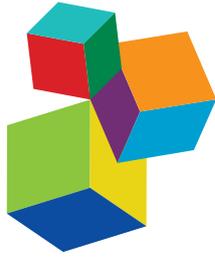


THE EVOLUTION OF NEUROPEPTIDES – A STROLL THROUGH THE ANIMAL KINGDOM: UPDATES FROM THE OTTAWA 2019 ICCPB SYMPOSIUM AND BEYOND

EDITED BY: Klaus H. Hoffmann and Elizabeth Amy Williams

PUBLISHED IN: Frontiers in Endocrinology and Frontiers in Neuroscience





frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88966-017-9

DOI 10.3389/978-2-88966-017-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

THE EVOLUTION OF NEUROPEPTIDES – A STROLL THROUGH THE ANIMAL KINGDOM: UPDATES FROM THE OTTAWA 2019 ICCPB SYMPOSIUM AND BEYOND

Topic Editors:

Klaus H. Hoffmann, University of Bayreuth, Germany

Elizabeth Amy Williams, University of Exeter, United Kingdom

Citation: Hoffmann, K. H., Williams, E. A., eds. (2020). The Evolution of Neuropeptides – a Stroll Through the Animal Kingdom: Updates from the Ottawa 2019 ICCPB Symposium and Beyond. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88966-017-9

Table of Contents

- 04** *Editorial: The Evolution of Neuropeptides - A Stroll Through the Animal Kingdom: Updates From the Ottawa 2019 ICCPB Symposium and Beyond*
Klaus H. Hoffmann and Elizabeth A. Williams
- 06** *Tachykinins: Neuropeptides That are Ancient, Diverse, Widespread and Functionally Pleiotropic*
Dick R. Nässel, Meet Zandawala, Tsuyoshi Kawada and Honoo Satake
- 32** *Global Neuropeptide Annotations From the Genomes and Transcriptomes of Cubozoa, Scyphozoa, Staurozoa (Cnidaria: Medusozoa), and Octocorallia (Cnidaria: Anthozoa)*
Thomas L. Koch and Cornelis J. P. Grimmelikhuijzen
- 46** *Evolution and Comparative Physiology of Luqin-Type Neuropeptide Signaling*
Luis Alfonso Yañez-Guerra and Maurice R. Elphick
- 61** *Expression Analysis of Cnidarian-Specific Neuropeptides in a Sea Anemone Unveils an Apical-Organ-Associated Nerve Net That Disintegrates at Metamorphosis*
Hannah Zang and Nagayasu Nakanishi
- 74** *Expression Profiling, Downstream Signaling, and Inter-subunit Interactions of GPA2/GPB5 in the Adult Mosquito Aedes aegypti*
David A. Rocco and Jean-Paul V. Paluzzi
- 89** *The Adipokinetic Peptides in Diptera: Structure, Function, and Evolutionary Trends*
Gerd Gäde, Petr Šimek and Heather G. Marco
- 105** *Evolution of Neuropeptide Precursors in Polyneoptera (Insecta)*
Marcel Bläser and Reinhard Predel
- 121** *Analysis of Pigment-Dispersing Factor Neuropeptides and Their Receptor in a Velvet Worm*
Christine Martin, Lars Hering, Niklas Metzendorf, Sarah Hormann, Sonja Kasten, Sonja Fuhrmann, Achim Werckenthin, Friedrich W. Herberg, Monika Stengl and Georg Mayer
- 142** *Comparative Aspects of Structure and Function of Cnidarian Neuropeptides*
Toshio Takahashi
- 153** *Function and Distribution of the Wamide Neuropeptide Superfamily in Metazoans*
Elizabeth A. Williams



Editorial: The Evolution of Neuropeptides - a Stroll Through the Animal Kingdom: Updates From the Ottawa 2019 ICCPB Symposium and Beyond

Klaus H. Hoffmann^{1*†} and Elizabeth A. Williams^{2*†}

¹ Faculty for Biology, Chemistry, and Earth Sciences, University of Bayreuth, Bayreuth, Germany, ² College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

Keywords: neuropeptides, neuropeptide evolution, neuropeptide signaling, neuropeptide function, cnidarians, arthropods, echinoderms, metazoan

Editorial on the Research Topic

The Evolution of Neuropeptides - a Stroll Through the Animal Kingdom: Updates From the Ottawa 2019 ICCPB Symposium and Beyond

OPEN ACCESS

Approved by:

Hubert Vaudry,
Université de Rouen, France

*Correspondence:

Klaus H. Hoffmann
Klaus.hoffmann@uni-bayreuth.de
Elizabeth A. Williams
e.williams2@exeter.ac.uk

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 09 June 2020

Accepted: 24 June 2020

Published: 28 July 2020

Citation:

Hoffmann KH and Williams EA (2020)
Editorial: The Evolution of
Neuropeptides - a Stroll Through the
Animal Kingdom: Updates From the
Ottawa 2019 ICCPB Symposium and
Beyond. *Front. Endocrinol.* 11:508.
doi: 10.3389/fendo.2020.00508

Signaling in the nervous and endocrine systems via neuropeptides is an ancient mechanism found in almost all animals. Active neuropeptides, generated by enzymatic cleavage of larger precursor peptides, play a role as neurotransmitters, neuromodulators, hormones, or growth factors and are involved in the regulation of a huge variety of biological systems. Through binding to their specific G protein-coupled receptors, neuropeptides regulate processes including animal development and growth, feeding, metabolism and digestion, behavior, diuresis and homeostasis, and ecdysis and metamorphosis. The chemical structure of neuropeptides is highly diverse and every peptide may be pleiotropic in function, making the study of neuropeptide signaling a complex yet fascinating subject.

Recent advances in genome/transcriptome sequencing, in concert with mass spectrometry and computational prediction and processing, have enabled the identification of active peptides, neuropeptide precursor proteins, and neuropeptide receptors in species from a growing variety of animal taxa. Also recently, techniques such as RNA interference, morpholino knockdown, and the CRISPR-Cas system for genome editing have been used for specific knockdown of neuropeptide precursor genes and neuropeptide receptors, to evaluate the function of neuropeptide signaling. These technological advancements have provided novel insights into neuropeptide function and evolution.

In August 2019, scientists from around the world gathered in Ottawa (Canada) to discuss all aspects of neuropeptide evolution, from structural diversity to functional condition and potential applications, as part of the 10th International Congress of Comparative Physiology and Biochemistry: mechanisms and evolutionary processes. This Research Topic collects the findings presented at the neuropeptide symposium plus recent associated research, including six original research papers and four review articles. The papers cover a wide range of metazoans, from cnidarians to echinoderms.

The origin of neuropeptide signaling in metazoans is currently a major focus in neuropeptide biology. Three papers in this collection extend our knowledge of neuropeptide diversity in the early branching cnidarians. Analysis of largescale transcriptome resources from several different classes of cnidaria by Koch and Grimmelikhuijzen identify two ancient cnidarian neuropeptide classes,

GRFamides and RPRSamides. This study also uncovers the distribution of different neuropeptide classes across largely understudied cnidarians including cubozoans, scyphozoans, staurozoans, and octocorallia. Complementing this study is a review article by Takahashi summarizing the structure and function of Cnidarian neuropeptides, with focus on hydrozoans and hexacorallians. Cnidarian neuropeptides function in a wide range of processes throughout the life cycle, including neuron differentiation, larval locomotion, metamorphosis, muscle contraction, feeding, sensory activity, and reproduction. Like other animal phyla, the cnidarian neuropeptide repertoire includes both novel and conserved neuropeptides. Zang and Nakanishi initiate investigations into cnidarian-specific neuropeptides by mapping expression of RPamide during development in the starlet sea anemone, *Nematostella vectensis*. This study reveals new features of the sea anemone nervous system, including the sensory larval apical organ, and provides insight into how the nervous system is reshaped during metamorphosis from larval to adult form.

Classic studies of neuropeptide structure and function in arthropods such as fruit fly, cockroach, cricket and crab were imperative in establishing the neuropeptide biology field. Here, four original research papers expand our knowledge of panarthropod neuropeptide signaling in diverse directions. Martin et al. characterize the expression of two pigment dispersing hormones (PDFs) and a PDF receptor in a velvet worm (Phylum Onychophora), enlightening the organization and function of PDF signaling, largely known for involvement in the circadian system, in a panarthropod ancestor. The authors also employ receptor deorphanization to uncover the difference in binding specificity of the two PDFs to this G protein-coupled receptor. Bläser and Predel carry out bioinformatic analysis of an impressive 200 insect species from the group Polyneoptera—cockroaches, termites, locusts, and stick insects—to study the evolution of single-copy neuropeptide precursors. These neuropeptides, including ASTC, CCAP, CCHamide, corazonin, elevenin, NPF, proctolin, SIFamide, and sNPF, show a relatively high degree of sequence conservation, although the extent of conservation differs between neuropeptide families. This provides clues regarding the conservation or diversification of function for polyneopteran neuropeptides. Gäde et al. report a mass spectrometry analysis of the presence and structure of adipokinetic hormones (AKHs) in Diptera (flies, mosquitoes) and Mecoptera (scorpion flies). The authors map how changes in single amino acid-increments from an ancestral peptide could have occurred during dipteran evolution to generate existing AKH diversity, providing a finescale look at neuropeptide

evolution. Rocco and Paluzzi investigate the expression of glycoprotein hormone subunits GPA2 and GPB5 and their G protein-coupled receptor LGR1 in the mosquito *Aedes aegypti*. The authors find that GPA2/GPB5 heterodimers activate the LGR1 receptor similar to glycoprotein receptor interactions in human and *Drosophila*. Importantly, the authors noted that each subunit may also interact with other unknown proteins to activate different receptors or signaling pathways for different functions. This study contributes to cross-phylum understanding of neuropeptide signaling as glycoprotein hormone is an ancient neurohormone found across metazoa.

Three review articles contribute to bridging our understanding of neuropeptide signaling across different metazoan phyla. Williams compares and contrasts Wamide neuropeptide family signaling across cnidarians and protostomes. Yañez-Guerra and Elphick survey the luqin family, a paralog of tachykinins, in metazoans. Luqins function in regulating feeding, diuresis, egg-laying, locomotion and lifespan. The identification of luqins in a starfish and hemichordate extends this family into deuterostomes. Nässel et al. review the organization and evolution of multi-functional tachykinin precursor peptides across protostomes and deuterostomes. The authors note that a subset of protostome tachykinins have been recruited for use in venom or salivary glands to affect prey. This interesting evolutionary path of neuropeptide recruitment for novel toxins is also recently revealed in *Nematostella vectensis* through discovery of an ShK-like venom component, which causes muscle contractions in *Nematostella* but is toxic to fish larvae (1).

Sincere thanks to the authors, reviewers, and editors for their valuable contributions to this Research Topic, and to the ICCBP Symposium presenters and participants for their enthusiastic participation. We look forward to moving toward practical applications utilizing knowledge of neuropeptide signaling in the development of pest and parasite control agents and novel drugs, the improvement of invertebrate culture for food, and in bio conservation, both on the land and in the sea.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

EW was supported by a BBSRC David Phillips Fellowship (BB/T00990X/1).

REFERENCES

1. Sachkova MY, Landau M, Surm JM, Macrander J, Singer S, Reitzel AM, et al. Toxin-like neuropeptides in the sea anemone *Nematostella* unravel recruitment from the nervous system to venom. *BioRxiv [Preprint]*. (2020). doi: 10.1101/2020.05.28.121442

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Hoffmann and Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Tachykinins: Neuropeptides That Are Ancient, Diverse, Widespread and Functionally Pleiotropic

Dick R. Nässel^{1*}, Meet Zandawala^{2†}, Tsuyoshi Kawada³ and Honoo Satake^{3†}

¹ Department of Zoology, Stockholm University, Stockholm, Sweden, ² Department of Neuroscience, Brown University, Providence, RI, United States, ³ Bioorganic Research Institute, Suntory Foundation for Life Sciences, Kyoto, Japan

OPEN ACCESS

Edited by:

Klaus H. Hoffmann,
University of Bayreuth, Germany

Reviewed by:

Liliane Schoofs,
KU Leuven, Belgium
Geoffrey Coast,
Birkbeck, University of London,
United Kingdom
Gáspár Jékely,
University of Exeter, United Kingdom

*Correspondence:

Dick R. Nässel
dnassel@zoologi.su.se

†ORCID:

Dick R. Nässel
orcid.org/0000-0002-1147-7766
Meet Zandawala
orcid.org/0000-0001-6498-2208
Honoo Satake
orcid.org/0000-0003-1165-3624

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Neuroscience

Received: 27 September 2019

Accepted: 06 November 2019

Published: 20 November 2019

Citation:

Nässel DR, Zandawala M,
Kawada T and Satake H (2019)
Tachykinins: Neuropeptides That Are
Ancient, Diverse, Widespread
and Functionally Pleiotropic.
Front. Neurosci. 13:1262.
doi: 10.3389/fnins.2019.01262

Tachykinins (TKs) are ancient neuropeptides present throughout the bilaterians and are, with some exceptions, characterized by a conserved FX₁GX₂Ramide carboxy terminus among protostomes and FXGLMamide in deuterostomes. The best-known TK is the vertebrate substance P, which in mammals, together with other TKs, has been implicated in health and disease with important roles in pain, inflammation, cancer, depressive disorder, immune system, gut function, hematopoiesis, sensory processing, and hormone regulation. The invertebrate TKs are also known to have multiple functions in the central nervous system and intestine and these have been investigated in more detail in the fly *Drosophila* and some other arthropods. Here, we review the protostome and deuterostome organization and evolution of TK precursors, peptides and their receptors, as well as their functions, which appear to be partly conserved across Bilateria. We also outline the distribution of TKs in the brains of representative organisms. In *Drosophila*, recent studies have revealed roles of TKs in early olfactory processing, neuromodulation in circuits controlling locomotion and food search, nociception, aggression, metabolic stress, and hormone release. TK signaling also regulates lipid metabolism in the *Drosophila* intestine. In crustaceans, TK is an important neuromodulator in rhythm-generating motor circuits in the stomatogastric nervous system and a presynaptic modulator of photoreceptor cells. Several additional functional roles of invertebrate TKs can be inferred from their distribution in various brain circuits. In addition, there are a few interesting cases where invertebrate TKs are injected into prey animals as vasodilators from salivary glands or paralyzing agents from venom glands. In these cases, the peptides are produced in the glands of the predator with sequences mimicking the prey TKs. Lastly, the TK-signaling system appears to have duplicated in Panarthropoda (comprising arthropods, onychophores, and tardigrades) to give rise to a novel type of peptides, natalisins, with a distinct receptor. The distribution and functions of natalisins are distinct from the TKs. In general, it appears that TKs are widely distributed and act in circuits at short range as neuromodulators or cotransmitters.

Keywords: substance P, neurokinin, neurokinin receptor, natalisin, G protein-coupled receptor, co-transmission, neuropeptide evolution, tachykinin-related peptide

INTRODUCTION

Substance P, the prototypic tachykinin (TK), was the first neuropeptide ever to be isolated from brain tissue already in 1931 (Von Euler and Gaddum, 1931). For a long time it was the sole brain neuropeptide known, and was joined only in the 1950s by the pituitary peptides, oxytocin and vasopressin (Turner et al., 1951; Du Vigneaud et al., 1953; Du Vigneaud, 1955). Today, the number of neuropeptides identified in the animal kingdom is huge and hard to overview [see (Jekely, 2013; Mirabeau and Joly, 2013; Nässel and Zandawala, 2019)]. Also, the number of TK family members has grown immensely over the years, and we know now that there often are several structural and functional representatives in each species.

Already from the outset it was recognized that substance P is produced both in the brain and the intestine (Von Euler and Gaddum, 1931; Hökfelt et al., 2001). Today, it is clear that TKs are utilized by neurons in the CNS, by neurons and enteroendocrine cells associated with the intestine (Otsuka and Yoshioka, 1993; Nässel, 1999; Hökfelt et al., 2001; Satake et al., 2013; Steinhoff et al., 2014), as well as by other cells in mammals such as hematopoietic cells (Zhang et al., 2000; Morteau et al., 2001), endothelial cells, Leydig cells and immune cells [see (Almeida et al., 2004)]. Thus, they are widespread and pleiotropic, and are not only neuropeptides, but also produced by other cell types.

Although the first TK was identified in 1931, it was not until 1971 that substance P was purified and sequenced from 20 kg bovine hypothalamus (Chang et al., 1971), and subsequently synthesized (Tregear et al., 1971). This enabled production of antisera and their application in radioimmunoassay (Powell et al., 1973) and immunocytochemistry (Hökfelt et al., 1975) to localize substance P and demonstrate its release. Thereafter, important experimental work ensued, including development of TK agonists and antagonists [see (Otsuka and Yoshioka, 1993; Hökfelt et al., 2001)], identification of TK receptors [see (Masu et al., 1987; Nakanishi, 1991)], developing genetic approaches and discovering important roles in health and disease [see (Otsuka and Yoshioka, 1993; Hökfelt et al., 2001; Onaga, 2014; Steinhoff et al., 2014)]. The discovery of the roles of substance P and other TKs in pain, inflammation, cancer, depressive disorder, immune function, gut function, hematopoiesis, sensory processing and hormone regulation [see (Hökfelt et al., 2001; Onaga, 2014; Steinhoff et al., 2014; Ziegglänsberger, 2019)] has led to extensive research into the pharmacology and molecular biology of this signaling system as a therapeutic target [see (Steinhoff et al., 2014)], resulting in a huge number of publications annually. However, it is hard to find recent comprehensive reviews on TKs that cover distribution and functions.

Tachykinins have also been explored outside mammals and other vertebrates. The first TK to be identified in an invertebrate was eleudoisin, isolated from salivary glands of the cephalopod *Eledone moschata* (Erspamer and Anastasi, 1962). Eleudoisin was actually the first TK to be sequenced, but since the sequence of substance P was not yet known, the structural relationship was realized only later. The authors, however, recognized that the action of eleudoisin on mammalian smooth muscle is similar to that of substance P (Erspamer and Anastasi, 1962;

Erspamer and Erspamer, 1962). Many years later, four TKs were isolated from the brain and retrocerebral glands of the locust *Locusta migratoria* (Schoofs et al., 1990a,b). Today, multiple TKs (more than 350 sequences) have been identified from over 50 insect species [see the DIneR database¹ (Yeoh et al., 2017)], and numerous ones from other invertebrates and protochordates [see e.g., Kawada et al. (2010), Veenstra (2010, 2011, 2016), Conzelmann et al. (2013), Palamiuc et al. (2017), Zandawala et al. (2017), Dubos et al. (2018), Koziol (2018) and **Figure 1** and **Supplementary Table S1**]. Also in invertebrates, the common TKs are produced by neurons of the CNS and by endocrine cells of the intestine, but the presence of invertebrate TKs in other cell types has not been reported thus far. Functional analysis has revealed that invertebrate TKs are also pleiotropic. Moreover, recent genetic work in *Drosophila*, suggests that many TK functions are conserved over evolution.

In this review, we first comment on TK terminology in invertebrates since at present it may seem somewhat complex and confusing. Furthermore, we discuss the evolution of genes encoding TK precursors and receptors as well as outline TK signaling systems in various phyla across the animal kingdom. Next, we discuss the distribution of TKs and functions of TK signaling systems; here, we are more comprehensive in dealing with invertebrates since vertebrate TK literature is very extensive. Furthermore, we highlight the functions of TK-signaling that are conserved across different animal phyla. Of note, TKs generally appear to signal over a relatively short range within defined neuronal circuits as neuromodulators or cotransmitters. Only a few examples of intestinal TKs acting as local circulating hormones are available. We also discuss a sister group of the TKs, the natalisins, that seems to have arisen by a gene duplication in the Panarthropoda (comprising arthropods, onychophores, and tardigrades) lineage and appears restricted to this group. The natalisins and their receptors constitute a distinct signaling system that has not been investigated in detail thus far.

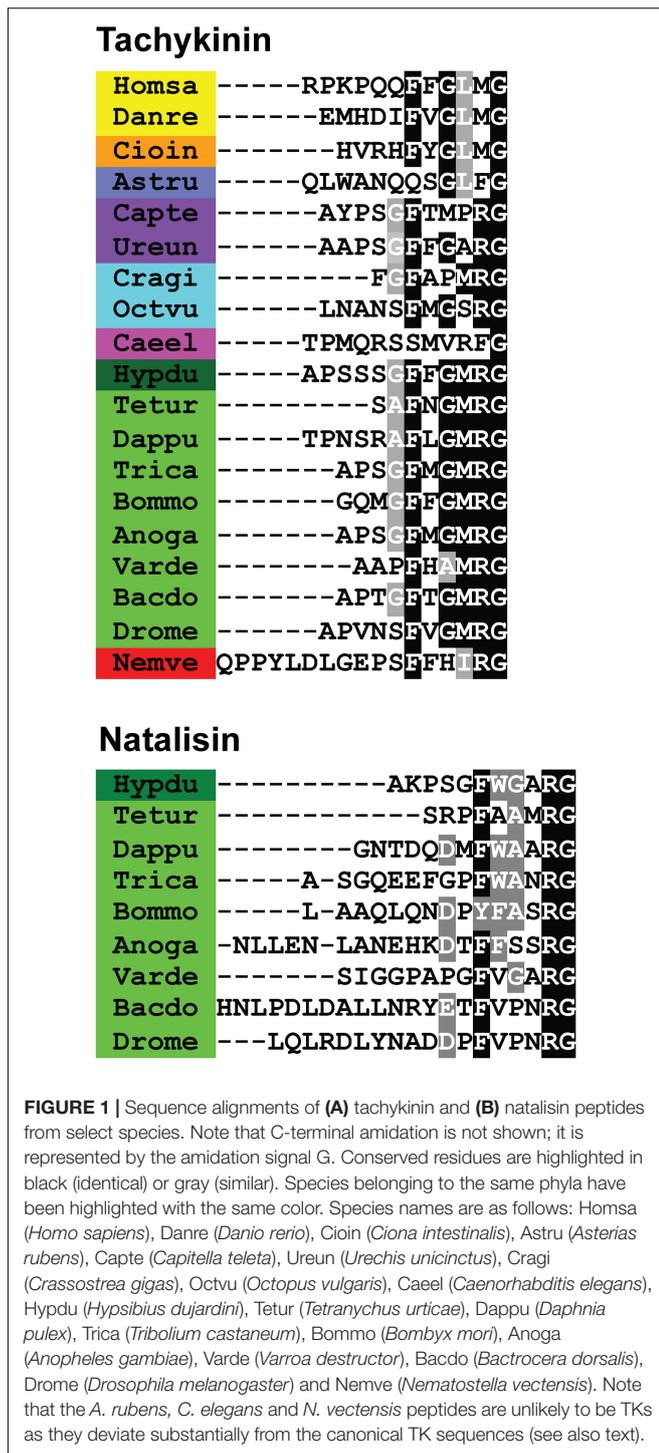
STRUCTURE OF TACHYKININ PEPTIDES AND ORGANIZATION OF GENES ENCODING THEIR PRECURSORS

We start this section with a commentary on TK terminology and continue with describing TK precursors, peptides and receptors in mammals where knowledge is the largest and then move on with other vertebrates and last invertebrates.

A Note on Major Types of Tachykinins and Terminology

Substance P (RPKPQQFFGLMamide), and other mammalian TKs, are characterized by an FXGLMamide carboxy terminus, and these peptides act on either of three TK receptors (GPCRs; NK1R – NK3R) (Otsuka and Yoshioka, 1993; Onaga, 2014; Steinhoff et al., 2014). The first invertebrate neuropeptides referred to as TKs (locustatachykinin I-IV; LomTK I-IV),

¹<http://www.neurostresspep.eu/diner/insectneuropeptides>



were isolated from the brain and retrocerebral complex of locusts, and have a different carboxy terminus, FX₁GX₂Ramide (Schoofs et al., 1990a,b). This sequence is shared by many other invertebrate TKs and only one type of insect TK receptor is known so far (Nässel, 1999; Van Loy et al., 2010; Satake et al., 2013). The structural difference in the active core of the two groups of TK peptides renders the FX₁GX₂Ramides inactive

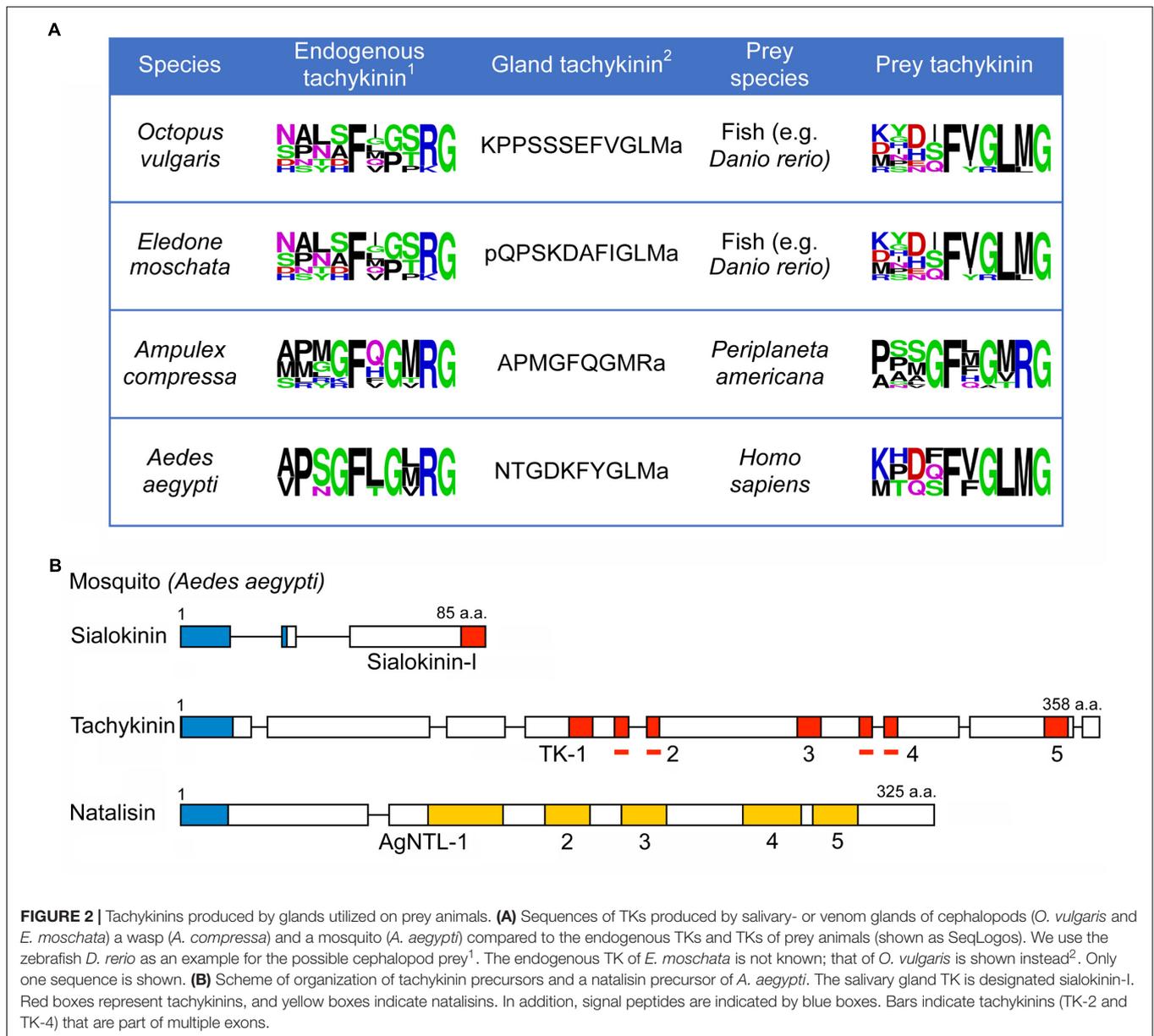
on the vertebrate-type TK receptors and conversely, vertebrate TKs do not activate invertebrate receptors (Satake et al., 2013). Thus, these authors suggested that the FX₁GX₂Ramides should be designated tachykinin-related peptides (TKRPs) to distinguish them from vertebrate TKs with FXGLMamide. In the literature, the individual TKs and TKRPs have been given many different names. In invertebrates, these commonly include a prefix indicating the species of origin (e.g., LomTK in *Locusta migratoria*) and then numbers if multiple peptide paracopies (isoforms) exist on the same precursor (LomTK-I, LomTK-II etc).

To further complicate the terminology of TKs, there are peptides with an FXGLMamide carboxy terminus produced by salivary glands of mosquitos, sialokinins (Champagne and Ribeiro, 1994) and in cephalopods, eledoisin and octopus-tachykinin (Erspamer and Anastasi, 1962; Kanda et al., 2003; Satake et al., 2003) (Figure 2 and Supplementary Table S2). These TKs are delivered to prey and meant to act on exogenous receptors, not within the “sender animal” (predator). The peptides of this kind were referred to as invertebrate TKs (Inv-TKs) (Satake et al., 2003), to distinguish them from TKRPs. Similarly, exocrine glands in amphibian skin produce FXGLMamide-type TKs that have been given different exotic names [see (Lazarus and Attila, 1993)]. A recent finding adds to the TK complexity; in the parasitoid Jewel wasp (*Nasonia vitripennis*), the toxin glands produce a precursor encoding multiple FQGMRamide containing peptides (Arvidson et al., 2019). The wasp injects the toxin that contains FQGMRamide peptide and other components into the cockroach brain to paralyze the host by acting on the cockroach TK receptor in circuits of the central complex. In summary, TKs for exogenous use are produced to act on receptors of target animals and the salivary gland ones deviate structurally from native TKs. We will discuss these in more detail later. In the present review, we use the names originally given to the different TKs when relevant (see Supplementary Tables 1, 2), but for the sake of simplicity we will henceforth use the term TKs for all FXGLMamide and FX₁GX₂Ramides when we discuss the peptides in general.

A related, but distinct, invertebrate peptide signaling system is constituted by the natalisins (NTL) and their receptors (Jiang et al., 2013). These will be discussed separately. Also, note that especially in early papers (but also some more recent ones) a family of neuropeptides designated leucokinins (LKs) has been considered related to TKs [see e.g., Holman et al. (1986), Nässel and Lundquist (1991), Al-Anzi et al. (2010)]. The LKs have an FXSWGamide carboxy terminus, and analysis based on precursor structure (and receptors) show that they are not homologous to TKs (Jekely, 2013; Mirabeau and Joly, 2013).

TKs and Their Receptors in Mammals

In mammals, including humans, there are three genes encoding precursors of tachykinins: preprotachykinin A (PPTA), preprotachykinin B (PPTB) and preprotachykinin C (PPTC), also known as Tac1, Tac3, and Tac4, respectively [see (Onaga, 2014; Steinhoff et al., 2014)]. These genes arose through two rounds of genome duplications in the vertebrate lineage followed by subsequent gene losses and diversification (Elphick et al., 2018). The Tac1 precursor gives rise to Substance



P (SP), Neurokinin A (NKA), Neuropeptide K (NPK), and neuropeptide γ (NP γ), Tac3 to Neurokinin B (NKB), and Tac4 to Hemokinin (HK), Endokinin-A (EKA) and Endokinin-B (EKB). Thus, there are nine different TKs in mice, rats and humans. The sequences of the Tac1 and Tac3 derived TKs are conserved in humans, mouse and rat, whereas the ones encoded on Tac4 differ between species (Steinhoff et al., 2014). The Tac4 encodes another two endokinins (EKC and EKD) that are not TKs (Steinhoff et al., 2014). The organization of mammalian TK precursors is shown in **Figure 3** and their sequences in **Table 1**. For comparison, TK precursors in other representative animals are shown in **Figures 4, 5**, and a cladogram with TK signaling components found in **Figure 6**. It is also worth noting that the Tac1 and Tac4 genes each give rise to 4 splice variants, α , β , γ , and δ (Onaga, 2014;

Steinhoff et al., 2014). The TKs have differential affinity for three different TK receptors, NK1R - NK3R (or TAC1R - TAC3R) (Nakanishi, 1991; Onaga, 2014; Steinhoff et al., 2014) as shown in **Table 1**; the ligand selectivity is as follows: SP > NKA > NKB for NK1R, NKA > NKB > SP for NK2R, and NKB > NKA > SP for NK3R (Satake et al., 2013; Steinhoff et al., 2014). HK and EKs exhibit the highest affinity to NK1R (Satake et al., 2013; Steinhoff et al., 2014). These are G-protein-coupled receptors (GPCRs) of the rhodopsin family (also known as family A GPCRs). The NK2R (neuropeptide K receptor) is of historical interest since it was the first neuropeptide receptor to be cloned (Masu et al., 1987). Signaling through the NK receptors is diverse and complex. For example, the ligand-activated NK1R initiates G-protein mediated signaling that can lead to (1) activation of phospholipase C (PLC), which results in formation

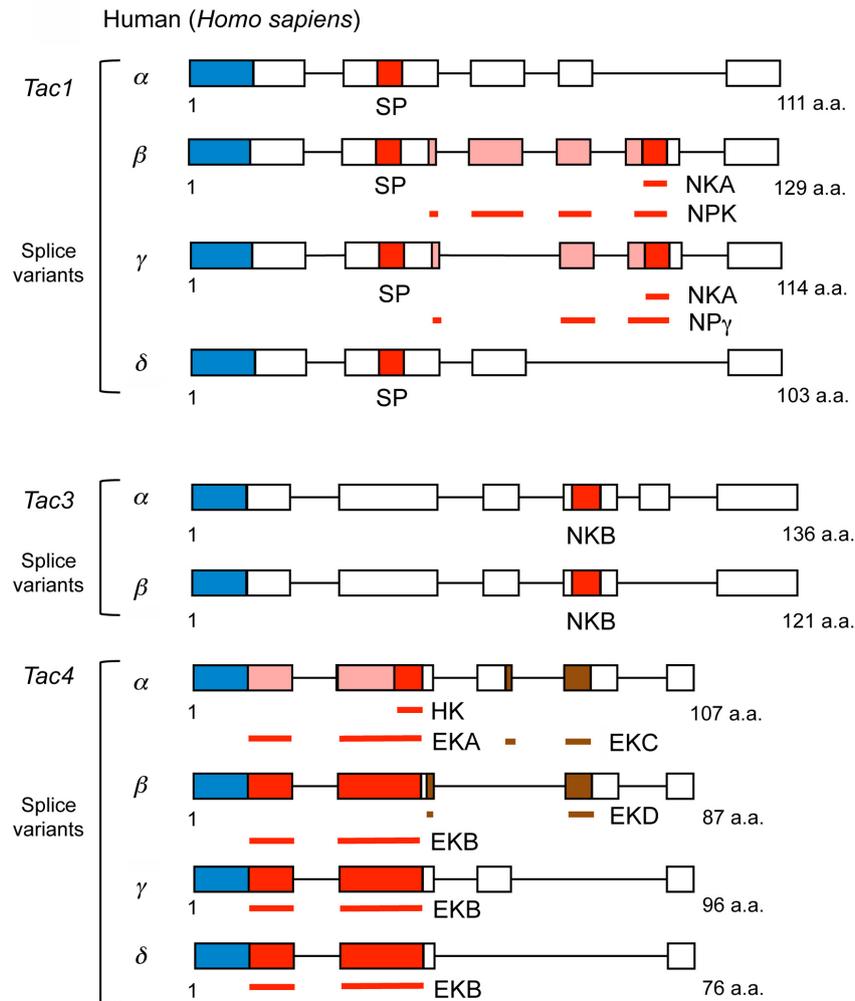


FIGURE 3 | Scheme of human tachykinin precursors (*Tac1*, *Tac3*, and *Tac4* and their splice variants). Boxes and lines show exons and introns, respectively. NPK and NP γ contain NKA sequences indicated by the red box at C-terminus, the N-termini of NPK and NP γ are represented by pink boxes, and brown boxes represent EKC or EKD that are not tachykinins. In addition, signal peptides are indicated by blue boxes. Bars indicate tachykinin or natalisin peptides that are part of multiple exons: NKA, NPK, NP γ , EKA, EKB, EKC, EKD, respectively. The *Tac4* peptides are designated endokinins A - D (EKA - EKD); hemokinin-1 (HK) represents the C-terminal portion of EKA. The N-terminus of EKA is shown as pink boxes. Primary sequence data from Steinhoff et al. (2014).

of inositol trisphosphate (IP₃) and diacylglycerol (DAG), mobilization of intracellular stores of Ca²⁺, and activation of PKC; (2) activation of adenylyl cyclase (AC), resulting in formation of cAMP, and activation of PKA; or (3) activation of phospholipase A₂ and production of arachidonic acid (Steinhoff et al., 2014).

In mammals, TKs play roles as neuromodulators/cotransmitters in central brain circuits, as well as in pain, stress, anxiety, depressive disorder, aggression, memory formation, inflammation, cancer, immune function, gut function, hematopoiesis, sensory processing, reproduction and cytokine and hormone regulation [see (Otsuka and Yoshioka, 1993; Felipe et al., 1998; Hökfelt et al., 2001; Holsboer, 2009; Onaga, 2014; Steinhoff et al., 2014; Lénárd et al., 2018; Zieglgänsberger, 2019)].

Substance P was mapped to neurons in the rat nervous system early on (Hökfelt et al., 1975, 1977; Ljungdahl et al., 1978).

Now, we know that the distribution of SP and other TKs, as well as their receptors, is widespread and plastic. Receptor expression is regulated by various transcription factors under different physiological states. For instance, the receptors can be upregulated during inflammation via the transcription factor NF- κ B (Onaga, 2014; Steinhoff et al., 2014). SP and NKA and their receptors are not only widely distributed throughout the central and peripheral nervous system, but also in many other tissues including dermal tissue, gastrointestinal tract, as well as the respiratory, urogenital and immune systems (Hökfelt et al., 2001; Onaga, 2014; Steinhoff et al., 2014). Whereas SP and NKA are expressed throughout the brain in mammals, NKB is found mainly in the hypothalamus and spinal cord. Furthermore, SP is present in brain circuits that are involved in the processing of anxiety, such as the amygdala, septum, mid-brain, periaqueductal gray, hippocampus, and hypothalamus (Holsboer, 2009).

TABLE 1 | Human tachykinins.¹

Tachykinin	Gene ²	Receptor	Amino acid sequence
Substance P	Tac1 ³	NK1R	RPKPQQFFGLMa
Neurokinin A	Tac1	NK2R	HKTDSFVGLMa
Neuropeptide K	Tac1	NK2R	DADSSIEKQVALLKALYGH GQISHKRHKTDTSFVGLMa
Neuropeptide γ	Tac1	NK2R	DAGHGQISHKR HKTDSFVGLMa
Neurokinin B	Tac3	NK3R	DMHDFVGLMa
Hemokinin-1	Tac4 ⁴	NK1R	TGKASQFFGLMa
Endokinin-A	Tac4	NK1R	DGGEEQTLSTEAEWVVALEEGAG PSIQLQLQEVKTGKASQFFGLMa
Endokinin-B	Tac4	NK1R	DGGEEQTLSTEAEWEGAQ LQLQEVKTGKASQFFGLMa

¹Sequence and receptor data from Onaga (2014); Steinhoff et al. (2014). The sequences of the Tac1 and Tac3 derived peptides are conserved in humans, mouse and rat. ²These are also known as Ppt-a, Ppt-b and Ppt-c. ³Tac1 gene gives rise to four splice forms: a, b, g, and d, see Figure 3. ⁴Tac4 gene also gives rise to four splice forms, a, b, g, and d (see Figure 3), and two additional peptides (Endokinin-C and D) that are not TKs.

As another example, in zebrafish, *Tac1* transcript has been mapped to neurons in the olfactory bulb, telencephalon, preoptic region, hypothalamus, mesencephalon, and rhombencephalon, whereas *Tac3a* was observed in the preoptic region, habenula and hypothalamus and *Tac3b* predominantly expressed in the dorsal mesencephalon (Ogawa et al., 2012). Additional details of SP and NKA distribution are beyond the scope of this review.

A few further examples of TK signaling are given here that are of interest for the discussion of TK functions in invertebrates in later sections. Certain taste cells express NK1R, and SP appears to regulate responses not only to toxins, but also to tastants: spicy foods stimulate SP release to enhance umami taste reception (Onaga, 2014). In the dorsal horn of the spinal cord, SP modulates nociceptive signals relayed to the brain, but also in pain-processing areas of the brain cortex (Felipe et al., 1998; Zieglgänsberger, 2019). NKB regulates hormone (e.g., gonadotropin-releasing hormone, GnRH) release in the hypothalamus (Steinhoff et al., 2014). NKR are densely distributed in the rat olfactory bulb and it was shown that SP acts to depress neuronal activity in glomerular neurons, by triggering release of GABA (Olpe et al., 1987), similar to TKs and GABA in the antennal lobe of *Drosophila* (Ignell et al., 2009; Ko et al., 2015). The intestine is supplied by processes from TK-expressing neurons in dorsal root ganglia, or from local neurons (Höckfelt et al., 2001). NKR are widely expressed in the intestine (in a cell-specific manner) by enteric neurons, intestinal muscle, epithelium, vasculature as well as immune system. TK signaling in the gut thus influences motility, electrolyte and fluid secretion, as well as vascular and immune functions (Höckfelt et al., 2001).

Recently, all NKR were also shown to be expressed in genital organs and cells including the testis, sperm, ovary, granulosa cells, cumulus cells, and the uterus, and shown to be involved in sperm motility and reproduction (Pinto et al., 2010, 2015; García-Ortega et al., 2014; Candenias et al., 2018; Blasco et al., 2019).

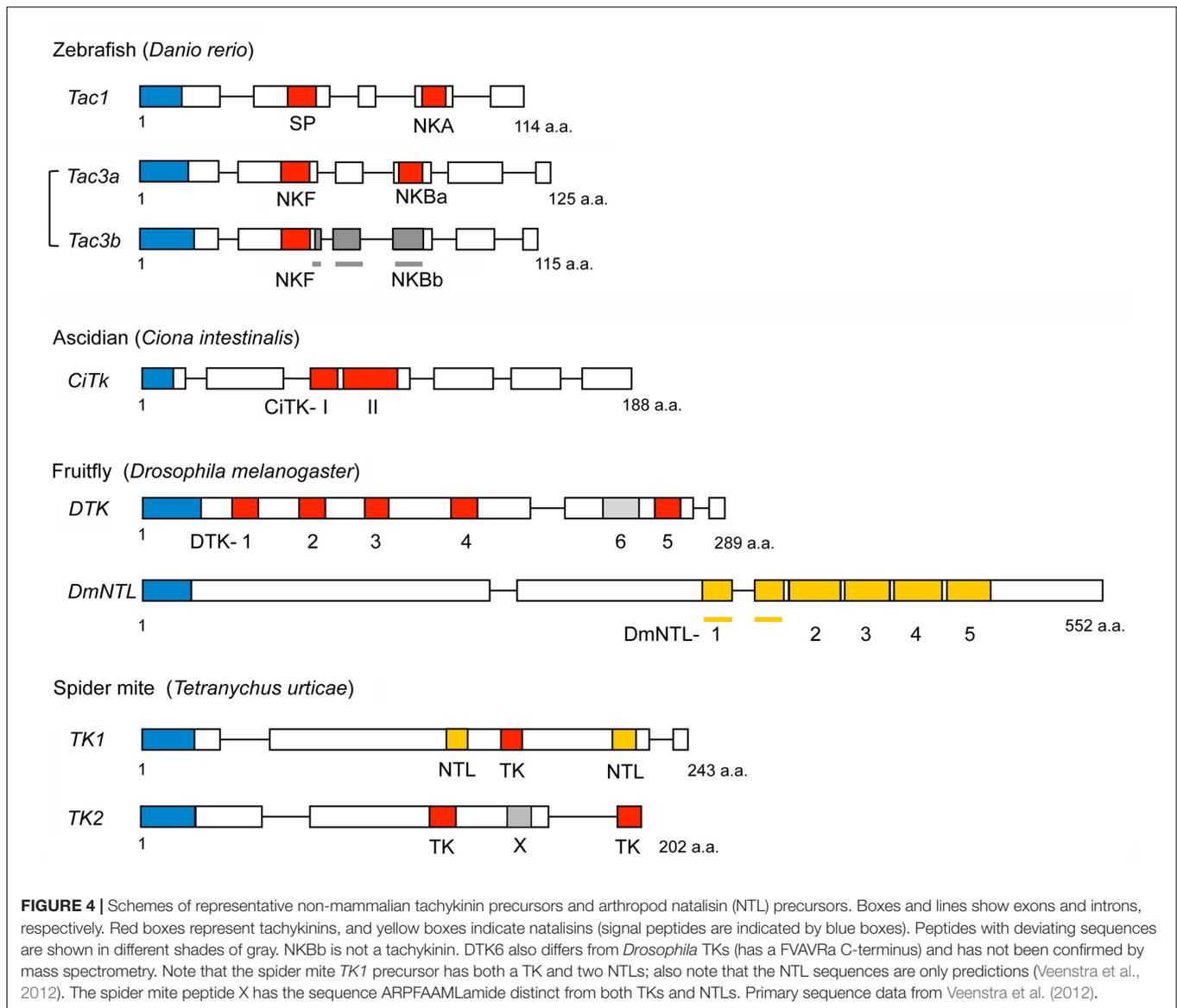
Substance P is also known to activate three Mas-related GPCRs (Mrgprs), a promiscuous group of receptors underlying

itch: human MRGPRX2, mouse MrgprA1, and mouse MrgprB2 (Bader et al., 2014). It is clear that the Mrgpr group is entirely separated from the TKR group (Bader et al., 2014). Interaction of SP with the Mrgprs induces elevation of intercellular Ca²⁺ (Azimi et al., 2016). The EC₅₀ value of SP for human MRGPRX2 is approximately 150 nM, while EC₅₀ values of SP for mouse MrgprA1 and MrgprB2 are about 5 μ M and 50 μ M, respectively (Azimi et al., 2016). Interestingly, analyses using *NK1R* knockout mice suggested that SP induces itch via Mrgprs rather than the NK1R (Azimi et al., 2017). Mrgprs are specific to mammals, suggesting that new SP-recognizing receptors arose during mammalian evolution.

TKs and Receptors in Protochordates and Non-mammalian Vertebrates

Genomes of non-mammalian vertebrates possess receptors that are homologous to NK1R – NK3R (Biran et al., 2012; Satake et al., 2013). Likewise, several genes encoding TK homologs are present in non-mammalian vertebrate species. *Tac1* and *Tac3* have also been identified from zebrafish, and the *Tac3* prototype gene appears to have duplicated to give rise to *Tac3a* and *Tac3b* (Biran et al., 2012) as shown in Figure 4. In addition, two *Tac3* have also been characterized from an eel, *Anguilla anguilla*, that is a basal vertebrate (Campo et al., 2018). Due to teleost-specific whole genome duplication multiple *Tac3* genes were generated in teleost fish (Moriyama and Koshihara-Takeuchi, 2018). *Tac3a* of zebrafish encodes not only an NKB (NKBa), but also an NKF that is a piscine-specific TK (Biran et al., 2012). *Tac3b* of zebrafish encodes both an NKB (NKBb) and an NKF, although NKBb contains an FVGLMamide sequence at the C-terminus that differs from TK consensus sequence FXGLMamide (Biran et al., 2012). Gene duplication of zebrafish TK receptor genes has also occurred, resulting in two *NK1R* genes (*Tacr1a*, *Tacr1b*) and three *NK3R* genes (*Tacr3a*, *Tacr3b*, and *Tacr3c*) (Biran et al., 2012). Both TACR3a and TACR3b are efficiently activated by zebrafish NKBa and NKF, and their interaction induces both elevation of intercellular Ca²⁺ and production of cAMP (Biran et al., 2012). EC₅₀ values of NKBb for TACR3a and TACR3b are 50–100 fold higher than that of NKBa for TACR3a and TACR3b (Biran et al., 2012). It is not clear whether homologs of HK and EKs are present in non-mammalian vertebrates, although it was proposed that *Tac4* is present in fish, including zebrafish, *Danio rerio* (Biran et al., 2012).

More than 20 TKs have been identified from skin secretion of frogs including, *Odorrana grahami*, *Rana chensinensis*, *Theioderma kwangsiensis*, *Kassina senegalensis*, and *Physalaemus fuscumaculatus* (Supplementary Table S3). These skin TKs possess the characteristic TK consensus sequence FXGLMamide (Bertaccini et al., 1965; Anastasi et al., 1977; Li et al., 2006; Wu et al., 2013; Zhang et al., 2013). The frog skin TKs are likely to act as exogenous factors, for instance as antimicrobial substances, rather than endogenous neuropeptides or hormones. An endogenous tachykinin has also been isolated from the brain of the frog, *Rana ridibunda* with a sequence homologous to that of NKB (O'Harte et al., 1991). Moreover, *Tac1* and *Tac3* of the frog, *Xenopus tropicalis*, have been predicted (registered



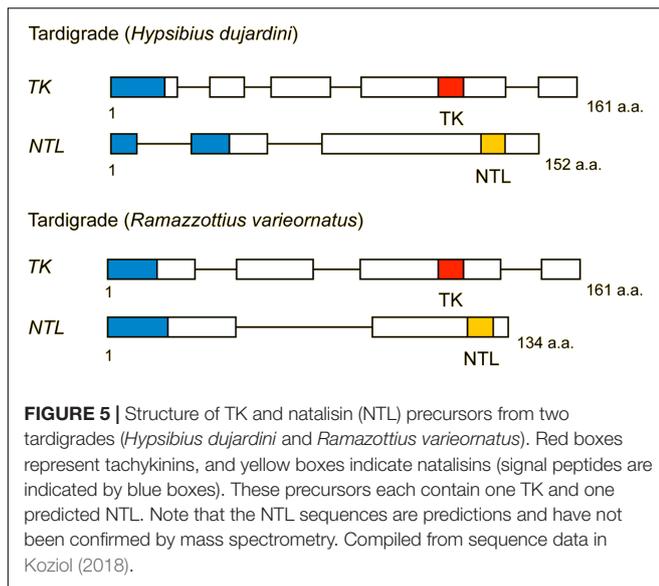
in NCBI databases). The *X. tropicalis Tac1* encodes an SP-like and an NKA-like peptide, while *Tac3* gene encodes an NKb-like and an NKF-like peptide. Interestingly, the sequence of the *T. kwangsiensis* skin TK is identical to that of the *X. tropicalis* SP-like peptide, except for two amino acid residues. The sequence of the *K. senegalensis* skin TK is also similar to that of SP (**Supplementary Table S3**). The skin TKs are amphibian-specific, suggesting that TKs acquired new functions in the amphibian lineage.

As shown in **Figure 4** and **Supplementary Table S1**, two TKs (CiTK-I, CiTK-II) were identified from the ascidian (protochordate), *Ciona intestinalis* (Satake et al., 2004). Like vertebrate TKs, CiTKs contain a C-terminal FXGLMamide (Satake et al., 2004), suggesting that the FXGLMamide sequence of TKs is highly conserved in Olfactores (vertebrates and ascidians). Since the ascidians are among the basal chordates, the *CiTk* gene is likely to correspond to a prototype of vertebrate

TK genes. This single *CiTk* gene in *Ciona* encodes CiTK-I and CiTK-II (Satake et al., 2004; **Figure 3B**), suggesting that gene duplications of a prototype *TK* gene have occurred in the vertebrate lineage, resulting in *Tac1*, *Tac3*, *Tac4*, and amphibian skin *TK* genes. Furthermore, CiTK-I and -II are located in the same exon of the *CiTk* gene (Satake et al., 2004), indicating that splice variants of the *CiTk* gene are absent (**Figure 4**). Therefore, alternative splicing of the *TK* gene also emerged during vertebrate evolution.

Homology search for mammalian NK1R - NK3R sequences using a *C. intestinalis* database², showed that only one homologous receptor is present in the ascidian (Satake et al., 2004). This receptor, CiTKR, can be activated by CiTKs (Satake et al., 2004), indicating that it is an authentic TK receptor. Phylogenetic analysis reveals that the CiTKR is sister

²<http://ghost.zool.kyoto-u.ac.jp/SearchGenomekh.html>



to the vertebrate TK receptor clade, which comprises NK1R - NK3R (Figure 7). These results suggest that NK1R - NK3R arose via duplication and diversification in the vertebrate lineage (Figure 7).

TKs and Their Receptors in Insects and Other Protostome Invertebrates

Tachykinins have been identified in a wide range of invertebrates, including annelids, mollusks, arthropods, tardigrades, echinoderms and tunicates, and tentatively in nematodes (see **Supplementary Table S1**). Several of these were isolated biochemically, others cloned, but most were identified by bioinformatics and subsequently confirmed by mass spectrometry. However, many TKs listed in the **Supplementary Table S1** have only been predicted from sequences identified from genomes and transcriptomes, based on similarity searches and await confirmation by mass spectrometry. There are some groups of invertebrates where TKs have not been identified or where sequences are only remotely similar to TKs. For instance in flatworms (Platyhelminthes) and cubomedusae (Cnidaria) no TKs have yet been discovered (McVeigh et al., 2009; Nielsen et al., 2019), and in sea anemones (Cnidaria) and nematodes (e. g. *C. elegans*) the “TK sequences” are not clearly TK-related (Ohno et al., 2017; Palamiuc et al., 2017; Hayakawa et al., 2019), see **Supplementary Table S1**. In fact, the proposed *C. elegans* TK-like receptor (NPR-22) is more closely related to RYamide/Luqin receptors than to TK-like receptors, and luqin-like peptides (from the LURY-1 precursor) are found in the worm and were shown to activate the receptor (Ohno et al., 2017; Yañez-Guerra et al., 2018). The previously proposed ligand (FMRFamide-like peptide 7, FLP-7) activates NPR-22 only at micromolar concentrations in a heterologous assay (Mertens et al., 2006; Palamiuc et al., 2017), suggesting that it is not a ligand [see (Ohno et al., 2017)]. However, it remains to be tested whether FLP-7 peptide is an NPR-22 ligand *in vivo*. Also, the *C. elegans*

TK-like peptide derived from NPL-8 (SFDRMGTEFGLM), does not activate the NPR-22 receptor (Mertens et al., 2006). Thus, the presence of a TK signaling system in *C. elegans* is still unresolved. However, TK-like receptors are found in Cnidaria (Anctil, 2009; Krishnan and Schiöth, 2015) (bioinformatics only), so the origin of TK signaling could possibly be traced to the common ancestor of Bilateria and Cnidaria. The presence of TK receptors has been demonstrated in the two major clades of Bilateria, the Nephrozoa (protostomes and deuterostomes) and its sister group the Xenacoelomorpha, that include Xenoturbella, Nemertodermatida, and Acoela (Thiel et al., 2018). In the Xenacoelomorpha, the presence is based on bioinformatics only.

With a few exceptions, each species has one gene encoding a TK precursor with multiple copies of TK peptides. As exceptions, two precursor genes were found in e.g., the limpet *Lottia gigantea* (Veenstra, 2010), the polychaete worm *Platynereis dumerilii* (Conzelmann et al., 2013), the crab *Carcinus maenas* (Christie, 2016), the tardigrade *Hypsibius dujardini* (Koziol, 2018) and the spider mite *Tetranychus urticae* (Veenstra et al., 2012; **Supplementary Table S1**). The organization of *Drosophila* and spider mite TK precursors is shown in **Figure 4** and that of the mosquito *Aedes aegypti* in **Figure 2B**. The number of peptides that can be cleaved from invertebrate TK precursors range from 1 in tardigrades (Koziol, 2018; **Figure 5**) to 15 in the cockroaches *Leucophaea maderae* and *Periplaneta americana* (Predel et al., 2005). Commonly, these peptides all have different, but related sequences (designated paracopies). In a few cases, the precursor only has several identical TKs, like in the crayfish *Procambarus clarkii*, which has seven CabTRP1 (Yasuda-Kamatani and Yasuda, 2004). The TKs are generally between 9 and 11 amino acids long, but a few have only 6, and others up to 18 as in cockroaches (Muren and Nässel, 1996; Predel et al., 2005), or even 37 residues as predicted in some scorpions and spiders (Veenstra, 2016). Interestingly, the N-terminally extended TKs of cockroaches have internal dibasic cleavage sites and it appears as if in the brain these are more likely to be processed and, thus, generate the shortened TKs, whereas the extended TKs are normally found in the midgut (Muren and Nässel, 1996, 1997; Winther et al., 1999; Predel et al., 2005).

All TKs are amidated, but only very few have been detected that may be N-terminally blocked by pyroglutamate (pQ), for instance in hemipteran bugs and the bivalve mollusk *Anodonta cygnea* (**Supplementary Table S1**). In insects and many other arthropods, it is common to find TKs with an N-terminal proline (P) in the second position (e.g., *Drosophila* DTK-2, APLAFVGLRa). This is likely to render the peptides sensitive to proline-specific dipeptidyl peptidase (DPP-IV) cleavage and inactivation (Nässel et al., 2000; Isaac et al., 2009). Thus, TKs can be specifically inactivated by DPP-IV selectively located in regions of the CNS or in the periphery (Nässel et al., 2000). Other peptidases that have been shown to inactivate TKs are nephriylsins, angiotensin converting enzymes and deaminases [see (Isaac et al., 2002, 2009)].

Two putative TK receptors were cloned from *Drosophila* before the endogenous ligands were known (Li et al., 1991; Monnier et al., 1992). Both receptors displayed significant similarities to mammalian TK receptors. One of these, designated

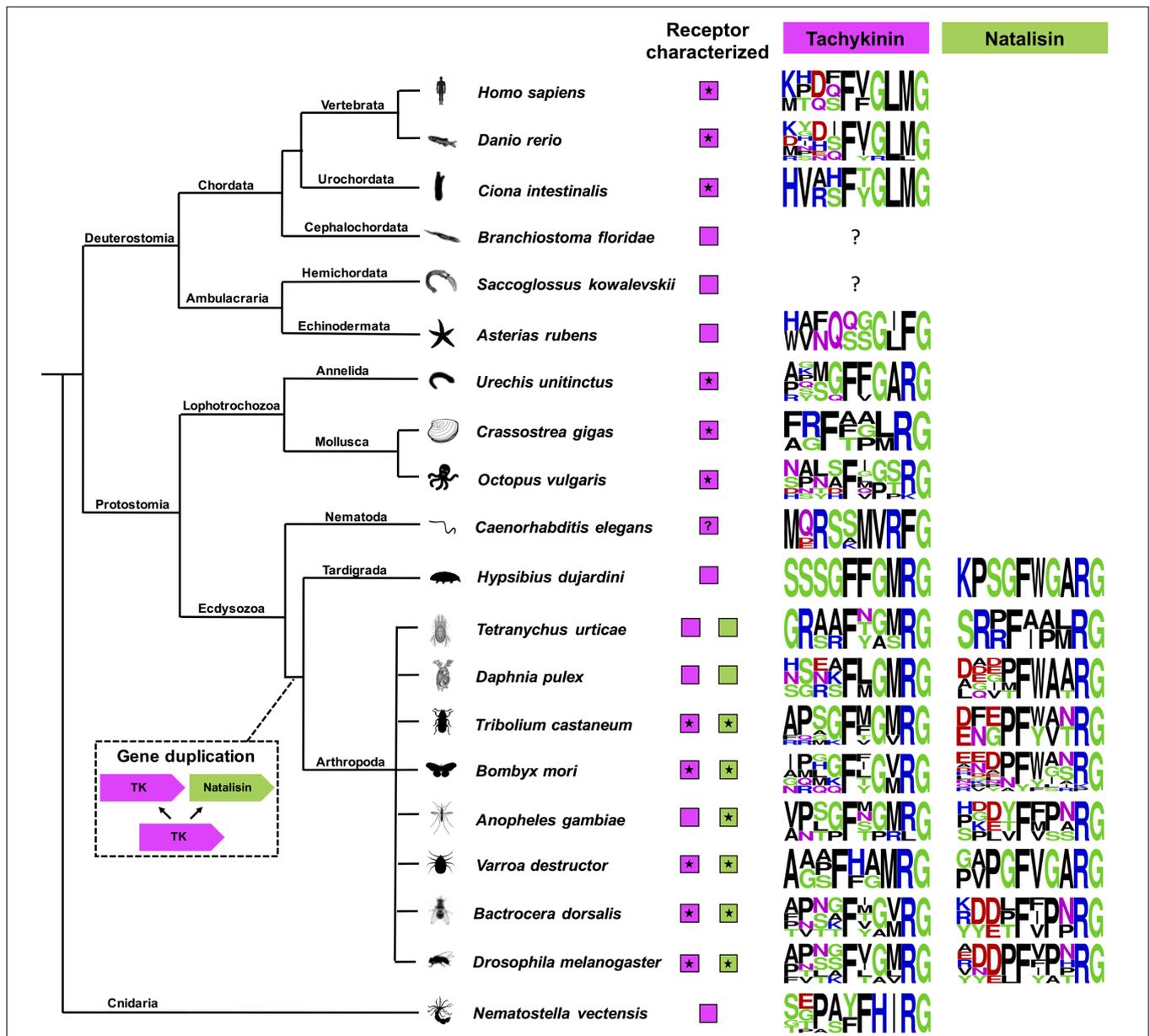


FIGURE 6 | A cladogram showing the occurrence of tachykinin and natalisin signaling systems in Bilateria and Cnidaria. Sequence logos of the peptides have also been provided. These were constructed using the final 10 amino acids at the C-terminus. Note that C-terminal amidation is not shown; it is represented by the amidation signal G. Species in which the receptors have been functionally characterized are indicated by asterisk. Note that the natalisin signaling system appears to have arisen in the lineage leading up to tardigrades and arthropods. The tachykinin-like peptide sequences in *Asterias rubens*, *Caenorhabditis elegans*, and *Nematostella vectensis* diverge from the canonical TK sequences in other phyla and should probably not be classified as TKs (see text). In *C. elegans* a TK receptor has not been functionally characterized, although an FLP-7/NPR-22 signaling system has been proposed (Palamiuc et al., 2017), but shown to represent a luqin/Ramide signaling system (Ohno et al., 2017). The peptide encoded by the TK2 precursor in *Hypsibius dujardini* looks more similar to arthropod natalisins. Precursors encoding TK-like peptides have not yet been identified in *Branchiostoma floridae* and *Saccoglossus kowalevskii*. However, TK-like receptors are present in all the animals presented here.

DTKR (CG7887) was confirmed as a receptor for endogenous *Drosophila* TKs (DTK-1-5) (Birse et al., 2006), the other NKD (CG6515) was first shown to respond to DTK-6, but not the other *Drosophila* TKs (Poels et al., 2009). DTK-6 has an FVAVRamide C-terminus instead of the common FX₁GX₂Ramide. Surprisingly, it turned out several years later

that NKD is a receptor for a novel family of neuropeptides called natalisins (NTL) that in *Drosophila* have a consensus sequence of FX₁X₂X₃Ramide (Jiang et al., 2013). One of these peptides, NTL4, has an FFATRamide, remotely similar to DTK-6, and indeed at high concentrations NTL4 activated the TK receptor DTKR (Jiang et al., 2013). The same authors also

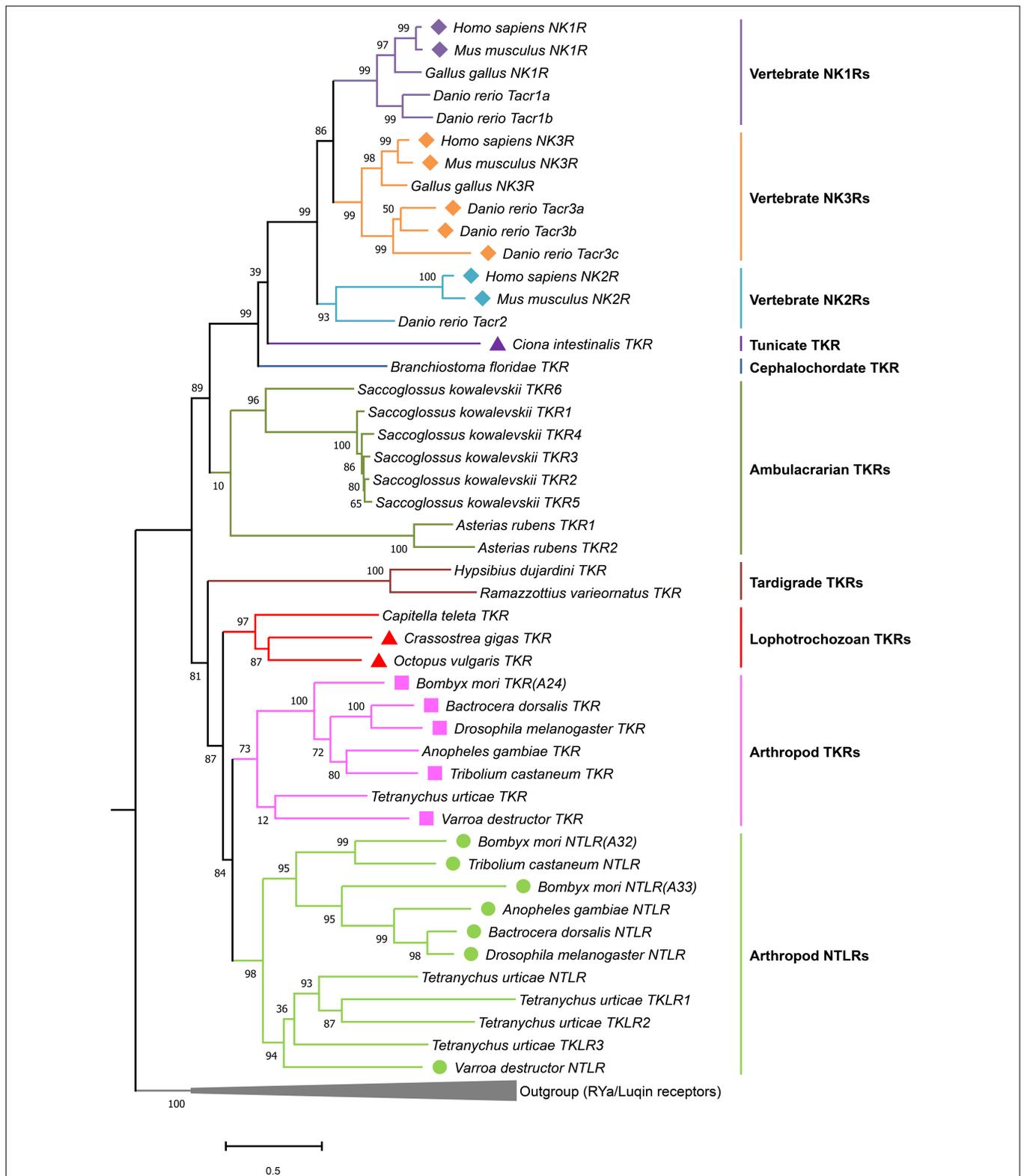
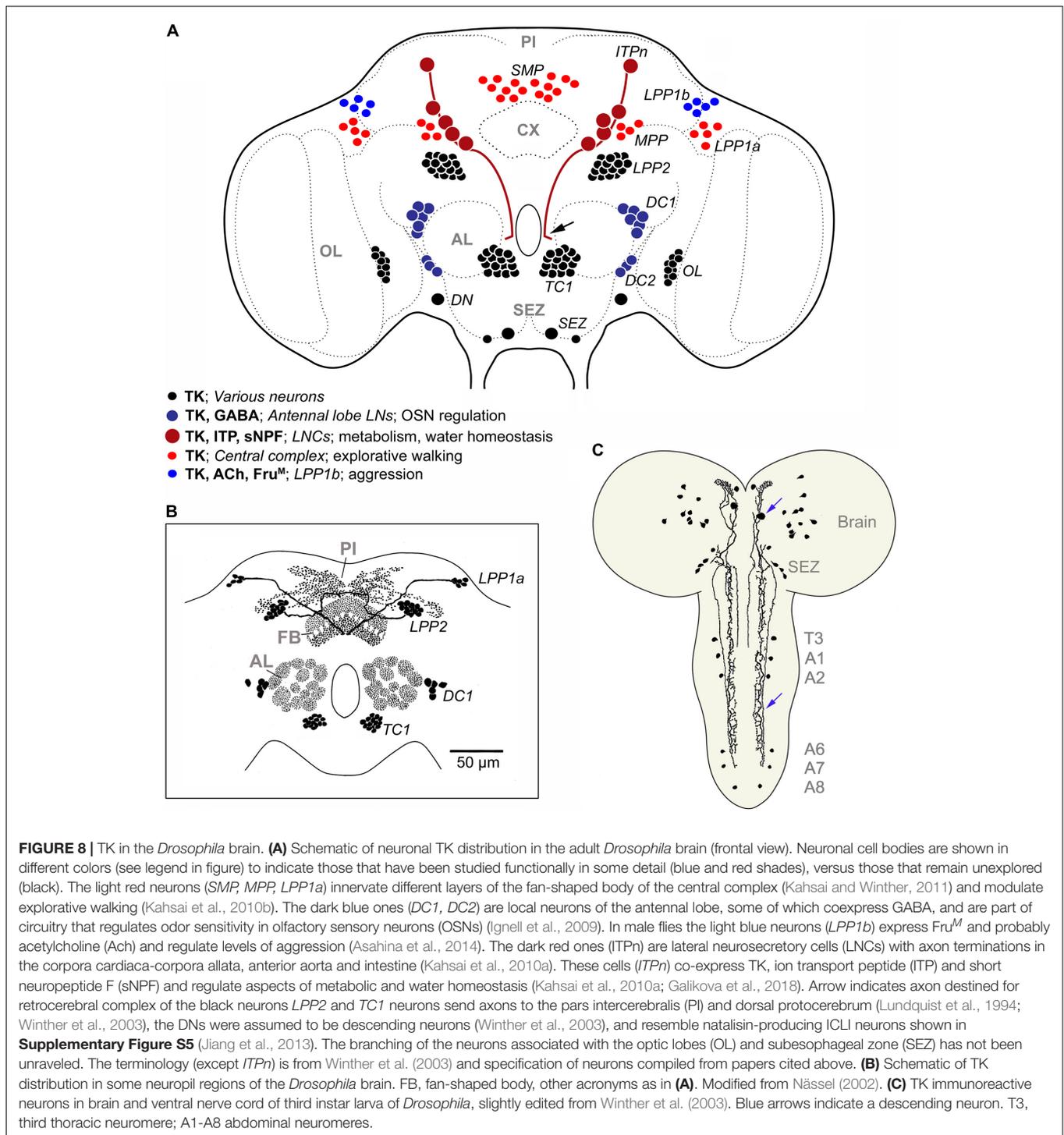


FIGURE 7 | Phylogeny of tachykinin and natalisin receptors. RYamide and luqin receptors were used as outgroup to root the tree. Amino acid sequences of full-length receptors were used for the analysis. Sequences were aligned using the MAFFT (E-INS-i algorithm and BLOSUM30 scoring matrix) and the phylogenetic tree constructed using the FastTree plugin in Geneious Prime (2019). Receptors that have been functionally characterized are indicated by a symbol before the species name. Note that in another annelid, *Urechis unicinctus* (among Lophotrochozoans), the TK receptor has been functionally characterized (not shown here). The figure was constructed in MEGAX. Sequences used to generate the phylogeny are provided in **Supplementary Material Text File S1**.



showed that DTK-6 at high concentrations activates the NTL receptor NTLR. In many insects, receptors of both DTKR- and NTLR-type have been identified (Jiang et al., 2013). In other invertebrates, such as for instance annelids (*Urechis*) and mollusks (*Octopus*) only receptors of DTKR type are known (Kawada et al., 2002; Kanda et al., 2007), suggesting that the NTL signaling system arose in the arthropod lineage (Jiang et al.,

2013). However, TK-like precursors with NTL-like peptides are found in the spider mite (chelicerate) as well as tardigrades; the latter suggesting that the NTL signaling might also be present outside arthropods. Natalisin signaling will be discussed in a separate section.

Some invertebrate TKs act on exogenous TK receptors in prey animals. TKs with FQGMRA C-termini (**Figure 2** and

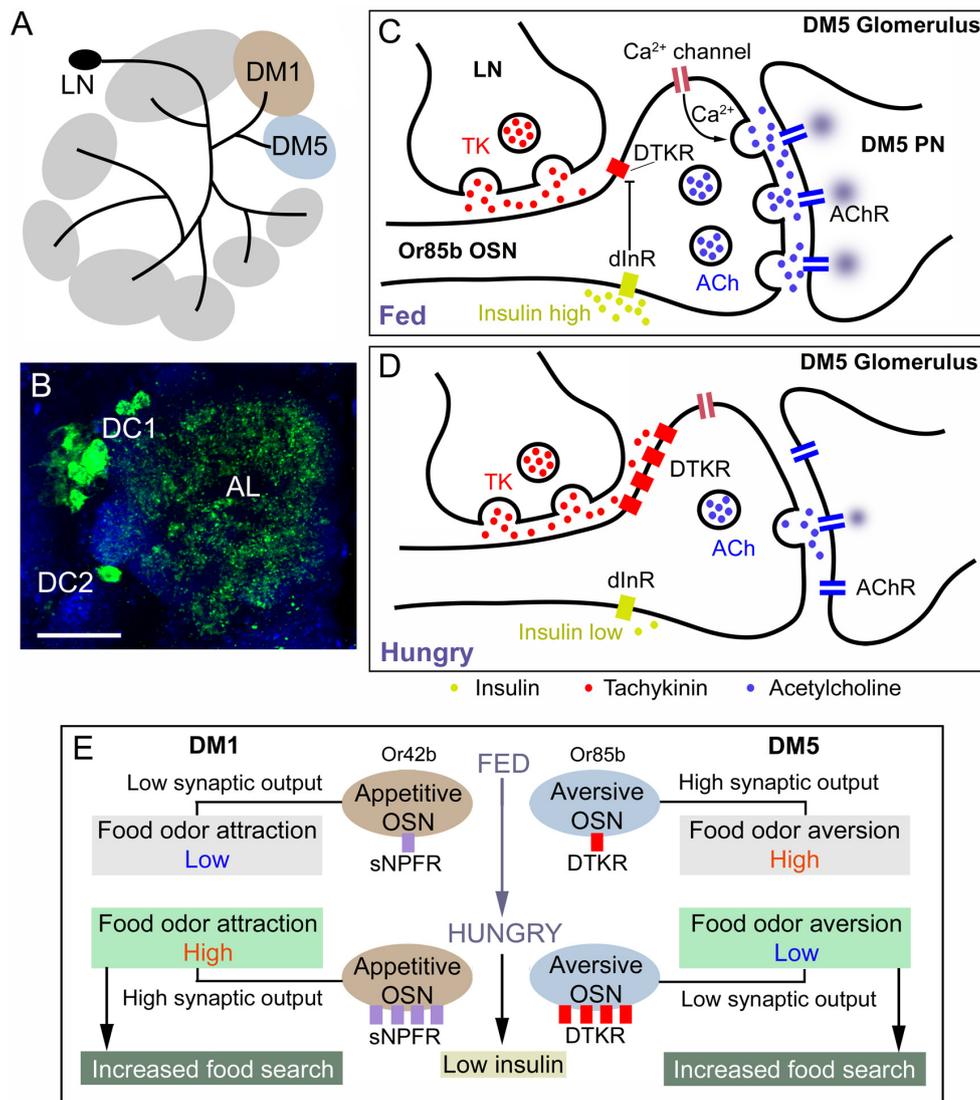


FIGURE 9 | Role of TK signaling in food odor sensing in the antennal lobe of *Drosophila*. **(A)** TK peptides are expressed in local neurons (LN) of the antennal lobe and innervate most glomeruli. Two glomeruli are shown here (DM1 and DM5). Of these DM1 is mediating food odor attraction (Or42b) and DM5 food odor aversion (Or85b). **(B)** Image of TK immunoreactive LNs (green) in the clusters DC1 and DC2 innervating antennal lobe (AL). **(C)** Role of TK signaling in the DM5 glomerulus, which relays aversive odor signals from olfactory sensory neurons (OSN) that express odorant receptors Or85b to DM5 projection neurons (PN) which in turn signal to higher order neurons that control food search. In the fed fly the circulating level of insulin-like peptide (DILP) is high, which suppresses expression (transcription) of the TK receptor DTKR. When DTKR signaling is low there is no suppression of Ca^{2+} channel activity and therefore release of acetylcholine (ACh) is strong when the OSN is activated and as a consequence the DM5 PN relays strong aversive signals and food search is reduced. **(D)** In the hungry fly the DILP level is low, DTKR expression is high and therefore TK signaling activates DTKR and the OSN releases less ACh resulting in suppressed activation of the aversive DM5 PN and therefore increase food search. **(E)** A scheme showing the combined signaling from the DM1 and DM2 signaling pathways that increase food search in hungry flies with low circulating insulin. In hungry flies the aversive DM5 odor pathway is inactivated resulting in increased food search (as detailed in **Figure 7C**). In the DM1 pathway (food odor attraction) signaling with short neuropeptide F (sNPFR) and its receptor sNPFR is increased in hungry flies due to low insulin and increased expression of the sNPFR. This leads to presynaptic potentiation of the ACh signaling and increased activation of DM1 PNs, resulting in increased food search. The panels **(A,C–E)** were redrawn from figures in Ko et al. (2015) and Jékely et al. (2018).

Supplementary Table S2) are produced by venom glands of the Jewel wasp *Ampulex compressa* and injected into the cockroach brain where action on the TK receptors leads to paralysis (Arvidson et al., 2019). Interestingly, these wasp venom TKs are injected as a precursor protein in the low pH venom, and not as cleaved peptides; only in the cockroach brain

with neutral pH they will be slowly liberated to act on TK receptors (Arvidson et al., 2019). Other TKs with C-terminal FXGLMamide are produced in salivary glands of the mosquito *Aedes aegypti* (sialokinins I and II), and the cephalopods *Eledone moschata* (eledoisin) and *Octopus vulgaris* (OctTK-I and II) (Erspamer and Anastasi, 1962; Champagne and Ribeiro, 1994;

Kanda et al., 2003). Presumably these TKs cause vasodilation in vertebrate prey animals [see (Erspamer and Erspamer, 1962; Beerntsen et al., 1999)]. The sequences of gland TKs are shown in **Figure 2** and **Supplementary Table S2**.

A Case of Novel Ligands for an Insect TK Receptor

Although most peptides and their cognate receptors co-evolve, there are a few interesting cases where receptors have adopted novel structurally unrelated ligands in addition to their original ligands. A well-known example is the *Drosophila* sex peptide, produced in male accessory glands and transferred to the female during copulation (Wolfner, 2002; Kubli, 2003). Sex peptide was adopted as an additional ligand for the myoinhibitory peptide (MIP) receptor (Kim et al., 2010). The *Drosophila* MIP receptor is the only known receptor for sex peptide. Another example, again in *Drosophila*, is the pigment-dispersing factor (PDF) receptor that has adopted DH31 as an additional ligand (Shafer et al., 2008; Goda et al., 2016). However, unlike sex peptide, DH31 also exerts its effects by binding to its own specific DH31 receptor (Johnson et al., 2005), suggesting that the PDF receptor is promiscuous. A similar phenomenon has been discovered for the silkworm *Bombyx mori* TK receptor (BNGR-A24), which seems to have adopted ion-transport peptide (extended ITPL isoform in particular) as a novel additional ligand (Nagai-Okatani et al., 2016). *Bombyx* ITPL is a large protein (79 amino acids), comprising 6 cysteine residues, which form three disulfide bridges, and is thus structurally very dissimilar to tachykinins (Roller et al., 2008; Nagai-Okatani et al., 2016). Nonetheless, *Bombyx* ITPL and TKs appear to be orthosteric ligands of the *Bombyx* TK receptor (BNGR-A24) based on heterologous and homologous cell culture experiments (He et al., 2014; Nagai-Okatani et al., 2016). Moreover, activation of BNGR-A24 by ITPL is coupled to the cGMP pathway, whereas BNGR-A24 activation by TKs can activate different second messenger pathways in a cell-type specific manner. More specifically, TK-mediated activation of BNGR-A24 in BmN cells has no effect on cAMP and cGMP levels, but if the receptor is expressed in HEK293 and Sf21 cells causes an increase in cAMP and Ca²⁺ levels (He et al., 2014; Nagai-Okatani et al., 2016). Thus, the two ligands of the *Bombyx* TK receptor may activate distinct second messenger pathways, at least *in vitro*. Since ITP or ITPL receptors have not been identified in any other species besides *Bombyx*, it remains to be determined if this phenomenon is widespread amongst insects, or whether it is just restricted to *Bombyx*.

TKs and Their Receptors in Deuterostome Invertebrates

TK-like receptors have also been mined from the genomes and transcriptomes of invertebrate deuterostome phyla such as Cephalochordata (e.g., *Branchiostoma floridae*), Hemichordata (e.g., *Saccoglossus kowalevskii*) and Echinodermata (e.g., *Asterias rubens*) (Jekely, 2013; Mirabeau and Joly, 2013; Yañez-Guerra et al., 2018). Phylogenetic analysis suggests that TK-like receptors and luqin/Ryamide-type receptors arose by gene duplication in a common ancestor of the Bilateria (Yañez-Guerra et al.,

2018). Precursors encoding TK-like peptides have also been predicted in the starfish, *Asterias rubens*, and brittle stars (**Figure 1** and **Supplementary Table S1**) (Semmens et al., 2016; Zandawala et al., 2017). However, the predicted TK-like peptides from *A. rubens* with an XXGL/IFamide C-terminus diverge substantially from FXGXRamide peptides of invertebrates, as well as from TKs of the protochordate *Ciona intestinalis* and vertebrates (FXGLMamide) (Semmens et al., 2016). Interestingly, the C-terminus of *A. rubens* (XXGL/IFamide) is somewhat similar to the proposed *C. elegans* tachykinin-like peptides (XMVRFamide) (Palamiuc et al., 2017), with peptides of both species having an Famide C-terminus and lacking the conserved phenylalanine residue in 5th position from the C-terminus. However, as mentioned above, the *C. elegans* TK-like receptor is more similar to RYamide/Luqin-like receptors, so it remains to be determined whether the predicted TK-like peptides in echinoderms are *bona fide* endogenous ligands for the echinoderm TK-like receptors. Moreover, sequences encoding TK-like receptors have been identified in the genomes and/or transcriptomes of hemichordates and cephalochordates (Jekely, 2013; Mirabeau and Joly, 2013; **Figure 6**). However, TK-like peptides have not yet been identified in these taxa. Perhaps the difficulty in discovering these peptides can be attributed to substantial diversification in the canonical sequence, which would render the homology-based search protocols ineffective.

EVOLUTION OF TACHYKININ SIGNALING COMPONENTS

We show a cladogram of TKs (**Figure 6**) and phylogenetic analysis their receptors (**Figure 7**) in the animal kingdom and the occurrence of natalisin signaling in some groups. TK signaling is evolutionary ancient (**Figure 6**) and is one of several peptide families that emerged before the split of deuterostomes and protostomes (Jekely, 2013; Mirabeau and Joly, 2013). A few recent studies suggest that it might be more ancient than previously thought. TK-like receptors were recently found in genomes of Xenacoelomorpha, which is a sister group of Nephrozoa (comprising deuterostomes and protostomes) (Thiel et al., 2018). However, no TK-like ligands were identified in these genomes. A TK-like GPCR has also been predicted in the genome of the sea anemone *Nematostella vectensis*, but this receptor is more closely related to other *Nematostella* neuropeptide GPCRs than it is to bilaterian TK receptors (Krishnan and Schiöth, 2015; Thiel et al., 2018). Most protostomes have a single TK receptor but protostomian TK precursors encode multiple TK peptides (**Figure 7**). Thus, it appears that protostomian TKRs can all be activated by the different TKs, as shown already in *Drosophila* and *Bombyx* (Birse et al., 2006; Jiang et al., 2013; Nagai-Okatani et al., 2016). Our phylogenetic analysis shows that a single TK-like receptor is also found in tardigrades (**Figure 7**; Koziol, 2018). Interestingly, tardigrades have 2 TK-like precursors, one of which encodes a peptide with sequence similarity to TKs and another one, which encodes a NTL-like peptide (**Figure 5**; Koziol, 2018). This suggests that NTL signaling may have arisen by the duplication of TK gene first

and a subsequent duplication and diversification of its receptor. Interestingly, the spider mite genome encodes multiple NTL-like receptors and a single TK receptor. In addition, it possesses two TK-like precursors, one of which contains TK-like and NTL-like peptides and another precursor with only TK-like peptides (Figure 5). This perhaps indicates a more advanced point in the diversification of the NTL-signaling, as the receptors now seem to have duplicated. Additional genomes of basal arthropods, onychophores and tardigrades need to be examined to determine the nature of TK-like and NTL-like signaling present in these animals before we can establish the precise lineage in which the NTL-signaling arose.

In deuterostomes, at least ancient Olfactores (vertebrates and ascidians) acquired a TK receptor that recognized a TK harboring the C-terminal FXGLMamide motif. The three subtypes of TK receptors, namely NK1R, NK2R, and NK3R, appear to have arisen following the whole genome duplications in the vertebrate lineage. Furthermore, these subfamilies might have acquired ligand selectivity during their diversification along with the generation of TK subtypes. The current missing pieces are echinoderm, acorn worm, and amphioxus counterparts, because canonical TKs have not yet been identified in genomes and/or transcriptomes of these deuterostome invertebrates. However, multiple TKRs are present in echinoderms and hemichordates suggesting additional independent gene duplication events within these lineages. In **Supplementary Figure S1** we show multiple sequence alignments of select TK and natalisin receptors.

TACHYKININS IN INVERTEBRATES, DISTRIBUTION AND FUNCTIONS

Overview of Functional Diversity From Early Studies

The first TKs isolated from insects were purified with the aid of a hindgut contraction assay that had also been utilized for first discovery of numerous other insect neuropeptides (Holman et al., 1990, 1991; Schoofs et al., 1990b). TKs from the annelid worm *Urechis unicinctus* were also purified with the aid of a muscle contraction assay (Ikeda et al., 1993). Thus, it was shown early that TKs are myostimulatory on a variety of muscles in the body wall, oviduct, foregut, hindgut, as well as heart [see (Ikeda et al., 1993; Schoofs et al., 1993; Nässel, 1999; Sliwowska et al., 2001)]. Examples of other functions established before employment of genetic tools are modulation of network activity in the stomatogastric ganglion of crustaceans [(Blitz et al., 1995), reviewed in Nusbaum and Blitz (2012), Nusbaum et al. (2017)], activation of dorsal unpaired median neurons in locust (Lundquist and Nässel, 1997), stimulation of release of adipokinetic hormone from locust corpora cardiaca (Nässel et al., 1995b), diuretic action on Malpighian tubules of locust *L. migratoria* and moth *Manduca sexta* (Skaer et al., 2002; Johard et al., 2003) and presynaptic inhibition of crayfish photoreceptors, likely as a co-transmitter of GABA (Glantz et al., 2000). From numerous *in vitro* studies, there is no evidence

that the different paracopies of TKs in a species have any major differential activities or functions, except possibly DTK-6 in *Drosophila*, but it should be noted that the presence of this mature peptide has not been verified by mass spectrometry. Additional functions of TKs discovered using various approaches are discussed separately in the context of different species in sections “Distribution and Function of TKs in Invertebrates” and “Functional Roles of TKs in *Drosophila*, Genetic Advances”.

Distribution and Function of TKs in Invertebrates

Early work used antisera to substance P and other vertebrate tachykinins to localize presumptive TK neurons in the CNS of several invertebrates summarized in Nässel (1999). It should be noted that the earliest of these studies were performed before neuronal/intestinal TKs had been isolated and sequenced in invertebrates. In retrospect, it appears that of the TK antibodies used during this era, one monoclonal antibody (Cuello et al., 1979) actually recognizes the invertebrate TKs [see (Blitz et al., 1995; Johansson et al., 1999; Nässel, 1999)], whereas the polyclonal ones, except anti-Kassinin (Lundquist et al., 1994), seem to label other epitopes, at least in insects. Thus, TK distribution in several crustaceans (Sandeman et al., 1990; Schmidt and Ache, 1994; Blitz et al., 1995; Johansson et al., 1999), and the horse shoe crab *Limulus polyphemus* (Chamberlain and Engbretson, 1982; Mancillas and Selverston, 1985) is likely to be correctly described in these earlier studies. The first antisera to invertebrate TKs were raised against locust LomTK-I (Nässel, 1993) and LomTK-II (Vitzthum and Homberg, 1998), blowfly CavTKII (Nässel et al., 1995a) and cockroach LemTRP1 (Winther and Nässel, 2001) and these were subsequently used in a large number of invertebrate species, some of which are outlined below.

Insects

The neuronal distribution of TK immunoreactivity in the CNS is in general fairly well conserved between insects studied, whereas in other arthropods only some features seem to be shared with insects. Characteristic of TK distribution in insects is presence in neuronal processes in antennal lobes, central complex, pars intercerebralis, dorsolateral protocerebrum, optic lobes and subesophageal zone. First, we will outline the neuronal localization of TK in *Drosophila* and a few other insects, then move on to crustaceans, and snails. For these organisms, except *Drosophila*, we also briefly describe TK functions. Functional roles of TK signaling in *Drosophila* are described in a separate section.

In the adult *Drosophila* brain, *in situ* hybridization and immunolabeling revealed that there are more than 160 TK-expressing neurons that can be divided up into 11 bilateral groups and one unique pair (Figure 8A; Winther et al., 2003). Ten large lateral neurosecretory cells (*ITPn*) express TK, as well as two other peptides (short neuropeptide F, sNPF, and ion transport peptide, ITP) (Kahsai et al., 2010a). The other TK neurons are interneurons of different kinds innervating the fan-shaped body of the central complex, the antennal lobes, the optic lobes, pars intercerebralis, dorsal lateral protocerebrum and the

subesophageal zone (Winther et al., 2003). Some of these TK neuron clusters have been functionally investigated by genetic manipulations (colored cells in **Figure 8A**), whereas the functions of other clusters (black cells in **Figure 8A**) remain obscure. Functional aspects will be discussed later in a separate section. Details of some of the TK neurons are shown in **Figure 8B**. In the third instar larva there are only about 44 neurons in the entire CNS that are consistently labeled by TK antisera; 32 of these are in the brain and SEZ (**Figure 8C**) (Siviter et al., 2000; Winther et al., 2003). In both larvae and adults, enteroendocrine cells of the midgut and anterior hindgut express TK (Siviter et al., 2000; Veenstra et al., 2008; Veenstra, 2009).

In some cells in *Drosophila*, TK is colocalized with other neuropeptides or GABA (**Supplementary Table S4**): in neurosecretory cells (ITPn) with sNPF and ITP, in local neurons of the antennal lobe with either GABA, allatostatin-A or myoinhibitory peptide (MIP; also known as allatostatin-B) and in midgut enteroendocrine cells with either neuropeptide F or diuretic hormone 31 (Veenstra et al., 2008; Ignell et al., 2009; Carlsson et al., 2010; Kahsai et al., 2010a). In addition, single cell transcriptome sequencing of brain neurons shows that TK is coexpressed with glycoprotein hormone beta 5 (GPB5) (Davie et al., 2018). A more systematic screen of colocalized substances in the insect CNS would probably make this list longer.

The blowfly *Calliphora vomitoria* displays a neuronal distribution of TK very similar to that in *Drosophila* (Lundquist et al., 1994). Studies of TK distribution in other insects reveal many similarities, except that the numbers of neurons in the different clusters vary between species, as shown next.

In the brain of the honeybee *Apis mellifera* TK distribution was mapped by *in situ* hybridization (Takeuchi et al., 2004). Neuronal cell bodies were revealed in association with the central complex, antennal lobes and optic lobes, as in other insects, but also associated with the mushroom body calyces. The intrinsic mushroom body neurons were identified as the small-type Kenyon cells (class I and II) (Takeuchi et al., 2004). Later, immunolabeling also confirmed presence of TK in Kenyon cells including their axons in the lobes (Heuer et al., 2012). The distribution of TK transcript is spatially similar irrespectively of sex, cast, or division of labor of workers: however, quantitatively transcript levels are higher in queens and foragers than in nurse and drone bees (Takeuchi et al., 2004). In bees, the TK in mushroom bodies may be involved in regulation of foraging and social behaviors (Takeuchi et al., 2004; Brockmann et al., 2009; Boerjan et al., 2010). Also in other hymenopterans (Oya et al., 2017) and in the beetle *Tribolium castaneum* (Binzer et al., 2014), TK was identified in major subpopulations of Kenyon cells, but in other studied insects there are so far no reports of such neurons producing TKs. In the honeybee, quantification of neuropeptides by mass spectrometry was performed after foraging nectar or pollen (Brockmann et al., 2009). TK was among the three peptides whose levels were most affected in association with foraging for nectar or pollen.

In the moth *Spodoptera litura*, at least 80 TK neurons were detected in the adult brain, and the innervation of the central complex, the antennal lobes, pars intercerebralis, dorsal lateral protocerebrum and the subesophageal zone is similar to that

in *Drosophila* (Kim et al., 1998; see **Supplementary Figure S2**). Also a pair of large descending neurons was identified. A special feature of the moth is the presence of TK expression in 8 median neurosecretory cells with axon terminations in the retrocerebral complex and anterior aorta (Kim et al., 1998). Similar median neurosecretory cells (MNCs) were also seen in the moth *Heliothis virescens* (Zhao et al., 2017), and the beetles *Tenebrio molitor* and *Zophobas atratus* (Sliwowska et al., 2001). In another moth *Manduca sexta*, the TK distribution in the brain (except the MNCs) and intestine was found similar to *S. litura* and it was shown that TK stimulates secretion in the Malpighian tubules *in vitro* (Skaer et al., 2002).

The brains of the cockroach *Leucophaea maderae* and the locust *Locusta migratoria* contain far larger numbers of TK neurons, but the innervation pattern of brain regions is similar to *Drosophila* and moth (Nässel, 1993; Muren et al., 1995; Vitzthum and Homberg, 1998). In the *L. maderae* brain (without optic lobes), about 360 TK neurons were found (Muren et al., 1995; **Supplementary Figure S3**), and about 800 in the entire brain of *L. migratoria* (Nässel, 1993). In contrast to *Drosophila*, there are efferent TK neurons in the cockroach abdominal ganglia that innervate the hindgut and TK neurons in the stomatogastric ganglia that supply extensive axon terminations over the foregut and midgut (Muren et al., 1995; Nässel et al., 1998). In locusts, TK neurons in the lateral neurosecretory cell group send axons to the corpora cardiaca where they contact cells producing adipokinetic hormone (AKH) and it was shown that TKs induce AKH release *in vitro* (Nässel et al., 1995b). These neurosecretory cells may be analogous to the ITPn neurons in *Drosophila* (Kahsai et al., 2010a), although a role in hormone release was not yet analyzed in the fly. In both locust and cockroach midgut, endocrine cells express TK, and in the locust there are also TK producing endocrine cells in the six midgut ampullae at the base of the Malpighian tubules (Muren et al., 1995; Winther and Nässel, 2001). TKs stimulate secretion in locust Malpighian tubules (Johard et al., 2003). Calcium-dependent release of TK from the cockroach and locust intestine could be induced by potassium application, and TK was demonstrated in hemolymph, suggesting that hormonal release of intestinal TK regulates tubules secretion (Winther and Nässel, 2001). In locusts, several cases of colocalization of TK and other peptides have been demonstrated. The endocrines of the ampullae (but not in the rest of the midgut) coexpress TK, diuretic hormone 44 (DH44) and FMRFamide-like peptide, and TK was shown to stimulate secretion in locust Malpighian tubules together with DH44 (Johard et al., 2003). In certain central complex neurons there is co-expression of TK and leucokinin and in others TK and octopamine or GABA (Vitzthum and Homberg, 1998). Finally, sensory neurons of the metathoracic legs co-express TK, allatotropin, FMRFamide-like peptide and probably acetylcholine (Persson and Nässel, 1999; **Supplementary Table S4**).

Another insect studied in some detail is the hemipteran blood-sucking bug *Rhodnius prolixus* where a total of about 250 TK immunoreactive neurons were found in the brain (Kwok et al., 2005). These are distributed in the optic lobes and in several other clusters in the midbrain. Interestingly, no TK containing enteroendocrine cells were detected in this species, in contrast

to many other studied insects, but the hindgut is innervated by TK axons (Kwok et al., 2005). These TK axons also express leucokinin, another myostimulatory peptide (Haddad et al., 2018). *Rhodnius* TKs were shown to increase the basal tonus of the hindgut, but also to increase the frequency and amplitude of peristaltic contractions of the salivary gland, a tissue that displays high levels of TK transcript, but no immunoreactive TK (Haddad et al., 2018).

Some further functions of TKs in insects other than *Drosophila* have been explored that might shed light on TK signaling in general. In female burying beetles, *Nicrophorus vespilloides*, neuropeptides were quantified in solitary virgins, individuals actively parenting or post-parenting solitary adults to identify neuropeptides associated with parenting (Cunningham et al., 2017). TK was found as one of the few peptides associated with active parenting. In several insects, including *Drosophila*, oriental fruitfly, cockroaches and moths, TK seems to play a role in modulation of olfactory sensory processing (Ignell et al., 2009; Jung et al., 2013; Fusca et al., 2015; Ko et al., 2015; Gui et al., 2017b; Lizbinski and Dacks, 2017). Other functions have not been investigated in multiple species; however, immunocytochemistry suggests some conservation of the distribution of TK in neurons of specific brain centers, and intestine in insects and crustaceans that might also reflect functional conservation.

Crustaceans

The TK distribution in the brain of a few crayfish, lobster and crab species has been studied, mostly with a monoclonal antibody to substance P (Goldberg et al., 1988; Sandeman et al., 1990; Schmidt and Ache, 1994; Schmidt, 1997a,b), but one study also employed antiserum to a cockroach TK that is nearly identical to crab TK (Johansson et al., 1999). These studies have mostly focused on TK neurons in the olfactory centers of the brain, but also the stomatogastric nervous system.

As seen in **Supplementary Figure S4**, brain of crayfishes possesses a pair of TK interneurons with large cell bodies and extensive processes in the anterior deutocerebrum and varicose branches among the cell bodies of a group of olfactory interneurons in the lateral deutocerebrum (Sandeman et al., 1990; Johansson et al., 1999). There is another pair of TK neurons with deutocerebral cell bodies and processes in the neuropil of the olfactory lobe, as well as larger numbers of small TK neurons with processes in the olfactory and accessory lobes (Sandeman et al., 1990; Johansson et al., 1999). TK neurons were shown in all the neuropils of the optic lobes of the crayfish and specifically a set of TK and GABA expressing amacrine cells were identified in the lamina ganglionaris (Glantz et al., 2000). This study shows that application of GABA and TK to photoreceptor terminals in the lamina induces a short-latency, dose-dependent hyperpolarization with a decay time of a few seconds. TK also acts over several minutes to reduce the photoreceptor potential to potentiate the action of GABA (Glantz et al., 2000). In the American lobster, the distribution of TK processes is in general similar to that seen in insects with TK immunolabeling in the protocerebral bridge, central body, olfactory (antennal) lobes, and anterior median protocerebral neuropil (Langworthy et al., 1997). Like

in insects, midgut enteroendocrine cells in crabs also express TK (Christie et al., 2007).

In decapod crustaceans the stomatogastric nervous system (STN) consists of 25–30 neurons (depending on species) and controls handling of ingested food. Most of these neurons contribute to the activity in one or both of the neural networks in the STN, which regulate (1) gastric mill (chewing) or (2) pyloric circuit (pumping and filtering of food that has been chewed) [see (Nusbaum et al., 2017)]. A pair of neurons (MCN1) that innervate the stomatogastric ganglion produces the TK CabTRP-Ia (Blitz et al., 1995; Christie et al., 1997). The MCN1s also produce GABA and the peptide proctolin (Nusbaum et al., 2017). It was shown that CabTRP-Ia and GABA released from MCN1 are critical for activation of the gastric mill rhythm, whereas MCN1 release of CabTRP-Ia and proctolin predominantly excites the pyloric rhythm (Nusbaum et al., 2017).

Mollusks

A few mollusks have been investigated with respect to TK distribution (using antiserum to locust TK). These include the pond snail *Lymnaea stagnalis*, the pulmonate terrestrial snail *Helix pomatia* and the freshwater bivalve, *Anodonta cygnea* (Elekes and Nässel, 1994; Elekes et al., 1995). In *L. stagnalis*, about 180 TK neurons were found, distributed in cerebral and pedal ganglia and TK axons were detected in the intestine (Elekes et al., 1995). About 900 TK neurons were seen in *H. pomatia* with about 80% of these in cerebral ganglia, whereas in *A. cygnea* only a smaller number of TK neurons was detected in cerebral, pedal and visceral ganglia (Elekes and Nässel, 1994; Elekes et al., 1995). In the snails, a large number of TK neurons are located in procerebrum of the cerebral ganglia (**Supplementary Figure S5**). The procerebrum is an association center for olfactory information similar to the mushroom bodies of insects and thus TKs seem to be involved in olfactory processing also in mollusks (Elekes and Nässel, 1994). Recently, a TK receptor related to the *Drosophila* DTKR and responding to endogenous TKs was identified in the bivalve mollusk *Crassostrea gigas* (Dubos et al., 2018). In the snail *Helix*, the neuronal membrane effects of locust LomTK-I, and anodontatachykinin, were either depolarizing or hyperpolarizing depending on neuron-type, and voltage-clamp experiments revealed a role of Ca²⁺- or K⁺-currents in these peptide effects (Elekes et al., 1995).

Nematode Worms

In *Caenorhabditis elegans*, the gene FLP-7 was considered to encode a TK precursor ortholog (Palamiuc et al., 2017). However, as mentioned in section “TKs and Their Receptors in Insects and Other Protostome Invertebrates,” this gene encodes FMRFamide-like peptides (Mertens et al., 2006) and the proposed receptor gene (NPR-22) is only remotely related to TK receptors, and more closely related to the RYamide/Luqin receptor (Ohno et al., 2017; Yañez-Guerra et al., 2018; see **Figure 6**). Nevertheless, since the signaling system was referred to as a TK system (Palamiuc et al., 2017) we summarize the findings here. It was shown that FLP-7 is expressed in several tissues, including the head, the nervous system, and the sensillum (wormbase.org). At the cellular level, a fluorescent transgenic reporter line revealed that FLP-7 is

expressed in the ALA motor- and the AVG interneurons and in the ASI sensory neurons, and that the reporter is secreted into the “circulation” (Palamiuc et al., 2017). The ALA motor neuron has been shown to regulate locomotion, the AVG neuron influences ventral cord development, and the ASI sensory neuron pair regulates whole body physiology during development, controls lifespan via neurohormones, and regulates 5-HT-induced fat loss (Palamiuc et al., 2017). FLP-7 was shown to act in the intestine to induce lipase activity and fat loss (Palamiuc et al., 2017). In another study, the ligands of NPR-22 were found to be the luqin-like peptides LURY-1 and 2 (AVLPRYa and PALLSRYa) encoded on the gene Y75B8A.11 (Ohno et al., 2017). The LURY peptides are secreted from pharyngeal neurons and regulate feeding, lifespan, egg-laying and locomotor activity (Ohno et al., 2017). In summary, no clear-cut TK signaling system has been discovered in *C. elegans* so far.

Ambulacraria

Tachykinin expression has not been mapped yet in echinoderms or hemichordates.

Functional Roles of TKs in *Drosophila*, Genetic Advances

With the introduction of the binary Gal4-UAS system (Brand and Perrimon, 1993) it became possible to genetically target components of the TK signaling system spatially and temporarily, and thus knock down or increase activity in a neuron-specific fashion. **Table 2** summarizes the known functions of TKs in *Drosophila*. A first study, utilized *Tk*-RNAi to broadly knock down TK production in neurons by means of ubiquitously expressed drivers (*Elav*- and *tubulin-Gal4*) and monitor effects on olfaction and locomotion (Winther et al., 2006). The flies with globally reduced TK signaling displayed decreased responses to certain odors and were hyperactive in locomotor assays. Subsequent studies describe more targeted manipulations where TK functions in smaller populations of neurons could be revealed.

It was found that the TK receptor DTKR is expressed by olfactory sensory neurons (OSNs) of the *Drosophila* antennae and TK in subpopulations of the local neurons (LNs) of the antennal lobe (Ignell et al., 2009). TK signaling from LNs to OSNs provides presynaptic inhibitory feedback by suppressing calcium and synaptic activity (Ignell et al., 2009). An ensuing study revealed further details on the role of TK signaling in olfaction and food search (Ko et al., 2015). It was shown that in hungry flies where circulating levels of insulin-like peptide (ILP) are low there is an upregulation of the DTKR in OSNs carrying specific odorant receptors (Or42b and Or85a) (**Figure 9**). In the antennal glomerulus DM5, which conveys food odor aversion (negative valence), upregulation of the inhibitory DTKR in a hungry fly leads to increased TK signaling and thus suppressed depolarization and as a consequence decreased synaptic activation of antennal lobe projection neurons (PNs) leading to increased food attraction (Ko et al., 2015). When the fly has fed, and circulating insulin is high, the DTKR expression decreases due to activation of the insulin receptor in OSNs, synaptic signaling increase and food aversion is augmented

TABLE 2 | Functions of TKs in *Drosophila*.

Neurons targeted ¹	Functional TK role indicated	References
Global Tk-RNAi	Modulation of odor sensitivity	Winther et al., 2006
Global Tk-RNAi	Modulation locomotor activity	Winther et al., 2006
OSNs in AL (Dtkr)	Presynaptic inhibitory feedback to OSNs	Ignell et al., 2009
OSNs in AL (Dtkr)	Starvation-induced increase in odor sensitivity	Ko et al., 2015
Neurons in SEZ (Tk)	Modulation of pheromone response (via Gr68a)	Shankar et al., 2015
Central complex (Tk)	Modulation of explorative walking	Kahsai et al., 2010a
Brain neurons (Tk)	Modulation of aggression level (fruitless neurons)	Asahina et al., 2014
IPCs (Dtkr)	Regulation of insulin production	Birse et al., 2011
Nociceptive cells (Dtkr)	Modulation of nociception in sensory cells	Im et al., 2015
ITPn (brain NSCs; Tk)	Regulation of metabolic stress responses	Kahsai et al., 2010a
ICNs (brain neurons) ²	Inhibit larval IPCs, affect growth via IIS and EGF	Meschi et al., 2019
Endocrines in gut (Tk)	Regulation of lipid metabolism in intestine	Song et al., 2014
TK endocrines in gut ³	IMD-mediated DILP upregulation; organismal growth	Kamareddine et al., 2018
<i>In vitro</i> TK application	Induces midgut contraction	Siviter et al., 2000
<i>In vitro</i> TK application	Modulation of heart contraction rate	Schiemann et al., 2018

¹In brackets: genetic interference with peptide (Tk) or receptor (Dtkr). ²Ablation of TK-expressing ICN neurons, activation or inactivation of neurons. ³Interference with signaling in TK expressing endocrines. OSN, olfactory sensory neurons; AL, antennal lobe; SEZ, subesophageal zone; IPC, insulin-producing cells; ITPn, ITP neurons; NSC, neurosecretory cell; ICN, IPC contacting neurons; EGF, epidermal growth factor; IMD, immune-deficiency pathway; DILP, *Drosophila* insulin-like peptide.

