Published in final edited form as:

J Proteome Res. 2015 November 6; 14(11): 4776–4791. doi:10.1021/acs.jproteome.5b00627.

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

Hui Ye^{†,‡}, Jingxin Wang[§], Zichuan Zhang[‡], Chenxi Jia[‡], Claire Schmerberg[‡], Adam D. Catherman[¶], Paul M. Thomas[¶], Neil L. Kelleher[¶], and Lingjun Li^{*,‡,§,⊥,||}

[†]State Key Laboratory of Natural Medicines, Key Laboratory of Drug Metabolism and Pharmacokinetics, China Pharmaceutical University, Nanjing 210009, China

[‡]School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin 53705, United States

[§]Neuroscience Training Program, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

[⊥]Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

[¶]Departments of Chemistry and Molecular Biosciences, Proteomics Center of Excellence and Chemistry of Life Processes Institute, Northwestern University, 2145 North Sheridan Road, Evanston, Illinois 60208, United States

^{||}School of Life Sciences, Tianjin University, No. 92 Weijin Road, Nankai District, Tianjin 300072, China

Abstract

Decapod crustaceans are important animal models for neurobiologists due to their relatively simple nervous systems with well-defined neural circuits and extensive neuromodulation by a diverse set of signaling peptides. However, biochemical characterization of these endogenous neuropeptides is often challenging due to limited sequence information about these neuropeptide genes and the encoded preprohormones. By taking advantage of sequence homology in neuropeptides observed in related species using a home-built crustacean neuropeptide database, we developed a semi-automated sequencing strategy to characterize the neuropeptidome of *Panulirus interruptus*, an important aquaculture species, with few known neuropeptide

*Corresponding Author lingjun.li@wisc.edu. Phone: +1-608-265-8491. Fax: +1-608-262-5345..

The authors declare no competing financial interest.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b00627.

Schematic workflow of the multifaceted MS-based platform for comprehensive neuropeptidome analysis; profiling of *P. interruptus* brain extract on MALDI-TOF/TOF (PDF)

HRAM MS/MS assignments for selected peptides (PDF) Selected HRAM MS/MS spectra for lobster *P. Interruptus* neuropeptides (PDF)

Selected QTOF MS/MS spectra for de novo sequencing of neuropeptides using PepSeq software (PDF)

Detailed MS/MS spectra information of identified neuropeptides (XLS)

preprohormone sequences. Our streamlined process searched the high mass accuracy and high-resolution data acquired on a LTQ-Orbitrap with a flexible algorithm in ProSight that allows for sequence discrepancy from reported sequences in our database, resulting in the detection of 32 neuropeptides, including 19 novel ones. We further improved the overall coverage to 51 neuropeptides with our multidimensional platform that employed multiple analytical techniques including dimethylation-assisted fragmentation, de novo sequencing using nanoliquid chromatography-electrospray ionization-quadrupole-time-of-flight (nanoLC-ESI-Q-TOF), direct tissue analysis, and mass spectrometry imaging on matrix-assisted laser desorption/ionization (MALDI)-TOF/TOF. The high discovery rate from this unsequenced model organism demonstrated the utility of our neuropeptide discovery pipeline and highlighted the advantage of utilizing multiple sequencing strategies. Collectively, our study expands the catalog of crustacean neuropeptides and more importantly presents an approach that can be adapted to exploring neuropeptidome from species that possess limited sequence information.

Keywords

neuropeptide; *Panulirus interruptus*; de novo sequencing; ProSight; high-resolution and high-accuracy mass spectrometry; dimethylation-assisted fragmentation; mass spectrometric imaging; peptidomics

INTRODUCTION

The activities of neural circuits are extensively modulated by cell–cell signaling molecules. Neuropeptides represent the largest and most diverse group of signaling molecules in nervous systems and affect a wide range of physiological processes like hunger, pain, stress, and reproduction.^{1–3} To understand the underlying mechanisms of such actions, an initial step is to comprehensively characterize the peptide neuromodulators present in the system. Biologically active neuropeptides are produced by initial cleavage from large protein precursors followed by multiple subsequent enzymatic processing steps in various orders, resulting in peptides with a diverse array of physical and chemical properties.^{4,5} The capability to unveil novel neuropeptides could enable a more advanced understanding of peptidergic signaling in the nervous system. Nonetheless, the detection and characterization of these fully processed endogenous neuropeptides remain to be challenging due to their low abundance, chemical diversity, and the lack of knowledge of processing enzymes involved during these steps.

Historically, Edman degradation has served as the predominant technique to sequence novel neuropeptides by sequential removal of amino acids from the N-terminus.⁶ However, this method suffers from low throughput and large sample amounts required for such analysis. In recent years, mass spectrometry (MS) has revolutionized the analysis of neuropeptides.^{6,7} MS allows for accurate measurements of the molecular weights of peptide precursor ions and derivation of amino acid sequence via gas-phase fragmentation in tandem MS (MS/MS) mode. The coupling of liquid chromatography (LC) to MS/MS has permitted analysis of complex samples including neuronal tissue homogenates. Subsequently, the patterns of peptide fragment ions acquired in MS/MS spectra are matched to theoretical peptide fragmentation spectra, which are predicted from preprohormone sequences arising from a cDNA library or a sequenced genome. Therefore, this workflow relies heavily on knowledge of the cleavage characteristics of the enzymes that process preprohormones to neuropeptides and, more crucially, the preprohormone sequences. To date, the workflow of LC–MS/MS analysis in conjunction with subsequent database search has become prevalent in neuropeptidomics,^{8,9} a thriving field that is devoted to the characterization of neuropeptides expressed in an organism, by precisely describing hundreds of diverse forms of neuropeptides from various species including rat,^{10–14} mouse,^{15–17} tree shrew,¹⁸ invertebrates including the house cricket *Acheta domesticus*,¹⁹ the American cockroach *Periplaneta americana*,²⁰ the fruit fly *Drosophila melanogaster*,²¹ the crabs *Cancer borealis*,^{9,22} *Callinectes sapidus*,²³ *Ocypode ceratophthalma*,²⁴ *Rhynchonelliform brachiopods*,²⁵ the giant river prawn *Macrobrachium rosenbergii*,²⁶ and the nematodes *Caenorhabditis elegans*²⁷ and *Ascaris suum*,²⁸ among others.⁷

However, applying this database searching pipeline to identify neuropeptides from MS/MS spectra is not always straightforward when the species of interest have few or virtually no preprohormone sequences available. For example, decapod crustaceans have served as important animal models for neurobiologists due to their relatively simple, well-characterized nervous system and intriguing behavior patterns mediated by various neuropeptides.^{29,5} Nevertheless, complete genome sequencing has not been attempted for these decapod crustacean species so far. Therefore, previously reported crustacean neuropeptide sequences were mostly obtained by interpreting MS/MS spectra acquired from crustacean neuronal tissue extract without known preprohormone sequences, which is termed de novo sequencing.^{30,31} This process is more challenging than assigning an experimental spectrum to a predicted peptide since a spectrum of rich fragment ions at high mass accuracy is a prerequisite to “read off” the sequence. Despite the challenges, researchers in this field have successfully exploited the LC–MS/MS method to identify hundreds of neuropeptides via de novo sequencing from the order Decapoda,^{29–35} and previously reported neuropeptides and cDNA sequences that encoded predicted or known crustacean neuropeptides (UniProt) have been compiled to assemble a crustacean neuropeptide (CNP) database (<https://uwmadison.box.com/lilabNP>). Interestingly, a significant degree of neuropeptide sequence conservation among species is observed within the CNP database and other publicly available neuropeptide databases such as Neuropedia³² and Neuropep.³³ A number of identical or similar motifs in neuropeptide families, such as tachykinin-related peptide (TRP), orcokinin, and SIFamide, are shared among different crustacean species. We thereby developed a semiautomated de novo

sequencing strategy by searching the high-resolution accurate mass (HRAM) LC–MS/MS data obtained from the brain extract of the spiny lobster *Panulirus interruptus* (*P. interruptus*), a decapod crustacean species commonly used for neurobiology and physiology studies,^{34–36} against the compiled CNP database. In this study, we coupled nanoLC to a hybrid LTQ–Orbitrap via electrospray ionization (ESI) for the acquisition of LC–MS/MS data and employed a database searching software designed for neuropeptide discovery, Pro-Sight,³⁷ to accommodate potential sequence variability and potential post-translational modifications (PTMs) present in *P. interruptus*. The combination of HRAM MS data and rich MS/MS fragments, in conjunction with the flexible scoring/searching software system that translates MS/MS spectra into peptide identifications, results in highly confident identification of neuropeptides from the complex *P. interruptus* brain homogenate with minimal manual intervention as compared to another well-recognized de novo sequencing software PEAKS.³⁸

We further improved the comprehensive neuropeptide discovery workflow by employing other sample preparation techniques like N-terminal dimethylation-assisted fragmentation (DAF) and direct tissue analysis, and multiple instrument platforms including ESI–quadrupole-time-of-flight (Q-TOF) and matrix-assisted laser desorption/ionization (MALDI) time-of-flight (TOF)/TOF to increase the coverage of *P. interruptus* neuropeptidome. Using this multifaceted workflow, we discovered 51 endogenous neuropeptides from the *P. interruptus* brain for the first time. Besides 20 neuropeptides that have been reported previously in other crustacean species *C. borealis* and *C. sapidus* yet new to *P. interruptus*, 31 novel neuropeptides belonging to nine neuropeptide families were sequenced for the first time with high confidence. This new information displays both the diversity and conservation of crustacean neuropeptide sequences and provides an important link to their potential biological functions. The inclusion of mass spectrometric imaging (MSI) to the workflow adds another dimension of information to our investigation of the *P. interruptus* brain neuropeptidome by providing spatial distribution and colocalization patterns of these endogenous signaling molecules. Overall, the multifaceted workflow highlighted in this study provides a basis for successful discovery of endogenous neuropeptides from *P. interruptus* as well as a universal approach that can be adapted to exploring neuropeptidomes from crustacean and other species that have only limited sequence information.

EXPERIMENTAL SECTION

Chemicals and Materials

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

Animals and Dissection

California spiny lobsters were purchased from Catalina Offshore Products (San Diego, CA) and maintained in a circulating artificial seawater tank at 14–16 °C before use. Lobsters

were anesthetized by packing them in ice for 30 min, after which the dorsal carapace was removed from each individual, and its supraesophageal ganglion (brain) were dissected free from surrounding muscle and connective tissues. The details of dissection were described elsewhere.³¹ Following dissection, brain samples were immediately placed in acidified MeOH (90% MeOH:9% glacial acetic acid:1% deionized water) and stored at -80°C until utilized for peptide extraction.

