Full Length Article

Structure determination of disease associated peak AAA from l-Tryptophan implicated in the eosinophilia-myalgia syndrome

Klaus Klarskov\textsuperscript{a}, Hugo Gagnon\textsuperscript{b}, Pierre-Luc Boudreau\textsuperscript{a}, Chad Normandin\textsuperscript{a}, Baptiste Plancq\textsuperscript{a}, Eric Marsault\textsuperscript{a}, Gerald J. Gleich\textsuperscript{c}, Stephen Naylor\textsuperscript{d,}\textsuperscript{*}

\textsuperscript{a} Department of Pharmacology and Physiology, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, Quebec J1H SN4, Canada
\textsuperscript{b} PhenoSwitch Bioscience, Sherbrooke, Quebec J1H SN4, Canada
\textsuperscript{c} Departments of Dermatology and Medicine, School of Medicine, University of Utah, Salt Lake City, UT 84132, USA
\textsuperscript{d} ReNeuroGen LLC, Milwaukee, WI 53122, USA

\textbf{A R T I C L E   I N F O}

\textbf{Keywords:}
Eosinophilia-myalgia syndrome
Peak AAA
Case-associated
Dietary supplements
Contaminants

\textbf{A B S T R A C T}

The eosinophilia-myalgia syndrome (EMS) outbreak of 1989 that occurred in the USA and elsewhere was caused by the ingestion of l-Tryptophan (L-Trp) solely manufactured by the Japanese company Showa Denko K.K (SD). Six compounds present in the SD L-Trp were reported to be case-associated contaminants. However, “one” of these compounds, Peak AAA has remained structurally uncharacterized, despite the fact that it was described as “the only statistically significant (p = 0.0014) contaminant”. Here, we employ on-line microcapillary-high performance liquid chromatography-electrospray ionization mass spectrometry (LC-MS), and tandem mass spectrometry (MS/MS) to determine that Peak AAA is in fact two structurally related isomers. Peak AAA\textsubscript{1} and Peak AAA\textsubscript{2} differed in LC retention times, and were determined by accurate mass-LC-MS to both have a protonated molecular ion (MH\textsuperscript{+}) of mass 343.239 Da (Da), corresponding to a molecular formula of C\textsubscript{21}H\textsubscript{30}N\textsubscript{2}O\textsubscript{2}, and possessing eight degrees of unsaturation (DoU) for the non-protonated molecule. By comparing the LC-MS and LC-MS-MS retention times and spectra with authentic synthetic standards, Peak AAA\textsubscript{1} was identified as the intermolecular condensation product of L-Trp with anteiso 7-methylnonanoic acid, to afford (S)-2-amino-3-((S,E)-7-methylnon-1-en-1-yl)-1H-indol-3-yl)propanoic acid. Peak AAA\textsubscript{2} was determined to be a condensation product of L-Trp with decanoic acid, which produced (S)-2-amino-3-(2-(S,E)-7-methylnon-1-en-1-yl)-1H-indol-3-yl)propanoic acid.

\textbf{1. Introduction}

The eosinophilia-myalgia syndrome (EMS) is a chronic, multisystemic disorder characterized by peripheral eosinophilia and subacute onset myalgia (Hertzman et al., 2001; Martin et al., 1990). The disease was first identified during October 1989 in three USA patients who were all taking a l-Tryptophan (L-Trp) dietary supplement (Hertzman et al., 1990). Initial epidemiological studies as well as a national surveillance program initiated by the USA based Centers for Disease Control and Prevention (CDC) revealed a strong association between onset of EMS and consumption of L-Trp (Belongia et al., 1990; Belongia 2004; Kilbourne 1992; Swygert et al., 1990; Varga et al., 1993). The USA Food and Drug Administration (FDA) promptly responded, and on November 11th, 1989 issued a nationwide alert advising consumers to stop consumption of manufactured L-Trp food products. They also requested a nationwide recall of all L-Trp supplements sold over-the-counter. The ensuing aftermath resulted in over 1500 patients being afflicted with EMS, and at least 36 deaths were directly attributed to consumption of L-Trp in the USA alone (Swygert et al., 1993). Numerous additional EMS cases were reported worldwide including the UK, Germany, Canada, Belgium, France, Israel and Japan (Hertzman et al., 1991; UK Committee on Toxicology, 2004). Subsequent analyses by individual USA State health departments and the CDC indicated that EMS was triggered by the consumption of L-Trp manufactured by a single company, Showa Denko K.K (SD) of Japan (Eidson et al., 1990; Philen et al., 1993; Slutsker et al., 1990).