(**Figure 9**). In glomerulus DM1 (positive valence; wired for food odor attraction) innervated by Or42b expressing OSNs, enhanced signaling with sNPF increases food attraction in hungry flies with low circulating insulin (Ko et al., 2015). This enhanced signaling is caused by up-regulation of sNPF receptor expression on OSNs and strengthened synaptic activation of PNs (**Figure 9**). Together, peptidergic neuromodulation of the two odor channels (DM1 and DM5) ensures that hungry flies increase food search. Whereas it has been shown that sNPF facilitates cholinergic transmission in OSNs to PNs (Ko et al., 2015), it is not clear whether TK acts to modulate inhibitory GABA transmission in LNs (Ignell et al., 2009). Also in the cockroach *Periplaneta americana* (Jung et al., 2013) and the oriental fruitfly *Bactrocera dorsalis* (Gui et al., 2017b), TK signaling modulates olfactory sensitivity, and the presence of TK in antennal lobe neurons in all studied insects may suggest a conserved role in olfaction.

In the central complex of *Drosophila*, TK is found in a few sets of neurons (light red neurons in **Figure 8A**) and in assays of explorative walking TK knockdown in some of these neurons resulted in flies with increased center zone avoidance, whereas knockdown in other neurons resulted in flies with increased activity-rest bouts (Kahsai et al., 2010b). Thus, TK in

the central complex seems to be important for modulation of spatial orientation, activity levels, and temporal organization of spontaneous walking.

A small set of protocerebral TK neurons (light blue in **Figure 8A**) have been shown to regulate levels of aggression in male *Drosophila* (Asahina et al., 2014). These TK neurons are a small subpopulation of the numerous neurons that express the male splice form of fruitless ($FruM^+$), a transcription factor that specifies male-specific behavior, including male aggression. Thus, a set of 4 pairs of neurons in the brain designated $Tk-GAL4^{FruM}$ neurons control the level of male-male aggression, but have no influence on male-female courtship behavior (Asahina et al., 2014). The same authors found that the $Tk-GAL4^{FruM}$ neurons also may be cholinergic (express marker for acetylcholine signaling) and that this neurotransmitter may thus play an additional role in the circuit.

Another male-specific TK circuit in *Drosophila* is involved in gustatory detection of an anti-aphrodisiac pheromone (CH503). Gustatory cells (Gr68a) in the forelegs respond to this pheromone and mediate signals to central brain circuits via 8 to 10 TK neurons located in the subesophageal zone and thereby suppress courtship (Shankar et al., 2015). It is not clear from this study to which specific neurons TK in **Figure 8** they correspond.

The insulin-producing cells (IPCs) of the *Drosophila* brain are modulated by several factors, including TK (Birse et al., 2011; Nässel and Vanden Broeck, 2016). The IPCs produce four insulin-like peptides (DILP1, 2, 3, and 5) and are known to regulate many aspects of development and adult physiology, such as growth, metabolism, stress responses, reproduction and lifespan reviewed in Owusu-Ansah and Perrimon (2014), Nässel and Vanden Broeck (2016). Knockdown of the receptor DTKR in IPCs affected levels of *dilp2* and *dilp3* transcripts in these cells, increased the fly lifespan and diminished carbohydrate levels during starvation (Birse et al., 2011). Knockdown of the natalisin receptor (NTLR; CG6115; earlier known as NKD) had no effect on IPC activity and the TK cells acting on the IPCs were not identified (Birse et al., 2011). In a more recent paper, a pair of TK neurons was demonstrated in the *Drosophila* larva, which connect functionally to the IPCs (Meschi et al., 2019). These TK neurons (ICNs) are inhibitory on IPCs. Under protein-rich diet conditions the ICNs respond to growth-blocking peptides secreted from the larval fat body and this alleviates the inhibitory action on IPCs, and DILPs can be released to stimulate growth (Meschi et al., 2019). It is not completely clear from the images of this paper, but it appears as if the ICNs are the same as the descending neurons shown in **Figure 8C** (blue arrows), which also exist in the adults (DN in **Figure 8A**).

In larvae, a nociceptive pathway mediating thermal tissue damage signals was identified and shown to include TK and the receptor DTKR (Im et al., 2015). The DTKR receptor is expressed in the nociceptive sensory neurons and required for mediation of thermal hypersensitivity after tissue damage. A set of TK expressing interneurons in the ventral nerve cord mediates this presynaptic modulation of nociceptive sensory neurons (Im et al., 2015). Substance P is known for its role in modulation of nociceptive sensory signals in the dorsal horn of the spinal cord [see (Hökfelt et al., 2001; Steinhoff et al., 2014)], suggesting a

conserved role of tachykinin signaling, although the pathway and mechanisms differ.

A set of five pairs of large neurosecretory cells (ITPn in **Figure 8A**) produces TK, as well as ITP and sNPF. Targeted knockdown of TK (or sNPF) in these cells result in flies that display decreased survival time when exposed to desiccation or starvation, and also suffer increased water loss at desiccation (Kahsai et al., 2010a). ITP is acting as an antidiuretic hormone (Galikova et al., 2018), but it is not likely that TK or sNPF are released as circulating hormones from the ITPn cells (Kahsai et al., 2010a). Instead, these peptides might act locally, either presynaptically on ITPn axon terminations, or on other brain neurons/neurosecretory cells to modulate antidiuretic signals or metabolic stress responses. A similar local action of sNPF released from lateral neurosecretory cells in the brain has been demonstrated; it was found that the IPCs in the brain and the AKH-producing cells in the CC were directly regulated by locally released sNPF (Kapan et al., 2012; Oh et al., 2019).

In gut endocrine cells (EECs) of *Drosophila*, TK was shown to influence lipid homeostasis by controlling lipid production in enterocytes of the midgut (Song et al., 2014). These TK- (and DH31-) producing EECs are nutrient-sensing and can be activated by the presence of circulating dietary proteins and amino acids (Park et al., 2016). The EECs have also been shown to play a role in the innate immune system and development of *Drosophila* (Kamareddine et al., 2018). Activating the immune deficiency (IMD) pathway in EECs triggers TK signaling leading to DILP3 upregulation in the gut and mobilization of lipids increased insulin signaling and effects on organismal development and growth. Thus, the gut microbiota can influence growth via the immune system and TK and insulin signaling (Kamareddine et al., 2018). TK was also shown to activate peristalsis in the midgut (Siviter et al., 2000), maybe by local paracrine signaling.

Functional Roles of TKs in Protochordates and Non-mammalian Vertebrates

Also in several non-insect invertebrates and non-mammalian vertebrates TKs were found to exhibit contractile activity on muscles in the digestive tract (Satake and Kawada, 2006; Satake et al., 2013; Steinhoff et al., 2014). In the following we will discuss additional roles of TKs in sea squirts (Ascidians) and fish, exemplified by Zebrafish.

Ascidians

Aoyama et al. (2012) demonstrated that CiTK induces growth of follicles in *Ciona* during late stage-II (vitellogenic stage) to stage-III (post-vitellogenic stage) via up-regulation of gene expression and the enzymatic activities of follicle-processing proteases: cathepsin D, chymotrypsin, and carboxypeptidase B1. This is consistent with the finding that CiTKR is expressed exclusively in test cells (functional counterparts of vertebrate granulosa cells) residing in late stage-II follicles (Aoyama et al., 2012). Moreover, Ci Cathepsin D, co-localized with CiTKR in test cells, is initially activated, and Ci Carboxypeptidase B1 and

Ci Chymotrypsin, localized in follicular cells, are activated 1 h later (Aoyama et al., 2012). These findings provide evidence for a novel tachykininergic follicle growth pathway. In addition, the CiTK-induced follicle growth is suppressed by a *Ciona*-specific neuropeptide, CiNTP6, via downregulation of the three aforementioned proteases (Kawada et al., 2011). It would be interesting to reveal whether the tachykininergic regulation of follicle growth is conserved, at least in part, in vertebrates or other invertebrates.

Teleost Fish

The roles of NKB in reproductive functions are of interest in both teleosts and mammals. NKB is expressed in the hypothalamus of mammals (Satake and Kawada, 2006; Steinhoff et al., 2014). Moreover, NKB is colocalized with kisspeptin and dynorphin A in KNDy neurons in the arcuate nuclei. KNDy neurons and NKB are responsible for the generation of GnRH pulsatility in the hypothalamus, which plays a central role in reproductive functions via induction of secretion of gonadotropins (LH and FSH) from the pituitary to the gonads (Lehman et al., 2010; Navarro, 2012). Interestingly, some mutations were detected in the genomic sequences of human *Tac3* and *Tacr3* in a portion of patients with hypogonadotropic hypogonadism (Topaloglu et al., 2009; Chen et al., 2018). NKB is likely to downregulate production of LH and FSH in zebrafish, tilapia and goldfish (Biran et al., 2012; Qi et al., 2015; Hu et al., 2017; Chen et al., 2018; Liu et al., 2019). Several of these studies also indicated that NKB and/or NKF (identical to NKB-related peptide) downregulate the expression of *kiss2* that is a homolog of mammalian kisspeptin. The role of mammalian kisspeptin in induction of GnRH synthesis and release may suggest that kisspeptin 2 also upregulates GnRH synthesis and release in teleost. However, conservation of such kisspeptin 2-directed GnRH regulation in teleost is not likely, since kisspeptin 2 seems not to be involved in reproduction in teleosts (Nakajo et al., 2017). Collectively, these findings suggest that biological roles of NKB and NKF in reproduction in teleost are distinct from those in mammals. In addition, SP and NKA were found to upregulate gene expression and release of LH, prolactin, and somatolactin α in carp pituitary cells (Hu et al., 2017). Interestingly, short-term SP treatment (3 h) induces LH release, but long-term SP treatment attenuated gene LH expression (Hu et al., 2017). Thus, in teleosts SP and NKA are important in reproductive functions.

CONSERVED ROLES OF TACHYKININ SIGNALING IN THE ANIMAL KINGDOM

Some of the functional roles of TKs that have been described in some detail in earlier sections are evolutionarily conserved, at least in general terms. By general terms we mean that for instance a role in nociception has been found for TKs both in *Drosophila* (Im et al., 2015) and in mammals [see (Onaga, 2014; Steinhoff et al., 2014; Ziegglänsberger, 2019)], but the neuronal pathways and mechanisms are quite different. In a similar fashion, TKs are acting as cotransmitters in many neuronal circuits and thus play roles in for instance: modulation of

olfactory sensory signaling together with GABA in *Drosophila* (Ignell et al., 2009; Ko et al., 2015) and mammals (Olpe et al., 1987), modulation of rhythm generating motor networks in crustaceans (Nusbaum et al., 2017) and lampreys (Parker and Grillner, 1998), aggression in *Drosophila* (Asahina et al., 2014) and mammals (Felipe et al., 1998; Katsouni et al., 2009), as well as roles in learning and memory circuits in honey bees (Takeuchi et al., 2004; Brockmann et al., 2009; Boerjan et al., 2010) and mammals (Lénárd et al., 2018). Furthermore, TKs are involved in regulation of several aspects of intestinal function, including electrolyte and fluid secretion in insects (Johard et al., 2003; Veenstra et al., 2008; Lemaitre and Miguel-Aliaga, 2013; Song et al., 2014) and mammals (Hökfelt et al., 2001; Steinhoff et al., 2014), in regulation of gustatory receptors in *Drosophila* (Shankar et al., 2015) and mammals (Onaga, 2014) and in control of hormone release in insects (Nässel et al., 1995b; Birse et al., 2011; Meschi et al., 2019) and vertebrates (Hu et al., 2014; Steinhoff et al., 2014; Zhang et al., 2019).

Roles of TKs in reproductive functions have been demonstrated in vertebrates (Satake et al., 2013; Steinhoff et al., 2014), but not yet in insects or other invertebrates. However, in insects and other arthropods natalisins seem to be important in reproductive behavior, as outlined in the next section (Jiang et al., 2013; Gui et al., 2018).

NATALISINS, A SISTER GROUP OF TACHYKININS IN ARTHROPODS

A novel peptide precursor gene that encodes multiple copies of peptides that were designated natalisins (NTLs) was discovered in *Drosophila*, *Tribolium castaneum*, and *Bombyx mori*; these have a consensus sequence FXXXXamide (Jiang et al., 2013). The name NTL is derived from the functional role of the peptide in reproduction (Latin word *natalis* for birth) (Jiang et al., 2013). The NTLs have so far only been identified in arthropods and tardigrades, and the peptides display minor similarities to TKs. However, the NTL receptor (NTLR) was previously identified as a TK receptor (CG6115; TakR86C; NKD) (Monnier et al., 1992; Poels et al., 2009), suggesting that NTLs are ancestrally related to arthropod TKs (Jiang et al., 2013). In fact, phylogenetic analysis suggests that NTL signaling arose through duplication of the TK signaling early in the arthropod lineage (Jiang et al., 2013; see **Figure 6**). However, TK-like precursors with NTL-like peptides are found in the spider mite (chelicerate) as well as tardigrades as shown in **Figures 4, 5**, suggesting that the NTL signaling might also be present outside arthropods (Veenstra et al., 2012; Koziol, 2018). It was noted that in the centipede *Strigamia maritima* (Myriapoda) there is no NTL gene (Veenstra, 2016) and in the spider mite *Tetranychus urticae*, there is no separate NTL gene in the genome (Veenstra et al., 2012). However, two TK genes were annotated in *T. urticae* and on these precursors two of the three putative mature peptides are similar to NTL, and one is a TK; the second precursor encodes two TKs and an unrelated peptide, but no NTL (**Figure 4**). Thus, with this mix of TKs and NTLs on the spider mite genes, it was suggested that TK and NTL divergence started by internal events on duplicated TK

genes and resulted in the evolution of a separate NTL signaling system (Jiang et al., 2013). Additional genomes/transcriptomes of basal arthropods need to be examined to substantiate this claim. As seen in **Supplementary Table S5**, there are five paracopies of natalisins in *Drosophila*, 7 in *Anopheles aegypti*, 11 in *Bombyx mori* and 15 in *Manduca sexta*, 2 in *Tibolium castaneum* and only one each in the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus* (Jiang et al., 2013; Koziol, 2018). The *Drosophila* peptides (DmNTL1-5) range from 15 to 24 residues and have a consensus C-terminus FXPXRamide (except DmNTL4).

In the brains of *Drosophila*, *B. mori*, *T. castaneum* and *Varroa destructor*, there are two pairs of identifiable NTL neurons with very similar locations and arborizations (designated ADLI and ICLI in each species) (Jiang et al., 2013, 2016). The *Drosophila* neurons are shown in **Supplementary Figure S6**. In the *B. mori* brain, there are two additional pairs of neurons in the subesophageal zone. Another study shows that in the oriental fruit fly *Bactrocera dorsalis* there are three pairs of NTL neurons (Gui et al., 2017a). Thus, in insects studied so far, the NTL system appears relatively simple and the neuronal branches do not seem to innervate any of the well-defined centers, such as antennal lobes, mushroom bodies, central complex or optic lobes (Jiang et al., 2013). There are a few additional segmental neurons in the ventral nerve cord. The brain ICLI neurons coexpress NTL, Ast-A and MIP (Diesner et al., 2018).

In *Drosophila*, genetic experiments revealed that NTL and the four NTL neurons are important for male mating success (Jiang et al., 2013). NTL-RNAi in NTL-Gal4 neurons reduces male copulation success rate. The courtship behavior was only affected in the latency of courtship initiation in males. NTL-RNAi females also displayed reduced mating frequency, but did not actively reject males (Jiang et al., 2013). Silencing of the NTL-Gal4 neurons resulted in complete repression of mating in males, but had no effect in females. Manipulations of NTL neurons had no effect on egg laying, however, in *T. castaneum* systemic NTL-RNAi in either sex resulted in reduced egg numbers after mating (Jiang et al., 2013). Also in the oriental fruit fly *Bactrocera dorsalis* NTL and its receptor (NTLR) play important roles in mating (Gui et al., 2017a, 2018). In this species the NTL signaling is required for regulation of mating frequency in both males and females.

CONCLUSION AND PERSPECTIVES

We have shown that TKs are neuropeptides that emerged early in bilaterian lineages, but it is not clear what their ancestral form is since cnidarians and other non-bilaterian do not possess typical TKs, although bioinformatics has indicated presence of TK receptors [see (Krishnan and Schiöth, 2015; Hayakawa et al., 2019)]. It is also puzzling that no typical TKs have been identified in echinoderms, acorn worms, or amphioxus. Possibly this is due to species-specific diversification of TK sequences. Thus, important questions regarding the evolution and diversification of this signaling system remain unanswered. Nevertheless, it is clear that TK signaling is widespread and diverse among bilaterians and contributes to many vital functions. Some of these functions appear conserved over evolution, at least in

general terms. It is important to stress that TKs seem to have multiple distributed (localized) functions in different neuronal circuits and commonly act as co-transmitters, and thus TK signaling is not likely to orchestrate global functions. Elucidation of TK functions in neglected phyla (such as echinoderms, xenacoelomorphs and cnidarians) can also provide clues on whether the mode of action of TK as a co-transmitter is an ancient or a more derived trait. This is important since besides TK, there are only a few other neuropeptides in protostomes (at least in arthropods) that seem to function mainly as co-transmitters, such as sNPF and proctolin (Nässel, 2018; Nässel and Zandawala, 2019). Did TKs evolve as primary paracrine peptide signals in basal phyla with simple nervous systems, and then diversified functionally to also confer plasticity to more complex neural circuits by providing neuromodulatory actions as co-transmitters? In organisms without nervous systems, such a *Trichoplax adherens*, peptides seem to act as primary messengers that induce simple behaviors (Nikitin, 2015; Senatore et al., 2017; Varoqueaux et al., 2018) and even in more evolved organisms many neuropeptides/peptide hormones seem to relay single global orchestrating actions [see (Nässel and Zandawala, 2019; Nässel et al., 2019)].

In mammals, TK signaling has received extensive attention due to its clinical importance with roles for instance in pain, inflammation, cancer, depressive disorder and immune system. Thus, the literature list is huge: searching PubMed for the term “substance P” renders more than 24,000 hits. This means that our coverage of mammalian TKs in this review is very superficial and incomplete. On the other hand, a search for e.g., “tachykinin in insects” yields about 200 hits and therefore our discussion of invertebrate TKs is somewhat more detailed, but certainly still providing a sketchy picture of TK signaling since functional studies are not yet that numerous.

The development of powerful genetic tools, not only in *Drosophila* and *C. elegans*, but also other organisms has improved the possibilities to analyze neuropeptide signaling down to single identified neurons or sets of neurons. Furthermore, with optogenetics and other strategies for temporal control of manipulations and elegant techniques for imaging neuronal connections or activity, we already see an increase in studies of invertebrate neuropeptides. For TKs, it is of importance to note that in the CNS these peptides seem to operate as local neuromodulators and/or co-transmitters. Many (if not most) TK expressing neurons may additionally signal with small molecule transmitters and, therefore, manipulations of TK signaling only remove one layer of the signal transfer.

Whereas many neuropeptides also have functions as circulating hormones [see (Nässel and Zandawala, 2019)], it seems like TKs do not in most organisms studied. In studies of cockroach, locust and *Drosophila* it was proposed that TKs are released into the circulation from gut endocrine cells to stimulate secretion in nearby Malpighian tubules (Winther and Nässel, 2001; Johard et al., 2003; Söderberg et al., 2011). A recent *Drosophila* study showed that gut TKs act locally and do not affect behavior, indicating that there is no signaling to the brain via the circulation (Song et al., 2014). If *bona fide* hormonal roles of TKs can be excluded, we can focus on their local actions,

but we still face some difficulties due to the diversity of TK expressing neuronal systems and the co-expression of small molecule transmitters. Hopefully this review will trigger interest in TK signaling in invertebrates in spite of these challenges. It is obvious from the literature that research on TK signaling in mammals is already very extensive, but certainly further basic research and clinical studies are urgently needed to unravel this important and interesting signaling system.

AUTHOR CONTRIBUTIONS

DN contributed to the conceptualization, prepared the first draft of the manuscript, wrote parts of the manuscript, prepared figures and tables, and coordinated the assembly of the manuscript. MZ contributed to the conceptualization, wrote parts of the manuscript, and prepared figures and tables. TK and HS wrote

parts of the manuscript, and prepared figures and tables. All authors edited and finally approved the manuscript.

FUNDING

This work was funded by the Swedish Research Council (Vetenskapsrådet), grant number 2015-04626 (DN) and the Japan Society for the Promotion of Science, grant number JP19K06752 (HS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2019.01262/full#supplementary-material>

REFERENCES

- Al-Anzi, B., Armand, E., Nagamei, P., Olszewski, M., Sapin, V., Waters, C., et al. (2010). The leucokinin pathway and its neurons regulate meal size in *Drosophila*. *Curr. Biol.* 20, 969–978. doi: 10.1016/j.cub.2010.04.039
- Almeida, T. A., Rojo, J., Nieto, P. M., Pinto, F. M., Hernandez, M., Martin, J. D., et al. (2004). Tachykinins and Tachykinin receptors: structure and activity relationships. *Curr. Med. Chem.* 11, 2045–2081. doi: 10.2174/0929867043364748
- Anastasi, A., Montecucchi, P., Erspamer, V., and Visser, J. (1977). Amino acid composition and sequence of kassinin, a tachykinin dodecapeptide from the skin of the African frog *Kassina senegalensis*. *Experientia* 33, 857–858. doi: 10.1007/bf01951242
- Anctil, M. (2009). Chemical transmission in the sea anemone *Nematostella vectensis*: a genomic perspective. *Comp. Biochem. Physiol. D* 4, 268–289. doi: 10.1016/j.cbd.2009.07.001
- Aoyama, M., Kawada, T., and Satake, H. (2012). Localization and enzymatic activity profiles of the proteases responsible for tachykinin-directed oocyte growth in the protochordate. *Ciona intestinalis Peptides* 34, 186–192. doi: 10.1016/j.peptides.2011.07.019
- Arvidson, R., Kaiser, M., Lee, S. S., Urenda, J.-P., Dail, C., Mohammed, H., et al. (2019). Parasitoid jewel wasp mounts multipronged neurochemical attack to hijack a host brain. *Mol. Cell. Proteomics* 18, 99–114. doi: 10.1074/mcp.RA118.000908
- Asahina, K., Watanabe, K., Duistermars, B. J., Hoopfer, E., Gonzalez, C. R., Eyjolfsson, E. A., et al. (2014). Tachykinin-expressing neurons control male-specific aggressive arousal in *Drosophila*. *Cell* 156, 221–235. doi: 10.1016/j.cell.2013.11.045
- Azimi, E., Reddy, V. B., Pereira, P. J. S., Talbot, S., Woolf, C. J., and Lerner, E. A. (2017). Substance P activates Mas-related G protein-coupled receptors to induce itch. *J. Allergy Clin. Immunology* 140, 447.e–453.e. doi: 10.1016/j.jaci.2016.12.980
- Azimi, E., Reddy, V. B., Shade, K.-T. C., Anthony, R. M., Talbot, S., Pereira, P. J. S., et al. (2016). Dual action of neurokinin-1 antagonists on mas-related GPCRs. *JCI Insight* 1:e89362. doi: 10.1172/jci.insight.89362
- Bader, M., Alenina, N., Andrade-Navarro, M. A., and Santos, R. A. (2014). Mas and its related G protein-coupled receptors. *Mrgprs. Pharmacol. Rev.* 66:1080. doi: 10.1124/pr.113.008136
- Beerntsen, B. T., Champagne, D. E., Coleman, J. L., Campos, Y. A., and James, A. A. (1999). Characterization of the sialokinin I gene encoding the salivary vasodilator of the yellow fever mosquito. *Aedes aegypti. Insect. Mol. Biol.* 8, 459–467. doi: 10.1046/j.1365-2583.1999.00141.x
- Bertaccini, G., Cej, J. M., and Erspamer, V. (1965). Occurrence of physalaemin in extracts of the skin of *Physalaemus fuscumaculatus* and its pharmacological actions on extravascular smooth muscle. *Br. J. Pharmacol. Chemother.* 25, 363–379. doi: 10.1111/j.1476-5381.1965.tb02056.x
- Binzer, M., Heuer, C. M., Kollmann, M., Kahnt, J., Hauser, F., Grimmelikhuijzen, C. J. P., et al. (2014). Neuropeptidome of *Tribolium castaneum* antennal lobes and mushroom bodies. *J. Comp. Neurol.* 522, 337–357. doi: 10.1002/cne.23399
- Biran, J., Palevitch, O., Ben-Dor, S., and Levavi-Sivan, B. (2012). Neurokinin Bs and neurokinin B receptors in zebrafish-potential role in controlling fish reproduction. *Proc. Natl. Acad. Sci.* 109:10269. doi: 10.1073/pnas.1119165109
- Birse, R. T., Johnson, E. C., Taghert, P. H., and Nässel, D. R. (2006). Widely distributed *Drosophila* G-protein-coupled receptor (CG7887) is activated by endogenous tachykinin-related peptides. *J. Neurobiol.* 66, 33–46. doi: 10.1002/neu.20189
- Birse, R. T., Söderberg, J. A., Luo, J., Winther, ÅM., and Nässel, D. R. (2011). Regulation of insulin-producing cells in the adult *Drosophila* brain via the tachykinin peptide receptor DTKR. *J. Exp. Biol.* 214, 4201–4208. doi: 10.1242/jeb.062091
- Blasco, V., Pinto, F. M., Fernández-Atucha, A., Prados, N., Tena-Sempere, M., Fernández-Sánchez, M., et al. (2019). Altered expression of the kisspeptin/KISS1R and neurokinin B/NK3R systems in mural granulosa and cumulus cells of patients with polycystic ovarian syndrome. *J. Assis. Reproduct. Genet.* 36, 113–120. doi: 10.1007/s10815-018-1338-7
- Blitz, D. M., Christie, A. E., Marder, E., and Nusbaum, M. P. (1995). Distribution and effects of tachykinin-like peptides in the stomatogastric nervous system of the crab. *Cancer borealis. J. Comp. Neurol.* 354, 282–294. doi: 10.1002/cne.903540209
- Boerjan, B., Cardoen, D., Bogaerts, A., Landuyt, B., Schoofs, L., and Verleyen, P. (2010). Mass spectrometric profiling of (neuro)-peptides in the worker honeybee. *Apis mellifera. Neuropharmacology* 58, 248–258. doi: 10.1016/j.neuropharm.2009.06.026
- Brand, A. H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415.
- Brockmann, A., Annangudi, S. P., Richmond, T. A., Ament, S. A., Xie, F., Southey, B. R., et al. (2009). Quantitative peptidomics reveal brain peptide signatures of behavior. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2383–2388. doi: 10.1073/pnas.0813021106
- Campo, A., Lafont, A.-G., Lefranc, B., Leprince, J., Tostivint, H., Kamech, N., et al. (2018). Tachykinin-3 genes and peptides characterized in a basal teleost, the european eel: evolutionary perspective and pituitary role. *Front. Endocrinol.* 9:304. doi: 10.3389/fendo.2018.00304
- Candenas, L., Pinto, F. M., Cejudo-Román, A., González-Ravina, C., Fernández-Sánchez, M., Pérez-Hernández, N., et al. (2018). Veratridine-sensitive Na⁺ channels regulate human sperm fertilization capacity. *Life Sci.* 196, 48–55. doi: 10.1016/j.lfs.2018.01.004
- Carlsson, M. A., Diesner, M., Schachtner, J., and Nässel, D. R. (2010). Multiple neuropeptides in the *Drosophila* antennal lobe suggest complex modulatory circuits. *J. Comp. Neurol.* 518, 3359–3380. doi: 10.1002/cne.22405

- Chamberlain, S. C., and Engbretson, G. A. (1982). Neuropeptide immunoreactivity in *Limulus*. I. Substance P-like immunoreactivity in the lateral eye and protocerebrum. *J. Comp. Neurol.* 208, 304–315. doi: 10.1002/cne.902080307
- Champagne, D. E., and Ribeiro, J. M. (1994). Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito *Aedes aegypti*. *Proc. Natl. Acad. Sci.* 91, 138. doi: 10.1073/pnas.91.1.138
- Chang, M. M., Leeman, S. E., and Niall, H. D. (1971). Amino-acid sequence of substance P. *Nat. New Biol.* 232, 86–87. doi: 10.1038/newbio232086a0
- Chen, H., Xiao, L., Liu, Y., Li, S., Li, G., Zhang, Y., et al. (2018). Neurokinin B signaling in hermaphroditic species, a study of the orange-spotted grouper (*Epinephelus coioides*). *Gen. Comp. Endocrinol.* 260, 125–135. doi: 10.1016/j.ygcen.2018.01.009
- Christie, A. E. (2016). Expansion of the neuropeptidome of the globally invasive marine crab *Carcinus maenas*. *Gen. Comp. Endocrinol.* 235, 150–169. doi: 10.1016/j.ygcen.2016.05.013
- Christie, A. E., Kutz-Naber, K. K., Stemmler, E. A., Klein, A., Messinger, D. I., Goiney, C. C., et al. (2007). Midgut epithelial endocrine cells are a rich source of the neuropeptides APSGFLGMRamide (*Cancer borealis* tachykinin-related peptide Ia) and GYRKPPFNGSIFamide (Gly-SIFamide) in the crabs *Cancer borealis*, *Cancer magister* and *Cancer productus*. *J. Exp. Biol.* 210:699. doi: 10.1242/jeb.02696
- Christie, A. E., Lundquist, C. T., Nässel, D. R., and Nusbaum, M. P. (1997). Two novel tachykinin-related peptides from the nervous system of the crab *Cancer borealis*. *J. Exp. Biol.* 200, 2279–2294.
- Conzelmann, M., Williams, E. A., Krug, K., Franz-Wachtel, M., Macek, B., and Jékely, G. (2013). The neuropeptide complement of the marine annelid *Platynereis dumerilii*. *BMC Genomics* 14:906. doi: 10.1186/1471-2164-14-906
- Cuello, A. C., Galfre, G., and Milstein, C. (1979). Detection of substance P in the central nervous system by a monoclonal antibody. *Proc. Natl. Acad. Sci. U.S.A.* 76, 3532–3536. doi: 10.1073/pnas.76.7.3532
- Cunningham, C. B., Badgett, M. J., Meagher, R. B., Orlando, R., and Moore, A. J. (2017). Ethological principles predict the neuropeptides co-opted to influence parenting. *Nat. Commun.* 8:14225. doi: 10.1038/ncomms14225
- Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Krefl, L., et al. (2018). A single-cell transcriptome atlas of the aging *Drosophila* brain. *Cell* 174, 982–998. doi: 10.1016/j.cell.2018.05.057
- Diesner, M., Predel, R., and Neupert, S. (2018). Neuropeptide mapping of dimmed cells of adult *Drosophila* brain. *J. Am. Soc. Mass. Spectrom.* 29, 890–902. doi: 10.1007/s13361-017-1870-1
- Du Vigneaud, V. (1955). Oxytocin, the principal oxytocic hormone of the posterior pituitary gland: its isolation, structure, and synthesis. *Experientia* (Suppl. 2), 9–26.
- Du Vigneaud, V., Ressler, C., and Trippett, S. (1953). The sequence of amino acids in oxytocin, with a proposal for the structure of oxytocin. *J. Biol. Chem.* 205, 949–957.
- Dubos, M.-P., Zels, S., Schwartz, J., Pasquier, J., Schoofs, L., and Favrel, P. (2018). Characterization of a tachykinin signalling system in the bivalve mollusc *Crassostrea gigas*. *Gen. Comp. Endocrinol.* 266, 110–118. doi: 10.1016/j.ygcen.2018.05.003
- Elekes, K., Hernadi, L., Kiss, T., Muneoka, Y., and Nässel, D. R. (1995). Tachykinin- and leucokinin-related peptides in the molluscan nervous system. *Acta Biol. Hung.* 46, 281–294.
- Elekes, K., and Nässel, D. R. (1994). Tachykinin-related neuropeptides in the central nervous system of the snail *Helix pomatia*: an immunocytochemical study. *Brain Res.* 661, 223–236. doi: 10.1016/0006-8993(94)91199-1
- Elphick, M. R., Mirabeau, O., and Larhammar, D. (2018). Evolution of neuropeptide signalling systems. *J. Exp. Biol.* 221:jeb151092. doi: 10.1242/jeb.151092
- Erspamer, V., and Anastasi, A. (1962). Structure and pharmacological actions of eledoisin, active endecapeptide of posterior salivary glands of eledone. *Experientia* 18, 58–59. doi: 10.1007/bf02138250
- Erspamer, V., and Erspamer, G. F. (1962). Pharmacological actions of eledoisin on extravascular smooth muscle. *Br. J. Pharmacol. Chemother.* 19, 337–337.
- Felipe, C. D., Herrero, J. F., O'Brien, J. A., Palmer, J. A., Doyle, C. A., Smith, A. J. H., et al. (1998). Altered nociception, analgesia and aggression in mice lacking the receptor for substance P. *Nature* 392, 394–397. doi: 10.1038/32904
- Fusca, D., Schachtner, J., and Kloppenburg, P. (2015). Colocalization of allatotropin and tachykinin-related peptides with classical transmitters in physiologically distinct subtypes of olfactory local interneurons in the cockroach (*Periplaneta americana*). *J. Comp. Neurol.* 523, 1569–1586. doi: 10.1002/cne.23757
- Galikova, M., Dircksen, H., and Nässel, D. R. (2018). The thirsty fly: Ion transport peptide (ITP) is a novel endocrine regulator of water homeostasis in *Drosophila*. *PLoS Genet.* 14:e1007618. doi: 10.1371/journal.pgen.1007618
- García-Ortega, J., Pinto, F. M., Fernández-Sánchez, M., Prados, N., Cejudo-Román, A., Almeida, T. A., et al. (2014). Expression of neurokinin B/NK3 receptor and kisspeptin/KISS1 receptor in human granulosa cells. *Hum. Reprod.* 29, 2736–2746. doi: 10.1093/humrep/deu247
- Glantz, R. M., Miller, C. S., and Nässel, D. R. (2000). Tachykinin-related peptide and GABA-mediated presynaptic inhibition of crayfish photoreceptors. *J. Neurosci.* 20, 1780–1790. doi: 10.1523/jneurosci.20-05-01780.2000
- Goda, T., Tang, X., Umezaki, Y., Chu, M. L., and Hamada, F. N. (2016). *Drosophila* DH31 neuropeptide and PDF receptor regulate night-onset temperature preference. *J. Neurosci.* 36, 11739–11754. doi: 10.1523/jneurosci.0964-16.2016
- Goldberg, D., Nusbaum, M. P., and Marder, E. (1988). Substance P-like immunoreactivity in the stomatogastric nervous systems of the crab *Cancer borealis* and the lobsters *Panulirus interruptus* and *Homarus americanus*. *Cell Tissue Res.* 252, 515–522. doi: 10.1007/bf00216638
- Gui, S. H., Jiang, H. B., Liu, X. Q., Xu, L., and Wang, J. J. (2017a). Molecular characterizations of natalisin and its roles in modulating mating in the oriental fruit fly. *Bactrocera dorsalis* (Hendel). *Insect Mol. Biol.* 26, 103–112. doi: 10.1111/imb.12274
- Gui, S.-H., Jiang, H.-B., Xu, L., Pei, Y.-X., Liu, X.-Q., Smagghe, G., et al. (2017b). Role of a tachykinin-related peptide and its receptor in modulating the olfactory sensitivity in the oriental fruit fly. *Bactrocera dorsalis* (Hendel). *Insect Biochem. Mol. Biol.* 80, 71–78. doi: 10.1016/j.ibmb.2016.12.002
- Gui, S.-H., Pei, Y.-X., Xu, L., Wang, W.-P., Jiang, H.-B., Nachman, R. J., et al. (2018). Function of the natalisin receptor in mating of the oriental fruit fly. *Bactrocera dorsalis* (Hendel) and testing of peptidomimetics. *PLoS One* 13:e0193058. doi: 10.1371/journal.pone.0193058
- Haddad, A. N. S., Defferrari, M. S., Hana, S., Szeto, S. G., and Lange, A. B. (2018). Expression and functional characterization of tachykinin-related peptides in the blood-feeding bug. *Rhodnius prolixus*. *Peptides* 99, 247–254. doi: 10.1016/j.peptides.2017.11.006
- Hayakawa, E., Watanabe, H., Menschaert, G., Holstein, T. W., Baggerman, G., and Schoofs, L. (2019). A combined strategy of neuropeptide predictions and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*. *bioRxiv* 593384[Preprint].
- He, X., Zang, J., Li, X., Shao, J., Yang, H., Yang, J., et al. (2014). Activation of BNGR-A24 by direct interaction with tachykinin-related peptides from the silkworm *Bombyx mori* leads to the Gq- and Gs-coupled signaling cascades. *Biochemistry* 53, 6667–6678. doi: 10.1021/bi5007207
- Heuer, C. M., Kollmann, M., Binzer, M., and Schachtner, J. (2012). Neuropeptides in insect mushroom bodies. *Arthropod Struct. Dev.* 41, 199–226. doi: 10.1016/j.asd.2012.02.005
- Hökfelt, T., Kellerth, J. O., Nilsson, G., and Pernow, B. (1975). Substance P localization in the central nervous system and in some primary sensory neurons. *Science* 190, 889–890. doi: 10.1126/science.242075
- Hökfelt, T., Ljungdahl, A., Terenius, L., Elde, R., and Nilsson, G. (1977). Immunohistochemical analysis of peptide pathways possibly related to pain and analgesia: enkephalin and substance P. *Proc. Natl. Acad. Sci. U.S.A.* 74, 3081–3085. doi: 10.1073/pnas.74.7.3081
- Hökfelt, T., Pernow, B., and Wahren, J. (2001). Substance P: a pioneer amongst neuropeptides. *J. Intern. Med.* 249, 27–40. doi: 10.1046/j.0954-6820.2000.00773.x
- Holman, G. M., Cook, B. J., and Nachman, R. J. (1986). Isolation, primary structure and synthesis of two neuropeptides from *Leucophaea maderae*: members of a new family of cephalotropins. *Comp. Biochem. Physiol.* 84C, 205–211. doi: 10.1016/0742-8413(86)90084-8
- Holman, G. M., Nachman, R. J., Schoofs, L., Hayes, T. K., Wright, M. S., and Deloof, A. (1991). The *Leucophaea maderae* hindgut preparation: a rapid and sensitive bioassay tool for the isolation of insect myotropins of other insect species. *Insect Biochem.* 21, 107–112. doi: 10.1016/0020-1790(91)90070-u
- Holman, G. M., Nachman, R. J., and Wright, M. S. (1990). Insect neuropeptides. *Annu. Rev. Entomol.* 35, 201–217. doi: 10.1146/annurev.ento.35.1.201

- Holsboer, F. (2009). "Neuropeptides: mental disease," in *Encyclopedia of Neuroscience*, ed. L. R. Squire, (Oxford: Academic Press), 923–930. doi: 10.1016/b978-008045046-9.01461-3
- Hu, G., He, M., Ko, W. K. W., Lin, C., and wong, a. o. l. (2014). novel pituitary actions of tac3 gene products in fish model: receptor specificity and signal transduction for prolactin and somatolactin α regulation by neurokinin B (NKB) and NKB-related peptide in carp pituitary cells. *Endocrinology* 155, 3582–3596. doi: 10.1210/en.2014-1105
- Hu, G., He, M., Ko, W. K. W., and Wong, A. O. L. (2017). TAC1 gene products regulate pituitary hormone secretion and gene expression in prepubertal grass carp pituitary cells. *Endocrinology* 158, 1776–1797. doi: 10.1210/en.2016-1740
- Ignell, R., Root, C. M., Birse, R. T., Wang, J. W., Nässel, D. R., and Winther, ÅM. (2009). Presynaptic peptidergic modulation of olfactory receptor neurons in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13070–13075. doi: 10.1073/pnas.0813004106
- Ikeda, T., Minakata, H., Nomoto, K., Kubota, I., and Muneoka, Y. (1993). Two novel tachykinin-related neuropeptides in the echiuroid worm. *Urechis unicinctus*. *Biochem. Biophys. Res. Commun.* 192, 1–6. doi: 10.1006/bbr.1993.1373
- Im, S. H., Takle, K. J., Babcock, D. T., Ma, Z., Xiang, Y., et al. (2015). Tachykinin acts upstream of autocrine Hedgehog signaling during nociceptive sensitization in *Drosophila*. *eLife* 4:e10735. doi: 10.7554/eLife.10735
- Isaac, R. E., Bland, N. D., and Shirras, A. D. (2009). Neuropeptidases and the metabolic inactivation of insect neuropeptides. *Gen. Comp. Endocrinol.* 162, 8–17. doi: 10.1016/j.ygcen.2008.12.011
- Isaac, R. E., Parkin, E. T., Keen, J. N., Nässel, D. R., Siviter, R. J., and Shirras, A. D. (2002). Inactivation of a tachykinin-related peptide: identification of four neuropeptide-degrading enzymes in neuronal membranes of insects from four different orders. *Peptides* 23, 725–733. doi: 10.1016/s0196-9781(01)00653-2
- Jekely, G. (2013). Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8702–8707. doi: 10.1073/pnas.1221833110
- Jekely, G., Melzer, S., Beets, I., Kadow, I. C. G., Koene, J., Haddad, S., et al. (2018). The long and the short of it – a perspective on peptidergic regulation of circuits and behaviour. *J. Exp. Biol.* 221:jeb166710. doi: 10.1242/jeb.166710
- Jiang, H., Kim, D., Dobesh, S., Evans, J. D., Nachman, R. J., Kaczmarek, K., et al. (2016). Ligand selectivity in tachykinin and natalisin neuropeptidergic systems of the honey bee parasitic mite *Varroa destructor*. *Sci. Rep.* 6:19547. doi: 10.1038/srep19547
- Jiang, H., Lkhagva, A., Daubnerova, I., Chae, H. S., Simo, L., Jung, S. H., et al. (2013). Natalisin, a tachykinin-like signaling system, regulates sexual activity and fecundity in insects. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3526–E3534. doi: 10.1073/pnas.1310676110
- Johansson, K. U. I., Lundquist, C. T., Hallberg, E., and Nässel, D. R. (1999). Tachykinin-related neuropeptide in the crayfish olfactory midbrain. *Cell Tissue Res.* 296, 405–415. doi: 10.1007/s004410051300
- Johard, H. A., Coast, G. M., Mordue, W., and Nässel, D. R. (2003). Diuretic action of the peptide locustatachykinin I: cellular localisation and effects on fluid secretion in Malpighian tubules of locusts. *Peptides* 24, 1571–1579. doi: 10.1016/j.peptides.2003.08.012
- Johnson, E. C., Shafer, O. T., Trigg, J. S., Park, J., Schooley, D. A., Dow, J. A., et al. (2005). A novel diuretic hormone receptor in *Drosophila*: evidence for conservation of CGRP signaling. *J. Exp. Biol.* 208, 1239–1246. doi: 10.1242/jeb.01529
- Jung, J. W., Kim, J.-H., Pfeiffer, R., Ahn, Y.-J., Page, T. L., and Kwon, H. W. (2013). Neuromodulation of olfactory sensitivity in the peripheral olfactory organs of the American cockroach. *Periplaneta americana*. *PLoS One* 8:e81361. doi: 10.1371/journal.pone.0081361
- Kahsai, L., Kapan, N., Dirksen, H., Winther, ÅM., and Nässel, D. R. (2010a). Metabolic stress responses in *Drosophila* are modulated by brain neurosecretory cells that produce multiple neuropeptides. *PLoS One* 5:e11480. doi: 10.1371/journal.pone.0011480
- Kahsai, L., Martin, J. R., and Winther, ÅM. (2010b). Neuropeptides in the *Drosophila* central complex in modulation of locomotor behavior. *J. Exp. Biol.* 213, 2256–2265. doi: 10.1242/jeb.043190
- Kahsai, L., and Winther, A. M. (2011). Chemical neuroanatomy of the *Drosophila* central complex: distribution of multiple neuropeptides in relation to neurotransmitters. *J. Comp. Neurol.* 519, 290–315. doi: 10.1002/cne.22520
- Kamareddine, L., Robins, W. P., Berkey, C. D., Mekalanos, J. J., and Watnick, P. I. (2018). The *Drosophila* immune deficiency pathway modulates enteroendocrine function and host metabolism. *Cell Metab.* 28, 449.e–462.e. doi: 10.1016/j.cmet.2018.05.026
- Kanda, A., Iwakoshi-Ukena, E., Takuwa-Kuroda, K., and Minakata, H. (2003). Isolation and characterization of novel tachykinins from the posterior salivary gland of the common octopus *Octopus vulgaris*. *Peptides* 24, 35–43. doi: 10.1016/s0196-9781(02)00274-7
- Kanda, A., Takuwa-Kuroda, K., Aoyama, M., and Satake, H. (2007). A novel tachykinin-related peptide receptor of *Octopus vulgaris* – evolutionary aspects of invertebrate tachykinin and tachykinin-related peptide. *FEBS J.* 274, 2229–2239. doi: 10.1111/j.1742-4658.2007.05760.x
- Kapan, N., Lushchak, O. V., Luo, J., and Nässel, D. R. (2012). Identified peptidergic neurons in the *Drosophila* brain regulate insulin-producing cells, stress responses and metabolism by coexpressed short neuropeptide F and corazonin. *Cell Mol. Life Sci.* 69, 4051–4066. doi: 10.1007/s00018-012-1097-z
- Katsouni, E., Sakkas, P., Zarros, A., Skandali, N., and Liapi, C. (2009). The involvement of substance P in the induction of aggressive behavior. *Peptides* 30, 1586–1591. doi: 10.1016/j.peptides.2009.05.001
- Kawada, T., Furukawa, Y., Shimizu, Y., Minakata, H., Nomoto, K., and Satake, H. (2002). A novel tachykinin-related peptide receptor. *Eur. J. Biochem.* 269, 4238–4246. doi: 10.1046/j.1432-1033.2002.03106.x
- Kawada, T., Ogasawara, M., Sekiguchi, T., Aoyama, M., Hotta, K., Oka, K., et al. (2011). Peptidomic analysis of the central nervous system of the protochordate, *Ciona intestinalis*: homologs and prototypes of vertebrate peptides and novel peptides. *Endocrinology* 152, 2416–2427. doi: 10.1210/en.2010-1348
- Kawada, T., Sekiguchi, T., Sakai, T., Aoyama, M., and Satake, H. (2010). Neuropeptides, hormone peptides, and their receptors in *Ciona intestinalis*: an update. *Zool. Sci.* 27:120.
- Kim, M. Y., Lee, B. H., Kwon, D., Kang, H., and Nässel, D. R. (1998). Distribution of tachykinin-related neuropeptide in the developing central nervous system of the moth *Spodoptera litura*. *Cell Tissue Res.* 294, 351–365. doi: 10.1007/s004410051185
- Kim, Y. J., Bartalska, K., Audsley, N., Yamanaka, N., Yapici, N., Lee, J. Y., et al. (2010). MIPs are ancestral ligands for the sex peptide receptor. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6520–6525. doi: 10.1073/pnas.0914764107
- Ko, K. I., Root, C. M., Lindsay, S. A., Zaninovich, O. A., Shepherd, A. K., Wasserman, S. A., et al. (2015). Starvation promotes concerted modulation of appetitive olfactory behavior via parallel neuromodulatory circuits. *eLife* 4:e08298. doi: 10.7554/eLife.08298
- Kozioł, U. (2018). Precursors of neuropeptides and peptide hormones in the genomes of tardigrades. *Gen. Comp. Endocrinol.* 267, 116–127. doi: 10.1016/j.ygcen.2018.06.012
- Krishnan, A., and Schiöth, H. B. (2015). The role of G protein-coupled receptors in the early evolution of neurotransmission and the nervous system. *J. Exp. Biol.* 218:562. doi: 10.1242/jeb.110312
- Kubli, E. (2003). Sex-peptides: seminal peptides of the *Drosophila* male. *Cell Mol. Life Sci.* 60, 1689–1704. doi: 10.1007/s00018-003-3052
- Kwok, R., Chung, D., Brugge, V. T., and Orchard, I. (2005). The distribution and activity of tachykinin-related peptides in the blood-feeding bug. *Rhodnius prolixus*. *Peptides* 26, 43–51. doi: 10.1016/j.peptides.2004.08.024
- Langworthy, K., Helluy, S., Benton, J., and Beltz, B. (1997). Amines and peptides in the brain of the American lobster: immunocytochemical localization patterns and implications for brain function. *Cell Tissue Res.* 288, 191–206. doi: 10.1007/s004410050806
- Lazarus, L. H., and Attila, M. (1993). The toad, ugly and venomous, wears yet a precious jewel in his skin. *Prog. Neurobiol.* 41, 473–507. doi: 10.1016/0301-0082(93)90027-p
- Lehman, M. N., Coolen, L. M., and Goodman, R. L. (2010). Minireview: kisspeptin/neurokinin B/dynorphin (KNDy) Cells of the arcuate nucleus: a central node in the control of gonadotropin-releasing hormone secretion. *Endocrinology* 151, 3479–3489. doi: 10.1210/en.2010-0022
- Lemaitre, B., and Miguel-Aliaga, I. (2013). The digestive tract of *Drosophila melanogaster*. *Ann. Rev. Genet.* 47, 377–404. doi: 10.1146/annurev-genet-111212-133343
- Lénárd, L., László, K., Kertes, E., Ollmann, T., Péczely, L., Kovács, A., et al. (2018). Substance P and neurotensin in the limbic system: Their roles in reinforcement

- and memory consolidation. *Neurosci. Biobehav. Rev.* 85, 1–20. doi: 10.1016/j.neubiorev.2017.09.003
- Li, J., Liu, T., Xu, X., Wang, X., Wu, M., Yang, H., et al. (2006). Amphibian tachykinin precursor. *Biochem. Biophys. Res. Commun.* 350, 983–986. doi: 10.1016/j.bbrc.2006.09.150
- Li, X. J., Wolfgang, W., Wu, Y. N., North, R. A., and Forte, M. (1991). Cloning, heterologous expression and developmental regulation of a *Drosophila* receptor for tachykinin-like peptides. *EMBO J.* 10, 3221–3229. doi: 10.1002/j.1460-2075.1991.tb04885.x
- Liu, Y., Wang, Q., Wang, X., Meng, Z., Liu, Y., Li, S., et al. (2019). NKB/NK3 system negatively regulates the reproductive axis in sexually immature goldfish (*Carassius auratus*). *Gen. Comp. Endocrinol.* 281, 126–136. doi: 10.1016/j.ygcen.2019.05.020
- Lizbinski, K. M., and Dacks, A. M. (2017). Intrinsic and extrinsic neuromodulation of olfactory processing. *Front. Cell Neurosci.* 11:424. doi: 10.3389/fncel.2017.00424
- Ljungdahl, Å., Hökfelt, T., and Nilsson, G. (1978). Distribution of substance P-like immunoreactivity in the central nervous system of the rat—I. Cell bodies and nerve terminals. *Neuroscience* 3, 861–943. doi: 10.1016/0306-4522(78)90116-1
- Lundquist, C. T., Clottens, F. L., Holman, G. M., Riehm, J. P., Bonkale, W., and Nässel, D. R. (1994). Locustatachykinin immunoreactivity in the blowfly central nervous system and intestine. *J. Comp. Neurol.* 341, 225–240. doi: 10.1002/cne.903410208
- Lundquist, C. T., and Nässel, D. R. (1997). Peptidergic activation of locust dorsal unpaired median neurons: depolarization induced by locustatachykinins may be mediated by cyclic AMP. *J. Neurobiol.* 33, 297–315. doi: 10.1002/(sici)1097-4695(199709)33:3<297::aid-neu8>3.0.co;2-x
- Mancillas, J. R., and Selverston, A. I. (1985). Substance P-like immunoreactivity is present in the central nervous system of *Limulus polyphemus*. *J. Comp. Neurol.* 238, 38–52. doi: 10.1002/cne.902380104
- Masu, Y., Nakayama, K., Tamaki, H., Harada, Y., Kuno, M., and Nakanishi, S. (1987). cDNA cloning of bovine substance-K receptor through oocyte expression system. *Nature* 329, 836–838. doi: 10.1038/329836a0
- McVeigh, P., Mair, G. R., Atkinson, L., Ladurner, P., Zamanian, M., Novozhilova, E., et al. (2009). Discovery of multiple neuropeptide families in the phylum Platyhelminthes. *Int. J. Parasitol.* 39, 1243–1252. doi: 10.1016/j.ijpara.2009.03.005
- Mertens, L., Clinckspoor, I., Janssen, T., Nachman, R., and Schoofs, L. (2006). FMRamide related peptide ligands activate the *Caenorhabditis elegans* orphan GPCR Y59H11AL.1. *Peptides* 27, 1291–1296. doi: 10.1016/j.peptides.2005.11.017
- Meschi, E., Léopold, P., and Delanoue, R. (2019). An EGF-responsive neural circuit couples insulin secretion with nutrition in *Drosophila*. *Dev. Cell* 48, 76.e–86.e.
- Mirabeau, O., and Joly, J. S. (2013). Molecular evolution of peptidergic signaling systems in bilaterians. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2028–E2037. doi: 10.1073/pnas.1219956110
- Monnier, D., Colas, J. F., Rosay, P., Hen, R., Borrelli, E., and Maroteaux, L. (1992). NKD, a developmentally regulated tachykinin receptor in *Drosophila*. *J. Biol. Chem.* 267, 1298–1302.
- Moriyama, Y., and Koshiba-Takeuchi, K. (2018). Significance of whole-genome duplications on the emergence of evolutionary novelties. *Brief. Funct. Genom.* 17, 329–338. doi: 10.1093/bfpg/ely007
- Morteau, O., Lu, B., Gerard, C., and Gerard, N. P. (2001). Hemokinin 1 is a full agonist at the substance P receptor. *Nat. Immunol.* 2, 1088–1088. doi: 10.1038/nm1201-1088
- Muren, J. E., Lundquist, C. T., and Nässel, D. R. (1995). Abundant distribution of locustatachykinin-like peptide in the nervous system and intestine of the cockroach *Leucophaea maderae*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 348, 423–444. doi: 10.1098/rstb.1995.0079
- Muren, J. E., and Nässel, D. R. (1996). Isolation of five tachykinin-related peptides from the midgut of the cockroach *Leucophaea maderae*: existence of N-terminally extended isoforms. *Regul. Pept.* 65, 185–196. doi: 10.1016/0167-0115(96)00092-4
- Muren, J. E., and Nässel, D. R. (1997). Seven tachykinin-related peptides isolated from the brain of the Madeira cockroach: evidence for tissue-specific expression of isoforms. *Peptides* 18, 7–15. doi: 10.1016/s0196-9781(96)00243-4
- Nagai-Okatani, C., Nagasawa, H., and Nagata, S. (2016). Tachykinin-related peptides share a G protein-coupled receptor with ion transport peptide-like in the silkworm *Bombyx mori*. *PLoS One* 11:e0156501. doi: 10.1371/journal.pone.0156501
- Nakajo, M., Kanda, S., Karigo, T., Takahashi, A., Akazome, Y., Uenoyama, Y., et al. (2017). Evolutionally conserved function of kisspeptin neuronal system is nonreproductive regulation as revealed by nonmammalian study. *Endocrinology* 159, 163–183. doi: 10.1210/en.2017-00808
- Nakanishi, S. (1991). Mammalian tachykinin receptors. *Annu. Rev. Neurosci.* 14, 123–136. doi: 10.1146/annurev.neuro.14.1.123
- Nässel, D. R. (1993). Insect myotropic peptides: differential distribution of locustatachykinin- and leucokinin-like immunoreactive neurons in the locust brain. *Cell Tissue Res.* 274, 27–40. doi: 10.1007/bf00327982
- Nässel, D. R. (1999). Tachykinin-related peptides in invertebrates: a review. *Peptides* 20, 141–158. doi: 10.1016/s0196-9781(98)00142-9
- Nässel, D. R. (2002). Neuropeptides in the nervous system of *Drosophila* and other insects: multiple roles as neuromodulators and neurohormones. *Prog. Neurobiol.* 68, 1–84. doi: 10.1016/s0301-0082(02)00057-6
- Nässel, D. R. (2018). Substrates for neuronal cotransmission with neuropeptides and small molecule neurotransmitters in *Drosophila*. *Front. Cell Neurosci.* 12:83. doi: 10.3389/fncel.2018.00083
- Nässel, D. R., Eckert, M., Muren, J. E., and Penzlin, H. (1998). Species-specific action and distribution of tachykinin-related peptides in the foregut of the cockroaches *Leucophaea maderae* and *Periplaneta americana*. *J. Exp. Biol.* 201, 1615–1626.
- Nässel, D. R., Kim, M. Y., and Lundquist, C. T. (1995a). Several forms of callitachykinins are distributed in the central nervous system and intestine of the blowfly *Calliphora vomitoria*. *J. Exp. Biol.* 198, 2527–2536.
- Nässel, D. R., Passier, P. C., Elekes, K., Dircksen, H., Vullings, H. G., and Cantera, R. (1995b). Evidence that locustatachykinin I is involved in release of adipokinetic hormone from locust corpora cardiaca. *Regul. Pept.* 57, 297–310. doi: 10.1016/0167-0115(95)00043-b
- Nässel, D. R., and Lundquist, C. T. (1991). Insect tachykinin-like peptide: distribution of leucokinin immunoreactive neurons in the cockroach and blowfly brains. *Neurosci. Lett.* 130, 225–228. doi: 10.1016/0304-3940(91)90402-f
- Nässel, D. R., Mentlein, R., Bollner, T., and Karlsson, A. (2000). Proline-specific dipeptidyl peptidase activity in the cockroach brain and intestine: partial characterization, distribution, and inactivation of tachykinin-related peptides. *J. Comp. Neurol.* 418, 81–92. doi: 10.1002/(sici)1096-9861(20000228)418:1<81::aid-cne6>3.0.co;2-b
- Nässel, D. R., Pauls, D., and Huetteroth, W. (2019). Neuropeptides in modulation of *Drosophila* behavior: how to get a grip on their pleiotropic actions. *Curr. Opin. Insect Sci.* 36, 1–8. doi: 10.1016/j.cois.2019.03.002
- Nässel, D. R., and 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. doi: 10.1007/s00018-015-2063-3
- Nässel, D. R., and Zandawala, M. (2019). Recent advances in neuropeptide signaling in *Drosophila*, from genes to physiology and behavior. *Prog. Neurobiol.* 179:101607. doi: 10.1016/j.pneurobio.2019.02.003
- Navarro, V. (2012). New insights into the control of pulsatile gnRH release: the role of Kiss1/Neurokinin B neurons. *Front. Endocrinol.* 3:48. doi: 10.3389/fendo.2012.00048
- Nielsen, S. K. D., Koch, T. L., Hauser, F., Garm, A., and Grimmelikhuijzen, C. J. P. (2019). De novo transcriptome assembly of the cubomedusa *Tripedalia cystophora*, including the analysis of a set of genes involved in peptidergic neurotransmission. *BMC Genomics* 20:175. doi: 10.1186/s12864-019-5514-7
- Nikitin, M. (2015). Bioinformatic prediction of *Trichoplax adhaerens* regulatory peptides. *Gen. Comp. Endocrinol.* 212, 145–155. doi: 10.1016/j.ygcen.2014.03.049
- Nusbaum, M. P., and Blitz, D. M. (2012). Neuropeptide modulation of microcircuits. *Curr. Opin. Neurobiol.* 22, 592–601. doi: 10.1016/j.conb.2012.01.003
- Nusbaum, M. P., Blitz, D. M., and Marder, E. (2017). Functional consequences of neuropeptide and small-molecule co-transmission. *Nat. Rev. Neurosci.* 18, 389–403. doi: 10.1038/nrn.2017.56
- Ogawa, S., Ramadasan, P. N., Goschorska, M., Anantharajah, A., Ng, K. W., and Parhar, I. S. (2012). Cloning and expression of tachykinins and their association

- with kisspeptins in the brains of zebrafish. *J. Comp. Neurol.* 520, 2991–3012. doi: 10.1002/cne.23103
- Oh, Y., Lai, J. S.-Y., Mills, H. J., Erdjument-Bromage, H., Giammarinaro, B., Saadipour, K., et al. (2019). A glucose-sensing neuron pair regulates insulin and glucagon in *Drosophila*. *Nature* 574, 559–564. doi: 10.1038/s41586-019-1675-4
- O'Harte, F., Burcher, E., Lovas, S., Smith, D. D., Vaudry, H., and Conlon, J. M. (1991). Ranakinin: a novel NK1 tachykinin receptor agonist isolated with neurokinin b from the brain of the frog *Rana ridibunda*. *J. Neurochem.* 57, 2086–2091. doi: 10.1111/j.1471-4159.1991.tb06426.x
- Ohno, H., Yoshida, M., Sato, T., Kato, J., Miyazato, M., Kojima, M., et al. (2017). Luqin-like RYamide peptides regulate food-evoked responses in *C. elegans*. *eLife* 6:e28877. doi: 10.7554/eLife.28877
- Olpe, H. R., Heid, J., Bittiger, H., and Steinmann, M. W. (1987). Substance P depresses neuronal activity in the rat olfactory bulb in vitro and in vivo: possible mediation via γ -aminobutyric acid release. *Brain Res.* 412, 269–274. doi: 10.1016/0006-8993(87)91133-4
- Onaga, T. (2014). Tachykinin: recent developments and novel roles in health and disease. *Biomol. Concepts* 5, 225–243. doi: 10.1515/bmc-2014-0008
- Otsuka, M., and Yoshioka, K. (1993). Neurotransmitter functions of mammalian tachykinins. *Physiol. Rev.* 73, 229–308. doi: 10.1152/physrev.1993.73.2.229
- Owusu-Ansah, E., and Perrimon, N. (2014). Modeling metabolic homeostasis and nutrient sensing in *Drosophila*: implications for aging and metabolic diseases. *Dis. Models Mech.* 7, 343–350. doi: 10.1242/dmm.012989
- Oya, S., Kohno, H., Kainoh, Y., Ono, M., and Kubo, T. (2017). Increased complexity of mushroom body Kenyon cell subtypes in the brain is associated with behavioral evolution in hymenopteran insects. *Sci. Rep.* 7:13785. doi: 10.1038/s41598-017-14174-6
- Palamiuc, L., Noble, T., Witham, E., Ratanpal, H., Vaughan, M., and Srinivasan, S. (2017). A tachykinin-like neuroendocrine signalling axis couples central serotonin action and nutrient sensing with peripheral lipid metabolism. *Nat. Commun.* 8:14237. doi: 10.1038/ncomms14237
- Park, J. H., Chen, J., Jang, S., Ahn, T. J., Kang, K., Choi, M. S., et al. (2016). A subset of enteroendocrine cells is activated by amino acids in the *Drosophila* midgut. *FEBS Lett.* 590, 493–500. doi: 10.1002/1873-3468.12073
- Parker, D., and Grillner, S. (1998). Cellular and synaptic modulation underlying substance P-mediated plasticity of the lamprey locomotor network. *J. Neurosci.* 18:8095. doi: 10.1523/jneurosci.18-19-08095.1998
- Persson, M. G., and Nässel, D. R. (1999). Neuropeptides in insect sensory neurones: tachykinin-, FMRFamide- and allatotropin-related peptides in terminals of locust thoracic sensory afferents. *Brain Res.* 816, 131–141. doi: 10.1016/s0006-8993(98)01139-1
- Pinto, F. M., Bello, A. R., Gallardo-Castro, M., Valladares, F., Almeida, T. A., Tena-Sempere, M., et al. (2015). Analysis of the expression of Tachykinins and Tachykinin receptors in the rat uterus during early pregnancy. *Biol. Reprod.* 92, 600–606. doi: 10.1093/biolre/biv066
- Pinto, F. M., Ravina, C. G., Subiran, N., Cejudo-Román, A., Fernández-Sánchez, M., Irazusta, J., et al. (2010). Autocrine regulation of human sperm motility by tachykinins. *Reprod. Biol. Endocrinol.* 8:104. doi: 10.1186/1477-7827-8-104
- Poels, J., Birse, R. T., Nachman, R. J., Fichna, J., Janecka, A., Vanden Broeck, J., et al. (2009). Characterization and distribution of NKD, a receptor for *Drosophila* tachykinin-related peptide 6. *Peptides* 30, 545–556. doi: 10.1016/j.peptides.2008.10.012
- Powell, D., Leeman, S. E., Tregear, G. W., Niall, H. D., and Potts, J. J. T. (1973). Radioimmunoassay for substance P. *Nature* 241, 252–254.
- Predel, R., Neupert, S., Roth, S., Derst, C., and Nässel, D. R. (2005). Tachykinin-related peptide precursors in two cockroach species. *FEBS J.* 272, 3365–3375. doi: 10.1111/j.1742-4658.2005.04752.x
- Qi, X., Zhou, W., Li, S., Liu, Y., Ye, G., Liu, X., et al. (2015). Goldfish neurokinin B: cloning, tissue distribution, and potential role in regulating reproduction. *Gen. Comp. Endocrinol.* 221, 267–277. doi: 10.1016/j.ygcen.2014.10.017
- Roller, L., Yamanaka, N., Watanabe, K., Daubnerova, I., Zitnan, D., Kataoka, H., et al. (2008). The unique evolution of neuropeptide genes in the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* 38, 1147–1157. doi: 10.1016/j.ibmb.2008.04.009
- Sandeman, R. E., Sandeman, D. C., and Watson, A. (1990). Substance P antibody reveals homologous neurons with axon terminals among somata in the crayfish and crab brain. *J. Comp. Neurol.* 294, 569–582. doi: 10.1002/cne.902940405
- Satake, H., Aoyama, M., Sekiguchi, T., and Kawada, T. (2013). Insight into molecular and functional diversity of tachykinins and their receptors. *Protein Pept. Lett.* 20, 615–627. doi: 10.2174/0929866511320060002
- Satake, H., and Kawada, T. (2006). Overview of the primary structure, tissue-distribution, and functions of Tachykinins and their receptors. *Curr. Drug Targets* 7, 963–974. doi: 10.2174/138945006778019273
- Satake, H., Kawada, T., Nomoto, K., and Minakata, H. (2003). Insight into tachykinin-related peptides, their receptors, and invertebrate tachykinins: a review. *Zool. Sci.* 20, 533–549. doi: 10.2108/zsj.20.533
- Satake, H., Ogasawara, M., Kawada, T., Masuda, K., Aoyama, M., Minakata, H., et al. (2004). Tachykinin and Tachykinin receptor of an ascidian, *Ciona intestinalis*: evolutionary origin of the vertebrate tachykinin family. *J. Biol. Chem.* 279, 53798–53805. doi: 10.1074/jbc.m408161200
- Schiemann, R., Lammers, K., Janz, M., Lohmann, J., Paululat, A., and Meyer, H. (2018). Identification and in vivo characterisation of cardioactive peptides in *Drosophila melanogaster*. *Int. J. Mol. Sci.* 20:E20. doi: 10.3390/ijms20010002
- Schmidt, M. (1997a). Distribution of centrifugal neurons targeting the soma clusters of the olfactory midbrain among decapod crustaceans. *Brain Res.* 752, 15–25. doi: 10.1016/s0006-8993(96)01441-2
- Schmidt, M. (1997b). Distribution of presumptive chemosensory afferents with FMRFamide- or substance P-like immunoreactivity in decapod crustaceans. *Brain Res.* 746, 71–84. doi: 10.1016/s0006-8993(96)01187-0
- Schmidt, M., and Ache, B. W. (1994). Descending neurons with dopamine-like or with substance P/FMRFamide-like immunoreactivity target the somata of olfactory interneurons in the brain of the spiny lobster. *Panulirus argus*. *Cell Tissue Res.* 278, 337–352. doi: 10.1007/s004410050224
- Schoofs, L., Holman, G. M., Hayes, T. K., Kochansky, J. P., Nachman, R. J., and De Loof, A. (1990a). Locustatachykinin III and IV: two additional insect neuropeptides with homology to peptides of the vertebrate tachykinin family. *Regul. Pept.* 31, 199–212. doi: 10.1016/0167-0115(90)90006-i
- Schoofs, L., Holman, G. M., Hayes, T. K., Nachman, R. J., and De Loof, A. (1990b). Locustatachykinin I and II, two novel insect neuropeptides with homology to peptides of the vertebrate tachykinin family. *FEBS Lett.* 261, 397–401. doi: 10.1016/0014-5793(90)80601-e
- Schoofs, L., Vanden Broeck, J., and De Loof, A. (1993). The myotropic peptides of *Locusta migratoria*: structures, distribution, functions and receptors. *Insect Biochem. Mol. Biol.* 23, 859–881. doi: 10.1016/0965-1748(93)90104-z
- Semmens, D. C., Mirabeau, O., Moghul, I., Pancholi, M. R., Wurm, Y., and Elphick, M. R. (2016). Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. *Open Biol.* 6:150224. doi: 10.1098/rsob.150224
- Senatore, A., Reese, T. S., and Smith, C. L. (2017). Neuropeptidergic integration of behavior in *Trichoplax adhaerens*, an animal without synapses. *J. Exp. Biol.* 220:3381. doi: 10.1242/jeb.162396
- Shafer, O. T., Kim, D. J., Dunbar-Yaffe, R., Nikolaev, V. O., Lohse, M. J., and Taghert, P. H. (2008). Widespread receptivity to neuropeptide PDF throughout the neuronal circadian clock network of *Drosophila* revealed by real-time cyclic AMP imaging. *Neuron* 58, 223–237. doi: 10.1016/j.neuron.2008.02.018
- Shankar, S., Chua, J. Y., Tan, K. J., Calvert, M. E., Weng, R., Ng, W. C., et al. (2015). The neuropeptide tachykinin is essential for pheromone detection in a gustatory neural circuit. *eLife* 4:e06914. doi: 10.7554/eLife.06914
- Siviter, R. J., Coast, G. M., Winther, Å.M., Nachman, R. J., Taylor, C. A., Shirras, A. D., et al. (2000). Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A. *J. Biol. Chem.* 275, 23273–23280. doi: 10.1074/jbc.m002875200
- Skaer, N. J., Nässel, D. R., Maddrell, S. H., and Tublitz, N. J. (2002). Neurochemical fine tuning of a peripheral tissue: peptidergic and aminergic regulation of fluid secretion by *Malpighian tubules* in the tobacco hawkmoth *M. sexta*. *J. Exp. Biol.* 205, 1869–1880.
- Sliwowska, J., Rosinski, G., and Nässel, D. R. (2001). Cardioacceleratory action of tachykinin-related neuropeptides and proctolin in two coleopteran insect species. *Peptides* 22, 209–217. doi: 10.1016/s0196-9781(00)00384-3
- Söderberg, J. A., Birse, R. T., and Nässel, D. R. (2011). Insulin production and signaling in renal tubules of *Drosophila* is under control of tachykinin-related peptide and regulates stress resistance. *PLoS One* 6:e19866. doi: 10.1371/journal.pone.0019866
- Song, W., Veenstra, J. A., and Perrimon, N. (2014). Control of lipid metabolism by tachykinin in *Drosophila*. *Cell Rep.* 9, 40–47. doi: 10.1016/j.celrep.2014.08.060

- Steinhoff, M. S., Von Mentzer, B., Geppetti, P., Pothoulakis, C., and Bunnett, N. W. (2014). Tachykinins and their receptors: contributions to physiological control and the mechanisms of disease. *Physiol. Rev.* 94, 265–301. doi: 10.1152/physrev.00031.2013
- Takeuchi, H., Yasuda, A., Yasuda-Kamatani, Y., Sawata, M., Matsuo, Y., Kato, A., et al. (2004). Prepro-tachykinin gene expression in the brain of the honeybee *Apis mellifera*. *Cell Tissue Res.* 316, 281–293. doi: 10.1007/s00441-004-0865-y
- Thiel, D., Franz-Wachtel, M., Aguilera, F., and Hejnol, A. (2018). Xenacoelomorph neuropeptidomes reveal a major expansion of neuropeptide systems during early bilaterian evolution. *Mol. Biol. Evol.* 35, 2528–2543. doi: 10.1093/molbev/msy160
- Topaloglu, A. K., Reimann, F., Guclu, M., Yalin, A. S., Kotan, L. D., Porter, K. M., et al. (2009). TAC3 and TACR3 mutations in familial hypogonadotropic hypogonadism reveal a key role for Neurokinin B in the central control of reproduction. *Nat. Genet.* 41, 354–358. doi: 10.1038/ng.306
- Tregear, G. W., Niall, H. D., Potts, J. T., Leeman, S. E., and Chang, M. M. (1971). Synthesis of substance P. *Nature* 232, 87–89.
- Turner, R. A., Pierce, J. G., and Du Vigneaud, V. (1951). The purification and the amino acid content of vasopressin preparations. *J. Biol. Chem.* 191, 21–28.
- Van Loy, T., Vandersmissen, H. P., Poels, J., Van Hiel, M. B., Verlinden, H., and Vanden Broeck, J. (2010). Tachykinin-related peptides and their receptors in invertebrates: a current view. *Peptides* 31, 520–524. doi: 10.1016/j.peptides.2009.09.023
- Varoqueaux, F., Williams, E. A., Grandemange, S., Truscillo, L., Kamm, K., Schierwater, B., et al. (2018). High cell diversity and complex peptidergic signaling underlie placozoan behavior. *Curr. Biol.* 28, 3495.e–3501.e. doi: 10.1016/j.cub.2018.08.067
- Veenstra, J. A. (2009). Peptidergic paracrine and endocrine cells in the midgut of the fruit fly maggot. *Cell Tissue Res.* 336, 309–323. doi: 10.1007/s00441-009-0769-y
- Veenstra, J. A. (2010). Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen. Comp. Endocrinol.* 167, 86–103. doi: 10.1016/j.ygcen.2010.02.010
- Veenstra, J. A. (2011). Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. *Gen. Comp. Endocrinol.* 171, 160–175. doi: 10.1016/j.ygcen.2011.01.005
- Veenstra, J. A. (2016). Neuropeptide evolution: chelicerate neurohormone and neuropeptide genes may reflect one or more whole genome duplications. *Gen. Comp. Endocrinol.* 229, 41–55. doi: 10.1016/j.ygcen.2015.11.019
- Veenstra, J. A., Agrícola, H. J., and Sellami, A. (2008). Regulatory peptides in fruit fly midgut. *Cell Tissue Res.* 334, 499–516. doi: 10.1007/s00441-008-0708-3
- Veenstra, J. A., Rombauts, S., and Grbic, M. (2012). In silico cloning of genes encoding neuropeptides, neurohormones and their putative G-protein coupled receptors in a spider mite. *Insect Biochem. Mol. Biol.* 42, 277–295. doi: 10.1016/j.ibmb.2011.12.009
- Vitzthum, H., and Homberg, U. (1998). Immunocytochemical demonstration of locustatachykinin-related peptides in the central complex of the locust brain. *J. Comp. Neurol.* 390, 455–469. doi: 10.1002/(sici)1096-9861(19980126)390:4<455::aid-cne1>3.0.co;2-#
- Von Euler, U. S., and Gaddum, J. H. (1931). An unidentified depressor substance in certain tissue extracts. *J. Physiol.* 72, 74–87. doi: 10.1113/jphysiol.1931.sp002763
- Winther, Å.M., Acebes, A., and Ferrus, A. (2006). Tachykinin-related peptides modulate odor perception and locomotor activity in *Drosophila*. *Mol. Cell Neurosci.* 31, 399–406. doi: 10.1016/j.mcn.2005.10.010
- Winther, Å.M., Siviter, R. J., Isaac, R. E., Predel, R., and Nässel, D. R. (2003). Neuronal expression of tachykinin-related peptides and gene transcript during postembryonic development of *Drosophila*. *J. Comp. Neurol.* 464, 180–196. doi: 10.1002/cne.10790
- Winther, Å.M. E., Muren, J. E., Ahlborg, N., and Nässel, D. R. (1999). Differential distribution of isoforms of *Leucophaea tachykinin*-related peptides (LemTRPs) in endocrine cells and neuronal processes of the cockroach midgut. *J. Comp. Neurol.* 406, 15–28. doi: 10.1002/(sici)1096-9861(19990329)406:1<15::aid-cne2>3.3.co;2-7
- Winther, Å.M. E., and Nässel, D. R. (2001). Intestinal peptides as circulating hormones: release of tachykinin-related peptide from the locust and cockroach midgut. *J. Exp. Biol.* 204, 1269–1280.
- Wolfner, M. F. (2002). The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in *Drosophila*. *Heredity* 88, 85–93. doi: 10.1038/sj.hdy.6800017
- Wu, Y., Renjie, L., Jie, M., Mei, Z., Lei, W., Timothy, I. M., et al. (2013). Ranachensinin: a novel aliphatic tachykinin from the skin secretion of the chinese brown frog. *Rana chensinensis*. *Protein Pept. Lett.* 20, 1217–1224.
- Yañez-Guerra, L. A., Delroisse, J., Barreiro-Iglesias, A., Slade, S. E., Scrivens, J. H., and Elphick, M. R. (2018). Discovery and functional characterisation of a lujin-type neuropeptide signalling system in a deuterostome. *Sci. Rep.* 8:7220. doi: 10.1038/s41598-018-25606-2
- Yasuda-Kamatani, Y., and Yasuda, A. (2004). APSGFLGMRamide is a unique tachykinin-related peptide in crustaceans. *Eur. J. Biochem.* 271, 1546–1556. doi: 10.1111/j.1432-1033.2004.04065.x
- Yeoh, J. G. C., Pandit, A. A., Zandawala, M., Nässel, D. R., Davies, S. A., and Dow, J. A. T. (2017). DIneR: database for insect neuropeptide research. *Insect Biochem. Mol. Biol.* 86, 9–19. doi: 10.1016/j.ibmb.2017.05.001
- Zandawala, M., Moghul, I., Yañez Guerra, L. A., Delroisse, J., Abylkassimova, N., Hugall, A. F., et al. (2017). Discovery of novel representatives of bilaterian neuropeptide families and reconstruction of neuropeptide precursor evolution in ophiuroid echinoderms. *Open Biol.* 7, 170129. doi: 10.1098/rsob.170129
- Zhang, H., Wei, L., Zou, C., Bai, J., Song, Y., and Liu, H. (2013). Purification and characterization of a tachykinin-like peptide from skin secretions of the tree frog. *Theloderma kwangsiensis*. *Zool. Sci.* 30, 529–533. doi: 10.2108/zsj.30.529
- Zhang, Y., Lu, L., Furlonger, C., Wu, G. E., and Paige, C. J. (2000). Hemokinin is a hematopoietic-specific tachykinin that regulates B lymphopoiesis. *Nat. Immunol.* 1, 392–397. doi: 10.1038/80826
- Zhang, Z., Wen, H., Li, Y., Li, Q., Li, W., Zhou, Y., et al. (2019). TAC3 gene products regulate brain and digestive system gene expression in the spotted sea bass (*Lateolabrax maculatus*). *Front. Endocrinol.* 10:556. doi: 10.3389/fendo.2019.00556
- Zhao, X. C., Xie, G. Y., Berg, B. G., Schachtner, J., and Homberg, U. (2017). Distribution of tachykinin-related peptides in the brain of the tobacco budworm *Heliothis virescens*. *J. Comp. Neurol.* 525, 3918–3934. doi: 10.1002/cne.24310
- Zieglgänsberger, W. (2019). Substance P and pain chronicity. *Cell Tissue Res.* 375, 227–241. doi: 10.1007/s00441-018-2922-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Nässel, Zandawala, Kawada and Satake. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Global Neuropeptide Annotations From the Genomes and Transcriptomes of Cubozoa, Scyphozoa, Staurozoa (Cnidaria: Medusozoa), and Octocorallia (Cnidaria: Anthozoa)

Thomas L. Koch and Cornelis J. P. Grimmelikhuijzen*

Section for Cell and Neurobiology, Department of Biology, University of Copenhagen, Copenhagen, Denmark

OPEN ACCESS

Edited by:

Elizabeth Amy Williams,
University of Exeter, United Kingdom

Reviewed by:

David Plachetzki,
University of New Hampshire,
United States
Meet Zandawala,
Brown University, United States

*Correspondence:

Cornelis J. P. Grimmelikhuijzen
cgrimmelikhuijzen@bio.ku.dk

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 15 July 2019

Accepted: 13 November 2019

Published: 06 December 2019

Citation:

Koch TL and Grimmelikhuijzen CJP
(2019) Global Neuropeptide
Annotations From the Genomes and
Transcriptomes of Cubozoa,
Scyphozoa, Staurozoa (Cnidaria:
Medusozoa), and Octocorallia
(Cnidaria: Anthozoa).
Front. Endocrinol. 10:831.
doi: 10.3389/fendo.2019.00831

During animal evolution, ancestral Cnidaria and Bilateria diverged more than 600 million years ago. The nervous systems of extant cnidarians are strongly peptidergic. Neuropeptides have been isolated and sequenced from a few model cnidarians, but a global investigation of the presence of neuropeptides in all cnidarian classes has been lacking. Here, we have used a recently developed software program to annotate neuropeptides in the publicly available genomes and transcriptomes from members of the classes Cubozoa, Scyphozoa, and Staurozoa (which all belong to the subphylum Medusozoa) and contrasted these results with neuropeptides present in the subclass Octocorallia (belonging to the class Anthozoa). We found three to six neuropeptide preprohormone genes in members of the above-mentioned cnidarian classes or subclasses, each coding for several (up to thirty-two) similar or identical neuropeptide copies. Two of these neuropeptide preprohormone genes are present in all cnidarian classes/subclasses investigated, so they are good candidates for being among the first neuropeptide genes evolved in cnidarians. One of these primordial neuropeptide genes codes for neuropeptides having the C-terminal sequence GRFamide (pQGRFamide in Octocorallia; pQWLRGRFamide in Cubozoa and Scyphozoa; pQFLRGRFamide in Staurozoa). The other primordial neuropeptide gene codes for peptides having RPRSamide or closely resembling amino acid sequences. In addition to these two primordial neuropeptide sequences, cnidarians have their own class- or subclass-specific neuropeptides, which probably evolved to serve class/subclass-specific needs. When we carried out phylogenetic tree analyses of the GRFamide or RPRSamide preprohormones from cubozoans, scyphozoans, staurozoans, and octocorallia, we found that their phylogenetic relationships perfectly agreed with current models of the phylogeny of the studied cnidarian classes and subclasses. These results support the early origins of the GRFamide and RPRSamide preprohormone genes.

Keywords: neuropeptide, evolution, nervous system, Cnidaria, phylogeny

TABLE 1 | Accession numbers for the different databases used.

Species	Class	Database type	Accession number
<i>Renilla reniformis</i>	Octocorallia	WGS	FXAL00000000.1
<i>Eleutherobia rubra</i>	Octocorallia	TSA	GHFI00000000.1
<i>Xenia</i> sp. KK-2018	Octocorallia	TSA	GHBC00000000.1
<i>Briareum asbestinum</i>	Octocorallia	TSA	GHBD00000000.1
<i>Clavularia</i> sp.	Octocorallia	TSA	GHAW00000000.1
<i>Heliopora coerulea</i>	Octocorallia	TSA	GFVH00000000.1 IABP00000000.1
<i>Acanthogorgia aspera</i>	Octocorallia	TSA	GETB00000000.1 GEXC00000000.1
<i>Rhopilema esculentum</i>	Scyphozoa	TSA	GEMS00000000.1
<i>Nemopilema nomurai</i>	Scyphozoa	TSA	GHAR00000000.1
<i>Aurelia aurita</i>	Scyphozoa	TSA	GBRG00000000.1
<i>Aurelia aurita</i>	Scyphozoa	WGS	REGM00000000.1
<i>Caladosia cruxmelitensis</i>	Staurozoa	TSA	HAHC00000000.1
<i>Calvadosia cruxmelitensis</i>	Staurozoa	WGS	OFHS00000000.1
<i>Haliclystus sanjuanensis</i>	Staurozoa	TSA	HAHB00000000.1
<i>Craterolophus convolvulus</i>	Staurozoa	TSA	HAGZ00000000.1
<i>Lucernaria quadricornis</i>	Staurozoa	TSA	HAHD00000000.1
<i>Haliclystus aricula</i>	Staurozoa	TSA	HAHA00000000.1

the early cnidarian nervous systems. In a recent paper we have developed a bioinformatics tool to predict neuropeptide preprohormone genes from several cubozoan transcriptomes (31). In our current paper we have applied this script to predict neuropeptide preprohormone genes in cnidarian species with publicly accessible genomes or transcriptomes that belong to three classes (Scyphozoa, Staurozoa, Cubozoa), all belonging to the Medusozoa. We have compared these data from Medusozoa with a prediction of neuropeptide genes present in Octocorallia, where this subclass was used as a kind of outgroup (see also **Figure 1**) to generate more contrasts in our results.

The aim of the current paper was to determine whether these cnidarian classes produce the same types of neuropeptides, or whether there exist class-specific neuropeptides. When common neuropeptides would be present in these classes, these peptides would be good candidates for being among the first neuropeptides that evolved during cnidarian evolution.

MATERIALS AND METHODS

Sequence Data

We investigated assembled genomes (WGSs) and transcriptomes (TSAs) from seven octocorallians (*Renilla reniformis*, *Eleutherobia rubra*, *Xenia* sp. KK-2018, *Briareum asbestinum*, *Clavularia* sp., *Heliopora coerulea* and *Acanthogorgia aspera*), three scyphozoans (*Aurelia aurita*, *Rhopilema esculentum*, and *Nemopilema nomurai*) and five staurozoans (*Calvadosia cruxmelitensis*, *Haliclystus aricula*, *Haliclystus sanjuanensis*, *Craterolophus convolvulus*, and *Lucernaria quadricornis*). The database accession numbers are shown in **Table 1**.

Identification of Neuropeptide Preprohormones

We screened the translated genomes and transcriptomes for neuropeptide preprohormones using a script that is extensively explained in Nielsen et al. (31). This script is based on the presence of at least three similar peptide sequences followed by classical preprohormone processing sites (“GR” and “GK”) in the proteins. This script has, of course, its limitations, because neuropeptide genes might be missed that code for two or one neuropeptide copies on their preprohormones. Proteins with at least three processing sites were manually curated and labeled as neuropeptide preprohormones based on the presence of a signal peptide, the presence of three or more potential neuropeptide sequences, and the overall structure of the protein. The C-terminal parts of the immature neuropeptide sequences are easy to identify, because they are often followed by GR, GRR, or GKR sequences, which are classical processing sites for prohormone convertases (R, RR, KR), while the G residues are a classical processing signal for C-terminal amidation (3, 4). The N-termini, however, are often more difficult to determine as, in Cnidaria, N-terminal processing occurs by an unknown unspecific aminopeptidase cleaving at multiple residues, but stopping at Q residues, which are converted into N-terminal pQ groups (3, 4). In addition, N-terminal processing stops at N-terminal P or X-P sequences, which are also resistant to N-terminal degradation. The residues that are preferred for N-terminal processing are E, D, S, T, N, G, A, L, V, Y, or F (3). These residues often form the spacings in between the immature neuropeptide sequences on cnidarian preprohormones.

The identified neuropeptide preprohormones were also used as queries in TBLASTN searches against the other data sets using standard settings.

The putative preprohormones were investigated for the presence of signal peptides using SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/>) (34).

Phylogenetic Analysis

The preprohormones were aligned using ClustalW (35). For phylogenetic tree analysis the aligned protein sequences were loaded in PAUP*¹ and the maximum parsimony tree was calculated using p-distance and visualized in figtree².

RESULTS

Annotation of Neuropeptide Preprohormones in Scyphozoa

Using our script for the discovery of neuropeptide preprohormones in cnidarian genomes and transcriptomes (31), we could detect six neuropeptide preprohormone genes in four publicly accessible databases from three scyphozoans: *Rhopilema esculentum* (transcriptome) *Nemopilema nomurai* (transcriptome), and *Aurelia aurita* (genome and transcriptome) (**Table 1**). The script detects neuropeptide genes that code for preprohormones that have three or more neuropeptide

¹<https://paup.phylosolutions.com/> (accessed September 20, 2019).

²<https://github.com/rambaut/figtree> (accessed September 20, 2019).

TABLE 2 | An overview of scyphozoan neuropeptide families.

Neuropeptide family number	Species number	Species name	Neuropeptide sequence	Minimal number of neuropeptide copies
1.	Similar to Cubozoa, Staurozoa, and Octocorallia			
	1.	<i>Nemopilema nomurai</i>	LPR ^S amide	10
	2.	<i>Rhopilema esculentum</i>	LPR ^S amide	12
	3.	<i>Aurelia aurita</i>	MPR ^S amide	6
	3.	<i>Aurelia aurita</i>	RPR ^A amide	12
2.	Similar or identical to Cubozoa and Staurozoa, but different from Octocorallia			
	1.	<i>Nemopilema nomurai</i>	pQWLRGRFamide	29
	2.	<i>Rhopilema esculentum</i>	pQWLRGRFamide	32
	3.	<i>Aurelia aurita</i>	pQWLRGRFamide	22
3.	Similar or identical to Cubozoa and Staurozoa, but absent in Octocorallia			
	1.	<i>Nemopilema nomurai</i>	pQPPGVWamide	4
	1.	<i>Nemopilema nomurai</i>	pQPPGIW	8
	2.	<i>Rhopilema esculentum</i>	pQPPGVWamide	4
	2.	<i>Rhopilema esculentum</i>	pQPPGVW	5
	3.	<i>Aurelia aurita</i>	pQPPGVWamide	1
	3.	<i>Aurelia aurita</i>	pQPPGTWamide	1
	3.	<i>Aurelia aurita</i>	pQPPGTW	5
4.	Similar to Cubozoa, but absent in Staurozoa and Octocorallia			
	1.	<i>Nemopilema nomurai</i>	CTSPMCWFRP-amide	1
	1.	<i>Nemopilema nomurai</i>	CNSPMCWFRG-amide	1
	1.	<i>Nemopilema nomurai</i>	CDSPMCWFRP-amide	1
	2.	<i>Rhopilema esculentum</i>	CTSPMCWFRP-amide	1
	2.	<i>Rhopilema esculentum</i>	CNSPMCWFRA-amide	1
	2.	<i>Rhopilema esculentum</i>	CDSPMCWFRP-amide	1
	3.	<i>Aurelia aurita</i>	CSSPMCWFRPDamide	1
	3.	<i>Aurelia aurita</i>	CASPMCWFRAamide	1
	3.	<i>Aurelia aurita</i>	CSSPMCWFRA-amide	1
5.	Absent in Cubozoa, Staurozoa, and Octocorallia			
	1.	<i>Nemopilema nomurai</i>	pQHLYamide	6
	1.	<i>Nemopilema nomurai</i>	pQ-LRYamide	3
	2.	<i>Rhopilema esculentum</i>	pQHVRamide	6
	3.	<i>Aurelia aurita</i>	PHVRYamide	3
	3.	<i>Aurelia aurita</i>	PHLYamide	1
6.	Absent in Cubozoa, Staurozoa, and Octocorallia			
	1.	<i>Nemopilema nomurai</i>	pQPLWSARFamide	24
	1.	<i>Nemopilema nomurai</i>	pQPLWTGRamide	2
	1.	<i>Nemopilema nomurai</i>	PPFWSGRamide	1
	1.	<i>Nemopilema nomurai</i>	PPLWIGRFamide	1
	2.	<i>Rhopilema esculentum</i>	pQPLWSRamide	4
	2.	<i>Rhopilema esculentum</i>	pQPLWNSRamide	1
	2.	<i>Rhopilema esculentum</i>	pQ-L- - -RPamide	3
	2.	<i>Rhopilema esculentum</i>	pQPLWNGRamide	1
	3.	<i>Aurelia aurita</i>	PFWKVRFamide	1
	3.	<i>Aurelia aurita</i>	PLWSARFamide	1
	3.	<i>Aurelia aurita</i>	PLWKSRYamide	1
	3.	<i>Aurelia aurita</i>	PPWASRYamide	1
3.	<i>Aurelia aurita</i>	PFWNGRamide	2	

Only the major neuropeptides located on a preprohormone are listed here. The preprohormones are given in **Supplementary Figures 1–6**. Amino acid residues that are in common with the first-mentioned neuropeptide sequence from each neuropeptide family are highlighted in yellow.

sequences, thus neuropeptide genes might be missed that code for two or one neuropeptide copies. **Table 2** gives an overview of the neuropeptides contained in these six

preprohormones. **Supplementary Figures 1–6** give all the preprohormone sequences identified with our script in the three scyphozoan species.

TABLE 3 | An overview of staurozoan neuropeptide families.

Neuropeptide family number	Species number	Species name	Neuropeptide sequence	Minimal number of neuropeptide copies
1.	Similar to Cubozoa, Scyphozoa, and Octocorallia			
	1.	<i>Calvadosia cruxmelitensis</i>	RPRSamide	11
	2.	<i>Halicyclustus auricula</i>	RPRSamide	15
	3.	<i>Halicyclustus sanjuanensis</i>	RPRSamide	16
	4.	<i>Craterolophus convolvulus</i>	RPRSamide	8
	5.	<i>Lucernaria quadricornis</i>	RPRSamide	3
2.	Similar or identical to Cubozoa and Scyphozoa, but different from Octocorallia			
	1.	<i>Calvadosia cruxmelitensis</i>	pQFLRGRFamide	6
	1.	<i>Calvadosia cruxmelitensis</i>	pQFLKGRFamide	2
	2.	<i>Halicyclustus auricula</i>	pQFLRGRFamide	9
	3.	<i>Halicyclustus sanjuanensis</i>	pQFLRGRFamide	10
	5.	<i>Lucernaria quadricornis</i>	pQFLRGRFamide	4
3.	Similar to Cubozoa and Scyphozoa, but absent in Octocorallia			
	1.	<i>Calvadosia cruxmelitensis</i>	pQPP-GAWamide	4
	1.	<i>Calvadosia cruxmelitensis</i>	pQPP-GVWamide	3
	1.	<i>Calvadosia cruxmelitensis</i>	pQP--GAWamide	2
	2.	<i>Halicyclustus auricula</i>	pQPP-GVWamide	9
	3.	<i>Halicyclustus sanjuanensis</i>	pQPP-GVWamide	15
	4.	<i>Craterolophus convolvulus</i>	pQPP-GVWamide	6
	5.	<i>Lucernaria quadricornis</i>	pQPPKGTWamide	2
5.	<i>Lucernaria quadricornis</i>	pQLPTGTWamide	1	

Only the major neuropeptides located on a preprohormone are listed here. The preprohormones are given in **Supplementary Figures 7–9**. Amino acid residues are highlighted as in **Table 2**.

The scyphozoan databases from *N. nomurai*, *R. esculentum*, and *A. aurita* all contain transcripts and genes coding for a preprohormone that produce multiple copies of either LPRSamide or closely related neuropeptide sequences (**Table 2**; neuropeptide family #1). These sequences are flanked by classical GKR or GRR processing sites at their C-termini, where cleavage occurs C-terminally of K or R, after which the C-terminal G residues are converted into a C-terminal amide group (3, 4). At the N-termini of the neuropeptide sequences are acidic (E or D), or S, T, G, N, A, M, L, or V residues, which are processing sites that are often used in cnidarians for N-terminal neuropeptide processing (3, 4) (**Supplementary Figure 1**). In the proposed mature neuropeptide sequences, the C-termini are protected by an amide bond (for example LPRSamide), while the N-termini are protected by a prolyl residue in the second position of the peptide (**Table 2**). Similar LPRSamide neuropeptide sequences can also be detected in databases from Staurozoa (**Table 3**), Cubozoa (**Table 4**), and Octocorallia (**Table 5**).

The Scyphozoan databases from *N. nomurai*, *R. esculentum*, and *A. aurita* also contain transcripts and genes that code for numerous copies of identical pQWLRGRFamide neuropeptides (**Table 2**; neuropeptide family #2). These GRFamide neuropeptides have classical C-terminal GR or GKR processing sites and, again acidic (D or E), K, N, G, or S N-terminal processing sites that are often used in cnidarian N-terminal preprohormone processing (3, 4) (**Supplementary Figure 2**).

Similar GRFamide peptides occur in Staurozoa (**Table 3**), and identical GRFamide neuropeptides occur in Cubozoa (**Table 4**). In Octocorallia GRFamide neuropeptides exist that have the C-terminal GRFamide sequence in common with the scyphozoan pQWLRGRFamide peptides (**Table 5**).

All four databases show that scyphozoans also produce a neuropeptide preprohormone that code for several copies of pQPPGVWamide and pQPPGTWamide and their non-amidated variants pQPPGVW and pQPPGTW (**Table 2**; neuropeptide family #3). These preprohormones have classical GKR, RR, RRR, RKR, or RKK processing sites at the C-termini of the neuropeptide sequences and N-terminal K, S, or N residues that in cnidarians are known to be involved in N-terminal processing (3, 4) (**Supplementary Figure 3**). Similar pQPPGVWamide peptides also occur in Staurozoa (**Table 3**; peptide family #3) and Cubozoa (**Table 4**; peptide family #3), but they are absent in Octocorallia (**Table 5**).

The *N. nomurai* transcriptome database also codes for a preprohormone that contains one copy of a probable cyclic neuropeptide CTSPMCWFRPamide and several other nearly identical peptides, where the two cysteine residues are likely forming a cystine bridge (**Table 2**; neuropeptide family #4; **Supplementary Figure 4**). Similar peptides can be found in the databases from *R. esculentum*, where a small number of amino acid residue exchanges occur without, however, changing the consensus sequence CXSPMCWFRXamide (**Table 2**; peptide

TABLE 4 | An overview of cubozoan neuropeptide families.

Neuropeptide family number	Species number	Species name	Neuropeptide sequence	Minimal number of neuropeptide copies
1.	Similar to Scyphozoa, Staurozoa, and Octocorallia			
	1.	<i>Tripedalia cystophora</i>	RPRAAmide	13
	2.	<i>Alatina alata</i>	RPRAAmide	14
	3.	<i>Carybdea xaymacana</i>	RPRAAmide	2
	4.	<i>Chironex fleckeri</i>	RPRAAmide	3
2.	Similar or identical to Scyphozoa and Staurozoa, but different from Octocorallia			
	1.	<i>Tripedalia cystophora</i>	pQWLRGRFamide	19
	2.	<i>Alatina alata</i>	pQWLRGRFamide	18
	3.	<i>Carybdea xaymacana</i>	pQWLRGRFamide	11
	4.	<i>Chironex fleckeri</i>	pQWLRGRFamide	7
	5.	<i>Chiropsalmus quadrumanus</i>	pQWLRGRFamide	1
3.	Similar or identical to Scyphozoa and Staurozoa, but absent in Octocorallia			
	1.	<i>Tripedalia cystophora</i>	pQPPGVWamide	6
	2.	<i>Alatina alata</i>	pQPPGVWamide	6
	3.	<i>Carybdea xaymacana</i>	pQPPGVWamide	5
	4.	<i>Chironex fleckeri</i>	pQPPGVWamide	4
4.	Similar to Scyphozoa, but absent in Staurozoa and Octocorallia			
	1.	<i>Tripedalia cystophora</i>	CKGQMCWFRamide	2
	1.	<i>Tripedalia cystophora</i>	CTGQMCWFRamide	4
	1.	<i>Tripedalia cystophora</i>	CVGQMCWFRamide	1
	2.	<i>Alatina alata</i>	CKGQMCWFRamide	1
	2.	<i>Alatina alata</i>	CTGQMCWFRamide	2
	2.	<i>Alatina alata</i>	CVGQMCWFRamide	2
	2.	<i>Alatina alata</i>	CEGQMCWFRamide	1
	3.	<i>Chironex fleckeri</i>	CKGQMCWFRamide	1
	3.	<i>Chironex fleckeri</i>	CTGQMCWFRamide	1
5.	Absent in Scyphozoa, Staurozoa, and Octocorallia			
	1.	<i>Tripedalia cystophora</i>	...GLWamide	5
	1.	<i>Tripedalia cystophora</i>	...GMWamide	1
	2.	<i>Alatina alata</i>	...GLWamide	2
	2.	<i>Alatina alata</i>	...GMWamide	1
	4.	<i>Chironex fleckeri</i>	...GMWamide	1
6.	Absent in Scyphozoa, Staurozoa, and Octocorallia			
	1.	<i>Tripedalia cystophora</i>	...GRYamide	3
	1.	<i>Tripedalia cystophora</i>	...QRYamide	1
	2.	<i>Alatina alata</i>	...GRYamide	2
	2.	<i>Alatina alata</i>	...QRYamide	1
	3.	<i>Carybdea xaymacana</i>	...GRYamide	2
	3.	<i>Carybdea xaymacana</i>	...QRYamide	1

This table is a shortened version of Table 1 from Nielsen et al. (31). Only the major neuropeptides are shown here. Amino acid residues are highlighted as in **Table 2**.

family #4). In *A. aurita*, however, the C-termini are extended by one or two amino acid residues (**Table 2**; peptide family #4). The C-termini of all neuropeptide sequences in the preprohormones are followed by classical GR, GKR, GKRR, or GKRR processing sites and the neuropeptide sequences are preceded by G, N, D, E, or S sequences, which are known processing sites in cnidarian preprohormones (3, 4) (**Supplementary Figure 4**). Similar cyclic neuropeptides can also be found in the transcriptomes of three cubozoans, which all have the consensus

sequence CXGQMCWFRamide (**Table 4**, **Figure 2**). Thus, compared to the scyphozoan neuropeptides, these cubozoan neuropeptides have the sequences CXXXMCWFRamide in common, including a common distance between the cysteine residues forming the presumed cystine bridge (**Figure 2**). These neuropeptides could not be found in Staurozoa (**Table 3**), or Octocorallia (**Table 5**).

Finally, Scyphozoans have two neuropeptide families that cannot be found in Staurozoa, Cubozoa, and Octocorallia and

TABLE 5 | An overview of octocorallian neuropeptide families.

Neuropeptide family number	Species number	Species name	Neuropeptide sequence	Minimal number of neuropeptide copies
1.	Similar to Cubozoa, Scyphozoa, and Staurozoa			
	1.	<i>Renilla reniformis</i>	GPRGamide	5
	2.	<i>Eleutherobia rubra</i>	FPRGamide	2
	3.	<i>Xenia</i> sp.	APRGamide	6
	4.	<i>Briareum asbestinum</i>	EPRGamide	6
	5.	<i>Clavularia</i> sp.	APRGamide	5
	6.	<i>Heliopora coerulea</i>	APRGamide	2
	7.	<i>Acanthogorgia aspera</i>	KPRGamide	2
2.	Different from Cubozoa, Scyphozoa, Staurozoa, but the C-termini are identical			
	1.	<i>Renilla reniformis</i>	pQGRFamide	5
	2.	<i>Eleutherobia rubra</i>	pQGRFamide	27
	3.	<i>Xenia</i> sp.	pQGRFamide	9
	4.	<i>Briareum asbestinum</i>	pQGRFamide	16
	5.	<i>Clavularia</i> sp.	pQGRFamide	15
	6.	<i>Heliopora coerulea</i>	pQGRFamide	12
	7.	<i>Acanthogorgia aspera</i>	pQGRFamide	27
3.	Absent in Cubozoa, Scyphozoa, and Staurozoa			
	1.	<i>Renilla reniformis</i>	pQLRGamide	7
	2.	<i>Eleutherobia rubra</i>	pQLRGamide	18
	3.	<i>Xenia</i> sp.	pQLRAamide	5
	4.	<i>Briareum asbestinum</i>	pQLRSamide	2
	5.	<i>Clavularia</i> sp.	pQLRAamide	7
	6.	<i>Heliopora coerulea</i>	pQLRGamide	16
	7.	<i>Acanthogorgia aspera</i>	pQLRGamide	26
4.	Absent in Cubozoa, Scyphozoa, and Staurozoa			
	1.	<i>Renilla reniformis</i>	PPFHamide	3
	2.	<i>Eleutherobia rubra</i>	pQPFHamide	4
	3.	<i>Xenia</i> sp.	pQPFHamide	9
	4.	<i>Clavularia</i> sp.	pQPFHamide	6
	5.	<i>Heliopora coerulea</i>	RPFHamide	10
5.	Absent in Cubozoa, Scyphozoa, and Staurozoa			
	1.	<i>Renilla reniformis</i>	GPRRamide	2
	2.	<i>Eleutherobia rubra</i>	GPRRamide	8
	3.	<i>Xenia</i> sp.	GPRRamide	15
	4.	<i>Briareum asbestinum</i>	GPRRamide	22
	5.	<i>Clavularia</i> sp.	GPRRamide	5
	6.	<i>Heliopora coerulea</i>	GPRRamide	5
	7.	<i>Acanthogorgia aspera</i>	GPRRamide	4

Only the major neuropeptides located on a preprohormone are listed here. The preprohormones are given in **Supplementary Figures 10–14**. Amino acid residues are highlighted as in **Table 2**.

which, therefore, appear to be scyphozoan-specific (**Table 2**; peptide families #5 and #6).

In *N. nomurai*, the first neuropeptide family (neuropeptide family #5 of **Table 2**; **Supplementary Figure 5**) consists of members having the sequence pQHLYamide or other very similar sequences. In *R. esculentum*, six copies of the sequence pQHRYamide can be identified (**Table 2**). In *A. aurita*, a prolyl residue, which also protects the N-terminus of a neuropeptide, is replacing the pyroglutamyl residue: PHVRYamide and PHLRYamide (**Table 2**). In the preprohormones for these neuropeptides the neuropeptide sequences have classical GR

or GK processing sites at their C-termini, while at their N-termini they have R, D, T, and A residues, which in cnidarians are known to be involved in N-terminal processing (3, 4) (**Supplementary Figure 5**).

