Insect Biochemistry and Molecular Biology 116 (2020) 103257

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Assessment of midgut enteroendocrine peptide complement in the honey bee, *Apis mellifera*

Andrew E. Christie

Békésy Laboratory of Neurobiology, Pacific Biosciences Research Center, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, 1993 East-West Road, Honolulu, HI, 96822, USA

ARTICLE INFO

Keywords:
Transcriptomics
Epithelial endocrine cell
Brain-gut peptide

ABSTRACT

Peptides modulate physiological/behavioral control systems in all animals. In arthropods, midgut epithelial endocrine cells are one of the largest sources of these signaling agents. At present, little is known about the identity of the peptides that form arthropod midgut enteroendocrine peptidomes. While many techniques can be used for peptide structural identification, *in silico* transcriptome mining is one that has been used extensively for arthropod neuropeptidome prediction; this strategy has yet to be used for large-scale arthropod enteroendocrine peptide discovery. Here, a tissue-specific transcriptome was used to assess putative enteroendocrine peptide complement in the honey bee, *Apis mellifera,* midgut. Searches for transcripts encoding members of 42 peptide families were conducted, with evidence of expression for 15 groups found in the assembly: adipokinetic hormone, allatostatin A, allatostatin C, bursicon, CCHamide, CNMamide, diuretic hormone 31, diuretic hormone 44, insulin-like peptide, myosuppressin, neuropeptide F, pigment dispersing hormone, pyrokinin, short neuropeptide F, and tachykinin-related peptide. The proteins deduced from the midgut transcripts are identical in sequence, or nearly so, to those of *Apis* pre/preprohormones deposited previously into NCBI, providing increased confidence in the accuracy of the reported data. Seventy-five peptides were predicted from the deduced precursor proteins, 26 being members of known peptide families. Comparisons to previously published mass spectrometric data support the existence of many of the predicted *Apis* peptides. This study is the first prediction of an arthropod midgut peptidome using transcriptomics, and provides a powerful new resource for investigating enteroendocrine peptide signaling within/from the *Apis* midgut, a species of significant ecological/economic importance.

1. Introduction

Many classes of chemicals have been shown to be important players in the modulatory control of arthropod physiology and behavior. By far the largest and most diverse single class of these compounds is peptides, with approximately 40 families now recognized as functioning as locally-released paracrines and/or circulating hormones in members of this taxon (*e.g.*, Altstein and Nässel, 2010; Christie et al., 2010; Clynen et al., 2010; Nässel and Winther, 2010). Work from many laboratories, and on many different arthropod species, has shown peptides to be critical for the control of diverse physiological/behavioral processes, including but not limited to reproduction, growth, feeding, and adaptation to environmental change (*e.g.*, Altstein and Nässel, 2010; Bendena, 2010; Christie et al., 2010; Clynen et al., 2010; Heuer et al., 2012; Nässel and Vanden Broeck, 2016; Nässel and Winther, 2010; Schoofs et al., 2017; Van Wielendaele et al., 2013; Verlinden et al., 2015; Webster et al., 2012). Key to understanding the roles played by peptides in any organism is knowledge of the structures of the native isoforms present in it, as well as their tissue distributions.

A number of arthropod tissues produce peptide paracrines and/or hormones, perhaps the best known and most thoroughly investigated being the nervous system and its associated neuroendocrine organs (*e.g.*, Altstein and Nässel, 2010; Bendena, 2010; Christie et al., 2010; Clynen et al., 2010; Heuer et al., 2012; Nässel and Winther, 2010; Schoofs et al., 2017; Van Wielendaele et al., 2013; Verlinden et al., 2015). Epithelial endocrine cells within the digestive tract, primarily in the midgut, are also major loci of peptide production and release in members of the Arthropoda (*e.g.*, Christie, 2011; Wegener and Veenstra, 2015), though far less work has focused on characterizing enteroendocrine peptides than on their neural counterparts. Many investigations of arthropod enteroendocrine peptides have involved the use of immunohistochemistry to provide information on the peptides present in the midgut of a species (*e.g.*, Christie et al., 2007; Veenstra et al., 1995, 2008, 2009a). While clearly of use for assessment of the

E-mail address: crabman@pbrc.hawaii.edu.

https://doi.org/10.1016/j.ibmb.2019.103257
Received 30 June 2019; Received in revised form 10 October 2019; Accepted 28 October 2019
Available online 31 October 2019
0965-1748/

putative peptide families present in the tissue, this technique provides no information on the identity of the native peptide isoforms.

While many techniques can be used to determine the identity of the peptides present in a given species or tissues, *e.g.*, biochemical isolation/sequencing (*e.g.*, Christie et al., 1997; Rao et al., 1987; Torfs et al., 2002; Wainwright et al., 1996), targeted molecular cloning (*e.g.*, Chung et al., 2009, 2015; Fodor et al., 2017; Klein et al., 1994), and/or mass spectrometry (*e.g.*, Boerjan et al., 2010; Caers et al., 2015; Strum et al., 2016; Ye et al., 2015), *in silico* genome/transcriptome mining is one technique for rapid peptidome prediction that has been exploited extensively for peptide discovery in arthropods (*e.g.*, Christie, 2015; Christie and Chi, 2015; Christie and Pascual, 2016; Christie and Yu, 2019; Christie et al., 2011, 2017, 2018a). Recent advances in technology and decreasing costs for high-throughput nucleotide sequencing have led to an explosion in the nucleotide sequence data that are publicly accessible for a broad array of arthropod species, including some tissue/tissue region-specific datasets (*e.g.*, Bao et al., 2015; Chintapalli et al., 2007; Christie et al., 2017a, 2018b, 2018c; Leader et al., 2018; Robinson et al., 2013; Shelomi et al., 2014; Tran et al., 2019; Waiho et al., 2017). These assemblies/databases have proven powerful resources for predicting area-specific peptidomes for portions of the nervous system (*e.g.*, Bao et al., 2015; Christie et al., 2017a), as well as for reproductive tissues (*e.g.*, Christie, 2016), for several arthropod species. Interestingly, this approach has not yet been applied for large-scale enteroendocrine peptide discovery in any member of the Arthropoda. Recently, a midgut-specific transcriptome for the honey bee, *Apis mellifera*, was generated and deposited for public use (**BioProject No. PRJNA238833**; Shelomi et al., 2014). This resource provides an opportunity to apply the *in silico* approach to enteroendocrine peptide discovery, and here such an investigation was undertaken using this resource.

2. Materials and methods

2.1. In silico transcriptome mining

Searches of the *A. mellifera* midgut transcriptome were conducted using a well-established protocol (*e.g.*, Christie, 2015; Christie and Chi, 2015; Christie and Pascual, 2016; Christie and Yu, 2019; Christie et al., 2017a, 2017b, 2018a). In brief, the database of the online program tblastn (National Center for Biotechnology Information, Bethesda, MD; http://blast.ncbi.nlm.nih.gov/Blast.cgi) was set to "Transcriptome Shotgun Assembly (TSA)" and restricted to data from the *A. mellifera* midgut assembly (**BioProject No. PRJNA238833**; Shelomi et al., 2014). Known arthropod peptide precursors, many from *A. mellifera* itself, were used as the query sequences for the BLAST searches. The complete list of peptide families searched for in this study is provided in Table 1. The specific queries used for the BLAST searches, as well as the accession numbers of the *A. mellifera* transcripts identified as encoding putative peptide precursors, are provided in Supplemental Table 1.

