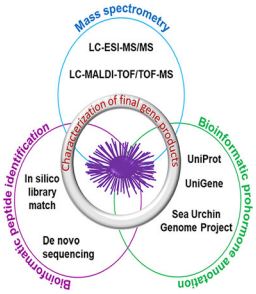
J. Am. Soc. Mass Spectrom. (2018) 29:923–934
DOI: 10.1007/s13361-018-1898-x

FOCUS: 29th SANIBEL CONFERENCE, PEPTIDOMICS; BRIDGING THE GAP BETWEEN PROTEOMICS AND METABOLOMICS BY MS: RESEARCH ARTICLE

Exploring the Sea Urchin Neuropeptide Landscape by Mass Spectrometry

Eric B. Monroe,[1] Suresh P. Annangudi,[1] Andinet A. Wadhams,[1] Timothy A. Richmond,[1] Ning Yang,[1] Bruce R. Southey,[2] Elena V. Romanova,[1] Liliane Schoofs,[3] Geert Baggerman,[4] Jonathan V. Sweedler[1]

[1]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2]Department of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3]Functional Genomics and Proteomics Unit, KU Leuven, 3000, Leuven, Belgium
[4]ProMeta Interfacultary Center for Proteomics and Metabolomics, KU Leuven, 3000, Leuven, Belgium



Abstract. Neuropeptides are essential cell-to-cell signaling messengers and serve important regulatory roles in animals. Although remarkable progress has been made in peptide identification across the Metazoa, for some phyla such as Echinodermata, limited neuropeptides are known and even fewer have been verified on the protein level. We employed peptidomic approaches using bioinformatics and mass spectrometry (MS) to experimentally confirm 23 prohormones and to characterize a new prohormone in nervous system tissue from *Strongylocentrotus purpuratus*, the purple sea urchin. Ninety-three distinct peptides from known and novel prohormones were detected with MS from extracts of the radial nerves, many of which are reported or experimentally confirmed here for the first time, representing a large-scale study of neuropeptides from the phylum Echinodermata. Many of the identified peptides and their precursor proteins have low homology to known prohormones from other species/phyla and are unique to the sea urchin. By pairing bioinformatics with MS, the capacity to characterize novel peptides and annotate prohormone genes is enhanced.

Keywords: Sea urchin, Mass spectrometry, Neuropeptides, Peptidomics, Bioinformatics

Received: 30 September 2017/Revised: 4 December 2017/Accepted: 16 December 2017/Published Online: 17 April 2018

Introduction

Peptides are the most diverse class of cell-to-cell signaling molecules found in animals. These prohormone-derived molecules act as neuropeptides, neurotransmitters, neurohormones, and neuromodulators and play important roles in the regulation of a variety of physiological processes, such as food intake, drug addiction, and circadian rhythm [1–5]. At present, more than 20,000 peptides originating from all classes of animals are known; however, this knowledge is not spread uniformly across the metazoan phyla [6]. For instance, our understanding of neuropeptides from the phylum Echinodermata, which includes animals like starfish (Asteroidea), sea cucumbers (Holothuroidea), and sea urchins (Echinoidea), is largely based on computational analyses of genomic and transcriptomic information from a few species (reviewed in detail by Semmens and Elphick [7]), even though echinoderms occupy an interesting position in the Animal Kingdom. Echinoderms are a unique group of species among deuterostomes that exhibit a broad diversity of body forms, but share unique characteristics that distinguish them from other clades of the Animal Kingdom. Unlike other deuterostomes, living echinoderms have a pentamerous (five-part) radial symmetry (with a few exceptions), a water vascular system derived from a central cavity,

Eric B. Monroe and Suresh P. Annangudi contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13361-018-1898-x) contains supplementary material, which is available to authorized users.

Correspondence to: Jonathan Sweedler; *e-mail:* jsweedle@illinois.edu

and are propelled by cilia movement. Recent studies of their molecular phylogeny, nuclear and mitochondrial genomes, and evolutionary developmental biology have unambiguously demonstrated that echinoderms and hemichordates form a distinct new superphylum among deuterostomes, Ambulacraria, segregating them from other chordates with a bilateral body plan [8]. Understanding the neuropeptide complement in this phylum may provide new insight into the origins of the nervous system and neurogenesis in Metazoa.

Several neuropeptides had been identified previously from echinoderms using mass spectrometry (MS) and radioimmunoassay, including the SALMFa-related peptides found in various starfish, sea cucumbers, and more recently, in the sea urchin *Strongylocentrotus purpuratus* [9–11], as well as the NGFFFa peptide family [12]. Additionally, genome and transcriptome-wide analyses of proteins involved in neuronal control in *S. purpuratus* uncovered 20 neuropeptide and hormone precursors [13, 14]. This is in contrast to the number of neuropeptide precursors identified in other sequenced invertebrates and vertebrates, including starfish [15, 16], which ranges between 30 and 100 [17–22]. In addition, 37 neuropeptide and hormone receptors have been reported based on the *S. purpuratus* genomic sequence [13]. The striking variance in the number of discovered neuropeptides between echinoderms and other sequenced animals and between receptors and known peptide genes suggests that a number of prohormone genes and their related peptides remain to be discovered. To address this information gap, we utilized a peptidomics approach [18, 23–27] to examine the peptidome of the *S. purpuratus* nervous system.

Theoretically, the availability of the entire genomic sequence from the sea urchin should allow putative peptide precursors to be mined from the genome via bioinformatic tools alone. However, in practice, the identification of peptide precursors solely via computational methods is challenging because bioactive signaling peptides are typically short in length, and their active cores are often limited to only a few amino acids, resulting in incomplete or missed annotations. Fortunately, this issue is alleviated by employing mass spectrometric techniques, particularly when paired with bioinformatic tools and resources, such as genomic/proteomic information and databases. De novo sequencing of unassigned tandem MS (MS/MS) spectra can lead to the structural characterization of unknown peptides and identification of new gene families that may otherwise be missed. Examples of this strategy were demonstrated in studies using MS and bioinformatic tools to successfully identify novel prohormones and characterize previously unassigned peptides in the honey bee (*Apis mellifera*) [20], planarian (*Schmidtea mediterranea*) [28], cichlid fish (*Astatotilapia burtoni*) [29], zebra finch (*Taeniopygia guttata*) [30], and starfish *(Acanthaster planci)* [16].

In this work, we extended our previous IggyPep (Indexed Genomes Gracefully Yield Peptide IDs) methodology [27], which relied on the translated genome in all six reading frames, to utilize two-dimensional liquid chromatography (LC) and MS/MS, combined with enhanced homology-based gene discovery (Fig. 1). We used this multiplatform approach to uncover new *S. purpuratus* prohormones and neuropeptides that were not identified during the initial IggyPep annotation.

Experimental

Tissue Dissection/Preparation

Wild adult *S. purpuratus* were collected from the Pacific Ocean by Santa Barbara Marine Biologicals (Santa Barbara, CA) and kept at 15 °C in an aquarium containing continuously circulating, aerated, and filtered artificial seawater prepared using Instant Ocean salts (Aquarium Systems Inc., Mentor, OH), according to manufacturer instructions. Animals were anesthetized by injection of isotonic $CaCl_2$ into the body cavity. Once spine movement was minimized, demonstrating anesthesia, the animal's test (or shell) was equatorially bisected using surgical scissors. The upper and lower portions of the test were separated, the gonads and internal organs removed, and the body cavity rinsed with ice-cold artificial sea water that consisted of 460 mM NaCl, 10 mM KCl, 10 mM $CaCl_2$, 22 mM $MgCl_2$, 6 mM $MgSO_4$, and 10 mM HEPES, adjusted to pH 7 by titration with 1 M NaOH. The dissections were performed on ice and as rapidly as possible to limit protein and peptide degradation. The water canals covering each radial nerve were removed with forceps; nerves were separated from surrounding tissue by cutting along their edges with a scalpel, allowing each nerve to be lifted with forceps and placed in a centrifuge tube for peptide extraction. Typically, two or three animals were dissected for each extraction, with five radial nerves collected from each animal.

Collected tissues were homogenized and extracted in acidified acetone (60:40:1 acetone/H_2O/HCl) at 4 °C as described elsewhere [27, 31]. Insoluble material was pelleted via centrifugation (5804R, Eppendorf, Hamburg, Germany) for 10 min at 15,000×*g*. The supernatant was then removed, dried with a Savant SpeedVac (Thermo Fisher Scientific, Waltham, MA), and resuspended in minimal 5% acetonitrile (ACN) prior to use of a 10 kDa molecular weight cut-off filter (Millipore, Billerica, MA) to remove large proteins from the sample. In order to increase the coverage of identified prohormones, several MS instruments were used in these experiments as differing ionization strategies regularly produce different sets of ions: LC-electrospray ionization (ESI)-MS systems with either ion trap (IT) or quadrupole time-of-flight (QTOF) analyzers, or a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF/TOF) instrument, as detailed below.

