

© 2021 David Mast

METHODS FOR THE STEREOCHEMICAL ANALYSES OF ENDOGENOUS CELL-CELL
SIGNALING PEPTIDES IN ANIMALS

BY

DAVID HAMILTON MAST

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois Urbana-Champaign, 2021

Urbana, Illinois

Doctoral Committee:

Professor Jonathan V. Sweedler, Chair
Associate Professor Mary L. Kraft
Professor Deborah E. Leckband
Assistant Professor Peter M. Yau

ABSTRACT

D-amino acid containing peptides (DAACPs) represent an unusual class of signaling molecules in animals. Characterized by the presence of a non-proteinogenic D-amino acid residue in the backbone of an otherwise all-L residue peptide, DAACPs have fundamentally challenged the current understanding of animal protein chirality. DAACPs are considered uncommon because relatively few have been discovered in just a handful of species. However, it is possible many animal DAACPs have gone unidentified because modern peptide characterization studies do not routinely evaluate peptide stereochemistry. With increasing availability of genome/transcriptome-derived protein databases, and the widespread adaptation of liquid (LC)-chromatography-tandem mass spectrometry (MS/MS) sequencing, it is possible for hundreds to thousands of endogenous peptides to be identified in a single experiment. The reporting of endogenous peptide sequences without any stereochemical characterization has become widely accepted. The justification for this practice is premised on the overwhelming empirical evidence for a highly stereoselective ribosomal protein synthesis pathway in animals.

D-amino acid residues can be incorporated into animal cell-cell peptides enzymatically, by a post-translational modification (PTM) known as L- to D-residue isomerization. Like most PTMs, predicting which substrates will be modified by L- to D-residue isomerization based solely on genomic information is challenging and therefore, the PTM must be measured directly. However, unlike most other PTMs, L- to D-residue isomerization is a zero Dalton mass change PTM that is not accompanied by the addition or subtraction of a function group, or a change in amino acid sequence. Consequently, peptides that undergo L- to D-residue isomerization can easily go

undetected during LC-MS/MS sequencing because MS/MS on its own cannot be used for the *ab initio* determination of amino acid residue chirality.

This dissertation asserts that the prevalence of post-translational L- to D-residue isomerization in animals has been underestimated and that many cell-cell signaling DAACPs in animals may remain undiscovered or have been overlooked by previous studies. An analytical framework was applied to search for cell-cell signaling DAACPs in *Aplysia californica* and *Rattus norvegicus* tissues.

This dissertation describes the development and application of non-targeted methods to discover DAACPs among singling peptides in animals. These methods were used to discover eight enzyme derived cell-cell signaling DAACPs in the *A. californica* central nervous system. These DAACPs were encoded by protein precursors that had previously not been identified as encoding isomerized peptides. The tissue specific, and single cell localization of DAACPs was also investigated in *A. californica* to reveal important insights into the biosynthesis of DAACPs. In addition, a preliminary search for DAACPs, and other isomeric peptides, was performed on *Rattus norvegicus* neuro-endocrine tissues. The findings from screening *R. norvegicus* tissues for DAACPs were compared to the findings from *A. californica*, and the implications they have on future DAACP discovery efforts in animals are discussed.

This work is dedicated to my family, friends, and Plutonium.

ACKNOWLEDGEMENTS

I could not have completed my Ph.D. without the guidance from several individuals including, Jonathan Sweedler, James Checco, Elena Romanova, Stanislav Rubakhin, Thanh Do, Kevin Clark, Hua-Chia Tai, Peter Yau, Brian Imai, Deborah Leckband, Mary Kraft. I couldn't have asked for a more understanding and supportive advisor than Jonathan Sweedler, and I am so fortunate to have studied in his group.

I also want to acknowledge previous undergraduate professors Nancy Levinger, Joseph DiVerdi, Pat McCurdy, and Chuck Henry from Colorado State University for encouraging me to pursue a graduate degree in chemistry, as well as their continued support throughout graduate school. They not only provided me a strong educational background in Chemistry, but also offered mentorship and guidance on my professional development throughout graduate school.

The time I spent away from lab was just as important for my success as the time I spent working in lab. Therefore, I would like to thank specific individuals who support outside of lab. I reserve a special thanks to members of the top-secret pinochle club: Tony, Caroline, and Josiah.

Finally, I want to acknowledge Steven Glade, who was truly an inspirational colleague and the best friend a person could have. Had we never met, I am not sure I would be writing this now. May you rest in peace.

TABLE OF CONTENTS

CHAPTER 1: DISSERTATION OVERVIEW	1
CHAPTER 2: ADVANCING D-AMINO ACID CONTAINING PEPTIDE DISCOVERY IN THE METAZOAN	3
CHAPTER 3: DIFFERENTIAL POST-TRANSLATIONAL AMINO ACID ISOMERIZATION FOUND AMONG NEUROPEPTIDES IN APLYSIA CALIFORNICA	47
CHAPTER 4: ANALYSIS OF PEPTIDE STEREOCHEMISTRY IN SINGLE CELLS BY CAPILLARY ELECTROPHORESIS-TRAPPED ION MOBILITY SPECTROMETRY-MASS SPECTROMETRY	119
CHAPTER 5: ENZYMATIC DIGESTIONS ON INDIVIDUAL APLYSIA CALIFORNICA NEURONS	170
CHAPTER 6: SCREENING FOR ISOMERIC PEPTIDE HORMONES IN RODENT PANCREATIC ISLETS OF LANGERHANS	195
CHAPTER 7: SCREENING FOR ISOMERIC NEUROPEPTIDES IN RODENT CENTRAL AND PERIPHERAL NERVOUS SYSTEM TISSUES	218
CHAPTER 8: CONCLUSION AND FUTURE DIRECTIONS	256
APPENDIX A: CHIRAL AMINO ACID ANALYSIS OF PEGVADIN A FROM STREPTOMYCES NOURSEI	264
APPENDIX B: SUPPLEMENTAL INFORMATION FOR CHAPTER 3 AND CHAPTER 4	271

CHAPTER 1

DISSERTATION OVERVIEW

Chapter 2 begins this dissertation with a comprehensive review of the known D-amino acid containing peptides (DAACPs) in animals and the evolution of the analytical methods that have been used to characterize peptide stereochemistry. Chapter 2 also provides important definitions and background information about the biosynthesis and functions of DAACPs.

In Chapter 3, the prevalence of DAACPs was investigated in the common neuroscience model organism, *Aplysia californica*. Specifically, using liquid chromatography (LC)-tandem mass spectrometry (MS/MS), peptide extracts of different regions of the *A. californica* central nervous system (CNS) were screened for novel DAACPs based on enzymatic stability and chromatographic, or mass spectrometric properties. Three peptides encoded by the *A. californica* pleurin precursor were identified as DAACPs. In addition, two of the pleurin-derived peptides were found to be isomerized into more than one diastereomer product, representing the first examples of differential neuropeptide isomerization. Finally, Chapter 3 also describes two previously uncharacterized prohormone precursors that encode isomerized peptides: the FMRFamide precursor, and the *Aplysia* isomerized peptides precursor (AIPP). In total, between pleurin, FMRFamide, and AIPP, eight new DAACPs were discovered.

The DAACPs that were discovered in Chapter 3 were investigated further at the single cell level in Chapter 4. Capillary electrophoresis (CE) trapped ion mobility spectrometry (TIMS) MS was employed to determine the stereochemistry of three neuropeptides in individual neurons. The measurement of DAACPs at the single cell level revealed an unusual distribution of DAACPs in

the *A. californica* CNS, which provided key insights into the kinetics of L- to D-residue isomerization of neuropeptides in the secretory pathway.

In Chapter 5, an assay was developed and applied to detect DAACPs single cells by matrix assisted laser desorption-ionization (MALDI) MS. The purpose of developing of this assay was to test *A. californica* neurons for cell-type specific isomerization. The assay was tested on individual neurons that express prohormones known to encode isomerized peptides.

Finally, Chapters 6 and 7 describe the preliminary results from non-targeted DAACP screening of *Rattus norvegicus* neuro-endocrine tissues. In Chapter 6, pancreatic Islets of Langerhans were tested for cell-cell signaling DAACPs. In Chapter 7, spinal cord was tested for DAACPs. Additionally, Chapter 7 explores the chromatographic and TIMS properties of endogenous neuropeptides in *R. norvegicus* hypothalamus and pituitary to identify potentially isomeric peptides. Screening for DAACPs in *R. norvegicus* tissues revealed several peptides of interest that may have isomeric forms, but these peptides could not be conclusively identified as DAACPs due to limitations of the screening procedure.

Appendix A describes the stereochemical characterization of a peptide isolated from *Streptomyces noursei*. Appendix B describes supplemental experiments related to the work in Chapter 3 and Chapter 4 pertaining to the *A. californica* pleurin-derived peptides. A series of synthetic pleurin-derived peptide diastereomers that might be present endogenously were tested by LC-MS/MS. The relative expression of *pleurin* mRNA was measured in various ganglia of the CNS. Finally, an experiment was performed the test whether vesicle transport inhibitors cause DAACPs to accumulate in the cell bodies of neurons.

CHAPTER 2

ADVANCING D-AMINO ACID CONTAINING PEPTIDE DISCOVERY IN THE METAZOAN

2.1 NOTES AND ACKNOWLEDGEMENTS

The text in this chapter is published in *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*: Mast, D. H.; Checco, J. W.; Sweedler, J. V., Advancing D-amino acid-containing peptide discovery in the metazoan. *Biochim Biophys Acta Proteins Proteom* **2021**, 1869 (1), 140553. DOI: <https://doi.org/10.1016/j.bbapap.2020.140553>

Support for this work was provided by Award Number P30 DA018310 from the National Institute on Drug Abuse (NIDA) and Award No. RO1 NS031609 from the National Institute of Neurological Disorders and Stroke (NINDS).

2.2 ABSTRACT

The discovery of enzyme-derived D-amino acid-containing peptides (DAACPs) that have physiological importance in the metazoan challenges previous assumptions about the homochirality of animal proteins while simultaneously revealing new analytical challenges in the structural and functional characterization of peptides. Most known DAACPs have been identified through laborious activity-guided purification studies or by homology to previously identified DAACPs. Peptide characterization experiments are increasingly dominated by high throughput mass spectrometry-based peptidomics, with stereochemistry rarely considered due to the technical challenges of identifying L/D isomerization. This chapter discusses the prevalence of enzyme-

derived DAACPs among animals and the physiological consequences of peptide isomerization. Also highlighted are the analytical methods that have been applied for structural characterization/discovery of DAACPs, including results of several recent studies using non-targeted discovery methods for revealing novel DAACPs, strongly suggesting that more DAACPs remain to be uncovered.

2.3 INTRODUCTION

Prior to the 1980s, it was well-accepted that free D-amino acids played important roles in animal biology, but the presence of D-amino acid residues in animal-derived peptides/proteins had not been reported.¹ Aside from age-related, spontaneous isomerization of amino acid residues in peptides and proteins, a long-standing scientific consensus was that animal proteins/peptides were comprised of L-amino acids. However, in 1981 a D-Ala residue was detected in dermorphin, a gene-derived opioid-like hepta-peptide isolated from the skin secretions of two frog species, *Phyllomedusa rhodeti*² and *Phyllomedusa sauvagei*.³ In the years following the discovery of dermorphin, physiologically active D-amino acid-containing peptides (DAACPs) were uncovered in a variety of species, including spiders, crustaceans, mollusks, and even in the platypus (Figure 2.1). Unlike previously identified DAACPs in animals, these DAACPs were found to be the product of an enzyme-catalyzed post-translational modification (PTM), and not a result of spontaneous, age-related isomerization.

Despite the increasing prevalence of animal DAACPs, modern peptide characterization workflows seldom measure stereochemistry. Even in cases when experimental results suggest isomerization, these are sometimes mis-interpreted as proline *cis-trans* isomerization⁴⁻⁵ or disulfide bond pairing isomers.⁶⁻⁷ Furthermore, it is a difficult analytical task to distinguish peptide

isomers that result from spontaneous isomerization/epimerization of Asp and Ser residues from enzyme-derived DAACPs, which have a different set of biological consequences than age-related DAACPs (for relevant publications on spontaneous peptide isomerization see references 8-11).⁸⁻

¹¹ This chapter focuses on presumed enzyme-derived DAACPs in the metazoan (although the specific isomerases are often yet to be described). Because greater numbers of enzyme-derived DAACPs with biological functions are being discovered, it becomes increasingly important for stereochemistry to be considered when identifying new peptides and proteins. For this reason, the routine identification and discovery of novel enzyme-derived DAACPs in animals requires understanding the biosynthesis, and the biological consequences, of isomerization in functional peptides.

2.4 DAACPs THROUGHOUT THE METAZOAN

DAACPs have been discovered in animals across three phyla: chordates, arthropods, and mollusks. Figure 2.1 depicts the phylogeny—including the kingdom, phylum, class, and order—of the organisms known to have DAACPs. The DAACPs discovered across phyla differ substantially in peptide sequence, length, position and identity of the D-residue, and their physiological properties/functions. Most of these DAACPs were found to be secretory peptides, encoded by protein precursors targeted to the secretory pathway.

The first indication that DAACPs are the product of an enzyme-catalyzed PTM came from the discovery that the standard codon for L-Ala was found to be present in the cDNA sequence of dermorphin,¹²⁻¹³ suggesting that the D-residue was not incorporated by the ribosome. Subsequently, the discoveries of peptide isomerase activity in the extracts of spider venom,¹⁴⁻¹⁶ frog skin secretions,¹⁷⁻¹⁹ and platypus venom,^{7, 20-22} provided more evidence that DAACPs in

animals are produced by post-translational isomerization, with the enzymes yet to be characterized. Isomerase activity has also been detected in extracts of male echidna venom²³ and even mouse heart,²⁴ providing significant evidence for the presence of L/D isomerases and DAACPs in several species.

Initial studies on the biosynthesis of DAACPs in frogs suggested that isomerization occurs at the early stages of post-translational processing, before the peptides are cleaved from the precursor.²⁵ However, subsequent studies on cellular and subcellular localization of DAACPs in frogs,²⁶ crustaceans,²⁷⁻²⁹ and mollusks,³⁰⁻³¹ have shaped the current understanding of isomerization as a late stage, enzymatic PTM that occurs in secretory granules, after the peptide has been proteolytically cleaved from its precursor by prohormone convertases.^{26, 29, 32} Importantly, it has been suggested that L/D isomerization of cone snail toxin peptides precedes proteolytic cleavage of the precursor.³³ Exactly where L/D isomerization occurs in the secretory pathway has not been the subject of rigorous study, and it could vary depending on the species and enzymes involved. Whether isomerization occurs in the secretory pathway before or after cleavage by prohormone convertases may indeed vary in different species.

To date, two protein sequences for L/D isomerases have been reported: one L/D isomerase from the venom of the funnel web spider³⁴ and a second from frog skin secretions.¹⁷ Although these two proteins have little sequence similarity and have different substrate specificities,^{15-16, 18} both appear to catalyze the L/D isomerization of amino acid residues in peptides by similar mechanisms.^{16, 19} The reported isomerase sequence for the funnel web spider has been questioned and may be a potential misidentification.³⁵ Similarly, a modest amount of enzymatic activity from the frog isomerase could be recapitulated in a *Xenopus* oocyte system;¹⁷ to date, no L/D isomerase has been successfully purified from a recombinant system.

DAACPs can generally be classified into two categories: DAACP toxins and cell-cell signaling DAACPs. These categories are defined by the organism/tissue the peptide was isolated from, and the function/physiological activity of the peptide. The biosynthetic mechanisms of cell-cell signaling peptides and peptide toxins are similar and it is assumed that DAACPS under both categories are products of enzymatic post-translational isomerization.

DAACP Toxins

For the purposes of this dissertation, a DAACP toxin can be generally defined as a gene-derived DAACP produced by an organism that has poisonous effects on another organism. DAACPs under this category have been isolated from mammalian platypus venom, frog skin secretions, spider venom, and marine cone snail venom, and comprise a significant proportion of discovered DAACPs. Table 2.1 lists the DAACPs classified as toxins that have been confirmed experimentally.

Toxin DAACPs can cause potent interspecies biological effects. For example, the DAACPs isolated from frog skin secretions, dermorphin and deltorphin, showed significantly greater affinity and selectivity for mammalian δ -opioid and μ -opioid receptors than their all-L counterparts.³⁶⁻³⁹ The DAACP isolated from the funnel web spider, ω -agatoxin, is able to block Purkinje neuron P-type calcium ion channels and has modestly higher potency than the all-L-form (4- to 90-fold, depending on the assay).^{14, 40}

While the D-residue in many toxins appears to be required for bioactivity of the toxin, this is not always the case. The bombinin H family of DAACPs from frog skin secretions have potent antimicrobial and hemolytic activity,⁴¹⁻⁴³ but the biological effects of the all-L and D-amino acid-containing bombinin H family of peptides do not differ significantly. Many toxins remain

functionally and structurally uncharacterized, and many DAACPs have yet to be identified in venoms/toxins produced by several species. As researchers continue to characterize the active components of venoms, it will be exciting to investigate if DAACPs toxins are found to be present in other species.

Cell-Cell Signaling DAACPs.

For the purposes of this thesis, cell-cell signaling DAACPs can be defined as peptides that have been either shown or are expected to be involved in physiological processes within the organism that produces them. Cell-cell signaling DAACPs have been identified in arthropods and mollusks, but to our knowledge, have not been identified in chordates. A list of cell-cell signaling DAACPs that have been confirmed in mollusks and arthropods are reported in Table 2.2.

For DAACPs involved in cell-cell signaling that have been functionally characterized, the D-residue is often critical for the peptides' biological activity. For example, DAACPs found in gastropod mollusks, including achatin and achatin-related peptides (G**F**AD, G**F**FD, and G**Y**FD; D-amino acid residues indicated in bold and underlined),⁴⁴⁻⁴⁸ fulicin (F**N**EFV-NH₂),⁴⁹⁻⁵⁰ and fulyal (Y**A**EFL-NH₂),⁵¹ show potent activity from physiological measurements of neurons, muscle tissue or behavior, whereas their all-L-residue counterparts show little or no activity in the same experiments. Similarly, the D-residue in the cardioactive DAACP Ocp-1 (G**F**GD) isolated from octopus was determined to be critical for bioactivity. However, the opposite pattern was observed with another DAACP isolated from octopus, Octp-4 (G**S**WD), wherein bioactivity was present only for the all-L form of the peptide.⁵²

Presently, receptors have been identified for only two families of endogenous cell-cell signaling DAACPs: the achatin family (achatin-1, ApALNP-1, ApALNP-2 and homologues)⁵³⁻⁵⁴

and *Aplysia* allatotropin-related polypeptides (ApATRP).⁵⁵ In the case of achatin family peptides, the DAACP form is a potent agonist of the receptor, while the all-L-residue analogues are devoid of biological activity.⁵³⁻⁵⁴ These findings demonstrate that the D-residue is critical for receptor activation in the achatin family, which is consistent with previous results from *in vivo* and *ex vivo* assays showing that the D-residue is required for biological activity.⁴⁴⁻⁴⁷ These functional studies suggest that in many cases the post-translational incorporation of a D-residue is a key step in generating the biologically active form of the peptide used for cellular communication.

In contrast to the results described above, there are also several cases of DAACPs that display similar biological activity to their all-L-residue counterparts. For example, both forms of *Aplysia* ApATRP are potent agonists of the identified receptor, and the all-L-residue form is somewhat more potent at activating this receptor than the DAACP form (Figure 2.2A).⁵⁵ Additionally, both the all-L and D-amino acid-containing forms of ApATRP showed similar effects in neuronal electrophysiology experiments,⁵⁵ and similar potency in myoactivity assays of both diastereomers of the bivalve mollusk FMRFa-related decapeptide were also observed.⁵⁶ As a final example, the DAACP and all-L-residue forms of the crustacean hyperglycemic hormone (CHH) demonstrated similar effects on hyperglycemia (albeit with different kinetics)⁵⁷ and on ecdysteroid synthesis.⁵⁸

Although physiological measurements of biological activity using *in vivo* or *ex vivo* approaches are powerful tools for elucidating peptide functions, they are somewhat limited in their ability to determine the full consequences of L/D isomerization. Physiological experiments output a response that is significantly downstream of the initial signaling event and dependent on many factors. In addition, the choice of bioactivity experiment for many newly discovered peptides is often somewhat arbitrary and may not reflect a primary biological function in the animal. In

contrast, studying a peptide's direct activation of cognate receptors can provide important information regarding their signaling capabilities that decouples receptor potency from other factors.

Alternative Roles for L/D Isomerization

Although the several DAACP toxins and cell-cell signaling DAACPS mentioned above do not differ substantially in bioactivity from their all-L counterparts, the presence of a D-residue can have other biological consequences. The presence of D-residue(s) within a sequence can slow native and exogenous protease action relative to an all-L-residue peptide, presumably because the substrate-altered backbone does not allow it to be positioned correctly into the protease active site⁵⁹⁻⁶⁰. For these reasons, DAACPs may have greater stability against proteolytic degradation and an increased biological lifetime. For some DAACP toxins that have been identified, studies have shown a higher protease stability for the DAACP relative to the all-L-residue peptide in the venom environment,^{14, 17, 37} suggesting the D-residue provides stability while the toxin is stored and waiting to be released. For cell-cell signaling DAACPs, increased resistance to protease degradation may allow these peptides to engage in hormone-like signaling, activating receptors at distances far from the point of release. Consistent with this hypothesis, in experiments measuring the degradation of naturally occurring *Aplysia* DAACPs in hemolymph plasma, the DAACP forms were found to be >70-fold more stable to degradation than their all-L-residue analogues.⁵⁵ These results are consistent with the fact that hemolymph is rich in aminopeptidases⁶¹ and for some aminopeptidases, activity is hindered by a single D-residue close to the N-terminus of a peptide sequence.⁶²⁻⁶³ The increased stability of DAACPs may indicate another possible role for peptide L/D isomerization: to modify peptide stability and thus, duration of action. It will be important for

future studies to carefully examine the effects of L/D isomerization on the lifetime of peptides in their natural context (e.g., upon release at the synapse for neuropeptides, or after exposure to prey/predators for toxins) to gain full insight into the consequences on stability.

Structural Diversity of DAACPs

Despite the large variation in DAACP structure, stability, and function in animals, one common feature that is shared by all known DAACPs among chordates, arthropods, and mollusks is the presence of one D-amino acid in the peptide backbone. In other words, the animal DAACPs discovered to date have all been epimers of the all-L-residue peptide (Tables 2.1 and 2.2). This important characteristic makes animal DAACPs largely distinct from those found in bacteria, which often have more than one D-residue.⁶⁴ Very recently, DAACPs were identified from the whole bodies of the annelid *Whitmania pigra* that contained two D-residues.⁶⁵ However, it is unclear if these DAACPs were gene-derived peptides and produced via the biosynthetic machinery of this animal, or through other mechanisms (e.g., spontaneous isomerization or bacterial derived).

The position and composition of the isomerized residue in the peptide backbone reflects the isomerase substrate specificity. Interestingly, the isomerized residue is oftentimes located near the amino or carboxy termini of the peptides, and the D-residues found across all animals tend to be hydrophobic or aromatic amino acid residues. Importantly, because peptide stereochemistry is not routinely evaluated in modern peptide characterization workflows, the structural diversity of DAACPs may be much greater. Although the functions of many DAACPs have not been fully explored, the structural features common to known DAACPs (e.g., position and composition of the D-residue(s)) can aid in designing approaches to facilitate future discovery efforts for novel DAACPs by targeting unique structural features common to many substrates of L/D isomerization.

2.5 HISTORY OF DAACP DISCOVERY

Many of the first animal DAACPs were discovered between the 1980s and early 2000s, before the widespread usage/availability of modern mass spectrometric technologies. At that time, peptide sequencing required large quantities of peptide to be purified, sometimes from hundreds of animals. Oftentimes, the bioactivity of the native peptide of interest was known before its structure, and this activity guided the purification.^{3, 33, 36-37, 44, 49, 52, 56, 66-67} Following purification of a native peptide, the amino acid composition and sequence were determined by techniques such as acid hydrolysis/amino acid analysis or Edman degradation. Finally, the structure of the native peptide was confirmed by comparing the bioactivity, chromatographic retention, and other properties of the purified native peptide to a synthetic peptide standard. Following this workflow, many of the first animal DAACPs were discovered after the biological activities,^{40, 44, 49, 52} chromatographic^{5, 7, 40, 44, 49, 56, 67-71} or spectroscopic properties^{5, 7, 44, 72} of the purified native peptide(s) were not reproduced by the synthetic all-L peptide standards.^{5, 7, 40, 44, 49, 52, 56, 67-72} The presence of a D-residue was only revealed upon further investigation into the native peptide structure using various chiral amino acid analysis techniques.⁷³⁻⁷⁶

Figure 2.3 is a schematic representation of the peptide characterization workflow that led to the discovery of many DAACPs, a process that has continued to evolve since the discovery of dermorphin. With the increasing capabilities and availability of mass spectrometry (MS) instrumentation, large-scale peptide purifications and chemical amino acid sequencing techniques have generally been replaced by high throughput liquid chromatography (LC)-tandem MS (MS/MS) sequencing. LC-MS/MS data can be searched against genomic- and transcriptomic-derived protein databases, enabling accurate identification of hundreds to thousands of sequences in a complex mixture in a single experiment. However, because MS cannot easily differentiate

peptide stereoisomers, stereochemistry is not routinely evaluated in modern peptidomics studies. As a result, DAACPs may be overlooked simply because the practice of confirming native peptide structures by comparing their properties to synthetic standards has become increasingly limited. Consequently, the abundance of DAACPs among animal peptides may be severely underestimated.

It is important to note that the major challenge facing modern DAACP discovery is not the characterization of stereochemistry. The real challenge is predicting which few peptides, among the thousands that are detected in LC-MS/MS experiments, likely contain a D-residue. As soon as a DAACP is suspected, identifying its stereochemistry is straightforward and can be accomplished with chiral amino acid analysis techniques, or by comparing physical and chemical properties to synthetic peptide standards. Various modern methods used to characterize the structure of DAACPs, including amino acid analysis, LC-MS/MS, ion mobility (IM)-MS, MS/MS, capillary electrophoresis, etc., have been discussed in recent reviews.⁷⁷⁻⁷⁹ Despite the number and quality of analytical tools now available for differentiating peptide stereoisomers, they are being underutilized by modern peptidomics workflows for the identification of native peptide stereoisomers.

Although native peptide diastereomers do not differ by mass, they can be differentiated based on other physical or chemical properties such as gas phase structure, enzymatic stability, electrophoretic mobility, chromatographic retention, and even MS/MS fragmentation. Putative DAACPs may also be identified based on sequence similarities to known DAACPs. Modernization of DAACP discovery requires analytical methods capable of predicting which peptides in a complex mixture are DAACPs before the biological activity is evaluated. As discussed further in

the “Modern DAACP discovery” section, the goal is to develop standardized workflows to identify novel DAACPs.

2.6 METHODS FOR IDENTIFYING DAACPS

Predicting DAACPs Based on Homology and Peptide Sequence

In some cases, peptides that undergo post-translational L/D isomerization can be predicted based on the L/D isomerase specificity (if known),¹⁸ the sequences of known DAACPs from the organism, or by sequence similarity to known DAACPs in other organisms. For example, if a DAACP is discovered in one organism, one might predict that similar peptide sequences from the same organism, or in a related species, may also exist as DAACPs. This homology-guided discovery can help narrow the list of potential DAACPs that will be subsequently characterized.

Until recently, most modern DAACP discovery efforts relied heavily on homology to find DAACPs in different organisms. For example, a D-residue has been detected at position 2 from the N-terminus in dermorphin, deltorphin, and related peptides from a variety of frog species.^{2, 12-13, 25, 66, 80-82} Because this peptide family has been confirmed to undergo post-translational L/D isomerization in so many instances across species, it is largely assumed that a D-residue is present in all of the dermorphin/deltorphin-related peptides in frog skin secretions.⁸⁰⁻⁸¹ Many cone snail DAACPs in the contryphan conotoxin family were also discovered based on homology.^{33, 83-88} In crustaceans, the DAACP forms of CHH and vitellogenesis inhibiting hormone (VIH) were originally discovered in the American lobster, *Homarus americanus*,^{29, 57} but the D-residue has been shown to be conserved in VIH and CHH peptides from a variety of other lobster and crayfish species.^{27, 58, 79, 89-93} The achatin-like peptides from *A. californica*, GFFD and GYFD, were both initially targeted and discovered because these peptides had similar sequences to the known

DAACP, GFAD from *A. fulica*, and were encoded by homologous prohormones.⁴⁶⁻⁴⁷ Similarly, the *A. californica* DAACP, allatotropin-related polypeptide ([D-Phe²]-ApATRP), was targeted because of its sequence similarity to the cone snail DAACP toxin, conomap-Vt.⁵⁵

Once a DAACP has been discovered in one organism, there is a high likelihood that other peptides in the organism (or related species) also have DAACPs. For instance, Mignogna et al.⁴¹ tested the bombinin H peptides in *Bombina variegata* for D-residues simply because “peptides from the skin of frogs belonging to the sub-family *Phyllomedusinae* contain a D-amino acid in position 2,” despite having virtually no sequence similarity with dermorphin-related peptides. If one peptide encoded by a prohormone is identified as a DAACP, it is likely (but not guaranteed) that similar peptides encoded by the same prohormone also exist with a D-amino acid-containing form. The finding that multiple peptides encoded by a common precursor undergo L/D isomerization was first observed with peptides from the *A. fulica* and *P. sauvagei*, and later observed in other organisms.^{25, 47, 51, 94} Recently, three peptide sequences encoded by the *A. californica* pleurin precursor, three peptides from the ApALNP precursor, and two peptides from the *Aplysia* isomerized peptides precursor (AIPP) were discovered to be DAACPs.^{46-47, 94} However, just because multiple similar peptides are encoded on the same precursor does not guarantee every peptide will have a D-amino acid-containing form.^{51, 94} The latter observation may be due to the L/D isomerase substrate specificity, or native substrate concentration.⁹⁴

Importantly, prohormone homology and sequence similarity are not always predictors of DAACPs between organisms. For example, although fulyal peptide (YAEFL-NH₂) was found to be a DAACP in *A. fulica*, the identical sequence in *A. californica* was only detected in the all-L-residue form.⁴⁶ As another example, the sequence of the DAACP from platypus venom, ovCNPb, closely resembles the sequence for the mammalian C-type natriuretic peptide (CNP).⁴⁻⁵ Although

the possibility that mammalian CNPs may exist in multiple stereoisomers has not been evaluated, it should not be assumed that they have a D-residue simply because of this homology. The conservation of D-residues in homologous peptides across species is likely dependent on the conservation of the L/D isomerase enzyme.

Although DAACPs are common among certain peptide families across species, homology-guided discovery alone significantly limits the ability to discover novel DAACPs. To uncover novel and more structurally diverse DAACPs in the metazoan, peptide characterization efforts would benefit from developing methods to swiftly characterize native peptide stereochemistry instead of relying on previous findings. In the future, cellular localization of peptides may be an important factor for determining whether a peptide may exist as a DAACP. As more DAACPs are discovered, and single cell characterization methods advance, it may be possible to predict novel DAACPs by targeting peptides that colocalize with known DAACPs, or with the L/D isomerase (if known) in single cells.

Analytical Methods for Identifying DAACPs

DAACPs oftentimes coexist in biological extracts with the all-L-residue form of the peptide.^{5, 30, 41, 44, 52, 57, 94-95} Therefore, substrates of L/D isomerization may be identified without the need for synthetic standards by attentively searching for evidence of multiple diastereomers of peptide sequences using separations techniques or other specialized approaches. Native peptide diastereomers are often observed in chromatographic separations as multiple peaks of the same sequence eluting at different retention times, a key signature indicating that a peptide could have a D-amino acid-containing form.^{7, 79, 94, 96} The DAACPs from platypus and lobsters/crayfish were

initially discovered because identical peptide sequences were found to elute at more than one retention time.^{4-5, 29, 57}

Six peptide sequences detected with split peak elution profiles by LC-MS/MS were recently discovered to have D-amino acid-containing forms (Figure 2.4).⁹⁴ Modern LC-MS/MS peptide characterization studies could implement routine identification of DAACP candidates by analyzing chromatograms for peptides that elute at multiple retention times, helping identify peptides to target for downstream structural characterization. Some peptidomics studies have already begun to report peptides with split eluting peaks, even when no follow-up chiral analysis was performed.⁹⁷ In cases where standard conditions fail to separate DAACPs (i.e., C18 stationary phase, water/acetonitrile mobile phase), a separation of peptide stereoisomers can be accomplished by modifying gradient conditions, solvent system, stationary phase, or chromatographic technique.^{49, 55, 94, 98} Future DAACP discovery workflows may include LC-MS/MS analysis using several separation conditions to identify peptide stereoisomers that coelute under one chromatographic condition but separate under different conditions. However, DAACPs could easily be overlooked if their discovery relies solely on chromatographic separation approaches. Co-elution of native peptide stereoisomers may be common. Therefore, alternative methods that can be easily worked into LC-MS based workflows must also be considered to assist in DAACP identification.

IM-MS can achieve rapid gas phase separation of peptide stereoisomers,^{77, 99-107} and is a potentially valuable approach for identifying native peptide stereoisomers. IM-MS methods have been widely applied for the structural characterization and separation and of known naturally occurring DAACPs and other peptide stereoisomers.^{99-104, 106-107} Recently, a workflow utilizing IM-MS in conjunction with MS/MS was developed to perform site-specific localization of D-

amino acid residues in a known DAACP.¹⁰⁵ The latter was a milestone for DAACP discovery techniques not only because native peptide diastereomers could be differentiated by IM-MS, but more importantly, the position of the D-residue in the peptide could be localized by comparing the mobilities of product ions between the two diastereomers. However, IM-MS methods have seldom been used in the direct discovery of native DAACPs.

Recently, trapped ion mobility spectrometry (TIMS) combined with LC-MS/MS was used to confirm and differentiate native stereoisomers of Plrn2 ([D-Phe²]Plrn2 and [D-Tyr³]Plrn2) in different regions of the *A. californica* central nervous system and validate stereochemical assignments.⁹⁴ At present, IM-MS is underutilized in native DAACP discovery workflows, but exciting new discoveries will certainly be made as IM-MS technology improves, resolving power increases, and the approach becomes more common. Future DAACP discovery efforts may integrate IM-MS into LC-MS/MS-based peptidomics workflows to perform gas phase separations of native peptide stereoisomers that may coelute by LC.

MS/MS can be a powerful tool for differentiating peptide stereoisomers by comparing the relative abundances of the fragment ions intensities produced by this measurement approach.¹⁰⁸⁻¹¹¹ DAACPs can be identified, confirmed or quantified by comparing the fragment ion spectra of native peptide stereoisomers to each other, or comparing the fragment ion ratios of native peptides and synthetic standards.^{31, 46, 72, 94, 112} The low detection limits of MS/MS-based methods is a benefit for confirming or identifying possible DAACPs. However, a drawback is that structural confirmation requires a synthetic standard for comparison, which is not ideal for non-targeted discovery efforts when hundreds of potential DAACPs are possible. MS/MS has been extremely useful in recent years for structural characterization of peptide isomers resulting from spontaneous/age-related peptide isomerization/epimerization. Looking ahead peptide

characterization protocols may increase confidence in peptide stereoisomer assignments using LC-MS/MS by comparing the product ion ratios of peptide sequences eluting at multiple retention times.

Another successful approach for identifying candidate DAACPs is by analysis of the native peptide's enzymatic stability. In combination with chromatographic separations and MS detection, enzymatic digestions can simultaneously aid in the identification of peptide stereoisomers and in localization of the position of the D-residue within a peptide sequence. The presence of D-amino acid residues in peptides can make them poorer substrates for proteolytic enzymes¹¹³. Therefore, DAACPs can be differentiated from their all-L counterparts by the rate of proteolysis, or by the degradation products of proteolytic enzymes.^{3-5, 57-58, 68-69, 71}

The exopeptidase, aminopeptidase M (APM), has been widely used for identifying native DAACPs. APM sequentially cleaves L-amino acid residues from the N-termini of peptides, but peptides with a D-residue present near the N-terminus tend to resist degradation by APM.⁶²⁻⁶³ Because APM has a well-defined stereoselectivity and substrate cleavage specificity, it has historically been used to help confirm the position of the D-residue in position 2 in naturally occurring DAACPs.^{3, 5, 41, 47, 55, 65, 94, 114} This was done simply by testing the resistance of the native peptide to APM digestion compared to the synthetic all-L peptide and synthetic DAACP standards. The DAACP is expected to be resistant to APM, while the all-L peptide is expected to degrade. Figure 2.4 shows examples of native DAACPS that were identified and differentiated from other native stereoisomers. While APM is commonly used to identify native DAACPs, future discovery efforts could benefit from using other proteolytic enzymes with different substrate specificities to aid in identifying native DAACPs.

Modern DAACP Discovery Methods

Modern DAACP discovery approaches seek to identify potential DAACPs in animals using the methods discussed above. Workflows have only just begun to be used to evaluate the chromatographic/mass spectrometric properties and enzymatic stability of peptides in biological extracts to identify possible DAACP candidates for follow up analysis (Figure 2.5). These approaches take into consideration peptide homology, enzyme stability, and MS properties to look for possible DAACPs in peptide extracts.

Recently, an LC-MS/MS based, non-targeted workflow termed “the DAACP discovery funnel” was developed that used APM to identify potential peptides with a D-residue in position 2 based on stability against APM digestion.⁴⁷ Peptides that resist degradation by APM—and thus may have a D-residue near the N-terminus—are purified and the stereochemistry evaluated using chiral amino acid analysis techniques. If a D-residue is detected in the peptide, the native peptide’s stereochemistry is confirmed by comparing the chromatographic retention of the native peptide with synthetic standards. More recently, the discovery funnel was improved upon by adapting the APM screening step to large scale LC-MS/MS peptidomics.⁹⁴ Using this optimized discovery funnel workflow, the resistance to APM of hundreds of peptides in peptide extracts was evaluated using LC-MS/MS and comparing multiple chromatographic and mass spectrometric properties of peptides before and after APM digestion. This led to the discovery of eight novel DAACPs, including three from previously uncharacterized prohormones in the *A. californica* CNS.⁹⁴ DAACP candidates were assigned if they met a strict set of criteria, e.g., peptide sequences needed to be resistant to APM degradation and be observed at multiple retention times before APM treatment. Because the position of the D-residue could be localized to the N-terminus based on the stereospecificity of APM, the purification of the peptides and the chiral amino acid analysis step

could be skipped and the predicted stereochemistry of the native peptides confirmed by comparing their chromatographic retention, mobilities (via TIMS), and MS/MS fragmentation patterns to synthetic peptide standards.⁹⁴ This method also showed that by considering multiple properties (i.e., mass spectrometric and chromatographic properties, and enzyme stability) it is possible to rapidly discover structurally diverse DAACPs in a peptidomics experiment.

2.7 CONCLUSION

Among the enzyme-derived DAACPs that have been identified, L/D isomerization has been shown to play a functional role in toxins, hormones, and neuropeptides. However, despite this functional importance, stereochemistry is rarely considered in modern peptide characterization experiments due to the technical challenges of differentiating L- and D-residues using LC-MS-based methods. Because of these challenges, animal DAACPs may be missed and thus, more diverse than has been currently reported. Importantly, aside from the DAACP toxins identified in the platypus, DAACPs have not been identified in mammals, although isomerase activity has been reported in the mouse heart. As more non-targeted methods are developed to detect DAACPs, it will be exciting to see if this modification is found in mammalian cell-cell signaling peptides. Characterization of the L/D isomerases responsible for producing native DAACPs will provide important insight into their biosynthesis, and the conservation of DAACPs throughout the metazoan could help identify new DAACPs. Further, while nearly all known DAACPs in the metazoan have been encoded by proteins in the secretory pathway, the enzymatic incorporation of D-residues into non-secretory, gene-derived proteins/peptides has not been studied.

It is interesting to speculate that proteins larger than those identified thus far might be isomerized, although this would be extremely challenging to detect. Modern peptide

characterization studies could benefit from integrating new approaches to evaluate stereochemistry, including methods outlined in this review. Future developments in non-targeted DAACP identification will ultimately help reveal the true prevalence and biological consequences of L/D isomerization among animals, and may even uncover new structurally diverse DAACPs in the metazoan.

2.8 FIGURES AND DATA

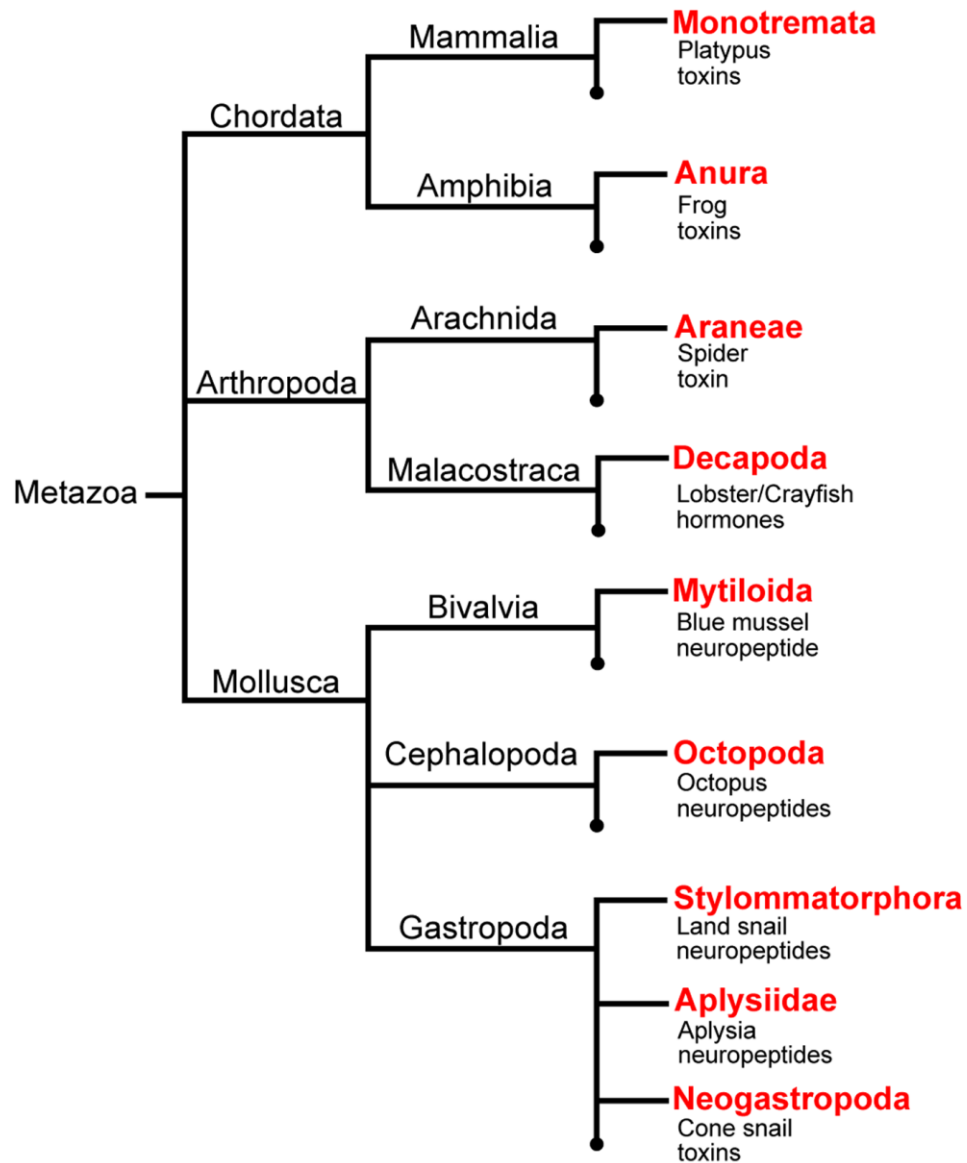


Figure 2.1. Phylogeny of organisms in which enzyme-derived DAACPs have been discovered, highlighting kingdom, phylum, class, and order. The roles of the DAACPs are presented beneath the order.

Table 2.1. DAACP venoms/toxins discovered in the metazoan. References correspond to the publication in which the DAACP was first confirmed experimentally. D-amino acid residues are indicated in bold and underlined (**X**).

Peptide Name	Sequence	Species	Ref
Leu-Contryphan-P	---GCVLLPWC	<i>Conus purpurascens</i>	88
Leu-contryphan-Tx	----CVLYPWC	<i>Conus textile</i>	83
Contryphan-R	---GCOWEPWC-NH ₂	<i>Conus radiatus</i>	67
Bromocontryphan	----COWEPwC-NH ₂	<i>Conus radiatus</i>	33
Contryphan-Sm	---GCOWQPWC-NH ₂	<i>Conus stercusmuscarum</i>	85
Contryphan-P	---GCOWDPWC-NH ₂	<i>Conus purpurascens</i>	85
Contryphan P/Am	---GCOWDPWC-NH ₂	<i>Conus amadis</i>	87
Contryphan-Lo	---GCPWDPWC-NH ₂	<i>Conus loroisii</i>	87
Contryphan-Vn	--GDCPWKPWC-NH ₂	<i>Conus ventricosus</i>	84
Contryphan-R/Tx	---GCOWEPWC-NH ₂	<i>Conus textile</i>	83
Contryphan-Tx	----COWQPYC-NH ₂	<i>Conus textile</i>	83
Glacontryphan-M	NySyCPWHpWC-NH ₂	<i>Conus marmoreus</i>	86
Gld-v*	-----AOANSvWS	<i>Conus gladiator</i>	72
Gld-v	-----AOANSVWS	<i>Conus gladiator</i>	72
Mus-V*	-----SOANSvWS	<i>Conus gladiator</i>	72
Mus-V	-----SOANSVWS	<i>Conus gladiator</i>	72
Conomorphin	DWEYHAHPKONSEWT	<i>Conus marmoreus</i>	68
R11a	----...GCSTSSFFKI	<i>Conus radiatus</i>	71
R11b	----...GCSTSSFFRI	<i>Conus radiatus</i>	115
R11c	----...GCSTNVFLT	<i>Conus radiatus</i>	115
Conomap vt	AFVKGSAQRVAHGYNH ₂	<i>Conus vitulinus</i>	69
Dermorphin related	Y I FHLMD	<i>Pachymedusa dactylos</i>	80
Dermorphin	Y A FGYPS-NH ₂	<i>Phyllomedusa rhodei</i>	2
[Hyp ⁶]-dermorphin	Y A FGYOS-NH ₂	<i>Phyllomedusa rhodei</i>	2
Dermorphin	Y A FGYPS-NH ₂	<i>Phyllomedusa sauvagei</i>	3,13
Met-Deltorphin	Y M FHLMD-NH ₂	<i>Phyllomedusa sauvagei</i>	36,116
Dermorphin-like1	Y A FWYPN	<i>Phyllomedusa bicolor</i>	12
Dermorphin-like2	Y A FGYPK	<i>Phyllomedusa bicolor</i>	12
pentapeptide	Y A FWN	<i>Phyllomedusa bicolor</i>	12
Deltorphin1	Y A FDVVG-NH ₂	<i>Phyllomedusa bicolor</i>	37,12
Deltorphin2	Y A FEVVG-NH ₂	<i>Phyllomedusa bicolor</i>	37,12
Leu-deltorphin	Y L FADVASTIGDFFHSI-NH ₂	<i>Phyllomedusa burmeisteri</i>	66
Bombinin H3	I I GPVLGMVGSALGGLKKI-NH ₂	<i>Bombina variegata</i>	41
Bombinin H4	L I GPVLGLVGSALGGLKKI-NH ₂	<i>Bombina variegata</i>	41
Bombinin H5	I I GPVLGLVGSALGGLKKI-NH ₂	<i>Bombina variegata</i>	41
GH-2	I L GPVLDLVGRALRGLKKI-NH ₂	<i>Bombina orientalis</i>	112
Phenylseptin	F F FDTLKNLAGKVIGALT-NH ₂	<i>Hypsiboas punctatus</i>	107
ovCNPb	L L HDHP...	<i>Ornithorhynchus anatinus</i>	5
DLP	I M FEMQ...	<i>Ornithorhynchus anatinus</i>	7
ω-agatoxin	...LIMEGL S FA	<i>Agelenopsis aperta</i>	40

w = 6-bromotyrosine; **y** = gamma D-hydroxy valine; O = hydroxyproline; **γ** = γ-carboxylated glutamate; **I** = D-allo-isoleucine. Note: this list is curated to include only DAACPs that have been confirmed experimentally. Some DAACPs may have been excluded unintentionally.

Table 2.2. Cell-cell signaling DAACPs discovered in the metazoan. References correspond to the publication in which the DAACP was first confirmed experimentally in the reported species. D-amino acid residues indicated in bold and underlined (**X**).

Peptide Name	Sequence	Species	Ref
Achatin-1	G F AD	<i>Achatina fulica</i>	44
Fulyl	F N EFV-NH ₂	<i>Achatina fulica</i>	49
Fulicin	Y A EFL-NH ₂	<i>Achatina fulica</i>	51
Ocp-1	G F GD	<i>Octopus minor</i>	52
Ocp-4	G S WD	<i>Octopus minor</i>	52
FMRF-related decapeptide	A L AGDHFFRF-NH ₂	<i>Mytilus edulis</i>	56
NdWFa	N W F-NH ₂	<i>Aplysia kurodai</i>	70
NdWFa	N W F-NH ₂	<i>Aplysia californica</i>	30, 114
ApALNP-1	G F FD	<i>Aplysia californica</i>	46
ApALNP-2	G Y FD	<i>Aplysia californica</i>	47
ApALNP-3	S Y ADSKDEESNAALSDFA (ED)	<i>Aplysia californica</i>	47
[D-Phe ²]ApATRP	G F RLNSASRVAHGY-NH ₂	<i>Aplysia californica</i>	55
[D-Phe ²]Plm1	M E YTKGSDSDYPRI-NH ₂	<i>Aplysia californica</i>	94
[D-Phe ²]Plm2	S F YTTGNGNHYPRI-NH ₂	<i>Aplysia californica</i>	94
[D-allo ²]Plm3	G I FTQSAYGSYPRI-NH ₂	<i>Aplysia californica</i>	94
[D-Leu ²]Ip1	Y L DHLGSSLV	<i>Aplysia californica</i>	94
[D-Leu ²]Ip2	Y L DGIASSLI	<i>Aplysia californica</i>	94
[D-Met ²]FMRGFa	F M RGF-NH ₂	<i>Aplysia californica</i>	94
[D-Tyr ³]Plm2	S F YTTGNGNHYPRI-NH ₂	<i>Aplysia californica</i>	94
[D-Phe ³]Plm3	G I FTQSAYGSYPRI-NH ₂	<i>Aplysia californica</i>	94
[D-Phe ³]CHH	q V F D QACK...	<i>Homarus americanus</i>	57
[D-Trp ⁴]VIH	ASA W F T ND...	<i>Homarus americanus</i>	29

q = pyro glutamic acid; CHH = crustacean hyperglycemic hormone, VIH = vitellogenesis inhibiting hormone; ApALNP = Aplysia achatin-like neuropeptide. Note: this list is curated by the authors to include only DAACPs that have been confirmed experimentally. Some DAACPs may have been excluded unintentionally.

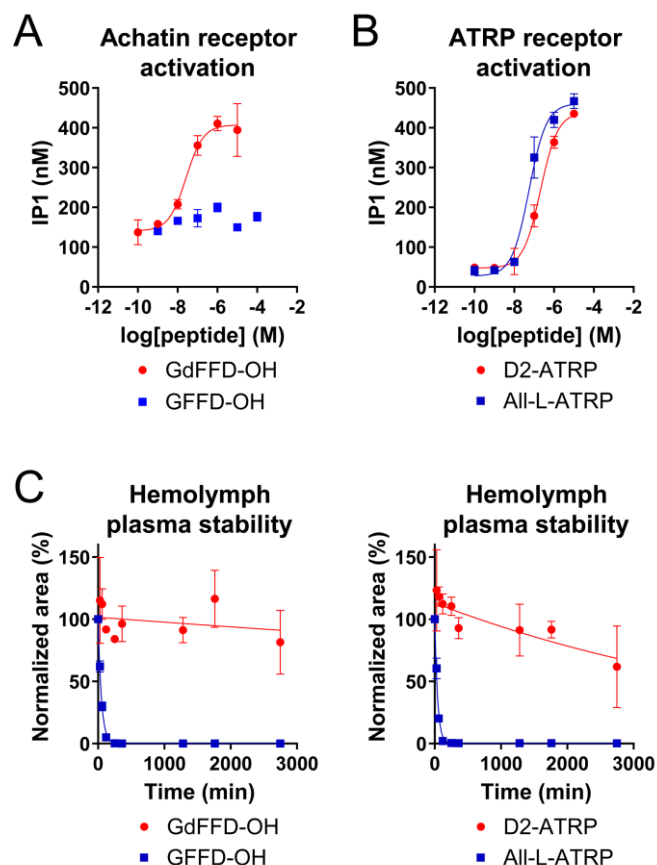


Figure 2.2. Biological consequences of L- to D-residue isomerization among *Aplysia* neuropeptides. (A, B) Representative dose-response curves showing the ability of peptides to activate either the (A) achatin receptor or (B) ApATRP receptor, when transiently transfected in CHO-K1 cells. Points represent mean $\pm$ standard deviation for duplicate wells in an experiment. (C) Time-course of peptide degradation in *Aplysia* hemolymph plasma, as assessed by LC-multiple reaction monitoring. Points represent mean $\pm$ standard deviation for two biological replicates, and curves show fit of the data to an exponential decay model. Panel (A) adapted from J.W. Checco, G. Zhang, W.D. Yuan, K. Yu, S.Y. Yin, R.H. Roberts-Galbraith, P.M. Yau, E.V. Romanova, J. Jing, J.V. Sweedler, Molecular and physiological characterization of a receptor for D-amino acid-containing neuropeptides, *ACS Chem. Biol.*, 13 (2018) 1343-1352. Copyright 2018, American Chemical Society⁵⁴. Panels (B) and (C) originally published in the *Journal of Biological Chemistry*, and adapted from J.W. Checco, G. Zhang, W.D. Yuan, Z.W. Le, J. Jing, J.V. Sweedler, *Aplysia* allatotropin-related peptide and its newly identified d-amino acid-containing epimer both activate a receptor and a neuronal target, *J. Biol. Chem.*, 293 (2018) 16862-16873, © the Authors.⁵⁵

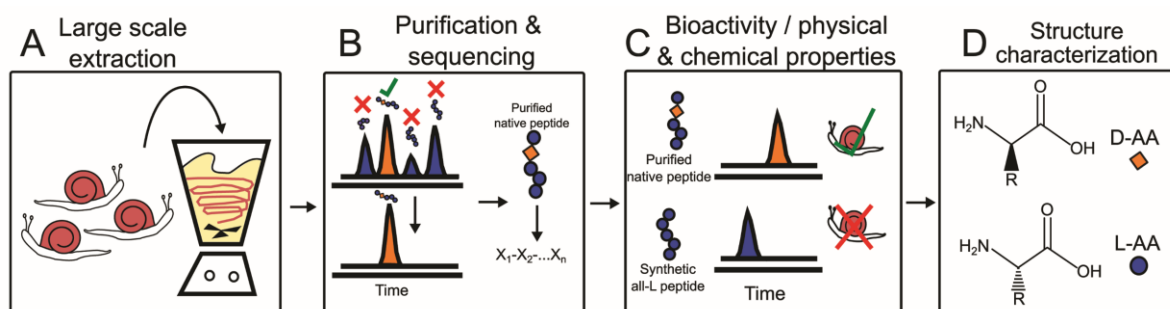


Figure 2.3. Peptide characterization workflow historically used that resulted in the discovery of DAACPs in animals. (A) Large-scale peptide extraction and (B) purification and amino acid sequencing was performed to determine amino acid sequence and composition without consideration of stereochemistry. (C) Once the sequence was known, physical/chemical properties or biological activities of the native peptide and synthetic peptide sequences (all-L residue) were compared. (D) If the properties of the native peptide were not reproduced by the synthetic peptide, further structural characterization of the native peptide was performed using chiral amino acid analysis techniques or other structural characterization methods, and the native peptide was discovered to contain a D-residue.

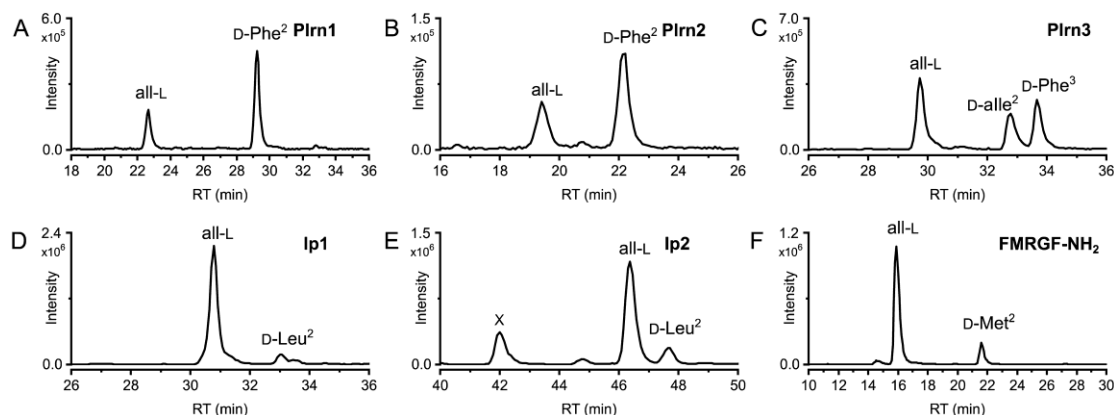


Figure 2.4. Identification of peptide stereoisomers in *A. californica* CNS extracts by LC-MS/MS based on multiple peak elution. Each plot shows LC-MS/MS extracted ion chromatograms corresponding to native peptide stereoisomers of (A) Plrn1, (B) Plrn2, (C) Plrn3 (D) Ip1, (E) Ip2, (F) FMRGF-NH₂ in a peptide extract of cerebral ganglia. Within each plot, each peak is a peptide stereoisomer, and the peaks that were identified as DAACPs are labeled with the D-amino acid residue that was detected in the peptides. Peaks labeled all-L correspond to the all-L-residue form of the peptide. The full sequences for the peptides can be found in Table 2.2. The peak labeled with an “X” in (E) was not determined to be a stereoisomer of Ip2. Adapted from D.H. Mast, J.W. Checco, J.V. Sweedler, Differential post-translational amino acid isomerization found among neuropeptides in *Aplysia californica*, ACS Chem. Biol., 15 (2020) 272-281. Copyright 2020, American Chemical Society.⁹⁴

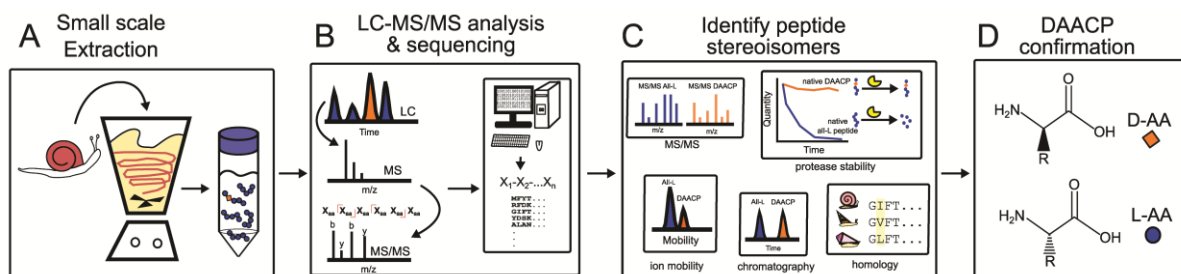


Figure 2.5. Proposed general workflow for modern DAACP discovery. (A) Peptide extraction followed by (B) LC-MS/MS analysis and sequencing by database search to identify native peptides. (C) Identification of putative isomeric peptides detected in LC-MS/MS analysis by analyzing ion mobility migration, MS/MS, protease stability, chromatographic retention, or homology. (D) Finally, after identifying peptide stereoisomers, the DAACPs are confirmed by comparing the properties of native peptides to using synthetic standards, or by chiral AA analysis.

2.9 REFERENCES

1. Corrigan, J. J., D-Amino Acids in Animals. *Science* **1969**, *164* (3876), 142.
2. Montecucchi, P. C.; de Castiglione, R.; Erspamer, V., Identification of dermorphin and Hyp6-dermorphin in skin extracts of the Brazilian frog *Phyllomedusa rhodei*. *Int. J. Pept. Protein Res.* **1981**, *17* (3), 316-21.
3. Montecucchi, P. C.; de Castiglione, R.; Piani, S.; Gozzini, L.; Erspamer, V., Amino acid composition and sequence of dermorphin, a novel opiate-like peptide from the skin of *Phyllomedusa sauvagei*. *Int. J. Pept. Protein Res.* **1981**, *17* (3), 275-83.
4. de Plater, G. M.; Martin, R. L.; Milburn, P. J., A C-type natriuretic peptide from the venom of the platypus (*Ornithorhynchus anatinus*): Structure and pharmacology. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **1998**, *120* (1), 99-110.
5. Torres, A. M.; Menz, I.; Alewood, P. F.; Bansal, P.; Lahnstein, J.; Gallagher, C. H.; Kuchel, P. W., D-Amino acid residue in the C-type natriuretic peptide from the venom of the mammal, *Ornithorhynchus anatinus*, the Australian platypus. *FEBS Lett.* **2002**, *524* (1-3), 172-6.
6. Torres, A. M.; de Plater, G. M.; Doverskog, M.; Birinyi-Strachan, L. C.; Nicholson, G. M.; Gallagher, C. H.; Kuchel, P. W., Defensin-like peptide-2 from platypus venom: member of a class of peptides with a distinct structural fold. *Biochem. J.* **2000**, *348 Pt 3* (Pt 3), 649-56.
7. Torres, A. M.; Tsampazi, C.; Geraghty, D. P.; Bansal, P. S.; Alewood, P. F.; Kuchel, P. W., D-amino acid residue in a defensin-like peptide from platypus venom: effect on structure and chromatographic properties. *Biochem. J.* **2005**, *391* (Pt 2), 215-20.

8. Lambeth, T. R.; Riggs, D. L.; Talbert, L. E.; Tang, J.; Coburn, E.; Kang, A. S.; Noll, J.; Augello, C.; Ford, B. D.; Julian, R. R., Spontaneous Isomerization of Long-Lived Proteins Provides a Molecular Mechanism for the Lysosomal Failure Observed in Alzheimer's Disease. *ACS Cent. Sci.* **2019**, 5 (8), 1387-1395.
9. Kumar, S.; Prakash, S.; Gupta, K.; Dongre, A.; Balaram, P.; Balaram, H., Unexpected functional implication of a stable succinimide in the structural stability of *Methanocaldococcus jannaschii* glutaminase. *Nat. Commun.* **2016**, 7, 12798.
10. Stephenson, R. C.; Clarke, S., Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J. Biol. Chem.* **1989**, 264 (11), 6164-70.
11. Geiger, T.; Clarke, S., Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* **1987**, 262 (2), 785-94.
12. Richter, K.; Egger, R.; Negri, L.; Corsi, R.; Severini, C.; Kreil, G., cDNAs encoding [D-Ala²]deltorphin precursors from skin of *Phyllomedusa bicolor* also contain genetic information for three dermorphin-related opioid peptides. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, 87 (12), 4836-9.
13. Richter, K.; Egger, R.; Kreil, G., D-alanine in the frog skin peptide dermorphin is derived from L-alanine in the precursor. *Science* **1987**, 238 (4824), 200-2.
14. Heck, S. D.; Siok, C. J.; Krapcho, K. J.; Kelbaugh, P. R.; Thadeio, P. F.; Welch, M. J.; Williams, R. D.; Ganong, A. H.; Kelly, M. E.; Lanzetti, A. J.; et al., Functional consequences of posttranslational isomerization of Ser46 in a calcium channel toxin. *Science* **1994**, 266 (5187), 1065-8.

15. Murkin, A. S.; Tanner, M. E., Dehydroalanine-based inhibition of a peptide epimerase from spider venom. *J. Org. Chem.* **2002**, *67* (24), 8389-94.
16. Heck, S. D.; Faraci, W. S.; Kelbaugh, P. R.; Saccomano, N. A.; Thadeio, P. F.; Volkmann, R. A., Posttranslational amino acid epimerization: enzyme-catalyzed isomerization of amino acid residues in peptide chains. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (9), 4036-9.
17. Jilek, A.; Mollay, C.; Tippelt, C.; Grassi, J.; Mignogna, G.; Müllegger, J.; Sander, V.; Fehrer, C.; Barra, D.; Kreil, G., Biosynthesis of a D-amino acid in peptide linkage by an enzyme from frog skin secretions. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (12), 4235-9.
18. Jilek, A.; Mollay, C.; Lohner, K.; Kreil, G., Substrate specificity of a peptidyl-aminoacyl-L/D-isomerase from frog skin. *Amino Acids* **2012**, *42* (5), 1757-64.
19. Gehmayr, V.; Mollay, C.; Reith, L.; Müller, N.; Jilek, A., Tight binding of transition-state analogues to a peptidyl-aminoacyl-L/D-isomerase from frog skin. *ChemBioChem* **2011**, *12* (13), 1996-2000.
20. Torres, A. M.; Tsampazi, M.; Tsampazi, C.; Kennett, E. C.; Belov, K.; Geraghty, D. P.; Bansal, P. S.; Alewood, P. F.; Kuchel, P. W., Mammalian L-to-D-amino-acid-residue isomerase from platypus venom. *FEBS Lett.* **2006**, *580* (6), 1587-91.
21. Torres, A. M.; Tsampazi, M.; Kennett, E. C.; Belov, K.; Geraghty, D. P.; Bansal, P. S.; Alewood, P. F.; Kuchel, P. W., Characterization and isolation of L-to-D-amino-acid-residue isomerase from platypus venom. *Amino Acids* **2007**, *32* (1), 63-8.
22. Bansal, P. S.; Torres, A. M.; Crossett, B.; Wong, K. K.; Koh, J. M.; Geraghty, D. P.; Vandenberg, J. I.; Kuchel, P. W., Substrate specificity of platypus venom L-to-D-peptide isomerase. *J. Biol. Chem.* **2008**, *283* (14), 8969-75.

23. Koh, J. M. S.; Haynes, L.; Belov, K.; Kuchel, P. W., L-to-D peptide isomerase in male echidna venom. *Aust. J. Zool.* **2011**, 58 (5), 284-288.
24. Koh, J. M.; Chow, S. J.; Crossett, B.; Kuchel, P. W., Mammalian peptide isomerase: platypus-type activity is present in mouse heart. *Chem. Biodivers.* **2010**, 7 (6), 1603-11.
25. Mor, A.; Delfour, A.; Nicolas, P., Identification of a D-alanine-containing polypeptide precursor for the peptide opioid, dermorphin. *J. Biol. Chem.* **1991**, 266 (10), 6264-70.
26. Auvynet, C.; Seddiki, N.; Dunia, I.; Nicolas, P.; Amiche, M.; Lacombe, C., Post-translational amino acid racemization in the frog skin peptide deltorphin I in the secretion granules of cutaneous serous glands. *Eur J Cell Biol* **2006**, 85 (1), 25-34.
27. Soye, D.; Laverdure, A. M.; Kallen, J.; Van Herp, F., Demonstration of a cell-specific isomerization of invertebrate neuropeptides. *Neuroscience* **1998**, 82 (3), 935-42.
28. Gallois, D.; Brisorgueil, M. J.; Conrath, M.; Mailly, P.; Soye, D., Posttranslational isomerization of a neuropeptide in crustacean neurosecretory cells studied by ultrastructural immunocytochemistry. *Eur J Cell Biol* **2003**, 82 (8), 431-40.
29. Ollivaux, C.; Gallois, D.; Amiche, M.; Boscameric, M.; Soye, D., Molecular and cellular specificity of post-translational aminoacyl isomerization in the crustacean hyperglycaemic hormone family. *FEBS J.* **2009**, 276 (17), 4790-802.
30. Sheeley, S. A.; Miao, H.; Ewing, M. A.; Rubakhin, S. S.; Sweedler, J. V., Measuring D-amino acid-containing neuropeptides with capillary electrophoresis. *Analyst* **2005**, 130 (8), 1198-203.
31. Bai, L.; Romanova, E. V.; Sweedler, J. V., Distinguishing endogenous D-amino acid-containing neuropeptides in individual neurons using tandem mass spectrometry. *Anal. Chem.* **2011**, 83 (7), 2794-800.

32. Soye, D.; Toullec, J. Y.; Ollivaux, C.; Geraud, G., L-to-D amino acid isomerization in a peptide hormone is a late post-translational event occurring in specialized neurosecretory cells. *J. Biol. Chem.* **2000**, *275* (48), 37870-5.
33. Jimenez, E. C.; Craig, A. G.; Watkins, M.; Hillyard, D. R.; Gray, W. R.; Gulyas, J.; Rivier, J. E.; Cruz, L. J.; Olivera, B. M., Bromocontryphan: post-translational bromination of tryptophan. *Biochemistry* **1997**, *36* (5), 989-94.
34. Shikata, Y.; Watanabe, T.; Teramoto, T.; Inoue, A.; Kawakami, Y.; Nishizawa, Y.; Katayama, K.; Kuwada, M., Isolation and characterization of a peptide isomerase from funnel web spider venom. *J. Biol. Chem.* **1995**, *270* (28), 16719-23.
35. Langenegger, N.; Koua, D.; Schürch, S.; Heller, M.; Nentwig, W.; Kuhn-Nentwig, L., Identification of a precursor processing protease from the spider *Cupiennius salei* essential for venom neurotoxin maturation. *J. Biol. Chem.* **2018**, *293* (6), 2079-2090.
36. Kreil, G.; Barra, D.; Simmaco, M.; Erspamer, V.; Erspamer, G. F.; Negri, L.; Severini, C.; Corsi, R.; Melchiorri, P., Deltorphan, a novel amphibian skin peptide with high selectivity and affinity for delta opioid receptors. *Eur J Pharmacol* **1989**, *162* (1), 123-8.
37. Erspamer, V.; Melchiorri, P.; Falconieri-Erspamer, G.; Negri, L.; Corsi, R.; Severini, C.; Barra, D.; Simmaco, M.; Kreil, G., Deltorphins: a family of naturally occurring peptides with high affinity and selectivity for delta opioid binding sites. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (13), 5188-92.
38. Broccardo, M.; Erspamer, V.; Falconieri Erspamer, G.; Improta, G.; Linari, G.; Melchiorri, P.; Montecucchi, P. C., Pharmacological data on dermorphins, a new class of potent opioid peptides from amphibian skin. *Br. J. Pharmacol.* **1981**, *73* (3), 625-31.

39. Negri, L.; Erspamer, G. F.; Severini, C.; Potenza, R. L.; Melchiorri, P.; Erspamer, V., Dermorphin-related peptides from the skin of *Phyllomedusa bicolor* and their amidated analogs activate two mu opioid receptor subtypes that modulate antinociception and catalepsy in the rat. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (15), 7203-7.
40. Kuwada, M.; Teramoto, T.; Kumagaye, K. Y.; Nakajima, K.; Watanabe, T.; Kawai, T.; Kawakami, Y.; Niidome, T.; Sawada, K.; Nishizawa, Y.; et al., Omega-agatoxin-TK containing D-serine at position 46, but not synthetic omega-[L-Ser46]agatoxin-TK, exerts blockade of P-type calcium channels in cerebellar Purkinje neurons. *Mol. Pharmacol.* **1994**, *46* (4), 587-93.
41. Mignogna, G.; Simmaco, M.; Kreil, G.; Barra, D., Antibacterial and haemolytic peptides containing D-alloisoleucine from the skin of *Bombina variegata*. *EMBO J* **1993**, *12* (12), 4829-32.
42. Simmaco, M.; Kreil, G.; Barra, D., Bombinins, antimicrobial peptides from *Bombina* species. *Biochim. Biophys. Acta* **2009**, *1788* (8), 1551-5.
43. Mangoni, M. L.; Grovale, N.; Giorgi, A.; Mignogna, G.; Simmaco, M.; Barra, D., Structure-function relationships in bombinins H, antimicrobial peptides from *Bombina* skin secretions. *Peptides* **2000**, *21* (11), 1673-9.
44. Kamatani, Y.; Minakata, H.; Kenny, P. T.; Iwashita, T.; Watanabe, K.; Funase, K.; Sun, X. P.; Yongsiri, A.; Kim, K. H.; Novales-Li, P.; et al., Achatin-I, an endogenous neuroexcitatory tetrapeptide from *Achatina fulica* Ferussac containing a D-amino acid residue. *Biochem. Biophys. Res. Commun.* **1989**, *160* (3), 1015-20.
45. Fujimoto, K.; Kubota, I.; Yasuda-Kamatani, Y.; Minakata, H.; Nomoto, K.; Yoshida, M.; Harada, A.; Muneoka, Y.; Kobayashi, M., Purification of achatin-I from the atria of the

- African giant snail, *Achatina fulica*, and its possible function. *Biochem. Biophys. Res. Commun.* **1991**, *177* (2), 847-53.
46. Bai, L.; Livnat, I.; Romanova, E. V.; Alexeeva, V.; Yau, P. M.; Vilim, F. S.; Weiss, K. R.; Jing, J.; Sweedler, J. V., Characterization of GdFFD, a D-amino acid-containing neuropeptide that functions as an extrinsic modulator of the *Aplysia* feeding circuit. *J. Biol. Chem.* **2013**, *288* (46), 32837-51.
 47. Livnat, I.; Tai, H. C.; Jansson, E. T.; Bai, L.; Romanova, E. V.; Chen, T. T.; Yu, K.; Chen, S. A.; Zhang, Y.; Wang, Z. Y.; Liu, D. D.; Weiss, K. R.; Jing, J.; Sweedler, J. V., A d-Amino Acid-Containing Neuropeptide Discovery Funnel. *Anal. Chem.* **2016**, *88* (23), 11868-11876.
 48. Yang, C. Y.; Yu, K.; Wang, Y.; Chen, S. A.; Liu, D. D.; Wang, Z. Y.; Su, Y. N.; Yang, S. Z.; Chen, T. T.; Livnat, I.; Vilim, F. S.; Cropper, E. C.; Weiss, K. R.; Sweedler, J. V.; Jing, J., *Aplysia* Locomotion: Network and Behavioral Actions of GdFFD, a D-Amino Acid-Containing Neuropeptide. *PLoS One* **2016**, *11* (1), e0147335.
 49. Ohta, N.; Kubota, I.; Takao, T.; Shimonishi, Y.; Yasuda-Kamatani, Y.; Minakata, H.; Nomoto, K.; Muneoka, Y.; Kobayashi, M., Fulicin, a novel neuropeptide containing a D-amino acid residue isolated from the ganglia of *Achatina fulica*. *Biochem. Biophys. Res. Commun.* **1991**, *178* (2), 486-93.
 50. Fujita, K.; Minakata, H.; Nomoto, K.; Furukawa, Y.; Kobayashi, M., Structure-activity relations of fulicin, a peptide containing a D-amino acid residue. *Peptides* **1995**, *16* (4), 565-8.

51. Yasuda-Kamatani, Y.; Kobayashi, M.; Yasuda, A.; Fujita, T.; Minakata, H.; Nomoto, K.; Nakamura, M.; Sakiyama, F., A novel D-amino acid-containing peptide, fulyal, coexists with fulicin gene-related peptides in Achatina atria. *Peptides* **1997**, *18* (3), 347-54.
52. Iwakoshi, E.; Hisada, M.; Minakata, H., Cardioactive peptides isolated from the brain of a Japanese octopus, Octopus minor. *Peptides* **2000**, *21* (5), 623-30.
53. Bauknecht, P.; Jekely, G., Large-Scale Combinatorial Deorphanization of Platynereis Neuropeptide GPCRs. *Cell Rep.* **2015**, *12* (4), 684-93.
54. Checco, J. W.; Zhang, G.; Yuan, W. D.; Yu, K.; Yin, S. Y.; Roberts-Galbraith, R. H.; Yau, P. M.; Romanova, E. V.; Jing, J.; Sweedler, J. V., Molecular and Physiological Characterization of a Receptor for d-Amino Acid-Containing Neuropeptides. *ACS Chem. Biol.* **2018**, *13* (5), 1343-1352.
55. Checco, J. W.; Zhang, G.; Yuan, W. D.; Le, Z. W.; Jing, J.; Sweedler, J. V., Aplysia allatotropin-related peptide and its newly identified d-amino acid-containing epimer both activate a receptor and a neuronal target. *J. Biol. Chem.* **2018**, *293* (43), 16862-16873.
56. Fujisawa, Y.; Ikeda, T.; Nomoto, K.; Yasuda-Kamatani, Y.; Minakata, H.; Kenny, P. T.; Kubota, I.; Muneoka, Y., The FMRFamide-related decapeptide of Mytilus contains a D-amino acid residue. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **1992**, *102* (1), 91-5.
57. Soye, D.; Van Herp, F.; Rossier, J.; Le Caer, J. P.; Tensen, C. P.; Lafont, R., Evidence for a conformational polymorphism of invertebrate neurohormones. D-amino acid residue in crustacean hyperglycemic peptides. *J. Biol. Chem.* **1994**, *269* (28), 18295-8.