2.2. Peptide structural prediction

A well-established workflow was used to predict the putative mature structures of *A. mellifera* enteroendocrine peptides (*e.g.* Christie, 2015; Christie and Chi, 2015; Christie and Pascual, 2016; Christie and Yu, 2019; Christie et al., 2017a, 2017b, 2018a). Specifically, all hits returned by a given BLAST search were translated using the Translate tool of ExPASy (http://web.expasy.org/translate/) and assessed for completeness. Proteins listed as "full-length" exhibit a start methionine and are flanked on their carboxyl (C)-terminus by a stop codon. Proteins described here as "partial" lacked a start methionine (referred to as C-terminal partial proteins) or a stop codon (referred to as amino [N]-terminal partial proteins). Next, each deduced full-length or N-terminal partial precursor protein was assessed for the presence of a signal peptide using the online program SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/; Bendtsen et al., 2004). Prohormone cleavage sites were identified based on the information presented in Veenstra (2000) and/or by homology to known arthropod pre/preprohormone processing schemes. When present, the sulfation state of tyrosine

Table 1
Comparison of putative midgut enteroendocrine peptide families thus far identified in the midguts of *Apis mellifera* (Apime), *Drosophila melanogaster* (Drome), *Aedes aegypti* (Aedae), *Delia radicum* (Delra), and *Glossina morsitans* (Glomo).

Peptide family	Midgut detection				
	Apime	Drome	Aedae	Delra	Glomo
Adipokinetic hormone	+	+			+
Adipokinetic hormone-corazonin-like peptide					
Allatostatin A	+	+	+	+	+
Allatostatin B		+	+		
Allatostatin C	+	+		+	
Allatotropin					
Bursicon	+	+			
CCHamide	+	+		+	
CNMamide	+				
Corazonin					
Crustacean cardioactive peptide		+			+
Crustacean hyperglycemic hormone/ion transport peptide		+			
Diuretic hormone 31	+	+			
Diuretic hormone 44	+	+			
Ecdysis triggering hormone		+			
Eclosion hormone		+			
Elevenin					
FMRFamide-like peptide					
Glycoprotein hormone		+			
GSEFLamide					
Inotocin					
Insulin-like peptide	+	+			
Leucokinin					
Myosuppressin	+	+			+
Natalisin		+			
Neuroparsin					
Neuropeptide F	+	+			
Neuropeptide-like precursor 1					
Neuropeptide-like precursor 2		+			
Neuropeptide-like precursor 3		+			
Neuropeptide-like precursor 4		+			
Orcokinin		+			
Pigment dispersing hormone	+	+			+
Proctolin					
Prothoracicotropic hormone		+			
Pyrokinin	+		+		+
RYamide					
Short neuropeptide F	+	+	+		+
SIFamide					
Sulfakinin					
Tachykinin-related peptide	+	+	+	+	
Trissin					
Total number of putative midgut peptide families	**15**	**25**	**5**	**4**	**7**

+, present in midgut.
Drosophila melanogaster data from Chen et al. (2016); Reiher et al. (2011); Scopelliti et al. (2014); Siviter et al. (2000); Veenstra (2009a); Veenstra and Ida (2014); Veenstra et al. (2008); Williamson et al. (2001); Yoon and Stay (1995), as well as from FlyAtlas/FlyAtlas 2 (Chintapalli et al., 2007; Leader et al., 2018; Robinson et al., 2013). It should be noted that if a peptide group was reported from any of the abovementioned sources, it is listed as "present" in *Drosophila*; not all sources report the same results for all peptides, *e.g.*, while reported at least one study, no evidence of adult or larval midgut expression was found for adipokinetic hormone, crustacean cardioactive peptide, ecdysis-triggering hormone, eclosion hormone, myosuppressin, or pigment dispersing hormone FlyAtlas.
Aedes aegypti data from Predel et al. (2010).
Delia radicum data from Zoephel et al. (2012).
Glossina morsitans morsitans data from Caers et al. (2015).

A. Alignment of midgut prepro-tachykinin-related peptide variants 1 and 2

```
Prepro-TRP-v1   MIIHSTFLLMVSITLVIAEESDNVLFDKRAPTGHQEMQGKEKNSASLNSENFGIFKRALM
Prepro-TRP-v2   MIIHSTFLLMVSITLVIAEESDNVLFDKRAPTGHQEMQGKEKNSASLNSENFGIFKRALM
                ************************************************************

Prepro-TRP-v1   GFQGVRGKKNSIINDVKNELFPEDINKRAPMGFQGMRGKKASFDDEYYKRAPMGFQGMRG
Prepro-TRP-v2   GFQGVRGKKNSIINDVKNELFPEDINKRAPMGFQGMRGKKASFDDEYYKRAPMGFQGMRG
                ************************************************************

Prepro-TRP-v1   KKSLEEILDEIKKKTTRFQDSRSKDVYLIDYPEDYGKRVLSMDGYQNILDKKDELLGEWE
Prepro-TRP-v2   KKSLEEILDEIKKKTTRFQDSRSKDVYLIDYPEDYGKRVLSMDGYQNILDKKDELLGEWE
                ************************************************************

Prepro-TRP-v1   KRAPMGFYGTRGKKIILDALEELDKRGVMDFQIVSKLFNKNG------------------
Prepro-TRP-v2   KRAPMGFYGTRGKKIILDALEELDKRGVMDFQIGLQRKKDTTFDDYLDYAINPFDYEKRS
                ********************************  :  :..

Prepro-TRP-v1   ------------------------------------------------------------
Prepro-TRP-v2   TDFQDVESGSESFKRARMGFHGMRGKRDAAGIYGSNSSTVGTIFGYQGTYVNLVIVKVFT

Prepro-TRP-v1   --
Prepro-TRP-v2   IK
```

B. Alignment of midgut prepro-tachykinin-related peptide variant 1 and BAC76400

```
Prepro-TRP-v1   MIIHSTFLLMVSITLVIAEESDNVLFDKRAPTGHQEMQGKEKNSASLNSENFGIFKRALM
BAC76400        MIIHSIFLLMVSITLVIAEESDNVLFDKRAPTGHQEMQGKQ-NSASLNSENFGIFKRALM
                ***** ********************************: ******************

Prepro-TRP-v1   GFQGVRGKKNSIINDVKNELFPEDINKRAPMGFQGMRGKKASFDDEYYKRAPMGFQGMRG
BAC76400        GFQGVRGKKNSIINDVKNELFPEDINKRAPMGFQGMRGKKASFDDEYYKRAPMGFQGMRG
                ************************************************************

Prepro-TRP-v1   KKSLEEILDEIKKKTTRFQDSRSKDVYLIDYPEDYGKRVLSMDGYQNILDKKDELLGEWE
BAC76400        KKSLEEILDEIKKKTTRFQDSRSKDVYLIDYPEDYGKRVLSMDGYQNILDKKDELLGEWE
                ************************************************************

Prepro-TRP-v1   KRAPMGFYGTRGKKIILDALEELDKRGVMDFQIVSKLFNKNG
BAC76400        KRAPMGFYGTRGKKIILDALEELDKRGVMDFQIVSI------
                **********************************
```