Abbreviations: CDC, USA Centers for Disease Control and Prevention; CE, collision energy; CES, collision energy spread; DoU, degree(s) of unsaturation; ESI, electrospray ionization; EMS, eosinophilia-myalgia syndrome; FDA, USA Food and Drug Administration; HPLC, high performance liquid chromatography; IDA, information-dependent acquisition; LC-MS, microcapillary HPLC mass spectrometry; LC-UV, HPLC with UV detection; L-Trp, L-Tryptophan; NMR, nuclear magnetic resonance; MS/MS, tandem mass spectrometry; MS\textsuperscript{n}, multistage tandem mass spectrometry; PAA, 3-(phenylamino)alanine; SD, Showa Denko K.K; TOF, time-of-flight

\* Corresponding author.
E-mail address: snaylor@rngen.com (S. Naylor).

http://dx.doi.org/10.1016/j.toxlet.2017.10.012
Received 4 June 2017; Received in revised form 19 August 2017; Accepted 10 October 2017
Available online 14 October 2017
0178-4274/ © 2017 Elsevier B.V. All rights reserved.
The SD L-Trp was produced by fermentation using a genetically engineered strain of Bacillus amyloliquefaciens (Belongia et al., 1992; Mayeno and Gleich, 1994). The epidemic was essentially curtailed when the FDA removed L-Trp from the retail market via a recall. Analysis of the SD L-Trp by high performance liquid chromatography (LC) and LC coupled on-line with mass spectrometry (LC–MS) revealed the presence of over sixty contaminants (Muller et al., 1991; Toyo-o’oka et al., 1991; Truckess, 1993; Williamson et al., 1997). Careful and exhaustive epidemiological studies combined with sample lot analyses of this L-Trp revealed six contaminants as being case-associated with the onset of EMS. These specific contaminants were labeled as Peaks UV-5; E; 200; C; FF; and AAA; as defined by their unique LC retention times (Philen et al., 1993; Hill et al., 1993). The ensuing comprehensive analysis of SD priority case lot samples of L-Trp by Hill and coworkers revealed that of the sixty-three contaminants detected by LC–UV, only AAA was statistically significant (p = 0.0014) in terms of association with EMS cases. They concluded that peak AAA “should receive a high priority for isolation and identification” based on their rigorous conditional logistic regression model analysis that was also stratified by time considerations (Hill et al., 1993).

Structural characterization of five of the six case-associated contaminants has been reported previously using Nuclear Magnetic Resonance (NMR), LC–MS and tandem mass spectrometry (MS/MS). Peak UV-5 was identified as 3-(phenylamino)alanine (PAA) after isolation of the contaminant from SD L-Trp (Mayeno et al., 1992; Goda et al., 1992). Peak E was determined to be an acetaldehyde-tryptophan condensation reaction product, namely 1,1′-ethylenebis(tryptophan) using a combination of MS, MS/MS, nuclear magnetic resonance (NMR) and synthetic organic chemistry (Mayeno et al., 1990; Smith et al., 1991). Peak 200 has been identified as 2-(3-indolymethyl)-tryptophan using both NMR (Muller et al., 1991), and utilization of LC–MS and LC–MS/MS (Williamson et al., 1997). Peak C was characterized by accurate mass LC–MS, LC–MS/MS and multistage mass spectrometry (MSn) to be 3a-hydroxy-1,2,3,3a,8,8a-hexahydropyrrolo-[2,3-b]-indole-2-carboxylic acid (Williamson et al., 1998). Peak FF was also subjected to the same analytical protocols as Peak C and identified as 2-(2-hydroxy indoline)-Trp (Williamson et al., 1998). In the case of the “high priority” contaminant peak AAA, to date and to the best of our knowledge, no complete structure determination has been reported. We report here the structural identification of the last, and highest priority case-associated contaminant peak AAA present in SD L-Trp. Peak AAA is actually, two structural aliphatic chain isomers, renamed AAA1 and AAA2.

2. Materials and methods

2.1. Chemicals and reagents

LC–MS grade water, methanol and acetonitrile were purchased from Millipore-Canada Ltd (Etobicoke, ON, Canada). Formic acid was obtained from either Sigma (Markham, ON, Canada) or Millipore-Canada. [Glu1]-fibrinopeptide B was purchased from Sigma. Solid phase Sep-Pak™ C-18 cartridges were obtained from Waters Corporation (Mississauga, ON, Canada). All commercially available synthetic starting materials and solvents were purchased from Sigma-Aldrich Canada (Oakville, Canada), Combi-Block Inc (San Diego, USA), Fisher Scientific Canada (Ottawa, Canada), and used as received. All reactions involving water-sensitive chemicals were carried out in oven-dried glassware with magnetic stirring under argon atmosphere.

2.2. Showa Denko L-Trp

Dr. Rossanne Philen (CDC) provided SD case implicated L-Trp. This sample lot was manufactured between January-June 1989, and had previously been demonstrated as case-implicated in EMS onset (Hill et al., 1993; Mayeno and Gleich, 1994; Philen et al., 1993). Sample storage and handling at the CDC has been described elsewhere (Hill et al., 1993; Philen et al., 1993). We received these samples on September 10th, 1996. All samples were kept in Fisher Scientific polypropylene centrifuge tubes with screw caps, under Nitrogen and further sealed with parafilm. These sample tubes were kept at −20 °C in assorted commercial freezers and out of contact with direct light except in brief instances of sample handling and preparation for analyses. In the case of sample analyses, all samples were prepared fresh on each occasion as described in Section 2.5 below.