The second scyphozoan-specific neuropeptide family (neuropeptide family #6 of **Table 2**; **Supplementary Figure 6**) consists of pQPLWSARFamide or related sequences in *N. nomurai* (**Table 2**). For some peptides the N-terminal pyroglutamyl group is lacking, but those peptides still have two sequential N-terminal prolyl residues, which protect the N-termini against enzymatic degradation (**Table 2**). *R. esculentum*

Species	Class	Peptide name	Peptide sequence
<i>T. cystophora</i>	Cubozoa	Tcy-FRamide-1	CKGQMCWFR-amide
<i>T. cystophora</i>	Cubozoa	Tcy-FRamide-2	CTGQMCWFR-amide
<i>T. cystophora</i>	Cubozoa	Tcy-FRamide-3	CVGQMCWFR-amide
<i>N. nomurai</i>	Scyphozoa	Nno-FRGamide	CNSPMCWFRGamide

FIGURE 2 | Alignment of the cyclic FRamide peptides from the cubozoan *T. cystophora* and the scyphozoan *N. nomurai*. Amino acid residues that are in common with the cubozoan sequence CKGQMCWFRamide, named Tcy-FRamide [see Table 1 from Nielsen et al. (31)] are highlighted in yellow. Please note that the peptides are cyclic after the cysteine residues have formed a cystine bridge and that the resulting “loops” are six-rings with the same size in all peptides.

have peptides that are very similar to the ones occurring in *N. nomurai* with the exception of three copies of a short peptide pQLRPamide that do not occur in *N. nomurai*. All peptides from *A. aurita* lack N-terminal pyroglutamyl residues, but, again, are N-terminally protected by prolyl residues (Table 2).

Annotation of Neuropeptide Preprohormones in Staurozoa

Using our script (31), we discovered three different neuropeptide preprohormone genes in staurozoans (Table 3). In *Calvadosia cruxmelitensis* we found a preprohormone that contained 11 copies of the neuropeptide sequence RPRSamide (Table 3, neuropeptide family #1; Supplementary Figure 7). These RPRSamide sequences have the classical C-terminal processing site GKR, while N-terminally of the RPRSamide sequences are D, V, I, F, and A residues, which in cnidarians are known preprohormone processing sites (3, 4). Similar preprohormones exist in the transcriptomes from *Haliclystus auricula* and *Haliclystus sanjuanensis* (Supplementary Figure 7), which contain 15, respectively, 16 copies of the RPRSamide sequence (Table 3). In the *Craterolophus convolvulus* transcriptome we could identify a preprohormone with 8 copies of the RPRSamide sequence. In the transcriptome from *Lucernaria quadricornis* we discovered a neuropeptide preprohormone with 3 copies of the RPRSamide and 6 copies of the KPRSamide sequence (Table 3; Supplementary Figure 7). As mentioned earlier, neuropeptide preprohormones having numerous copies of RPRSamide or similar peptides, also occur in Scyphozoa (Table 2), Cubozoa (Table 4), and Octocorallia (Table 5).

The transcriptomes from the four Staurozoa species also contain transcripts coding for GRFamide preprohormones (Table 3, neuropeptide family #2). For all species these preprohormones contain numerous copies of the neuropeptide pQFLRGRFamide (Table 3; Supplementary Figure 8). This neuropeptide is very similar to the one from scyphozoans (Table 2, neuropeptide family #2), but it contains an F residue at position 2 instead of a W residue in the scyphozoan peptide. The same is true for cubozoans (Table 4, neuropeptide family #2), where the GRFamide peptides also have a W residue at position 2. However, compared to Octocorallia (Table 5, neuropeptide family #2) the staurozoan GRFamide peptides are much longer, being N-terminally elongated by three amino acid residues.

The third neuropeptide preprohormone that we discovered in staurozoans contains numerous copies

of a pQPPGAWamide neuropeptide or closely related sequences (Table 3, neuropeptide family #3; Supplementary Figure 9). All peptides are protected by amino-terminal pQ, pQP, or pQPP sequences against unspecific aminoterminal enzymatic degradation. Identical or similar neuropeptides occur in scyphozoans (Table 2, neuropeptide family #3) or cubozoans (Table 4, neuropeptide family #3). The peptides, however, are absent in Octocorallia (Table 5).

Annotation of Neuropeptide Preprohormones in Cubozoa

We have recently annotated neuropeptide preprohormones in five different cubozoan species (31). A short summary of these results is given in Table 4, while the amino acid sequences of the preprohormones are given in Nielsen et al. (31).

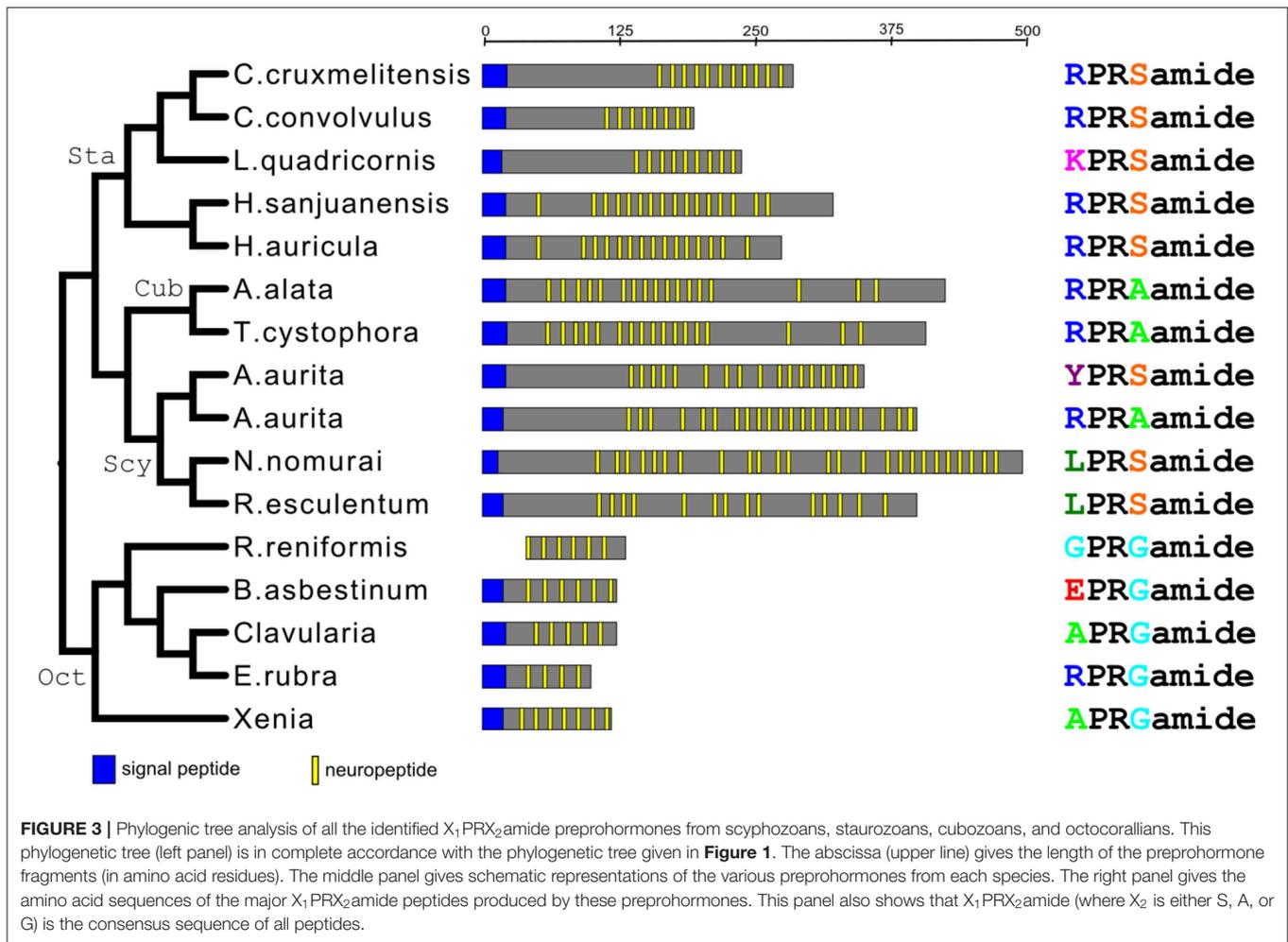
Cubozoans produce RPRamide neuropeptides (Table 4, neuropeptide family #1) that are very similar to neuropeptides occurring in scyphozoans (Table 1, neuropeptide family #1), staurozoans (Table 3, neuropeptide family #1) or octocorallians (Table 5, neuropeptide family #1).

Cubozoans also have preprohormones that produce numerous copies of pQWLRGRFamide (Table 4, neuropeptide family #2). Identical or very similar neuropeptides can also be found in scyphozoans (Table 2, neuropeptide family #2), and staurozoans (Table 3, neuropeptide family #2). However, in octocorallians there is a shorter version of these peptides (Table 5, neuropeptide family #2).

Cubozoans produce the neuropeptide pQPPGVWamide (Table 4, neuropeptide family #3) that is identical or very similar to neuropeptides produced in scyphozoans (Table 2, neuropeptide family #3), and staurozoans (Table 3, neuropeptide family #3). These peptides, however, do not occur in octocorallians (Table 5).

Cubozoans have peptides with the sequence CXGQMCWFRamide, which are probably cyclic after the two cysteine residues have formed a cystine bridge (Table 4, neuropeptide family #4). Scyphozoans have similar peptides (Table 2, neuropeptide family #4). However, these peptides are lacking in Staurozoa (Table 3) and Octocorallia (Table 5).

Finally, cubozoans have two peptide families that do not occur in Scypho- and Staurozoa. These are neuropeptides with the C-terminal sequence GLWamide (Table 4, neuropeptide family #5), and neuropeptides with the C-terminal sequence GRYamide



(**Table 4**, neuropeptide family #6). The C-termini from the GLWamides are relatively well-conserved, but their N-termini are quite variable [see **Table 1** from Nielsen et al. (31)]. The same is true for the GRYamide peptides (31).

Annotation of Neuropeptide Preprohormones From Octocorallia

We investigated the transcriptomes from seven Octocorallia species for the presence of neuropeptide genes (**Tables 1, 5**).

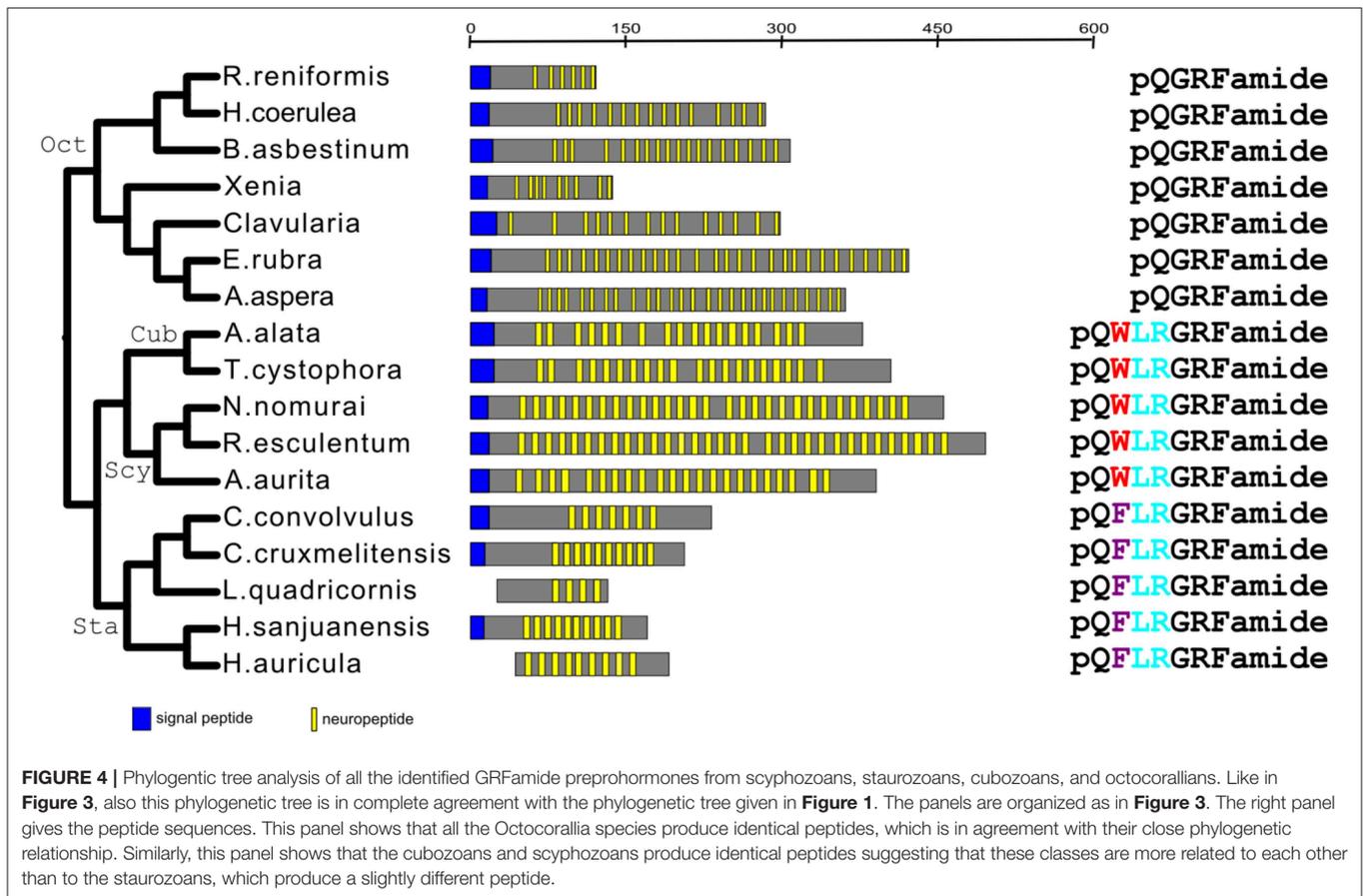
Octocorallia produce GPRGamide and closely related peptides (**Table 5**, neuropeptide family #1; **Supplementary Figure 10**) that resemble very much the LPRSamide and RPRAamide peptides from scyphozoans (**Table 2**, neuropeptide family #1), staurozoans (**Table 2**, neuropeptide family #1), and cubozoans (**Table 4**, neuropeptide family #1).

All octocorallians produce a preprohormone that carry numerous copies of the neuropeptide pQGRFamide (**Table 5**, neuropeptide family #2; **Supplementary Figure 11**). This peptide, dubbed Antho-RFamide, was the first cnidarian neuropeptide to be chemically isolated and sequenced from Anthozoa (36), including the octocoral *Renilla* (37). The presence of acidic residues, preceding the peptide sequence in the *Renilla* preprohormone (**Supplementary Figure 11**

illustrates, again, that processing must occur at E or D residues. Antho-RFamide does not occur in Scyphozoa, Staurozoa, and Cubozoa, but these medusozoans have N-terminally elongated forms that have the C-terminal GRFamide sequence in common with Antho-RFamide (neuropeptide families #2 from **Tables 2–4**). Thus, GRFamide neuropeptides are widespread in cnidarians.

In contrast to the two neuropeptide families discussed above that appear to be ubiquitous in cnidarians, octocorallians also produce Octocorallia-specific neuropeptides. The first group (**Table 5**, neuropeptide family #3) has the structure pQLRGamide or a very similar sequence. Their preprohormones are given in **Supplementary Figure 12**.

The second group of neuropeptides (**Table 5**, neuropeptide family #4) has the structure PPFHamide, or pQPFHamide. Both sequences are N-terminally protected by either the N-terminal PP or pQP sequences. Their preprohormones are given in **Supplementary Figure 13**. In addition, we discovered a preprohormone in *H. coerulea* that produces multiple copies of an RPFLamide sequence. Also these peptides are N-terminally protected by prolyl residues at position 2, but they are only 50% identical with the other peptides from this family (**Table 5**, peptide family #4; **Supplementary Figure 13**).



The third group of Octocorallia-specific neuropeptides (**Table 5**, neuropeptide family #5; **Supplementary Figure 14**) has the sequence GPRRamide, or a closely related sequence. Again, these neuropeptide sequences are protected against N-terminal enzymatic degradation by a prolyl residue at position 2 of the peptides. The *R. reniformis* preprohormone has at least two copies of GPRRamide, the *E. rubra* preprohormone produces eight copies of GPRRamide, the *Xenia* sp. preprohormone contains 15 GPRRamide copies, the *B. asbestinum* preprohormone 22 copies, the *Clavularia* sp. preprohormone 5 copies, and the *H. coerulea* preprohormone 4 copies of GPRRamide (**Table 5**, neuropeptide family #5; **Supplementary Figure 14**).

Phylogenetic Tree Analyses of the XPRXamide and GRFamide Preprohormones

We carried out a phylogenetic tree analysis of all the XPRSamide/XPRAamide/XPRGamide preprohormones investigated in this paper [neuropeptide families #1, from **Tables 2–5**; **Supplementary Figures 1, 7, 10**; (31)]. These studies (**Figure 3**) showed that the structural relationships between these preprohormones very precisely followed the established phylogenetic relationships of the classes and subclasses they belong to (**Figure 1**). These findings show that

all XPRSamide/XPRAamide/XPRGamide preprohormones are derived from a common ancestor.

When we carried out the same analysis for the GRFamide preprohormones, we came to the same conclusion (**Figures 1, 4**). Aligning the mature neuropeptide sequences themselves (right panel of **Figure 4**), showed that the octocorallian peptide pQGRFamide is farthest away from the other GRFamide peptides, while the pQWLRGRFamides are identical in both Cubo- and Scyphozoa and only slightly different from the pQFLRGRFamides that occur in Staurozoa. These findings are, again in complete agreement with the phylogenetic relationships between the classes and subclasses to which these peptides belong (**Figures 1, 4**), showing that all GRFamide preprohormones are derived from a common ancestor.

DISCUSSION

In our paper we have analyzed the neuropeptide preprohormones from three cnidarian classes (Scyphozoa, Cubozoa, Staurozoa) and one subclass (Octocorallia). We did not include the remaining classes (Hydrozoa and Myxozoa) and subclasses (Hexacorallia) in our study, because already the current study includes a large amount of data (**Supplementary Figures 1–14**) with altogether 66 preprohormones, each of which contains a varying number of different neuropeptides (**Tables 2–5**).

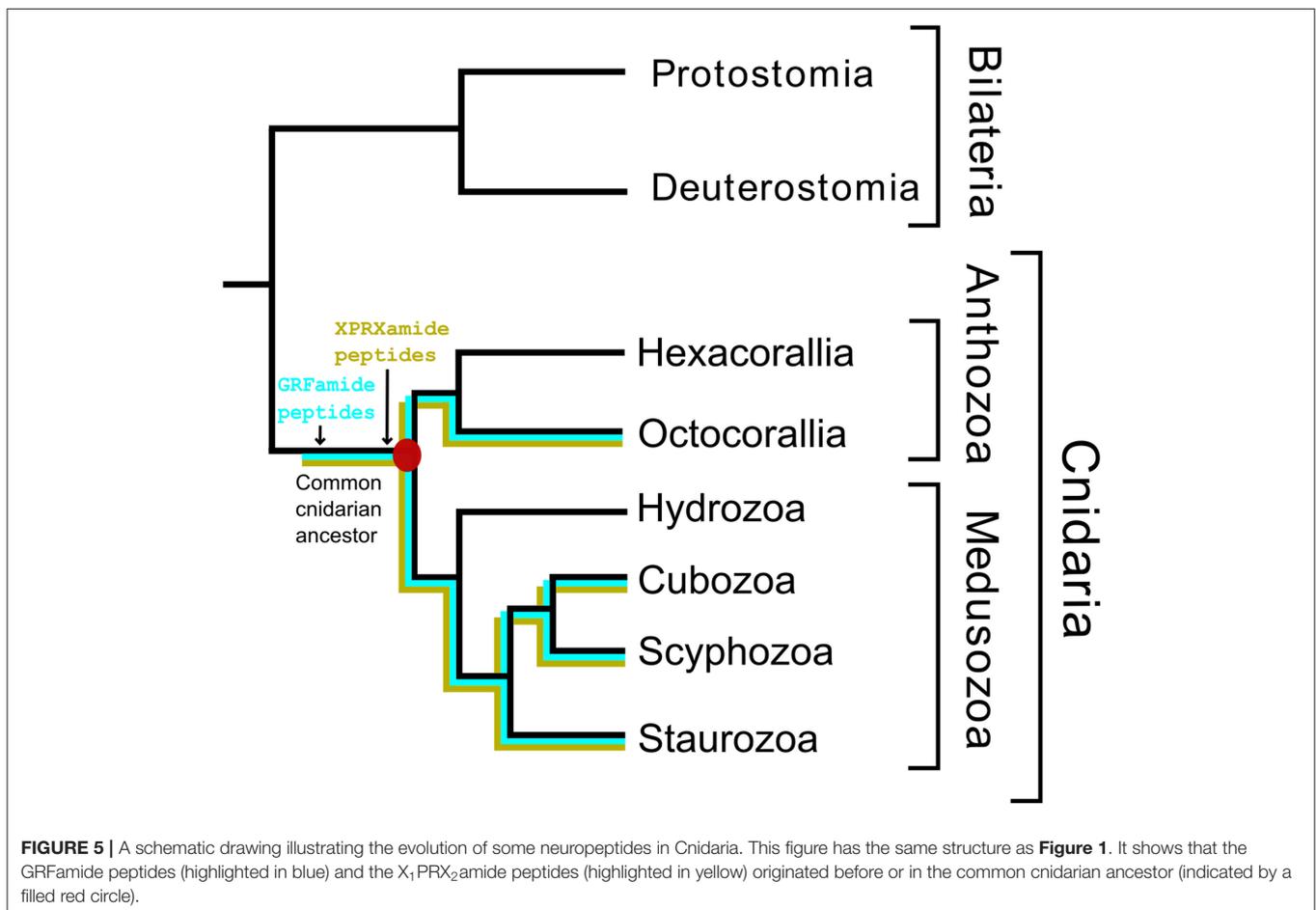
These large numbers of preprohormones and neuropeptides are difficult to analyze and present in an understandable and concise way. Yet, although not all cnidarian classes have been included in our analyses, we can already now draw conclusions related to the major questions that we asked at the end of the Introduction: (i) Do all cnidarian classes produce the same types of neuropeptides; or (ii) are there class-specific neuropeptides?

We found that Scyphozoa, Cubozoa, Staurozoa, and Octocorallia all produced GRFamide peptides (**Figure 4**), which is an answer to the above-mentioned question (i). This finding also means that the common cnidarian ancestor (red filled circle in **Figure 5**) must have produced GRFamides, suggesting that GRFamides are ancient neuropeptides that probably evolved together with the first cnidarians. pQGRFamide (Antho-RFamide) is a well-established neuropeptide that has been isolated, sequenced and cloned from the sea anemones *Anthopleura elegantissima* and *Calliactis parasitica* (Hexacorallia) and from the sea pansy *Renilla koellikeri* (Octocorallia) (3, 4, 36–39). Using immunocytochemistry, dense nets of Antho-RFamide producing neurons have been found in *C. parasitica* and *R. koellikeri* (2, 40), showing that these peptides are genuine neuropeptides.

N-terminally elongated forms of Antho-RFamide have been isolated and sequenced from Scyphozoa, such as pQWLRGRFamide from *Cyanea lamarckii* (compare right panel of **Figure 4**; and **Table 2**, peptide family #2) and the use of antibodies showed its presence in nerve nets of *C. lamarckii* (41, 42). Thus, both Antho-RFamide and its N-terminally elongated forms (right panel of **Figure 4**) are well-established neuropeptides in Cnidaria.

There is another peptide-family that occurs in all Octocorallia, Cubozoa, Scyphozoa, and Staurozoa species that we have investigated. These peptides have the general structure X_1PRX_2 amide, where X_1 is quite variable, while X_2 is confined to S, A, or G (neuropeptide family #1 from **Tables 2–5**; right panel of **Figure 3**). Because these peptides occur in both Octocorallia and the three classes that belong to the Medusozoa (**Figure 5**), it is likely that they also were present in the common cnidarian ancestor (red filled circle in **Figure 5**).

After submission of our manuscript and two reviewing rounds, a paper was published on neuropeptides present in the hexacoral *N. vectensis* (43). This paper confirms the presence of GRFamide and X_1PRX_2 peptides in Hexacorallia, which supports our conclusion that these two neuropeptide families originated in the common cnidarian ancestor (**Figure 5**). Furthermore,



this paper confirms the presence of class-specific neuropeptides (see below).

In contrast to the GRFamides, not much is known about the XPRXamides. However, two peptides related to this peptide family (WPRPamide and RPRPamide) have recently been identified in the hydrozoan jellyfishes *Clytia hemispherica* and *Cladonema pacificum* as the endogenous neuropeptides inducing oocyte maturation, and oocyte and sperm release (44). The peptides have also been localized in neurons and, therefore, are genuine neuropeptides (44).

If one looks at the structures of the various X_1PRX_2 amide peptides (right panel of **Figure 3**), one can see several interesting features. First, all peptides have a prolyl residue at position 2. This residue might help protecting the neuropeptide against non-specific enzymatic N-terminal degradation, because the X-P bond is not an amide, but an imide bond, which likely gives resistance against enzymatic hydrolysis. Prolyl residues have also been found at the N-termini of many other peptides described in this paper (for example **Table 2**, peptide family #6; **Table 5**, peptide families #4 and #5). Second, none of the X_1PRX_2 amide peptides are protected by an N-terminal pQ group, while most of the other neuropeptide families (**Tables 2–5**) have such protecting groups, suggesting that the X_1PRX_2 amide peptides need an N-terminal positive charge (by the protonation of the N-terminal primary amine group) for binding to their G-protein-coupled receptor (GPCRs). Third, the R residue at position 3 of the peptide is conserved in all peptides, creating, again, a positive charge in the middle of the peptide and making the overall charge of all peptides in the family quite positive, especially when the residues in position 1 are R or K (**Figure 3**, right panel).

In Octocorallia there is, in addition to the X_1PRX_2 amide peptides (where X_2 is S, A, or G), another peptide family, of which most members have the GPRRamide sequence (**Table 5**, neuropeptide family #5). However, we do not think that these peptides belong to the same family as the X_1PRX_2 amide family, because their preprohormones have a different organization. While in many X_1PRX_2 amide preprohormones, most neuropeptide sequences are followed by a dibasic (KR) processing site [**Supplementary Figures 1, 7, 10**; (31)], there are exclusively single basic (R) processing sites at these positions in the GPRRamide precursor (**Supplementary Figure 14**). Furthermore, in all GPRRamide preprohormones, the neuropeptide sequence is frequently preceded by the sequence DEIT (**Supplementary Figure 14**), whereas this sequence is absent in all X_1PRX_2 preprohormones [**Supplementary Figures 1, 7, 12**; (31)]. We assume, therefore, that the GPRRamide preprohormones might not be evolutionarily closely related to the X_1PRX_2 amide preprohormones.

Besides the GRFamides and X_1PRX_2 amides all other peptide families described in this paper are novel, with exception of the peptides given in **Table 4**, which have been published earlier (31), and we do not really know whether they are localized in neurons and, thus, are genuine neuropeptides. These peptides are not

ubiquitous in cnidarians, but often occur in more than one class. For example, the presumed cyclic peptides from scyphozoans (**Table 2**, neuropeptide family #4) do also occur in cubozoans (**Table 4**, neuropeptide family #4). These results establish the close relationships between scyphozoans and cubozoans, which again is in accordance with current models of the phylogeny of cnidarian classes (**Figure 1**).

Also the pQPPGVWamide peptide family (**Table 2**, neuropeptide family #3; **Table 3**, neuropeptide family #3; **Table 4**, neuropeptide family #3) occurs in scypho-, cubo-, and staurozoans. These results confirm the close phylogenetic relationships between these classes, which is in full agreement with the current models for cnidarian phylogeny (**Figure 1**).

In addition to these peptide families that occur in more than one classes, there are neuropeptides that are confined to a single class (**Table 2**, neuropeptide families #5 and #6; **Table 4**, neuropeptide families #5 and #6; **Table 5**, neuropeptide families #3, #4, #5). These neuropeptides may serve class-specific physiological processes.

None of the above-mentioned peptides identified in the four cnidarian classes/subclasses have significant structural similarities with any of the known bilaterian neuropeptides.

DATA AVAILABILITY STATEMENT

The datasets for *T. cystophora* can be found in the GenBank GGWE01000000.

AUTHOR'S NOTE

This paper is part of the article collection “The Evolution of Neuropeptides: A Stroll through the Animal Kingdom: Updates from the Ottawa 2019 ICCPB Symposium and Beyond” hosted by Dr. Klaus H. Hoffmann and Dr. Elisabeth Amy Williams.

AUTHOR CONTRIBUTIONS

TK and CG conceived and designed the project, and analyzed the data. TK carried out the experiments. CG wrote the paper with inputs from TK. All authors approved the final manuscript.

FUNDING

We thank the Danish Council for Independent Research (grant number 7014-00088 to CG) for financial support. This funding body played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2019.00831/full#supplementary-material>

REFERENCES

- Douzery EJ, Snell EA, Bapteste E, Delsuc F, Philippe H. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc Natl Acad Sci USA*. (2004) 101:15386–91. doi: 10.1073/pnas.0403984101
- Grimmelikhuijzen CJP, Carstensen K, Darmer D, Moosler A, Nothacker H-P, Reinscheid RK, et al. Coelenterate neuropeptides: structure, action and biosynthesis. *Amer Zool*. (1992) 32:1–12. doi: 10.1093/icb/32.1.1
- Grimmelikhuijzen CJP, Leviev I, Carstensen K. Peptides in the nervous systems of cnidarians: structure, function and biosynthesis. *Int Rev Cytol*. (1996) 167:37–89. doi: 10.1016/S0074-7696(08)61345-5
- Grimmelikhuijzen CJP, Williamson M, Hansen GN. Neuropeptides in cnidarians. *Can J Zool*. (2002) 80:1690–702. doi: 10.1139/z02-137
- Grimmelikhuijzen CJP. FMRFamide immunoreactivity is generally occurring in the nervous systems of coelenterates. *Histochemistry*. (1983) 78:361–81. doi: 10.1007/BF00496623
- Smith CL, Varoqueaux F, Kittelmann M, Azzam RN, Cooper B, Winters CA, et al. Novel cell types, neurosecretory cells, and body plan of the early-diverging metazoan *Trichoplax adhaerens*. *Curr Biol*. (2014) 24:1565–72. doi: 10.1016/j.cub.2014.05.046
- Jékely G. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci USA*. (2013) 110:8702–7. doi: 10.1073/pnas.1221833110
- Nikitin M. Bioinformatic prediction of *Trichoplax adhaerens* regulatory peptides. *Gen Comp Endocrinol*. (2015) 212:145–55. doi: 10.1016/j.ygcn.2014.03.049
- Technau U, Steele RE. Evolutionary crossroads in developmental biology: Cnidaria. *Development*. (2011) 138:1447–58. doi: 10.1242/dev.048959
- Zapata F, Goetz FE, Smith SA, Howison M, Siebert S, Church SH, et al. Phylogenomic analysis support traditional relationships within Cnidaria. *PLoS ONE*. (2015) 10:e0139068. doi: 10.1371/journal.pone.0139068
- Kayal E, Bentlage B, Pankey MS, Ohdera AH, Medina M, Plachetzki DC, et al. Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol Biol*. (2018) 18:68. doi: 10.1186/s12862-018-1142-0
- Grimmelikhuijzen CJP, Spencer AN. FMRFamide immunoreactivity in the nervous system of the medusa *Polyorchis penicillatus*. *J Comp Neurol*. (1984) 230:361–71. doi: 10.1002/cne.902300305
- Grimmelikhuijzen CJP. Antisera to the sequence Arg-Phe-amide visualize neuronal centralization in hydroid polyps. *Cell Tissue Res*. (1985) 241:171–82. doi: 10.1007/BF00214639
- Grimmelikhuijzen CJP, Spencer AN, Carré D. Organization of the nervous system of physonectid siphonophores. *Cell Tissue Res*. (1986) 246:463–79. doi: 10.1007/BF00215186
- Takahashi T, Muneoka Y, Lohmann J, Lopez de Haro MS, Solleder G, Bosch TCG, et al. Systematic isolation of peptide signal molecules regulating development in *Hydra*; LWamide and PW families. *Proc Natl Acad Sci USA*. (1997) 94:1241–6. doi: 10.1073/pnas.94.4.1241
- Takahashi T, Takeda N. Insight into the molecular and functional diversity of cnidarian neuropeptides. *Int J Mol Sci*. (2015) 16:2610–25. doi: 10.3390/ijms16022610
- Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, et al. The dynamic genome of *Hydra*. *Nature*. (2010) 464:592–6. doi: 10.1038/nature08830
- Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, et al. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature*. (2011) 476:320–3. doi: 10.1038/nature10249
- Gold DA, Katsuki T, Li Y, Yan X, Reguluski M, Ibberson D, et al. The genome of the jellyfish *Aurelia* and the evolution of animal complexity. *Nat Ecol Evol*. (2019) 3:96–104. doi: 10.1038/s41559-018-0719-8
- Jeon Y, Park SG, Lee N, Weber JA, Kim HS, Hwang SJ, et al. The draft genome of an octocoral, *Dendronephthya gigantea*. *Genome Biol Evol*. (2019) 11:949–53. doi: 10.1093/gbe/evz043
- Leclère L, Horin C, Chevalier S, Lapébie P, Dru P, Peron S, et al. The genome of the jellyfish *Clytia hemisphaerica* and the evolution of the cnidarian life-cycle. *Nat Ecol Evol*. (2019) 3:801–10. doi: 10.1038/s41559-019-0833-2
- Ohdera A, Lewis Ames C, Dikow RB, Kayal E, Chiodin M, Busby B, et al. Box, stalked, and upside-down? Draft genomes from diverse jellyfish (Cnidaria, Acraspeda) lineages: *Alatina alata* (Cubozoa), *Calvadosia cruxmelitensis* (Staurozoa), and *Cassiopea xamachana* (Scyphozoa). *Gigascience*. (2019) 8:giz069. doi: 10.1093/gigascience/giz069
- Khalturin K, Shinzato C, Khalturina M, Hamada M, Fujie M, Koyanagi R, et al. Mudosozoan genomes inform the evolution of the jellyfish body plan. *Nat Ecol Evol*. (2019) 3:811–22. doi: 10.1038/s41559-019-0853-y
- Sunagawa S, Wilson EC, Thaler M, Smith ML, Caruso C, Pringle JR, et al. Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. *BMC Genomics*. (2009) 10:258. doi: 10.1186/1471-2164-10-258
- Stefanik DJ, Lubinski TJ, Granger BR, Byrd AL, Reitzel AM, DeFilippo L, et al. Production of a reference transcriptome and transcriptomic database (EdwardsiellaBase) for the lined sea anemone, *Edwardsiella lineata*, a parasitic cnidarian. *BMC Genomics*. (2014) 15:71. doi: 10.1186/1471-2164-15-71
- Brekman V, Malik A, Haas B, Sher N, Lotan T. Transcriptome profiling of the dynamic life cycle of the scyphozoan jellyfish *Aurelia aurita*. *BMC Genomics*. (2015) 16:74. doi: 10.1186/s12864-015-1320-z
- Kitchen SA, Crowder CM, Poole AZ, Weis VM, Meyer E. *De novo* assembly and characterization of four anthozoan (phylum Cnidaria) transcriptomes. *G3*. (2015) 5:2441–52. doi: 10.1534/g3.115.020164
- Lewis Ames C, Ryan JF, Bely AE, Cartwright P, Collins AG. A new transcriptome and transcriptome profiling of adult and larval tissue in the box jellyfish *Alatina alata*: an emerging model for studying venom, vision, and sex. *BMC Genomics*. (2016) 17:650. doi: 10.1186/s12864-016-3305-y
- Ge J, Liu C, Tan J, Bian L, Chen S. Transcriptome analysis of scyphozoan jellyfish *Rhopilema esculentum* from polyp to medusa identifies potential genes regulating strobilation. *Dev Genes Evol*. (2018) 228:243–54. doi: 10.1007/s00427-018-0621-z
- Veglia AJ, Hammerman NM, Revera-Vicéns RE, Schizas NV. *De novo* transcriptome assembly of the coral *Agaricia lamarcki* (Lamarck's sheet coral) from mesophotic depth in southwest Puerto Rico. *Mar Genomics*. (2018) 41:6–11. doi: 10.1016/j.margen.2018.08.003
- Nielsen SKD, Koch TL, Hauser F, Garm A, Grimmelikhuijzen CJP. *De novo* transcriptome assembly of the cubomedusa *Tripedalia cystophora*, including the analysis of a set of genes involved in peptidergic neurotransmission. *BMC Genomics*. (2019) 20:175. doi: 10.1186/s12864-019-5514-7
- Rivera-García L, Rivera-Vicéns RE, Veglia AJ, Schizas NV. *De novo* transcriptome assembly of the digitate morphotype of *Briareum asbestinum* (Octocorallia: Alcyonacea) from the southwest shelf of Puerto Rico. *Mar Genomics*. (2019) 47:100676. doi: 10.1016/j.margen.2019.04.001
- Wang C, Wang B, Wang Q, Liu G, Wang T, et al. Unique diversity of sting-related toxins based on transcriptomic and proteomic analysis of the jellyfish *Cyanea capillata* and *Nemopilema nomurai* (Cnidaria: Scyphozoa). *J Proteome Res*. (2019) 18:436–48. doi: 10.1021/acs.jproteome.8b00735
- Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. (2019) 37:420–3. doi: 10.1038/s41587-019-0036-z
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. (2007) 23:2947–8. doi: 10.1093/bioinformatics/btm404
- Grimmelikhuijzen CJP, Graff D. Isolation of <Glu-Gly-Arg-Phe-NH₂> (Antho-RFamide), a neuropeptide from sea anemones. *Proc Natl Acad Sci USA*. (1986) 83:9817–21. doi: 10.1073/pnas.83.24.9817
- Grimmelikhuijzen CJP, Groeger A. Isolation of the neuropeptide pGlu-Gly-Arg-Phe-amide from the pennatulid *Renilla köllikeri*. *FEBS Lett*. (1987) 211:105–8. doi: 10.1016/0014-5793(87)81283-8
- Darmer D, Schmutzler C, Diekhoff D, Grimmelikhuijzen CJP. Primary structure of the precursor for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH₂>). *Proc Natl Acad Sci USA*. (1991) 88:2555–9. doi: 10.1073/pnas.88.6.2555
- Schmutzler C, Darmer D, Diekhoff D, Grimmelikhuijzen CJP. Identification of a novel type of processing sites in the precursor for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH₂>) from *Anthopleura elegantissima*. *J Biol Chem*. (1992) 267:22534–41.

40. Pernet V, Anctil M, Grimmelikhuijzen CJP. Antho-RFamide-containing neurons in the primitive nervous system of the anthozoan *Renilla koellikeri*. *J Comp Neurol.* (2004) 472:208–20. doi: 10.1002/cne.20108
41. Moosler A, Rinehardt KL, Grimmelikhuijzen CJP. Isolation of three novel neuropeptides, the Cyanea-RFamides I-III, from scyphomedusae. *Biochem Biophys Res Commun.* (1997) 236:743–9. doi: 10.1006/bbrc.1997.7022
42. Anderson PAV, Moosler A, Grimmelikhuijzen CJP. The presence and distribution of Antho-RFamide-like material in scyphomedusae. *Cell Tissue Res.* (1992) 267:67–74. doi: 10.1007/BF00318692
43. Hayakawa E, Watanabe H, Menschaert G, Holstein TW, Baggerman G, Schoofs L. A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*. *PLoS ONE.* (2019) 14:e0215185. doi: 10.1371/journal.pone.0215185
44. Takeda N, Kon Y, Quiroga Artigas G, Lapébie P, Barreau C, Koizumi O, et al. Identification of jellyfish neuropeptides that act directly as oocyte maturation-inducing hormones. *Development.* (2018) 145. doi: 10.1242/dev.156786

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Koch and Grimmelikhuijzen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Evolution and Comparative Physiology of Luqin-Type Neuropeptide Signaling

Luis Alfonso Yañez-Guerra**† and Maurice R. Elphick*

School of Biological and Chemical Sciences, Faculty of Science and Engineering, Queen Mary University of London, London, United Kingdom

OPEN ACCESS

Edited by:

Klaus H. Hoffmann,
University of Bayreuth, Germany

Reviewed by:

Lindy Holden-Dye,
University of Southampton,
United Kingdom
Liliane Schoofs,
KU Leuven, Belgium

*Correspondence:

Luis Alfonso Yañez-Guerra
L.Yanez-Guerra@exeter.ac.uk
Maurice R. Elphick
m.r.elphick@qmul.ac.uk

† Present address:

Luis Alfonso Yañez-Guerra,
Living Systems Institute, University
of Exeter, Exeter, United Kingdom

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Neuroscience

Received: 22 October 2019

Accepted: 31 January 2020

Published: 18 February 2020

Citation:

Yañez-Guerra LA and Elphick MR
(2020) Evolution and Comparative
Physiology of Luqin-Type
Neuropeptide Signaling.
Front. Neurosci. 14:130.
doi: 10.3389/fnins.2020.00130

Luqin is a neuropeptide that was discovered and named on account of its expression in left upper quadrant cells of the abdominal ganglion in the mollusc *Aplysia californica*. Subsequently, luqin-type peptides were identified as cardio-excitatory neuropeptides in other molluscs and a cognate receptor was discovered in the pond snail *Lymnaea stagnalis*. Phylogenetic analyses have revealed that orthologs of molluscan luqin-type neuropeptides occur in other phyla; these include neuropeptides in ecdysozoans (arthropods, nematodes) that have a C-terminal RYamide motif (RYamides) and neuropeptides in ambulacrarians (echinoderms, hemichordates) that have a C-terminal RWamide motif (RWamides). Furthermore, precursors of luqin-type neuropeptides typically have a conserved C-terminal motif containing two cysteine residues, although the functional significance of this is unknown. Consistent with the orthology of the neuropeptides and their precursors, phylogenetic and pharmacological studies have revealed that orthologous G-protein coupled receptors (GPCRs) mediate effects of luqin-type neuropeptides in spiralian, ecdysozoans, and ambulacrarians. Luqin-type signaling originated in a common ancestor of the Bilateria as a paralog of tachykinin-type signaling but, unlike tachykinin-type signaling, luqin-type signaling was lost in chordates. This may largely explain why luqin-type signaling has received less attention than many other neuropeptide signaling systems. However, insights into the physiological actions of luqin-type neuropeptides (RYamides) in ecdysozoans have been reported recently, with roles in regulation of feeding and diuresis revealed in insects and roles in regulation of feeding, egg laying, locomotion, and lifespan revealed in the nematode *Caenorhabditis elegans*. Furthermore, characterization of a luqin-type neuropeptide in the starfish *Asterias rubens* (phylum Echinodermata) has provided the first insights into the physiological roles of luqin-type signaling in a deuterostome. In conclusion, although luqin was discovered in *Aplysia* over 30 years ago, there is still much to be learnt about luqin-type neuropeptide signaling. This will be facilitated in the post-genomic era by the emerging opportunities for experimental studies on a variety of invertebrate taxa.

Keywords: luqin, cardio-excitatory peptide, RYamides, RWamides, neuropeptide evolution, G-protein coupled receptors

INTRODUCTION

Neuropeptides are evolutionarily ancient neuronal signaling molecules that typically exert their effects on target cells by binding to cognate G-protein coupled receptors (GPCRs) (Elphick et al., 2018; Jékely et al., 2018). Phylogenetic studies have revealed that the evolutionary origin of at least 30 neuropeptide signaling systems can be traced back to the bilaterian common ancestor of deuterostomes and protostomes. However, some neuropeptide signaling systems have been lost in one or more bilaterian phyla/sub-phyla (Jékely, 2013; Mirabeau and Joly, 2013; Elphick et al., 2018). Luqin-type neuropeptide signaling, which is the focus of this review, is one of the bilaterian neuropeptide signaling systems that have been lost in chordates/vertebrates. This may in part explain why less is known about luqin-type neuropeptide signaling than other bilaterian neuropeptide systems that have been retained in vertebrates. Nevertheless, several advances in our knowledge of the evolution and comparative physiology of luqin-type signaling in invertebrates have been made recently and therefore writing of this the first review article on luqin-type neuropeptide signaling is timely.

The neuropeptide luqin and its cognate GPCR were first discovered in molluscs (Shyamala et al., 1986; Fujimoto et al., 1990; Tensen et al., 1998). Subsequently, luqin-like neuropeptides known as RYamides, on account of a C-terminal Arg-Tyr-NH₂ motif, were discovered in the arthropod *Cancer borealis* (Li et al., 2003). Furthermore, receptors for RYamides were identified in the fruitfly *Drosophila melanogaster* and in the red flour beetle *Tribolium castaneum* (Collin et al., 2011; Ida et al., 2011). In 2013, evidence that molluscan luqin-type signaling and arthropod RYamide-type signaling are orthologous was reported. Thus, use of pairwise-based clustering methods revealed that luqin precursors and RYamide precursors form part of the same protein cluster (Jékely, 2013). Furthermore, phylogenetic analysis of G-protein coupled neuropeptide receptors revealed that molluscan luqin receptors are orthologs of arthropod RYamide receptors (Mirabeau and Joly, 2013). In addition, these two studies also reported for the first time the discovery of a luqin-type signaling system in the nematode *Caenorhabditis elegans*, which has subsequently been functionally characterized experimentally (Jékely, 2013; Mirabeau and Joly, 2013; Ohno et al., 2017).

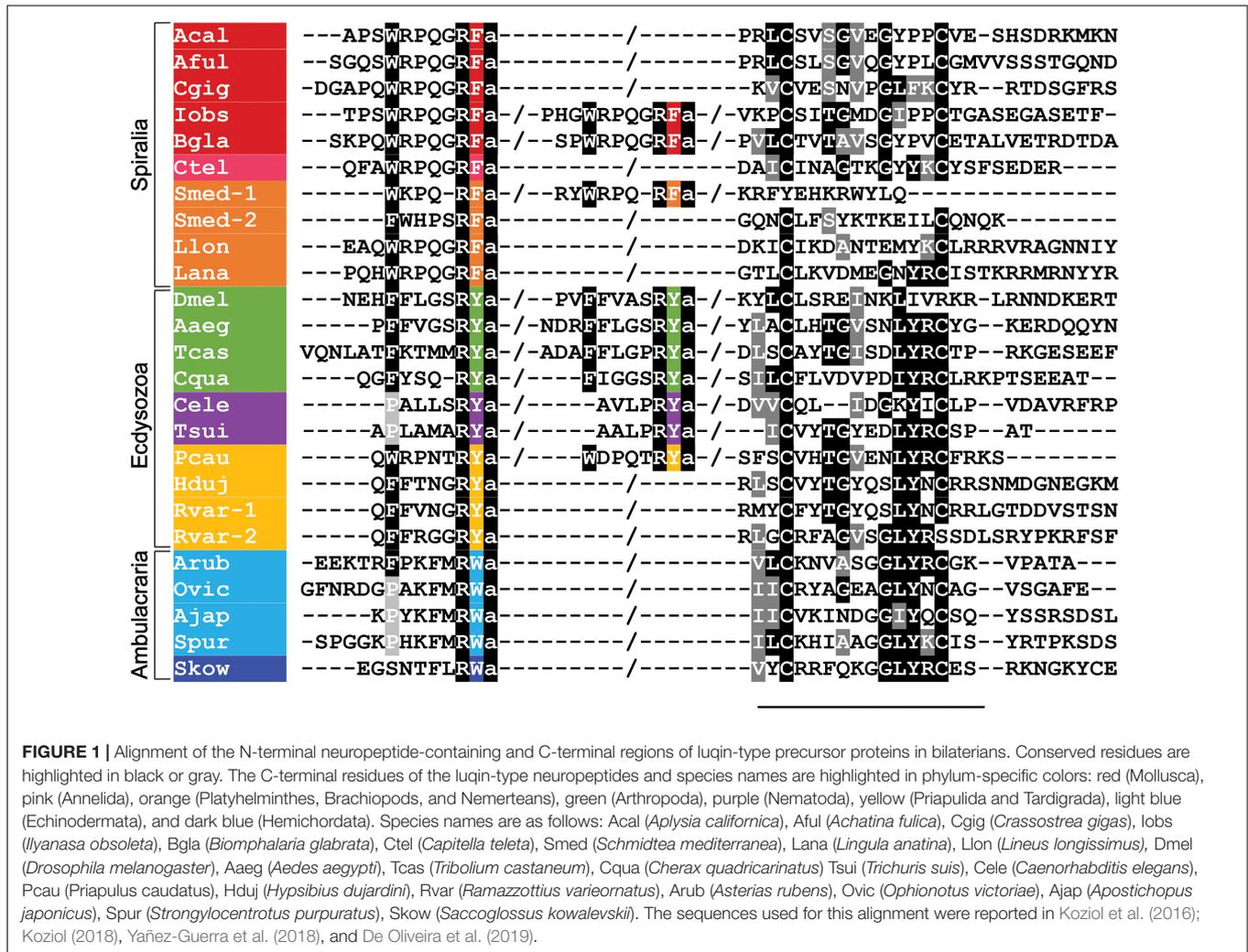
Importantly, phylogenetic analysis of transcriptome/genome sequence data has revealed that while luqin-type signaling has been lost in vertebrates and other chordates (urochordates, cephalochordates), genes encoding luqin-type precursors and receptors are present in ambulacrarian deuterostomes (echinoderms and hemichordates) (Jékely, 2013; Mirabeau and Joly, 2013). Thus, it was established for the first time that the evolutionary origin of luqin-type neuropeptide signaling predates the divergence of protostomes and deuterostomes, but with differential loss in the deuterostome branch of the Bilateria. Luqin-type neuropeptides in ambulacrarians are characterized by a predicted C-terminal RWamide motif (Elphick and Mirabeau, 2014; Rowe et al., 2014; Semmens et al., 2016). Furthermore, biochemical characterization of luqin-type neuropeptide signaling in the starfish *Asterias rubens* (phylum Echinodermata)

demonstrated that a neuropeptide with the confirmed structure EKGRFPKFMRW-NH₂ acts as a ligand for two luqin-type receptors in this species (Yañez-Guerra et al., 2018). Thus, the luqin-type neuropeptides in spiralian protostomes, the luqin-type RYamides in ecdysozoan protostomes, and the luqin-type RWamides in ambulacrarian deuterostomes have been unified as members of a bilaterian family of neuropeptides (Jékely, 2013; Mirabeau and Joly, 2013; Yañez-Guerra et al., 2018). It is against the backdrop of these important recent findings that we review here our knowledge of the evolution and comparative physiology of luqin-type neuropeptide signaling.

THE DISCOVERY AND FUNCTIONAL CHARACTERIZATION OF LUQIN-TYPE NEUROPEPTIDE SIGNALING IN MOLLUSCS

The discovery of the neuropeptide luqin was enabled by identification of transcripts expressed in the L5 neuron of the mollusc *Aplysia californica*. This was decades before the development of contemporary single cell transcriptomic methodologies and was facilitated by the large size of the cell body of the L5 neuron (Shyamala et al., 1986). Analysis of the expression of the cardio-excitatory neuropeptide FMRFamide in the central nervous system of *A. californica* using antibodies to FMRFamide revealed immunostaining in many neurons, including the L5 neuron that is located in the upper quadrant of the left abdominal ganglion (Brown et al., 1985; Schaefer et al., 1985). However, it was found that the FMRFamide gene is not expressed in the L5 neuron. Therefore, to determine the identity of the neuropeptide(s) responsible for FMRFamide-like immunoreactivity in L5, transcripts expressed in this neuron were sequenced. Sequencing of a transcript named L5-67 revealed that it encodes a 112 amino acid residue protein comprising an N-terminal signal peptide followed by a neuropeptide with a predicted C-terminal RFamide motif, which therefore could be cross-reactive with FMRFamide antibodies (Shyamala et al., 1986). Subsequent biochemical analysis revealed that processing of this precursor gives rise to the amidated decapeptide APSWRPQGRF-NH₂, which was named luqin because it is expressed in the left upper quadrant cells of the abdominal ganglion (Aloyz and DesGroseillers, 1995). Furthermore, a 76 amino acid peptide corresponding to the region of the precursor protein C-terminal to the luqin neuropeptide was also detected and named proline-rich mature peptide (PRMP) (Aloyz and DesGroseillers, 1995). PRMP contains two cysteines separated by a 10 amino acid-residue sequence and subsequent studies have revealed that this is an evolutionarily conserved feature of luqin-type precursor proteins (**Figure 1**) (Jékely, 2013; Mirabeau and Joly, 2013; Yañez-Guerra et al., 2018).

In addition to luqin and PRMP, a mass spectrometric survey of the left upper quadrant neurons revealed that two other peptides are derived from the *A. californica* luqin precursor, which were named luqin-B and luqin-C (Li et al., 1998). The luqin-B fragment contains part of the mature luqin neuropeptide



and the luqin-C fragment contains a shorter version of PRMP (Li et al., 1998), but it is not known if PRMP, luqin-B, or luqin-C are biologically active molecules. However, it has been shown that alternative splicing by exon skipping of one of the exons of the gene encoding the luqin precursor results in a frame shift and production of a precursor protein comprising the complete mature amidated decapeptide luqin and a short C-terminal region that does not contain the two cysteines characteristic of PRMP. This suggests that PRMP may not be essential for biosynthesis of the mature luqin neuropeptide. Still, from a functional perspective, it is noteworthy that while the full-length transcript is widely expressed in *A. californica*, the alternative transcript is specifically expressed in the kidney (Angers and DesGroseillers, 1998).

Sequencing of transcripts encoding luqin-type precursors and mass spectrometric identification of neuropeptides derived from them has revealed a high level of sequence conservation of luqin-type neuropeptides in molluscs. For example, in the giant snail *Achatina fullica*, a luqin-type peptide was identified as an amidated undecapeptide SGQSWRPQGRF-NH₂, the C-terminal region of which (underlined) is identical to

Aplysia luqin (Fujimoto et al., 1990), and in the pond snail *Lymnaea stagnalis* the luqin-type peptide TPHWRPQGRF-NH₂ was identified. It is noteworthy that the luqin precursors in *A. californica*, *L. stagnalis*, and *A. fullica* comprise a single luqin-type neuropeptide, whereas in some other gastropod molluscs, such as the eastern mudsnail *Ilyanassa obsoleta* and the freshwater snail *Biomphalaria glabrata*, the precursor comprises two luqin-type neuropeptides (Figure 1). The first insight into the characteristics of luqin receptors was made with the discovery that the luqin-type peptide TPHWRPQGRF-NH₂ is a ligand for an orphan G protein-coupled receptor, GRL106, in the pond snail *L. stagnalis*. It was also reported that this receptor is closely related to vertebrate tachykinin receptors and the *Drosophila* neuropeptide-Y-type receptor (Tensen et al., 1998).

Insights into the physiological roles of luqin-type neuropeptides were facilitated by the identification of a luqin-type neuropeptide in the snail *A. fullica*—a pulmonate gastropod. The peptide was discovered on account of its bioactivity as a cardioactive peptide that causes an increase in the frequency of beating when applied to auricle preparations from *A. fullica*. Hence, this peptide was named *Achatina* cardioexcitatory peptide

(ACEP) (Fujimoto et al., 1990). However, ACEP also affects other muscle systems in *A. fullica*, increasing the amplitude of tetanic contraction of the penis retractor muscle and buccal muscle in response to electrical stimulation. Furthermore, ACEP was found to induce depolarization and rhythmic firing of a motor neuron (B4) that innervates buccal muscle (Fujimoto et al., 1990), an effect indicative of a physiological role in regulation of feeding behavior. Consistent with findings from *A. fullica*, the luqin-type peptide identified as a ligand for the *Lymnaea* receptor GRL106 was also found to be cardioexcitatory in this pulmonate gastropod species. Hence, the peptide was named *Lymnaea* cardioexcitatory peptide (LyCEP). Furthermore, the sequence similarity that LyCEP and ACEP share with *Aplysia* luqin was noted (Tensen et al., 1998). In accordance with the cardioexcitatory actions of LyCEP, immunohistochemical analysis revealed that LyCEP-immunoreactivity is present in nerve fibers ending in the pericardial cavity of the heart, indicating that the peptide is released into the pericardial cavity as a neurohormone (Tensen et al., 1998). However, LyCEP is not, as its name implies, specifically a cardioactive peptide because it is also expressed in nerve fibers associated with inhibition of the egg-laying hormone-producing caudodorsal cells in *L. stagnalis*. Accordingly, transcripts encoding the LyCEP receptor are present in the caudodorsal cells and LyCEP causes hyperpolarization of these cells *in vitro* (Tensen et al., 1998).

Analysis of the expression of the luqin precursor in the ophistobranch gastropod *A. californica* revealed transcripts in approximately 100 neurons of the central nervous system, but predominantly in neurons that innervate the circulatory and reproductive systems. In peripheral tissues, transcripts were detected in the intestine and in the kidneys (Giardino et al., 1996). Whole mount immunolabeling experiments with an antibody directed against the luqin precursor revealed immunoreactive fibers in different regions of the circulatory system of *Aplysia*, including the auricle, the ventricle, and the aorta. In the reproductive system, immunoreactive fibers were detected in the small and large hermaphroditic ducts and in the ovotestis (Giardino et al., 1996). The kidney also displayed immunoreactivity, located on the inner surface of the kidney wall. Strong immunoreactivity was also seen in neurites located in a large nerve associated with muscles of the renal pore, a sphincter that controls urine efflux (Angers et al., 2000). Altogether, these findings suggest roles of luqin-type peptides in regulation of the reproductive and circulatory systems, in fluid mobilization, and water homeostasis in gastropod molluscs.

While what is known about luqin expression and function in molluscs is largely based on experimental studies on selected gastropod species, as detailed above, genes/transcripts encoding luqin-type precursors have been identified in other gastropod species and in species belonging to other molluscan classes, including Bivalvia, Scaphopoda, Cephalopoda, Monoplacophora, Polyplacophora, Chaetodermomorpha, and Neomeniomorpha (De Oliveira et al., 2019). Thus, the occurrence of luqin-type neuropeptide precursors throughout the phylum Mollusca has been established, providing a basis for extending analysis of luqin-type neuropeptide function beyond gastropods to other classes.

LUQIN-TYPE SIGNALING IN ANNELIDS

Luqin-type precursors with a similar organization to those in molluscs have been identified in annelids. Analysis of genomic sequence data from *Capitella teleta* revealed the occurrence of a luqin-type precursor comprising the predicted mature peptide QFAWRPQGRF-NH₂ (Veenstra, 2011) (**Figure 1**). Later, a partial precursor transcript was identified in the annelid *Platynereis dumerilii* and the structure of the luqin-type peptide derived from this precursor was confirmed as WRPQGRF-NH₂ using mass spectrometry (Conzelmann et al., 2013). Furthermore, a transcript encoding a luqin-type receptor was identified in *P. dumerilii* and pharmacological characterization of this receptor demonstrated that the luqin-type peptide WRPQGRF-NH₂ acts as a ligand for this receptor (Bauknecht and Jékely, 2015). Currently, there are no data available that provide insights into the physiological roles of luqin-type neuropeptides in annelids.

LUQIN-TYPE SIGNALING IN OTHER SPIRALIANS

Analysis of genome/transcriptome sequence data has revealed the occurrence of genes/transcripts encoding luqin-type precursors in a variety of species belonging to the phylum Platyhelminthes. For example, in the planarian *Schmidtea mediterranea*, an expanded family of genes encoding four luqin-type precursors has been identified (Koziol et al., 2016). The predicted neuropeptides derived from these precursors share sequence similarities with luqin-type neuropeptides from molluscs and annelids, including a C-terminal RFamide motif and the N-terminal motif WRPQ, which is conserved with only conservative substitutions (**Figure 1**). Interestingly, only two of the four precursors have a C-terminal pair of cysteines separated by 10 amino acids, which is typically a conserved feature of luqin-type precursors in other phyla (Koziol et al., 2016). The functional significance of the loss of this feature in two precursors is unknown, but it may be reflective of a loss of selection pressure after gene duplications gave rise to the four precursor genes.

Genes encoding luqin-type precursors have also been identified in parasitic platyhelminths, including the fox tapeworm *Echinococcus multilocularis*, the salmon fluke *Gyrodactylus salaris*, the rodent tapeworm *Hymenolepis microstoma*, the broad fish tapeworm *Diphyllobothrium latum*, and the cestode *Mesocestoides corti* (Koziol et al., 2016). In all these species, a single luqin-type precursor comprising a single predicted neuropeptide was identified, with the peptides having the N-terminal motif WRPQ that is similar to the WRPQ motif that is a feature of luqin-type neuropeptides in molluscs and annelids. Interestingly, however, the luqin-type peptides in these species are positioned in the C-terminal region of the precursor protein. This contrasts with precursors in other taxa, where luqin-type peptides are located N-terminally and proximal to the signal peptide. Furthermore, none of the luqin-type precursors in the parasitic platyhelminth species listed above have a C-terminal pair of cysteines separated by 10 amino

acids. Thus, the luqin-type precursors identified in parasitic platyhelminthes are highly divergent by comparison with other luqin-type precursors. This makes sequence alignment difficult and for this reason luqin-type precursors from parasitic platyhelminths are not included in **Figure 1** but instead they are shown in **Supplementary Figure 1B**. Nevertheless, evidence that luqin-type neuropeptides in parasitic platyhelminths are functional can be found in the identification of genes encoding candidate luqin-type receptors; for example, in *E. multilocularis* (Kozioł et al., 2016). Furthermore, a comprehensive analysis of GPCRs encoded in the genome of the parasitic helminth *Fasciola hepatica* has revealed the presence of a receptor that is clearly an ortholog of the luqin-type receptor from the annelid *P. dumerilii* and the RYamide receptor from *D. melanogaster* (McVeigh et al., 2018). Further studies are now needed to confirm the predicted luqin-type ligand-receptor partners in platyhelminths and to investigate the expression and pharmacological actions of luqin-type neuropeptides in platyhelminths.

Luqin-type precursors have also been identified in the brachiopod *Lingula anatina* and in the nemertean *Lineus longissimus* (bootlace worm) (De Oliveira et al., 2019). In both species, a single luqin-type precursor was identified that contains a predicted mature neuropeptide with the C-terminal sequence WRPQGRF-NH₂, which is the same motif identified in molluscan and annelid luqin-type peptides. The precursors identified in these species also have the typical C-terminal region containing the two cysteines separated by 10 amino acid residues (**Figure 1**). Furthermore, two proteins in *L. anatina* have been annotated as luqin-type receptors (XP_013402794.1, XP_013402807.1).

RYamides: LUQIN-TYPE NEUROPEPTIDES IN ARTHROPODS

Arthropodan neuropeptides with a C-terminal RYamide motif (RYamides) were first identified in the decapod *C. borealis* by *de novo* post source decay sequencing of peptides in extracts of the pericardial organs of this species. Five different peptides were identified, all of them sharing the conserved C-terminal motif FXXXRY-NH₂, where X is variable (Li et al., 2003). RYamides sharing the same C-terminal motif were subsequently identified by analysis of tissue extracts from other decapods, including *Cancer productus* (Fu et al., 2005), *Pugettia producta* (Stemmler et al., 2007), *Carcinus maenas* (Ma et al., 2009), and in the Pacific white shrimp *Litopenaeus vannamei* (Ma et al., 2010).

Genes/transcripts encoding precursors of RYamides have been identified in several crustacean species, including the water flea *Daphnia pulex* (Dirksen et al., 2011), the isopod *Proasellus cavaticus* (Christie, 2017), the red swamp crayfish *Procambarus clarkii* (Veenstra, 2015), the Australian crayfish *Cherax quadricarinatus* (Nguyen et al., 2016), and the freshwater amphipod *Hyalella azteca* (Christie et al., 2018). Interestingly, while most of these precursors comprise two RYamides, the *D. pulex* and *P. clarkii* precursors comprise three predicted RYamides (**Supplementary Figure 1A**). In the case of the *D. pulex* RYamide precursor, mass spectrometric analysis of brain tissue enabled identification of two of the three

RYamides predicted to be derived from this precursor. The first peptide, located immediately after the signal peptide, was identified in its post-translationally modified form as pQTFFFTNGRY-NH₂, with both C-terminal amidation and N-terminal conversion of a glutamine residue to pyroglutamate. The second RYamide predicted to be derived from this precursor was not detected by mass spectrometry. Two different forms of the third RYamide peptide were detected—the 27 residue peptide SGNGGIVLGNSELDARNPERFFIGSRY-NH₂ and a C-terminal fragment of this peptide (NPERFFIGSRY-NH₂) generated by cleavage at the underlined arginine residue in the longer peptide (Dirksen et al., 2011).

Genes/transcripts encoding precursors of RYamides have also been identified in a variety of insects, including six *Drosophila* species, the red flour beetle *T. castaneum*, the silkworm *Bombyx mori*, the honey bee *Apis mellifera*, the pea aphid *Acyrtosiphon pisum*, the yellow fever mosquitoes *Aedes aegypti* and *Culex pipiens*, and the stick insect *Carausius morosus* (Hauser et al., 2010; Liessem et al., 2018), and these typically comprise two predicted RYamides. Interestingly, in the RYamide precursor of the parasitic wasp *Nasonia vitripennis* paracopy expansion has occurred to give rise to a precursor comprising seven predicted RYamide-type neuropeptides (**Supplementary Figure 1A**). However, only one of these peptides has thus far been characterized biochemically by mass spectrometry (Hauser et al., 2010).

The first arthropod RYamide receptors to be characterized were from *T. castaneum* and *D. melanogaster*. In the case of the red flour beetle *T. castaneum*, a receptor was identified (GenBank Accession Number: HQ709383) and shown to be activated in a dose-dependent manner by the two RYamide peptides derived from the *T. castaneum* RYamide precursor (Collin et al., 2011). In *D. melanogaster*, the RYamide receptor was characterized independently by two laboratories, revealing that the two RYamides derived from the *D. melanogaster* RYamide precursor activate, in a dose-dependent manner, a receptor encoded by the gene CG5811 (Collin et al., 2011; Ida et al., 2011). Ida et al. (2011) also reported that injection of RYamide-1 suppresses the proboscis extension reflex (PER) in the blowfly *Phormia regina*, indicating a physiological role in regulation of feeding behavior (Ida et al., 2011). Subsequent experimental studies on this species revealed the presence of 26 RYamide-immunoreactive neurons in the brain and showed that injection of RYamide-1 or -2 has no effect on the volume of sucrose solution intake when feeding occurs but causes a reduction in the percentage of flies exhibiting the PER. Furthermore, injection of RYamide-1 was found to cause a significant decrease in the responsiveness to sucrose solution of sugar receptor neurons located on the labellum of the proboscis. Thus, it was concluded that RYamides suppress feeding motivation and sucrose responsiveness in the blow fly *P. regina* (Maeda et al., 2015). Analysis of the expression of the RYamide precursor in the silkworm *B. mori* using mRNA *in situ* hybridization revealed expression in the brain, terminal abdominal ganglion, and midgut. In the larval and adult brain, four to seven pairs of RYamide precursor-expressing neurons were identified in the protocerebrum and tritocerebrum. Expression was also revealed in a pair of posterior dorsomedial neurons in the terminal abdominal ganglion of adults and

larvae. Lastly, RYamide precursor expression was revealed in enteroendocrine cells of the anterior midgut of larvae, pupae, and adult specimens of *B. mori*. This pattern of expression indicates that RYamides are involved in regulation of feeding and digestion in *B. mori* (Roller et al., 2016).

Consistent with the hypothesis that RYamides may also regulate feeding behavior in other arthropods, use of quantitative real-time PCR revealed that expression of the RYamide precursor gene is significantly downregulated in the brain after starvation in the decapod *Marsupenaeus japonicus* (kuruma shrimp). Furthermore, injection of RYamides into the muscle of juvenile *M. japonicus* caused suppression of food intake in some experiments, but this was not consistently reproducible (Mekata et al., 2017). Foregut activity in decapods is controlled by motor neurons in the stomatogastric ganglion (STG), the activity of which is regulated by a variety of different neuropeptide types (Marder and Bucher, 2007). To gain further insights into the complexity of neuropeptide signaling in the STG, transcriptomic analysis of the STG in the crab *C. borealis* revealed the presence of 46 transcripts encoding receptors for 27 different neuropeptide types, but interestingly RYamide receptor transcripts were not detected. Consistent with this finding, *in vitro* application of RYamides to *C. borealis* STG preparations had no consistent modulatory effect on the motor outputs of the ganglion (Dickinson et al., 2019). Thus, further studies are now required to investigate the mechanisms by which RYamides may affect feeding in decapod crustaceans.

RYamides are not only involved in regulation of feeding behavior in arthropods. Analysis of RYamide expression in several *Drosophila* species revealed that the RYamide precursor gene is expressed in two abdominal neurons of the adult central nervous system that project to the rectal papillae, organs that mediate water re-absorption in flies (Veenstra and Khammassi, 2017). Consistent with this expression pattern, injection of female mosquitoes with RYamides delays postprandial diuresis (Veenstra and Khammassi, 2017). Thus, it appears that RYamides may act as regulators of urine production in insects. Interestingly, this is consistent with immunohistochemical evidence of a similar role in the mollusc *Aplysia*, where luqin-immunoreactivity has been localized in a nerve associated with muscles of the renal pore, a sphincter that controls urine efflux (Angers et al., 2000) (see above).

DETAILED FUNCTIONAL CHARACTERIZATION OF LUQIN-TYPE SIGNALING IN THE NEMATODE *Caenorhabditis elegans*

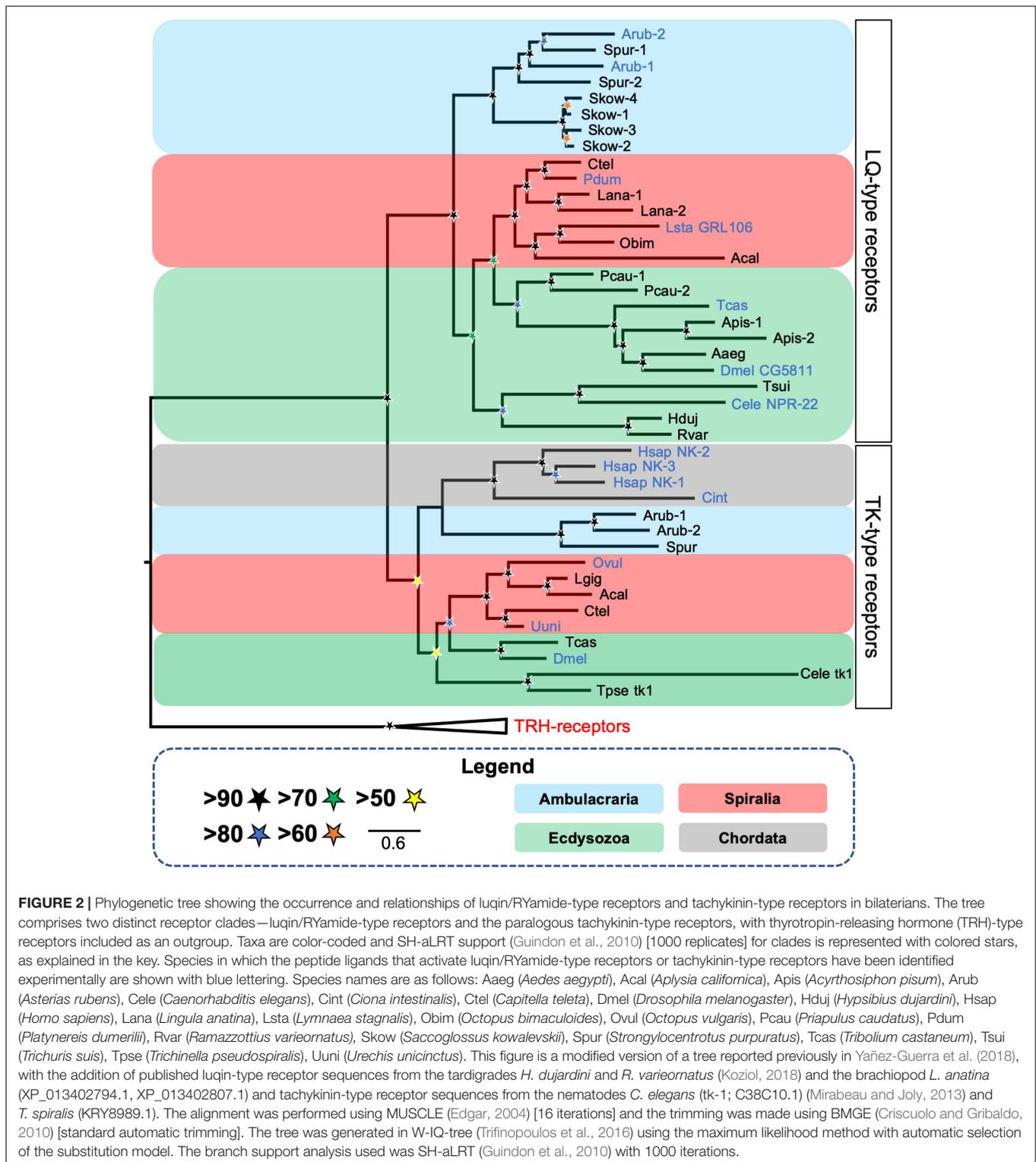
In 2003, Keating et al. reported an extensive functional analysis of GPCRs in the nematode *C. elegans*, employing use of RNA interference (RNAi) methods to knockdown GPCR gene expression. Included in this study was the gene *Y59H11AL.1* (also now known as *npr-22*) and a phylogenetic analysis revealed that the protein encoded by *Y59H11AL.1* is most closely related to the *L. stagnalis* luqin receptor GRL106. However, both the receptor encoded by *Y59H11AL.1* and GRL106 were classified

by the authors under the general descriptor of “tachykinin-like receptors” (Keating et al., 2003). Subsequently, efforts were made to identify the neuropeptide(s) that act as a ligand for the *Y59H11AL.1*-encoded GPCR (NPR-22) by screening a library of synthetic neuropeptides predicted from analysis of the sequences of *C. elegans* neuropeptide precursors (Mertens et al., 2006). It was discovered that FMRFamide related peptides derived from the FLP-7 precursor can activate NPR-22, with the peptide FLP-7.3 (SPMERSAMVRF-NH₂) being the most potent, albeit with an EC₅₀ of ~1 μM (Mertens et al., 2006). The authors also noted that NPR-22 (*Y59H11AL.1*) is closely related to the *D. melanogaster* receptor CG5811. However, at the time the peptides that act as ligands for CG5811 were unknown and it was not until 5 years later that it was discovered that CG5811 is a RYamide receptor (Collin et al., 2011; Ida et al., 2011).

In 2013, an extensive phylogenetic analysis of GPCRs and neuropeptide precursors identified a luqin-type precursor in *C. elegans* (Mirabeau and Joly, 2013). Furthermore, alignment of the *C. elegans* precursor with luqin-type precursors from other taxa revealed many conserved features. Thus, the precursor comprises two putative luqin-type peptides that have an RYamide motif and the C-terminal region of the precursor contains two cysteines, which are separated by eight amino acid residues (Mirabeau and Joly, 2013). The spacing of the two cysteines residues in the *C. elegans* luqin-type precursor is atypical of luqin-type precursors where, as highlighted above and shown in **Figure 1**, the two cysteine residues are usually separated by ten residues. However, this feature of the *C. elegans* luqin-type precursor is probably a derived characteristic because in another nematode species, the parasite *Trichuris suis*, the luqin-type precursor has two C-terminal cysteines separated by 10 residues (**Figure 1**) (Yañez-Guerra et al., 2018).

Recently, a detailed analysis of the phylogenetic relationships of luqin-type receptors revealed that the *C. elegans* receptor NPR-22 (*Y59H11AL.1*) is an ortholog of the *L. stagnalis* luqin receptor (GRL106) and the *D. melanogaster* receptor CG5811 and belongs to a clade of luqin-type receptors that is quite distinct from the closely related tachykinin-type receptors and neuropeptide-Y-type receptors (Yañez-Guerra et al., 2018) (**Figure 2**). Furthermore, the neuropeptides LURY-1.1 (PALLSRY-NH₂) and LURY-1.2 (AVLPRY-NH₂) derived from the *C. elegans* luqin/RYamide precursor (LURY-1) were shown to act as ligands for NPR-22 at nanomolar concentrations (Ohno et al., 2017). Additionally, the authors showed that, as reported previously by Mertens et al. (2006), the FLP-7.3 peptide (SPMERSAMVRF-NH₂) derived from the FLP-7 precursor also acts as a ligand for NPR-22, but only at micromolar concentrations (Ohno et al., 2017). Therefore, it is likely that LURY-1.1 and LURY-1.2 are the natural ligands for NPR-22 and the ability of FLP-7 derived-peptides to activate NPR-22 at micromolar concentrations may be an *in vitro* pharmacological and non-physiological phenomenon due to the presence of a C-terminal RYamide-like RFamide motif in these peptides. However, evidence that NPR-22 mediates effects of FLP-7 derived peptides *in vivo* has been reported in an investigation of neuroendocrine mechanisms of serotonin-induced fat loss in *C. elegans* (Palamiuc et al., 2017).

Having identified the molecular components of a luqin/RYamide-type neuropeptide signaling system in



C. elegans, Ohno et al. (2017) performed a detailed functional characterization of this signaling system. Analysis of the expression of the LURY-1 precursor and NPR-22 in *C. elegans* revealed that LURY-1 is expressed by the pharyngeal neurons M1 and M2, which regulate feeding in *C. elegans*. Thus, the M1

neuron stimulates spitting behavior whereas the M2 neuron stimulates pharyngeal pumping (Bhatla et al., 2015). With a much wider pattern of expression, NPR-22 is expressed in head muscles, I2 pharyngeal neurons, feeding pacemaker MC neurons, the RIH neuron, the interneurons AIA and AUA, the

ASK neurons, the ASI neurons, a few B-type motoneurons in the posterior ventral nerve cord, pharyngeal muscles, body wall muscles, the intestine, and a few unidentified cells anterior to the nerve ring (Ohno et al., 2017; Palamiuc et al., 2017).

To gain insights into the physiological roles of LURY-1-derived peptides, the *lury-1* gene was overexpressed in *C. elegans*, with several phenotypes being observed. First, there was an increase in the number of unlaidd eggs in the uterus but the rate of egg-laying was not affected, indicating that the rate of ovulation is normal but egg-laying is facilitated and embryos are laid prematurely. Accordingly, microinjection of synthetic LURY-1 peptides (10 μ M) caused the same phenotype. Second, pharyngeal pumping, which is required for food intake, was reduced and microinjection of synthetic LURY-1 peptides (10 μ M) caused the same phenotype. Third, adult lifespan was extended by as much as 21–50%. Finally, a reduction in locomotor activity was observed (Ohno et al., 2017). Importantly, all of these phenotypes were largely suppressed by the deletion of *npr-22*, indicating that LURY-1 peptides act upstream of NPR-22. Furthermore, these findings were consistent with a previous analysis of *npr-22* knockdown, which revealed a phenotype in which animals have a reduced body size and brood size (Ceron et al., 2007). More specifically, cell-specific rescue experiments indicated that NPR-22 acts in the feeding pacemaker MC neurons to control feeding and lifespan and NPR-22 acts upstream of the serotonin-uptaking RIH neuron to control egg-laying (Ohno et al., 2017). The authors conclude that food-evoked activation of the pharynx triggers MC neurons to release of LURY-1 peptides, which then act as hormones via NPR-22-dependent mechanisms to control feeding, egg-laying, and roaming behavior (Ohno et al., 2017). Thus, use of *C. elegans* as a model experimental system has provided the first whole-animal perspective on the physiological/behavioral roles of luqin-type neuropeptide signaling.

RYamide-TYPE NEUROPEPTIDE PRECURSORS IN OTHER ECDYSOZOANS

Analysis of transcriptome sequence data from the penis worm *Priapulus caudatus* (phylum Priapulida) revealed the existence of a luqin-type neuropeptide precursor in this species (Yañez-Guerra et al., 2018). This precursor comprises two putative luqin-type neuropeptides with a C-terminal RYamide motif and has the conserved C-terminal region with two cysteines separated by 10 amino acid residues (Figure 1). Interestingly, the *P. caudatus* neuropeptides share similarities with both arthropod/nematode RYamides and spiralian luqins. Thus, while the C-terminal RYamide motif is a conserved feature of this neuropeptide family in ecdysozoans, the N-terminal region of the *P. caudatus* luqin-type neuropeptides shares more sequence similarity with mollusc/annelid luqins than with arthropod/nematode RYamides. For example, the N-terminal sequence QWRP in one of the *P. caudatus* RYamides is also a feature of several molluscan luqin-type peptides (Figure 1). These “intermediate” characteristics of the *P. caudatus* RYamides are also reflected in

a phylogenetic analysis of the relationships of luqin/RYamide-type precursors, where the *P. caudatus* RYamide precursor is not positioned in a clade comprising arthropod/nematode precursors, as would be expected based on animal phylogenetic relationships, but instead it is positioned at the base of a clade comprising mollusc/annelid precursors. This suggests that the *P. caudatus* RYamide precursor may have retained many of the ancestral characteristics of protostome luqin-type precursors, whereas the arthropod/nematode luqin-type precursors appear to be more divergent (Yañez-Guerra et al., 2018). In *P. caudatus*, two candidate receptors for luqin-type neuropeptides have been identified based on their phylogenetic relationship with luqin-type receptors that have been characterized in other protostomes (Figure 2) (Yañez-Guerra et al., 2018). Experimental studies are now needed to determine if the two *P. caudatus* luqin-type neuropeptides are effective as ligands for both receptors or if the receptors exhibit preferential ligand binding. Furthermore, experimental studies are needed to investigate the physiological roles of luqin-type signaling in priapulids.

Analysis of the genome sequences of the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus* (phylum Tardigrada) has revealed the occurrence of luqin-type precursors and receptors in both of these species (Kozioł, 2018). One luqin-type precursor was identified in *H. dujardini* and two luqin-type precursors were identified in *R. varieornatus* (Kozioł, 2018). The predicted neuropeptides derived from these precursors have a C-terminal RYamide motif, consistent with other members of this neuropeptide family in ecdysozoans. However, atypical of ecdysozoan luqin-type precursors, the tardigrade precursors comprise only one predicted neuropeptide (Figure 1). The C-terminal region of the precursor contains two cysteines separated by 10 amino acid residues in *H. dujardini* and in one of the two precursors from *R. varieornatus*, while the second precursor in *R. varieornatus* lacks the second cysteine (Figure 1). Genes encoding a single luqin-type receptor have been identified in both *H. dujardini* and *R. varieornatus* (Kozioł, 2018) but the ligand-binding properties of these receptors remain to be investigated experimentally. Furthermore, nothing is known about the physiological roles of luqin-type neuropeptides in tardigrades and so this will be an interesting area for investigation in the future, particularly in the context of their remarkable capacity to withstand extreme environmental conditions, including radiation tolerance, desiccation, and both high and low temperature and pressure (Jönsson, 2019; Jönsson et al., 2019; Kamilari et al., 2019).

DISCOVERY OF LUQIN-TYPE SIGNALING IN AMBULACRARIAN DEUTEROSTOMES REVEALS THE URBILATERIAN ORIGIN OF THIS NEUROPEPTIDE SIGNALING SYSTEM

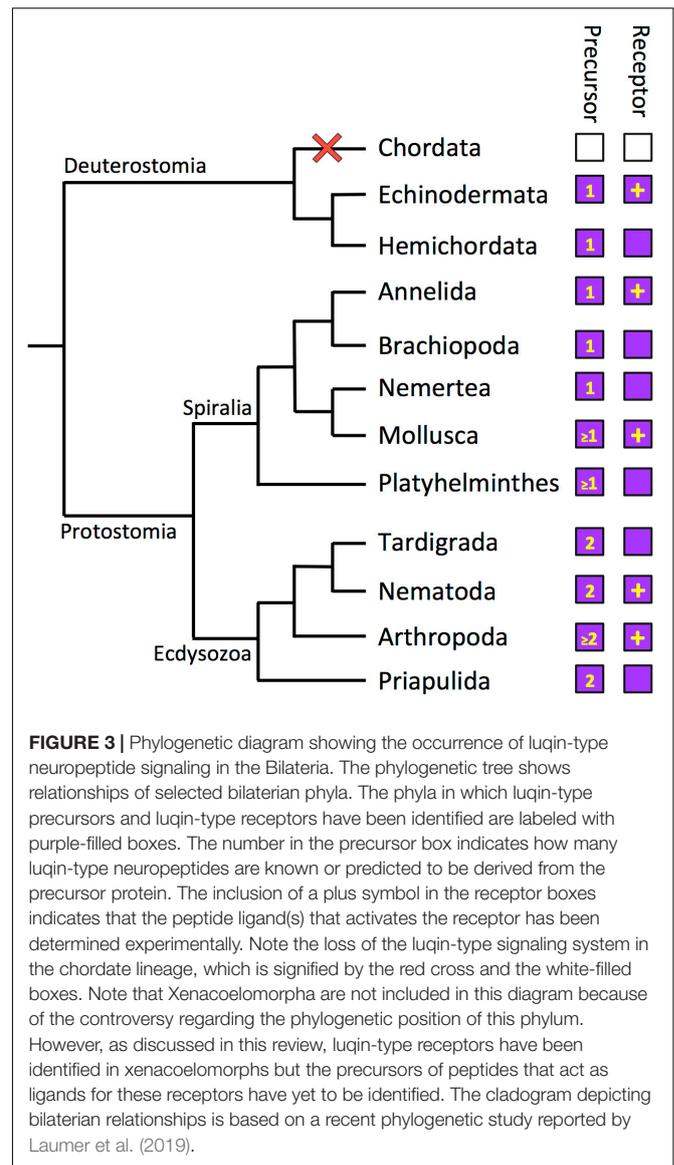
The discovery of precursors of luqin-type peptides in deuterostomian invertebrates was first reported in 2013. They were identified in the hemichordate *Saccoglossus kowalevskii*

and in the echinoderm *Strongylocentrotus purpuratus*, which was facilitated by the presence of the aforementioned conserved C-terminal region containing two cysteines separated by 10 amino acid residues (Jékely, 2013). The *S. purpuratus* and *S. kowalevskii* precursor proteins comprise the putative neuropeptides EIRSPGGKPHKFMWR-NH₂ and EGSNTFLRW-NH₂, respectively. Thus, the presence of a C-terminal FXRW-NH₂ motif (where X is L or M) was identified as a characteristic feature of luqin-type peptides in the ambulacrarian clade of the deuterostomes (Elphick and Mirabeau, 2014).

Through analysis of transcriptome/genome sequence data, luqin-type neuropeptide precursors have subsequently been identified in other echinoderm classes, including Holothuroidea (sea cucumbers), Asteroidea (starfish or sea stars), and Ophiuroidea (brittle stars). In the sea cucumbers *Apostichopus japonicus*, *Holothuria glaberrima*, *Holothuria scabra*, and *Holothuria leucospilota*, the precursor comprises a single neuropeptide with the same predicted structure in all four species—KPYKFMWR-NH₂ (Rowe et al., 2014; Suwansa-Ard et al., 2018; Chen et al., 2019; Chieu et al., 2019). Luqin-type precursors identified in the starfish species *A. rubens* and *Acanthaster planci* comprise a single putative neuropeptide with the amino acid sequence EKGRFPKFMWR-NH₂ and EEKTRFPKFMWR-NH₂, respectively (Semmens et al., 2016; Smith et al., 2017). Ophiuroid luqin-type precursors also comprise a single putative neuropeptide, which has the predicted sequence QGFNRDGPAPKFMWR-NH₂ in *Ophionotus victoriae*, QGFNRGEGPAKFMWR-NH₂ in *Ophiopsila aranea*, and QGFSRDGPAPKFMWR-NH₂ in *Amphiura filiformis* (Zandawala et al., 2017). Thus, the C-terminal motif KFMWR-NH₂ appears to be a conserved feature of luqin-type neuropeptides in echinoderms.

A large-scale analysis of the phylogenetic distribution of G-protein coupled neuropeptide receptors in bilaterian animals revealed the presence of luqin-type receptors in ambulacrarians (hemichordates and echinoderms) (Mirabeau and Joly, 2013), consistent with the identification of luqin-type neuropeptide precursors in these taxa. However, luqin-type receptors were not identified in vertebrates or other chordates (urochordates and cephalochordates) and accordingly luqin-type neuropeptide precursors have not been identified in these taxa. Thus, it was concluded that the evolutionary origin of luqin-type receptors can be traced to the common ancestor of protostomes and deuterostomes, but with subsequent loss in the chordate lineage (Mirabeau and Joly, 2013). Furthermore, the phylogenetic analysis of neuropeptide receptor relationships reported by Mirabeau and Joly (2013) revealed that luqin-type receptors are paralogs of tachykinin-type receptors and this finding was confirmed recently by a phylogenetic analysis specifically focused on luqin-type receptors and closely related neuropeptide receptors (Yañez-Guerra et al., 2018) (Figure 2). Thus, it can be inferred that gene duplication in a common ancestor of the Bilateria gave rise to the paralogous luqin-type and tachykinin-type signaling systems, but with subsequent loss of luqin-type signaling in chordates (Figure 3).

A detailed analysis of luqin-type receptors in ambulacrarians revealed the presence of four genes/transcripts in the



hemichordate *S. kowalevskii* and two genes/transcripts in the echinoderms *S. purpuratus* (sea urchin) and *A. rubens* (starfish) that encode members of this family of neuropeptide receptors (Yañez-Guerra et al., 2018). Furthermore, the starfish *A. rubens* was selected a model experimental system in which to functionally characterize luqin-type neuropeptide signaling for the first time in a deuterostome. A cDNA encoding the *A. rubens* luqin-type precursor ArLQP was cloned and the structure of the mature peptide derived from this precursor was determined using mass spectrometry as a 12 amino acid residue peptide that is C-terminally amidated—EEKTRFPKFMWR-NH₂ (ArLQ). Cloning, sequencing, and heterologous expression of cDNAs encoding two *A. rubens* luqin-type receptors (ArLQR1 and ArLQR2) facilitated testing of synthetic ArLQ as a candidate ligand for these receptors. This revealed that ArLQ is a potent ligand for both ArLQR1 and ArLQR2, with

EC₅₀ values of 2.4×10^{-8} and 7.8×10^{-10} M, respectively (Yañez-Guerra et al., 2018).

To gain insights into the physiological roles of luqin-type neuropeptide signaling in *A. rubens*, mRNA *in situ* hybridization methods were employed to investigate the expression pattern of ArLQP in adult starfish. ArLQP-expressing cells were revealed in the central nervous system, including the circumoral nerve ring and the radial nerve cords. However, expression was limited to the ectoneural region, which contains sensory and interneurons, with no expression detected in motoneuron cell bodies located in the hyponeuronal region. ArLQP-expressing cells were also revealed in starfish locomotor organs—tube feet—specifically located in close proximity to the basal nerve ring in the disk region. Lastly, in the digestive system, ArLQP-expressing cells were revealed in the cardiac stomach and pyloric stomach (Yañez-Guerra et al., 2018). Efforts to generate antibodies to ArLQ were also made to enable immunohistochemical analysis of ArLQ expression in *A. rubens*, but these were unsuccessful. Nevertheless, informed by the pattern of ArLQP expression revealed by use of mRNA *in situ* hybridization, synthetic ArLQ was tested as a potential myoactive peptide on *in vitro* preparations of tube feet and cardiac stomach. No effects on cardiac stomach preparations were observed but, interestingly, ArLQ was found to cause dose-dependent relaxation of tube foot preparations (Yañez-Guerra et al., 2018). Furthermore, the relaxing effect of ArLQ was similar in potency and magnitude to SALMFamide-2 (S2), a neuropeptide that has been identified and functionally characterized previously as a myorelaxant in *A. rubens* (Elphick et al., 1991; Melarange and Elphick, 2003). Clearly, further studies are now needed to gain broader insights into the physiological roles of luqin-type neuropeptides in starfish and other echinoderms. Further investigation of the physiological roles of luqin-type neuropeptide signaling in echinoderms could also be extended beyond adult animals to the free-swimming larval stage of these animals. Detailed anatomical analyses of neuropeptide precursor gene expression in larvae of *A. rubens* and *S. purpuratus* have been reported recently (Mayorova et al., 2016; Wood et al., 2018) but these studies did not incorporate analysis of the expression of luqin-type precursors. Therefore, this is also an important objective for future work on luqin-type neuropeptide signaling in echinoderms.

LUQIN-TYPE NEUROPEPTIDE SIGNALING IN XENACOELOMORPHS: RECEPTORS WITH MISSING LIGANDS

The phylum Xenacoelomorpha comprises an assemblage of marine worms that have a simple body plan without a through-gut (Hejnol and Pang, 2016; Telford and Copley, 2016; Gavilán et al., 2019). They are of particular interest for evolutionary studies because of controversy regarding their phylogenetic position in the animal kingdom. On the one hand, they have been placed as a sister group to all other bilaterian animals [Nephrozoa hypothesis] (Cannon et al., 2016; Rouse et al., 2016). Alternatively, they are considered to be closely related to the

Ambulacraria, forming a clade known as the Xenambulacraria [Xenambulacraria hypothesis] (Bourlat et al., 2006; Philippe et al., 2011). The most recent analysis of the phylogenetic position of xenacoelomorphs, including a strategy devoted to mitigate the effects of systematic errors, has supported the Xenambulacraria hypothesis (Philippe et al., 2019).

Analysis of transcriptome sequence data from 13 xenacoelomorph species revealed the occurrence of luqin-type receptors in this phylum (Thiel et al., 2018). Transcripts encoding luqin-type receptors were identified in species belonging to each of the three xenacoelomorph sub-phyla; Xenoturbellida (*Xenoturbella bocki*), Nemertodermatida (*Nemertoderma westbladi*), and Acoela (*Hofstenia miamia*). Furthermore, phylogenetic analysis revealed that these receptors form part of a clade of luqin-type receptors that include spiralian luqin receptors and ecdysozoan RYamide receptors. Interestingly, the xenacoelomorph luqin-type receptors are positioned within a branch that also includes ambulacrarian luqin-type receptors (Thiel et al., 2018). Therefore, this may be additional evidence in support of the Xenambulacraria hypothesis. Thus far, precursors of luqin-type neuropeptides have yet to be identified in xenacoelomorphs and so this represents an important objective for future work. In particular, it would be interesting to determine the C-terminal motif of luqin-type neuropeptides in xenacoelomorphs. If the peptides have a C-terminal RWamide motif, which is a characteristic of ambulacrarian luqin-type neuropeptides, then this would be further evidence of a close relationship between xenacoelomorphs and ambulacrarians. Furthermore, discovery of luqin-type neuropeptides in xenacoelomorphs would provide a basis for functional investigation of physiological roles of these neuropeptides in this phylum.

GENERAL CONCLUSIONS AND SPECULATIONS

The discovery and naming of luqin as a neuropeptide that is expressed in left upper quadrant cells of the abdominal ganglion of the mollusc *Aplysia* was perhaps a rather esoteric beginning to a new field of neuropeptide research. However, gradually over a period of more than three decades, it has become apparent that in fact this finding has broad relevance to neuropeptide signaling in bilaterian animals, with the notable exception of chordates. Although the name luqin was highly specific in its derivation, it nevertheless provides a useful generic name for the neuropeptide family. Thus, it is preferable to RFamides, RYamides, or RWamides because these C-terminal motifs are not unique to or even generally applicable to the neuropeptide family as whole. Therefore, our recommendation is that all members of this neuropeptide family are referred to as “luqins,” recognizing of course that the derivation of the name is meaningless beyond *Aplysia*.

Comparison of the sequences of luqin-type precursors has revealed variability in the number of luqin-type neuropeptides derived from these proteins. All the luqin-type precursors identified thus far in ambulacrarians comprise a single luqin-type

neuropeptide and this is also a feature of some luqin-type precursors in spiralian and ecdysozoans (Figure 1). Therefore, it could be inferred that this may reflect the characteristics of the luqin-type precursor in the common ancestor of the Bilateria. However, the existence of precursors that comprise two luqin-type neuropeptides is a feature of many ecdysozoans and some spiralian species, indicating perhaps that this characteristic has evolved independently in both lineages (Figure 1). Further expansions in the number of luqin-type neuropeptides derived from precursor proteins are seen in some crustacean species; for example, in *D. pulex* and *P. cavaticus*, where there are three predicted mature peptides (Dirksen et al., 2011; Christie, 2017) (Supplementary Figure 1A). However, the most extreme example is seen in the wasp *N. vitiprenis*, where the precursor comprises seven predicted luqin-type neuropeptides (Hauser et al., 2010) (Supplementary Figure 1A). The functional significance of this expansion in some arthropods is as yet unknown. Furthermore, the existence of multiple genes encoding luqin-type neuropeptides is a feature of some platyhelminth species (Kozioł et al., 2016).

Thus far, genes/transcripts encoding luqin-type precursors and/or luqin-type receptors have been discovered in all of the non-chordate bilaterian phyla that have been investigated, but not in chordates (Figure 3). The loss of LQ-type signaling in chordates is of interest from a functional perspective, as discussed below. However, loss of LQ-type signaling in chordates has also influenced efforts to classify neuropeptide receptors. Hence, the original classification of the *C. elegans* and *L. stagnalis* LQ-type receptors as “tachykinin-like receptors” (Keating et al., 2003). More recently, in a phylogenetic analysis of gonadotropin-inhibitory hormone (GnIH)-type signaling, it was concluded that GnIH-type receptors have a “strong evolutionary relationship” with the *C. elegans* receptor NPR-22 (Ubuka and Tsutsui, 2018). This conclusion was a consequence of a phylogenetic analysis that was restricted to comparison of human GnIH-type receptors with a variety of *C. elegans* neuropeptide receptors, while more comprehensive phylogenetic analyses clearly show that the *C. elegans* receptor NPR-22 is an LQ-type receptor (Ohno et al., 2017; Yañez-Guerra et al., 2018).

While genes/transcripts encoding luqin-type precursors and/or luqin-type receptors have been discovered in all of the non-chordate bilaterian phyla that have been investigated, there are still many phyla that remain to be examined and so we cannot rule out the possibility that luqin-type neuropeptide signaling has been lost in some non-chordate phyla. Nevertheless, based on what is currently known, the loss of luqin signaling in chordates appears to be singular and therefore notable. Why then was luqin signaling lost in the chordates? To address this question, we need to consider, first, what is known about the physiological roles of luqin-type neuropeptide signaling and, second, the paralogous relationship of luqin-type signaling with tachykinin-type signaling.

Although luqin-type signaling was first discovered and then functionally characterized in molluscs, it was the recent detailed analysis of this signaling system in the nematode *C. elegans*, employing use of reverse genetic techniques, that has provided the most comprehensive insights into the physiological

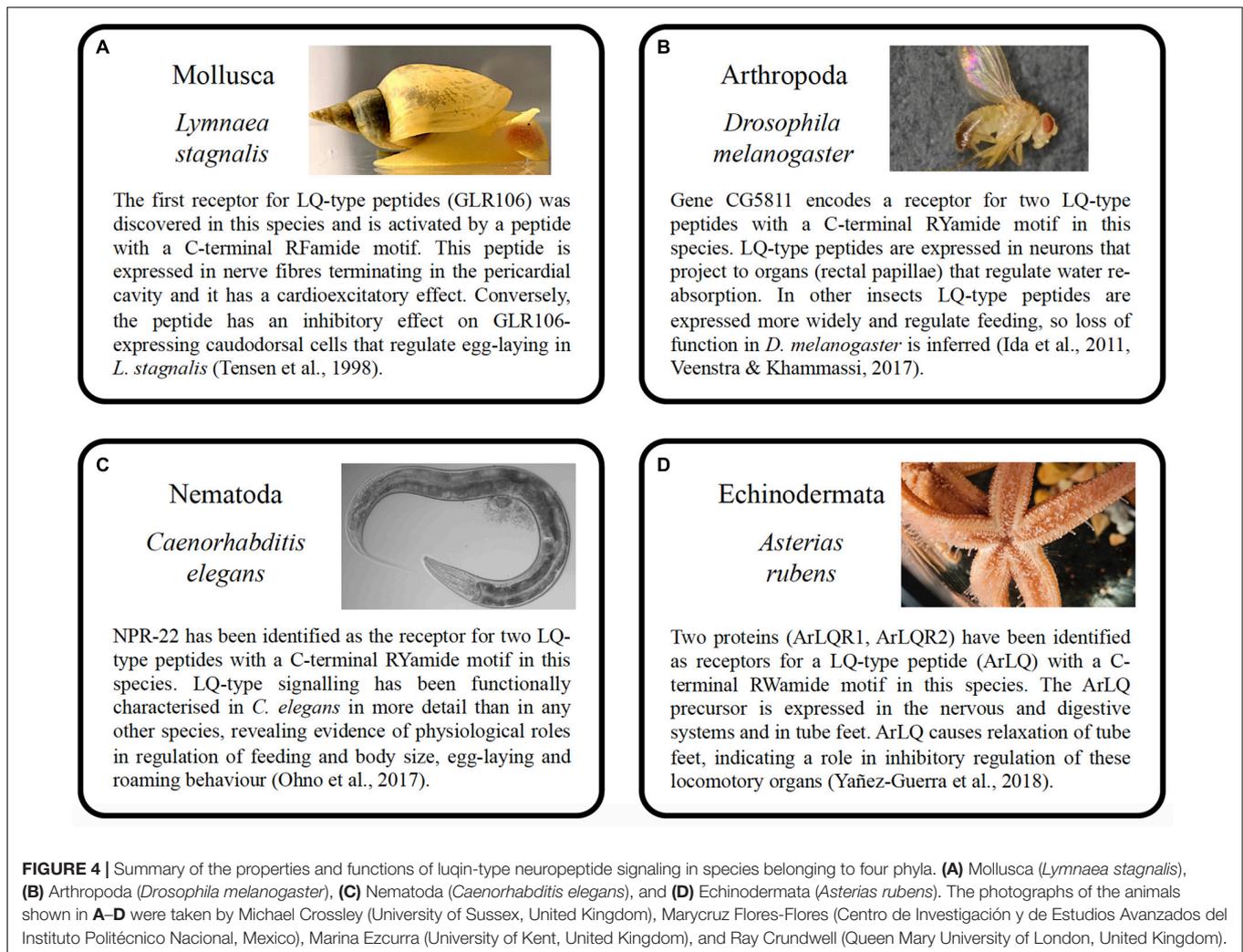
roles of luqin-type neuropeptide signaling. Furthermore, the findings from *C. elegans* provide a broad functional context for comparison with experimental findings from other less intensely studied taxa (Figure 4).

A key finding from *C. elegans* was that luqin-type neuropeptides are secreted by a pair of pharyngeal neurons (M1 and M2) and act as hormones to suppress feeding (Ohno et al., 2017). It is noteworthy, therefore, that luqin-type RYamides suppress feeding motivation and sucrose responsiveness in an insect (Maeda et al., 2015) and brain expression of the luqin-type RYamide precursor gene is significantly downregulated after starvation in a crustacean (Mekata et al., 2017). Thus, in ecdysozoans, there is evidence of an evolutionarily conserved role of luqin-type neuropeptide signaling as an inhibitory regulator of feeding in association with changes in food availability. Further studies are now required to investigate if luqin-type neuropeptides also regulate feeding activity in spiralian and ambulacrarians. Nevertheless, it is noteworthy that in the mollusc *A. fullica* luqin has excitatory effects on buccal neurons and muscles (Fujimoto et al., 1990) and in the starfish *A. rubens* a luqin-type neuropeptide precursor is expressed in the cardiac stomach, a region of the digestive system that is everted when starfish feed (Yañez-Guerra et al., 2018).

Luqin-type neuropeptides also regulate egg laying in *C. elegans*, acting via the luqin-type receptor NPR-22 upstream of the serotonin-uptaking RIH neuron. Thus, by comparison with wild-type animals, mutant worms lacking expression of the luqin-type LURY-1 precursor or NPR-22 exhibit reduced egg-laying during a period of refeeding after starvation (Ohno et al., 2017). Accordingly, detection of luqin-immunoreactivity in nerve fibers associated with the hermaphroditic ducts and the ovotestis and in nerve fibers associated with inhibition of the egg-laying hormone-producing caudodorsal cells in *L. stagnalis*, which express luqin-type receptor transcripts (Tensen et al., 1998), suggest that luqin-type signaling may likewise regulate egg laying in molluscs.

Overexpression of luqin-type neuropeptides in *C. elegans* produced a phenotype where worms exhibited reduced roaming activity and, importantly, this was not observed in animals lacking the luqin-type receptor NPR-22. It was concluded that this action may reflect a physiological role of luqin signaling in attenuating locomotory activity when food is abundant (Ohno et al., 2017). It is noteworthy, therefore, that the relaxing effect of the luqin-type neuropeptide ArLQ on starfish tube feet may likewise be consistent with a physiological role in inhibitory regulation of locomotor activity in an echinoderm (Yañez-Guerra et al., 2018). Investigation of the effects of ArLQ on locomotor activity will therefore be an interesting objective for future studies, employing use of methods that have been reported recently to examine the effects of other neuropeptides on starfish locomotor activity (Tinoco et al., 2018).

The discovery that luqin-type signaling is paralogous to tachykinin-type signaling was based on phylogenetic analysis of the relationships of G-protein coupled neuropeptide receptors in the Bilateria (Mirabeau and Joly, 2013; Thiel et al., 2018; Yañez-Guerra et al., 2018) (Figure 3). Therefore, it can be inferred that duplication of genes encoding a common ancestral neuropeptide



precursor and receptor occurred in a common ancestor of the Bilateria to give rise to the paralogous luqin-type and tachykinin-type signaling systems. Analysis of the phylogenetic distribution of tachykinin-type signaling indicates that it has been retained in all bilaterian phyla that have been analyzed (Mirabeau and Joly, 2013; Elphick et al., 2018).