C. Alignment of midgut prepro-tachykinin-related peptide variant 2 and BAC76399

```
Prepro-TRP-v2   MIIHSTFLLMVSITLVIAEESDNVLFDKRAPTGHQEMQGKEKNSASLNSENFGIFKRALM
BAC76399        MIIHSIFLLMVSITLVIAEESDNVLFDKRAPTGHQEMQGKQ-NSASLNSENFGIFKRALM
                ***** ********************************: ******************

Prepro-TRP-v2   GFQGVRGKKNSIINDVKNELFPEDINKRAPMGFQGMRGKKASFDDEYYKRAPMGFQGMRG
BAC76399        GFQGVRGKKNSIINDVKNELFPEDINKRAPMGFQGMRGKKASFDDEYYKRAPMGFQGMRG
                ************************************************************

Prepro-TRP-v2   KKSLEEILDEIKKKTTRFQDSRSKDVYLIDYPEDYGKRVLSMDGYQNILDKKDELLGEWE
BAC76399        KKSLEEILDEIKKKTTRFQDSRSKDVYLIDYPEDYGKRVLSMDGYQNILDKKDELLGEWE
                ************************************************************

Prepro-TRP-v2   KRAPMGFYGTRGKKIILDALEELDKRGVMDFQIGLQRKKDTTFDDYLDYAINPFDYEKRS
BAC76399        KRAPMGFYGTRGKKIILDALEELDKRGVMDFQIGLQRKKDTTFDDYLDYAINPFDYEKRS
                ************************************************************

Prepro-TRP-v2   TDFQDVESGSESFKRARMGFHGMRGKRDAAGIYGSNSSTVGTIFGYQGTYVNLVIVKVFT
BAC76399        TDFQDVESGSESFKRARMGFHGMRGKRDAAGIYGSNSSTVGTIFGYQGTYVNLVIVKVFT
                ************************************************************

Prepro-TRP-v2   IK
BAC76399        IK
                **
```

Fig. 1. MAFFT alignment of *Apis mellifera* tachykinin-related peptide precursor proteins. (**A**) Alignment of *A. mellifera* prepro-TRP midgut variants 1 (deduced from **GAZV01011906**) and 2 (deduced from **GAZV01011907**). (**B**) Alignment of *A. mellifera* prepro-TRP midgut variant 1 and **BAC76400**, a TRP precursor previously deposited in GenBank. (**B**) Alignment of *A. mellifera* prepro-TRP midgut variant 2 and **BAC76399**, a second TRP precursor previously deposited in GenBank. In this figure, signal peptides are shown in gray, while all mono/dibasic cleavage loci are shown in black. Isoforms of TRP are shown in red, with all linker/precursor related peptides shown in blue. In the line below each sequence grouping, amino acids that are identically conserved between proteins are indicated by "*", while conservative amino acid substitutions are noted by ":" or ".". (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

residues was predicted by homology to known peptide isoforms or by using the online program Sulfinator (http://www.expasy.org/tools/sulfinator/; Monigatti et al., 2002). Disulfide bonding between cysteine residues was predicted by homology to known peptide isoforms or by using the online program DiANNA (http://clavius.bc.edu/~clotelab/DiANNA/; Ferrè and Clote, 2005). Other post-translational modifications, *i.e.*, cyclization of N-terminal glutamine/glutamic acid residues and C-terminal amidation at glycine residues, were predicted by homology to known arthropod peptides. All protein/peptide alignments were done using the online program MAFFT version 7 (http://mafft.cbrc.jp/alignment/software/; Katoh and Standley, 2013). To determine amino acid conservation between selected precursor proteins or peptides, the sequences in question were aligned using MAFFT, and amino acid identity/similarity subsequently determined using the alignment output. Specifically, percent identity was calculated as the number of identical amino acids divided by the total number of residues in the longest sequence (x100), while amino acid similarity was calculated as the number of identical and similar amino acids divided by the total number of residues in the longest sequence (x100).

Table 2
Amino acid conservation between deduced *Apis mellifera* midgut peptide precursors and *A. mellifera* protein sequences previously deposited in NCBI.

Midgut precursor	NCBI protein accession number	Percent identical[a]
Prepro-adipokinetic hormone	AEW68342	100
Prepro-allatostatin A	NP_001161181	100
Prepro-allatostatin C II variant 1	XP_006570428	100
Prepro-allatostatin C II variant 2	XP_016770378	100
Prepro-allatostatin C III	No homolog in dataset	–
Pre-bursicon α subunit	XP_006570932	64
Pre-bursicon β subunit	NP_001035352	100[b]
Prepro-CCHamide-2	XP_001120020	100
Prepro-CNMamide	XP_001121373	100
Prepro-diuretic hormone 31	XP_026296832	100
Prepro-diuretic hormone 44	XP_026299101	100[c]
Prepro-insulin-like peptide	XP_026300673	100
Prepro-myosuppressin	XP_006566614	100
Prepro-neuropeptide F	XP_006559366	100
Prepro-pigment dispersing hormone	XP_006570344	100[d]
Prepro-pyrokinin	NP_001104182	100
Prepro-short neuropeptide F	XP_003250155	100
Prepro-tachykinin-related peptide variant 1	BAC76400	95
Prepro-tachykinin-related peptide variant 2	BAC76399	99

[a] Unless otherwise noted, percent identity calculated as the number of identical residues in the two proteins divided by to number of amino acids in the longest sequence (x100).
[b] Amino- and carboxyl-termini only.
[c] Percent identity calculated for the region of overlap only; the midgut protein is amino-terminally extended by 14 amino acids relative to XP_026299101.
[d] Percent identity calculated for the region of overlap only; the midgut protein is a carboxyl-terminal partial protein.