Tissue Extraction

Brain was homogenized and extracted with acidified MeOH buffer that was also used for tissue storage. The resulting extract was then concentrated in a Savant SC 110 Speedvac concentrator (Thermo Electron Corporation, West Palm Beach, FL, USA) and resuspended in 0.1% formic acid. The resuspended extracts were then vortexed and briefly centrifuged. The resulting solution was purified and concentrated with C₁₈ ZipTip (Millipore, Billerica, MA, USA). Briefly, the C₁₈ ZipTip was first wetted using ACN and then pre-equilibrated for sample binding with 0.1% formic acid in water. Subsequently, the tissue extract was loaded on the C₁₈ ZipTip. After being rinsed three times with 0.1% formic acid in water, the sample was eluted with 5 μL of ACN/water/formic acid solution (50:49.9:0.1; vol/vol/vol). Next, the eluent was dried and resuspended in 10 μL of 0.1% formic acid in water and subjected to future LC–MS/MS and MALDI-MS analysis. One brain was used per MS run in this study thanks to the high sensitivity offered by the state-of-the-art MS-based platform. Nevertheless, multiple brains can be pooled and subjected to analysis based on specific instruments the users employ.

N-Terminal Dimethylation via Formaldehyde Labeling

For formaldehyde labeling, the neuropeptide extract was resuspended in 5 μL of water followed by the addition of 1.5 μL of borane pyridine (120 mM in MeOH) and 1.5 μL of formaldehyde (4% in H₂O). The labeling mixture was then incubated in a 37°C water bath for 15 min. Excess formaldehyde was quenched via the addition of 3 μL of ammonium bicarbonate buffer (0.2 M). The resulting solution was dried and resuspended in 10 μL of 0.1% formic acid in water and subjected to future LC–MS/MS analysis.

Tissue Preparation for MALDI-MS

Direct tissue analysis was performed as reported previously.³¹ Briefly, the accessory lobe (AcN) and olfactory lobe (ON) were designated as region 1 (R1) and region 2 (R2)³⁹ and dissected out followed by a brief rinse in acidified MeOH buffer and subsequent desalting in 10 mg/mL aqueous DHB solution.⁴⁰ The processed tissues were then placed on the MALDI target with 0.3 μL of 100 mg/mL DHB matrix solution (50:49.9:0.1 MeOH/water/formic acid, v/v) deposited on top of the tissue, which allows extraction of neuropeptides from tissue prior to forming analyte-doped crystals.

Sample Preparation for MSI

The freshly dissected brain was rinsed briefly in deionized water to eliminate excess salt, embedded in an aqueous solution of gelatin (100 mg/mL), and then snap-frozen in -80°C freezer for further processing. Brain sections were acquired at a thickness of 12 μm on a

cryostat (HM525, Thermo Fisher Scientific, Waltham, MA) at -20°C and thaw-mounted onto indium tin oxide (ITO)-coated conductive glass slides (Delta Technologies, Loveland, CO). The sections were then allowed to dry for 30 min under vacuum. For imaging purposes, 150 mg/mL DHB dissolved in 50% MeOH, 0.1% formic acid (v/v) was sprayed onto brain sections homogeneously using an airbrush (Paasche Airbrush Company, Chicago, IL, USA). Five coats were applied, and the spray duration for each coat was 30 s with 1 min dry time between each cycle.

Online Top-down MS/MS on Nano-LC–ESI-LTQ-Orbitrap Elite and Data Analysis

The brain neuropeptide extract was further analyzed by online top-down MS with an Ultimate 3000 RSLCnano system coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). An aliquot of 1 μL was injected onto a 2 cm, 150 μm i.d. PLRP-S d_p 5 μm , pore size 1000 $\text{\AA}$, trap column. A 10 cm, 75 μm i.d. PLRP-S column was used for separation. The gradient was delivered at 300 nL/min starting at 95% A (0.2% formic acid in water), 5% B (95% acetonitrile and 0.2% formic acid) and rose to 10% B at 7 min, 50% B at 50 min, and 85% B at 58 min. The mass spectrometer was operated in a data-dependent mode, performing higher-energy collision dissociation (HCD)-MS² (scan 1), collision-induced dissociation (CID)-MS² (scan 2), and electron transfer dissociation (ETD)-MS² (scan 3) on each of the top three precursors (selected by intact mass) in a FT-MS precursor scan. Charge state preferences were set as 2+, 3+, and 4+. The specific conditions for the three fragmentation methods are as follows. HCD: isolation width, 15 m/z ; normalized collision energy, 30; activation time, 0.1 ms. CID: isolation width, 15 m/z ; normalized collision energy, 41; activation q , 0.4; activation time, 100 ms. ETD: isolation width, 15 m/z ; normalized collision energy, 35; activation q , 0.25; activation time, 100 ms. Dynamic exclusion was set as 10 s.

Data were deisotoped with Xtract using the cRAWler algorithm (ThermoFisher, Bremen, Germany) and searched with a custom 168-core ProSightPC 3.0 (ThermoFisher, Bremen, Germany) cluster using an iterative search tree. ProSightPC Warehouses were created against our home-built CNP database (<https://uwmadison.box.com/lilabNP>) with all the crustacean neuropeptides previously reported in literature and cDNA sequences filtered in Uniprot at subphylum Crustacea (<http://www.uniprot.org/taxonomy/6657>). The data were first searched in an “absolute mass” method, in which the mass tolerance for precursor ions was set at 200 and 20 Da, and that for fragment ions was 10 ppm with “ m mode” on. A minimum match of five fragment ions was required. Then a “biomarker” search mode was performed, in which the mass tolerances were set at 200 and 20 Da for precursor ions, and fragment ions were set at 10 ppm with “ m mode” on. A minimum match of four fragment ions was required. The semiautomated de novo sequencing was performed by manually reviewing all the peptide hits and modifying the sequences with possible PTMs or amino acids when necessary.

NanoLC–ESI-Q-TOF

A Waters nanoAcquity UPLC system was coupled to a Synapt G2 HDMS mass spectrometer (Waters Corp., Milford, MA) for LC–MS/MS analysis of the brain neuropeptide extract. Chromatographic separations were performed on a Waters BEH 130 $\text{\AA}$

C18 reversed-phase capillary column (150 mm $\times$ 75 μ m, 1.7 μ m). The mobile phases used were: 0.1% formic acid in deionized water (A); 0.1% formic acid in ACN (B). An aliquot of 3 μ L of brain neuropeptide extract was injected and loaded onto the Waters NanoACQ 2G-V/M SymC18 (20 mm $\times$ 180 μ m, 5 μ m) using 99% mobile phase A and 1% mobile phase B at a flow rate of 5 μ L/min for 1 min. Following this, a linear gradient from 5 to 45% mobile phase B at a flow rate of 300 nL/min was performed over 90 min at 35 °C. Data-dependent acquisition (DDA) was employed with three precursors selected for MS/MS at once. The MS scan range was from m/z 300–2000, and the MS/MS scan was from m/z 50–2000. Charge state preferences were set as 2+, 3+, and 4+. Dynamic exclusion was set as 10 s. The MS/MS sequencing was performed with a combination of manual sequencing and PepSeq software (Waters Corp., Milford, MA) to assist in de novo sequencing.

CE Fractionation for MALDI-MS Analysis

The offline CE-MALDI-MS was performed with a home-built apparatus featuring a sheathless membrane-covered joint. Briefly, a 65 cm long fused-silica capillary (75 μ m i.d./190 μ m o.d.) was employed with a cellulose acetate membrane-coated porous joint made 3 cm from the outlet ends of the capillary.⁷ This joint was inserted into a buffer cell filled with 0.5% ammonium acetate at pH 4.9, with the negative electrode connected. About 2 cm of capillary was stretched out of the buffer cell from a small hole with a screw on the bottom. For the inlet end, a 0.6 mL plastic vial was filled with the same CE buffer with positive electrode and capillary inserted. A 10 kV high voltage was supplied by TriSep-2100 HV power supplier from Unimicro Technologies (Pleasanton, CA) for CE separation, while a high voltage of 9 kV was applied for sample loading. The CE flow was then collected on the MALDI plate every 1 min for 25 min.

MALDI-MS Analysis

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

For analysis of CE fractions, each spectrum was summed from 1000 laser shots acquired in positive reflectron mode. Spectrum smoothing and baseline subtraction were performed to process the spectra using flexAnalysis software (Bruker Daltonics, Billerica, MA, USA), and the resulting peaks with $S/N > 3$ were matched with LC-MS/MS assignments within a

mass error of 0.03 Da for neuropeptide identification from resulting spectra of direct tissue and CE fraction analysis. The imaging files were processed, and MS images of neuropeptides were generated using the flexImaging software package (Bruker Daltonics, Billerica, MA, USA).

RESULTS AND DISCUSSION

Comprehensive Neuropeptidomic Analysis of the Brain of *Panulirus interruptus* by a Multidimensional MS-Based Platform

In this study, we employed three commonly used mass spectrometers, LC–ESI–LTQ–Orbitrap, LC–ESI–Q–TOF, and MALDI–TOF/TOF, for endogenous neuropeptide discovery. The results obtained from different mass spectrometers were compared and combined (Table 1) since each instrument offers overlapping yet distinct peptide coverage due to their specific ionization preferences and sequencing capabilities. For instance, although LC–ESI–LTQ–Orbitrap and LC–ESI–Q–TOF share the same ionization platform, the former instrument offers higher ion transmission efficiency and more versatile MS/MS fragmentation techniques that enable a more comprehensive coverage of high-mass peptides and proteins.³¹ Furthermore, HRAM spectra contributed by a high-end instrument like LTQ–Orbitrap are a prerequisite for our semiautomated de novo sequencing strategy as a mass tolerance of 10 ppm was set when querying the observed m/z values of *P. interruptus* neuropeptide precursors and fragments against the CNP database in ProSight. This permitted precise determination of possible PTMs and amino acid substitutions alone and in combination with others. This semiautomated de novo sequencing approach revolutionizes the de novo sequencing process for decapod crustacean species by dramatically decreasing the amount of time required for analysis without sacrificing identification confidence, resulting in discovery of 32 neuropeptides. On the other hand, ESI–Q–TOF provides spectra of decent quality and excels at detection of low-mass immonium ions, resulting in the identification of a complementary set of neuropeptides compared to LTQ–Orbitrap. In addition to the LC–ESI–MS-based platform, the characterization of *P. interruptus* neuropeptides was further facilitated by the MALDI–MS-based approach due to rapid analysis, good sensitivity, and moderate tolerance to salt, affirming the presence of 22 neuropeptides directly from the brain tissue. In total, we characterized and identified 51 neuropeptides, the majority of which were not reported in previous *P. interruptus* studies^{41–44} and 31 of which were sequenced in decapod crustacean species for the first time in this study. In Table 1, neuropeptides detected using different instruments are described, highlighting the combinatorial utilization of multiple powerful mass spectral techniques for the discovery of novel neuropeptides from the spiny lobster *P. interruptus* brain. A schematic workflow illustrating such multidimensional MS-based platform for comprehensive neuropeptidomic analysis is shown in Figure S1 (Supporting Information).