2.3. Analytical methods for synthetic L-Trp analogs

The analytical techniques used to purify and characterize the final synthetic L-Trp analog products (1–3) were:

2.3.1. UPHPLC–UV and UPHPLC–MS

Purity was determined on a Waters UPLC H-Class (Mississauga, Canada) with UV detection, equipped with an Acquity UPLC CSH C-18 1.7 μm 2.1 x 50 mm column. MS spectra were recorded on a Waters SQD 2 detector (electrospray) instrument from Waters (Mississauga, Canada) with a linear gradient of 5–95% acetonitrile and water containing 0.1% formic acid. Final products were purified to > 95% purity (UPLC–UV) using a Waters Preparative LC (Sample Manager 2767 Fraction collector), Binary gradient module 2545, with two 515 HPLC pump and a System Fluidics Organizer, Photodiode Array Detector 2998: column X Select CSH Prep C18 5 μm OBD 19 x 250 mm column, buffer: A: 0.1% formic acid in water, B: 0.1% formic acid in acetonitrile, at a flow rate of 20 mL/min.

2.3.2. Proton nuclear magnetic resonance NMR 1H

Proton NMR spectra were recorded with a Bruker Ascend™ 400 (Rheinstetten, Germany) operating at 400 MHz. Samples were dissolved in the indicated deuterated solvent and run at 25 °C. NMR chemical shifts are reported in ppm (parts per million) with reference to the residual peaks of the solvent. The abbreviations for peak multiplicities are described as follows: s (singlet), d (doublet), dd (doublet of doublets), t (triplet), q (quartet), qt (quintet), m (multiplet) and bs (broad singlet) for 1H NMR.

2.4. Synthesis of L-Trp analogs

Details of the syntheses of the L-Trp analogs are described in the supplemental materials. The last steps and characterization of each synthetic compound are described here.

2.4.1. (S)-2-Amino-3-(2-(S,E)-7-methyl-naphthalen-1-yl)-1H-indol-3-yl) propanoic acid (1)

Trifluoroacetic acid (TFA) (1 mL) was added directly to tert-butyl-3-((S)-(S,E)-7-methyl-3-(S,E)-7-methyl-3-oxopropyl)-2-(S,E)-7-methyl-3-oxopropyl)-2-(S,E)-7-methyl-3-oxopropyl)-2-(S,E)

LC–MS retention time: 1.465 min; accurate mass ESI–MS, MS retention time: 1.465 min; accurate mass ESI–MS, final
2.4.1. (S)-2-Amino-3-((E)-8-methylene-1-en-1-yl)-1H-indol-3-yl) propanoic acid (2)

(S,E)-methyl 2-amino-3-(2-(8-methylene-1-en-1-yl)-1H-indol-3-yl) propanoate (60 mg, 0.168 mmol) was dissolved in dioxane and treated with a NaOH (20 mg, 0.378 mmol) solution (1.5:1 dioxane/H₂O) for 30 min. The mixture was neutralized using 1 M HCl, purified by preparative HPLC and then lyophilized to yield 13 mg of (2) as a white solid.

1H NMR (400 MHz, CD₃OD) δ (ppm): 7.61 (d, J = 8.0 Hz, 1H), 7.41 (d, J = 8.0 Hz, 1H), 7.24 (t, 1H), 7.12 (dd, J = 8.0, 7.4 Hz, 1H), 7.01 (dd, J = 7.2, 7.2 Hz, 2H), 3.86 (dd, J = 7.1, 5.3 Hz, 1H), 3.30 (dd, J = 15.0, 5.3 Hz, 1H), 3.17 (dd, J = 15.0, 7.1 Hz, 1H), 2.06-1.97 (m, 2H), 1.97-1.89 (m, 2H), 1.77 (tt, J = 7.2, 7.2 Hz, 2H), 1.33-1.14 (m, 6H), 0.84 (t, J = 6.7 Hz, 3H).

LC-MS retention time: 1.46 min; accurate mass ESI–MS MH⁺ = 343.2307, calculated for C₂₁H₃₁N₂O₂. Purity by LC–UV–MS, 100%.

2.4.3. (S)-2-Amino-3-((E)-dec-1-en-1-yl)-1H-indol-3-yl)propanoic acid (3)

TFA (1 mL) was added directly to tert-butyl-3-((S)-2-((di-tert-butoxycarbonyl)amino)-3-methoxy-3-oxopropyl)-2(5H)-methylene-1-yl)-1H-indole-1-carboxylate (62 mg, 0.047 mmol), stirred for 30 min at room temperature, and then treated with a NaOH (5.4 mg, 0.135 mmol) solution (1.5:1 dioxane/H₂O) for 30 min. The mixture was neutralized using 1 M HCl and purified by preparative HPLC and lyophilized to yield 20 mg of (3) as a white solid.