58. Yasuda, A.; Yasuda, Y.; Fujita, T.; Naya, Y., Characterization of crustacean hyperglycemic hormone from the crayfish (*Procambarus clarkii*): multiplicity of molecular forms by stereoinversion and diverse functions. *Gen. Comp. Endocrinol.* **1994**, *95* (3), 387-98.
59. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H., Comparison of the proteolytic susceptibilities of homologous L-amino acid, D-amino acid, and N-substituted glycine peptide and peptoid oligomers. *Drug Dev. Res.* **1995**, *35* (1), 20-32.
60. Milton, R. C.; Milton, S. C.; Kent, S. B., Total chemical synthesis of a D-enzyme: the enantiomers of HIV-1 protease show reciprocal chiral substrate specificity [corrected]. *Science* **1992**, *256* (5062), 1445-8.
61. Rothman, B. S.; Dekruyff, S.; Talebian, M.; Menon, J. G.; Squire, C. R.; Yeh, C. H.; Lee, T. D., Aplysia peptide neurotransmitters beta-bag cell peptide, Phe-Met-Arg-Phe-amide, and small cardioexcitatory peptide B are rapidly degraded by a leucine aminopeptidase-like activity in hemolymph. *J. Biol. Chem.* **1992**, *267* (35), 25135-40.
62. Wachsmuth, E. D.; Fritze, I.; Pfeleiderer, G., An aminopeptidase occurring in pig kidney. I. An improved method of preparation. Physical and enzymic properties. *Biochemistry* **1966**, *5* (1), 169-74.
63. Wachsmuth, E. D.; Fritze, I.; Pfeleiderer, G., An aminopeptidase occurring in pig kidney. II. A study on the mechanism of the hydrolysis. *Biochemistry* **1966**, *5* (1), 175-82.
64. Ogasawara, Y.; Dairi, T., Peptide Epimerization Machinerics Found in Microorganisms. *Front. Microbiol.* **2018**, *9*, 156.

65. Zhang, Z. Y.; Ma, N.; Tao, L. J.; Gong, X. Y.; Ye, W. C.; Wang, L., Linear Peptides Containing d-Leucine with Neuroprotective Activities from the Leech *Whitmania pigra* Whitman. *J. Nat. Prod.* **2019**, *82* (8), 2349-2353.
66. Barra, D.; Mignogna, G.; Simmaco, M.; Pucci, P.; Severini, C.; Falconieri-Erspamer, G.; Negri, L.; Erspamer, V., [D-Leu²]deltorphan, a 17 amino acid opioid peptide from the skin of the Brazilian hyliid frog, *Phyllomedusa burmeisteri*. *Peptides* **1994**, *15* (2), 199-202.
67. Jimenez, E. C.; Olivera, B. M.; Gray, W. R.; Cruz, L. J., Contryphan is a D-tryptophan-containing *Conus* peptide. *J. Biol. Chem.* **1996**, *271* (45), 28002-5.
68. Han, Y.; Huang, F.; Jiang, H.; Liu, L.; Wang, Q.; Wang, Y.; Shao, X.; Chi, C.; Du, W.; Wang, C., Purification and structural characterization of a D-amino acid-containing conopeptide, conomarphin, from *Conus marmoreus*. *FEBS J.* **2008**, *275* (9), 1976-87.
69. Dutertre, S.; Lumsden, N. G.; Alewood, P. F.; Lewis, R. J., Isolation and characterisation of conomap-Vt, a D-amino acid containing excitatory peptide from the venom of a vermivorous cone snail. *FEBS Lett.* **2006**, *580* (16), 3860-6.
70. Morishita, F.; Nakanishi, Y.; Kaku, S.; Furukawa, Y.; Ohta, S.; Hirata, T.; Ohtani, M.; Fujisawa, Y.; Muneoka, Y.; Matsushima, O., A novel D-amino-acid-containing peptide isolated from *Aplysia* heart. *Biochem. Biophys. Res. Commun.* **1997**, *240* (2), 354-8.
71. Buczek, O.; Yoshikami, D.; Bulaj, G.; Jimenez, E. C.; Olivera, B. M., Post-translational amino acid isomerization: a functionally important D-amino acid in an excitatory peptide. *J. Biol. Chem.* **2005**, *280* (6), 4247-53.
72. Pisarewicz, K.; Mora, D.; Pflueger, F. C.; Fields, G. B.; Mari, F., Polypeptide chains containing D-gamma-hydroxyvaline. *J Am Chem Soc* **2005**, *127* (17), 6207-15.

73. Marfey, P., Determination of D-amino acids. II. Use of a bifunctional reagent, 1,5-difluoro-2,4-dinitrobenzene. *Carlsberg Res. Commun.* **1984**, 49 (6), 591.
74. Einarsson, S.; Josefsson, B.; Moller, P.; Sanchez, D., Separation of amino acid enantiomers and chiral amines using precolumn derivatization with (+)-1-(9-fluorenyl)ethyl chloroformate and reversed-phase liquid chromatography. *Anal. Chem.* **1987**, 59 (8), 1191-5.
75. Scaloni, A.; Simmaco, M.; Bossa, F., Determination of the chirality of amino acid residues in the course of subtractive Edman degradation of peptides. *Anal. Biochem.* **1991**, 197 (2), 305-10.
76. Szokan, G.; Mezo, G.; Hudecz, F., Application of Marfey's reagent in racemization studies of amino acids and peptides. *J. Chromatogr.* **1988**, 444, 115-22.
77. Jansson, E. T., Strategies for analysis of isomeric peptides. *J. Sep. Sci.* **2018**, 41 (1), 385-397.
78. Bai, L.; Sheeley, S.; Sweedler, J. V., Analysis of Endogenous D-Amino Acid-Containing Peptides in Metazoa. *Bioanal. Rev.* **2009**, 1 (1), 7-24.
79. Soye, D.; Toullec, J. Y.; Montagne, N.; Ollivaux, C., Experimental strategies for the analysis of D-amino acid containing peptides in crustaceans: a review. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2011**, 879 (29), 3102-7.
80. Wechselberger, C.; Severini, C.; Kreil, G.; Negri, L., A new opioid peptide predicted from cloned cDNAs from skin of *Pachymedusa dactylos* and *Agalychnis saltator*. *FEBS Lett.* **1998**, 429 (1), 41-3.
81. Thompson, A. H.; Bjorson, A. J.; Orr, D. F.; Shaw, C.; McClean, S., Amphibian skin secretomics: application of parallel quadrupole time-of-flight mass spectrometry and

- peptide precursor cDNA cloning to rapidly characterize the skin secretory peptidome of *Phyllomedusa hypochondrialis azurea*: discovery of a novel peptide family, the hyposins. *J. Proteome Res.* **2007**, *6* (9), 3604-13.
82. Rates, B.; Silva, L. P.; Ireno, I. C.; Leite, F. S.; Borges, M. H.; Bloch, C., Jr.; De Lima, M. E.; Pimenta, A. M., Peptidomic dissection of the skin secretion of *Phasmahyla jandaia* (Bokermann and Sazima, 1978) (Anura, Hylidae, Phyllomedusinae). *Toxicon* **2011**, *57* (1), 35-52.
 83. Jimenez, E. C.; Watkins, M.; Juszczak, L. J.; Cruz, L. J.; Olivera, B. M., Contryphans from *Conus textile* venom ducts. *Toxicon* **2001**, *39* (6), 803-8.
 84. Massilia, G. R.; Schinina, M. E.; Ascenzi, P.; Polticelli, F., Contryphan-Vn: a novel peptide from the venom of the Mediterranean snail *Conus ventricosus*. *Biochem. Biophys. Res. Commun.* **2001**, *288* (4), 908-13.
 85. Jacobsen, R.; Jimenez, E. C.; Grilley, M.; Watkins, M.; Hillyard, D.; Cruz, L. J.; Olivera, B. M., The contryphans, a D-tryptophan-containing family of *Conus* peptides: interconversion between conformers. *J. Pept. Res.* **1998**, *51* (3), 173-9.
 86. Hansson, K.; Ma, X.; Eliasson, L.; Czerwiec, E.; Furie, B.; Furie, B. C.; Rorsman, P.; Stenflo, J., The first gamma-carboxyglutamic acid-containing contryphan. A selective L-type calcium ion channel blocker isolated from the venom of *Conus marmoreus*. *J. Biol. Chem.* **2004**, *279* (31), 32453-63.
 87. Gowd, K. H.; Sabareesh, V.; Sudarshani, S.; Iengar, P.; Franklin, B.; Fernando, A.; Dewan, K.; Ramaswami, M.; Sarma, S. P.; Sikdar, S.; Balaram, P.; Krishnan, K. S., Novel peptides of therapeutic promise from Indian Conidae. *Ann. N. Y. Acad. Sci.* **2005**, *1056*, 462-73.

88. Jacobsen, R. B.; Jimenez, E. C.; De la Cruz, R. G.; Gray, W. R.; Cruz, L. J.; Olivera, B. M., A novel D-leucine-containing Conus peptide: diverse conformational dynamics in the contryphan family. *J. Pept. Res.* **1999**, *54* (2), 93-9.
89. Aguilar, M. B.; Soye, D.; Falchetto, R.; Arnott, D.; Shabanowitz, J.; Hunt, D. F.; Huberman, A., Amino acid sequence of the minor isomorph of the crustacean hyperglycemic hormone (CHH-II) of the Mexican crayfish *Procambarus bouvieri* (Ortmann): presence of a D-amino acid. *Peptides* **1995**, *16* (8), 1375-83.
90. Bulau, P.; Meisen, I.; Reichwein-Roderburg, B.; Peter-Katalinic, J.; Keller, R., Two genetic variants of the crustacean hyperglycemic hormone (CHH) from the Australian crayfish, *Cherax destructor*: detection of chiral isoforms due to posttranslational modification. *Peptides* **2003**, *24* (12), 1871-9.
91. Serrano, L.; Blanvillain, G.; Soye, D.; Charmantier, G.; Grousset, E.; Aujoulat, F.; Spanings-Pierrot, C., Putative involvement of crustacean hyperglycemic hormone isoforms in the neuroendocrine mediation of osmoregulation in the crayfish *Astacus leptodactylus*. *J. Exp. Biol.* **2003**, *206* (Pt 6), 979-88.
92. Ollivaux, C.; Vinh, J.; Soye, D.; Toullec, J. Y., Crustacean hyperglycemic and vitellogenesis-inhibiting hormones in the lobster *Homarus gammarus*. *FEBS J.* **2006**, *273* (10), 2151-60.
93. Redeker, V.; Toullec, J. Y.; Vinh, J.; Rossier, J.; Soye, D., Combination of peptide profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and immunodetection on single glands or cells. *Anal. Chem.* **1998**, *70* (9), 1805-11.

94. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Differential post-translational amino acid isomerization found among neuropeptides in *Aplysia californica*. *ACS Chem. Biol.* **2020**, *15* (1), 272-281.
95. Song, Y.; Liu, Y. M., Quantitation of cardioexcitatory Asn-D-Trp-Phe-NH₂ diastereomers in *Aplysia*'s central nervous system by nanoscale liquid chromatography-tandem mass spectrometry. *J. Mass Spectrom.* **2008**, *43* (9), 1285-90.
96. Pinkse, M.; Evaristo, G.; Pieterse, M.; Yu, Y.; Verhaert, P., MS approaches to select peptides with post-translational modifications from amphibian defense secretions prior to full sequence elucidation. *EuPA Open Proteomics* **2014**, *5*, 32-40.
97. Vijayasathy, M.; Basheer, S. M.; Franklin, J. B.; Balaram, P., Contryphan Genes and Mature Peptides in the Venom of Nine Cone Snail Species by Transcriptomic and Mass Spectrometric Analysis. *J. Proteome Res.* **2017**, *16* (2), 763-772.
98. Klodt, J.; Kuhn, M.; Marx, U. C.; Martin, S.; Rosch, P.; Forssmann, W. G.; Adermann, K., Synthesis, biological activity and isomerism of guanylate cyclase C-activating peptides guanylin and uroguanylin. *J. Pept. Res.* **1997**, *50* (3), 222-30.
99. Li, G.; Delafield, D. G.; Li, L., Improved structural elucidation of peptide isomers and their receptors using advanced ion mobility-mass spectrometry. *TrAC, Trends Anal. Chem.* **2020**, *124*, 115546.
100. Zimnicka, M. M.; Troc, A., Acid-based approach for separation of peptide epimers using IM-MS. *J. Mass Spectrom.* **2019**, *54* (7), 620-628.
101. Li, G.; DeLaney, K.; Li, L., Molecular basis for chirality-regulated A β self-assembly and receptor recognition revealed by ion mobility-mass spectrometry. *Nat. Commun.* **2019**, *10* (1), 5038.

102. Do, T. D.; Checco, J. W.; Tro, M.; Shea, J. E.; Bowers, M. T.; Sweedler, J. V., Conformational investigation of the structure-activity relationship of GdFFD and its analogues on an achatin-like neuropeptide receptor of *Aplysia californica* involved in the feeding circuit. *Phys. Chem. Chem. Phys.* **2018**, *20* (34), 22047-22057.
103. Pang, X.; Jia, C.; Chen, Z.; Li, L., Structural Characterization of Monomers and Oligomers of D-Amino Acid-Containing Peptides Using T-Wave Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 110-118.
104. Jeanne Dit Fouque, K.; Garabedian, A.; Porter, J.; Baird, M.; Pang, X.; Williams, T. D.; Li, L.; Shvartsburg, A.; Fernandez-Lima, F., Fast and Effective Ion Mobility-Mass Spectrometry Separation of d-Amino-Acid-Containing Peptides. *Anal. Chem.* **2017**, *89* (21), 11787-11794.
105. Jia, C.; Lietz, C. B.; Yu, Q.; Li, L., Site-specific characterization of D-amino acid containing peptide epimers by ion mobility spectrometry. *Anal. Chem.* **2014**, *86* (6), 2972-81.
106. Wu, C.; Siems, W. F.; Klasmeier, J.; Hill, H. H., Jr., Separation of isomeric peptides using electrospray ionization/high-resolution ion mobility spectrometry. *Anal. Chem.* **2000**, *72* (2), 391-5.
107. de Magalhães, M. T.; Barbosa, E. A.; Prates, M. V.; Verly, R. M.; Munhoz, V. H.; de Araújo, I. E.; Bloch, C., Jr., Conformational and functional effects induced by D- and L-amino acid epimerization on a single gene encoded peptide from the skin secretion of *Hypsiboas punctatus*. *PLoS One* **2013**, *8* (4), e59255.

108. Serafin, S. V.; Maranan, R.; Zhang, K.; Morton, T. H., Mass spectrometric differentiation of linear peptides composed of L-amino acids from isomers containing one D-amino acid residue. *Anal. Chem.* **2005**, *77* (17), 5480-7.
109. Adams, C. M.; Zubarev, R. A., Distinguishing and quantifying peptides and proteins containing D-amino acids by tandem mass spectrometry. *Anal. Chem.* **2005**, *77* (14), 4571-80.
110. Adams, C. M.; Kjeldsen, F.; Zubarev, R. A.; Budnik, B. A.; Haselmann, K. F., Electron capture dissociation distinguishes a single D-amino acid in a protein and probes the tertiary structure. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (7), 1087-98.
111. Lambeth, T. R.; Julian, R. R., Differentiation of peptide isomers and epimers by radical-directed dissociation. *Methods Enzymol.* **2019**, *626*, 67-87.
112. Koehbach, J.; Gruber, C. W.; Becker, C.; Kreil, D. P.; Jilek, A., MALDI TOF/TOF-Based Approach for the Identification of d- Amino Acids in Biologically Active Peptides and Proteins. *J. Proteome Res.* **2016**, *15* (5), 1487-96.
113. Tugyi, R.; Uray, K.; Iván, D.; Fellingner, E.; Perkins, A.; Hudecz, F., Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (2), 413-8.
114. Ewing, M. A.; Wang, J.; Sheeley, S. A.; Sweedler, J. V., Detecting D-amino acid-containing neuropeptides using selective enzymatic digestion. *Anal. Chem.* **2008**, *80* (8), 2874-80.
115. Buczek, O.; Yoshikami, D.; Watkins, M.; Bulaj, G.; Jimenez, E. C.; Olivera, B. M., Characterization of D-amino-acid-containing excitatory conotoxins and redefinition of the I-conotoxin superfamily. *FEBS J.* **2005**, *272* (16), 4178-88.

116. Mor, A.; Delfour, A.; Sagan, S.; Amiche, M.; Pradelles, P.; Rossier, J.; Nicolas, P.,
Isolation of dermenkephalin from amphibian skin, a high-affinity delta-selective opioid
heptapeptide containing a D-amino acid residue. *FEBS Lett.* **1989**, 255 (2), 269-74.

CHAPTER 3

DIFFERENTIAL POST-TRANSLATIONAL AMINO ACID ISOMERIZATION FOUND AMONG NEUROPEPTIDES IN APLYSIA CALIFORNICA

3.1 NOTES AND ACKNOWLEDGEMENTS

The text in this chapter is published in ACS Chemical Biology: Mast, D. H.; Checco, J. W.; Sweedler, J. V., Differential Post-Translational Amino Acid Isomerization Found among Neuropeptides in *Aplysia californica*. *ACS Chem Biol* **2020**, *15* (1), 272-281. DOI: <https://doi.org/10.1021/acscchembio.9b00910>.

Special thanks to P. Yau and B. Imai at the University of Illinois Roy. J. Carver Biotechnology Protein Sciences Facility for helpful discussions, as well as Z. Li at the Roy J. Carver Center for Metabolomics for assistance with LC–MRM experiments. Thanks to K. Perkins for contributing artwork. Funding and support was provided by the National Institutes of Health, National Institute of Neurological Disorders and Stroke under Award No. R01NS031609, and the National Institute on Drug Abuse under Award No. P30DA018310. J.W.C. was funded in part by the Beckman Institute Postdoctoral Fellows program. The National Resource for *Aplysia* (Miami, FL) is funded by PHS grant P40OD010952.

3.2 ABSTRACT

D-amino acid-containing peptides (DAACPs) are a class of post-translationally modified peptides in animals that play important roles as cell-to-cell signaling molecules. Despite the

functional importance of L- to D-residue isomerization, little is known about its prevalence, mostly due to difficulties associated with detecting differences in peptide stereochemistry. Prior efforts to discover DAACPs have been largely focused on pursuing peptides based on homology to known DAACPs or DAACP-encoding precursors. Here, we used a combination of enzymatic screening, mass spectrometry, and chromatographic analysis to identify novel DAACPs in the central nervous system (CNS) of *Aplysia californica*. We identified five new DAACPs from the pleurin precursor, and three DAACPs from previously uncharacterized proteins. Further, two peptides from the pleurin precursor, Plrn2 and Plrn3, exist as DAACPs with the D-residue found at either position 2 or position 3. These differentially modified forms of Plrn2 and Plrn3 are located in specific regions of the animal's CNS. Plrn2 and Plrn3 appear to be the first animal DAACPs where the D-residue is found at more than one position and suggests that L- to D-residue isomerization may be a more variable / dynamic modification than previously thought. Overall, this study demonstrates the utility of non-targeted DAACP discovery approaches to identify new DAACPs and demonstrates that isomerization is prevalent throughout the CNS of *A. californica*.

3.3 INTRODUCTION

D-amino acid-containing peptides (DAACPs) can be formed by an age-related isomerization of amino acid residues, usually aspartate, and are most commonly found in long-lived peptides and proteins such as lens crystallin proteins and amyloid-beta peptide.¹⁻⁴ In the metazoan, DAACPs may also be the products of an enzyme-catalyzed post-translational modification (PTM), which isomerizes an L-amino acid residue in a peptide into a D-amino acid residue.⁵⁻⁶ The functions of DAACPs in animals range from defense (i.e., venoms and toxins) to cell-cell signaling (i.e., hormones and neuropeptides).⁵⁻⁶ For DAACPs with identified bioactivity,

the D-residue can be critical for physiological functions or behaviors, with the all-L-stereoisomer displaying weaker potency or being completely inactive.⁵⁻¹¹ In fact, most DAACPs were discovered only after the biological activity of a purified endogenous peptide was not reproduced by the synthetic all-L-residue stereoisomer. However, several cases do exist in which a DAACP and the all-L-residue stereoisomer share similar receptor potency or physiological activity, but may differ in other functional properties, such as biological half-life.¹²⁻¹⁵

The biological relevance of animal cell-cell signaling DAACPs is not well known, likely because relatively few (~10–20) have been characterized and shown to be involved in this process.¹² In contrast to commonly identified neuropeptide PTMs, such as acetylation or amidation, peptide isomerization may be understudied because modern “peptidomics” characterization approaches do not routinely evaluate stereochemistry.¹⁶⁻²³ Unlike most PTMs, isomerization does not change the resulting peptide’s molecular mass or sequence, making this modification difficult to detect with commonly used liquid chromatography (LC)–tandem mass spectrometry (MS/MS) techniques.

We developed an approach to identify DAACPs in tissue extracts using a stereoselective enzymatic digestion.²⁴⁻²⁵ Peptides with a D-residue near the N-terminus can be identified and differentiated from peptides containing only L-residues in tissue extracts based on differences in the rate of degradation by the stereoselective exopeptidase, aminopeptidase M (APM). Variations of this approach have been previously used by our group and others to either discover or confirm DAACPs in animals.^{12, 24-26} When APM is added to a peptide mixture, the undigested peptides remaining after an extended incubation are putative DAACPs. However, choosing which peptides to pursue for DAACP confirmation studies based solely on a list of APM-resistant sequences is

challenging because there are other peptide sequences and modifications that resist APM digestion, independent of the presence of a D-residue.

In our previous work with *Aplysia californica*, **GdYFD** (**dX**, in bold and underlined indicates the D-residue) and **SdY**ADSKDEESNAALSDFA from the *A. californica* achatin-like neuropeptide precursor,²⁴ and **GdF**RLNSASRVAHGY-NH₂ from allatotropin-related peptide precursor,¹² were found to be DAACPs based on their resistance to APM digestion. However, these previous studies relied on homology-guided discovery. As an example, **GdY**FD and **SdY**ADSKDEESNAALSDFA were targeted because they are derived from the same prohormone as a well-studied DAACP, **GdF**FD.⁷ All peptides from the same precursor were expected to be exposed to the same processing enzymes in the secretory pathway, including a putative isomerase. Similarly, the peptide **GdF**RLNSASRVAHGY-NH₂ was investigated because of its sequence similarities to the previously characterized DAACP from cone snail venom, conomap-Vt.^{12, 27}

Here for the first time, the DAACP discovery funnel²⁴ was applied in a completely non-targeted manner, with the goal of identifying novel DAACPs from precursors other than those already known to encode isomerized peptides in the central nervous system (CNS) of *A. californica*. Initially, the DAACPs were screened for based on enzymatic stability, followed by evaluation of the chromatographic and mass spectrometric properties of the DAACP candidates. Using this non-targeted approach, eight novel DAACPs were able to be identified. Additionally, two peptide sequences from the pleurin precursor²⁸ were found to be isomerized into multiple epimer forms, with a D-residue at either position 2 or position 3 (from the N-termini), and the relative abundance of these epimers differed across brain regions of the *A. californica* CNS. Finally, three of the DAACPs that were discovered are encoded by previously uncharacterized proteins that likely represent novel neuropeptide prohormones. These results suggest that DAACPs

should be accounted for in future peptidomics studies, and that the discovery of novel DAACPs leads to uncovering novel prohormones.

3.4 METHODS

Chemicals

Solvents and reagents used, unless specified otherwise, were purchased from Thermo Fisher Scientific or MilliporeSigma.

Peptide Extraction

A. californica weighing 70–160 g were purchased from the NIH/University of Miami National Resource for *Aplysia*, and maintained at 14 °C in an aquarium filled with Instant Ocean (Aquarium Systems Inc.). Before dissection, animals were injected (66% volume/body weight) with a solution of 74 g/L $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ to anesthetize, after which the buccal, cerebral, left pleural, right pleural, and abdominal ganglia from 3–5 animals were dissected and separated into 50 mL beakers filled with artificial sea water (ASW) cooled on ice. Following dissection, individual ganglia were removed from the ASW and placed directly into a 500 μL solution of 70/30 acetone/water (v/v) with HCl at a final concentration of 0.2 N (acidified acetone). Ganglia were homogenized manually using mechanical homogenization, and sonicated for 1 min in an ice bath before centrifugation at $15,000 \times g$, 10 min, 4 °C. The supernatant was transferred to a clean tube and the remaining pellet was washed once with 200 μL of LC–MS-grade water and centrifuged at $15,000 \times g$, 5 min at 4 °C. The resulting supernatant was combined with the acidified acetone extract and dried. Samples were desalted using C18 solid-phase extraction (SPE). The SPE samples were dried and stored at –80 °C until analysis.

APM Digestion

APM from porcine kidney was purchased from EMD Millipore. APM screening was performed as previously described, with minor alterations.¹² The approximate peptide concentration of ganglia extracts from five animals was estimated using bicinchoninic acid. Following quantitation, each ganglia extract was dried and reconstituted in a solution of 50 mM Tris, 500 mM NaCl, pH 7.5, to a final peptide concentration of 25 µg/mL. Each ganglia extract was then divided into two separate tubes: APM was added to one tube to a final concentration of 0.3 U/mL (APM-treated); an equal volume of water was added to another tube (no-APM). The tubes were then incubated for 15 h at 37 °C on a shaker. Following incubation, the reaction was quenched by bringing each tube to 5% CH₃CN and 0.1% formic acid (FA) and placing the reaction on ice, and the quenched reactions were desalted by SPE. After desalting, the samples were dried, and each reconstituted in 40 µL of 5% CH₃CN, 0.1 % FA. For each ganglia extract, 3–6 µL of extract was used for LC–MS/MS analysis.

LC–MS/MS Analysis

LC was carried out on an Ultimate 3000 Nano-RSLC system (Thermo Fisher Scientific) coupled to a Bruker Impact quadrupole time-of-flight mass spectrometer with a CaptiveSpray ionization source, operated in positive ionization mode. MS/MS peptide sequencing was performed by collision induced dissociation fragmentation. Separation was performed using an Acclaim™ Pepmap™ 100 C18 nano-LC column, 75 µm i.d. × 15 cm, 2 µm particle size, 100 Å pore size (P/N 164534, Thermo Fisher Scientific) at 35 °C with a flow rate of 300 nL/min. Mobile phase compositions were as follows: 95% H₂O, 5% CH₃CN, 0.1% FA (solvent A) and 95%

CH₃CN, 5% H₂O, 0.1% FA (solvent B). The gradient used was 4–50% B over 90 min, 50–90% B over 5 min, 90% B for 10 min, 90–4% B over 5 min, 4% B for 10 min.

Peptide Identification

Peptides present in the no-APM control and the APM-treated extracts of buccal, cerebral, abdominal, pleural left, and pleural right ganglia were identified using PEAKS Studio 8.0 (Bioinformatics Solutions Inc.), searching against the NCBI *A. californica* protein database. EICs, MS1, and MS2 spectra of identified DAACPs are shown in Figures 3.22-3.27. MS/MS fragment ion identifications of endogenous peptides are provided in Figures 3.28-3.33.

Relative Quantitation of Plrn3 Stereoisomers

Cerebral and buccal ganglia extracts were prepared as above and run under the same LC–MS/MS analysis conditions described above. Relative abundances of Plrn3 were calculated using the peak areas from an EIC of m/z 772.9 ± 0.1 . Significance was tested using a two-sample t-test with a significance of $\alpha < 0.05$. Relative abundances were an average of N=4 biological sets of buccal and cerebral ganglia, where each biological set contained ganglia pooled from 3–5 animals.

LC Separations of Plrn2 Stereoisomers with MRM

Cerebral and buccal ganglia extracts were prepared as above, reconstituted in 50 μ L of 0.1% FA in water and sonicated for 5 min; 30 μ L of the resulting solution was injected onto an Agilent 1200 series HPLC system (Agilent Technologies). The LC separation of Plrn2 stereoisomers was performed on a Supelco Aztec Chirobiotic T2 column, 4.6 mm i.d. $\times$ 25 cm, 5 μ m particle size, 200 Å pore size (P/N 16024AST, MilliporeSigma). Mobile phase compositions

were as follows: 0.1% FA in water (solvent A) and 0.1% FA in acetonitrile (solvent B). The separation was performed at 35 °C using a flow rate of 1 mL/min and running a linear gradient as follows: 15% B over 2 min, 15–25% B over 15 min, 70% B over 2 min, and 15% B over 3 min. MRM MS analysis was performed using a benchtop hybrid triple quadrupole-linear accelerator trap LC–MS/MS system (5500 QTRAP; Sciex) located in the Metabolomics Lab of the Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign. Mass spectra were acquired under positive electrospray ionization. The parent ion m/z monitored was m/z 814.4 for ^{13}C -labeled synthetic peptide standards, and 813.4 for the endogenous peptide. Fragment ions monitored were m/z 547.1, 384.1, and 207.0 for both synthetic and endogenous Plrn2.

LC–MS/MS with TIMS

For the TIMS measurements of Plrn2 stereoisomers, LC was carried out on a nanoElute system (Bruker) using a nanoElute FIFTEEN C18 ReproSil AQ reversed-phase analytical column, 75 μm i.d. $\times$ 15 cm, 1.9 μm particle size, 120 Å pore size (P/N 1842621, Bruker) coupled to a timsTOF Pro mass spectrometer with a CaptiveSpray ionization source (Bruker). Separation was carried out at 40 °C at a flowrate of 30 nL/min. Mobile phase compositions were as follows: H_2O , 0.1% FA (solvent A) and CH_3CN , 0.1% FA (solvent B). The LC gradient conditions used were 4–50% B in 50 min, 50–95% B in 2 min, 95% B for 10 min. MS parameters used for TIMS: MS was collected in positive mode, with TIMS enabled and the ion charge control target set to 2.00 Mio. The imex setting was set to custom resolution. The reduced mobility ($1/k_0$) range was set to 0.9–1.2 $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$, ramp time 350 ms, spectra rate 2.82 Hz, and the lock accumulation to mobility range was selected. The TIMS tuning delta values were as follows: $\Delta 1 = -20.0$ V, $\Delta 2 = -80.0$ V, $\Delta 3 = 110.0$ V, $\Delta 4 = 110.0$ V, $\Delta 5 = 0.0$ V, $\Delta 6 = 120.0$ V. The tunnel out pressure was 8.06×10^{-1} mBar

and the tunnel in pressure was 2.57 mBar. For TIMS MS measurements, the TOF and TIMS were calibrated using Agilent low concentration tune mix introduced through the CaptiveSpray source.

Peptide Synthesis

Synthetic Plrn1, Plrn2, Plrn3, FMRGF-NH₂, Ip1, and Ip2 peptides, labeled with ¹³C-Gly, were synthesized using Fmoc solid-phase peptide synthesis, as described previously.¹² ¹³C-Fmoc-Gly-OH was purchased from Anaspec Inc. and labeled with one ¹³C at the carbonyl carbon. MS/MS spectra of synthetic peptides with labeled fragment ion identifications are shown in Figures 3.28-3.33.

Spiking Synthetic Standards

Peptide standards were diluted to 1 µM in water with 0.1% FA. For spiking, each standard was then combined with cerebral ganglia extracts (prepared as described above) to a final concentration between 0.01–0.5 µM, and 3–6 µL of extract was used for LC–MS/MS analysis using the method described above.

APM Digestions on Synthetic Peptide Standards

10 µM of ¹³C-labeled synthetic Plrn1, Plrn2, and Plrn3 peptide standards were prepared in a solution of 50 mM Tris, 500 mM NaCl, pH = 7.5 in 0.7 mL Eppendorf tubes. Aminopeptidase M (APM) was then added at a concentration of 1 U/mL and the tubes were incubated at 37 °C for 43 h. 1 µL aliquots were collected at several time points, combined with 1 µL of 50 mg/mL dihydroxybenzoic acid in 70/30 CH₃CN/H₂O with 0.1% trifluoroacetic acid and spotted on a polished steel target for matrix-assisted laser desorption/ionization (MALDI) mass spectrometry

(MS). Spots were allowed to dry in air. The samples were then analyzed on an ultrafleXtreme MALDI mass spectrometer (Bruker).

Determination of DAACP Candidates

Peptides present in the no-APM control, and the APM-treated samples were identified using PEAKS Studio 8.0 (Bioinformatics Solutions Inc.) by searching against the NCBI *A. californica* protein database. If the same peptide sequence was detected in both the APM-treated and no-APM control samples, extracted ion chromatograms (EICs) were generated and evaluated for the presence of two peaks in the no-APM control. A peptide was considered a DAACP candidate if the same sequence appeared as two peaks in the no-APM control bearing identical m/z values and highly similar tandem MS (MS/MS) fragmentation, but not the APM-treated sample. Fresh extracts were then prepared as described in the main text and peptide identification was performed by liquid chromatography (LC)–MS/MS to confirm that split eluting peaks of the peptides were not a result of isomerization during extended incubation at neutral pH. The EICs, and MS1 and MS2 spectra of Plrn1, Plrn2, Plrn3, FMRGF-NH₂, Ip1, and Ip2 are shown in Figures 3.22-3.27.

Synthetic Peptide Purification, Sequence Confirmation, and Purity Evaluation

Peptides were purified using reversed phase HPLC as described previously.¹² Peptide purity was assessed with capillary scale reversed phase LC with UV detection at 214 nm. An Acclaim PepMap RSLC C18 reversed phase column, 300 μm i.d. $\times$ 150 mm, 3 μm particle size, 100 Å pore size (P/N 164571, Thermo Fisher) was used for capillary LC purity check. All peptides were >90% purity. For purity checks, the solvent compositions were as follows: water with 4%

acetonitrile and 0.1% formic acid (Solvent A) and 80% acetonitrile with 0.1% formic acid (solvent B). A linear gradient from 4–50% B in 30 min was used at a flowrate of 4 μ L/min and a column temp of 35 °C. 30–100 pmol of each standard was injected onto the capillary LC. The sequences were confirmed by nano-scale LC–MS/MS using the LC–MS/MS method described in the main text. 0.05–0.5 pmol of synthetic peptide was injected for MS/MS confirmation of each synthetic peptide sequence (Figures 3.28–3.33).

Data Handling

Raw LC–MS/MS data (Bruker .d format) were exported from Bruker DataAnalysis software as .mzML files and loaded into openMS TOPPview software version 2.1.⁴³ EICs were exported from openMS as .mzML files and loaded into Origin Pro 2019 software version 9.6.0. MS2 and MS1 spectra were exported directly from Bruker DataAnalysis as .xy files and loaded into Origin Pro for further processing. Plots were assembled into figures using Inkscape or Adobe Illustrator software.

Sequences Resisting APM Digestion

Consistent with previous studies that have reported the specificity of APM,^{29–31} we noticed that peptides with PyroGlu, Glu-Xaa, Asp-Xaa, Xaa-Pro, and Pro-Xaa at the N-terminus were not well hydrolyzed by APM. In our study, 741 peptides were detected in the no-APM control experiments and 269/741 peptides were resistant to APM degradation, while 472/741 were degraded. 63% of the APM resistant peptides were expected to resist degradation and the N-terminal residues were as follows: Glu-Xaa (9%), Asp-Xaa (8%), Pro-Xaa (5%), Xaa-Pro (13%),

pyroGlu-Xaa (28%). 37% of the APM-resistant peptides did not have these sequence characteristics, and thus would be expected to degrade.

Trapped Ion Mobility Spectrometry (TIMS) Detection

For TIMS measurements, the TIMS device was calibrated using an Agilent low concentration ESI tuning mix (Agilent Technologies, P/N G1969-85010) over a $1/k_0$ range of 0.5–2.0 $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$. In order to optimize the resolution for the TIMS separation of Plrn2 stereoisomers, after the calibration, the $1/k_0$ range of the measurement was adjusted to 0.9–1.2 $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$ and we found this resulted in a slight shift in the $1/k_0$ of the ^{13}C -labeled Plrn2 standards and calibrant peaks. Importantly, the purpose of using TIMS was only to separate the endogenous Plrn2 stereoisomers and compare their mobilities to the synthetic standards. Modeling of gas phase structures and collisional cross sections was not relevant to the analysis.

MS/MS Differentiation of [D-Phe²]-Plrn2 and [D-Tyr³]-Plrn2.

Reversed-phase, Nano LC–MS/MS analysis was performed using the methods described in the main text to collect MS2 spectra on endogenous Plrn2, and the MS2 spectra of synthetic Plrn2 peptides were obtained as described above. Previous studies have shown peptide stereoisomers can be differentiated based on the relative abundance of fragment ions.^{31,44–45} In order to test whether [D-Tyr³]-Plrn2 may be present endogenously, we analyzed the MS2 spectra of the second eluting Plrn2 peaks (m/z 542.6, $z = 3+$) from the no-APM control samples of buccal and cerebral ganglia extracts for differences in product ion ratios. The MS2 spectra of second eluting Plrn2 peaks collected by LC–MS/MS on fresh extracts of buccal and cerebral ganglia were also compared and included as biological replicates. Collision induced dissociation (CID) was

used for fragmentation and all measurements were made with a collision energy of 21.3 eV. The 3+ charge state of Plrn2 ions was used for all measurements. MS2 spectra were normalized to the most intense peak and the y_{10}^{2+}/y_{11}^{2+} ratio was calculated after normalization because this ratio showed the greatest differences. The y_{10}^{2+}/y_{11}^{2+} ratio of the synthetic standards was then compared to the y_{10}^{2+}/y_{11}^{2+} ratio of the second eluting Plrn2 peak from buccal and cerebral ganglia extracts. The results from these experiments are shown in Figure 3.18.

3.5 RESULTS AND DISCUSSION

APM screening

The *A. californica* CNS is comprised of five primary ganglia: buccal, cerebral, pedal, pleural, and abdominal (Figure 3.1). Expression of most prohormones is known to be ganglia specific and certain peptides can be present at significant quantities only in specific ganglia.²⁸ LC–MS/MS peptide identification was performed on APM-treated extracts from pleural, cerebral, abdominal, and buccal ganglia from the *A. californica* CNS to construct a list of peptide sequences that were resistant to APM degradation. Pedal ganglia were omitted from this study because they were the focus of our previous work.²⁴ The list of APM-resistant peptides was compared to a list of sequences present in extracts not treated with APM (no-APM) to ensure the peptides in the APM-treated list were present in the extract before enzyme treatment. Peptides observed in both lists were considered DAACP candidates. Ideally, the all-L residue peptides would be degraded by APM, leaving only peptides with a D-residue near the N-terminus in the extract. However, not all peptides resistant to APM digestion have a D-residue because APM tends not to cleave certain residues, such as Asp, Glu, Pro or modified N-termini.²⁹⁻³¹

DAACPs are often detected alongside the all-L residue peptides in biological extracts, and the two species are typically separable by reversed-phase LC.^{5, 32-33} Therefore, we examined whether any of the candidates possibly existed as both the all-L and D-amino acid-containing forms in the extract without APM treatment by checking the LC–MS/MS chromatograms for multiple eluting peaks corresponding to the same peptide sequence. DAACP candidate sequences were selected for follow-up analysis if they were observed to elute at two distinct retention times in the no-APM extracts and the same sequence was only detected at one retention time after APM digestion. The APM-resistant peak likely corresponded to a DAACP, while the peak degraded by APM likely corresponded to the all-L peptide. Any peptides that had N-terminal Asp, Glu, Pro, Xaa-Pro, or modified N-termini were excluded from our analysis because resistance to APM by these peptides may not be the result of a D-residue.

Following our analysis, Plrn1, Plrn2, and Plrn3 from the pleurin prohormone, Ip1 and Ip2 from an uncharacterized protein (NCBI accession: XP_012946424.1), and FMRGF-NH₂ from one of three possible uncharacterized proteins (NCBI accessions: XP_005094941.1, XP_005094942.1, XP_005094943.1), eluted at two (or more) retention times in the no-APM extracts, and had only one peak resistant to degradation in the APM-treated extracts (Figure 3.2). Amino acid sequences of the identified DAACP candidates are reported in Table 3.1. The six APM-resistant peptides mentioned above were detected in the cerebral ganglia extracts, but some of these DAACP candidates were not observed in the extracts from other ganglia (Table 3.2). Extracted ion chromatograms (EICs) of putative DAACPs obtained from the buccal, pleural, and abdominal ganglia extracts are shown in Figure 3.3.

The resistant peptides mentioned above were expected to likely have a D-residue near the N-terminus based on the stereoselectivity of APM, and more specifically, have a D-residue at the

second position from the N-terminus because the D-residue for all previously reported *A. californica* DAACPs has been at position 2.^{7, 9, 12, 24}

The stereochemistry of Plrn1, Plrn2, Plrn3, Ip1, Ip2, and FMRGF-NH₂ was tested by comparing the chromatographic retention of synthetic standards and endogenous peptides using reversed-phase LC with MS detection. Peptide standards were synthesized with isotopically labeled ¹³C-Gly; the sequences are reported in Table 3.3. Figure 3.4 shows the EICs of endogenous Plrn1, Plrn2, Plrn3, Ip1, Ip2, and FMRGF-NH₂, as well as the synthetic ¹³C-labeled D2-peptide standards spiked into the cerebral ganglia extracts. The synthetic D2-peptide standards co-eluted with the peaks corresponding to the endogenous APM-resistant Plrn1, Plrn2, Plrn3, Ip1, Ip2, and FMRGF-NH₂ (Figure 3.4). Co-elution of the synthetic D2-standards with the endogenous APM-resistant peptides supports the idea that the endogenous peptides are DAACPs with a D-residue in position 2. Further, the all-L-residue synthetic standards did not co-elute with the suspected DAACPs but did co-elute with peaks that were susceptible to APM digestion (Figure 3.5).

Identification of [D-Phe³]-Plrn3

As seen in Figure 3.4C, the EIC of Plrn3 shows three major peaks sharing the same parent mass and MS/MS fragmentation ions: the first peak (30.5 min) was identified as L-Plrn3 (Figure 3.5) and the second peak (33.5 min) was identified as [D-²Ala]-Plrn3, as described above. The third peak (34.5 min) was not fully resistant to APM (Figure 3.2C) and did not co-elute with either of the ¹³C-[D-²Ala]-Plrn3 or ¹³C-L-Plrn3 synthetic standards (Figure 3.4C and Figure 3.5C, respectively). It was plausible that the third peak might be a stereoisomer of Plrn3 with a D-residue at an amino acid position further from the N-terminus than position 2 because it did not co-elute with the synthetic all-L or D2-peptides but was still degraded by APM. Analysis of the APM

digestion experiments showed an increase in partially digested Plrn3(2-14) lacking the Gly residue in position 1 after a 15 h digestion with APM that was not present in the controls (Figure 3.6). Because Plrn3(2-14) remained resistant to APM for the duration of the APM digestion, we expected that the identity of the third eluting Plrn3 peak (34.5 min) in Figure 3.4C corresponded to a stereoisomer bearing a D-Phe at position 3. Consistent with this supposition, when the synthetic ^{13}C -[D-Phe³]-Plrn3 was spiked into the cerebral ganglia extracts, it co-eluted with the peak at 34.5 min (Figure 3.7C).

To ensure the origin of the endogenous Plrn3(2-14) observed in the APM-treated ganglia extracts is from [D-Phe³]-Plrn3 and not from endogenous L- or D2-peptide, an APM digestion was performed on synthetic standards (Figure 3.8). ^{13}C -L-Plrn3 and ^{13}C -[D-alIe²]-Plrn3 were both treated with APM, but neither ^{13}C -L-Plrn3(2-14) or ^{13}C -[D-aIle²]-Plrn3(2-14) were detected in the digests of these peptides (Figure 3.8). The synthetic ^{13}C -[D-Phe³]-Plrn3 was also treated with APM and it was degraded into ^{13}C -[D-Phe³]-Plrn3(2-14), which remained stable for as long as 43 h. The latter result indicates the Plrn3(2-14) observed in the APM-treated ganglia extracts likely originated from [D-Phe³]-Plrn3.

Identification of [D-Tyr³]-Plrn2

Because endogenous Plrn3 was found to have a D-residue in position 3, we tested whether the other two peptides from the pleurin prohormone also existed with a D-residue at position 3. As seen in Figure 3.4A, B endogenous Plrn1 and Plrn2 eluted at only two retention times, indicating either a third stereoisomer was not present, or that a third stereoisomer co-eluted with one of the two peaks. We tested this by comparing the retention times of synthetic ^{13}C -[D-Tyr³]-Plrn1 and ^{13}C -[D-Tyr³]-Plrn2 to those of endogenous peptides by spiking. As shown in Figure 3.7A, ^{13}C -[D-

Tyr³]-Plrn1 did not co-elute with either of the endogenous Plrn1 peaks. Only the L and D2 forms of Plrn1 could be identified in this study.

We found that ¹³C-[D-Tyr³]-Plrn2 and ¹³C-[D-Phe²]-Plrn2 both co-eluted with the endogenous APM-resistant Plrn2 peptide on a C18 stationary phase (Figure 3.7B). In a control experiment, we found that APM degraded the synthetic ¹³C-[D-Tyr³]-Plrn2 into ¹³C-[D-Tyr³]-Plrn2(2-14) (Figure 3.9). Thus, if [D-Tyr³]-Plrn2 was present in the cerebral ganglia extract, we would have expected it to be partially degraded by APM into [D-Tyr³]-Plrn2(2-14). However, we did not detect the *m/z* corresponding to Plrn2(2-14) in APM-treated cerebral ganglia extracts (Figure 3.10A). In contrast, the second eluting Plrn2 peak in the APM-treated buccal extract was mostly degraded, and this degradation corresponded to an increase in the presence of Plrn2(2-14) (Figure 3.10B). These results suggest that [D-Tyr³]-Plrn2 may be present and localized to the buccal ganglion. The detection of Plrn2(2-14) in APM-treated ganglia extracts does not unambiguously demonstrate that [D-Tyr³]-Plrn2 is present because we did find that some peptides with a D-residue in position 2 can still be partially degraded by APM. For example, we observed partial degradation of [D-Phe²]-Plrn1 into [D-Phe²]-Plrn1(2-14) in the presence of APM (Figures 3.11 and 3.12).

To obtain further evidence for the presence of [D-Tyr³]-Plrn2, the endogenous Plrn2 stereoisomers were separated by LC using a chiral stationary phase, and their retention times were compared to the retention times of the synthetic standards. LC with multiple reaction monitoring (MRM) MS was used for detection of endogenous peptides. Extracts from both buccal and cerebral ganglia were tested for [D-Tyr³]-Plrn2 because APM screening had indicated [D-Tyr³]-Plrn2 would be detected in the buccal ganglia extracts, but not detected in the cerebral ganglia extracts. Three Plrn2 peaks were observed in both the cerebral and buccal ganglia extracts that eluted at the same

retention times as the Plrn2 synthetic standards (Figure 3.13). The elution order of the synthetic Plrn2 standards on the chiral column was determined in a separate experiment (Figure 3.14). Consistent with the results obtained by APM screening, the LC–MRM results showed [D-Tyr³]-Plrn2 is present in the buccal ganglia extracts. Importantly, the relative abundance of [D-Phe²]-Plrn2 was found to be substantially reduced in the buccal ganglion compared to the cerebral ganglion. Additionally, the relative abundance of [D-Tyr³]-Plrn2 was increased in the buccal ganglion compared to the cerebral ganglion. This further confirms that [D-Tyr³]-Plrn2 is localized to the buccal ganglion.

Evaluation of [D-Tyr³]-Plrn2 by Trapped Ion Mobility Spectrometry (TIMS)

DAACPs can adopt different gas phase structures than their all-L-residue counterparts, allowing them to be separated via TIMS.³⁴⁻³⁷ As a complementary method to the LC–MS experiments described above, we next combined LC–MS/MS TIMS to evaluate whether [D-Tyr³]-Plrn2 was present endogenously. Initially a TIMS separation was performed on synthetic Plrn2 standards. The ¹³C-[D-Phe²]-Plrn2 and ¹³C-[D-Tyr³]-Plrn2 [m+2H]²⁺ ions were found to have slightly different inverse reduced mobilities (1/k₀) during TIMS. As shown in Figure 3.15A, the D2 and D3 stereoisomers of Plrn2 each adopted two gas phase conformations. The two ¹³C-[D-Tyr³]-Plrn2 peaks were only partially resolved but two distinct peaks at 1.010 V·s·cm⁻² and 1.020 V·s·cm⁻² were observed (Figure 3.15). ¹³C-[D-Phe²]-Plrn2 had a major peak observed at 1.027 V·s·cm⁻² and a minor peak at 1.048 V·s·cm⁻². The extracted ion mobilogram (EIM) obtained for ¹³C-L-Plrn2 is provided in Figure 3.16.

After optimizing the TIMS separation on synthetic standards, we measured the 1/k₀ of the endogenous peptides. As mentioned above, APM screening indicated [D-Tyr³]-Plrn2 was detected

in the buccal ganglia, but not the cerebral ganglia. Therefore, we measured the stereochemistry of Plrn2 from both the buccal and cerebral ganglia by TIMS. The EIM obtained from the cerebral ganglia extracts showed peaks at $1.030 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ and $1.053 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$, indicating [D-Phe²]-Plrn2 was present. However, peaks at $1.010 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ or $1.020 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ were not observed in the EIM, indicating [D-Tyr³]-Plrn2 was not detected in the cerebral ganglia extracts. In contrast, the EIM obtained from buccal ganglia extracts showed partially resolved peaks at $1.010 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ and $1.020 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$, consistent with $1/k_0$ values measured for the synthetic ¹³C-[D-Tyr³]-Plrn2 (Figures 3.15B and 3.16B). Additionally, a third, partially resolved peak was observed at $1.029 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ in the EIM from buccal ganglia extracts, indicating [D-Phe²]-Plrn2 and [D-Tyr³]-Plrn2 coexist in the buccal ganglion. Note, the $1/k_0$ for ¹³C-[D-Phe²]-Plrn2 standard was shifted by roughly $0.003 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ compared to the $1/k_0$ measured for the endogenous [D-Phe²]-Plrn2 from cerebral ganglia extracts. We expect this shift was the result of drift in the TIMS calibration over time. Importantly, the $1/k_0$ values reported here were not intended to be used to calculate the collisional cross sections of the Plrn2 stereoisomers.

Relative Abundance of D2 and D3 Stereoisomers Between Regions of the CNS

The TIMS results described above were consistent with the results obtained from APM screening and LC–MRM, indicating [D-Tyr³]-Plrn2 is present in the buccal ganglion. Furthermore, both TIMS and LC–MRM showed a reduced relative abundance of [D-Phe²]-Plrn2 in buccal ganglia extracts compared to cerebral ganglia extracts.

The observed differences in the relative abundance of D2 and D3 stereoisomers between regions of the CNS were not unique to Plrn2. We also observed distinct differences in the relative abundance of Plrn3 stereoisomers between the buccal and cerebral ganglia extracts by LC–MS.

Specifically, [D- AlIe^2]-Plrn3 was not detected in buccal ganglia extracts whereas both [D- AlIe^2]-Plrn3 and [D- Phe^3]-Plrn3 were detected in cerebral ganglia extracts. Because the three Plrn3 stereoisomers were well resolved by reversed-phase LC, we were able to quantify their relative abundances. Figure 3.17A shows EICs of Plrn3 obtained from the buccal or cerebral ganglia extracts, and Figure 3.17B shows the averaged relative abundance of Plrn3 stereoisomers from four biological sets. In four replicates, [D- AlIe^2]-Plrn3 was not detected in buccal ganglia extracts, demonstrating that the absence of [D- AlIe^2]-Plrn3 was not the result of biological variability.

Though we did not quantify the relative abundance of Plrn2 stereoisomers, both LC–MRM and TIMS showed a decreased relative abundance of [D- Phe^2]-Plrn2 in buccal ganglia extracts compared to cerebral ganglia extracts. Additionally, we were able to show the [D- Tyr^3]-Plrn2 stereoisomer comprises a greater relative abundance in the buccal ganglia extracts than in cerebral ganglia extracts by comparing the ion ratios obtained from collision induced dissociation (Figure 3.18). The region-specific distribution of Plrn2 and Plrn3 we observed may indicate that the PTM processing of these peptides is different in the buccal ganglion than in the cerebral ganglion.

Differential Isomerization of Plrn2 and Plrn3

Plrn2 and Plrn3 are the first examples of isomerization occurring at the third position of a peptide in *A. californica*. To our knowledge, Plrn2 and Plrn3 also represent the first reported examples (in animals) of L- to D-residue isomerization occurring at different positions within the same peptide sequence. The closest known example of this observed differential isomerization is found among hormones from crustaceans. It was previously found that the crustacean hyperglycemic hormone (CHH) is isomerized at [Phe³] and the vitellogenesis-inhibiting hormone (VIH) is isomerized at [Trp⁴].^{14, 38} Studies on the cellular distribution of VIH and CHH found that

isomerization of these peptides occurs in unique cell populations, indicating differential cellular expression of one or more isomerases.⁶ However, unlike VIH and CHH from crustaceans, Plrn2 and Plrn3 from *A. californica* are differentially isomerized, with the D-residue occurring at either of two positions within the same peptide sequence.

The unexpected discovery of [D-Tyr³]-Plrn2 and [D-Phe³]-Plrn3 might indicate the L/D-isomerase responsible for the biogenesis of these DAACPs has somewhat loose specificity for isomerization of positions 2 versus 3. The five previously identified DAACPs in *A. californica* are isomerized at aromatic amino acids in position 2. Isomerization of Plrn2 and Plrn3 at positions 2 and 3 may result because the isomerase favors aromatic amino acid residues. The lack of [D-Tyr³]-Plrn1 in biological extracts was somewhat surprising, given that Plrn1 also derives from the pleurin prohormone and shares a similar N-terminal sequence to Plrn2 (Table 3.1). This difference in the specificity of the L/D-isomerase towards residue positions may arise from subtle differences in the peptide sequences. Alternatively, it is possible that there exist two distinct L/D-isomerases responsible for isomerization at position 2 vs. position 3.

Identification of Aplysia Isomerized Peptides Precursor and FMRGFamide Precursors

As mentioned above, the two newly identified DAACPs, Ip1 and Ip2, are encoded on an uncharacterized predicted protein (NCBI accession: XP_012946424.1) supported by transcript and expressed sequence tag evidence in *A. californica*. In prior experiments our lab identified the Ip1 sequence by *de novo* sequencing of whole CNS lysates.³⁹ In the current study, a total of 11 peptides, including Ip1 and Ip2, were detected in screening experiments that derived from the same uncharacterized protein. Consistent with the proteolytic processing expected of peptide prohormones, all detected peptides were flanked by dibasic cleavage sites on the precursor protein.

According to SignalP,⁴⁰ there is a 50% probability of signal peptide cleavage between residues 31 and 32 of the precursor, suggesting the protein can be targeted to the secretory pathway. A BLAST search of this protein showed it has 32% sequence similarities to *A. californica* pedal peptides 2 precursor (NP_001191623.1), and 38% similarity to feeding circuit activating peptides precursor (FCAP, NP_001191518.1). We refer to this uncharacterized protein as the *Aplysia* isomerized peptides precursor (AIPP) to describe the PTM observed on two of the encoded peptides. The sequence coverage of AIPP by the detected peptides is reported in Figure 3.19, Table 3.4); and the complete annotation of the processing of AIPP will be the subject of a follow-up study.

Another novel DAACP identified in this work, FMRGF-NH₂, is one of 40 peptides matching sequences encoded in three uncharacterized protein isoforms (XP_005094941.1, XP_005094942.1, XP_005094943.1) predicted from the *A. californica* transcripts (XM_005094884.2, XM_005094885.2, XM_005094886.2). The three proteoforms have 95% signal peptide cleavage probability between residues 29 and 30 of the precursor, and numerous conventional proteolytic cleavage sites characteristic of neuropeptide precursors. A BLAST search showed the proteins have >90% sequence similarity between themselves and ~40% similarity to the *A. californica* FMRFamide precursor (AAA27752.1, P08021.3) and mytilus inhibitory peptide-related peptides precursor (NP_001191614.1). We refer to these proteins as the *A. californica* FMRGFamide precursor(s). The novel prohormone sequence(s) is supported by LC–MS/MS detection of numerous predicted peptides (Figure 3.20, Table 3.5). Peptides detected by LC–MS/MS matched sequences common in all three predicted proteoforms and could have originated from any of the three proteins. It was beyond the scope of this work to determine the exact transcript variant that leads to the observed FMRGF-NH₂ peptides. An interesting question

is whether only one of the three FMRGFamide proteins produces the [D-Met²]-FMRGF-NH₂, and this will be a subject for future studies.

The evidence presented above indicates that the AIPP and FMRGFamide-related peptides detected during this screening were likely the result of post-translational proteolytic processing of the precursor by prohormone convertases.

Prior studies have indicated that multiple peptides from the same prohormone are often isomerized.^{24, 41-42} Consistent with previously characterized prohormones encoding DAACPs, we identified three peptides from the pleurin precursor and two peptides from AIPP (Ip1 and Ip2) as DAACPs. This is consistent with a mechanism by which peptides are first cleaved from their precursor by prohormone convertases, and then isomerized within the dense core vesicles. However, in APM screening and peptide identification experiments, nine peptides were detected from AIPP that all shared similar sequences (Table 3.4), but only Ip1 and Ip2 were found to be DAACPs. This was surprising given the sequence similarities between these peptides. Similarly, among numerous similar peptides predicted and detected from FMRGFamide precursor(s), [D-Met²]-FMRGF-NH₂ was the only DAACP identified from the prohormone (Table 3.5).

While the selective isomerization may be caused by enzyme processing, it may be technical in nature. Although FMRGFamide and AIPP encode numerous similar peptide sequences, Ip1, Ip2, and FMRGF-NH₂ have multiple repeats of the same sequence compared to the other peptides encoded on their precursors. Following proteolytic processing, Ip1, Ip2, and FMRGF-NH₂ exist at much higher concentrations within the dense core vesicles relative to the other possible isomerase substrates encoded on the prohormones. Therefore, although the other possible isomerase substrates may be present from the same prohormone, the isomerase is more likely to encounter

Ip1 and Ip2 or FMRGF-NH₂, or perhaps the isomerized peptides present at lower levels are not detectable.

3.6 CONCLUSION

Using a combination of enzymatic screening, analysis of LC–MS/MS properties, and synthetic peptide standards, we identified eight novel DAACPs from three different prohormone precursors—two of which were previously uncharacterized and are described for the first time in this study. By systematically evaluating potential DAACPs based on both their APM resistance and chromatographic properties, we were able to identify DAACPs purely based on their chemical properties without regard to homology to other known DAACPs. Importantly, our approach also allowed us to identify the first animal peptides to undergo differential post-translational isomerization within the same sequence, and we found that this processing differed across different regions of the CNS. This study represents the most comprehensive discovery effort for endogenous DAACPs in a single organism to date. However, the DAACP discovery method reported here is restricted to finding D-residues near the N-terminus and for sequences having N-terminal Asp, Glu, Pro, Xaa-Pro- or modified N-termini. Thus, a fraction of the DAACPs may have been overlooked and future work will require new methods to address these sequence constraints. Despite the latter limitations, our findings indicate that DAACPs are more common than previously thought in the *A. californica* CNS. They also suggest that bioactivity studies on all-L-peptides in *A. californica* may have missed activity because of the possible presence of unrecognized D-residues. Isomerization of amino acid residues in peptides appears to be an underreported PTM and may have a broader scope than *A. californica*. Future studies should be carried out to determine the

prevalence of peptide isomerization in a range of animals. An important next step is the identification and characterization of the isomerase enzyme responsible for this modification.

3.7 FIGURES AND DATA

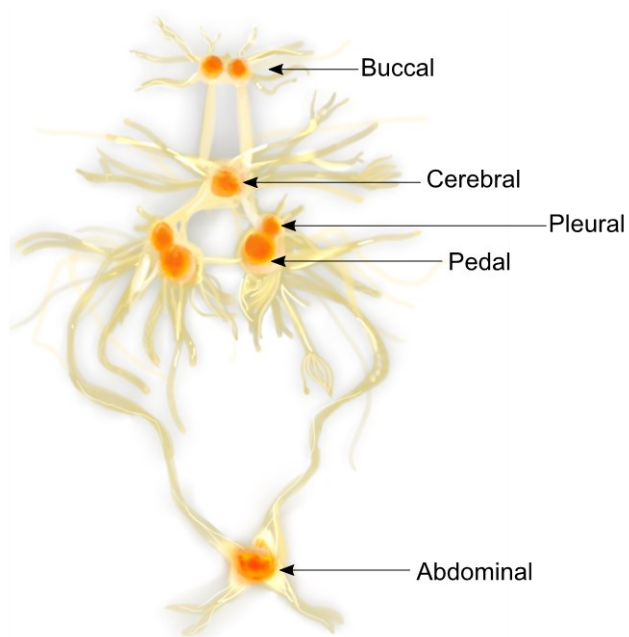


Figure 3.1. Schematic representation of the *A. californica* central nervous system (not drawn to scale), highlighting each of the major ganglia. Image created by K. Perkins.

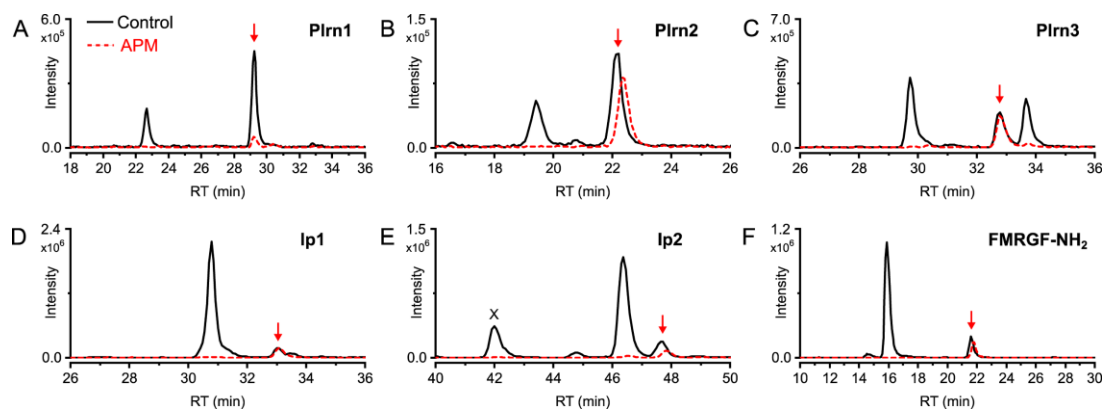


Figure 3.2. APM screening results from cerebral ganglia extracts showing the APM-resistant peptides. Plots show LC–MS EICs of (A) Plrn1 (m/z 560.6, $z = 3+$), (B) Plrn2 (m/z 542.6, $z = 3+$), (C) Plrn3 (m/z 772.9, $z = 2+$), (D) Ip1 (m/z 552.3, $z = 2+$), (E) Ip2 (m/z 1051.6, $z = 1+$), and (F) FMRGF-NH₂ (m/z 328.3, $z = 2+$) incubated for 15 h without APM (black trace) and incubated with APM for 15 h (dashed red trace). The red arrows indicate the APM-resistant peak. The peak labeled with an X in (E) was identified as an unrelated peptide sequence.

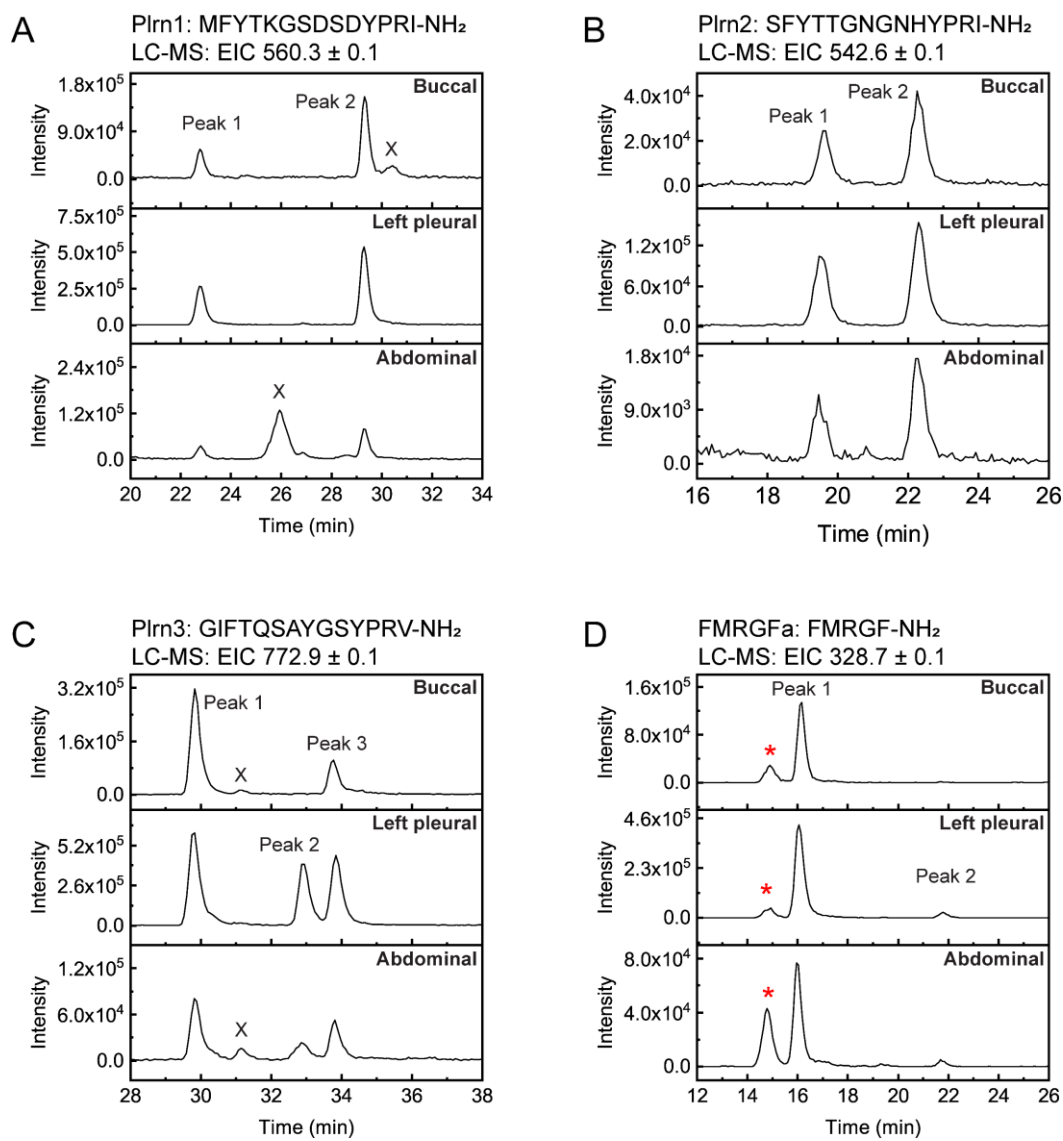


Figure 3.3. (A–D). EICs of DAACP sequences from no-APM control extracts of the buccal, left pleural, and abdominal ganglia regions of the *A. californica* CNS as indicated in the plot. EICs of (A) Plrn1 (m/z 560.3 ± 0.1, z = 3+) Peak 1 (L-Plrn1), Peak 2 ([D-Phe²]-Plrn1); (B) Plrn2 (m/z 542.6 ± 0.1, z = 3+) Peak 1 (L-Plrn2), Peak 2 ([D-Phe²]-Plrn2; (C) Plrn3 (m/z 772.9 ± 0.1, z = 2+) Peak 1 (L-Plrn3) and Peak 2 ([D-alle²]-Plrn3), Peak 3 ([D-Phe³]-Plrn3; (D) FMRGF-NH₂ (m/z 328.7 ± 0.1, z = 2+) Peak 1 (L-FMRGF-NH₂) and Peak 2 ([D-Met²]-FMRGF-NH₂). The peaks labeled with X's indicate a peak that exhibited a different MS1 and/or MS2 from the sequence of interest and thus is not the same peptide primary sequence. Peaks labeled with a red * indicate the m/z is the same as the suspected DAACP, but the identity and/or stereochemistry of the peak was not confirmed.

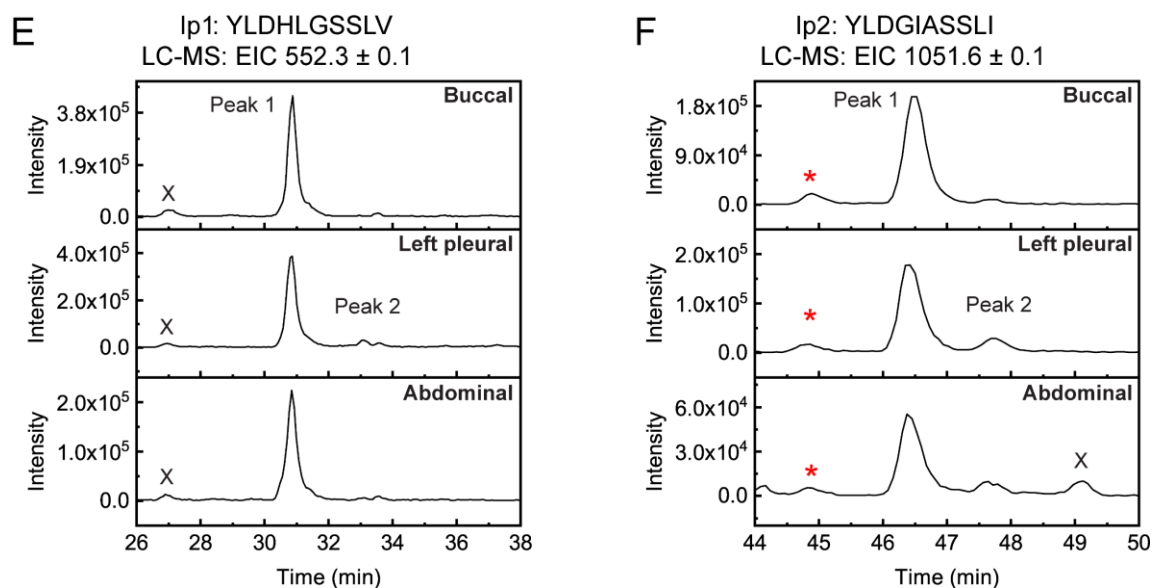


Figure 3.3. cont. (E–F). EICs of DAACP sequences from no-APM control extracts of the buccal, left pleural, and abdominal ganglia regions of the *A. californica* CNS as indicated in the plot. (E) Ip1 (m/z 552.3 $\pm$ 0.1, $z = 2+$) Peak 1 (L-Ip1) and Peak 2 ([D-Leu²]-Ip1); (F) Ip2 (m/z 1051.6 $\pm$ 0.1, $z = 1+$) Peak 1 (L-Ip2) and Peak 2 ([D-Leu²]-Ip2). The peaks labeled with X's exhibited a different MS1 and/or MS2 from the sequence of interest and thus are not the same peptide primary sequence. Peaks labeled with a red * indicate the m/z is the same as the suspected DAACP, but the identity and/or stereochemistry of the peak was not confirmed.

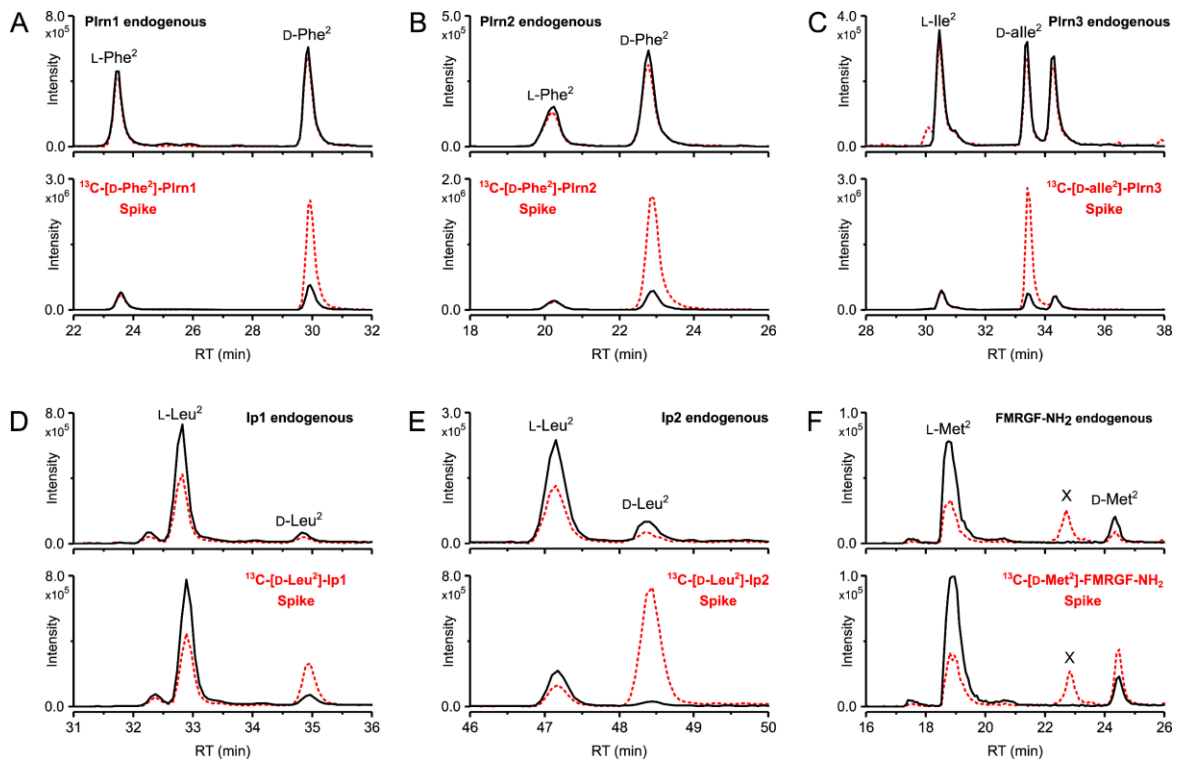


Figure 3.4. Confirmation of endogenous peptide stereochemistry identifications by LC-MS using ^{13}C -labeled synthetic D2 peptides spiked into *A. californica* cerebral ganglia extracts. For each plot, the black trace shows EICs for the monoisotopic m/z predicted for the endogenous peptide, while the dashed red trace shows the EIC for $(m+1)/z$, corresponding to the m/z of the ^{13}C -labeled synthetic standard. For each plot, the upper trace shows the EICs of the endogenous peptides in cerebral ganglia extracts spiked with water, and the lower trace shows the synthetic D2 standard spiked into cerebral ganglia extracts. (A) Plrn1, m/z 560.3 ± 0.1 , $z = 3+$ (black trace), $(m+1)/z$ 560.6 ± 0.1 , $z = 3+$ (red trace), (B) Plrn2, m/z 542.6 ± 0.1 , $z = 3+$ (black trace), $(m+1)/z$ 542.9 ± 0.1 , $z = 3+$ (red trace), (C) Plrn3, m/z 772.9 ± 0.1 , $z = 2+$ (black trace), $(m+1)/z$ 773.4 ± 0.1 , $z = 2+$ (red trace), (D) Ip1, m/z 552.3 ± 0.1 , $z = 2$ (black trace), $(m+1)/z$ 552.8 ± 0.1 , $z = 2+$ (red trace), (E) Ip2, m/z 1051.6 ± 0.1 , $z = 1+$ (black trace), $(m+1)/z$ 1052.6 ± 0.1 , $z = 1+$ (red trace), (F) FMRGF-NH₂, m/z 328.7 ± 0.1 , $z = 2+$ (black trace), $(m+1)/z$ 329.2 ± 0.1 , $z = 2+$ (red trace). For each peptide, the $(m+1)/z$ isotope is present in the endogenous sample due to the natural isotopic pattern of peptides in this mass range. However, when the ^{13}C -labeled synthetic standard is spiked into the extract, the increase in the $(m+1)/z$ is due to the presence of the isotope-labeled peptide co-eluting with the endogenous peptide. The peak labeled with an X in (F) was identified as an unrelated endogenous peptide sharing the same m/z as the synthetic standard.

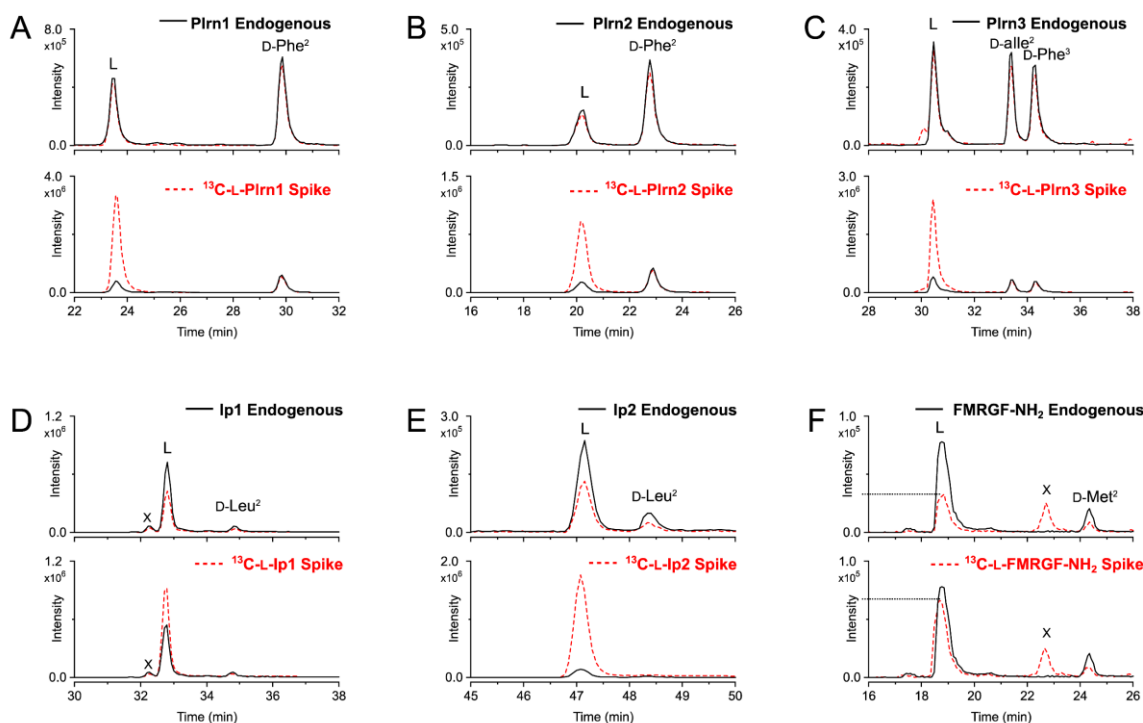


Figure 3.5. Spiking of ^{13}C -labeled synthetic all-L peptide standards into cerebral ganglia extracts. For each plot, the black trace shows the EICs for the monoisotopic m/z predicted for the endogenous peptide, while the dashed red trace shows the EIC for $(m+1)/z$, corresponding to the m/z of the ^{13}C -labeled synthetic standard. Each of the all-L peptide standards were labeled with one ^{13}C -Gly. The upper traces show the EICs of the endogenous peptides in cerebral ganglia extracts spiked with water, and the lower trace shows the synthetic all-L standard spiked into cerebral ganglia extracts. (A) Endogenous Plrn1, m/z 560.3 ± 0.1 , $z = 3+$ (black trace), $(m+1)/z$ 560.6 ± 0.1 , $z = 3+$ (dashed red trace); (B) endogenous Plrn2, m/z 542.6 ± 0.1 , $z = 3+$ (black trace), $(m+1)/z$ 542.9 ± 0.1 , $z = 3+$ (red trace); (C) Plrn3, m/z 772.9 ± 0.1 , $z = 2+$ (black trace), $(m+1)/z$ 773.4 ± 0.1 , $z = 2+$ (red trace); (D) Ip1, m/z 552.3 ± 0.1 , $z = 2+$ (black trace), $(m+1)/z$ 552.8 ± 0.1 , $z = 2+$ (red trace); (E) endogenous Ip2, m/z 1051.6 ± 0.1 , $z = 1+$ (black trace), $(m+1)/z$ 1052.6 ± 0.1 , $z = 1+$ (red trace); (F) endogenous FMRGF-NH₂, m/z 328.7 ± 0.1 , $z = 2+$ (black trace), $(m+1)/z$ 329.2 ± 0.1 , $z = 2+$ (red trace). For each peptide, the $(m+1)/z$ isotope is present in the endogenous sample due to the natural isotopic pattern of peptides in this mass range. However, when the ^{13}C synthetic standard is spiked into the extract, the increase in the $(m+1)/z$ is due to the presence of the isotope-labeled peptide co-eluting with the endogenous peptide. Note: endogenous peptide EICs in plots A–E (upper) are from Figure 1 for comparison purposes. The peaks labeled with an X in (E) and (D) indicate the peak was identified as an unrelated endogenous peptide.

