Molecular and mass spectral identification of the broadly conserved decapod crustacean neuropeptide pQIRYHQCYFNPISCF: the first PISCF-allatostatin (*Manduca sexta*- or C-type allatostatin) from a non-insect

Elizabeth A. Stemmler[1,*], **Emily A. Bruns**[1], **Christopher R. Cashman**[1,2], **Patsy S. Dickinson**[2], and **Andrew E. Christie**[3,*]

[1]Department of Chemistry, Bowdoin College, 6600 College Station, Brunswick, Maine 04011 USA

[2]Department of Biology, Bowdoin College, 6500 College Station, Brunswick, Maine 04011 USA

[3]Center for Marine Functional Genomics, Mount Desert Island Biological Laboratory, P.O. Box 35, Old Bar Harbor Road, Salisbury Cove, Maine 04672 USA

Abstract

The PISCF-allatostatins (*Manduca sexta*- or C-type allatostatins) are a family of pentadecapeptides characterized by a pyroglutamine blocked N-terminus, an unamidated –PISCF C-terminus, and a disulfide bridge between two internal Cys residues. Several isoforms of PISCF-AST are known, all from holometabolous insects. Using a combination of transcriptomics and mass spectrometry, we have identified the first PISCF-type peptides from a non-insect species. *In silico* analysis of crustacean ESTs identified several *Litopenaeus vannamei* (infraorder Penaeidea) transcripts encoding putative PISCF-AST precursors. Translation of these ESTs, with subsequent prediction of their putative post-translational processing, revealed the existence of as many as three PISCF-type peptides, including pQIRYHQCYFNPISCF (disulfide bridging between Cys^7 and Cys^{14}). Although none of the predicted isoforms was detected by mass spectrometry in *L. vannamei*, MALDI-FTMS mass profiling identified an *m/z* signal corresponding to pQIRYHQCYFNPISCF (disulfide bridge present) in neural tissue from 28 other decapods, which included members of six infraorders (Stenopodidea, Astacidea, Thalassinidea, Achelata, Anomura and Brachyura). Further characterization of the peptide using SORI-CID and chemical derivatization/enzymatic digestion supported the theorized structure. In both the crab *Cancer borealis* and the lobster *Homarus americanus*, MALDI-based tissue surveys suggest that pQIRYHQCYFNPISCF is broadly distributed in the nervous system; it was also detected in the posterior midgut caecum. Collectively, our data show that members of the PISCF-AST family are not restricted to the holometabolous insects, but instead may be broadly conserved within the Pancrustacea. Moreover, our data suggest that one highly conserved PISCF-type peptide, pQIRYHQCYFNPISCF, is present in decapod crustaceans, functioning as a brain-gut paracrine/ hormone.

[*]Correspondence to either: Dr. Elizabeth A. Stemmler, Department of Chemistry, Bowdoin College, 6600 College Station, Brunswick, ME 04011. Phone: 207-725-3633; FAX: 207-725-3405; Email: estemmle@bowdoin.edu, Dr. Andrew E. Christie, Center for Marine Functional Genomics, Mount Desert Island Biological Laboratory, P.O. Box 35, Old Bar Harbor Road, Salisbury Cove, ME 04672. Phone: 207-288-9880 ext. 284; FAX: 207-288-2130; Email: achristi@mdibl.org.

Keywords

C-type allatostatin (C-AST); *Manduca sexta*-type allatostatin; PISCF-allatostatin; expressed sequence tag (EST); matrix-assisted laser desorption/ionization Fourier transform mass spectrometry (MALDI-FTMS); neurohormone; neuropeptide; sustained off-resonance irradiation collision-induced dissociation (SORI-CID)

1. Introduction

In insects, three structurally distinct families of neuropeptides have been shown to inhibit the production of juvenile hormone by the corpora allata. Due to this shared activity, these peptide families are commonly known referred to as allatostatins (ASTs; Stay and Tobe, 2007). Both the A-type (typified by the carboxy [C]-terminal motif –YXFGLamide) and the B- or cricket-type (typified by the C-terminal motif –WX_6Wamide) ASTs appear to be broadly conserved within the Arthropoda, including numerous isoforms of both groups described from decapod crustaceans (Duve et al., 1997; Dircksen et al., 1999; Duve et al., 2002; Huybrechts et al., 2003; Fu et al., 2005; Yasuda-Kamatani and Yasuda, 2006; Yin et al., 2006; Fu et al., 2007; Christie et al., 2008a; Christie et al., 2008b; Ma et al., 2008; Gard et al., 2009; Ma et al., 2009). In contrast, authentic PISCF-ASTs, also known as *Manduca sexta*- or C-type allatostatins, (typified by a pyroglutamine blocked amino [N]-terminus, an unamidated -PISCF C-terminus and a disulfide bridge between two internal Cys residues) have thus far been identified only in insects, and there, only from holometablous species, *i.e.* dipterans and lepidopterans (Stay and Tobe, 2007). Interestingly, in the honeybee *Apis mellifera*, several hemimetabolous insects, and a number of crustaceans (including numerous members of the Decapoda), where no authentic PISCF-ASTs have thus far been identified, a PISCF-AST-like peptide, SYWKQCAFNAVSCFamide (disulfide bridge present between the two Cys residues), is present (Hummon et al., 2006; Dickinson et al., 2009; Gard et al., 2009), leading to the hypothesis that it may be the functional homolog of the PISCF-AST family in these animals (Dickinson et al., 2009).

In the study presented here, we utilized a strategy combining *in silico* transcriptome mining and high-mass-resolution matrix-assisted laser desorption/ionization Fourier transform mass spectrometry (MALDI-FTMS), in combination with chemical derivatization/enzymatic digestion and sustained off-resonance irradiation collision-induced dissociation (SORI-CID), to test the hypothesis that PISCF-ASTs are also present in members of the Decapoda. Specifically, the sequence of a known insect prepro-PISCF-AST was used to query the extant database of crustacean expressed sequence tags (ESTs) for putative orthologs, with the peptides encoded by the identified transcripts predicted using several online peptide processing programs and by homology to known insect PISCF-type isoforms. Using this approach, several putative PISCF-ASTs were predicted, including pQIRYHQCYFNPISCF (disulfide bridging between Cys[7] and Cys[14]), from the penaeid shrimp *Litopenaeus vannamei*. Using MALDI-FTMS, we searched for masses corresponding to those of the predicted *L. vannamei* PISCF-type peptides in a wide range of decapod crustaceans. While none of the PISCF-type isoforms was detected in *L. vannamei* itself, a mass corresponding to the theoretical *m/z* of pQIRYHQCYFNPISCF (disulfide bridge present) was seen in neural tissues of 28 other decapod species, which included members of six of the infraorders that comprise this taxon. MALDI-based sequencing of the *m/z* peak corresponding to pQIRYHQCYFNPISCF from the crab *Cancer borealis* pericardial organ (PO), conducted using SORI-CID, supports the predicted structure of this peptide, including its disulfide bridge, as does chemical derivatization and enzymatic digestion of the peptide from the PO of the crab *Pugettia producta*. In *C. borealis* and the lobster *Homarus americanus*, extensive tissue surveys were carried out, identifying pQIRYHQCYFNPISCF both in regions of

synaptic neuropil and in neuroendocrine organs in both species, as well as in midgut epithelial tissue, suggesting that this peptide functions both as a locally-released paracrine and as a circulating hormone. Taken collectively, the data presented here illustrate that authentic members of the PISCF-AST family are not restricted to the holometabolous insects, but instead are shared with members of the Decapoda, where a single, perhaps ubiquitously conserved, brain-gut isoform appears present.

2. Materials and methods

2.1. Animals

The animals used in our study included the following: infraorder Penaeidae - *Farfantepenaeus duorarum* (purchased from Gulf Specimen Marine Laboratories, Panacea, FL, USA) and *Litopenaeus vannamei* (purchased from Greene Prairie Aquafarms, Boligee, AL, USA, or Island Aquaculture, Kaneohe, HI, USA); infraorder Caridea - *Crangon septemspinosa* (collected by hand at Mount Desert Island Biological Laboratory, Salisbury Cove, ME, USA), *Pandalus danae* (collected by hand at Friday Harbor Laboratories, Friday Harbor, WA, USA) and *Pandalus platyceros* (collected by trap or dredge, San Juan Island area, WA, USA); infraorder Stenopodidea - *Stenopus hispidus* (purchased from That Pet Place, Lancaster, PA, USA); infraorder Astacidea - *Cherax quadricarinatus* (purchased from Stick-Fins Fish Farm, Elkton, FL, USA)*, Homarus americanus* (purchased from local suppliers; Brunswick/Bar Harbor, ME, USA)*, Homarus gammarus* (purchased from Scottish Wild Harvest, Plainfield, NJ, USA)*, Nephrops norvegicus* (purchased from Scottish Wild Harvest)*, Pacifastacus leniusculus* (collected by trap from Lake Washington, Seattle, WA, USA) and *Procambarus clarkii* (purchased from Carolina Biological Supply Company, Burlington, NC, USA); infraorder Thalassinidea - *Callianassa californiensis* (purchased from Ray's Bait Works, Snohomish, WA, USA); infraorder Achelata - *Panulirus interruptus* (purchased from Tomlinson Commercial Fishing, San Diego, CA, USA)*, Panulirus versicolor* (purchased from That Pet Place) and *Scyllarides latus* (purchased from That Pet Place); infraorder Anomura - *Clibanarius vittatus* (purchased from Gulf Specimens Marine Laboratories)*, Lithodes maja* (purchased from local suppliers; Brunswick/Bar Harbor, ME)*, Pachycheles rudis* (collected by hand at Neah Bay, WA, USA)*, Pagurus acadianus* (collected by hand at Mount Desert Island Biological Laboratory)*, Pagurus granosimanus* (collected by hand throughout the great Puget Sound area), *Pagurus pollicaris* (collected by hand at Rocky Neck State Park, Niantic, CT, USA), *Petrolisthes cinctipes* (collected by hand at Neah Bay) and *Petrolisthes eriomerus* (collected by hand throughout the great Puget Sound area, WA, USA); infraorder Brachyura - *Cancer antennarius* (collected by hand at Neah Bay), *Cancer borealis* (purchased from local suppliers, Brunswick/Bar Harbor, ME)*, Cancer gracilis* (collected by hand at False Bay, San Juan Island, WA, USA)*, Cancer irroratus* (purchased from local suppliers, Brunswick/Bar Harbor, ME)*, Cancer magister* (collected by hand or trap throughout the great Puget Sound area)*, Cancer productus* (collected by hand or trap throughout the great Puget Sound area)*, Carcinus maenas* (collected by hand at Mount Desert Island Biological Laboratory)*, Hemigrapsus nudus* (collected by hand throughout the great Puget Sound area)*, Lophopanopeus bellus* (collected by hand throughout the great Puget Sound area)*, Ovalipes ocellatus* (collected by hand at Rocky Neck State Park)*, Pugettia gracilis* (collected by hand at Friday Harbor Laboratories)*, Pugettia producta* (collected by hand at Friday Harbor Laboratories) and *Scyra acutifrons* (collected by hand at Friday Harbor Laboratories).

