).
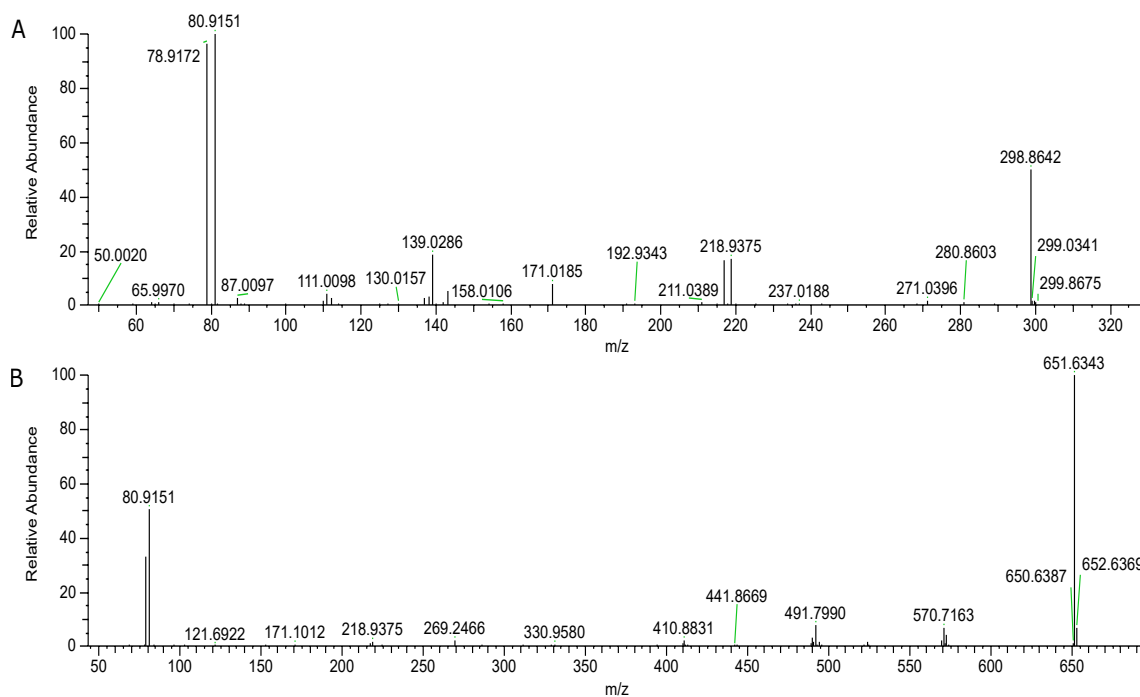
Figure A.3.10.: (**A**) MS/MS data of 5,7-dibromo-3-carbonitrile indole (**5**). (**B**) MS/MS data of AETX.

A.4. Supplementary information to chapter 3.3

Table A.4.1.: Environmental samples investigated in present study, with summarized results of the PCR and HPLC-MS analysis. Covington Reservoir (COV), Long Branch Reservoir (LB), Jackson Lake on Tussahaw Creek (TUS), Lake Sinclair (SIN), and J. Strom Thurmond reservoir (THU).

N	Sample ID	PCR	LC-MS	Peak Area
1	COV-H1-1	+	+	3.33×10^6
2	COV-H1-2	+	+	1.35×10^7
3	COV-H1-3	+	+	1.47×10^7
4	COV-H1-4	+	+	4.04×10^6
5	COV-H1-5	+	+	1.20×10^6
6	COV-H2-1	+	+	1.39×10^7
7	COV-H2-2	+	+	5.28×10^6
8	COV-H2-3	+	+	1.49×10^7
9	COV-H2-4	+	+	9.07×10^5
10	COV-H2-5	+	+	1.97×10^7
11	COV-H3-1	+	+	6.42×10^5
12	COV-H3-2	+	+	2.11×10^6
13	COV-H3-3	+	+	2.60×10^5
14	COV-H3-4	+	-	-
15	COV-H3-5	+	+	1.93×10^5
16	LB-H1-1	+	+	6.14×10^7
17	LB-H1-2	+	+	1.59×10^8
18	LB-H1-3	+	+	4.03×10^7
19	LB-H1-4	+	+	4.58×10^7
20	LB-H1-5	+	+	3.36×10^7
21	LB-H2-1	+	+	7.50×10^7
22	LB-H2-2	+	+	5.83×10^7
23	LB-H2-3	+	+	3.30×10^7
24	LB-H2-4	+	+	4.06×10^7
25	LB-H2-5	+	+	6.52×10^7
26	LB-H3-1	+	+	1.09×10^8
27	LB-H3-2	+	+	1.03×10^7
28	LB-H3-3	+	+	5.70×10^7
29	LB-H3-4	+	+	3.44×10^7
30	LB-H3-5	+	+	2.80×10^7
31	TU-H1-1	+	+	2.04×10^7
32	TU-H1-2	+	+	1.11×10^7
33	TU-H1-3	+	+	2.43×10^7
34	TU-H1-4	+	+	2.61×10^7
35	TU-H1-5	+	+	8.15×10^7
36	TU-H2-1	+	+	1.53×10^7
37	TU-H2-2	+	+	2.51×10^5
38	TU-H2-3	+	+	2.40×10^6
39	TU-H2-4	+	+	7.79×10^5
40	TU-H2-5	+	+	1.56×10^6
41	TU-H3-1	+	+	2.62×10^7
42	TU-H3-2	+	+	2.50×10^7
43	TU-H3-3	+	+	2.73×10^7
44	TU-H3-4	+	+	1.30×10^7
45	TU-H3-5	+	+	1.72×10^7
46	SIN-H1-1	-	-	-
47	SIN-H1-2	-	-	-
48	SIN-H1-3	-	-	-
49	SIN-H1-4	-	-	-
50	SIN-H1-5	-	-	-
51	SIN-H2-1	-	-	-
52	SIN-H2-2	-	-	-
53	SIN-H2-3	-	-	-
54	SIN-H2-4	-	-	-
55	SIN-H2-5	-	-	-
56	SIN-H3-1	-	-	-
57	SIN-H3-2	-	-	-
58	SIN-H3-3	-	-	-
59	SIN-H3-4	-	-	-
60	SIN-H3-5	-	-	-
61	THU-W1-1	+	+	1.32×10^5
62	THU-W1-2	+	+	3.16×10^5
63	THU-W1-3	+	+	4.21×10^4

A.5. Publications

[1]: More than just an Eagle Killer: The freshwater cyanobacterium *Aetokthonos hydrillicola* produces highly toxic dolastatin derivatives

Markus Schwark, José A. Martínez Yerena, Kristin Röhrborn, Pavel Hrouzek, Petra Divoká, Lenka Štenclová, Kateřina Delawská, Heike Enke, Christopher Vorreiter, Faith Wiley, Wolfgang Sippl, Roman Sobotka, Subhasish Saha, Susan B. Wilde, Jan Mareš, Timo H. J. Niedermeyer, More than just an Eagle Killer: The freshwater cyanobacterium *Aetokthonos hydrillicola* produces highly toxic dolastatin derivatives. *PNAS* 120, e2219230120 (2023), DOI: 10.1073/pnas.2219230120.

PNAS

More than just an eagle killer: The freshwater cyanobacterium *Aetokthonos hydrillicola* produces highly toxic dolastatin derivatives

Markus Schwark[a,1], José A. Martínez Yerena[b,c,d,1], Kristin Röhrborn[a,2], Pavel Hrouzek[d], Petra Divoká[d], Lenka Štenclová[b], Kateřina Delawská[c,d], Heike Enke[e], Christopher Vorreiter[f], Faith Wiley[g,3], Wolfgang Sippl[f], Roman Sobotka[c,d], Subhasish Saha[d,4], Susan B. Wilde[h], Jan Mareš[b,c,d,5], and Timo H. Niedermeyer[a,i,5]

Cyanobacteria are infamous producers of toxins. While the toxic potential of planktonic cyanobacterial blooms is well documented, the ecosystem level effects of toxigenic benthic and epiphytic cyanobacteria are an understudied threat. The freshwater epiphytic cyanobacterium *Aetokthonos hydrillicola* has recently been shown to produce the "eagle killer" neurotoxin aetokthonotoxin (AETX) causing the fatal neurological disease vacuolar myelinopathy. The disease affects a wide array of wildlife in the southeastern United States, most notably waterfowl and birds of prey, including the bald eagle. In an assay for cytotoxicity, we found the crude extract of the cyanobacterium to be much more potent than pure AETX, prompting further investigation. Here, we describe the isolation and structure elucidation of the aetokthonostatins (AESTs), linear peptides belonging to the dolastatin compound family, featuring a unique modification of the C-terminal phenylalanine-derived moiety. Using immunofluorescence microscopy and molecular modeling, we confirmed that AEST potently impacts microtubule dynamics and can bind to tubulin in a similar matter as dolastatin 10. We also show that AEST inhibits reproduction of the nematode *Caenorhabditis elegans*. Bioinformatic analysis revealed the AEST biosynthetic gene cluster encoding a nonribosomal peptide synthetase/polyketide synthase accompanied by a unique tailoring machinery. The biosynthetic activity of a specific N-terminal methyltransferase was confirmed by in vitro biochemical studies, establishing a mechanistic link between the gene cluster and its product.

cyanotoxin | cytotoxicity | aetokthonostatin | dolastatin | biosynthesis

Cyanobacteria are a prolific source of biologically active specialized metabolites (1–3). They are infamous for their potent toxins, such as the neurotoxic alkaloids anatoxins or the hepatotoxic nonribosomal peptides microcystins (4, 5). Harmful cyanobacterial blooms (mass proliferation of toxic cyanobacteria in stagnant waters) can pose a threat for ecosystem health (6). While planktic cyanobacteria have been studied intensely in the past, the effects of benthic or epiphytic cyanobacteria remain largely unknown, although their cyanotoxin repertoire is equally high, and they have been linked to numerous animal poisonings (7). Recently, a novel cyanobacterial toxin was discovered from the filamentous epiphytic cyanobacterium *Aetokthonos hydrillicola* (8). The biindole alkaloid aetokthonotoxin (AETX, Fig. 1*A*) has been found to be the cause of the disease Vacuolar Myelinopathy (VM), affecting wildlife in the southeastern United States. *A. hydrillicola* grows on submersed aquatic plants, especially the invasive *Hydrilla verticillata*, which is consumed by waterfowl, fish, and snails, which in turn are consumed by birds of prey. AETX consumption resulted in VM throughout every trophic level of this aquatic food chain (8).

Biologically active cyanobacterial specialized metabolites provide opportunities for drug substance development (1, 9, 10). Prominent examples that have advanced into clinical studies are saxitoxins as long-lasting local anesthetics (11, 12) or cryptophycins as potential antineoplastic drugs (13, 14). However, the only cyanobacterial natural products of which derivatives are clinically used to date are the dolastatins (1, 9). Linear dolastatins like dolastatin 10 are pentapeptides that contain two uncommon amino acids characteristic for the dolastatin compound family, dolaproine (Dap) and dolaisoleuine (Dil) (15). Linear dolastatins as well as structurally related peptides of cyanobacterial origin, e.g., symplostatin 1 (16) or malevamide D (17), are highly cytotoxic (low to sub-nM EC_{50}). Binding to the vinca binding site of the β1 subunit of tubulin, they interrupt microtubule assembly, resulting in cell cycle arrest in the G_2/M phase (18–20). Synthetic derivatives of dolastatin 10, monomethylauristatin E (MMAE) and monomethylauristatin F (MMAF), have been introduced into the clinic as highly cytotoxic payloads of antineoplastic antibody–drug conjugates. The

Significance

Cyanotoxins have adverse effects on ecosystems. Our understanding of their potential risk has recently been expanded by the discovery of aetokthonotoxin, produced by the cyanobacterium *Aetokthonos hydrillicola* growing on invasive plants. Via trophic transfer, it acts as a neurotoxin causing mortality in animals including top predators like bald eagles. Closer examination of *A. hydrillicola* revealed that it also produces highly toxic dolastatin derivatives. *A. hydrillicola* is the only known cultured cyanobacterium producing dolastatin derivatives, which allowed uncovering biosynthetic gene clusters of this compound family. In contrast to all other known dolastatin producers, which are marine cyanobacteria, *A. hydrillicola* thrives in freshwater reservoirs, making it a potential threat also for human health. Monitoring of the cyanobacterium and its toxins is strongly recommended.

Fig. 1. *A. hydrillicola* produces dolastatin analogs. (*A*) Structures of AETX, AEST, MMAEST, DAAEST, as well as MMAE, MMAF and dolastatin 10. (*B*) AEST cluster in a GNPS feature-based MS/MS networking analysis of an HPLC-MS/MS analysis of an *A. hydrillicola* extract. Node size proportional to ion intensity, exact mass of [M+H]⁺ indicated next to the respective node. Red: AEST, orange: pentapeptide AEST derivatives like MMAEST, cyan: tetrapeptide AEST derivatives like DAAEST.

first of these, brentuximab vedotin, has been approved by the US Food and Drug Administration in 2011 (9).

The dolastatins were first described from the sea hare *Dolabella auricularia* by Petitt et al. in 1987 (21). However, later studies found that sea hares acquire the dolastatins by their diet, which consists largely of cyanobacteria of the genus *Symploca*, the actual producers of these compounds (22, 23). Notably, all cytotoxic linear derivatives of dolastatin 10 known to date have been isolated from marine cyanobacteria of the genus *Symploca*/*Caldora* (24). Marine benthic filamentous cyanobacteria have been intensively studied as an extremely rich source of bioactive compounds (25). However, no dolastatin-producing strain has been cultured yet, hampering the elucidation of the genetic basis for the compound's biosynthesis as well as synthetic biology approaches for their modification. As well-established total synthesis routes to the dolastatins are available, ongoing research on the medicinal chemistry of dolastatins focuses on synthetic analogs (26).

Profiling the bioactivity of purified AETX and crude extracts of the cyanobacterium *A. hydrillicola*, we observed that the cytotoxicity of a crude extract of *A. hydrillicola* biomass was much stronger than the cytotoxicity of a corresponding amount of pure AETX. Thus, there had to be another compound in the extract with higher cytotoxicity than AETX. Intrigued by these results, we set out to isolate and characterize the cytotoxic compound produced by *A. hydrillicola*. Here, we present the isolation and structure elucidation of aetokthonostatin (AEST) and AEST derivatives. The AESTs are linear peptides belonging to the dolastatin compound family, featuring a yet undescribed C-terminal phenylalanine-derived building block. Immunofluorescence microscopy and molecular modeling studies suggest that AEST, like dolastatin 10, exerts its strong cytotoxicity by binding to tubulin. We show that AEST inhibits reproduction of the nematode *Caenorhabditis elegans*. Bioinformatic analysis of the *A. hydrillicola* genome revealed the biosynthetic gene cluster (BGC) of AEST, which we confirmed by biochemical experiments. Our study

106

highlights that the cyanobacterium *A. hydrillicola* might pose a serious threat to drinking-water supplies, as it is able to produce two specialized metabolites with pronounced toxicity.

Results and Discussion

Toxicity of *A. hydrillicola* Extract and AETX. Early investigation of a crude extract of the cyanobacterium *A. hydrillicola* for cytotoxicity showed a remarkably high activity of the extract (EC_{50} 0.12 μg/mL, HeLa, *SI Appendix,* Fig. S2). Focusing on the isolation of the toxin causing VM, we assumed that the toxin causing VM would also be responsible for the extract's cytotoxicity. After isolation of AETX and testing the pure compound for cytotoxicity, to our surprise, we found that AETX is only moderately cytotoxic (EC_{50} 1 μM = 0.65 μg/mL, HeLa). Quantification of the AETX content in a crude extract assayed for cytotoxicity showed that based on the determined AETX content, the extract was about fivefold more cytotoxic than expected if AETX was the sole cytotoxic compound in the extract. This suggested that another, more cytotoxic compound must be present. Indeed, time-based fractionation of the extract and subsequent cytotoxicity testing of the individual fractions revealed that not AETX but the main compound observed in the extract's HPLC chromatogram was highly cytotoxic (*SI Appendix,* Fig. S3).

Isolation and Structure Elucidation of AEST. In contrast to AETX, which is only produced by *A. hydrillicola* when the strain's cultivation medium is supplemented with bromide salts (8), we found the cytotoxic metabolite to be produced under all tested cultivation conditions. Thus, the cyanobacterium was cultivated in standard BG-11 medium to generate biomass for subsequent processing. After harvest and lyophilization, the biomass was extracted with methanol. The crude extract was fractionated using flash chromatography. Fractions were assayed for cytotoxicity, and the cytotoxic fractions were subjected to semipreparative HPLC to isolate the active compounds. We found that the main peak observed in the chromatogram (*SI Appendix,* Fig. S3) consisted of two individual compounds, which we subsequently separated using ammonium acetate buffer (pH 9) in the mobile phase.

