A Mass Spectrometry Imaging-based Platform: Towards Discovery in Metabolomics, Neuropeptidomics and Proteomics

By
Hui Ye

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Pharmaceutical Sciences)

at the
UNIVERSITY OF WISCONSIN-MADISON
2013

Date of final oral examination: 05/06/13

The dissertation is approved by the following members of the Final Oral Committee:
Lingjun Li, Professor, School of Pharmacy and Department of Chemistry
Jean-Michel Ané, Associate Professor, Department of Agronomy
Richard E. Peterson, Professor, School of Pharmacy
Antony O.W. Stretton, Professor, Department of Zoology
Tim Bugni, Assistant Professor, School of Pharmacy

Acknowledgement

I would like to acknowledge my advisor Dr. Lingjun Li-- none of the work described in this thesis could be accomplished without her. I first want to express my gratitude to her for allowing me to join her lab. Dr. Li has led our lab to become a world-class bioanalytical and neurochemistry research group with development of numerous creative and practical analytical strategies that bridge the neuroscience and analytical chemistry fields. As an undergraduate who only had minimal exposure and training in mass spectrometry, I was introduced to the wonderful world of biological mass spectrometry and its cutting edge applications in neuroscience. With Dr. Li's encouragement and guidance I have become an independent scientist in the field of bioanalytical neuroscience research that I discover my passion and aspiration. Dr. Li has always been supportive and patient, who provides me enormous tenacity to persevere and overcome the challenges encountered in my research. I also appreciate the freedom Dr. Li gave me to navigate the projects and delve into the directions that I find interesting as the projects proceed, which I would certainly cherish as a future principal investigator in academia upon graduation. I am also thankful for her emotional support and motivation she provided at certain points that I felt challenging to carry on during my Ph.D. career. It has been an extremely productive and pleasant experience of transformation from a student to a scientific researcher under Dr. Li's guidance and I will forever cherish and benefit from the opportunity of being part of the Li group.

I also would like to thank my committee members, Dr. Jean-Michel Ané, Dr. Richard Peterson, Dr. Antony Stretton, Dr. Tim Bugni, for providing valuable insights

and suggestions at every critical stage of my Ph.D. process. I wish to thank Dr. Jean-Michel Ané, my first collaborator, for constructive criticism and assistance in experimental design, data analysis and biological interpretation of the metabolite imaging platform we developed to investigate the plant-microbe interactions. I am very grateful for him to have faith in the approach we took and in me to answer the biological questions we aim to address. Dr. Muthusubramanian Venkateshwaran and my labmate Erin Gemperline have also contributed tremendously to the project in every aspect they could.

In addition to my committee members, I would also like to thank my other collaborators and colleagues who have contributed to my work. I am first indebted to Dr. Chrysanthy Ikonomidou and Rakesh Mandal, who have given me the chance to apply the analytical platform I developed to a brand new biological model and gain novel information regarding neurodevelopmental disorders. I then want to extend my gratefulness to my other collaborators, Dr. Katherine Kellersberger from Bruker and Dr. Kerstin Strupat from Thermo Fisher Scientific, who provided initial instrument access and helpful input in the applications of high resolution mass spectrometry imaging in the applications of neuropeptides and metabolites. Dr. Paul Martin Thomas, Adam Catherman, Dr. Neil Kelleher and Dr. Deborah Baro aided in the work of discovering novel neuropeptides from the spiny lobster *Panulirus interruptus*. Dr. Robert Thorne, Mohan Gautam, Dr. James Dowell, Quentin Bremer and Dr. Brian Baldo provided me the opportunity to work on mammalian neuropeptidomics in two other projects not included in the thesis.

Furthermore, I would like to thank my former labmates Dr. Ruibing Chen, Dr. Limei Hui, Dr. Feng Xiang, Dr. Junhua Wang and Dr. Xiaoyue Jiang for their guidance when I initially started my graduate research. They first led me to appreciate the wonderfulness of the research our lab is doing. I also want to thank my labmates Jingxin Wang and Bingming Chen, who have assisted in gathering quality data for our joint projects. Tyler Greer, Dr. Claire Schmerberg, Chris Lietz, Chenxi Jia and Zichuan Zhang have kindly helped with projects related to this work. In addition, I have benefited from inspiring discussions and day-to-day interactions with Dr. Xuefei Zhong, Dr. Yan Liu, Dr. Di Ma, Chenxi Yang, Chuanzi Ouyang and Nicole Woodards. I would also like to thank Dr. Cameron Scarlett, Molly Hahn from the Analytical Instrumentation Center of the School of Pharmacy and Dr. Amy Harms and Dr. Grezegorz Sabat from the Biotechnology Center for maintaining and troubleshooting instruments used in this work.

Last but not least, I would like to thank my parents Zhen Gao and Yongjun Ye, Dr. Xiandeng Hou from Sichuan University, Dr. Shiming Ye and Dr. Qin Lu for their support of me pursuing Ph.D. degree in the U.S.. I also feel incredibly lucky to have enormous support and good distractions provided from friends and other family members, Zengyi Jiang, Li Gu, Yan Xiong, Miaosi Li from old times and Shiyu Chen, Zhidan Liang, Na Liu among others in Madison. I would also like to dedicate this work to my grandmother Xiuying Li, who unfortunately passed away last month and could not see me complete this work. She and my grandfather Taiming Gao would be happy to know my achievements in heaven.

Table of Contents

A Mass Spectrometry Imaging-based Platform: Towards Discovery in Metabolomics, Neuropeptidomics and Proteomics

Hui Ye

Under the supervision of Professor Lingjun Li

At the University of Wisconsin-Madison

Abstract

Mass spectrometry imaging (MSI) has evolved as a promising technology to map a broad range of molecules ranging from metabolites, peptides to proteins in an anatomical context. In this work, a multi-dimensional MSI-based platform was developed and applied to investigate several intriguing biological problems.

This dissertation first introduces the rationale and recent progress of MSI, reviews its emerging applications to various fields and outlines one limitation that the detection of low molecular weight (MW) species is usually hindered by matrix-derived peaks. An ionless matrix referred as the proton sponge is therefore investigated as a solution to eliminate the interferences. It allows for probing metabolites in roots and root nodules of leguminous plants *Medicago truncatula* that possess symbiotic interactions with nitrogen-fixing rhizobia, highlighting differences in metabolic distribution between roots and nodules. In addition to exploring novel matrices, high resolution and high accuracy MSI (HRMSI) is also developed to circumvent matrix interferences by resolving endogenous analytes from background peaks. Novel spatial information of multiple metabolites and neurotransmitters is gleaned using rodent and crustacean brain specimens in this proof-of-

principle study. Beyond metabolites, the HRMSI-based approach is applied to map neuropeptides in crustacean stomatogastric ganglia. Complementary to MALDI-MS, electrospray ionization (ESI)-MS-based experiments deliver another suite of neuropeptides. We therefore combine multiple MS instruments and compared their performances for characterization of novel neuropeptides from a marine model organism, the spiny lobster *Panulirus interruptus*. In addition to peptides, MALDI-MSI has also been applied to mapping protein distributions in the brain under pharmacological treatment. Specifically, alterations of protein expression patterns in the brains of infant rats were studied following neonatal exposure to the N-methyl-D-aspartate receptor antagonist dizocilpine (MK 801). A strategy combining profiling and imaging MS and top-down high-resolution tandem MS is utilized to discover putative protein biomarkers that may serve as potential targets of treatments for neurodevelopmental disorders.

This work not only improves upon MSI by developing new methods, but also presents a useful platform that integrates MALDI-MSI with ESI-MS in exploring the underlying chemistry of various biological models exemplified by rodent, crustacean and plant-bacteria systems, and demonstrates wide applicability to future research.

Chapter 1

Introduction and Research Summary

Introduction and Research Summary

1.1 Introduction

This work focuses on the development of cutting-edge analytical tools that improves upon current mass spectrometry imaging (MSI) technique, and the construction of an MSI-based platform that combines matrix-assisted laser desorption/ionization (MALDI)-MSI with conventional liquid chromatography (LC)-ESI-MS/MS approach. This multi-dimensional MSI-based platform combines the advantages provided by both approaches, attaining spatial information of select compounds while gleaning the identity information of analytes along the process. In addition to method development, this work also demonstrates the usefulness of developed tools and platform in relevant research spanning from metabolomics detailed in Chapter 3 and 4, neuropeptidomics described in Chapter 5 and 6 to proteomics highlighted in Chapter 7. Chapter 1 serves as a general introduction of the work described in this thesis and summarizes the projects chapter by chapter. Chapter 2 provides an overview of key elements associated with MSI, and highlights the applications of MSI particularly in neuropeptide signaling and clinical research.

1.2 Mass Spectrometry Imaging: An Emerging Technology for Probing Neuropeptide Signaling and Clinical Diagnostics

In the past decade, mass spectrometry imaging (MSI) is one of the most exciting fields in analytical science that grants the ability to study a broad mass range of molecular species ranging from small molecules to large proteins by creating distribution maps of select compounds.[1, 2] To date, MSI has demonstrated its versatility in a number of

applications by accurately and simultaneously determining a number of analytes directly from tissues and cells; among those its usefulness in biomarker discovery towards clinical applications and in probing neuropeptide signaling are highlighted in Chapter 2.[3, 4] Although the underlying concept is simple, a number of factors such as sample preparation procedures, choices of mass analyzers, and software suites must all be taken into account for successful applications of MSI.[5] Therefore, this section provides a brief overview of key elements of MSI: history, rationale, sampling protocols, instrumentation, data analysis tools, yet focuses on the technical advances that are facilitating the improvement of this field and current MSI applications that provide significant insights in neuropeptide signaling and biomarker discovery.

1.3 Visualizing Neurotransmitters and Metabolites in the Central Nervous System by High Resolution and High Accuracy Mass Spectrometric Imaging

The spatial localization and molecular identity of metabolites and neurotransmitters within biological organisms is of tremendous interest to neuroscientists. In this section, a novel MALDI-MSI-based approach to map metabolites and neurotransmitters is introduced. MALDI-MSI has demonstrated its unique advantage by directly localizing the distributions of a wide range of biomolecules simultaneously from a tissue specimen compared to conventional imaging techniques like immunohistochemistry.[6] Nevertheless, its application to metabolites and neurotransmitters is hindered by numerous matrix-derived peaks as outlined in Chapter 2, high-resolution and high-accuracy mass spectrometers (HRMS) address this challenge by unambiguous differentiation of

endogenous analytes from matrix peaks. In this study, we utilized rodent and crustacean central nervous systems specimens to validate the HRMSI approach in mapping neurotransmitters and metabolites from neurobiological samples. The HRMSI results were compared with those obtained from a medium-resolution mass spectrometer (MRMS), tandem time-of-flight instrument, to demonstrate the power and unique advantages of HRMSI and reveal how this new tool would benefit molecular imaging applications in neuroscience.

1.4 MALDI Mass Spectrometry-assisted Molecular Imaging of Metabolites during Nitrogen Fixation in the *Medicago truncatula – Sinorhizobium meliloti* Symbiosis

In addition to the use of HRMSI, regular MALDI MSI employing an alternative matrix that efficiently ionizes analytes yet introduces few interfering peaks to mass spectra could also reduce matrix interference, thus enable imaging of small molecule metabolites. In this section, an ionless matrix 1,8-bis(dimethyl-amino) naphthalene (DMAN) previously reported by the Svatoš group was applied to MSI for the first time.[7] Together with the conventional matrix 2,5-dihydroxybenzoic acid (DHB), an improved detection coverage of metabolites was achieved. This novel combination of matrices was then applied to probe the symbiotic associations between leguminous plants *Medicago truncatula* and nitrogen-fixing rhizobia *Sinorhizobium meliloti* in a specialized organ called root nodules, in which the rhizobia fix atmospheric nitrogen and transfer it to the plant.[8] The information of metabolite distribution in roots and root nodules of *M. truncatula* during nitrogen fixation could shed light on the nitrogen-fixing symbiosis. As

expected, the combination of this efficient, novel MALDI matrix DMAN with DHB permitted the detection of a large array of organic acids, amino acids, sugars, lipids, flavonoids and their conjugates with great coverage. Ion density maps of representative metabolites were presented and correlated with the nitrogen fixation process. Moreover, metabolic differences were manifested between roots and nodules, and also between fixing and non-fixing nodules produced by plant and bacterial mutants due to the availability of genetic information in both symbiotic partners.[8, 9] In summary, this study showed the benefits of using MSI with the combination of DMAN and DHB for mapping metabolic differences in plant biology.

1.5 Mapping of Neuropeptides in the Crustacean Stomatogastric Nervous System by Imaging Mass Spectrometry

Beyond metabolites, MALDI-MSI has also been applied to determine the spatial distribution of neuropeptides in cells and tissues. Neuropeptides remain the largest and most diverse group of endocrine signaling molecules in the nervous system. Tremendous efforts have been devoted in the past to the studies of crustacean stomatogastric nervous system (STNS), a rich source of numerous neuropeptides and neurotransmitters with greater simplicity compared to mammalian nervous system.[10, 11] In this chapter, we have developed a multi-faceted MSI-based strategy combining profiling and imaging techniques to unequivocally characterize and map neuropeptides from the blue crab *Callinectes sapidus* STNS at the network level. Overall, 55 neuropeptides from 10 families were identified from the major ganglia in the *C. sapidus* STNS for the first time, including the stomatogastric ganglion (STG), the paired commissural ganglia (CoG), the

esophageal ganglion (OG), and the connecting nerve stomatogastric nerve (*stn*) on MALDI-TOF/TOF. The STG, as the center of the STNS, is a ganglion of minute size and contains merely 25-30 motor neurons, thus posing a significant analytical challenge to neuropeptide imaging. Previously, immunohistochemistry (IHC) has been applied to this endeavor yet with identification accuracy and throughput compromised.[12] In contrast, our work successfully documented the locations of multiple neuropeptides at a spatial resolution of 25 μm along with confident identification in the STG using a combination of MALDI-TOF/TOF and an HRMS MALDI-Fourier transform ion cyclotron resonance (FT-ICR) instrument. Different isoforms from the same family were simultaneously and unambiguously mapped, facilitating the functional exploration of neuropeptides present in the crustacean STNS and exemplifying the revolutionary role of this platform in neuronal network studies.

1.6 Defining the Neuropeptidome of the Spiny Lobster *Panulirus interruptus* Brain Using a Multi-Dimensional Mass Spectrometry-Based Platform

Distinct from Chapter 5, this project is focused on the development of a toolkit in crustacean neuropeptide research using liquid-phase analysis. A major challenge to the characterization of endogenous neuropeptides in decapod crustacean nervous systems is the limited amount of information about known preprohormones for decapods. Nevertheless, we observed great conservation among decapod species when compiling the home-built decapod neuropeptide database. We thereby characterized the neuropeptidome of a brain extract from *P. interruptus*, an important aquaculture decapod species with few known preprohormones, with an HRMS LTQ-Orbitrap instrument and

conducted a ProSight database search using a flexible algorithm that allows for sequence discrepancy from known sequence against our database. In addition to this streamlined semi-automated sequencing strategy, we employed a suite of analytical techniques including dimethylation-assisted fragmentation, manual *de novo* sequencing, direct tissue analysis and capillary electrophoresis (CE)-MALDI screening to improve the detection coverage, resulting in an overall detection of 55 neuropeptides with 34 novel ones reported first time in this study. The high discovery rate from this species demonstrated the usefulness of the neuropeptide discovery pipeline we developed in this work and highlighted the advantage of utilizing multiple mass spectrometers with complementary capabilities. The localization of brain neuropeptides were further elucidated with MALDI-MSI. Collectively, our study expands the catalog of crustacean neuropeptides present in *P. interruptus*, and more importantly, presents an approach that can be adapted to explore neuropeptidomes from other species that possess limited sequence information.

1.7 Top-down Proteomics with Mass Spectrometry Imaging: A Novel Approach towards Discovery of Biomarkers for Neurodevelopmental Disorders

A final application of MSI is presented in Chapter 7, in which the brain proteome of infant rats was characterized. In the developing mammalian brain, inhibition of N-methyl-D-aspartate receptor (NMDAR) can induce widespread neuroapoptosis, inhibit neurogenesis and cause impairment of learning and memory.[13, 14] Nevertheless, our understanding of the full spectrum of developmental events affected by early exposure to these chemical agents in the brain is still limited. Therefore, we attempted to gain insight into the impact of pharmacologically induced excitatory/inhibitory imbalance in infancy

on the brain proteome using MSI. We first analyzed the protein expression profiles of rat brains exposed to vehicle or an NMDAR antagonist MK801 expression in postnatal day 10 (P10) as a high throughput screening of putative biomarkers, and compared the MALDI MS images of rat brain proteins that revealed differential abundances. We then identified these markers filtered by MS profiling and imaging as proteins such as ubiquitin, purkinje cell protein 4 (PEP-19), cytochrome c oxidase subunits and calmodulin, by a combination of reversed-phase (RP) HPLC fractionation and top-down tandem MS platform. The findings revealed by MALDI-MS profiling and imaging indicate that a brief neonatal exposure to an NMDAR antagonist that alters excitatory/inhibitory balance in the brain has a long term effect on protein expression patterns during subsequent development, and demonstrate the utility of MALDI-MSI as a discovery tool for potential biomarkers.

1.8 Conclusions and Future Directions

The last chapter consists of conclusions and future directions of various ongoing projects. This section intends to aid in relevant research in the future by analyzing the limitation of currently attained data and proposing promising techniques and alternative strategies that could improve upon the methods currently used. In summary, this dissertation presents a well-integrated MSI-based platform that contains multiple advanced analytical tools involved in sample preparation, separation, tandem MS-based fragmentation and database search. The utility of such an integrated platform is demonstrated with wide-ranging applications including the study of metabolites, peptides and proteins using systems biology approach.

1.9 References

1. Amstalden van Hove, E. R.; Smith, D. F.; Heeren, R. M. (2010) A concise review of mass spectrometry imaging. *J Chromatogr. A* 1217, 3946-54.
2. Cornett, D. S.; Reyzer, M. L.; Chaurand, P.; Caprioli, R. M. (2007) MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat. Methods* 4, 828-33.
3. Ye, H.; Gemperline, E.; Li, L. (2013) A vision for better health: Mass spectrometry imaging for clinical diagnostics. *Clin. Chim. Acta*, doi: 10.1016/j.cca.2012.10.018.
4. Ye, H.; Greer, T.; Li, L. (2012) Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *J. Proteomics* 75, 5014-26.
5. Ye, H.; Greer, T.; Li, L. (2011) From pixel to voxel: a deeper view of biological tissue by 3D mass spectral imaging. *Bioanalysis* 3, 313-32.
6. Jones, B. E.; Beaudet, A. (1987) Distribution of acetylcholine and catecholamine neurons in the cat brainstem: a choline acetyltransferase and tyrosine hydroxylase immunohistochemical study. *J. Comp. Neurol.* 261, 15-32.
7. Shroff, R.; Svatoš, A. (2009) Proton sponge: a novel and versatile MALDI matrix for the analysis of metabolites using mass spectrometry. *Anal. Chem.* 81, 7954-59.
8. Mitra, R. M.; Long, S. R. (2004) Plant and bacterial symbiotic mutants define three transcriptionally distinct stages in the development of the *Medicago truncatula*/*Sinorhizobium meliloti* symbiosis. *Plant Physiol.* 134, 595-604.
9. Starker, C. G.; Parra-Colmenares, A. L.; Smith, L.; Mitra, R. M.; Long, S. R. (2006) Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression. *Plant Physiol.* 140, 671-80.
10. Marder, E.; Bucher, D. (2001) Central pattern generators and the control of rhythmic movements. *Curr. Biol.* 11, R986-96.
11. Nusbaum, M. P. (2002) Regulating peptidergic modulation of rhythmically active neural circuits. *Brain Behav. Evolut.* 60, 378-87.
12. Stemmler, E. A.; Peguero, B.; Bruns, E. A.; Dickinson, P. S.; Christie, A. E. (2007) Identification, physiological actions, and distribution of TPSGFLGMRamide: a novel tachykinin-related peptide from the midgut and stomatogastric nervous system of Cancer crabs. *J. Neurochem.* 101, 1351-66.
13. Ikonomidou, C. (2009) Triggers of apoptosis in the immature brain. *Brain Dev.* 31, 488-92.
14. Kaindl, A. M.; Koppelstaetter, A.; Nebrich, G.; Stuwe, J.; Sifringer, M.; Zabel, C.; Klose, J.; Ikonomidou, C. (2008) Brief alteration of NMDA or GABA(A) receptor-mediated neurotransmission has long term effects on the developing cerebral cortex. *Mol. Cell. Proteomics* 7, 2293-310.

Chapter 2

A Vision for Better Health: Mass Spectrometry Imaging for Clinical Diagnostics

Adapted from **Hui Ye**, Erin Gemperline, Lingjun Li. A vision for better health: mass spectrometry imaging for clinical diagnostics. *Clin. Chim. Acta*, 2013, doi: 10.1016/j.cca.2012.10.018.

Hui Ye, Tyler Greer, Lingjun Li. Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *J. Proteomics*, **2012**, 75, 5014-26.

Abstract

Mass spectrometry imaging (MSI) is a powerful tool that grants the ability to investigate a broad mass range of molecules from metabolites to large proteins by creating detailed distribution maps of selected compounds in an anatomical context. To date, MSI has demonstrated its versatility in clinical applications, such as biomarker diagnostics of different diseases, prognostics of disease severities and metabolic response to drug treatment, etc. These studies have provided significant insight in clinical studies over the years and current technical advances are further facilitating the improvement of this field. It is also currently one of the most useful techniques to determine the spatial distribution of neuropeptides in cells and tissues. Although the underlying concept is simple, factors such as choice of ionization method, sample preparation, instrumentation and data analysis must be taken into account for successful applications of MSI. Herein, we briefly reviewed these key elements yet focused on the applications of MSI in clinical studies and neuropeptide probing that cannot be addressed by other means. Challenges and future perspectives in this field are also discussed to conclude that the ever-growing applications with continuous development of this powerful analytical tool will lead to a better understanding of clinical diseases and neurochemistry.

2.1 Introduction

The integration of gel electrophoresis and liquid chromatography coupled to MS enables high throughput characterization of complex proteomes to detect disease biomarkers [1-4] or evaluate the response to the exposure of drugs or stress. [5-8] Yet spatial localization of the detected proteins and their corresponding expression changes to specific disease or treatment cannot be attained. Conventional immunohistochemical staining (IHC) allows for obtaining high resolution distribution images of targeted proteins.

IHC is also employed frequently for neuroscientists to map neuropeptides. Neuropeptides represent the major class of chemical messengers in organisms with nervous systems and affect multiple physiological processes and behaviors like hunger, thirst, pain, stress and reproduction. [9-15] Neuropeptides are cleaved from large preprohormones, resulting in biologically active forms after numerous enzymatic processing steps in various orders. [16, 17]

However, a significant limitation of this standard IHC method is the need for labeling, which means that the target molecules must be known prior to the experiment. Alternatively, MSI has evolved as a powerful tool for the analysis of a wide range of molecules, mainly using matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), but also using other ionization methods such as desorption electrospray ionization mass spectrometry (DESI-MS) and secondary ion mass spectrometry (SIMS). MSI has enormous advantages over conventional IHC techniques in that not only is it

label-free, but it also enables simultaneous mapping of numerous molecules in tissue samples with great sensitivity and chemical specificity. MALDI-MSI has proven to be a valuable technology with numerous applications in localizing proteins, [18, 19] examining lipid distributions [20] and mapping neuropeptides, [21, 22] at the organ and cellular domains, by varying the experimental conditions. Since no prior knowledge of molecular identities is required for MSI applications, it has also become a standard discovery tool to compare analyte expression pattern changes by analyzing multiplexed data sets.

Herein we review current publications that underscore the critical role of MSI, especially MALDI-MSI, play in the study of molecular dynamics in the context of clinical and neurochemical researches. For clinical research, a better understanding of the molecular pathology of various diseases such as cancer and neurodegenerative diseases can be obtained, and this information can be used for more efficient diagnoses and improved treatments. On the other hand, the ability to visualize neuropeptide distributions in cells, clusters and organs also allows the experimenter to determine if neuropeptides from distinct families or multiple isoforms from the same family co-localize in specific regions. Such spatial information about neuropeptides can help elucidate the underlying mechanism of cell-cell signaling interactions via neuropeptide messengers.

2.2 Methodology

2.2.1 Ionization

MALDI has shown its revolutionary power with its capability of analyzing a wide mass range of intact molecules spanning from large proteins and peptides to small metabolites and lipids with a 'soft' and efficient ionization source. Other than the wide mass range, MALDI also delivers several other unique features in comparison to other ionization techniques, such as a great tolerance for salts, producing mostly singly charged ions and low femtomole to attomole sensitivity. Moreover, its capability to acquire biomolecules' mass to charge ratio (*m/z)* and additional sequence information utilizing post source decay (PSD) or LIFT (a short form of "potential lift") [23] demonstrated its power for unraveling and understanding molecular complexity. N_2 or neodymium-doped yttrium aluminum garnet (Nd:YAG) lasers are usually employed to perform MSI experiments. The size of the laser beam and the matrix crystal combinatorially determine the spatial resolution of MALDI-MS images that can be attained. To date, a commercially available MALDI-TOF/TOF with a regular matrix sprayer is usually able to provide a spatial resolution of 30~75 μm. The resolution could be further improved down to cellular scale with the advances discussed in **2.2.2** single-cell MSI.

SIMS is another technique that has long been established for MSI applications. In SIMS, ionization takes place by first bombarding a solid sample under a high vacuum with high energy "primary" ions. [24] Following the primary ion impact, secondary ions are sputtered from the sample surface and then be drawn into a mass analyzer, typically a TOF analyzer, for surface chemistry analysis. SIMS provides excellent spatial resolution due to its highly focused ion beam. Nevertheless, the primary ion source is fairly limited to small molecules due to its high energy that is prone to fragment large molecules in the

ablation/ionization process. Recent advances in SIMS, such as metal cluster and polyatomic ion sources, provide significantly greater ion yields of ions in the *m/z* 400–3000 range with a reasonable resolution of 200 nm. [25] However, as suggested by Jones *et al.*, while SIMS can easily achieve spatially resolved images of small molecules for MSI applications, it will not likely be able to match MALDI's analytical capabilities for proteins. [25]

MALDI and SIMS-MSI both require vacuum for specimen analyses, complicating MSI procedure and limiting the applications to live biological samples. DESI, in comparison, is a simple, ambient ionization technique. [26-28] DESI channels charged solvent droplets and ions from an electrospray source onto the surface of interest. [29] The surface is impacted by the charged particles, yielding gaseous ions of the originally present biomolecules. By rastering the electrospray liquid jet across the surface of interest, DESI-MSI is accomplished. DESI's ambient nature and softness allow for the examination of various natural surfaces with no need for matrix.[30] Moreover, its perspectives of miniaturization for field applications strongly poised this technique to be an innovative, portable tool since vacuum is no longer necessitated. [31] The downfall DESI has over vacuum MS methods is its spatial resolution that is generally reported to be approximately 180–200 μm. [29] Solutions have been undertaken to increase this resolution to 12 μm by controlled desorption of analytes present in a restricted region of specimen using a minute amount of solvent between two capillaries comprising the nano-DESI probe. [28]

2.2.2 Sample preparation for MSI

A typical workflow of MSI generally consists of sample preparation, MSI acquisition and data analysis as simplified in **Figure 1**. These elements strongly determine the outcome of MSI experiments and are thereby elaborated in the following sections.

2.2.2.1 Sample preparation for tissues

Sample preparation for fresh tissues

Sample preparation in imaging experiments aim to generate reproducible and reliable MS images directly from tissue sections or cells. The structural integrity and morphology of tissues must be maintained after sample treatment without delocalization and degradation of analytes. The most common procedure after harvesting tissue is snap-freezing in powdered dry ice or liquid nitrogen, followed by storage at -80ºC until use. [32] Another method involves loosely wrapping the tissue in aluminum foil and gently placing it into liquid nitrogen, ice-cold ethanol or isopropanol bath for 30-60 seconds. [33] We recommend the latter approach since the gentler, longer freezing process avoids tissue cracking and fragmentation. Prior to long-term storage, tissue stabilization is recommended to minimize the sample aging effect. Tissue stabilization methods, such as microwave irradiation [34] and heat denaturation by Denator Stabilizor T1 (Gothenburg, Sweden), have been reported to effectively deactivate proteolytic enzymes, preventing post mortem degradation of proteins or peptides of interest. [35]

MSI experiments typically require 10-20 μm thick tissue sections. [36] Embedding tissues in supporting media such as gelatin [21] or sucrose [37] allows for easy handling and

precise sectioning of tissue samples without introducing interferences to mass spectrometer, whereas polymer-containing material, such as optimal cutting temperature (OCT) compound, Tissue-tek and carboxymethylcellulose (CMC), should be avoided. [33]

Tissue sections are then transferred and attached onto a stainless steel conductive plate or indium-tin-oxide (ITO)-coated conductive glass slides by thaw-mounting. [33] ITO-coated glass slides are more routinely used to analyze biomedical samples these days since it allows microscopic analyses of the tissue in MSI experiments with clinical diagnostic purpose.

Washing tissue sections with organic solvents is another recommended step to increase ion yields of protein/peptide signal by fixing tissues and removing ion-suppressing salts and lipids. [38] Through comparison with various solvents, [39, 40] graded ethanol has become a routine procedure for MSI of clinical proteomics by enhancing protein signals effectively and reproducibly. However, concerns arise that such washing step might result in diffusion of small, soluble neuropeptides. A new procedure developed by van Hove et al. uses solvent wetted fiber-free paper to enable local washing of tissue sections for MSI applications. [41]

In contrast, SIMS and DESI-MS require relatively simple sample preparation, mainly involving tissue sectioning, due to their limited applicability to low molecular weight (MW) analytes instead of peptides and proteins.

Sample preparation for formaldehyde-fixed paraffin-embedding (FFPE) tissues

Although many successes have been achieved using fresh frozen tissues for MSI, the most commonly used preservation technique for clinical tissue specimens is FFPE. FFPE allows long-term storage at room temperature and is a preferred tissue preservation technique by clinical researchers. Nevertheless, the use of formalin leads to protein/peptide cross-linking and tissue embedding in paraffin wax brings severe interferences to MS analysis. [39, 42] The Caprioli group reported a novel technique to re-gain access to the proteins through a process of antigen retrieval by heating. [43] This process first denatures proteins, thus allowing for enzymatic digestion of these accessible proteins to non-crosslinked peptides that can be analyzed by MS. The peptides of interest are then isolated for MS/MS sequencing and searched against a database to retrieve the parent protein. With this approach, the distribution information of proteins can be obtained from FFPE tissues and correlated to the histology. This method opens up new possibilities of discovering disease biomarkers by analyzing a larger pool of samples from tissue bank, compared to available fresh frozen tissues, for statistical confidence. Unfortunately, FFPE tissues are not amenable to analyses of lipids, which are involved in various biological and disease processes, with current technologies, due to the repetitive washing of the tissues.

Laser capture microdissection

A major challenge of MSI is the lack of selectivity of the analytes during the ionization process and thus the ion suppression effect could impact the detection of low-level analytes of interest. One solution to this problem is to purify cells of interest from

the entire tissue section by laser capture microdissection (LCM) prior to MSI. LCM is a technique that allows for isolation of samples down to single cell scale from thin tissue sections by irradiating a focused laser beam to the target region that is adhered to a heat-sensitive polymer film and subsequently removed upon irradiation. [44] The incorporation of LCM into MSI pipeline for spatially-resolved sampling has been demonstrated by mapping of proteins in the mouse epididymis, [45] ocular lens, [46] human breast tissue [47] and rat kidneys, [48] etc.

In situ tryptic digestion

Although MALDI-MSI is powerful in mapping proteins throughout tissue sections, it is significantly more difficult to identify the potential proteins of interest via MALDI-MS alone due to limited fragmentation obtainable from the singly-charged, high-mass ions produced by MALDI-MS. Therefore, the ability to digest the proteins and perform MS/MS sequencing directly on the tissue allows us to identify proteins *in situ* with high confidence without losing spatial resolution, thus facilitating the discovery of protein disease biomarkers. Groseclose *et al.* carried out *in situ* digestion by robotically spotting trypsin solution onto a coronal rat brain section in a well-defined microspotted array followed by automatic deposition of matrix. [49] Subsequent collection of MS and MS/MS spectra enables sequencing of tryptic fragments and thereby protein identification. This technique is fairly flexible in the workflow of MSI, compatible with upstream preparation like FFPE-tissue processing [50] or downstream coupling to ion mobility analysis. [51, 52]

2.2.2.2 Single-cell MSI

While MSI offers a powerful tool for clinical diagnosis and disease prognosis evaluation, single-cell MSI presents unique merits and challenges that deserve special attention. One limiting factor of single-cell MSI is instrument sensitivity, [53] since the amount of analytes decreases as the square of the spatial resolution, as reviewed by McDonnell. [54] Induction-based fluidics (IBF) reported by Tu *et al.* alleviated this issue by using a charged capillary tip to charge the nL-sized matrix solution followed by launching the droplet onto the specimen. [55] This technique is amenable to single-cell MSI since the analyte diffusion is minimized and the resulting intensity is increased by 10-fold compared to conventional manual spotting by pipette.

An adaptable and cost-effective setup, the stretched sample method, was developed by Zimmerman *et al.* to improve the spatial resolution for single-cell scale MSI of neuropeptides. [56] Briefly, dissected, individual cells, like neurons, were placed on a stretchable substrate, Parafilm, embedded with a monolayer of bead array Then Parafilm was manually stretched and placed on conductive ITO glass slides for further MALDI-MSI experiments. [56, 57] The MS images of the cells are registered to their original locations based on the optical image taken prior to stretching.

The Sweedler group further improved the spatial resolution by reducing the incremental movement of the laser beam to smaller than the laser diameter. [58] They introduced this setup as an "oversampling" method by desorbing/ionized from a much smaller area with each incremental step after sample material is ablated away at the initial spot. Such oversampling procedures are relatively simple and do not need any additional

hardware. Nevertheless, this method flaws in that the total sample has to be consumed, excluding the possibility to re-analyze the samples.

On the other hand, several custom-designed mass spectrometers permit single-cell resolution. For example, Spengler *et al.* fabricated special confocal-type objectives, which were furthered modified by Rompp and Guenther *et al.,* offering sub-micron spatial resolution. [59, 60] This ion optical and laser setup was then coupled to a high-end Q-Exactive Orbitrap Fourier transform mass spectrometer (FTMS) by Schober *et al.*, delivering metabolites and lipids' MS images from Hela cells. [61] Resolution as low as 7 μm was reported by Chaurand *et al.* by introducing a coaxial laser illumination ion source to a MALDI-TOF. [62] Moreover, a proprietary smartbeam-II MALDI laser developed by Bruker Daltonics allows beam diameters to be focused down to 10 μm and permits routine MSI analyses to be performed at the cellular level. [63]

An alternative approach to increasing single-cell spatial resolution is "microscope" MALDI-MSI by employing a setup that integrates a defocused UV laser, high-quality ion optical and a position-sensitive detector to record the position of the stigmatic projection of ions. [64] Luxembourg *et al.* reached a pixel size of 500 nm and a resolving power of 4 μm using this configuration. [64] However, the use of this method is limited due to technical constraints like the specialized ion optics, a fast detector necessitated to achieve high magnification and computing software to reconstruct ion images.

MSI has been extensively used in differentiation between healthy and diseased tissues with the purpose of clinical prognostics and diagnostics. Its cutting-edge

applications to single-cell analyses will provide detailed biochemical information at a cellular scale for mechanistic understanding and, ultimately, development of therapeutic treatments. The recent advances of single-cell MALDI-MSI and its clinical impact are presented by Boggio *et al.*. [53]

Single-cell is also critical to neuropeptide studies. Distinct neuropeptide expression patterns have been observed at the distal regions and the soma of a neuron or among adjacent neurons. Because of this chemical heterogeneity, information about neuropeptide identities and locations in individual neurons is valuable to neurobiologists. Determining neuropeptide complements at the cellular level can help researchers understand the role of neuropeptides in intracellular regulation and intercellular communication. Additional sample preparative steps must be taken to achieve single-cell MSI from neurons compared to tissue specimens. A representative study by the Sweedler group using mollusk to develop and improve sample preparation techniques in single-cell MSI is useful. Their approach to single-cell MSI began with the isolation of single neurons. The ganglion was incubated in protease solution first to remove the sheath. The neurons were then dissected using sharp needles and transferred onto a glass slide. Following dissection, the neurons were placed and cultured in a dish filled with artificial sea water (ASW)-antibiotic solution. [65] For MALDI-MS profiling, neurons can be placed onto a MALDI target plate after briefly rinsing to remove salts. Methods for MSI were investigated to preserve cell morphology, prevent neuropeptide redistribution and replace the cell matrix. A 30% glycerol-ASW mixture was found to substitute ASW-antibiotic media well and stabilize neuron morphology without ion-suppression. After exposure to

glycerol for 1-5 minutes, the extracellular media was removed by vacuum suction from the desired cell culture regions. [65] The extracellular media must be optimized for specific types of neurons to preserve cell morphology and prevent neuropeptide redistribution.

2.2.3 Matrix application

Choosing a matrix and its application method are critical to MSI results. Other than conventional matrices such as α-cyano-4-hydroxycinnamic acid (CHCA) and 2,5-dihydroxy benzoic acid (DHB), [21, 66] ionic matrices made by mixing conventional matrices with organic bases are also widely used and reported to improve spectral quality, crystallization and vacuum stability. [67]

Matrix application is another area under continuous innovation. Several types of matrix application apparatus exist by deposited matrix either as homogeneous layers (spray coating) or discrete spots (microspotting). Pneumatic spray device such as pneumatic sprayer, airbrush, TM sprayer system from HTX imaging and thin layer chromatography sprayer, and vibrational spray apparatus like ImagePrep Device from Bruker Daltonics are all capable of applying a uniform layer of small to medium-sized matrix droplets, [21, 22] whereas microspotters like CHIP produced by Shimadzu deposit pL-sized droplets of matrix according to a predefined array and require multiple rounds of spotting for sufficient matrix coverage. In comparison to the wet application, solvent free methods like sublimation or dry coating yield very fine crystals amenable to high-spatial-resolution-MSI. However, it suffers from relatively low sensitivities due to the limited analyte-matrix interactions. [68] Deutskens *et al.* modified this procedure by

rehydrating the sections following dry-coating and improved the detection sensitivity of proteins from rat cerebellum sections. [69]

2.2.4 MS Instrumentation

An ion source can be coupled to most state-of-the-art mass analyzers for MSI capabilities. Time-of-flight (TOF) analyzers are most extensively used in MALDI-MSI applications. In TOF analyzers, desorbed ions are accelerated to the same kinetic energy and the *m/z* is determined by the time the ions take to travel through the TOF tube. [70] This design provides high sensitivity, a wide mass range (2~30 kDa) and fast analysis speed, thereby favored by the applications of MSI to clinical proteomics. Nevertheless, the characterization of peptides and low MW compounds necessitate the addition of another TOF tube, allowing for accurate measurement of peptides' and small molecules' masses and tandem mass fragmentation for sequence validation. [71]

The integration of ion mobility (IM) with MALDI-TOF is a breakthrough to MALDI-MS-based analyses. IM-TOF is a two-dimensional gas-phase separation technique that discriminates ions based on their *m/z* and collisional cross-section. Upon desorption/ionization from the tissue surface in MALDI source, ions travel inside an IM drift cell, which is equipped with an applied electric field and a carrier buffer gas that opposes ion motion. The ion's mass, charge, size and shape combinatorially determine its migration time inside the drift cell, consequently managing to discriminate isobaric ions according to their collisional cross-section. [72] For instance, IM-TOF has proven its usefulness in MSI of *in situ* digested proteins. Stauber *et al.* successfully separated and identified isobaric ions of Tublin and ubiquitin peptide fragments at *m/z* 1039 directly

from a rat brain section based on different drift times, whereas the MS images of the two peptides would be merged without the additional separation by IM. [52] The benefit of performing IM separation prior to MS measurements is enormous and promises to find extensive applicability to MALDI-MSI-based research.

Compared to the relatively low resolving power offered by a TOF/TOF mass analyzer, Fourier transform ion cyclotron resonance (FT-ICR) mass analyzers provide superior resolution and accuracy for unambiguous discrimination of analytes without sacrificing spatial resolution. [73] Moreover, it also provides unique ion trapping and storage capabilities, which are utilized by Kutz *et al.* for in-cell accumulation [74] and by Bruker Daltonics for continuous accumulation of selected ions (CASI) [75] to improve the detection sensitivity of trace-level analytes. The primary drawback of FT-ICR is the slow scan speed, which lowers the throughput for serial MSI experiments. A more powerful magnetic field could partially alleviate this problem by increasing the ion cyclotron frequency without affecting spectral quality. In addition, the Orbitrap manufactured by Thermo Fisher Scientific provides comparable resolution to an FT-ICR. Orbitrap determines the *m/z* by measuring the axial oscillation frequency of ions back and forth along a spindle-like electrode within an electrostatic field, [76] which is proportional to the square root of the electrical-field strength, whereas the cyclotron frequency measured by FT-ICR is related to the magnetic field strength. [77] To obtain high resolving power on an FT-ICR, high-field magnets are needed, concomitantly at a high cost. Predictably, the Orbitrap is gaining more popularity in future MSI applications due to its excellent performance and relatively low cost compared with FT-ICR.

Besides mass analyzer, laser optics advancements have also improved MSI measurements in multiple categories. Conventional MSI experiments are based on the microprobe approach. In the "microprobe" mode, the laser beam is focused to a small region from which a single spectrum is acquired. By rastering the surface according to a predefined Cartesian grid, an array of mass spectra is acquired and reconstructed to a cohesive MS image where each pixel contains the data from the corresponding spectrum. The laser beam focus determines the pixel density of a MS image, i.e. the spatial resolution. Alternatively, a "microscope" mode could be applied to MALDI-MSI applications. This setup employs a defocused UV laser beam to envelop a large area, from which a stigmatic projection of ions were desorbed/ionized. The high-quality ion optics transfer and magnify the ion packet produced by the single laser shot, relying on a position-sensitive detector to record the arrival time and position of the ions. [64, 78] An increased number of pixels are simultaneously acquired, which dramatically increases the throughput of high resolution imaging. The resulting stigmatic images of highest magnification reported by Luxembourg et al. reached a pixel size of 500 nm and a resolving power of 4 μm. [64, 79] In this configuration the limiting factor for spatial resolution is no longer the size of the desorption/ionization area, but rather, the quality of the ion optics, the ion optical magnification, and the pixel size of the detector. Unfortunately, the method is not widely used because of technical constraints like the need for specialized ion optics, a fast detector and computing software to reconstruct ion images.

2.2.5 Histology-guided MSI

One of MSI's advantages is that spatial localization of analytes within a tissue is preserved in comparison to homogenization. It allows for differentiation of tissue regions based on molecular features. Therefore, MSI can be integrated with histology to directly target for diseased regions and correlate the peaks that are up-regulated/down-regulated in the specific regions with the reference information provided by histology. To date, histology-guided imaging has become the standard for applying MSI to clinical diagnostics. [53, 80, 81] Hematoxylin and eosin (H&E) staining is commonly performed on clinical tissue samples. This technique allows the visualization of cells with bright field microscopy by labeling the nucleus and cytoplasm of the cells. [53] A trained pathologist can then use the H&E stained tissue to establish a diagnosis. [40] Methods of histology guided MSI have been developed that allow both the histological features and the MS image of the tissue to be observed. In such methods, the histological stain is applied to a serial section of the tissue that is to be analyzed by MSI, which dissipates the concern for mass spectrometry-friendly stains. [80, 82] One downfall to this method is that it can be difficult to obtain serial sections, reproducibly, without tearing or folding the tissue on the cryostat. [83] This can cause a misalignment of the sections, causing a loss of correlation of small features. Another method is to apply the histological stain directly to the tissue that is to be analyzed by MSI. This method requires MS-compatible stains, such as cresyl violet or methylene blue, or performing imaging first and washing off the matrix before applying the histological stain. [53, 83] The downfall of this method is that H&E staining is not compatible with MSI. Therefore the standard workflow would have to be changed in order to include non-standard histological stains and non-routine evaluation of the results.

[83] The future of MSI in clinical diagnostics relies on incorporating it into the current diagnostic workflow, and must be compared and correlated with histological results. [81]

2.2.6 Data analysis

MSI data analysis software, such as BioMap (http://www.maldi-msi.org, available for free downloading) and proprietary programs for MSI systems (e.g., FlexImaging from Bruker Daltonics, ImageQuest from Thermo Fisher Scientific and TissueView from Applied Biosystems/MDS), are mostly employed to produce distribution maps for selected analytes. These software packages allow the user to adjust color scales, overlay ion density maps, and integrate MS images with acquired histological pictures.

Other than image processing, software that provides statistical support are also applied to MSI data analysis. For example, biomarker discovery studies usually involve comparing MSI data sets sampled from a control group with those from a treated group. Data-mining software, like ClinProTools by Bruker Daltonics, capable of performing principal component analysis (PCA) and hierarchical clustering of multiple MSI data sets to extract differentially expressed or distributed molecules, are widely used for identification of potential disease biomarker candidates.

2.3 Clinical Applications of MSI

Over the past decade, MSI has become a powerful tool that has been extensively applied to various clinical applications. In this section, we present current advances of this technique and its novel applications in clinical setting, including drug response

measurement (**3.1**), lipid and protein biomarker discovery (**3.2**) and several other novel applications (**3.3**) as summarized in **Figure 1**.

2.3.1 Measuring Drug Response and Metabolites

Histology guided MSI has been extensively applied to the study of drug response and distribution. [84-86] Drexler *et al.* combined quantitative whole-body autoradiography (QWBA) and imaging mass spectrometry to simulate an *in vivo* phototoxicity study of a proprietary drug candidate (Bristol-Myers Squibb Company, Princeton, NJ) within ocular tissue. [85] QWBA reveals quantitative and tissue distribution information of radiolabeled analytes. [87] However, QWBA cannot be used to obtain information about the drug and its metabolites or degradation products. MSI is notorious for giving poor quantitative data, but can reveal information about the analyte and its products; therefore it was advantageous to combine QWBA and MSI for this study. The radio-labeled and non-labeled drug candidate, BMS-X, was dosed to two different test groups of rats. Two hours after the dose was administered, the animals were sacrificed for QWBA and MALDI-MSI experiments. The QWBA experiment showed that radio-labeled analyte(s) preferentially distributed to the back of the eye, not the cornea or lens, but revealed no information about the molecular species of the radio-labeled analyte(s). The MSI experiments detected only BMS-X, again in the back of the eye, therefore suggesting that there were no metabolites or degradation products of the drug present in the tissue. Employing complementary technologies of MALDI-MSI and QWBA, Drexler *et al.* were able to assess the distribution of a new drug candidate and its potential metabolites which

is essential for phototoxicity studies of pharmaceuticals before they can be prescribed in a clinical setting.

For pharmacokinetic and toxicity studies, the quantitative analysis of the *in vivo* distribution of a drug after it has been administered is essential. One of the major limitations to MSI is the lack of ability for successful quantification, as mentioned previously. A quantitative research method using MSI is greatly desired and could be very insightful for the pharmaceutical industry and for clinical diagnostics. Takai *et al.* developed a quantitative MALDI-MSI approach to analyze the drug, raclopride (RCP), in multiple mouse organs and study drug distribution/ accumulation. [84] RCP (*m/z* 347) is a dopamine D2 receptor-selective antagonist. RCP was initially spotted onto tissue sections and MS/MS experiments were performed in order to choose a fragment ion to monitor for the quantitation studies. The daughter ion *m/*z 129 was analyzed for the quantitation study. RCP was dosed to mice intravenously. The mice were sacrificed and whole-body sections of the mice were imaged with MALDI-MSI at different time points post-dosage:10 minutes, 30 minutes, and 60 minutes post-dosage. The average signal intensities (n=3) of RCP were calculated for each organ are each time point. For more precise quantification, Takai *et al.* used a normalization method by using the signal intensity of DHB peaks as an internal standard. The RCP signal intensities were divided by the DHB signal intensity within the specific organ of interest in order to normalize the results. Additionally, RCP was administered to mice and the concentrations of the drug in each organ studied (liver, heart, spleen, brain, and plasma) were determined by LC/MS/MS for comparison with the quantitative MSI procedure. There was a strong

correlation observed between the concentrations determined from the LC/MS/MS and the intensities calculated in the MSI method suggesting the future potential of MSI for quantitative analysis in multiple organs simultaneously.

Ambient ionization methods have gained interest and popularity for clinical MSI. [86] Wiseman *et al.* utilized desorption electrospray ionization (DESI) MS, an ambient ionization technique, to image drugs and their metabolites in histological tissue sections. Their study includes monitoring the antipsychotic drug, clozapine (*m/z* 327.1) and its dominant, N-desmethylclozapine, metabolite (*m/z* 313.1) in rat brain, lung, kidney, and testis tissue samples. By using their developed DESI-MSI method, and imaging in MS/MS mode, Wiseman *et al.* were able to detect Clozapine not only in brain tissue, but also the accumulation of the drug in lung, kidney, and testis tissue sections, which was previously unknown. The research team demonstrated the ability of DESI-MSI to detect small molecule pharmaceuticals and their metabolites simultaneously, directly from histological tissue sections. They were able to rapidly detect the drug and metabolite and display spatial distributions as well as provide relative quantitation with minimal sample preparation, giving this ambient ionization technique potential for use in the clinical setting.

There have been many reports of drug and metabolite imaging in the literature. [88-90] Using multi-faceted approaches to distinguish between the drug and its metabolite provides more detailed and useful information for potential use in clinical studies. The

ability to quantify active drugs simultaneously in multiple organs provides another key layer of information for toxicology studies for new potential drug candidates.

2.3.2 Biomarker Discovery and Validation

The discovery and validation of biomarkers, along with other laboratory and clinical evaluations, contribute to the assessment of disease severity, disease progression, and treatment response. [91] Gene expression profiling was the prominent source of biomarker discovery, but more recently, protein expression profiling has exploded in this field, [92] and lipidomics has been gaining popularity. [93]

2.3.2.1 Protein Biomarkers

Changes in protein levels in a tissue can correlate with disease state. These protein levels can be monitored in order to reveal, not only the type of disease, but also the severity of the disease. MSI has emerged as a powerful tool for biomarker discovery largely because of its ability to probe the proteomics of the tissue while maintaining the spatial distribution, which allows for direct comparison with histology. [80] MSI can be applied simultaneously to multiple tissues, allowing a more in depth study of the target disease and its potential biomarkers. [81]

Six common types of cancer (Barrett's cancer, breast cancer, colon cancer, hepatocellular carcinoma, gastric cancer, and thyroid carcinoma) have been probed with MALDI-MSI by Meding et al. Patient diagnosis begins with tumor origin identification and classification and when a primary tumor cannot be identified, the sample is

diagnosed as cancer of unknown primary (CUP). [92] The researchers used MALDI-MSI to establish distinct protein biomarkers for each type of known cancer. Cancer cell specific spectra were extracted and classified based on their proteomic differences with high confidence. These data were applied toward the identification of cancer from CUP samples. It is extremely important for the development of an individual patient treatment regimen that the secondary tumor (metastasis) be identified even if the primary tumor cannot be found. Meding *et al.* introduced colon cancer liver metastases into the study in order to test the ability of MALDI-MSI to distinguish between colon cancer primary tumors, colon cancer liver metastases, and hepatocellular carcinomas (liver cancer). Their results indicated that MSI can be used to distinguish between, and classify these closely related cancer entities. This research team generated a classifier system based on MALDI-MSI methods that provides accurate tumor classification with high confidence levels. This method could become a valuable addition to the workflow for clinical tumor diagnostics.

Not only are biomarkers useful for diagnosing and classifying different types of cancer, they are also being used to distinguish between diseases with similar histological characteristics. One such example is Spitz Nevi and Spitzoid malignant melanoma which is found primarily in children. [94] Spitz Nevi is a benign skin lesion, whereas Spitzoid malignant melanoma requires surgery and chemotherapy. [95] These two diseases are very difficult to distinguish and can result in misdiagnoses when relying on histological criteria alone. [94, 95] Lazova *et al.* performed a study using MALDI-MSI to compare the protein profiles of Spitz Nevi and Spitzoid malignant melanoma. [95] They have found 5

peptide peaks (*m/z* 976.49, 1060.18, 1336.72, 1410.74, and 1428.77) that were able to best discriminate between the two diseases, and a total of 12 discriminatory peaks that were used to build a classification model. Similarly to Dill *et al.* , as mentioned previously, Lazova *et al.* developed their model with a training group, and validated it with a test group. The classification method for the tumor showed a sensitivity of 97% and specificity of 90%. Interestingly, the research team was able to correctly classify 28 of 31 samples based solely on the proteomic information found in the dermis and not the tumor itself. These results show great promise for improving the diagnostic accuracy of these diseases in conjunction with standard histological evaluations.

Biomarkers can also show signs of effective treatment of a disease, not just identify the disease. Kim *et al.* utilized MSI to study drug distribution and potential biomarkers of response to therapy in prostate cancer. [96] The research team sought to study the spatial pharmacokinetics of prostate tumors treated with a novel tyrosine kinase inhibitor, AEE788. AEE788 potentially inhibits VEGFR and EGFR in nanomolar concentrations, which could enhance radiotherapy. Mice with prostate tumor xenografts were split into 4 groups: treated with AEE788, treated with radiation therapy, treated with AEE788 and radiation therapy, or untreated. The tumors were imaged for drugs or potential biomarker proteins. The drug and protein distributions were altered when the tumors were treated with irradiation. One potential biomarker was found at *m/z* 7765.4 which was present in the tumors treated with AEE788, but not when treated with radiation therapy alone. Radiation treated tumors had increased expression of multiple proteins that were not present in the AEE788 treated tumors. These could be potential biomarkers for treated

vs. untreated tumors, which could provide means for evaluating efficacy of different therapeutics and optimizing individual patient treatment regimes.

2.3.2.2 Lipid Biomarkers

Lipidomics has recently gained popularity in clinical applications. Similarly to protein biomarker discovery, lipidomics aims to characterize lipid molecules in tissues and relate this information to disease states, among other applications. [93] Proteomics aims to characterize and quantify the cellular performers reflecting gene expression, whereas lipidomics (a subset of metabolomics) studies the molecular products of metabolism and is closest to patient phenotype, making lipidomics an important field of study. [97] Traumatic brain injuries can have a great effect on the lipid distribution and composition of the brain. Cox *et al.* used MSI to study brain lipids after cortical impact injury. [26] For this study, rats were subjected to controlled cortical impact injury and the brains were removed at different time points of 24 hours, 3 days, and 7 days post injury. The distribution of lipid species in the brain showed large variation in the MS images, demonstrating time dependent changes in the lipid profile of the injured areas after traumatic injury. Not only do these results show lipid biomarker distributions that can reveal the extent and time post-injury, which can be used to monitor patient recovery after treatment, but also the observed changes in lipid abundances shed light on the mechanisms of injury and repair on the molecular level.

On a related note, Hanada *et al.* studied the alteration of phospholipids and prostaglandins after a spinal cord injury. [98] Alterations in lipid metabolism may play a

key role in neurological disorders; therefore it is important to study these alterations to the lipid profiles during the occurrence and progression of a spinal cord injury. For this study, rats with applied spinal cord injuries were used, and the lipid distributions were studied after 12 hours, 1 day, or 1, 2, or 8 weeks post-injury. Unique distribution patterns were observed for different types of phospholipids as the tissue damage resulting from the spinal cord injury progressed. Spatiotemporal changes in phosphatidylcholines (PCs) were examined with MSI. Both temporal and irreversible alterations in distinct PC species were observed, especially in species that contained 3-docosahexaenoic acid (DHA). **Figure 2** shows the spatio-temporal images of two different DHA-containing PCs. At the impact site, an irreversible change in DHA-containing PCs was observed over time. This observation may result from an irreversible deficit of the neurons and could lead to motor dysfunction, meaning that DHA-PC reduction could be a potential indicator of the pathology of spinal cord injuries to be used in clinical settings.

2.3.2.3 Small Molecule Biomarkers

Ambient ionization methods can also be used to study and identify biomarkers in cancerous tissues. [26, 99, 100] Dill *et al.* used DESI-MSI to study human bladder cancer and develop a statistical method for distinguishing between cancerous and non-cancerous tissue samples in conjunction with standard histological identification methods. [26] The researchers used a training set of samples to establish the best predictive features and a test/validation set to evaluate the performance of their statistical model on representative samples. There were 20 pairs of cancerous and adjacent normal human bladder tissue

samples used in this study. DESI-MSI data show significant changes in glycerophosphoinositols, glycerophosphoserines, and fatty acids between the cancerous and normal tissue samples. A series of DESI-MSI images were used to visually characterize the distribution of particular molecular species across the set of tissue samples. These images were visually compared to optically scanned images of the H&E-stained histological sections. Their results agreed with the pathological diagnosis of cancer and normal tissue for 15 of the tissue pairs. There was excellent agreement between the H&E-stained sections and the DESI images. Interestingly, for one sample, a border of tumor was detected on the normal tissue section, demonstrating the utility of DESI-MSI for determining the margins of the tumor before surgery. Overall, these results are very encouraging for the development of a method that could be used in a clinical setting for the diagnosis of cancer.

2.3.3 Unique Applications for Clinical Diagnostics

Apart from drug distribution and biomarker discovery, MSI has been used for several other clinical applications. [101-103] Osteoporosis is well-known disease for which biomarkers have been extensively studied [99], but little attention has been devoted to bone material quality. [102] Routine clinical fracture risk assessments do not consider the quality of the bone mineral matrix. Zoehrer *et al.* used SIMS-MSI to investigate the spatial distribution and relationship of phosphorus and calcium in bone. [102] Calcium (Ca) and phosphate (P) are the main elements of the bone mineral building block, hydroxyapatite, and therefore support bone tissue mineralization. Bone material properties of the femoral

head in male patients, age 65-80, with fragility fractures were compared to male, age-matched, non-fracture controls. SIMS-MSI clearly showed a greater frequency of areas on high Ca ion intensity in the tissue samples with fragility fractures, when compared to the control group. In the fragility fracture samples, a distinct, ~25 μm wide line of Ca ions was observed along the surface of the endosteum portion of the trabecular bone. In the control group, this line of Ca ions was distributed more diffusely, over a larger area of ~50 μm. Regions of high P ion intensity were observed in the control samples. The overall results show that a significant decrease in the P levels is associated with fragility fractures and the Ca/P ratio and distribution on the trabecular bone could be valuable parameters to consider during therapeutic and diagnostic trials.

Assessing the compatibility of treatments with the body is a key step necessary to ensure the safety of the patients. Biodegradable polymers are of high interest in the medical field because of their potential applications in the field of tissue engineering and as drug delivery carriers. [103] Supramolecular polymers have gained much interest as potential drug delivery carriers because of their compatibility with poorly water soluble compounds and their potential for controlled drug release. [27, 104] In drug delivery systems, these polymers are designed to locally deliver active pharmaceutical ingredients to a localized place in the body, that needs the drug, and expose other parts of the body at a much lower dose. Klerk *et al.* has used SIMS-MSI to elucidate molecular distributions in and around such polymer. A supramolecular polymeric hydrogel (containing no drugs) was implanted into rat renal tissue and the rat kidneys were harvested 15 days after implantation. Half of the kidneys were prepared for MSI and half

were used for histological comparison. The MSI images show lipids from the kidneys, the polymer capsule, and cholesterol distribution around the implantation site, as well as silicone contaminants from the surgical tools. The images show that lipids from the kidney tissue entered into the polymer capsule, as shown in **Figure 3**, demonstrating biological activity within the polymer. When the study is continued with a polymer implant that does contain active pharmaceutical ingredients, the researchers will have the opportunity to study drug release by imaging the drug distribution in the surrounding tissue at various time points. Overall, this research shows the great potential of SIMS-MSI for understanding controlled drug distribution.

Biodegradable polymers are also being applied to the field of tissue engineering, requiring extensive study of the compatibility of these materials with the body. In tissue engineering, scaffolds made of biodegradable polymers can be designed to be slowly degraded by the body as their function is taken over by newly generated tissue. During renal failure, it may be necessary for a patient to undergo hemodialysis, which requires the doctors to surgically form an arteriovenous fistulae (AVF), using vein grafts. [101] In this procedure, the patients artery and vein are sutured together with a graft connecting them, which allows for the blood to bypass the capillaries and the vein to be enlarged so that it can accommodate the volume of blood being transferred during hemodialysis. [105] These grafts are typically made, not with polymers, but with the patient's existing veins. Although the grafts come from the patient themselves, the veins in and around the graft degrade over time due to biological changes in the tissue. [101] Tanaka *et al.* used MALDI-MSI to study the lipid profile of AVF tissues. [101] In this study, there were three types of

vein samples: AVF tissues, control veins (CV), and peripheral artery occlusive disease (PAD) tissue. PAD is a disease that causes veins to deteriorate and is caused naturally, rather than through surgery like AVF. The control veins were segmental cephalic vein tissues taken from patients who underwent AVF creation, but the CV tissues were not part of the AVF. MALDI-MSI was performed on all three tissue samples, and the distributions of lysophosphatidylcholine (LPC) and phosphatidylcholine (PC) were determined. The results show that the distributions of LPC and PC differed between the three tissue types as shown in **Figure 4**. LPC was localized in the intima of the CV and PAD tissues, but distributed through the media and adventitia of the AVF tissue. The intensities of PC were much higher in the AVF and PAD tissues than it was in the CV tissue, suggesting an abnormal accumulation of these lipid molecules in both tissues. These results were compared with histological data, and both revealed complementary patterns of lipid accumulation in the AVF and PAD tissue samples. Overall, the MALDI-MSI results of this study provided the first evidence of the characteristic lipid distribution of AVF tissue on the molecular level which suggests an association between molecule-induced inflammation and tissue degeneration. This research provides valuable insight into the cause of vein degradation that can be applied to future procedures developed for hemodialysis or for the continued monitoring of patients with AVF's already in place.

2.4. Neuropeptide Signaling

2.4.1 MSI of neuropeptides at the organ level

The successful application of MALDI-MSI to proteins contributed to its extension into signaling peptide localization in neuroendocrine organs. Verhaert et al. used a MALDI LTQ-Orbitrap to image secretory neuropeptides from the American cockroach (*Periplaneta americana*). [37] They also differentiated neuropeptides in *Acheta domesticus* (the house cricket) from lipids by examining the diagnostic tandem MS fragments of the two classes. [106] Similarly, Dekeyser et al. investigated the differential distribution of neuropeptides in the brain and pericardial organ from the crab *Cancer borealis*. [22] Over 30 neuropeptides from 10 neuropeptide families were detected using a MALDI-TOF/TOF and characterized by *in situ* fragmentation to verify peptide identities. Spatial relationships between multiple neuropeptide isoforms of the same family and the relative distributions of neuropeptide families were elucidated in this study. Intriguingly, both of the studies found that neuropeptides appeared to be more differentially localized than phospholipids.

The chemical heterogeneity of neuropeptides in complex neural tissues like the brain necessitates the interrogation of neuropeptide expression patterns in a tissue volume. Three-dimensional (3D) images of neuropeptides reconstructed from 2D images of serial sections deliver contextual information to planar mapping and lower the possibility of neglecting small anatomical structures. Chen *et al.* established a protocol to examine the volumetric distribution of various neuropeptides from a crustacean brain. [21] By using seven sequential sections with an equal interval of 132 μm on the z-axis of the brain, over twenty neuropeptides belonging to eight families were mapped throughout 3D structures. A reconstructed 3D image of a crustacean neuropeptide *Cancer borealis* tachykinin-

related peptide (CabTRP) 1a is shown in **Figure 5a**, displaying high abundance on the surface of olfactory lobes in the brain. In contrast, a 3D representation of phosphatidylcholine 38:6 (**Figure 5b**) shows that it is distributed relatively evenly throughout the whole brain structure. Another example shows the 3D distribution of an RFamide-like peptide SMPSLRLRFa. These results demonstrate that 3D imaging reveals more detailed molecular distribution data than 2D imaging because of the extra dimension. A comprehensive review about 3D MSI applications is a good resource for further reading. [107]

Researchers are exploring new ionization techniques for neuropeptide MSI studies. Although SIMS does not yield intact peptide or protein signals above 1 kDa from biological samples, it can produce submicron MS images. [108] Recently, an integrative study aimed to map neuropeptides, metabolites and lipids from rat spinal cord and demonstrated their colocalization patterns. [109] This multi-faceted MS platform provides an overall picture of molecules utilized by the nervous system. To enhance the desorption/ionization yield of SIMS for neuropeptide imaging, a hybrid technique named matrix-enhanced (ME)-SIMS was developed. [110] The matrix significantly enhanced intensity values, leading to the imaging of three peptides from the cerebral ganglia of the freshwater snail *Lymnaea stagnalis* at a spatial resolution of 3μm. [108]

Advancements in sample-handling and mass spectrometers have allowed imaging at cellular-length scales in several studies. As previously described, Altelaar and Luxembourg et al. developed the microscope approach and modified a triply focusing TOF for this purpose (**Figure 6a**). [64, 79] In this instrument, desorbed ions within a large

defocused laser spot are stigmatically transmitted through the mass analyzer and focused onto a position-sensitive detector. The resulting MS images had a pixel size of 500 nm and an effective resolving power of 4 μm. A series of pituitary neuropeptides were examined at the same magnification, producing MS images with superior spatial resolution. [79] **Figure 6b** shows a high resolution MS image obtained with the microscope approach of the diacetylated neuropeptide α- melanocyte stimulating hormone (MSH) from a mouse pituitary section. In comparison, the digitally resampled microprobe images have a pixel size of 80 μm × 80 μm (red) (**Figure 6c**) and 250 μm × 250 μm (red) (**Figure 6d**) and lack the resolution required to resolve anatomical features on a cellular scale. High spatial resolution images of other neuropeptides like oxytocin and vasopressin from pituitary gland sections were also shown, highlighting the ability of label-free MSI to deliver images at a biologically relevant length scale. Interestingly, neuropeptides of mouse pituitary gland section were again imaged in another proof-of-principle experiment aiming to improve spatial resolution. In that study, a linear ion trap Orbitrap was coupled with an in-house developed atmospheric pressure MALDI imaging ion source equipped with a highly focused laser. MS images of neuropeptides, such as oxytocin, vasopressin, copeptin etc. were obtained with a cellular resolution of 5 μm. [60] Nevertheless, these modifications are not widely practiced because of their complexity.

2.4.2 MSI of neuropeptides at the cellular level

Invertebrates are useful model systems when developing sample preparation techniques for neuropeptide MSI studies of single cells. Many invertebrate organisms have been characterized, including the freshwater snail (*Lymnaea stagnalis*), [111] sea

mollusk (*Aplysia californica* and *Aplysia vaccaria*), [112] crayfish (*Orconectes limosus* and *Procambarus clarkii*), [113] cockroach (*Periplaneta americana*), [37] ticks (*Ixodes ricinus* and *Boophilus microplus*), [114] and a parasitic nematode (*Ascaris suum*). [115] Using vesicles sampled from the exocrine atrial gland of *Aplysia californica* as a model, Rubakhin et al. detected a wide range of bioactive peptides within individual vesicles (1–2 μm). [116] Later, they optimized a MS-friendly glycerol fixation protocol and reported different neuropeptide complements probed from the dendrites and soma of an isolated intact neuron from *Aplysia*. [65] It is anticipated that the development of improved micro-sampling protocols in combination with continuous advancements of high spatial resolution techniques should enable the sub-cellular mapping of neuropeptides from single cells with MALDI-MSI in the near future.

2.5 Future Perspectives

2.5.1 3D MSI

One of the more exciting, recent advances in the field of MSI is 3D MSI. 3D MSI grants the ability to study a broad mass range of molecular species by creating a lateral and vertical distribution map of select compounds. [107] This technique serves as a powerful discovery tool for pathologists and to the pharmaceutical industry by allowing a more complete visualization of tissue samples, which improves the ability for identification of distinct molecular signatures and drug distribution throughout the entire tissue. Advances involving imaging acquisition speed, image resolution, and data

processing are always on-going in the biological sciences. [117] By adding the third dimension to the traditional 2D MSI method, there is a greater need for improvements in sample preparation, data acquisition, and data file transfer. [107] MSI generates extremely large data files; advances in computational tools will need to be made in order to improve the efficiency of data transfer and processing that will be much more time consuming when collecting a third dimension of imaging data. 3D imaging shows great promise for advancing clinical diagnostics and neural network studies, but there are still many areas for potential improvements before it becomes a widespread technique.

2.5.2 Alternative to Formalin-Fixed and FFPE Tissues

The FFPE technique preserves the cellular and morphological details of the tissue. The analysis of human tissue specimen is key to the identification of novel biomarkers that can be used to create more specific therapies and treatments. Recent progress has been made in the analysis of FFPE tissues, but proteomic analysis from these tissues is still very difficult because formaldehyde causes protein-protein cross-linking. PreAnalytiX (Hilden, Germany) developed an alcohol-based, formalin-free tissue fixation system, PAXgene that is commercially available for research use. The PAXgene system is a two-step approach composed of the PAXgene Tissue Fixation Reagent and the PAXgene Tissue Stabilization Reagent. The fixation is carried out without the cross-linking of biomolecules, allowing the stabilization of proteins and nucleic acids while still retaining the histomorphology of the tissue. Ergin *et al.* compared MALDI-MSI capabilities on FFPE-fixed, cryopreserved, and PAXgene-fixed tissue samples. [118] The

protein recovery efficiency of these three types of tissue fixation methods was compared, and the results showed that PAXgene fixation allows for high quantity of protein from mouse and human tissue samples. The protein pattern of these three tissue fixation methods was also examined, and the results showed that PAXgene and cryo-preserved tissue samples revealed similar proteomic signatures when examined by MALDI-MSI, and no protein peaks were observed from the FFPE tissue samples. Overall, these data demonstrate the potential of the PAXgene fixation system to become an integral part of protein biomarker discovery, which will facilitate advancing the clinical applications of MSI, if it becomes more widely used.

2.5.3 Quantitative Imaging

Quantitative imaging mass spectrometry is of great interest to the field of mass spectrometry and to the application of MSI. Advances are slowly being made on the development of a reproducible, quantitative MSI method. Takai *et al.*, [84] as mentioned previously, have developed a quantitative MALDI-MSI approach to analyze the concentration of a drug in mouse tissue after specific time points. This method is based on generating a calibration curve for the drug by spiking different concentrations of the drug of interest directly onto a tissue section and analyzing it with MSI. Their results were compared to trusted quantitation results obtained by LC/MS/MS, and had good agreement. Drexler *et al.* obtained quantitative information by combining MSI with QWBA. The downfall of this approach is that the analyte must be synthesized to contain radiolabels for QWBA, which is often time consuming and costly that are undesirable in

the early stages of drug development. Several other groups combine MSI with a variety of other techniques in order to obtain quantitative data along with high spatial resolution images. [119] There is no universal technique for obtaining high resolution images and robust quantitative information, but progress is being made on this front, and there is potential for future improvements that would greatly impact the field of clinical diagnostics as well as many other scientific fields.

2.5.4 Incorporation of Ambient Ionization Techniques

MSI had been mainly viewed as an invasive process until the development of ambient ionization techniques, such as DESI. Ambient ionization techniques allow for direct analysis in real time and also remove the limitation imposed from requiring vacuum pressure conditions in traditional MSI. [107] DESI and other ambient techniques, which can be performed on untreated histological samples in the open lab environment, have the potential to promote *in situ* analysis in the near future. [31] Miniature mass spectrometers have been developed that are capable of performing DESI-MSI for disease diagnostics and paper-spray ionization MS for therapeutic drug monitoring. [31] These techniques have significant potential to apply real-time diagnostic information in order, for example, to guide surgery. With the current advancements in MSI technology in the areas of miniaturization and ambient sampling, future improvements may involve configuring the instrument in the most efficient and useful way for use by doctors and surgeons in the clinic. The ability to perform *in situ* analysis and the convenience of the portable mass spectrometers suggests the potential role of DESI-MSI and other ambient

ionization techniques in guiding therapy in parallel with standard histological methods in the clinical setting. [31]

2.6 Conclusions

Over the past decade, MSI has obtained increasing attention from biologists and become more routinely employed to map various classes of biomolecules from biological specimen. Its novel applications to biomarker discovery in clinical settings have gained us valuable knowledge regarding disease mechanisms and related reparative processes. Moreover, its usefulness is probing neuropeptide signaling has also been well demonstrated. These studies offer significant insight in clinical and neuroscience studies and hold promise for our continuous search of effective treatment for diseases and imprvoed understanding of neurochemical signaling. Moreover, exciting technical advances are further improving this analytical tool and its clinical applications from various aspects such as sample preparation and instrumentation.

Acknowledgements:

Preparation of this manuscript was supported in part by National Science Foundation (CHE- 0957784) and National Institutes of Health through grant 1R01DK071801. L.L. acknowledges an H. I. Romnes Faculty Fellowship.

2.7 References

1. Craig-Schapiro, R.; Perrin, R. J.; Roe, C. M.; Xiong, C. J.; Carter, D.; Cairns, N. J.; Mintun, M. A.; Peskind, E. R.; Li, G.; Galasko, D. R.; Clark, C. M.; Quinn, J. F.; D'Angelo, G.; Malone, J. P.; Townsend, R. R.; Morris, J. C.; Fagan, A. M.; Holtzman, D. M. (2010) YKL-40: a novel prognostic fluid biomarker for preclinical Alzheimer's disease. *Biol. Psychiat.* 68, 903-912.
2. Hu, S.; Arellano, M.; Boontheung, P.; Wang, J. H.; Zhou, H.; Jiang, J.; Elashoff, D.; Wei, R.; Loo, J. A.; Wong, D. T. (2008) Salivary proteomics for oral cancer biomarker discovery. *Clin. Cancer Res.* 14, 6246-6252.
3. Lowenthal, M. S.; Mehta, A. I.; Frogale, K.; Bandle, R. W.; Araujo, R. P.; Hood, B. L.; Veenstra, T. D.; Conrads, T. P.; Goldsmith, P.; Fishman, D.; Petricoin, E. F.; Liotta, L. A. (2005) Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clin. Chem.* 51, 1933-1945.
4. Teng, P. N.; Hood, B. L.; Sun, M.; Dhir, R.; Conrads, T. P. (2011) Differential proteomic analysis of renal cell carcinoma tissue interstitial fluid. *J. Proteome Res.* 10, 1333-1342.
5. Flint, M. S.; Hood, B. L.; Sun, M.; Stewart, N. A.; Jones-Laughner, J.; Conrads, T. P. (2010) Proteomic analysis of the murine liver in response to a combined exposure to psychological stress and 7,12-Dimethylbenz(a)anthracene. *J. Proteome Res.* 9, 509-520.
6. Ishimura, R.; Ohsako, S.; Kawakami, T.; Sakaue, M.; Aoki, Y.; Tohyama, C. (2002) Altered protein profile and possible hypoxia in the placenta of 2,3,7,8-tetrachlorodibenzo-p-dioxin-exposed rats. *Toxicol. Appl. Pharm.* 185, 197-206.
7. Li, T. A.; Xu, S. L.; Oses-Prieto, J. A.; Putil, S.; Xu, P.; Wang, R. J.; Li, K. H.; Maltby, D. A.; An, L. H.; Burlingame, A. L.; Deng, Z. P.; Wang, Z. Y. (2011) Proteomics analysis reveals post-translational mechanisms for cold-induced metabolic changes in *Arabidopsis*. *Mol. Plant* 4, 361-374.
8. Ruepp, S. U.; Tonge, R. P.; Shaw, J.; Wallis, N.; Pognan, F. (2002) Genomics and proteomics analysis of acetaminophen toxicity in mouse liver. *Toxicol. Sci.* 65, 135-150.
9. Gundlach, A. L.; Burazin, T. C. D.; Larm, J. A. (2001) Distribution, regulation and role of hypothalamic galanin systems: Renewed interest in a pleiotropic peptide family. *Clin. Exp. Pharmacol. P.* 28, 100-105.
10. Audsley, N.; Weaver, R. J. (2009) Neuropeptides associated with the regulation of feeding in insects. *Gen. Comp. Endocr.* 162, 93-104.
11. Jensen, J. (2001) Regulatory peptides and control of food intake in non-mammalian vertebrates. *Comp. Biochem. Phys. A* 128, 471-479.
12. Kieffer, B. L.; Gaveriaux-Ruff, C. (2002) Exploring the opioid system by gene knockout. *Prog. Neurobiol.* 66, 285-306.
13. Levin, E. R.; Hu, R. M.; Rossi, M.; Pickard, M. (1992) Arginine vasopressin stimulates atrial-natriuretic-peptide gene-expression and secretion from diencephalic neurons. *Clin. Res.* 40, A47-A47.
14. Okada, Y.; Tsuda, Y.; Bryant, S. D.; Lazarus, L. H. (2002) Endomorphins and related opioid peptides. *Vitam. Horm.* 65, 257-279.
15. Stefano, G. B.; Fricchione, G. L.; Goumon, Y.; Esch, T. (2005) Pain, immunity, opiate and opioid compounds and health. *Med. Sci. Monit.* 11, Ms47-Ms53.
16. Emanuelsson, O.; Brunak, S.; von Heijne, G.; Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* 2, 953-971.
17. Emanuelsson, O.; Nielsen, H.; Brunak, S.; von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005-1016.

18. Cazares, L. H.; Troyer, D.; Mendrinos, S.; Lance, R. A.; Nyalwidhe, J. O.; Beydoun, H. A.; Clements, M. A.; Drake, R. R.; Semmes, O. J. (2009) Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 2 discriminates cancer from uninvolved prostate tissue. *Clin. Cancer Res.* 15, 5541-51.
19. Deininger, S. O.; Ebert, M. P.; Futterer, A.; Gerhard, M.; Rocken, C. (2008) MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *J. Proteome Res.* 7, 5230-6.
20. Murphy, R. C.; Hankin, J. A.; Barkley, R. M. (2009) Imaging of lipid species by MALDI mass spectrometry. *J. Lipid Res.* 50, S317-S322.
21. Chen, R.; Hui, L.; Sturm, R. M.; Li, L. (2009) Three dimensional mapping of neuropeptides and lipids in crustacean brain by mass spectral imaging. *J. Am. Soc. Mass Spectrom.* 20, 1068-77.
22. DeKeyser, S. S., Kutz-Naber, K.K., Schmidt, J.J., Barrett-Wilt, G.A.and Li, L. (2007) Mass spectral imaging of neuropeptides in crustacean nervous tissue by MALDI TOF/TOF. *J. Proteome Res.* 6, 1782-1791.
23. Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A. (2003) A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* 376, 952-65.
24. Fletcher, J. S.; Lockyer, N. P.; Vickerman, J. C. (2011) Developments in molecular SIMS depth profiling and 3D imaging of biological systems using polyatomic primary ions. *Mass. Spectrom. Rev.* 30, 142-74.
25. Jones, E. A.; Lockyer, N. P.; Vickerman, J. C. (2007) Mass spectral analysis and imaging of tissue by ToF-SIMS - The role of buckminsterfullerene, $C_{60}^{(+)}$, primary ions. *Int. J. Mass Spectrom.* 260, 146-157.
26. Dill, A. L.; Eberlin, L. S.; Costa, A. B.; Zheng, C.; Ifa, D. R.; Cheng, L. A.; Masterson, T. A.; Koch, M. O.; Vitek, O.; Cooks, R. G. (2011) Multivariate statistical identification of human bladder carcinomas using ambient ionization imaging mass spectrometry. *Chem-Eur J* 17, 2897-2902.
27. Gao, Y. H.; Yang, C. H.; Liu, X.; Ma, R. J.; Kong, D. L.; Shi, L. Q. (2012) A multifunctional manocarrier based on nanogated mesoporous silica for enhanced tumor-specific uptake and intracellular delivery. *Macromol. Biosci.* 12, 251-259.
28. Laskin, J.; Heath, B. S.; Roach, P. J.; Cazares, L.; Semmes, O. J. (2012) Tissue imaging using nanospray desorption electrospray ionization mass spectrometry. *Anal. Chem.* 84, 141-148.
29. Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G. (2004) Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science* 306, 471-473.
30. Ifa, D. R.; Wu, C. P.; Ouyang, Z.; Cooks, R. G. (2010) Desorption electrospray ionization and other ambient ionization methods: current progress and preview. *Analyst* 135, 669-681.
31. Cooks, R. G.; Manicke, N. E.; Dill, A. L.; Ifa, D. R.; Eberlin, L. S.; Costa, A. B.; Wang, H.; Huang, G. M.; Zheng, O. Y. (2011) New ionization methods and miniature mass spectrometers for biomedicine: DESI imaging for cancer diagnostics and paper spray ionization for therapeutic drug monitoring. *Faraday Discuss.* 149, 247-267.
32. Jehl, B.; Bauer, R.; Dorge, A.; Rick, R. (1981) The use of propane-isopentane mixtures for rapid freezing of biological specimens. *J. Microsc-Oxford.* 123, 307-309.
33. Schwartz, S. A.; Reyzer, M. L.; Caprioli, R. M. (2003) Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J.Mass Spectrom.* 38, 699-708.

34. Che, F. Y.; Lim, J.; Pan, H.; Biswas, R.; Fricker, L. D. (2005) Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell. Proteomics* 4, 1391-1405.
35. Svensson, M.; Boren, M.; Skold, K.; Falth, M.; Sjogren, B.; Andersson, M.; Svenningsson, P.; Andren, P. E. (2009) Heat stabilization of the tissue proteome: a new technology for improved proteomics. *J. Proteome Res.* 8, 974-981.
36. Crossman, L.; McHugh, N. A.; Hsieh, Y. S.; Korfmacher, W. A.; Chen, J. W. (2006) Investigation of the profiling depth in matrix-assisted laser desorption/ionization imaging mass spectrometry. *Rapid Commun. Mass Spectrom.* 20, 284-290.
37. Verhaert, P. D.; Pinkse, M. W.; Strupat, K.; Conaway, M. C. (2010) Imaging of similar mass neuropeptides in neuronal tissue by enhanced resolution MALDI MS with an ion trap - Orbitrap hybrid instrument. *Methods Mol. Biol.* 656, 433-49.
38. Kaletas, B. K.; van der Wiel, I. M.; Stauber, J.; Dekker, L. J.; Guzel, C.; Kros, J. M.; Luider, T. M.; Heeren, R. M. A. (2009) Sample preparation issues for tissue imaging by imaging MS. *Proteomics* 9, 2622-2633.
39. Lemaire, R.; Wisztorski, M.; Desmons, A.; Tabet, J. C.; Day, R.; Salzet, M.; Fournier, I. (2006) MALDI-MS direct tissue analysis of proteins: improving signal sensitivity using organic treatments. *Anal. Chem.* 78, 7145-7153.
40. Seeley, E. H.; Oppenheimer, S. R.; Mi, D.; Chaurand, P.; Caprioli, R. M. (2008) Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *J. Am. Soc. Mass Spectrom.* 19, 1069-1077.
41. van Hove, E. R. A.; Smith, D. F.; Fornai, L.; Glunde, K.; Heeren, R. M. A. (2011) An alternative paper based tissue washing method for mass spectrometry imaging: localized washing and fragile tissue analysis. *J. Am. Soc. Mass Spectrom.* 22, 1885-1890.
42. Scicchitano, M. S.; Dalmas, D. A.; Boyce, R. W.; Thomas, H. C.; Frazier, K. S. (2009) Protein extraction of formalin-fixed, paraffin-embedded tissue enables robust proteomic profiles by mass spectrometry. *J. Histochem. Cytochem.* 57, 849-860.
43. Casadonte, R.; Caprioli, R. M. (2011) Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. *Nat. Protoc.* 6, 1695-1709.
44. EmmertBuck, M. R.; Bonner, R. F.; Smith, P. D.; Chuaqui, R. F.; Zhuang, Z. P.; Goldstein, S. R.; Weiss, R. A.; Liotta, L. A. (1996) Laser capture microdissection. *Science* 274, 998-1001.
45. Chaurand, P.; Fouchecourt, S.; DaGue, B. B.; Xu, B. G. J.; Reyzer, M. L.; Orgebin-Crist, M. C.; Caprioli, R. M. (2003) Profiling and imaging proteins in the mouse epididymis by imaging mass spectrometry. *Proteomics* 3, 2221-2239.
46. Wang, Z.; Han, J.; Schey, K. L. (2008) Spatial differences in an integral membrane proteome detected in laser capture microdissected samples. *J. Proteome Res.* 7, 2696-2702.
47. Xu, B. G. J.; Caprioli, R. M. (2002) Direct analysis of laser capture microdissected cells by MALDI mass spectrometry. *J. Am. Soc. Mass Spectrom.* 13, 1292-1297.
48. Xu, B. J.; Shyr, Y.; Liang, X. B.; Ma, L. J.; Donnert, E. M.; Roberts, J. D.; Zhang, X. Q.; Kon, V.; Brown, N. J.; Caprioli, R. M.; Fogo, A. B. (2005) Proteomic patterns and prediction of glomerulosclerosis and its mechanisms. *J. Am. Soc. Nephrol.* 16, 2967-2975.
49. Groseclose, M. R.; Andersson, M.; Hardesty, W. M.; Caprioli, R. M. (2007) Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry. *J. Mass Spectrom.* 42, 254-262.
50. Ronci, M.; Bonanno, E.; Colantoni, A.; Pieroni, L.; Di Ilio, C.; Spagnoli, L. G.; Federici, G.; Urbani, A. (2008) Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: Application to MALDI-TOF imaging MS investigations. *Proteomics* 8, 3702-3714.

51. Djidja, M. C.; Francese, S.; Loadman, P. M.; Sutton, C. W.; Scriven, P.; Claude, E.; Snel, M. F.; Franck, J.; Salzet, M.; Clench, M. R. (2009) Detergent addition to tryptic digests and ion mobility separation prior to MS/MS improves peptide yield and protein identification for in situ proteomic investigation of frozen and formalin-fixed paraffin-embedded adenocarcinoma tissue sections. *Proteomics* 9, 2750-2763.
52. Stauber, J.; MacAleese, L.; Franck, J.; Claude, E.; Snel, M.; Kaletas, B. K.; Wiel, I. M.; Wisztorski, M.; Fournier, I.; Heeren, R. M. (2010) On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry. *J. Am. Soc. Mass Spectrom.* 21, 338-47.
53. Boggio, K. J.; Obasuyi, E.; Sugino, K.; Nelson, S. B.; Agar, N. Y. R.; Agar, J. N. (2011) Recent advances in single-cell MALDI mass spectrometry imaging and potential clinical impact. *Expert Rev. Proteomics* 8, 591-604.
54. McDonnell, L. A.; Corthals, G. L.; Willems, S. M.; van Remoortere, A.; van Zeijl, R. J. M.; Deelder, A. M. (2010) Peptide and protein imaging mass spectrometry in cancer research. *J. Proteomics* 73, 1921-1944.
55. Tu, T.; Sauter, A. D.; Sauter, A. D.; Gross, M. L. (2008) Improving the signal intensity and sensitivity of MALDI mass spectrometry by using nanoliter spots deposited by induction-based fluidics. *J. Am. Soc. Mass Spectrom.* 19, 1086-1090.
56. Zimmerman, T. A.; Monroe, E. B.; Sweedler, J. V. (2008) Adapting the stretched sample method from tissue profiling to imaging. *Proteomics* 8, 3809-3815.
57. Zimmerman, T. A.; Rubakhin, S. S.; Sweedler, J. V. (2011) MALDI Mass Spectrometry Imaging of Neuronal Cell Cultures. *J. Am. Soc. Mass Spectrom.* 22, 828-836.
58. Jurchen, J. C.; Rubakhin, S. S.; Sweedler, J. V. (2005) MALDI-MS imaging of features smaller than the size of the laser beam. *J. Am. Soc. Mass Spectrom.* 16, 1654-1659.
59. Koestler, M.; Kirsch, D.; Hester, A.; Leisner, A.; Guenther, S.; Spengler, B. (2008) A high-resolution scanning microprobe matrix-assisted laser desorption/ionization ion source for imaging analysis on an ion trap/Fourier transform ion cyclotron resonance mass spectrometer. *Rapid Commun. Mass Spectrom.* 22, 3275-3285.
60. Guenther, S.; Rompp, A.; Kummer, W.; Spengler, B. (2011) AP-MALDI imaging of neuropeptides in mouse pituitary gland with 5 μm spatial resolution and high mass accuracy. *Int. J. Mass Spectrom.* 305, 228-237.
61. Schober, Y.; Guenther, S.; Spengler, B.; Rompp, A. (2012) Single cell matrix-assisted laser desorption/ionization mass spectrometry imaging. *Anal. Chem.* 84, 6293-7.
62. Chaurand, P.; Schriver, K. E.; Caprioli, R. M. (2007) Instrument design and characterization for high resolution MALDI-MS imaging of tissue sections. *J. Mass Spectrom.* 42, 476-489.
63. Schafer, R. (2009) ultrafleXtreme: redefining MALDI mass spectrometry performance. *LC GC N. Am.*, 14-15.
64. Luxembourg, S. L.; Mize, T. H.; McDonnell, L. A.; Heeren, R. M. A. (2004) High-spatial resolution mass spectrometric imaging of peptide and protein distributions on a surface. *Anal. Chem.* 76, 5339-5344.
65. Rubakhin, S. S.; Greenough, W. T.; Sweedler, J. V. (2003) Spatial profiling with MALDI MS: distribution of neuropeptides within single neurons. *Anal. Chem.* 75, 5374-5380.
66. Schwamborn, K.; Caprioli, R. M. Molecular imaging by mass spectrometry--looking beyond classical histology. *Nat Rev Cancer* 10, 639-46.
67. Lemaire, R.; Tabet, J. C.; Ducoroy, P.; Hendra, J. B.; Salzet, M.; Fournier, I. (2006) Solid ionic matrixes for direct tissue analysis and MALDI Imaging. *Anal. Chem.* 78, 809-819.
68. Thomas, A.; Charbonneau, J. L.; Fournaise, E.; Chaurand, P. (2012) Sublimation of new matrix candidates for high spatial resolution imaging mass spectrometry of lipids: enhanced

information in both positive and negative polarities after 1,5-diaminonapthalene deposition. *Anal. Chem.* 84, 2048-2054.
69. Deutskens, F.; Yang, J. H.; Caprioli, R. M. (2011) High spatial resolution imaging mass spectrometry and classical histology on a single tissue section. *J. Mass Spectrom.* 46, 568-571.
70. Cohen, M. Z. (1987) A historical overview of the phenomenologic movement. *Image J. Nurs. Sch.* 19, 31-4.
71. Vestal, M. L.; Campbell, J. M. (2005) Tandem time-of-flight mass spectrometry. *Methods Enzymol.* 402, 79-108.
72. Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H. (2008) Ion mobility-mass spectrometry. *J. Mass Spectrom.* 43, 1-22.
73. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom. Rev.* 17, 1-35.
74. Kutz, K. K.; Schmidt, J. J.; Li, L. J. (2004) *In situ* tissue analysis of neuropeptides by MALDI FTMS in-cell accumulation. *Anal. Chem.* 76, 5630-5640.
75. Fuchser, J.; Cornett, S.; Becker, M., High resolution molecular imaging of pharmaceuticals at theraputic levels. In *Bruker Daltonik GmbH*, 2008.
76. Makarov, A. (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* 72, 1156-1162.
77. Hu, Q. Z.; Noll, R. J.; Li, H. Y.; Makarov, A.; Hardman, M.; Cooks, R. G. (2005) The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* 40, 430-443.
78. Klerk, L. A.; Altelaar, A. F. M.; Froesch, M.; McDonnell, L. A.; Heeren, R. M. A. (2009) Fast and automated large-area imaging MALDI mass spectrometry in microprobe and microscope mode. *International Journal of Mass Spectrometry* 285, 19-25.
79. Altelaar, A. F. M.; Taban, I. M.; McDonnell, L. A.; Verhaert, P. D. E. M.; de Lange, R. P. J.; Adan, R. A. H.; Mooi, W. J.; Heeren, R. M. A.; Piersma, S. R. (2007) High-resolution MALDI imaging mass spectrometry allows localization of peptide distributions at cellular length scales in pituitary tissue sections. *Int. J. Mass Spectrom.* 260, 203-211.
80. Seeley, E. H.; Caprioli, R. M. (2008) Imaging mass spectrometry: towards clinical diagnostics. *Proteomics Clin. Appl.* 2, 1435-1443.
81. Cazares, L. H.; Troyer, D. A.; Wang, B. H.; Drake, R. R.; Semmes, O. J. (2011) MALDI tissue imaging: from biomarker discovery to clinical applications. *Anal. Bioanal. Chem.* 401, 17-27.
82. Seeley, E. H.; Caprioli, R. M. (2011) MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives. *Trends Biotechnol.* 29, 136-143.
83. Rauser, S.; Deininger, S. O.; Suckau, D.; Hofler, H.; Walch, A. (2010) Approaching MALDI molecular imaging for clinical proteomic research: current state and fields of application. *Expert Rev. Proteomics* 7, 927-941.
84. Takai, N.; Tanaka, Y.; Inazawa, K.; Saji, H. (2012) Quantitative analysis of pharmaceutical drug distribution in multiple organs by imaging mass spectrometry. *Rapid Commun. Mass Spectrom.* 26, 1549-1556.
85. Drexler, D. M.; Tannehill-Gregg, S. H.; Wang, L. F.; Brock, B. J. (2011) Utility of quantitative whole-body autoradiography (QWBA) and imaging mass spectrometry (IMS) by matrix-assisted laser desorption/ionization (MALDI) in the assessment of ocular distribution of drugs. *J. Pharmacol. Toxicol. Methods* 63, 205-208.
86. Wiseman, J. M.; Ifa, D. R.; Zhu, Y. X.; Kissinger, C. B.; Manicke, N. E.; Kissinger, P. T.; Cooks, R. G. (2008) Desorption electrospray ionization mass spectrometry: Imaging drugs and metabolites in tissues. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18120-18125.

87. Markowitz, L.; Anis, S. I.; Linehan, S. T. (2010) QWBA - current methodologies and future advancements positioning of image. *J. Labelled Comp. Radiopharm.* 53, 330-331.
88. Cornett, D. S.; Frappier, S. L.; Caprioli, R. M. (2008) MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. *Anal. Chem.* 80, 5648-5653.
89. Wittig, A.; Arlinghaus, H. F.; Kriegeskotte, C.; Moss, R. L.; Appelman, K.; Schmid, K. W.; Sauerwein, W. A. G. (2008) Laser postionization secondary neutral mass spectrometry in tissue: a powerful tool for elemental and molecular imaging in the development of targeted drugs. *Mol. Cancer Ther.* 7, 1763-1771.
90. Nilsson, A.; Fehniger, T. E.; Gustavsson, L.; Andersson, M.; Kenne, K.; Marko-Varga, G.; Andren, P. E. (2010) Fine mapping the spatial distribution and concentration of unlabeled drugs within tissue micro-compartments using imaging mass spectrometry. *PLoS One* 5.
91. Young, S. P.; Piraud, M.; Goldstein, J. L.; Zhang, H. Y.; Rehder, C.; Laforet, P.; Kishnani, P. S.; Millington, D. S.; Bashir, M. R.; Bali, D. S. (2012) Assessing disease severity in Pompe disease: The roles of a urinary glucose tetrasaccharide biomarker and imaging techniques. *Am. J. Med. Genet. C Semin. Med. Genet.* 160C, 50-58.
92. Meding, S.; Nitsche, U.; Balluff, B.; Elsner, M.; Rauser, S.; Schone, C.; Nipp, M.; Maak, M.; Feith, M.; Ebert, M. P.; Friess, H.; Langer, R.; Hofler, H.; Zitzelsberger, H.; Rosenberg, R.; Walch, A. (2012) Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. *J.Proteome Res.* 11, 1996-2003.
93. Postle, A. D. (2012) Lipidomics. *Curr. Opin. Clin. Nutr. Metab. Care* 15, 127-133.
94. Walsh, N.; Crotty, K.; Palmer, A.; McCarthy, S. (1998) Spitz nevus versus spitzoid malignant melanoma: an evaluation of the current distinguishing histopathologic criteria. *Hum. Pathol.* 29, 1105-12.
95. Lazova, R.; Seeley, E. H.; Keenan, M.; Gueorguieva, R.; Caprioli, R. M. (2012) Imaging mass spectrometry-a new and promising method to differentiate Spitz Nevi from spitzoid malignant melanomas. *Am. J. Dermatopathol.* 34, 82-90.
96. Kim, D. W.; Huamani, J.; Reyzer, M. L.; Mi, D.; Caprioli, R. M.; Hallahan, D. E. (2006) Imaging mass spectrometry to map distribution of radiation enhancing vasculature targeted drug and protein biomarkers of response to therapy in prostate cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 66, S554-S554.
97. Thomas, A.; Lenglet, S.; Chaurand, P.; Deglon, J.; Mangin, P.; Mach, F.; Steffens, S.; Wolfender, J. L.; Staub, C. (2011) Mass spectrometry for the evaluation of cardiovascular diseases based on proteomics and lipidomics. *Thromb. Haemost.* 106, 20-33.
98. Hanada, M.; Sugiura, Y.; Shinjo, R.; Masaki, N.; Imagama, S.; Ishiguro, N.; Matsuyama, Y.; Setou, M. (2012) Spatiotemporal alteration of phospholipids and prostaglandins in a rat model of spinal cord injury. *Anal. Bioanal.Chem.* 403, 1873-1884.
99. Weaver, C. M.; Liebman, M. (2002) Biomarkers of bone health appropriate for evaluating functional foods designed to reduce risk of osteoporosis. *Br. J. Nutr.* 88, S225-S232.
100. Gamez-Pozo, A.; Sanchez-Navarro, I.; Calvo, E.; Agullo-Ortuno, M. T.; Lopez-Vacas, R.; Diaz, E.; Camafeita, E.; Nistal, M.; Madero, R.; Espinosa, E.; Lopez, J. A.; Vara, J. A. F. (2012) PTRF/Cavin-1 and MIF proteins are identified as non-small cell lung cancer biomarkers by label-free proteomics. *PLoS One* 7.
101. Tanaka, H.; Zaima, N.; Yamamoto, N.; Suzuki, M.; Mano, Y.; Konno, H.; Unno, N.; Setou, M. (2011) Distribution of phospholipid molecular species in autogenous access grafts for hemodialysis analyzed using imaging mass spectrometry. *Anal. Bioanal. Chem.* 400, 1873-1880.
102. Zoehrer, R.; Perilli, E.; Kuliwaba, J. S.; Shapter, J. G.; Fazzalari, N. L.; Voelcker, N. H. (2012) Human bone material characterization: integrated imaging surface investigation of male fragility fractures. *Osteoporos. Int.* 23, 1297-1309.

103. Klerk, L. A.; Dankers, P. Y. W.; Popa, E. R.; Bosman, A. W.; Sanders, M. E.; Reedquist, K. A.; Heeren, R. M. A. (2010) TOF-secondary ion mass spectrometry imaging of polymeric scaffolds with surrounding tissue after *in vivo* implantation. *Anal. Chem.* 82, 4337-4343.
104. Ooya, T.; Mori, H.; Terano, M.; Yui, N. (1995) Synthesis of a biodegradable polymeric supramolecular assembly for drug delivery. *Macromol. Rapid Commun.* 16, 259-263.
105. Lin, C. H.; Mardini, S.; Lin, Y. T.; Yeh, J. T.; Wei, F. C.; Chen, H. C. (2004) Sixty-five clinical cases of free tissue transfer using long arteriovenous fistulas or vein grafts. *J. Trauma.* 56, 1107-1117.
106. Verhaert, P. D.; Conaway, M. C. P.; Pekar, T. M.; Miller, K. (2007) Neuropeptide imaging on an LTQ with vMALDI source: the complete 'all-in-one' peptidome analysis. *Int. J. Mass Spectrom.* 260, 177-184.
107. Ye, H.; Greer, T.; Li, L. J. (2011) From pixel to voxel: a deeper view of biological tissue by 3D mass spectral imaging. *Bioanalysis* 3, 313-332.
108. Altelaar, A. F. M.; van Minnen, J.; Jimenez, C. R.; Heeren, R. M. A.; Piersma, S. R. (2005) Direct molecular Imaging of *Lymnaea stagnalis* nervous tissue at subcellular spatial resolution by mass spectrometry. *Anal. Chem.* 77, 735-741.
109. Monroe, E. B.; Annangudi, S. R.; Hatcher, N. G.; Gutstein, H. B.; Rubakhin, S. S.; Sweedler, J. V. (2008) SIMS and MALDI MS imaging of the spinal cord. *Proteomics* 8, 3746-3754.
110. McDonnell, L. A.; Piersma, S. R.; Altelaar, A. F. M.; Mize, T. H.; Luxembourg, S. L.; Verhaert, P. D. E. M.; van Minnen, J.; Heeren, R. M. A. (2005) Subcellular imaging mass spectrometry of brain tissue. *J. Mass Spectrom.* 40, 160-168.
111. Jimenez, C. R.; Li, K. W.; Dreisewerd, K.; Spijker, S.; Kingston, R.; Bateman, R. H.; Burlingame, A. L.; Smit, A. B.; van Minnen, J.; Geraerts, W. P. M. (1998) Direct mass spectrometric peptide profiling and sequencing of single neurons reveals differential peptide patterns in a small neuronal network. *Biochemistry* 37, 2070-2076.
112. Li, L. J.; Garden, R. W.; Romanova, E. V.; Sweedler, J. V. (1999) *In situ* sequencing of peptides from biological tissues and single cells using MALDI-PSD/CID analysis. *Anal. Chem.* 71, 5451-5458.
113. Yasuda, A.; Yasuda-Kamatani, Y.; Nozaki, M.; Nakajima, T. (2004) Identification of GYRKPPFNGSIFamide (crustacean- SIFamide) in the crayfish *Procambarus clarkii* by topological mass spectrometry analysis. *Gen. Comp. Endocrinol.* 135, 391-400.
114. Neupert, S.; Predel, R.; Russell, W. K.; Davies, R.; Pietrantonio, P. V.; Nachman, R. J. (2005) Identification of tick periviscerokinin, the first neurohormone of Ixodidae: single cell analysis by means of MALDI-TOF/TOF mass spectrometry. *Biochem. Bioph. Res. Commun.* 338, 1860-1864.
115. Jarecki, J. L.; Andersen, K.; Konop, C. J.; Knickelbine, J. J.; Vestling, M. M.; Stretton, A. O. (2010) Mapping neuropeptide expression by mass spectrometry in single dissected identified neurons from the dorsal ganglion of the nematode *Ascaris suum*. *ACS Chem. Neurosci.* 1, 505-519.
116. Rubakhin, S. S.; Garden, R. W.; Fuller, R. R.; Sweedler, J. V. (2000) Measuring the peptides in individual organelles with mass spectrometry. *Nat. Biotechnol.* 18, 172-175.
117. Jones, E. A.; van Zeijl, R. J. M.; Andren, P. E.; Deelder, A. M.; Wolters, L.; McDonnell, L. A. (2012) High speed data processing for imaging MS-based molecular histology using graphical processing units. *J. Am. Soc. Mass Spectrom.* 23, 745-752.
118. Ergin, B.; Meding, S.; Langer, R.; Kap, M.; Viertler, C.; Schott, C.; Ferch, U.; Riegman, P.; Zatloukal, K.; Walch, A.; Becker, K. F. (2010) Proteomic analysis of PAXgene-fixed tissues. *J. Proteome Res.* 9, 5188-5196.

119. Wang, H. A. O.; Grolimund, D.; Van Loon, L. R.; Barmettler, K.; Borca, C. N.; Aeschimann, B.; Gunther, D. (2011) Quantitative chemical imaging of element diffusion into heterogeneous media using laser ablation inductively coupled plasma mass spectrometry, synchrotron micro-X-ray fluorescence, and extended X-ray absorption fine structure spectroscopy. *Anal. Chem.* 83, 6259-6266.

Figures

Figure 1. A schematic representation of the clinical applications of MSI. MSI spectra arising from the diseased regions are recorded, and the molecular MS images are reconstructed. The molecules like lipids or proteins that differentiate the diseased regions from the normal ones are potential biomarkers of the disease. Moreover, the metabolites that changed corresponding to drug treatment are also investigated based on this workflow.

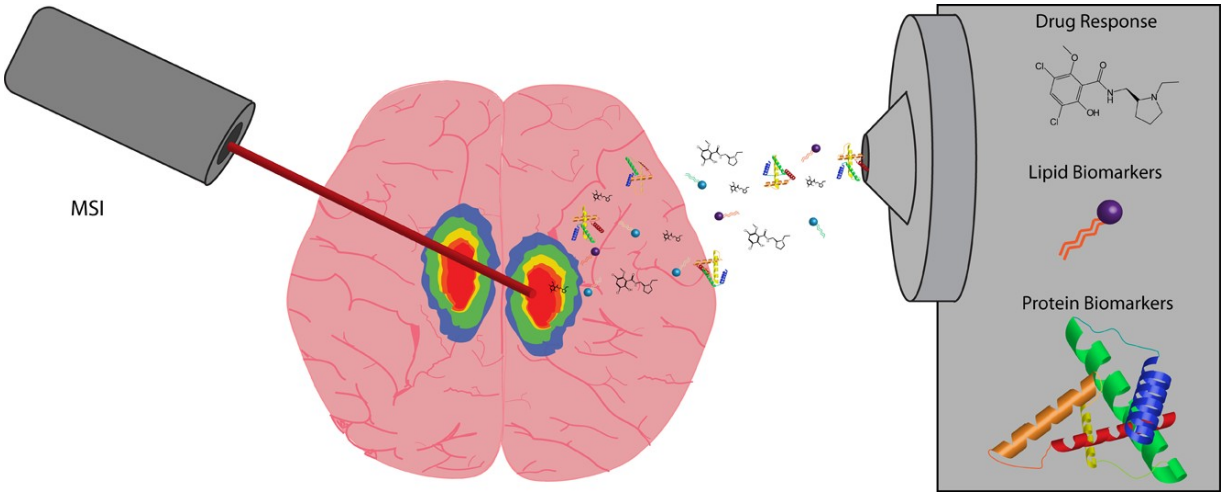


Figure 2. DHA-containing PCs exhibited impact site-specific irreversible reductions from 1 day to 8 weeks post-SCI. The MSI results for DHA-containing PCs, i.e.: PC(diacyl-16:0/22:6) and PC (diacyl-18:0/22:6), are detailed. In particular, the 24 ion images for each DHA-PC from sections of normal (sham-operated) and SCI-treated samples at five different time points are shown. The distribution of DHA-PCs was unaltered at 12 h post-SCI in comparison with the control. The primary reduction was observed around the central canal and gray commissure region to a severe extent at 1 d post-SCI, whereas the decreases at the anterior and posterior horns were moderate (arrowheads). However, at 1 week post-SCI, DHA-PCs were also lost from these tissue regions, and these reductions evolved at later time points and the DHA-PCs had almost disappeared by 8 weeks post-SCI. [98] Reprinted with permission.

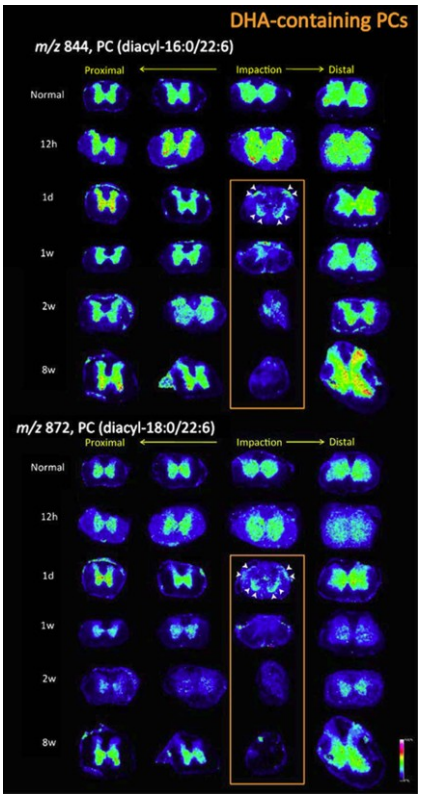


Figure 3. MS images of lipids from the kidney tissue entered into the polymer capsule. (a) Large area image of the hydrogel implant under the renal capsule of a rat, 15 days after implantation. Various localizations are indicated, based on PCA+VARIMAX results (see spectra in (b)) and in the overlay with an optical microscope image. The presence of lipids inside the polymer area shows cellular infiltration in the drug delivery carrier. Some smearing artifact is visible at the bottom region of the polymer. The respective spectral results are given in (b). The first and second PCs gave non-informative distributions. PC 3 shows signal for various lipids, including Diacyl glycerols (*m/z* 550-620), PC 4 shows cholesterol ($(M\text{-}OH)^+$ at *m/z* 369.4 and M^+ at *m/z* 385.4), PC 5 shows silicone contamination ($C_7H_{21}O_2Si_3^+$ at *m/z* 221.1, further identified from low-mass peaks in the corresponding region), PC 7 shows the polymer distribution, readily recognize from the *m/z* 44 spacing between the peaks, which exactly corresponds to the mass of one PEG unit. (c) Shows the PEG distribution in detail with characteristic 16 Da (K^+ and Na^+ difference or O loss) and 14 Da (CH_2 loss) intervals. The change from 0.2 to 0.3 values is due to binning down and, thus, rounding to 0.1 Da. [103] Reprinted with permission from Klerk LA, Dankers PYW, Popa ER, et al. TOF-secondary ion mass spectrometry imaging of polymeric scaffolds with surrounding tissue after in vivo implantation. Reprinted with permission.

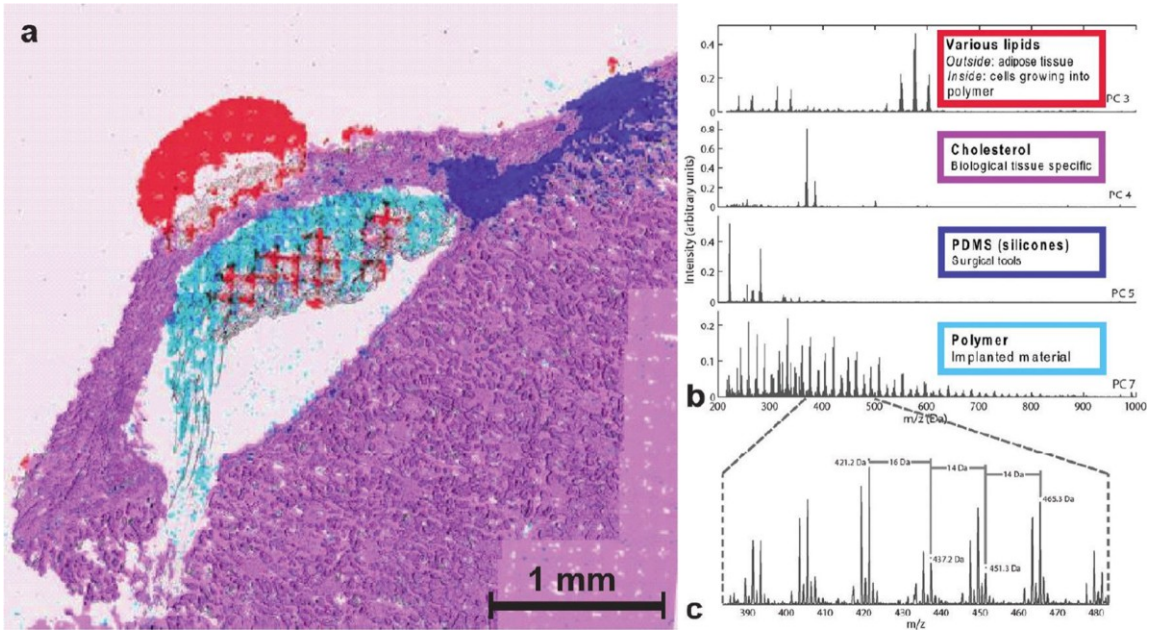

a
1 mm
Various lipids
Outside: adipose tissue
Inside: cells growing into polymer
PC 3
Cholesterol
Biological tissue specific
PC 4
PDMS (silicones)
Surgical tools
PC 5
Polymer
Implanted material
PC 7
Intensity (arbitrary units)
b
m/z (Da)
c
m/z
421.2 Da
16 Da
14 Da
14 Da
465.3 Da
437.2 Da
451.3 Da

Figure 4. Imaging mass spectrometry and optical images of frozen section (8 μm) of the control vein (CV), arteriovenous fistula (AVF), and peripheral artery occlusive disease (PAD) samples. Scale bar=200 μm. Ion intensity was normalized by total ion current. MSI of human AVF revealed the characteristic distribution of phospholipid molecules in the intima and media compared with that in the CV. [101] Reprinted with permission from Tanaka H, Zaima N, Yamamoto N, et al. Distribution of phospholipid molecular species in autogenous access grafts for hemodialysis analyzed using imaging mass spectrometry. Anal Bioanal Chem 2011; 400:1873-1880.

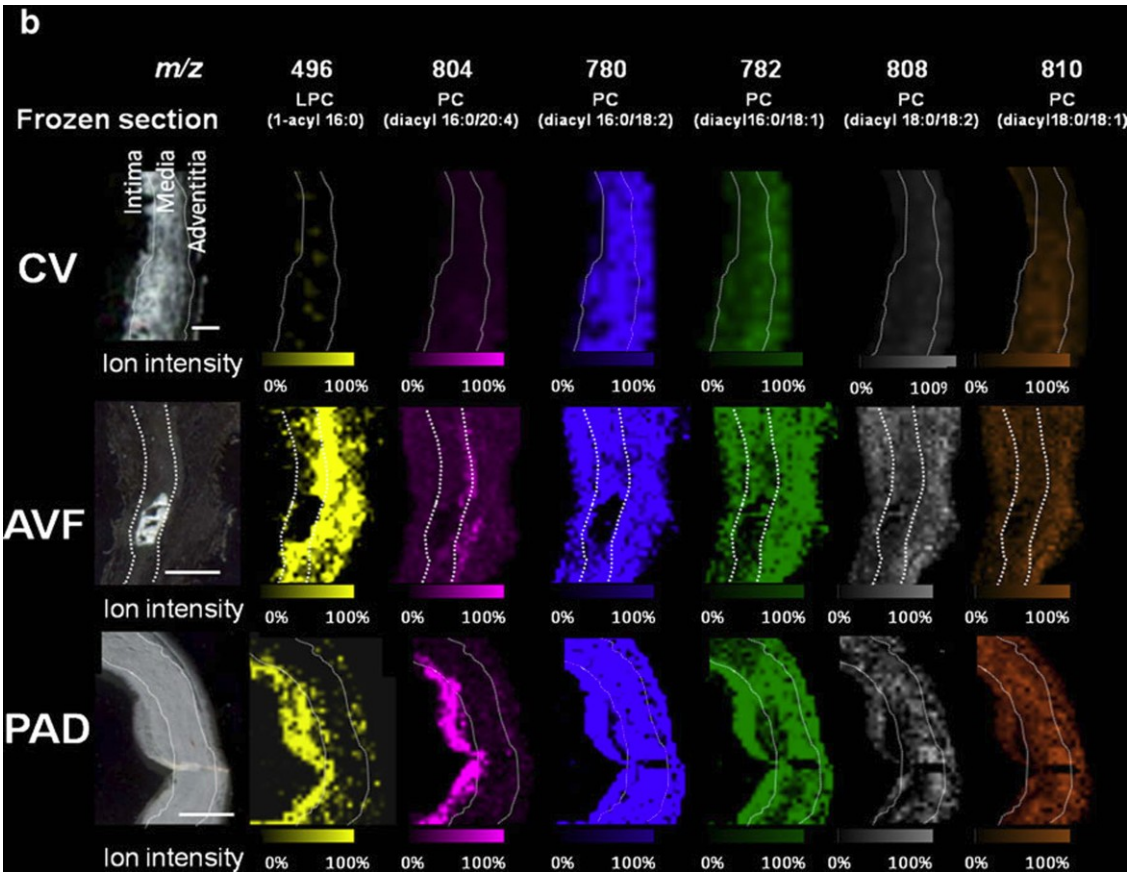


Figure 5. 3D MSI images obtained from serial crab brain sections. (a) 3D reconstructed ion image of CabTRP 1a by MALDI MSI prepared using regular matrix coating. (b) 3D reconstructed ion image of lipid PC 38:6 by MALDI MSI prepared using dry matrix spraying technique for lipid detection. [21] Adapted with permission.