With the following exceptions, animals were maintained in aerated natural seawater aquaria at 8-12 °C. For *F. duorarum, L. vannamei, S. hispidus, P. versicolor, S. latus* and *C. vittatus*, aerated natural seawater aquaria were held at 18-20 °C, while those for *P. interruptus* were maintained at approximately 15 °C. *C. californiensis* were maintained in seawater-moistened wood shavings at 10 °C. *C. quadricarinatus* and *P. clarkii* were held in aerated tanks of aged

tap water at 18-20 °C, while *P. leniusculus* were maintained in aged tap water tanks at 10 °C.

All animals were anesthetized by packing in ice for 30-60 minutes prior to dissections. Tissues were removed in cold (approximately 10 °C) physiological saline appropriate to the species (compositions in mM): for Penaeidea, Caridea, Stenopodidea and Achelata, 479 NaCl, 12.8 KCl, 13.7 CaCl$_2$, 3.9 Na$_2$SO$_4$, 10 MgSO$_4$, 11 Trizma base, and 4.8 maleic acid; pH 7.5–7.6; for salt water Astacidea (*Homarus* and *Nephrops* species) and Anomura, 479 NaCl, 12.8 KCl, 13.7 CaCl$_2$, 20 MgSO$_4$, 3.9 Na$_2$SO$_4$, and 4.8 HEPES (pH 7.4–7.5); for freshwater Astacidea (crayfish species), 200 NaCl, 5.4 KCl, 17.2 CaCl$_2$, 5.5 MgCl$_2$, 22 Tris base and 4.7 maleic acid, pH 7.2–7.4; for Thalassinidea and Brachyura, 442 NaCl, 11 KCl, 13 CaCl$_2$, 26 MgCl$_2$, 12 Trizma base, and 1.2 maleic acid (pH 7.4–7.6).

2.2. Functional genomics

2.2.1. Database searches—Transcriptome searches were conducted using methods modified from several recent publications (Christie, 2008a; Christie, 2008b; Christie et al. 2008; Dickinson et al., 2009; Gard et al., 2009; Ma et al., 2009). Specifically, the online program tblastn (NCBI; http://www.ncbi.nlm.nih.gov/BLAST/) was used to mine for ESTs encoding putative crustacean PISCF-AST precursors via queries using known insect prepro-hormone sequences. For all searches, the program database was set to non-human, non-mouse ESTs (EST_others) and was restricted to crustacean sequences (taxid:6657). All hits were fully translated and checked manually for homology to the target query (see below).

2.2.2. Peptide prediction—Translation of the nucleotide sequences of the identified ESTs was performed using the Translate tool of ExPASy (Swiss Institute of Bioinformatics, Basel, Switzerland; http://www.expasy.ch/tools/dna.html). Signal peptide prediction was done via the online program SignalP 3.0, using both the Neural Networks and the Hidden Markov Models algorithms (Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark; http://www.cbs.dtu.dk/services/SignalP/ (Bendtsen et al., 2004). Pro-hormone convertase cleavage sites were predicted based on the information presented in Veenstra (2000). Post-translational modifications, *e.g.* cyclization of N-terminal Gln residues and disulfide bridging between Cys residues, were predicted by homology to known PISCF-type peptides.

2.3. Mass spectrometry

2.3.1. Sample preparation

2.3.1.1. Direct tissue analyses: To prepare samples for direct tissue MALDI-FTMS, we first isolated either small pieces of a larger tissue sample, *i.e.* the supraoesophageal ganglion (brain), the pericardial organ (PO), the eyestalk ganglia (including the sinus gland [SG]) or the posterior midgut caecum (PMC), or the entire tissue, *i.e.* the commissural ganglion (CoG), the stomatogastric ganglion (STG) or the SG itself, using manual micro-dissection techniques. The ganglionic sheath surrounding the brain, CoG and STG was removed with further manual microdissection. The isolated tissue was then removed from the saline with fine forceps, rinsed sequentially in two 20 µL droplets of 0.75 M fructose (SigmaAldrich, St. Louis, MO, USA; 99%) and placed on one face of a ten-faceted stainless steel probe tip, minimizing co-transfers of solution. The tissue was then sliced 10-20 times with a 0.2 mm needle, gathered together and covered with a 0.5 µL droplet of 1.0 M 2,5-dihydroxybenzoic acid (DHB; SigmaAldrich; 98%, sublimed prior to use) prepared in 1:1 acetonitrile (Fisher Scientific, Pittsburg, PA, USA; HPLC grade):water containing 2 % (v/v) phosphoric acid (SigmaAldrich, 99%).

2.3.1.2. Analyses of tissue extracts: To prepare tissue extracts for mass spectral analyses, a small piece of PO, paired, desheathed CoGs, an entire eyestalk, or brain were removed from saline, rinsed sequentially in two 20 μL droplets of 0.75 M fructose and placed in a 0.6 mL tube with 30 μL of extraction solvent (7% acetic acid, 64% methanol, 29% deionized H_2O). The tissue was homogenized by cutting with spring scissors. The homogenate was sonicated for 2 minutes and centrifuged at 2200 g for 5 minutes in a microcentrifuge (Fischer Scientific). The supernatant was saved, and the pellet resuspended with 5 μL of deionized water. The sonication, centrifugation, and resuspension steps were repeated two additional times. The supernatants of all cycles were combined. Deionized H_2O (20 μL) and $CDCl_3$ (25 μL, SigmaAldrich) were added to the solution. The organic layer was removed, and the aqueous layer was evaporated to dryness. For most samples, the resultant extracts were desalted using C18 ZipTip pipette tips (Millipore, Billerica, MA, USA). After their preparation, 0.5 μL of extract was mixed with 0.5 μL of DHB matrix on one face of the MALDI probe and the extract-matrix mixture was allowed to co-crystallize.

2.3.2. Chemical and enzymatic reactions—Methyl esterification was performed by adding 10 μL of methanolic HCl to the evaporated tissue extracts. The methanolic HCl was prepared immediately prior to its use by adding 400 μL acetyl chloride (Alltech, Deerfield, IL, USA) dropwise to 2.5 mL methanol (Alltech) on ice. The methanolic HCl/tissue extract mixture was allowed to react for 2 hours at room temperature. Acetylation was achieved by adding acetic anhydride (5 μl; Alltech) and nanopure water (2.5 μl) to a dried extract from a single PO for one hour. Enzymatic digestion used pyroglutamate amino-peptidase (10 mU; SigmaAldrich) reconstituted in 50 ml of buffer solution (50 mM sodium phosphate, pH 7.0, 10 mM DDT, 1 mM EDTA). Enzyme solution (5 μl) was added to a dried extract from a single PO and reacted overnight at 37 °C. Following chemical or enzymatic treatment, samples were dried and reconstituted in 5 μl 1:1 acetonitrile: nanopure water. In some cases, the treated extracts were desalted using C18 ZipTip pipette tips. Samples were prepared for MALDI-FTMS as described in Section *2.3.1.2*

2.3.3. Instrumentation—Samples were analyzed using a HiResMALDI Fourier transform mass spectrometer (IonSpec, Lake Forest, CA, USA) equipped with a Cryomagnetics (Oak Ridge, TN, USA) 4.7 Tesla actively-shielded superconducting magnet. Ions were generated using a pulsed nitrogen laser (337 nm) and were transported from the external ion source to the closed cylindrical ICR cell using a quadrupole ion guide. The ion guide radio frequency potential and trapping delay time were optimized to transmit and trap ions of a selected mass range (optimized for *m/z* 1500 or 2500). A pulse of argon was introduced to the vacuum system during trapping to elevate the system pressure transiently for collisional cooling. All spectra were measured using ion accumulation techniques, where ions from seven to thirty successive laser shots were accumulated in the cell. Exact mass measurements were calibrated using either the internal calibration on adjacent samples (InCAS) technique (O'Connor and Costello, 2000), modified to include the accumulation of mass-selected calibrant ions or calibration with previously identified peptides (GYRKPPFNGSIFamide, VYRKPPFNGSIFamide, APSGFLGMRamide, and pQDLDHVFLRFamide; Stemmler et al., 2007a). For calibration of samples from brachyurans and anomurans, NFDEIDRSGFGFA and its fragment ions were also used for calibration (Dickinson et al., 2009). A delay of 5-10 s preceded ion detection, which occurred with analyzer pressures of 1 to 5×10^{-10} Torr. Transients from direct tissue spectra were apodized using a Blackman function and zero-filled prior to fast Fourier transformation.