1H NMR (400 MHz, CD₃OD) δ (ppm): 7.62 (d, 1H), 7.30 (d, 1H), 7.10 (t, 1H), 7.01 (t, 1H), 6.64 (d, 1H), 6.29 (m, 1H), 3.86 (dd, 1H), 3.54 (dd, 1H), 3.13 (dd, 1H), 2.30 (dd, 2H), 1.58-1.47 (m, 2H), 1.45-1.24 (m, 9H), 0.90 (t, 3H).

LC–MS retention time: 1.47 min; accurate mass ESI–MS MH⁺ = 343.2307, calculated for C₂₁H₂₁N₂O₂. Purity by LC–UV–MS, 100%.

2.5. L-Trp sample extraction

SD case-implicated L-Trp was dissolved in 50% aqueous methanol to a concentration of 20 mg/mL. Following sonication for 30 min the sample was centrifuged at 13,000 rpm using a Heraeus Biofuge (Fisher Scientific, Canada). The supernatant was diluted 10 x in water and loaded onto a methanol pre-equilibrated and water rinsed C18 solid phase cartridge. After an initial rinsing of the cartridge with 5% (v/v) aqueous methanol, L-Trp contaminants were eluted with 0.5 mL of methanol and vacuum dried. The sample was dissolved in 100 μL 50% (v/v) aqueous acetonitrile containing 0.2% (v/v) formic acid prior to LC–MS/MS analysis.

2.6. MS and MS/MS of L-Trp synthetic isomers

MS and MS/MS analyses of synthetic N-linked (S,Z)-2-Amino-3-1(dec-4-1-en-1-yl)-1H-indol-3-yl)propanoic acid (N-Alk-Trp), the (S,Z)-2-Amino-3-2(dec-4-1-en-1-yl)-1H-indol-3-yl)propanoic acid (2Z,2-Alk-Trp) analog, as well as the C-2-linked antioiso- and linear L-Trp isomers (1–3) were performed by direct infusion on a Micromass Q-TOF-2 (Waters Corp. ON, Canada). Each synthetic L-Trp sample was diluted to 10 μM in 50% (v/v) aqueous acetonitrile containing 0.2% (v/v) formic acid prior to LC–MS/MS analysis.

2.7. Mass spectrometry of Showa Denko L-Trp

Separations were done online using an Eksigent micro-UHPLC 200 (Sciex, California, US). Samples were injected onto a C-18 reversed phase HALO column, 100 x 0.5 mm ID, 2.7 μm particle size, 90 Å pore size (Sciex, California, US) by loop overfilling using a 5 μL injection loop. A LC gradient consisting of 0.1% (v/v) aqueous formic acid in water as Phase A, and acetonitrile with 0.1% formic acid as Phase B at a flow rate of 35 μL/min. was used with following gradient conditions: (%B/min) 15/0, 15/0.2, 45/5, 5.6/100, 100/6.3, 15/6.4/15/7. All LC-MS and LC-MS/MS analyses were done on a Triple-TOF 5600 (Sciex, California, US) equipped with an electrospray ion source. High resolution time-of-flight (TOF)-MS survey scans (200 ms/sweep, m/z 40–650) were followed by a series of 100 ms product ion scans (0.7 Da Q1 precursor window) with collision energy (CE) set to 35 V and collision energy spread (CES) of 20. The total cycle time was 750 ms. Curtain, ion source 1 and 2 gas were respectively 28, 14 and 17 L/min. In addition it is possible to combine LC-MS survey scans and up to 30 dependent MS/MS scans (accumulation time of 25 ms). This enables the selection of specific precursor ions for accurate mass LC–MS/MS determination directly from the survey scan data.

3. RESULTS & DISCUSSION

Previously we have reported that SD L-Trp contains over sixty different contaminants present at 0.1–0.5% of the parent compound concentration. This determination was predicated on a LC–MS analysis of the SD L-Trp, in which previous analytical conditions reported by Hill were modified (Hill et al., 1993). Such changes involved the replacement of 0.1% TFA with 1% acetic acid in the LC mobile phase and the use of a narrower inner diameter LC column to ensure better interfacing with the ESI-MS interface (Williamson et al., 1997, 1998).

3.1. LC–MS detection of AAA₁ and AAA₂

We employed such analytical changes in a preliminary characterization of AAA in SD L-Trp, using a combination of accurate mass-LC–MS and LC–MS/MS (Klarskov et al., 2000). We determined that AAA had a protonated mass (MH⁺) of 343.2381 Da (Da), and a LC retention time of 50.65 min (Supplemental Material Fig. S1). An accurate mass single ion monitoring LC–MS analysis of MH⁺ = 343.21 ± 343.25 only revealed a single ion response throughout the ion chromatogram. It should also be noted that there were other significant co-eluting ions at m/z 329 and 290 that represent fatty acid homologs of AAA₁ and AAA₂, and these compounds will be discussed in detail elsewhere (Klarskov et al., 2017).