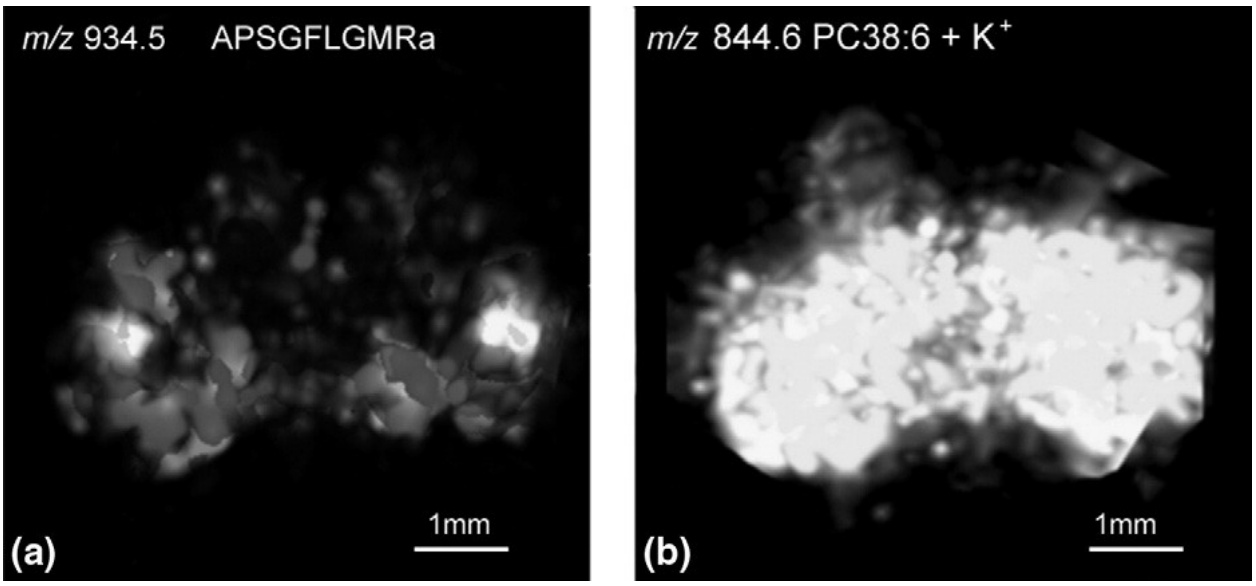
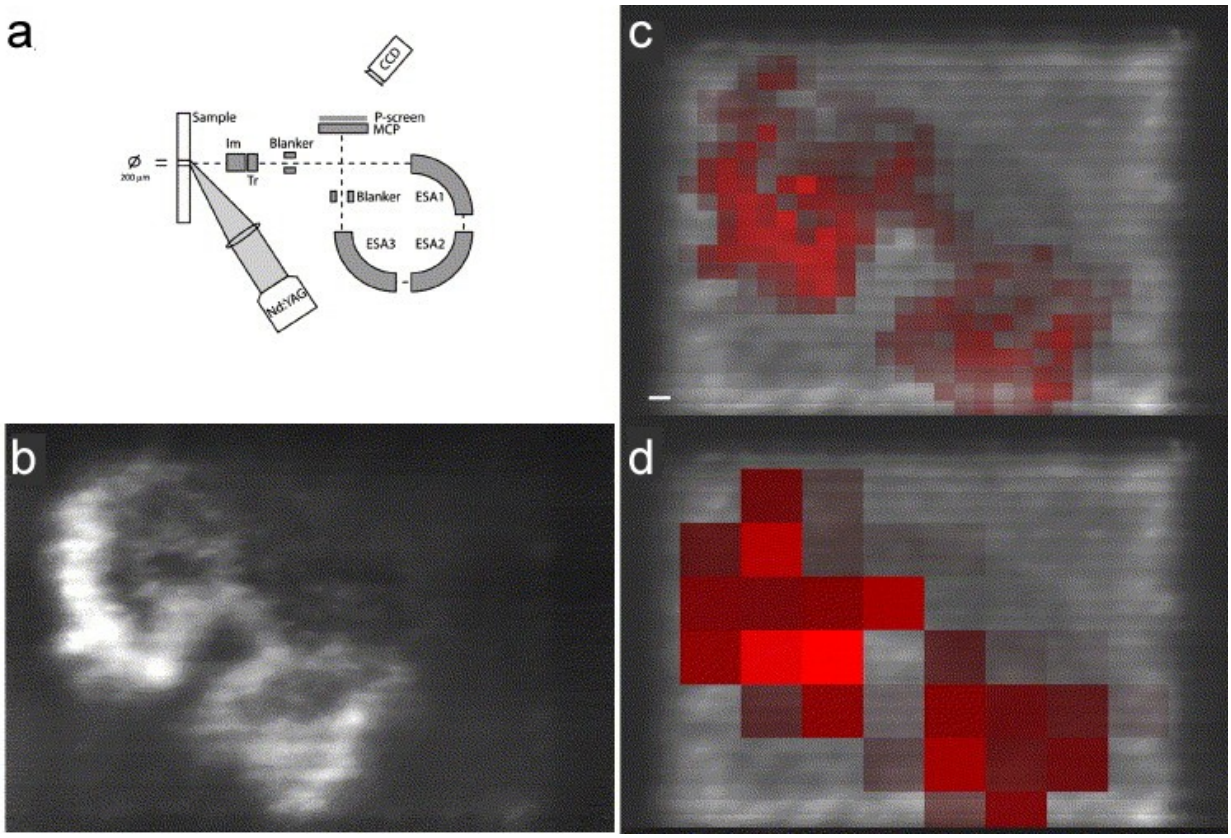


Figure 6. Instrumentation of the stigmatic mass microscope and a comparison with microprobe MSI. (a) Schematic representation of the Physical Electronics Trift II mass spectrometer. (b) The stigmatic ion image of the neuropeptide α-MSH (diacetylated) is compared to the distribution of the diacetylated α-MSH in microprobe imaging, depicted with a resampled pixel size of (c) 80 μm × 80 μm (red) and (d) 250 μm × 250 μm (red) and overlaid on the stigmatic TIC image. [79] Scale bar is 100 μm. Adapted with permission.



Chapter 3

Visualizing Neurotransmitters and Metabolites in the Central Nervous Systems by High Resolution and High Accuracy Mass Spectrometric Imaging

Adapted from **Hui Ye**, Jingxin Wang, Tyler Greer, Kerstin Strupat, Lingjun Li. Visualizing neurotransmitters and metabolites in central nervous systems by high resolution and accuracy mass spectrometric imaging, *ACS Chem. Neurosci*., in press.

Abstract

The spatial localization and molecular distribution of metabolites and neurotransmitters within biological organisms is of tremendous interest to neuroscientists. In comparison to conventional imaging techniques like immunohistochemistry, matrix-assisted laser desorption/ionization (MALDI) mass spectrometric imaging (MSI) has demonstrated its unique advantage by directly localizing the distribution of a wide range of biomolecules simultaneously from a tissue specimen. Although MALDI-MSI of metabolites and neurotransmitters is hindered by numerous matrix-derived peaks, high-resolution and high-accuracy mass spectrometers (HRMS) allow differentiation of endogenous analytes from matrix peaks, unambiguously obtaining biomolecular distributions. In this study, we presented MSI of metabolites and neurotransmitters in rodent and crustacean central nervous systems acquired on HRMS. Results were compared with those obtained from a medium-resolution mass spectrometer (MRMS), tandem time-of-flight instrument, to demonstrate the power and unique advantages of HRMSI and reveal how this new tool would benefit molecular imaging applications in neuroscience.

3.1 Introduction

Metabolites are the end products of regulatory processes in cells and could directly reveal the physiological states of cells. Therefore, large scale analyses of metabolites have been used to discriminate cell populations of different physiological conditions, such as diseased *vs.* healthy,[1] young *vs.* old[2] and normal *vs.* stressed,[3] to potentially identify specific biomarkers.[4] Currently, changes in metabolic profiles have been clinically utilized to characterize ~200 inherited enzymatic disorders.[4] Neurotransmitters (NTs) are a class of endogenous small molecules that transmit intercellular signals from a neuron to a target cell at chemical synapses.[5] Neurotransmitters include amino acids such as serine and glycine, monoamines like dopamine, and other classical NTs like acetylcholine (ACh).[6] Abnormal neurotransmitter expression levels have also been associated with various diseases. For example, decreased synthesis of ACh has been proposed as the cause of Alzheimer's disease (AD),[7] and the loss of dopaminergic neurons is implicated in Parkinson's disease.[8] Studies of metabolites and NTs could not only deepen our understanding of related pathogenic mechanisms but could also lead to potential therapeutic treatments.

Examination of metabolites and NTs from cells and tissues of homogenized extracts can be performed via well-established gas chromatography (GC) or liquid chromatography (LC) mass spectrometry (MS) methods.[9] Unfortunately, these techniques do not account for the spatial distributions of endogenous metabolites and

NTs. Therefore, a complementary approach that directly characterizes and localizes metabolites and NTs from tissues is needed.

Previously, metabolite and NT imaging was usually achieved by imaging antibodies against the targets' conjugates or against the enzymes involved in synthesis or degradation of the low molecular weight (MW) targets. One example is immunohistochemical (IHC) imaging of ACh by localizing the ACh-synthesizing enzyme choline acetyltransferase ChAT.[10] Although high-resolution images of low MW targets can be provided via IHC, several drawbacks exist: sample preparation is tedious, a limited number of targets can be imaged simultaneously, highly specific antibodies are difficult to obtain, and discovery of unknown molecules is impossible because molecule-specific antibodies are used.[11] In contrast, mass spectrometric imaging (MSI) examines the bio-distribution of multiple molecules without prior knowledge of their identities through direct detection from tissue samples.[11, 12] Its unique capability enables mapping of low-MW species with minimal sample preparation and high specificity, revolutionizing our understanding of biological processes by visualizing the dynamics of metabolites and NTs.[13-17]

MALDI is a sensitive, soft ionization technique capable of ionizing a wide range of molecules,[11, 18, 19] and its recent application to drug and metabolite imaging has gained enormous attention.[19-22] However, MALDI-MSI of metabolites and NTs present at physiological concentrations in tissues is often complicated by numerous high-abundance matrix-derived peaks. High selectivity is therefore essential to successfully implement

MSI in the low-MW region. One effective approach to increase selectivity is to perform tandem MS on selected low-MW metabolites and NTs.[23] Instead of monitoring intact molecules in MS mode, the distribution of target molecules is measured in tandem MS mode by generating ion density maps of fragment ions exclusively produced by the targeted precursor ions.[23] However, this method requires one tandem MSI analysis for each target, meaning N times more sample material and acquisition time to map N ions, whereas HRMSI solves this problem without compromising throughput and sensitivity.[17, 24] The high resolution, accurate mass measurements provided by HRMS discriminate low-MW chemical noise from endogenous metabolites and NTs, thereby verifying MS assignments with high confidence and increasing the information obtained from a biologically complex specimen. [16, 17, 24, 25]

In this study, we developed HRMSI applications to measure the localizations of metabolites and NTs in rodent and crustacean central nervous systems (CNS) using a MALDI LTQ Orbitrap XL mass spectrometer. A comparison between results from complex CNS specimens obtained on a HRMS LTQ Orbitrap XL and a MRMS, tandem time-of-flight (TOF/TOF), was also made. Although the MALDI-TOF/TOF (error < 10 ppm and resolution of up to 20,000 for the peptide range) is the most widely used instrument for MSI applications, the high mass-accuracy (< 3 ppm) and mass-resolution (of up to 130,000 @ m/z 400) of HRMS enables unambiguous, confident identification of a number of metabolites and neurotransmitters directly from biologically complex tissue samples.

3.2 Experimental Section

Animal dissection

Animal experiments were performed according to institutional guidelines (UW-Madison IACUC) and can be found in detail in **Supplemental Information (SI)**.

Sample Preparation

The CNS tissue was placed in gelatin solution, flash-frozen and sectioned on a cryostat. The resulting sections were then dehydrated and applied with matrix DHB. The details can be found in **SI**.

MALDI LTQ Orbitrap XL

The MALDI LTQ Orbitrap XL instrument (Thermo Scientific, Bremen, Germany) equipped with a commercial 60 Hz N_2 laser at 337 nm (Lasertechnik Berlin GmbH, Berlin, Germany) was used for MSI.[36] Automatic gain control (AGC) was on. MSI experiments were acquired using a step-size of 50 µm for crab brain and 500 µm for rat brain. The details can be found in **SI**.

MALDI-TOF/TOF

An Autoflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA) equipped with a 200 Hz smartbeam™ was employed for MALDI-MSI and tandem MSI. The mass spectra data were acquired over a mass range of *m/z* 40-1000.

Each spectrum consists of 200 laser shots, and a step-size of 500 μm was used. Positive LIFT™ mode was employed for tandem MSI acquisition. The details can be found in **SI**.

Data analysis

The regional MS spectra and MS images obtained on the MALDI LTQ Orbitrap XL were processed using ImageQuest™ (Thermo Scientific). Briefly, the overall spectra of the rat or crab brain sections were generated by selecting the whole section as a ROI. The representative spectrum of the rat brain section was acquired by randomly selecting a region in the cortex as an ROI. A list of peaks with accurate mass information was obtained and verified by comparing the peaks from the off-the-tissue region, which is only covered with matrix, to ascertain the origin of the ions. The peak lists were then searched against the hmdb (http://www.hmdb.ca) and Metlin (http://metlin.scripps.edu/) databases with a mass accuracy window of 3 ppm. The identified metabolites were then searched in MetaCyc (http://metacyc.org/) and KEGG (http://www.genome.jp/kegg/kegg2.html) to ensure their biological roles in the CNS. For identified metabolites, distribution maps of selected metabolites were generated as spatially-resolved color-coded extracted ion chromatograms with narrow extraction windows. The smallest window set in ImageQuest software has a width of 2 mDa (+/- 1mDa). This equals to a +/- 3 ppm window width at *m/z* 330. Overlaying metabolites' distribution is accomplished via the "combo" feature in ImageQuest by selecting up to three metabolites' MS images extracted from a single dataset.

The overall mass spectra and MS images of the rat brain sections obtained on the Autoflex III TOF/TOF were processed by flexImaging (Bruker Daltonics). A representative spectrum was generated by selecting a ROI at an approximately identical region on a serial section of the rat brain. For tandem MSI of the putative metabolite guanosine monophosphate (GMP), the image file was also processed with flexImaging by monitoring the selected fragment ion within a mass window of 0.5 Da.

3.3 Results and Discussion

The difference in performance between HRMS and MRMS instrumentation is exhibited in **Figure 1**. In **Figure 1a**, the averaged mass spectrum acquired on the LTQ Orbitrap corresponds to the imaged region-of-interest (ROI) in rat brain cortex designated in **Figure 1b**. **Figure 1c** displays the averaged TOF/TOF spectrum of the ROI labeled in **Figure 1d**. Both ROIs are located in approximately the same regions of two adjacent sections, and the analytes desorbed/ionized from the two areas are assumed to produce similar mass spectra. The peaks present in the two representative spectra, **Figure 1a** and **1c**, exhibited similar masses and relative intensities. However, the spectral resolution showed a clear distinction between the data generated by HRMS and MRMS. The inserts display the explicit difference between the HRMS- and MRMS-produced spectra by enlarging the peak at *m/z* 364.07. The relatively low-abundance *m/z* 364.07 peak (at ~3% of base peak height) had a high signal to noise ratio (S/N) in both zoomed-in spectra, indicating the high sensitivity of both instruments. Nevertheless, the ion obtained from

the LTQ Orbitrap was detected at *m/z* 364.0655, whereas the ion detected by the TOF/TOF could only be assigned to *m/z* of 364.07 due to its limited resolution. The accurate, well-resolved peak in the LTQ Orbitrap spectrum contrasted the wide, asymmetrical peak at *m/z* 364.07 in the TOF/TOF spectrum and was confidently identified as guanosine monophosphate (GMP) with a mass error of 0.5 ppm. Additionally, peaks of small *m/z* differences at *m/z* 363.9300 or of low abundance compared to the base peak (**Figure 1a** insert at *m/z* 364.3265) were detected and well-resolved in the zoomed-in LTQ Orbitrap spectrum. Conversely, those peaks were merged in the shoulder of the *m/z* 364.07 peak in the enlarged TOF/TOF spectrum. This comparison demonstrates that HRMS can resolve nearly isobaric species and identify metabolites unequivocally from rat CNS specimen.

More chemical information generated from HRMS is shown in **Table 1**, and the resulting images of the identified metabolites and NTs are shown in **Figure 2**. Since we measured analytes' distribution as a proof-of-principle study to demonstrate the usefulness of HRMSI–based platform in imaging metabolites and NTs, we set a relatively large step-size to reduce the analysis time. **Figure 2a** shows an optical image of a rat brain section prepared for MSI experiments. Anatomical regions such as the olfactory bulb, cortex, medulla and pons, can be visualized clearly. The optical image was compared to MS images to interpret identified analyte localization. For example, **Figure 2b** displays the localization of choline (*m/z* at 104.1070) over almost the entire rat brain section. Choline's ubiquitous distribution could be explained by its role as an essential nutrient reported in many plants and animal organs.[26] It is also the precursor to a classic,

synaptic neurotransmitter, ACh. Mapping ACh is of interest to numerous researchers due to its involvement in various normal and abnormal neuronal functions, association with learning, memory and diseases like AD. [27] Conventional IHC methods studied the distributions of ACh-synthesizing enzyme choline acetyltransferase [28] or ACh-metabolizing enzyme acetylcholinesterase [29] to indirectly represent the localization of ACh. In this study, ACh was directly imaged from a rat brain section via single-stage MSI for the first time. As shown in **Figure 2c**, the abundance of the ACh peak was higher in the cortex and brainstem but lower in the olfactory bulb and cerebellum. The ACh localization image acquired via HRMSI correlated well with the distribution map of acetylcholinesterase from the Allen brain atlas (www.brain-map.org). Although this study attained low-spatial-resolution images due to the relatively large step-size (500 μm) during acquisition, high-spatial-resolution MS images of ACh can be obtained by implementing the HRMS-based platform with a highly focused laser beam and smaller step-size on tissue prepared with dryer matrix application method like sublimation. Previously, Sugiura et al. reported the application of tandem MSI to study the localization of ACh in mouse brain because of an unresolved, interfering matrix-related peak close to the *m/z* of ACh in MS using a MRMS MALDI-LTQ linear ion trap. [21] The image of a signature ACh fragment ion instead of intact ACh was used to represent the distribution of ACh. Although their method successfully circumvented matrix interferences and specifically imaged ACh from tissue samples, its throughput was compromised because only one precursor ion could be selected for monitoring in a single experiment. Moreover, tandem MSI results are based on the assumption that the production of

fragment ions can be correlated to the abundance of the precursor ion. However, this could be complicated due to factors like instrument instability and sample inhomogeneity. Additionally, the low intensity of the precursors might result in poor fragmentation, reducing the sensitivity of MSI for selected molecules. In contrast, the HRMSI-based approach is simple, relatively fast, and eliminates limitations resulting from the reliance on tandem MS events.

Figure 2d-e shows the distribution of two nucleotides, adenosine monophosphate (AMP) (*m/z* 348.0706) and GMP (*m/z* 364.0655), from the same rat brain section. AMP and GMP are both ubiquitously distributed on the rat brain with a higher abundance in the cortex. To further confirm the HRMSI results, we performed tandem MSI on the TOF/TOF by selecting *m/z* 364.07 as the precursor ion and monitoring the distribution of its major fragment ion at *m/z* 135.03. **Supplemental Figure (Figure S1a)** depicts the fragmentation pathway of this transition. The tandem MS image shown in **Figure S1c** confirms that the fragment ion is distributed throughout the brain region with a higher abundance in the cortex when correlated with the optical image in **Figure S1b**. **Figure S1d** corresponds to the image of an interfering ion produced simultaneously during fragmentation and is assigned as matrix-derived due to its distribution outside the tissue section. The difference in origin is shown in **Figure S1e**. Other than primary metabolites, lipid distributions are of interest to researchers because they have been reported to regulate a variety of biological functions in the CNS, including energy storage, membrane composition and chemical signaling.[25, 30, 31] In **Figure 2f**, the ion at *m/z* 369.3519 was identified as cholesterol with a neutral loss of water, agreeing with the

previous observation of an *m/z* 369.3 peak by Jackson et al. using a MRMS MALDI-TOF.[32] The abundance of the cholesterol peak was higher in brainstem regions like the medulla, pons and midbrain, as shown in **Figure 2f**. This result agrees with previous work showing that cholesterol is more concentrated in areas that are rich in myelinated axons.[33] **Figure 2g** displays the localization of a potassiated lipid phosphotidylcholine (PC) (32:0) at *m/z* 772.5258. Although both of the molecules are lipids, they belong to different lipid families and exhibited distinct distributions, indicating their different biological roles involved in rat CNS functioning. **Figure 2h** shows the distribution of a coenzyme extensively used in cells' redox reactions, nicotinamide adenine dinucleotide, at *m/z* 664.1177. This coenzyme is distributed differently than other metabolites shown in **Figure 2** by its elevated abundance in the diencephalon and decreased expression levels in the forebrain. HRMSI also enables direct mapping of peptide neurotransmitters in the CNS and makes it possible to investigate the co-localization of low-MW and peptide NTs. **Figure 2i** displays the MS image of a truncated peptide of myelin basic protein detected at *m/z* 2028.076, which shares a similar distribution pattern with cholesterol in myelin-rich regions. **Figure 2j** is an overlay of **Figure 2c** and **2f**, showing the distinct distribution of these two biomolecules (ACh and cholesterol-derived ions) from a single rat brain section. These representative images demonstrate the ability of HRMSI to map metabolites and NTs, and investigate their biological interactions through spatial correlation. It is worth mentioning that the HRMS LTQ-Orbitrap is also capable of tandem MSI compared to TOF/TOF, whereas it outperforms TOF/TOF by isolating multiple precursors in a single experiment. A multiplexing imaging approach has been

developed to attain untargeted and targeted imaging data simultaneously. [17, 24] By incorporating both full MS scan and multiple tandem MS scans as one experiment, one could maximize the information gained from precious sample material with minimal acquisition time. [17, 24]

In order to further validate the HRMSI platform, we used it to map metabolites from another CNS model, blue crab *C. sapidus* brain, for the first time. **Figure 3** shows the HRMS spectra of metabolites recorded from the crab brain. **Figure 3a** shows the overall HRMS spectrum averaged across the whole crab brain section, and **Figure 3b-3d** are three enlarged mass spectra, exemplifying the importance of high-resolution and high-accuracy in metabolite identification. Specifically, **Figure 3b** examines the *m/z* range of 147.06-147.12. Two peaks at *m/z* 147.0763 and 147.1127 were well-resolved (**Figure 3b**) and identified as glutamine and lysine, respectively. **Figure 3c** demonstrates that our HRMSI approach provides superior mass resolution by separating two peaks differing by only 0.23 mDa at *m/z* 176.0084 and 176.0107. The corresponding ion density maps are shown in **Figure 3e-f**, where the potassiated aminobenzoic acid (*m/z* 176.0107) was distributed distinctly from another ion (*m/z* 176.0084), illustrating the necessity of acquiring accurate *m/z* values to resolve low-MW compounds in MSI experiments. Otherwise, the two peaks would have been observed as a solitary peak in MRMS data, leading to a shift in observed *m/z* and incorrect MS images of the two molecules as one. **Figure 3d** displays the zoomed-in spectrum encompassing *m/z* 385.94-386.14, in which two peaks with close *m/z* values (*m/z* 386.0263 and 386.0478) were identified as potassiated AMP and sodiated GMP.

Supplemental Figure 2a is an illustration of a crab brain with the main anatomical regions annotated.[34] **Supplemental Figure 2b** shows the MS image of a major matrix-related ion at *m/z* 137.0233 that was identified as $[M\text{-}H_2O+H]^+$ and was detected over the entire imaged region. Conversely, the MS image in **Supplemental Figure 2c** shows that the endogenous metabolites were localized exclusively in the brain section, facilitating the differentiation between matrix peaks and endogenous molecules. Various classes of metabolites' MS images were collected by the HRMSI-based technique and shown here for the first time (**Figure 4**). Specifically, amino acids like glutamine (*m/z* 147.0763), lysine (*m/z* 147.1127), histidine (*m/z* 156.0766), arginine (*m/z* 175.1190), and carnitine (*m/z* 162.1122) were found to be distributed over the entire crab brain (**Figure 4a**). In comparison, an *m/z* 191.1138 ion corresponding to hydroxyarginine was localized to different regions of the brain, including the protocerebrum and median antenna I neuropil (MAN) from the deutocerebrum. Moreover, HRMS images of other amino acids such as proline, leucine, phenylalanine and tyrosine were also obtained. Nucleobases were also imaged, including adenine (*m/z* 136.0620) and guanine (*m/z* 152.0568), as shown in **Figure 4b**. Interestingly, adenine and guanine were both localized to the large, spherical olfactory lobes (ON) in the deutocerebrum, whereas AMP and GMP ions (*m/z* 348.0707 and 364.0657, respectively) displayed nearly identical distributions. Based on HRMS measurements, the well-separated peaks in **Figure 3d** were identified as potassiated AMP and sodiated GMP, and their distributions agreed with the protonated forms of AMP and GMP accordingly. Organic acids and their adducts were also detected by HRMSI, as shown in **Figure 4c**. For example, the protonated, sodiated and potassiated forms of

aminobenzoic acid, *m/z* 138.0547, 160.0368 (**Figure 4c**) and 176.0107, (**Figure 3f**) were shown to co-localize in the outer, whitish sheath of the crab brain section, which validated our HRMS-based assignment. Another example is aminopentanoic acid. Its sodiated peak, *m/z* 140.0680, was also concentrated in the outer, fatty layer outside the neuronal structure. **Figure 4d** overlays three metabolites that exhibited drastically different localizations on the same crab brain section, aminopentanoic acid (red), hydroxyarginine (green) and GMP (blue). Their different localizations suggest each metabolite has a different function in the crustacean CNS. Using crustacean CNS as a model, HRMSI's role in MSI of metabolites and NTs is once again validated by assigning molecular identities (detailed in **Table 2** and **Supplemental Table 1**) and quickly delivering accurate spatial images of multiple ions from biological samples simultaneously. Notably, unambiguous identification has to be verified via structural elucidation by MS/MS. Besides the presence of isobaric ions, in-source decay (ISD) fragments that are formed during precursor ion acceleration process thermally or by radicals could also complicate the molecular assignment if merely relying on HRMS. [35]

HRMSI has also become integral in neuropeptide imaging studies of neural tissue.[11] **Figure 4f** is an overlaid MS image of the metabolite AMP (red) and a protocerebrum-concentrated neuropeptide SIFamide at *m/z* 1381.735 (**Figure 4e**) (green). Small molecule and peptide distribution data may assist investigations that study the dynamic interaction and synergetic relationship between neurotransmitters and neuropeptides.

3.4 Conclusions

Our study presents a novel HRMSI-based platform to investigate the distribution of metabolites and NTs in biologically complex and important specimen. The high spectral resolution and mass accuracy provided by HRMS instrumentation employed in this study was advantageous when identifying analytes of interest in comparison to a MRMS instrument. We successfully generated MS images of multiple metabolites and NTs in rat and crustacean CNS, demonstrating the potential of HRMSI in biomedical applications. We anticipate that the novel knowledge of metabolite and NT distributions gained via HRMSI in this study will provide insight into related neuroscience research and present a powerful approach for future studies.

Acknowledgments

This work is supported in part by the National Institutes of Health grants (1R01DK071801, 1R56DK071801, and 1S10RR029531) and the National Science Foundation grant (CHE-0957784). L. Li acknowledges an H.I. Romnes Faculty Research Fellowship.

3.5 References

1. Orešič, M.; Hyötyläinen, T.; Herukka, S. K.; Sysi-Aho, M.; Mattila, I.; Seppänan-Laakso, T.; Julkunen, V.; Gopalacharyulu, P. V.; Hallikainen, M.; Koikkalainen, J.; Kivipelto, M.; Helisalmi, S.; Lötjönen, J.; Soininen, H. (2011) Metabolome in progression to Alzheimer's disease. *Transl. Psychiatry* 1, e57.
2. Houtkooper, R. H.; Argmann, C.; Houten, S. M.; Canto, C.; Jeninga, E. H.; Andreux, P. A.; Thomas, C.; Doenlen, R.; Schoonjans, K.; Auwerx, J. (2011) The metabolic footprint of aging in mice. *Sci. Rep.* 1, 134.
3. Krasensky, J.; Jonak, C. (2012) Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *J. Exp. Bot.* 63, 1593-1608.

4. Want, E. J.; Cravatt, B. F.; Siuzdak, G. (2005) The expanding role of mass spectrometry in metabolite profiling and characterization. *Chembiochem* 6, 1941-51.
5. Marc, D. T.; Ailts, J. W.; Campeau, D. C. A.; Bull, M. J.; Olson, K. L. (2011) Neurotransmitters excreted in the urine as biomarkers of nervous system activity: validity and clinical applicability. *Neurosci. Biobehav. Rev.* 35, 635-644.
6. Westerink, B. H. C.; Hofsteede, H. M.; Damsma, G.; de Vries, J. B. (1988) The significance of extracellular calcium for the release of dopamine, acetylcholine and amino-acids in conscious rats, evaluated by brain microdialysis. *Naunyn Schmiedebergs Arch. Pharmacol.* 337, 373-378.
7. Watanabe, T.; Yamagata, N.; Takasaki, K.; Sano, K.; Hayakawa, K.; Katsurabayashi, S.; Egashira, N.; Mishima, K.; Iwasaki, K.; Fujiwara, M. (2009) Decreased acetylcholine release is correlated to memory impairment in the Tg2576 transgenic mouse model of Alzheimer's disease. *Brain Res.* 1249, 222-8.
8. Braak, H.; Del Tredici, K.; Rüb, U.; de Vos, R. A.; Jansen Steur, E. N.; Braak, E. (2003) Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* 24, 197-211.
9. Lei, Z.; Huhman, D. V.; Sumner, L. W. (2011) Mass spectrometry strategies in metabolomics. *J. Biol. Chem.* 286, 25435-42.
10. Jones, B. E.; Beaudet, A. (1987) Distribution of acetylcholine and catecholamine neurons in the cat brain stem: a choline acetyltransferase and tyrosine hydroxylase immunohistochemical study. *J. Comp. Neurol.* 261, 15-32.
11. Ye, H.; Greer, T.; Li, L. (2012) Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *J. Proteomics* 75, 5014-26.
12. McDonnell, L. A.; Heeren, R. M. A.; Andren, P. E.; Stoeckli, M.; Corthals, G. L. (2012) Going forward: increasing the accessibility of imaging mass spectrometry. *J. Proteomics* 75, 5113-5121.
13. Ifa, D. R.; Wu, C.; Ouyang, Z.; Cooks, R. G. (2010) Desorption electrospray ionization and other ambient ionization methods: current progress and preview. *Analyst* 135, 669-681.
14. Liu, Q.; Xiao, Y. S.; Pagan-Miranda, C.; Chiu, Y. M.; He, L. (2009) Metabolite Imaging using matrix-enhanced surface-assisted laser desorption/ionization mass spectrometry (ME-SALDI-MS). *J. Am. Soc. Mass Spectrom.* 20, 80-88.
15. Liu, J. J.; Cooks, R. G.; Ouyang, Z. (2011) Biological tissue diagnostics using needle biopsy and spray ionization mass spectrometry. *Anal. Chem.* 83, 9221-9225.
16. Jun, J. H.; Song, Z.; Liu, Z.; Nikolau, B. J.; Yeung, E. S.; Lee, Y. J. (2010) High-spatial and high-mass resolution imaging of surface metabolites of *Arabidopsis thaliana* by laser desorption-ionization mass spectrometry using colloidal silver. *Anal. Chem.* 82, 3255-65.
17. Yagnik, G. B.; Korte, A. R.; Lee, Y. J. (2013) Multiplex mass spectrometry imaging for latent fingerprints. *J. Mass Spectrom.* 48, 100-4.
18. Ye, H.; Greer, T.; Li, L. J. (2011) From pixel to voxel: a deeper view of biological tissue by 3D mass spectral imaging. *Bioanalysis* 3, 313-332.
19. Lee, Y. J.; Perdian, D. C.; Song, Z. H.; Yeung, E. S.; Nikolau, B. J. (2012) Use of mass spectrometry for imaging metabolites in plants. *Plant J.* 70, 81-95.
20. Pirman, D. A.; Reich, R. F.; Kiss, A.; Heeren, R. M.; Yost, R. A. (2012) Quantitative MALDI tandem mass spectrometric imaging of cocaine from brain tissue with a deuterated internal standard. *Anal. Chem.* published online Dec 7, 2012, DOI: 10.1021/ac302960j.
21. Cornett, D. S.; Frappier, S. L.; Caprioli, R. M. (2008) MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. *Anal. Chem.* 80, 5648-5653.
22. Nilsson, A.; Forngren, B.; Bjurstrom, S.; Goodwin, R. J.; Basmaci, E.; Gustafsson, I.; Annas, A.; Hellgren, D.; Svanhagen, A.; Andren, P. E.; Lindberg, J. (2012) In situ mass

spectrometry imaging and ex vivo characterization of renal crystalline deposits induced in multiple preclinical drug toxicology studies. *PLoS One* 7, e47353.
23. Sugiura, Y.; Zaima, N.; Setou, M.; Ito, S.; Yao, I. (2012) Visualization of acetylcholine distribution in central nervous system tissue sections by tandem imaging mass spectrometry. *Anal. Bioanal. Chem.* 403, 1851-1861.
24. Perdian, D. C.; Lee, Y. J. (2010) Imaging MS methodology for more chemical information in less data acquisition time utilizing a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.* 82, 9393-9400.
25. Manicke, N. E.; Dill, A. L.; Ifa, D. R.; Cooks, R. G. (2010) High-resolution tissue imaging on an orbitrap mass spectrometer by desorption electrospray ionization mass spectrometry. *J. Mass Spectrom.* 45, 223-226.
26. Zeisel, S. H.; da Costa, K. A. (2009) Choline: an essential nutrient for public health. *Nutr. Rev.* 67, 615-23.
27. Vanderwolf, C. H. (1988) Cerebral activity and behavior: control by central cholinergic and serotonergic systems. *Int. Rev. Neurobiol.* 30, 225-340.
28. Armstrong, D. M.; Saper, C. B.; Levey, A. I.; Wainer, B. H.; Terry, R. D. (1983) Distribution of cholinergic neurons in rat brain: demonstrated by the immunocytochemical localization of choline acetyltransferase. *J. Comp. Neurol.* 216, 53-68.
29. Jacobowitz, D. M.; Palkovits, M. (1974) Topographic atlas of catecholamine and acetylcholinesterase-containing neurons in the rat brain. *J. Comp. Neurol.* 157, 13-28.
30. Johansson, B. (2006) ToF-SIMS imaging of lipids in cell membranes. *Surf. Interface Anal.* 38, 1401-1412.
31. Fernandis, A. Z.; Wenk, M. R. (2007) Membrane lipids as signaling molecules. *Curr. Opin. in Lipidol.* 18, 121-128.
32. Jackson, S. N.; Wang, H. Y. J.; Woods, A. S. (2005) Direct profiling of lipid distribution in brain tissue using MALDI-TOFMS. *Anal. Chem.* 77, 4523-4527.
33. Nemes, P.; Woods, A. S.; Vertes, A. (2010) Simultaneous imaging of small metabolites and lipids in rat brain tissues at atmospheric pressure by laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 82, 982-988.
34. Sandeman, D.; Sandeman, R.; Derby, C.; Schmidt, M. (1992) Morphology of the brain of crayfish, crabs, and spiny lobsters: a common nomenclature for homologous structures. *Biol. Bull.* 183, 304-326.
35. Debois, D.; Bertrand, V.; Quinton, L.; De Pauw-Gillet, M. C.; De Pauw, E. (2010) MALDI-in source decay applied to mass spectrometry imaging: a new tool for protein identification. *Anal. Chem.* 82, 4036-45.
36. Strupat, K.; Kovtoun, V.; Bui, H.; Viner, R.; Stafford, G.; Horning, S. (2009) MALDI produced ions inspected with a linear ion trap-orbitrap hybrid mass analyzer. *J. Am. Soc. Mass Spectrom.* 20, 1451-1463.

Table 1. Metabolites and NTs identified from the rat CNS specimen.

Name	Formula	Theoretical Monoisotopic *m/z*	Orbitrap XL Measurement *m/z*	Δm (ppm)
Choline	$C_5H_{14}NO$	104.1070	104.1070	0
Acetylcholine	$C_7H_{16}NO_2$	146.1176	146.1175	-0.7
AMP	$C_{10}H_{15}N_5O_7P$	348.0704	348.0706	0.6
GMP	$C_{10}H_{15}N_5O_8P$	364.0653	364.0655	0.5
Cholestrol	$C_{27}H_{45}$	369.3516	369.3519	0.8
Nicotinamide adenine dinucleotide	$C_{21}H_{28}N_7O_{14}P_2$	664.1164	664.1177	2.0
PC(32:0)+K	$C_{40}H_{80}NO_8PK$	772.5253	772.5258	0.6

Table 2. Metabolites identified from the crab brain specimen based on HRMSI.

Name	Formula	Theoretical Monoisotopic m/z	Orbitrap XL Measurement *m/z*	Δm (ppm)
Glutamine	$C_5H_{11}N_2O_3$	147.0764	147.0763	-0.7
Lysine	$C_6H_{15}N_2O_2$	147.1128	147.1127	-0.7
Histidine	$C_6H_{10}N_3O_2$	156.0768	156.0766	-1.3
Carnitine	$C_7H_{16}NO_3$	162.1125	162.1122	-1.9
Arginine	$C_6H_{15}N_4O_2$	175.1190	175.1190	0
Hydroxyarginine	$C_6H_{15}N_4O_3$	191.1139	191.1138	-0.5
Adenine	$C_5H_6N_5$	136.0618	136.0620	1.5
Guanine	$C_5H_6N_5O$	152.0567	152.0568	0.7
AMP	$C_{10}H_{15}N_5O_7P$	348.0704	348.0707	0.9
GMP	$C_{10}H_{15}N_5O_8P$	364.0653	364.0657	1.1
AMP+K	$C_{10}H_{14}N_5O_7PK$	386.0262	386.0263	0.3
GMP+Na	$C_{10}H_{14}N_5O_8PNa$	386.0472	386.0478	1.6
Aminobenzoic acid	$C_7H_8NO_2$	138.0550	138.0547	-2.2
Aminobenzoic acid+Na	$C_7H_7NO_2Na$	160.0369	160.0368	-0.6
Aminopentanoic acid+Na	$C_5H_{11}NO_2Na$	140.0682	140.0680	-1.4

Figures:

Figure 1. Comparison between HRMS- and MRMS-produced spectra acquired from serial rat CNS sections. (a) An HRMS-generated spectrum corresponding to the region of interest (ROI) localized in the cortex region of a rat brain section as highlighted in (b). The insert shows an enlarged *m/z* range of 362-366 obtained on a MALDI-LTQ-Orbitrap XL. (c) An MRMS-generated mass spectrum corresponding to an identical ROI from the cortex region of a serial rat brain section as shown in (d). The insert shows an enlarged *m/z* range of 362 to 366 obtained on a MALDI-TOF/TOF.

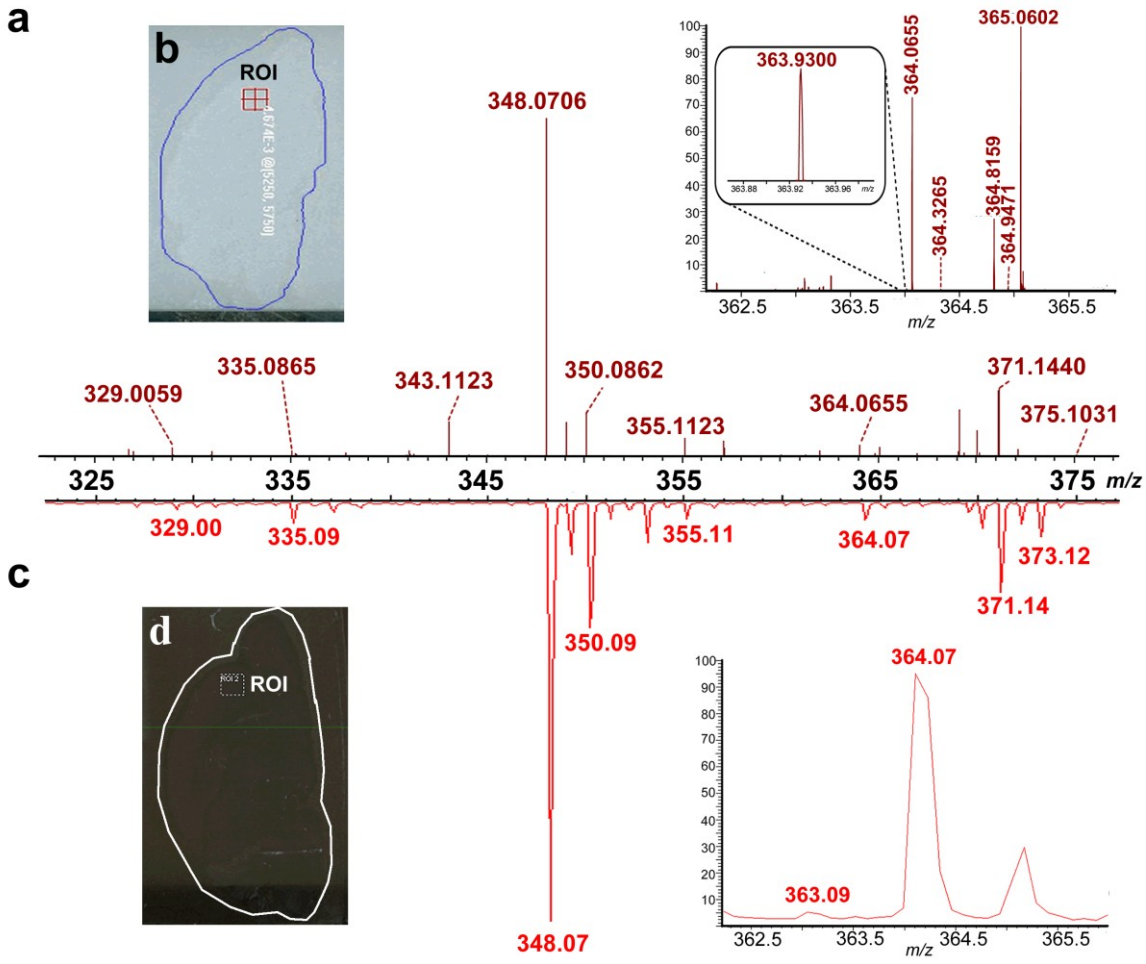


Figure 2. MS images of metabolites and NTs identified from rat CNS. (a) An optical image of a rat brain section subjected to MSI sample preparation. (b)-(g) MS image of metabolites and NTs, including (b) choline at *m/z* 104.1070, (c) ACh at *m/z* 146.1175, (d) AMP at *m/z* 348.0706, (e) GMP at *m/z* 364.0655, (f) cholesterol with neutral loss at *m/z* 369.3519, (g) potassiated PC(32:0) at *m/z* 772.5258 and (h) nicotinamide adenine dinucleotide at *m/z* 664.1177. Other than low MW molecules, (i) MS image of the acetylated peptide ASQKRPSQRHGSKYLATA at *m/z* 2028.076 was also shown. (j) An overlaid image of a NT, ACh, as in (c) and cholesterol-derived ion as in (f).

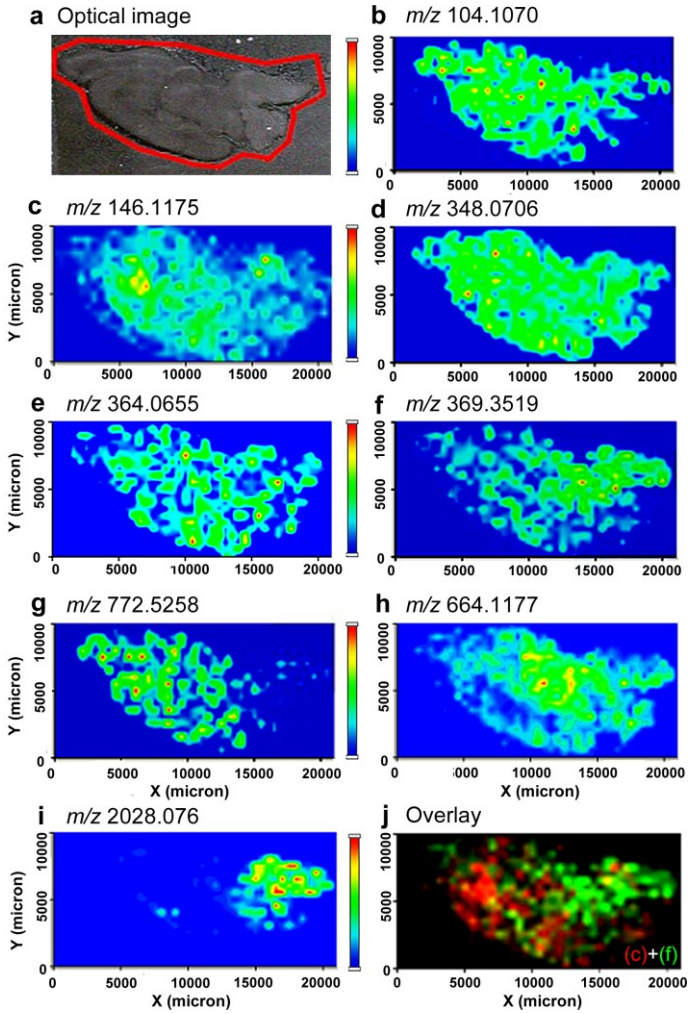


Figure 3. HRMSI spectra averaged from a blue crab brain section acquired on a MALDI-LTQ-Orbitrap XL. (a) An averaged HRMS spectra displaying the *m/z* range of 90-500. (b)-(d) The zoomed-in spectra shown in (a) over the range of (b) *m/z* 147.06-147.12, (c) *m/z* 175.98-176.04 and (d) *m/z* 385.94-386.14. (e) and (f) display different distributions of two ions present in (c) from blue crab brain.

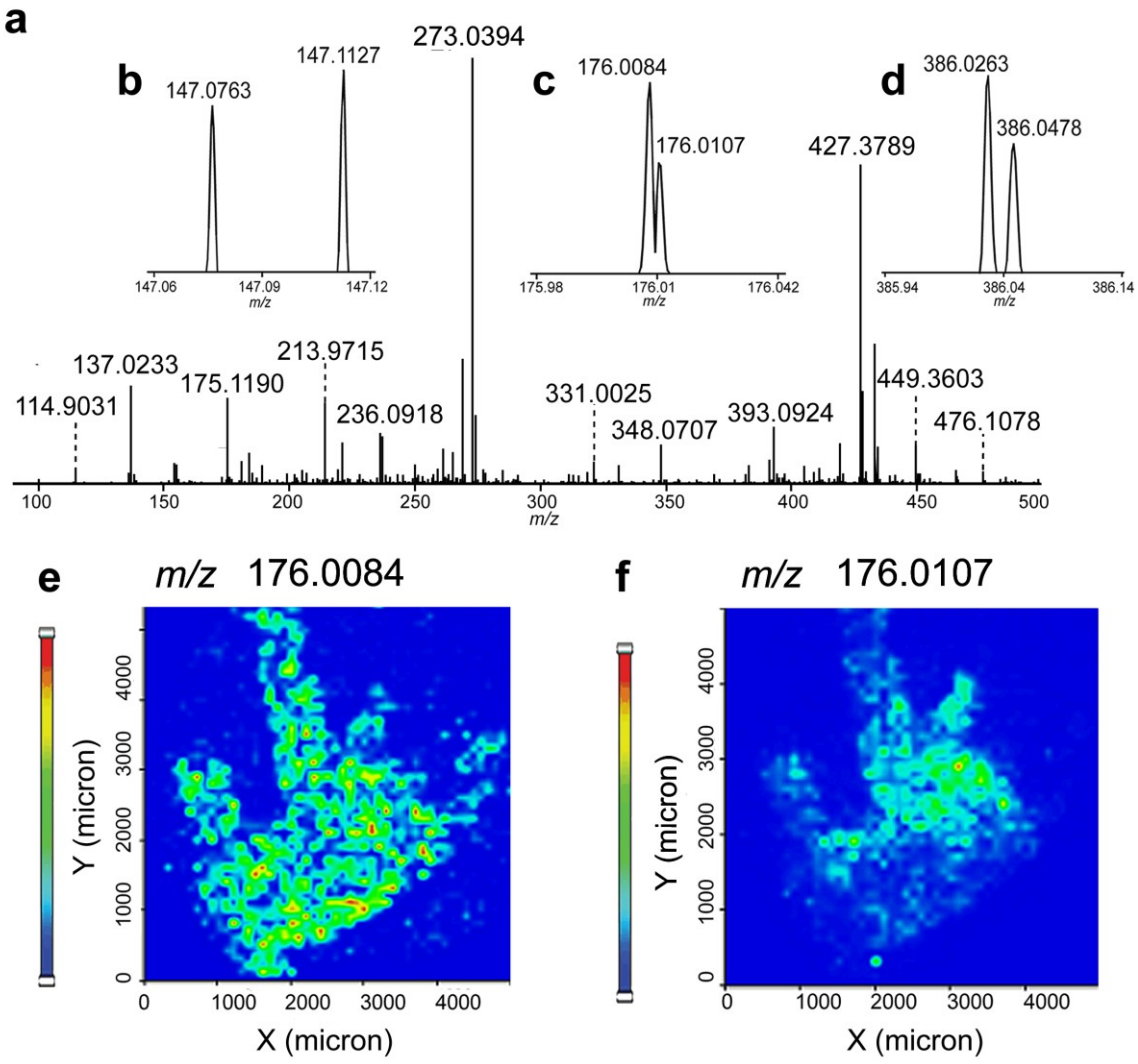
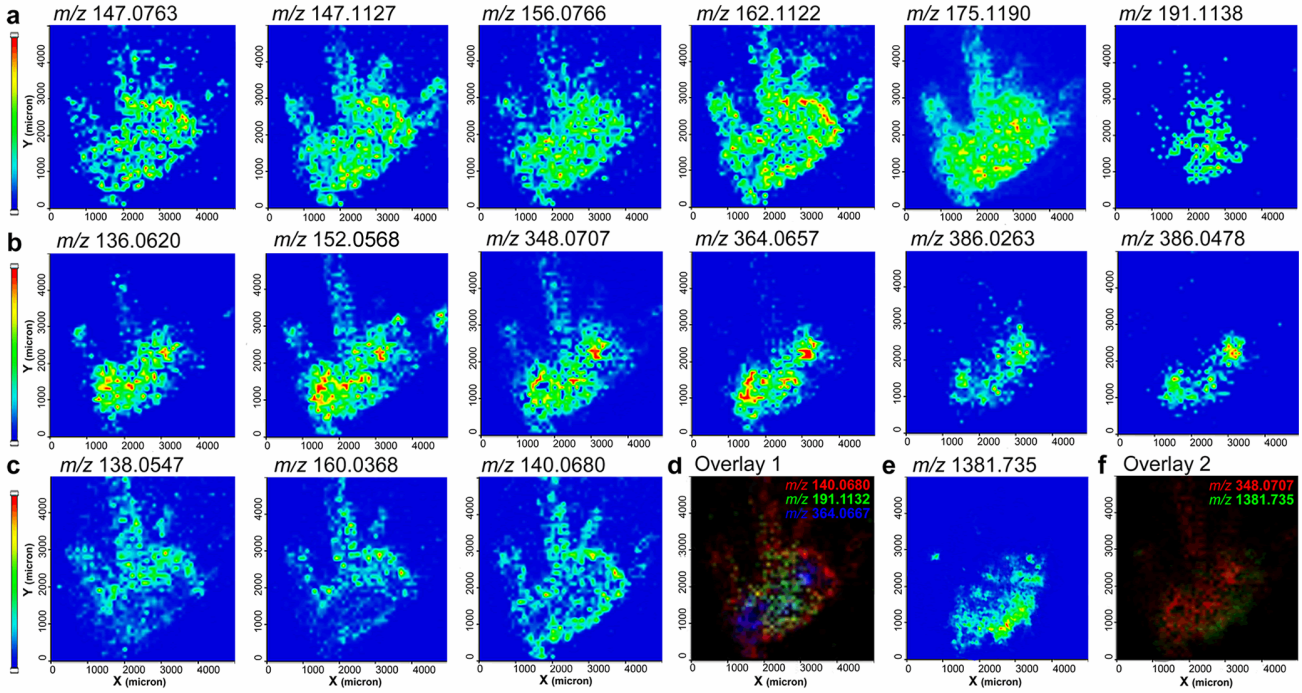


Figure 4. MS images of metabolites identified from crustacean CNS. (a) MS images of amino acids, including glutamine, lysine, histidine, carnitine, arginine and hydroxyarginine, respectively. (b) MS images of nucleobase-derived metabolites, including adenine, guanine, AMP, GMP, potassiated AMP and sodiated GMP, respectively. (c) MS images of organic acids, including the protonated and sodiated aminobenzoic acid and the sodiated aminopentanoic acid. (d) An overlaid image of the sodiated aminopentoic acid as in red, the hydroxyarginine as in green and GMP as in blue. Other than small molecules, (e) MS image of the neuropeptide GYRKPPFNGSIFa displayed a distinct distribution from all the metabolites shown above. (f) An overlaid



Supplemental Information

Animal dissection

Animal experiments were performed according to institutional guidelines (UW-Madison IACUC). Blue crabs, C. *sapidus*, were purchased from the local market and maintained in recirculating artificial seawater aquaria at approximately 12-13 °C before use. Crabs were anesthetized by packing them on ice for 15 min, after which the dorsal carapace was removed. The crab's brain was dissected free from the connective stomach in chilled physiological saline. Wistar rats were terminally anesthetized with 5% isoflurane, transcardially perfused with saline solution and decapitated on P10 according to the approved experiment protocol. The details of dissection were described elsewhere.[32]

Sample Preparation

The crab brain was rinsed briefly in deionized water immediately following dissection to reduce the salt content, after which the excessive water was removed by KimWipes™. The brain was placed in a 100 mg/mL gelatin aqueous solution contained in a tissue block box. The rat brain was also embedded in the gelatin aqueous solution immediately after dissection. The tissue blocks were then flash-frozen in dry ice. The completely frozen tissues were sectioned to 12 μm sections that were mounted onto indium tin oxide (ITO) coated glass slides (Delta, Loveland, CO, USA). Optical images of the crab or rat brain sections were taken by scanning the glass slide with an office scanner (Epson V300 PHOTO) in professional mode with 2400 dpi. The crab or rat brain sections were then dehydrated in a desiccator at -20 °C for 30 min prior to matrix

application. For matrix application, 150 mg/ml DHB (50:50 methanol: water, v/v) was used and the airbrush (Paasche airbrush company, Chicago, IL, USA) was held 35 cm from the plate for regular spray. Five coats were applied, and the spray duration for each coat was 30 s with 1 min dry time between each cycle.

MALDI-TOF/TOF

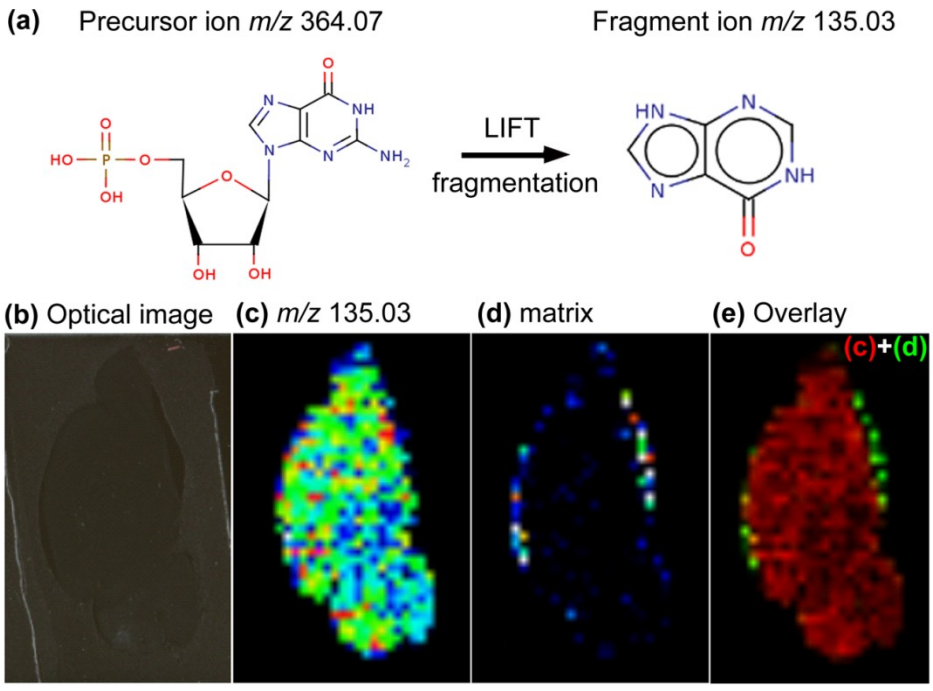
An Autoflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA) equipped with a 200 Hz smartbeam™ was employed for MALDI-MSI and tandem MSI. The following parameters were input in the positive reflectron mode for imaging acquisition: ion source 1 voltage 19.00 kV, ion source 2 voltage 16.62 kV, reflector 1 voltage 20.90 kV, reflector 2 voltage 9.64 kV and lens voltage 8.70 kV. The mass spectra data were acquired over a mass range of *m/z* 40-1000. Each spectrum consists of 200 laser shots, and the array of spectra was collected at 500 μm intervals in both x and y dimensions across the surface of the rat brain section. The following parameters were adopted in the positive LIFT™ mode for tandem MSI acquisition: precursor ion at *m/z* 364.07, ion isolation window ± 2Da, ion source 1 voltage 6.00 kV, ion source 2 voltage 5.30 kV, lens voltage 2.85 kV, reflector voltage 27.00 kV, Reflector 2 voltage 11.50 kV, Lift 1 voltage 19.00 kV and Lift 2 voltage 4.40 kV. The array of spectra was collected at a step-size of 500 μm, and each spectrum consists of 200 laser shots and was acquired over the *m/z* 40-400 mass range.

MALDI LTQ Orbitrap XL

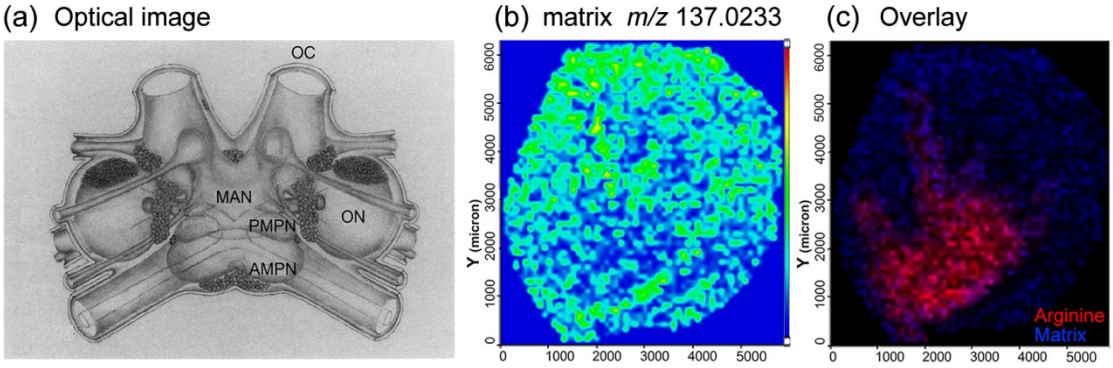
The MALDI LTQ Orbitrap XL instrument (Thermo Scientific, Bremen, Germany) equipped with a commercial 60 Hz N_2 laser at 337 nm (LTB Lasertechnik Berlin GmbH, Berlin, Germany) was used for MSI. [31] All experiments were performed with automatic gain control (AGC) turned on. MSI experiments were acquired using a step-size of 50 μm for crab brain and 500 μm for rat brain. Qualitative data were obtained using Thermo Scientific Tune 2.5.5 SP 2 instrument control software and Thermo Scientific Xcalibur™ software. The Orbitrap analyzer was fully tuned and calibrated with the aid of two calibration peptide mixtures (MSCal4, Sigma Aldrich, St. Louis, MO) for optimization in two mass ranges: *m/z* 50-1100 for small molecules and *m/z* 900-3000 for peptides. Data has been internally calibrated using a matrix peak $[DHB+H-H_2O]^+$ at *m/z* 137.0233 to improve mass accuracy using RecalOffline (Thermo Scientific) in a post-processing manner.

Supplemental Figures

Supplemental Figure 1. Tandem MSI of the *m/z* 364.07 ion on a LRMS MALDI-TOF/TOF. (a) Fragmentation pathway of the GMP precursor ion at *m/z* 364.07 to the monitored fragment ion at *m/z* 135.03. (b) Optical image of the rat brain section used for pseudo SRM imaging. (c) Tandem MS image of the major fragment at *m/z* 135.03. (d) Tandem MS image of an interfering ion resulted from fragmentation of matrix introduced simultaneously with the targeted precursor ion. (e) Overlay of the major AMP fragment and the matrix fragment. The presence of the interfering fragment ion demonstrated the interference of matrix in single-stage MSI when employing LRMS.



Supplemental Figure 2. (a) Illustration of the dorsal view of a crab brain, which reveals the main structure, including AMPN and PMPN from protocrebrum, MAN and ON from deutocrebrum and tracts that link neuropils like OC.[30] (b) MS image of a matrix-derived ion at *m/z* 137.0233. (c) An overlaid image of arginine in red and the *m/z* 137.0233 ion in blue. Their origins are revealed by their distinct distribution maps provided by HRMSI.



Supplemental Table 1. Metabolites identified from the crab brain specimen based on HRMSI.

Name	Formula	Theoretical Monoisotopic m/z	Orbitrap XL Measurement *m/z*	Δm (ppm)
Phenylalanine	$C_9H_{12}NO_2$	166.0863	166.0859	-2.4
Tyrosine	$C_9H_{12}NO_3$	182.0812	182.0814	1.1
NAD	$C_{21}H_{28}N_7O_{14}P_2$	664.1164	664.1177	2.0
Aminobenzoic acid+K	$C_7H_7NO_2K$	176.0108	176.0107	-0.6
Aminopentanoic acid	$C_5H_{12}NO_2$	118.0863	118.0863	0
Aminopentanoic acid+K	$C_5H_{11}NO_2K$	156.0421	156.0418	-1.9
GDP	$C_{10}H_{16}N_5O_{11}P_2$	444.0316	444.0317	0.2
ADP-ribose	$C_{15}H_{24}N_5O_{14}P_2$	560.0790	560.0797	1.3

Chapter 4

MALDI Mass Spectrometry-assisted Molecular Imaging of Metabolites During Nitrogen Fixation in the *Medicago truncatula – Sinorhizobium meliloti* Symbiosis

Adapted from **Hui Ye**, Erin Gemperline, Muthusubramanian Venkateshwaran, Ruibing Chen, Pierre-Marc Delaux, Maegen Howes-Podoll, Jean-Michel Ané, Lingjun Li. MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the *Medicago truncatula – Sinorhizobium meliloti* symbiosis. *Plant J.*, doi: 10.1111/tpj.12191

Abstract

Symbiotic associations between leguminous plants and nitrogen-fixing rhizobia culminate in the formation of specialized organs called root nodules, in which the rhizobia fix atmospheric nitrogen and transfer it to the plant. Efficient biological nitrogen fixation depends on metabolites produced by and exchanged between both partners. The *Medicago truncatula–Sinorhizobium meliloti* association is an excellent model for dissecting this nitrogen-fixing symbiosis because of the availability of genetic information in both symbiotic partners. Here, we employed a powerful imaging technique, matrix-assisted laser desorption/ionization (MALDI)-mass spectrometric imaging (MSI), to study metabolite distribution in roots and root nodules of *M. truncatula* during nitrogen fixation. The combination of an efficient, novel MALDI matrix 1,8-bis(dimethyl-amino) naphthalene (DMAN, $pK_{a(conjugate\text{-}acid)}$ = 12.5) with a conventional matrix 2,5-dihydroxybenzoic acid (DHB) allowed the detection of a large array of organic acids, amino acids, sugars, lipids, flavonoids and their conjugates with improved coverage. Ion density maps of representative metabolites are presented and correlated with the nitrogen fixation process. We demonstrate differences in metabolite distribution between roots and nodules, and also between fixing and non-fixing nodules produced by plant and bacterial mutants. Our study highlights the benefits of using MSI for detecting differences in metabolite distributions in plant biology.

4.1 Introduction

Legumes are able to develop a very efficient nitrogen fixing symbiosis with soil bacteria collectively referred to as rhizobia. Every year, legume nodulation produces a quantity of available nitrogen that is equivalent to the amount synthesized by the fertilizer industry throughout the world. [1, 2] This symbiotic association leads to the development of specialized organs called root nodules in which the rhizobia reduce atmospheric dinitrogen into ammonium and transfer it to the host plant. [3] Biological nitrogen fixation is catalyzed by the nitrogenase complex and is energetically expensive for the bacteria that in turn require an abundant carbon supply from the plant partner. [4] Moreover, the nitrogenase is irreversibly inhibited by oxygen so root nodules create micro-aerobic conditions in part through the production of a hemeprotein called leghemoglobin. [5, 6] Therefore, maintaining this efficient symbiosis requires a fine coordination between legume and rhizobial metabolic processes. [7]

Medicago truncatula (Medicago) has emerged as model legume for studies pertaining not only to plant-microbe symbioses, but also to general legume biology. [8-13] Transcriptomic, proteomic and phosphoproteomic studies have been extensively pursued in this model legume. [14-16] By contrast, most non-targeted large scale metabolomic studies have often been limited to metabolite profiling of total extracts obtained from Medicago suspension cell cultures. [17-20] One notable exception is a recent untargeted

metabolomic study that used entire Medicago seedlings and identified oxylipins as important regulators of early symbiotic signaling. [21]

Plant metabolites have mainly been investigated using conventional techniques, such as liquid chromatography-mass spectrometry (LC-MS), gas chromatography-mass spectrometry (GC-MS) and capillary electrophoresis-mass spectrometry (CE-MS). [22-26] Various amino acids, organic acids, sugars and polyols were shown to be differentially represented in roots and nodules. [22, 23, 27-29] While impressive, these experiments used homogenized metabolite extracts, which were not able to resolve the spatial distribution of the metabolic components at the organ or tissue levels.

MALDI-MSI has emerged as a powerful tool to investigate the distribution of a wide range of molecules through direct analysis of plant specimen to visualize the distribution of multiple metabolites among different organs and tissues, gaining enormous interest in plant biology. [30-32] This innovative technique thereby holds promise for a mechanistic understanding of the metabolic differentiation that defines legume nodulation. However, the matrix involved in desorption/ionization process usually generates abundant peaks at low *m/z* (mass to charge ratio), thus interfering with the detection of small molecules present in the samples. Matrix-free laser desorption/ionization (LDI) has therefore been investigated and proven applicable to image UV-absorbing metabolites from plant tissues. [33] Moreover, various matrix-free surface

modified/functionalized techniques including metal [34] and carbon materials, such as graphite powder, graphite solution, colloidal graphite or nanostructure-initiator functionalized chips were developed to improve the performance of metabolite imaging applications. [35-37] Alternative ionization techniques have been applied to the field of small molecule MSI. Secondary ion mass spectrometry, which provides comparatively good resolution, has been coupled to MSI to investigate metabolites' distribution, but is limited by its poor detection efficiency in the low mass range due to relatively low yields of secondary ion production. [38] Some ambient ion sources have been developed recently, including atmospheric pressure infrared-MALDI, [39, 40] laser ablation electrospray ionization [41], direct analysis in real time [42] and direct electrospray ionization. [43, 44] Nonetheless, all of these ambient ion sources require modified instrument setup and provide limited spatial resolution while MALDI-MSI instruments have been widely employed and extensively developed for high spatial resolution tissue imaging with excellent ionization efficiency for metabolites. Therefore, alternatives to conventional matrices have been sought to alleviate the matrix-related interferences. Based on the Bronsted-Lowry acid-base theory, which states that extremely basic or acidic organic molecules do not produce matrix-related ions, [45] a strong base DMAN proton sponge was examined for metabolite detection. Spectra devoid of matrix-related peaks were shown

in low mass regions, enabling matrix-free metabolite detection with routine MALDI-MSI operations. [46]

Herein, we investigated the utilization of the novel matrix DMAN for MALDI-MSI applications to study Medicago root and nodule metabolome during nitrogen fixation. Together with the conventional matrix DHB, metabolites of various chemical species, including amino acids, sugars, organic acids, lipids, flavonoids and their conjugates, were characterized and mapped on Medicago roots and nodules. In addition to wild-type (WT) plants and rhizobia, we utilized plant and bacterial mutants defective in nitrogen fixation to detect the metabolic differences relevant to nitrogen fixation, generating valuable information to understanding the underlying mechanism of nitrogen-fixing process.

4.2 Experimental Section

Plant growth and inoculation with rhizobia

Medicago seeds were acid scarified, surface sterilized and germinated as described.[47] One-day-old seedlings were placed on plates containing nitrogen-free modified Fahraeus medium [47] overlaid with sterile germination paper, and grown for 7 days at 22°C, 16h of 130 to 200 mmol m^{-2} s^{-1} light and 8 h of dark in a growth chamber.

The roots were grown in dark as the root portion was covered by aluminum foil. The roots were inoculated with *S. meliloti* Rm1021 (WT) and *fixJ* mutant with rhizobial suspension at OD_{600nm} value of 0.01, and incubated in the growth chamber for 3 weeks for nodule development. About 3 weeks old nodules were selected from each sample for metabolite imaging. Information of plant lines and bacterial strains are provided in *Supplemental Information* (**Method S1**). The experimental set up is shown in Figure S6.

Acetylene Reduction Assays (ARA)

Acetylene reduction assays [48] were performed to quantify the amount of nitrogen fixed in the nodules formed on Medicago roots. The details are described in (**Method S2**).

Tissue extraction

Nodulated roots were detached from the plant, flash-frozen, ground to powder in liquid nitrogen and extracted using acidified ice-cold methanol (92:7.9:0.1 MeOH:water:formic acid, v/v). The detailed procedure can be found in (**Method S3**).

Sample preparation for imaging

The nodules were excised along with flanking root tissues, embedded in 100 mg/ml gelatin aqueous solution and flash-frozen in liquid nitrogen. The frozen tissue was then sectioned into 12 μm slices at -20^{o}C. The 12 μm thick intact longitudinal

sections of Medicago nodules along with the root tissue were subjected to metabolite MSI. Corresponding methylene blue-stained nodule sections (**Figure 1a**) were used to correlate the MS images acquired. The sections for MSI were thaw-mounted either onto a MALDI plate or on ITO glass slides, depending on specific instruments. A representative optical image of a cross-section thaw-mounted on a MALDI plate is shown (**Figure 1b**). Subsequently, these sections were dehydrated, matrix coated and subjected to MS imaging/profiling as illustrated (**Figure 1c**). Different matrix application procedures were utilized for positive and negative mode MSI experiments. The details are described in (**Method S4**).

MALDI-TOF/TOF for MSI

A 4800 MALDI-tandem-time-of-flight (TOF/TOF) analyzer (Applied Biosystems, Framingham, MA) was equipped with a 200 Hz, 355 nm Nd:YAG laser (spot diameter 50 μm) for negative mode MS profiling and imaging. Positive mode MSI was acquired on an autoflex III MALDI-TOF/TOF (Bruker Daltonics, Billerica, MA, USA). The instrumental details can be found in (**Method S5**).

Image files obtained on 4800 TOF/TOF were processed, and extracted ion images were created using TissueView (Applied Biosystems). Image files obtained on autoflex III were processed using flexImaging (Bruker Daltonics).

ESI-Q-TOF

The direct infusion data was acquired by reconstituting the lyophilized Medicago root nodule extract, followed by injection on a UHR-Q-TOF (Bruker Daltonics, Billerica, MA, USA) in positive mode and negative mode, respectively. To reduce sample complexity and acquire tandem MS spectra for isomeric metabolites, the lyophilized Medicago root nodule extract was reconstituted in 0.1% formic acid and injected onto a Waters NanoAcquity Ultra Performance capillary LC system coupled to a ESI-Q-TOF SYNAPT G2 (Waters Corp., Milford, MA). The detailed experimental setup can be found in (**Method S6**).

Molecular identification of metabolites

Molecular identification of the metabolites involves comparing the masses acquired on MALDI-TOF/TOF from Medicago sections with those recorded on UHR-Q-TOF from Medicago extract and searching the matched mass against database. The metabolite list generated by mass-matching was further confirmed by comparing with purchased standards, database or previous literature. METLIN, KNApSAcK and the Human Metabolome Database were used. The detailed experimental setup can be found in (**Method S7**).

4.3 Results

In situ *profiling of metabolites in Medicago roots and nodules in the positive mode*

Medicago roots with nodules were sampled three weeks after inoculation with *S. meliloti*. The nodules were excised, embedded in gelatin and sectioned followed by matrix application prior to MSI analyses (Figure 1). To study the metabolic distribution in the roots and nodules, *in situ* profiling was performed in the positive mode on the cross-sections of Medicago root and nodule tissues. The "snapshots" of the metabolic profiles of the two distinct organs were compared with the MS profile obtained from the matrix alone (**Figure 2a**). The matrix-derived trace indicates that DHB gives rise to a number of abundant peaks, especially in the mass range of 150-400 Da, complicating metabolite detection within this range. Nevertheless, the matrix background is relatively clean outside this range, rendering the possibility to examine the endogenous metabolites present in Medicago. For instance, the usefulness of DHB was exemplified by the enlargement of the overlaid spectrum between *m/z* 100 and 160. A number of ions were detected above the matrix background (**Figure 2b**). Those masses were searched against the database and assigned to putative metabolites. The reliability of the peak assignments was thereby indicated as the absolute deviation of observed and theoretical monoisotopic mass values (Δm). Moreover, the assignments that agree with previous

metabolite studies in legumes were given higher fidelity. Multiple peaks in **Figure 2b** were assigned as amino acids with a threshold of Δm = 0.03 Da based on this method. For example, the *m/z* 118.09 and the *m/z* 132.10 ions were assigned as two non-polar amino acid, valine and leucine, respectively, both with Δm = 0. Several polar amino acids, including proline, asparagine and glutamine, were assigned to the peaks at *m/z* 116.07 (Δm = 0), 133.06 (Δm = 0) and 147.09 (Δm = 0.01 Da), respectively. In the legume-rhizobia symbiosis, free amino acids are not only the reduction products of atmospheric nitrogen catalyzed by nitrogenase in bacteroids, but are also involved in amino acid cycles to balance the mutualism. Therefore, it is not surprising to detect free amino acids, one major class of primary metabolites, appearing prominently in profiling spectra of Medicago roots and nodules. In addition, several other free amino acids, including arginine and histidine, were also detected among the matrix clusters due to their relatively high abundance in Medicago. Other than standard amino acids, an osmoprotectant derived from proline, proline betaine, was assigned to the peak at *m/z* 144.10. Besides amino acids, the peak at *m/z* 136.07 was identified as adenine, a nucleotide base with a variety of biological functions, such as respiration and protein synthesis. [49, 50] Moreover, its detection in nodulated alfalfa by MS reinforced our confidence in this assignment. [23]

Accurate mass and tandem mass spectrometry-assisted identification of metabolites

In order to confirm the mass-based identification obtained on MALDI-TOF/TOF, extracts of roots with nodules were directly infused into an electrospray ionization (ESI)- ultra high resolution-quadrupole-time-of-flight (UHR-Q-TOF) mass spectrometer (mass error <5 ppm). Besides the benefits of high resolution and accuracy provided by UHR-Q-TOF, the matrix-free ionization method, ESI, excludes the possibility of introducing interfering matrix peaks to the resulting spectrum, confirming the existence of the detected ions in the extract. The extracts were also separated and analyzed by coupling C_{18} reversed phase liquid chromatography (RPLC) with an ESI-Q-TOF for analysis of complex isobaric secondary metabolites in the extracts. As shown in **Table 1**, the high-accuracy mass measurements acquired on the UHR-Q-TOF confirmed the MALDI-TOF/TOF results, facilitating the molecular identification of metabolites. For example, the high mass resolution and accuracy of the UHR-Q-TOF was manifested by the differentiation of two close peaks at *m/z* 104.0706 and 104.1071 assigned respectively as gamma-aminobutyric acid (GABA) (Δm = 0 mDa) and choline (Δm = 0.1 mDa). In contrast, the modest resolution of MALDI-TOF/TOF delivered a convoluted peak at *m/z* 104.10, resulting a Δm = -0.03 Da for GABA and a Δm = 0.01 Da for choline. The presence of GABA could have been overlooked if merely relying on MALDI-TOF/TOF. Despite the interference from matrix ions in the

low mass region, DHB serves as an excellent matrix for MALDI MSI by promoting ionization of analyte molecules in the higher mass range. **Figure 2c** shows two abundant peaks at *m/z* 616.18 and 664.12 in the enlarged nodule-specific spectrum. Through a database search, the *m/z* 616.18 ion corresponds to heme and the *m/z* 664.10 ion is likely to be nicotinamide adenine dinucleotide (NAD). Unfortunately, one *m/z* value of the putative metabolite obtained on a moderate resolution MS could be close or equal to several potential candidate metabolites. Therefore, tandem mass (MS/MS) experiments were conducted on selected precursor ions for unequivocal assignment directly from the regions that were rich in the metabolites of interest by MALDI-TOF/TOF. **Figure 2d-f** demonstrates the utilization of *in situ* MS/MS for structural elucidation and, therefore, molecular identification. The fragments of the *m/z* 664.12 precursor ion sampled from Medicago nodule region matched well with the theoretical fragments of NAD, indicating that the majority of the *m/z* 664.10 peak correspond to NAD. Similarly, the MS/MS spectrum of the *m/z* 616.18 ion obtained from the nodule region gave rise to the fragments that correspond to sequential losses of $-CH_2COOH$ from heme [51] in **Figure 2e**. In **Figure 2f**, MS/MS again aided in the identification of the root-specific *m/z* 517.13 peak that was mass-matched to an isoflavonoid conjugate, formononetin glucoside malonate (MalGlc). Its signature fragment at *m/z* 269.08 agreed well with what was observed on an ESI-Q-TOF (**Figure**

S1a), increasing our confidence in MALDI-MS-based molecular assignments. **Figure S1b-c** show the tandem mass spectra of isoflavonoid formononetin and flavonoid conjugate aformonisin MalGlc at *m/z* 269.08 and 547.14 acquired via a conventional LC-MS/MS approach. The retention time and MS/MS spectra of putative metabolites observed on LC-ESI-Q-TOF further confirmed the identities assigned by MALDI-MS mass matching. In summary, the combined use of MALDI-MS with conventional metabolite characterization tool such as ESI-Q-TOF enables mapping of multiple metabolites from nodulated Medicago with unambiguous assignment (**Table 1**).

ClinProTools-processed profiling of root and nodule spectra

By comparing the root and nodule MS profiling, we clearly observed that a number of ions were differentially represented in the two regions, indicating that a characteristic metabolic pattern might exist for roots and nodules. In order to quickly extract the signature metabolites from the complex dataset which consists of 12 root and nodule samples (3 mass spectra acquired from each root and nodule regions derived from 4 Medicago sections), the raw spectra from the root and nodule regions were divided into two classes and loaded to ClinProTools (Bruker) sequentially. With ClinProTools, we were able to process and analyze raw spectra, including spectra recalibration and average spectra calculation, to pick peaks and to calculate peak statistics. Processed spectra of

the two loaded classes arranged in a "gel" view with nodules displayed at the bottom of the view and roots above it are shown in **Figure 3a**. Principal component analysis (PCA), a multivariate data mining technique, was then applied to spectra of the two classes. Peak masses and peak heights from each technical replicate were averaged and translated to a point (x, y, and z) in a 3D coordinate. Therefore, each point shown in the 3D PCA score plot represents a biological replicate averaged from three technical replicates, and each class consists of four points. The four biological replicates of nodules (red point) and the four biological replicates of roots (green point) segregate into two distinct groups as highlighted in the dashed circles (**Figure 3b**). To determine the significance of this segregation, an unsupervised clustering using a hierarchical clustering algorithm was then performed on PCA-transformed spectra and generated a dendrogram for all the data points (**Figure S2**). Two clades are shown, each corresponding to spectra with root or nodule origin, providing statistical support to the differentiation between roots and nodule metabolite spectra. Semi-quantitative results were obtained utilizing the peak statistics function embedded in ClinProTools. The peaks of *m/z* 137.02, 155.03, 177.02 and 214.97 that correspond to $[DHB\text{-}H_2O+H]^+$, $[DHB+H]^+$, $[DHB+Na]^+$ and $[DHB\text{-}H+K+Na]^+$, all show similar intensity between the nodule and root classes, demonstrating the ionization efficiency of the analytes from the roots and nodules are comparable. Additionally, the signal intensity of metabolites likely reflects

their presence in root and nodule regions and allows regional comparison. The metabolites showing distinct intensities in the two classes were thereby recorded, which potentially contributes to the metabolic differences between Medicago roots and nodules. As indicated in **Figure 3d**, representative metabolites at *m/z* 147.09, 381.05, 616.18 and 664.12 were abundant in nodules with significant differences ($p < 0.01$), whereas the intensities of the metabolites of *m/z* 517.12 and 849.47 were higher in roots than in nodules ($p < 0.01$ and 0.05, respectively). The identities of these marker metabolites were assigned as glutamine, the potassiated sucrose, heme, NAD, and formononetin MalGlc and a glycerophospholipid (GP) 36:5, respectively (**Table 1**).

Imaging of metabolites in Medicago roots and nodules in the positive mode

Although MALDI-MS-based metabolite profiling reveals differential distribution pattern of metabolites in Medicago roots and nodules, spatial resolution was sacrificed. To visualize the organ-specific patterns, MSI was performed on Medicago sections. Representative metabolites that exemplified distinct distribution patterns in roots and nodules are shown in **Figure 4a-h**. The molecule at *m/z* 616.18, previously assigned as heme, was localized in the nodule region and particularly concentrated in the fixation zone (**Figure 4a**). Moreover, the peak of *m/z* 664.12 identified as NAD was also abundant in the nodule region (**Figure 4b**). This observation agrees with the higher

abundance of these two peaks in nodules (**Figure 3d**), demonstrating the consistency between profiling and imaging results. On the other hand, the *m/z* 144.10 peak previously identified as proline betaine was more abundant in roots than in nodules (**Figure 4c**). In addition, the ion of *m/z* 849.47, a putative GP(36:5), displays a ubiquitous distribution pattern over the entire root/nodule section, yet exhibiting relatively higher abundance in the root (**Figure 4d**).

An interesting class of secondary metabolites, flavonoids and their conjugates, were mapped with high resolution in nodulated Medicago. The *m/z* 269.08 ion assigned as formononetin exhibited ubiquitous distribution on the entire section with high abundance in the root region (**Figure 4e**), whereas its gluco-conjugate, formononetin MalGlc, at *m/z* 517.12 also exhibited higher abundance in roots compared to nodules (**Figure 4f**). Nevertheless, flavonoids and their conjugates are a group of diverse molecules and have various functions in legume nodulation. Therefore, it is not surprising to find that the ion assigned as afrormosin MalGlc at *m/z* 547.16 exhibited a distinct localization from formononetin MalGlc. The metabolites that defined the nodule and root regions by localization were overlaid for visual contrast in **Figure 4h**, exhibiting heme (*m/z* 616.18) in red and formononetin MalGlc (*m/z* 517.13) in green. PCA was applied to the MSI results, showing significant differences between nodule and root in **Figure 4i**. **Figure 4j** shows the loading plots of PC1 and PC2 The

highlighted points (**Figure 4j**) corresponding to the peaks at *m/z* 547.16, 616.18 and 664.12 contribute significantly to the segregation of nodule and root regions on PC2, agreeing well with their nodule-concentrated localization. By performing MSI with corroborative statistical analyses, detailed localization of metabolites was obtained without prior knowledge, enabling dissection of nodule chemistry and related mechanistic studies.

Although metabolic differences of the root and nodule regions were visualized by MSI in Medicago, identifying the metabolites which are differentially regulated in fixing nodules over non- fixing nodules can further facilitate our investigation of key metabolites involved in nitrogen fixation. Herein, we utilized mutants which are defective in nitrogen fixation and performed MSI on the sections of the WT and mutant plants, respectively. Medicago *defective in nitrogen fixation 1* (*dnf1*) mutants are affected in bacteroid and symbiosome development, hence *dnf1* mutants form non-functional nodules which are unable to fix nitrogen. [52, 53] Similarly, *fixJ* mutants of *S. meliloti* can differentiate into elongated bacteroids, but quickly degenerate leading to nodule senescence. These *fixJ* mutants cannot encode nitrogenase enzyme, hence forming non-fixing nodules. Thus, we utilized WT plants and WT rhizobia in combination with mutant plant and rhizobia (*dnf1-1* and *fixJ*) to conduct differential MSI experiments. Nitrogen fixing nodules formed on the WT plants interacting with WT

rhizobia were pinkish and elongated (**Figure 5a**), while the non-fixing nodules formed on WT/*fixJ*, *dnf1*/WT and *dnf1*/*fixJ* were pale, small and more globular in shape (**Figure 5b-d**). No significant difference was observed in the rate of nodulation and amount of nodule tissue per root in three of the four combinations of plants and rhizobia tested (WT/WT, WT/*fixJ* and *dnf1*/WT) (**Figure S3**). As expected, the nodules formed on WT/WT interaction showed high nitrogen fixation activity in the acetylene reduction assay, whereas nodules formed on the WT/*fixJ*, *dnf1*/WT, *dnf1*/*fixJ* showed negligible to nil nitrogen fixation activity (**Figure 5e)**. Distinct metabolite distributions were observed in the nodules formed on WT/WT interaction over the nodules formed on other combinations (WT/*fixJ*, *dnf1*/WT, *dnf1*/*fixJ*; **Figure 5f-i**). The *m/z* 616.18 ion corresponding to the heme-moiety, possibly of leghemoglobin, was present at high abundance in the fixation zone of nodules formed on WT/WT samples, whereas it was completely absent from the nitrogen fixation defective nodules. In contrast, the ion at *m/z* 517.13, formononetin MalGlc, was particularly concentrated in the roots of WT/WT Medicago, showing negligible differences with the other mutant samples (**Figure 5j-5m**).

Profiling and imaging of metabolites in Medicago roots and nodules using DMAN in the negative mode

Although DHB has manifested its great capability in ionizing metabolites in the positive mode, the high-abundance matrix-derived peaks made endogenous metabolite detection more challenging. To overcome this problem, we utilized DMAN and developed its application to metabolite imaging in the negative mode. Since the ionization efficiency of DMAN has been shown to be effective using metabolite standards (**Figure S4**), its utility for MALDI-MSI on biological tissues was investigated. Spectra acquired from the nodule region, the root region and matrix alone using DMAN were compared, revealing numerous putative Medicago-specific metabolites with clean background (**Figure 6**). Based on the mass matching and MS/MS-assisted identification approach described previously, different classes of molecules were identified, including 6 amino acids, 14 organic acids and 4 carbohydrate-containing compounds. In contrast to the amino acids identified in the positive mode, negatively charged side chain groups, such as aspartate at *m/z* 132.03 and glutamate at *m/z* 146.01 were identified. Several other amino acids with relatively low isoelectric points (p*I*) were also detected, including alanine, serine and threonine. In addition, GABA was detected in the negative mode with no interference, circumventing the issues of potential misidentification and overlapping with choline in the positive mode. All of the amino acid masses obtained on UHR-Q-TOF were within Δ m = 1.5 mDa. In addition to amino acids, organic acids are critical for nitrogen fixation, serving as a source of carbon and energy delivered from

the host plant cells to the bacteroids. Multiple organic acids were detected using DMAN rather than DHB due to its high proton affinity. As detailed in **Table 2**, a significant number of organic acids were detected in the negative mode, greatly expanding the metabolite coverage compared with DHB alone. Carbohydrates are another class of molecules pivotal to the mutual exchange of nutrients between legumes and rhizobia. In positive mode, sucrose is prone to be detected in its sodiated form at *m/z* 365.11 and potassiated form at *m/z* 381.08. In contrast, sucrose was detected at *m/z* 341.07 (Δ m = -0.04 Da) in its deprotonated form in negative mode, providing a straightforward evidence for its presence in Medicago roots and nodules. Moreover, ions corresponding to the hexoses (glucose and fructose) were observed at *m/z* 179.05 (Δ m = -0.01 Da) in its deprotonated form. The ion at *m/z* 149.04 was identified as pentose with the accurate mass of 149.0439 (Δm = -0.6 mDa) further provided by the UHR-Q-TOF. Sugar phosphates are also used in biological systems to store or transfer energy. For example, hexose-6-phosphate was deprotonated and gave rise to a peak at *m/z* 259.04 in the negative mode using DMAN. Its detection on UHR-Q-TOF further increased the confidence of assignment.

Similar as the positive-mode workflow, MS/MS fragmentation patterns aid in metabolite identification (Figure 7). The presence of diagnostic fragment peaks at *m/z* 27 and 71 shown in the top panel of **Figure 7a** suggest an identity of maleic acid when

being compared to the MS/MS fragments of the standard in the bottom panel. Similarly, the major fragments ions at *m/z* 59, 87, 129 and 173 from the citric acid standard and the peaks at *m/z* 59, 71, 89, 119 and 179 from the sucrose standard confirmed the identities of the metabolites *in situ* fragmented from the Medicago sections as in **Figure 7b-c**. The MS/MS spectra confirm the metabolite identification based on accurate masses. Moreover, the utility of DMAN in generating MS/MS spectra of quality for metabolite identification is also demonstrated.

DMAN has herein been applied to MSI of metabolites that cannot be achieved otherwise. **Figure 8** shows representative MS images of deprotonated metabolite peaks, including maleic acid (*m/z* 115.03), citric acid (*m/z* 191.02) and sucrose (*m/z* 341.07). Maleic acid exhibited more universal distribution over the entire section with higher concentration in the nitrogen-fixation zone, whereas citric acid and sucrose were localized more to the root area and the fixation zone, yet displayed the highest concentration at the meristematic zone.

4.4 Discussion

To date, MALDI-MSI has become one of the most promising imaging techniques with ever-increasing sensitivity and specificity. Previous applications of

MSI were dominated by the investigation of large molecules, until recently MALDI-MSI of metabolites has gained continued interest in plant biology. Herein, we have adapted MSI to describe the metabolome in Medicago roots and nodules. A number of metabolites were identified based on accurate mass measurements and MS/MS fragmentation, including amino acids, carbohydrates, organic acids and flavonoids family. With statistical support by ClinProTools, we compared the differential abundances of metabolites in Medicago root and nodule. For example, glutamine and sucrose showed significantly higher intensities in Medicago nodule. Interestingly, an independent study using GC-MS in *Medicago sativa* also reported significant increase in the abundance of glutamine and sucrose in nodules compared to roots using GC-MS. [23] Sucrose is the primary source of carbon energy synthesized in the leaves and exported to sinks like nodules via phloem, where it is metabolized by glycolytic pathway into phosphoenol pyruvate or malate, the likely substrates of bacteroids to fuel biological nitrogen fixation (conversion of atmospheric dinitrogen into ammonia). The fixed ammonia is assimilated into glutamine which is subsequently utilized in the synthesis of amides (in the case of Medicago) and ureides, and transported via xylem to the aerial part. [54] Since nodules are considered as the main nitrogen source of nodulated legume plants, the elevated levels of glutamine together with other amino acids probably reflects the efficient nitrogen metabolism occurring inside the nodule. The power of MALDI-MSI in

identifying region-specific distribution of metabolites was further supported with the detection of heme-moiety in the mature fixing nodules, where heme-moiety constitutes a key component of both leghemoglobin of plant origin and FixL of rhizobial origin. [5, 55, 56] Leghemoglobins are present in legume nodules at concentrations of 2-3 mM in the host cell cytoplasm and help maintaining free oxygen concentrations around 20-40 nM in the cytosol. [57] Such microaerobic conditions allow an adequate supply of ATP through respiration for nitrogen fixation but prevents inactivation of nitrogenase complex [58]. By comparing MS images obtained from nodulated roots of Medicago (WT and *dnf1*) inoculated with rhizobia (WT and *fixJ*), we showed that the heme is exclusively present in fixing nodules formed on WT plants inoculated with WT rhizobia. In contrast, nodules formed by plant or bacterial mutants were non- fixing and were devoid of the heme moiety. In line with our observations, Starker et al. [59] observed a complete absence of nitrogen fixation in *dnf1-1* mutant compared to the WT plants. Medicago *dnf1* mutants form infection threads similar to WT plants, but bacteroid differentiation is arrested at stage 1. [53, 59] Similarly, *fixJ* rhizobia mutants cannot make the nitrogenase enzyme, but are released into plant cells where they differentiate into elongated bacteroids. However, these bacteria typically degenerate and the nodule senesces prematurely. [52, 60] In comparison, the heme-moiety, possibly of leghemoglobin was found to be highly enriched in the nitrogen fixation zone (**Figure 4a**), where the

differentiation of bacteroids occurs in the fixing nodules. In accordance with our observation, it was shown that the transcript for coproporphyrinogen III oxidase, one of the terminal enzymes of heme synthesis, is highly elevated in soybean and pea nodules compared with roots, particularly in the infected cells of the nodules where the protein moiety of leghemoglobin is also synthesized. [61]

It is also possible that the heme-moiety may belong to FixL of rhizobial origin, an oxygen-binding hemoprotein with kinase and phosphatase activities, which senses the oxygen levels directly and transmits this signal to FixJ via phosphorylation-dephosphorylation reactions. [5] FixJ controls the expression of other regulatory genes, including *nifA* and *fixK*, that regulate the transcription of genes required for symbiotic nitrogen fixation. [62] With the current techniques and knowledge, we could not differentiate the heme-moiety of leghemoglobin from that of FixL. Proline and glycine betaines are nitrogenous osmolytes accumulated under osmotic stress conditions in plants. [63, 64] It was shown that *S. meliloti* uses proline betaine as a carbon and nitrogen source or as an osmoprotectant. [65, 66] Given the competition among soil bacteria for plant carbon source, the ability of *S. meliloti* to utilize proline betaine provides selective advantage in the colonization of legume roots. [67] In addition, proline betaine can also act as an inducer of nodulation (*nod*) genes in *S. meliloti*. [67] Our MALDI-MSI strategy identified a strong accumulation of proline betaine in root tissues

compared to nodules. It is possible that higher rate of proline betaine catabolism (as an energy source) by *S. meliloti* inside the nodule may be responsible for the lower proline betaine content in nodule than in root tissues.

Several metabolites, like NAD or adenine, were also detected using DHB in the positive mode. However, these ions were not observed on the UHR-Q-TOF employed for this study. Despite the high concentration of these ions in nodules, their dilution in the homogenized root tissues probably prevented us from detecting them on an ESI-MS instrument. Therefore, the *in situ* tandem MS spectra of selected metabolites acquired on MALDI-TOF/TOF are critical to the determination of their identities. A comparison between the MS/MS spectra acquired on MALDI-TOF/TOF and conventional LC-ESI-Q-TOF displayed identical fragments as in **Figure 2f** and **Figure S1a**, increasing our confidence in assignments based on MALDI-TOF/TOF-generated spectra. Moreover, the MALDI-TOF/TOF-produced MS/MS fragments could all be well-matched to the theoretical structures (Figure 2), further supporting their identifications.

Flavonoids are polyphenolics constitute one of the most diverse classes of metabolites in higher plants. Here, we directly localized the distribution of flavonoids and their conjugates, formononetin and formononetin MalGlc, in Medicago roots and nodules. Flavonoids and isoflavonoids play essential roles at different steps of legume nodulation from the recruitment of compatible rhizobia to the regulation of nodule

development [68]. Flavonoids and isoflavonoids are released by the host plants which act as chemo-attractants for the rhizobia to recognize their host. [69-71] These compounds activate the expression of several *nod* genes in rhizobia leading to Nod factor production and secretion. [72, 73] Localized inhibition of auxin transport mediated by flavonoids is also necessary for nodule development in the roots of legumes with indeterminate nodules, such as Medicago. [68] Auxin transport is arrested at the site of infection by rhizobia, which is mediated by flavonoids. [74] RNAi silencing of the isoflavone biosynthesis pathway resulted in increased auxin transport and defective nodule formation. [75] Although the function of isoflavonoid conjugates remains uncertain [76], the higher level of afrormosin MalGlc in nodules compared to roots suggest a distinct role from formononetin MalGlc in Medicago.

Although conventional methods might provide a good coverage of metabolites detected from tissue extracts, MSI enables visualization of metabolites through direct analysis of tissue sections, as demonstrated here with Medicago roots and nodules. In this study, we pioneered the use of a novel matrix, DMAN, in exploring the metabolome of plant tissues and showed an excellent compatibility with MS/MS and MSI. The interference-free matrix enabled direct mapping of acidic metabolites by MALDI-TOF/TOF in the negative mode, significantly expanding the coverage of the metabolites detected. A representative acidic molecule is malic acid, which was only

detected in the negative mode using DMAN and is involved in the tricarboxylic acid (TCA) cycle, glyoxylate and dicarboxylate metabolic pathway and carbon fixation in photosynthetic organisms. This dicarboxylic acid is also the major energy source for the bacteroids and is used as carbon skeleton in the glutamine synthetase/glutamate synthase pathway. [77, 78] Another organic acid, pyruvic acid, is the end product of glycolysis and the starting point of gluconeogenesis. Similarly, succinic acid was found to be an excellent substrate that supports the highest level of nitrogen fixation by isolated bacteroids. [79, 80] Other than being intermediates in the key metabolic pathways, organic acids have also been proposed to be involved in many processes, including nutrient acquisition and uptake, metal detoxification, alleviation of anaerobic stress in roots, mineral weathering, and interaction with microbes in the rhizosphere. [81] Actually, most key organic acids involved in TCA cycle have been detected using DMAN, as illustrated in **Supplemental Figure S5**. This observation will help our understanding not only of the legume-rhizobia symbiosis but also of many other metabolic processes in plants [46]. Nevertheless, the MSI-based approach often lacks the selectivity to separate the metabolites desorbed simultaneously from Medicago tissue. Therefore, caution must be taken when evaluating the distribution images of metabolites that have isobaric ions, such as sucrose and anhydrous trehalose. Given their identical *m/z* and highly similar CID fragmentation patterns, it is possible that both species were observed as one peak with

MALDI-TOF/TOF in this study. Nevertheless, these two isobaric disaccharide species can be distinguished by a combination of ion mobility spectrometry and vacuum ultraviolet photodissociation techniques as recently reported by Clemmer's group. [82] In future investigation, we will employ ion mobility spectrometry to further separate these structural isomers to allow definitive characterization. Besides isomers, isotopic contribution of metabolites with adjacent masses could also result in merging of two species into one peak such as the second isotopic peak of oxalacetate $[C_4H_4O_5\text{-}H]^-$ at *m/z* 132.0009 and the third isotopic peak of aspartate $[C_4H_7NO_4\text{-}H]^-$ at *m/z* 132.0291 on MALDI-TOF/TOF, whereas a high resolution MALDI-MS instrument could fully resolve the nearly isobaric peaks and produce their corresponding distribution images unambiguously. Exploration of the metabolic differences on nodule chemistry via MALDI-MSI will be continued and attempted with high-resolution MALDI-MS in the future.

4.5 Conclusions

We have demonstrated the benefits of using MALDI-MSI in obtaining unique and valuable information on the identity and spatial distribution of plant metabolites that cannot be accomplished otherwise. The conventional matrix DHB and the novel matrix DMAN were complementary in profiling and imaging metabolites directly from tissue

sections. Furthermore, MS/MS fragmentations were performed to facilitate the identification of region-specific metabolites. The knowledge gained through the comparison of metabolite profiles and molecular ion images obtained from nitrogen fixing and non-fixing nodules highlighted the benefits of MALDI-MSI in understanding metabolites' roles in symbiotic nitrogen-fixation.

Acknowledgement

This work was supported by a grant from the National Science Foundation (NSF#0701846) to J.M.A and by funding from the University of Wisconsin Graduate School and the Wisconsin Alumni Research Foundation (WARF) and Romnes Faculty Research Fellowship program to L.L.

4.6 References

1. Graham, P. H.; Vance, C. P. (2003) Legumes: importance and constraints to greater use. *Plant Physiol.* 131, 872-877.
2. Zahran, H. H. (1999) *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* 63, 968-989.
3. Venkateshwaran, M.; Ané, J. M., Legumes and nitrogen fixation: physiological, molecular, evolutionary perspective and applications. In *In The Molecular Basis of Nutrient Use Efficiency in Crops*, Hawkesford, M.; Barraclough, P., Eds. John Wiley & Sons.: 2011; pp 443-89.
4. Masson-Boivin, C.; Giraud, E.; Perret, X.; Batut, J. (2009) Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol.* 17, 458-466.
5. Gilles-Gonzalez, M. A.; Ditta, G. S.; Helinski, D. R. (1991) A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. *Nature* 350, 170-2.

6. Ott, T.; Sullivan, J.; James, E. K.; Flemetakis, E.; Gunther, C.; Gibon, Y.; Ronson, C.; Udvardi, M. (2009) Absence of symbiotic leghemoglobins alters bacteroid and plant cell differentiation during development of *Lotus japonicus* root nodules. *Mol. Plant Microbe In.* 22, 800-808.

7. Udvardi, M. K.; Day, D. A. (1997) Metabolite transport across symbiotic membranes of legume nodules. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48, 493-523.

8. Penmetsa, R. V.; Cook, D. R. (2000) Production and characterization of diverse developmental mutants of *Medicago truncatula*. *Plant Physiol.* 123, 1387-1397.

9. Benloch, R.; Navarro, C.; Beltran, J. P.; Canas, L. A. (2003) Floral development of the model legume *Medicago truncatula*: ontogeny studies as a tool to better characterize homeotic mutations. *Sex. Plant Reprod.* 15, 231-241.

10. Branca, A.; Paape, T. D.; Zhou, P.; Briskine, R.; Farmer, A. D.; Mudge, J.; Bharti, A. K.; Woodward, J. E.; May, G. D.; Gentzbittel, L.; Ben, C.; Denny, R.; Sadowsky, M. J.; Ronfort, J.; Bataillon, T.; Young, N. D.; Tiffin, P. (2011) Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proc. Natl. Acad. Sci. USA* 108, E864-E870.

11. Gallardo, K.; Le Signor, C.; Vandekerckhove, J.; Thompson, R. D.; Burstin, J. (2003) Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol.* 133, 664-682.

12. Samac, D. A.; Peñuela, S.; Schnurr, J. A.; Hunt, E. N.; Foster-Hartnett, D.; Vandenbosch, K. A.; Gantt, J. S. (2011) Expression of coordinately regulated defence response genes and analysis of their role in disease resistance in *Medicago truncatula*. *Mol. Plant Pathol.* 12, 786-798.

13. Wang, H. L.; Chen, J. H.; Wen, J. Q.; Tadege, M.; Li, G. M.; Liu, Y.; Mysore, K. S.; Ratet, P.; Chen, R. J. (2008) Control of compound leaf development by FLORICAULA/LEAFY ortholog SINGLE LEAFLET1 in *Medicago truncatula*. *Plant Physiol.* 146, 1759-1772.

14. Benedito, V. A.; Torres-Jerez, I.; Murray, J. D.; Andriankaja, A.; Allen, S.; Kakar, K.; Wandrey, M.; Verdier, J.; Zuber, H.; Ott, T.; Moreau, S.; Niebel, A.; Frickey, T.; Weiller, G.; He, J.; Dai, X.; Zhao, P. X.; Tang, Y.; Udvardi, M. K. (2008) A gene expression atlas of the model legume *Medicago truncatula*. *Plant J.* 55, 504-13.

15. Grimsrud, P. A.; den Os, D.; Wenger, C. D.; Swaney, D. L.; Schwartz, D.; Sussman, M. R.; Ané, J. M.; Coon, J. J. (2010) Large-scale phosphoprotein analysis in *Medicago truncatula* roots provides insight into *in vivo* kinase activity in legumes. *Plant Physiol.* 152, 19-28.

16. Rose, C. M.; Venkateshwaran, M.; Grimsrud, P. A.; Westphall, M. S.; Sussman, M. R.; Coon, J. J.; Ané, J. M. (2012) Medicago PhosphoProtein Database: a repository for *Medicago truncatula* phosphoprotein data. *Front. Plant Sci.* 3, 122.

17. Broeckling, C. D.; Huhman, D. V.; Farag, M. A.; Smith, J. T.; May, G. D.; Mendes, P.; Dixon, R. A.; Sumner, L. W. (2005) Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. *J. Exp. Bot.* 56, 323-36.

18. Farag, M. A.; Huhman, D. V.; Dixon, R. A.; Sumner, L. W. (2008) Metabolomics reveals novel pathways and differential mechanistic and elicitor-specific responses in phenylpropanoid and isoflavonoid biosynthesis in *Medicago truncatula* cell cultures. *Plant Physiol.* 146, 387-402.

19. Farag, M. A.; Huhman, D. V.; Lei, Z. T.; Sumner, L. W. (2007) Metabolic profiling and systematic identification of flavonoids and isoflavonoids in roots and cell suspension cultures of *Medicago truncatula* using HPLC-UV-ESI-MS and GC-MS. *Phytochemistry* 68, 342-354.

20. Suzuki, H.; Reddy, M. S.; Naoumkina, M.; Aziz, N.; May, G. D.; Huhman, D. V.; Sumner, L. W.; Blount, J. W.; Mendes, P.; Dixon, R. A. (2005) Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. *Planta* 220, 696-707.

21. Zhang, N.; Venkateshwaran, M.; Boersma, M.; Harms, A.; Howes-Podoll, M.; den Os, D.; Ané, J. M.; Sussman, M. R. (2012) Metabolomic profiling reveals suppression of oxylipin biosynthesis during the early stages of legume-rhizobia symbiosis. *FEBS Lett.* 586, 3150-8.

22. Desbrosses, G. G.; Kopka, J.; Udvardi, M. K. (2005) *Lotus japonicus* metabolic profiling. Development of gas chromatography-mass spectrometry resources for the study of plant-microbe interactions. *Plant Physiol.* 137, 1302-1318.

23. Barsch, A.; Tellström, V.; Patschkowski, T.; Küster, H.; Niehaus, K. (2006) Metabolite profiles of nodulated alfalfa plants indicate that distinct stages of nodule organogenesis are accompanied by global physiological adaptations. *Mol. Plant Microbe In.* 19, 998-1013.

24. Edwards, E. L.; Rodrigues, J. A.; Ferreira, J.; Goodall, D. M.; Rauter, A. P.; Justino, J.; Thomas-Oates, J. (2006) Capillary electrophoresis-mass spectrometry characterisation of secondary metabolites from the antihyperglycaemic plant *Genista tenera*. *Electrophoresis* 27, 2164-2170.

25. Harada, K.; Fukusaki, E. (2009) Profiling of primary metabolite by means of capillary electrophoresis-mass spectrometry and its application for plant science. *Plant Biotechnol.* 26, 47-52.

26. Kueger, S.; Steinhauser, D.; Willmitzer, L.; Giavalisco, P. (2012) High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *Plant J.* 70, 39-50.

27. Colebatch, G.; Desbrosses, G.; Ott, T.; Krusell, L.; Montanari, O.; Kloska, S.; Kopka, J.; Udvardi, M. K. (2004) Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. *Plant J.* 39, 487-512.

28. Barsch, A.; Carvalho, H. G.; Cullimore, J. V.; Niehaus, K. (2006) GC-MS based metabolite profiling implies three interdependent ways of ammonium assimilation in *Medicago truncatula* root nodules. *J Biotechnol.* 127, 79-83.

29. Brechenmacher, L.; Lei, Z.; Libault, M.; Findley, S.; Sugawara, M.; Sadowsky, M. J.; Sumner, L. W.; Stacey, G. (2010) Soybean metabolites regulated in root hairs in response to the symbiotic bacterium *Bradyrhizobium japonicum*. *Plant Physiol.* 153, 1808-22.

30. Lee, Y. J.; Perdian, D. C.; Song, Z. H.; Yeung, E. S.; Nikolau, B. J. (2012) Use of mass spectrometry for imaging metabolites in plants. *Plant J.* 70, 81-95.

31. Stoeckli, M.; Chaurand, P.; Hallahan, D. E.; Caprioli, R. M. (2001) Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* 7, 493-496.

32. Kaspar, S.; Peukert, M.; Svatoš, A.; Matros, A.; Mock, H. P. (2011) MALDI-imaging mass spectrometry - An emerging technique in plant biology. *Proteomics* 11, 1840-1850.

33. Hölscher, D.; Shroff, R.; Knop, K.; Gottschaldt, M.; Crecelius, A.; Schneider, B.; Heckel, D. G.; Schubert, U. S.; Svatoš, A. (2009) Matrix-free UV-laser desorption/ionization (LDI) mass spectrometric imaging at the single-cell level: distribution of secondary metabolites of Arabidopsis thaliana and Hypericum species. *Plant J.* 60, 907-918.

34. Jun, J. H.; Song, Z.; Liu, Z.; Nikolau, B. J.; Yeung, E. S.; Lee, Y. J. (2010) High-spatial and high-mass resolution imaging of surface metabolites of *Arabidopsis thaliana* by laser desorption-ionization mass spectrometry using colloidal silver. *Anal. Chem.* 82, 3255-65.

35. Northen, T. R.; Yanes, O.; Northen, M. T.; Marrinucci, D.; Uritboonthai, W.; Apon, J.; Golledge, S. L.; Nordström, A.; Siuzdak, G. (2007) Clathrate nanostructures for mass spectrometry. *Nature* 449, 1033-U3.

36. Cha, S. W.; Zhang, H.; Ilarslan, H. I.; Wurtele, E. S.; Brachova, L.; Nikolau, B. J.; Yeung, E. S. (2008) Direct profiling and imaging of plant metabolites in intact tissues by using colloidal graphite-assisted laser desorption ionization mass spectrometry. *Plant J.* 55, 348-360.

37. Woo, H. K.; Northen, T. R.; Yanes, O.; Siuzdak, G. (2008) Nanostructure-initiator mass spectrometry: a protocol for preparing and applying NIMS surfaces for high-sensitivity mass analysis. *Nat. Protoc.* 3, 1341-1349.

38. Pachuta, S. J.; Cooks, R. G. (1987) Mechanisms in molecular SIMS. *Chem Rev.* 87, 647-669.

39. Li, Y.; Shrestha, B.; Vertes, A. (2007) Atmospheric pressure molecular imaging by infrared MALDI mass spectrometry. *Anal. Chem.* 79, 523-532.

40. Li, Y.; Shrestha, B.; Vertes, A. (2008) Atmospheric pressure infrared MALDI imaging mass spectrometry for plant metabolomics. *Anal. Chem.* 80, 407-420.

41. Nemes, P.; Vertes, A. (2007) Laser ablation electrospray ionization for atmospheric pressure, *in vivo*, and imaging mass spectrometry. *Anal. Chem.* 79, 8098-8106.

42. Vaclavik, L.; Cajka, T.; Hrbek, V.; Hajslova, J. (2009) Ambient mass spectrometry employing direct analysis in real time (DART) ion source for olive oil quality and authenticity assessment. *Anal. Chim. Acta* 645, 56-63.

43. Takáts, Z.; Wiseman, J. M.; Cooks, R. G. (2005) Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *J Mass Spectrom.* 40, 1261-1275.

44. Wiseman, J. M.; Ifa, D. R.; Song, Q. Y.; Cooks, R. G. (2006) Tissue imaging at atmospheric pressure using desorption electrospray ionization (DESI) mass spectrometry. *Angew. Chem. Int. Edit. Engl.* 45, 7188-7192.

45. Shroff, R.; Svatoš, A. (2009) Proton sponge: a novel and versatile MALDI matrix for the analysis of metabolites using mass spectrometry. *Anal. Chem.* 81, 7954-7959.

46. Shroff, R.; Rulisek, L.; Doubsky, J.; Svatoš, A. (2009) Acid-base-driven matrix-assisted mass spectrometry for targeted metabolomics. *Proc. Nat. Acad. Sci. USA* 106, 10092-10096.

47. Catoira, R.; Galera, C.; de Billy, F.; Penmetsa, R. V.; Journet, E. P.; Maillet, F.; Rosenberg, C.; Cook, D.; Gough, C.; Denarie, J. (2000) Four genes of *Medicago truncatula* controlling components of a Nod factor transduction pathway. *Plant Cell* 12, 1647-66.

48. Turner, G. L.; Gibson, A. H., Measurement of nitrogen fixation by indirect means. In *Methods for Evaluating Biological Nitrogen Fixation*, Bergersen, F. J., Ed. John Wiley & Sons: Chichester, UK, 1980; pp 111-138.

49. Lodwig, E. M.; Hosie, A. H. F.; Bourdès, A.; Findlay, K.; Allaway, D.; Karunakaran, R.; Downie, J. A.; Poole, P. S. (2003) Amino-acid cycling drives nitrogen fixation in the legume-*Rhizobium* symbiosis. *Nature* 422, 722-726.

50. Borowiec, A.; Lechward, K.; Tkacz-Stachowska, K.; Skladanowski, A. C. (2006) Adenosine as a metabolic regulator of tissue function: production of adenosine by cytoplasmic 5'-nucleotidases. *Acta Biochim. Pol.* 53, 269-278.

51. Shimma, S.; Seto, M. (2007) Mass microscopy to reveal distinct localization of Heme B (*m/z* 616) in colon cancer liver metastasis. *J. Mass Spectrom. Soc. Jpn.* 55, 145-148.

52. Mitra, R. M.; Long, S. R. (2004) Plant and bacterial symbiotic mutants define three transcriptionally distinct stages in the development of the *Medicago truncatula*/*Sinorhizobium meliloti* symbiosis. *Plant Physiol.* 134, 595-604.

53. Wang, D.; Griffitts, J.; Starker, C.; Fedorova, E.; Limpens, E.; Ivanov, S.; Bisseling, T.; Long, S. (2010) A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* 327, 1126-9.

54. Gordon, A. J.; Thomas, B. J.; James, C. L. (1995) The location of sucrose synthase in root nodules of white clover. *New Phytol.* 130, 523-530.

55. Garrocho-Villegas, V.; Gopalasubramaniam, S. K.; Arredondo-Peter, R. (2007) Plant hemoglobins: what we know six decades after their discovery. *Gene* 398, 78-85.

56. Lois, A. F.; Ditta, G. S.; Helinski, D. R. (1993) The oxygen sensor FixL of *Rhizobium Meliloti* is a membrane protein containing four possible transmembrane segments. *J. Bacteriol.* 175, 1103-1109.

57. Becana, M.; Klucas, R. V. (1992) Oxidation and reduction of leghemoglobin in root nodules of leguminous plants. *Plant Physiol.* 98, 1217-21.

58. Bergersen, F. J.; Kennedy, C.; Hill, S. (1982) Influence of low oxygen concentration on derepression of nitrogenase in *Klebsiella pneumoniae*. *J. Gen. Microbiol.* 128, 909-915.

59. Starker, C. G.; Parra-Colmenares, A. L.; Smith, L.; Mitra, R. M.; Long, S. R. (2006) Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression. *Plant Physiol.* 140, 671-680.

60. Vasse, J.; de Billy, F.; Camut, S.; Truchet, G. (1990) Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* 172, 4295-4306.

61. Santana, M. A.; Pihakaski-Maunsbach, K.; Sandal, N.; Marcker, K. A.; Smith, A. G. (1998) Evidence that the plant host synthesizes the heme moiety of leghemoglobin in root nodules. *Plant Physiol.* 116, 1259-1269.

62. Reyrat, J. M.; David, M.; Blonski, C.; Boistard, P.; Batut, J. (1993) Oxygen-regulated in vitro transcription of *Rhizobium meliloti nifA* and *fixK* genes. *J. Bacteriol.* 175, 6867-72.

63. Hasegawa, P. M.; Bressan, R. A.; Zhu, J. K.; Bohnert, H. J. (2000) Plant cellular and molecular responses to high salinity. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51, 463-499.

64. Trinchant, J. C.; Boscari, A.; Spermato, G.; Van de Sype, G.; Le Rudulier, D. (2004) Proline betaine accumulation and metabolism in alfalfa plants under sodium chloride stress. Exploring its compartmentalization in nodules. *Plant Physiol.* 135, 1583-1594.

65. Gloux, K.; Lerudulier, D. (1989) Transport and catabolism of proline betaine in salt-stressed *Rhizobium meliloti*. *Arch. Microbiol.* 151, 143-148.

66. Goldmann, A.; Boivin, C.; Fleury, V.; Message, B.; Lecoeur, L.; Maille, M.; Tepfer, D. (1991) Betaine use by rhizosphere bacteria: genes essential for trigonelline, stachydrine, and carnitine catabolism in *Rhizobium meliloti* are located on pSym in the symbiotic region. *Mol. Plant Microbe In.* 4, 571-578.

67. Phillips, D. A.; Joseph, C. M.; Maxwell, C. A. (1992) Trigonelline and stachydrine released from alfalfa seeds activate NodD2 protein in *Rhizobium meliloti*. *Plant Physiol.* 99, 1526-1531.

68. Subramanian, S.; Stacey, G.; Yu, O. (2007) Distinct, crucial roles of flavonoids during legume nodulation. *Trends Plant Sci.* 12, 282-285.

69. Peters, N. K.; Frost, J. W.; Long, S. R. (1986) A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* 233, 977-980.

70. Maxwell, C. A.; Hartwig, U. A.; Joseph, C. M.; Phillips, D. A. (1989) A chalcone and 2 related flavonoids released from alfalfa roots induce *nod* genes of *Rhizobium meliloti*. *Plant Physiol.* 91, 842-847.

71. Hartwig, U. A.; Maxwell, C. A.; Joseph, C. M.; Phillips, D. A. (1990) Effects of alfalfa *nod* gene-inducing flavonoids on *nodABC* transcription in *Rhizobium meliloti* strains containing different *nodD* genes. *J. Bacteriol.* 172, 2769-2773.

72. Hartwig, U. A.; Maxwell, C. A.; Joseph, C. M.; Phillips, D. A. (1990) Chrysoeriol and luteolin released from alfalfa seeds induce *nod* genes in *Rhizobium meliloti*. *Plant Physiol.* 92, 116-122.

73. D'Haeze, W.; Holsters, M. (2002) Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiol.* 12, 79R-105R.

74. Subramanian, S.; Stacey, G.; Yu, O. (2006) Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. *Plant J.* 48, 261-273.

75. Wasson, A. P.; Pellerone, F. I.; Mathesius, U. (2006) Silencing the flavonoid pathway in *Medicago truncatula* inhibits root nodule formation and prevents auxin transport regulation by rhizobia. *Plant Cell* 18, 1617-1629.

76. Kessmann, H.; Edwards, R.; Geno, P. W.; Dixon, R. A. (1990) Stress responses in alfalfa (*Medicago sativa* L.) : V. Constitutive and elicitor-induced accumulation of isoflavonoid conjugates in cell suspension cultures. *Plant Physiol.* 94, 227-32.