A possible explanation for the loss of a neuropeptide signaling system in a taxon could be functional redundancy with respect to a paralogous signaling system. However, given (i) the length of time elapsed since the gene duplications that gave rise to the paralogous luqin-type and tachykinin-type signaling systems is likely to be in excess of 650 million years, based on the estimated time of divergence of protostomes and deuterostomes (Erwin et al., 2011) and (ii) the preservation of both signaling systems in most bilaterian phyla that have been analyzed, it seems unlikely that this explains the loss of luqin signaling in chordates. Further insights may be obtained as part of a broader investigation of the functional significance of the loss of bilaterian neuropeptide signaling systems. Examples of other neuropeptide types that have been lost in chordates include the leucokinin-type and pigment dispersing factor (PDF)-type signaling systems

(Mirabeau and Joly, 2013) while more specifically corazonin-type signaling has been lost in vertebrates and urochordates but not in cephalochordates (Tian et al., 2016; Zandawala et al., 2018). Comparisons with these neuropeptide systems may therefore be informative in future efforts to draw general conclusions on the evolutionary and functional significance of neuropeptide loss in chordates.

AUTHOR CONTRIBUTIONS

Both authors contributed equally to the planning and writing of the review article. LY-G produced the figures.

FUNDING

LY-G was supported by a Leverhulme Trust grant (RPG-2016-353) and by a Ph.D. studentship awarded by the Mexican Council of Science and Technology (CONACyT studentship No. 418612) and Queen Mary University of London. ME was

supported by grants from the Biotechnology and Biological Sciences Research Council (BB/M001644/1) and the Leverhulme Trust (RPG-2016-353).

ACKNOWLEDGMENTS

We thank Michael Crossley (University of Sussex, United Kingdom), Marycruz Flores-Flores (Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico), Marina Ezcurra (University of Kent,

United Kingdom), and Ray Crundwell (Queen Mary University of London, United Kingdom) for providing copyright free photographs of *Lymnaea stagnalis*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Asterias rubens*, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2020.00130/full#supplementary-material>

REFERENCES

- Aloyz, R. S., and DesGroseillers, L. (1995). Processing of the L5-67 precursor peptide and characterization of LUQIN in the LUQ neurons of *Aplysia californica*. *Peptides* 16, 331–338. doi: 10.1016/0196-9781(94)00140-5
- Angers, A., and DesGroseillers, L. (1998). Alternative splicing and genomic organization of the L5-67 gene of *Aplysia californica*. *Gene* 208, 271–277. doi: 10.1016/s0378-1119(98)00009-2
- Angers, A., Zappulla, J. P., Zollinger, M., and DesGroseillers, L. (2000). Gene products from LUQ neurons in the abdominal ganglion are present at the renal pore of *Aplysia californica*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 126, 435–443. doi: 10.1016/s0305-0491(00)00217-0
- Bauknecht, P., and Jékely, G. (2015). Large-scale combinatorial deorphanization of Platynereis neuropeptide GPCRs. *Cell Rep.* 12, 684–693. doi: 10.1016/j.celrep.2015.06.052
- Bhatla, N., Droste, R., Sando, S. R., Huang, A., and Horvitz, H. R. (2015). Distinct neural circuits control rhythm inhibition and spitting by the myogenic pharynx of *C. elegans*. *Curr. Biol.* 25, 2075–2089. doi: 10.1016/j.cub.2015.06.052
- Boulat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Aronowicz, J., Kirschner, M., et al. (2006). Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* 444, 85–88. doi: 10.1038/nature05241
- Brown, R. O., Gusman, D., Basbaum, A. I., and Mayeri, E. (1985). Identification of *Aplysia* neurons containing immunoreactive FMRFamide. *Neuropeptides* 6, 517–526. doi: 10.1016/0143-4179(85)90113-1
- Cannon, J. T., Vellutini, B. C., Smith, J., Ronquist, F., Jondelius, U., and Hejnol, A. (2016). Xenacoelomorpha is the sister group to Nephrozoa. *Nature* 530, 89–93. doi: 10.1038/nature16520
- Ceron, J., Rual, J.-F., Chandra, A., Dupuy, D., Vidal, M., and van den Heuvel, S. (2007). Large-scale RNAi screens identify novel genes that interact with the *C. elegans* retinoblastoma pathway as well as splicing-related components with synMuv B activity. *BMC Dev. Biol.* 7:30. doi: 10.1186/1471-213X-7-30
- Chen, M., Talarovicova, A., Zheng, Y., Storey, K. B., and Elphick, M. R. (2019). Neuropeptide precursors and neuropeptides in the sea cucumber *Apostichopus japonicus*: a genomic, transcriptomic and proteomic analysis. *Sci. Rep.* 9:8829. doi: 10.1038/s41598-019-45271-3
- Chieu, H. D., Suwansa-Ard, S., Wang, T., Elizur, A., and Cummins, S. F. (2019). Identification of neuropeptides in the sea cucumber *Holothuria leucospilota*. *Gen. Comp. Endocrinol.* 283:113229. doi: 10.1016/j.ygcen.2019.113229
- Christie, A. E. (2017). Neuropeptide discovery in *Proasellus cavaticus*: prediction of the first large-scale peptidome for a member of the Isopoda using a publicly accessible transcriptome. *Peptides* 97, 29–45. doi: 10.1016/j.peptides.2017.09.003
- Christie, A. E., Cieslak, M. C., Roncalli, V., Lenz, P. H., Major, K. M., and Poynton, H. C. (2018). Prediction of a peptidome for the ecotoxicological model *Hyalella azteca* (Crustacea: Amphipoda) using a de novo assembled transcriptome. *Mar. Genomics* 38, 67–88. doi: 10.1016/j.margen.2017.12.003
- Collin, C., Hauser, F., Krogh-Meyer, P., Hansen, K. K., Gonzalez de Valdivia, E., Williamson, M., et al. (2011). Identification of the *Drosophila* and *Tribolium* receptors for the recently discovered insect RYamide neuropeptides. *Biochem. Biophys. Res. Commun.* 412, 578–583. doi: 10.1016/j.bbrc.2011.07.131
- Conzelmann, M., Williams, E. A., Krug, K., Franz-Wachtel, M., Macek, B., and Jékely, G. (2013). The neuropeptide complement of the marine annelid *Platynereis dumerilii*. *BMC Genomics* 14:906. doi: 10.1186/1471-2164-14-906
- Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10:210. doi: 10.1186/1471-2148-10-210
- De Oliveira, A. L., Calcino, A., and Wanninger, A. (2019). Extensive conservation of the proneuropeptide and peptide prohormone complement in mollusks. *Sci. Rep.* 9:4846. doi: 10.1038/s41598-019-40949-0
- Dickinson, P. S., Hull, J. J., Miller, A., Oleisky, E. R., and Christie, A. E. (2019). To what extent may peptide receptor gene diversity/complement contribute to functional flexibility in a simple pattern-generating neural network? *Comp. Biochem. Physiol. Part D Genomics Proteomics* 30, 262–282. doi: 10.1016/j.cbd.2019.03.002
- Dirksen, H., Neupert, S., Predel, R., Verleyen, P., Huybrechts, J., Strauss, J., et al. (2011). Genomics, transcriptomics, and peptidomics of *Daphnia pulex* neuropeptides and protein hormones. *J. Proteome Res.* 10, 4478–4504. doi: 10.1021/pr200284e
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Elphick, M. R., and Mirabeau, O. (2014). The evolution and variety of RFamide-type neuropeptides: insights from deuterostomian invertebrates. *Front. Endocrinol.* 5:93. doi: 10.3389/fendo.2014.00093
- Elphick, M. R., Mirabeau, O., and Larhammar, D. (2018). Evolution of neuropeptide signalling systems. *J. Exp. Biol.* 221(Pt 3):jeb151092. doi: 10.1242/jeb.151092
- Elphick, M. R., Price, D. A., Lee, T. D., and Thorndyke, M. C. (1991). The SALMFamides: a new family of neuropeptides isolated from an echinoderm. *Proc. Biol. Sci.* 243, 121–127. doi: 10.1098/rspb.1991.0020
- Erwin, D. H., Laflamme, M., Tweedt, S. M., Sperling, E. A., Pisani, D., and Peterson, K. J. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* 334, 1091–1097. doi: 10.1126/science.1206375
- Fu, Q., Kutz, K. K., Schmidt, J. J., Hsu, Y.-W. A., Messinger, D. I., Cain, S. D., et al. (2005). Hormone complement of the *Cancer productus* sinus gland and pericardial organ: an anatomical and mass spectrometric investigation. *J. Comp. Neurol.* 493, 607–626. doi: 10.1002/cne.20773
- Fujimoto, K., Ohta, N., Yoshida, M., Kubota, I., Muneoka, Y., and Kobayashi, M. (1990). A novel cardio-excitatory peptide isolated from the atria of the African giant snail, *Achatina fulica*. *Biochem. Biophys. Res. Commun.* 167, 777–783. doi: 10.1016/0006-291x(90)92093-f
- Gavilán, B., Sprecher, S. G., Hartenstein, V., and Martínez, P. (2019). The digestive system of xenacoelomorphs. *Cell Tissue Res.* 377, 369–382. doi: 10.1007/s00441-019-03038-2
- Giardino, N. D., Aloyz, R. S., Zollinger, M., Miller, M. W., and DesGroseillers, L. (1996). L5-67 and LUQ-1 peptide precursors of *Aplysia californica*: distribution and localization of immunoreactivity in the central nervous system and in peripheral tissues. *J. Comp. Neurol.* 374, 230–245. doi: 10.1002/(sici)1096-9861(19961014)374:2<230::aid-cne6>3.0.co;2-3
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Hauser, F., Neupert, S., Williamson, M., Predel, R., Tanaka, Y., and Gimmelikhuijzen, C. J. P. (2010). Genomics and peptidomics of neuropeptides

- and protein hormones present in the parasitic wasp *Nasonia vitripennis*. *J. Proteome Res.* 9, 5296–5310. doi: 10.1021/pr100570j
- Hejnol, A., and Pang, K. (2016). Xenacoelomorpha's significance for understanding bilaterian evolution. *Curr. Opin. Genet. Dev.* 39, 48–54. doi: 10.1016/j.cde.2016.05.019
- Ida, T., Takahashi, T., Tominaga, H., Sato, T., Kume, K., Ozaki, M., et al. (2011). Identification of the novel bioactive peptides dRYamide-1 and dRYamide-2, ligands for a neuropeptide Y-like receptor in *Drosophila*. *Biochem. Biophys. Res. Commun.* 410, 872–877. doi: 10.1016/j.bbrc.2011.06.081
- Jékely, G. (2013). Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8702–8707. doi: 10.1073/pnas.1221833110
- Jékely, G., Melzer, S., Beets, I., Kadow, I. C. G., Koene, J., Haddad, S., et al. (2018). The long and the short of it - a perspective on peptidergic regulation of circuits and behaviour. *J. Exp. Biol.* 221(Pt 3):jeb166710. doi: 10.1242/jeb.166710
- Jönsson, K. I. (2019). Radiation tolerance in tardigrades: current knowledge and potential applications in medicine. *Cancers* 11:1333. doi: 10.3390/cancers11091333
- Jönsson, K. I., Holm, I., and Tassidis, H. (2019). Cell biology of the tardigrades: current knowledge and perspectives. *Results Probl. Cell Differ.* 68, 231–249. doi: 10.1007/978-3-030-23459-1_10
- Kamilari, M., Jørgensen, A., Schiott, M., and Møbjerg, N. (2019). Comparative transcriptomics suggest unique molecular adaptations within tardigrade lineages. *BMC Genomics* 20:607. doi: 10.1186/s12864-019-5912-x
- Keating, C. D., Kriek, N., Daniels, M., Ashcroft, N. R., Hopper, N. A., Siney, E. J., et al. (2003). Whole-genome analysis of 60 G protein-coupled receptors in *Caenorhabditis elegans* by gene knockout with RNAi. *Curr. Biol.* 13, 1715–1720. doi: 10.1016/j.cub.2003.09.003
- Kozioł, U. (2018). Precursors of neuropeptides and peptide hormones in the genomes of tardigrades. *Gen. Comp. Endocrinol.* 267, 116–127. doi: 10.1016/j.ygcn.2018.06.012
- Kozioł, U., Kozioł, M., Preza, M., Costabile, A., Brehm, K., and Castillo, E. (2016). De novo discovery of neuropeptides in the genomes of parasitic flatworms using a novel comparative approach. *Int. J. Parasitol.* 46, 709–721. doi: 10.1016/j.ijpara.2016.05.007
- Laumer, C. E., Fernández, R., Lemer, S., Combosch, D., Kocot, K. M., Riesgo, A., et al. (2019). Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. Biol. Sci.* 286:20190831. doi: 10.1098/rspb.2019.0831
- Li, L., Kelley, W. P., Billimoria, C. P., Christie, A. E., Pulver, S. R., Sweedler, J. V., et al. (2003). Mass spectrometric investigation of the neuropeptide complement and release in the pericardial organs of the crab, *Cancer borealis*. *J. Neurochem.* 87, 642–656. doi: 10.1046/j.1471-4159.2003.02031.x
- Li, L., Moroz, T. P., Garden, R. W., Floyd, P. D., Weiss, K. R., and Sweedler, J. V. (1998). Mass spectrometric survey of interganglionically transported peptides in *Aplysia*. *Peptides* 19, 1425–1433. doi: 10.1016/s0196-9781(98)00094-1
- Liessem, S., Ragionieri, L., Neupert, S., Büschges, A., and Predel, R. (2018). Transcriptomic and neuropeptidomic analysis of the stick insect, *Carausius morosus*. *J. Proteome Res.* 17, 2192–2204. doi: 10.1021/acs.jproteome.8b00155
- Ma, M., Bors, E. K., Dickinson, E. S., Kwiatkowski, M. A., Sousa, G. L., Henry, R. P., et al. (2009). Characterization of the *Carcinus maenas* neuropeptidome by mass spectrometry and functional genomics. *Gen. Comp. Endocrinol.* 161, 320–334. doi: 10.1016/j.ygcn.2009.01.015
- Ma, M., Gard, A. L., Xiang, F., Wang, J., Davoodian, N., Lenz, P. H., et al. (2010). Combining *in silico* transcriptome mining and biological mass spectrometry for neuropeptide discovery in the Pacific white shrimp *Litopenaeus vannamei*. *Peptides* 31, 27–43. doi: 10.1016/j.peptides.2009.10.007
- Maeda, T., Nakamura, Y., Shiotani, H., Hojo, M. K., Yoshii, T., Ida, T., et al. (2015). Suppressive effects of dRYamides on feeding behavior of the blowfly, *Phormia regina*. *Zool. Lett.* 1:35. doi: 10.1186/s40851-015-0034-z
- Marder, E., and Bucher, D. (2007). Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annu. Rev. Physiol.* 69, 291–316. doi: 10.1146/annurev.physiol.69.031905.161516
- Mayorova, T. D., Tian, S., Cai, W., Semmens, D. C., Odekunle, E. A., Zandawala, M., et al. (2016). Localization of neuropeptide gene expression in larvae of an echinoderm, the starfish *Asterias rubens*. *Front. Neurosci.* 10:553. doi: 10.3389/fnins.2016.00553
- McVeigh, P., McCammick, E., McCusker, P., Wells, D., Hodgkinson, J., Paterson, S., et al. (2018). Profiling G protein-coupled receptors of *Fasciola hepatica* identifies orphan rhodopsins unique to phylum Platyhelminthes. *Int. J. Parasitol. Drugs Drug Resist.* 8, 87–103. doi: 10.1016/j.ijddr.2018.01.001
- Mekata, T., Kono, T., Satoh, J., Yoshida, M., Mori, K., Sato, T., et al. (2017). Purification and characterization of bioactive peptides RYamide and CCHamide in the kuruma shrimp *Marsupenaeus japonicus*. *Gen. Comp. Endocrinol.* 246, 321–330. doi: 10.1016/j.ygcn.2017.01.008
- Melarange, R., and Elphick, M. R. (2003). Comparative analysis of nitric oxide and SALMFamide neuropeptides as general muscle relaxants in starfish. *J. Exp. Biol.* 206, 893–899. doi: 10.1242/jeb.00197
- Mertens, I., Clinckspoor, I., Janssen, T., Nachman, R., and Schoofs, L. (2006). FMRFamide related peptide ligands activate the *Caenorhabditis elegans* orphan GPCR Y59H11AL.1. *Peptides* 27, 1291–1296. doi: 10.1016/j.peptides.2005.11.017
- Mirabeau, O., and Joly, J.-S. (2013). Molecular evolution of peptidergic signaling systems in bilaterians. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2028–E2037. doi: 10.1073/pnas.1219956110
- Nguyen, T. V., Cummins, S. F., Elizur, A., and Ventura, T. (2016). Transcriptomic characterization and curation of candidate neuropeptides regulating reproduction in the eyestalk ganglia of the Australian crayfish, *Cherax quadricarinatus*. *Sci. Rep.* 6:38658. doi: 10.1038/srep38658
- Ohno, H., Yoshida, M., Sato, T., Kato, J., Miyazato, M., Kojima, M., et al. (2017). Luqin-like RYamide peptides regulate food-evoked responses in *C. elegans*. *eLife* 6:e28877. doi: 10.7554/eLife.28877
- Palamiuc, L., Noble, T., Witham, E., Ratanpal, H., Vaughan, M., and Srinivasan, S. (2017). A tachykinin-like neuroendocrine signalling axis couples central serotonin action and nutrient sensing with peripheral lipid metabolism. *Nat. Commun.* 8:14237. doi: 10.1038/ncomms14237
- Philippe, H., Brinkmann, H., Copley, R. R., Moroz, L. L., Nakano, H., Poustka, A. J., et al. (2011). Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. *Nature* 470, 255–258. doi: 10.1038/nature09676
- Philippe, H., Poustka, A. J., Chiodin, M., Hoff, K. J., Dessimoz, C., Tomiczek, B., et al. (2019). Mitigating anticipated effects of systematic errors supports sister-group relationship between Xenacoelomorpha and Ambulacraria. *Curr. Biol.* 29, 1818–1826.e6. doi: 10.1016/j.cub.2019.04.009
- Roller, L., Čížmár, D., Bednár, B., and Žitňan, D. (2016). Expression of RYamide in the nervous and endocrine system of *Bombyx mori*. *Peptides* 80, 72–79. doi: 10.1016/j.peptides.2016.02.003
- Rouse, G. W., Wilson, N. G., Carvajal, J. I., and Vrijenhoek, R. C. (2016). New deep-sea species of *Xenoturbella* and the position of Xenacoelomorpha. *Nature* 530, 94–97. doi: 10.1038/nature16545
- Rowe, M. L., Achhala, S., and Elphick, M. R. (2014). Neuropeptides and polypeptide hormones in echinoderms: new insights from analysis of the transcriptome of the sea cucumber *Apostichopus japonicus*. *Gen. Comp. Endocrinol.* 197, 43–55. doi: 10.1016/j.ygcn.2013.12.002
- Schaefer, M., Picciotto, M. R., Kreiner, T., Kaldany, R. R., Taussig, R., and Scheller, R. H. (1985). *Aplysia* neurons express a gene encoding multiple FMRFamide neuropeptides. *Cell* 41, 457–467. doi: 10.1016/s0092-8674(85)80019-2
- Semmens, D. C., Mirabeau, O., Moghul, I., Pancholi, M. R., Wurm, Y., and Elphick, M. R. (2016). Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. *Open Biol.* 6:150224. doi: 10.1098/rsob.150224
- Shyamala, M., Fisher, J. M., and Scheller, R. H. (1986). A neuropeptide precursor expressed in *Aplysia* neuron L5. *DNA* 5, 203–208. doi: 10.1089/dna.1986.5.203
- Smith, M. K., Wang, T., Suwansa-Ard, S., Motti, C. A., Elizur, A., Zhao, M., et al. (2017). The neuropeptidome of the crown-of-thorns starfish, *Acanthaster planci*. *J. Proteomics* 165, 61–68. doi: 10.1016/j.jprot.2017.05.026
- Stemmler, E. A., Bruns, E. A., Gardner, N. P., Dickinson, P. S., and Christie, A. E. (2007). Mass spectrometric identification of pEGFYSQRYamide: a crustacean peptide hormone possessing a vertebrate neuropeptide Y (NPY)-like carboxy-terminus. *Gen. Comp. Endocrinol.* 152, 1–7. doi: 10.1016/j.ygcn.2007.02.025
- Suwansa-Ard, S., Chaiyamon, A., Talarovicova, A., Tinikul, R., Tinikul, Y., Poomtong, T., et al. (2018). Transcriptomic discovery and comparative analysis of neuropeptide precursors in sea cucumbers (Holotheuroidea). *Peptides* 99, 231–240. doi: 10.1016/j.peptides.2017.10.008
- Telford, M. J., and Copley, R. R. (2016). Zoology: war of the worms. *Curr. Biol.* 26, R335–R337. doi: 10.1016/j.cub.2016.03.015
- Tensen, C. P., Cox, K. J., Smit, A. B., van der Schors, R. C., Meyerhof, W., Richter, D., et al. (1998). The lymnaea cardioexcitatory peptide (LyCEP) receptor: a

- G-protein-coupled receptor for a novel member of the RFamide neuropeptide family. *J. Neurosci.* 18, 9812–9821. doi: 10.1523/jneurosci.18-23-09812.1998
- Thiel, D., Franz-Wachtel, M., Aguilera, F., and Hejnal, A. (2018). Xenacoelomorph neuropeptidomes reveal a major expansion of neuropeptide systems during early bilaterian evolution. *Mol. Biol. Evol.* 35, 2528–2543. doi: 10.1093/molbev/msy160
- Tian, S., Zandawala, M., Beets, I., Baytemur, E., Slade, S. E., Scrivens, J. H., et al. (2016). Urbilaterian origin of paralogous GnRH and corazonin neuropeptide signalling pathways. *Sci. Rep.* 6:28788. doi: 10.1038/srep28788
- Tinoco, A. B., Semmens, D. C., Patching, E. C., Gunner, E. F., Egertová, M., and Elphick, M. R. (2018). Characterization of NGFFYamide signaling in starfish reveals roles in regulation of feeding behavior and locomotory systems. *Front. Endocrinol.* 9:507. doi: 10.3389/fendo.2018.00507
- Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44, W232–W235. doi: 10.1093/nar/gkw256
- Ubuka, T., and Tsutsui, K. (2018). Comparative and evolutionary aspects of gonadotropin-inhibitory hormone and FMRFamide-like peptide systems. *Front. Neurosci.* 12:747. doi: 10.3389/fnins.2018.00747
- Veenstra, J. A. (2011). Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. *Gen. Comp. Endocrinol.* 171, 160–175. doi: 10.1016/j.ygcen.2011.01.005
- Veenstra, J. A. (2015). The power of next-generation sequencing as illustrated by the neuropeptidome of the crayfish *Procambarus clarkii*. *Gen. Comp. Endocrinol.* 224, 84–95. doi: 10.1016/j.ygcen.2015.06.013
- Veenstra, J. A., and Khammassi, H. (2017). Rudimentary expression of RYamide in *Drosophila melanogaster* relative to other *Drosophila* species points to a functional decline of this neuropeptide gene. *Insect Biochem. Mol. Biol.* 83, 68–79. doi: 10.1016/j.ibmb.2017.03.001
- Wood, N. J., Mattiello, T., Rowe, M. L., Ward, L., Perillo, M., Arnone, M. I., et al. (2018). Neuropeptidergic systems in pluteus larvae of the sea urchin *Strongylocentrotus purpuratus*: neurochemical complexity in a “simple” nervous system. *Front. Endocrinol.* 9:628. doi: 10.3389/fendo.2018.00628
- Yañez-Guerra, L. A., Delroisse, J., Barreiro-Iglesias, A., Slade, S. E., Scrivens, J. H., and Elphick, M. R. (2018). Discovery and functional characterisation of a luqin-type neuropeptide signalling system in a deuterostome. *Sci. Rep.* 8:7220. doi: 10.1038/s41598-018-25606-2
- Zandawala, M., Moghul, I., Yañez Guerra, L. A., Delroisse, J., Abylkassimova, N., Hugall, A. F., et al. (2017). Discovery of novel representatives of bilaterian neuropeptide families and reconstruction of neuropeptide precursor evolution in ophiuroid echinoderms. *Open Biol.* 7:170129. doi: 10.1098/rsob.170129
- Zandawala, M., Tian, S., and Elphick, M. R. (2018). The evolution and nomenclature of GnRH-type and corazonin-type neuropeptide signaling systems. *Gen. Comp. Endocrinol.* 264, 64–77. doi: 10.1016/j.ygcen.2017.06.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Yañez-Guerra and Elphick. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Expression Analysis of Cnidarian-Specific Neuropeptides in a Sea Anemone Unveils an Apical-Organ-Associated Nerve Net That Disintegrates at Metamorphosis

Hannah Zang^{1,2} and Nagayasu Nakanishi^{2*}

¹ Lyon College, Batesville, AR, United States, ² Department of Biological Sciences, University of Arkansas, Fayetteville, AR, United States

OPEN ACCESS

Edited by:

Elizabeth Amy Williams,
University of Exeter, United Kingdom

Reviewed by:

Chiara Sinigaglia,
UMR5242 Institut de Génétique
Fonctionnelle de Lyon (IGFL), France
Meet Zandawala,
Brown University, United States

*Correspondence:

Nagayasu Nakanishi
nnakanis@uark.edu

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 12 November 2019

Accepted: 31 January 2020

Published: 19 February 2020

Citation:

Zang H and Nakanishi N (2020)
Expression Analysis of
Cnidarian-Specific Neuropeptides in a
Sea Anemone Unveils an
Apical-Organ-Associated Nerve Net
That Disintegrates at Metamorphosis.
Front. Endocrinol. 11:63.
doi: 10.3389/fendo.2020.00063

Neuropeptides are ancient neuronal signaling molecules that have diversified across Cnidaria (e.g., jellyfish, corals, and sea anemones) and its sister group Bilateria (e.g., vertebrates, insects, and worms). Over the course of neuropeptide evolution emerged lineage-specific neuropeptides, but their roles in the evolution and diversification of nervous system function remain enigmatic. As a step toward filling in this knowledge gap, we investigated the expression pattern of a cnidarian-specific neuropeptide—RPamide—during the development of the starlet sea anemone *Nematostella vectensis*, using *in situ* hybridization and immunohistochemistry. We show that RPamide precursor transcripts first occur during gastrulation in scattered epithelial cells of the aboral ectoderm. These RPamide-positive epithelial cells exhibit a spindle-shaped, sensory-cell-like morphology, and extend basal neuronal processes that form a nerve net in the aboral ectoderm of the free-swimming planula larva. At the aboral end, RPamide-positive sensory cells become integrated into the developing apical organ that forms a bundle of long cilia referred to as the apical tuft. Later during planula development, RPamide expression becomes evident in sensory cells in the oral ectoderm of the body column and pharynx, and in the developing endodermal nervous system. At metamorphosis into a polyp, the RPamide-positive sensory nerve net in the aboral ectoderm degenerates by apoptosis, and RPamide expression begins in ectodermal sensory cells of growing oral tentacles. In addition, we find that the expression pattern of RPamide in planulae differs from that of conserved neuropeptides that are shared across Cnidaria and Bilateria, indicative of distinct functions. Our results not only provide the anatomical framework necessary to analyze the function of the cnidarian-specific neuropeptides in future studies, but also reveal previously unrecognized features of the sea anemone nervous system—the apical organ neurons of the planula larva, and metamorphosis-associated reorganization of the ectodermal nervous system.

Keywords: neuropeptide, evolution, cnidaria, sea anemone, neural development, metamorphosis

INTRODUCTION

Neuropeptides are short polypeptide hormones that are generated from a larger precursor protein via proteolytic cleavage in neurons and neuroendocrine cells [reviewed in (1)]. The cleaved peptides, in some cases, undergo peptide α -amidation involving post-translational conversion of a C-terminal glycine into an amide group. Processed neuropeptides are stored in membrane-bound secretory vesicles that are distributed throughout the cell. Upon receiving stimuli, the secretory vesicles fuse with the cell membrane, releasing neuropeptides into the extracellular space and/or circulation. The neuropeptides then diffuse and bind to receptors—typically G-protein-coupled receptors—in the cell membrane of target cells, triggering intracellular signaling cascades. Neuropeptides, but not small molecule neurotransmitters such as glutamate, GABA and acetylcholine, are commonly found in the nervous systems of cnidarians (e.g., jellyfishes and sea anemones) and bilaterians (e.g., vertebrates and insects), and thus are thought to be the first neurotransmitter molecules in nervous system evolution (2, 3). In an attempt to better understand how the primordial nervous system may have functioned, efforts have been directed toward resolving deeply conserved functions of neuropeptides that are shared across Cnidaria and Bilateria. These studies suggest a likely ancestral role of deeply conserved Wamide neuropeptides in modulating life cycle transition (4–6). Comparably little is known, however, about how lineage-specific neuropeptides contribute to the evolution and diversification of nervous system function, particularly in Cnidaria. As a step toward understanding the role of lineage-specific neuropeptides in the evolution of neural function in Cnidaria, we investigated the expression pattern of a cnidarian-specific neuropeptide—RPamide—during development of the starlet sea anemone *Nematostella vectensis*.

Cnidaria is an early-evolving animal group that diverged from its sister group Bilateria over 600 million years ago in the Precambrian (7–11). Cnidaria represents a diverse clade in which Medusozoa forms a sister group to Anthozoa (12, 13); Medusozoa consists of Hydrozoa (e.g., Portuguese man o' war), Staurozoa (stalked jellyfish), Scyphozoa (e.g., moonjelly), and Cubozoa (e.g., sea wasp), and Anthozoa comprises Octocorallia (e.g., sea fans) and Hexacorallia (e.g., hard corals and sea anemones). Cnidarians are diploblastic, being composed of ectoderm and endoderm separated by an extracellular matrix called the mesoglea, and are characterized by having a phylum-defining stinging cell type, the cnidocyte. Cnidarian development typically entails multiple life cycle stages, whereby gastrulation generates a two-layered, free-swimming planula larva that metamorphoses into a sessile polyp with oral tentacles. The polyp sexually matures in Anthozoa, Staurozoa and some hydrozoan cnidarians (e.g., *Hydra*); in most medusozoans, the polyp undergoes another round of metamorphosis via transverse fission/strobilation (in Scyphozoa and Cubozoa) or lateral budding (in Hydrozoa) to generate free-swimming medusae that grow and reach sexual maturity. The nervous system develops in the ectoderm [and, in some cases, in the endoderm; e.g., the sea anemone *N. vectensis* (14, 15), the hydrozoan

Hydra littoralis (16), and the scyphozoan *Cyanea capillata* (17)] at each of the life cycle stages [but see (18) for electron microscopic evidence for the lack of the nervous system in cubozoan planulae]. It is composed of epithelial sensory cells and basally localized ganglion cells, which extend basal neurites that form a basiepithelial network (19). This basic organization is modified to generate neural structures of varying morphological complexities, from the oral nerve ring of *Hydra* (20) to the lensed eyes of cubozoan medusae (21). Neuropeptides, but not small molecule neurotransmitters, are ubiquitously expressed in cnidarian nervous systems, and thus are thought to be the primary neurotransmitters and neurohormones of Cnidaria [reviewed in (2, 3)].

Among cnidarians, the starlet sea anemone *N. vectensis* is a useful model to examine neuropeptide function during development, because genomic and transcriptomic resources (8, 22, 23), molecular genetic tools [e.g., *in situ* hybridization and CRISPR-mediated mutagenesis (24, 25)], and data on neural anatomy and development (14, 15) are available. In *N. vectensis*, gastrulation occurs by invagination (26, 27), and the site of gastrulation—the blastopore—becomes the mouth of the animal (28). Both sensory cells and ganglion cells begin to develop in the outer ectoderm during gastrulation (14, 29). The gastrula embryo then develops into a free-swimming planula larva that forms a bundle of long cilia at the aboral pole, which is referred to as the apical tuft (14, 30). The ectodermal structure that houses the apical tuft is termed the apical organ, and is often assumed to be a sensory structure that is used to perceive sensory signals for settlement and metamorphosis [e.g., (30)]. Early-born neurons in the ectoderm send basal neurites toward the base of the apical organ, forming a basiepithelial nerve net (14). During planula development, sensory cells develop in the pharyngeal ectoderm, as well as in the endoderm (14). At metamorphosis, the planula larva transforms into a sessile polyp by developing circumoral tentacles with mechanosensory hair cells and losing the apical tuft (14). We note that *N. vectensis* neurons have been reported to show diverse transcriptome profiles (31), and thus transcriptionally distinct neural cell types may exist among morphologically similar neurons.

Isolated originally from the green aggregating anemone *Anthopleura elegantissima* (32, 33) and recently from the starlet sea anemone *N. vectensis* (34), RPamide represents one of the cnidarian-specific peptides. RPamide precursor-encoding genes have been recovered from the genomes of anthozoans [*A. elegantissima* (35) and *N. vectensis* (36); but absent in Octocorallia (37)] and medusozoans [hydrozoans *Clytia hemisphaerica* and *Cladonema pacificum* (38, 39); scyphozoans *Nemopilema nomurai* and *Rhopilema esculentum* (37); but absent in staurozoans and cubozoans (37)], but not from those of other metazoan groups. Immunohistochemical evidence suggests that RPamide is a neuropeptide; immunostaining with an antiserum against RPamide has shown that RPamide is expressed in ectodermal sensory cells of the body wall in the sea anemone *Calliactis parasitica* (32) and in ectodermal photosensory cells in gonads of the hydrozoan *C. hemisphaerica* (38). Treatment of isolated sea anemone tentacles with synthetic RPamide changed the rates of spontaneous contractions, consistent with a role

in neurotransmission (32, 33). In addition, recent experimental evidence suggests that RPamide regulates oocyte maturation and spawning in hydrozoans (39). However, whether RPamide has any role during cnidarian development is not known.

In this paper, we have investigated the developmental expression pattern of RPamide in *N. vectensis* in order to build a neuroanatomical framework for understanding RPamide function during development. By using *in situ* hybridization and immunohistochemistry, we show that RPamide is dynamically expressed during *N. vectensis* development. RPamide expression begins at gastrulation in scattered epithelial cells in the aboral ectoderm. These RPamide-positive epithelial cells show a spindle-shaped, sensory cell-like morphology and extend basal neurites that form an aboral nerve net of the planula larva. A subset of the RPamide-positive sensory cells located at the aboral end become integrated into the developing apical organ. Later during planula development, RPamide-positive sensory cells become evident in oral ectoderm of the body column and the pharynx, as well as in the endoderm. At metamorphosis, RPamide-positive aboral sensory nerve net disintegrates by apoptosis, and RPamide becomes expressed in ectodermal sensory cells of growing oral tentacles. In addition, we find that expression of RPamide and that of the conserved neuropeptide GLWamide occurs in distinct subsets of neurons in planulae, suggesting that RPamide and GLWamide may perform non-overlapping functions in *N. vectensis* planulae.

MATERIALS AND METHODS

Animal Culture

N. vectensis were cultured as previously described (40, 41).

RNA Extraction, cDNA Synthesis, and Gene Cloning

Total RNA was extracted from a mixture of planulae and primary polyps using TRIzol (Thermo Fisher Scientific). cDNAs were synthesized using the BD SMARTer RACE cDNA Amplification Kit (Cat. No. 634858; BD Biosciences, San Jose, CA, USA). RPamide gene sequences [Nv37852 and Nv244953; (36)] were retrieved from the Joint Genome Institute genome database (*N. vectensis* v1.0; <http://genome.jgi-psf.org/Nemve1/Nemve1.home.html>). 5' and 3' RACE were conducted, following manufacturer's recommendations (BD SMARTer RACE cDNA Amplification Kit, BD Biosciences, San Jose, CA, USA), in order to confirm *in silico* predicted gene sequences. Gene specific primer sequences used for RACE PCR are: 3' RACE Forward 5'-GCTCGGTACAGAGCCGAAACCTGAGACAC-3'; 5' RACE primary PCR Forward 5'-CATGGGCAACGGTCAGCGGCAGATCGATG-3'. RACE PCR fragments were cloned into the pGEM-T plasmid vector using the pGEM-T Vector Systems (Cat. No. A3600; Promega), and were sequenced at Macrogen Corp., Maryland. To generate templates for RNA probes for *in situ* hybridization experiments, RTPCR was performed to amplify a 406 bp NvRPa (Nv244953) gene fragment. Gene specific primer sequences used for RTPCR are: Forward 5'-CGAAGGACCTTGAAAGTGGACTGTTCTCGG-3'; Reverse 5'-CATGGGCAACGGTCAGCGGCAGATCGATG-3'. The

PCR product was cloned into a pCR4-TOPO TA vector using the TOPO TA Cloning kit (Cat. No. K457501; ThermoFisher Scientific), and sequenced at Genewiz, New Jersey.

Generation of an Antibody Against RPamide

An antibody against a synthetic peptide CEDSSNYEFPPGFHRPamide corresponding in amino acid sequence to Nv-RPamide IV [*sensu* (36); **Figure 1**] was generated in rabbit (YenZym Antibodies, LLC); a recent mass spectrometry study has confirmed that the C-terminal FPPGFHRPamide is secreted by adult *N. vectensis* (34). The antibody was generated against the predicted Nv-RPamide IV sequence because mass spectrometry data were not available when the antibody was produced. Following immunization, the resultant antiserum was affinity purified with the CEDSSNYEFPPGFHRPamide peptide. The affinity purified antibody was then affinity-absorbed with KWSCSLRPamide and KWSCCLRPamide, which correspond to predicted RPamide peptides encoded by Nv37852—Nv-RPamide I and Nv-RPamide II, respectively [*sensu* (36)]—in order to generate a Nv-RPamide IV-specific antibody. Immunostaining with the antibody preadsorbed with the synthetic Nv-RPamide IV (CEDSSNYEFPPGFHRPamide) for 2 h at 37°C showed little immunoreactivity at the mid planula and primary polyp stages (**Supplementary Figure 4**), suggesting that the antibody reacts with Nv-RPamide IV. However, this antibody likely cross-reacts with other RPamide-like peptides, as we have observed a population of neurons that are immunoreactive with the antibody but do not express NvRPa transcripts in the tentacular ectoderm at the tentacle-bud and primary polyp stages (e.g., **Figure 4C**).

CRISPR-Cas9-Mediated Mutagenesis

20 nt-long sgRNA target sites were manually identified in the genomic locus for NvRPa. To minimize off-target effects, target sites that had 17 bp-or-higher sequence identity elsewhere in the genome (*N. vectensis* v1.0; <http://genome.jgi.doe.gov/Nemve1/Nemve1.home.html>) were excluded. Selected target sequences were as follows.

5'-ACTGATGGCGCTTGTGTGCT-3' (Cr1)

5'-CTATTCGGCTTTATTAGGTA-3' (Cr2)

5'-GGTATAACTATCCCGTTCCT-3' (Cr3)

The sgRNA species were synthesized *in vitro* (Synthego), and mixed at equal concentrations. The sgRNA mix and Cas9 endonuclease (PNA Bio, PC15111, Thousand Oaks, CA, USA) were co-injected into fertilized eggs at concentrations of 500 ng/μl and 1,000 ng/μl, respectively.

Genotyping of Embryos

Genomic DNA from single embryos was extracted by using a published protocol (24), and the targeted genomic locus was amplified by nested PCR. Primer sequences used for nested genomic PCR are: "1" Forward 5'-CGAAGGACCTTGAAAGTGGACTGTTCTCGG-3', "1" Reverse 5'-TGTCTGGGACTAGTTACCTACAGCG-3', "2" Forward 5'-GTGGTATGAGGCACAACGTAGATGG-3', "2" Reverse 5'-CATGGGCAACGGTCAGCGGCAGATCGATG-3'.

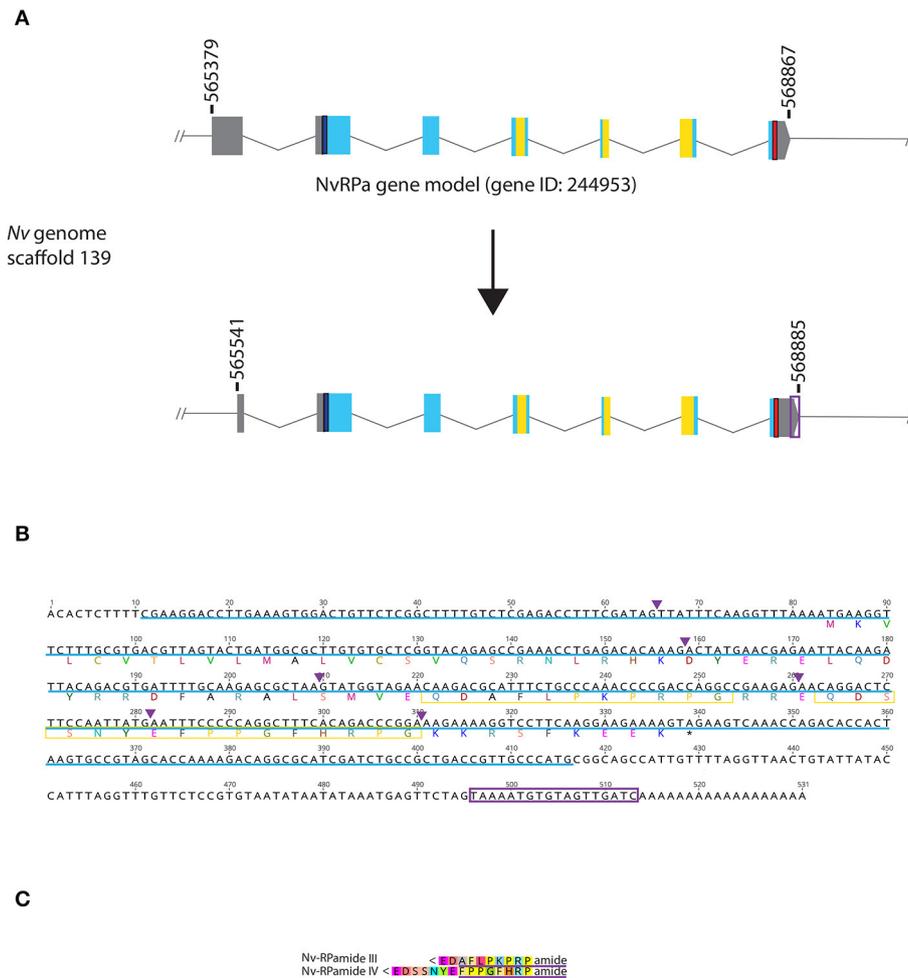


FIGURE 1 | Genomic organization, cDNA sequence, and predicted mature peptides of RPamide gene. **(A)** Schematic view of the NvRPa genomic locus Nv244953 located in scaffolds 139 of the *N. vectensis* genome (v.1.0; <http://genome.jgi.doe.gov/Nemve1/Nemve1.home.html>). Gene structure based on gene model prediction is shown at the top, and the revised structure based on the comparison of the *N. vectensis* genome and cDNA sequences is shown at the bottom. The number above the schematic indicate genomic coordinates within the scaffold. Gray boxes are untranslated regions; aqua-colored boxes are translated regions. Blue bars show predicted translation start sites; red bars show predicted translation termination sites; yellow bars show predicted RPamide-encoding regions described below. Note that the revised model has a shorter exon 1 and a longer exon 7 (boxed in purple) relative to the predicted gene model. **(B)** cDNA and deduced amino acid sequence of NvRPa. Short peptides predicted to be released by proteolytic cleavages in a previous study by Anciale (36) are boxed in yellow. Single or multiple acidic and basic residues are assumed to become cleaved as suggested in another sea anemone *Anthopleura elegantissima* (32, 33, 35). Purple arrowheads indicate intron positions. The sequence used as a template to generate riboprobes for *in situ* hybridization is highlighted in light blue. **(C)** Amino acid sequence alignment for RPamide, “Nv-RPamide III,” and “Nv-RPamide IV”; nomenclature of these peptides was adopted from a previous genomic study (36). It is assumed that upon release of the peptide, an N-terminal glutamine is spontaneously converted into a pyroglutamic acid residue (<Glu), and a C-terminal glycine becomes amidated by an α -amidating enzyme [cf. (35)]. A single precursor protein is predicted to generate one copy of each peptide. We note the possibility that, given the presence of additional acidic residues within regions predicted to encode Nv-RPamide III and IV—namely, aspartic acid in Nv-RPamide III, and glutamic acid in Nv-RPamide IV—the processed peptides could be shorter; Nv-RPamide III may have the sequence AFLPKPRFamide and Nv-RPamide IV FPPGFHRPamide (purple underlines), the latter of which has been recently confirmed to occur in adult *N. vectensis* using mass spectrometry (34).

Immunofluorescence, TUNEL and *in situ* Hybridization

Animal fixation, immunohistochemistry, and *in situ* hybridization were performed as previously described (14, 25). For immunohistochemistry, we used primary antibodies against RPamide (rabbit, 1:200), GLWamide [rabbit, 1:200; (5)], acetylated β -tubulin (mouse, 1:500, Sigma T6793) and tyrosinated β -tubulin (mouse, 1:500, Sigma T9028), and

secondary antibodies conjugated to AlexaFluor 568 [1:200, Molecular Probes A-11031 (anti-mouse) or A-11036 (anti-rabbit)] or AlexaFluor 647 [1:200, Molecular Probes A-21236 (anti-mouse) or A-21245 (anti-rabbit)]. Nuclei were labeled using fluorescent dyes DAPI (1:1,000, Molecular Probes D1306), and filamentous actin was labeled using AlexaFluor 488-conjugated phalloidin (1:25, Molecular Probes A12379). TUNEL assay was carried out after immunostaining, by using *in situ* Cell Death

Detection Kit (TMR red cat no. 1684795, Roche, Indianapolis, IN, USA) or Click-iT Plus TUNEL Assay for *in situ* Apoptosis Detection (Alexa Fluor 488, cat no. C10617, Molecular Probes) according to manufacturer's recommendation; both TUNEL assay kits showed similar results, and thus only data generated by using the Roche kit are reported. For *in situ* hybridization, antisense and sense digoxigenin-labeled riboprobes were synthesized by using the MEGascript transcription kits according to manufacturer's recommendation (Ambion; T7, AM1333; T3, AM1338), and were used at the final concentration of 1 ng/ μ l. Fluorescent images were recorded using a Leica SP5 Confocal Microscope or a Zeiss LSM900. Images were viewed using ImageJ.

RESULTS

It has been previously reported that the genome of the starlet sea anemone *N. vectensis* contains two RPamide-precursor genes [Nv37852 and Nv244953; (36)]. However, Reverse Transcriptase PCR by using planula and polyp cDNAs showed detectable expression for Nv244953, but not for Nv37852 (**Supplementary Figure 1**). Thus, we have focused our present analyses on Nv244953, which we will herein refer to as NvRPa (*Nematostella vectensis* RPamide). Consistent with *in silico* gene prediction ([https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=\\$Nemve1&id=\\$244953](https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=$Nemve1&id=$244953)), comparison of the *N. vectensis* genome and NvRPa cDNA sequence indicates that NvRPa is a 7-exon gene. We note that the cDNA sequence has a shorter exon 1 and a longer exon 7 relative to the *in silico* predicted gene model (**Figures 1A,B**); however, the shorter exon 1 could have resulted from failure to synthesize full-length cDNA.

Based on the location of diagnostic endoproteolytic cleavage sites (i.e., acidic residues, aspartic/glutamic acid; basic residues, lysine/arginine) and C-terminal amidation sites (i.e., glycine residue) [reviewed in (2)], NvRPa is predicted to encode single copies of two RPamide peptides that differ in N-terminal sequence [**Figure 1C**; Nv-RPamide III and IV, *sensu* (36)].

RPamide Is Expressed in the Aboral Sensory Nerve Net of the Planula Larva

Next we examined the spatiotemporal expression patterns of NvRPa during *N. vectensis* development. We used *in situ* hybridization and immunostaining to localize NvRPa transcripts and RPamide peptides, respectively. *In situ* hybridization using a sense probe did not result in cell-type-specific staining (**Supplementary Figure 2**). An antibody against RPamide peptides was made against Nv-RPamide IV predicted to be generated by NvRPa (see section Materials and Methods), and was validated by CRISPR-Cas9-mediated mutagenesis; a significantly reduced average number of anti-RPamide-immunoreactive neurons was observed in NvRPa F0 mosaic mutants (2; $n = 5$) relative to the wildtype control (20; $n = 5$) (two tailed *t*-test, $p < 0.0001$; **Supplementary Figure 3**). In addition, immunostaining with the antibody that was preadsorbed with Nv-RPamide IV peptides did not result in cell-type-specific staining (**Supplementary Figure 4**), indicating that the antibody indeed reacts with Nv-RPamide IV. *In situ*

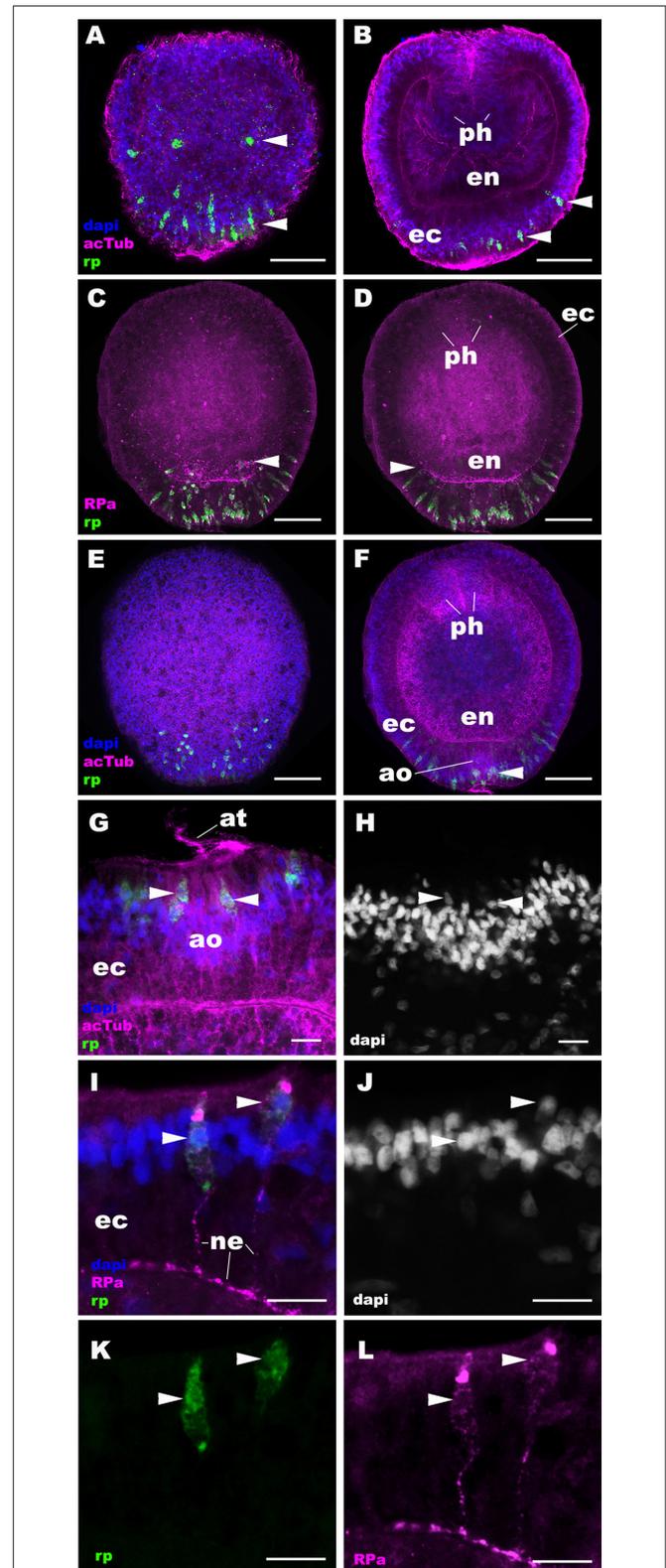


FIGURE 2 | RPamide-positive sensory epithelial cells form a nerve net in the aboral ectoderm during gastrulation and early planula development in a sea anemone. Z-projections of confocal sections of *Nematostella vectensis* at gastrula (**A,B**) and early planula (**C-L**) stages, labeled with an antisense
(Continued)

FIGURE 2 | riboprobe against NvRPa transcript (“rp”) and antibodies against RPamide (“RPa”) and acetylated β -tubulin (“acTub”). Nuclei are labeled with DAPI. **(A–F)** are side views of animals with the blastopore/oral opening facing up. **(A, C, E)** show superficial planes of section at the level of the ectoderm, and **(B, D, F)** show longitudinal sections through the center of the animal. **(G, H)** depict an apical organ (ao) and aboral-lateral ectoderm, respectively; apical surface is to the top. During gastrulation, scattered epithelial cells in the aboral ectoderm begin to express NvRPa transcripts (arrowheads in **A, B**), which show spindle-shaped morphologies characteristic of cnidarian sensory cells (19). At the early planula stage, NvRPa-expressing sensory cells extend basal RPamide-positive neurites forming an aboral basiepithelial nerve net (arrowheads in **C, D**), and a subset of NvRPa-expressing sensory cells become integrated into the developing apical organ (ao; arrowhead in **F**). Note that the cell bodies of NvRPa-expressing sensory cells are located in the superficial stratum of the apical organ (arrowheads in **G, H**) and of the lateral ectoderm (arrowheads in **I–L**). ph, pharynx; ec, ectoderm; en, endoderm; at, apical tuft; ne, neurite. Scale bar: 50 μ m (**A–F**); 10 μ m (**G–L**).

hybridization shows that NvRPa transcripts first occur at gastrulation in scattered epithelial cells of the aboral ectoderm, which exhibit a spindle-shaped morphology characteristic of cnidarian sensory cells (arrowheads in **Figures 2A, B**). During early planula development, these NvRPa-expressing epithelial cells extend RPamide-positive basal neuronal processes that form an interconnected basiepithelial network in the aboral ectoderm (**Figures 2C, D**). A subset of NvRPa-expressing sensory cells become integrated into the apical organ developing at the aboral end (**Figures 2E–H**). The cell bodies of RPamide-positive neurons are located in the superficial stratum of the ectoderm (arrowheads in **Figures 2G–L**). Consequently, within the apical organ, RPamide-positive sensory cells can be distinguished from apical-tuft cells based on the position of nuclei; the former have superficial nuclei, while the latter have basal nuclei [**Figures 2G, H**; (14)]. Thus, RPamide appears active in the aborally concentrated sensory nerve net of the planula larva.

At later stages of planula development, NvRPa becomes expressed in sensory cells of the oral body column ectoderm, pharyngeal ectoderm, as well as in the developing endodermal nervous system (**Figures 3A, B**). Some NvRPa-expressing sensory cells of the oral ectoderm extend basal neurites in an aboral direction (ne_j in **Figure 3C**), while others extend in a transverse orientation (ne_t in **Figure 3C**). NvRPa-expressing neurons in the pharyngeal ectoderm extend neurites preferentially in an aboral direction (**Figures 3D, E**). In the endoderm, NvRPa expression primarily occurs in neurons with longitudinally oriented neurites (**Figures 3D, F**), which become part of longitudinal neuronal bundles of the polyp endodermal nervous system (see below).

RPamide-Positive Ectodermal Nervous System Is Reorganized at Metamorphosis, Involving Removal of the Aboral Sensory Nerve Net by Apoptosis, and *de novo* Development of RPamide-Positive Tentacular Sensory Cells

At metamorphosis, the planula larva transforms into a polyp by developing tentacles around the mouth. During this process,

the NvRPa-expressing aboral sensory nerve net disappears (**Figures 4A, B**) along with the apical organ (14). Concomitantly, a subset of sensory cells in the tentacular ectoderm begin to express NvRPa (SC_t in **Figure 4C**). In addition, NvRPa expression is detected in some of the endodermal neurons with transverse neurites that connect with longitudinal neuronal bundles (blue arrowhead in **Figure 4D**). We note that at this developmental stage (i.e., tentacle bud and primary polyp stages) we observed a population of neurons that are immunoreactive with the antibody but do not express NvRPa transcripts in the tentacular ectoderm (**Figure 4C**); over 92% of RPa-immunoreactive neurons in oral tentacles ($n = 77$) did not show detectable levels of NvRPa expression. This suggests that RPamide-like peptides that are not generated by NvRPa are expressed in some of the tentacular ectodermal neurons. Alternatively, NvRPa could be expressed in these neurons at levels not detectable by *in situ* hybridization. The lack of colocalization of NvRPa transcripts and RPa-immunoreactivity is rarely found prior to metamorphosis.

We then examined the mechanism by which the NvRPa-expressing aboral sensory nerve net was removed at metamorphosis in *N. vectensis*. Specifically, we have considered the possibility that NvRPa-expressing aboral sensory cells undergoes programmed cell death at metamorphosis. To test this hypothesis, we labeled apoptotic DNA fragmentation at metamorphosis by using the terminal deoxynucleotidyl transferase dUTP nick end-labeling (TUNEL) assay. We have found that, indeed, NvRPa-expressing aboral sensory cells contain TUNEL-positive nuclei and/or nuclear fragments at the tentacle-bud stage (69.6%, $n = 46$; **Figure 5**). This finding supports the hypothesis that the NvRPa-expressing sensory nerve net in the aboral ectoderm of the planula larva disintegrates by apoptosis during metamorphosis in *N. vectensis*. We note that apoptosis in the aboral ectoderm at metamorphosis is not restricted to RPamide-positive neurons, but include other peptidergic neurons such as GLWamide- and RFamide-immunoreactive neurons (**Supplementary Figure 5**).

The Expression Pattern of RPamide Is Distinct From That of GLWamide

Next we asked whether the expression pattern of NvRPa was distinct from that of GLWamide—a neuropeptide thought to be conserved across Cnidaria and Bilateria, which regulates the timing of life cycle transition in *N. vectensis* (5). To address this, we combined *in situ* hybridization and immunostaining to examine the spatial localization of NvRPa transcripts and GLWamide peptides during planula development in *N. vectensis*. We found no evidence of colocalization of NvRPa and GLWamide expression in ectodermal and endodermal nervous systems in *N. vectensis* planulae (**Figure 6**). Thus, RPamide and GLWamide are expressed in non-overlapping subsets of neurons in the planula nervous systems in *N. vectensis*, suggesting that RPamide and GLWamide may perform non-overlapping neural functions in planulae. These data are also consistent with multiple peptidergic neuron types having complementary roles in specific biological processes such as metamorphosis.

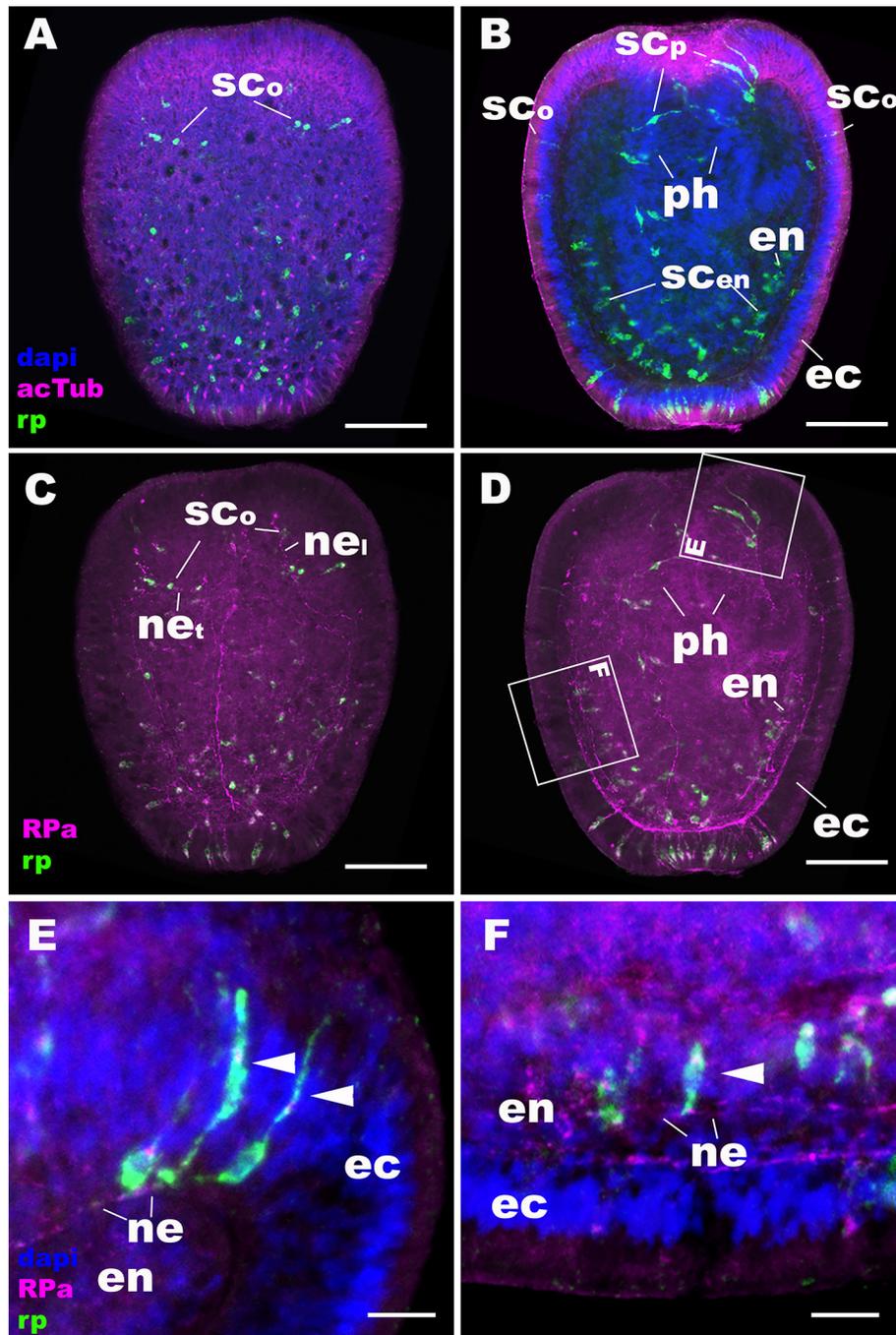


FIGURE 3 | During sea anemone planula development, RPamide becomes expressed in the nervous system of the oral body column ectoderm, pharyngeal ectoderm, and endoderm. Z-projections of confocal sections of *Nematostella vectensis* at mid-planula stages, labeled with an antisense riboprobe against NvRPa transcript ("rp") and antibodies against RPamide ("RPa") and acetylated β -tubulin ("acTub"). Nuclei are labeled with DAPI. **(A–D)** are side views of animals with the oral opening facing up. **(A,C)** show superficial planes of section at the level of the ectoderm; **(B,D)** show longitudinal sections through the center. **(E,F)** are magnified views of boxed regions in **(D)** showing NvRPa-expressing cells with their apical side facing up. During planula development, NvRPa becomes expressed in epithelial cells with the sensory-cell-like morphology and RPa-positive basal neurites. These NvRPa-expressing sensory cells reside in the body column oral ectoderm (sc_o in **A–C**), pharyngeal ectoderm (sc_p in **B**; arrowheads in **E**) and endoderm (sc_{en} in **B**; an arrowhead in **F**). Note in **C** that NvRPa-expressing sensory cells in the body column oral ectoderm can have aborally oriented longitudinal neurites (ne), or transverse neurites (ne_t). NvRPa-expressing sensory cells in the pharynx have aborally oriented longitudinal neurites (ne in **E**), and those in the endoderm send longitudinal neurites in both oral and aboral directions (ne in **F**). ph, pharynx; ec, ectoderm; en, endoderm. Scale bar: 50 μ m (**A–D**); 10 μ m (**E,F**).

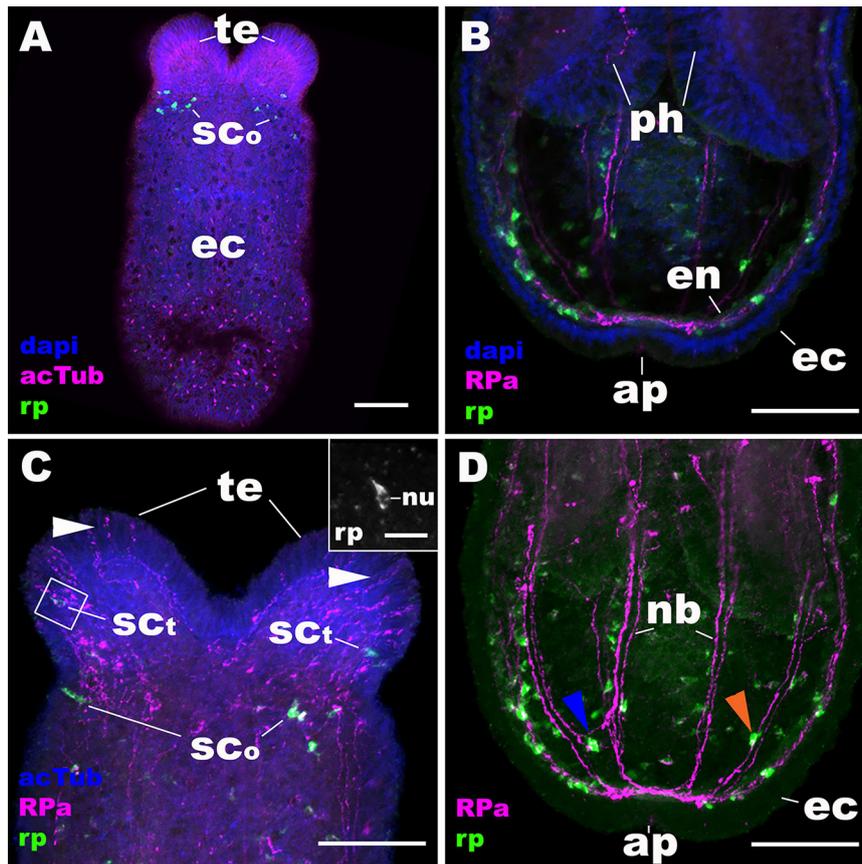


FIGURE 4 | During metamorphosis of a planula larva into a sea anemone polyp, the RPamide-positive sensory nerve net of the aboral ectoderm disappears, and RPamide becomes expressed in ectodermal sensory cells of developing oral tentacles. Z-projections of confocal sections of *Nematostella vectensis* at the tentacle-bud/primary polyp stage, labeled with an antisense riboprobe against NVRPa transcript (“rp”) and antibodies against RPamide (“RPa”) and acetylated β -tubulin (“acTub”). Nuclei are labeled with DAPI. All panels show side views of the animal with the oral opening facing up. **(A,C)** show superficial planes of section at the level of surface ectoderm (ec); **(B,D)** show longitudinal sections through the center. **(B,D)** depict an aboral half of the animal, while **(C)** shows an oral half of the animal. The inset in **(C)** is a magnified view of the boxed region showing a NVRPa-expressing cell with their apical side facing up; note that NVRPa transcripts primarily localize to the cytoplasm so that the nucleus (nu) appears NVRPa-negative. At metamorphosis, RPamide expression becomes undetectable in the aboral ectoderm at the transcript and peptide levels (compare **A,B** with **Figures 2A–D**). Concomitantly, NVRPa transcripts become expressed in ectodermal sensory cells of the developing oral tentacles that send RPamide-positive longitudinal neurites in an aboral direction (sc_t in **C**); note in the inset in **(C)** the sensory cell-like spindle-shaped morphology of an NVRPa-expressing tentacular cell. White arrowheads in **(C)** show anti-RPamide-immunoreactive sensory cells in the tentacular ectoderm that do not express detectable levels of NVRPa transcripts, which suggests that the immunoreactive materials in these cells are not produced by NVRPa. Note in **(D)** that RPamide expression is evident in endodermal neurons that are integrated within longitudinal neuronal bundles (nb) (orange arrowhead), as well as in those that reside between neighboring neuronal bundles (blue arrowhead); neurites from the latter neurons connect with those from neurons in the neuronal bundles. te, tentacle; sc_o, sensory cell in the oral body column ectoderm; ph, pharynx; en, endoderm; ap, aboral pole. Scale bar: 50 μ m (**A–D**); 10 μ m (inset in **C**).

DISCUSSION

In this paper, we reported the expression pattern of a cnidarian-specific neuropeptide, RPamide, during development of the starlet sea anemone *N. vectensis*. We have found that early RPamide-positive neurons form a sensory nerve net in the aboral ectoderm of the planula larva, and that a subset of RPamide-positive sensory cells located at the aboral end become part of the apical organ. During planula development, RPamide becomes expressed in sensory cells of oral ectoderm in the body column and pharynx, as well as in endodermal neurons. At metamorphosis, the RPamide-positive sensory nerve net in aboral ectoderm disappears by apoptosis, and RPamide becomes

expressed in a subset of ectodermal sensory cells in growing oral tentacles. RPamide expression and GLWamide expression do not co-occur in the same neurons in planulae, indicative of functional segregation between the two neuropeptides.

The Sea Anemone Apical Organ Consists of Multiple Cell Types, Including Apical Tuft Cells and Peptidergic Sensory Cells

Our study sheds light on the diversity of cell types that constitute the apical organ in a sea anemone planula. Chia and Koss (30) used electron microscopy to describe the structure of an apical organ in the sea anemone *A. elegantissima*. They reported that

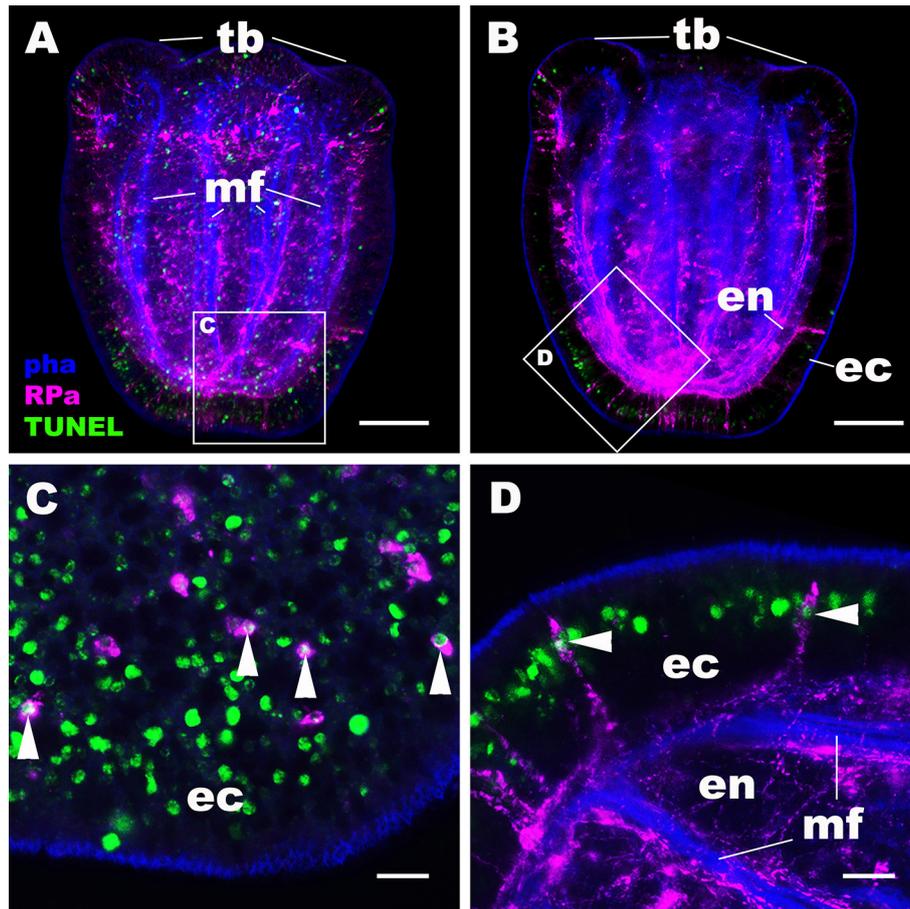


FIGURE 5 | RPamide-positive sensory cells of the aboral ectoderm undergo apoptosis at metamorphosis. Z-projections of confocal sections of *Nematostella vectensis* at the tentacle-bud stage, labeled with an antibody against RPamide (“RPa”). Filamentous actin is labeled with a fluorescent dye phalloidin (“pha”), and DNA fragmentation is detected by TUNEL assay (“TUNEL”). (A,B) are side views of the animal with the oral opening facing up; (A) represents a superficial plane of section at the level of surface ectoderm and parts of the endoderm, while B depicts a longitudinal section through the center. (C,D) are sections from boxed areas in (A,B), respectively; note in D that the orientation has been altered relative to B so that the apical surface of the ectoderm faces up. Arrowheads in (C,D) point to TUNEL-positive DNA fragmentation in RPamide-positive sensory cells of the aboral ectoderm, evidencing that these cells are undergoing programmed cell death. tb, tentacle bud; mf, endodermal muscle fibers; ec, ectoderm; en, endoderm. Scale bar: 50 μm (A,B); 10 μm (C,D).

the apical organ consists of columnar epithelial cells, each of which has a long cilium and a collar of microvilli on the apical cell surface and a basally localized nucleus. Long cilia from the columnar epithelial cells form a ciliary bundle known as the apical tuft, and their ciliary rootlet characteristically reach the basal processes of the cells. In addition, Nakanishi et al. (14) identified two types of gland cells in the apical organ of *N. vectensis*: one with electron-lucent granules and the other with electron-dense granules. In this study, we have discovered that a subset of early-born RPamide-positive sensory cells in the aboral ectoderm become part of the apical organ in *N. vectensis*. RPamide-positive sensory cells of the apical organ are morphologically distinct from other cell types that are housed in this structure. Specifically, in contrast to the apical-tuft epithelial cells that are column-shaped and have basal nuclei, these RPamide-positive sensory cells are spindle-shaped and contain nuclei located close to the epithelial surface. To our

knowledge, this is the first evidence that the sea anemone apical organ contains a peptidergic sensory cell type. Thus, the apical organ is composed of at least three cell types: apical-tuft cells, gland cells, and peptidergic sensory cells.

Motivated by the question of whether the aboral domain of cnidarians is homologous to the anterior domain of bilaterians, several studies investigated the molecular mechanism of apical organ development in *N. vectensis* [e.g., (42–44)]. It has been proposed that apical organ development in *N. vectensis* occurs in two phases, involving a set of regulatory factors—*Six3*, *FoxQ2* and *FGFs*—whose bilaterian homologs play a conserved role in controlling anterior development (43). First, during gastrulation, *six3/6* specifies the identity of the broad aboral domain in the ectoderm by positively regulating the expression of *fgfa1* and *foxQ2a*. Subsequently, during planula development, *FGFa1* signaling drives the differentiation of apical organ cells by downregulating *six3/6* and *foxQ2a* expression at the aboral pole.

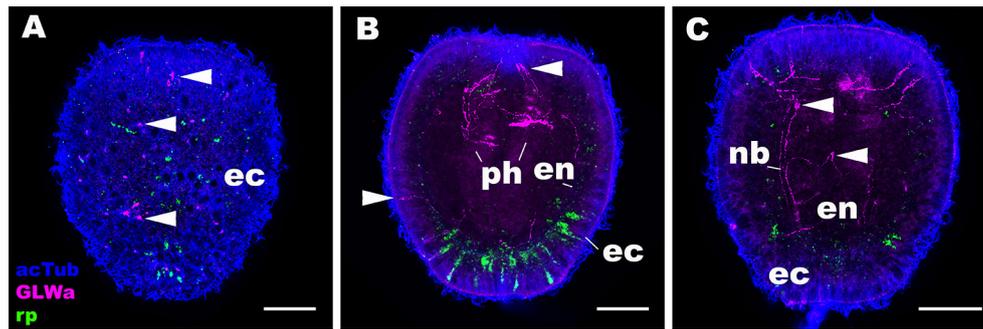


FIGURE 6 | RPamide and GLWamide are expressed in distinct, non-overlapping subsets of sensory cells in the ectoderm of a sea anemone planula larva. Z-projections of confocal sections of *Nematostella vectensis* at the mid-planula stage, labeled with an antisense riboprobe against NvRPa transcript (“rp”) and antibodies against GLWamide (“GLWa”) and acetylated β -tubulin (“acTub”). All panels show side views of animals with the oral opening facing up. **(A)** shows a superficial plane of section at the level of surface ectoderm (ec). **(B,C)** Show longitudinal sections at the level of pharynx (ph) and longitudinal neuronal bundles (nb), respectively. Arrowheads point to GLWamide-positive neurons that are NvRPa-negative, which include sensory cells in the body column ectoderm **(A,B)**, in pharyngeal ectoderm **(B)**, and in endoderm (en; **C**). Note also that none of the NvRPa-expressing cells are GLWamide-positive. Scale bar: 50 μ m.

Consistent with this model, experimental perturbation of *six3/6* and FGf1 expression results in planula larvae without apical tufts (42, 43). However, it has not been clear whether all apical organ cell types—besides the apical-tuft cells—develop in this way. In the present study, we have found that development of RPamide-positive sensory cells at the aboral end occurs during gastrulation, and precedes the mid-planula stage at which downregulation of *six3/6* is reported to occur (43). This suggests that neurogenesis that generates apical organ neurons takes place prior to differentiation of apical tuft cells, and therefore does not depend on downregulation of *six3/6* at the aboral pole. Our data therefore imply that neurogenesis and ciliogenesis in apical organ development are mechanistically decoupled. Consistent with this hypothesis, positive regulators of neurogenesis, such as *soxB(2)*, *ath*, and *ashA*, are expressed in the aboral domain during gastrulation, prior to apical tuft formation (29, 45–47). How the mechanisms of neurogenesis and ciliogenesis in the context of sea anemone apical organ development differ from each other, and how they relate to anterior development of bilaterians, are to be addressed in future studies.

Reorganization of the Ectodermal Nervous System at Metamorphosis Represents an Ancestral Condition in Cnidaria

Nervous system development in the sea anemone *N. vectensis* has been viewed as a gradual process that entails progressive addition of new features; neurons begin development in ectoderm at gastrulation, then in endoderm during planula development, and finally in the polyp tentacular ectoderm at metamorphosis (14, 15). Only apical tuft cells have been reported to disappear at metamorphosis in *N. vectensis* (14). In contrast, reorganization of peptidergic ectodermal nervous systems at the planula-polyp transition is commonly observed in non-sea anemone cnidarians such as the coral *Acropora* (48), moonjelly *Aurelia* (49), and hydrozoans [*Pennaria disticha* (formerly *Pennaria tiarella*) (50); *Hydractinia echinate* (51,

52); *Clava multicornis* (53)]. This process typically involves disappearance of the planula peptidergic nervous system—housed exclusively in the ectoderm in these cnidarians—likely via apoptosis (49, 53–56), followed by the development of the orally concentrated polyp nervous system. Therefore, reorganization of the ectodermal nervous system at metamorphosis appears to be an ancestral condition in Cnidaria. In this paper, we have discovered that the early-born RPamide-positive sensory nerve net of the aboral ectoderm is removed at metamorphosis by programmed cell death. We also note that early-born GLWamide- and RFamide-positive sensory cells in the aboral ectoderm similarly disintegrate by undergoing DNA fragmentation at metamorphosis (Figure 5). Hence, we suggest that sea anemones have in fact retained the ancestral pattern of cnidarian neural development to undergo reorganization of the peptidergic ectodermal nervous system at the planula-polyp transition.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

NN conceived and designed the study. HZ and NN carried out the experiments and drafted the manuscript.

FUNDING

This work was supported by funds from Arkansas IDeA Network of Biomedical Research Excellence (INBRE) (to HZ and NN), Arkansas Bioscience Institute (to NN), and the University of Arkansas (to NN).

ACKNOWLEDGMENTS

We thank Ethan Ozment for animal husbandry, and Arianna Tamvacakis for comments on an earlier version of the manuscript. We are also grateful to Mark Q. Martindale for his encouragement to explore every neuropeptide encoded by the *Nematostella* genome, which was the impetus for this study. We would also like to thank reviewers for comments on the earlier version of the manuscript, which greatly improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2020.00063/full#supplementary-material>

Supplemental Figure 1 | Reverse Transcriptase (RT) PCR of RPamide-encoding genes. cDNA PCR results showing the presence of Nv244953 transcripts (expected size: 291 bp) and the absence of Nv37852 transcripts (expected size: 207 bp). “–” denotes a negative control without cDNA. 100 bp ladders were used. cDNA was generated from RNA extracted from planulae and primary polyps. Primer sequences used for PCR are: “244953” Forward 5'-GCTCGGTACAGAGCCGAAACCTGAGACAC-3', Reverse 5'-CATGGGCAACGGTCAGCGGCAGATCGATG-3', “37852” Forward 5'-GCCTTGGTCGTGTTGCCTTCGCCCTGGTC-3', “2” Reverse 5'-CTATACGACCAGGGCGAGGCTACACGACC-3'.

Supplemental Figure 2 | *In situ* hybridization using a sense riboprobe. Z-projections of confocal sections of *Nematostella vectensis* at mid-planula (A,B) and tentacle-bud/primary polyp (C,D) stages, labeled with a sense NvRPa riboprobe (“sense probe”) and an antibody against acetylated β -tubulin (“acTub”). All panels are side views of animals with the oral opening facing up. (A,C) show superficial planes of section at the level of the ectoderm; (B,D) show longitudinal sections through the center. No cell-type-specific staining is observable. Scale bar: 50 μ m.

Supplemental Figure 3 | CRISPR mutagenesis and validation of an anti-RPamide antibody. (A) Schematic view of the NvRPa locus. (B) Genomic DNA nested PCR results of Cas9-injected wildtype control embryos (“WT”) and F0 embryos injected with NvRPa locus-specific sgRNAs together with Cas9 (“RPa”).

REFERENCES

- Strand FL. *Neuropeptides: Regulators of Physiological Processes*. Cambridge: MIT Press (1999). doi: 10.7551/mitpress/4950.001.0001
- Grimmelikhuijzen CJP, Williamson M, Hansen GN. Neuropeptides in cnidarians. *Can J Zool.* (2002) 80:1690–702. doi: 10.1139/z02-137
- Grimmelikhuijzen CJP, Hauser F. Mini-review: The evolution of neuropeptide signaling. *Regul Pept.* (2012) 177:S6–9. doi: 10.1016/j.regpep.2012.05.001
- Conzelmann M, Williams EA, Tunaru S, Randel N, Shahidi R, Asadulina A, et al. Conserved MIP receptor-ligand pair regulates *Platynereis* larval settlement. *Proc Natl Acad Sci USA.* (2013) 110:8224–9. doi: 10.1073/pnas.1220285110
- Nakanishi N, Martindale MQ. CRISPR knockouts reveal an endogenous role for ancient neuropeptides in regulating developmental timing in a sea anemone. *Elife.* (2018) 7:e39742. doi: 10.7554/eLife.39742
- Schoofs L, Beets I. Neuropeptides control life-phase transitions. *Proc Natl Acad Sci USA.* (2013) 110:7973–4. doi: 10.1073/pnas.1305724110
- Medina M, Collins AG, Silberman JD, Sogin ML. Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proc Natl Acad Sci USA.* (2001) 98:9707–12. doi: 10.1073/pnas.171316998
- Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science.* (2007) 317:86–94. doi: 10.1126/science.1139158
- Hejnal A, Obst M, Stamatakis A, Ott M, Rouse GW, Edgecombe GD, et al. Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc Biol Sci.* (2009) 276:4261–70. doi: 10.1098/rspb.2009.0896
- Chen JY, Oliveri P, Gao F, Dornbos SQ, Li CW, Bottjer DJ, et al. Precambrian animal life: Probable developmental and adult cnidarian forms from Southwest China. *Dev Biol.* (2002) 248:182–96. doi: 10.1006/dbio.2002.0714
- Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science.* (2011) 334:1091–7. doi: 10.1126/science.1206375
- Collins AG, Schuchert P, Marques AC, Jankowski T, Medina M, Schierwater B. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. *Syst Biol.* (2006) 55:97–115. doi: 10.1080/10635150500433615
- Zapata F, Goetz FE, Smith SA, Howison M, Siebert S, Church SH, et al. Phylogenomic analyses support traditional relationships within cnidaria. *PLoS ONE.* (2015) 10:e0139068. doi: 10.1101/017632
- Nakanishi N, Renfer E, Technau U, Rentsch F. Nervous systems of the sea anemone *Nematostella vectensis* are generated by ectoderm and

Results of secondary PCR are shown. (C,D) Medial-level confocal sections of Cas9-injected wildtype control (C) and NvRPa F0 mutant (D) mid-planulae at 4 day-post-fertilization (dpf), labeled with an anti-RPamide (“RPa”). Nuclei are labeled with DAPI. Oral side is up. (E) A histogram showing the number of anti-RPamide-positive neurons in Cas9-injected wildtype control animals (“WT”; $n = 5$) and NvRPa F0 mutants (“CR-RPa”; $n = 5$) at the 4 dpf mid-planula stage. In (A), a blue bar shows a predicted translation start site; a red bar shows a predicted translation termination site; yellow bars show predicted RPamide-encoding regions as in Figure 1. Purple arrowheads show sgRNA target sites. Blue arrows mark regions targeted in the PCR analysis shown in (B). Note in B that genomic PCR of the WT embryo shows the expected size of PCR fragments (1690 bp for secondary nested PCR), while F0 mutant embryos show additional bands of smaller sizes (arrowhead), indicating that targeted deletions of different sizes have occurred mosaically in each embryo. Over 88% of PCR genotyped embryos showed clear mutant bands (i.e., PCR fragments that are smaller than expected from a wildtype allele; $n = 34$). DNA sequencing of mutant bands (e.g., arrowhead) has confirmed that excision of RPamide-encoding regions can be induced by this CRISPR-mediated mutagenesis approach. In (C,D), arrowheads show anti-RPa-positive neurons. Note in (D) that an anti-RPa-positive cell is observed in the endoderm (en) but not in the ectoderm (ec). In (E), * denotes a statistically significant difference ($\alpha = 0.05$). Scale bar: 50 μ m.

Supplemental Figure 4 | Immunostaining with a preadsorbed anti-RPamide antibody. Z-projections of confocal sections of *Nematostella vectensis* at mid-planula (A,C) and primary polyp (B,D) stages, labeled with an anti-RPamide antibody (“anti-RPa”; A,B) and an anti-RPamide antibody preadsorbed with Nv-RPamide IV (CEDSSNYEFPFGFHRPamide) (“Preadsorbed anti-RPa”; C,D). All panels are side views of animals with the oral opening facing up. (A,C) show longitudinal sections through the center; (B,D) show sections through the entire animals. No cell-type-specific staining is observable with the preadsorbed antibody, indicating that the anti-RPamide indeed reacts with Nv-RPamide IV. Scale bar: 50 μ m.

Supplemental Figure 5 | GLWamide- and RFamide-immunoreactive sensory cells of the aboral ectoderm undergo apoptosis at metamorphosis. Z-projections of confocal sections of *Nematostella vectensis* at the tentacle-bud stage, labeled with antibodies against acetylated β -tubulin (“acTub”) and GLWamide (“GLWa”; (5)) or RFamide (“Rfa”; unpublished). DNA fragmentation is detected by TUNEL assay (“TUNEL”). All panels show side views of the animal with the oral opening facing up; (A,C) represent superficial planes of section at the level of surface ectoderm, while (B,D) depict longitudinal sections near the center. Boxed areas are magnified in insets; note that in insets in (B,D) the apical surface of the ectoderm faces up. Arrowheads point to TUNEL-positive DNA fragmentation, evidencing programmed cell death. Scale bar: 50 μ m.

- endoderm and shaped by distinct mechanisms. *Development*. (2012) 139:347–57. doi: 10.1242/dev.071902
15. Marlow HQ, Srivastava M, Matus DQ, Rokhsar D, Martindale MQ. Anatomy and development of the nervous system of *Nematostella vectensis*, an anthozoan cnidarian. *Dev Neurobiol*. (2009) 69:235–54. doi: 10.1002/dneu.20698
 16. Lentz TL, Barnett RJ. Fine structure of the nervous system of hydra. *Am Zool*. (1965) 5:341–56. doi: 10.1093/icb/5.3.341
 17. Carlberg M, Alfredsson K, Nielsen SO, Anderson PAV. Taurine-like immunoreactivity in the motor nerve net of the jellyfish *Cyanea capillata*. *Biol Bull*. (1995) 188:78–82. doi: 10.2307/1542069
 18. Nordstrom K, Wallen R, Seymour J, Nilsson D. A simple visual system without neurons in jellyfish larvae. *Proc R Soc London Ser B Biol Sci*. (2003) 270:2349–54. doi: 10.1098/rspb.2003.2504
 19. Thomas M, Edwards N. Cnidaria: Hydrozoa. In: Harrison FW, editor. *Microscopic Anatomy of Invertebrates*. New York, NY: Wiley-Liss (1991). p. 91–183.
 20. Koizumi O, Itazawa M, Mizumoto H, Minobe S, Javois LC, Grimmelikhuijzen CJP, Bode HR. Nerve ring of the hypostome in hydra. 1. Its structure, development, and maintenance. *J Comp Neurol*. (1992) 326:7–21. doi: 10.1002/cne.903260103
 21. Skogh C, Garm A, Nilsson DE, Ekstrom P. Bilaterally symmetrical rhopalial nervous system of the box jellyfish *Tripedalia cystophora*. *J Morphol*. (2006) 267:1391–405. doi: 10.1002/jmor.10472
 22. Helm RR, Siebert S, Tulin S, Smith J, Dunn CW. Characterization of differential transcript abundance through time during *Nematostella vectensis* development. *BMC Genomics*. (2013) 14:266. doi: 10.1186/1471-2164-14-266
 23. Tulin S, Aguiar D, Istrail S, Smith J. A quantitative reference transcriptome for *Nematostella vectensis* early embryonic development: a pipeline for *de novo* assembly in emerging model systems. *Evodevo*. (2013) 4:16. doi: 10.1186/2041-9139-4-16
 24. Ikmi A, McKinney SA, Delventhal KM, Gibson MC. TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan *Nematostella vectensis*. *Nat Commun*. (2014) 5:5486. doi: 10.1038/ncomms6486
 25. Martindale MQ, Pang K, Finnerty JR. Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa). *Development*. (2004) 131:2463–74. doi: 10.1242/dev.01119
 26. Kraus Y, Technau U. Gastrulation in the sea anemone *Nematostella vectensis* occurs by invagination and immigration: an ultrastructural study. *Dev Genes Evol*. (2006) 216:119–32. doi: 10.1007/s00427-005-0038-3
 27. Magie CR, Daly M, Martindale MQ. Gastrulation in the cnidarian *Nematostella vectensis* occurs via invagination not ingression. *Dev Biol*. (2007) 305:483–97. doi: 10.1016/j.ydbio.2007.02.044
 28. Fritzenwanker JH, Genikhovich G, Kraus Y, Technau U. Early development and axis specification in the sea anemone *Nematostella vectensis*. *Dev Biol*. (2007) 310:264–79. doi: 10.1016/j.ydbio.2007.07.029
 29. Richards GS, Rentzsch F. Transgenic analysis of a SoxB gene reveals neural progenitor cells in the cnidarian *Nematostella vectensis*. *Development*. (2014) 141:4681–9. doi: 10.1242/dev.112029
 30. Chia FS, Koss R. Fine-Structural studies of the nervous-system and the apical organ in the planula larva of the sea-anemone *Anthopleura-elegantissima*. *J Morphol*. (1979) 160:275–97. doi: 10.1002/jmor.1051600303
 31. Sebe-Pedros A, Saudemont B, Chomsky E, Plessier F, Mailhe MP, Renno J, et al. Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-Seq. *Cell*. (2018) 173:1520–34.e20. doi: 10.1016/j.cell.2018.05.019
 32. Carstensen K, Rinehart KL, Mcfarlane ID, Grimmelikhuijzen CJP. Isolation of Leu-Pro-Pro-Gly-Pro-Leu-Pro-Argpro-Nh2 (Antho-Rpamide), an N-terminally protected, biologically-active neuropeptide from sea-Anemones. *Peptides*. (1992) 13:851–7. doi: 10.1016/0196-9781(92)90040-A
 33. Carstensen K, Mcfarlane ID, Rinehart KL, Hudman D, Sun FR, Grimmelikhuijzen CJP. Isolation of Less-Than-Glu-Asn-Phe-His-Leu-Arg-Pro-Nh2 (Antho-Rpamide-Ii), a novel, biologically-active neuropeptide from sea-anemones. *Peptides*. (1993) 14:131–5. doi: 10.1016/0196-9781(93)90020-H
 34. Hayakawa E, Watanabe H, Menschaert G, Holstein TW, Baggerman G, Schoofs L. A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*. *PLoS ONE*. (2019) 14:e0215185. doi: 10.1371/journal.pone.0215185
 35. Grimmelikhuijzen CJP, Leviev IK, Carstensen K. Peptides in the nervous systems of cnidarians: Structure, function, and biosynthesis. *Int Rev Cytol*. (1996) 167:37–89. doi: 10.1016/S0074-7696(08)61345-5
 36. Antcil M. Chemical transmission in the sea anemone *Nematostella vectensis*: a genomic perspective. *Comp Biochem Physiol Part D Genomics Proteomics*. (2009) 4:268–89. doi: 10.1016/j.cbd.2009.07.001
 37. Koch TL, Grimmelikhuijzen CJP. Global neuropeptide annotations from the genomes and transcriptomes of cubozoa, scyphozoa, staurozoa (Cnidaria: Medusozoa), and Octocorallia (Cnidaria: Anthozoa). *Front Endocrinol*. (2019) 10:831. doi: 10.3389/fendo.2019.00831
 38. Artigas GQ, Lapebie P, Leclere L, Takeda N, Deguchi R, Jekely G, et al. A gonad-expressed opsin mediates light-induced spawning in the jellyfish *Clytia*. *Elife*. (2018) 7:e29555. doi: 10.7554/eLife.29555
 39. Takeda N, Kon Y, Artigas GQ, Lapebie P, Barreau C, Koizumi O, et al. Identification of jellyfish neuropeptides that act directly as oocyte maturation-inducing hormones. *Development*. (2018) 145:dev156786. doi: 10.1101/140160
 40. Fritzenwanker JH, Technau U. Induction of gametogenesis in the basal cnidarian *Nematostella vectensis*(Anthozoa). *Dev Genes Evol*. (2002) 212:99–103. doi: 10.1007/s00427-002-0214-7
 41. Hand C, Uhlinger KR. The culture, sexual and asexual reproduction, and growth of the sea-anemone *Nematostella-vectensis*. *Biol Bull*. (1992) 182:169–76. doi: 10.2307/1542110
 42. Rentzsch F, Fritzenwanker JH, Scholz CB, Technau U. FGF signalling controls formation of the apical sensory organ in the cnidarian *Nematostella vectensis*. *Development*. (2008) 135:1761–9. doi: 10.1242/dev.020784
 43. Sinigaglia C, Busengdal H, Leclere L, Technau U, Rentzsch F. The bilaterian head patterning gene *six3/6* controls aboral domain development in a cnidarian. *PLoS Biol*. (2013) 11:e1001488. doi: 10.1371/journal.pbio.1001488
 44. Sinigaglia C, Busengdal H, Lerner A, Oliveri P, Rentzsch F. Molecular characterization of the apical organ of the anthozoan *Nematostella vectensis*. *Dev Biol*. (2015) 398:120–33. doi: 10.1016/j.ydbio.2014.11.019
 45. Layden MJ, Boekhout M, Martindale MQ. *Nematostella vectensis* achaete-scute homolog *NvashA* regulates embryonic ectodermal neurogenesis and represents an ancient component of the metazoan neural specification pathway. *Development*. (2012) 139:1013–22. doi: 10.1242/dev.073221
 46. Richards GS, Rentzsch F. Regulation of *Nematostella* neural progenitors by SoxB, Notch and bHLH genes. *Development*. (2015) 142:3332–42. doi: 10.1242/dev.123745
 47. Magie CR, Pang K, Martindale MQ. Genomic inventory and expression of Sox and Fox genes in the cnidarian *Nematostella vectensis*. *Dev Genes Evol*. (2005) 215:618–30. doi: 10.1007/s00427-005-0022-y
 48. Attenborough RMF, Hayward DC, Wiedemann U, Foret S, Miller DJ, Ball EE. Expression of the neuropeptides RFamide and LWamide during development of the coral *Acropora millepora* in relation to settlement and metamorphosis. *Dev Biol*. (2019) 446:56–67. doi: 10.1016/j.ydbio.2018.11.022
 49. Nakanishi N, Yuan D, Jacobs DK, Hartenstein V. Early development, pattern, and reorganization of the planula nervous system in *Aurelia* (Cnidaria, Scyphozoa). *Dev Genes Evol*. (2008) 218:511–24. doi: 10.1007/s00427-008-0239-7
 50. Martin VJ. Reorganization of the nervous system during metamorphosis of a hydrozoan planula. *Invertebr Biol*. (2000) 119:243–53. doi: 10.1111/j.1744-7410.2000.tb00011.x
 51. Plickert G. Proportion-altering factor (Paf) stimulates nerve-cell formation in hydractinia-echinata. *Cell Diff Dev*. (1989) 26:19–27. doi: 10.1016/0922-3371(89)90780-6
 52. Schmich J, Trepel S, Leitz T. The role of GLWamides in metamorphosis of *Hydractinia echinata*. *Dev Genes Evol*. (1998) 208:267–73. doi: 10.1007/s004270050181
 53. Pennati R, Dell'Anna A, Pagliara P, Scari G, Piraino S, De Bernardi F. Neural system reorganization during metamorphosis in the planula larva of *Clava multicornis* (Hydrozoa, Cnidaria). *Zoomorphology*. (2013) 132:227–37. doi: 10.1007/s00435-013-0188-1

54. Yuan D, Nakanishi N, Jacobs DK, Hartenstein V. Embryonic development and metamorphosis of the scyphozoan Aurelia. *Dev Genes Evol.* (2008) 218:525–39. doi: 10.1007/s00427-008-0254-8
55. Seipp S, Wittig K, Stiening B, Bottger A, Leitz T. Metamorphosis of *Hydractinia echinata* (Cnidaria) is caspase-dependent. *Int J Dev Biol.* (2006) 50:63–70. doi: 10.1387/ijdb.052012ss
56. Seipp S, Schmich J, Will B, Schetter E, Plickert G, Leitz T. Neuronal cell death during metamorphosis of *Hydractinia echinata* (Cnidaria, Hydrozoa). *Invertebr Neurosci.* (2010) 10:77–91. doi: 10.1007/s10158-010-0109-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Zang and Nakanishi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Expression Profiling, Downstream Signaling, and Inter-subunit Interactions of GPA2/GPB5 in the Adult Mosquito *Aedes aegypti*

David A. Rocco*† and Jean-Paul V. Paluzzi*†

Laboratory of Integrative Vector Neuroendocrinology, Department of Biology, York University, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Elizabeth Amy Williams,
University of Exeter, United Kingdom

Reviewed by:

Frank Hauser,
University of Copenhagen, Denmark
Mark R. Brown,
University of Georgia, United States

*Correspondence:

David A. Rocco
davrocco@yorku.ca
Jean-Paul V. Paluzzi
paluzzi@yorku.ca

†ORCID:

David A. Rocco
orcid.org/0000-0001-9545-7713
Jean-Paul V. Paluzzi
orcid.org/0000-0002-7761-0590

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 18 December 2019

Accepted: 06 March 2020

Published: 31 March 2020

Citation:

Rocco DA and Paluzzi J-PV (2020)
Expression Profiling, Downstream
Signaling, and Inter-subunit
Interactions of GPA2/GPB5 in the
Adult Mosquito *Aedes aegypti*.
Front. Endocrinol. 11:158.
doi: 10.3389/fendo.2020.00158

GPA2/GPB5 and its receptor constitute a glycoprotein hormone-signaling system native to the genomes of most vertebrate and invertebrate organisms. Unlike the well-studied gonadotropins and thyrotropin, the exact function of GPA2/GPB5 remains elusive, and whether it elicits its functions as heterodimers, homodimers or as independent monomers remains unclear. Here, the glycoprotein hormone signaling system was investigated in adult mosquitoes, where GPA2 and GPB5 subunit expression was mapped and modes of its signaling were characterized. In adult *Aedes aegypti* mosquitoes, GPA2 and GPB5 transcripts co-localized to bilateral pairs of neuroendocrine cells, positioned within the first five abdominal ganglia of the central nervous system. Unlike GPA2/GPB5 homologs in human and fly, GPA2/GPB5 subunits in *A. aegypti* lacked evidence of heterodimerization. Rather, cross-linking analysis to determine subunit interactions revealed *A. aegypti* GPA2 and GPB5 subunits may form homodimers, although treatments with independent subunits did not demonstrate receptor activity. Since mosquito GPA2/GPB5 heterodimers were not evident by heterologous expression, a tethered fusion construct was generated for expression of the subunits as a single polypeptide chain to mimic heterodimer formation. Our findings revealed *A. aegypti* LGR1 elicited constitutive activity with elevated levels of cAMP. However, upon treatment with recombinant tethered GPA2/GPB5, an inhibitory G protein (Gi/o) signaling cascade is initiated and forskolin-induced cAMP production is inhibited. These results further support the notion that heterodimerization is a requirement for glycoprotein hormone receptor activation and provide novel insight to how signaling is achieved for GPA2/GPB5, an evolutionary ancient neurohormone.

Keywords: GPA2/GPB5, mosquito, glycoprotein hormone, leucine-rich repeat-containing G protein coupled-receptor 1, homodimer, heterodimer, thyrostimulin

INTRODUCTION

Members of the cystine knot growth factor (CKGF) superfamily, which are characterized with a CKGF domain as their primary structural feature, include (i) the glycoprotein hormones, (ii) the invertebrate bursicon hormone, (iii) the transforming growth factor beta (TGFβ) family, (iv) the bone morphogenetic protein (BMP) antagonist family, (v) the platelet-derived growth factor (PDGF) family and (vi) the nerve growth factor (NGF) family. Of the members of the CKGF

superfamily, the glycoprotein hormones are of fundamental importance in the regulation of both vertebrate and invertebrate physiology.

In vertebrates, members of the glycoprotein hormone family include follicle-stimulating hormone (FSH), luteinizing hormone (LH), thyroid-stimulating hormone (TSH) as well as chorionic gonadotropin (CG), which are implicated in governing several aspects of physiology including reproduction, energy metabolism along with growth and development. Structurally, these hormones are formed by the heterodimerization of two cystine-knot glycoprotein subunits, an α subunit that is structurally identical for each hormone (GPA1), and a hormone-specific β subunit (GPB1-4) (1, 2).

Two additional glycoprotein hormone subunits were more recently identified in the human genome, glycoprotein $\alpha 2$ (GPA2) and glycoprotein $\beta 5$ (GPB5), and were found to heterodimerize (GPA2/GPB5) and act on the same receptor as TSH. As a result, GPA2/GPB5 was coined the name thyrostimulin to differentiate it from TSH in vertebrates (3, 4). Unlike other glycoprotein hormones which are restricted to the vertebrate lineage, homologous genes encoding GPA2/GPB5 subunits exist in all bilaterian organisms, where its function appears to be pleiotropic (3–5). In vertebrates, the function of GPA2/GPB5 has been implicated, or at least suggested, to be involved in reproduction (6), thyroxine production (4, 7), skeletal development (8), immunoregulation (9, 10) and the proliferation of ovarian cancer cell lines (11). For invertebrate species, GPA2/GPB5 has been implicated or demonstrated to function in development (12–15), hydromineral balance (13–18) as well as in reproduction (15, 17, 19).

Heterodimerization by non-covalent interactions between the subunits forming FSH, LH, TSH, and CG is required for their respective biological functions (2, 20). However, whether heterodimerization is required for GPA2/GPB5 to activate its receptor and exert its physiological role in vertebrates and invertebrates is debated. For FSH, LH, TSH, and CG, the subunits are co-expressed in the same cells and each hormone is released into circulation as heterodimers (21, 22). On the other hand, GPA2 and GPB5 subunit expression profiles, both for vertebrate and invertebrate organisms, do not always occur in the same cells, and GPA2 is often expressed more widely and abundantly than GPB5 in some tissues (7, 12, 23–25). Additionally, unlike the beta subunits of the classic glycoprotein hormones, the structure of the GPB5 subunit lacks an extra pair of cysteine residues that form an additional disulfide linkage, referred to as the “seatbelt,” which strengthens and stabilizes its heterodimeric association with GPA2 (26).

Relative to other G protein-coupled receptors (GPCRs), members of the leucine-rich repeat-containing G protein-coupled receptor (LGR) family are often characterized with a large extracellular amino terminal domain responsible for the selective binding of their large hormone ligands (27). Following its initial genomic and molecular characterization (28), an invertebrate receptor for GPA2/GPB5, called LGR1, was functionally deorphanized in the fruit fly *Drosophila melanogaster*, where GPA2/GPB5 heterodimers were found to activate LGR1 increasing cyclic AMP (cAMP) levels (29).

Interestingly, stimulatory G protein (Gs) coupling and signaling to elevate cAMP was also shown with GPA2/GPB5-TSH receptor activation in humans (4, 6).

In the mosquito *Aedes aegypti*, genes encoding GPA2, GPB5 and LGR1 were identified and shown to be expressed in all developmental life stages, with expression enriched in adults compared to juvenile stages (13). In adults, LGR1 was found localized to epithelia throughout the gut where GPA2/GPB5 could regulate feeding-related processes and hydromineral balance (17). Notably, LGR1 transcript expression was also observed in the reproductive organs of males and females (19). In adult male mosquitoes, knockdown of LGR1 expression led to abnormal spermatogenesis with spermatozoa displaying malformations such as shortened flagella and consequently, LGR1-knockdown males had 60% less spermatozoa as well as significantly reduced fecundity relative to control mosquitoes (19).

With an interest in better understanding GPA2/GPB5 signaling in *A. aegypti* mosquitoes, our work herein set out to characterize the tissue-specific and cellular distribution expression profile of GPA2/GPB5 in mosquitoes. As well, we sought to determine if GPA2/GPB5 could heterodimerize, whether heterodimers were required to activate LGR1, and determine downstream signaling events upon receptor activation using a heterologous system. Combining various molecular techniques, we demonstrate GPA2/GPB5 cellular co-expression in the central nervous system of adult mosquitoes, and that both subunits are indeed required to activate LGR1, which exhibits ligand-dependent G protein-coupling activity. Moreover, our results also provided evidence for GPA2 and GPB5 homodimers, with individually-expressed subunits being incapable of activating LGR1. Overall, these findings appreciably advance our understanding of GPA2/GPB5 signaling in mosquitoes and provide novel directions to uncover the functions of homologous systems in other organisms.

MATERIALS AND METHODS

Animals

Adult *A. aegypti* (Liverpool) were derived from an established laboratory-reared colony raised under conditions described previously (17).

