This is the author's accepted manuscript of: Jarmusch, S.A., Feldmann, I., Blank-Landeshammer, B. et al. Cutting the Gordian knot: early and complete amino acid sequence confirmation of class II lasso peptides by HCD fragmentation. J Antibiot (2020). https://doi.org/10.1038/s41429-020-00369-z

Cutting the Gordian knot: Early and complete amino acid sequence confirmation of class II lasso peptides by HCD fragmentation

Scott A. Jarmusch[1†], Ingo Feldmann[2], Bernhard Blank-Landeshammer[2], Carlos Cortés-Albayay[3], Jean Franco Castro[1,3,4‡], Barbara Andrews[3], Juan Asenjo[3], Albert Sickmann[2], Rainer Ebel[1], Marcel Jaspars[1*]

[1] Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Old Aberdeen AB24 3UE, Scotland, U.K.

[2] Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Dortmund, Germany

[3] Centre for Biotechnology and Bioengineering, University of Chile, Beauchef 850, Santiago, Chile.

[4] Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, UK

* Address correspondence to Marcel Jaspars, m.jaspars@abdn.ac.uk

† Present address: Uppsala University, Department of Medicinal Chemistry – Pharmacognosy, Uppsala, Sweden

‡ Present address: Instituto de Investigaciones Agropecuarias, Chillán, Chile

ABSTRACT

Lasso peptides are a diverse class of ribosomally synthesized and post-translationally modified peptides (RiPPs). Their proteolytic and thermal stability alongside their growing potential as therapeutics has increased attention to these antimicrobial peptides. With the advent of genome mining, the discovery of RiPPs allows for the accurate prediction of putatively encoded structures, however, MS^n experiments only provide partial sequence confirmation, therefore 2D NMR experiments are necessary for characterisation. Multiple MS/MS techniques were applied to two structurally characterized lasso peptides, huascopeptin and leepeptin, and one uncharacterized lasso peptide, citrulassin C, which was not isolable in sufficient quantity for NMR analysis. We have shown that MS^2 can be used to elucidate the full amino acid sequences previously predicted with genome mining for this compound class. HCD was able to open the macrocycles and fragment the newly opened linear peptides, confirming the complete amino acid sequences of the characterised lasso peptides. Additionally, to determine if this technique could be applied at the earliest stages of the isolation process, we targeted a lasso peptide found by genome mining, citrulassin C, and were able to fully elucidate the amino acid sequence using only MS^2 from a semi-crude extract of *Streptomyces huasconensis* $HST28^T$.

DEDICATION

This study is dedicated to Professor William Fenical in recognition of his work on establishing marine microbial natural products as a force to be reckoned with.

KEYWORDS

RiPPs/lasso peptides/mass spectrometry/HCD/fragmentation/*Streptomyces*

INTRODUCTION

Lasso peptides are gaining increased attention due to their range of biological activities. Of the four classes of lasso peptides, class II is the largest and displays a wide variety of biological activities, including antimicrobial, anticancer and anti-HIV (1). *Actinobacteria* are the most prolific sources of antimicrobial lasso peptides making them ideal target organisms for the isolation of bioactive lasso peptides (2-12). Even though today between 70-80 lasso peptides have been described, bioinformatic tools like RODEO and RiPPER are showing that we have barely scratched the surface when it comes to appreciating their structural diversity (13-15). These tools, along with others like antiSMASH (16), have made it possible to predict a strain's biosynthetic potential prior to any chemical experimentation, yet, these are just predictions which should not be mistaken for confirmation of a compound being produced. Tietz et al. showed the power of predicting lasso peptide biosynthetic gene clusters and subsequent confirmation using accurate mass and MS^2, but with limitations in MS^n to fragment lasso peptides, these are only partial sequence confirmations showing only fragmentation of the tail (15).

MS^n experiments are an essential tool in cyclic peptide studies but have yet to be fully utilised when studying lasso peptides. Fragmentation of linear peptides is well defined by Peptide Fragmentation Pathways (PFPs) (17) but cyclic peptides tend to open randomly when fragmented (18), resulting in complex MS^n spectra. This phenomenon plagues *de novo* peptide sequencing but has become more manageable with the introduction of bioinformatics and automated mass spectrometric data analysis approaches (19-23). However, putative structures of RiPPs can be predicted using genome mining, therefore, these complex spectra can be somewhat deconvoluted by known fragment ions produced in PFPs. The combination of genome mining and automated mass spectrometry identification culminated in RiPPquest (21), yet without ways to open the single or sometimes multiple cycles characteristic of RiPPs, sequence elucidation by mass spectrometry still requires further structure elucidation and confirmation via NMR based techniques.

Mass spectrometry has been the primary tool for evaluating production of lasso peptides in genome mining approaches as well as in studies involving heterologous expression, but limitations in MS^2 due to the stable macrolactam ring has prevented mass spectrometry from becoming a definitive technique in the characterization of lasso peptides. Full structure characterisation of lasso peptides is a process reliant on 2D NMR experiments (1,24). In

both publications first describing the lasso topology of microcin J25, investigators utilized MS^3 studies to fragment the macrocycle further. Although this did open the ring, it was still inefficient at completely breaking each amino acid bond, resulting in non-sequential fragments that would make *de novo* sequencing difficult (25,26). Recently, it was shown that MS^n experiments using CID and ETD can provide topological information through a combination of identifying interlocked fragments and higher instances of hydrogen migration events, respectively, showing MS^n can be a more powerful tool in lasso peptide characterisation (27). Additionally, a separate methodology was established using trapped ion mobility spectrometry – mass spectrometry (TIMS-MS) to efficiently and quickly differentiate lasso peptides from their unthreaded branched-cyclic topoisomers (28). Both studies show the expanded use of MS^n for further structure characterisation of lasso peptides.