77. Stitt, M.; Muller, C.; Matt, P.; Gibon, Y.; Carillo, P.; Morcuende, R.; Scheible, W. R.; Krapp, A. (2002) Steps towards an integrated view of nitrogen metabolism. *J. Exp. Bot.* 53, 959-70.

78. Sulieman, S. (2011) Does GABA increase the efficiency of symbiotic N_2 fixation in legumes? *Plant Signal Behav.* 6, 32-6.

79. Kahn, M. L.; Kraus, J.; Somerville, J. E., A model of nutrient exchange in the Rhizobium-legume symbiosis. In *Nitrogen Fixation Research Progress*, Evans, H. J.; P.J., B.; Newton, W. E., Eds. Martinus Nijhoff Dordrecht, The Netherlands., 1985; pp 193-9.

80. Sulieman, S.; Schulze, J. (2010) Phloem-derived γ-aminobutyric acid (GABA) is involved in upregulating nodule N_2 fixation efficiency in the model legume *Medicago truncatula*. *Plant Cell Environ.* 33, 2162-72.

81. Jones, D. L. (1998) Organic acids in the rhizosphere - a critical review. *Plant Soil* 205, 25-44.

82. Lee, S.; Valentine, S. J.; Reilly, J. P.; Clemmer, D. E. (2012) Analyzing a mixture of disaccharides by IMS-VUVPD-MS. *Int. J. Mass Spectrom.* 309, 161-167.

83. Oke, V.; Long, S. R. (1999) Bacterial genes induced within the nodule during the Rhizobium-legume symbiosis. *Mol. Microbiol.* 32, 837-849.

Table 1. Peak assignments in positive mode profiling mass spectra of Medicago root and nodules.

Name of metabolite	Theoretical [M+H]+	UHR Q-TOF measured [M+H]+	UHR Q-TOF measurement error; Δm (mDa)	TOF/TOF measured [M+H]+
γ-aminobutyric acid[a]	104.0706	104.0706	0	104.10
Choline[a]	104.1070	104.1071	0.1	104.10
Proline	116.0706	116.0706	0	116.07
Valine	118.0863	118.0865	0.2	118.09
Leucine[a]	132.1019	132.1022	0.3	132.10
Asparagine[a]	133.0608	133.0602	-0.6	133.06
Adenine[a]	136.0618	NA	NA	136.07
Proline betaine[a]	144.1019	144.1024	0.5	144.10
Glutamine[a]	147.0764	147.0768	0.4	147.09
Histidine[a]	156.0768	156.0771	0.3	156.06
Arginine[a]	175.1190	175.1167	-2.3	175.13
Sucrose+K	381.0794	381.0791	-0.3	381.05
Heme[a]	616.1768	NA	NA	616.15
NAD[a]	664.1164	NA	NA	664.10
Formononetin[a]	269.0808	269.0803	-0.5	269.08
Chrysoseriol GlcA[a]	477.1028	477.1008	-2.0	477.10
Formononetin MalGlc[a]	517.1341	517.1337	-0.4	517.13
Afrormosin MalGlc[a]	547.1446	547.1427	-1.9	547.16
GP(36:5)	849.4677	NA	NA	849.47

[a] Tandem MS fragmentation performed for identification.

Table 2. Peak assignments in negative mode profiling mass spectra of Medicago root and nodules.

Name of metabolite	Monoisotopic [M-H]-	UHR Q-TOF measured [M-H]-	UHR Q-TOF measurement error; Δm (mDa)	TOF/TOF measured [M-H]-
Pyruvic acid	87.0077	NA	NA	87.00
Alanine[a]	88.0393	88.0381	-1.2	88.01
Lactic acid[a]	89.0233	89.0244	1.1	89.01
Phosphoric acid	96.9685	96.9671	-1.4	96.96
2-ketobutyric acid	101.0233	101.0227	-0.6	101.02
γ-aminobutyric acid[a]	102.0550	102.0561	1.1	102.04
Serine	104.0342	104.0334	-0.8	104.02
Maleic/fumaric acid[a]	115.0026	115.0027	0.1	115.03
Succinic acid[a]	117.0182	117.0184	0.2	117.01
Threonine	118.0499	118.0491	-0.8	118.04
Oxalacetic acid	130.9975	130.9965	-1.0	131.03
Aspartic acid[a]	132.0291	132.0305	1.4	132.03
Malic acid[a]	133.0132	133.0131	-0.1	133.02
Salicyclic acid	137.0233	137.0231	-0.2	137.02
α-ketoglutaric acid	145.0132	145.0127	-0.5	145.02
Glutamic acid[a]	146.0448	146.0435	-1.3	146.01
Pentose	149.0445	149.0439	-0.6	149.04
Aconitic acid[a]	173.0081	173.0083	0.2	173.03
Ascorbic acid[a]	175.0237	175.0212	-2.5	175.05
Hexose	179.0550	179.0531	-1.9	179.05
Citric/isocitric acid[a]	191.0186	191.0186	0	191.02
Palmitic acid	255.2319	255.2318	-0.1	255.22
Hexose-6-phosphate[a]	259.0213	259.0219	0.6	259.04

Stearic acid	283.2632	283.2639	0.7	283.26
Sucrose[a]	341.1078	341.1079	0.1	341.07

[a] Tandem MS fragmentation performed for identification. The UHR Q-TOF masses were internally calibrated for mass accuracy.

Bold type indicates major metabolites in tricarboxylic acid cycle.

Figures

Figure 1. MALDI-MSI of metabolites in Medicago roots and nodules. Optical images of (a) a longitudinal section of a nodule stained with methylene blue and (b) a thaw-mounted Medicago section on a MALDI plate. (c) Workflow of mapping metabolites in Medicago by MALDI-MSI.

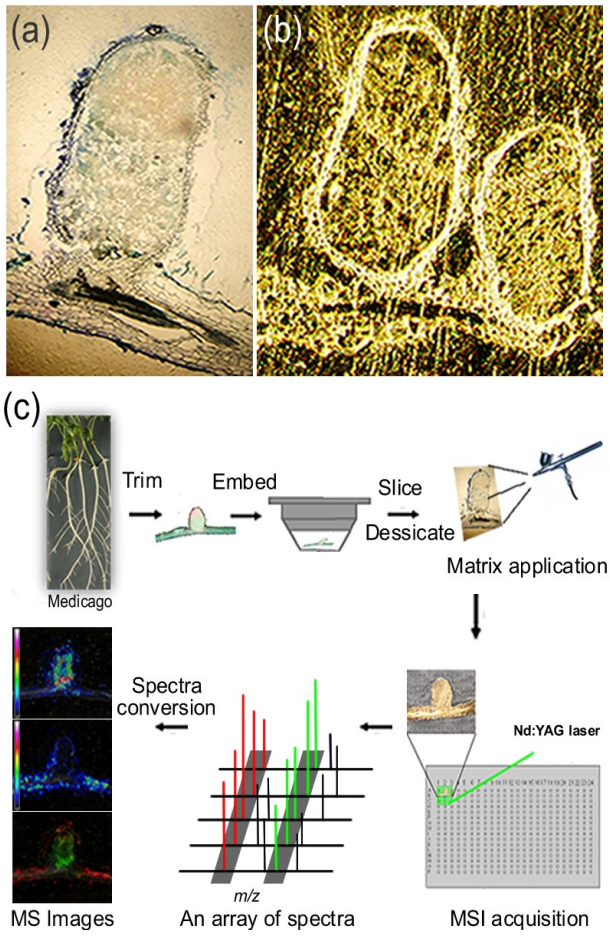


Figure 2. Profiling spectra of Medicago roots and nodules. (a) Overlaid profiling spectra from the nodule region (red), the root region (green) and the matrix DHB (blue). (b) and (c) are two enlarged spectra from (a). Metabolites detected on MALDI-TOF/TOF are annotated with their accurate masses. (d)-(f) MS/MS spectrum of selected precursors acquired in positive mode for molecular identification.

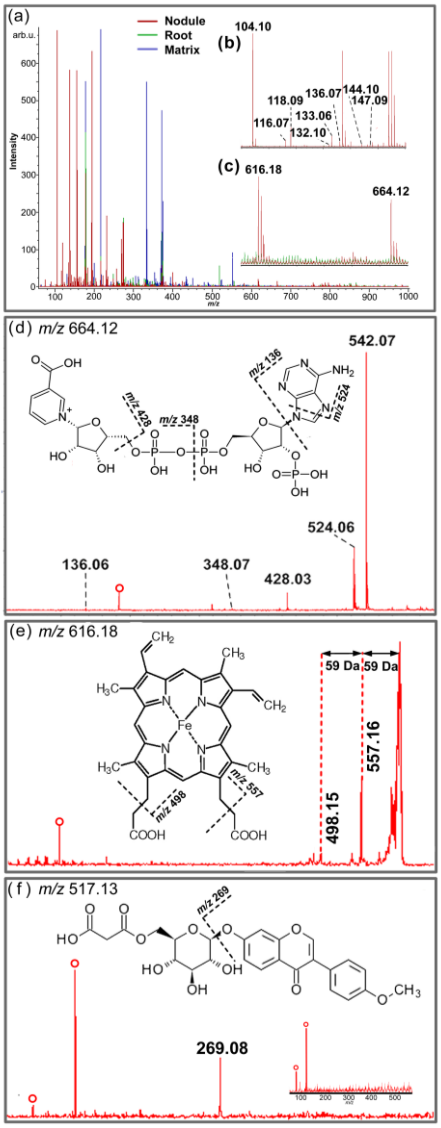


Figure 3. Statistical analysis of root and nodule profiling spectra. (a) ClinProTools-processed profiling spectra loaded as two classes arranged in a "gel" view with nodules displayed at the bottom and roots above. (b) PCA score plot of the nodules (red) and roots (green) classes. (c) Averaged signal intensities of the matrix-derived peaks between the nodule and root classes. (d) Averaged signal intensities of representative metabolites in Medicago. The peaks acquired in nodule regions are coded in red and those in root regions are coded in green. ** represents $p < 0.01$ and * represents $p < 0.05$ when comparing the nodule with the root specific spectra using Student's t-test.

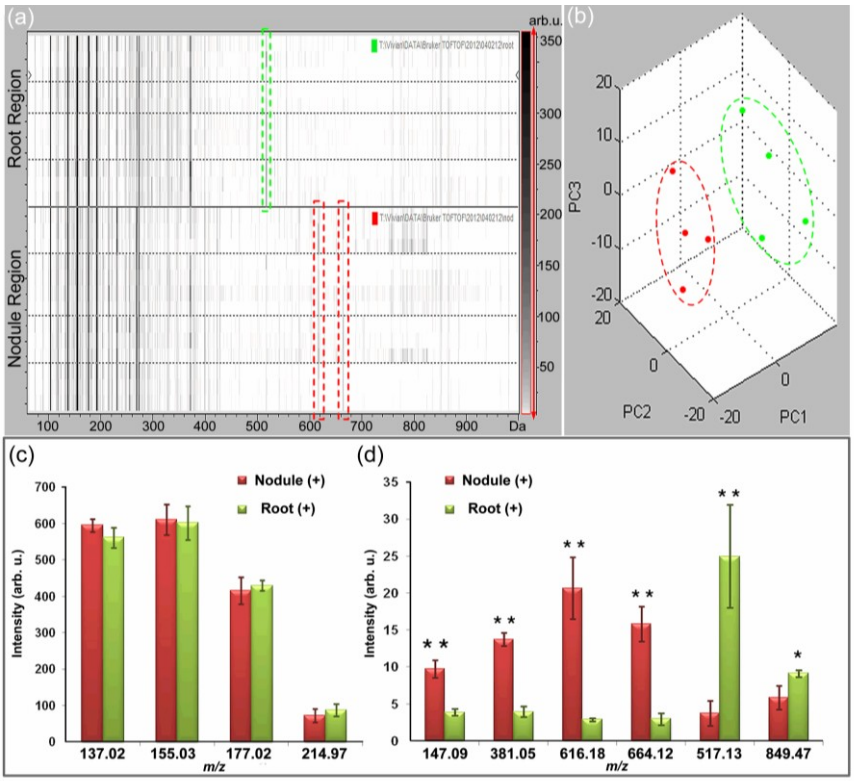


Figure 4. Representative metabolite distribution in Medicago section revealed by MALDI-MSI. (a) *m/z* 616.15 as heme-moiety, (b) *m/z* 664.10 as NAD, (c) *m/z* 144.10 as proline betaine, (d) *m/z* 849.47 as a putative sodiated lipid, (e) *m/z* 269.08 as formononetin, (f) *m/z* 517.13 as formononetin MalGlc, (g) *m/z* 547.16 afrormosin MalGlc displaying distinct distribution patterns in roots and nodules. (h) An overlaid image of (a) and (f). (i) PCA score plot of the two groups of MSI results arising from nodule and root by selecting the nodule as region of interest 1 (ROI 1) and root as ROI 2. The red points belonging to nodules were plotted apart from the green points representing roots, demonstrating significant differences between the two ROIs. (j) Loading plot of PC1 and PC2. Each point within the plot represents a peak detected in the MSI data set, whereas points that are far away from the center are responsible for the variance. The highlighted points corresponding to metabolites at *m/z* 547.16, 616.18 and 664.12 are furthest away from zero on PC2 and thereby indicate their significant contribution to the segregation of ROI 1 and ROI 2 on PC2.

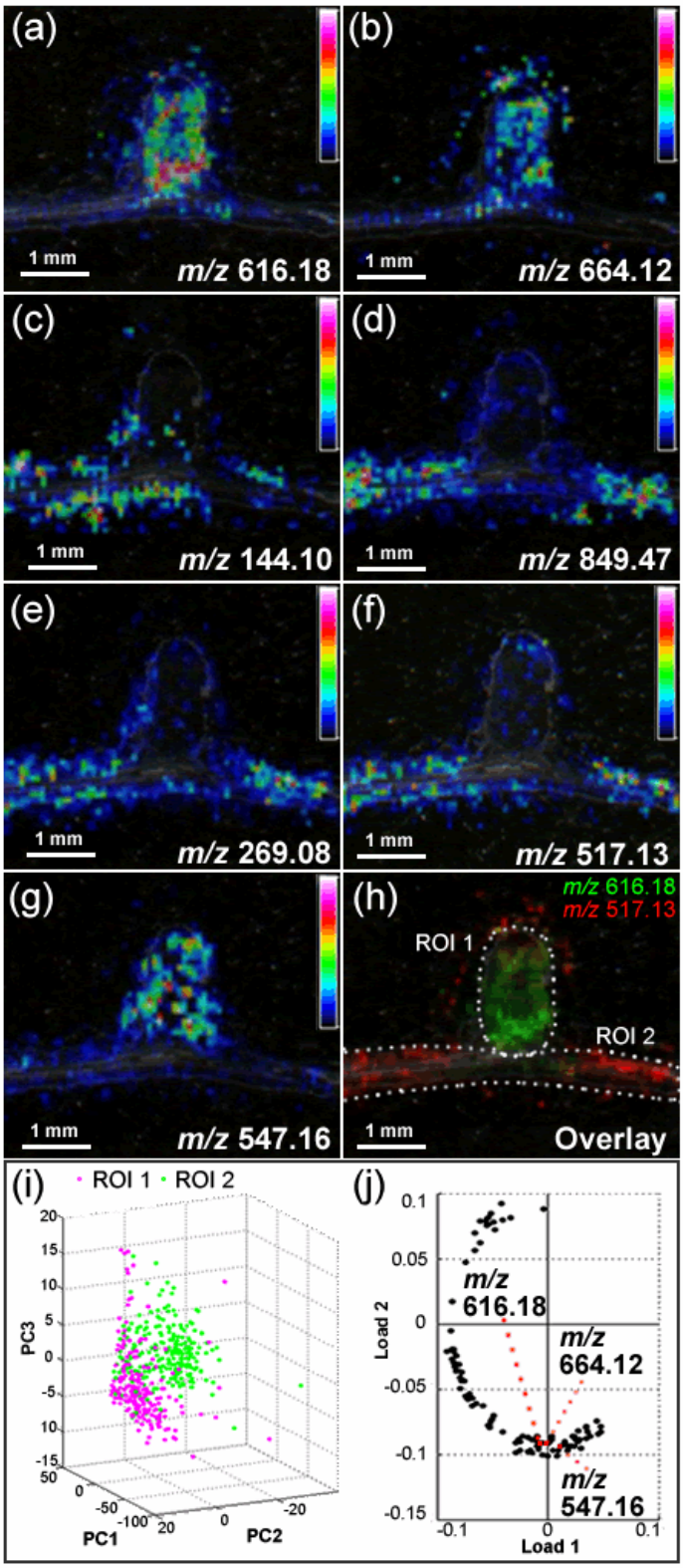

(a)
1 mm
m/z 616.18
(b)
1 mm
m/z 664.12
(c)
1 mm
m/z 144.10
(d)
1 mm
m/z 849.47
(e)
1 mm
m/z 269.08
(f)
1 mm
m/z 517.13
(g)
1 mm
m/z 547.16
(h)
m/z 616.18
m/z 517.13
ROI 1
ROI 2
1 mm
Overlay
(i)
ROI 1
ROI 2
PC3
PC1
PC2
(j)
Load 2
Load 1
m/z 616.18
m/z 664.12
m/z 547.16

Figure 5. Comparison of metabolite distribution in nitrogen fixing and non-fixing Medicago nodules. (a) Nodules on WT Medicago plant inoculated with WT strain of rhizobia (WT/WT). (b) Nodules on WT plant inoculated with *fixJ* mutant of rhizobia (WT/*fixJ*). (c) Nodules on Medicago *dnf1-1* mutant inoculated with WT strain of rhizobia (*dnf1*/WT). (d) Nodules on Medicago *dnf1-1* mutant inoculated with *fixJ* mutant of rhizobia (*dnf1/fixJ*). (e) Nitrogen fixation activity. Each bar corresponds to the average value of acetylene reduction for eight plants assayed three weeks after inoculation. Error bars indicate standard error (SE). Distribution of heme-moiety at *m/z* 616.18 in the nodules on four different nodule types, including (f) WT/WT, (g) WT/*fixJ,* (h) *dnf1*/WT and (i) *dnf1/fixJ*. Heme-moiety was detected only in fixing nodules in 5f. Distribution of metabolite formononetin MalGlc at *m/z* 517.13 on (j) WT/WT, (k) WT/*fixJ,* (l) *dnf1*/WT and (m) *dnf1/fixJ*, displaying great similarity in contrast to (f)-(i). Scale bar = 1 mm.

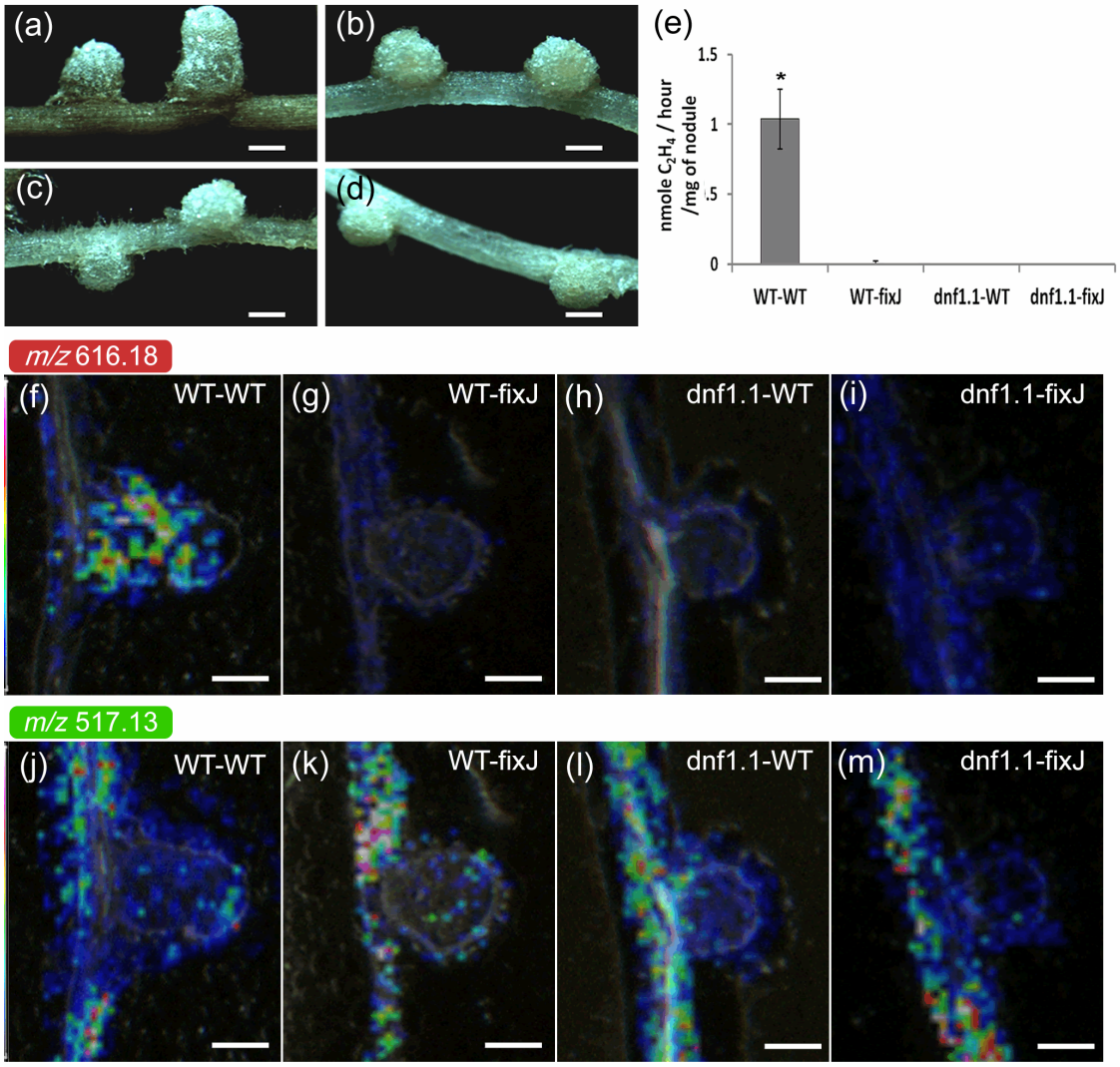

(a)
(b)
(c)
(d)
(e)
nmole C_2H_4 / hour /mg of nodule
1.5
1
.5
0
*
WT-WT
WT-fixJ
dnf1.1-WT
dnf1.1-fixJ
m/z 616.18
(f) WT-WT
(g) WT-fixJ
(h) dnf1.1-WT
(i) dnf1.1-fixJ
m/z 517.13
(j) WT-WT
(k) WT-fixJ
(l) dnf1.1-WT
(m) dnf1.1-fixJ

Figure 6. Comparison of *in situ* profiled mass spectra from Medicago roots, nodules and DMAN matrix in negative mode. The insert shows an enlarged view of the nodule-specific spectrum. The identified peaks were annotated with masses and assigned identities. The star **(*)**indicates the major peaks arising from DMAN.

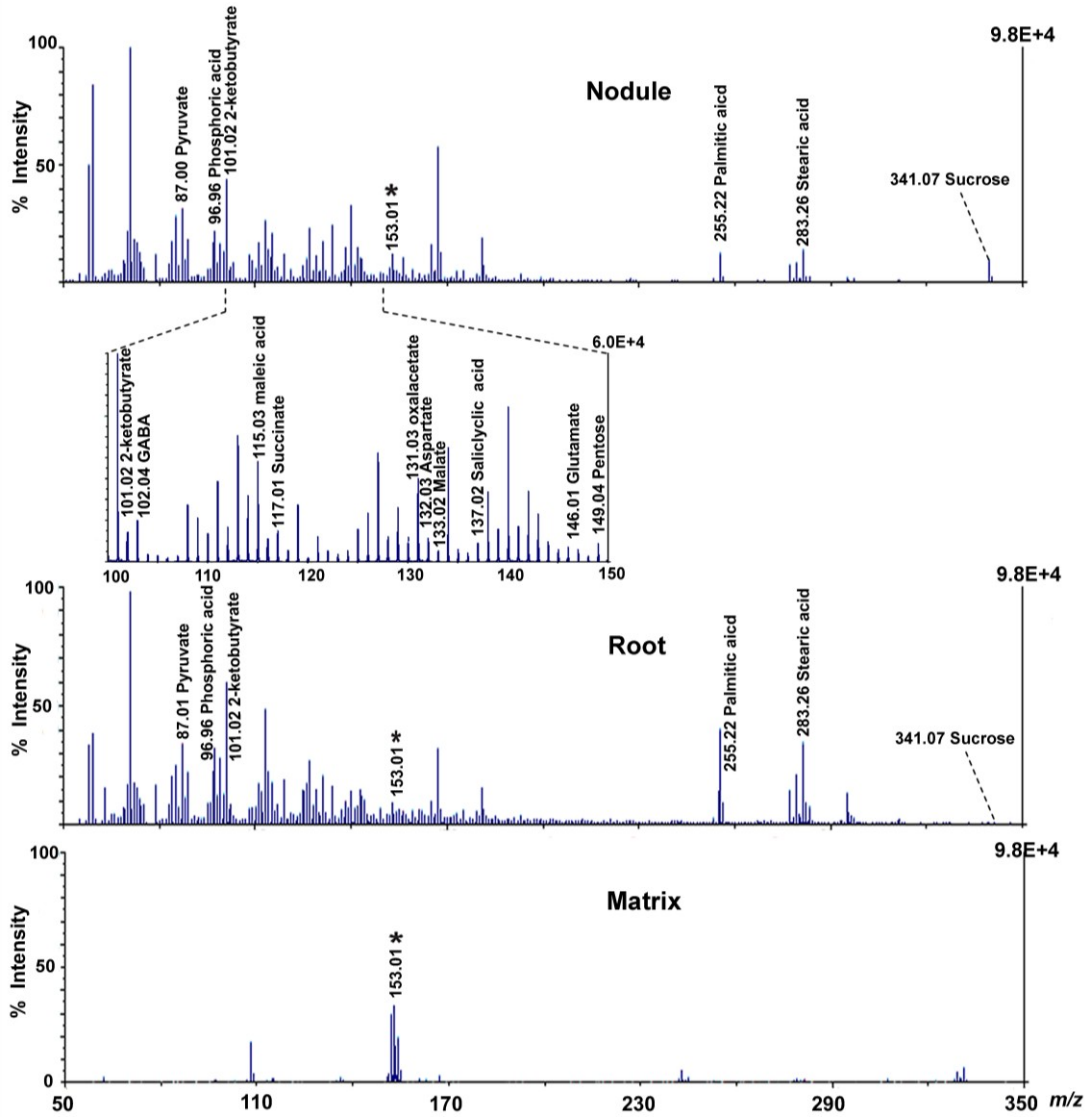


Figure 7. *In situ* MS/MS spectra of metabolites. *In situ* tandem MS spectra of (a) maleic acid, (b) citric acid and (c) sucrose conducted on Medicago sections in negative mode. MS/MS spectra of (d) maleic acid, (e) citric acid and (f) sucrose standards. The possible fragmentation pathways were proposed in inserts.

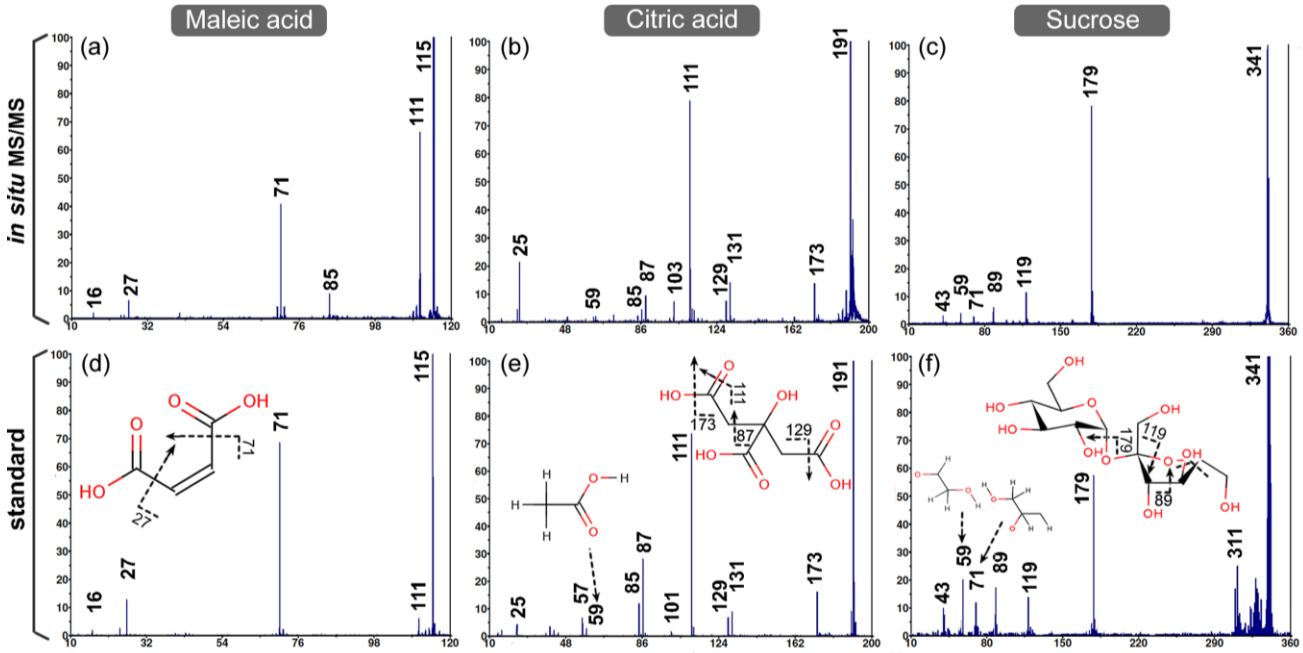
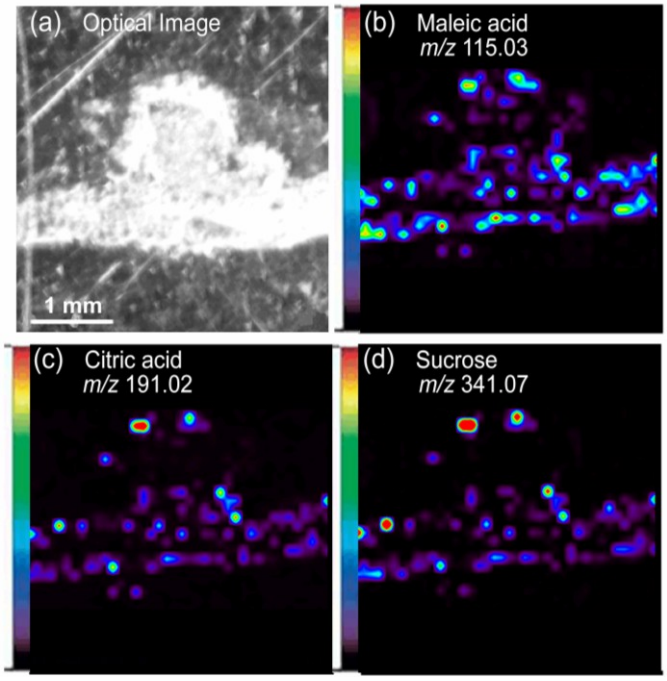


Figure 8. Spatial distribution of metabolites in Medicago nodules. (a) An optical image of Medicago section and its representative negative mode MS images of deprotonated metabolite peaks, including (b) maleic acid (*m/z* 115.03), (c) citric acid (*m/z* 191.02) and (d) sucrose (*m/z* 341.07). Scale bar = 1 mm.



Supplemental Information

Method S1: Plant lines, rhizobial strains and chemicals

Medicago truncatula (Medicago) Jemalong A17 (WT plant) and the mutant *dnf1-1* [52] were used in this study. *Sinorhizobium meliloti* strain Rm1021 and *S. meliloti* mutant strain VO2683 (*fixJ2.3*::Tn5-233) [83] were used for inoculation. DMAN, DHB, citric acid, maleic acid and aspartic acid were all purchased from Sigma-Aldrich (St. Louis, MO).

Method S2: Acetylene Reduction Assays (ARA)

Three weeks after inoculation with *S. meliloti* WT or *fixJ* mutant, the nodulated roots of Medicago (WT and *dnf1-1*) were excised and introduced into glass vials (14.5 ml) containing 500 µl of liquid Fahraeus medium and tightly sealed with rubber stopper and aluminum flip-off caps. 1 ml of acetylene (Airgas) was injected in each vial and was incubated in the dark. After 48 hours, 1 ml of the gas phase was injected on a Gas Chromatography (GC-2010 Shimadzu) equipped with an Rt-Alumina BOND/KCL column (Restek). The number and fresh weight of nodules were measured for each plant immediately after ARA measurement. The mean numbers for each treatment were compared by the Kruskal-Wallis test and, when significant ($p<0.05$), boxplot analysis. Statistical analyses were performed with the R Software.

Method S3: Tissue Extraction

92% ice-cold methanol acidified with 0.1% (v/v) formic acid was added to the plant powder at a 4:1 (w/w) ratio. The ground suspension was incubated with shaking for 20 minutes at 4^{o}C and then without shaking for 15 minutes at -20^{o}C. The sample was then centrifuged at 20,000 xg for 10 minutes. The resulting supernatant was collected and lyophilized overnight. The sample was stored at -80^{o}C for future use.

Method S4: Sample preparation for imaging

The nodules were excised along with flanking root tissues, embedded in 100 mg/ml gelatin aqueous solution and flash-frozen in liquid nitrogen. The frozen tissue was then sectioned into 12 μm slices at -20^{o}C. Sections were thaw-mounted onto an opti-TOF 384-well plate for negative mode imaging or an indium tin oxide (ITO)-coated glass slide for positive mode imaging. Tissue sections were dehydrated in a desiccator at -20^{o}C for at least 1 hour prior to matrix application. For positive mode imaging, 150 mg/ml DHB (in 0.1% formic acid and 50% methanol) was used. Ten coatings were applied; the spray duration was 30s with a 1 min interval between coatings. For negative mode imaging, 15 mg/ml DMAN (in 15% water and 85% methanol) was applied. Fifteen coatings were applied; the spray duration was 20s with a 1 min interval between coatings. These sections were stained with 0.01% aqueous solution of

methylene blue and observed under a bright field microscope to verify the tissue integrity.

Method S5: MALDI-TOF/TOF for MSI

At least four Medicago roots with nodules were used and three sections from each nodule were examined for reproducibility. The detailed MSI setup on 4800 MALDI-TOF/TOF is as follows: the tissue region to be imaged and the raster step size were controlled using the 4800 Imaging Application (Novartis, Basel, Switzerland). To generate images, spectra were collected at 50 μm intervals in both the x and y dimensions across the surface of the sample. Each mass spectrum was generated by averaging 250 laser shots over the *m/z* 50-800 mass range. Mass spectra were externally calibrated using standard mixtures applied directly to the MALDI target. Tandem MS fragmentation was accomplished by turning on 1 kV collisional-induced-dissociation (CID) and the optimized metastable suppressor. Air was used as the collision gas at a medium pressure. Ion peaks with signal to noise ratio over 10 were considered to be potential metabolite ion peaks, except for peak clusters that derived from the matrix or matrix adducts.

The following parameters were adopted in the positive reflectron mode for imaging acquisition on Bruker autoflex III: ion source 1 voltage 19.00 kV, ion source 2 voltage 16.62 kV, reflector 1 voltage 20.90 kV, reflector 2 voltage 9.64 kV and lens

voltage 8.70 kV. The array of spectra was collected at 50 μm intervals in both the x and y dimensions across the surface of the longitudinal section, and each spectrum consisted of 250 laser shots over the mass range *m/z* 80-1000. Tandem mass spectra were obtained in the LIFT mode.

Method S6: ESI-Q-TOF

In order to acquire the accurate masses of metabolites, the lyophilized Medicago root nodule extract was reconstituted in 100 μL of ACN:H_2O (50:50) with 0.1% formic acid for direct infusion in positive mode and 100 μL of ACN:H_2O (50:50) with 2% ammonium bicarbonate for direct infusion in negative mode on maXis.

MS direct infusion data were acquired on a maXis ultrahigh resolution (UHR) quadrupole-time-of-flight (Q-TOF) mass spectrometer (Bruker Daltonics, Bremen, Germany) in positive and negative mode, respectively. Prior to data acquisition, external calibration in positive mode was performed using a tune mix. Following spectral calibration, an aliquot of 100 μL extract was injected via a syringe pump at a flow rate of 0.3 mL/min.

To acquire LC-ESI-Q-TOF MS/MS data, the Medicago root and nodule extract underwent separation on a C18 reversed phase capillary column (75 μm internal diameter × 150 mm length, 3 μm particle size; Micro-Tech Scientific Inc., Vista, CA). The

mobile phases were: deionized H_2O with 0.1% formic acid (A); acetonitrile with 0.1% formic acid (B). The re-constituted extract was injected and loaded onto the trap column (Zorbax 300SB-C18 Nano trapping column, Agilent Technologies, Santa Clara, CA) using 95% mobile phase A and 5% mobile phase B at a flow rate of 10 μL/min for 10 min. Following this, the stream select module was switched to align the trap column with the analytical column, and a linear gradient of mobile phases A and B was carried out, progressing from 3% to 40% over 60 min. A data dependent acquisition was employed for the MS survey scan and the selection of three precursor ions and subsequent MS/MS of the selected parent ions. The MS scan range was from *m/z* 50–1000 and the MS/MS scan was from *m/z* 50–1000. The MS/MS sequences were interpreted manually.

Method S7: Molecular identification of metabolites

Molecular identification of the metabolites involves several steps by first searching the masses acquired on MALDI-TOF/TOF from Medicago root and nodule sections against database with a mass error within 3 mDa. The masses obtained from MALDI-TOF/TOF were then compared with those of Medicago root and nodule extract recorded on maXis, which offers excellent mass measurement accuracy with a low mass error within 5 ppm. The metabolite list generated by mass-matching was further

confirmed by comparing signature tandem MS fragment peaks obtained on MALDI-TOF/TOF or LC-ESI-Q-TOF with those from purchased standards, database or previous literature. As for low-abundance ions for which tandem MS analyses failed, identifications were conducted solely relying on accurate masses. METLIN, KNApSAcK and the Human Metabolome Database were used for metabolite mass matching and MS/MS fragment assignment.

Figure S1. Tandem mass spectra of metabolites identified from Medicago root and nodule extract via a conventional LC-MS/MS approach. (a) MS/MS fragmentation spectrum of the *m/z* 517.1337 ion assigned as formononetin MalGlc. (b) MS/MS fragmentation spectrum of the *m/z* 269.0803 ion assigned as formononetin. (c) MS/MS fragmentation spectrum of the *m/z* 547.1427 ion assigned as afrormosin MalGlc. The accurate mass, MS/MS spectra and retention time altogether confirm the molecular assignment based on MALDI-MS.

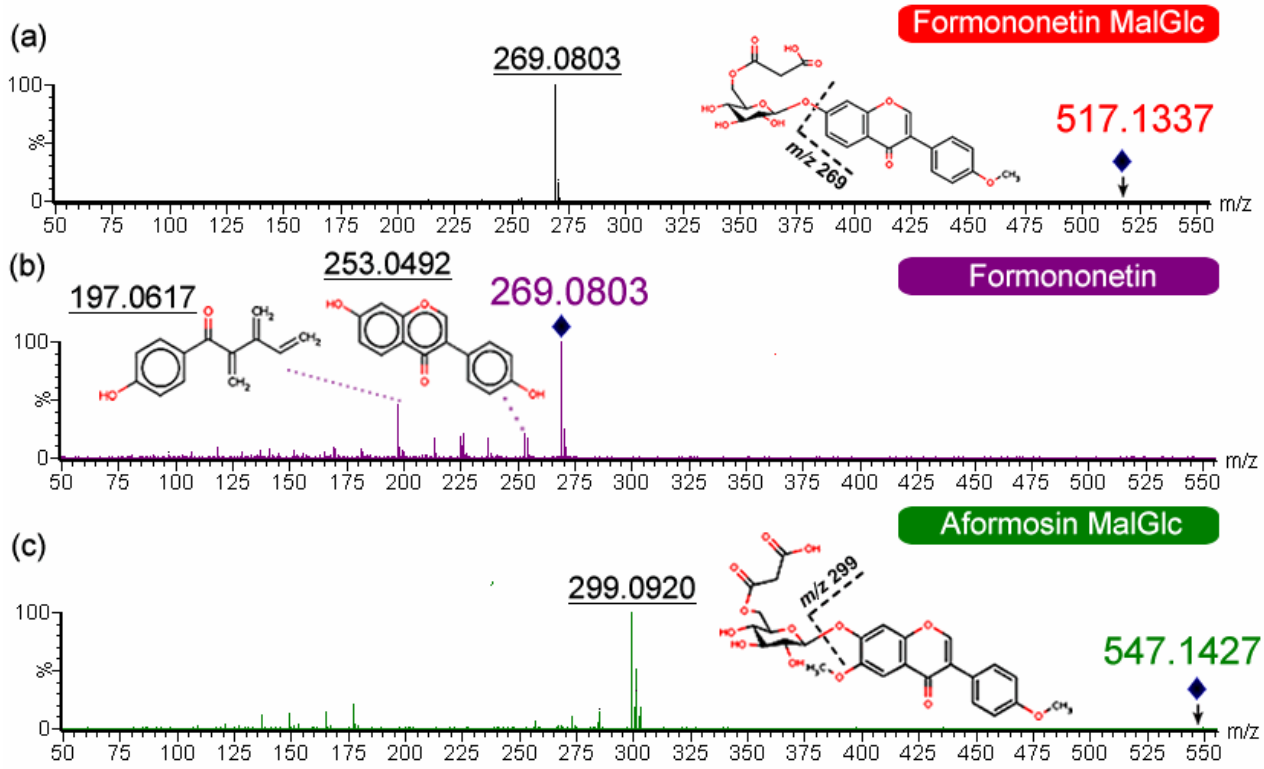


Figure S2. Dendrogram representation of the PCA-transformed spectra. A dendrogram was generated for all the PCA-transformed spectra using an unsupervised clustering function enabled by ClinProTools. Two branches are shown with the green dots corresponding to spectra with root and the red dots corresponding to nodule origin, providing statistical support to the differentiation between roots and nodule metabolite spectra.

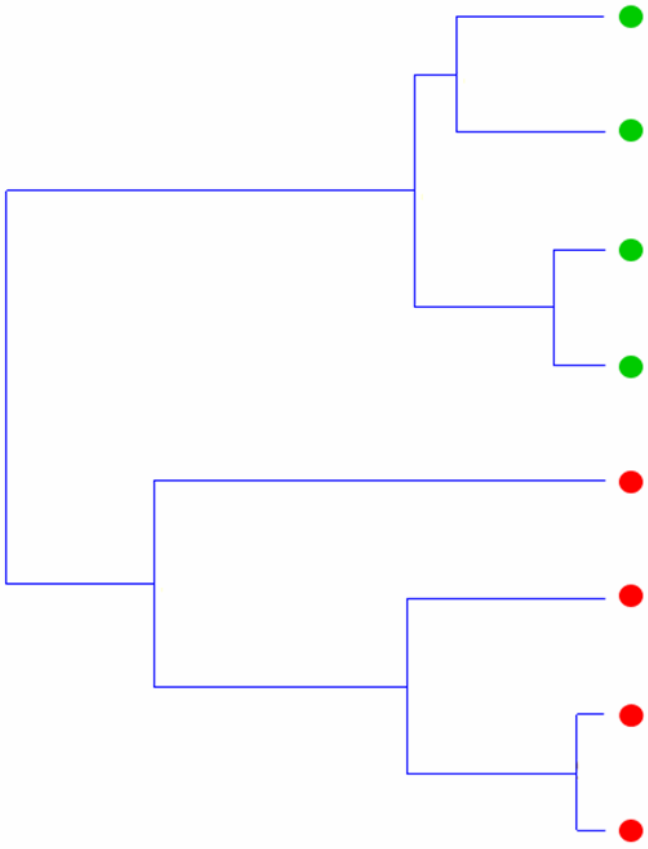


Figure S3. Comparison of Medicago nodules in WT and mutant backgrounds. (a) Nodulation rates for each different association (n=8). (b) Fresh weight of nodule tissues (n=8). Error bars indicate SE. a and b indicate statistically supported groups according to Kruskal-Wallis test ($p<0.05$) and boxplot analysis.

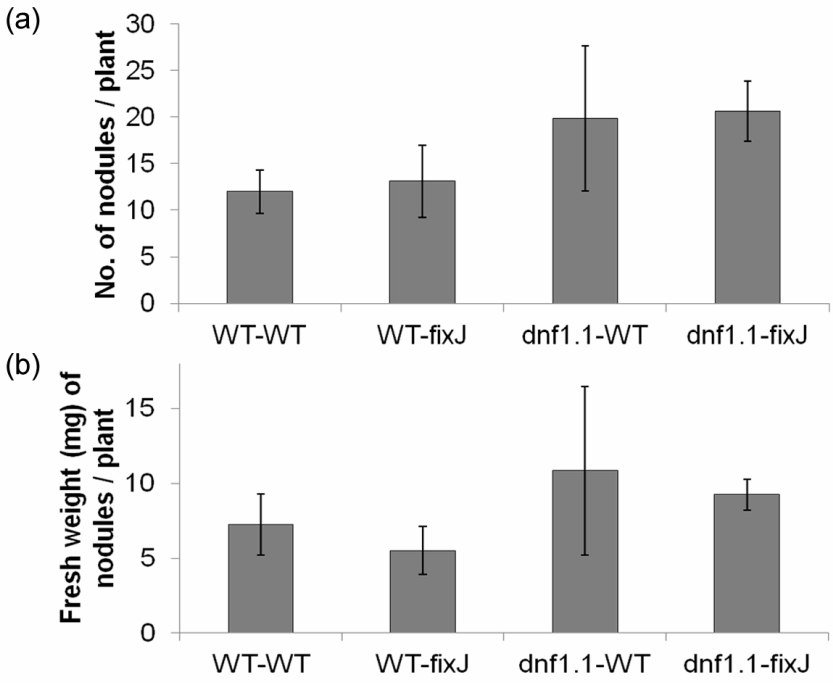


Figure S4. Comparison of profiling spectra with DMAN or DHB as matrix. Spectra of maleic acid using (a) DMAN and (d) DHB, citric acid using (b) DMAN and (e) DHB and aspartic acid using (c) DMAN and (f) DHB. Detected metabolite standards are highlighted in the spectra. The concentration of standard was 0.1 mg/ml, and both matrices were made in 10 mg/ml.

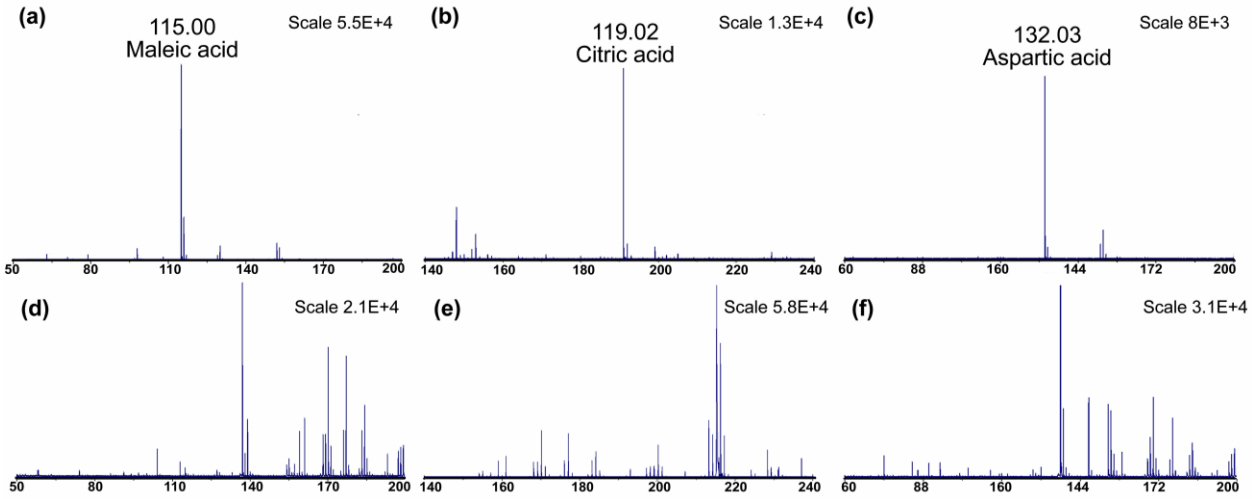


Figure S5. Schematic representation of the citric acid cycle and detected metabolites. The metabolites with black-coded *m/z* were designated by the measured *m/z* values obtained from both MALDI-TOF/TOF, maXis UHR-Q-TOF and partially by tandem mass fragmentation patterns from MALDI-TOF/TOF. The metabolite with red-coded *m/z*, pyruvate, was only detected by MALDI-TOF/TOF.

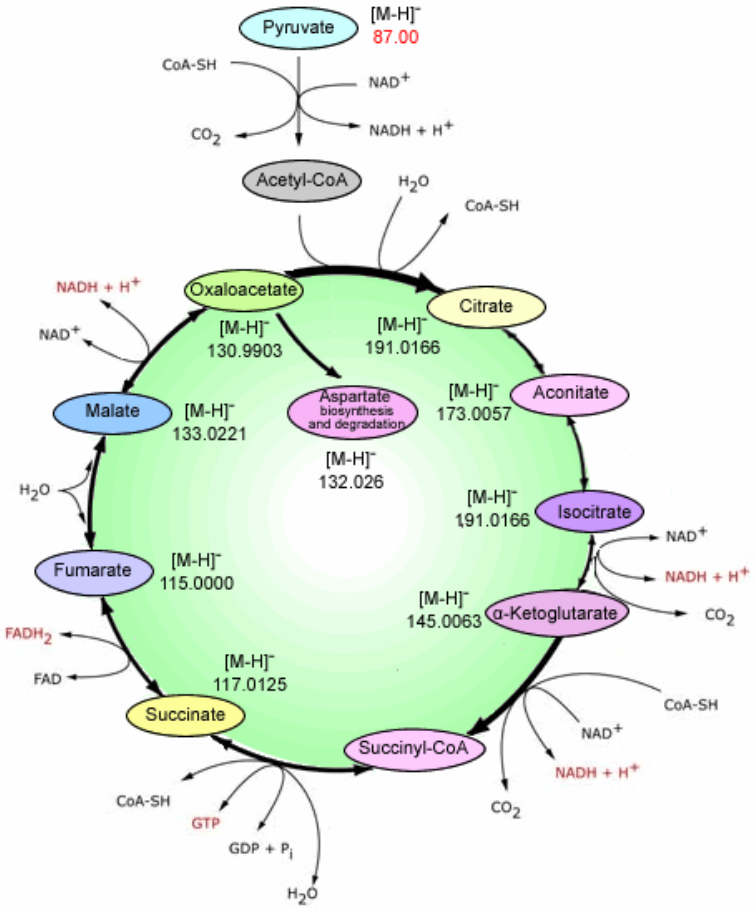
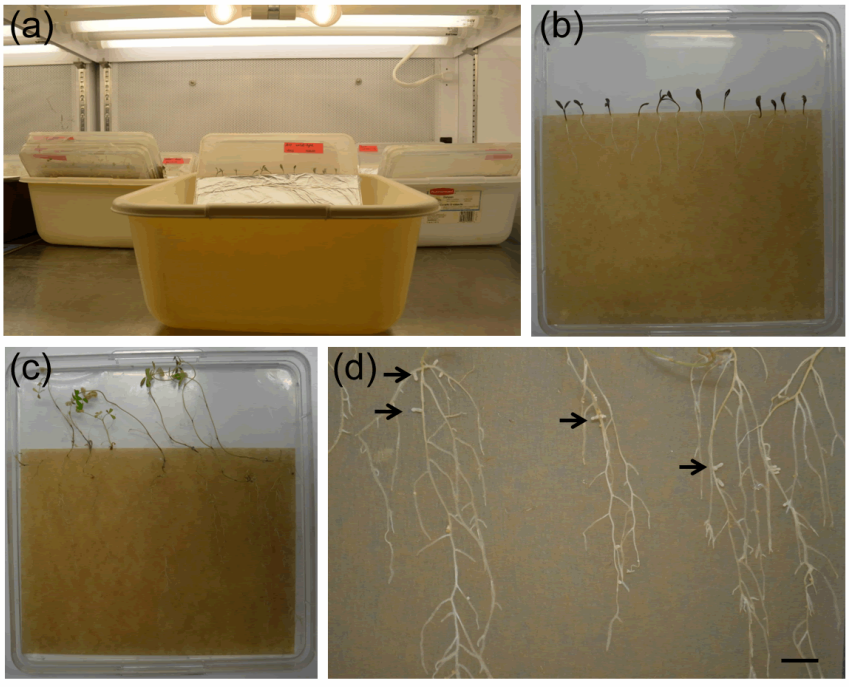


Figure S6. Plant growth conditions and experimental setup. (a) Medicago plants grown on Fahreaus medium plates in growth chamber. (b) Young 5-day old seedlings growing on Fahreaus medium plate. (c) Medicago plants three weeks after inoculation with rhizobia in Fahreaus medium plate. (d) Medicago roots with nodules three weeks after inoculation with rhizobia. Black arrows indicate mature nodules. Scale bar = 1 cm.



Chapter 5

Mapping of Neuropeptides in the Crustacean Stomatogastric Nervous System by Imaging Mass Spectrometry

Adapted from **Hui Ye**, Limei Hui, Katherine Kellersberger, Lingjun Li. Mapping of neuropeptides in the crustacean stomatogastric nervous system by imaging mass spectrometry. *J. Am. Soc. Mass Spectrom*. 2013, 24:134-47.

Abstract

Considerable effort has been devoted to characterizing the crustacean stomatogastric nervous system (STNS) with great emphasis on comprehensive analysis and mapping distribution of its diverse neuropeptide complement. Previously, immunohistochemistry (IHC) has been applied to this endeavor yet with identification accuracy and throughput compromised. Therefore, molecular imaging methods are pursued to unequivocally determine the identity and location of the neuropeptides at a high spatial resolution. In this work, we developed a novel multi-faceted mass spectrometric strategy combining profiling and imaging techniques to characterize and map neuropeptides from the blue crab *Callinectes sapidus* STNS at the network level. In total, 55 neuropeptides from 10 families were identified from the major ganglia in the *C. sapidus* STNS for the first time, including the stomatogastric ganglion (STG), the paired commissural ganglia (CoG), the esophageal ganglion (OG), and the connecting nerve stomatogastric nerve (*stn*) using matrix-assisted laser desorption/ionization tandem time-of-flight (MALDI-TOF/TOF) and the MS/MS capability of this technique. In addition, the locations of multiple neuropeptides were documented at a spatial resolution of 25 μm in the STG and upstream nerve using MALDI-TOF/TOF and high-mass-resolution and high-mass-accuracy MALDI-Fourier transform ion cyclotron resonance (FT-ICR) instrument. Furthermore, distributions of neuropeptides in the whole *C. sapidus* STNS were examined by imaging mass spectrometry (IMS). Different isoforms from the same family were simultaneously and unambiguously mapped, facilitating the functional exploration of neuropeptides

present in the crustacean STNS and exemplifying the revolutionary role of this novel platform in neuronal network studies.

5.1 Introduction

Neuropeptides are the largest and most diverse group of endocrine signaling molecules in the nervous system. The crustacean stomatogastric nervous system (STNS) is a rich source of numerous neuropeptides and neurotransmitters. [1-5] It contains four ganglia, including a paired CoG and the unpaired OG and the STG. [6] The neural circuits in the STNS provide an ideal model to study how a rich repertoire of neuropeptides modulate the rhythmic motor patterns that can control ingestive and masticatory movements of the foregut [6-10] to manage the different aspects of feeding including swallowing, food storage, chewing and the filtering of chewed food. [11] The STG, as the center of the STNS, is a ganglion of minute size and contains merely 25-30 motor neurons.

Many neuropeptides as well as neuromodulators related with rhythmic motor pattern generation in the STNS have been investigated using electrophysiological studies. For example, three families of allatostatins (AST) can reduce pyloric network burst frequency in a state-dependent manner, while application of proctolin on the STG shows robust excitatory actions. [12] However, the functions of many other neuropeptides, such as RYamide and YRamide, have not been characterized. The study of the complement and distribution of neuropeptides in this important nervous system is essential to discover the underlying mechanism of their functions. [13, 14] Previously, the CoGs and the STG of

several crustacean species have been characterized by mass spectrometry (MS) via tissue homogenization [15, 16] or direct tissue analysis. [17-19] Nevertheless, spatial information of the detected neuropeptides were missing in those approaches, [20] whereas a complete understanding of the functional roles of these signaling peptides in the neural circuit requires detailed characterization of all the molecular players involved in the system. A powerful technique to achieve both goals, the identification and discovery of neuropeptides and mapping the distribution of neuropeptides in the crustacean nervous system, is therefore highly demanded.

The ability to visualize neuropeptide distributions in organs or cells is usually achieved by combining immunochemistry with high-resolution microscopy. [21] However, this technique provides fairly limited chemical information and often suffers from cross-reactivity. These limitations can be overcome by imaging mass spectrometry (IMS), which integrates the high sensitivity and chemical specificity of a mass analyzer with the imaging capability of different ion sources. [22] For example, ambient ion source such as infrared laser ablation with electrospray ionization (LAESI), [23] atmospheric pressure infrared MALDI (AP IR-MALDI), [24] desorption electrospray ionization (DESI), [25] infrared laser ablation metastable-induced chemical ionization (IR-LAMICI) [26] have demonstrated their imaging capabilities and thereby present great potential in *in vivo* imaging of biomolecules. Unfortunately, their IMS applications are mostly limited by relatively low lateral-imaging resolution (typically 200-300 μm), limited sensitivity in peptide mass range or requirement of special instrumental modifications. [22] Alternatively, matrix-assisted laser desorption/ionization (MALDI) equipped with state-of-the-art mass

analyzers has been extensively employed to map a wide array of biomolecules ranging from proteins, [27, 28] neuropeptides [29, 30] to drugs and metabolites [31] with excellent throughput, robustness and sensitivity. [32] A major drawback of MALDI-IMS is its spatial resolution which is usually limited by the laser spot diameter (30~100 μm) on commercial instruments. [33] Nevertheless, advancements in sample-handling and mass spectrometers such as specially-designed confocal type objectives, [34] defocused laser beam employed in "microscope" mode MALDI-MS coupled with a position-sensitive detector [35] and an imaging acquisition method "oversampling" have allowed imaging at cellular-length scales in several studies. [36] Despite the limitation of spatial resolution, MALDI-IMS has exhibited its power in neuropeptide mapping from complex neural tissues ranging from mammalian tissues such as rat spinal cord, [30] brain, [29, 37] pituitary gland [38] to invertebrates like the crab *Cancer borealis* [39, 40] and insects like the house cricket *Acheta domesticus.* [41] These studies exemplified the advantages of IMS in neuropeptide mapping applications: it does not require prior knowledge about the specific signaling molecules; it can distinguish between multiple isoforms in the same peptide family, and enables simultaneous mapping of numerous neuropeptides in a single experiment. [33] Mass analyzers are critical for unequivocal identification of neuropeptides from complex samples. MALDI-TOF/TOF has become a major instrument employed in MALDI-IMS applications, offering high sensitivity, wide mass range and fast analysis speed. [33] In the current study, these advantages and the *in situ* MS/MS capability were demonstrated. Nevertheless, high-performance instruments like FT-ICR [18, 37, 38] or orbitrap [41, 42] outshine by allowing for imaging multiple peptides of close masses

unambiguously from tissue. The high measurement accuracy and resolving power of the FT-ICR employed in this study significantly enhanced our confidence in assigning neuropeptides based on mass-matching and resolved MS images of the neuropeptides with close masses present in tissue. Subsequent tandem MS measurements by ESI-Q-TOF further confirmed previous neuropeptide identifications obtained from the MALDI-MS platform. This novel multi-faceted mass spectrometry (MS)-based strategy combines high sensitivity, high resolving power, high mass accuracy and fast analysis speed, enabling a large-scale characterization and mapping of endogenous neuropeptides at a spatial resolution of 25 μm from a small yet complex neuronal network such as the *C. sapidus* STG exemplified in this report.

For the first time, a combination of MALDI-TOF/TOF and MALDI-FT-ICR was employed to study the identity and localization of neuropeptides in the STG, an organ of minute size, with high spatial resolution and high spectral quality. The colocalization of multiple peptide isoforms from several peptide families was demonstrated in this ganglion. IMS was also performed to map the whole STNS. In addition to imaging, multiple sample preparation techniques including direct tissue analysis, tissue extract analysis and online-UPLC analysis followed by tandem MS were implemented in this multi-faceted MS-based strategy. A total of 55 neuropeptides were unequivocally identified from the *C. sapidus* STNS in a high throughput fashion and verified by tandem MS sequencing.

5.2 Experimental Section

Chemicals and materials

Methanol, acetonitrile and formic acid (FA) were purchased from Fisher Scientific (Pittsburgh, PA). 2, 5-dihydroxybenzoic acid (DHB) was obtained from MP Biomedicals, Inc. (Solon, OH). Methylene blue staining solution was purchased from Sigma Aldrich (St. Louis, MO). Acidified methanol was prepared using 90% methanol, 9% glacial acetic acid, and 1% water. Composition of physiological saline is 440 mM NaCl; 11 mM KCl; 13 mM $CaCl_2$; 26 mM $MgCl_2$; 10 mM HEPES acid; pH 7.4-7.5. All water used in this study was Milli-Q water from a Millipore filtration system (Bedford, MA).

Animal dissection

Blue crabs C. *sapidus* were purchased from the local market and maintained in recirculating artificial seawater aquaria at ambient temperature (approximately 12-13 °C) before use. Crabs were anesthetized by packing them on ice for 15 min, after which the dorsal carapace were removed from each individual and its STNS were dissected free from surrounding muscle and connective tissues in chilled (approximately 10°C) physiological saline. The details of dissection were described elsewhere. [43] Following dissection, the STNS were immediately prepared for direct tissue analysis and IMS or saved in acidified methanol at -80°C for future use in peptide extraction.

Direct tissue analysis

For direct tissue analysis of the STNS, the STG, the *stn*, the CoG and the OG were removed from the STNS by microdissection scissors (Fine Science Tools, Inc., Foster city, CA, USA), briefly rinsed in acidified methanol and desalted in 10 mg/ml aqueous 2,5-dihydroxybenzoic acid (DHB) solution as described previously. [17] The tissue (N=3)

was then placed on the MALDI plate, and 0.3 µl of 150 mg/ml DHB solution (50:50 methanol: water, vol:vol) was applied on top of the tissue. Direct tissue mass spectrometric analysis was performed using a 4800 MALDI-TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA, USA).

MALDI-IMS sample preparation

The STNS was rinsed briefly in deionized water immediately following dissection to reduce the salt content, after which the excessive water was removed by KimWipes™ and the STNS was placed on a MALDI plate. For high-resolution imaging experiment, the STG was removed from the STNS by microdissection scissors (Fine Science Tools, Inc., Foster city, CA, USA) and placed on an indium tin oxide (ITO)-coated glass slide. Optical images of the STNS or STG were taken by scanning the glass slide with an office scanner (Epson V300 PHOTO) in professional mode with 2400 dpi resolution. The STNS or STG was then dehydrated in a desiccator in -20°C for 30 min prior to matrix application. For matrix application, 150 mg/ml DHB (50:50 methanol: water, vol:vol) was used and the airbrush (Paasche airbrush company, Chicago, IL, USA) was held 35 cm from the plate for regular spray. Five coats were applied, and the spray duration for each coat was 30 s with 1 min dry time between each cycle. Alternatively for IMS of the STG (N=3), the airbrush was held 45 cm from the plate for dryer spray. This modification allowed more matrix solvent to evaporate before reaching the plate and the formation of smaller matrix crystals on tissue surface for high resolution IMS. IMS of the STNS (N=3) was performed using a 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA), whereas IMS of the STG was conducted using

autoflex III MALDI-TOF/TOF (Bruker Daltonics, Billerica, MA, USA) and solariX^{TM} MALDI FT-ICR (Bruker Daltonics, Billerica, MA, USA).

MALDI-TOF/TOF MS

A model 4800 MALDI TOF/TOF analyzer equipped with a 200 Hz, 355 nm Nd:YAG laser (spot diameter of 75 µm) was employed for direct tissue analysis of the STG, *stn*, CoG and OG, and IMS of the STNS. Instrument parameters were set using the 4000 Series Explorer Software (Applied Biosystems, Framingham, MA). All mass spectra were collected in the positive ion reflectron mode. Mass calibration was performed externally using a mixture of synthetic peptide standards. Each representative profiling mass spectrum of the STG, *stn*, CoG and OG was averaged from 500 laser shots, and the mass spectra data was extracted over a mass range of *m/z* 600-1800 where majority of the neuropeptides were detected. Tandem mass spectra (MS/MS) were achieved *in situ* by 2 kV collision induced dissociation (CID) using air as the collision gas. 750 laser pulses were averaged for each MS/MS spectrum, and sequence interpretation was performed manually.

IMS acquisition of the STNS was controlled using the 4800 Imaging application software (Novartis, Basel, Switzerland) available through the MALDI IMS website (www.maldi-msi.org). Arrays of spectra were collected at 100 µm intervals in both x and y dimensions, and each spectrum was averaged from 200 laser shots. The mass spectra data was extracted over a mass range of *m/z* 600-1800. The image files were processed and MS images of neuropeptides were generated using the TissueView software package (Applied Biosystems, Framingham, MA, USA).

IMS of the STG was performed at high lateral resolution on autoflex III MALDI-TOF/TOF mass spectrometer equipped with a 200 Hz smartbeam^{TM} (diameter as small as 25 μm). The following parameters were adopted in the positive reflectron mode for imaging acquisition: ion source 1 voltage 19.00 kV, ion source 2 voltage 16.62 kV, reflector 1 voltage 20.90 kV, reflector 2 voltage 9.64 kV and lens voltage 8.70 kV. The mass spectra data was acquired over a mass range of *m/z* 600-1800 where majority of the neuropeptides were detected. Each spectrum consists of 500 laser shots and the array of spectra was collected at 25 μm intervals in both x and y dimensions across the surface of the STG. The image file was processed and MS images of neuropeptides were generated using the flexImaging software package (Bruker Daltonics, Billerica, MA, USA).

MALDI-FT-ICR MS

MALDI-FTMS measurements were performed at high spectral resolution and high lateral resolution on a 12T solariX™ (Bruker Daltonics, Billerica, MA) FT-ICR MS equipped with dual ESI-MALDI source and smartbeam^{TM} II 1kHz, 355 nm solid state Nd:YAG laser focused to a diameter of ~ 25 μm. Mass spectra were accumulated in positive ion mode and each spectrum consisted of 1 scan of 1000 laser shots from each spot, accumulated externally prior to measurement in the ICR cell over the mass range of *m/z* 250 – 4000. The array of spectra was collected at 25 μm intervals in both x and y dimensions across the surface of the STG. Data were acquired and analyzed using the flexImaging software (Bruker Daltonics, Billerica, MA), and digital images of the prepared sections were acquired prior to matrix coating.

Tissue extraction

Fifteen STGs were pooled, homogenized, and extracted with acidified methanol as described previously. [44] Extracts were concentrated in a Savant SC 110 Speedvac concentrator (Thermo Electron Corporation, West Palm Beach, FL, USA). The extract was subsequently re-suspended in 0.1% formic acid. The re-suspended extracts were then vortexed and briefly centrifuged. The resulting solution was purified and concentrated with C18 ZipTip (Millipore, Billerica, MA, USA). Briefly, the ZipTip C18 was first wetted using a ACN:water:formic acid (50:49.9:0.1;v/v/v) solution and then pre-equilibrated for sample binding with 0.1% formic acid in water. Subsequently, the tissue extract was loaded on the ZipTip C18. After washing with 0.1% formic acid in water for three times, the sample was eluted with 5 μL ACN:water:formic acid (50:49.9:0.1;v/v/v). Next, the eluent was dried and re-suspended in 0.1 % formic acid in water, and subjected to analysis by liquid chromatography-electrospray ionization-quadruple-time of flight (LC-ESI-Q-TOF).

NanoLC-ESI-Q-TOF MS/MS

Nanoscale LC-ESI-Q-TOF MS/MS was performed by a Waters nanoAcquity Ultra Performance capillary LC system coupled to a Q-TOF Micro mass spectrometer (Waters Corporation, Milford, MA, USA). The nanoflow ESI source conditions were set as follows: capillary voltage 3500 V, sample cone voltage 35 V, extraction cone voltage 1 V, source temperature 120°C, cone gas (N_2) 10L/hr. Chromatographic separations were performed on a homemade C18 reversed phase capillary column (75 μm internal diameter, 150 mm length, 3 μm particle size). The mobile phases used were: deionized water with 0.1% formic acid (A); acetonitrile with 0.1% formic acid (B). The re-

constituted extract was injected and loaded onto the trap column (Zorbax 300SB-C18 Nano trapping column, Agilent Technologies, Santa Clara, CA) using 95% mobile phase A and 5% mobile phase B at a flow rate of 10 μL/min for 10 min. Following this, the stream select module was switched to align the trap column with the analytical column, and a line gradient of mobile phases A and B was carried out, progressing from 5% to 45% over 60 min. A data dependent acquisition was employed for the MS survey scan and the selection of three precursor ions and subsequent MS/MS of the selected parent ions. The MS scan range was from *m/z* 400–1800 and the MS/MS scan was from *m/z* 50–1800. The MS/MS *de novo* sequencing was performed with a combination of manual sequencing and automatic sequencing by PepSeq software (Waters Corp., Milford, MA, USA).

Data analysis

Briefly, the peak lists of mass spectra from 4800 MALDI TOF/TOF were exported from the Data Explorer software, and pasted into Microsoft Excel. The peak lists were compared to an in-house database of crustacean neuropeptides, and filtered manually for known neuropeptide peaks. In this home-built neuropeptide database, we included both our own *de novo* sequencing results and neuropeptide sequences reported by other groups. Specifically, tblastn (http://www.ncbi.nlm.nih.gov/BLAST/) was used to mine for expressed sequence tags (ESTs) encoding putative *C. sapidus* peptide precursors (taxid: 6763) via queries using known arthropod preprohormone sequences. The hits were translated using the Translate tool of ExPASy (http://www.expasy.ch/tools/dna.html). The translated preprohormone sequences were used to predict mature peptide sequences

and assist our *de novo* sequencing process. The presence or absence of a peak was determined from each spectrum using 30 ppm as mass error tolerance. The peak list of mass spectra from solariX^{TM} MALDI-FTMS was processed by the flexImaging software, compared to an in-house database of crustacean neuropeptides, and filtered manually for known neuropeptide peaks with a mass error tolerance of 3 ppm.

5.3 Results and Discussion

Direct tissue profiling of neuropeptides in C. sapidus *STNS*

Characterization of neuropeptides in the STNS is inherently challenging due to the low pico- to sub-femtomolar *in vivo* concentrations of neuropeptides expressed in such small neural tissue. [45-47] The STNS lies between the brain and the suboesophageal ganglion (SOG) and consists of four ganglia together with connecting and motor nerves as shown in **Figure 1a**. The ganglia are the paired CoGs, OG, and STG. The OG lies between the two CoGs and is connected with them by the inferior esophageal nerves (*ion*). The STG is also connected with the CoGs by the superior esophageal nerves (*son*) and the *stn*. The STG is located within the ophthalmic artery and the substances released by the neurohaemal organs, the pericardial organs (PO), are transported to the STG via the artery by the heart. The STNS freshly dissected out of a blue crab *C. sapidus* and subjected to subsequent preparation for profiling and imaging experiments is shown in **Figure 1b**. A black sand paper was added beneath the STG as background for increased contrast during optical image scanning.

Direct tissue analysis of these ganglia provides a snapshot of the neuropeptides present in the tissues. However, the tube-like sheath wrapping around the ganglion

significantly hindered the effective detection of neuropeptides. To improve the sensitivity, we removed the sheath around the ganglia in the STNS. In **Figure 1c**, the center neuropil and the surrounding methylene blue-stained motorneurons are easily visualized under microscope after removing the sheath around the STG. Moreover, the neuropeptides are extracted more efficiently out of the ganglion by matrix solution, facilitating tissue profiling.

Figures 2a, b and c show representative annotated mass spectra collected by direct tissue analysis from the CoGs, STG and *stn* on MALDI-TOF/TOF, respectively. Peptides belonging to 10 different families were detected, including A-type and B-type allatostatins (AST), RFamide, orcokinin, tachykinin-related peptide (TRP), SIFamide, crustacean hyperglycemic hormone precursor related peptides (CPRP), pyrokinin, RYamide and YRamide. Distinct neuropeptide profiles were observed at different locations of the STNS, including both qualitative and relative quantitative differences. For instance, *Cancer borealis* TRP (CabTRP) 1a APSGFLGMRa (*m/z* 934.49) was detected as one of the most abundant peptides at all three locations; however, it only dominated the profiling spectrum of CoGs. Its isoform YPSGFLGMRa (*m/z* 1026.52), a recently discovered *Callinectes sapidus* TRP (CalsTRP), was also more concentrated in the CoGs than the STG or *stn*. The expression of TRP families in the STG, *stn* and CoG is in agreement with previous immunohistochemical studies visualizing their presence in the STNS of *Cancer borealis*. Moreover, both TRPs were reported to regulate feeding behavior in the crab *Callinectes sapidus* by exciting the gastric mill (food chewing) and pyloric (food filtering) rhythm. [48] An YRamide, HIGSLYRa (*m/z* 844.48), was the

second dominant peptide in the CoGs, whereas it exhibited relatively the same peak intensity as the other abundant RFamide and orcokinin peptides in the STG profiling spectrum. A SIFamide, GYRKPPFNGSIFa (*m/z* 1381.74), also displayed a similar trend as the YRamide in terms of relative intensity among the three locations. Interestingly, the RYamide family, including SGFYANRYa (*m/z* 976.46) and pEGFYSQRYa (*m/z* 1030.45), was detected exclusively in the CoGs, suggesting their significantly higher abundance in CoGs than STG. In contrast, B-type AST (AST-B), RFamide and orcokinin families were generally detected with higher signal intensities in the STG and *stn* than in the CoG, such as the AST-B peptides, AGWSSMRGAWa (*m/z* 1107.52), SGDWSSLRGAWa (*m/z* 1220.58) and VPNDWAHFRGSWa (*m/z* 1470.70). Nevertheless, exceptions were noticed; the orcokinin peptides NFDEIDRSGFGFA (*m/z* 1474.66) and NFDEIDRSSFGFA (*m/z* 1504.67) were more abundant in the STG, whereas the orcokinin NFDEIDRSSFGFN (*m/z* 1547.68) exhibited a significantly higher expression level than the other RFamides, AST-Bs or orcokinins in the CoGs. In contrast to AST-B, another family belonging to allatostatin, A-type allatostatin (AST-A) neuropeptides were detected with relatively low abundance. In addition, none of C-type ASTs (AST-C) were observed in this study, probably due to their low expression levels in *C. sapidus* STNS. The differential expression patterns of neuropeptides in the STNS demonstrate the complex structure of crustacean STNS and the heterogeneous distributions of multiple neuropeptides, indicating diverse roles of neuropeptides played in various physiological activities. Other than the STG, CoGs and *stn*, other neural organs and interconnecting nerves were also characterized. Nevertheless, only a few

neuropeptides were detected in the OG, probably due to the relatively low abundance of neuropeptides expressed in the OG and fewer neurons and much smaller tissue size. The nerves *son* that connect STG with CoG and the nerves *ion* that connect CoGs with OG, were also analyzed. Very limited intensity of signals arose from either *son* or *ion* (data not shown).

Although MALDI-TOF/TOF is capable to generate quick profiles of neuropeptides directly from tissue samples and even single cells, the high mass measurement accuracy and high mass resolving power achievable with FTMS-based instrumentation enables identification of neuropeptides from the complex neurosecretory organs and nerves in the crustacean model systems with improved confidence. Herein a high performance 12 Tesla MALDI-FT-ICR was employed to analyze neuropeptides expressed in the STG. The high quality MS data complemented and confirmed the MALDI-TOF/TOF results as detailed in **Table 1**. For example, the putative neuropeptide peak at *m/z* 1019.54 cannot be assigned to the known neuropeptide APRNFLRFamide (*m/z* 1019.59) based on a mass error tolerance of 30 ppm by mass matching with an in house *C. sapidus* neuropeptide database. However, the ultra-high resolving power of FT-ICR realized the separation of two ions with a 0.09 Da mass difference and detected two peaks at *m/z* of 1019.50 and 1019.59 Da, respectively. Therefore, the peak detected at *m/z* 1019.54 on MALDI-TOF/TOF could be attributed to an unresolved peak containing two molecular species that could only be unambiguously distinguished on an ultra-high performance mass analyzer. The utilization of FT-ICR can resolve closely related masses from a highly

complex neural tissue; moreover, it provides the requisite specificity for elucidating molecular species without a sacrifice in sensitivity.

A number of peptide peaks were detected on both instruments, whereas some were uniquely identified by one of the instruments, due to preferential detection of certain peptide ions with different mass analyzers. For example, the peptide GSNWSNLRGAWamide (*m/z* 1246.61) was detected on FT-ICR rather than TOF/TOF, whereas the TRP peptides, CabTRP 1a and CalsTRP, were not observed on FT-ICR but showed at relatively high intensity on TOF/TOF.

Peptide identification is greatly facilitated by the accurate mass measurements provided by FT-ICR. In addition, the neuropeptidome of other neural organs of *C. sapidus* has been characterized, providing a reservoir of potential neuropeptides that might be present in the STNS. Nevertheless, further confirmation by MS/MS fragmentation experiments is desirable for neuropeptide identification based solely upon accurate mass measurements. **Supplemental Figure S1a** shows the *in situ* collision-induced-dissociation (CID) MS/MS spectra of an orcokinin peptide NFDEIDRSGFGFA (*m/z* 1474.66) performed directly on the STG. As seen, almost complete series of fragment ions were detected, confirming high fidelity for peptide identification. **Figure S1b** is another MS/MS spectrum of a peptide at *m/z* 844.50. Based on its accurate mass and known sequence, the sequence was confirmed as HIGSLYRa. It is worth mentioning that the Ile/Leu ambiguity resulting from the isobaric nature of these amino acids has been identified through EST homology search from *Carcinus maenas*. [49] Intriguingly, several signature peaks typically belonging to lipid species were also observed in the

spectrum, such as neutral loss of trimethylamine group Δ59Da, product ion from the phosphocholine head chain (*m/z* 184.08) and fragments from choline (*m/z* 86.13 and 104.13). The accurate mass and the characteristic fragmentation pattern of phosphatidylcholine (PC) species allow the molecule to be sequenced as a K^+ adduct form of PC (38:6). The inclusion of the PC lipid at such a close mass (*m/z* 844.53) during direct tissue analysis also explains the detected mass to be 844.50 Da rather than the theoretical value of YRamide (*m/z* 844.48) alone. In comparison to *in situ* MS/MS performed on TOF/TOF, biochemical separation coupled to MS/MS sequencing via nanoLC-ESI-Q-TOF MS/MS was commonly exploited to confirm the identity of the known neuropeptides and to *de novo* sequence novel neuropeptides. **Supplemental Figure S2** shows two examples of MS/MS spectra acquired on ESI-Q-TOF, including CabTRP 1a (*m/z* 934.49, **Figure S2a**) and a SIFamide peptide (*m/z* 1381.74, **Figure S2b**). This classic approach combined with *in situ* MS/MS sequencing and direct tissue profiling resulted in the identification of 55 peptides from 10 families, including 50 neuropeptides discovered in other neural organs of *C. sapidus* like CalsTRP and 5 peptides previously identified in other species, but new to the *C. sapidus* species. Collectively, this work demonstrated the advantages of implementing combined MALDI-TOF/TOF and MALDI-FT-ICR platforms for enhanced sensitivity, speed and high mass resolution, for direct tissue profiling. The identification based on accurate mass matching and limited *in-situ* MS/MS via MALDI platform is complemented and validated by the employment of nanoLC-ESI tandem MS *de novo* peptide sequencing. The comprehensive characterization and mapping of various neuropeptides from *C. sapidus* STNS will

provide strong framework for future physiological investigations of neuropeptides in crustacean STNS.

High spatial resolution imaging of neuropeptides in C. sapidus *STG by MALDI-TOF/TOF*

Removal of the sheath from the STG is shown to improve the extraction and detection of neuropeptides for direct tissue analysis. Nevertheless, existence of the sheath maintains the integrity and intactness of the tissue as shown in **Figure 1d**, hindering the release of neuropeptides from the ganglion and preventing severe analyte diffusion during matrix application in subsequent IMS experiments. Consequently, sufficient extraction of the peptides from intact STG was achieved after multiple matrix coating. MS images that can be accurately co-registered with the optical image were obtained with intact STG, yet most neuropeptides distributed outside the edge of STG after matrix coating without the sheath. Hence, the sheath is kept intact in the application of STG IMS. The humidity of matrix application was adjusted accordingly to achieve maximal neuropeptide extraction out of the sheath while retaining their original localization.

A wealth of biochemical information is provided in a single MALDI-TOF/TOF IMS measurement and various neuropeptides exhibit significantly different or similar spatial distribution patterns. The utilization of autoflex III MALDI-TOF/TOF equipped with a highly focused laser smartbeam^{TM} allows for acquisition of multiple high resolution neuropeptide images from the STG using 25 μm as step size. Images of most neuropeptide isoforms detected in direct tissue analysis were acquired, consuming merely 85 min of instrument time. **Figure 3a** shows the optical image of a *C. sapidus* STG.

Representative MS images of the neuropeptides, including CabTRP 1a, three AST-Bs, three RFamides, a SIFamide and an orcokinin, are shown in **Figures 3b-3k**. All of the neuropeptides were localized within the STG tissue as shown in the optical image **Figure 3a**, suggesting negligible analyte diffusion due to the protection from the sheath. As seen in **Figure 2b**, the AST family dominated the profiling spectrum of the STG. The ASTs are neuropeptides that have important neuromodulatory roles in arthropods. Specifically, AST-B peptides are known for their myoinhibitory role and have been shown to regulate neuronal function in the STNS of *C. borealis*, such as inhibiting the pyloric rhythm in the STG. [12, 50] Most AST-B neuropeptides, such as AGWSSMRGAWa (*m/z* 1107.52), SGDWSSLRGAWa (*m/z* 1220.58) and VPNDWAHFRGSWa (*m/z* 1470.70), are expressed in high abundance in the bulbous structure of the neuropil which contains a large number of descending projection fibers and processes from upper modulatory ganglia. Moreover, the relative intensities of the three AST-B peptides (**Figures 3b-3d**) appear to be higher at the end towards *stn*. This is consistent with the fact that the AST-B peptides are detected with high intensity in *stn* shown in **Figure 2c** but not detectable in the nerve extended posteriorly toward the *dvn* (data not shown). A recent immunohistochemical study of VPNDWAHFRGSWa confirmed our observation. No immunoreactivity to VPNDWAHFRGSWa (named CbAST-B1 in that report) was detected in the STG somata, yet extensive staining in the STG neuropil and the fibers that projected from the *stn* into the neuropil was observed. [12] The antibody-based staining is capable to achieve high spatial resolution mapping of single cell and fibers. However, it provides fairly limited chemical information and lacks specificity to determine the

distribution of the different isoforms among AST-Bs or possibly all the AST family of peptides, even assuming the antibody employed is likely to have a higher affinity for the CbAST-B1. The utilization of IMS can generate images for each AST-B isoform unambiguously and combine these images to produce multiplexed images useful for determining peptide isoform co-localization and distinct distribution. In this case, the three AST-B peptides have similar structure and show colocalization within the STG, suggesting that the AST-B receptor might not distinguish among these peptides. Besides AST-Bs, the SIFamide peptide GYRKPPFNGSIFamide shows similar expression pattern in the STG neuropil as shown in **Figure 3e**. This Gly^1-SIFamide has been broadly reported in nervous system of crustacea, such as the giant tiger prawn *Penaeus monodon*, [51] Jonah crab *Cancer borealis* [52] and the crayfish family. [53] Nevertheless, immunohistochemical staining results have only been reported for Val^1-SIFamide in the STNS but not for Gly^1-SIFamide. Our work for the first time reported the localization of this specific SIFamide among many other neuropeptides, filling the gap for neurobiologists. Moreover, the co-existence of the peptides from two families possibly implies their synergistic interactions in crustacean STNS biological circuits. Other than AST-Bs, most RFamides were also expressed in high abundance in the STG and localized in a family colocalization pattern, although exceptions exist. As shown in **Figures 3f-h**, distribution of NRNFLRFa (*m/z* 965.54) resembled the AST-Bs and Gly^1-SIFamide, whereas SQPSKNYLRFa (*m/z* 1238.66) was localized more evenly within the STG neuropil, with pQDLDHVFLRFa (*m/z* 1271.64) being concentrated in the neuropil but more at the end close to *dvn*. This phenomenon suggests that neuropeptides belonging

to the same family yet with slightly different sequences could display varied localizations within the STNS, probably indicating their distinct biological functions. Clearly, IMS technique serves as the most suitable tool to examine the individual members from a complex composition of each neuropeptide family, facilitating the elucidation of possible distinct actions of various peptide isoforms within a superfamily. CabTRP 1a (*m/z* 934.49) in **Figure 3i** and the orcokinin peptide NFDEIDRSSFGFN (*m/z* 1547.68) in **Figure 3j** both evenly distributed within the STG neuropil region and appear to co-localize. Both peptides were expressed in high abundances in the STG, yet their distributions in the STG were examined for the first time in this IMS study. Despite the ultra-high resolution images provided by immunohistochemical staining (IHC), IMS complements this technique by enabling the mapping of multiple neuropeptides unambiguously with no prior knowledge at a spatial resolution that is high enough to differentiate various peptides within such a small organ, of which the length is around 1 mm and the width is less than 0.5 mm. Moreover, current advancements in reducing laser beam size will facilitate future applications of neuropeptide imaging at cellular level. [34]

Lipids were previously considered as interfering species for neuropeptide detection; however, increased interest has been drawn to understand potential roles of lipids present in the nervous system. **Figure 3k** shows the MS image of a lipid (*m/z* 806.57) that displays significantly higher intensity on the nerves and the sheath than that in the STG neuropil. This observation agrees well with previous knowledge that the sheath enclosing axons in invertebrates displayed remarkable similarities with vertebrate myelin sheath in terms of ultrastructure. [54] As expected, lipids distributed mostly in this whitish, fatty layer

rather than the neuropil. This observation illustrates a clearly distinct overall spatial localization patterns from majority of neuropeptides. Using the overlay function of the imaging software, a superimposed MS image of various molecular species, including the AST-B neuropeptide at *m/z* 1220.58 false-colored in red, the RFamide peptide of *m/z* 1271.64 coded in green and the lipid PC(38:6) shown in blue, is obtained, displaying the distinct distribution between neuropeptides and lipids within such a small ganglion. This intriguing observation implicates potentially different functions of the differentially-localized neuropeptides in the STG as well as the distinction between the peptides and lipids in the *C. sapidus* nervous system. The application of IMS at high spatial resolution offers a new level of understanding of the STNS by these well-resolved images of multiple molecular species and co-localization patterns of potentially interacting molecular players.

High mass spectral resolution imaging of neuropeptides in C. sapidus *STG by MALDI-FT-ICR*

High spatial resolution MS images are generated by TOF/TOF instruments rapidly, but these analyzers suffer from modest mass resolving power and mass accuracies. FT-ICR mass analyzer can provide comparable spatial resolutions to a TOF/TOF analyzer while offering superior resolution and accuracy for unequivocal assignment of analytes. [55] Therefore, MALDI-FT-ICR has herein been applied to MALDI IMS of the STG as a great alternative to TOF-based tissue imaging experiments. **Figure 4** demonstrates the essence of utilizing high-mass-accuracy and high-mass-resolution FT-ICR for neuropeptide imaging. **Figure 4a** is an *m/z* 1000-1500 spectrum averaging data acquired

throughout the entire STG, whereas insets **Figure 4b** and **4c** show two magnified mass ranges. In **Figure 4b**, the two different compounds at *m/z* 1019.50 and 1019.59 are resolved unequivocally, revealing the high resolution and quality of the FT-ICR-based IMS data. Moreover, these two species exhibit a different spatial distribution in the STG as shown in the images in **Figure 4d-e**. The lower-abundance peak of *m/z* 1019.50 was concentrated in the *stn* area as shown in **Figure 4d**, and the peak of *m/z* 1019.59 with higher signal intensity distributed mainly in the STG neuropil as displayed in **Figure 4e**. The overlaid image of these two compounds clearly demonstrated their distinct locations by false-color-coding in red and green, respectively. Based on the accurate mass obtained on FT-ICR and its concentrated distribution in the STG neuropil, the peak at *m/z* 1019.59 was assigned as the neuropeptide APRNFLRFamide with high confidence based on a 0.3 mDa mass error. In contrast, the modest mass spectral resolution of TOF/TOF cannot resolve the two compounds at 1019.50 Da and 1019.59 Da, and thus the discrimination cannot be made, resulting in an image combining the spatial distribution of these two compounds. The mass accuracy is also compromised by the lower spectral resolution of TOF/TOF, generating an unresolved peak at *m/z* 1019.54 in **Figure 2b**. Hence, the neuropeptide cannot be identified on a TOF platform, given that the mass error tolerance for neuropeptide assignment by mass measurement was set at 30 ppm. Another example involved the two peaks at *m/z* 1474.66 and 1474.74 that were separated in this very narrow mass range shown in **Figure 4c**. **Figure 4g** shows that the putative peptide of *m/z* 1474.66 distributed primarily in the STG neuropil and more concentrated at the end close to the *stn*, whereas the peak at *m/z* 1474.74 was documented to be localized in the STG

neuropil and also the connecting nerve in **Figure 4h**. The peak at *m/z* 1474.66 was assigned as an orcokinin peptide NFDEIDRSGFGFA by the mass and the distribution that resembles other orcokinin neuropeptides; nevertheless, the peak at *m/z* 1474.74 did not match any sequenced neuropeptides. In **Figure 4i**, the overlaid FT-ICR-MS image of the orcokinin neuropeptide and the putative peptide displayed overlapping but not identical expression patterns that cannot be differentiated by TOF/TOF. Moreover, the utilization of TOF/TOF resulted in a detected mass of *m/z* 1474.71 as previously shown in **Figure 4b**, probably due to the co-existence of the two analytes at *m/z* 1474.66 and 1474.74, respectively. The orcokinin could have been mis-assigned as an interference peak due to the 5 mDa mass error. These two examples clearly demonstrate the contribution of high mass resolution and accuracy provided by FT-ICR to the successful identification of neuropeptides from minute quantities of complex biological samples. Its combination with MALDI-TOF/TOF in the application of neuropeptide IMS provides unequivocal detailed localization information (as low as 25 μm lateral resolution) with the most comprehensive coverage of neuropeptides from the crustacean STG to date. This novel multi-faceted MS-based strategy can also be easily adapted to image other neural tissues, making this approach a valuable tool for neuropeptidomics studies.

An inherent characteristic of all FT-based analyses is that achievable mass resolving power is directly proportional to analysis time; hence the greater the desired mass resolving power, the longer one must measure ions in the detector. However, recent gains in FT-ICR technology, including the ability to desorb and accumulate ions externally during the detection event (Accumulate During Detect) has brought FT-ICR

measurement speeds close to those achieved with TOF. [56] As an example, the measurement region shown here consists of 5200 pixels analyzed at a rate of 1.33 seconds per pixel, or 45 pixels per minute with an achieved mass resolving power of 400,000 m/Δm at *m/z* 400, reflecting a 100-fold increase in spectral resolution over the TOF measurement (estimated 4000 m/Δm at *m/z* 400). This brings the total measurement time for the STG image to 115 minutes, or just under 2 hours for the 12 Tesla FT-ICR employed in the study, comparable to 85 min acquisition time on TOF/TOF. While higher magnetic field can utilize shorter transients without a sacrifice of spectral resolution, similar results could be achieved at lower field strengths with only a marginal increase in total measurement time. It is worth noting that for any given field strength, spectral resolution can be effectively doubled through application of a recently reported absorption mode phasing as an alternative to traditional magnitude mode signal processing [57].

As for the overlapping yet some distinct peptide detection results obtained from the TOF/TOF and FT-ICR instruments as shown in **Table 1**, this observation is not too surprising and has been previously documented in the studies of neuropeptides via direct tissue analysis. [15, 58] Although in theory one would expect that the same sample analyzed by two mass analyzers should reveal the same species, in practice the resulting mass spectral profiles can be affected by numerous factors including the instrumental configuration that impacts the ion transmission efficiency, ion transport time and the delay time between ion desorption/ionization and detection. Specifically, MALDI TOF/TOF enables fast acquisition and offers sensitive detection of a broad mass range of

peptides. The generated ions are quickly transferred and detected upon their formation in the MALDI source. In contrast, MALDI FT-ICR provides superior resolving power, but the relatively long pulse sequence and the space-charge effect could bias the detection of some low-abundance species. The ions produced by the external MALDI source undergo collisional cooling by nitrogen gas, storage in a hexapole followed by being transferred and analyzed in the ICR cell. This leads to the preservation and detection of more stable peptide ions [59] and an increase of the number of non-assigned peaks due to the fragmentation of less stable peptide ions. [60] All the factors discussed above can cause the differences in peptides we observed on two different instrumental platforms. However, the discrepancies can be interpreted as complementarities when results obtained by MALDI-TOF/TOF and MALDI-FT-ICR are combined. The complementarities between the TOF and FT-ICR-based instruments are appreciated by Römpp et al in a study of tryptic peptides that were differentially expressed in breast cancer patients. [61] da Silva et al also noticed an increased number of the detected BSA tryptic peptides and an improvement of sequence coverage when evaluating the combination of MALDI-TOF and MALDI-FT-ICR in peptide mass fingerprinting. [60]

IMS of C. sapidus STNS

The STNS of *C. sapidus* consists of the STG, the OG and the paired CoGs as well as a number of interconnecting and motor nerves. In the section of STNS profiling, we examined the neuropeptides from different regions of STNS; however, only the STG, CoG and *stn* generated neuropeptide-rich mass spectra. Only two orcokinin peptides, NFDEIDRSGFa (*m/z* 1198.55) and NFDEIDRSSFa (*m/z* 1228.56), and an AST-B

peptide VPNDWAHFRGSWa (*m/z* 1470.70) were detected in the OG. Moreover, no other putative neuropeptide peaks were detected from other regions in the STNS, probably due to their relatively low level of neuropeptide expression in the interconnecting nerves. Nevertheless, the first IMS study of neuropeptides from the crustacean STNS still demonstrated the potential of mapping multiple trace-level neuropeptides unambiguously and simultaneously from complex neural network. Representative neuropeptide images are shown in **Supplemental Figure S3**. **Figure S3a** presents the localization of the YRamide HIGSLYRa (*m/z* 844.48) that shows higher abundance in the STG and lower abundance in the CoG. This observation is also in agreement with the relative abundance of this neuropeptide observed in the STG and CoG profiling spectra in **Figure 2a** and **2b**. In **Figure S3b**, the orcokinin NFDEIDRSSFa (*m/z* 1228.56) is distributed throughout the STNS of *C. sapidus*, which agrees well with our previous profiling results. **Figure S3c** presents the localization of another orcokinin NFDEIDRSSFGFV (*m/z* 1532.70) that is particularly concentrated in the STG and the CoG. Other than neuropeptides, **Figure S3d** displays a potassiated lipid PC (36:1) at *m/z* 826.57, the abundance of which is higher in connecting nerves than ganglia in comparison to neuropeptides in the STNS. This observation is in agreement with the high-spatial-resolution images of the neuropeptides and lipids expressed in the STG as shown in **Figure 3**.

In summary, this integrative IMS platform presents a sensitive, reliable and high-throughput analytical tool for simultaneous mapping of signaling peptides and other types of biomolecules like lipids expressed in low abundance from tissue of minute size.

Knowledge of the complement and localization of neuropeptides in the *C. sapidus* STNS was directly obtained without tedious sample preparation step for liquid phase extraction or immunohistochemical techniques. Moreover, the wealth of information gained from IMS of the STNS provides unprecedented molecular details regarding chemical identities and anatomical distributions of these signaling peptides in this classical model nervous system.

5.4 Conclusions

For the first time, MALDI-TOF/TOF and MALDI-FT-ICR were employed for mapping the localizations of neuropeptides at the network and cellular resolution using the *C. sapidus* STNS as a model system. A novel multi-faceted MS strategy which combined direct tissue analysis, *in situ* MS/MS and micro-separation coupled to ESI-Q-TOF tandem MS peptide sequencing was employed here. In total, 55 neuropeptides were characterized from *C. sapidus* STNS. The central component of the STNS, STG, was imaged for neuropeptide distribution with both TOF/TOF and FT-ICR at a spatial resolution of 25 μm, generating detailed distribution maps of multiple neuropeptides and lipids simultaneously from this ganglion of minute size. The advantages of FT-ICR added to this platform were manifested by the well-resolved MS images of neuropeptides that could not be distinguished on TOF/TOF. The mass spectral imaging of small STG by FT-ICR combined with imaging of the STNS by TOF/TOF makes the approach a great tool to map neuropeptides in a high-throughput fashion with detailed information gained for the center of the nervous system. In addition, IMS has been applied to image the entire

STNS using TOF/TOF due to its fast analysis speed. MS/MS analysis of neuropeptides further elucidated the peptide sequences, rendering a comprehensive study of the complement and localization of neuropeptides in crustacean STNS. The multi-faceted MS-based platform established in this study will benefit the field of neuropeptide imaging and facilitate future investigations on the interplay of a myriad of signaling peptides and neurotransmitters in a neuronal network.

Acknowledgments

This work is supported in part by the National Science Foundation grant (CHE-0967784), and National Institutes of Health grants (1R01DK071801, 1R56DK071801). The authors thank Bruker Daltonics for graciously loaning the Autoflex III MALDI TOF/TOF mass spectrometer. L. Li acknowledges an H.I. Romnes Faculty Research Fellowship.

References

1. Marder, E.; Bucher, D. (2001) Central pattern generators and the control of rhythmic movements. *Curr. Biol.* 11, R986-R996.
2. Stemmler, E. A.; Peguero, B.; Bruns, E. A.; Dickinson, P. S.; Christie, A. E. (2007) Identification, physiological actions, and distribution of TPSGFLGMRamide: a novel tachykinin-related peptide from the midgut and stomatogastric nervous system of Cancer crabs. *J. Neurochem.* 101, 1351-1366.
3. Christie, A. E.; Stemmler, E. A.; Peguero, B.; Messinger, D. I.; Provencher, H. L.; Scheerlinck, P.; Hsu, Y. W. A.; Guiney, M. E.; de la Iglesia, H. O.; Dickinson, P. S. (2006) Identification, physiological actions, and distribution of VYRKPPFNGSIFamide (Val(1)-SIFamide) in the stomatogastric nervous system of the American lobster *Homarus americanus*. *J. Comp. Neurol.* 496, 406-421.
4. Skiebe, P.; Dreger, M.; Meseke, M.; Evers, J. F.; Hucho, F. (2002) Identification of orcokinins in single neurons in the stomatogastric nervous system of the crayfish, *Cherax destructor*. *J. Comp. Neurol.* 444, 245-259.
5. Nusbaum, M. P. (2002) Regulating peptidergic modulation of rhythmically active neural circuits. *Brain Behav. Evolut.* 60, 378-387.
6. Nusbaum, M. P.; Beenhakker, M. P. (2002) A small-systems approach to motor pattern generation. *Nature* 417, 343-350.
7. Marder, E. (2000) Motor pattern generation. *Curr.Opin.in Neurobiol.* 10, 691-698.
8. Marder, E.; Calabrese, R. L. (1996) Principles of rhythmic motor pattern generation. *Physiol. Rev.* 76, 687-717.
9. Dickinson, P. S.; Mecsas, C.; Marder, E. (1990) Neuropeptide Fusion of 2 Motor-Pattern Generator Circuits. *Nature* 344, 155-158.
10. Marder, E.; Bucher, D. (2007) Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annu. Rev. Physiol.* 69, 291-316.
11. Sahley, C. L. (1993) Neural integration. *Science* 260, 244.
12. Szabo, T. M.; Chen, R. B.; Goeritz, M. L.; Maloney, R. T.; Tang, L. S.; Li, L. J.; Marder, E. (2011) Distribution and Physiological Effects of B-Type Allatostatins (Myoinhibitory Peptides, MIPs) in the Stomatogastric Nervous System of the Crab Cancer borealis. *J. Comp. Neurol.* 519, 2658-2676.
13. Li, C.; Kim, K. (2009) *Neuropeptides*. In WormBook,
14. Nassel, D. R. (2002) Neuropeptides in the nervous system of Drosophila and other insects: multiple roles as neuromodulators and neurohormones. *Prog. Neurobiol.*68, 1-84.
15. Ma, M.; Chen, R.; Sousa, G. L.; Bors, E. K.; Kwiatkowski, M. A.; Goiney, C. C.; Goy, M. F.; Christie, A. E.; Li, L. (2008) Mass spectral characterization of peptide transmitters/hormones in the nervous system and neuroendocrine organs of the American lobster *Homarus americanus*. *Gen. Comp. Endocrinol.* 156, 395-409.
16. Ma, M. M.; Wang, J. H.; Chen, R. B.; Li, L. J. (2009) Expanding the crustacean neuropeptidome using a multifaceted mass spectrometric approach. *J. Proteome Res.* 8, 2426-2437.
17. Kutz, K. K.; Schmidt, J. J.; Li, L. J. (2004) *In situ* tissue analysis of neuropeptides by MALDI FTMS in-cell accumulation. *Anal. Chem.* 76, 5630-5640.
18. Stemmler, E. A.; Cashman, C. R.; Messinger, D. I.; Gardner, N. P.; Dickinson, P. S.; Christie, A. E. (2007) High-mass-resolution direct-tissue MALDI-FTMS reveals broad conservation of three neuropeptides (APSGFLGMRamide, GYRKPPFNGSIFamide and

pQDLDHVFLRFamide) across members of seven decapod crustaean infraorders. *Peptides* 28, 2104-15.
19. Stemmler, E. A.; Provencher, H. L.; Guiney, M. E.; Gardner, N. P.; Dickinson, P. S. (2005) Matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for the identification of orcokinin neuropeptides in crustaceans using metastable decay and sustained off-resonance irradiation. *Anal. Chem.* 77, 3594-3606.
20. Defelipe, J. (1993) Neocortical neuronal diversity - chemical heterogeneity revealed by colocalization studies of classic neurotransmitters, neuropeptides, calcium-binding proteins, and cell-surface molecules. *Cereb. Cortex* 3, 273-289.
21. Skiebe, P. (2003) Neuropeptides in the crayfish stomatogastric nervous system. *Microsc. Res. Technique* 60, 302-312.
22. Ye, H.; Greer, T.; Li, L. J. (2011) From pixel to voxel: a deeper view of biological tissue by 3D mass spectral imaging. *Bioanalysis* 3, 313-332.
23. Nemes, P.; Vertes, A. (2007) Laser ablation electrospray ionization for atmospheric pressure, in vivo, and imaging mass spectrometry. *Anal.Chem.* 79, 8098-8106.
24. Li, Y.; Shrestha, B.; Vertes, A. (2008) Atmospheric pressure infrared MALDI imaging mass spectrometry for plant metabolomics. *Anal. Chem.* 80, 407-420.
25. Wiseman, J. M.; Ifa, D. R.; Song, Q. Y.; Cooks, R. G. (2006) Tissue imaging at atmospheric pressure using desorption electrospray ionization (DESI) mass spectrometry. *Angew. Chem. Int. Edit.* 45, 7188-7192.
26. Galhena, A. S.; Harris, G. A.; Nyadong, L.; Murray, K. K.; Fernandez, F. M. (2010) Small molecule ambient mass spectrometry imaging by infrared laser ablation metastable-induced chemical ionization. *Anal. Chem.* 82, 2178-2181.
27. Stoeckli, M.; Chaurand, P.; Hallahan, D. E.; Caprioli, R. M. (2001) Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat. Med*. 7, 493-6.
28. Skold, K.; Svensson, M.; Nilsson, A.; Zhang, X. Q.; Nydahl, K.; Caprioli, R. M.; Svenningsson, P.; Andren, P. E. (2006) Decreased striatal levels of PEP-19 following MPTP lesion in the mouse. *J. Proteome Res.* 5, 262-269.
29. Hanrieder, J.; Ljungdahl, A.; Falth, M.; Mammo, S. E.; Bergquist, J.; Andersson, M. (2011) L-DOPA-induced Dyskinesia is Associated with Regional Increase of Striatal Dynorphin Peptides as Elucidated by Imaging Mass Spectrometry. *Mol. Cell. Proteomics* 10, doi: 10.1074/mcp.M111.009308.
30. Monroe, E. B.; Annangudi, S. R.; Hatcher, N. G.; Gutstein, H. B.; Rubakhin, S. S.; Sweedler, J. V. (2008) SIMS and MALDI MS imaging of the spinal cord. *Proteomics* 8, 3746-3754.
31. Cornett, D. S.; Frappier, S. L.; Caprioli, R. M. (2008) MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. *Anal. Chem.* 80, 5648-5653.
32. McDonnell, L. A.; Heeren, R. M. A. (2007) Imaging mass spectrometry. *Mass Spectrom. Rev.* 26, 606-643.
33. Ye, H.; Greer, T.; Li, L. Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *J Proteomics*. 75, 5014-26.
34. Römpp, A.; Guenther, S.; Schober, Y.; Schulz, O.; Takats, Z.; Kummer, W.; Spengler, B. (2010) Histology by Mass Spectrometry: Label-Free Tissue Characterization Obtained from High-Accuracy Bioanalytical Imaging. *Angew. Chem. Int. Edit.* 49, 3834-3838.
35. Klerk, L. A.; Altelaar, A. F. M.; Froesch, M.; McDonnell, L. A.; Heeren, R. M. A. (2009) Fast and automated large-area imaging MALDI mass spectrometry in microprobe and microscope mode. *Int. J. Mass Spectrom.* 285, 19-25.

36. Jurchen, J. C.; Rubakhin, S. S.; Sweedler, J. V. (2005) MALDI-MS imaging of features smaller than the size of the laser beam. *J. Am. Soc. Mass Spectrom.* 16, 1654-1659.
37. Altelaar, A. F. M.; Taban, I. M.; McDonnell, L. A.; Verhaert, P. D. E. M.; de Lange, R. P. J.; Adan, R. A. H.; Mooi, W. J.; Heeren, R. M. A.; Piersma, S. R. (2007) High-resolution MALDI imaging mass spectrometry allows localization of peptide distributions at cellular length scales in pituitary tissue sections. *Int. J. Mass Spectrom.* 260, 203-211.
38. Taban, I. M.; Altelaar, A. F. M.; Van der Burgt, Y. E. M.; McDonnell, L. A.; Heeren, R. M. A.; Fuchser, J.; Baykut, G. (2007) Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* 18, 145-151.
39. DeKeyser, S. S.; Kutz-Naber, K. K.; Schmidt, J. J.; Barrett-Wilt, G. A.; Li, L. J. (2007) Imaging mass spectrometry of neuropeptides in decapod crustacean neuronal tissues. *J. Proteome Res.* 6, 1782-1791.
40. Chen, R. B.; Hui, L. M.; Sturm, R. M.; Li, L. J. (2009) Three Dimensional Mapping of Neuropeptides and Lipids in Crustacean Brain by Mass Spectral Imaging. *J. Am. Soc. Mass Spectrom.* 20, 1068-1077.
41. Verhaert, P. D.; Conaway, M. C. P.; Pekar, T. M.; Miller, K. (2007) Neuropeptide imaging on an LTQ with vMALDI source: The complete 'all-in-one' peptidome analysis. *Int. J. Mass Spectrom.* 260, 177-184.
42. Verhaert, P. D.; Pinkse, M. W.; Strupat, K.; Conaway, M. C. Imaging of similar mass neuropeptides in neuronal tissue by enhanced resolution MALDI MS with an ion trap - Orbitrap hybrid instrument. *Methods Mol. Biol.* 656, 433-49.
43. Selverston, A. I.; Russell, D. F.; Miller, J. P. (1976) The stomatogastric nervous system: structure and function of a small neural network. *Prog. Neurobiol.* 7, 215-90.
44. Cape, S. S.; Rehm, K. J.; Ma, M.; Marder, E.; Li, L. J. (2008) Mass spectral comparison of the neuropeptide complement of the stomatogastric ganglion and brain in the adult and embryonic lobster, Homarus americanus. *J. Neurochem.*105, 690-702.
45. Li, L. J.; Sweedler, J. V. (2008) Peptides in the Brain: Mass Spectrometry-Based Measurement Approaches and Challenges. *Ann.Rev. Anal. Chem.* 1, 451-483.
46. Kendrick, K. M. (1990) Microdialysis measurement of in vivo neuropeptide release. *J. Neurosci. Meth.* 34, 35-46.
47. Bungart, D.; Dircksen, H.; Keller, R. (1994) Quantitative determination and distribution of the myotropic neuropeptide orcokinin in the nervous system of astacidean crustaceans. *Peptides* 15, 393-400.
48. Hui, L. M.; Zhang, Y. Z.; Wang, J. H.; Cook, A.; Ye, H.; Nusbaum, M. P.; Li, L. J. (2011) Discovery and Functional Study of a Novel Crustacean Tachykinin Neuropeptide. *ACS Chem.Neurosci.* 2, 711-722.
49. Christie, A. E.; Cashman, C. R.; Brennan, H. R.; Ma, M. M.; Sousa, G. L.; Li, L. J.; Stemmler, E. A.; Dickinson, P. S. (2008) Identification of putative crustacean neuropeptides using in silico analyses of publicly accessible expressed sequence tags. *Gen. Comp. Endocrinol.* 156, 246-264.
50. Nassel, D. R.; Winther, A. M. E. (2010) Drosophila neuropeptides in regulation of physiology and behavior. *Prog. Neurobiol.* 92, 42-104.
51. Sithigorngul, P.; Pupuem, J.; Krungkasem, C.; Longyant, S.; Chaivisuthangkura, P.; Sithigorngul, W.; Petsom, A. (2002) Seven novel FMRFamide-like neuropeptide sequences from the eyestalk of the giant tiger prawn Penaeus monodon. *Comp. Biochem. Phys. B* 131, 325-337.
52. Huybrechts, J.; Nusbaum, M. P.; Vanden Bosch, L.; Baggerman, G.; De Loof, A.; Schoofs, L. (2003) Neuropeptidomic analysis of the brain and thoracic ganglion from the Jonah crab, Cancer borealis. *Biochem. Bioph. Res. Co.* 308, 535-544.

53. Dickinson, P. S.; Stemmler, E. A.; Cashman, C. R.; Brennan, H. R.; Dennison, B.; Huber, K. E.; Peguero, B.; Rabacal, W.; Goiney, C. C.; Smith, C. M.; Towle, D. W.; Christie, A. E. (2008) SIFamide peptides in clawed lobsters and freshwater crayfish (Crustacea, Decapoda, Astacidea): A combined molecular, mass spectrometric and electrophysiological investigation. *Gen. Comp. Endocrinol.* 156, 347-360.
54. Pipa, R. L. (1961) Studies on the hexapod nervous system. III. Histology and histochemistry of cockroach neuroglia. *J. Comp. Neurol.* 116, 15-26.
55. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. (1998) Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* 17, 1-35.
56. Senko, M. W.; Hendrickson, C. L.; Emmett, M. R.; Shi, S. D. H.; Marshall, A. G. (1997) External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* 8, 970-976.
57. Qi, Y. L.; Barrow, M. P.; Van Orden, S. L.; Thompson, C. J.; Li, H. L.; Perez-Hurtado, P.; O'Connor, P. B. (2011) Variation of the Fourier transform mass spectra phase function with experimental parameters. *Anal. Chem.* 83, 8477-8483.
58. Chen, R. B.; Jiang, X. Y.; Conaway, M. C. P.; Mohtashemi, I.; Hui, L. M.; Viner, R.; Li, L. J. (2010) Mass Spectral Analysis of Neuropeptide Expression and Distribution in the Nervous System of the Lobster *Homarus americanus*. *J. Proteome Res.* 9, 818-832.
59. O'Connor, P. B.; Costello, C. E. (2001) A high pressure matrix-assisted laser desorption/ionization Fourier transform mass spectrometry ion source for thermal stabilization of labile biomolecules. *Rapid Commun. Mass Spectrom.* 15, 1862-1868.
60. da Silva, D.; Wasselin, T.; Carre, V.; Chaimbault, P.; Bezdetnaya, L.; Maunit, B.; Muller, J. F. (2011) Evaluation of combined matrix-assisted laser desorption/ionization time-of-flight and matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance mass spectrometry experiments for peptide mass fingerprinting analysis. *Rapid Commun. Mass Spectrom.* 25, 1881-1892.
61. Römpp, A.; Dekker, L.; Taban, I.; Jenster, G.; Boogerd, W.; Bonfrer, H.; Spengler, B.; Heeren, R.; Smitt, P. S.; Luider, T. M. (2007) Identification of leptomeningeal metastasis-related proteins in cerebrospinal fluid of patients with breast cancer by a combination of MALDI-TOF, MALDI-FTICR and nanoLC-FTICR MS. *Proteomics* 7, 474-481.

Table 1. Peptides identified from *C. sapidus* STG by MALDI-TOF/TOF, MALDI-FTICR and ESI-Q-TOF MS platforms.

m/z	Sequence	TOF/TOF	FT-ICR	Q-TOF	Other locations[a]
AST-A					
909.4941	ARPYSFGLa	+	-	-	CoG/*stn*
937.4890	PRDYAFGLa	+	-	-	CoG/*stn*
962.5094	APQPYAFGLa	-	+	-	*stn*
AST-B					
1107.5153	AGWSSMRGAWa	+	+	+	CoG/*stn*
1123.5147	AGWSSM(O)RGAWa	-	+	-	
1165.5538	NWNKFQGSWa	-	+	-	
1182.5691	TSWGKFQGSWa	+	-	-	*stn*
1209.5800	TGWNKFQGSWa	+	-	-	*stn*
1220.5807	SGDWSSLRGAWa	+	+	+	*stn*
1222.5752	GNWNKFQGSWa	+	-	+	*stn*
1246.6076	GSNWSNLRGAWa	-	+	+	
1252.5858	NNWSKFQGSWa	+	-	-	*stn*
1293.6335	STNWSSLRSAWa	+	+	+	CoG/*stn*
1470.7025	VPNDWAHFRGSWa	+	+	+	CoG/*stn/OG*
Orcokinin					
1062.5578	TPRDIANLYa	-	+	-	
1156.5269	FDEIDRSGFA	+	+	-	
1186.5375	FDEIDRSSFA	+	+	+	CoG/*stn*
1198.5487	NFDEIDRSGFa	+	-	-	OG
1228.5593	NFDEIDRSSFa	+	-	-	*Stn/OG*
1256.5542	NFDEIDRSGFG	+	+	-	
1286.5648	NFDEIDRSSFG	+	-	-	
1403.6226	NFDEIDRSGFGF	+	-	+	CoG/*stn*
1433.6332	NFDEIDRSSFGF	+	-	+	*stn*
1474.6597	NFDEIDRSGFGFA	+	+	+	CoG/*stn*
1502.6910	NFDEIDRSGFGFV	+	-	+	CoG
1504.6703	NFDEIDRSSFGFA	+	-	+	CoG/*stn*
1532.7016	NFDEIDRSSFGFV	+	-	+	CoG/*stn*
1547.6761	NFDEIDRSSFGFN	+	-	+	CoG/*stn*
1554.6972	NFDEIDRTGFGFH	+	-	-	CoG/*stn*
RFamide					

948.5414	PGVNFLRFa	+	+	+	
965.5428	NRNFLRFa	+	-	+	CoG/*stn*
1005.5741	GRPNFLRFa	+	-	-	*stn*
1019.5897	APRNFLRFa	+	+	-	*stn*
1022.5643	GNRNFLRFa	+	-	-	CoG/*stn*
1073.5527	TNYGGFLRFa	+	+	-	CoG/*stn*
1104.6061	GAHKNYLRFa	+	+	-	CoG/*stn*
1119.6455	SMPTLRLRFa	+	-	-	CoG/*stn*
1124.6323	GLSRNYLRFa	+	-	-	CoG/*stn*
1147.6483	APQRNFLRFa	+	-	+	CoG/*stn*
1158.6167	YGNRSFLRFa	+	+	-	CoG
1238.6640	SQPSKNYLRFa	-	+	-	
1271.6394	pQDLDHVFLRFa	+	+	+	CoG
1287.6844	GYSVGLNYLRFa	+	+	+	
1288.6797	QDLDHVFLRFa	+	-	+	CoG
1314.7753	DARTAPLRLRFa	+	-	+	CoG/*stn*
CabTRP					
934.4927	APSGFLGMRa	+	-	+	CoG/*stn*
1026.5189	YPSGFLGMRa	+	-	+	CoG
SIFa					
1381.7375	GYRKPPFNGSIFa	+	+	+	CoG/*stn*
CPRP					
1033.5207	RSAEGLGRMG	+	-	-	*stn*
1478.7485	TPLGDLSGSLGHPVE	+	+	-	
Pyrokinin					
878.5247	LYFAPRLamide	+	-	-	CoG
1037.5527	SGGFAFSPRLa	-	+	-	*stn*
RYamide					
976.4635	SGFYANRYa	-	-	-	CoG
1030.4508	pEGFYSQRYa	-	-	-	CoG
YRamide					
844.4788	HIGSLYRa	+	-	+	CoG

[a]The neuropeptides identified from other locations of STNS were conducted on MALDI-TOF/TOF.

Figures:

Figure 1. Images of *C. sapidus* STNS and STG. (a) A schematic drawing of the *C. sapidus* STNS, with the major ganglia and nerves labeled (figure courtesy Dr. Nadim and Dr. Marder labs). (b) An optical image of the isolated *C. sapidus* STNS mounted on a glass slide and subjected to subsequent IMS experiments. (c) An optical micrograph of a desheathed STG that clearly revealed its neuropil and surrounding motorneurons after methylene blue staining. (d) An optical image of the isolated STG mounted on a glass slide under microscope. The corresponding scale bar is annotated in each panel.