3. Results and discussion

3.1. Identification of peptide precursor-encoding transcripts and pre/preprohormones in the Apis mellifera midgut

To identify peptide groups likely to function as enteroendocrine signaling agents within and/or from the *A. mellifera* midgut, a tissue-specific assembly was searched for transcripts encoding putative precursor proteins for members of 42 generally recognized arthropod peptide families: adipokinetic hormone (AKH), adipokinetic hormone-corazonin-like peptide (ACP), allatostatin A (AST-A), allatostatin B, allatostatin C (AST-C), allatotropin, bursicon, CCHamide, CNMamide, corazonin, crustacean cardioactive peptide (CCAP), crustacean hyperglycemic hormone/ion transport peptide (CHH/ITP), diuretic hormone 31 (DH31), diuretic hormone 44 (DH44), ecdysis-triggering hormone (ETH), eclosion hormone (EH), elevenin, FMRFamide-like peptide (FLP), glycoprotein hormone (GPH), GSEFLamide, inotocin, insulin-like peptide (ILP), leucokinin, myosuppressin, natalisin, neuroparsin, neuropeptide F (NPF), neuropeptide-like precursor 1 (NPLP1), neuropeptide-like precursor 2 (NPLP2), neuropeptide-like precursor 3 (NPLP3), neuropeptide-like precursor 4 (NPLP4), orcokinin, pigment dispersing hormone (PDH), proctolin, prothoracicotropic hormone (PTTH), pyrokinin, RYamide, short neuropeptide F (sNPF), SIFamide, sulfakinin, tachykinin-related peptide (TRP), and trissin (Table 1 and Supplemental Table 1). Transcripts putatively encoding 15 of these groups were identified in the *A. mellifera* midgut transcriptome, *i.e.*, AKH, AST-A, AST-C (AST-C-II and AST-C-III, but not AST-C-I), bursicon, CCHamide (CCHamide-2, but not CCHamide-1), CNMamide, DH31, DH44, ILP, myosuppressin, NPF, PDH, pyrokinin, sNPF, and TRP (Table 1 and Supplemental Table 1). Translation of the identified transcripts revealed single full-length pre/preprohormones for AKH, AST-A, bursicon α, CCHamide-2, CNMamide, DH31, DH44, ILP, myosuppressin, NPF, pyrokinin, and sNPF (Supplemental Table 1 and Supplemental Fig. 1). Two full-length preprohormones were identified for both AST-C-II (Supplemental Table 1 and Supplemental Fig. 1) and TRP (Fig. 1, Supplemental Table 1 and Supplemental Fig. 1); for both groups, precursor variation appears to be due to alternative splicing of single genes. For AST-C-III and PDH, single C-terminal partial precursors were identified, while for the bursicon β, one N- and one C-terminal partial protein were deduced from midgut transcripts (Supplemental Table 1 and Supplemental Fig. 1).

3.2. Comparison of deduced midgut precursors with Apis mellifera protein sequences deposited previously in NCBI

To provide increased confidence in the accuracy of the sequence data reported here, each precursor protein deduced from the midgut transcriptome was aligned with *A. mellifera* pre/preprohormone sequences previously deposited in NCBI. In all but one case, these alignments revealed the midgut-derived protein to be identical in amino acid sequence, or nearly so, to at least one previously reported/submitted *A. mellifera* protein (Table 2). Specifically, the deduced midgut AKH, AST-A, both AST-C-IIs, CCHamide-2, CNMamide, DH31, ILP, myosuppressin, NPF, pyrokinin, and sNPF precursors are identical in amino acid sequence to at least one previously submitted *A. mellifera* protein. Similarly, the N- and C-terminal portions of the bursicon β subunit precursor, and the C-terminal portion of the PDH preprohormone, are identical in their region of overlap to previously submitted full-length precursor proteins. With the exception of a 14 amino acid N-terminal extension, the DH44 precursor deduced from the midgut transcriptome is also identical to a previously submitted sequence. Both of the midgut-derived TRP precursors are also highly similar to previously submitted sequences, differing from them in 10 or fewer substituted/missing amino acid residues (Fig. 1). In fact, only the putative midgut bursicon α-subunit precursor shows significant sequence variation from previously submitted *A. mellifera* proteins that are annotated as bursicon α prehormones; the midgut protein and the NCBI sequence annotated as a bursicon α precursor differ largely in the length of their N-terminus and presence/absence of one small and one large insertion/deletion (INDEL), with the precursor predicted from the midgut transcriptome having a shorter N-terminus and possessing the small and missing the large INDEL.

3.3. Prediction of a peptidome for the Apis mellifera midgut

Seventy-five distinct mature peptides were predicted for the *A. mellifera* midgut using the pre/preprohormones discussed above (Table 3). This peptidome includes 26 isoforms from generally recognized arthropod peptide families, as well as 49 linker/precursor-related peptides that may or may not represent additional bioactive moieties. The predicted midgut peptides included one isoform of AKH, five isoforms of AST-A, two isoforms of AST-C, one bursicon (composed of an α- and β-subunit heterodimer), one isoform of CCHamide, one isoform of CNMamide, one isoform of DH31, one isoform of DH44, one ILP (composed of an A- and B-chain heterodimer), one isoform of myosuppressin, one isoform of NPF, one isoform of PDH, four isoforms of pyrokinin, one isoform of sNPF, and four isoforms of TRP. Prior mass spectral studies conducted on portions of the *A. mellifera* nervous system (*e.g.* Audsley and Weaver, 2006; Boerjan et al., 2010; Hummon et al., 2006; Takeuchi et al., 2003) and/or the endocrine corpora cardiaca-corpora allata complex (*e.g.*, Audsley and Weaver, 2006; Sturm et al., 2016) provide support for the proposed structures of 19 of 26 (73%) of the midgut peptides from known families, and 14 of 49 (29%) of the linker/precursor-related peptide predicted from the deduced midgut preprohormones (Table 3). For example, mass spectrometry supports for the existence of the midgut AST-A isoforms LPVYN-FGIamide, GRDYSFGLamide, GRQPYSFGLamide, and PNDMLSQRYH-FGLamide and the midgut AST-A linker/precursor-related peptides

Table 3
Predicted *Apis mellifera* midgut enteroendocrine peptidome.