The main compound was isolated as a white amorphous powder. High-resolution mass spectrometry suggested the molecular formula $C_{43}H_{76}O_7N_5$ ([M+H]⁺ at *m/z* 774.5759). NMR analysis (*SI Appendix,* Figs. S5–S10) was complicated by a considerable signal overlap in the ¹H NMR spectrum, caused by the presence of two conformers in solution (ratio main to minor conformer about 3:2, *SI Appendix,* Fig. S5). Analysis of the ¹³C and ¹H NMR data suggested four amide or ester carbonyls and two NH protons (NMR data see *SI Appendix,* Table S1). The presence of amide bonds was confirmed by four signals agreeing with α-carbonyl protons of amino acids, giving us confidence that the compound was a peptide. Furthermore, we observed signals of six *N*- or *O*-methyl groups. Evaluation of HSQC-DEPT, TOCSY, and COSY spectra allowed us to assemble the spin systems of the amino acid building blocks in the peptide (for a more detailed discussion, see *SI Appendix*). In addition to the amino acids valine and isoleucine, we found that the compound contains the known nonproteinogenic amino acids Dil and Dap, indicating that the compound belongs to the dolastatin compound family. The presence of Dap explained the existence of two conformers in solution, which has also been described for other compounds in the dolastatin compound family (27). We could assemble the fifth monomer as 3-methoxy-1-phenylbutan-2-amine, a yet undescribed monomer in peptide natural products, that we give the trivial name

aetophenine (Aph). The building block sequence as well as the position of the *N*- and *O*-methyl groups was deduced from an HMBC spectrum and confirmed by a ROESY spectrum. The planar structure of the compound was thus established to be *N,N*-dimethyl-Ile-Val-Dil-Dap-Aph. This sequence was supported by the fragmentation pattern in MS/MS experiments (*SI Appendix,* Fig. S11). As this planar structure differs from the structure of symplostatin 1 only by the C-terminal monomer (22), we re-recorded a ¹³C-NMR spectrum of the compound in CD_2Cl_2 (*SI Appendix,* Fig. S6 *B* and *C*). Indeed, the NMR data of the four N-terminal monomers matched to those reported for symplostatin 1 (22), indicating that these four monomers have the same relative and absolute stereochemistries in both compounds (*SI Appendix,* Table S2). To confirm this assumption, the absolute stereochemistries of the fourth monomer, Dap, were determined by Marfey's analysis (*SI Appendix,* Fig. S12). As no reference material for Aph is available, we first tried to hydrolyze the compound and isolate Aph for subsequent VCD spectroscopic analysis. However, we found that under the strong acidic or basic conditions needed for hydrolysis, the methyl ether was hydrolyzed as well, obfuscating the original configuration. Crystallization for X-ray crystallography was not successful. The absolute configuration of Aph could thus not be determined. The structure of the compound is shown in Fig. 1*A*. This main compound produced by *A. hydrillicola* was named AEST.

The closely eluting second compound was isolated as a white amorphous powder. Its molecular formula was established as $C_{42}H_{74}O_7N_5$ ([M+H]⁺ at *m/z* 760.5571), differing from AEST only by the absence of one methyl group. Analysis of the MS/MS and NMR data confirmed that, compared to AEST, this compound lacks one of the two *N*-methyl groups at the N-terminal isoleucine (*SI Appendix,* Figs. S13–S19 and Table S3). Thus, it has been identified as monomethylaetokthonostatin (MMAEST).

A third compound, isolated as a white amorphous powder, eluted earlier than AEST and MMAEST, indicating higher polarity or smaller size. Indeed, its molecular formula, $C_{32}H_{61}O_7N_4$ ([M+H]⁺ at *m/z* 613.4535), suggested that it is constituted of only four amino acids. The mass difference between this compound and AEST corresponds to the C-terminal monomer of AEST, Aph. MS/MS data as well as NMR data supported the hypothesis that this monomer was missing, as no signals were observed in the aromatic region in the ¹H spectrum while the other signals agree with those in the spectra of AEST (*SI Appendix,* Figs. S20–S23 and Table S4), and the prominent fragment ion at *m/z* 180.13 could not be detected (*SI Appendix,* Fig. S24). This third compound was thus confirmed to be *N,N*-dimethyl-Ile-Val-Dil-Dap, des-Aph-aetokthonostatin (DAAEST).

AEST Derivatives in an *A. hydrillicola* Extract. In addition to these three isolated compounds, feature-based MS/MS networking analysis of HPLC-MS/MS data using the "Global Natural Product Social Molecular Networking" software (GNPS) (28) allowed us to detect additional AEST derivatives in an *A. hydrillicola* extract (Fig. 1*B*). These derivatives differ in the methylation pattern and the number of amino acids. Based on the ion intensity, AEST (*m/z* 774.5726) is the most abundant statin derivate in the extract, followed by its N-terminal desmethyl derivative MMAEST (*m/z* 760.5568), its des-Aph derivative DAAEST (*m/z* 613.4524), and the respective N-terminal desmethyl-des-Aph derivative (*m/z* 599.4367). Interestingly, four possible des-methyl derivatives of AEST as well as two of DAAEST could be detected as well as four derivatives of AEST and one derivative of DAAEST lacking two methyl groups. One derivative with an additional methyl group was observed (*m/z* 788.5893). While the pentapeptides represent

A. Appendix

the major part of the AEST derivatives, significant amounts of the respective tetrapeptides are present in the extract, as well. No di- or tripeptides could be detected. The structures of most of these compounds could be deduced based on their MS/MS spectra (*SI Appendix,* Figs. S25–S42).

Bioactivity of AEST. Structure elucidation of AEST already explained the exquisite cytotoxicity of the compound: It belongs to the dolastatin compound family. Dolastatin 10 and related compounds are known for their ability to arrest cells in G_2/M phase due to binding to tubulin and disrupting its assembly to microtubules (18–20). The cytotoxicity of AEST was assessed against HeLa cells and the triple-negative breast cancer cell line MDA-MB 231 and compared to the cytotoxicity of monomethyllauristatin E (MMAE) and monomethyllauristatin F (MMAF), which are used as payloads in approved antibody–drug conjugates (29, 30). We found that AEST (EC_{50} 1 ± 0.2 nM) is equally potent as MMAE (EC_{50} 3 ± 0.4 nM) against HeLa cells (31). Both are about 170-fold more active than MMAF (EC_{50} 170 ± 30 nM; *SI Appendix,* Fig. S4*A*) against MDA-MB 231 cells (31). In HeLa cells, this effect was less distinct; AEST and MMAE were 10-fold more active (Fig. 2*A*).

Due to its negative charge under physiological pH conditions, auristatins possessing a carboxy group at the C-terminal monomer have lower cell permeability, explaining the lower in vitro potency of MMAF compared to MMAE (32). AEST, like MMAE, features an uncharged C terminus. Therefore, it belongs to the MMAE group of dolastatin analogs, which show higher membrane permeability and thus lower EC_{50} values. Lack of the methyl group at the N-terminal nitrogen in MMAEST does not have noticeable influence on potency against HeLa cells (EC_{50} 4 ± 0.01 nM; Fig. 2*A*). We found that the C-terminal monomer of AEST is not mandatory for cytotoxicity, as DAAEST is still strongly cytotoxic (EC_{50} 25 ± 1.4 nM; Fig. 2*A*). This finding was surprising, as it is assumed that a phenyl moiety at the C terminus is mandatory for cytotoxicity (33, 34). However, interestingly, Pettit et al. found that the methyl ester of the N-terminal tetrapeptide of dolastatin 10 was about 30-fold less cytotoxic against L1210 cells compared to dolastatin 10, although it interacted comparably with tubulin (35), agreeing well with the potency difference of AEST and DAAEST we found in this study. This indicates that the lower potency of DAAEST might not only be due to reduced uptake caused by its free carboxylic acid.

The residual survival rate was about 50% in the case of AEST even at high concentrations (Fig. 2*A*). This dose–response behavior is similar to that observed for MMAE (30). Interestingly, MMAEST, and even more pronounced DAAEST, had a lower residual cell viability (about 30% and 20%, resp.).

In early experiments with *C. elegans* and *Danio rerio*, we observed that the crude *A. hydrillicola* extract seemed to be more toxic than isolated AETX. Having identified AEST as a second toxin in the extract, we wondered whether it contributed to the extract's toxicity. However, in stark contrast to AETX, but in agreement with previous findings on dolastatin 15 (36), AEST showed no acute toxic effect on *C. elegans* even at high concentrations of up to 56 μM. Instead, a decrease in the nematode's reproduction rate could be observed at low concentration (EC_{50} 0.8 ± 0.2 μM, Fig. 2*B*). Thus, *A. hydrillicola* produces two toxins displaying different types of toxicity.

As dolastatin 10 and related compounds are well-known tubulin inhibitors binding at the vinca domain, we performed a docking study to test whether AEST can bind to tubulin in a similar fashion to known dolastatin derivatives. As the absolute configuration of the two stereocenters of Aph could not yet be determined, we considered all four possible stereoisomers in this study. Using a crystal structure of a dolastatin analog cocrystallized with tubulin as template (Protein Data Bank ID 4X1K), we found that both AEST and MMAF can nicely be fit into the respective tubulin binding site (Fig. 2*B*; for a detailed discussion of the docking study, see *SI Appendix*). The carbonyl oxygen atoms of Dil and Dap form hydrogen bonds to the backbone amide NH groups of Tyr β224 and Gly β225, respectively. The carbonyl oxygen as well as the amide NH of Val form a bifocal hydrogen bond network with the side chain of Asn α329 (Fig. 2*B*). The charged N-terminal dimethylamino group shows electrostatic interactions with Asp β179. The C-terminal benzyl group showed different orientations in the docking poses (either buried toward Gly α225 or cation-π interaction with Arg β278). The flexibility of the C-terminal group was also observed in the molecular dynamics simulations of AEST and structurally similar cocrystalized peptides, where the C-terminal benzyl moiety is moving significantly during the simulations (*SI Appendix,* Fig. S46). This agrees with our finding that DAAEST, lacking the Aph unit, is still strongly cytotoxic. Also, it suggests that the absolute configuration of the stereocenters of Aph is not important for tubulin binding.

We subsequently confirmed inhibition of tubulin polymerization by AEST experimentally by immunofluorescence staining (Fig. 2*D*). Both AEST and MMAF were assessed in this assay. As expected, the tubulin network appeared affected by AEST at a concentration well below the EC_{50} value (observable effect at 0.1 nM) and proceeding toward higher concentrations, while for MMAF, the tubulin network morphology was comparable to untreated control cells at concentrations around 1 nM. Similar data were obtained with MDA-MB 231 cells, with the activity windows shifted toward higher concentrations (*SI Appendix,* Fig. S4*B*).

Biosynthesis of AEST. Surprisingly, although dolastatin 10 derivatives have been part of the first and to date only approved drugs based on cyanobacterial specialized metabolites, their biosynthesis has not yet been studied. In the genome of *A. hydrillicola* Thurmond2011, we located a 36.94 kbp long nonribosomal peptide synthetase/polyketide synthase (NRPS/PKS) gene cluster (GenBank accession number ON840103) consisting of twelve genes with deduced functions in AEST biosynthesis (Fig. 3*A* and *SI Appendix,* Table S5). It was located in the middle of a 472 kbp long contig of clearly cyanobacterial origin. The putative *aes* BGC included four NRPS genes (*aesA, C, F,* and *G*), harboring altogether five amino acid incorporating modules (Fig. 3*B*). The predicted substrate specificities of all the respective adenylation domains (*SI Appendix,* Table S6) and the occurrence of two additional *N*-methyltransferase domains agreed with the observed amino acids in AEST. Specifically, a starter unit of AesF was predicted to activate isoleucine and methylate it to form an N-terminal *N*-methylisoleucine residue, followed by the incorporation of valine and another *N*-methylisoleucine by two downstream NRPS modules of AesG. According to our prediction, the subsequent biosynthetic steps are not fully colinear with the arrangement of *aes* genes, analogously to other reported NRPS/PKS biosyntheses (37–39). We suggest that in the next step, the PKS enzyme AesJ, which contains a ketoreductase (KR) domain, accomplishes elongation of the nascent acyl chain by a reduced malonyl (*SI Appendix,* Table S7), as present in Dil. Enzymes encoded in a cassette of genes *aesA-D* likely accomplish the remaining steps in the formation of the AEST backbone. AesA was predicted to incorporate proline, followed by another elongating and ketoreducing PKS (AesB) responsible for incorporation of the reduced malonyl residue of Dap (*SI Appendix,* Table S7). The predicted stereo-configuration of ketide substrates of the KR

108
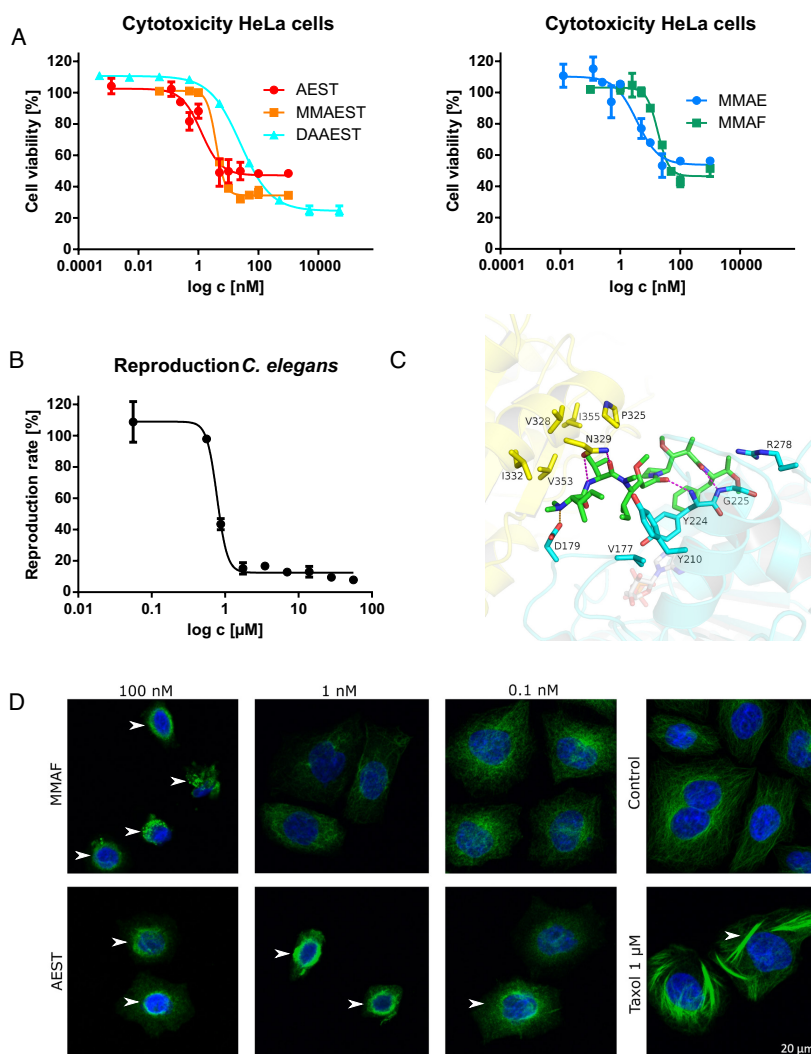
Fig. 2. AEST is a cytotoxic tubulin binder. (*A*) Cytotoxicity of AEST (EC$_{50}$ 1 ± 0.2 nM), MMAEST (EC$_{50}$ 4 ± 0.2 nM), DAAEST (EC$_{50}$ 25 ± 1.4 nM), MMAE (EC$_{50}$ 3 ± 0.4 nM), and MMAF (EC$_{50}$ 18 ± 0.7 nM) on HeLa cells (n = 3). Data are represented as average ± SEM. (*B*) Decrease of the reproduction rate of *C. elegans* treated with AEST (EC$_{50}$ 0.8 ± 0.2 μM, n = 3). Data are represented as average ± SEM. (*C*) Predicted binding of AEST (Aph as 2*S*,3*R*) to tubulin; tubulin α-subunit of one heterodimer visualized as yellow cartoon, the β-subunit of its neighboring heterodimer in cyan. Interacting binding site residues represented as sticks colored the same way. Magenta: hydrogen bonds, orange: salt bridges, red: cation–π interactions. Cofactor GDP shown as white, transparent sticks. (*D*) Immunofluorescence microscopy of HeLa cells showing the effect of AEST and MMAF on the tubulin network (green); nuclei stained with DAPI (blue). Arrows indicate tubulin network depolymerization after AEST and MMAF treatment (disruption of thread-like organization of microtubules). Taxol treatment, used as technical control to show detectability of tubulin interactions, results in tubulin stabilization and bundle formation typical for taxol.