For SORI-CID experiments, argon was used as the collision gas, the frequency offset was set equal to -1.5% of the reduced cyclotron frequency, and the voltage amplitude was 6.5 V_{bp}. Transients from SORI-CID spectra were processed without apodization and calibrated with a one-point calibration using the $[M+H]^+$ ion. Synthetic pQIRYHQCYFNPISCF

(disulfide bridge present between the cysteine residues) was custom synthesized by GenScript Corporation (Piscataway, NJ, USA). The Boston University Data Analysis (B.U.D.A.) software, provided by Dr. Peter O'Connor (Boston University School of Medicine), was used for the analysis of FT mass spectral data. Predicted isotopic distributions of the $[M+H]^+$ ion for pQIRYHQCYFNPISCF (disulfide bridge present between the Cys residues isotopic distribution for the $[M+H]^+$ ion for pQIRYHQCYFNPISCF were calculated using the program Exact Mass Calculator, Version 8.0.28, from (IonSpec Corporation).

3. Results

3.1. Identification of PISCF-allatostatins in Litopenaeus vannamei using transcriptomics

Four *L. vannamei* ESTs (Table 1) were identified as encoding putative PISCF-AST precursors via queries with the sequence of a *Drosophila melanogaster* prepro-PISCF-AST (accession no. AAK40100; Williamson et al., 2001). Translation of these shrimp transcripts revealed FE182974 and FE175093 to encode nearly identical, 139 amino acid, putative full-length prepro-hormones (differing only in a Lys vs Phe residue at position 106), with FE182975 and FE180026 encoding similar, though not identical, putative C-terminal partial pro-hormones of 26 and 18 amino acids, respectively. Specifically the partial pro-hormones encoded by FE182975 and FE180026 differed from the full-length precursors at one and two residues, respectively (*i.e.* Asp^{114} for Gly^{114} in FE182975 and Ile^{131} for Asn^{131} and Val^{133} for Ile^{133} in FE180026; numbering based on the sequence of the full-length prepro-hormones; Fig. 1). SignalP analysis of the full-length precursors identified the first 20 amino acids of each as a signal peptide, with the cleavage locus predicted between Ala^{20} and Ser^{21} in both proteins (Fig. 1). Within the remaining pro-hormone, seven prohormone convertase processing sites were identified (three Lys-Arg, three Arg-Arg, and one Lys-Lys; Fig. 1), cleavage at which, followed by carboxypeptidase activity, is predicted to produce six peptides, listed in their order of appearance in the pro-hormone: SPAPQDKPLEAAPNQPQHPVHLQ, AAPQDSSPEQDLAAFKDLLVAQVAAEL, SWQEPSAL, TLVDEDGLEEEQAAQS, MLAPLSGLPGELPTI (in FE182974) or MFAPLSGLPGELPTI (in FE175093), and QIRYHQCYFNPISCF, the latter peptide likely undergoing cyclization of its N-terminal Gln to become pQIRYHQCYFNPISCF. pQIRYHQCYFNPISCF is also predicted from the partial precursor deduced from FE182975, with a second PISCF-type isoform, pQIRYHQCYFIPVSCF, predicted from the partial pro-hormone deduced from FE180026. Homology to the known insect PISCF-ASTs predicts disulfide bridging between the Cys residues present in both pQIRYHQCYFNPISCF and pQIRYHQCYFIPVSCF, which would result in both peptides possessing all of the hallmarks of authentic PISCF-ASTs, *i.e.* a pyroglutamine blocked N-terminus, the unamidated C-terminal motif –PI/VSCF and a disulfide bridge between the Cys residues located at positions 7 and 14.

In addition to the four above-mentioned ESTs, a fifth *L. vannamei* transcript, FE180025, was also identified in our bioinformatics search for putative crustacean prepro-PISCF-ASTs (Table 1). Unlike the precursors deduced from the ESTs discussed earlier, the 138 amino acid PISCF-AST-containing sequence encoded by FE180025 was bounded by stop codons but contained no functional start codon that would suggest a protein destined for secretion (a Met residue is present, but not suggested by SignalP to produce a signal peptide). It is possible that this lack of a functional start codon results from a sequencing error, but at present, this is conjecture. Regardless, a large portion of the sequence shows a high level of amino acid identity with the putative prepro-PISCF-ASTs described above, including a putative isoform of PISCF-AST, predicted here to be pQIRYHQCYFNPVSCF (disulfide bridging predicted between Cys^7 and Cys^{14}), which is distinct from those predicted from the other transcripts.

3.2. Mass spectral assessment of PISCF-ASTs in L. vannamei tissues

Using MALDI-FTMS, we analyzed both direct tissue samples and tissue extracts from *L. vannamei*. The tissues that were examined included single CoGs, analyzed as direct tissue samples, and single brain, thoracic ganglion, and eyestalk ganglia, analyzed as single tissue extracts. Our analysis of these tissues (data not shown) failed to reveal mass spectral peaks corresponding to the *m/z* values predicted for disulfide bond-containing pQIRYHQCYFNPISCF (*m/z* = 1899.83052), pQIRYHQCYFIPVSCF (*m/z* = 1884.85601), or pQIRYHQCYFNPVSCF (m/z = 1885.81487).

3.3. Mass spectral identification of pQIRYHQCYFNPISCF in the pericardial organs (POs) of four brachyuran crabs, one anomuran and one astacidean

Although the PISCF-AST peptides predicted for *L. vannamei* were not detected by MALDI-FTMS in tissues from this species, we recognized that the peptide pQIRYHQCYFNPISCF (*m/z* = 1899.83052) showed mass and chemical similarities to a peptide detected, but not reported, in an earlier study that used chemical and enzymatic reactions to verify the N-terminal sequence of the peptide pEGFYSQRYamide, which was found in the POs from several brachyuran species (Stemmler et al., 2007b). In this previous study, PO extracts from the crab *P. producta* were subjected to chemical and enzymatic treatments that included methyl esterification, to convert acidic residues and unamidated C-terminal carboxylate groups to methyl esters, acetylation, to convert free N-terminal amino groups to acyl derivatives, and reaction with the enzyme pyroglutamate aminopeptidase, to deblock the pyroglutamate-containing N-terminus of peptides. The PO tissue extracts from *P. producta* showed a peak that agreed with that predicted for the peptide pQIRYHQCYFNPISCF (see Table 2). When extracts containing this peptide were subjected to the treatments described above, the subsequent shifts in *m/z* value were consistent with the sequence pQIRYHQCYFNPISCF. Specifically, a mass shift to *m/z* 1913.8473 was observed upon methyl esterification, which is consistent with a peptide containing only a C-terminal carboxylate group and no acidic residues (see Table 2). When the tissue extract was subjected to acetylation, the peptide mass remained unchanged, consistent with a blocked N-terminus and no Lys residue in the sequence (data not shown). Finally, when the extract was subjected to enzymatic degradation, a peak appearing at *m/z* 1788.8011 was observed under conditions corresponding to partial deblocking (see Table 2). After a longer reaction time, a shift to *m/z* 1790.8181 was found (see Table 2). These mass shifts are consistent with initial loss of the N-terminal pyroglutamate residue (loss of 111.032 Da, predicted), with disulfide bond reduction taking place over longer reaction times (2.016 Da, predicted). The disulfide bond reduction results from reaction with dithiothreitol (DTT), which is a component of the pyroglutamate aminopeptidase solution.

When we examined the direct tissue MALDI-FTMS spectra for POs from three *Cancer* crabs (infraorder Brachyura), *C. borealis, C. productus, and C. irroratus*, one anomuran, *L. maja*, and one astacidean, *H. americanus*, we detected a peak corresponding to pQIRYHQCYFNPISCF in each sample (Table 2). A representative spectrum from *C. borealis* is shown in Figure 2A. Comparison of the predicted mass and isotopic distribution for the [M+H]$^+$ ion for pQIRYHQCYFNPISCF (disulfide bridge present between the Cys residues; Fig. 2B) with that of the measured isotopic distribution of the putative pQIRYHQCYFNPISCF (disulfide bridge present between the Cys residues; Fig. 2C) showed excellent agreement, strongly supporting the presence and predicted structure of the peptide in the PO of this species.

To conclusively determine if the detected peptide is, in fact, pQIRYHQCYFNPISCF, we measured an MS/MS spectrum for the [M+H]$^+$ peak originating from the MALDI-FTMS analysis of a small piece of PO tissue from *C. borealis*. When the [M+H]$^+$ ion was subjected

to fragmentation using SORI-CID, the mass spectrum shown in Fig. 3A was measured. Although the presence of the disulfide bond limits the number of sequence ions that are observed, the detected fragment ions and their calibrated masses fully support the proposed sequence. When the SORI-CID mass spectrum of the putative peptide was compared with that of a synthesized peptide standard, the two MS/MS spectra showed excellent agreement (see Fig. 3B). This provides strong support for the identification of pQIRYHQCYFNPISCF.