Peptide family	Predicted peptide structure
AKH	**pQLNFSTGWa**
AKH-PRP	SQRIGVVEWGVRTECATQAKPSVEQLLSVYHLIQIEARKMLDCRKLNE
AST-A	**LPVYNFGIa**
	GRDYSFGLa
	pQYSFGLa
	GRQPYSFGLa
	PNDMLSQRYHFGLa
AST-A-PRP	MEETPASSMNLQHYNNMLNPMVFDDTMPE
	AYTYVSEY
	WIDTNDN
	NDNADYPLRLNLDYLPVDNPAFHSQENTDDFLEE
	[a]**AVHY(SO3H)SGGQPLGS**
	MSEDEEESSQ
AST-C	**SYWKQCAFNAVSCFa**
	GQAKGRVYWRCYFNAVTCF
AST-C-PRP	**IPAADKERLLNEVDLVDDDGSIETALINYLFTKQIV**
	LRNQLDIGDLQ
	KVVDNQNKGELDGSRPFARSPHLPIENSQIIPSTDEQQTIKMDLQ
Bursicon α	1
Bursicon β	2
CCHa	GCSAFGHSCFGGHa
CCHa-PRP	FDPNIREKILQDDDTITNREIEDLNSRNEFDVSE
	FGGQTEILSSQSRHQDSSRFNPFALSFIVRQWLTSH
	LHQPDMELNN
CNMa	TMISYMTLCHFKICNMa
CNMa-PRP	3
	pQFH
DH31	**GLDLGLSRGFSGSQAAKHLMGLAAANYAGGPa**
DH31-PRP	APQQGYWSQFDEEDPEALMEIITRLGHTMI
	NPELENN
	SEQA
DH44	**IGSLSIVNSMDVLRQRVLLELARRKALQDQAQIDANRRLLETIa**
DH44-PRP	HPISYNTYDERELSRDHPPLLLLVDHRIPDLENEMFDSGNDPGSTVVRT
	LES
	SLPLYGGNMSKTGDSRLKSEFEYVLEQQGKNHDRNVAPERISERLQDWLHNDDSAFRERQDDQT
	AQTNELRLL
ILP	MIRFRRNSKGIHEECCLKSCTTEELRSYCGA
	DVFQYGQKGQTVTEMHQYCGRTLSSTLQIMCGSVYNSRF
ILP-PRP	SNQEMEMDDYMAFYGYDLYPYKSIKNA
MS	**pQDVDHVFLRFa**
MS-PRP	ALPTQCNPGFLDDLPPRIRKVCVALSRIYELGSEMESYIGDKENHITGFHESIPLLDSGV
NPF	[b]**pEPEPMARPTRPEIFTSPEELRRYIDHVSDYYLLSGKARYa**
NPF-PRP	GNVLYSVPDVNYPWDTMKTVVENSQRSQQLKLE
	pQKDSELLGEHETYGA
	pETSRIDTRPCHVLDSIERYYDDVQ
PDH	**NSELINSLLGLPKNMNNAa**
PDH-PRP	+NNDLQRIAKLLLLPPCLCHS
PK	**TSQDITSGMWFGPRLa**
	pQITQFTPRLa
	[a] [b]**pESGEDY(SO3H)FSY(SO3H)GFPKDQEELYTEEQIYLPLFASRLa**
	VPWTPSPRLa
PK-PRP	pEYDGRDSSSGSNNDRAPSNEFGSCTDGKCI
	ADRKPEINSDIEAFANAFEEPHWAIVTIPETE
	pQLHNIVDKPRQNFNDP
sNPF	[c]**SQRSPSLRLRFa**
sNPF-PRP	TENY(SO3H)IDY(SO3H)SDEIPEKMPIENIQELYRLLMQ
	NALENARLGETPFEHLMI
	SDPHLSILSKPMSAIPSYKFDD
TRP	**ALMGFQGVRa**
	APMGFQGMRa
	APMGFYGTRa
	ARMGFHGMRa
TRP-PRP	pEESDNVLFD
	APTGHQEMQGKEKNSASLNSENFGIF
	NSIINDVKNELFPEDIN
	ASFDDEYY
	SLEEILDEI
	TTRFQDS
	SKDVY(SO3H)LIDYPEDYa
	[d]**VLSMDGYQNILD**
	[d]**DELLGEWE**
	IILDALEELD
	GVMDFQIVSKLFNKNa
	GVMDFQIGLQ

(*continued on next page*)

Table 3 (*continued*)

Peptide family	Predicted peptide structure
	DTTFDDY$_{(SO3H)}$LDY$_{(SO3H)}$AINPFDY$_{(SO3H)}$E STDFQDVESGSESF DAAGIYGSNSSTVGTIFGYQGTYVNLVIVKVFTI

Peptide family abbreviations: AKH, adipokinetic hormone; AST-A, allatostatin A; AST-C, allatostatin C; CCHa, CCHamide; CNMa, CNMamide; DH31, diuretic hormone 31; DH44, diuretic hormone 44; ILP, insulin-like peptide; MS, myosuppressin; NPF, neuropeptide F; PDH, pigment dispersing hormone; PK, pyrokinin; sNPF, short neuropeptide F; TRP, tachykinin-related peptide; PRP, precursor-related peptide.

Abbreviations in peptide structures: pQ/pE, amino-terminal pyroglutamic acid; Y$_{(SO3H)}$, sulfated tyrosine; C, cysteine residue participating in a disulfide bond; a, carboxyl-terminal amide group' +, missing amino- and/or carboxyl-terminal amino acids.

Peptide structures too large to fit into the table.

[1] IIGVDECQATPVIHFLQYPGCVPKPIPSYACRGRCSSYLFLQVSGSKIWQMERSCMCCQESGEREASVSLFCPRAKPGEKKFRKVKLLL (disulfide bridging between the first and eighth, second and sixth, third and fifth, and fourth and seventh cysteine residues).

[2] pQVTDDENCETLQSEVHITKDEYDEIGRLKRTCSG*XXXXXXXXXXXXXXXXXXXXXXXXXXXX*KECYCCRESYLKERHITLHHCYDADGIKLMNEENGVMEIKIREPVECKCIKCGDISQ.

[3] pEPIPSEFYGKTSDNDFILLNKLKELFEEKQYVAEREKELDKERMKIQSIIMEGKESEIQSDSDY$_{(SO3H)}$Y$_{(SO3H)}$AEKLPTPNAVVRQEPQHSa.

Peptides shown in bold font have been confirmed via mass spectrometry to exist in *A. mellifera*. Mass spectrometric confirmation data from: Audsley and Weaver, 2006; Boerjan et al., 2010; Hummon et al., 2006; Sturm et al., 2016; Takeuchi et al., 2003.

[a] Identified in an unsulfated form.

[b] Identified in an uncyclized form.

[c] A portion of the predicted peptide identified, *i.e.*, SPSLRLRFamide.

[d] Identified as a single partially processed peptide, *i.e.,* VLSMDGYQNILDKKDELLGEWE.

AYTYVSEY, NDNADYPLRLNLDYLPVDNPAFHSQENTDDFLEE, and AVHY$_{(SO3H)}$SGGQPLGS (unsulfated) in *A. mellifera* (Audsley and Weaver, 2006; Boerjan et al., 2010; Hummon et al., 2006). Similarly, all four of the predicted midgut TRPs, ALMGFQGVRamide, APMGFQGMRamide, APMGFYGTRamide, and ARMGFHGMRamide, have been identified in *A. mellifera* via mass spectrometry (Audsley and Weaver, 2006; Boerjan et al., 2010; Hummon et al., 2006; Takeuchi et al., 2003), as have the TRP linker/precursor-related peptides NSIINDVKNELFPEDIN, ASFDDEYY, SLEEILDEI, VLSMDGYQNILD, DELLGEWE, IILDALEELD, and GVMDFQIGLQ (Hummon et al., 2006). The identifications of the AST-C-III and CCHamide isoforms GQAKGRVYWRCYFNAVTCF and GCSAFGHSCFGGHamide (disulfide bridging between the two cysteine residues in each peptide), respectively, from the midgut appear to be the first reports of members of these two peptide groups in *A. mellifera*; neither family was described in prior genomic/mass spectrometric analyses of *A. mellifera* (*e.g.*, Audsley and Weaver, 2006; Boerjan et al., 2010; Hummon et al., 2006; Sturm et al., 2016; Takeuchi et al., 2003). This said, an *A. mellifera* protein identical to the putative CCHamide precursor (*i.e.*, Accession No. XP_001120020), though not annotated as such, has been deposited in GenBank (see above and Table 2), and sequences encoding GQAKGRVYWRCYFNAVTCF and GCSAFGHSCFGGHamide can be found in the *A. mellifera* whole-genome shotgun contigs dataset (*e.g.*, Accession No. AADG06001906 for the former and Accession No. AADG06006859 for the latter), providing additional support for the existence of both peptides in the honey bee. Given the relatively recent identification of both peptide groups (Roller et al., 2008; Veenstra, 2009b), it seems likely that these two peptides were simply missed, or not searched for, in the earlier genomic/mass spectral analyses (*e.g.*, Audsley and Weaver, 2006; Boerjan et al., 2010; Hummon et al., 2006; Takeuchi et al., 2003), which were largely conducted contemporaneously with or before the discoveries of the AST-C-III and CCHamide families.