LC-ESI-IT-MS/MS

Extracts were separated and analyzed with a capLC (Micromass, Manchester, UK) HPLC system in-line with a Bruker HCT Ultra ESI-IT mass spectrometer (Bruker, Billerica, MA). Portions (10 μL) of the extract were injected using a manual injector (Valco Instruments Co, Inc., Houston, TX),

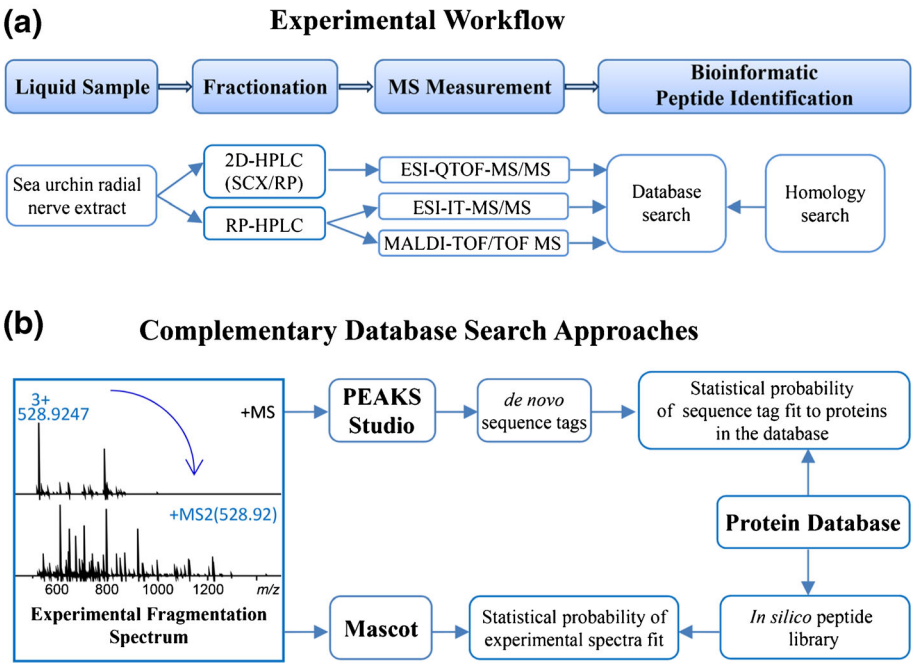


Figure 1. Multi-faceted peptidomics approach to characterization of neuropeptides in the *S. purpuratus* nervous system. **a** A general workflow. **b** Differences between the bioinformatics methods employed for peptide identification

loaded onto a trap column (PepMap, C18, 5 µm, 100 Å, Dionex, Sunnyvale, CA), and washed for 5 min. The trapped peptides were then eluted onto a reversed-phase capillary column (LC Packings 300-µm i.d. × 15 cm, C18 PepMap100, 100 Å, Dionex) at a 2-µL/min flow rate. A 70-min gradient run for LC separation was used to separate the peptides. MS data acquisition and the subsequent MS/MS of selected peaks were performed in a data-dependent manner using the Esquire software (Bruker). The capillary voltage was set to 3800 v, and the ion transmission optics were optimized for mass ranges between *m/z* 800 and 1000. For each MS scan, three peptides were selected to be fragmented, for 300 to 500 ms, based on their charge (preferably 2+) and intensity. Previously fragmented precursor ions were dynamically excluded for a period of 60 s to maximize the number of peptides fragmented. The MS and MS/MS scans were performed in the range of *m/z* 300 to 1500 and 50 to 2000, respectively.

LC-ESI-QTOF-MS/MS

Two-dimensional capillary LC-MS/MS experiments were conducted with an Ultimate HPLC pump, a column-switching device (Switchos), and a Famos autosampler (all LC Packings, the Netherlands) coupled to a hybrid QTOF mass spectrometer (Micromass, UK). The sample was dissolved in 25 µL of 5% ACN in water (Sigma-Aldrich, St. Louis, MO) and filtered. A 20-µL volume of this sample was injected onto a 500 µm × 15 mm Bio-SCX column (Dionex, Sunnyvale, CA) in-line with a reversed-phase C18 precolumn (µ-guard column, MGU-30 C18, LC Packings, San Francisco, CA) at a flow rate of 30 µL/min for 10 min with water containing 2% ACN and 0.1% formic acid (FA). The SCX column was switched off-line while the reversed-phase precolumn was rinsed for 5 min. The reversed-phase trapping column was then switched in-line with a 75 µm × 150 mm C18 capillary column (Thermo Fisher Scientific). The first fraction comprised those peptides that did not bind to the SCX column and were immediately trapped by the reversed-phase column, using a gradient from 2 to 50% ACN containing 0.1% FA and a flow rate at 200 nL/min over 50 min. The second fraction of peptides was eluted from the SCX column by injection of 20 µL of 20 mM ammonium acetate solution. The eluted peptides were again concentrated and desalted on the C18 precolumn prior to MS analysis. This cycle was repeated 10 times, including a breakthrough cycle (no salt), with successive concentrations of 20, 50, 100, 200, 400, 600, 800, 1000, and 2000 mM ammonium acetate for each cycle.

The LC system was connected in series with the ESI interface of the mass spectrometer, and the column eluent was directed through a stainless steel emitter (Proteon, Denmark). The needle voltage was set at 1650 V and the cone voltage at 35 V, with nitrogen as the nebulizing gas. Parent ions with two or three charges were automatically selected for fragmentation as they eluted from the column. Argon was used as the collision gas and the collision energy was set at 25–40 eV, depending on the mass and charge state of the selected ion. The detection window in the survey scan was set from *m/z* 400 to 1500. Fragmentation spectra were acquired from *m/z* 50 to 2500.

LC-MALDI-TOF/TOF-MS

Samples were dissolved in 15 μL of 5% ACN in water and filtered. Ten microliters of the sample were loaded onto the guard column with an isocratic flow of 2% ACN in ULC/MS grade water (Biosolve, Lexington, MA) and 0.1% FA, at a flow rate of 10 μL/min. After 2 min, the guard column was switched online with a PepMap C18, 3 μm × 150 mm nanocolumn (Thermo Fisher Scientific) analytical capillary column. Separation was conducted using a linear gradient from 95% solvent A, 5% solvent B to 5% solvent A, 65% solvent B in 35 min (solvent A: H_20/ACN/FA (94.9:5:0.1, *v*/*v*/*v*); solvent B: ACN/FA (99.9:0.1, v/v)), at a flow rate of 200 nL/min.

Fractions were collected manually using 15 s intervals on a ground steel MALDI target plate (MTP 384, Bruker). Prior to fraction collection, the sample plate was prespotted with α-cyano-4-hydroxycinnamic acid (CHCA) using 0.5 μL of a 5× dilution of a saturated solution of CHCA in 100% acetone. Fractions were collected in the center of each target spot within 10 to 35 min of the LC separation and left to dry, and then analyzed with the ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker). Mass spectra were acquired in reflectron mode from *m/z* 750 to 4000. Ions of sufficient intensity were selected manually for fragmentation analysis.

Genome/Protein Sequences

S. purpuratus nucleic and protein sequences were obtained from the Sea Urchin Genome Project [32]. Prohormone protein sequences from vertebrates and mollusks were obtained from NCBI GenBank [33] and UniProt [34] public sequence depositories and previously published articles [18, 27].

Similarity searches were conducted using the Basic Local Alignment Search Tool (BLAST) [35]. The default BLAST settings were modified depending on the length of the query. For a long query, only the expectation value (e-value) was increased. Occasionally, the filtering option was turned off in order to allow low-complexity sequences marked by biased representation of certain residues (including acidic, basic, or proline stretches) in the query sequence. For short queries (e.g., only when the query was a short peptide), the e-value was increased to 20,000, the word size decreased to 2, the gap penalty reduced to 9, and filtering turned off. Proteins identified from the similarity search were curated into a prohormone database. The existence and length of each signal peptide were determined in the predicted prohormone proteins using SignalP version 4.1 (http://www.cbs.dtu.dk/services/SignalP/) [36]. Cleavage sites were predicted using mollusk models [37] implemented in NeuroPred [38].

MS Data Analysis

Fragmentation spectra from both the LC-ESI-MS/MS and LC-MALDI-MS/MS analyses were sent to an in-house Mascot server [39] (Matrix Science Inc., Boston, MA) for protein and peptide identification. Pyroglutamic acid (Q), carboxyterminal amidation, and oxidation of Met were selected as variable modifications for all identifications due to their prevalence in known neuropeptides; no processing enzyme was selected. Pyroglutamination at the N-terminus or amidation at the C-terminus are two post-translational modifications (PTMs) that are hallmarks for the secretory pathway and are present in many bioactive neuropeptides [40, 41].