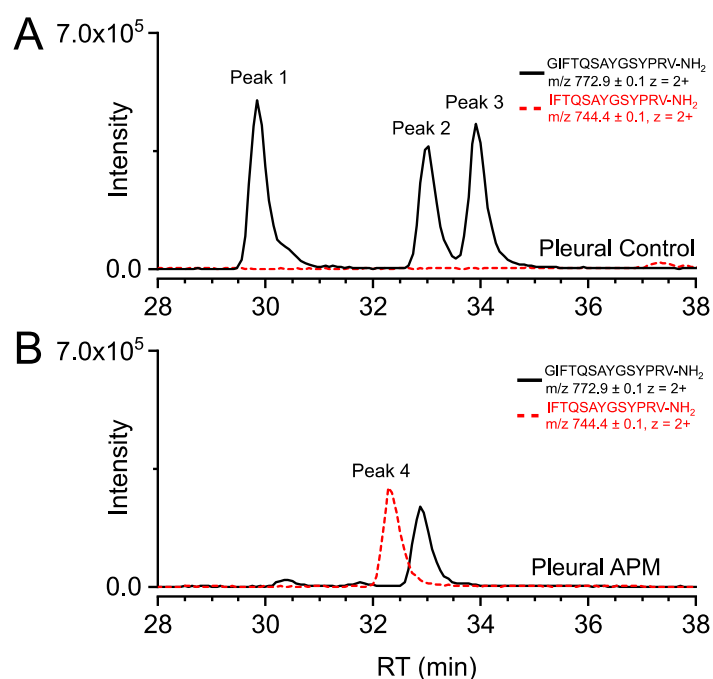


Figure 3.6. Evidence for [D-Phe³]-Pln3 by APM screening. Plot shows LC–MS EICs of Pln3 (black trace, m/z 772.9 $\pm$ 0.1, z = 2+) and Pln3(2-14) (dashed red trace, m/z 744.4, z = 2+) in (A) right pleural ganglia extracts incubated for 15 h at 37 °C without APM and (B) right pleural ganglia extracts incubated for 15 h with APM. Peak 1 (L-Pln3) and Peak 3 ([D-Phe³]-Pln3) were not present after incubation with APM, while Peak 2 ([D-aIle²]-Pln3) was present in the APM-treated extract. Peak 4 (Pln3(2-14)) was detected in APM-treated extracts. Pln3(2-14) is observed eluting at 32.5 minutes in the APM treated extract but not the control extract. The Gly¹ of [D-Phe³]-Pln3 was removed by APM and the truncated [D-Phe³]-Pln3(2-14) remained stable throughout the analysis. Pln3(2-14) was also observed in APM-treated extracts of cerebral, buccal, abdominal, and left pleural ganglia.

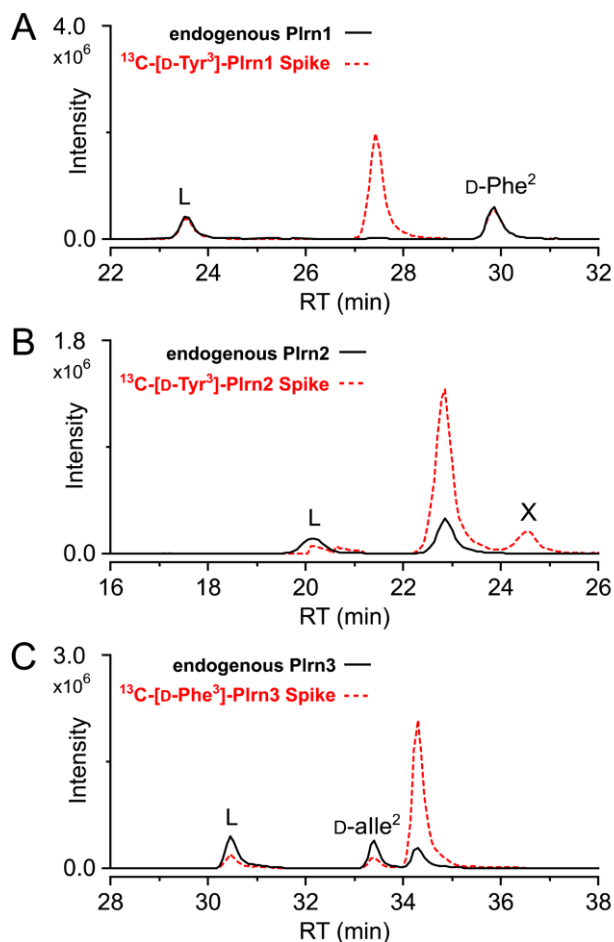


Figure 3.7. Testing for the presence of endogenous Plrn1, Plrn2, and Plrn3 D3-stereoisomers by spiking of ¹³C-labeled synthetic D3 forms of Plrn1, Plrn2, and Plrn3 peptides into *A. californica* cerebral ganglia extracts. Plots show LC-MS EICs of (A) endogenous Plrn1, m/z 560.3, $z = 3+$ (black trace) and synthetic ¹³C-[D-Tyr³]-Plrn1, $(m+1)/z$ 560.6, $z = 3+$ (dashed red trace), (B) endogenous Plrn2 m/z 542.6, $z = 3+$ (black trace) and ¹³C-[D-Tyr³]-Plrn2, $(m+2)/z$ 543.3, $z = 3+$ (dashed red trace), and (C) endogenous Plrn3, m/z 772.9 $\pm$ 0.1, $z = 2+$ (black trace), and ¹³C-[D-Phe³]-Plrn3, $(m+2)/z$ 773.9 $\pm$ 0.1, $z = 2+$ (dashed red trace). Shown in (B), ¹³C-[D-Tyr³]-Plrn2 co-eluted with the peak that was earlier identified as [D-Phe²]-Plrn2 based on spiking and APM resistance. The peak labeled with an “X” in (B) indicates an unrelated endogenous peptide that has the same m/z as the synthetic standard.

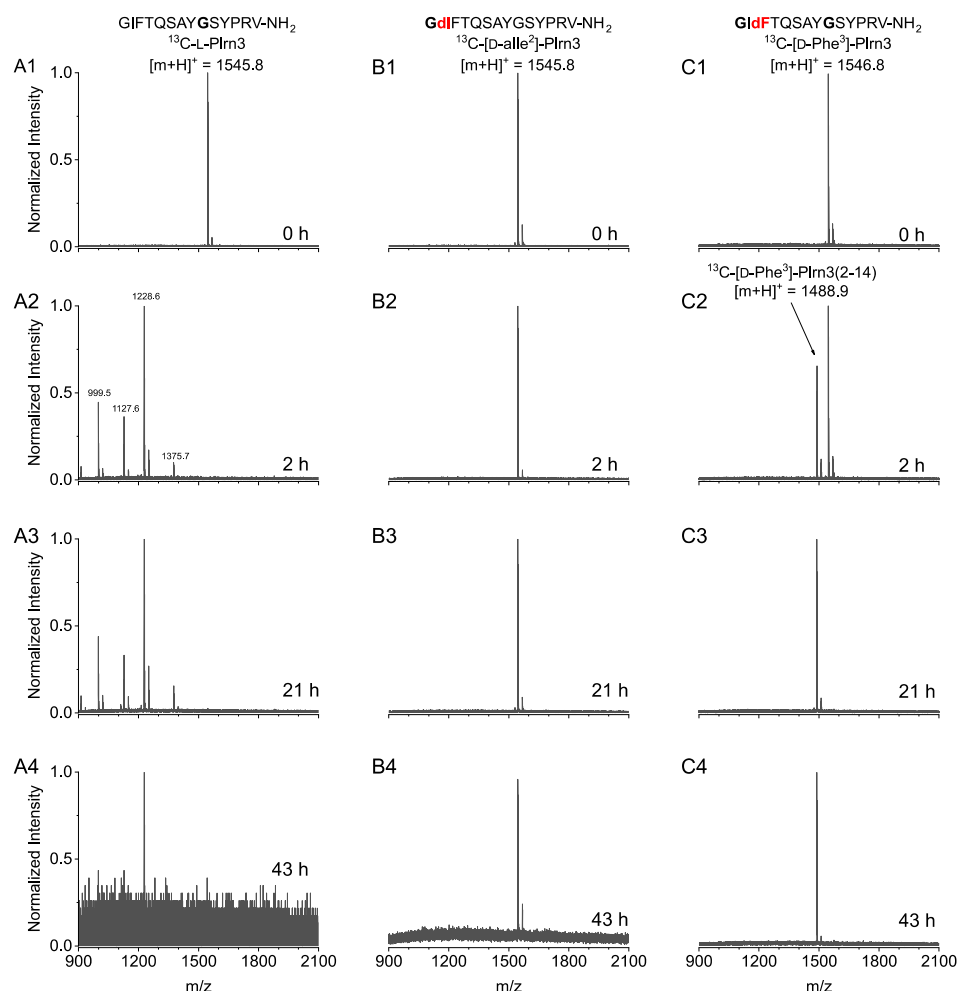


Figure 3.8. Time course showing the APM digestion of synthetic Pln3 peptide standards. Plots show MALDI MS spectra of ^{13}C -L-Pln3 (column A), ^{13}C -[D-allo²]-Pln3 (column B) and ^{13}C -[D-Phe³]-Pln3 (column C) collected at four time points of an APM digestion: before APM was added at 0 h (Row 1); incubation with APM for 2 h (Row 2), 21 h (Row 3), and 41 h (Row 4). The $[m+H]^+$ values of the truncated synthetic ^{13}C -L-Pln3 peptides are labeled in plot (A2) ^{13}C -L-Pln3(3-14) $[m+H]^+ = 1375.7$, ^{13}C -L-Pln3(4-14) $[m+H]^+ = 1228.6$, ^{13}C -L-Pln3(5-14) $[m+H]^+ = 1127.6$, ^{13}C -L-Pln3(6-14) $[m+H]^+ = 999.5$. The position of the ^{13}C -Gly is indicated in the peptide sequences by a bold “G”. The D-residues are denoted in red font.

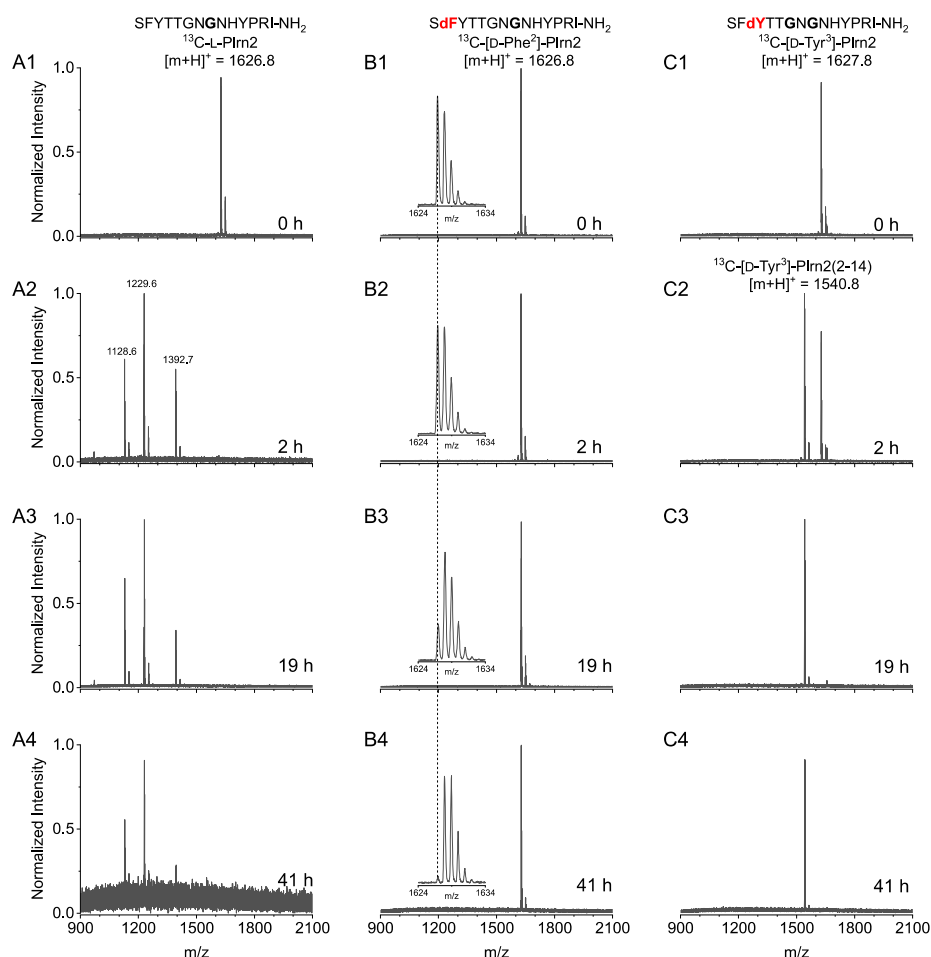


Figure 3.9. APM digestion time course on synthetic Plrn2 peptide standards. Plots show MALDI MS spectra of ^{13}C -L-Plrn2 (column A), ^{13}C -[D-Phe²]-Plrn2 (column B), and ^{13}C -[D-Tyr³]-Plrn2 (column C) collected at four time points of an APM digestion: before adding APM 0 h (Row 1); incubation with APM for 2h (Row 2), 19h (Row 3), and 41 h (Row 4). Note the inset spectra in Column B shows a shift by +1 in the isotope pattern for ^{13}C -[D-Phe²]-Plrn2 over the time course that resulted from deamidation of ^{13}C -[D-Phe²]-Plrn2 during extended incubation. In plot A2, the $[m+H]^+$ values of truncated ^{13}C -L-Plrn2 are labeled, ^{13}C -L-Plrn2(3-14) $[m+H]^+ = 1392.7$, ^{13}C -L-Plrn2(4-14) $[m+H]^+ = 1229.6$, and ^{13}C -L-Plrn2(5-14) $[m+H]^+ = 1128.6$. The position of the ^{13}C -Gly is indicated in the peptide sequences by a bold “G”. The D-residues are denoted in red font.

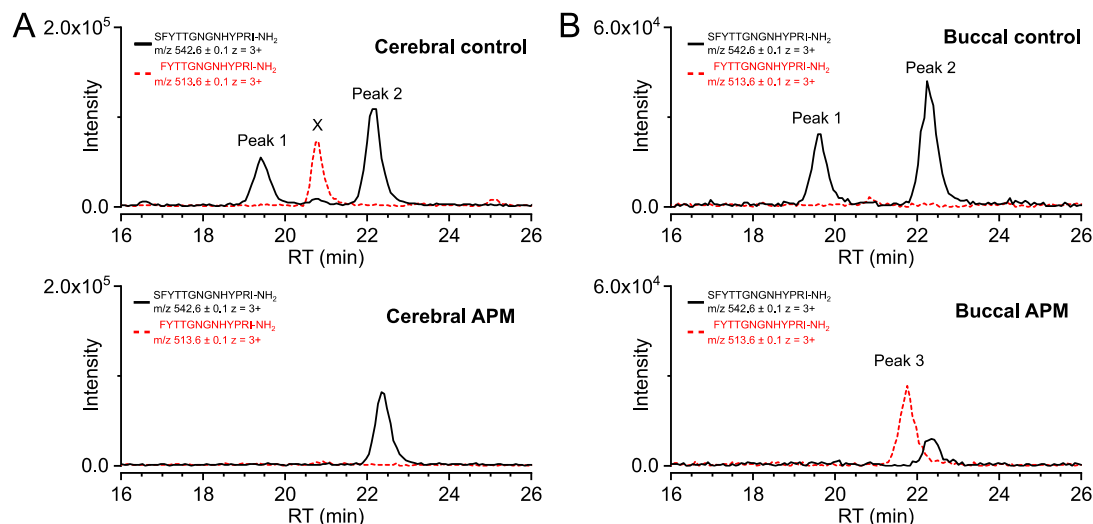


Figure 3.10. Evidence for [D-Tyr³]-Plrn2 in buccal ganglia extracts by partial APM digestion. Plots show LC–MS EICs of Plrn2 (black trace, $m/z\ 542.6 \pm 0.1$, $z = 3+$) and Plrn2(2-14) (dashed red trace, $m/z\ 513.6 \pm 0.1$, $z = 3+$) in (A) cerebral and (B) buccal ganglia extracts incubated for 15 h without APM added (upper) and incubated for 15 h with APM (lower). Peak 1 (L-Plrn2) was degraded by APM in both cerebral and buccal ganglia extracts. Peak 3 (Plrn2(2-14)) was observed in APM-treated buccal ganglia extracts but not in APM-treated cerebral ganglia extracts. Note, the EICs of Plrn2 cerebral control and cerebral APM are from Figure 1B in the main text and plotted here for comparison purposes. The EIC of Plrn2 shown in the buccal ganglia control is from Figure 3.3B and plotted here for comparison purposes. The X over the Peak at 21 min in (A) indicates the peak is not related to Plrn2(2-14) as determined by MS/MS.

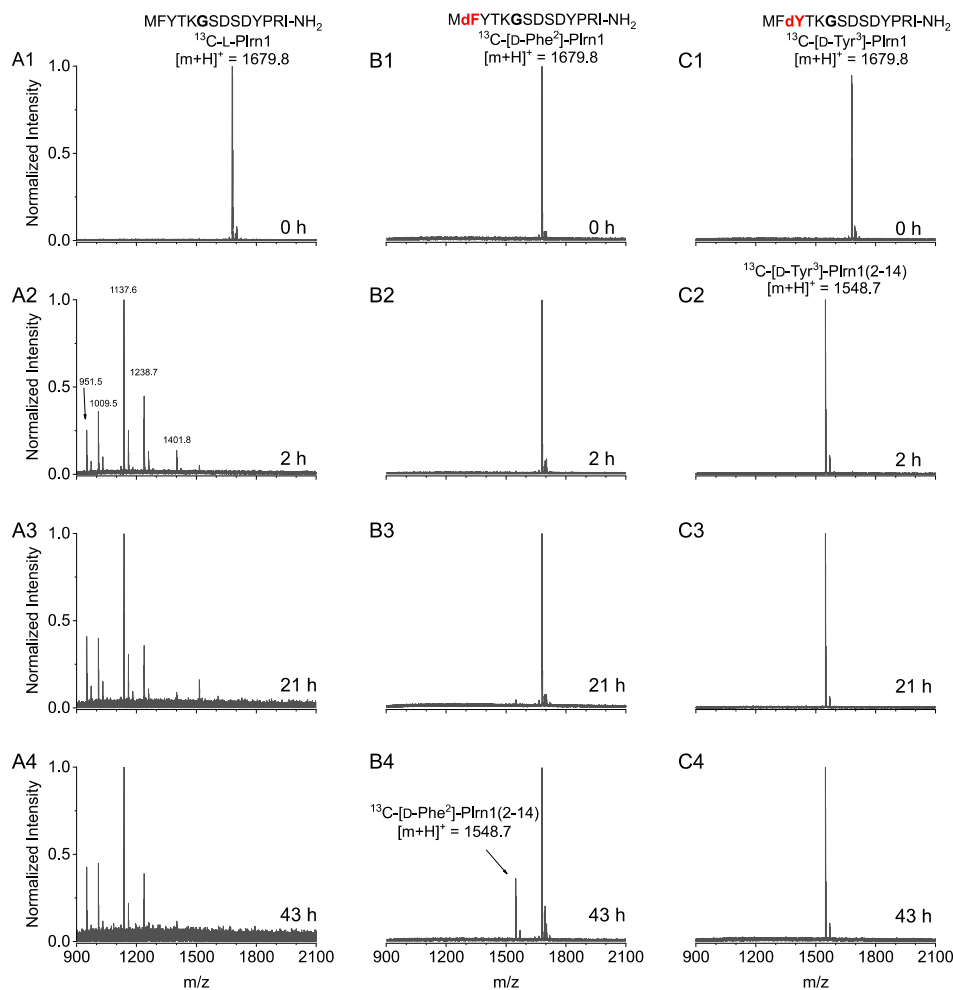


Figure 3.11. Time course showing the APM digestion of synthetic Plrn1 peptide standards. Plots show MALDI MS spectra of ^{13}C -L-Plrn1 (column A), ^{13}C -[D-Phe²]-Plrn1 (column B), and ^{13}C -[D-Tyr³]-Plrn1 (column C) collected at four time points of an APM digestion: before APM was added at 0 h (Row 1); incubation with APM for 2 h (Row 2), 21 h (Row 3), and 41 h (Row 4). Note, ^{13}C -[D-Tyr³]-Plrn1(2-14) begins to appear after 2 h (C2), while ^{13}C -[D-Phe²]-Plrn1(2-14) is observed after 43 h (B4). The $[m+H]^+$ values of the truncated synthetic ^{13}C -L-Plrn1 sequences labeled in plot (A2) are ^{13}C -L-Plrn1(3-14) $[m+H]^+ = 1401.8$, ^{13}C -L-Plrn1(4-14) $[m+H]^+ = 1238.7$, ^{13}C -L-Plrn1(5-14) $[m+H]^+ = 1137.6$, ^{13}C -L-Plrn1(6-14) $[m+H]^+ = 1009.5$, ^{13}C -L-Plrn1(7-14) $[m+H]^+ = 951.5$. The position of the ^{13}C -Gly is indicated in the peptide sequences by a bold “G”. The D-residues are denoted in red font.

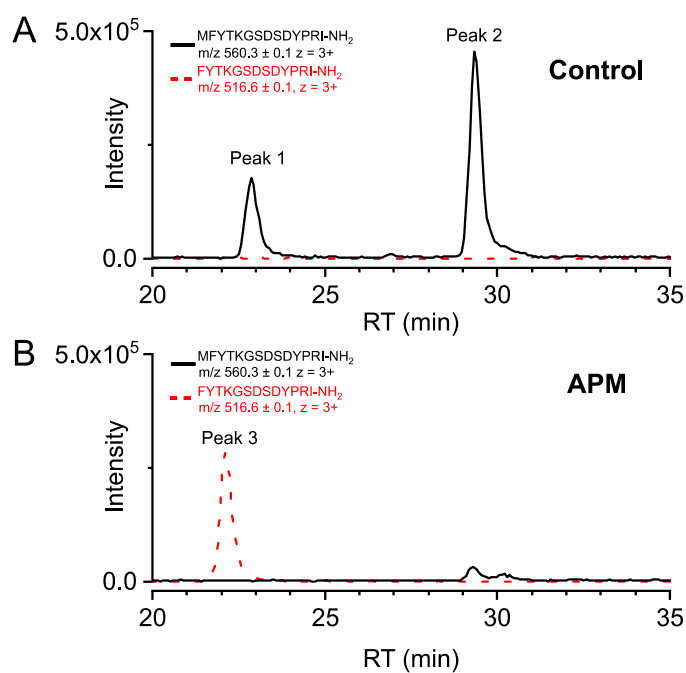


Figure 3.12. Partial degradation of [D-Phe²]-Pln1 observed during APM screening experiments. Plots show EICs of Pln1 (solid black trace, $m/z 560.3$, $z = 3+$) and Pln1(2-14) (dashed red trace, $m/z 516.6$, $z = 3+$) detected in extracts from pleural ganglia incubated for 15 h (A) without APM and (B) with APM. Both L-Pln1 (Peak 1) and [D-Phe²]-Pln1 (Peak 2) appear to degrade after APM treatment. In the APM-treated sample, Pln1(2-14) (Peak 3) was observed. The stereochemistry of Peaks 1 and 2 were determined by spiking.

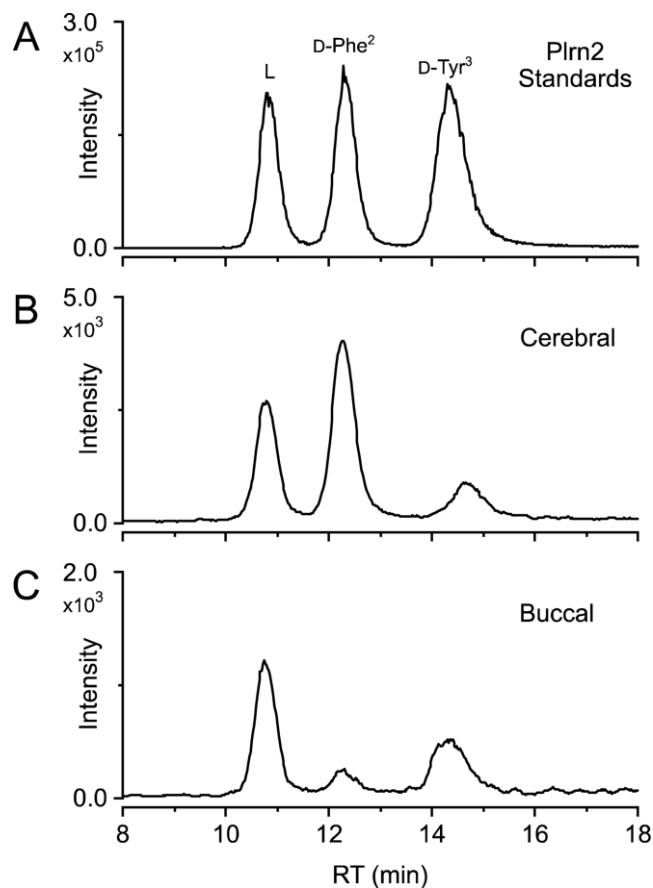


Figure 3.13. Separation of endogenous Plrn2 peptides using chiral LC–MRM MS. Chromatograms show (A) ^{13}C -labeled synthetic standards, ^{13}C -L-Plrn2 (10.7 min), ^{13}C -[D-Phe²]-Plrn2 (12.3 min), and ^{13}C -[D-Tyr³]-Plrn2 (14.4 min), (B) endogenous Plrn2 stereoisomers in cerebral ganglia extracts, and (C) buccal ganglia extracts. For ^{13}C synthetic Plrn2 stereoisomers, the channel monitored parent m/z 814.4 and fragment m/z 547.1. For the endogenous Plrn2, the channel monitored parent m/z 813.4 and fragment m/z 547.1.

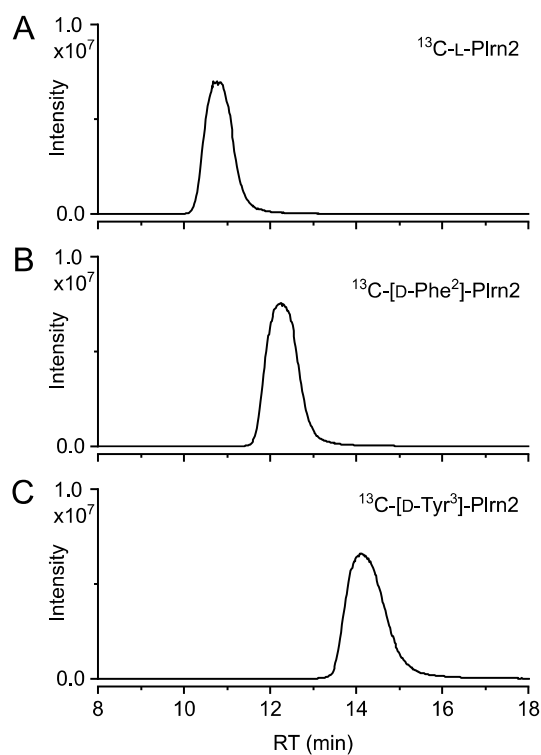


Figure 3.14. Elution order of ^{13}C -L-Plrn2, ^{13}C -[D-Phe²]-Plrn2, and ^{13}C -[D-Tyr³]-Plrn2 by chiral chromatography. Plots show LC MRM chromatograms of (A) ^{13}C -L-Plrn2, parent ion m/z 813.9, and fragment ion m/z 547.1; (B) ^{13}C -[D-Phe²]-Plrn2 parent ion m/z 813.9, and fragment ion m/z 547.1; (C) ^{13}C -[D-Tyr³]-Plrn2 parent ion m/z 814.4, and fragment ion m/z 547.1.

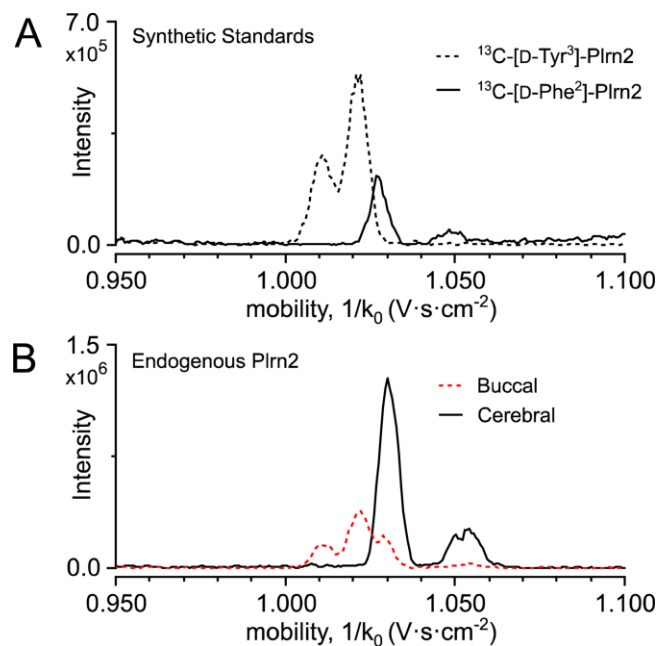


Figure 3.15. Evidence for [D-Tyr³]-Plrn2 in buccal ganglia extracts using TIMS. (A) EIMs of synthetic Plrn2 standards, ^{13}C -[D-Phe²]-Plrn2 (m/z 813.9, $z = 2+$, solid black trace) and ^{13}C -[D-Tyr³]-Plrn2 (m/z 814.4, $z = 2+$, dashed black trace). (B) EIMs of endogenous Plrn2 (m/z 813.4, $z = 2+$) from cerebral ganglia extracts (solid black trace) and buccal ganglia extracts (dashed red trace).

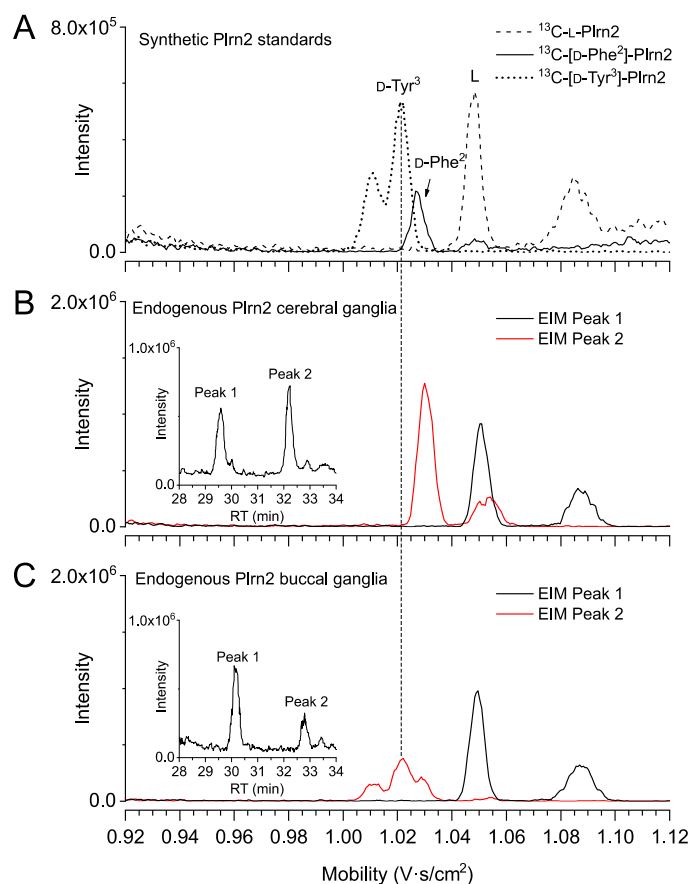


Figure 3.16. (A) Extracted ion mobilograms (EIMs) of synthetic Plrn2 peptide standards: ^{13}C -L-Plrn2 (dashed trace, $(m+1)/z$ 813.9 ± 0.05 , $z = 2+$), ^{13}C -[D-Phe²]-Plrn2 (solid trace, $(m+1)/z$ 813.9 ± 0.05 , $z = 2+$) and ^{13}C -[D-Tyr³]-Plrn2 (dotted trace, $(m+2)/z$ 814.4 ± 0.05 , $z = 2+$). (B) EIMs of endogenous Plrn2 (m/z 813.4 ± 0.05 , $z = 2+$) in cerebral ganglia extracts. (C) EIMs of endogenous Plrn2 (m/z 813.4 ± 0.05 , $z = 2+$) in buccal ganglia extracts. Inset in B and C show the EICs of Plrn2 (m/z 813.4 ± 0.1 , $z = 2+$). The EIMs were measured by averaging the mobility over Peak 1 (L-Plrn2) and Peak 2 (potentially coeluting [D-Phe²]-Plrn2/[D-Tyr³]-Plrn2) shown in the inset chromatograms. The shift in retention times of Peak 1 and Peak 2 between buccal and cerebral ganglia is likely a chromatographic artifact caused by differences in sample matrixes. Note, all traces were smoothed using a 7-point adjacent averaging filter; EIMs of ^{13}C -[D-Phe²]-Plrn2, ^{13}C -[D-Tyr³]-Plrn2, and [D-Phe²]-Plrn2, [D-Tyr³]-Plrn2 are from Figure 5 in the main text and plotted here for comparison purposes.

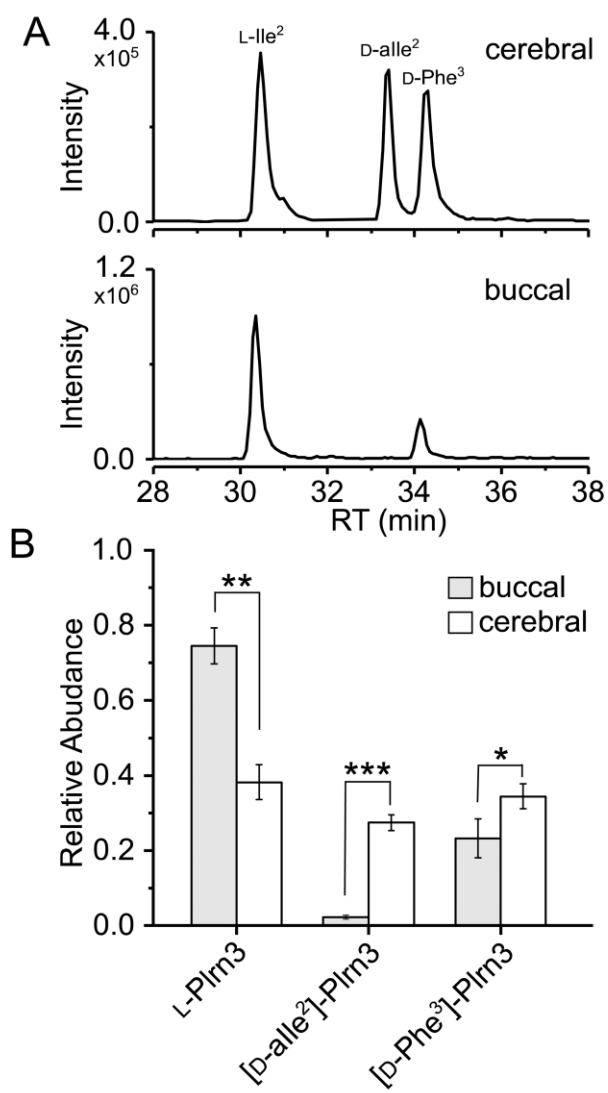


Figure 3.17. Relative abundance comparison of Pln3 between buccal and cerebral ganglia. (A) LC–MS EICs of Pln3 (m/z 772.9 $\pm$ 0.1, z = 2+) from cerebral ganglia extracts (upper) and buccal ganglia extracts (lower). The chromatogram from the cerebral ganglia is from Figure 2C for comparison purposes. (B) Quantitative comparison of the relative abundance of Pln3 stereoisomers in the buccal and cerebral ganglia measured by LC–MS. The bars represent averaged relative abundances $\pm$ standard deviation from N=4 biological sets. *** p = 3.8×10^{-7} , ** p = 3.6×10^{-5} , * p = 0.01. Significance was tested using a two-sample t-test.

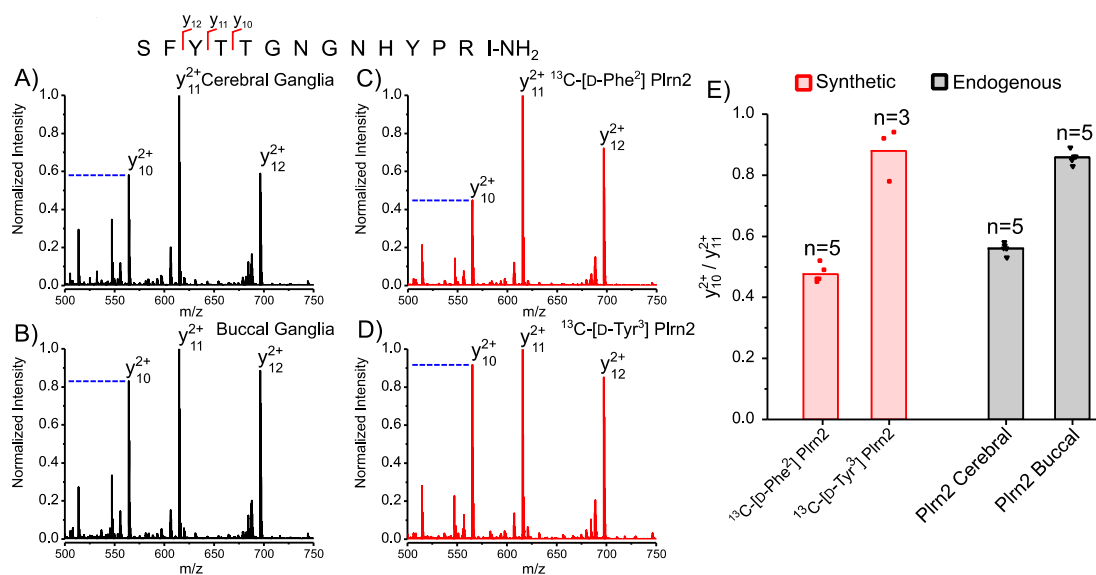


Figure 3.18. Evidence for [D-Tyr³]-Pln2 in buccal ganglia extracts by MS/MS. (A–D) Representative MS/MS spectra collected over co-eluting [D-Phe²]-Pln2 and [D-Tyr³]-Pln2 (m/z 542.6, $z = 3+$) from (A) cerebral ganglia extracts; (B) buccal ganglia extracts; (C) synthetic ¹³C-[D-Phe²]-Pln2; (D) synthetic ¹³C-[D-Tyr³]-Pln2. (E) Plot comparing the average endogenous Pln2 y₁₀²⁺/y₁₁²⁺ ratio of the second eluting Pln2 peak from buccal and cerebral ganglia extracts, to the y₁₀²⁺/y₁₁²⁺ ratio of the synthetic ¹³C labeled D2 and D3 standards. Endogenous [D-Phe²]-Pln2 and [D-Tyr³]-Pln2 were found to have different y₁₀²⁺/y₁₁²⁺ ion ratios in the different ganglia. The y₁₀²⁺/y₁₁²⁺ ratio of the second eluting Pln2 peak measured in buccal ganglia extracts was closer to the ratio observed for ¹³C-[D-Tyr³]-Pln2. In contrast, the y₁₀²⁺/y₁₁²⁺ ratio of the second eluting Pln2 peak measured in cerebral ganglia extracts was closer to the ratio measured for ¹³C-[D-Phe²]-Pln2.

>Aplysia Isomerized peptides precursor (XP_012946424.1)

MFAGSDVTLSTAGFRSWGLVYLTLTAVTA VSLNEHTTGPVSGRHTDGVNKLSSKVIQGEVTPAGQGVREQKDK
RFIESTAGDILYRRAGGGGHSSSGSNLGPVVTHAIAPPLGPGVGFQPPQQPHNDAGAGRGPSTGQKFTKFNDSPF
ERWVEWREEQANKRFLEAAAAERQGGGQIDGQDGGLDGGQDGGGEIVTQDGGEEKR FLESIAGHII RREEGGDVPYG
QDGGREKR FIESIAGHII RRGEGGASGDEGASGGEGADRTQEKR YLDGIASSLI KRQDVLGSDVAIGSTGNSGPDK
R YLDHLGSSLV KKS DGQDKR YLDHLGSSLV KKS DGQDKR YLDHLGSSLV KKWDEGDQEKR YLDHLGSSLV RKSDSQ
DKR YLDHLGSSLV KKS DGQEKR YLDHLGSSLV KKS DGIDQEKR YLDHLGSSLV KKS DTEEKR YLDHLGSSLV KKS D
GQEKR YLDHLGSSLV KKS DGIDQEKR YLDHLGSSLV KKS DTEEKR YLDHLGSSLV KKS DTEEKR YLDHLGSSLV KKS
SDTEEKR YLDHLGSSLV KKS DTEEKR YLDHLGSSLV KKS DTEEKR YLDHLGSSLV KKS DSQEKR YLDHLGSSLV KKS
SDREEKR YLDHLGSSLV KKS ESPELARFEQRMFTDYNNDLVAMEKDFLEKNRELQQSSPDVTSQGNKR YLDRISS
LV KKEDDEDETEVDGVLGFLIDLEKDEDENDLRQMAGTRAGNHKR YLDRVGSSLI KKQAQAQKR YLDGIASSLI KK
EDRGQKR YLDGIASSLI KKEDRGQKR YLDGIASSLI KKEDRGQKR YLDGIASSLI KKEDGFVGDI PNGQEKR YLDG
IASSLI KKKNELKSEKR YLDGIASSLI KKDNVQEEKR YLDGIASSLI KKEVRPGQKR YLDGIASSLI KKDSTNENKR
YLDGIASSLI KRQDSVSDEEKRYLDSIGSSLV RKSSSTTPYKR YLDQVGSSLV KKGASEMGPEVIPEAVYVGEIGI
SAGGEGPVYVTIDGQLPILAMSETDLPASPRKMYEVTDDVMTGSGHQQLVDDMAAAIHVS AKWAGLTGTSPSRQT
HPRI RRDTPDTPTSHSAHPLTSSPAHSLTSSPAHLRQKR FIESIAGNILY KRTRPELSALHQALVSDVTS PGPAQ
REAMSEPGFVDSPAGPGDDQSGGNT HIGDNQFVTSTTEFDFFPGTTHFSVSPA HFSVPAYFLDSPTHHQTRPLF
ETFSPPEVMGQGHVFQKR YLD DVASSLLH KR YLD TVASSLF KR LSNL PARPVGLGSGVPIH KR FTTESHRPLKAGN
NRHMSSI KRYIDSIASDLIYRKRPTLDLHRLRAMSPFQMANGELLTRPRDLAGSRERYLENNRGRGTGHVSTVWSG
RPTAMDAVRASGVAPGRESSKAYLREALREYVRGVATEVLDRTKVLGQSLTDGGQRGEESEEQAGSQGEGSVEGA
GSPSGWTNEKPGWGNEAYYSQGEETRDFFDLSLAGGILT

Figure 3.19. Amino acid sequence of *Aplysia* isomerized peptides precursor (AIPP). Sequences detected by LC–MS are highlighted in yellow and underlined. The signal peptide sequence predicted by SignalP⁹ is highlighted in gray.

>Aplysia FMRGFamide Precursor (XP_005094941.1)

MRVARLSGRMSGWSCACVLLVALMVAVVASEVSGADQQATNESPSPTGAEDKDLLIPSKAGKDNAVTSPE
EDTGSSADVQSRELQTSDDSSASKMAAGEATPGGSDVEDKDTETQPEDDNETDNEDSNSNSAQQQQPE
VDKSSSSDAGELSGDLVGKRDVSGLVDDKRFLRGFNQFARRSGVDAVPSDVDKRFMRGFGKRFMRGFRSD
NDGEAEKRFRMRGLRDSLSASSPNDESDFWKWDSYLNVPNHMDNLLPLDFSTPDSGPGADKRFMRGFRNY
VTSGEDEKRFMRGFHKKDPILAGDISDVLAEDKRFLRGLNNYAKKDSLDDVDGISDSADKHFMPGLIKNS
SSSSSDNVETSPDKRFIHGLQLYSRSYGGLGDSPVEEKRFMRGFGKKNEVYSDEEDLQTVADKRFMRGFN
KRFMRGFGGKRFMRGFNKRFIMGLNNYADKRFMRGFQKKEDAFSESDDEIVDDKRFMRGFGKRELVEGND
DEKRFMRGFGKRDLIGDDTDEKRFMRGFGKRELVESLGGEDKRFMRGFGKREPVDDVTDDKRFMRGFGKR
EPVDDVTDDKRFMRGFGKREPVEDVTDDKRFMRGFGKREPVDDVTDDKRFMRGFGKREPVDDVTDDKRFM
RGFGKREPVDDVTDDKRFMRGFGKREPVDDVTDDKRFMRGFGKREPIDDVTDDKRFMRGFGKRLPTEDDT
EEKRFMRGFGKRLPTAETENDTEEKRFMRGFGKRLPTEDDTEEKRFMRGFGKRLPTAETENDTEEKRFMR
GFGKRLPTEDDTDEKRFMRGFGKRLPTAETENDTEEKRFMRGFGKRLPTEDDTDEKRFMRGFQGKRFMRG
FQKKADEINELEKKFMRGFNKRSAGEHQRTKREVRVYADVDEGDLVEKLLQDLFGEPKSEKTELWYDEPI
IYPGEVNNEKRFMRGFSNYAKRGGSSLGANFEVGDDIVNDGPVNKRFMRGFQNFVNMAARRRYSYPREGDS
QDKRFMRGLSAYKRGAGGSLDEDISKRFIKGLERYALPASKRFMRGLQNWQKRNSLEGSGADKRFIRGL
HLYHKGKGGDMSSVDKRFMRGLNLYAKRYLGWDIPDSKNRFPASSVSVEPVLEVSQEDGPPSSVWEGEYDQP
KRDYMF⁹TSSLGNSSPKSGSRQFMSWRGGDYVLG

Figure 3.20. Amino acid sequence of one predicted isoform of the *Aplysia* FMRGFamide precursor (XP_005094941.1). Sequences detected by LC–MS/MS are highlighted in yellow and underlined. The sequences highlighted in orange indicate the m/z of the peptide was detected and the sequence is supported by MS/MS but was not picked up by a database search. The signal peptide sequence predicted by SignalP⁹ is highlighted in gray.

```

>NP_001191654.1 pleurin precursor [Aplysia californica]

MQSHVSSSLALCVLTGSACAMFYTKGSDSDYPRIGRRSFYTTGNGNHYPRIGRRDSPGAAAGTDILYPN
      Signal Peptide          Plrn1          Plrn2          cPlrn

TASGFSDFSAGAPAVKRGIFTQSAYGSYPRVGRRSQAGDDGSQEGLTARSKIFQKITGLDGPVEEELGD
                        Plrn3

GSAERLEADNQIGLELMFLAFDLGDQALSKSEFQSGMEKFRRHNPLC
                        EF hand domain

```

Figure 3.21. Amino acid sequence of the pleurin precursor. Plrn1, Plrn2, and Plrn3 are highlighted in blue and underlined. The peptide that is encoded between Plrn2 and Plrn3 is labeled cPlrn and is highlighted in pale purple and underlined. Signal peptide predicted by SignalP⁴⁰ is highlighted in gray. cPlrn was detected in this study but no experiments were done to determine if cPlrn is DAACP. A conserved domain (EF hand domain) was identified in a BLAST search of the prohormone and is shown highlighted in yellow.

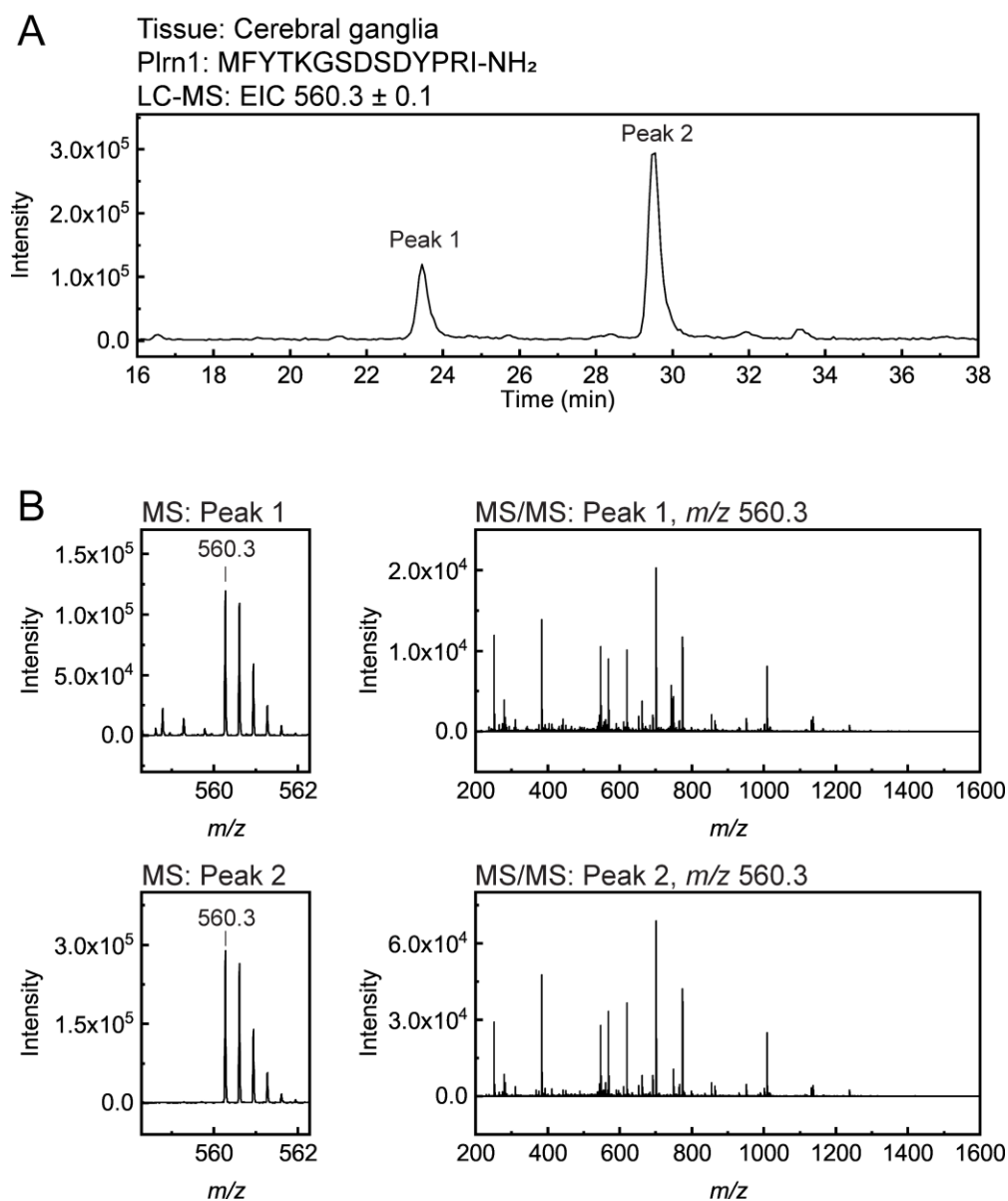


Figure 3.22. (A) EIC of endogenous Plrn1 (m/z 560.3 ± 0.1, $z = 3$) detected in cerebral ganglia extracts; Peak 1 (L-Plrn1) and Peak 2 ([D-Phe²]-Plrn1). (B) MS1 (left) and MS2 by CID (right) of Peak 1 and Peak 2 of the endogenous Plrn1 stereoisomers. The two peaks have the same MS/MS fragmentation patterns, indicating they have the same sequence.

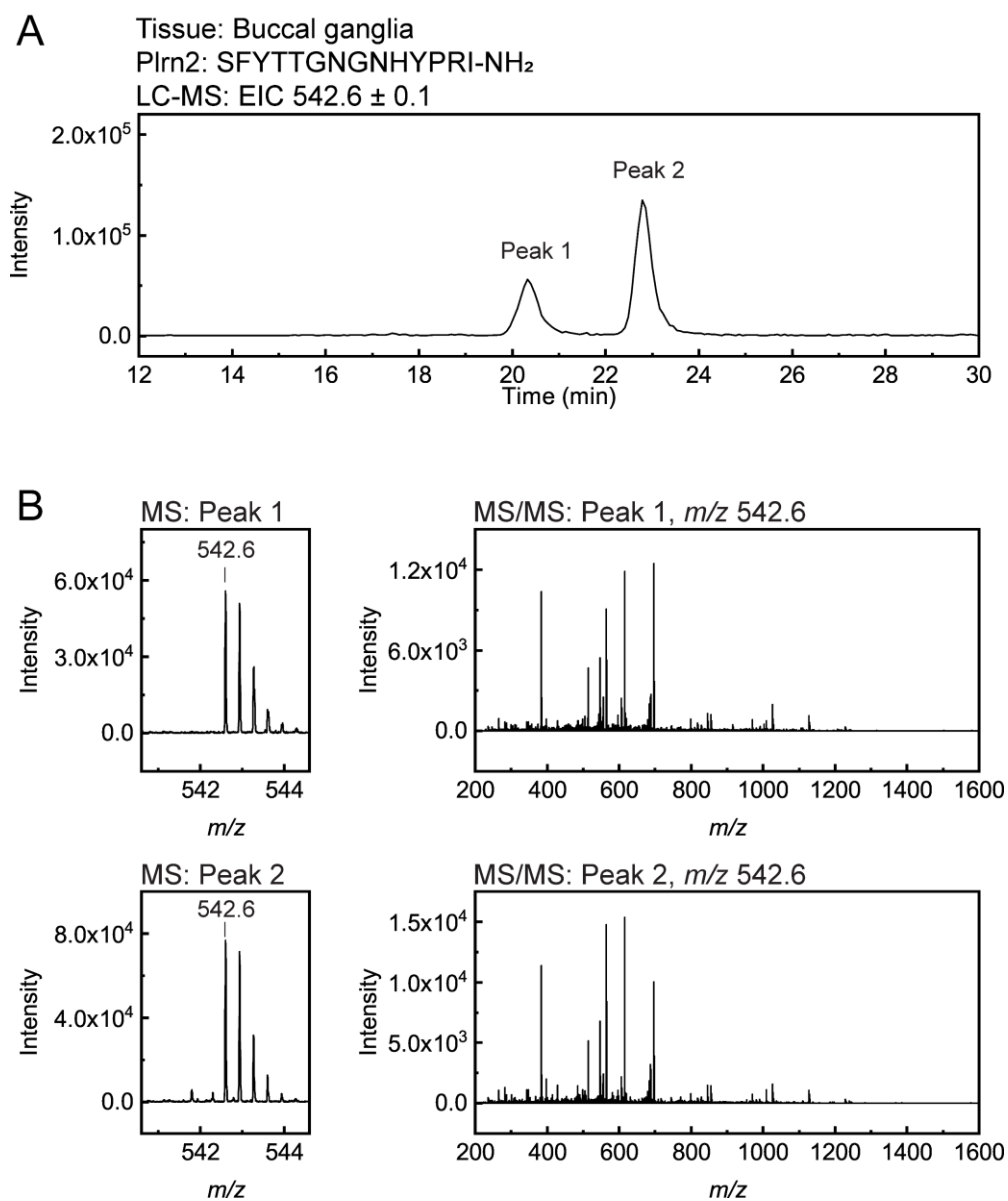


Figure 3.23. (A) EIC of endogenous Plrn2 (m/z 542.6 ± 0.1, $z = 3+$) detected in cerebral ganglia extracts. Peak 1 was identified as L-Plrn2 and Peak 2 was identified as co-eluting [D-Tyr³]-Plrn2 and [D-Phe²]-Plrn2 (B) MS1 (left) and MS2 by CID (right) of Peak 1 and Peak 2 of the endogenous Plrn2 stereoisomers. The two peaks have the same MS/MS fragmentation patterns, indicating they have the same sequence.

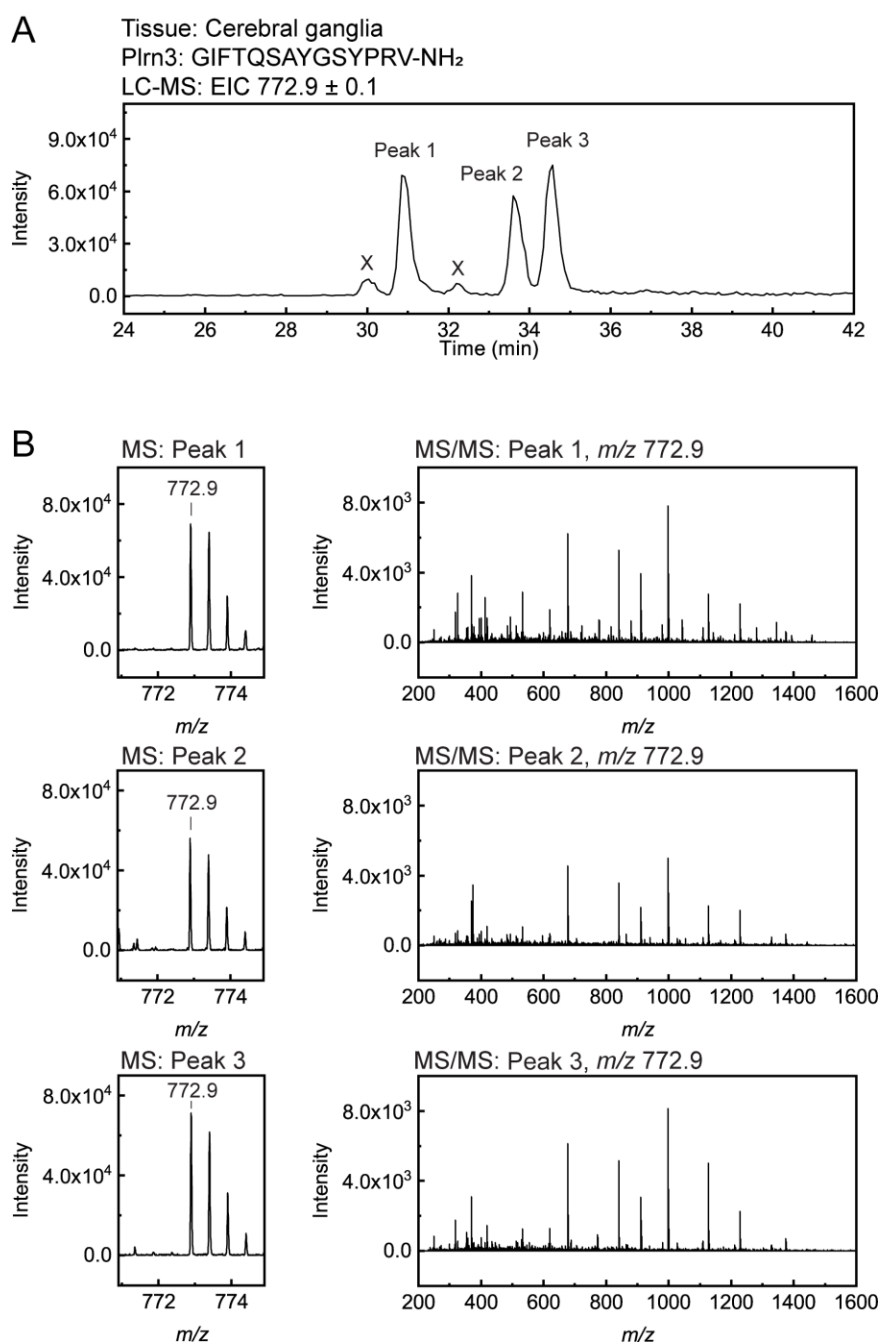


Figure 3.24. (A) EIC of Plrn3 (m/z 772.9 ± 0.1, $z = 2+$) in extracts from cerebral ganglia; Peak 1 (L-Plrn3), Peak 2 ([D-allo²]-Plrn3), and Peak 3 ([D-Phe³]-Plrn3). (B) MS1 (left) and MS2 by CID (right) collected over Peak 1, Peak 2, and Peak 3. The three peaks have the same MS/MS fragmentation patterns, indicating they have the same sequence. The X's in (A) indicate a peak that exhibited a different MS1 and/or MS2 from the sequence of interest.

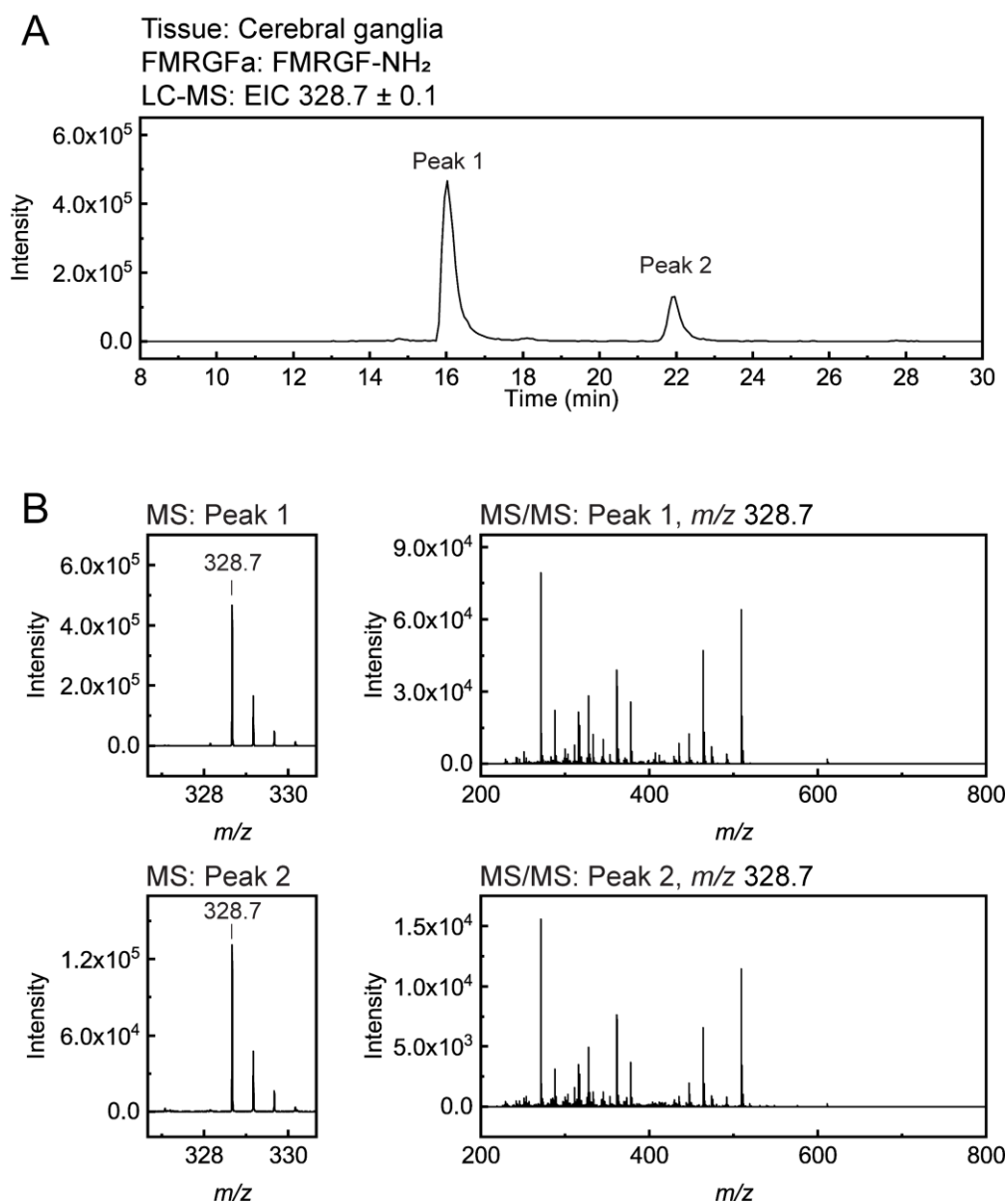


Figure 3.25. (A) EIC of endogenous FMRGF-NH₂ (m/z 328.7 ± 0.1, z = 2) detected in cerebral ganglia extracts; Peak 1 (L-FMRGF-NH₂) and Peak 2 ([D-Met²]-FMRGF-NH₂). (B) MS1 (left) and MS2 by CID (right) collected over Peak 1 and Peak 2. The two peaks have the same MS/MS fragmentation patterns, indicating they have the same sequence.

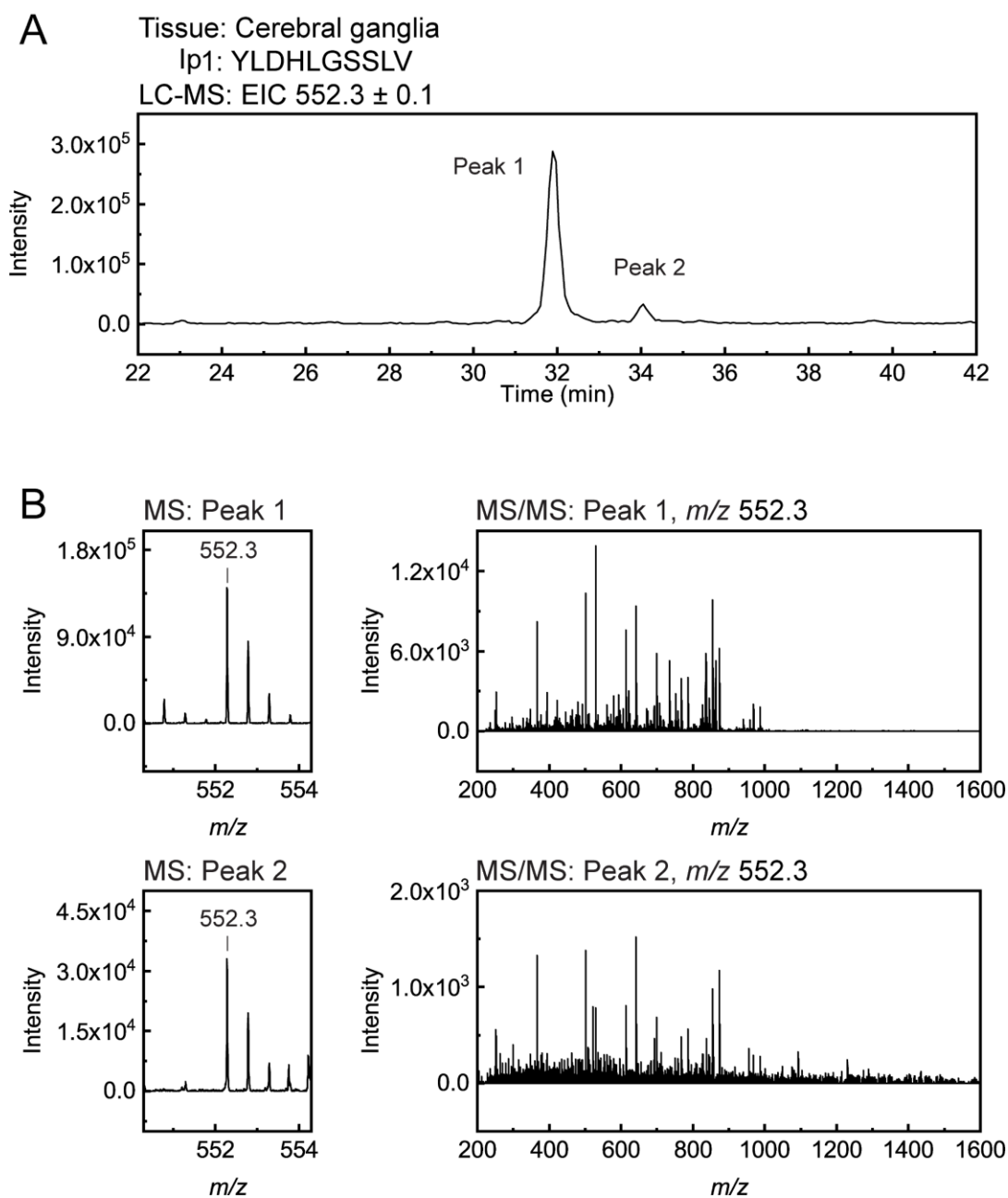


Figure 3.26. (A) EIC of endogenous Ip1 (m/z 552.3 $\pm$ 0.1, z = 2) detected in cerebral ganglia extracts; Peak 1 (L-Ip1) and Peak 2 ([D-Leu²]-Ip1). (B) MS1 (left) and MS2 by CID (right) collected over Peak 1 and Peak 2. The two peaks have the same MS/MS fragmentation patterns, indicating they have the same sequence.

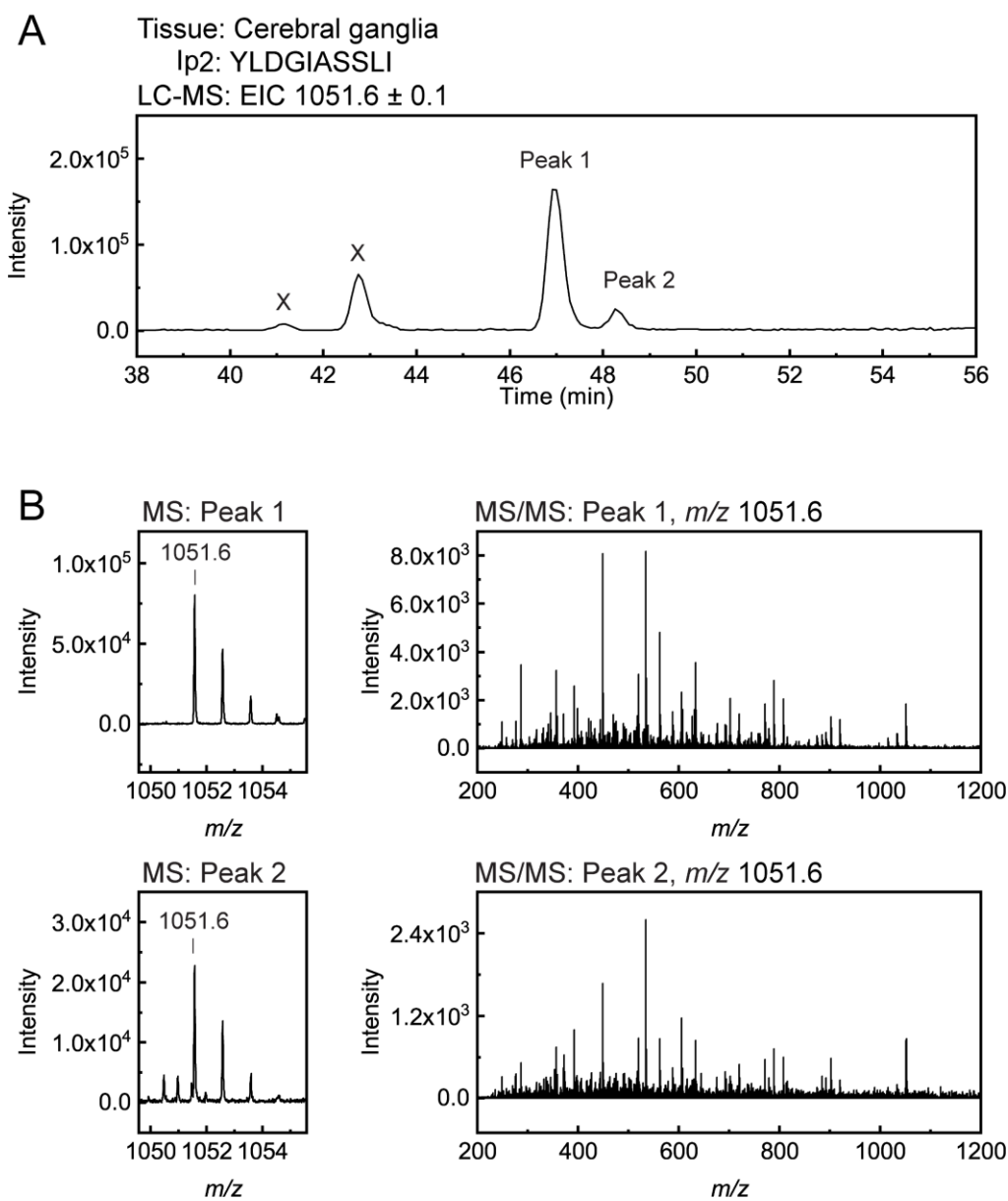


Figure 3.27. (A) EIC of endogenous Ip2 (m/z 1051.6 $\pm$ 0.1, $z = 2+$) detected in cerebral ganglia extracts; Peak 1 (L-Ip2) and Peak 2 ([D-Leu²]-Ip2). (B) MS1 (left) and MS2 by CID (right) collected over Peak 1 and Peak 2. The two peaks have the same MS/MS fragmentation patterns, indicating they have the same sequence. The X's in (A) indicate a peak that exhibited a different MS1 and/or MS2 from the sequence of interest.

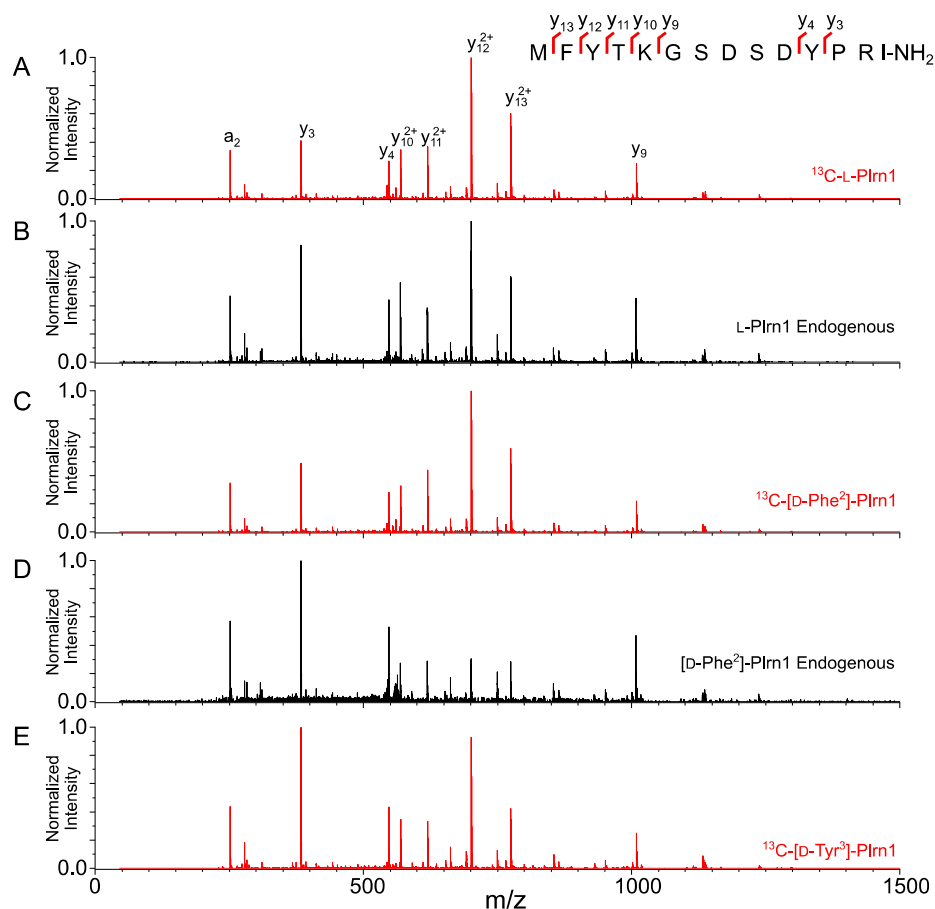


Figure 3.28. Product ion spectra of endogenous Plrn1 (black traces) and synthetic Plrn1 standards (red traces). CID collision energy was 21.8 eV. Product ion spectra are taken of the 3+ charge state of (A) synthetic ^{13}C -L-Plrn1 ($(m+1)/z$ 560.6); (B) endogenous L-Plrn1 (m/z 560.3) from no-APM treated pleural ganglia extracts; (C) synthetic ^{13}C -[D-Phe²]-Plrn1 ($(m+1)/z$ 560.6); (D) endogenous [D-Phe²]-Plrn1 from APM-treated pleural ganglia extracts; (E) synthetic ^{13}C -[D-Tyr³]-Plrn1 ($(m+1)/z$ 560.6). The most abundant fragment ions are labeled.