3.4. Comparison of *Apis mellifera* midgut peptide families with those of other insects

As stated earlier, while much work has focused on the identification of arthropod neuropeptides, little is known about the identity of enteroendocrine peptides in members of this taxon. This said, mass spectral, molecular and/or immunohistochemical studies conducted on the fruit fly, *Drosophila melanogaster* (*e.g.*, Chen et al., 2016; Reiher et al., 2011; Scopelliti et al., 2014; Siviter et al., 2000; Veenstra, 2009a; Veenstra and Ida, 2014; Veenstra et al., 2008; Williamson et al., 2001; Yoon and Stay, 1995), the mosquito, *Aedes aegypti* (Predel et al., 2010), the cabbage root fly, *Delia radicum* (Zoephel et al., 2012), and the tsetse fly, *Glossina morsitans* (Caers et al., 2015), do allow for at least limited comparisons of conservation/variation in the peptide groups that putatively serve as midgut-derived enteroendocrine signaling agents in these insects with those predicted here for *A. mellifera*. Additionally, searches of FlyAtlas/FlyAtlas 2 (Chintapalli et al., 2007; Leader et al., 2018; Robinson et al., 2013) for peptide paracrines/hormones expressed in the adult and/or larval midgut provide additional insight into the enteroendocrine peptide complement of *D. melanogaster*. As can be seen from Table 1, the only peptide family that was identified in the midgut of all five species was AST-A. However, a number of other groups were found in three or more species, *i.e.*, AST-C, CCHamide, PDH, pyrokinin, sNPF, and TRP, suggesting significant conservation as enteroendocrine peptides for members of these families as well. Interestingly, several peptide groups appear, at least as of now, species-specific in their midgut expression, *i.e.*, *A. mellifera* for CNMamide, *D. melanogaster* for CHH/ITP, ETH, EH, GPH, natalisin, NPLP2, NPLP3, NPLP4, orcokinin, and PTTH, and *G. morsitans* for CCAP, though again, the number of species thus far examined systematically is very small. In none of the five species is there currently evidence for the expression of ACP, allatotropin, corazonin, elevenin, FLP, GSEFLamide, inotocin, leucokinin, neuroparsin, NPLP1, proctolin, RYamide, SIFamide, sulfakinin, or trissin in the midgut (Table 1). As large-scale peptidomic studies are conducted on the midgut of a phylogenetically diverse set of arthropods, it will be interesting to see which peptide groups continue to show broad conservation as enteroendocrine signaling agents, and which are more limited in terms of their midgut expression.

4. Conclusions

In the study presented here, a publicly accessible transcriptome was used to assess putative enteroendocrine peptide complement in the

midgut of *A. mellifera*. Evidence of midgut expression for 15 different peptide families was found in *A. mellifera*, including AKH, AST-A, AST-C, bursicon, CCHamide, CNMamide, DH31, DH44, ILP, myosuppressin, NPF, PDH, pyrokinin, sNPF, and TRP. Seventy-five distinct midgut peptides were predicted, including 26 whose structures place them into a known family, and 49 linker/precursor-related peptides. Prior mass spectral analyses confirm the presence of many of these peptides in *A. mellifera.* The *A. mellifera* midgut peptidome is the first large-scale arthropod enteroendocrine peptidome predicted using transcriptomic data, and is one of the largest collections of midgut-derived peptides currently described for any arthropod species. The transcripts, precursor proteins, and peptides reported here provide a powerful new resource for initiating molecular/physiological investigations of peptidergic signaling within and from the midgut of *A. mellifera,* including what roles midgut-derived peptides may play in controlling enzyme release, gut motility, ion transport etc. within the midgut, as well as providing other organ systems, including the nervous system, with information concerning the status of digestive processing, hence providing a means for coordinating multiple physiological/behavioral control systems.

Funding

The National Science Foundation (IOS-1856307) Cades Foundation (Honolulu, Hawaii) provided funding for this study.

Acknowledgements

Lisa Baldwin is thanked for reading and editing an earlier version of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmb.2019.103257.