High-Resolution and Accurate Mass (HRAM) MS Measurement and MS/MS Sequencing

Neuropeptide extracts of *P. interruptus* brain were analyzed on an HRAM mass spectrometer, ESI–LTQ Orbitrap, with the top three precursors being subjected to top-down HRAM MS/MS using a complementary suite of fragmentation techniques: higher-energy C-trap dissociation (HCD), collision-induced dissociation (CID), and electron-transfer

dissociation (ETD). The ultrahigh resolution (120 000 in MS¹ and 60 000 in MS² at m/z 400) and high mass accuracy (mass error less than 5 ppm) measurements of precursor ions together with the rich sequence-specific fragment ions produced by the complementary fragmentation techniques were used to search the CNP database within ProSightPC 3.0.

Two types of searches were carried out in ProSight. The first type of search, “biomarker”, is similar to a “no-enzyme” search (employed by other search engines) to find any peptides from subsequences in the CNP database. Figure 1 shows the fragmentation spectra and assignment of neuropeptides derived from a known *P. interruptus* tachykinin related peptide (TRP) preprohormone included in the CNP database. Figure 1, panel a displays a HCD MS/MS spectrum arising from the m/z 1179.866 (4+) ion that can be matched with partial sequence of preproTRP,

DAAAPLNEVDEASANDYPILPDPIAARLYAFRNGNAPVGLAVPLRa, by ProSight (Figure 1b). The accurately measured and completely resolved precursor m/z resulted in a mass error of merely 0.85 ppm from the expected m/z at 1179.867, as shown in Figure 1, panel c. Over 48 fragment ions including 15 b- and y-ion pairs were observed in the representative MS/MS spectrum that could be matched to theoretical fragments of the TRP precursor-related peptide (TPRP), confirming the identity of ions assigned by ProSight search. The combinatorial use of different fragmentation techniques is also highlighted in the study. A fragmentation map showing b- and y-ions acquired in HCD, CID, and ETD MS/MS spectra is shown in Figure 1, panel d, with the illustration of how these fragments were produced in Figure 1, panel e. As expected, the most comprehensive coverage was obtained with the use of all three fragmentation techniques. The LTQ-Orbitrap employed in this study also features wide dynamic range and higher ion transmission efficiency of analytes, allowing for comprehensive coverage of neuropeptides spanning the mass range of detection. In comparison, none of the TPRPs or their truncated forms was observed on LC-ESI-Q-TOF, highlighting the importance of utilizing different instrumentation platforms for improved peptidome coverage.

Besides the capability to detect and fragment relatively large neuropeptides as shown in Figure 1, panel a, the capability of the hybrid LTQ Orbitrap to identify peptides of small size was demonstrated in Figure 1, panel f. Here, extensive fragmentation along the peptide backbone was observed with the use of HCD for the TRP peptide APSGFLGMRa. Therefore, the presence of the TRP, previously proposed as a single, conserved neuropeptide across decapod crustacean species,⁴⁴ was demonstrated in *P. interruptus* brain. Indeed, no other TRP isoforms like TPSGFLGMRa—previously detected in the crab *Cancer borealis*—were seen in this study, consistent with the fact that seven identical copies of APSGFLGMRa are encoded in the *P. interruptus* preproTRP.⁴¹ The *P. interruptus* TRP precursor undergoes several proteolytic processing steps, including cleavage at prohormone convertase processing sites, subsequent removal of the dibasic residues, and the final amidation of the exposed Gly residue at the C-terminus, to the final peptide product.⁴¹ Incomplete enzymatic processing cascade or artifact from the MS approach could possibly explain the detection of the truncated form of TRP, APSGFLGM, and the partially processed intermediate APSGFLGMRG, although the latter could also result from the C-terminal mRNA coding sequence APSGFLGMR that is directly followed by a stop codon.⁴⁴

In addition, a methionine-oxidized form of TRP, APSGFLGM-(Ox)Ra, was also identified, which could be an artifact of the sample preparation steps since no reducing agents were added to the extract of *P. interruptus* brain and methionine residue is prone to oxidation in air. In addition to variants of TRP, five truncated forms of the relatively large TPRP identified in Figure 1, panel a were detected. The presence of multiple forms of the conserved decapod TRP or TPRP is postulated to have important impacts on the physiological functions of the neuropeptides due to their different affinities for their receptors.⁴⁴ All the *P. interruptus* preproTRP-derived peptides identified via ProSight “biomarker” search are underlined with different colors in the deduced TRP precursor sequence (accession number AB 113378 in GeneBank/EMBL/DBJ), as shown in Figure 1, panel g. More TPRPs are predicted to exist than those were identified here, which could be explained by their relatively low abundance, comparably poor ionization efficiency, or insufficient chromatographic separation.⁴⁵ Nevertheless, these newly discovered peptides, especially those without C-terminus amidation, are likely processing intermediates or degradation products and therefore must be examined carefully before they can be assigned as authentic neuropeptides.

Besides TPRP and TRP, the neuropeptide belonging to the AST-A family, PRNYAFGLa, and four RFamides (PSLRLRFa, PSMRLRFa, SMPSLRLRFa, APQRNFLRFa) were identified based on the HRAM MS and rich MS/MS fragment ions produced by the complementary fragmentation techniques using the “biomarker” search. Intriguingly, a recent publication reported the identification of several AST-A and AST-B-type neuropeptides from sinus glands of *P. interruptus*.⁴⁶ This discrepancy of identification could be attributed to the differential distribution of neuropeptides at various neural organs. The relatively low concentration of AST-type neuropeptides, along with the low ionization efficiency of AST-A family neuropeptides, possibly contributed to the lack of detection of the AST family neuropeptides from *P. interruptus* brain in this study.

Semiautomated De Novo Sequencing of Novel *P. interruptus* Neuropeptides

Although the ProSight “biomarker” search of high-quality HRAM MS data against the known neuropeptides or sequenced preprohormones generates a number of neuropeptide hits, the second type of search, the “absolute mass” search, enhances novel neuropeptide discovery rate for *P. interruptus*. This mode consists of two searches that tolerate a 20 and 200 Da mass difference, respectively, between the detected mass of unknown peptides with the theoretical mass of peptides in the CNP database, provided that at least five fragment ions are matched within a mass error of 10 ppm with *m* mode on. This flexible scoring algorithm thus retrieves known neuropeptides from the CNP database that displayed similar fragmentation patterns as the unknown ions, accommodating the sequence variations that arise from amino acid variation or substitution, PTMs, or truncation of the previously known peptide entries. After the “absolute mass” search displayed the matches, the results were taken to Sequence Gazer, a manual peptide/protein characterization tool, to investigate the site of variation including substitution, addition, or deletion of amino acid sequence of the *P. interruptus* neuropeptide compared to the known one present in CNP database. This semiautomated de novo sequencing strategy took advantage of great sequence conservation in decapod neuropeptide isoforms, significantly reducing the labor-intensive manual de novo

sequencing and improving the confidence of identification because a cross-species comparison is also performed. In addition, peptides of large size, such as the TPRP in Figure 1, panel a, and Neuropeptide F, that pose challenges for complete manual de novo sequencing could also be matched with peptides in the CNP database as long as consensus sequences are identified within the query peptide.

A representative example is shown in Figure 2, where the full sequence of decapod calcitonin-like diuretic hormone (CLDH) was first characterized by MS without the knowledge of its cDNA sequence. Figure 2, panel a displays a representative HCD MS/MS spectrum of the *P. interruptus* CLDH detected at the m/z of 986.1756 ($z = 3$). The rich fragment ions produced from this ion and the flexible scoring algorithm employed in ProSight “absolute mass” search led to the resulting match with the American lobster *Homarus americanus* CLDH cDNA sequence of GLDLGLGRGFSGSQAAKHLMGLAAANFAGGPa. When compared with the observed fragmentation pattern, a series of bions, from b_2 – b_{30} with the exception of b_{17} , predicted from the *H. americanus* CLDH sequence were all matched to the dominant peaks in the MS/MS spectrum. Nevertheless, a peak with an m/z difference of 81.533 was observed following the b_{26} ion, as shown in Figure 2, panel b. The zoomed-in view in Figure 2, panel c exhibits the ion's completely resolved isotopic pattern and its monoisotopic mass calculated as 2656.349 Da, delivering an incremental mass of 163.066 Da (corresponding to tyrosine) rather than 147.068 Da (corresponding to phenylalanine) following b_{26} . This m of 7.999 between the observed and predicted doubly charged b_{27} ion could thereby be explained by the substitution of amino acid residue phenylalanine (F) as encoded in *H. americanus* CLDH sequence by tyrosine (Y). Moreover, the consecutively assigned b_{28} , b_{29} , and b_{30} ions as in Figure 2, panel a all show an m/z difference of 7.993, 7.992, and 7.993 from the predicted values, validating the site of amino acid variation as the 27th residue. Therefore, the *P. interruptus* CLDH sequence was deduced as GLDLGLGRGFSG-SQAAKHLMGLAAANYAGGPa, whereas the monoisotopic mass of its b_{27} ion is expected as 2656.349 Da, agreeing well with the observed mass (Figure 2c). This change could correspond to a single nucleotide change in the gene encoding CLDH, as tyrosine is encoded by UAU or UAC, and phenylalanine is encoded by UUU or UUC, since such a vital hormone is unlikely to be changed drastically between these two related species. Our result presented here is the first observation of mature CLDH without prior knowledge of the corresponding cDNA sequence, demonstrating the conservation of neuro-peptide sequences in decapods and the robustness of the HRAM MS/MS data paired with the powerful semiautomated de novo sequencing strategy.