3.4. Assessment of the tissue distribution of pQIRYHQCYFNPISCF in C. borealis and H. americanus via direct tissue MALDI-FTMS

To determine whether the peptide pQIRYHQCYFNPISCF was present in regions of the nervous system in addition to the PO, we analyzed a range of tissue samples from the crab, *C. borealis*, and the lobster, *H. americanus*. In both species, we detected signals with *m/z* values and an isotopic distribution characteristic of the [M+H]$^+$ ion for pQIRYHQCYFNPISCF in the brain, the CoGs, as well as the POs (reported above). In the lobster, *H. americanus*, we detected signals for pQIRYHQCYFNPISCF in extracts of an entire eyestalk (which included the SG); however, we did not detect signals for pQIRYHQCYFNPISCF in the SG itself in either *H. americanus* or *C. borealis*. In addition, in both species, signals for pQIRYHQCYFNPISCF were detected in tissue samples taken from the PMC, a known gut endocrine tissue (Christie et al., 2007).

3.5. Mass spectral detection of pQIRYHQCYFNPISCF in multiple decapod species

To determine whether pQIRYHQCYFNPISCF (with disulfide bridge) is a broadly conserved decapod neuropeptide, we searched for the *m/z* value and isotopic distribution associated with this peptide in MALDI-FTMS spectra of directly analyzed CoG tissues. The spectra that we accessed for this analysis were acquired as part of a prior study that constituted a broad survey to detect highly conserved neuropeptides (Stemmler et al., 2007a). The spectra were collected using conditions optimized for the detection of *m/z* 1500 and peptide masses were established using internal calibration with a polymer calibrant. A summary of all species where we detected signals consistent with pQIRYHQCYFNPISCF (with disulfide bond) can be found in Table 4 and representative spectra are shown in Fig. 4. In general, the intensities of the pQIRYHQCYFNPISCF peaks observed for CoG tissues were lower (in absolute and relative terms) than the signals observed for PO tissues.

Because the spectra used in our previous study were not optimized for the detection of this higher *m/z* peptide, the ions characteristic of pQIRYHQCYFNPISCF were not always detected in spectra calibrated with polymer calibrant. For that reason, some of the peaks reported in Table 4 were calibrated with other known peptides whose identities had been previously established (Stemmler et al., 2007a). In summary, our results support the presence of pQIRYHQCYFNPISCF (with disulfide bridge) in six infraorders of decapod crustaceans.

4. Discussion

4.1. Functional genomic and mass spectral analyses show that PISCF-type allatostatins are not limited to holometabolous insects

In insects, three structurally distinct families of allatostatins are known: the A-ASTs, exhibiting a –Y*X*FGLamide C-terminus, the B- or cricket-type ASTs, possessing a –W*X*$_6$Wamide C-terminus, and the PISCF-, *Manduca sexta*- or C-type ASTs, typified by a pyroglutamine blocked N-terminus, an unblocked –PISCF C-terminus, and a disulfide bridge between two internal Cys residues (Stay and Tobe, 2007). In crustaceans, multiple isoforms of the A- and B-type peptides have been identified, with members of both families shown to serve neuro/myomodulatory roles (Skiebe and Schneider, 1994; Duve et al., 1997;

Jorge-Rivera and Marder, 1997; Jorge-Rivera et al., 1998; Dircksen et al., 1999; Kreissl et al., 1999; Duve et al., 2002; Birmingham et al., 2003; Huybrechts et al., 2003; Fu et al., 2005; Billimoria et al., 2006; Yasuda-Kamatani and Yasuda, 2006; Yin et al., 2006; Cruz-Bermúdez and Marder, 2007; Fu et al., 2007; Christie et al., 2008a; Christie et al., 2008b; Ma et al., 2008; Gard et al., 2009; Ma et al., 2009). In contrast, PISCF-ASTs have been identified only in insects, and there, only from holometabolous species, leading to the question of whether or not they are unique to members of this insect grouping (Stay and Tobe, 2007). Here, using functional genomics, we have identified transcripts from the penaeid shrimp *L. vannamei* that encode PISCF-type isoforms, the theorized mature structures of these peptides being pQIRYHQCYFNPISCF and pQIRYHQCYFIPVSCF (and potentially pQIRYHQCYFNPVSCF as well; see *3.1*), with disulfide bridging predicted between the two Cys residues present in each sequence. The hypothesized *L. vannamei* peptides possess all of the hallmarks of the PISCF-AST family (with the exception of a conservative Val for Ile substitution at position 12 in pQIRYHQCYFIPVSCF and pQIRYHQCYFNPVSCF), and thus appear to be authentic members of this peptide family.

In an attempt to determine if the predicted *L. vannamei* PISCF-ASTs are actually produced, we conducted high-mass-resolution direct tissue MALDI-FTMS on neural tissues isolated from *L. vannamei*, as well as from a number of other decapod species. In *L. vannamei* we were unable to identify any of the theorized PISCF-type peptides. It is unclear whether our lack of detection results from a true lack of production of the peptides in this shrimp, or whether the peptides were present, but in quantities below the detection limits of our instrumentation. It is also possible that we simply did not assay the correct tissue for detection in *L. vannamei*. However, we were able to detect a mass corresponding to one of the predicted isoforms, pQIRYHQCYFNPISCF, in neural tissues from 28 other decapods. Moreover, in the POs of the crabs *C. borealis* and *P. producta*, where the peaks corresponding to the peptide were of the highest intensity, we were able to obtain confirmation of the peptide's predicted structure using SORI-CID sequencing and chemical derivatization/enzymatic cleavage, respectively.

Given our prediction of multiple PISCF-ASTs in *L. vannamei*, and our mass spectral confirmation of pQIRYHQCYFNPISCF in numerous other decapods, it is clear that the presence of PISCF-type family members is not limited solely to holometabolous insect species. However, the extent to which members of this peptide family are conserved remains to be determined. Interestingly, preliminary screening of ESTs from other phyla suggest that authentic PISCF-ASTs may also be present in at least some members of the Platyhelminthes, where the peptide pQIRYRQCYFNPISCF can be deduced from the *Schistosoma mansoni* transcript CD076673 (A.E. Christie, unpublished). The identification of this platyhelminth transcript supports the hypothesis that the PISCF-ASTs may be a broadly conserved peptide family. As additional experiments are conducted, it will be interesting to see if this hypothesis is borne out.

4.2. What are the origins of the multiple PISCF-type allatostatins predicted in Litopenaeus vannamei?

As stated in the previous section, multiple ESTs putatively encoding PISCF-ASTs were identified from the shrimp *L. vannamei* by functional genomics. Three of the five identified transcripts encoded the peptide pQIRYHQCYFNPISCF, while one of the remaining two ESTs encoded pQIRYHQCYFIPVSCF, and the other pQIRYHQCYFNPVSCF. At present, the origin of these different isoforms is unclear. In all species thus far examined, only one PISCF-AST is encoded within any given precursor (*e.g.* Williamson et al., 2001; Li et al., 2006; Sheng et al., 2006). Furthermore, in each of the insect species in which PISCF-type peptides have been identified, only a single PISCF-AST is known (*e.g.* Williamson et al., 2001; Li et al., 2006; Sheng et al., 2006). Thus, it seems likely to us that a similar situation

should exist in the shrimp, and since pQIRYHQCYFNPISCF is the major isoform identified by transcriptomics in *L. vannamei*, we propose that it is "the" isoform present in this species.

If the "one peptide" hypothesis posed above is true, there are a number of scenarios that could account for the additional isoforms of PISCF-AST predicted from *L. vannamei*. One possibility is that the additional PISCF-type peptides result from sequencing errors in the ESTs deposited in the NCBI database. ESTs are often single pass sequences (*e.g.* Towle and Smith, 2006), and thus errors/uncalled nucleotides in them are not uncommon. Misassigned nucleotides in an EST could result in the generation of a codon for a different the amino acid than that truly present in the transcript, thereby producing an alteration in the sequence of the protein translated from it. Given that the theoretical protein putatively liberating pQIRYHQCYFNPVSCF was bounded by stop codons but contained no functional start codon, it seems likely that a sequencing error is the origin of this predicted peptide (see *3.1*). Alternatively, multiple alleles of the gene(s) encoding PISCF-AST may be present in *L. vannamei* and/or the minor isoforms may be the result of individual specific mutations present in the animals used for the initial mRNA collection, situations that have recently been documented for several other decapod peptides (Cashman et al., 2007; Stemmler et al., 2007c; Hsu et al., 2008a). For the Ile for Asn and Val for Ile substitutions at positions 10 and 12, respectively, single nucleotide polymorphisms could convert the RNA codon for one amino acid to the other (*i.e.* AUU or AUC for Ile to AAU or AAC, respectively, for Asn, and GUU, GUC or GUA for Val to AUU, AUC or AUA, respectively for Ile). Clearly additional experimentation will be needed to clarify the origin of the multiple *L. vannamei* peptides unambiguously, and at present we cannot conclusively determine whether or not a single or multiple PISCF-ASTs are present in this species, if they are produced by it all.

4.3. Is pQIRYHQCYFNPISCF the sole, ubiquitously conserved decapod PISCF-AST?

Using a combination of transcriptomics and mass spectrometry, pQIRYHQCYFNPISCF was predicted/detected in one or more members of seven of the currently extant decapod infraorders. In fact, in only nine of the 37 species investigated was this peptide not detected in a MALDI-FTMS survey of CoG samples. Moreover, no other isoforms of PISCF-AST were identified by either transcriptomics or mass spectrometry, with the exceptions of those just discussed in Section *4.2.* While more species will need to be examined in order to definitively determine if pQIRYHQCYFNPISCF is the sole PISCF-AST isoform present in members of the Decapoda, the picture that is emerging from our data is that it may well be ubiquitously conserved within this taxon, regardless of whether or not it is the only isoform present in a given species, as may be the case in *L. vannamei*. Similar levels of conservation have been noted for several other decapod peptides, *e.g.* the myosuppressin pQDLDHVFLRFamide, the SIFamide GYRKPPFNGSIFamide and the tachykinin-related peptide APSGFLGMRamide (Stemmler et al., 2007a). However, to the best of our knowledge, this level of conservation has not been seen in the insect PISCF-ASTs.