domains embedded in AesJ and AesB match the configuration observed in the produced compounds (Fig. 1 and *SI Appendix,* Fig. S49 and Table S7). The last amino acid, phenylalanine, was predicted to be incorporated by AesC, and elongated by two additional carbons by the activity of the terminal PKS (AesD). AesD further included a thioesterase domain, which was predicted

to catalyze the cleavage of the acyl intermediate from the NRPS/PKS megasynthetase. The remaining biosynthetic step(s) leading to formation of the Aph residue of AEST were partly obscure. Such modification would likely involve a decarboxylation and reduction step (Fig. 3*B*); however, these enzymatic activities are not found in AesD. We hypothesize these reactions could be achieved by a
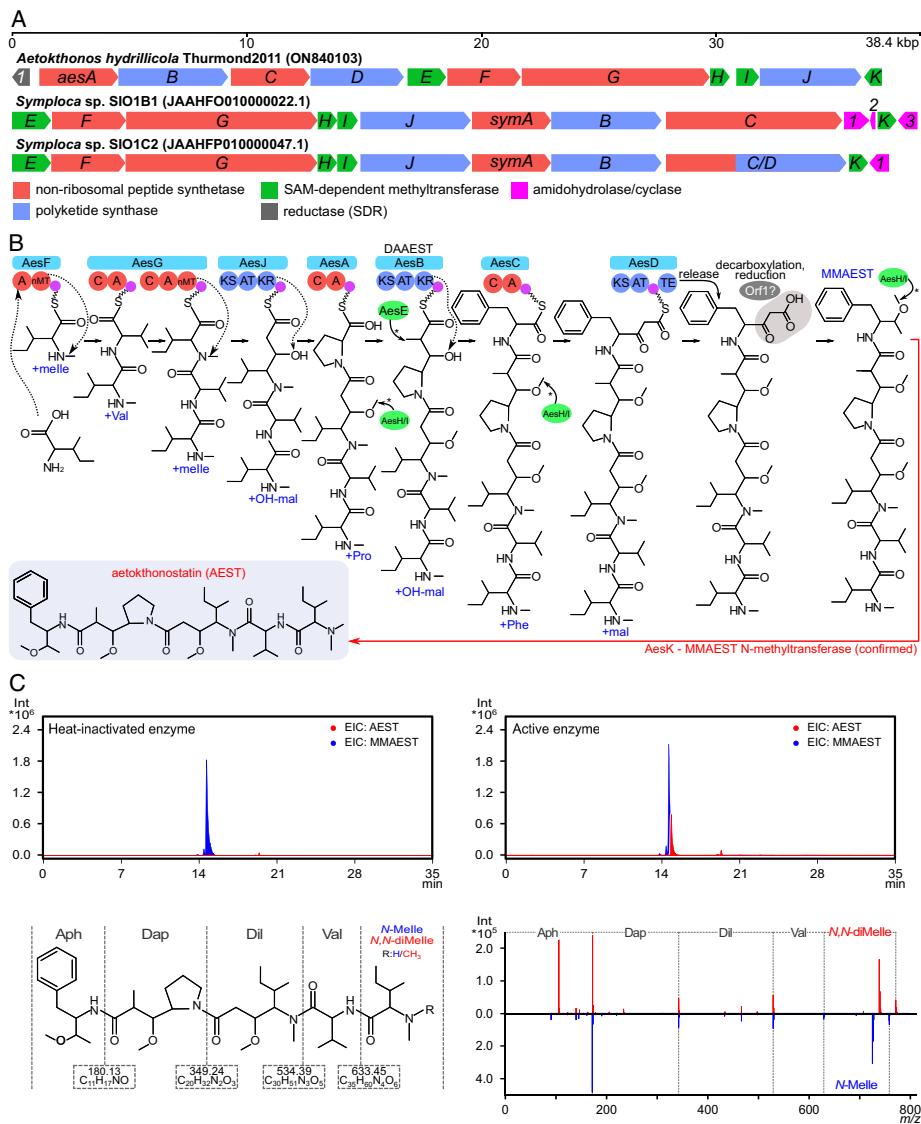
Fig. 3. AESTs are synthesized via an NRPS/PKS pathway. (*A*) Organization of the AEST BGC compared to homologous putative BGCs of members of the dolastatin compound family found in publicly available metagenome-assembled genomes of environmental marine *Symploca* sp. The *aes* orthologs identified in *Symploca* sp. are labeled *symA–K* as they presumably do not code for AESTs but rather symplostatins. The order of *symA–K* genes is collinear with the proposed biosynthesis of the symplostatin backbone scaffold. (*B*) Scheme of the predicted biosynthetic assembly of AEST. The chronology of the tailoring methylation reactions catalyzed by AesE, H, and I (marked with asterisks) is unknown; the reactions are depicted at their first theoretically suitable substrate. The specific target moieties of AesH/I (the Dil/Dap/Aph OH groups) remain to be assigned experimentally. The proposed decarboxylation/reduction of the C-terminal AEST residue (gray box) was not fully explained by bioinformatic analysis. The N-terminal methylation (red arrow) of MMAEST was reconstituted by in vitro enzymatic activity of AesK. (*C*) HPLC-HRMS/MS analysis (extracted ion chromatograms) of AEST (red) in reaction mixtures of MMAEST (blue) incubated with heat-inactivated or fresh Strep-AesK (*Top Left/Right*) show that the inactivated enzyme had no effect on the substrate while the fresh Strep-AesK methylated MMAEST. MMAEST and AEST differ only in the degree of methylation of the N-terminal Ile (*Bottom Left*). Comparison of the MS/MS spectra of MMAEST and the product of the fresh enzyme reaction (*Bottom Right*) shows that AesK methylated the N terminus of MMAEST, producing AEST. Abbreviations: AEST, aetokthonostatin; Aph, aetophenine; BGC, biosynthetic gene cluster; Dap, Dolaproine; Dil, Dolaisoleuine; mal, malonyl-CoA; MMAEST, monomethylaetokthonostatin; *N,N*-di-Melle, *N,N*-dimethylisoleucine; OH-mal, hydroxymalonyl-CoA; SAM, *S*-adenosylmethionine; SDR, short-chain dehydrogenase/reductase.

110

putative short-chain dehydrogenase/reductase (SDR) encoded by a gene (*orf1*) located at the 5′ terminus of the BGC.

The production of minor amounts of a C-terminally truncated NRPS product (DAAEST) could not be explained by our bioinformatics analyses. However, similar cases of partial premature chain termination and release of truncated peptides have been previously reported for other compounds such as the cyanobacterial linear pentapeptide microginins (40). Spontaneous release of NRPS/PKS intermediates has also been demonstrated experimentally, e.g., for bacillaene (41). Furthermore, the last module involved in DAAEST biosynthesis, AesB, ends with a KR domain (Fig. 3*B*). Since SDR domains such as KRs are among known alternative chain-release catalysts in NRPS pathways (42), one could speculate that DAAEST is released by the KR domain of AesB.

In addition to the NRPS/PKS genes, the BGC contained four genes (*aesE, H, I,* and *K*) annotated as class I S-adenosylmethionine (SAM)-dependent methyltransferases (MTs). This observation was in line with multiple methylations found in AEST, specifically the second *N*-methylation of the N-terminal isoleucine, *O*-methylation of hydroxyl groups of the Dil, Dap, and Aph residues, and an additional *C*-methyl group at C7 of AEST. As all four MTs were coded by standalone genes, it was impossible to conclusively infer their target atoms in AEST and the precise chronology of their activity in AEST tailoring—the methylation could occur at the nascent acyl chain or ex post after cleavage of the intermediate from the NRPS/PKS assembly line. We hypothesized that AesE could more likely catalyze the introduction of C7 methyl of AEST, as its closest functionally annotated hit in MIBiG was the *C-C* methyltranserase GphF (*SI Appendix,* Table S5) from the gephyronic acid biosynthetic pathway (43), and it clustered with other C-MTs in our phylogenetic analysis (*SI Appendix,* Fig. S50).

To provide a mechanistic link between the hypothetical *aes* BGC and its putative product, the three MTs AesH, I, and K were heterologously expressed in *Escherichia coli* (*SI Appendix,* Figs. S51 and S52) and subsequently used in in vitro enzymatic assays to elucidate if any of them is involved in the biosynthesis of the *N,N*-dimethylamino group at the N terminus of AEST. To test this hypothesis, aliquots containing the individual MTs and SAM were incubated with the monomethylated, putative late-stage AEST biosynthetic intermediate MMAEST as well as with its synthetic analogue MMAF. MMAF was used as an appropriate alternative substrate to test the specificity of the enzymes, as it likewise possesses a single methylation on the N-terminal residue, in this case *N*-methylvaline.

We detected no modification of the two substrates in the assays with Strep-AesH and Strep-AesI (*SI Appendix,* Figs. S53–S54). However, in the assay with Strep-AesK, we detected the formation of significant amounts of products at *m/z* 760.05, matching with the mass of AEST, in the assay containing MMAEST (Fig. 3*C*), and at *m/z* 746.05 (matching with auristatin F) in the assay containing MMAF (*SI Appendix,* Fig. S55). In both cases, the detected product matched the expected molecular mass corresponding to a single methylation of the provided substrate. HRMS/MS analysis showed that this methylation, as expected, had occurred at the N terminus of the respective substrates. Indeed, the MS/MS spectrum of the reaction product of MMAEST was identical to the MS/MS spectrum of AEST, proving that this compound was formed. (Fig. 3*C* and *SI Appendix,* Fig. S56). We suggest that MMAEST is the true substrate of AesK, as we found AesK unable to methylate the N-terminal amino acids of MMAEST, *N*-methylisoleucine, or MMAF, *N*-methylvaline, when provided as monomeric substrates (*SI Appendix,* Fig. S57).

As AesK only methylates the N terminus of both late-stage biosynthesis analogs, we could confirm its predicted role as the last enzyme of the AEST biosynthetic pathway (Fig. 3 *B* and *C*), demonstrating that the *aes* cluster indeed is responsible for the biosynthesis of AEST. This finding provides additional experimental proof for the cyanobacterial origin of dolastatins (44, 45), offering a cultured and genome-sequenced strain for further laboratory studies. A subsequent search in bioinformatic databases identified highly homologous BGCs in metagenomic assemblies of tropical marine cyanobacteria of the genus *Symploca* (*SI Appendix,* Table S8), from which several dolastatin analogs (e.g., symplostatins) were isolated previously (16, 17, 22, 46). Interestingly, in contrast to the noncolinear arrangement of genes observed in the AEST BGC, the genes in these BGCs from *Symploca* are arranged colinearly with the statin biosynthesis (Fig. 3*A*).

It was intriguing to see that AesK exhibited sufficient substrate specificity to avoid methylation of monomeric *N*-methylisoleucine and *N*-methylvaline, while it was capable to efficiently methylate MMAF, a synthetic dolastatin analog used in antibody–drug conjugate anticancer therapy (47). AesK was originally predicted to be an *O*-MT using automatic bioinformatics tools and phylogenetic analysis (*SI Appendix,* Fig. S50). The substrate specificity of MTs associated with NRPS/PKS clusters offers an example of convergent functional evolution that leads to unexpected enzymatic capabilities (48). Future in-depth studies of AEST biosynthesis, especially the formation of the unique Aph residue, may broaden our knowledge of the cyanobacterial biosynthetic capabilities and be valuable in synthetic biology and medical biotechnology.

Materials and Methods Summary

Full experimental details as well as additional discussions and references for the chemical, biological, and computational methods can be found in *SI Appendix*.

Chemistry. NMR data were recorded on an Agilent/Varian VNMRS 600 MHz spectrometer (5 mm cryoprobe, DMSO-d_6). HRESIMS[2] data were acquired on a Q Exactive Plus mass spectrometer coupled to an UltiMate 3000 HPLC system (Thermo Fisher Scientific). Lyophilized cyanobacteria biomass was homogenized, and extracted three times using methanol/H_2O mixtures (50/50, 80/20, 100 % v/v; sonication, agitation, centrifugation, collection, and drying of the combined supernatants). The extract was dry-loaded (Celite) for flash chromatography with an ACN-H_2O gradient (5 to 100% in 25 min, RP-18 cartridge). Cytotoxic fractions were submitted to RP-HPLC (UltiMate 3000, Phenomenex Luna C18 5 μm, 250 × 10 mm, 6.3 mL/min, ACN-H_2O with 10 mM ammonium acetate, pH 9, linear gradient 60 to 85% in 15 min). The absolute configuration of Dap was determined using Marfey's method.

Biology. *A. hydrillicola* was cultivated in BG11 medium (28 °C, av. light intensity 35 μEm^{-2}s^{-1}, 5% CO_2) and continuously harvested. HeLa cells were cultured in Dulbecco's modified Eagle medium low glucose supplemented with 10% fetal bovine serum and 2 mM glutamine solution at 37 °C (humidified incubator, 5% CO_2). A sulforhodamine B assay was used to evaluate cytotoxicity (2 to 3 triplicates, 96-well plates, 10,000 cells per well). After incubation, staining, washing steps, and photometric read-out, the IC$_{50}$ was determined by nonlinear regression analysis. For immunofluorescence microscopy, HeLa and MDA-MB231 cells were grown in 6-well plates on coated coverslips. As positive technical control, 1 μM taxol was used. After treatment, fixation, staining with formaldehyde, and permeabilization with Triton X-100, the cells were incubated with alpha-tubulin mouse monoclonal antibody, Nuc Blue™ Fixed Cell Stain solution, and Alexa Fluor 488 Rabbit Anti-Mouse IgG Secondary Antibody. After attaching the coverslips to a slide, images were acquired with a laser scanning confocal microscope. *C. elegans* were cultured at 22 °C in petri dishes on NGM agar seeded with *E. coli* OP50. Synchronous populations were obtained by treating gravid worms with 5M NaOH and 5% NaOCl, centrifugation of the lysate, and enrichment of the eggs by density gradient centrifugation (sucrose solution/water). The upper layer was transferred into sterile water (1:3) and centrifuged. The eggs were allowed to hatch in M9 buffer. For the reproduction assay, a 24-well plate was prepared with NGM agar and an *E. coli*

A. Appendix

OP50 solution and incubated. After addition of the test compound solutions and preincubation, three age synchronized L4 larvae were placed in each well. After 72 h, the total number of healthy adult worms were counted, and the IC_{50} was determined by nonlinear regression analysis. The recombinant MTs were codon optimized and transformed in *E. coli* BL21(DE3) cells using an in-house pET-28a-derived vector with N-terminal StrepII-tag. To produce the enzymes, the *E. coli* was grown in autoinducible LB broth medium with kanamycin. After lysis and sonication, the soluble strep-tagged proteins were purified by affinity chromatography (Strep-Tactin Superflow high-capacity resin). Purity of the proteins was confirmed by SDS-PAGE. The activity of the MTs was tested in triplicates, with heat-denatured enzyme and blank (no enzyme) as negative controls, as follows: 100 µM of the substrate and 3 µM of the enzyme were incubated for 3 h at 28 °C in 100 µL assay buffer containing 100 mM Tris–HCl, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, and 1 mM SAM. The reaction was quenched with methanol, and the supernatant obtained by centrifugation was analyzed using HPLC-MS.

Computational Methods. GNPS Molecular Networking Analysis was conducted on HPLC-MS data as described in *SI Appendix*. The previously published genome assembly of *A. hydrillicola* Thurmond2011 was improved in terms of completeness and contiguity by additional sequencing using Illumina MiSeq and Oxford Nanopore MinION, followed by reassembly of all available data (GenBank accession JAALHA020000000.2). The genome assembly was analyzed using antiSMASH v 6.0, and putative NRPS/PKS clusters were further investigated. Substrate specificity of A-domains was predicted using the Stachelhaus code; specificity of PKS AT domains and MT domains was assigned using dedicated pHMM analyses, both available through antiSMASH. A phylogenetic tree (Maximum Likelihood, GAMMA LG model, 1,000 bootstrap iterations) of MT domains was reconstructed using an alignment of bacterial MTs with known function and MTs mined from the BGCs analyzed in this study. Stereoconfiguration of PKS KR domain substrates was predicted based on an alignment of bacterial KR domains of various types together with KR domains encoded in the aes gene cluster, using an analysis of conserved residues within the active site groove. Molecular docking was carried out using four crystal structures of tubulin retrieved from the Protein Data Bank (PDB IDs 4X1I, 4X1K, 4X1Y, 4X20). The protein chains involved in forming the binding pocket were prepared using Schrödinger's Protein Preparation Wizard. Ligprep and Confgen were used to prepare the structures of respective cocrystallized inhibitors as well as AEST and MMAF for docking. All peptides were docked using Glide in Standard-Precision-Peptide mode and utilizing a core constraint

for docking pose restriction. Molecular dynamics simulations were carried out using Amber 18. The preparation steps for all protein–ligand complexes included partial charge assignments, parameterization according to selected force fields, complex solvation, and system neutralization. Prior to the actual production phase comprising 50 ns of MD simulation, each system underwent two stages of energy minimization as well as a heating and a pressure equilibration step.

Data, Materials, and Software Availability. Genomic data of *A. hydrillicola* have been deposited in GenBank (JAALHA000000000.2) (49). NMR and MS raw data are available at figshare (50).