In this study, MS^n experiments were carried out on three lasso peptides, two previously characterized by mass spectrometry and NMR, and one detected in a semi-crude extract from *Streptomyces huasconensis* $HST28^T$, confirming genome mining predictions. All lasso peptides and extracts originated from previous studies working with Atacama Desert *Streptomyces* spp. Collisional-induced dissociation (CID), electron-transfer dissociation (ETD), and higher energy collisional dissociation (HCD) for MS^2 and electron-transfer/higher-energy collisional dissociation (EThcD) for MS^3 were all carried out using stepped collision energies in order to see if one or more of the techniques would carry out macrocycle opening, leading to observable complete fragmentation of the lasso peptides. MS^2 HCD was the only technique that provided enough information to confirm the entire amino acid sequence of all three lasso peptides test.

RESULTS

Fragmentation of huascopeptin

The fragmentation of huascopeptin - Cyclo(GYGNAWD)-SKNGLF-, performed on the $[M+2H]^{2+}$ ion (*m/z* 705.80), was initially evaluated using 35% CID. The resulting spectrum matches the previously reported MS^2 spectra (29), showing sequential fragmentation of the tail and an ion for the macrocycle (*m/z* 746.28). ETchD showed some potential for opening the ring (Figure S1) but low ion counts were observed when compared to HCD. It was observed that HCD with 23-25% relative collision energy (RCE) exhibited the optimal fragmentation of the macrolactam ring and the linear peptide. Huascopeptin's macrocycle (GYGNAWD) was opened randomly, forming 7 linear peptides (Scheme 1). These

7 linear peptides yield a theoretical 42 *b* ions, out of which only 35 are observable due to the mass range cut-off. Out of these 35 ions, 33 were observed in addition to 27 out of 35 total *a* ions. The HCD spectrum of huascopeptin is displayed in Figure 1, along with the zoomed spectrum that displays the mass range for the macrocycle fragment ions. ETD/HCD showed some fragmentation of the ring (Figure S2), but significantly lower ion counts were observed for Trp_6, the primary indicator in the low mass range that the ring is being opened for huascopeptin (Figure S2). A full table of observed fragment ions is available in Table S1.

Overall, there were three observable ring opening events that we can detect in the spectrum of huascopeptin. (1) Cleavage of the tail residues first and subsequent ring opening at random positions as shown in Scheme 1, (2) formation of a branched or "T-shaped" structure where the macrocycle opens in the middle and the tail remains intact (Figure 1), and (3) ring opening at the isopeptide bond first, leading to fragmentation of a linear peptide (Figure 2). While 1 and 3 are difficult to distinguish experimentally, ions resulting from option 2 are indeed observed, but in our experience are less favoured.

Fragmentation of leepeptin

Similar to huascopeptin, CID with 35 % RCE on the $[M+2H]^{2+}$ ion (*m/z* 1170.02) of leepeptin, Cyclo(LYGVRNDE)-EINWHFDYWT, showed results previously described with the tail being sequentially fragmented (30). Comparable fragmentation using the different techniques was observed for leepeptin as was observed for huascopeptin (Figure S3).

The same HCD conditions for huascopeptin were used and exhibited the optimal fragmentation of the macrolactam ring. The macrolactam ring (LYGVRNDE) fragments like other cyclic peptides, opening randomly at each peptide bond, causing eight linear peptides to form (Scheme S1). Previous CID fragmentation results showed the cleavage of Glu9 from the tail region and the intact macrocycle (*m/z* 946.46), leading to the conclusion that Leu1 and Glu8 form the macrocycle, which was backed further by 2D NMR results in a previous study (30) and our HCD results (Figure S4 and Figure S5)

Overall, HCD fragmentation exhibited enough energy to open the macrocycle of leepeptin while also leaving sufficient structural information to observe fragmentation of the tail (Figure S4). The macrocycle of leepeptin randomly opens at each amino acid linkage forming 8 linear peptides (Scheme S1). Of these 8 linear peptides, 56 total *b* ions are

possible. All b1 ions were not detected due to the mass range cut off at *m/z* 150, therefore a total of 48 *b* ions were theoretically predicted (Scheme S1). Of these, 25 were experimentally observed in the spectrum. For linear peptides 7 and 8, no *b* ions were detected but *a* and *y* ions were present in the spectrum (Figure S5), which in some cases also prove to be diagnostic. However, this approach was complicated by separate fragmentation pathways yielding identical nominal masses for fragments. For instance, y2ED (*m/z* 230.0771) overlaps with b2NR (*m/z* 230.0771). Upon closer inspection, these two ions correspond for the same amino acid cleavage of Asn6-Asp7, therefore their overlap does not hinder spectral interpretation. A full table of observed fragment ions is available in Table S2.

Complete amino acid sequence elucidation of citrulassin C

At the time of running these pure compounds, we knew that the genome of *Streptomyces huasconensis* HST28^T contained a gene cluster and core peptide for another lasso peptide, citrulassin C. In order to determine the capability of the method for determining lasso peptide amino acid sequences *de novo*, we attempted to elucidate the full amino acid sequence of this previously detected lasso peptide by Tietz et al (15).

When the SPE fraction of the *Streptomyces huasconensis* HST28^T crude extract was analysed by LCMS (Figure S6) we observed an $[M+2H]^{2+}$ ion at *m/z* 855.4866, suggesting the presence of citrulassin C – Cyclo(LLQRSGND)-Cit-LILSKN – , which was previously reported from the culture medium of *Streptomyces albulus* NRRL S-146 but never isolated for NMR analysis (15). We applied the LCMS protocol, with HCD fragmentation, as described above to confirm the sequence of amino acids directly to an SPE fraction of the crude extract that still displayed an ion count deemed sufficient for the ion of interest. HCD using 35-40% normalized collision energy was optimal for opening the macrocycle (Figure 3). As expected, lower ion counts were observed in the semi-crude extract than for the pure compounds, therefore, ions for the varying linear peptides generated during macrocycle opening were not observed. Nevertheless, sequence confirmation of citrulassin C was observed through the fragmentation of the putative genome-mined amino acid sequence.

The b_6 and its complimentary fragment ion, y_9, were the only gap in the sequence but due to the unique mass loss of glycine, no other amino acid fits into this position. Based on the rules for lasso peptide macrocycles and the observed fragment ions of the linear peptide corresponding to the anticipated sequence (LLQRSGND), we show that

citrulassin C is the first lasso peptide to have its full amino acid sequence confirmed using only mass spectrometry. All observed fragments for citrulassin C are present in Table S3.