References

Altstein, M., Nässel, D.R., 2010. Neuropeptide signaling in insects. Adv. Exp. Med. Biol. 692, 155–165.
Audsley, N., Weaver, R.J., 2006. Analysis of peptides in the brain and corpora cardiaca-corpora allata of the honey bee, *Apis mellifera* using MALDI-TOF mass spectrometry. Peptides 27, 512–520.
Bao, C., Yang, Y., Huang, H., Ye, H., 2015. Neuropeptides in the cerebral ganglia of the mud crab, *Scylla paramamosain*: transcriptomic analysis and expression profiles during vitellogenesis. Sci. Rep. 5, 17055.
Bendena, W.G., 2010. Neuropeptide physiology in insects. Adv. Exp. Med. Biol. 692, 166–191.
Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795.
Boerjan, B., Cardoen, D., Bogaerts, A., Landuyt, B., Schoofs, L., Verleyen, P., 2010. Mass spectrometric profiling of (neuro)-peptides in the worker honeybee, *Apis mellifera.* Neuropharmacology 58, 248–258.
Caers, J., Boonen, K., Van Den Abbeele, J., Van Rompay, L., Schoofs, L., Van Hiel, M.B., 2015. Peptidomics of neuropeptidergic tissues of the tsetse fly *Glossina morsitans morsitans*. J. Am. Soc. Mass Spectrom. 26, 2024–2038.
Chen, J., Kim, S.M., Kwon, J.Y., 2016. A systematic analysis of *Drosophila* regulatory peptide expression in enteroendocrine cells. Mol. Cells 39, 358–366.
Chintapalli, V.R., Wang, J., Dow, J.A., 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat. Genet. 39, 715–720.
Christie, A.E., 2011. Crustacean neuroendocrine systems and their signaling agents. Cell Tissue Res. 345, 41–67.
Christie, A.E., 2015. Neuropeptide discovery in *Symphylella vulgaris* (Myriapoda, Symphyla): *In silico* prediction of the first myriapod peptidome. Gen. Comp. Endocrinol. 223, 73–86.
Christie, A.E., 2016. Prediction of *Scylla olivacea* (Crustacea; Brachyura) peptide hormones using publicly accessible transcriptome shotgun assembly (TSA) sequences. Gen. Comp. Endocrinol. 230–231, 1–16.
Christie, A.E., Chi, M., 2015. Neuropeptide discovery in the Araneae (Arthropoda, Chelicerata, Arachnida): elucidation of true spider peptidomes using that of the Western black widow as a reference. Gen. Comp. Endocrinol. 213, 90–109.
Christie, A.E., Hull, J.J., Richer, J.A., Geib, S.M., Tassone, E.E., 2017b. Prediction of a peptidome for the western tarnished plant bug *Lygus hesperus*. Gen. Comp. Endocrinol. 243, 22–38.
Christie, A.E., Kutz-Naber, K.K., Stemmler, E.A., Klein, A., Messinger, D.I., Goiney, C.C., Conterato, A.J., Bruns, E.A., Hsu, Y.W., Li, L., Dickinson, P.S., 2007. Midgut epithelial endocrine cells are a rich source of the neuropeptides APSGFLGMRamide (*Cancer borealis* tachykinin-related peptide Ia) and GYRKPPFNGSIFamide (Gly^1-SIFamide) in the crabs *Cancer borealis, Cancer magister* and *Cancer productus*. J. Exp. Biol. 210, 699–714.
Christie, A.E., Lundquist, C.T., Nässel, D.R., Nusbaum, M.P., 1997. Two novel tachykinin-related peptides from the nervous system of the crab *Cancer borealis*. J. Exp. Biol. 200, 2279–2294.
Christie, A.E., McCoole, M.D., Harmon, S.M., Baer, K.N., Lenz, P.H., 2011. Genomic analyses of the *Daphnia pulex* peptidome. Gen. Comp. Endocrinol. 171, 131–150.
Christie, A.E., Pascual, M.G., 2016. Peptidergic signaling in the crab *Cancer borealis*: tapping the power of transcriptomics for neuropeptidome expansion. Gen. Comp. Endocrinol. 237, 53–67.
Christie, A.E., Pascual, M.G., Yu, A., 2018a. Peptidergic signaling in the tadpole shrimp *Triops newberryi*: a potential model for investigating the roles played by peptide paracrines/hormones in adaptation to environmental change. Mar. Genomics 39, 45–63.
Christie, A.E., Roncalli, V., Cieslak, M.C., Pascual, M.G., Yu, A., Lameyer, T.J., Stanhope, M.E., Dickinson, P.S., 2017a. Prediction of a neuropeptidome for the eyestalk ganglia of the lobster *Homarus americanus* using a tissue-specific *de novo* assembled transcriptome. Gen. Comp. Endocrinol. 243, 96–119.
Christie, A.E., Stemmler, E.A., Dickinson, P.S., 2010. Crustacean neuropeptides. Cell. Mol. Life Sci. 67, 4135–4169.
Christie, A.E., Yu, A., 2019. Identification of peptide hormones and their cognate receptors in *Jasus edwardsii* – a potential resource for the development of new aquaculture management strategies for rock/spiny lobsters. Aquaculture 503, 636–662.
Christie, A.E., Yu, A., Pascual, M.G., Roncalli, V., Cieslak, M.C., Warner, A.N., Lameyer, T.J., Stanhope, M.E., Dickinson, P.S., Hull, J.J., 2018b. Circadian signaling in *Homarus americanus*: region-specific *de novo* assembled transcriptomes show that both the brain and eyestalk ganglia possess the molecular components of a putative clock system. Mar. Genomics 40, 25–44.
Christie, A.E., Yu, A., Roncalli, V., Pascual, M.G., Cieslak, M.C., Warner, A.N., Lameyer, T.J., Stanhope, M.E., Dickinson, P.S., Hull, J.J., 2018c. Molecular evidence for an intrinsic circadian pacemaker in the cardiac ganglion of the American lobster, *Homarus americanus* - is diel cycling of heartbeat frequency controlled by a peripheral clock system? Mar. Genomics 41, 19–30.
Chung, J.S., Ahn, I.S., Yu, O.H., Kim, D.S., 2015. Crustacean hyperglycemic hormones of two cold water crab species, *Chionoecetes opilio* and *C. japonicus*: isolation of cDNA sequences and localization of CHH neuropeptide in eyestalk ganglia. Gen. Comp. Endocrinol. 214, 177–185.
Chung, J.S., Bembe, S., Tamone, S., Andrews, E., Thomas, H., 2009. Molecular cloning of the crustacean hyperglycemic hormone (CHH) precursor from the X-organ and the identification of the neuropeptide from sinus gland of the Alaskan Tanner crab, *Chionoecetes bairdi*. Gen. Comp. Endocrinol. 162 1291-33.
Clynen, E., Reumer, A., Baggerman, G., Mertens, I., Schoofs, L., 2010. Neuropeptide biology in *Drosophila*. Adv. Exp. Med. Biol. 692, 192–210.
Ferrè, F., Clote, P., 2005. DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res. 33, W230–W232.
Fodor, J., Köblös, G., Kákai, Á., Kárpáti, Z., Molnár, B.P., Dankó, T., Bozsik, G., Bognár, C., Szőcs, G., Fónagy, A., 2017. Molecular cloning, mRNA expression and biological activity of the pheromone biosynthesis activating neuropeptide (PBAN) from the European corn borer, *Ostrinia nubilalis*. Insect Mol. Biol. 26, 616–632.
Heuer, C.M., Kollmann, M., Binzer, M., Schachtner, J., 2012. Neuropeptides in insect mushroom bodies. Arthropod Struct. Dev. 41, 199–226.
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., 2006. From the genome to the proteome: uncovering peptides in the *Apis* brain. Science 314, 647–649.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Klein, J.M., Mohrherr, C.J., Sleutels, F., Riehm, J.P., Rao, K.R., 1994. Molecular cloning of two pigment-dispersing hormone (PDH) precursors in the blue crab *Callinectes sapidus* reveals a novel member of the PDH neuropeptide family. Biochem. Biophys. Res. Commun. 205, 410–416.
Leader, D.P., Krause, S.A., Pandit, A., Davies, S.A., Dow, J.A.T., 2018. FlyAtlas 2: a new version of the *Drosophila melanogaster* expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. Nucleic Acids Res. 46, D809–D815.
Monigatti, F., Gasteiger, E., Bairoch, A., Jung, E., 2002. The Sulfinator: predicting tyrosine sulfation sites in protein sequences. Bioinformatics 18, 769–770.
Nässel, D.R., Vanden Broeck, J., 2016. Insulin/IGF signaling in *Drosophila* and other insects: factors that regulate production, release and post-release action of the insulin-like peptides. Cell. Mol. Life Sci. 73, 271–290.
Nässel, D.R., Winther, A.M., 2010. *Drosophila* neuropeptides in regulation of physiology and behavior. Prog. Neurobiol. 92, 42–104.
Predel, R., Neupert, S., Garczynski, S.F., Crim, J.W., Brown, M.R., Russell, W.K., Kahnt, J., Russell, D.H., Nachman, R.J., 2010. Neuropeptidomics of the mosquito *Aedes aegypti*. J. Proteome Res. 9, 2006–2015.
Rao, K.R., Mohrherr, C.J., Riehm, J.P., Zahnow, C.A., Norton, S., Johnson, L., Tarr, G.E., 1987. Primary structure of an analog of crustacean pigment-dispersing hormone from the lubber grasshopper *Romalea microptera*. J. Biol. Chem. 262, 2672–2675.
Reiher, W., Shirras, C., Kahnt, J., Baumeister, S., Isaac, R.E., Wegener, C., 2011. Peptidomics and peptide hormone processing in the *Drosophila* midgut. J. Proteome Res. 10, 1881–1892.
Robinson, S.W., Herzyk, P., Dow, J.A., Leader, D.P., 2013. FlyAtlas: database of gene expression in the tissues of *Drosophila melanogaster*. Nucleic Acids Res. 41,