Author affiliations: ªInstitute of Pharmacy, Pharmacognosy, Martin-Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; ᵇBiology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, České Budějovice 37005, Czech Republic; ᶜFaculty of Science, University of South Bohemia, České Budějovice 37005, Czech Republic; ᵈCentre Algatech, Institute of Microbiology of the Czech Academy of Sciences, Třeboň 37901, Czech Republic; ᵉSimris Biologics GmbH, Berlin 12489, Germany; ᶠInstitute of Pharmacy, Medicinal Chemistry, Martin-Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; ᵍMarine Biotoxins Program, Center for Coastal Environmental Health and Biomolecular Research, National Oceanic and Atmospheric Administration/National Ocean Service, Charleston, SC 29412; ʰWarnell School of Forestry and Natural Resources, Fisheries and Wildlife, University of Georgia, Athens, GA 30602; and ⁱInstitute of Pharmacy, Pharmaceutical Biology, Free University of Berlin, Berlin 14195, Germany

1. T. Niedermeyer, M. Brönstrup, "Natural product drug discovery from microalgae" in *Microalgal Biotechnology: Integration and Economy*, W. Posten, Ed. (de Gruyter, 2012), pp. 169–200.
2. K. Tidgewell, B. R. Clark, W. H. Gerwick, "The natural products chemistry of cyanobacteria" in *Comprehensive Natural Products II*, H.-W. Liu, L. Mander, Eds. (Elsevier, 2010), pp. 141–188.
3. J. K. Nunnery, E. Mevers, W. H. Gerwick, Biologically active secondary metabolites from marine cyanobacteria. *Curr. Opin. Biotechnol.* **21**, 787–793 (2010).
4. H. K. Hudnell, Ed., *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs* (Springer, 2008).
5. I. Chorus, J. Bartram, Eds., *Toxic Cyanobacteria in Water: A Guide to their Public Health Consequences, Monitoring, and Management* (CRC Press, 2021).
6. J. M. O'Neil, T. W. Davis, M. A. Burford, C. J. Gobler, The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae* **14**, 313–334 (2012).
7. S. A. Wood *et al.*, Toxic benthic freshwater cyanobacterial proliferations: Challenges and solutions for enhancing knowledge and improving monitoring and mitigation. *Freshw. Biol.* **65**, 1824–1842 (2020).
8. S. Breinlinger *et al.*, Hunting the eagle killer: A cyanobacterial neurotoxin causes vacuolar myelinopathy. *Science* **371**, eaax9050 (2021).
9. D. J. Newman, G. M. Cragg, Marine-sourced anti-cancer and cancer pain control agents in clinical and late preclinical development. *Mar. Drugs* **12**, 255–278 (2014).
10. L. A. Salvador-Reyes, H. Luesch, Biological targets and mechanisms of action of natural products from marine cyanobacteria. *Nat. Prod. Rep.* **32**, 478–503 (2015).
11. L. E. Llewellyn, Saxitoxin, a toxic marine natural product that targets a multitude of receptors. *Nat. Prod. Rep.* **23**, 200–222 (2006).
12. A. P. Thottumkara, W. H. Parsons, J. Du Bois, Saxitoxin. *Angew. Chem. Int. Ed. Engl.* **53**, 5760–5784 (2014).
13. M. Eggen, G. I. Georg, The cryptophycins: Their synthesis and anticancer activity. *Med. Res. Rev.* **22**, 85–101 (2002).
14. J. Rohr, Cryptophycin anticancer drugs revisited. *ACS Chem. Biol.* **1**, 747–750 (2006).
15. R. Bai, G. R. Pettit, E. Hamel, Dolastatin 10, a powerful cytostatic peptide derived from a marine animal. *Biochem. Pharmacol.* **39**, 1941–1949 (1990).
16. G. G. Harrigan *et al.*, Symplostatin 1: A dolastatin 10 analogue from the marine cyanobacterium *Symploca hydnoides*. *J. Nat. Prod.* **61**, 1075–1077 (1998).

17. F. D. Horgen, E. B. Kazmierski, H. E. Westenburg, W. Y. Yoshida, P. J. Scheuer, Malevamide D: Isolation and structure determination of an isodolastatin H analogue from the marine cyanobacterium *Symploca hydnoides*. *J. Nat. Prod.* **65**, 487–491 (2002).
18. M. Beckwith, W. J. Urba, D. L. Longo, Growth inhibition of human lymphoma cell lines by the marine products, dolastatins 10 and 15. *J. Nat. Cancer Inst.* **85**, 483–488 (1993).
19. A. Cormier, M. Marchand, R. B. G. Ravelli, M. Knossow, B. Gigant, Structural insight into the inhibition of tubulin by vinca domain peptide ligands. *EMBO Rep.* **9**, 1101–1106 (2008).
20. R. L. Bai, G. R. Pettit, E. Hamel, Binding of dolastatin 10 to tubulin at a distinct site for peptide antimitotic agents near the exchangeable nucleotide and vinca alkaloid sites. *J. Biol. Chem.* **265**, 17141–17149 (1990).
21. G. R. Pettit *et al.*, The isolation and structure of a remarkable marine animal antineoplastic constituent: Dolastatin 10. *J. Am. Chem. Soc.* **109**, 6883–6885 (1987).
22. H. Luesch, R. E. Moore, V. J. Paul, S. L. Mooberry, T. H. Corbett, Isolation of dolastatin 10 from the marine cyanobacterium Symploca species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1. *J. Nat. Prod.* **64**, 907–910 (2001).
23. H. Luesch, G. G. Harrigan, G. Goetz, F. D. Horgen, The cyanobacterial origin of potent anticancer agents originally isolated from sea hares. *Curr. Med. Chem.* **9**, 1791–1806 (2002).
24. N. Engene, A. Tronholm, L. A. Salvador-Reyes, H. Luesch, V. J. Paul, Caldora penicillata gen. nov., comb. nov. (Cyanobacteria), a pantropical marine species with biomedical relevance. *J. Phycol.* **51**, 670–681 (2015).
25. T. Leão *et al.*, A multi-omics characterization of the natural product potential of tropical filamentous marine cyanobacteria. *Mar. Drugs* **19**, 20 (2021).
26. A. Maderna, C. A. Leverett, Recent advances in the development of new auristatins: Structural modifications and application in antibody drug conjugates. *Mol. Pharmaceut.* **12**, 1798–1812 (2015).
27. M. P. Johansson, H. Maaheimo, F. S. Ekholm, New insight on the structural features of the cytotoxic auristatins MMAE and MMAF revealed by combined NMR spectroscopy and quantum chemical modelling. *Sci. Rep.* **7**, 15920 (2017).
28. L.-F. Nothias *et al.*, Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **17**, 905–908 (2020).
29. S. B. Singh, Discovery and development of dolastatin 10-derived antibody drug conjugate anticancer drugs. *J. Nat. Prod.* **85**, 666–687 (2022).

112

30. S. O. Doronina *et al.*, Development of potent monoclonal antibody auristatin conjugates for cancer therapy. *Nat. Biotechnol.* **21**, 778–784 (2003).

31. S. O. Doronina *et al.*, Enhanced activity of monomethylauristatin F through monoclonal antibody delivery: Effects of linker technology on efficacy and toxicity. *Bioconjug. Chem.* **17**, 114–124 (2006).

32. P. N. Moquist *et al.*, Novel auristatins with high bystander and cytotoxic activities in drug efflux-positive tumor models. *Mol. Cancer Ther.* **20**, 320–328 (2021).

33. K. Miyazaki *et al.*, Synthesis and antitumor activity of novel dolastatin 10 analogs. *Chem. Pharm. Bull.* **43**, 1706–1718 (1995).

34. J. Poncet *et al.*, Synthesis and biological activity of chimeric structures derived from the cytotoxic natural compounds dolastatin 10 and dolastatin 15. *J. Med. Chem.* **41**, 1524–1530 (1998).

35. G. R. Pettit *et al.*, Antineoplastic agents 365. Dolastatin 10 SAR probes. *Anticancer Drug Des.* **13**, 243–277 (1998).

36. I. Zubovych, T. Doundoulakis, P. G. Harran, M. G. Roth, A missense mutation in Caenorhabditis elegans prohibitin 2 confers an atypical multidrug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15523–15528 (2006).

37. N. Traitcheva, H. Jenke-Kodama, J. He, E. Dittmann, C. Hertweck, Non-colinear polyketide biosynthesis in the aureothin and neoaureothin pathways: An evolutionary perspective. *Chembiochem* **8**, 1841–1849 (2007).

38. J. Mareš *et al.*, Alternative biosynthetic starter units enhance the structural diversity of cyanobacterial lipopeptides. *Appl. Environ. Microbiol.* **85**, e02675-18 (2019).

39. N. A. Moss *et al.*, Nature's combinatorial biosynthesis produces vatiamides A-F. *Angew. Chem. Int. Ed. Engl.* **58**, 9027–9031 (2019).

40. N. Eusébio *et al.*, Discovery and heterologous expression of microginins from Microcystis aeruginosa LEGE 91341. *ACS Synth. Biol.* **11**, 3493–3503 (2022).

41. J. Moldenhauer, X.-H. Chen, R. Borriss, J. Piel, Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family. *Angew. Chem. Int. Ed. Engl.* **46**, 8195–8197 (2007).

42. R. F. Little, C. Hertweck, Chain release mechanisms in polyketide and non-ribosomal peptide biosynthesis. *Nat. Prod. Rep.* **39**, 163–205 (2021).

43. J. Young *et al.*, Elucidation of gephyronic acid biosynthetic pathway revealed unexpected SAM-dependent methylations. *J. Nat. Prod.* **76**, 2269–2276 (2013).

44. R. Ratnayake *et al.*, Dolastatin 15 from a marine cyanobacterium suppresses HIF-1α mediated cancer cell viability and vascularization. *ChemBioChem* **21**, 2356–2366 (2020), 10.1002/cbic.202000180.

45. H. Luesch, G. G. Harrigan, G. Goetz, F. D. Horgen, The cyanobacterial origin of potent anticancer agents originally isolated from sea hares. *Curr. Med. Chem.* **9**, 1791–1806 (2002).

46. H. Luesch *et al.*, Symplostatin 3, a new dolastatin 10 analogue from the marine cyanobacterium *Symploca* sp. VP452. *J. Nat. Prod.* **65**, 16–20 (2002).

47. G. Lassiter *et al.*, Belantamab mafodotin to treat multiple myeloma: A comprehensive review of disease, drug efficacy and side effects. *Curr. Oncol.* **28**, 640–660 (2021).

48. M. Z. Ansari, J. Sharma, R. S. Gokhale, D. Mohanty, In silico analysis of methyltransferase domains involved in biosynthesis of secondary metabolites. *BMC Bioinform.* **9**, 454 (2008).

49. J. Mareš *et al.*, *Aetokthonos hydrillicola* Thurmond2011, whole genome shotgun sequencing project. NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/JAALHA000000000.2. Deposited 24 June 2022.

50. M. Schwark, T. H. J. Niedermeyer, Analytical data of Aetokthonostatins. Figshare. https://doi.org/10.6084/m9.figshare.22578616.v1. Deposited 3 September 2023.

113

[2]: Total Synthesis of Aetokthonotoxin, the Cyanobacterial Neurotoxin Causing Vacuolar Myelinopathy

This article was published Open Access via Chemistry Europe and can be found under the above mentioned DOI.

Check for updates

Chemistry Europe
European Chemical
Societies Publishing

Total Synthesis of Aetokthonotoxin, the Cyanobacterial Neurotoxin Causing Vacuolar Myelinopathy

Manuel G. Ricardo,*[a, c] Markus Schwark,[b] Dayma Llanes,[a] Timo H. J. Niedermeyer,[b] and Bernhard Westermann*[a]

Dedicated to Prof. Ludger A. Wessjohann on the occasion of his 60th birthday.

Abstract: Aetokthonotoxin has recently been identified as the cyanobacterial neurotoxin causing Vacuolar Myelinopathy, a fatal neurologic disease, spreading through a trophic cascade and affecting birds of prey such as the bald eagle in the USA. Here, we describe the total synthesis of this specialized metabolite. The complex, highly brominated 1,2′-biindole could be synthesized via a Somei-type Michael reaction as key step. The optimised sequence yielded the natural product in five steps with an overall yield of 29%.

Aetokthonotoxin (AETX) is a pentabrominated biindole alkaloid produced by the cyanobacterium *Aetokthonos hydrillicola*, which grows epiphytically on an invasive plant, *Hydrilla verticillata*, in freshwater lakes.[1] The toxin traverses the food chain from animals consuming the plant (and thus also the cyanobacterium) like waterfowl or snails and then on to raptors such as eagles (e.g. American bald eagle) and kites. While the mode-of-action of AETX is still unknown, it has been shown to elicit Vacuolar Myelinopathy in birds, which is characterized by a widespread vacuolization of the myelinated axons in the white matter of the brain and spinal cord of affected animals.[1] AETX has also been found to be highly toxic to the nematode *C. elegans* (LC_{50} 40 nM),[1] but susceptibility of mammals to this toxin has not been studied, yet.

[a] Dr. M. G. Ricardo, D. Llanes, Prof. Dr. B. Westermann
Department of Bioorganic Chemistry
Leibniz Institute of Plant Biochemistry
Weinberg 3, 06120, Halle/Saale (Germany)
E-mail: Bernhard.Westermann@ipb-halle.de

[b] M. Schwark, Prof. Dr. T. H. J. Niedermeyer
Department of Pharmaceutical Biology
Martin-Luther-University Halle-Wittenberg
Institute of Pharmacy/Pharmacognosy
Hoher Weg 8, 06120 Halle (Germany)

[c] Dr. M. G. Ricardo
present address: Department of Biomolecular Systems
Max-Planck Institute of Colloids and Interfaces
Am Mühlenberg 1, 14476 Potsdam (Germany)
E-mail: Manuel.GarciaRicardo@mpikg.mpg.de

In Figure 1, the structure of Aetokthonotoxin (**1**) is depicted, exhibiting several structural features rarely (if at all) found in natural products: i) AETX is the only known naturally produced 1,2′-bi-1*H*-indole. ii) Only one other natural product contains an indole-3-carbonitrile.[2] iii) Most prominent, also from the isotope pattern that is observed in mass spectrometry analyses, are the five bromo substituents accounting for 61% of the molecular weight.

To enable detailed studies on the biological activity of this novel cyanotoxin, and thus to assess its potential impact on human health, a total synthetic route to **1** is a prerequisite. According to our retrosynthetic analysis (Scheme 1), the most obvious challenge is the coupling of two readily prepared indole subunits. Methods for this uncommon 1,2′-biindole linkage have been explored in a very limited extent, using metal-catalyzed ligation or nucleophilic displacement protocols.[3–5] Due to side-reactions based on participation of
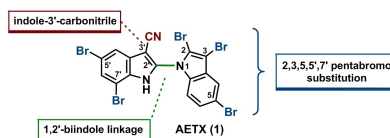

Figure 1. Structure of Aetokthonotoxin (AETX, **1**) with structural features rarely found in natural products.
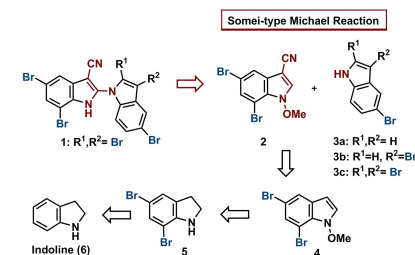

Scheme 1. Retrosynthetic analysis of AETX (**1**).

the bromo-phenylic moieties in these reactions, the use of such procedures seemed inappropriate for our target. Therefore, we devised that a Somei-type Michael coupling fusing advanced indole intermediates could offer a straightforward alternative to access this 1,2'-biindole.[6–8] This envisioned coupling would be the first example of nitrile-modified Michael acceptors and extend the use of this method largely. In addition, in previous literature discussions, it was believed that this reaction could be a biomimetic approach to achieve this hitherto unknown indole-linkage.[5]

Implementing the Somei coupling strategy, the two building blocks would have to be indoles **2** and **3a**–**c** (Scheme 1). The key intermediate **2**, which holds the Michael-acceptor reactivity (3-cyano-*N*-methoxy substitution) demanded by the coupling process, was planned to be obtained either by indole synthesis starting from appropriate brominated precursors, or already functionalized indoles to be brominated subsequently. For the complementary Michael-donors, the brominated indoles **3a**–**c** were suggested, depending on the reactivity in the coupling reaction. While coupling with **3c** would lead directly to the fully substituted natural product, coupling with **3a,b** would require further bromination steps.

For the synthesis of the Michael acceptor **2**, we initially considered a protocol by Penoni et al. to be particularly convenient for the creation of *N*-methoxy-3-substituted indoles.[9,10] As depicted in Scheme 2 (strategy A), this approach consisted of the cycloaddition of nitroso benzenes with electron

deficient alkynes affording desired indole **2**. Preceding the cycloaddition, 2,4-dibromo aniline **7** was efficiently transformed into the nitroso derivative **8**.[11] The cyano acetylene **9** was synthesized from commercially available propiol amide by dehydration (see Supporting Information).[12] Conversely, attempts to use cyano acetylene **9** for the synthesis of **2** led to difficult to separate product mixtures after the cycloaddition reaction due to the formation of polymerized byproducts (see Supporting Information).