DISCUSSION

HCD fragmentation provided the only MS^2 information that makes full sequence elucidation possible for huascopeptin and leepeptin. The results clearly demonstrate that HCD not only fragments the tail sequentially from the *C*-terminus but also is able to open the macrolactam ring and induce diagnostic fragmentation. In the case of citrulassin C, sequence elucidation was possible but confirmation of the resultant 8 linear peptides from fragmentation was not due to low quantity of this lasso peptide in semi-crude extract. Prior to this study, chemical cleavage or proteolytic digestion followed by MS^n has been used to open only two lasso peptide macrocycles as the overwhelming majority of lasso peptides are resistant to both forms of degradation (1). Thermolysin digestion was used to open the macrocycle of streptomonomicin and using MALDI-TOF, the linear peptide sequence was partially elucidated (31). Chemical cleavage was used to open the benenodin-1 ΔC5 macrocycle as well as other cyclic peptides and rotaxanes with subsequent MS/MS to determine the amino acid sequences (32), but this also required an additional step that MS^2 HCD fragmentation does not require. Purification of the compounds also has been made an optional step in the process, although the results clearly show that purified compounds allow for increased observation of fragment ions.

The normal complexity of random cyclic peptide fragmentation causes increased but manageable difficulty in spectral deconvolution as seen in Figures 1, 3 and S4. Similar to other RiPPs, lasso peptide biosynthetic gene clusters are well characterised and follow the standard motif of a leader and core peptide [11], therefore, putative structures of lasso peptides can predicted based on certain rules on ring size, isopeptide bond formation, and overall peptide size. However, bioinformatic prediction does not equate to confirmation, as biosynthetic gene clusters can be 'silent' or 'cryptic' and as a consequence, potentially not expressed.

The two fully characterized lasso peptides represent two structurally different variants of class II lasso peptides. Leepeptin (30) contains 18 amino acids including many aromatic tail residues and represents a very small percentage of lasso peptides that have a leucine at the *N*-terminus (15). Huascopeptin contains 13 residues and was the first lasso peptide with a seven-residue Gly-Asp macrocycle (29). Additionally, citrulassin C also consists of a leucine residue at

171 the *N*-terminus and the ring forming aspartic acid, with no aromatic amino acids. These three structurally different class
172 II lasso peptides provide good evidence for the method being more widely applicable to class II lasso peptides. The
173 study focused on class II lasso peptides only as they are the most well described, their chemical space is exponentially
174 greater compared to other classes, and they represent the most biological relevant class of lasso peptides at the
175 moment.

176 HCD provided no topological information as far as we are able to interpret the spectra. No interlocked and/or
177 hydrogen migration fragments were observed, nor would we expect them. These typically only occur in 'softer'
178 fragmentation techniques like ETD or ECD but are not observable in HCD (33). All three HCD spectra include extra ions
179 associated with internal fragmentation, common in HCD spectra, that were uninformative for the scope of this work
180 (33). We see this advancement as a compliment to the CID/ETD technique previously described (27). Each technique
181 fills the gaps left by the other (i.e., HCD opens the macrocycle where CID/ETD does not and CID/ETD allows for lasso
182 topology of the lower plug to be determined, where HCD cannot). Overall, different fragmentation techniques and
183 mass analysers are continually being explored for their potential use for characterising lasso peptides.

184 Additionally, other MS/MS studies involving lasso peptides evaluate the branched-cyclic topoisomers but
185 generating these were outside of the general scope of this project. In the aim to evaluate these fragmentation
186 techniques as a means to open the macrocycles to confirm genomic predictions at the earliest stages of fermentation,
187 the lasso topology is overwhelmingly the most relevant confirmation to study. Without the aid of mutasynthesis and
188 alterations to biosynthetic gene clusters, only one lasso peptide, lassomycin, has ever been found with a branched-
189 cyclic topology outside of unthreading at elevated temperature (1). Studies have shown that when lasso peptides lose
190 their lasso topology, they also lose their biological activity and their thermal stability, making it biologically extraneous
191 to produce the branched-cyclic topology (1). This is reinforced by the fact that lasso peptide biosynthetic gene clusters
192 that do not contain ABC transporters, contain isopeptidase enzymes responsible for breaking the macrocycle, removing
193 the lasso topology and thus removing the biological activity (34).

194 MS^3 experiments seem to be unnecessary as we have shown MS^2 is sufficient for observation of all relevant
195 ions. Overall, mass spectrometry requires a fraction of the sample NMR requires for structural elucidation as well as no
196 special sample preparation besides desalting and the removal of compounds harmful to mass spectrometers, allowing

for detection of lasso peptides in crude or semi-crude extracts and subsequent MS^2 with HCD to fully elucidate the amino acid sequence. Additionally, early mass spectrometry-based analysis of lasso peptides and other RiPPs can inform about post-translational modifications which are typically difficult to predict with genome mining alone. A prime example of this are the citrulassins, which were initially determined by MS^2 to contain a citrulline post translational modification in the tail residues (15). The methodology described here unlocks the further potential of detecting post translational modifications within the macrocycle.

The producing strains, *Streptomyces leeuwenhoekii* $C34^T$ (35) and *Streptomyces huasconensis* (36) both originate from Atacama Desert high-altitude salt flats, Salar de Atacama and Salar de Huasco, respectively, that are seasonally covered in water. The extreme environments these isolates arise from have been shown to contain a rich biodiversity in actinomycetes (35,37-39) as well as rich in secondary metabolites (8,37,38,40). The interface of the extremes from the desert biosphere (i.e., high UV radiation, high salinity gradients, high altitude) coupled to the marine environment, makes these bacteria a potentially interesting starting point for continuing drug discovery efforts.