The preliminary characterization of AAA was carried out during the period 1999–2000 and a tentative, unconfirmed structure reported (Klarskov et al., 2000). Since that time there have been further numerous analytical improvements in the LC–MS and LC–MS/MS of complex mixtures of drug metabolites, toxins and trace contaminants. The advent of reduced-flow rate capillary LC (Snyder et al., 2012), micro/nano/pico ESI-MS (Branerjee and Mazumdar, 2012), and routine accurate mass MS and MS/MS (Brenton and...
Godfrey, 2010) have resulted in a myriad of enhanced performance characteristics. They include significantly improved chromatographic resolution of complex clinical or toxicological mixtures, superior MS sensitivity and limit of detection values especially for accurate mass analysis, and facilitation of structural characterization capabilities (Pitt, 2009). We took advantage of these new improvements by employing a microcapillary LC column coupled to a microspray ion source on a Sciex 5600 accurate mass MS and MS/MS instrument in the current analyses.

In the current instance we utilized an accurate mass LC–MS survey scan in conjunction with targeted M/MS and extracted an ion analysis for MH⁺ = 343. The new, enhanced analytical performance capabilities revealed an ion chromatogram containing two distinct ions at MH⁺ = 343 Da, with LC retention times of 5.02 min and 5.30 min (see Fig. 1a). The accurate mass for protonated AAA₁ (shown in Fig. 1b) and AAA₂ (shown in Fig. 1c) was determined to be 343.2388 Da, and 343.2386 Da respectively, indicating, within the margin of error, identical masses. No other ions in the mass range m/z 343.21-343.25, with the appropriate hydrophobic LC retention time properties, were detected in this LC-MS analysis (Fig. 1a). These data clearly indicate that AAA₁ and AAA₂ correspond to the chromatographically unresolved “compound AAA” detected in the original analyses and shown in Supplemental Fig. S1. In the case of AAA₁, there was a smaller co-eluting contaminant ion present at MH⁺ = 343.2969 Da. The possible molecular formulae attributable to the accurate mass and the limited number of degrees of unsaturation (DoU) suggest an unknown contaminant derived from the column that co-eluted only with AAA₁.

The accurate mass determination of AAA₁ and AAA₂ suggests the same molecular formula of C₂₁H₃₀N₂O₂ (excluding the additional H⁺ that provides the protonated positive charge) (Patiny and Borel, 2013). The molecular formula indicates that both AAA₁ and AAA₂ have eight DoU. It is clear from these data that AAA₁ and AAA₂ do not contain a second L-Trp indole ring (as compared to Peak E, 200 and FF) (see Williamson et al., 1997, 1998) but appear to possess aliphatic hydrocarbon functionality, as indicated by increased LC retention time. In addition, based on the accurate mass LC–MS data AAA₁ and AAA₂ are either structural and/or stereoisomers of one another. Finally, the ratio of AAA₁:AAA₂ was determined to be ~3:1, based on the accurate mass ion chromatogram measurements of each peak (Fig. 1a).

3.2. LC–MS/MS and LC–MSⁿ of AAA₁ and AAA₂

We analyzed the SD L-Trp by accurate mass LC–MS/MS in order to...
elucidate the structures of AAA1 and AAA2. All precursor ions at MH⁺ = 343 were subjected to identical ionization and collision conditions. The two resultant product ion spectra for AAA1 (Fig. 2a), and AAA2 (Fig. 2b) were almost identical (note that the contaminant ion MH⁺ = 343.2969 present in Fig. 1b, afforded no discernible product ions under such MS/MS conditions). Since the product ion data was acquired in accurate mass MS/MS mode, it was possible to determine both the molecular ion formula and DoU for each specific product ion (Patiny and Borel, 2013). This is all summarized in Table 1 for both AAA1 and AAA2, in which the measured accurate mass, as well as the calculated molecular mass (predicated on molecular ion formula) is listed. The product ion spectra of AAA1 and AAA2 (Fig. 2a and (ed-highlight b in blue?) b respectively and Table 1) are discussed in detail below.

### 3.2.2. Product ion series at m/z 144, 130 and 118

The product ions at m/z 144 (C8H7N, DoU = 7), 130 (C6H7N, DoU = 7) and 118 (C6H6N, DoU = 6) confirm the existence of an intact indole ring of L-Trp (El Aribi et al., 2004). A concomitant loss of −OH affords the product ion at m/z 130 (C6H7NO, DoU = 10). In addition the product ions at m/z 280 (C20H25N, DoU = 9) and m/z 270 (C19H25N, DoU = 7) represent the loss of NH2/COOH and HCN/COOH respectively, from the precursor ion (El Aribi et al., 2004; Lioe et al., 2004). This ion series and the associated fragmentation losses clearly indicate that both the L-Trp amino and carboxylic acid functional groups in AAA1 and AAA2 are not modified. (see Scheme 1).

### 3.2.3. Product ion series at m/z 326, 308, 280 and 270

This product ion series provides structural information about the amino and carboxylic acid groups of the side-chain of L-Trp. The product ion at m/z 326 (C21H27NO2, DoU = 9) is the result of a facile loss of the −NH3 group from the precursor ion at MH⁺ = 343 (El Aribi et al., 2004; Lioe et al., 2004). A concomitant loss of −OH affords the product ion at m/z 308 (C21H25NO, DoU = 10). In addition the product ions at m/z 280 (C20H25N, DoU = 9) and m/z 270 (C19H25N, DoU = 7) represent the loss of NH2/COOH and HCN/COOH respectively, from the precursor ion (El Aribi et al., 2004; Lioe et al., 2004). This ion series and the associated fragmentation losses clearly indicate that both the L-Trp amino and carboxylic acid functional groups in AAA1 and AAA2 are not modified. (see Scheme 1).