Two rounds of database searches were performed using different protein databases and two search engines, Mascot [39] and PEAKS Studio [42] (Bioinformatics Solutions Inc., Canada). For the first round, the protein database consisted of annotated *S. purpuratus* proteins obtained from the Sea Urchin Genome Project [32]. The mass accuracy for precursor ions was set to 0.3 Da for ion trap spectra and 0.1 Da for spectra from the QTOF and MALDI-TOF/TOF instruments. The mass tolerance for fragment ions was set at 0.5 Da for ion trap spectra, 0.1 Da for spectra from the QTOF instrument, and 0.3 Da for spectra from MALDI-TOF/TOF instrument. For the second round, all of the data were searched against a smaller, targeted database of annotated sea urchin proteins and neuropeptide precursors obtained from the previous bioinformatics analyses. Less stringent parameters were used; 0.5 and 0.3 Da were used as precursor and fragment mass tolerances, respectively, for data acquired from all instruments because the presence of these prohormones in the investigated region is known. This targeted database contained prohormones from which the neuropeptides identified in the first round were derived. Criteria to evaluate the quality of identifications from the MS/MS data included a Mascot peptide score greater than the significance threshold at an expected value of less than 0.05.

High-quality MS/MS spectra that did not yield any identification in the Mascot searches were subjected to automated partial de novo sequencing through annotation using PEAKS Studio version 5.2 and manual annotation using Bruker DataAnalysis software (Bruker). Partial sequences obtained from automated de novo sequencing with a > 80% probability score are listed in Table S1. These sequences were used to search the public database resources for homology, both in the annotated protein database and in the crude genome database.

Results

An MS-based strategy was implemented to investigate the *S. purpuratus* central nervous system peptidome. Peptides from 23 precursor proteins were detected; all prohormones were supported by gene predictions and 17 were supported by expressed sequence tag information in the UniGene database (see Table 1). Most of these proteins are expected to be prohormones due to the presence of a predicted signal peptide and peptides resulting from cleavage at NeuroPred-predicted sites (see Supplementary Material). Spurp_10 (secretogranin V |neuroendocrine protein 7B2 precursor) and Spurp_23 (thymosin beta) lack signal peptides and technically are not prohormones. It is likely that the deposited Spurp_1 is only partial because no signal peptide was detected, the predicted

Table 1. Sea urchin proteins and prohormones experimentally confirmed by mass spectrometry

Protein symbol	Signal peptide	Mascot peptides	PEAKS peptide	NCBI Protein	Previously reported		NCBI Gene	UniGene	EchinoBase
Spurp_1	0	4	1	GNOMON:39491338.m	M				
Spurp_2	28	2	0	XP_011662958.1	M		105437599		
Spurp_3	23	1	0	XR_973258.1	M		105445038		
Spurp_4	21	2	2	XR_972762.1	M	R	753167	Spu.14862	
Spurp_5	19	13	1	XP_003727926.1	M	R	100889131	Spu.15355	SPU_024381
Spurp_6	27	6	1	XP_001177887.1	M		753529	Spu.24513	SPU_003170
Spurp_7	24	1	0	XR_971714.1	M	R	753700	Spu.24538	
Spurp_8	24	2	0	XP_001175944.1	M	R	752246	Spu.36930	
Spurp_9	23	2	2	XR_973850.1		R	752149	Spu.24531	
Spurp_10	15	3	1	XP_799858.1	M		575887	Spu.14943	SPU_015798
Spurp_11	15	3	1	NP_001116993.1	M	R	580381	Spu.7138	SPU_008352
Spurp_12	29	8	4	XP_785647.1	M	R	580501	Spu.9558	SPU_003108
Spurp_13	24	9	7	XP_001199000.1	M	R	763123	Spu.15620	SPU_018666
Spurp_14	21	2	3	XP_001198950.2	M		763080	Spu.36923	SPU_021555
Spurp_15	19	1	0	XP_011670672.1			757218	Spu.36964	SPU_009905
Spurp_16	21	0	3	XR_973214.1	M	R	752016	Spu.30718	
Spurp_17	24	0	1	XP_011677983.1			105444880		
Spurp_18	22	0	1	XP_001176371.1	M	R	752454	Spu.17285	
Spurp_19	21	0	1	XP_011666330.1			764662		
Spurp_20	0	0	1	XP_011682716.1			105446953		
Spurp_21	22	0	1	XR_972749.1		R	752090	Spu.14945	
Spurp_22	25	0	1	XR_971124.1			753482	EC439240.1	
Spurp_23	0	0	2	NP_999791.1	M		373484	Spu.3149	SPU_026031
Spurp_24	17	0	0						

M: Menschaert, G. et al., A hybrid, de novo based, genome-wide database search approach applied to the sea urchin neuropeptidome. *J. Proteome Res*. 9, 990-996 (2010)
R: Rowe, M. L. and M. R. Elphick, The neuropeptide transcriptome of a model echinoderm, the sea urchin *Strongylocentrotus purpuratus*. *Gen. Comp. Endocrinol*. 179, 331-344 (2012)

neuropeptide cleavage sites are located near the end of the N-terminus, and only two peptides were experimentally detected from the protein sequence. Finally, Spurp_20 may not be a prohormone as no signal peptide was predicted and only an N-terminal peptide was mass-matched to this protein. For several prohormones, we detected the linker peptide immediately adjacent to the signal peptide; in case of Spurp_8 and Spurp_13, this allowed us to determine the actual signal peptide cleavage site, and thus, SignalP predictions were corrected as indicated in the Supplementary Material. Out of 22 prohormones, four encode numerous putative amidated peptides (Spurp_11, 13, 14, 24), and the majority of the precursors encode both putative amidated and non-amidated peptides.

A total of 93 unique peptides were identified, with 59 peptides assigned using Mascot (Table 2). Only eight peptides were identified with both LC-ESI-MS/MS and LC-MALDI-MS/MS, 41 peptides identified with LC-ESI-MS/MS, and 10 identified with LC-MALDI-MS/MS. There were 34 tentative peptides identified from partial de novo tags by PEAKS (Table S1), and 23 of these peptides originated from proteins also identified using Mascot, while 11 peptides were matched to eight proteins not found in Mascot searches. There were 19 peptides with PTMs, including 12 amidated peptides (10 from Mascot and two from PEAKS), and six peptides with pyroglutamination at the N-terminus (four from Mascot and two from PEAKS).

All of the detected peptides were unique to the matched prohormone, and most peptides were present in a single unique copy in the prohormone. However, some prohormones contained multiple copies of the same identified peptide or peptides with very similar sequences within the respective protein sequence. Both Spurp_12 (pedal peptide 1/orcokinin-type neuropeptide precursor) and Spurp_13 (putative tachykinin-type neuropeptide precursor) contained multiple copies of the same identified peptide within the respective protein sequence. For example, the ANMFRSRLRG sequence was repeated 10 times in Spurp_13, and GFNSGAMEPLGAGFF was repeated eight times in Spurp_12. These prohormones and other prohormones, such as Spurp_5 (epsin 1-like protein), Spurp_12, and Spurp_13, provided different peptides with similar sequences within the respective protein sequence. We were able to identify eight and four peptides from Spurp_12 using Mascot and PEAKS, respectively, resulting in > 75% sequence coverage (see Supplementary Material). Interestingly, we were able to identify 12 unique peptides (plus one peptide with oxidation of Met), and one putative peptide by mass match from another protein annotated as Spurp_5, with nearly a complete sequence coverage of this protein (see Supplementary Material). In addition, peptides from Spurp_5 were located between the dibasic cleavage sites. These features support Spurp_5 as a prohormone. Spurp_11 encodes numerous copies of QYPG-amide peptide and multiple other four-amino-acid-long variants of amidated peptides. Unfortunately, sequencing of such small peptides on the employed MS equipment was limited, but we did detect multiple longer linker peptides from this prohormone.