GPA2/GPB5 Transcript Analysis by RT-qPCR

Total RNA was isolated and purified from select *A. aegypti* tissues and organs, reverse transcribed into cDNA. GPA2 and GPB5 transcript abundance was quantified using a StepOnePlus™ Real-Time PCR System (Applied Biosystems) (see **SI Materials and Methods** for details) and transcript abundance was normalized to transcript levels of two reference genes; ribosomal protein 49 (GenBank accession: AY539746) and 60S ribosomal protein S18 (GenBank accession: XM_001660270), following the $\Delta\Delta C_t$ method as previously described (13). Experiments were repeated using a total of three technical replicates per sample and three biological replicates for each tissue/organ.

Fluorescence *in situ* Hybridization

Using gene-specific primers (Table S1), *A. aegypti* GPA2 and GPB5 sequences were amplified from previously prepared constructs (13) that contained the GPA2 and GPB5 complete open reading frames (ORFs). Sense and antisense probes were then generated following a similar protocol as recently reported (19). Briefly, cDNA fragments were ligated to pGEM T Easy vector (Promega, Madison, WI, USA) and used to transform NEB 5- α competent *Escherichia coli* cells (New England Biolabs, Whitby, ON, Canada). After screening plasmid constructs for directionality using T7 promoter oligonucleotide and gene-specific primers (Table S1), template sense or anti-sense cDNA strands for GPA2 and GPB5 probe synthesis were created by PCR amplification and verified by Sanger sequencing for base accuracy (The Centre for Applied Genomics, Sick Kids Hospital, Toronto, ON, Canada). Digoxigenin (DIG)-labeled anti-sense and sense RNA probes corresponding to GPA2 and GPB5 subunits were synthesized using the HiScribe T7 High Yield RNA Synthesis kit (New England Biolabs, Whitby, ON, Canada). Fluorescence *in situ* hybridization (FISH) was then used to detect GPA2 and/or GPB5 transcript in the mosquito central nervous system using 4 ng μl^{-1} (GPB5) and/or 6 ng μl^{-1} (GPA2) RNA sense/antisense probes, following a previously established protocol (30). Preparations were analyzed with a Lumen Dynamics X-CiteTM 120Q Nikon fluorescence microscope (Nikon, Mississauga, ON, Canada), or a Yokogawa CSU-XI Zeiss Cell Observer Spinning Disk confocal microscope, and images were processed using Zeiss Zen and ImageJ software. All microscope settings were kept identical when acquiring images of control and experimental preparations.

Wholemount Immunohistochemistry

GPB5 immunoreactivity in the abdominal ganglia of adult mosquitoes was examined in newly emerged and four-day old *A. aegypti* that were lightly CO₂ anesthetized, and dissected in PBS at RT. Tissues were fixed, permeabilized and incubated in a custom affinity-purified rabbit polyclonal GPB5 antibody (1 $\mu\text{g ml}^{-1}$; Genscript, Piscataway, NJ) designed against an antigen sequence (CDSNEISDWRFP) positioned at residues 69–80 of the deduced *A. aegypti* GPB5 protein sequence (13) for 48 h at 4°C rocking. Control treatments involved pre-incubated *A. aegypti* GPB5 primary antibody solution containing 100:1 peptide antigen:antibody (mol:mol). After several washes, tissues were incubated overnight at 4°C in Alexa Fluor 488-conjugated goat anti-rabbit Ab (1:200) secondary antibody (Life Technologies, Carlsbad, CA) in PBS containing 10% normal sheep serum. The next day, samples were washed and then mounted onto coverslips using mounting media containing 5 $\mu\text{g ml}^{-1}$ Diamidino-2-phenylindole dihydrochloride (DAPI), and analyzed using a Lumen Dynamics X-CiteTM 120Q Nikon fluorescence microscope (Nikon, Mississauga, ON, Canada), or optically sectioned using a Yokogawa CSU-XI Zeiss Cell Observer Spinning Disk confocal microscope. All images were processed using Zeiss Zen and ImageJ software. Further details concerning the wholemount immunohistochemical protocol were reported in an earlier study (17).

Plasmid Expression Constructs

Plasmid expression constructs were designed to study *A. aegypti* and *H. sapiens* GPA2/GPB5 subunit dimerization patterns and receptor signaling. Using previously available hexa-histidine-tagged *A. aegypti* GPA2-His (13), as well as FLAG (DYKDDDDK)-tagged (-FLAG) *H. sapiens* GPB5-FLAG (Genscript, Clone OHu31847D) plasmid vectors as template, the full ORF of each (*A. aegypti* and *H. sapiens*) GPA2 and GPB5 subunit coding sequence, including a consensus Kozak translation initiation sequence, was amplified and a hexa-histidine or FLAG tag sequence was incorporated on the carboxyl-terminus of subunits to produce the following fusion proteins; *A. aegypti* GPA2-FLAG, *H. sapiens* GPB5-His (Table S2). A pcDNA3.1⁺ mammalian expression construct containing mCherry, which was a gift from Scott Gradia (Addgene plasmid # 30125), was utilized to verify cell transfection efficiency. Experiments also utilized previously prepared pcDNA3.1⁺ constructs with *A. aegypti* GPB5-His and *A. aegypti* LGR1 coding sequences and dual promoter vector pBudCE4.1 containing both *A. aegypti* GPA2-His and GPB5-His (13). Additionally, pcDNA 3.1⁺ mammalian expression vector construct containing FLAG tagged *H. sapiens* thyrotropin receptor (TSHR-FLAG) (Genscript USA Inc., Clone OHu18318D), *H. sapiens* GPA2-FLAG (Genscript USA Inc., Clone OHu31847D), *H. sapiens* GPB5-FLAG (Genscript USA Inc., Clone OHu55827D) and pGlosensorTM-22F cyclic adenosine monophosphate (cAMP) biosensor plasmid (Promega Corp., Madison, WI), which were used for receptor activation and intracellular signaling assays.

Generation of Tethered *A. aegypti* GPA2/GPB5 Fusion Construct

The ORFs of *A. aegypti* GPA2 and GPB5 sequences were tethered together in order to promote heterodimer interactions for testing in receptor activity assays with mammalian cell lines. A hexa-histidine tagged artificial linker sequence involving three glycine-serine repeats was used to fuse the amino-terminus of *A. aegypti* GPA2 propeptide sequence to the carboxyl-terminus of *A. aegypti* GPB5 prepropeptide sequence, using multiple PCR amplifications with several primer sets (Table S2) as performed previously using lamprey and *Amphioxus* GPA2 and GPB5 sequences (31, 32) (see SI Materials and Methods for details).

Transient Transfection of HEK 293T Cells

Human embryonic kidney (HEK 293T) cells were grown in complete growth media (Dulbecco's modified eagles medium: nutrient F12 (DMEM) media, 10% heat inactivated fetal bovine serum (Wisent, St. Bruno, QC) and 1X antimycotic-antibiotic) and maintained in a water-jacketed incubator at 37°C, 5% CO₂. When cells reached ~80–90% confluency, they were transfected with mammalian expression plasmid constructs in 6-well tissue culture plates (Thermo Fisher Scientific, Burlington, ON) using Lipofectamine 3000 transfection reagent (Life Technologies, Carlsbad, CA) with 3:1 ($\mu\text{L}:\mu\text{g}$) transfection reagent to DNA ratio. Before transfection, culture media was replaced with either serum-free medium (DMEM and 1X antimycotic-antibiotic) for experiments that collected secreted proteins, or fresh complete

growth medium for experiments that dually-transfected cells with pGlosensorTM-22F cAMP biosensor plasmid and either *H. sapiens* TSHR, *A. aegypti* LGR1, or mCherry.

Preparation of Protein Samples

At 48 h post-transfection, serum-free culture media containing secreted proteins were collected and concentrated in 0.5 mL 3-kDa molecular weight cut-off centrifugal filters (VWR North America). In some experiments, cells were dislodged with PBS containing 5 mM ethylenediaminetetraacetic acid (EDTA; Life Technologies) (PBS-EDTA) pH 8.0, pelleted at $400 \times g$ for 5 min, resuspended in PBS and transferred to 1.5 mL centrifuge tubes for a subsequent centrifugation. Cell lysates were then prepared by resuspending and sonicating cells for 5 s in cell lysis buffer containing 37.5 mM Tris, pH = 7.5, 1.5 mM EDTA, pH 8.0, 3% sodium dodecyl sulfate, 1.5% protease inhibitor cocktail (v/v), and 1.5 mM dithiothreitol (DTT). For receptor activity assays, in order to prevent carry-over of lysis buffer, cell lysates were concentrated in 3-kDa molecular weight cut-off centrifugal filters and re-constituted back to initial volumes with serum-free media for a total of three repetitions. Proteins were then used for cross-linking analysis, deglycosylation, and immunoblotting or used as ligands for functional receptor activation using the cAMP signaling biosensor assays.

Dissuccinimidyl suberate (DSS), a chemical cross-linker that is primarily reactive toward amino groups providing stabilization of weak or transient protein intermolecular interactions, was employed to study GPA2 and GPB5 protein-protein interactions. *A. aegypti* and *H. sapiens* GPA2-FLAG and GPB5-His proteins were kept separate or combined together (GPA2-FLAG/GPB5-His) and then treated with DSS (Sigma Aldrich, Oakville, ON). In some experiments, *A. aegypti* GPA2-His/GPB5-His were co-expressed using a dual promoter plasmid and as such, media containing both His-tagged subunits were directly treated with DSS. To treat secreted concentrates of culture media containing *H. sapiens* GPA2 and GPB5 proteins, 0.68 mM DSS was used whereas both 0.68 mM (data not shown) and 2.04 mM DSS was used to test the dimerization characteristics of *A. aegypti* GPA2-His/GPB5-FLAG proteins. Cross-linking was performed for 30 min at RT and reactions were quenched with 50 mM Tris, pH 7.4 for 10 min under constant mixing. To remove N-linked sugars, protein samples were treated with peptide-N-Glycosidase F (PNGase) (New England Biolabs, Whitby, ON) following manufacturers guidelines, with the only modification being that protein samples were not heated to 100°C before enzymatic deglycosylation. Experiments aimed to determine the effects of protein glycosylation on cross-linking ability first treated *A. aegypti* GPA2-His/GPB5-His subunits with PNGase and subsequently cross-linked samples with DSS after deglycosylation.

Western Blot Analysis

Samples were prepared in 2x Laemmli buffer (Sigma Aldrich, Oakville, ON) containing 4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125 M Tris HCl, pH ~6.8, and resolved on 10 or 15% SDS-polyacrylamide gels under reducing conditions at 120 V for 90–110 min.

Using a wet transfer system, proteins were then transferred to polyvinylidene difluoride (PVDF) membranes at 100 V for 75 min. For GPA2-FLAG/GPB5-His heterodimerization experiments, samples were run in duplicate on the same gel and following transfer, membranes were cut in half for separate primary antibody incubations. Membrane blots containing protein samples were blocked for 1 h in PBS containing 0.1% Tween-20 (Bioshop, Burlington, ON, Canada) and 5% skim milk powder (PBSTB) rocking at RT. After blocking, membranes were incubated overnight at 4°C on a rocking platform in PBSTB containing mouse monoclonal anti-His (1:500 dilution) or mouse monoclonal anti-FLAG (1:500) primary antibody solutions. The next day, membranes were washed three times with PBS containing 0.1% Tween-20 (PBST) for 15 min each wash and then were incubated in PBSTB containing anti-mouse HRP conjugated secondary antibody (1:2000–1:3000 dilution) for 1 h rocking at RT before washing again with PBST (3 × 15 min washes). Finally, blots were incubated with Clarity Western ECL substrate and images were developed using a ChemiDoc MP Imaging System (Bio-Rad Laboratories, Mississauga, ON) and molecular weight measurements and analysis were performed using Image Lab 5.0 software (Bio-Rad Laboratories, Mississauga, ON). Western blots used to study recombinant GPA2 and GPB5 dimerization were repeated at least three times.

Receptor Functional Activation Bioluminescence Assays

HEK 293T cells were co-transfected to express (i) *H. sapiens* TSHR, *A. aegypti* LGR1, or mCherry along with (ii) pGlosensorTM-22F cAMP biosensor plasmid (Promega Corp., Madison, WI), which encodes a modified form of firefly luciferase with a fused cAMP binding moiety providing a biosensor for the direct detection of cAMP signaling in live cells. At 48 h post transfection, recombinant cells were dislodged with PBS-EDTA, pelleted at $400 \times g$ for 5 min, resuspended in assay media [DMEM:F12 media with 10% fetal bovine serum (v/v)] containing cAMP GloSensor reagent (2% v/v), and incubated for 3 h rocking at RT shielded from light. White 96-well luminescence plates (Greiner Bio-One, Germany) were loaded under low light with previously prepared secreted or cell lysate protein concentrates (described above), forskolin or assay media alone, and incubated at 37°C for 30 min prior to performing receptor activity assays. For stimulatory G-protein (Gs) pathway detection, recombinant cells expressing either *H. sapiens* TSHR, *A. aegypti* LGR1, or mCherry along with the cAMP biosensor were pre-treated with 0.25 mM 3-isobutyl-1-methylxanthine (IBMX) for 30 min at RT with rocking and shielded from light. Using an automatic injector unit (BioTek Instruments Inc., Winooski VT), co-transfected HEK 293T cells were then loaded into wells containing various treatments, including 250 nM forskolin as a positive control or concentrates of protein fractions from mCherry-transfected cells as a negative control. For inhibitory G-protein (Gi/o) pathway testing, *A. aegypti* GPA2 and/or GPB5 was tested for the ability to inhibit a forskolin-induced cAMP-mediated bioluminescent response. Using an automatic injector unit (BioTek Instruments Inc., Winooski

VT), cells expressing LGR1 or mCherry (LGR1 activation and signaling negative control) were loaded into wells that contained various ligand treatments. Subsequently, 250 nM or 1 μ M forskolin was added to wells immediately after the addition of cells, using a second automatic injector unit. For bioluminescent assays with tethered GPA2/GPB5 fusion proteins, recombinant cells expressing or not expressing LGR1 were equally divided to test Gs and Gi signaling pathways simultaneously using the same batch of cells per biological replicate. In all assays, luminescence was measured every 2 min for 20 min at 37°C using a Synergy 2 Multimode Microplate Reader (BioTek, Winooski, VT) and averaged over 4–8 technical replicates for each treatment. To calculate the relative luminescent response, data were normalized to luminescent values recorded from treatments with 250 nM or 1 μ M forskolin alone. Assays were performed repeatedly and involved 3–6 independent biological replicates. Relative luminescent values were compiled on Excel and transferred into GraphPad Prism 8.0 for figure preparation and statistical analysis. Data were analyzed using one-way or two-way ANOVA with Tukey's multiple comparison post-test ($p < 0.05$).

RESULTS

A. *aegypti* GPA2 and GPB5 Subunit Expression Localization

To determine the distribution of GPA2 and GPB5 subunit expression in the central nervous system and peripheral tissues, adult mosquito organs were analyzed using RT-qPCR. GPA2 and GPB5 subunit transcript was detected in the central nervous system of adult male and female mosquitoes, with significantly enriched expression in the abdominal ganglia relative to other regions of the nervous system and peripheral tissues (Figures 1A,B). Fluorescence *in situ* hybridization was employed to localize cell-specific expression of the GPA2 and GPB5 transcripts in the abdominal ganglia. GPA2 and GPB5 anti-sense RNA probes identified two bilateral pairs of neuroendocrine cells (Figures 1C,D) within each of the first five abdominal ganglia in male and female mosquitoes whereas the control sense probes did not detect cells in the nervous system (Figures 1E,F). Within each of these first five abdominal ganglia, GPA2 transcript localized to similar laterally-positioned cells as the GPB5 transcript (Figures 1G,H). To determine if cells expressing GPA2 transcript were the same cells expressing GPB5 transcript, abdominal ganglia were simultaneously treated with both GPA2 and GPB5 anti-sense RNA probes. Using this dual probe approach, results confirmed the detection of only two bilateral pairs of cells (Figures 1I,J) that displayed a greater staining intensity compared to the intensity of cells detected when using either the GPA2 or GPB5 anti-sense probe alone (Figures 1C,D).

Using a custom antibody targeting *A. aegypti* GPB5, we next sought to immunolocalize GPB5 protein in the abdominal ganglia. GPB5 immunoreactivity localized to two bilateral pairs of cells (Figure 2A) within the first five ganglia, which were in similar positions to cells expressing GPA2 and GPB5 transcript (Figures 1C,D,G–J). Regardless of sex, in 30% of

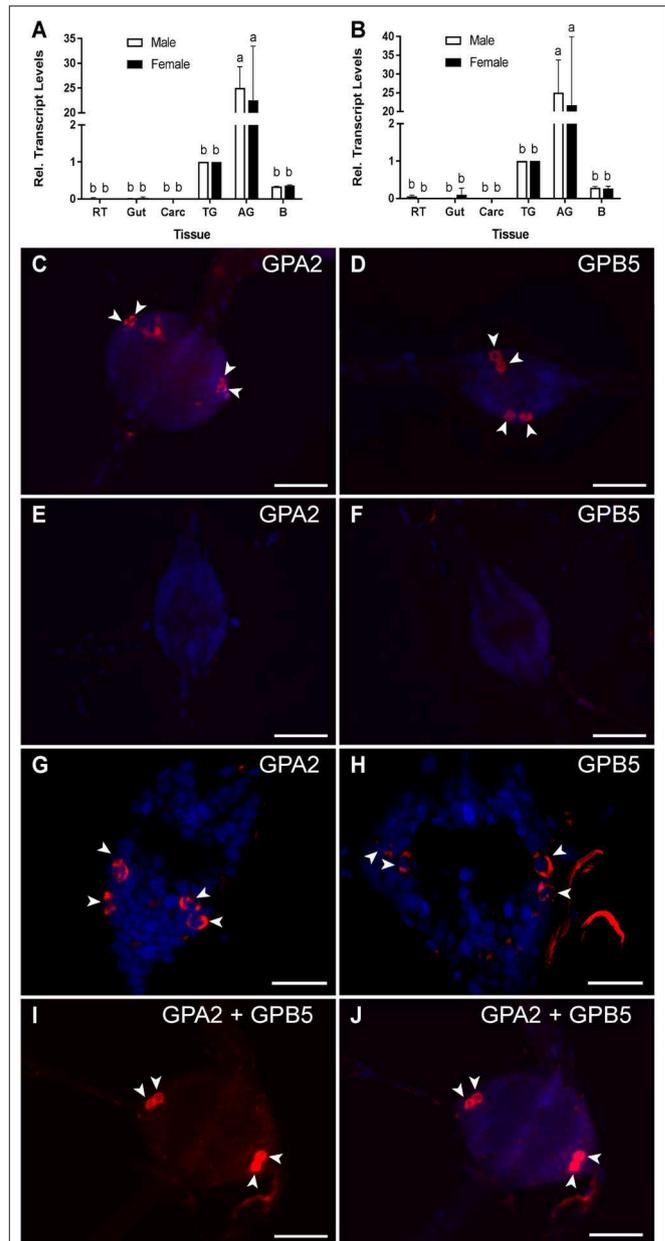


FIGURE 1 | GPA2/GPB5 subunit transcript expression and localization in adult *A. aegypti*. (A,B) RT-qPCR examining GPA2 (A) and GPB5 (B) transcript expression in the central nervous system of adult mosquitoes, with significant enrichment in the abdominal ganglia (AG). Subunit transcript abundance is shown relative to their expression in the thoracic ganglia (TG). Mean \pm SEM of three biological replicates. Columns denoted with different letters are significantly different from one another. Multiple comparisons two-way ANOVA test with Tukey's multiple comparisons ($P < 0.05$) to determine sex- and tissue-specific differences. Reproductive tissues (RT), alimentary canal (Gut), carcass (Carc), brain (B). Fluorescence *in situ* hybridization anti-sense (C,D,G–J) and sense (E,F) probes to determine GPA2 and GPB5 transcript localization (GPA2 and/or GPB5 transcript, red; nuclei, blue) in the abdominal ganglia of adult mosquitoes. Unlike sense probe controls (E,F), two bilateral pairs of cells (arrowheads) were detected with GPA2 (C,G) and GPB5 (D,H) anti-sense probes in the first five abdominal ganglia of adult male and female *A. aegypti*. Co-localization of the GPA2 (G) and GPB5 (H) transcript was

(Continued)

FIGURE 1 | verified by treating abdominal ganglia dually with a combination of GPA2 and GPB5 anti-sense probes (**I,J**) that revealed two, intensely-stained bilateral pairs of cells. In (**C–F,I,J**), microscope settings were kept identical when acquiring images of control and experimental ganglia. The second abdominal ganglion is depicted in each image as a representative, given that no differences in the number nor staining intensity of cells were observed between ganglia of a given mosquito. Scale bars are 50 μm in (**C–F,I,J**) and 40 μm in (**G,H**). Experimental procedures were repeated three-four times with five mosquitoes (of each sex) per trial.

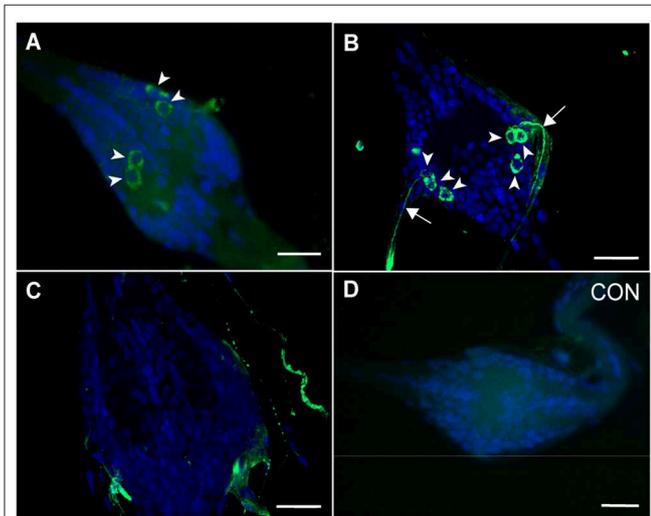


FIGURE 2 | Immunolocalization of GPB5 subunit expression in the abdominal ganglia of adult *A. aegypti*. Experimental treatments demonstrated GPB5 immunoreactivity (green) in two (**A**), and in some mosquitoes, three (**B**) bilateral pairs of neuroendocrine cells (arrowheads) within the first five abdominal ganglia of the ventral nerve cord (DAPI, blue). Optical sections of ganglia revealed axonal projections emanate from GPB5-immunoreactive cells through the lateral nerves (arrows). No cells were detected in the sixth, terminal ganglion (**C**) and in control treatments (CON) where GPB5 antibody was preabsorbed with GPB5 synthetic antigen (**D**). In (**A,D**) or (**B,C**), microscope settings were kept identical when acquiring images of control and experimental ganglia. The second abdominal ganglion is depicted in (**A,B,D**) as a representative, given that no differences in the number nor staining intensity of cells were observed between ganglia of a given mosquito. Scale bars are 25 μm in (**A–D**) and 20 μm in (**B,C**). Experimental procedures were repeated three-four times with five mosquitoes (of each sex) per trial.

mosquitoes examined, three instead of two bilateral pairs of cells immunolocalized to each of the first five abdominal ganglia of the ventral nerve cord (**Figure 2B**), whereas no cells were ever detected in the sixth terminal ganglion (**Figure 2C**). For a given mosquito, there were no differences in the number of GPB5 immunoreactive cells detected between different ganglia. Along the lateral sides of each of the first five abdominal ganglia, GPB5 immunoreactive processes were observed to closely associate into a tract of axons that emanated through the lateral nerve (**Figure 2B**). Control treatments with GPB5 antibody preabsorbed with the GPB5 immunogenic antigen did not detect any cells in ganglia (**Figure 2D**).

Cross-Linking Analyses to Determine *A. aegypti* GPA2 and GPB5 Subunit Interactions

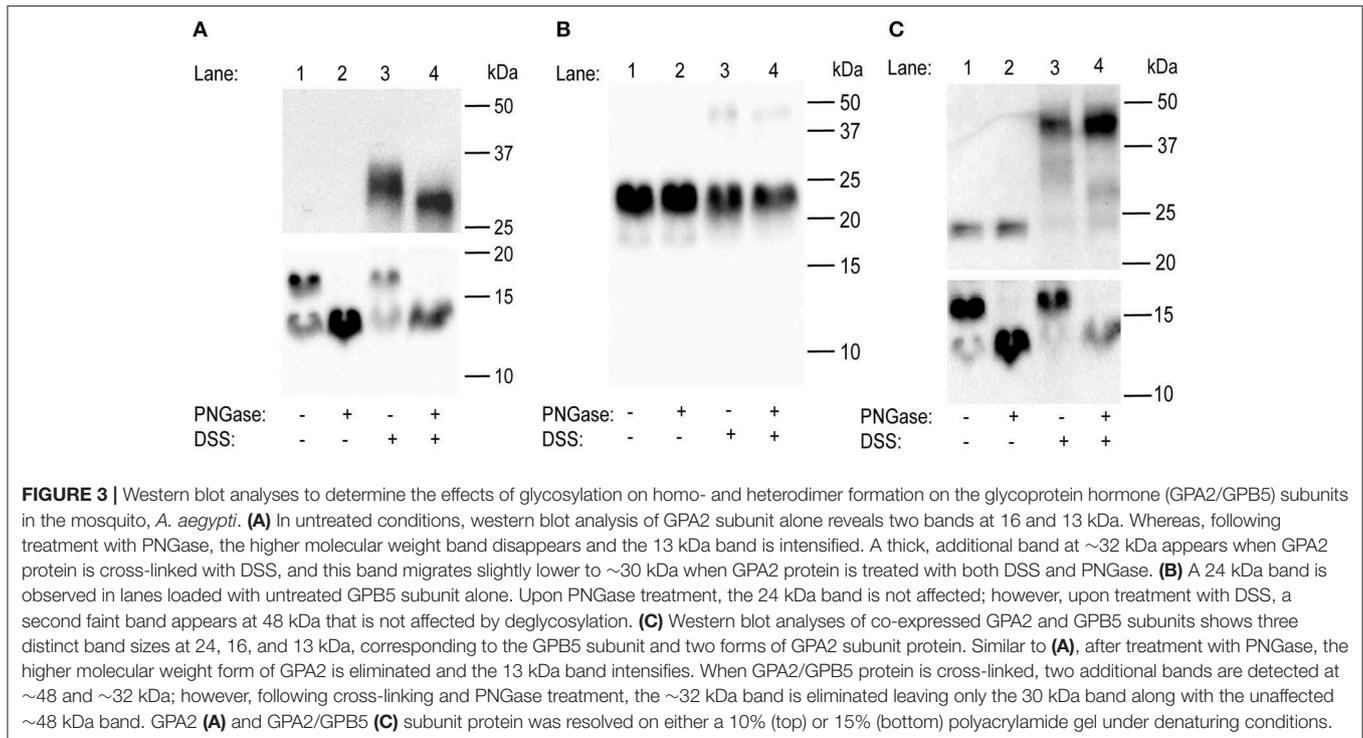
A. aegypti GPA2 and GPB5 subunit protein interactions were studied using western blot analysis of recombinant proteins from HEK 293T cells expressing each subunit independently

or co-expressing both subunits in the same cells using a dual promoter plasmid. Under control conditions, GPA2 protein is represented as two bands at 16 and 13 kDa, which correspond to the glycosylated and non-glycosylated forms of *A. aegypti* GPA2, respectively (**Figure 3A**). Following deglycosylation with PNGase, the higher molecular weight band of GPA2 is eliminated, and the non-glycosylated lower molecular weight band intensifies (**Figure 3A**). Interestingly, when the GPA2 subunit was tested to examine potential homodimerization, an additional strong band at ~ 32 kDa was detected, which migrates to ~ 30 kDa when cross-linked protein samples were deglycosylated using PNGase (**Figure 3A**). Under control conditions, GPB5 protein is represented as a band size at 24 kDa and the migration pattern is not affected by PNGase treatment (**Figure 3B**). Following treatment with DSS, a faint second band appears at 48 kDa, which does not change in molecular weight after treatment with PNGase (**Figure 3B**). Three independent band sizes at 24 kDa (GPB5), 16 kDa (glycosylated GPA2) and 13 kDa (non-glycosylated GPA2) were detected in lanes loaded with protein isolated from HEK 293T cells co-expressing GPA2 and GPB5 subunits (**Figure 3C**). After removal of N-linked sugars, the 24 kDa band is not affected but the 16 kDa band disappears and 13 kDa band intensifies (**Figure 3C**), as observed when assessing the GPA2 subunit independently (**Figure 3A**). Cross-linked samples show the addition of two higher molecular weight bands at ~ 48 and ~ 32 kDa, the latter of which migrates lower to ~ 30 kDa when subjected to PNGase treatment (**Figure 3C**). Given the resolution and intensity of the detected bands, the molecular weight band at ~ 45 kDa could indicate GPA2/ GPB5 heterodimeric interactions (13 or 16 kDa + 24 kDa = 37–40 kDa). Alternatively, these bands could reflect GPB5 (24 + 24 kDa = 48 kDa) homodimers. As a result of this uncertainty, additional experiments were performed to clarify whether *A. aegypti* GPA2 and GPB5 subunits are heterodimeric partners.

Heterodimerization of Mosquito and Human GPA2/GPB5

Using yeast two-hybrid analyses and cross-linking experiments, it was previously shown that human GPA2 (hGPA2) and GPB5 (hGPB5) subunits are capable of heterodimerization (4). As a result, to verify whether *A. aegypti* GPA2 and GPB5 subunit proteins are heterodimeric candidates, experiments were performed alongside hGPA2/hGPB5 subunit proteins, using the latter as a positive experimental control for heterodimer detection.

Initially, single-promoter expression constructs were designed to incorporate a FLAG-tag and His-tag on the C-terminus of (human and mosquito) GPA2 and GPB5 subunits, respectively (i.e., GPA2-FLAG and GPB5-His). When probed with an anti-His antibody, no bands were detected in lanes containing only GPA2-FLAG (human



and mosquito) protein (**Figures 4A,B**). Similarly, when an anti-FLAG antibody, no bands were detected in lanes containing only GPB5-His protein (human and mosquito) (**Figures 4A,B**).

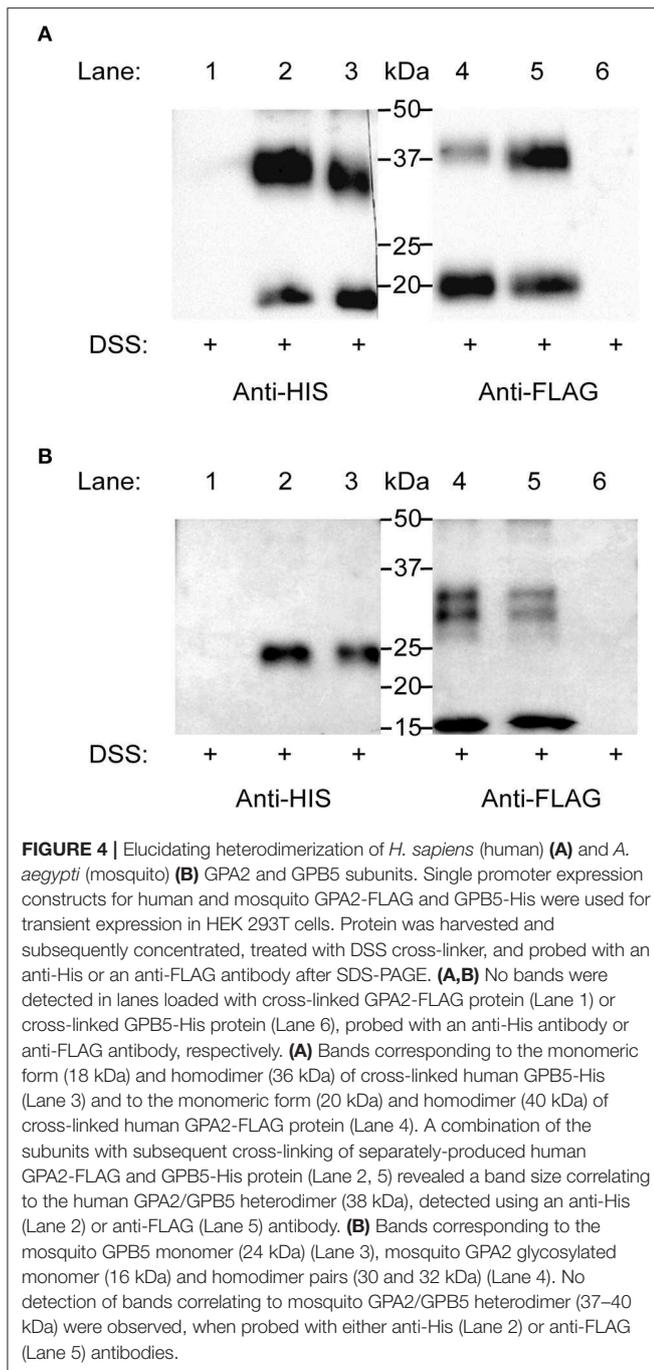
Lanes loaded with cross-linked hGPB5-His protein revealed two bands at 18 kDa (monomer) and 36 kDa (homodimer) (**Figure 4A**). Similarly, in lanes loaded with cross-linked hGPA2-FLAG protein, two bands at 20 kDa (monomer) and 40 kDa (homodimer) were detected (**Figure 4A**). To determine subunit heterodimerization, hGPA2-FLAG protein was combined with hGPB5-His protein (i.e., produced separately in different cell batches) and subsequently crosslinked using DSS. Results showed an intense 38 kDa band size that correlated to the molecular weight of the hGPA2-FLAG/hGPB5-His heterodimers; detected with both anti-His and anti-FLAG primary antibody solutions (**Figure 4A**).

Identical experiments were conducted using *A. aegypti* GPA2-FLAG/GPB5-His protein but using three-fold higher concentration of DSS cross-linker to help improve detection of inter-subunit interactions. Results demonstrated that lanes loaded with DSS-treated GPB5-His protein resulted in a 24 kDa monomer band (**Figure 4B**), whereas lanes containing cross-linked GPA2-FLAG detected a 16 kDa (glycosylated monomer), 30 and 32 kDa (homodimers) band size (**Figure 4B**). As a result, unlike immunoblots containing hGPA2-FLAG and hGPB5-His protein (**Figure 4A**), lanes loaded with cross-linked *A. aegypti* GPA2-FLAG and GPB5-His failed to provide evidence of bands correlating to the predicted molecular weight of an *A. aegypti* GPA2/GPB5

heterodimer, which would be expected at 37–40 kDa (**Figure 4B**).

A. aegypti GPA2/GPB5 Unable to Activate LGR1-Mediated Gs and Gi/o Signaling Pathways

Bioluminescent assays were employed to confirm LGR1 interaction with mosquito GPA2 and/or GPB5 subunits and elucidate downstream signaling pathways upon receptor activation. Since human GPA2/GPB5 (hGPA2/hGPB5) has previously been shown to bind and activate human thyrotropin receptor (hTSHR) mediating a stimulatory G protein (Gs) signaling pathway (4), we first validated our experimental design using hGPA2/hGPB5 and hTSHR as a positive control. Recombinant hGPA2 and hGPB5 subunit proteins were produced in HEK 293T cells with single promoter expression constructs containing the hGPA2 or hGPB5 sequences. Conditioned culture media containing secreted proteins were subsequently concentrated, and crude extracts containing hGPA2, hGPB5 or a combination of hGPA2 and hGPB5 subunits were tested as ligands on HEK 293T cells expressing the hTSHR and a cAMP-sensitive firefly luciferase biosensor, which produces bioluminescence upon interacting with cAMP. Control treatments involved the incubation of hTSHR/luciferase-expressing cells with concentrated media collected from mCherry-transfected cells (negative control), and data was normalized to treatments with 250 nM forskolin (positive control) (**Figures 5A,B**). As expected, the results indicate that incubation with extracts containing a combination of both



hGPA2 and hGPB5 were required to stimulate a cAMP-mediated luminescent response from hTSHR-expressing cells, but not when incubated with extracts containing individual subunits (Figure 5A). We next performed identical experiments using *A. aegypti* GPA2/GPB5 and LGR1. Given the availability of a dual promoter plasmid (pBudCE4.1), an additional treatment was performed whereby GPA2/GPB5 subunits were co-expressed within the same cells and conditioned media was concentrated as described above. Unlike the results using hGPA2/hGPB5 subunit homologs on HEK 293T cells expressing the hTSHR

(Figure 5A), no combination of mosquito GPA2 and/or GPB5 subunit proteins led to an increase in the cAMP-based luminescent response in HEK 293T cells expressing mosquito LGR1 (Figure 5B).

Using an *in silico* analysis to predict coupling specificity of *A. aegypti* LGR1 and human TSHR to different families of G-proteins (33), we determined that *A. aegypti* LGR1 is strongly predicted to couple to inhibitory (Gi/o) G proteins Table S3). As a result, to determine whether *A. aegypti* GPA2 and/or GPB5 activate a Gi/o signaling pathway, various combinations of GPA2 and GPB5 were tested for their ability to inhibit a 250 nM forskolin-induced rise in cAMP measured by changes in bioluminescence (Figures 5C,D). Results revealed that sole treatments of GPB5 proteins alone significantly inhibited a forskolin-induced luminescent response, relative to control treatments with mCherry proteins, when incubated with cells expressing LGR1 (Figure 5C). However, similar inhibitory effects of GPA2 and GPB5 proteins were also observed with cells lacking LGR1 expression (Figure 5D).

Characterization of Tethered *A. aegypti* GPA2/GPB5

Activity on the hTSHR was only observed when both hGPA2 and hGPB5 subunits were coapplied for receptor activation (Figure 5A) and, unlike the heterodimerization of hGPA2/hGPB5 observed in our experiments, mosquito GPA2/GPB5 lacked evidence of heterodimerization (Figure 4). In light of these observations, we hypothesized that the activation of *A. aegypti* LGR1 also required subunit heterodimerization. To mimic GPA2/GPB5 heterodimers using the heterologous expression system, both GPA2 and GPB5 mosquito subunits were expressed as a tethered, single-chain polypeptide by fusing the C-terminus of the GPB5 prepropeptide sequence with the N-terminus of the GPA2 propeptide sequence, using a tagged linker sequence composed of twelve amino acids, involving three glycine-serine repeats and six histidine residues.

HEK 293T cells were transfected to transiently express a single promoter plasmid construct containing the tethered GPA2/GPB5 sequence, or the red fluorescent protein (mCherry) as a negative control. To verify expression of the construct, at 48 h post-transfection, cell lysates along with the conditioned culture media, the latter of which contains secreted proteins, were collected for immunoblot analysis. No bands were detected in lanes containing cell lysate or secreted protein fractions of mCherry transfected cells (Figure 6A). However, in the lysates of cells transfected to express tethered GPA2/GPB5, an intense band at 32 kDa, and less intense band at 37 kDa were detected; the latter of which corresponds to the predicted molecular weight of non-glycosylated GPA2 (13 kDa) plus GPB5 (24 kDa) (Figure 6A). In lanes containing secreted fractions of tethered GPA2/GPB5-transfected cells, a band at 37 kDa was again detected, as well as a stronger 40 kDa band size that correlates to the predicted molecular weight of glycosylated GPA2 (16 kDa) plus GPB5 (24 kDa) (Figures 6A,B). After secreted protein extracts containing tethered GPA2/GPB5 proteins were treated with PNGase, the higher 40 kDa molecular

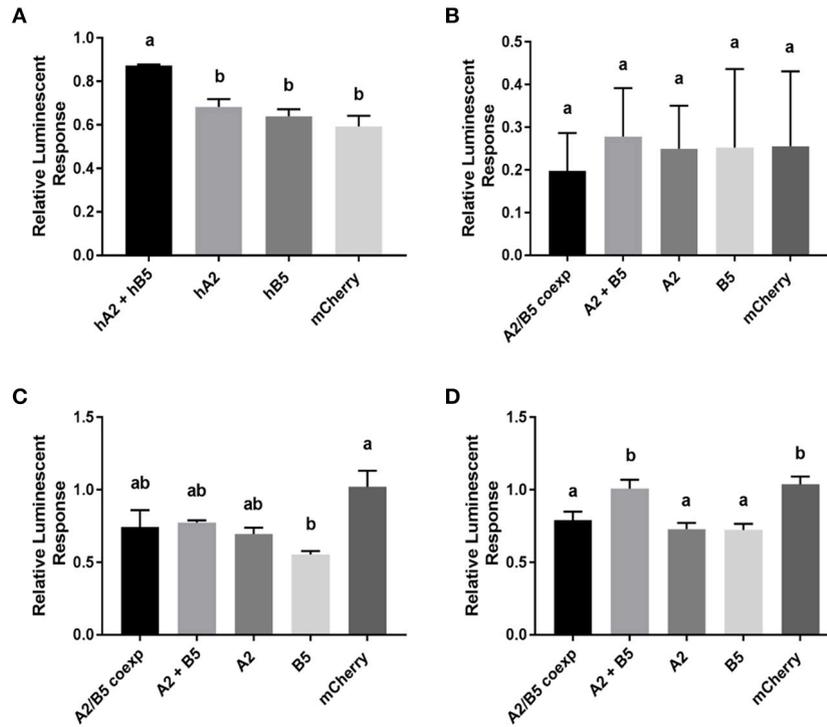


FIGURE 5 | cAMP-mediated bioluminescence assays to determine the effect of GPA2, GPB5, and GPA2/GPB5 on G-protein signaling of *H. sapiens* (human) TSHR (A), *A. aegypti* (mosquito) LGR1 (B,C), or cells expressing a red fluorescent protein, mCherry (D). Secreted protein fractions for each subunit were prepared separately from HEK 293T cells expressing human GPA2 (hA2), human GPB5 (hB5), mosquito GPA2 (A2), mosquito GPB5 (B5), mCherry, or co-expressing mosquito GPA2 and GPB5 in a dual promoter plasmid (A2/B5 coexp). Secreted protein fractions were then tested separately or combined (A2 + B5) and then incubated with cells co-expressing the cAMP biosensor along with either (A) human TSHR, (B,C) mosquito LGR1 or (D) mCherry, the latter of which was used as a negative control in the functional assay and also served to verify transfection efficiency of HEK 293T cells. (A–D) Luminescent values were recorded and normalized to luminescence response from treatments with 250 nM forskolin. (A) Unlike treatments with human GPA2 (hA2) or human GPB5 (hB5) applied singly, a significant increase in cAMP-mediated luminescence was observed when TSHR-expressing cells were incubated with culture media containing both human GPA2 and human GPB5 (hA2 + hB5), relative to incubations with mCherry controls. (B) No differences in luminescence were observed when LGR1-expressing cells were incubated with media containing mosquito GPA2/GPB5 subunits, compared to mCherry controls. (C,D) To test Gi/o signaling pathways, experiments were performed in the presence of 250 nM forskolin to determine the ability for each ligand to inhibit a forskolin-induced cAMP response. (C) The addition of GPB5 on LGR1-expressing cells significantly inhibited forskolin-induced luminescent response, compared to treatments with mCherry controls; (D) however, this inhibition was also observed when GPA2 and GPB5 proteins were incubated with HEK 293T cells in the absence of LGR1. Mean \pm SEM of three (A,B,D) or six (C) biological replicates. Columns denoted with different letters are significantly different from one another. Multiple comparisons one-way ANOVA test with Tukey's multiple comparisons ($P < 0.05$).

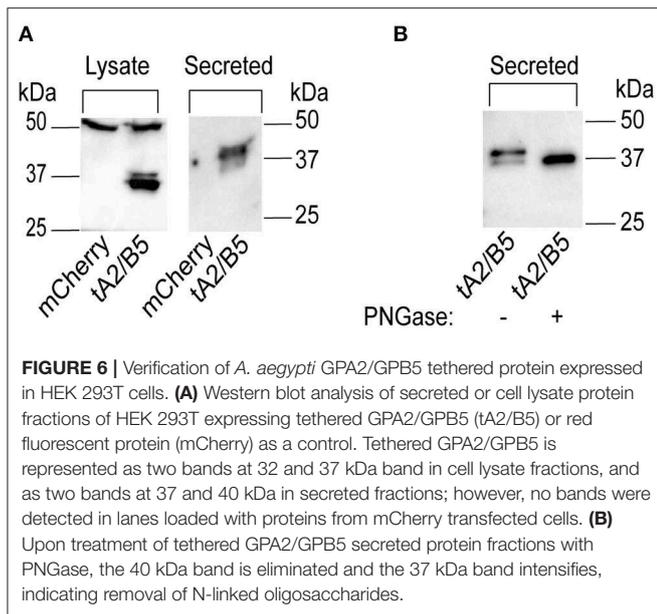
weight band disappears and the 37 kDa band size intensifies, which confirms the observed molecular weight shift results from removal of N-linked oligosaccharides (Figure 6B). Thus, the tethered GPA2/GPB5 undergoes similar post-translational processing as observed for the subunits expressed individually.

Tethered *A. aegypti* GPA2/GPB5 Activates LGR1

The activity of tethered GPA2/GPB5 proteins on LGR1 activation was examined. Cell lysate or secreted protein fractions collected from mCherry- or tethered GPA2/GPB5-transfected cells were incubated with HEK 293T cells co-expressing the cAMP luciferase biosensor and either *A. aegypti* LGR1 or mCherry (i.e., not expressing LGR1). Whether tethered GPA2/GPB5 proteins (cell lysate or secreted fractions) could elevate cAMP or inhibit a forskolin-induced rise in cAMP was assessed and compared

to negative control treatments with proteins harvested from mCherry-transfected cells.

Without the addition of tethered GPA2/GPB5 proteins, the basal luminescent levels were higher in LGR1-transfected cells compared to mCherry-transfected cells (Figure 7), suggesting constitutive activity of *A. aegypti* LGR1 elevating cAMP levels. However, the application of neither secreted protein fractions nor cell lysates of tethered GPA2/GPB5-transfected cells elicited an increase in the cAMP-mediated luminescent response relative to control treatments with mCherry proteins (Figures 7A,B). Secreted protein fractions containing tethered GPA2/GPB5 protein had no effect on the forskolin-induced cAMP-dependent luminescence, compared to control treatments with mCherry secreted proteins in LGR1-expressing cells (Figure 7C). Notably, however, treatments of LGR1-expressing cells with cell lysates containing tethered GPA2/GPB5 significantly inhibited the



luminescent response owing to the forskolin-induced rise in cAMP relative to treatments with mCherry-transfected cell lysates, which was not observed in cells lacking LGR1 expression (Figure 7D).

DISCUSSION

GPA2/GPB5: A Neuropeptide Produced in the Abdominal Ganglia of Mosquitoes

The central nervous system (CNS) of adult mosquitoes is comprised of a brain and a ventral nerve cord, consisting of the subesophageal ganglion, three fused thoracic ganglia and six abdominal ganglia. GPA2 and GPB5 transcripts are significantly enriched in the abdominal ganglia of adult mosquitoes relative to peripheral tissues and other regions of the CNS. Although low levels of GPA2 and GPB5 transcripts were detected in the thoracic ganglia and brain using RT-qPCR, fluorescence *in situ* hybridization (FISH) techniques used to localize GPA2 and GPB5 transcripts, along with immunohistochemical detection of GPB5, did not identify specific cells in these regions of the nervous system. Instead, GPA2 and GPB5 transcripts, as well as GPB5 immunoreactivity was identified in 2–3 laterally-localized bilateral pairs of neuroendocrine cells within the first five abdominal ganglia. These findings are consistent with previous findings in the fruit fly *D. melanogaster*, where GPA2 and GPB5 subunit transcripts were localized to four bilateral pairs of neuroendocrine cells within the abdominal neuromeres in the fused ventral nerve cord, that were distinct from cells expressing other neuropeptides including leucokinin, bursicon, crustacean cardioactive peptide or calcitonin-like diuretic hormone (18).

Both FISH and immunohistochemical techniques revealed GPA2 and GPB5 expression within two bilateral pairs of neuroendocrine cells, that were positioned slightly posterior to where the lateral nerves emanate from these ganglia. It

is possible that GPB5 immunoreactive axons follow similar projection patterns as ovary ecdysteroidogenic hormone (OEH I), a gonadotropin produced in lateral neurosecretory cells in the abdominal ganglia of mosquitoes (34). OEH I positive cells emanate through the lateral nerves, similar to our findings with GPB5 immunoreactive cells, and terminate on perivisceral organs that function as neurohaemal release sites (34).

In some abdominal ganglia preparations, a third bilateral pair of cells immunoreactive for GPB5 protein was detected; however these additional cells were not detected using FISH, which suggests GPB5 transcript may be differentially regulated between different bilateral pairs of cells. Given that the same number of cells were observed to express GPA2 transcript, and these cells localized to similar positions as GPB5 expressing cells, the glycoprotein hormone subunits are likely co-expressed in the same cells. To investigate cellular co-expression of GPA2/GPB5, abdominal ganglia were simultaneously treated with both GPA2- and GPB5-targeted anti-sense RNA probes. From this analysis, again only two bilateral pairs of cells were detected and these were more intensely stained compared to preparations treated with either probe alone, which confirms that GPA2 and GPB5 are indeed co-expressed within the same neurosecretory cells of the first five abdominal ganglia in adult mosquitoes. The cellular co-expression of GPA2 and GPB5 proteins implies that, upon a given stimulus, both subunits are regulated in a similar manner and are likely simultaneously released following the appropriate stimulus. Importantly, since co-expression and heterodimerization of the classic vertebrate glycoprotein hormone subunits takes place within the same cells (21, 22), these findings indicate that the mosquito GPA2/GPB5 subunits may be produced and released as heterodimers *in vivo*.

Heterodimerization and Homodimerization of GPA2/GPB5

To study the interactions of *A. aegypti* GPA2 and GPB5 subunits *in vitro*, hexa-histidine tagged proteins secreted into the culture media of transfected HEK 293T cells were collected and analyzed under denaturing conditions after cross-linking treatments, which had been utilized previously to show GPA2/GPB5 heterodimerization in other organisms (4, 6, 29). Cross-linked protein samples were then deglycosylated to identify whether the removal of N-linked sugars affected dimerization. Treatment of mosquito GPA2 and GPB5 subunits individually with cross-linker resulted in the detection of bands with sizes corresponding to homodimers of GPA2 (~32 kDa) and GPB5 (48 kDa). GPA2 homodimer bands migrated lower to ~30 kDa following deglycosylation treatment with PNGase. However, experiments performed with cross-linked GPA2/GPB5 protein were not able to confirm heterodimerization since the detected band sizes could also reflect GPB5 homodimeric interactions. Similarly, previous studies that demonstrated GPA2/GPB5 heterodimerization in human (4, 6) and fruit fly (29) did not provide evidence on the interactions of each subunit alone to determine if homodimerization is possible. In these earlier studies, molecular weight band sizes that were identified as heterodimers could have been the result of homodimeric

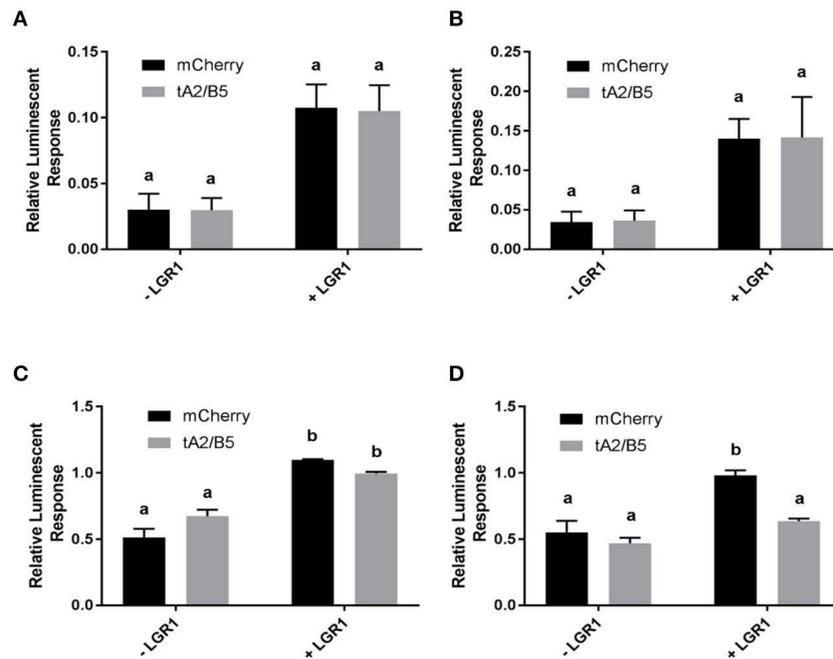


FIGURE 7 | cAMP-mediated bioluminescence assay determining the effects of tethered GPA2/GPB5 on receptor activation and G protein signaling of LGR1. Secreted protein fractions (A,C) and cell lysates (B,D) derived from cells transiently expressing tethered GPA2/GPB5 (tA2/B5) or red fluorescent protein (mCherry) were tested for their ability to stimulate (Gs signaling) (A,B) or inhibit 1 μ M forskolin-induced (Gi/o signaling) (C,D) cAMP-mediated luminescence. (A–D) Luminescence response was recorded and shown as normalized to the luminescence from sole treatments with 1 μ M forskolin on LGR1-expressing cells. (C,D) In treatments involving mCherry proteins, the relative luminescence response was greater in LGR1-expressing cells (+ LGR1) compared to cells not expressing LGR1 (–LGR1). (A,B) Incubation of LGR1-expressing cells with tA2/B5 secreted (A) or cell lysate (B) proteins, failed to increase cAMP-mediated luminescence above background levels from incubations with mCherry proteins. (C,D) 1 μ M forskolin along with either mCherry proteins or tA2/B5 proteins was added to cells in the presence or absence of LGR1 expression. The ability for each ligand treatment to reduce a forskolin-induced increase in cAMP luminescence was examined. (C) The tA2/B5 secreted protein samples incubated with LGR1-expressing cells did not significantly affect the forskolin-induced cAMP luminescence, compared to control treatments with secreted proteins from mCherry expressing cells. (D) Relative to incubations with mCherry cell lysate proteins, cell lysates containing tA2/B5 protein significantly inhibited forskolin-induced elevations of cAMP-mediated luminescence, and this inhibition was specific to LGR1-expressing cells. Mean \pm SEM of three biological replicates. Columns denoted with different letters are significantly different from one another. Multiple comparisons two-way ANOVA test with Tukey's multiple comparisons ($P < 0.05$).

interactions (4, 6, 29). As a result, the findings herein with mosquito GPA2/GPB5 indicate additional experiments are required to confirm GPA2/GPB5 heterodimerization in these organisms.

To clarify whether *A. aegypti* (mosquito) GPA2/GPB5 heterodimerize, each subunit was differentially tagged (GPA2-FLAG and GPB5-His), and immunoblots containing various combinations of cross-linked subunits were probed with either anti-FLAG or anti-His antibody. As a positive control, experiments were first performed using *H. sapiens* (human) hGPA2/hGPB5 subunit proteins. Similar to mosquito GPA2 and GPB5 subunits, results showed hGPA2 and hGPB5 subunits are capable of homodimerization. To study heterodimeric interactions, GPA2 and GPB5 subunit proteins were expressed separately in HEK 293T cells. Upon combining and treating protein samples with DSS, a molecular weight band size at 38 kDa, corresponding to the molecular weight of hGPA2/hGPB5 heterodimers (hGPA2-FLAG and hGPB5-His tagged), was detected and migrated differently than bands corresponding to hGPA2 (40 kDa) and hGPB5 (36 kDa) homodimers. Taken

together, our results confirm hGPA2 and hGPB5 subunits are indeed capable of heterodimerization *in vitro*. Further, an induction of cAMP was observed when both subunits were present for TSHR functional activation, whereas treatments with individual subunits failed to significantly increase cAMP-mediated luminescence. As a result, hGPA2/hGPB5 is capable of heterodimerization, and since a combination of both subunits were required to signal a TSHR-mediated elevation in cAMP, these heterodimers are required to functionally activate its cognate glycoprotein hormone receptor.

Analogous experiments using the same concentration of DSS cross-linker (data not shown) as well as three-fold higher concentrations were performed with mosquito GPA2-FLAG and GPB5-His subunit proteins. Irrespective of the DSS concentration used, no bands corresponding to the expected molecular weights of mosquito GPA2/GPB5 heterodimers (37–40 kDa) were observed, but rather, only GPA2 and GPB5 homodimers were detected. Moreover, mosquito GPA2/GPB5, either each subunit alone, mixed from different cell batches or co-expressed in the same cells using a dual promoter vector,

was unable to stimulate a cAMP-mediated luminescent response in HEK 293T cells expressing LGR1. Since we identified *A. aegypti* LGR1 is predicted to couple a Gi/o signaling pathway (Table S3), we examined if mosquito GPA2/GPB5 could inhibit a forskolin-induced cAMP response. Sole treatments of GPA2 and GPB5 subunit proteins inhibited a forskolin-induced rise in cAMP, however these inhibitory actions were not owed to G protein signaling events related to *A. aegypti* LGR1, since inhibition was observed in control cell lines in the absence of LGR1. These results suggest mosquito GPA2 and GPB5 subunit proteins may non-specifically interact with other endogenously expressed proteins, like the orphan receptors LGR4 and LGR5 that are highly expressed in HEK 293T cells (35).

A. aegypti GPA2/GPB5 Heterodimers Activate LGR1 and Initiate a Switch From Gs to Gi Coupling

Activation of the human thyrotropin receptor was only observed when both hGPA2 and hGPB5 subunits were present which, unlike data obtained involving mosquito subunits, demonstrated hGPA2/hGPB5 subunit heterodimerization *in vitro* using the mammalian heterologous system. The inability of mosquito GPA2 and GPB5 subunits to successfully heterodimerize could result from improper protein folding of insect-derived secretory proteins in HEK 293T cells used for heterologous expression. As a result, future experiments should test the expression and heterodimerization of *A. aegypti* GPA2/GPB5 in other heterologous systems such as insect cell lines, which could provide a more appropriate environment for tertiary and quaternary protein structure formation. Nonetheless, given the observed co-localization of the subunits within bilateral pairs of cells in the first five abdominal ganglia in *A. aegypti*, which is comparable to cellular colocalization shown earlier in *D. melanogaster* (18), we proposed that *A. aegypti* GPA2/GPB5 heterodimers would be required to functionally activate LGR1 *in vitro*. To confirm this possibility, a tethered construct was designed linking the C-terminus of GPB5 to the N-terminus of GPA2 using a histidine tagged glycine/serine-rich linker sequence. Natural and synthetic linkers function as spacers that connect multidomain proteins, and are commonly used to study unstable or weak protein-protein interactions (36). The incorporation of a linker sequence between glycoprotein hormone subunits has been performed previously, and does not affect the assembly, secretion or bioactivity of human FSH (37), TSH (38), and CG (39). The conversion of two independent glycoprotein hormone subunits into a single polypeptide chain using a glycine-serine repeat linker sequence has also been performed recently with lamprey GPA2/GPB5 (31), which was shown to induce a cAMP response. Interestingly, similar proteins involving TSH alpha fused to TSH beta with carboxyl-terminal peptide (CTP) as a linker promoted a three-fold higher induction of cAMP compared to wild-type TSH, likely because the addition of a CTP linker increases protein stability and flexibility (40).

Thus, tethered *A. aegypti* GPA2/GPB5 was expressed in HEK 293T cells and secreted protein fractions along with cell lysates were collected for expression studies. In cell lysates,

immunoblot studies revealed the detection of two bands at 32 and 37 kDa which correlate to immature (incompletely processed) forms of tethered GPA2/GPB5 proteins and non-glycosylated GPA2 (13 kDa) plus GPB5 (24 kDa), respectively. In secreted fractions, a faint band corresponding to non-glycosylated GPA2/GPB5 as well as an intense band at 40 kDa, glycosylated GPA2 (16 kDa) plus GPB5 (24 kDa), was detected. Moreover, treatments with PNGase verified the higher molecular weight band size in secreted fractions at 40 kDa is glycosylated, as observed for GPA2 expressed independently in earlier experiments herein and in previous studies (13). Bands in cell lysates indicate the retention of immature and non-glycosylated GPA2/GPB5 heterodimers. For this reason, cell lysates and secreted protein fractions of tethered GPA2/GPB5-expressing cells were separately tested for their ability to activate *A. aegypti* LGR1. Unlike secreted fractions that were more rich in glycosylated GPA2/GPB5 proteins, cell lysates that lacked glycosylated GPA2/GPB5 were able to activate LGR1, leading to a reduction in cAMP-mediated luminescence. Thus, our results suggest that post-translational modifications like glycosylation may impact ligand-receptor activity. Alternatively, it is possible that a lack of inhibitory activity of secreted fractions could be owed to smaller concentrations of tethered GPA2/GPB5 compared to lysate fractions, given that the amounts of secreted and cell lysate fractions were not equalized. Future experiments should aim to create novel constructs that improve expression and secretion of tethered GPA2/GPB5.

Evolutionary Conservation and Divergence of GPA2/GPB5 Receptor Signaling

In humans, GPA2/GPB5-TSHR signaling stimulates adenylyl cyclase activity to increase intracellular cAMP via interaction with a Gs protein (4, 6, 11), and these results were confirmed in our studies. Comparatively, GPA2/GPB5 signaling was also shown to increase levels of cAMP upon binding LGR1 in *D. melanogaster* (29). Interestingly, our experiments demonstrate low level constitutive activity of adenylyl cyclase in LGR1 expressing cells since cAMP luminescent response was moderately greater in LGR1-transfected cells compared to cells not expressing LGR1. High basal constitutive activity of LGR1 has been observed previously in *D. melanogaster* (29). Moreover, constitutive activity of glycoprotein hormone receptors is well-known in vertebrates and has been demonstrated to be stronger for the thyrotropin receptor than for the LH/CG receptor (41). Given that GPA2/GPB5 receptor homologs in vertebrate and invertebrate model organisms have been shown previously, and herein, to promote constitutive synthesis of cAMP, it may be suggestive that a common downstream effector and/or function exists for LGR1 and homologs in different organisms.

Surprisingly, however, our experiments indicate that incubations of LGR1-expressing cells with mosquito GPA2/GPB5 tethered protein triggers a switch from low level constitutive Gs coupling to Gi/o coupling for *A. aegypti* LGR1, given that tethered GPA2/GPB5 significantly inhibited the forskolin-induced increase in cAMP in LGR1-expressing HEK 293T cells but not in control cells lacking LGR1 expression. This

finding, while highly interesting, is not entirely unusual since promiscuous G protein coupling has been reported for glycoprotein hormone receptors like the TSH receptor (Gs and Gq) and LH/CG (Gi and Gs) (42–44).

Regulation by GPA2/GPB5 Heterodimers

To help stabilize heterodimerization, the beta subunit sequences of the classic glycoprotein hormones (FSH, LH, TSH, and CG) contain two additional cysteine residues that form an additional disulfide bridge which wraps around and “buckles” the alpha subunit (26). Though heterodimerization can occur with mutated forms of this “seatbelt” structure, there is a dramatic decrease in heterodimer stability (21, 22). GPB5 in vertebrates and invertebrates lack the seatbelt structure required to stabilize heterodimerization (26). Thus, the hypothesis that GPA2/GPB5 functions as a heterodimer in a physiological situation (i.e., without chemical cross-linking) is challenged. The dissociation constant (Kd) associated with heterodimerization of the classic glycoproteins hormone subunits and GPA2/GPB5 is 10^{-7} to 10^{-6} M, which indicates heterodimeric interactions are favored at these concentrations (45, 46). However, since the classic beta subunits contain an additional disulfide bridge that strengthens its association with the common alpha subunit, heterodimeric interactions are stabilized in circulation at physiological concentrations as low as 10^{-11} to 10^{-9} M (21). Without this seatbelt structure, GPA2/GPB5 heterodimeric interactions are possible only at micromolar concentrations, which are not typically observed in circulation (21, 26). Together, the limited evidence so far using heterologous expression challenges the possibility of endocrine regulation by GPA2/GPB5 heterodimers. Nonetheless, it was argued in *D. melanogaster* that the large neurosecretory cells co-expressing the glycoprotein hormone subunits, along with their corresponding axonal projections that localized distinctly from organs that express the GPA2/GPB5 receptor (LGR1), did support that this system is indeed endocrine in nature (18). Alternatively, the subunits could function independently or regulate physiology as a heterodimer in a paracrine/autocrine fashion. In rats, GPA2/GPB5 is expressed in oocytes and may act as a paracrine regulator of TSHR-expressed granulosa cells in the ovary to regulate reproductive processes (6). Another possibility to consider is that additional endogenous co-factors may be involved, but remain unidentified, which help to strengthen interaction between the GPA2 and GPB5 subunits, since the tethered mosquito GPA2/GPB5 was indeed capable of activating LGR1 *in vitro* inducing a Gi/o signaling cascade.

GPA2 and GPB5 Homodimerization

Our results establish that human and mosquito GPA2 and GPB5 subunits can weakly and strongly, respectively, homodimerize. However, whether these homodimers have a physiological function *in vivo* is unknown. Treatments of either mosquito or human GPA2 and GPB5 subunits alone did not stimulate specific downstream signaling in LGR1- or TSHR-expressing cells, upholding that only GPA2/GPB5 heterodimers can activate their cognate glycoprotein hormone receptors. However, it is possible that GPA2 and GPB5 homodimers may target other unidentified

receptors. In insects, the molting hormone bursicon is a heterodimer of two subunits called burs and pburs. Burs/pburs heterodimers act via a glycoprotein hormone receptor (i.e., LGR2) to regulate processes such as tanning and sclerotization of the insect cuticle as well as wing inflation after adult emergence (47). Recently, it was demonstrated that bursicon subunits can homodimerize (i.e., burs/burs and pburs/pburs) and these homodimers mediate actions independently of LGR2 to regulate immune responses in *A. aegypti* and *D. melanogaster* (48, 49).

In addition to the human and mosquito GPA2/GPB5 homodimers observed in our studies, human GPA2 was also shown to interact with the beta subunits of CG and FSH (4). Lastly, the expression patterns of GPA2 and GPB5 in a number of organisms do not always strictly co-localize, since GPA2 expression exhibits a much wider distribution and is expressed more abundantly than GPB5 in a number of vertebrate and invertebrate organisms (7, 12, 23, 24, 50, 51). Taken together, this raises the possibility that GPA2 and GPB5 subunits may interact with other unknown proteins that could activate different receptors or signaling pathways and elicit distinct functions.

CONCLUDING REMARKS

Although much is known about the classic vertebrate glycoprotein hormones including LH, FSH, TSH, and CG along with their associated receptors, little progress has been made thus far toward better understanding the function of GPA2/GPB5 signaling and subunit interactions, particularly for the invertebrate organisms. To our knowledge, this is the first study to demonstrate *A. aegypti* and *H. sapiens* GPA2 and GPB5 subunit homodimerization *in vitro*. Our results also confirm that heterodimerization of *A. aegypti* and *H. sapiens* GPA2/GPB5 are required for the activation of their cognate receptors LGR1 and TSHR, respectively. In contrast to previous reports showing GPA2/GPB5-induced LGR1 activation elevates intracellular cAMP by coupling a Gs pathway, the current findings provide novel information supporting that *A. aegypti* LGR1 couples to a Gi/o protein to inhibit cAMP levels following application of the GPA2/GPB5 fusion polypeptide. Further, our results revealed that mosquito LGR1 is constitutively active when overexpressed in the absence of its ligand, GPA2/GPB5, inducing a Gs signaling pathway that raises levels of cAMP levels, which is consistent with previous observations with overexpression of fruit fly LGR1 (29, 52) as well as mammals including dog and human TSH receptor (53, 54).

In the mosquito nervous system, our results confirm GPA2 and GPB5 subunits are co-expressed within the same neurosecretory cells of the first five abdominal ganglia where their coordinated release and regulation are likely. As a result, whether GPA2/GPB5 are secreted as heterodimers, like the classic glycoprotein hormones, and/or as homodimers remains to be determined *in vivo*. While homodimers were inactive in the heterologous assay used herein, whether these homodimers are functional *in vivo* and what physiological role they play (if any) is a research direction that should be addressed in future studies. All in all, this investigation has provided novel information

for an invertebrate GPA2/GPB5 and LGR1 signaling system and contributes toward advancing our understanding and the functional elucidation of this ancient glycoprotein hormone signaling system common to nearly all bilaterian organisms.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

DR and J-PP contributed to experimental design, wrote the paper, and aided with data analysis. DR performed all the experiments.

REFERENCES

- Hartree AS, Renwick AGC. Molecular structures of glycoprotein hormones and functions of their carbohydrate components. *Biochem J.* (1992) 287(Pt 3):665–79. doi: 10.1042/bj2870665
- Pierce JG, Parsons TF. Glycoprotein hormones: structure and function. *Annu Rev Biochem.* (1981) 50:465–95. doi: 10.1146/annurev.bi.50.070181.002341
- Hsu SY, Nakabayashi K, Bhalla A. Evolution of glycoprotein hormone subunit genes in bilateral metazoa: identification of two novel human glycoprotein hormone subunit family genes, GPA2 and GPB5. *Mol Endocrinol.* (2002) 16:1538–51. doi: 10.1210/mend.16.7.0871
- Nakabayashi K, Matsumi H, Bhalla A, Bae J, Mosselman S, Hsu SY, et al. Thyrostimulin, a heterodimer of two new human glycoprotein hormone subunits, activates the thyroid-stimulating hormone receptor. *J Clin Invest.* (2002) 109:1445–52. doi: 10.1172/JCI0214340
- Dos Santos S, Mazan S, Venkatesh B, Cohen-Tannoudji J, Quérat B. Emergence and evolution of the glycoprotein hormone and neurotrophin gene families in vertebrates. *BMC Evol Biol.* (2011) 11:332. doi: 10.1186/1471-2148-11-332
- Sun SC, Hsu PJ, Wu FJ, Li SH, Lu CH, Luo CW. Thyrostimulin, but not thyroid-stimulating hormone (TSH), acts as a paracrine regulator to activate the TSH receptor in mammalian ovary. *J Biol Chem.* (2010) 285:3758–65. doi: 10.1074/jbc.M109.066266
- Okada SL, Ellsworth JL, Durnam DM, Haugen HS, Holloway JL, Kelley ML, et al. A glycoprotein hormone expressed in corticotrophs exhibits unique binding properties on thyroid-stimulating hormone receptor. *Mol Endocrinol.* (2006) 20:414–25. doi: 10.1210/me.2005-0270
- Bassett JHD, van der Spek A, Logan JG, Gogakos A, Chakraborty JB, Murphy E, et al. Thyrostimulin regulates osteoblastic bone formation during early skeletal development. *Endocrinology.* (2015) 159:3098–113. doi: 10.1210/en.2014-1943
- Suzuki C, Nagasaki H, Okajima Y, Suga H, Ozaki N, Arima H, et al. Inflammatory cytokines regulate glycoprotein subunit beta5 of thyrostimulin through nuclear factor-kappaB. *Endocrinology.* (2009) 150:2237–43. doi: 10.1210/en.2008-0823
- van Zeijl CJJ, Surovtseva OV, Kwakkel J, van Beeren HC, Duncan Bassett JH, Williams GR, et al. Thyrostimulin deficiency does not alter peripheral responses to acute inflammation-induced nonthyroidal illness. *Am J Physiol Endocrinol Metab.* (2014) 307:E527–37. doi: 10.1152/ajpendo.00266.2014
- Huang W, Li Z, Lin T, Wang S, Wu F, Luo C. Thyrostimulin-TSHR signaling promotes the proliferation of NIH : OVCAR-3 ovarian cancer cells via trans-regulation of the EGFR pathway. *Nat Publishing Group.* (2016) 6:1–13. doi: 10.1038/srep27471
- Heyland A, Plachetzki D, Donnelly E, Gunaratne D, Bobkova Y, Jacobson J, et al. Distinct expression patterns of glycoprotein hormone subunits in the lophotrochozoan *Aplysia*: implications for the evolution of neuroendocrine systems in animals. *Endocrinology.* (2012) 153:5440–51. doi: 10.1210/en.2012-1677

FUNDING

This study was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and an Ontario Ministry of Research and Innovation Early Researcher Award to J-PP. This manuscript has been released as a Pre-Print on BioRxiv available at <https://www.biorxiv.org/content/10.1101/694653v2>.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2020.00158/full#supplementary-material>

- Paluzzi J-P, Vanderveken M, O'Donnell MJ. The heterodimeric glycoprotein hormone, GPA2/GPB5, regulates ion transport across the hindgut of the adult mosquito, *Aedes aegypti*. *PLoS ONE.* (2014) 9:e86386. doi: 10.1371/journal.pone.0086386
- Rocco DA, Paluzzi J-PV. Functional role of the heterodimeric glycoprotein hormone, GPA2/GPB5, and its receptor, LGR1: An invertebrate perspective. *Gen Comp Endocrinol.* (2016) 234:20–7. doi: 10.1016/j.ygcen.2015.12.011
- Vandersmissen HP, Van Hiel MB, Van Loy T, Vleugels R, Vanden Broeck J. Silencing D. *melanogaster* *lgr1* impairs transition from larval to pupal stage. *Gen Comp Endocrinol.* (2014) 209:135–47. doi: 10.1016/j.ygcen.2014.08.006
- Jourjine N, Mullaney BC, Mann K, Scott K. Coupled sensing of hunger and thirst signals balances sugar and water consumption. *Cell.* (2016) 166:855–66. doi: 10.1016/j.cell.2016.06.046
- Rocco DA, Kim DH, Paluzzi JPV. Immunohistochemical mapping and transcript expression of the GPA2/GPB5 receptor in tissues of the adult mosquito, *Aedes aegypti*. *Cell Tissue Res.* (2017) 369:313–30. doi: 10.1007/s00441-017-2610-3
- Sellami A, Agricola HJ, Veenstra JA. Neuroendocrine cells in *Drosophila melanogaster* producing GPA2/GPB5, a hormone with homology to LH, FSH and TSH. *Gen Comp Endocrinol.* (2011) 170:582–8. doi: 10.1016/j.ygcen.2010.11.015
- Rocco DA, Garcia ASG, Scudeler EL, Santos DC, Nóbrega RH, Paluzzi JPV. Glycoprotein hormone receptor knockdown leads to reduced reproductive success in male *Aedes aegypti*. *Front Physiol.* (2019) 10:266. doi: 10.3389/fphys.2019.00266
- Combarrous Y. Molecular basis of the specificity of binding of glycoprotein hormones to their receptors. *Endocr Rev.* (1992) 13:670–91. doi: 10.1210/edrv-13-4-670
- Galet C, Lecompte F, Combarrous Y. Association/dissociation of gonadotropin subunits involves disulfide bridge disruption which is influenced by carbohydrate moiety. *Biochem Biophys Res Commun.* (2004) 324:868–73. doi: 10.1016/j.bbrc.2004.09.143
- Xing Y, Myers R, Cao D, Lin W, Jiang M, Bernard M, et al. Glycoprotein hormone assembly in the endoplasmic reticulum: II. Multiple roles of a redox sensitive (beta)-subunit disulfide switch. *J Biol Chem.* (2004) 279:35437–48. doi: 10.1074/jbc.M403053200
- Dos Santos S, Bardet C, Bertrand S, Escrive H, Habert D, Querat B. Distinct expression patterns of glycoprotein hormone-alpha2 and -beta5 in a basal chordate suggest independent developmental functions. *Endocrinology.* (2009) 150:3815–22. doi: 10.1210/en.2008-1743
- Nagasaki H, Wang Z, Jackson VR, Lin S, Nothacker HP, Civelli O. Differential expression of the thyrostimulin subunits, glycoprotein alpha2 and beta5 in the rat pituitary. *J Mol Endocrinol.* (2006) 37:39–50. doi: 10.1677/jme.1.01932
- Tando Y, Kubokawa K. A homolog of the vertebrate thyrostimulin glycoprotein hormone alpha subunit (GPA2) is expressed in *Amphioxus* neurons. *Zoological Sci.* (2009) 26:409–14. doi: 10.2108/zsj.26.409
- Alvarez E, Cahoreau C, Combarrous Y. Comparative structure analyses of cystine knot-containing molecules with eight aminoacyl ring including

- glycoprotein hormones (GPH) alpha and beta subunits and GPH-related A2 (GPA2) and B5 (GPB5) molecules. *Reprod Biol Endocrinol.* (2009) 7:90. doi: 10.1186/1477-7827-7-90
27. Krause G, Kreuchwig A, Kleinau G. Extended and structurally supported insights into extracellular hormone binding, signal transduction and organization of the thyrotropin receptor. *PLoS ONE.* (2012) 7:e52920. doi: 10.1371/journal.pone.0052920
 28. Hauser F, Nothacker HP, Grimmlikhuijzen CJ. Molecular cloning, genomic organization, and developmental regulation of a novel receptor from *Drosophila melanogaster* structurally related to members of the thyroid-stimulating hormone, follicle-stimulating hormone, luteinizing hormone/choriogonadotropin. *J Biol Chem.* (1997) 272:1002–10. doi: 10.1074/jbc.272.2.1002
 29. Sudo S, Kuwabara Y, Park J, Il Sheau YH, Hsueh AJW. Heterodimeric fly glycoprotein hormone- α 2 (GPA2) and glycoprotein hormone- β 5 (GPB5) activate fly leucine-rich repeat-containing G protein-coupled receptor-1 (DLGR1) and stimulation of human thyrotropin receptors by chimeric fly GPA2 and human GPB5. *Endocrinology.* (2005) 146:3596–604. doi: 10.1210/en.2005-0317
 30. Paluzzi JP, Russell WK, Nachman RJ, Orchard I. Isolation, cloning, and expression mapping of a gene encoding an antidiuretic hormone and other CAPA-related peptides in the disease vector, *Rhodnius prolixus*. *Endocrinology.* (2008) 149:4638–46. doi: 10.1210/en.2008-0353
 31. Sower SA, Decatur WA, Hausken KN, Marquis TJ, Barton SL, Gargan J, et al. Emergence of an ancestral glycoprotein hormone in the pituitary of the Sea Lamprey, a Basal Vertebrate. *Endocrinology.* (2015) 156:3026–37. doi: 10.1210/en.2014-1797
 32. Wang P, Liu S, Yang Q, Liu Z, Zhang S. Functional characterization of thyrostimulin in *Amphioxus* suggests an ancestral origin of the TH signaling pathway. *Endocrinology.* (2018) 159:3536–48. doi: 10.1210/en.2018-00550
 33. Sgourakis NG, Bagos PG, Papasaikas PK, Hamodrakas SJ. A method for the prediction of GPCRs coupling specificity to G-proteins using refined profile Hidden Markov Models. *BMC Bioinformatics.* (2005) 12:1–12. doi: 10.1186/1471-2105-6-104
 34. Brown MR, Cao C. Distribution of ovary ecdysteroidogenic hormone I in the nervous system and gut of mosquitoes. *J Insect Sci.* (2001) 1:1–11. doi: 10.1673/031.001.0301
 35. Atwood B, Lopez J, Wager-Miller J, Mackie K, Straiker A. Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. *BMC Genomics.* (2011) 12:14. doi: 10.1186/1471-2164-12-14
 36. Priyanka V, Chichili R, Kumar V, Sivaraman J. Linkers in the structural biology of protein – protein interactions. *Protein Sci.* (2013) 22:153–167. doi: 10.1002/pro.2206
 37. Fares F, Suganuma N, Nishimori K, LaPolt P, Hsueh A, Biome I. Design of a long-acting follitropin agonist by fusing the C-terminal sequence of the chorionic gonadotropin β subunit to the follitropin β subunit. *Proc Natl Acad Sci USA.* (1992) 89:4304–8. doi: 10.1073/pnas.89.10.4304
 38. Joshi L, Murata Y, Wondisford F, Szkudlinski M, Desai R, Weintraub B. Recombinant thyrotropin containing a β -subunit chimera with the human chorionic gonadotropin- β carboxy terminal is biologically active with a prolonged plasma half-life: role of carbohydrate in bioactivity and metabolic clearance. *Endocrinology.* (1995) 136:3839–948. doi: 10.1210/endo.136.9.7544273
 39. Furuhashi M, Shikone T, Fares F, Sugahara T, Hsueh A, Boime I. Fusing the carboxy-terminal peptide of the chorionic gonadotropin (CG) β -subunit to the common α -subunit: retention of O-linked glycosylation and enhanced *in vivo* bioactivity of chimeric human CG. *Mol Endocrinol.* (1995) 9:54–63. doi: 10.1210/mend.9.1.7539107
 40. Azzam N, Bar-Shalom R, Fares F. Conversion of TSH heterodimer to a single polypeptide chain increases bioactivity and longevity. *Endocrinology.* (2012) 153:954–60. doi: 10.1210/en.2011-1856
 41. Cetani F, Tonacchera M, Vassart G. Differential effects of NaCl concentration on the constitutive activity of the thyrotropin and the luteinizing hormone/chorionic gonadotropin receptors. *FEBS Lett.* (1996) 378:27–31. doi: 10.1016/0014-5793(95)01384-9
 42. Grasberger H, Van Sande J, Hag-Dahood M, Tenenbaum-Rakover Y, Refetoff S. A familial thyrotropin (TSH) receptor mutation provides *in vivo* evidence that the inositol phosphates/Ca²⁺ cascade mediates TSH action on thyroid hormone synthesis. *J Clin Endocrinol Metab.* (2007) 92:2816–20. doi: 10.1210/jc.2007-0366
 43. Herrlich A, Kuhn B, Grosse R, Schmid A, Schultz G, Gudermann T. Involvement of Gs and Gi proteins in dual coupling of the luteinizing hormone receptor to adenylyl cyclase and phospholipase C. *J Biol Chem.* (1996) 271:16764–72. doi: 10.1074/jbc.271.28.16764
 44. Kleinau G, Jaeschke H, Worth C, Mueller S, Gonzalez J, Paschke R, et al. Principles and determinants of G-protein coupling by the rhodopsin-like thyrotropin receptor. *PLoS ONE.* (2010) 5:e9745. doi: 10.1371/journal.pone.0009745
 45. Ingham K, Weintraub B, Edelhock H. Kinetics of recombination of the subunits of human chorionic gonadotropin. Effect of subunit concentration. *Biochemistry.* (1976) 15:1720–6. doi: 10.1021/bi00653a020
 46. Strickland TW, Puett D. The kinetic and equilibrium parameters of subunit association and gonadotropin dissociation. *J Biol Chem.* (1982) 257:2954–60.
 47. Luo C-W, Dewey EM, Sudo S, Ewer J, Hsu SY, Honegger H-W, et al. Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2. *Proc Natl Acad Sci USA.* (2005) 102:2820–5. doi: 10.1073/pnas.0409916102
 48. An S, Dong S, Wang Q, Li S, Gilbert LI, Stanley D, et al. Insect neuropeptide bursicon homodimers induce innate immune and stress genes during molting by activating the NF- κ B transcription factor relish. *PLoS ONE.* (2012) 7:e34510. doi: 10.1371/journal.pone.0034510
 49. Zhang H, Dong S, Xi C, Stanley D, Beerntsen B. Relish2 mediates bursicon homodimer-induced prophylactic immunity in the mosquito *Aedes aegypti*. *Sci Rep.* (2017) 7:43163. doi: 10.1038/srep43163
 50. Tando Y, Kubokawa K. Expression of the gene for ancestral glycoprotein hormone β subunit in the nerve cord of *Amphioxus*. *Gen Comp Endocrinol.* (2009) 162:329–39. doi: 10.1016/j.ygcn.2009.04.015
 51. Trudeau VL. Really old hormones up to new tricks: glycoprotein hormone subunits may have roles in development. *Endocrinology.* (2009) 150:3446–7. doi: 10.1210/en.2009-0465
 52. Nishi S, Hsu SY, Zell K, Hsueh AJ. Characterization of two fly LGR (leucine-rich repeat-containing, G protein-coupled receptor) proteins homologous to vertebrate glycoprotein hormone receptors. *Endocrinology.* (2000) 141:4081–90. doi: 10.1210/endo.141.11.7744
 53. Sugawa ZML, Kosugi S, Mori T. Constitutive activation of the thyrotropin receptor by deletion of a portion of the extracellular domain. *Biochem Biophys Res Commun.* (1995) 211:205–10. doi: 10.1006/bbrc.1995.1797
 54. Van Sande J, Swillens S, Gerard C, Allgeier A, Massart C, Vassart G, et al. In Chinese hamster ovary K1 cells dog and human thyrotropin receptors activate both the cyclic AMP and the phosphatidylinositol 4,5-bisphosphate cascades in the presence of thyrotropin and the cyclic AMP cascade in its absence. *Eur J Biochem.* (1995) 229:338–43. doi: 10.1111/j.1432-1033.1995.0338k.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Rocco and Paluzzi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Adipokinetic Peptides in Diptera: Structure, Function, and Evolutionary Trends

Gerd Gäde^{1*}, Petr Šimek^{2*} and Heather G. Marco¹

¹ Department of Biological Sciences, University of Cape Town, Cape Town, South Africa, ² Biology Centre, Czech Academy of Sciences, České Budejovice, Czechia

OPEN ACCESS

Edited by:

Klaus H. Hoffmann,
University of Bayreuth, Germany

Reviewed by:

Christian Wegener,
Julius Maximilian University of
Würzburg, Germany
Mark R. Brown,
University of Georgia, United States

*Correspondence:

Gerd Gäde
gerd.gade@uct.ac.za
Petr Šimek
simek@bclab.eu

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 03 January 2020

Accepted: 04 March 2020

Published: 31 March 2020

Citation:

Gäde G, Šimek P and Marco HG
(2020) The Adipokinetic Peptides in
Diptera: Structure, Function, and
Evolutionary Trends.
Front. Endocrinol. 11:153.
doi: 10.3389/fendo.2020.00153

Nineteen species of various families of the order Diptera and one species from the order Mecoptera are investigated with mass spectrometry for the presence and primary structure of putative adipokinetic hormones (AKHs). Additionally, the peptide structure of putative AKHs in other Diptera are deduced from data mining of publicly available genomic or transcriptomic data. The study aims to demonstrate the structural biodiversity of AKHs in this insect order and also possible evolutionary trends. Sequence analysis of AKHs is achieved by liquid chromatography coupled to mass spectrometry. The corpora cardiaca of almost all dipteran species contain AKH octapeptides, a decapeptide is an exception found only in one species. In general, the dipteran AKHs are order-specific- they are not found in any other insect order with two exceptions only. Four novel AKHs are revealed by mass spectrometry: two in the basal infraorder of Tipulomorpha and two in the brachyceran family Syrphidae. Data mining revealed another four novel AKHs: one in various species of the infraorder Culicimorpha, one in the brachyceran superfamily Asiloidea, one in the family Diopsidae and in a Drosophilidae species, and the last of the novel AKHs is found in yet another *Drosophila*. In general, there is quite a biodiversity in the lower Diptera, whereas the majority of the cyclorraphan Brachycera produce the octapeptide Phote-HrTH. A hypothetical molecular peptide evolution of dipteran AKHs is suggested to start with an ancestral AKH, such as Glomo-AKH, from which all other AKHs in Diptera to date can evolve via point mutation of one of the base triplets, with one exception.

Keywords: diptera, adipokinetic peptides, mass spectrometry, adipokinetic and hypertrahalosemic biological assays, fly phylogeny

INTRODUCTION

There are surely a few facts about the importance of Diptera for ecology, medicine and agriculture that are not so well-known to the general public and non-dipteran specialists alike. Most associations with “flies” in general are about the role of mosquitoes as vectors of terribly infectious diseases, such as malaria or yellow fever, or about the use of the vinegar fly *Drosophila melanogaster* as model organism for genetic experiments into development, aging, behavior, metabolism and diseases. Mosquitoes (Family: Culicidae) are widely known as “detrimental” dipteran species, for they are vectors for Dengue-, West Nile- Zika- and yellow fever, as well as for malaria and encephalitis (https://www.who.int/neglected_diseases/vector_ecology/mosquito-borne-diseases/en/). Certain house flies (Family: Muscidae) and blow flies (Family: Calliphoridae)

transmit disease microorganisms that lead to, e.g., cholera, dysentery and gastroenteritis in humans (1), and from horse flies (Family: Tabanidae), diseases may be effected in other animals, such as equine infectious anemia, and anthrax in cattle and sheep (2). Next to such detrimental effectors, it is perhaps understandable that many people overlook the fact that there are also beneficial fly species, e.g., hover flies (Family: Syrphidae); these are the second most important pollinators after the Hymenoptera (3). Additionally, there are specialist pollinators known from other dipteran families, such as the long-tongued horse flies (Family: Tabanidae) or the cocoa-tree pollinator midges (Family: Ceratopogonidae) (3). Other benign functions are fulfilled by certain fly larvae of the family Calliphoridae, used for medicinal purposes to clean wounds by consuming necrotic tissues, and as producers of pharmacologically active substances such as antimicrobial, antiviral, and antitumor compounds (4). Also, fly larvae of the black soldier fly *Hermetia illucens* (Family: Stratiomyidae), feed voraciously on household and agricultural waste products. These insect decomposers are now reared at industrial scale, and their pupae used for the economic production of oils and dried protein feed for chicken and fish (5). Finally, in forensic entomology the succession order of larvae of various flies on/in the decomposing human corpse is indicative of the time of death (6) and provides yet another example of the positive usefulness of Diptera.

From the above remarks it should be clear that “the flies” are quite a diverse order with respect to being classified as “pest” or “beneficial.” This is not surprising when considering the vast number of fly species: the order Diptera comprises about 152 000 described species (7). Until recently the major divide has been (a) the Nematocera [mostly crane flies (Family: Tipulidae) and mosquitoes, characterized by long antennae] and (b) Brachycera which harbor *inter alia* the above-mentioned families—Tabanidae, Drosophilidae and Muscidae—which are all characterized by short antennae. The newest phylogenomic research takes 149 out of 157 families of Diptera into account and interprets molecular data from 14 nuclear loci and from many complete mitochondrial genomes, as well as 371 morphological characters (8). The comprehensive study presented a phylogenetic tree that shows (i) the monophyly of Diptera with Mecoptera (scorpion flies) and Siphonaptera (fleas) as closest relatives; (ii) paraphyly in the former Nematocera which are now rather called lower Diptera; the rare and species-poor families Deuterophlebiidae (mountain midges) and Nymphomyiidae are at the base of the fly tree, followed by the lower dipteran infraorders Tipulomorpha and Culicomorpha; (iii) the clade Neodiptera comprising of the infraorder Bibionomorpha (marsh flies and gall midges) is the sister group of Brachycera; (iv) the major clades of Brachycera are the Orthorrhapha [Tabanomorpha, Stratiomorpha (such as soldier flies) and Asiloidea (such as robber flies)], the Schizophora in the Cyclorrhapha [Tetratoidea (fruit flies) and Ephydroidea (relatives of *Drosophila*)] and the Calyptrate inside the Cyclorrhapha (such as tsetse flies, house flies and blow flies)] (8).

In the current study, we determine the primary structure of an important metabolic hormone from the so-called

adipokinetic hormone (AKH) family in major families of lower Diptera and Brachycera with the aim to establish how the biodiversity and distribution of various AKH sequences follows the afore-mentioned phylogeny. Previously, we have successfully implemented the use of the primary structure of AKHs in various insect orders to verify certain phylogenetic trends and ancestral relationships [e.g., (9–13)]. The AKH is one of many biologically active neuropeptides in insects. Its major function is regulation of metabolism and it can be compared in a functional respect with the vertebrate hormone, glucagon. Structurally, however, the AKH peptide and its G-protein coupled receptor is related to the vertebrate gonadotropin releasing hormone (GnRH) system, and it is suggested to form a large peptide superfamily together with two other insect neuropeptide systems: corazonin, and adipokinetic hormone/corazonin-related peptide (14–16). AKHs are synthesized and released from the corpus cardiacum (CC), a neurohemal organ. The peptides are characterized by a chain length of 8 to 10 amino acids and post-translationally modified N- and C-termini (pyroglutamate and carboxyamidation, respectively). At position two from the amino terminal, AKHs have either an aliphatic amino acid (leucine, isoleucine, or valine) or an aromatic amino acid (phenylalanine); position three is either a threonine or asparagine residue; at position four one finds either the aromatic phenylalanine or tyrosine; position five has either threonine or serine; position eight is always the aromatic tryptophan and position nine comprises of a simple glycine, whereas at position six, seven, and ten there is a large variability of amino acids possible (17, 18).

The first two AKH peptides of Diptera fully known by primary structure were isolated from the CC of the horse fly *Tabanus atratus* (Family: Tabanidae) and had the sequence of an octapeptide [pELTFTPGW amide] and a decapeptide [pELTFTPGWGY amide; thus, the octapeptide extended by two amino acids, (19)]. The code-names for these peptides were derived from their biological activity ascertained in heterologous biological assays: Tabat-AKH for the octapeptide and Tabat-HoTH for the decapeptide because these peptides, respectively, increased the lipids in the hemolymph (adipokinetic effect) or decreased the concentration of trehalose (hypotrehalosemic effect) in the hemolymph of the face fly *Musca autumnalis* (dipteran Family: Muscidae) (19). Homologous bioassays in the horse fly *Tabanus lineolata* set the record straight and came up with some strikingly different results: the octapeptide, which was called an “adipokinetic” peptide, had no hyperlipemic effect and caused only a slight increase of carbohydrates in the hemolymph of *T. lineolata*, whereas the decapeptide, called a “hypotrehalosemic peptide,” was a very effective adipokinetic and hyperglycemic hormone (20). Regardless of the specific action of these peptides in different insect species, they are generically known as members of the adipokinetic hormone family, in short AKHs.

The next dipteran AKH was elucidated a year after the tabanid peptides, isolated and sequenced from CC of the blowfly *Phormia* (= *Protophormia*) *terraenovae* (Family: Calliphoridae) as an octapeptide [pELTFSPDW amide; (21)] that regulates carbohydrate metabolism and is released during induced stress

(22) and was, hence, code-named Phote-HrTH. This was the first member of the AKH family that had a charged amino acid; all other AKHs up to then were neutral molecules. A peptide with the same sequence was found at the same time in the vinegar fly *D. melanogaster* [Family: Drosophilidae; (23)], where biological assays established its activity as cardioacceleratory in prepupae of *D. melanogaster* (24). Phote-HrTH has subsequently been identified in a number of Dipteran species (including in larval specimens of *Drosophila melanogaster*, see **Table 1**) and is also known for its modulatory action on specific crop muscles of *Phormia regina* (44) that enables delivery of ingested food from the crop to the midgut for digestion. In the post-genomic era, *D. melanogaster* has become the established model insect for research on energy metabolism at the molecular level, with engineered AKH mutants, transgenic insect lines with specific modified cellular components (e.g., transcription factors, ATP-sensitive K⁺ channels, cGMP-dependent protein kinase), ablated endocrine cells, nutrient stress and other parameters, providing information to elucidate the pathways involved in glucose and lipid homeostasis, as well as disease states such as obesity, diabetes and aging [see recent reviews by Gálíková and Klepsatel (45) and Marco and Gäde (18)].

In the malaria mosquito, *Anopheles gambiae* (Family: Culicidae), genomic information, gene cloning and physiological experimentation led to the identification of a hypertrehalosemic hormone as an octapeptide with the sequence pELFTPAW amide [= Anoga-HrTH, (27, 46, 47)]. Similar genomic and immunocytochemical methods identified an AKH with the sequence pELFTPSW amide (= Aedae-AKH) in the CC of the yellow fever mosquito, *Aedes aegypti* (Family: Culicidae), confirmed by mass spectrometric measurements (25, 26, 48). An AKH gene cloned from the West Nile virus vector, *Culex pipiens* = *C. quinquefasciatus* yielded an identical AKH, i.e., Aedae-AKH (25). Interestingly, Aedae-AKH as mature peptide was detected and sequenced by mass spectrometry in the alder fly *Sialis lutaria* which belongs not to Diptera but to the order Megaloptera (49). Kaufmann et al. (25) identified two AKH genes when mining publicly accessible genetic data from the tsetse fly *Glossina morsitans* (Diptera, Family: Glossinidae), the deduced mature peptides is the well-known Phote-HrTH and a novel octapeptide with the sequence pELTFSPGW amide (= Glomo-AKH); both of these genes were later cloned (36), and annotated during whole genome sequencing (50) of *G. morsitans*, while the mature peptide sequences were verified by mass spectrometry (37). Functionally, Phote-HrTH increased the amount of lipids released from fat body tissue of adult female *G. morsitans in vitro* (36); the function of Glomo-AKH was not physiologically investigated.

From the above, it is evident that there is some sort of group-specificity in the AKH family peptide sequences of the Diptera investigated to date, but there is also variation in the number of sequences per species and even in the length of the peptide sequence. Further elaboration of AKH sequences may also be useful in ascertaining the overlap of sequence identity in beneficial and pest dipteran species for the potential of finding a suitable lead for developing peptide mimetics that would target specifically pest dipterans. An additional aim of the current study

therefore is to provide a comprehensive list of putative AKH sequences from Diptera by mining publicly-accessible data bases.

MATERIALS AND METHODS

Insects

Adult specimens of unknown age and both sexes were used in this study. Fly species were either caught in the field by netting, were purchased from breeders or were received as a gift from a research institution. In total 19 species of Diptera were studied. Details of the fly species and the taxonomic affiliations are given below. For the latter, the phylogenetic outline given by Pape et al. (51) and Wiegmann et al. (8) were followed (see also section Introduction).

Suborder Lower Diptera (Formerly: Nematocera), Infraorder Tipulomorpha

Two species of crane fly (Family: Tipulidae) were investigated. Specimens of *Tipula paludosa* were caught on a meadow in Groß Raden in Mecklenburg-Western Pomerania (Mecklenburg-Vorpommern), Germany, by netting. Specimens of the other species (very likely *Tipula oleracea* based on morphological characters that we could ascertain) were caught on the wall of a hotel complex in northern Crete close to Heraklion.

Suborder Brachycera, Infraorder Stratiomyomorpha

One species of soldier fly (Family: Stratiomyidae) was investigated. Specimens of *Hermetia illucens* were purchased from Illucens GmbH, Ahaus, Germany.

Suborder Brachycera, Superfamily Syrphoidea

Six species of hover fly (Family: Syrphidae) were investigated. Specimens of *Volucella pellucens*, *Volucella zonaria*, *Chrysotoxum cautum*, and *Eristalis* ssp. were collected by netting from flowers of wild oregano (*Origanum vulgare*) in July and August 2017 and 2019 around Bad Iburg and Hagen, Germany. Specimens of *Eristalis tenax* and *Allograpta fuscotibialis* were netted on flowers in a private garden in Cape Town, South Africa.

Suborder Brachycera, Superfamily Tephritoidea

One species of fruit fly (Family: Tephritidae) was investigated. Specimens of the Mediterranean fruit fly (“medfly”) *Ceratitis capitata* were a gift from the Fruitfly Africa Medfly rearing facility in Stellenbosch, South Africa.

Suborder Brachycera, Superfamily Hippoboscoidea

Three species of tsetse fly (Family: Glossinidae) were investigated. Specimens of *Glossina fuscipes* and *G. austeni* came from the Veterinary Institute in Onderstepoort, South Africa, whereas *G. morsitans morsitans* were a gift from the International Atomic Energy Agency in Vienna, Austria.

Suborder Brachycera, Superfamily Muscoidea

Two species of the house fly (Family: Muscidae) and one species of kelp fly (Family: Anthomyiidae) were investigated.

TABLE 1 | The distribution of AKH peptides in the Diptera, to date: primary sequence and calculated protonated mass.

Higher taxonomy	Family	Species	AKH name	AKH Sequence*	[M+H] ⁺	References**	
Lower Diptera, Tipulomorpha	Tipulidae	<i>Tipula paludosa</i>	Glomo-AKH	pELTFSPGWa	917.4516	This study	
			Tippa-CC-I	pELTYSPSWa	963.4571		
			Tippa-CC-II	pELTFSPSWa	947.4621		
		<i>Tipula oleracea</i> <i>Tipula oleracea?***</i>	Tippa-CC-II	pELTFSPSWa	947.4621	JXPP01112122.1	
			Glomo-AKH	pELTFSPGWa	917.4516	This study	
			Tippa-CC-I	pELTYSPSWa	963.4571		
Lower Diptera, Psychodomorpha	Psychodidae	<i>Lutzomyia longipalpis</i> <i>Phlebotomus papatasi</i>	Tabat-AKH	pELTFTPGWa	931.4672	JH689332.1	
			Tabat-AKH	pELTFTPGWa	931.4672	A0A1B0DLM1	
Lower Diptera, Culicomorpha	Chironomidae	<i>Clunio marinus</i>	Aedae-AKH	pELTFTPSWa	961.4778	CRK93994.1	
Lower Diptera, Culicomorpha	Culicidae	<i>Culex quinquefasciatus</i> / <i>pipiens</i> <i>Aedes aegypti</i> <i>Aedes albopictus</i> <i>Anopheles gambiae</i> 15 Other Anopheles species ^a <i>Anopheles dirus</i> complex <i>Anopheles farauti</i> complex	Aedae-AKH	pELTFTPSWa	961.4778	(25); B0W1A8	
			Aedae-AKH	pELTFTPSWa	961.4778	(25, 26); AAEL011996-PA XP_029714753.1	
			Anoga-HrTH	pELTFTPAWa	945.4829	(27); A0NGF4	
			Anoga-HrTH	pELTFTPAWa	945.4829		
			NOVEL 1	pELTFTPTWa	975.4934	A0A182NZ26	
			NOVEL 1	pELTFTPTWa	975.4934	A0A182R1D8	
			NOVEL 1	pELTFTPTWa	975.4934	JXOU01007118.1	
			NOVEL 1	pELTFTPTWa	975.4934	JXPH01036237.1	
			NOVEL 2	pELTFTPVWa	973.5142	QYTT01077274.1	
Brachycera, Orthorapha, Tabanomorpha	Tabanidae	<i>Tabanus atratus</i>	Tabat-AKH	pELTFTPGWa	931.4672	(19); P14595.2 & P14596.2	
			Tabat-HoTH	pELTFTPGWYa	1151.5520		
Brachycera, Orthorapha, Stratiomyomorpha	Stratiomyidae	<i>Hermetia illucens</i> <i>Hermetia illucens</i>	Tabat-AKH	pELTFTPGWa	931.4672	This study (28); JXPW01153780.1	
				pELTFTGQWa	962.4730		
Brachycera, Orthorapha, Asiloidea	Asilidae	<i>Dasyopogon diadema</i>	NOVEL 2	pELTFTPVWa	973.5142	QYTT01077274.1	
Brachycera, Cyclorapha, - Syrphoidea	Syrphidae	<i>Volucella pellucens</i> <i>Volucella zonaria</i> <i>Chrysotoxum cautum</i> <i>Allograpta fuscotibialis</i> <i>Eristalis ssp mixture</i> <i>Eristalis tenax</i> <i>Eristalis ssp 1</i> (not <i>E.</i> <i>tenax</i>) <i>Eristalis ssp 2</i> (not <i>E.</i> <i>tenax</i>) <i>Eristalis dimidiata</i>	Glomo-AKH	pELTFSPGWa	917.4516	This study	
			Phote-HrTH	pELTFSPDWa	975.4571		
			Volpe-CC	pELTFSPYWa	1023.4934		
			Glomo-AKH	pELTFSPGWa	917.4516	This study	
				Phote-HrTH	pELTFSPDWa		975.4571
				Volpe-CC	pELTFSPYWa		1023.4934
			Glomo-AKH	pELTFSPGWa	917.4516	This study	
				Volpe-CC	pELTFSPYWa		1023.4934
				Volpe-CC	pELTFSPYWa		1023.4934
			Eriss-CC	pELTFSPYWa	1023.4934	This study	
				pELTFSPYWa	1023.4934		
				pELTFSPYWa	1023.4934		
			Eriss-CC	pELTFSPYWa	1023.4934	This study	
				pELTFSPYWa	1023.4934		
				pELTFSPYWa	1023.4934		
			Glomo-AKH	pELTFSPGWa	917.4516	This study	
				Phote-HrTH	pELTFSPDWa		975.4571
Volpe-CC	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					
Glomo-AKH	pELTFSPGWa	917.4516	This study				
	Volpe-CC	pELTFSPYWa		1023.4934			
	Volpe-CC	pELTFSPYWa		1023.4934			
Eriss-CC	pELTFSPYWa	1023.4934	This study				
	pELTFSPYWa	1023.4934					
	pELTFSPYWa	1023.4934					

(Continued)

TABLE 1 | Continued

Higher taxonomy	Family	Species	AKH name	AKH Sequence*	[M+H] ⁺	References**
Brachycera, Cyclorrapha, Schizophora	Sepsidae	<i>Themira minor</i>	Phote-HrTH	pELTFSPDWa	975.4571	JXPZ01029543.1
Brachycera, Cyclorrapha, Schizophora, Tephritoidea	Tephritidae	<i>Ceratitis capitata</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study; XP_004526727.1
		<i>Batrocera dorsalis</i>	Phote-HrTH	pELTFSPDWa	975.4571	XP_011211021.1
		<i>Batrocera oleae</i>	Phote-HrTH	pELTFSPDWa	975.4571	XP_014099190.1
		<i>Batrocera latifrons</i>	Phote-HrTH	pELTFSPDWa	975.4571	XP_018793135.1
		<i>Zeugodacus cucurbitae</i>	Phote-HrTH	pELTFSPDWa	975.4571	XP_011184633.1
		<i>Rhagoletis zephyra</i>	Phote-HrTH	pELTFSPDWa	975.4571	XP_017484104.1
		<i>Rhagoletis cerasi</i>	Phote-HrTH	pELTFSPDWa	975.4571	(29)
		<i>Sphyracephala brevicornis</i>	NOVEL 3	pELTFSPNWa	974.4730	JXPL01104182.1; (28)
Brachycera, Cyclorrapha, Schizophora	Diopsidae	<i>Teleopsis dalmanni</i>	NOVEL 3	pELTFSPNWa	974.4730	NLCU01016345.1; (28)
		<i>Liriomyza trifolii</i>	Phote-HrTH	pELTFSPDWa	975.4571	JXHJ01005265.1
Brachycera, Cyclorrapha, Schizophora, Ephydroidea	Ephydriidae	<i>Cirrula hians</i>	Phote-HrTH	pELTFSPDWa	975.4571	JXOS01046117.1
		Brachycera, Cyclorrapha, Schizophora Ephydroidea	Drosophilidae	<i>Drosophila melanogaster</i>	Phote-HrTH	pELTFSPDWa
17 Other <i>Drosophila</i> species ^b	Phote-HrTH			pELTFSPDWa	975.4571	(34, 35); and data bases
<i>Drosophila grimshawi</i>	NOVEL 3			pELTFSPNWa	974.4730	(34); XP_001983486.1
<i>Drosophila ficusphila</i>	NOVEL 4			pELTYSPDWa	991.4520	XP_017059889.1
<i>Scaptodrosophila lebanonensis</i>	Phote-HrTH			pELTFSPDWa	975.4571	XP_030380311.1
Brachycera, Cyclorrapha, Schizophora, Calyptrate	Glossinidae	<i>Glossina morsitans</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study; (25, 36, 37); AEH25942.1 & AEH25941.1
			Glomo-AKH	pELTFSPGWa	917.4516	
		<i>Glossina fuscipes</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study
			Glomo-AKH	pELTFSPGWa	917.4516	
		<i>Glossina austeni</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study
			Glomo-AKH	pELTFSPGWa	917.4516	
		Other <i>Glossina</i> species ^c	Phote-HrTH	pELTFSPDWa	975.4571	(38)
Glomo-AKH	pELTFSPGWa	917.4516				
Brachycera, Cyclorrapha, Schizophora, Calyptrate	Muscidae	<i>Musca domestica</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study; XP_005178897.1
		<i>Musca autumnalis</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study
		<i>Stomoxys calcitrans</i>	Phote-HrTH	pELTFSPDWa	975.4571	AOA118QC11
		<i>Haematobia irritans</i>	Phote-HrTH	pELTFSPDWa	975.4571	PGFW01000414.1
Brachycera, Cyclorrapha, Schizophora, Calyptrate	Anthomyiidae	<i>Fucellia capensis</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study
		<i>Delia radicum</i>	Phote-HrTH	pELTFSPDWa	975.4571	(39, 40); B3EWM7.1
Brachycera, Cyclorrapha, Schizophora, Calyptrate, Oestroidea	Sarcophagidae	<i>Sarcophaga (carnaria?)</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study
		<i>Sarcophaga (= Neobellieria) bullata</i>	Phote-HrTH	pELTFSPDWa	975.4571	(30, 41, 42); TMW51522.1
		<i>Sarcophaga crassipalpis</i>	Phote-HrTH	pELTFSPDWa	975.4571	(17)

(Continued)

TABLE 1 | Continued

Higher taxonomy	Family	Species	AKH name	AKH Sequence*	[M+H] ⁺	References**
Brachycera, Cyclorhapha, Schizophora, Calyptrate, Oestroidea	Rhinophoridae	<i>Paykullia maculate</i>	Phote-HrTH	pELTFSPDWa	975.4571	NDXZ01142068.1
Brachycera, Cyclorhapha, Schizophora, Calyptrate, Oestroidea	Calliphoridae	<i>Calliphora vicina</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study; JXOT01195109.1
		<i>Cochliomyia hominivorax</i>	Phote-HrTH	pELTFSPDWa	975.4571	PYHX01001830.1
		<i>Lucilia cuprina</i>	Phote-HrTH	pELTFSPDWa	975.4571	This study; (42) XP_023304778.1
		<i>Lucilia sericata</i>	Phote-HrTH	pELTFSPDWa	975.4571	JXPF01042013.1
		<i>Protophormia</i> (= <i>Phormia</i>) <i>terraenovae</i>	Phote-HrTH	pELTFSPDWa	975.4571	(21, 43)
		<i>Phormia regina</i>	Phote-HrTH	pELTFSPDWa	975.4571	(44); MINK01172052.1

*Mature peptide sequence: in the case of sequences derived from nucleotide databases, the post-translational modifications (i.e., blocked termini) are deduced from the characteristic features of previously-sequenced AKHs. Note that the N-terminal pyroglutamic acid (pE) can arise from the cyclization of a glutamine (Q) or, more rarely, from a glutamate (E) residue. In the case of prepro-AKH sequences translated from mRNA, all have a glutamine amino acid residue at position 1 of the uncleaved AKH sequence. Novel peptide sequences deduced from data mining are consecutively numbered (1–4)—note that we have not confirmed the data experimentally.