4.3. The tissue distribution of pQIRYHQCYFNPISCF suggests that it serves both local-modulatory and hormonal roles in decapod species

In all organisms with both nervous and circulatory systems, neuropeptides can serve as locally-released paracrines (in some cases autocrines) and/or as circulating hormones (Kastin, 2006). In decapods, most of the currently studied peptides appear to serve both functions (Christie et al., 1995; Skiebe, 2001), although there are cases in which a peptide appears to exert its actions via only one or the other pathway. In the crab *C. productus*, for example, it has been suggested that pigment dispersing hormone (PDH) I functions as a locally-released neuromodulator, while PDH II signaling is solely via a hormonal route (Hsu et al., 2008b). Here, using high-resolution MALDI-FTMS mass profiling, we surveyed a

number of neural tissues of the crab *C. borealis* and the lobster *H. americanus* for pQIRYHQCYFNPISCF, identifying it in the brain, the STNS (specifically the CoG), the eyestalk ganglia (but not in the SG itself), and the PO. As the tissues in which pQIRYHQCYFNPISCF was detected include both regions with well-known synaptic neuropil (*e.g.* the brain and CoGs) and neuroendocrine tissues (*e.g.* the PO), it appears that this peptide serves both as a locally-released neuromodulator and as a circulating neurohormone in these species. Given the extreme conservation of the peptide among the decapods, it seems likely that pQIRYHQCYFNPISCF may serve similar roles in other members of this taxon as well. The functional roles served by locally-released and/or circulating pQIRYHQCYFNPISCF in *C. borealis* and *H. americanus* remain to be determined, though its presence in the ganglia of the STNS of these species, which contain the neural circuits responsible for controlling the musculature the foregut, as well as in the PMC of the midgut, strongly suggest a role in modulation of feeding-related behavior.

4.4. Crustacean PISCF-AST vs. PISCF-AST-like peptide

Recently, we identified a broadly conserved decapod neuropeptide with the structure SYWKQCAFNAVSCFamide (Dickinson et al., 2009). This peptide, originally described from honeybee *A. mellifera* (Hummon et al., 2006), and recently predicted via transcriptomics from the cladoceran crustacean *Daphnia pulex* (Gard et al., 2009), exhibits several structural elements in common with PISCF-ASTs, *i.e.* the C-terminal sequence – QC*X*FN/L*X*V/I SCF (*X*'s indicating variable residues) and disulfide bridging between the two internal Cys residues. In addition, both SYWKQCAFNAVSCFamide and each of the known authentic PISCF-type peptides is the only PISCF-AST/PISCF-AST-like peptide isoform encoded within its respective precursor. Moreover, each of these is the C-terminal most peptide encoded within its prepro-hormone. In contrast to these similarities, the authentic PISCF-ASTs are N-terminally blocked by pyroglutamine and possess free C-termini, whereas the PISCF-AST-like peptide is N-terminally capped by Ser and is C-terminally amidated.

At the time of our earlier study (Dickinson et al., 2009), the insects and crustaceans in which the PISCF-AST-like peptide SYWKQCAFNAVSCFamide had been found were distinct from those that possess authentic PISCF-ASTs. However, in at least the honeybee, a putative receptor for authentic PISCF-AST had been identified via genome mining (Dickinson et al., 2009). Collectively, these data led us to propose that SYWKQCAFNAVSCFamide might be the honeybee/hemimetabolous insect/crustacean equivalent of PISCF-AST, thus implying that SYWKQCAFNAVSCFamide was evolutionarily related to the PISCF-ASTs. With the results presented here, it is now clear that, in at least the decapods, both the authentic PISCF-AST pQIRYHQCYFNPISCF and the PISCF-AST-like peptide SYWKQCAFNAVSCFamide are present in most, if not all, species. While this finding does not disqualify either an evolutionary link between the two peptides or the possibility that SYWKQCAFNAVSCFamide is the honeybee/ hemimetabolous insect PISCF-AST, it does require that these hypotheses be revisited.

Acknowledgments

References

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol. 2004; 340:783–795. [PubMed: 15223320]

Billimoria CP, DiCaprio RA, Birmingham JT, Abbott LF, Marder E. Neuromodulation of spike-timing precision in sensory neurons. J Neurosci. 2006; 26:5910–5919. [PubMed: 16738233]

Birmingham JT, Billimoria CP, DeKlotz TR, Stewart RA, Marder E. Differential and history-dependent modulation of a stretch receptor in the stomatogastric system of the crab, *Cancer borealis*. J Neurophysiol. 2003; 90:3608–3616. [PubMed: 12944539]

Cashman, CR.; Hsu, YA.; Messinger, DI.; Christie, AE.; Dickinson, PS.; de la Iglesia, HO.; Stemmler, EA. Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience; 2007. Identification of individual-specific variations in peptide complement of crustacean neural tissues using direct tissue MALDI-FTMS. Program No 140.4. 2007. Online

Christie AE. Neuropeptide discovery in Ixodoidea: an *in silico* investigation using publicly accessible expressed sequence tags. Gen Comp Endocrinol. 2008a; 157:174–185. [PubMed: 18495123]

Christie AE. *In silico* analyses of peptide paracrines/hormones. in Aphidoidea Gen Comp Endocrinol. 2008b; 159:67–79.

Christie AE, Cashman CR, Brennan HR, Ma M, Sousa GL, Li L, Stemmler EA, Dickinson PS. Identification of putative crustacean neuropeptides using *in silico* analyses of publicly accessible expressed sequence tags. Gen Comp Endocrinol. 2008a; 156:246–264. [PubMed: 18321503]

Christie AE, Kutz-Naber KK, Stemmler EA, Klein A, Messinger DI, Goiney CC, Conterato AJ, Bruns EA, Hsu YW, Li L, Dickinson PS. Midgut epithelial endocrine cells are a rich source of the neuropeptides APSGFLGMRamide (*Cancer borealis* tachykinin-related peptide Ia) and GYRKPPFNGSIFamide (Gly1-SIFamide) in the crabs *Cancer borealis, Cancer magister* and *Cancer productus*. J Exp Biol. 2007; 210:699–714. [PubMed: 17267655]

Christie AE, Skiebe P, Marder E. Matrix of neuromodulators in neurosecretory structures of the crab *Cancer borealis*. J Exp Biol. 1995; 198:2431–2439. [PubMed: 8576680]

Christie AE, Sousa GL, Rus S, Smith CM, Towle DW, Hartline DK, Dickinson PS. Identification of A-type allatostatins possessing -YXFGI/Vamide carboxy-termini from the nervous system of the copepod crustacean *Calanus finmarchicus*. Gen Comp Endocrinol. 2008b; 155:526–533. [PubMed: 17950732]

Cruz-Bermúdez ND, Marder E. Multiple modulators act on the cardiac ganglion of the crab, *Cancer borealis*. J Exp Biol. 2007; 210:2873–2884. [PubMed: 17690236]

Dickinson PS, Wiwatpanit T, Gabranski ER, Ackerman RJ, Stevens JS, Cashman CR, Stemmler EA, Christie AE. Identification of SYWKQCAFNAVSCFamide: a broadly conserved crustacean C-type allatostatin-like peptide with both neuromodulatory and cardioactive properties. J Exp Biol. 2009; 212:1140–1152. [PubMed: 19423507]

Dircksen H, Skiebe P, Abel B, Agricola H, Buchner K, Muren JE, Nässel DR. Structure, distribution, and biological activity of novel members of the allatostatin family in the crayfish *Orconectes limosus*. Peptides. 1999; 20:695–712. [PubMed: 10477125]

Duve H, Johnsen AH, Maestro JL, Scott AG, Jaros PP, Thorpe A. Isolation and identification of multiple neuropeptides of the allatostatin superfamily in the shore crab *Carcinus maenas*. Eur J Biochem. 1997; 250:727–734. [PubMed: 9461295]

Duve H, Johnsen AH, Scott AG, Thorpe A. Allatostatins of the tiger prawn, *Penaeus monodon* (Crustacea: Penaeidea). Peptides. 2002; 23:1039–1051. [PubMed: 12126730]

Fu Q, Kutz KK, Schmidt JJ, Hsu YW, Messinger DI, Cain SD, de la Iglesia HO, Christie AE, Li L. Hormone complement of the *Cancer productus* sinus gland and pericardial organ: an anatomical and mass spectrometric investigation. J Comp Neurol. 2005; 493:607–626. [PubMed: 16304631]

Fu Q, Tang LS, Marder E, Li L. Mass spectrometric characterization and physiological actions of VPNDWAHFRGSWamide, a novel B type allatostatin in the crab, *Cancer borealis*. J Neurochem. 2007; 101:1099–1107. [PubMed: 17394556]

Gard AL, Lenz PH, Shaw JR, Christie AE. Identification of putative peptide paracrines/hormones in the water flea *Daphnia pulex* (Crustacea: Branchiopoda; Cladocera) using transcriptomics and immunohistochemistry. Gen Comp Endocrinol. 2009; 160:271–287. [PubMed: 19135444]

Hsu YW, Stemmler EA, Messinger DI, Dickinson PS, Christie AE, de la Iglesia HO. Cloning and differential expression of two β-pigment-dispersing hormone (β-PDH) isoforms in the crab *Cancer productus*: Evidence for authentic β-PDH as a local neurotransmitter and β-PDH II as a humoral factor. J Comp Neurol. 2008b; 508:197–211. [PubMed: 18311785]

Hsu YW, Weller JR, Christie AE, de la Iglesia HO. Molecular cloning of four cDNAs encoding prepro-crustacean hyperglycemic hormone (CHH) from the eyestalk of the red rock crab *Cancer productus:* identification of two genetically encoded CHH isoforms and two putative post-translationally derived CHH variants. Gen Comp Endocrinol. 2008a; 155:517–525. [PubMed: 17961562]

Hummon AB, Richmond TA, Verleyen P, Baggerman G, Huybrechts J, Ewing MA, Vierstraete E, Rodriguez-Zas SL, Schoofs L, Robinson GE, Sweedler JV. From the genome to the proteome: uncovering peptides in the *Apis* brain. Science. 2006; 314:647–649. [PubMed: 17068263]

Huybrechts J, Nusbaum MP, Bosch LV, Baggerman G, De Loof A, Schoofs L. Neuropeptidomic analysis of the brain and thoracic ganglion from the Jonah crab, *Cancer borealis*. Biochem Biophys Res Commun. 2003; 308:535–344. [PubMed: 12914784]

Jorge-Rivera JC, Marder E. Allatostatin decreases stomatogastric neuromuscular transmission in the crab *Cancer borealis*. J Exp Biol. 1997; 200:2937–2946. [PubMed: 9359878]

Jorge-Rivera JC, Sen K, Birmingham JT, Abbott LF, Marder E. Temporal dynamics of convergent modulation at a crustacean neuromuscular junction. J Neurophysiol. 1998; 80:2559–2570. [PubMed: 9819263]

Kastin, AJ. Handbook of Biologically Active Peptides. First. Academic Press; 2006.