CLDH has been previously identified in insects, including *Rhodnius prolixus*,⁴⁷ *Daploptera punctata*,⁴⁸ *Drosophila melanogaster*,⁴⁹ etc., and crustaceans such as *Balanus amphitrite*, *Daphnia pulex*,⁵⁰ and *H. americanus*.⁴² Functional studies of CLDH have shown that this neuropeptide family is implicated in the control of diuresis in insects. CLDH also intrinsically modulates the cardiac neuromuscular system in *H. americanus* and thus has been proven to be bioactive in Decapoda.⁴² Interestingly, a high degree of sequence similarity and a C-terminal amidation are observed in different isoforms of this 31 amino acid peptide hormone in insects and crustaceans, as illustrated in Figure 2, panel d. The

extensive sequence homology among different species supports the postulation that not only CLDHs, but also their receptors are conserved and might play critical roles throughout evolution.

Similarly, two ions were detected and matched with the preproneuropeptide F (NPF)-derived peptides (NPFP) predicted from *Penaeus vannamei* and *Melicertus marginatus* (UniProt: F6KM62 and F6KM63) EST transcripts, SDYPMPSGDALMEASERLLET and SDYPMPSGDALMEASERLLETA,⁵¹ both with a *m* of 7.94 Da. Manual characterization was performed in Sequence Gazer on these two ions, deducing the actually detected peptides' sequences as SDYPLPPGDALMEASERLLET and SDYPLPPGDALMEASERLLETA, both with a substitution of M to L at the fifth position and S to P at the seventh position. Prior to our study, no authentic NPF isoforms or NPFP have been identified or validated from crustacean species using a MS-based platform. The NPF peptides are viewed as the invertebrate homologues of the vertebrate peptide hormone neuropeptide Y (NPY), which is known to be associated with the regulation of appetite and feeding behavior.⁵² It is noteworthy that the administration of vertebrate NPY dramatically increased food intake in invertebrate *Marsupenaeus japonicus* as well,⁵³ suggesting a conserved role for NPY and possibly a degree of sequence homology in crustacean species, most likely in the part of the hormone that interacts with its receptor. Intriguingly, conservation of the NPF sequence in crustaceans is further evidenced by the sequence homology of the truncated NPF and NPFP between *P. interruptus* and two other species, *Penaeus vannamei* and *Melicertus marginatus*, that represent both a derived and a basal taxon. The physiological significance of this conservation awaits discovery and further investigation. It will be interesting to investigate the impact of endogenous NPF molecules on food intake and growth in crustacean species, especially those of economical value to the aquaculture industry. In addition to this potentially important discovery, our work provides a framework for future investigations of NPF and its physiological functions in crustacean species in that it characterizes this important neuropeptide without prior knowledge of its sequence.

Besides the flexible searching algorithm to account for the unknown sequence variation that does not exist in the CNP database, our semiautomated de novo sequencing strategy also enables "multiplexed searches", allowing for identification of multiple peptide precursors using multiplexed MS/MS spectra,⁵⁴ also known as chimeric spectra. A mass window of ± 7.5 Da was chosen to isolate peptide precursors subjected to MS/MS. This significantly wide isolation window improves sensitivity in hybrid LTQ Orbitrap without sacrificing identification confidence since the ProSight multiplexing search option is designed to handle chimeric MS/MS spectra with the aid of HRAM MS/MS data. Figure 3 highlights the identification of two novel neuropeptide isoforms of SIFamide simultaneously from one multiplexed search in ProSight. In Figure 3, panel a, two peptides of *m/z* 465.9091 and 466.5807 with overlapping isotopic distributions and another peptide at *m/z* 471.9124 were observed in a MS precursor scan. Appropriate assignments of their charge state and the resulting monoisotopic masses were automatically done by cRAWler. The resulting deconvoluted masses of 1396.722, 1394.709, and 1412.716 Da were then searched independently in ProSight using all the HRAM MS/MS fragments acquired in the single

ETD MS/MS spectrum, as shown in Figure 3, panel b, all leading to a homologous match with a highly conserved decapod crustacean neuropeptide GYRKPPFNGSIFa.⁵⁵ With manual examination, the fragment ions of the peptide at 1396.722 Da were matched to c₆ to c₁₁ of GYRKPPFNGSIFa. For the z-ions, a consistent increase of 16.00 Da was observed for all of the assigned z-ions, including z₆, z₉, z₁₀, and z₁₁. This suggests an amino acid substitution of F to Y at the 12th position, and thus the deduced sequence is GYRKPPFNGSIYa as in Figure 3, panel c. Similarly, the other SIFamide isoform of 1412.716 Da is suggested to have another Y at the seventh position and sequenced as GYRKPPYNGSIYa due to the observation of a constant *m* 15.994 Da between several pairs of fragment ions from c₇–c₁₁ but an overlapping c₆ when compared to the fragment ions of GYRKPPFNGSIYa. In addition, the mass difference of 15.994 Da between all the observed z-ions including z₆, z₉, z₁₀, and z₁₁ when compared to GYRKPPFNGSIYa further validates our deduced sequence based on c-ions. The peptide with a mass of 1394.709 Da could be explained as an intermediate with loss of water from GYRKPPYNGS(dehydrated)IYa based on the observation of c₁₁, z₉, z₁₀, and z₁₁ ions. This work therefore demonstrates that multiplexing fragmentation and tailored software for chimeric spectra interpretation allow for conclusive identification of two peptides from a single MS/MS isolation. The HRAM MS and MS/MS scans thereby allow one to confidently identify multiple peptides per DDA MS/MS spectrum, increasing the absolute number of peptides identified without a compromise in duty cycle.

Intriguingly, the previously characterized neuropeptide GYRKPPFNGSIFa (Gly-SIFamide) was also identified in *P. interruptus*, agreeing well with its conserved presence in other decapod crustaceans including crabs and crayfish.⁵⁵ It is striking that the substitution of F to Y was observed in the two SIFamides as well as CLDH in *P. interruptus*. Future studies that isolate and sequence the cDNA encoding the SIFamides from this species would help to validate the amino acid substitution observed from *P. interruptus* neuropeptidome using a MS-based approach in this work. In addition to Gly¹-SIFamide, another SIFamide isoform, N-term acetylated (Ac)GYRKPPFNGSIFa, a truncated form RKPPFNGSIFa, and two incompletely processed forms GYRKPPFNGSIF and GYRKPPFNGSIFG were detected in *P. interruptus*. All the other raw spectra and fragmentation maps of identified neuropeptides are compiled in the Supporting Information. Collectively, our strategy not only significantly expanded the catalog of the SIFamide family neuropeptides in decapod crustaceans to those neuropeptides containing C-terminal sequence motif SIYamide, but also laid a foundation for future functional studies of various *P. interruptus* neuropeptide isoforms in this important species.

Discovery of a Novel Motif of the Orcokinin Neuropeptide Family

Although the hybrid LTQ Orbitrap and the tailored database search algorithm enable semiautomated de novo sequencing of 32 neuropeptides, including 19 novel ones, manual de novo sequencing of data acquired on SYNAPT G2 HDMS with the aid of proprietary software PepSeq also generates a list of 34 neuropeptides with 18 novel ones that have never been reported before (all shown in the Supporting Information). Taken together with the novel identifications achieved on LTQ Orbitrap, this work has revealed a novel sequence motif for the decapod crustacean orcokinin neuropeptide family. A myriad of orcokinins

have previously been discovered in crustaceans, insects, and other invertebrate species with a conserved N-terminal motif of NFDEIDR. More specifically, decapod orckokinins share a consensus sequence of NFDEIDRSX(G/S)FGFX(H/N/V/A) (S-orckokinin) as in Figure 4, panel a, with the exception of the truncated Asp-Orckokinins observed in the blue crab *Callinectes sapidus*.⁵⁶ Interpretation of MS/MS spectra produced by orckokinin peptides relied heavily on the y-ion series due to this N-terminal consensus sequence. After sequential loss of the conserved residues from the N-terminus that can be viewed as fingerprints of the orckokinin family, information regarding the variable sequence appearing at the C-terminus is retained in the remaining y-ions. On the basis of accurate mass measurement and the characteristic fingerprints provided by y-ions, we identified nine orckokinin peptides sharing a motif of NFDEIDRAGX(F/L/I/V)X(G/A/V)FX(H/N) (A-orckokinin) as summarized in Figure 4, panel b (L is used in the figure for the purpose of alignment). In Figure 4, panel c, the clean y₂–y₇ ion series made direct reading of the N-terminus sequence NFDEID from the MS/MS spectrum possible. The accurate mass measurement of the y₂ ion determines the C-terminus combination as RA, indicating a novel truncated form of orckokinin, NFDEIDRA. The fingerprints of unique fragmentation patterns belonging to orckokinin family members further facilitated our identification in the representative MS/MS spectra shown in Figure 4, panels d–f. In Figure 4, panel d, the high-abundance y₅–y₁₀ ions immediately reveal its identity as orckokinin with an N-terminus sequence NFDEID, whereas the remaining y₁–y₄ ions determined that the rest of C-terminus variable sequence is RAGL/IA, despite these ions being detected at relatively low abundances. Similarly, in Figure 4, panel e, the y₄–y₁₀ ions again exhibit the fingerprint fragmentation pattern of orckokinins, delivering the conserved N-terminus sequence NFDEIDRA, whereas y₁–y₃ ions aid in deducing the complete sequence as NFDEIDRAGL/IGF. A complete series of y-ions with relatively high abundance starting from y₁–y₁₂ is also displayed in Figure 4, panel f, supporting the identity of this peptide as another orckokinin variant with a sequence of NFDEIDRAGL/IGFH. It is worth noticing that the MS/MS fragmentation pattern of the orckokinin NFDEIDRAGL/IGFH appears to be quite distinct from the orckokinin with a single amino acid truncation at the C-terminus, NFDEIDRAGL/IGF. As shown in Figure 4, panel f, the b-ion series from b₄–b₁₁ were not detected in NFDEIDRAGL/IGFH, although a number of internal fragments were observed and assigned based on the deduced sequence. These internal fragments complicated the MS/MS spectrum, raising the challenge to interpret the fragment ions even with the assistance of the de novo sequencing software PepSeq.