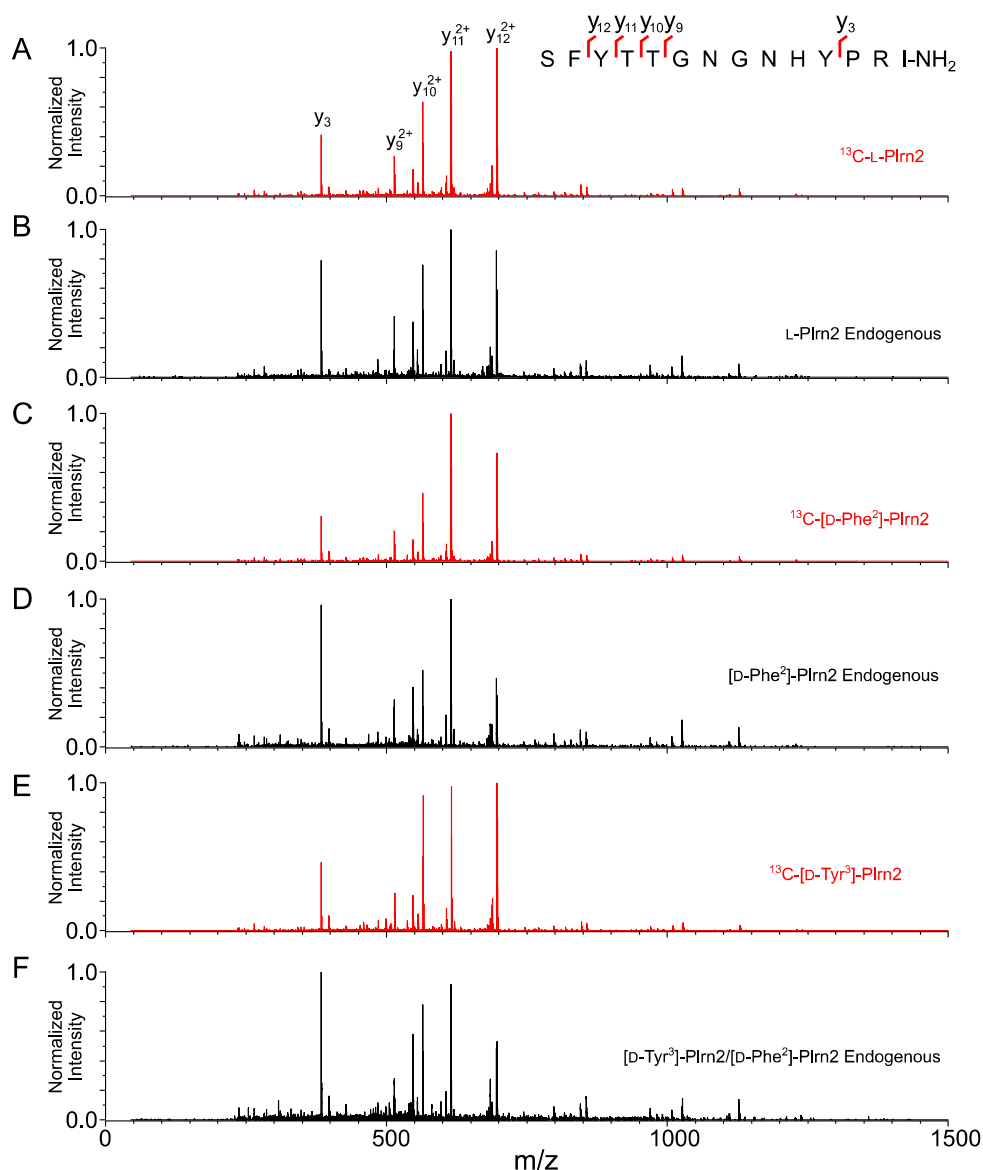


Figure 3.29. Product ion spectra comparison of endogenous Plrn2 (black traces) and synthetic Plrn2 standards (red traces). CID collision energy was 21.3 eV. Product ion spectra are taken of the 3+ charge state of (A) synthetic ¹³C-L-Plrn2 ((m+1)/z 542.9); (B) endogenous L-Plrn2 (m/z 542.6) from no-APM-treated cerebral ganglia extracts; (C) synthetic ¹³C-[D-Phe²]-Plrn2 ((m+1)/z 542.9); (D) endogenous [D-Phe²]-Plrn2, detected in the APM-treated cerebral ganglia extracts; (E) and synthetic ¹³C-[D-Tyr³]-Plrn2 ((m+2)/z 543.3). (F) Co-eluting endogenous [D-Phe²]-Plrn2/[D-Tyr³]-Plrn2 taken from no-APM-treated buccal ganglia extracts. The most abundant fragment ions are labeled.

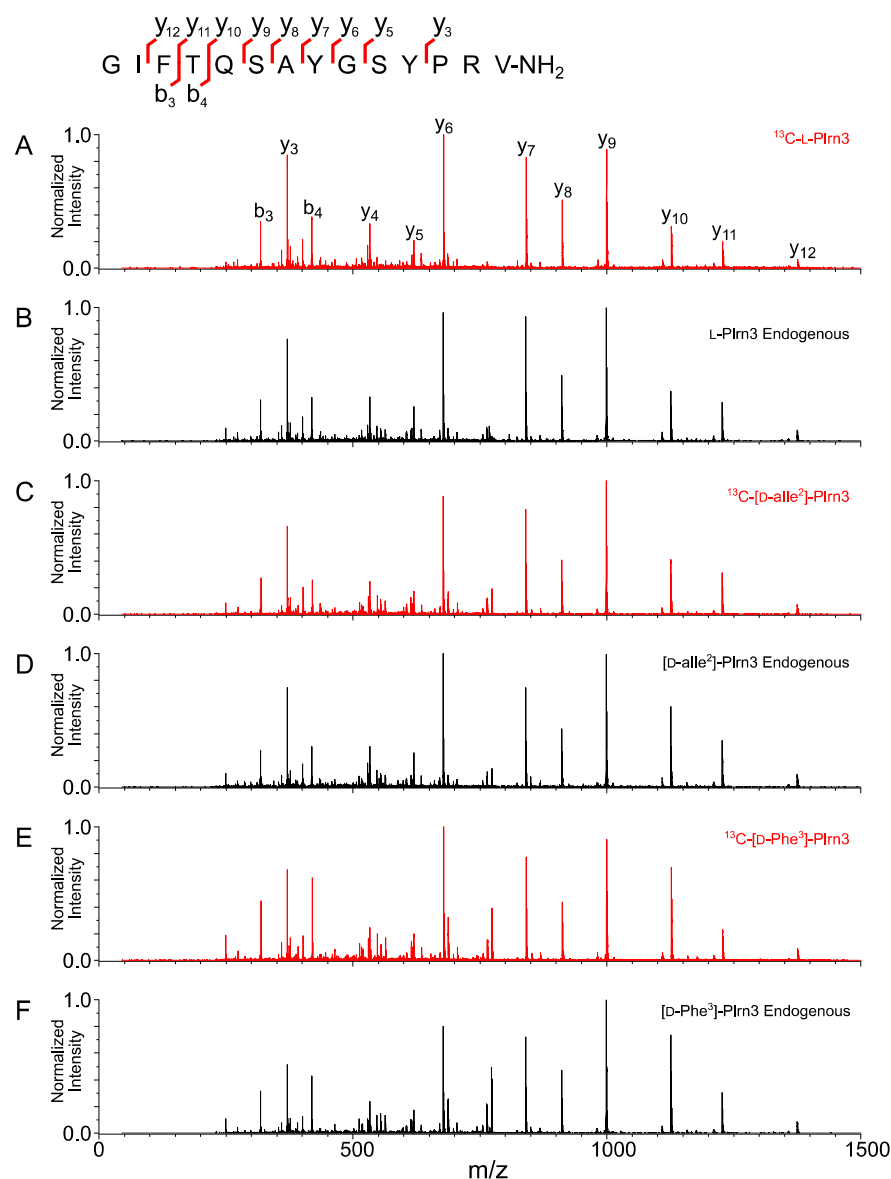


Figure 3.30. Product ion spectra comparison of endogenous Plrn3 (black traces) and synthetic Plrn3 standards (red Traces). CID collision energy was 33.2 eV. Product ion spectra are taken of the 2+ charge state of (A) synthetic ¹³C-L-Plrn3 ((*m*+1)/*z* 773.4); (B) endogenous L-Plrn3 (*m/z* 772.9) from no-APM-treated pleural ganglia extracts; (C) synthetic ¹³C-[D-alle²]-Plrn3 ((*m*+1)/*z* 773.4); (D) endogenous [D-alle²]-Plrn3 from APM-treated pleural ganglia extracts; (E) synthetic ¹³C-[D-Phe³]-Plrn3 ((*m*+2)/*z* 773.9); (F) endogenous [D-Phe³]-Plrn3 from no-APM-treated pleural ganglia extracts. The most abundant fragment ions are labeled.

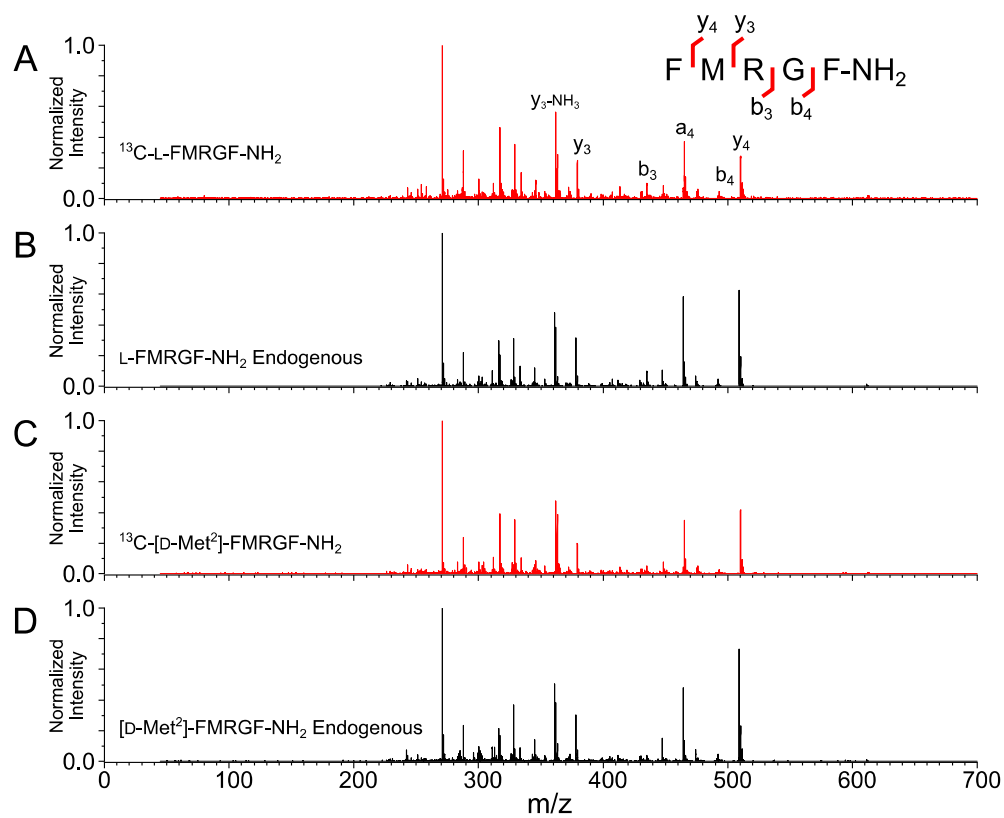


Figure 3.31. Product ion spectra comparison of endogenous FMRGF-NH₂ (black traces) and synthetic FMRGF-NH₂ standards (red traces). CID collision energy was 25.0 eV. Product ion spectra are taken of the 2+ charge state of (A) synthetic ^{13}C -L-FMRGF-NH₂ ($(m+1)/z$ 329.2); (B) endogenous L-FMRGF-NH₂ (m/z 328.7) from no-APM-treated cerebral ganglia extracts; (C) synthetic ^{13}C -[D-Met²]-FMRGF-NH₂ ($(m+1)/z$ 329.2); (D) endogenous [D-Met²]-FMRGF-NH₂ from APM-treated cerebral ganglia extracts. The most abundant fragment ions are labeled.

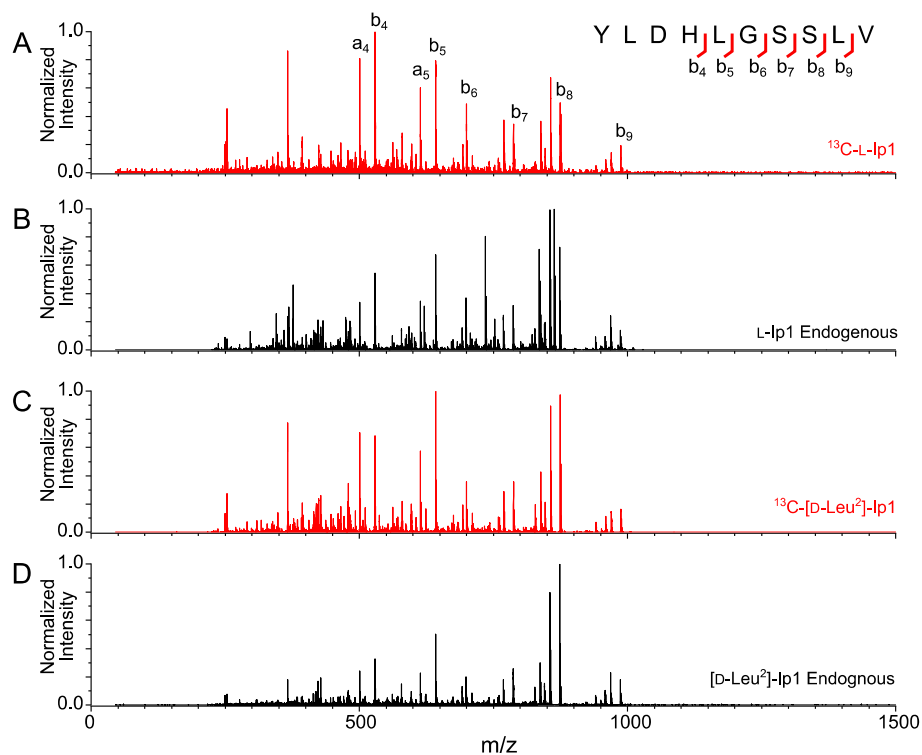


Figure 3.32. Product ion spectra comparison of endogenous Ip1 (black traces) and synthetic Ip1 standards (red traces). CID collision energy was 26.6 eV. Product ion spectra are taken of the 2+ charge state of (A) synthetic ^{13}C -L-Ip1 standard ($(m+1)/z$ 552.8); (B) endogenous L-Ip1 (m/z 552.3) from no-APM-treated cerebral ganglia extracts; (C) synthetic ^{13}C -[D-Leu²]-Ip1 standard ($(m+1)/z$ 552.8); (D) endogenous [D-Leu²]-Ip1 from APM-treated cerebral ganglia extracts. The most abundant fragment ions are labeled.

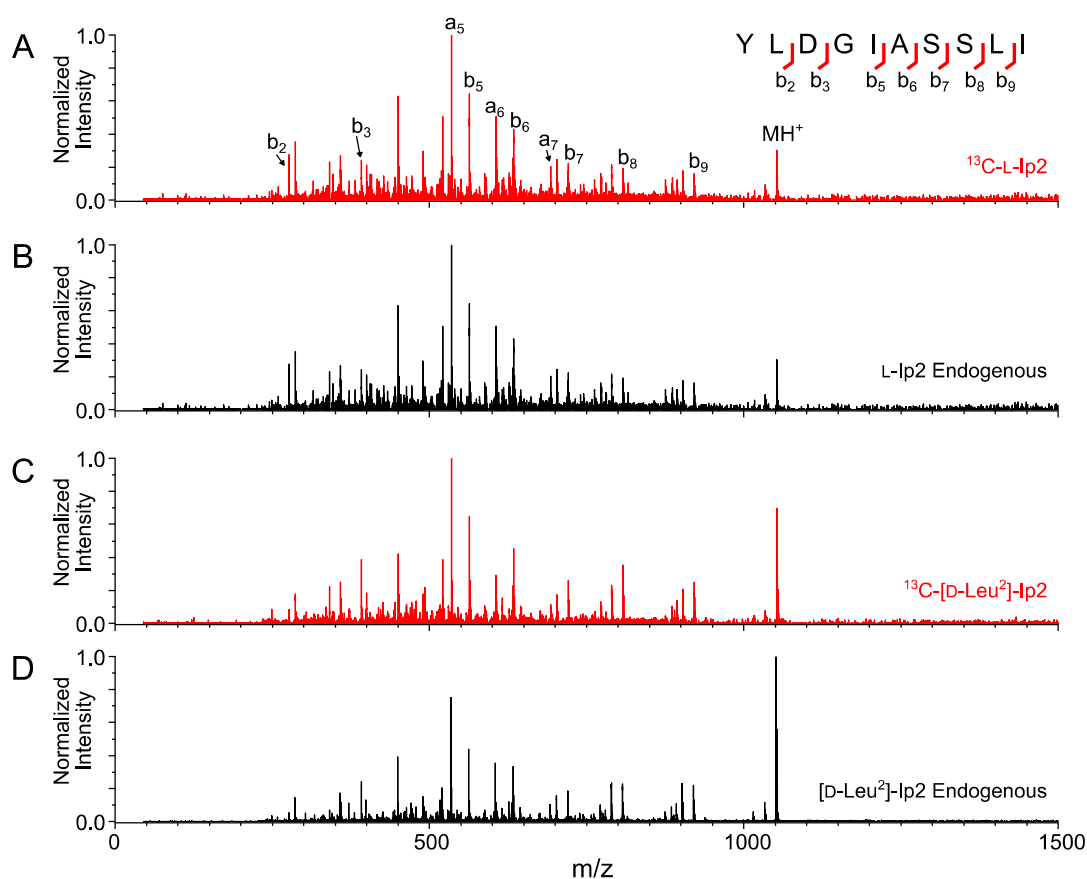


Figure 3.33. Product ion spectra comparison of endogenous Ip2 (black traces) and synthetic Ip2 standards (red traces). CID collision energy was 50.0 eV. Product ion spectra are taken of the 1+ charge state of (A) synthetic ^{13}C -L-Ip2 ($(m+1)/z$ 1052.6); (B) endogenous L-Ip2 (m/z 1051.6) from no-APM-treated cerebral ganglia extracts; (C) synthetic ^{13}C -[D-Leu²]-Ip2 ($(m+1)/z$ 1052.6); (D) endogenous [D-Leu²]-Ip2 from APM-treated cerebral ganglia extracts. The most abundant fragment ions are labeled.

Table 3.1. Peptide stereoisomers identified by APM screening. The predicted D-residue in each peptide is indicated in the single amino acid code preceded by a lowercase “d” (**dX**, bold font and underlined).

Peptide name	Peptide Sequence	Prohormone
L-Plrn1	MFYTKGSDSDYPRI-NH ₂	Pleurin
L-Plrn2	SFYTTGNGNHYPRI-NH ₂	Pleurin
L-Plrn3	GIFTQSAYGSYPRV-NH ₂	Pleurin
L-Ip1	YLDHLGSSLV	AIPP ^a
L-Ip2	YLDGIASSLI	AIPP
L-FMRGF-NH ₂	FMRGF-NH ₂	FMRGFamide ^b
DAACPs		
[D-Phe ²]-Plrn1	M <u>d</u> FYTKGSDSDYPRI-NH ₂	Pleurin
[D-Phe ²]-Plrn2	S <u>d</u> FYTTGNGNHYPRI-NH ₂	Pleurin
[D-allo ²]-Plrn3 ^c	G <u>d</u> IFTQSAYGSYPRV-NH ₂	Pleurin
[D-Leu ²]-Ip1	Y <u>d</u> LHLGSSLV	AIPP
[D-Leu ²]-Ip2	Y <u>d</u> LDGIASSLI	AIPP
[D-Met ²]-FMRGF-NH ₂	F <u>d</u> MGRGF-NH ₂	FMRGFamide
[D-Tyr ³]-Plrn2	SF <u>d</u> YTTGNGNHYPRI-NH ₂	Pleurin
[D-Phe ³]-Plrn3	G <u>d</u> IFTQSAYGSYPRV-NH ₂	Pleurin

^a*Aplysia* isomerized peptides precursor.

^bThree possible proteins (XP_005094941.1, XP_005094942.1, XP_005094943.1).

^cD-allo refers to (2R,3S)-2-amino-3-methylpentanoic acid.

Table 3.2. Summary of the endogenous peptide stereoisomers detected in the different ganglia during APM screening experiments. The D-residue in each peptide is shown as a single amino acid code preceded by a lowercase “d” (**dX**, bold, red font and underlined). Note [D-Phe²]-Plrn2 and [D-Tyr³]-Plrn2 coeluted on the column used for APM screening experiments, therefore, we qualified [D-Tyr³]-Plrn2 to be present only if Plrn2(2-14) was observed in the APM-treated ganglia extract.

Peptide name	Assigned Peptide Sequence	Ganglion Detected During APM experiments			
		Abd	Plr	Crb	Buc
L-Plrn1	MFYTKGSDSDYPRI-NH ₂	Y	Y	Y	Y
[D-Phe ²]-Plrn1	M <u>dF</u> YTKGSDSDYPRI-NH ₂	Y	Y	Y	Y
L-Plrn2	SFYTTGNGNHYPRI-NH ₂	Y	Y	Y	Y
[D-Phe ²]-Plrn2	S <u>dF</u> YTTGNGNHYPRI-NH ₂	Y	Y	Y	Y
[D-Tyr ³]-Plrn2	SF <u>dY</u> TTGNGNHYPRI-NH ₂	N	N	N	Y
L-Plrn3	GIFTQSAYGSYPRV-NH ₂	Y	Y	Y	Y
[D-allo ²]-Plrn3	G <u>dI</u> FTQSAYGSYPRV-NH ₂	Y	Y	Y	N
[D-Phe ³]-Plrn3	GI <u>dF</u> TQSAYGSYPRV-NH ₂	Y	Y	Y	Y
L-FMRGF-NH ₂	FMRGF-NH ₂	Y	Y	Y	Y
[D-Met ²]-FMRGF-NH ₂	F <u>dM</u> RGF-NH ₂	Y	Y	Y	Y
L-Ip1	YLDHLGSSLV	Y	Y	Y	Y
[D-Leu ²]-Ip1	Y <u>dL</u> DHLGSSLV	N	Y	Y	Y
L-Ip2	YLDGIASSLI	Y	Y	Y	Y
[D-Leu ²]-Ip2	Y <u>dL</u> DGIASSLI	Y	Y	Y	N

N denotes not detected
 Y denotes yes detected
 Abd = abdominal ganglion
 Plr = pleural ganglion
 Crb = cerebral ganglion
 Buc = buccal ganglion

Table 3.3. Synthetic peptide sequences. Bolded, italicized **G** indicates a ^{13}C -Gly labeled with one ^{13}C at the carbonyl carbon. The D-residue in each peptide is shown as a single amino acid code preceded by a lowercase “d” (**dX**, bold, red font and underlined).

Peptide name	Peptide Sequence	Monoisotopic Mass
^{13}C -L-Plrn1	MFYTK G SDSDYPRI-NH ₂	1678.79
^{13}C -[D-Phe ²]-Plrn1	M <u>dF</u> YTK G SDSDYPRI-NH ₂	1678.79
^{13}C -[D-Tyr ³]-Plrn1	MF <u>dY</u> TK G SDSDYPRI-NH ₂	1678.79
^{13}C -L-Plrn2	SFYTTGN G NHYPRI-NH ₂	1625.78
^{13}C -[D-Phe ²]-Plrn2	S <u>dF</u> YTTGN G NHYPRI-NH ₂	1625.78
^{13}C -[D-Tyr ³]-Plrn2	SF <u>dY</u> TTGN G NHYPRI-NH ₂	1626.78
^{13}C -L-Plrn3	GIFTQSAY G SYPRV-NH ₂	1544.79
^{13}C -[D-alle ²]-Plrn3	<u>GdI</u> FTQSAYGSYPRV-NH ₂	1544.79
^{13}C -[D-Phe ³]-Plrn3	<u>GIdF</u> TQSAYGSYPRV-NH ₂	1545.79
^{13}C -L-FMRGF-NH ₂	FMR G F-NH ₂	656.33
^{13}C -[D-Met ²]-FMRGF-NH ₂	F <u>dM</u> R G F-NH ₂	656.33
^{13}C -L-Ip1	YLDHL G SSLV	1103.57
^{13}C -[D-Leu ²]-Ip1	Y <u>dL</u> DHL G SSLV	1103.57
^{13}C -L-Ip2	YLD G IASLI	1051.57
^{13}C -[D-Leu ²]-Ip2	Y <u>dL</u> D G IASLI	1051.57

Table 3.4. Peptide products derived from the *Aplysia* isomerized peptides precursor (AIPP) that were detected in screening experiments. Table values were taken directly from PEAKS output. Unless otherwise specified, reported peptides were detected in the “no-APM” controls of screening experiments. Green highlighted residues indicate identical amino acids shared between Ip1, Ip2, and Ip-like sequences. Yellow highlighted amino acids indicate highly similar residues shared between Ip1, Ip2, and Ip-like sequences. Peptides were named arbitrarily.

Name	Peptide	-10lgP	Mass	ppm	Observed <i>m/z</i>	<i>z</i>
Ip1	YLDHLGSSLV	34.25	1102.5658	16.0	552.2990	2
Ip2	YLDGIASSLI	44.52	1050.5597	17.5	1051.5853	1
Ip-like1	YLDDVASSLLH	43.42	1231.6084	18.3	616.8228	2
Ip-like2	YLDTVASSLF	41.75	1114.5546	18.2	1115.5822	1
Ip-like3	YLDRVGSLLI	22.03	1121.6080	22.0	561.8236	2
Ip-like4	YLDRIGSSLV	17.89	1121.6080	18.5	561.8217	2
Ip-like5	YLDQVGSSLV	35.99	1079.5498	20.3	1080.5790	1
Ip-like6a	FLESIAGHII	36.70	1098.6073	17.8	550.3207	2
Ip-like6b	FIESIAGHII ^a	--	--	--	--	--
Ip-like7	FIESIAGNIIY	34.55	1238.6547	18.1	1239.6843	1
Un-1	LSNLPARPVGLGSGVPIH	39.58	1783.0104	16.0	595.3548	3
Un-2	FTTESHRPLKAGNNRHMSI	20.57	2282.1338	19.0	571.5516	4

^aIp-like6b is a structural isomer of Ip-like6a, and either could have been detected. It was not determined which was present.

Table 3.5. Peptides detected in LC–MS/MS screening experiments from the *Aplysia* FMRGFamide precursor(s). Table values were taken directly from PEAKS output. Unless otherwise specified, reported peptides were detected in the no-APM controls and the detected peptides are encoded in the precursor flanked by monobasic or dibasic amino acid residues. Peptides were named arbitrarily.

Name	Peptide	-10lgP	Mass	ppm	<i>m/z</i>	<i>z</i>
FMRGFamide	FMRGF (- .98) .G	25.05	655.3264	21.9	656.3480	1
FMRGF	FMRGF	15.74	656.3104	19.9	657.3307	1
FMRGFGamide	FMRGFG (- .98) .G ^a	-	713.3600	18.7	357.1879	2
FMRGFH	FMRGFH	21.05	793.3693	19.4	397.6996	2
FMRGFN	FMRGFN	18.12	770.3534	18.9	386.1913	2
FMRGFQ	FMRGFQ ^a	-	784.3800	19.6	393.1995	2
FMRGFQamide	FMRGFQ (- .98) .G ^a	-	783.3900	18.6	392.7071	2
FMRGFxMA	FMRGFQNFVNMA (+15 .99) A	36.14	1476.6642	18.7	739.3532	2
FMRGFxMA	FMRGFQNFVNMA	34.24	1460.6693	21.2	731.3574	2
FMRGFxYA	FMRGFSNYA	30.17	1091.4858	23.1	546.7628	2
FMRGLxYA	FMRGLNLYA	15.36	1083.5535	20.8	542.7953	2
FMRGLxWQ	FMRGLQNWQ	27.34	1178.5654	18.7	590.3010	2
FMRGLSAY	FMRGLSAY	35.67	943.4586	20.1	472.7460	2
FIHGLxYS	FIHGLQLYS	16.31	1076.5654	19.4	539.3004	2
FIHGLxEE	FIHGLQLYSRSYGGLGDSPEE	19.34	2423.1758	20.3	808.7490	3
FIKGLxAS	FIKGLERYALPAS	53.87	1463.8136	19.9	732.9286	2
--	FIMGLNNYA .D	50.31	1041.4954	17.0	1042.5204	1
FIMGLxAD	FIMGLNNYAD	47.71	1156.5222	18.2	1157.5505	1
--	FIRGLHLY .H	34.28	1017.5759	21.1	509.8060	2
FIRGLxYH	FIRGLHLYH	33.21	1154.6349	18.8	385.8928	3
FLRGLxYA	FLRGLNNYA	35.22	1066.5559	19.9	534.2958	2
	GAGGDSLDEDIS	39.59	1134.4677	18.3	1135.4957	1
	GGSSLGANFEVGDDIVNDGPVN	39.11	2131.9658	17.8	1067.0092	2
	LPTAETENDTEE	51.75	1347.5677	20.0	674.8046	2
	LPTEDDTDE	52.05	1033.4087	17.8	1034.4343	1
	LPTEDDTEE	47.63	1047.4244	18.3	1048.4509	1
	NYVTSGEDE	33.55	1012.3985	19.2	1013.4252	1
	SGVDAVPSDVD	43.05	1059.4720	18.0	1060.4984	1
	SYGGLGDSPEE	51.37	1208.5197	16.5	1209.5469	1
	YLGWDIPDSKNRFPASSVSVEPVLE VSQEDGPSSVWEGEYDQPK ^b	44.62	4922.3198	12.6	1231.6028	4
	ADEINELE	25.68	931.4135	17.3	932.4368	1
	DPILAGDISVDLAEL .D	32.46	1541.7460	18.7	771.8947	2
	DPILAGDISVDLAED	39.82	1656.7729	16.9	829.4077	2
	EDAFSESDDIEVDD	26.24	1584.5951	18.9	793.3198	2
	NEVYSDEEDLQTVAD	37.28	1725.7217	19.2	863.8846	2
	DLIGDDTDE	41.91	991.3982	17.4	992.4227	1
	DVSGLVDD	36.26	818.3657	18.8	819.3884	1
	E (-18 .01) LVESLGGED ^b	38.6	1028.4662	14.3	1029.4800	3
	ELVEGNDDE	32.49	1018.4091	12.2	1019.4288	1
	ELVESLGGED	50.07	1046.4768	17.5	1047.5024	1
	EPIDDVTD	37.58	1017.4138	17.5	1018.4389	1
	EPVDDVTDD	41.56	1003.3982	18.1	1004.4236	1
	EPVEDVTDD	39.73	1017.4138	18.8	1018.4402	1

^a Peptide was not picked up by a database search, but the *m/z* was detected by LC–MS/MS and the sequence supported by MS/MS evidence.

^b Peptide detected only in the APM-treated sample.

Table 3.6. Comparison of the relative abundance of DAACPs between buccal and cerebral ganglia extracts. %D was calculated using EICs from m/z provided in the table by taking the sum of the peak area of all confirmed DAACP stereoisomers divided by the sum of the peak area of all stereoisomers detected. %D reported as the average $\pm$ standard deviation of the peak areas measured in $N \geq 3$ biological sets where each biological set consisted of pooled cerebral or buccal ganglia extracts from 3-5 animals.

Peptide	Cerebral %D	Buccal %D	m/z	z
Plrn1 ^a	67 $\pm$ 6	70 $\pm$ 3	560.3 $\pm$ 0.1	3
Plrn2	68 $\pm$ 3	64 $\pm$ 2	542.6 $\pm$ 0.1	3
Plrn3	62 $\pm$ 4	23 $\pm$ 5	772.9 $\pm$ 0.1	2
FMRGF-NH ₂ ^a	19 $\pm$ 5	3 $\pm$ 1	328.7 $\pm$ 0.1	2
Ip1 ^{a, b}	--	--	552.3 $\pm$ 0.1	2
Ip2 ^{a, b}	--	--	1051.6 $\pm$ 0.1	1

^a only D2 and L forms were present for this peptide

^b Relative abundance could not be accurately calculated due to the presence of a coeluting species with an m/z (but different charge) overlapping with the m/z of the D2 peptide.

3.8 REFERENCES

1. Riggs, D. L.; Gomez, S. V.; Julian, R. R., Sequence and solution effects on the prevalence of D-Isomers produced by deamidation. *ACS. Chem. Biol.* **2017**, *12* (11), 2875-2882.
2. Lambeth, T. R.; Riggs, D. L.; Talbert, L. E.; Tang, J.; Coburn, E.; Kang, A. S.; Noll, J.; Augello, C.; Ford, B. D.; Julian, R. R., Spontaneous isomerization of long-lived proteins provides a molecular mechanism for the lysosomal failure observed in alzheimer's disease. *ACS Cent. Sci.* **2019**, *5* (8), 1387-1395.
3. Hooi, M. Y.; Truscott, R. J., Racemisation and human cataract. D-Ser, D-Asp/Asn and D-Thr are higher in the lifelong proteins of cataract lenses than in age-matched normal lenses. *Age (Dordrecht, Netherlands)* **2011**, *33* (2), 131-41.
4. Roher, A. E.; Lowenson, J. D.; Clarke, S.; Wolkow, C.; Wang, R.; Cotter, R. J.; Reardon, I. M.; Zurcher-Neely, H. A.; Heinrichson, R. L.; Ball, M. J.; et al., Structural alterations in the peptide backbone of beta-amyloid core protein may account for its deposition and stability in Alzheimer's disease. *J. Biol. Chem.* **1993**, *268* (5), 3072-83.
5. Bai, L.; Sheeley, S.; Sweedler, J. V., Analysis of endogenous D-amino acid-containing peptides in metazoa. *Bioanal Rev* **2009**, *1* (1), 7-24.
6. Ollivaux, C.; Soye, D.; Toullec, J. Y., Biogenesis of D-amino acid containing peptides/proteins: where, when and how? *J. Pept. Sci.* **2014**, *20* (8), 595-612.
7. Bai, L.; Livnat, I.; Romanova, E. V.; Alexeeva, V.; Yau, P. M.; Vilim, F. S.; Weiss, K. R.; Jing, J.; Sweedler, J. V., Characterization of GdFFD, a D-amino acid-containing neuropeptide that functions as an extrinsic modulator of the aplysia feeding circuit. *J. Biol. Chem.* **2013**, *288* (46), 32837-32851.

8. Kamatani, Y.; Minakata, H.; Kenny, P. T. M.; Iwashita, T.; Watanabe, K.; Funase, K.; Sun, X. P.; Yongsiri, A.; Kim, K. H.; Novalesli, P.; Novales, E. T.; Kanapi, C. G.; Takeuchi, H.; Nomoto, K., Achatin-I, an endogenous neuroexcitatory tetrapeptide from *Achatina fulica ferrussac* containing a D-amino acid residue. *Biochem. Biophys. Res. Commun.* **1989**, *160* (3), 1015-1020.
9. Morishita, F.; Nakanishi, Y.; Kaku, S.; Furukawa, Y.; Ohta, S.; Hirata, T.; Ohtani, M.; Fujisawa, Y.; Muneoka, Y.; Matsushima, O., A novel D-amino-acid-containing peptide isolated from *Aplysia* heart. *Biochem. Biophys. Res. Commun.* **1997**, *240* (2), 354-358.
10. Checco, J. W.; Zhang, G.; Yuan, W.-D.; Yu, K.; Yin, S.-Y.; Roberts-Galbraith, R. H.; Yau, P. M.; Romanova, E. V.; Jing, J.; Sweedler, J. V., Molecular and physiological characterization of a receptor for D-amino acid-containing neuropeptides. *ACS. Chem. Biol.* **2018**, *13* (5), 1343-1352.
11. Buczek, O.; Yoshikami, D.; Bulaj, G.; Jimenez, E. C.; Olivera, B. M., Post-translational amino acid isomerization: a functionally important D-amino acid in an excitatory peptide. *J. Biol. Chem.* **2005**, *280* (6), 4247-4253.
12. Checco, J. W.; Zhang, G.; Yuan, W. D.; Le, Z. W.; Jing, J.; Sweedler, J. V., *Aplysia* allatotropin-related peptide and its newly identified d-amino acid-containing epimer both activate a receptor and a neuronal target. *J. Biol. Chem.* **2018**, *293* (43), 16862-16873.
13. Fujisawa, Y.; Ikeda, T.; Nomoto, K.; Yasudakamatani, Y.; Minakata, H.; Kenny, P. T. M.; Kubota, I.; Muneoka, Y., The FMRFamide-related decapeptide of *Mytilus* contains a D-amino acid residue. *Comp. Biochem. Physiol. C-Pharmacol. Toxicol. Endocrinol.* **1992**, *102* (1), 91-95.

14. Soyez, D.; Van Herp, F.; Rossier, J.; Le Caer, J. P.; Tensen, C. P.; Lafont, R., Evidence for a conformational polymorphism of invertebrate neurohormones. D-amino acid residue in crustacean hyperglycemic peptides. *J. Biol. Chem.* **1994**, *269* (28), 18295-8.
15. Yasuda, A.; Yasuda, Y.; Fujita, T.; Naya, Y., Characterization of crustacean hyperglycemic hormone from the crayfish (*Procambarus clarkii*): multiplicity of molecular forms by stereoinversion and diverse functions. *Gen. Comp. Endocrinol.* **1994**, *95* (3), 387-98.
16. Buchberger, A.; Yu, Q.; Li, L., Advances in mass spectrometric tools for probing neuropeptides. *Annu. Rev. Anal. Chem.* **2015**, *8*, 485-509.
17. Secher, A.; Kelstrup, C. D.; Conde-Frieboes, K. W.; Pyke, C.; Raun, K.; Wulff, B. S.; Olsen, J. V., Analytic framework for peptidomics applied to large-scale neuropeptide identification. *Nat. Commun.* **2016**, *7*, 11436.
18. Dallas, D. C.; Guerrero, A.; Parker, E. A.; Robinson, R. C.; Gan, J.; German, J. B.; Barile, D.; Lebrilla, C. B., Current peptidomics: applications, purification, identification, quantification, and functional analysis. *Proteomics* **2015**, *15* (5-6), 1026-1038.
19. Hook, V.; Lietz, C. B.; Podvin, S.; Cajka, T.; Fiehn, O., Diversity of neuropeptide cell-cell signaling molecules generated by proteolytic processing revealed by neuropeptidomics mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2018**, *29* (5), 807-816.
20. Fricker, L. D., Analysis of mouse brain peptides using mass spectrometry-based peptidomics: implications for novel functions ranging from non-classical neuropeptides to microproteins. *Mol. BioSyst.* **2010**, *6* (8), 1355-65.
21. Decaillot, F. M.; Che, F. Y.; Fricker, L. D.; Devi, L. A., Peptidomics of Cpefat/fat mouse hypothalamus and striatum: effect of chronic morphine administration. *J. Mol. Neurosci.* **2006**, *28* (3), 277-284.

22. Ma, M.; Szabo, T. M.; Jia, C.; Marder, E.; Li, L., Mass spectrometric characterization and physiological actions of novel crustacean C-type allatostatins. *Peptides* **2009**, *30* (9), 1660-8.
23. Cape, S. S.; Rehm, K. J.; Ma, M.; Marder, E.; Li, L., Mass spectral comparison of the neuropeptide complement of the stomatogastric ganglion and brain in the adult and embryonic lobster, *Homarus americanus*. *J. Neurochem.* **2008**, *105* (3), 690-702.
24. Livnat, I.; Tai, H. C.; Jansson, E. T.; Bai, L.; Romanova, E. V.; Chen, T. T.; Yu, K.; Chen, S. A.; Zhang, Y.; Wang, Z. Y.; Liu, D. D.; Weiss, K. R.; Jing, J.; Sweedler, J. V., A D-amino acid-containing neuropeptide discovery funnel. *Anal. Chem.* **2016**, *88* (23), 11868-11876.
25. Ewing, M. A.; Wang, J.; Sheeley, S. A.; Sweedler, J. V., Detecting D-amino acid-containing neuropeptides using selective enzymatic digestion. *Anal. Chem.* **2008**, *80* (8), 2874-2880.
26. Torres, A. M.; Menz, I.; Alewood, P. F.; Bansal, P.; Lahnstein, J.; Gallagher, C. H.; Kuchel, P. W., D-amino acid residue in the C-type natriuretic peptide from the venom of the mammal, *Ornithorhynchus anatinus*, the Australian platypus. *FEBS Lett.* **2002**, *524* (1-3), 172-176.
27. Dutertre, S.; Lumsden, N. G.; Alewood, P. F.; Lewis, R. J., Isolation and characterisation of conomaph-Vt, a D-amino acid containing excitatory peptide from the venom of a vermivorous cone snail. *FEBS Lett.* **2006**, *580* (16), 3860-3866.
28. Moroz, L. L.; Edwards, J. R.; Puthanveetil, S. V.; Kohn, A. B.; Ha, T.; Heyland, A.; Knudsen, B.; Sahni, A.; Yu, F.; Liu, L.; Jezini, S.; Lovell, P.; Iannuccilli, W.; Chen, M.; Nguyen, T.; Sheng, H.; Shaw, R.; Kalachikov, S.; Panchin, Y. V.; Farmerie, W.; Russo, J.

- J.; Ju, J.; Kandel, E. R., Neuronal transcriptome of Aplysia: neuronal compartments and circuitry. *Cell* **2006**, *127* (7), 1453-67.
29. Wachsmuth, E. D.; Fritze, I.; Pfeleiderer, G., An aminopeptidase occurring in pig kidney. I. An improved method of preparation - physical and enzymic properties. *Biochemistry* **1966**, *5* (1), 169-174.
30. Wachsmuth, E. D.; Fritze, I.; Pfeleiderer, G., An aminopeptidase occurring in pig kidney. II. A study on mechanism of hydrolysis. *Biochemistry* **1966**, *5* (1), 175-182.
31. Turner, A. J., *Handbook of Proteolytic Enzymes*. 2 ed.; Elsevier Science Bv: Amsterdam, 2004; Vol. 1, p 1-3932.
32. Soye, D.; Toullec, J. Y.; Montagne, N.; Ollivaux, C., Experimental strategies for the analysis of D-amino acid containing peptides in crustaceans: a review. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2011**, *879* (29), 3102-7.
33. Iida, T.; Santa, T.; Toriba, A.; Imai, K., Amino acid sequence and D/L-configuration determination methods for D-amino acid-containing peptides in living organisms. *Biomed. Chromatogr.* **2001**, *15* (5), 319-327.
34. Do, T. D.; Checco, J. W.; Tro, M.; Shea, J. E.; Bowers, M. T.; Sweedler, J. V., Conformational investigation of the structure-activity relationship of GdFFD and its analogues on an achatin-like neuropeptide receptor of Aplysia californica involved in the feeding circuit. *Phys. Chem. Chem. Phys.* **2018**, *20* (34), 22047-22057.
35. Jia, C.; Lietz, C. B.; Yu, Q.; Li, L., Site-specific characterization of D-amino acid containing peptide epimers by ion mobility spectrometry. *Anal. Chem.* **2014**, *86* (6), 2972-81.

36. Jeanne Dit Fouque, K.; Garabedian, A.; Porter, J.; Baird, M.; Pang, X.; Williams, T. D.; Li, L.; Shvartsburg, A.; Fernandez-Lima, F., Fast and effective ion mobility-mass spectrometry separation of D-amino-acid-containing peptides. *Anal. Chem.* **2017**, *89* (21), 11787-11794.
37. Li, G.; Delafield, D. G.; Li, L., Improved structural elucidation of peptide isomers and their receptors using advanced ion mobility-mass spectrometry. *TrAC Trends Anal. Chem.* **2019**.
38. Ollivaux, C.; Gallois, D.; Amiche, M.; Boscameric, M.; Soye, D., Molecular and cellular specificity of post-translational aminoacyl isomerization in the crustacean hyperglycaemic hormone family. *FEBSj.* **2009**, *276* (17), 4790-4802.
39. Xie, F. Neuropeptidomics Across Metazoan Life. Dissertation, University of Illinois at Urbana-Champaign, 2009.
40. Nielsen, H., Predicting secretory proteins with signalP. *Methods Mol. Bio.* **2017**, *1611*, 59-73.
41. Richter, K.; Egger, R.; Negri, L.; Corsi, R.; Severini, C.; Kreil, G., cDNAs encoding [D-Ala²]deltorphan precursors from skin of *Phyllomedusa bicolor* also contain genetic information for three dermorphin-related opioid peptides. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (12), 4836-4839.
42. YasudaKamatani, Y.; Kobayashi, M.; Yasuda, A.; Fujita, T.; Minakata, H.; Nomoto, K.; Nakamura, M.; Sakiyama, F., A novel D-amino acid-containing peptide, fulyal, coexists with fulicin gene-related peptides in *Achatina* atria. *Peptides* **1997**, *18* (3), 347-354.
43. Rost, H. L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weisser, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H. C.; Gutenbrunner, P.; Kenar, E.; Liang, X.; Nahnsen, S.; Nilse, L.; Pfeuffer, J.; Rosenberger, G.; Rurik, M.; Schmitt, U.; Veit, J.; Walzer, M.; Wojnar, D.; Wolski, W.

- E.; Schilling, O.; Choudhary, J. S.; Malmstrom, L.; Aebersold, R.; Reinert, K.; Kohlbacher, O., OpenMS: a flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **2016**, *13*, 741-748.
44. Adams, C. M.; Zubarev, R. A., Distinguishing and quantifying peptides and proteins containing D-amino acids by tandem mass spectrometry. *Anal. Chem.* **2005**, *77*, 4571-4580.
45. Koehbach, J.; Gruber, C. W.; Becker, C.; Kreil, D. P.; Jilek, A., MALDI TOF/TOF-Based Approach for the Identification of D-Amino Acids in Biologically Active Peptides and Proteins. *J. Proteome Res.* **2016**, *15*, 1487-1496.

CHAPTER 4

ANALYSIS OF PEPTIDE STEREOCHEMISTRY IN SINGLE CELLS BY CAPILLARY ELECTROPHORESIS-TRAPPED ION MOBILITY SPECTROMETRY-MASS SPECTROMETRY

4.1 NOTES AND ACKNOWLEDGEMENTS

The text in this chapter is published in Analytical Chemistry: Mast, D. H.; Liao, H. W.; Romanova, E. V.; Sweedler, J. V., Analysis of Peptide Stereochemistry in Single Cells by Capillary Electrophoresis-Trapped Ion Mobility Spectrometry Mass Spectrometry. *Anal Chem* **2021**, 93 (15), 6205-6213. DOI: <https://doi.org/10.1021/acs.analchem.1c00445>.

Special thanks Professor T. Do at the University of Tennessee Knoxville for useful discussions about ion mobility spectrometry. Special thanks to C. Mullens, and T. Srikumar at Bruker Corp. for guidance and assistance on TIMS measurements and the operation of the timsTOF pro mass spectrometer. This work was supported by the National Institutes of Health, National Institute of Neurological Disorders and Stroke by Award No. R01NS031609, and the National Institute on Drug Abuse by Award No. P30DA018310. The National Resource for *Aplysia* (Miami, FL) is funded by PHS Grant P40OD010952. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

4.2 ABSTRACT

Single cell analysis strives to probe molecular heterogeneity in morphologically similar cell populations through quantitative or qualitative measurements of genetic, proteomic, or

metabolic products. Here, we apply mass analysis of single neurons to investigate cell-cell signaling peptides. The multiplicity of endogenous cell-cell signaling peptides is a common source of chemical diversity among cell populations. Certain peptides can undergo post-translational isomerization of select residues, which has important physiological consequences. The limited number of single cell analysis techniques that are sensitive to peptide stereochemistry make it challenging to study isomerization at the individual cell level. We performed capillary electrophoresis (CE) with mass spectrometry (MS) detection to characterize the peptide content of single cells. Using complementary trapped ion mobility spectrometry (TIMS) separations, we measured the stereochemical configurations of three neuropeptide gene products derived from the pleurin precursor in individual neurons ($N = 3$) isolated from the central nervous system of *Aplysia californica*. An analysis of the resultant mobility profiles indicated >98% of the detectable pleurin-derived peptides exist as the non-isomerized, all-L forms in individual neuron cell bodies. However, we observed 44% of the Pln2 peptide from the pleurin precursor was present as the isomerized, D-residue containing form in the nerve tissue. These findings demonstrate an unusual distribution of isomerized peptides in *A. californica* and establish CE–TIMS MS as a powerful analytical tool for investigating peptide stereochemistry at the single cell level.

4.3 INTRODUCTION

The inversion of amino acid residue stereochemistry can manifest substantial alterations to the normal physiochemical or biological properties of endogenous peptides.¹ Enzymatic L- to D-residue isomerization is an unusual post-translational modification (PTM) in the secretory pathway that contributes to the diversification of peptide structures and functions.¹⁻² For decades, the products of L- to D-residue isomerization, D-amino acid containing peptides (DAACPs), have been

isolated from numerous animal species;² however, because the enzymes that catalyze isomerization are not known, conventional molecular biology or biochemistry approaches cannot be used to measure DAACPs in individual cells. As a result, peptide isomerization remains an underexplored PTM.² Improved methods for measuring peptide diastereomers are essential for understanding the biosynthetic pathways involved in neuropeptide isomerization.

Differentiating and quantifying peptide diastereomers in single cells requires methods that provide both low detection limits and high stereoselectivity. Mass spectrometry (MS) excels at distinguishing the individual chemical components in complex mixtures and satisfies the low detection-limit requirements needed for single cell analysis. MS has been successfully used for single cell chemical profiling,³⁻⁴ but because MS alone is not inherently stereospecific, it must be coupled with complementary techniques to analyze endogenous DAACPs. High sensitivity and stereoselective tandem MS (MS/MS)-based methods have been developed to distinguish peptide diastereomers,⁵⁻⁸ but have only occasionally been useful for detecting DAACPs in bulk tissue extracts;^{2, 9-12} moreover, fragmentation-based techniques have seldom been applied for the relative quantification of DAACPs in individual cells.¹³ This is in part because differentiation and quantitation of peptide diastereomers by MS/MS is not straight forward, and relies on product ion ratios rather than a physical separation of diastereomers.^{5, 7, 13}

Solution-phase separation methods such as liquid chromatography (LC) have been used to distinguish native DAACPs in biological systems,^{2, 11} but LC requires sample dilution and clean-up steps that tend to be incompatible with handling single cells, even when coupled with MS detection. In contrast, capillary electrophoresis (CE) is convenient for small-volume analysis, and when used in combination with MS detection, is well suited for the biomolecular characterization of individual cells. CE can accomplish efficient molecular separations on small sample-volumes

(pL–nL of sample),¹⁴⁻¹⁵ which minimizes sample dilution steps and improves the ionization of low-abundance analytes for subsequent MS analysis. Thus, CE–MS has been widely applied for profiling small molecules, metabolites, proteins, and peptides in animal tissues, including at or near the single cell level.¹⁶⁻²⁴

Despite the advantages of CE–MS for single cell analysis, there is still a need for further improvements to existing sampling techniques and electrophoretic separations of peptide stereoisomers.²⁵⁻²⁶ Specifically, CE separations of peptide stereoisomers may require specialized buffer conditions and chiral selectivity reagents that are incompatible with MS detection.²⁷⁻²⁹ As a result, characterization of DAACPs by CE–MS remains challenging.

Coupling CE with ion mobility spectrometry (IMS) MS is a promising approach that can circumvent several of the challenges facing the analysis of DAACPs in single cells by CE–MS. Electrospray ionization (ESI)–IMS MS analysis has been applied to achieve lower detection limits and increase MS spectral feature detection compared to ESI–MS alone.³⁰⁻³² This is possible because IMS serves as an additional dimension of ion separation in the gas phase, allowing for the partitioning of isomeric/isobaric species based on molecular geometry. Furthermore, IMS MS can distinguish peptide diastereomers with subtle structural variations that might not easily separate by solution-phase separation approaches.^{2, 33-36}

Conveniently, mass spectrometers capable of IMS measurements can be outfitted with commercial or custom CE devices for single cell measurements. While there are several commercially available IMS systems, the resolving powers achievable by trapped ion mobility spectrometry (TIMS) relative to the timescale of the mobility measurements makes TIMS well suited to coupling with CE separations.³⁷⁻³⁸ Parallel accumulation TIMS enables acquisition cycle speeds and mobility measurements on the order of tens to hundreds of milliseconds, which are

compatible with high-efficiency CE separations where analytes can migrate over a narrow migration time frame (on the order of seconds). Additionally, with TIMS, ions are preconcentrated (trapped) in an electric field and this enables increased MS signal-to-noise ratios and lower detection limits with minimal ion losses.³⁹ For these reasons, TIMS is rising in popularity for the analysis of small-volume biological samples.⁴⁰ Taken together, CE in combination with TIMS MS has potential for characterizing DAACPs at the single cell level.

We created a CE–TIMS MS workflow that allowed us to accurately assign the stereochemistry of selected endogenous neuropeptides in individual neurons (Scheme 4.1). More specifically, sampling techniques and CE–MS approaches adapted to work with individual cells were combined with TIMS and used to characterize isomerized peptides.^{2, 41} CE enabled us to separate the peptides from individual neurons with minimal sample dilution, and TIMS allowed us to accurately assign the native peptide stereochemistry. The combined approaches facilitated the detection and simultaneous stereochemical characterization of three known isomerized neuropeptides derived from the pleurin precursor (NP_001191654.1) in single neurons isolated from the *Aplysia californica* central nervous system (CNS).¹¹

The *A. californica* CNS is an excellent model for developing single cell methods because the locations of individual neurons that are known to express DAACP-encoding precursors have been identified by *in-situ* hybridization,⁴²⁻⁴³ and these live neurons can be manually isolated for subsequent analysis.⁴⁴ In this study, we specifically targeted neurons expressing the pleurin precursor because it was recently discovered to encode three peptides that undergo post-translational L- to D-residue isomerization: Plrn1, Plrn2, and Plrn3 (Plrn1–Plrn3, sequences are listed in Table 4.1).¹¹

Based on our prior work,¹¹ we know that in the *A. californica* CNS, Plrn1 exists as two possible stereoisomers, L-Plrn1 and D2-Plrn1 (D2 indicates a D-residue at the second residue from the N-terminus); whereas Plrn2 and Plrn3 each exist as three possible stereoisomers, L-Plrn2/Plrn3, D2-Plrn2/Plrn3, or D3-Plrn2/Plrn3 (D3 indicates a D-residue at the third residue from the N-terminus) (Table 4.1).

Using CE–TIMS MS, we sought to determine the percentage of pleurin-derived peptides that are isomerized in individual neurons. To accomplish this, live pleurin-expressing neurons were manually isolated from the *A. californica* CNS according to an *in-situ* hybridization staining map.⁴³ Individual cell extracts were analyzed by CE–TIMS MS and the mobility profiles of the native Plrn1–Plrn3 ions were compared to the mobility profiles of synthetic ¹³C-labeled Plrn1–Plrn3 standards to assess the native peptide stereochemistry. Further, we used TIMS mobility profiles to quantify the relative abundance of Plrn2 diastereomers in both neuron cell bodies, and in the nerves connecting the central ganglia (referred to as the ganglia connectives). Our findings reveal that the D-residue-containing forms of pleurin-derived peptides are not homogenously distributed between these two locations of the CNS.

4.4 METHODS

Reagents and Standards

All solvents and reagents, LC–MS grade or higher, used for the CE–TIMS MS measurements were purchased from Thermo Fisher Scientific (Waltham, MA). Unless otherwise specified, reagents used for *A. californica* dissections and neuron isolations were >98% purity and purchased from MilliporeSigma (St. Louis, MO) or Thermo Fisher Scientific. Synthetic peptide

standards of ^{13}C -labeled Plrn1, Plrn2, and Plrn3 were previously synthesized and characterized,² with the sequences provided in Table 4.1.

Lab-Made Micropipettes

Lab-made micropipettes were constructed from disposable Eppendorf (Hamburg, Germany) Microloader™ pipette tips, which were cut using a razor blade to create an inlet with an approximately 200 μm inner diameter. The micropipette tip was then capped with a rubber bulb made from 1 mm rubber tubing melted at one end to form a seal.

Animals and Single Cell Isolation.

A. californica (100–200g) were obtained from the National Institutes of Health/University of Miami National Resource for *Aplysia* (Miami, FL). Animals were kept in an aquarium containing aerated and filtered Instant Ocean seawater (Aquarium Systems Inc., Mentor, OH) at 14°C until use. Prior to dissection, the animals were weighed and anesthetized by injecting 50% (volume/body weight) of 333 mM MgCl_2 into the body cavity. To aid the isolation of individual neurons, CNS tissues were incubated for 45–60 min at 34°C in a solution containing 1% neutral protease, Dispase (100 units/mL, Worthington Biochemical Corp., Lakewood, NJ), and antibiotic cocktail (penicillin G, 100 $\mu\text{g/mL}$ streptomycin, and 100 $\mu\text{g/mL}$ gentamicin) in artificial sea water (ASW: 460 mM NaCl, 10mM KCl, 11 mM CaCl_2 , 55 mM MgCl_2 , 5 mM NaHCO_3 , pH 7.8). Following incubation, CNS tissues were rinsed in ASW/antibiotic cocktail for 20 min at 4°C, and the ganglionic sheaths surgically removed. Individual neurons from the pleural ganglia were manually isolated with electrolytically sharpened tungsten needles using pleurin *in-situ* staining images as a guide.⁴²⁻⁴³ For sampling ganglia connectives, approximately 50–100 μm -long portions

were cut using microdissection scissors. Scheme 4.2 depicts the sampling of pleurin neurons and ganglia connectives. Isolated neurons and ganglia connectives were transferred individually onto the inner side of a 0.25 mL PCR tube filled with 5 μ L of acidified methanol (90:9:1 methanol: water: glacial acetic acid) using lab-made plastic micropipettes filled with deionized water. Excess liquid was aspirated, and neurons were forced down the tube wall into the bottom of the tube by flowing extraction media pipetted from the same tube. Sample tubes with and without cells were kept on ice for the length of the isolation procedure. Sample tubes were stored at -20°C and analyzed by CE-TIMS MS within 24 h. After all individual neurons and ganglia connective samples were collected, 0.3 μ L of the extracts were sampled for peptide validation by an orthogonal technique, matrix-assisted laser desorption/ionization (MALDI)-time-of-flight (TOF) MS and the remaining extract was used for CE-TIMS MS analysis.

MALDI-TOF MS Analysis of Single Neurons

Individual pleurin-expressing neurons were isolated and analyzed by MALDI-TOF MS using methods described previously but with minor alterations.⁴³⁻⁴⁴ Briefly, each neuron was isolated from the *A. californica* pleural ganglia as described above, but instead of depositing the neuron into an extract solution, it was deposited directly onto a MALDI target plate (Bruker Corp., Billerica, MA) and excess liquid removed. The deposited neuron was then mixed with ~ 1 μ L of 50 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 50% CH_3CN with 0.1% trifluoroacetic acid and allowed to air dry. Peptide profiles were measured using an ultrafleXtreme MALDI-TOF mass spectrometer (Bruker Corp.) in positive reflection mode. The Bruker ultrafleXtreme MALDI-TOF mass spectrometer was equipped with a Smartbeam II frequency tripled Nd:YAG solid state laser. Mass spectra were manually acquired in positive reflectron mode and calibrated externally using

Bruker Peptide Mix II standards. Signals from 500–4000 laser shots fired at 1000 Hz frequency at multiple locations within each sample spot were summed into a representative cell spectrum. Mass spectra were processed using flexAnalysis 3.4 software (Bruker Corp.). Pleurin peptides were assigned by mass matching to a theoretical mass of predicted mature peptides within 150 ppm.

CE Separations

For CE, the sheath liquid was 50% methanol in water solution with 0.1% formic acid (FA). The sheath liquid was supplied through the emitter at a 750 nL/min flow rate. The CE separation was performed on a fused silica capillary, 75–80 cm, 40 μ m inner diameter and 105 μ m outer diameter (CM Scientific, Silsden, UK), with a separation voltage of +20 kV. The background solution was 0.75% FA and 25% CH₃CN in water solution.⁴⁵ The field-amplified sample injection was performed by using electrokinetic injection of the sample solution by +20 kV for 30 s. A lab-built CE–ESI source^{19, 46} was coupled to a timsTOF Pro mass spectrometer (Bruker Corp.). The emitter was grounded, and the ion guide capillary voltage of the mass spectrometer was set at 2000 V to establish the cone-jet spray. The dry nitrogen gas was set at 180°C with a flow rate of 3 L/min. The single cell extracts (~5 μ L) were dried using vacuum centrifugation and reconstituted in 2 μ L of acidified methanol; 1 μ L of extract was deposited into lab-made stainless steel nano-vials⁴⁶⁻⁴⁷ and injected into the capillary by electrokinetic injection, as described above. At the end of each set of single cell measurements, 1 μ L solutions of 10 nM ¹³C-Plrn1–Plrn3 were electrokinetically injected into the capillary to provide a comparison of the 1/k₀ values of the synthetic standards to the native peptides.

Optimization of TIMS Separations for Synthetic ^{13}C -labeled Plrn1–Plrn3 Diastereomers

Optimized TIMS parameters for the separation of synthetic ^{13}C -labeled Plrn1–Plrn3 were similar to the conditions reported previously,¹¹ with minor alterations. Following the TIMS calibration, synthetic peptide standards mixes of 1 μM ^{13}C -L-Plrn1–Plrn3, ^{13}C -D2-Plrn1–Plrn3, and ^{13}C -D3-Plrn1–Plrn3 in 50/50 v% CH_3CN /water with 0.1% formic acid (FA) were introduced to the mass spectrometer by direct infusion. The $[\text{m}^*+2\text{H}]^{2+}$ (m^* indicates monoisotopic mass of the ^{13}C -labeled peptide) ions of synthetic ^{13}C -Plrn1–Plrn3 peptides were monitored (see Table 4.1) because the $[\text{m}^*+2\text{H}]^{2+}$ ions showed the greatest differences in reduced mobility between stereoisomers. The mobility range and ramp time were the two most important parameters that affected the TIMS resolution. Increasing the ramp time and decreasing the mobility range improved the resolving powers (Figures 4.1, 4.2). The optimal separation resolution was determined by incrementally stepping the ramp time between 50–350 ms (Figure 4.1). A ramp time of 350 ms was found to be optimal and used for all subsequent measurements. To maintain the duty cycle at 100%, the accumulation time was also set to 350 ms. The TIMS resolution was further improved by narrowing the mobility range from the default $0.5 < 1/k_0 < 2.0 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ down to $0.9 < 1/k_0 < 1.2 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$ (Figure 4.2). Note, during optimization, changing the mobility range and ramp time resulted in a cumulative positive shift in the observed $1/k_{0,\text{max}}$ values of the synthetic peptides by about 0.02–0.03 $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$ (Figures 4.1A, 4.2A). This shift was not unexpected because adjusting the TIMS parameters typically demands a recalibration.^{49,60} However, a linear recalibration of the TIMS analyzer was not performed after optimizing the parameters because the calibrant masses 322.0, 922.0, and 1222.0 have mobilities outside of the mobility range of $0.9 < 1/k_0 < 1.2 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$. Therefore, because a linear recalibration could not be performed, this may have contributed to a systematic error in our calculated collisional cross section (CCS) values. For

the purposes of these studies, the observed shifts were inconsequential because we compared the extracted ion mobilograms (EIMs) of the native peptides directly to synthetic standards and the CCS values were not used for structural modeling.

TIMS MS Parameters for Single Neurons and Ganglia Connectives

The timsTOF Pro calibration was performed using a low-concentration ESI tune mix (P/N G1969-85010, Agilent Technologies, Santa Clara, CA,) as described previously.^{39,48} Nitrogen was the buffer gas for all measurements. Following the instrument calibration, the standard ESI source was replaced by a custom, lab-built CE-ESI source and the acquisition was adjusted to the optimized TIMS MS method. For data acquisition on single cells, the optimized TIMS MS method was operated with the PASEF mode off. Imex settings were set to custom resolution, set $1/k_0$ End for accumulation to $2.00 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$, ramp time was 350 ms and lock duty cycle to 100% was selected. The mobility range was optimized by narrowing the mobility range to $0.9 < 1/k_0 < 1.2 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$. For single cells, the mass spectrometer was operated in full scan mode with an m/z range of 100–1700. For TIMS MS measurements on ganglia connectives, the Plrn1–Plrn3 peptide $[\text{m}+2\text{H}]^{2+}$ ions were not detected using the full m/z scan mode that was used for single cells. We therefore analyzed the ganglia connectives using multiple reaction monitoring (MRM) mode, specifically focusing on detecting the Plrn2 peptide, and this resulted in improved signal. For MRM measurements of Plrn2, an m/z window of 813.4 ± 0.2 was monitored, corresponding to the $[\text{m}+2\text{H}]^{2+}$ of native Plrn2.

Curve Fitting of EIMs and Calculations

EIMs of Plrn1–Plrn3 ions were generated in, and exported from, DataAnalysis 5.0 (Bruker Corp.) as .txt files and loaded into Origin Pro 2019 (v9.6.0.172) for further analysis. The EIMs were first normalized to the maximum measured intensity in the mobilogram, and the intensity normalized EIMs were fit with a multiple Gaussian distribution curve using the Origin multiple peak fit function (Eq. 4.1) as follows:

$$y = y_0 + \frac{A}{w \sqrt{\frac{\pi}{4 \ln(2)}}} e^{\frac{-4 \ln(2)(x-x_c)^2}{w^2}} + \dots \quad \text{Eq. 4.1}$$

where y_0 is the base line intensity offset, A is the peak area, w is the full width at half maximum (FWHM), x is the inverse reduced mobility ($1/k_0$ in $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$), and x_c is the $1/k_0$ value at the maximum peak intensity ($1/k_{0,\text{max}}$).

After fitting the experimental EIMs with Eq. 4.1, the experimental resolving powers, R_p , were calculated for individual peaks in Plrn1–Plrn3 EIMs by using Eq. 4.2.³⁷ The experimental peak-to-peak resolving powers, R_{pp} , were calculated using Eq. 4.3.³⁷ The dispersion axis used to calculate R_p , and R_{pp} was $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$.⁴⁹

$$R_p = \frac{x}{\Delta x} = \frac{1/k_{0,\text{max}}}{w} \quad \text{Eq. 4.2}$$

$$R_{pp} = 1.18 \times \frac{x_2 - x_1}{\Delta x_2 + \Delta x_1} = 1.18 \times \frac{1/k_{0,\text{max}2} - 1/k_{0,\text{max}1}}{w_1 + w_2} \quad \text{Eq. 4.3}$$

The calculated Gaussian curve fit coefficients and R_p values for synthetic L, D2 and D3 ^{13}C -Plrn1–Plrn3 stereoisomers are reported in Table 4.2. Curve fit plots are shown in Figure 4.3.

Generating Simulated Mobility Profiles of ^{13}C -L-Plrn2 Diastereomers

The purpose of simulating EIMs of ^{13}C -L-Plrn2 and ^{13}C -D2-Plrn2 diastereomers was to calculate the approximate ratio of D2-Plrn2 to L-Plrn2 detected in cerebral-pleural ganglia connectives. Simulated EIMs were generated from experimentally measured EIMs of synthetic ^{13}C -L-Plrn2 and ^{13}C -D2-Plrn2 diastereomers. Each EIM was fit with a multiple Gaussian peak function as described above (Eq. 4.1). The fitted curves were then intensity normalized by dividing the curves by the maximum intensity. Fitted curves were then baseline subtracted and normalized by peak area.

Simulating EIMs of L-Plrn2 and D2-Plrn2 Mixtures

To simulate the mobility profiles expected for mixtures of Plrn2 diastereomers, linear combinations of simulated curves of D2-Plrn2 and L-Plrn2 were performed. Let $\bar{f}(x)$ and $\bar{g}(x)$ be the area normalized simulated EIMs describing L-Plrn2 and D2-Plrn2, respectively. Where x is the inverse reduced mobility ($1/k_0$ in $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$). Let $h(x)$ be the simulated EIMs describing the linear combination of L-Plrn2 and D2-Plrn2 as follows:

$$h(x) = \alpha \bar{f}(x) + \beta \bar{g}(x) \quad \text{Eq. 4.4}$$

where α and β are scalars to simulate the relative concentrations of L-Plrn2 and D2-Plrn2. Let $\bar{h}(x)$ be the area normalized simulated EIM describing a mixture of L-Plrn2 and D2-Plrn2. $\bar{h}(x)$ was calculated as follows:

$$\bar{h}(x) = \frac{\alpha \bar{f}(x) + \beta \bar{g}(x)}{\alpha + \beta} \quad \text{Eq. 4.5}$$

Figure 4.4. shows the simulated $\bar{f}(x)$, $\bar{g}(x)$ and an example of $\bar{h}(x)$.

Eq. 4.5 can be interpreted as follows: if $\alpha = 1$ and $\beta = 1$, then this simulates a mixture of L-Plrn2 and D2-Plrn2 being present in the sample at a 1:1 ratio. If $\alpha = 16$ and $\beta = 33$, this simulates L-Plrn2 and D2-Plrn2 being present at a 1:2.06 ratio in the sample, and so on. This assumes the ionization efficiency of L-Plrn2 and D2-Plrn2 are equal. By plotting $\bar{h}(x)$ using varying values for α and β , the expected normalized EIMs for mixtures of stereoisomers can be predicted using Eq. 4.5. For simplification, the expected mobility profiles of $\bar{h}(x)$ were generated by setting α as a constant ($\alpha = 1$), and only varying β . Figure 4.5A shows examples of the simulated mobility profiles of mixtures of L-Plrn2 and D2-Plrn2 at various values of β .

The plots of the relative peak areas for each peak (P1, P2 or P3) in $\bar{h}(x)$ as a function β are shown in Figure 4.5B. The equation describing the relative peak area (C_{sim}) for a given peak (P1, P2 or P3) $\bar{h}(x)$ function of β is given by Eq. 4.6.

$$C_{sim} = \frac{A_{sim}}{1 + \beta} + \frac{\beta B_{sim}}{1 + \beta} \quad \text{Eq. 4.6}$$

Where C_{sim} is the relative peak area of one of the peaks (P1, P2 or P3) in $\bar{h}(x)$. A_{sim} is the relative area of the peak contributed by $\bar{f}(x)$, and B_{sim} is the relative area of the peak contributed by $\bar{g}(x)$. Each curve in Figure 4.5B was fit with Eq. 4.6. Note, only the relative area of P2 has contributions from both $\bar{f}(x)$ and $\bar{g}(x)$ in this case.

The simulated calibration curve (Eq. 4.6) was used to estimate the β (proportion of D2-Plrn2 to L-Plrn2) in a biological sample. This was done by comparing the relative peak areas of P1, P2 and P3 in the measured native Plrn2 EIM, to those in the simulated calibration curve.

Estimating the Value of β for the Native Plrn2 EIM

To find the relative peak areas of each peak in the native Plrn2 EIM from the cerebro-pleural connective, the relative peak areas in the EIM of native Plrn2 from the sample were calculated by fitting the EIMs with a multi-Gaussian fit function (Eq. 4.1), and area normalizing the fitted EIM.

To determine the proportion of D2-Plrn2 to L-Plrn2 for the native Plrn2, Eq. 4.6 was solved for β , and the simulated relative peak area (C_{sim}) for a given peak (P1, P2 or P3) was replaced by the measured relative peak areas for the native Plrn2 EIM (C_{biol}) as shown in Eq. 4.7 below:

$$\beta = \frac{C_{biol} - A_{sim}}{B_{sim} - C_{biol}}, B_{sim} \neq C_{biol} \text{ and } A_{sim} \neq B_{sim} \quad \text{Eq. 4.7}$$

Where C_{biol} in Eq. 4.7 is equal to the measured relative peak area of a given peak (P1, P2 or P3) in the mobility profile of native Plrn2. β was calculated by inputting the values of A_{sim} and B_{sim} coefficients (calculated from the curve fit with Eq. 4.6), and C_{biol} into Eq. 4.7. We calculated the value of β from all three peaks in the native Plrn2 EIM (Table 4.3). Finally, the $\beta_{connective}$ estimating the relative amount D2-Plrn2 to L-Plrn2 was calculated using the averaged β values determined for each peak in the native Plrn2 EIM. See Table 4.3 for the values of A_{sim} and B_{sim} , C_{biol} , β and $\beta_{connective}$.

Figure 4.5C shows the simulated mobility profile of the mixed L-Plrn2 and D2-Plrn2 for $\beta = 0.8$ compared to the Gaussian fitted native mobility profile of Plrn2 from the cerebral-pleural ganglia connective (Figure 4.5D). The simulated and native Plrn2 EIMs were in reasonable

agreement, indicating the model used was effective for estimating the relative abundance of D2-Plrn2 to L-Plrn2 in the biological sample.

This method served to estimate the relative abundance of D2-Plrn2 in the biological sample. Ideally, an experimental calibration curve would be generated with solutions containing mixtures of diastereomers to determine the relative abundance of D2-Plrn2 in a biological sample, rather than a simulated curve. However, for the purposes of this study, a simulated curve was enough to estimate the relative abundance of D2-Plrn2.

Minimum Detectable Percentage of D2-Plrn2 from a Mobilogram

Solutions of 10 μM ^{13}C -D2-Plrn2 were diluted to concentrations of 0–25 fmol/ μL with 100 fmol/ μL of ^{13}C -L-Plrn2 in 50% MeOH, 0.1% FA. Each concentration was introduced to the mass spectrometer by direct infusion at 3 $\mu\text{L}/\text{min}$ using a standard ESI source. Mobilograms were acquired using the optimized TIMS parameters described in the main text. The mobilograms used were from the summed intensities of $[\text{m}^*+2\text{H}]^{2+} = 813.89 \pm 0.05$ over a 0.4 min acquisition time. The area of the ^{13}C -D2-Plrn2 peak at 1.06 $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$ in the EIM was calculated in Origin Pro 2019 (v9.6.0.172) and plotted vs. fmol of ^{13}C -D2-Plrn2 peptide injected. The resulting calibration curve is shown in Figure 4.6A-B. A linear fit was used to determine the detection limit and the minimum percentage of isomerized peptide that can be detected.

4.5 RESULTS AND DISCUSSION

Characterizing Plrn1–Plrn3 Standards

Before the stereochemistry of Plrn1–Plrn3 was analyzed in single cells, the synthetic ^{13}C -labeled Plrn1–Plrn3 standards were characterized by CE and TIMS MS to determine the

effectiveness of our experimental conditions for separating the diastereomers. Injections of individual and combined ^{13}C -Plrn1–Plrn3 diastereomers showed Plrn1, Plrn2, and Plrn3 sequences migrate differently, but the L-, D2-, and D3-forms of each peptide co-migrate (Figure 4.7). ^{13}C -Plrn1–Plrn3 had long migration times (20–30 min), which may have contributed to peak broadening and reduced the CE separation efficiency of the diastereomers. Overall, the CE conditions used in this study were suitable for separating peptides, but did not allow for the separation of ^{13}C -Plrn1–Plrn3 diastereomers. The subsequent TIMS separations compensated for the sub-optimal CE separations of stereoisomers.

In contrast to CE, TIMS allowed us to better distinguish between ^{13}C -Plrn1–Plrn3 diastereomers. Extracted ion mobilograms (EIMs) obtained from individual injections of the synthetic ^{13}C -Plrn1–Plrn3 standards are shown in Figure 4.8. For TIMS separations, we achieved peak resolving powers ranging from 106 to 162 for ^{13}C -Plrn1 and ^{13}C -Plrn2 diastereomers, and 37 to 142 for ^{13}C -Plrn3 diastereomers (Table 4.2). The greatest resolution was achieved for ^{13}C -Plrn2 diastereomers; Δ collisional cross section (CCS) values as low as $5 \text{ \AA}^2$ could be differentiable (Figure 4.8B).⁴⁹ The structural changes to a peptide resulting from the presence of a D-residue may differ depending on the sequence or length of the peptide, and the position and structure of the D-residue.⁵⁰ In the cases of ^{13}C -Plrn1–Plrn3, each diastereomer had a unique mobility fingerprint with varying degrees of complexity. The presence of multiple peaks in the EIMs suggests that these peptides can adopt more than one stable gas phase conformation during TIMS,^{38, 51} but exploring the exact physical explanation for why these features arose was beyond the scope of this work. It may be *cis-trans* proline isomerization (Appendix B, Figure B.5). The important takeaway is that we expected the native Plrn1–Plrn3 EIMs to exhibit similar features to the synthetic standards.

MALDI-TOF and CE-TIMS MS of Individual Neurons

Locating and isolating the target neurons for analysis was challenging. For cell isolation, pleurin neurons were located in live ganglia by referencing previously published images of pleurin neurons in preserved ganglia.^{41,42} The available reference images showed the expression of the *pleurin* gene transcript is localized to clusters of ~10 neurons in the pleural ganglia.⁴²⁻⁴³ But in the live CNS, there were no morphological features unique to pleurin neurons that distinguished them from the hundreds of surrounding neurons. Manual manipulation ensured precision isolation of solitary cells.