We experimentally confirmed peptides from Spurp_10 (secretogranin V, also known as 7B2 protein). We found that

Table 2. Sea urchin peptides detected by complimentary mass spectrometry methods using mascot

Symbol	NCBI Protein	Peptide sequence	Copy #	Exp. Mass[a]	δM	Mascot Score	Expect value[b]	LC-ESI-MS	LC-MALDI MS	PTM[c]
Spurp_1	GNOMON:39491338.m	R.GSTPEDIAELV.S	1	1129.55	−0.03	51	0.00022	X		
Spurp_1	GNOMON:39491338.m	R.GSTPEDIAELVS.R	1	1216.58	−0.19	41	0.0023	X		
Spurp_1	GNOMON:39491338.m	R.GSTPEDIAELVSRN.R	1	1486.72	−0.29	100	0.000000002	X	X	
Spurp_1	GNOMON:39491338.m	Q.QKQDLAAILDQLHNTYQM.A	1	2112.03	−0.01	70	0.0000018	X	X	pyro-Glu
Spurp_2	XP_011662958.1	R.GYPRNSVVADPVL.R	1	1385.73	−0.19	38	0.0028	X		
Spurp_2	XP_011662958.1	R.SPVTYKSMSKYLQGLASRRFV.R	1	2417.19	−0.1	61	0.000015		X	
Spurp_3	XR_973258.1	R.LYDALKNAQV.K	1	1133.61	−0.06	44	0.0011	X		
Spurp_4	XR_972762.1	R.RRSVDDLPQVNDAETE.-	1	1842.87	−0.3	63	0.0000012	X	X	
Spurp_4	XR_972762.1	R.SVDDLPQVNDAETE.-	1	1530.67	−0.18	32	0.014	X		
Spurp_5	XP_003727926.1	R.FGGAMEPMSSGFY.K	1	1379.55	−0.21	49	0.00017	X		
Spurp_5	XP_003727926.1	R.FGGANEPMRSG.F	1	1121.49	−0.19	53	0.00011	X		
Spurp_5	XP_003727926.1	R.FGGANEPmRSGFF.K	1	1431.62	0.02	46	0.0007	X		Oxidation
Spurp_5	XP_003727926.1	R.FGGANEPMRSGFF.K	1	1415.62	−0.09	80	0.0000018	X	X	
Spurp_5	XP_003727926.1	R.FGSGLDSMQSGFY.K	1	1394.58	−0.09	82	0.00000012	X		
Spurp_5	XP_003727926.1	R.FGSGSLEPMSSGFY.K	2	1464.62	−0.09	63	0.000023	X		
Spurp_5	XP_003727926.1	R.FGSMNMEPLVSGFY.K	1	1577.69	−0.23	84	0.00000009	X		
Spurp_5	XP_003727926.1	T.LPIEDKDGLDIEDQEE.A	1	1856.85	0.03	83	0.000000088	X		
Spurp_5	XP_003727926.1	T.LPIEDKDGLDIEDQEEAE.K	1	2056.93	0.03	53	0.000075	X		
Spurp_5	XP_003727926.1	K.NFGGSLDAMQSGFY.K	1	1492.63	−0.17	83	0.00000011	X		
Spurp_5	XP_003727926.1	K.NFGGSLEPMQSGFY.K	1	1532.66	−0.09	64	0.0000077	X		
Spurp_5	XP_003727926.1	K.NFGGSMEPMQSGFY.K	1	1550.62	−0.27	60	0.000038	X		
Spurp_5	XP_003727926.1	K.NFGSGLNMEPMQSGFY.K	1	1777.74	−0.27	48	0.0003	X		
Spurp_6	XP_001177887.1	R.GFETPASSRIN.S	1	1177.57	−0.08	45	0.00052	X		
Spurp_6	XP_001177887.1	R.GFETPASSRINS.R	1	1264.6	−0.08	55	0.000045	X		
Spurp_6	XP_001177887.1	R.GFRVLPQLNDDD.N	1	1387.67	−0.28	62	0.000011	X		
Spurp_6	XP_001177887.1	R.GFRVLPQLNDDDN.-	1	1501.71	−0.08	85	0.000000046	X	X	
Spurp_6	XP_001177887.1	P.QDEQIDLRYLLQNFLDNRDT.R	1	2491.2	−0.16	62	0.000015	X		pyro-Glu
Spurp_6	XP_001177887.1	R.SPQDEQIDLRYLLQNFLDNRDT.R	1	2692.31	0.48	41	0.0015	X		
Spurp_7	XR_971714.1	R.GAAENALDEQEIYEIIESLEHAM.S	1	2574.18	−0.15	72	0.00000038	X	X	
Spurp_8	XP_001175944.1	S.LQFETTQDRVPA.K	1	1403.7	−0.23	86	0.000000039	X		
Spurp_8	XP_001175944.1	C.SLQFETTQDRVPA.K	1	1490.74	−0.25	84	0.000000073	X	X	
Spurp_9	XR_973850.1	R.GGKKNMGSIHSHSGIHF.GK	1	1791.89	−0.05	73	0.00000073		X	Amidation
Spurp_9	XR_973850.1	R.MRLHPGLLF.GK	1	1081.62	−0.04	54	0.000061		X	Amidation
Spurp_10	XP_799858.1	K.APVYSGAKPIM.-	1	1132.59	−0.22	62	0.012	X		
Spurp_10	XP_799858.1	R.SINSYLPGDMVRHVS.K	1	1673.82	0.01	39	0.0023	X		
Spurp_10	XP_799858.1	R.SLKNRQLFTQTRNKY.S	1	1335.74	−0.06	28	0.027		X	
Spurp_11	NP_001116993.1	R.QFVGGEALEQESNIN.K	1	1616.73	−0.2	84	0.00000011	X		pyro-Glu
Spurp_11	NP_001116993.1	R.QFVGGELIPSPEL.R	1	1367.69	−0.21	52	0.000081	X		pyro-Glu
Spurp_11	NP_001116993.1	R.SEDDQDLLPMEI.R	1	1403.61	−0.3	54	0.00015	X		
Spurp_12	XP_785647.1	R.FLTGALEPLSSGFI.K	1	1450.77	−0.12	61	0.000015	X		
Spurp_12	XP_785647.1	K.GFHAGAMEPLSSGFID.GK	1	1633.75	−0.12	82	0.00000012	X		Amidation
Spurp_12	XP_785647.1	K.GFHNGAMEPLKSGFL.K	1	1603.78	−0.14	63	0.000012	X		
Spurp_12	XP_785647.1	K.GFNSGAMEPLGAGFF.K	8	1500.57	−0.1	83	0.00000011	X		
Spurp_12	XP_785647.1	K.GFNSGAMEPLGSGFI.K	5	1482.68	−0.16	76	0.00000038	X		
Spurp_12	XP_785647.1	K.GFNTGAMEPLGSGFI.K	2	1496.69	−0.23	65	0.0000054	X		
Spurp_12	XP_785647.1	R.GFYNGAMEPLSAGFHQ.GK	1	1723.78	−0.1	92	6.6E-09	X		Amidation
Spurp_12	XP_785647.1	K.RFLTGALEPLSSGFI.K	1	1606.87	0.03	56	0.000039	X		
Spurp_13	XP_001199000.1	R.ANFRARQRPKL.GK	1	1354.81	−0.05	28	0.023		X	Amidation
Spurp_13	XP_001199000.1	R.ANMFRSRLR.GK	10	1148.63	−0.03	37	0.0034		X	Amidation
Spurp_13	XP_001199000.1	R.ANMFRSRLRGK.GK	8	1333.75	−0.03	40	0.0019		X	Amidation
Spurp_13	XP_001199000.1	R.ANYFRGRGRKP.GK	1	1319.73	−0.06	53	0.000079		X	Amidation

Table 2 (continued)

Symbol	NCBI Protein	Peptide sequence	Copy #	Exp. Mass[a]	δM	Mascot Score	Expect value[b]	LC-ESI-MS	LC-MALDI MS	PTM[c]
Spurp_13	XP_001199000.1	R.ANYFRGRGRRP.GK	1	1347.74	−0.04	60	0.00002	X	X	Amidation
Spurp_13	XP_001199000.1	R.DDPDAALVDEFM.D	1	1336.54	−0.24	37	0.0058	X		
Spurp_13	XP_001199000.1	R.DDPDAALVDEFMDEE.K	1	1709.66	−0.25	56	0.000034	X		
Spurp_13	XP_001199000.1	R.DEPDAALVEDE.K	1	1201.5	−0.12	40	0.0023	X		
Spurp_13	XP_001199000.1	D.PDAALVDEFMDEE.K	1	1479.6	−0.38	41	0.003	X		
Spurp_14	XP_001198950.2	R.ARAQYAARRPPVTT.R	1	1556.85	−0.04	32	0.0096		X	
Spurp_14	XP_001198950.2	R.PHGGSAFVF.GR	1	916.45	−0.16	25	0.048	X		Amidation
Spurp_15	XP_011670672.1	R.NKEDLRAFINSMPTLEGL.R	1	2047.04	−0.05	47	0.00032		X	

[a]Monoisotopic experimental mass
[b]Mascot expected value
[c]Post-translational modification: pyro-Glu: pyroglutamination at the N-terminus; Amidation: glycine-mediated amidation at the C-terminus

S. purpuratus secretogranin V includes the highly conserved PPNPCP motif (indicated with an asterisk in Fig. 2a), which is shared among phyla from Platyhelminthes to Arthropoda and Chordata, including fishes, frogs, birds, and mammals. As expected, the cladogram (Fig. 2b) shows that species from the same phylum are more similar than species from different phyla.