Interestingly, a previous study by Stemmler et al. reported the use of metastable decay and sustained off-resonance irradiation collision-induced dissociation (SORI-CID) to identify orckokinin neuropeptides in crustaceans.⁴³ They identified the orckokinin isoform NFDEIDRAGLGF in *P. interruptus*, which is in agreement with our newly reported motif, and predicted the presence of more orckokinin variants that might show distinct C-terminus variation from previously reported orckokinin peptides.⁴³ To the best of our knowledge, our study with multiple mass spectral techniques presented here describes the most comprehensive repertoire of crustacean orckokinin variants that contain this novel motif, NFDEIDRAG, including NFDEIDRAGLG, NFDEIDRAGVV, NFDEIDRAGFG, NFDEIDRAGFA, a truncated form NFDEIDRAGLa, and the four orckokinins shown in

Figure 4. The specificity of this orcokinin motif to *P. interruptus* is noteworthy, and a plausible explanation is a single codon alteration in orcokinins for this species from UCU/C/A/G (corresponding to S) to GCU/C/A/G (corresponding to A). Simply changing the first nucleotide of the codon encoding the amino acid at that position from a U to a G could lead to substitution of A for S in the final sequence. This may suggest that A-orcokinins have another physiological function different from the S-orcokinins, that is, that *P. interruptus* has developed another pathway in which they are involved in signaling. Alternatively, this could also mean that the A-orcokinins are simply redundant extra copies of the orcokinin gene with a small error that occurred in DNA replication in an ancestor that has been propagated. In this case, the extra set of A-orcokinins might simply amplify the signal of the other orcokinins. This would need further exploration on whether the A-orcokinins have different receptor affinities or different receptors or downstream targets of the receptors. This is expected to generate interesting knowledge regarding the impact of structural variance on neuropeptide biological functions in decapod crustaceans for people who study molecular evolution and speciation.

Dimethylation-Assisted Fragmentation (DAF) and De Novo Sequencing

A reductive N-terminal dimethylation labeling method has previously been shown to facilitate the fragmentation and ultimately de novo sequencing of unknown neuropeptides.^{57,58} A schematic illustration of the reaction with formaldehyde and reducing reagent borane–pyridine complex is shown in Figure 5, panel a. Briefly, the efficient and quick derivatization method labels a peptide's N-terminus and ϵ -amino group of lysine, if any are present, through reductive dimethylation, resulting in an incremental mass of 28.03 Da for each derivatized site.⁵⁸ Moreover, substantial signal enhancement of the a_1 ion after labeling can be utilized to resolve sequence ambiguity in the N-terminus since the b_1 ion for a native peptide is usually missing in a MS/MS spectrum.⁵⁷ In addition, the labeling method has also been employed to facilitate sequencing of singly charged peptides by increasing the signal intensities of the a - and b -ion series and reducing the complexity of fragmentation patterns.⁵⁷ We observed the same simplification effect in the interpretation of native orcokinin NFDEIDRAGL/IGFH with the approach of DAF. As previously shown in Figure 4, panel f, a highly complex MS/MS spectrum was obtained for this peptide. Upon labeling, the mass of the native peptide at m/z 745.85 ($z = +2$) was shifted to m/z 759.86 (Figure 5b), corresponding to the mass difference induced by N-terminus dimethylation. Intriguingly, the form-aldehyde labeled orcokinin displayed dramatically cleaner MS/MS spectrum, containing primarily b - and y -ion series, as shown in Figure 5, panel b. Other than the complete y -ion series from y_1 – y_{12} , the b -ion series were significantly enhanced, resulting in detection of b_4 , b_5 , b_7 , b_8 , b_{10} , b_{11} , and b_{12} ions, all of which were missing in the unlabeled sample. In addition to enhancing b -ions, internal fragments were suppressed, and thus a simplified fragmentation profile was obtained, enabling the identification of this novel orcokinin with remarkable confidence and ease.

Unambiguous identification of amino acid residues at the N-terminus of the sequence could be achieved via the DAF method as well.⁵⁹ As an example, Figure 5, panel c shows the MS/MS sequencing of an RFamide QDL DHVFLRFa, a peptide that has a close mass to KDL DHVFLRFa and is isobaric with AG(or GA)DL DHVFLRFa. A 28.05 Da difference

was detected upon labeling, as indicated by the mass shift from m/z 430.22 to 439.58 ($z = +3$) shown as the inset. This excludes the possible existence of lysine in the peptide, as it too would have been dimethylated for a total mass shift of 56.06 Da. Although the combined mass of two amino acid residues A+G is equal to that of Q (129.10 Da), no a_1 ion corresponding to G (58.07 Da) or A (72.08 Da) was seen even though the dimethylated a_1 ion should be significantly enhanced upon labeling. In contrast, an a_1 ion with m/z corresponding to dimethylated Q was detected with high abundance at m/z 129.10, supporting our assignment of this peptide as N-terminally labeled QDL DHVFLRFa.

N-terminal pyroglutamate cyclization modification is among the most common PTMs in neuropeptides. The reductive dimethylation method aids in the characterization of neuropeptides with N-terminal blockage. Peptides with pyroglutamate modification, as shown in Figure 5, panel d, can be therefore easily assigned based on their inability to react at the N-terminus due to the lack of free N-terminal amine group. After labeling, identical mass and similar retention times were observed compared to the native sample for the peptide at m/z 636.32 in the labeled sample, suggesting the presence of an N-terminal blockage in this peptide. Together with the accurate mass measurement of b- and y-ion series, particularly a_1 , the peptide was confidently assigned as pQDL DHVFLRFa. Interestingly, this pQDL DHVFLRFa is very similar to insect myosuppressin family neuropeptides that possess highly conserved sequences such as pEDVDHVFLRFa and TDVDHVFLRFa. Such high degrees of sequence homology and conservation suggest a concomitantly conserved function of myosuppressins. Since myosuppressins are reported to play a functional role in insects by modulating muscle contraction and regulating food intake, it would be interesting to validate such functions in *P. interruptus* as well.

In addition to facilitating the identification of pyro-Glu/Gln modification due to N-terminal blockage, the application of reductive dimethylation also helps to discern peptide with N-terminal acetylation due to its inability to produce mass shift at the N-terminus. A representative example is the novel peptide sequenced as (Ac)GYRKPPFNGSIFa with a mass of 1422.74 Da. However, it possesses a mass close to the previously reported VYRKPPFNGSIFa, which is found at 1422.78 Da. This ambiguity could be clarified by the observation that it is the only peptide with a single dimethylation at K rather than the expected dimethylation at both N-terminus and K due to the N-terminal blockage. Nevertheless, it is worth noting that this peptide was identified unambiguously without labeling relying on the HRAM data provided by LTQ-Orbitrap.

The observed enhancement and simplification in MS/MS fragmentation after reductive dimethylation could possibly be explained by the adjusted proton affinities (PA) of the fragments.⁶⁰ Upon labeling, the primary amine group at the N-terminus becomes a tertiary amine, possessing higher PA compared to the native form with primary amine. This helps to stabilize the b-ions. Also, the N-terminal dimethylation prevents the formation of cyclic intermediate, thus reducing sequence scrambling effect.⁶¹ Therefore, the abundances of b-ions are generally improved in the resulting MS/MS spectra of labeled peptides compared to the unlabeled one. Moreover, internal fragments with a minimum of one b-type and one y-type bond cleavage are produced with greater frequency when samples are not labeled.⁵⁷ With the increased PA in the N-terminus, the fragmentation of y-type bonds is suppressed,

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

To further improve our detection coverage and strengthen the confidence of our assignments, an offline CE-MALDI-MS platform was utilized to fractionate the brain extract to reduce signal suppression introduced by the large number of peptides present with sample consumption of merely nanoliters. As shown in Figure S2, the sample background was much cleaner upon separation as demonstrated by the representative fractions #8 and #14 in Figure S2a,b compared to the direct profiling of the extract in Figure S2c. All the peptides profiled via CE-MALDI-MS were indicated in Table 1, whereas those also detected by direct tissue analysis were noted as “d”. All the peptides identified from *P. interruptus* brain were reported in this study. Nevertheless, it is worth noting that our results also include several peptides possibly resulting from incomplete processing of neuropeptide precursors such as orckinins and SIFamides. Therefore, these peptides must be examined carefully via electrophysiological investigations before they can be assigned as neuropeptides. We include these peptides in our list to show the power of our platform by successfully characterizing these low-concentration intermediates and also to shed light on the enzymatic processing involved in neuropeptide production.

Regiospecific Localization of Neuropeptides in the Brain of *P. interruptus*

Using our multifaceted MS-based strategy, we have discovered a number of novel neuropeptides from *P. interruptus* that have shared homology yet variations in amino acid residues or PTMs from previously reported neuropeptides or preprohormones. Examples of these changes include the amino acid substitution of S to A in *P. interruptus* orckinins, and F to Y substitutions in *P. interruptus* SIFamide and CLDH. To provide further evidence that the origins of these sequence variation result from genetic differences among species rather than artifacts from sample preparation procedures,^{31,56} we performed direct tissue analysis on *P. interruptus* brain to examine the resulting peptide profiles with minimal sample preparation. As highlighted in Figure 6, panel a, eight orckinins were observed, including peptides that possess the common motif NFDEIDRSG such as NFDEIDRSGFG (m/z 1256.55), NFDEIDRSGFA (m/z 1270.57), NFDEIDRSGFGFNa (m/z 1516.68), and NFDEIDESGFGFH (m/z 1540.68), and the orckinins containing the novel motif NFDEIDRAG such as NFDEIDRAGLG (m/z 1206.57), NFDEIDRAGFG (m/z 1240.56), NFDEIDRAGLGF (m/z 1353.64), and NFDEIDRAGLGFH (m/z 1490.70). Unfortunately, not all of the sequenced orckinins were seen via direct tissue analysis on MALDI-TOF/TOF, probably due to their relatively low abundances and the complex sample matrix associated with the unprocessed tissue leading to analyte suppression. The complexity of this biological matrix is proven by the myriad of lipid peaks shown in Figure 6, panel a. Nevertheless, the observation of these novel orckinins directly from tissue validated their existence in the *P. interruptus* brain and thereby the novel motif we discovered. In addition, the substitution of F to Y exemplified by CLDH at m/z 2956.50 was also detected via direct

tissue analysis of the *P. interruptus* brain in the linear mode, excluding the possibility of the sequence variation due to sample processing procedures.