Precise and careful manual cell isolation was also important for reducing sample complexity and mitigating sample loss. The neurons were isolated from live ganglia in ASW, so it was important to minimize the transfer of excess ASW to avoid potential complications with electrophoretic separations and the electrospray process. Live dissected neurons were manipulated using a micropipette filled with deionized water. When performed quickly, this provided a means to wash away salts while preserving cellular integrity. Careful manual manipulation allowed us to isolate neurons without the need for any subsequent desalting procedures, thereby reducing sample handling and associated sample losses.

Pleurin neurons were initially characterized pleurin by MALDI-TOF MS to determine the peptide fingerprint expected from a typical pleurin neuron. MALDI-TOF MS spectra showed that Plrn1–Plrn3 [m+H]⁺ ions were among the most-abundant ions detected in the cells, but peptides derived from the urotensin II precursor (ApUII),⁵² *Aplysia* mytilus inhibitory peptides related peptides (AMRPS) precursor,⁵³ and thymosin beta (TMSB),⁵⁴ were also detected (Figure 4.9A). Full sequences and observed *m/z* values of peptides colocalizing with Plrn1–Plrn3 in neurons are

reported in Table 4.4. These results establish that the peptide fingerprint of pleurin neurons is largely consistent between animals, and from cell-to-cell ($N \geq 10$ neurons from $N \geq 2$ animals).

For CE-TIMS MS, isolated neurons were analyzed using the same CE conditions and TIMS MS parameters that were used for the synthetic peptides. Although the TIMS parameters were tuned for the optimal resolution of Plrn1–Plrn3 $[m+2H]^{2+}$ ions, the simultaneous detection of other colocalizing peptides was still possible. This allowed us to directly compare peptide profiles observed by CE-TIMS MS and MALDI-TOF MS. A representative base peak electropherogram (BPE) from an individual pleurin-containing neuron is shown in Figure 4.9B. Including Plrn1–Plrn3, a total of 8 peptides were identified that matched peptides also detected by MALDI-TOF MS (Figure 4.9A, B; Table 4.4). In contrast to MALDI TOF MS, however, Plrn1–Plrn3 were not the most abundant ions detected in neurons by CE-TIMS MS, but this can be explained by differences in the ionization techniques (MALDI vs. ESI).

A total of $N = 7$ neurons were isolated for CE-TIMS MS analysis, and $N = 3$ showed the characteristic peptide profile expected for a pleurin neuron. The detection of multiple expected colocalizing neuropeptides helped validate the individual cells were isolated from the intended cell population. But there was a trade-off between the number of peptides that could be simultaneously analyzed by CE-TIMS MS vs. the TIMS resolution. Ions with mobilities outside the set mobility range of our experiment ($0.9 < 1/k_0 < 1.2 \text{ V} \cdot \text{s} \cdot \text{cm}^{-2}$) were filtered out by the TIMS analyzer, and thus, did not appear in the final mass spectrum. Expanding the mobility range of our measurement may have resulted in the detection of more peptides, but at the expense of TIMS resolution.

Assignment of Native Plrn1–Plrn3 Stereochemistry

To determine the stereochemistry of Plrn1–Plrn3 in individual neurons, the mobility profiles of synthetic and native Plrn1–Plrn3 peptides were directly compared; all showed multiple isomeric peaks and closely resembled those obtained from the pure synthetic ^{13}C -L-Plrn1–Plrn3 standards (Figure 4.9C-E). An overlay of the intensity-normalized EIMs of the synthetic ^{13}C -L-Plrn1–Plrn3 and the native peptides showed significant congruency between the traces (Figure 4.9C-E). The native Plrn1–Plrn3 mobility profiles did not resemble those expected for D2- or D3-Plrn1–Plrn3 (Figure 4.8A-C), or mixtures of synthetic diastereomers (Figure 4.9F-H). Further, even though we observed the L-, D2-, and D3- ^{13}C -Plrn1–Plrn3 stereoisomers were not fully resolved by TIMS, the EIMs of the synthetic diastereomer mixtures were still distinctive from the pure all-L peptides (Figure 4.9F-H). The slight variations between the mobility fingerprints of the pure synthetic peptides and mixtures of synthetic diastereomers were essential for assigning the stereochemistry of the native neuropeptides. We emphasize the importance of our initial analysis of the synthetic ^{13}C -Plrn1–Plrn3 standards for assigning the native neuropeptide stereochemistry. The EIMs of the native Plrn1–Plrn3 showed multiple isomeric peaks, which could have easily been misinterpreted as resulting from different native peptide diastereomers. By analyzing the synthetic peptides, we established the isomeric peaks in the Plrn1–Plrn3 EIMs are likely gas phase conformers.

Based on the significant agreement between the mobility profiles of synthetic ^{13}C -L-Plrn1–Plrn3 and the native peptides ($N = 3$ sampled neurons), we concluded that negligible quantities of D2- or D3-Plrn1–Plrn3 contributed to the native peptide EIMs. In a separate experiment, we determined a minimum percentage of isomerized peptide that can be detected from an EIM by spiking solutions of 100 fmol/ μL synthetic ^{13}C -L-Plrn2 with synthetic ^{13}C -D2-Plrn2 at

concentrations ranging from 0–25 fmol/μL. We were able to detect < 2% of ^{13}C -D2-Plrn2 for an amount of 120 fmol ^{13}C -L-Plrn2 injected (Figure 4.6). Similar limits were expected for the ^{13}C -D3-Plrn2, as well as the ^{13}C -Plrn1 and ^{13}C -Plrn3 diastereomers. We concluded that > 98% of the Plrn1–Plrn3 detected in pleurin neurons exist in the all-L configurations.

We had initially predicted the all-L forms of pleurin-derived peptides would colocalize with the D2 and/or D3 forms in individual neurons, but our stereochemical analysis did not support our initial expectation. This expectation was based on our previous work, where we reported that more than 50% of Plrn1–Plrn3 exist as either the D2 or D3 forms in whole ganglia homogenates.¹¹ But neural ganglia are complex anatomical structures having a conglomerate of cell types encircling a large, synaptically dense region called the neuropil. Thus, the chemical complexity of an entire ganglion homogenate is substantially greater than an individual neuron. Consequently, the relative abundance of DAACPs in whole ganglia homogenates is expected to differ compared to single cells. For example, the relative abundance of another known DAACP in *A. californica*, NdWF-NH₂, has been previously quantified in both single cells and bulk tissue extracts. It was found that the relative abundance of NdWF-NH₂ in whole ganglia homogenates was between 60–80%,⁵⁵ whereas in single cells it was closer to 90%.^{13, 27}

Our finding that D2- or D3-Plrn1–Plrn3 were not detectable in individual pleurin neurons is distinct from previous observations indicating significant conversion of NWF-NH₂ to NdWF-NH₂ in individual neurons from *A. californica*.^{13, 27} We considered whether experimental factors such as ion suppression, matrix effects, electrokinetic injection bias, or selective adsorption of DAACPs to the surfaces of the capillary, tubes, pipette tips etc., prevented the detection of isomerized Plrn1–Plrn3 in single cell samples. However, our parallel analysis on synthetic ^{13}C -Plrn1–Plrn3 diastereomers excluded these factors.

A possible explanation for the absence of isomerized Plrn1–Plrn3 from individual pleurin neurons is that the isomerized peptides localize in the neuronal processes (axons) rather than the cell bodies. Immunohistochemical staining with stereospecific antibodies has shown the subcellular distribution of DAACPs in neurons from the crayfish species, *Orconectes limosus*.⁵⁶ These studies revealed the isomerization of the crustacean hyperglycemic hormone (CHH) to D3-CHH increased distally from cell bodies to the axon terminal at the release site.⁵⁶ The latter observations have shaped the current understanding of L- to D-residue isomerization as a late stage PTM that occurs gradually during anterograde neuropeptide transport.¹⁻² In *A. californica*, neuropeptides can be transported several centimeters from cell bodies along cellular processes to the sites of release, at rates of 100 mm/day.⁵⁷⁻⁵⁸ Because L- to D-residue isomerization has previously been shown to occur gradually during neuropeptide transport in crustaceans, it is plausible that the isomerization of pleurin-derived peptides in *A. californica* occurs predominantly along cellular processes. When neurons are isolated from *A. californica*, the fragile neuronal processes detach from the cell body and remain in the ganglion. This is depicted in Scheme 4.2, which shows an isolated pleurin neuron and its neuronal process (in red) remaining in the ganglion. By isolating the cell bodies without the cellular processes, we sampled and detected newly synthesized all-L peptides as opposed to the slower forming isomerized products, D2- and/or D3-Plrn1–Plrn3.

CE–TIMS MS of Ganglia Connectives

Our anatomical knowledge of pleurin neurons in the *A. californica* CNS allowed us to design an experiment to measure the Plrn1–Plrn3 stereochemistry in cellular processes. It has been shown that cellular process from pleurin neurons located in the pleural ganglia project to adjacent

ganglia (either cerebral, abdominal, or pedal) via ganglia connectives.⁵⁹ Scheme 4.2 illustrates a neuronal process extending to a neural connective. These ganglia connectives serve as “highways” for delivering secretory vesicles containing neuropeptides to their release sites.⁵⁸ Theorizing that the isomerization of pleurin-derived peptides occurs in the cellular processes, we expected the isomerized pleurin-derived peptides to accumulate in the ganglia connectives. To validate our hypothesis, we used CE–TIMS MS to characterize the stereochemistry of pleurin-derived peptides in ganglia connectives. Cerebro-pleural connectives were chosen for CE–TIMS MS analysis because sample screening by MALDI-TOF MS indicated that the pleurin-derived peptides were among the most abundant ions in these samples (Figure 4.10) However, while Plrn1, Plrn2 and Plrn3 were all detected in the cerebro-pleural connectives by MALDI-TOF MS, only the stereochemistry of Plrn2 was successfully measured by CE–TIMS MS in the samples (see the Experimental Section above: TIMS MS Parameters for Single Neurons and Ganglia Connectives).

Figure 4.11 shows the EIMs obtained from synthetic ¹³C-L-Plrn2 and ¹³C-D2-Plrn2 standards compared to an EIM obtained for native Plrn2 in the cerebro-pleural connective. As anticipated, native D2-Plrn2 was colocalized with all-L Plrn2 in extracts from the cerebro-pleural connectives (N = 3 sampled sections) as indicated by the presence of an additional peak at $1/k_0 = 1.06 \text{ V} \cdot \text{s} \cdot \text{cm}^{-2}$ that was not observed in the EIMs obtained from single neuron samples (Figure 4.11). No peaks were detected at the $1/k_0$ values expected for D3-Plrn2 in the EIM from the cerebro-pleural connectives (N = 3 sampled sections), but this was not surprising because in our previous work,¹¹ D3-Plrn2 was only detected in buccal ganglia extracts but not in either pleural or cerebral ganglia extracts. Using the ¹³C-L-Plrn2 and ¹³C-D2-Plrn2 TIMS measurements shown in Figure 4.11, we estimated the ratio of D2-Plrn2 to L-Plrn2 in the connectives to be approximately 0.8:1. This corresponds to 44% D2-Plrn2 and 66% L-Plrn2 in the ganglia connective. The presence

of D2-Plrn2 in the cerebro-pleural connective supports our revised hypothesis that isomerized forms of pleurin peptides accumulate in cellular projections rather than cell bodies, and this in agreement with published findings from the crayfish.⁵⁶

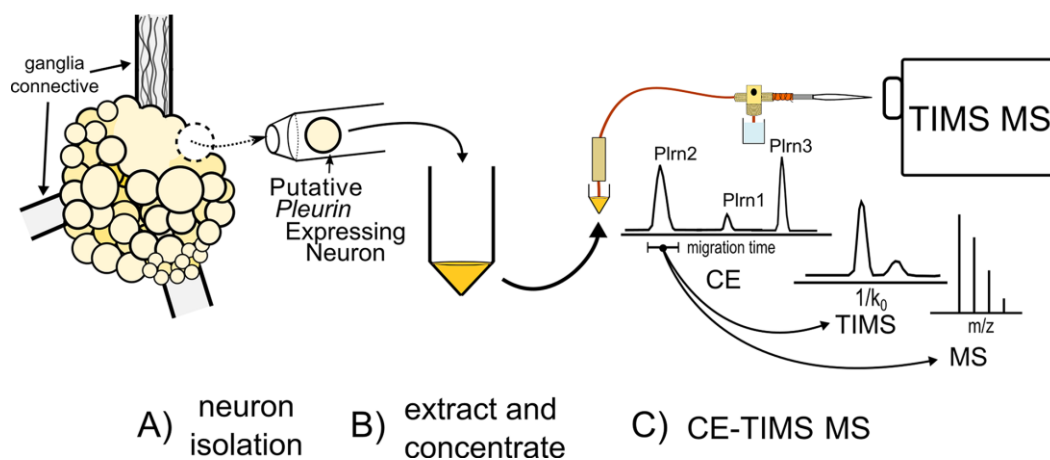
Our results show that the L-Plrn1–Plrn3 peptides in the *A. californica* CNS can be spatially separated from their DAACP counterparts. The localization of isomerized Plrn1–Plrn3 within the central nerves indicates that in our previous work,¹¹ a majority of the D2-and/or D3-Plrn1–Plrn3 we detected in ganglia homogenates may have been from the ganglia connectives and/or neuropil rather than the neuron cell bodies. This biological insight demonstrates the significant advantages of performing measurements at the single cell level as opposed to bulk tissue.

4.6 CONCLUSION

This work has established CE–TIMS MS as an effective technique for the analysis of peptides and peptide diastereomers at the single cell level. The combination of CE and TIMS MS enabled the stereochemical characterization and relative quantitation of peptide diastereomers in single cells and nerve tissue. Direct comparisons of the mobility profiles of synthetic and endogenous peptides allowed us to unambiguously assign native neuropeptide stereochemistry. Our stereochemical analysis of Plrn1–Plrn3 in individual neurons indicated L-Plrn1–Plrn3, but not D2- and/or D3-Plrn1–Plrn3, are localized in the cell bodies. In contrast, D2-Plrn2 comprised 44% of the Plrn2 within the cerebro-pleural ganglia connectives. These observations are consistent with DAACP localization patterns observed in crustacean neurons. Future CE–TIMS MS studies will focus on applying the protocols reported here to investigate the effects of external chemical stimuli on the levels of DAACPs in individual pleurin neurons in cell culture. CE–TIMS MS may also be useful for measuring other known DAACPs in *A. californica* or identifying novel, low-abundance

DAACPs present only in rare cell-types. The extension of these CE–TIMS MS protocols to studying DAACPs in single cells from a variety of species could shed further light on the molecular mechanisms underlying L- to D-residue isomerization.

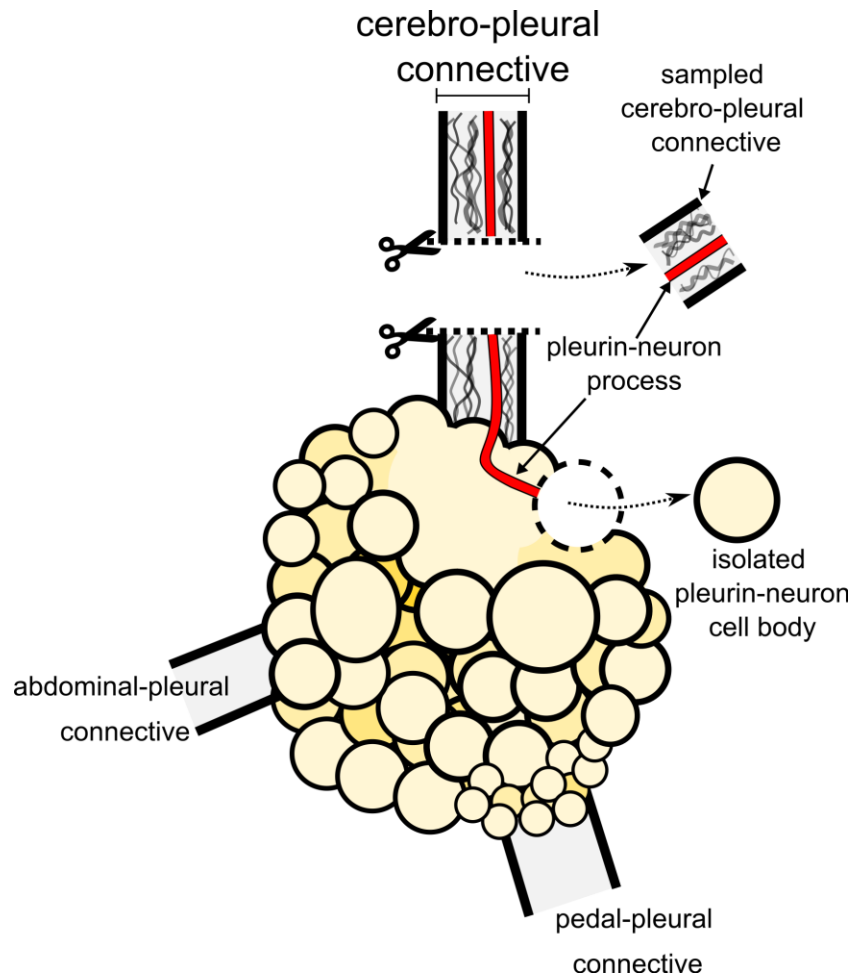
4.7 FIGURES AND DATA



Scheme 4.1. The workflow used to measure neuropeptide stereochemistry in individual neurons.

Table 4.1. Native and synthetic pleurin precursor-derived peptide sequences and m/z values monitored. D-residue is indicated by single letter amino acid code preceded by a lowercase “d” (**dX**, bold font and underlined). ^{13}C -labeled Gly residues (where one ^{13}C is at the α -carbon) are indicated by “***G***” (***G*** italicized and bold font). Note, m^* indicates the monoisotopic mass of the ^{13}C -Gly labeled synthetic peptides.

Name	Sequence	Native $[m+2H]^{2+}$	Synthetic $[m^*+2H]^{2+}$
L-Plrn1	MFYTK G SDSDYPRI-NH ₂	839.9	840.4
D2-Plrn1	M <u>dF</u> YTK G SDSDYPRI-NH ₂	839.9	840.4
D3-Plrn1	MF <u>dY</u> TK G SDSDYPRI-NH ₂	839.9	840.4
L-Plrn2	SFYTTGN G NHYPRI-NH ₂	813.4	813.9
D2-Plrn2	S <u>dF</u> YTTGN G NHYPRI-NH ₂	813.4	813.9
D3-Plrn2	SF <u>dY</u> TTGN G NHYPRI-NH ₂	813.4	814.4
L-Plrn3	GIFTQSAY G SYPRV-NH ₂	772.9	773.4
D2-Plrn3	<u>GdI</u> FTQSAYGSYPRV-NH ₂	772.9	773.4
D3-Plrn3	<u>GldF</u> TQSAYGSYPRV-NH ₂	772.9	773.9



Scheme 4.2. Right pleural ganglion (not drawn to scale) illustrating the sampling of an individual neuron and cerebro-pleural ganglia connective. An isolated pleurin neuron is shown removed from the cell body. A pleurin neuronal process is depicted in red, extending through the cerebro-pleural connective.

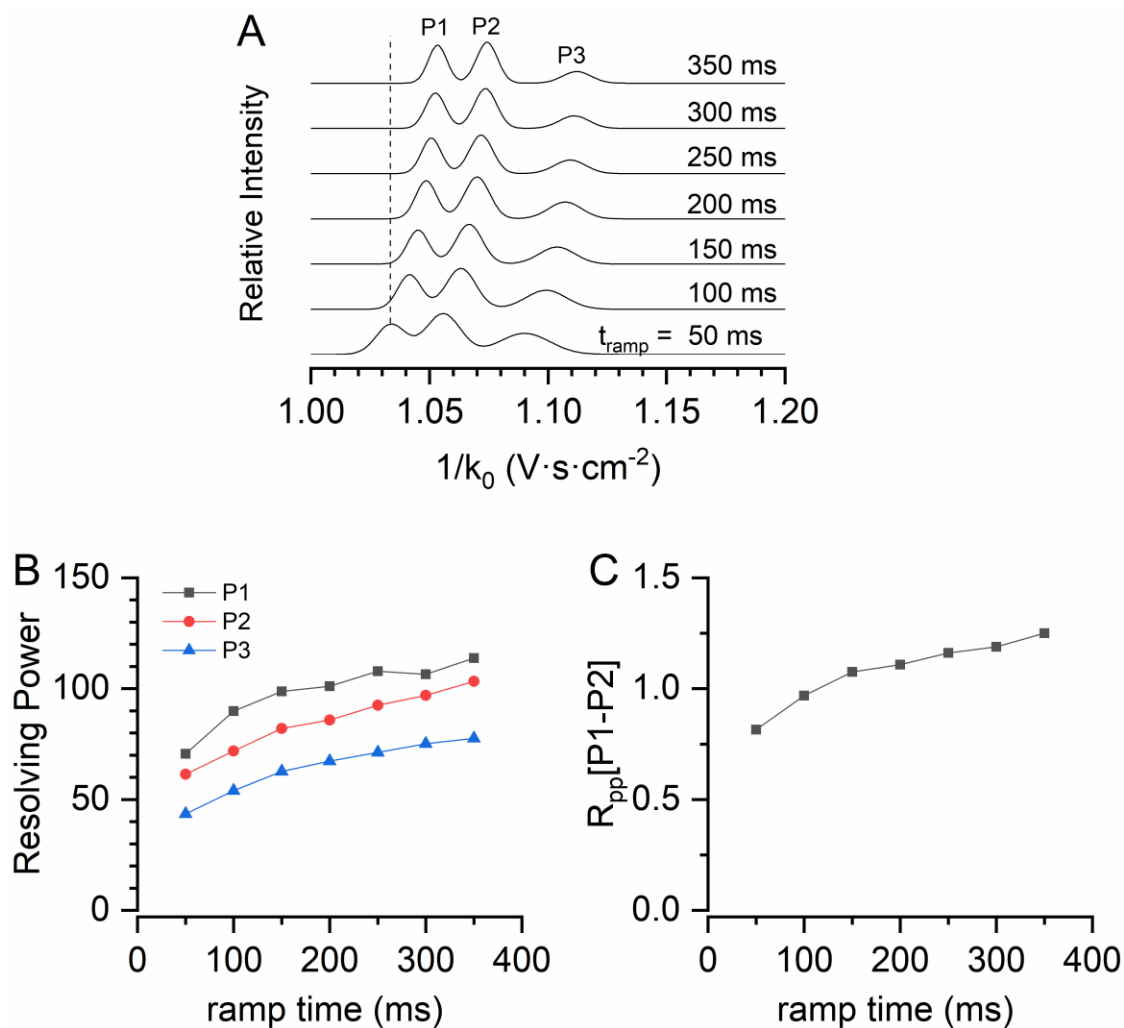


Figure 4.1. TIMS optimization of ramp time using $^{13}\text{C-L-Plrn2}$ and $^{13}\text{C-D2-Plrn2}$ standards. (A) Stacked intensity-normalized, Gaussian-fitted EIMS of a 1:1 mixture of $^{13}\text{C-L-Plrn2}$ and $^{13}\text{C-D2-Plrn2}$, $[\text{m}^*+2\text{H}]^{2+}$ ions at varying ramp times (t_{ramp}) as indicated next to each trace. The dashed line is present to emphasize the shift in the $1/k_{0,\text{max}}$ observed as the ramp time is changed. (B) Calculated resolving powers of peaks P1, P2, and P3 from panel A as plotted at varying ramp times. Resolving powers (R_p) were calculated using Eq. 4.2. (C) Peak-Peak resolving powers (R_{pp}) of P1 ($^{13}\text{C-D2-Plrn2}$) and P2 ($^{13}\text{C-L-Plrn2}$) as a function of ramp time. R_{pp} values were calculated using Eq. 4.3.

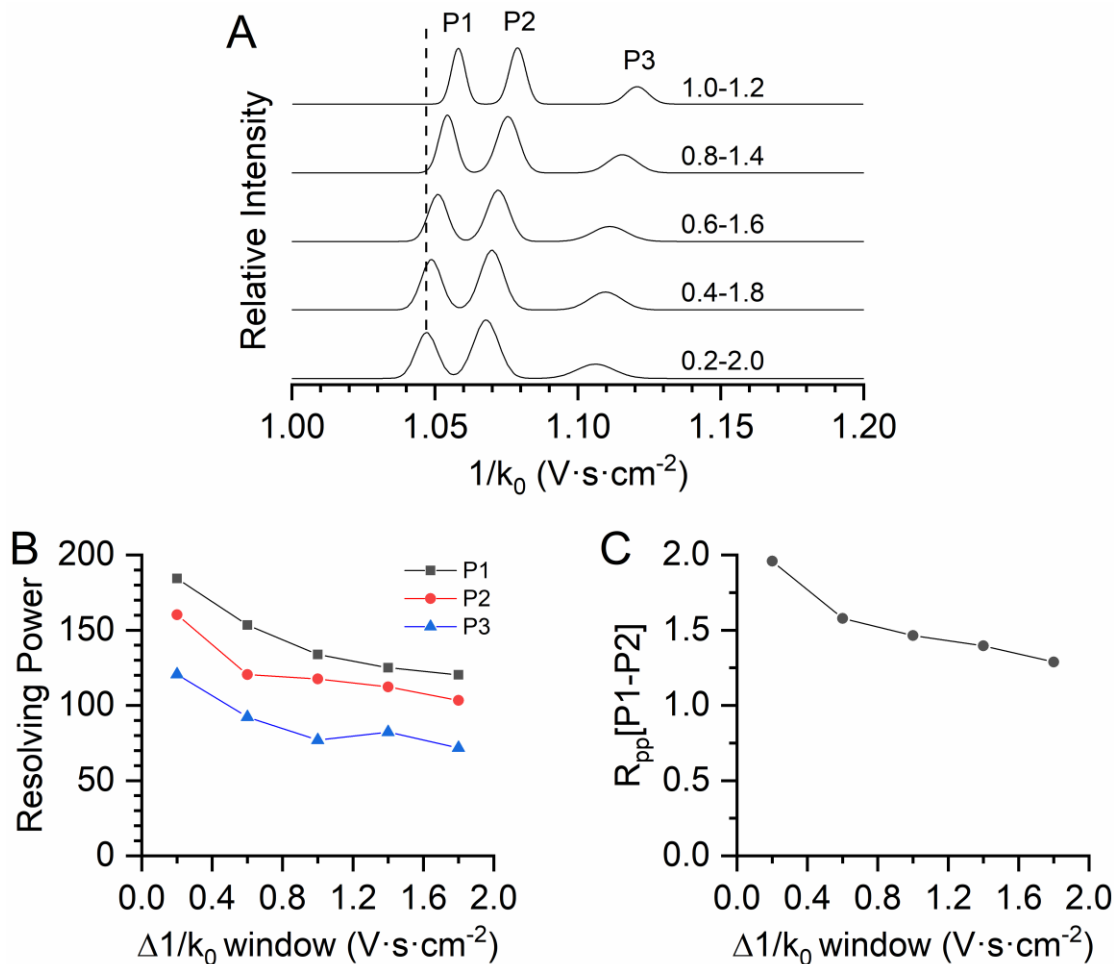


Figure 4.2. TIMS parameter optimization of mobility range using ^{13}C -L-Plrn2 and ^{13}C -D2-Plrn2 standards. (A) Stacked intensity-normalized, Gaussian-fitted EIMs of a 1:1 mixture of ^{13}C -L-Plrn2 and ^{13}C -D2-Plrn2 $[\text{m}^*+2\text{H}]^{2+}$ ions at various $1/k_0$ window ranges as indicated on the plots. The $1/k_0$ range in $\text{V}\cdot\text{s}\cdot\text{cm}^{-2}$ is displayed next to each plot. The dashed line is present to emphasize the shift in the $1/k_{0,\text{max}}$ observed as the window changed. (B) Calculated resolving powers of peaks P1, P2, and P3 from panel A as a function of $1/k_0$ window. Resolving powers were calculated using Eq. 4.2. (C) Peak-Peak resolving powers (R_{pp}) of P1 (^{13}C -D2-Plrn2) and P2 (^{13}C -L-Plrn2) as a function of $1/k_0$ window. R_{pp} values were calculated using Eq. 4.3.

Table 4.2. EIM Gaussian curve fit coefficients $\pm$ standard error (SE) from Origin Pro 2019 (as defined in Eq. 4.1) and the peak resolving powers (Rp) calculated for synthetic ^{13}C -Plrn1-Plrn3 diastereomers. Curve fit coefficients and Rp correspond to plots shown in Figure 4.3.

Peptide	Peak	R_p	Fit Coefficients				Adj. R^2
			$y_0 \pm \text{SE}$	$xc \pm \text{SE}$	$A \pm \text{SE}$	$w \pm \text{SE}$	
^{13}C -L-Plrn1	P1	136	$0.0122 \pm 8 \times 10^{-4}$	$1.10023 \pm 3 \times 10^{-5}$	$0.00817 \pm 3 \times 10^{-5}$	$0.00807 \pm 7 \times 10^{-5}$	0.9854
	P2	143	$0.0122 \pm 8 \times 10^{-4}$	$1.1135 \pm 1 \times 10^{-4}$	$0.00177 \pm 6 \times 10^{-5}$	$0.0078 \pm 3 \times 10^{-4}$	
^{13}C -D2-Plrn1	P1	135	$0.0084 \pm 8 \times 10^{-4}$	$1.09776 \pm 3 \times 10^{-5}$	$0.00781 \pm 6 \times 10^{-5}$	$0.00813 \pm 7 \times 10^{-5}$	0.9810
^{13}C -D3-Plrn1	P1	126	$0.0077 \pm 7 \times 10^{-4}$	$1.07027 \pm 3 \times 10^{-5}$	$0.00890 \pm 6 \times 10^{-5}$	$0.00853 \pm 6 \times 10^{-5}$	0.9879
^{13}C -L-Plrn2	P1	149	$0.0060 \pm 4 \times 10^{-4}$	$1.09290 \pm 1 \times 10^{-5}$	$0.00767 \pm 3 \times 10^{-5}$	$0.00733 \pm 3 \times 10^{-5}$	0.9965
	P2	107	$0.0060 \pm 4 \times 10^{-4}$	$1.13523 \pm 4 \times 10^{-5}$	$0.00435 \pm 3 \times 10^{-5}$	$0.01064 \pm 9 \times 10^{-5}$	
^{13}C -D2-Plrn2	P1	162	$0.0089 \pm 5 \times 10^{-4}$	$1.07394 \pm 2 \times 10^{-5}$	$0.00675 \pm 3 \times 10^{-5}$	$0.00662 \pm 4 \times 10^{-5}$	0.9911
	P2	118	$0.0089 \pm 5 \times 10^{-4}$	$1.1011 \pm 1 \times 10^{-4}$	$0.00168 \pm 4 \times 10^{-5}$	$0.0093 \pm 3 \times 10^{-4}$	
^{13}C -D3-Plrn2	P1	148	$0.0055 \pm 5 \times 10^{-4}$	$1.05211 \pm 5 \times 10^{-5}$	$0.00286 \pm 4 \times 10^{-5}$	$0.0071 \pm 1 \times 10^{-4}$	0.9920
	P2	151	$0.0055 \pm 5 \times 10^{-4}$	$1.06305 \pm 2 \times 10^{-5}$	$0.00703 \pm 4 \times 10^{-5}$	$0.00702 \pm 5 \times 10^{-5}$	
^{13}C -L-Plrn3	P1	101	$0.0003 \pm 7 \times 10^{-4}$	$1.02231 \pm 8 \times 10^{-5}$	$0.0039 \pm 2 \times 10^{-4}$	$0.0101 \pm 2 \times 10^{-4}$	0.9977
	P2	38	$0.0003 \pm 7 \times 10^{-4}$	$1.0425 \pm 1 \times 10^{-4}$	$0.0272 \pm 4 \times 10^{-4}$	$0.0275 \pm 4 \times 10^{-4}$	
	P3	98	$0.0003 \pm 7 \times 10^{-4}$	$1.05852 \pm 5 \times 10^{-5}$	$0.0063 \pm 3 \times 10^{-4}$	$0.0108 \pm 2 \times 10^{-4}$	
^{13}C -D2-Plrn3	P1	142	$0.0006 \pm 4 \times 10^{-4}$	$1.0196 \pm 1 \times 10^{-4}$	$0.006 \pm 1 \times 10^{-3}$	$0.0072 \pm 2 \times 10^{-4}$	0.9976
	P2	97	$0.0006 \pm 4 \times 10^{-4}$	$1.025 \pm 1 \times 10^{-3}$	$0.005 \pm 1 \times 10^{-3}$	$0.010 \pm 1 \times 10^{-3}$	
	P3	108	$0.0006 \pm 4 \times 10^{-4}$	$1.04052 \pm 3 \times 10^{-5}$	$0.00677 \pm 6 \times 10^{-5}$	$0.00960 \pm 8 \times 10^{-5}$	
^{13}C -D3-Plrn3	P1	142	$0.0011 \pm 4 \times 10^{-4}$	$1.0386 \pm 1 \times 10^{-4}$	$0.0049 \pm 8 \times 10^{-4}$	$0.0073 \pm 2 \times 10^{-4}$	0.9990
	P2	99	$0.0011 \pm 4 \times 10^{-4}$	$1.0441 \pm 5 \times 10^{-4}$	$0.0074 \pm 8 \times 10^{-4}$	$0.01056 \pm 5 \times 10^{-4}$	
	P3	91	$0.0011 \pm 4 \times 10^{-4}$	$1.06239 \pm 2 \times 10^{-5}$	$0.01127 \pm 4 \times 10^{-5}$	$0.01169 \pm 5 \times 10^{-5}$	

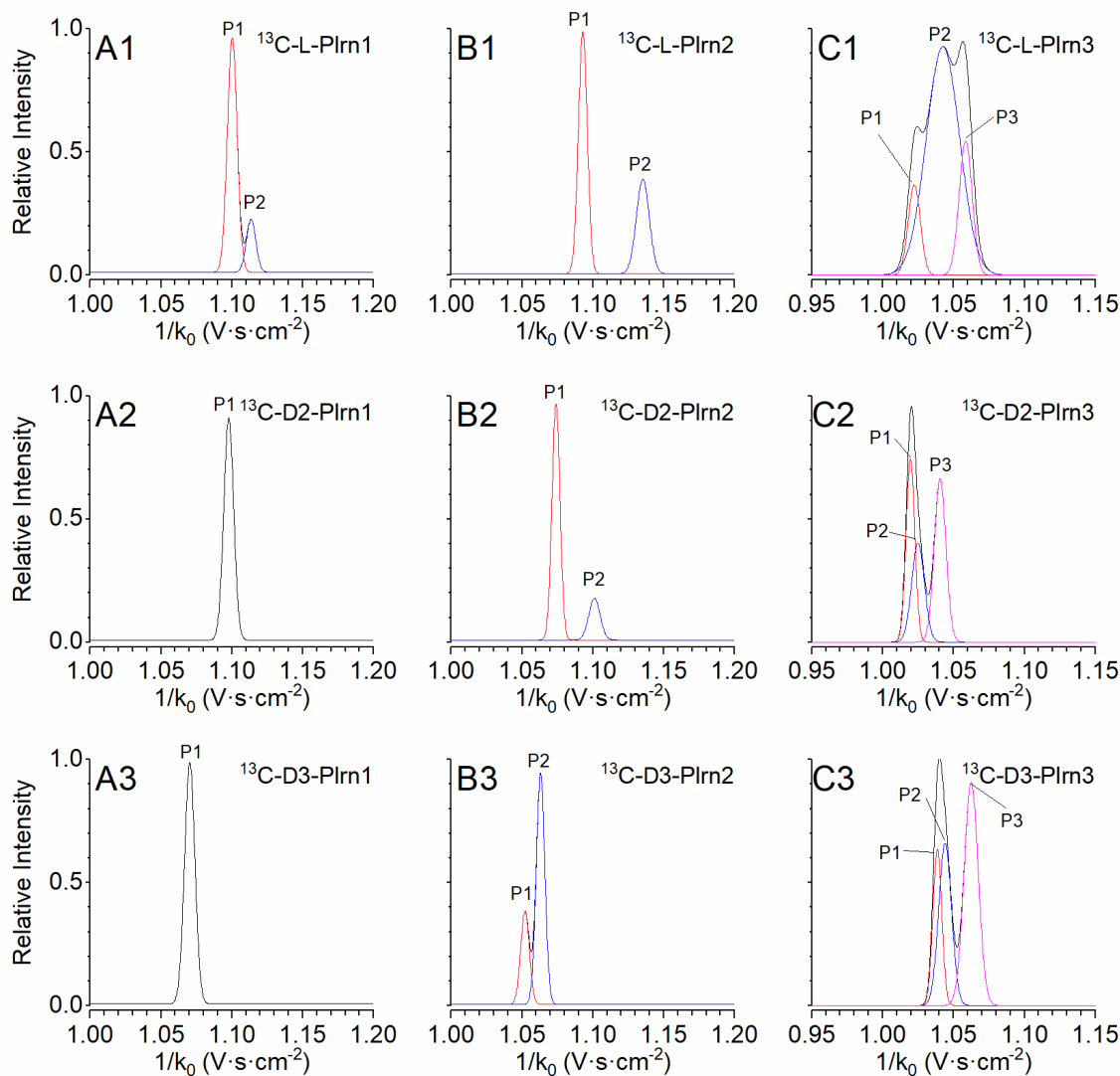


Figure 4.3. Gaussian curve fit plots of synthetic ^{13}C -Plrn1–Plrn3 standards EIMs. Peaks labeled according to the Origin function output, in order from left to right: P1 (red), P2 (blue), P3 (purple), and the cumulative curve fits are shown in black. For peptides with only one peak (panels A2 and A3), P1 is the same as the cumulative fit and shown in black instead of red. The cumulative curve fit may not be visible in all plots because P1–P3 overlay the cumulative fit.

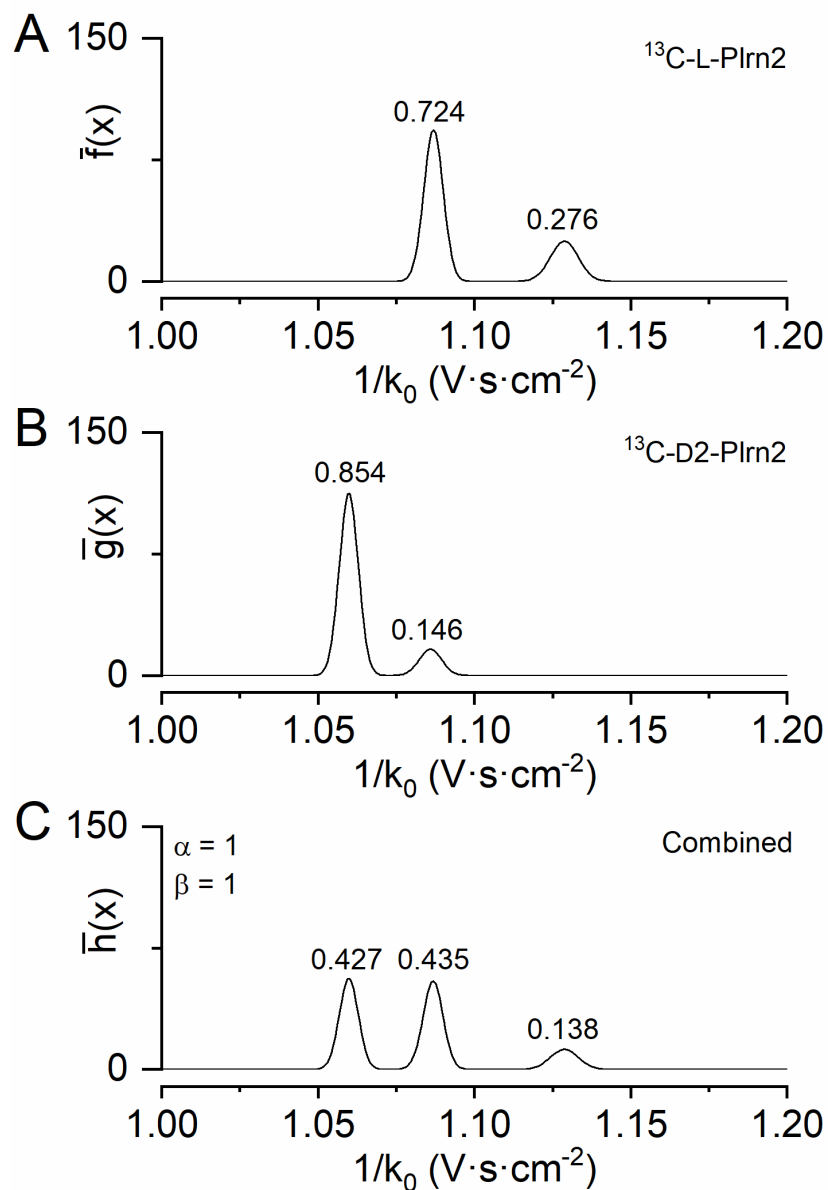


Figure 4.4. Area normalized simulated EIMs of $^{13}\text{C-L-Plrn2}$ and $^{13}\text{C-D2-Plrn2}$. The relative peak areas are displayed above each peak. (A) Simulated $^{13}\text{C-L-Plrn2}$ EIM, $\bar{f}(x)$, (B) simulated $^{13}\text{C-D2-Plrn2}$ EIM, $\bar{g}(x)$, and (C) simulated combined $^{13}\text{C-L-Plrn2}$ and $^{13}\text{C-D2-Plrn2}$ EIM, $\bar{h}(x)$ when $\alpha = 1$ and $\beta = 1$.

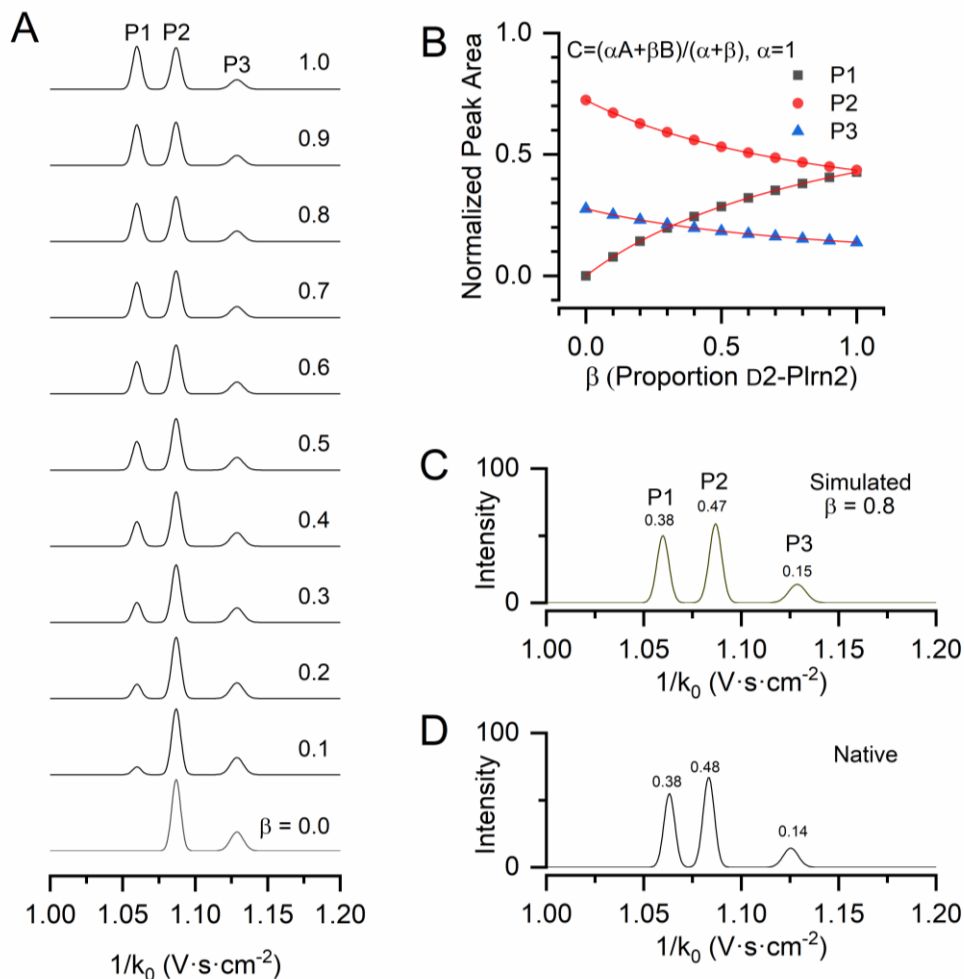


Figure 4.5. Estimation of the relative abundance of D2-Plrn2 in the cerebro-pleural ganglia connective using simulated EIMs. (A) Simulated EIMs of combinations of L-Plrn2 and D2-Plrn2 with increasing β (proportion of D2-Plrn2). The value of β is shown in the plots. (B) Relative peak areas of P1, P2, and P3 in simulated EIMs shown in panel A as a function of β . Each of the plots is fitted to Eq. 4.6 (red traces). (C) Simulated EIM of L-Plrn2 and D2-Plrn2 mixture where β was set to 0.8 (0.8:1 ratio). Relative peak areas are displayed above each peak. (D) Gaussian curve fitted, and area normalized EIM of Plrn2 obtained from the cerebro-pleural ganglia connective. Relative peak areas are displayed above each peak. The D2-Plrn2:L-Plrn2 ratio was estimated to be 0.8:1. Thus the relative percentage of D2-Plrn2 is $0.8/(1+0.8) \times 100 = 44.4\%$.

Table 4.3. Calculated relative abundances of D2-Plrn2 in the cerebro-pleural connective. Curve fit coefficients, B_{sim} and A_{sim} , calculated by fitting Eq. 4.6 to the plot of β vs. the relative peak area, and the β value(s) for P1, P2, and P3 in the cerebro-pleural connective, calculated by plugging the measured relative peaks areas (C_{biol}) into Eq. 4.7. Note, P1, P2 and P3 correspond to labels in Figure 4.5. Notes: A_{sim} for P1, and B_{sim} for P3 should theoretically be 0. Assumptions: 1) Ionization efficiencies of Plrn2 diastereomers are equal. 2) Intensities/areas of the peaks in the EIMs of the stereoisomers scale linearly with concentration. 3) If multiple peaks are present in the EIM of the diastereomer, the ratio of the peaks remains constant and does not change with concentration. The TIMS analyzer was not saturated.

Peak	A_{sim}	B_{sim}	Cerebro-Pleural Connective Plrn2 Measured Peak Area C_{biol}	β calculated
P1	-4.36×10^{-5}	0.855	0.38	0.8 ₀
P2	0.724	0.146	0.48	0.7 ₃
P3	0.27	-8.57×10^{-4}	0.14	0.9 ₆
$\beta_{connective}$ (Average $\pm$ std dev)				0.8 ₃ $\pm$ 0.1 ₂

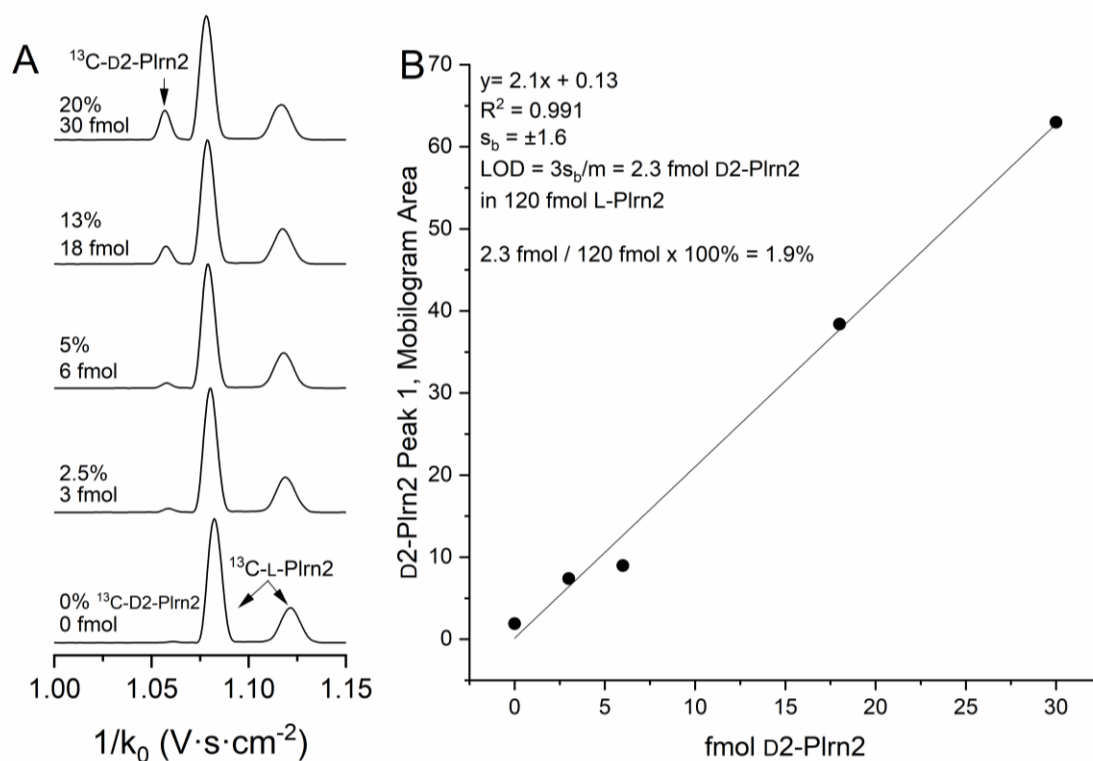


Figure 4.6. Determination of a minimum detectable percentage of $^{13}\text{C-D2-Plrn2}$. (A) Normalized EIMs of $^{13}\text{C-L-Plrn2}$ (constant 120 fmol) with varying amounts of $^{13}\text{C-D2-Plrn2}$ as indicated on the plot. (B) Linear calibration curve of the $^{13}\text{C-D2-Plrn2}$ peak area as a function the fmol $^{13}\text{C-D2-Plrn2}$. The detection limit for $^{13}\text{C-D2-Plrn2}$ was found to be 2.3 fmol.

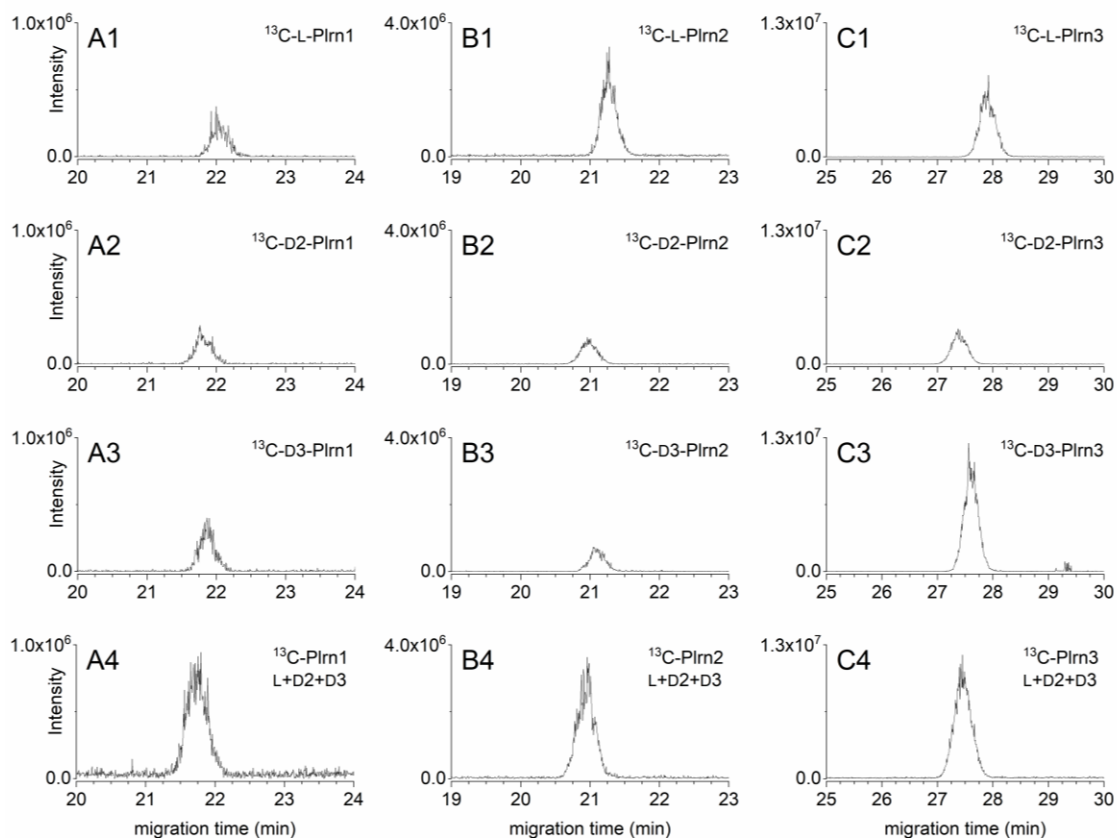


Figure 4.7. CE migration time analysis on pure and mixed synthetic ^{13}C -Plrn1-Plrn3 diastereomers. (A1) ^{13}C -L-Plrn1, (A2) ^{13}C -D2-Plrn1, and (A3) ^{13}C -D3-Plrn1. (B1) ^{13}C -L-Plrn2, (B2) ^{13}C -D2-Plrn2, and (B3) ^{13}C -D3-Plrn2. (C1) ^{13}C -L-Plrn3, (C2) ^{13}C -D2-Plrn3, and (C3) ^{13}C -D3-Plrn3. (A4-C4) Co-migration of 1:1:1 mixtures (10 nM ea.) of L, D2 and D3 (A4) ^{13}C -Plrn1, (B4) ^{13}C -Plrn2, and (C4) ^{13}C -Plrn3 by CE. All plots show extracted ion electropherograms (EIEs) of ^{13}C -Plrn1-Plrn3 $[\text{m}^*+2\text{H}]^{2+}$ ions. Each row corresponds to a separate injection. EIEs were generated from the summed intensity of $[\text{m}^*+2\text{H}]^{2+}$ ions over the mobility range of $0.9 < 1/k_0 < 1.2 \text{ V}\cdot\text{s}\cdot\text{cm}^{-2}$.

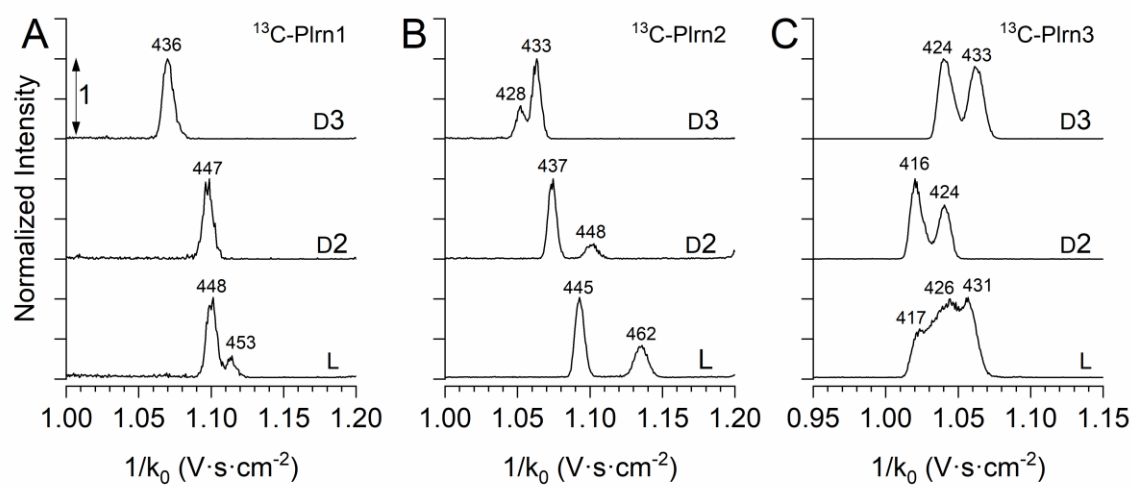


Figure 4.8. Differentiating synthetic neuropeptide stereoisomers by TIMS. Stacked EIMs of the $[\text{m}^*+2\text{H}]^{2+}$ ions of pure (A) ^{13}C -L-Plrn1, ^{13}C -D2-Plrn1 and ^{13}C -D3-Plrn1; (B) EIMs of ^{13}C -L-Plrn2, ^{13}C -D2-Plrn2 and ^{13}C -D3-Plrn2; (C) EIMs of ^{13}C -L-Plrn3, ^{13}C -D2-Plrn3 and ^{13}C -D3-Plrn3. The CCS values (in $\text{\AA}^2$) are displayed above each peak. Individual electropherograms corresponding to each peptide are shown in Figure 4.7.

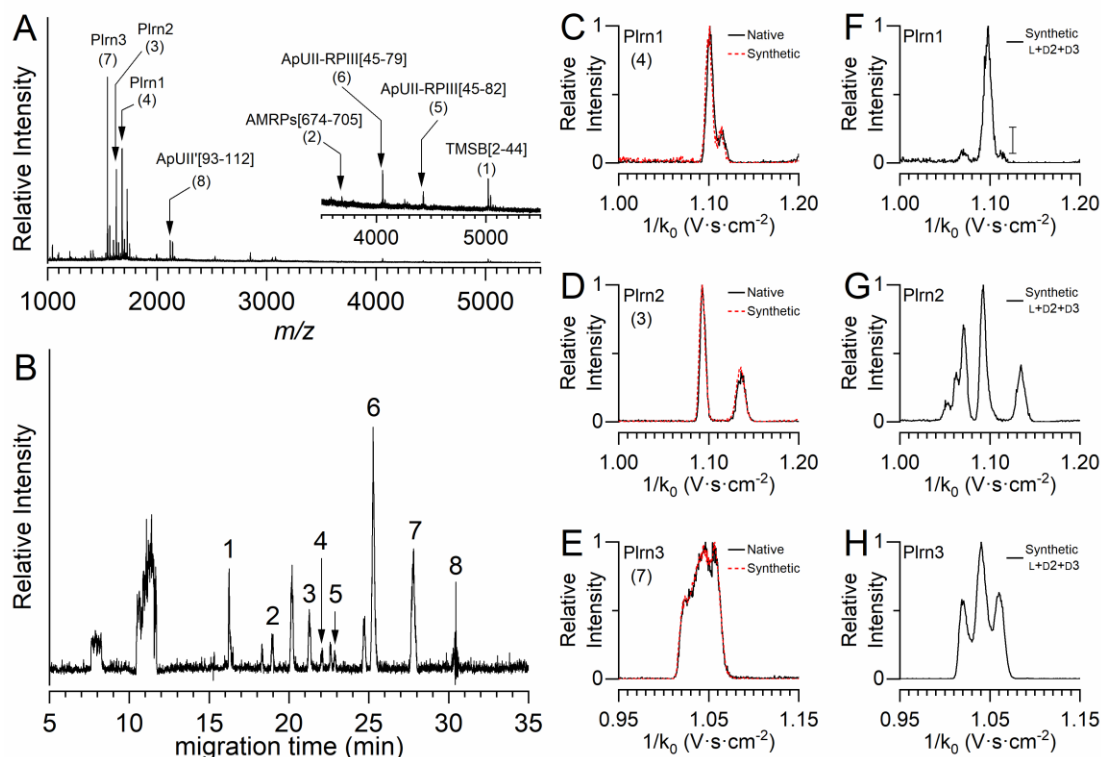


Figure 4.9. Single cell analysis of pleurin neurons by MALDI-TOF MS and CE-TIMS MS. (A) Representative MALDI-TOF MS spectrum of an individual pleurin neuron and native peptide identifications. Inset in panel A is the zoomed view of peaks between m/z 3500–5500. (B) Representative CE-TIMS MS base peak electropherogram (m/z 550–1200, mobility range 0.9–1.2 $V \cdot s \cdot cm^{-2}$) obtained from a different single pleurin neuron. Peaks matching to peptides also identified by MALDI MS are numbered according to Panel A. Unlabeled peaks correspond to unknown ions. (C–E) Assignment of native Plrn1–Plrn3 stereochemistry in the pleurin neuron by TIMS. Plots show overlaid EIMs of native (C) Plrn1, (D) Plrn2, and (E) Plrn3 with the corresponding synthetic ^{13}C -labeled all-L standard. Note that the ^{13}C -L-Plrn1–Plrn3 EIMs were replotted from Figure 1 for comparison purposes. (F–H) EIMs of 1:1:1 mixtures of synthetic ^{13}C -labeled L-, D2-, D3- (F) Plrn1, (G) Plrn2 and (H) Plrn3. EIMs in panels C–H are of the $[m+2H]^{2+}$ or $[m^*+2H]^{2+}$ ions. The vertical bar in (H) serves to highlight the differences between the mobility fingerprints of the pure L-Plrn1 in panel C, and the combined L-, D2-, and D3-Plrn1.

Table 4.4. Peptide sequences detected in pleurin neurons by CE–TIMS MS and MALDI MS described in Figure 4.9. Sequences were determined by mass matching to predicted peptide masses and published data.

#	Precursor/N ame	Sequence	Theor. mass	TIMS MS <i>m/z</i> obs.	z	ppm ^a	MALDI [m+H] ⁺ Obs.	ppm ^b	PTM
1	TMSB[2-44]	SDKHDKPDISEVTKFDKSKLKKT- ETHEKNPLPTKETIDQEKQG	5018.642	628.332	8	-2.7	5019.95	-61.5	Acetyl ation
2	AMRPs[674- 705]	GEDVSERDYAQLLEALSRLQAI- KQIKARIQNE	3683.996	737.799	5	-0.5	3685.31	-85.2	None
3	pleurin/Plm2	SFYTTGNGNHYPRI-NH ₂	1624.790	813.396	2	-2.1	1626.07	-174.7	Amida tion C-term
4	pleurin/Plm1	MFYTKGSDSDYPRI-NH ₂	1677.797	839.901	2	-2.6	1679.07	-161.2	Amida tion C-term
5	apUII-RPII[45- 82]	SMLLSKYMALMNRLQEQRQQEG- DLSDLVGSPYVQQRSQ	4426.220	886.246	5	-2.0	4427.44	-49.8	None
6	apUII-RPII[45- 79]	SMLLSKYMALMNRLQEQRQQE- GDLSDLVGSPYVQQ	4055.020	1014.759	4	-4.2	4056.23	-51.7	None
7	pleurin/Plm3	GIFTQSAYGSYPRV-NH ₂	1543.793	772.897	2	-0.1	1545.05	-169.2	Amida tion C-term
8	apUII*[93-112]	SGGMSLC(-1.01)LWKVC(-1.01)PA- APWLIS	2116.041	1059.028	2	-7.2	2117.22	-84.6	C-C disulfi de bonded

^a CE-TIMS-MS measurement

^b MALDI-MS measurement

Pleurin, NP_001191654.1

TMSB = thymosin beta, NP_001191437.1

AMRPs = *Aplysia* mytilus inhibitory peptide-related peptides, NP_001191614.1

ApUII = *Aplysia* urotensin II, XP_012945633.1

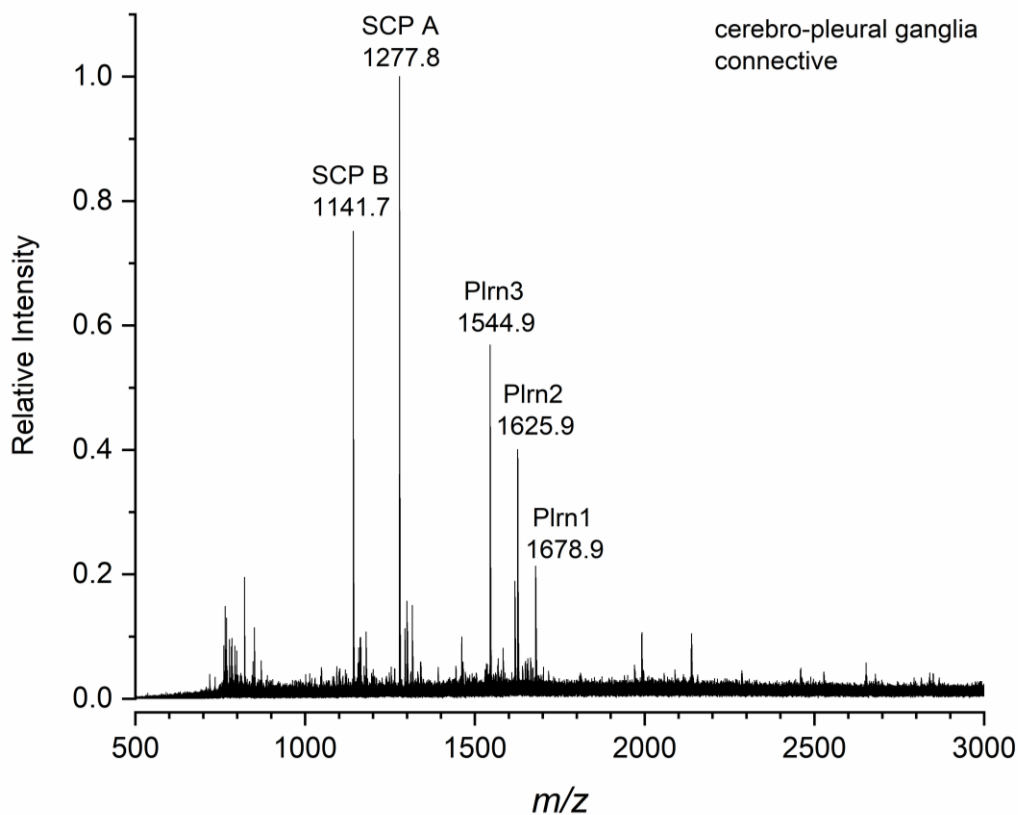


Figure 4.10. MALDI MS spectrum obtained from a section of cerebro-pleural ganglia connective. In addition to Plrn1–Plrn3, the $[m+H]^+$ ions of the most-abundant peptides detected in the connective are labeled as follows: small cardioactive peptide A (SCP A: ARPGYLAFPRM-NH₂) and small cardioactive peptide B (SCP B: MNYLAFPRM-NH₂) from the *A. californica* small cardioactive peptides precursor (NCBI accession: XP_012943902.1). Detection of SCP A and SCP B, and other peptides (unlabeled) co-localizing with Plrn1–Plrn3 in cerebro-pleural connectives is because neuronal processes from different, non-pleurin-containing neuron populations also pass through the cerebro-pleural ganglia connective.

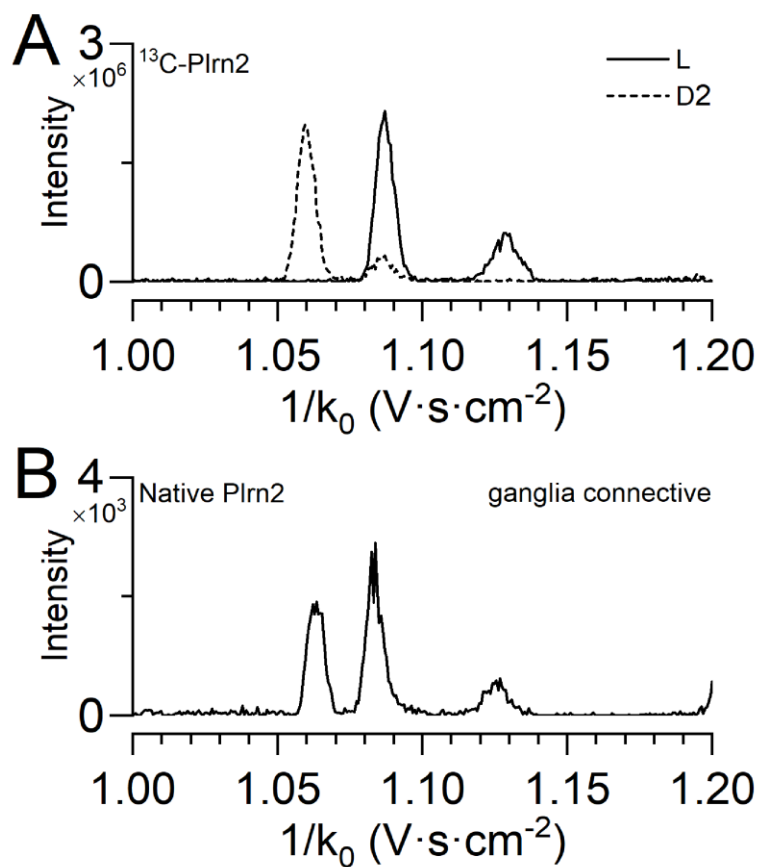


Figure 4.11. EIMs obtained by CE-MRM-TIMS MS showing the localization of D2-Plrn2 to the cerebro-pleural connective. (A) EIMs of synthetic ^{13}C -L-Plrn2 (solid trace) and ^{13}C -D2-Plrn2 $[\text{m}^*+2\text{H}]^{2+}$ ions (dashed trace). (B) EIM of native Plrn2 $[\text{m}+2\text{H}]^{2+}$ detected in the cerebro-pleural ganglia connective.

4.8 REFERENCES

1. Ollivaux, C.; Soye, D.; Toullec, J. Y., Biogenesis of D-amino acid containing peptides/proteins: where, when and how? *J. Pept. Sci.* **2014**, *20* (8), 595-612.
2. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Advancing d-amino acid-containing peptide discovery in the metazoan. *Biochim. Biophys. Acta Proteins Proteom.* **2021**, *1869* (1), 140553.
3. Yin, L.; Zhang, Z.; Liu, Y.; Gao, Y.; Gu, J., Recent advances in single-cell analysis by mass spectrometry. *Analyst* **2019**, *144* (3), 824-845.
4. Comi, T. J.; Do, T. D.; Rubakhin, S. S.; Sweedler, J. V., Categorizing Cells on the Basis of their Chemical Profiles: Progress in Single-Cell Mass Spectrometry. *J. Am. Chem. Soc.* **2017**, *139* (11), 3920-3929.
5. Adams, C. M.; Zubarev, R. A., Distinguishing and quantifying peptides and proteins containing D-amino acids by tandem mass spectrometry. *Anal. Chem.* **2005**, *77* (14), 4571-80.
6. Sachon, E.; Clodic, G.; Galanth, C.; Amiche, M.; Ollivaux, C.; Soye, D.; Bolbach, G., D-amino acid detection in peptides by MALDI-TOF-TOF. *Anal. Chem.* **2009**, *81* (11), 4389-96.
7. Adams, C. M.; Kjeldsen, F.; Zubarev, R. A.; Budnik, B. A.; Haselmann, K. F., Electron capture dissociation distinguishes a single D-amino acid in a protein and probes the tertiary structure. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (7), 1087-98.
8. Serafin, S. V.; Maranan, R.; Zhang, K.; Morton, T. H., Mass spectrometric differentiation of linear peptides composed of L-amino acids from isomers containing one D-amino acid residue. *Anal. Chem.* **2005**, *77* (17), 5480-7.

9. Bai, L.; Livnat, I.; Romanova, E. V.; Alexeeva, V.; Yau, P. M.; Vilim, F. S.; Weiss, K. R.; Jing, J.; Sweedler, J. V., Characterization of GdFFD, a D-amino acid-containing neuropeptide that functions as an extrinsic modulator of the Aplysia feeding circuit. *J. Biol. Chem.* **2013**, 288 (46), 32837-51.
10. Koehbach, J.; Gruber, C. W.; Becker, C.; Kreil, D. P.; Jilek, A., MALDI TOF/TOF-Based Approach for the Identification of d- Amino Acids in Biologically Active Peptides and Proteins. *J. Proteome Res.* **2016**, 15 (5), 1487-96.
11. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Differential Post-Translational Amino Acid Isomerization Found among Neuropeptides in Aplysia californica. *ACS Chem. Biol.* **2020**, 15 (1), 272-281.
12. Pisarewicz, K.; Mora, D.; Pflueger, F. C.; Fields, G. B.; Marí, F., Polypeptide chains containing D-gamma-hydroxyvaline. *J. Am. Chem. Soc.* **2005**, 127 (17), 6207-15.
13. Bai, L.; Romanova, E. V.; Sweedler, J. V., Distinguishing endogenous D-amino acid-containing neuropeptides in individual neurons using tandem mass spectrometry. *Anal. Chem.* **2011**, 83 (7), 2794-800.
14. Zhang, Z.; Hebert, A. S.; Westphall, M. S.; Coon, J. J.; Dovichi, N. J., Single-Shot Capillary Zone Electrophoresis-Tandem Mass Spectrometry Produces over 4400 Phosphopeptide Identifications from a 220 ng Sample. *J. Proteome Res.* **2019**, 18 (8), 3166-3173.
15. Qu, Y.; Sun, L.; Zhu, G.; Zhang, Z.; Peuchen, E. H.; Dovichi, N. J., Sensitive and fast characterization of site-specific protein glycosylation with capillary electrophoresis coupled to mass spectrometry. *Talanta* **2018**, 179, 22-27.

16. Evers, T. M. J.; Hochane, M.; Tans, S. J.; Heeren, R. M. A.; Semrau, S.; Nemes, P.; Mashaghi, A., Deciphering Metabolic Heterogeneity by Single-Cell Analysis. *Anal. Chem.* **2019**, *91* (21), 13314-13323.
17. Onjiko, R. M.; Portero, E. P.; Moody, S. A.; Nemes, P., In Situ Microprobe Single-Cell Capillary Electrophoresis Mass Spectrometry: Metabolic Reorganization in Single Differentiating Cells in the Live Vertebrate (*Xenopus laevis*) Embryo. *Anal. Chem.* **2017**, *89* (13), 7069-7076.
18. Nemes, P.; Knolhoff, A. M.; Rubakhin, S. S.; Sweedler, J. V., Metabolic differentiation of neuronal phenotypes by single-cell capillary electrophoresis-electrospray ionization-mass spectrometry. *Anal. Chem.* **2011**, *83* (17), 6810-7.
19. Liao, H. W.; Rubakhin, S. S.; Philip, M. C.; Sweedler, J. V., Enhanced single-cell metabolomics by capillary electrophoresis electrospray ionization-mass spectrometry with field amplified sample injection. *Anal. Chim. Acta* **2020**, *1118*, 36-43.
20. Wang, J.; Ma, M.; Chen, R.; Li, L., Enhanced neuropeptide profiling via capillary electrophoresis off-line coupled with MALDI FTMS. *Anal. Chem.* **2008**, *80* (16), 6168-77.
21. Lombard-Banek, C.; Yu, Z.; Swiercz, A. P.; Marvar, P. J.; Nemes, P., A microanalytical capillary electrophoresis mass spectrometry assay for quantifying angiotensin peptides in the brain. *Anal. Bioanal. Chem.* **2019**, *411* (19), 4661-4671.
22. Lombard-Banek, C.; Moody, S. A.; Nemes, P., Single-Cell Mass Spectrometry for Discovery Proteomics: Quantifying Translational Cell Heterogeneity in the 16-Cell Frog (*Xenopus*) Embryo. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (7), 2454-8.
23. Lombard-Banek, C.; Moody, S. A.; Manzini, M. C.; Nemes, P., Microsampling Capillary Electrophoresis Mass Spectrometry Enables Single-Cell Proteomics in Complex Tissues:

- Developing Cell Clones in Live *Xenopus laevis* and Zebrafish Embryos. *Anal. Chem.* **2019**, *91* (7), 4797-4805.
24. Kawai, T.; Ota, N.; Okada, K.; Imasato, A.; Owa, Y.; Morita, M.; Tada, M.; Tanaka, Y., Ultrasensitive Single Cell Metabolomics by Capillary Electrophoresis–Mass Spectrometry with a Thin-Walled Tapered Emitter and Large-Volume Dual Sample Preconcentration. *Anal. Chem.* **2019**, *91* (16), 10564-10572.
25. Zhao, S. S.; Zhong, X.; Tie, C.; Chen, D. D., Capillary electrophoresis-mass spectrometry for analysis of complex samples. *Proteomics* **2012**, *12* (19-20), 2991-3012.
26. Jansson, E. T., Strategies for analysis of isomeric peptides. *J. Sep. Sci.* **2018**, *41* (1), 385-397.
27. Sheeley, S. A.; Miao, H.; Ewing, M. A.; Rubakhin, S. S.; Sweedler, J. V., Measuring D-amino acid-containing neuropeptides with capillary electrophoresis. *Analyst* **2005**, *130* (8), 1198-203.
28. Scriba, G. K., Recent developments in peptide stereoisomer separations by capillary electromigration techniques. *Electrophoresis* **2009**, *30 Suppl 1*, S222-8.
29. Ali, I.; Al-Othman, Z. A.; Al-Warthan, A.; Asnin, L.; Chudinov, A., Advances in chiral separations of small peptides by capillary electrophoresis and chromatography. *J. Sep. Sci.* **2014**, *37* (18), 2447-66.
30. Zhang, L.; Foreman, D. P.; Grant, P. A.; Shrestha, B.; Moody, S. A.; Villiers, F.; Kwak, J. M.; Vertes, A., In situ metabolic analysis of single plant cells by capillary microsampling and electrospray ionization mass spectrometry with ion mobility separation. *Analyst* **2014**, *139* (20), 5079-85.

31. Zhang, L.; Vertes, A., Energy Charge, Redox State, and Metabolite Turnover in Single Human Hepatocytes Revealed by Capillary Microsampling Mass Spectrometry. *Anal. Chem.* **2015**, *87* (20), 10397-405.
32. Zhang, L.; Sevinsky, C. J.; Davis, B. M.; Vertes, A., Single-Cell Mass Spectrometry of Subpopulations Selected by Fluorescence Microscopy. *Anal. Chem.* **2018**, *90* (7), 4626-4634.
33. Li, G.; DeLaney, K.; Li, L., Molecular basis for chirality-regulated A β self-assembly and receptor recognition revealed by ion mobility-mass spectrometry. *Nat. Comm.* **2019**, *10* (1), 5038.
34. Jeanne Dit Fouque, K.; Garabedian, A.; Porter, J.; Baird, M.; Pang, X.; Williams, T. D.; Li, L.; Shvartsburg, A.; Fernandez-Lima, F., Fast and Effective Ion Mobility-Mass Spectrometry Separation of d-Amino-Acid-Containing Peptides. *Anal. Chem.* **2017**, *89* (21), 11787-11794.
35. Jia, C.; Lietz, C. B.; Yu, Q.; Li, L., Site-specific characterization of (D)-amino acid containing peptide epimers by ion mobility spectrometry. *Anal. Chem.* **2014**, *86* (6), 2972-81.
36. Pang, X.; Jia, C.; Chen, Z.; Li, L., Structural Characterization of Monomers and Oligomers of D-Amino Acid-Containing Peptides Using T-Wave Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (1), 110-118.
37. Dodds, J. N.; May, J. C.; McLean, J. A., Correlating Resolving Power, Resolution, and Collision Cross Section: Unifying Cross-Platform Assessment of Separation Efficiency in Ion Mobility Spectrometry. *Anal. Chem.* **2017**, *89* (22), 12176-12184.