Kreissl S, Weiss T, Djokaj S, Balezina O, Rathmayer W. Allatostatin modulates skeletal muscle performance in crustaceans through pre- and postsynaptic effects. Eur J Neurosci. 1999; 11:2519–2530. [PubMed: 10383641]

Li Y, Hernandez-Martinez S, Fernandez F, Mayoral JG, Topalis P, Priestap H, Perez M, Navare A, Noriega FG. Biochemical, molecular, and functional characterization of PISCF-allatostatin, a regulator of juvenile hormone biosynthesis in the mosquito *Aedes aegypti*. J Biol Chem. 2006; 281:34048–34055. [PubMed: 16968697]

Ma M, Bors EK, Dickinson ES, Kwiatkowski MA, Sousa GL, Henry RP, Smith CM, Towle DW, Christie AE, Li L. Characterization of the Carcinus maenas neuropeptidome by mass spectrometry and functional genomics. Gen Comp Endocrinol. 2009 In press.

Ma M, Chen R, Sousa GL, Bors EK, Kwiatkowski M, Goiney CC, Goy MF, Christie AE, Li L. Mass spectral characterization of peptide transmitters/hormones in the nervous system and neuroendocrine organs of the American lobster *Homarus americanus*. Gen Comp Endocrinol. 2008; 156:395–409. [PubMed: 18304551]

O'Connor PB, Costello CE. Internal calibration on adjacent samples (InCAS) with Fourier transform mass spectrometry. Anal Chem. 2000; 72:5881–5885. [PubMed: 11140752]

Roepstorff P. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom. 1984; 11:601. [PubMed: 6525415]

Sheng Z, Ma L, Cao MX, Li S, Jiang RJ. Biochemical and molecular characterization of allatotropin and allatostatin from the Eri silkworm, *Samia cynthia ricini*. Insect Biochem Mol Biol. 2007; 37:90–96. [PubMed: 17175449]

Skiebe P. Neuropeptides are ubiquitous chemical mediators: using the stomatogastric nervous system as a model system. J Exp Biol. 2001; 204:2035–2048. [PubMed: 11441046]

Skiebe P, Schneider H. Allatostatin peptides in the crab stomatogastric nervous system: inhibition of the pyloric motor pattern and distribution of allatostatin-like immunoreactivity. J Exp Biol. 1994; 194:195–208. [PubMed: 7964402]

Stay B, Tobe SS. The role of allatostatins in juvenile hormone synthesis in insects and crustaceans. Annu Rev Entomol. 2007; 52:277–299. [PubMed: 16968202]

Stemmler EA, Bruns EA, Gardner NP, Dickinson PS, Christie AE. Mass spectrometric identification of pEGFYSQRYamide: a crustacean peptide hormone possessing a vertebrate neuropeptide Y (NPY)-like carboxy-terminus. Gen Comp Endocrinol. 2007b; 152:1–7. [PubMed: 17420018]

Stemmler EA, Cashman CR, Messinger DI, Gardner NP, Dickinson PS, Christie AE. High mass resolution direct-tissue MALDI-FTMS reveals broad conservation of three neuropeptides (APSGFLGMRamide, GYRKPPFNGSIFamide and pQDLDHVFLRFamide) across members of seven decapod crustacean infraorders. Peptides. 2007b; 28:2104–2115. [PubMed: 17928104]

Stemmler EA, Hsu YW, Cashman CR, Messinger DI, de la Iglesia HO, Dickinson PS, Christie AE. Direct tissue MALDI-FTMS profiling of individual *Cancer productus* sinus glands reveals that one

of three distinct combinations of crustacean hyperglycemic hormone precursor-related peptide (CPRP) isoforms are present in individual crabs. Gen Comp Endocrinol. 2007c; 154:184–192. [PubMed: 17698069]

Towle DW, Smith CM. Gene discovery in *Carcinus maenas* and *Homarus americanus* via expressed sequence tags. Int Comp Biol. 2006; 46:912–918.

Veenstra JA. Mono-and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Arch Insect Biochem Physiol. 2000; 43:49–63. [PubMed: 10644969]

Williamson M, Lenz C, Winther AM, Nässel DR, Grimmelikhuijzen CJ. Molecular cloning, genomic organization, and expression of a C-type (*Manduca sexta*-type) allatostatin preprohormone from *Drosophila melanogaster*. Biochem Biophys Res Commun. 2001; 282:124–130. [PubMed: 11263981]

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]

Yin GL, Yang JS, Cao JX, Yang WJ. Molecular cloning and characterization of FGLamide allatostatin gene from the prawn, *Macrobrachium rosenbergii*. Peptides. 2006; 27:1241–1250. [PubMed: 16376458]

```
(FE182974)    MTPTKVHVLAAALLLALALASPAPQDKPLEAAPNQPQHPVHLQKRAAPQDSSPEQDLAAF
(FE175093)    MTPTKVHVLAAALLLALALASPAPQDKPLEAAPNQPQHPVHLQKRAAPQDSSPEQDLAAF
(FE182975)    ------------------------------------------------------------
(FE180026)    ------------------------------------------------------------

(FE182974)    KDLLVAQVAAELRRSWQEPSALRRTLVDEDGLEEEQAAQSKKKRMLAPLSGLPGELPTIK
(FE175093)    KDLLVAQVAAELRRSWQEPSALRRTLVDEDGLEEEQAAQSKKKRMFAPLSGLPGELPTIK
(FE182975)    -------------------------------------------------DELPTIK
(FE180026)    ------------------------------------------------------------

(FE182974)    RQIRYHQCYFNPISCFRRK*
(FE175093)    RQIRYHQCYFNPISCFRRK*
(FE182975)    RQIRYHQCYFNPISCFRRK*
(FE180026)    -QIRYHQCYFIPVSCFRRK*
```