**Published work or database accession number.

^aOther *Anopheles* species (accession no.): *A. coluzzii* (A0A182LR92); *A. stephensi* (A0A182Y983); *A. albimanus* (A0A182FZY6); *A. arabiensis* (A0A182IIN3); *A. atroparvus* (A0A182J042); *A. christyi* (A0A182K374); *A. culicifacies* (A0A182MX76); *A. darlingi* (A0A158N7M0-1); *A. epiroticus* (A0A182PMF71); *A. funestus* (A0A182S506); *A. melas* (A0A182TPX4); *A. merus* (A0A182VPI3); *A. minimus* (A0A182WR25); *A. quadriannulatus* (A0A182XUQ9); *A. sinensis* (A0A084VYF4-1).

^bOther *Drosophila* species (accession no.): *D. elegans* (XP_017127434.1); *D. virilis* (XP_002046309.1); *D. novamexicana* (XP_030561075.1); *D. serrata* (Sequence ID: XP_020807712.1); *D. pseudoobscura pseudoobscura* (XP_001353031.1); *D. guanche* (SPP77431.1); *D. obscura* (XP_022220454.1); *D. kikawai* (XP_017019402.1); *D. willistoni* (XP_002068218.1); *D. ananassae* (XP_001957722.1); *D. arizonae* (XP_017860946.1); *D. busckii* (XP_017842867.1); *D. mojavensis* (XP_002012181.1); *D. hydei* (XP_023165811.1); *D. bipunctata* (XP_017101344.1); *D. takahashii* (XP_017009635.1); *D. rhopaloa* (XP_016978805.1); *D. suzukii* (XP_016943637.1); *D. yakuba* (XP_002093648.1); *D. eugracilis* (XP_017071536.1); *D. biarmipes* (XP_016966107.1); *D. sechellia* (XP_002035290.1); *D. erecta* (XP_001971844.1); *D. simulans* (XP_002083583.1).

^cOther *Glossina* species (gene identification number): *G. pallidipes* (GPAI036121; GPAI049064); *G. austeni* (GAUT013261; GAUT013267); *G. palpalis* (GPPI030617; GPPI030614); *G. fuscipes* (GFUI054167; GFUI054166); *G. brevipalpis* (GBRI027509; GBRI045557).

***A tipulid species that is not *T. paludosa*, and most likely *T. oleracea* (although not confirmed by a dedicated taxonomist) was collected in Crete, and the AKH complement was elucidated via MS.

The phylogenetic outline is as per Wiegmann et al. (3).

Specimens of *Musca domestica* were a gift from Clintech (Pty), Ltd., Bryanston, South Africa; *Musca autumnalis* was caught in the wild in August 2019 around Bad Iburg, Germany, and the kelp fly *Fucellia capensis* was caught by netting on rotting kelp on the beach of Muizenberg, South Africa, in August 2019.

Suborder Brachycera, Superfamily Oestroidea

Two species of blowfly (Family: Calliphoridae) and one of flesh fly (Family: Sarcophagidae) were investigated. Specimens of *Calliphora vicina* and *Sarcophaga* ssp. (very likely *caritaria*) were collected during July and August 2017 around Osnabrück, Germany, and *Lucilia cuprina* was a gift of Clintech (Pty), Ltd., Bryanston, South Africa.

Species of the Order Mecoptera

Adults of both sexes of the common scorpion fly, *Panorpa communis*, were collected by netting during July 2017 around Bad Iburg, Germany.

Biological Assays

Conspecific metabolic bioassays were performed with adult specimens of both sexes of the soldier fly *Hermetia illucens*, of which a larger number of individuals was available. Flies were acclimatized for about 1 h before experimentation at

ambient temperature ($22 \pm 1^\circ\text{C}$), each in a 15 ml plastic container with water-soaked cotton wool pieces. After this resting period, 0.5 μl of hemolymph was sampled dorsally from the thorax with a disposable glass microcapillary (Hirschmann Laborgeräte, Eberstadt, Germany), and the hemolymph was blown into concentrated sulfuric acid to measure vanillin-positive material (= total lipids) or anthrone-positive material (= total carbohydrates) according to the phosphovanillin method (52) and anthrone method (53), respectively, as modified by Holwerda et al. (54). Subsequently, the flies were injected ventrolaterally into the abdomen with 3 μl of either water (a control for assessing stress effects of handling), a crude corpora cardiaca extract, or a synthetic peptide using a Hamilton fine-bore 10 μl syringe. A second hemolymph sample was taken 90 min post-injection from the same individuals under resting conditions.

Dissection of Corpora Cardiaca, Peptide Extraction, Mass Spectrometry, and Sequence Analysis

Corpora cardiaca (CCs) were microdissected from the head/neck region of adult flies from each species under investigation using a stereomicroscope at about 20- to 30-fold magnification; CCs were placed into 80% v/v methanol. Peptide material were extracted

TABLE 2 | Biological activity of a crude methanolic extract of corpora cardiaca from the black soldier fly (*Hermetia illucens*), and the synthetic peptide Tabat-AKH in homologous bioassays.

Treatment	Hemolymph carbohydrates (mg ml ⁻¹) ^a					Hemolymph lipids (mg ml ⁻¹) ^a				
	n	0 min	90 min	Difference	p ^b	n	0 min	90 min	Difference	p ^b
Control (3 µl distilled water)	18	12.9 ± 4.7	13.0 ± 4.9	0.2 ± 1.7	NS	11	14.3 ± 6.7	13.4 ± 7.4	-0.9 ± 2.2	NS
<i>H. illucens</i> CC (0.5 gland pair equiv.)	6	11.5 ± 3.5	15.7 ± 5.0	4.2 ± 2.0	0.002			Not determined		
Tabat-AKH (10 pmol)	23	13.1 ± 4.6	17.4 ± 6.5	4.3 ± 3.8	0.00001	11	14.0 ± 7.7	13.6 ± 7.5	-0.5 ± 1.6	NS

^aData are presented as Mean ± SD.

^bPaired t-test was used to calculate the significance between pre- and post-injection. NS, not significant.

from the dissected CCs as outlined previously (55). The vacuum-centrifuged dried extracts were dissolved in 50 µl of aqueous formic acid for liquid chromatography tandem positive ion electrospray mass spectrometry (LC-MSⁿ) on an LTQ XL linear ion trap instrument (Thermo Fisher Scientific, San Jose, CA), as described in detail previously (56).

For conspecific biological assays with the black soldier fly (see section Biological Assays), the dried CCs extracted from *H. illucens* were reconstituted in distilled water for injection into the flies.

Synthetic Peptides

The following peptides were previously synthesized by Peninsula Laboratories (Belmont, CA, USA), Pepmic Co. Ltd. (Suzhou, China), or custom-synthesized by Dr. Kevin Clark (University of Georgia, Athens, USA): Pote-HrTH (= Drome-HrTH; pELTFSPDW amide), Glomo-AKH (pELTFSPGW amide), Tabat-AKH (pELTFTPGW amide), Tabat-HoTH (pELTFTPGWGY amide), Anoga-HrTH (pELTFTPAW amide), and Aedae-AKH (pELTFPSW amide). The novel peptides of this study, Eriss-CC (pELTFPSAGW amide), Volpe-CC (pELTFSPYW amide), Tippa-CC-I (pELTYSPSW amide) and Tippa-CC-II (pELTFSPSW amide) were all custom-synthesized by Pepmic Co. Ltd.

Mining of AKH Sequences From Publicly Available Data Bases

The primary sequence of AKH family peptides in dipteran species were investigated by mass spectrometry (see Section Dissection of Corpora Cardiaca, Peptide Extraction, Mass Spectrometry, and Sequence Analysis). To compare and analyse Dipteran AKHs in a phylogenetic manner, we identified further AKH sequences from other dipteran species via literature searches (i.e., from previously published texts), as well as via bioinformatics. Such, *in silico* searches of Dipteran protein, genomic and/or EST databases were conducted to identify translated amino acid sequences and transcripts encoding putative AKH peptide precursors.

Some AKH sequences were retrieved from VectorBase (<https://www.vectorbase.org/>), which is a Bioinformatics Resource for invertebrate vectors of human pathogens. The searches for AKH genes were conducted using the “search term” function, or via the BLAST (Basic Local Alignment Search Tool) search function whereby the nucleotide sequence

of the *Glossina morsitans* AKH was inserted. From the search results, the AKH peptide sequence contained within the deduced prepro-hormones were predicted from homology to known insect AKH isoforms. The associated UniProt accession numbers (<https://www.uniprot.org/uniprot>) were recorded. The UniProt Knowledgebase (UniProtKB) is the publicly available central hub for the collection of functional information on proteins.

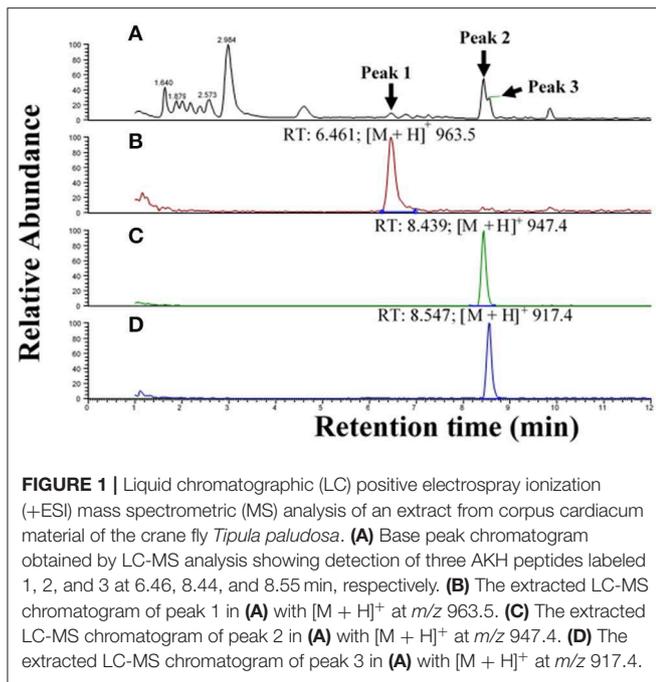
The remainder of putative AKH sequences were obtained via homology searches using BLAST from the National Center for Biotechnology Information site (<https://blast.ncbi.nlm.nih.gov/>); AKH peptide precursors from *G. morsitans* were used as BLAST query. For all searches resulting in sequence identifications, the BLAST score and BLAST-generated E-value for significant alignment were considered.

RESULTS AND DISCUSSION

Function of Dipteran AKH: The Black Soldier Fly as a Test Case

We conducted a conspecific biological assay with only one dipteran species to show proof of principle that the CC extracts of dipteran species do have hypertrehalosemic biological activity. Unambiguous results reveal that conspecific CC extracts of the black soldier fly *H. illucens* injected into resting adults had a small but statistically significant hypertrehalosemic effect (Table 2). Such a hypertrehalosemic effect was also shown when the endogenous AKH octapeptide of the black soldier fly, Tabat-AKH, was injected, whereas the lipid concentration was not affected (Table 2). These results confirm previous data on the action of AKH peptides in a blowfly (22) and the malaria mosquito (47), viz. the mobilization of carbohydrates. When working with a crude CC extract, one could argue that the extract contains, of course, all other putative neuropeptide hormones. Although we are not aware of any neuropeptide that is in higher concentration in the CC and has a clear effect of increasing carbohydrates or lipids except the AKHs, we can only assume that the measured effect was induced by an AKH.

However, it is clear from studies on other Diptera that these peptides can also mobilize lipids in certain fly species (20, 36). Thus, as reported for other insect orders, it is the metabolic machinery of the species under investigation and the metabolic needs of an insect that



determines activation of a lipase and/or a phosphorylase by AKHs (18, 57).

Mass Spectral Analyses of AKHs Reveal Octapeptides in Major Clades of Diptera

Methanolic CC extracts from all 19 fly species were analyzed by mass spectrometry. Here we will show a few exemplary cases; we chose species from the infraorders Tipulomorpha and Stratiomyomorpha, as well as from the superfamilies Syrphoidea and Tephritoidea.

Tipulomorpha *Tipula paludosa*

The CC extract of the crane fly *T. paludosa* was fractionated (separated) by reversed-phase liquid chromatography (LC) and the peptides detected by positive ion electrospray mass spectrometry (MS). **Figure 1A** shows the base peak chromatogram; **Figures 1B–D** depict extracted mass peaks of AKHs at 6.46, 8.44, and 8.55 min, respectively, with the corresponding $[M + H]^+$ mass ions at m/z 963.5, 947.4, and 917.4, respectively. The primary structures of these peak materials were deduced from the tandem MS² spectra obtained by collision-induced dissociation (CID) of the respective m/z ions. The spectrum of m/z 963.4 (**Figure 2**) with clearly defined b , y , $b-H_2O$, $y-NH_3$, and other product ions allowed an almost complete assignment of a typical octapeptide member of the AKH family under the assumption that the peptide has a characteristic pyroglutamate residue at the N-terminus (see schematic inset in **Figure 2**). All other amino acids are assigned except at position two where the remaining mass of 113 can be accredited to one of the isomers leucine or isoleucine. Such a peptide with the sequence pGlu-Ile/Leu-Thr-Tyr-Ser-Pro-Ser-Trp amide had never been found in any insect, to date. It was hence code-named Tippa-CC-I (*Tipula paludosa* CC peptide I)

and awaited clarification of the amino acid residue in position 2 through co-elution experiments with a synthetically made Leu² analog of the peptide (see below). The spectrum of m/z 947.4 (**Figure 3**) led to the interpretation of another octapeptide with the sequence pGlu-Ile/Leu-Thr-Phe-Ser-Pro-Ser-Trp amide which is also novel and, thus, called Tippa-CC-II with the same ambiguity about position 2 in the primary sequence. The spectrum of the third peptide at m/z 917.4 (**Figure 4**) resulted in the sequence interpretation pGlu-Ile/Leu-Thr-Phe-Ser-Pro-Gly-Trp amide which, with Leu², is well-known under the name Glomo-AKH as one of the two peptides found in the tsetse fly (see section Introduction; **Table 1**). All three peptides were synthesized as Leu² isomer and co-elution experiments were performed; previously we had established that the isobaric Leu/Ile peptides have different LC retention times (58, 59). As depicted in **Figures S1A–I** all three synthetic peptides had identical retention times to the natural peptides in the CC extract and have, therefore, the correctly assigned primary peptide sequence, confirming that leucine is the second amino acid residue also of the novel Tippa-CC-I and II. The same three masses were also identified from the tipulid species caught on Crete and support the finding of these 3 octapeptides in this infraorder.

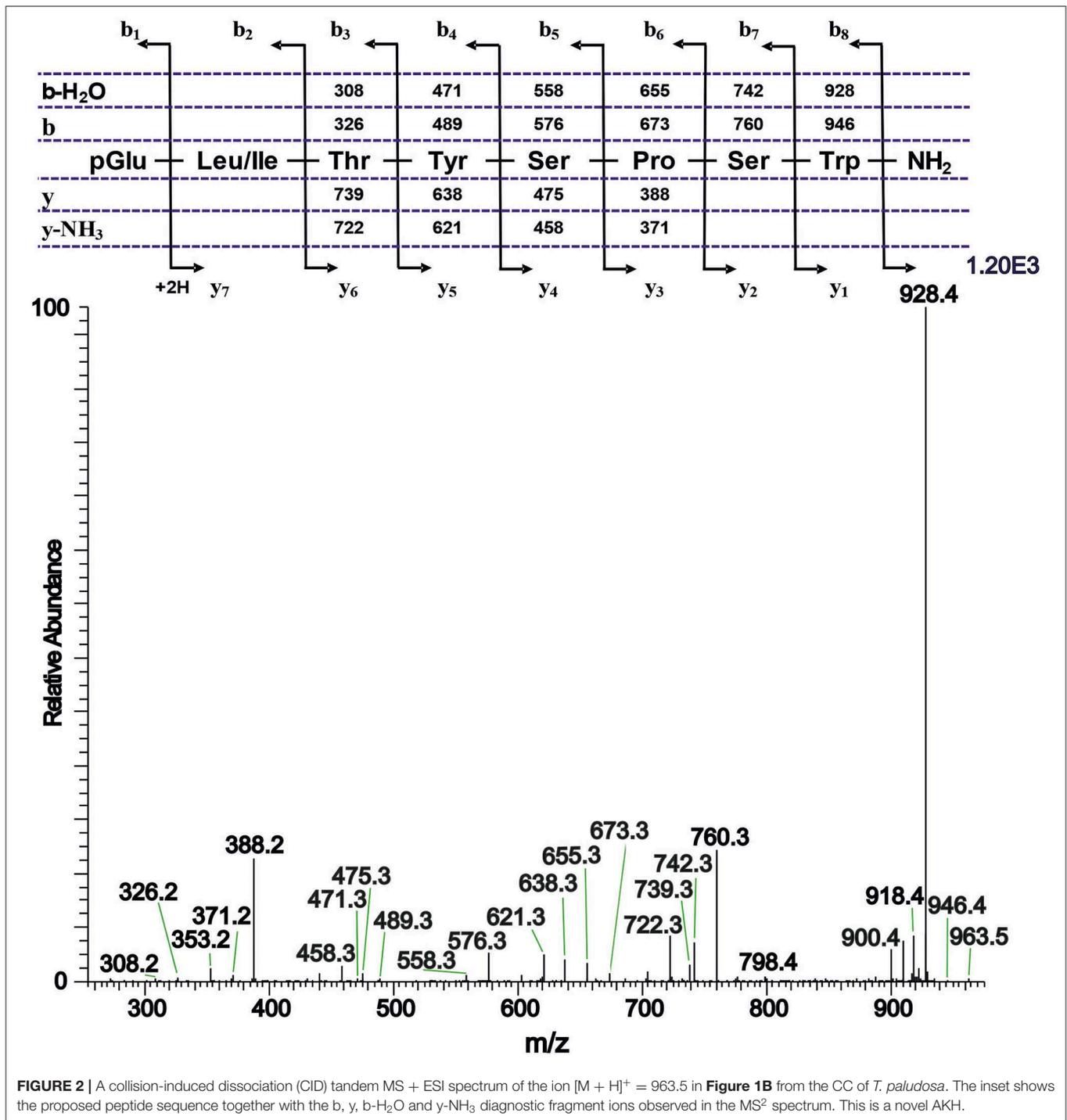
When we look at the primary structure of the three AKHs of *T. paludosa* and the other tipula species, it is obvious that they are closely related to each other and to the AKH known as Phote-HrTH from *Phormia* and *Drosophila* (**Table 1**).

Stratiomorpha *Hermetia illucens*

An extract from the CC of the black soldier fly shows a peak with a retention time of 8.88 min (**Figure S2A**) which corresponds in MS analysis to an $[M + H]^+$ ion of m/z 931.4 (**Figure S2B**). The CID spectrum gave clear product ions and led to the interpretation of an AKH with the primary structure pGlu-Ile/Leu-Thr-Phe-Thr-Pro-Gly-Trp-amide (**Figure S2C**) of which the Leu² form was established by a co-elution experiment (**Figures S2D–F**). This peptide is known as Tabat-AKH and found in *Tabanus atratus* (19). Surprisingly, genetic work has proposed another peptide for this species, i.e., pGlu-Leu-Thr-Phe-Thr-Gly-Gln-Trp-amide, which differs at position 6 (Gly instead of Pro) and 7 (Gln instead of Gly), and in its calculated protonated mass of 962.4730 [see **Table 1**; (28)]. Although there were many other prominent peak material eluting from the CC extract in the current study (**Figure S2A**), these peaks did not correspond with the calculated mass of 962.47, nor did they relate to an AKH peptide sequence. Since there has never been a glycine residue found at position 6 in any of the 90 AKHs known so far (G. Gäde personal communication), we suggest that there may be a misread in the genetic code and, hence, do not currently accept this sequence as a novel one until it has been confirmed (or refuted) by mass spectrometric methods.

Syrphoidea: *Eristalis* ssp

An extract from the CC of a number of *Eristalis* specimens, possibly a mixture of a few species that are difficult to distinguish



taxonomically, gave information of four clearly identified AKHs. The extracted chromatograms reveal three AKHs with near-identical retention times but with different $[M + H]^+$ ions of m/z 891.4, 917.4, and 975.5, the fourth AKH with $[M + H]^+$ ion of m/z 1023.5 had a longer retention time (**Figures S3A–E**). The CID spectra gave clear evidence for the sequence of the respective AKHs. Here we show only the CIDs for $[M + H]^+$ 891.4 and 1023.5 because they represent novel AKHs. The peptide with $[M$

$+ H]^+$ 917.4 (= Glomo-AKH) has been dealt with in the earlier example (see Tipulomorpha: *Tipula paludosa*) while $[M + H]^+$ 975.5 (= Phote-HrTH) will be discussed later (see Tephritoidea: *Ceratitis capitata*). The primary structure of an AKH representing $[M + H]^+$ 891.4 was deduced as pGlu-Leu/Ile-Thr-Phe-Ser-Ala-Gly-Trp amide (**Figure 5A**) and the one with $[M + H]^+$ 1023.5 was assigned the sequence pGlu-Leu/Ile-The-Phe-Ser-Pro-Tyr-Trp amide (**Figure 5B**). The ambiguity at position two was solved

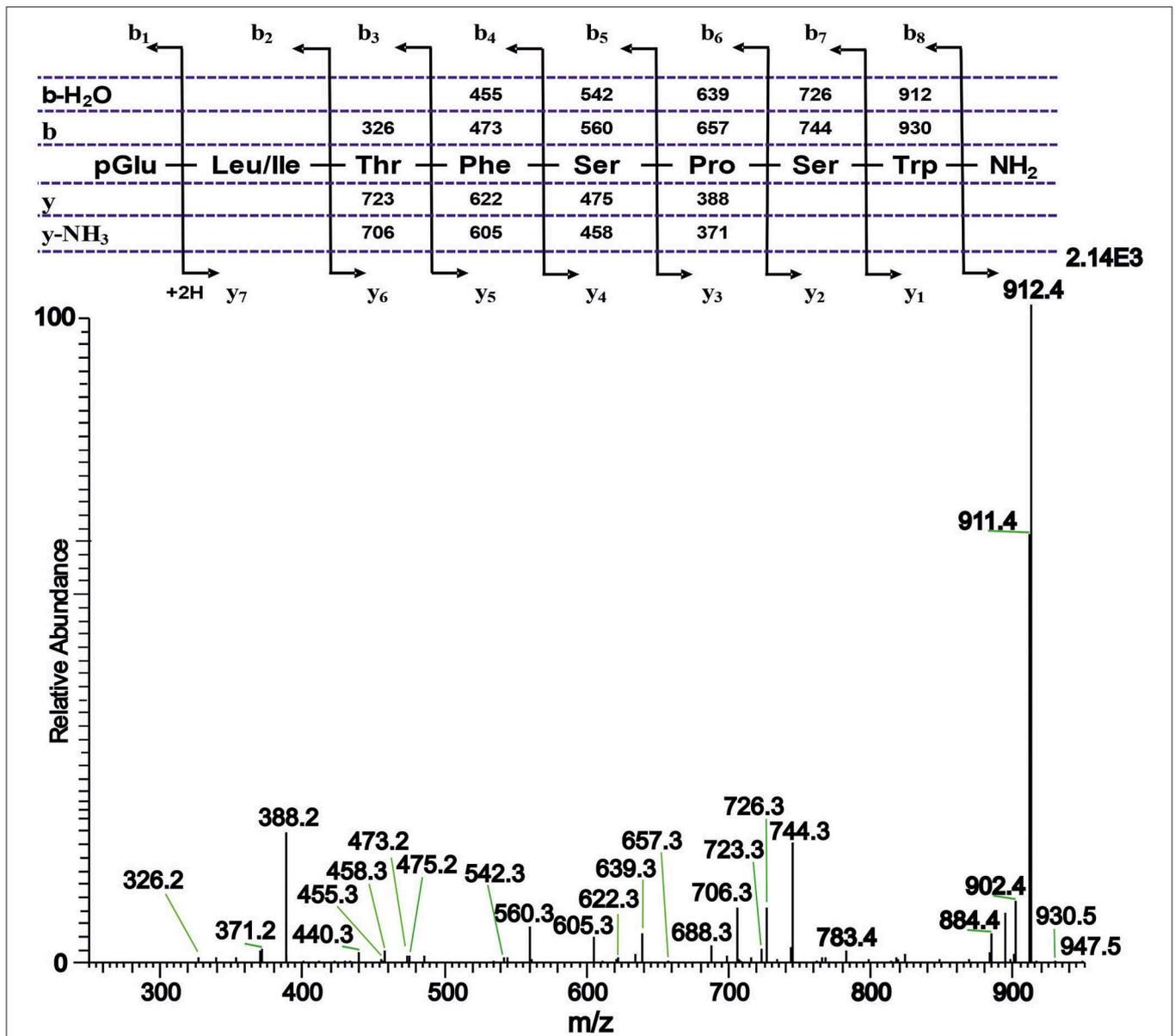


FIGURE 3 | A collision-induced dissociation (CID) tandem MS + ESI spectrum of the ion $[M + H]^+ = 947.4$ in **Figure 1C** from the CC of *T. paludosa*. The inset shows the proposed peptide sequence together with the b, y, b-H₂O and y-NH₃ diagnostic fragment ions observed in the MS² spectrum. This is a novel AKH.

to the presence of Leu in both cases by co-elution experiments with the synthetic peptide (**Figures S3F–Q**). Both peptides are novel, thus have not been found in any insect before and we assign them the code-name Eriss-CC for the peptide with MH⁺ 891.4 because it occurs in an *Eristalis* subspecies (see **Table 1**: *Eristalis* subspecies 2), and Volpe-CC for the peptide with MH⁺ 1023.5 as this peptide was discovered in the unambiguously defined species *Volucella pellucens* and all other syrphid species in the current study (**Table 1**).

Tephritoidea *Ceratitis capitata*

An extract from the CC of the fruit fly shows only one peak with a retention time of 8.81 min (**Figure S4A**) which corresponds in

MS analysis to an $[M + H]^+$ ion of m/z 975.4 (**Figure S4B**); the CID (**Figure S4C**) determines a sequence of pGlu-Leu/Ile-Thr-Phe-Ser-Pro-Asp-Trp amide. Co-elution with synthetic peptide (**Figures S4D–F**) unequivocally determines Leu at position 2 and, hence, characterizes this peptide as Phote-HrTH found previously in certain flies [(21, 23); see **Table 1**].

AKHs of Diptera Are Order-Specific

The CC extracts of all the other Diptera investigated in the current study were analyzed in the same way as in the above examples. The results are given in **Table 1**. In this table we present the taxonomic units of Diptera in the phylogenetic relationship as outlined by Wiegmann et al. (8).

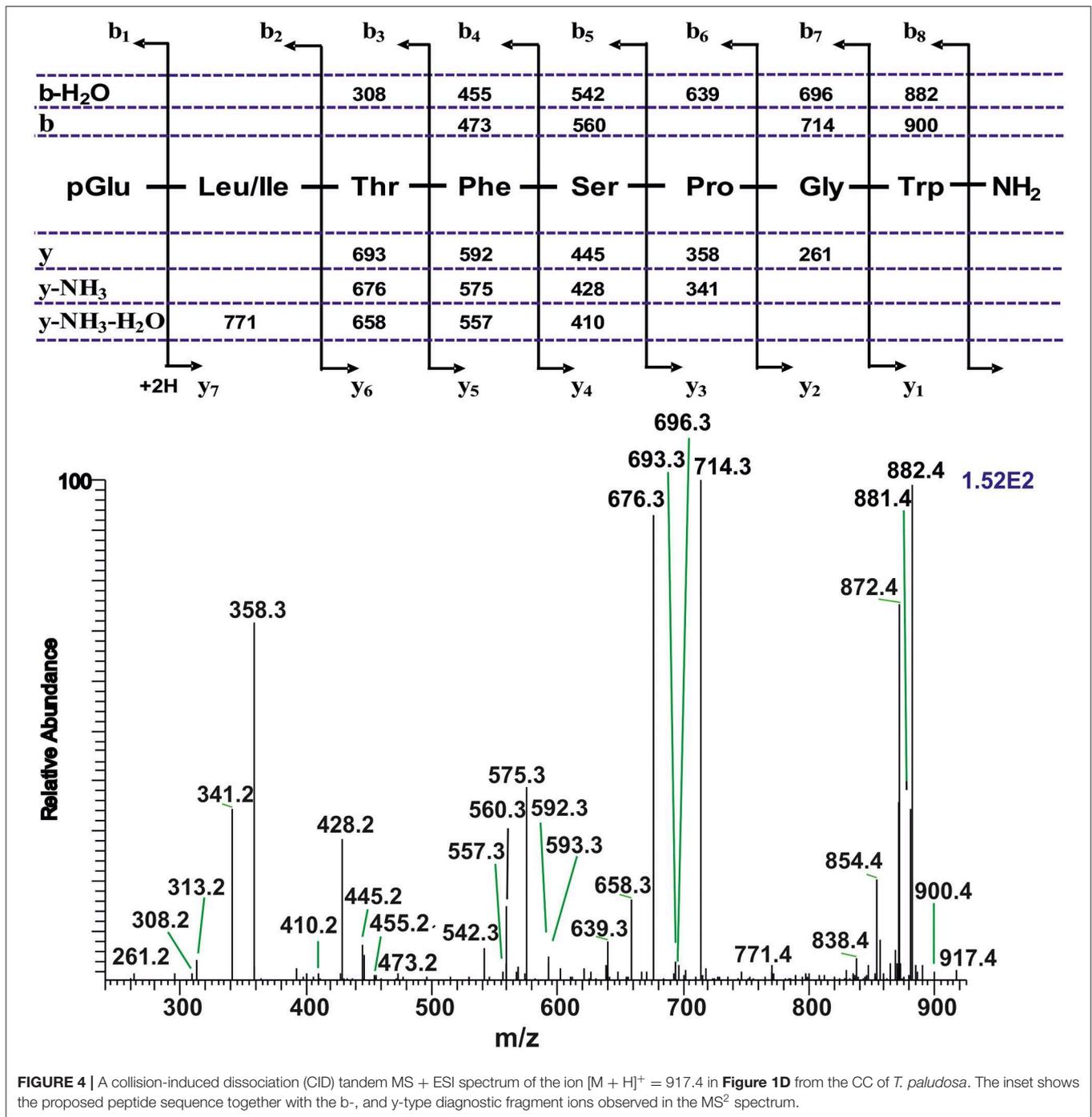


FIGURE 4 | A collision-induced dissociation (CID) tandem MS + ESI spectrum of the ion $[M + H]^+ = 917.4$ in **Figure 1D** from the CC of *T. paludosa*. The inset shows the proposed peptide sequence together with the b-, and y-type diagnostic fragment ions observed in the MS^2 spectrum.

Additionally, previously published data on AKH sequences are incorporated in this table as are sequences that were “mined” from genomic or transcriptomic data sets as outlined in Materials and Methods.

It is apparent from the combined data sets in **Table 1** that there are a few main points to be made about the AKHs in Diptera which will be discussed in more detail hereafter: (i) Dipteran AKHs are octapeptides. There is only one decapeptide exception. (ii) Current mass spectral investigations find 4 novel octapeptides

from fly corpora cardiaca; an additional 4 novel octapeptides are found through data mining. (iii) Dipteran AKHs are specific for the order. There are only two exceptions.

To date, we know roughly 90 sequences of AKHs from insects. There is a clear bias toward the production of octapeptides as only one third of the known AKHs are decapeptides and only three are nonapeptides. Most of the decapeptide AKHs are found in the orders Hymenoptera, Hemiptera and Caelifera, while the nonapeptides are in Lepidoptera and Hemiptera. The

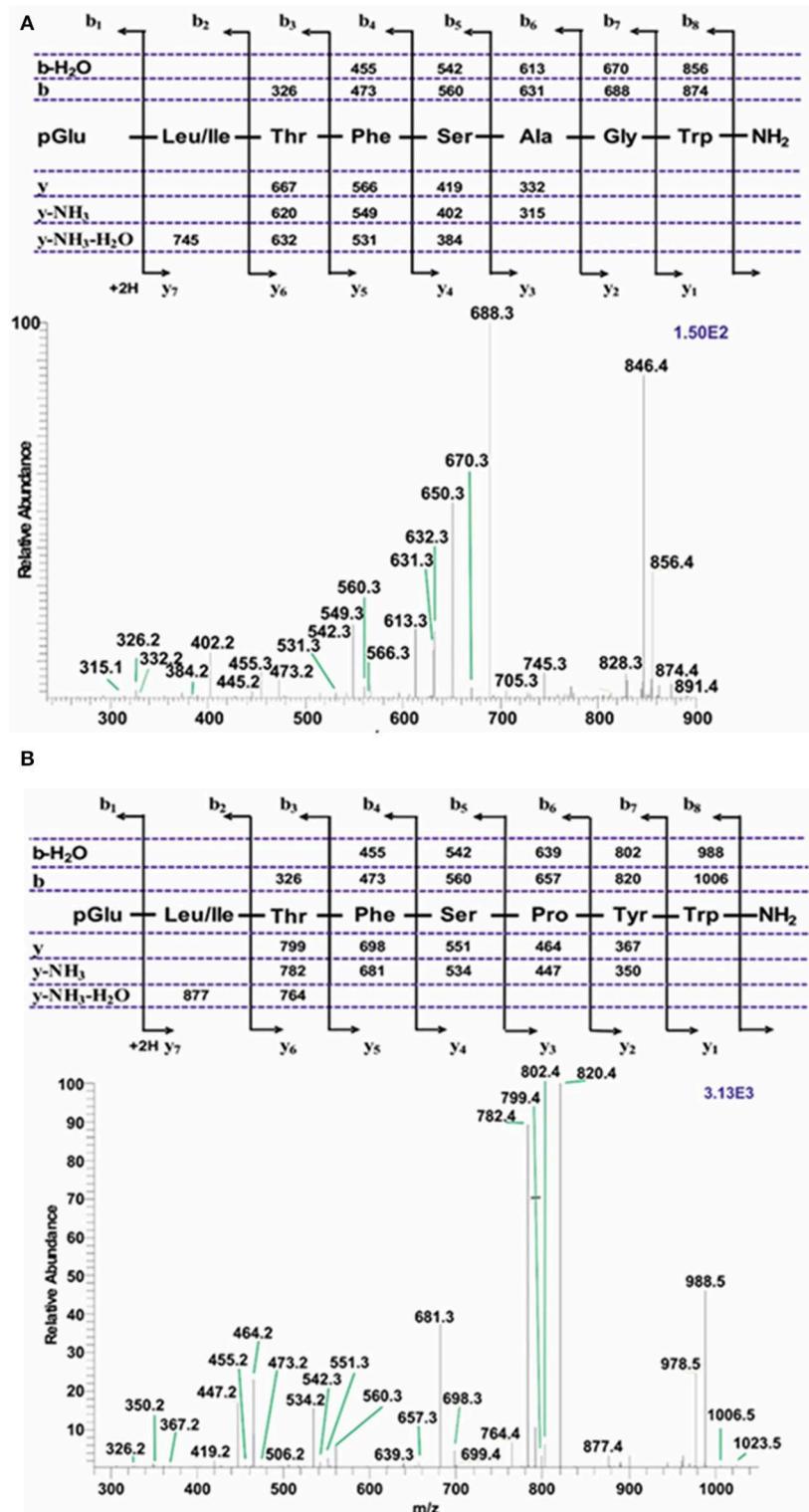


FIGURE 5 | Liquid chromatographic (LC) positive electrospray ionization (+ESI) mass spectrometry (MS) analysis of an extract from corpus cardiacum material of a mixture of hoverfly species from the genus *Eristalis*. **(A)** A collision-induced dissociation (CID) tandem MS + ESI spectrum of the ion $[M + H]^+ = 891.4$. The inset shows the proposed peptide sequence together with the b- and y-type diagnostic fragment ions observed in the MS² spectrum. This is a novel AKH. **(B)** A collision-induced dissociation (CID) tandem MS + ESI spectrum of the ion $[M + H]^+ = 1023.5$. The inset shows the proposed peptide sequence together with the b- and y-type diagnostic fragment ions observed in the MS² spectrum. This is a novel AKH.

order under scrutiny here, the Diptera, contain up to now only a single AKH decapeptide, viz. Tabat-HoTH in the horse flies. The presence of Tabat-HoTH can be well-explained by gene duplication since this decapeptide has the same amino acid sequence as the accompanying octapeptide in the horse fly but is elongated by two amino acids at the C- terminus. It remains to be seen whether another genus of the tabanids also has such a complement of AKHs. Unfortunately, our efforts with an extract prepared from long-frozen CCs of the common horse fly, *Haematopota fluvialis*, did not yield any mass spectral data for inclusion in the current work. Although gene duplication has taken place in other dipteran species, such as in crane flies and hover flies as shown in the current work, this has not led to decapeptides but mutations to other octapeptides.

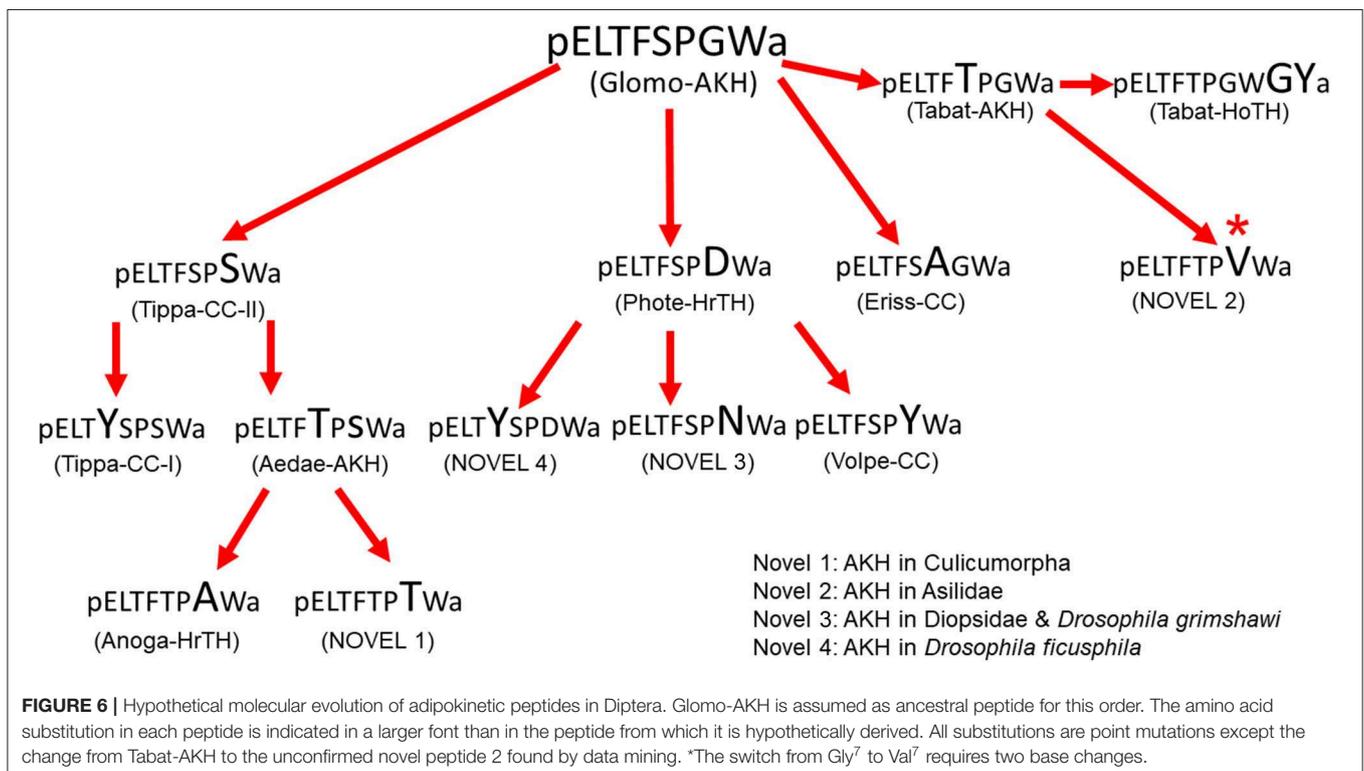
Before we started this study, 6 AKHs were known from the order Diptera (see history in section **Introduction**). The current study adds 8 further sequences: 4 novel AKH peptides (2 in the tipulids and 2 in the syrphids) found by unequivocal mass spectrometric studies, and 4 novel sequences (distributed in Culicomorpha, Stratiomyomorpha, Asiloidea, Diopsidae, and Drosophilidae) from data mining of public data bases—although it should be cautioned that the “mined” sequences have not been substantiated by mass spectrometry or peptide chemistry. Three of the four characterized novel AKHs and 1 of the mined AKHs share the same N-terminal pentapeptide pELTFS, while the fourth novel AKH share pELTYS with one of the mined AKHs (**Table 1**). The remaining 2 of the 4 novel peptides found by data mining have the pentapeptide sequence pELTFT (**Table 1**). A scan of the dipteran AKH sequences in **Table 1** shows that

the most prevalent amino acid residue in position 6 is proline, and there is a clear distribution pattern of AKHs with pELTFSP in lower diptera Tipulomorpha, while pELTFTP is present from the lower diptera Psychodomorpha until its last appearance in orthorrhaphan Brachycera Asiloidea, after which AKHs with pELTFSP reappear in Brachycera Syrphoidea onwards).

In contrast to almost all other insect orders it appears that AKHs in Diptera are relatively order-specific. Of the 14 AKHs now known from Diptera, only two are found outside of this order, viz. Aedae-AKH which is found in the only Megaloptera species investigated to date (49), and a new discovery from the current work: Glomo-AKH is found in a species from Mecoptera (see Putative Molecular Evolution of Dipteran AKHs). This is quite remarkable when considering the diversity and species-richness of Diptera. Such order-specificity is at the moment also known from Lepidoptera (60) but not from Coleoptera (59, 61) and Orthoptera (13, 62). We cannot, at this point, exclude the possibility that future investigations may reveal AKH structures that are common to Diptera and various other orders.

Putative Molecular Evolution of Dipteran AKHs

If we want to speculate about the molecular evolution of dipteran AKHs we first have to find out what the putative ancestor AKH may be. For this it would be advantageous to know the AKHs of the closest relatives to the Diptera. According to Wiegmann et al. (63) the orders Mecoptera (scorpion flies) and Siphonaptera (fleas) are evolutionary closest and not the parasitic Strepsiptera as once postulated. Data mining revealed genomic information



for a member of the Siphonaptera proposing a mature AKH of the cat flea, *Ctenocephalides felis*, with the sequence pELTFTPVW amide (XP_026477356.1), which is also a predicted AKH for the robber fly *Dasygogon diadema* (QYTT01077274.1; see **Table 1**). We did not find any data base information on Mecoptera and thus decided to include the common scorpion fly, *Panorpa communis*, into our study. The mass spectrometric data from the CC of this insect are clear: the AKH has the ionized mass (MH⁺) of 917.5 and the sequence of pELTFSPGW amide, thus Glomo-AKH (results not shown). Thus, both related orders have AKHs with sequences that also occur in Diptera, viz. Glomo-AKH in members of the Tipulidae, Syrphidae, and Glossinidae and the novel peptide (novel 2) in a member of the Asilidae (see **Table 1**). This may point to a common ancestor of the three orders with the sequence pELTFXPXW amide. Unfortunately, we do not have data on AKHs from the most basal fly families, that are species-poor and not very accessible, i.e., the Deuterophlebiidae and Nymphomyiidae (8). Currently, the AKHs known from Tipulidae are the ones from the most basal dipteran family. Since Glomo-AKH does also occur in the close relative order Mecoptera and in a number of dipteran families, we currently assume this molecule as parent to speculate and sketch a molecular evolution of the AKHs occurring in the various families of Diptera, as inferred by single point mutations. **Figure 6** clearly shows that this hypothetical peptide evolution could have occurred without introducing any new sequences to the ones listed in **Table 1**. Additionally, only in one case indicated in the figure would two nucleotides of the triplet have to mutate. Although speculation, this scheme of putative evolution has all the known AKHs of the lower Diptera (Tipulomorpha and Culicomorpha) on one side (left) of the flow chart giving it some credibility.

As previously shown, e.g., for Odonata, sequences of mature AKHs do not allow an insight into deeper phylogenetic relationships but give some indications of relatedness (10, 64). For example, the three peptides of the Tipulidae are very similar, as are the four of the Syrphidae and also the three in Drosophilidae (see **Table 1** and **Figure 6**). It is also clear that the vast majority of the cycloraphan flies, which are monophyletic (8), only produce Phote-HrTH. It appears that quite a radiation of AKHs occurred early in the evolution of the Diptera, thus in the Tipulomorpha and Culicomorpha which are thought to have radiated in the Triassic (8). The next wave of radiation according to Wiegmann et al. (8) occurred in the early Jurassic by the lower Brachycera including the clades Tabanomorpha, Stratiomyomorpha and Asiloidea all of which

have their AKHs clustered together at the right side of **Figure 6**. The latest radiation involves the clade Schizophora between earliest Paleocene to Tertiary and includes the species in the Tephritoidea, Ephydroidea and Oestroidea which all produce Phote-HrTH and its closest relatives (see middle of **Figure 6**).

DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/**Supplementary Material**. Any other requests can be directed to Gerd Gäde, gerd.gade@uct.ac.za for metabolic data and Petr Šimek, simek@bclab.eu for mass spectrometric data.

AUTHOR CONTRIBUTIONS

GG: concept and design of the study, acquisition of insect species and synthetic peptides, interpretation of the data and writing the draft manuscript. HM: co-designed the study, data acquisition (biological assays, dissection of insect corpora cardiaca and preparation of extracts, mining data bases for AKH sequences), interpretation and analyses of data, writing and refining the draft manuscript. PS: mass spectrometric analyses, data interpretation, and drafting of MS Figures for the manuscript.

FUNDING

This work is based on the research supported in part by the National Research Foundation of South Africa: grant numbers 85768 [IFR13020116790] to GG and 109204 [IFR170221223270] to HM, and the University of Cape Town (Block Grants to GG and HM). PS was supported by the Czech Science Foundation No. 17-22276S.

ACKNOWLEDGMENTS

The authors thank Ms. Pavla Kružberská for excellent technical support with MS measurements, all suppliers of insects mentioned in the main text, and colleagues for identification of certain species (Wolfgang Rutkies and Frank Wolf).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2020.00153/full#supplementary-material>

REFERENCES

1. Khamesipour F, Lankarani KB, Honarvar B, Kwenti TE. A systematic review of human pathogens carried by the housefly (*Musca domestica* L.). *BMC Public Health*. (2018) 18:1049. doi: 10.1186/s12889-018-5934-3
2. Cheng TC. *General Parasitology*. České Budějovice: Elsevier Science. (2012). p. 660.
3. Ssymank A, Kearns CA, Pape T, Thompson FC. Pollinating flies (Diptera): a major contribution to plant diversity and agricultural production. *Biodiversity*. (2008) 9:86–9. doi: 10.1080/14888386.2008.9712892
4. Yakovlev AY, Kruglikova AA, Chernysh SI. Calliphoridae flies in medical biotechnology. *Entomol Rev*. (2019) 99:292–301. doi: 10.1134/S0013873819030023
5. Rumpold BA, Schlüter OK. Potential and challenges of insects as an innovative source for food and feed production. *Innovat Food Sci Emerg Technol*. (2013) 17:1–11. doi: 10.1016/j.ifset.2012.11.005
6. Joseph I, Mathew DG, Sathyan P, Vargheese G. The use of insects in forensic investigations: an overview on the scope of forensic entomology. *J Foren Dental Sci*. (2011) 3:89–91. doi: 10.4103/0975-1475.92154

7. Grimaldi D, Engel MS. *Evolution of the Insects*. New York, NY: Cambridge University Press (2005).
8. Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim J-W, Lambkin C, et al. Episodic radiations in the fly tree of life. *Proc Natl Acad Sci USA*. (2011) 108:5690–5. doi: 10.1073/pnas.1012675108
9. Gäde G. The hypertrehalosaemic peptides of cockroaches: a phylogenetic study. *Gen Compar Endocrinol*. (1989) 75:287–300. doi: 10.1016/0016-6480(89)90082-8
10. Gäde G, Marco HG. The adipokinetic hormones of Odonata: a phylogenetic approach. *J Insect Physiol*. (2005) 51:333–41. doi: 10.1016/j.jinsphys.2004.12.011
11. Marco HG, Šimek P, Gäde G. Adipokinetic hormones of the two extant apterygotan insect orders, Archaeognatha and Zygentoma. *J Insect Physiol*. (2014) 60, 17–24. doi: 10.1016/j.jinsphys.2013.11.002
12. Gäde G, Marco HG. The adipokinetic hormone of the coleopteran suborder Adephaga: Structure, function, and comparison of distribution in other insects. *Arch Insect Biochem Physiol*. (2017) 95:e21399. doi: 10.1002/arch.21399
13. Gäde G, Marco HG, Desutter-Grandcolas L. A phylogenetic analysis of the adipokinetic neuropeptides of Ensifera. *Physiol Entomol*. (2003) 28:283–9. doi: 10.1111/j.1365-3032.2003.00344.x
14. Hansen KK, Stafflinger E, Schneider M, Hauser F, Cazzamali G, Williamson M, et al. Discovery of a novel insect neuropeptide signaling system closely related to the insect adipokinetic hormone and corazonin hormonal systems. *J Biol Chem*. (2010) 285:10736–47. doi: 10.1074/jbc.M109.045369
15. Roch GJ, Busby ER, Sherwood NM. Evolution of GnRH: diving deeper. *Gen Comp Endocrinol*. (2011) 171:1–16. doi: 10.1016/j.ygcen.2010.12.014
16. Gäde G, Šimek P, Marco HG. An invertebrate [hydroxyproline]-modified neuropeptide: further evidence for a close evolutionary relationship between insect adipokinetic hormone and mammalian gonadotropin hormone family. *Biochem Biophys Res Commun*. (2011) 414:592–7. doi: 10.1016/j.bbrc.2011.09.127
17. Gäde G. Peptides of the adipokinetic hormone/red pigment-concentrating hormone family. A new take on biodiversity. *Ann N Y Acad Sci*. (2009) 1163:125–36. doi: 10.1111/j.1749-6632.2008.03625.x
18. Marco H, Gäde G. Chapter 3: Adipokinetic hormone: a hormone for all seasons? In: Saleuddin S, Lange AB, and Orchard I, editors. *Advances in Invertebrate (Neuro)Endocrinology: A Collection of Reviews in the Post-Genomic Era, Vol 2*. Burlington, ON; Palm Bay, FL: Apple Academic Press (2020) P.129–175.
19. Jaffe H, Raina AK, Riley CT, Fraser BA, Nachman RJ, Vogel VW, et al. Primary structure of two neuropeptide hormones with adipokinetic and hypotrehalosemic activity isolated from the corpora cardiaca of horse flies (Diptera). *Proc Natl Acad Sci USA*. (1989) 86:8161–4. doi: 10.1073/pnas.86.20.8161
20. Woodring J, Leprince DJ. The function of corpus cardiacum peptides in horse flies. *J Insect Physiol*. (1992) 38:775–82. doi: 10.1016/0022-1910(92)90030-H
21. Gäde G, Wilps H, Kellner R. Isolation and structure of a novel charged member of the red–pigment-concentrating hormone-adipokinetic hormone family of peptides isolated from the corpora cardiaca of the blowfly *Phormia terraenovae* (Diptera). *Biochem J*. (1990) 269:309–13. doi: 10.1042/bj2690309
22. Wilps H, Gäde G. Hormonal regulation of carbohydrate metabolism in the blowfly *Phormia terraenovae*. *J Insect Physiol*. (1990) 36:441–50. doi: 10.1016/0022-1910(90)90062-K
23. Schaffer MH, Noyes BE, Slaughter CA, Thorne GC, Gaskell SJ. The fruitfly *Drosophila melanogaster* contains a novel charged adipokinetic-hormone-family peptide. *Biochem J*. (1990) 269:315–20. doi: 10.1042/bj2690315
24. Noyes BE, Katz FN, Schaffer MH. Identification and expression of the *Drosophila* adipokinetic hormone gene. *Mol Cell Endocrinol*. (1995) 109:133–41. doi: 10.1016/0303-7207(95)03492-P
25. Kaufmann C, Merzendorfer H, Gäde G. The adipokinetic hormone system in Culicinae (Diptera: Culicidae): molecular identification and characterization of two adipokinetic hormone (AKH) precursors from *Aedes aegypti* and *Culex pipiens* and two putative AKH receptor variants from *A. aegypti*. *Insect Biochem Mol Biol*. (2009) 39:770–81. doi: 10.1016/j.ibmb.2009.09.002
26. Predel R, Neupert S, Garczynski SF, Crim JW, Brown MR, Russell WK, et al. Neuropeptidomics of the mosquito *Aedes aegypti*. *J Proteome Res*. (2010) 9:2006–15. doi: 10.1021/pr901187p
27. Kaufmann C, Brown MR. Adipokinetic hormones in the African malaria mosquito, *Anopheles gambiae*: identification and expression of genes for two peptides and a putative receptor. *Insect Biochem Mol Biol*. (2006) 36:466–81. doi: 10.1016/j.ibmb.2006.03.009
28. Vicoso B, Bachtrog D. Numerous transitions of sex chromosomes in Diptera. *PLoS Biol*. (2015) 13:e1002078. doi: 10.1371/journal.pbio.1002078
29. Wegener C, Köppler K. Peptide profiling of the retrocerebral complex and identification of an adipokinetic hormone and short neuropeptide F peptides in diapausing and non-diapausing cherry fruit flies *Rhagoletis cerasi* (Diptera: Tephritidae). *Mitt Dtsch Ges Allg Angew Ent*. (2014) 19:265–8.
30. Baggerman G, Cerstiaens A, De Loof A, Schoofs L. Peptidomics of the larval *Drosophila melanogaster* central nervous system. *J Biol Chem*. (2002) 277:40368–74. doi: 10.1074/jbc.M206257200
31. Baggerman G, Boonen K, Verleyen P, De Loof A, Schoofs L. Peptidomic analysis of the larval *Drosophila melanogaster* central nervous system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. *J Mass Spectrom*. (2005) 40:250–60. doi: 10.1002/jms.744
32. Predel R, Wegener C, Russell WK, Tichy SE, Russell DH, Nachman RJ. Peptidomics of CNS-associated neurohemal systems of adult *Drosophila melanogaster*: a mass spectrometric survey of peptides from individual flies. *J Comp Neurol*. (2004) 474:379–92. doi: 10.1002/cne.20145
33. Wegener C, Reinl T, Jansch L, Predel R. Direct mass spectrometric peptide profiling and fragmentation of larval peptide hormone release sites in *Drosophila melanogaster* reveals tagma-specific peptide expression and differential processing. *J Neurochem*. (2006) 96:1362–74. doi: 10.1111/j.1471-4159.2005.03634.x
34. *Drosophila* 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*. (2007) 450:203–18. doi: 10.1038/nature06341
35. Wegener C, Gorbashov A. Molecular evolution of neuropeptides in the genus *Drosophila*. *Genome Biol*. (2008) 9:R131. doi: 10.1186/gb-2008-9-8-r131
36. Attardo GM, Benoit JB, Michalkova V, Yang GX, Roller L, Bohova J, et al. Analysis of lipolysis underlying lactation in the tsetse fly, *Glossina morsitans*. *Insect Biochem Mol Biol*. (2012) 42:360–70. doi: 10.1016/j.ibmb.2012.01.007
37. Caers J, Boonen K, Van Den Abbeele J, Van Rompay L, Schoofs L, Van Hiel MB. Peptidomics of neuropeptidergic tissues of the tsetse fly *Glossina morsitans morsitans*. *J Am Soc Mass Spectrom*. (2015) 26:2024–38. doi: 10.1007/s13361-015-1248-1
38. Attardo GM, Abd-Alla AMM, Acosta-Serrano A, Allen JE, Bateta R, Benoit JB, et al. Comparative genomic analysis of six *Glossina* genomes, vectors of African trypanosomes. *Genome Biol*. (2019) 20:1–30. doi: 10.1186/s13059-019-1768-2
39. Audsley N, Matthews HJ, Down RE, Weaver RJ. Neuropeptides associated with the central nervous system of the cabbage root fly, *Delia radicum* (L.). *Peptides*. (2011) 32:434–40. doi: 10.1016/j.peptides.2010.08.028
40. Zoephel J, Reiher W, Rexer KH, Kahnt J, Wegener C. Peptidomics of the agriculturally damaging larval stage of the cabbage root fly *Delia radicum* (Diptera: Anthomyiidae). *PLoS ONE*. (2012) 7:e41543. doi: 10.1371/journal.pone.0041543
41. Verleyen P, Huybrechts J, Sas F, Clynen E, Baggerman G, De Loof A, et al. Neuropeptidomics of the grey flesh fly *Neobellieria bullata*. *Biochem Biophys Res Commun*. (2004) 316:763–70. doi: 10.1016/j.bbrc.2004.02.115
42. Rahman MM, Neupert S, Predel R. Neuropeptidomics of the Australian sheep blowfly *Lucilia cuprina* (Wiedemann) and related Diptera. *Peptides*. (2013) 41:31–7. doi: 10.1016/j.peptides.2012.12.021
43. Inosaki A, Yasuda A, Shinada T, Ohfuné Y, Numata H, Shiga S. Mass spectrometric analysis of peptides in brain neurosecretory cells and neurohemal organs in the adult blowfly *Protophormia terraenovae*. *Comp Biochem Physiol A*. (2010) 155:190–9. doi: 10.1016/j.cbpa.2009.10.036
44. Stoffolano JG Jr, Croke K, Chambers J, Gäde G, Solari P, Liscia A. Role of Phote-HrTH. (*Phormia terraenovae* hypertrehalosemic hormone) in modulating the supercontractile muscles of the crop of adult *Phormia regina* Meigen. *J Insect Physiol*. (2014) 71:147–55. doi: 10.1016/j.jinsphys.2014.10.014

45. Gálíková M, Klepsatel P. Obesity and aging in the *Drosophila* model. *Int J Mol Sci.* (2018) 19:1896. doi: 10.3390/ijms19071896
46. Riehle MA, Garczynski SF, Crim JW, Hill CA, Brown MR. Neuropeptides and peptide hormones in *Anopheles gambiae*. *Science.* (2002) 298:172–5. doi: 10.1126/science.1076827
47. Kaufmann C, Brown MR. Regulation of carbohydrate metabolism and flight performance by a hypertrehalosaemic hormone in the mosquito *Anopheles gambiae*. *J Insect Physiol.* (2008) 54:367–77. doi: 10.1016/j.jinsphys.2007.10.007
48. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, et al. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science.* (2007) 316:1718–23. doi: 10.1126/science.1138878
49. Gäde G, Šimek P, Marco HG. The first identified neuropeptide in the insect order Megaloptera: a novel member of the adipokinetic hormone family in the alderfly *Sialis lutaria*. *Peptides.* (2009) 30:477–82. doi: 10.1016/j.peptides.2008.07.022
50. International Glossina Genome Consortium. Genome sequence of the tsetse fly (*Glossina morsitans*): vector of African trypanosomiasis. *Science.* (2014) 344:380–86. doi: 10.1126/science.1249656
51. Pape T, Blagoderov VA, Mostovski MB. Order Diptera Linnaeus, 1758. *Zootaxa.* (2011) 3148:222–9. doi: 10.11646/zootaxa.3148.1.42
52. Zöllner N, Kirsch K. Über die quantitative Bestimmung von Lipoiden. (Mikromethode) mittels der vielen natürlichen Lipoiden (allen bekannten Plasmalipoiden) gemeinsamen Sulfo-phosphovanillin-Reaktion. *Zeitschrift für die gesamte experimentelle Medizin.* (1962) 135:545–61. doi: 10.1007/BF02045455
53. Spik G, Montreuil J. Deux causes d'erreur dans les dosages colorimétriques des oses neutres totaux. *Bull Soc Chimie Biol.* (1964) 46:739–49.
54. Holwerda DA, van Doorn J, Beenackers AMT. Characterization of the adipokinetic and hyperglycaemic substances from the locust corpus cardiacum. *Insect Biochem.* (1977) 7:151–7. doi: 10.1016/0020-1790(77)90008-7
55. Gäde G, Goldsworthy GJ, Kegel G, Keller R. Single step purification of locust adipokinetic hormones I and II by reversed-phase high-performance liquid-chromatography, and amino acid composition of the hormone II. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie.* (1984) 365:393–8. doi: 10.1515/bchm2.1984.365.1.393
56. Kodrík D, Marco HG, Šimek P, Socha R, Štys P, Gäde G. The adipokinetic hormones of Heteroptera: a comparative study. *Physiol Entomol.* (2010) 35:117–127. doi: 10.1111/j.1365-3032.2009.00717.x
57. Gäde G. The explosion of structural information on insect neuropeptides. In: Herz W, Kirby GW, Moore RE, Steglich W, Tamm CH, editors. *Progress in the Chemistry of Organic Natural Products, Vol 71*. New York, NY: Springer (1997). p. 1–128. doi: 10.1007/978-3-7091-6529-4_1
58. Gäde G, Auerswald L, Šimek P, Marco HG, Kodrík D. Red pigment-concentrating hormone is not limited to crustaceans. *Biochem Biophys Res Commun.* (2003) 309:967–73. doi: 10.1016/j.bbrc.2003.08.107
59. Gäde G, Šimek P, Marco HG. Novel members of the adipokinetic hormone family in beetles of the superfamily Scarabaeoidea. *Amino Acids.* (2016) 48:2785–98. doi: 10.1007/s00726-016-2314-0
60. Marco HG, Gäde G. Structure and function of adipokinetic hormones of the large white butterfly *Pieris brassicae*. *Physiol. Entomol.* (2017) 42:103–12. doi: 10.1111/phen.12175
61. Gäde G, Šimek P, Marco HG. Structural diversity of adipokinetic hormones in the hyperdiverse coleopteran Cucujiformia. *Arch Insect Biochem Physiol.* (2019) 2019:e21611. doi: 10.1002/arch.21611
62. Gäde G, Marco HG. Peptides of the adipokinetic hormone/red pigment-concentrating hormone family with special emphasis on Caelifera: primary sequences and functional considerations contrasting grasshoppers and locusts. *Gen Comp Endocrinol.* (2009) 162:59–68. doi: 10.1016/j.ygcen.2008.06.007
63. Wiegmann BM, Trautwein MD, Kim J-W, Cassel BK, Bertone MA, Winterton SL, et al. Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol.* (2009) 7:34. doi: 10.1186/1741-7007-7-34
64. Gäde G, Šimek P, Fescemyer HW. Adipokinetic hormones provide inference for the phylogeny of Odonata. *J Insect Physiol.* (2011) 57:174–8. doi: 10.1016/j.jinsphys.2010.11.002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Gäde, Šimek and Marco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Evolution of Neuropeptide Precursors in Polyneoptera (Insecta)

Marcel Bläser and Reinhard Predel*

Department of Biology, Institute for Zoology, University of Cologne, Cologne, Germany

OPEN ACCESS

Edited by:

Elizabeth Amy Williams,
University of Exeter, United Kingdom

Reviewed by:

Sheila Ons,
National University of La
Plata, Argentina
Sven Bradler,
University of Göttingen, Germany

*Correspondence:

Reinhard Predel
rpredel@uni-koeln.de

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 30 January 2020

Accepted: 19 March 2020

Published: 15 April 2020

Citation:

Bläser M and Predel R (2020)
Evolution of Neuropeptide Precursors
in Polyneoptera (Insecta).
Front. Endocrinol. 11:197.
doi: 10.3389/fendo.2020.00197

Neuropeptides are among the structurally most diverse signaling molecules and participate in intercellular information transfer from neurotransmission to intrinsic or extrinsic neuromodulation. Many of the peptidergic systems have a very ancient origin that can be traced back to the early evolution of the Metazoa. In recent years, new insights into the evolution of these peptidergic systems resulted from the increasing availability of genome and transcriptome data which facilitated the investigation of the complete neuropeptide precursor sequences. Here we used a comprehensive transcriptome dataset of about 200 species from the 1KITE initiative to study the evolution of single-copy neuropeptide precursors in Polyneoptera. This group comprises well-known orders such as cockroaches, termites, locusts, and stick insects. Due to their phylogenetic position within the insects and the large number of old lineages, these insects are ideal candidates for studying the evolution of insect neuropeptides and their precursors. Our analyses include the orthologs of 21 single-copy neuropeptide precursors, namely ACP, allatotropin, AST-CC, AST-CCC, CCAP, CCHamide-1 and 2, CNMamide, corazonin, CRF-DH, CT-DH, elevenin, HanSolin, NPF-1 and 2, MS, proctolin, RFLamide, SIFamide, sNPF, and trissin. Based on the sequences obtained, the degree of sequence conservation between and within the different polyneopteran lineages is discussed. Furthermore, the data are used to postulate the individual neuropeptide sequences that were present at the time of the insect emergence more than 400 million years ago. The data confirm that the extent of sequence conservation across Polyneoptera is remarkably different between the different neuropeptides. Furthermore, the average evolutionary distance for the single-copy neuropeptides differs significantly between the polyneopteran orders. Nonetheless, the single-copy neuropeptide precursors of the Polyneoptera show a relatively high degree of sequence conservation. Basic features of these precursors in this very heterogeneous insect group are explained here in detail for the first time.

Keywords: neuropeptides, transcriptome, Polyneoptera, insect evolution, Blattodea, Orthoptera, Phasmatodea, Dermaptera

INTRODUCTION

Neuropeptides are among the structurally most diverse signaling molecules in multi-cellular animal organisms. As such, they participate in intercellular information transfer from neurotransmission to intrinsic or extrinsic neuromodulation and regulate physiological processes including growth, reproduction, development, and behavior. For a single insect species, up to 50

neuropeptide genes can be expected coding for single or multiple copies of neuropeptides (1, 2). The sequences of single-copy neuropeptides, which are the focus of our study, are on average better conserved than those of multiple-copy peptides because amino acid (AA) substitutions in the single ligand of a particular neuropeptide receptor are potentially more likely to lead to a general loss of function than substitutions involving only one of several related neuropeptides produced from the same precursor. Thus, mutations that alter the sequences of single-copy neuropeptides must either be accompanied by parallel mutations in receptor genes that maintain the binding properties of the respective receptors or should not alter the steric properties of the peptides to maintain functionality (3). Most neuropeptides activate peptide-specific G-protein coupled receptors and many of these peptidergic systems have a very ancient origin that can be traced back to the early evolution of the Metazoa [e.g., (4)]. In fact, in many cases orthologies between neuropeptide and/or corresponding receptor genes of distantly related lineages can be identified (5–7). The identification of cockroach sulfakinins (8) with an already then suspected relationship to the cholecystokinins of vertebrates was a first strong indication of the conservation of peptidergic systems across protostomes and deuterostomes, these taxa diverged more than 700 million years ago (9, 10). Mirabeau and Joly (6) later described eight conserved peptidergic systems (vasopressin, neuropeptide Y/F, tachykinin, gonadotropin-releasing hormone/adipokinetic hormone, cholecystokinin/sulfakinin, neuromedin U/pyrokinin, corticotropin-releasing factor, calcitonin) which probably already occurred in the last common ancestor of Bilateria. More recently, Elphick et al. (4) have described as much as 30 neuropeptide signaling systems with orthologs in protostomes and deuterostomes.

Insects have always been in the focus of neuropeptide research (11). Today, the fruitfly *Drosophila melanogaster* is the model organism also for the study of neuropeptide functions (12). However, due to the large number of harmful insects that have an impact on human health or reduce yields in agriculture and forestry, other groups of insects, in particular true bugs, beetles, lepidopterans, and various groups of flies have also been intensively examined [e.g., (13–18)]. Beneficial insects such as honey bees and predators or parasites of pest insects [e.g., (19–21)] were also investigated in detail. Nevertheless, many neuropeptides in insects were first described from Polyneoptera such as locusts, cockroaches and stick insects; including proctolin, the first ever identified insect neuropeptide (22). In recent years, most of the insights into the evolution of peptidergic systems then resulted from the increasing availability of genome and transcriptome data which facilitated the investigation of the complete neuropeptide precursor sequences. Such data have been used in several comprehensive studies, mainly to compile neuropeptide precursor sequences within higher taxa (23–25). In a recent study a large dataset of neuropeptide precursors from Blattodea was used to demonstrate the considerable phylogenetic information contained in these sequences (3). A particular focus on the general evolution of neuropeptide precursors has been placed in a study on the precursors of 12 *Drosophila* species (26). In this study the authors also discussed the potentially

different evolution of single-copy and multiple-copy precursors, with mutations in the neuropeptide sequences of single-copy precursors being exposed to stronger stabilizing selection.

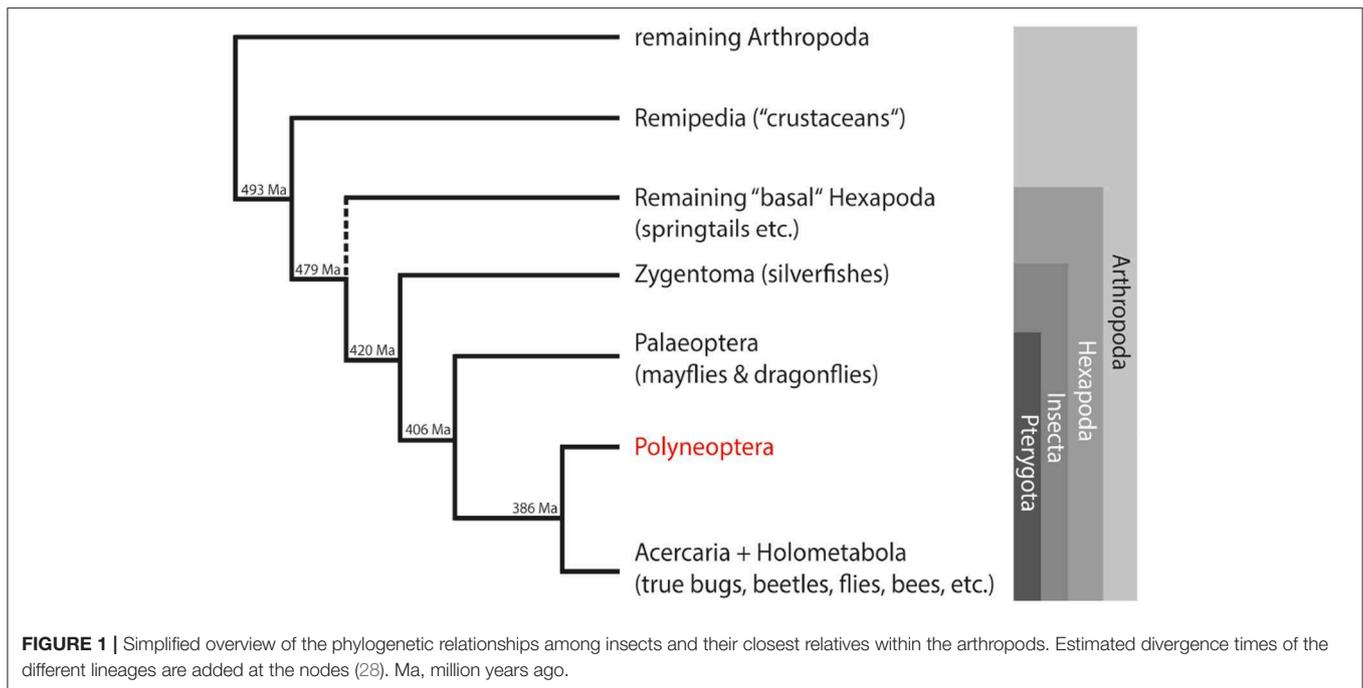
In our study, we used a comprehensive transcriptome dataset of about 200 species from the 1KITE initiative (<http://www.1kite.org/>) to study the evolution of single-copy neuropeptide precursors in Polyneoptera. This group comprises well-known orders such as cockroaches and termites (Blattodea), locusts (Orthoptera) and stick insects (Phasmatodea), but also rather unknown orders with few species such as ice crawlers (Grylloblattodea), heel walkers (Mantophasmatodea), and angel insects (Zoraptera). The internal relationships of Polyneoptera were recently resolved as part of the 1KITE project (27). Due to their phylogenetic position within the Insecta (**Figure 1**) and the large number of old lineages, these insects are ideal candidates for studying the evolution of insect neuropeptides and their precursors. Furthermore, the taxon sampling of the 1KITE initiative includes multiple species from all higher lineages of Polyneoptera (with the exception of Zoraptera with only one species). This enabled us to discuss changes in single-copy neuropeptide precursor sequences with respect to insect evolution.

Our analyses include the orthologs of 21 single-copy neuropeptide precursors. The neuropeptides of several of these precursors were first described from Polyneoptera (cockroaches: proctolin, corazonin, myosuppressin; stick insects: HanSolin, RFLamide), but were later also found in other insects. Based on the sequences obtained, the degree of sequence conservation between and within the different polyneopteran lineages is discussed. Furthermore, the data are used to postulate the individual neuropeptide sequences that were present at the time of the insect emergence more than 400 million years ago, as well as to detect taxon-specific losses of peptidergic systems. The dataset from the 1KITE initiative has become “historic” and contains a number of transcriptomes with partly insufficient coverage of neuropeptide precursors. The few new transcriptomes we have prepared specifically for this study (Mantophasmatodea) show much better data coverage. Nevertheless, the comprehensive taxon sampling of the 1KITE initiative allows a sufficient validation of almost all statements we provide about the evolution of neuropeptide precursors in Polyneoptera.

MATERIALS AND METHODS

Orthology Assessment and Alignment of Neuropeptide Precursor Sequences

Orthology assessment and alignments were performed as described in Bläser et al. (3). Briefly, we mined transcriptome sequences, provided by the 1KITE initiative (GenBank Umbrella BioProject ID PRJNA183205), for single-copy neuropeptide precursors in the datasets from each order of Polyneoptera; starting with neuropeptide precursor sequences of *Carausius morosus* (2), *Locusta migratoria* (1) and Blattodea (3). Once a full set of single-copy neuropeptide precursors was obtained, we used this information to search for precursors in the remaining



species of this order or species from related orders. Assembled transcripts were analyzed with the *tblastn* algorithms provided by NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Identified candidate nucleotide precursor gene sequences were translated into AA sequences using the ExpPASy Translate tool (29) with the standard genetic code. Orthologous neuropeptide precursors were aligned using the MAFFT-L-INS-i algorithm (30) (*dvtditr* (aa) Version 7.299b *alg=A*, *model=BLOSUM62*, *1.53*, *-0.00*, *-0.00*, *noshift*, *amax=0.0*). Alignments generated with the MAFFT-L-INS-i algorithm were then manually checked for misaligned sequences using N-termini of signal peptides and conserved AA residues (cleavage signals, Cys as target for disulphide bridges) as anchor points. Incompletely translated transcripts of neuropeptide precursors and transcripts of questionable quality were either combined to generate complete precursors when possible, or labeled with question marks at the respective AA positions.

Assessment of Precursor Characteristics

Individual AA alignments of each group of orthologous neuropeptide precursors from each order were merged in BioEdit 7.2.5 (31). The coverage of single-copy neuropeptide precursors in our dataset (Additional File 1), minimal and maximal length of precursors as well as number of identified transcripts and the position of the conserved neuropeptide sequences in the precursor were manually determined for each neuropeptide precursor in each order of Polyneoptera, respectively. Additionally, the predicted neuropeptide sequences in the individual AA alignments were determined and further analyzed using BioEdit 7.2.5. The lengths of these sequences as well as N-terminal and C-terminal cleavage sites were manually determined.

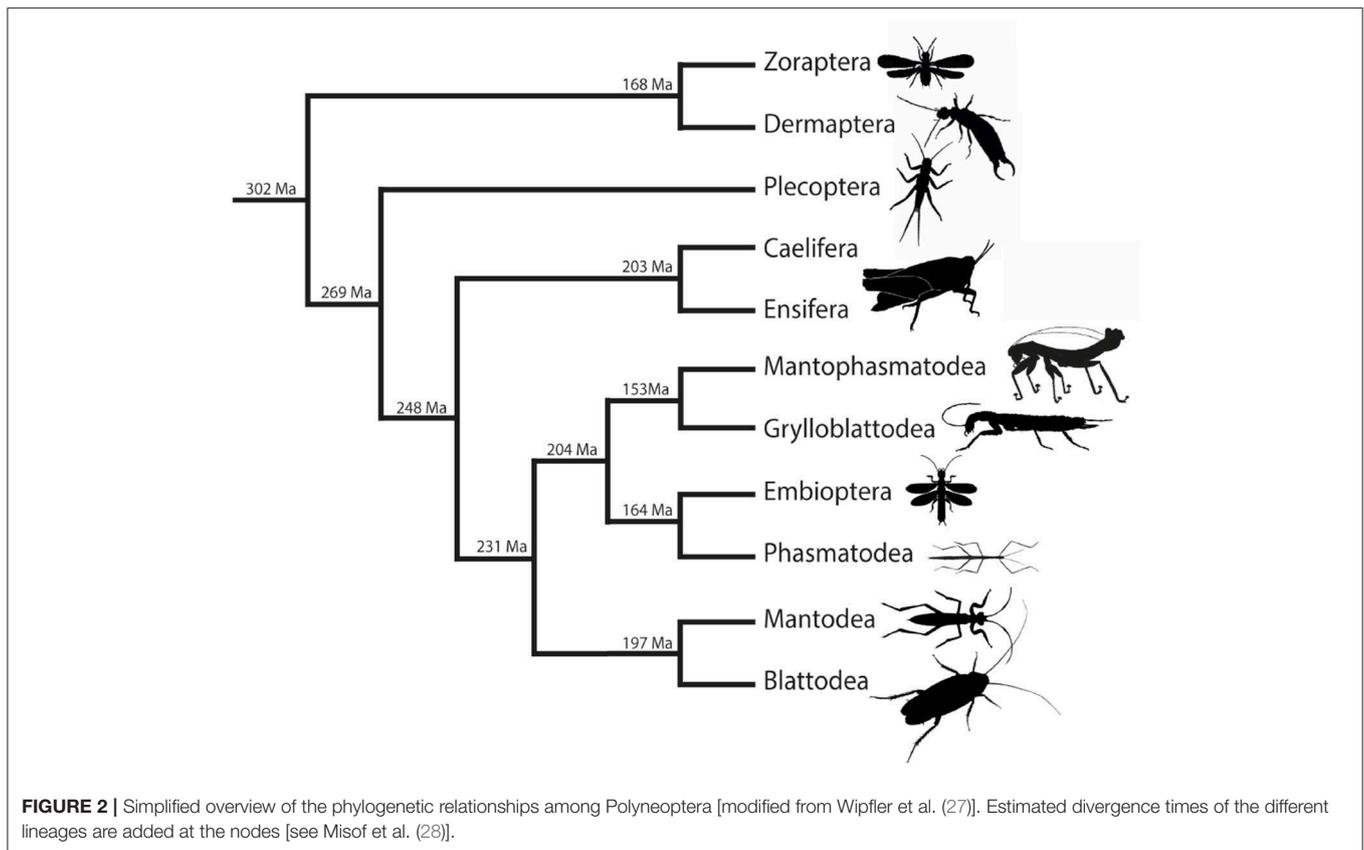
Alignments of single-copy neuropeptide precursors of the individual polyneopteran orders as well as combined alignments

of all polyneopteran lineages were used to estimate the average evolutionary divergence (AED) over all sequence pairs in Mega X (32). Standard error estimates were obtained by implementing 500 bootstrap replicates using the Poisson correction model (33). We used the pairwise deletion option to remove all ambiguous sites. The results of these analyses are shown in Additional File 3. The median AED of each single-copy neuropeptide precursor was calculated with Microsoft Excel and compared to the overall AED value of the respective single-copy neuropeptide precursor. To calculate the overall AED value, complete sequences of each single-copy neuropeptide from each polyneopteran order (see Additional File 1) were merged into a single file and aligned again. Furthermore, the alignments of the predicted neuropeptide sequences were used to calculate the overall AED value for the conserved neuropeptide sequences for each neuropeptide. Finally the overall median AED of all single-copy neuropeptides for each order was calculated using Microsoft Excel.

These analyses enabled us to compare internal sequence variation for each neuropeptide in all polyneopteran orders (AED for each order), as well as between orders (overall AED). The median AED also allowed an assessment of relative levels of sequence conservation. The overall AED of the predicted neuropeptide sequences enables a comparison of sequence conservation between neuropeptide sequences and the complete precursor sequence.

Sequence Logo Generation and Topology Mapping

Sequence logos of the aligned neuropeptide precursor orthologs were generated using the tool WebLogo version 2.8.2 (34). Each stack represents one position in the multiple sequence alignment. The overall height of a stack indicates the sequence conservation at this AA position; the height of letters within



the stack indicates the relative frequency of each AA at that position. For the color scheme of AA residues, the default settings were selected. Resulting sequence logos were manually mapped (Adobe Illustrator CS6; version 16.0.0.) on a simplified tree showing the phylogeny of Polyneoptera (27).

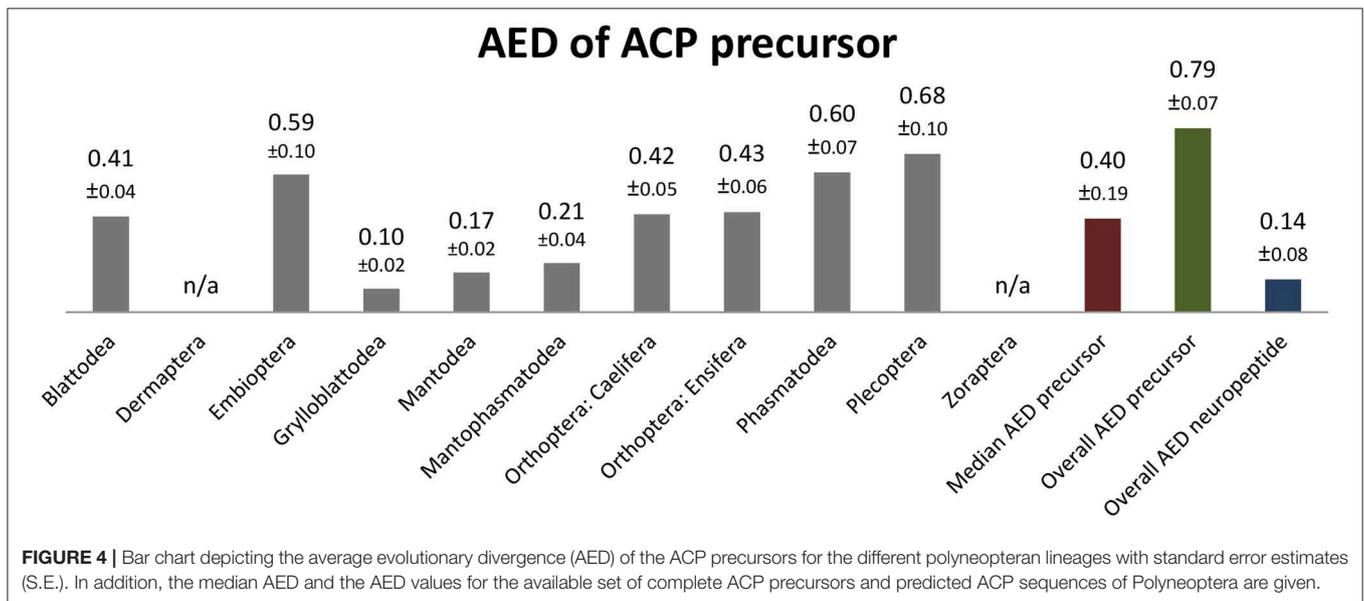
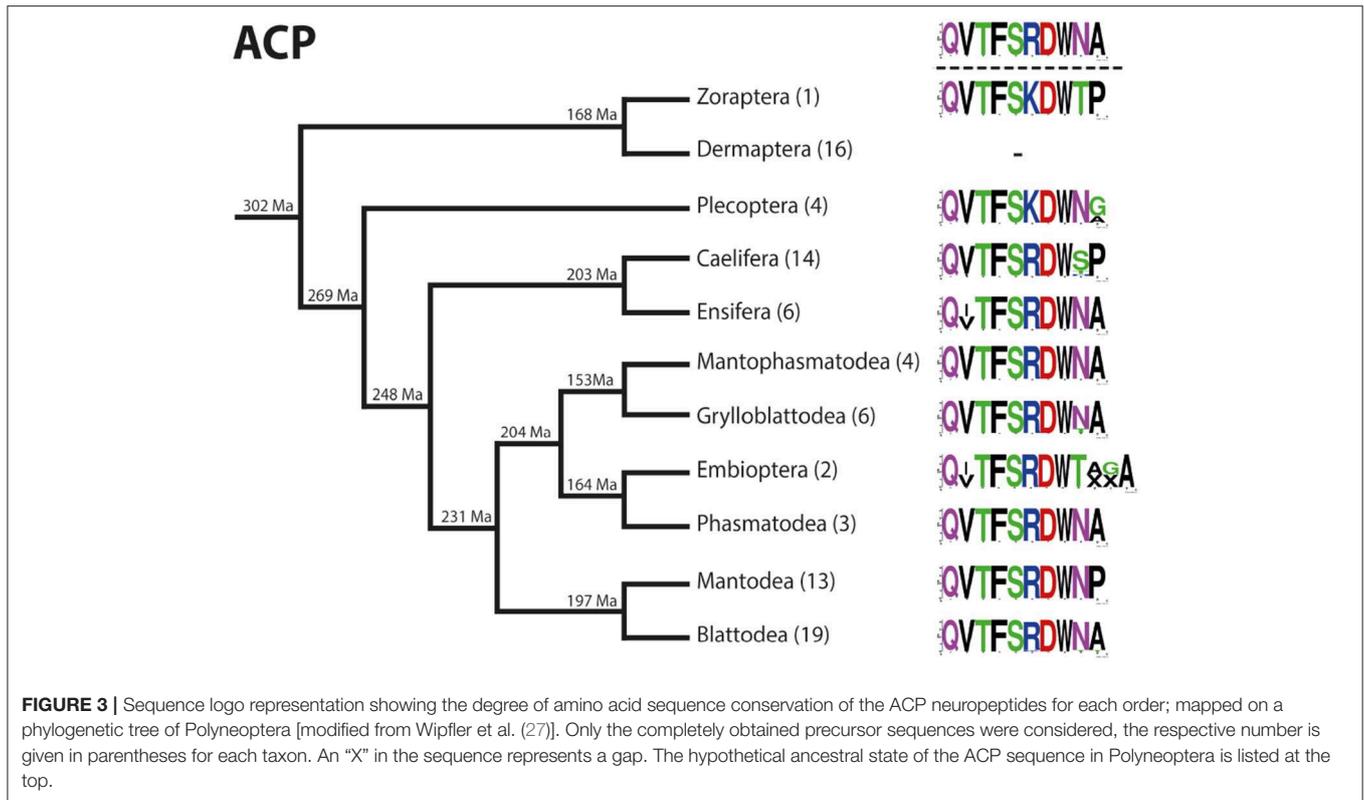
RESULTS AND DISCUSSION

The BLAST searches in the Polyneoptera transcriptome assemblies of the 1KITE initiative were performed with single-copy neuropeptide precursor sequences of *C. morosus* (2), *L. migratoria* (1), and Blattodea (3). Due to the varying quality of the transcriptome data and the generally low quantity of several neuropeptide-coding RNA sequences in whole body transcriptomes (see **Additional File 1**), the number of identified precursor sequences is significantly lower than the number of species analyzed. In addition, the yield of neuropeptide precursors is different for the different precursors; for example, much fewer CCHamide-1 precursors could be identified across the different lineages than precursors for other neuropeptides such as proctolin and NPF-1 (see **Additional File 1**). Nevertheless, the extensive material of the 1KITE initiative guaranteed sufficient information for almost all orders of Polyneoptera. The only exception was Zoraptera, where only a single transcriptome is available. In total, we have included 21 different single-copy precursors in our analysis. The precursors for adipokinetic hormones (AKHs) were not included in our study because the number of AKH genes varies

considerably between and within the polyneopteran orders and the orthologies could not be resolved. All neuropeptide precursor sequences identified in this study are listed in **Additional File 2**, sorted by the different polyneopteran lineages. The phylogenetic relationships in Polyneoptera are illustrated in **Figure 2** which also shows the estimated divergence times of the different orders. The estimated divergence times between orders vary between 300 and 150 Ma, which corresponds to the time scale of the parallel (independent) evolution of the respective precursors. Due to the long separate history of the two orthopteran lineages Caelifera and Ensifera (**Figure 2**), we have treated them separately in our analyses. The information about the neuropeptide sequences for each lineage is used in the following to determine the respective ancestral neuropeptide sequences for the Polyneoptera. A comparison with orthocopies of *Zygentoma* and *Remipedia* (24) allows in many cases also a statement about the possible ancestral neuropeptide sequences of the insects or even of the hexapods.

ACP (Additional File 3A)

A single adipokinetic hormone/ corazonin-like peptide (ACP) precursor with a length of 85–109 AA is present in almost all polyneopteran orders. The only exception was found in Dermaptera, where the ACP precursor is absent. It is noteworthy that of the 37 phasmatodean species analyzed, an ACP precursor was found only in 5 species of Oriophasmata (35). However, the transcriptomes of Phasmatodea in general show a rather high percentage of missing data (see **Additional File 1**) and therefore the low number of ACP precursors in Phasmatodea might also



be a result of this incomplete dataset. Each precursor contains a usually very well conserved ACP motif with an amidation site. The ACP sequence immediately follows the signal peptide and terminates upstream of a dibasic KR cleavage site.

The ACP sequences are mostly decapeptides; only in one species of Embioptera (*Rhagadochir virgo*) ACP is a duodecapeptide. The sequence (p)QVTF SRD WNA-NH₂, which also occurs in the remipedian *X. tulumensis*, was found in

various orders of Polyneoptera (Figure 3). This sequence is likely ancestral for all Hexapoda. Amino acid substitutions in ACPs of Polyneoptera are largely limited to substitutions from Val² to Ile² (few Ensifera and Embioptera), Arg⁶ to Lys⁶ (Zoraptera, all Plecoptera), and several substitutions of the two C-terminal AA (Figure 3). The median average evolutionary divergence (AED) for the ACP precursor is 0.40 (Figure 4). Grylloblattodea show the lowest ACP precursor variation, while Plecoptera

possess the most variable ACP precursor sequences. The overall AED for the ACP precursors is 0.79, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.14; **Figure 4**).

AST-CC (Additional File 3B)

A single allatostatin-CC (AST-CC) precursor with a length of 108–172 AA is present in all polyneopteran orders. The precursors contain a well conserved AST-CC motif without a C-terminal amidation site. Only in *Aptoperla tikumana* (Plecoptera) a second partial precursor with an alternative AST-CC motif was found. The AST-CC sequence is always located C-terminal in the precursor, N-terminally flanked by a dibasic RR cleavage site [monobasic Arg in *Tryonicus parvus* (Blattodea)] and terminates upstream of variable C-terminal cleavage sites or the AST-CC precursor sequence ends directly with the AST-CC sequence.

Most AST-CC sequences of Polyneoptera are non-decapeptides, only some derived AST-CC sequences of Orthoptera have 18 (few Caelifera) or 20 AA [few Ensifera and *T. parvus* (Blattodea)]. The sequence GQKGRVYWRVCFNAVTCF-OH which also occurs in the silverfish *T. domestica* was found in most orders of Polyneoptera (not in Dermaptera, Caelifera, Embioptera). This sequence might therefore be regarded as ancestral for all Pterygota. While the C-terminal motif YWRVCFNAVTCF-OH is highly conserved in all Polyneoptera, the N-terminus shows a number of lineage-specific AA substitutions, particularly at position 2 and 3. The median AED for the AST-CC precursor is 0.24. Grylloblattodea and Mantophasmatodea show the lowest AST-CC precursor variation, while Embioptera possess the most variable AST-CC precursors. The overall AED for the AST-CC precursor is 0.49, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.08).

AST-CCC (Additional File 3C)

A single allatostatin-CCC (AST-CCC) precursor with a length of 93–121 AA is present in all polyneopteran orders. The precursors contain a highly conserved AST-CCC motif with an amidation site. The AST-CCC sequence is always located C-terminal in the precursor, N-terminally flanked by a dibasic KR cleavage site and terminates upstream of a monobasic K cleavage site (KK in few species of Caelifera).

All AST-CCC sequences of Polyneoptera are tetradecapeptides. The sequence SYWKQCAFNAVSCF-NH₂ which also occurs in the remipedian *X. tulumensis* was found in all orders of Polyneoptera. This sequence might therefore be regarded as ancestral for all Hexapoda. Amino acid substitutions in AST-CCC are limited to a substitution from Lys⁴ to Arg⁴ in few species of Dermaptera and Embioptera and all species of Mantodea. The median AED for the AST-CCC precursor is 0.23. Mantophasmatodea show the lowest AST-CCC precursor variation, while Ensifera possess the most variable AST-CCC precursors. The overall AED for the AST-CCC precursor is 0.36, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.03).

In our study we could not find an AST-C precursor in any transcriptome. The presence of an AST-C precursor was previously suggested for *L. migratoria* (36, 37), but the corresponding AST-C neuropeptide has not been confirmed biochemically (e.g., fragment analysis).

AT (Additional File 3D)

A single allatotropin (AT) precursor with a length of 102–142 AA is present in all polyneopteran orders. The precursor contains a usually well-conserved AT motif with a C-terminal amidation site. The AT sequence follows a precursor peptide (17–21 AA) inserted between the signal peptide and the AT sequence. Only in the Dermaptera the AT sequence is directly C-terminal of the signal peptide. In all other taxa, the AT sequence is N-terminally flanked by a monobasic Arg cleavage site, while all sequences terminate upstream of a dibasic KR cleavage site. Specific features of AT precursors were found in Embioptera, where the species *Haploembia palaui* has a second AT precursor and another species (*Ptilocerembia catherinae*) has a second, longer AT motif immediately C-terminal of the first AT sequence.

With a single exception (Ensifera: *Comicus calcaris*; 12 AA), the AT sequences (AT-1 of *P. catherinae*) are tridecapeptides. The sequence GFKNVALSTARGF-NH₂ which also occurs in the silverfish *T. domestica* was found in many orders of Polyneoptera (not in Zoraptera, Dermaptera, Mantophasmatodea, Grylloblattodea, Embioptera). This sequence might therefore be regarded as ancestral for all Pterygota. Common AA substitutions of AT sequences affect the positions 5 and/or 6 from the N-terminus, resulting in lineage-specific AA at these positions. Highly derived sequences of AT are typical of all Dermaptera and most Embioptera; in Dermaptera these substitutions even affect the N- and C-terminal AA, which are conserved in all other polyneopteran orders. The median AED for the ACP precursor is 0.3. Grylloblattodea and Mantophasmatodea show the lowest ACP precursor variation in Polyneoptera, while Embioptera possess the most variable ACP precursors. The overall AED for the ACP precursors is 0.48, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.10).

CCAP (Additional File 3E)

A single crustacean cardioactive peptide (CCAP) precursor with a length of 143–174 AA is present in all polyneopteran orders. The precursor contains a fully conserved CCAP motif with a C-terminal amidation site. The CCAP sequence follows a precursor peptide (24–27 AA) inserted between the signal peptide and the CCAP sequence. The CCAP sequences are N-terminally flanked by dibasic KR cleavage sites and terminate upstream of KKR or RKR (few Orthoptera and Embioptera) cleavage sites. The sequence PFCNAFTGC-NH₂ which also occurs in the remipedian *X. tulumensis* was found in all orders of Polyneoptera. This sequence might therefore be regarded as ancestral for all Hexapoda. A single AA substitution from Phe⁶ to Leu⁶ was found in *Creoxylus spinosus* (Phasmatodea). The median AED for the CCAP precursor is 0.22. Grylloblattodea and Mantophasmatodea show the lowest CCAP precursor variation in Polyneoptera, while Plecoptera possess the most variable CCAP precursors.

The overall AED for the ACP precursors is 0.45, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.00).

CCHamide-1 (Additional File 3F)

A single CCHamide-1 precursor with a length of 115–241 AA is present in almost all polyneopteran orders. The only exception was found in Ensifera, where the CCHamide-1 precursor is absent. The precursor contains a usually well-conserved CCHamide-1 motif with a C-terminal amidation site. The CCHamide-1 sequence in the precursor immediately follows the signal peptide and terminates upstream of a dibasic KR cleavage site.

With very few exceptions (N-terminally extended in Embioptera and possibly also in Caelifera), the predicted CCHamide-1 sequences are always tetradecapeptides. The sequence GSCLSYGHSCWGAH-NH₂ which also occurs in the silverfish *T. domestica* was found in most orders of Polyneoptera (not in Zoraptera). This sequence might therefore be regarded as ancestral for all Pterygota. Significant intraordinal variation is only present in Dermaptera, Plecoptera, Blattodea, and Mantodea. The most common AA substitution across different orders was that of Ala¹³ to Gly¹³ (Zoraptera and few Mantophasmatodea, Mantodea and Blattodea). The median AED for the CCHamide-1 precursor is 0.52. Mantophasmatodea show the lowest CCHamide-1 precursor variation, while Plecoptera have the most variable CCHamide-1 precursors. The overall AED for the CCHamide-1 precursors is 0.81, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.07).

CCHamide-2 (Additional File 3G)

A single CCHamide-2 precursor with a length of 97–174 AA is present in all polyneopteran orders. The precursors contain a well-conserved CCHamide-2 motif with a C-terminal amidation site. The CCHamide-2 sequence in the precursor follows immediately the signal peptide and terminates upstream of a dibasic KR cleavage site. An analysis of the *C. morosus* peptidome (2) has shown that the N-terminal KR of CCHamide-2 is not recognized as a cleavage site and we hypothetically assume that the same N-terminus occurs in all polyneopteran CCHamide-2. Two precursors with different CCHamide-2 sequences were found in the species *H. palau* (Embioptera). Additionally, in several species of Caelifera a second transcript of the CCHamide-2 precursor was found. In these species, 16 AA (P/SYGVRR/TPGD/AIQI/TRRAG) are inserted following the N-terminal KR in the respective CCHamide-2 sequences.

With few exceptions in Blattodea (*Nocticola* sp.: 16 AA; *Catara rugosicollis*, *Coptotermes* sp.: 14 AA), the predicted CCHamide-2 sequences are always pentadecapeptides. The sequence KRGCsAFGHSCFGGH-NH₂ which also occurs in several silverfish species, but not *T. domestica*, was found in most orders of Polyneoptera (not in Plecoptera, Dermaptera). This sequence might therefore be regarded as ancestral for all Pterygota. Common AA substitutions in the CCHamide-2 sequences are Ala⁴ to Ser⁴ (Phasmatodea, several Blattodea, few Plecoptera) and Phe¹⁰ to Tyr¹⁰ (few Embioptera, Plecoptera, and Dermaptera). The AA at position 3 of the N-terminus

show lineage-specific AA substitutions in the majority of Dermaptera (Ser to Gln), Plecoptera (Ser to Thr) and Caelifera (Ser to Met). The median AED for the CCHamide-2 precursor is 0.37. Mantophasmatodea show the lowest CCHamide-2 precursor variation in Polyneoptera, while Dermaptera possess the most variable CCHamide-2 precursors. The overall AED for the CCHamide-2 precursors is 0.72, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.10).

CNMamide (Additional File 3H)

CNMamide precursors with a length of 127–179 AA are present in all polyneopteran orders. Two transcripts with different CNMamide sequences were found in several Blattodea. The less commonly found longer transcripts (7 species of Blattodea) are orthologs of the single CNMamide precursor of Mantodea (sister group of Blattodea). In Caelifera, we identified two precursors in the species *Haplotropis brunneriana* and *Pielomastax soochowensis*. One of these precursors is very similar to the single CNMamide precursors of other Polyneoptera, while the second precursor, found in the majority of caeliferan species (27), has a rather variable and N-terminally extended sequence. In Ensifera, we have also identified two very different precursor sequences, but these sequences were always found in different species. It therefore remains unclear whether these sequences represent different transcripts or result from the rapid sequence diversification of CNMamide precursors. All precursors contain the CNMamide motif with a C-terminal amidation site. The CNMamide sequence in the precursor is located C-terminal in the precursor, N-terminally flanked by a dibasic KR cleavage site (KK in the ensiferan *Acheta domesticus* and *Phaeophila crisbredoides*), and terminates upstream of variable C-terminal cleavage sites (mostly RKR).

Most CNMamides are tetradecapeptides, but the full variation is from 13 to 18 AA. The sequence GSYMSLCHFkICNM-NH₂ which also occurs in the silverfish *T. domestica* was found in many orders of Polyneoptera (not in Dermaptera, Caelifera, Mantophasmatodea, Embioptera, Mantodea). This sequence might therefore be regarded as ancestral for all Pterygota. Common AA substitutions in the CNMamide sequences are Gly¹ to Asn¹ (all Mantophasmatodea), Gly¹ to Thr¹ (all Embioptera), and Ser⁵-Leu⁶ to Thr⁵-Met⁶ (all Mantodea). Particularly the CNMamide sequences of Caelifera, Blattodea, and Ensifera in which two different precursors are present are quite variable at the N-terminus. A similar sequence variation was also found in Dermaptera and Plecoptera. The median AED for the CCHamide-2 precursor is 0.44. Mantophasmatodea show the lowest CCHamide-2 precursor variation in Polyneoptera, while Plecoptera possess the most variable CCHamide-2 precursors. The overall AED for the CCHamide-2 precursors is 0.95, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.26).

Corazonin (Additional File 3I)

A corazonin precursor with a length of 85–140 AA was found in almost all polyneopteran orders. The only exception was obtained in Zoraptera, where the corazonin precursor is absent. For two species, *Nippancistroger testaceus* (Ensifera) and *Medauroidea*

extradentata (Phasmatodea), we identified a second precursor with different corazonin sequences. Otherwise, the corazonin precursors contain highly conserved corazonin motifs with C-terminal amidation sites. The corazonin sequence in the precursor follows immediately the signal peptide and terminates upstream of a RKR cleavage site.