D744–D750.
Roller, L., Yamanaka, N., Watanabe, K., Daubnerová, I., Zitnan, D., Kataoka, H., Tanaka, Y., 2008. The unique evolution of neuropeptide genes in the silkworm *Bombyx mori*. Insect Biochem. Mol. Biol. 38, 1147–1157.
Schoofs, L., De Loof, A., Van Hiel, M.B., 2017. Neuropeptides as regulators of behavior in insects. Annu. Rev. Entomol. 62, 35–52.
Scopelliti, A., Cordero, J.B., Diao, F., Strathdee, K., White, B.H., Sansom, O.J., Vidal, M., 2014. Local control of intestinal stem cell homeostasis by enteroendocrine cells in the adult *Drosophila* midgut. Curr. Biol. 24, 1199–1211.
Shelomi, M., Jasper, W.C., Atallah, J., Kimsey, L.S., Johnson, B.R., 2014. Differential expression of endogenous plant cell wall degrading enzyme genes in the stick insect (Phasmatodea) midgut. BMC Genomics 15, 917.
Siviter, R.J., Coast, G.M., Winther, A.M., Nachman, R.J., Taylor, C.A., Shirras, A.D., Coates, D., Isaac, R.E., Nässel, D.R., 2000. Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian pre-protachykinin A. J. Biol. Chem. 275, 23273–23280.
Sturm, S., Ramesh, D., Brockmann, A., Neupert, S., Predel, R., 2016. Agatoxin-like peptides in the neuroendocrine system of the honey bee and other insects. J. Proteomics 132, 77–84.
Takeuchi, H., Yasuda, A., Yasuda-Kamatani, Y., Kubo, T., Nakajima, T., 2003. Identification of a tachykinin-related neuropeptide from the honeybee brain using direct MALDI-TOF MS and its gene expression in worker, queen and drone heads. Insect Mol. Biol. 12, 291–298.
Torfs, P., Baggerman, G., Meeusen, T., Nieto, J., Nachman, R.J., Calderon, J., De Loof, A., Schoofs, L., 2002. Isolation, identification, and synthesis of a disulfated sulfakinin from the central nervous system of an arthropods the white shrimp *Litopenaeus vannamei*. Biochem. Biophys. Res. Commun. 299, 312–320.
Tran, N.M., Mykles, D.L., Elizur, A., Ventura, T., 2019. Characterization of G-protein coupled receptors from the blackback land crab *Gecarcinus lateralis* Y organ transcriptome over the molt cycle. BMC Genomics 20, 74.
Van Wielendaele, P., Badisco, L., Vanden Broeck, J., 2013. Neuropeptidergic regulation of reproduction in insects. Gen. Comp. Endocrinol. 188, 23–34.
Veenstra, J.A., 2000. Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. Arch. Insect Biochem. Physiol. 43, 49–63.
Veenstra, J.A., 2009a. Peptidergic paracrine and endocrine cells in the midgut of the fruit fly maggot. Cell Tissue Res. 336, 309–323.
Veenstra, J.A., 2009b. Allatostatin C and its paralog allatostatin double C: the arthropod somatostatins. Insect Biochem. Mol. Biol. 39, 161–170.
Veenstra, J.A., Agricola, H.J., Sellami, A., 2008. Regulatory peptides in fruit fly midgut. Cell Tissue Res. 334, 499–516.
Veenstra, J.A., Ida, T., 2014. More *Drosophila* enteroendocrine peptides: orcokinin B and the CCHamides 1 and 2. Cell Tissue Res. 357, 607–621.
Veenstra, J.A., Lau, G.W., Agricola, H.J., Petzel, D.H., 1995. Immunohistological localization of regulatory peptides in the midgut of the female mosquito *Aedes aegypti*. Histochem. Cell Biol. 104, 337–347.
Verlinden, H., Gijbels, M., Lismont, E., Lenaerts, C., Vanden Broeck, J., Marchal, E., 2015. The pleiotropic allatoregulatory neuropeptides and their receptors: a mini-review. J. Insect Physiol. 80, 2–14.
Waiho, K., Fazhan, H., Shahreza, M.S., Moh, J.H., Noorbaiduri, S., Wong, L.L., Sinnasamy, S., Ikhwanuddin, M., 2017. Transcriptome analysis and differential gene expression on the testis of orange mud crab, *Scylla olivacea*, during sexual maturation. PLoS One 12, e0171095.
Wainwright, G., Webster, S.G., Wilkinson, M.C., Chung, J.S., Rees, H.H., 1996. Structure and significance of mandibular organ-inhibiting hormone in the crab, *Cancer pagurus*. Involvement in multihormonal regulation of growth and reproduction. J. Biol. Chem. 271, 12749–12754.
Webster, S.G., Keller, R., Dircksen, H., 2012. The CHH-superfamily of multifunctional peptide hormones controlling crustacean metabolism, osmoregulation, moulting, and reproduction. Gen. Comp. Endocrinol. 175, 217–233.
Wegener, C., Veenstra, J.A., 2015. Chemical identity, function and regulation of enteroendocrine peptides in insects. Curr. Opin. Insect Sci. 11, 8–13.
Williamson, M., Lenz, C., Winther, A.M., Nässel, D.R., Grimmelikhuijzen, C.J., 2001. Molecular cloning, genomic organization, and expression of a C-type (*Manduca sexta*-type) allatostatin preprohormone from *Drosophila melanogaster*. Biochem. Biophys. Res. Commun. 282, 124–130.
Ye, H., Wang, J., Zhang, Z., Jia, C., Schmerberg, C., Catherman, A.D., Thomas, P.M., Kelleher, N.L., Li, L., 2015. Defining the neuropeptidome of the spiny lobster *Panulirus interruptus* brain using a multidimensional mass spectrometry-based platform. J. Proteome Res. 14, 4776–4791.
Yoon, J.G., Stay, B., 1995. Immunocytochemical localization of *Diploptera punctata* allatostatin-like peptide in *Drosophila melanogaster*. J. Comp. Neurol. 363, 475–488.
Zoephel, J., Reiher, W., Rexer, K.H., Kahnt, J., Wegener, C., 2012. Peptidomics of the agriculturally damaging larval stage of the cabbage root fly *Delia radicum* (Diptera: Anthomyiidae). PLoS One 7, e41543.