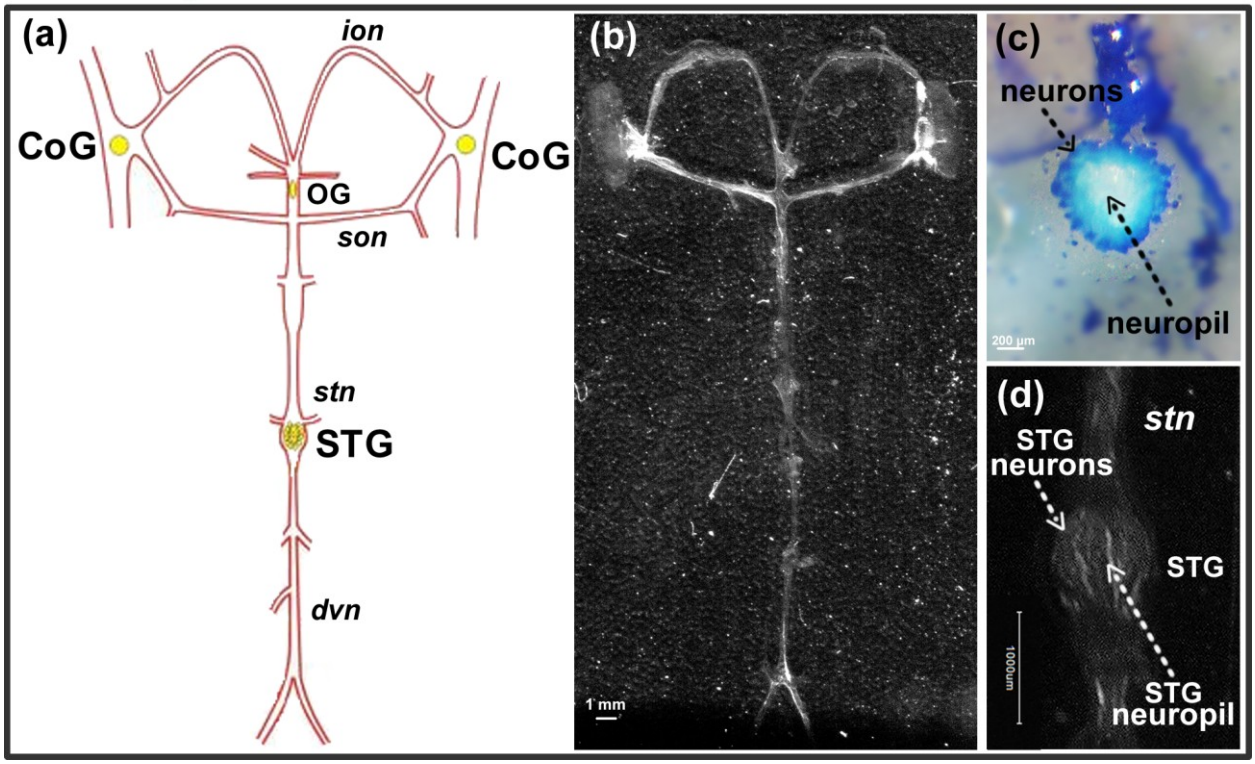


Figure 2. Comparison of direct tissue mass spectra from three discrete locations of *C. sapidus* STNS, including (a) CoG, (b) STG and (c) *stn*. Peptides are annotated with molecular weight and their corresponding families. Note that not all neuropeptides detected via direct tissue analysis are evident in these representative spectra. The peaks labeled with asterisks are assigned as interferents.

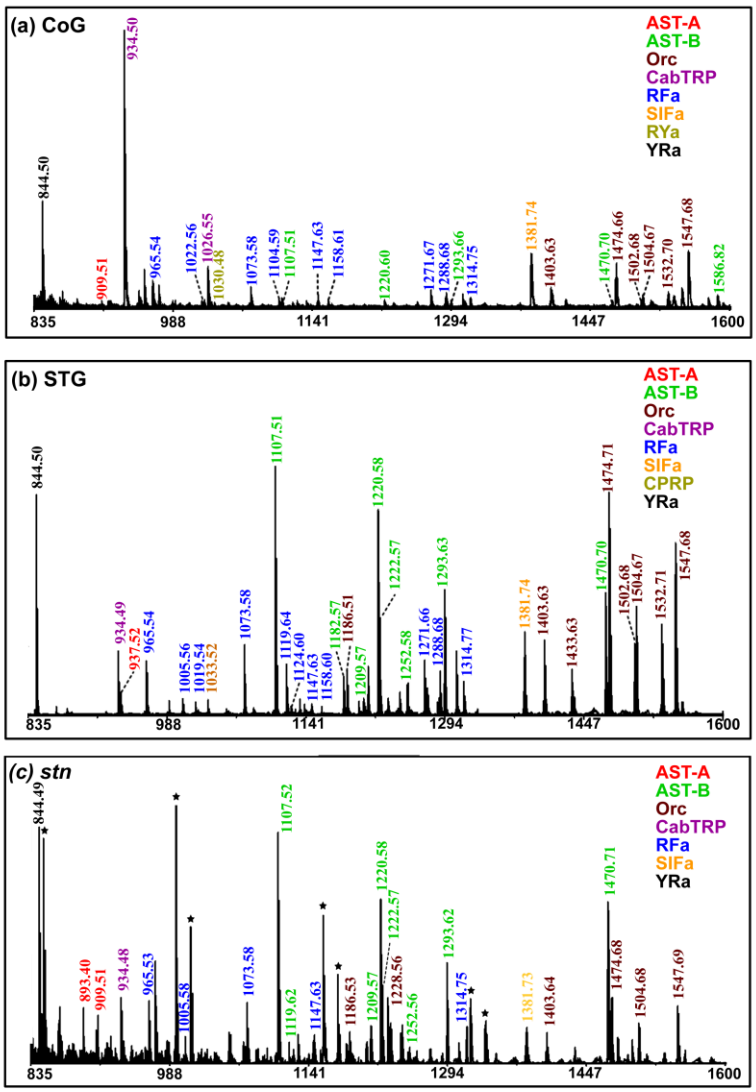


Figure 3. Distribution of neuropeptides in *C. sapidus* STG revealed by IMS on MALDI-TOF/TOF. (a) An optical image of the STG subjected to subsequent IMS preparation. Nine neuropeptides were shown, including three B-type ASTs: (b) AGWSSMRGAWa (*m/z* 1107.52), (c) SGDWSSLRGAWa (*m/z* 1220.58) and (d) VPNDWAHFRGSWa (*m/z* 1470.70); (e) SIFamide GYRKPPFNGSIFa (*m/z* 1381.74); three RFamides: (f) NRNFLRFa (*m/z* 965.54), (g) SQPSKNYLRFa (*m/z* 1238.66) and (h) pQDLDHVFLRFa (*m/z* 1271.64); (i) CabTRP 1a APSGFLGMRa (*m/z* 934.49); (j) Orcokinin NFDEIDRSSFGFN (*m/z* 1547.68). In addition, the distribution of (k) a lipid PC (38:6) (*m/z* 806.57) was shown. The observed masses are labeled in white above the MS images. Different distribution patterns are shown between different families and also isoforms within individual families. (l) is an overlaid image of (c), (h) and (k) as displayed in green, red and blue, respectively.

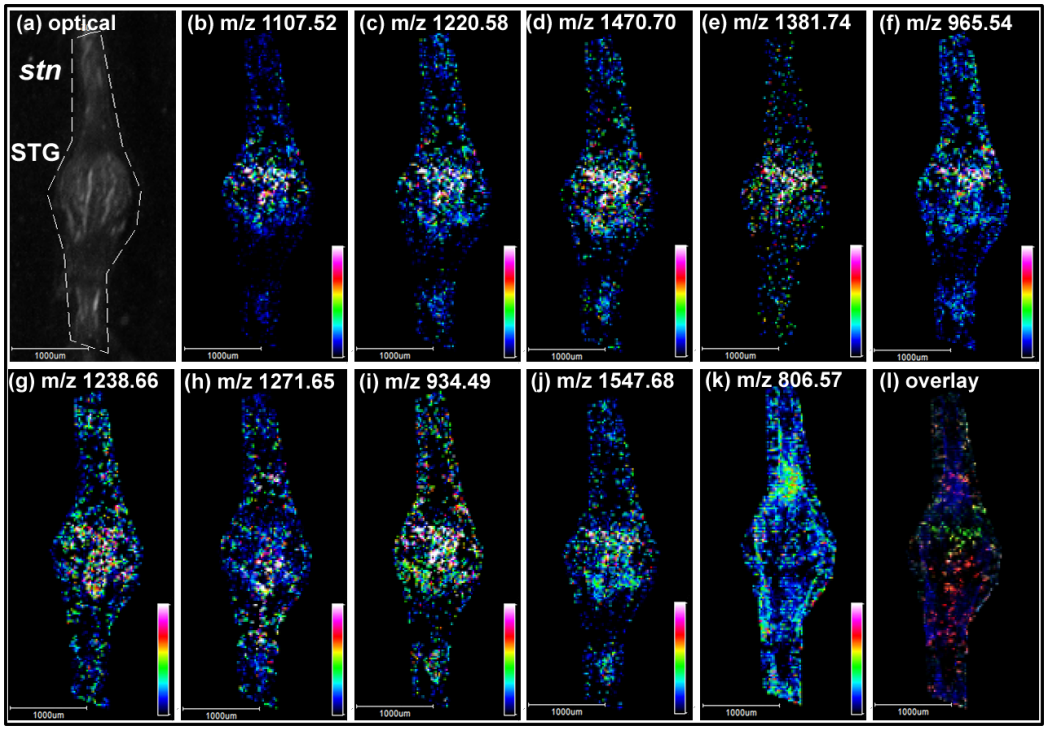
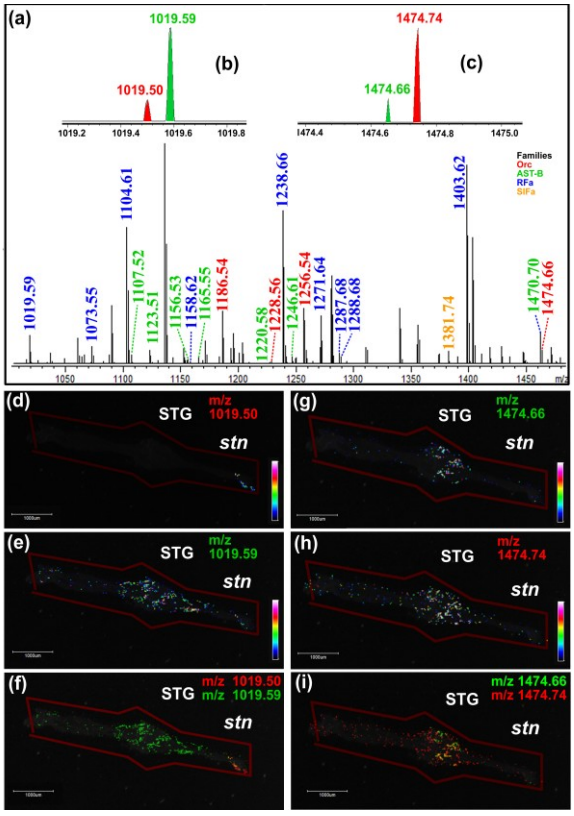
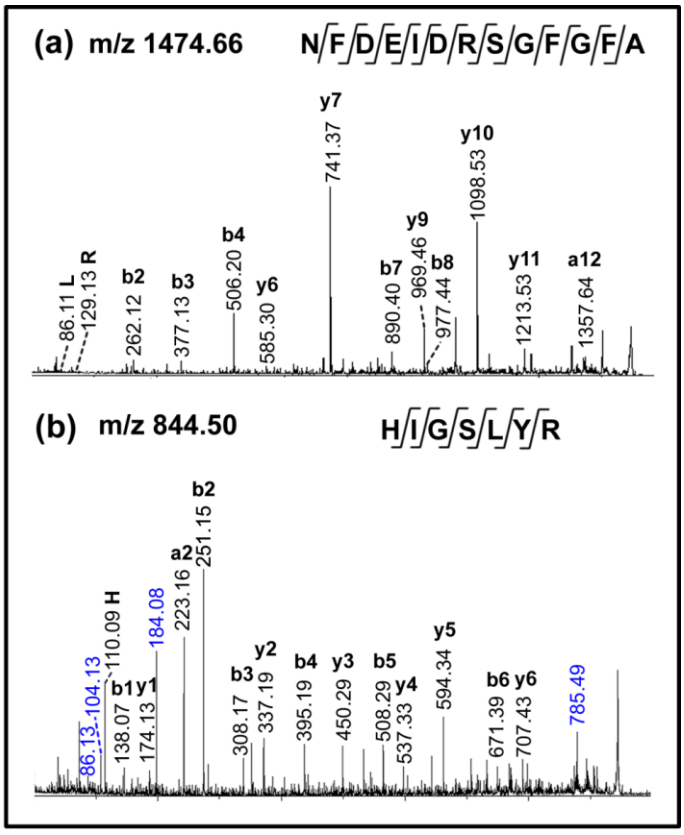


Figure 4. Localization of neuropeptides in *C. sapidus* STG revealed by IMS on MALDI-FT-ICR. (a) Total MALDI-FT-ICR mass spectrum of the *C. sapidus* STG. Peptides are annotated with molecular weight and color-coded according to their corresponding families. (b) and (c) are two inserts showing the zoom-in spectra of two ion clusters *m/z* 1019-1020 and 1474-1475. (d) and (e) show the distribution of the ions at *m/z* 1019.50 and 1019.59. (f) shows an overlaid image of (d) and (e), visualizing the different distribution of the *m/z* 1019.50 ion displayed in green and the *m/z* 1019.59 ion displayed in red. Similarly, (g) and (h) manifested the localization of the peaks at *m/z* 1474.66 and 1474.74, respectively. (i) is a superimposed image of the 1474.66 Da ion displayed in green and the 1474.74 Da ion displayed in red, visually demonstrating their distinct localization in the STG.

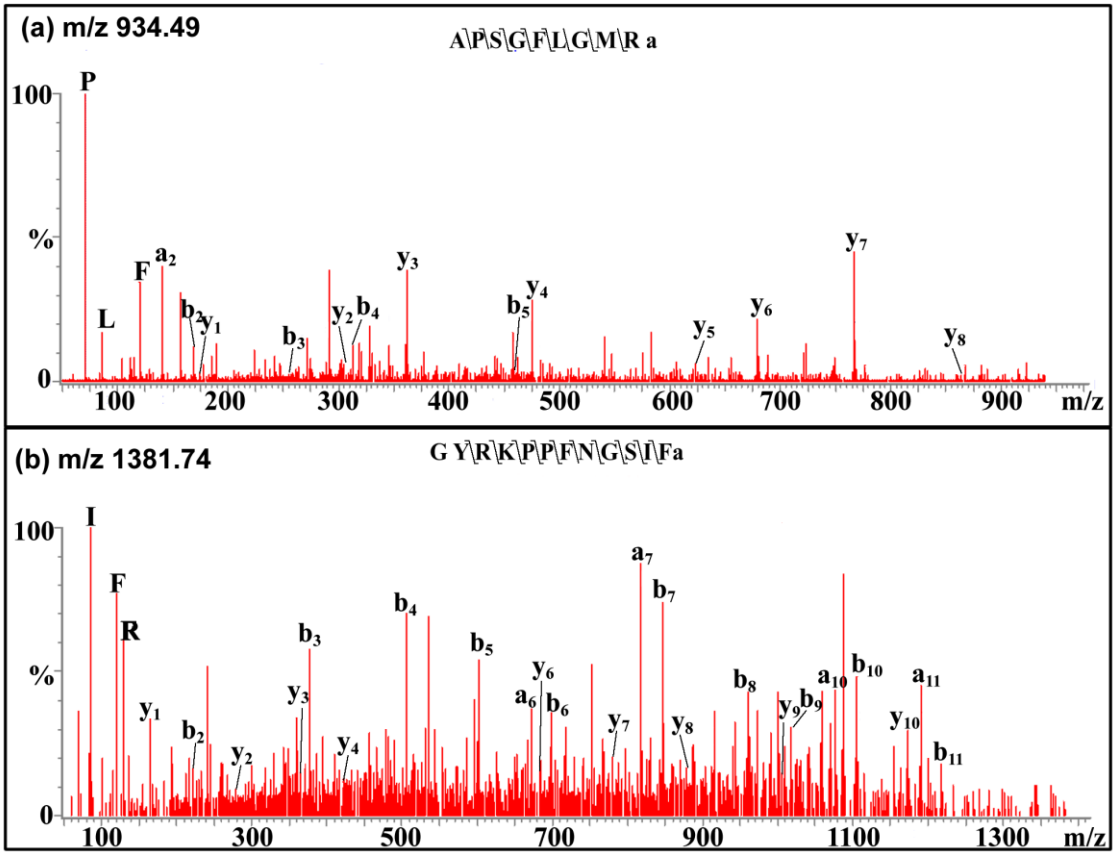


Supplemental Information

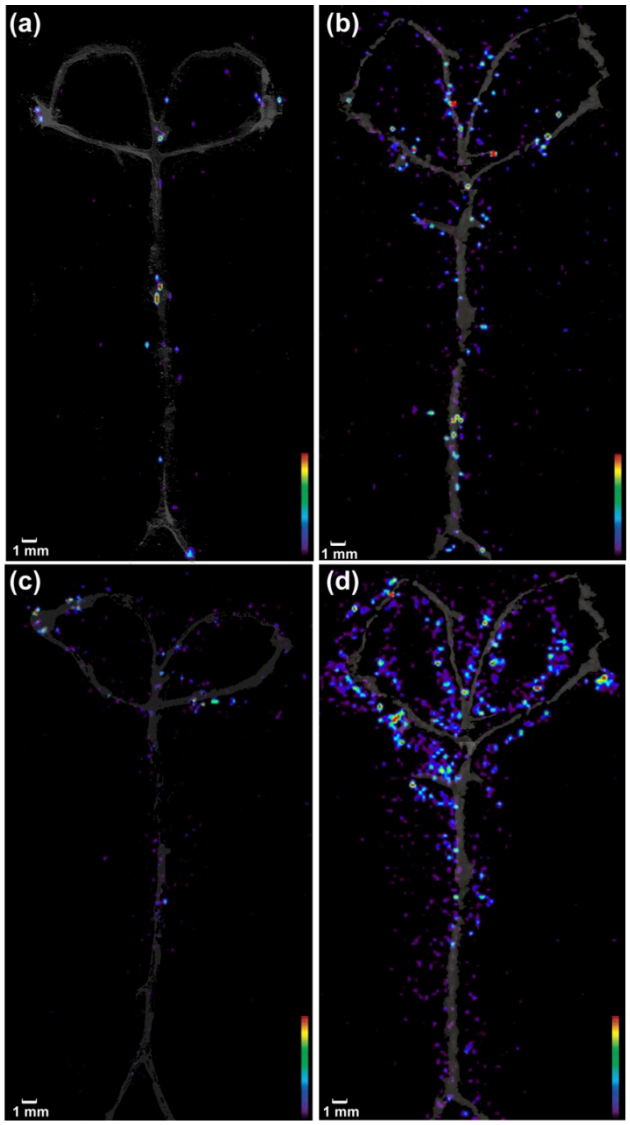
Supplemental Figure S1. Representative tandem mass spectra highlight MS/MS sequencing of two neuropeptides obtained directly from the *C. sapidus* STG. The amino acid sequence of each peptide is given above the spectrum. (a) MS/MS fragmentation of orcokinin NFDEIDRSGFGFA (*m/z* 1474.66). (b) MS/MS fragmentation of YRamide HIGSLYRa (*m/z* 844.48). The presence of b and y ions is indicated by horizontal lines above (y ions) or below (b ions) the corresponding amino acid residues in the peptide sequence. The fragment ions resulting from a lipid with close mass to the YRamide are labeled in blue.



Supplemental Figure S2. Collision-induced dissociation spectra of neuropeptides from the STG extract of *C. sapidus* on nanoLC-ESI-Q-TOF instrument. (a) MS/MS fragmentation of CabTRP 1a APSGFLGMRa (*m/z* 934.49). (b) MS/MS fragmentation of SIFamide GYRKPPFNGSIFa (*m/z* 1381.74). The presence of b and y ions is indicated by horizontal lines above (b ions) or below (y ions) the corresponding amino acid residues in the peptide sequence.



Supplemental Figure S3. MALDI TOF/TOF imaging of the *C. sapidus* STNS. Distribution of three neuropeptides are shown, including (a) YRamide HIGSLYRa (*m/z* 844.48) and two orcokinins: (b) NFDEIDRSSFa (*m/z* 1228.56) and (c) NFDEIDRSSFGFV (*m/z* 1532.70); and the MS image of a potassiated lipid PC(36:1) at *m/z* 826.57 is displayed in (d).



Chapter 6

Defining the Neuropeptidome of the Spiny Lobster *Panulirus interruptus* Brain Using a Multi-Dimensional Mass Spectrometry-Based Platform

Adapted from **Hui Ye**, Jingxin Wang, Zichuan Zhang, Chenxi Jia, Claire Schmerberg, Adam Catherman, Paul Martin Thomas, Neil Kelleher and Lingjun Li. Defining the Neuropeptidome of the spiny lobster *Panulirus interruptus* brain using a multi-dimensional mass spectrometry-based platform. *To be submitted*.

Abstract

Decapod crustaceans have long served as important animal models for neurobiologists and electrophysiologists, whereas characterization of their endogenous neuropeptides is challenging due to limited information about preprohormone sequences. Nevertheless, great conservation is observed among decapod species shown by our home-built decapod neuropeptide database. We thereby characterized the brain extract from *P. interruptus*, an important aquaculture decapod species with few known preprohormone sequences, with high mass accuracy and resolution data acquired on a LTQ-Orbitrap and conducted a ProSight database search using a flexible algorithm that tolerates sequence discrepancy from known sequence against our database. In addition to this streamlined semi-automated sequencing strategy, we characterized *P. interruptu*s by dimethylation-assisted fragmentation and manual *de novo* sequencing using nanoLC-ESI-Q-TOF, and direct tissue analysis on MALDI-TOF-TOF, further improving the coverage and resulting in an overall detection of 55 neuropeptides with 34 novel ones. The high discovery rate from this species demonstrated the usefulness of the neuropeptide discovery pipeline we developed and highlighted the advantage of utilizing multiple mass spectrometers with complementary capabilities. The localization patterns of brain neuropeptides were further elucidated with MALDI-MSI. Collectively, our study expands the catalog of crustacean neuropeptides present in *P. interruptus*, and more importantly, presents an approach that can be adapted to explore neuropeptidomes from other species that possess limited sequence information.

Abbreviations

MS	mass spectrometry
LC	liquid chromatography
ESI	electrospray ionization
MALDI	matrix-assisted laser desorption/ionization
Q-TOF	quadrupole-time-of-flight
TOF/TOF	time-of-flight/time-of-flight
MSI	mass spectrometry imaging
HRAM	high-resolution and accurate mass measurement
DAF	N-terminus dimethylation-assisted fragmentation

6.1 Introduction

The actions of neural circuits regulate and modify an organism's functions and behaviors, whereas these circuits are extensively modulated by cell-cell signaling molecules. Neuropeptides represent the largest and most diverse group of signaling molecules in nervous system and have shown to affect a wide range of physiological processes like hunger, pain, stress and reproduction.[1-3] To understand the underlying mechanisms of such actions, an initial step is to comprehensively characterize the peptide neuromodulators present in the system. Biologically active neuropeptides are produced by initial cleavage from large protein precursors followed by multiple subsequent enzymatic processing steps in various orders, resulting in a diverse array of physical and chemical properties.[4, 5] Therefore, significant challenges are present in the detection and identification of endogenous neuropeptides, although the capability to unveil novel neuropeptides could enable a more advanced understanding of peptidergic signaling in the nervous system.

Historically, Edman degradation has served as the predominant technique to sequence novel neuropeptides by sequential removal of amino acids from the N-terminus[6]. However, this method suffers from low throughput and large sample amount required for such analysis. In recent years, mass spectrometry (MS) has revolutionized the analysis of neuropeptides.[6, 7] MS allows for accurate measurements of the molecular weights of the peptide precursor ions and derivation of amino acid sequence via gas-phase fragmentation in tandem MS (MS/MS) mode. The coupling of liquid chromatography (LC) to MS/MS has permitted analysis of complex samples, including

neuronal tissue homogenate. Subsequently, the pattern of peptide fragment ions acquired in MS/MS spectra are matched to theoretical peptide spectra, which are predicted from preprohormone sequences arising from cDNA or a sequenced genome. Therefore, this workflow relies heavily on knowledge of the cleavage characteristics of the enzymes that process preprohormones to neuropeptides and, more crucially, the preprohormone sequences. To date, the workflow of LC-MS/MS analysis in conjunction with subsequent database search has become prevalant in neuropeptidomics, a thriving field that is devoted to the characterization of neuropeptides expressed in an organism, by precisely describing hundreds of diverse forms of neuropeptides from various species including rat,[8-11] mouse,[12-14] invertebrates including the house cricket *Acheta domesticus*,[15] the American cockroach *Periplaneta americana*,[16] the fruit fly *Drosophila melanogaster*,[17] the crabs *Cancer borealis*,[18] *Callinectes sapidus*[19] and *Ocypode ceratophthalma*,[20] and the nematodes *Caenorhabditis elegans*,[21] *Ascaris suum*,[22] among others.[6]

However, applying this database searching pipeline to identify neuropeptides from MS/MS spectra is not always straightforward when the species of interest have few or virtually no preprohormone sequences available. For example, decapod crustaceans are favored animal models for neurobiologists due to their relatively simple, well-characterized nervous system and intriguing behavior patterns mediated by various neuropeptides.[5, 23] Nevertheless, complete genome sequencing has not been attempted for these decapod crustacean species so far. Therefore, previously reported crustacean neuropeptide sequences were mostly obtained by interpreting MS/MS spectra acquired from crustacean neuronal tissue extract without known preprohormone sequences, which

is termed *de novo* sequencing.[24, 25] This process is more challenging than assigning an experimental spectrum to a predicted peptide, since a spectrum of rich fragment ions at high mass accuracy is a prerequisite to read off the sequence. Despite the these challenges, researchers in this field have successfully exploited LC-MS/MS method to identify hundreds of neuropeptides via *de novo* sequencing from Decapoda,[18-20, 24-30] and previously reported neuropeptides and cDNA sequences that encoded predicted or known crustacean neuropeptides (UniProt) have been compiled to assemble a decapod crustacean neuropeptide (CNP) database. Interestingly, a significant degree of NP sequence conservation among species is observed within the CNP database. A number of identical or similar motifs in neuropeptide families, such as tachykinin-related peptide (TRP), orcokinin and SIFamide, are shared among different crustacean species. We thereby developed a semi-automated *de novo* sequencing strategy by searching the LC-MS/MS data obtained from the spiny lobster *Panulirus interruptus*, a decapod crustacean species that has been seldomly characterized, against the compiled CNP database as a representative model here. In this study, we coupled nanoLC to a hybrid LTQ-Orbitrap via electrospray ionization (ESI) for the acquisition of LC-MS/MS data, and employed a database searching software designed for neuropeptide discovery, ProSight PC 2.0, to accommodate potential sequence variability and potential post-translational modifications (PTMs) present in *P. interruptus*. The combination of high resolution accurate mass (HRAM) MS data and rich peptide MS/MS fragments, in conjunction with the flexible scoring/searching software system that translates MS/MS spectra into peptide identifications results in highly confident identification of neuropeptides from the

complex *P. interruptus* brain homogenate with minimal manual intervention.

We further improved the comprehensive neuropeptide discovery workflow by employing other sample preparation techniques like N-terminal dimethylation-assisted fragmentation (DAF) and direct tissue analysis, and multiple instruments such as ESI-quadrupole-time-of-flight (Q-TOF) and matrix-assisted laser desorption/ionization (MALDI) time of flight (TOF)/TOF to increase the coverage of *P. interruptus* neuropeptidome. Using this workflow, we discovered 55 endogenous neuropeptides from the *P. interruptus* brain. Although 21 neuropeptides have been reported previously in the crustacean species *C. borealis* and *C. sapidus*, 34 novel neuropeptides were sequenced for the first time from *P. interruptus* with high confidence. This new information displays the diversity and conservation of crustacean neuropeptide sequences, and provides an important link to their potential biological functions. The inclusion of mass spectrometry imaging (MSI) to the workflow adds another dimension of information to our investigation of the *P. interruptus* brain neuropeptidome by providing spatial distribution and co-localization patterns of these endogenous signaling molecules. Overall, the multi-faceted workflow demonstrated in this study serves as a basis for successful discovery of endogenous neuropeptides from *P. interruptus*, as well as a universal approach that can be adapted to explore neuropeptidomes from crustacean and other species that have only limited sequence information.

6.2 Experimental Section

Chemical and Materials

Glacial acetic acid, borane pyridine, formaldehyde, 2,5-dihydroxybenzoic acid (DHB) and ammonium bicarbonate were obtained from Sigma-Aldrich (St. Louis, MO). Optima grade formic acid, acetonitrile (ACN), water, and methanol (MeOH) were purchased from Fisher Scientific (Pittsburgh, PA).

Animals and Dissection

California spiny lobsters were purchased from Catalina Offshore Products (San Diego, CA) and maintained in a circulating artificial seawater tank at 14–16 °C before use. Lobsters were anesthetized by packing them in ice for 30 minutes, after which the dorsal carapace was removed from each individual and its supraoesophageal ganglia (brain) were dissected free from surrounding muscle and connective tissues. The details of dissection were described elsewhere.[25] Following dissection, brain samples were immediately placed in acidified MeOH (90% MeOH: 9% glacial acetic acid: 1% deionized water) and stored at -80 °C until utilized for peptide extraction.

Tissue Extraction

Brain was homogenized and extracted with acidified MeOH buffer that was also used for tissue storage. The resulting extract was then concentrated in a Savant SC 110 Speedvac concentrator (Thermo Electron Corporation, West Palm Beach, FL, USA) and re-suspended in 0.1 % formic acid. The re-suspended extracts were then vortexed and briefly centrifuged. The resulting solution was purified and concentrated with C_{18} ZipTip (Millipore, Billerica, MA, USA). Briefly, the C_{18} ZipTip was first wetted using ACN and then pre-equilibrated for sample binding with 0.1 % formic acid in water. Subsequently, the tissue extract was loaded on the C18 ZipTip. After rinsing with 0.1 % formic acid in

water for three times, the sample was eluted with 5 μL ACN:water:formic acid solution (50:49.9:0.1; vol/vol/vol). Next, the eluent was dried and re-suspended in 10 μL of 0.1 % formic acid in water, and subjected to future LC-MS/MS and MALDI-MS analysis.

N-terminal Dimethylation via Formaldehyde labeling

For formaldehyde labeling, the neuropeptide extract was re-suspended in 5 μL of 0.1 % formic acid in water followed by the addition of 1.5 μL of borane pyridine (120 mM in MeOH) and 1.5 μL of formaldehyde (4% in H_2O). The labeling mixture was then incubated in 37°C water bath for 15 min. Excess formaldehyde was quenched via the addition of 3 μL of ammonium bicarbonate buffer (0.2 M). The resulting solution was dried and re-suspended in 10 μL of 0.1 % formic acid in water, and subjected to future LC-MS/MS analysis.

Tissue Preparation for MALDI-MS

Sample preparation for direct tissue analysis: Direct tissue analysis was performed as reported previously. Briefly, the accessory lobe (AcN) and olfactory lobe (ON) were designated as region 1 (R1) and region 2 (R2),[31] and dissected out followed by a brief rinse in acidified MeOH buffer and subsequent desalting in 10 mg/mL aqueous DHB solution. The processed tissues were then placed on the MALDI target with 0.3 μL of 100 mg/ml DHB matrix solution (50:49.9:0.1 MeOH:water:formic acid, v/v) deposited on top of the tissue, which allows extraction of neuropeptides from tissue prior to forming analyte-doped crystals.

Sample preparation for MSI

The freshly dissected brain was rinsed briefly in deionized water to eliminate excess

salt, embedded in an aqueous solution of gelatin (100 mg/ml) and then snap-frozen in -80° C freezer for further processing. Brain sections were acquired at a thickness of 12 µm on a cryostat (HM525, Thermo Fisher Scientific, Waltham, MA) at -20 °C and thaw-mounted onto indium tin oxide (ITO) coated conductive glass slides (Delta Technologies, Loveland, CO). The sections were then allowed to dry for 30 min under vacuum. For imaging purposes, 150 mg/ml DHB dissolved in 50% MeOH, 0.1% formic acid (v/v) was sprayed onto brain sections homogenously using an airbrush (Paasche Airbrush Company, Chicago, IL, USA). Five coats were applied and the spray duration for each coat was 30 s with 1 min dry time between each cycle.

On-line Top-down MS/MS on Nano-LC-ESI-LTQ-Orbitrap Elite and Data Analysis

The brain neuropeptide extract was further analyzed by on-line top-down MS on an Ultimate 3000 RSLCnano system coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). An aliquot of 1 µL was injected onto a 2 cm, 150 µm i.d. PLRP-S d_p 5 µm, pore size 1000Å) trap column. A 10 cm, 75 µm i.d. PLRP-S column was used for separation. The gradient was delivered at 300 nL/min starting at 95% A (0.2% formic acid in water), 5% B (95% acetonitrile and 0.2% formic acid) and rose to 10% B at 7 min, 50% B at 50 min, and 85% B at 58 min. The mass spectrometer was operated in a data-dependent mode, performing higher-energy collision dissociation (HCD)-MS^2 (scan 1), collision-induced dissociation (CID)-MS^2 (scan 2) and electron transfer dissociation (ETD)-MS^2 (scan 3) on each of the Top 3 precursors (selected by intact mass) in a FT-MS precursor scan. The specific conditions for the three fragmentation methods are as follows: HCD: Isolation width: 15 *m/z*, normalized

collision energy: 30, activation time: 0.1 ms; CID: Isolation width: 15 *m/z*, normalized collision energy: 41, activation q: 0.4, activation time: 100 ms; ETD: Isolation width: 15 *m/z*, normalized collision energy: 35, activation q: 0.25, activation time: 100 ms.

Data were deisotoped with Xtract using the cRAWler algorithm (ThermoFisher, Bremen, Germany) and searched with a custom 168-core ProSightPC 2.0 cluster using an iterative search tree. ProSightPC Warehouses were created against our home-built CNP database (https://uwmadison.box.com/lilabNP) with all the crustacean neuropeptides previously reported in literature and cDNA sequences filtered in UniProt for *crustaceans* (TaxId: 6657). The data was first searched in an "absolute mass" method, in which the mass tolerance for precursor ions was set at 200 Da and that for fragment ions was 10 ppm with Δ m mode on. A minimum match of 5 fragment ions was required. Then a "biomarker" search was performed, in which the mass tolerances for precursor ions and fragment ions are both set at 10 ppm with Δ m mode off. A minimum match of 4 fragment ions was required. The semi-automatic *de novo* sequencing was performed by manually reviewing all the peptide hits and modifying the sequences with possible PTMs or amino acids when necessary.

NanoLC-ESI-Q-TOF

A Waters nanoAcquity UPLC system was coupled to a Synapt G2 HDMS mass spectrometer (Waters Corp., Milford, MA) for LC-MS/MS analysis of the brain neuropeptide extract. Chromatographic separations were performed on a Waters BEH 130Å C18 reversed phase capillary column (150 mm X 75 μm, 1.7 μm). The mobile phases used were: 0.1% formic acid in deionized water (A); 0.1% formic acid in ACN

(B). An aliquot of 3 μL of brain neuropeptide extract was injected and loaded onto the Waters NanoACQ 2G-V/M SymC18 (180 um X 20mm, 5 μm) using 99% mobile phase A and 1% mobile phase B at a flow rate of 5 μL/min for 1 min. Following this, a linear gradient from 5 to 45% mobile phase B at a flow rate of 300 nL/min was performed over 90 minutes at 35°C. Data dependent acquisition (DDA) was employed with 3 precursors selected for MS/MS at once . The MS scan range was from *m/z* 300–2000 and the MS/MS scan was from *m/z* 50–2000. The MS/MS sequencing was performed with a combination of manual sequencing and PepSeq software (Waters Corp., Milford, MA) to assist in *de novo* sequencing.

CE fractionation for MALDI-MS analysis

The offline CE-MALDI-MS was performed with a home-built apparatus featuring a sheathless membrane-covered joint. Briefly, a 65 cm long fused-silica capillary (75 μm i.d./190 μm o.d.) was employed with a cellulose acetate membrane-coated porous joint made 3 cm from the outlet ends of the capillary.[32] This joint was inserted into a buffer cell filled with 0.5% ammonium acetate in pH 4.9, with the negative electrode connected. About 2 cm of capillary was stretched out of the buffer cell from a small hole with screw on the bottom. For the inlet end, a 0.6 mL plastic vial was filled with the same CE buffer with positive electrode and capillary inserted. A 10 kV high voltage was supplied by TriSep-2100 HV power supplier from Unimicro Technologies (Pleasanton, CA) for CE separation, while a high voltage of 9 kV was applied for sample loading. The CE flow was then collected on the MALDI plate every 1 min for 25 min.

MALDI-MS analysis

Direct tissue analysis and MSI of the *P. interruptus* brain together with the analysis of CE fractions were all performed on an autoflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA) equipped with a 200 Hz smartbeam. The following parameters were adopted in the positive reflectron mode for acquisition: ion source 1 voltage 19.00 kV, ion source 2 voltage 16.62 kV, reflector 1 voltage 20.90 kV, reflector 2 voltage 9.64 kV, and lens voltage 8.70 kV. The mass spectra data were acquired over a mass range of *m/z* 600–2000 where the majority of the neuropeptides were detected. For direct issue analysis, each spectrum was accumulated over 2000 laser shots for the profiled region. Additionally, the following parameters were adopted in the positive linear mode at a mass range of 3-10 kDa to profile the larger neuropeptides for direct tissue analysis: ion source 1 voltage 20.00 kV, ion source 3 voltage 18.55 kV, lens voltage 6.80 kV and pulsed ion extraction 130 ns. For MSI acquisition, each spectrum consisted of 500 laser shots acquired in positive reflectron mode and the array of spectra was collected at 50 μm intervals in both x and y dimensions across the surface of the brain section.

For analysis of CE fractions, each spectrum was summed from 1000 laser shots acquired in positive reflectron mode. Spectrum smoothing and baseline subtraction were performed to process the spectra using flexAnalysis software (Bruker Daltonics, Billerica, MA, USA), and the resulting peaks with S/N > 3 were matched with LC-MS/MS assignments within a mass error of 0.03 Da for neuropeptide identification from resulting spectra of direct tissue and CE fraction analysis. The imaging files were processed and MS images of neuropeptides were generated using the flexImaging

software package (Bruker Daltonics, Billerica, MA, USA).

6.3 Results and Disccusions

Comprehensive Neuropeptidomics Analysis of the Brain of Panulirus interruptus by a Multi-Dimensional MS-Based Platform

In this study, we employed three commonly used mass spectrometers, LC-ESI LTQ-Orbitrap, LC-ESI Q-TOF and MALDI TOF/TOF, for endogenous neuropeptide discovery. The results obtained from different mass spectrometers were compared and combined (**Table 1**) since each instrument offers overlapping yet distinct peptide coverage due to their specific ionization preferences and sequencing capabilities. For instance, although LC-ESI LTQ Orbitrap and LC-ESI Q-TOF share the same ionization platform, the former instrument offers higher ion transmission efficiency and more versatile MS/MS fragmentation techniques that enable a more comprehensive coverage of high-mass peptides and proteins.[25] Furthermore, spectra of high spectral resolution and mass accuracy contributed by a high-end instrument like LTQ-Orbitrap is a prerequisite for our semi-automated *de novo* sequencing strategy as a mass tolerance of 10 ppm was set when querying the observed *m/z* values of *P. interruptus* neuropeptide precursors and fragments against the CNP database in ProSight. This permitted precise determination of possible PTMs and amino acid substitutions, alone and in combination with others. This semi-automated *de novo* sequencing approach revolutionizes the *de novo* sequencing process for decapod crustacean species by dramatically decreasing the amount of time required for analysis without sacrificing identification confidence,

resulting in discovery of 34 neuropeptides. On the other hand, ESI-Q-TOF provides spectra of decent quality and excels at detection of low-mass immonium ions, delivering a complementary set of neuropeptides compared to LTQ-Orbitrap. In addition to the LC-ESI-MS-based platform, the characterization of *P. interruptus* neuropeptides was further facilitated by the MALDI-MS-based approach due to rapid analysis, good sensitivity and high tolerance to salt, providing the affirmed presence of 24 neuropeptides directly from the brain tissue. In total, we characterized and identified 55 neuropeptides, the majority of which were not reported in previous *P. interruptus* studies[33-36] and 34 of which are first sequenced in decapod crustacean species in this study. In **Table 1**, neuropeptides detected using different instruments are detailed, highlighting the combinatorial utilization of multiple powerful mass spectral techniques for the discovery of novel neuropeptides from the spiny lobster *P. interruptus* brain.

High Spectral Resolution and Mass Accuracy MS and MS/MS Sequencing

Neuropeptide extracts of *P. interruptus* brain were analyzed on a high spectral resolution and mass accuracy HRAM mass spectrometer, ESI-LTQ Orbitrap, with the top 3 precursors subjected to top-down HRAM MS/MS using a complementary suite of fragmentation techniques: Higher-energy C-trap dissociation (HCD), collision-induced dissociation (CID) and electron-transfer dissociation (ETD). The ultra-high resolution (120,000 in MS^1 and 60,000 in MS^2 at *m/z* 400) and accuracy (mass error lower than 5 ppm) measurements of precursor mass together with the rich sequence-specific fragment ions produced by the complementary fragmentation techniques were used to search the CNP database within ProSightPC 2.0.

Two types of searches were carried out in ProSight. The first type of search, "biomarker", was employed to find any peptides that were identical as they appeared in the CNP database. **Figure 1** shows the fragmentation spectra and assignment of neuropeptides derived from a known *P. interruptus* tachykinin related peptide (TRP) preprohormone included in the CNP database. **Figure 1a** displays a CID MS/MS spectrum arising from the *m/z* 1179.866 (4+) ion that can be matched with partial sequence of preproTRP, DAAAPLNEVDEASANDYPILPDPIA-ARLYAFRNGNAPVGLAVPLRa, by ProSight (**Figure 1b)**. The accurately measured and completely resolved precursor *m/z* resulted in a mass error of merely 0.85 ppm from the expected *m/z* at 1179.867 as shown in **Figure 1c**. Over 48 fragment ions including 15 b- and y- ion pairs were observed in the representative MS/MS spectrum that could be matched to theoretical fragments of the TRP precursor-related peptide (TPRP), confirming the identity of ions assigned by ProSight search. The combinatorial use of different fragmentation techniques is also highlighted in the study. A fragmentation map showing fragment ions acquired in HCD, CID and ETD MS/MS spectra is shown in **Figure 1d**, with the illustration of how these fragments were produced in **Figure 1e**. As expected, the most comprehensive coverage was obtained with the use of all three fragmentation techniques. The LTQ-Orbitrap employed in this study also features wide dynamic range and higher ion transmission efficiency of analytes, allowing for comprehensive coverage of neuropeptides spanning the mass range of detection. In comparison, none of the TPRPs or their truncated forms were observed on LC-ESI Q-TOF, highlighting the importance of utilizing different instrumentation platforms for

improved coverage.

Other than the capability to detect and fragment the relatively large neuropeptides as shown in **Figure 1a**, its capability to identify peptides of small size was demonstrated in **Figure 1f**. In **Figure 1f**, extensive fragmentation along the peptide backbone was observed with the use of HCD for the TRP APSGFLGMRa. Therefore, the presence of the TRP, previously proposed as a single, conserved neuropeptide across decapod crustacean species,[36] was demonstrated in *P. interruptus* brain. Indeed, no other TRP isoforms like TPSGFLGMRa—previously detected in the crab *Cancer borealis*—were seen in this study, consistent with the fact that seven identical copies of APSGFLGRMa are encoded in the *P. interruptus* preproTRP.[33] The *P. interruptus* TRP precursor undergoes several proteolytic processing steps, including cleavage at pro-hormone convertase processing sites, subsequent removal of the dibasic residues and the final amidation of the exposed Gly residue at the C-terminus, to the final peptide product.[33] Interference with this enzymatic processing cascade could possibly explain the detection of the truncated form of TRP, APSGFLGM, and the incompletely processed APSGFLGMRG, although the latter could alternatively have originated from the C-terminal mRNA coding sequence APSGFLGMR that is directly followed by a stop codon.[36] In addition, a methionine-oxidized form of TRP, APSGFLGM(O)Ra, was also identified, which could be an artifact of the sample preparation steps since no reducing agents have been added to the extract since the isolation of *P. interruptus* brain. In addition to variants of TRP, five truncated forms of the relatively large TPRP identified in **Figure 1a** were detected. The presence of multiple forms of the conserved decapod TRP

or TPRP are postulated to have important impacts on the physiological functions of the neuropeptides due to different affinity to their receptors.[36] All the *P. interruptus* preproTRP-derived peptides identified via ProSight "biomarker" search are underlined with different colors in the deduced TRP precursor sequence (accession number AB 113378 in GeneBank/EMBL/DDBJ) as shown in **Figure 1g**. More TPRP are predicted to exist than were identified here, which could be explained by their relatively low abundance, comparably poor ionization efficiency or insufficient chromatographic separation.[37]

Besides TPRP and TRP, two neuropeptides belonging to the AST-A family, ADPYAFGLa and PRNYAGFLa, and five RFamides (PSLRLRFa, PSMRLRFa, SMPSLRLRFa, APQRNFLRFa and QDLDHVFLRFa) were identified based on the HRAM MS together with the rich fragment ions delivered by the complementary fragmentation techniques using "biomarker" search.

Semi-Automated de novo Sequencing of Novel P. interruptus Neuropeptides

Although the ProSight "biomarker" search of high quality HRAM MS data against the known neuropeptides or sequenced preprohormones generates a number of neuropeptide hits, the second type of search, the "absolute mass" search, enhances novel neuropeptide discovery rate for *P. interruptus*. This mode consists of two searches that tolerate a 20 Da and 200 Da mass difference, respectively, between the detected mass of unknown peptides with the theoretical mass of peptides in CNP database, provided at least 5 fragment ions are matched within a mass error of 10 ppm. This flexible scoring algorithm thus retrieves known neuropeptides from the CNP database that displayed

similar fragmentation patterns as the unknown ions, accommodating the sequence variations that arise from amino acid variation, PTMs or truncation of the previously known peptide entries. After the "absolute mass" search displayed the matches, the results were taken to Sequence Gazer, a manual peptide/protein characterization tool, to investigate the site of variation including substitution, addition or deletion of amino acid sequence of the *P. interruptus* neuropeptide compared to the known one present in CNP database. This semi-automated *de novo* sequencing strategy took the advantage of great sequence conservation in decapod neuropeptide isoforms, significantly reducing the labor-intensive workloads of manual *de novo* sequencing and improving the confidence of identification because a cross-species comparison is also performed. In addition, peptides of large size, such as the TPRP in **Figure 1a,** that pose challenges for manual *de novo* sequencing could also be matched with peptides in CNP database as long as consensus sequences are identified within the query peptide.

A representative example is shown in **Figure 2**, where the full sequence of decapod calcitonin-like diuretic hormone (CLDH) was first characterized by MS without the knowledge of its cDNA sequence. **Figure 2a** displays a representative HCD MS/MS spectrum of the *P. interruptus* CLDH detected at the *m/z* of 986.1756 (z=3). The rich fragment ions produced from this ion and the flexible scoring algorithm employed in ProSight "absolute mass" search led to the resulting match with the American lobster *Homarus americanus* CLDH cDNA sequence of GLDLGLGRGFSGSQAAKHLMGLAAANFAGGPa. When compared with the observed fragmentation pattern, a series of b ions, from b_2 to b_{30} with the exception of

b_{17}, predicted from the *H. americanus* CLDH sequence were all matched to the dominant peaks in the MS/MS spectrum. Nevertheless, a peak with an *m/z* difference of 81.533 was observed following the b_{26} ion as shown in **Figure 2b**. The zoomed-in view in **Figure 2c** exhibits the ion's completely resolved isotopic pattern and its monoisotopic mass calculated as 2656.349 Da, delivering an incremental mass of 163.066 Da (corresponding to tyrosine) rather than 147.068 Da (corresponding to phenylalanine) following b_{26}. This Δm of 7.999 between the observed and predicted doubly charged b_{27} ion could thereby be explained by the substitution of amino acid residue phenylalanine (F) as encoded in *H. americanus* CLDH sequence by tyrosine (Y). Moreover, the consecutively assigned b_{28}, b_{29} and b_{30} ions as in **Figure 2a** all show an *m/z* difference of 7.993, 7.992 and 7.993 from the predicted values, validating the site of amino acid variation as the 27th residue. Therefore, the *P. interruptus* CLDH sequence was deduced as GLDLGLGRGFSGSQAAKHLMGLAAAN**Y**AGGPa, whereas the monoisotopic mass of its b_{27} ion is expected as 2656.349 Da, agreeing well with the observed mass (**Figure 2c**). This change could correspond to a single nucleotide change in the gene encoding CLDH, as tyrosine is encoded by UAU or UAC and phenylalanine is encoded by UUU or UUC, since such a vital hormone is unlikely to be changed drastically between these two related species. Our result presented here is the first observation of mature CLDH without prior knowledge of the corresponding cDNA sequence, demonstrating the conservation of neuropeptide sequences in decapods and the robustness of the HRAM MS/MS technique paired with the powerful semi-automated *de novo* sequencing strategy.

CLDH has been previously identified in insects including *Rhodnius prolixus*,[38]

Daploptera punctata,[39] *Drosophila melanogaster*[40] etc., and crustaceans such as *Balanus amphitrite*, *Daphnia pulex*[41] and *H. americanus*.[34] Functional studies of CLDH have shown that this family is implicated in the control of diuresis in insects. CLDH also intrinsically modulates the cardiac neuromuscular system in *H. americanus*, and thus has been proven to be bioactive in Decapoda.[34] Interestingly, a high degree of sequence similarity and a C-terminal amidation are observed in different isoforms of this 31 amino acid peptide hormone in insects and crustaceans, as illustrated in **Figure 2d**. The extensive sequence homology among different species supports the postulation that not only CLDHs but also their receptors are conserved, and might play critical roles throughout evolution.

Similarly, two ions were detected and matched with the preproneuropeptide F (NPF)-derived peptides (NPFP) predicted from *Penaeus vannamei* and *Melicertus marginatus* (UniProt: F6KM62 and F6KM63) EST transcripts, SDYPMPSGDALMEASERLLET and SDYPMPSGDALMEASERLLETA,[26] both with a Δm of 7.94 Da. Manual characterization was performed in Sequence Gazer on these two ions, deducing the actually detected peptides' sequences as SDYP**L**P**P**GDALMEASERLLET and SDYP**L**P**P**GDALMEASERLLETA, both with a substitution of M to L at the 5th position and S to P at the 7th position. Prior to our study, no authentic NPF isoforms or NPFP have been identified or validated from crustacean species using a MS-based platform. The NPF are viewed as the invertebrate homologs of the vertebrate peptide hormone neuropeptide Y (NPY), which has been associated with the regulation of appetite and feeding behavior.[42] It is noteworthy that the administration

of vertebrate NPY dramatically increased food intake in invertebrate *Marsupenaeus japonicus* as well,[43] suggesting a conserved role for NPY and possibly a degree of sequence homology in crustacean species, most likely in the part of the hormone that interacts with its receptor. Intriguingly, conservation of the NPF sequence in crustaceans is further evidenced by the sequence homology of the truncated NPF and NPFP between *P. interruptus* and two other species, *Penaeus vannamei* and *Melicertus marginatus*, that represent both a derived and a basal taxon. The physiological significance of this conservation awaits discovery. It will be interesting to investigate the impact of endogenous NPF molecules on food intake and growth in crustacean species, especially those of economical value to the aquaculture industry. In addition to this potentially important discovery, our work provides a framework for future investigations of NPF and its physiological functions in crustacean species in that it characterizes this important neuropeptide without prior knowledge of its sequence.

Besides the flexible searching algorithm to account for the unknown sequence variation not housed in the CNP database, our semi-automated *de novo* sequencing strategy also enables "multiplexed search," allowing for identification of multiple peptide precursors using multiplexed MS/MS spectra,[44] also known as chimeric spectra. A mass window of ± 7.5 Da was chosen to isolate peptide precursors subjected to MS/MS. This significantly wide isolation window improves sensitivity in hybrid LTQ Orbitrap without sacrificing identification confidence, since the ProSight multiplexing search option is designed to handle chimeric MS/MS spectra with the aid of HRAM MS/MS data. **Figure 3** highlights the identification of two novel neuropeptide isoforms of SIFamide

simultaneously from one multiplexed search in ProSight. In **Figure 3a**, two peptides of *m/z* 465.9091 and 466.5807 with overlapping isotopic distributions and another peptide at *m/z* 471.9124 were observed in a MS precursor scan. Appropriate assignments of their charge state and the resulting monoisotopic masses were automatically done by cRAWler. The resulting deconvoluted masses of 1396.722, 1394.709 and 1412.716 Da were then searched independently in ProSight using all the HRAM MS/MS fragments acquired in the single ETD MS/MS spectrum as shown in **Figure 3b**, all leading to a homologous match with a highly conserved decapod crustacean neuropeptide GYRKPPFNGSIFa.[30] With manual examination, the fragment ions of the peptide at 1396.722 Da were matched to c_6 to c_{11} of GYRKPPFNGSIFa. For the z ions, a consistent increase of 16.00 Da was observed for all of the assigned z ions, including z_6, z_9, z_{10} and z_{11}. This suggests an amino acid substitution of F to Y at the 12^{th} position, and thus the deduced sequence is GYRKPPFNGSI**Y**a as in **Figure 3c**. Similarly, the other SIFamide isoform of 1412.716 Da is suggested to have another Y at the 7^{th} position and sequenced as GYRKPP**Y**NGSI**Y**a due to the observation of a constant Δm 15.994 Da between several pairs of fragment ions from c_7 to c_{11} but an overlapping c_6 when compared to the fragment ions of GYRKPPFNGSIYa. In addition, the mass difference of 15.994 Da between all the observed z ions including z_6, z_9, z_{10} and z_{11} when compared to GYRKPPFNGSIYa further validates our deduced sequence based on c ions. No other sequences or fragment ions could possibly explain the peptide with a mass of 1394.709 Da, suggesting it as an intermediate with loss of water from GYRKPPYNGSIYa. This work therefore, demonstrates that multiplexing fragmentation and tailored software for

chimeric spectra interpretation allow for conclusive identification of two peptides from a single MS/MS isolation. The HRAM MS and MS/MS scans thereby allow one to confidently identify multiple peptides per DDA MS/MS spectra, increasing the absolute number of peptides identified without a compromise in duty cycle.

Intriguingly, the previously characterized neuropeptide GYRKPPFNGSIFa (Gly-SIFamide) was also identified in *P. interruptus*, agreeing well with its conserved presence in other decapod crustaceans including crabs and crayfish.[30] It is striking that the substitution of F to Y was observed in the two SIFamides as well as CLDH in *P. interruptus*. Future studies that isolate and sequence the cDNA encoding the SIFamides from this species could validate the amino acid substitution observed from *P. interruptus* neuropeptidome using a MS-based approach in this work. In addition to Gly-SIFamide, another SIFamide isoform, N-term acetylated (Ac)GYRKPPFNGSIFa, a truncated form RKPPFNGSIFa, and two incompletely processed forms GYRKPPFNGSIF and GRYKPPFNGSIFG were detected in *P. interruptus*. All the other raw spectra and fragmentation maps of identified neuropeptides are complied as **Appendix II**. Collectively, our strategy not only significantly expanded the catalog of SIFamide family peptides in decapod crustaceans to contain NPs with the C-terminal sequence motif SIYamide, but also laid a foundation for future functional studies of various *P. interruptus* neuropeptide isoforms in this important species.

Discovery of a Novel Motif of Orcokinin Family

Although the hybrid LTQ Orbitrap and the tailored database search algorithm enables semi-automated *de novo* sequencing of 37 neuropeptides, including 21 novel

ones, manual *de novo* sequencing of data acquired on SYNAPT G2 HDMS with the aid of proprietary software PepSeq also generates a list of 38 neuropeptides with 19 ones that have not been reported before. Taken together with the novel identifications achieved on LTQ Orbitrap, this work has revealed a novel sequence motif for the decapod crustacean orcokinin neuropeptide family. A myriad of orcokinins have previously been discovered from crustaceans, insects and other invertebrate species with a conserved N-terminal motif of NFDEIDR. More specifically, decapod orcokinins share a consensus sequence of NFDEISRSX(G/S)FGFX(H/N/V/A) (S-orcokinin) as in **Figure 4a,** with the exception of the Asp-Orcokinins observed in the blue crab *Callinectes sapidus*.[28] Interpretation of MS/MS spectra produced by orcokinin peptides relied heavily on the y ion series, due to this N-terminal consensus sequence. After sequential loss of the conserved residues from the N-terminus that can be viewed as fingerprints of the orcokinin family, information regarding the variable sequence appearing at the C-terminus is retained in the remaining y ions. Based upon accurate mass measurement and the characteristic fingerprints provided by y ions, we identified nine orcokinin peptides sharing a motif of NFDEIDRAGX(F/L/I/V)X(G/A/V)FX(H/N) (A-orcokinin) as summarized in **Figure 4b** (L is used in the figure for the purpose of alignment). In **Figure 4c**, the clean y_2-y_7 ion series made direct reading of the N-terminus sequence NFDEID from the MS/MS spectrum possible. The accurate mass measurement of the y_2 ion determines the C-terminus combination as RA, indicating a novel truncated form of orcokinin, NFDEIDRA. The fingerprints of unique fragmentation patterns belonging to orcokinin family members further facilitated our identification in the representative MS/MS spectra

shown in **Figure 4d-f**. In **Figure 4d**, the high-abundance y_5 to y_{10} ions immediately reveal its identity as orcokinin with an N-terminus sequence NFDEID, whereas the remaining y_1 to y_4 ions determined that the rest of C-terminus variable sequence is RAGL/IA, despite these ions being detected at relatively low abundances. Similarly, in **Figure 4e**, the y_4 to y_{10} ions again exhibit the fingerprint fragmentation pattern of orcokinins, delivering the conserved N-terminus sequence NFDEIDRA, whereas y_1 to y_3 ions aid in deducing the complete sequence as NFDEIDRAGL/IGF. A complete series of y ions with relatively high abundance starting from y_1 to y_{12} is also displayed in **Figure 4f**, supporting the identity of this peptide as another orcokinin variant with a sequence of NFDEIDRAGL/IGFH. It is worth noticing that the MS/MS fragmentation pattern of the orcokinin NFDEIDRAGL/IGFH appears to be quite distinct from the orcokinin with a single amino acid truncation at the C-terminus, NFDEIDRAGL/IGF. As shown in **Figure 4f**, the b ion series from b_4 to b_{11} were not detected in NFDEIDRAGL/IGFH, although a number of internal fragments were observed and assigned based on the deduced sequence. These internal fragments complicated the MS/MS spectrum, raising the challenge to interpret the fragment ions even with the assistance of the *de novo* sequencing software PepSeq.

Interestingly, a previous study by Stemmler et al. reported the approach of use metastable decay and sustained off-resonance irradiation collision-induced dissociation (SORI-CID) to identify orcokinin neuropeptides in crustaceans.[35] They identified the orcokinin isoform NFDEIDRAGLGF in *P. interruptus*, which is in agreement with our newly reported motif, and predicted the presence of more orcokinin variants that might

show distinct C-terminus variation from previously reported orcokinin peptides.[35] To the best of our knowledge, our study with multiple mass spectral techniques presented here describes the most comprehensive repertoire of crustacean orcokinin variants that contain this novel motif, NFDEIDRAG, including NFDEIDRAGLG, NFDEIDRAGLA, NFDEIDRAGVV, NFDEIDRAGFG, NFDEIDRAGFA, NFDEIDRAGLAFNa, a truncated form NFDEIDRAGLa and the three orcokinins shown in **Figure 4**. The specificity of this orcokinin motif to *P. interruptus* is noteworthy and a plausible explanation is a single codon changed on orcokinins for this species from UCU/C/A/G (corresponding to S) to GCU/C/A/G (corresponding to A). Simply changing the first nucleotide of the codon encoding the amino acid at that position from a U to a G could lead to substitution of A for S in the final sequence. This may suggest that A-orcokinins have another physiological function different from the S-orcokinins, i.e., that *P. interruptus* has developed another pathway in which they are involved in signaling. Alternatively, this could also mean that the A-orcokinins are simply redundant extra copies of the orcokinin gene with a small error that occured in DNA replication in an ancestor that has been propagated. In this case, the extra set of A-orcokinins might simply amplify the signal of the other orcokinins. This would need further exploration on whether the A-orcokinins have different receptor affinities and/or different receptors and/or downstream targets of the receptors. This is expected to generate interesting information regarding the impact of structural variance on neuropeptide biological functions in decapod crustaceans for people who study molecular evolution and speciation.

Dimethylation-Assisted Fragmentation (DAF) and de novo Sequencing

A reductive N-terminal dimethylation labeling method has previously been shown to facilitate the fragmentation and ultimately *de novo* sequencing of unknown neuropeptides.[27, 45] A schematic illustration of the reaction with formaldehyde and reducing reagent borane-pyridine complex is shown in **Figure 5a**. Briefly, the efficient and quick derivatization method labels a peptide's N-terminus and ε-amino groups of lysine, if any are present, through reductive dimethylation, resulting in an incremental mass of 28.03 Da for each derivatized site.[45] Moreover, substantial signal enhancement of the a_1 ion after labeling can be utilized to resolve sequence ambiguity in the N-terminus, since the b_1 ion is usually missing in a MS/MS spectrum.[27] In addition, the labeling method has also been employed to facilitate sequencing of singly charged peptides by increasing the signal intensities of the a- and b-ion series and reducing the complexity of fragmentation patterns.[27] We observed the same simplification effect in the interpretation of native orcokinin NFDEIDRAGL/IGFH with the approach of DAF. As previously shown in **Figure 4f**, a highly complex MS/MS spectrum was obtained for this peptide. Upon labeling, the mass of the native peptide at *m/z* 745.85 (z = +2) was shifted to *m/z* 759.86 (**Figure 5b**), corresponding to the mass difference induced by N-terminus dimethylation. Intriguingly, the formaldehyde labeled orcokinin displayed dramatically cleaner MS/MS spectrum, containing primarily b- and y-ion series, as shown in **Figure 5b**. Other than the complete y-ion series from y_1 to y_{12}, the b-ion series were significantly enhanced, resulting in detection of b_4, b_5, b_7, b_8, b_{10}, b_{11}, b_{12} ions, all of which were missing in the unlabeled sample. In addition to enhancing b ions, internal

fragments were suppressed, and thus a simplified fragmentation profile was obtained, enabling the identification of this novel orcokinin with remarkable confidence and ease.

Unambiguous identification of amino acid residues at the N-terminus of the sequence could be achieved via the DAF method as well. As an example, **Figure 5c** shows the MS/MS sequencing of an RFamide QDLDHVFLRFa, a peptide that has a close mass to KDLDHVFLRFa and isobaric mass with AG(or GA)DLDHVFLRFa. A 28.05 Da difference was detected upon labeling, as indicated by the mass shift from *m/z* 430.22 to 439.58 (z = +3) shown as the insert. This excludes the possible existence of lysine in the peptide, as it too would have been dimethylated for a total mass shift of 56.06 Da. Although the combined mass of two amino acid residues A+G is equal to that of Q (129.10 Da), no a_1 ion corresponding to G (58.07 Da) or A (72.08 Da) was seen even though the dimethylated a_1 ion should be significantly enhanced upon labeling. In contrast, an a_1 ion with *m/z* corresponding to dimethylated Q was detected with high abundance at *m/z* 129.10, supporting our assignment of this peptide as N-terminally labeled QDLDHVFLRFa.

N-terminal pyroglutamate and pyroaspartate modifications are among the most common PTMs in neuropeptides. The reductive dimethylation method aids in the characterization of neuropeptides with N-terminal blockage. Peptides with pyroglutamate modification, as shown in **Figure 5d**, can be therefore easily assigned based on their inability to react at the N-terminus due to the lack of free N-terminal amine group. After labeling, identical mass and similar retention times were observed compared to the native sample for the peptide at *m/z* 636.32 in the labeled sample, suggesting the

presence of an N-terminal blockage in this peptide. Together with the accurate mass measurement of b- and y-ion series, particularly a_1, the peptide was confidently assigned as pQDLDHVFLRFa. Other than pyro- modification, the application of reductive dimethylation also applies to N-terminal acetylation. A representative example is the novel peptide sequenced as (Ac)GYRKPPFNGSIFa with a mass of 1422.74 Da. However, it possesses a mass close to the previously reported VYRKPPFNGSIFa, which is found at 1422.78 Da. This ambiguity could be clarified by the observation that it is the only peptide with a single dimethylation at K rather than the expected dimethylation at both N-terminus and K, due to the N-terminal blockage. Nevertheless, it is worth noticing that this peptide was identified unambiguously without labeling relying on the HRAM data provided by LTQ-Orbitrap.

The observed enhancement and simplification in MS/MS fragmentation after reductive dimethylation could possibly be explained by the adjusted proton affinities (PA) of the fragments.[46] Upon labeling, the primary amine group at the N-terminus becomes a tertiary amine, possessing higher PA compared to the native form with primary amine. This helps to stabilize the b ions. Also, the N-terminal dimethylation prevents the formation of cyclic intermediate, therefore reducing sequence scrambling effect.[47] Therefore, the abundances of b-ions are generally improved in the resulting MS/MS spectra of labeled peptides compared to the unlabeled one. Moreover, internal fragments with a minimum of one b-type and one y-type bond cleavage are produced with greater frequency when samples are not labeled.[27] With the increased PA in the N-terminus, the fragmentation of y-type bonds is suppressed, resulting in a cleaner MS/MS fragmentation

pattern. Overall, the utilization of DAF for *de novo* sequencing not only provided validation for our assignments of *P. interruptus* neuropeptides based on their unmodified forms, but also resolved some sequence ambiguity at the N-terminus or in the middle of a peptide, yielding identification with improved confidence on a mass spectrometer with medium spectral resolution and mass accuracy.

To further improve our detection coverage and strengthen the confidence of our assignments, an offline CE-MALDI-MS platform was utilized to fractionate the brain extract to reduce signal suppression induced by the large number of peptides present with sample consumption of merely nanoliters. As shown in Supplemental Information **Figure S1**, the sample background was much cleaner upon separation as demonstrated by the representative fractions #8 and # 14 in **Figure S1a** and **1b** compared to the direct profiling of the extract in **Figure S1c**. All the peptides profiled via CE-MALDI-MS were indicted in Table 1, whereas those also detected by direct tissue analysis were noted as "d".

Region-Specific Localization of Neuropeptides in the Brain of P. interruptus

Using our multi-faceted MS-based strategy, we have discovered a number of novel neuropeptides from *P. interruptus* that have shared homology yet variations in amino acid residues or PTMs from previously reported neuropeptides or preprohormones. Examples of these changes include the amino acid substitution of S to A in *P. interruptus* orcokinins, and F to Y substitutions in *P. interruptus* SIFamide and CLDH. To provide further evidence that the origins of these sequence variation result from genetic differences among species rather than artifacts from sample preparation procedures,[25, 28]

we performed direct tissue analysis on *P. interruptus* brain to examine the resulting peptide profiles with minimal sample preparation. As highlighted in **Figure 6a**, 8 orcokinins were observed, including peptides that possess the common motif NFDEIDRSG such as NFDEIDRSGFG (*m/z* 1256.55), NFDEIDRSGFA (*m/z* 1270.57), NFDEIDRSGFGFNa (*m/z* 1516.68) and NFDEIDESGFGFH (*m/z* 1540.68), and the orcokinins containing the novel motif NFDEIDRAG such as NFDEIDRAGLG (*m/z* 1206.57), NFDEIDRAGFG (*m/z* 1240.56), NFDEIDRAGLGF (*m/z* 1353.64), and NFDEIDRAGLGFH (*m/z* 1490.70). Unfortunately, not all of the sequenced orcokinins were seen via direct tissue analysis on MALDI-TOF/TOF, probably due to their relatively low abundances and the complex sample matrix associated with the unprocessed tissue leading to analyte suppression. The complexity of this matrix is proven by the myriad of lipid peaks shown in **Figure 6a**. Nevertheless, the observation of these novel orcokinins directly from tissue validated their existence in the *P. interruptus* brain and thereby the novel motif we discovered. In addition, the substitution of F to Y exemplified by CLDH at *m/z* 2956.50 was also detected via direct tissue analysis of the *P. interruptus* brain in linear mode, excluding the possibility of the sequence variation due to sample processing procedures.

Using knowledge gained from pioneering studies of the brains of spiny lobsters dating back to the 1960's,[48] substructures of the *P. interruptus* brain were examined in detail using direct tissue profiling. Crustacean brains consist of more complex structures compared to other neuronal organs in crustacean species, including numerous areas of neuropil and neuronal clusters[31]. Previously, neuropeptides that displayed distinguishable

distributions have been linked to different biological functions in crustacean species like *C. sapidus*[49] and *C. borealis*.[24, 50] Therefore, it is crucial to study the neuropeptides' localization in *P. interruptus* brain due to the correlations that can often be made between brain location and function. One interesting finding is that neuropeptides desorbed/ionized from region 1 (R1) in **Figure 6a** exhibited overlapping yet distinct peptide profiles compared to those from region 2 (R2) in **Figure 6b**. In **Figure 6a**, orcokinin peptides were detected at high abundances, whereas the SIFamide peptide at *m/z* 1381.74 was observed at a modest level when compared to the base peak APSGFLGMRa at *m/z* 934.49 in R1. In contrast, SIFamide rose to a dominant peak in the profiling spectrum of R2, whereas orcokinins were detected at a low level in this region compared to R1. The differential neuropeptide patterns from these two regions' profiling spectra demonstrate the complexity of the *P. interruptus* brain structure and the potentially distinct regulatory roles neuropeptides play in different regions of the brain.

Mapping Distribution of Various Brain Neuropeptides via MALDI-MSI

Direct tissue analysis yields "snapshots" of neuropeptide profiles corresponding to targeted regions in *P. interruptus* brain, providing information about spatial distribution of neuropeptides in a high throughput manner. Nonetheless, the need to document detailed and accurate localizations of neuropeptides in a complex and intricate structure like brain cannot be satisfied with direct tissue analysis. Alternatively, MSI, a newly emerged technique, has demonstrated its capability to characterize and localize neuropeptides from neural tissue of heterogeneous structure at high spatial resolution.[6, 19, 51-53] The high-resolution mapping of *P. interruptus* brain neuropeptides is thus

accomplished here via MSI with representative neuropeptides from major families, tachykinin, SIFamide, orcokinin, AST-A and RFamide, shown in **Figure 7**. In **Figure 7a**, the TRP neuropeptide of *m/z* 934.49, detected at high abundance in both regions in **Figure 6,** is shown to be distributed throughout the major regions in brain, including anterior medial protocerebral neuropils (AMPN), accessory lobes (AcN), and olfactory lobes (ON) as illustrated by the rostral view of *Panulirus* brain (**Figure 7**).[31] Its distribution in AcN and ON agrees well with our previous knowledge that decapod TRP has been identified as a neuromodulator in the olfactory neural pathway.[54] Interestingly, two SIFamide peptides, GYRKPPFNGSIFa at *m/z* 1381.74 (**Figure 7b**) and (Ac)GYRKPPFNGSIFa at *m/z* 1423.75 (**Figure 7c**), show concentrated localization in the AcN, which agrees well with its significantly higher abundance in R2 compared to R1 as displayed in **Figure 6**. In contrast, orcokinin peptides at *m/z* 1254.47 and 1490.70, also revealed by regional profiling in **Figure 6**, exhibited higher levels of abundance in the ON compared to AcN and AMPN in **Figure 7d-e**. The co-localization of various neuropeptide isoforms from the same family, particularly in the AcN and ON, and potential links to function warrant further investigation. Moreover, the AST-A peptide ADPYAFGLa at *m/z* 852.43 also displays a higher concentration in ON, illustrated in **Figure 7f**. In addition, the RFamide SMPSLRLRFa at *m/z* 1105.63 localized mostly in AcN and ON, yet present in lower quantities in the AMPN in **Figure 7g.** This observation is also in agreement with its presence in both R1 and R2 profiling spectra.

The MALDI-MSI results successfully mapped neuropeptides of different families in the brain, presenting useful and interesting spatial information about endogenous

neuropeptides and providing insights into their regulatory functions in the brain of *P. interruptus*. OL and AcN are dominant deutocerebral neuropils (DCN), and are best visualized in the rostral view. A unique feature of *P. interruptus* brain is that the large AcNs are located medial to the slightly smaller ONs, an organization that is distinct from that previously characterized in the lobster *Homarus americanus*[25] and other decapod species like crayfish or crabs.[31] OL receives primary olfactory input from olfactory receptor neurons on antennae, whereas AcN not only receives secondary olfactory signals from the OL but also higher multimodal signals through synaptic contact with neurons derived from tactile and visual sensory systems.[54] Therefore, the distributions of neuropeptides within DCN suggest their possible role as neuromodulators involved in the function of the olfactory system, the tactile system, and/or the visual sensory system. An alternative function could be in the integration of multiple sensory modalities. Moreover, the information about colocalization of neuropeptides gained in this study implies a relationship between their functions—those that are found in close proximity to each other may have similar or related roles in cell-cell signaling. Further in-depth study might provide a clear answer regarding the possibly synergistic, complementary, or antagonistic relationships that could be present among these co-localized peptide modulators.

To the best of our knowledge, our study presents the first investigation of the spatial distribution of neuropeptides in the brain of *P. interruptus*. Although MSI generally has a limitation in resolution compared to conventional immunohistochemical methods, the capability to distinguish neuropeptide isoforms, exemplified by SIFamide peptides in

Figure 7b-c and orcokinin peptides in **Figure 7d-e**, in a high throughput manner enables unambiguous mapping of neuropeptides in large-scale. In addition, the overlay feature enables highly multiplexed simultaneous study of co-localization patterns of numerous neuropeptides (**Figure 7h**) and thus may provide information on potential relationship of their functions.

6.4 Conclusions

In this study, we employed a suite of mass spectrometric approaches for the discovery of neuropeptides in the brain of *P. interruptus*, an important aquaculture species that has not been extensively characterized. Collectively, 55 neuropeptides were sequenced in this work, including 34 novel ones that are *de novo* sequenced. The use of HRAM MS/MS data with various fragmentation methods (CID, HCD and ETD) and tailored ProSight search against our home-built CNP database highlighted the possibility of streamlining the peptide discovery process with highly confidnet identifications. Although the hybrid LTQ-Orbitrap excels in the detection of larger peptides, Q-TOF provides complementary coverage in interpretation of medium-sized neuropeptides. Furthermore, the application of DAF aids in the *de novo* sequencing process and improves identification confidence of neuropeptides on a medium-resolution instrument. In addition, direct tissue analysis was applied to analyze the brain tissue with minimal sample preparation, better preserving native neuropeptide expression profiles in *P. interruptus* brain. Finally, MALDI-MSI was employed to map the distribution of multiple neuropeptides simultaneously from a brain slice to yield information on

localization that may be important in determining the function of these NPs. In summary, our study not only presents comprehensive characterization of neuropeptide expression and distribution in the brain of *P. interruptus* providing insights into physiological functions of these endogenous neuropeptides, but also demonstrates the application of a multi-dimensional MS-based platform that will be useful for future neuropeptidome discovery studies on crustacean species without the prerequisite of known preprohormones.

Acknowledgement

This work is supported in part by the National Institutes of Health grants (R01DK071801 and 1S10RR029531) and the National Science Foundation grant (CHE-0957784).

6.5 References

1. Morton, G. J.; Cummings, D. E.; Baskin, D. G.; Barsh, G. S.; Schwartz, M. W. (2006) Central nervous system control of food intake and body weight. *Nature* 443, 289-95.
2. Hokfelt, T.; Bartfai, T.; Bloom, F. (2003) Neuropeptides: opportunities for drug discovery. *Lancet Neurol.* 2, 463-72.
3. Solway, B.; Bose, S. C.; Corder, G.; Donahue, R. R.; Taylor, B. K. (2011) Tonic inhibition of chronic pain by neuropeptide Y. *Proc. Natl. Acad. Sci. U. S. A.* 108, 7224-9.
4. Devi, L. (1991) Consensus Sequence for Processing of Peptide Precursors at Monobasic Sites. *FEBS Lett.* 280, 189-194.
5. Li, L.; Sweedler, J. V. (2008) Peptides in the brain: mass spectrometry-based measurement approaches and challenges. *Annu. Rev. Anal. Chem.* 1, 451-83.
6. Ye, H.; Greer, T.; Li, L. J. (2012) Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *J. Proteomics* 75, 5014-5026.
7. Hummon, A. B.; Amare, A.; Sweedler, J. V. (2006) Discovering new invertebrate neuropeptides using mass spectrometry. *Mass Spectrom. Rev.* 25, 77-98.
8. Karlsson, O.; Kultima, K.; Wadensten, H.; Nilsson, A.; Roman, E.; Andrén, P. E.; Brittebo, E. B. (2013) Neurotoxin-Induced Neuropeptide Perturbations in Striatum of Neonatal Rats. *J Proteome Res*, Epub ahead of print.
9. Romanova, E. V.; Lee, J. E.; Kelleher, N. L.; Sweedler, J. V.; Gulley, J. M. (2012) Comparative peptidomics analysis of neural adaptations in rats repeatedly exposed to amphetamine. *J. Neurochem.* 123, 276-287.
10. Mabrouk, O. S.; Kennedy, R. T. (2012) Simultaneous oxytocin and arg-vasopressin measurements in microdialysates using capillary liquid chromatography-mass spectrometry. *J. Neurosci. Methods* 209, 127-133.
11. Ljungdahl, A.; Hanrieder, J.; Falth, M.; Bergquist, J.; Andersson, M. (2011) Imaging mass spectrometry reveals elevated nigral levels of dynorphin neuropeptides in L-DOPA-induced dyskinesia in rat model of Parkinson's disease. *PLoS One* 6, e25653.
12. Guenther, S.; Römpp, A.; Kummer, W.; Spengler, B. (2011) AP-MALDI imaging of neuropeptides in mouse pituitary gland with 5 µm spatial resolution and high mass accuracy. *Int. J. Mass Spectrom.* 305, 228-237.
13. Wardman, J. H.; Berezniuk, I.; Di, S.; Tasker, J. G.; Fricker, L. D. (2011) ProSAAS-derived peptides are colocalized with neuropeptide Y and function as neuropeptides in the regulation of food Intake. *PLoS One* 6, e28152.
14. Zhang, X. Z.; Petruzziello, F.; Zani, F.; Fouillen, L.; Andrén, P. E.; Solinas, G.; Rainer, G. (2012) High Identification Rates of Endogenous Neuropeptides from Mouse Brain. *J. Proteome Res.* 11, 2819-2827.
15. Verhaert, P. D.; Conaway, M. C. P.; Pekar, T. M.; Miller, K. (2007) Neuropeptide imaging on an LTQ with vMALDI source: The complete 'all-in-one' peptidome analysis. *Int. J. Mass Spectrom.* 260, 177-184.
16. Verhaert, P. D.; Pinkse, M. W.; Strupat, K.; Conaway, M. C. (2010) Imaging of similar mass neuropeptides in neuronal tissue by enhanced resolution MALDI MS with an ion trap - Orbitrap hybrid instrument. *Methods. Mol. Biol.* 656, 433-49.
17. Wegener, C.; Neupert, S.; Predel, R. (2006) Direct MALDI-TOF mass spectrometric peptide profiling of neuroendocrine tissue of Drosophila. *Methods Mol. Biol.* 615, 117-27.
18. DeKeyser, S. S.; Kutz-Naber, K. K.; Schmidt, J. J.; Barrett-Wilt, G. A.; Li, L. J. (2007) Imaging mass spectrometry of neuropeptides in decapod crustacean neuronal tissues. *J. Proteome Res.* 6, 1782-1791.

19. Ye, H.; Hui, L. M.; Kellersberger, K.; Li, L. J. (2013) Mapping of neuropeptides in the crustacean stomatogastric nervous system by imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.* 24, 134-147.
20. Hui, L.; D'Andrea, B. T.; Jia, C.; Liang, Z.; Christie, A. E.; Li, L. (2013) Mass spectrometric characterization of the neuropeptidome of the ghost crab *Ocypode ceratophthalma* (Brachyura, Ocypodidae). *Gen. Comp. Endocrinol.* 184, 22-34.
21. Husson, S. J.; Clynen, E.; Boonen, K.; Janssen, T.; Lindemans, M.; Baggerman, G.; Schoofs, L. (2010) Approaches to identify endogenous peptides in the soil nematode *Caenorhabditis elegans*. *Methods Mol. Biol.* 615, 29-47.
22. Jarecki, J. L.; Frey, B. L.; Smith, L. M.; Stretton, A. O. (2011) Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J. Proteome Res.* 10, 3098-3106.
23. Christie, A. E.; Stemmler, E. A.; Dickinson, P. S. (2010) Crustacean neuropeptides. *Cell. Mol. Life Sci.* 67, 4135-4169.
24. Cape, S. S.; Rehm, K. J.; Ma, M.; Marder, E.; Li, L. J. (2008) Mass spectral comparison of the neuropeptide complement of the stomatogastric ganglion and brain in the adult and embryonic lobster, *Homarus americanus*. *J. Neurochem.* 105, 690-702.
25. Chen, R. B.; Jiang, X. Y.; Conaway, M. C. P.; Mohtashemi, I.; Hui, L. M.; Viner, R.; Li, L. J. (2010) Mass spectral analysis of neuropeptide expression and distribution in the nervous system of the lobster *Homarus americanus*. *J. Proteome Res.* 9, 818-832.
26. Christie, A. E.; Cashman, C. R.; Brennan, H. R.; Ma, M. M.; Sousa, G. L.; Li, L. J.; Stemmler, E. A.; Dickinson, P. S. (2008) Identification of putative crustacean neuropeptides using in silico analyses of publicly accessible expressed sequence tags. *Gen. Comp. Endocrinol.* 156, 246-264.
27. Fu, Q.; Li, L. J. (2005) *De novo* sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal. Chem.* 77, 7783-7795.
28. Hui, L.; Cunningham, R.; Zhang, Z.; Cao, W.; Jia, C.; Li, L. (2011) Discovery and characterization of the crustacean hyperglycemic hormone precursor related peptides (CPRP) and orcokinin neuropeptides in the sinus glands of the blue crab *Callinectes sapidus* using multiple tandem mass spectrometry techniques. *J. Proteome Res.* 10, 4219-29.
29. Li, L. J.; Pulver, S. R.; Kelley, W. P.; Thirumalai, V.; Sweedler, J. V.; Marder, E. (2002) Orcokinin peptides in developing and adult crustacean stomatogastric nervous systems and pericardial organs. *J. Comp. Neurol.* 444, 227-244.
30. Vázquez-Acevedo, N.; Rivera, N. M.; Torres-Gonzalez, A. M.; Rullan-Matheu, Y.; Ruíz-Rodríguez, E. A.; Sosa, M. A. (2009) GYRKPPFNGSIFamide (Gly-SIFamide) modulates aggression in the freshwater prawn Macrobrachium rosenbergii. *Biol. Bull.* 217, 313-26.
31. Sandeman, D.; Sandeman, R.; Derby, C.; Schmidt, M. (1992) Morphology of the brain of crayfish, crabs, and spiny lobsters: a common nomenclature for homologous structures. *Biol. Bull.* 183, 304-326.
32. Zhang, Z. C.; Ye, H.; Wang, J. H.; Hui, L. M.; Li, L. J. (2012) Pressure-assisted capillary electrophoresis coupling with matrix-assisted laser desorption/ionization-mass spectrometric imaging for quantitative analysis of complex peptide mixtures. *Anal. Chem.* 84, 7684-7691.
33. Christie, A. E.; Cashman, C. R.; Stevens, J. S.; Smith, C. M.; Beale, K. M.; Stemmler, E. A.; Greenwood, S. J.; Towle, D. W.; Dickinson, P. S. (2008) Identification and cardiotropic actions of brain/gut-derived tachykinin-related peptides (TRPs) from the American lobster *Homarus americanus*. *Peptides* 29, 1909-18.

34. Christie, A. E.; Stevens, J. S.; Bowers, M. R.; Chapline, M. C.; Jensen, D. A.; Schegg, K. M.; Goldwaser, J.; Kwiatkowski, M. A.; Pleasant, T. K.; Shoenfeld, L.; Tempest, L. K.; Williams, C. R.; Wiwatpanit, T.; Smith, C. M.; Beale, K. M.; Towle, D. W.; Schooley, D. A.; Dickinson, P. S. (2010) Identification of a calcitonin-like diuretic hormone that functions as an intrinsic modulator of the American lobster, *Homarus americanus*, cardiac neuromuscular system. *J. Exp. Biol.* 213, 118-127.
35. Stemmler, E. A.; Provencher, H. L.; Guiney, M. E.; Gardner, N. P.; Dickinson, P. S. (2005) Matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for the identification of orcokinin neuropeptides in crustaceans using metastable decay and sustained off-resonance irradiation. *Anal. Chem.* 77, 3594-3606.
36. Yasuda-Kamatani, Y.; Yasuda, A. (2004) APSGFLGMRamide is a unique tachykinin-related peptide in crustaceans. *Eur. J. Biochem.* 271, 1546-56.
37. Gelman, J. S.; Dasgupta, S.; Berezniuk, I.; Fricker, L. D. (2013) Analysis of peptides secreted from cultured mouse brain tissue. *Biochim. Biophys. Acta*, doi: 10.1016/j.bbapap.2013.01.043.
38. Brugge, V. A.; Schooley, D. A.; Orchard, I. (2008) Amino acid sequence and biological activity of a calcitonin-like diuretic hormone (DH31) from *Rhodnius prolixus*. *J. Exp. Biol.* 211, 382-90.
39. Furuya, K.; Milchak, R. J.; Schegg, K. M.; Zhang, J.; Tobe, S. S.; Coast, G. M.; Schooley, D. A. (2000) Cockroach diuretic hormones: characterization of a calcitonin-like peptide in insects. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6469-74.
40. Coast, G. M.; Webster, S. G.; Schegg, K. M.; Tobe, S. S.; Schooley, D. A. (2001) The *Drosophila melanogaster* homologue of an insect calcitonin-like diuretic peptide stimulates V-ATPase activity in fruit fly *Malpighian tubules*. *J. Exp. Biol.* 204, 1795-804.
41. Gard, A. L.; Lenz, P. H.; Shaw, J. R.; Christie, A. E. (2009) Identification of putative peptide paracrines/hormones in the water flea *Daphnia pulex* (Crustacea; Branchiopoda; Cladocera) using transcriptomics and immunohistochemistry. *Gen. Comp. Endocrinol.* 160, 271-87.
42. Xi, L.; Jin, Y.; Parker, E. A.; Josh, P.; Jones, A.; Wijffels, G.; Colgrave, M. L. (2012) Challenges in mass spectrometry-based quantification of bioactive peptides: a case study exploring the neuropeptide Y family. *Biopolymers* 98, 357-66.
43. Kiris, I. G.; Eroldogan, O. T.; Kir, M.; Kumlu, M. (2004) Influence of neuropeptide Y (NPY) on food intake and growth of penaeid shrimps *Marsupenaeus japonicus* and *Penaeus semisulcatus* (Decapoda: Penaeidae). *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 139, 239-44.
44. Boyne, M. T.; Garcia, B. A.; Li, M.; Zamdborg, L.; Wenger, C. D.; Babai, S.; Kelleher, N. L. (2009) Tandem mass spectrometry with ultrahigh mass accuracy clarifies peptide identification by database retrieval. *J. Proteome Res.* 8, 374-9.
45. Hsu, J. L.; Huang, S. Y.; Shiea, J. T.; Huang, W. Y.; Chen, S. H. (2005) Beyond quantitative proteomics: signal enhancement of the a1 ion as a mass tag for peptide sequencing using dimethyl labeling. *J. Proteome Res.* 4, 101-8.
46. Paizs, B.; Suhai, S. (2005) Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* 24, 508-48.
47. Jia, C.; Qi, W.; He, Z. (2007) Cyclization reaction of peptide fragment ions during multistage collisionally activated decomposition: an inducement to lose internal amino-acid residues. *J. Am. Soc. Mass Spectrom.* 18, 663-78.
48. Maynard, D. M. (1966) Integration in crustacean ganglia. *Symp. Soc. Exp. Biol.* 20, 111-49.

49. Hui, L. M.; Zhang, Y. Z.; Wang, J. H.; Cook, A.; Ye, H.; Nusbaum, M. P.; Li, L. J. (2011) Discovery and functional study of a novel crustacean tachykinin neuropeptide. *ACS Chem. Neurosci.* 2, 711-722.
50. Chen, R. B.; Hui, L. M.; Sturm, R. M.; Li, L. J. (2009) Three dimensional mapping of neuropeptides and lipids in crustacean brain by mass spectral imaging. *J. Am. Soc. Mass Spectrom.* 20, 1068-1077.
51. Hanrieder, J.; Ljungdahl, A.; Falth, M.; Mammo, S. E.; Bergquist, J.; Andersson, M. (2011) L-DOPA-induced dyskinesia is associated with regional increase of striatal dynorphin peptides as elucidated by imaging mass spectrometry. *Mol. Cell. Proteomics* 10.
52. Römpp, A.; Guenther, S.; Schober, Y.; Schulz, O.; Takats, Z.; Kummer, W.; Spengler, B. (2010) Histology by mass spectrometry: label-free tissue characterization obtained from high-accuracy bioanalytical imaging. *Angew. Chem. Int. Edi.* 49, 3834-3838.
53. Zimmerman, T. A.; Rubakhin, S. S.; Sweedler, J. V. (2011) MALDI mass spectrometry imaging of neuronal cell cultures. *J. Am. Soc.Mass Spectrom.* 22, 828-836.
54. Yasuda-Kamatani, Y.; Yasuda, A. (2006) Characteristic expression patterns of allatostatin-like peptide, FMRFamide-related peptide, orcokinin, tachykinin-related peptide, and SIFamide in the olfactory system of crayfish *Procambarus clarkii*. *J. Comp. Neurol.* 496, 135-47.

Table 1. Neuropeptides detected in the brain of the California spiny lobster *Panulirus interruptus*.

Family	$[M+H]^+$	Sequence	Q/TOF	Orbitrap	TOF/TOF
AST-A	852.43	ADPYAFGLa	+		d
	936.51	PRNYAFGLa	+	+	+
RFa	887.56	PSLRLRFa	+	+	
	905.51	PSMRLRFa		+	
	1104.61	GAHKNYLRF	+		
	1105.63	SMPSLRLRFa	+	+	+; d
	1147.65	APQRNFLRFa	+	+	
	1271.65	pQDLDHVFLRFa	+		d
	1288.68	QDLDHVFLRFa	+		
Orcomytropin	976.44	FDAFTTGFA		+	
	1186.52	FDAFTTGFGHS	+	+	
Orcokinin	979.45	NFDEIDRA	+		
	1148.57	NFDEIDRAGLa		+	+
	1206.57	NFDEIDRAGLG		+	+; d
	1220.59	NFDEIDRAGLA	+		+; d
	1234.57	NFDEIDRAGVV		+	+
	1240.56	NFDEIDRAGFG	+		+; d
	1254.57	NFDEIDRAGFA	+	+	+; d
	1256.55	NFDEIDRSGFG		+	+; d
	1270.57	NFDEIDRSGFA	+		+; d
	1353.64	NFDEIDRAGLGF	+		+; d
	1490.70	NFDEIDRAGLGFH	+	+	+; d
	1516.68	NFDEIDRSGFGFNa	+		+; d
	1540.68	NFDEIDRSGFGFH	+		+; d
SIFa	1161.65	RKPPFNGSIFa	+		+
	1381.74	GYRKPPFNGSIFa	+	+	+; d
	1382.72	GYRKPPFNGSIF	+		

241	1395.75	GYRKPPYNGSIYa(-H_2O)		+	
	1397.78	GYRKPPFNGSIYa	+	+	
	1413.74	GYRKPPYNGSIYa		+	
	1423.75	(Ac)GYRKPPFNGSIFa	+	+	+; d
	1431.72	GYRKPPFNGSYFa	+		
	1439.74	GYRKPPFNGSIFG	+	+	+; d
PDH	887.52	NSLLGISAL	+		
	1012.62	ELAAQILRV		+	
	1111.56	KYQAREM(O)VA	+		+; d
	1133.56	AREM(O)VAELAQ	+		
	1927.03	NSELINSILGLPKVMNDAa	+	+	
	1928.02	NSELINSILGLPKVMNDA	+		
	1890.01	QQIYRVAQAPQAGAVGPHa	+		
Tachykinin	605.30	APSGFQa	+		
	779.38	APSGFLGM	+		
	934.49	APSGFLGMRa	+	+	+; d
	950.49	APSGFLGM(O)Ra	+		+; d
	992.50	APSGFLGMRG	+	+	
	1650.96	AFRNGNAPVGLAVPLRa		+	+; d
	1651.94	AFRNGNAPVGLAVPLR			+
	1814.02	YAFRNGNAPVGLAVPLRa		+	
	3084.52	DAAAPLNEVDEASANDYPILPDPIAARLY		+	
	3155.53	DAAAPLNEVDEASANDYPILPDPIAARLYA		+	
	4716.45	DAAAPLNEVDEASANDYPILPDPIAARLYAFRNGNAPVGLAVPLRa		+	d
Neuropeptide F	2304.11	SDYPLPPGDALMEASERLLET		+	
	2488.23	SDYPLPPGDALMEASERLLETLA		+	
CLDH	2956.52	GLDLGLGRGFSGSQAAKHLMGLAAANYAGGPa	+	+	+; d

Figures

Figure 1. Sequencing of *P. interruptus* TRP and TPRP with the aid of ProSight. (a) A HRAM MS/MS spectrum of *P. interruptus* TPRP acquired in HCD. (b) Fragment ion assignment of the TPRP MS/MS spectrum shown in (a). (c) Zoomed-in raw MS spectrum of the precursor TPRP that shows the HRAM *m/z* information. (d) Sequence coverage of the TPRP, combining fragments observed in HCD, CID and ETD spectra. (e) An illustration of how the fragment ions of HCD and CID (b, y ions) and ETD (c,z ions) are produced and annotated in the fragmentation map in (b) and (f). (f) An HCD spectrum of *P. interruptus* TRP and its assignment of fragmentation. (g) Previously reported *P. interruptus* preprotachykinin cDNA sequence with detected peptides highlighted in different colors.

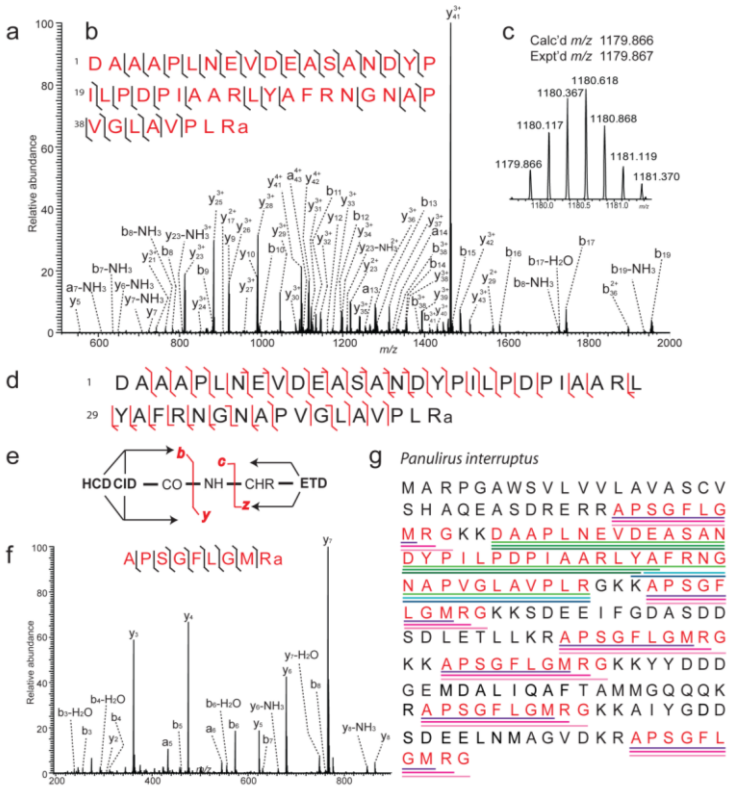


Figure 2. *De novo* sequencing of *P. interruptus* CLDH. (a) An HRAM CID MS/MS spectrum of *P. interruptus* CLDH observed at *m/z* 986.1756. Its fragmentation map and sequence coverage are shown in the insert. (b) The accurate assignment of amino acid substitution at position 27 in CLDH by registering the mass discrepancy observed in ProSight to b_{27} ion in the raw MS/MS spectrum. (c) A zoomed-in view of the well-resolved b_{27} ion detected in the original CID spectrum. (d) Sequence alignment of the newly sequenced *P. interruptus* CLDH with previously reported CLDH from other species, highlighting the sequence conservation of this peptide family. Species names in purple shade indicate insects, blue shade indicates crustacean. Sequence letters highlighted in green shade are shared among various species.

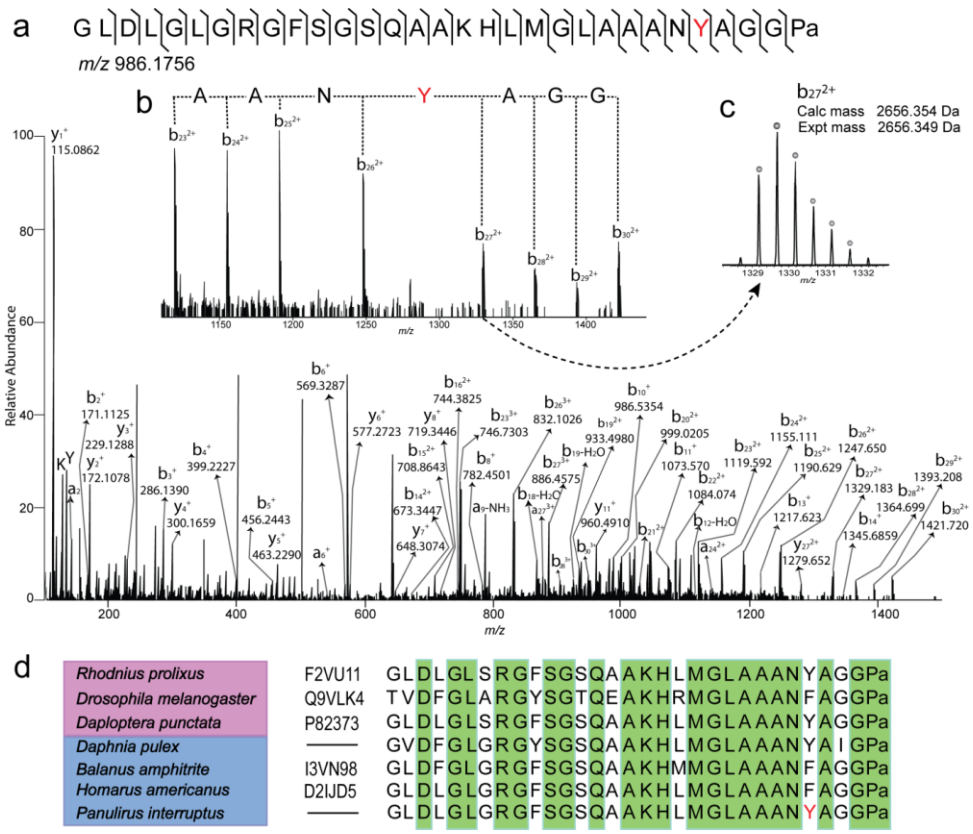


Figure 3. *De novo* sequencing of two novel *P. interruptus* neuropeptides as homologues of SIFamide. (a) A zoomed-in MS spectra showing the calulcated deconvoluted masses of the three precursors identified in single MS/MS event with ProSight search with the ion at *m/z* 465.9091 assigned as the product of neutral loss from the ion at *m/z* 471.9124. Isotopic peaks from different precursors are indicated with colored dots, whereas the overlapped fragment ions are shown in green. (b) The multiplexed MS/MS spectra that exhibits the fragmentation pattern of the two neuropeptides, with the fragments coming from each peptide identified by corresponding colors. (c) The fragmentation map and sequence coverage of the two SIFamide homologues.

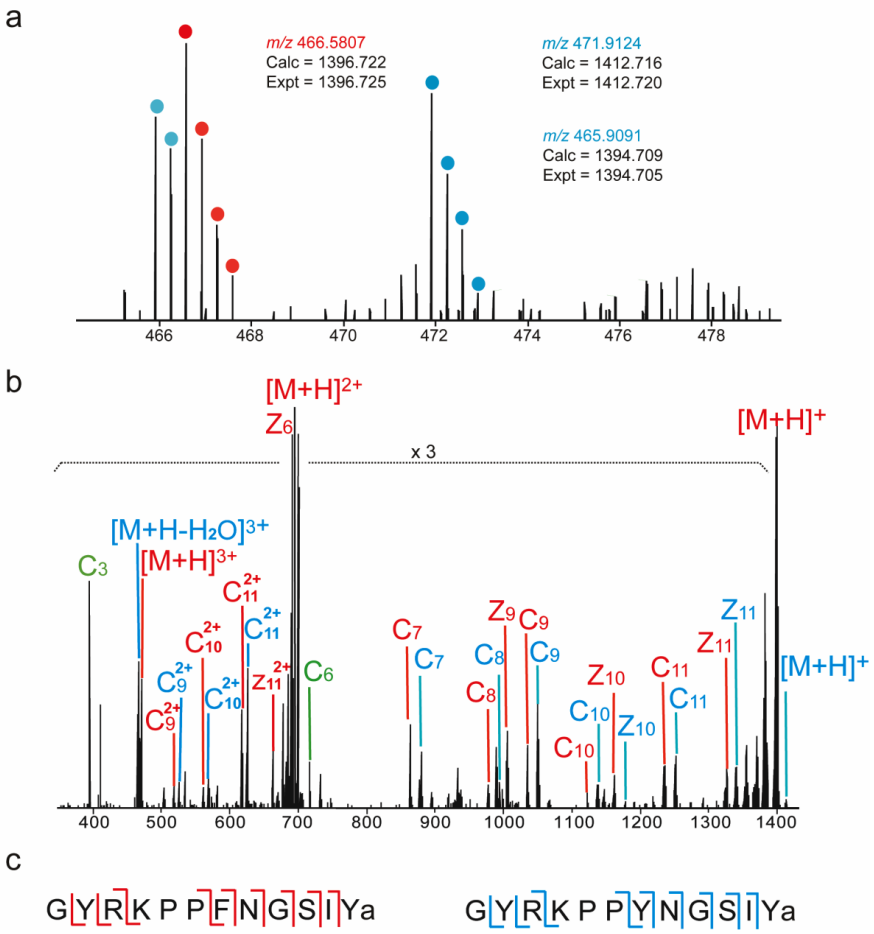


Figure 4. Discovery of a novel crustacean orcokinin motif from *P. interruptus* orcokinins. (a) A consensus sequence of NFDEISRSX(G/S)FGFX(H/N/V/A) shared by crustacean orcokinin neuropeptides. (b) The novel motif NFDEIDRAG summarized based on *P. interruptus* orcokinin neuropeptides. (c)-(f) *De novo* sequencing of representative A-orcokinins NFDEIDRA, NFDEIDRAGLA and NFDEIDRAGLAF, NFDEIDRAGLGFH with MS/MS spectra acquired on Q-TOF.

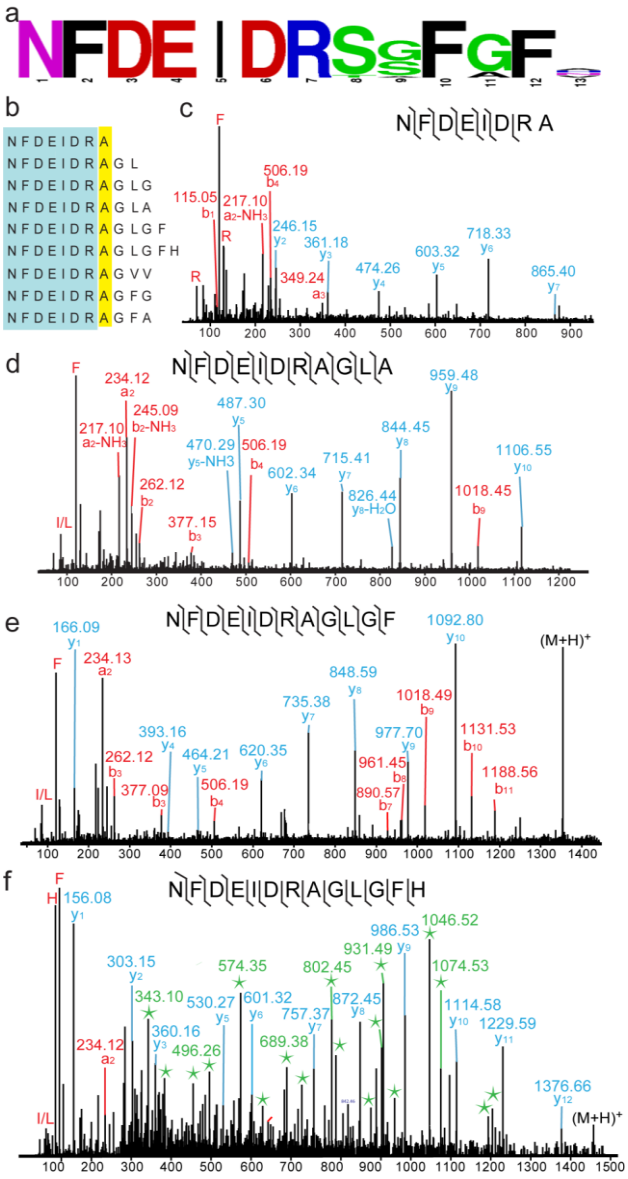


Figure 5. Reductive dimethylation-assisted *de novo* sequencing of *P. interruptus* neuropeptides. (a) A schematic illustration of the formaldehyde labeling reaction. (b) *De novo* sequencing of dimethylated orcokinin NFDEIDRAGLGFH. (c) *De novo* sequencing of dimethylated RFamide QDLDHVFLRFa. (d) *De novo* sequencing of RFamide pQDLDHVFLRFa that cannot be labeled due to N-terminus blockage.

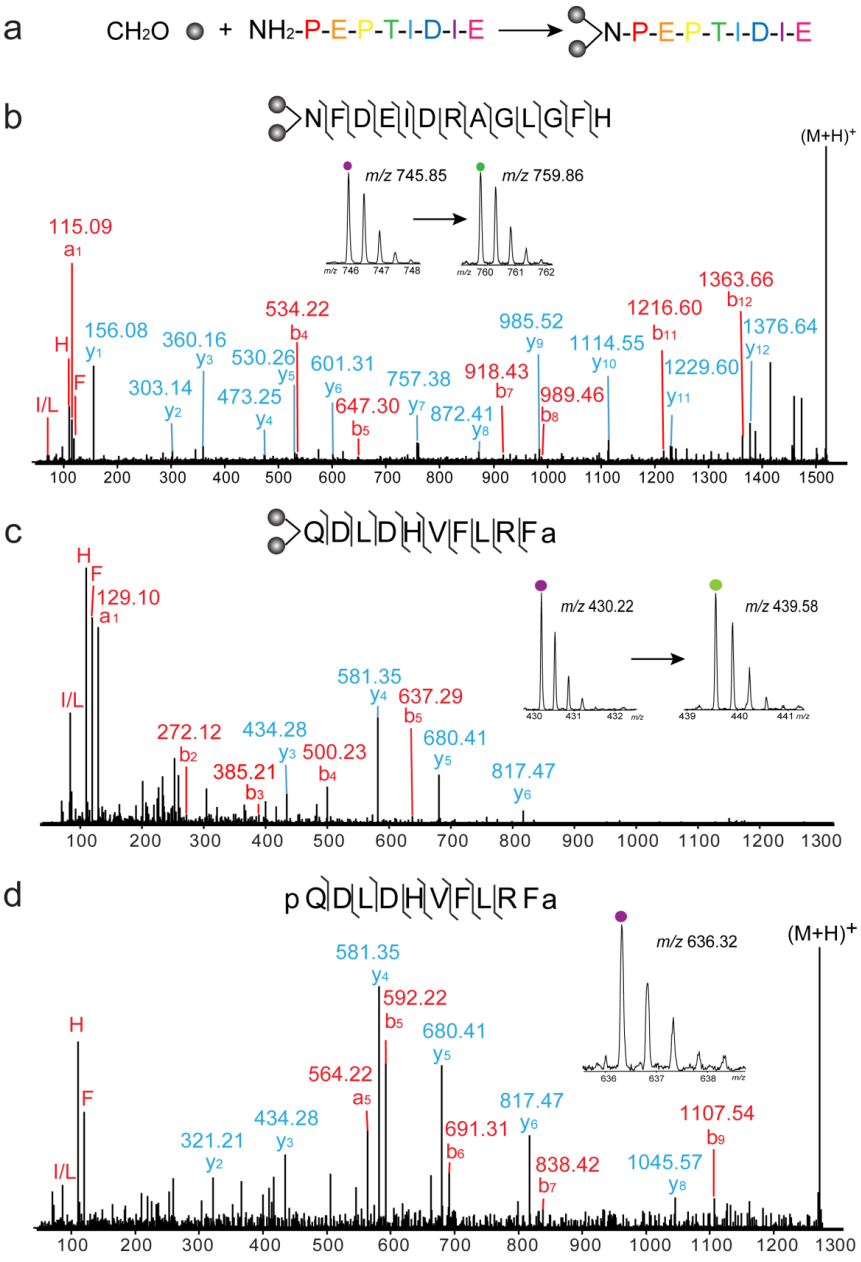


Figure 6. Direct tissue analysis of *P. interruptus* brain region 1 and 2. (a) Profiling spectrum of neuropeptides desorbed/ionized from region 1 (R1) that exhibited overlapped yet distinct peptide profiles to that from region 2 (R2) in (b).

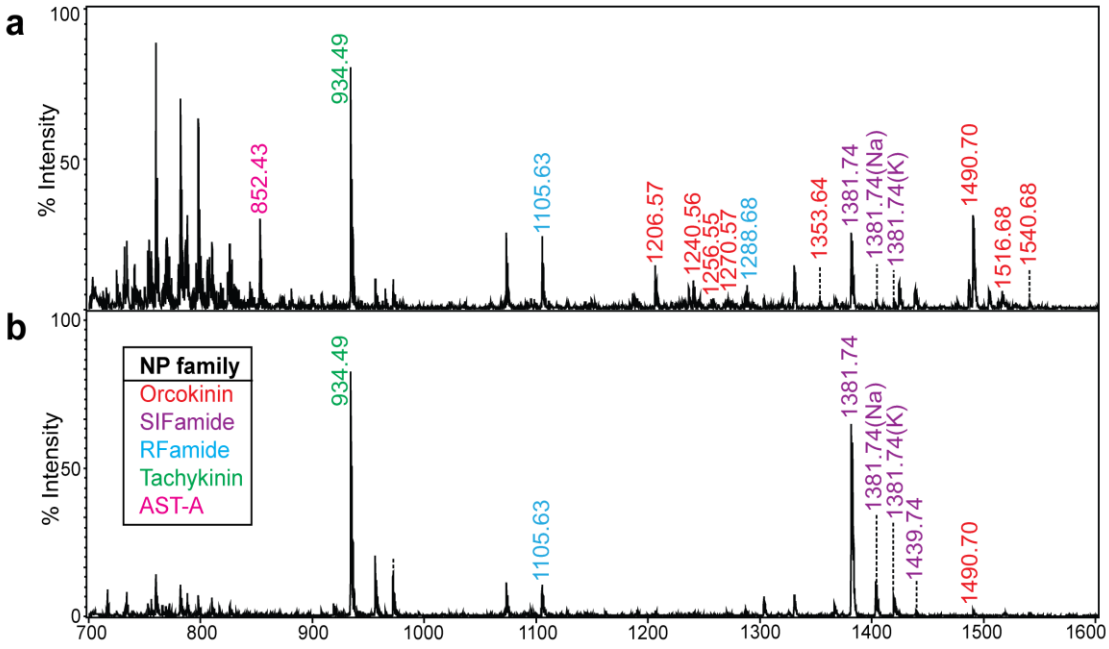
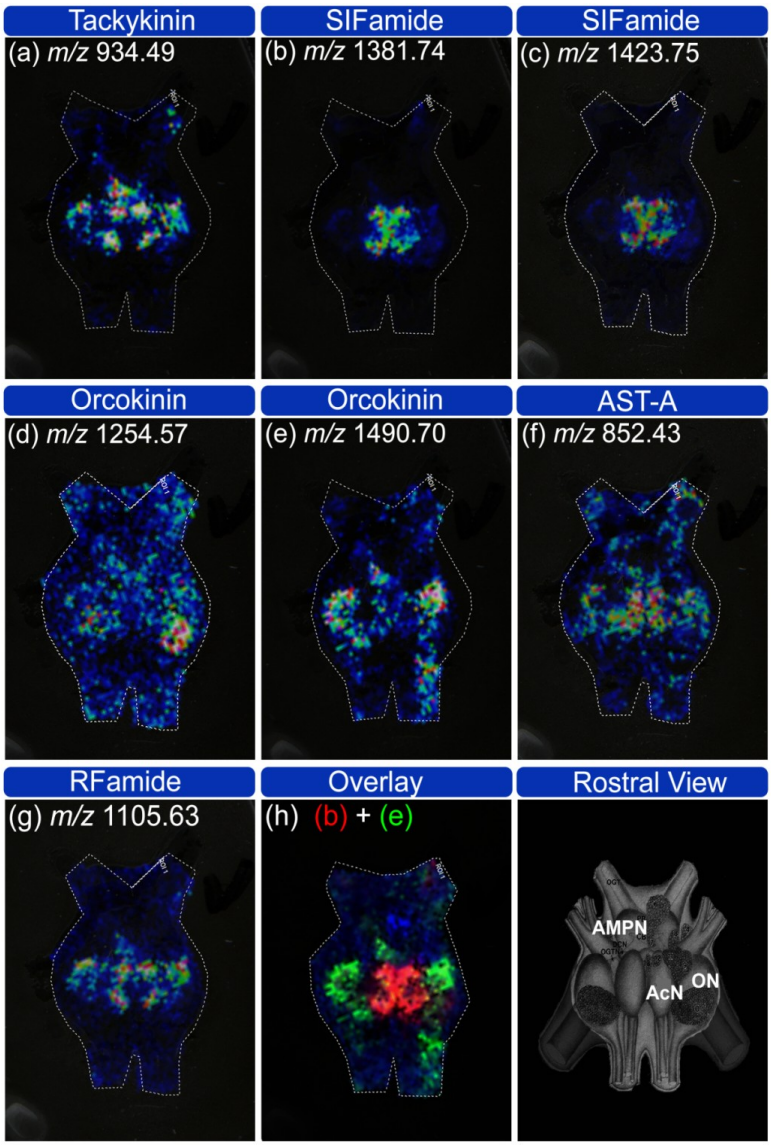
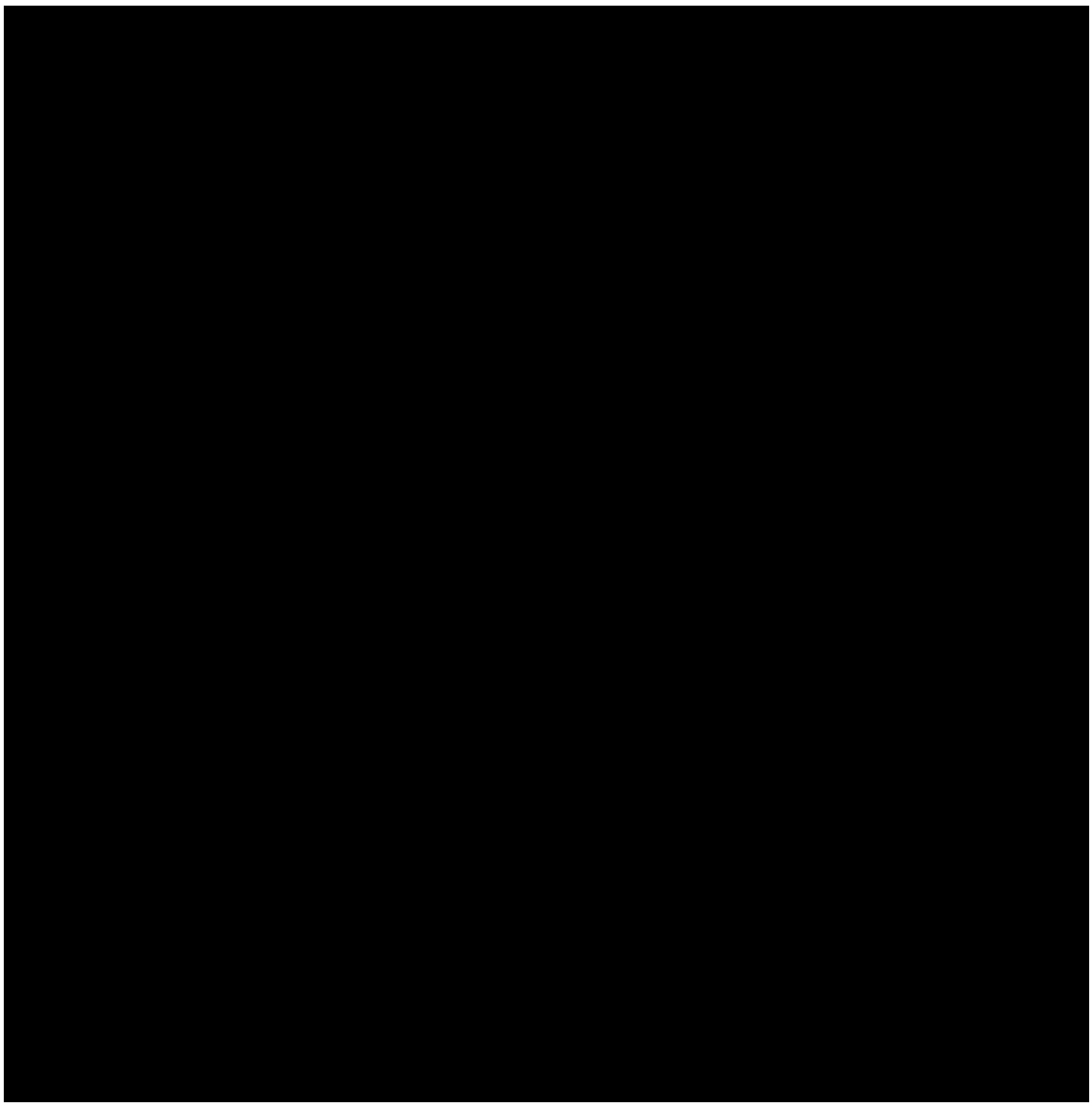


Figure 7. MSI of *P. interruptus* brain shows the localization of several neuropeptides from tachykinin, SIFamide, orcokinin, AST-A and RFamide families. The distribution maps of neuropeptides visualized by MSI technique shown in (a)-(g) could be correlated with the schematic illustration of *Panulirus* brain in rostral view. (h) An overlay of the distribution maps of a SIFamide at *m/z* 1381.74 in (b) and an orcokinin at *m/z* 1490.70 in (e).