A prospective prohormone protein was identified through homology search using bioinformatics without MS support. Prohormones from the mollusk *Aplysia californica* were used here to search the sea urchin contigs because many of the *A. californica* prohormones have been well characterized and contain repeated short peptides separated by basic residues. Matches with low expectation values (e value < 1 e^{-4}), or those that contained a pattern of conventional cleavage sites, such as mono- and di-basic residues, were selected as potential hits. A sequence match to a molluscan myomodulin prohormone was identified. The complete prohormone protein sequence was obtained as Spurp_24, which includes five predicted RFamides peptides, such as YGSDNRFamide (two copies within the predicted protein sequence). The putative predicted peptides were not experimentally detected in radial nerve samples, however.

Discussion

The interconnected structure of the adult *S. purpuratus* central nervous system has been compared to internet cable networks [13]. Dissecting specific anatomically correct sections of such structures can be challenging. Nonetheless, the five radial nerves that line the interior of the test and comprise a significant portion of its nervous system can be quickly and accurately isolated in a reproducible manner, minimizing delays during dissection that lead to increased protein degradation. Because radial nerve structures are covered by a layer of neurons and are thought to be involved in overall control, coordination [13, 44], and photoreception [45], we expected the radial nerves to be a source of neuropeptide synthesis and major signaling events; hence, we used these structures for our peptidomic measurements.

We used both LC-ESI-MS and LC-MALDI-MS to investigate the peptides in the sea urchin radial nerves. The advantage of using a multiplatform MS approach is that it increases the number of identified peptides because differing ionization and detection strategies oftentimes produce overlapping but different sets of detected ions from complex biological samples [46, 47]. Eight peptides were detected via both MS methods, and most proteins were identified by at least one confident peptide characterized by either MS approach. As shown in Table 1, by performing complementary MS analyses, the sequence coverage for several prohormones was increased. For example, in the case of the putative tachykinin-type neuropeptide precursor, LC-MALDI-MS/MS and LC-ESI-MS/MS each detected four out of nine peptides, with one peptide detected by both approaches. Moreover, structurally verified peptides from

(a)

```
Human                NIPNIVAELTGDN--IPKDFSEDQGYPDPPNPCPVGKTAD
Mouse                NIPNIVAELTGDN--IPKDFSEDQGYPDPPNPCPLGKTAD
Chimpanzee           NIPNIVAELTGDN--IPKDFSEDQGYPDPPNPCPVGKT-D
Drosophila           --PNRYPTIVKND-------AGLPAYCNPPNPCPEGYDME
Honey_bee            SVKNIKNEK-ENA--LP-------AYCTPPNPCPVGYTSK
Chicken              NIPNIVAELTGDN--IPKDFSEDQGYPDPPNPCPIGKTVD
Zebrafish            NIPNIVAELTGDS--VPKDFSEDHGYPDPPNPCPLGKTAA
Xenopus              NIPNILAELTGDN--IPKNFREDKGYPNPPNPCPVGKT-G
Ant                  DRHRIQPTGLKGI-KDEKAENPLPAYCTPPNPCPVGYTSE
Daphnia              LKISETASTTSGSSSSGKSENVLPAYCNPPNPCPIGYTAE
Sea_urchin           DIPNRQNPIPEVS----------SQYDNPPNPCPPGGELT
Acorn_worm           NIPNEKLALPELD---------SPEYEVPPNPCPMELP-K
Amphioxus            NIPNLWAAMVESD-----NNRKPDAYPNPPNPCPKGYTAE
Alligator            NIPNIVAELTGDN--IPKDFSEDQGYPDPPNPCPIGKTVD
Aplysia_californica  VPPN---------PKQVKSDKQLPEYCNPPNPCPVGKTAK
Starfish             KIPNKQVAIPEIH----------SGYSTPPNPCPKTMDEA
Sea_cucumber         KIPNKQVAIPEIQ----------SGYDTPPNPCPVNAKVK
                                             ******
```

(b)

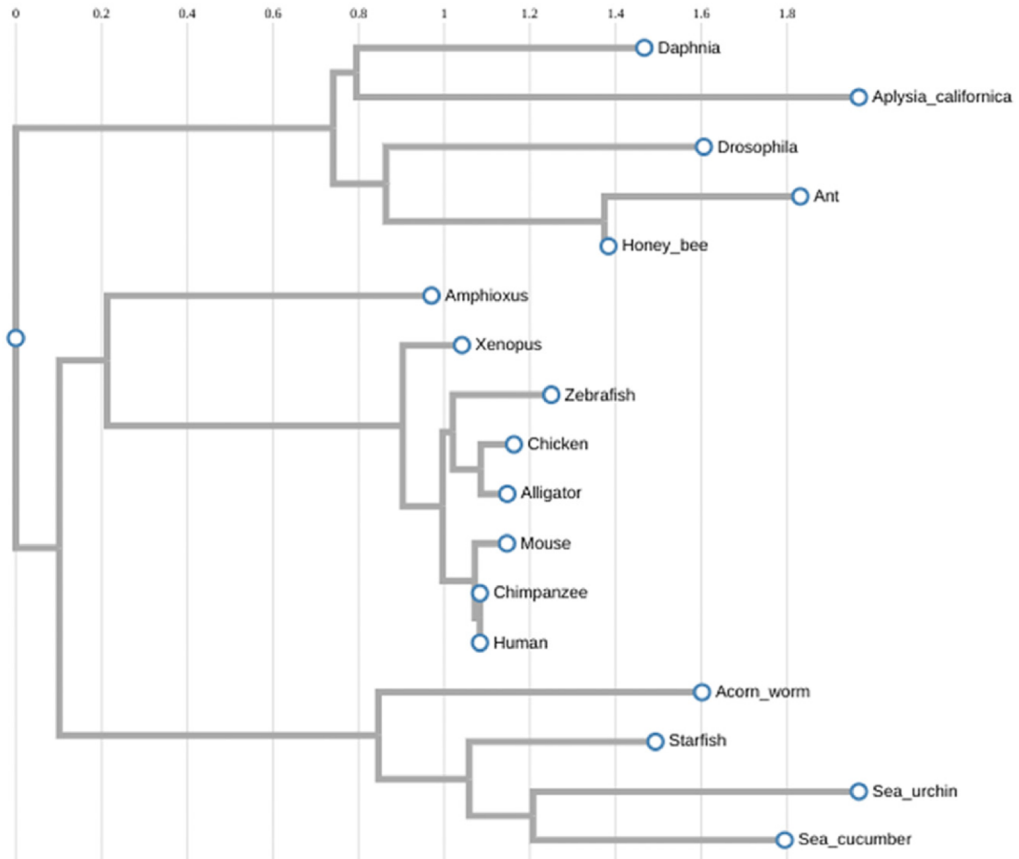


Figure 2. Evolutionary conservation of secretogranin V proteins among phyla using the ETE3 phylogenetic analysis pipeline (http://www.genome.jp/tools-bin/ete) with default settings [43]. **a** Partial sequence alignment reveals that sea urchin secretogranin V shares a highly conserved motif (marked with asterisks) with mammals, other chordates and arthropods. **b** Phylogenetic tree for secretogranin V proteins. Sequence accession numbers: Acorn worm:NP_001161654.1; Alligator:XP_006018554.1; Amphioxus:XP_019625604.1; Ant:XP_011256505.1; Aplysia_californica:NP_001191628.1; Chicken:NP_001278687.1; Chimpanzee: NP_001092019.1; Daphnia:KZS15816.1; Drosophila:NP_001014608.1; Honey bee:XP_392155.2; Human:NP_001138229.1; Mouse:NP_033188.3; Sea_cucumber:PIK61457.1; Sea urchin:XP_799858.1 (Spurp_10); Starfish:XP_022104829.1; Xenopus:NP_001085080.1; Zebrafish:NP_957020

Spurp_3, Spurp_11, and Spurp_12 were exclusively identified from the LC-ESI-MS data, while peptides from Spurp_9 and Spurp_15 were exclusively identified from the LC-MALDI-MS data.

We confirmed 16 proteins detected in our previous study [27] and found one novel protein; of the 16 proteins, nine were also identified previously by transcriptomic analysis [14]. However, we did not identify peptides from the previously identified prohormone, New_Precursor_DAGPAWYG, and this prohormone was also not reported in the sea urchin transcriptome study by Rowe and Elphick [14]. We speculate that peptides from this prohormone could have been either below the detection level of this study, localized to other parts of the nervous system not included in the current study, or expressed at a different developmental stage of the sea urchin. In addition, we confirmed that two of the predicted prohormones from the Rowe and Elphick informatics study [14], Spurp_9 and Spurp_21, are indeed translated and post-translationally processed into potential bioactive peptides.