We have shown in this study, using the examples of leepeptin and huascopeptin, that full amino acid sequence elucidation of lasso peptides is possible using only MS^2 without any prior sample preparation or proteolytic digestion. Additionally, confirmation of predicted amino acid sequences at the earliest stages of bacterial fermentation was shown using citrulassin C, adding to the growing use of mass spectrometry in lasso peptide studies. Using MS^2 HCD fragmentation removes the need for NMR techniques for amino acid sequence confirmation of these RiPPs and increases the scope of mass spectrometric based analysis. With further development, HCD and CID/ETD could have the potential to fully elucidate lasso peptide sequences and lasso topology, respectively, rendering NMR as the complimentary technique and mass spectrometry the main technique for lasso peptide identification and characterisation. Additionally, lasso peptides represent a unique group of peptide antimicrobial compounds that show promise as potent and narrow spectrum antibiotics that are amenable to engineering as therapeutics. The ability to accurately predict putative structures using genome mining, the detection of lasso peptides in the crude extract stage and to subsequently be able to confirm their complete amino acid sequence makes this a major tool in the continued biodiscovery of lasso peptides.

EXPERIMENTAL

Leepeptin and huascopeptin were isolated in previous studies (29,30). A SPE fraction of the crude extract containing citrulassin C was obtained from large scale fermentation of *Streptomyces huasconensis* HST28^T using the same conditions described for the isolation of huascopeptin. Nomenclature for the linear peptides formed via random ring opening during HCD follows the work of Ngoka and Gross (18). Use of this nomenclature system is shown in Figures 1 and S1. Interpretation of lasso peptide fragmentation was aided by the use of ProteinProspector MS-Product (http://prospector.ucsf.edu). Linear cyclic peptides were input into this online fragmentation prediction software in order to determine all of the possible *a*, *b*, and *y* ions resulting from the opening of each lasso peptide macrocycle.

Sample preparation and MS^n acquisition

Both leepeptin and huascopeptin (100 μg each) were dissolved in 0.1% TFA to a final concentration of about 280 pmol/μL. To improve solvation samples were vortexed, ultra-sonicated and stored overnight in a refrigerator (4 °C). Subsequently, each sample was centrifuged, and an aliquot of the supernatant was diluted with a solution containing 30% acetonitrile acidified with 0.1% formic acid. Finally, leepeptin and huascopeptin were measured by direct-infusion experiments on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific) using a flex-source coupled with a 25 μL syringe and spray voltage was set to 1.8 kV. Four different fragmentation methods were attempted (CID, EThcD, HCD, HCD/ETD using 100 and 200 ms activation time for ETD). The precursor masses of *m/z* 705.8 (huascopeptin) and *m/z* 1170.5 (leepeptin) were selected for all MS^2 experiments. The resolution of the Orbitrap was fixed to 60,000 using an isolation width of 2 Da for all MS^n experiments. The relative collision energy for HCD was manually varied between 15 to 25% and up to 35% for all MS^3 experiments.

The SPE fraction (2.5 mg) containing citrulassin C was dissolved in 1 mL of 10% acetonitrile acidified with 0.1% TFA After centrifugation, the supernatant was diluted 1/200 in 0.1% formic acid to give a final concentration of about 12.5 ng/mL. 15 μL was injected into a nano-HPLC system (Ultimate 3000, Thermo Scientific) using a precolumn (Acclaim Pepmap, 100 μm × 2 cm, Thermo Scientific) for pre-concentration and an Acclaim Pepmap 100 C18 column (3 μm, 75 μm × 50 cm) for separation. The gradient used was: 30 min from 3% to 50% of eluent B (eluent A: aqueous 0.1% formic

acid, eluent B: 84% acetonitrile/ aqueous 0.1% formic acid). The nano LC system was coupled to an LTQ Orbitrap Velos Pro mass spectrometer (Thermo Scientific) for detection. HCD was used for fragmentation using a Top 5 method. Different normalized collision energy settings were used in each run (CE 25/30/35/40). With respect to the long cycle time of this instrument, the mass resolution for MS1 was set to 30,000 and to 7,500 for MS^2. The spray voltage applied to the PicoTip emitter (New Objective) was set to 1.5 kV.

ACKNOWLEDGEMENTS

SAJ would like to thank the University of Aberdeen for an Elphinstone Scholarship. CCA thanks CONICYT PFCHA/DOCTORADO BECAS CHILE/2016 (#21160585) fellowship and CONICYT Basal Centre Grant for the Centre for Biotechnology and Bioengineering, CeBiB (FB0001). JFC also thanks CONICYT for a National PhD Scholarship (#21110356) and a Visiting Student Scholarship.

REFERENCES

(1) Hegemann JD, Zimmermann M, Xie X, Marahiel MA. Lasso peptides: an intriguing class of bacterial natural products. Acc Chem Res. 2015;48(7):1909-1919.

(2) Knappe TA, Linne U, Xie X, Marahiel MA. The glucagon receptor antagonist BI-32169 constitutes a new class of lasso peptides. FEBS Lett. 2010;584(4):785-789.

(3) Iwatsuki M, Tomoda H, Uchida R, Gouda H, Hirono S, Omura S. Lariatins, antimycobacterial peptides produced by Rhodococcus sp. K01-B0171, have a lasso structure. J Am Chem Soc. 2006;128(23):7486-7491.

(4) Detlefsen DJ, Hill SE, Volk KJ, Klohr SE, Tsunakawa M, Furumai T, et al. Siamycins I and II, new anti-HIV-1 peptides: II. Sequence analysis and structure determination of siamycin I. J Antibiot. (Tokyo) 1995;48(12):1515-1517.

(5) Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, et al. Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. Chem Biol. 2014;21(4):509-518.

(6) Metelev M, Tietz JI, Melby JO, Blair PM, Zhu L, Livnat I, et al. Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso Peptide from an understudied halophilic actinomycete. Chem Biol. 2015;22(2):241-250.

(7) Um S, Kim YJ, Kwon H, Wen H, Kim SH, Kwon HC, et al. Sungsanpin, a lasso peptide from a deep-sea streptomycete. J Nat Prod. 2013;76(5):873-879.

(8) Elsayed SS, Trusch F, Deng H, Raab A, Prokes I, Busarakam K, et al. Chaxapeptin, a Lasso Peptide from Extremotolerant Streptomyces leeuwenhoekii Strain C58 from the Hyperarid Atacama Desert. J Org Chem. 2015;80(20):10252-10260.

(9) Weber W, Fischli W, Hochuli E, Kupfer E, Weibel EK. Anantin--a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. J Antibiot. (Tokyo) 1991;44(2):164-171.