A second alternative, depicted in Scheme 2 (strategy B), was envisioned to be based on the protocol developed by Somei et al. for the concerted oxidative transformation of indoline derivatives into *N*-methoxy indoles.[13–15] Here, 5,7-dibromoindoline **5** (available from indoline **6** in 83 % yield) is transformed to **4**, in 72 % yield following slightly modified protocols.[14] Very gratifyingly, the procedure developed by Vorbrüggen et al. using CSI led to conversion of **4** to the wanted nitrile **2** in an isolated yield of 95 %.[16] Overall, the synthesis of **2** could be achieved in gram scale in only three steps starting from commercially available educts.

In the best case, the synthesis of AETX (**1**) would have been achieved by coupling **2** with tribrominated indole **3c** via intermediate **12** by an initial Michael-reaction (Scheme 3). To achieve the 1,4-C–N coupling, carbonyl moieties have been reported on the Micheal acceptor,[14] here for the first time a cyano group is used. The biindole is restored by a loss of the methoxy moiety and rearomatization. The synthesis of this tribrominated indole, however, proved to be quite tedious, as most of the common bromination reagents gave only mixtures of multi-brominated products of di (3,5; **3b**), tri (2,3,5; **3c**) and

Scheme 2. Synthesis of building blocks **2**, and **3b,c**. Reagents and conditions: a) (NH₄)₂SO₄ (5 equiv.), H₂SO₄ (5 equiv.), 0 °C, 4 h, 61 %; b) **9** (5 equiv.), K₂CO₃ (10 equiv.), (CH₃)₂SO₄ (10 equiv.), toluene, 60 °C, 6 h, no reaction; c) PyHBr₃ (2.2 equiv.), DCM, 48 h, 83 %; d) i: H₂O₂ (30 %, 10 equiv.), Na₂WO₄ (cat.), MeOH/H₂O/THF 6 : 1 : 2, 0–5 °C, 1 h; ii: K₂CO₃ (3 equiv.), (CH₃)₂SO₄, (4 equiv.), 0–5 °C, 1 h, 72 % (two steps); e) CSI (1 equiv.), MeCN, 0 °C, 1 h, then DMF, 0 °C, 1 h, 95 %; f) NaH (1.2 equiv.), ClCO₂Me (1.2 equiv.), DMF, 0 °C, 2 h, 94 %; g) Br₂ (8 equiv.), CCl₄, r. t., 12 h, 55 %, h) NaH (1.5 equiv.), MeOH, reflux; 2 h, 88 %; i) PyHBr₃ (1,1 equiv.), DCM, 12 h, 77 %.

Scheme 3. Somei-Michael coupling attempts. Reagents and conditions: a) General conditions of Somei reaction: **3** (1.2 equiv.), NaH (1.1 equiv.), **2** (1 equiv.), DMF, 0 °C to r. t., 12 h; b) NH₄Cl, **13a** (65 %), **13b** (53 %), for **3c**: **14** detected instead of biindole; c) All bromination attempts modify only C-3 in **13a** to form **13b**, when NBS in THF are used, C-6' is then brominated to form **15**.
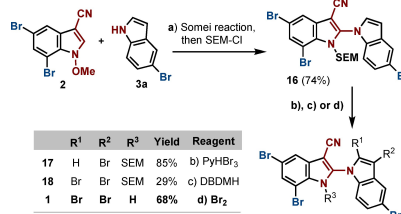
tetra (2,3,5,6)-substituted bromoindoles. Thus, using bromine as bromination agent, **3c** could be obtained in 45% overall yield starting from **3a** (Scheme 2).[17] The synthesis of indole **3b** has been evaluated following numerous established protocols, revealing that the method with PyHBr$_3$ was most efficient in our hands (Scheme 2).[18–20]

For the crucial biindole formation, the coupling step establishing the N1–C2′ linkage could not be achieved when using indole **3c**; steric or stereoelectronic effects may be accountable for this. Different conditions were applied unsuccessfully, with the undesired 2-methoxy-3-cyano-5,7-dibromoindole **14** being the only compound observed, as reported in the literature when poor nucleophiles are used (Scheme 3, Supporting Information).[6,8]

Consequently, indoles **3a** and **3b** were employed in the coupling reactions. Considered as alternatives to the desired coupling, further bromination of the subsequent dimer would be necessary. The Somei couplings performed with monobrominated **3a** and di-brominated **3b**, using NaH as base, confirmed the expected reactivity of the indole **2**, affording the desired N1–C2′ biindoles **13a** and **13b** in 65% and 53% yield, respectively (Scheme 3). With these results, the choice of using the Somei-Michael reaction for the synthesis of AETX became a realistic strategy.

Having in hand advanced biindoles, lacking bromine atoms only at C-2 (**13b**) or C-2 and C-3 (**13a**) with respect to the natural product; brominations at these most reactive positions were considered achievable due to methods already carried out in our studies (e.g. **11**).[17,19,21] However, the very poor solubility of **13a,b** in most common organic solvents reduced the possibilities for the final bromination substantially, as the polarity of the solvent is reported to be crucial in the selectivity of these reactions.[17,19] All bromination attempts, regardless of the agent used, did not afford **1** starting from **13a,b** using THF or DMF as solvents. Position C-3 of **13a** was always smoothly brominated to form **13b**, while the missing bromine substitution at C-2 neither occurred nor took place at other positions rather than C-2. We speculated that the electronic influence of the highly ionizable NH group (¹H NMR signal over 12 ppm) along with a different puckering of the two indole units in such polar solvents would lead to form considerably electron enriched biindole species that guide the bromination to other, unwanted positions. The presence of a bromine atom at C-6′ in a pentabromo biindole obtained in an experiment conducted with **13a**, using NBS in THF, agreed with this hypothesis (Scheme 3).[22]

To extend the bromination alternatives, building block **2** was coupled with **3a**, and subsequently, instead of acidic "work-up", intermediate **12** was quenched with SEM-chloride to yield **16** in 74% yield (Scheme 4). The SEM group was chosen as ideal protecting group to avoid interference of the NH-proton in the required bromination steps and to boost hydrophobicity and, therefore, increase solubility and facilitate purification. According with our previous experiences in brominating procedures, we tried PyHBr$_3$ in DCM first, observing a clean reaction to tetra-brominated **17**. At this stage, attempts to selectively brominate at position C-2 also failed. In further

	R¹	R²	R³	Yield	Reagent
17	H	Br	SEM	85%	b) PyHBr$_3$
18	Br	Br	SEM	29%	c) DBDMH
1	**Br**	**Br**	**H**	**68%**	d) Br$_2$

Scheme 4. Synthesis of AETX (**1**). Reaction conditions: a) i: **3a** (1.2 equiv.), NaH (1.1 equiv.), **2** (1 equiv.), DMF, 0 °C to r. t., 12 h; ii: SEM–Cl (2.5 equiv.), Cs$_2$CO$_3$, (2 equiv.), r. t., 12 h, 74% (two steps); b) PyHBr$_3$ (3 equiv.), DCM, 12 h, 85% (**17**); c) DBDMH (3 equiv.), DCE, reflux, 12 h, 29% (**18**); d) Br$_2$ (5 equiv.), DCE, r. t., 1 h, 68% (**1**).

attempts, we turned our attention to DBDMH in DCE, as it was reported to achieve C-2 and C-3 dibrominaton of indoles.[17] To our delight, we were able, for the first time, to isolate a pentabromo biindole **18**. Subsequent SEM deprotection using 50% TFA in DCM afforded the natural product **1**. Unfortunately, under these conditions, the yield was quite low. Optimization efforts yielded the final product in only 29% despite full consumption of the starting material, as side reactions were observed. However, this result was inspiring, as it indicated the feasibility of the coupling/bromination strategy. Finally, applying 5 equivalents of Br$_2$ in DCE,[19,23–24] a clean reaction could be obtained, leading to the satisfactory bromination at C-2 and C-3 at the same time. Most pleasingly, the reaction conditions also cleaved the SEM-protecting group (triggered by HBr formed in situ) to afford **1** directly and in a one-pot reaction. The bromination/deprotection sequence could be accomplished in 68% yield. All spectroscopic and analytical data of synthesized AETX were in accordance with the isolated natural product.

In summary, we could achieve the first total synthesis of Aetokthonotoxin in five steps with an overall yield of 29%. The crucial biindole-linkage could be realized via a Somei-Michael reaction, the protection/deprotection reactions were both carried out on the fly as quenching reactions. The highly convergent and flexible synthesis presented here may facilitate remarkable opportunities for further biological studies upon derivatization of this alkaloid.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

A. Appendix

[1] S. Breinlinger, T. J. Phillips, B. Haram, J. Mareš, J. A. Martínez Yerena, P. Hrouzek, R. Sobotka, W. M. Henderson, P. Schmieder, S. Williams, J. D. Lauderdale, H. D. Wilde, W. Gerrin, A. Kust, J. W. Wahington, C. Wagner, B. Geier, M. Liebeke, H. Enke, T. H. J. Niedermeyer, S. B. Wilde, *Science* **2021**, *371*, eaax9050.

[2] X. Fu, F. J. Schmith, *J. Nat. Prod.* **1995**, *58*, 1950–1954.

[3] X. Tian, L. Song, M. Rudolph, F. Rominger, T. Oeser, A. S. K. Hashmi, *Angew. Chem. Int. Ed.* **2019**, *58*, 3589–3593; *Angew. Chem.* **2019**, *131*, 3627–3631.

[4] M. F. Comber, C. J. Moody, *Synthesis* **1992**, 731–733.

[5] T. Newhouse, C. A. Lewis, K. J. Eastman, P. S. Baran, *J. Am. Chem. Soc.* **2010**, *132*, 7119–7137.

[6] F. Yamada, D. Shinmyo, M. Somei, *Heterocycles* **1994**, *38*, 273–276.

[7] M. Somei, A. Tanimoto, H. Orita, F. Yamada, T. Ohta, *Heterocycles* **2001**, *54*, 425–432.

[8] F. Yamada, D. Shinmyo, M. Nakajou, M. Somei, *Heterocycles* **2012**, *86*, 435–453.

[9] A. Penoni, G. Palmisano, G. Broggini, A. Kadowaki, K. M. Nicholas, *J. Org. Chem.* **2006**, *71*, 823–825.

[10] G. Ieronimo, A. Mondelli, F. Tibiletti, A. Maspero, G. Palmisano, S. Galli, S. Tollari, N. Masciocchi, K. M. Nicholas, S. Tagliapietra, G. Cravotto, A. Penoni, *Tetrahedron* **2013**, *69*, 10906–10920.

[11] A. M. Churakov, O. Y. Smirnov, S. L. Ioffe, Y. A. Strelenko, V. A. Tartakovsky, *Russ. Chem. Bull.* **1994**, *43*, 1532–1535.

[12] R. J. Halter, R. L. Fimmen, R. J. McMahon, S. A. Peebles, R. L. Kuczkowski, J. F. Stanton, *J. Am. Chem. Soc.* **2001**, *123*, 12353–12363.

[13] M. Somei, T. Kawasaki, *Heterocycles* **1989**, *29*, 1251–1254.

[14] K. Aoki, Y. Nagahama, K. Sugaya, Y. Maeda, H. Sato, K. Nakagawa(Goto), M. Somei, *Heterocycles* **2019**, *98*, 236–270.

[15] M. Somei, K. Yamada, M. Hasegawa, M. Tabata, Y. Nagahama, H. Morikawa, F. Yamada, *Heterocycles* **1996**, *43*, 1855–1858.

[16] H. Vorbrüggen, K. Krolikiewicz, *Tetrahedron* **1994**, *50*, 6549–6558.

[17] J. Yan, T. Ni, F. Yan, *Tetrahedron Lett.* **2015**, *56*, 1096–1098.

[18] V. Bocchi, G. Palla, *Synthesis* **1982**, 1096–1097.

[19] O. R. Suárez-Castillo, L. Beiza-Granados, M. Meléndez-Rodríguez, A. Álvarez-Hernández, M. S. Morales-Ríos, P. Joseph-Nathan, *J. Nat. Prod.* **2006**, *69*, 1596–1600.

[20] M. R. Brennan, K. L. Erickson, F. S. Szmalc, M. J. Tansey, J. M. Thornton, *Heterocycles* **1986**, *24*, 2879–2885.

[21] S. Tang, J. H. Li, Y. X. Xie, N. X. Wang, *Synthesis* **2007**, 1535–1541.

[22] M. Hussain, S. M. Tengho Toguem, R. Ahmad, D. Thanh Tùng, I. Knepper, A. Villinger, P. Langer, *Tetrahedron* **2011**, *67*, 5304–5318.

[23] I. Saikai, A. J. Borah, P. Phukan, *Chem. Rev.* **2016**, *116*, 6837–7042.

[24] K. Piers, C. Meimaroglou, R. V. Jardine, R. K. Brown, *Can. J. Chem.* **1963**, *41*, 2399–2401.

[3]: Occurrence of aetokthonotoxin producer in natural samples - A PCR protocol for easy detection

Lenka Štenclová, Susan B. Wilde, Markus Schwark, Jeffrey L. Cullen, Seth A. McWorther, Timo H. J. Niedermeyer, W. Matthew Henderson, Jan Mareš, Occurrence of aetokthonotoxin producer in natural samples - A PCR protocol for easy detection. *Harmful Algae* 125, 102425-102433 (2023), DOI: 10.1016/j.hal.2023.102425.

A. Appendix

Occurrence of aetokthonotoxin producer in natural samples – A PCR protocol for easy detection

Lenka Štenclová [a,b], Susan B. Wilde [c], Markus Schwark [d], Jeffrey L. Cullen [c], Seth A. McWhorter [c,e], Timo H.J. Niedermeyer [d], W. Matthew Henderson [e], Jan Mareš [a,b,f,*]

[a] Biology Centre of the CAS, Institute of Hydrobiology, České Budějovice, 370 05 Czechia
[b] University of South Bohemia, Faculty of Science, České Budějovice, 370 05 Czechia
[c] Daniel B. Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, 30602 USA
[d] Institute of Pharmacy, Martin Luther University Halle-Wittenberg, Halle (Saale), 06120 Germany
[e] U.S. Environmental Protection Agency, Office of Research and Development, Center for Environmental Measurement and Modeling, Athens, GA, 30605 USA
[f] Centre Algatech, Institute of Microbiology of the CAS, Třeboň, 379 01 Czechia

ABSTRACT

Cyanobacteria are well known producers of bioactive metabolites, including harmful substances. The recently discovered "eagle killer" neurotoxin aetokthonotoxin (AETX) is produced by the epiphytic cyanobacterium *Aetokthonos hydrillicola* growing on invasive water thyme (*Hydrilla verticillata*). The biosynthetic gene cluster of AETX was previously identified from an *Aetokthonos* strain isolated from the J. Strom Thurmond Reservoir, Georgia, USA. Here, a PCR protocol for easy detection of AETX-producers in environmental samples of plant-cyanobacterium consortia was designed and tested. Three different loci of the AETX gene cluster were amplified to confirm the genetic potential for AETX production, along with two variable types of rRNA ITS regions to confirm the homogeneity of the producer's taxonomic identity. In samples of *Hydrilla* from three *Aetokthonos*-positive reservoirs and one *Aetokthonos*-negative lake, the PCR of all four loci provided results congruent with the *Aetokthonos* presence/absence detected by light and fluorescence microscopy. The production of AETX in the *Aetokthonos*-positive samples was confirmed using LC-MS. Intriguingly, in J. Strom Thurmond Reservoir, recently *Hydrilla* free, an *Aetokthonos*-like cyanobacterium was found growing on American water-willow (*Justicia americana*). Those specimens were positive for all three *aet* markers but contained only minute amounts of AETX. The obtained genetic information (ITS rRNA sequence) and morphology of the novel *Aetokthonos* distinguished it from all the *Hydrilla*-hosted *A. hydrillicola*, likely at the species level. Our results suggest that the toxigenic *Aetokthonos* spp. can colonize a broader array of aquatic plants, however the level of accumulation of the toxin may be driven by host-specific interactions such as the locally hyper-accumulated bromide in *Hydrilla*.

1. Introduction

Cyanobacteria are an ancient group of photosynthetic prokaryotes known for their extraordinary repertoire of specialized bioactive metabolites, including potent hepato-, neuro-, and cytotoxins (Tidgewell et al., 2010). In aquatic environments, the growing incidence of planktic cyanobacterial blooms accumulating cyanotoxins has been linked to ecosystem changes induced by human activities. The symptoms of aquatic ecosystem deterioration include eutrophication and pollution (Wurtsbaugh et al., 2019), climate change (rising average temperatures and CO_2 levels) (Huisman et al., 2018), and the spread of non-native

species (Sukenik et al., 2012). Altogether, these are predicted to increase the occurrence of harmful cyanobacterial species (Paerl and Barnard, 2020).