### 3.2.4. Product ion at m/z 168, and related ions at m/z 266 and 182

The MS/MS spectra of both AAA1 and AAA2 contain a prominent product ion at m/z 168 (C13H11N, DoU = 9). The relative abundance of this ion indicates a facile fragmentation process that produces a stable product ion. In order to further characterize this ion we subjected it to LC-MSn. ESI ion source fragmentation of either AAA1 or AAA2 resulted in the formation of m/z 168, which was subjected to further collision-induced dissociation in the collision cell. The resulting MS³ spectrum for AAA1 is shown in Fig. 3. An identical MS⁴ spectrum was produced by AAA2 (data not shown). A search against MS and MS/MS databases including MassBank, the collection of tandem mass spectra of the “NIST/NIH/EPMA Mass Spectral Library 2012”, METLIN, and the ‘Wiley Registry of Tandem Mass Spectral Data, MSforID revealed that the MS⁴ spectrum of AAA1 was very similar to that of the aromatic heterocycle, carbazole (Oberacher, 2013).

It is noteworthy that the product ions at m/z 266 (C14H13N, DoU = 9) and 182 (C13H11N, DoU = 9) are structurally related to m/z 168. MS⁵ analyses of the m/z 266 and 182 ions both afford the carbazole ion at m/z 168 (data not shown). All these MS⁵ data for the product ions m/z 168 and 266/182 indicate that the aliphatic hydrocarbon chain is attached at the C-2 carbon of the indole ring. Furthermore these data suggest that the ninth DoU is due to a double bond at the C1-C2 position of the aliphatic chain. Only this regiochemistry and double bond location facilitates an energetically favoured intramolecular Diels-Alder ring formation (Demarque et al., 2016) of product ions at m/z 266 and 308 to afford the stable tricyclic ion at m/z 168, as well as the as the tricyclic product ions at m/z 266 and 182. This is all captured in Scheme 1.

In order to support the tentative conclusions that the aliphatic chain was attached at the C-2 indole carbon and a double bond was present at the C1-C2 carbons of the hydrocarbon chain, we synthesized two structural analogs of AAA1 and AAA2. The N-linked aliphatic chain isomer, (S,Z)-2-Amino-3-(1-(dec-4-en-1-yl)-1H-indol-3-yl)propanoic...
acid (N-Alk-Trp) was synthesized and subjected to MS/MS analysis. The product ion spectrum (shown in Supplemental Material Fig. S2) did not contain the stable carbazole tricyclic ring fragment at m/z 168. Instead the predominant product ion was a facile fragmentation of the heteroatom indole nitrogen and C1′ of the aliphatic chain (Demarque et al., 2016). This analysis confirms that the regiochemistry of the aliphatic chain was not at the indole nitrogen but at the C-2 of the indole ring. In addition we also synthesized a stereoisomer of AAA1/AAA2 in which the indole ring C-2 was linked to the aliphatic chain but with the double bond at the C4′-C5′ position of the aliphatic chain. This compound, (S,Z)-2-Amino-3-(2-(dec-4-en-1-yl)-1H-indol-3-yl)propionic acid (C2-Alk-Trp) was also subjected to MS/MS analysis as shown in Supplemental Material Fig. S2 and discussed in the Results section.