(10) Potterat O, Wagner K, Gemmecker G, Mack J, Puder C, Vettermann R, et al. BI-32169, a bicyclic 19-peptide with strong glucagon receptor antagonist activity from Streptomyces sp. J Nat Prod. 2004;67(9):1528-1531.

(11) Cheung-Lee WL, Parry ME, Jaramillo Cartagena A, Darst SA, Link AJ. Discovery and structure of the antimicrobial lasso peptide citrocin. J Biol Chem. 2019; 294(17):6822-6830.

(12) Kaweewan I, Hemmi H, Komaki H, Harada S, Kodani S. Isolation and structure determination of a new lasso peptide specialicin based on genome mining. Bioorganic & Medicinal Chemistry. 2018;26(23):6050-6055.

(13) Santos-Aberturas J, Chandra G, Frattaruolo L, Lacret R, Pham TH, Vior NM, et al. Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. Nucleic Acids Res. 2019;47(9):4624-4637.

(14) Hudson G, Burkhart B, DiCaprio A, Schwalen C, Kille B, Pogorelov T, et al. Bioinformatic Mapping of Radical S-Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. J Am Chem Soc. 2019;141(20):8228-8238.

(15) Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, et al. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. Nat Chem Biol. 2017;13(5):470-478.

(16) Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 2019;47(W1):W81-W87.

(17) Paizs B, Suhai S. Fragmentation pathways of protonated peptides. Mass Spectrom Rev. 2005;24(4):508-548.

(18) Ngoka LC, Gross ML. Multistep tandem mass spectrometry for sequencing cyclic peptides in an ion-trap mass spectrometer. J Am Soc Mass Spectrom. 1999;10(8):732-746.

(19) Medzihradszky KF, Chalkley RJ. Lessons in de novo peptide sequencing by tandem mass spectrometry. Mass Spec Rev. 2013;34(1):43-63.

(20) Mohimani H, Gurevich A, Mikheenko A, Garg N, Nothias L, Ninomiya A, et al. Dereplication of peptidic natural products through database search of mass spectra. Nat Chem Biol. 2017; 13(1):30-37.

(21) Mohimani H, Kersten RD, Liu WT, Wang M, Purvine SO, Wu S, et al. Automated genome mining of ribosomal peptide natural products. ACS Chem Biol. 2014;9(7):1545-1551.

(22) Mohimani H, Yang YL, Liu WT, Hsieh PW, Dorrestein PC, Pevzner PA. Sequencing cyclic peptides by multistage mass spectrometry. Proteomics. 2011;11(18):3642-3650.

(23) Niedermeyer, Timo H J, Strohalm M. mMass as a Software Tool for the Annotation of Cyclic Peptide Tandem Mass Spectra. PLoS ONE 7(9): e44913.

(24) Xie X, Marahiel MA. NMR as an effective tool for the structure determination of lasso peptides. Chem biochem. 2012;13(5):621-625.

(25) Rosengren KJ, Clark RJ, Daly NL, Göransson U, Jones A, Craik DJ. Microcin J25 Has a Threaded Sidechain-to-Backbone Ring Structure and Not a Head-to-Tail Cyclized Backbone. J Am Chem Soc. 2003;125(41):12464-12474.

(26) Wilson K, Kalkum M, Ottesen J, Yuzenkova J, Chait BT, Landick R, et al. Structure of Microcin J25, a Peptide Inhibitor of Bacterial RNA Polymerase, is a Lassoed Tail. J Am Chem Soc. 2003;125(41):12475-12483.

(27) Jeanne Dit Fouque K, Lavanant H, Zirah S, Hegemann JD, Fage CD, Marahiel MA, et al. General rules of fragmentation evidencing lasso structures in CID and ETD. Analyst. 2018;143(5):1157-1170.

(28) Dit Fouque KJ, Moreno J, Hegemann JD, Zirah S, Rebuffat S, Fernandez-Lima F. Identification of Lasso Peptide Topologies Using Native Nanoelectrospray Ionization-Trapped Ion Mobility Spectrometry - Mass Spectrometry. Anal Chem. 2018;90(8):5139-5146.

(29) Cortés-Albayay C, Jarmusch S, Trusch F, Ebel R, Andrews B, Jaspars M, et al. Downsizing Class II Lasso Peptides: Genome Mining-Guided Isolation of Huascopeptin Containing the First Gly1-Asp7 Macrocycle. J. Org. Chem. 2020,85(3):1661-1667

(30) Gomez-Escribano JP, Castro JF, Razmilic V, Jarmusch SA, Saalbach G, Ebel R, Jaspars M, Andrews B, Asenjo JA, Bibb MJ. Heterologous expression of a cryptic gene cluster from Streptomyces leeuwenhoekii C34T yields a novel lasso peptide, leepeptin. Appl Environ Microbiol. 2019;85:e01752-19

(31) Metelev M, Tietz JI, Melby JO, Blair PM, Zhu L, Livnat I, et al. Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso Peptide from an understudied halophilic actinomycete. Chem Biol. 2015;22(2):241-250.

(32) Elashal HE, Cohen RD, Elashal HE, Zong C, Link AJ, Raj M. Cyclic and Lasso Peptides: Sequence Determination, Topology Analysis, and Rotaxane Formation. Angew Chem Int Ed Engl. 2018;57(21):6150-6154.

(33) Brodbelt, JS. Ion Activation Methods for Peptides and Proteins. Anal. Chem. 2016, 88, 1, 30-51.

(34) Chekan JR, Koos JD, Zong C, Maksimov MO, Link AJ, Nair SK. Structure of the Lasso Peptide Isopeptidase Identifies a Topology for Processing Threaded Substrates. J Am Chem Soc. 2016;138(50):16452-16458.

(35) Okoro CK, Brown R, Jones AL, Andrews BA, Asenjo JA, Goodfellow M, et al. Diversity of culturable actinomycetes in hyper-arid soils of the Atacama Desert, Chile. Antonie Van Leeuwenhoek. 2009;95(2):121-133.