38. Silveira, J. A.; Ridgeway, M. E.; Park, M. A., High resolution trapped ion mobility spectrometry of peptides. *Anal. Chem.* **2014**, *86* (12), 5624-7.
39. Meier, F.; Brunner, A. D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Räther, O.; Mann, M., Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Mol. Cell. Proteomics* **2018**, *17* (12), 2534-2545.
40. Vasilopoulou, C. G.; Sulek, K.; Brunner, A. D.; Meitei, N. S.; Schweiger-Hufnagel, U.; Meyer, S. W.; Barsch, A.; Mann, M.; Meier, F., Trapped ion mobility spectrometry and PASEF enable in-depth lipidomics from minimal sample amounts. *Nat. Commun.* **2020**, *11* (1), 331.
41. Cecala, C.; Sweedler, J. V., Sampling techniques for single-cell electrophoresis. *Analyst* **2012**, *137* (13), 2922-9.
42. Moroz, L. L.; Edwards, J. R.; Puthanveetil, S. V.; Kohn, A. B.; Ha, T.; Heyland, A.; Knudsen, B.; Sahni, A.; Yu, F.; Liu, L.; Jezzini, S.; Lovell, P.; Iannuccilli, W.; Chen, M.; Nguyen, T.; Sheng, H.; Shaw, R.; Kalachikov, S.; Panchin, Y. V.; Farmerie, W.; Russo, J. J.; Ju, J.; Kandel, E. R., Neuronal transcriptome of Aplysia: neuronal compartments and circuitry. *Cell* **2006**, *127* (7), 1453-67.
43. Romanova, E. V.; Aerts, J. T.; Croushore, C. A.; Sweedler, J. V., Small-volume analysis of cell-cell signaling molecules in the brain. *Neuropsychopharmacology* **2014**, *39* (1), 50-64.
44. Li, L.; Garden, R. W.; Sweedler, J. V., Single-cell MALDI: a new tool for direct peptide profiling. *Trends Biotechnol.* **2000**, *18* (4), 151-60.

45. Lombard-Banek, C.; Reddy, S.; Moody, S. A.; Nemes, P., Label-free Quantification of Proteins in Single Embryonic Cells with Neural Fate in the Cleavage-Stage Frog (*Xenopus laevis*) Embryo using Capillary Electrophoresis Electrospray Ionization High-Resolution Mass Spectrometry (CE-ESI-HRMS). *Mol. Cell. Proteomics* **2016**, *15* (8), 2756-68.
46. Nemes, P.; Rubakhin, S. S.; Aerts, J. T.; Sweedler, J. V., Qualitative and quantitative metabolomic investigation of single neurons by capillary electrophoresis electrospray ionization mass spectrometry. *Nat. Protoc.* **2013**, *8* (4), 783-99.
47. Fuller, R. R.; Sweedler, J. V., Factors affecting quantitative electrokinetic injections from submicroliter conductive vials in capillary electrophoresis. *Anal. Chem.* **1999**, *71* (18), 4014-22.
48. Naylor, C. N.; Reinecke, T.; Ridgeway, M. E.; Park, M. A.; Clowers, B. H., Validation of Calibration Parameters for Trapped Ion Mobility Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (10), 2152-2162.
49. Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J., Jr.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinou, K.; Valentine, S. J.; Wytenbach, T., Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom. Rev.* **2019**, *38* (3), 291-320.
50. Li, G.; Delafield, D. G.; Li, L., Improved structural elucidation of peptide isomers and their receptors using advanced ion mobility-mass spectrometry. *TrAC, Trends Anal. Chem.* **2020**, *124*, 115546.

51. Garabedian, A.; Bolufer, A.; Leng, F.; Fernandez-Lima, F., Peptide Sequence Influence on the Conformational Dynamics and DNA binding of the Intrinsically Disordered AT-Hook 3 Peptide. *Sci. Rep.* **2018**, *8* (1), 10783.
52. Romanova, E. V.; Sasaki, K.; Alexeeva, V.; Vilim, F. S.; Jing, J.; Richmond, T. A.; Weiss, K. R.; Sweedler, J. V., Urotensin II in invertebrates: from structure to function in *Aplysia californica*. *PLoS One* **2012**, *7* (11), e48764.
53. Fujisawa, Y.; Furukawa, Y.; Ohta, S.; Ellis, T. A.; Dembrow, N. C.; Li, L.; Floyd, P. D.; Sweedler, J. V.; Minakata, H.; Nakamaru, K.; Morishita, F.; Matsushima, O.; Weiss, K. R.; Vilim, F. S., The *Aplysia mytilus* inhibitory peptide-related peptides: identification, cloning, processing, distribution, and action. *J. Neurosci.* **1999**, *19* (21), 9618-34.
54. Romanova, E. V.; Roth, M. J.; Rubakhin, S. S.; Jakubowski, J. A.; Kelley, W. P.; Kirk, M. D.; Kelleher, N. L.; Sweedler, J. V., Identification and characterization of homologues of vertebrate beta-thymosin in the marine mollusk *Aplysia californica*. *J. Mass Spectrom.* **2006**, *41* (8), 1030-40.
55. Song, Y.; Liu, Y. M., Quantitation of cardioexcitatory Asn-D-Trp-Phe-NH₂ diastereomers in *Aplysia*'s central nervous system by nanoscale liquid chromatography-tandem mass spectrometry. *J. Mass Spectrom.* **2008**, *43* (9), 1285-90.
56. Soye, D.; Toullec, J. Y.; Ollivaux, C.; Géraud, G., L to D amino acid isomerization in a peptide hormone is a late post-translational event occurring in specialized neurosecretory cells. *J. Biol. Chem.* **2000**, *275* (48), 37870-5.
57. Pearson, W. L.; Lloyd, P. E., Immunocytochemical localization of pedal peptide in the central nervous system and periphery of *Aplysia*. *J. Neurosci.* **1989**, *9* (1), 318-25.

58. Lloyd, P. E., Interganglionic axonal transport of neuropeptides in Aplysia. *J. Neurosci.* **1989**, *9* (9), 3243-9.
59. Ierusalimsky, V. N.; Boguslavsky, D. V.; Belyavsky, A. V.; Balaban, P. M., Helix peptide immunoreactivity pattern in the nervous system of juvenile aplysia. *Brain Res. Mol. Brain Res.* **2003**, *120* (1), 84-9.
60. Chai, M.; Young, M. N.; Liu, F. C.; Bleiholder, C. A Transferable, Sample-Independent Calibration Procedure for Trapped Ion Mobility Spectrometry (TIMS). *Anal. Chem.* **2018**, *90*, 9040-9047.

CHAPTER 5

ENZYMATIC DIGESTIONS ON INDIVIDUAL *APLYSIA CALIFORNICA* NEURONS

5.1 NOTES AND ACKNOWLEDGEMENTS

This work was completed with the assistance from E. Romanova who helped with instruction on animal discussions, anatomy, and single cell sampling. Thanks to S. Rubakhin for instruction on the operation of various instruments, and general guidance and suggestions. Thanks to K. Anapindi for providing supplementary *Aplysia californica* neuropeptide sequence libraries. Also, a special thanks to J. Checco for guidance and mentorship.

5.2 INTRODUCTION

L- to D-residue isomerization is an unusual enzyme catalyzed post-translational modification (PTM) to signaling peptides that converts all-L residue peptides into D-amino acid containing peptides (DAACPs), and this may be a cell-type specific PTM in animals.¹⁻³ Many previous efforts to analyze DAACPs have sacrificed single cell heterogeneity by pooling large cell populations, often from several animals, to obtain concentrations sufficient for detection (see Chapter 2). While this type of bulk tissue analysis is a reasonable approach for non-targeted discovery of DAACPs, it is not useful for studying cell-specific localization of DAACPs. Previous approaches that have been used to detect DAACPs in single cells have been largely targeted and required method optimization specific to the DAACP analyte of interest.⁴⁻⁶ To test for cell-type

specific isomerization, there is a need for a high throughput method capable unbiased detection of a broad range of DAACPs in single cells.

In this chapter, aminopeptidase M (APM) digestions were performed on single cells from the *Aplysia californica* central nervous system (CNS) to test for the presence of known DAACPs in three different cell types. Matrix assisted laser desorption-ionization (MALDI) mass spectrometry (MS) was used to detect peptides, and DAACPs were differentiated from all-L peptides based on resistance to APM. APM kinetically favors the successive cleavage of L-amino acid residues from peptides, but peptides with a D-amino acid residue at the primary or secondary positions from the N-terminus are more stable against proteolysis by APM. Neurons were first isolated and directly deposited onto a microscope slide and their peptide profiles were measured using MALDI MS. The same neuron was then treated with APM to presumably degrade away most of the all-L residue peptides, while the more stable DAACPs remain undigested. After the APM digestion, the peptide profile was measured again by MALDI MS and compared to the peptide profile that was measured before the APM digestion. Using this procedure, it was possible to detect DAACPs in the single cell after the APM digestion because the all-L peptides were degraded away. This protocol allowed for direct detection of DAACPs in single cells by MALDI MS without the need for solution phase separations.

Single cell APM digestions were performed on neurons that express the *Aplysia* achatin-like neuropeptide precursor (ApALNP), the pleurin precursor, and the FMRFamide precursor because these prohormones are known to encode isomerized peptides. The peptide profiles of these neurons were examined before and after the APM digestion to determine whether known DAACPs localize to the cell bodies of isolated neurons.

5.3 METHODS

Animal Dissections, Neuron Isolation and Cell Isolation

A. californica were dissected and the CNS was surgically removed and desheathed as described in Chapter 4.⁵ Pedal and pleural neurons were isolated directly from ganglia without dispersion. The precision of the single cell isolation is shown in Figure 5.1. Pedal, or pleural neurons were isolated removed using a lab-made micropipette filled with deionized water (see Chapter 4),⁵ and deposited onto indium tin oxide coated $25 \times 75 \times 1.1$ mm microscope slides (P/N CG-901N-S115; Delta technologies, Loveland, CO) in an array. Excess liquid was quickly removed from each spot after depositing the neuron, and 0.5 μ L saturated DHB in water with 0-0.1% trifluoroacetic acid was deposited onto spots. The droplets were then allowed to dry in air under a tungsten lamp. Each spot was then analyzed by matrix assisted laser desorption-ionization (MALDI) mass spectrometry (MS) using a Bruker Ultraflex MALDI MS (Bruker Daltonics, Billerica, MA) before screening procedure started.

MALDI TOF MS

MALDI MS analysis was performed an ultrafleXtreme MALDI-TOF mass spectrometer (Bruker Corp.) in positive reflection mode. A Smartbeam II frequency tripled Nd:YAG solid state laser was used for ionization. To minimize the sample depletion, laser shots were fired in 50 shot bursts at a frequency of 1000 Hz, fired randomly around perimeter of the spots to a maximum of 1000 cumulative shots. The summed MS spectra were then used for subsequent analyses.

Aminopeptidase M Preparation

Aminopeptidase M from porcine kidney (APM, CAS#: 9054-63-1) was purchased from Calbiochem (Darmstadt, Germany). The APM suspension (170 U/mL, 2.7 mg/mL) was stored under manufacturer recommended conditions until use. For preparation of APM in NH_4HCO_3 , the APM suspension was gently swirled in its container and then 50 μL was transferred to a 0.5 mL PCR tube and centrifuged at low speed on a tabletop centrifuge for 1 min. After centrifugation, the supernatant was removed and the remaining sediment containing the protein enzyme was reconstituted in 50 μL 200 mM NH_4HCO_3 and subsequently diluted to the desired activity with 200 mM NH_4HCO_3 . APM solutions were prepped fresh before each experiment and stored on ice until use.

Matrix Neutralization and Preparation for APM Digestion

After the initial MALDI MS measurement was performed on deposited neurons, spots in the array were reconstituted in 2 μL of 50-200 mM NH_4HCO_3 to neutralize the droplet containing the cellular contents mixed with DHB to pH ~ 8 . To ensure each spot was sufficiently neutralized, the spots were titrated with 200 mM NH_4HCO_3 in 0.2 μL intervals until CO_2 was no longer observed evolving from the droplets. After neutralizing the DHB, 0.5 μL of 60-170 U/mL APM in 200 mM NH_4HCO_3 was added to each droplet and the ITO slides were incubated in a humidified chamber at 37 $^\circ\text{C}$ for 3-6 h in a warm room. The final volume of each spot was 2-2.5 μL .

APM Digestion and Incubation Conditions

The ITO slides containing neutralized droplets (2-2.5 μL ea.) with APM were incubated inside of a lab-made humidified chamber. The chamber included aqueous 50 mM NH_4HCO_3

solution in an open vessel within the chamber to maintain an atmosphere consistent with the composition of the neutralized droplets. Maintaining a humid atmosphere was important to prevent the droplets from drying during the APM digestion and including NH_4HCO_3 in the open vessel served to maintain the solution at the intended pH. After incubation, the droplets were dried in a vacuum chamber.

Matrix Reacidification

After the APM digestion, it was necessary to reacidify the neutralized droplets in order to perform the MALDI MS measurements on the digests. To reacidify the MALDI matrix, 10% acetic acid (AcOH) in 50/50 H_2O /methanol was added to the spots in 0.5 μL intervals to react with excess NH_4HCO_3 . Spots were titrated with additional AcOH in 50/50 methanol/water in 0.25 μL intervals until CO_2 was no longer observed evolving from the spots, indicating the droplets were sufficiently neutralized. Droplets were then dried under vacuum to remove excess volatile buffers and AcOH from the droplets. Following the reacidification, when each spot was sufficiently dry, the target was removed from the vacuum and 0.5-1 μL of DHB solution prepared in 50/50 CH_3CN /water with 0.1% TFA was added to the droplets, and allowed to dry under a tungsten lamp and measured again by MALDI MS.

Peptide Identification

Previous studies have used MALDI MS to identify endogenous prohormone processing products at the single cell level, and assignment of peptide sequences in this chapter was performed in accordance with this literature precedent.⁷⁻¹¹ Endogenous peptides were identified in MALDI MS spectra by mass matching ions in MALDI MS spectra to an in-lab *A. californica* neuropeptide

library assembled by liquid chromatography tandem mass spectrometry (LC-MS/MS) sequencing on ganglia extracts. In addition, the prohormone processing of specific precursors was predicted using neuropred prediction software, to identify any additional peptides that may not have been detected by LC-MS/MS or to help identify the most likely peptide product if two colocalizing precursors encode isobaric peptides.²² Generally, the detection of multiple $[m+H]^+$ ions corresponding to peptides derived from a single precursor increases the confidence of the peptide detected.

5.4 RESULTS AND DISCUSSION

APM Digestions on ApALNP-Neurons

At least three DAACPs were expected to be present in ApALNP containing neurons: **GdFFD**, **GdYFD** and D2-ApALNP-3 (**SdY**ADSKDEESNAALSDF~~A~~).¹²⁻¹³ Three pedal neurons were isolated where ApALNP-3 was detected. However, the GFFD $[m+H]^+$ ions were only detected in 2/3 of the neurons before the APM digestions, and the GYFD $[m+H]^+$ was not detected in any of the isolated neurons. GFFD was not detected after the APM digestions on any of the isolated neurons. The full length ApALNP-3 peptide (SYADSKDEESNAALSDF~~A~~ED) and two of its truncated forms (SYADSKDEESNAALSDF~~A~~E, SYADSKDEESNAALSDF~~A~~) were detected in the cells before the APM digestion. After the APM digestion, three additional truncated forms of the full length ApALNP-3 were observed, including YADSKDEESNAALSDF~~A~~, ADSKDEESNAALSDF~~A~~, and DSKDEESNAALSDF~~A~~. APM digestions sequentially removes amino acids from the N-termini of peptides and these additional truncated forms of Ap-ALNP-3 are likely the partially degraded all-L peptide. The other peptides detected in the cell before the

digestion were fully degraded (Figure 5.2), but the identities of all these peptides were not able to be determined.

APM Digestions on Pleurin-Neurons

Pleurin-neurons, were also tested for DAACPs using the on-target APM digestion method. N = 3 pleurin-neurons were isolated. The pleurin precursor encodes three peptides, which can be isomerized into at least 5 possible DAACPs: D2-Plrn1 (**MdFY**TKGSDSYPRI-NH₂), D2-Plrn2 (**SdFY**TTGNGNYPRI-NH₂), D3-Plrn2 (**SFdY**TTGNGNHYPRI-NH₂), D2-Plrn3 (**GdIF**TQSAYGSYPRV-NH₂) and D3-Plrn3 (**GIdF**TQSAYGSYPRV-NH₂).¹⁴

The MALDI MS spectra of a pleurin-neurons before and after APM digestion is shown in Figure 5.3. Plrn1, Plrn2 and Plrn3 (Plrn1-Plrn3) were detected in the neurons before the digestions. After APM treatment, the Plrn1-Plrn3 [m+H]⁺ ions were not detected, indicating the D2-diastereomers of Plrn1-Plrn3 were not present, or were present below the detection limit. It was demonstrated in Chapter 3 that APM cleaves the N-terminal Ser and Gly residues from D3-Plrn2 and D3-Plrn3, respectively, and the truncated D2-Plrn2(2-14) and D3-Plrn3(2-14) remain stable against digestion.¹⁴ Plrn2(2-14) and Plrn3(2-14) were not detected after APM digestions, indicating D3-forms were also not present or were below the limit of detection. The latter finding indicates that the DAACP forms of Plrn1, Plrn2 and Plrn3 were not present in the cell bodies of Pleurin neurons (N=3) and is consistent with the results reported in Chapter 4.⁵ Peptides from other prohormones that colocalized with pleurin-peptides were described in Chapter 4, and none of these peptides were observed resisting APM in the single cell digestion.

APM Digestions on FMRGFamide-Neurons

FdMRGF-NH₂ was previously determined to be a DAACP in *A. californica* (see Chapter 3).¹⁴ Therefore, FMRGF-NH₂ containing neurons were selected to be tested by the on-target APM screening procedure. However, locating the neurons that express the FMRGFamide precursor for isolation was particularly challenging because these cells have not previously been identified in the CNS by *in-situ* hybridization staining. Before these neurons could be isolated it was necessary to determine where in the CNS they are located.

A total of N = 20 neurons were isolated at random from the pedal or pleural ganglia of a single animal. FMRGF-NH₂ ([m+H]⁺ = 656.4) and other FMRGFamide-derived peptides were detected in N = 7 of the neurons. Peptides that derived from the pedal peptides 1 (PP1)¹⁵⁻¹⁷ and the buccalin¹⁸⁻²¹ precursors were also detected colocalizing with FMRGF-NH₂ in all N = 7 neurons. In several neurons that were isolated, buccalin and PP1 derived peptides were detected but FMRGFamide-derived peptides were not, but this was not unexpected. These measurements represent the first reports of direct detection of FMRGFamide-related peptides in single neurons from *A. californica*.

After determining the relative location of FMRGFamide-neurons, a second animal was dissected. Pedal neurons were isolated and tested for FMRGF-NH₂ by MALDI MS. In total N = 4 neurons were isolated and tested that contained FMRGFamide-derived peptides. These neurons were then subjected to APM digestion to test for the presence of **FdMRGF**-NH₂. Figures 5.4-5.9 shows the APM digestion results for an FMRGFamide-neuron and represents the result for all N = 4 replicates. FMRGF-NH₂ ([m+H]⁺ = 656.4) was not detected after the digestion in any of the cells tested. FM(ox)RGF-NH₂ ([m+16+H]⁺ = 672.4) was also not detected and no truncated forms

of were detected. These results suggest that, like the Plrn1-Plrn3 peptides (see Chapter 4), the DAACP form of FMRGF-NH₂ does not localize to cell bodies.

FdMRGF-NH₂ was not detected (based on APM resistance) in single neurons and this could be because it accumulates outside the cell bodies (as was the case with Plrn1-Plrn3 in Chapter 4). Alternatively, it is possible that ion suppression in the low *m/z* range (<1000 Da) prevented the detection of any APM resistant FMRGF-NH₂ that may have been present. Follow-up analysis of FMRGFamide-neurons using the methods described in Chapter 4 could be performed to increase the confidence in the interpretation that FMRGF-NH₂ does not localize to neuron cell bodies.

The colocalization of buccalin and PP1 derived peptides with FMRGFamide-derived peptides is an interesting observation to note because it suggests these peptides are also exposed to the same enzyme that isomerizes FMRGF-NH₂. Furthermore, several of the buccalin peptides have N-terminal sequence similarities to many of the known DAACPs in *A. californica*, which could indicate they too are substrates for the *A. californica* isomerase(s).ⁱ Therefore, the peptides derived from PP1 and buccalin were also examined after APM digestion. If any of these sequences were resistant to APM it could suggest they have a D-residue containing form. Most of the buccalin-derived peptides were degraded by APM. Only one buccalin peptide, PNVDPSYLSV-NH₂ ([m+H]⁺ = 1349.7), was observed after the APM digestion. The resistant buccalin peptide had an N-terminal proline, and thus, its stability against hydrolysis could not be confidently attributed to the presence of a D-residue. Of the PP1-derived peptides detected, only those with N-terminal proline residues resisted APM digestion (Figure 5.8). Peptides derived from PP1 (PLDSVYGTHGLSGFA, PLDSVYGTHGMSGFA, PFDSISQGEGLSGFA) were resistant

ⁱ Notably, buccalin peptides share the Gly-Xaa- motif similar to other Aplysia peptides (See Chapter 2). Others have Met at position 2, like FMRGF-NH₂.

to APM in every cell tested. These peptides were expected to resist APM because in the experiments reported in Chapter 3, 1-2% remained undigested (as determined by LC-MS). The likely reason for the incomplete digestion of the buccalin peptide, and the PP1 peptides is because they have an N-terminal proline, making them more resistant to digestion. Because these sequences have proline at the N-terminus, the reason they resisted APM digestion cannot be confidently attributed to a D-residue. Therefore, no new DAACPs could be identified by performing APM digestions on single FMRGFamide-neurons.

Given these observations, it is perplexing that D-residue containing forms of the buccalin peptides were not identified in the screening experiments performed in Chapter 3 or in the results presented here. These puzzling observations highlight several gaps in the current understanding of the substrate specificity and cellular localization of the *A. californica* isomerase(s) and DAACPs.ⁱⁱ

Finally, the m/z corresponding to another known isomerized peptide from *Aplysia* allatotropin related polypeptide (ApATRP) in *A. californica* (GFRLNSASRVAHGY-NH₂, $[m+H]^+ = 1533.8$), was detected in the pedal neurons.²³ However, it is more likely that m/z 1532.8 is just the sodium adduct of PFDSISQGEGLSGFA ($[m+Na]^+ = 1533.7$). The sodium adduct of PFDSISQGEGLSGFA and the ApATRP $[m+H]^+$ have a $\Delta m/z = 0.1$, which could not distinguished in these MALDI MS measurements.

Method Limitations and Additional Considerations

Although this method was used to perform APM digests on individual neurons, in its present state of development, it is not complete. It could not be used to accurately quantify the relative abundance of peptides before and after the APM digestion. Initial attempts to quantify

ⁱⁱ It could also reflect a limitation of APM screening to identify DAACPs based on APM resistance, if some DAACPs are able to be degraded by APM.

peptides before and after APM digestions by MALDI MS found high variability from sample to sample. Quantitation was challenging because the MALDI matrix composition was fundamentally changed after performing the matrix neutralization/reacidification procedure and the APM digestion.

After the digestion, the low mass range tended to have low signal-to-noise ratios. It is possible that low m/z ions (approximately $m/z < 1000$) were more affected by ion suppression after applying the neutralization/reacidification procedure. If ion suppression did occur after the digestion, this could make it appear as though APM resistant peptides were degraded.

Additionally, endogenous proteolytic enzymes were found to reactivate after matrix neutralization and incubation for 3 h at 37 °C. The reactivation of proteolytic enzymes can degrade native peptides (including potential DAACPs). Degradation of the APM by reactivated enzymes is a major problem because it can result in an incomplete digestion of the all-L peptides. This could lead to the false assignment of undigested all-L peptides as DAACPs. However, the reactivation of endogenous enzymes was determined to be more of an issue in digestions of groups of neurons, as opposed to individual cells but was not explored further in this dissertation.

5.5 CONCLUSION

In this chapter, APM digestions were performed on single cells directly on MALDI MS targets. The enzymatic digestion was accomplished in the presence of DHB MALDI matrix by neutralizing the matrix to the optimal conditions for APM activity with a volatile buffer. The on-target screening was applied to test individual *A. californica* neurons that express prohormones known to encode isomerized peptides from the pedal and pleural ganglia.

The results from APM digestions on pleurin-containing neurons indicated that the DAACP forms for pleurin-peptides do not localize to the cell bodies, and this was consistent with the results in Chapter 4. Furthermore, preliminary results suggest that F**d**MRGF-NH₂ also does not localize to cell bodies of neurons expressing the FMRGFamide precursor. Conversely, APM digestions on ApALNP-containing neurons showed ApALNP-3 resisting APM in single cells. This suggests D2-ApALNP-3 does localize to the cell bodies of neurons. These preliminary observations indicate the on-target single cell APM protocol can be used in future studies to screen for the cell-type specific localization of DAACPs in the *A. californica* CNS.

5.6 FIGURES AND DATA

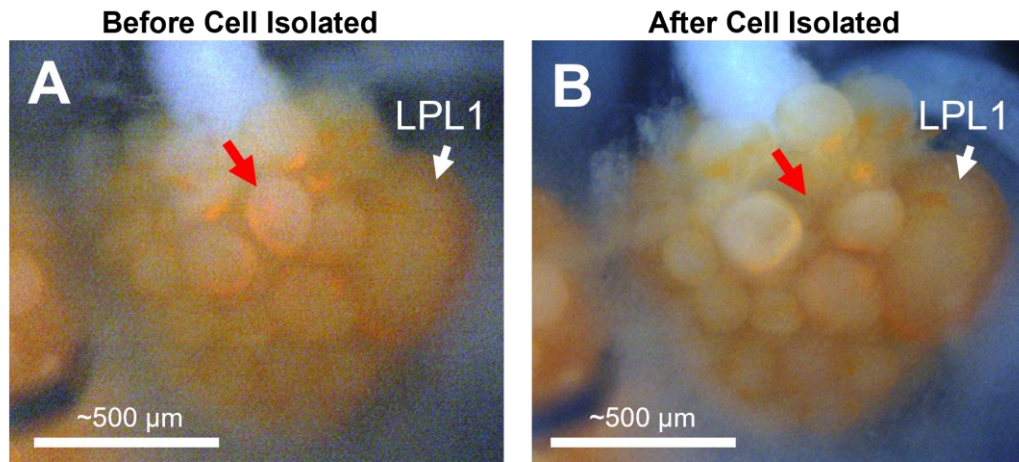


Figure 5.1. Precision of single cell isolation for on-target APM screening. Isolation of a single pleural neuron from the left pleural ganglion. The red arrow in (A) points to the neuron that was targeted for isolation, and the red arrow in (B) points to the same location after the neuron was removed. The location of LPL1 is shown for reference and indicated by the white arrow. Note the image displayed in B was modified by increasing the brightness to improve visibility. Scale bars in Panels A and B represent the length of the LPL1 neuron. Note, the ~500 μm length displayed above the scale bars is only an estimate. The LPL1 neuron in *A. californica* could range from 300-1000 μm in length.

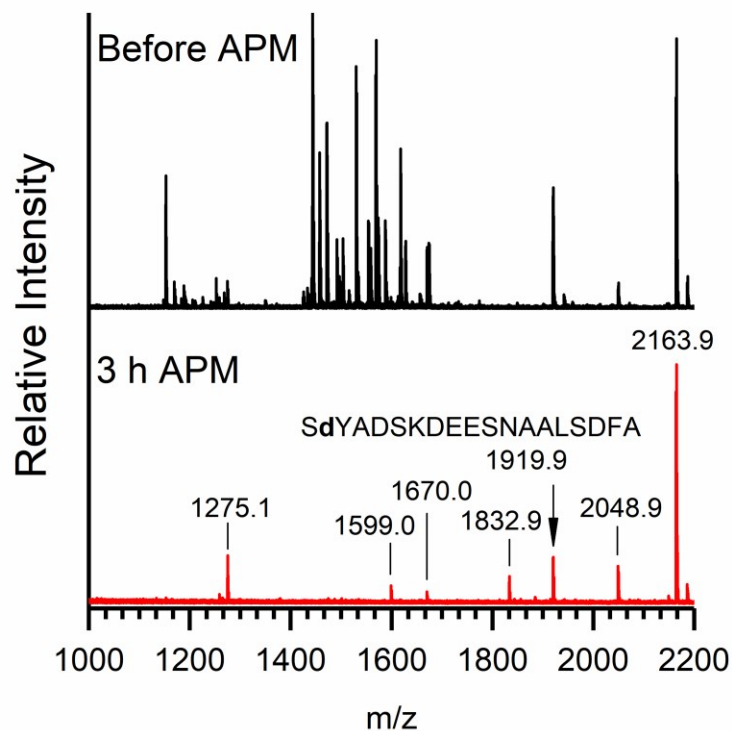


Figure 5.2. APM digestion on a single ApALNP containing pedal neuron. Representative of N = 3 single cell replicates. Before APM (black trace); after 3 h APM (red trace). The arrow shows the known DAACP, ApALNP-3 resisted APM. The $[m+H]^+$ of other peptides detected in APM treated neuron are labeled and correspond to the following peptides: SYADSKDEESNAALSDFAED ($[m+H]^+ = 2163.9$); SYADSKDEESNAALSDFAE ($[m+H]^+ = 2048.9$); SYADSKDEESNAALSDFA ($[m+H]^+ = 1919.9$); YADSKDEESNAALSDFA ($[m+H]^+ = 1832.9$); ADSKDEESNAALSDFA ($[m+H]^+ = 1670.0$); DSKDEESNAALSDFA ($[m+H]^+ = 1599.0$). $[m+H]^+ = 1275.1$ is an unknown ion.

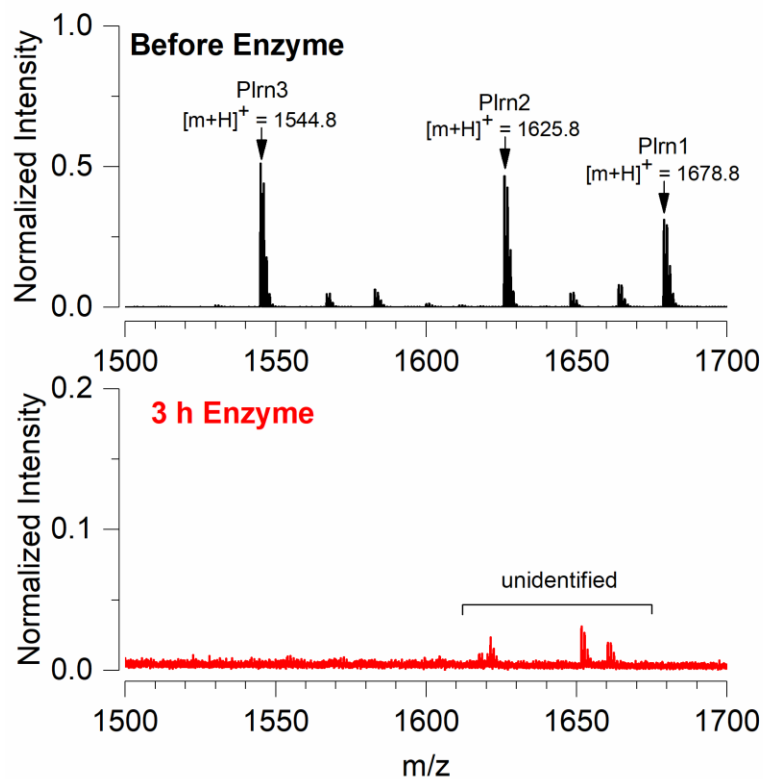


Figure 5.3. APM digestion on a single pleurin-containing neuron from the pleural ganglion. Representative of N=3 cells. The pleurin derived peptides of interest are indicated on the plot. Upper plot shows the MALDI MS spectrum of a pleural neuron before the APM digestion (black trace) and the lower plot shows the MALDI MS spectrum of the same neuron after APM digestion (Red trace).

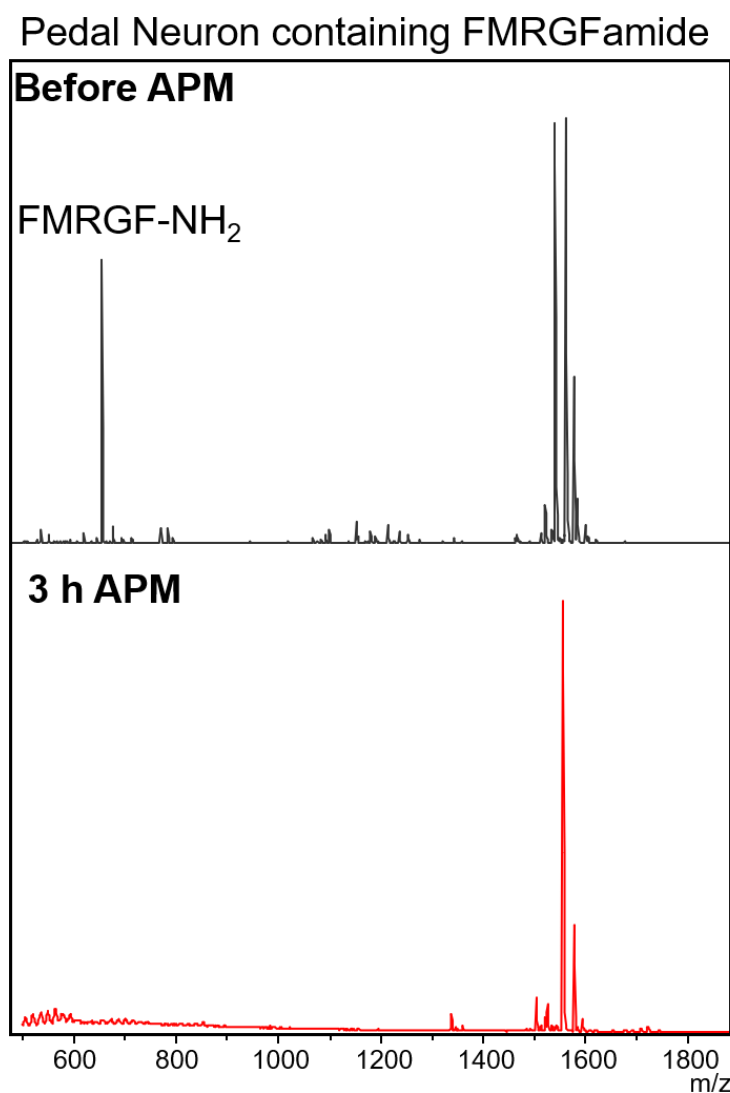


Figure 5.4. APM digestion on a single neuron isolated from the pedal ganglion containing FMRGF-NH₂. Upper trace shows the full MALDI MS spectrum of the FMRGFamide-neuron before APM digestion (black trace) and the lower trace shows the full MALDI MS spectrum of the same neuron after the 3 h APM digestion (red trace). Representative of N = 4 cells.

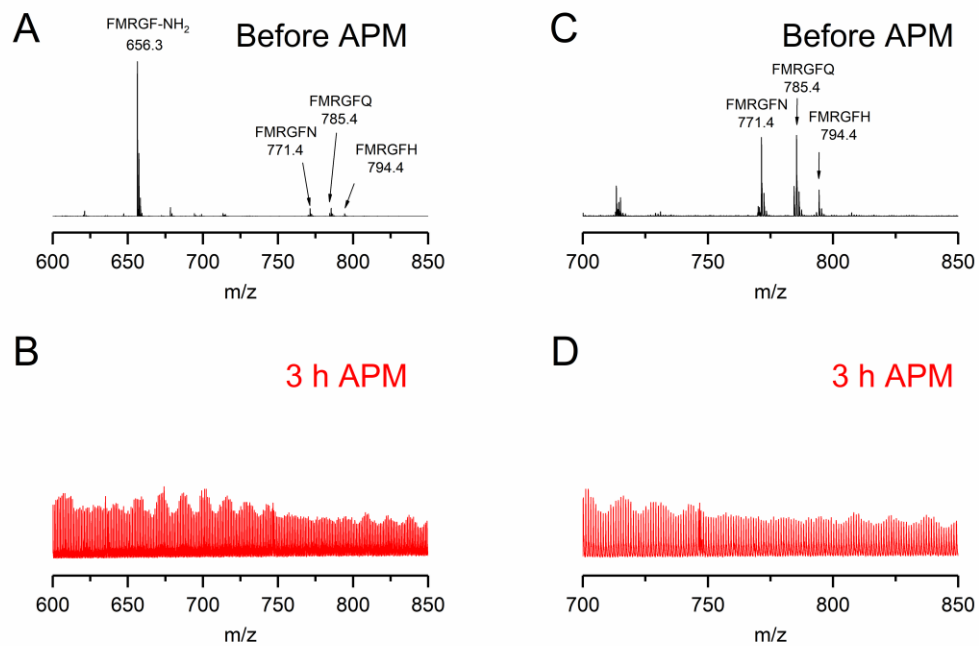


Figure 5.5. APM digestion results from an FMRGFamide-neuron. MALDI MS spectra of a the FMRGFamide neurons (A, C) before the APM digestion (black traces) at two different m/z windows; (B, D) after the APM digestion (red trace) at two different m/z windows. This result is representative of $N = 4$ replicates. Only peptides derived from the FMRGFamide precursor are labeled. FMRGFQ-NH₂ ($[m+H]^+ = 784.4$) is not labeled, but was also observed, and degraded by APM.

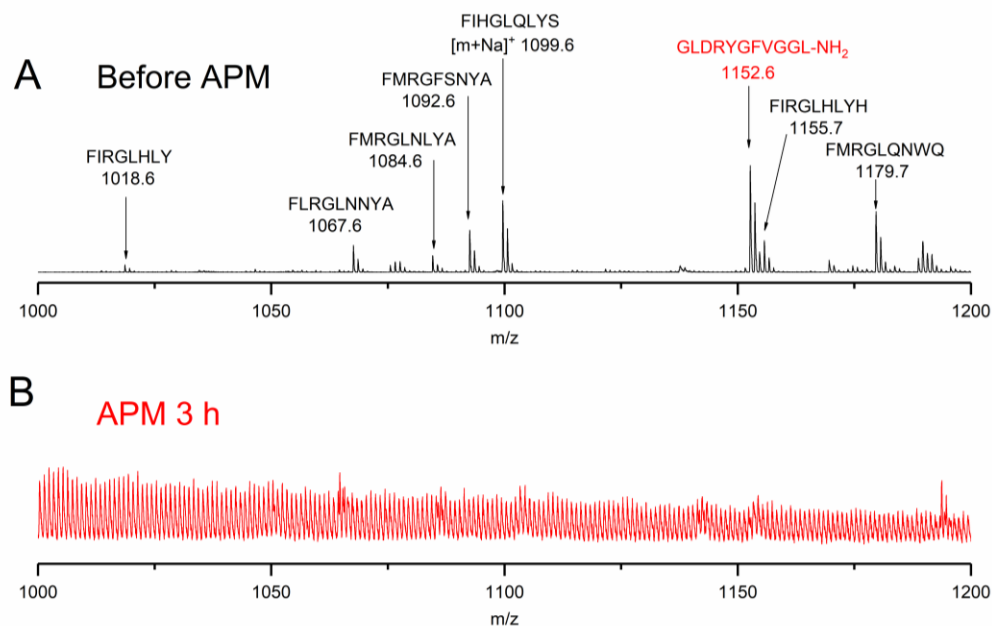


Figure 5.6. APM digestion results from an FMRGFamide-neuron. MALDI MS spectra of a the FMRGFamide neurons (A) before the APM digestion (black trace); (B) after the APM digestion (red trace). This result is representative of N = 4 replicates. Peptides derived from the FMRGFamide precursor are labeled in black and include (FIRGLHLY, FLRGLNNYA, FMRGLNLYA, FMRGFSNYA, FIHGLQLYS, FIRGLHLYH, FMRGLQNWQ). One peptide from buccalin (NP_001191649.1) is labeled in red (GLDRYGFVGGGL-NH₂). Unless otherwise specified, the [m+H]⁺ is displayed below the sequences. Note, [m+H]⁺ = 1092.5 is labeled as FMRGFSNYA from the FMRGFamide precursor, but it is isobaric with YDSFAYSAGL-NH₂ from buccalin and could be either one. [m+H]⁺ = 1155.6 labeled as FIRGLHLYH from the FMRGFamide precursor, but is isobaric with GMDDFAFSPGL-NH₂ from the buccalin precursor, and could be either one. [m+H]⁺ = 1053.6 (GMDSLAFSGGL-NH₂) from buccalin is present but not marked. [m+H]⁺ = 1024.5 (MDGFAFAPGL-NH₂) from buccalin was present but not marked. [m+H]⁺ = 1054.5 (MDSFAFAPGL-NH₂) may be present but overlaps with other peaks.

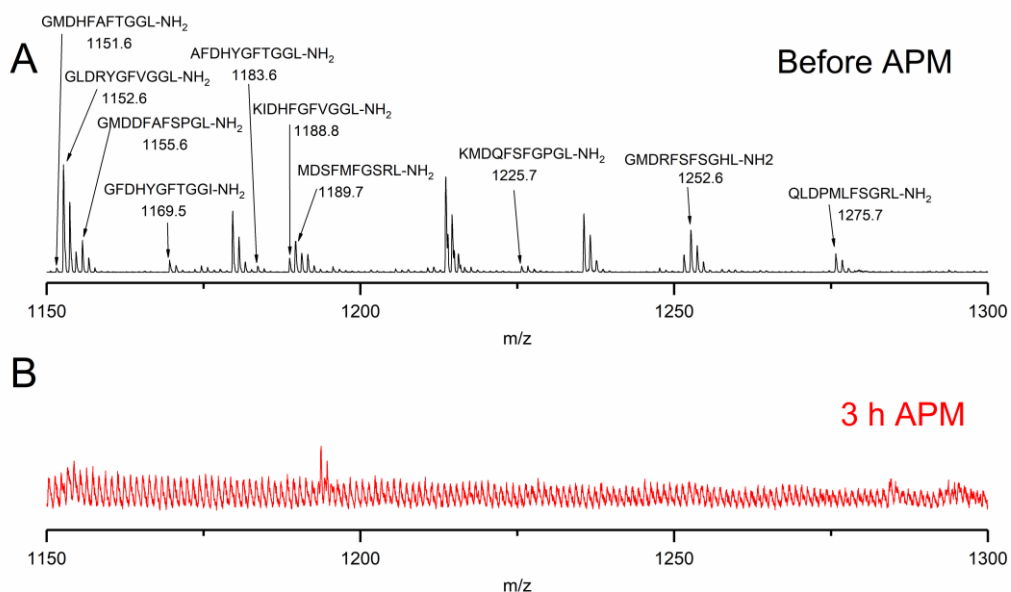


Figure 5.7. APM digestion results from an FMRGFamide-neuron. MALDI MS spectra of a the FMRGFamide neurons (A) before the APM digestion (black trace); (B) after the APM digestion (red trace). This result is representative of N = 4 replicates. Only peptides derived from the buccalin precursor (NP_001191649.1) are annotated. Unless otherwise specified, the $[m+H]^+$ is displayed below the sequences.

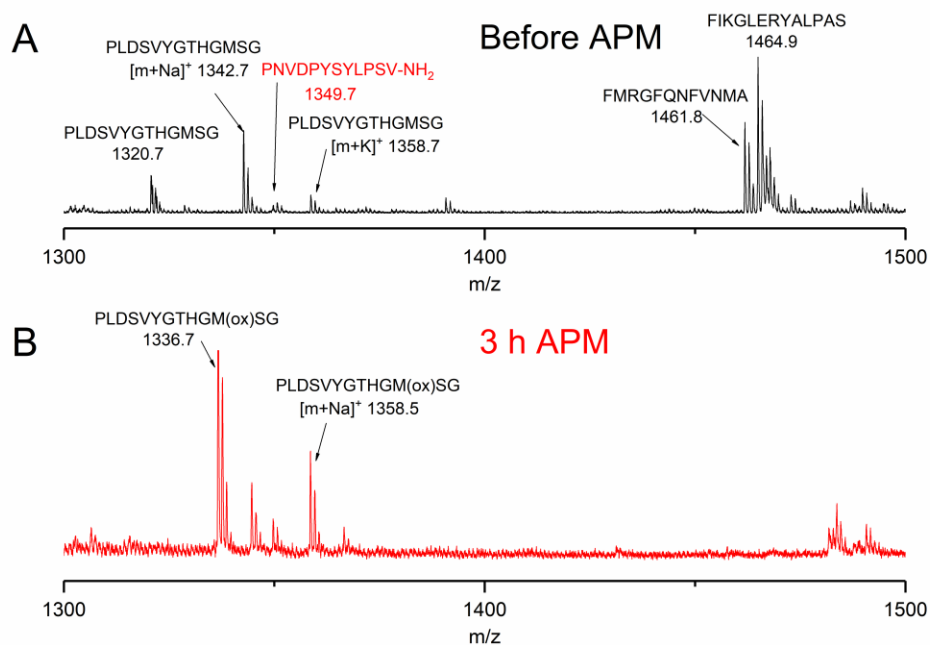


Figure 5.8. APM digestion results from an FMRGFamide-neuron. MALDI MS spectra of a the FMRGFamide neurons (A) before the APM digestion (black trace); (B) after the APM digestion (red trace). This result is representative of N=4 replicates. The PP1 (NP_001191585.1) derived peptide that is annotated is PLDSVYGTHGMSG. The buccalin (NP_001191649.1) derived peptide that is annotated is PNVDPYSYLPVS-NH₂. Two peptides derived from FMRGFamide are labeled FMRGFQNFVNMA, and FIKGLERYALPAS. Note, [m+H]⁺ = 1349.7 was detected after the APM digestion in Panel B, but is not labeled. Unless otherwise specified, the [m+H]⁺ values are labeled under the sequences.

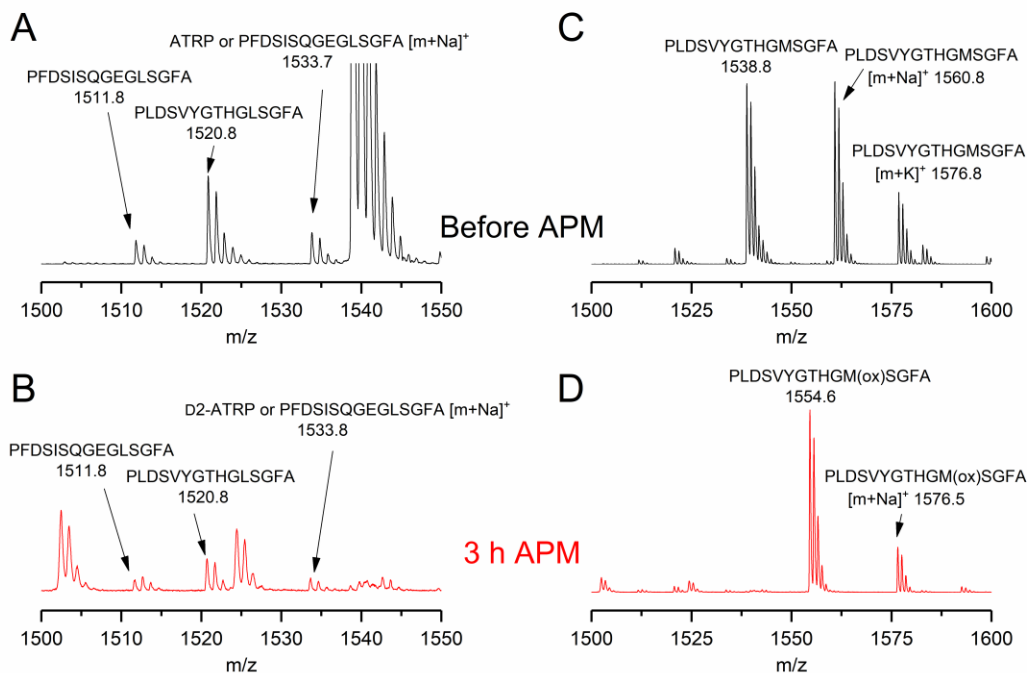


Figure 5.9. APM digestion results from an FMRGFamide-containing neuron. Representative of $N = 4$ replicates. (A, C) show MALDI MS spectra of the name neuron at different m/z ranges before the APM digestion (black traces), while (B, D) show the MALDI MS spectrum of the same neuron after the APM digestion (red traces). Only peptides derived from the PP1 precursor (NP_001191585.1) are annotated. Note, ApATRP from the *Aplysia* allatotropin-like precursor (NP_001191666.1) and PFDSISQGEGLSAGFA have a nominal mass difference of 22.1. PFDSISQGEGLSAGFA $[m+H]^+$ and PFDSISQGEGLSAGFA $[m+Na]^+$ have a nominal mass difference of 22.0. The m/z 1533.8 in both Panel A and Panel B is most likely PFDSISQGEGLSAGFA $[m+Na]^+$ and not ApATRP. Also note, while it may appear that m/z 1538.8 was completely digested, this was not actually the case. Rather, the Met residue oxidized during the APM digestion. The presence of m/z 1554.6 in Panel D indicates the peptide was resistant to APM.

Table 5.1. Putative assignments of peptides derived from the FMRGFamide precursor that were detected by MALDI MS on pedal neurons neurons. Observed $[m+H]^+$ is the average from N = 4 individual neurons. Stddev indicates standard deviation of the mean from N = 4 individual neurons. Ppm was calculated from the averaged observed $[m+H]^+$. Note, there are three possible proteoforms of FMRGFamide precursor: XP_005094941.1, XP_005094942.1, XP_035824739.1. The exact proteoform which is present could not be determined.

sequence	Observed $[m+H]^+$ avg.	Std. dev.	Expected $[m+H]^+$	ppm
FMRGF-NH ₂	656.378	0.029	656.3337	67.2
FMRGFG-NH ₂	713.386	0.026	713.3552	42.9
FMRGFN	771.392	0.030	771.3607	39.9
FMRGFQ-NH ₂	784.425	0.031	784.3923	41.7
FMRGFQ	785.413	0.034	785.3763	46.7
FMRGFH	794.403	0.031	794.3766	33.2
FMRGLSAY	944.490	0.041	944.4658	25.6
FIRGLHLY	1018.642	0.052	1018.583	57.9
FLRGLNNYA	1067.597	0.040	1067.563	31.8
FIHGLQLYS	1077.586	0.075	1077.573	12.1
FMRGLNLYA	1084.593	0.035	1084.561	29.5
FMRGFSNYA	1092.536	0.047	1092.493	39.4
FMRGLQNWQ	1179.632	0.055	1179.573	50.0

5.7 REFERENCES

1. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Advancing D-amino acid-containing peptide discovery in the metazoan. *Biochim Biophys Acta Proteins Proteom* **2021**, 1869 (1), 140553.
2. Ollivaux, C.; Soye, D.; Toullec, J. Y., Biogenesis of D-amino acid containing peptides/proteins: where, when and how? *J Pept Sci* **2014**, 20 (8), 595-612.
3. Bai, L.; Sheeley, S.; Sweedler, J. V., Analysis of Endogenous D-Amino Acid-Containing Peptides in Metazoa. *Bioanal Rev* **2009**, 1 (1), 7-24.
4. Bai, L.; Romanova, E. V.; Sweedler, J. V., Distinguishing endogenous D-amino acid-containing neuropeptides in individual neurons using tandem mass spectrometry. *Anal Chem* **2011**, 83 (7), 2794-800.
5. Mast, D. H.; Liao, H.-W.; Romanova, E. V.; Sweedler, J. V., Analysis of Peptide Stereochemistry in Single Cells by Capillary Electrophoresis–Trapped Ion Mobility Spectrometry Mass Spectrometry. *Analytical Chemistry* **2021**, 93 (15), 6205-6213.
6. Sheeley, S. A.; Miao, H.; Ewing, M. A.; Rubakhin, S. S.; Sweedler, J. V., Measuring D-amino acid-containing neuropeptides with capillary electrophoresis. *Analyst* **2005**, 130 (8), 1198-203.
7. Garden, R. W.; Shippey, S. A.; Li, L.; Moroz, T. P.; Sweedler, J. V., Proteolytic processing of the Aplysia egg-laying hormone prohormone. *Proc Natl Acad Sci U S A* **1998**, 95 (7), 3972-7.
8. Jiménez, C. R.; van Veelen, P. A.; Li, K. W.; Wildering, W. C.; Geraerts, W. P.; Tjaden, U. R.; van der Greef, J., Neuropeptide expression and processing as revealed by direct

- matrix-assisted laser desorption ionization mass spectrometry of single neurons. *J Neurochem* **1994**, 62 (1), 404-7.
9. Jansson, E. T.; Comi, T. J.; Rubakhin, S. S.; Sweedler, J. V., Single Cell Peptide Heterogeneity of Rat Islets of Langerhans. *ACS Chem Biol* **2016**, 11 (9), 2588-95.
 10. Li, L.; Garden, R. W.; Sweedler, J. V., Single-cell MALDI: a new tool for direct peptide profiling. *Trends Biotechnol* **2000**, 18 (4), 151-60.
 11. Rubakhin, S. S.; Romanova, E. V.; Nemes, P.; Sweedler, J. V., Profiling metabolites and peptides in single cells. *Nat Methods* **2011**, 8 (4 Suppl), S20-9.
 12. Bai, L.; Livnat, I.; Romanova, E. V.; Alexeeva, V.; Yau, P. M.; Vilim, F. S.; Weiss, K. R.; Jing, J.; Sweedler, J. V., Characterization of GdFFD, a D-amino acid-containing neuropeptide that functions as an extrinsic modulator of the Aplysia feeding circuit. *J Biol Chem* **2013**, 288 (46), 32837-51.
 13. Livnat, I.; Tai, H. C.; Jansson, E. T.; Bai, L.; Romanova, E. V.; Chen, T. T.; Yu, K.; Chen, S. A.; Zhang, Y.; Wang, Z. Y.; Liu, D. D.; Weiss, K. R.; Jing, J.; Sweedler, J. V., A d-Amino Acid-Containing Neuropeptide Discovery Funnel. *Anal Chem* **2016**, 88 (23), 11868-11876.
 14. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Differential Post-Translational Amino Acid Isomerization Found among Neuropeptides in Aplysia californica. *ACS Chem Biol* **2020**, 15 (1), 272-281.
 15. Hall, J. D.; Lloyd, P. E., Involvement of pedal peptide in locomotion in Aplysia: modulation of foot muscle contractions. *J Neurobiol* **1990**, 21 (6), 858-68.
 16. Hall, J. D.; Lloyd, P. E., Release of pedal peptide from Aplysia neurons in primary culture. *J Neurobiol* **1991**, 22 (6), 583-9.

17. Lloyd, P. E.; Connolly, C. M., Sequence of pedal peptide: a novel neuropeptide from the central nervous system of *Aplysia*. *J Neurosci* **1989**, *9* (1), 312-7.
18. Cropper, E. C.; Miller, M. W.; Tenenbaum, R.; Kolks, M. A.; Kupfermann, I.; Weiss, K. R., Structure and action of buccalin: a modulatory neuropeptide localized to an identified small cardioactive peptide-containing cholinergic motor neuron of *Aplysia californica*. *Proc Natl Acad Sci U S A* **1988**, *85* (16), 6177-81.
19. Miller, M. W.; Beushausen, S.; Cropper, E. C.; Eisinger, K.; Stamm, S.; Vilim, F. S.; Vitek, A.; Zajc, A.; Kupfermann, I.; Brosius, J.; et al., The buccalin-related neuropeptides: isolation and characterization of an *Aplysia* cDNA clone encoding a family of peptide cotransmitters. *J Neurosci* **1993**, *13* (8), 3346-57.
20. Li, L.; Moroz, T. P.; Garden, R. W.; Floyd, P. D.; Weiss, K. R.; Sweedler, J. V., Mass spectrometric survey of interganglionically transported peptides in *Aplysia*. *Peptides* **1998**, *19* (8), 1425-33.
21. Rubakhin, S. S.; Li, L.; Moroz, T. P.; Sweedler, J. V., Characterization of the *Aplysia californica* cerebral ganglion F cluster. *J Neurophysiol* **1999**, *81* (3), 1251-60.
22. Southey, B. R.; Amare, A.; Zimmerman, T. A.; Rodriguez-Zas, S. L.; Sweedler, J. V., NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Res* **2006**, *34* (Web Server issue), W267-72.
23. Checco, J. W.; Zhang, G.; Yuan, W. D.; Le, Z. W.; Jing, J.; Sweedler, J. V., *Aplysia* allatotropin-related peptide and its newly identified d-amino acid-containing epimer both activate a receptor and a neuronal target. *J Biol Chem* **2018**, *293* (43), 16862-16873.

CHAPTER 6

SCREENING FOR ISOMERIC PEPTIDE HORMONES IN RODENT PANCREATIC ISLETS OF LANGERHANS

6.1 NOTES AND ACKNOWLEDGEMENTS

This work was completed with help from J. Checco, and H. Tai. Thanks to S. Rubakhin for performing rodent dissections, and thanks to P. Yau and B. Imai at the University of Illinois Roy. J. Carver Biotechnology Protein Sciences Facility for performing Orbitrap measurements. Support for this work was provided by Award Number P30 DA018310 from the National Institute on Drug Abuse (NIDA) and Award No. RO1 NS031609 from the National Institute of Neurological Disorders and Stroke (NINDS).

6.2 INTRODUCTION

This chapter represents a preliminary exploration for peptides that undergo post-translational L- to D-residue isomerization in pancreatic Islets of Langerhans from *Rattus norvegicus*. Islet peptide extracts were screened for D-amino acid containing peptides (DAACPs) using liquid chromatography (LC) tandem mass spectrometry (MS/MS) based workflows outlined in Chapter 2 and applied in Chapter 3. The work presented in this chapter represent initial attempts at screening islet peptides. It does not aim to classify peptides as DAACPs, but simply to report the peptides that meet (partially or in full) the criteria for assigning DAACP candidates (i.e., APM resistance, split peak elution, and MS properties), discussed in Chapters 2 and 3 of this dissertation.

The general goal of DAACP screening is to identify all the potential isomeric peptides, then work backwards to determine which peptides are the most promising DAACP candidates, and which are likely false positives. The work reported in this chapter does not prove any islet peptides are DAACPs.

Pancreatic Islets of Langerhans are an endocrine structure in the pancreas responsible for producing several important hormones that regulate the glucose metabolism and other important physiological processes.¹ Islets are a composite structure of five endocrine cell types, each differentiated by the hormone(s) it secretes: α -cells produce glucagon (Gluc), β -cells produce Insulin (Ins), δ -cells produce somatostatin (Sms), γ -cells produce pancreatic polypeptide (Ppy) and ϵ -cells produce ghrelin.² But islets also express a host of other important prohormones. The ancient and conserved role of these peptides across the metazoan, and the abundance of peptide hormones makes pancreatic islets a desirable tissue to explore for potentially rare or unusual post-translational modifications (PTMs).

There are evolutionary reasons for exploring mammalian pancreatic tissue for DAACPs. As discussed in Chapter 2, most of the cell-cell signaling DAACPs discovered to-date have been discovered among brain structures in invertebrate species.³⁻⁵ Although invertebrates lack a pancreas, many peptide hormones ancestral to those secreted by vertebrate pancreatic tissues, are produced by cells in the central nervous system (CNS) and gut of invertebrate species.⁶ As an example, in crustaceans, the crustacean hyperglycemic hormone, a peptide known to undergo L- to D-residue isomerization and secreted by the X-organ, triggers the breakdown of glycogen in the hepatopancreas to release glucose.⁷ The participation of DAACPs in the invertebrate glucose metabolism is a compelling reason to screen for DAACPs amongst hormones involved in similar biological pathways in mammals.

Peptides and proteins of mammalian pancreatic islets have been intensely studied by peptidomics and proteomics,⁸⁻¹² but an in-depth evaluation of peptide stereoisomers has not been reported.

The trends and similarities between DAACPs across species were also discussed in Chapter 2. These trends showed DAACPs among both vertebrates and invertebrates, commonly have a D-residue at position 2 from the N-terminus. Therefore, it can be hypothesized that if *R. norvegicus* pancreatic islets have DAACPs, the likely position of a D-residue would be at position 2. Thus, the same methods that were applied to detected DAACPs in the *A. californica* CNS described in Chapter 3, can be applied to pancreatic islets to test for DAACPs.

Like in Chapter 3, in this chapter, an emphasis was placed on identifying isomeric peptides of *R. norvegicus* pancreatic islets by examining LC-MS chromatograms of peptide extracts for sequences that elute at multiple retention times. The detection of a peptide at two (or more) retention times by LC-MS/MS could indicate it has an endogenous diastereomer. After identifying peptides that elute at more than one retention time, LC-MS/MS analysis was performed on aminopeptidase M (APM) treated extracts. APM was used to degrade away most of the all-L residue peptides, leaving only the peptides that are stable against hydrolysis, such as DAACPS, undigested in the extract. Finally, the LC-MS chromatograms of the peptides that eluted at more than one retention time before the APM digestion were examined after the APM digestion. Any peptides that eluted at more than one retention time before the APM digestion and eluted at only one retention time after the APM digestion, were assigned as promising DAACP candidates. This general workflow for DAACP screening of a complex peptide mixture is the same workflow that was applied in Chapter 3 to successfully discover 8 novel DAACPs in *A. californica*.

6.3 METHODS

Pancreatic Islet Isolation.

Islets were isolated from rat pancreas using procedures based on previously published methods.¹³ Modified Grey's balanced salt solution (mGBSS) was prepared 1.5 mM CaCl₂, 4.9 mM KCl, 0.2 mM KH₂PO₄, 11 mM MgCl₂, 0.3 mM MgSO₄, 138 mM NaCl, 27.7 mM NaHCO₃, 0.8 mM NaH₂PO₄, and 25 mM HEPES dissolved in Milli-Q water pH =7.4. Pancreases were removed from rats into a 15 mL centrifuge tube containing 10-15 mL of 1 mg/mL collagenase P in MGBSS with 5 mM glucose and kept on ice. The tube was incubated at 37 °C for 30-40 min in a water bath, and every 10 min the pancreas was agitated by vigorously shaking the tube manually. This process was repeated until the solution had a consistency of pea soup. After incubation, the tube was placed in ice, and centrifuged at 0.3 ×g for 3 min at 4 °C. The centrifugation step was repeated until a pellet formed with a transparent supernatant. The supernatant was removed, and the pellet was washed with 7 mL of 5 mM glucose in MGBSS with 1% BSA by gently pipetting the pelleted material in and out of disposable transfer pipette until the pellet was fully dispersed. The material was then pelleted again by centrifugation at 0.3 ×g for 3 min at 4 °C and this was repeated 3 times. After the final wash, the supernatant was removed, and the pellet was reconstituted in 4 mL of 5 mM glucose in MGBSS with 1% BSA and dispersed using a disposable transfer pipette. The solution containing islets and acinar tissue was then poured directly into a petri dish and added islets were manually picked using a lab-made micropipette (see Chapter 4) into a separate petri dish containing fresh 5 mM glucose in MGBSS with 1% BSA and counted.

Tissue Extractions

Approximately 100-150 pancreatic islets were transferred directly from the MGBSS solution into 500 μ L acidified methanol (90:9:1 methanol: glacial acetic acid:water) using the lab-made micropipette,¹⁴ and about 100-200 μ L of MGBSS. The islets were mechanically homogenized at 4 °C, sonicated for 5 min at 4 °C, then centrifuged at 15,000 $\times$ g for 10 min. the supernatant was transferred into a new 1.5 mL centrifuge tube and the solvent was removed with a vacuum centrifuge. Islet extracts were reconstituted in 200 μ L of 5% methanol in water with 0.1-0.5% formic acid (FA) and desalted by solid phase extraction (SPE) using Pierce C18 spin columns (Thermo-Fisher scientific). Peptides were eluted off the spin column with 50% methanol in water. Eluent containing the peptides was dried using vacuum centrifugation and reconstituted in 5% CH₃CN in water with 0.1% FA, and the peptide content was quantified using a Pierce BCA protein assay kit (Thermo-Fisher Scientific).

Aminopeptidase M preparation

APM from porcine kidney (APM, CAS#: 9054-63-1) was purchased from Calbiochem (Darmstadt, Germany). The APM suspension (170 U/mL, 2.7 mg/mL) was stored under manufacturer recommended conditions until use. For preparation of APM for digestion of pancreatic islet extracts, the APM suspension was gently swirled in its container and then 20 μ L was transferred to a 0.5 mL PCR tube and centrifuged at low speed on a tabletop centrifuge for 1 min. After centrifugation, the supernatant was removed and the remaining sediment containing the APM was reconstituted in 40 μ L of 50 mM ammonium formate (NH₄HCO₂) buffer in water. The latter conditions were used for APM digestions on Rat pancreatic islet extracts.

APM Digestion Islets Extracts

Pancreatic islets were isolated, and the peptides were extracted as described in the previous sections. APM screening of islet peptides were carried out under the following conditions. Islet peptide extracts (20 µg/mL) in 50 mM NH₄HCO₂ at pH 7-8 with 1-2 U/mL APM. Prepping the solution. Dried islet peptide extracts were reconstituted in 5% CH₃CN, 0.1 % FA in water to a final concentration of 100 µg/mL. Then, 130 µL of the islet peptide extract was diluted to 20 µg/mL with 520 µL of 50 mM NH₄HCO₂ buffer (pH 7.5). The pH of the diluted extracts dropped to pH ~5 as determined by pH paper and was adjusted back to pH 7-8 with ammonium hydroxide (NH₄OH) solution (~2 µL of 1 M NH₄OH) as determined by pH paper. The solution was then divided into 3 separate 1.5 mL centrifuge tubes (200 µL/tube). To tube 1 (“No-APM control” sample), 1.87 µL of NH₄HCO₂ buffer was added. To tubes 2 and 3 (“APM-treated”), 1.87 µL of 2.7 mg/mL APM (170 U/mL) solution was added to a final activity of 1-2 U/mL. Each tube was incubated at 37 °C for 24 h. At 16 h, 100 µL was collected from each tube and the reaction quenched by cooling to 4 °C and adding 1 µL for formic acid and stored in at -20 °C for 8 h. At 24 h, the remaining reaction solutions in tube 1, tube 2 and tube 3 were removed from the warm room, the reaction was quenched on ice, and by adding 1 µL of formic acid. Each solution from the 16 h and 24 h time points were dried using a vacuum centrifuge. Each tube containing approximately 2.5 µg of peptide was reconstituted in 20 µL 5% CH₃CN in water, and 4 µL was of each sample was injected for LC-MS/MS analysis using an orbitrap.

LC-MS/MS Analysis

Nano LC-MS/MS analysis was performed using a Dionex Nano-RSLC system (Thermo Fisher Scientific) a Thermo-Fisher Scientific Orbitrap Fusion mass spectrometer. The column used

for the chromatographic separation was an Acclaim Pepmap 100 C18 nano-LC column [75 μ m inside diameter $\times$ 15 cm, 2 μ m particle size, 100 Å pore size (P/N 164534, Thermo Fisher Scientific)] at 35 °C with a flow rate of 300 nL/min. For reversed phase LC, solvent A was 5% acetonitrile (CH₃CN), 0.1% formic acid (FA), and solvent B was 95% CH₃CN, 0.1 % FA. Chromatographic separation condition was a 55 min linear gradient from 5% B to 55% B. For Nano-LC-MS/MS analysis 0.5-1 μ g of peptide extract was injected onto the column. The Nano-LC was connected to a Thermo-Fisher orbitrap fusion (Thermo-fisher Scientific) Bruker ImpactHD qTOF (Bruker Daltonics). Product ion spectra for both qTOF and Orbitrap MS were generated by collision induced dissociation (CID). A mascot database search against the NCBI *R. norvegicus* proteome was performed to identify peptides. The significance threshold was $p < 0.05$, and standard scoring was used.

Matrix Assisted Laser Desorption Ionization (MALDI) MS

APM digests on islet peptides were desalted using C18 zip tips, and 0.5 μ L was deposited onto an anchorchip polished steel 384 well MALDI target and mixed with 50 mg/mL dihydroxybenzoic acid (DHB) in 50/50 CH₃CN/water with 0.1% trifluoroacetic acid (TFA). For MALDI MS on individual islets, the islets were transferred directly from the petri dish onto the anchorchip target with the lab-made micropipette and excess liquid was removed. 1 μ L of saturated DHB in 0.1% TFA was then deposited over the islet and allowed to dry in air under a tungsten lamp. MALDI MS was calibrated using Bruker Peptide Calibration Standard II (P/N 8222570) for mass range between 700 and 3200 Da. The MS was operated in reflectron positive mode, with an m/z range from 400-6000 m/z . Averaged spectra acquired from 100-1000 laser shots were used.

MALDI MS was analyzed in FlexAnalysis 4 (Bruker Daltonics). MS spectra were exported in .txt format and opened in Origin Pro for plotting figures.

6.4 RESULTS AND DISCUSSION

Islet Peptides Characterization

Individual rat islets were analyzed by MALDI MS. The peptides detected were identified by mass matching $[m+H]^+$ ions detected by MALDI MS to peptide libraries assembled by LC-MS/MS sequencing. A representative MALDI MS spectrum of pancreatic islet is shown in Figure 6.1 and the sequences of peptides detected by MALDI MS in individual islets throughout the course of the experiments is provided in Tables 6.1 and 6.2.

Islets were transferred directly onto MALDI targets immediately after isolation, excess buffer was removed, and DHB solution was deposited over the islet and allowed to dry. This method of depositing in-tact cells directly onto MALDI MS targets has previously been applied for studying endogenous prohormone processing in *A. californica* neurons (see Chapter 5). Direct deposition of live cells onto the target minuses sample handling and reduces peptide degradation that can occur during extractions and desalting steps. This allows for a more accurate assessment of the endogenous prohormone processing products while minimizing extraction artifacts.

Notably, despite utilizing the protocol described above, Tables 6.1 and 6.2 show a significant number of N-terminal and C-terminal truncated latter sequences that could be artifactual. During the islet transfer from solution to the MALDI target, it is possible for cell lysis to occur, exposing endogenous peptides to extracellular proteases. This may result the detection of truncated sequences that are not the products of normal proteolytic processing in the secretory pathway. However, after the DHB solution is applied, endogenous enzymes are effectively

inactivated. Another factor that may affect the integrity of islets is the sample handling. From dissection to isolation, the preparation protocol exposes islets to a significant amount of mechanical, thermal, and chemical stress over a short, 2 h time frame, but the effects these might have on the islet peptide profiles was not investigated in detail. Images showing the morphologies of isolated islets are shown in Figure 6.2. MALDI MS spectra of islets consistently showed the same peptide profile, independent of morphological structure. MALDI MS measurements on cultured islets might reveal a peptide profile that more accurately reflects the natural prohormone processing in islets. Finally, although MALDI MS is considered a “soft” ionization technique, it is still possible for peptides to undergo post-source decay resulting in the detection of truncated sequences.

APM Screening

One of the primary reasons for screening pancreatic islets for DAACPs was to determine if the rat Ins1 C-peptide (EVEDPQVPQLELGGGPEAGDLQTLALEVARQ) and Ins2 C-peptide (EVEDPQVAQLELGGGPGAGDLQTLALEVARQ) have a D-residue. Preliminary results from previous Sweedler group members had found the rat Ins C-peptides were resistant to APM digestion, suggesting the possibility of a D-residue. However, the resistance of the Ins C-peptides to proteolysis can be explained by the sequence substrate specificity of APM toward peptides with acidic N-terminal residues. APM tends to degrade peptides with N-terminal Glu and Asp residues slower than hydrophobic or aliphatic residues at the N-terminus.¹⁵ The Ins1 and Ins2 C-peptides were partially purified by HPLC and hydrolyzed and Marfey’s chiral amino acid analysis was performed as described previously.¹⁶ L-Glu, L-Leu, and L-Val were detected in the hydrolysate, but D-Glu, D-Leu, or D-Val were not detected. Acid hydrolysis results supports the explanation

that the Ins C-peptides are resistant to APM proteolysis because of the N-terminal Glu-residues, and not because of a D-residue.

Figure 6.3 shows extracted ion chromatograms (EICs) of select peptides before and after the APM digestion. Truncated Ins1 and Ins2 C-peptide sequences with N-terminal Glu residues resisted APM digestion (as expected), while other truncated form of the Ins1 C-peptide with hydrophobic Leu at the N-terminus completely degraded. In addition, glicentin related polypeptide (GRPP) was nearly completely degraded (Figure 6.3). Just 8% of the GRPP signal remained after 40 h digestion and this is likely attributed to an incomplete digestion.

Figure 6.4 shows MALDI MS spectra of islet peptide extract spiked with synthetic YMGWMDF-NH₂, CDPGYIGSR, and human secretin at over 40 h APM digestion. The synthetic peptides were degraded away, while the Ins1 and Ins2 C-peptides resisted digestion. The APM resistance of the Ins1 and Ins2 C-peptides is likely explained in full by the N-terminal Glu residues. Glicentin related polypeptide (GRPP) resisted APM digestion in Figure 6.4, but this is likely because APM degrades longer peptides at a slower rate.¹⁵

Aside from the Ins C-peptides, the other peptides detected resisting APM in islets were not great candidates for DAACPs because most had N-terminal amino acid residues that make them resistant to hydrolysis by APM. The breakdown of peptide sequences observed before and after APM digestion based on N-terminal sequences is shown in Figure 6.5. Table 6.3 lists peptides that were detected after 40 h APM treatment that were expected to degrade. None of the APM resistant peptides in Table 6.3 were detected eluting at more than one retention time by LC-MS before the APM digestion. The presence of two eluting peaks by LC-MS before APM, and only one peak after APM is an important condition for identifying DAACPs (see Chapter 3).

Just because a peptide is resistant to APM, even if it was expected to degrade, does not necessarily indicate there is a D-residue at position 2. Complementary observations, such as multiple peak elution, help to increase confidence that a peptide is a DAACP (see Chapters 2 and 3). However, multiple peak elution was not observed before APM digestion for any of the APM resistant islet peptides reported in Table 6.3. The most likely explanation these peptides were partially resistant to APM is most likely because the APM digestion was incomplete.

It is difficult to precisely control or predict the kinetics of APM hydrolysis for a complex mixture of hundreds of peptides. The rate of peptide hydrolysis by APM can decrease substantially in a mixture of peptides compared to the pure peptides digested independently.¹⁷ To help drive the digestions to completion, APM can be added at a higher concentration, or the digestion can be increased, but modifying these variables can have unintended consequences. For example, increasing the incubation time allows for increased chances of Asn or Gln deamidation to occur.¹⁸⁻
²¹ Long term incubation of peptide extracts in solutions with pH > 7 at 37 °C, can increase the probability of spontaneous isomerization, or degradation of peptides.^{18, 20-22} Lastly, just because DAACPs are more resistant to APM than all-L peptides, it was demonstrated in Chapter 3 that even DAACPs with a D-residue in position 2 can be degraded by APM with enough time.

It is important to recognize that MS has such low detection limits that even peptides were >99% degraded, and the MS1 was barely detectable above the signal-to-noise, an MS/MS spectrum still had enough sequence coverage to be identified by a database search. Because of this, it is important to calculate the percentage of peptide that resisted digestion when assessing APM resistant peptides.

Peptides derived from other precursors such as secretogranin III (Scg3), chromogranin A (CGA), PCSK1, Urocortin III (UCNIII) were also observed in pancreatic islets during these

experiments, but were present at such low abundance, it was difficult to accurately evaluate the resistance to APM for these peptides by LC-MS because they were not consistently detected (Table 6.2). The high concentration of peptides derived from the most abundant hormones produced by islets (Ins, Gluc, Ppy, SMS and IAPP) made detection of lower abundance hormones by LC-MS more challenging because of LC column loading limits.

6.5 CONCLUSION

In this chapter, pancreatic islets were screened for DAACPs based on resistance to APM, and multiple peak elution by LC-MS/MS. No islet peptides could be confidently detected eluting at multiple retention times by LC-MS before the APM digestion. This suggests that DAACP forms of peptide could not be detected or were inseparable under the experimental conditions used. In addition, although several peptides resisted the APM digestion, the stability of these peptides could not confidently be attributed to the presence of a D-residue. Therefore, in this chapter, no DAACP candidates were identified in pancreatic islets. These results indicate the methods applied to discover DAACPs in Chapter 3, have not identified any DAACPs in *R. norvegicus* pancreatic islets. This does not prove DAACPs are absent from islets but does suggest alternative approaches may be necessary to discover DAACPs in Islets of Langerhans.