### Table 1

<table>
<thead>
<tr>
<th>Calculated Mass</th>
<th>Measured Mass</th>
<th>Measured Mass</th>
<th>Measured Mass</th>
<th>Measured Mass</th>
<th>Measured Mass</th>
<th>DoU</th>
<th>Molecular Compositiona</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA1</td>
<td>43.0548</td>
<td>43.0564</td>
<td>43.0563</td>
<td>43.0556</td>
<td>43.0558</td>
<td>1</td>
<td>C3 H7</td>
</tr>
<tr>
<td>AAA2</td>
<td>57.0704</td>
<td>57.0712</td>
<td>57.0699</td>
<td>57.0714</td>
<td>57.0720</td>
<td>1</td>
<td>C4 H9</td>
</tr>
<tr>
<td>Anteiso</td>
<td>71.0861</td>
<td>71.0857</td>
<td>71.0865</td>
<td>71.0856</td>
<td>71.0858</td>
<td>1</td>
<td>C5 H11</td>
</tr>
<tr>
<td>Iso</td>
<td>118.0657</td>
<td>118.0659</td>
<td>118.0645</td>
<td>118.0652</td>
<td>118.0643</td>
<td>6</td>
<td>C8 H8 N</td>
</tr>
<tr>
<td>Linear</td>
<td>130.0657</td>
<td>130.0652</td>
<td>130.0651</td>
<td>130.0656</td>
<td>130.0654</td>
<td>7</td>
<td>C9 H8 N</td>
</tr>
<tr>
<td></td>
<td>144.0813</td>
<td>144.0812</td>
<td>144.0799</td>
<td>144.0808</td>
<td>144.0809</td>
<td>144.0809</td>
<td>C10 H10 N</td>
</tr>
<tr>
<td></td>
<td>168.0813</td>
<td>168.0806</td>
<td>168.0808</td>
<td>168.0807</td>
<td>168.0813</td>
<td>168.0807</td>
<td>C12 H10 N</td>
</tr>
<tr>
<td></td>
<td>182.0970</td>
<td>182.0971</td>
<td>182.0969</td>
<td>182.0958</td>
<td>182.0972</td>
<td>182.0972</td>
<td>C13 H12 N</td>
</tr>
<tr>
<td></td>
<td>266.1909</td>
<td>266.1902</td>
<td>266.1914</td>
<td>266.1903</td>
<td>266.1915</td>
<td>9</td>
<td>C19 H24 N</td>
</tr>
<tr>
<td></td>
<td>270.2222</td>
<td>270.2224</td>
<td>270.2221</td>
<td>270.2211</td>
<td>270.2221</td>
<td>7</td>
<td>C19 H28 N</td>
</tr>
<tr>
<td></td>
<td>280.2065</td>
<td>280.2066</td>
<td>280.2057</td>
<td>280.2078</td>
<td>280.2067</td>
<td>9</td>
<td>C20 H26 N</td>
</tr>
<tr>
<td></td>
<td>326.2120</td>
<td>326.2124</td>
<td>326.2125</td>
<td>326.2120</td>
<td>326.2116</td>
<td>9</td>
<td>C21 H28 N O</td>
</tr>
<tr>
<td></td>
<td>343.2386</td>
<td>343.2389</td>
<td>343.2388</td>
<td>343.2396</td>
<td>343.2390</td>
<td>8</td>
<td>C21 H31 N2 O2</td>
</tr>
<tr>
<td></td>
<td>240.2333c</td>
<td>240.2327</td>
<td>240.2316</td>
<td>240.2333</td>
<td>240.2308</td>
<td>2</td>
<td>C15 H30 N O</td>
</tr>
</tbody>
</table>

DoU for the Uncharged Molecular Species. Molecular Formula for the Protonated Charged Molecular Species.

The ion at MH+ = 240 is a contaminant from the LC Column.

### Scheme 1.

Product ion spectral analysis of AAA1 and AAA2. Proposed product ion fragmentation pathways and structures based on MS/MS and MSn analyses. Note that the masses shown in the Scheme represent the calculated MH+ product ion species. The actual measured product ions for authentic AAA1, AAA2 and synthetic compounds (1–3) are shown in Table 1.
in Supplemental Material Fig. S3. There was no indication of the presence of the product ion at m/z 168, only an abundant product ion at m/z 202 indicative of the intact L-Trp. These data suggest that the assignment of the double bond at the C1’-C2’ position of the aliphatic chain was correct.

3.3. LC–MS and LC–MS/MS of synthetic versus SD L-Trp AAA1 and AAA2

We have discussed above that the aliphatic hydrocarbon chain is linked at the C-2 indole ring with a double bond at the C1’-C2’ alkyl chain. However, the stereochemistry of the C1’-C2’ double bond as well
as identification of the AAA1 and AAA2 aliphatic chain structural isomers needs to be addressed. In that regard consideration of the origin of the aliphatic chains of AAA1 and AAA2 is useful. It is most likely that the aliphatic chains were derived from either the free fatty acids or triglycerides present in the fermentation broth. Bacterial fatty acids are primarily found as the acyl components of phospholipids in the lipid membrane bilayer. Kaneda has stated that membrane fatty acids can be divided into two major families, namely the linear (straight chain) and branched chain fatty acids (Kaneda, 1991). In the latter case, the predominant fatty acids found in Bacillus species are the iso and anteiso structural isomers (see Scheme 1 for a representation of the structural differences of the two isomers). In order to control bacterial cell viability, Bacillus species can change membrane fluidity and permeability by altering fatty acid composition, chain length, and the ratio of branched-to-linear chain structural isomers (Diomande et al., 2015; Murínová and Dercová, 2014). These well described processes and the fatty acid constituents of Bacillus sp. suggest that the aliphatic chains of AAA1 and AAA2 were derived from C10 anteiso, iso or linear fatty acids.

3.3.1. LC-MS/MS

Based on consideration of all the LC-MS, LC-MS/MS and MS<sup>n</sup> data, as well as Bacillus fatty acid biochemistry, we synthesized the anteiso (1), iso (2) and linear (3) L-Trp modified isomers (see Supplemental Materials for detailed synthetic protocols and characterization of (1–3)). The purified synthetic standards were then subjected to accurate mass MS/MS analysis under similar conditions carried out for the analyses of AAA1 and AAA2. The product ion spectra for anteiso (1), iso (2) and linear (3) L-Trp modified isomers (see Supplemental References of the two isomers). In order to control bacterial cell viability, Bacillus species can change membrane fluidity and permeability by altering fatty acid composition, chain length, and the ratio of branched-to-linear chain structural isomers. However, a simple comparative pattern recognition analysis between AAA1/AAA2 and the synthetic standard MS/MS data tentatively suggests AAA1 is the anteiso isomer and AAA2 may be the linear isomer. Nevertheless we still concluded that the MS/MS and MS<sup>n</sup> analyses alone could not provide structural data to definitively characterize AAA1 and AAA2.