(36) Cortés-Albayay C, Dorador C, Schumann P, Andrews B, Asenjo J, Nouioui I. Streptomyces huasconensis sp. nov., an haloalkalitolerant actinobacterium isolated from a high altitude saline wetland at the Chilean Altiplano. International Journal of Systematic and Evolutionary Microbiology. 2019;69(8):2315-2322.

(37) Rateb ME, Houssen WE, Harrison WT, Deng H, Okoro CK, Asenjo JA, et al. Diverse metabolic profiles of a Streptomyces strain isolated from a hyper-arid environment. J Nat Prod. 2011;74(9):1965-1971.

(38) Goodfellow M, Nouioui I, Sanderson R, Xie F, Bull AT. Rare taxa and dark microbial matter: novel bioactive actinobacteria abound in Atacama Desert soils. Antonie Van Leeuwenhoek. 2018;111(8):1315-1332.

(39) Manfio GP, Atalan E, Zakrzewska-Czerwinska J, Mordarski M, Rodriguez C, Collins MD, et al. Classification of novel soil streptomycetes as Streptomyces aureus sp. nov., Streptomyces laceyi sp. nov. and Streptomyces sanglieri sp. nov. Antonie Van Leeuwenhoek. 2003;83(3):245-255.

(40) Rateb ME, Houssen WE, Arnold M, Abdelrahman MH, Deng H, Harrison WT, et al. Chaxamycins A-D, bioactive ansamycins from a hyper-arid desert Streptomyces sp. J Nat Prod. 2011;74(6):1491-1499.

FIGURE LEGENDS

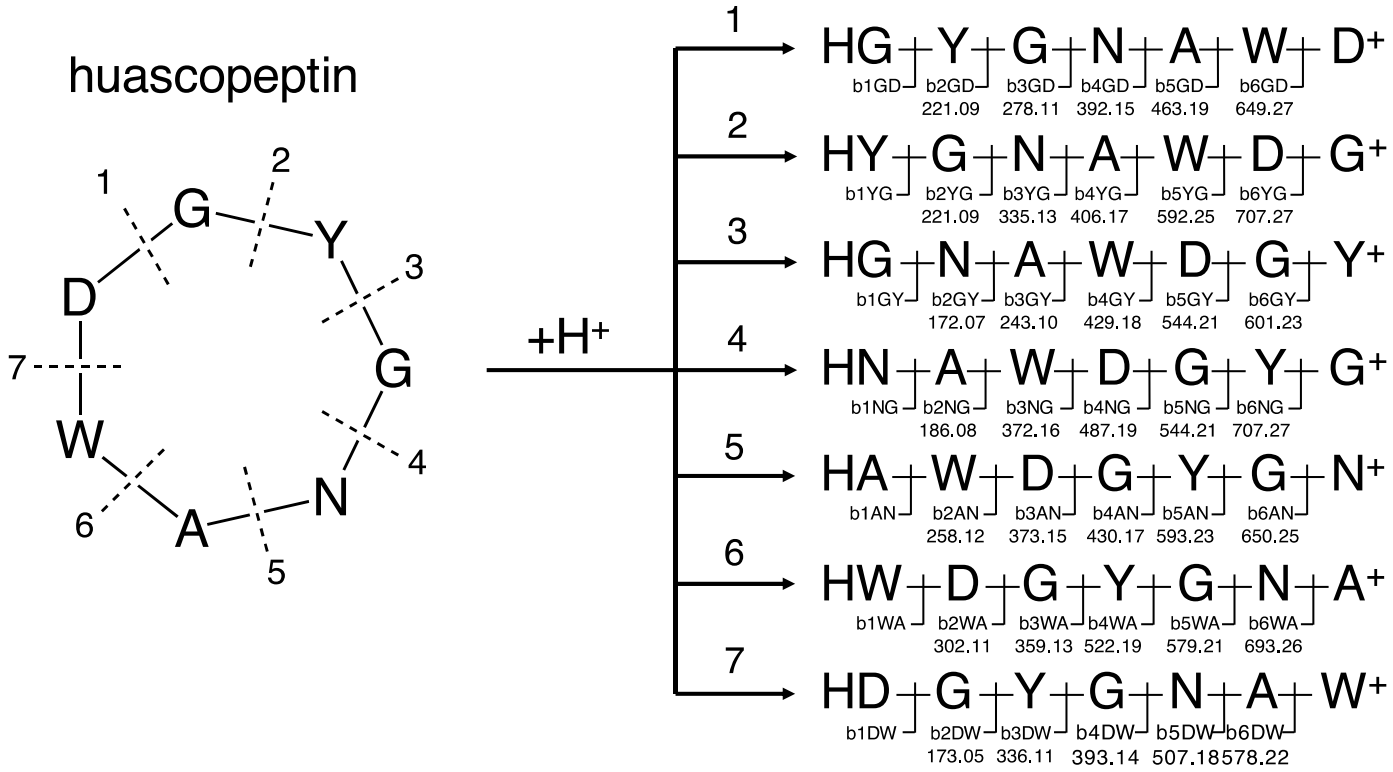
Scheme 1. Sequence of *b* ions from the fragmentation of the macrolactam ring of huascopeptin. According to conventions of cyclic peptide fragmentation, random ring opening events occur yielding 7 linear peptides. C terminus cleavages occur sequentially, yielding 35 *b* ions. All b1 ions are below the *m/z* threshold of 150 and therefore are not considered for this study.

Figure 1. (Top) MS^2 spectrum of huascopeptin fragmented with HCD at 25% RCE. (Bottom) MS^2 spectrum of huascopeptin macrolactam ring linear peptide 1 (see Scheme S1). Presence of *a* and *b* ions is indicated in the spectrum. Additionally, linear peptide spectra can be found in Figures S2.

Figure 2. MS^2 spectrum of huascopeptin: Evidence for branched or "T-shaped" fragments. Opening first occurs at a peptide bond within a macrocycle followed by subsequent cleavage of macrocycle residues.

Figure 3. MS^2 spectrum of citrulassin C fragmented with HCD at 35% NCE. * indicates the ion was only observed at 40% NCE.

368 **Scheme 1.**



369

Figure 1.