Figure 1.
Deduced amino acid sequences *of Litopenaeus vannamei* PISCF-allatostatin (PISCF-AST) prepro-hormones. Accession numbers of the ESTs from which the prepro-hormones were predicted are shown on the left with the deduced amino acid sequences of the precursor proteins shown on the right. Signal peptides (when present) are shown in grey, with prohormone convertase cleavage loci shown in black. Isoforms of PISCF-AST are shown in red, with other precursor-related peptides shown in blue. Amino acid residues that vary between the precursors are highlighted in yellow. Asterisks indicate the presence of a stop codon in the translated sequence.
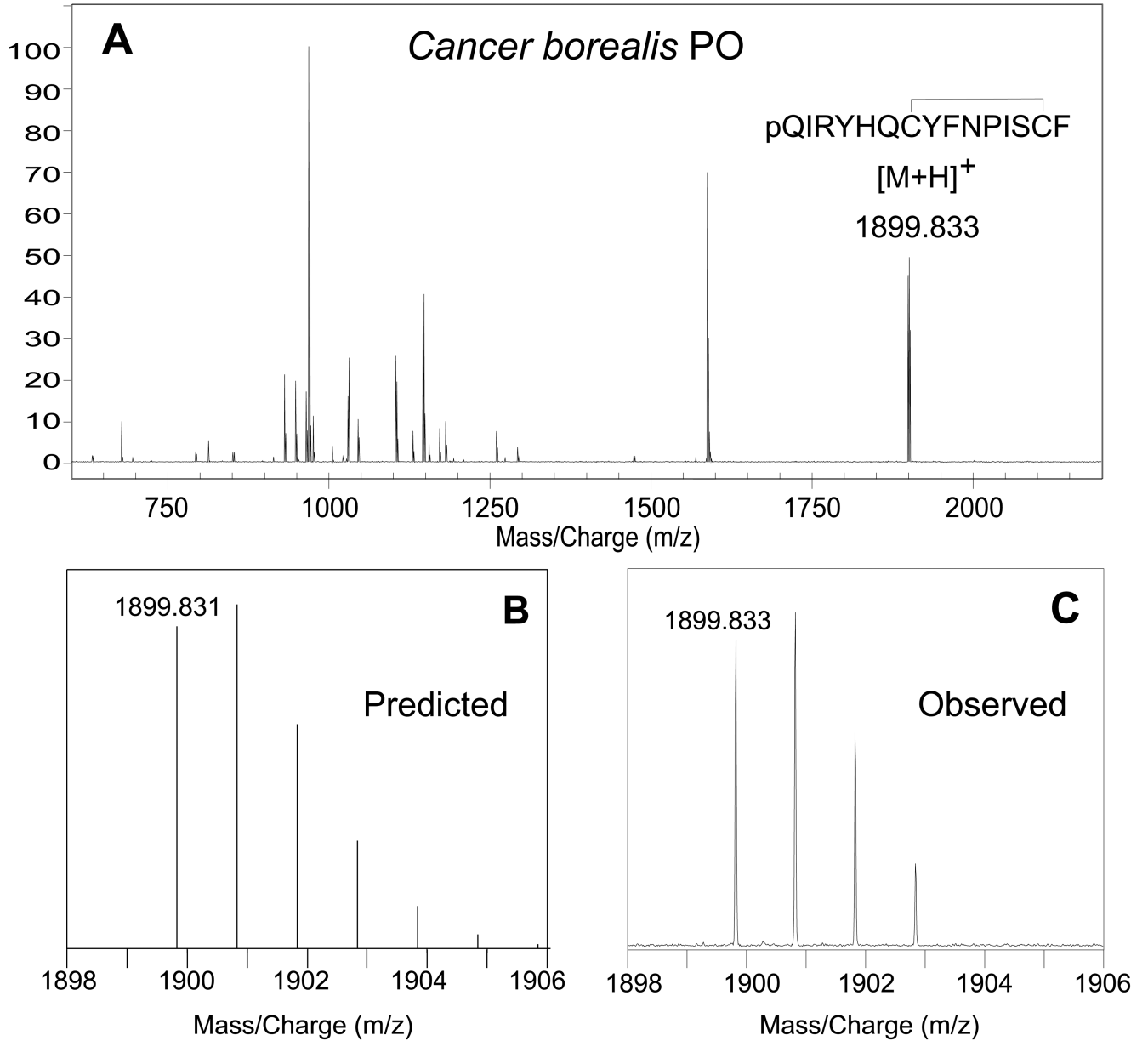
Figure 2.
(A) Direct MALDI-FTMS spectrum of a small piece of freshly dissected *Cancer borealis* pericardial organ (PO). This spectrum was measured using DHB as the matrix, with conditions optimized for the accumulation of *m/z* 2500. As can be seen from this spectrum, a peak corresponding to that of the [M+H]+ ion for pQIRYHQCYFNPISCF (disulfide bridge present between the cysteine residues) was seen at *m/z* 1898.833 (1.5 ppm error from the theoretical *m/z* of 1899.831). (B) Predicted mass and isotopic distribution for the [M +H]+ ion for pQIRYHQCYFNPISCF (disulfide bridge present between the cysteine residues). (C) An expansion of the measured isotopic distribution for putative pQIRYHQCYFNPISCF, showing that the measured mass and isotopic distribution strongly support the existence of this peptide in *C. borealis* PO.
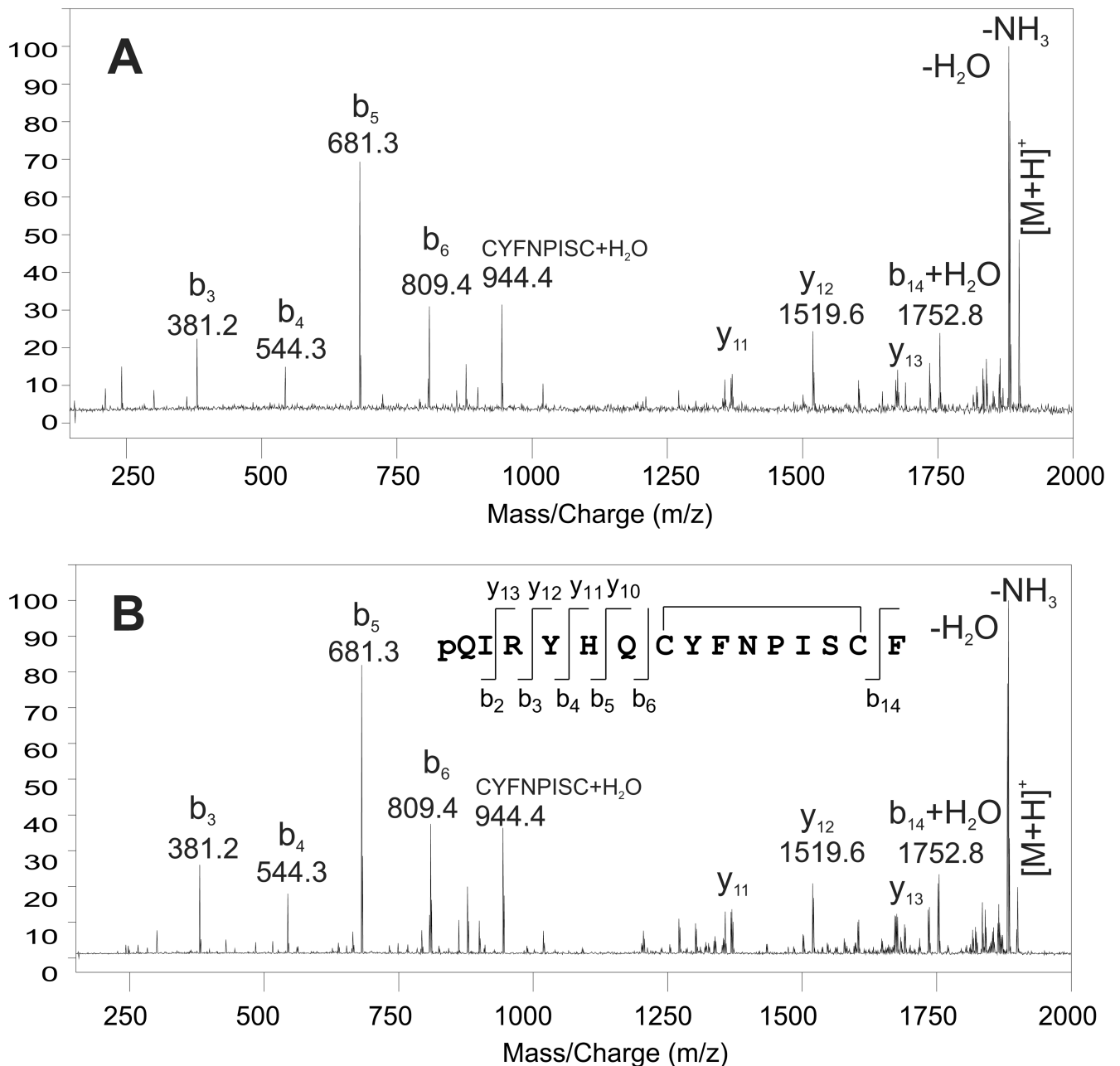
Figure 3.
SORI-CID MALDI-FTMS spectra of putative and synthetic pQIRYHQCYFNPISCF
(disulfide bridge present between the cysteine residues) with a calculated $m/z = 1899.8305$.
(A) SORI-CID of the m/z 1899.83 peak detected in the direct tissue MALDI-FT mass
spectrum of a pericardial organ from *Cancer borealis*. The m/z 1899.83 peak was dissociated
following ion isolation using argon as the collision gas with an excitation amplitude of 6.5
V_{bp}, n = 10; (B) SORI-CID of the m/z 1899.83, $[M+H]^+$ peak, from a synthetic standard of
pQIRYHQCYFNPISCF (disulfide bridge present between the cysteine residues). The m/z
1899.83 peak was dissociated following ion isolation using argon as the collision gas with
an excitation amplitude of 6.5 V_{bp}, n = 5. The y- and b-type fragment ions are identified
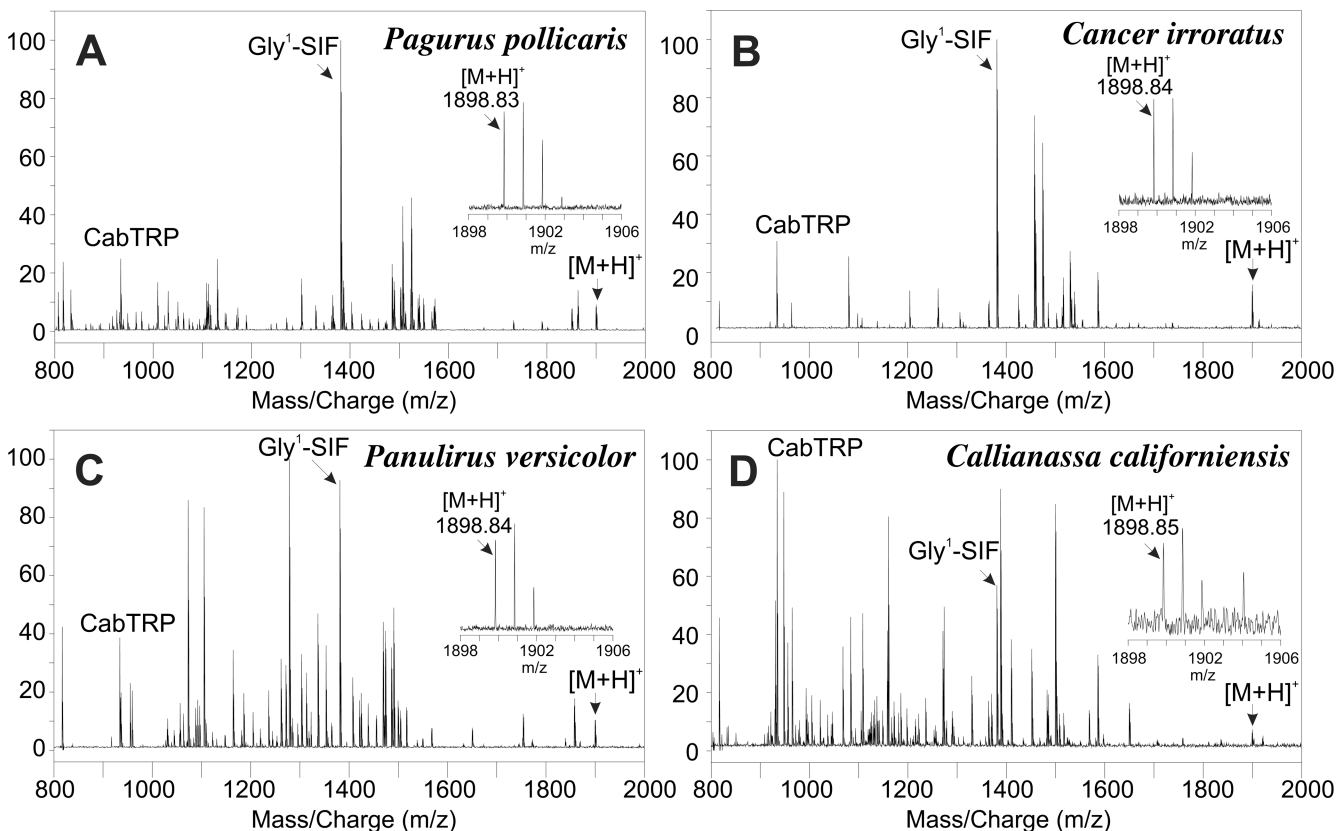using the nomenclature established by Roepstorff (1984).