Using knowledge gained from pioneering studies of the brains of spiny lobsters dating back to the 1960s,⁶² substructures of the *P. interruptus* brain were examined in detail using direct tissue profiling. Crustacean brains consist of complex structures including numerous areas of neuropil and neuronal clusters.³¹ Previously, neuropeptides that displayed distinct distributions have been linked to different biological functions in crustacean species such as *C. sapidus*⁶³ and *C. borealis*.^{30,64} Therefore, it is crucial to study the neuropeptides' localization in *P. interruptus* brain due to the correlations that can often be made between brain location and function. One interesting finding is that neuropeptides desorbed/ionized from region 1 (R1) in Figure 6, panel a exhibited overlapping yet distinct peptide profiles compared to those from region 2 (R2) in Figure 6, panel b. In Figure 6, panel a, orcokinin peptides were detected in high abundance, whereas the SIFamide peptide at m/z 1381.74 was observed at a modest level when compared to the base peak APSGFLGMRa at m/z 934.49 in R1. In contrast, SIFamide became a dominant peak in the profiling spectrum of R2, whereas orcokinins were detected at a low level in this region compared to R1. The differential neuropeptide patterns from these two regions' profiling spectra demonstrate the complexity of the *P. interruptus* brain structure and the potentially distinct regulatory roles neuropeptides play in different regions of the brain.

Mapping Distribution of Various Brain Neuropeptides via MALDI-MSI

Direct tissue analysis yields “snapshots” of neuropeptide profiles corresponding to targeted regions in *P. interruptus* brain, providing information about spatial distribution of neuropeptides in a high throughput manner. Nonetheless, the need to document detailed and accurate localizations of neuropeptides in a complex and intricate structure like brain cannot be satisfied with direct tissue analysis. Alternatively, MSI, a newly emerged technique, has demonstrated its capability to characterize and localize neuropeptides from neural tissue of heterogeneous structure at high spatial resolution.^{65,66,23,7,67} The high-resolution mapping of *P. interruptus* brain neuropeptides is thus accomplished here via MSI with representative neuropeptides from major NP families, tachykinin, SIFamide, orcokinin, AST-A, and RFamide, shown in Figure 7. In Figure 7, panel a, the TRP neuropeptide of m/z 934.49, detected at high abundance in both regions in Figure 6, is shown to be distributed throughout the major regions in brain, including anterior medial protocerebral neuropils (AMPN), accessory lobes (AcN), and olfactory lobes (ON), as illustrated by the rostral view of *Panulirus* brain (Figure 7).³⁹ Its distribution in AcN and ON agrees well with our previous knowledge that decapod TRP has been identified as a neuromodulator in the olfactory neural pathway.⁶⁸ Interestingly, two SIFamide peptides, GYRKPPFNGSIFa at m/z 1381.74 (Figure 7b) and (Ac)GYRKPPFNGSIFa at m/z 1423.75 (Figure 7c), show concentrated localization in the AcN, which agrees well with its significantly higher abundance in R2 compared to R1 as displayed in Figure 6. In contrast, orcokinin peptides at m/z 1254.47 and 1490.70, also revealed by regional profiling in Figure 6, exhibited higher levels of abundance in the ON compared to AcN and AMPN in Figure 7, panels d and e. The colocalization of various neuropeptide isoforms from the same family, particularly in the AcN and ON, and potential links to function warrant further investigation. Moreover, the AST-A peptide ADPYAFGLa

at m/z 852.43 also displays a higher concentration in ON, illustrated in Figure 7, panel f. In addition, the RFamide SMPSLRLRFa at m/z 1105.63 localized mostly in AcN and ON, yet was present in lower quantities in the AMPN in Figure 7, panel g. This observation is also in agreement with its presence in both R1 and R2 profiling spectra.

The MALDI-MSI results successfully mapped neuropeptides of different families in the brain, presenting useful and interesting spatial information about endogenous neuropeptides and providing insights into their regulatory functions in the brain of *P. interruptus*. OL and AcN are dominant deutocerebral neuropils (DCN) and are best visualized in the rostral view. A unique feature of *P. interruptus* brain is that the large AcNs are located medial to the slightly smaller ONs, an organization that is distinct from that previously characterized in the lobster *Homarus americanus*³¹ and other decapod species such as crayfish or crabs.³⁹ OL receives primary olfactory input from olfactory receptor neurons on antennae, whereas AcN not only receives secondary olfactory signals from the OL, but also higher multimodal signals through synaptic contact with neurons derived from tactile and visual sensory systems.⁶⁸ Therefore, the distributions of neuropeptides within DCN suggest their possible role as neuromodulators involved in the function of the olfactory system, the tactile system, or the visual sensory system. An alternative function could be in the integration of multiple sensory modalities. Moreover, the information about colocalization of neuropeptides gained in this study implies a relationship between their functions—those that are found in close proximity to each other may have similar or related roles in cell–cell signaling. Further in-depth study might provide a clear answer regarding the possibly synergistic, complementary, or antagonistic relationships that could be present among these colocalized peptide modulators.

To the best of our knowledge, our study presents the first investigation of the spatial distribution of neuropeptides in the brain of *P. interruptus*. Although MSI generally has a limitation in resolution compared to conventional immunohistochemical methods, the capability to distinguish neuropeptide isoforms, exemplified by SIFamide peptides in Figure 7, panels b and c and orcokinin peptides in Figure 7, panels d and e, in a high-throughput manner enables unambiguous mapping of neuropeptides in large-scale. In addition, the overlay feature enables highly multiplexed study of colocalization patterns of numerous neuropeptides (Figure 7h) and thus may provide information on potential interactions and relationship of their functions.

CONCLUSIONS

In this study, we employed a suite of mass spectrometric approaches for the discovery of neuropeptides in the brain of *P. interruptus*, an important aquaculture species that has not been extensively characterized. Collectively, 51 neuropeptides were sequenced in this work, including 31 novel ones that were de novo sequenced for the first time. The use of HRAM MS/MS data with various fragmentation methods (CID, HCD, and ETD) and tailored ProSight searching against our home-built CNP database highlighted the possibility of streamlining the peptide discovery process with highly confident identifications. Although the hybrid LTQ-Orbitrap excels in the detection of larger peptides, Q-TOF provides complementary coverage in the interpretation of medium-sized neuropeptides. Furthermore,

the application of DAF aids in the de novo sequencing process and improves neuropeptide identification confidence on a medium-resolution instrument. In addition, direct tissue analysis was applied to analyze the brain tissue with minimal sample preparation, better preserving native neuropeptide expression profiles in *P. interruptus* brain. Finally, MALDI-MSI was employed to map the distribution of multiple neuropeptides simultaneously from a brain slice to yield information on localization that may be important in determining the function of these NPs. In summary, our study not only presents comprehensive characterization of neuropeptide expression and distribution in the brain of *P. interruptus* providing insights into physiological functions of these endogenous neuropeptides, but also demonstrates the application of a multidimensional MS-based platform that will be useful for future neuropeptidome discovery studies on crustacean species without the prerequisite of known preprohormones.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors wish to thank Professor Deborah Baro at the Georgia State University for the critical reading and insightful suggestions of the manuscript. This work is supported in part by the National Institutes of Health grants (R01DK071801 and 1S10RR029531 to LL) and the National Science Foundation grant (CHE-1413596 to LL). N.L.K. thanks support from NIH Grant No. R01 GM067193. L.L. acknowledges an H. I. Romnes Faculty Research Fellowship and a Vilas Distinguished Achievement Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy. The authors also wish to thank Bruker Daltonics for graciously loaning the Autoflex III MALDITOF/TOF mass spectrometer.

ABBREVIATIONS

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

REFERENCES

1. Hokfelt T, Bartfai T, Bloom F. Neuropeptides: opportunities for drug discovery. *Lancet Neurol.* 2003; 2:463–472. [PubMed: 12878434]
2. Morton GJ, Cummings DE, Baskin DG, Barsh GS, Schwartz MW. Central nervous system control of food intake and body weight. *Nature.* 2006; 443:289–295. [PubMed: 16988703]
3. Solway B, Bose SC, Corder G, Donahue RR, Taylor BK. Tonic inhibition of chronic pain by neuropeptide Y. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:7224–7229. [PubMed: 21482764]

4. Devi L. Consensus sequence for processing of peptide precursors at monobasic sites. *FEBS Lett.* 1991; 280:189–194. [PubMed: 2013311]
5. Li L, Sweedler JV. Peptides in the brain: mass spectrometry-based measurement approaches and challenges. *Annu. Rev. Anal. Chem.* 2008; 1:451–483.
6. Hummon AB, Amare A, Sweedler JV. Discovering new invertebrate neuropeptides using mass spectrometry. *Mass Spectrom. Rev.* 2006; 25:77–98. [PubMed: 15937922]
7. Zhang ZC, Ye H, Wang JH, Hui LM, Li LJ. Pressure-assisted capillary electrophoresis coupling with matrix-assisted laser desorption/ionization-mass spectrometric imaging for quantitative analysis of complex peptide mixtures. *Anal. Chem.* 2012; 84:7684–7691. [PubMed: 22891936]
8. Podvin S, Bunday R, Toneff T, Ziegler M, Hook V. Profiles of secreted neuropeptides and catecholamines illustrate similarities and differences in response to stimulation by distinct secretagogues. *Mol. Cell. Neurosci.* 2015; 68:177–185. [PubMed: 26092702]
9. Ma MM, Wang JH, Chen RB, Li LJ. Expanding the Crustacean Neuropeptidome Using a Multifaceted Mass Spectrometric Approach. *J. Proteome Res.* 2009; 8:2426–2437. [PubMed: 19222238]
10. Karlsson O, Kultima K, Wadensten H, Nilsson A, Roman E, Andrén PE, Brittebo EB. Neurotoxin-induced neuropeptide perturbations in striatum of neonatal rats. *J. Proteome Res.* 2013; 12:1678–1690. [PubMed: 23410195]
11. Ljungdahl A, Hanrieder J, Falth M, Bergquist J, Andersson M. Imaging mass spectrometry reveals elevated nigral levels of dynorphin neuropeptides in L-DOPA-induced dyskinesia in rat model of Parkinson's disease. *PLoS One.* 2011; 6:e25653. [PubMed: 21984936]
12. Mabrouk OS, Kennedy RT. Simultaneous oxytocin and argvasopressin measurements in microdialysates using capillary liquid chromatography-mass spectrometry. *J. Neurosci. Methods.* 2012; 209:127–133. [PubMed: 22710285]
13. Petruzzello F, Falasca S, Andren PE, Rainer G, Zhang X. Chronic nicotine treatment impacts the regulation of opioid and nonopioid peptides in the rat dorsal striatum. *Mol. Cell. Proteomics.* 2013; 12:1553–1562. [PubMed: 23436905]
14. Romanova EV, Lee JE, Kelleher NL, Sweedler JV, Gulley JM. Comparative peptidomics analysis of neural adaptations in rats repeatedly exposed to amphetamine. *J. Neurochem.* 2012; 123:276–287. [PubMed: 22860605]
15. Guenther S, Römpf A, Kummer W, Spengler B. AP-MALDI imaging of neuropeptides in mouse pituitary gland with 5 µm spatial resolution and high mass accuracy. *Int. J. Mass Spectrom.* 2011; 305:228–237.
16. Wardman JH, Berezniuk I, Di S, Tasker JG, Fricker LD. ProSAAS-derived peptides are colocalized with neuropeptide Y and function as neuropeptides in the regulation of food Intake. *PLoS One.* 2011; 6:e28152. [PubMed: 22164236]
17. Zhang XZ, Petruzzello F, Zani F, Fouillen L, Andrén PE, Solinas G, Rainer G. High identification rates of endogenous neuropeptides from mouse brain. *J. Proteome Res.* 2012; 11:2819–2827. [PubMed: 22424378]
18. Fouillen L, Petruzzello F, Veit J, Bhattacharyya A, Kretz R, Rainer G, Zhang X. Neuropeptide alterations in the tree shrew hypothalamus during volatile anesthesia. *J. Proteomics.* 2013; 80:311–319. [PubMed: 23228960]
19. Verhaert PD, Conaway MCP, Pekar TM, Miller K. Neuropeptide imaging on an LTQ with vMALDI source: The complete 'all-in-one' peptidome analysis. *Int. J. Mass Spectrom.* 2007; 260:177–184.
20. Verhaert PD, Pinkse MW, Strupat K, Conaway MC. Imaging of similar mass neuropeptides in neuronal tissue by enhanced resolution MALDI MS with an ion trap - Orbitrap hybrid instrument. *Methods Mol. Biol.* 2010; 656:433–449. [PubMed: 20680606]
21. Wegener C, Neupert S, Predel R. Direct MALDI-TOF mass spectrometric peptide profiling of neuroendocrine tissue of *Drosophila*. *Methods Mol. Biol.* 2010; 615:117–127. [PubMed: 20013204]
22. DeKeyser SS, Kutz-Naber KK, Schmidt JJ, Barrett-Wilt GA, Li LJ. Imaging mass spectrometry of neuropeptides in decapod crustacean neuronal tissues. *J. Proteome Res.* 2007; 6:1782–1791. [PubMed: 17381149]

