JOURNAL OF PROTEOMICS 91 (2013) 1–12

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

A multi-scale strategy for discovery of novel endogenous neuropeptides in the crustacean nervous system

*Chenxi Jia, Christopher B. Lietz, Hui Ye, Limei Hui, Qing Yu, Sujin Yoo, Lingjun Li**

School of Pharmacy, University of Wisconsin–Madison, WI, USA
Department of Chemistry, University of Wisconsin–Madison, WI, USA

ARTICLE INFO

Article history:
Received 18 March 2013
Accepted 16 June 2013
Available online 24 June 2013

Keywords:
Neuropeptide
Peptidomics
De novo sequencing
Mass spectrometry
Crustacean

ABSTRACT

The conventional mass spectrometry (MS)-based strategy is often inadequate for the comprehensive characterization of various size neuropeptides without the assistance of genomic information. This study evaluated sequence coverage of different size neuropeptides in two crustacean species, blue crab *Callinectes sapidus* and Jonah crab *Cancer borealis* using conventional MS methodologies and revealed limitations to mid- and large-size peptide analysis. Herein we attempt to establish a multi-scale strategy for simultaneous and confident sequence elucidation of various sizes of peptides in the crustacean nervous system. Nine novel neuropeptides spanning a wide range of molecular weights (0.9–8.2 kDa) were fully sequenced from a major neuroendocrine organ, the sinus gland of the spiny lobster *Panulirus interruptus*. These novel neuropeptides included seven allatostatin (A- and B-type) peptides, one crustacean hyperglycemic hormone precursor-related peptide, and one crustacean hyperglycemic hormone. Highly accurate multi-scale characterization of a collection of varied size neuropeptides was achieved by integrating traditional data-dependent tandem MS, improved bottom-up sequencing, multiple fragmentation technique-enabled top-down sequencing, chemical derivatization, and *in silico* homology search. Collectively, the ability to characterize a neuropeptidome with vastly differing molecule sizes from a neural tissue extract could find great utility in unraveling complex signaling peptide mixtures employed by other biological systems.

Biological significance

Mass spectrometry (MS)-based neuropeptidomics aims to completely characterize the neuropeptides in a target organism as an important first step toward a better understanding of the structure and function of these complex signaling molecules. Although liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS) with data-dependent acquisition is a powerful tool in peptidomic research, it often lacks the capability for *de novo* sequencing of mid-size and large peptides due to inefficient fragmentation of peptides larger than 4 kDa. This study describes a multi-scale strategy for complete and confident sequence elucidation of various sizes of neuropeptides in the crustacean nervous system. The aim is to fill a technical gap where the conventional strategy is inefficient for comprehensive characterization of a complex neuropeptidome

* *Corresponding author at:* School of Pharmacy and Department of Chemistry, University of Wisconsin–Madison, 777 Highland Avenue, Madison, Wisconsin 53705-2222, USA. Fax: +1 608 262 5345.
E-mail address: lli@pharmacy.wisc.edu (L. Li).

1874-3919/$ – see front matter
http://dx.doi.org/10.1016/j.jprot.2013.06.021

without assistance of genomic information. Nine novel neuropeptides in a wide range of molecular weights (0.9–8.2 kDa) were fully sequenced from a major neuroendocrine organ of the spiny lobster, *P. interruptus*. The resulting molecular information extracted from such multi-scale peptidomic analysis will greatly accelerate functional studies of these novel neuropeptides.

1. Introduction

Neuropeptides, including endogenous peptide neuromodulators and hormones, mediate or modulate neuronal communication by acting on cell surface receptors and are involved in a broad range of physiological and behavioral processes [1–4]. Mass spectrometry (MS)-based neuropeptidomics aims to completely characterize the neuropeptides in a target organism as an important first step toward a better understanding of the structure and function of these complex signaling molecules [5–17]. A significant challenge to this goal is that many of these endogenous neuropeptides display large diversity at the molecular and cellular level, such as various molecular sizes [15], extensive and multiple post-translational modifications (PTMs) [18], different hydrophobicities [12], and a wide dynamic range of concentrations [19]. Because of this, a uniform approach for comprehensive neuropeptide characterization is difficult to engineer.