3.3.2. LC-MS

Previously Borghi and coworkers have demonstrated that it is possible to separate Teicoplanin T-A-2 isomers using reverse phase HPLC (Borghi et al., 1991). Teicoplanin is a cyclic glycopeptide that contains seven amino acids. It can be isolated in five major isomeric forms designated T-A2-1 through T-A2-5. The isomers only differ in the structure of a single fatty acid linked via a glucosamine residue to the cyclic peptide backbone. The five fatty acid residues are 4-n-decenoc acid (n-C10:1), 8-methylnonoenoic acid (iso-C10:0), n-decanolic acid (n-C10:0), 8-methyldodecanic acid (anteiso-C11:0) and 9-methyldodecanic acid (iso-C11:0). Since there is a striking structural homology to the aliphatic side chains of the synthetic standards (1–3), we developed LC conditions to separate the anteiso (1), iso (2) and linear (3) isomers that were compatible with ESI-MS analysis (Johnson et al., 2013).
Each of the synthetic standards (1–3) was injected separately at an approximate equimolar concentration, and comparative retention times were determined. This is shown in a composite LC–MS ion chromatogram (MH+ = 343) with retention times of 5.01 min. (anteiso (1)), 5.15 min. (iso (2)) and 5.37 min. (linear (3)) for each respective isomer (Fig. 5a). There was a relatively minor contaminant present in the anteiso (1) compound and this can be seen in Fig. 5a, and is denoted by an asterisk. Finally it should be noted that the relative retention time elution profile of the synthetic anteiso: iso: linear isomers (1–3) was the same as reported for the Teicoplanin T-A2 isomers (Borghi et al., 1991).

We subsequently injected an approximate equimolar amount of authentic SD L-Trp (AAA1, (1:3):1; 5.31 min. (AAA2)) respectively. Comparison of retention times (Christie and Han, 2010). Therefore we can further re-assigned as (S). In addition, it has been reported that in the case of Bacillus amyloliquefaciens (Hau et al., 1982). Hence the absolute configuration of the fatty acid derived aliphatic chains in AAA1 and AAA2, contains the linear derived fatty acid chain and confirmed the identity of the fatty acid derived fatty acid chain. In the case of AAA2, it contains the linear derived fatty acid chain and confirms the identity as 2-amino-3-(2-((E)-7-methylnon-1-en-1-yl)-1H-indol-3-yl)propanoic acid. The final structural consideration for AAA1 and AAA2 involves the chiral carbon containing references in retention times of 5.01 min. (anteiso (1)), 5.15 min. (iso (2)) and 5.37 min. (linear (3)) for each respective isomer (Fig. 5a).

1. "Confident identifications are based upon a minimum of two different pieces of confirmatory data relative to an authentic standard." We have provided three such confirmatory data sets relative to characterized synthetic standards. They include accurate mass MS, accurate mass MS/MS and MS^n, and LC–MS retention time data.

2. "Putatively annotated compounds and putatively characterized compound classes" should be compared to the metabolite or other small molecule We have discussed in detail the MS/MS spectra of authentic SD L-Trp derived AAA1 and AAA2 and compared and contrasted these data to that published in the literature/databases for component elements of these contaminants. Consideration of all of these elements satisfy the current criteria for the structure determination of, in this case, two case-associated small molecule contaminants found in SD L-Trp.

4. Conclusions

The thorough structure determination of the two isomers AAA1 and AAA2, finally completes the identification of the "six" original case-associated contaminants identified by Hill and coworkers (Hill et al., 1993; Philen et al., 1993). Simat has argued that the biotechnological manufacturing of L-Trp produces six different types of contaminants that includes i. metabolites; ii. oxidation products; iii. carbonyl condensation compounds; iv. 2-substituent –Trp derivatives; v. 1-substituent –Trp derivatives and vi. PAA and related compounds (Simat et al., 1999). AAA1 and AAA2 are clearly 2-substituent–Trp derivatives. However the presence of the fatty acid derived aliphatic chains in AAA1 and AAA2 will result in very different metabolic and distribution pathways through the body of a person consuming SD L-Trp. Whether this is of relevance in ascertaining the causal onset of EMS is still to be determined.

Conflict of interests

None

Acknowledgements

SN and GJG thank Mayo Foundation and NIH (Grant AI-31155) for funding the original work carried out at Mayo Clinic. KK acknowledges the Natural Sciences and Engineering Research Council of Canada
Supplementary data associated with this article can be found in the online version, at https://doi.org/10.1016/j.toxlet.2017.10.012.

References