Many of the bioinformatically identified precursor proteins in the public databases are uncharacterized or show little homology to known proteins from other species. The vast majority of the sea urchin peptides detected in this work were consistent with the molluscan prediction models [37] that were implemented using NeuroPred [38]. In contrast with other NeuroPred models, these molluscan models failed to predict cleavage at R-R sites, which is necessary to produce the five characterized peptides in this study.

Two SALMFamide precursors were identified by Elphick and coworkers [9, 48]. We confirmed two peptides via Mascot search and three peptides via PEAKS from the F-type SALMFamide precursor (Spurp_14). The F-type SALMFamide prohormone exists in the GLEAN protein collection as an LFRFamide-like precursor [11]. SALMFamide peptides act as muscle relaxants in starfish and the edible or common sea urchin *Echinus esculentus* [49, 50]. Moreover, the SALMFamide peptides have different potencies as muscle relaxants in starfish [50]. The role of F-type SALMFamide in *S. purpuratus* remains to be investigated. It will be promising and interesting to extend a physiological function study of SALMFamide peptides in *S. purpuratus*, given its proximity to starfish phylogenetically. Other than acting as muscle relaxants, the SALMFamide peptides were also found to inhibit reproductive processes in echinoderms when its sequences were compared with other reproduction-related peptides, the gonadotropin-inhibitory hormone in chordates, and SIFamide-type neuropeptides in protostomes [51]. Sequence similarities were observed across these three types of peptides, suggesting they might originate and evolve from a common ancestral signaling system regulating reproduction [48].

Many of the precursor proteins have been listed as uncharacterized proteins in the NCBI database; thus, we first looked at annotated proteins. For example, pedal peptide, which has been found to regulate locomotion in the mollusk, may be expected to be present in the radial nerves of the sea urchin and have a similar function [52]. The pedal peptide was first discovered in *A. californica* and is involved in the control of muscle contraction and cilia movement [52, 53]. Therefore, it is not surprising that similar peptides were found in our samples since we studied the nerves going to the urchin tube feet. The effects of pedal peptides on *S. purpuratus* muscle contraction and movement merit further investigation. The presence of the two pedal peptide-type prohormones in *S. purpuratus* suggests that pedal peptides originated from a common ancestor of the Mollusca and Echinodermata. In addition, Rowe and Elphick [14] found that pedal peptide-type precursors in nematodes have significant sequence similarities with the pedal peptide precursor in Mollusca and orcokinin-type neuropeptides in arthropods. This indicates the presence of pedal/orcokinin-like precursors throughout the Bilateria. Their findings sparked interest in the study of the functions of pedal/orcokinin-like peptides in different invertebrates, including deuterostomes [54].

The first thyrotropin-releasing hormone (TRH)-like protein (Spurp_11; NP_001116993.1) in invertebrates was discovered in our previous study [27], although no peptides were experimentally confirmed, and also by Rowe and Elphick [14] in their sea urchin transcriptome study. We have now detected four peptides from this prohormone using MS, including three peptides with pyroglutamination. This protein contains multiple copies of highly similar peptides with the Q[YWF]PGG sequence between dibasic sites, which closely mirrors the shorter mammalian thyroliberin peptide sequence, QHPG, also located between dibasic sites. Similar to thyroliberin, the Q[YWF]PGG peptides are predicted to be post-translationally modified with both amidation of the C-terminus and pyroglutamination of the N-terminus. N-terminal glutamine is known to undergo pyroglutamination in neuropeptides in other species [4, 55–57]. However, the modified or unmodified Q[YWF]PGG peptides were not experimentally detected here as their molecular masses are below the mass range examined in this study. TRH has been reported to control the release of thyrotropin-stimulating hormone and production of prolactin in the pituitary [58, 59], but little is known about its function in invertebrates. Burke and coworkers [13] predicted a gene encoding the TRH receptor in *S. purpuratus*, suggesting that this prohormone may produce thyroliberin-like peptides that are ligands to this receptor.

Few functions of secretogranin V in echinoderms have been investigated, but its involvement in the regulation of pituitary hormone processing and secretion has been reported in mammals [60, 61]. In addition to physiological functions, the study of secretogranin V has been extended into evolutionary aspects. Mbikay and coworkers [62] built a phylogenetic tree using sequences from human, pig, rat, *Drosophila*, *Caenorhabditis elegans*, and *Xenopus* and revealed that secretogranin V evolved at a faster rate in invertebrates than in vertebrates. According to the cladogram shown in Fig. 2b, the structure of Echinodermata secretogranin V is slightly closer to vertebrates than arthropods.

We also identified Spurp_24, a potential myomodulin, which was not originally annotated as a gene product in the

NCBI *S. purpuratus* protein database. Myomodulin peptides are known to encode the inhibitory peptides acting in muscles of Mollusca and Arthropoda [63, 64]. Whether or not peptides from the Spurp_24 protein have similar effects in sea urchin remains to be investigated.

Conclusions

The sea urchin, *S. purpuratus*, is the first animal within the phylum Echinodermata to have its genome sequenced [32]. However, a limited number of neuropeptides and prohormones have been reported, in contrast to other sequenced species such as *Drosophila melanogaster* and *S. mediterranea* [19, 28, 65]. In this work, we characterized 93 peptides and 24 precursor proteins from the radial nerves of *S. purpuratus* by combining bioinformatics and MS-based peptide discovery tools to facilitate the discovery and annotation of new and/or obscure genes, as well as structural characterization of final gene products. As we only used radial nerves, neuropeptides found exclusively in other structures such as the ring nerves would not have been characterized in our work. Nevertheless, this work represents the most comprehensive peptidomics study of the sea urchin *S. purpuratus* to date.

Acknowledgments

The projected described was supported by the National Institute on Drug Abuse under Award No. P30 DA018310, and the National Institute of Neurological Disorders and Stroke under Award No. R01 NS031609. The content is solely the responsibility of the authors and does not necessarily represent the official views of the awarding agencies.