6.6 FIGURES AND DATA

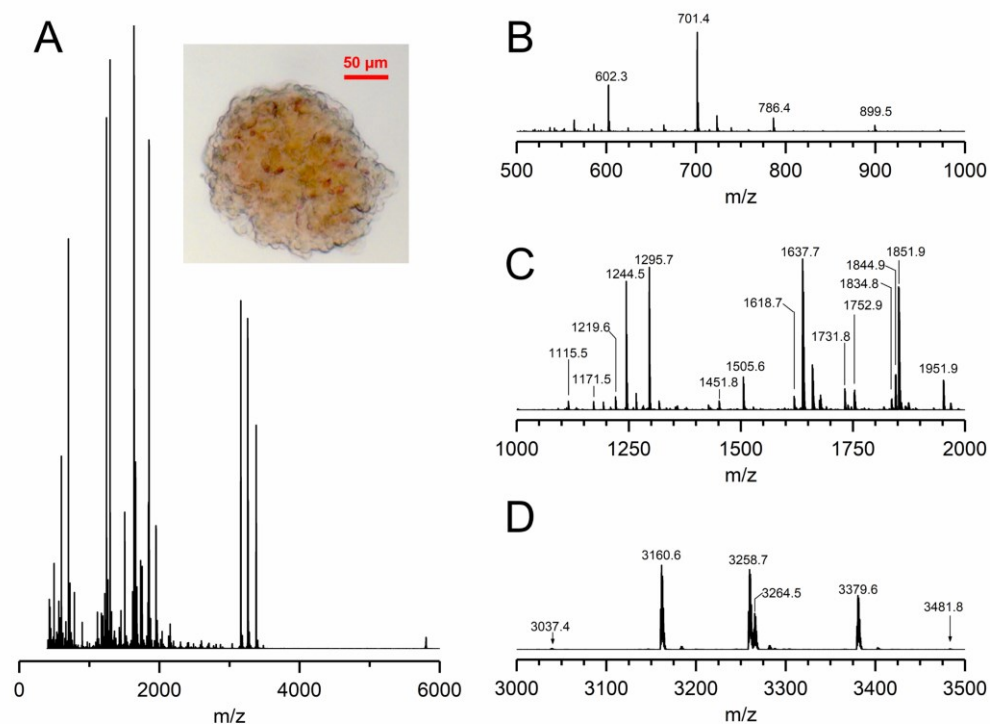


Figure 6.1. MALDI MS spectrum of an individual islet. (A) Representative MALDI MS spectrum obtained from a single *R. norvegicus* pancreatic islet. Inset in Panel A shows a representative islet isolated from surrounding acinar tissues using the isolation protocol. (B-D) zoomed in views of the MS spectrum shown in Panel A at different mass ranges. Each peak is labeled with the nominal $[m+H]^+$ observed. The peptides detected in individual islets by MALDI MS are provided in Tables 6.1 and 6.2.

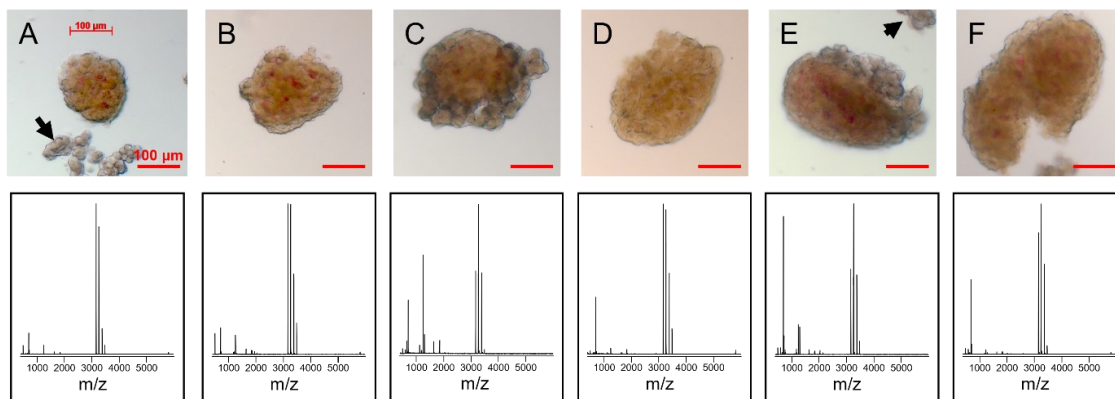


Figure 6.2. Morphology of pancreatic islets after performing the isolation protocol. The islet shown in (A) displays an islet with very little mechanical damage (indicated by the smooth rounded edges). Cells originating from surrounding acinar tissue are indicated in Panel A and Panel E by black arrows. Islets in (B-F) have signs of structural or cellular damage that likely resulted as an artifact of the isolation. In (C) and (E), dark cells appear attached islets may be related to cellular damage or the presence of acinar cells that were not completely removed from the surface. Although each islet displayed significantly different morphologies, the same peptides were generally detected consistently (albeit at different relative abundances) from islet to islet as demonstrated by the corresponding MALDI MS profiles displayed below each panel.

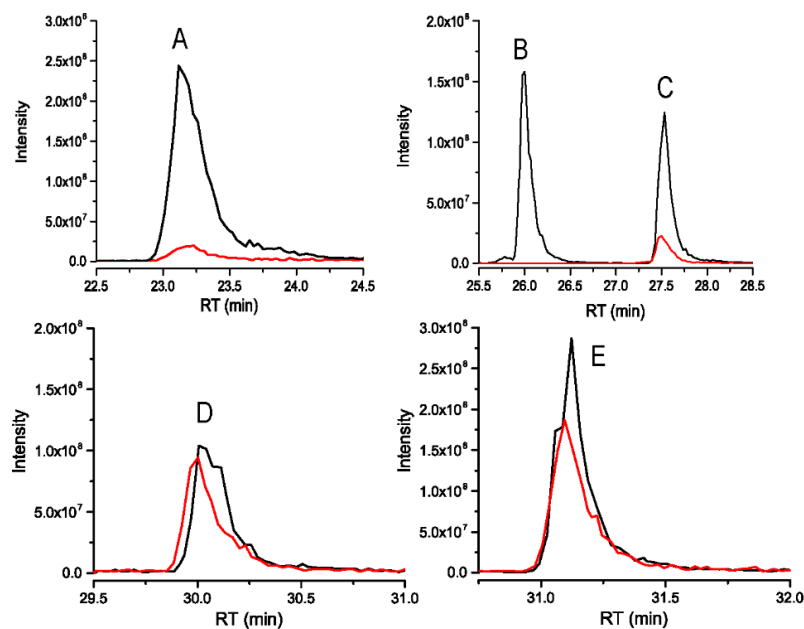


Figure 6.3. APM screening results showing select peptides derived from the Gluc and Ins precursors in the pancreatic islets. Plots show extracted ion chromatograms of peptides after 40 h incubation with no APM (Black Trace), and after incubation for 40h with APM (Red trace). (A) Full length glicentin related polypeptide (HAPQDTEENARSFPASQTEPLEDPNQINED, m/z 1127.17); (B) Ins2 C-peptide fragment (LELGGGPGAGDLQT, m/z 642.82); (C) Ins2 C-peptide fragment (ELGGGPGAGDLQTL, m/z 642.82); (D) Ins2 C-peptide fragment (EVEDPQVAQLELGGGPGAGDLQ m/z 1090.29); (E) Ins1 C-peptide fragment (EVEDPQVPQLELGGGPEAGDLQ m/z 1139.04).

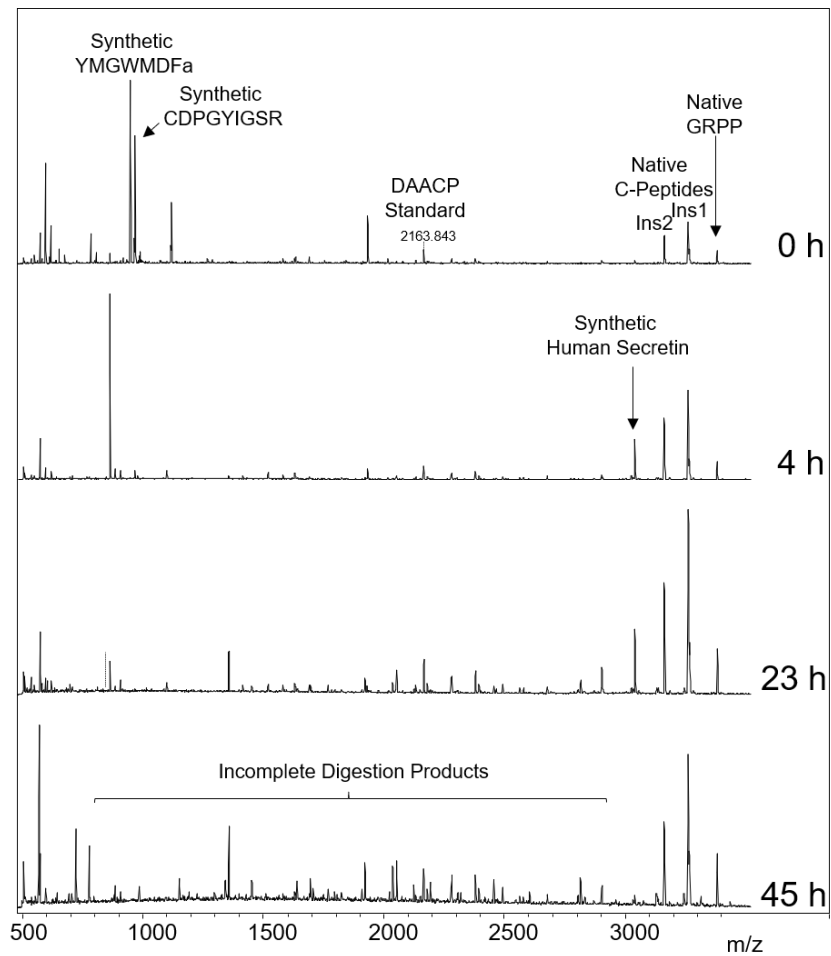


Figure 6.4. APM digestion on islets extracts spiked with synthetic peptide standards. Plots show MALDI MS spectra of the APM treated extract at various time points during the digestion. Native Ins1 C-Peptide ($[m+H]^+ = 3258.6$); native Ins2 C-peptide ($[m+H]^+ = 3160.6$); native glicentin related polypeptide (GRPP, $[m+H]^+ = 3379.5$); synthetic peptide YMGWMDF-NH₂ ($[m+H]^+ = 948.4$); synthetic peptide CDPGYIGSR ($[m+H]^+ = 967.4$); synthetic peptide human secretin HSDGTFTSELSRLREGARLQRLQLQGLV-NH₂ ($[m+H]^+ = 3038.7$); the synthetic DAACP standard sequence is SYADSKDEESNAALSDFEAD (ApALNP-3, $[m+H]^+ = 2163.8$).

Table 6.1. Peptides detected using MALDI MS on individual islets that matched sequences also observed in islets extracts by LC-MS/MS derived from the Ins1 and Ins2 C-peptides. Note, latter sequences and truncated sequences that are not flanked by dibasic or monobasic cleavage sites in the precursor may be present due to an artifact of sample preparation and do not necessarily represent endogenous peptides.

Peptide Sequence	Nominal Monoisotopic Mass	Prohormone
EVEDPQVAQLELGGGPGAGDLQ	2178.0	Ins 2
EVEDPQVAQLELGGGPGAGDLQT	2279.1	Ins 2
EVEDPQVAQLELGGGPGAGDLQTL	2392.2	Ins 2
EVEDPQVAQLELGGGPGAGDLQTLA	2463.2	Ins 2
EVEDPQVAQLELGGGPGAGDLQTLAL	2576.3	Ins 2
EVEDPQVAQLELGGGPGAGDLQTLALEVARQ (full length insulin 2 C peptide)	3159.6	Ins 2
PQVAQLELGGGPGAGDLQTLALEVARQ	2687.4	Ins 2
VAQLELGGGPGAGDLQTLALEVARQ	2462.3	Ins 2
ELGGGPGAGDLQTLALEVARQ	2051.1	Ins 2
LGGGPGAGDLQTLALEVARQ	1922.0	Ins 2
GGGPGAGDLQTLALEVARQ	1808.9	Ins 2
GPGAGDLQTLALEVARQ	1694.9	Ins 2
PQVAQLELGGGPGAGDLQTL	1919.9	Ins 2
LELGGGPGAGDLQT	1283.6	Ins 2
EDPQVAQLELGGGPGAGDLQTL	2164.1	Ins 2
EVEDPQVPQL	1152.6	Ins 1
EVEDPQVPQLELGGGP	1662.8	Ins 1
EVEDPQVPQLELGGGPE	1791.9	Ins 1
EVEDPQVPQLELGGGPEAG	1919.9	Ins 1
EVEDPQVPQLELGGGPEAGDLQ	2276.1	Ins 1
EVEDPQVPQLELGGGPEAGDLQT	2377.1	Ins 1
EVEDPQVPQLELGGGPEAGDLQTL	2490.2	Ins 1
EVEDPQVPQLELGGGPEAGDLQTLA	2561.3	Ins 1
EVEDPQVPQLELGGGPEAGDLQTLAL	2674.3	Ins 1
EVEDPQVPQLELGGGPEAGDLQTLALEVARQ (full length insulin 1 C peptide)	3257.6	Ins 1
PQVPQLELGGGPEAGDLQTLALEVARQ	2785.4	Ins 1
ELGGGPEAGDLQTLALEVARQ	2123.1	Ins 1
GGGPEAGDLQTLALEVARQ	1881.0	Ins 1
GGPEAGDLQTLALEVARQ	1823.9	Ins 1
PEAGDLQTLALEVARQ	1709.9	Ins 1
PQVPQLELGGGPEAGDLQTLALEVA	2501.3	Ins 1
LELGGGPEAGDLQT	1355.7	Ins 1
AGDLQTLALEVARQ	1483.8	Ins 1 & 2
GDLQTLALEVARQ	1412.8	Ins 1 & 2
DLQTLALEVARQ	1355.7	Ins 1 & 2
LQTLALEVARQ	1240.7	Ins 1 & 2
LALEVARQ	898.5	Ins 1 & 2
ALEVARQ	785.4	Ins 1 & 2
EVARQ	601.3	Ins 1 & 2

Table 6.2. Peptides detected using MALDI MS on individual islets that matched sequences also observed in islets extracts by LC-MS/MS derived from other prohormones. Note, latter sequences and truncated sequences that are not flanked by dibasic or monobasic cleavage sites in the precursor may be present due to an artifact of sample preparation and do not necessarily represent endogenous peptides.

Peptide Sequence	Nominal Monoisotopic Mass	Prohormone
NVAEDPNRESLDFLL	1843.9	IAPP
NVAEDPNRESLDFLL	1730.8	IAPP
NVAEDPNRESLDFL	1617.8	IAPP
NVAEDPNRESLDF	1504.7	IAPP
EDPNRESLDFLL	1559.8	IAPP
DPNRESLDFLL	1430.7	IAPP
DPNRESLDFLL	1317.7	IAPP
KC(-1.01)NTATC(-1.01)ATQRLANFLVRSSNNLGPVLPPTNVGSNTY _a	3918.0	IAPP
KC(-1.01)NTATC(-1.01)ATQRLANFLV	1850.9	IAPP
KC(-1.01)NTATC(-1.01)ATQRLANFL	1751.8	IAPP
LVRSSNNLGPVLPPTNVGSNTY _a	2297.2	IAPP
SSNNLGPVLPPTNVGSNTY _a	1928.9	IAPP
TPVGSNTNPQVD	1170.5	IAPP
HSQGTFTSDYSKYLDS	1834.8	Gluc
AQDFVQWLMNT	1351.6	Gluc
HSQGTFTSDYSKYLDSRRAQDFVQWLMNT	3480.6	Gluc
SWQHAPQDTEENARSFPASQTEPLEDPNQINED	3779.7	Gluc
SWQHAPQDTEENARSFPASQTEPLEDPNQINE	3664.6	Gluc
HAPQDTEENARSFPASQTEPLEDPNQINED	3378.5	Gluc
HAPQDTEENARSFPASQTEPLEDPNQINE	3263.5	Gluc
DFPEEVAIAEEL _a	1359.6	Gluc
SFPASQTEPLEDPNQINED	2130.0	Gluc
SFPASQTEPLEDPNQINE	2014.9	Gluc
HAPQDTEENA	1110.5	Gluc
NRNNIA	700.4	Gluc
SWQHAPQDTEENA	1511.6	Gluc
APLEPMYPGDYATHEQRAQYETQLRRYINTLTRPRY _a	4396.1	Ppy
APLEPMYPGDYATHEQRAQYETQL	2807.3	Ppy
APLEPMYPGDYATHEQ	1817.8	Ppy
AQYETQLRRYINTLTRPRY _a	2440.3	Ppy
RYINTLTRPRY _a	1451.8	Ppy
YINTLTRPRY _a	1294.7	Ppy
DEDTAGLPGRQLPPC(-1.01)TSLLVGLMPC(-1.01)AAARS	3036.5	Ppy
SANSNPAMAPRE	1243.6	SMS
SANSNPAMAPR	1114.50	SMS
SANSNPAMAPRERKAGCKNFFWKFTFTSC	3146.5	SMS
AGC(-1.01)KNFFWKFTFTSC(-1.01)	1636.7	SMS
LPQAAEQDEMRLQL	1769.9	SMS
APSDPRLRQFLQ	1426.8	SMS
APSDPRLRQF	1185.6	SMS
AVDQDLGPEVPPENVLGALLRV	2300.2	PCSK1
SLSAASAPLAETSTPLRL	1784.0	PCSK1
LENSSPQAPARRLLPP	1745.0	PCSK1
AYGFRDPGPQL	1219.6	CMGA
WSRMDQLAKELTAE	1676.8	CMGA
WSRMDQLAKELTA	1547.8	CMGA
WSRMDQLAKELT	1476.7	CMGA
LEGEDDPDRSMKLSFRA	1964.93	CMGA
ELSAERPLNEQIAEAEADKI	2225.1	SCGIII
ELSAERPLNEQIAEAEAD	1983.9	SCGIII
NKLEDVPVLS	1112.6	UCN III
NFGYLPTQDPSGEEDEQKHIKN	2674.2	UCN III
NFGYLPTQDPSGEEDEQKHI	2432.1	UCN III

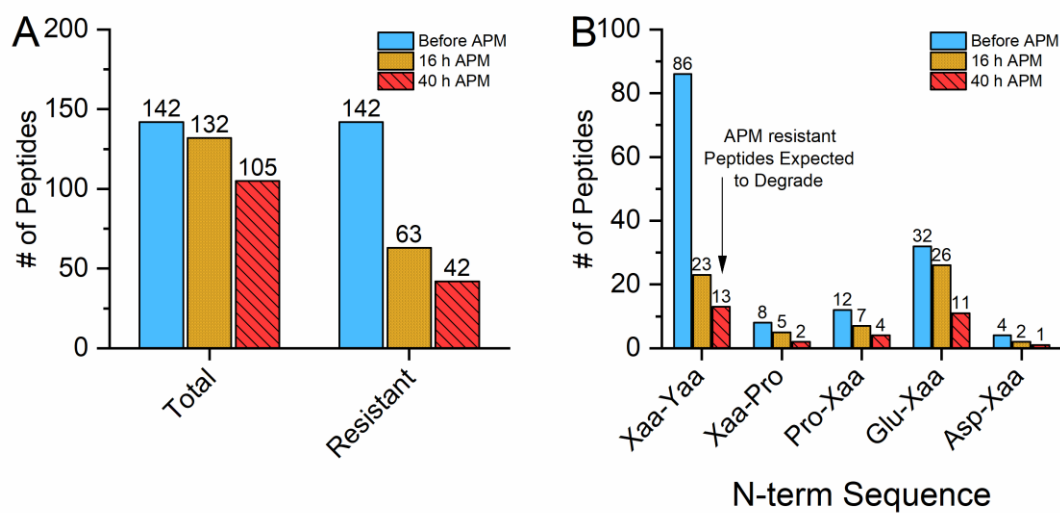


Figure 6.5. Breakdown of APM screening results. (A) Total number of peptides detected by LC-MS/MS before APM (blue bars), 16 h with APM (orange bars), and 40 h with APM (striped, red bars). (B) Number of sequences detected with specific N-terminal sequences before APM (blue bars), after 16 h with APM (orange bars), and 40 h with APM (striped, red bars). The number of peptides plotted in this figure excludes sequences that were unique the incubated samples. The percentage of peptide signal that remained after the digestion is evaluated in Table 6.3. MS has such low detection limits that even small amounts may be detected and result in a false positive “DAACP candidate.” Yaa indicates any amino acid residue other than Pro. Xaa indicates any amino acid residue other than Glu, Asp or Pro.

Table 6.3. Peptides detected before APM digestion and after 40 h APM digestion that were expected to be fully degraded based on the N-terminal residues. These results were acquired from APM screening using orbitrap for detection. % Degraded was calculated from the 40 h “no-APM control” and the 40 h “APM-treated” samples using the chromatogram peak intensities of EICs of the *m/z* observed. The ppm mass error, Mascot Score and Expect values are also reported. The sequence in this table corresponds to the 13 APM resistant peptides in Figure 6.5B. The sequences likely represent false positive DAACP candidates. Note certain full-length peptides (i.e., Ins 1 and Ins2 C-peptides) were not detected before or after APM digestion despite having been determined to be resistant because these peptides were not retained during the extraction and desalting procedure used.

Sequence	<i>m/z</i> obs.	ppm	Score	Expect	% Degraded	Precursor
HAPQDTEENARSFPASQTEPLEDPNQINED	1127.17	-0.7	112	9.5×10^{-9}	92	Gluc
HAPQDTEENARSFPASQTEPLEDPNQINE ^a	1088.83	-1.4	95	6.4×10^{-7}	77	
HAPQDTEENARSFPASQTEPLEDPNQIN ^a	1045.81	-0.9	80	2.7×10^{-5}	42	
SFPASQTEPLEDPNQINED ^b	1065.98	-0.1	49	2.2×10^{-2}	>99	
SFPASQTEPLEDPNQINE ^a	1008.46	-0.2	57	5.4×10^{-3}	89	
SQTEPLEDPNQINE ^a	807.37	-1.0	47	4.5×10^{-2}	79	
TEPLEDPNQINE ^a	699.82	-0.7	84	8.7×10^{-6}	85	
LELGGGPEAGDLQ ^a	628.31	-0.7	77	6.6×10^{-5}	>99	Ins1
LELGGGPEAGDLQT ^{a,c}	678.83	-1.8	69	3.6×10^{-4}	- ^c	
LELGGGPGAGDLQ ^a	592.30	-0.6	55	8.6×10^{-3}	>99	Ins2
QLELGGGPGAGDLQ ^a	656.33	-1.1	56	9.4×10^{-3}	93	
NVAEDPNRESLDF ^a	620.36	-0.1	63	1.1×10^{-3}	84	IAPP
NFEDPKFEVLDPKQS ^{a,c,d}	598.30	-0.8	65	1.2×10^{-3}	- ^d	ATP-scf6m

^a Peptide is a truncated sequence, and not necessarily an endogenously produced peptide.

^b MS1 signal below signal to noise in the “40 h APM-treated sample,” but high confidence MS2

^c Observed “before APM”, and “40 h APM-treated” samples but not in the “40 h no-APM control”

^d High confidence MS2, but MS1 signal below signal to noise in all samples. % Resistant unreliable

Gluc: NP_036839.2

Ins1: AAA41439.1

Ins2: AAA41440.1

IAPP: NP_036718.1

ATP-scf6m (ATP synthase-coupling factor 6, mitochondrial precursor): NP_446054.1

6.7 REFERENCES

1. Date, Y.; Nakazato, M.; Hashiguchi, S.; Dezaki, K.; Mondal, M. S.; Hosoda, H.; Kojima, M.; Kangawa, K.; Arima, T.; Matsuo, H.; Yada, T.; Matsukura, S., Ghrelin is present in pancreatic alpha-cells of humans and rats and stimulates insulin secretion. *Diabetes* **2002**, *51* (1), 124-9.
2. Campbell, J. E.; Newgard, C. B., Mechanisms controlling pancreatic islet cell function in insulin secretion. *Nat Rev Mol Cell Biol* **2021**, *22* (2), 142-158.
3. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Advancing d-amino acid-containing peptide discovery in the metazoan. *Biochim Biophys Acta Proteins Proteom* **2021**, *1869* (1), 140553.
4. Bai, L.; Sheeley, S.; Sweedler, J. V., Analysis of Endogenous D-Amino Acid-Containing Peptides in Metazoa. *Bioanal Rev* **2009**, *1* (1), 7-24.
5. Ollivaux, C.; Soye, D.; Toullec, J.-Y., Biogenesis of d-amino acid containing peptides/proteins: where, when and how? *Journal of Peptide Science* **2014**, *20* (8), 595-612.
6. Slack, J. M., Developmental biology of the pancreas. *Development* **1995**, *121* (6), 1569-80.
7. Chen, H. Y.; Toullec, J. Y.; Lee, C. Y., The Crustacean Hyperglycemic Hormone Superfamily: Progress Made in the Past Decade. *Front Endocrinol (Lausanne)* **2020**, *11*, 578958.
8. Taylor, S. W.; Nikoulina, S. E.; Andon, N. L.; Lowe, C., Peptidomic profiling of secreted products from pancreatic islet culture results in a higher yield of full-length peptide hormones than found using cell lysis procedures. *J Proteome Res* **2013**, *12* (8), 3610-9.

9. Boonen, K.; Baggerman, G.; D'Hertog, W.; Husson, S. J.; Overbergh, L.; Mathieu, C.; Schoofs, L., Neuropeptides of the islets of Langerhans: a peptidomics study. *Gen Comp Endocrinol* **2007**, *152* (2-3), 231-41.
10. Metz, T. O.; Jacobs, J. M.; Gritsenko, M. A.; Fontès, G.; Qian, W. J.; Camp, D. G., 2nd; Poitout, V.; Smith, R. D., Characterization of the human pancreatic islet proteome by two-dimensional LC/MS/MS. *J Proteome Res* **2006**, *5* (12), 3345-54.
11. Waanders, L. F.; Chwalek, K.; Monetti, M.; Kumar, C.; Lammert, E.; Mann, M., Quantitative proteomic analysis of single pancreatic islets. *Proc Natl Acad Sci U S A* **2009**, *106* (45), 18902-7.
12. Petyuk, V. A.; Qian, W. J.; Hinault, C.; Gritsenko, M. A.; Singhal, M.; Monroe, M. E.; Camp, D. G., 2nd; Kulkarni, R. N.; Smith, R. D., Characterization of the mouse pancreatic islet proteome and comparative analysis with other mouse tissues. *J Proteome Res* **2008**, *7* (8), 3114-26.
13. Carter, J. D.; Dula, S. B.; Corbin, K. L.; Wu, R.; Nunemaker, C. S., A practical guide to rodent islet isolation and assessment. *Biol Proced Online* **2009**, *11*, 3-31.
14. Mast, D. H.; Liao, H. W.; Romanova, E. V.; Sweedler, J. V., Analysis of Peptide Stereochemistry in Single Cells by Capillary Electrophoresis-Trapped Ion Mobility Spectrometry Mass Spectrometry. *Anal Chem* **2021**, *93* (15), 6205-6213.
15. Turner, A. J., Aminopeptidase N. *Handbook of Proteolytic Enzymes* **2013**, 397-403.
16. Livnat, I.; Tai, H. C.; Jansson, E. T.; Bai, L.; Romanova, E. V.; Chen, T. T.; Yu, K.; Chen, S. A.; Zhang, Y.; Wang, Z. Y.; Liu, D. D.; Weiss, K. R.; Jing, J.; Sweedler, J. V., A d-Amino Acid-Containing Neuropeptide Discovery Funnel. *Anal Chem* **2016**, *88* (23), 11868-11876.

17. Palmieri, F. E.; Bausback, H. H.; Ward, P. E., Metabolism of vasoactive peptides by vascular endothelium and smooth muscle aminopeptidase M. *Biochemical Pharmacology* **1989**, 38 (1), 173-180.
18. Geiger, T.; Clarke, S., Deamidation, isomerization, and racemization at asparaginy and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *Journal of Biological Chemistry* **1987**, 262 (2), 785-794.
19. Robinson, N. E.; Robinson, Z. W.; Robinson, B. R.; Robinson, A. L.; Robinson, J. A.; Robinson, M. L.; Robinson, A. B., Structure-dependent nonenzymatic deamidation of glutaminy and asparaginy pentapeptides. *J Pept Res* **2004**, 63 (5), 426-36.
20. Li, X.; Lin, C.; O'Connor, P. B., Glutamine deamidation: differentiation of glutamic acid and gamma-glutamic acid in peptides by electron capture dissociation. *Anal Chem* **2010**, 82 (9), 3606-15.
21. Wright, H. T., Nonenzymatic deamidation of asparaginy and glutaminy residues in proteins. *Crit Rev Biochem Mol Biol* **1991**, 26 (1), 1-52.
22. Aswad, D. W.; Paranandi, M. V.; Schurter, B. T., Isoaspartate in peptides and proteins: formation, significance, and analysis. *J Pharm Biomed Anal* **2000**, 21 (6), 1129-36.

CHAPTER 7

SCREENING FOR ISOMERIC NEUROPEPTIDES IN RODENT CENTRAL AND PERIPHERAL NERVOUS SYSTEM TISSUES

7.1 NOTES AND ACKNOWLEDGEMENTS

Trapped ion mobility measurements on rat hypothalamus and pituitary extracts were completed in collaboration with S. Okyem and E. Romanova. APM screening of spinal cord and peptide synthesis was completed with assistance from H. Chiu. J. Checco shared LC-MS/MS peptide identification results for hypothalamus and spinal cord for the isomeric peptide screening. This work was supported by the National Institutes of Health, National Institute of Neurological Disorders and Stroke by Award No. R01NS031609, and the National Institute on Drug Abuse by Award No. P30DA018310.

7.2 INTRODUCTION

In this chapter, the methods described in Chapter 3 have been employed to screen for D-amino acid containing peptides (DAACPs) in the model organism *Rattus norvegicus*. This chapter describes a preliminary examination of *R. norvegicus* for DAACPs in spinal cord tissue using the aminopeptidase M (APM) digestion protocols employed in Chapter 3.

Briefly, liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis was performed on *R. norvegicus* spinal cord extracts that were treated with APM and extracts that were not treated with APM. Then, the peptides present in each were compared to identify APM resistant peptides. Importantly, Chapter 3 defined a set of LC-MS/MS and chemical properties a peptide

must exhibit during this screening protocol to be considered a DAACP candidate. Specifically, peptides were only considered a possible DAACP if it met both of the following conditions: 1) the peptide must be detected at two (or more) retention times (with well-defined peaks) by LC-MS/MS before the APM digestion and 2) at least one of the peaks that was observed before the APM digestion must be absent from the LC-MS chromatogram after the APM digestion.

In Chapter 6, no peptides were detected in *R. norvegicus* pancreatic islets that exhibited all the properties that were expected of a DAACP described above. Therefore, in Chapter 6, no DAACPs could be identified.

In this chapter, peptides that met the criteria for identifying DAACPs (in part or in full) established in previous Chapter 3 (i.e., APM resistance, elution at multiple retention times by chromatography) were reported. They represent the most promising candidates for future follow-up analysis. As discussed in Chapter 2, two mammalian DAACPs have been isolated from platypus venom, but functional cell-cell signaling DAACPs have not been discovered in mammals to-date.¹ The prevalence of enzyme derived cell-cell signaling DAACPs in mammalian model organisms has not been explored.

Many mammalian neuropeptides and hormones were sequenced and characterized in the mid-to late 20th century. Owing to technological constraints, sequencing and analysis of peptides was time consuming and labor intensive. But also, a simple lack of sensitive analytical methods made identifying certain post-translational modifications (PTMs) difficult, including isomerization.¹ Additionally, other peptides have been isolated, but elicit no significant bioactivity or receptor binding/activation when it is expected.¹⁻² The undetected presence of a D-residue in certain endogenous peptides could explain the lack of bioactivity, or the detection of multiple

forms. However, DAACPs are rarely considered during peptide characterization experiments and could be hidden in plain sight.

While there have not yet been any cell-cell signaling DAACPs discovered in mammals, L- to D-residue isomerase activity has been detected in rodent heart,³⁻⁴ lung, spinal cord, and hippocampus protein extracts.⁴ To date, the detection of L- to D- residue isomerase activity in rodent tissue homogenates is some of the best evidence for the existence of endogenous cell-cell DAACPs in mammals, but a native cell-cell signaling peptide substrate for this unidentified enzyme has not been identified. The synthetic peptide, neuropeptide FF, was previously found to be a substrate for an unidentified L- to D-residue isomerase enzyme in mouse and rat tissues that converts the L-Leu2 in position 2 of neuropeptide FF into a D-Leu2.⁴ However, an endogenous D-leu2 containing neuropeptide FF diastereomer has not been observed experimentally.

In this chapter, tissues from the central nervous system (CNS) and peripheral nervous system (PNS) of *Rattus norvegicus* were screened for DAACPs using the methods described in previous chapters. Aminopeptidase M (APM) screening was performed to identify DAACP candidates in *R. norvegicus* spinal cord extracts based on enzymatic stability, as was done in Chapter 3 and Chapter 6. In addition, *R. norvegicus* hypothalamus and pituitary extracts were analyzed by liquid chromatography (LC)-trapped ion mobility spectrometry (TIMS) mass spectrometry (MS) to identify possible isomeric peptides by screening chromatograms for peptides that elute at multiple retention times and have slightly different collisional cross sections. Neuropeptides and peptides derived from secretory proteins were prioritized. A list of several possible isomeric peptides was assembled based on enzymatic stability, chromatographic and ion mobility properties.

The possibility of acid extraction-induced Asp residue dehydration and isomerization is also discussed as a possible explanation for why certain peptides were observed eluting at multiple retention times by LC-MS.

7.3 METHODS

Tissue Extractions

For peptide extractions and LC-MS/MS analysis, all peptides all solvents were LC-MS grade and purchased from Thermofisher scientific or Sigma Aldrich, unless otherwise specified.

For peptide extractions, heat denatured⁵ rat hypothalamus, hippocampus, and spinal cord were mechanically homogenized in 250-500 μ L acidified methanol (10% glacial acetic acid in methanol) at 0 °C in 1.5 mL centrifuge tubes using Bio Plas disposable polypropylene pestles. The tissue homogenate was centrifuged at 20,000 $\times$ g for 15 min at 4 °C and the supernatant was collected into a new centrifuge tube. The pellets were then washed with 150 μ L water and centrifuged at 20,000 $\times$ g for 10 min at 4 °C. The pellet wash supernatants were combined with the acidified methanol extract supernatants and concentrated to about 50 μ L in a vacuum centrifuge at 25 °C for approximately 1 h.

For desalting, Pierce C-18 spin columns (Thermo-Fisher Scientific) were used. Spin column sorbent was activated with 50% methanol in water and equilibrated with 5% methanol in 0.5% Formic acid water. Spin columns were loaded with 50 μ L of extracts and centrifuged at 1500 $\times$ g for 1 min. The flow through was recovered and re-loaded onto column to ensure efficient binding. Peptide clean-up was carried out using a wash solution consisting of 95% water, 5% methanol and 0.5% formic acid. Washing was performed by adding 150 μ L of wash solution to column and centrifuging for 1 min at 1500 $\times$ g and repeated twice. Finally, peptides were eluted

from the columns using 40 μ L of 70% methanol in water. Eluent was concentrated and stored at -20 °C. Desalted peptide extracts were reconstituted in 50-100 μ L of 5% CH₃CN in water with 0.1% formic acid (FA) and protein content was quantified using a Pierce BCA protein assay kit (Thermo-Fisher Scientific).

Aminopeptidase M Enzyme Preparation

Aminopeptidase M (APM) from porcine kidney (APM, CAS#: 9054-63-1) was purchased from Calbiochem (Darmstadt, Germany). The APM suspension (170 U/mL, 2.7 mg/mL) was stored under manufacturer recommended conditions until use. The APM suspension was gently swirled and transferred directly to the reaction vials without any buffer exchange. APM digestions on Spinal cord tissue was performed using the same procedure instrumentation and described in Chapter 3.

APM digestions Spinal cord

APM digestions of spinal cord extracts were performed as described in Chapter 3 Briefly, extracts were desalted, and the protein content was quantified by BCA protein assay kit. 25 μ g/mL of peptide and 0.3 U/mL APM. APM digestions were carried out for 20 h. Followed by LC-MS/MS analysis. LC-MS/MS analysis of Spinal cord APM screening experiments, extract was split into to vials a “no APM control” which was incubated without APM for 20 h, and an “APM treated” sample which was incubated with APM for 20 h. A small amount of fresh extract was set aside for LC-MS/MS analysis of fresh tissue extract that was not subjected to APM screening conditions.

LC-TIMS-MS/MS analysis

For Nano LC-TIMS MS measurements on fresh *R. norvegicus* hypothalamus, hippocampus, and pituitary extracts, LC was carried out on a nanoElute system (Bruker Daltonics) using a nanoElute FIFTEEN C18 ReproSil AQ reversed-phase analytical column, 75 μm i.d. $\times$ 15 cm, 1.9 μm particle size, 120 Å pore size (P/N 1842621, Bruker) coupled to a timsTOF Pro mass spectrometer with a CaptiveSpray ionization source (Bruker). Separation was carried out at 40 °C at a flowrate of 300 nL/min. Mobile phase compositions were as follows: H₂O, 0.1% FA (solvent A) and CH₃CN, 0.1% FA (solvent B). The LC gradient was a 90 min gradient 5-10 %B in 5 min, then 10–40% B in 85 min. For TIMS MS measurements, the TIMS device was calibrated using an Agilent low concentration ESI tuning mix (Agilent Technologies, P/N G1969-85010) over a $1/k_0$ range of 0.5-2.0 V·s·cm⁻². The TIMS MS was operated using parallel accumulation serial fragmentation (PASEF) method. LC-MS/MS was carried out using the same procedure described in Chapter 3.

7.4. RESULTS AND DISCUSSION

Definition of “DAACP Candidates”

In this work, the DAACP discovery methods that were applied in Chapters 3 and Chapter 6 were employed to screen for potential DAACPs in spinal cord tissue from *R. norvegicus*. Spinal cord tissue was specifically selected because L- to D-residue isomerase activity was previously reported in this tissue.⁴ In addition, LC-TIMS-MS was also employed to search for potentially isomeric peptides in hypothalamus and pituitary.

Peptides were labeled as “DAACP candidates” only if they met both of the following criteria: 1) the peptide must be detected eluting at two (or more) retention times by LC-MS in the

control extractⁱⁱⁱ 2) for peptides that met the first condition, at least one of the peaks must be detected in the LC-MS chromatogram of the APM-treated extract (must be resistant to APM), and at least one peak must be susceptible to APM digestion (must be degraded by APM). While analyzing screening results for DAACP candidates, the reason the chromatogram of an APM resistant peptide must be accompanied by a second, APM susceptible isomeric peptide (that eluted at a different retention time in the control), is because this indicates that an all-L form of the peptide was also present but was fully degraded by APM. Therefore, the APM susceptible peak serves as an endogenous control, which allows for the APM resistant isomer to be reasonably attributed to a D-residue containing form of the peptide.

Finally, as in Chapters 3 and 6, peptides with N-terminal acetylation, or N-terminal Glu, Asp, Pyro-Glu (pGlu), Pro, or Xaa-Pro- were expected to resist APM digestion independent of a D-residue, and thus, peptides with these N-terminal residues/modifications were not categorized as potential DAACP candidates based on resistance to APM.

Spinal Cord APM-Screening Results

Not a single peptide that was detected in the APM screening experiments on spinal cord met both the necessary conditions described in the previous section. Therefore, no DAACP candidates could be confidently identified. This result is similar to the results from the screening of *R. norvegicus* pancreatic islets reported in Chapter 6. However, this result does not indicate that *R. norvegicus* spinal cord does not have DAACPs, but it does suggest that the criteria used to assign DAACP candidates in Chapter 3 may not be as effective at identifying DAACPs in *R.*

ⁱⁱⁱ Control extract refers to the extract that was incubated in the APM digestion buffer but was not treated with APM.

norvegicus or other mammals. There are both analytical and biological reasons for why this may be the case, and these are described in a later section.

Although no peptides were observed in spinal cord extracts that met both the necessary conditions to be categorized as a DAACP candidate, several of the peptides met at least one of the conditions. 16% of the peptides that were detected in the APM treated sample were expected to be fully degraded by APM (Figure 7.1). However, this 16% of APM resistant peptides did not meet the second condition, that is, they were only observed with a single peak at retention time by LC-MS. The sequences that resisted APM (but were expected to degrade) are shown in Table 1. For these sequences, because they did not have a second, APM susceptible peak, it was impossible to confidently attribute their stability against APM hydrolysis to the presence of a D-residue. This is a critical interpretation to emphasize—just because a peptide is resistant to APM, this does not prove it has a D-residue at the second position from the N-terminus. The latter interpretation is especially important to keep in mind for the sequences in Table 7.1 because they did not have a second, APM susceptible peak eluting at a different retention time in the control.

The presence of a D-residue at position 2 is not the only reason why a peptide (that would be expected to degrade) may resist APM hydrolysis. Some peptides may resist digestion because their solution phase conformation prevents the N-terminus from accessing the APM active site. Other peptides may appear to resist APM degradation because the digestion did not go to completion. It is difficult to predict the enzyme kinetics when performing APM digestions on a complex mixture. This was discussed in more detail in Chapter 6.

The other 84% of peptides that resisted the APM digestion were expected to be resistant because they had N-terminal Asp, Glu, pGlu, Pro, Xaa-Pro- (where Xaa is any amino acid) residues, or N-terminal acetylation (Figure 7.1). The individual chromatograms of APM resistant

peptides that were expected to resist digestion, were evaluated for sequences which eluted at two or more retention times. Two peptides were found that eluted at two retention times: PENK[198-209] from the proenkephalin A precursor (PENK) and pGlu-Scg1[597-611] from the secretogranin 1 precursor (Scg1).

Both the PENK[198-209] and pGlu-Scg1[597-611] peptides were observed at two retention times, but every peak was resistant to APM. Because these peptides were detected eluting at two retention times by LC-MS, this indicates they may have an isomeric form. But just because the same peptide sequence is observed eluting at two retention times by LC-MS, this is not proof that one of the peaks corresponds to the D-amino acid containing isomer. This is because the split peaks could be a chromatographic artifact (i.e., column overload), mass spectrometric artifact (i.e., in source/post source chemical reactions) or separation of stable conformers (i.e., *cis-trans* proline conformers) or a different structural isomer. As mentioned previously, one of the two peaks needs to be susceptible to APM hydrolysis for the presence of a D-residue containing form to be inferred. Both peptides resisted APM because of their N-terminal sequences. PENK[198-209] has a proline at position 2 from the N-terminus, and pGlu-Scg1[597-611] has a pyro-Glu residue at the N-terminus. Because of this, APM screening provided no additional information about the stereochemistry of these peptides. Therefore it was impossible to identify which (if any) of the two (PENK[198-209] and pGlu-Scg1[597-611]) isomers was a DAACP based on APM resistance. A plausible explanation for the observation of isomeric forms of PENK[198-209] and pGlu-Scg1[597-611]) is discussed in later sections of this chapter.

Extended Discussion of Spinal Cord Peptides Detected at Two Retention Times by LC-MS

Of the “DAACP candidates” reported in Table 7.1, only two were assigned based on multiple peak elution: PENK[198-209] and pGlu-Scg1[597-611]. Because these peptides were detected eluting at two retention times by LC-MS, this indicates a second isomer of these peptides may be present. Both peptides resisted APM because of their N-terminal sequences. PENK[198-209] has a proline at position 2 from the N-terminus, and pGlu-Scg1[597-611] has a pGlu residue at the N-terminus. Therefore, APM screening provided no additional information about the stereochemistry of these peptides. Therefore it was impossible to identify which (if any) of the two forms of PENK[198-209] and pGlu-Scg1[597-611]) was a DAACP based on APM resistance.

The unmodified form of Scg1[597-611] has a Gln residue at the N-terminus, it was only observed at one retention time and was degraded by APM after 20 hours (Figure 7.2A). It is unclear why the pGlu-Scg1[597-611] eluted at two retention times while Scg1[597-611] did not.

The unmodified Scg1[597-611] eluted 10 min earlier than both of the pGlu-Scg1[597-611] peaks by LC-MS/MS (Figure 7.2A-B), indicating the second peak is likely not the result of in-source Gln cyclization.⁶ pGlu-Scg1[597-611] was also observed with split peaks in pituitary and hypothalamus. The possibility that the two eluting peaks of this peptide was related to a labile PTM such as Tyr sulfation was considered. Tyr sulfation can produce artifactual split peak elution because the SO₃ group is acid labile, and can undergo an in-source neutral loss during electrospray. Previous studies have reported that some peptides derived from Scg1 have sulfo-Tyr residues, but sulfation of Tyr598 or Tyr611 of Scg1[597-611] has not been previously reported.⁷ Here, an *m/z* corresponding to sulfated pGlu-Scg1[597-611] peptide was not detected in spinal cord extracts. Therefore, the idea that Tyr sulfation may be responsible for the second chromatographic peak was excluded.

PENK[198-209] was also detected at two retention times in spinal cord extracts and both peaks resisted APM digestion. PENK[198-209] was expected to resist APM because it has a Pro residue at position two from the N-terminus. However, because PENK[198-209] was detected at two retention times, this suggests two isomeric forms of this peptide are present. The lower intensity peak had a relative abundance of just 0.8% but the retention time was nearly 4 min earlier than the higher intensity peak (Figure 7.3). Preliminary results suggest possible Asp isomerization could be occurring to PENK[198-209] and this is discussed further in a later section.

Extended Discussion of APM Resistant Peptides from Spinal Cord

Because so few peptides were detected with split eluting peaks, DAACPs candidates were assigned based on APM resistance regardless of whether split peak elution was observed. Table 7.1 lists the “DAACP candidates” that resisted APM but did not exhibit multiple peak elution but were expected to degrade. However, just because a peptide resists APM digestion, this does not unambiguously indicate it is a DAACP. While APM degrades a broad range of peptides, there is always the possibility that certain peptide sequences are just not easily hydrolyzed by APM.

It is still worth highlighting the peptides that resisted APM digestion because a D-residue is still one of the possible explanations for resistance. The most notable APM resistant peptides from this work included peptides encoded by PENK, proSAAS and secretogranin 2 (Scg2) (Table 7.1.). EICs of PENK-, and ProSAAS- derived peptides incubated with and without APM for 20 h are shown in Figure 7.4.

In Chapter 2, it was discussed that peptide sequence homology can sometimes be used to predict DAACPs across species. Chapter 2 also discussed multiple peptides derived from a single precursor are often DAACPs. Because PENK[114-133] and PENK[229-259] were resistant to

APM all the peptides from PENK were analyzed in further detail. Several peptides from PENK have short sequences (< 8 residues) that have sequences similar to **FdMR**GF-NH₂ from *A. californica*. Database searches on LC-MS/MS results may not pick up peptides with shorter sequences shorter peptides. EICs of the PENK derived opioid peptides, YGGFM, YGGFL, YGGFMRGL, and YGGFMRF were manually inspected for and APM resistance, and elution at multiple retention times by LC-MS. However, these four peptides were completely digested by APM and were not detected at multiple retention times in spinal cord extracts by LC-MS. This is good evidence that D-amino acid residue containing forms of these peptides were not detected in spinal cord extracts.

The most probably explanations for why PENK[114-133] and PENK[229-259] were resistant to APM is because their sequences are not good substrates for APM, or the APM digestion did not go to completion. However, it is interesting that these peptides derived from PENK, because, as discussed in the previous section, PENK[198-209] appears to be present in extracts as two isomeric forms.

In *R. norvegicus*, YGGFM, YGGFL, YGGFMRF and other opioid peptides with similar N-terminal sequences, are known to be native substrates for endogenous Rat aminopeptidases.⁸ Although YGGFM was observed fully degraded by APM here, the extended form of the peptide, PENK[107-133] (YGGFMKKMDELYPVEPEEEANGGEILA) was resistant to APM despite sharing the YGGFM-motif. This is consistent with previous work which reported β -endorphin(1-5) was completely degraded by leucine aminopeptidase,^{iv} while β -endorphin(1-17) was resistant, despite both peptides sharing the same N-terminal YGGFM-motif.⁹ The k_{cat}/K_M of APM reportedly decreases with increases peptide length.¹⁰ It could be that lower activity of APM

^{iv}Aminopeptidase M and leucine aminopeptidase refer to different enzymes, but both belong to the same family of aminopeptidases and have similar specificity.

towards longer peptides explains why peptides in Table 7.1 were resistant to APM. However, the presence of a D-residue in the peptides reported in Table 7.1 is still one possible explanation for APM resistance and follow-up analysis will be an important next step. The best initial follow-up experiment will be to compare the retention time of the all-L synthetic standard to the native peptide.

Aside from PENK derived peptides, several sequences were detected only partially degraded by APM in spinal cord (Figure 7.4). If it is assumed that APM completely degraded away the all-L peptides from the extract, this could be an indication that the all-L and DAACP forms of the peptides coeluted. However, it seems unlikely that D-residue containing forms every single peptide in Figure 7.4 and Table 7.1 would coincidentally coelute with the all-L peptides. As discussed in Chapter 6, encountering partial or incomplete degradation of peptides in a complex mixture of peptides is virtually unavoidable. Another important thing to keep in mind when screening for DAACPs is the presence of endogenous APM inhibitors in tissue extracts. Spinal cord extracts may competitive aminopeptidase inhibitors that reduce the activity of APM towards other native peptides in the extract. For example, certain endogenous peptides that are likely present in spinal cord (or other regions of the CNS and PNS) are known to inhibit APM competitively including substance P, neurotensin, dynorphin(1-17) and angiotensin III.¹¹⁻¹² The inclusion of the external standard may also contribute to competitive inhibition. Competitive inhibition may result in partial degradation of peptides. The potential for competitive inhibition is reduced when digests are performed over longer incubation times (on the order of several hours) but again, the kinetics of APM hydrolysis towards a specific sequence in a complex mixture of hundreds of peptides is difficult to predict or control. Therefore, it is also very possible (and likely) that the peptides reported in Table 7.1 and Figure 7.4 consist of all-L peptides that did not get fully

hydrolyzed by APM. These ambiguities emphasize why relying on APM resistance alone to prove a peptide has a residue is irresponsible and highlights the importance of performing follow-up structural characterization of APM resistant peptides.

A final thing to consider is that many peptide sequences found in mammals may have evolved to naturally resist mammalian APM (or APM has evolved in a way so as not to degrade certain sequences), in future work, it may be beneficial to screen for DAACPs using an aminopeptidase with similar stereospecificity, but from a different (perhaps non-mammalian) species.

The Limitations of APM Screening on Mammalian Tissues

In general, LC-MS chromatograms showed peptides in spinal cord extracts lacked the frequency of multiple eluting peaks by LC-MS/MS that was seen in *A. californica* CNS in Chapter 3.¹³ Searching LC-MS/MS chromatograms for peptides exhibiting multiple peak elution is a useful factor for identifying isomeric peptides, but it is most useful in DAACP discovery when coupled with APM digestion.¹ The DAACP discovery protocol employed in Chapter 3¹ relied on the coexistence of the endogenous all-L residue peptide and the DAACP in the extract. The presence of the native all-L form served as internal control that gets degraded by APM, allowing for the APM resistant peak be assigned as a likely DAACP. However, the detection of APM resistant peptides that did not have multiple peak elution by LC (Table 7.1) makes it difficult to correlate their stereochemistry and their resistance to APM digestion. In addition, in this chapter, for the peptides that were observed eluting at more than one retention time, the N-terminal residues make the peptides particularly resistant to APM, and APM digestions could not provide any additional useful information about the structure of the two eluting isomers.

One plausible analytical explanation for the absence of peptides with multiple eluting peaks may be that the measurement conditions were ineffective at separating peptide diastereomers because the DAACP forms coincidentally coelute with the all-L forms.

Any biological explanations for why none of the peptides detected in spinal cord exhibited the properties that would be expected of DAACPs (based on the work done in Chapter 3) can only be left up to speculation. However, these biological explanations warrant a brief discussion as it relates to experimental design considerations for future DAACP discovery efforts in mammals.

The DAACPs may not be present in significant quantities in spinal cord and may be localized to a different tissue. This is not an unreasonable hypothesis as it has been demonstrated previously that tissue specific localization of DAACPs is possible.¹³⁻¹⁴ Another plausible explanation is the hypothetical L- to D-residue isomerase colocalizes with a only small percentage of cells that express the isomerase substates, leading to a relative abundance of DAACPs that is below the limit of detection. Finally, assuming the DAACP and all-L forms are separable by LC, the presence of DAACPs may not be immediately evident based on multiple peak elution because endogenous proteolytic enzymes in *R. norvegicus* had degraded the all-L forms of the peptides more efficiently after biosynthesis, or even during extraction protocols, leaving only the more stable DAACPs remaining in the extract. It has been shown that the number of peptides detected in extracts can be heavily dependent on the tissue stabilization methods used to inactivate endogenous enzymes.⁵ While this has not been demonstrated in studies on DAACPs, it is plausible that unintended exposure to endogenous enzymes at any step in the extraction could result in the degradation of all-L peptides, but not the DAACP. These alternative explanations are important considerations so as not to dismiss the absence of peptides that meet the DAACP candidate criteria as evidence that mammals do not have DAACPs. Until the rodent L- to D- residue isomerase is

fully characterized, and more complete DAACP screening studies have been performed, biological explanations for the prevalence of DAACPs (or lack thereof), is a matter of conjecture.

LC-MS/MS Analysis of Synthetic pGlu-Scg1[597-611] Diastereomers

To further elucidate the structure of the native pGlu-Scg1[597-611] peptide isomers that were detected in spinal cord, ^{13}C -Gly labeled pGlu-D2-Scg1*[597-611] and pGlu-L-Scg1*[597-611] were synthesized (The “*” in Scg1*[597-611] indicates the synthetic peptide(s) has a ^{13}C -labeled Gly residue).

pGlu-L-Scg1*[597-611] was spiked into spinal cord extracts and was found to perfectly coelute with the earlier eluting native pGlu-Scg1[597-611] peak in Figure 7.5. Because the synthetic ^{13}C -L-Scg1[597-611] didn't have split peak elution LC-MS when spiked into the spinal cord extracts, this is a good indication the second eluting native species is unique, and not a result of dimerization, non-covalent interactions, or interconversion of conformers during LC.

pGlu-D2-Scg1*[597-611] was not spiked into spinal cord extracts, but its retention time was compared to the retention time of synthetic pGlu-L-Scg1*[597-611] (Figure 7.5B). The synthetic pGlu-D2-Scg1*[597-611] was partially resolved from pGlu-L-Scg1*[597-611], but clearly eluted slightly earlier. This indicated the second eluting peak present in the spinal cord extract likely does not correspond to a native diastereomer of pGlu-Scg1[597-611] with a D-residue in position 2.

From the above experiments it was determined that the earlier eluting native pGlu-Scg1[597-611] is comprised of only L-residues and that the second eluting pGlu-Scg1[597-611] peak is likely an isomer, but which kind is unknown. However, the isomer is not a native diastereomer of pGlu-Scg1[597-611] with a D-residue in position 2. The site of a modified

stereocenter or isomerized residue in the second eluting native pGlu-Scg1[597-611] was not fully elucidated but the most plausible explanation is described in the following section.

Possible Succinimide Formation at Asp600 of pGlu-Scg1[597-611]

The Asp600 in the sequence of pGlu-Scg1[598-611] is followed by a Gly residue. Peptides containing -Asp-Gly- in their sequences have been shown to be more susceptible to spontaneous Asp cyclization and isomerization.¹⁵ Because the Asp600 is followed by a Gly, it is possible that it is prone to spontaneous isomerization. Therefore, the later eluting native pGlu-Scg1[597-611] described in the previous section may correspond to an L-IsoAsp600, D-Asp600, D-IsoAsp600 residue containing isomer.

Spontaneous isomerization of Asp residues proceeds via Asp cyclization and side chain dehydration to form a succinimide intermediate. Although the succinimide is considered intermediary, the amino succinyl residue can remain stable in acidic solutions (pH < 7), and can also remain stable enough to be analyzed by LC-MS/MS.¹⁶ If a succinimide-containing form of the native pGlu-Scg1[597-611] could be detected in spinal cord extracts it would support the hypothesis that the later eluting pGlu-Scg1[597-611] peak is has an L/D-IsoAsp, or D-Asp.

Spinal cord extracts were tested for the presence of a succinimide containing intermediate formed by dehydration of Asp600 in pGlu-Scg1[597-611]. To accomplish this, fresh spinal cord extracts (extracted using the conditions described in the methods section) were analyzed by LC-MS/MS. The pH of the extract solution was not increased above pH 3.0. EICs corresponding to the m/z of the dehydrated pGlu-Scg1[597-611] were then plotted. A peak corresponding to the dehydrated m/z of pGlu-Scg1[597-611] was detected eluting slightly before the pGlu-Scg1[597-611] peptide in the spinal cord extracts and the product ion spectrum. The dehydrated pGlu-

Scg1[597-611] was not detected in spinal cord extracts incubated at pH 7.4 for 20 h. succinimides are unstable at $\text{pH} > 7$,¹⁷⁻¹⁸ so the absence of the dehydrated peptide from incubated extracts is addition evidence that dehydrated pGlu-Scg1[597-611] ion detected in the fresh extracts does correspond to a succinimide. It should be noted that the detection of the dehydrated pGlu-Scg1[597-611] does not unambiguously prove the presence of a succinimide, or that spontaneous Asp isomerization is occurring. However, based on the observations discussed above, Asp isomerization is the most-likely source of the isomeric pGlu-Scg1[597-611] peak detected.

Possible Succinimide Formation at Asp203 of PENK[198-209]

PENK[198-209] has an Asp residue, and one of the peaks be an L/D-IsoAsp or D-Asp isomer. In fresh spinal cord extracts an EIC corresponding to m/z of the dehydrated-PENK[198-209] showed a chromatographic peak eluting in between the two PENK[198-209] peaks. MS/MS indicated that the dehydration was occurring at an internal residue. Because the dehydrated PENK[198-209] eluted at a different retention time than the two non-dehydrated PENK[198-209] peaks, this indicated the dehydrated PENK[198-209] was not a result in-source/post-source dehydration of unmodified peptide.¹⁶

During APM screening on spinal cord extracts, this dehydrated PENK[198-209] was not observed in the control extract incubated for 20 h without APM, or in the APM treated sample. As mentioned in the precious section, succinimides are more stable in acidic solution than at neutral or basic pH,¹⁷⁻¹⁸ which may explain why dehydrated-PENK[198-209] was not observed in the spinal cord extract incubated at neutral pH for 20 h at 37 °C, but was observed in the fresh spinal cord extract.

The intensity of the dehydrated-PENK[198-209] peak from spinal cord extracts was too low intensity to obtain a quality product ion spectrum on the qTOF spectrometer, but the dehydrated-PENK[198-209] peak was also detected with the timsTOF pro at a similar relative retention time in fresh hypothalamus and pituitary extracts with good product ion spectra (Figure 7.3). Based on the MS/MS result obtained from the timsTOF spectrometer, the dehydrated-PENK[198-209] likely corresponds to a water loss at an internal Asp204 residue (Figure 7.3 C-D). This observation is consistent with the presence of an amino succinyl residue formed by cyclization of Asp.¹⁹⁻²⁰ However, because the b_6-H_2O and b_7-H_2O ions were not detected in the MS/MS spectrum of the dehydrated PENK[198-209] and the dehydration could also be occurring at the Glu residue.

Trapped Ion Mobility Spectrometry Analysis of Galanin in Hypothalamus Extracts.

LC-MS/MS results deposited in the proteome exchange database (www.proteomexchange.org) from published studies on rodent central nervous system

Hypothalamus extracts were analyzed by LC-TIMS-MS/MS and chromatograms were evaluated for peptides that elute at multiple retention times. The Mobility profiles of peptides detected at more than one retention time were then compared to look for differences in collisional cross sections.

Among the peptides detected at more than one retention time was Galanin (GWTLSAGYLLGPHAIDNHRFSFDKHGLT-NH₂) and the chromatograms are shown in Figure 7.6). The mobility profiles of the galanin peptide at both retention times were compared. Figure 7.7. shows the isotope ratios, and the mobility profiles of Galanin in the 3+, 4+, 5+, 6+ charge states. In each charge state, the mobility profile of the first eluting species was slightly

different than that of the second eluting species. This is strong evidence that the two species are structurally unique.

The presence of a D-residue containing isomer is just one possible explanation for why Galanin was observed at two retention times by LC-TIMS-MS. Alternative hypotheses for the isomeric peaks include *cis-trans* proline isomers, or Asp isomers from spontaneous isomerization. Iso-Asp containing forms of Galanin resulting from Asn deamidation have been previously reported.²¹ Asn-deamidated forms of Galanin were not detected in the work presented here, and the MS1 profiles of the two eluting peaks exclude Asn deamidation as the explanation.

Like pGlu-Scg1[597-611] and PENK[198-209], the *m/z* corresponding to dehydrated Galanin was detected in hypothalamus by LC-TIMS-MS, and was observed eluting an earlier retention time than both the unmodified galanin peaks. This indicated the dehydration product was not the result of in-source dehydration. MS/MS spectra of the dehydrated galanin peptide indicated the dehydration occurred at an internal residue, but the specific site of dehydration could not be confirmed because of low sequence coverage. Galanin does have two Asp residues and it is possible the dehydrated ion corresponds to a succinimide. The most-likely explanation for the second eluting Galanin peak in hypothalamus extracts by LC-TIMS-MS is the presence of an Asp-isomer containing form. However, further structural characterization is necessary before drawing any conclusions.

Implications of Dehydrated Peptides

In the sections above, PENK[198-209] and pGlu-Scg1[597-611], and galanin were observed eluting at two retention times by LC-TIMS-MS. In addition, dehydrated forms PENK[198-209] and pGlu-Scg1[597-611] and galanin were also detected. The dehydrated peptides may correspond to the succinimide-containing PENK[198-209] and pGlu-Scg1[597-611] and galanin, but this still needs to be confirmed in future studies. This could indicate asp isomerization is responsible for the multiple peak elution of these peptides.

Methanol acidified with acetic acid was used for peptide extraction from PNS and CNS tissues. Computational modeling has suggested that acetic acid can catalyze Asp residue cyclization,²² but this hasn't been supported by experimental evidence to back this up yet. However, Asp residue cyclization and isomerization can also naturally occur endogenously in long lived proteins and peptides. It remains ambiguous whether dehydration of PENK[198-209], pGlu-Scg1[597-611] and galanin occurred endogenously or if the dehydration was just an artifact of peptide extraction and sample handling.

One counterargument against the hypothesis that PENK[198-209], pGlu-Scg1[597-611] and galanin undergo spontaneous Asp-isomerization is that LC-MS chromatograms only separated two isomeric peaks of these peptides—whereas four isomeric peaks would be expected if Asp isomerization had occurred. However, favored hydrolysis of the succinimide into just one isomeric form has been observed and was explained by steric effects.²³ So presence of only two isomeric peaks is not evidence against spontaneous Asp isomerization.

The presence of an iso-Asp residue in these peptides may be elucidated in future LC-MS/MS experiments by using electron transfer dissociation (ETD),²⁴⁻²⁵ or electron capture

dissociation (ECD),²⁶ because these fragmentation methods can more easily identify isoAsp residues compared to collision induced dissociation (CID).²⁵

Performing the extraction in ¹⁸O-labeled water can be done to demonstrate Asp isomerization is occurring during the extraction. Similarly, the presence of the succinimide containing form of the peptide after extraction could be determined by incubating the extract at neutral pH in ¹⁸O-labeled water.²⁷

Targeted Analysis of Neuropeptide FF

Synthetic neuropeptide FF (FLFQPQRF-NH₂) was previously found to be a substrate of an unidentified L- to D-residue isomerase from rodent tissues.⁴ Spinal cord was tested to determine if a D-residue containing form of neuropeptide FF exists because previous reports have identified neuropeptide FF in rat spinal cord extracts.²⁸ Although neuropeptide FF was not detected in spinal cord extracts here. The extended form of neuropeptide FF, neuropeptide SF (NPAFLFQPQRF-NH₂), was detected, and this peptide has also been detected in previous studies on rat in spinal cord and CNS tissues.^{5, 29} Neuropeptide SF was not observed eluting at two retention times by LC-MS, it was fully degraded by APM, and no stable truncated forms were detected in the APM treated sample. Taken together, these observations indicate a D-residue containing form of NPFF could not be identified by the DAACP screening. Neuropeptide SF was also detected in rat hypothalamus, and pituitary extracts by LC-MS/MS, but was only detected at one retention time.

7.5 CONCLUSION

APM screening was performed on *R. norvegicus* spinal cord extracts and several possible DAACPs were identified based on resistance to APM digestions and or presence of multiple

isomeric peaks by chromatography. In addition, hypothalamus and pituitary extracts were evaluated by LC-TIMS-MS/MS and peptides and the mobility profile of peptides with isomeric peaks were compared to evaluate structural differences of the isomers. Overall, several peptides were stable against enzymatic digestion or displayed chromatographic properties consistent with peptide diastereomers. The findings here represent promising preliminary results suggesting the presence of DAACPs or unidentified isomeric peptides in the *R. norvegicus* nervous system.

7.6 FIGURES AND DATA

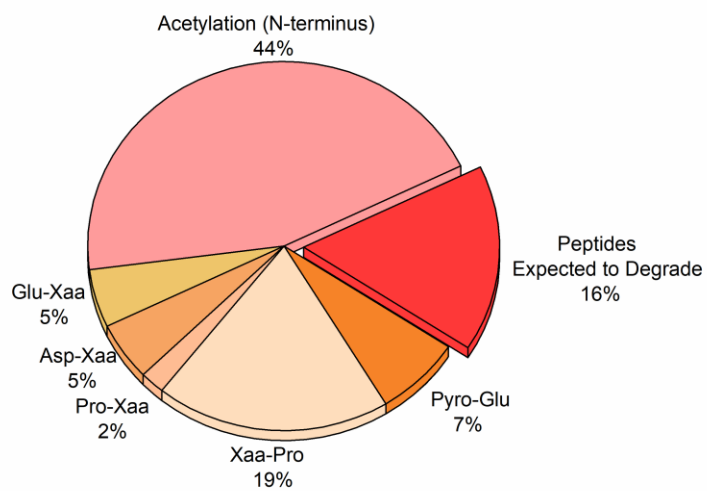


Figure 7.1. Breakdown of peptides resistant to APM. Peptides are categorized based on N-terminal amino acid residues or modifications. Peptides that were expected to degrade represented 16% of the peptides observed after the digestion.

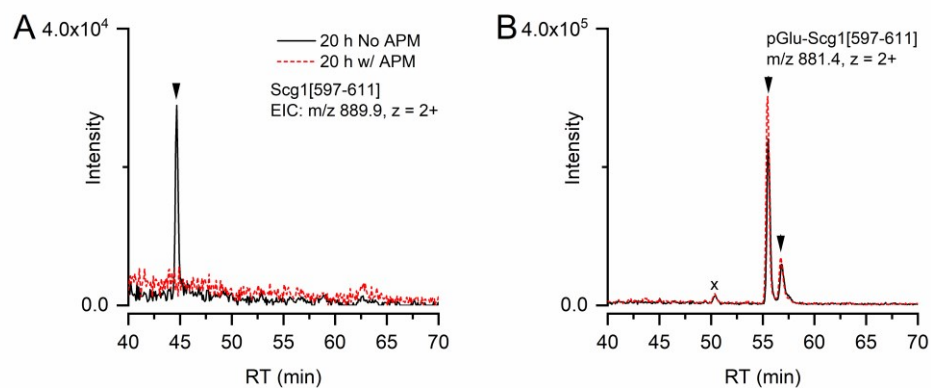


Figure 7.2. APM digestion results for (A) Scg1[597-611] and (B) pGlu-Scg1[597-611] in *R. norvegicus* spinal cord extracts. Plots show LC-MS EICs of the $[m+2H]^{2+}$ ions intensity after incubation with no APM for 20 h (black trace) and after incubation with APM for 20 h (dashed red trace). The “x” over the peak at 50 min in panel B indicates an unrelated ion.

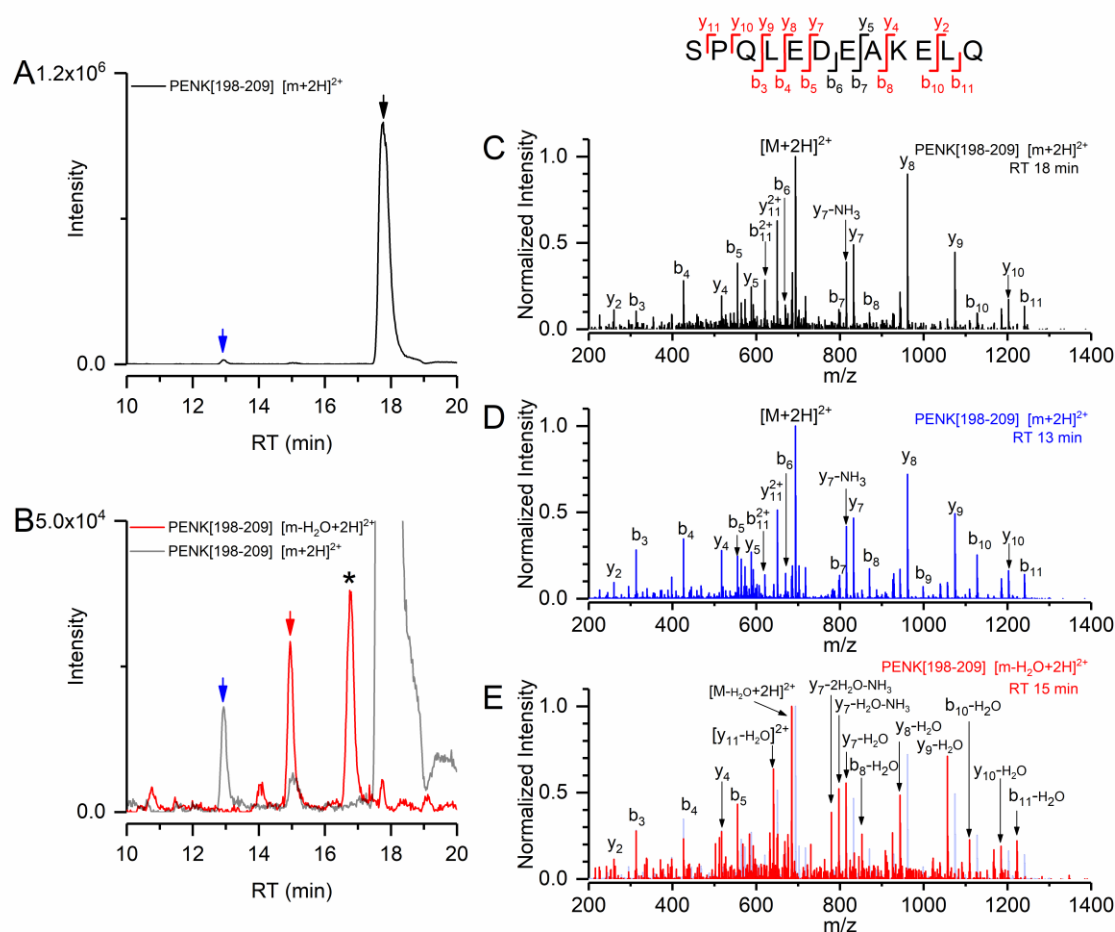


Figure 7.3. Evidence for isomeric PENK[198-209] peptides in rat hypothalamus. (A) EIC of PENK[198-209] ($[m+2H]^{2+} = 693.8$). The black and blue arrows indicate the two peaks corresponding to PENK[198-209] at 18 min and 13 min, respectively. (B) EIC of dehydrated PENK[198-209] ($[m-H_2O+2H]^{2+} = 684.8$) in red. The zoomed in view of the unmodified PENK[198-209] from panel A is shown in black. The red arrow indicates the peak corresponding to dehydrated PENK[198-209]. The blue arrow shows the zoomed in view of PENK[198-209] from panel A. The * over the peak at 17 min in panel B indicates an unrelated ion with an isobaric monoisotopic m/z to the dehydrated PENK. (C) Product ion spectrum obtained from the peak eluting at 18 min in panel A. (D) Product ion spectrum of the peak eluting at 13 min in panels A & B. (E) Product ion spectrum of the dehydrated PENK[198-209] (red trace). The MS2 of unmodified PENK[198-209] is shown in blue for comparison. Fragment ion assignments are labeled in each product ion spectrum. Above Panel C shows the fragment ion sequence coverage with fragment ions common to both the dehydrated and unmodified PENK[198-209] in red, and the fragment ions only detected in the unmodified [PENK 198-209] shown in black.

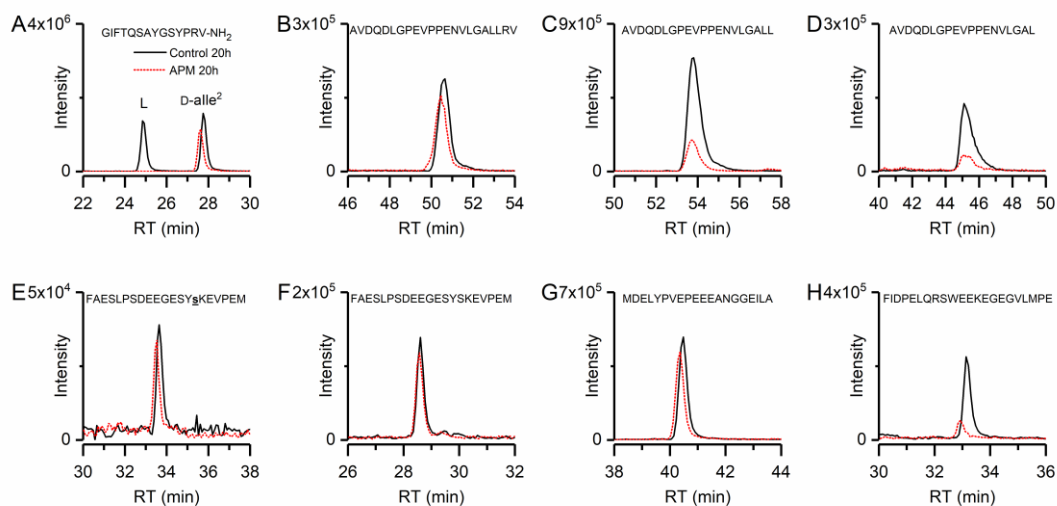


Figure 7.4. APM screening results for select peptides in *R. norvegicus* spinal cord extracts. Plots show LC-MS EICs from the control extract, incubated without APM for 20 h (Solid black trace) and in extract treated with APM for 20 h (dotted red trace). (A) EIC of synthetic *A. californica* peptide standards, ^{13}C -L-Plrn3 and ^{13}C -D2-Plrn3 ($[m+2H]^{2+} = 773.4$) which was added to the solution as an all-L peptide and DAACP control prior to incubation. (B-D) EICs of peptides derived from ProSAAS. (B) ProSAAS[221-242] ($[m+2H]^{2+} = 1151.2$); (C) ProSAAS[221-240] ($[m+2H]^{2+} = 1023.5$); (D) ProSAAS[221-239] ($[m+2H]^{2+} = 967.0$). (E-G) EICs of APM resistant peptides derived from Proenkephalin A; (E) Phospho-Ser253-PENK[239-259] ($[m+2H]^{2+} = 1220.5$); (F) PENK[239-259] ($[m+2H]^{2+} = 1180.5$); (G) PENK[114-133] ($[m+2H]^{2+} = 1103.0$); (H) TRH[188-199] ($[m+3H]^{3+} = 873.4$) from prothyrotropin releasing hormone (TRH) precursor.

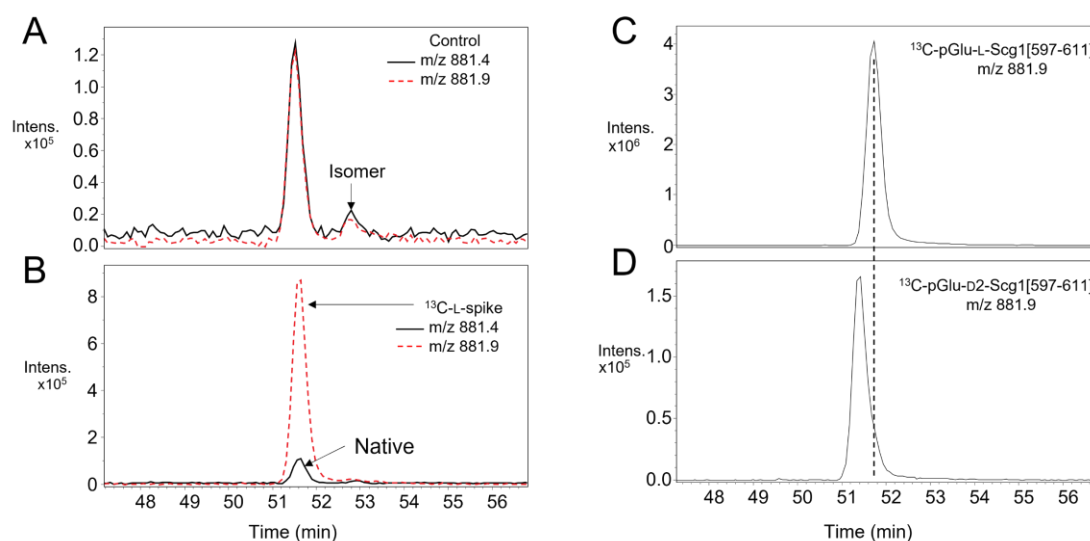


Figure 7.5. LC-MS analysis of synthetic and native pGlu-Scg1[597-611]. (A-B) Spiking of synthetic pGlu-L-Scg1*[597-611] into spinal cord extracts. (A) EIC of native pGlu-Scg1[597-611] $[m+2H]^{2+} = 881.4$ (solid black trace) and native $[m+1+2H]^{2+} = 881.9$ (dashed red trace) in spinal cord extract without the synthetic standard. (B) EIC of native pGlu-Scg1[597-611], $[m+2H]^{2+} = 881.4$ (solid black trace) spiked with synthetic pGlu-L-Scg1*[597-611], $[m^*+2H]^{2+} = 881.9$ (dashed red trace). (C-D) LC-MS retention time comparison of synthetic pGlu-Scg1*[597-611]. (C) EIC of synthetic pGlu-L-Scg1*[597-611] ($[m^*+2H]^{2+} = 881.9$); (D) EIC of synthetic pGlu-D2-Scg1*[597-611], $[m^*+2H]^{2+} = 881.9$. The black dashed line in C-D is shown to highlight the retention time difference between the synthetic diastereomers.