The research focus has extended to non-planktic cyanobacteria and an incredible variety of bioactive compounds have been documented in benthic and soil filamentous cyanobacteria (Costa et al., 2012, Hrouzek et al., 2016). Focusing on neurotoxins, numerous genera of freshwater benthic cyanobacteria were found to produce anatoxin-a, which is responsible for recurrent animal fatalities (Quiblier et al., 2013) as well as saxitoxins (Wood et al., 2020) and ß-methylamino-L-alanine (BMAA) whose toxic and bioaccumulation potential is well documented in

marine food webs (Costa, 2016, Davis et al., 2019).

Furthermore, symbiotic (endophytic) terrestrial cyanobacteria are assumed to accumulate neurotoxic agents transferred and biomagnified along a food chain. An intriguing case is the neurodegenerative disease amyotrophic lateral sclerosis-parkinsonism dementia complex, observed in human population of Guam since the 1950s (Koerner, 1952), and tentatively linked to the neurotoxic cyanobacterial BMAA. According to the hypothesis, BMAA is transferred through a food chain starting with symbiotic cyanobacteria in cycas (*Cycas micronesica*) roots, through flying foxes consuming their fruits and seeds, and eventually impacting the local human population (Kurland, 1988, Cox et al., 2003).

Recently, the novel cyanobacterial neurotoxin aetokthonotoxin (AETX), a peculiar pentabrominated biindole alkaloid implicated in fatal Vacuolar Myelinopathy (VM), was discovered by our team (Breinlinger et al., 2021). This neurodegenerative disease was first recorded in 1994 during an outbreak of bald-eagle poisonings at De Gray Lake in Arkansas, USA (Thomas et al., 1998), and later confirmed in aquatic animals of multiple phyla connected through a shared food web (Birrenkott et al., 2004, Dodd et al., 2016). AETX was experimentally confirmed to be produced by the true branching heterocytous cyanobacterium *Aetokthonos hydrillicola* Wilde and Johansen 2014 (Breinlinger et al., 2021), growing epiphytically on the bottom side of water thyme (*Hydrilla verticillata* (L. f.) Royle) leaves (Wilde et al., 2014). The production of AETX is dependent on bromide (Br⁻) availability, and likely linked to its hyper-accumulation by the host plant (Breinlinger et al., 2021). However, the origin of the increased Br⁻ levels at affected localities is not well understood and may include industrial pollution such as coal power plants and bromine production from brine (Good and VanBriesen, 2017; Vainikka and Hupa, 2012) as well as herbicide treatment (Glomski et al., 2005) and other anthropogenic sources.

H. verticillata is considered strongly invasive in most of the southeastern USA. While its global range of distribution includes vast areas of the temperate and the subtropical zones on both north and south hemispheres (Rojas-Sandoval, 2022), its symbiotic relationship with *A. hydrillicola* was reported thus far only from artificial reservoirs in southeastern USA (Breinlinger et al., 2021). Hence, the poisonings by AETX might be a consequence of multiple combined anthropogenic stressors (pollution, herbicide treatment, introduction of non-native species), possibly with some specific conditions present in these semi-artificial environments. However, a future large-scale expansion of *A. hydrillicola* still cannot be excluded. *Hydrilla's* expansion in the US reservoirs has been monitored since the 1970s (Haller and Sutton, 1975) and mostly unsuccessful attempts to mitigate its spread are almost as old (Haller, 2014; Gerrin et al., 2022). Unfortunately, the reservoirs not only host abundant wildlife but also serve as resources of drinking water to humans. Therefore, regular monitoring of *A. hydrillicola* (accompanied by assessment of Br⁻ and AETX levels) is highly advisable to predict the possible threat of further VM outbreaks.

The growing concern of toxic cyanobacteria, especially in freshwater resources, has stimulated the development of standardized methods for their routine detection (Jaramillo and OShea, 2019). These include, in particular, analytical assessment using LC-MS techniques, accompanied by enzymatic activity assays (Gaget et al., 2017b). In parallel, the cyanotoxin producers can be detected in environmental samples using PCR and qPCR-based methods (Al-Tebrineh et al., 2012), including highly sensitive loop-mediated isothermal amplification (Ramya et al., 2018). The successful detection of cyanotoxin biosynthetic genes can provide early-warning information about the toxigenic potential of natural cyanobacterial populations even in the absence of the actual cyanotoxin production.

The present study aimed to design a simple and fast PCR protocol suitable for routine monitoring of the AETX biosynthetic gene cluster (BGC) in environmental samples of plant-cyanobacterium consortia. The cluster consisting of six genes was previously identified in two *A. hydrillicola* strains isolated from one of the VM outbreak sites (Breinlinger et al., 2021). For five of the genes, their predicted

biosynthetic function was successfully reconstituted *in vitro* (Adak et al., 2022). *AetA* encodes a FAD-dependent halogenase, *aetB* a cytochrome P450, *aetD* a nitrile synthase, *aetE* a tryptophanase, and *aetF* another FAD-dependent tryptophan halogenase. The function of *aetC* remains unknown (Adak et al., 2022). The unique sequence of fine-tuned mechanism of bromination, biaryl coupling, and carbonitrile synthesis is assumed to offer sufficient specificity in its nucleotide sequence for design of a highly specific PCR protocol.

2. Material and methods

2.1. Environmental sampling

The months of October and November 2021 were selected as an appropriate sampling period based on previously reported seasonal VM occurrence in fall and winter (Wiley et al., 2007). Four artificial reservoirs in the area of occurrence of *A. hydrillicola* in Georgia, USA, were chosen for sampling: Covington Reservoir (COV), Longbranch Reservoir (LB), Tussahaw Reservoir (TUS) and Lake Sinclair (SIN) (Fig. 1, Table 1). At each lake, three independent sampling positions were selected (Fig. 1). The biomass of *Hydrilla* was collected using a custom sampling device consisting of a rake on a rope. At each position, five replicate samples of *Hydrilla* biomass were collected into a 3.785 L (= 1 gallon) collecting plastic bag, which altogether made fifteen individual samples per lake. In three of the sampled lakes (COV, LB and TUS) *Hydrilla* was expected to co-occur with *A. hydrillicola*, the remaining lake (SIN) was expected to host *Hydrilla* without *A. hydrillicola* based on previous observations and was therefore applied as a negative control in the present study. The fifth visited lake was J. Strom Thurmond reservoir (THU), from which the original taxonomic description of *A. hydrillicola* was done, and the model strain called Thurmond2011 was isolated. No *Hydrilla* was found in this lake, however three specimens of American water willow (*Justicia americana* (L.) Vahl) with macroscopically visible epiphytic cyanobacterial colonies were collected for further investigation.

The biomass was stored on ice and immediately transported to the laboratory. In the laboratory, a subsample for microscopic analysis and taxonomic identification was taken from each sample. *Hydrilla* leaves from approximately ten stems from each *Hydrilla* sample (and an adequate amount of *Justicia* specimens containing a mixture of plant parts with visible epiphytic cyanobacteria) were collected and transferred to a 50 mL whirl-pack. Subsequently, the biomass was freeze-dried and stored at -80°C.

2.2. Light microscopy observations

The presence of *A. hydrillicola* on the *Hydrilla* leaves was first confirmed by light and fluorescent microscopy as described in Wilde et al. (2014), using the fresh biomass. The detailed observation of specimens collected at THU (cyanobacterial growths on *Justicia*) was conducted using an Olympus BX-100 microscope equipped with DIC optics. Photographs were made using the Olympus DP-72 digital camera and CellSens software at 400 × magnification (Fig. 2). The morphology of examined specimens was compared to the original species description of *A. hydrillicola* (Wilde et al., 2014).

2.3. Strains, cultivation, DNA extraction

Two available cultured strains of *A. hydrillicola* were used as positive controls containing the AETX BGC: Thurmond2011 and CCALA 1050. The first was found to be contaminated by a high number of heterotrophic bacteria, the latter was almost axenic at the time of analysis. Another set of 32 AETX-nonproducing cyanobacterial strains across multiple taxonomic orders (Strunecký et al., 2022) was compiled to test the specificity of PCR primers (Table S1). All strains were cultivated in liquid BG-11 medium (Rippka et al., 1979) in 50 mL Erlenmeyer flasks at
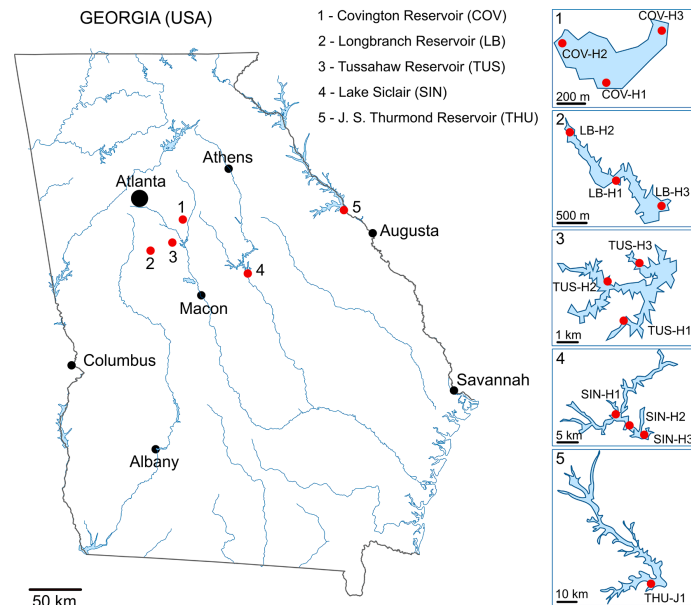
A. Appendix

Fig. 1. Localities examined in present study. The detailed map of the collection sites of individual lakes.

18°C under constant irradiance of 21 µmol m^{-2} s^{-1} photosynthetically active radiation and a 16/8 h light/dark cycle. The biomass was harvested and either dried on silica-gel for 4 days, or directly kept frozen at -80 °C prior to DNA extraction.

In total, biomass of 34 single-species cyanobacterial strains and 63 specimens of lyophilized biomass of natural samples of approximatly 35 mg of dried *Hydrilla* leaves (or *Justicia* specimens) were prepared for DNA extraction. Samples were homogenized in 2 mL Eppendorf tubes using a Retsch MM 200 mill (Retsch, Haan, Germany) and wolfram carbide beads (2 mm diameter) shaken at maximum frequency (25 s^{-1}) for 7 minutes. Subsequently, total gDNA extraction was performed on the whole mixture of homogenized plant and cyanobacterial biomass using NucleoSpin Soil Mini kit (Macherey Nagel, Düren, Germany) according to manufacturers' instructions (including the additional glass bead-beating step), using the lysis buffer S1 and Enhancer SX options.

2.4. Primer design, PCR amplification and sequencing

The PCR protocols were developed based on the previously published BGC of AETX (*aetA-F*, GenBank accession no. MT225528.1) derived from the sequenced genomes of the two *A. hydrillicola* strains used as positive controls.

Primer design was conducted using the Primers tool in Geneious Prime 2022.1.1 software to achieve sequence specificity and optimal melting temperatures, and check for the absence of hairpin and dimer formation at working temperatures. Altogether, 18 primer combinations scattered throughout the whole BGC were tested on the set of 32 AETX-nonproducing cyanobacterial strains (Table S1) to evaluate the specificity of the individual primer pairs towards the AETX BGC. The three

best performing primer pairs were selected (Table 2) based on PCR product specificity and high amplification yield, each amplifying a different locus of the BGC: the genes *aetA*, *aetB*, and the region *aetC-D*, respectively (Fig. 3). Specific primers amplifying the *Aetokthonos* rRNA ITS region were designed (Table 2) based on the data obtained from cloned rRNA operon sequences of the strain Thurmond2011 with comparison of ITS sequences of closely related taxa.

All PCR reactions had a total volume of 25 µL and consisted of 12.5 µL of Plain PP Master Mix Combi (Top Bio, Prague, Czech Republic), 10 pmol of the forward and the reverse primers, 0.5 µL of DNA, and sterile PCR water filling up the remaining volume.

First, the integrity of DNA and content of cyanobacterial DNA was verified in the cultured strains by PCR amplification of the 16S rRNA and ITS region, using universal cyanobacterial primers VFR1 and VFR2 (Boyer et al., 2001). The applied thermocycler settings were: initial denaturation (94°C, 5 min), followed by 40 cycles of denaturation (94°C, 1 min), annealing (54°C, 1 min) and elongation (72°C, 2 min), and a terminal elongation step (72°C for 10 min). PCR product of the *Aetokthonos* strain Thurmond2011 was cloned according to a standard procedure (Johansen et al., 2017) to obtain both variants of the rRNA operon. These data were used as a background for design of *Aetokthonos*-specific ITS primers for both rRNA operon variants.

Amplifications of all three *aet* loci and the *Aetokthonos*-specific ITS were accomplished with the following thermocycler settings: initial denaturation (94°C, 5 min), 40 cycles of denaturation (94°C, 30 s), annealing (60°C, 30 s) and elongation (72°C, 40 s), and terminal elongation (72°C for 10 min). Additionally, the sensitivity of the *aet* BGC primers was tested on a sample set of dilution series of the DNA isolated from the almost axenic *Aetokthonos* strain CCALA 1050 (Fig. S1-F). ITS

3

122

Table 1

Environmental samples investigated in the present study, with summarized results of the molecular and chemical analyses. N – serial number of the sample applied as a unique identifier throughout the present study. Peak area – AETX relative peak area in the MS analysis (normalized to sample biomass dry weight).

N	Sample ID	Positive PCR	LC-MS	Peak area (mg^{-1})	N	Sample ID	Positive PCR	LC-MS	Peak area (mg^{-1})
1	COV-H1-1	A,B,C, ITS1,2	+	3.33×10^6	33	TU-H1-3	A,B,C, ITS1,2	+	2.43×10^7
2	COV-H1-2	A,B,C, ITS1,2	+	1.35×10^7	34	TU-H1-4	A,B,C, ITS1,2	+	2.61×10^7
3	COV-H1-3	A,B,C, ITS1,2	+	1.47×10^7	35	TU-H1-5	A,B,C, ITS1,2	+	8.15×10^7
4	COV-H1-4	A,B,C, ITS1,2	+	4.04×10^6	36	TU-H2-1	A,B,C, ITS1,2	+	1.53×10^7
5	COV-H1-5	A,B,C, ITS1,2	+	1.20×10^6	37	TU-H2-2	A,B,C, ITS1,2	+	2.51×10^5
6	COV-H2-1	A,B,C, ITS1,2	+	1.39×10^7	38	TU-H2-3	A,B,C, ITS1,2	+	2.40×10^6
7	COV-H2-2	A,B,C, ITS1,2	+	5.28×10^6	39	TU-H2-4	A,B,C, ITS1,2	+	7.79×10^5
8	COV-H2-3	A,B,C, ITS1,2	+	1.49×10^7	40	TU-H2-5	A,B,C, ITS1,2	+	1.56×10^6
9	COV-H2-4	A,B,C, ITS1,2	+	9.07×10^5	41	TU-H3-1	A,B,C, ITS1,2	+	2.62×10^7
10	COV-H2-5	A,B,C, ITS1,2	+	1.97×10^7	42	TU-H3-2	A,B,C, ITS1,2	+	2.50×10^7
11	COV-H3-1	A,B,C, ITS1,2	+	6.42×10^5	43	TU-H3-3	A,B,C, ITS1,2	+	2.73×10^7
12	COV-H3-2	A,B,C, ITS1,2	+	2.11×10^6	44	TU-H3-4	A,B,C, ITS1,2	+	1.30×10^7
13	COV-H3-3	A,B,C, ITS1,2	+	2.60×10^5	45	TU-H3-5	A,B,C, ITS1,2	+	1.72×10^7
14	COV-H3-4	A,B,C, ITS1,2	-	-	46	SIN-H1-1	-	-	-
15	COV-H3-5	A,B,C, ITS1,2	+	1.93×10^5	47	SIN-H1-2	-	-	-
16	LB-H1-1	A,B,C, ITS1,2	+	6.14×10^7	48	SIN-H1-3	-	-	-
17	LB-H1-2	A,B,C, ITS1,2	+	1.59×10^8	49	SIN-H1-4	-	-	-
18	LB-H1-3	A,B,C, ITS1,2	+	4.03×10^7	50	SIN-H1-5	-	-	-
19	LB-H1-4	A,B,C, ITS1,2	+	4.58×10^7	51	SIN-H2-1	-	-	-
20	LB-H1-5	A,B,C, ITS1,2	+	3.36×10^7	52	SIN-H2-2	-	-	-
21	LB-H2-1	A,B,C, ITS1,2	+	7.50×10^7	53	SIN-H2-3	-	-	-
22	LB-H2-2	A,B,C, ITS1,2	+	5.83×10^7	54	SIN-H2-4	-	-	-
23	LB-H2-3	A,B,C, ITS1,2	+	3.30×10^7	55	SIN-H2-5	-	-	-
24	LB-H2-4	A,B,C, ITS1,2	+	4.06×10^7	56	SIN-H3-1	-	-	-
25	LB-H2-5	A,B,C, ITS1,2	+	6.52×10^7	57	SIN-H3-2	-	-	-
26	LB-H3-1	A,B,C, ITS1,2	+	1.09×10^8	58	SIN-H3-3	-	-	-
27	LB-H3-2	A,B,C, ITS1,2	+	1.03×10^7	59	SIN-H3-4	-	-	-
28	LB-H3-3	A,B,C, ITS1,2	+	5.70×10^7	60	SIN-H3-5	-	-	-
29	LB-H3-4	A,B,C, ITS1,2	+	3.44×10^7	61	THU-J1-1	A,B,C, ITS2	+	1.32×10^5
30	LB-H3-5	A,B,C, ITS1,2	+	2.80×10^7	62	THU-J1-2	A,B,C, ITS2	+	3.16×10^5
31	TU-H1-1	A,B,C, ITS1,2	+	2.04×10^7	63	THU-J1-3	A,B,C, ITS2	+	4.21×10^4
32	TU-H1-2	A,B,C, ITS1,2	+	1.11×10^7					

PCR products of *Justicia* (THU) specimens were also cloned (Johansen et al., 2017) to avoid neglecting minor content of *A. hydrillicola* or another *Aetokthonos* species. Raw nucleotide sequences were assembled using the *de novo* method in Geneious Prime, all alignments and pairwise p-distance analysis of nucleotide sequences was performed in the same software. The obtained sequences were deposited in GenBank – NCBI under the accession numbers OQ325307, OQ325311–OQ325313, and OQ338813–OQ338815, in case of identical sequences, only a single representative was uploaded.