23. Ye H, Hui LM, Kellersberger K, Li LJ. Mapping of neuropeptides in the crustacean stomatogastric nervous system by imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.* 2013; 24:134–147. [PubMed: 23192703]
24. Hui L, D'Andrea BT, Jia C, Liang Z, Christie AE, Li L. Mass spectrometric characterization of the neuropeptidome of the ghost crab *Ocypode ceratophthalma* (Brachyura, Ocypodidae). *Gen. Comp. Endocrinol.* 2013; 184:22–34. [PubMed: 23298572]
25. Immel F, Gaspard D, Marie A, Guichard N, Cusack M, Marin F. Shell proteome of rhynchonelliform brachiopods. *J. Struct. Biol.* 2015; 190:360–366. [PubMed: 25896726]
26. Suwansa-Ard S, Thongbuakaew T, Wang T, Zhao M, Elizur A, Hanna PJ, Sretarugsa P, Cummins SF, Sobhon P. In silico Neuropeptidome of Female *Macrobrachium rosenbergii* Based on Transcriptome and Peptide Mining of Eyestalk, Central Nervous System and Ovary. *PLoS One.* 2015; 10:e0123848. [PubMed: 26023789]
27. Husson SJ, Clynen E, Boonen K, Janssen T, Lindemans M, Baggerman G, Schoofs L. Approaches to identify endogenous peptides in the soil nematode *Caenorhabditis elegans*. *Methods Mol. Biol.* 2010; 615:29–47. [PubMed: 20013198]
28. Jarecki JL, Frey BL, Smith LM, Stretton AO. Discovery of neuropeptides in the nematode *Ascaris suum* by database mining and tandem mass spectrometry. *J. Proteome Res.* 2011; 10:3098–3106. [PubMed: 21524146]
29. Christie AE, Stemmler EA, Dickinson PS. Crustacean neuropeptides. *Cell. Mol. Life Sci.* 2010; 67:4135–4169. [PubMed: 20725764]
30. Cape SS, Rehm KJ, Ma M, Marder E, Li LJ. Mass spectral comparison of the neuropeptide complement of the stomatogastric ganglion and brain in the adult and embryonic lobster. *J. Neurochem.* 2008; 105:690–702. [PubMed: 18088365]
31. Chen RB, Jiang XY, Conaway MCP, Mohtashemi I, Hui LM, Viner R, Li LJ. Mass spectral analysis of neuropeptide expression and distribution in the nervous system of the lobster *Homarus americanus*. *J. Proteome Res.* 2010; 9:818–832. [PubMed: 20025296]
32. Kim Y, Bark S, Hook V, Bandeira N. NeuroPedia: neuropeptide database and spectral library. *Bioinformatics.* 2011; 27:2772–2773. [PubMed: 21821666]
33. Wang Y, Wang M, Yin S, Jang R, Wang J, Xue Z, Xu T. NeuroPep: a comprehensive resource of neuropeptides. *Database.* 2015; 2015:bav038. [PubMed: 25931458]
34. Buchberger A, Yu Q, Li L. Advances in mass spectrometric tools for probing neuropeptides. *Annu. Rev. Anal. Chem.* 2015; 8:485–509.
35. OuYang C, Liang Z, Li L. Mass spectrometric analysis of spatio-temporal dynamics of crustacean neuropeptides. *Biochim. Biophys. Acta, Proteins Proteomics.* 2015; 1854:798–811.
36. Rodgers EW, Krenz WD, Jiang X, Li L, Baro DJ. Dopaminergic tone regulates transient potassium current maximal conductance through a translational mechanism requiring D1Rs, cAMP/PKA, Erk and mTOR. *BMC Neurosci.* 2013; 14:143. [PubMed: 24225021]
37. LeDuc RD, Taylor GK, Kim YB, Januszkyk TE, Bynum LH, Sola JV, Garavelli JS, Kelleher NL. ProSight PTM: an integrated environment for protein identification and characterization by top-down mass spectrometry. *Nucleic Acids Res.* 2004; 32:W340–W345. [PubMed: 15215407]
38. Zhang J, Xin L, Shan B, Chen W, Xie M, Yuen D, Zhang W, Zhang Z, Lajoie GA, Ma B. PEAKS DB. de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell. Proteomics.* 2012; 11:M111.010587.
39. Sandeman D, Sandeman R, Derby C, Schmidt M. Morphology of the brain of crayfish, crabs, and spiny lobsters: a common nomenclature for homologous structures. *Biol. Bull.* 1992; 183:304–326.
40. Kutz KK, Schmidt JJ, Li L. In situ tissue analysis of neuropeptides by MALDI FTMS in-cell accumulation. *Anal. Chem.* 2004; 76:5630–5640. [PubMed: 15456280]
41. Christie AE, Cashman CR, Stevens JS, Smith CM, Beale KM, Stemmler EA, Greenwood SJ, Towle DW, Dickinson PS. Identification and cardiotropic actions of brain/gut-derived tachykinin-related peptides (TRPs) from the American lobster *Homarus americanus*. *Peptides.* 2008; 29:1909–1918. [PubMed: 18706463]
42. Christie AE, Stevens JS, Bowers MR, Chapline MC, Jensen DA, Schegg KM, Goldwaser J, Kwiatkowski MA, Pleasant TK, Shoenfeld L, Tempest LK, Williams CR, Wiwatpanit T, Smith

- CM, Beale KM, Towle DW, Schooley DA, Dickinson PS. Identification of a calcitonin-like diuretic hormone that functions as an intrinsic modulator of the American lobster, *Homarus americanus*, cardiac neuromuscular system. *J. Exp. Biol.* 2010; 213:118–127. [PubMed: 20008368]
43. Stemmler EA, Provencher HL, Guiney ME, Gardner NP, Dickinson PS. Matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for the identification of orckinin neuropeptides in crustaceans using metastable decay and sustained off-resonance irradiation. *Anal. Chem.* 2005; 77:3594–3606. [PubMed: 15924394]
 44. Yasuda-Kamatani Y, Yasuda A. APSGFLGMRamide is a unique tachykinin-related peptide in crustaceans. *Eur. J. Biochem.* 2004; 271:1546–1556. [PubMed: 15066180]
 45. Gelman JS, Dasgupta S, Berezniuk I, Fricker LD. Analysis of peptides secreted from cultured mouse brain tissue. *Biochim. Biophys. Acta, Proteins Proteomics.* 2013; 1834:2408–2417.
 46. Jia CX, Lietz CB, Ye H, Hui LM, Yu Q, Yoo S, Li LJ. A multi-scale strategy for discovery of novel endogenous neuropeptides in the crustacean nervous system. *J. Proteomics.* 2013; 91:1–12. [PubMed: 23806756]
 47. Brugge VA, Schooley DA, Orchard I. Amino acid sequence and biological activity of a calcitonin-like diuretic hormone (DH31) from *Rhodnius prolixus*. *J. Exp. Biol.* 2008; 211:382–390. [PubMed: 18203994]
 48. Furuya K, Milchak RJ, Schegg KM, Zhang J, Tobe SS, Coast GM, Schooley DA. Cockroach diuretic hormones: characterization of a calcitonin-like peptide in insects. *Proc. Natl. Acad. Sci. U. S. A.* 2000; 97:6469–6474. [PubMed: 10841553]
 49. Coast GM, Webster SG, Schegg KM, Tobe SS, Schooley DA. The *Drosophila melanogaster* homologue of an insect calcitonin-like diuretic peptide stimulates V-ATPase activity in fruit fly Malpighian tubules. *J. Exp. Biol.* 2001; 204:1795–1804. [PubMed: 11316500]
 50. Gard AL, Lenz PH, Shaw JR, Christie AE. Identification of putative peptide paracrine/hormones in the water flea *Daphnia pulex* (Crustacea; Branchiopoda; Cladocera) using transcriptomics and immunohistochemistry. *Gen. Comp. Endocrinol.* 2009; 160:271–287. [PubMed: 19135444]
 51. Christie AE, Cashman CR, Brennan HR, Ma MM, Sousa GL, Li LJ, Stemmler EA, Dickinson PS. Identification of putative crustacean neuropeptides using in silico analyses of publicly accessible expressed sequence tags. *Gen. Comp. Endocrinol.* 2008; 156:246–264. [PubMed: 18321503]
 52. Xi L, Jin Y, Parker EA, Josh P, Jones A, Wijffels G, Colgrave ML. Challenges in mass spectrometry-based quantification of bioactive peptides: a case study exploring the neuropeptide Y family. *Biopolymers.* 2012; 98:357–366. [PubMed: 23193600]
 53. Kiris IG, Eroldogan OT, Kir M, Kumlu M. Influence of neuropeptide Y (NPY) on food intake and growth of penaeid shrimps *Marsupenaeus japonicus* and *Penaeus semisulcatus* (Decapoda: Penaeidae). *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* 2004; 139:239–244.
 54. Boyne MT, Garcia BA, Li M, Zamdborg L, Wenger CD, Babai S, Kelleher NL. Tandem mass spectrometry with ultrahigh mass accuracy clarifies peptide identification by database retrieval. *J. Proteome Res.* 2009; 8:374–379. [PubMed: 19053528]
 55. Vázquez-Acevedo N, Rivera NM, Torres-Gonzalez AM, Rullan-Matheu Y, Ruíz-Rodríguez EA, Sosa MA. GYRKPPFNGSIFamide (Gly-SIFamide) modulates aggression in the freshwater prawn *Macrobrachium rosenbergii*. *Biol. Bull.* 2009; 217:313–326. [PubMed: 20040755]
 56. Hui L, Cunningham R, Zhang Z, Cao W, Jia C, Li L. Discovery and characterization of the crustacean hyperglycemic hormone precursor related peptides (CPRP) and orckinin neuropeptides in the sinus glands of the blue crab *Callinectes sapidus* using multiple tandem mass spectrometry techniques. *J. Proteome Res.* 2011; 10:4219–4229. [PubMed: 21740068]
 57. Fu Q, Li LJ. De novo sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal. Chem.* 2005; 77:7783–7795. [PubMed: 16316189]
 58. Hsu JL, Huang SY, Shiea JT, Huang WY, Chen SH. Beyond quantitative proteomics: signal enhancement of the a₁ ion as a mass tag for peptide sequencing using dimethyl labeling. *J. Proteome Res.* 2005; 4:101–108. [PubMed: 15707364]
 59. Spit J, Badisco L, Verlinden H, Van Wielendaele P, Zels S, Dillen S, Vanden Broeck J. Peptidergic control of food intake and digestion in insects. *Can. J. Zool.* 2012; 90:489–506.