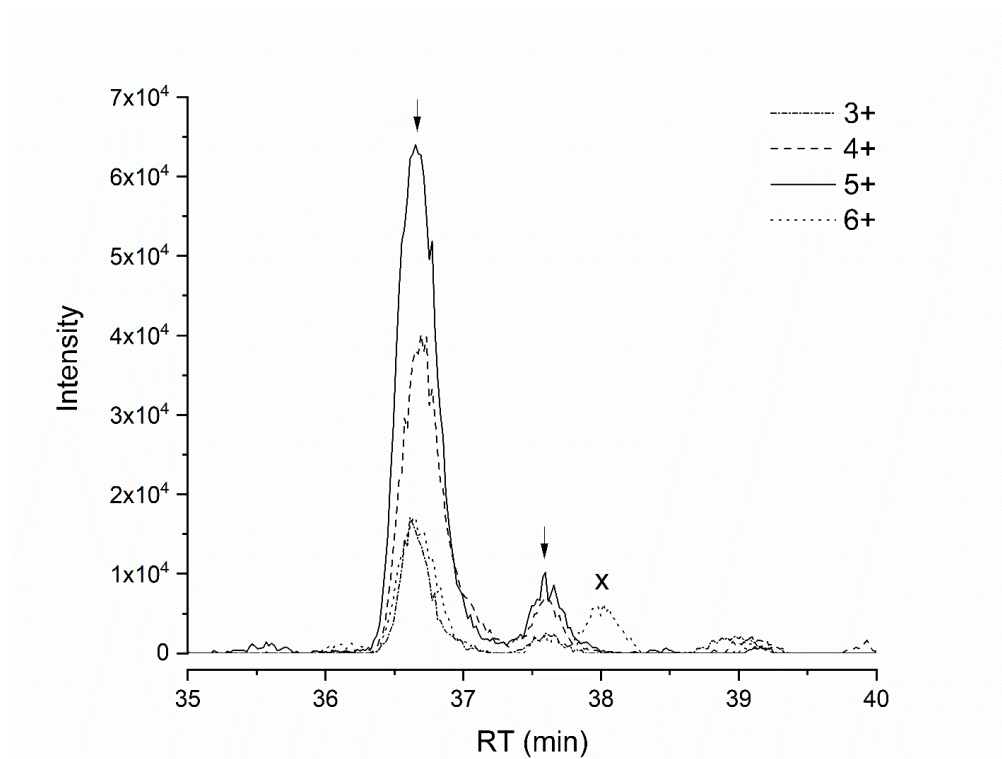


Figure 7.6. LC-TIMS-MS EIC of galanin from hypothalamus extracts at four different charge states as indicated on the plot: $[m+3H]^{3+} = 1055.20$; $[m+4H]^{4+} = 791.65$; $[m+5H]^{5+} = 633.52$; $[m+6H]^{6+} = 528.10$. Arrows indicate peaks corresponding to Galanin, while the x over the peak at 38 minutes indicates an unrelated ion.

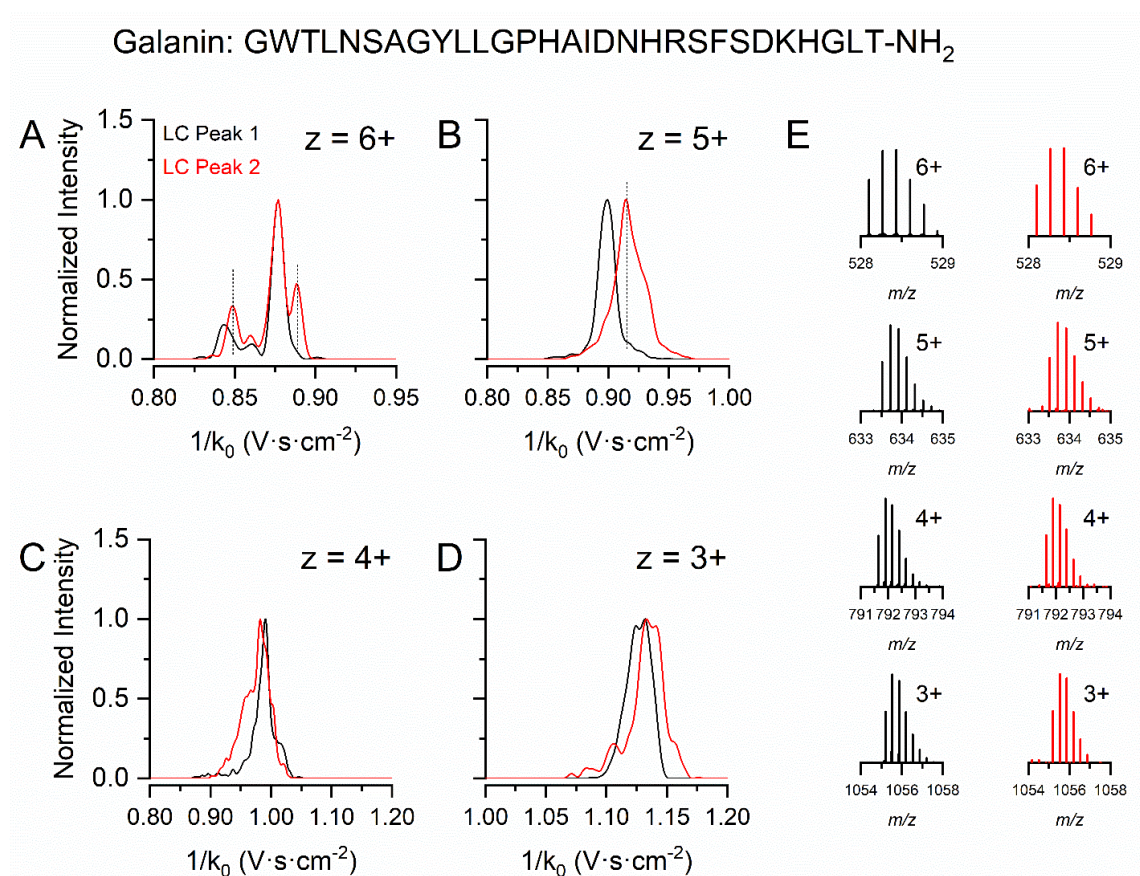


Figure 7.7. Comparison of mobility and MS1 profiles of Galanin in hypothalamus extracts by TIMS MS at four charge states. Black traces correspond to Galanin ions eluting at 36.5 min while red traces correspond to Galanin ion eluting at 37.5 min. (A-D) TIMS TOF MS extracted ion mobilograms of Galanin at four different charge states. (A) $[m+6H]^{6+} = 528.10$; (B) $[m+5H]^{5+} = 633.52$; (C) $[m+4H]^{4+} = 791.65$; (D) $[m+3H]^{3+} = 1055.20$. (E) MS1 profiles of galanin at different charge states as indicated on the plot. Black plots on left correspond to LC peak 1. Red plots on right correspond to LC peak 2.

Table 7.1. DAACP candidates in spinal cord assigned based on APM resistance or multiple peak elution. Peptides with split eluting peaks are indicated in the plot by a footnote. Unless otherwise noted, all peptides were resistant to APM digestion.

Peptide	Sequence	Precursor Accession
Scg1[597-611]	qYDDGVAELDQLLHY ^{a,b}	NP_036658.2
Scg2[184-216]	TNEIVEEQYTPQSLATLESVFQELGKLTGPSNQ	NP_073160.1
Scg2[185-216]	NEIVEEQYTPQSLATLESVFQELGKLTGPSNQ	
AAK1[777-796]	SDPFIPLQVPDAPEKLIEGL	XP_017448306.1
PENK[198-209]	SPQLEDEAKELQ ^{b 30-32}	NP_058835.2
PENK[239-259] (Enkelytin-like[1-21])	FAESLPSDEEGESY ^{<u>s</u>} KEVP ^c EM ^{c, 30, 33}	
PENK[239-259] (Enkelytin-like[1-21])	FAESLPSDEEGESYSKEVP ^c EM	
PENK[239-260] (Enkelytin-like[1-22])	FAESLPSDEEGESYSKEVP ^c EME	
PENK[114-133] (Amidorphin-like[6-27])	MDELYPVEPEEEEANGGEILA	
PENK[107-133] (Amidorphin-like)	YGGFMKKMDELYPVEPEEEEANGGEILA	
PENK[107-132] (Amidorphin-like[1-26])	YGGFMKKMDELYPVEPEEEEANGGEIL	
TKNK[95-115]	NSQPDTPADVVEENTPSFGVL	NP_062035.1
TRH[178-199] ³⁴	FIDPELQRSWEEKEGEGVLMPE ^d	NP_037178.1
TRH[179-199]	IDPELQRSWEEKEGEGVLMPE	
ProSAAS[221-242] (PEN)	AVDQDLGPEVPPENVLGALLRV	NP_062152.1
ProSAAS[221-240] (PEN[1-20])	AVDQDLGPEVPPENVLGALL	
ProSAAS[221-239] (PEN[1-19])	AVDQDLGPEVPPENVLGAL	
Sp3a-like[311-317]	FVAPAVT	NP_001005539.1
NFL[63-84]	SLMPSLENLDLSQVAAISNDLK	NP_113971.1

^a lowercase “q” indicates pyro-Glu at N-term
^b peptide had split eluting peaks
^c bold underlined lowercase “s” indicates phosphoserine
^d Peptide mostly degraded by APM

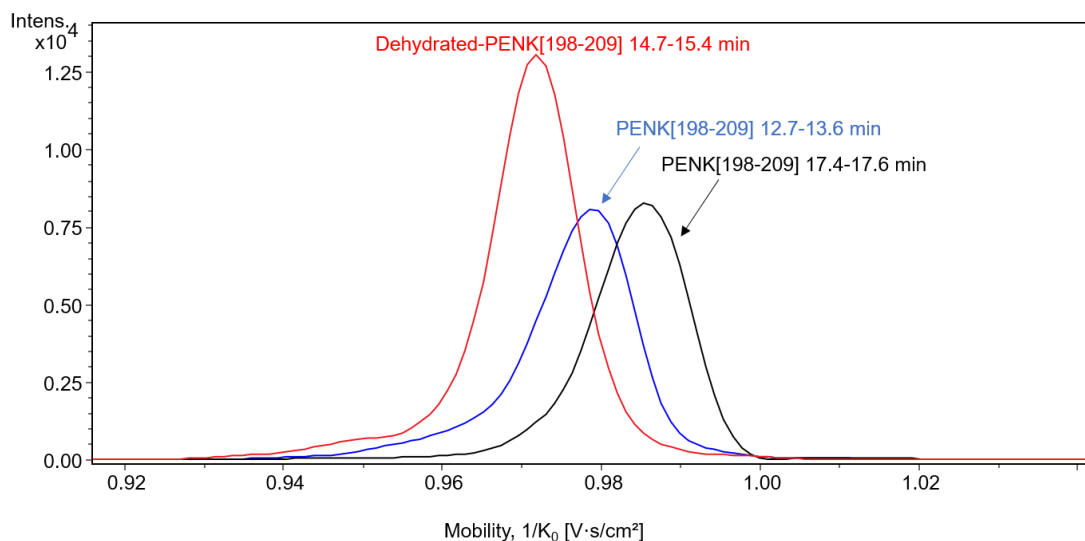


Figure 7.8. Mobility profiles of PENK[198-209] (black and blue traces, $[m+2H]^{2+} = 693.8$) and dehydrated PENK[198-209] (red trace, $[m-H_2O+2H]^{2+} = 684.8$). Profiles are the sum of the mobility profiles obtained over retention times indicated in the plot, corresponding to the chromatograms in Figure 7.3. The mobility profiles of both the unmodified PENK[198-209] peaks were similar, but showed slightly different maximum collisional cross sections at 400.0 Å² (Peak 1 at 13 min) and 403.0 Å² (Peak 2 at 17 min). Moreover, the mobility profile of the dehydrated PENK[198-209] was slightly shifted compared to the unmodified peptides, having a maximum CCS at 397.2 Å². The shift indicates the dehydration is associated with a decrease in CCS.

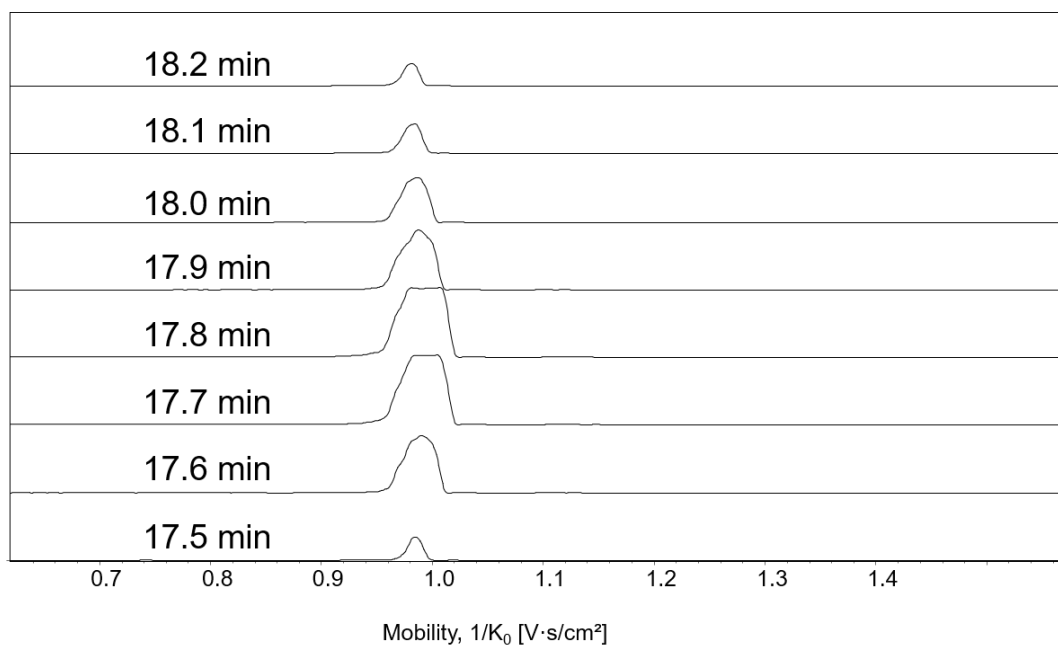


Figure 7.9. Mobility profile of PENK[198-209] sampled at various time points along the elution profile. Because the second eluting PENK[198-209] peak was so intense compared to the earlier PENK[198-209] peak, it was important to compare the mobility profiles of the two ions when they are at similar concentrations in the TIMS analyzer. The differences in mobility profile peak shape at each retention time can be attributed to space charge effects, which result from high abundance of ions in the TIMS analyzer.³⁵ Increasing the number of ions in the TIMS device can distort the mobility profile. This plot specifically demonstrates that to make comparisons between two eluting peaks, the retention time from which the mobility profile is sampled must be considered. This is especially important to consider when comparing the mobility profiles of a high abundance isomer to a low abundance isomer.

7.7 REFERENCES

1. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Advancing D-amino acid-containing peptide discovery in the metazoan. *Biochim Biophys Acta Proteins Proteom* **2021**, 1869 (1), 140553.
2. Schwartz, T. W.; Hansen, H. F.; Håkanson, R.; Sundler, F.; Tager, H. S., Human pancreatic icosapeptide: isolation, sequence, and immunocytochemical localization of the COOH-terminal fragment of the pancreatic polypeptide precursor. *Proc Natl Acad Sci U S A* **1984**, 81 (3), 708-12.
3. Koh, J. M.; Chow, S. J.; Crossett, B.; Kuchel, P. W., Mammalian peptide isomerase: platypus-type activity is present in mouse heart. *Chem Biodivers* **2010**, 7 (6), 1603-11.
4. Tai, H. Analysis of enzymatic L/D-peptide isomerization in animals Ph.D. Dissertation. University of Illinois at Urbana-Champaign, Online 2018 <http://hdl.handle.net/2142/101796> (accessed July 5th. 2021).
5. Yang, N.; Anapindi, K. D. B.; Romanova, E. V.; Rubakhin, S. S.; Sweedler, J. V., Improved identification and quantitation of mature endogenous peptides in the rodent hypothalamus using a rapid conductive sample heating system. *Analyst* **2017**, 142 (23), 4476-4485.
6. Purwaha, P.; Silva, L. P.; Hawke, D. H.; Weinstein, J. N.; Lorenzi, P. L., An artifact in LC-MS/MS measurement of glutamine and glutamic acid: in-source cyclization to pyroglutamic acid. *Anal Chem* **2014**, 86 (12), 5633-7.
7. Taylor, S. W.; Sun, C.; Hsieh, A.; Andon, N. L.; Ghosh, S. S., A sulfated, phosphorylated 7 kDa secreted peptide characterized by direct analysis of cell culture media. *J Proteome Res* **2008**, 7 (2), 795-802.

8. Gros, C.; Giros, B.; Schwartz, J. C., Identification of aminopeptidase M as an enkephalin-inactivating enzyme in rat cerebral membranes. *Biochemistry* **1985**, *24* (9), 2179-85.
9. Li, C. H.; Chung, D., Distinct behavior of beta-endorphin and corticotropin toward leucine aminopeptidase action. *Int J Pept Protein Res* **1985**, *26* (2), 113-7.
10. Gibson, A. M.; McDermott, J. R.; Lauffart, B.; Mantle, D., Specificity of action of human brain alanyl aminopeptidase on Leu-enkephalin and dynorphin-related peptides. *Neuropeptides* **1989**, *13* (4), 259-62.
11. Bausback, H. H.; Ward, P. E., Degradation of low-molecular-weight opioid peptides by vascular plasma membrane aminopeptidase M. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1986**, *882* (3), 437-444.
12. Churchill, L.; Bausback, H. H.; Gerritsen, M. E.; Ward, P. E., Metabolism of opioid peptides by cerebral microvascular aminopeptidase M. *Biochim Biophys Acta* **1987**, *923* (1), 35-41.
13. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Differential Post-Translational Amino Acid Isomerization Found among Neuropeptides in *Aplysia californica*. *ACS Chem Biol* **2020**, *15* (1), 272-281.
14. Mast, D. H.; Liao, H. W.; Romanova, E. V.; Sweedler, J. V., Analysis of Peptide Stereochemistry in Single Cells by Capillary Electrophoresis-Trapped Ion Mobility Spectrometry Mass Spectrometry. *Anal Chem* **2021**, *93* (15), 6205-6213.
15. Aswad, D. W.; Paranandi, M. V.; Schurter, B. T., Isoaspartate in peptides and proteins: formation, significance, and analysis. *J Pharm Biomed Anal* **2000**, *21* (6), 1129-36.

16. Kumar, M.; Chatterjee, A.; Khedkar, A. P.; Kusumanchi, M.; Adhikary, L., Mass spectrometric distinction of in-source and in-solution pyroglutamate and succinimide in proteins: a case study on rhG-CSF. *J Am Soc Mass Spectrom* **2013**, *24* (2), 202-12.
17. Ouellette, D.; Chumsae, C.; Clabbers, A.; Radziejewski, C.; Correia, I., Comparison of the in vitro and in vivo stability of a succinimide intermediate observed on a therapeutic IgG1 molecule. *MAbs* **2013**, *5* (3), 432-44.
18. Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S., Spontaneous cyclization of the aspartic acid side chain to the succinimide derivative. *Journal of the Chemical Society, Chemical Communications* **1992**, (12), 919-921.
19. Yu, X. C.; Joe, K.; Zhang, Y.; Adriano, A.; Wang, Y.; Gazzano-Santoro, H.; Keck, R. G.; Deperalta, G.; Ling, V., Accurate determination of succinimide degradation products using high fidelity trypsin digestion peptide map analysis. *Anal Chem* **2011**, *83* (15), 5912-9.
20. Stephenson, R. C.; Clarke, S., Succinimide formation from aspartyl and asparaginyll peptides as a model for the spontaneous degradation of proteins. *J Biol Chem* **1989**, *264* (11), 6164-70.
21. Norberg, Å.; Griffiths, W. J.; Hjelmqvist, L.; Jörnvall, H.; Rökaeus, Å., Identification of variant forms of the neuroendocrine peptide galanin. *Rapid Communications in Mass Spectrometry* **2004**, *18* (14), 1583-1591.
22. Takahashi, O.; Kirikoshi, R.; Manabe, N., Acetic acid can catalyze succinimide formation from aspartic acid residues by a concerted bond reorganization mechanism: a computational study. *Int J Mol Sci* **2015**, *16* (1), 1613-26.
23. Janetzko, J.; Walker, S., Aspartate Glycosylation Triggers Isomerization to Isoaspartate. *J Am Chem Soc* **2017**, *139* (9), 3332-3335.

24. Ni, W.; Dai, S.; Karger, B. L.; Zhou, Z. S., Analysis of isoaspartic Acid by selective proteolysis with Asp-N and electron transfer dissociation mass spectrometry. *Anal Chem* **2010**, 82 (17), 7485-91.
25. O'Connor, P. B.; Cournoyer, J. J.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A., Differentiation of aspartic and isoaspartic acids using electron transfer dissociation. *J Am Soc Mass Spectrom* **2006**, 17 (1), 15-19.
26. Cournoyer, J. J.; Pittman, J. L.; Ivleva, V. B.; Fallows, E.; Waskell, L.; Costello, C. E.; O'Connor, P. B., Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Sci* **2005**, 14 (2), 452-63.
27. Xiao, G.; Bondarenko, P. V.; Jacob, J.; Chu, G. C.; Chelius, D., 18O Labeling Method for Identification and Quantification of Succinimide in Proteins. *Analytical Chemistry* **2007**, 79 (7), 2714-2721.
28. Bonnard, E.; Burlet-Schiltz, O.; Francés, B.; Mazarguil, H.; Monsarrat, B.; Zajac, J. M.; Roussin, A., Identification of neuropeptide FF-related peptides in rodent spinal cord. *Peptides* **2001**, 22 (7), 1085-92.
29. Ye, H.; Wang, J.; Tian, Z.; Ma, F.; Dowell, J.; Bremer, Q.; Lu, G.; Baldo, B.; Li, L., Quantitative Mass Spectrometry Reveals Food Intake-Induced Neuropeptide Level Changes in Rat Brain: Functional Assessment of Selected Neuropeptides as Feeding Regulators. *Mol Cell Proteomics* **2017**, 16 (11), 1922-1937.
30. Bernay, B.; Gaillard, M. C.; Guryca, V.; Emadali, A.; Kuhn, L.; Bertrand, A.; Detraz, I.; Carcenac, C.; Savasta, M.; Brouillet, E.; Garin, J.; Elalouf, J. M., Discovering new bioactive neuropeptides in the striatum secretome using in vivo microdialysis and versatile proteomics. *Mol Cell Proteomics* **2009**, 8 (5), 946-58.

31. Bora, A.; Annangudi, S. P.; Millet, L. J.; Rubakhin, S. S.; Forbes, A. J.; Kelleher, N. L.; Gillette, M. U.; Sweedler, J. V., Neuropeptidomics of the supraoptic rat nucleus. *J Proteome Res* **2008**, 7 (11), 4992-5003.
32. Haskins, W. E.; Watson, C. J.; Cellar, N. A.; Powell, D. H.; Kennedy, R. T., Discovery and neurochemical screening of peptides in brain extracellular fluid by chemical analysis of in vivo microdialysis samples. *Anal Chem* **2004**, 76 (18), 5523-33.
33. Lundby, A.; Secher, A.; Lage, K.; Nordsborg, N. B.; Dmytriiev, A.; Lundby, C.; Olsen, J. V., Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nat Commun* **2012**, 3, 876.
34. Bulant, M.; Delfour, A.; Vaudry, H.; Nicolas, P., Processing of thyrotropin-releasing hormone prohormone (pro-TRH) generates pro-TRH-connecting peptides. Identification and characterization of prepro-TRH-(160-169) and prepro-TRH-(178-199) in the rat nervous system. *J Biol Chem* **1988**, 263 (32), 17189-96.
35. Fernandez de la Mora, J., Space Charge Effects on Ion Mobility Spectrometry. *Journal of the American Society for Mass Spectrometry* **2019**, 30 (6), 1082-1091.

CHAPTER 8

CONCLUSION AND FUTURE DIRECTIONS

8.1 CONCLUSION

Enzymatic L- to D-residue isomerization of cell-cell signaling peptides exemplifies the unexplored complexity of biochemical pathways that animals have evolved. It seems counterintuitive that animals—which have evolved a highly stereoselective protein synthesis pathway—would possess an enzyme to convert all-L peptides into D-amino acid containing peptides (DAACPs). At present, post-translational L- to D-residue isomerization has only been discovered in relatively few animals. Consequently, DAACPs are considered, “rare” among animals, and unique to only a small number of signaling peptides.

The work that has been presented was premised upon the idea that DAACPs have been considered “rare” because modern mass spectrometry (MS)-based peptidomics studies do not look for them; as a result, the prevalence of L- to D-residue isomerization, and DAACPs, may be underestimated in the metazoan. This dissertation has demonstrated that, isomeric peptides, including DAACPs, can be rapidly identified using modern peptidomics and bioinformatics strategies through the critical examination of the enzymatic stability, or chromatographic, mass spectrometric, and ion mobility spectrometry properties of endogenous peptides.

Summary of the Dissertation Findings

In this dissertation, stereoselective enzymatic digestions with aminopeptidase M (APM) allowed for the differentiation of DAACPs from all-L peptides in tissue homogenates by liquid-chromatography (LC) tandem mass spectrometry (MS/MS). The increased stability of DAACPs

against proteolysis by APM, allowed them to be identified and differentiated from all-L residue peptides. Here, this general approach, while not new, was refined and successfully incorporated into a generalized peptidomics workflow. To distinguish a peptide as a plausible DAACP, it needed to exhibit a strict set of chromatographic, mass spectrometric, or chemical properties (see Chapter 3 for details). A peptide was on considered a DAACP candidate if it met all of the strict, predefined conditions. And if a peptide met those conditions, it was subsequently tested for the presence of a D-residue.

This systematic approach for detecting DAACPs in peptide extracts was described and applied in Chapter 3. It is distinguished from other approaches because it was the first successful study to intentionally to discovery DAACPs, without relying on bioactivity, or sequence homology to previously discovered DAACPs in other organisms.¹

Prior work had already reported the existence of several DAACPs in the *A. californica* CNS. But the application of unbiased DAACP screening reported in Chapter 3 revealed eight novel DAACPs that had been missed by previous studies. These eight DAACPs were encoded by neuropeptide precursor proteins that had previously not been known to encode isomerized peptides in any species. In addition, three of the DAACPs discovered in *A. californica* were encoded novel prohormone precursors, that were characterized, at least in part, for the first time. Several of these DAACPs were studied at the single cell level in Chapters 4 and 5. It was discovered that some DAACPs in *A. californica* do not localize the cell-bodies of individual neurons.

Unlike *A. californica*, enzyme-derived cell-cell signaling DAACPs have never been discovered in a mammalian species.¹ This was reviewed in Chapter 2. However, L- to D-residue isomerase activity, similar to that observed in *A. californica*, has been reported in rodent tissue

homogenates.²⁻³ The latter observations inspired a re-evaluation of endogenous peptides in *R. norvegicus* tissues to identify any DAACPs that may have been previously overlooked.

In Chapters 6 and 7 of this dissertation, the same unbiased methods that were used to discover DAACPs in *A. californica*, were ambitiously applied to search for cell-cell signaling DAACPs in *R. norvegicus*. However, in preliminary explorations of pancreatic islets (Chapter 6), and spinal cord (Chapter 7), no DAACPs could be identified. This does not necessarily reflect a failure of the unbiased screening protocol to identify DAACPs, and it does not prove mammals do not have DAACPs. The reason DAACPs could not be identified in the rodent tissues tested is because none of the peptides met the predefined criteria that was used for categorizing DAACP candidates. While the criteria were successful for DAACP discovery in Chapter 3, it is possible that DAACPs in *R. norvegicus* (or mammals in general) do not follow the same patterns as the DAACPs in *A. californica* (or invertebrates in general)—this was discussed in Chapter 7.

To discover DAACPs in mammals, it may be necessary to refine or broaden the criteria for identifying DAACP candidates. Therefore, in Chapter 7, rather than categorizing a peptide as a DAACP candidate, peptides were categorized as possible isomeric peptides if they exhibited some, but not all of the properties expected of DAACPs. For example, peptides that were fully resistant to APM, but did not have multiple eluting peaks by LC-MS/MS were categorized as potentially being isomeric because of their enzymatic stability. Similarly, peptides that had split eluting peaks in general, regardless of their stability against APM digestion, were also assigned as possible isomeric peptides because two indistinguishable sequences of the peptide were separable by LC. Using these more lenient criteria, several peptides of interest were identified in *R. norvegicus* spinal cord and CNS tissue that could have isomeric forms. But the peptides listed as potentially isomeric peptides in Chapter 7, may not necessarily be DAACPs. There are many different

explanations for why they exhibited properties expected of peptide isomers. Chapter 7 specifically focused on the possibility of spontaneous Asp isomerization. Certain chemical reactions that may be indicative Asp isomers are present (i.e., dehydration) were discussed to help elucidate the most likely reason for the detection of isomeric peaks by LC-MS/MS.

In Chapter 7, three peptides were found that were categorized as potentially isomeric peptides and showed signs of undergoing spontaneous Asp isomerization. These peptides include Galanin, pGlu-Scg1[597-611], and PENK[198-209]. Finally, a simple spiking experiment with synthetic standards revealed the all-L form of pGlu-Scg1[597-611] was present in spinal cord extracts, but the second isomer that was detected did not have a D-residue at the second position from the N-terminus. This suggested that the second isomeric form of pGlu-Scg1[597-611] detect in spinal cord extracts was not a DAACP product of *A. californica*-like L- to D-residue isomerase activity. The results from Chapter 7 indicate alternative methods may need to be developed to screen for DAACPs in mammalian model organisms.

Future Directions

Accompanying the discoveries made in *A. californica* are a plethora of unresolved questions, that open the door for numerous future studies. Perhaps the most interesting question to ask is whether DAACPs are as common in other species as they are in *A. californica*.

Chapter 2 discussed the general trend that, if a peptide is a DAACP in one organism, then homologues of that peptide in related organisms are most likely also DAACPs. The prohormones that encode isomerized peptides in *A. californica* are highly conserved across mollusk species.⁴ Therefore, the assertion from Chapter 3, that DAACPs are more common than previously thought in *A. californica*, can be expanded more broadly to include mollusk species in general. This

generalization is not meant to suggest that every possible DAACP in mollusks has been discovered. In fact, it is quite the contrary. It is more likely that there are many more DAACPs in other mollusks that remain undiscovered, and this dissertation establishes an experimental framework that will enable future studies to explore this hypothesis. However, the discovery of new DAACPs in mollusk species (including *A. californica*) will no longer be unexpected. It would almost be more surprising if DAACPs were not as common in other mollusk species as they are in *A. californica*.

Most of the cell-cell signaling DAACPs discovered to-date have been discovered in mollusks (Chapter 2).¹ While DAACPs are likely prevalent in mollusks, the same cannot yet be assumed for species from other phyla. For example, only two cell-cell signaling DAACPs have been discovered in crustaceans, and no cell-cell signaling DAACPs have been discovered in mammalian species. Further exploration of other phyla will be necessary to elucidate the true prevalence of cell-cell signaling DAACPs in the metazoan.

Other species need to be explored for DAACPs. Cell-cell signaling DAACPs have also been discovered in crustaceans, but the prevalence of DAACPs in arthropods is still unknown. In crustaceans, only two cell-cell signaling DAACPs have been discovered, the crustacean hyperglycemic hormone (CHH) and the vitellogenesis inhibiting hormone (VIH).^{1,5} VIH and CHH share little sequence resemblance to any known DAACPs in mollusks. VIH has a D-residue at position 4 from the N-terminus, while CHH has a D-residue at position 3 from the N-terminus.⁵ The isomerization of VIH and CHH at residues 4 and 3 from their N-termini, respectively, suggests that the specificity of isomerases may differ in other phyla compared to *A. californica*. However, it is also possible that isomerization of neuropeptides at position 2 from the N-terminus can also occur in crustaceans but this has not been explored yet. A complete investigation of endogenous

DAACPs in crustaceans may help to further understand the evolution and specificity of the unidentified L- to D-residue isomerases. Interestingly, while the *A. californica* pleurin peptides share virtually no sequence similarity with CHH and VIH, the differential isomerization of Plrn2 and Plrn3 (see Chapter 3) hints at an unexpected connection between mollusks and crustaceans, related to isomerization.

Among vertebrate species, DAACPs have been discovered in frogs, and in platypus, but these DAACPs are not considered cell-cell signaling DAACPs.¹ Chapter 6 and Chapter 7 offer an initial glimpse into the prevalence of DAACPs in a mammal. Continued exploration of mammals for DAACPs will be important. Although Chapters 6 and 7 were only a preliminary exploration of *R. norvegicus*, DAACPs could not be confidently identified in spinal cord or pancreatic islets. This could indicate cell-cell signaling DAACPs are much less prevalent in the mammalian CNS and endocrine system than they are in the *A. californica* CNS. However, more comprehensive studies need to be performed. It is possible alternative methods will be required to discover DAACPs in mammalian tissues. Future work to help elucidate the prevalence of DAACPs in mammals may consist of the continuation of APM screening for DAACPs in other neuro-endocrine tissues of *R. norvegicus* and other species. The importance of screening various tissues cannot be understated because in Chapter 3, and Chapter 4, it was shown that DAACPs can be highly localized to specific tissues in *A. californica*. Similar heterogeneous distributions of DAACPs may be possible in mammals. Therefore, just because a peptide is not detected as a DAACP in one tissue, does not exclude it from possibly being a DAACP in another tissue.

Chapter 2 provided a comprehensive list of animals where DAACPs have been found. There is a high likelihood that many DAACPs have gone undiscovered, especially in animals where DAACPs have previously already been identified. For example, DAACPs have been isolated from

arachnid venoms, but arachnid nervous systems have never been tested for cell-cell signaling DAACPs. Similarly, DAACP toxins have been found in frog skin secretions, but the frog central nervous system has never been tested for cell-cell signaling DAACPs. But future work should not be limited to organisms where DAACPs have already been discovered. Many organisms have yet to be searched for DAACPs. For example, birds, lizards, snakes, bees, flies, worms (annelids), etc., have not been tested for the presence of endogenous cell-cell signaling DAACPs. The extension of the methods described here to search for DAACPs in other model organisms will ultimately reveal how common L- to D-residue isomerization is in animals.

DAACPs may not actually be as rare as they appear in animals. This dissertation has demonstrated that enzymatic screening for DAACPs can be easily incorporated into LC-MS/MS based peptide characterization workflows. Widespread adaptation and standardization of APM screening into peptidomics workflows will undoubtedly reveal the true prevalence of DAACPs among cell-cell signaling peptides in animals.

8.2 REFERENCES

1. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Advancing D-amino acid-containing peptide discovery in the metazoan. *Biochim Biophys Acta Proteins Proteom* **2021**, 1869 (1), 140553.
2. Tai, H. Analysis of enzymatic L/D-peptide isomerization in animals Ph.D. Dissertation. University of Illinois at Urbana-Champaign, Online 2018 <http://hdl.handle.net/2142/101796> (accessed July 5th. 2021).
3. Koh, J. M.; Chow, S. J.; Crossett, B.; Kuchel, P. W., Mammalian peptide isomerase: platypus-type activity is present in mouse heart. *Chem Biodivers* **2010**, 7 (6), 1603-11.

4. De Oliveira, A. L.; Calcino, A.; Wanninger, A., Extensive conservation of the proneuropeptide and peptide prohormone complement in mollusks. *Sci Rep* **2019**, *9* (1), 4846.
5. Ollivaux, C.; Soyez, D.; Toullec, J.-Y., Biogenesis of d-amino acid containing peptides/proteins: where, when and how? *Journal of Peptide Science* **2014**, *20* (8), 595-612.

APPENDIX A

CHIRAL AMINO ACID ANALYSIS OF PEGVADIN A FROM STREPTOMYCES NOURSEI

A.1 NOTES AND ACKNOWLEDGEMENTS

The work reported here was previously published. Georgiou, M. A.; Dommaraju, S. R.; Guo, X.; Mast, D. H.; Mitchell, D. A., Bioinformatic and Reactivity-Based Discovery of Linaridins. *ACS Chemical Biology* **2020**, *15* (11), 2976-2985. DOI: <https://doi.org/10.1021/acscchembio.0c00620> .Only the experimental methods used for the stereochemical analysis of Pegvadin A are described in this appendix. The published work in its entirety is not described.

A.2 INTRODUCTION

Linaridins are linear, dehydrated peptides.¹ They represent a class of ribosomally synthesized and post-translationally modified peptides (RiPPs) characterized by the presence of dehydro-amino acid residues formed enzymatically. Linardins are encoded by a precursor protein in the genome prokaryotic organisms and are heavily post-translationally modified. Some of the enzymes responsible for modifying these peptides have not been identified, making the structural characterization of this class of peptide difficult.¹

The linaridin Pegvadin A (PegA), isolated from *Streptomyces noursei*, is the subject of this appendix. PegA was discovered using a bioinformatics approach, described in detail in the full publication.² Although the presence of D-amino acid residues in RiPPs is relatively uncommon, the first D-amino acid containing peptide linaridin, Salinpeptin isolated from *Streptomyces*, was

recently discovered.³ The finding of the unusual D-amino acid residue containing linaridin, isolated from from *Streptomyces* named salinpeptin, prompted the stereochemical investigation of PegA. D-amino acids were not expected to be present in PegA, but Chiral amino acid analysis was performed, and it was found to be comprised almost entirely of D-amino acid residues.

A.3 METHODS

Marfey's Analysis of pegvadin A stereochemistry

The sequence of PegA is as follows: LTPLATPEGVAVTSATAYALAHEVTKDIQR. The N-terminal Leu is dimethylated, and the threonine's are dehydrobutyrines. The stereochemistry of the N-terminal Leu was not determined.

Alongside pegvadin A (PegA), labeling blanks, consisting of two empty hydrolysis tubes, were tested. After microwave assisted hydrolysis these blanks were reacted with FDAA to evaluate baseline levels of amino acids present in the background. A peptide containing all L-amino acids, YLDHLGSSLV, was also hydrolyzed under the same conditions. This was done to ensure amino acid racemization did not occur during the reaction.

50 µg of PegA was reconstituted in 200 µL of LC-MS grade methanol. The reconstituted PegA was then transferred to a glass hydrolysis tube and methanol was removed by vacuum centrifugation. Microwave assisted acid hydrolysis was performed using a CEM Discover microwave (CEM corporation).

5 mL of 6 M ^2HCl in $^2\text{H}_2\text{O}$ with 0.1% phenol (unstablized) was used added to the reaction vessel. The vessel was then sealed, and microwave assisted hydrolysis was performed under the following conditions: temperature and pressure were ramped to 165 °C and 60 PSI in 5 minutes, and then held for 30 minutes. Following the hydrolysis, the vessel was cooled, and excess

condensed liquid in glass reaction vials was removed by vacuum centrifugation. The hydrolysate was reconstituted in 55 μL of 0.5 M NaHCO_3 and sonicated for 30 s. For amino acid derivatization, 20 μL of 1 mg/mL 1-fluoro-2-4-dinitrophenyl-5-L-alanine amide (Marfey's reagent, FDAA) in CH_3CN was combined with 25 μL of hydrolysate and incubated for 3 h at 63 ± 3 $^\circ\text{C}$. Following incubation, the reaction mixture was cooled to room temperature and 10 μL was diluted to 200 μL with 25 mM NH_4HCO_2 (ammonium formate) in LC-MS grade H_2O and 20 μL of injected onto the LC. For labeling of amino acid standards, 25 μL of amino acid standards (200 μM each amino acid) was combined with 20 μL of 1mg/mL of FDAA and incubated/analyzed as above. After incubation with FDAA, the labeled L/D AA standards were diluted to 1 μM with 25 mM NH_4HCO_2 and 20 μL was then injected onto the LC-MS. For labeling blanks, tube contents were reconstituted in 55 μL of 500 mM sodium bicarbonate and 25 μL was then combined with 20 μL of 1 mg/mL FDAA in CH_3CN . After incubation, the *amino* acid standards 10 μL of the FDAA reacted labeling blank solution was diluted to 200 μL with 25 mM NH_4HCO_2 and 20 μL was injected. Liquid chromatography (LC)-multiple reaction monitoring (MRM)-mass spectrometry (MS) analysis was performed using an EVOQ elite, triple quad mass spectrometer equipped with an Elute HT UHPLC (Bruker Daltonics) and a Kinetex Phenyl-Hexyl column (100 Å pore size, 2.6 μm particle size, 100 mm x 2.1 mm). Amino acids were purified under the following conditions: solvent A (25 mM NH_4HCO_2) and solvent B (MeOH), flow rate of 0.3 mL/min and the following gradient: t=0 min, 5% B, t=2 min, 5% B, t=7 min, 20% B, t=12 min, 55% B, t=15 min, 55% B, t=18 min, 90% B, t=21, 90% B. Ions monitored by MRM are shown in Table A.1.

A.4 FIGURE AND DATA

Table A.1. Ions monitored by MRM MS				
Amino Acid	Precursor <i>m/z</i>	Product <i>m/z</i>	Voltage (V)	Mode
Ala	340	278.1	14	(-)
Arg	426.9	70.1	20	(+)
	426.9	130.1	20	
Asp	384.0	267.9	30	(-)
	384.0	207.0	30	
Glu	398.0	202.0	24	(-)
Gly	326.1	162.0	32	(-)
His	658.2	549.1	19	(-)
Tyr	684.1	351.9	28	(-)
Tyr +2 ^a	686.1	353.9	27	(-)
Pro	365.9	321.9	15	(-)
Leu/Ile	382.0	319.8	18	(-)
Val	367.9	306.0	15	(-)
Ser	355.9	263.9	16	(-)
Lys	649.1	479.1	36	(-)
^a the (m+2)/z ion was used to detect Tyr in the hydrolysate.				

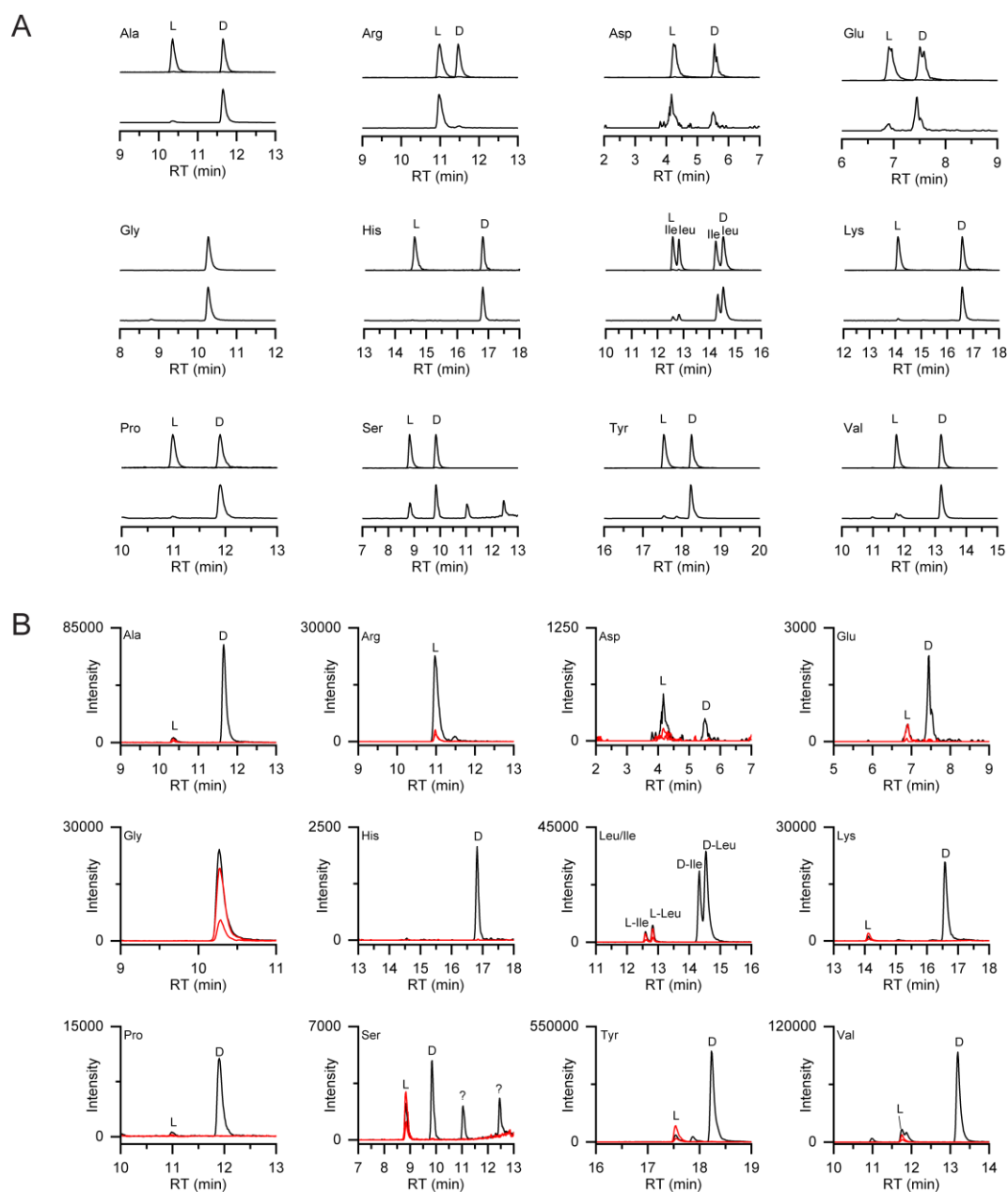


Figure A.1 Marfey's Stereochemical analysis of pegvadin A. (A) Intensity normalized LC-MRM MS chromatograms of FDAA-labeled amino acid standards (top) and pegvadin A hydrolysate (bottom). (B) Non-normalized LC-MRM-MS chromatograms of FDAA-labeled hydrolyzed pegvadin A (black) and a blank hydrolysis reaction.

Table A.2 Marfey's Analysis results summary.							
Amino acid	L-peak area	D-Peak area	D/(L+D)×100 (%D uncorrected)	L-peak area labeling blank 1	L-peak area labeling blank 2	Prediction	Notes
Ala	524	9017	94.5	336	149	D-isomer	
Asx	97	38	28.1	40	20	Unable to determine	Low intensity in PegA hydrolysate but detectable No D-Asx peak detected in labeling blanks. Uncertain whether L-Asp or D-Asp present in peptide.
Glx	57	242	80.9	51	9	D-isomer	Low intensity, no D-Glx peak detected in labeling blanks. Likely only D-Glx in peptide.
His	0	206	100.0	0	0	D-isomer	Low intensity, but it appears that D-His is likely the stereoisomer present. No L-His or D-His signal in labeling blanks.
Ile	404	2998	88.1	320	100	D-isomer	No D-Ile peak detected in labeling blanks.
Leu	612	5425	89.9	456	162	D-isomer	Observed ratio is 1.8:1 Leu:Ile, rather than predicted 3:1 Leu:Ile. No D-Leu peak detected in labeling blanks.
Lys	141	3045	95.6	297	74	D-isomer	No D-Lys peak detected in labeling blanks.
Val	1416	11234	88.8	825.1	284	D-isomer	No D-Val peak detected in labeling blanks.
Ser	287	590	67.3	425	136	D-isomer	Presence of L-Ser in hydrosylate and blanks can be attributed to background. No D-Ser peak detected in labeling blanks.
Arg	3316	213	6.0	325	237	L-isomer	No D-Arg peak detected in labeling blanks.
Pro	78	1312	94.4	44	17	D-isomer	No D-Pro peak detected in labeling blanks.
Tyr	3414	43716	92.8	8544	1800	D-isomer	No D-Tyr peak detected in labeling blanks.

A.5 REFERENCES

1. Ma, S.; Zhang, Q., Linaridin natural products. *Nat Prod Rep* **2020**, *37* (9), 1152-1163.
2. Georgiou, M. A.; Dommaraju, S. R.; Guo, X.; Mast, D. H.; Mitchell, D. A., Bioinformatic and Reactivity-Based Discovery of Linaridins. *ACS Chem Biol* **2020**, *15* (11), 2976-2985.
3. Shang, Z.; Winter, J. M.; Kauffman, C. A.; Yang, I.; Fenical, W., Salinipeptins: Integrated Genomic and Chemical Approaches Reveal Unusual D-Amino Acid-Containing Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs) from a Great Salt Lake Streptomyces sp. *ACS Chem Biol* **2019**, *14* (3), 415-425.

APPENDIX B

SUPPLEMENTAL INFORMATION FOR CHAPTER 3 AND CHAPTER 4

B.1 NOTES AND ACKNOWLEDGEMENTS

Work involving qPCR was completed with assistance from J. Checco. For work involving *Aplysia californica* ganglia cultures, E. Romanova performed animal dissections, ganglia cultures and cell isolations. For capillary electrophoresis was performed with assistance from H. Liao.

B.2 INTRODUCTION

Pleurin is a neuropeptide precursor in mollusks. In *A. californica*, it was discovered that pleurin encodes three amidated peptides (Plrn1, Plrn2, and Plrn3) that undergo post-translational L- to D-residue isomerization into D-amino acid containing peptides. As discussed in Chapter 3, two of the peptides encoded by pleurin (Plrn2, and Plrn3) undergo a form of differentiation processing wherein the all-L peptide is isomerized at the amino acid residue at position 2 from the N-terminus or position 3 from the N-terminus. Interestingly, the differentially isomerized forms of Plrn2 and Plrn3 appeared to localize to different regions of the central nervous system.

Plrn2 (SFYTTGNGNHYPRI-NH₂), exists in two different epimer forms, [D-Phe²]-Plrn2 (S**d**FYTTGNGNHYPRI-NH₂) and [D-Tyr³]-Plrn2 (SF**d**YTTGNGNHYPRI-NH₂). These forms of Plrn2 differ in that one has a D-residue in position 2, while the other has a D-residue in position 3 from the N-terminus. Furthermore, these epimers were found to be localized in different regions of the *A. californica* CNS. [D-Tyr³]-Plrn2 was detected in the buccal ganglia region while [D-Phe²]-Plrn2 was detected in the cerebral ganglion. like Plrn2, the Plrn3 peptide was also found to be

isomerized into two different D-amino acid containing epimers, [D-allo²]-Plrn3 (GdIFTQSAYGSYPRI-NH₂) and [D-Phe³]-Plrn3 (GIdFFTQSAYGSYPRI-NH₂). The epimers of Plrn3 are also localized to different regions of the CNS. Pleurin also encodes Plrn1 (MFYTKGSDSDYPRI-NH₂) which was not found to be differentially isomerized. This is significant because it suggests a single peptide can be processed in different ways depending on the tissue which it is localized in, and the different forms could have unique functions.

The purpose of this appendix section is to report the results and observations from experiments designed to elucidate the mechanisms behind the differential isomerization of pleurin. The questions asked, and the experiments designed to answer them, are discussed in the following sections of the introduction.

Is there a D2D3 Diastereomer of Plrn2 or Plrn3?

One question that was not answered in Chapter 3 was the following: Is it possible Plrn2 and Plrn3 have a third DAACP form, with D-residues at both position 2 and position 3 (i.e. SdFdYTTGNGNHYPRI-NH₂, and GIdIdFFTQSAYGSYPRI-NH₂). Spiking experiments with synthetic peptides were performed to answer this question. The “D2D3” epimers of the pleurin derived peptides were synthesized and compared to the retention times of the synthetic D2 and D3 diastereomers.

Do DAACPs Accumulate in Neuron Cell Bodies if Vesicle Transport is Inhibited?

In Chapter 4, the stereochemistry of the pleurin derived peptides was measured in single neurons using capillary electrophoresis (CE) coupled to trapped ion mobility spectrometry (TIMS) mass spectrometry (MS). It was determined that Plrn1, Plrn2 and Plrn3 are not present in the

DAACP forms in the cell bodies of neurons expressing the pleurin precursor. However, they were detected in ganglia connectives, which are comprised of axons/cellular processes of interneurons. The localization of the DAACP forms to the connectives, but not the cell bodies was suggested as evidence that L- to D-residue isomerization occurs slowly during anterograde vesicle transport. In this appendix, ganglia culture experiments were designed and performed to test this hypothesis.

Here the following question was asked: Does preventing vesicle transport inhibit cause DAACPs to accumulate in cell bodies? This question was tested by culturing isolated ganglia in the presence of the vesicle transport inhibitor, colchicine, and measuring the stereochemistry of Pleurin-derived peptides in neuron cell bodies that were cultured with or without the vesicle transport inhibitor.

Is there a Pleurin Expressing Neuron Population in the Buccal Ganglia that Exclusively Isomerizes Plrn2 and Plrn3 into D3 the Diastereomers?

In Chapter 3, the D3 forms, but not the D2 forms of Plrn2 and Plrn3 were detected in the buccal ganglion. One hypothesis to explain this unique distribution of Plrn2 and Plrn3 diastereomers in the *A. californica* CNS may be that the buccal ganglia has a pleurin-expressing cell population the exclusively isomerizes Plrn2 and Plrn3 into the D3 forms.

This hypothesis was initially tested simply by testing whether *pleurin* mRNA is expressed in the buccal ganglion. The relative *pleurin* mRNA expression in various regions of the *A. californica* CNS was measured by qPCR. The results found pleurin mRNA could not be identified at all in the buccal ganglion suggesting there is not a Pleurin-expressing cell population in the buccal ganglion at all.

Why Does the Mobilogram of the Plrn2 [m+2H]²⁺ Ion Have Two Peaks?

In Chapter 3 and Chapter 4 the Plrn2 [m+2H]²⁺ ion mobilogram obtained by TIMS MS was observed with two peaks. This suggests two stable gas phase conformers were present. One possibility is that the conformers are the result of *cis-trans* proline isomerization. To test this, the Pro residue in Plrn2 was substituted for an Ala residue and tested by trapped ion mobility.

B.3 METHODS

LC-MS analysis of synthetic Plrn1-Plrn3 diastereomers.

LC gradient conditions, and LC-MS methods were the same those used for testing synthetic Plrn1-Plrn3 diastereomers reported in Chapter 3. Individual solutions of 1 μ M synthetic Plrn1-Plrn3 diastereomers were prepared in 5% CH₃CN in water with 0.1% formic acid. 1-3 pmol was injected for LC-MS analysis. Mixtures of diastereomer combinations, 0.25-0.5 μ M each, were also prepped, and 1 pmol was injected for analysis.

Ganglia Culture with Colchicine

Inhibition of axonal transport to increase accumulation of potential pleurin peptides in the pleural neurons was carried out on symmetrical isolated and desheathed pairs of pleuro-pedal hemiganglia from two animals. The CPICs were cut off in vicinity of pleural hemiganglia; cerebro-pedal connective were severed. The hemiganglia were incubated for ~18h at 14°C in either ASW or in ASW containing the vesicle transport blocker colchicine (10 μ M) (ref). Incubation solutions for the left and right ganglionic pairs were reversed among the two animals. After incubation, putative pleurin expressing neurons (n = 12 colchicine-incubated + 13 ASW-incubated = 25 total from both animals) were manually isolated from left and right pleural ganglia according

to the pleurin in situ staining map and processed for peptide extraction by acidified methanol, as described above. Presence of Plrn2 peptide isomers was assessed by CE-TIMS-MS using the methods described in Chapter 4.

qPCR

For each biological set, three *Aplysia* (100-150 g) were dissected, and RNA was extracted from ganglia, as previously described.¹ cDNA synthesis for each synthesis was then carried out using iScript cDNA synthesis kit (Bio-Rad, 1708890) using 1 µg of RNA in a 20 µL reaction, following the manufacturer's instructions. After synthesis, each reaction solution was diluted to 50 µL with water and stored at -20 °C until qPCR analysis. Comparison of gene expression among various tissues by qPCR was performed using Power SYBR Green PCR Master Mix (Applied Biosystems, 4368708) using GAPDH as a reference gene.¹ For each tissue type and biological set, 10 µL of the cDNA stock was diluted to 40 µL with water, and 4 µL of this diluted sample was added to 16 µL of Master Mix with primers (final primer concentration = 160 nM each). The following primers were used for analysis: Pleurin (AY833131.1, Forward 5'-CAT GCG CAA TGT ACA CC-3', reverse, 5'-AAA TCG GAA AAC CCT GAA GC-3') GAPDH (NM_001280826.1, forward 5'-ACC TGC AGA TTG GAG AAA GG-3', reverse 5'-GAC ATC GTC CTC GGT GTA GG-3'). Primers were validated by measuring amplification of serial dilutions of cDNA from pleural ganglia, and Ct values were plotted against log[template dilution] (Figure B.4). The slope of the resulting linear fits was used to calculate the amplification efficiency (E) of each primer set, using the relationship $E = 10^{-1/\text{slope}}$. Primers were validated at 160 nM, with E values within 10% of each other. The relative *Pleurin* expression was calculated using $2^{-\Delta C_t}$, where ΔC_t is the difference between the Ct values (mean of technical triplicate) for *Pleurin* versus

GAPDH for the same template sample. For Figure B.3, bars represent the mean $\pm$ standard deviation for three biological sets of three animals. Graphs were prepared using GraphPad Prism 7.

TIMS MS on the Ala12-Plrn2 Analogue

Separate solutions of Ala12-Plrn2 (SFYTTGNGNHYARI-NH₂) and ¹³C-L-Plrn2 (SFYTTGNGNHYPRI-NH₂) were prepared at concentrations of 5 μ M in 20 mM ammonium acetate. The solutions were analyzed on a timsTOF Pro mass spectrometer (Bruker Daltonics) by direct infusion at 1 μ L/min. MS was operated in positive mode. The TIMS parameters, including ramp time and mobility range, were the same as described in Chapter 3 and Chapter 4.

B.4 RESULTS AND DISCUSSION

Retention Time Comparison of L-, D2-, D3- and D2D3- Plrn1-Plrn3

The retention times of the synthetic L, D2-, D3- and D2D3- Plrn1-Plrn3 were compared and the results are shown in Figure B.1. The synthetic D2D3-Plrn1 coeluted with synthetic D3-Plrn1. Synthetic D2D3-Plrn2 eluted later than but, was not fully resolved from L-Plrn2, but did not coelute with D2 or D3-Plrn2. Isomeric peptides at these relative retention times were not observed in the results in Chapter 3, this indicated that D2D3 forms of Plrn1 and Plrn2 were not detected in those experiments. Finally, the retention time of the synthetic D2D3-Plrn3 peptide was compared to the retention times of the synthetic L-, D2- and D3-Plrn3 peptides, and it was found that D2D3-Plrn3 elutes later than but was not fully resolved from L-Plrn3.

In Chapter 3, chromatograms from APM digestions on Pleural ganglia extracts showed that Plrn3 in pleural ganglia has four peaks, and two were resistant to aminopeptidase M digestion. A

smaller APM resistant peak was observed at slightly later retention time than the L-Plrn3 peptide, and the peak was observed partially coeluting with L-Plrn3 in the control. The synthetic D2D3-Plrn3 eluted at a similar retention time relative to the synthetic L-Plrn3 as the small, earlier eluting aminopeptidase M resistant peak in the pleural ganglia. This suggests D2D3-Plrn3 may be present at low levels in pleural ganglia extracts. However, no spiking experiments were performed to test whether D2D3-Plrn3 coelutes with the minor APM resistant peak in the pleural ganglion.

Culture of Pleurin Neurons with Colchicine

Table 1 summarizes the results from cell culture experiments. In N = 6 measurements on pleurin neurons incubated with colchicine, negligible quantities of D2- or D3-Plrn3 were detected. Only the L-Plrn2 could be detected. Several possible factors may have prevented this experiment from working as intended, including the possibility that colchicine was old, or and may not have been used at a high enough level. Colchicine could have other biological effects that indirectly affect isomerization. Ultimately, the culture was repeated twice, and both times yielded the same result: negligible quantities of D2- or D3-Plrn2 were detected by CE-TIMS-MS in colchicine cultured neurons. Future experiments may explore alternative vesicle transport inhibitors.

qPCR on A. californica Ganglia

Figure B.3 shows the qPCR expression of *pleurin* mRNA in the various ganglia. qPCR showed that *pleurin* mRNA is negligible in the buccal ganglion, whereas the highest expression of *pleurin* mRNA is in the pleural ganglia, followed by the abdominal ganglion. Very low levels of mRNA were detected in the cerebral and pedal ganglia. Negligible quantities of *pleurin* mRNA were detected in the buccal ganglion, and this result is not consistent with the hypothesis that a

unique *pleurin* expressing cell population exists in the buccal ganglion that exclusively produces the D3-Plrn2/Plrn3 epimers. Therefore, a unique pleurin-containing cell population in the buccal ganglion does not explain the unusual distribution of Plrn2 and Plrn3 diastereomers observed in Chapter 3. Notably, this experiment does not disprove the broader hypothesis of cell-type specific isomerization. It only demonstrates that a *pleurin* expressing cell population is likely not present in the buccal ganglion.

TIMS MS on the Ala12-Plrn2 Analogue

The mobilogram of the Ala12-Plrn2 $[m+2H]^{2+}$ ion showed only one uniform peak by TIMS MS (Figure B.5A). Conversely, the ^{13}C -L-Plrn2 $[m+2H]^{2+}$ ion showed two distinct peaks by TIMS MS (Figure B.5B). Substitution of the Pro12 for an Ala residue appears to have eliminated the ability of this peptide to adopt two stable gas phase conformations during TIMS MS. This is good evidence indicating *cis-trans* proline isomerization explains why native Plrn2 $[m+2H]^{2+}$ ions have two gas phase conformations during TIMS MS.

B.5 FIGURES AND DATA

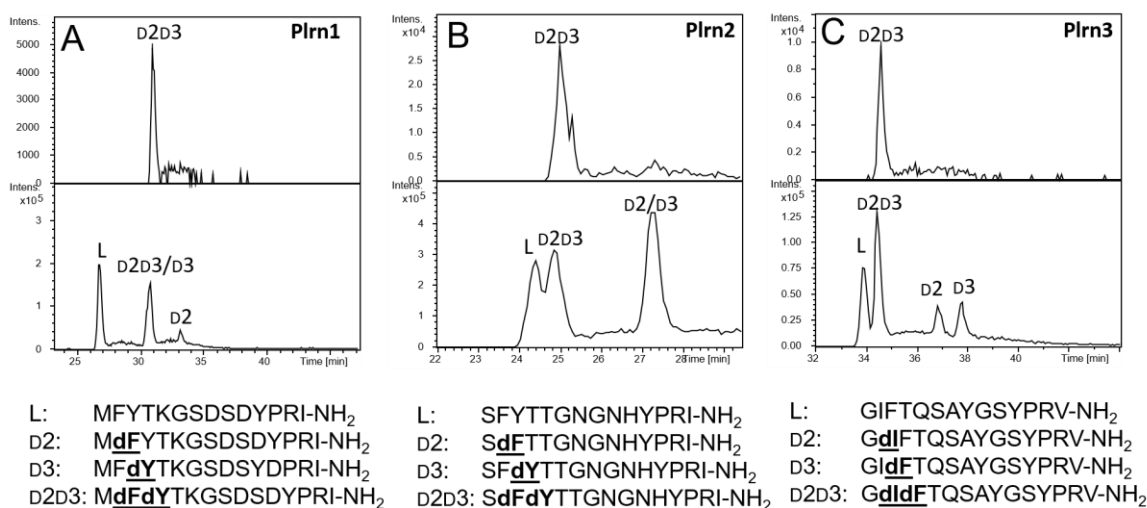


Figure B.1. Retention time comparison of synthetic ^{13}C -Plrn1-Plrn3 diastereomers. (A) ^{13}C -Plrn1 ($[\text{m}^*+3\text{H}]^{3+}$), (B) ^{13}C -Plrn2 ($[\text{m}^*+3\text{H}]^{3+}$) and (C) ^{13}C -Plrn3 ($[\text{m}^*+2\text{H}]^{2+}$) diastereomers. Upper panels show the LC-MS/MS extracted ion chromatograms (EICs) of pure ^{13}C -D2D3-Plrn1, ^{13}C -D2D3-Plrn2, ^{13}C -D2D3-Plrn3. The ^{13}C -D2D3-Plrn1 was synthesized with one ^{13}C -labeled Gly residue ($\text{m}^* = \text{m}+1$ for L- and D2- ^{13}C -Plrn1-Plrn3, and ^{13}C -D3-Plrn1). The ^{13}C -D2D3-Plrn2 and D2D3- ^{13}C -Plrn3 were synthesized with two ^{13}C -labeled Gly residues ($\text{m}^* = \text{m}+2$ for D2D3- and D3- ^{13}C -Plrn2 and ^{13}C -Plrn3). Lower panels show the EICs of the combined diastereomers.

Table B.1. Cell isolation success rates for isolation of Pleurin containing neurons. Table reports the descriptions of the cells that were isolated, and if Plrn2 was detected by MALDI MS and CE-TIMS-MS. A “+” indicates the peptide in the column description was detected, while a “-” indicates the peptide in the column description was not detected. An “x” indicates the sample was not tested by the technique. N is the number of cells tested by MALDI MS, M is the number of cells tested by CE-TIMS MS.

Tube	Description of cell isolated	Plrn2 detected MALDI MS (+/-)	L-Plrn2 CE-TIMS-MS (+/-)	D2-Plrn2 CE-TIMS-MS (+/-)
	Cholchicine Animal #1	N = 7	M = 4	M = 4
42	Putative Pleurin Neuron RPI	-	x	x
43	Putative Pleurin Neuron RPI ^a	+	-	-
44	Putative Pleurin Neuron RPI	+	+	-
45	Putative Pleurin Neuron RPI	+	+	-
46	Putative Pleurin Neuron RPI	+	+	-
47	Putative Pleurin Neuron RPI	-	x	x
48	Putative Pleurin Neuron RPI	-	x	x
49	Blank	-	-	-
	Success Rates	<u>4</u> /7	3/ <u>4</u> (3/4)	0
	Control Animal #1	N = 6	M = 4	M = 4
50	Putative Pleurin Neuron LPI ^a	+	+	-
51	Putative Pleurin Neuron LPI	+	-	-
52	Putative Pleurin Neuron LPI	-	x	x
53	Putative Pleurin Neuron LPI	+	-	-
54	Putative Pleurin Neuron LPI	+	+	-
55	Putative Pleurin Neuron LPI ^a	-	x	x
	Success Rates	<u>4</u> /6	2/ <u>4</u> (2/4)	0
	Cholchicine Animal #2	N = 5	M = 5	M = 5
56	Putative Pleurin Neuron LPI	+	+	-
57	Putative Pleurin Neuron LPI	+	+	-
58	Putative Pleurin Neuron LPI	+	-	-
59	Putative Pleurin Neuron LPI ^a	+	-	-
60	Putative Pleurin Neuron LPI	+	+	-
	Success Rates	<u>5</u> /5	3/ <u>5</u> (3/5)	0
	Control Animal #2	N = 7	M = 7	M = 7
61	Putative Pleurin Neuron RPI	-	-	-
62	Putative Pleurin Neuron RPI	-	-	-
63*	Putative Pleurin Neuron RPI ^b	+	+	-
64	Putative Pleurin Neuron RPI	+	+	-
65*	Putative Pleurin Neuron RPI	+	+	-
66	Putative Pleurin Neuron RPI	-	-	-
67	Putative Pleurin Neuron RPI	+	+	-
	Success Rates	<u>4</u> /7	4/ <u>4</u> (4/7)	0

^a Cell burst during isolation.

^b Problems removing ASW, may have high salt content.

*low signal to noise detection not confident.

```

>NM_001204725.1 Aplysia californica pleurin precursor
(LOC100533448), mRNA
GAAGTGAGTGAAATCAGTCCACTCGCCGCACCAGTTTCTGCCACCATTCCCCTGAGAAAAGTCACGTGA
CGCTTGACTGCTGATCCGGACAAATACAGGCTAGTGGTACGGACATTCACGAAACGCCTTGGTACCCCT
GAGAAGTGAACTCACCAACATAAGTCAACTCCCACCAAATAAACCGCTCAAGATGCAGTCTCACGTGA
GCAGCAGCCTACTGGCCCTGTGCGTCCTGACGGGCAGCGCATGCGCAATGTTCTACACCAAAGGCAGC
GACAGCGACTACCCCCGGATCGGGCGTCGGAGCTTCTACACAACCGGAAACGGAAACCATTTACCCGCG
AATAGGACGACGGGATTCCCCTGGCGCAGCTGCTGGCACGGACATTCTTTACCCGAACACGGCTTCAG
GGTTTTCCGATTTCTCGGCAGGCGCTCCGGCGGTGAAGAGAGGGATTTTCACGCAGTCGGCGTACGGT
TCTTACCCGAGGGTAGGACGAAGGAGTCAGGCGGGAGATGATGGTTCCAGGAAGGTCTTCAGACTGC
GAGATCTAAGATTTTCCAGAAGATTACGGGGCTGGACGGGCCGGTGGAAGAAGAGCTCGGGGATGGAT
CGGCCGAGAGATTGGAAGCTGACAACCAGATTGGGCTTGAGCTCATGTTCCCTCGCCTTTGACCTAGAT
GGTGATCAAGCTTTGTCAAAATCAGAGTTCCAGTCCGGGATGGAAAAGTTCCGACGACACAACCCACT
TTGCTAACGACAGCACACGTTCTGATGACGTCACTGAGCATCTAAAGAGGGACAAGACAGCCAATAG
ATCTACTCTCGAGGAAAAAACTCTATGTGATTTTGGCACACTCAGAATCTGCTCAGTATAGAATGGC
GACACATTGGTTATCAAAATGGTTATGAATGCCCTTTACGGGCTGACCTCTTTATCCCTAACAGATTGT
GATTTTTCAAGTATTTCTGTTTACTATCTTGCTTTTTTGAAATAATATACAATGCCTTAATTGCAATGA
TTTTGTGGGGACGCTGAGCCATAAAATAACGACCATCACTTAAAAAAA

```

Figure B.2. *Pleurin* mRNA sequence from the NCBI database. Highlighted in orange and blue indicate the regions where the primers targeted. The underlined region was amplified by qPCR.

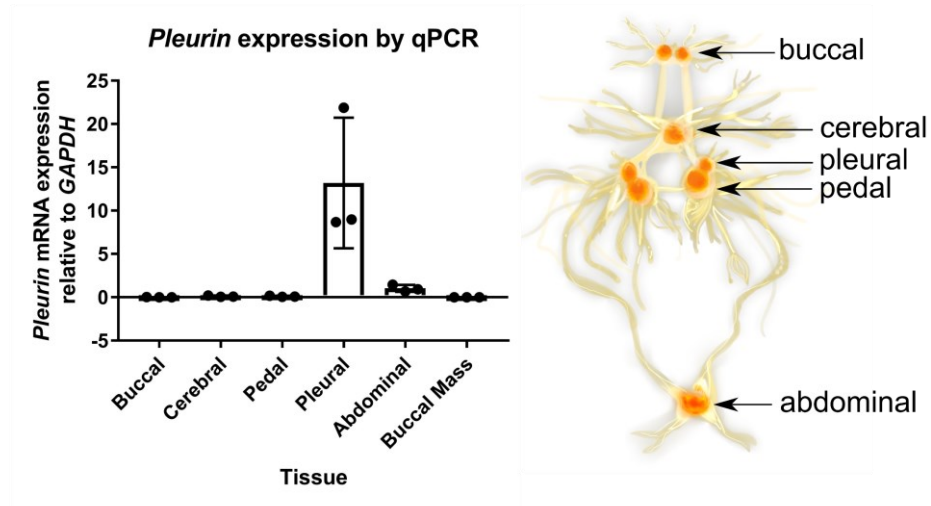


Figure B.3. (Left) *Pleurin* mRNA expression in different regions of the *A. californica* CNS relative to *GAPDH*. Data is from N = 3 biological replicates where each biological replicate comprised the ganglia from 5 animals. (Right) A cartoon depiction of the *A. californica* CNS (not drawn to scale) with each ganglion labeled.

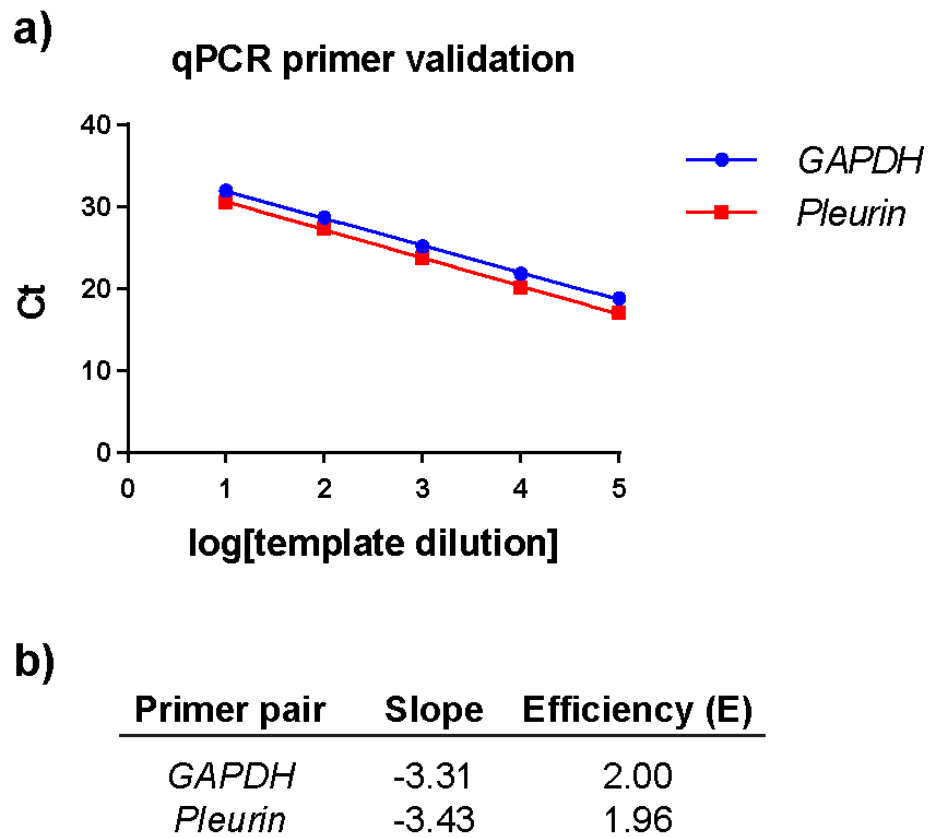


Figure B.4. Validation of qPCR primers. (a) Dilution series of cDNA from pleural ganglia, amplified with 160 nM each primer. (b) Experimental slope and efficiency values.

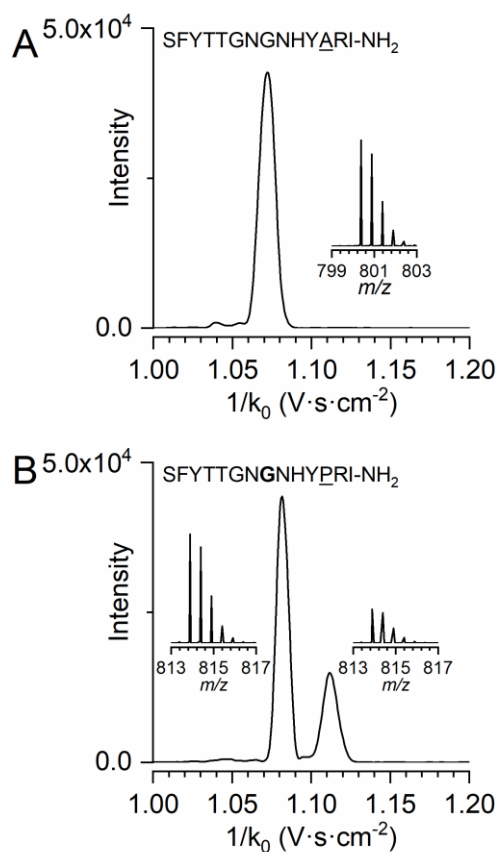


Figure B.5. TIMS MS extracted ion mobilograms (EIMs) of synthetic all-L residue (A) Ala12-L-Pln2 ($[m+2H]^{2+} = 800.4$) and (B) ¹³C-L-Pln2 ($[m+2H]^{2+} = 813.9$). The MS1 profiles corresponding to the peaks are inset in Panel A and Panel B. Sequences are displayed above each Panel. The Ala residue in sequence displayed above Panel A is underlined to emphasize the site of the substitution. The Pro residue is underlined in the sequence displayed above Panel B for emphasis. The Gly residue in the sequence displayed in Panel B is in bold font to indicate it is a ¹³C-labeled Gly residue.

B.6 REFERENCES

1. Checco, J. W.; Zhang, G.; Yuan, W. D.; Le, Z. W.; Jing, J.; Sweedler, J. V., Aplysia allatotropin-related peptide and its newly identified d-amino acid-containing epimer both activate a receptor and a neuronal target. *J Biol Chem* **2018**, *293* (43), 16862-16873.
2. Mast, D. H.; Checco, J. W.; Sweedler, J. V., Differential Post-Translational Amino Acid Isomerization Found among Neuropeptides in *Aplysia californica*. *ACS Chem Biol* **2020**, *15* (1), 272-281.
3. Mast, D. H.; Liao, H. W.; Romanova, E. V.; Sweedler, J. V., Analysis of Peptide Stereochemistry in Single Cells by Capillary Electrophoresis-Trapped Ion Mobility Spectrometry Mass Spectrometry. *Anal Chem* **2021**, *93* (15), 6205-6213.