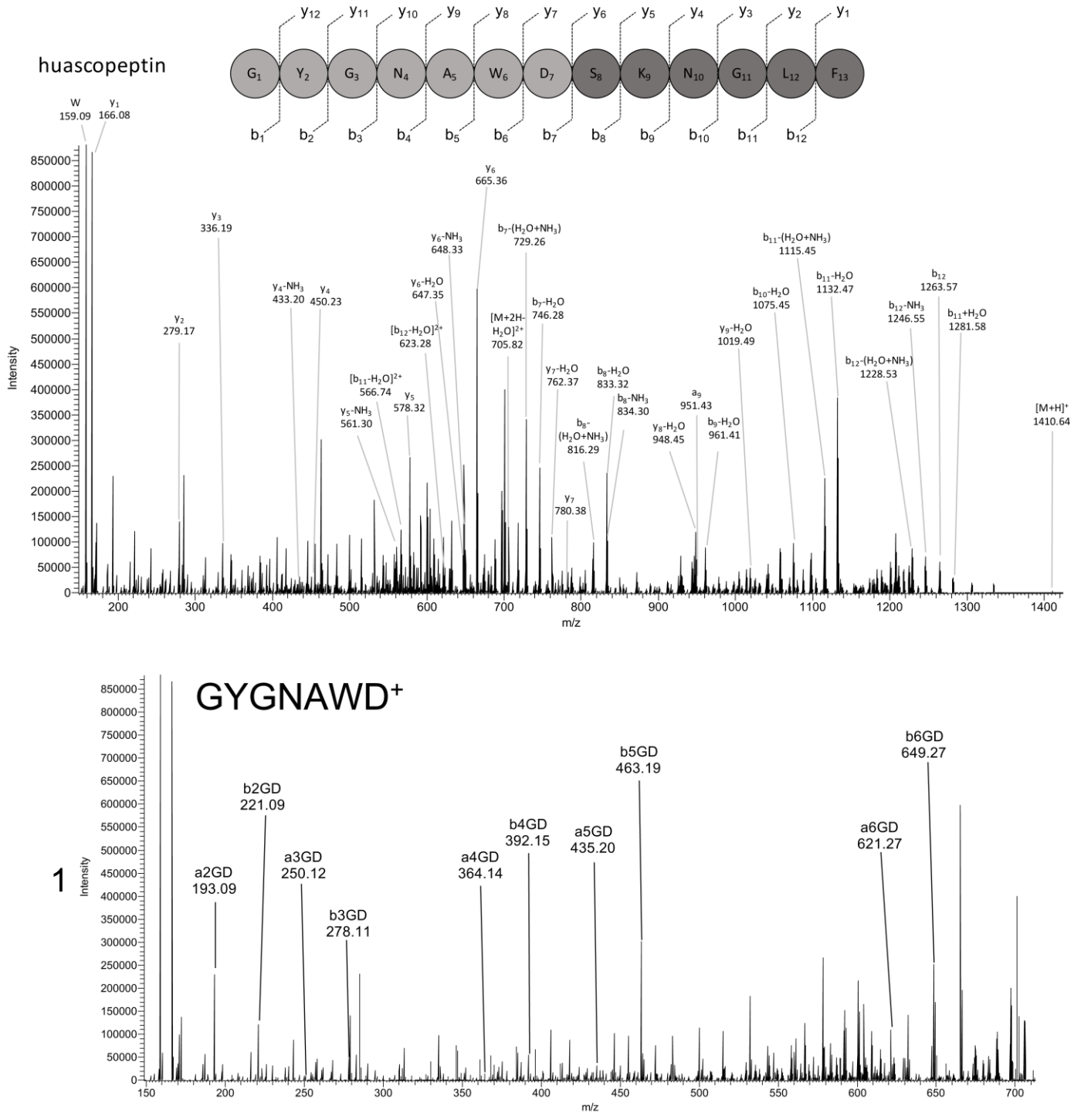

huascopeptin
G1 Y2 G3 N4 A5 W6 D7 S8 K9 N10 G11 L12 F13
Intensity
m/z
GYGNAWD+
1
a2GD 193.09
b2GD 221.09
a3GD 250.12
b3GD 278.11
a4GD 364.14
b4GD 392.15
a5GD 435.20
b5GD 463.19
a6GD 621.27
b6GD 649.27

Figure 2.