Supplemental Information

Supplemental Figure S1. Profiling of *P. interruptus* brain extract on MALDI-TOF/TOF. MALDI-MS profiling spectra of (a) CE Fraction #8 that shows the tachykinin APSGFLGMRa at *m/z* 934.49 and (b) CE Fraction #14 that is dominated by the SIFamide GYRKPPFNGSIFa at *m/z* 1381.74 displayed significantly reduced background compared to the regular profiling of (c) brain extract without CE fractionation.



Chapter 7

Top-down Proteomics with Mass Spectrometry Imaging: A Novel Approach Towards Discovery of Biomarkers for Neurodevelopmental Disorders

Adapted from **Hui Ye**, Rakesh Mandal, Adam Whitehorse, Paul Martin Thomas, Neil Kelleher, Chrysanthy Ikonomidou, Lingjun Li. Top-down proteomics with mass spectrometry imaging: a novel approach towards discovery of biomarkers for neurodevelopmental disorders. *To be submitted*.

Abbreviations

ASD	autism spectrum disorders
GABA	γ-aminobutyric acid
MSI	mass spectrometry imaging
MALDI-MS	matrix-assisted laser desorption/ionization mass spectrometry
MK801	dizocilpine
NMDA	N-methyl-D-aspartate
P2, P4, P10	postnatal day 2, 4, 10
PEP-19	purkinje cell protein 4
m/z	mass-to-charge ratio
RP	reverse-phase
HPLC	high performance liquid chromatography

Abstract

In the developing mammalian brain, inhibition of NMDAR can induce widespread neuroapoptosis, inhibit neurogenesis and cause impairment of learning and memory. Although some mechanistic insights into adverse neurological actions of these NMDAR antagonists exist, our understanding of the full spectrum of developmental events affected by early exposure to these chemical agents in the brain is still limited. Here we attempt to gain insight into the impact of a pharmacologically induced excitatory/inhibitory imbalance in infancy on the brain proteome using mass spectrometry imaging (MSI). Our goal was to study changes in protein expression in postnatal day 10 (P10) rat brains following neonatal exposure to the NMDA receptor antagonist dizocilpine (MK801). Analysis of rat brains exposed to vehicle or MK801 and comparison of their MALDI MS images revealed differential abundances of several proteins. We then identified these markers such as ubiquitin, purkinje cell protein 4 (PEP-19), Cytochrome c oxidase subunits and calmodulin, by a combination of reversed-phase (RP) HPLC fractionation and top-down tandem MS platform. Our findings indicate that a brief neonatal exposure to a compound that alters excitatory/inhibitory balance in the brain has a long term effect on protein expression patterns during subsequent development, highlighting the utility of MALDI-MSI as a discovery tool for potential biomarkers.

7.1 Introduction

Sedatives, anaesthetics and anticonvulsants, such as ketamine, nitrous oxide (laughing gas), propofol, sevofluran, benzodiazepines and barbiturates are used frequently in obstetric and pediatric medicine. Neurological abnormalities associated with developmental exposure to such drugs and to environmental toxins, such as ethanol, have been reported. It has been suggested that disturbances in neurotransmitter signalling during critical periods can redirect subsequent development. Detrimental long-term effects observed in humans range from mild neurobehavioral disturbances (hyperactivity/attention deficit and learning disabilities) to severe mental retardation and autism spectrum disorders (ASD). [1, 2] Interestingly, a potential unifying explanation for the pathogenesis of ASD is the disruption of excitatory/inhibitory circuit balance during critical periods of development. [3] Excitatory/inhibitory neurotransmission balance is important for single neurons to acquire multiple functional properties through an experience-dependent maturation process, which occurs during critical periods of brain development.

Previously we provided evidence that excitation/inhibition imbalance in early development can cause long lasting changes of the brain proteome. [4] In addition, compounds which inhibit glutamate N-methyl-D-asparate (NMDA) receptors or activate γ-aminobutyric acid A ($GABA_A$) receptors and cause excitation/inhibition imbalance, can induce widespread neuroapoptosis in the developing rodent brain, inhibit neurogenesis [5] and cause impairment of learning and memory. [6] Interestingly, early disruption of excitatory/inhibitory balance in infant mice (P0-P7) via blockade of NMDA receptors or activation of $GABA_A$ receptors caused very similar acute and long-term changes of the cerebral proteome. [4] Changes of the cortex proteome one

day (P7), one week (P14) and four weeks (P35) following treatment at P6 suggest that short lasting suppression of synaptic neurotransmission during critical periods of brain development causes long lasting dysregulation of proteins associated with apoptosis, oxidative stress, inflammation, proliferation and neuronal circuit formation. [4] These observations support the hypothesis that early disruption of excitatory/inhibitory balance in the brain interferes with many subsequent critical developmental events.

The integration of gel electrophoresis and liquid chromatography coupled to MS enables more comprehensive characterization of complex proteomes to detect disease biomarkers [7-10] or evaluate the response to stress or to the exposure of compounds of interest. [11-14] Previous studies of Kaindl et al. examined the whole cortex proteome after treatment with MK801 in infancy using this conventional approach, [4] yet more precise localization of the detected protein expression changes to specific brain regions has not been conducted. Conventional immunohistochemical (IHC) staining allows for obtaining high resolution distribution images of targeted proteins. However, a significant limitation to this method is the need for labeling, which means that the target molecules must be known prior to the experiment. Alternatively, mass spectrometric imaging (MSI) has evolved as a powerful tool based on matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). This technique allows for the determination of the mass and location of biomolecules directly from a tissue sample. [15-18] MALDI MSI has enormous advantages over conventional protein imaging techniques in that not only it is label-free but also it enables simultaneous mapping of numerous molecules in tissue samples with great sensitivity and chemical specificity. MALDI MSI has proven to be a valuable technology with numerous applications in localizing proteins [19-22], examining lipid distributions, [23, 24] mapping

neuropeptides [25, 26] and imaging drugs or metabolites [27] at both organ and cellular levels by varying the experimental conditions. Since no prior knowledge of molecular identities is required for MSI applications, it has become a useful biomarker discovery tool to compare analyte expression pattern changes by analyzing multiplexed data sets. For example, Minerva et al. evaluated the capability of MALDI-MS to study peptide expression patterns in the mouse pancreas under normal and pathological conditions and identified a peptide marker, the C-peptide of insulin, which was dramatically up-regulated in obese mice. [28] Sköld et al. observed a significant decrease of a large neuron-specific peptide, PEP-19, in a mouse model of Parkinson's disease (PD) using MALDI-MSI. [29] Hanrieder et al. quantified relative abundances of striatal neuropeptides in a rat model of L-DOPA-induced dyskinesia (LID) and revealed several upregulated dynorphins and enkephalins in the dorsolateral striatum of high-dyskinetic rats. [30] Two prodynorphin derived peptides, dynorphin B and α-neoendorphin, were positively associated with LID severity. Jones et al. also revealed substantial neuropeptide changes following transient electrophysiological events such as cortical spreading depression (CSD) in mice brain. [31] Moreover, Meistermann unambiguously identified transthyretin (Ser^{28}-Gln^{146}) as a potential toxicity biomarker in kidneys of rats that were administered the nephrotoxin gentamicin. [32] Lagarrigue identified seven potential biomarkers of childhood absence epilepsy as Synapsin-1 fragments. [31] Li and Hummon further explored the application of MALDI-MSI to image proteins from 3D colon carcinoma cell cultures and visualized cytochrome C and histone H4 in the cell spheroid. [33] In clinical settings, MALDI MSI has successfully discovered tissue biomarkers for diseases such as myxoid sarcomas, [34] adenocarcinoma, [35] gastric cancer, [36] hypertension, [37] metastatic

melanoma [38] and prostate cancer. [20] These findings underscore the critical role MALDI-MSI plays in the study of peptide/protein dynamics in the context of human diseases. With MSI, a better understanding of the molecular pathology of various diseases such as cancer and neurodegenerative diseases can be obtained.

In the present study we used MSI with the aim to investigate region specific changes in the brain proteome following blockade of NMDA receptors in infancy. We first pinpointed 22 potential protein biomarkers using both a profiling and an imaging strategy. We then implemented a protein identification strategy combining a protein extraction step of tissues followed by RP HPLC fractionation and top-down sequencing by high resolution tandem MS.

7.2 Experimental Section

Animal Experiments

Animal experiments were performed according to institutional guidelines. Wistar rats were treated with the NMDA receptors antagonist dizocilpine ((+)MK801; Tocris, Bristol, UK) 1 mg/kg i.p. on P2 and P4. An equal number of sex-matched littermates that received vehicle (normal saline) served as controls. Rats were terminally anesthetized with 5% isoflurane, transcardially perfused with saline solution and decapitated on P10 (n=4 per group).

Sample Preparation for MALDI-MS

Rat brains were dissected out freshly and cut along the midline to halves. Then the hemispheres were immediately embedded into gelatin solution (100 mg/ml) and snap-frozen in -80° C freezer for further processing. Sagittal sections were acquired at a thickness of 12 μm on a cryostat (HM525, Thermo Fisher Scientific, Waltham, MA) and deposited onto indium tin oxide (ITO) coated conductive glass slides (Delta

Technologies, Loveland, CO). Tissue sections were immersed in 70% ethanol for 1 min and another 95% ethanol bath for 1 min for protein fixation and lipid removal. The sections were then allowed to dry for 30 min under vacuum. For profiling purpose, matrix deposition was conducted manually with a pipette directly on tissue sections. 300 nL of freshly prepared 20 mg/ml sinapinic acid (Sigma Aldrich, St. Louis, MO) dissolved in 50% acetonitrile, 0.1% formic acid were sequentially deposited onto 14 regions of a brain section. For imaging purposes, 20 mg/ml sinapinic acid dissolved in 50% acetonitrile, 0.1% formic acid was sprayed onto a brain section homogenously using an airbrush held 35 cm from the tissue. Five coats were applied and the spray duration for each coat was 30 s with 1 min dry time between each cycle.

MALDI-MS Profiling and Imaging

MALDI-MS profiling spectra of the manually spotted tissue sections and imaging data of intact brain sections were acquired on an Autoflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Billerica, MA) equipped with a 200 Hz smart beam laser. The following parameters were adopted in the positive linear mode at a mass range of 3-25 kDa for spectral acquisition with a delayed extraction of 50 ns: ion source 1 voltage 20.00 kV, ion source 2 voltage 18.55 kV, lens voltage 6.80 kV and pulsed ion extraction 130 ns. Protein calibration standard I (Bruker Daltonics, Billerica, MA) was used to externally calibrate the instrument before data acquisition. Briefly, 2000 consecutive laser shots were accumulated for each deposited matrix droplet. The profiling spectra were smoothed and baseline subtracted using flexAnalysis (Bruker Daltonics, Billerica, MA). As for imaging, automated MSI acquisitions of the brain sections were controlled by flexImaging software (Bruker

Daltonics, Billerica, MA). Array of spectra was collected at 200 μm intervals in both x and y dimensions, and each spectrum consists of 200 laser shots.

MALDI-MS Data Analysis

The MATLAB-supported software ClinProTools 2.2 enables baseline subtraction, normalization, peak calibration, spectral alignment and statistical analyses of the resulting profiling spectra. The profiling spectra were divided as the vehicle-treated and the MK801-treated class and loaded into the software. By averaging, a single representative spectrum composed of 14 discrete spots from each section was generated for each class, which was used in all of the statistical analyses. The profiling spectra from the two classes were also compared region-wise for evaluating the sensitivity of cells in response to the MK801 treatment. Four serial sections were collected and profiled for each animal (N=4). PCA plot and loading plot were generated for each comparison group.

For image processing, the Bruker proprietary software flexImaging 3.0 was utilized to generate MS images and overlay the optical image with the MS images for good visualization. One section from each sacrificed animal was collected (N=4) and MS images were acquired from sections collected from animals from two classes. MSI data were also imported into ClinProTools to analyze differences in abundance of proteins in the two animal groups. For comparison, specific regions of interest (ROI) were selected manually and the spectra resulting from the ROI were averaged and compared in ClinProTools. PCA plots and loading plots were generated for each comparison group.

Liquid-phase Protein Extraction and Fractionation

Hemisphere of rat brain was manually homogenized in acidified methanol

(1:9:90 acetic acid/water/methanol) at a 1:5 ratio (tissue/buffer) on ice. The extract was centrifuged at 16,100 rcf for 10 min in an Eppendorf 5415 D microcentrifuge (Brinkman Instruments), and the resulting supernatant was collected followed by vacuum drying. The mixture was then resuspended in 30 µL of water containing 0.1% formic acid and separated on a 2.1×100 mm Phenomenex Kinetex 2.6 µm particle C18 reversed phase (RP) HPLC column using a 60 min gradient with a flow rate at 0.3 mL/min. Fractions were collected every three minutes and concentrated in a Savant SC 110 Speedvac concentrator (Thermo Electron Corporation, West Palm Beach, FL). The intact protein fractions were reconstituted in 0.1% formic acid and screened by MS profiling using MALDI-TOF/TOF via on-plate mixing with sinapinic acid. The spots were analyzed according to previous instrumentation settings and matched to *m/z* values obtained with *in situ* tissue profiling.

On-line Top-down MS/MS using Nano-LC-ESI-LTQ-Orbitrap Elite

The fractions that contained peaks of interest were further analyzed by on-line top-down MS on an Ultimate 3000 RSLCnano system coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Fractions of interest (0.5 µL) were injected onto a 2 cm, 150 µm i.d. PLRP-S (d_p 5 µm, pore size 1000Å) trap column. A 10 cm, 75 µm i.d. PLRP-S column was used for separation. The gradient was delivered at 300 nL/min starting at 5% B (95% acetonitrile and 0.2% formic acid) and rose to 10% B at 7 min, 50% B at 50 min, and 85% B at 58 min. The mass spectrometer was operated in a data-dependent mode, performing higher-energy C-trap collision induced dissociation (HCD)-MS^2 (scan 1), conventional collision induced dissociation (CID)-MS^2 (scan 2) and electron transfer dissociation (ETD)-MS^2 (scan 3) on each of the Top 3 precursors (selected by intact

mass) in a FT-MS precursor scan. More details on mass spectrometric acquisition parameters can be found in **Supplemental Information**.

Data were deisotoped with Xtract using the cRAWler algorithm (ThermoFisher, Bremen, Germany) and searched with a custom 168-core ProSightPC 3.0 cluster using an iterative search tree. The detailed information on search tree can be found in Supplemental Materials section. ProSightPC Warehouses were created against the UniProt 2011_04 build filtered for *Rattus norvegicus* (TaxId: 10116). The "simple" rat database was created with all base sequences (signal peptide cleavage products, alternative splice forms, etc) and N-terminal acetylation applied. The simple database has 33,998 base sequences and 109,973 modified forms. The "complex" rat database contains all of the information in the simple database in addition to the annotated post-translational modifications in UniProt. The complex database had 40,990 base sequences and 3,616,723 modified forms. All retrieved proteins are considered as positive identification if: (a) an expect score (E value) less than 10^{-5}, and (b) mass error (Δm) less than 5 ppm between the calculated and expected intact masses [39]. The protein *m/z* observed on MALDI-MS, entry name, accession number, protein identity, calculated mass deconvoluted from the *m/z* observed on LTQ Orbitrap, expected mass, average mass, post translational modifications (PTMs) present in the sequence and E value of the search results are all listed in **Table 1-2** and **Supplemental Information**. The annotated MS/MS spectra were recorded in **Appendix III**.

7.3 Results

In this study, we used a combined profiling and imaging strategy as illustrated in **Figure 1** to identify the differentially expressed proteins in the vehicle and MK801

treated rat brains. Profiling involves manually depositing matrix onto the regions of interest of the brain section and analyzing the section region by region, whereas imaging requires the whole tissue section to be coated with a homogenous matrix layer and scanned according to a predefined array.

In situ *profiling of the vehicle- and MK801-treated infant rat brains*

For high-throughput profiling, four pairs of sex-matched P10 infant rat brains that had undergone treatment with saline as the control group (N=4) and MK801 as the treated group (N=4) were compared. Four sections were used from each brain as technical replicates, and each section was manually profiled with fourteen spots representing the corresponding major anatomical structures in a rat brain sagittal section. Specifically, the fourteen major regions were divided based on rat brain atlas (www.brainmaps.org) as designated in **Fig. 1C**, including olfactory bulb as region #1, accumbens nucleus (Acb) as region #3, cortex as region #2, 5, 9 and 12, corpus callosum (cc) as region #4, hypothalamus (Hypo) as region #6, thalamus (TH) as region #7, hippocampus (Hippo) as region #8, medulla as region #10, colliculi as region #11, pons as region #13 and cerebellum as region #14. Spectra produced from profiling the fourteen brain regions clearly differentiate the anatomical regions as previously reported. More importantly, the MS profiles of proteins from each of the regions were compared between infant rat brains treated with vehicle and MK801 with statistical support provided by ClinProTools. The differences of peaks were evaluated by *m/z*, signal intensity and p value. Although most of the masses found in the two classes resemble each other, the overlaid representative spectra from the two classes displayed differential abundance levels of several peaks in the control as the green trace and the treated samples as the red trace shown in **Fig. 2A**. The enlarged

spectra are displayed in **Fig. 2B-E**. **Fig. 2B** displays the peaks at *m/z* 8468 and 8566. In this panel, the former peak exhibited relatively the same abundance between the two groups, whereas the latter was significantly upregulated in the averaged spectra acquired from MK801-treated rat brains. **Figure 2C** shows another peak at *m/z* 6718, the abundance of which was also greatly increased in the MK801 treated group, nevertheless, the peak at *m/z* 6649 was detected at almost the same intensity in the two groups. In contrast with **Fig. 2B** and **2C**, **Fig. 2D** and **Fig. 2E** exhibit two ions at *m/z* 15825 and *m/z* 16792 that were both down-regulated in the MK801-treated group. More proteins differentially expressed in the control compared with the treated samples are summarized in **Table 1** with the information of observed *m/z*, p-value of t-test between the two groups and their relative abundance ratio. Based on previous MALDI-MS literature [29, 32, 40] and database search in UniProt, several peaks that exhibited differential abundance levels after MK801 treatment were mass-matched to proteins with a mass error tolerance of ± 4 Da. For example, the proteins displayed higher abundances in the treated group were assigned as ubiquitin at *m/z* 8566 as shown in **Figure 2B** and PEP-19 at *m/z* 6718 as displayed in **Figure 2C**, whereas the protein at *m/z* 6649 that displayed the same abundance level following the treatment was assigned as cytochrome c oxidase subunit 6a-L as in **Fig. 2C**. In contrast, the protein that exhibited decreased abundance in the MK801-treated rat brains was identified as calmodulin with *m/z* 16792 in **Fig. 2E**. Although MALDI-MS has demonstrated its capability of providing accurate *m/z* of the protein species present in rat brain sections for identification, the identities of the molecules that have not been reported, such as the peaks at *m/z* 8468 and 15825 in **Fig. 2B** and **2D**, could not be assigned due to the lack of structural information. A comprehensive strategy that

combines MALDI-MS with high resolution top-down MS/MS sequencing was employed for protein identification and will be discussed in the following section.

Mapping the spatial distribution of peaks of differential expression patterns from the vehicle- and MK801-treated infant rat brains

Although the profiling strategy provides high throughput information regarding regional differences, MSI allows for mapping of the ions of interest in the entire rat brain section with much higher spatial resolution than regional profiling. Overall, the average spectra for the imaged brain sections of MK801 treated rats exhibited great similarities to those of vehicle-treated rats, agreeing well with the average profiling spectra. Nevertheless, the differential features highlighted in the profiling approach were confirmed through MSI using one rat brain section from each group as depicted in **Fig. 1E**. Additionally, statistical software also aids in validating the potential biomarkers that displayed differential expression levels in the vehicle and treated groups. For example, the profiling spectra showed that the ion at *m/z* 6718 exhibited significant up-regulation by 1.8 fold in the MK801-treated group compared to the vehicle-treated one as in **Table 1**. More specifically, regions #1, #2, #5, #9, #7 and #14 displayed the most drastic changes after the treatment (**Fig. 3A)**. **Figure 3B** is an optical image of a rat brain section that underwent MSI experiments. As highlighted in **Figure 3B,** region #2, #5 and # 9 correspond to the frontal and cingulate cortex, and region #1, #7 and #14 are registered to the olfactory bulb, TH and cerebellum, respectively. Among all the regions, olfactory bulb as region #1 exhibited the most prominent changes with a 2.5 fold increase comparing to roughly 1.5 fold as shown by other regions. This observation was confirmed by imaging data as shown in **Figure 3C. Figure 3C** shows two MS spectra averaged specifically from ROI 1 and

ROI 2 in **Figure 3E**, which correspond to the olfactory bulb regions from the vehicle and the MK801 treated rat brain sections, respectively. The peak at *m/z* 6718 was elevated to roughly 3 fold from the MK801-treated spectrum in comparison to that of a vehicle-treated one, consistent with the profiling results. Moreover, the loading plot shown in **Figure 3D** suggests that the *m/z* 6718 peak contributes significantly to the distinction of the two ROIs, especially in load 2, providing statistical evidence to the differentiating role of this ion. **Figure 3E** displays more straightforward proof by showing the MS images of the *m/z* 6718 ion obtained from the vehicle and MK801 treated section. The difference in color suggested that the *m/z* 6718 ion was significantly more concentrated in the olfactory bulb of the MK801 treated section. The cortex also exhibited explicit difference in the vehicle and the treated group. Moreover, the abundance of the *m/z* 6718 ion in TH and cerebellum were also increased as indicated by the MS images. The differential expression levels of *m/z* 6718 in the vehicle and MK801-treated rat brains were reproducibly proven by the integration of profiling and imaging data, demonstrating consistency of our combined strategy.

Other than the *m/z* 6718 ion, the abundance of the peak at *m/z* 8566 in the MK801-treated group was elevated by 1.6 fold compared to the vehicle-treated one. Profiling spectra of regions #2, #5, #3, #8, #9 and #12 displayed the most distinguishable changes in expression levels of the *m/z* 8566 ion as shown in **Fig. 4A**. As annotated in **Figure 4B**, regions #2 and #5 are registered as cortex; regions #7 and #14 correspond to TH and cerebellum, respectively; and regions #3 and #6 are assigned as Acb and Hypo. Accordingly, the intensity of the *m/z* 8566 ion in the MS spectra of the TH regions from the MK801-treated rat was elevated by around 1.4 fold

comparing to the vehicle-treated one as indicated in **Figure 4C**, which agreed well with the increase displayed in **Figure 4A**. The spectra were averaged specifically from the TH regions of the vehicle and MK801-treated rat brain sections, which were designated as ROI 1 and ROI 2, respectively. **Figure 4D** shows the statistical loading plot of the two ROIs, indicating that the *m/z* 8566 ion contributes most significantly to differentiate the TH regions of the two groups in Load 1. The statistical evidence was further supported by the MS images displayed in **Figure 4E**. Acb as region #3, Hypo as region #6 and cortex as region #2 and #5 all showed significant increase in relative abundance of the *m/z* 8566 ion in the image obtained from the MK801-treated section compared with the vehicle-treated section, as indicated by the color scale in **Figure 4E**. TH as region #7 and cerebellum as region #14 also exhibited noticeable increase of concentration in the MK801-treated section compared with the vehicle one. In summary, the utilization of the MALDI-MS profiling and imaging techniques was exemplified by the differentiating peaks at *m/z* 6718 and 8566 and have been applied to discern the proteins that changed significantly following MK801 treatment due to the reproducibility and consistency of the combined strategy.

Protein Identification

Since knowledge about the identity and chemical form of the molecular species of interest is essential to investigate their biological roles, efforts were focused on identifying the peaks that display significantly different expression levels after MK801 treatment. Although the discrimination could be made by combining the imaging and the profiling techniques utilizing MALDI-TOF/TOF MSI, it is significantly more difficult to identify the potential proteins of interest via MALDI-MS alone due to limited fragmentation obtainable from the singly charged high-mass

ions produced by MALDI-MS. The initial step of identification was fulfilled by the conventional fractionation, which helped eliminating ion suppression effects, reducing sample complexity and thus increasing the dynamic range of the analysis. The rat brains were pooled and fractionated in RP HPLC to concentrate the targeted proteins and reduce the complexity arising from high abundance proteins of tissues. All fractions were then screened on MALDI-TOF/TOF with the aim to ascertain if the proteins of interest observed on MALDI-MS from the brain sections were extracted by tissue homogenization and purified by crude fractionation. As shown in **Fig. 5A**, the peaks of *m/z* 6718 and 8566 were detected at high abundance in fraction #7 and #10, respectively. The fractions containing the targeted proteins selected via MSI were then analyzed on a nanoLC-ESI-Orbitrap Elite system. The accurate masses provided by this high-end instrument were then deconvoluted, and the averaged masses were compared to the masses detected on MALDI-TOF/TOF for mass-matching. **Fig. 5B** shows a representative MS spectrum, which contains several multiply charged ions, including *m/z* 611.392 (z=11), *m/z* 672.432 (z=10), *m/z* 747.036 (z=9), *m/z* 840.290 (z=8), *m/z* 960.183 (z=7), *m/z* 1120.05 (z=6) and *m/z* 1343.86 (z=5). After spectrum deconvolution performed by software Xtract 3.0, these ions were all attributed to the molecule with monoisotopic mass at 6714.25 Da. Moreover, the well-resolved isotopic distribution in the insert of **Fig. 5B** suggests its neutral average mass at 6717 Da, which could be matched with the singly-charged *m/z* 6718 peak generated by MALDI-MS. The mass discrepancy between the peaks observed on TOF/TOF and Orbitrap further demonstrates the different resolution provided by two instruments. Followed by mass-matching with the targeted proteins, the selected ions were then subjected to top-down high resolution tandem MS, using a

complementary suite of fragmentation techniques: HCD, CID and ETD. **Fig. 5C** displays the multiple fragmentation spectra arising from the *m/z* 1120.05 (z=6) ion using these various techniques. The rich sequence-specific fragment ions produced by the complementary fragmentation techniques were used to search the rat protein database with ProSightPC 3.0. Based on the ultra-high mass resolution (120,000 in MS^1 and 60,000 in MS^2 at *m/z* 400) and accuracy (mass error smaller than 5 ppm) together with the high quality MS/MS spectra, we identified the peak at *m/z* 6718 on the MALDI-TOF/TOF as N-term acetylated PEP-19 and *m/z* 8566 as ubiquitin-40S ribosomal protein S27a (Ubiquitin). The best fragmentation spectra corresponding to the two ions are displayed in **Fig. 6A** and **6C**, validating the previous identifications assigned based on masses. The insert in **Fig. 6A** is the deconvoluted spectrum of the precursor ion at *m/z* 1120.05, showing its monoisotopic mass of 6714.25 Da. The original tandem MS spectrum is enlarged to exhibit the most abundant fragment ion y_{27} at *m/z* 997.8751, which demonstrates the high mass accuracy and resolution of the Orbitrap Elite system we employed in this study. In the inserts of **Fig. 6C**, the precursor ion at *m/z* 1070.96 with a monoisotopic mass of 8559.61 Da gave rise to an abundant ion y_{40} at *m/z* 1141.37. The deconvoluted mass of this ion is calculated as 4561.45 Da, exhibiting merely a 0.01 Da mass error from the theoretical value. The high-quality tandem MS spectra led to high-confidence protein identification, facilitating future mechanistic studies on these proteins and related disease.

Inevitably, the same protein forms of high abundance were repeatedly selected for fragmentation within a single run. Therefore, increased cleavage site and improved fragmentation coverage were observed for all of the three fragmentation techniques. For example, the ion shown in **Fig. 6B** corresponds to acetylated PEP-19,

identified with an E-value of 1.5E-77. Among all the fragmentation methods, CID provided the most complete coverage for PEP-19; HCD delivered less fragment ions in comparison to CID yet accurately located the acetylation site at the Serine of N-terminal; whereas ETD produced ~1/3 of the number of the HCD fragment ions. As for ubiquitin, HCD produced the most efficient fragmentation covering almost all of the possible fragmentation sites, and CID produced slightly less fragment ions as displayed in **Fig. 6D**. Nevertheless, ETD again displayed merely ~1/3 of the fragment ions obtained from either HCD or CID. **Fig. 6E** illustrates the product ions generated by CID, HCD and ETD. ETD is a radical driven fragmentation technique and cleaves at N-Cα bond, which is distinct from those of HCD and CID. Therefore, ETD is complementary to HCD and CID fragmentation methods. By comparing the fragmentation patterns produced by the three fragmentation techniques as shown in **Fig. 6B** and **6D**, ETD predominantly produced low mass fragment ions, whereas CID and HCD gave rise to fragment ions spanning the whole mass range. The utilization of all three fragmentation techniques significantly improves the coverage of the cleavage sites from the targeted proteins. The differentiating proteins quantified by MALDI-MS and verified by top-down tandem MS sequencing were listed in **Table 1** with details presented in **Supplemental Table 1**, whereas the proteins that displayed relatively the same abundance among groups via MALDI-MS yet identified by top-down sequencing were summarized in **Table 2**. A number of proteins, such as mitochondria import inner membrane translocase at *m/z* 10283, adenylate cyclases type 10 at *m/z* 9193, polyubiquitin at *m/z* 8468 and superoxide dismutase [Cu-Zn] at *m/z* 15825 were reported from the rat brain in a MALDI-MS setting for the first time. The identification of these proteins could provide valuable *m/z* information for other

researchers that rely on MALDI-MSI to discover protein biomarkers. A full list of identified proteins and corresponding fragment ions' annotation based on high resolution top-down MS sequencing are detailed in **Supplemental Information** and **Appendix III.**

7.4 Discussion

A combined profiling and imaging strategy for biomarker discovery and screening

Our MALDI MS-based strategy combines MS profiling and imaging techniques with the goal to screen for changes in the distribution and relative abundance of proteins in the brains of rats that had experienced early developmental exposure to the NMDA antagonist MK801 on postnatal days 2 and 4. The profiling approach, which consists of manually depositing matrix on predefined substructures on a rat brain section, clearly highlights its advantages of speed, robustness and high reproducibility. Nevertheless, profiling also has the limitations that low spatial resolution is obtained for each region of interest, and molecular information of the un-profiled regions is lost. Alternatively, the imaging approach delivers high spatial resolution MS images, allowing for further investigation of structural detail of the entire rat brain section. The utilization of software ClinProTools aids in raw data processing and statistical analysis to filter the spectral features that displayed significant changes in the MK801 and vehicle treated samples. Relying on the combined profiling and imaging strategy, elevated levels of ubiquitin, thymosin-β 4, PEP-19, acyl-CoA-binding protein subunits, cytochrome c oxidase subunits, mitochondria import inner membrane translocase and adenylate cyclase type 10 were detected in the MK801-treated rat brains when compared to the vehicle-treated ones in the profiling spectra of selected

brain regions. Moreover, MALDI-MSI affirmed increased abundance of these proteins by delivering high spatial resolution and high throughput MS images directly from the rat brain sections obtained from rats subjected to MK801 treatment. On the other hand, proteins that show decreased abundance levels in the MK801 treated rats were also observed, such as ATP synthase-coupling factor 6, ProSAAS, secretogranin, calmodulin and Superoxide dismutase [Cu-Zn]. Intriguingly, the observation of these discriminating peaks in the MK801- and vehicle-treated rats might suggest involvement of the respective proteins in the pathological mechanisms of observed neurological deficits or in reparative processes. In contrast, the detection of proteins that displayed the same expression levels in the MK801 treated rats might imply that proteins from the treated subjects function normally as in the control group.

The spatial distribution of these differentially expressed proteins, determined by MS profiling and imaging techniques, corroborated with previous literature reporting massive apoptosis in the frontal and cingulate cortex following MK801 treatment of infant rats [6]. As shown in **Figure 3E** and **4E**, significant differences could be observed in frontal and cingulate cortex for both ions, indicating that NMDA receptor blockade induced by MK801 during development might have led to severe neuronal loss in this region and triggered significant response of brain tissue in the form of protein expression level changes. Selective vulnerability of brain regions to MK801-induced imbalance could be attributed to a high level of NMDA receptors on neuronal populations within those regions. [41] The significant difference observed in the Acb region could also be attributed to a high concentration of NMDA receptors in Acb within the basal ganglia. However, receptor density cannot solely account for neuronal vulnerability. The NMDA receptor density together with additional factors,

such as stage in synaptogenesis during treatment and the subunit composition of the NMDA receptor complex, combinatorially determine the degree of neuronal sensitivity. [6] This probably explains the significant changes observed in other regions, such as thalamus, cerebellum and olfactory bulb as shown in **Figure 3E** and **4E**.

MALDI-MS combined with high resolution top-down tandem MS for protein identification

The subsequent step after biomarker screening is to unequivocally identify the protein targets of interest. We employed an MSI approach in combination with top-down tandem MS identification for this purpose. In a top-down proteomics experiment, protein identification is performed through intact mass measurement followed by tandem MS sequencing with high resolution high mass accuracy FTMS-based platform. Comparing to bottom-up tandem MS sequencing, top-down tandem MS approach directly analyzes intact proteins [42] and best preserves the biological information observed by MALDI-MS. [43] It avoids the tedious mass correlation between intact protein masses observed on MSI and digested peptide masses detected by ESI-MS/MS. Moreover, it provides unique advantages in assessing protein modifications, such as PTMs and sequence variants, present *in vivo* with no prior knowledge. [44, 45] The high-end instrument employed in this study offers superior mass accuracy in MS and MS/MS, [46] which is highly advantageous for precise characterization of PTMs, and ultra high mass resolution, exemplified by the well-resolved whole protein isotopic distribution as shown in **Fig. 6A** and **6C**. Combining with a standard fractionation step, we further purified the proteins of interest and unambiguously determined their identities by top-down tandem MS sequencing. Most of the protein targets observed by MALDI-MS were verified by the top-down

approach and reliably mapped with full sequence coverage and PTMs. Other proteins were detected and identified, and are shown in the **Supplemental Information.** The innovative protein identification strategy applied in this study, by integrating the advantages of MALDI-MSI and top-down high resolution tandem MS, enables the identification of potential protein biomarkers for neuropsychiatric syndromes which result from excitatory/inhibitory neurotransmission imbalance in early development.

The ion signal at *m/z* 16792 in this study was identified as calmodulin and was also found to be N-terminally acetylated. In contrast to those upregulated proteins, MK801 treatment decreased the relative intensity of calmodulin in the rat brain compared to the vehicle-treated rats. Calmodulin is a major Ca^{2+}-binding protein in cells, where it complexes with Ca^{2+} and activates several types of intracellular enzymes, including kinases, phosphatases and adenyl cyclases [47]. In its Ca^{2+} free form, calmodulin can also regulate the function of various proteins, like actin-binding and cytoskeletal proteins, [48] making it essential in the regulation of cell structure. Failure of neuronal calcium homeostasis has been associated with cytotoxic events in the aging process and the pathogenesis of neurodegenerative diseases, [49] since Ca^{2+}/calmodulin signaling is involved in controlling processes such as muscle contraction, neurotransmitter release, transcriptional regulation, and cell death. [50, 51] More intriguingly, calmodulin has also been linked to the regulation of ion channels coupled to NMDA receptors. [47]

Several intracellular peptides that contained an IQ motif were bestowed a unique ability to bind calcium-poor calmodulin, such as neuromodulin, neurogranin and PEP-19. [47] PEP-19 can modulate Ca^{2+}/calmodulin homeostasis by altering the calcium-binding dynamics of free calmodulin and calmodulin bound to other target proteins.

Interestingly, expression level changes of PEP-19 have been linked to several neurodegenerative conditions, such as Parkinson's disease, [29] Alzheimer's disease [52] and Huntington's disease. [34] Among the biological function studies of PEP-19, a study showing that PEP-19 can inhibit apoptotic cell death due to UV irradiation or treatment with staurosporine is of special interest. [47] Hence, the upregulation of PEP-19 in the MK801 treated samples could contribute to neuronal reparative processes.

Other than Ca^{2+}/calmodulin signaling, the ubiquitin pathway might also be involved in the molecular pathology of NMDA receptor antagonist neurotoxicity in developing rat brains given the increased abundance of ubiquitin observed in the MK801-treated group. Aberrations in the ubiquitin system have been linked with neurodegenerative diseases, such as Parkinson's disease, [53] Alzheimer's disease, [52, 54] Huntington's disease, [55] Down's Syndrome [56] or progressive supranuclear palsy. [57] Recently, ubiquitin gene networks expressed within the central nervous system have been postulated to contribute to the genetic susceptibility to ASD. [58] Previous literature also suggested that accumulation of ubiquitin could potentially result in dominant inhibition of the ubiquitin–proteasome system, leading to insufficient degradation of toxic proteins with neuropathological consequences.

The abundance of thymosin beta-4 was also elevated in rat brains subjected to MK801 treatment. Thymosin beta 4 ($T\beta_4$) detected at *m/z* 4964 with N-term acetylation is a major actin monomer-sequestering molecule in mammalian cells. It has various biological functions, including regulation of inflammatory chemokines and cytokines, cellular migration, blood vessel formation and apoptosis. [59] It has been shown to prevent cell death from infections in the eye, [60] loss of blood supply in the heart after myocardial infarction [61] and toxins in nerve cells. [62] For example,

administration of thymosin beta has been shown to prevent the loss of hippocampal neurons after kainic acid treatment. [62] The significance of elevated abundance of $T\beta_4$ following MK801 treatment is unknown. It may be speculated that there is possible involvement of $T\beta_4$ in reparative processes, but its value as a potential therapeutic target ought to be evaluated in future research.

In this study, our results also revealed an intriguing increase of three cytochrome c oxidase (COX) subunits, i.e. the smallest subunit 7c at *m/z* 5486, the subunit 6c at *m/z* 8368 and the largest subunit 6b at *m/z* 9979, in the MK801 treated group, although the subunit 6c was not confirmed by top-down MS/MS spectra. The COX subunits reported here are mitochondria respiratory chain (MRC) components, involved in electron transfer to oxygen. [35] COX is known to modulate apoptosis. [63, 64] Previous studies have shown that an induction of MRC proteins like COX precede apoptosis induced by multiple stimuli. Specifically, increased expression of COX mRNAs and proteins were observed in apoptotic model systems such as Jurkat cells treated with camptothecin, [65] breast cells treated with teniposide [65] and apoptotic epithelial cancer cells induced by chemical BMD188. [66] The increased synthesis of MRC proteins, possibly representative of a more global mitochondrial activation response, is suggested to be responsible for subsequent disruption of MRC functions and ultimately cell death in response to apoptotic stimuli. [63] The up-regulation of COX subunits detected here could possibly be explained by this mechanism.

7.5 Conclusions

We have demonstrated that MALDI-MSI is capable of mapping the spatial distribution and relative abundance changes of multiple proteins simultaneously and

directly from rat brain sections. The profiling strategy is fast, robust, and highly reproducible for producing proteome "snapshots" of the anatomical regions of interest in the infant rat brain sections. Complementary to this, the MSI technology enables more detailed molecular mapping with enhanced spatial resolution while increasing instrument time as a compromise. A combination of the two techniques allowed localization and assessment of the relative abundances of the biomolecules expressed in the MK801-treated and control samples with relatively high spatial resolution and reproducibility. The subsequent high mass spectral resolution top-down tandem MS approach allowed protein identification through accurate mass measurement of intact proteins followed by complementary fragmentation for tandem MS analysis. This hybrid approach showcases the unique abilities of MSI in the discovery of potential disease biomarkers. Moreover, this study opens up new avenues for future applications in the field of biomarker discovery and for mechanistic investigations delineating pathophysiology of neurodevelopmental disorders.

Acknowledgement

This work was supported in part by University of Wisconsin grants PRJ-26CH, PRJ-23FA, PRJ-25AH and Waisman grant PRJ 233-23CK to C.I., a core grant to the Waisman Center from the National Institute of Child Health and Human Development (P30 HD03352), the National Institutes of Health grant (R01DK071801 to LL). NLK thanks support from NIH grants R01 GM067193 and P30 DA018310 and the Searle Funds at the Chicago Community Trust (to Chicago Biomedical Consortium). LL acknowledges an H. I. Romnes Faculty Research Fellowship. The authors thank Bruker Daltonics for graciously loaning the Autoflex III MALDI TOF/TOF mass spectrometer.

7.6 References

1. Kaindl, A. M.; Asimiadou, S.; Manthey, D.; Hagen, M. V. D.; Turski, L.; Ikonomidou, C. (2006) Antiepileptic drugs and the developing brain. *Cell. Mol. Life Sci.* 63, 399-413.
2. Kaindl, A. M.; Ikonomidou, C. (2007) Glutamate antagonists are neurotoxins for the developing brain. *Neurotox. Res.* 11, 203-218.
3. Gogolla, N.; LeBlanc, J. J.; Quast, K. B.; Sudhof, T. C.; Fagiolini, M.; Hensch, T. K. (2009) Common circuit defect of excitatory-inhibitory balance in mouse models of autism. *J. Neurodev. Disord.* 1, 172-181.
4. Kaindl, A. M.; Koppelstaetter, A.; Nebrich, G.; Stuwe, J.; Sifringer, M.; Zabel, C.; Klose, J.; Ikonomidou, C. (2008) Brief Alteration of NMDA or GABA(A) receptor-mediated neurotransmission has long term effects on the developing cerebral cortex. *Mol. Cell. Proteomics* 7, 2293-2310.
5. Stefovska, V. G.; Uckermann, O.; Czuczwar, M.; Smitka, M.; Czuczwar, P.; Kis, J.; Kaindl, A. M.; Turski, L.; Turski, W. A.; Ikonomidou, C. (2008) Sedative and anticonvulsant drugs suppress postnatal neurogenesis. *Ann. Neurol.* 64, 434-445.
6. Ikonomidou, C.; Bosch, F.; Miksa, M.; Bittigau, P.; Vockler, J.; Dikranian, K.; Tenkova, T. I.; Stefovska, V.; Turski, L.; Olney, J. W. (1999) Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. *Science* 283, 70-74.
7. Teng, P. N.; Hood, B. L.; Sun, M.; Dhir, R.; Conrads, T. P. (2011) Differential proteomic analysis of renal cell carcinoma tissue interstitial fluid. *J. Proteome Res.* 10, 1333-1342.
8. Lowenthal, M. S.; Mehta, A. I.; Frogale, K.; Bandle, R. W.; Araujo, R. P.; Hood, B. L.; Veenstra, T. D.; Conrads, T. P.; Goldsmith, P.; Fishman, D.; Petricoin, E. F.; Liotta, L. A. (2005) Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clin. Chem.* 51, 1933-1945.
9. Hu, S.; Arellano, M.; Boontheung, P.; Wang, J. H.; Zhou, H.; Jiang, J.; Elashoff, D.; Wei, R.; Loo, J. A.; Wong, D. T. (2008) Salivary proteomics for oral cancer biomarker discovery. *Clin. Cancer Res.* 14, 6246-6252.
10. Craig-Schapiro, R.; Perrin, R. J.; Roe, C. M.; Xiong, C. J.; Carter, D.; Cairns, N. J.; Mintun, M. A.; Peskind, E. R.; Li, G.; Galasko, D. R.; Clark, C. M.; Quinn, J. F.; D'Angelo, G.; Malone, J. P.; Townsend, R. R.; Morris, J. C.; Fagan, A. M.; Holtzman, D. M. (2010) YKL-40: a Novel prognostic fluid biomarker for preclinical Alzheimer's disease. *Biol. Psychiatry.* 68, 903-912.
11. Flint, M. S.; Hood, B. L.; Sun, M.; Stewart, N. A.; Jones-Laughner, J.; Conrads, T. P. (2010) Proteomic analysis of the murine liver in response to a combined exposure to psychological stress and 7,12-Dimethylbenz(a)anthracene. *J. Proteome Res.* 9, 509-520.
12. Ruepp, S. U.; Tonge, R. P.; Shaw, J.; Wallis, N.; Pognan, F. (2002) Genomics and proteomics analysis of acetaminophen toxicity in mouse liver. *Toxicol. Sci.* 65, 135-150.
13. Ishimura, R.; Ohsako, S.; Kawakami, T.; Sakaue, M.; Aoki, Y.; Tohyama, C. (2002) Altered protein profile and possible hypoxia in the placenta of 2,3,7,8-tetrachlorodibenzo-p-dioxin-exposed rats. *Toxicol. Appl. Pharmacol.* 185, 197-206.
14. Li, T. A.; Xu, S. L.; Oses-Prieto, J. A.; Putil, S.; Xu, P.; Wang, R. J.; Li, K. H.; Maltby, D. A.; An, L. H.; Burlingame, A. L.; Deng, Z. P.; Wang, Z. Y. (2011) Proteomics analysis reveals post-translational mechanisms for cold-induced metabolic changes in Arabidopsis. *Mol. Plant* 4, 361-374.
15. MacAleese, L.; Stauber, J.; Heeren, R. M. (2009) Perspectives for imaging mass spectrometry in the proteomics landscape. *Proteomics* 9, 819-34.
16. Seeley, E. H.; Caprioli, R. M. (2008) Molecular imaging of proteins in tissues by mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 105, 18126-31.
17. Caprioli, R. M.; Farmer, T. B.; Gile, J. (1997) Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* 69, 4751-4760.
18. Chaurand, P.; Schwartz, S. A.; Caprioli, R. M. (2002) Imaging mass spectrometry: a new

tool to investigate the spatial organization of peptides and proteins in mammalian tissue sections. *Curr. Opin. Chem. Biol.* 6, 676-681.
19. Lagarrigue, M.; Becker, M.; Lavigne, R.; Deininger, S. O.; Walch, A.; Aubry, F.; Suckau, D.; Pineau, C. (2011) Revisiting rat spermatogenesis with MALDI imaging at 20-µm resolution. *Mol. Cell. Proteomics* 10, M110.005991.
20. Cazares, L. H.; Troyer, D.; Mendrinos, S.; Lance, R. A.; Nyalwidhe, J. O.; Beydoun, H. A.; Clements, M. A.; Drake, R. R.; Semmes, O. J. (2009) Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 2 discriminates cancer from uninvolved prostate tissue. *Clin. Cancer Res.* 15, 5541-51.
21. Deininger, S. O.; Ebert, M. P.; Futterer, A.; Gerhard, M.; Rocken, C. (2008) MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *J. Proteome Re.s* 7, 5230-6.
22. Blaze, M. T. M.; Aydin, B.; Carlson, R. P.; Hanley, L. (2012) Identification and imaging of peptides and proteins on *Enterococcus faecalis* biofilms by matrix assisted laser desorption ionization mass spectrometry. *Analyst* 137, 5018-5025.
23. Murphy, R. C.; Hankin, J. A.; Barkley, R. M. (2009) Imaging of lipid species by MALDI mass spectrometry. *J. Lipid Res.* 50, S317-S322.
24. Chughtai, K.; Jiang, L.; Greenwood, T. R.; Glunde, K.; Heeren, R. M. (2012) Mass spectrometry images acylcarnitines, phosphatidylcholines, and sphingomyelin in MDA-MB-231 breast tumor models. *J. Lipid Res.* 54, 333-44.
25. Chen, R.; Hui, L.; Sturm, R. M.; Li, L. (2009) Three dimensional mapping of neuropeptides and lipids in crustacean brain by mass spectral imaging. *J. Am. Soc. Mass Spectrom.* 20, 1068-77.
26. DeKeyser, S. S., Kutz-Naber, K.K., Schmidt, J.J., Barrett-Wilt, G.A.and Li, L. (2007) Mass spectral imaging of neuropeptides in crustacean nervous tissue by MALDI TOF/TOF. *J. Proteome Res.* 6, 1782-1791.
27. Pirman, D. A.; Reich, R. F.; Kiss, A.; Heeren, R. M.; Yost, R. A. (2012) Quantitative MALDI tandem mass spectrometric imaging of cocaine from brain tissue with a deuterated internal standard. *Anal. Chem.* 85, 1081-9.
28. Minerva, L.; Clerens, S.; Baggerman, G.; Arckens, L. (2008) Direct profiling and identification of peptide expression differences in the pancreas of control and ob/ob mice by imaging mass spectrometry. *Proteomics* 8, 3763-3774.
29. Sköld, K.; Svensson, M.; Nilsson, A.; Zhang, X. Q.; Nydahl, K.; Caprioli, R. M.; Svenningsson, P.; Andren, P. E. (2006) Decreased striatal levels of PEP-19 following MPTP lesion in the mouse. *J. Proteome Res.* 5, 262-269.
30. Hanrieder, J.; Ljungdahl, A.; Falth, M.; Mammo, S. E.; Bergquist, J.; Andersson, M. (2011) L-DOPA-induced dyskinesia is associated with regional increase of striatal dynorphin peptides as elucidated by imaging mass spectrometry. *Mol. Cell. Proteomics* 10, M111.009308.
31. Lagarrigue, M.; Alexandrov, T.; Dieuset, G.; Perrin, A.; Lavigne, R.; Baulac, S.; Thiele, H.; Martin, B.; Pineau, C. (2012) New analysis workflow for MALDI imaging mass spectrometry: application to the discovery and identification of potential markers of childhood absence epilepsy. *J. Proteome Res.* 11, 5453-5463.
32. Meistermann, H.; Norris, J. L.; Aerni, H. R.; Cornett, D. S.; Friedlein, A.; Erskine, A. R.; Augustin, A.; De Vera Mudry, M. C.; Ruepp, S.; Suter, L.; Langen, H.; Caprioli, R. M.; Ducret, A. (2006) Biomarker discovery by imaging mass spectrometry: transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat. *Mol. Cell. Proteomics* 5, 1876-86.
33. Li, H. H.; Hummon, A. B. (2011) Imaging mass spectrometry of three-dimensional cell culture systems. *Anal. Chem.* 83, 8794-8801.
34. Utal, A. K.; Stopka, A. L.; Roy, M.; Coleman, P. D. (1999) PEP-19 immunohistochemistry defines the basal ganglia and associated structures in the adult human

brain, and is dramatically reduced in Huntington's disease. *Neuroscience* 88, 993-993.
35. Elsner, M.; Rauser, S.; Maier, S.; Schone, C.; Balluff, B.; Meding, S.; Jung, G.; Nipp, M.; Sarioglu, H.; Maccarrone, G.; Aichler, M.; Feuchtinger, A.; Langer, R.; Jutting, U.; Feith, M.; Kuster, B.; Ueffing, M.; Zitzelsberger, H.; Hofler, H.; Walch, A. (2012) MALDI imaging mass spectrometry reveals COX7A2, TAGLN2 and S100-A10 as novel prognostic markers in Barrett's adenocarcinoma. *J. Proteomics* doi:10.1016/j.jprot.2012.02.012.
36. Balluff, B.; Rauser, S.; Meding, S.; Elsner, M.; Schone, C.; Feuchtinger, A.; Schuhmacher, C.; Novotny, A.; Jutting, U.; Maccarrone, G.; Sarioglu, H.; Ueffing, M.; Braselmann, H.; Zitzelsberger, H.; Schmid, R. M.; Hofler, H.; Ebert, M. P.; Walch, A. (2011) MALDI imaging identifies prognostic seven protein signature of novel tissue markers in intestinal-type gastric cancer. *Am. J. Pathol.* 179, 2720-2729.
37. Delbosc, S.; Haloui, M.; Louedec, L.; Dupuis, M.; Cubizolles, M.; Podust, V. N.; Fung, E. T.; Michel, J. B.; Meilhac, O. (2008) Proteomic analysis permits the identification of new biomarkers of arterial wall remodeling in hypertension. *Mol. Med.* 14, 383-94.
38. Hardesty, W. M.; Kelley, M. C.; Mi, D.; Low, R. L.; Caprioli, R. M. (2011) Protein signatures for survival and recurrence in metastatic melanoma. *J. Proteomics* 74, 1002-14.
39. Bora, A.; Annangudi, S. P.; Millet, L. J.; Rubakhin, S. S.; Forbes, A. J.; Kelleher, N. L.; Gillette, M. U.; Sweedler, J. V. (2008) Neuropeptidomics of the supraoptic rat nucleus. *J. Proteome Res.* 7, 4992-5003.
40. Pierson, J.; Norris, J. L.; Aerni, H. R.; Svenningsson, P.; Caprioli, R. M.; Andren, P. E. (2004) Molecular profiling of experimental Parkinson's disease: direct analysis of peptides and proteins on brain tissue sections by MALDI mass spectrometry. *J. Proteome Res.* 3, 289-295.
41. Monaghan, D. T.; Cotman, C. W. (1985) Distribution of N-methyl-D-aspartate-sensitive L-[H-3]glutamate-binding sites in rat brain. *J. Neurosci.* 5, 2909-2919.
42. Kelleher, N. L. (2004) Top-down proteomics. *Anal Chem* 76, 196a-203a.
43. Sampson, J. S.; Murray, K. K.; Muddiman, D. C. (2009) Intact and top-down characterization of biomolecules and direct analysis using infrared matrix-assisted laser desorption electrospray ionization coupled to FT-ICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* 20, 667-673.
44. Ferguson, J. T.; Wenger, C. D.; Metcalf, W. W.; Kelleher, N. L. (2009) Top-down proteomics reveals novel protein forms expressed in *Methanosarcina acetivorans*. *J. Am. Soc. Mass Spectrom.* 20, 1743-1750.
45. Zhang, J. A.; Dong, X. T.; Hacker, T. A.; Ge, Y. (2010) Deciphering modifications in swine cardiac troponin I by top-down high-resolution tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* 21, 940-948.
46. Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Muller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J. J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. (2012) Ultra high resolution linear ion trap orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Mol. Cell. Proteomics* 11, O111.013698.
47. Johanson, R. A.; Sarau, H. M.; Foley, J. J.; Slemmon, J. R. (2000) Calmodulin-binding peptide PEP-19 modulates activation of calmodulin kinase II In situ. *J Neurosci* 20, 2860-6.
48. Beckingham, K.; Lu, A. Q.; Andruss, B. F. (1998) Calcium-binding proteins and development. *Biometals* 11, 359-73.
49. Pereira, C. F.; Oliveira, C. R. (2000) Oxidative glutamate toxicity involves mitochondrial dysfunction and perturbation of intracellular Ca^{2+} homeostasis. *Neurosci. Res.* 37, 227-36.
50. Berridge, M. J.; Bootman, M. D.; Roderick, H. L. (2003) Calcium signalling: dynamics, homeostasis and remodelling. *Nat. Rev. Mol. Cell. Biol.* 4, 517-529.
51. Monteith, G. R.; McAndrew, D.; Faddy, H. M.; Roberts-Thomson, S. J. (2007) Calcium and cancer: targeting Ca2+ transport. *Nat. Rev. Cancer.* 7, 519-530.
52. Slemmon, J. R.; Hughes, C. M.; Campbell, G. A.; Flood, D. G. (1994) Increased levels of

hemoglobin-derived and other peptides in Alzheimers's disease cerebellum. *J. Neurosci.* 14, 2225-2235.
53. McNaught, K. S.; Olanow, C. W.; Halliwell, B.; Isacson, O.; Jenner, P. (2001) Failure of the ubiquitin-proteasome system in Parkinson's disease. *Nat. Rev. Neurosci.* 2, 589-94.
54. Lam, Y. A.; Pickart, C. M.; Alban, A.; Landon, M.; Jamieson, C.; Ramage, R.; Mayer, R. J.; Layfield, R. (2000) Inhibition of the ubiquitin-proteasome system in Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A*. 97, 9902-6.
55. Bennett, E. J.; Shaler, T. A.; Woodman, B.; Ryu, K. Y.; Zaitseva, T. S.; Becker, C. H.; Bates, G. P.; Schulman, H.; Kopito, R. R. (2007) Global changes to the ubiquitin system in Huntington's disease. *Nature* 448, 704-8.
56. van Leeuwen, F. W.; de Kleijn, D. P.; van den Hurk, H. H.; Neubauer, A.; Sonnemans, M. A.; Sluijs, J. A.; Koycu, S.; Ramdjielal, R. D.; Salehi, A.; Martens, G. J.; Grosveld, F. G.; Peter, J.; Burbach, H.; Hol, E. M. (1998) Frameshift mutants of beta amyloid precursor protein and ubiquitin-B in Alzheimer's and Down patients. *Science* 279, 242-7.
57. Fergusson, J.; Landon, M.; Lowe, J.; Ward, L.; van Leeuwen, F. W.; Mayer, R. J. (2000) Neurofibrillary tangles in progressive supranuclear palsy brains exhibit immunoreactivity to frameshift mutant ubiquitin-B protein. *Neurosci. Lett.* 279, 69-72.
58. Glessner, J. T.; Wang, K.; Cai, G.; Korvatska, O.; Kim, C. E.; Wood, S.; Zhang, H.; Estes, A.; Brune, C. W.; Bradfield, J. P.; Imielinski, M.; Frackelton, E. C.; Reichert, J.; Crawford, E. L.; Munson, J.; Sleiman, P. M.; Chiavacci, R.; Annaiah, K.; Thomas, K.; Hou, C.; Glaberson, W.; Flory, J.; Otieno, F.; Garris, M.; Soorya, L.; Klei, L.; Piven, J.; Meyer, K. J.; Anagnostou, E.; Sakurai, T.; Game, R. M.; Rudd, D. S.; Zurawiecki, D.; McDougle, C. J.; Davis, L. K.; Miller, J.; Posey, D. J.; Michaels, S.; Kolevzon, A.; Silverman, J. M.; Bernier, R.; Levy, S. E.; Schultz, R. T.; Dawson, G.; Owley, T.; McMahon, W. M.; Wassink, T. H.; Sweeney, J. A.; Nurnberger, J. I.; Coon, H.; Sutcliffe, J. S.; Minshew, N. J.; Grant, S. F.; Bucan, M.; Cook, E. H.; Buxbaum, J. D.; Devlin, B.; Schellenberg, G. D.; Hakonarson, H. (2009) Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 459, 569-73.
59. Philp, D.; Kleinman, H. K. (2010) Animal studies with thymosin beta 4, a multifunctional tissue repair and regeneration peptide. *Ann. N. Y .Acad. Sci.* 1194, 81-86.
60. Sosne, G.; Siddiqi, A.; Kurpakus-Wheater, M. (2004) Thymosin-beta4 inhibits corneal epithelial cell apoptosis after ethanol exposure in vitro. *Invest. Ophthalmol. Vis. Sci.* 45, 1095-100.
61. Badamchian, M.; Fagarasan, M. O.; Danner, R. L.; Suffredini, A. F.; Damavandy, H.; Goldstein, A. L. (2003) Thymosin beta(4) reduces lethality and down-regulates inflammatory mediators in endotoxin-induced septic shock. *Int. Immunopharmacol.* 3, 1225-1233.
62. Popoli, P.; Pepponi, R.; Martire, A.; Armida, M.; Pezzola, A.; Galluzzo, M.; Domenici, M. R.; Potenza, R. L.; Tebano, M. T.; Mollinari, C.; Merlo, D.; Garaci, E. (2007) Neuroprotective effects of thymosin beta 4 in experimental models of excitotoxicity. *Ann. N. Y. Acad. Sci.* 1112, 219-224.
63. Chandra, D.; Liu, J. W.; Tang, D. G. (2002) Early mitochondrial activation and cytochrome c up-regulation during apoptosis. *J Biol Chem* 277, 50842-54.
64. Huttemann, M.; Lee, I.; Grossman, L. I.; Doan, J. W.; Sanderson, T. H. (2012) Phosphorylation of mammalian cytochrome c and cytochrome c oxidase in the regulation of cell destiny: respiration, apoptosis, and human disease. *Adv. Exp. Med. Biol.* 748, 237-64.
65. Sanchez-Alcazar, J. A.; Khodjakov, A.; Schneider, E. (2001) Anticancer drugs induce increased mitochondrial cytochrome c expression that precedes cell death. *Cancer Res.* 61, 1038-44.
66. Joshi, B.; Li, L.; Taffe, B. G.; Zhu, Z.; Wahl, S.; Tian, H.; Ben-Josef, E.; Taylor, J. D.; Porter, A. T.; Tang, D. G. (1999) Apoptosis induction by a novel anti-prostate cancer compound, BMD188 (a fatty acid-containing hydroxamic acid), requires the mitochondrial respiratory chain. *Cancer Res.* 59, 4343-55.

Figures

Figure 1. An overall strategy for mapping proteins in brain tissues by combining MS profiling and imaging. The animals treated with MK801 were sacrificed and handled pairwise (A). The brains were removed and embedded in gelatin (B) followed by sectioning (C). Four sections from each brain were used and 14 regions as annotated in (C) from each section were profiled (D). An entire section was imaged from one brain (E). Four pairs of brains (vehicle *vs.* MK801) were profiled and imaged (N=4), respectively, on a MALDI-TOF/TOF instrument (Bruker autoflex III) (F). The profiling spectra were compared pairwise for each specific region (G), and the imaging data were compared as well (H).

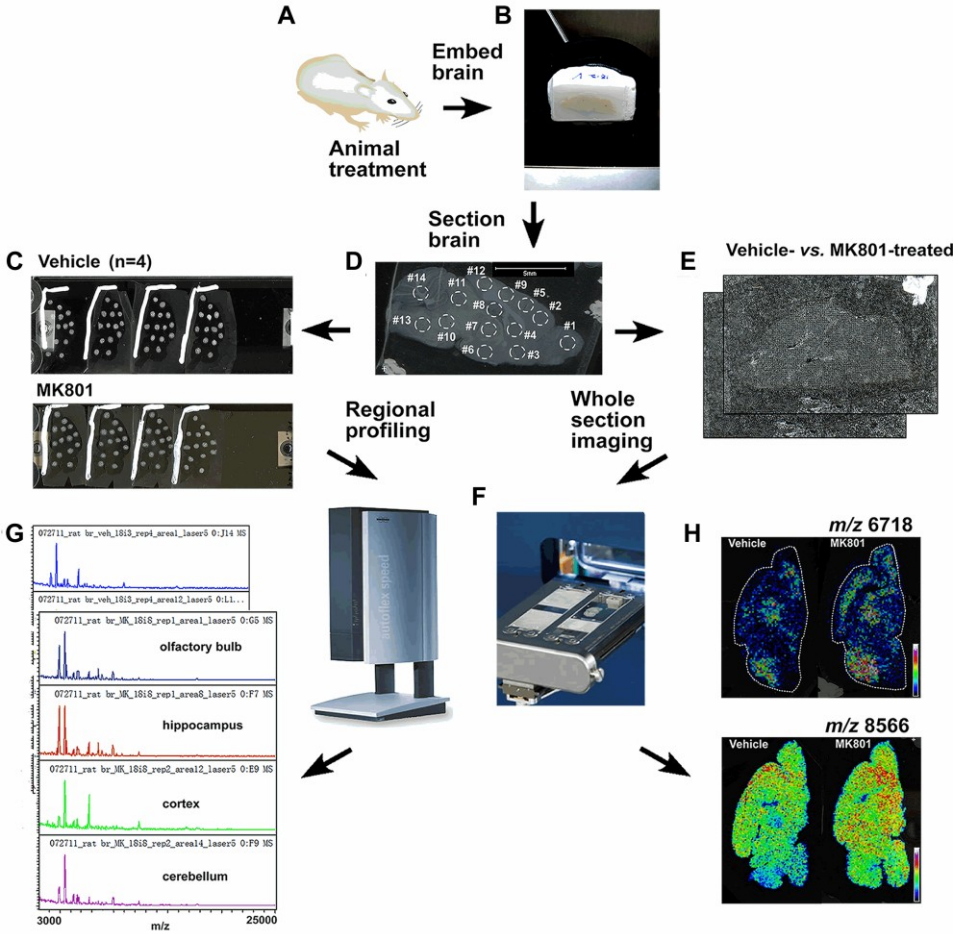


Figure 2. Regional profiling of the rat brain sections treated with vehicle and MK801. (A) Representative overall profiling spectra averaged from 14 regions of 4 serial sections from each animal (N=4, two classes, i.e. vehicle-treated and MK801-treated rats as displayed in green and red, respectively). (B)-(E) are the zoomed-in spectra of several ions displaying differential expression levels in the two groups. The blue shades indicate the peaks picked by ClinProTools (signal to noise threshold higher than 5 in the average spectra) and included for statistical comparison between the two classes.

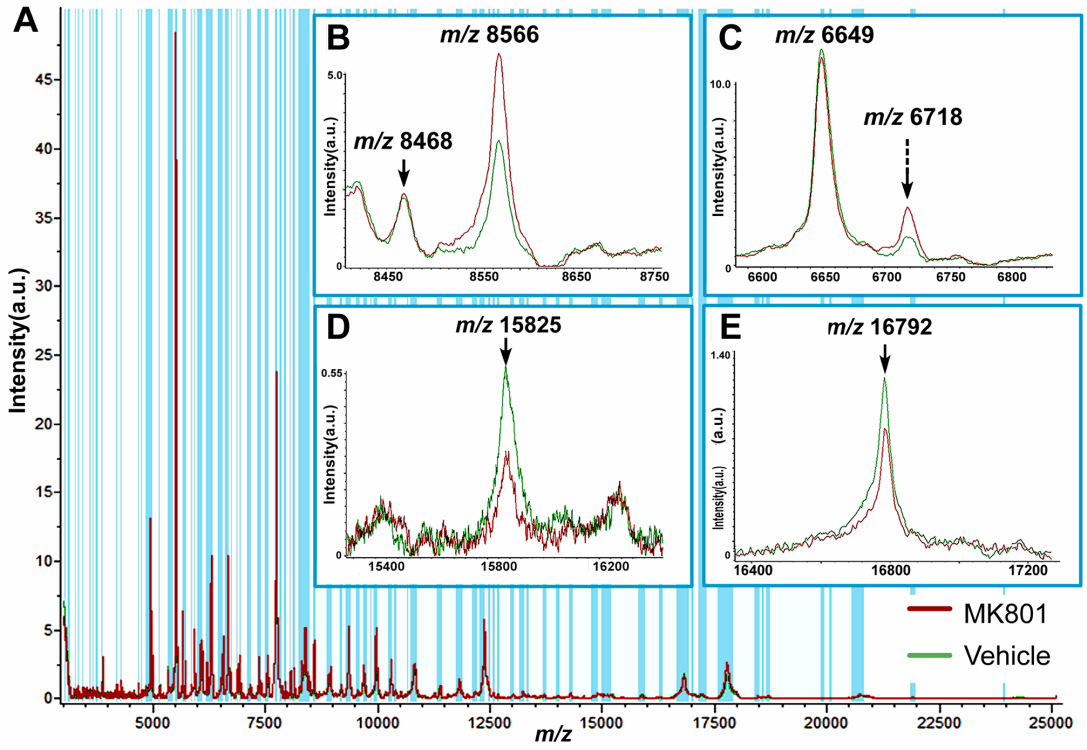


Figure 3. Elevated level of the *m/z* 6718 ion quantified by the strategy combining MS profiling and imaging. A) The ratios of the intensities of the ion at *m/z* 6718 in the profiling spectra of specific regions, regions #1, #2, #5, #9, #7 and #14, where the ion was concentrated. The averaged intensities and p-values were calculated by ClinProTools and manually confirmed. ＊, $p < 0.05$; ＊＊, $p < 0.01$, unpaired student's t-test. (B) An optical image of a rat brain section washed after MSI experiments. (C) The *m/z* 6718 ion was up-regulated in the averaged MSI spectrum of the olfactory bulbs, ROI 1 and ROI2, of the rat brain section treated with vehicle and MK801 as shown in (E). (D) A loading plot of ROI 1 and ROI 2 analyzed by ClinProTools. The *m/z* 6718 ion was shown to contribute significantly in differentiating the two ROIs in Load 2. (E) MS images of the ions at *m/z* 6718 for the vehicle and the MK801-treated rat brain sections. The *m/z* 6718 ion was more concentrated in the olfactory bulb, cortex, thalamus (TH) and cerebellum regions in the MK801-treated section compared to the vehicle-treated one.

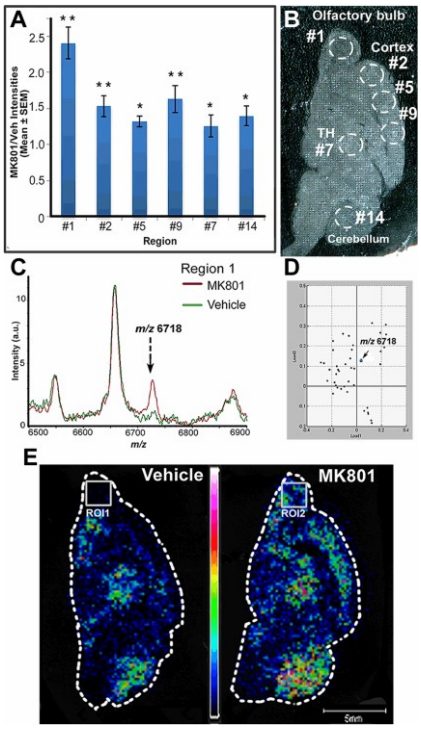


Figure 4. Elevated level of the *m/z* 8566 ion quantified by the strategy combining MS profiling and imaging. A) The ratios of the intensities of the ion at *m/z* 8566 in the profiling spectra of specific regions, regions #2, #5, #3, #6, #7 and #14, where the ion was concentrated. ＊, p＜ 0.05; ＊＊, p＜0.01, unpaired student's t-test. (B) An optical image of a rat brain section washed after MSI experiments. (C) The *m/z* 8566 ion was up-regulated in the averaged MSI spectrum of the thalamus (TH) regions, ROI 1 and ROI2, of the rat brain section treated with vehicle and MK801 as shown in (E). (D) A loading plot of ROI 1 and ROI 2 analyzed by ClinProTools. The *m/z* 8566 ion was shown to contribute significantly in differentiating the two ROIs in Load 1. (E) MS images of the ions at *m/z* 8566 for the vehicle and the MK801-treated rat brain sections. The *m/z* 8566 ion was more concentrated in the frontal cortex, accumbens nucleus (Acb), thalamus (TH) and cerebellum regions in the MK801-treated section compared to the vehicle-treated one.

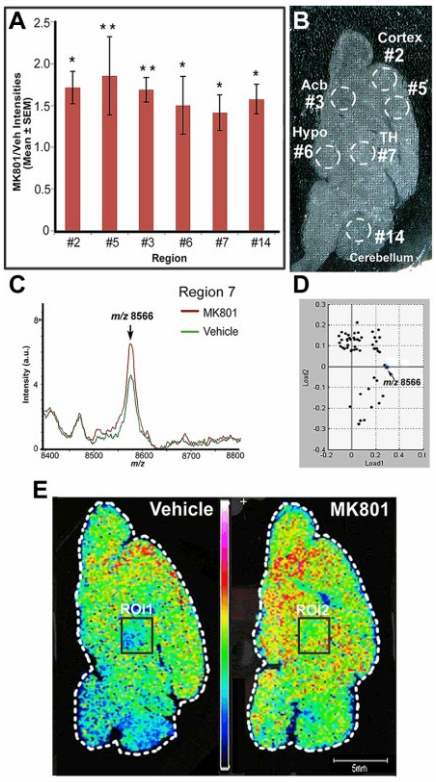


Figure 5. The workflow to identify the proteins that displayed changes in the MK801-treated samples. (A) MALDI-MS screening spectra of the HPLC fractions that contained the concentrated proteins of interest at *m/z* 6718 and 8566. (B) Representative MS spectrum acquired from a HPLC fraction containing the *m/z* 6718 ion on an ESI-Orbitrap Elite system, which provides ultra-high mass resolution and accuracy. The insert is the deconvoluted spectrum averaged from several multiply charged ion forms as highlighted in blue and red. (C) HCD, CID and ETD tandem MS fragmentation spectra of the particular ion at *m/z* 1120.05 (z=6) highlighted in red.

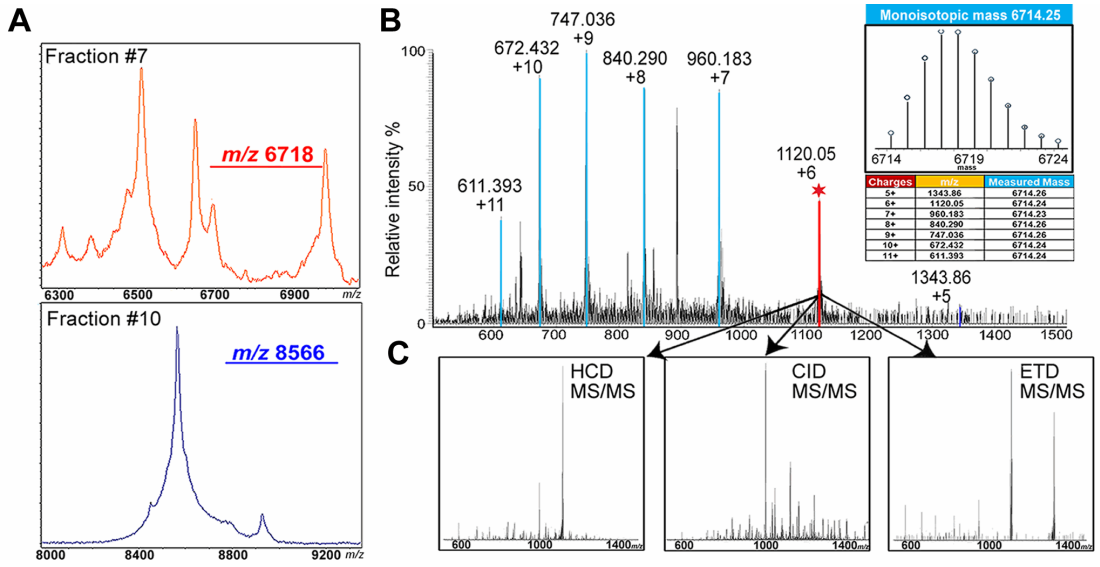
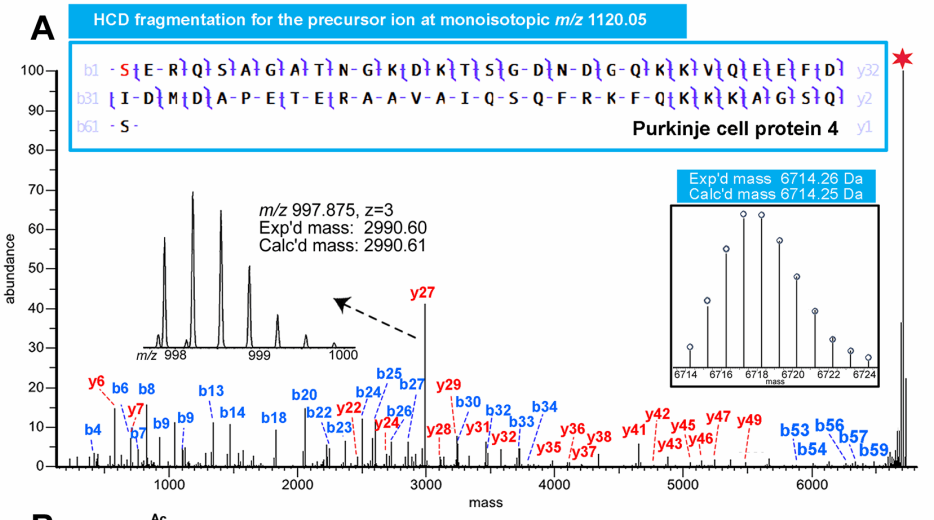


Figure 6. Fragmentation spectra and maps of the *m/z* 6718 and 8566 ions by HCD, CID and ETD. (A) HCD fragmentation spectrum of the *m/z* 6718 ion at monoisotopic *m/z* 1120.05 (z=6). The ion at *m/z* 1120.05 (z=6) was assigned as PEP-19 based on accurate mass and fragmentation pattern obtained on Orbitrap Elite. The boxed insert on the right is the deconvoluted spectrum of the *m/z* 1120.05 (z=6) ion. The original HCD tandem MS spectrum containing the y_{27} fragment ion is also zoomed-in as an insert on the left. (B) Fragmentation map of PEP-19 by three fragmentation techniques, HCD, CID and ETD. (C) HCD fragmentation spectrum of the *m/z* 8566 ion at monoisotopic *m/z* 1070.96 (z=8). The ion at *m/z* 1070.96 (z=8) was assigned as ubiquitin based on accurate mass and fragmentation pattern obtained on Orbitrap Elite. The boxed insert on the right is the deconvoluted spectrum of the *m/z* 1070.96 (z=8) ion. The original HCD tandem MS spectrum containing the y_{40} ion is also enlarged as an insert on the left. (D) Fragmentation map of Ubiquitin by HCD, CID and ETD. The number of the fragment ions produced by each of the fragmentation techniques is listed below the sequence. (E) An illustration of how the fragment ions of HCD (b, y ions), CID (b, y ions) and ETD (c, z ions) are produced and annotated in the fragmentation map is shown in (B) and (D).

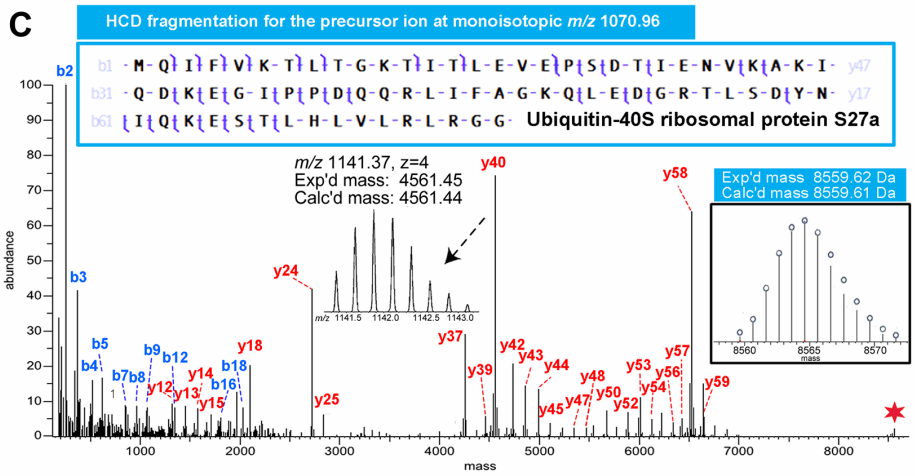


B

Ac

1	S E R Q S A G A T N G K D K T S G D N D	42
21	G Q K K V Q E E F D I D M D A P E T E R	22
41	A A V A I Q S Q F R K F Q K K K A G S Q	2
61	S	1

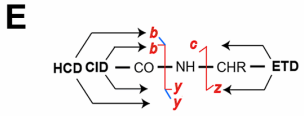
HCD: 19 b ions, 46 y ions；CID: 37 b ions, 42 y ions；ETD: 15 c ions, 12 z ions



D

1	M Q I F V K T L T G K T I T L E V E P S	57
21	D T I E N V K A K I Q D K E G I P P D Q	37
41	Q R L I F A G K Q L E D G R T L S D Y N	17
61	I Q K E S T L H L V L R L R G G	1

HCD: 38 b ions, 36 y ions；CID: 31 b ions, 37 y ions；ETD: 17 c ions, 6 z ions



Supplemental Information

Tables

Table 1. List of differentially-expressed proteins in rat brain sections observed on MALDI-TOF/TOF and identified by mass matching. The proteins that were further identified by top-down MS/MS were highlighted in red and annotated with UniProt entry, accession number, calculated mass based on the *m/z* observed on Orbitrap Elite, expected mass based on the matched protein sequence and the mass difference between the calculated mass and expected mass. All MALDI-MS profiling spectra have been processed by ClinProTools and summarized here by "peak statistic" function. PTTA represents p-value of t-test (2 classes); Ratio represents the relative abundance changes of these proteins in the brain sections of rats treated with MK801 and vehicle, calculated based on the average intensity of the ion observed in the MK801-treated class over the vehicle-treated class.

m/z on MALDI-MS	Entry	Accession	Protein Identity	Calc'd Mass (Da)	Exp'd Mass (Da)	Δ Mass (ppm)	Average Mass (Da)	PTTA	Ratio (MK/Veh)
6718	PCP4_RAT	P63055	PEP-19	6714.25	6714.26	-1	6718.19	< 0.001	1.8
11367			Unknown					< 0.001	1.8
8566	RS27A_RAT	P62982	Ubiquitin	8559.61	8559.62	-1	8564.76	< 0.001	1.6
5634			Acyl-CoA-binding protein					< 0.001	1.5
4964	TYB10_RAT	P63312	Thymosin β-4	4960.48	4960.49	-2	4963.45	< 0.001	1.4
9939	ACBP_RAT	P11030	Acyl-CoA-binding protein	9932.12	9932.12	0	9938.19	< 0.001	1.4
9979	D3ZD09_RAT	D3ZD09	Cytochrome c oxidase subunit 6b	9971.80	9971.82	-2	9978.17	< 0.001	1.4

5486	B2RYT3_RAT	B2RYT3	Cytochrome c oxidase subunit 7c	5481.87	5481.87	0	5485.40	< 0.001	1.3
8368	CX6C2_RAT	P11951	Cytochrome c oxidase subunit 6C					< 0.001	1.2
10283	TIM9_RAT	Q9WV97	**Mitochondria import inner membrane translocase**	10276.08	10276.09	-1	10282.66	< 0.05	1.2
9193	TIM8B_RAT	P62078	**Adenylate cyclase type 10**	9187.44	9187.45	-1	9193.22	< 0.01	1.2
9727	ACBP_RAT	P11030	Acyl-CoA-binding protein					< 0.001	1.1
8040	UCRI_RAT	P20788	**Cytochrome b-c1 complex subunit**	8035.38	8035.38	0	8040.30	< 0.05	1.1
4936	TYB10_RAT	P63312	Thymosin beta-10	4933.51	4933.52	-2	4936.47	< 0.05	1.1
3430	ATP5J_RAT	P21571	**ATP synthase-coupling factor 6**	3426.89	3426.90	-3	3428.94	< 0.05	0.9
16792	CALM_RAT	P62161	Calmodulin	16779.82	16779.81	1	16790.32	< 0.01	0.9
3717	Q5U2U9_RAT	Q5U2U9	**CCR4-NOT transcription complex, subunit 8**	3713.87	3713.88	-3	3716.25	< 0.01	0.9
3891	PCSK1_RAT	Q9QXU9	**ProSAAS**	3888.04	3888.04	0	3890.32	< 0.05	0.9
4800	SCG2_RAT	P10362	**Secretogranin-2**	4796.37	4796.37	0	4799.25	< 0.01	0.9
4854	VGF_RAT	P20156	**Neurosecretory protein VGF**	4850.41	4850.41	0	4853.24	< 0.05	0.9
15825	SODC_RAT	P07632	**Superoxide dismutase [Cu-Zn]**	15810.82	15810.75	4	15820.33	< 0.05	0.8
11348			Unknown					< 0.001	0.6

[a] Protein sequences in red have been verified by top-down tandem mass spectrometry. The rest of the identity assignments are solely based on mass matching. The proteins that can be identified by neither approach were shown as unknown.
[b] The bolded sequences represent the proteins that were reported from rat brain in a MALDI-MS-based platform for the first time.

Table 2. List of proteins that displayed consistent abundance levels in MK801-treated and vehicle-treated rat brain sections via MALDI-MS. These proteins were all identified by means of top-down tandem MS sequencing on a nanoLC-ESI-LTQ-Orbitrap Elite system.

m/z on MALDI-MS	Entry	Accession	Protein Identity	Calc'd Mass (Da)	Exp'd Mass (Da)	Δ Mass (ppm)	Average Mass (Da)	PTM	E value
3327	VIP_RAT	P01283	VIP peptides	3323.75	3323.76	-3	3325.80	C-term amidation	8.3E-14
3437	7B2_RAT	P27682	Neuroendocrine protein 7B 2	3433.78	3433.79	-3	3435.87		7.2E-39
3464	NPY_RAT	P07808	Pro-Neuropeptide Y	3460.66	3460.66	0	3462.82		1.6E-18
3593	7B2_RAT	P27682	Neuroendocrine protein 7B 2	3589.89	3589.89	0	3592.06		2.5E-60
3653	SCG2_RAT	P10362	Secretogranin 2	3649.80	3649.80	0	3651.94		3.8E-67
3676	VGF_RAT	P20156	Neurosecretory protein VGF	3672.77	3672.78	-3	3674.90		8.0E-08
3871	VGF_RAT	P20156	Neurosecretory protein VGF	3867.86	3867.87	-3	3870.12		2.9E-37
4273	NPY_RAT	P07808	Proneuropeptide Y	4269.08	4269.08	0	4271.69	C-term amidation	3.6E-42
4331	NPY_RAT	P07808	Proneuropeptide Y	4327.08	4327.09	-2	4329.73		1.6E-42
4522	TYB10_RAT	P63312	Thymosin beta-10	4517.33	4517.33	0	4520.04	N-term acetylation	2.2E-70
4738	TYB10_RAT	P63312	Thymosin beta-10	4733.40	4733.41	-2	4736.23	N-term acetylation	1.7E-89
4871	SCG2_RAT	P10362	Secretogranin-2	4867.40	4867.41	-2	4870.33		1.2E-77
7114	FILUV9_RAT	FILUV9	Uncharacterized	7108.52	7108.51	1	7112.62		4.9E-42
8468	UBB_RAT	P0CG51	Polyubiquilin	8461.56	8461.57	-1	8466.65	Oxidation	2.9E-35
8928	ATP5J_RAT	P21571	ATP synthase coupling factor 6	8921.55	8921.55	0	8927.06		1.7E-52
10918	DLRB1_RAT	P62628	Dynein light chain roadblock-type 1	10909.71	10909.72	-1	10916.41	N-term acetylation, oxidation	6.2E-36
12204	ARP19_RAT	Q712U5	cAMP-regulated phosphorylation 19	12196.24	12196.26	-2	12203.67	N-term acetylation	2.3E-13
14428	Q6MGC4_RAT	Q6MGC4	H2-K region expressed gene 2, rat orthologue	14417.76	14417.73	2	14426.26	N-term acetylation, water adduct	3.1E-24

Chapter 8

Conclusions and Future Directions

Conclusions and Future Directions

8.1 Conclusions

In this dissertation, I constructed a MSI-based platform that enables investigation of different types of molecules including metabolites and neurotransmitters, neuropeptides and neuroproteins with spatial and chemical information. The novel MSI platform encompasses the following elements: (1) new methods to improve the detection of low MW species using MALDI-MSI-based approach were described in Chapter 3 and 4; (2) high spatial resolution MSI of neuropeptides at the network level and comparison of MRMS and HRMS for MSI applications were described in Chapter 5; and (3) an integrated use of top-down proteomics with MALDI-MSI complements MSI-based characterization of protein features from complex samples with accurate structural elucidation enabled by HRMS/MS in Chapter 7. Although the analytical methods developed in this study are mostly intended to advance current MALDI-MSI platform for tissue analysis, new tools are also developed for biochemical analysis of liquid-phase extract, such as semi-automated *de novo* sequencing strategy and dimethylation-assisted fragmentation in Chapter 6.

In addition to the impact on current methodology of analytical biochemistry, this work also reveal novel functional information gained from the studies conducted. Specifically, Chapter 3 presents a novel HRMSI-based platform to investigate the distribution of metabolites and neurotransmitters (NTs) unambiguously from biologically complex and important specimen. It demonstrates the advantages of HRMSI when identifying analytes of interest in comparison to a MRMS instrument. In this work, we

successfully generated MS images of a myriad of metabolites and NTs simultaneously in rat and crustacean brain specimen, such as neurotransmitter acetylcholine (Ach), amino acids, nucleotides, the spatial information of which are usually challenging to attain. The novel knowledge of identified metabolites and NTs' distributions generated via HRMSI in this study will provide insight into related neuroscience research and present a powerful approach for future studies of disease that are associated with abnormal levels of metabolites and NTs.[1-3]

In Chapter 4, we further demonstrated the benefits of using MALDI-MSI in obtaining unique and valuable information on the identity and spatial distribution of plant metabolites that cannot be accomplished otherwise. A novel aspect of this work is the first use of ionless matrix DMAN for MSI applications, which delivers no matrix ions in the negative mode. Therefore, the conventional matrix DHB and the novel matrix DMAN were complementary in profiling and imaging metabolites directly from tissue sections when both positive and negative modes were employed. Furthermore, MS/MS experiments on putative metabolites were performed to confirm the assignments of region-specific metabolites. In addition, new knowledge was obtained regarding the putative roles of key metabolites involved in symbiotic nitrogen-fixation through the comparison of metabolite profiles and molecular ion images obtained from nitrogen fixing and non-fixing nodules with MALDI-MSI.

Chapter 5 details the application of MSI technology to neuropeptidomics analysis. In this chapter, we described the first documentation of neuropeptide distribution in *C. sapidus* STNS at the network and cellular resolution using an MRMS MALDI-TOF/TOF

and an HRMS MALDI-FT-ICR. A multi-faceted MS strategy that combined direct tissue analysis, *in situ* MS/MS and micro-separation coupled to ESI-Q-TOF tandem MS peptide sequencing was included in this toolkit for confident and unequivocal mapping of endogenous neuropeptides, enabling detection of 55 neuropeptides from the STNS. In addition, the central component of the STNS, STG, was imaged at a spatial resolution of 25 μm, generating detailed distribution maps of multiple neuropeptides and lipids simultaneously from this ganglion of minute size. A comparison between MRMS and HRMS that are both employed in this platform manifested where the HRMS outperforms MRMS by the well-resolved MS images of neuropeptides that could not be distinguished on an MRMS TOF/TOF. Overall, this study highlights the utility of a suite of analytical tools to localize neuropeptides in a high-throughput fashion with enhanced confidence in identification, rendering a comprehensive study of the complement and localization of neuropeptides in crustacean STNS. More importantly, the multi-faceted MSI-based platform established in this study will benefit the field of neuropeptide imaging and has the potential of facilitating future investigations on the interplay of a myriad of signaling peptides and neurotransmitters in a neuronal network.

As shown in Chapter 6, the methodology developed for novel neuropeptide discovery from the spiny lobster *Panulirus interruptus* will be applicable to the neuropeptidomic studies of decapod crustacean species with few known preprohormone sequences. The high mass accuracy and resolution data acquired on a LTQ-Orbitrap and the flexible ProSight search allows for accurate assignments of the sites where the sequence discrepancy arise from known crustacean neuropeptides, significantly

improving the identification efficiency and confidence. In addition to this streamlined semi-automated sequencing strategy, other analytical chemistry tools for neuropeptide assignment are introduced, including dimethylation-assisted fragmentation, manual *de novo* sequencing and direct tissue analysis, on multiple instruments, including nanoLC-ESI-Q-TOF and MALDI-TOF-TOF. Collectively, 55 neuropeptides are sequenced in this work, including 34 novel ones that are *de novo* sequenced and reported for the first time here. The high discovery rate from this poorly charaterized species demonstrates the usefulness of the neuropeptide discovery pipeline we developed. Finally, MALDI-MSI is employed to map neuropeptides present in *P. interruptus* brain to yield information about localization that may be important in determining the function of these charaterized neuropeptides. In summary, our study not only presents comprehensive characterization of neuropeptide expression and distribution in the brain of *P. interruptus* providing important insights into physiological functions of these endogenous neuropeptides, but also demonstrates the application of a multi-dimensional MS-based platform that will be of benefit to future neuropeptidome discovery studies on crustacean species without the knowledge of preprohormone sequences.

Finally, we demonstrate the utility of MSI application to proteomics in Chapter 7, where the MALDI-MSI-based strategy is employed to map spatial distribution and expression pattern of proteins from a rat model of neurodevelopmental disorder. The profiling strategy is fast, robust, and highly reproducible for producing proteome "snapshots" of multiple anatomical regions of interest in the infant rat brain sections, whereas the complementary MSI approach enables more detailed molecular mapping

with enhanced spatial resolution. The combination of both approaches enable localization and assessment of the relative abundances of brain proteins expressed in the MK801-treated and control rats with relatively high spatial resolution and reproducibility, leading to the discovery and identification of 22 putative biomarker targets. The subsequent top-down MS/MS approach allows the identifications of these putative marker proteins through accurate mass measurement of intact proteins followed by complementary fragmentation via tandem MS analysis. This hybrid approach showcases the unique abilities of MSI in the discovery of potential disease biomarkers and HRMS/MS sequencing for accurate assignment. Moreover, this study opens up new avenues for future applications in the field of biomarker discovery and for mechanistic investigations delineating pathophysiology of neurodevelopmental disorders.

8.2 Future Directions

This work lays the foundation for future research efforts on several on-going projects in this lab. In Chapter 3, we demonstrate the feasibility to image NTs with peptide signaling molecules, neuropeptides, from serial sections, whereas simultaneous imaging of the two classes of molecules can be achieved on single tissue section by performing multiple types of scans at one “pixel”. This would deliver invaluable information regarding the co-localization of NTs and NPs, raising questions regarding their possibly synergistic, complementary, or antagonistic relations that could be present among these co-localized modulators in the neural network. Moreover, a limitation of HRMSI is the lack of MS/MS spectra for definitive assignment of molecular structures. This limitation could also be addressed by incorporating multiple scans in one set of

experiments;[4-6] for example, one full MS scan of low MW range with one targeted MS/MS scan of the precursor like ACh. The HRMS image of the precursor along with the MS/MS image of the resulting signature fragment ion would provide solid evidence for the assignment of the precursor via HRMSI alone.

For the study described in Chapter 4, further refinement of matrix application parameters and methods would be beneficial to acquire MS images at higher spatial resolution. Currently, a project that compares the crystallization and extraction efficiency of matrices with airbrush (original method used in this study), an automatic TM sprayer and sublimation apparatus is underway. With more efforts devoted to optimizing the matrix application process, better-resolved MS images of metabolites with greater coverage would be obtained. Another possible direction is to carry out label-free quantitation experiments on the metabolites with a conventional LC-MS approach, because more secondary metabolites are expected to be discovered from this biological system as compared to previous report from other leguminous plant.[7, 8] Moreover, the differential abundance of metabolites observed with LC-MS is expected to agree with MALDI-MSI data.

In Chapter 5, we extended the MSI applications to the area of neuropeptidomics. With the smartbeam laser, we successfully acquired MS images of spatial resolution at 25 μm. Nevertheless, the resolution could be further improved by the oversampling technique. In addition, the desired matrix crystals smaller than observable tissue features could be achieved with the solutions discussed in Chapter 4 on the premise that the

extraction efficiency is not sacrificed due to the fact that small molecules usually fly more easily compared to large peptides and proteins.

A lot remains to be discovered for the *P. interruptus* neuropeptidome, especially the functions of the novel neuropeptides discovered. For example, we have found some novel homologous neuropeptides such as A-orcokinin and SIYamide along with S-orcokinin and SIFamide, which have been previously charaterized from species such as *C. borealis* or *C. sapidus*. Further in-depth studies that explore the physiological functions of these homologs such as A-orcokinin and S-orcokinin would help to elucidate whether *P. interruptus* has developed another signaling pathway in which the newly discovered A-orcokinins are involved in signaling. Future studies that collect more *P. interruptus* brain tissues followed by HPLC fractionation and analysis of each individual fraction should significantly enhance the number of novel neuropeptides discovered from this species. In addition, neuropeptidomic studies with more uncharacterized crustacean species would further validate our semi-automated sequencing strategy and lead to the expansion of our decapod crustacean neuropeptide database.

In addition to presenting a novel approach towards discovery of biomarkers for neurodevelopmental disorders, the work described in Chapter 7 also provides insights into possible pathway involved in various neuropsychiatric syndromes that could result from excitatory/inhibitory neurotransmission imbalance in early development. Biological assays such as immunohistochemical staining and western blot are expected to validate the putative biomarker proteins that displayed differential abundances from the samples

subjected to MK801 treatment compared to vehicle as observed via our MALDI-MSI-based approach.

In summary, the work described in this dissertation has showcased an advanced and powerful MSI-based platform that has been demonstrated to facilitate the mapping and identification of various types of biomolecules from tissues and extracts. Its versatility designates its future use in a myriad of biological models under various physiological or behavioral manipulations. Moreover, the research conducted in this thesis also contains a useful biological information that help to improve our understanding of several challenging neuroscience questions. It is anticipated that the analytical methodology we developed and experimental data we obtained in this study would impact future research in both analytical chemistry and neuroscience fields.

8.3 References

1. Braak, H.; Del Tredici, K.; Rub, U.; de Vos, R. A.; Jansen Steur, E. N.; Braak, E. (2003) Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* 24, 197-211.
2. Want, E. J.; Cravatt, B. F.; Siuzdak, G. (2005) The expanding role of mass spectrometry in metabolite profiling and characterization. *Chembiochem.* 6, 1941-1951.
3. Watanabe, T.; Yamagata, N.; Takasaki, K.; Sano, K.; Hayakawa, K.; Katsurabayashi, S.; Egashira, N.; Mishima, K.; Iwasaki, K.; Fujiwara, M. (2009) Decreased acetylcholine release as correlated to memory impairment in the Tg2576 transgenic mouse model of Alzheimer's disease. *Brain Res.* 1249, 222-228.
4. Yagnik, G. B.; Korte, A. R.; Lee, Y. J. (2013) Multiplex mass spectrometry imaging for latent fingerprints. *J. Mass Spectrom* 48, 100-104.
5. Lee, Y. J.; Perdian, D. C.; Song, Z.; Yeung, E. S.; Nikolau, B. J. (2012) Use of mass spectrometry for imaging metabolites in plants. *Plant J.* 70, 81-95.
6. Perdian, D. C.; Lee, Y. J. (2010) Imaging MS methodology for more chemical information in less data acquisition time utilizing a hybrid linear ion trap-orbitrap mass spectrometer. *Anal. Chem.* 82, 9393-400.

7. Farag, M. A.; Huhman, D. V.; Dixon, R. A.; Sumner, L. W. (2008) Metabolomics reveals novel pathways and differential mechanistic and elicitor-specific responses in phenylpropanoid and isoflavonoid biosynthesis in *Medicago truncatula* cell cultures. *Plant Physiol.* 146, 387-402.
8. Farag, M. A.; Huhman, D. V.; Lei, Z.; Sumner, L. W. (2007) Metabolic profiling and systematic identification of flavonoids and isoflavonoids in roots and cell suspension cultures of *Medicago truncatula* using HPLC-UV-ESI-MS and GC-MS. *Phytochemistry* 68, 342-354.

Appendix I

List of Publications and Presentations

Publications

1. Hui Ye, Jingxin Wang, Tyler Greer, Kerstin Strupat, Lingjun Li. Visualizing neurotransmitters and metabolites in central nervous systems by high resolution and accuracy mass spectrometric imaging, *ACS Chem. Neurosci*., in press.

2. Hui Ye, Erin Gemperline, Muthusubramanian Venkateshwaran, Ruibing Chen, Pierre-Marc Delaux, Maegen Howes-Podoll, Jean-Michel Ané, Lingjun Li. MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the *Medicago truncatula* – *Sinorhizobium meliloti* symbiosis. *Plant J.*, doi: 10.1111/tpj.12191.

3. Hui Ye, Limei Hui, Katherine Kellersberger, Lingjun Li. Mapping of neuropeptides in the crustacean stomatogastric nervous system by imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 20, 1068-77.

4. Hui Ye, Erin Gemperline, Lingjun Li. A vision for better health: mass spectrometry imaging for clinical diagnostics. *Clin. Chim. Acta,* doi: 10.1016/j.cca.2012.10.018.

5. Hui Ye, Tyler Greer, Lingjun Li. Probing neuropeptide signaling at the organ and cellular domains via imaging mass spectrometry. *J. Proteomics*, **2012**, 75, 5014-26.

6. Hui Ye, Tyler Greer, Lingjun Li. From pixel to voxel: a deeper view of biological tissue by 3D mass spectral imaging. *Bioanalysis*, **2011**, 3, 313-32.

7. Hui Ye, Rakesh Mandal, Adam Catherman, Paul Martin Thomas, Neil Kelleher, Chrysanthy Ikonomidou, Lingjun Li. Top-down proteomics with mass spectrometry imaging: a novel approach towards discovery of biomarkers for neurodevelopmental disorders. *PLoS One*, to be submitted.

8. Hui Ye, Jingxin Wang, Zichuan Zhang, Chenxi Jia, Claire Schmerberg, Adam Catherman, Paul Martin Thomas, Neil Kelleher and Lingjun Li. Defining the Neuropeptidome of the spiny lobster *Panulirus interruptus* brain using a multi-dimensional mass spectrometry-based platform. *J. Proteome Res.*, to be submitted.

9. Hui Ye, Jingxin Wang, James Dowell, Quentin Bremer, Brian Baldo, Lingjun Li. Quantitation of neuropeptides in discrete rat brain regions after food with mass spectrometry, manuscript in preparation.

10. Zichuan Zhang, **Hui Ye**, Junhua Wang, Limei Hui and Lingjun Li. Pressure-assisted capillary electrophoresis coupling with matrix-assisted laser desorption/ionization-mass spectrometric imaging for quantitative analysis of complex peptide mixtures. *Anal. Chem.*, 2012, 84, 7684-91.

11. Junhua Wang, **Hui Ye**, Zichuan Zhang, Feng Xiang, Gary Girdaukas, Lingjun Li. Advancing matrix-assisted laser desorption/ionization-mass spectrometric imaging for capillary electrophoresis analysis of peptides. *Anal. Chem.*, **2011**, 83, 3462-69.

12. Feng Xiang, **Hui Ye**, Ruibing Chen, Qiang Fu and Lingjun Li. N,N-dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics. *Anal. Chem.*, **2010**, 82, 2817–2825.

13. Limei Hui, Yuzhuo Zhang, Junhua Wang, Aaron Cook, **Hui Ye**, Michael P. Nusbaum, and Lingjun Li. Discovery and functional study of a novel crustacean tachykinin neuropeptide. *ACS Chem. Neurosci.*, **2011**, 2, 711-722.

14. Chenxi Jia, Chris Lietz, **Hui Ye**, Limei Hui, Qing Yu, Sujin Yoo and Lingjun Li. A multi-scale strategy for discovery of novel endogenous neuropeptides in crustacean nervous system. *J. Proteomics*, pending acceptance.

15. Man Jiang, Lin-Tao Yu, **Hui Ye** and Xiang-Ge Zhou. A new salen ligand: 2, 2'-[(spiro [4.4] nonane-1, 6-diyl) dinitrilomethylidyne] bis (4, 6-dichlorophenol), *Acta Crystallogr.*, **2007**, E63, o3049.

Selected Presentations

1. Hui Ye, Cameron Scarlett, Lingjun Li. "Research resources for neuroendocrine system characterization using mass spectrometry," oral presentation, campus-wide **ICTR seminar**, Madison, WI, 2012. （http://videos.med.wisc.edu/videoInfo.php?videoid=42224）

2. Hui Ye, Rakesh Mandal, Chen Chen, Chrysanthy Ikonomidou, Lingjun Li. "Mass spectrometric
imaging of brain proteins following neonatal exposure to an NMDA antagonist and ethanol", poster presentation, 60th **ASMS**, Vancouver, Canada, 2012.

3. Hui Ye, Rakesh Mandal, Lingjun Li, Chrysanthy Ikonomidou. "Mass spectrometric imaging of brain proteins following neonatal exposure to an NMDA antagonist MK801," poster presentation, **Neuroscience**, Washington DC, 2011.

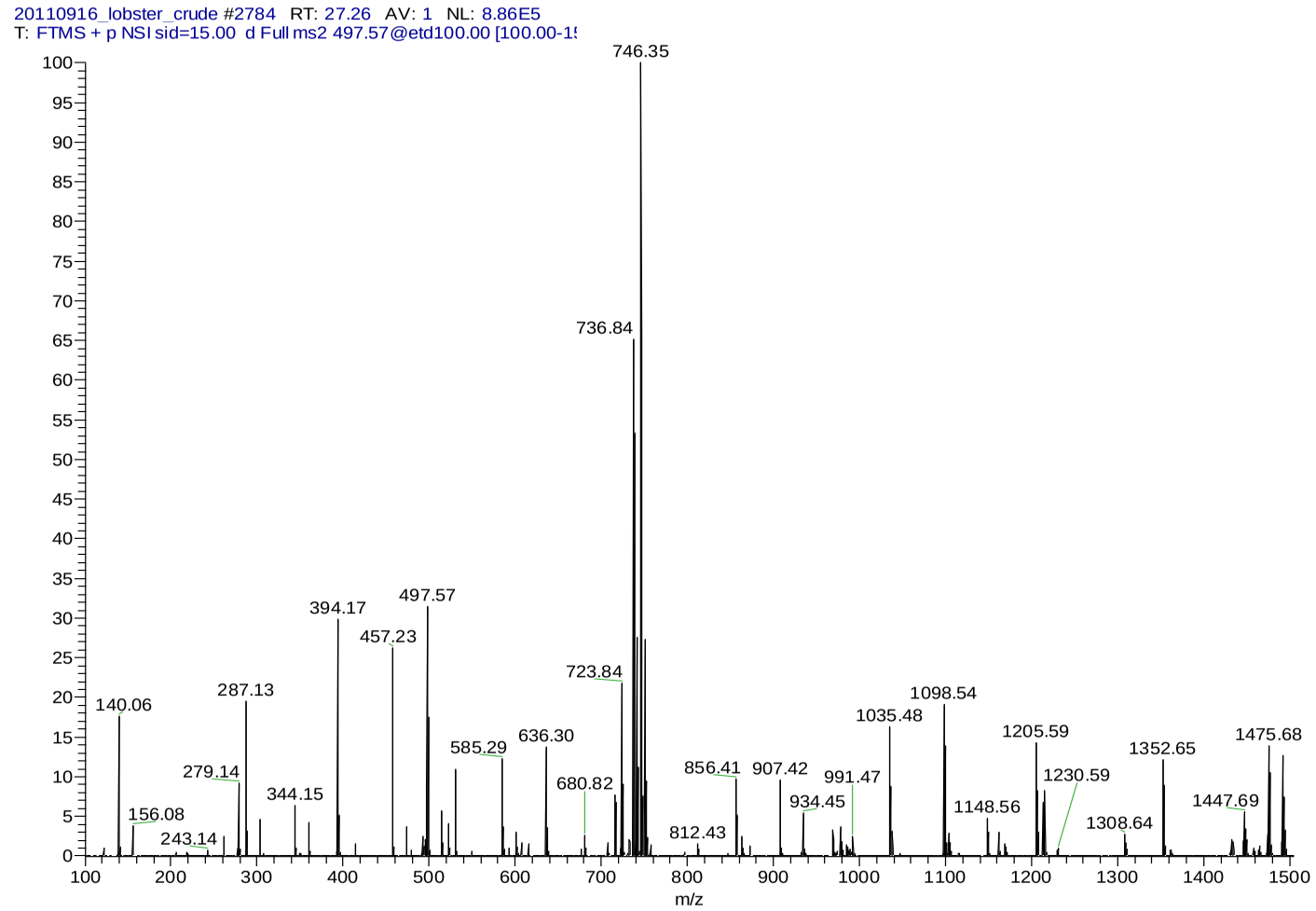
4. Hui Ye, Yuzhuo Zhang, Matt Spencer, Lingjun Li. "Qualitative and Quantitative Analysis of Neurotransmitters and Neuropeptides in the decapod crustacean nervous System by Gold-assisted Laser Desorption/Ionization," oral presentation, **Pittcon**, Atlanta, GA, 2011.