Figure 4.
Direct tissue MALDI-FTMS spectra of freshly dissected commissural ganglia from (A) *Pagurus pollicaris* (infraorder Anomura), (B) *Cancer irroratus* (infraorder Brachyura), (C) *Panulirus versicolor* (infraorder Achelata) and (D) *Callianassa californiensis* (infraorder Thalassinidea). All spectra were measured using DHB as the matrix and conditions optimized for m/z 1500. The peaks identified as [M+H]$^+$ correspond to the m/z for pQIRYHQCYFNPISCF (disulfide bridge present between the cysteine residues) with a calculated m/z = 1899.8305. The inserts show an expansion of the [M+H]+ peak region to show the measured mass and the isotopic distributions. Spectra were calibrated using known peptide peaks, including APSGFLGMRamide (CabTRP) and GYRKPPFNGSIFamide (Gly^1-SIFamide) at m/z 934.4927 and m/z 1381.7375, respectively (Stemmler et al., 2007a).

Table 1

Bioinformatics of putative decapod PISCF-type allatostatin encoding expressed sequence tags (ESTs)

Species	Library tissue	Accession no.	Blast score	E-value
Litopenaeus vannamei[1]	Nerve cord	FE182974	47.0	2e-05
		FE175093	47.0	3e-05
		FE182975	45.8	6e-05
		FE180025	45.4	7e-05
		FE180026	33.5	0.31

[1]ESTs identified using the sequence of accession no. AAK40100 (Williamson et al., 2001b) as a query.

Table 2

Exact mass measurements for putative pQIRYHQCYFNPISCF (with disulfide bridge) and derivatives detected in the pericardial organs (POs) using MALDI-FTMS

Species	Sequence[a]	Expected mass, [M+H]+	Measured mass[b], [M+H]+	Error (ppm)
P. producta	pQIRYHQ**C**YFNPIS**C**F[c]	1899.8305	1899.8308	0.2
	pQIRYHQ**C**YFNPIS**C**F-OMe[d]	1913.8462	1913.8473	0.6
	IRYHQ**C**YFNPIS**C**F[e]	1788.7983	1788.8011	1.5
	IRYHQCYFNPISCF[f]	1790.8141	1790.8181	2.2
C. borealis[g]	pQIRYHQ**C**YFNPIS**C**F	1899.8305	1899.8302	-0.2
C. productus[g]	pQIRYHQ**C**YFNPIS**C**F	1899.8305	1899.8250	-2.9
C. irroratus[g]	pQIRYHQ**C**YFNPIS**C**F	1899.8305	1899.8367	3.3
L. maja[g]	pQIRYHQ**C**YFNPIS**C**F	1899.8305	1899.8238	-3.6
H. americanus[g]	pQIRYHQ**C**YFNPIS**C**F	1899.8305	1899.8360	2.9

[a]The presence of a disulfide bond is indicated by cysteine residues that have been underlined and appear in bold;

[b]Monoisotopic *m/z, m*ass measured using internal calibration;

[c]Measurements made using PO tissue extracts and direct tissue analysis;

[d]Putative peptide formed following methyl esterification of tissue extract;

[e]Putative peptide formed following partial deblocking with pyroglutamate aminopeptidase of tissue extract;

[f]Putative peptide formed following full deblocking with pyroglutamate aminopeptidase of tissue extract with disulfide bond reduction;

[g]Measurements made using the direct analysis of freshly dissected tissue samples.

Table 3

Exact mass measurements for pQIRYHQ**C**YFNPIS**C**F (with disulfide bridge) from the analysis of freshly dissected *H. americanus* tissue samples or tissue extracts using MALDI-FTMS

Peptide Identity	pQIRYHQCYFNPISCF[a] (with disulfide bridge)	
Ion Identity	[M+H]$^+$	
Calculated *m/z*	*m/z* 1899.83052[b]	
Species	*C. borealis*	*H. americanus*
Tissue[c]	Measured *m/z* (Error, ppm)	Measured *m/z* (Error, ppm)
Supraoesophageal ganglia (brain)	1899.8289 [d] (-0.9)	1899.8333 (1.5)
Eyestalk	---[e]	1899.8250 [d] (-2.9)
Sinus gland (SG)	ND[f]	ND[f]
Commisural ganglion (CoG)	1899.8325 (1.0)	1899.8321 (0.8)
Stomatogastric ganglion (STG)	---[e]	ND[f]
Pericardial organ (PO)	1899.8302 (-0.2)	1899.8360 (2.9)
Posterior midgut caecum (PMC)	1899.8307 (0.1)	1899.8291 (-0.8)

[a] The presence of a disulfide bond is indicated by cysteine residues that have been underlined and appear in bold;

[b] Monoisotopic *m/z*, calibration using either poly(propylene glycol) or other known peptides;

[c] Unless noted, small pieces of tissue were analyzed directly;

[d] Tissue extract was analyzed;

[e] Tissue was not analyzed;

[f] The peptide was not detected.

Table 4

Exact mass measurements for pQIRYHQ**C**YFNPIS**C**F (with disulfide bridge) from the analysis of freshly dissected and desheathed commissural ganglia (CoG) tissue samples using MALDI-FTMS

	Peptide Identity	pQIRYHQ**C**YFNPIS**C**F (with disulfide bridge)	
	Ion Identity	[M+H]$^+$	
	Calculated *m/z*	*m/z* 1899.83052[a]	
Infraorder	Species	Measured *m/z*[b]	Error (ppm)[c]
Penaeidae	*Farfantepenaeus duorarum*	ND	
	Litopenaeus vannamei	ND	
Caridea	*Crangon septemspinosa*	ND	
	Pandalus danae	ND	
	Pandalus platyceros	ND	
Stenopodidea	*Stenopus hispidus*	1899.82820	-1.2
Astacidea	*Cherax quadricarinatus*	1899.83900	4.5
	Homarus americanus	1899.83210	0.8
	Homarus gammarus	1899.83831	4.1
	Nephrops norvegicus	1899.84194	6.0
	Pacifastacus leniusculus	1899.83408	1.9
	Procambarus clarkii	1899.83425	2.0
Thalassinidea	*Callianassa californiensis*	1899.83311	1.4
Achelata	*Panulirus interruptus*	ND	
	Panulirus versicolor	1899.82738	-1.6
	Scyllarides latus	1899.82613	-2.3
Anomura	*Clibanarius vittatus*	1899.83171	0.6
	Lithodes maja	1899.84010	5.0
	Pachycheles rudis	ND	
	Pagurus acadianus	1899.83036	-0.1
	Pagurus granosimanus	ND	
	Pagurus pollicaris	1899.83572	2.7
	Petrolisthes cinctipes	1899.83344	1.5
	Petrolisthes eriomerus	ND	
Brachyura	*Cancer antennarius*	1899.83800	4.0
	Cancer borealis	1899.83252	1.0
	Cancer gracilis	1899.83126	0.4
	Cancer irroratus	1899.83635	3.1
	Cancer magister	1899.83757	3.7
	Cancer productus	1899.83467	2.2
	Carcinus maenas	1899.83064	0.1
	Hemigrapsus nudus	1899.83408	1.9
	Lophopanopeus bellus	1899.83255	1.1

	Peptide Identity	pQIRYHQCYFNPISCF (with disulfide bridge)	
	Ion Identity	[M+H]$^+$	
	Calculated *m/z*	*m/z* **1899.83052**[a]	
Infraorder	Species	Measured *m/z*[b]	Error (ppm)[c]
	Ovalipes ocellatus	1899.84186	6.0
	Pugettia gracilis	1899.83685	3.3
	Pugettia producta	1899.83842	4.2
	Scyra acutifrons	1899.82610	-2.3

[a]Monoisotopic *m/z*;

[b]Average of calibrated mass measurements. Spectra were calibrated using either poly(propylene glycol), NFDEIDRSGFGFA (Anomura and Brachyura) and APSGFLGMRamide (*m/z* 934.4927), pQDLDHVFLRFamide (*m/z* 1271.6531 and *m/z* 817.4832), and GYRKPPFNGSIFamide (*m/z* 1381.77375) or VYRKPPFNGSIFamide (*m/z* 1423.7845). ND = not detected;

[c]Mass measurement error in ppm