Corazonin sequences are almost exclusively undecapeptides. Only in one species of Dermaptera (*Nesogaster amoenus*), corazonin has 9 AA and the C-terminal dipeptide is missing. The sequence (p)QTFQYSRGWTN-NH₂, which also occurs in Malacostraca and even Myriapoda, was found in most orders of Polyneoptera (not in Phasmatodea and Dermaptera). This sequence might therefore be regarded as ancestral for all Hexapoda. Peptidomics confirmed for different polyneopteran taxa that the N-terminal Gln of corazonin is almost completely converted to pGlu (38). Amino acid substitutions in corazonin of Polyneoptera are largely limited to substitutions from Arg⁷ to His⁷ (many Phasmatodea, Caelifera and Dermaptera). Considering the phylogenetic position of the respective insect taxa (**Figure 1**), the His⁷-corazonins probably evolved several times independently of each other. Significant intraordinal variation is only present in Dermaptera. In Mantophasmatodea, all species of a single lineage (Austrophasmatidae) have a unique corazonin sequence with two AA substitutions (Gln⁴ to His⁴; Arg⁷ to Gln⁷), while in all other Mantophasmatodea the original sequence is retained (39). The median AED for the corazonin precursor is 0.37. Mantophasmatodea show the lowest corazonin precursor variation, while Dermaptera possess the most variable corazonin precursor sequences. The overall AED for the corazonin precursors is 0.82, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.07).

CRF-DH (Additional File 3J)

A precursor for the corticotropin-releasing factor-like diuretic hormone (CRF-DH) was found with a length of 138-284 AA in almost all polyneopteran orders. A partial sequence (N-terminal) of a possible CRF-DH was identified for the single species of Zoraptera. This sequence is not further considered here. The CRF-DH precursors contain a mostly well-conserved CRF-DH motif with C-terminal amidation site. The CRF-DH sequence is located in the middle of the precursor and is flanked by dibasic KR cleavage sites in most polyneopteran taxa. Notable exceptions are the precursors of Dermaptera which have a C-terminal RKR cleavage site (not in *Parapsalis infernalis*) and lack the N-terminal KR cleavage motif. Therefore, the sequences of the mature CRF-DHs of Dermaptera cannot be predicted with certainty and require biochemical confirmation first.

Most CRF-DHs of Polyneoptera consist of 46 AA, shorter sequences are indicated for several Plecoptera (42-46 AA), Ensifera (45-46 AA), Caelifera (44-46 AA), Mantophasmatodea (44 AA), and a single species of Blattodea (45 AA in *Nocticola*). Dermaptera have N-terminal extended CRF-DHs (see above). Due to considerable sequence variations, particularly in Dermaptera, an ancestral sequence of CRF-DH for Polyneoptera cannot be determined with certainty. All species contain a consensus sequence of PLSIVNxxDVLQRxxLExxRxRMR within the CRF-DH. The variable AA (x) within this sequence

decrease significantly if the CFR-DHs of Dermaptera are not considered (PSLSIVNxxDVLQRLLLExxARRRMR). The median AED for the CFR-DH precursor is 0.30. Mantophasmatodea show the lowest CFR-DH precursor variation, while Ensifera possess the most variable CFR-DH precursor sequences. The overall AED for the CFR-DH precursors is 0.62, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.31).

CT-DH (Additional File 3K)

A precursor for the calcitonin-like diuretic hormone (CT-DH) was found with a length of 107-178 AA in all polyneopteran orders. The CT-DH precursors contain a highly conserved CT-DH motif with C-terminal amidation site. The CT-DH sequence is located in the middle of the precursor, N-terminally flanked by a dibasic KR cleavage site, and terminates upstream of an RRRR cleavage site (RKRR in Plecoptera).

All CT-DHs of Polyneoptera consist of 31 AA. The sequence GLDLGLSRGFSGSQAAKHLMGLAAANYAGGP-NH₂ which also occurs in the silverfish *T. domestica* was found in most orders of Polyneoptera (not in Zoraptera, Dermaptera, Caelifera). This sequence might therefore be regarded as ancestral for all Pterygota. The remarkable sequence conservation of CT-DH is unique for such a long neuropeptide. The few AA substitutions in CT-DHs of Polyneoptera are often lineage-specific and cover all species within the corresponding insect orders: Phe¹⁰ to Tyr¹⁰ and Tyr²⁷ to Phe²⁷ (Zoraptera), Leu⁶ to Met⁶ (Dermaptera), Ser⁷ to Asn⁷, and Ser¹³ to Ala/Thr¹³ (Caelifera). Significant intraordinal variation is only present in Caelifera. The median AED for the CFR-DH precursor is as low as 0.15. Mantophasmatodea and Grylloblattodea show the lowest CT-DH precursor variation, while Dermaptera possess the most variable CT-DH precursor sequences. The overall AED for the CT-DH precursors is 0.44, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.05).

Elevenin (Additional File 3L)

An elevenin precursor with a length of 99-170 AA was found in all polyneopteran orders. For a single species, *Paratemnopteryx coulouiana* (Blattodea), we have identified a second and sequence-related precursor. The elevenin precursors contain quite variable elevenin motifs without C-terminal amidation site. The elevenin sequence in the precursor follows immediately the signal peptide and terminates upstream of a dibasic KR cleavage site (RKR or KKR in some Caelifera). In Caelifera, the predicted elevenin sequence contains a Lys²Arg³ motif that might be used as cleavage site. However, as noted above for CCHamide-2, KR motifs that immediately follow the signal peptide sequence do not necessarily function as cleavage signals for prohormone convertases.

The elevenins of Polyneoptera are variable in length and consist of 17 (Embioptera) up to 22 AA (Dermaptera, multiple species of Blattodea). Due to considerable sequence variations (**Figure 5**), an ancestral sequence of elevenin for Polyneoptera cannot be determined. Most polyneopteran taxa have species with a consensus C-terminus of CRGVAA-OH (CRGASA-OH in Mantophasmatodea) and a conserved position of the two

Cys residues; specific features also found in *T. domestica*. Significant intraordinal variation is present in Dermaptera, Caelifera, Embioptera, Phasmatodea, and Blattodea. The median AED for the elevenin precursor is 0.42. Grylloblattodea and Mantophasmatodea show the lowest elevenin precursor variation, while Caelifera possess the most variable elevenin precursor sequences. The overall AED for the elevenin precursors is 0.87, the actual neuropeptide sequence is better conserved (overall AED: 0.45).

HanSolin (Additional File 3M)

HanSolin was recently described from *C. morosus* (2). A subsequent search for HanSolin in Coleoptera (25) also revealed orthologous precursors in these holometabolous insects. Here we found a single HanSolin precursor with a length of 88-139 AA in all polyneopteran orders. The HanSolin precursors contain quite variable HanSolin motifs with a conserved C-terminus; including a C-terminal amidation site. HanSolin is always located C-terminal in the precursor, N-terminally mostly flanked by a monobasic Arg cleavage site (RR in Mantophasmatodea, multiple Caelifera and Ensifera), and terminates upstream of a dibasic C-terminal RR cleavage site (occasionally KR or monobasic R for different orders).

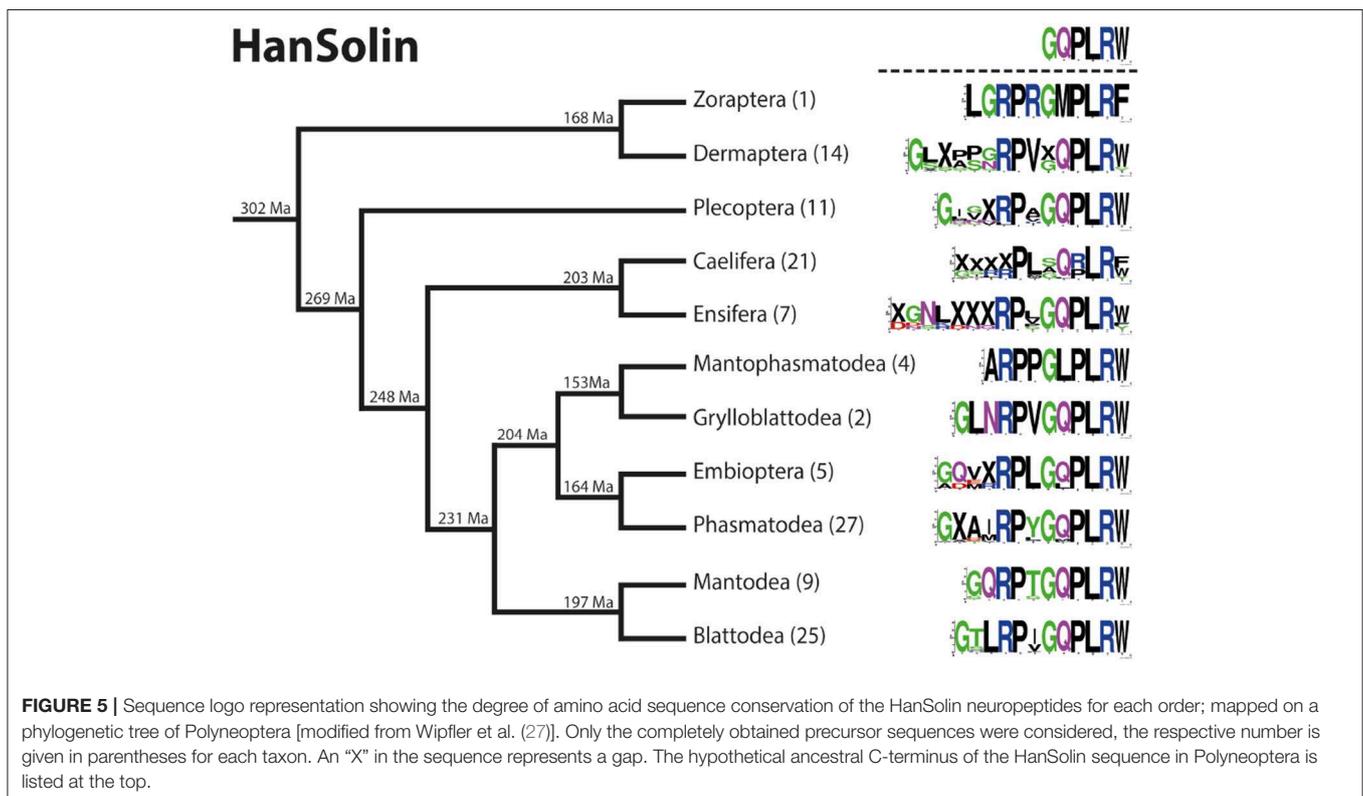
The HanSolins of Polyneoptera seem to be very variable in length and consist of 8 (Caelifera) up to 16 AA (Ensifera). In many cases, the N-terminal sequences of the predicted mature peptides require biochemical confirmation. Due to considerable sequence variations, an ancestral sequence of

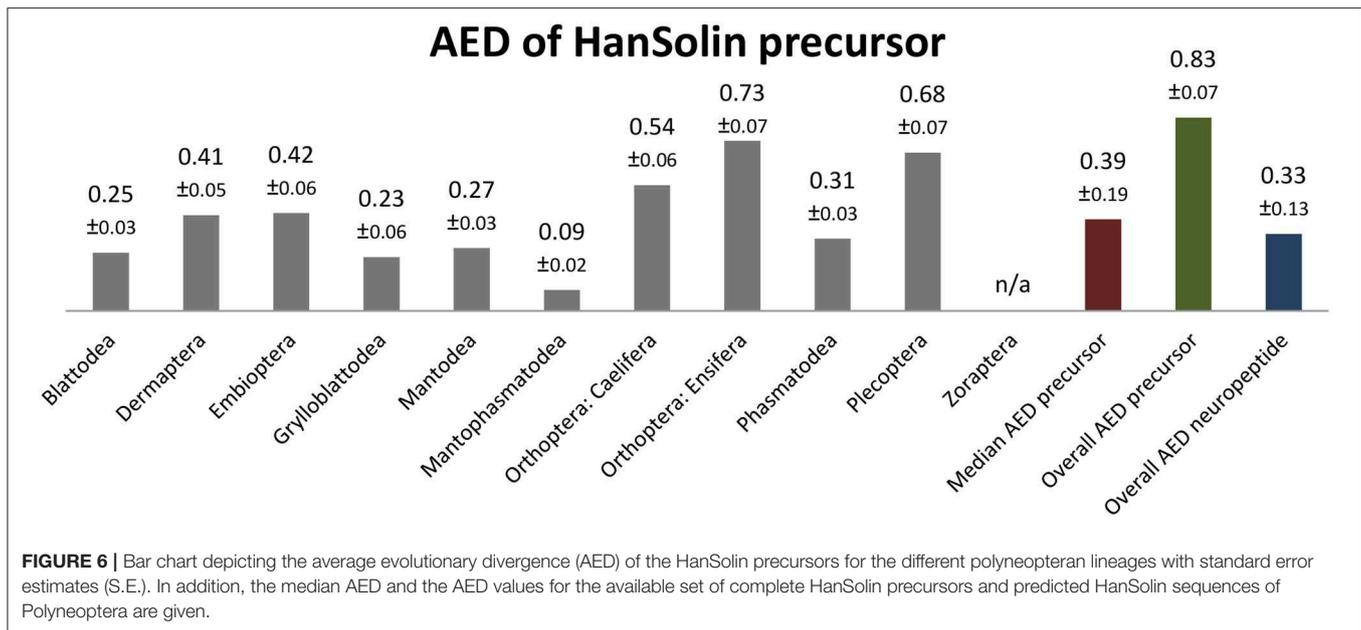
HanSolin for Polyneoptera cannot be determined. Most polyneopteran taxa have species with a consensus C-terminal hexapeptide of GQPLRW-NH₂ (GMPLRF-NH₂ in Zoraptera, GLPLRW-NH₂ in Mantophasmatodea); this C-terminus is also found in *T. domestica* (Bläser and Predel, unpublished). Significant intraordinal variation is present in Dermaptera, Plecoptera, Caelifera, Ensifera, Embioptera, and Phasmatodea. The median AED for the HanSolin precursor is 0.39 (Figure 6). Mantophasmatodea show the lowest HanSolin precursor variation, while Ensifera and Plecoptera possess the most variable HanSolin precursor sequences. The overall AED for the HanSolin precursors is 0.83, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.33).

MS (Additional File 3N)

A myosuppressin (MS) precursor with a length of 143-174 AA is present in all polyneopteran orders. All of these precursors contain a highly conserved MS motif with a C-terminal amidation site. For three species, *L. migratoria*, *Prosarthria teretirostris* (both Caelifera) and *Hemimerus* sp. (Dermaptera), we identified a second precursor with related MS sequences. We also found a second transcript with identical MS sequences in five species of Mantodea, but with an insertion of 39 AA in the middle part of the precursor. The MS sequence is always located C-terminal in the precursor, N-terminally flanked by a dibasic KR cleavage site and terminates upstream of an RRR cleavage motif.

MS sequences are almost exclusively decapeptides. Only in one species of Blattodea (*Reticulitermes santonensis*) MS





is a highly derived undecapeptide (KEDSQHMFLRF-NH₂). The sequence (p)QDVDHVFLRF-NH₂, which also occurs in Remipedia, was found in most orders of Polyneoptera (not in Caelifera and Dermoptera). This sequence might therefore be regarded as ancestral for all Hexapoda. Peptidomics confirmed for different polyneopteran taxa that the N-terminal Gln of MS is only partially converted to pGlu [e.g., (2, 40)]. With the exception of the MS of *R. santonensis* (see above) AA substitutions are restricted to the N-terminal AA (P/T in Caelifera, H in several Ensifera) and the position 6 (Val⁶ to Ile⁶ in Dermoptera). The second precursor of few species contains additional AA substitutions (EDVGHVFLRF-NH₂ in *L. migratoria*; KDIEHVFLRF-NH₂ in *P. teretirostris*, QDVHHNFLRF-NH₂ in *Hemimerus* sp.). The median AED for the MS precursor is 0.26. Grylloblattodea and Mantophasmatodea show the lowest MS precursor variation in Polyneoptera, while Ensifera possess the most variable MS precursors. The overall AED for the MS precursors is 0.57, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.05).

NPF-1 (Additional File 30)

The insect neuropeptide F-1 (NPF-1) precursor with a length of 81–97 AA was found in all polyneopteran orders. In most taxa we have also identified a second and longer transcript showing an insertion of about 40 AA in the middle of the NPF-1 neuropeptide (= NPF-1_b). Such transcripts are also known from *T. domestica* and are therefore a basic feature of Pterygota. The remaining sequences of NPF-1 are identical in both transcripts. The NPF-1 precursors contain a well-conserved NPF-1 motif with C-terminal amidation site. To a slightly lesser extent, this also applies to the insertion in the long transcript. The NPF-1 sequence in the precursor follows immediately the signal peptide and terminates upstream of a dibasic KR cleavage site.

Most NPF-1_a neuropeptides of Polyneoptera consist of 33 AA, but the full range is 30–36 AA. The C-terminus LQQLDRYYSQVARPRF-NH₂ is fully conserved in the majority of Polyneoptera (E to M/R in Embioptera and Mantophasmatodea; V to N/K in Plecoptera). Particularly significant intraordinal variation of the N-terminus is present in Dermoptera, Plecoptera, Caelifera, Grylloblattodea, and Embioptera. The median AED for the CFR-DH precursor is 0.27. Mantophasmatodea and Mantodea show the lowest CT-DH precursor variation, while Ensifera and Plecoptera possess the most variable CT-DH precursor sequences. The overall AED for the CT-DH precursors is 0.45, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.19).

NPF-2 (Additional File 3P)

An insect neuropeptide F-2 (NPF-2) precursor with a length of 85–134 AA was found in all polyneopteran orders. The NPF-2 precursors contain a quite variable NPF-2 motif with a well-conserved C-terminus and C-terminal amidation site. The NPF-2 sequence in the precursor follows immediately the signal peptide and terminates upstream of a dibasic KR cleavage site.

The NPF-2 neuropeptides of Polyneoptera are variable in length and consist of 43–47 AA. Only the C-terminus PRF-NH₂ is fully conserved in all analyzed species while a C-terminal RPRF-NH₂ was at least found in members of all order of Polyneoptera. Thus, the information about the C-terminal AA is not sufficient to distinguish between the neuropeptides NPF-1 and NPF-2. Particularly significant intraordinal variation of the N-terminus is present in Dermoptera, Plecoptera, Ensifera, and Embioptera. The median AED for the NPF-2 precursor is 0.29. Mantophasmatodea and Mantodea show the lowest NPF-2 precursor variation, while Plecoptera possess the most variable NPF-2 precursor sequences. The overall AED for the NPF-2

precursors is 0.65, the actual neuropeptide sequence is slightly better conserved (overall AED: 0.41).

Proctolin (Additional File 3Q)

A proctolin precursor with a length of 74–104 AA is present in almost all polyneopteran orders. The only exception was found in Dermaptera, where the proctolin precursor is absent. For the single species of Zoraptera (*Zorotypus caudelli*) we identified a second precursor. All these precursors contain a highly conserved proctolin motif without a C-terminal amidation site. The proctolin sequence immediately follows the signal peptide and terminates upstream of a monobasic Arg cleavage site.

Proctolin sequences are exclusively pentapeptides. The sequence RYLPT-OH, which also occurs in Remipedia and even Myriapoda, was found in all orders of Polyneoptera. This sequence might therefore be regarded as ancestral for Hexapoda. With the exception of the proctolin of *Systella rafflesii* (Pro⁴ to His⁴) and *Zubovskia* sp. (Thr⁵ to Val⁵; both Caelifera), all species possess the original sequence RYPLT-OH. The median AED for the proctolin precursor is 0.31. Grylloblattodea, Mantophasmatodea, and Mantodea show the lowest proctolin precursor variation in Polyneoptera, while Ensifera and Caelifera possess the most variable proctolin precursors. The overall AED for the ACP precursors is 0.52, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.00).

RFLamide (Additional File 3R)

RFLamides were only recently described from *C. morosus* (2). An RFLamide precursor with a length of 122–211 AA is present in all polyneopteran orders. Most precursors contain a well-conserved RFLamide with a C-terminal amidation site. The RFLamide sequence is always located C-terminal in the precursor, N-terminally flanked by an RKR or RRR cleavage site (dibasic RR in *Protonemura ausonia*, Plecoptera) and terminates upstream of quite variable cleavage motifs (monobasic Arg up to 5 basic AA which terminate the precursor sequence).

RFLamides are mostly duodecapeptides. Only in Mantophasmatodea and Grylloblattodea the RFLamides are 14 mers with an extended C-terminus. In these taxa the first Arg of the original cleavage motif is replaced by a Met, resulting in RFLamides with two additional AA without a C-terminal amidation site. The unique C-terminus of these insects is a remarkable synapomorphy of Mantophasmatodea and Grylloblattodea. The sequence PASAIFTNIRFL-NH₂ was found in most orders of Polyneoptera (not in Zoraptera, Mantophasmatodea, Grylloblattodea). This sequence might therefore be regarded as ancestral for all Polyneoptera (Figure 7). Amino acid substitutions in RFLamides of Polyneoptera are largely limited to substitutions of Ser³, Ala⁴, and Ile⁵. Mantophasmatodea and Grylloblattodea have the most derived sequences, each with several lineage-specific features; in addition to the distinct C-terminus, which is identical in both taxa. Significant intraordinal variation is present in Plecoptera, Ensifera, and Embioptera. The median AED for the RFLamide precursor is 0.28 (Figure 8). Grylloblattodea (only 2 species) show the lowest RFLamide precursor variation in Polyneoptera, while Plecoptera possess the most variable RFLamide precursors.

The overall AED for the RFLamide precursors is 0.62, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.14).

SIFamide (Additional File 3S)

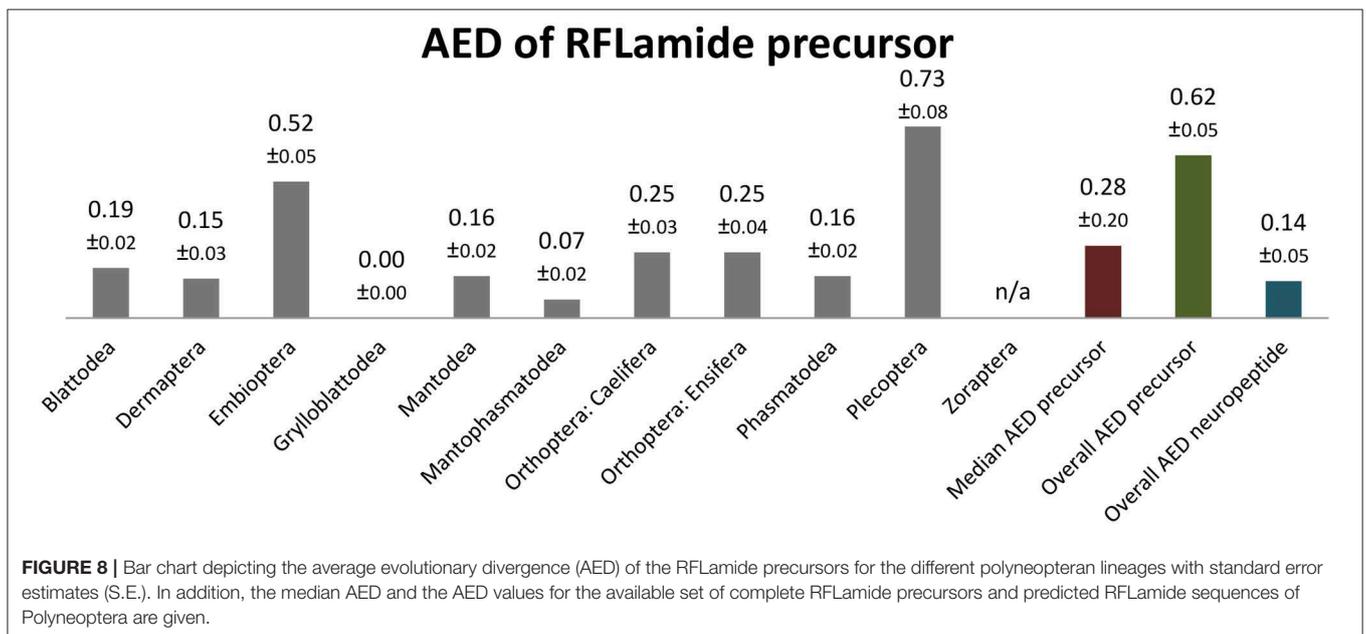
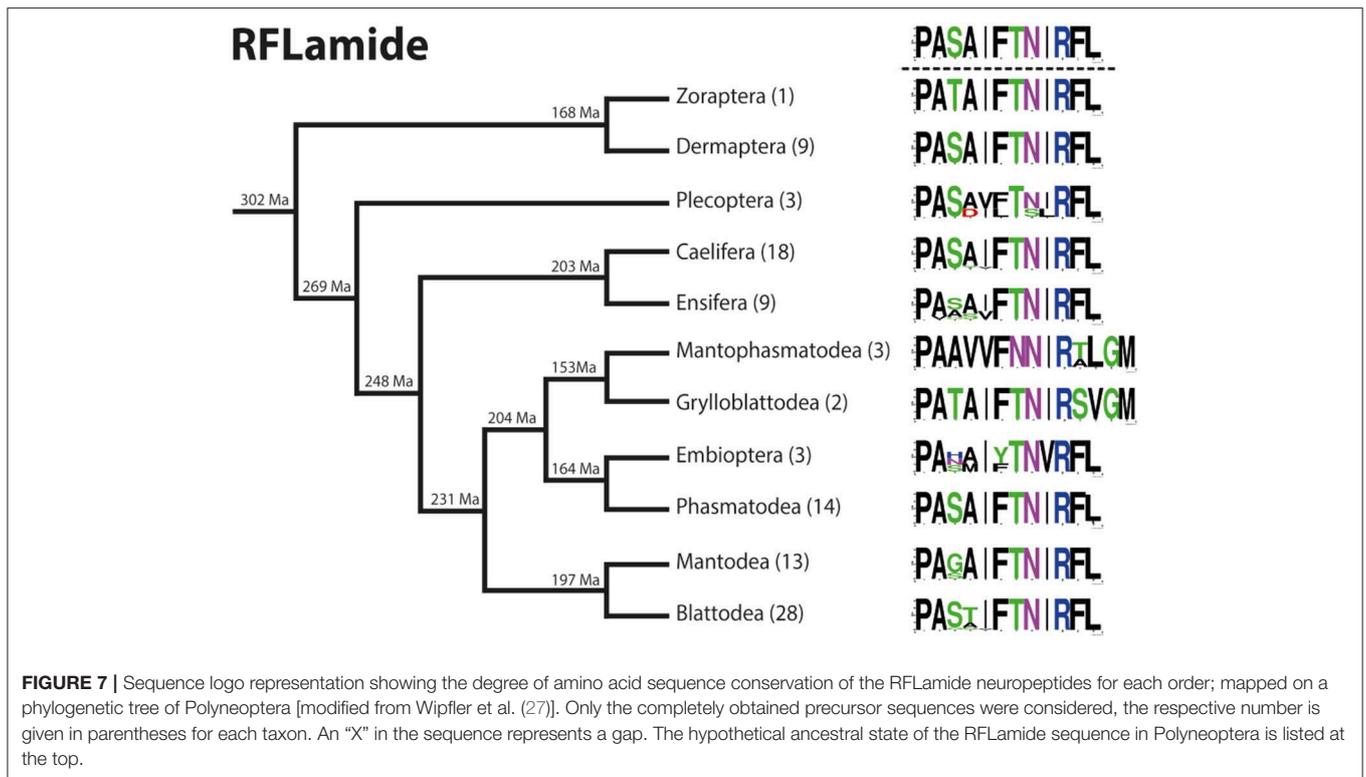
A SIFamide precursor with a length of 71–103 AA is present in all polyneopteran orders. In two blattodean species (*Prorhinotermes* and *Schultesia*) we found a second SIFamide precursor with identical SIFamide sequences; probably slightly different alleles. Most precursors contain a very well-conserved SIFamide with a C-terminal amidation site. An exception was found in *R. virgo* (Embioptera), which has three consecutive copies of SIFamide in the precursor and therefore cannot not be treated as a single-copy peptide. The SIFamides (actually SIYamides) of this species do indeed have more derived sequences and are not considered in our analyses. Whether the transition to multiple copies is a specific feature of Embioptera cannot yet be determined, since we could not find SIFamide precursors in the other Embioptera species. The SIFamide sequence in the precursors of all other species follow immediately the signal peptide and terminates upstream of a dibasic KR cleavage site.

The SIFamides are mostly duodecapeptides or longer. Only in the SIFamides of Dermaptera the N-terminal AA is missing (= undecapeptides). Since the length of the signal peptide of SIFamide precursors cannot always be predicted with certainty (41), the N-terminus of SIFamides should be confirmed by peptidomics in taxa not yet examined. The sequence TYRKPPFNGSIF-NH₂ was found in most orders of Polyneoptera (not in Dermaptera). This sequence might therefore be regarded as ancestral for all Polyneoptera. In the SIFamide of *T. domestica* (Zygentoma) only the N-terminal AA is different (Thr¹-Gly¹). Amino acid substitutions in SIFamides of Polyneoptera are largely limited to substitutions of Thr¹ and Tyr². The median AED for the SIFamide precursor is 0.22. Grylloblattodea show the lowest SIFamide precursor variation in Polyneoptera, while Caelifera possess the most variable SIFamide precursors. The overall AED for the SIFamide precursors is 0.5, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.09).

A second gene coding for a related neuropeptide, SMYamide, was described for *L. migratoria* (Caelifera) and *Zootermopsis nevadensis* [Blattodea; (1)]. We found SMYamide precursors in Caelifera, Ensifera, Embioptera, Phasmatodea, Mantodea, and Blattodea. The phylogenetic position of these taxa suggests that the SMYamide gene has evolved within the Polyneoptera. This is corroborated by the fact that so far no orthologs of SMY genes have been reported from any other insect.

sNPF (Additional File 3T)

A short neuropeptide F (sNPF) precursor with a length of 86–134 AA is present in all polyneopteran orders. All of these precursors contain a highly conserved sNPF motif with a C-terminal amidation site. A specific feature of most Caelifera is the presence of a second and longer sNPF neuropeptide immediately after the first sNPF sequence in the precursor. While the two species of *Xya* (Caelifera) still show the original pattern with a single sNPF sequence, *P. teretrirostris* (Caelifera) even has three



consecutive sNPFs in the precursor. The sequence of the N-terminal sNPF of the Caelifera with multiple sNPFs is highly conserved and closely resembles the orthologous sNPFs of the other polyneopteran orders. Therefore, it was included in our analyses. The sNPF sequence is always located in the middle of the precursor, N-terminally flanked by a dibasic RK cleavage site, and terminates upstream of a dibasic RR cleavage site.

Short NPF sequences (in Caelifera only the N-terminal sNPF sequence) are exclusively undecapeptides with a potential secondary cleavage site (Arg³). The sequence SNRSPSLRLRF-NH₂ which also occurs in *T. domestica*, was found in several orders of Polyneoptera (Zoraptera, Dermaptera, Plecoptera, Caelifera and Ensifera). This sequence might therefore be regarded as ancestral for all Pterygota. Apparently the

sister group of Zoraptera + Dermaptera (i.e., the remaining polyneopteran orders) originally had two alleles coding for Ser or Ala as the N-terminal AA. Several orders of this group (Plecoptera, Caelifera, Ensifera, Grylloblattodea, Phasmatodea) still have species either with Ser¹ or Ala¹, while in Mantophasmatodea, Embioptera, Mantodea, and Blattodea the sNPF with Ala¹ has completely replaced the original Ser at this position. Other AA substitutions are restricted to the second AA (Gln² to Ser²) and have been detected in a few Grylloblattodea and Caelifera and in all Embioptera. The median AED for the sNPF precursor is 0.20. Grylloblattodea and Mantophasmatodea show the lowest sNPF precursor variation in Polyneoptera, while Embioptera and Plecoptera possess the most variable sNPF precursors. The overall AED for the sNPF precursors is 0.44, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.05).

Trissin (Additional File 3U)

A trissin precursor with a length of 88–117 AA is present in almost all polyneopteran orders. The only exceptions were found in Dermaptera and Zoraptera, where trissin precursors are absent. For two species of the genus *Xya* (Caelifera), we have identified two trissin precursors with moderately (trissin 1) or strongly modified N-termini (trissin 2). Otherwise, the trissin precursors contain usually a well-conserved trissin motif without C-terminal amidation site. The trissin sequence in the precursor follows immediately after the signal peptide and terminates upstream of a tribasic RKR cleavage site (KKR in 2 of 26 Mantodea species and dibasic KR in 2 of 6 Ensifera species).

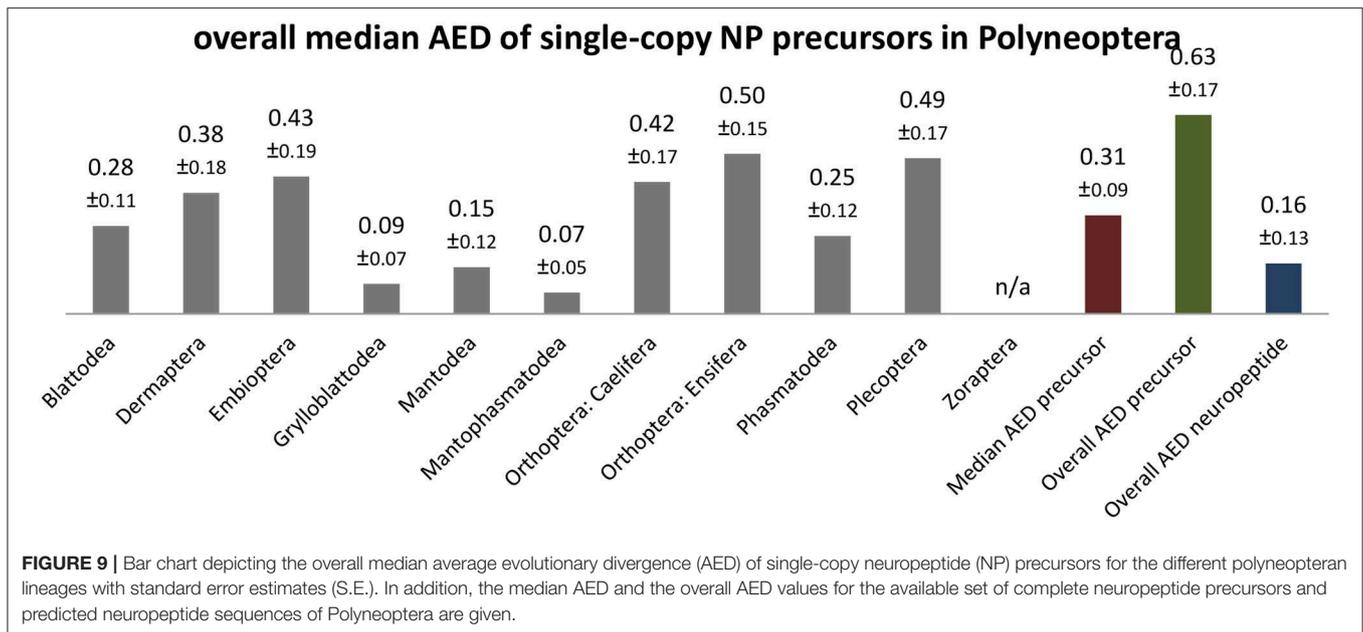
Most trissins of Polyneoptera consist of 27 AA. For most Caelifera a truncated sequence of trissin with a single AA (instead of two) preceding the N-terminal Cys is predicted by SignalP-5.0; trissin 1 of *Xya* probably starts directly with the N-terminal Cys. Trissins 1 of *Xya variegata* and *X. japonica* additionally show an insertion of Ser downstream of the N-terminal Cys. In Ensifera, the N-terminal of trissin is not always clearly predicted; probably it starts directly with the N-terminal Cys. Generally, the N-terminal cleavage of trissins should be confirmed by peptidomics. However, trissin has not been detected biochemically from any polyneopteran species so far. In two species of Ensifera (*Ceuthophilus* sp. and *Diestrammena asynamora*) the first Arg of the C-terminal cleavage motif is replaced by Ser, which probably leads to an extended C-terminus (NYLS-OH instead of NYL-OH). All species of the ensiferan infraorder Gryllidea whose trissin sequence has been identified (*Ceuthophilus* sp., *Gryllotalpa* sp., *Neonetus* sp.) show an insertion of Asp in the middle of the sequence, indicating a synapomorphy. The sequence LSCDSCGRECXXCGRNFRNRTCCFNLYL-OH (XXX: no ancestral AA assigned) was found in *T. domestica* and most orders of Polyneoptera. This sequence might therefore be regarded as ancestral for all Pterygota. Amino acid substitutions in trissins of Polyneoptera are mainly limited to substitutions of AA at positions 11–13 and 18. Significant intraordinal variation is present in Plecoptera, Caelifera, Ensifera, and Blattodea. Distinct lineage-specific features are substitutions of Ser⁴ to Phe⁴/Val⁴ (Caelifera) or Ile⁴ (Ensifera), Phe²⁴ to Leu²⁴/Tyr²⁴ (Caelifera/Ensifera), and Arg²¹ to Val²¹/His²¹ (Ensifera). The

median AED for the trissin precursor is 0.23. Mantophasmatodea show the lowest trissin precursor variation, while Ensifera possess the most variable trissin precursor sequences. The overall AED for the trissin precursors is 0.59, the actual neuropeptide sequence is significantly better conserved (overall AED: 0.20).

CONCLUSIONS

In our analysis we examined the single-copy precursor sequences of 21 neuropeptide genes of Polyneoptera. The neuropeptides of 17 of these precursors are C-terminally amidated (not AST-CC, elevenin, proctolin, trissin), which prevents rapid degradation by exopeptidases and thus supports their functions as hormones. Only very few neuropeptide genes coding for single-copy neuropeptides are completely missing in a given polyneopteran order. Dermaptera have no ACP, proctolin and trissin, Ensifera do not have CCHamide-1, and in most Embioptera we could not detect any SIFamide precursor (with the exception of a multiple-copy SIFa precursor in *R. virgo*, see above). Furthermore, we did not find precursors for corazonin, CRF-DH and trissin in Zoraptera, but only one species of this order could be analyzed. Therefore, the absence of the respective neuropeptide genes has yet to be confirmed for Zoraptera. For most orders and also for the individual species within these orders, we have found all single-copy precursors, a feature already documented for the “basal” hexapods, which represent the sister group of the Pterygota [winged insects; (24)]. In contrast, peptide gene losses are more frequent in the much more species-rich and ecologically significant Holometabola. The fruit fly *Drosophila melanogaster*, which is used as a model organism in molecular biology, neurobiology and also physiology, is a good example in this context as it lacks not less than 6 of the 21 peptidergic systems analyzed here (ACP, AT, Elevenin, Hansolin, NPF-2, RFLamide) (12).

The sequence conservation of the precursor sequences, including the signal peptides, varies for the different neuropeptide genes. Low overall AED values (AST-C: 0.36; sNPF, CT-DH: 0.44; CCAP, NPF-1: 0.45; see **Additional File 3**) contrast with high AED values (CNMamide: 0.95; Elevenin: 0.87; Hansolin: 0.83; see **Additional File 3**), which are significantly above the average value of 0.63 calculated for all neuropeptide precursors (**Figure 9**). As expected, the sequences of single-copy neuropeptides within the precursors are much better conserved (overall AED 0.16; **Figure 9**). However, the extent of sequence conservation across Polyneoptera is remarkably different between the different neuropeptides. Neuropeptides such as proctolin, CCAP, AST-C, sNPF, MS, and CT-DH (overall AED ≤ 0.05) are almost identical in all taxa and the most common sequence always represents the predicted ancestral sequence of Pterygota (sNPF, CT-DH) or even the ancestral sequence of Hexapoda (proctolin, CCAP, AST-C, MS). For all neuropeptides with very high AED values (Elevenin: 0.45; NPF-2: 0.41; Hansolin: 0.33), the sequence ancestral to Polyneoptera could not be determined. Many of the neuropeptides with high AED values have long sequences, but this does not necessarily



lead to high AED values, as is shown for example with CT-DH (31 AA; AED: 0.05).

The overall median AEDs for the single-copy neuropeptides and precursors differ significantly between the polyneopteran orders. This was to be expected, since the different lineages evolved independently of each other over different periods of time (see **Figure 2**). In addition, several polyneopteran orders (e.g., Grylloblattodea and Mantophasmatodea) represent relict groups with only a few extant and rather closely related taxa. Thus, these orders show particularly low AEDs, while Dermaptera, Plecoptera and Orthoptera (Ensifera + Caelifera) have much higher intra-ordinal sequence diversity (**Figure 9**). The relatively high AEDs for Embioptera were somewhat unexpected in this context. Although the AEDs for the various neuropeptide precursors of the different Polyneoptera are mostly in the range of the median AED for all neuropeptide precursors, there are striking exceptions. This is especially true for Mantodea (significantly lower AEDs for NPF-1 and -2) and Dermaptera (significantly lower AED for AST-CC). A comparison of AEDs in the orthopteran sister groups Ensifera and Caelifera also shows very different AEDs for the different neuropeptide precursors, either in favor of Ensifera or Caelifera (**Additional File 3**). This means that in the evolution of the sequences of neuropeptide precursors there have been some striking increases or decreases in the AA substitution rate, which cannot be directly related to a uniform development of the peptidergic system of a given taxon or to a specific neuropeptide gene.

A number of derived neuropeptide sequences were found, showing sequence motifs (= synapomorphies) typical only for representatives of a specific polyneopteran lineage. This has to be separated from intra-ordinal variation. Within the respective lineages, the derived sequences are often well-conserved (**Additional File 3**). However, surprisingly few

examples of derived sequences have been found that are typical of two or more polyneopteran orders. One clear example is the substitution within the C-terminal cleavage motif of RLFamides, which probably occurred in the last common ancestor of Mantophasmatodea and Grylloblattodea. This substitution prevents the C-terminal amidation and is typical of all Mantophasmatodea and Grylloblattodea. Furthermore, the absence of trissin in both Dermaptera and Zoraptera (here, however, only a single transcriptome was available) indicates that the loss of this neuropeptide already occurred in the last common ancestor of these two lineages. Typical for most Dictyoptera (Mantodea + Blattodea) is Gln¹¹ of trissin, which is only found in this taxon.

Overall, the single-copy neuropeptide precursors of the Polyneoptera show a relatively high degree of sequence conservation. Basic features of these precursors in this very heterogeneous insect group are explained here in detail for the first time. Further insights into the evolution of neuropeptides can be expected from future analyzes of the much more variable multiple-copy neuropeptides.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MB and RP contributed to the conception and design of the study. MB mined the transcriptomes for neuropeptide precursors and wrote the first draft of the manuscript. All authors contributed to final version of the manuscript and approved the submitted manuscript.

FUNDING

This study was funded by the Deutsche Forschungsgemeinschaft (RP 766/11-1).

ACKNOWLEDGMENTS

We are grateful to the entire 1KITE consortium, especially to Bernhard Misof, Alexander Donath, Benjamin Wipfler (ZFMK Bonn, Germany) and Karen Meusemann (University of Freiburg,

Germany) for providing early access to the 1KITE transcriptome assemblies and raw data. Furthermore, we thank Sander Liessem and Lapo Ragionieri (University of Cologne, Germany) for supporting the transcriptome analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2020.00197/full#supplementary-material>

REFERENCES

- Veenstra JA. The contribution of the genomes of a termite and a locust to our understanding of insect neuropeptides and neurohormones. *Front Physiol.* (2014) 5:454. doi: 10.3389/fphys.2014.00454
- Liessem S, Ragionieri L, Neupert S, Büschges A, Predel R. Transcriptomic and neuropeptidomic analysis of the stick insect, *Carausius morosus*. *J Proteome Res.* (2018) 17:2192–204. doi: 10.1021/acs.jproteome.8b00155
- Bläser M, Misof B, Predel R. The power of neuropeptide precursor sequences to reveal phylogenetic relationships in insects: a case study on Blattodea. *Mol Phylogenet Evol.* (2020) 66:479–506. doi: 10.1016/j.ympev.2019.106686
- Elphick MR, Mirabeau O, Larhammar D. Evolution of neuropeptide signalling systems. *J Exp Biol.* (2018) 221:jeb151092. doi: 10.1242/jeb.151092
- Jékely G. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci USA.* (2013) 110:8702–7. doi: 10.1073/pnas.1221833110
- Mirabeau O, Joly J-S. Molecular evolution of peptidergic signaling systems in bilaterians. *Proc Natl Acad Sci USA.* (2013) 110:E2028–37. doi: 10.1073/pnas.1219956110
- Bauknecht P, Jékely G. Large-scale combinatorial deorphanization of platynereis neuropeptide GPCRs. *Cell Rep.* (2015) 12:684–93. doi: 10.1016/j.celrep.2015.06.052
- Nachman RJ, Holman GM, Haddon WF, Ling N. Leucosulfakinin, a sulfated insect neuropeptide with homology to gastrin and cholecystokinin. *Science.* (1986) 234:71–3. doi: 10.1126/science.3749893
- Blair JE, Hedges SB. Molecular phylogeny and divergence times of deuterostome animals. *Mol Biol Evol.* (2005) 22:2275–84. doi: 10.1093/molbev/msi225
- Simakov O, Kawashima T, Marlétaz F, Jenkins J, Koyanagi R, Mitros T, et al. Hemichordate genomes and deuterostome origins. *Nature.* (2015) 527:459–65. doi: 10.1038/nature16150
- Schoofs L, De Loof A, Van Hiel MB. Neuropeptides as regulators of behavior in insects. *Annu Rev Entomol.* (2017) 62:35–52. doi: 10.1146/annurev-ento-031616-035500
- Nässel DR, Zandawala M. Recent advances in neuropeptide signaling in *Drosophila*, from genes to physiology and behavior. *Prog Neurobiol.* (2019) 179:101607. doi: 10.1016/j.pneurobio.2019.02.003
- Li B, Predel R, Neupert S, Hauser F, Tanaka Y, Cazzamali G, et al. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. *Genome Res.* (2008) 18:113–22. doi: 10.1101/gr.6714008
- Ons S, Richter F, Urlaub H, Pomar RR. The neuropeptidome of *Rhodnius prolixus* brain. *Proteomics.* (2009) 9:788–92. doi: 10.1002/pmic.200800499
- Huybrechts J, Bonhomme J, Sebastian M, Prunier-Leterme N, Dombrovsky A, Abdel-Latif M, et al. Neuropeptide and neurohormone precursors in the pea aphid *Acyrtosiphon pisum*. *Insect Mol Biol.* (2010) 19(Suppl 2):87–95. doi: 10.1111/j.1365-2583.2009.00951.x
- Predel R, Neupert S, Garczyński SE, Crim JW, Brown MR, Russell WK, et al. Neuropeptidomics of the mosquito *Aedes aegypti*. *J Proteome Res.* (2010) 9:2006–15. doi: 10.1021/pr901187p
- Predel R, Neupert S, Derst C, Reinhardt K, Wegener C. Neuropeptidomics of the bed bug *Cimex lectularius*. *J Proteome Res.* (2018) 17:440–54. doi: 10.1021/acs.jproteome.7b00630
- Pandit AA, Ragionieri L, Marley R, Yeoh JGC, Inward DJG, Davies S-A, et al. Coordinated RNA-Seq and peptidomics identify neuropeptides and G-protein coupled receptors (GPCRs) in the large pine weevil *Hylobius abietis*, a major forestry pest. *Insect Biochem Mol Biol.* (2018) 101:94–107. doi: 10.1016/j.ibmb.2018.08.003
- Hummon AB, Richmond TA, Verleyen P, Baggerman G, Huybrechts J, Ewing MA, et al. From the genome to the proteome: uncovering peptides in the *Apis* brain. *Science.* (2006) 314:647–9. doi: 10.1126/science.1124128
- Hauser F, Neupert S, Williamson M, Predel R, Tanaka Y, Grimmelikhuijzen CJ. Genomics and peptidomics of neuropeptides and protein hormones present in the parasitic wasp *Nasonia vitripennis*. *J Proteome Res.* (2010) 9:5296–310. doi: 10.1021/pr100570j
- Ragionieri L, Predel R. The neuropeptidome of *Carabus* (Coleoptera, Adephaga: Carabidae). *Insect Biochem Mol Biol.* (2020) 20:103309. doi: 10.1016/j.ibmb.2019.103309
- Brown BE, Starratt AN. Isolation of proctolin, a myotropic peptide, from *Periplaneta americana*. *J Insect Phys.* (1975) 23:1879–81. doi: 10.1016/0022-1910(75)90257-7
- Yeoh JGC, Pandit AA, Zandawala M, Nässel DR, Davies SA, Dow JAT. DIneR: database for insect neuropeptide research. *Insect Biochem Mol Biol.* (2017) 86:9–19. doi: 10.1016/j.ibmb.2017.05.001
- Derst C, Dirksen H, Meusemann K, Zhou X, Liu S, Predel R. Evolution of neuropeptides in non-ptyerygote hexapods. *BMC Evol Biol.* (2016) 16:51. doi: 10.1186/s12862-016-0621-4
- Veenstra JA. Coleoptera genome and transcriptome sequences reveal numerous differences in neuropeptide signaling between species. *PeerJ.* (2019) 7:e7144. doi: 10.7717/peerj.7144
- Wegener C, Gorbashov A. Molecular evolution of neuropeptides in the genus *Drosophila*. *Genome Biol.* (2008) 9:R131. doi: 10.1186/gb-2008-9-8-r131
- Wipfler B, Letsch H, Frandsen PB, Kapli P, Mayer C, Bartel D, et al. Evolutionary history of Polyneoptera and its implications for our understanding of early winged insects. *Proc Natl Acad Sci USA.* (2019) 116:3024–9. doi: 10.1073/pnas.1817794116
- Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. *Science.* (2014) 346:763–67. doi: 10.1126/science.1257570
- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, Castro E, et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* (2012) 40:W597–603. doi: 10.1093/nar/gks400
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* (2013) 30:772–80. doi: 10.1093/molbev/mst010
- Hall TA. "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Sympos Seri.* (1999) 41:c1979–c2000.
- Kumar S, Stecher G, Li M, Nkay C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* (2018) 35:1547–9. doi: 10.1093/molbev/msy096
- Zuckerkindl E, Pauling L. Evolutionary divergence and convergence in proteins. In: *Edited In Evolving Genes and Proteins*. Bryson V, Vogel HJ, editors. New York, NY: Academic Press. (1965). p. 97–166. doi: 10.1016/B978-1-4832-2734-4.50017-6

34. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. *Genome Res.* (2004) 14:1188–90. doi: 10.1101/gr.849004
35. Simon S, Letsch H, Bank S, Buckley TR, Donath A, Liu S, et al. Old world and new world phasmatodea: phylogenomics resolve the evolutionary history of stick and leaf insects. *Front Ecol Evol.* (2019) 7:744. doi: 10.3389/fevo.2019.00345
36. Clynen E, Huybrechts J, Verleyen P, De Loof A, Schoofs L. Annotation of novel neuropeptide precursors in the migratory locust based on transcript screening of a public EST database and mass spectrometry. *BMC Genom.* (2006) 7:201. doi: 10.1186/1471-2164-7-201
37. Veenstra JA. Allatostatins C, double C and triple C, the result of a local gene triplication in an ancestral arthropod. *Gen Comp Endocr.* (2016) 230–231:153–7. doi: 10.1016/j.ygcen.2016.04.013
38. Predel R, Neupert S, Russell WK, Scheibner O, Nachman RJ. Corazonin in insects. *Peptides.* (2007) 28:3–10. doi: 10.1016/j.peptides.2006.10.011
39. Predel R, Neupert S, Huetteroth W, Kahnt J, Waidelich D, Roth S. Peptidomics-based phylogeny and biogeography of Mantophasmatodea (Hexapoda). *Syst Biol.* (2012) 61:609–29. doi: 10.1093/sysbio/sys003
40. Neupert S, Fusca D, Schachtner J, Kloppenburg P, Predel R. Toward a single-cell-based analysis of neuropeptide expression in *Periplaneta americana* antennal lobe neurons. *J Comp Neurol.* (2012) 520:694–716. doi: 10.1002/cne.22745
41. Gellerer A, Franke A, Neupert S, Predel R, Zhou X, Liu S, et al. Identification and distribution of SIFamide in the nervous system of the desert locust *Schistocerca gregaria*. *J Comp Neurol.* (2015) 523:108–25. doi: 10.1002/cne.23671

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Bläser and Predel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Analysis of Pigment-Dispersing Factor Neuropeptides and Their Receptor in a Velvet Worm

Christine Martin^{1†}, Lars Hering^{1†}, Niklas Metzendorf¹, Sarah Hormann¹, Sonja Kasten¹, Sonja Fuhrmann¹, Achim Werckenthin², Friedrich W. Herberg³, Monika Stengl² and Georg Mayer^{1*}

¹ Department of Zoology, Institute of Biology, University of Kassel, Kassel, Germany, ² Department of Animal Physiology, Institute of Biology, University of Kassel, Kassel, Germany, ³ Department of Biochemistry, Institute of Biology, University of Kassel, Kassel, Germany

OPEN ACCESS

Edited by:

Elizabeth Amy Williams,
University of Exeter, United Kingdom

Reviewed by:

Bo Joakim Eriksson,
University of Vienna, Austria
Dusan Zitnan,
Slovak Academy of Sciences, Slovakia

*Correspondence:

Georg Mayer
georg.mayer@uni-kassel.de

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 17 February 2020

Accepted: 14 April 2020

Published: 12 May 2020

Citation:

Martin C, Hering L, Metzendorf N,
Hormann S, Kasten S, Fuhrmann S,
Werckenthin A, Herberg FW, Stengl M
and Mayer G (2020) Analysis of
Pigment-Dispersing Factor
Neuropeptides and Their Receptor in
a Velvet Worm.
Front. Endocrinol. 11:273.
doi: 10.3389/fendo.2020.00273

Pigment-dispersing factor neuropeptides (PDFs) occur in a wide range of protostomes including ecdysozoans (= molting animals) and lophotrochozoans (mollusks, annelids, flatworms, and allies). Studies in insects revealed that PDFs play a role as coupling factors of circadian pacemaker cells, thereby controlling rest-activity rhythms. While the last common ancestor of protostomes most likely possessed only one *pdf* gene, two *pdf* homologs, *pdf-I* and *pdf-II*, might have been present in the last common ancestors of Ecdysozoa and Panarthropoda (Onychophora + Tardigrada + Arthropoda). One of these homologs, however, was subsequently lost in the tardigrade and arthropod lineages followed by independent duplications of *pdf-I* in tardigrades and decapod crustaceans. Due to the ancestral set of two *pdf* genes, the study of PDFs and their receptor (PDFR) in Onychophora might reveal the ancient organization and function of the PDF/PDFR system in panarthropods. Therefore, we deorphanized the PDF receptor and generated specific antibodies to localize the two PDF peptides and their receptor in the onychophoran *Euperipatoides rowelli*. We further conducted bioluminescence resonance energy transfer (BRET) experiments on cultured human cells (HEK293T) using an Epac-based sensor (Epac-L) to examine cAMP responses in transfected cells and to reveal potential differences in the interaction of PDF-I and PDF-II with PDFR from *E. rowelli*. These data show that PDF-II has a tenfold higher potency than PDF-I as an activating ligand. Double immunolabeling revealed that both peptides are co-expressed in *E. rowelli* but their respective levels of expression differ between specific cells: some neurons express the same amount of both peptides, while others exhibit higher levels of either PDF-I or PDF-II. The detection of the onychophoran PDF receptor in cells that additionally express the two PDF peptides suggests autoreception, whereas spatial separation of PDFR- and PDF-expressing cells supports hormonal release of PDF into the hemolymph. This suggests a dual role of PDF peptides—as hormones and as neurotransmitters/neuromodulators—in Onychophora.

Keywords: PDF, PDFR, BRET, Epac, *E. rowelli*, Onychophora, Panarthropoda, Ecdysozoa

INTRODUCTION

Pigment-dispersing factors (PDFs) and pigment-dispersing hormones (PDHs) are neuropeptides that occur in various protostomes and have been mainly studied in nematodes, crustaceans and insects [reviewed in (1, 2)]. In decapod crustaceans, PDH controls a circadian relocation of the retinal pigment to shield the photoreceptor cells from light and also orchestrates circadian migrations of pigment within the integumental chromatophores (3–6). The crustacean PDHs, however, seem to not only play a hormonal role in mediating pigment dispersion but might additionally act as neurotransmitters or neuromodulators in the nervous system (7–12). The same holds true for the insect PDFs—homologs of the crustacean PDHs (4, 13). PDFs are synthesized by the circadian clock neurons in various insect species and are involved in different aspects of circadian timing (1, 14–34).

The PDFs of the fruit fly *Drosophila melanogaster* and the cockroach *Rhyarobia maderae* are analogs of the vasoactive intestinal peptide (VIP) of mammals (35). It must be noted, however, that despite their similar function in the circadian pacemaker systems, the insect PDFs (and crustacean PDHs) do not share a common ancestry with VIP. The ancestral *pdf* gene rather evolved in the protostome lineage, as it occurs in both spiralian and ecdysozoans but not in deuterostomes (36–38). The report of potential PDF precursors in echinoderms and an enteropneust (39) should be taken with caution due to the low sequence similarity with the PDFs/PDHs of insects and crustaceans [cf. Figure 7 in (39)] and the lack of a phylogenetic analysis. Interestingly, while the last common ancestor of protostomes possessed only one *pdf* gene, which has been retained at least in mollusks and annelids [(36, 37); cf. supporting information Supplementary Figures 1 and 2 in (38)], a duplication of this gene might have occurred in the ecdysozoan lineage (Figure 1). Two homologs, *pdf-I* and *pdf-II*, have been identified thus far in priapulids, nematodes and onychophorans (38, 41, 42), whereas *pdf-II* seems to have been subsequently lost in tardigrades and arthropods. Even more intriguing is the independent duplication of the retained *pdf-I* gene in tardigrades, which show three in-paralogs, and decapod crustaceans, which express two to three PDH isoforms but seem to possess only two *pdh* genes (11, 38, 43–48).

Since two PDF peptides were most likely encoded in the genome of the last common ancestor of Ecdysozoa (Figure 1), the question arises of whether their ancestral role was in light-dependent pigment dispersion, in coupling of circadian clocks or in other processes. Alternatively, these peptides might have played divergent or multiple roles. Research on nematodes revealed that their two *pdf* genes express three PDFs that most likely also have an impact on circadian timing (49). There, the same peptides serve several functions including the control of locomotion, mate searching, mechano- and chemosensation as well as the sensation of oxygen (41, 50). Accordingly, in the nematode nervous system PDFs occur in nearly all sensory neurons that also control mate searching and locomotor activity patterns (42). Insects instead possess fewer PDF-expressing cells organized in 2–4 clusters associated with each compound eye

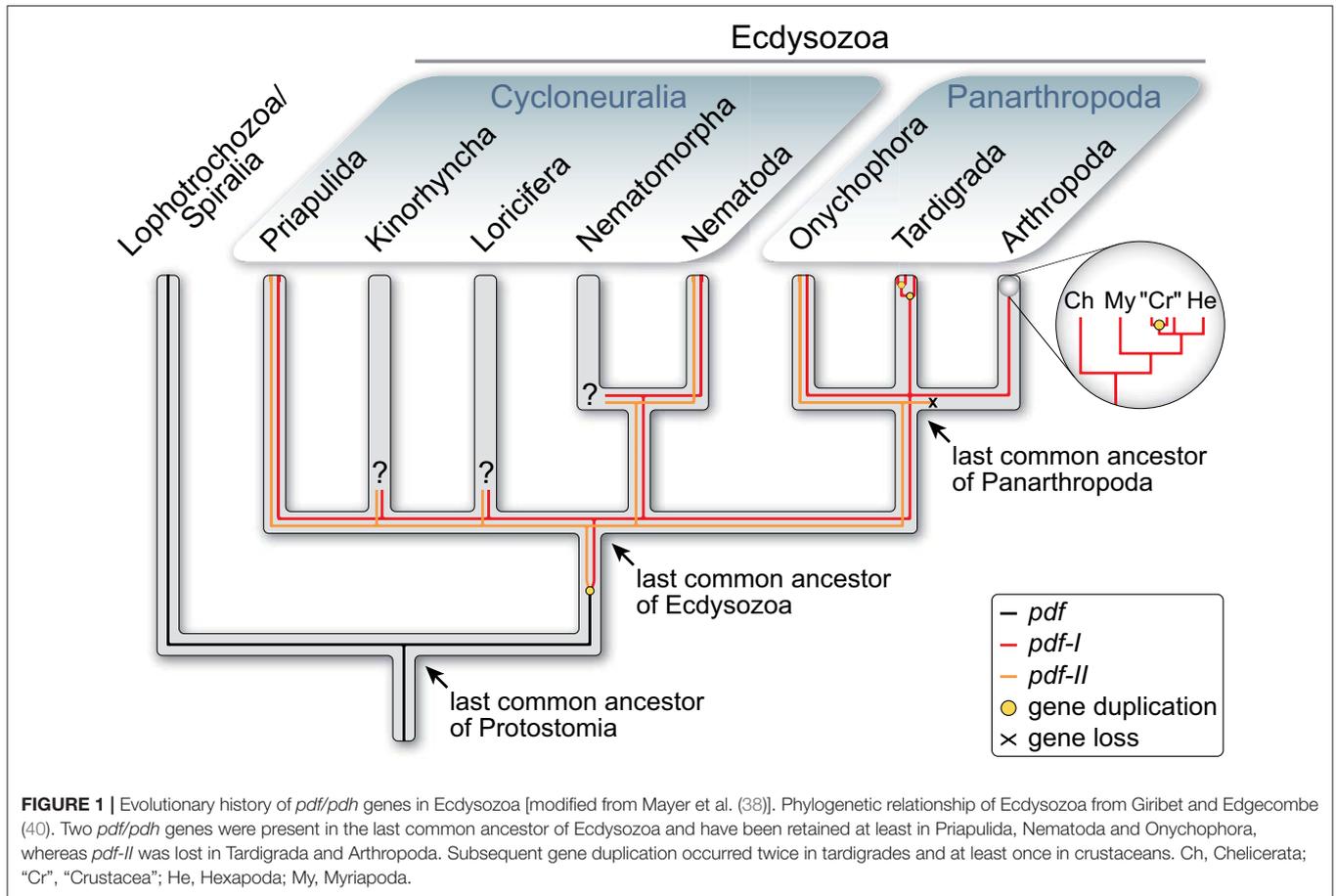
and few additional somata showing a species-specific distribution in the protocerebrum [e.g., (14–16, 22, 23, 25, 29, 30, 51–56)]. The situation in crustaceans seems to be more complex, but they also possess PDH-immunoreactive somata in the median protocerebrum, in addition to those associated with the eye stalks (7, 8, 11).

Similar to crustaceans, Mayer et al. (38) detected numerous immunoreactive somata in the median protocerebrum of six distantly related onychophoran species using a broadly reactive antiserum raised against the synthetic β -PDH peptide of the crustacean *Uca pugilator* (7). These and other results indicate a broad and elaborate distribution of PDFs in the peripheral and central nervous system of onychophorans. However, since the applied antiserum recognized both onychophoran peptides, Ony-PDF-I and Ony-PDF-II, despite its higher affinity to Ony-PDF-I [83.3% sequence similarity as opposed by 55.6–61.1% to Ony-PDF-II; cf. Figure 1 in (38)], it remains unclear whether these two peptides are co-expressed or show distinct distributions in various tissues and cells. To clarify this issue, we obtained specific antibodies against each PDF peptide of the onychophoran *Euperipatoides rowelli* (Peripatopsidae) and performed double immunolabeling. The gathered data provide insights into the ancestral distribution and potential roles of the two PDF peptides in the last common ancestors of Panarthropoda and Ecdysozoa.

Another important goal of our study was to deorphanize and immunolocalize the PDF receptor (PDFR) of onychophorans. Investigations of PDFRs are generally scarce. These membrane proteins belong to the ancient family of G protein-coupled receptors (GPCRs) and have a similar function to the VPAC₂ receptors in the circadian system of mammals [reviewed in (35)]. However, while PDFRs are activated by PDF/PDH peptides, the VPAC₂ receptors rather bind the aforementioned vasoactive intestinal peptide (VIP). Despite their important functions, only little is known about the origin and evolutionary history of PDFRs and VPAC₂ receptors.

The PDFR was initially deorphanized in the fruit fly *D. melanogaster* by three simultaneous studies: one of them examined a range of GPCRs for their sensitivity to PDF (57) and the other two reported that mutants lacking the *groom-of-PDF* (*gop*) or the *han* gene (referred to as *pdfr* thereafter) imitate phenotypical behavior of *pdf* null mutants (58, 59). Subsequently, PDFR was shown to be expressed in 60% of pacemaker neurons (60), a few additional, clock-unrelated neurons as well as within the visual system of the fruit fly (61). Like *D. melanogaster*, the nematode *Caenorhabditis elegans* possesses one *pdfr* gene but three PDFR splice variants (41), all of which show a dose-dependent sensitivity to the three nematode PDFs (PDF-1a, PDF-1b, and PDF-2). The nematode PDFRs have been localized in the motor neurons of the nervous system as well as in nearly all muscle cells (41). Apart from *D. melanogaster* and *C. elegans*, to our knowledge there are no immunohistochemical data on the local distribution of PDFRs in other animals.

To improve our understanding of the PDF/PDFR system in onychophorans, we screened the transcriptome of *E. rowelli* (62) for GPCR genes and identified a single *pdfr* candidate.



We further performed functional analyses using synthetic PDF-I and PDF-II peptides (based on sequences from *E. rowelli*) in conjunction with bioluminescence resonance energy transfer (BRET) to confirm the identity of the putative receptor. Finally, we generated a specific antibody against the newly detected PDFR of *E. rowelli* and localized this receptor protein in the animal. The obtained data provide insights into the onychophoran PDF/PDFR network and contribute to a better understanding of its evolution in panarthropods and ecdysozoans.

MATERIALS AND METHODS

Collection and Maintenance of Specimens

Specimens of *Euperipatoides rowelli* Reid, 1996 (63) were obtained from decaying logs and leaf litter in the Tallaganda State Forest (New South Wales, Australia; 35°30'31"S, 149°36'14.3"E, 934 m) in October 2016. They were collected under the permit number SPPR0008 issued by the Forestry Commission of New South Wales and exported under the permit number WT2012-8163 provided by the Department of Sustainability, Environment, Water, Population and Communities. The collected specimens were maintained in the laboratory as described previously (64).

Transcriptome *de novo* Assembly and Identification of Pigment-Dispersing Factor Receptor Genes

Illumina short reads from two specimens (male and female) of *E. rowelli*, sequenced as part of the i5K initiative (65, 66), were obtained from the short read archive (SRA) of GenBank (ER9 female: accession number SRR1946792; ER10 male: accession number SRR1946791). Prior to the assembly step, sequencing adapters were trimmed using cutadapt v1.8.1 (67) (parameters: $-\text{max-n} = 0 -\text{u} 10 -\text{u-5} -\text{U} 10 -\text{U-5}$) and short reads with more than five bases below a Phred quality score of 15 were removed using ConDeTri v2.2 (68) ($-\text{hq} = 14 -\text{lq} = 10 -\text{sc} = 33 -\text{frac} = 0.95 -\text{minlen} = 85$). Both data sets were then assembled *de novo* using the software IDBA-Tran v1.1.1 (69) ($-\text{mink} 20 -\text{maxk} 85 -\text{step} 5 -\text{max_isoforms} 1 -\text{min_contig} 100$) yielding 92,426 unique contigs for ER9 (mean contig length = 717, N50 = 851) and 59,131 for ER10 (mean contig length = 707, N50 = 817), respectively. Both assemblies are available upon request.

To obtain the sequences of putative *pdfr* genes from the transcriptome of *E. rowelli*, two bait sequences of known pigment-dispersing factor receptors from the kuruma shrimp *Marsupenaeus japonicus* (accession number BAH85843.1) and the fruit fly *D. melanogaster* (NP_570007.2) were used in tBLASTn v2.4.0+ searches (70) against three assemblies in total

[ER9 and ER10 from this study; Filter15 assembly from Hering et al. (62)] yielding 27 candidate sequences with an E value less than $1e^{-5}$ as a threshold.

Cluster Analysis and Phylogenetic Reconstruction

After screening for putative pigment-dispersing factor receptor genes in *E. rowelli* transcriptomes, a cluster analysis of the 27 candidate sequences together with 18,337 bilaterian G protein-coupled receptors (GPCRs) obtained from the GPCR database (71) was performed using CLANS (72). In short, all sequences were clustered based on pairwise sequence similarities obtained from all-vs.-all BLAST searches. Because the best BLAST hit does not necessarily imply the closest relationship (73), a phylogenetic tree was reconstructed afterwards from all formerly obtained class B GPCRs, to which the PDFRs most likely belong (57, 59).

For phylogenetic analyses, the transmembrane domains of all class B GPCRs obtained from the cluster analysis [993 in total, mainly hormone receptors, including PDF receptors; reviewed in (74, 75)] were predicted using Pfam-A v29 (76, 77) and aligned using the MAFFT online version v7.452 (78) with the most accurate option L-INS-i and default parameters. To remove homoplastic and random-like positions, the alignment was masked with the software Noisy rel. 1.15.12 (79) (-seqtype = P -shuffles = 10,000) prior to the analyses. Two Maximum likelihood analyses were conducted with the Pthreads version of RAxML v8.2.10 (80). For each run, the best tree was obtained from a combined analysis (-f a option) of 10 independent inferences and GAMMA correction of the final tree under either the empirical JTT substitution model or a dataset-specific GTR substitution matrix, including calculation of bootstrap support values from 100 pseudoreplicates using the rapid bootstrapping algorithm implemented in RAxML. The JTT model was automatically selected by RAxML (PROTGAMMAAUTO option) as best-fitting substitution model. Notably, the analysis using the dataset-specific GTR+G model yielded a better log likelihood score (-87,233.49) for the best tree than using the best obtained empirical model JTT+G (-88,041.12). The phylogenetic tree was visualized with iTol v2 (81) and edited with Adobe Illustrator CS 5.1 (Adobe Systems Incorporated, San Jose, CA, USA).

Amplification of Gene Fragments and Directional Cloning

The whole coding sequence (CDS) of the identified *Er-pdf* gene was amplified from previously synthesized cDNA of *E. rowelli* [see (82)] using gene-specific primers (*Er_pdfr_HindIII_F*: 5'-tcgaagcttgccaccATGTGGTATTTAATTTATTGTATTTTTCCTTC-3'; *Er_pdfr_XhoI_R*: 5'-tgcgctcgagCTAAATACACAATTCCTCCAACAC-3'), which contained restriction sites (HindIII, XhoI) required for subsequent directional cloning into the mammalian expression vector pcDNA3.1(+) (Invitrogen, Carlsbad, CA, USA) to generate the plasmid pcDNA3.1-*ErPDFR*-3K10. The plasmid DNA was purified from 100 mL of transformed *E. coli* bacterial cell culture using the PureYield™ Plasmid Midiprep System (Promega GmbH,

Walldorf, Germany) including an endotoxin removal step. The sequence and orientation of the cloned insert was verified by Sanger sequencing (Eurofins Genomics GmbH, Ebersberg, Germany) and deposited in GenBank under the accession number *Er-pdf* MT080366.

Antibody Generation

Polyclonal antibodies against *Er*-PDF-I and *Er*-PDF-II were newly generated by using HPLC-purified synthetic *Er*-PDF-I (CNAELINSLGLPKMMNDA-NH₂) and *Er*-PDF-II peptides (CNAELINSLNLPKQLQEA-NH₂; peptides&elephants GmbH (Hennigsdorf, Germany). Prior to immunization in rabbits, the peptides were coupled via an additional cysteine at the N-terminus to Sulfo-SMCC-activated bovine thyroglobulin as carrier. The anti-*Er*-PDF antibodies (anti-*Er*-PDF-I: IG-P1035, LOT# 2078B, 14 μg/mL; anti-*Er*-PDF-II: IG-P1036, LOT# 2079B2, 21 μg/mL) were purified from sera of immunized rabbits by two-fold depletion of one antibody against the antigen of the other and subsequent affinity purification on its own antigen to minimize cross-reactivity (ImmunoGlobe GmbH, Himmelstadt, Germany).

A HPLC-purified synthetic peptide from the C-terminal region of *Er*-PDFR (Ac-LRDHQGGQLSEDRRDC-NH₂, Schafer-N ApS, Copenhagen, Denmark) was coupled to keyhole limpet hemocyanin (KLH) and used to newly generate a polyclonal antibody (anti-*Er*-PDFR; IG-P1044, LOT# 2130-5.6, 25 μg/mL), which was affinity purified from blood serum of an immunized goat (ImmunoGlobe GmbH).

BRET Assay

HEK293T cells were seeded on 96-well Nunc plates (Thermo Fisher Scientific, Waltham, MA, USA) at a density of 2×10^4 cells/well and cultured in DMEM high glucose (Sigma-Aldrich Chemie GmbH, Munich, Germany) supplemented with 10% fetal bovine serum and 5% CO₂ at 37°C. After 24 h, cells were transfected using polyethyleneimine (25 kDa PEI, linear; Polysciences Europe GmbH, Hirschberg an der Bergstrasse, Germany) with GFP2-hEpac1E157-P881-Rluc8 [Epac-L; modified from (83–85)] and pcDNA3.1-*ErPDFR*-3K10 in a 1:1 ratio (100 ng total DNA/well). Cells were washed with HBSS (Thermo Fisher Scientific) 2 days after transfection and stimulated with different concentrations of synthetic PDF peptides (10^{-11} to 10^{-5} M final concentration) together with the substrate coelenterazine 400A ("DeepBlueC"; BIOTREND Chemikalien GmbH, Cologne, Germany) at a final concentration of 5 μM in a total volume of 50 μL HBSS. Control experiments were conducted with cells expressing Epac-L alone, stimulation without PDF peptides, stimulation using an unrelated peptide (myoinhibitory peptide 7 from the cockroach *Rhyarobia maderae*), and stimulation with forskolin (50 μM final concentration; Sigma-Aldrich Chemie GmbH) + IBMX (100 μM final concentration; Sigma-Aldrich Chemie GmbH), respectively. The raw luminescence was measured immediately after stimulation at wavelengths of 410 ± 80 nm for the donor and 515 ± 30 nm for the acceptor in eight biological replicates for each tested PDF peptide concentration ($n = 8$) using a POLARstar Omega microplate reader (BMG Labtech, Cary, NC,

USA). Their ratios [emission (410 nm)/emission (515 nm)] were box plotted as relative luminescence against the respective PDF concentration in dose response curves.

Western Blots and Specificity Tests

Western blots for goat anti-Er-PDF-I and goat anti-Er-PDF-II were performed using synthetic Er-PDF-I and Er-PDF-II peptides. Western blots for goat anti-Er-PDFR were conducted using synthetic peptide coupled to bovine serum albumin (BSA) as well as tissue lysate from dissected heads. For lysate preparation, five specimens of *E. rowelli* were anesthetized in chloroform vapor and dissected in cold phosphate-buffered saline (PBS; 0.1 M, pH 7.4). Tissues of onychophoran heads were homogenized on ice in a solution containing 500 μ L NP-40 lysis buffer (pH 8.0) and the protease inhibitor cOmplete™ Mini (Roche Diagnostics GmbH, Mannheim, Germany). Tissue was incubated in lysis buffer for 2 h at 4°C. Subsequently, the lysed tissue was centrifuged and the supernatant stored at -80°C.

Synthetic Er-PDF-I and Er-PDF-II peptides were diluted in 4x Laemmli sample buffer (pH 6.8), heated for 5 min at 95°C and applied to an acrylamide Tris-tricine gel (10%) for 30 min at 30 V and 120 min at 50 V and then transferred to a Roti®-pvd 0.2 membrane (Carl Roth GmbH & Co. KG, Karlsruhe, Germany) via semi-dry blot for 120 min at 84 mA. The synthetic Er-PDFR peptides were diluted in 2x Laemmli sample buffer, heated for 5 min at 95°C and, together with the proteins of the lysate (lysed 1:1 in 2x Laemmli sample buffer), separated using SDS-PAGE on a 10% gel for 15 min at 30 V and 40 min at 200 V and then transferred to a Porablot NCP nitrocellulose membrane (Macherey-Nagel GmbH & Co. KG, Düren, Germany) via semi-dry blot for 60 min at 150 mA. Membranes were blocked for 30 min using 4% powdered milk in PBS at room temperature and then incubated with either rabbit anti-Er-PDF-I and rabbit anti-Er-PDF-II (1 μ g/mL in PBS each) or goat anti-Er-PDFR (1 μ g/mL in PBS) overnight at 4°C. Following several washing steps with PBS, membranes were incubated with either goat anti-rabbit or donkey anti-goat antibodies, respectively, conjugated with alkaline phosphatase (1:500; dianova GmbH, Hamburg, Germany) and washed again in PBS. The signal was developed using a solution containing 175 μ g/mL BCIP (5-bromo-4-chloro-3-indolyl phosphate; Thermo Fisher Scientific) in dimethylformamide and the reaction was stopped with PBS.

In order to further test the specificity of the generated antibodies, additional western blots were conducted as described above with the following changes: Er-PDF-I antibody was tested against the Er-PDF-II peptide and, vice versa, the Er-PDF-II antibody was tested against the Er-PDF-I peptide.

Whole-Mount Preparation, Vibratome Sectioning, and Immunohistochemistry

For immunohistochemistry, specimens were anesthetized in chloroform vapor and fixed according to Stefanini et al. (86) in 4% paraformaldehyde (PFA) containing 7.5% picric acid in phosphate buffered saline (PBS) overnight. Specimens were washed in several rinses of PBS, cut in halves and embedded in 31.25% albumin from chicken egg (Sigma-Aldrich Chemie

GmbH) and 4.17% gelatin from porcine skin (Type A; Sigma-Aldrich Chemie GmbH) in distilled water. Albumin-gelatin blocks were hardened for 4 h at 4°C and then fixed in 10% PFA overnight. After several rinses in PBS, the blocks were cut in 80–100 μ m thick sections using a vibratome (MICROM 650 V; Microm International GmbH, part of Thermo Fisher Scientific, Walldorf, Germany). The sections used for anti PDFR-labeling were treated with a mixture of collagenase/dispace (Roche Diagnostics GmbH; 1 mg/mL each) and hyaluronidase (Sigma-Aldrich Chemie GmbH; 1 mg/mL) in PBS for 40 min at 37°C and subsequently washed in PBS. Sections were blocked either in 10% normal goat serum (NGS; Sigma-Aldrich Chemie GmbH) for anti-Er-PDF-I and anti-Er-PDF-II labeling or in 10% bovine albumin serum (BSA; Carl Roth GmbH & Co. KG) for anti-Er-PDFR labeling in PBS containing 1% Triton X-100 (PBS-Tx, Sigma-Aldrich Co) for 1.5 h, followed by an incubation with the primary antibodies, either (i) rabbit anti-Er-PDF-I, (ii) rabbit anti-Er-PDF-II or (iii) goat anti-Er-PDFR (0.2 μ g/mL in PBS-Tx, 1% NGS) for 2 days (rabbit anti-Er-PDF-I and rabbit anti-Er-PDF-II) or 5 days (goat anti-Er-PDFR) at 4°C. After several changes of PBS for 8–24 h, sections were incubated with either goat anti-rabbit Alexa Fluor® 488 (1:500; Thermo Fisher Scientific) or donkey anti-goat Alexa Fluor® 488 or Alexa Fluor® 680 (1:500; Thermo Fisher Scientific) for 2 days at 4°C. After washing steps in PBS-Tx, sections were further washed using PBS alone and then incubated in a solution containing 4',6-diamidino-2-phenylindole (DAPI; 1 ng/mL; Carl Roth GmbH & Co. KG) and phalloidin-rhodamine (25 μ g/mL in PBS; Thermo Fisher Scientific) for 2 h at room temperature. After rinsing the sections several times in PBS they were mounted between two coverslips in ProLong™ Gold Antifade Reagent (Invitrogen).

For double labeling of both PDF peptides, the procedure was the same as for single labeling with the following changes: the sections were incubated in 10% BSA in PBS-Tx for 1.5 h. Thereafter, the sections were first incubated with the rabbit anti-Er-PDF-II (0.2 μ g/mL in PBS-Tx, 1% BSA) at 4°C for 2 days. After several changes of PBS-Tx for 12 h, the sections were then incubated in a solution containing the fab fragments goat anti-rabbit (Jackson Immuno Research Europe Ltd., Cambridgeshire, UK; 40 μ g/mL in PBS-Tx, 1% BSA) for 2 days at 4°C. After changes of PBS-Tx for 8–24 h, the sections were incubated in donkey anti-goat Alexa Fluor® 594 (1:500, Thermo Fischer Scientific). Thereafter, the sections were incubated with rabbit anti-Er-PDF-I (0.2 μ g/mL in PBS-Tx, 1% BSA) at 4°C for 2 days, followed by an incubation in goat anti-rabbit Alexa Fluor® 488 (1:500; Thermo Fischer Scientific) and counterstained with DAPI (Carl Roth GmbH & Co. KG).

For double labeling with anti-Er-PDFR and either rabbit anti-Er-PDF-I or rabbit anti-Er-PDF-II, the procedure was as described for single labeling with the following changes: the incubations with the primary antibodies anti-Er-PDFR with either rabbit anti-Er-PDF-I or rabbit anti-Er-PDF-II (0.2 μ g/mL in PBS-Tx) as well as the secondary antibodies donkey anti-goat Alexa Fluor® 488 and donkey anti-rabbit® 568 (1:500; Thermo Fisher Scientific) were performed simultaneously for 5 days at 4°C.

Additional double labeling of both PDF peptides was performed on whole mount-preparations of the brain. Specimens of *E. rowelli* were first anesthetized in chloroform vapor and their brains were dissected in physiological saline (87). Following the fixation after Stefanini et al. (86) for 3 h at room temperature, the brains were washed in PBS, dehydrated through an ascending ethanol series (70, 90, 95, 100, and 100%, 10 min each), cleared in xylene for 10 min twice, rehydrated through a descending ethanol series (100, 100, 95, 90, 70 and 50%, 5 min each) and washed several times in PBS. The samples were then treated with a mixture of collagenase/dispase (Roche Diagnostics GmbH; 1 mg/mL each) and hyaluronidase (Sigma-Aldrich Chemie GmbH; 1 mg/mL) in PBS for 40 min at 37°C and washed in several rinses of PBS. The staining procedure was as described for sections with the following modifications: the incubation times of the primary antibodies, the Fab fragments and the secondary antibodies were increased to 4 days. After the incubation with the secondary antibodies, the samples were washed in PBS-Tx, dehydrated through an ascending ethanol series (70, 90, 95, 100, and 100%, 10 min each), cleared and mounted between two coverslips in methyl salicylate using custom-made carbon fiber-reinforced polymer microscope slides as spacers.

Confocal Laser Scanning Microscopy and Image Processing

Whole mount preparations and vibratome sections were analyzed using a confocal laser scanning microscope (Zeiss LSM 880; Carl Zeiss Microscopy GmbH, Jena, Germany) equipped with an Airyscan module. Images were acquired and raw Airyscan datasets processed using the ZEN 2 (Black edition) imaging software (Carl Zeiss Microscopy GmbH). Selected substacks were created and adjusted to optimal brightness and contrast using Fiji v.1.52 (88, 89). Panels and diagrams were designed using Illustrator CS5.1 and the video was processed with Adobe Premiere Pro CS5.1 (Adobe Systems Incorporated).

RESULTS

Identification of the Onychophoran PDF Receptor

The initial BLAST search for PDF receptor homologs in three transcriptomes of *E. rowelli* yielded 27 candidate sequences. We included these in subsequent phylogenetic analyses of known bilaterian GPCRs to obtain the onychophoran PDFR ortholog. After prescreening the most likely candidate sequences in an all-vs.-all BLAST cluster analysis (Figure 2A), a maximum likelihood tree was reconstructed using the formerly obtained class B GPCR members, to which PDF receptors most likely belong (57, 59). Using this approach, we identified one PDF receptor gene (*Er-pdfr*) in each analyzed transcriptome of *E. rowelli*; only in one of the transcriptomes (ER10) the respective transcript was fractured in two contigs (transcripts 10,426 and 39,667; see Supplementary Figures 1, 2). The detected *Er-pdfr* sequence clearly falls into the monophyletic group of protostome PDFRs (Figure 2B; Supplementary Figures 1, 2).

In vitro Analysis of the Functionality of the Identified PDF Receptor

To test the functionality of the recognized Er-PDF receptor, we further performed bioluminescence resonance energy transfer (BRET) experiments on cultured human cells (HEK293T) using an Epac-based cAMP-sensor (Epac-L). The cAMP response data revealed that the transfected Er-PDFR is activated by both PDF peptides of *E. rowelli*, Er-PDF-I and Er-PDF-II, in a dose-dependent manner (Figures 3A,B). In contrast, the myoinhibitory peptide Rm-MIP7, which is co-localized with PDF in circadian pacemaker cells of the cockroach *Rhyparobia maderae* [see (91)] and which we used as a control, did not stimulate the receptor of *E. rowelli*. The minimum concentrations required for PDFR stimulation differs between the two PDFs: while Er-PDF-I showed an activation threshold at 10^{-6} M, the receptor sensitivity to Er-PDF-II was lower by one order of magnitude at 10^{-7} M (Figures 3A,B).

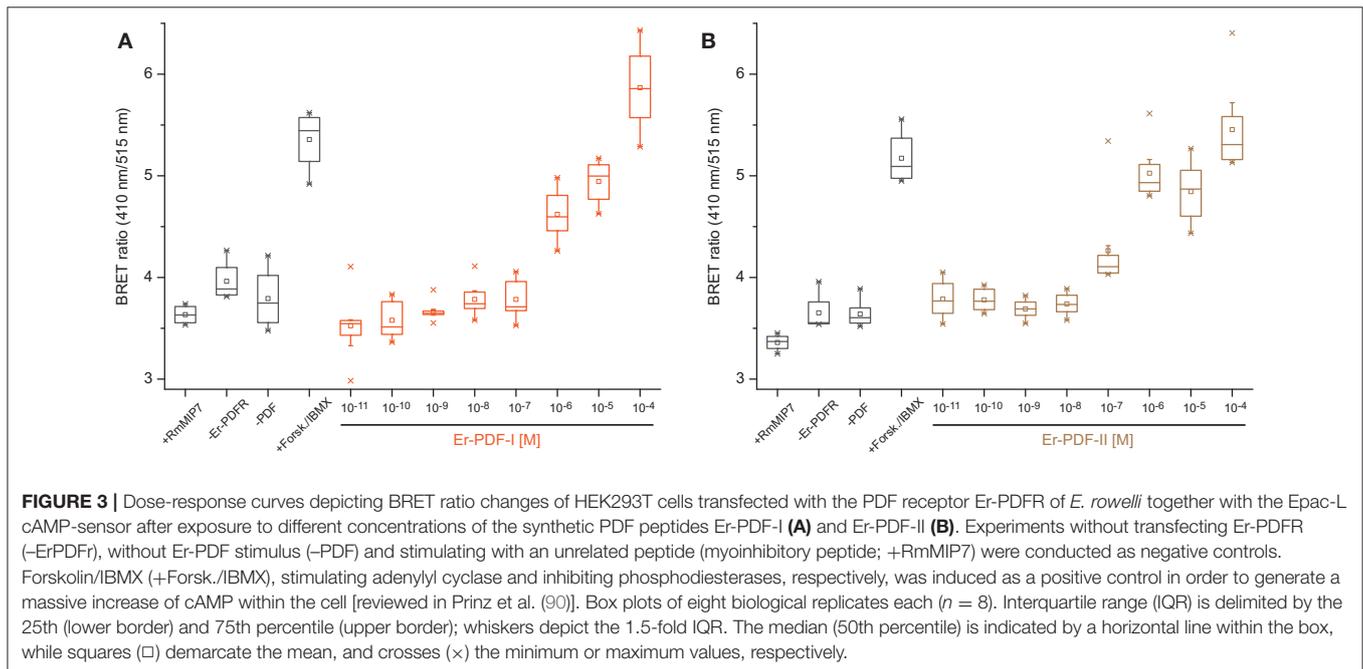
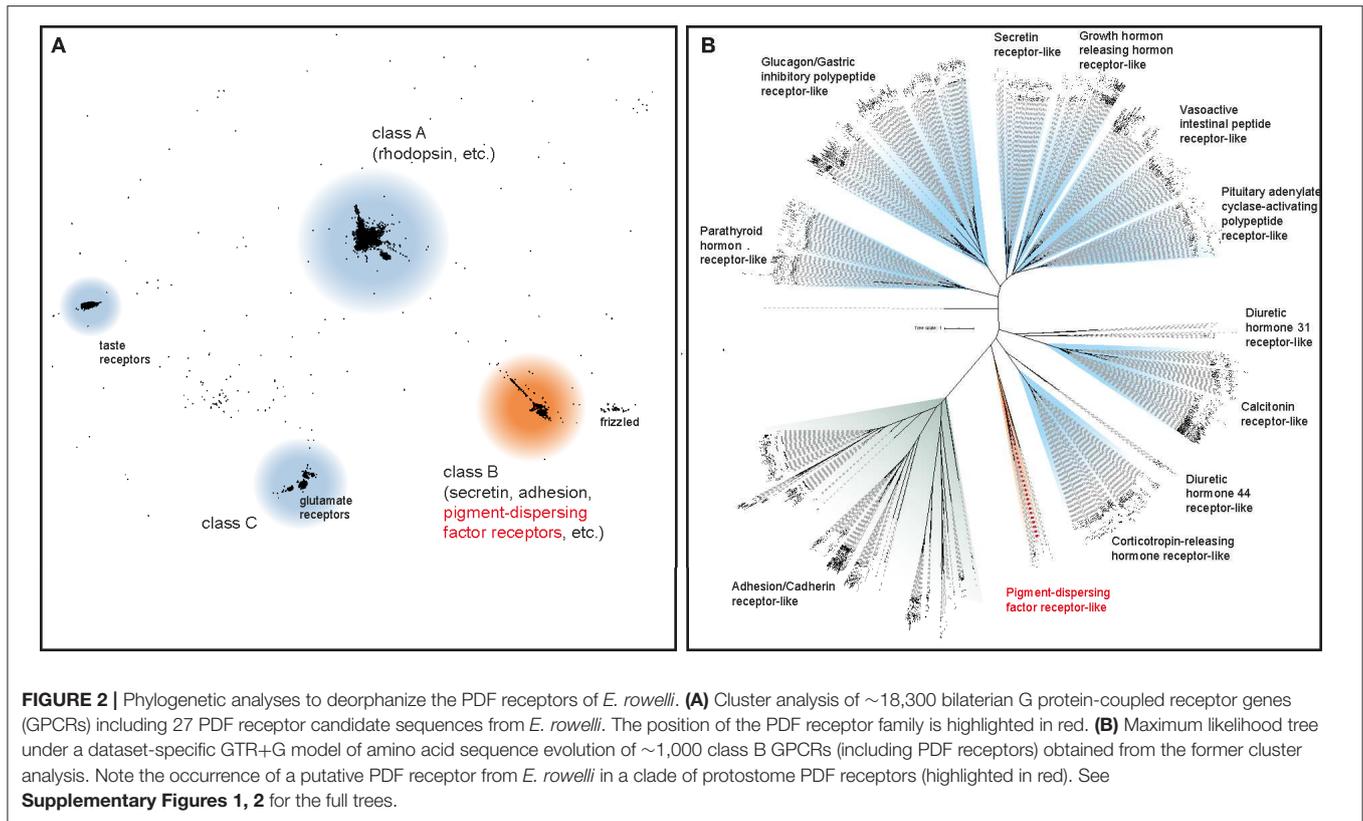
Specificity Tests of Er-PDF-I, Er-PDF-II, and Er-PDFR Antibodies

To test the specificity of the generated Er-PDF-I and Er-PDF-II antibodies, we performed western blot analyses using different concentrations of synthetic peptides (Figures 4A–D). The western blots show a distinct band below 4.6 kDa for the Er-PDF-I antibody (tested on 500 ng of synthetic Er-PDF-I peptide) and the Er-PDF-II antibody (using 50 ng and 500 ng of synthetic Er-PDF-II peptide), respectively (Figures 4A,D). The control experiments reciprocally using the non-corresponding peptides revealed no signal (Figures 4B,C).

The specificity of the Er-PDFR antibody was tested using both the synthetic peptide as well as lysates of entire heads of *E. rowelli* at different dilutions (Figure 5). The western blot employing the synthetic PDFR peptide coupled to BSA exhibits a distinct band at ~70 kDa and an additional band above 100 kDa. The western blots using different dilutions of lysates show distinct bands at ~70 kDa and a considerably weaker band at ~120 kDa in all three lanes (Figure 5). The lane carrying 15 μ l of lysate exhibits an additional weak band at ~250 kDa.

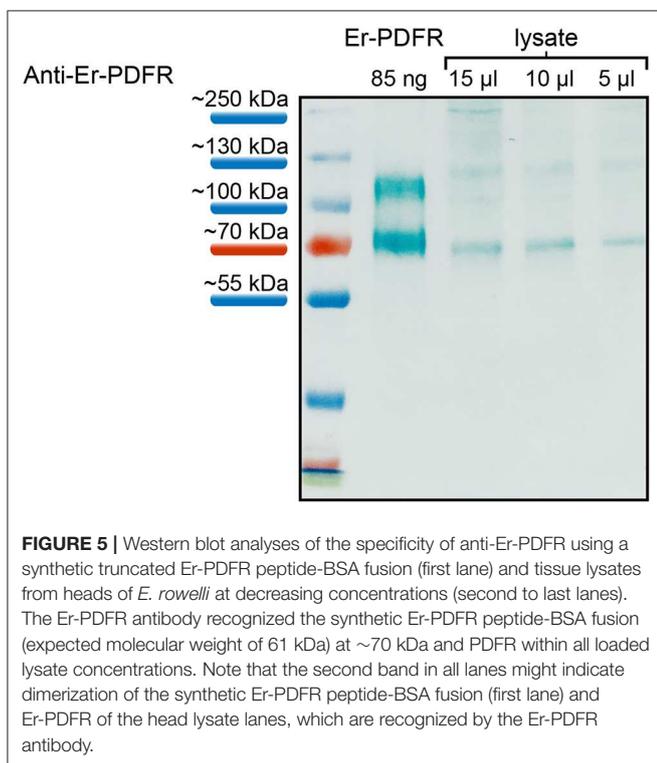
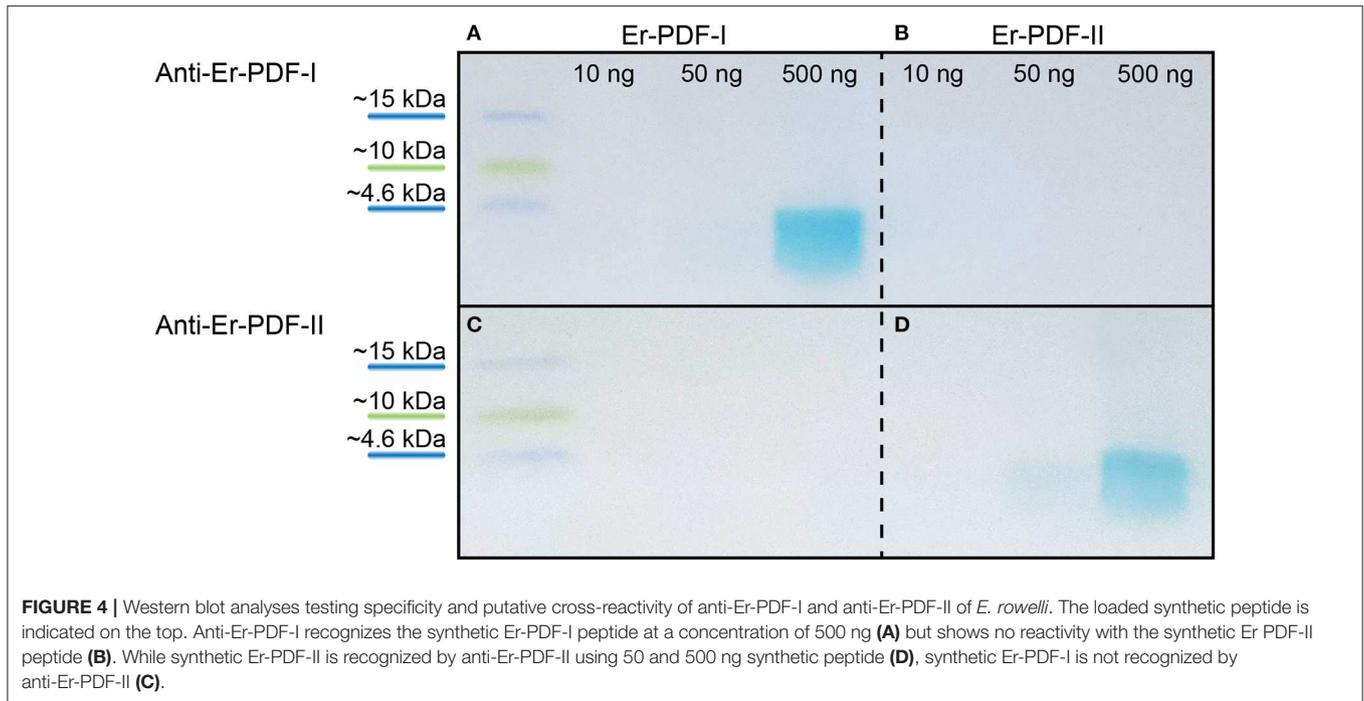
Immunolocalization of PDF-I and PDF-II Peptides and Their Receptor

Immunolabeling against the Er-PDF-I and Er-PDF-II peptides revealed high numbers of somata and fibers in the brain and the ventral nerve cords of *E. rowelli* (Figures 6A–D, 7A–C, A'–C'). In each brain hemisphere, Er-PDF-I and Er-PDF-II immunoreactive (ir) somata occur mainly in one large ventromedian (Figures 7A'–C') and two dorsal groups, the latter comprising a smaller anterior and a larger median group (Figures 7A–C). While the somata of the anterior group innervate the anterior neuropil, the larger median group is associated with the central body and the central neuropil and is located closer to the dorsal surface of the brain (Figures 6A,B; Supplementary Video 1). The somata of the ventromedian group occupy mainly the median portion of the ventral perikaryal layer, in which they are situated at different



depths (**Figures 6A,B, 7A',B'**). Notably, no Er-PDF-I-ir or Er-PDF-II-ir somata are found in the so-called hypocerebral organs-vesicle-like structures associated with the ventral cortex of each brain hemisphere (**Supplementary Figures 3A,B**). Additional

groups of Er-PDF-I-ir and Er-PDF-II-ir somata occur in the ventral part of the deutocerebrum as well as in the connecting cords, which link the brain with the ventral nerve cords (**Figure 6C**). Within the ventral nerve cords,



the somata are mainly located in the ventral and median perikaryal layers (**Figure 6D**). While most somata with a strong Er-PDF-II signal appear ventrally, those with a strong Er-PDF-I signal show a wide mediolateral distribution in each nerve cord.

In addition to somata, our immunolabeling revealed numerous Er-PDF-I-ir and Er-PDF-II-ir fibers, which are characterized by typical varicosities and occur in all major neuropils of the *E. rowelli* brain including the central neuropil, the central body, the mushroom bodies, the olfactory lobes, and the optic tracts and neuropils (**Figures 6A–C**). Despite the nearly ubiquitous distribution of Er-PDF-I-ir/Er-PDF-II-ir fibers in all brain neuropils, their density differs among specific regions. While the central neuropil shows a dense mesh of internal Er-PDF-I-ir and Er-PDF-II-ir fibers, the olfactory glomeruli are rather surrounded by them (**Figures 6A,B**). The lowest density of Er-PDF-I-ir and Er-PDF-II-ir fibers is found in the antennal tracts and the mushroom bodies, both of which exhibit fibers mainly in the periphery, except for the median lobe of the mushroom body, which additionally shows strong internal Er-PDF-I immunoreactivity (**Figure 6B**). In contrast, the hypocerebral organs do not contain any Er-PDF-I-ir or Er-PDF-II-ir fibers (**Supplementary Figures 3A,B**). Like in the brain, numerous Er-PDF-I-ir and Er-PDF-II-ir fibers are evident in the neuropils of each nerve cord (**Figure 6D**). Most fibers with a prominent Er-PDF-II immunoreactivity are confined to the ventral portion of the ventral cord neuropil, whereas those exhibiting strong Er-PDF-I signal are more widely distributed, thus reflecting the wide distribution of the corresponding somata in the perikaryal layer of the nerve cord.

To determine whether Er-PDF-I and Er-PDF-II are co-expressed at least in some cells and regions of the nervous system, we performed double labeling. Our data indeed show that both peptides are found in the same sets of neuronal somata and fibers within the brain and the ventral nerve cords (**Figures 7C, 8A–H**). However, the signal intensity differs markedly between

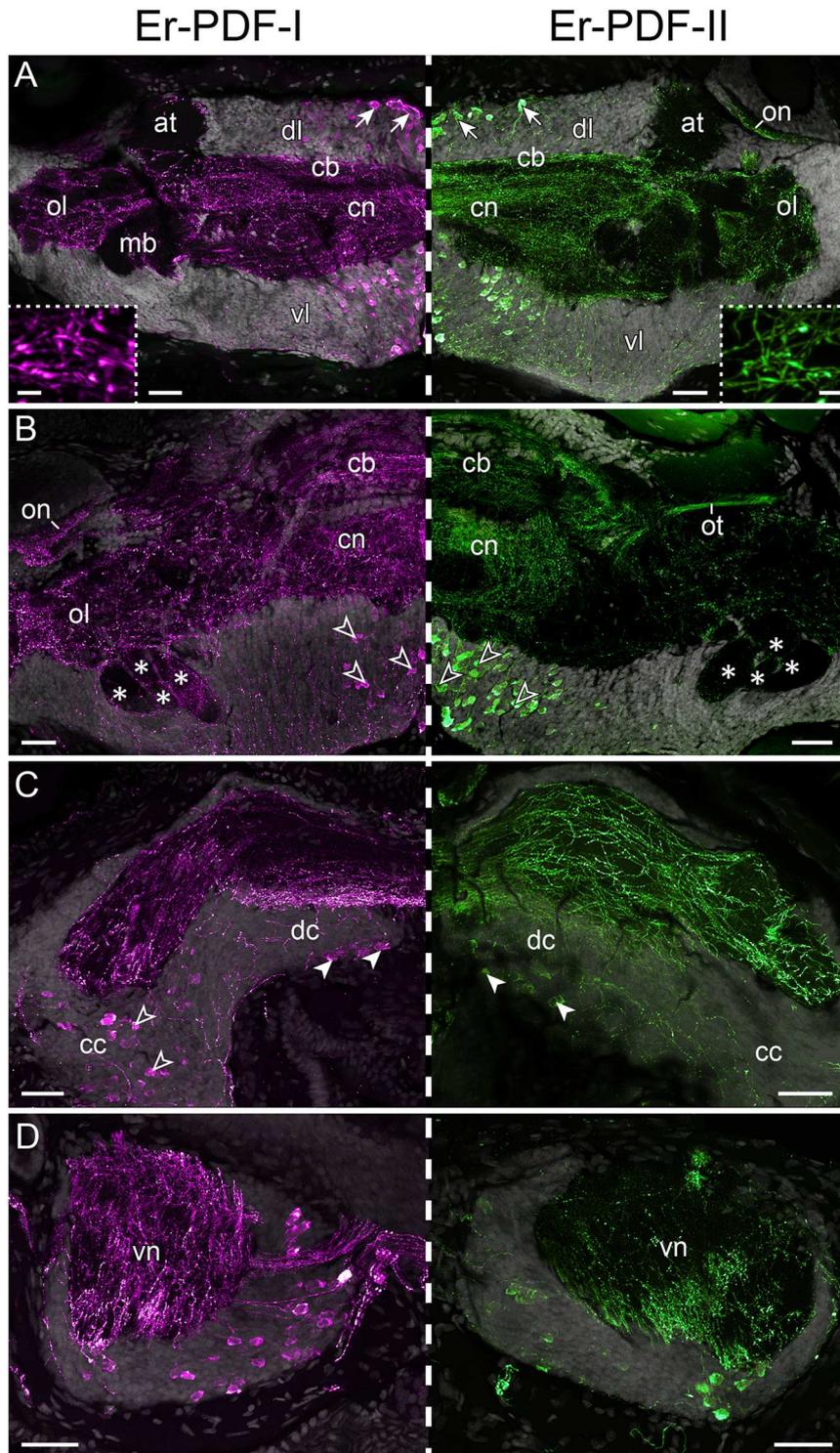
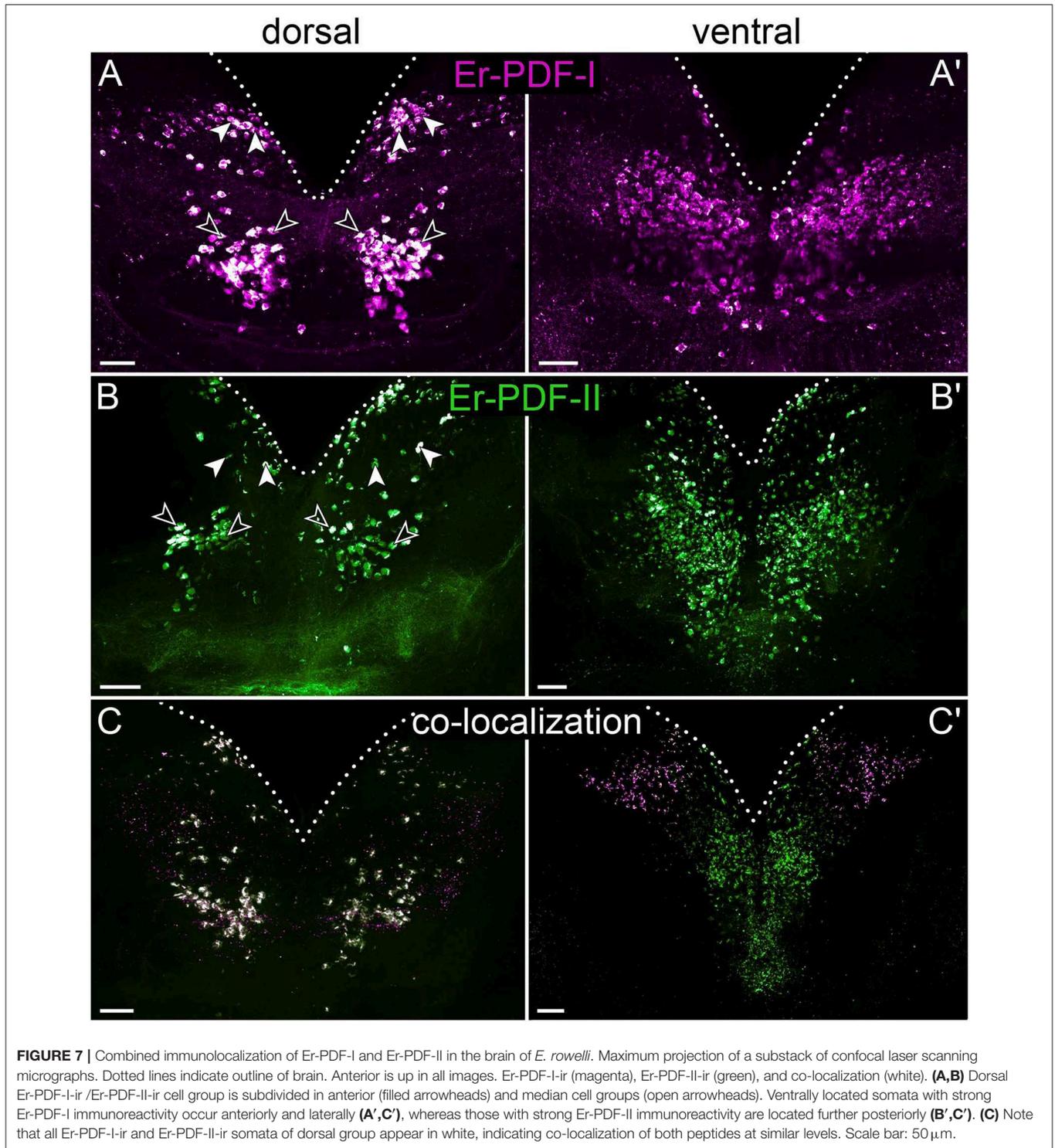


FIGURE 6 | Immunolocalization of Er-PDF-I and Er-PDF-II in *E. rowelli*. Confocal laser scanning micrographs of vibratome sections. Dorsal is up in all images. Hatched line indicates median regions. Er-PDF-I-ir (magenta), Er-PDF-II-ir (green) and DNA-labeling (gray) from anterior to posterior through head (**A–C**) and trunk (**D**). Note similar distribution of either peptide in brain and ventral nerve cords. (**A**) Arrows indicate dorsal groups of somata in protocerebrum. Insets show large varicosities in PDF-immunoreactive fibers. (**B**) Arrowheads point to large ventral groups of somata in protocerebrum. Asterisks indicate four lobes of mushroom bodies. (**C**) Filled arrowheads point to somata in deutocerebrum. Empty arrowheads demarcate somata in connecting cords. (**D**) Cross sections of ventral nerve cords. at, antennal tract; cb, central body; cc, connecting cord; cn, central neuropil; dc, deutocerebrum; dl, dorsal perikaryal layer; ol, olfactory lobe; on, optic neuropil; ot, optic tract; vl, ventral perikaryal layer; vn, neuropil of ventral nerve cord. Scale bars: 50 μ m (**A–D**) and 500 nm (insets).



tissues and cells. For example, while Er-PDF-I and Er-PDF-II are both expressed at a high level in the dorsal groups of somata within the brain (**Figures 7C, 8A–C**), the ventromedian groups show a more sophisticated pattern, with Er-PDF-I exhibiting a stronger signal in the anterolateral and Er-PDF-II in the median regions (**Figures 7C', 8A,B,D**). A differentiated signal is

evident even within individual cells. While some somata show a completely overlapping signal, others seem to express either Er-PDF-I or Er-PDF-II (**Figures 7C,C', 8A–D**). However, high resolution imaging confirmed the presence of both peptides in all labeled somata, although they seem to be expressed at different levels (**Supplementary Figures 4, 5**).