5. Hui Ye, Ruibing Chen, Robert Sturm, Maegen Howes-Podoll, Jean-Michel Ané, Lingjun Li. Exploring the metabolite distribution in *Medicago truncatula* Roots and Nodules by MALDI Mass Spectrometry Imaging (MSI)," poster presentation, 58th **ASMS** Annual Conference, Salt Lake City, Utah, 2010. （presented at **ASMS** conferences from 2010 to 2012）

Appendix II

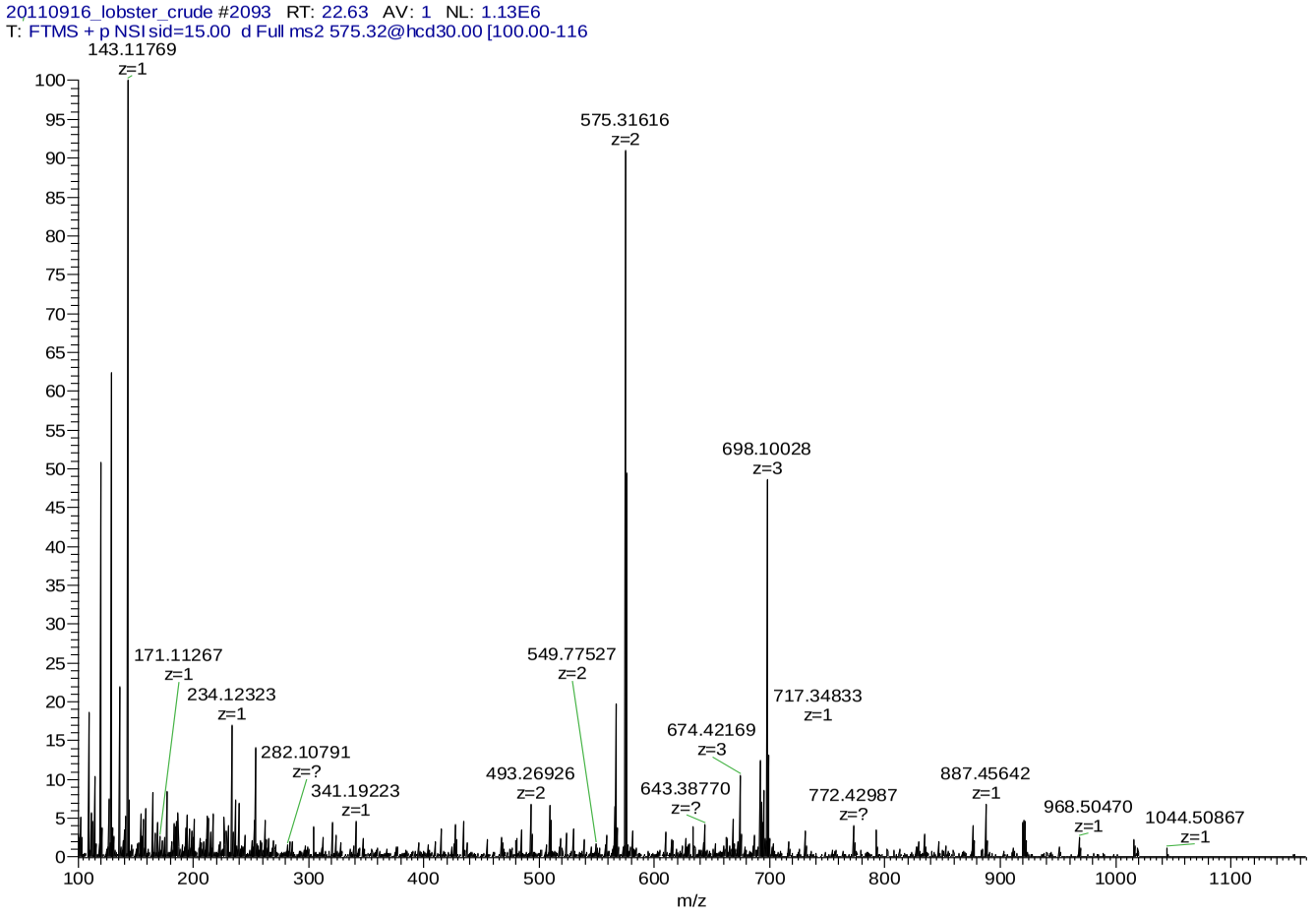
MS/MS Spectra and Fragmentation Maps of Neuropeptides in Chapter 6

Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2784	1793	Orcokinin	1489.6909	1489.694	-2.34948	NFDEIDRTGFGFH	NFDEIDRAGLGFH		3.77E-47

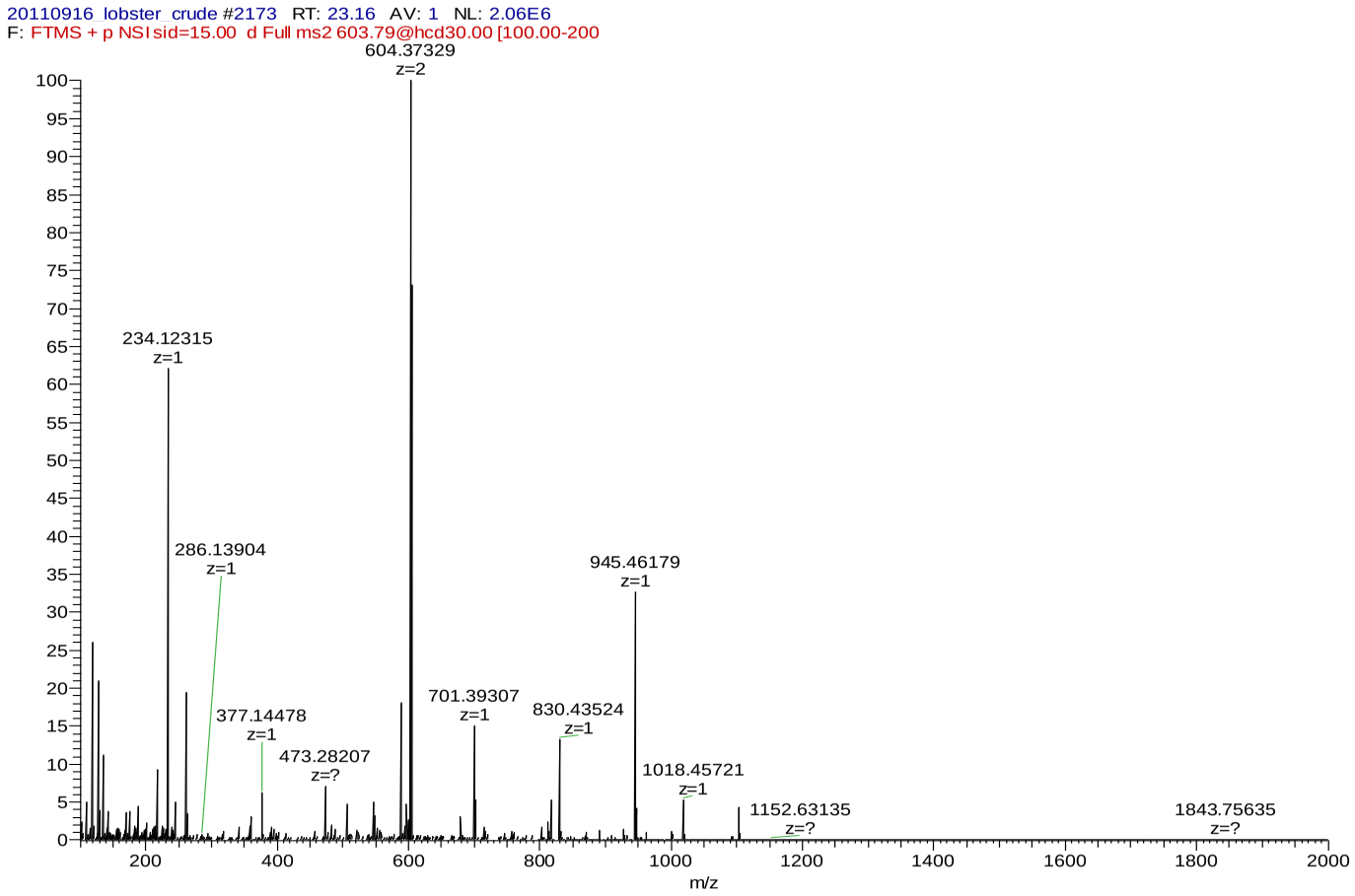


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2093	1348	Orcokinin	1147.559	1147.562	-2.5271	NFDEIDRSSFA	NFDEIDRAGLa	Amidation	4.94E-08

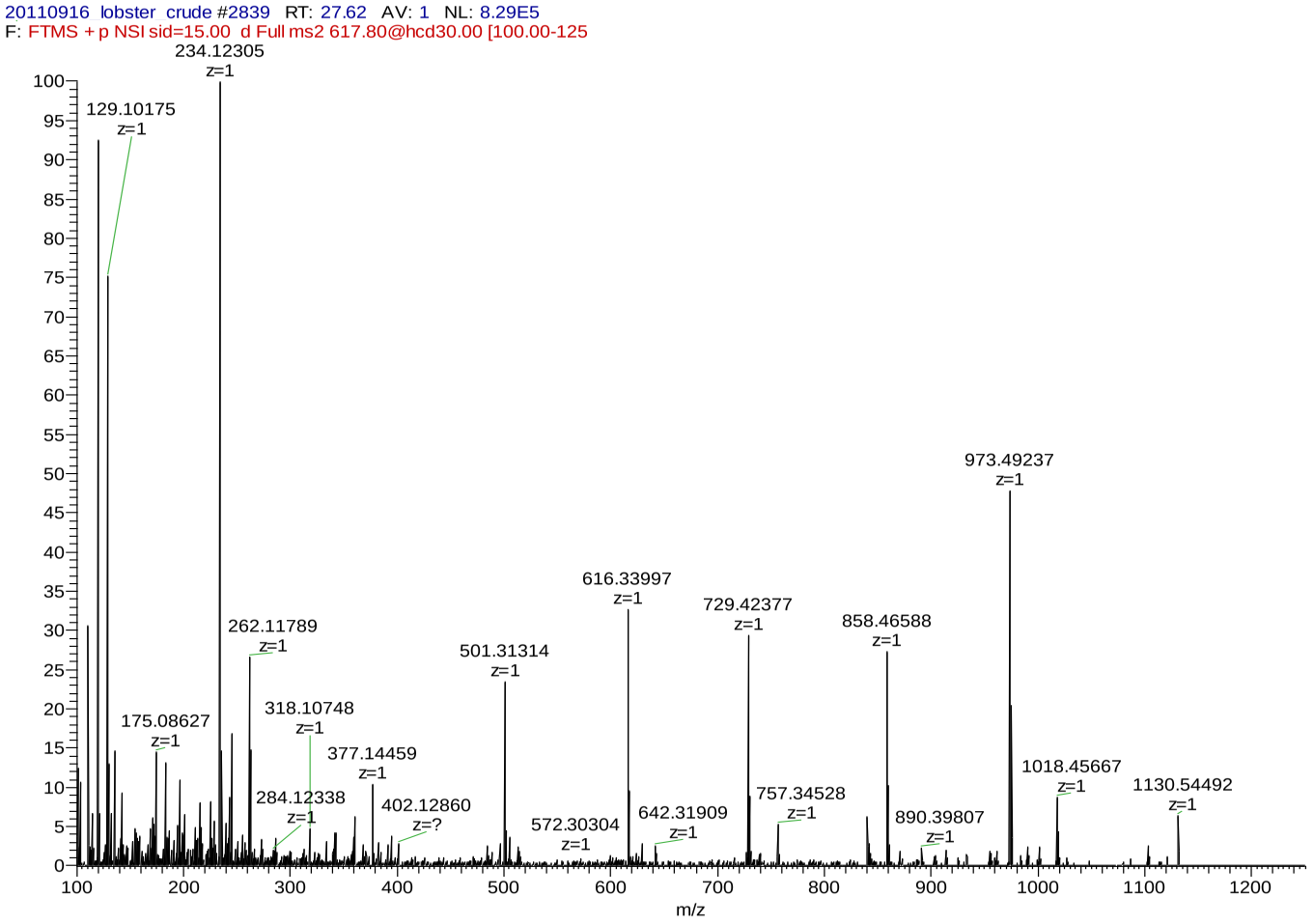
-N-F-[D]-E-I-[D]-R-A-G-[L]-a



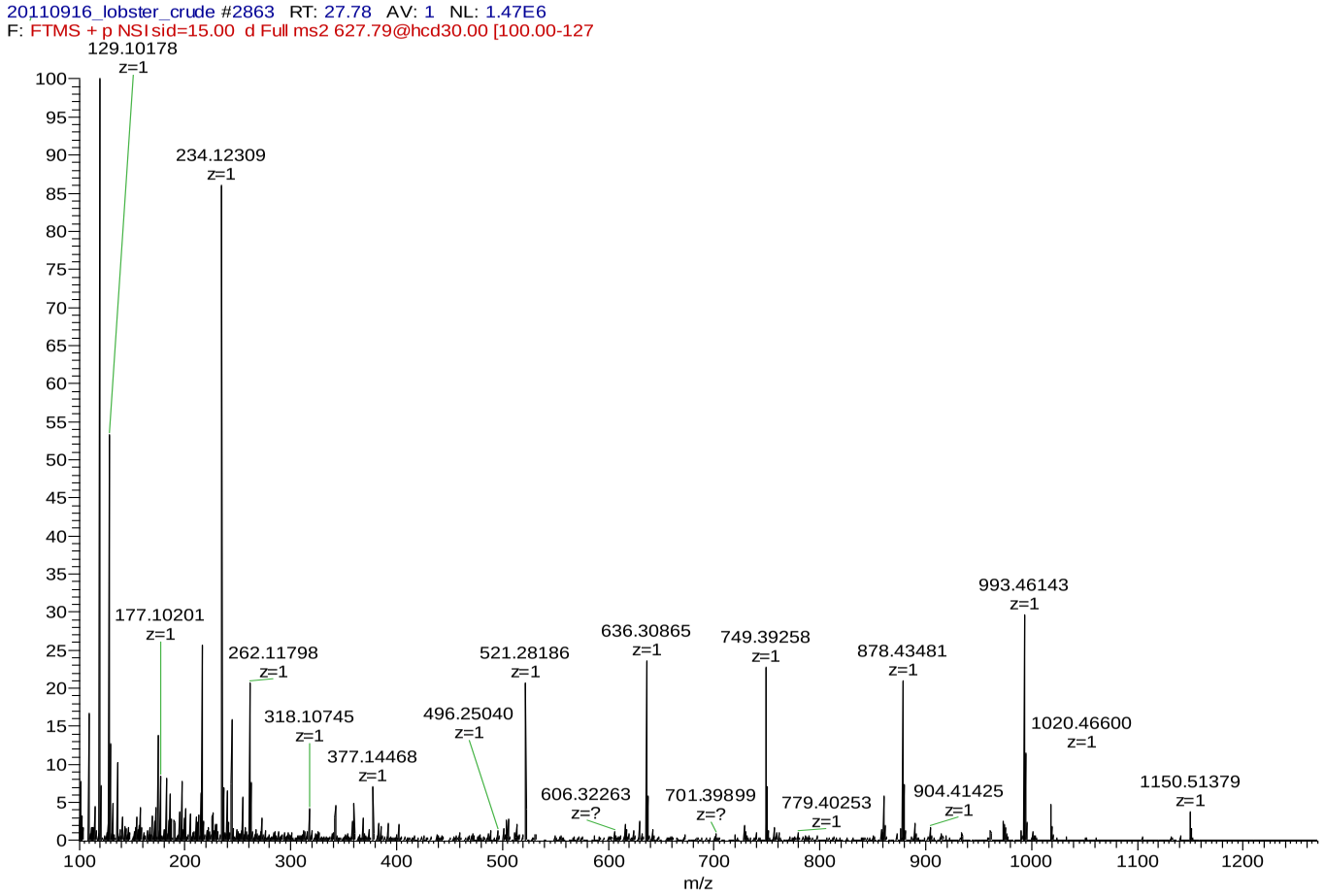
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2173	1399	Orcokinin	1205.564	1205.567	-2.32256	NFDEIDRSGFG	NFDEIDRAGLG		9.22E-13



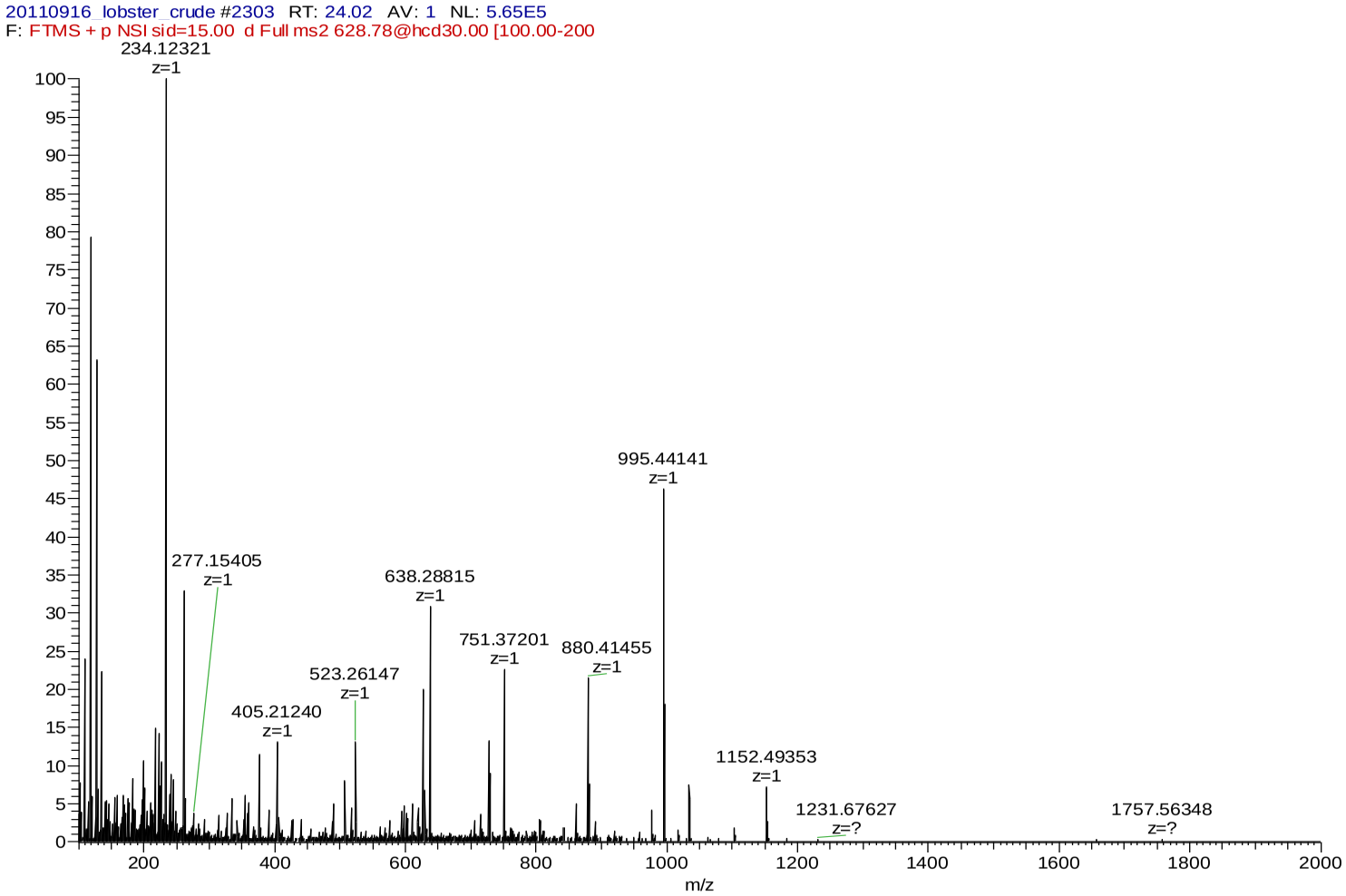
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2839	1831	Orcokinin	1233.595	1233.598	-2.43191	NFDEIDRSSFGF	NFDEIDRAGVV		1.41E-26



Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2863	1840	Orcokinin	1253.564	1253.567	-2.15385	NFDEIDRSSFA	NFDEIDRAGFA		3.48E-20

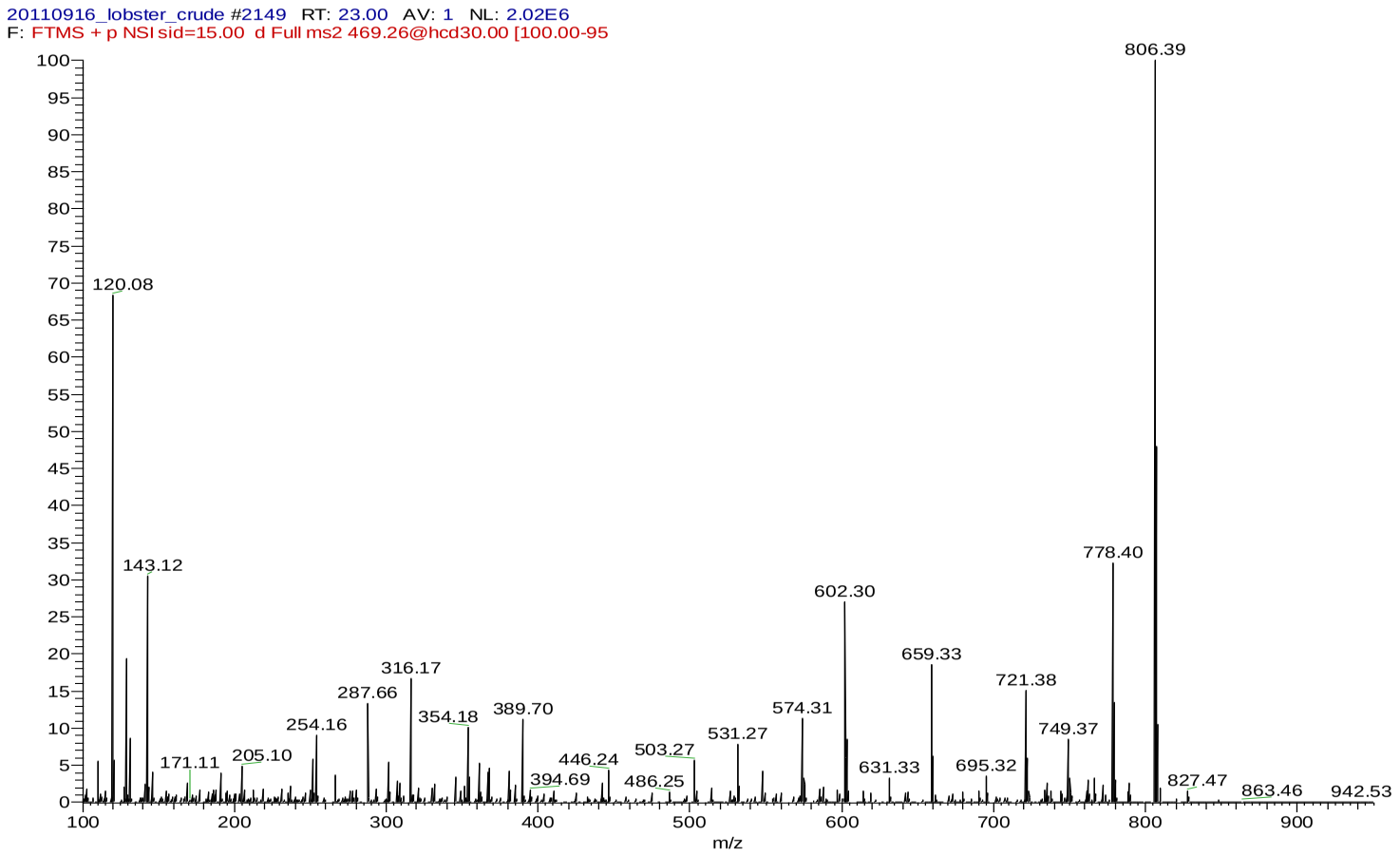


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2303	1486	Orcokinin	1255.543	1255.546	-2.46904	NFDEIDRSGFGF	NFDEIDRSGFG		9.05E-23

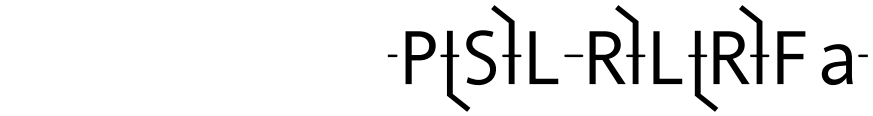


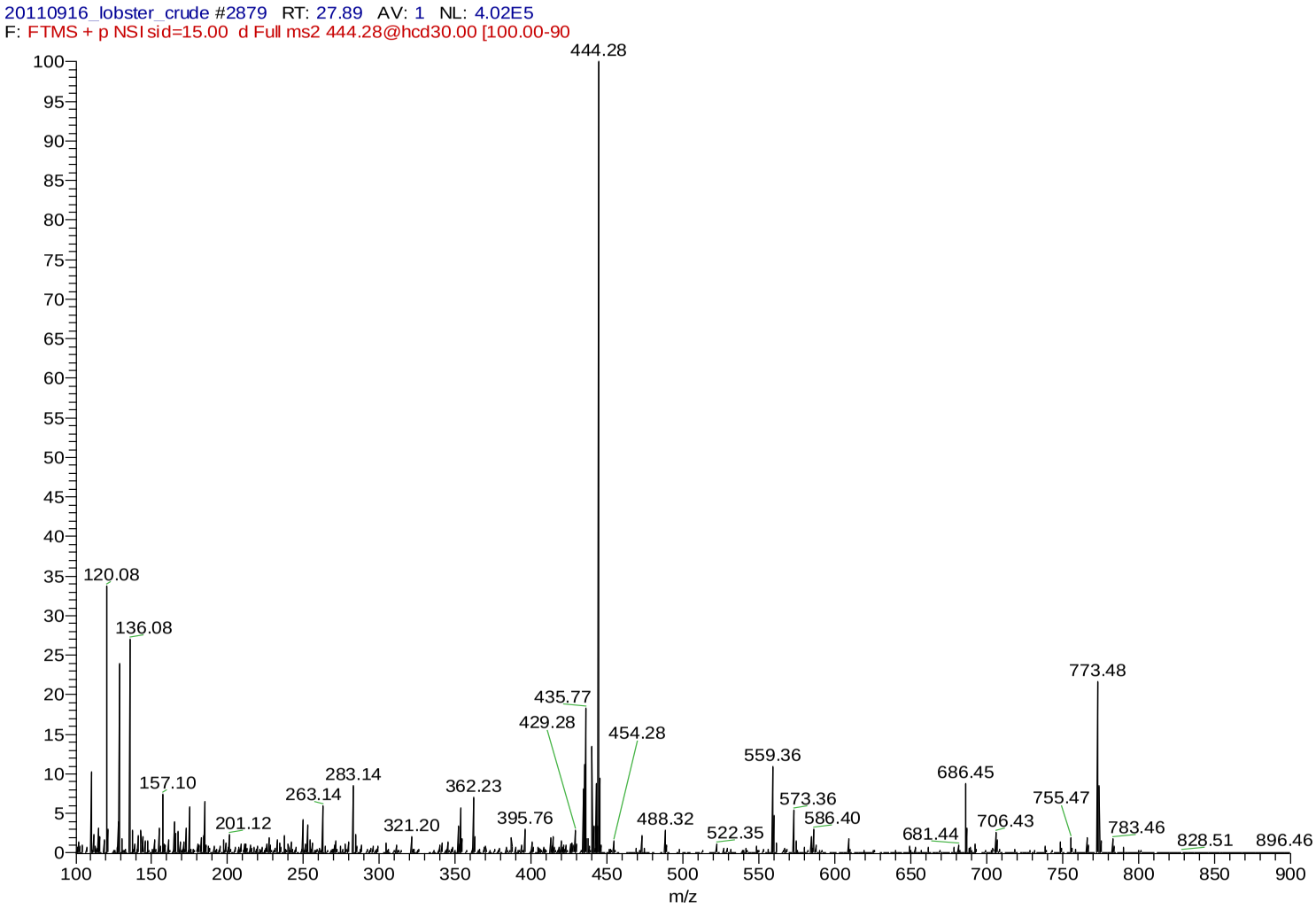
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2149	1387	AST-A	935.4949	935.4972	-2.45859	PRNYAFGL	PRNYAFGLa	Amidation	9.70E-14

-P-R-N-Y-A-F-G-L a-

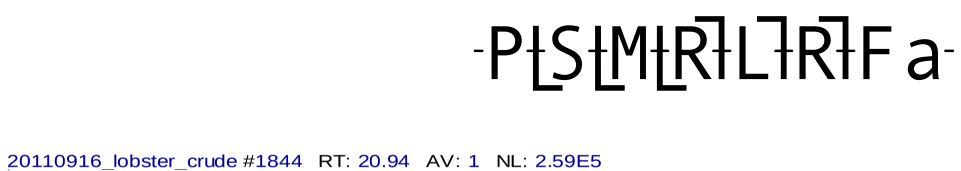


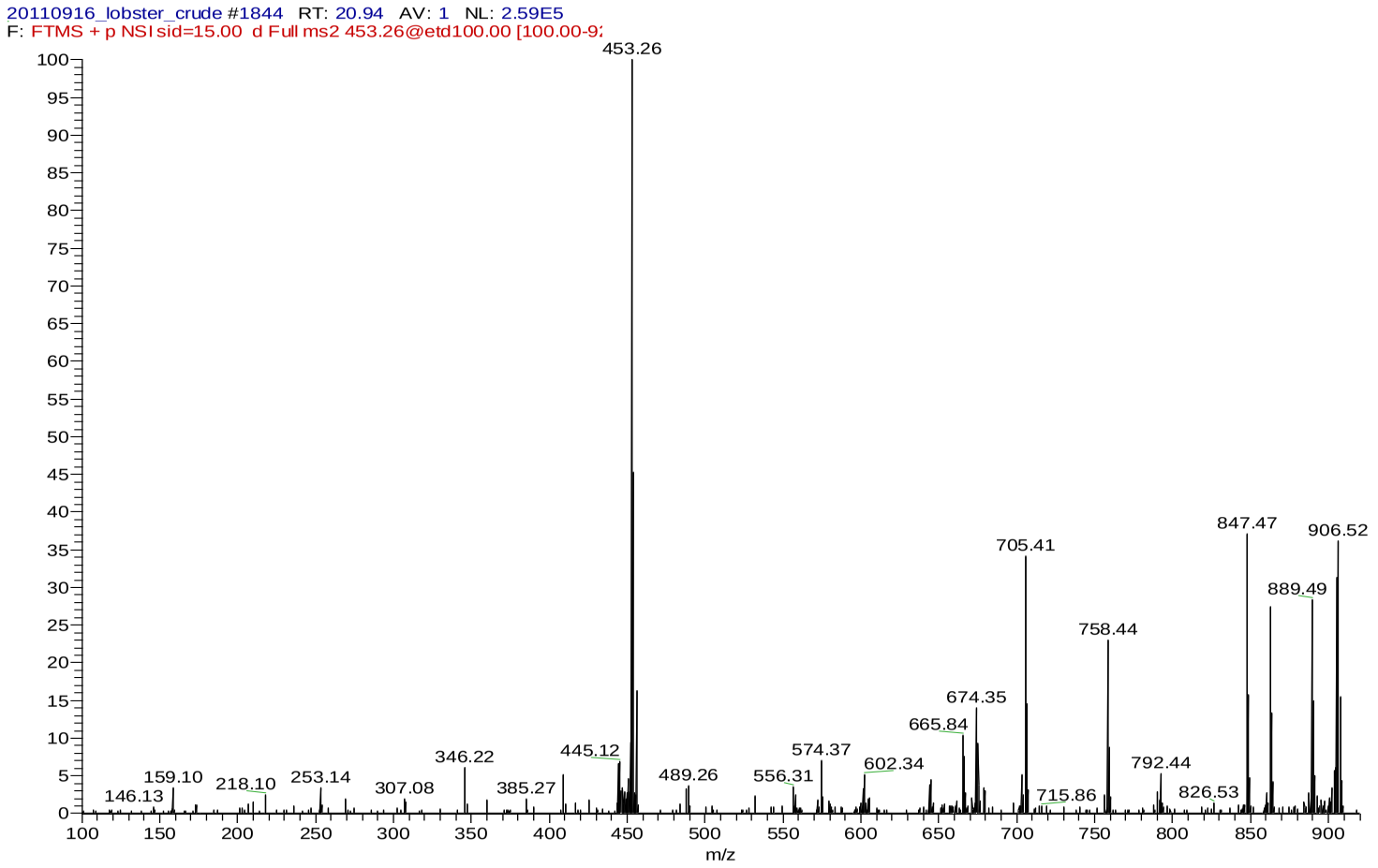
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2879	1855	Rfa	886.5473	886.5496	-2.59433	PSLRLRF	PSLRLRFa	Amidation	6.95E-12



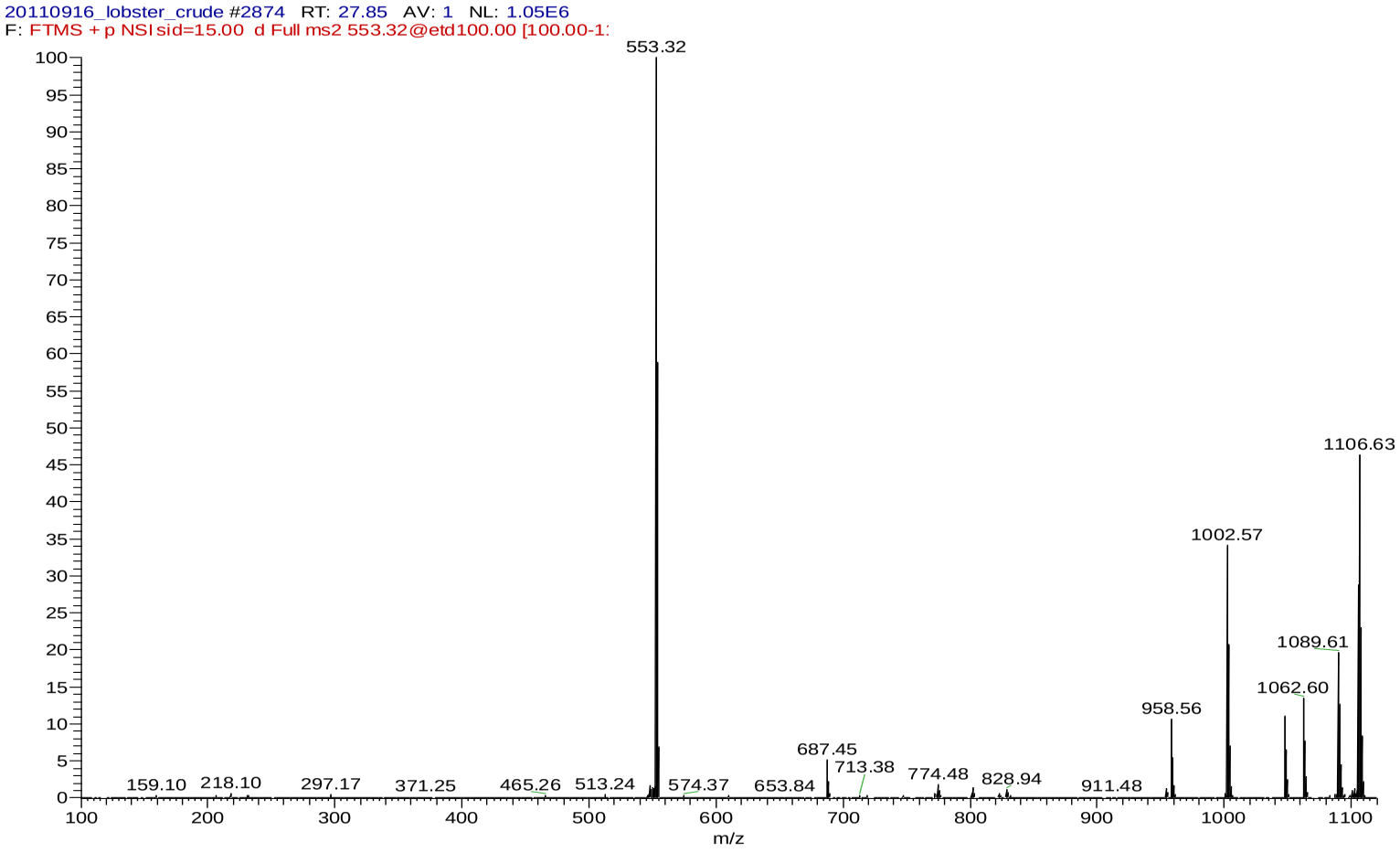


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
1844	1196	Rfa	904.5036	904.506	-2.65338	PSMRLRF	PSMRLRFa	Amidation	1.34E-13

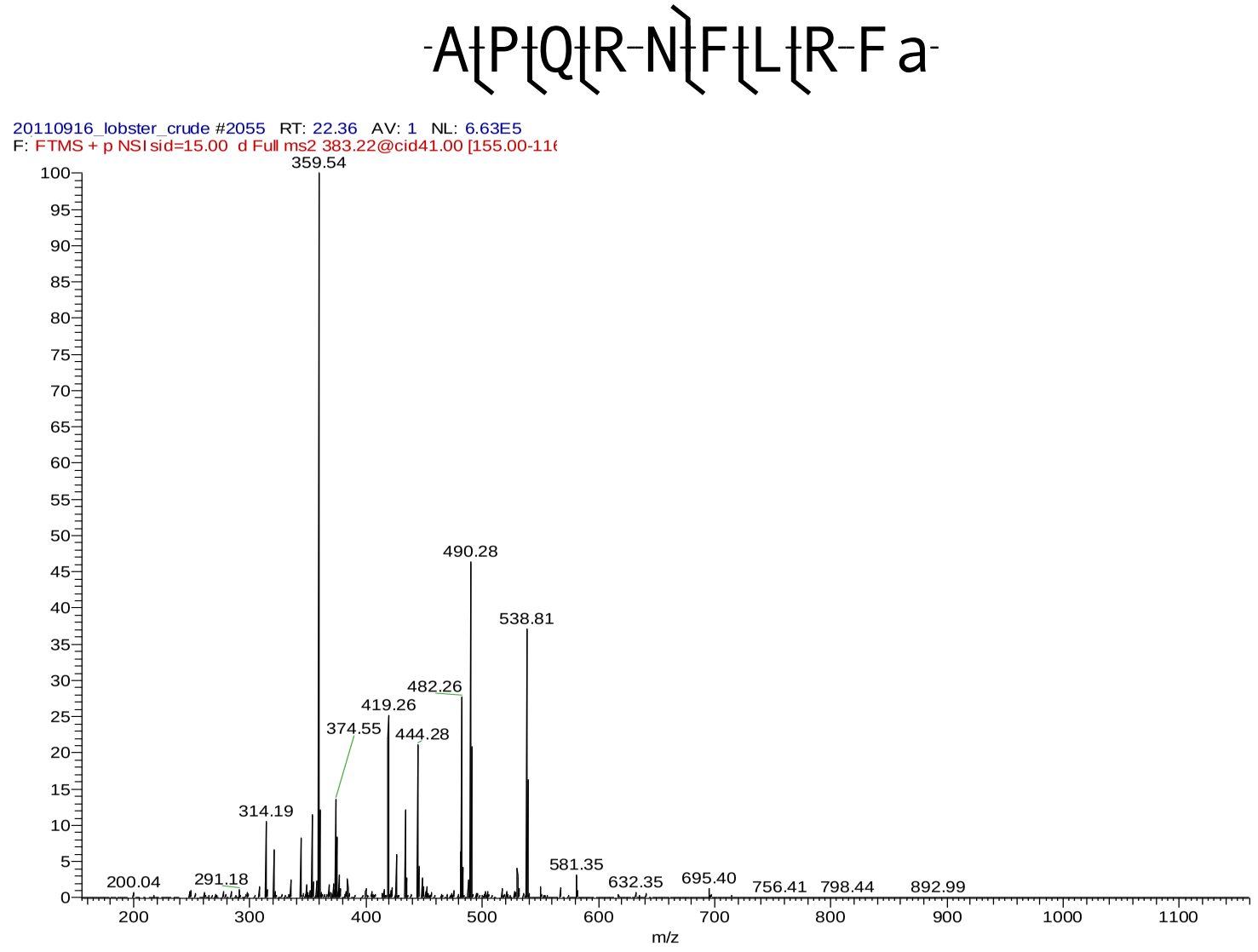




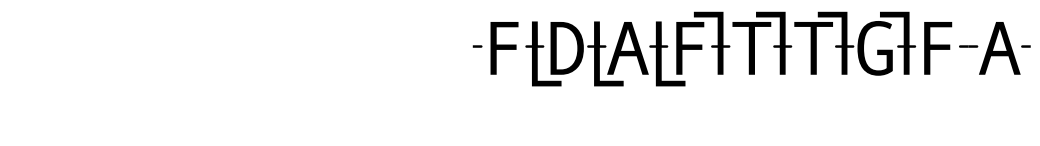
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2874	1850	Rfa	1104.62	1104.622	-2.35375	SMPSLRLRF	SMPSLRLRFa	Amidation	5.96E-16

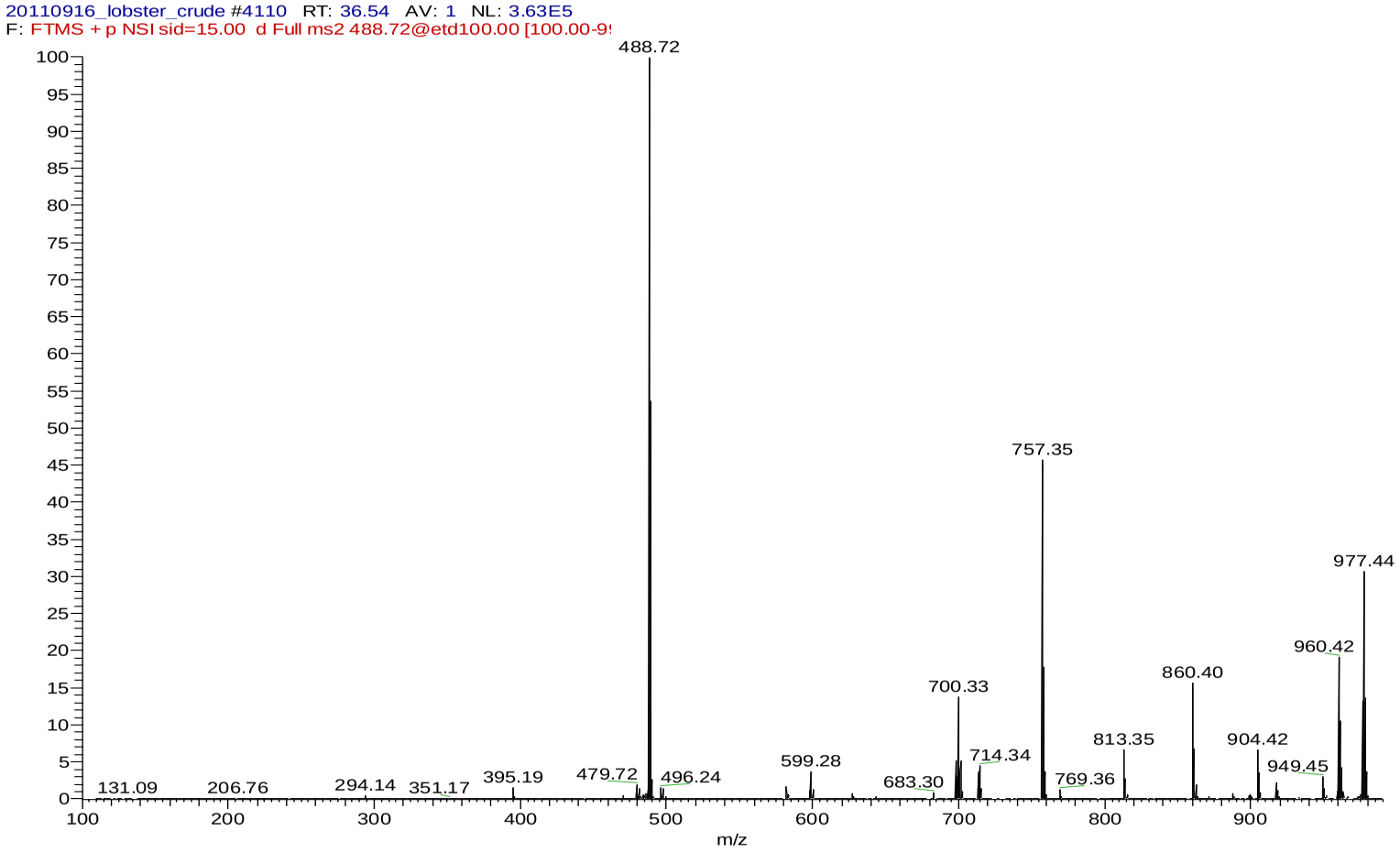


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2055	1329	Rfa	1146.638	1146.641	-2.52913	APQRNFLRF	APQRNFLRFa	Amidation	1.04E-12

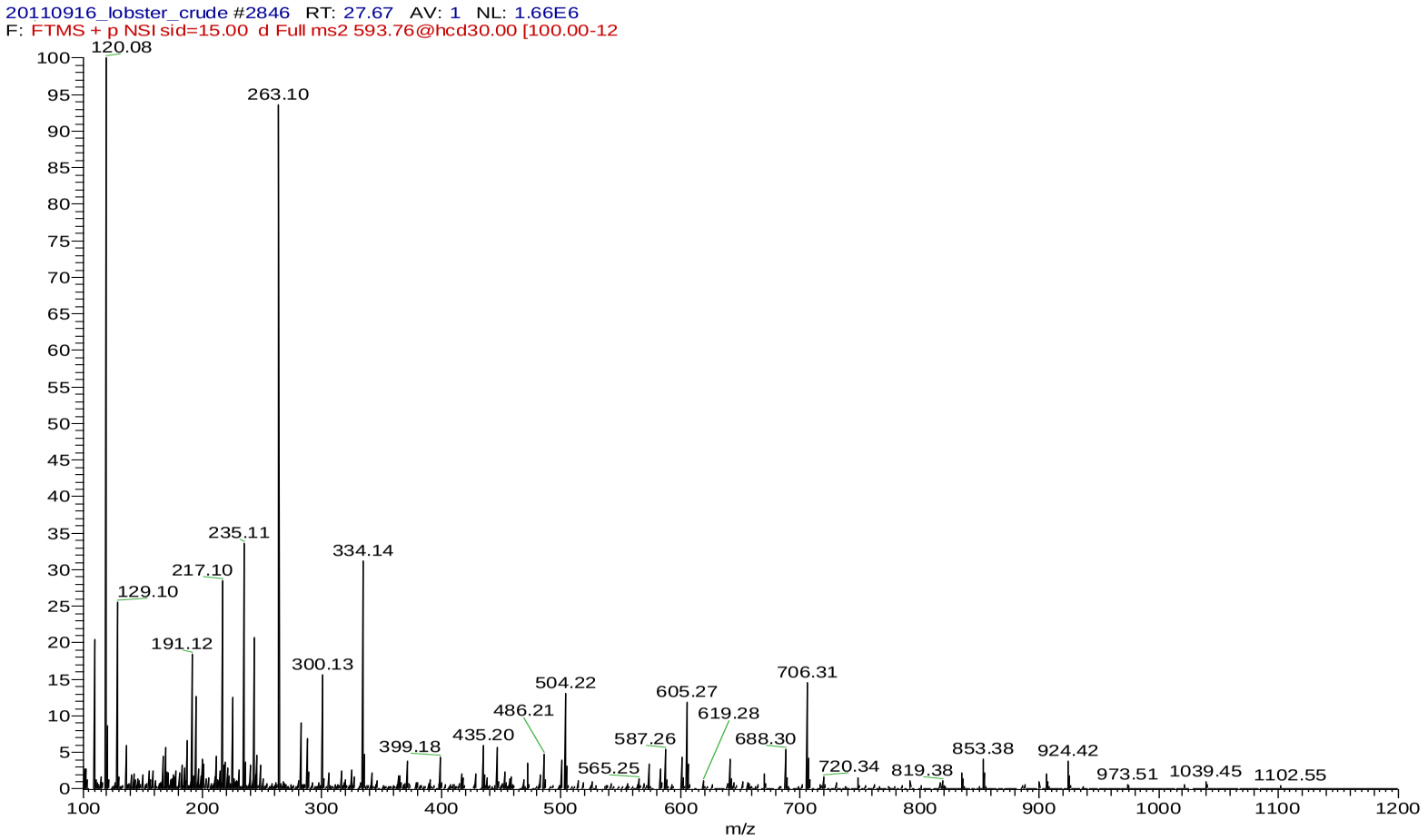


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
3699	2666	Orcomytropin	975.4309	975.4332	-2.35793	FDAFTTGFGHS	FDAFTTGFA		1.02E-17

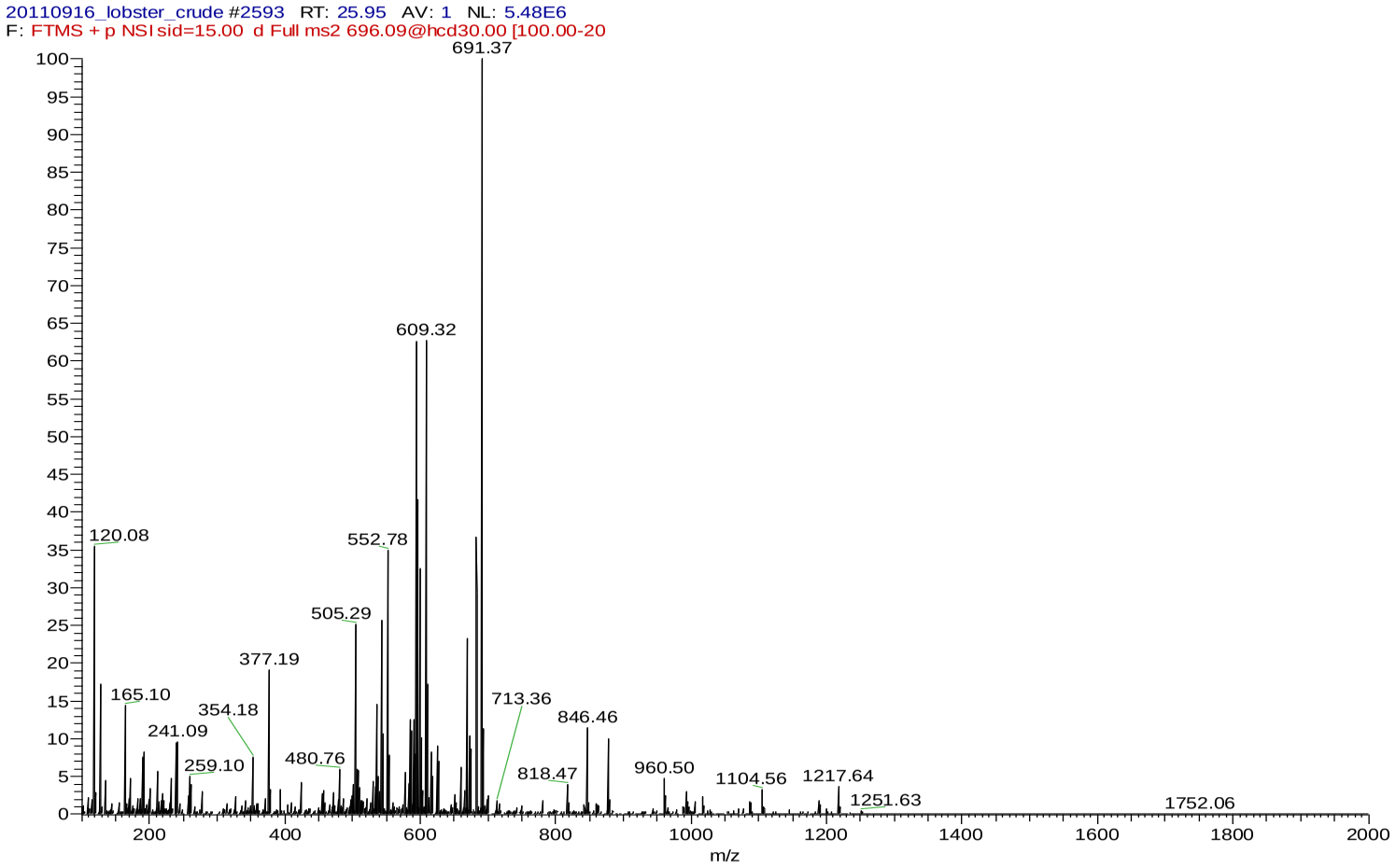




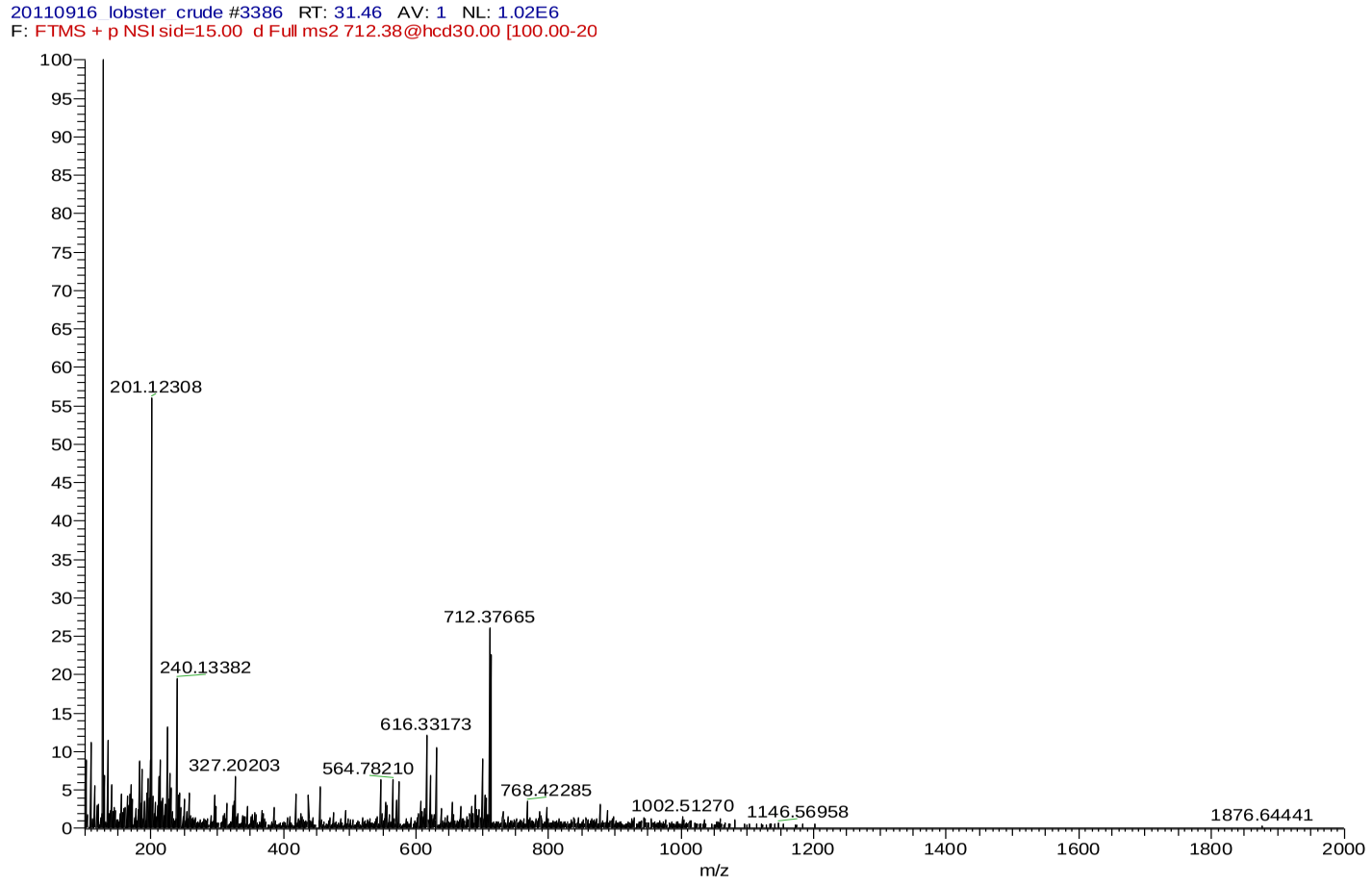
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2846	2828	Orcomytropin	1185.505	1185.509	-2.86797	FDAFTTGFGHS	FDAFTTGFGHS		2.06E-18



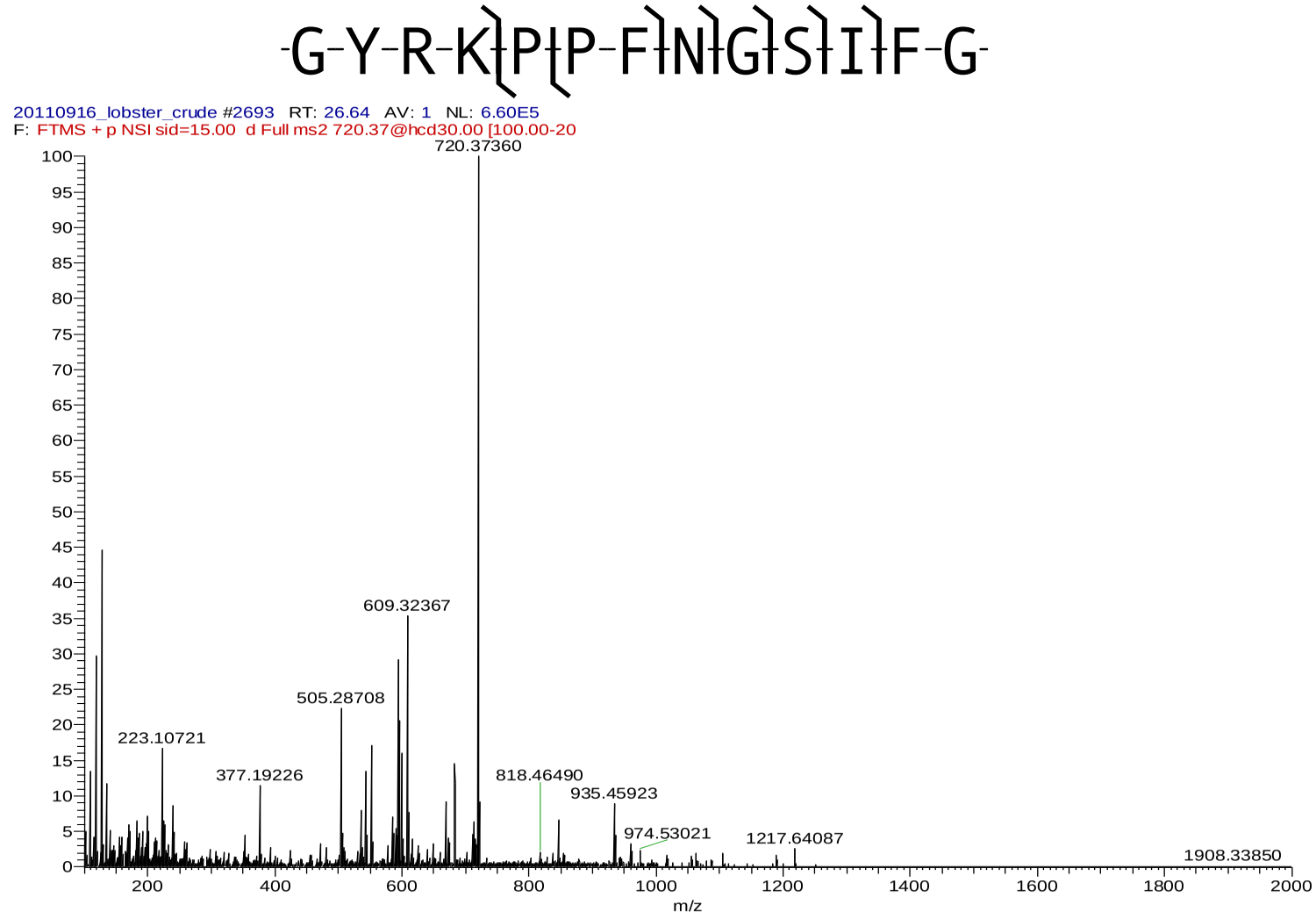
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2593	1651	SiFa	1380.727	1380.73	-2.02791	GYRKPPFNGSIF	GYRKPPFNGSIFa	Amidation	1.99E-26



Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
3386	2173	SIFa	1422.738	1422.777	-26.9895	GYRKPPFNGSIF	VYRKPPFNGSIFa	Acetylation, Amidation	2.64E-13

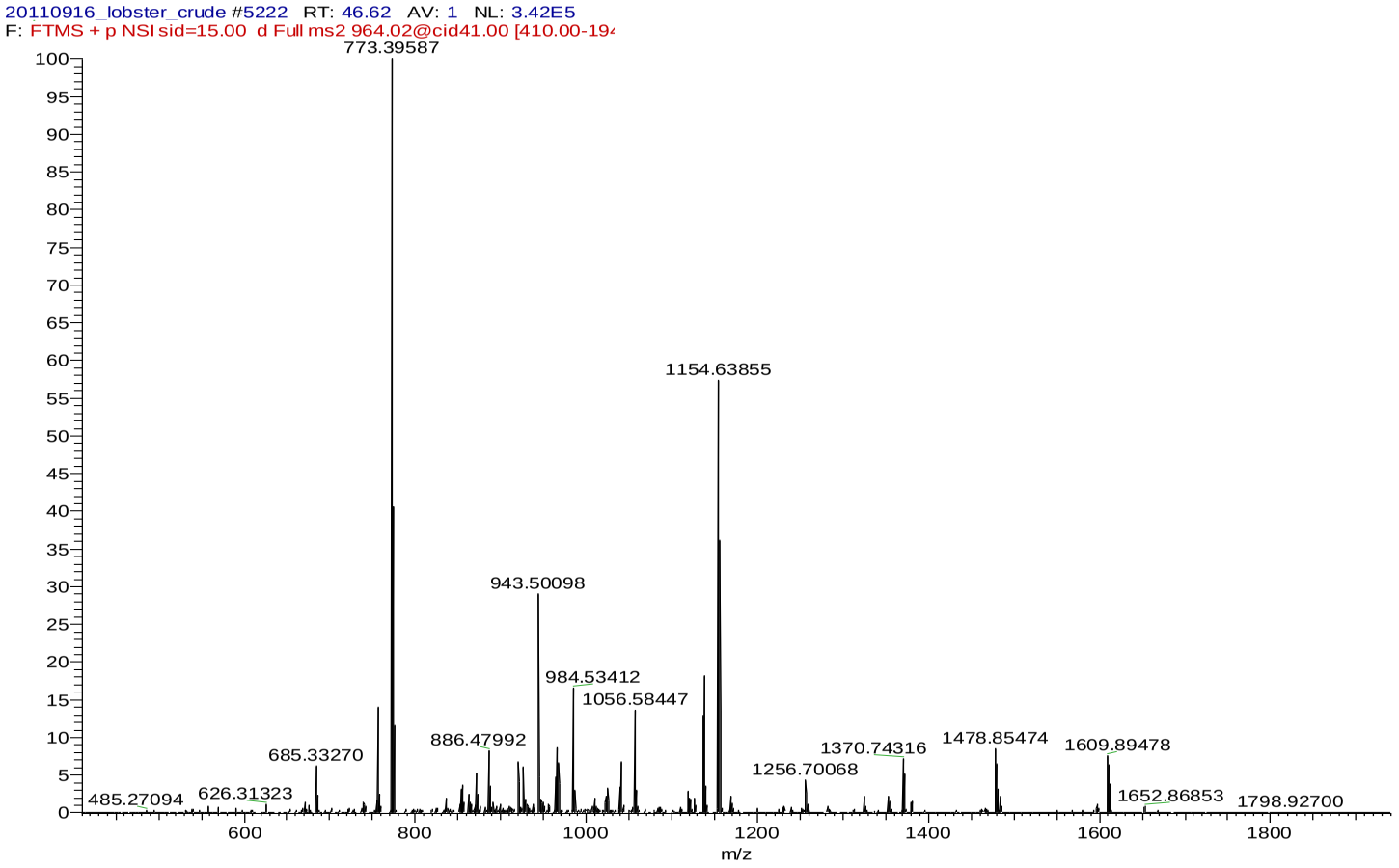


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2693	1735	SIFa	1438.732	1438.735	-2.29368	GYRKPPFNGSIF	GYRKPPFNGSIFG		1.06E-17

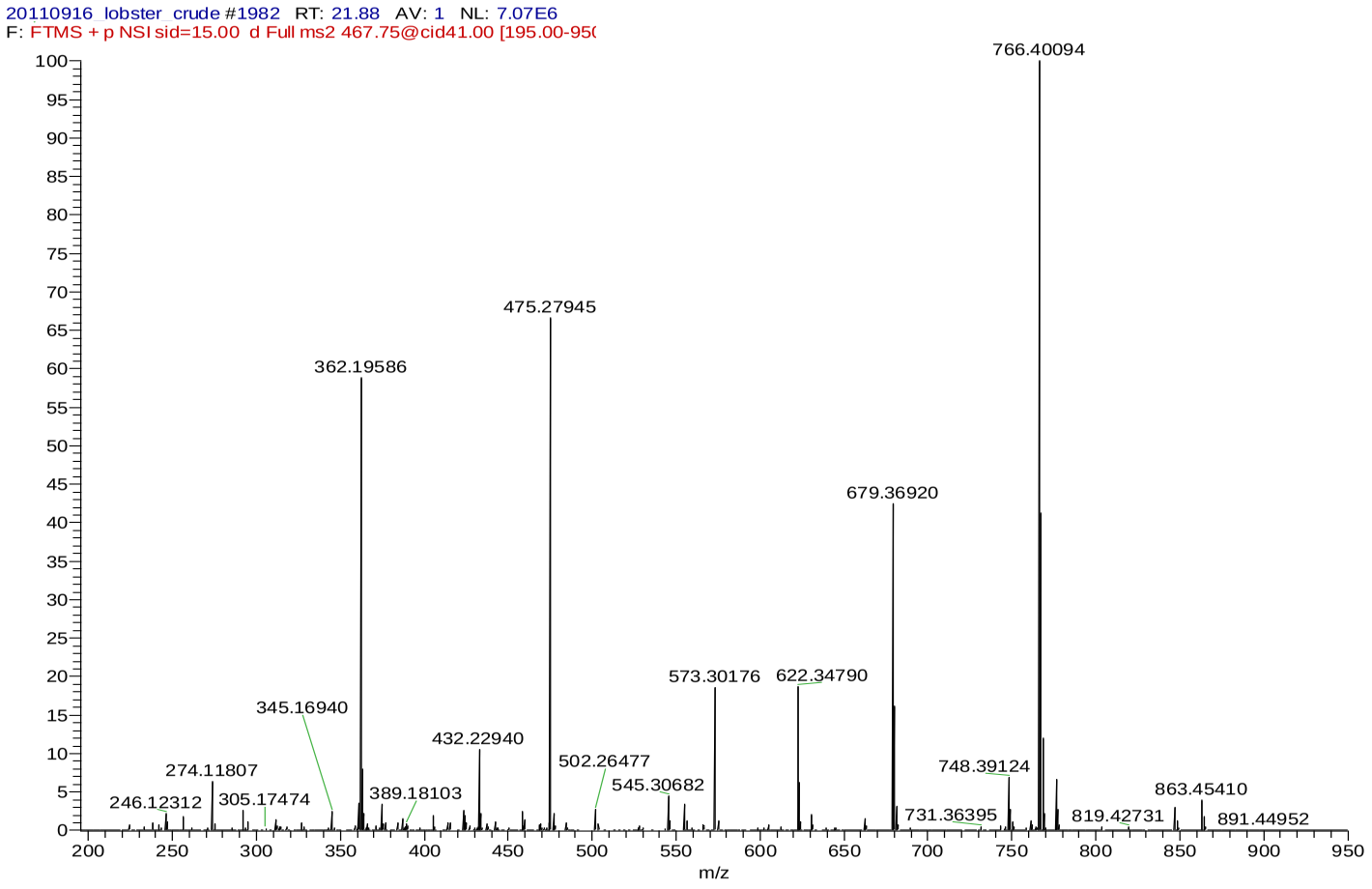


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
5222	3420	PDH	1926.02	1926.024	-2.18066	NSELINSILGLPKVMNDA	NSELINSILGLPKVMNDAa	Amidation	3.46E-33

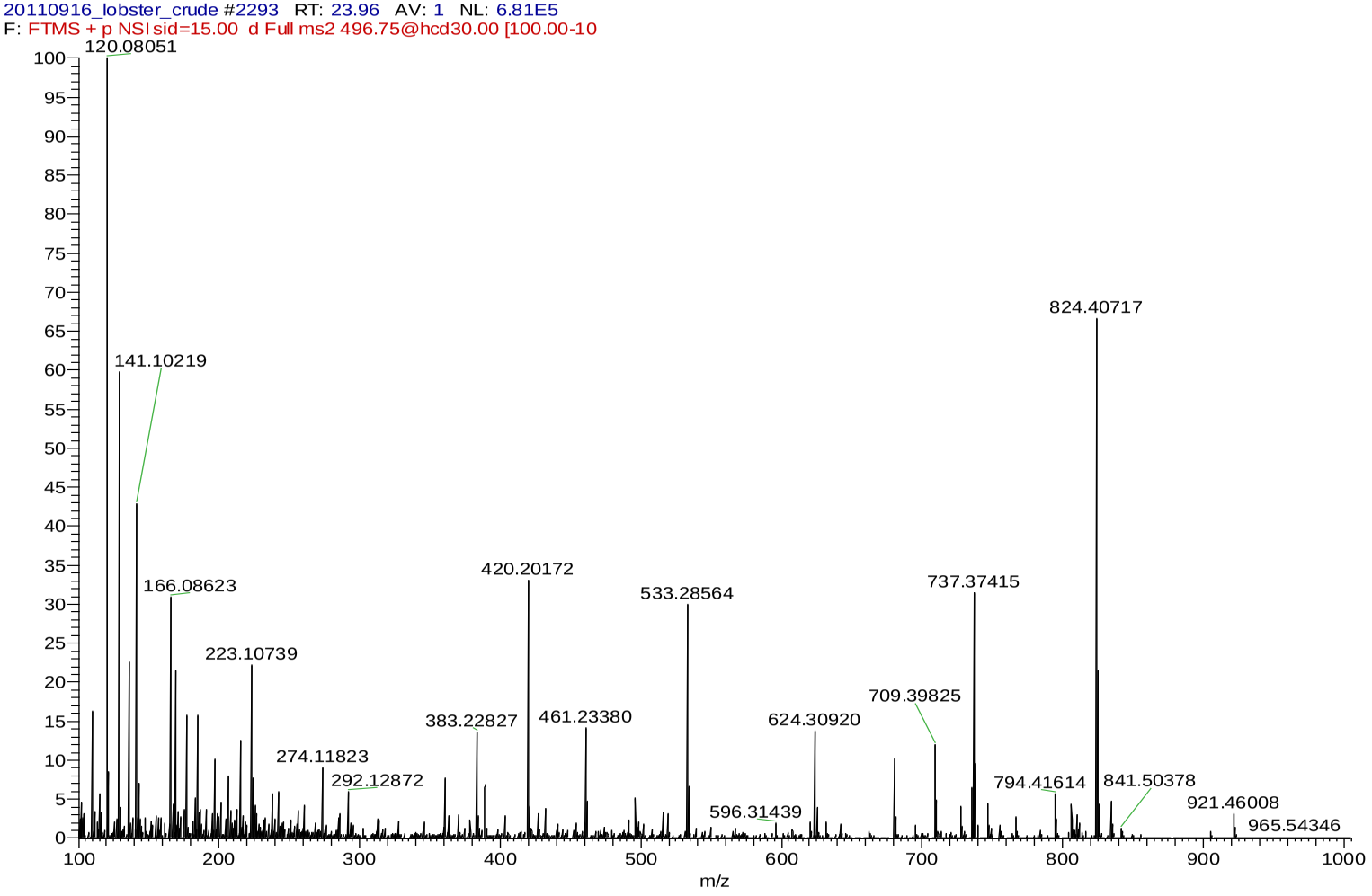
-N-S-E-L-I-N-S-I-L-G-L-P-K-V-M-N-D-A a-



Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
1982	1287	Preprotackykinin	933.4819	933.4849	-3.21376	APSGFLGMR	APSGFLGMRa	Amidation	1.45E-19

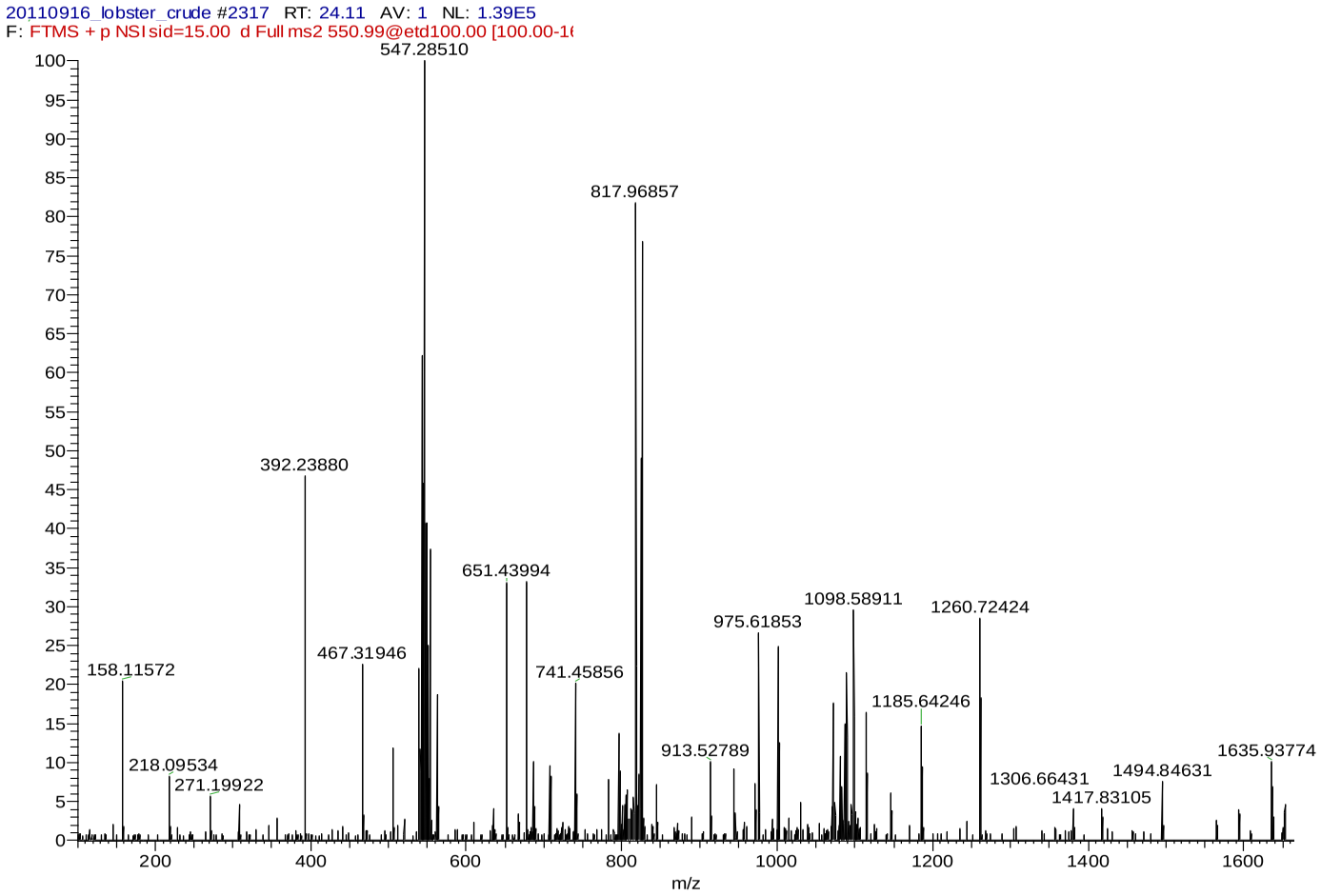


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2293	1483	Preprotackykinin	991.4877	991.4904	-2.72317	APSGFLGMR	APSGFLGMRG		7.91E-16



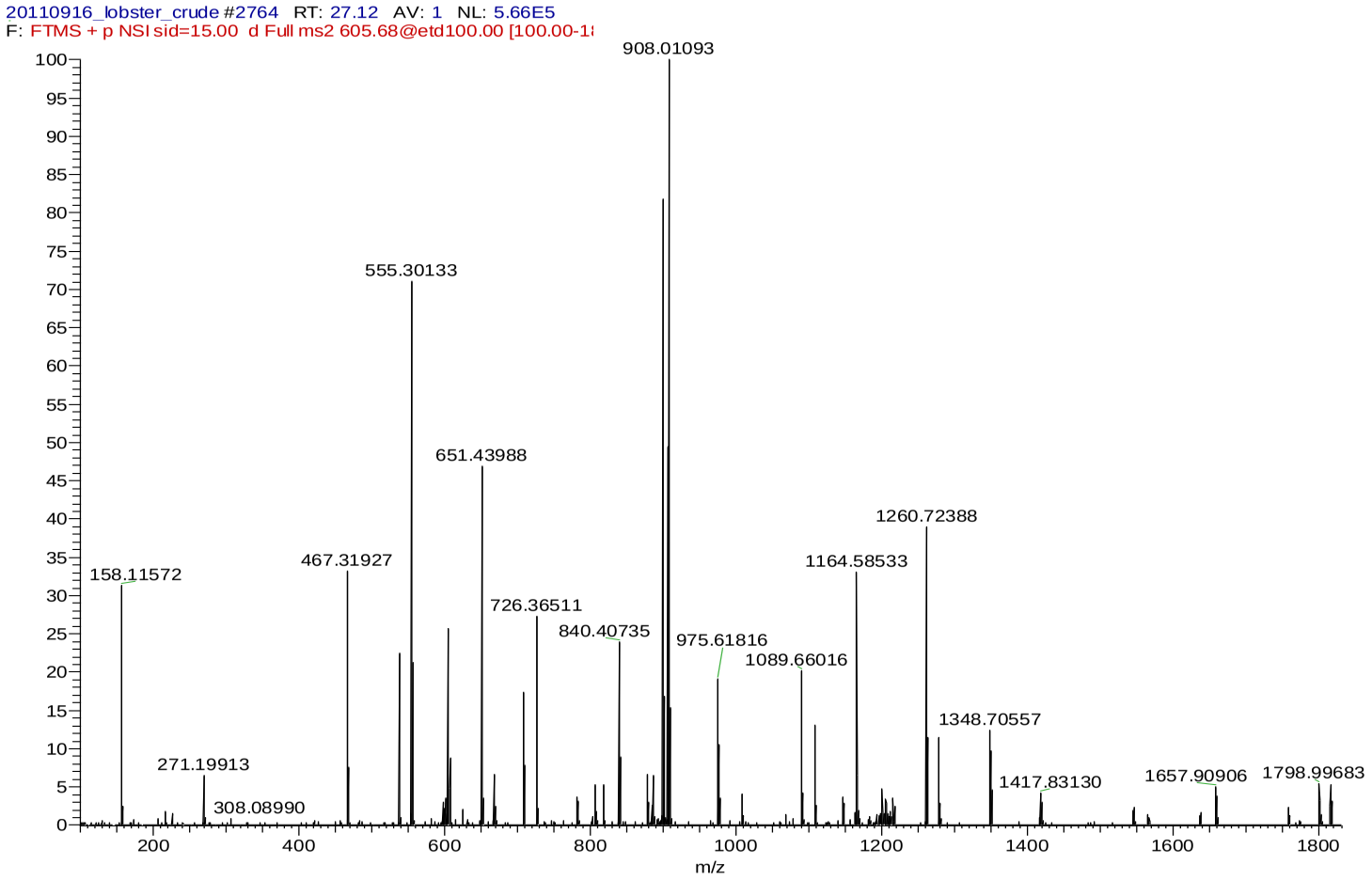
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2317	1493	Preprotackykinin	1649.943	1649.947	-2.48493	AFRNGNAPVGLAVPLR	AFRNGNAPVGLAVPLRa	Amidation	5.98E-32

-A[F[R[N-G-N[A-P[V[G[L[A[V-P[L[R-

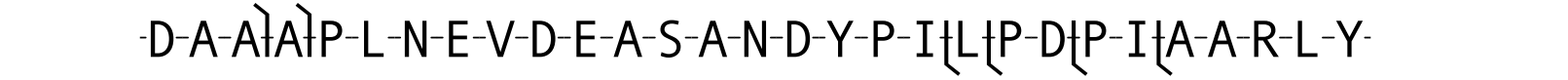


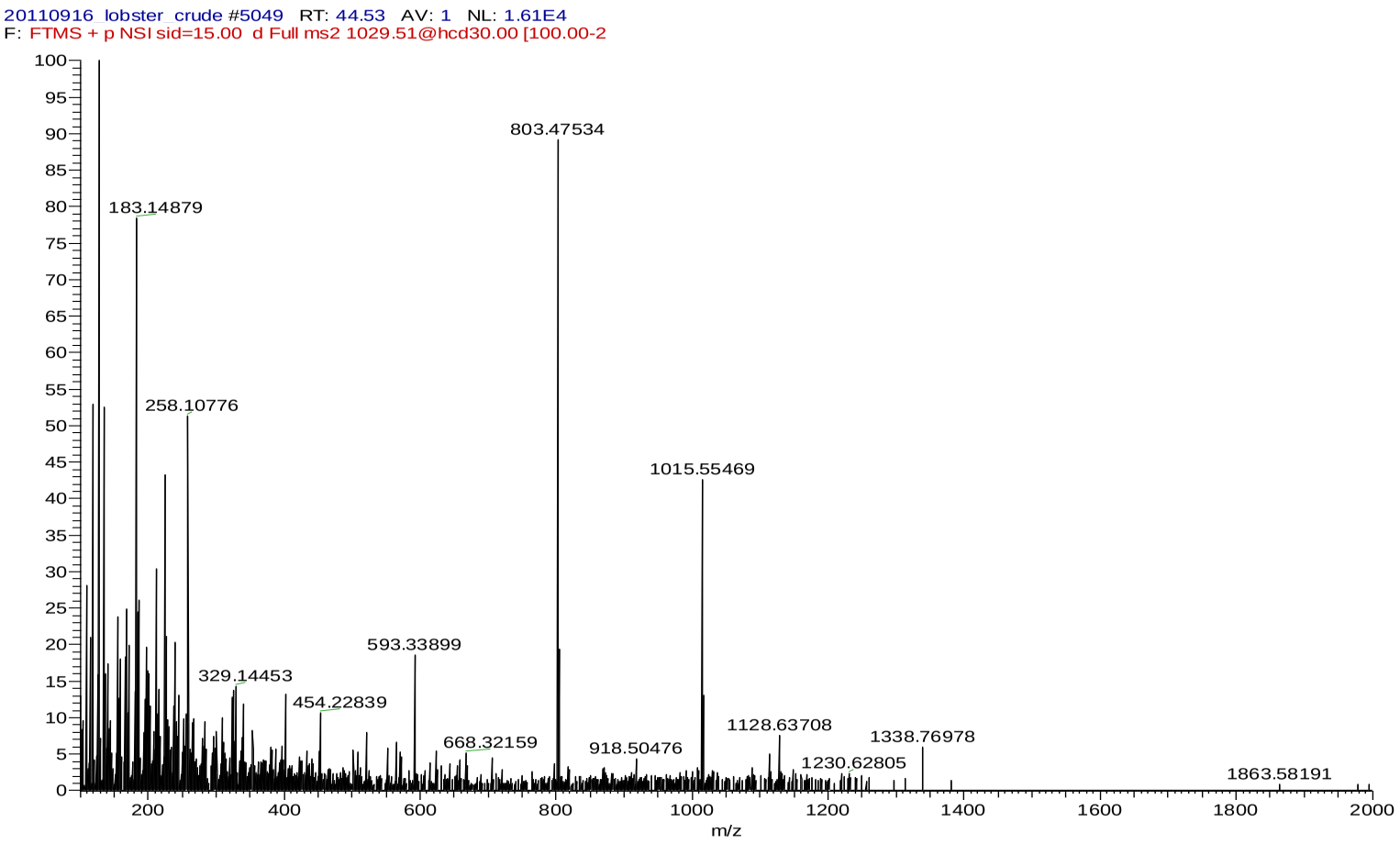
Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
2764	1784	Preprotackykinin	1813.006	1813.011	-2.59237	YAFRNGNAPVGLAVPLR	YAFRNGNAPVGLAVPLRa	Amidation	7.40E-49

-Y[A[F-R[N[G[N[A-P[V[G[L[A[V-P[L[R-

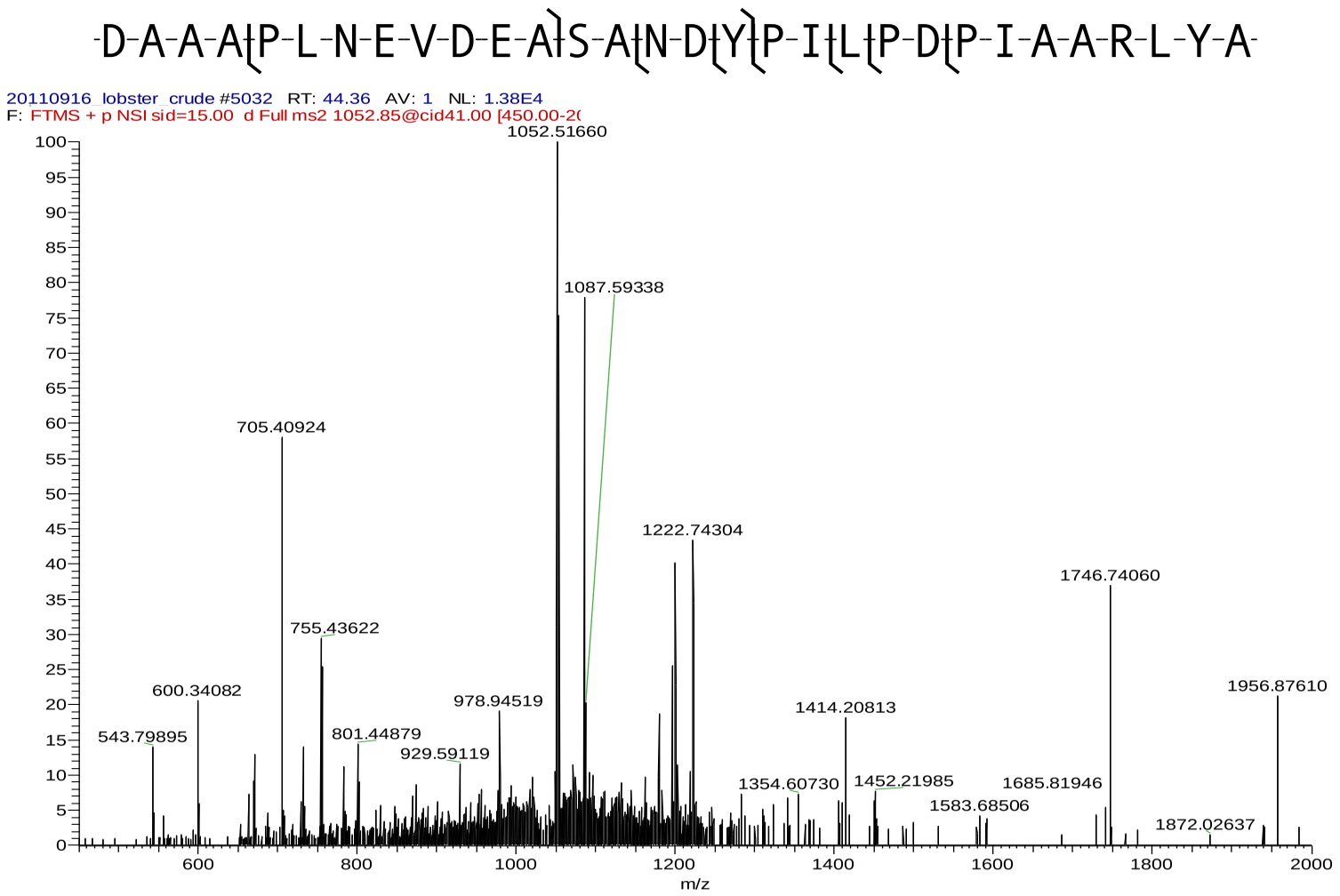


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
5049	3301	Preprotackykinin	3083.498	3083.508	-3.1782	DAAAPLNEVDEASANDYPILPDPIAARLY	DAAAPLNEVDEASANDYPILPDPIAARLY		1.48E-11



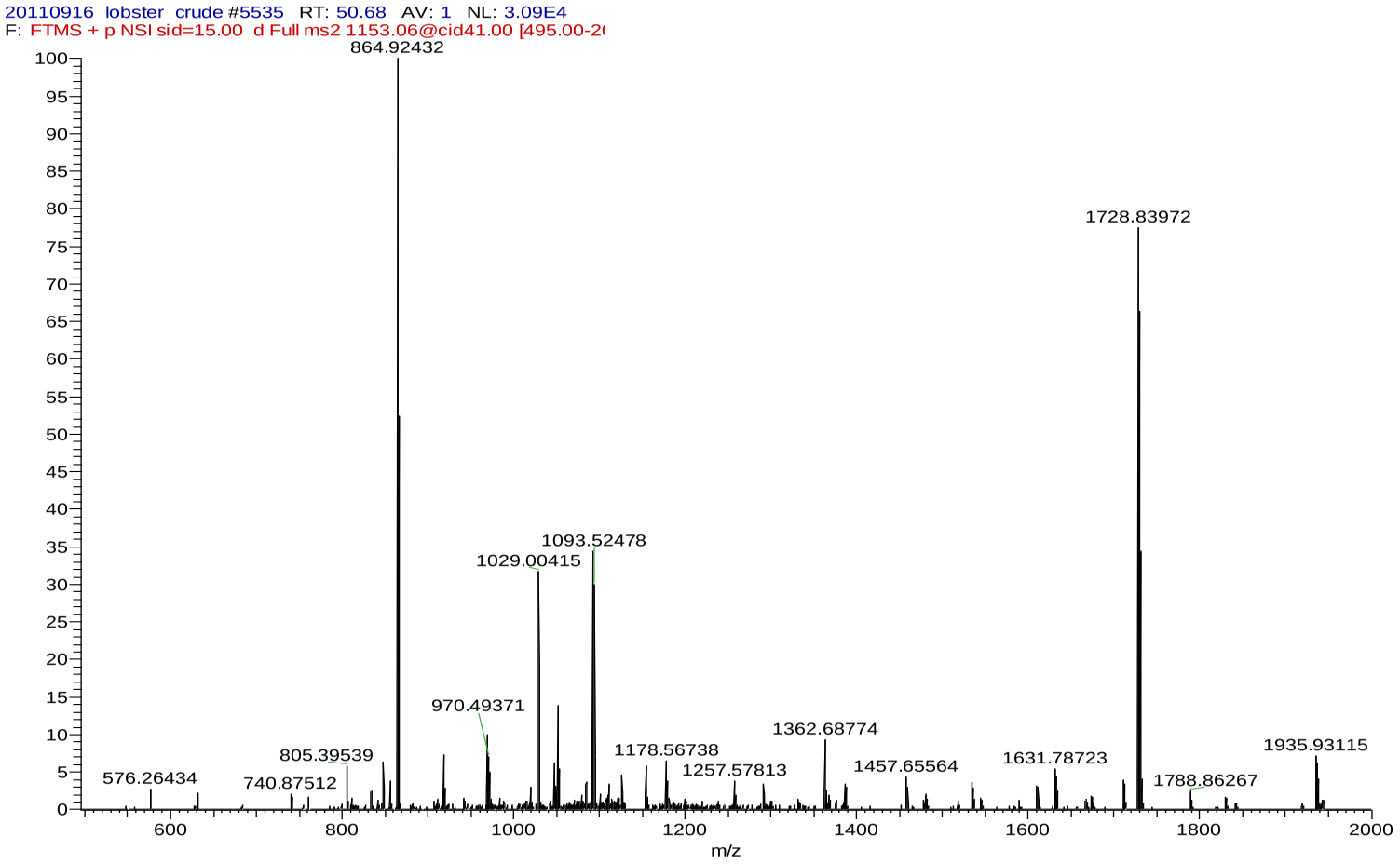


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
5032	3282	Preprotachykinin	3154.543	3154.545	-0.69741	DAAAPLNEVDEASANDYPILPDPIAARLYA	DAAAPLNEVDEASANDYPILPDPIAARLYA		1.24E-10

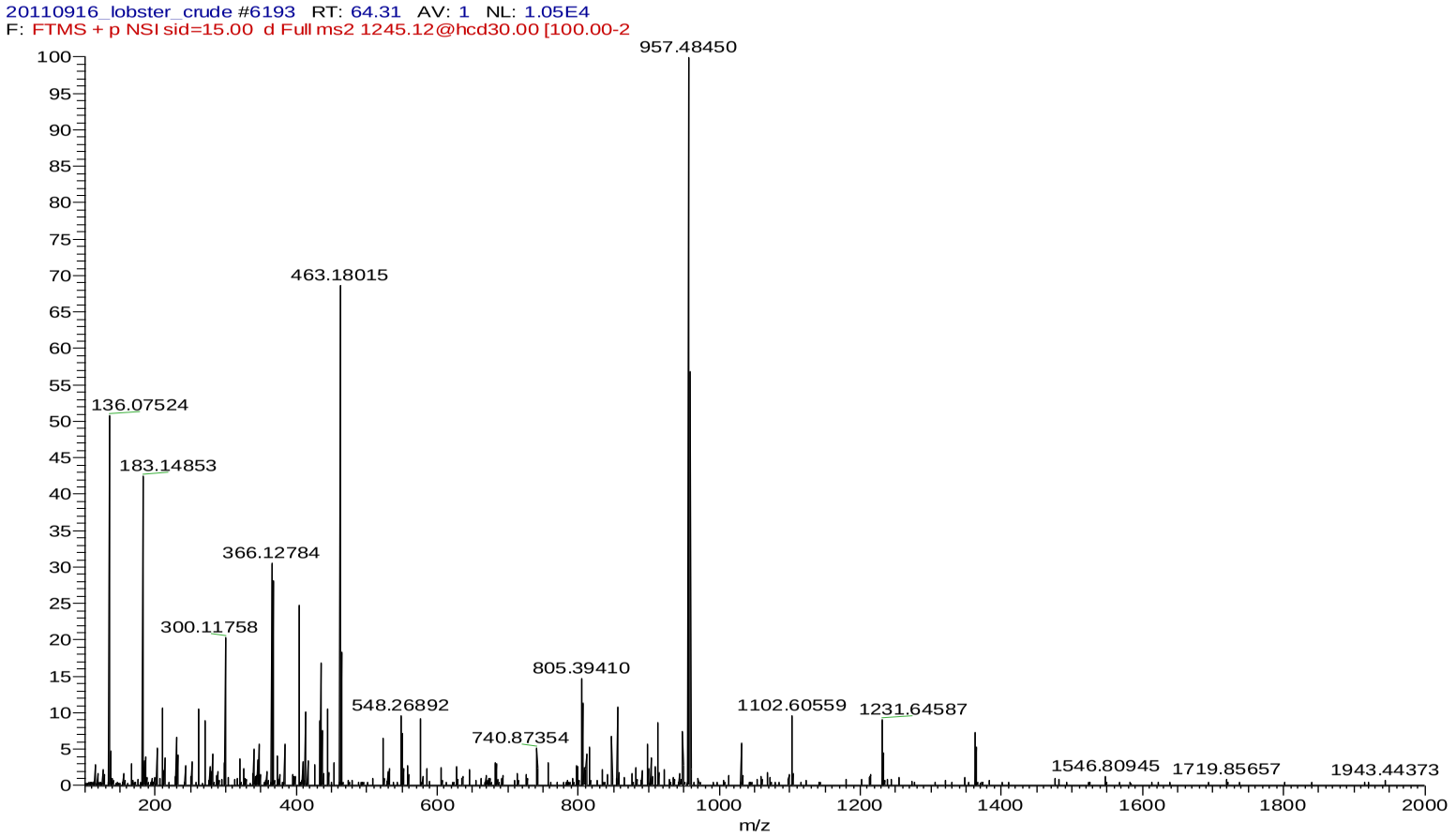


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
5535	3651	Preproneuropeptide F	2303.094	2303.099	-2.12757	SDYAMPSGDALMEASERLLET	SDYPLPPGDALMEASERLLET		2.44E-38

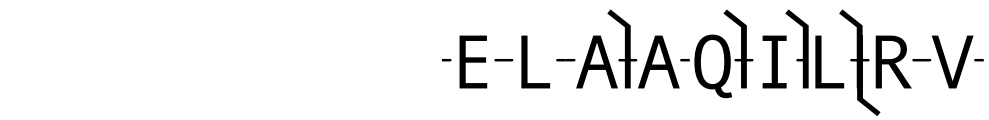
-S-D[Y[P-L[P[P[G[D[A[L[M[E[A[S[E[R[L[L[E[T-

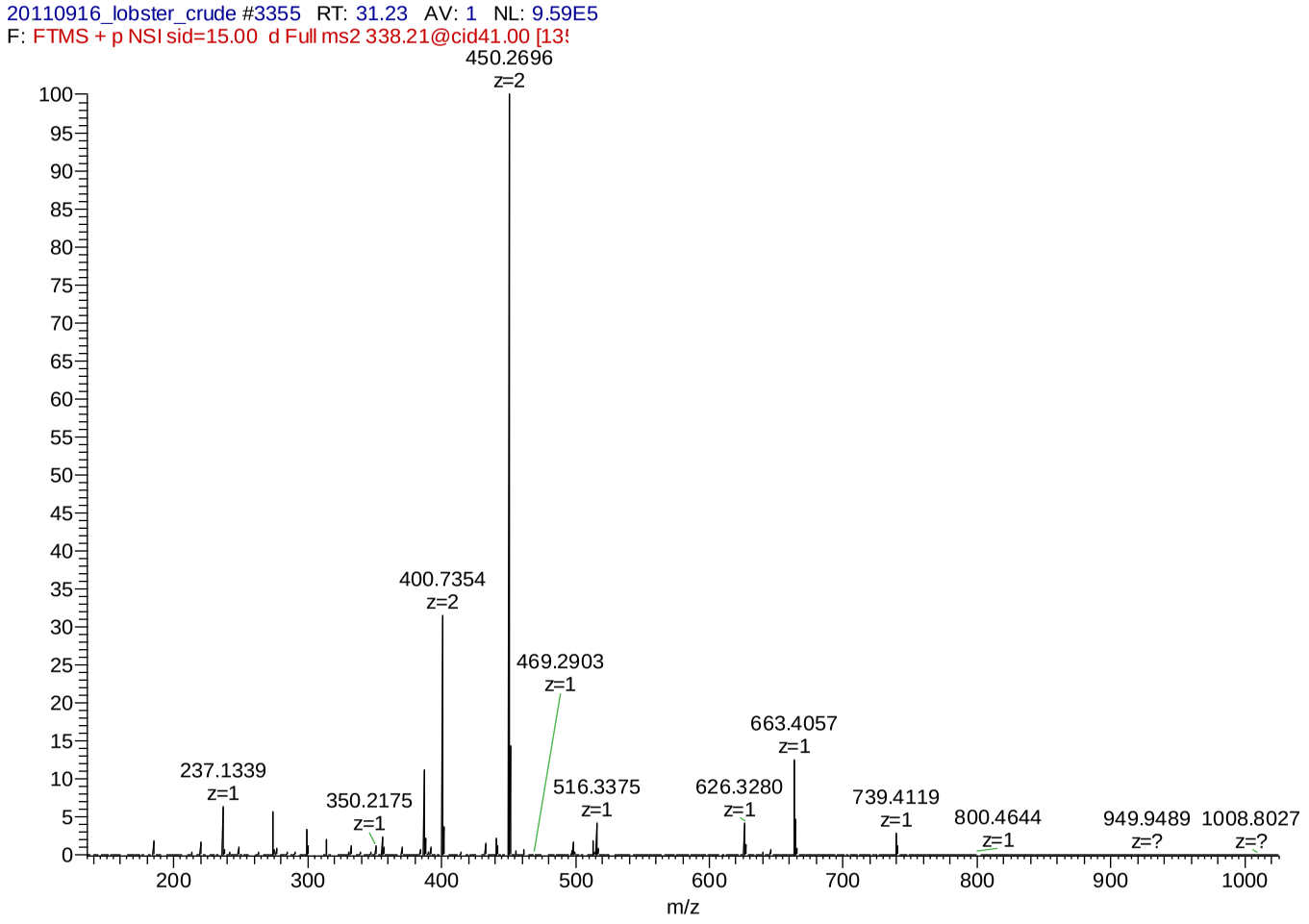


Scan #	Exp ID	Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	Reported Sequence	De novo Sequence	PTM	E-value
6193	4110	Preproneuropeptide F	2487.216	2487.22	-1.60822	SDYPMPSGDALMEASERLLETLA	SDYPLPPGDALMEASERLLETLA		3.83E-21



Peptide	Observed Mass	Exp'd mass	Mass Difference (PPM)	De novo Sequence
PDH	1011.6071	1011.6079	-0.7908	ELAAQILRV





Appendix III

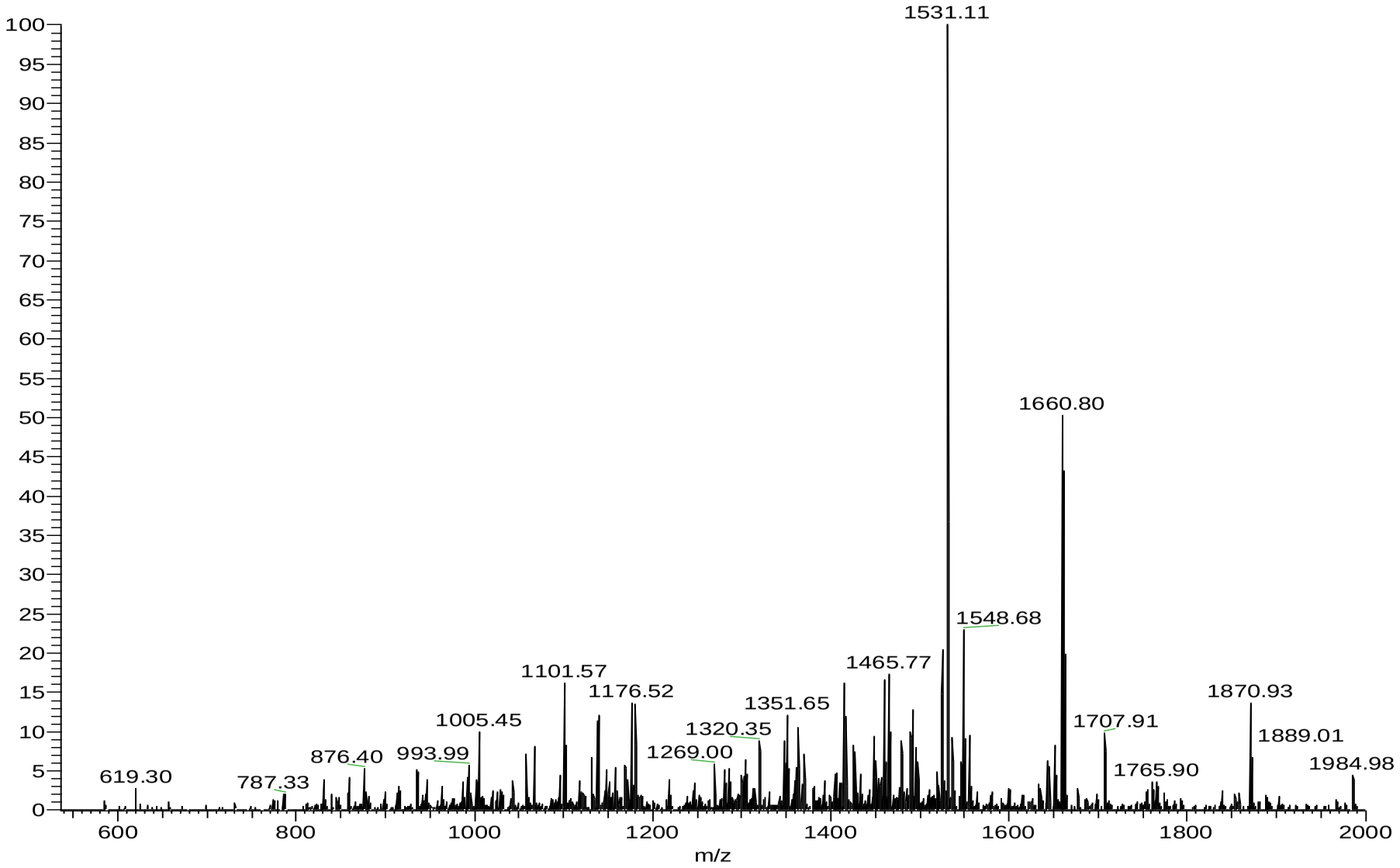
MS/MS Spectra and Fragmentation Maps of Proteins in Chapter 7

334

4964	TYB10_RAT	P63312	Thymosin β-4	4960.48	4960.49	-2	N-term Acetylation	5.0E-77

b1 - S - D - K - P - D - M - A - E - I - E - K - F - D - K - S - K - L - K - K - T - E - T - Q - E - K y19

b26 - N - P - L - P - S - K - E - T - I - E - Q - E - K - Q - A - G - E - S - y1



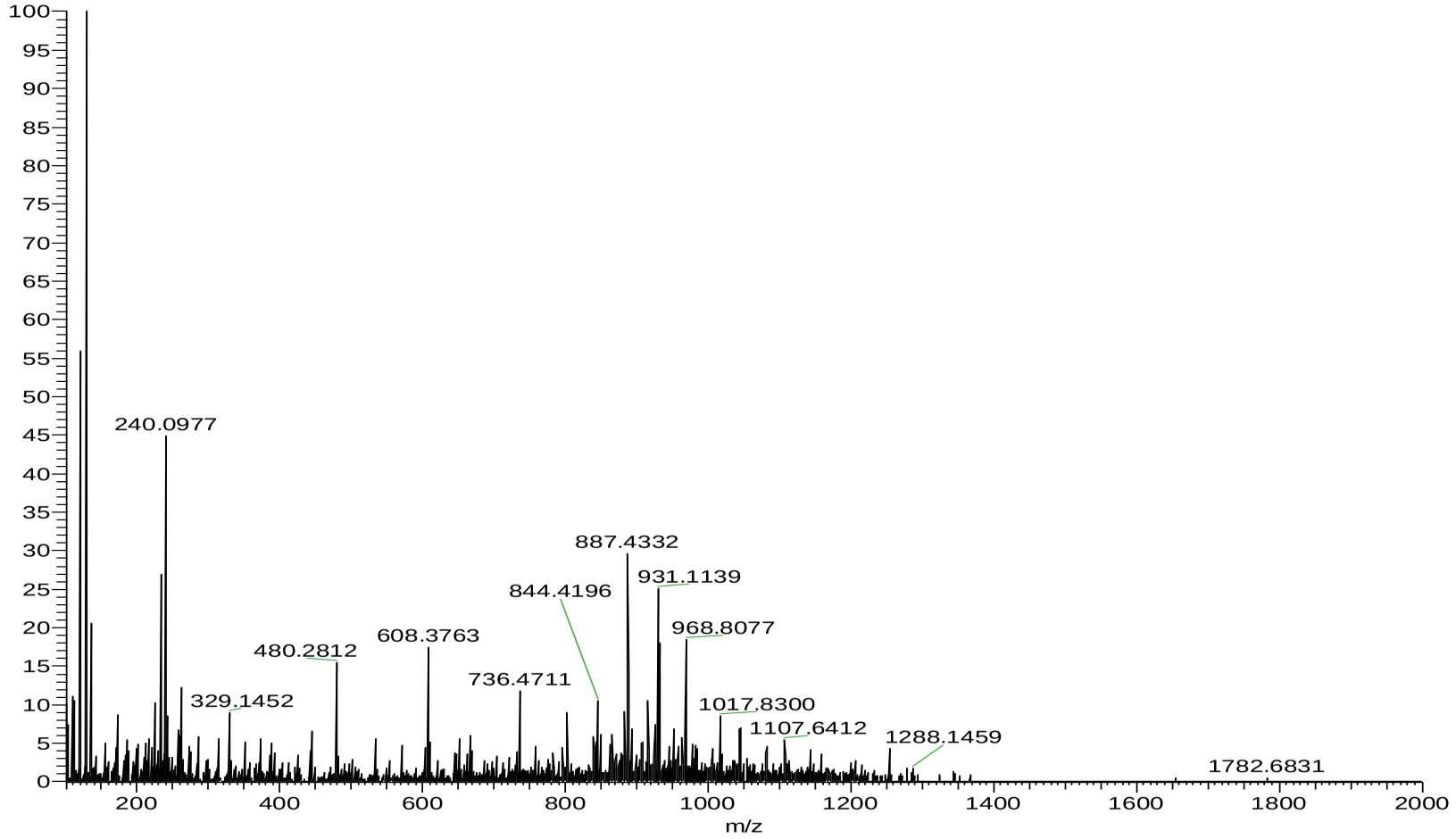
9939	ACBP_RAT	P11030	Acyl-CoA-binding protein	9932.12	9932.12	0	N-term acetylation	1.2E-44

b1 - S - Q - A - D - F - D - K - A - A - E - E - V - K - R - L - K - T - Q - P - T - D - E - E - M - L - y62

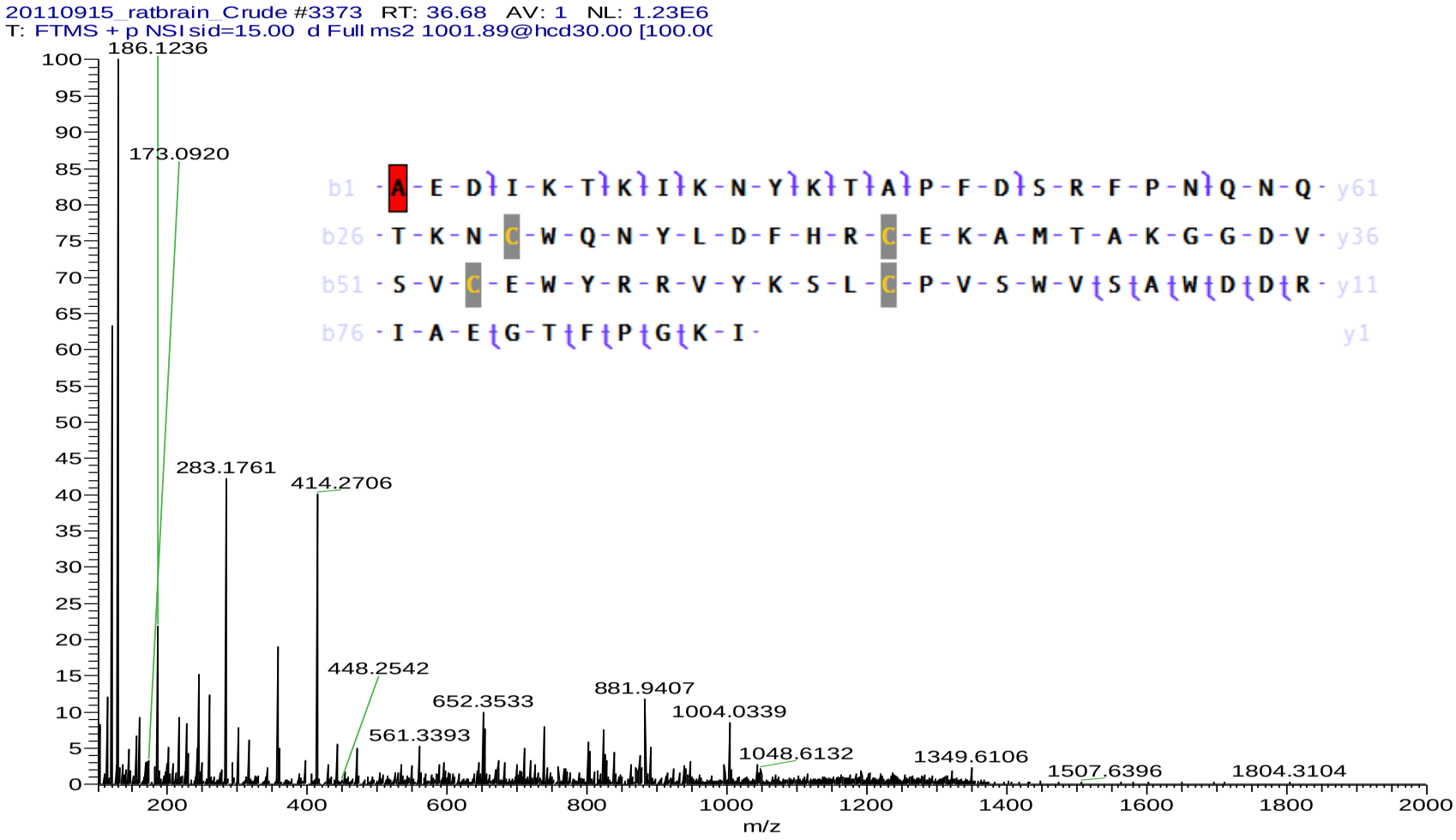
b26 - F - I - Y - S - H - F - K - Q - A - T - V - G - D - V - N - T - D - R - P - G - L - L - D - L - K - y37

b51 - G - K - A - K - W - D - S - W - N - K - L - K - G - T - S - K - E - N - A - M - K - T - Y - V - E - y12

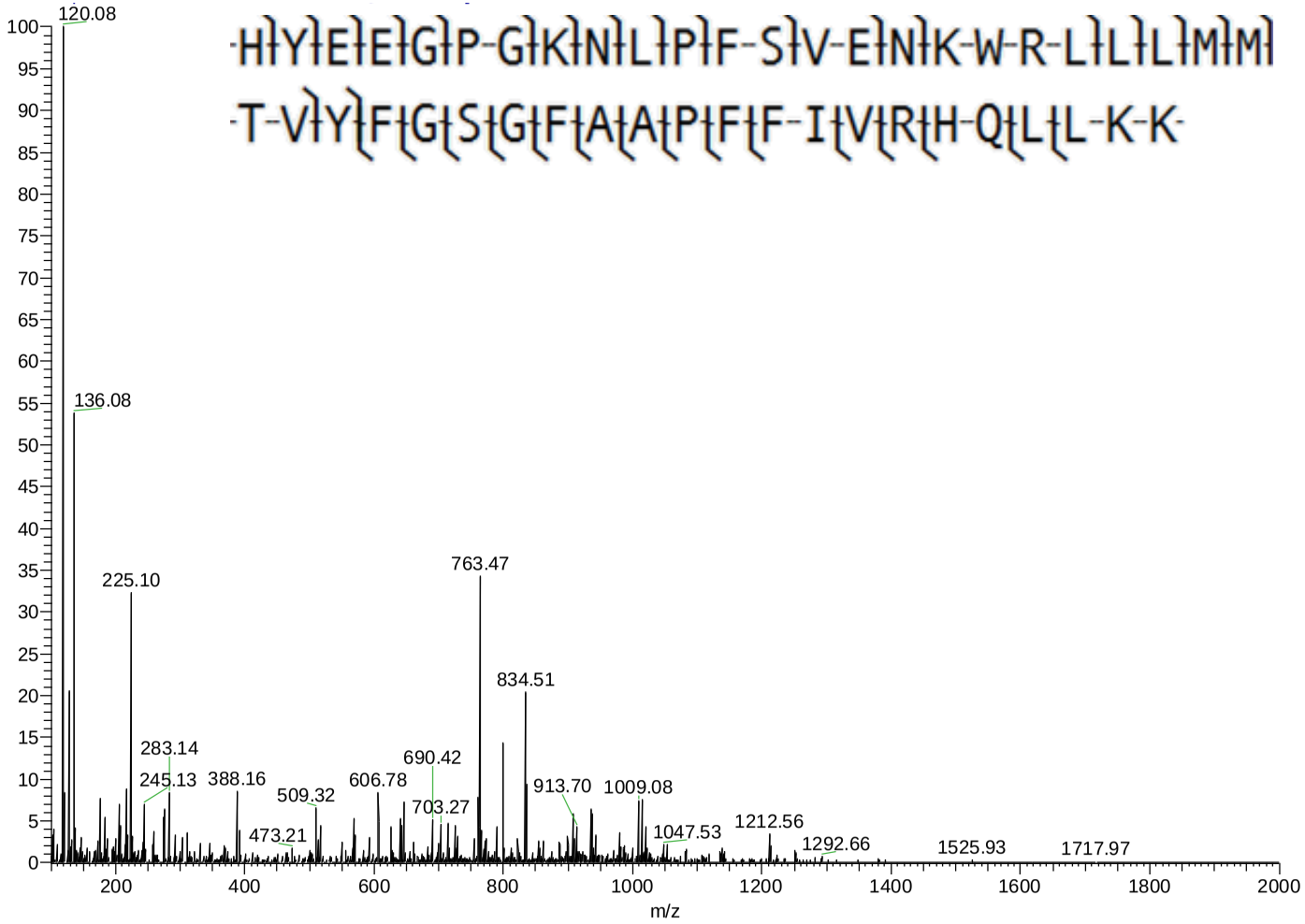
b76 - K - V - E - E - L - K - K - K - Y - G - I - y1



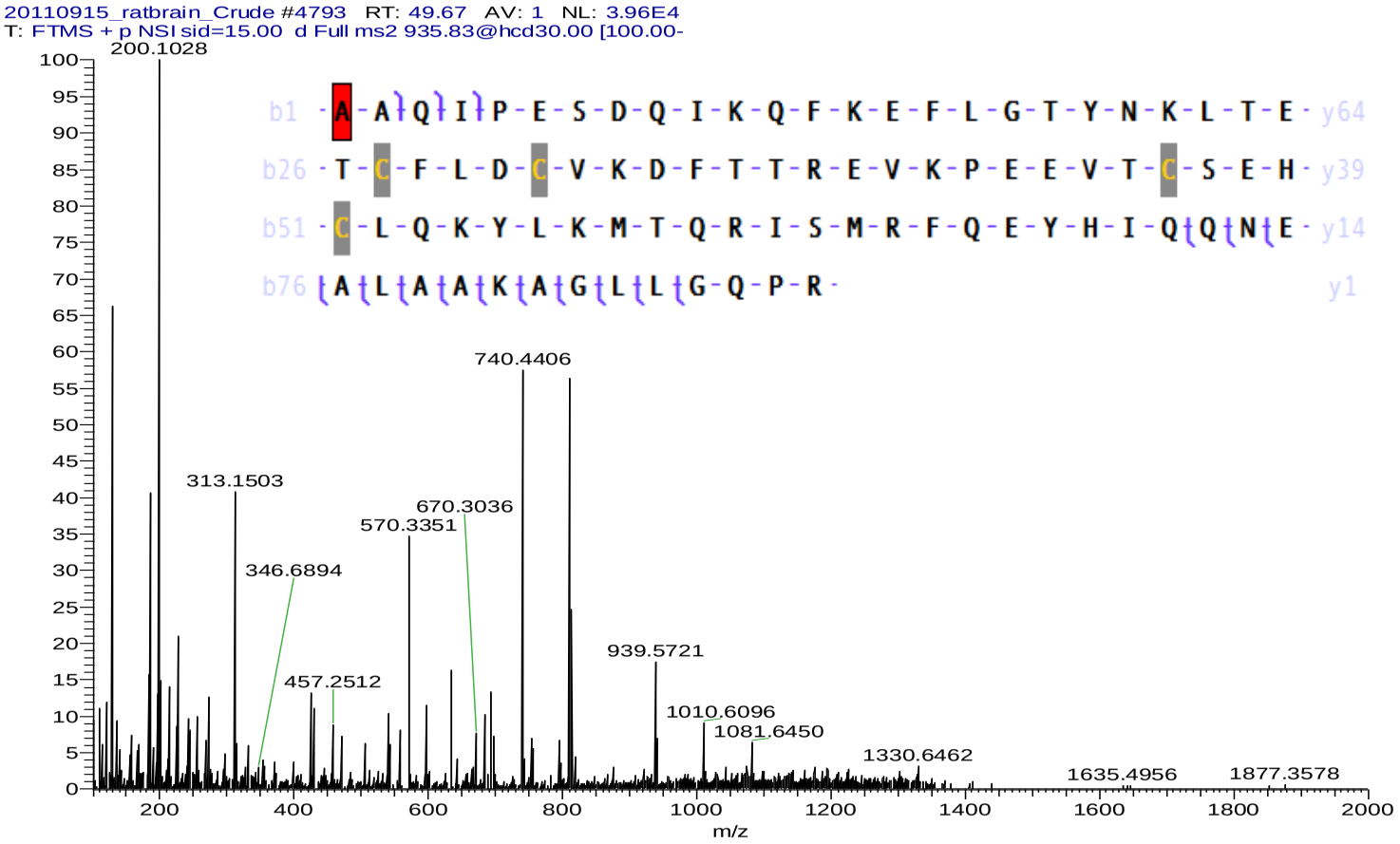
9979	D3ZD09_RAT	D3ZD09	cytochrome c oxidase polypeptide VI b	9971.80	9971.82	-2	N-term acetylation, 2 disulfide bonds	1.8E-41



5486	B2RYT3_RAT	B2RYT3	Cox7c protein	5481.87	5481.87	0		1.5E-61



10283	TIM9_RAT	Q9WV97	Mitochondria import inner membrane translocase	10276.08	10276.09	-1	N-term acetylation, 2 disulfide bonds	8.0E-24