2.5. Chemical analysis

Approximately 10 mg of lyophilized *Hydrilla* leaves or *Justicia* specimens were mechanically homogenized with a spatula in a reaction tube, resuspended in 3 mL of methanol, and subsequently treated with an ultra sonotrode (Bandelin Sonificator 250, 2 min, 90 % duty cycle) to break up cell walls. To prevent degradation, the samples were maintained on ice during sonification. Samples were placed on an overhead shaker for 0.5 h, centrifuged, and the supernatant was collected. The residues were extracted two more times (1 min sonification, 0.5 h overhead shaking) with the same volume of methanol (9 mL of methanol per 10 mg of dry biomass in total). Supernatants from respective samples were combined and dried using a vacuum centrifuge.

LC-MS analysis was carried out on a Phenomenex Kinetex C18 column (2.1 × 50 mm, 2.6 μm particle size) using a standard gradient from 5-100% (v/v) MeCN in water (+0.1 % formic acid each) in 14 min with a flow rate of 0.4 mL min^{-1}. Final conditions were maintained for an additional 6 min. ESI parameters of general experimental procedure: HRESIMS^2 data were acquired on a Q Exactive Plus mass spectrometer equipped with a heated ESI Interface coupled to an UltiMate 3000 HPLC system (both Thermo Fisher Scientific, Waltham, MA, USA). The following parameters were used for the data acquisition: pos. mode, ESI spray voltage 4.5 kV; neg. mode, ESI spray voltage 2.5 kV, scan range 50 – 2000 *m/z*.

The acquired data set was analyzed with the software Freestyle™ (Thermo Fisher Scientific, Version 1.6). The presence or absence of AETX was determined by comparing the EIC of the *m/z* range 649.6 – 649.7 against an in-house AETX standard.

For semiquantitative analysis, the area under the AETX peak was determined. The obtained values were normalized to the dry weight of extracted biomass from each sample. Statistical single factorial ANOVA analysis was performed using the software GraphPad Prism (Dotmatics, Version 6.01). The alpha level was set to 0.05 and a Turkey's multiple comparison test was used to compare the reservoirs against each other.

3. Results

The specimens of invasive *Hydrilla* with epiphytic cyanobacterium were found growing abundantly throughout three independent freshwater reservoirs in Georgia (COV, TUS, LB), an additional reservoir (SIN) contained *Hydrilla* without macroscopically visible *A. hydrillicola* colonies. Intriguingly, *Aetokthonos* sp. were found growing on the surface of native *Justicia* in THU, a previous VM site where *Hydrilla* has been recently eradicated.

3.1. Microscopic observations

The presence of *A. hydrillicola* on the bottom surface of *Hydrilla* leaves was first confirmed by light- and epifluorescent microscopy and the observations were documented. As expected, the samples from COV, TUS, LB, contained *A. hydrillicola*, whereas SIN was *Aetokthonos*-free. Three specimens of *Justicia* from THU hosted a diverse epiphytic cyanobacterial community on the surface of their leaves, stem and rhizomes. The microbial community was dominated by a true branching heterocytous cyanobacterium, highly reminiscent of *A. hydrillicola*. However, it differed from *A. hydrillicola* in particular morphological
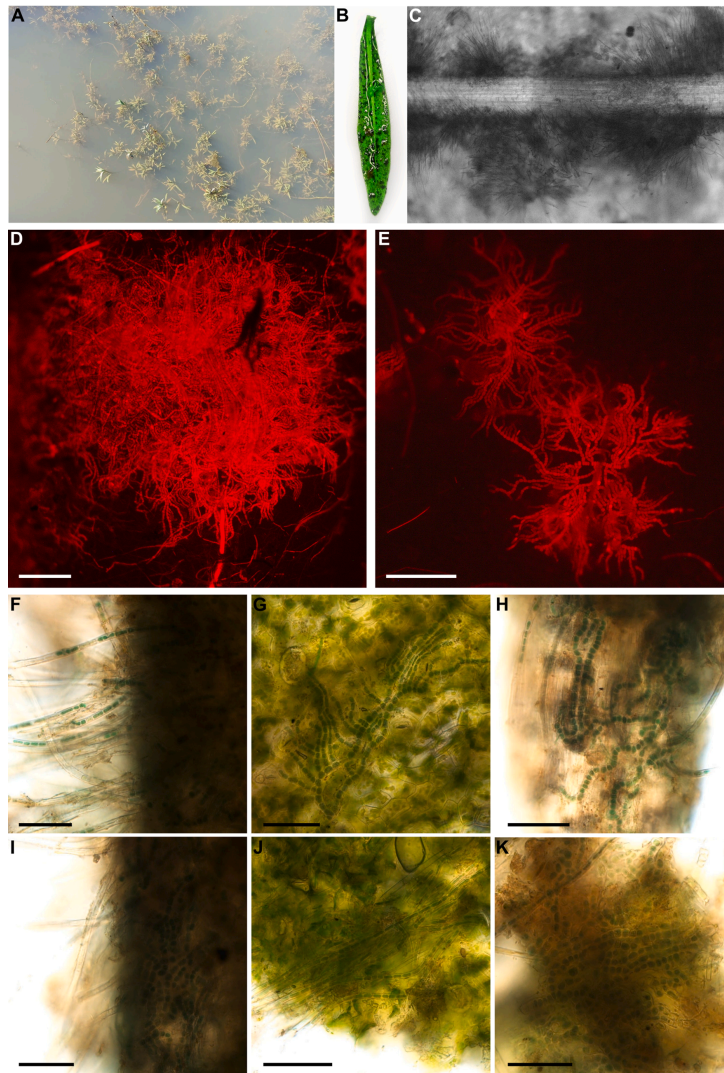
A. Appendix

Fig. 2. Novel *Aetokthonos* sp. growing epiphytically on *Justicia americana*. **A:** submerged vegetation of *J. americana* in J. Strom Thurmond Reservoir. **B-C:** Macroscopically visible colonies of *Aetokthonos* sp. **D-E:** Epifluorescent micrographs of *Aetokthonos* sp.; the scalebars represent 100 μm. **F-K:** Morphology of *Aetokthonos* sp. growing on rhizoids (F,I), leaves (G,J) and stems (H,K) of *J. americana*. The scalebars represent 50 μm. Long lateral branches are clearly visible (C,D,F,J).

5

124

Table 2

Primers used in present study. Tm responds to the melting temperatures predicted by Geneious Prime software. Position represents the site of primer annealing counted from the beginning of the AETX gene cluster or from the beginning of the *Aetokthonos hydrillicola* ITS of the type 1 rRNA operon.

Name	Sequence (5′ to 3′)	Gene	Tm	Position (bp)	Band length (bp)
AETA-F	ACACGAACCACACTTTGCCA	*aetA*	57.3	0813 – 0832	599
AETA-R	CGCTTTCCATTCGCTCGTG	*aetA*	58.8	1412 – 1394	599
AETB-F	TAAGACCATCACAGAAGCTT	*aetB*	54.0	2343 – 2362	608
AETB-R	TGTCAGACTCACTTAGATCG	*aetB*	54.0	2931 – 2950	608
AETCD-F	TCTGGGCTTGAATTTTTCTG	*aetC-D*	54.0	3929 – 3910	530
AETCD-R	TTTGATAGTAGTCGCTCGAA	*aetC-D*	54.0	4439 – 4420	530
A-ITS1-F	GACCTAACCCATCACAAAAC	ITS1	54.2	17 – 36	362
A-ITS1-R	CTTTCTAGCTTGACCACTTTA	ITS1	53.2	362 – 382	362
A-ITS2-F	TCATTACCCCAACTATTCGT	ITS2	53.7	-101 – -81	360
A-ITS2-R	CCTACACTTTCCTATCTGCT	ITS2	53.9	240 – 59	360
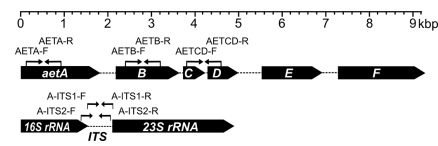
Fig. 3. Gene map of the aetokthonotoxin biosynthetic gene cluster (*aetA-F*) and the rRNA operon of *Aetokthonos hydrillicola* Thurmond2011 with annealing sites of the PCR primers designed in this study.

traits, most remarkably by its disproportionately long lateral branches, with gradually tapering and elongating cells (Fig. 2).

3.2. PCR detection of aet genes

Among the set of tested PCR primers for the AETX BGC, three pairs (Table 2) spanning the genes *aetA-D* (Fig. 3) were selected based on PCR product specificity and high amplification yield to be the best candidates for routine monitoring. These primers unambiguously amplified a single PCR product in positive controls as well as *Aetokthonos*-positive environmental *Hydrilla* samples (Fig. 4), and no PCR product was obtained in the array of negative control cyanobacterial strains (Table S1, Fig. S1). The primers exhibited high sensitivity, successfully amplifying the target loci at gDNA concentrations in the final extract at least as low as 5.10^{-4} ng.μL^{-1} (Fig. S1-F). All the specimens from COV, TUS, LB, were detected as positive in all three tested loci of the AETX BGC, which was confirmed by sequencing of the PCR products. The same applied to the specimens of *Justicia* from THU, yielding PCR products with identical nucleotide sequences of the amplified *aet* loci to the *Hydrilla* samples (Fig. 4, Table 1).

3.3. ITS rRNA analysis

The presence of two different rRNA operon variants and identical ITS sequences within these variants confirmed the genetic homogeneity of *A. hydrillicola* growing on *Hydrilla* across the three reservoirs.

In *Justicia* samples from THU, both primer combinations amplified just one of the ITS variants, a homologue of the *A. hydrillicola* ITS2 operon. The ITS1 primer pair product was very weakly amplified; thus, sequence verification was needed (Fig. S1-A). Therefore, both ITS1 and ITS2 PCR products of sample 63 were analyzed by sequencing a clone library. All 46 obtained sequences of THU specimens were identical to each other (with rare single nucleotide polymorphisms as an artefact of cloning) but they slightly differed from the *Hydrilla*-hosted *A. hydrillicola* sequences, found to be only 92.14 % identical (Fig. S2). No sequence with a higher identity to *A. hydrillicola* was detected among the clones. The results indicated the presence of a unique, although closely related, genotype.

3.4. Analysis of AETX content

LC-MS analysis confirmed the presence of AETX in the biomass of the plant-cyanobacteria samples from all four water bodies (COV, TUS, LB, THU) containing *Aetokthonos* (Table 1), whereas the toxin was not detected in samples from the negative control lake (SIN). Only a single sample of all forty-five *Aetokthonos*-positive and PCR-positive samples was detected as AETX-negative (COV-H3-4). As the semi-quantitative analysis showed a relatively moderate presence of AETX across the samples obtained from lakes COV and TUS (Fig. S3), concentration of the toxin in this sample was perhaps under the detection limit. In the lake LB, LC-MS analysis indicated approximately ten-fold higher amounts of AETX compared to COV and TUS (Fig. S3). Interestingly, in the samples of *Justicia* hosting a cyanobacterial community including the *Aetokthonos* morphotype and PCR-positive for *aet* genes, relatively low AETX production was detected (Fig. S3-4).

4. Discussion

The recently discovered cyanotoxin AETX, responsible for the fatal neurodegenerative animal disease VM, represents a significant concern for wildlife protection and possibly public health (Breinlinger et al., 2021). However, reliable molecular methods to monitor its producer have not yet been developed. Following the results of this study, a set of three pairs of reliable and highly sensitive PCR primers are provided (Table 2) to amplify the biosynthetic genes of AETX from environmental samples. The PCR protocols were found to be specific enough to selectively amplify the AETX loci in the presence of excess plant biomass (when whole leaves with the cyanobacterial biofilm are processed) as well as in a mixed cyanobacterial and microbial community growing on the various plant parts in case of THU samples of *Justicia* (Fig. 2). The PCR tests were negative in a set of AETX-nonproducing cyanobacterial strains across various phylogenetic lineages (Table S1). Furthermore, the sensitivity of the PCR primers (up to 5.10^{-4} ng.L^{-1}) was far below typical DNA extraction yields from environmental samples (Gaget et al., 2017a), even when assuming that cyanobacterial DNA is less abundant compared to the excess plant DNA (Fig. S1-F). Therefore, we suggest that the presented method is simple and accurate enough to be directly applied in routine field monitoring by researchers as well as environmental protection agencies and water management authorities.

The sequence analysis of the obtained PCR products demonstrated the genetic uniformity of *A. hydrillicola* growing on *Hydrilla* at three independent reservoirs, and also two cultured *A. hydrillicola* strains. The sequences were identical in all amplified regions, including four AETX biosynthetic genes and two paralogues of the rRNA ITS. Thus, it is reasonable to expect that the *aet* genes are likely going to be specifically and reliably amplified at other potential VM sites in the USA.

Detection of AETX in the sampled biomass, as provided by LC-MS analysis, was in good correlation with the PCR results – all PCR-negative samples were AETX-negative, and all but one PCR-positive samples contained detectable amounts of AETX (Table 1). As a side
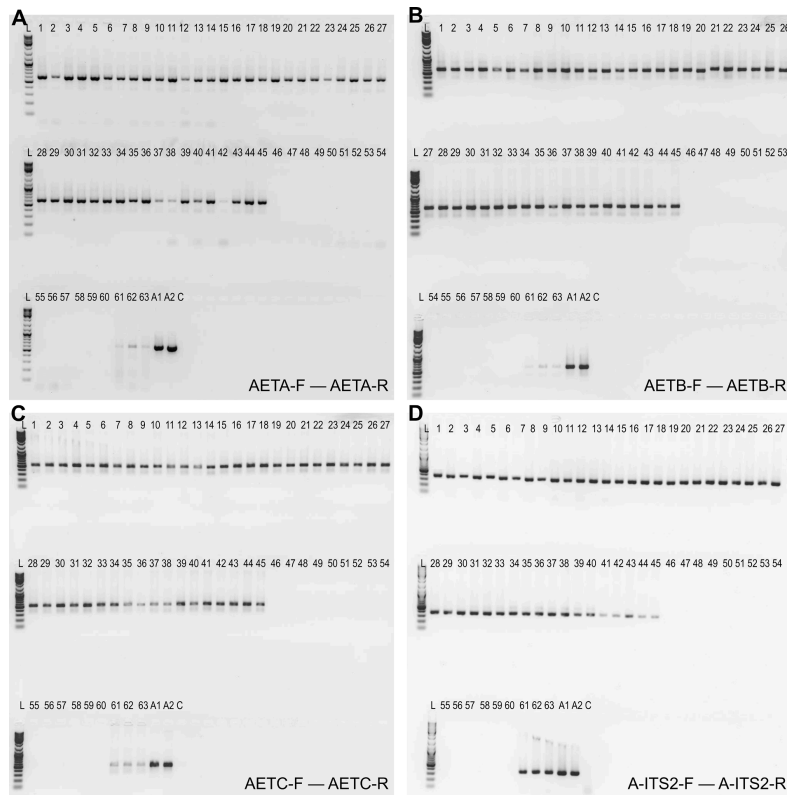
A. Appendix

Fig. 4. Gel electrophoresis (1.5% agarose, 110 V for 45 min, 3.5 μL per sample) of PCR products obtained using primers newly designed in this study. L = ladder (GeneRuler DNA Ladder Mix, Thermo Fisher Scientific). A1 = *Aetokthonos hydrillicola* CCALA 1050, A2 = *Aetokthonos hydrillicola* Thurmond2011 applied as positive controls. C = negative control. Numbers 1-63 represent individual natural samples (Table 1). **A:** *aetA* gene amplified by primer combination AETA-F and AETA-R. **B:** *aetB* gene amplified by primer combination AETA-F and AETA-R. **C:** *aetC – aetD* region amplified by primer combination AETCD-F and AETCD-R. **D:** ITS region of the ribosomal operon type 2 of *Aetokthonos hydrillicola* amplified using primer combination A-ITS2-F and A-ITS2-R. Exact expected sizes of the particular amplified bands are in Table 2.

product of testing the new primers, the presence of the toxigenic genotype and production of AETX at three *Hydrilla* sites previously not studied by molecular or analytical methods (COV, LB, TUS) was discovered. Thus, PCR screening of plant-cyanobacterial associations for *aet* genes may serve as a good predictor for the occurrence of VM in the future, especially if concentrations of available Br⁻ in the water column, sediment, and the host plant tissue are determined simultaneously (Breinlinger et al., 2021). VM was observed in LB previously (Wilde at al., 2014), therefore it is not surprising that *aet* genes were successfully detected here (Fig. 4) and relatively high concentrations of AETX in the *Hydrilla/ Aetokthonos* samples were confirmed (Fig. S3-4).