60. Paizs B, Suhai S. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* 2005; 24:508–548. [PubMed: 15389847]
61. Jia C, Qi W, He Z. Cyclization reaction of peptide fragment ions during multistage collisionally activated decomposition: an inducement to lose internal amino-acid residues. *J. Am. Soc. Mass Spectrom.* 2007; 18:663–678. [PubMed: 17234429]
62. Maynard DM. Integration in crustacean ganglia. *Symp. Soc. Exp. Biol.* 1966; 20:111–149. [PubMed: 5958359]
63. Hui LM, Zhang YZ, Wang JH, Cook A, Ye H, Nusbaum MP, Li LJ. Discovery and functional study of a novel crustacean tachykinin neuropeptide. *ACS Chem. Neurosci.* 2011; 2:711–722. [PubMed: 22247794]
64. Chen RB, Hui LM, Sturm RM, Li LJ. Three dimensional mapping of neuropeptides and lipids in crustacean brain by mass spectral imaging. *J. Am. Soc. Mass Spectrom.* 2009; 20:1068–1077. [PubMed: 19264504]
65. Hanrieder J, Ljungdahl A, Falth M, Mammo SE, Bergquist J, Andersson M. L-DOPA-induced dyskinesia is associated with regional increase of striatal dynorphin peptides as elucidated by imaging mass spectrometry. *Mol. Cell. Proteomics.* 2011; 10:M111.009308. [PubMed: 21737418]
66. Römpf A, Guenther S, Schober Y, Schulz O, Takats Z, Kummer W, Spengler B. Histology by mass spectrometry: label-free tissue characterization obtained from high-accuracy bioanalytical imaging. *Angew. Chem., Int. Ed.* 2010; 49(22):3834–3838.
67. Zimmerman TA, Rubakhin SS, Sweedler JV. MALDI mass spectrometry imaging of neuronal cell cultures. *J. Am. Soc. Mass Spectrom.* 2011; 22:828–836. [PubMed: 21472517]
68. Yasuda-Kamatani Y, Yasuda A. Characteristic expression patterns of allatostatin-like peptide, FMRFamide-related peptide, orcokinin, tachykinin-related peptide, and SIFamide in the olfactory system of crayfish *Procambarus clarkii*. *J. Comp. Neurol.* 2006; 496:135–147. [PubMed: 16528723]

Figure 1.

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

Figure 2. De novo sequencing of *P. interruptus* CLDH. (a) An HRAM CID MS/MS spectrum of *P. interruptus* CLDH observed at m/z 986.1756. Its fragmentation map and sequence coverage are shown in the inset. (b) The accurate assignment of amino acid substitution at position 27 in CLDH by registering the mass discrepancy observed in ProSight to b₂₇ ion in the raw MS/MS spectrum. (c) A zoomed-in view of the well-resolved b₂₇ ion detected in the original CID spectrum. (d) Sequence alignment of the newly sequenced *P. interruptus* CLDH with previously reported CLDH from other species, highlighting the sequence conservation of this peptide family. Species names in purple shade indicate insects; blue shade indicates crustacean. Sequence letters highlighted in green shade are shared among various species.

Figure 3.

De novo sequencing of two novel *P. interruptus* neuropeptides as homologues of SIFamide.

(a) A zoomed-in MS spectra showing the calculated deconvoluted masses of the three precursors identified in single MS/MS event with ProSight search with the ion at m/z 465.9091 assigned as the product of neutral loss from the ion at m/z 471.9124. Isotopic peaks from different precursors are indicated with colored dots, whereas the overlapped fragment ions are shown in green. (b) The multiplexed MS/MS spectra that exhibits the fragmentation pattern of the two neuropeptides, with the fragments coming from each peptide identified by corresponding colors. (c) The fragmentation map and sequence coverage of the two SIFamide homologues. The peaks annotated in red and blue correspond to the fragment ions of the peptides labeled in red and blue, respectively, whereas the peaks labeled in green are those shared by the two peptides.

Figure 4.

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

Figure 5. Reductive dimethylation-assisted de novo sequencing of *P. interruptus* neuropeptides. (a) A schematic illustration of the form-aldehyde labeling reaction. (b) De novo sequencing of dimethylated orckinin NFDEIDRAGLG[F]H. (c) De novo sequencing of dimethylated RFamide QDLDDHVFLRFa. (d) De novo sequencing of RFamide pQDLDDHVFLRFa that cannot be labeled due to N-terminus blockage.

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

Figure 7.

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

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

family	[M + H] ⁺	sequence	Q-TOF	Orbitrap	TOF/TOF
AST-A	852.43	ADPYAFGLa	+		d
	936.51	PRNYAFGLa	+	+	+
RFa	887.56	PSLRLRFa	+	+	
	905.51	PSMRLRFa		+	
	1105.63	SMPSLRLRFa	+	+	+, d
	1147.65	APQRNFLRFa	+	+	
	1271.65	pQDLHDVFLRFa	+		d
	1288.68	QDLHDVFLRFa	+		
Orcomyotropin	976.44	FDAFTTGFA		+	
	1186.52	FDAFTTGFGHS	+	+	
Orcokinin	979.45	NFDEIDRA	+		
	1148.57	NFDEIDRAGLa		+	+
	1206.57	NFDEIDRAGLG		+	+, d
	1220.59	NFDEIDRAGLA	+		+, d
	1234.57	NFDEIDRAGVV		+	+
	1240.56	NFDEIDRAGFG	+		+, d
	1254.57	NFDEIDRAGFA	+	+	+, d
	1256.55	NFDEIDRSGFG		+	+, d
	1270.57	NFDEIDRSGFA	+		+, d
	1353.64	NFDEIDRAGLGF	+		+, d
	1490.7	NFDEIDRAGLGFH	+	+	+, d
	1516.68	NFDEIDRSGFGFNa	+		+, d
	1540.68	NFDEIDRSGFGFH		+	+, d
	1161.65	RKPPFNGSIFa	+		+
	1381.74	GYRKPPFNGSIFa	+	+	+, d
SIFa	1382.72	GYRKPPFNGSIF	+		
	1395.75	GYRKPPYNGSIYa(-H ₂ O)		+	
	1397.78	GYRKPPFNGSIYa	+	+	
	1413.74	GYRKPPYNGSIYa		+	
	1423.75	(Ac)GYRKPPFNGSIFa	+	+	+, d
	1431.72	GYRKPPFNGSYFa	+		
	1439.74	GYRKPPFNGSIFG	+	+	+, d
PDH	887.52	NSLLGISAL	+		
	1012.62	ELAAQILRV		+	
	1111.56	KYQAREM(O)VA	+		+, d
	1133.56	AREM(O)VAELAQ	+		
	1927.03	NSELINSILGLPKVMNDAA	+	+	
Tachykinin	605.3	APSGFQa	+		
	779.38	APSGFLGM	+		

family	[M + H] ⁺	sequence	Q-TOF	Orbitrap	TOF/TOF
	934.49	APSGFLGMRa	+	+	+, d
	950.49	APSGFLGM(O)Ra	+		+, d
	992.5	APSGFLGMRG	+	+	
	1650.96	AFRNGNAPVGLAVPLRa	+	+	+, d
	1651.94	AFRNGNAPVGLAVPLR			+
	1814.02	YAFRNGNAPVGLAVPLRa		+	
	3084.52	DAAAPLNEVDEASANDYPILPDPIAARLY		+	
	3155.53	DAAAPLNEVDEASANDYPILPDPIAARLYA		+	
	4716.45	DAAAPLNEVDEASANDYPILPDPIAARLYAFRNGNAPVGLAVPLRa		+	d
Neuropeptide F	2304.11	SDYPLPPGDALMEASERLLET		+	
	2488.23	SDYPLPPGDALMEASERLLETLA		+	
CLDH	2956.52	GLDLGLGRGFGSQAAKHLMGLAAANYAGGPa	+	+	+, d

All peptides reported in this study are new to this species, and those novel neuropeptides sequenced using our multifaceted strategy are highlighted in red; d denotes neuropeptides detected by direct tissue analysis