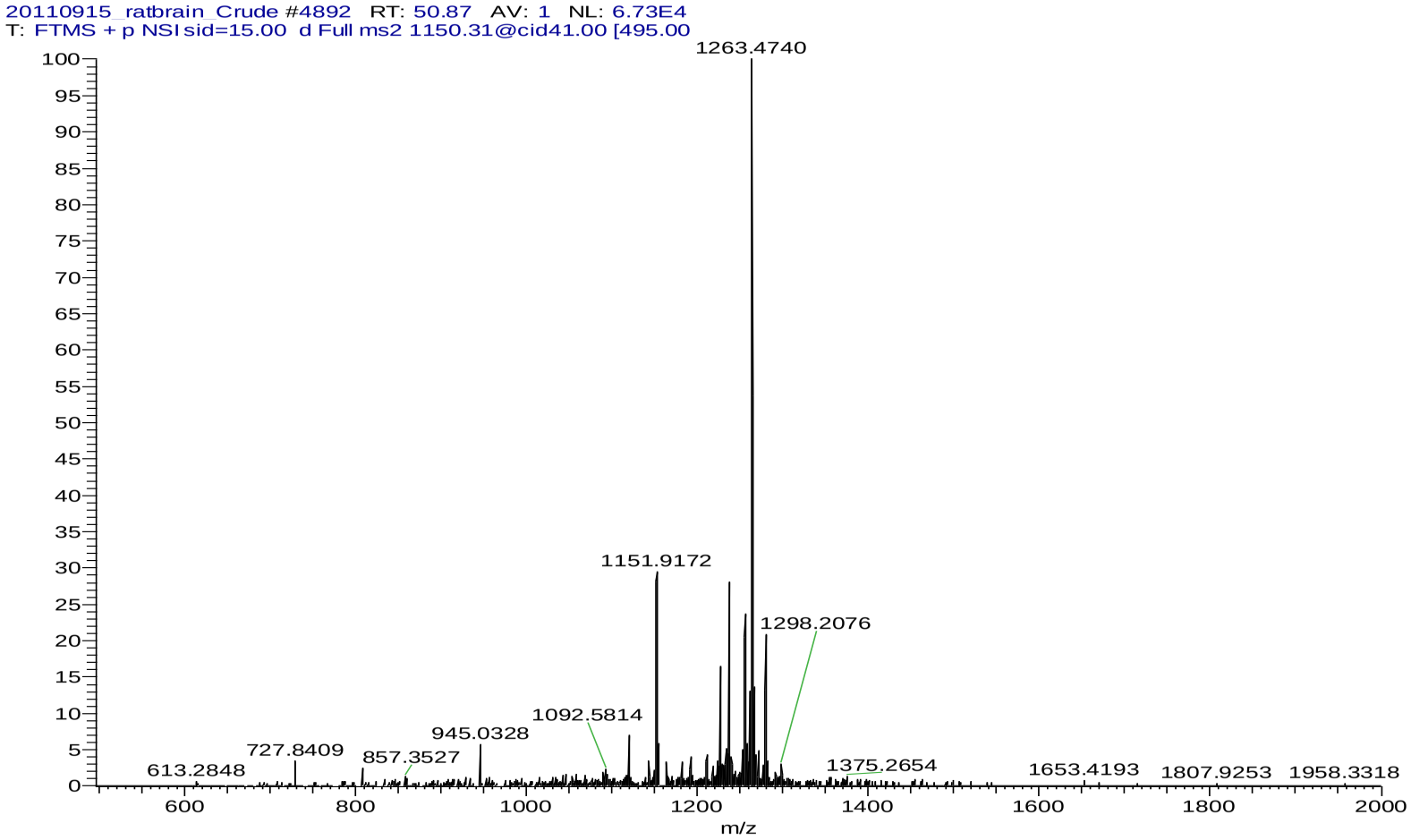
9193	TIM8B_RAT	P62078	Adenylate cyclase type 10	9187.44	9187.45	-1	N-term acetylation, 2 disulfide bonds	2.3E-25

b1 - A - E - L - G - E - A - D - E - A - E - L - Q - R - L - V - A - A - E - Q - Q - K - A - Q - F - T - y58

b26 - A - Q - V - H - H - F - M - E - L - C - W - D - K - C - V - E - K - P - G - S - R - L - D - S - R - y33

b51 - T - E - N - C - L - S - S - C - V - D - R - F - I - D - T - T - L - A - I - T - G - R - F - A - Q - y8

b76 - I - V - Q - K - G - G - Q - y1



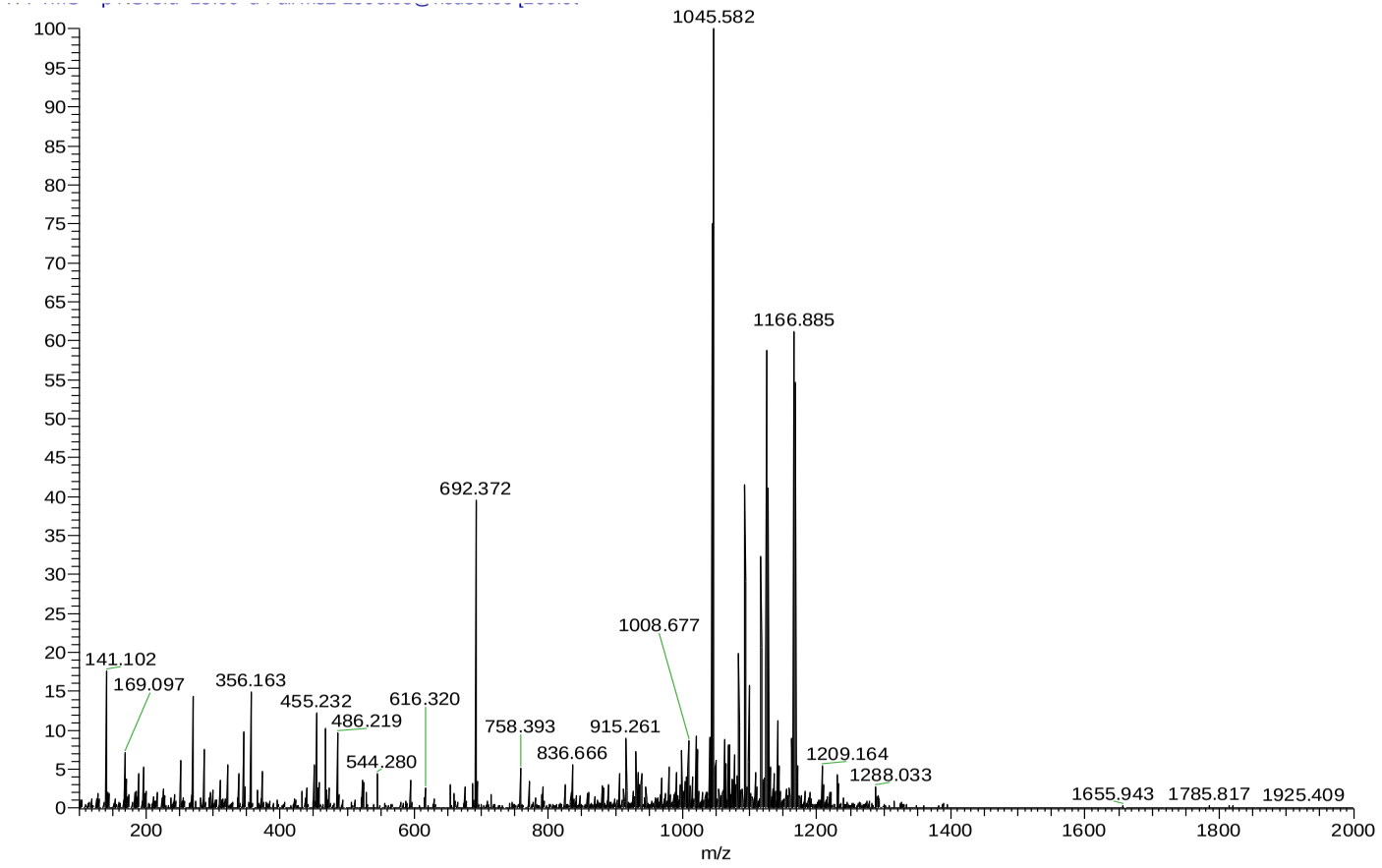
8040	UCRI_RAT	P20788	Cytochrome b-c1 complex subunit	8035.38	8035.38	0	N-term acetylation	4.4E-39

b1 - M - L - S - V - A - A - R - S - G - P - F - A - P - V - L - S - A - T - S - R - G - V - A - G - A - y54

b26 - L - R - P - L - L - Q - S - A - V - P - A - T - S - E - P - P - V - L - D - V - K - R - P - F - L - y29

b51 - C - R - E - S - L - S - G - Q - A - A - T - R - P - L - V - A - T - V - G - L - N - V - P - A - S - y4

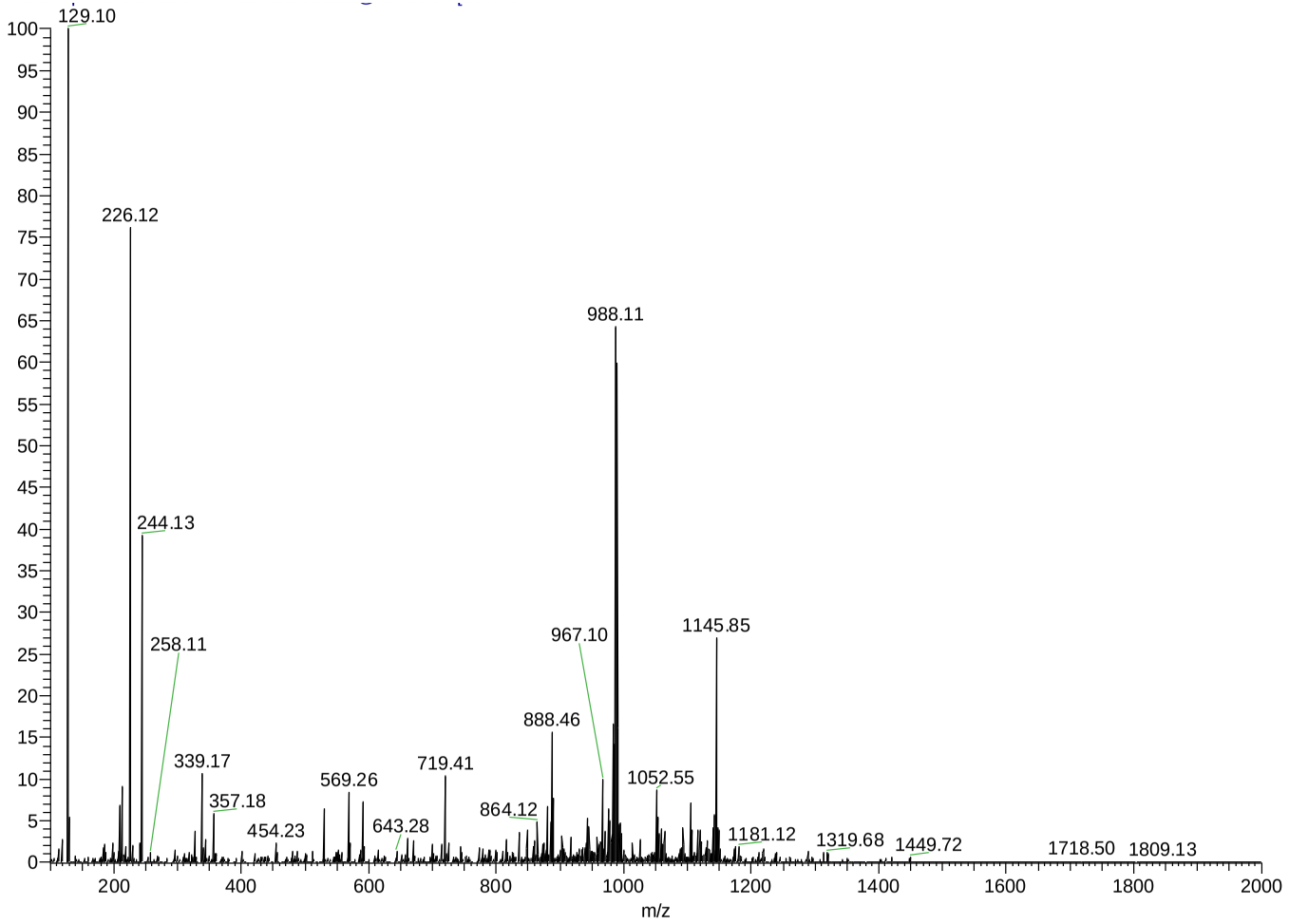
b76 - V - R - Y - y1



4936	TYB10_RAT	P63312	Thymosin beta-10	4933.51	4933.52	-2	N-term acetylation	7.5E-87

b1 -A-D-K-P-D-M-G-E-I-A-S-F-D-K-A-K-L-K-K-T-E-T-Q-E-K- y19

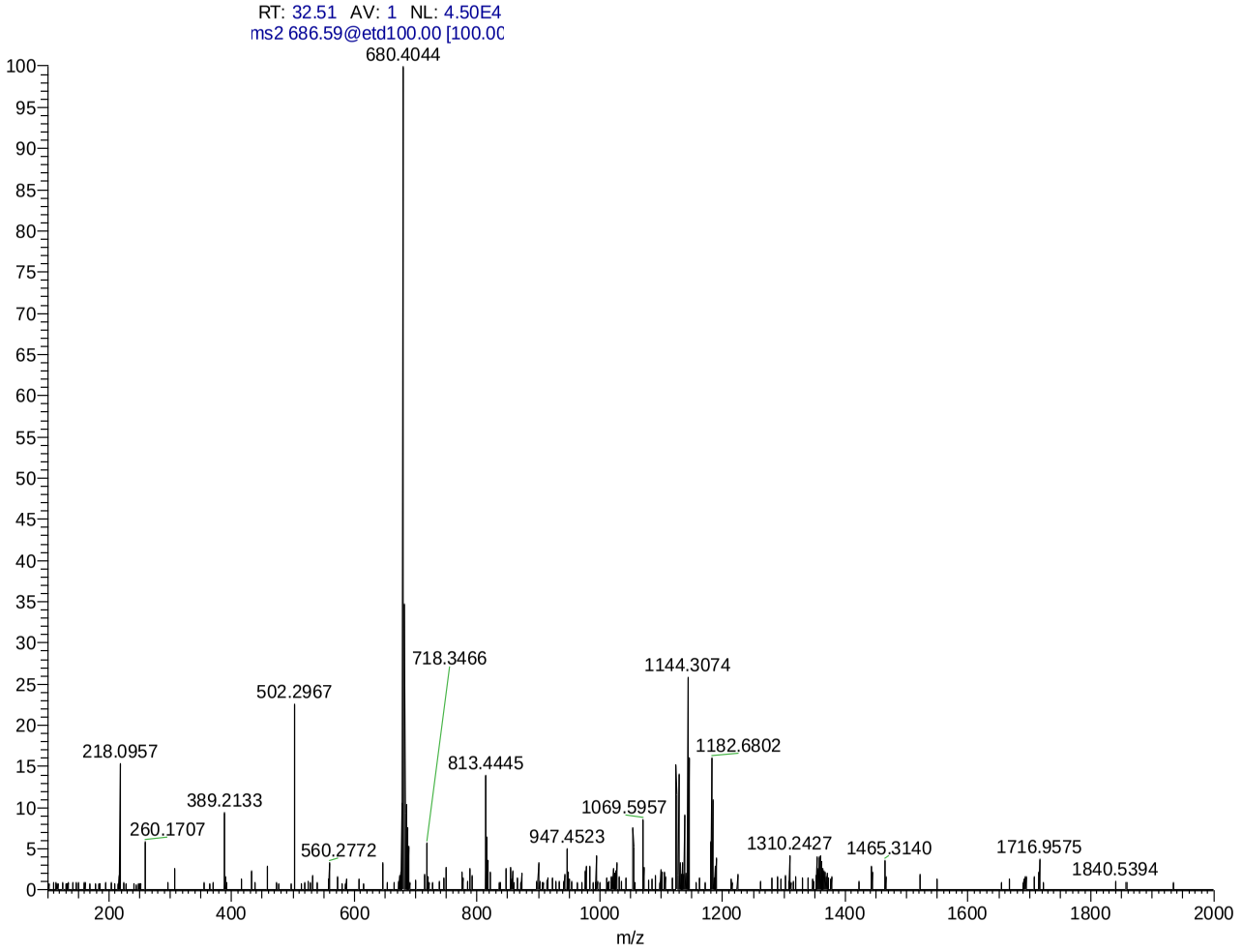
b26 -N-T-L-P-T-K-E-T-I-E-Q-E-K-R-S-E-I-S- y1



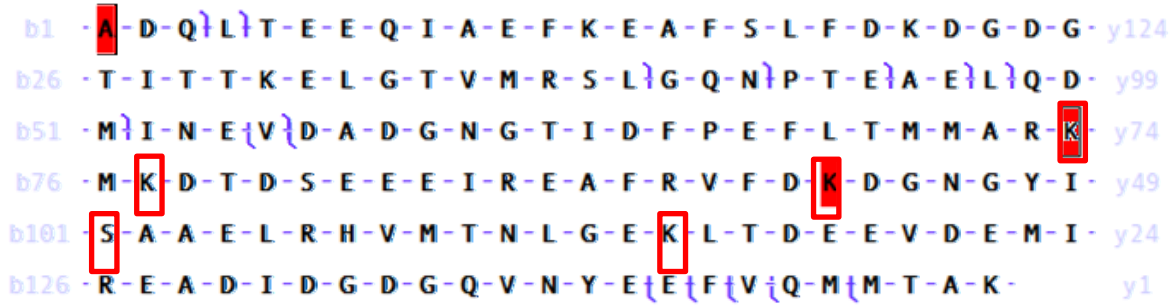
3430	ATP5J_RAT	P21571	ATP synthase-coupling factor 6	3426.89	3426.90	-3		4.3E-13

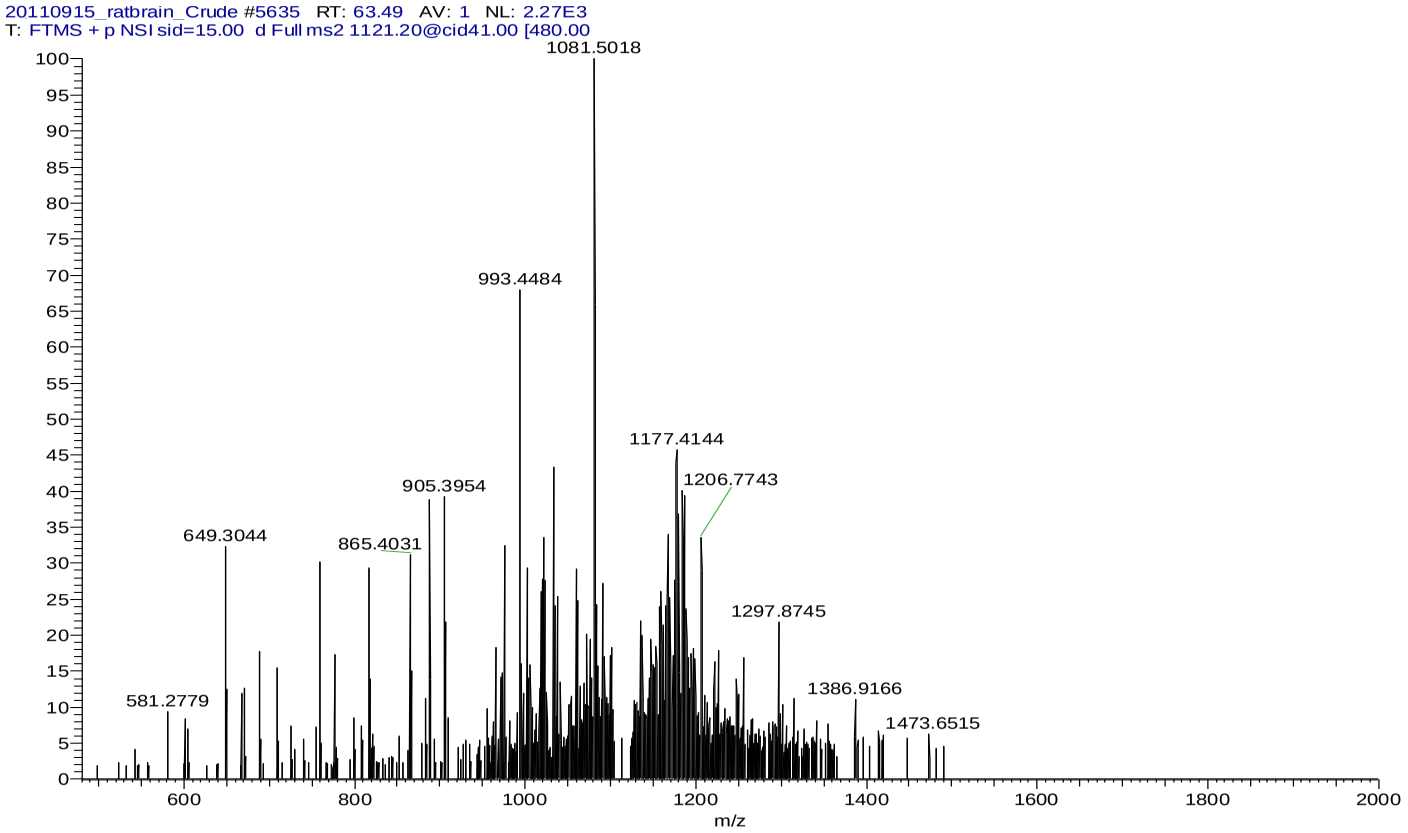
c1 -N-K-E-L-D-P-V-Q-K-L-F-L-D-K-I-R-E-Y-K-A-K-R-L-A-S- z6

c26 -G-G-P-V-D- z1



16792	CALM_RAT	P62161	Calmodulin	16779.82	16779.81	1	N-term acetylation, unfixed acetylation	1.3E-10

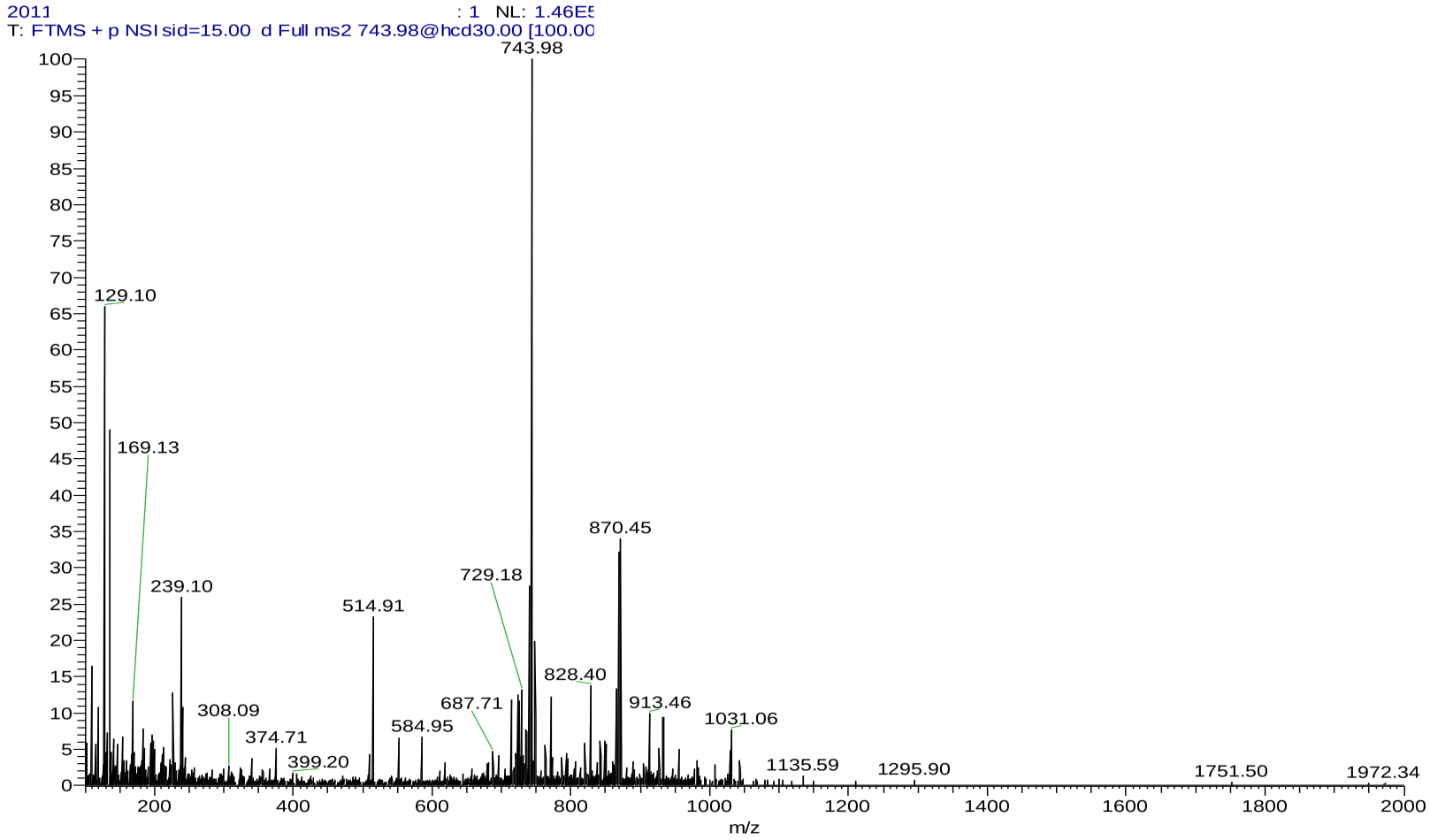




3717	Q5U2U9_RAT	Q5U2U9	CCR4-NOT transcription complex, subunit 8	3713.87	3713.88	-3		7.9E-06

b1 - E - E - E - M - R - K - I - R - E - T - V - L - S - Y - S - Y - I - A - M - D - T - E - F - P - G - y8

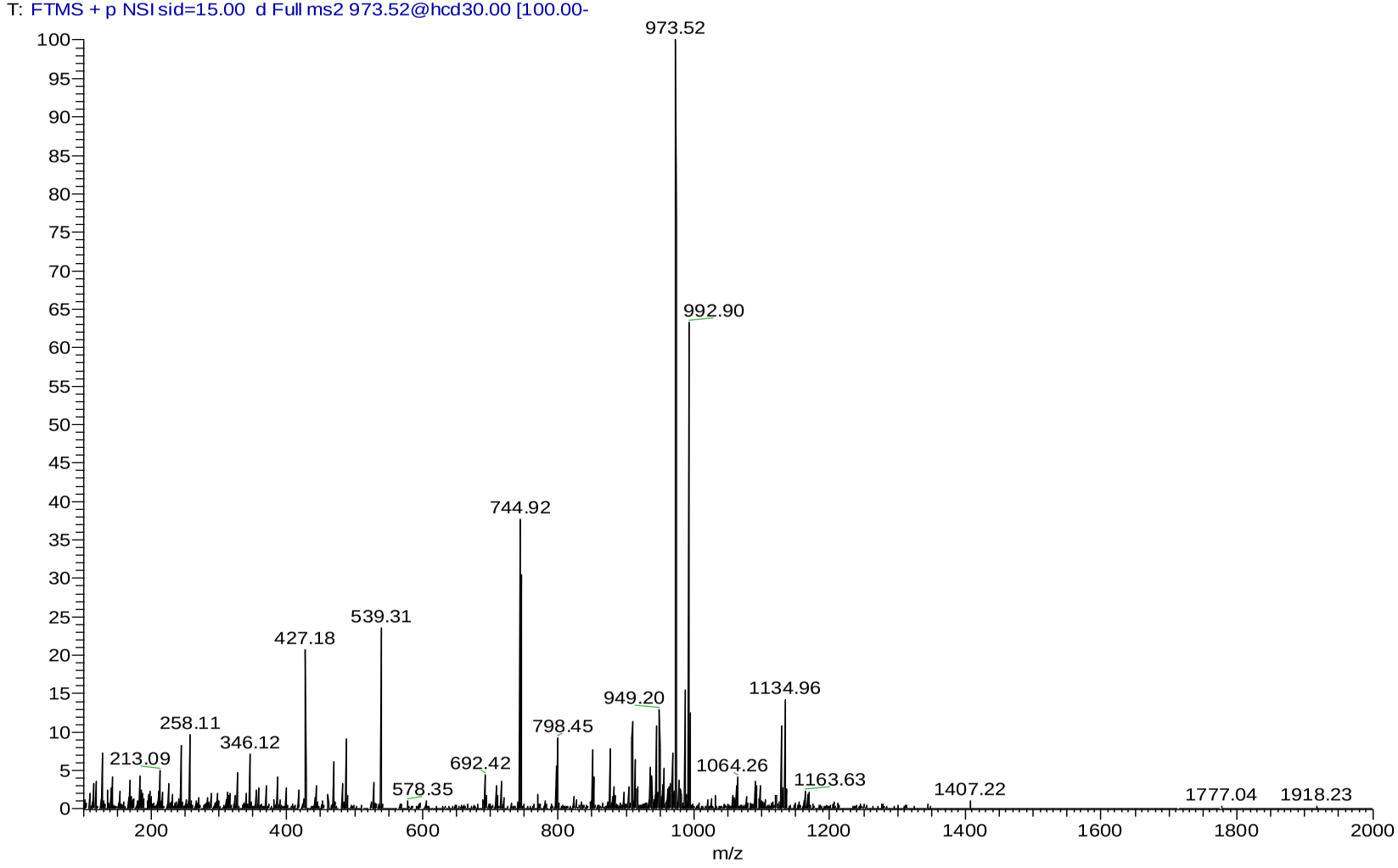
b26 - V - V - V - R - P - I - G - y1



3891	PCSK1_RAT	Q9QXU9	ProSAAS	3888.04	3888.04	0		1.5E-33

b1 - A | A | D | E | T | P | D | V | D | P - E | L - L - R | Y - L - L - G - R | I - L - T | G - S - S - y12

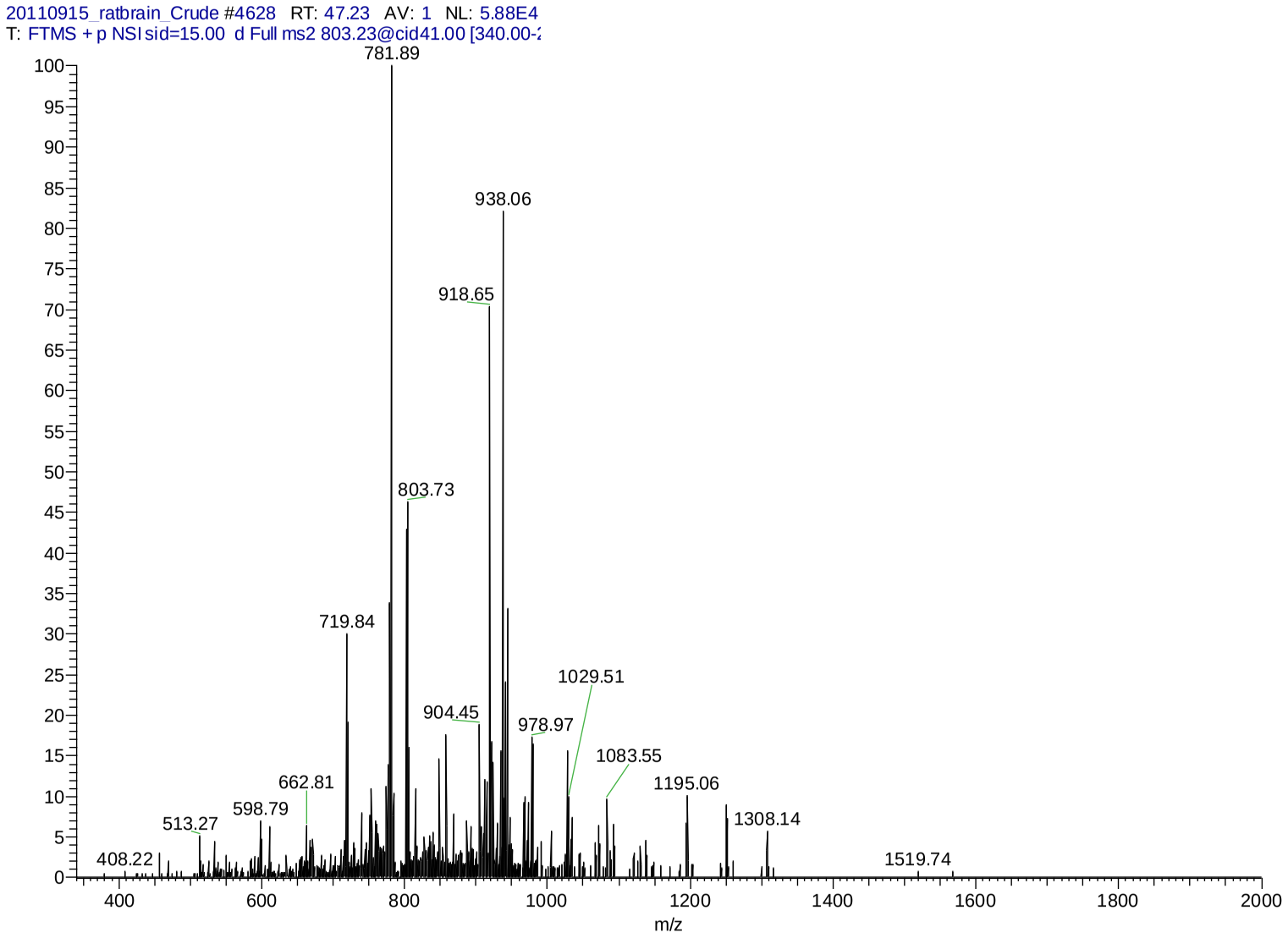
b26 - E | P - E | A | A | P - A - P - R - R | L - y1



4800	SCG2_RAT	P10362	Secretogranin-2	4796.37	4796.37	0		5.8E-72

b1 -I-P-A-G-S-L-K-N-E-D-T-P-N-R-Q-Y-L-D-E-D-M-L-L-K-V- y17

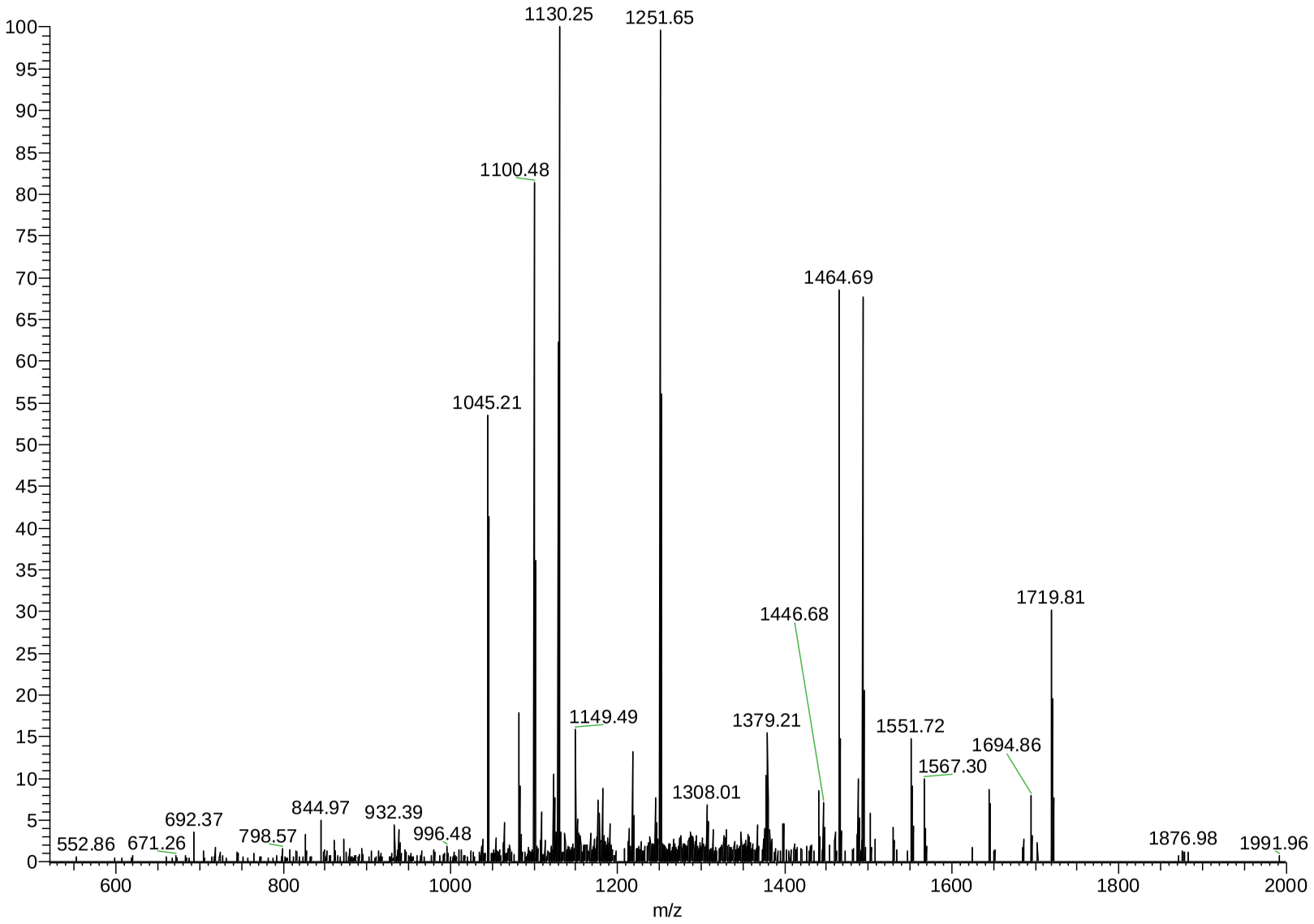
b26 -L-E-Y-L-N-Q-E-Q-A-E-Q-G-R-E-H-L- y1



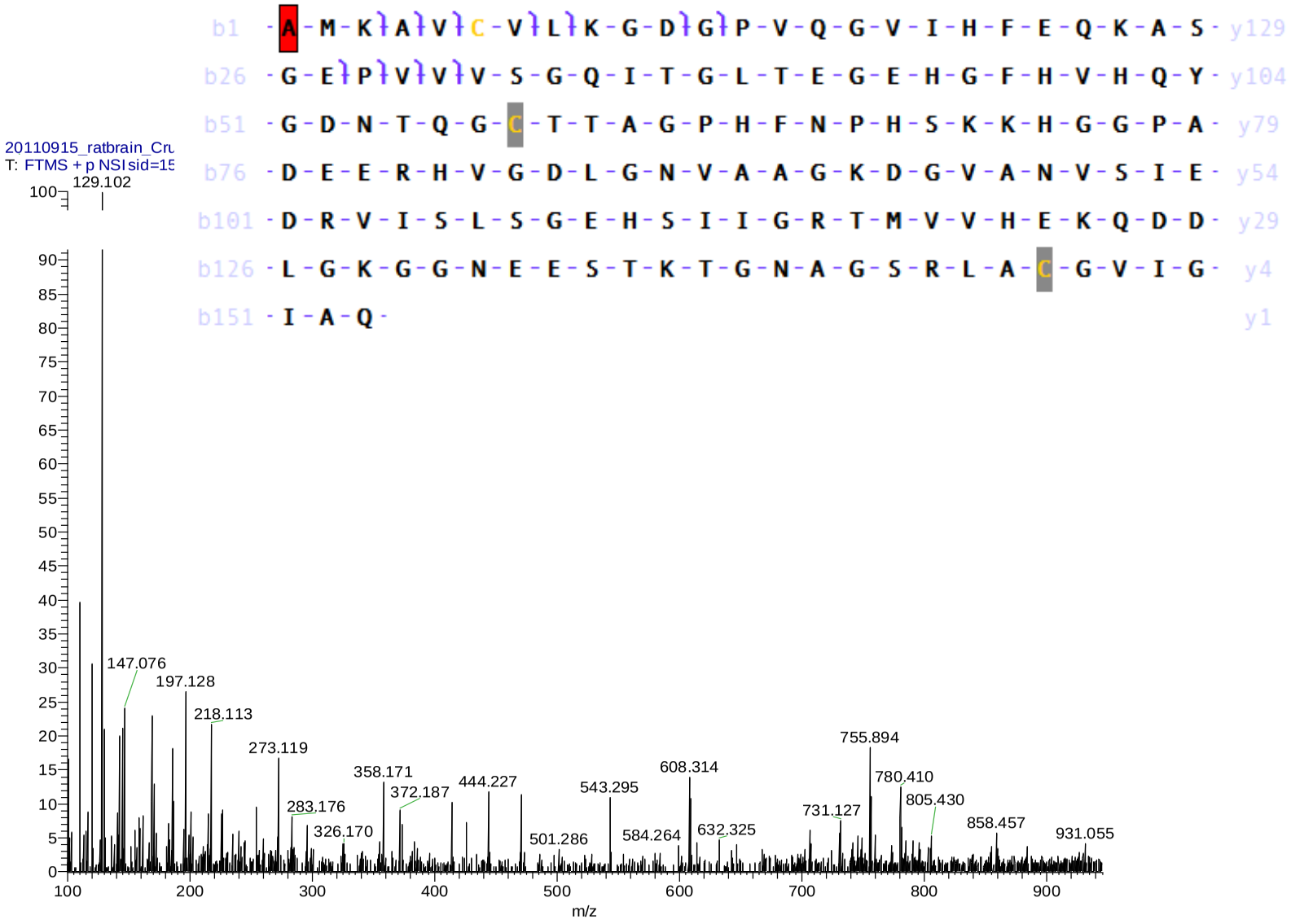
4854	VGF_RAT	P20156	Neurosecretory protein VGF	4850.41	4850.41	0		1.4E-24

b1 - N - S - E - P - Q - D - Q - G - E - L - F - Q - G - V) D) P - R - A - L - A - A V - L - L - Q - y23

b26 - A - L - D) R) P - A) S) P - P - A) V) P - A) G - S - Q - Q) G - T) P - E) E - y1



15825	SODC_RAT	P07632	Superoxide dismutase [Cu-Zn]	15810.82	15810.75	4	N-term acetylation, disulfide bond	3.2E-21



3327	VIP_RAT	P01283	VIP peptides	3323.75	3323.76	-3	C-term amidation	8.3E-14

b1 - H - S - D - A - V - F - T - D - N - Y - T - R - L - R - K - Q - M - A - V - K - K - Y - L - N - S - y4

b26 - I - L - N - y1

20110915_ratbrain_Crude #3013 RT: 33.66 AV: 1 NL: 4.70E4
T: FTMS + p NSI sid=15.00 d Full ms2 665.96@hcd30.00 [100.00-

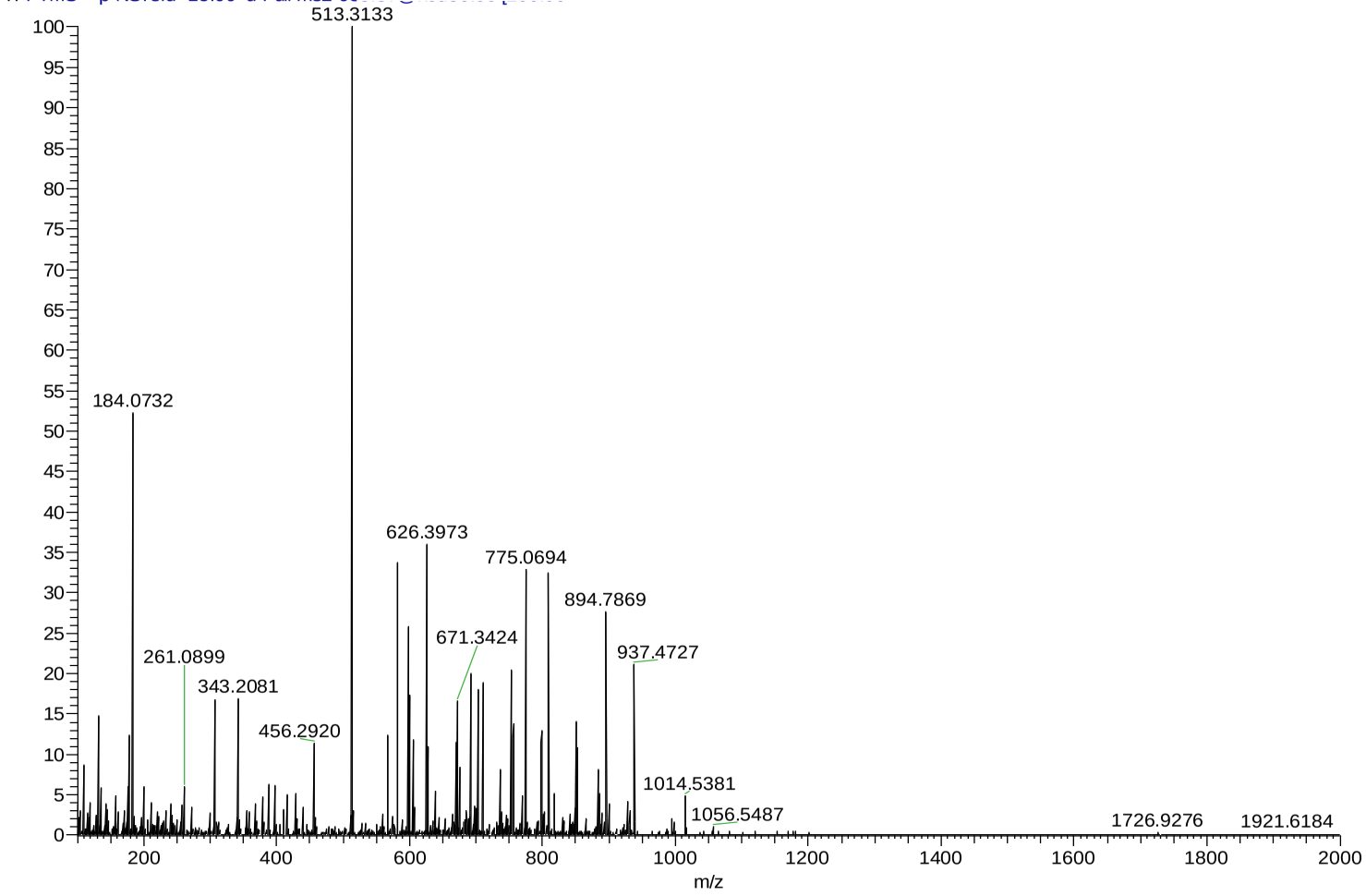
3437	7B2_RAT	P27682	Neuroendocrine protein 7B 2	3433.78	3433.79	-3		7.2E-39

b1 - Y - S - P - R - T - P - D - R - V - S - E - T - D - I - Q - R - L - L - H - G - V - M - E - Q - L - y6

b26 - G - I - A - R - P - y1

20110915_ratbrain_Crude #5006 RT: 5

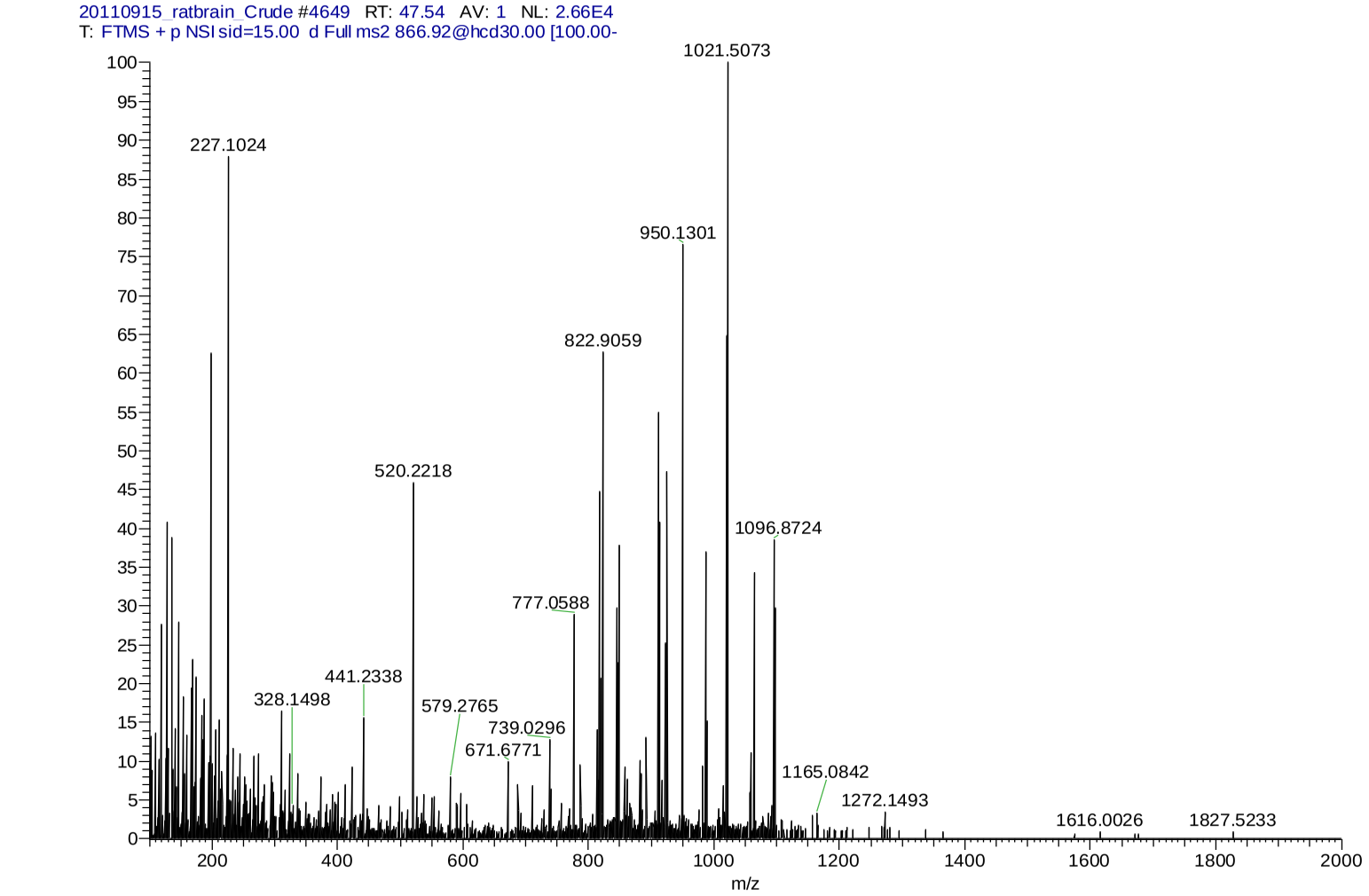
T: FTMS + p NSI sid=15.00 d Full ms2 6



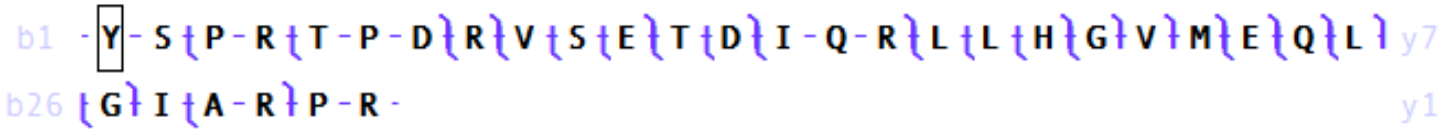
3464	NPY_RAT	P07808	Pro-Neuropeptide Y	3460.66	3460.66	0		1.6E-18

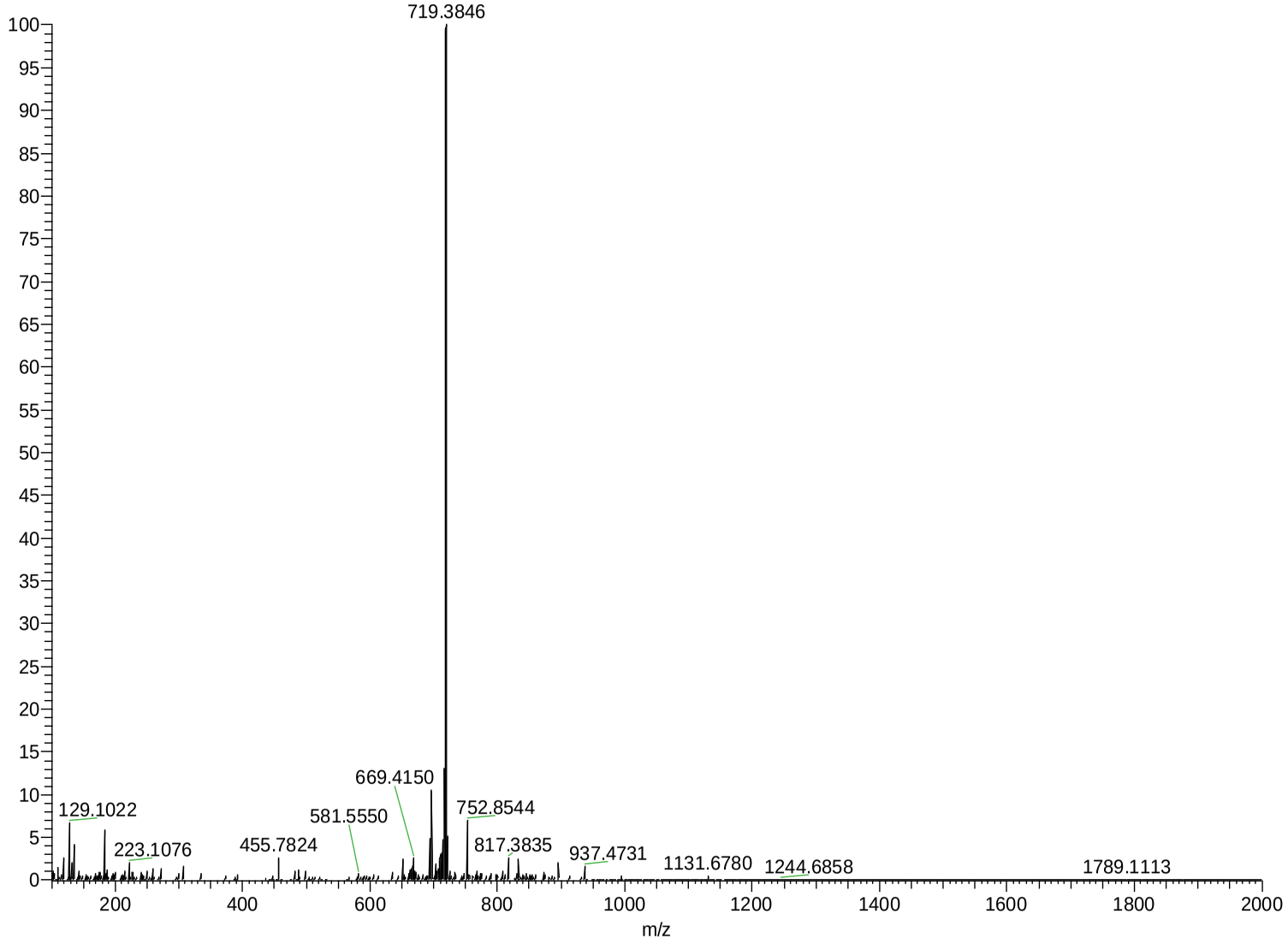
b1 - S - S - P - E - T - L - I - S - D - L - L - M - R - E - S - T - E - N - A - P - R - T - R - L - E - y6

b26 - D - P - S - M - W - y1



3593	7B2_RAT	P27682	Neuroendocrine protein 7B 2	3589.89	3589.89	0		2.5E-60



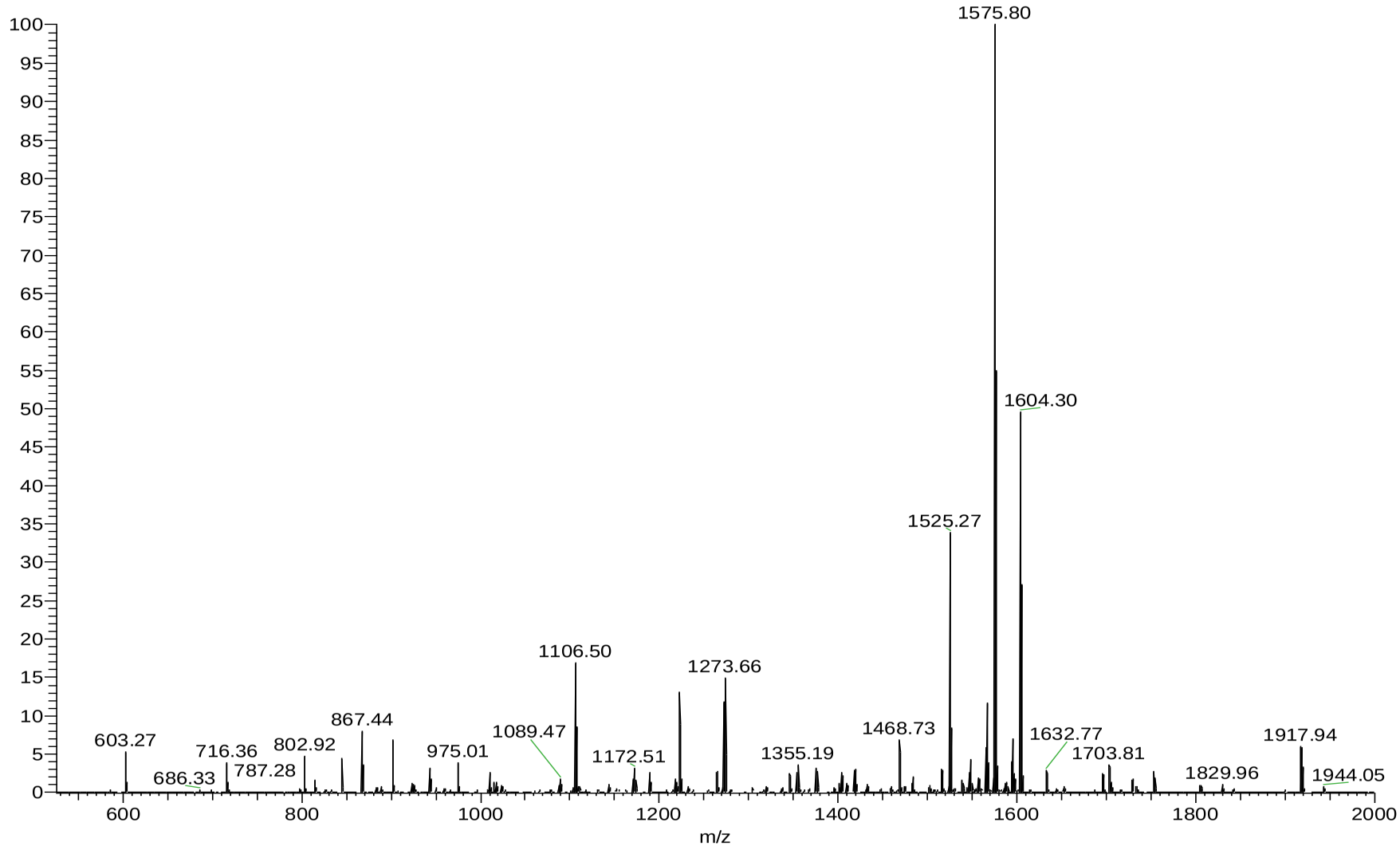


353

3653	SCG2_RAT	P10362	Secretogranin 2	3649.80	3649.80	0		3.8E-67

b1 -T-N-E-I-V-E-E-Q-Y-T-P-Q-S-L-A-T-L-E-S-V-F-Q-E-L-G- y9

b26 -K-L-T-G-P-S-N-Q- y1

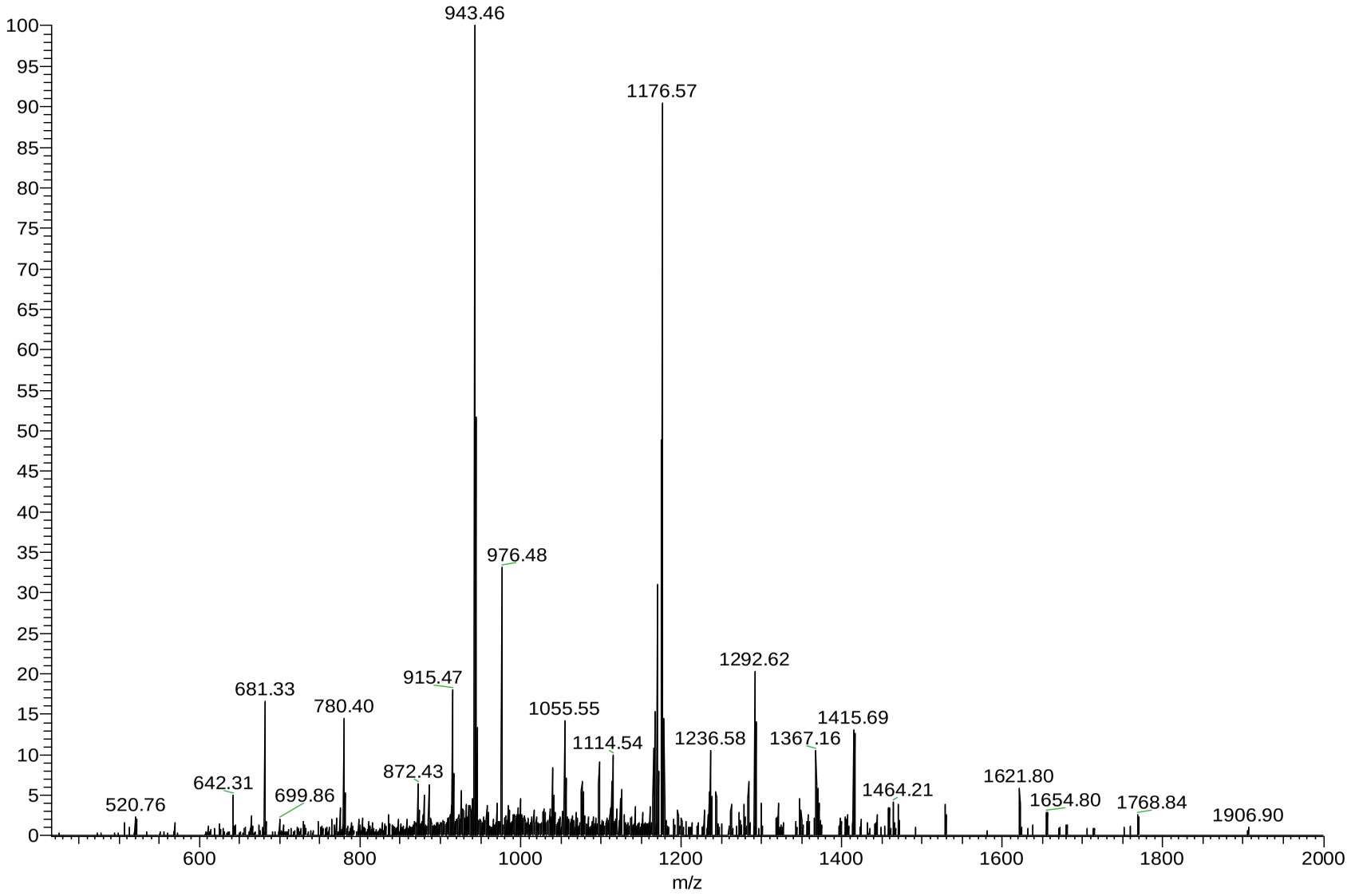


3676
VGF_RAT
P20156
Neurosecretory protein VGF
3672.77
3672.78
-3
8.0E-08
b1 -A-Q-E-E-A-D-A-E-E-R-R-L-Q-E-Q-E-E-L-E-N-Y-I-E-H-V- y6
b26 L-L-H-R-P- y1
129.10
187.14
260.11
300.12
401.17
522.31
612.33
635.40
661.41
725.93
789.98
869.68
913.90
1000.57
1181.53
1303.09
1850.82
m/z

3871	VGF_RAT	P20156	Neurosecretory protein VGF	3867.86	3867.87	-3		2.9E-37

b1 -A-P-P-G-R-S-D-V-Y-P-P-P-L-G-S-E-H-N-G-Q-V-A-E-D-A- y16

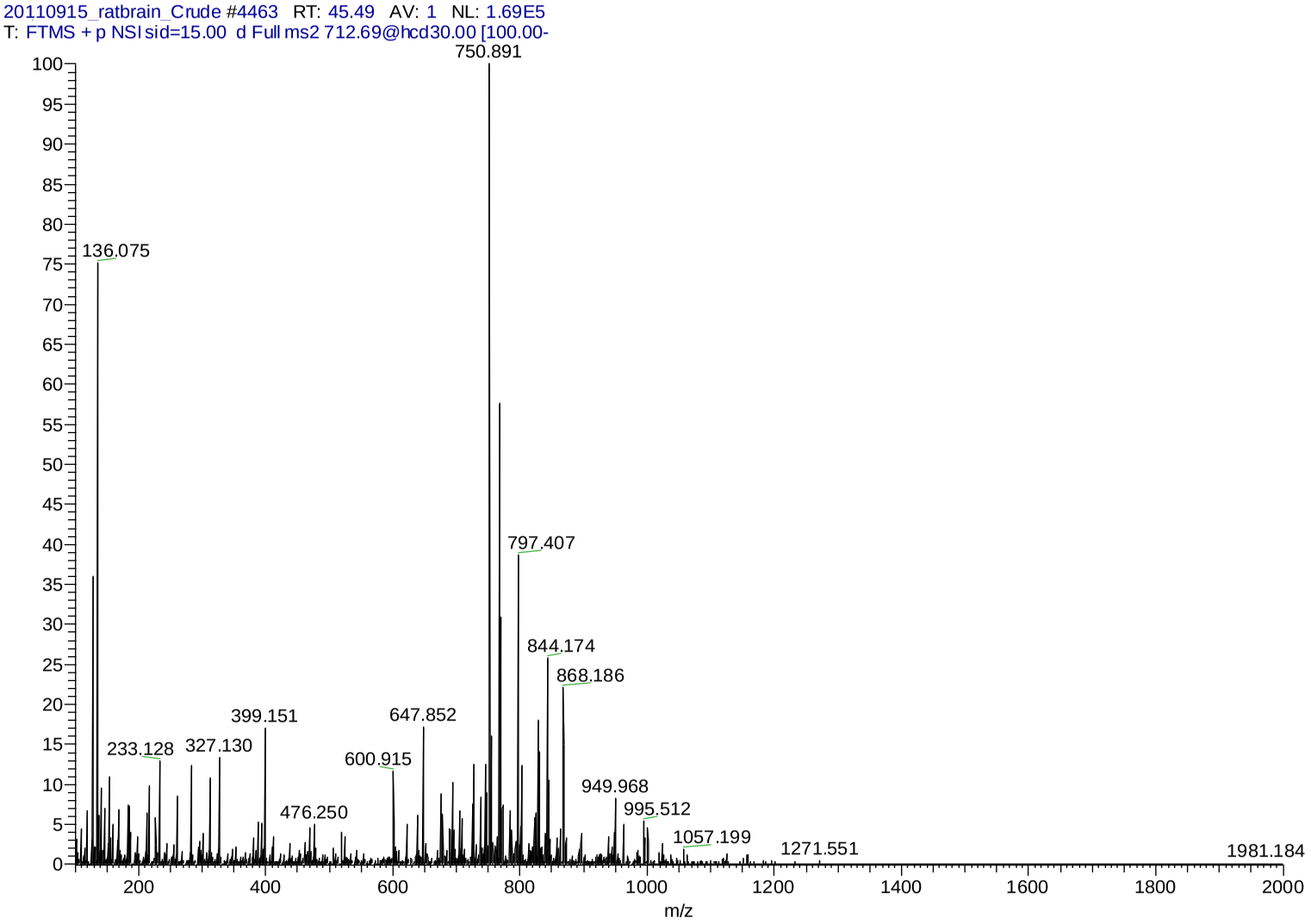
b26 -V-S-R-P-K-D-D-S-V-P-E-V y1



4273	NPY_RAT	P07808	Proneuropeptide Y	4269.08	4269.08	0	C-term amidation	3.6E-42

b1 -Y-P-S-K-P-D-N-P-G-E-D-A-P-A-E-D-M-A-R-Y-Y-S-A-L-R- y12

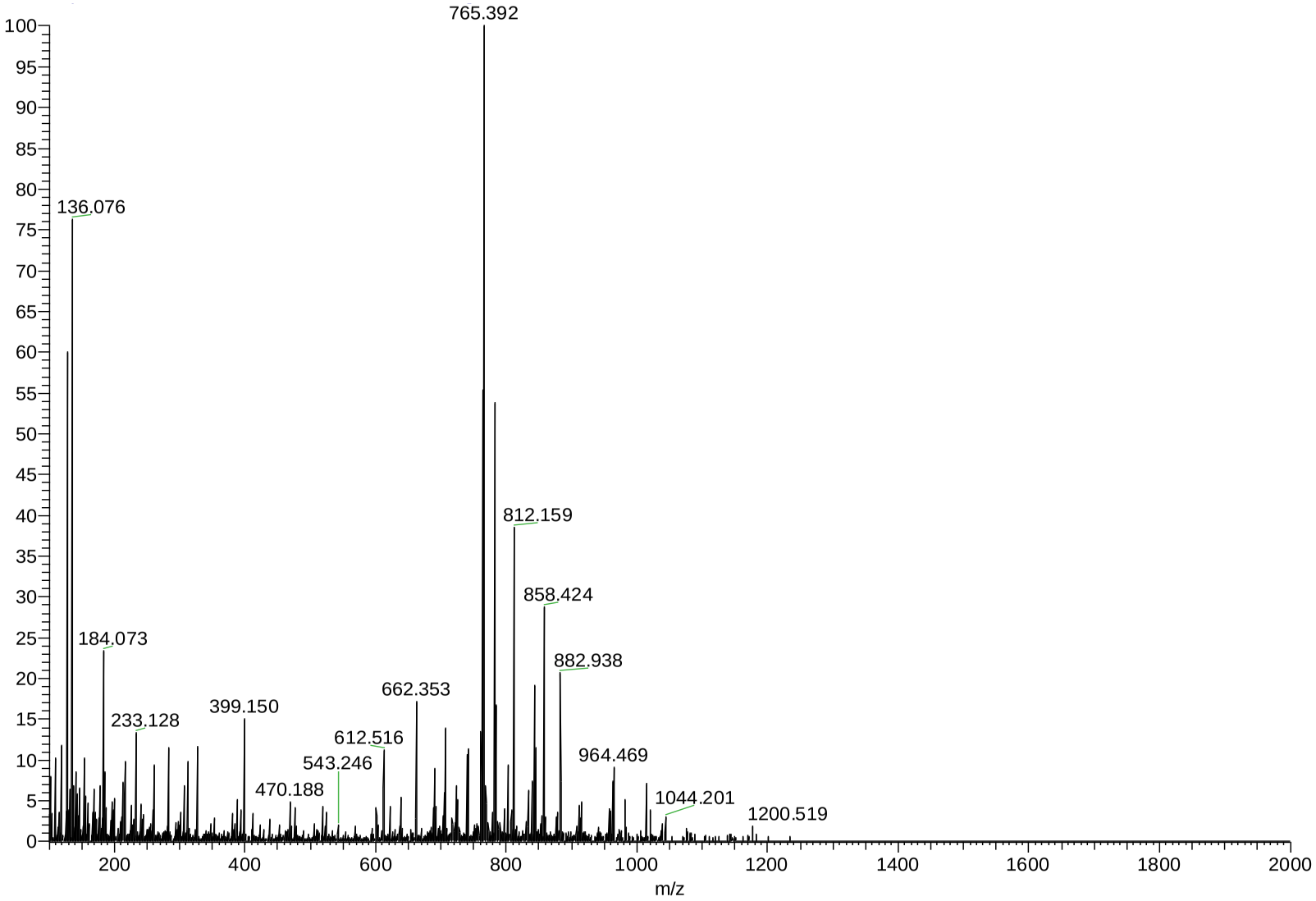
b26 -H-Y-I-N-L-I-T-R-Q-R-Y- y1



4331	NPY_RAT	P07808	Proneuropeptide Y	4327.08	4327.09	-2		1.6E-42

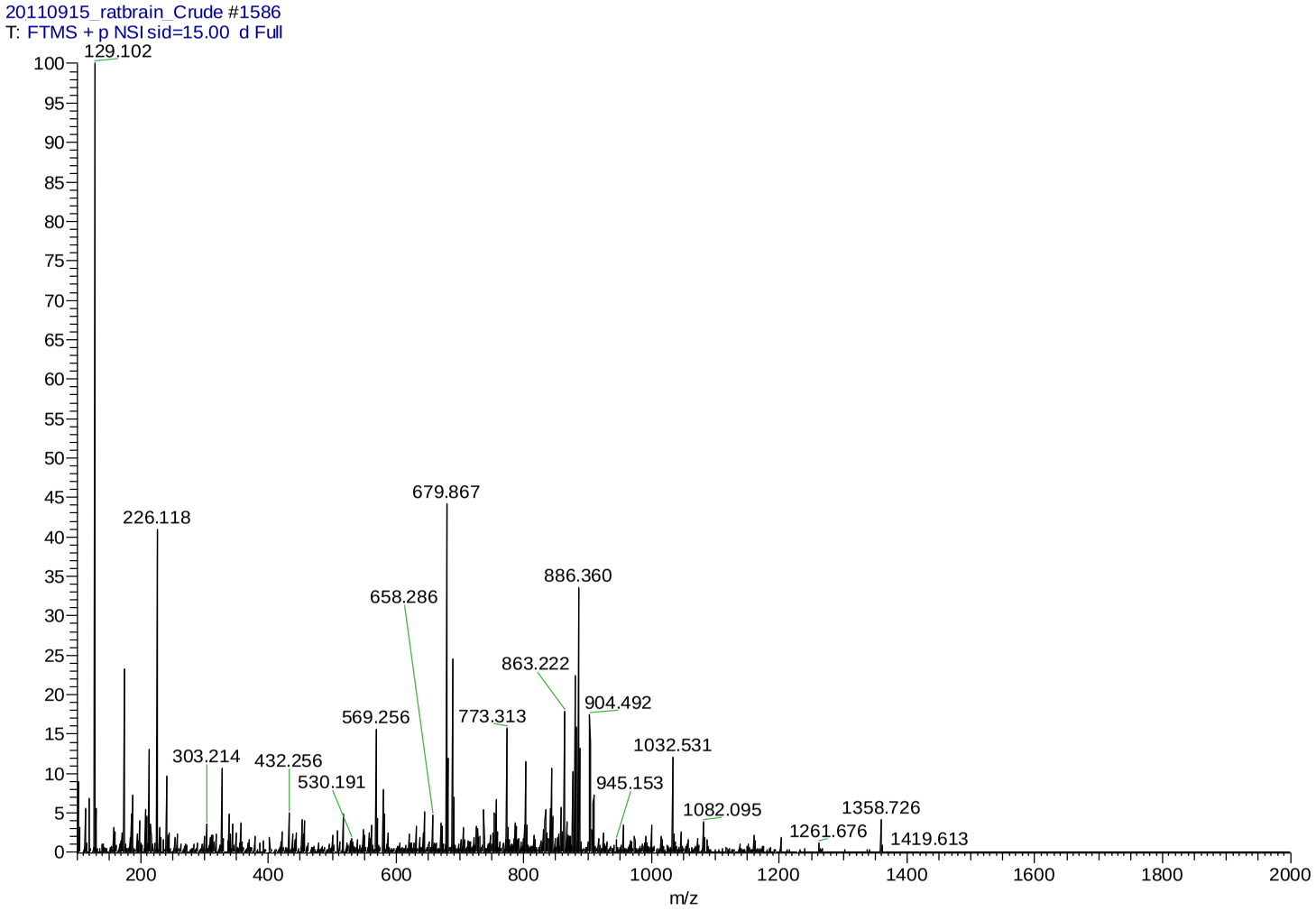
b1 -Y-P-S-K-P-D-N-P-G-E-D-A-P-A-E-D-M-A-R-Y-Y-S-A-L-R- y12

y1



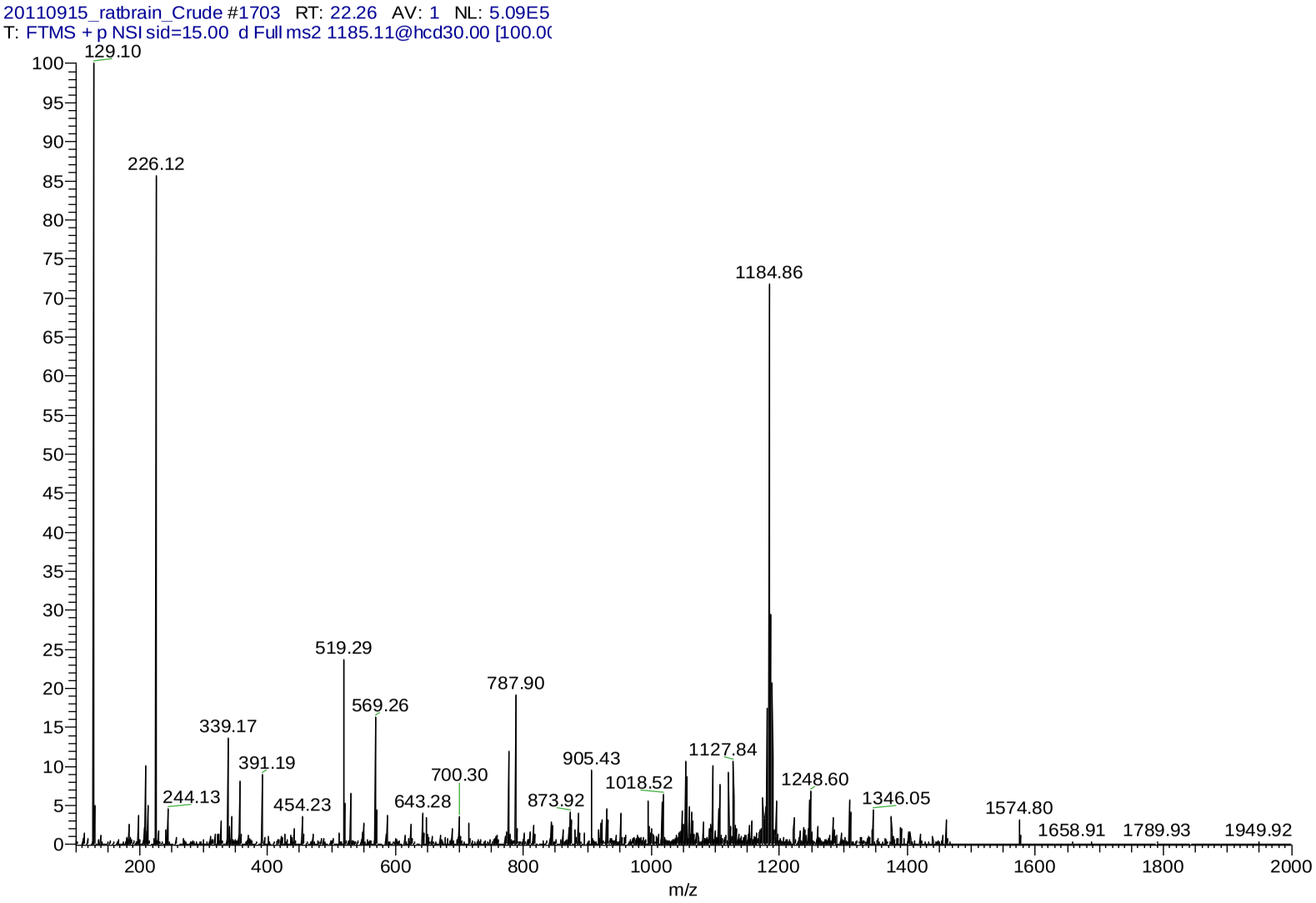
4522	TYB10_RAT	P63312	Thymosin beta-10	4517.33	4517.33	0	N-terminal acetylation	2.2E-70

-A-DKPDMGEIASFDKAKLKKTETQEK-
NTLPTKETIEQE-K-R



4738	TYB10_RAT	P63312	Thymosin beta-10	4733.40	4733.41	-2	N-term acetylation	1.7E-89

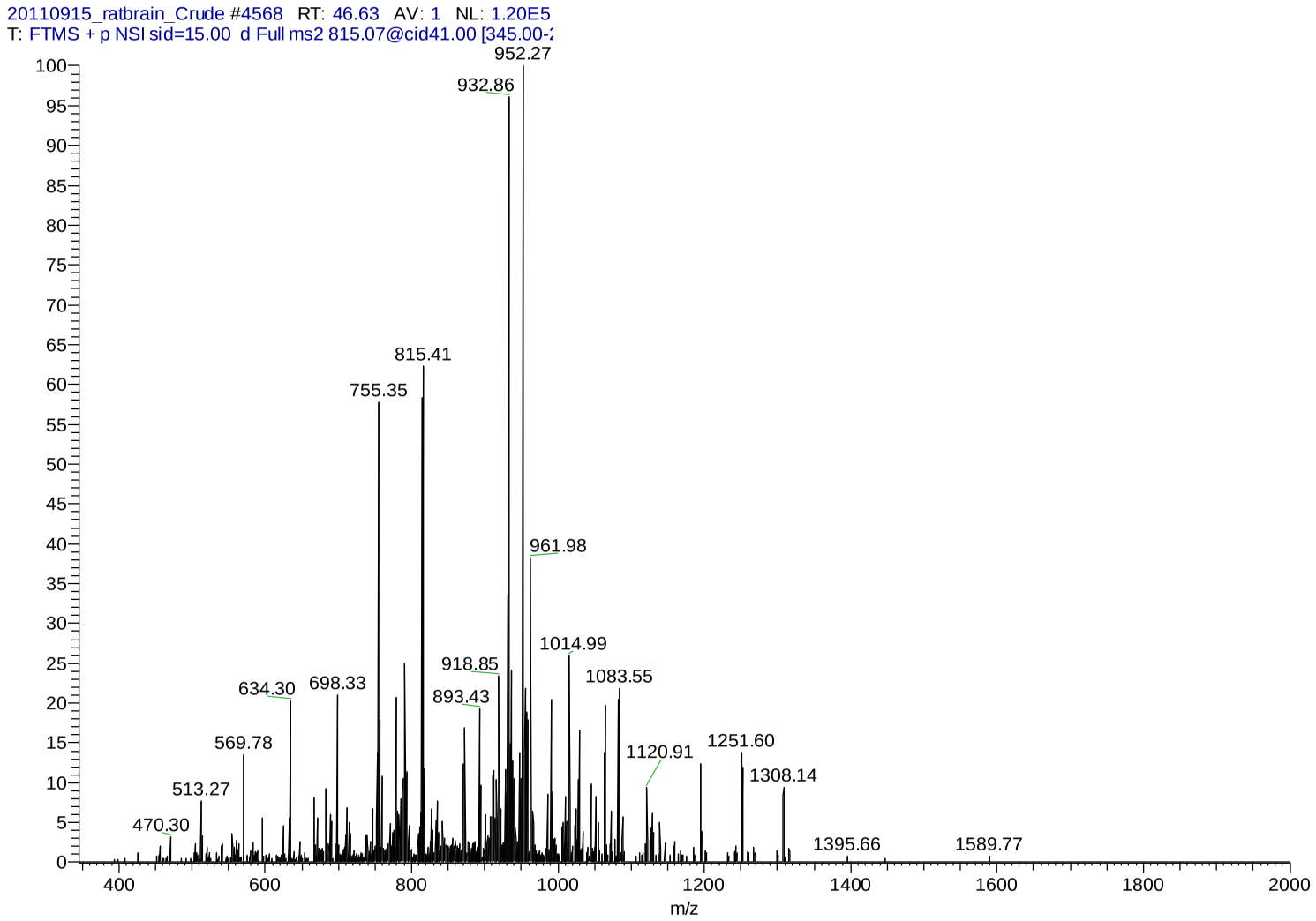
-A-D-K-P-D-M-G-E-I-A-S-F-D-K-A-K-L-K-K-T-E-T-Q-E-K-N-T-L-P-T-K-E-T-I-E-Q-E-K-R-S-E



4871	SCG2_RAT	P10362	Secretogranin-2	4867.40	4867.41	-2		1.2E-77

b1 -I}P}A}G-S}L-K}N}E}D}T}P-N-R-Q-Y}L}D}E}D}M}L}L}K}V} y18

}H}L-A- y1



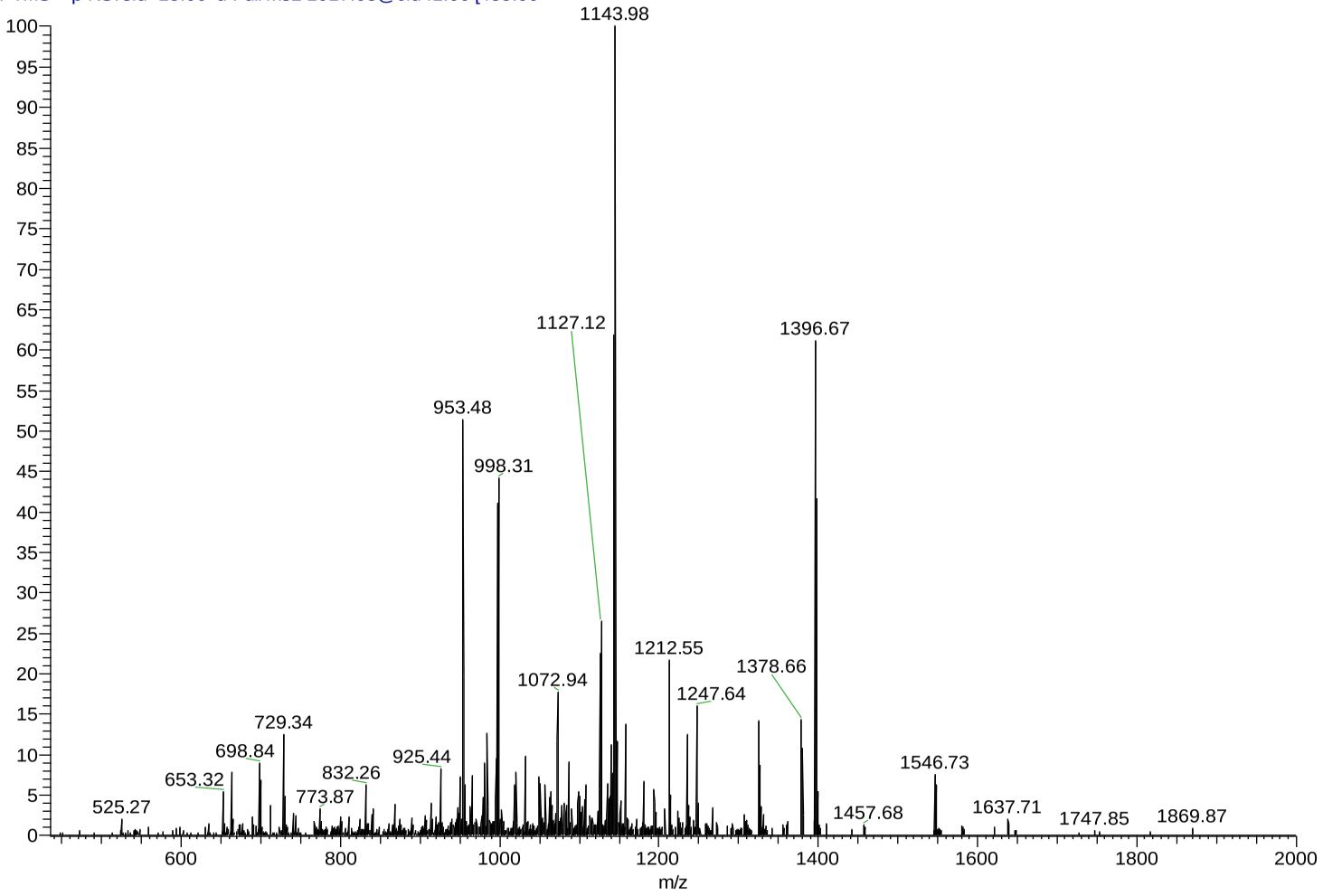
7114	FILUV9_RAT	FILUV9	Uncharacterized	7108.52	7108.51	1		4.9E-42

b1 - S - P - A - P - A - T - G - A - S - G - Q - A - S - E - L - A - P - S - T - A - D - S - A - V - P - y47

b26 - P - A - P - A - K - T - E - K - G - P - V - E - T - K - S - E - P - Q - E - S - E - A - K - P - A - y22

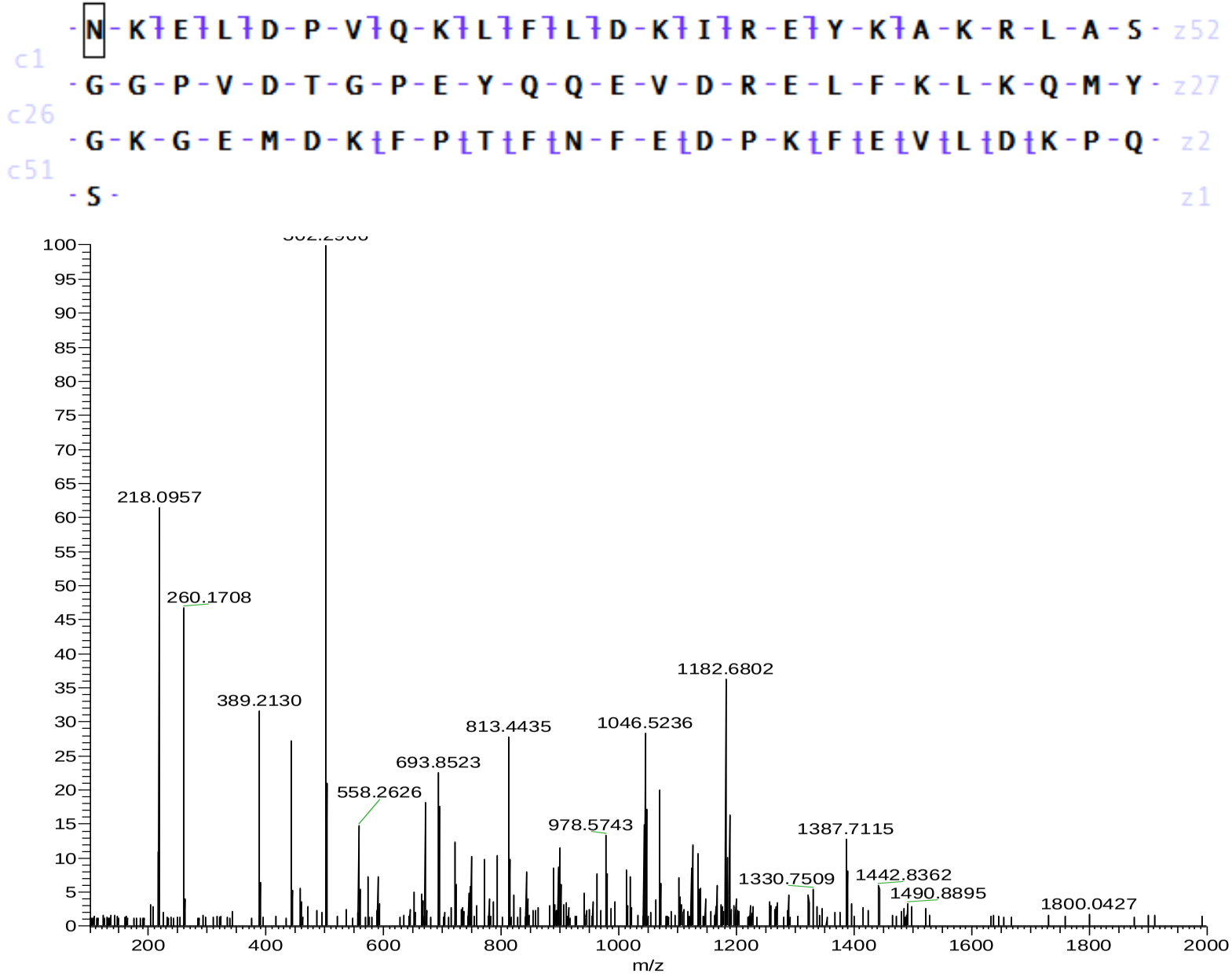
b51 - P - T - E - V - K - T - V - P - N - E - A - T - Q - T - K - E - N - E - S - K - A - y1

20110915_ratbrain_Crude #1015 RT: 17.87 AV: 1 NL: 9.67E4
T: FTMS + p NSI sid=15.00 d Full ms2 1017.08@cid41.00 [435.00



8468
UBB_RAT
P0CG51
Polyubiquitin
8461.56
8461.57
-1
Oxidation
2.9E-35
b1 - M - Q - I - F - V - K - T - L - T - G - K - T - I - T - L - E - V - E - P - S - D - T - I - E - N - y50
b26 - V - K - A - K - I - Q - D - K - E - G - I - P - D - Q - Q - R - L - I - F - A - G - K - Q - L - y25
b51 - E - D - G - R - T - L - S - D - Y - N - I - Q - K - E - S - T - L - H - L - V - L - R - L - R - y1
20110915_ratbrain_Crude #3676 RT: 38.84 AV: 1 NL: 4.18E5
T: FTMS + p NSI sid=15.00 d Full ms2 770.79@hcd30.00 [100.00-
136.0757
233.1646
276.1010
436.1537
604.3842
732.9537
891.0907
982.4921
1113.3600
1253.6090
m/z

8928	ATP5J_RAT	P21571	ATP synthase coupling factor 6	8921.55	8921.55	0		1.7E-52



10918	DLRB1_RAT	P62628	Dynein light chain roadblock-type 1	10909.71	10909.72	-1	N-term acetylation, unfixed oxidation	6.2E-36

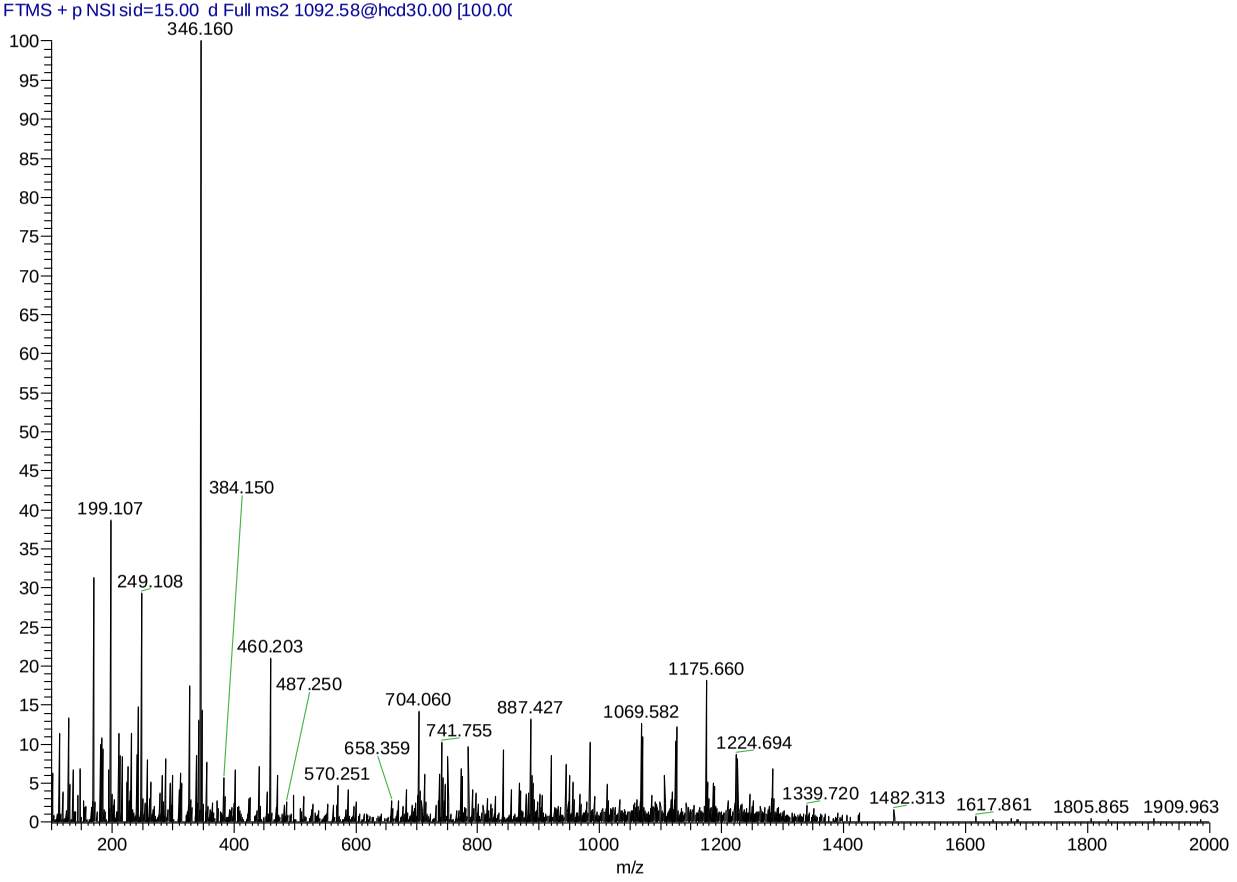
b1 - A - E - V - E - E - T - L - K - R - L - Q - S - Q - K - G - V - Q - G - I - I - V - V - N - T - E - y71

b26 - G - I - P - I - K - S - T - M - D - N - P - T - T - T - Q - Y - A - N - L - M - H - N - F - I - L - y46

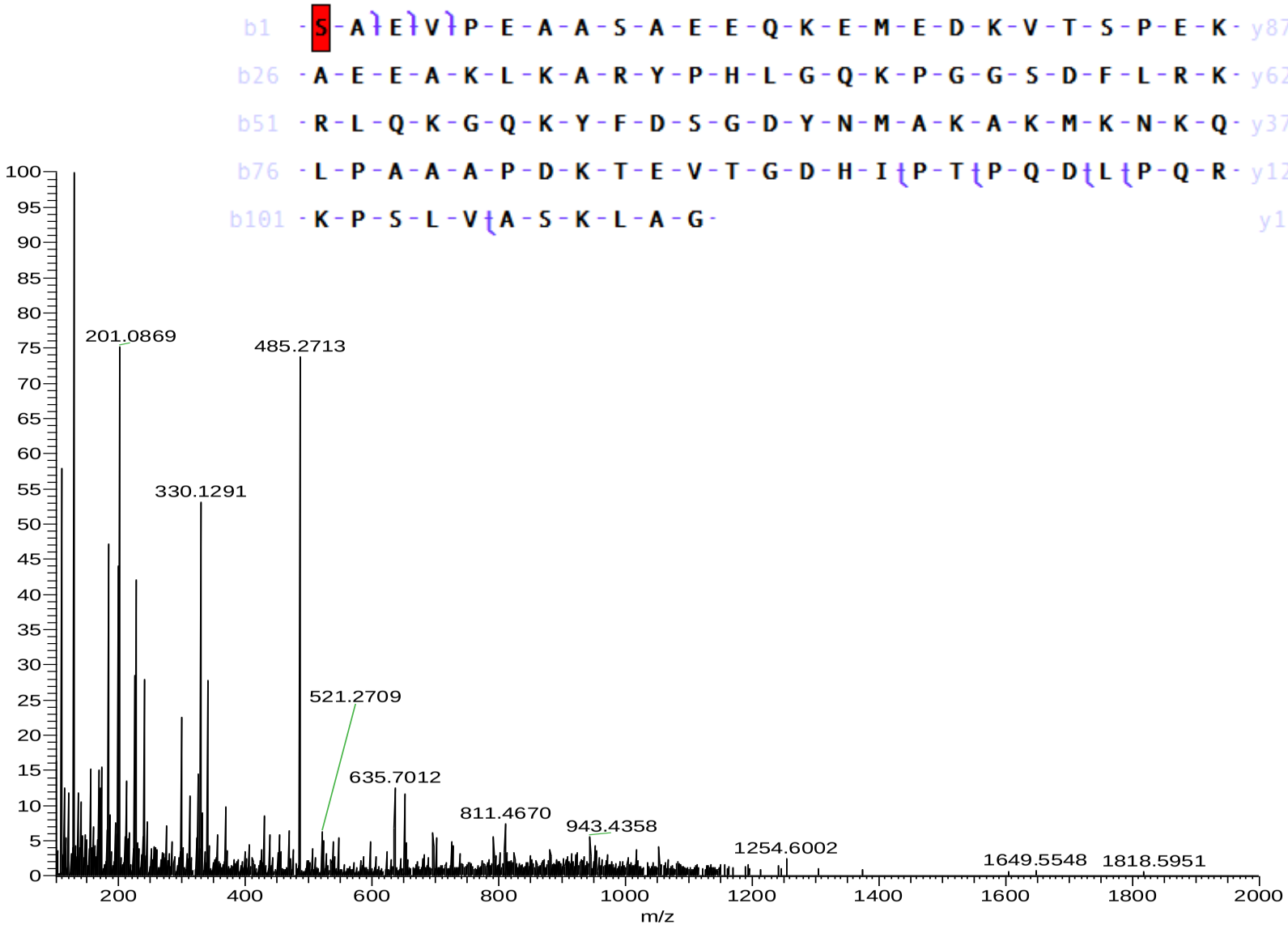
b51 - K - A - R - S - T - V - R - E - I - D - P - Q - N - D - L - T - F - L - R - I - R - S - K - K - N - y21

b76 - I - M - V - A - P - D - K - D - Y - F - L - I - V - I - Q - N - P - T - E - y1

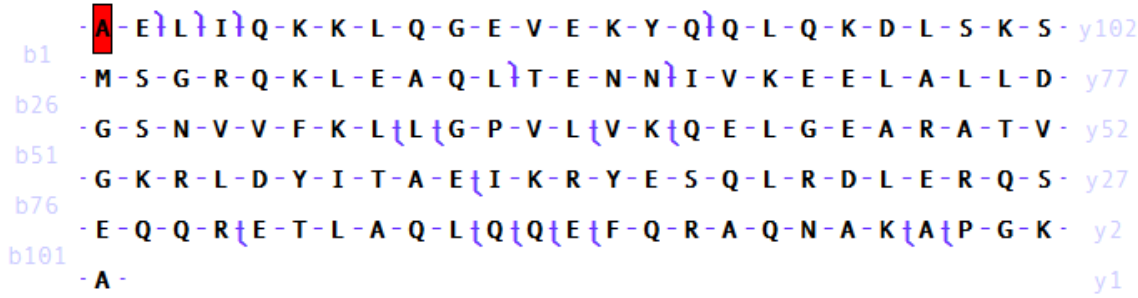
20110915_ratbrain_Crude #4946 RT: 51.35 AV: 1 NL: 7.55E4
T: FTMS + p NSI sid=15.00 d Full ms2 1092.58@hcd30.00 [100.0(

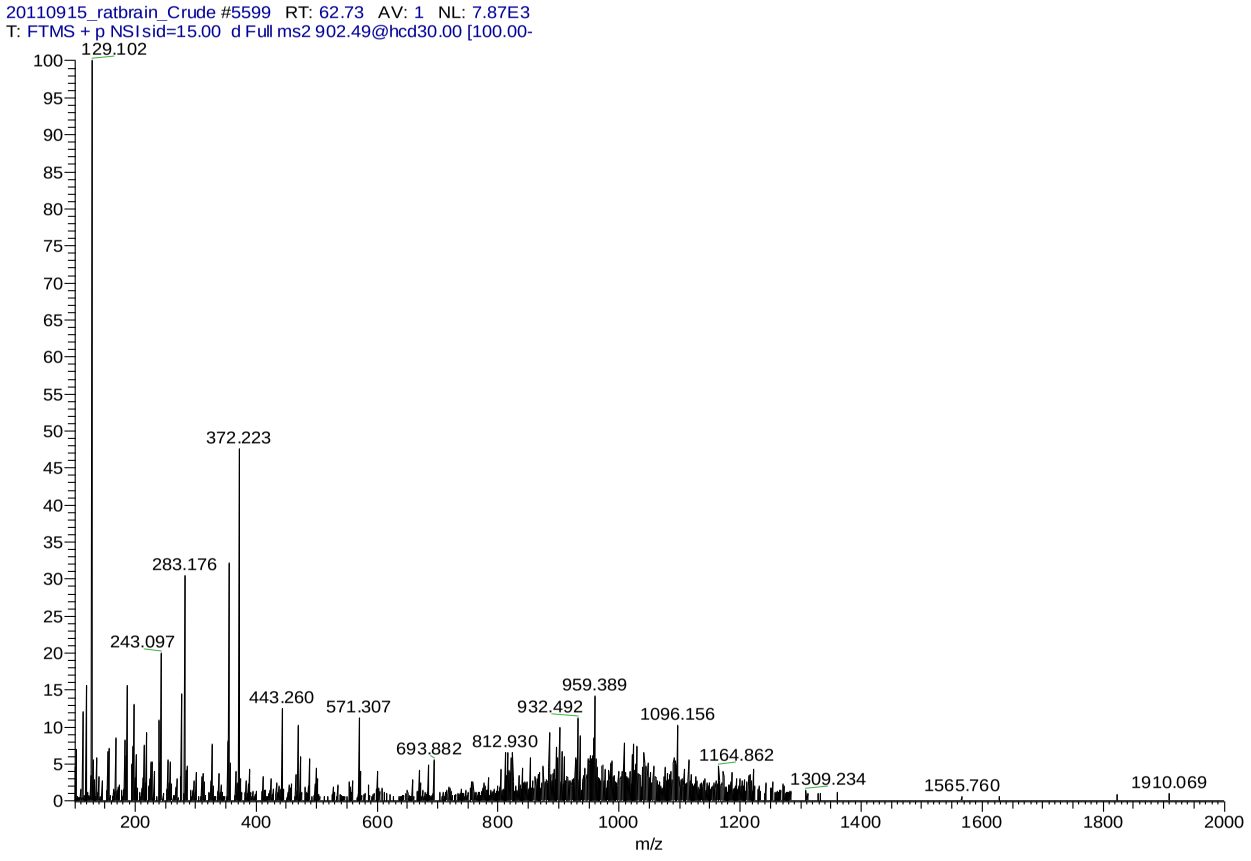


12204	ARP19_RAT	Q712U5	cAMP-regulated phosphorylation 19	12196.24	12196.26	-2	N-term acetylation	2.3E-13



14428	Q6MGC4_RAT	Q6MGC4	H2-K region expressed gene 2, rat orthologue	14417.76	14417.73	2	N-term acetylation, H_2O adduct	3.1E-24

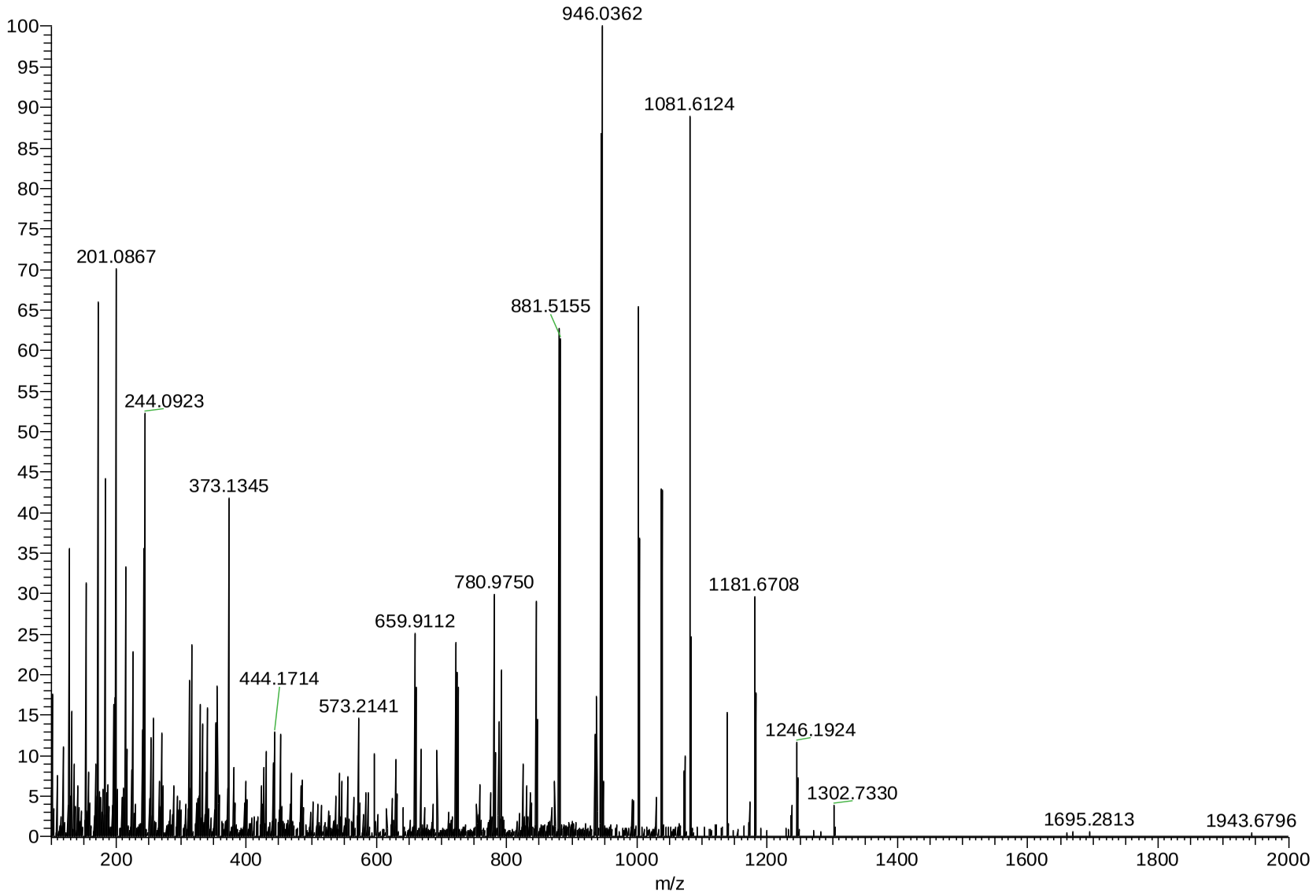




CMGA_RAT	P10354	Chromagranin-A	3174.66	3174.66	0	C-term amidation	6.8E-34

b1 - A - E - D - Q - E - L - E - S - L - S - A - I - E - A - E - L - E - K - V - A - H - Q - L - Q - A - y4

b26 - L - R - R - y1

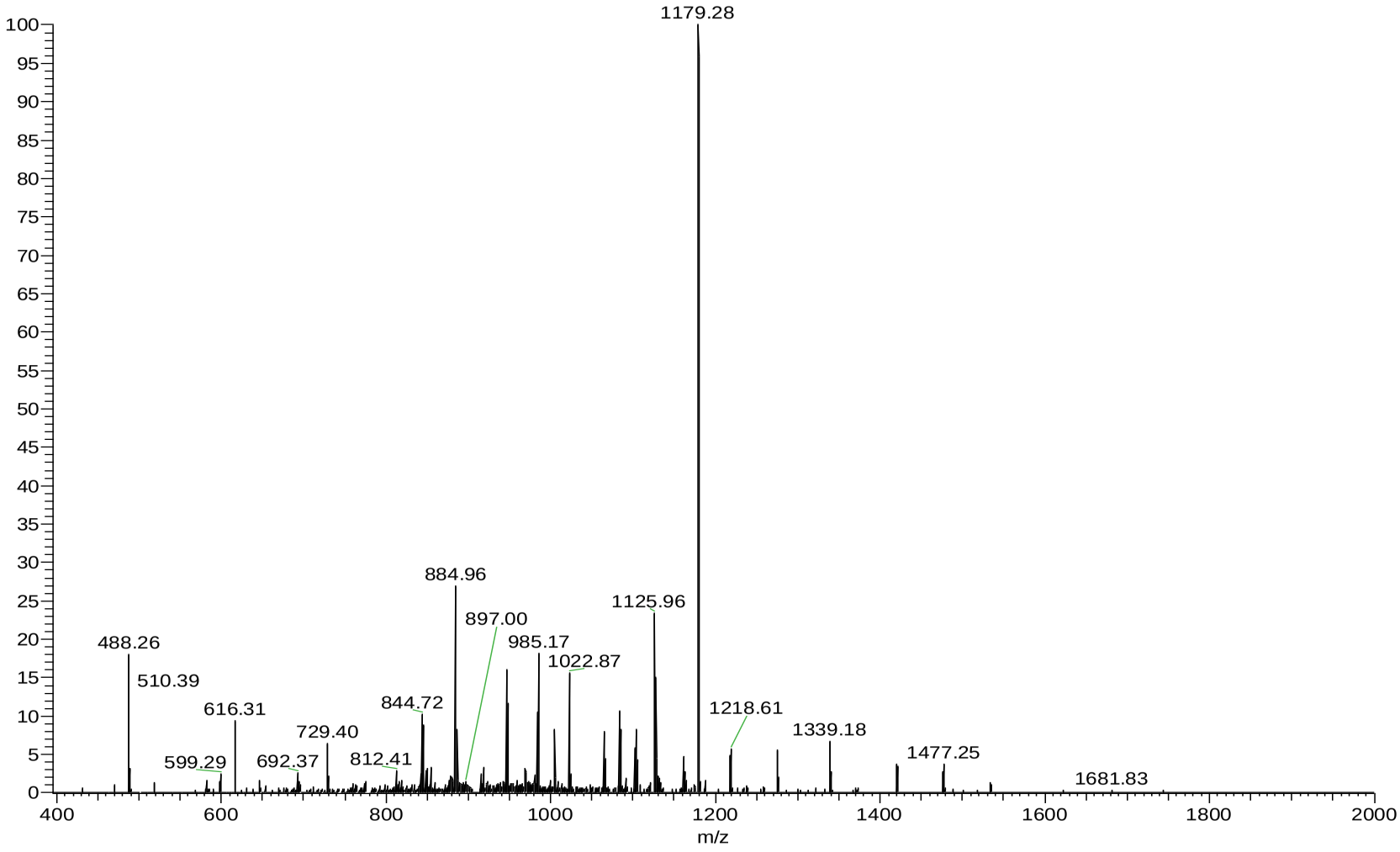


CCKN_RAT
P01355
Cholecystokinin
3392.80
3392.81
-3
2.5E-13
b1 -Y-I-Q-Q-V-R-K-A-P-S-G-R-M-S-V-L-K-N-L-Q-G-L-D-P-S- y15
b26 H-R-I-S-D-R-D-Y-M-G-W-M-D-F- y1
20110915_ratbrain_Crude #2288 RT: 26.84 AV: 1 NL: 9.84E4
T: FTMS + p NSI sid=15.00 d Full ms2 682.76@cid41.00 [290.00-
683.1688
780.4224
811.4047
862.1444
926.4307
1096.5354
1042.5519
1147.5496
1269.3987
1780.2806
635.3842
406.2059
494.2900
584.3395
m/z

F1LP33_RAT	F1LP33	Oxysterol-binding protein	3679.89	3679.87	5		1.2E-06

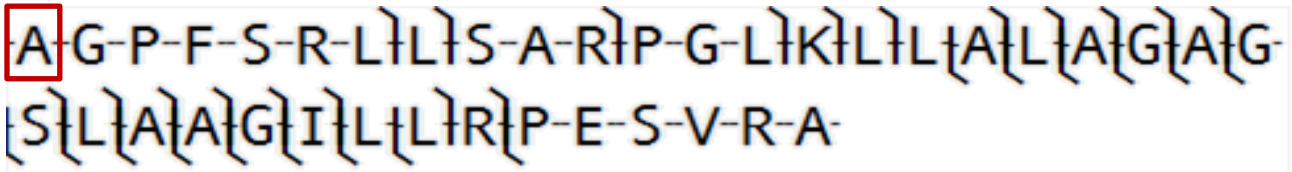
b1 - K - L - T - E - A - D - A - Y - L - Q - I - L - I - E - Q - L - K - L - F - D - D - K - L - Q - N - y7

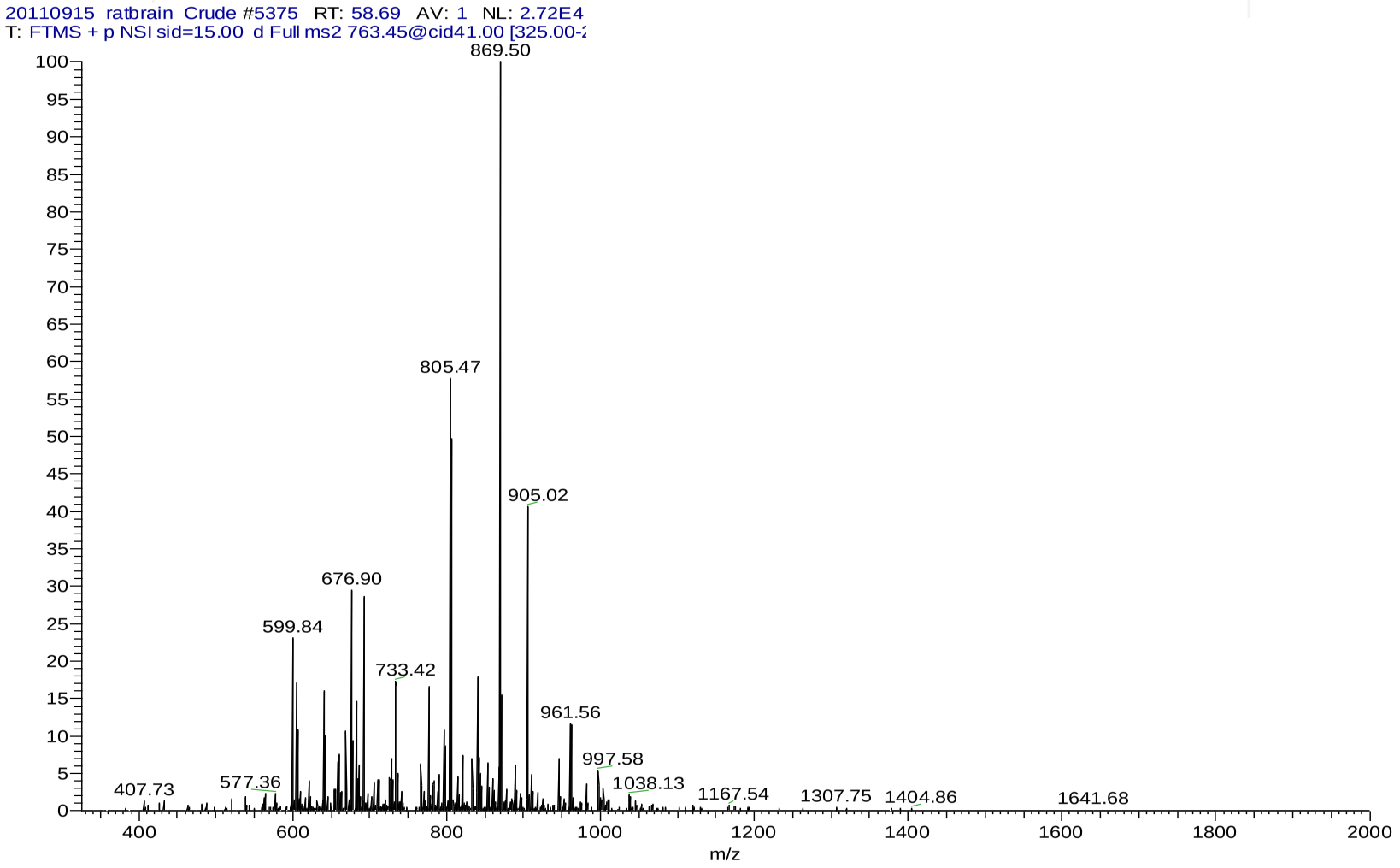
b26 - C - K - D - D - E - Q - y1



370

KCRU_RAT	P25809	Creatine kinase U-type	3811.23	3811.23	0	N-term acetylation	4.2E-57