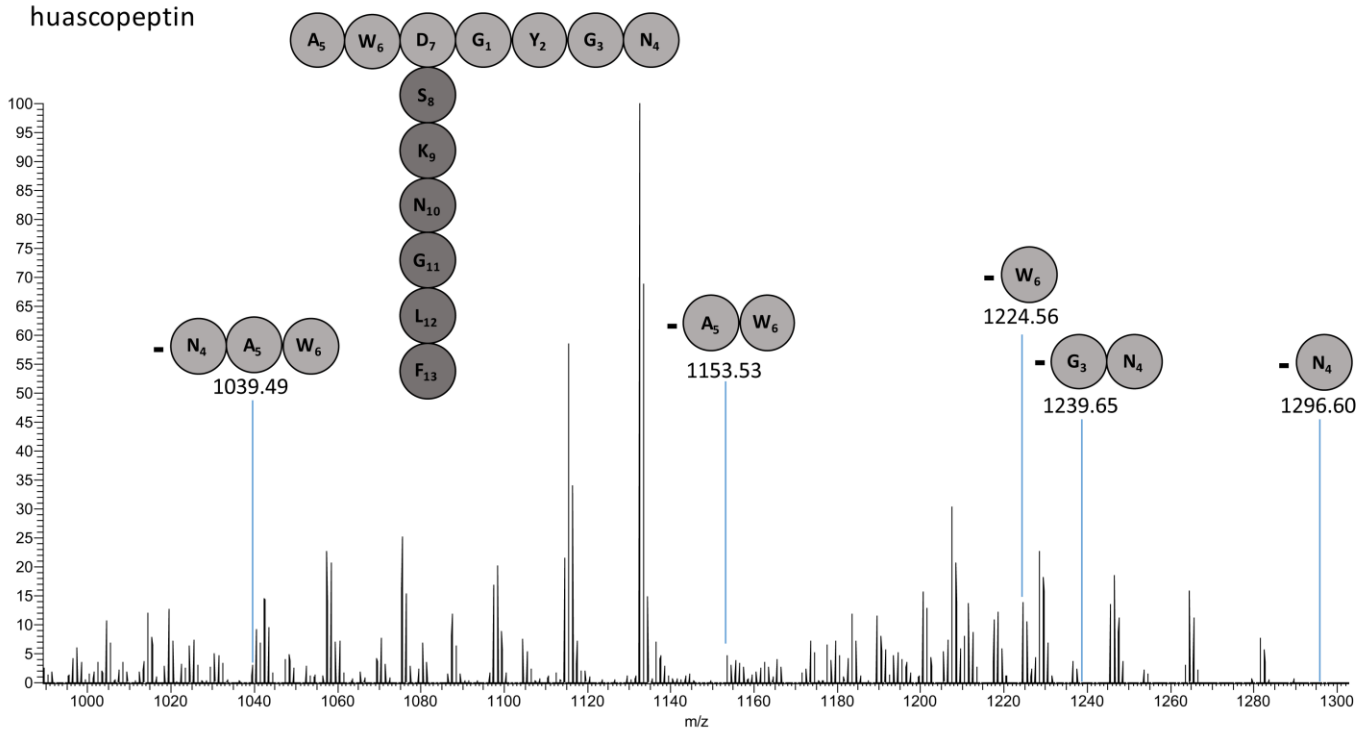
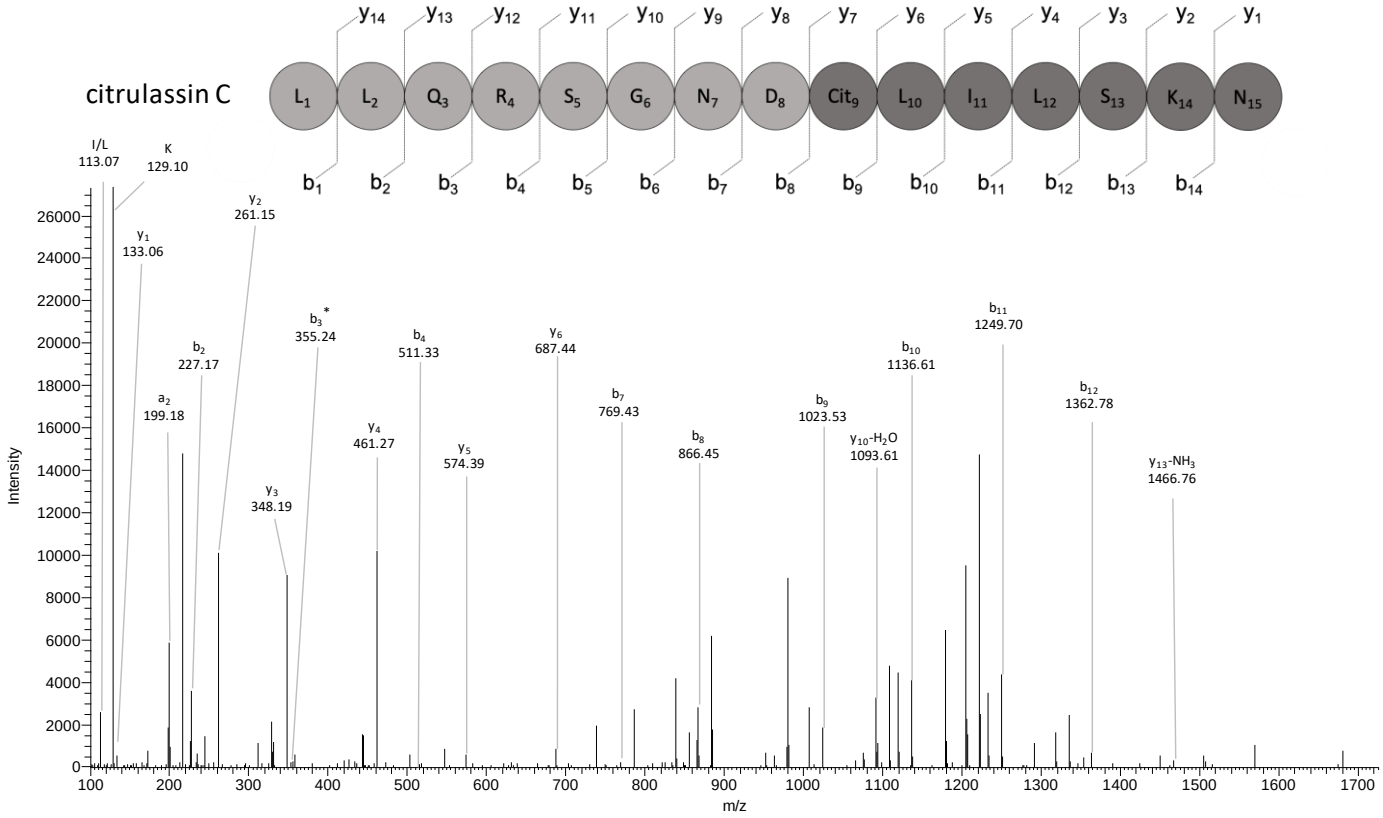

huascopeptin
A5 W6 D7 G1 Y2 G3 N4
S8
K9
N10
G11
L12
F13
- N4 A5 W6
1039.49
- A5 W6
1153.53
- W6
1224.56
- G3 N4
1239.65
- N4
1296.60
m/z

Figure 3.


citrulassin C
L1 L2 Q3 R4 S5 G6 N7 D8 Cit9 L10 I11 L12 S13 K14 N15
y14 y13 y12 y11 y10 y9 y8 y7 y6 y5 y4 y3 y2 y1
b1 b2 b3 b4 b5 b6 b7 b8 b9 b10 b11 b12 b13 b14
I/L 113.07
K 129.10
y2 261.15
y1 133.06
b3* 355.24
b2 227.17
b4 511.33
y6 687.44
b11 1249.70
b10 1136.61
a2 199.18
b7 769.43
b12 1362.78
b9 1023.53
y4 461.27
y10-H2O 1093.61
y5 574.39
b8 866.45
y13-NH3 1466.76
y3 348.19
Intensity
m/z