References

1. Jing, J., Vilim, F.S., Horn, C.C., Alexeeva, V., Hatcher, N.G., Sasaki, K., Yashina, I., Zhurov, Y., Kupfermann, I., Sweedler, J.V., Weiss, K.R.: From hunger to satiety: reconfiguration of a feeding network by *Aplysia* neuropeptide y. J. Neurosci. **27**, 3490–3502 (2007)
2. Sweedler, J.V., Li, L., Rubakhin, S.S., Alexeeva, V., Dembrow, N.C., Dowling, O., Jing, J., Weiss, K.R., Vilim, F.S.: Identification and characterization of the feeding circuit-activating peptides, a novel neuropeptide family of *Aplysia*. J. Neurosci. **22**, 7797–7808 (2002)
3. Romanova, E., Rubakhin, S., Ossyra, J., Zombeck, J., Nosek, M., Sweedler, J., Rhodes, J.: Differential peptidomics assessment of strain and age differences in mice in response to acute cocaine administration. J. Neurochem. **135**, 1038–1048 (2015)
4. Lee, J.E., Atkins Jr., N., Hatcher, N.G., Zamdborg, L., Gillette, M.U., Sweedler, J.V., Kelleher, N.L.: Endogenous peptide discovery of the rat circadian clock: a focused study of the suprachiasmatic nucleus by ultrahigh performance tandem mass spectrometry. Mol. Cell. Proteomics. **9**, 285–297 (2010)
5. Hatcher, N.G., Atkins Jr., N., Annangudi, S.P., Forbes, A.J., Kelleher, N.L., Gillette, M.U., Sweedler, J.V.: Mass spectrometry-based discovery of circadian peptides. Proc. Natl. Acad. Sci. U. S. A. **105**, 12527–12532 (2008)
6. Liu, F., Baggerman, G., Schoofs, L., Wets, G.: The construction of a bioactive peptide database in *Metazoa*. J. Proteome Res. **7**, 4119–4131 (2008)
7. Semmens, D.C., Elphick, M.R.: The evolution of neuropeptide signalling: insights from echinoderms. Brief. Funct. Genomics. **16**, 288–298 (2017)
8. Satoh, N., Rokhsar, D., Nishikawa, T.: Chordate evolution and the three-phylum system. Proc. Biol. Sci. **281**, 20141729 (2014)
9. Rowe, M.L., Elphick, M.R.: Discovery of a second SALMFamide gene in the sea urchin *Strongylocentrotus purpuratus* reveals that L-type and F-type SALMFamide neuropeptides coexist in an echinoderm species. Mar. Genomics. **3**, 91–97 (2010)
10. Yun, S.-S., Thorndyke, M.C., Elphick, M.R.: Identification of novel SALMFamide neuropeptides in the starfish *Marthasterias glacialis*. Comp. Biochem. Physiol. **147**, 536–542 (2007)
11. Elphick, M.R., Price, D.A., Lee, T.D., Thorndyke, M.C.: The SALMFamides: A new family of neuropeptides isolated from an Echinoderm. Proc. Biol. Sci. **243**, 121–127 (1991)
12. Elphick, M.R.: Ng peptides: a novel family of neurophysin-associated neuropeptides. Gene. **458**, 20–26 (2010)
13. Burke, R.D., Angerer, L.M., Elphick, M.R., Humphrey, G.W., Yaguchi, S., Kiyama, T., Liang, S., Mu, X., Agca, C., Klein, W.H., Brandhorst, B.P., Rowe, M., Wilson, K., Churcher, A.M., Taylor, J.S., Chen, N., Murray, G., Wang, D., Mellott, D., Olinski, R., Hallböök, F., Thorndyke, M.C.: A genomic view of the sea urchin nervous system. Dev. Biol. **300**, 434–460 (2006)
14. Rowe, M.L., Elphick, M.R.: The neuropeptide transcriptome of a model echinoderm, the sea urchin *Strongylocentrotus purpuratus*. Gen. Comp. Endocrinol. **179**, 331–344 (2012)
15. Semmens, D.C., Mirabeau, O., Moghul, I., Pancholi, M.R., Wurm, Y., Elphick, M.R.: Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. Open Biol. **6**, 150224 (2016)
16. Smith, M.K., Wang, T., Suwansa-Ard, S., Motti, C.A., Elizur, A., Zhao, M., Rowe, M.L., Hall, M.R., Elphick, M.R., Cummins, S.F.: The neuropeptidome of the crown-of-thorns starfish, Acanthaster Planci. J. Proteome. **165**, 61–68 (2017)
17. Amare, A., Sweedler, J.V.: Neuropeptide precursors in *Tribolium castaneum*. Peptides. **28**, 1282–1291 (2007)
18. Amare, A., Hummon, A.B., Southey, B.R., Zimmerman, T.A., Rodriguez-Zas, S.L., Sweedler, J.V.: Bridging neuropeptidomics and genomics with bioinformatics: prediction of mammalian neuropeptide prohormone processing. J. Proteome Res. **5**, 1162–1167 (2006)
19. Hewes, R.S., Taghert, P.H.: Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. Genome Res. **11**, 1126–1142 (2001)
20. Hummon, A.B., Richmond, T.A., Verleyen, P., Baggerman, G., Huybrechts, J., Ewing, M.A., Vierstraete, E., Rodriguez-Zas, S.L., Schoofs, L., Robinson, G.E., Sweedler, J.V.: From the genome to the proteome: uncovering peptides in the *Apis* brain. Science. **314**, 647–649 (2006)
21. Southey, B.R., Rodriguez-Zas, S.L., Sweedler, J.V.: Characterization of the prohormone complement in cattle using genomic libraries and cleavage prediction approaches. BMC Genomics. **10**, 228–228 (2009)
22. Riehle, M.A., Garczynski, S.F., Crim, J.W., Hill, C.A., Brown, M.R.: Neuropeptides and peptide hormones in *Anopheles gambiae*. Science. **298**, 172–175 (2002)
23. Romanova, E.V., Sweedler, J.V.: Peptidomics for the discovery and characterization of neuropeptides and hormones. Trends Pharmacol. Sci. **36**, 579–586 (2015)
24. Schulz-Knappe, P., Schrader, M., Zucht, H.D.: The peptidomics concept. Comb. Chem. High Throughput Screen. **8**, 697–704 (2005)
25. Boonen, K., Landuyt, B., Baggerman, G., Husson, S.J., Huybrechts, J., Schoofs, L.: Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis. J. Sep. Sci. **31**, 427–445 (2008)
26. Menschaert, G., Vandekerckhove, T.T., Baggerman, G., Schoofs, L., Luyten, W., Van Criekinge, W.: Peptidomics coming of age: a review of contributions from a bioinformatics angle. J. Proteome Res. **9**, 2051–2061 (2010)
27. Menschaert, G., Vandekerckhove, T., Baggerman, G., Landuyt, B., Sweedler, J., Schoofs, L., Luyten, W., Van Criekinge, W.: A hybrid, de novo based, genome-wide database search approach applied to the sea urchin neuropeptidome. J. Proteome Res. **9**, 990–996 (2010)
28. Collins, J.J., Hou, X., Romanova, E.V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J.V., Newmark, P.A.: Genome-wide analyses reveal a role for peptide hormones in planarian germline development. PLoS Biol. **8**, e1000509 (2010)
29. Hu, C.K., Southey, B.R., Romanova, E.V., Maruska, K.P., Sweedler, J.V., Fernald, R.D.: Identification of prohormones and pituitary