Molecular detection of cyanotoxin producers using (q)PCR has been widely applied to identify potentially toxigenic species at low concentrations in natural samples, and even under toxin non-producing conditions (Gaget et al. 2017b). At THU, *Hydrilla* was recently completely eradicated by a successful biomanipulation strategy using herbivorous sterile grass carp (*Ctenopharyngodon idella*) (Gerrin at al. 2022). Remarkably, in the current study, cyanobacteria morphologically similar to *A. hydrillicola* were found growing as a component of a cyanobacterial community colonizing the surface of *Justicia* (Fig. 2).

This putatively novel *Aetokthonos* material was very similar in the creeping part of the thallus to *A. hydrillicola* but additionally displayed long erect branches with elongated cells (Fig. 2). Even though truebranched cyanobacteria are known to exhibit extreme morphological plasticity (during their life cycle or in response to environmental conditions), hindering their taxonomic interpretation (Casamatta et al., 2020), the morphological differences detected here could be significant enough to establish a novel species. In the absence of a cultured strain of

126

the *Justicia's* epibiont, it was difficult to conclusively assess whether it represents a novel species or just a morphotype of *A. hydrillicola* growing on a different host plant. However, the absence of the type 1 rRNA operon (or a sequence difference sufficient to avoid amplification by the *Aetokthonos* specific ITS PCR protocol) may indicate substantial genetic diversification. Moreover, the results of sequencing the rRNA ITS2 locus of *Aetokthonos* sp. from *Justicia* revealed a genotype related to the canonical *A. hydrillicola* rRNA ITS2 sequence but exhibiting a pairwise p-distance of 7.86 % (Fig. S2). In cyanobacterial taxonomy, such a level of diversity in the homologous ITS regions has recently been considered as an indicator of variability among two different species (Osorio-Santos et al., 2014). Notably, in the Scytonemataceae family that includes *Aetokthonos* (Strunecký et al., 2022), rare examples of high intragenomic rRNA divergence, content of particular operons (and their total number) were previously documented (Johansen et al., 2017).

Irrespective of the taxonomic questions, *Aetokthonos* growing on *Justicia* contained all the monitored AETX loci in our PCR analysis, exhibiting identical nucleotide sequences. It is worrying that the toxigenic *Aetokthonos* genotype can colonize other host plants than *Hydrilla*, including the indigenous aquatic flora of the USA. Since *Aetokthonos* was never reported or sequenced outside the USA (Breinlinger et al., 2021), it is likely a native cyanobacterium that existed here before the introduction of *Hydrilla*, and we should not be surprised to find members of this genus on other submersed substrates in the future.

Importantly, AETX production was found to be very low in the *Justicia* specimens colonized by *Aetokthonos* sp. (Table 1, Fig. S3-4). The cyanobacterium was relatively less abundant in the samples of *Justicia* plant compared to *Hydrilla* samples, furthermore, the PCR amplification was less pronounced (Fig. 4). Another cause for the low measured AETX concentrations might be that the biosynthesis of AETX is dependent on sufficient supply of dissolved Br⁻. Although THU is one of the previous VM sites (Wilde et al., 2014), and Br⁻ is expected to be available there, the possible explanation may lie in different physiological traits of the host plants. It was shown that *Hydrilla* can hyper-accumulate Br⁻ from the sediment (Breinlinger et al., 2021), and the locally increased Br⁻ levels on the surface of its leaves may therefore be the trigger for AETX production. To test this hypothesis, following studies should further focus on quantification of Br⁻ in water, sediments, and biomass of the host plants (including *Justicia*). However, we can already suggest that elimination of bromide pollution likely represents the most direct and suitable prevention against VM.

Data availability

Sequencing data from this study is available at the NCBI under accession numbers OQ325307, OQ325311–OQ325313 and OQ338813–OQ338815.

Contribution to the manuscript

LS: Participated in sampling and biomass preparation, partly conducted light microscopic observations, developed the PCR protocol, conducted the molecular and taxonomic analyses, had a leading role in the preparation of the manuscript. SBW: Participated in sampling and biomass preparation. Conducted light and fluorescent microscopic observations and participated in their documentation. MS: Conducted LC-MS analyses, their interpretation, and data visualization. JLC: Participated in sampling design, sampling procedures, and biomass preparation. SAM, THJN and WMH: Assisted with design of LC-MS analysis, provided important advice and assisted in interpreting the LC-MS data. JM: Designed, coordinated, and gained funding for the study, assisted with development of the PCR protocol and data analysis, substantially contributed to the manuscript preparation. All authors contributed to the manuscript writing and review.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The sequencing data are available online as explained and other data are available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.hal.2023.102425.

References

Adak, S., Lukowski, A.L., Schafer, R.J.B., Moore, B.S., 2022. From tryptophan to toxin: nature's convergent biosynthetic strategy to aetokthonotoxin. J. Am. Chem. Soc. 144, 2861–2866.

Al-Tebrineh, J., Pearson, L.A., Yasar, S.A., Neilan, B.A., 2012. A multiplex qPCR targeting hepato- and neurotoxigenic cyanobacteria of global significance. Harmful Algae 15, 19–25.

Birrenkott, A.H., Wilde, S.B., Hains, J.J., Fischer, J.R., Murphy, T.M., Hope, C.P., Parnell, P.G., Bowerman, W.W., 2004. Establishing a food-chain link between aquatic plant material and avian vacuolar myelinopathy in mallards (*Anas platyrhynchos*). J. Wildlife Dis. 40, 485–492.

Boyer, S.L., Flechtner, V.R., Johansen, J.R., 2001. Is the 16S-23S rRNA internal transcribed spacer region a good tool for use in molecular systematics and population genetics? A case study in cyanobacteria. Mol. Biol. Evol. 18, 1057–1069.

Breinlinger, S., Phillips, T.J., Haram, B.N., Mareš, J., Yerena, J.A.M., Hrouzek, P., Sobotka, R., Henderson, W.M., Schmieder, P., Williams, S.M., Lauderdale, J.D., Wilde, H.D., Gerrin, W., Kust, A., Washington, J.W., Wagner, C., Geier, B., Liebeke, M., Enke, H., Niedermeyer, T.H.J., Wilde, S.B., 2021. Hunting the eagle killer: a cyanobacterial neurotoxin causes vacuolar myelinopathy. Science 371, 1335.

Casamatta, D.A., Villanueva, C.D., Garvey, A.D., Stocks, H.S., Vaccarino, M., Dvořák, P., Hašler, P., Johansen, J.R., 2020. *Reptodigitus chapmanii* (Nostocales, Hapalosiphonaceae) gen. nov.: a unique nostocalean (Cyanobacteria) genus based on a polyphasic approach. J. Phycol. 56, 425–436.

Costa, M., Costa-Rodrigues, J., Fernandes, M.H., Barros, P., Vasconcelos, V., Martins, R., 2012. Marine cyanobacteria compounds with anticancer properties: a review on the implication of apoptosis. Mar. Drugs 10, 2181–2207.

Costa, P.R., 2016. Impact and effects of paralytic shellfish poisoning toxins derived from harmful algal blooms to marine fish. Fish Fish. 17, 226–248.

Cox, P.A., Banack, S.A., Murch, S.J., 2003. Biomagnification of cyanobacterial neurotoxins and neurodegenerative disease among the Chamorro people of Guam. P. Natl. Acad. Sci. USA 100, 13380–13383.

Davis, D.A., Mondo, K., Stern, E., Annor, A.K., Murch, S.J., Coyne, T.M., Brand, L.E., Niemeyer, M.E., Sharp, S., Bradley, W.G., Cox, P.A., Mash, D.C., 2019. Cyanobacterial neurotoxin BMAA and brain pathology in stranded dolphins. PLoS One 14, e0213346.

Dodd, S.R., Haynie, R.S., Williams, S.M., Wilde, S.B., 2016. Alternate food-chain transfer of the toxin linked to avian vacuolar myelinopathy and implications for the endangered Florida snail kite (*Rostrhamus sociabilis*). J. Wildlife Dis. 52, 335–344.

A. Appendix

Gaget, V., Keulen, A., Lau, M., Monis, P., Brookes, J.D., 2017a. DNA extraction from benthic Cyanobacteria: comparative assessment and optimization. J. Appl. Microbiol. 122, 294–304.

Gaget, V., Lau, M., Sendall, B., Froscio, S., Humpage, A.R., 2017b. Cyanotoxins: which detection technique for an optimum risk assessment? Water Res. 118, 227–238.

Gerrin, W.L., Haram, B., Jennings, C.A., Brandon, G., Wilde, S.B., 2022. Factors affecting movement and habitat use of grass carp in a mainstem reservoir. Fisheries Manag. Ecol. 29, 100–103.

Glomski, L.A.M., Skogerboe, J.G., Getsinger, K.D., 2005. Comparative efficacy of diquat for control of two members of the hydrocharitaceae: Elodea and hydrilla. J. Aquat. Plant Manage. 43, 103–105.

Good, K.D., VanBriesen, J.M., 2017. Power plant bromide discharges and downstream drinking water systems in Pennsylvania. Environ. Sci. Technol. 51, 11829–11838.

Haller, W.T., 2014. Hydrilla. In: Gettys, L.A., Haller, W.T., Petty, D.G. (Eds.), Biology and control of aquatic plants, a best management practices handbook. Aquatic Ecosystem Restoration Foundation, Marietta, pp. 29–34.

Haller, W.T., Sutton, D.L., 1975. Community structure and competition between *Hydrilla* and *Vallisneria*. Fla. Agr. Exp. Sta. 13, 48–50.

Hrouzek, P., Kapuscik, A., Vacek, J., Voráčová, K., Paichlová, J., Kosina, P., Voloshko, L., Ventura, S., Kopecký, J., 2016. Cytotoxicity evaluation of large cyanobacterial strain set using selected human and murine in vitro cell models. Ecotox. Environ. Safe. 124, 177–185.

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M.H., Visser, P.M., 2018. Cyanobacterial blooms. Nat. Rev. Microbiol. 16, 471–483.

Jaramillo, M., O'Shea, K.E., 2019. Analytical methods for assessment of cyanotoxin contamination in drinking water sources. Curr. Opin. Environ. Sci. Health 7, 45–51.

Johansen, J.R., Mareš, J., Pietrasiak, N., Bohunická, M., Zima Jr., J., Štenclová, L., Hauer, T., 2017. Highly divergent 16S rRNA sequences in ribosomal operons of *Scytonema hyalinum* (Cyanobacteria). PLoS One 12, e0186393.

Koerner, D.R., 1952. Amyotrophic lateral sclerosis on Guam - a clinical study and review of the literature. Ann. Intern. Med. 37, 1204–1220.

Kurland, L.T., 1988. Amyotrophic lateral sclerosis and Parkinsons-disease complex on Guam linked to an environmental neurotoxin. Trends Neurosci. 11, 51–54.

Osorio-Santos, K., Pietrasiak, N., Bohunická, M., Miscoe, L.H., Kováčík, L., Martin, M.P., Johansen, J.R., 2014. Seven new species of *Oculatella* (Pseudanabaenales, Cyanobacteria): taxonomically recognizing cryptic diversification. Eur. J. Phycol. 49, 450–470.

Paerl, H.W., Barnard, M.A., 2020. Mitigating the global expansion of harmful cyanobacterial blooms: moving targets in a human- and climatically-altered world. Harmful Algae 96, 101845.

Quiblier, C., Wood, S., Echenique-Subiabre, I., Heath, M., Villeneuve, A., Humbert, J.F., 2013. A review of current knowledge on toxic benthic freshwater cyanobacteria - Ecology, toxin production and risk management. Water Res. 47, 5464–5479.

Ramya, M., Kayalvizhi, M., Haripriya, G., Rathinasabapathi, P., 2018. Detection of microcystin-producing cyanobacteria in water samples using loop-mediated isothermal amplification targeting *mcyB* gene. 3 Biotech 8, 378.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., Stanier, R.Y., 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J. Gen. Microbiol. 111, 1–61.

Rojas-Sandoval, J., 2022. *Hydrilla verticillata* (hydrilla). CABI Compendium, CABI International. https://doi.org/10.1079/cabicompendium.28170.

Strunecký, O., Ivanova, A.P., Mareš, J., 2022. An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. J. Phycol. https://doi.org/10.1111/jpy.13304 online ahead of print.

Sukeník, A., Hadas, O., Kaplan, A., Quesada, A., 2012. Invasion of Nostocales (cyanobacteria) to subtropical and temperate freshwater lakes - physiological, regional, and global driving forces. Front. Microbiol. 3, 86.

Thomas, N.J., Meteyer, C.U., Sileo, L., 1998. Epizootic vacuolar myelinopathy of the central nervous system of bald eagles (*Haliaeetus leucocephalus*) and American coots (*Fulica americana*). Vet. Pathol. 35, 479–487.

Tidgewell, K., Clark, B.R., Gerwick, W.H., 2010. The natural products chemistry of Cyanobacteria. In: Mander, L., Liu, H.-W. (Eds.), Comprehensive Natural Products II. Elsevier Science, Kidlington, pp. 142–183.

Vainikka, P., Hupa, M., 2012. Review on bromine in solid fuels – Part 2: Anthropogenic occurrence. Fuel 94, 34–51.

Wilde, S.B., Johansen, J.R., Wilde, H.D., Jiang, P., Bartelme, B.A., Haynie, R.S., 2014. *Aetokthonos hydrillicola* gen. et sp. nov.: epiphytic cyanobacteria on invasive aquatic plants implicated in Avian Vacuolar Myelinopathy. Phytotaxa 181, 243–260.

Wiley, F.E., Wilde, S.B., Birrenkott, A.H., Williams, S.K., Murphy, T.M., Hope, C.P., Bowerman, W.W., Fischer, J.R., 2007. Investigation of the link between avian vacuolar myelinopathy and a novel species of cyanobacteria through laboratory feeding trials. J. Wildlife Dis. 43, 337–344.

Wood, S.A., Kelly, L.T., Bouma-Gregson, K., Humbert, J.F., Laughinghouse, H.D., Lazorchak, J., McAllister, T.G., McQueen, A., Pokrzywinski, K., Puddick, J., Quiblier, C., Reitz, L.A., Ryan, K.G., Vadeboncoeur, Y., Zastepa, A., Davis, T.W., 2020. Toxic benthic freshwater cyanobacterial proliferations: challenges and solutions for enhancing knowledge and improving monitoring and mitigation. Freshwater Biol. 65, 1824–1842.

Wurtsbaugh, W.A., 2019. Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. WIRES Water 6, e1373.

128

Curriculum vitae

Markus Schwark

Apotheker

Persönliche Angaben

Nachname, Vorname   Schwark, Markus

Beruflicher Werdegang

Mai 2019 – Juli 2023   **Doktorand**, *Martin-Luther-Universität Halle-Wittenberg*, Halle (Saale).
Doktorarbeit "The cyanobacterium *A. hydrillicola* - more than just an "eagle killer"?

Februar 2019 - April 2019   **Apotheker**, *Apotheke an der hohen Leuchte*, Marburg.

Januar 2019   **Approbation zum Apotheker**
3. Staatsexamen - Landesapothekerkammer Hessen, Marburg.

Mai 2018 – Oktober 2018   **Pharmazeut im Praktikum**, *Philipps-Universität*, Marburg.

November 2017 – April 2018   **Pharmazeut im Praktikum**, *Apotheke an der hohen Leuchte*, Marburg.

Bildungsweg

Oktober 2013 – Oktober 2017   **Pharmaziestudium**, *Philipps-Universität*, Marburg.
Staatsexamenstudiengang, Abschluss mit dem 2. Staatsexamen.

August 2004 – Juli 2013   **Allg. Hochschulreife**, *Städtisches Gymnasium*, Lünen.

Halle (Saale), den 20.07.2023

Acknowledgments

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die im Literaturverzeichnis angeführten Quellen genutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen wurden als solche gekennzeichnet. Die vorgelegte Arbeit ist weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt worden. Ich habe keine früheren erfolglosen Promotionsversuche unternommen.

20.07.2023   Unterschrift des Antragstellers