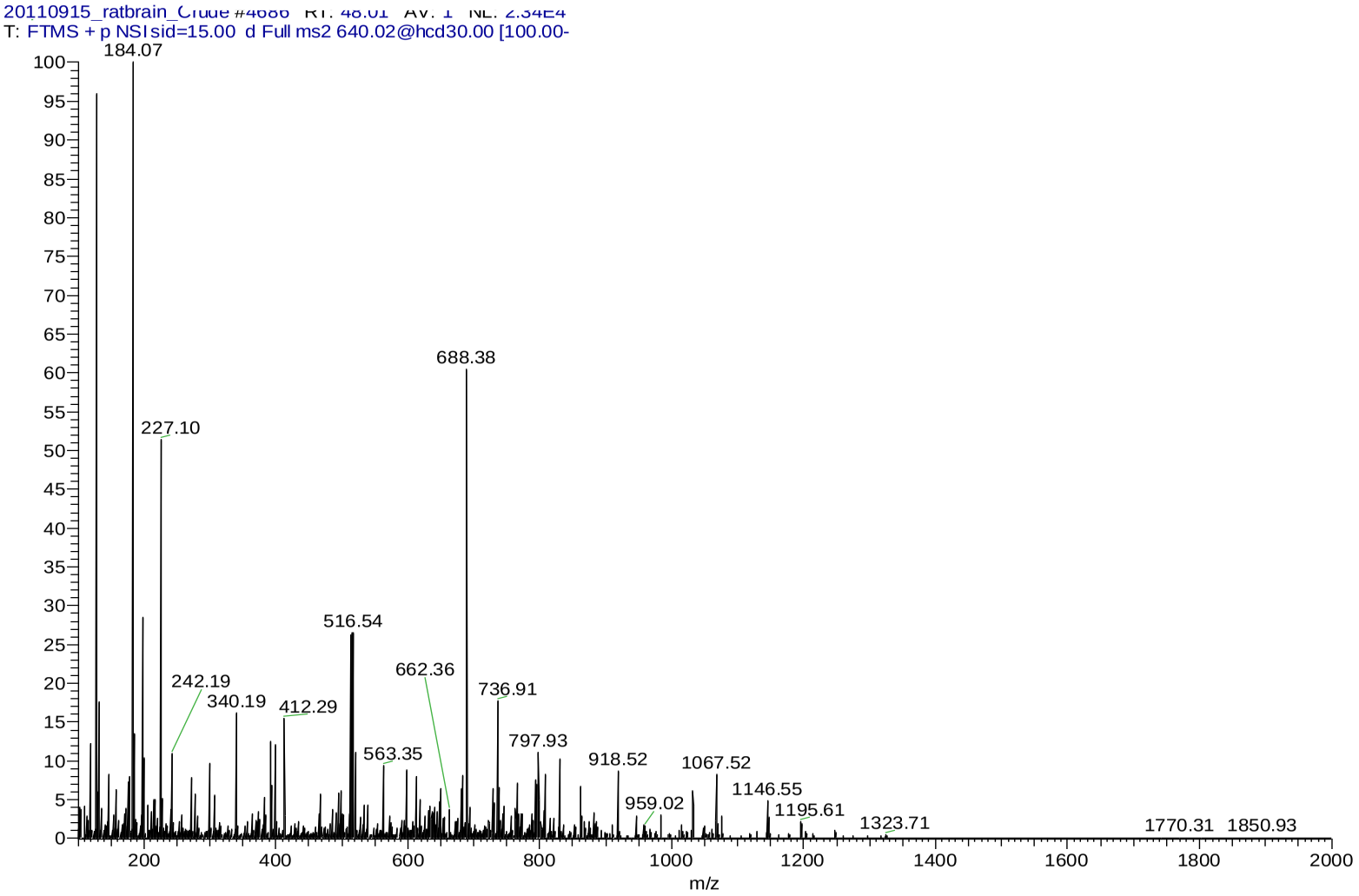




MDHM_RAT	P04636	Malate dehydrogenase	3832.09	3832.09	0		7.0E-50

b1 -I-G-K-I-T-P-F-E-E-K-M-I-A-E-A-I-P-E-L-K-A-S-I-K-K- y10

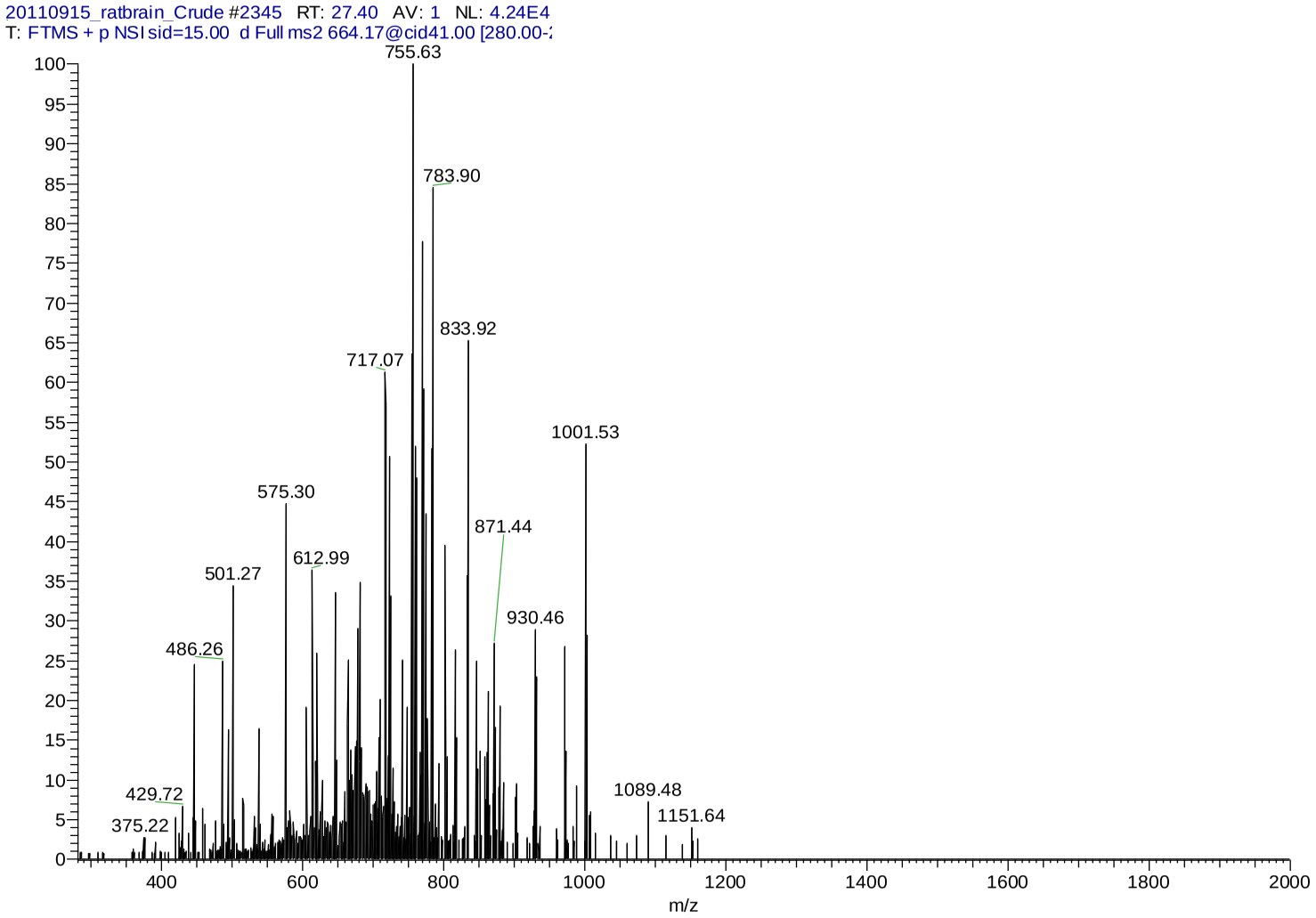
b26 -G-E-D-F-V-K-N-M-K- y1



D4AAI2_RAT	D4AAI2	Histocompatibility (minor) HA-1, isoform CRA_b	3975.99	3975.99	0		8.2E-05

b1 - F - Y - H - E - L - V - G - L - A - K - D - S - L - K - A - E - A - E - A - K - A - A - S - R - G - y14

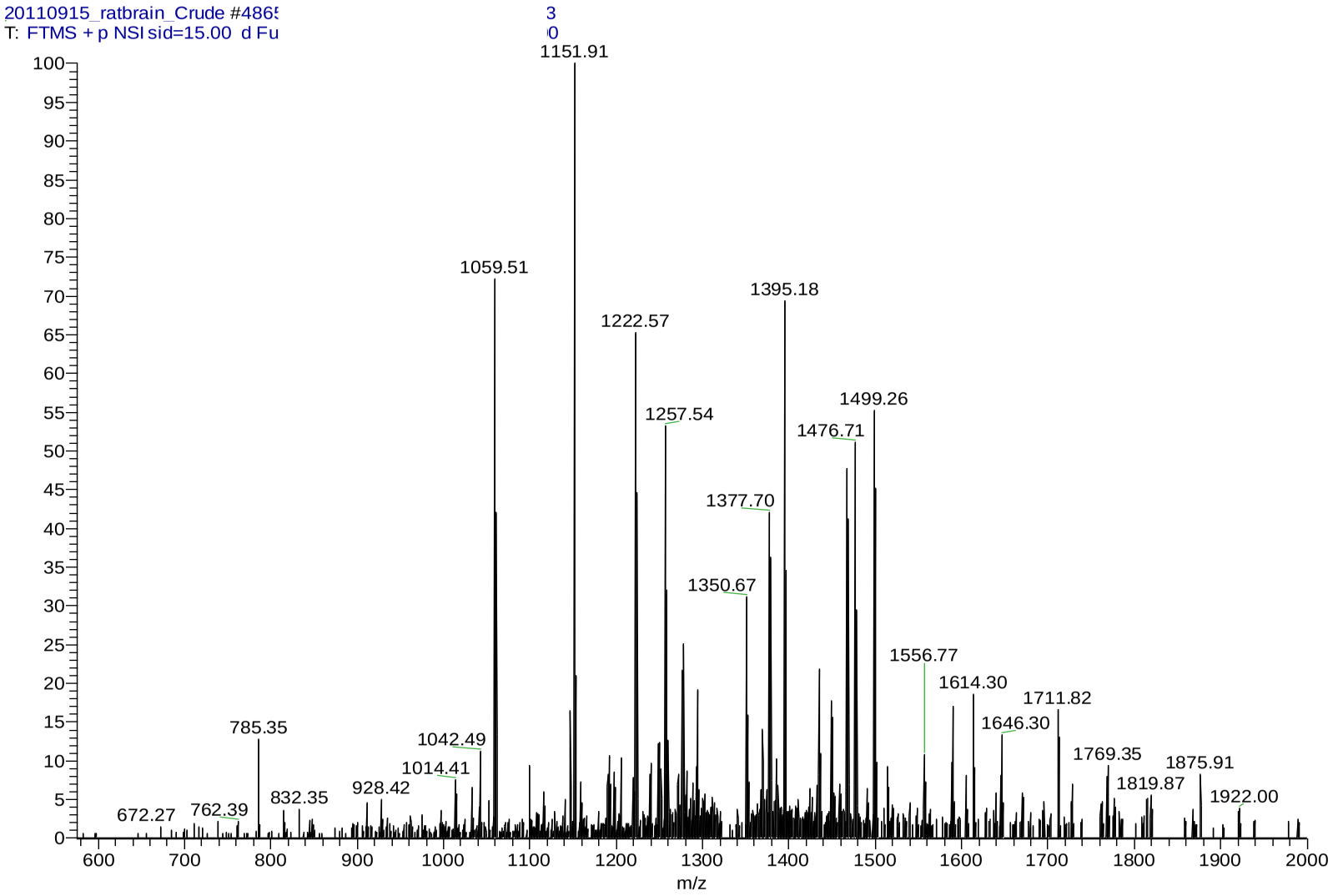
b26 - R - Q - D - G - S - E - S - E - A - A - T - L - A - y1



D3ZMX3_RAT	D3ZMX3	Neutralized-like (drosophila), isoform CRA_b	4008.92	4008.91	2		3.3E-07

b1 - S - L - C - D - L - N - V - P - G - A - D - G - E - D - G - A - P - P - A - G - C - P - I - P - Q - y16

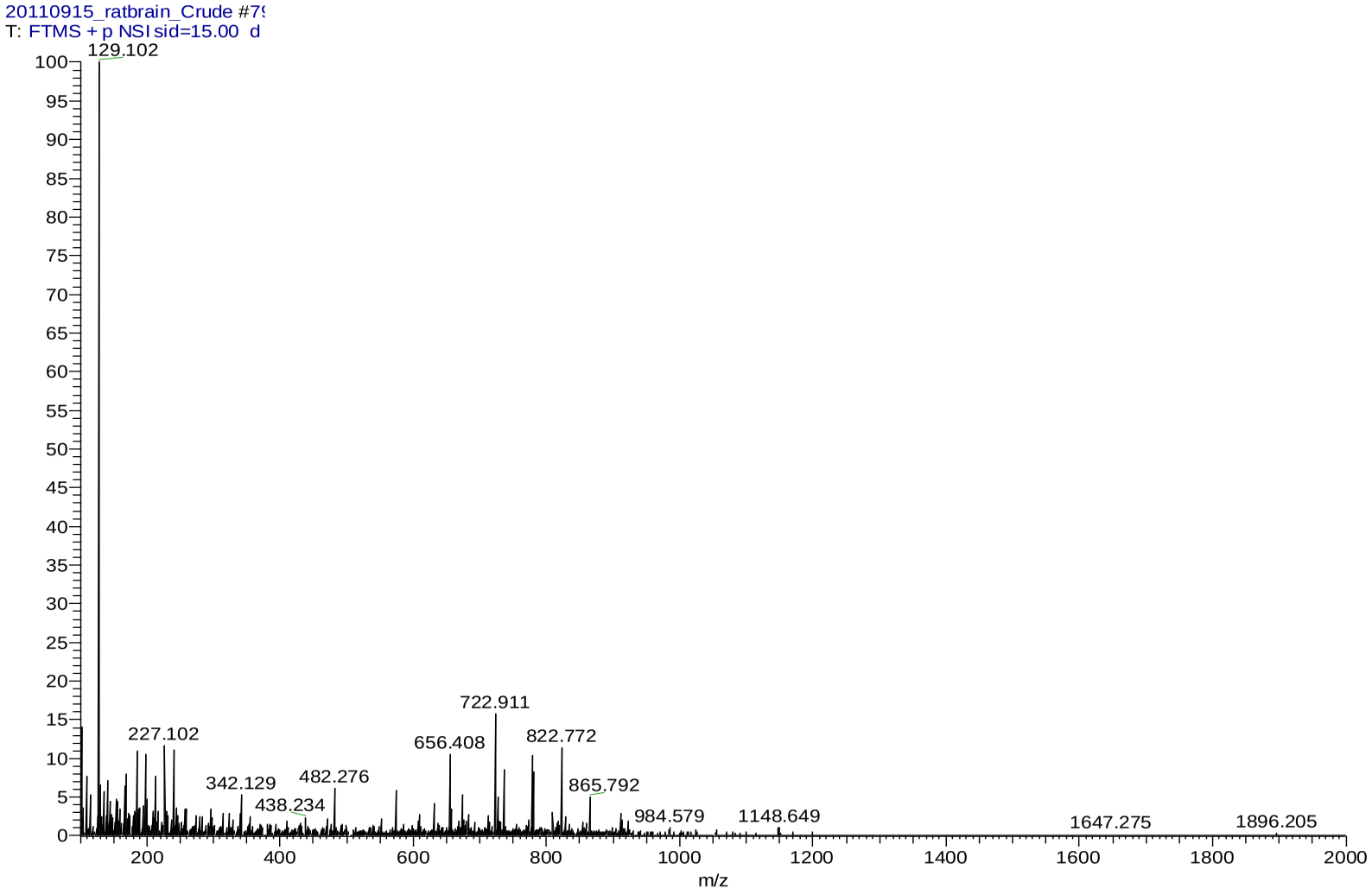
b26 - N - S - I - N - S - O - H - S - R - A - L - P - A - Q - L - y1



TMM35_RAT	Q6JAM9	Transmembrane protein 35	4036.15	4036.16	-2		1.2E-36

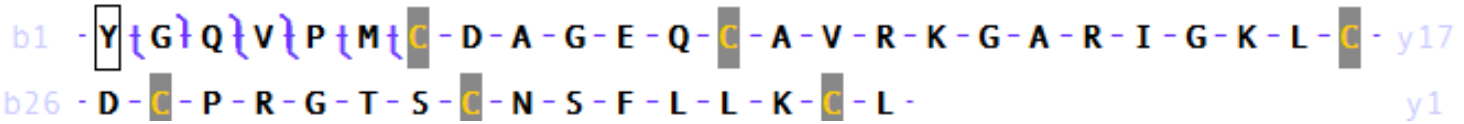
b1 -L-I-A-R-K-P-E-D-R-S-S-E-K-K-A-L-P-E-S-A-E-E-Q-P-S- y12

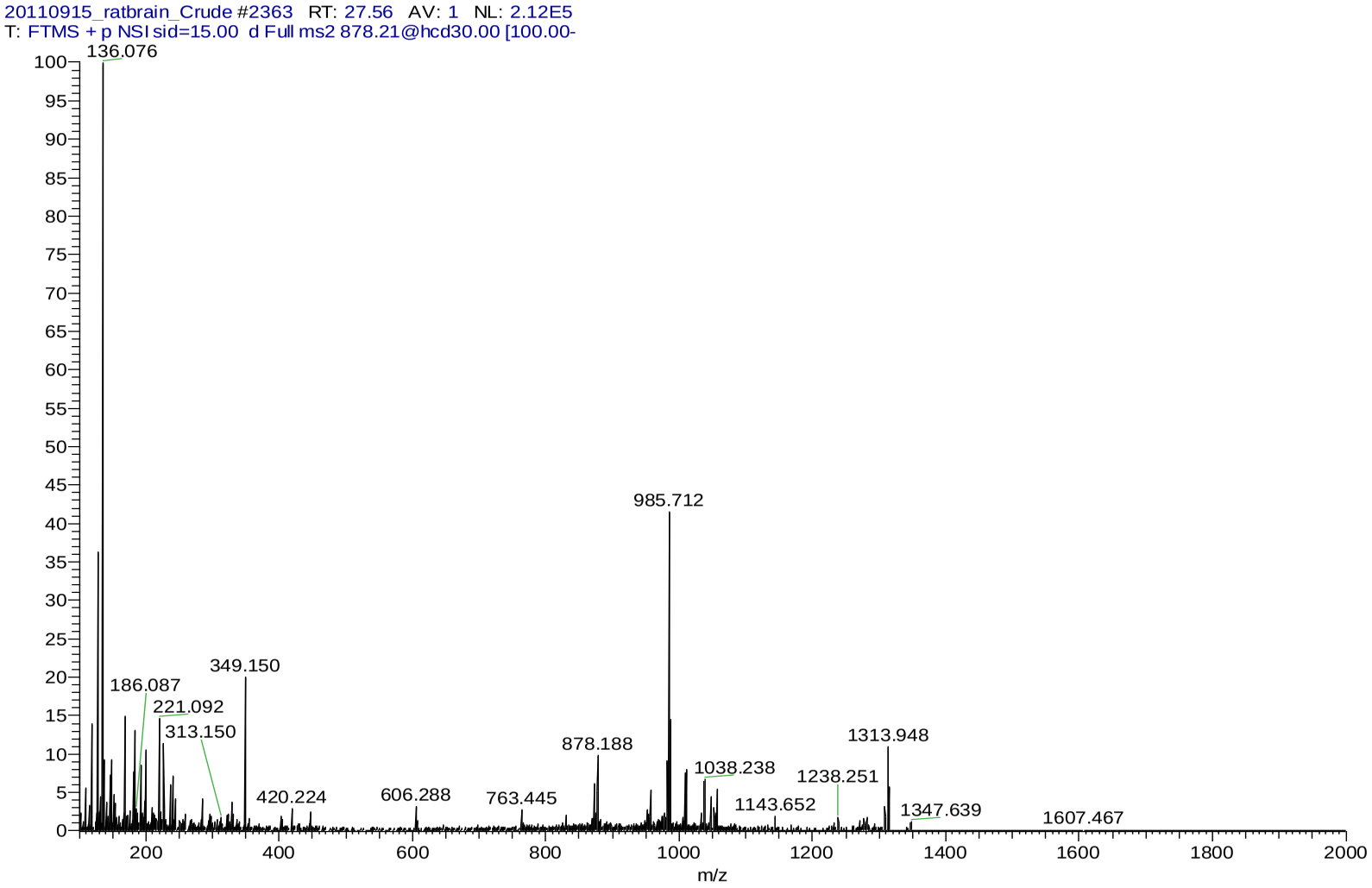
b26 -L-Y-E-K-A-P-Q-G-K-V-K- y1



NPY_RAT
P07808
Proneuropeptide Y
4285.07
4285.08
-2
Oxidation, C-term amidation
6.2E-34
b1 -Y-P-S-K-P-D-N-P-G-E-D-A-P-A-E-D-M-A-R-Y-Y-S-A-L-R- y12
b26 H-Y-I-N-L-I-T-R-Q-R-Y- y1
136.075
155.082
233.128
327.129
399.151
470.188
519.230
601.115
647.852
750.891
768.650
797.407
844.173
949.967
1024.531
1195.256
1907.784
m/z

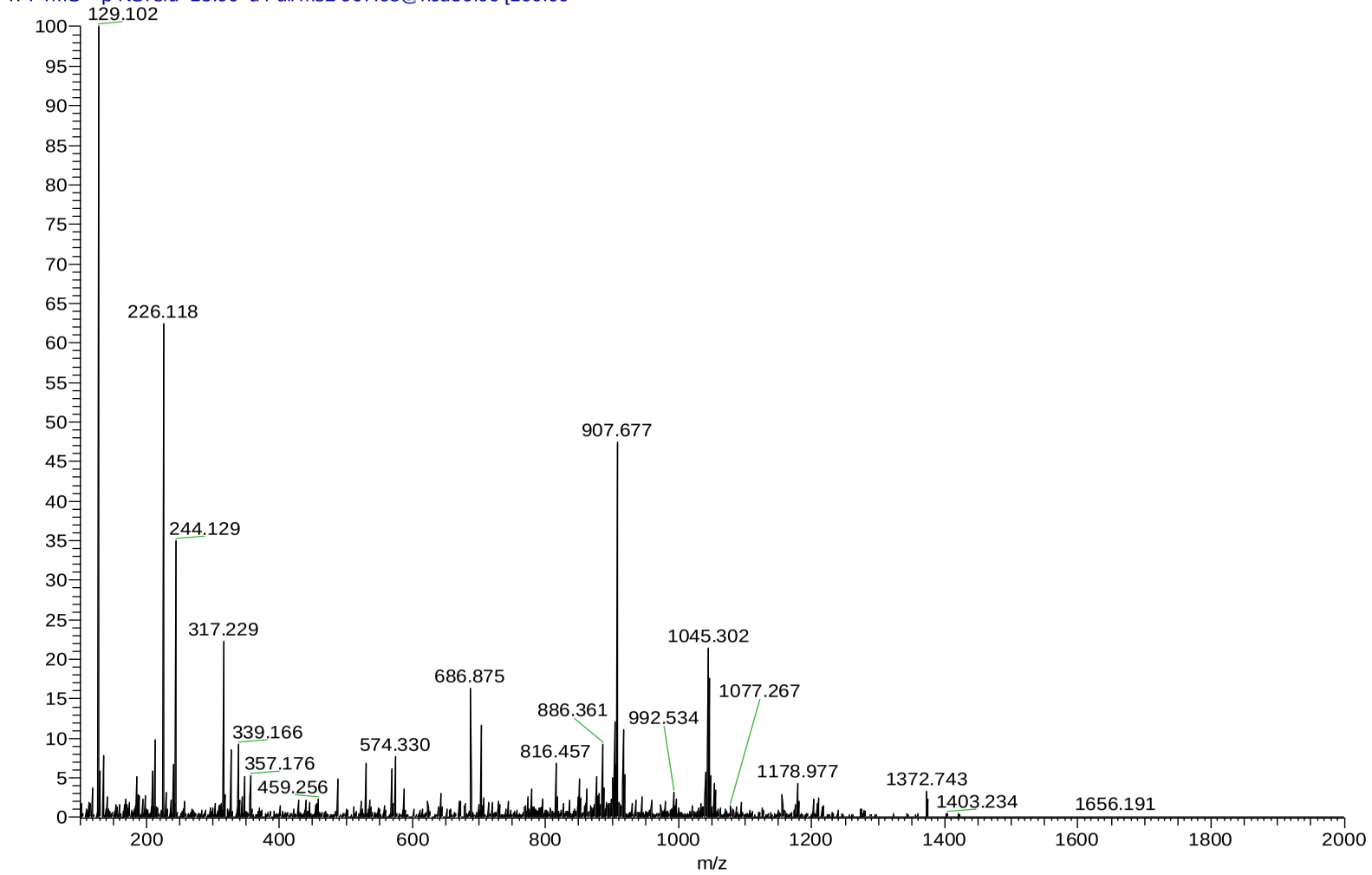
CART_RAT	P49192	Cocaine- and amphetamine-regulated transcript protein	4384.02	4384.03	-2	3*Disulfide bonds	1.3E-11





TYB10_RAT	P63312	Thymosin beta-10	4531.34	4531.35	-2	N-term acetylation, C-term methylation	1.7E-69

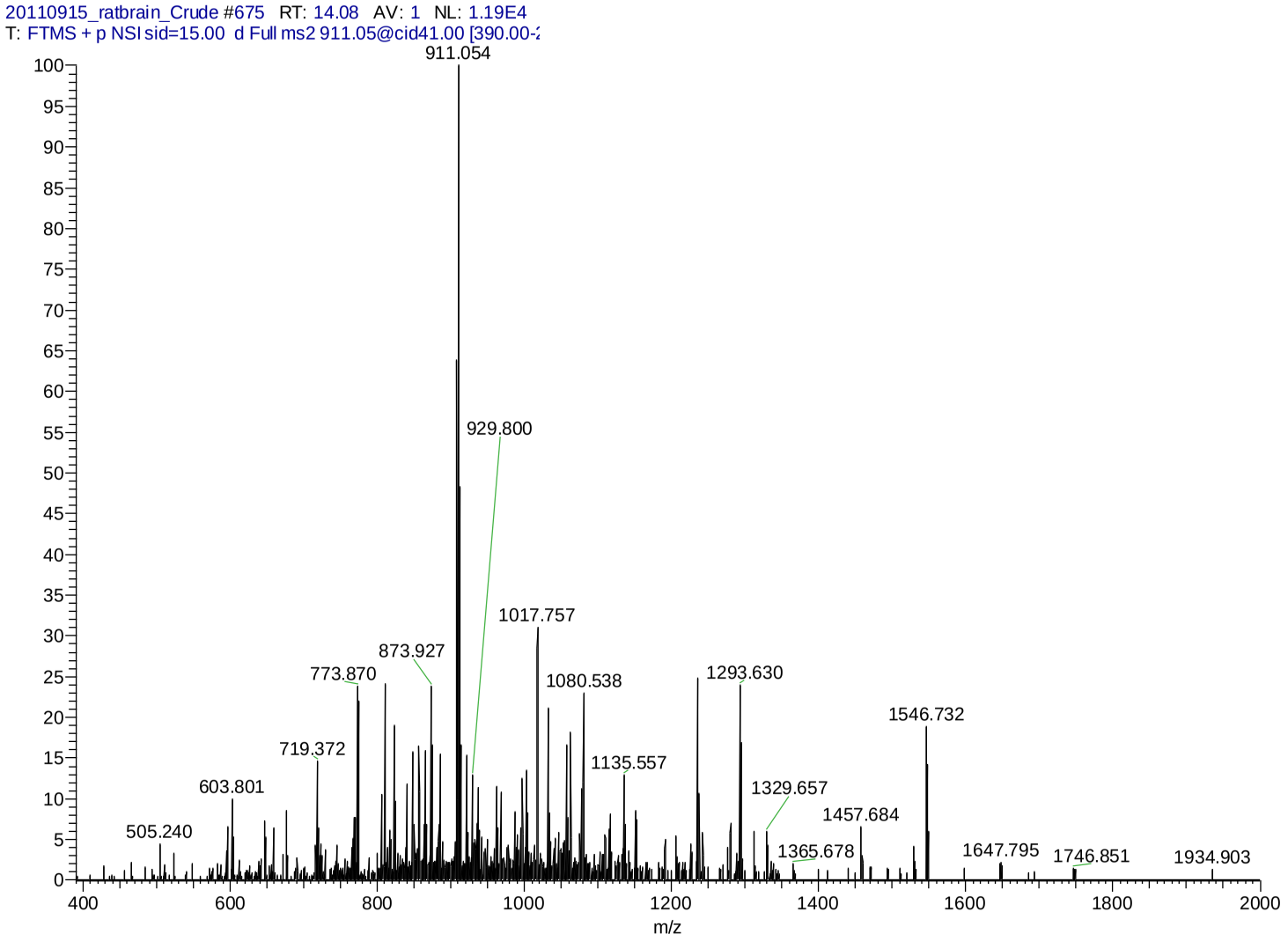
-A-DKPDMGEIASFDKAKLKKT-ETQEK
NTLP-TKETIEQEK-R



NCAM1_RAT	P13596	Neural cell adhesion molecule 1	4548.23	4548.23	0		7.5E-41

b1 - E P E K G - P - V - E T - K S - E P - Q E S E - A K P - A P - T - E V - y18

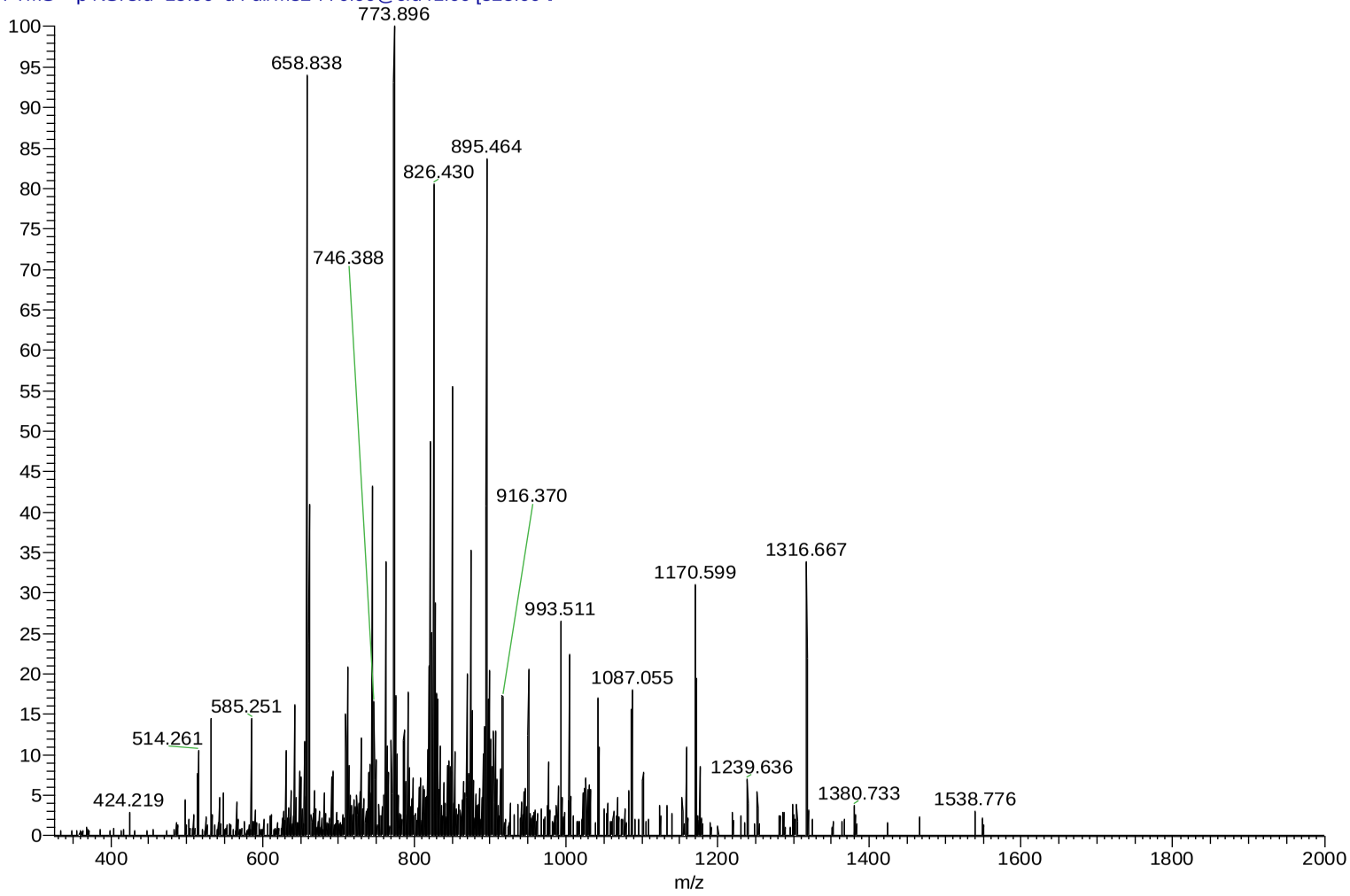
b26 - K T - V P - N - E A T Q - T - K E N - E S - K - A - y1



TYB4_RAT	P62329	Thymosin beta-4	4616.35	4616.35	0	N-term acetylation	5.0E-62

-S-D⟊K⟊P-D⟊M-A-E⟊I⟊E⟊K⟊F-D⟊K-S-K⟊L⟊K⟊K⟊T-E⟊T-Q⟊E⟊K⟊
⟊N⟊P⟊L⟊P-S-K⟊E⟊T⟊I⟊E⟊Q-E⟊K⟊Q

20110915_ratbrain_Crude #1298 RT: 19.68 AV: 1 NL: 4.52E5
T: FTMS + p NSI sid=15.00 d Full ms2 770.90@cid41.00 [325.00-

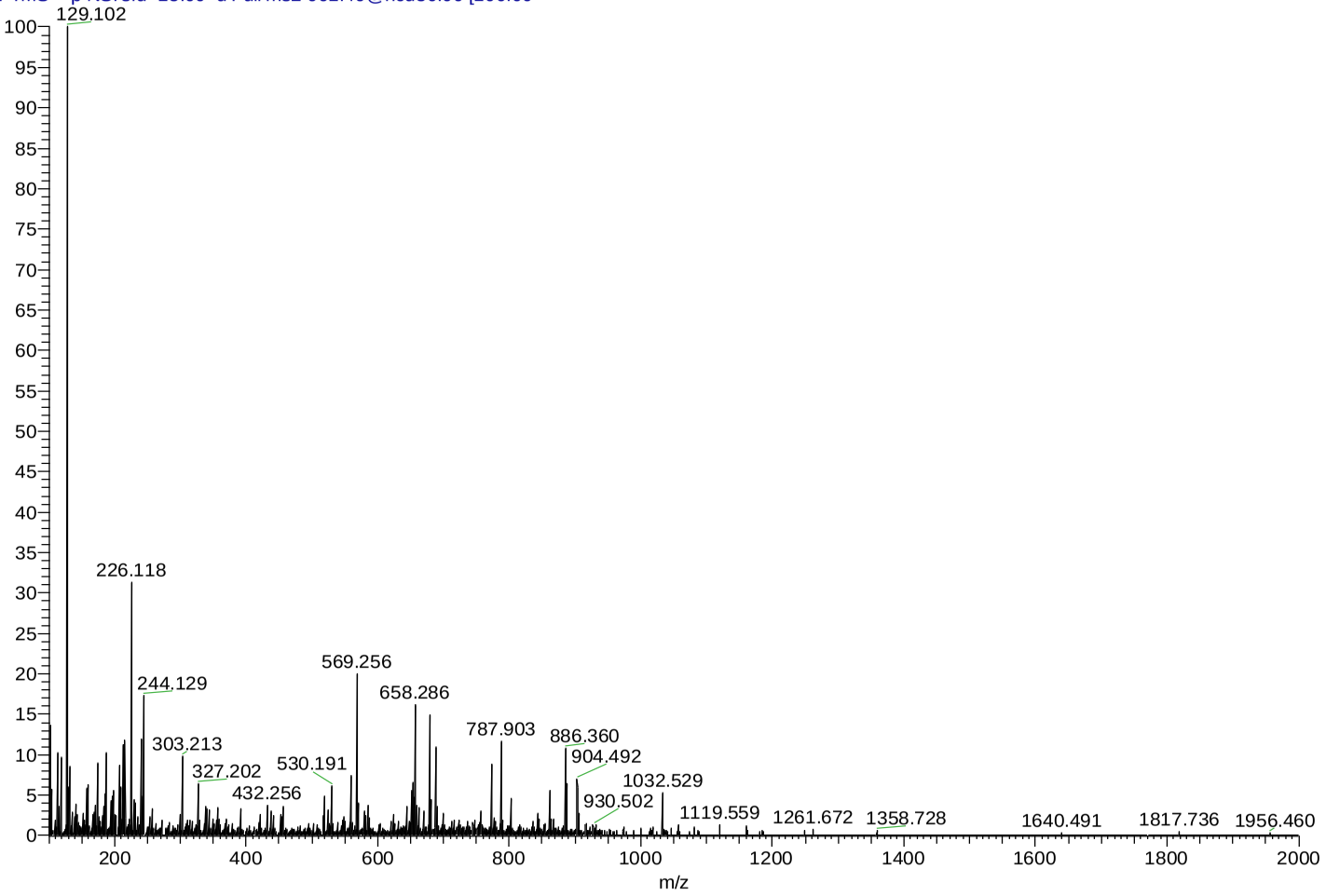


TYB10_RAT	P63312	Thymosin beta-10	4620.36	4620.36	0		9.4E-13

b1 - D - K ǀ P ǀ D - M - G - E ǀ I - A - S - F - D - K - A - K - L - K - K - T - E - T - Q - E - K - N - y16

b26 - T ǀ L ǀ P - T - K ǀ E ǀ T - I ǀ E - Q - E - K - R - S - E - y1

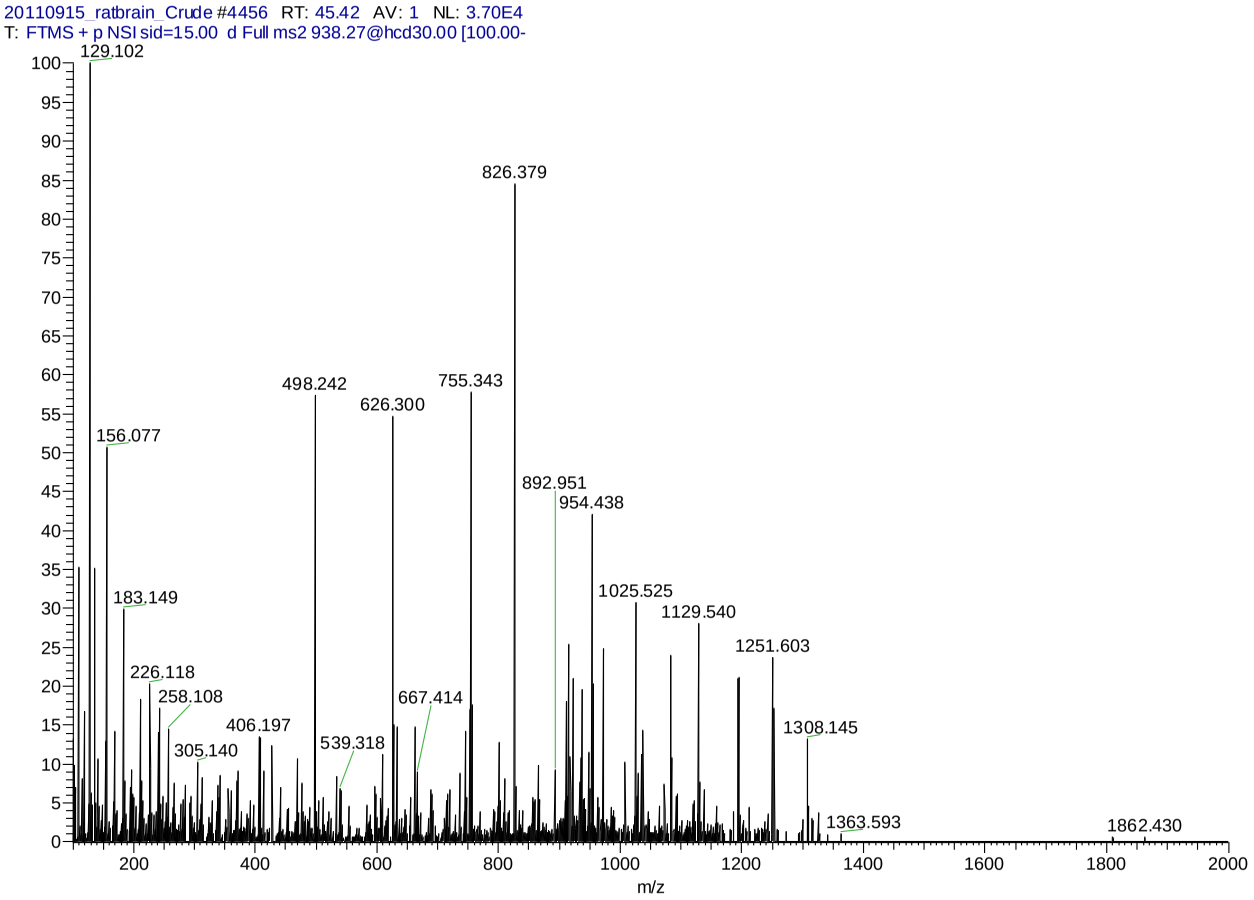
20110915_ratbrain_Crude #1713 RT: 22.32 AV: 1 NL: 6.33E5
T: FTMS + p NSI sid=15.00 d Full ms2 661.49@hcd30.00 [100.00-



SCG2_RAT	P10362	Secretogranin-2	4683.28	4683.29	-2		2.1E-70

b1 - I - P - A - G - S - L - K - N - E - D - T - P - N - R - Q - Y - L - D - E - D - M - L - L - K - V - y16

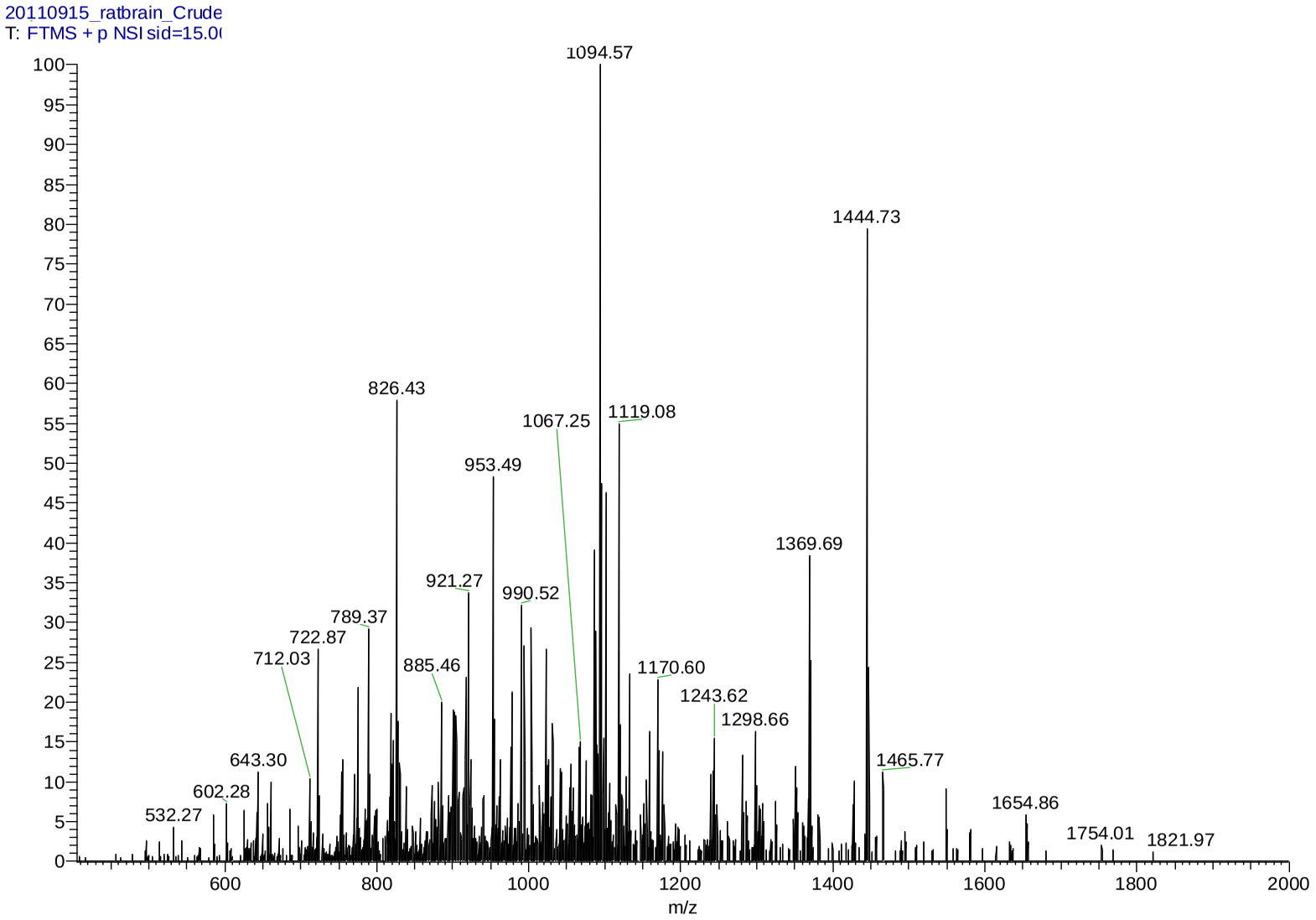
b26 - L - E - Y - L - N - Q - E - Q - A - E - Q - G - R - E - H - y1



TYB4_RAT	P62329	Thymosin beta-4	4744.40	4744.41	-2	N-term acetylation	3.1E-68

b1 -S-D-K-P-D-M-A-E-I-E-K-F-D-K-S-K-L-K-K-T-E-T-Q-E-K-

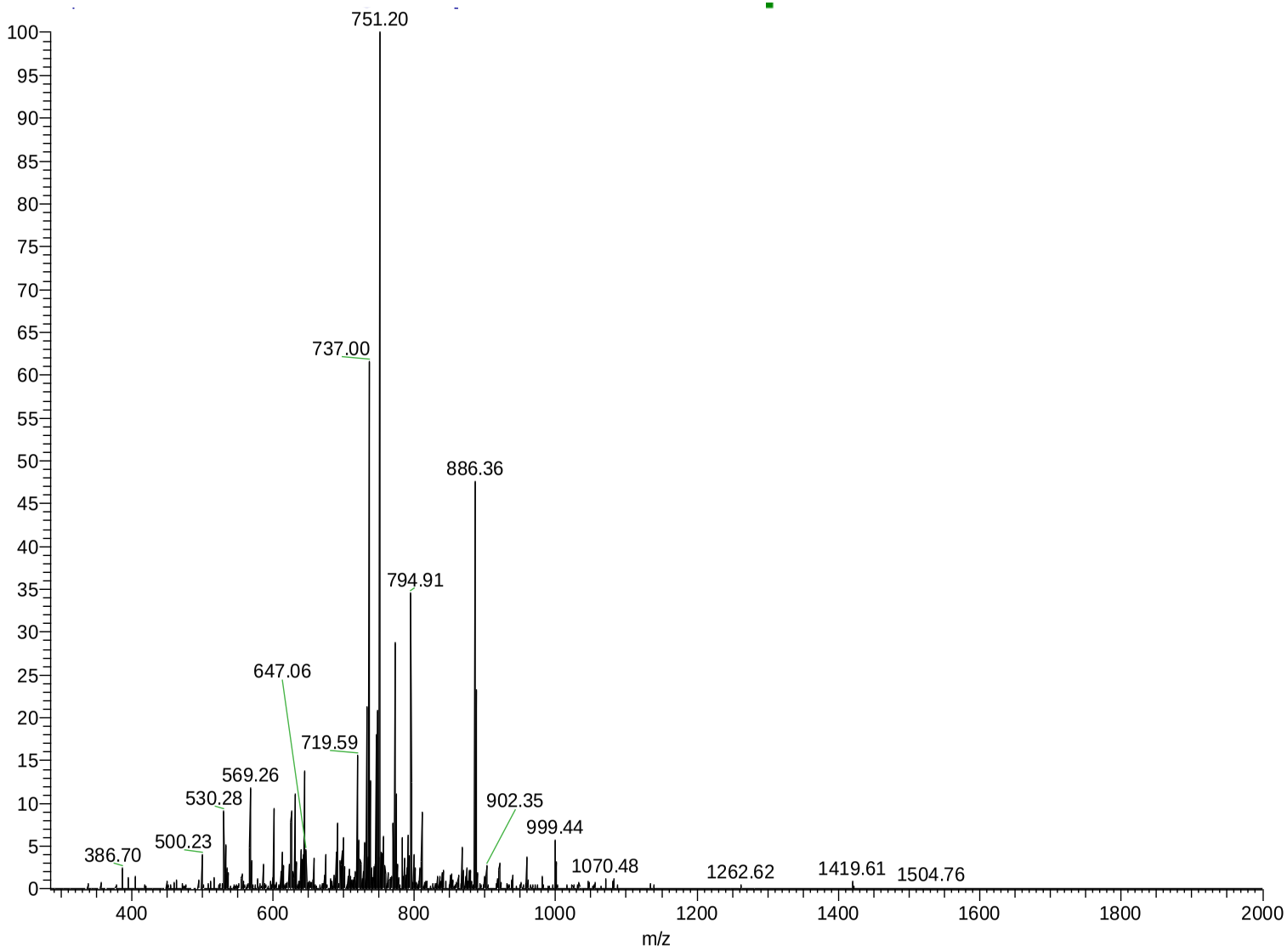
b26 -N-P-L-P-S-K-E-T-I-E-Q-E-K-Q-A-G



TYB10_RAT	P63312	Thymosin beta-10	4747.42	4747.42	0	N-term acetylation, C-term/R methylation	6.9E-72

b1 -A-D-K-P-D-M-G-E-I-A-S-F-D-K-A-K-L-K-K-T-E-T-Q-E-K-

b26 -N-T-L-P-T-K-E-T-I-E-Q-E-K-R-S-E-I-S-

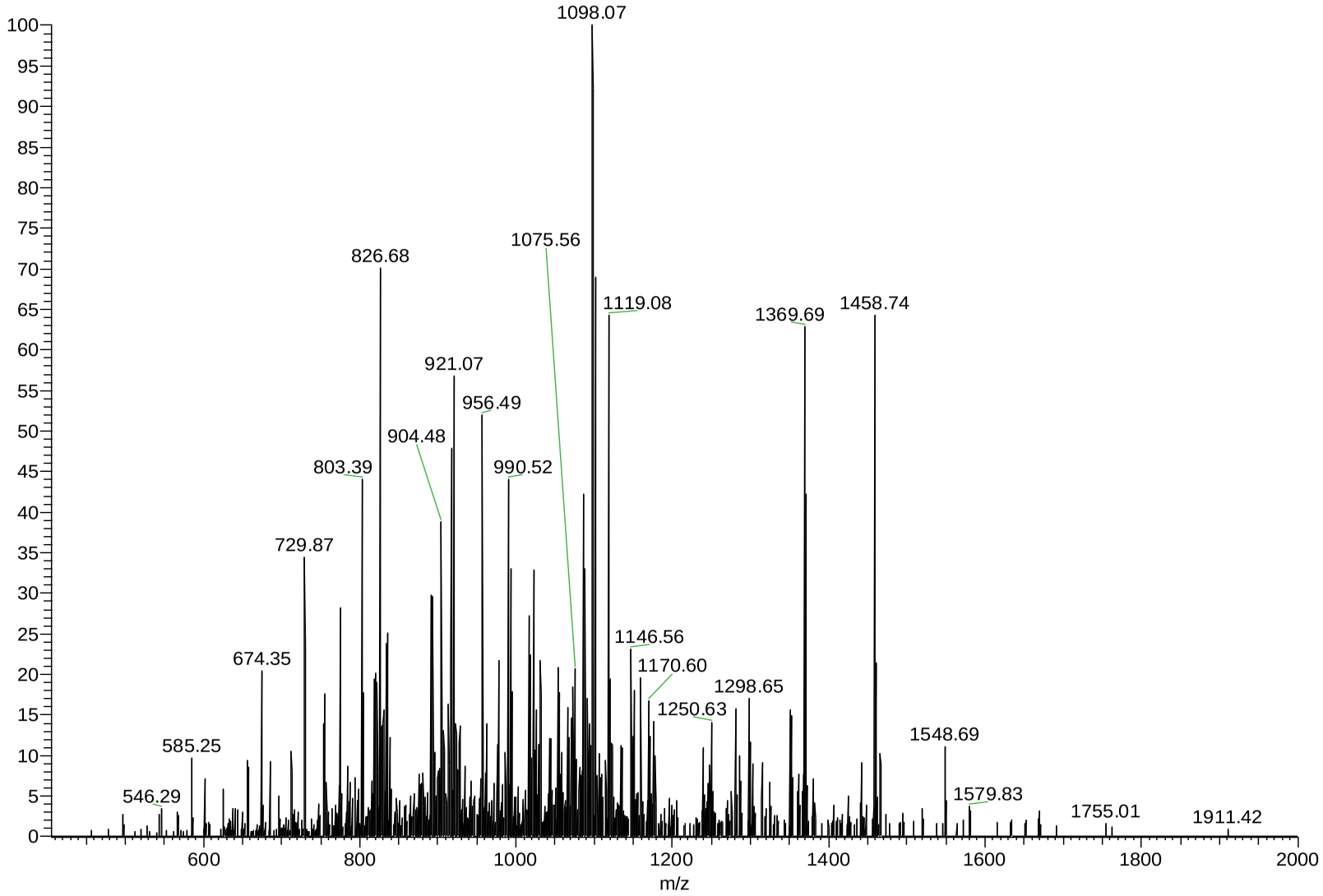


384

TYB10_RAT	P63312	Thymosin beta-4	4758.42	4758.43	-2	N-term acetylation, C-term methylation	1.2E-81

b1 -S-D-K-P-D-M-A-E-I-E-K-F-D-K-S-K-L-K-K-T-E-T-Q-E-K-

b26 -N-P-L-P-S-K-E-T-I-E-Q-E-K-Q-A-G-

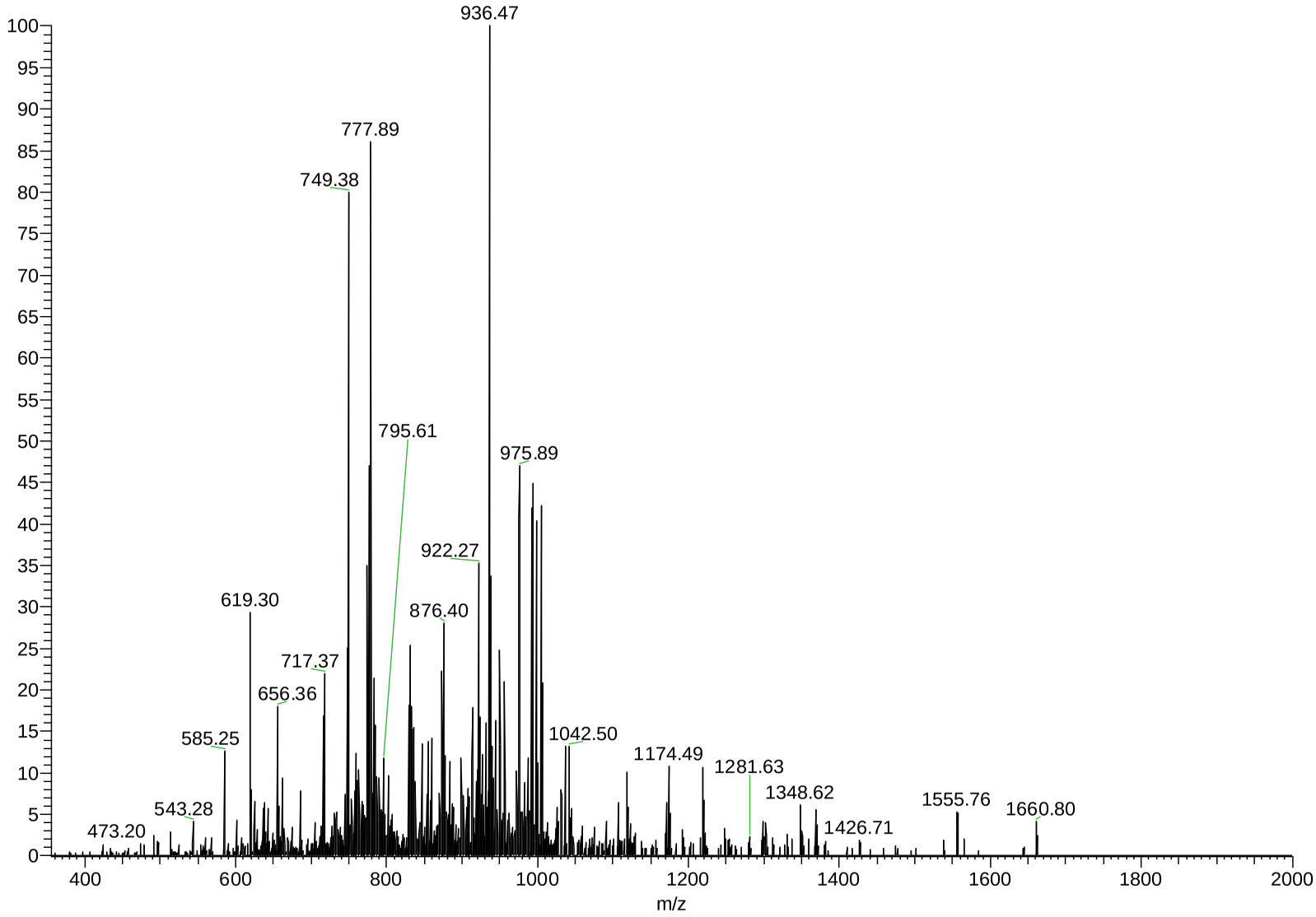


D3ZWN0_RAT
D3ZWN0
Protein Plekha6
4774.41
4774.42
-2
1.6E-14
b1 -K-G-Q-P-K-T-D-Y-E-P-S-K-K-D-P-G-Q-T-S-P-L{D}T{H-R- y19
b26 -D-I-S{L-V-P-T-R-Q-E{V-E-A{E}K}Q}A}A- y1
129.10
226.12
209.09
271.14
355.16
417.24
546.29
585.25
674.35
729.87
803.39
954.45
1031.19
1107.24
1195.11
1298.65
1369.69
1458.74
1564.68
1668.89
1881.48
m/z

TYB10_RAT	P63312	Thymosin beta-4	4976.47	4976.48	-2	N-term acetylation, oxidation	7.7E-86

b1 -S-D⟩K⟩P-D⟩M-A-E⟩I-E⟩K⟩F-D⟩K⟩S-K⟩L⟩K⟩K⟩T⟩E⟩T-Q-E⟩K⟩ y19

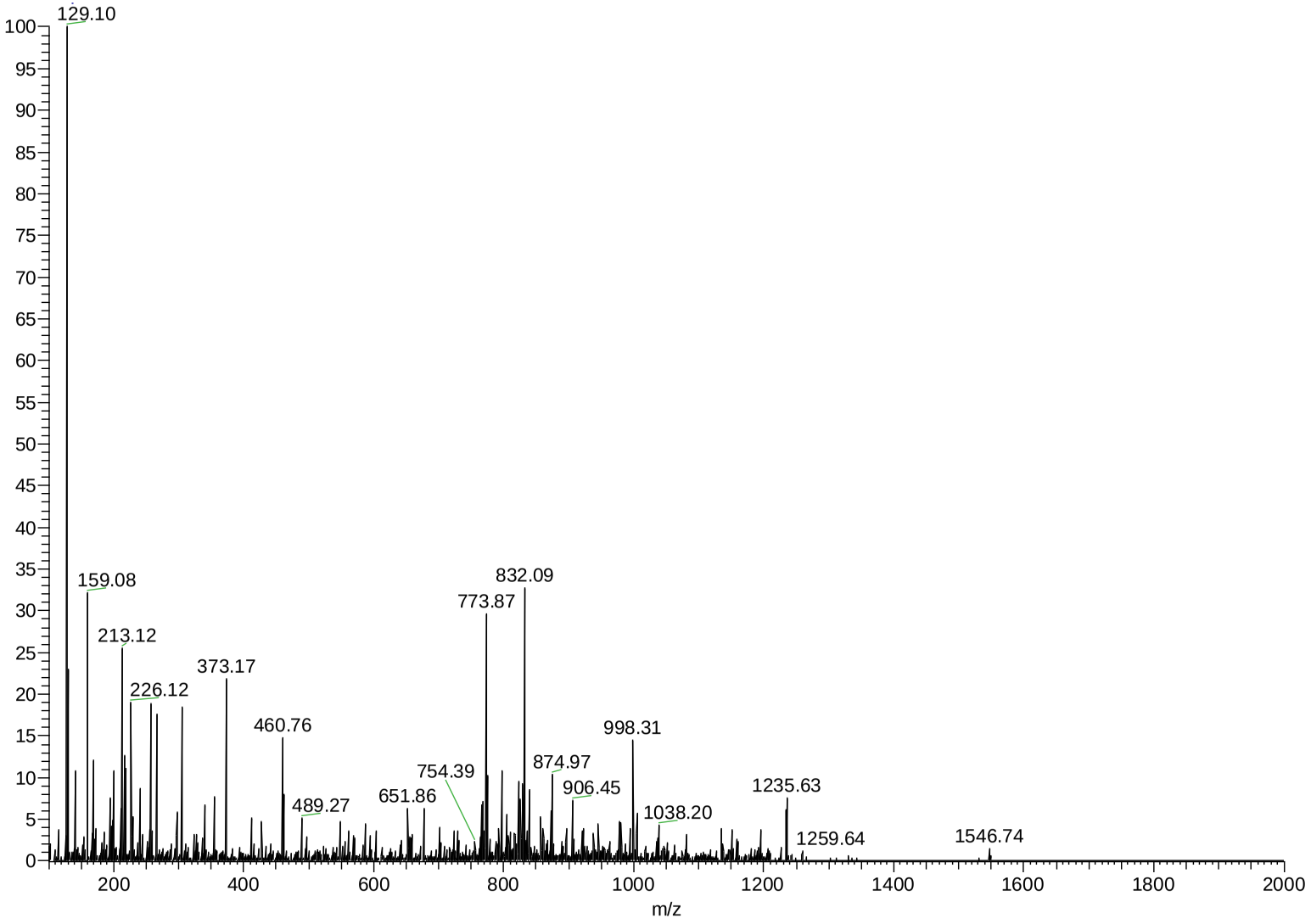
b26 ⟨N⟩P⟩L⟩P-S⟩K⟩E⟩T⟩I⟩E⟩Q⟩E⟩K⟩Q⟩A⟩G⟩E⟩S- y1



F1LUV9_RAT	F1LUV9	Uncharacterized	5241.65	5241.65	0		3.0E-39

b1 -S-A-V-P-P-A-P-A-K-T-E-K-G-P-V-E-T-K-S-E-P-Q-E-S-E- y26

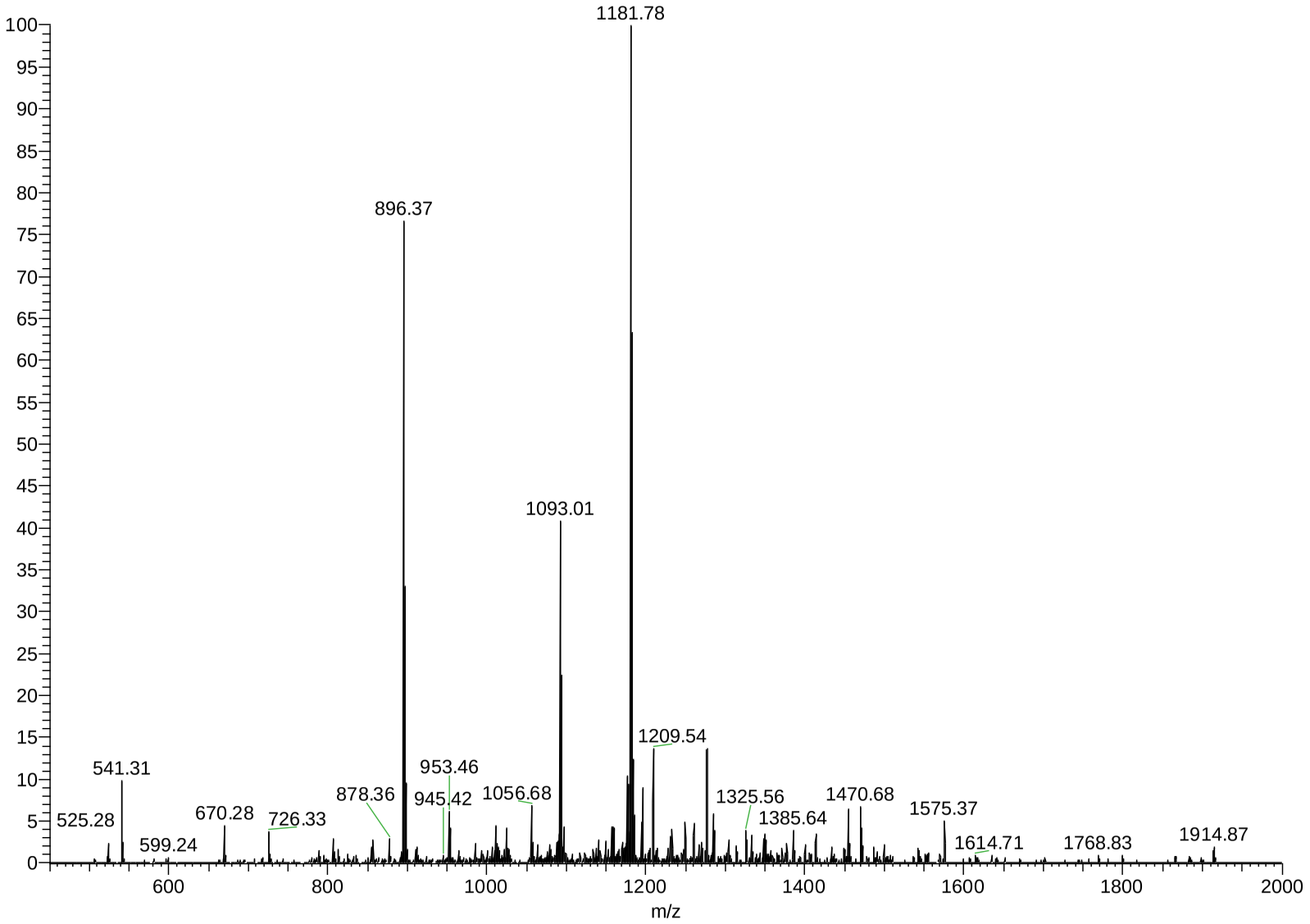
b26 -A-K-P-A-P-T-E-V-K-T-V-P-N-E-A-T-Q-T-K-E-N-E-S-K-A- y1



NEUM_RAT	P07936	Neuromodulin	5261.39	5261.40	-2		5.7E-86

b1 -A-A-K-A-A-Q-P-P-T-E-T-A-E-S-S-Q-A-E-E-E-K-E-A-V-D- y25

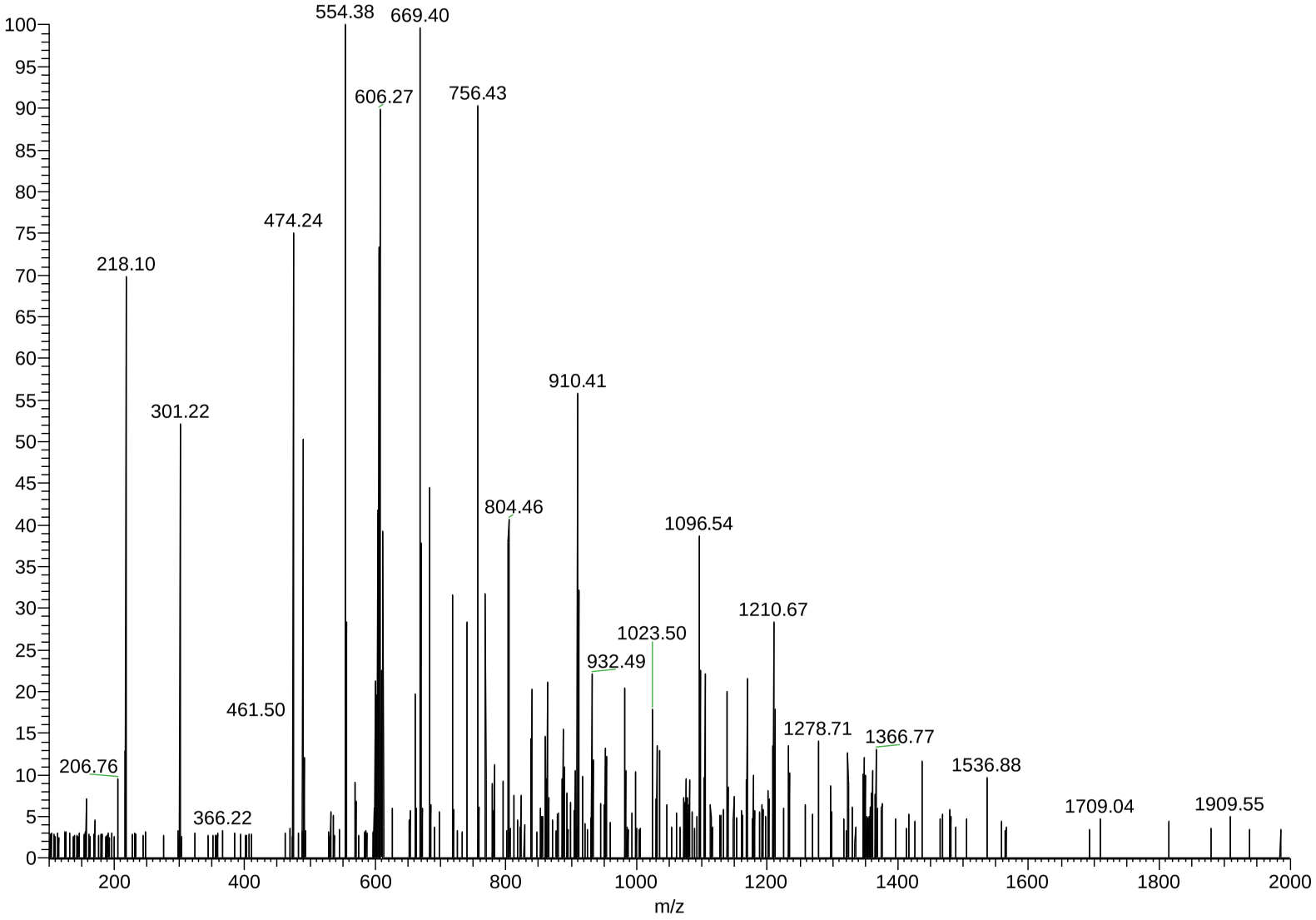
b26 -E-A-K-P-K-E-S-A-R-Q-D-E-G-K-E-D-P-E-A-D-Q-E-H-A- y1



CCKN_RAT	P01355	CCK	5434.99	5435.00	-2		7.7E-44

c1 -A-V-L-R-P-D-S-E-P-R-A-R-L-G-A-L-L-A-R-Y-I-Q-Q-V-R-

c26 -K-A-P-S-G-R-M-S-V-L-K-N-L-Q-G-L-D-P-S-H-R-I-S-D-R-

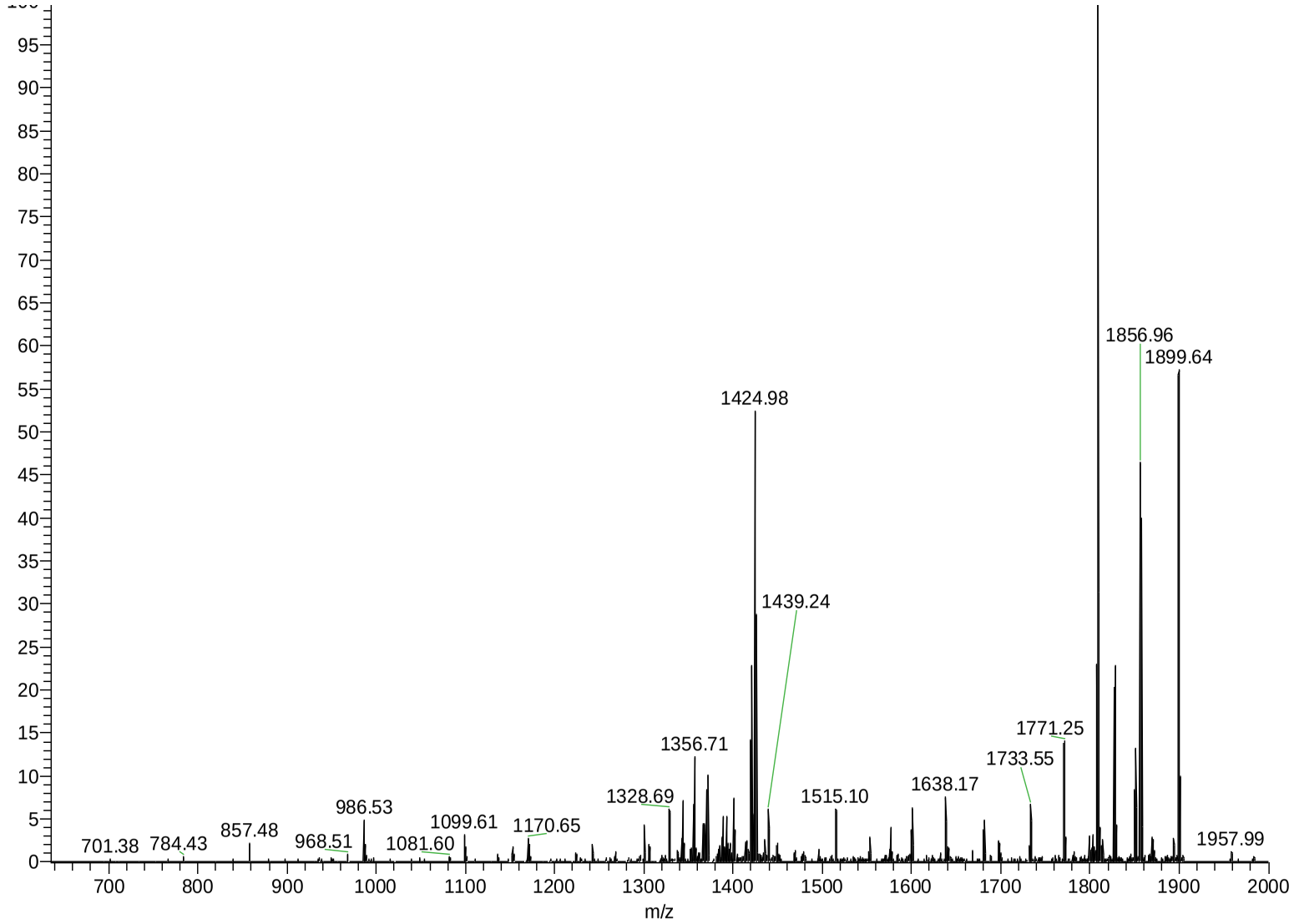


D4AA63_RAT	D4AA63	Ubiquitin	5895.01	5895.01	0		1.0E-55

b1 - G - S - P - P - Q - P - P - N - P - E - V - R - F - Q - Q - Q - L - E - Q - L - N - A - M - G - F - y31

b26 - L - N - R - E - A - N - L - Q - A - L - I - A - T - G - G - D - I - N - A - A - I - E - R - L - L - y6

b51 - G - S - Q - P - S - y1

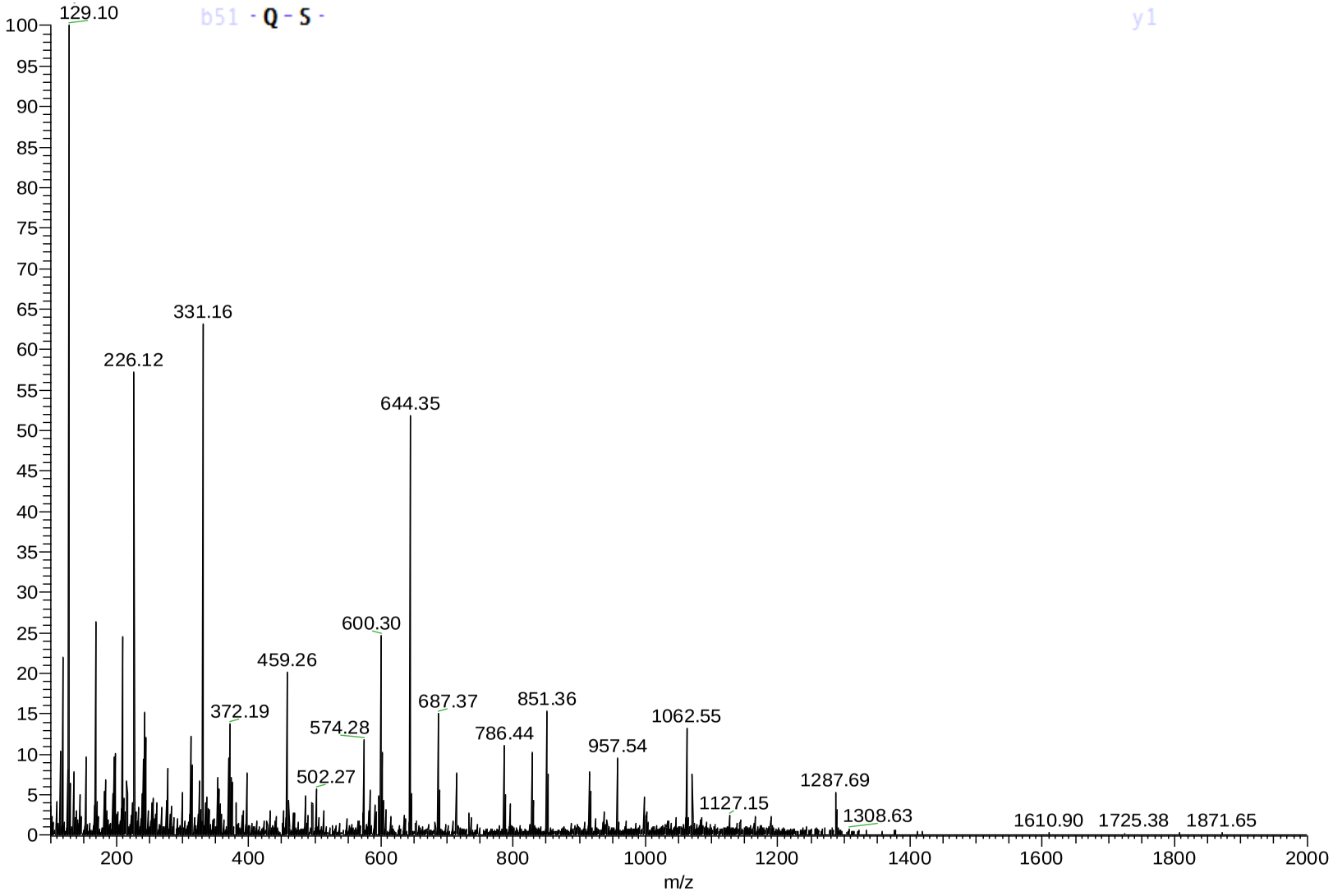


ATP5J_RAT	P21571	ATP synthase-coupling factor 6	6024.88	6024.89	-2		4.0E-15

b1 - S - G - G - P - V - D - T - G - P - E - Y - Q - Q - E - V - D - R - E - L - F - K - L - K - Q - M - y28

b26 - Y - G - K - G - E - M - D - K - F - P - T - F - N - F - E - D - P - K - F - E - V - L - D - K - P - y3

b51 - Q - S - y1

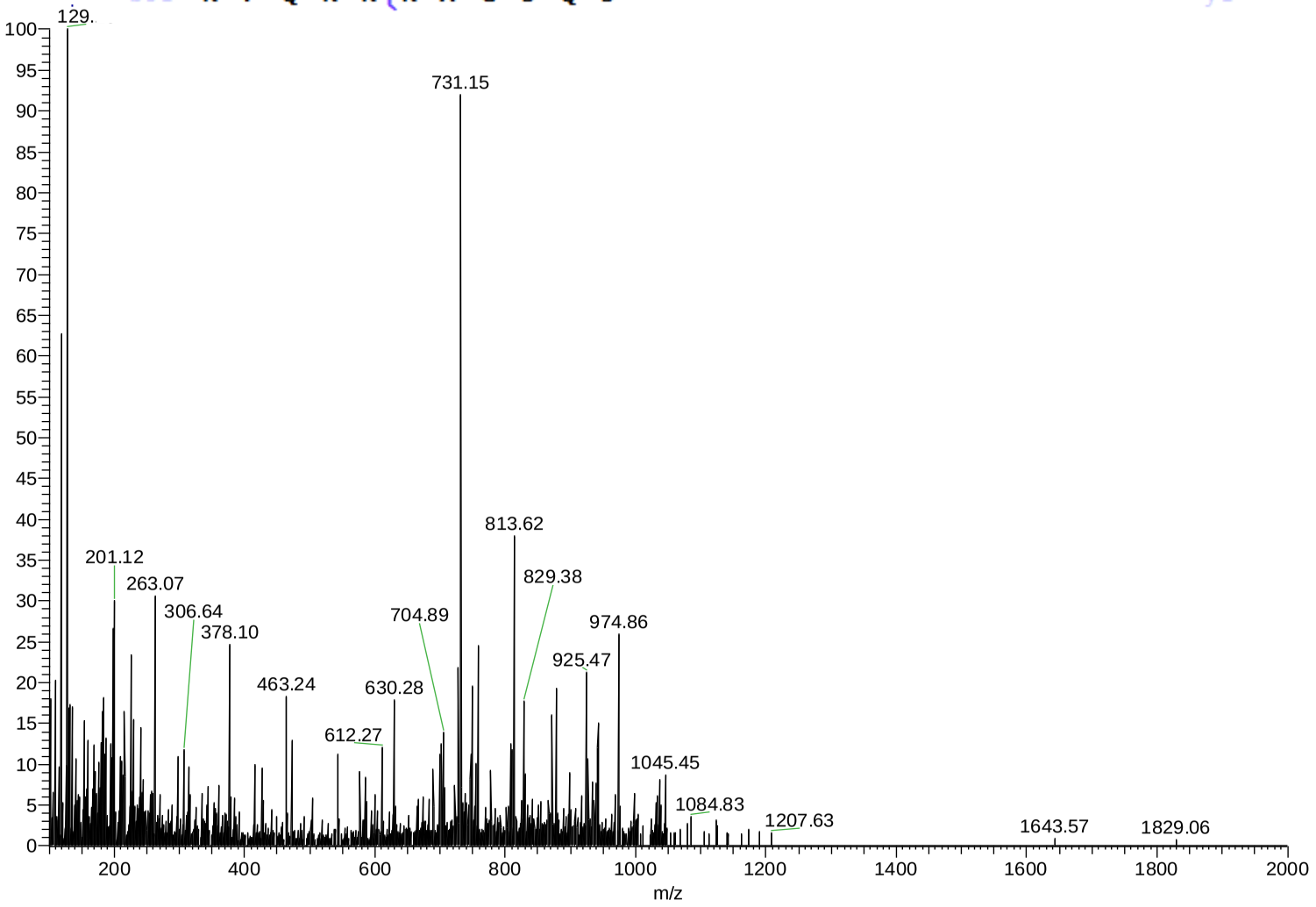


PCP4_RAT	P63055	PEP-19	6730.24	6730.25	-1	N-term acetylation, oxidation	4.4E-32

b1 - S - E - R - Q - S - A - G - A - T - N - G - K - D - K - T - S - G - D - N - D - G - Q - K - K - V - y37

b26 - Q - E - E - F - D - I - D - M - D - A - P - E - T - E - R - A - A - V - A - I - Q - S - Q - F - R - y12

b51 - K - F - Q - K - K - K - A - G - S - Q - S - y1



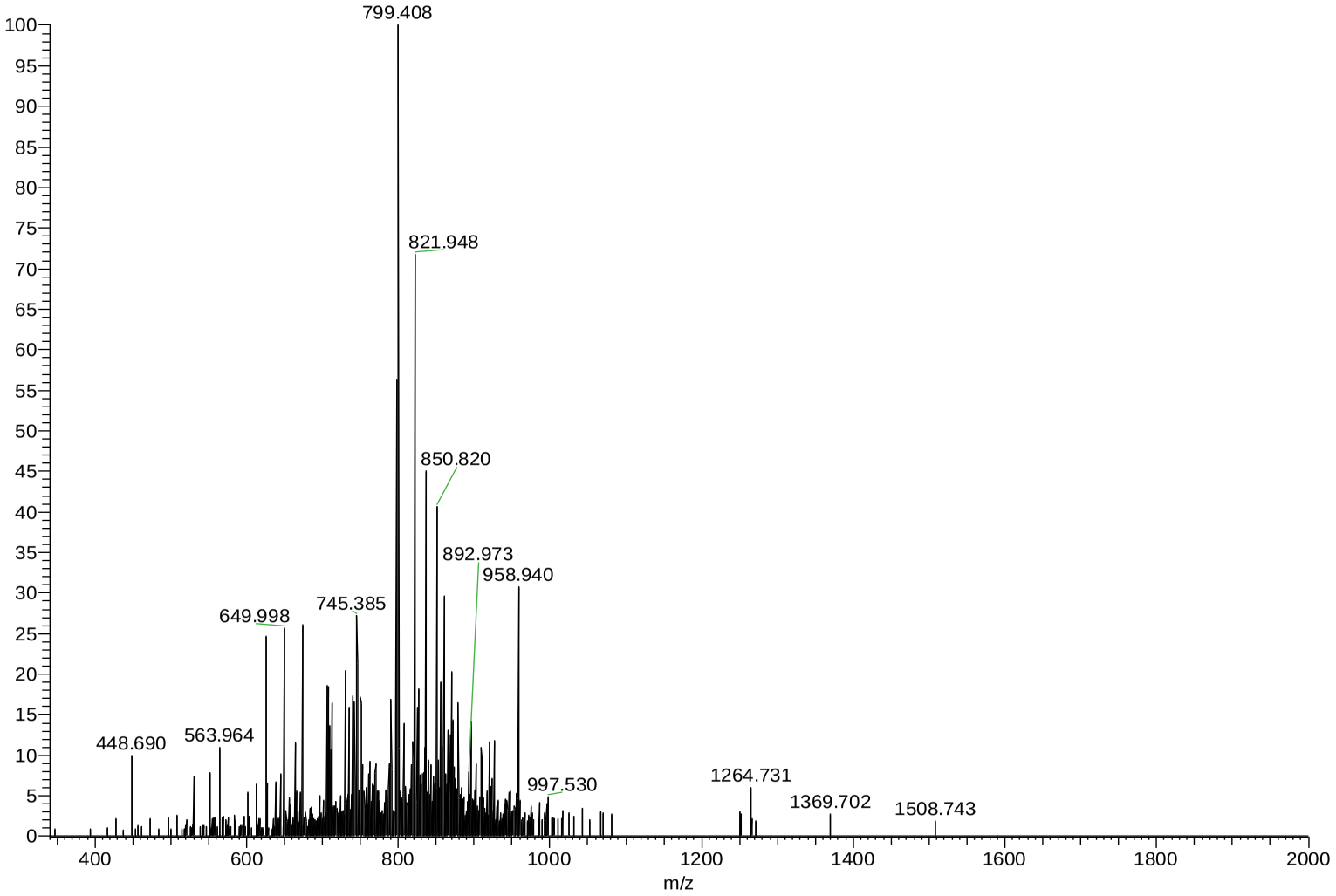
NEUM_RAT	P07936	Neuromodulin	7963.66	7963.70	-5		1.7E-25

b1 - G - P - A - K - E - E - P - K - Q - A - D - V | P - A | A | V | T | D | A | A | A | T | T | P - A | y53

b26 - A - E - D - A - A - K - A - A - Q - P - P - T - E - T - A | E - S - S - Q - A - E - E - E - K - E - y28

b51 - A - V - D - E - A - K - P - K - E - S - A - R - Q - D - E - G - K - E - D | P - E - A - D - Q - E - y3

b76 - H - A - y1



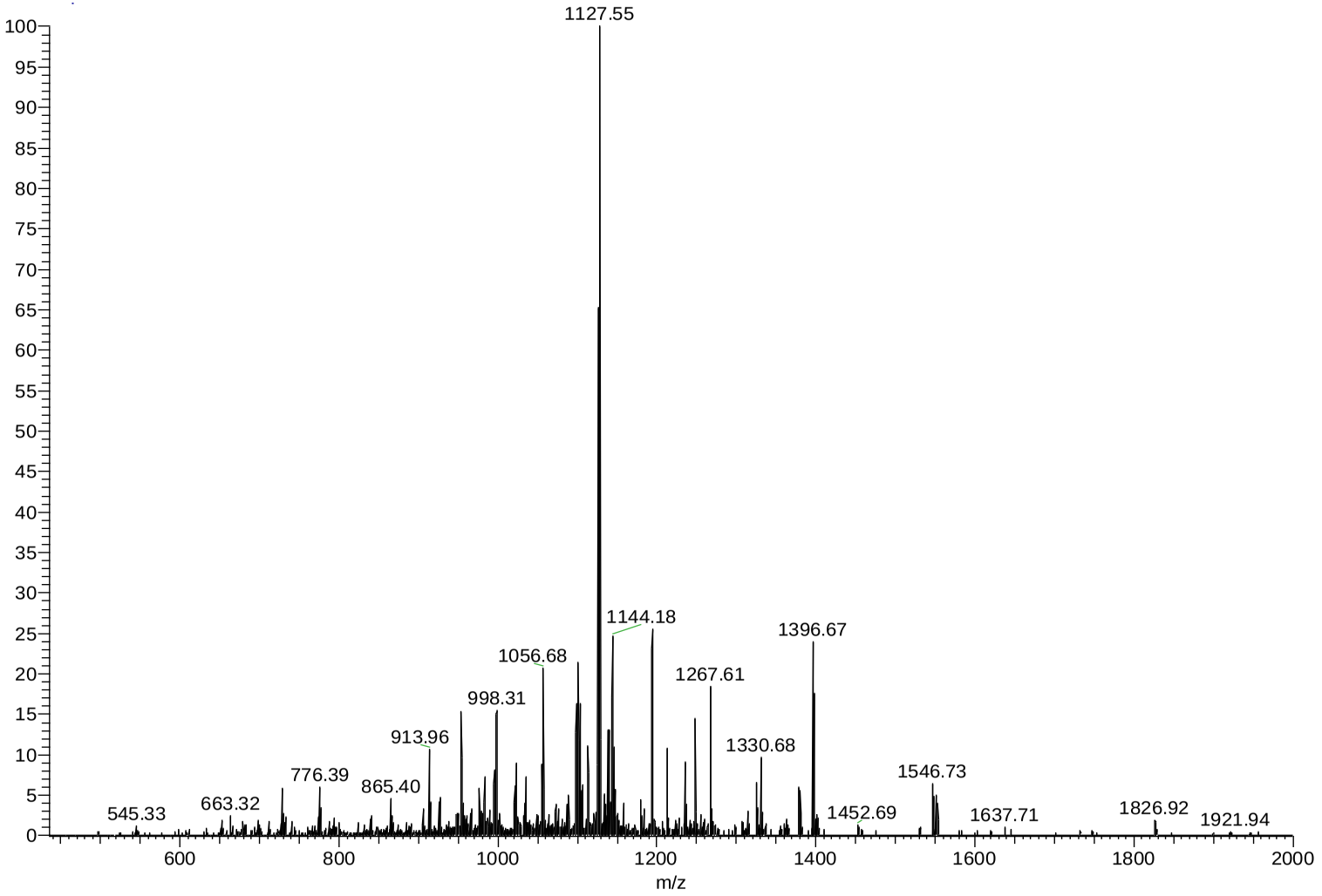
F1LNN9_RAT	F1LNN9	Uncharacterized protein	8155.96	8155.94	2		3.9E-28

b1 - G - A - P - Q - E - E - G - E - A - K - K - T - E - A - P - A - A - G - P - E - A - K - S - D - A - y62

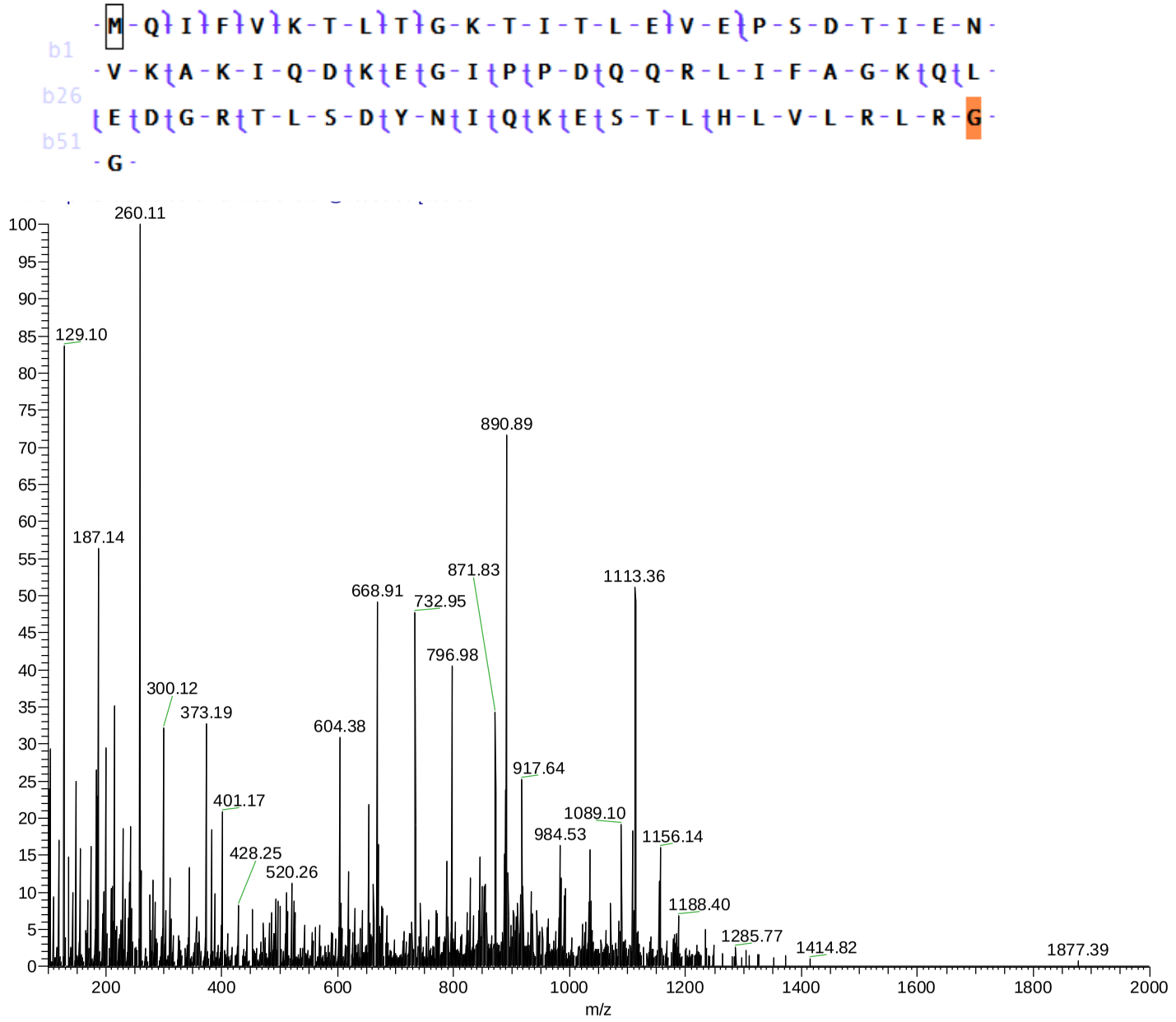
b26 - A - P - A - A - S - D - S - K - P - S - S - A - E - P - A - P - S - S - K - E - T - P - A - A - S - y37

b51 - E - A - P - S - S - A - A - K - A - P - A - P - A - A - P - A - A - E - P - Q - A - E - A - P - V - y12

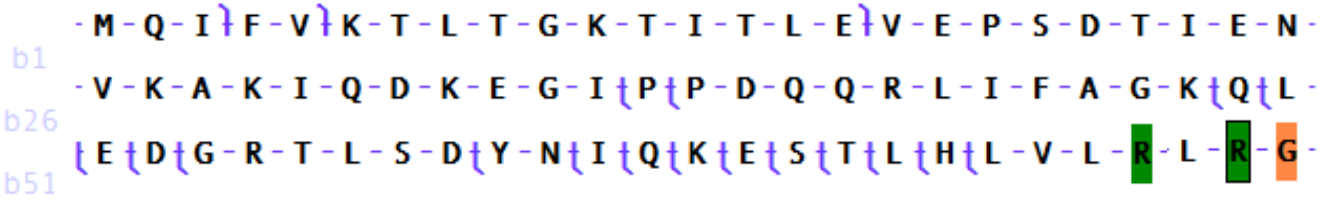
b76 - A - S - S - E - Q - S - V - A - V - K - E - y1

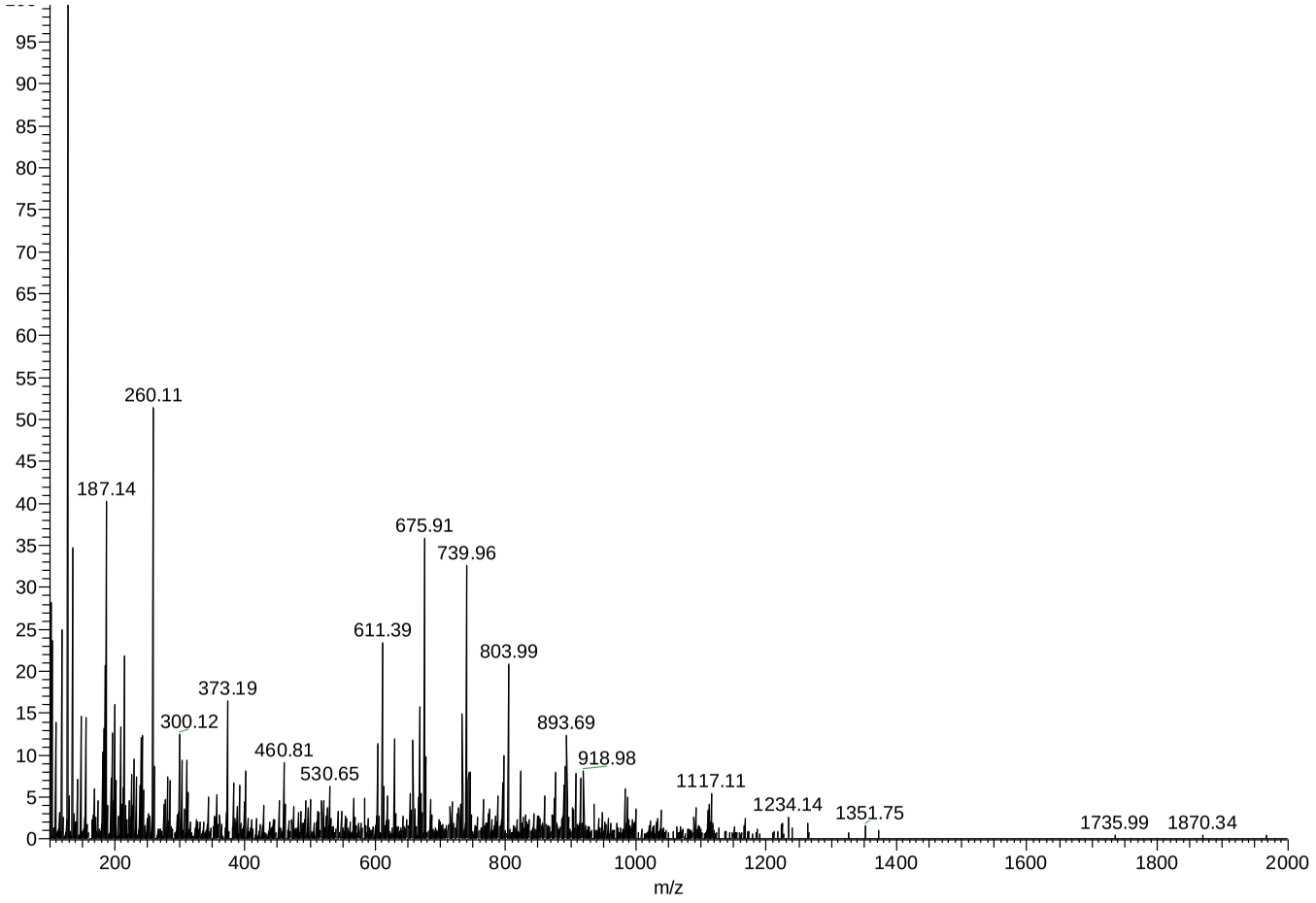


RS27A_RAT	P62982	Ubiquintin-40S ribosomal protein S27a	8445.57	8445.57	0		2.1E-43

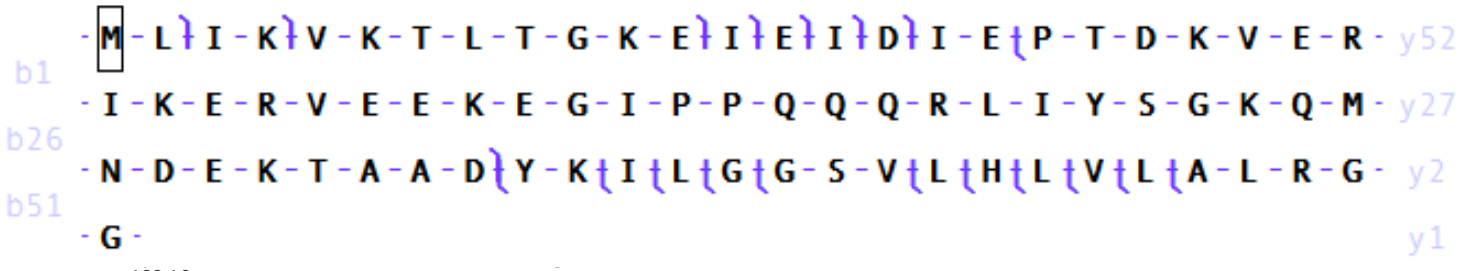


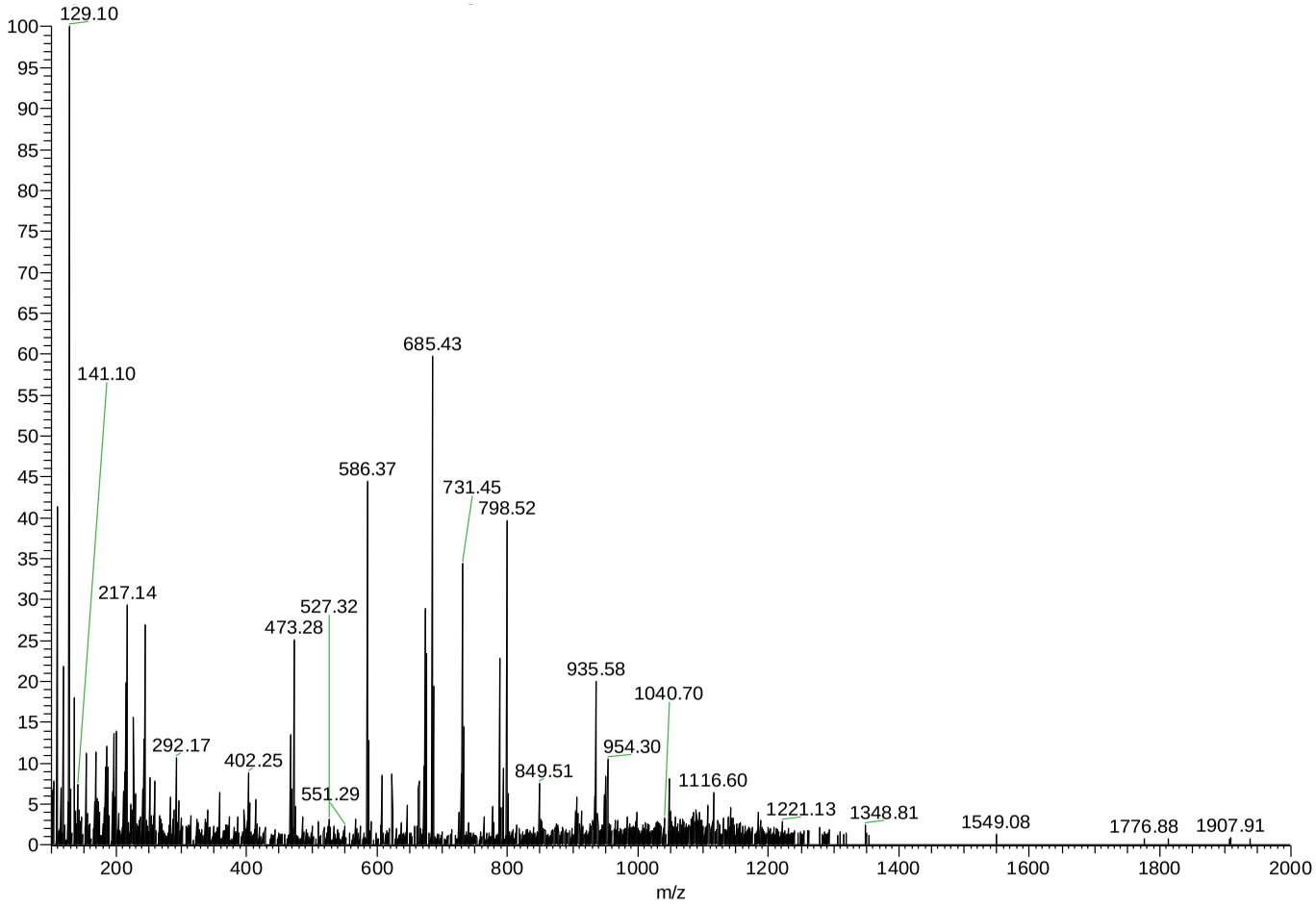
RS27A_RAT	P62982	Ubiquintin-40S ribosomal protein S27a	8459.59	8459.59	0	C-terminal/R methylation	2.7E-32





NEDD8_RAT	Q71UE8	NEDD8	8554.67	8554.67	0.00		2.1E-34





RS27A_RAT	P62982	Ubiquitin-40S ribosomal protein S27a	8575.60	8575.61	-1	Oxidation	2.0E-48

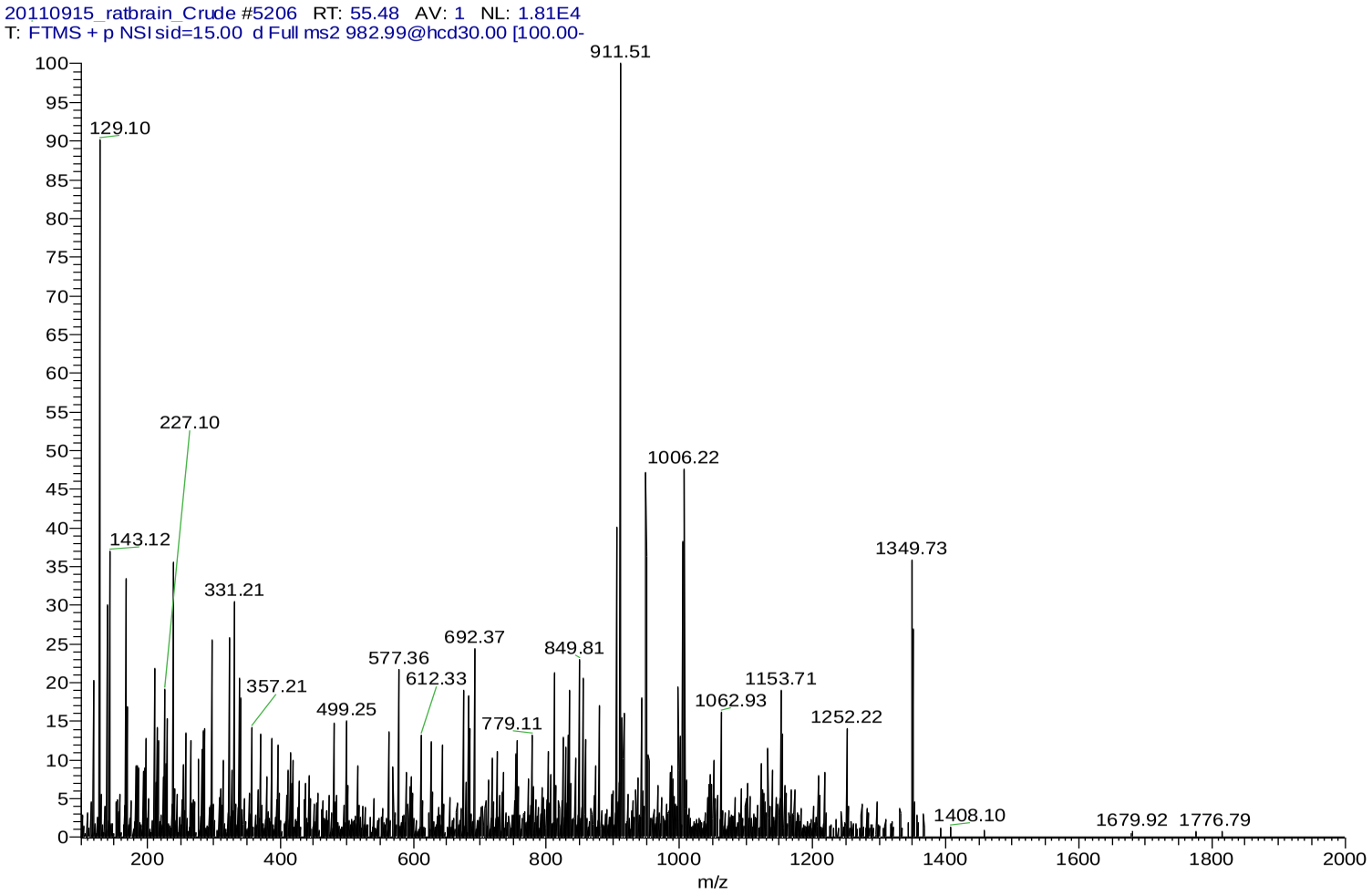
UFM1_RAT	Q5BJP3	Ubiquitin-fold modifier 1	8832.80	8832.8	0	N-terminal acetylation	1.1E-41

b1 - S - K - V - S - F - K | I | T | L | T - S - D | P - R - L | P - Y | K | V | L | S - V | P - E - S - y58

b26 - T - P - F - T - A - V | L | K - F - A - A - E - E - F - K - V - P - A - A - T - S - A - I - I - T - y33

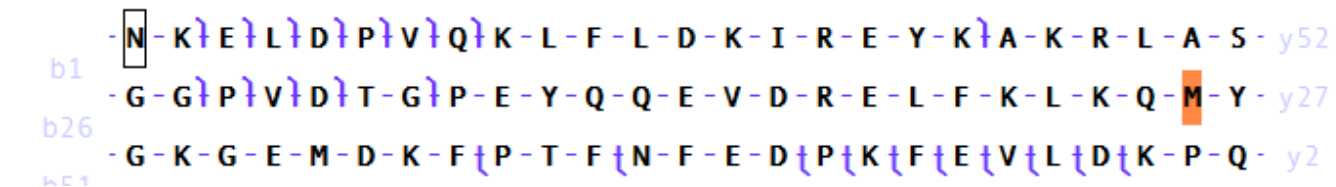
b51 - N | D | G | I | G | I | N | P | A | Q | T | A | G - N - V | F | L - K | H | G - S - E - L - R - L - y8

b76 - I - P - R - D - R - V - G - y1

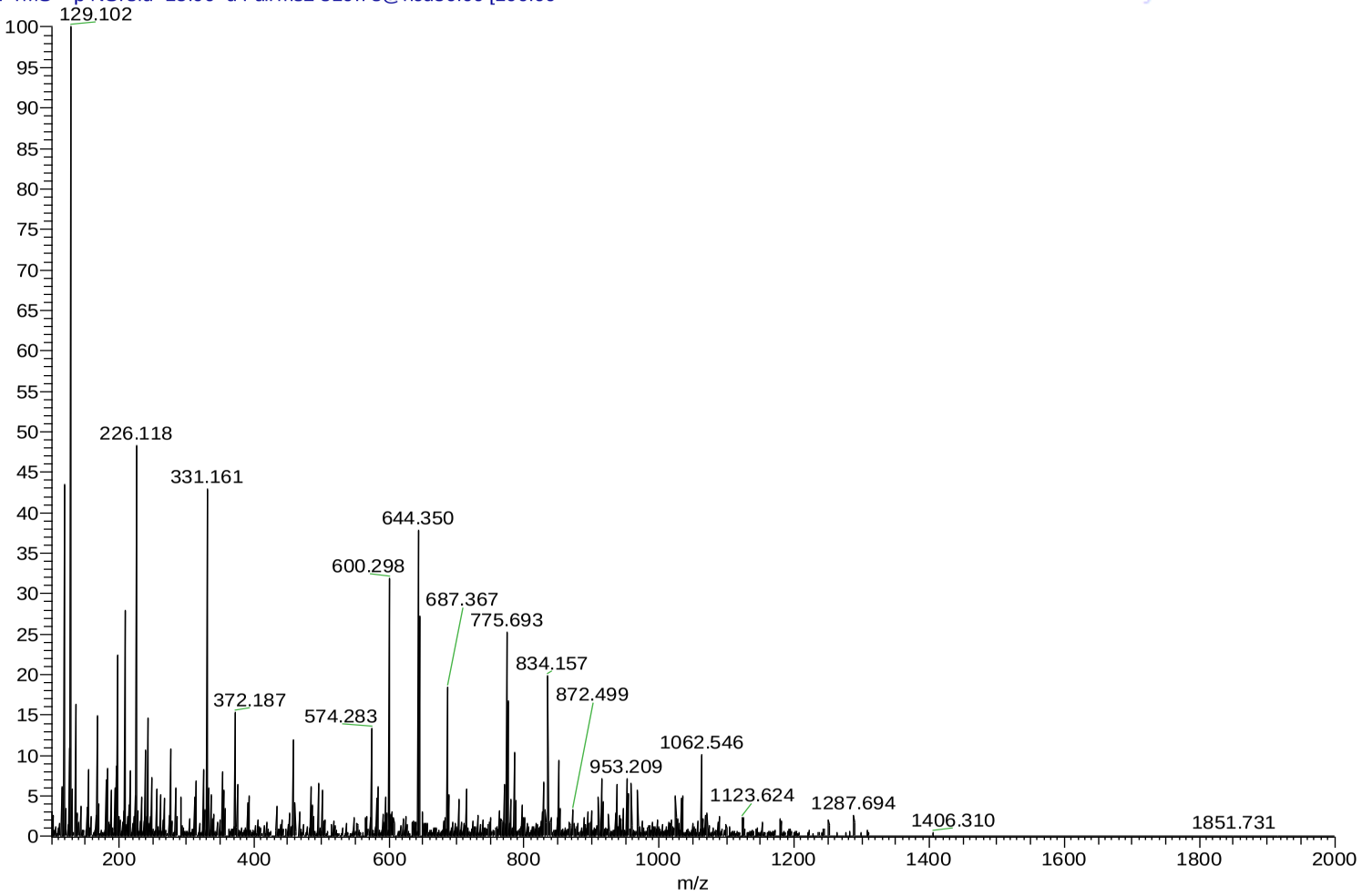


400

ATP5J_RAT	P21571	ATP synthase-coupling factor 6	8937.52	8937.55	-3	Oxidation	1.5E-37



20110915_ratbrain_Crude #4139 RT: 41.92 AV: 1 NL: 7.90E5
T: FTMS + p NSI sid=15.00 d Full ms2 819.78@hcd30.00 [100.00-



ACBP_RAT	P11030	Acyl-CoA-binding protein	9762.01	9762.02	-1	N-term acetylation	9.4E-13

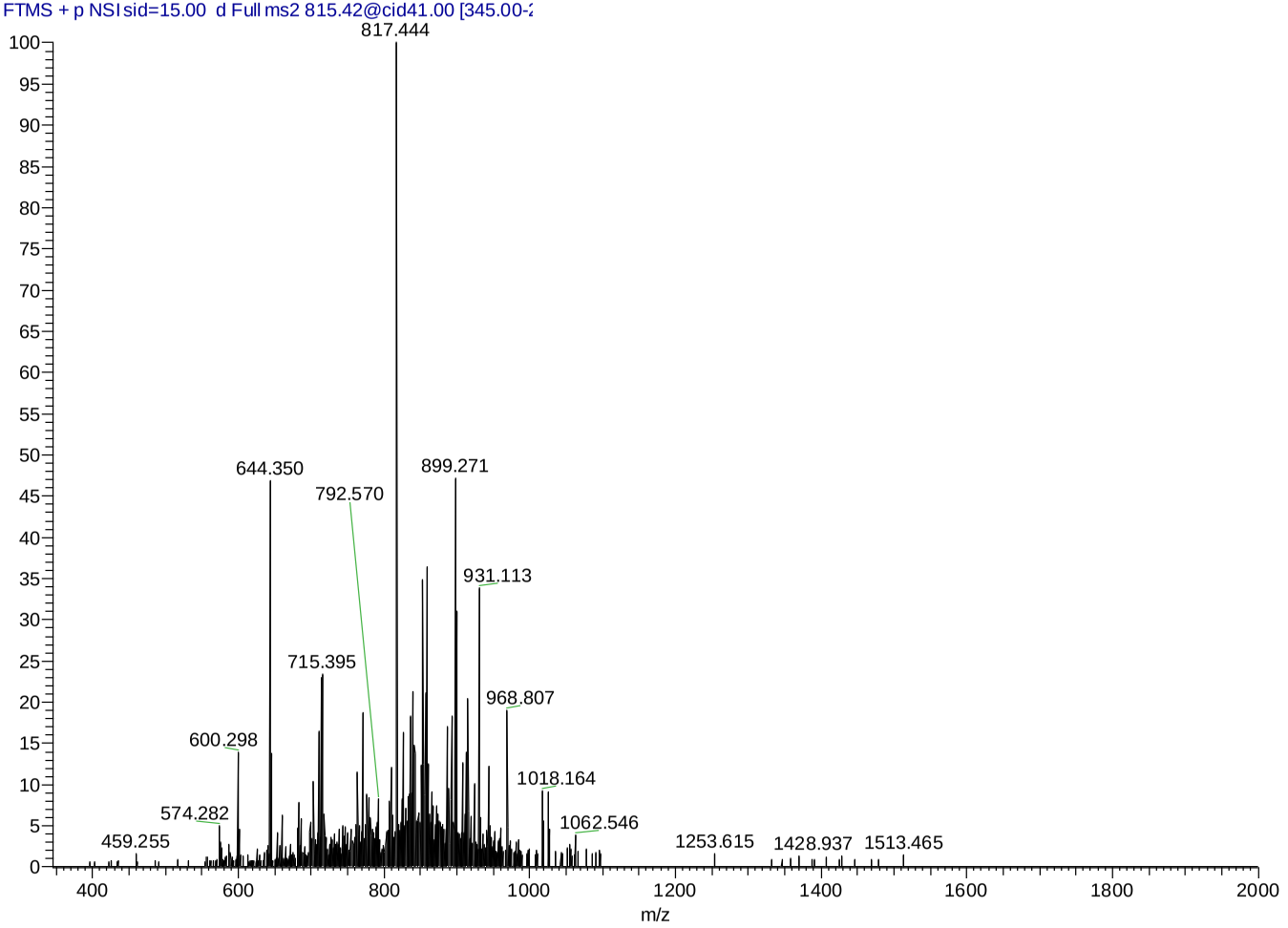
b1 - S - Q - A - D - F - D - K - A - A - E - E - V - K - R - L - K - T - Q - P - T - D - E - E - M - L -

b26 - F - I - Y - S - H - F - K - Q - A - T - V - G - D - V - N - T - D - R - P - G - L - L - D - L - K -

b51 - G - K - A - K - W - D - S - W - N - K - L - K - G - T - S - K - E - N - A - M - K - T - Y - V - E -

b76 - K - V - E - E - L - K - K - K - Y

20110915_ratbrain_Crude #3682 RT: 38.89 AV: 1 NL: 1.98E5
T: FTMS + p NSI sid=15.00 d Full ms2 815.42@cid41.00 [345.00-



DLRB1_RAT	P62628	Dynein light chain roadblock-type 1	10893.73	10893.73	0	N-term acetylation	2.0E-53

b1 - A - E - V - E - E - T - L - K - R - L - Q - S - Q - K - G - V - Q - G - I - I - V - V - N - T - E - y71

b26 - G - I - P - I - K - S - T - M - D - N - P - T - T - T - Q - Y - A - N - L - M - H - N - F - I - L - y46

b51 - K - A - R - S - T - V - R - E - I - D - P - Q - N - D - L - T - F - L - R - I - R - S - K - K - N - y21

b76 - E - I - M - V - A - P - D - K - D - Y - F - L - I - V - I - Q - N - P - T - E - y1

20110915_ratbrain_Crude #4935 RT: 51.23 AV: 1 NL: 1.37E4
T: FTMS + p NSI sid=15.00 d Full ms2 1363.47@cid41.00 [590.00