neuropeptides in the African cichlid, *Astatotilapia burtoni*. BMC Genomics. **17**, 660 (2016)
30. Xie, F., London, S.E., Southey, B.R., Annangudi, S.P., Amare, A., Rodriguez-Zas, S.L., Clayton, D.F., Sweedler, J.V.: The zebra finch neuropeptidome: prediction, detection and expression. BMC Biol. **8**, 28–28 (2010)
31. Yin, P., Hou, X., Romanova, E.V., Sweedler, J.V.: Neuropeptidomics: mass spectrometry-based qualitative and quantitative analysis. Methods Mol. Biol. **789**, 223–236 (2011)
32. Sea Urchin Genome Sequencing Consortium, Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R.C., Angerer, L.M., Arnone, M.I., Burgess, D.R., Burke, R.D., Coffman, J.A., Dean, M., Elphick, M.R., Ettensohn, C.A., Foltz, K.R., Hamdoun, A., Hynes, R.O., Klein, W.H., Marzluff, W., McClay, D.R., Morris, R.L., Mushegian, A., Rast, J.P., Smith, L.C., Thorndyke, M.C., Vacquier, V.D., Wessel, G.M., Wray, G., Zhang, L., Elsik, C.G., Ermolaeva, O., Hlavina, W., Hofmann, G., Kitts, P., Landrum, M.J., Mackey, A.J., Maglott, D., Panopoulou, G., Poustka, A.J., Pruitt, K., Sapojnikov, V., Song, X., Souvorov, A., Solovyev, V., Wei, Z., Whittaker, C.A., Worley, K., Durbin, K.J., Shen, Y., Fedrigo, O., Garfield, D., Haygood, R., Primus, A., Satija, R., Severson, T., Gonzalez-Garay, M.L., Jackson, A.R., Milosavljevic, A., Tong, M., Killian, C.E., Livingston, B.T., Wilt, F.H., Adams, N., Bellé, R., Carbonneau, S., Cheung, R., Cormier, P., Cosson, B., Croce, J., Fernandez-Guerra, A., Geneviève, A.-M., Goel, M., Kelkar, H., Morales, J., Mulner-Lorillon, O., Robertson, A.J., Goldstone, J.V., Cole, B., Epel, D., Gold, B., Hahn, M.E., Howard-Ashby, M., Scally, M., Stegeman, J.J., Allgood, E.L., Cool, J., Judkins, K.M., McCafferty, S.S., Musante, A.M., Obar, R.A., Rawson, A.P., Rossetti, B.J., Gibbons, I.R., Hoffman, M.P., Leone, A., Istrail, S., Materna, S.C., Samanta, M.P., Stolc, V., Tongprasit, W., Tu, Q., Bergeron, K.-F., Brandhorst, B.P., Whittle, J., Berney, K., Bottjer, D.J., Calestani, C., Peterson, K., Chow, E., Yuan, Q.A., Elhaik, E., Graur, D., Reese, J.T., Bosdet, I., Heesun, S., Marra, M.A., Schein, J., Anderson, M.K., Brockton, V., Buckley, K.M., Cohen, A.H., Fugmann, S.D., Hibino, T., Loza-Coll, M., Majeske, A.J., Messier, C., Nair, S.V., Pancer, Z., Terwilliger, D.P., Agca, C., Arboleda, E., Chen, N., Churcher, A.M., Hallböök, F., Humphrey, G.W., Idris, M.M., Kiyama, T., Liang, S., Mellott, D., Mu, X., Murray, G., Olinski, R.P., Raible, F., Rowe, M., Taylor, J.S., Tessmar-Raible, K., Wang, D., Wilson, K.H., Yaguchi, S., Gaasterland, T., Galindo, B.E., Gunaratne, H.J., Juliano, C., Kinukawa, M., Moy, G.W., Neill, A.T., Nomura, M., Raisch, M., Reade, A., Roux, M.M., Song, J.L., Su, Y.-H., Townley, I.K., Voronina, E., Wong, J.L., Amore, G., Branno, M., Brown, E.R., Cavalieri, V., Duboc, V., Duloquin, L., Flytzanis, C., Gache, C., Lapraz, F., Lepage, T., Locascio, A., Martinez, P., Matassi, G., Matranga, V., Range, R., Rizzo, F., Röttinger, E., Beane, W., Bradham, C., Byrum, C., Glenn, T., Hussain, S., Manning, G., Miranda, E., Thomason, R., Walton, K., Wikramanayake, A., Wu, S.-Y., Xu, R., Brown, C.T., Chen, L., Gray, R.F., Lee, P.Y., Nam, J., Oliveri, P., Smith, J., Muzny, D., Bell, S., Chacko, J., Cree, A., Curry, S., Davis, C., Dinh, H., Dugan-Rocha, S., Fowler, J., Gill, R., Hamilton, C., Hernandez, J., Hines, S., Hume, J., Jackson, L., Jolivet, A., Kovar, C., Lee, S., Lewis, L., Miner, G., Morgan, M., Nazareth, L.V., Okwuonu, G., Parker, D., Pu, L.-L., Thorn, R., Wright, R.: The genome of the sea urchin *Strongylocentrotus purpuratus*. Science. **314**, 941–952 (2006)
33. NCBI Resource Coordinators: Database resources of the national center for biotechnology information. Nucleic Acids Res. **45**, D12–d17 (2017)
34. The UniProt Consortium: Uniprot: The universal protein knowledgebase. Nucleic Acids Res. **45**, D158–d169 (2017)
35. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J.: Basic local alignment search tool. J. Mol. Biol. **215**, 403–410 (1990)
36. Dyrløv Bendtsen, J., Nielsen, H., von Heijne, G., Brunak, S.: Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340**, 783–795 (2004)
37. Hummon, A.B., Hummon, N.P., Corbin, R.W., Li, L., Vilim, F.S., Weiss, K.R., Sweedler, J.V.: From precursor to final peptides: a statistical sequence-based approach to predicting prohormone processing. J. Proteome Res. **2**, 650–656 (2003)
38. Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., Sweedler, J.V.: Neuropred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. Nucleic Acids Res. **34**, W267–W272 (2006)
39. Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. **20**, 3551–3567 (1999)
40. Merkler, D.J.: C-terminal amidated peptides: production by the in vitro enzymatic amidation of glycine-extended peptides and the importance of the amide to bioactivity. Enzym. Microb. Technol. **16**, 450–456 (1994)
41. Prigge, S.T., Mains, R.E., Eipper, B.A., Amzel, L.M.: New insights into copper monooxygenases and peptide amidation: structure, mechanism and function. Cell. Mol. Life Sci. **57**, 1236–1259 (2000)
42. Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A., Ma, B.: Peaks db: de novo sequencing assisted database search for sensitive and accurate peptide identification. Mol. Cell. Proteomics. **11**, M111.010587 (2012)
43. Huerta-Cepas, J., Serra, F., Bork, P.: Ete 3: reconstruction, analysis, and visualization of phylogenomic data. Mol. Biol. Evol. **33**, 1635–1638 (2016)
44. Motokawa, T., Fuchigami, Y.: Coordination between catch connective tissue and muscles through nerves in the spine joint of the sea urchin Diadema Setosum. J. Exp. Biol. **218**, 703–710 (2015)
45. Ullrich-Luter, E.M., Dupont, S., Arboleda, E., Hausen, H., Arnone, M.I.: Unique system of photoreceptors in sea urchin tube feet. Proc. Natl. Acad. Sci. U. S. A. **108**, 8367–8372 (2011)
46. Ma, M., Wang, J., Chen, R., Li, L.: Expanding the crustacean neuropeptidome using a multifaceted mass spectrometric approach. J. Proteome Res. **8**, 2426–2437 (2009)
47. Ye, H., Wang, J., Zhang, Z., Jia, C., Schmerberg, C., Catherman, A.D., Thomas, P.M., Kelleher, N.L., Li, L.: Defining the neuropeptidome of the spiny lobster *Panulirus interruptus* brain using a multidimensional mass spectrometry-based platform. J. Proteome Res. **14**, 4776–4791 (2015)
48. Elphick, M.R., Achhala, S., Martynyuk, N.: The evolution and diversity of SALMFamide neuropeptides. PLoS One. **8**, e59076 (2013)
49. Elphick, M.R., Melarange, R.: Neural control of muscle relaxation in echinoderms. J. Exp. Biol. **204**, 875–885 (2001)
50. Elphick, M.R., Thorndyke, M.C.: Molecular characterisation of SALMFamide neuropeptides in sea urchins. J. Exp. Biol. **208**, 4273–4282 (2005)
51. Elphick, M.R.: From gonadotropin-inhibitory hormone to SIFamides: are echinoderm SALMFamides the "missing link" in a bilaterian family of neuropeptides that regulate reproductive processes? Gen. Comp. Endocrinol. **193**, 229–233 (2013)
52. Hall, J.D., Lloyd, P.E.: Involvement of pedal peptide in locomotion in *Aplysia*: modulation of foot muscle contractions. J. Neurobiol. **21**, 858–868 (1990)
53. Lloyd, P.E., Connolly, C.M.: Sequence of pedal peptide: a novel neuropeptide from the central nervous system of *Aplysia*. J. Neurosci. **9**, 312–317 (1989)
54. Lin, M., Egertova, M., Zampronio, C.G., Jones, A.M., Elphick, M.R.: Pedal peptide/orcokinin-type neuropeptide signaling in a deuterostome: The anatomy and pharmacology of starfish myorelaxant peptide in *Asterias rubens*. J. Comp. Neurol. **525**, 3890–3917 (2017)
55. Jimenez, C.R., Spijker, S., de Schipper, S., Lodder, J.C., Janse, C.K., Geraerts, W.P., van Minnen, J., Syed, N.I., Burlingame, A.L., Smit, A.B., Li, K.: Peptidomics of a single identified neuron reveals diversity of multiple neuropeptides with convergent actions on cellular excitability. J. Neurosci. **26**, 518–529 (2006)
56. Sigafoos, J., Chestnut, W.G., Merrill, B.M., Taylor, L.C., Diliberto Jr., E.J., Viveros, O.H.: Novel peptides from adrenomedullary chromaffin vesicles. J. Anat. **183**(Pt 2), 253–264 (1993)
57. Romanova, E.V., McKay, N., Weiss, K.R., Sweedler, J.V., Koester, J.: Autonomic control network active in *Aplysia* during locomotion includes neurons that express splice variants of R15-neuropeptides. J. Neurophysiol. **97**, 481–491 (2007)
58. Tashjian Jr., A.H., Barowsky, N.J., Jensen, D.K.: Thyrotropin releasing hormone: direct evidence for stimulation of prolactin production by pituitary cells in culture. Biochem. Biophys. Res. Commun. **43**, 516–523 (1971)
59. Schalch, D.S., Gonzalez-Barcena, D., Kastin, A.J., Schally, A.V., Lee, L.A.: Abnormalities in the release of TSH in response to thyrotropin-releasing hormone (TRH) in patients with disorders of the pituitary, hypothalamus and basal ganglia. J. Clin. Endocrinol. Metab. **35**, 609–615 (1972)
60. Westphal, C.H., Muller, L., Zhou, A., Zhu, X., Bonner-Weir, S., Schambelan, M., Steiner, D.F., Lindberg, I., Leder, P.: The neuroendocrine protein 7B2 is required for peptide hormone processing in vivo and provides a novel mechanism for pituitary Cushing's disease. Cell. **96**, 689–700 (1999)

61. Huttner, W.B., Natori, S.: Regulated secretion: helper proteins for neuroendocrine secretion. Curr. Biol. **5**, 242–245 (1995)
62. Mbikay, M., Seidah, N.G., Chrétien, M.: Neuroendocrine secretory protein 7B2: structure, expression and functions. Biochem. J. **357**, 329–342 (2001)
63. Jékely, G.: Global view of the evolution and diversity of metazoan neuropeptide signaling. Proc. Natl. Acad. Sci. U. S. A. **110**, 8702–8707 (2013)
64. Orekhova, I.V., Alexeeva, V., Church, P.J., Weiss, K.R., Brezina, V.: Multiple presynaptic and postsynaptic sites of inhibitory modulation by myomodulin at ARC neuromuscular junctions of *Aplysia*. J. Neurophysiol. **89**, 1488–1502 (2003)
65. Vanden Broeck, J.: Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*. Peptides. **22**, 241–254 (2001)