A neuropeptidome usually contains peptides of various sizes [1]. Although liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS) with data-dependent acquisition is a powerful tool in peptidome research [5,6,11–13,17], it lacks the capability for *de novo* sequencing of mid-size and large peptides due to inefficient fragmentation of peptides larger than 4 kDa [20]. Innovation of MS-based instrumentation has greatly advanced our capability to analyze complex peptide mixtures, including those large peptides and intact proteins with extensive modifications [21]. For example, emergence of ultra-high resolution Fourier transform mass spectrometers greatly facilitates the bioinformatics-assisted peptidome research, as mass accuracy is a critical factor in scoring algorithm for peptide sequence assignment and protein identification [12,17]. The observed mass values of peptide precursors and their resulting fragment ions are matched to the theoretical values arising from a genome/cDNA sequence, and thus partial peptide sequence coverage from interpretation of MS/MS spectra can usually produce confident hits [10]. However, there are many valuable animal models whose genomes have not been sequenced yet, thus no genomic database or large cDNA database is available for database searching strategy commonly used [1,3]. Peptide discovery in these model organisms would need to rely on obtaining full sequence coverage, including enhanced local identification confidence on each amino acid residue.

The California spiny lobster *Panulirus interruptus* has long served as an important animal model for many areas of research in endocrinology and neurobiology [22–24], but its genome has not been sequenced yet and there is no protein/cDNA database. In particular, its stomatogastric ganglion has been utilized as a powerful model system to understand the cellular mechanisms of rhythmic pattern generation in neuronal networks [22]. Many studies reported that neurotransmitters and neuropeptides regulate the functional output of these well-defined neuronal circuits [25–28]. Therefore, it is highly desirable to obtain accurate molecular information on neuropeptides in this species.

In a previous study [29], we established a strategy by combining bottom-up, off-line top-down, and on-line top-down MS methods for confident *de novo* sequencing of crustacean hyperglycemic hormone (CHH)-family neuropeptides with molecular weight (MW) around 8.4–9.2 kDa. The current work aims to discover and confidently identify various sizes of novel neuropeptides (MW 0.9–8.2 kDa) in the crustacean nervous system. We evaluate the possibility for improvement of current approaches on sequencing of small-, middle- and large-molecular sizes of neuropeptides. A multi-scale strategy is established by rational optimization of methodology and further validated by sequencing of nine novel neuropeptides in *P. interruptus* sinus gland, a major neuroendocrine structure that secretes peptide hormones to regulate many essential functions of the animal.

2. Materials and methods

2.1. Chemicals

Methanol, glacial acetic acid, borane pyridine and formaldehyde were obtained from Sigma-Aldrich (St. Louis, MO). Optima grade formic acid, acetonitrile (ACN), water, and methanol were purchased from Fisher Scientific (Pittsburgh, PA).

2.2. Animals and tissue dissection

Blue crabs *Callinectes sapidus* and Jonah crabs *Cancer borealis* were shipped from the Fresh Lobster Company (Gloucester, MA); and the California spiny lobsters were purchased from Catalina Offshore Products (San Diego, CA). Tissue dissection was performed according to our previous report [30]. Briefly, ten animals were anesthetized in ice, and the sinus glands were dissected and collected in chilled acidified methanol and stored in −80 °C freezer prior to further sample processing.

2.3. Tissue extraction and LC fractionation

The tissues were homogenized and extracted with 100 μL of acidified methanol (methanol:H_2O:acetic acid, 90:9:1, v:v:v) [30,31]. After centrifugation, supernatant fractions were combined and the extraction procedure was repeated three times. After dryness in a Savant SC 110 SpeedVac concentrator (Thermo Electron Corporation, CA), the sample was re-suspended in 100 μL of deionized water for further analysis. High-performance liquid chromatography (HPLC) separations were performed with a Waters Alliance HPLC system (Milford, MA) according to our previous report [31]. The mobile phases included solution A