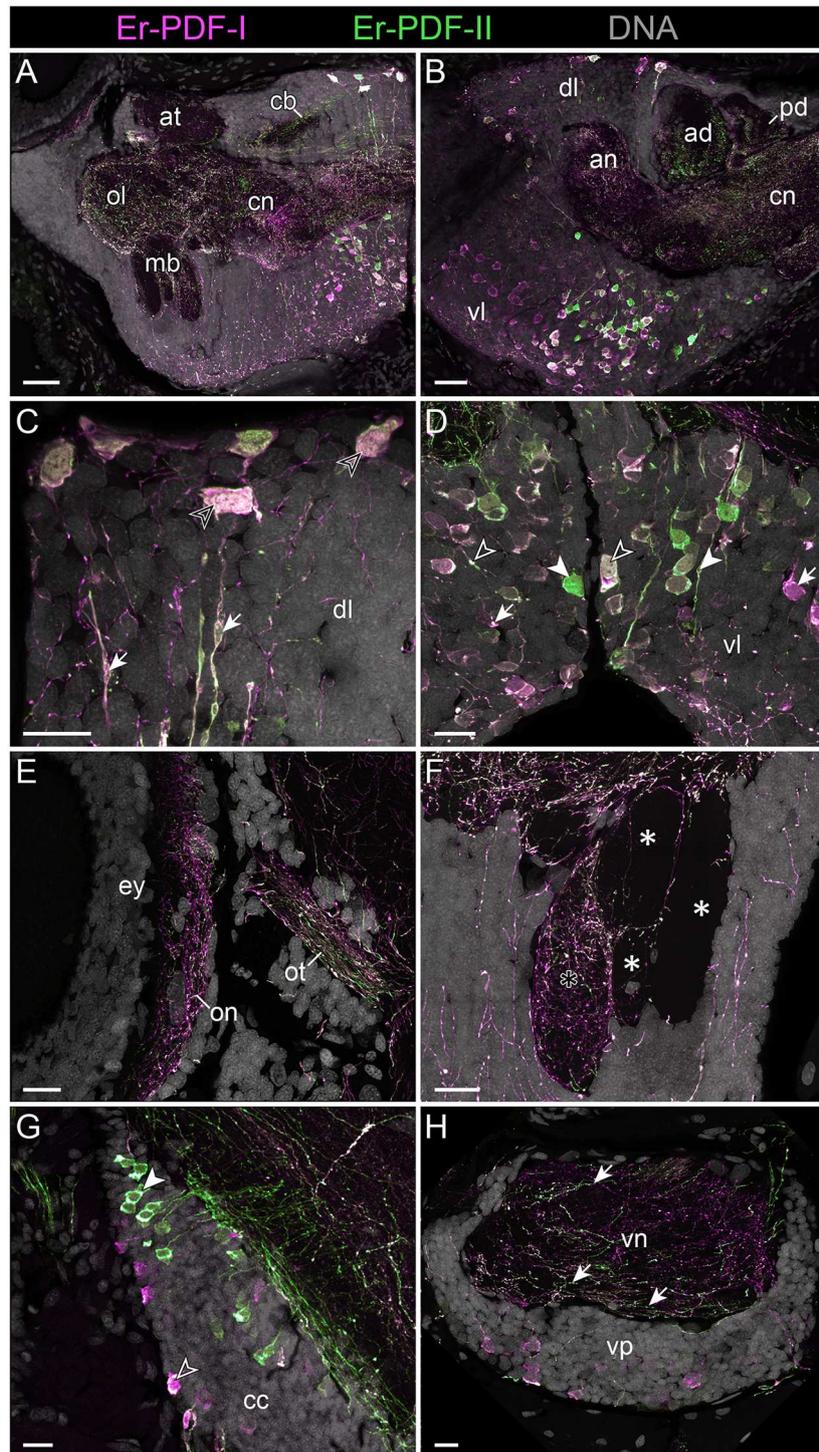


FIGURE 8 | Combined immunolocalization of Er-PDF-I and Er-PDF-II in the brain of *E. rowelli*. Confocal laser scanning micrographs of vibratome sections. Dorsal is up in all images and anterior is left in **(B,G)**. Er-PDF-I-ir (magenta), Er-PDF-II-ir (green) and DNA (gray). **(A,B)** Er-PDF-I and Er-PDF-II are co-localized in neuropil and somata, albeit at different intensities. **(C)** Somata (arrowheads) and processes (arrows) of dorsal cell group exhibit equal intensity levels of Er-PDF-I-ir and Er-PDF-II-ir. **(D)** In contrast, somata and processes of ventral cell group occur in three variants: (i) Er-PDF-I-ir and Er-PDF-II-ir at equal levels (open arrowheads), (ii) Er-PDF-I-ir at higher level (arrows), and (iii) Er-PDF-II-ir at higher level (filled arrowheads). **(E–H)** Differences in expression levels of Er-PDF-I-ir and Er-PDF-II-ir are also seen in optic neuropil **(E)**, inner lobe (black asterisk) as opposed to remaining lobes of the mushroom bodies (white asterisks) **(F)**, somata of connecting cords **(G)**, and somata and neuropil of nerve cords **(H)**. an, anterior neuropil; ad, anterior division of central body; at, antennal tract; cb, central body; cc, connecting cord; cn, central neuropil; dl, dorsal perikaryal layer; ey, eye; mb, mushroom body; ol, olfactory lobe; on, optic neuropil; ot, optic tract; pd, posterior division of the central body; vl, ventral perikaryal layer; vn, neuropil of ventral nerve cord; vp, perikaryal layer of ventral nerve cord. Scale bars: 50 μm **(A,B)** and 20 μm **(C–H)**.

Different signal intensities are also evident among the fibers. Like the somata, the fibers show three different expression patterns with either both peptides expressed similarly or one or the other occurring at a higher level (Figures 8A–H, 9A). Differences are evident, for example, in the central body, which shows a stronger Er-PDF-II signal (Figure 8B), whereas the optic neuropil exhibits a higher expression of Er-PDF-I (Figure 8E). Double labeling of the mushroom body confirms the predominant occurrence of Er-PDF-I inside the median lobe (Figure 8F), which was detected by separately labeling the individual peptides (cf. Figure 6B). Likewise, the double labeling confirms the differences in the distribution of Er-PDF-I and Er-PDF-II in the neuropil of the ventral nerve cords, which shows a condensation of fibers with a strong Er-PDF-II signal in its ventral portion but a more ubiquitous distribution of fibers with a strong Er-PDF-I signal (Figure 8H).

Although the somata of the Er-PDF-I-ir and Er-PDF-II-ir neurons are exclusively found in the central nervous system of *E. rowelli*, it is worth mentioning that at least some of their fibers extend into the peripheral nervous system including various peripheral nerves (such as leg nerves, slime papilla nerves, tongue nerve, oral and pharyngeal nerves), the ring commissures, and the heart nerve (Figure 9A). In the heart wall, the fibers show the same differential expression pattern of the two peptides as in the remaining nervous system. While the fibers of the longitudinal, dorsomedian heart nerve express both peptides at a similar level, the associated transverse fibers exhibit a higher intensity of Er-PDF-II signal than Er-PDF-I. Remarkably, the high resolution confocal microscopy used for imaging reveals diverging patterns of Er-PDF-I and Er-PDF-II immunoreactivity even within the individual fibers in our specimens (Figures 8C,D, 9A).

In addition to the Er-PDF-I and Er-PDF-II peptides, we immunolocalized the newly identified and deorphanized pigment-dispersing factor receptor (PDFR) of *E. rowelli*. As expected from its nature as a membrane receptor protein, the Er-PDFR immunoreactivity is generally found in the periphery of each cell (e.g., Figures 9B,C). Like the two analyzed PDF peptides, Er-PDFR is localized in a dorsal and a ventral group of somata within the brain (Figures 10A–C, 11A–C). However, double labeling with either Er-PDF-I or Er-PDF-II revealed no co-expression of Er-PDFR in the dorsal groups of neuronal somata, although at least some PDFR-ir cells are either directly adjacent to or located in close proximity to the Er-PDF-I-ir/Er-PDF-II-ir somata (Figures 10B, 11B). In contrast to the dorsal group, the Er-PDFR-ir somata of the ventral group show three different conditions: they are either (i) not closely associated with the Er-PDF-I-ir/Er-PDF-II-ir somata, (ii) directly adjacent to them, or (iii) co-express the two peptides and the receptor (Figures 10C, 11C). The hypocerebral organs exhibit no PDFR-ir somata (Supplementary Figure 3C).

Besides the dorsal and ventral perikaryal layers of the brain, additional Er-PDFR-ir somata occur, for example, next to the olfactory glomeruli (Supplementary Figure 6A) and the median lobe of the mushroom body (Supplementary Figure 6B). Moreover, Er-PDFR is expressed in a few dorsolateral somata of the brain near the eye as well as in some somata of the optic

ganglion itself (Figures 10A, 11A, 12A–C). Er-PDFR-expressing somata are absent in the ventral nerve cords, but are abundant in the connecting cords (Supplementary Figures 6C,D). Here, double labeling solely revealed somata with either Er-PDFR or Er-PDF-I/Er-PDF-II signal, whereas co-expression of both the peptide and the receptor is not evident.

In addition to the described Er-PDFR immunoreactivity in the neuronal somata, the receptor is identifiable in the neuropils of the brain, the eyes and the ventral nerve cords (Figures 10A, 11A, Supplementary Figures 6A–C). However, the signal in the neuronal processes is generally weaker and less defined than in the somata. Single fibers are thus difficult to follow and appear as numerous dots in each neuropil region (Supplementary Figures 6A–C).

Notably, in contrast to the PDF signal, Er-PDFR immunoreactivity is not restricted to the nervous system of *E. rowelli* but is additionally found in other cells and tissues (Figures 9B,C, 10A, 11A, 12A–D). The most prominent signal outside the nervous system occurs in the visual system including the microvilli of the photoreceptor processes (Figures 12A,B) and the pigment granules [cf. 86]) of the supportive cells (Figures 12C,D). Er-PDFR immunoreactivity is further associated with membranes of certain blood cells (= hemocytes), which appear blue under the bright-field microscope (Figures 9B,C, Supplementary Figures 7A,B). These blueish Er-PDFR-ir hemocytes do not only occur in the heart lumen but also in other regions of the circulatory system of *E. rowelli* (Figure 9C).

DISCUSSION

Deorphanization of the Onychophoran PDF Receptor

Using a combination of transcriptomic and phylogenetic analyses, we were able to identify a putative PDF receptor candidate in the onychophoran *E. rowelli*. PDFRs are generally involved in cAMP-mediated signaling pathways and—in response to PDF—increase intracellular cAMP levels *in vivo* (92) and also *in vitro*, for example, when expressed in human embryonic kidney cells [HEK293; (57)]. We used the latter approach for clarifying whether or not the identified receptor is functional and whether it responds to both peptides by measuring its activity in the presence of either Er-PDF-I or Er-PDF-II *in vitro*. This was done by detecting the cAMP-induced activation of an Epac-based sensor [“Exchange protein activated by cAMP”; (93)] using bioluminescence resonance energy transfer [BRET; (84, 90, 94)]. Our data revealed that Er-PDFR is activated by both peptides, although Er-PDF-II seems to be one order of magnitude more potent than Er-PDF-I. A similar result was obtained from nematodes (41), but the reason for the observed differences in stimulation of the PDF receptor by different ligands remains unknown. Taken together, the results of our phylogenetic analyses and BRET experiments suggest that the identified receptor is indeed functional and most likely represents the sole PDF receptor of the onychophoran *E. rowelli*. Our BRET data further suggest that

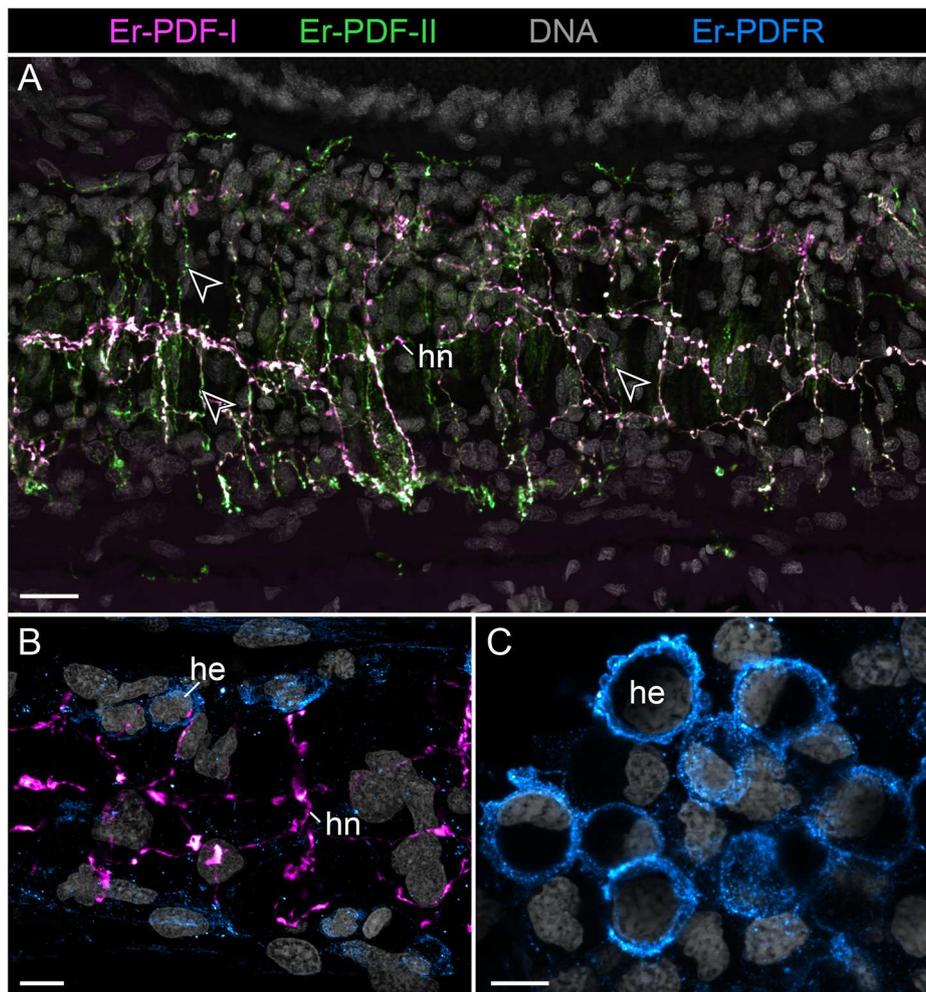


FIGURE 9 | Combined immunolocalization of either Er-PDF-I or Er-PDF-II with Er-PDFR in the vascular system of *E. rowelli*. Confocal laser scanning micrographs of vibratome sections. Anterior is left in all images. Er-PDF-I (magenta), Er-PDF-II (green), Er-PDFR (cyan), and DNA (gray). **(A)** Er-PDF-I and Er-PDF-II are co-localized in heart nerve with varying intensities, especially within fine branches (arrowheads). **(B,C)** PDFR immunoreactivity in cell membranes of hemocytes found in heart lumen **(B)** and body cavity **(C)**. he, hemocyte; hn, heart nerve. Scale bars: 25 μm **(A)**, 10 μm **(B)**, and 5 μm **(C)**.

Er-PDFR is activated by Er-PDF-I and Er-PDF-II in a distinct and dose-dependent manner.

Specificity of the Generated Antibodies

The western blots conducted for testing the affinity of the newly generated antibodies against Er-PDF-I and Er-PDF-II revealed specific bands below 4.6 kDa, corresponding to the expected molecular weight of both peptides of 1.9–2.0 kDa each. Specificity tests using the two antibodies against each non-corresponding synthetic peptide (anti Er-PDF-I and anti Er-PDF-II) were negative, suggesting that the two antibodies are specific and do not cross-react.

Additional western blots for testing the affinity of the antibody generated against the PDF receptor of *E. rowelli* yielded more intricate but consistent results. For these tests, we used a synthetic PDFR peptide coupled to BSA together with a lysate of dissected heads at different dilutions. All of these tests revealed a distinct

band at ~ 70 kDa corresponding well to the expected molecular weight of the synthetic PDFR peptide-BSA conjugate (which is ~ 6 kDa: 1.9 kDa fractionized Er-PDFR plus 66 kDa BSA) and the entire receptor protein (61.06 kDa) in the head lysate. The appearance of additional bands at ~ 120 kDa and ~ 250 kDa may have been induced by the formation of dimers and multimers by the receptor protein itself/the BSA coupled to the synthetic peptide. We believe that this result does not compromise the specificity of the Er-PDFR antibody generated for this study, as it might be due to a technical artifact.

Differential Expression of PDF-I and PDF-II Peptides in the Onychophoran Nervous System

Our immunolabeling using two specific antibodies against the PDF-I and PDF-II peptides of *E. rowelli* largely confirms the

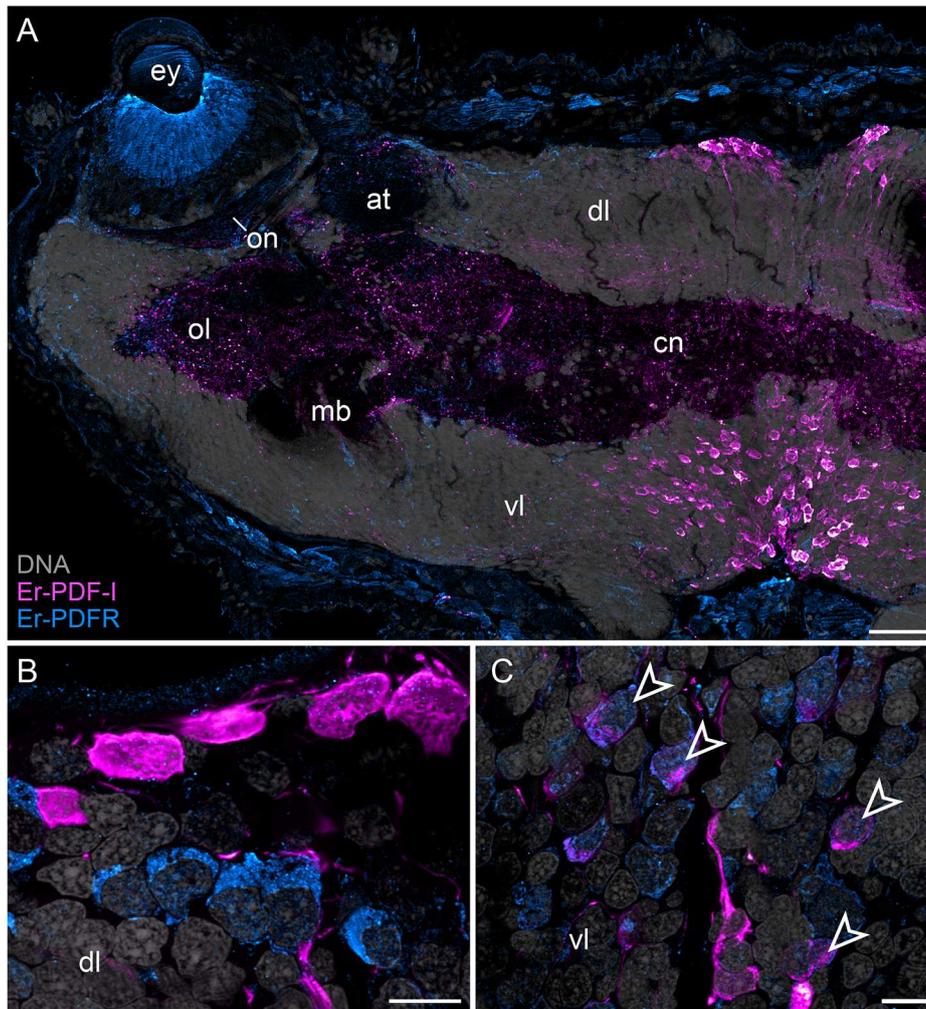


FIGURE 10 | Combined immunolocalization of Er-PDF-I and Er-PDFR in *E. rowelli*. Confocal laser scanning micrographs of vibratome sections. Dorsal is up in all images. Er-PDF-I (magenta), Er-PDFR (cyan), and DNA (gray). Note that cuticle is autofluorescent. **(A)** Overview of protocerebrum. **(B)** Detailed view of dorsal perikaryal layer. **(C)** Detailed view of ventral perikaryal layer. Note that Er-PDF-I and Er-PDFR are co-localized only in some cells of ventral group (arrowheads). at, antennal tract; cn, central neuropil; dl, dorsal perikaryal layer; ey, eye; mb, mushroom body; ol, olfactory lobe; on, optic neuropil; vl, ventral perikaryal layer; Scale bars: 50 μm **(A)** and 10 μm **(B,C)**.

broad distribution of PDFs in the onychophoran nervous system demonstrated previously (38). In accord with these previous findings, we encountered PDF immunoreactivity in groups of neuronal somata in the ventral and dorsal protocerebrum, the deutocerebrum, the ventral nerve cords, as well as an elaborate fiber network within the brain, the ventral nerve cords and the peripheral nervous system including the heart nerve. Interestingly, we found that both peptides are co-expressed in all immunoreactive structures of *E. rowelli*, although their respective levels of expression differ at least in some cells and tissues. We were able to recognize three major patterns: (i) equal staining intensities of both peptides, (ii) higher Er-PDF-I signal, and (iii) higher Er-PDF-II signal. These different patterns were particularly evident in specific groups of somata within the brain (Figure 13).

Currently, we can only speculate about the functional significance of these observations. One possible reason for

the differential expression of Er-PDF-I and Er-PDF-II might be their different functionality, as evidenced by our BRET assays with the deorphanized PDF receptor. Interestingly, the nematode *C. elegans* also possesses two *pdf* genes (albeit three PDF isoforms), which are co-expressed in certain cells including interneurons, and chemosensory and motor neurons (42). The PDF peptides in these cells seem to have opposing effects on locomotion, as PDF-1 deletion mutants and animals with overexpressed PDF-2 show the same atypical locomotory behavior (41, 49). Irrespective of whether or not similar opposing effects of the two peptides also exist in onychophorans, their co-expression in all immunoreactive cells of *E. rowelli* and at least in some cells of *C. elegans* suggests that this feature might have been present in the last common ancestor of Ecdysozoa.

The co-localization of the two PDF peptides in onychophorans and nematodes contrasts with what is known from arthropods that exhibit more than one PDF/PDH isoform

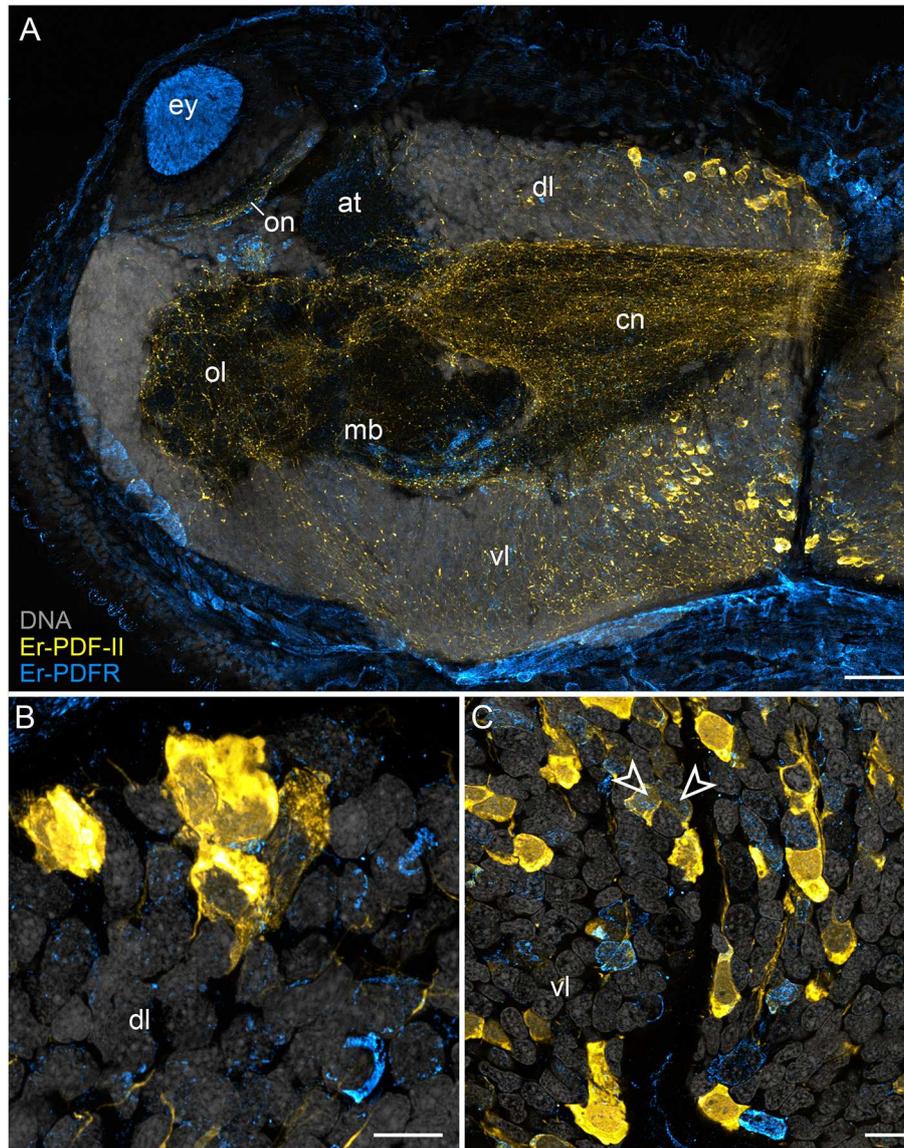


FIGURE 11 | Combined immunolocalization of Er-PDF-II and Er-PDFR in *E. rowelli*. Confocal laser scanning micrographs of vibratome sections. Dorsal is up in all images. Er-PDF-II (yellow), Er-PDFR (cyan), and DNA (gray). Note that cuticle is autofluorescent. **(A)** Overview of protocerebrum. **(B)** Detailed view of dorsal perikaryal layer. **(C)** Detailed view of ventral perikaryal layer. Note that Er-PDF-II and Er-PDFR are co-localized in some cells of the ventral group (arrowheads). at, antennal tract; cn, central neuropil; dl, dorsal perikaryal layer; ey, eye; mb, mushroom body; ol, olfactory lobe; on, optic neuropil; vl, ventral perikaryal layer; Scale bars: 50 μm **(A)** and 10 μm **(B,C)**.

such as decapod crustaceans. These animals express two to three PDH peptides (isoforms of either α -PDH or β -PDH), which are assumed to fulfill different functions based on their distinct localization (11, 43, 44, 46–48, 95–100). For example, while β -PDH II is expressed in the sinus gland of *Cancer productus*, and therefore most likely acts as a neurohormone, β -PDH is localized in the eyestalk, thus serving as a neuromodulator or neurotransmitter (11). It must be noted, however, that all crustacean PDH isoforms are derivatives and in-paralogs of *pdf-I* (**Figure 1**), whereas onychophorans and nematodes most likely inherited two *pdf* genes, *pdf-I* and *pdf-II* (*pdf-1* and *pdf-2 sensu*

41), from the last common ancestor of Ecdysozoa (38). From the evolutionary point of view, the multiple isoforms of decapod crustaceans are a derived feature of this group.

Immunolocalization of the Onychophoran PDF Receptor Indicates Different Signal Transduction Mechanisms and Supports Hormonal Control

Localization of the deorphanized PDF receptor in *E. rowelli* using a specific Er-PDFR antibody revealed prominent

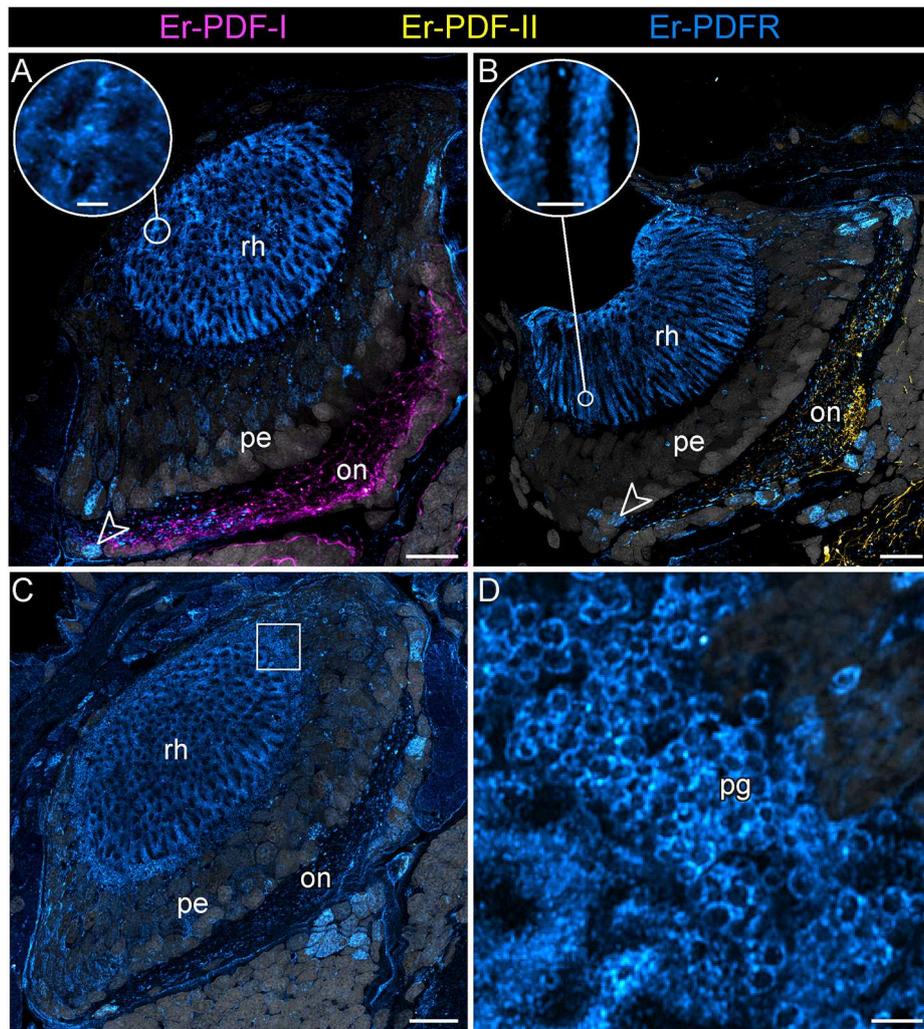


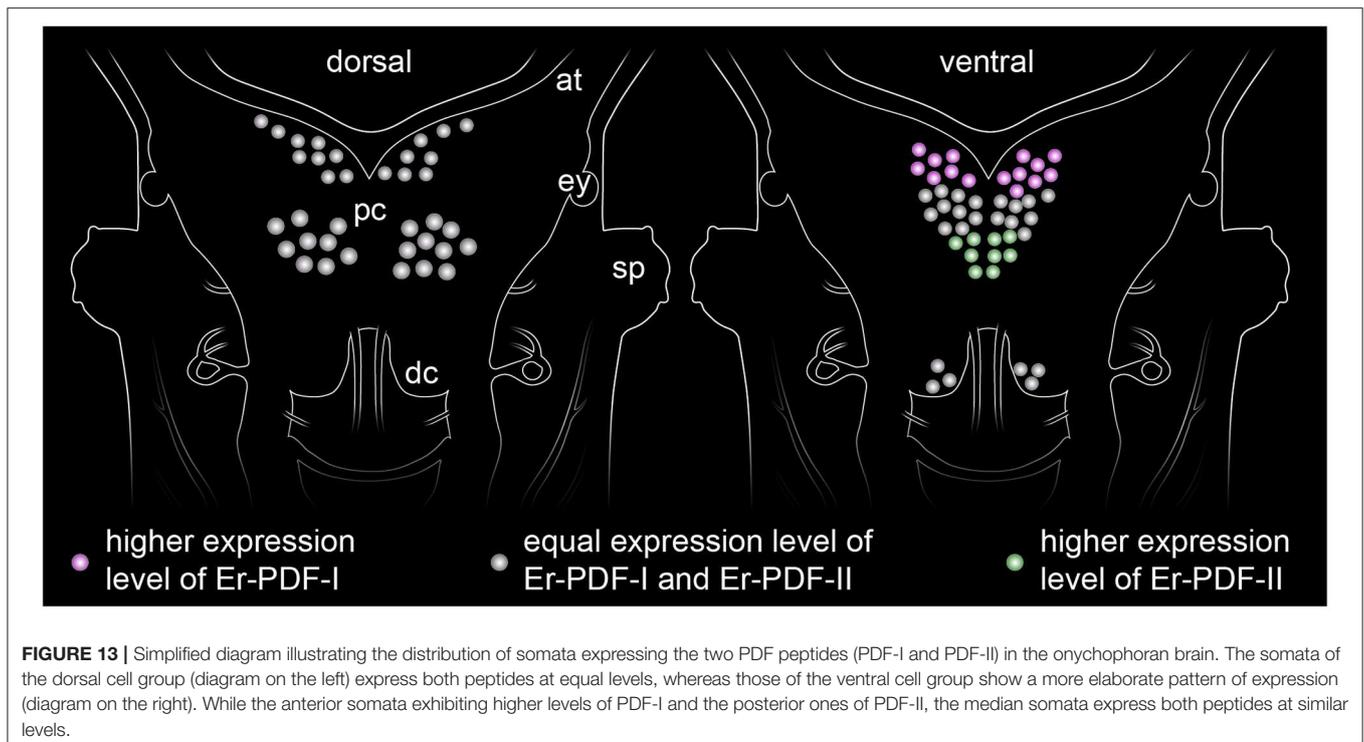
FIGURE 12 | Combined immunolocalization of either Er-PDF-I or Er-PDF-II with Er-PDFR in the visual system *E. rowelli*. Confocal laser scanning micrographs of vibratome sections. Dorsal is up in all images. Er-PDF-I (magenta), Er-PDF-II (yellow), Er-PDFR (cyan) and DNA (gray). **(A–D)** PDFR immunoreactivity occurs in optic neuropil, few somata of optic ganglion (arrowheads), rhabdomeric layer **(A–C)** and insets for detail, and pigment granules **(C,D)**. Note the lack of co-localization of Er-PDFR with Er-PDF-I **(A)** or Er-PDF-II **(B)**. on, optic neuropil; pe, perikaryal layer of eye; pg, pigment granules rh, rhabdomeric layer. Scale bars: 20 μm **(A–C)** and 2 μm (insets of **A, B,D**).

immunoreactivity associated with membranes of specific cells in different organ systems. In contrast to the two PDF peptides, PDFR is not expressed in the ventral nerve cords but only in the brain and the connecting cords (101) within the central nervous system. There are three types of PDFR-expressing cells in the brain including those (i) co-expressing PDFR and PDF-I/PDF-II, (ii) not co-expressing but directly adjacent to PDF-I/PDF-II-expressing cells, and (iii) not closely associated with PDF-I/PDF-II immunoreactive cells. This indicates three different mechanisms of signal transduction:

(i) Co-localization of PDFR and the two PDF peptides suggests autoreception, i.e., feedback on its own cell, which has also been reported from fruit flies and Madeira cockroaches

(34, 102). In insects, however, the majority of PDF-expressing cells are autoreceptive and they are typically outnumbered by PDFR-expressing cells (60), whereas PDF-immunoreactive cells are more numerous than PDFR-expressing cells in the onychophoran brain.

- (ii) In cells that do not co-express the PDF peptides and their receptor, the signal transduction might be accomplished through wiring transmission via the synaptic contacts or membrane juxtapositions of cells with direct contact to each other (103).
- (iii) Finally, cells that are not directly associated with each other might communicate via the hormonal release of peptides (i.e., volume transmission), like in decapod crustaceans (43, 95). The PDF-ir cells of onychophorans are



indeed characterized by numerous varicosities and axonal terminals, which are reminiscent of typical release sites (38).

The so-called hypocerebral organs—enigmatic, vesicle-like structures associated with the onychophoran brain—have been repeatedly proposed to play a neurosecretory role (104, 105). Although we cannot rule out this potential function, the lack of PDF-I, PDF-II, and PDFR immunoreactivity in their tissues suggests that the hypocerebral organs do not produce or secrete PDF peptides; neither do they seem to be involved in the PDF/PDFR system.

Apart from the central nervous system, prominent PDFR signal is associated with the visual system of *E. rowelli*. In particular, membranes surrounding the microvilli of the photoreceptor cell processes (rhabdoms) and the pigment granules of the supportive cells [cf. (106)] are highly immunoreactive. PDH has been described to induce retinal pigment dispersion in the eyes of crustaceans under bright light conditions (1, 3, 5, 6). Whether or not such pigment dispersion occurs in the onychophoran eye is unclear, but it seems unlikely due to the generally low resolution of vision in these animals (107) and the position of pigment granules at the base of the rhabdomeric layer (106, 108, 109). Regardless of in which direction the pigment granules would move within the supportive cells, their basal position prevents shading/protection of photoreceptor processes from incident light. Additional PDFR immunoreactivity within the eye occurs in a few somata of the optic ganglion [cf. (109)]. Whether this signal is associated with glial cells, as reported from *D. melanogaster* (61), remains to be clarified.

Most intriguingly, besides the nervous and visual systems of *E. rowelli*, we detected PDFR in membranes of specific hemocytes. These hemocytes appeared blueish under the bright-field microscope and were localized in different parts of the circulatory system including the heart. Up to five major types of hemocytes have been described from onychophorans based on their ultrastructure (110, 111), but beyond this only little is known about their possible functions. At this point, we can only speculate that the blue-pigmented granules might be due to storage of hemocyanin (112) or, alternatively, might be associated with a role of these cells in immune defense, as hemocytes have been proposed to absorb dead ectodermal cells (113), which in *E. rowelli* contain blue pigment granules. Irrespective of whether these hemocytes are involved in the potential storage of hemocyanin or innate immune response, each of these roles might be controlled by the PDF/PDFR system depending on the amount of PDF peptides released into the hemolymph. This potentially novel function of PDF peptides associated with hemocytes requires further investigation. Nonetheless, the demonstrated occurrence of PDF receptor in cells of the visual and circulatory systems of *E. rowelli* clearly supports the suggested (38) hormonal role of PDF peptides in onychophorans.

CONCLUSIONS

In this study, we deorphanized and immunolocalized the onychophoran PDF receptor and performed double labeling using specific antibodies against the two onychophoran PDF peptides. We further explored potential differences in the stimulation of PDFR by each peptide, revealing that PDF-II has

a tenfold higher potency than PDF-I as an activating ligand. Double immunolabeling demonstrates that both onychophoran PDF peptides are co-expressed but their respective levels of expression show cell-specific variation. For example, some neurons express the same amount of both peptides, while others exhibit higher levels of either PDF-I or PDF-II. Whether this variation is due to cyclic changes, as for example in insects (33, 114, 115), remains to be clarified.

The detection of the onychophoran PDF receptor in cells that additionally express the two PDF peptides suggests autoreception, whereas spatial separation of PDFR- and PDF-expressing cells confirms hormonal release into the hemolymph (38). Hence, the PDF peptides of onychophorans might play a dual role—as hormones and as neurotransmitters or neuromodulators—similar to the PDH peptides of decapod crustaceans (7–12). Whether the PDF-releasing cells of onychophorans are light-responsive ultradian and circadian oscillators, as for example in the Madeira cockroach (116–119), is unknown. Future studies should therefore focus on clarifying whether there are any cycling patterns in the expression of the two peptides. Establishment of cell cultures (e.g., for *in vivo* calcium imaging) and corresponding behavioral assays would contribute to a better understanding of the PDF/PDFR system in Onychophora and the last common ancestors of Panarthropoda and Ecdysozoa.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

LH, CM, AW, FH, MS, and GM designed the research. LH conducted the BRET experiments and the phylogenetic

analyses. CM, NM, and SH performed the immunolabeling experiments. NM, SH, SK, and SF carried out the specificity tests. CM, LH, NM, SH, and GM analyzed the data and wrote the first draft. All authors have read and approved the final manuscript.

FUNDING

This work was supported by German Research Foundation (DFG: MA 4147/10-1) to GM. University of Kassel, Programmlinie Aufbau Graduiertenprogramme (Biological Clocks: 1.10.01.EWN/E1) to MS, FH, and GM.

ACKNOWLEDGMENTS

The authors are thankful to Noel N. Tait and Dave M. Rowell for help with the permits. Dave M. Rowell, Ivo de Sena Oliveira, Isabell Schumann, and Alexander Baer are acknowledged for their support with collecting the specimens. We gratefully acknowledge Alexander Meier for providing the Epac-L cAMP-sensor and all members of the Department of Zoology, University of Kassel, for their assistance with animal husbandry. We are thankful to Thordis Arnold, Erik Machal, and Irmtraud Hammerl-Witzel for assistance in laboratory work. Vladimir Gross kindly proofread the manuscript. We thank the staff of the Department of Sustainability, Environment, Water, Population and Communities of the Australian Government for providing collecting and export permits.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2020.00273/full#supplementary-material>

REFERENCES

- Helfrich-Förster C. Neuropeptide PDF plays multiple roles in the circadian clock of *Drosophila melanogaster*. *Sleep Biol Rhythms*. (2009) 7:130–43. doi: 10.1111/j.1479-8425.2009.00408.x
- Meelkop E, Temmerman L, Schoofs L, Janssen T. Signalling through pigment dispersing hormone-like peptides in invertebrates. *Prog Neurobiol*. (2011) 93:125–47. doi: 10.1016/j.pneurobio.2010.10.004
- Larimer JL, Smith JT. Circadian rhythm of retinal sensitivity in crayfish: modulation by the cerebral and optic ganglia. *J Comp Physiol*. (1980) 136:313–26. doi: 10.1007/BF00657351
- Rao KR, Riehm JP. Pigment-dispersing hormones: a novel family of neuropeptides from arthropods. *Peptides*. (1988) 9:153–9. doi: 10.1016/0196-9781(88)90239-2
- Verde M, Barriga-Montoya C, Fuentes-Pardo B. Pigment dispersing hormone generates a circadian response to light in the crayfish, *Procambarus clarkii*. *Comp Biochem Phys A*. (2007) 147:983–92. doi: 10.1016/j.cbpa.2007.03.004
- Strauss J, Dircksen H. Circadian clocks in crustaceans: identified neuronal and cellular systems. *Front Biosci*. (2010) 15:1040–74. doi: 10.2741/3661
- Dircksen H, Zahnow CA, Gaus G, Keller R, Rao KR, Riehm JP. The ultrastructure of nerve endings containing pigment-dispersing hormone (PDH) in crustacean sinus glands: identification by an antiserum against a synthetic PDH. *Cell Tissue Res*. (1987) 250:377–87. doi: 10.1007/BF00219082
- Mangerich S, Keller R, Dircksen H, Rao RK, Riehm JP. Immunocytochemical localization of pigment-dispersing hormone (PDH) and its coexistence with FMRFamide-immunoreactive material in the eyestalks of the decapod crustaceans *Carcinus maenas* and *Orconectes limosus*. *Cell Tissue Res*. (1987) 250:365–75. doi: 10.1007/BF00219081
- Mangerich S, Keller R. Localization of pigment-dispersing hormone (PDH) immunoreactivity in the central nervous system of *Carcinus maenas* and *Orconectes limosus* (Crustacea), with reference to FMRFamide immunoreactivity in *O. limosus*. *Cell Tissue Res*. (1988) 253:199–208. doi: 10.1007/BF00221755
- Nussbaum T, Dircksen H. Neuronal pathways of classical crustacean neurohormones in the central nervous system of the woodlouse, *Oniscus asellus* (L.). *Philos Trans R Soc Lond B Biol Sci*. (1995) 347:139–54. doi: 10.1098/rstb.1995.0018
- Hsu Y-WA, 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.* (2008) 508:197–211. doi: 10.1002/cne.21659
12. Harzsch S, Dircksen H, Beltz BS. Development of pigment-dispersing hormone-immunoreactive neurons in the *American lobster*: homology to the insect circadian pacemaker system? *Cell Tissue Res.* (2009) 335:417–29. doi: 10.1007/s00441-008-0728-z
 13. Rao KR, Mohrherr CJ, Riehm JP, Zahnow CA, Norton S, Johnson L, et al. Primary structure of an analog of crustacean pigment-dispersing hormone from the lubber grasshopper *Romalea microptera*. *J Biol Chem.* (1987) 262:2672–5.
 14. Homberg U, Würden S, Dircksen H, Rao KR. Comparative anatomy of pigment-dispersing hormone-immunoreactive neurons in the brain of orthopteroid insects. *Cell Tissue Res.* (1991) 266:343–57. doi: 10.1007/bf00318190
 15. Nässel DR, Shiga S, Wikstrand EM, Rao KR. Pigment-dispersing hormone-immunoreactive neurons and their relation to serotonergic neurons in the blowfly and cockroach visual system. *Cell Tissue Res.* (1991) 266:511–23. doi: 10.1007/BF00318593
 16. Nässel DR, Shiga S, Mohrherr CJ, Rao KR. Pigment-dispersing hormone-like peptide in the nervous system of the flies *Phormia* and *Drosophila*: immunocytochemistry and partial characterization. *J Comp Neurol.* (1993) 331:183–98. doi: 10.1002/cne.903310204
 17. Helfrich-Förster C, Homberg U. Pigment-dispersing hormone-immunoreactive neurons in the nervous system of wild-type *Drosophila melanogaster* and of several mutants with altered circadian rhythmicity. *J Comp Neurol.* (1993) 337:177–90. doi: 10.1002/cne.903370202
 18. Stengl M, Homberg U. Pigment-dispersing hormone-immunoreactive neurons in the cockroach *Leucophaea maderae* share properties with circadian pacemaker neurons. *J Comp Physiol A.* (1994) 175:203–13. doi: 10.1007/bf00215116
 19. Petri B, Stengl M, Würden S, Homberg U. Immunocytochemical characterization of the accessory medulla in the cockroach *Leucophaea maderae*. *Cell Tissue Res.* (1995) 282:3–19. doi: 10.1007/BF00319128
 20. Würden S, Homberg U. Immunocytochemical mapping of serotonin and neuropeptides in the accessory medulla of the locust, *Schistocerca gregaria*. *J Comp Neurol.* (1995) 362:305–19. doi: 10.1002/cne.903620302
 21. Reischig T, Stengl M. Morphology and pigment-dispersing hormone immunocytochemistry of the accessory medulla, the presumptive circadian pacemaker of the cockroach *Leucophaea maderae*: a light- and electron-microscopic study. *Cell Tissue Res.* (1996) 285:305–19. doi: 10.1007/s004410050648
 22. Reischig T, Stengl M. Optic lobe commissures in a three-dimensional brain model of the cockroach *Leucophaea maderae*: a search for the circadian coupling pathways. *J Comp Neurol.* (2002) 443:388–400. doi: 10.1002/cne.10133
 23. Reischig T, Stengl M. Ultrastructure of pigment-dispersing hormone-immunoreactive neurons in a three-dimensional model of the accessory medulla of the cockroach *Leucophaea maderae*. *Cell Tissue Res.* (2003) 314:421–35. doi: 10.1007/s00441-003-0772-7
 24. Petri B, Stengl M. Pigment-dispersing hormone shifts the phase of the circadian pacemaker of the cockroach *Leucophaea maderae*. *J Neurosci.* (1997) 17:4087–93. doi: 10.1523/JNEUROSCI.17-11-04087.1997
 25. Helfrich-Förster C, Stengl M, Homberg U. Organization of the circadian system in insects. *Chronobiol Int.* (1998) 15:567–94. doi: 10.3109/07420529808993195
 26. Renn SCP, Park JH, Rosbash M, Hall JC, Taghert PH. A *pdf* neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in *Drosophila*. *Cell.* (1999) 99:791–802. doi: 10.1016/s0092-8674(00)81676-1
 27. Persson MGS, Eklund MB, Dircksen H, Muren JE, Nässel DR. Pigment-dispersing factor in the locust abdominal ganglia may have roles as circulating neurohormone and central neuromodulator. *J Neurobiol.* (2001) 48:19–41. doi: 10.1002/neu.1040
 28. Bloch G, Solomon SM, Robinson GE, Fahrbach SE. Patterns of PERIOD and pigment-dispersing hormone immunoreactivity in the brain of the European honeybee (*Apis mellifera*): age- and time-related plasticity. *J Comp Neurol.* (2003) 464:269–84. doi: 10.1002/cne.10778
 29. Sehadová H, Sauman I, Sehnal F. Immunocytochemical distribution of pigment-dispersing hormone in the cephalic ganglia of polyneopteran insects. *Cell Tissue Res.* (2003) 312:113–25. doi: 10.1007/s00441-003-0705-5
 30. Reischig T, Petri B, Stengl M. Pigment-dispersing hormone (PDH)-immunoreactive neurons form a direct coupling pathway between the bilaterally symmetric circadian pacemakers of the cockroach *Leucophaea maderae*. *Cell Tissue Res.* (2004) 318:553–64. doi: 10.1007/s00441-004-0927-1
 31. Wei H, Yasar H, Funk NW, Giese M, Baz E-S, Stengl M. Signaling of pigment-dispersing factor (PDF) in the Madeira cockroach *Rhyarobia maderae*. *PLoS ONE.* (2014) 9:e108757. doi: 10.1371/journal.pone.0108757
 32. Stengl M, Arendt A. Peptidergic circadian clock circuits in the Madeira cockroach. *Curr Opin Neurobiol.* (2016) 41:44–52. doi: 10.1016/j.conb.2016.07.010
 33. Beer K, Kolbe E, Kahana NB, Yayon N, Weiss R, Menegazzi P, et al. Pigment-dispersing factor-expressing neurons convey circadian information in the honey bee brain. *Open Biol.* (2018) 8:170224. doi: 10.1098/rsob.170224
 34. Gestrich J, Giese M, Shen W, Zhang Y, Voss A, Popov C, et al. Sensitivity to pigment-dispersing factor (PDF) is cell-type specific among PDF-expressing circadian clock neurons in the Madeira cockroach. *J Biol Rhythms.* (2018) 33:35–51. doi: 10.1177/0748730417739471
 35. Vosko AM, Schroeder A, Loh DH, Colwell CS. Vasoactive intestinal peptide and the mammalian circadian system. *Gen Comp Endocrinol.* (2007) 152:165–75. doi: 10.1016/j.ygcen.2007.04.018
 36. Veenstra JA. Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen Comp Endocrinol.* (2010) 167:86–103. doi: 10.1016/j.ygcen.2010.02.010
 37. Veenstra JA. Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. *Gen Comp Endocrinol.* (2011) 171:160–75. doi: 10.1016/j.ygcen.2011.01.005
 38. Mayer G, Hering L, Stosch JM, Stevenson PA, Dircksen H. Evolution of pigment-dispersing factor neuropeptides in Panarthropoda: insights from Onychophora (velvet worms) and Tardigrada (water bears). *J Comp Neurol.* (2015) 523:1865–85. doi: 10.1002/cne.23767
 39. Semmens DC, Mirabeau O, Moghul I, Pancholi MR, Wurm Y, Elphick MR. Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. *Open Biol.* (2016) 6:150224. doi: 10.1098/rsob.150224
 40. Giribet G, Edgecombe GD. Current understanding of Ecdysozoa and its internal phylogenetic relationships. *Integr Comp Biol.* (2017) 57:455–66. doi: 10.1093/icb/ixc072
 41. Janssen T, Husson SJ, Lindemans M, Mertens I, Rademakers S, Donck KV, et al. Functional characterization of three G protein-coupled receptors for pigment dispersing factors in *Caenorhabditis elegans*. *J Biol Chem.* (2008) 283:15241–9. doi: 10.1074/jbc.M709060200
 42. Janssen T, Husson SJ, Meelkop E, Temmerman L, Lindemans M, Verstraelen K, et al. Discovery and characterization of a conserved pigment dispersing factor-like neuropeptide pathway in *Caenorhabditis elegans*. *J Neurochem.* (2009) 111:228–41. doi: 10.1111/j.1471-4159.2009.06323.x
 43. Rao KR, Riehm JP. Pigment-dispersing hormones. *Ann N Y Acad Sci.* (1993) 680:78–88. doi: 10.1111/j.1749-6632.1993.tb19676.x
 44. Klein JM, Mohrherr CJ, Sleutels F, Riehm JP, Rao KR. 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.* (1994) 205:410–6. doi: 10.1006/bbrc.1994.2680
 45. Desmoucelles-Carette C, Sellos D, Van Wormhoudt A. Molecular cloning of the pigment dispersing hormone in a crustacean. *Ann N Y Acad Sci.* (1998) 839:395–6. doi: 10.1111/j.1749-6632.1998.tb10810.x
 46. Yang W-J, Aida K, Nagasawa H. Characterization of chromatophoretropic neuropeptides from the kuruma prawn *Penaeus japonicus*. *Gen Comp Endocrinol.* (1999) 114:415–24. doi: 10.1006/gcen.1999.7266
 47. Ohira T, Nagasawa H, Aida K. Molecular cloning of cDNAs encoding two pigment-dispersing hormones and two corresponding genes from the kuruma prawn (*Penaeus japonicus*). *Mar Biotechnol.* (2002) 4:463–70. doi: 10.1007/s10126-002-0042-9

48. Ohira T, Tsutsui N, Kawazoe I, Wilder MN. Isolation and characterization of two pigment-dispersing hormones from the whiteleg shrimp, *Litopenaeus vannamei*. *Zool Sci.* (2006) 23:601–6. doi: 10.2108/zsj.23.601
49. Herrero A, Romanowski A, Meelkop E, Caldart C, Schoofs L, Golombek DA. Pigment-dispersing factor signaling in the circadian system of *Caenorhabditis elegans*. *Genes Brain Behav.* (2015) 14:493–501. doi: 10.1111/gbb.12231
50. Barrios A, Ghosh R, Fang C, Emmons SW, Barr MM. PDF-1 neuropeptide signaling modulates a neural circuit for mate-searching behavior in *C. elegans*. *Nat Neurosci.* (2012) 15:1675–82. doi: 10.7554/eLife.36547.001
51. Sato S, Chuman Y, Matsushima A, Tominaga Y, Shimohigashi Y, Shimohigashi M. A circadian neuropeptide, pigment-dispersing factor-PDF, in the last-summer cicada *Meimuna opalifera*: cDNA cloning and immunocytochemistry. *Zool Sci.* (2002) 19:821–8. doi: 10.2108/zsj.19.821
52. Honda T, Matsushima A, Sumida K, Chuman Y, Sakaguchi K, Onoue H, et al. Structural isoforms of the circadian neuropeptide PDF expressed in the optic lobes of the cricket *Gryllus bimaculatus*: immunocytochemical evidence from specific monoclonal antibodies. *J Comp Neurol.* (2006) 499:404–21. doi: 10.1002/cne.21112
53. Wen C-J, Lee H-J. Mapping the cellular network of the circadian clock in two cockroach species. *Arch Insect Biochem Physiol.* (2008) 68:215–31. doi: 10.1002/arch.20236
54. Závodská R, Wen C-J, Hrdý I, Sauman I, Lee H-J, Sehna F. Distribution of corazonin and pigment-dispersing factor in the cephalic ganglia of termites. *Arthropod Struct Dev.* (2008) 37:273–86. doi: 10.1016/j.asd.2008.01.005
55. Wei H, El Jundi B, Homberg U, Stengl M. Implementation of pigment-dispersing factor-immunoreactive neurons in a standardized atlas of the brain of the cockroach *Leucophaea maderae*. *J Comp Neurol.* (2010) 518:4113–33. doi: 10.1002/cne.22471
56. Ikeno T, Numata H, Goto SG, Shiga S. Involvement of the brain region containing pigment-dispersing factor-immunoreactive neurons in the photoperiodic response of the bean bug, *Riptortus pedestris*. *J Exp Biol.* (2014) 217:453–62. doi: 10.1242/jeb.091801
57. Mertens I, Vandingenen A, Johnson EC, Shafer OT, Li W, Trigg JS, et al. PDF receptor signaling in *Drosophila* contributes to both circadian and geotactic behaviors. *Neuron.* (2005) 48:213–9. doi: 10.1016/j.neuron.2005.09.009
58. Hyun S, Lee Y, Hong S-T, Bang S, Paik D, Kang J, et al. *Drosophila* GPCR Han is a receptor for the circadian clock neuropeptide PDF. *Neuron.* (2005) 48:267–78. doi: 10.1016/j.neuron.2005.08.025
59. Lear BC, Merrill CE, Lin J-M, Schroeder A, Zhang L, Allada R. A G protein-coupled receptor, *groom-of-PDF*, is required for PDF neuron action in circadian behavior. *Neuron.* (2005) 48:221–7. doi: 10.1016/j.neuron.2005.09.008
60. Shafer OT, Kim DJ, Dunbar-Yaffe R, Nikolaev VO, Lohse MJ, Taghert PH. Widespread receptivity to neuropeptide PDF throughout the neuronal circadian clock network of *Drosophila* revealed by real-time cyclic AMP imaging. *Neuron.* (2008) 58:223–37. doi: 10.1016/j.neuron.2008.02.018
61. Im SH, Taghert PH. PDF receptor expression reveals direct interactions between circadian oscillators in *Drosophila*. *J Comp Neurol.* (2010) 518:1925–45. doi: 10.1002/cne.22311
62. Hering L, Henze MJ, Kohler M, Kelber A, Bleidorn C, Leschke M, et al. Opsins in Onychophora (velvet worms) suggest a single origin and subsequent diversification of visual pigments in arthropods. *Mol Biol Evol.* (2012) 29:3451–8. doi: 10.1093/molbev/mss148
63. Reid AL. Review of the Peripatopsidae (Onychophora) in Australia, with comments on peripatopsid relationships. *Invertebr. Taxon.* (1996) 10:663–936. doi: 10.1071/IT9960663
64. Baer A, Mayer G. Comparative anatomy of slime glands in Onychophora (velvet worms). *J Morphol.* (2012) 273:1079–88. doi: 10.1002/jmor.20044
65. I5k C. The i5K Initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *J Hered.* (2013) 104:595–600. doi: 10.1093/jhered/est050
66. Thomas GWC, Dohmen E, Hughes DST, Murali SC, Poelchau M, Glastad K, et al. Gene content evolution in the arthropods. *Genome Biol.* (2020) 21:15. doi: 10.1186/s13059-019-1925-7
67. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* (2011) 17:10–2. doi: 10.14806/ej.17.1.200
68. Smeds L, Künstner A. ConDeTri-a content dependent read trimmer for Illumina data. *PLoS ONE.* (2011) 6: e26314. doi: 10.1371/journal.pone.0026314
69. Peng Y, Leung HC, Yiu S-M, Lv M-J, Zhu X-G, Chin FY. IDBA-tran: a more robust de novo de Bruijn graph assembler for transcriptomes with uneven expression levels. *Bioinformatics.* (2013) 29:i326–34. doi: 10.1093/bioinformatics/btt219
70. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* (1997) 25:3389–402. doi: 10.1093/nar/25.17.3389
71. Pándy-Szekeres G, Munk C, Tsonkov TM, Mordalski S, Harpsøe K, Hauser AS, et al. GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Res.* (2018) 46:D440–6. doi: 10.1093/nar/gkx1109
72. Frickey T, Lupas A. CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics.* (2004) 20:3702–4. doi: 10.1093/bioinformatics/bth444
73. Koski LB, Golding GB. The closest BLAST hit is often not the nearest neighbor. *J. Mol. Evol.* (2001) 52:540–2. doi: 10.1007/s002390010184
74. Hoare SR. Mechanisms of peptide and nonpeptide ligand binding to class B G-protein-coupled receptors. *Drug Discov Today.* (2005) 10:417–27. doi: 10.1016/S1359-6446(05)03370-2
75. Hollenstein K, De Graaf C, Bortolato A, Wang M-W, Marshall FH, Stevens RC. Insights into the structure of class B GPCRs. *Trends Pharmacol Sci.* (2014) 35:12–22. doi: 10.1016/j.tips.2013.11.001
76. Sonnhammer EL, Eddy SR, Durbin R. Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins.* (1997) 28:405–20. doi: 10.1002/(SICI)1097-0134(199707)28:3<405::AID-PROT10>3.0.CO;2-L
77. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* (2016) 44:D279–85. doi: 10.1093/nar/gkv1344
78. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* (2017) 20:1160–6. doi: 10.1093/bib/bbx108
79. Dress AWM, Flamm C, Fritsch G, Grünwald S, Kruspe M, Prohaska SJ, et al. Noisy: identification of problematic columns in multiple sequence alignments. *Algorithm Mol Biol.* (2008) 3:7. doi: 10.1186/1748-7188-3-7
80. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* (2014) 30:1312–3. doi: 10.1093/bioinformatics/btu033
81. Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* (2011) 39:W475–8. doi: 10.1093/nar/gkr201
82. Treffkorn S, Kahnke L, Hering L, Mayer G. Expression of NK cluster genes in the onychophoran *Euperipatoides rowelli*: implications for the evolution of NK family genes in nephrozoans. *EvoDevo.* (2018) 9:17. doi: 10.1186/s13227-018-0105-2
83. Nikolaev VO, Bünemann M, Hein L, Hannawacker A, Lohse MJ. Novel single chain cAMP sensors for receptor-induced signal propagation. *J Biol Chem.* (2004) 279:37215–8. doi: 10.1074/jbc.C400302200
84. Ponsioen B, Zhao J, Riedl J, Zwartkruis F, Van Der Krogt G, Zaccolo M, et al. Detecting cAMP-induced Epac activation by fluorescence resonance energy transfer: Epac as a novel cAMP indicator. *EMBO Rep.* (2004) 5:1176–80. doi: 10.1038/sj.embor.7400290
85. Klarenbeek J, Goedhart J, Van Batenburg A, Groenewald D, Jalink K. Fourth-generation Epac-based FRET sensors for cAMP feature exceptional brightness, photostability and dynamic range: characterization of dedicated sensors for FLIM, for ratiometry and with high affinity. *PLoS ONE.* (2015) 10:e0122513. doi: 10.1371/journal.pone.0122513
86. Stefanini M, De Martino C, Zamboni L. Fixation of ejaculated spermatozoa for electron microscopy. *Nature.* (1967) 216:173–4. doi: 10.1038/216173a0
87. Robson E, Lockwood A, Ralph R. Composition of the blood in Onychophora. *Nature.* (1966) 209:533. doi: 10.1038/209533a0
88. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Meth.* (2012) 9:676–82. doi: 10.1038/nmeth.2019

89. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Meth.* (2012) 9:671–5. doi: 10.1038/nmeth.2089
90. Prinz A, Diskar M, Herberg FW. Application of bioluminescence resonance energy transfer (BRET) for biomolecular interaction studies. *ChemBioChem.* (2006) 7:1007–12. doi: 10.1002/cbic.200600048
91. Schulze J, Neupert S, Schmidt L, Predel R, Lamkemeyer T, Homberg U, et al. Myoinhibitory peptides in the brain of the cockroach *Leucophaea maderae* and colocalization with pigment-dispersing factor in circadian pacemaker cells. *J Comp Neurol.* (2012) 520:1078–97. doi: 10.1002/cne.22785
92. Nery LEM, De Lauro Castrucci AM. Pigment cell signalling for physiological color change. *Comp Biochem Physiol A.* (1997) 118:1135–44. doi: 10.1016/S0300-9629(97)00045-5
93. De Rooij J, Zwartkruis FJT, Verheijen MHG, Cool RH, Nijman SMB, Wittinghofer A, et al. Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature.* (1998) 396:474–7. doi: 10.1038/24884
94. Nikolaev VO, Lohse MJ. Monitoring of cAMP synthesis and degradation in living cells. *Physiology.* (2006) 21:86–92. doi: 10.1152/physiol.00057.2005
95. Rao KR, Riehm JP. The pigment-dispersing hormone family: chemistry, structure-activity relations, and distribution. *Biol Bull.* (1989) 177:225–9. doi: 10.2307/1541937
96. De Kleijn DPV, Linck B, Klein JM, Weidemann WM, Keller R, Van Herp F. Structure and localization of mRNA encoding a pigment dispersing hormone (PDH) in the eyestalk of the crayfish *Orconectes limosus*. *FEBS Lett.* (1993) 321:251–5. doi: 10.1016/0014-5793(93)80119-f
97. Desmoucelles-Carette C, Sellos D, Van Wormhoudt A. Molecular cloning of the precursors of pigment dispersing hormone in crustaceans. *Biochem Biophys Res Commun.* (1996) 221:739–43. doi: 10.1006/bbrc.1996.0666
98. Rao KR. Crustacean pigmentary-effector hormones: chemistry and functions of RPCH, PDH, and related peptides. *Am Zool.* (2001) 41:364–79. doi: 10.1093/icb/41.3.364
99. Bulau P, Meisen I, Schmitz T, Keller R, Peter-Katalinić J. Identification of neuropeptides from the sinus gland of the crayfish *Orconectes limosus* using nanoscale on-line liquid chromatography tandem mass spectrometry. *Mol Cell Proteomics.* (2004) 3:558–64. doi: 10.1074/mcp.M300076-MCP200
100. Fu Q, Goy MF, Li L. Identification of neuropeptides from the decapod crustacean sinus glands using nanoscale liquid chromatography tandem mass spectrometry. *Biochem Biophys Res Commun.* (2005) 337:765–78. doi: 10.1016/j.bbrc.2005.09.111
101. Martin C, Gross V, Hering L, Tepper B, Jahn H, De Sena Oliveira I, et al. The nervous and visual systems of onychophorans and tardigrades: learning about arthropod evolution from their closest relatives. *J Comp Physiol A.* (2017) 203:565–90. doi: 10.1007/s00359-017-1186-4
102. Taghert PH, Nitabach MN. Peptide neuromodulation in invertebrate model systems. *Neuron.* (2012) 76:82–97. doi: 10.1016/j.neuron.2012.08.035
103. Zoli M, Jansson A, Syková E, Agnati LF, Fuxe K. Volume transmission in the CNS and its relevance for neuropsychopharmacology. *Trends Pharmacol Sci.* (1999) 20:142–50. doi: 10.1016/S0165-6147(99)01343-7
104. Sanchez S. Cellules neurosécrétrices et organes infracérébraux de *Peripatopsis moseleyi* Wood (Onychophores) et neurosécrétion chez *Nymphon gracile* Leach (Pycnogonides). *Arch Zool Exp Gen.* (1958) 96:57–62.
105. Eriksson BJ, Tait NN, Norman JM, Budd GE. An ultrastructural investigation of the hypocerebral organ of the adult *Euperipatoides kanangrensis* (Onychophora, Peripatopsidae). *Arthropod Struct Dev.* (2005) 34:407–18. doi: 10.1016/j.asd.2005.03.002
106. Eakin RM, Westfall JA. Fine structure of the eye of peripatus (Onychophora). *Z Zellforsch Mikrosk Anat.* (1965) 68:278–300. doi: 10.1007/BF00342434
107. Kirwan JD, Graf J, Smolka J, Mayer G, Henze MJ, Nilsson DE. Low-resolution vision in a velvet worm (Onychophora). *J Exp Biol.* (2018) 221:jeb175802. doi: 10.1242/jeb.175802
108. Dakin WJ. The eye of peripatus. *Q J Microsc Sci.* (1921) 65:163–72.
109. Mayer G. Structure and development of onychophoran eyes—what is the ancestral visual organ in arthropods? *Arthropod Struct Dev.* (2006) 35:231–45. doi: 10.1016/j.asd.2006.06.003
110. Lavallard R, Campiglia S. Contribution à l'hématologie de *Peripatus acacioi* Marcus et Marcus (Onychophora) — I. Structure et ultrastructure des hémocytes. *Ann Sci Nat Zool Biol Anim.* (1975) 17:67–92.
111. Ravindranath MH. Onychophorans and myriapods. In: Ratcliffe NA, Rowley AF, editors. *Invertebrate Blood Cells: Arthropods to Urochordates, Invertebrates and Vertebrates Compared*. London: Academic Press (1981). p. 329–52.
112. Kusche K, Ruhberg H, Burmester T. A hemocyanin from the Onychophora and the emergence of respiratory proteins. *Proc Natl Acad Sci USA.* (2002) 99:10545–8. doi: 10.1073/pnas.152241199
113. Manton SM. Studies on the Onychophora, IV. The passage of spermatozoa into the ovary in Peripatopsis and the early development of the ova. *Philos Trans R Soc B Biol Sci.* (1938) 228:421–44. doi: 10.1098/rstb.1938.0001
114. Park JH, Helfrich-Förster C, Lee G, Liu L, Rosbash M, Hall JC. Differential regulation of circadian pacemaker output by separate clock genes in *Drosophila*. *Proc Natl Acad Sci USA.* (2000) 97:3608–13. doi: 10.1073/pnas.97.7.3608
115. Vafopoulou X, Terry KL, Steel CGH. The circadian timing system in the brain of the fifth larval instar of *Rhodnius prolixus* (hemiptera). *J Comp Neurol.* (2010) 518:1264–82. doi: 10.1002/cne.22274
116. Schneider N-L, Stengl M. Pigment-dispersing factor and GABA synchronize cells of the isolated circadian clock of the cockroach *Leucophaea maderae*. *J Neurosci.* (2005) 25:5138–47. doi: 10.1523/jneurosci.5138-a-04.2005
117. Schneider N-L, Stengl M. Extracellular long-term recordings of the isolated accessory medulla, the circadian pacemaker center of the cockroach *Leucophaea maderae*, reveal ultradian and hint circadian rhythms. *J Comp Physiol A.* (2007) 193:35–42. doi: 10.1007/s00359-006-0169-7
118. Wei H, Stengl M. Light affects the branching pattern of peptidergic circadian pacemaker neurons in the brain of the cockroach *Leucophaea maderae*. *J Biol Rhythms.* (2011) 26:507–17. doi: 10.1177/0748730411419968
119. Wei H, Stengl M. Ca²⁺-dependent ion channels underlying spontaneous activity in insect circadian pacemaker neurons. *Eur J Neurosci.* (2012) 36:3021–9. doi: 10.1111/j.1460-9568.2012.08227.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Martin, Hering, Metzendorf, Hormann, Kasten, Fuhrmann, Werckenthin, Herberg, Stengl and Mayer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Comparative Aspects of Structure and Function of Cnidarian Neuropeptides

Toshio Takahashi*

Suntory Foundation for Life Sciences, Bioorganic Research Institute, Kyoto, Japan

Cnidarians are early-branching animals in the eukaryotic tree of life. The phylum Cnidaria are divided into five classes: Scyphozoa (true jellyfish), Cubozoa (box jellyfish), Hydrozoa (species, *Hydra* and *Hydractinia*), Anthozoa (sea anemone, corals, and sea pen), and Staurozoa (stalked jellyfish). Peptides play important roles as signaling molecules in development and differentiation in cnidaria. For example, cnidaria use peptides for cell-to-cell communication. Recent discoveries show that *Hydra* neuropeptides control several biological processes including muscle contraction, neuron differentiation, and metamorphosis. Here, I describe the structure and functions of neuropeptides in *Hydra* and other cnidarian species. I also discuss that so-called primitive nervous system of *Hydra* is in more complex than generally believed. I also discuss how cnidaria use peptides for communication among cells rather than in higher animals.

OPEN ACCESS

Edited by:

James A. Carr,
Texas Tech University, United States

Reviewed by:

Hervé Tostivint,
Muséum National d'Histoire
Naturelle, France
Joao Carlos dos Reis Cardoso,
University of Algarve, Portugal

*Correspondence:

Toshio Takahashi
takahashi@sunbor.or.jp

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 03 March 2020

Accepted: 30 April 2020

Published: 27 May 2020

Citation:

Takahashi T (2020) Comparative
Aspects of Structure and Function of
Cnidarian Neuropeptides.
Front. Endocrinol. 11:339.
doi: 10.3389/fendo.2020.00339

Keywords: *Hydra*, Cnidaria, neuropeptide, metamorphosis, myoactivity, interstitial stem cell, neuron differentiation

INTRODUCTION

Molecular phylogenetic studies show that Cnidaria are the sister group of Bilateria. Ancestral Cnidarians diverged over 500 million years ago in animal evolution. Despite the long course of evolution, the nervous systems of cnidarians are differentiated (1). Cnidarian species are also mainly classified into two groups according to the unique life cycle, the anthozoans and medusozoans (1). Anthozoa lives exclusively as polyps. Among medusozoans, Cubozoa and Scyphozoa predominantly live as medusae. On the other hand, Hydrozoa usually follows a life cycle where the species alternate between these two forms except for *Hydra* and *Hydractinia*. Staurozoa lives exclusively as polyps.

Cnidaria such as *Hydra* are composed of multiple cell types that represent the fundamental architecture of multicellular organisms. *Hydra* exhibits a simple body plan with a head and tentacles on one end and a foot on the opposite end of a hypostome. The gastric region is located between the head and foot. The body is composed of two layers, ectoderm and endoderm, which are separated by an extracellular matrix, the mesoglea. The cells of both epithelial layers also function as muscle cells. *Hydra* also have multipotent interstitial stem cells, which differentiate into nerve cells (2), nematocytes (2), gland cells (3), and germ cells (4). *Hydra* as a member of cnidaria represents an attractive model to understand axial pattern formation into head- and foot-specific tissues.

The nervous system of *Hydra* is simple and is composed of a nerve net that extends throughout the animal. The cnidarian nervous system is mainly peptidergic (5). Classical molecules such as acetylcholine also contribute to the *Hydra* nervous system (6).

Peptides play important roles as hormones and neurotransmitters and they are involved in the maintenance of a variety of developmental stages. However, little is known about whether they

are involved in differentiation and development. In *Hydra*, theoretical model suggests that small molecules such as peptides are transported to establish morphogenetic gradients that regulate patterning processes. To systematically identify and characterize peptide signaling molecules, we started the Hydra Peptide Project (7). By using the strategy illustrated in **Figure 1**, many peptides were extracted and purified with successive steps of high performance liquid chromatography (HPLC). Signaling peptides were identified by their effect on the gene expression profile of *Hydra* by using differential display (DD)-PCR. Positive peptides were chemically synthesized, the synthetic peptides were used for biological assays including behavioral (muscle contraction), neuron differentiation, and others. Furthermore, introduction of the *Hydra* Expressed Sequence Tag (EST) Project has enabled us to identify transcripts for novel peptides even more efficiently (**Figure 1**) (8).

The primary aim of the present review is to describe the structures and functions of peptide signaling molecules such as neuropeptides in cnidarians, especially in *Hydra*.

CNIDARIAN NEUROPEPTIDES

FMRFamide-Like Peptides (FLPs)

The peptide FMRFamide was originally purified from the cerebral ganglion of the clam *Macrocallista nimbosa* (9, 10). Other mollusks and members of most other phyla express peptides with a similar sequence. FMRFamides are categorized into two groups depending on the structural similarity with FMRFamide. The first category consists of FMRFamide-related peptides (FaRPs), which include encode for multiple peptides with the C-terminal FMRFamide or FLRFamide (11). The second category of FMRFamides includes FLPs, which are peptides that have only the RFamide sequence at C-termini (12). Therefore,

FaRPs and all other RFamide peptides are considered FLPs. Krajniak (13) excellently reviewed FaRPs in invertebrates. This overview primarily focuses on cnidarian FLPs.

A variety of FLPs are expressed in the evolutionarily ancient nervous system of cnidarians (**Table 1**). Peptides with GRFamide at the C-terminus have been found in a scyphozoan (the jellyfish *Cyanea lamarckii*) (15), three hydrozoans (*Hydra magnipapillata*, the hydromedusa *Polyorchis penicillatus*, and *Hydractinia echinata*) (16–21), and an anthozoan (the sea anemone *Anthopleura elegantissima*) (14), whereas peptides with TRFamide and/or RRFamide at the C-terminus have been described in another anthozoan (the sea anemone *Nematostella vectensis*) (22). All mature neuropeptides are controlled by highly regulated secretion pathways. Usually, a precursor of a neuropeptide is incorporated as a preprohormone in the endoplasmic reticulum, where it is converted into a prohormone. Next, prohormones move to the Golgi apparatus for endoproteolysis and/or amidation at the C-terminus, which results in the final active peptide. FLPs have been identified in numerous cnidarians. A *Calliactis parasitica* cDNA includes 19 copies of Antho-RFamide (**Table 1**), two copies of FQGRFamide, and one copy of YVPGRYamide (24). Two cDNAs have been isolated from *Anthopleura elegantissima*; one cDNA includes 13 copies of Antho-RFamide (**Table 1**) and nine other FLPs; the second cDNA includes 14 copies of Antho-RFamide and eight other FLPs (25). *Renilla koellikeri* has 36 copies of Antho-RFamide (26). A *Polyorchis penicillatus* cDNA includes one copy of Pol-RFamide I (**Table 1**) and 11 copies of Pol-RFamide II (**Table 1**), in addition to another predicted FLP (27). In *Hydra*, RFamides are spliced from three different preprohormones called A, B, and C (**Figure 2A**). Preprohormone-A includes six Hydra-RFamides (Hydra-RFamide I–VI) (**Figure 2A**) (**Table 1**) (19). Preprohormone-B has one copy of Hydra-RFamide I and

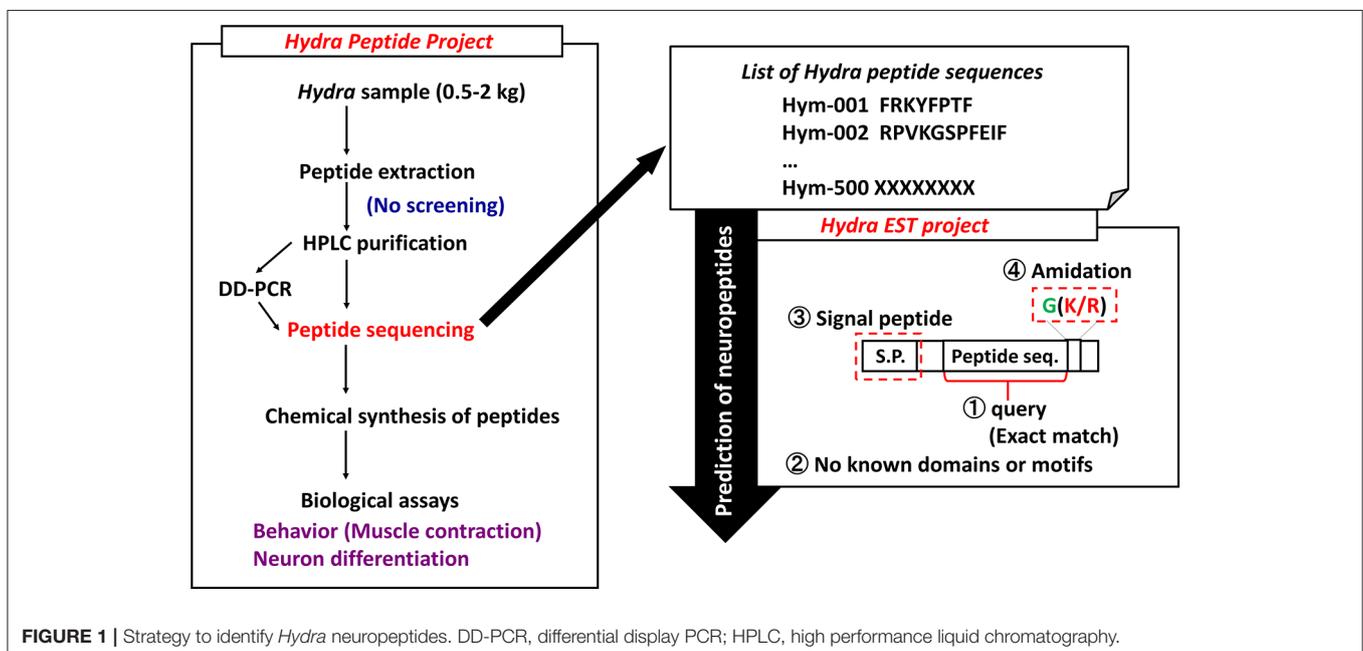


FIGURE 1 | Strategy to identify *Hydra* neuropeptides. DD-PCR, differential display PCR; HPLC, high performance liquid chromatography.

TABLE 1 | FLPs in cnidarians.

Name	Peptide sequence	Species	Reference
Antho-RFamide	pQGRFamide	Anthopleura elegantissima	(14)
Cyanea-RFamide I	pQWLRGRFamide	Cyanea lamarckii	(15)
Cyanea-RFamide II	pQPLWSGRFamide		
Cyanea-RFamide III	GRFamide		
Pol-RFamide I	pQLLGGRFamide	Polyorchis penicillatus	(16)
Pol-RFamide II	pQWLKGRFamide		(17)
Hydra-RFamide I	pQWLGGRFamide	Hydra magnipapillata	(18)
Hydra-RFamide II	pQWFNGRFamide		
Hydra-RFamide III	KPHLRGRFamide		
Hydra-RFamide IV	HLRGRFamide		
Hydra-RFamide V	pQLMSGRFamide	Hydra magnipapillata	(19)
Hydra-RFamide VI	pQLMRGRFamide		
Hydra-RFamide VII	pQLLRGRFamide		
Hydra-RFamide VIII	KPHYRGRFamide		
Hydra-RFamide IX	HYRGRFamide		
Hydra-RFamide X	KPHLIGRFamide	Hydra magnipapillata	(20)
Hydra-RFamide XI	pQLMTGRFamide		
He-RFamide	pQWLKGRFamide	Hydractinia echinata	(21)
Nv-RFamide I	pQITRFamide	Nematostella vectensis	(22)
Nv-RFamide II	VPRRFamide		
RFamide (ID:17)	pQGRFGREDQGRFamide		(23)

pQ, pyroglutamate.

Hydra-RFamide II and three probable Hydra-RFamides (Hydra-RFamide V, VII, and VIII) (**Figure 2A**) (19). Preprohormone-C has one copy of Hydra-RFamide I and seven copies of additional neuropeptide sequences (one copy of pQWFSGRFamide and six copies of pQWLSGRFamide) (**Figure 2A**) (19). In *Hydractinia echinata*, one copy of He-RFamide is present (**Table 1**) (21). In *Nematostella vectensis*, three FLPs [Nv-RFamide I and II and RFamide (ID:17)] are present (**Table 1**) (22, 23). Collectively, precursor-encoding cnidarian FLP cDNAs yield many neuropeptides with great structural diversity, indicating that they have great functional diversity as well.

Cnidarian FLPs control several functions, such as muscle contraction, feeding, sensation, reproduction, metamorphosis, and movement of larvae. Treatment of the sea anemone *Calliactis parasitica* with Antho-RFamide increases muscle tone, contraction amplitude, and contraction of slow muscles (28). In individual autozooid polyps of *Renilla koellikeri*, Antho-RFamide also leads to tonic contractions in the rachis and peduncle (29). In *Hydra*, Hydra-RFamide III mediates pumping of the peduncle in a dose-dependent manner (30).

FMRFamide activates a Na⁺ channel identified in snails (31, 32). Three cation channel subunits of the degenerin (DEG)/epithelial Na⁺ channel (ENaC) gene family were cloned from the freshwater polyp *Hydra magnipapillata* and designated *Hydra* Na⁺ channel (HyNaC)2–4 (33). Subsequently, a novel subunit, designated HyNaC5, was cloned, and expression of the gene was shown to be co-localized with HyNaC2 and HyNaC3 at the base of the tentacles (34). Co-injection of HyNaC5 with HyNaC2 and HyNaC3 genes in *Xenopus* oocytes strongly enhances the current amplitude after peptide application and increases the affinity of the channel for Hydra-RFamide I and II (34). HyNaC2/3/5 is assembled into a functional heterotrimeric channel that is activated by Hydra-RFamide I with high affinity. The experimental data of HyNaCs suggested that secretion of Hydra-RFamide I and/or II induces tentacle contraction, perhaps during feeding (33, 34). Seven additional HyNaC subunits, HyNaC6–HyNaC12, were cloned, and all belong to the DEG/ENaC gene family (35). These subunits and the four originally identified subunits self-assemble in *Xenopus* oocytes to create 13 different ion channels that show high-affinity binding of Hydra-RFamide I and II. The HyNaC inhibitor, diminazene, slows tentacle movement in *Hydra*. Because *Hydra* express multiple peptide-gated ion channels with a restricted number of FLPs as ligands (35), FLPs may be important for fast transmission at neuro-muscular junction in cnidarians. The function of Hydra-RFamide IV in *Hydra* is unknown.

Highly specialized mechanoreceptor cells, called stinging cells or nematocytes, that are important for capturing prey and defense are present in cnidarians (36). Two- and three-cell synaptic pathways, including synapses between nematocytes and nearby nerve cells, are present in the epidermis of the sea anemone tentacles (37, 38). Cnidarian sensory function is probably mediated by FLPs, as evidenced by anti-FMRFamide and anti-RFamide antibody staining in the tentacles of four classes of cnidaria. Thus, FLPs probably mediate chemosensory regulation of cnidocyte discharge (39). The epidermal sensory cells of the spot ocellus in *Aurelia* are also positive for FMRFamide (40), which may inhibit spontaneous firing of nematocytes.

FLPs also play a key role in cnidarian reproduction, larval movement, and metamorphosis. Reproduction of colonial octocorals such as *Renilla koellikeri* occurs via spawning and exfoliation. Intact gamete follicles are released into the water during spawning. These follicles rupture during exfoliation, releasing the gametes. Antho-RFamide is present in ciliated neurons in the epithelium of follicles of *Renilla koellikeri* and induces exfoliation of the epithelium and subsequent release of the gametes into water (41). Light enhances the potency of Antho-RFamide (41).

The colony-forming marine hydroid, *Hydractinia echinata*, is closely related to freshwater *Hydra*. Fertilized eggs of this species undergo rapid cleavage divisions for about 1 day and develop into spindle-shaped planula larvae in about 3 days (42). Planula larvae are capable of migrating toward light (43), and they metamorphose into adult polyps when they receive appropriate environmental stimuli (44, 45). Hydra-RFamide I inhibits the migration of planula larvae, thus modulating phototaxis by

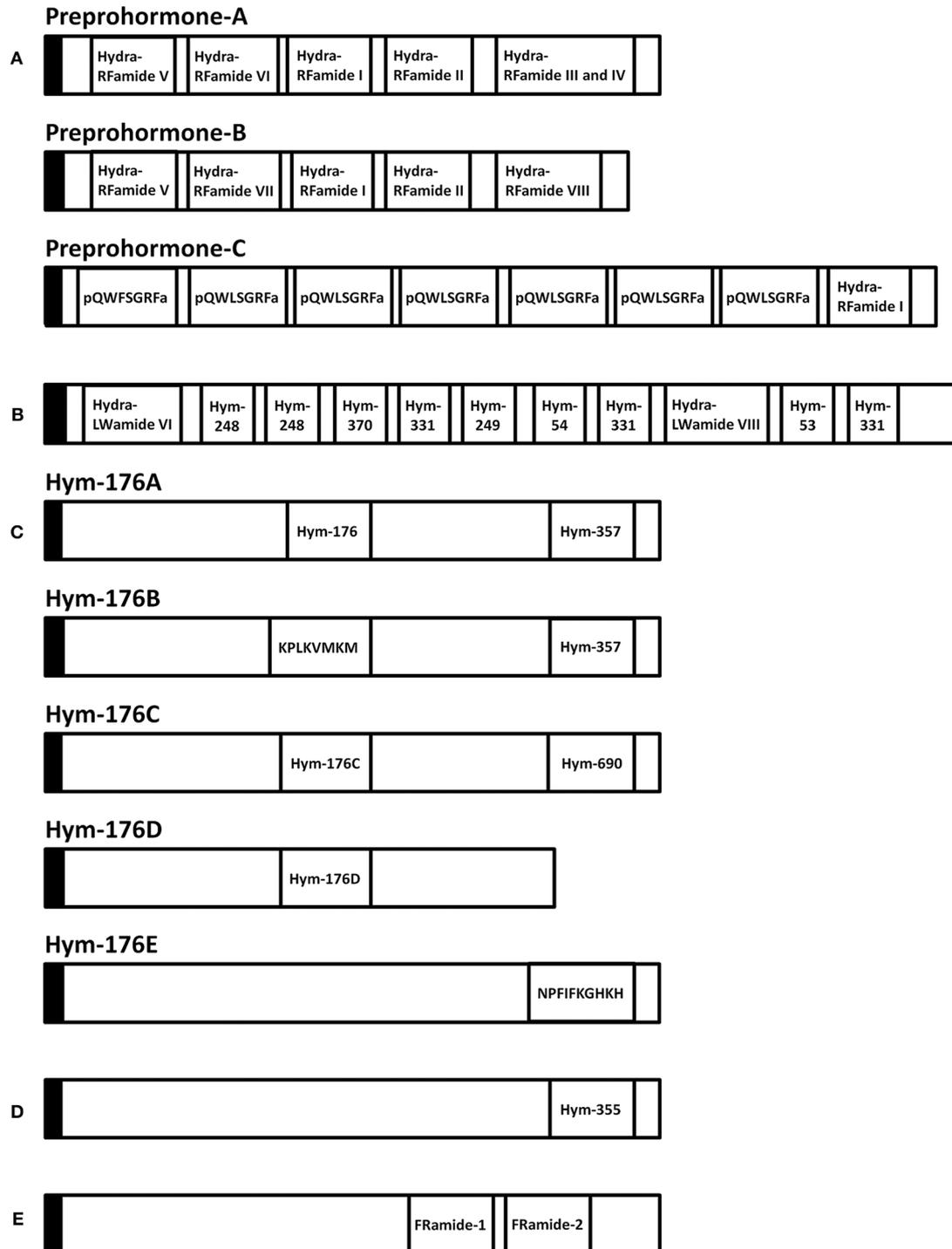


FIGURE 2 | Schematic representation of eleven preprohormones in *Hydra*. **(A)** Preprohormone-A contains unprocessed Hydra-RFamide I, II, III, IV, V, and VI. Preprohormone-B contains unprocessed Hydra-RFamide I, II, V, VII, and VIII. Preprohormone-C contains Hydra-RFamide I and two putative neuropeptide sequences (pQWFSGRFa and pQWLSGRFa). **(B)** The GLWamide precursor contains unprocessed GLWamides (Hym-53, 54, 248, 249, 331, and 370) and two putative neuropeptide sequences (Hydra-LWamide VI and VIII). **(C)** The Hym-176 precursor (Hym-176A) contains one copy of unprocessed Hym-176 and Hym-357. Hym-176B contains Hym-357 and one putative neuropeptide (KPLKVMKM). Hym-176C and Hym-176D contain one copy of Hm-176-homologous peptide (Hym-176C and Hym-176D), respectively. Hym-176C also contains one unprocessed Hym-690. Hym-176E contains one putative neuropeptide sequence (NPFIFKGHKH). **(D)** The Hym-355 precursor contains one unprocessed Hym-355. **(E)** The FRamide precursor contains unprocessed FRamide-1 and -2. Black box: signal sequence, a: amide.

TABLE 2 | GLWamide family peptides in cnidarians.

Name	Peptide sequence	Species	Reference
MMA	pQQPGLWamide	<i>Anthopleura elegantissima</i>	(48)
Hym-53	NPYPGLWamide	<i>Hydra magnipapillata</i>	(7, 49, 50)
Hym-54	GPMTGLWamide		
Hym-248	EPLPIGLWamide		
Hym-249	KPIPGLWamide		
Hym-331	GPPPGLWamide		
Hym-338	GPP ^h PGLWamide		
Hym-370	KPNAYKGKLIPLWamide		
Hydra-LWamide VI	RLPLGLWamide		
Hydra-LWamide VIII	pQPPIGMWamide		
He-LWamide I	pQRPPGLWamide	<i>Hydractinia echinata</i>	(51)
He-LWamide II	KPPGLWamide		
Ae-LWamide I	pQQHGLWamide	<i>Actinia equine</i>	(51)
Ae-LWamide II	pQNPGLWamide		
Ae-LWamide III	pQPGLWamide		
Ae-LWamide IV	pQKAGLWamide		
Ae-LWamide V	pQLGLWamide		
Ae-LWamide VI	RSRIGLWamide		
Ae-MWamide	pQDLDIGMWamide		
MMA	pQQPGLWamide		
As-LWamide I	pQQAGLWamide	<i>Anemonia sulcata</i>	(51)
As-LWamide II	pQHPGLWamide		
As-IWamide	pQERIGLWamide		
Ae-LWamide II	pQNPGLWamide		
MMA	pQQPGLWamide		

pQ, pyroglutamate; ^hP, hydroxyproline.

inhibiting myomodulation (43). Metamorphosis is also inhibited by this peptide, leading to the suggestion that the function of endogenous FLPs is to stabilize the larval stage (46). Thus, FLPs may play a role in regulating the movement of planula larvae prior to metamorphosis, possibly linking movement to chemotactic or phototactic processes (47). Sensory neurons that express FLPs are present in planula larvae, suggesting that migration and metamorphosis of these animals may be mediated by secretion of endogenous neuropeptides in response to environmental stimuli.

GLWamides

GLWamides are characterized by certain features at their N- and C-termini. Most GLWamides have a GLWamide motif at the C-terminus (Table 2). Seven GLWamide peptides are found in *Hydra*, and they include X-Pro or X-Pro-Pro at their N-termini (Table 2) (7, 49). In the anthozoan *Anthopleura elegantissima*, Metamorphosin A (MMA) that is a member of the GLWamide family has an N-terminal pyroglutamine (Table 2) (48). Both N-terminal modifications produce resistance to aminopeptidase (52).

GLWamide cDNAs are found in other cnidarians as well. A cDNA encoding a preprohormone with 11 immature peptide sequences, nine of which are unique, was cloned from *Hydra magnipapillata* (Figure 2B) (50). The corresponding gene includes one copy of Hym-53 (NPYPGLWamide), Hym-54 (GPMTGLWamide), Hym-249 (KPIPGLWamide), and Hym-370 (KPNAYKGKLIPLWamide); two copies of Hym-248 (EPLPIGLWamide); and t copies of Hym-331 (GPPPGLWamide), as well as two additional putative GLWamides (Hydra-LWamide VI and VIII) (Table 2). Hydra-LWamide VIII is predicted from this cDNA and probably includes GMWamide at the C-terminus (50). A cDNA encoding GLWamides has been cloned from *Hydractinia echinata* (51) and includes one copy of He-LWamide I and 17 copies of He-LWamide II (Table 2). Two unique cDNAs have been cloned from the anthozoans *Actinia equine* and *Anemonia sulcata* (51). The *Actinia* gene includes one copy of MMA, Ae-LWamide IV, Ae-LWamide V, Ae-LWamide VI, and Ae-MWamide; two copies of Ae-LWamide I and Ae-LWamide III; and four copies of Ae-LWamide II (Table 2). In contrast, the *Anemonia* gene has one copy of MMA, Ae-LWamide II, and As-IWamide; two copies of As-LWamide II; and four copies of As-LWamide I (Table 2) (51). The preprohormones of anthozoans but not hydrozoans include MMA. The peptide is probably a prototype of the family (53). Two other peptides that are possibly generated from the preprohormones of *Actinia* and *Anemonia* are likely processed into -GMWamide (Ae-MWamide) and -GIWamide (As-IWamide) at their C-terminus (Table 2). Whether these two peptides and Hydra-LWamide VIII belong to the GLWamide family is uncertain, as substitution of the Leu residue in GLWamide with Met or Ile results in deactivation of contractile activity in the retractor muscle of the anthozoan *Anthopleura fuscoviridis* (54).

The various species of *Hydractinia* generally live on hermit crab shells. The *Hydractinia* life cycle includes a planula larval stage but no medusa stage. After attaching to snail shells, planula larvae undergo MMA-induced metamorphosis and become polyps after about 1 week (48, 55). MMA thus works as a neurohormone to mediate development in addition to its roles as a neurotransmitter and neuromodulator. In *Hydractinia serrata*, *Hydra* GLWamides also cause polyp development from planula larvae (7, 49). A common GLWamide sequence is required to induce metamorphosis in *Hydractinia*, and the GLWamide terminus and amidation are essential and specific for inducing metamorphosis (56). Substitution of Gly in GLWamide with another common amino acid (except Cys) decreases or completely inhibits potency of the peptide, and substitution of Leu or Trp in GLWamide with another common amino acid (except Cys) partially or completely blocks its potency for muscle contraction in *Anthopleura fuscoviridis* (54). The precise mechanism of how these peptides induce metamorphosis remains to be determined. Bacteria in the environment produce a chemical that can induce larvae to undergo metamorphosis (48). This chemical signal probably affects sensory neurons in the planula larvae that secrete endogenous GLWamides to induce a phenotypic change in the surrounding epithelial cells. *Hydra* lack a larval stage and develop directly into adults from

embryos, and thus, how GLWamide peptides function during early development in *Hydra* is unclear.

Motile planula larvae play a role in sexual reproduction in reef-building corals. These larvae undergo complex metamorphosis after adhering to a substrate, and a juvenile coral colony results. In *Acropora*, Hym-248 induces dose-dependent metamorphosis of nearly 100% of planula larvae into polyps (57). However, the effect of Hym-248 on metamorphosis is species-specific (57, 58). A Hym-248-specific receptor appears to exist in *Acropora*. The receptor may serve as a barrier to ensure specification in corals. In *Hydractinia*, the peptide for their receptors is loose. The possible receptors may share certain common sequences and binding sites. Hym-248-related peptide(s) are expected to be identified in *Acropora*.

In *Hydra*, all GLWamide peptides serve as myoactive peptides to activate sphincter muscle contraction and bud detachment (7). The sphincter muscle is involved in bud detachment. To test myoactivity in *Hydra*, nerve-free tissue of epithelial hydra is typically used (59, 60). When normal *Hydra* that contains nerve cells is treated with the peptides, they exhibit the same effect as epithelial *Hydra*. GLWamides are synthesized and expressed in nerve cells (49) and thus function as neurotransmitters or neuromodulators at the neuromuscular junction. Hym-248, which is a *Hydra* GLWamide, induces both bud detachment and body elongation (49). Muscle tissue in *Hydra* runs perpendicular to the ectodermal and endodermal epithelial cells. Hym-248 may bind to two different types of receptors, one that binds all types of GLWamides and one that specifically binds to Hym-248. Substance P (SP) is a highly conserved member of the tachykinin peptide family that is widely expressed throughout the animal kingdom (61). It binds to tachykinin receptors [neurokinin-1, 2, and 3 receptor (NK1R, NK2R, and NK3R)] that belong to G-protein-coupled receptors (GPCRs). SP preferentially activates NK1R. This difference of specificity against other tachykinin peptides can be accounted for the conformational flexibility of the short and linear peptides and ligand binding affinity for the receptors (62). Probably, the features of both receptors for Hym-248 may depend on the ligand structure and binding affinity for receptors.

All GLWamide family peptides enhance retractor muscle contraction of *Anthopleura* (49). Nerve cells in the sea anemone retractor muscle stain strongly with a GLWamide motif-specific antibody, similar to the nervous system of *Hydra* (49).

In *Hydractinia echinata*, GLWamide and RFamide neuropeptides modulate planula larva migration. He-LWamide II, which is a GLWamide, induces migration by extending the active period (43). GLWamides and FLPs antagonize one another to modulate migration of *Hydractinia echinata* planula larvae.

In hydrozoan jellyfish, maturation of oocytes and spawning are initiated by light-dark cycles in natural conditions within 1 second (63). Exposure to Hym-53 for < 2 min is sufficient for oocyte maturation and spawning (64). Thus, neuropeptides function as hormones that modulate the first step that determines whether oocytes undergo irreversible meiosis after light exposure.

TABLE 3 | Hym-176, Hym-357, and their related peptides in *Hydra*.

Name	Peptide sequence	Species	Reference
Hym-176	APFIFPGPKVamide	<i>Hydra magnipapillata</i>	(65)
Hym-176C	YPFYNQNPKVamide		(66)
Hym-176D	NPKNKNFMIFVGPVKVamide		(66)
Hym-357	KPAFLFKGYKPamide	<i>Hydra magnipapillata</i>	(20, 66)
Hym-690	KPLYLFGYKPamide		(20, 66)

Hym-176 (APFIFPGPKVamide)

Hym-176 was a newly identified as a neuropeptide (Table 3) (7, 65). The gene that encodes Hym-176 is strongly expressed in the neurons of the lower peduncle and weakly expressed in the gastric region (67). This peptide induces contraction of the ectodermal muscle in *Hydra* (65). This region-specific neuron subset correlates with the myoactivity of the peptide. Hym-176 has no effects on muscle contraction in *Anhtopleura*, metamorphosis in *Hydractinia*, and oocyte maturation and spawning in *Cytaeis*. And also, the gene encoding the peptide (Hym-176A) is just isolated from *Hydra* (Figure 2C) (66, 67). Thus, the peptide is species-specific.

The gene that encodes Hym-176 also encodes a second peptide, Hym-357 (KPAFLFKGYKPamide) (Figure 2C) (Table 3). This neuropeptide was identified in a screen for myoactive peptides (20). Detailed observations suggest that Hym-357 neurons activate other neurons to release neurotransmitters for induction of muscle contraction.

To identify the homologous gene that encodes Hym-176, Noro and coworkers found four candidate genes in the freshwater polyp *Hydra magnipapillata* (66). No authentic Hym-176 is present in the four paralogues (Figure 2C) (66). The cDNAs, Hym-176C and Hym-176D, encode one copy of a Hym-176-homologous peptide (Figure 2C) (Table 3). Hym-357 is encoded in both the gene that encodes Hym-176 and the gene that encodes Hym-176B (Figure 2C) (66). Hym-176C encodes Hym-690 (KPLYLFGYKPamide), which is closely related to Hym-357 (Figure 2C) (Table 3) (20). Hym-176E appears not to have Hym-176- and Hym-357-related peptides (Figure 2C). The function of Hym-176C and D and Hym-690 has not yet been characterized in *Hydra*.

Hym-355 (FPQSFLPRGamide)

Hym-355 is a member of the PRXamide family of peptides that have PRXamide at their C-terminal region (Figure 2D) (Table 4) (68) and are subdivided into three groups in invertebrates: (a) neuropeptides that induce pheromone biosynthesis (70) and similar molecules, (b) small cardioactive peptides (71–73), and (c) antho-RPamide (52) and similar molecules. Antho-RPamide (LPPGPLPRamide) is located in neurons of sea anemones and induces tentacle contraction. Thus, the peptide is involved in neurotransmission. PRXamide peptides have been identified in many invertebrates. Hym-355 is homologous to members of sub-group (c), including LPPGPLPRamide (*Anthopleura elegantissima*) (Table 4), AAPLPRamide (*Urechis unicinctus*) (74), QPPLPRamide

(*Helix pomatia*), and pQPPLRYamide (*Helix pomatia*) (75). GPRGGRATEFGPRGamide and GPRGGREVNLEGPRGamide both have PRGamide at their C-termini and are expressed in the sea anemone *Nematostella vectensis* (Table 4) (23). The gene encoding the PRGamides is expressed in neurons (23), indicating that the PRGamides are neuropeptides.

Oxytocin-vasopressin superfamily peptides are neuropeptides synthesized in the hypothalamus and secreted from the posterior pituitary gland in mammals. Whether cnidarians express oxytocin/vasopressin superfamily peptides remains an open question in the field of comparative physiology

TABLE 4 | PRXamide peptides in cnidarians.

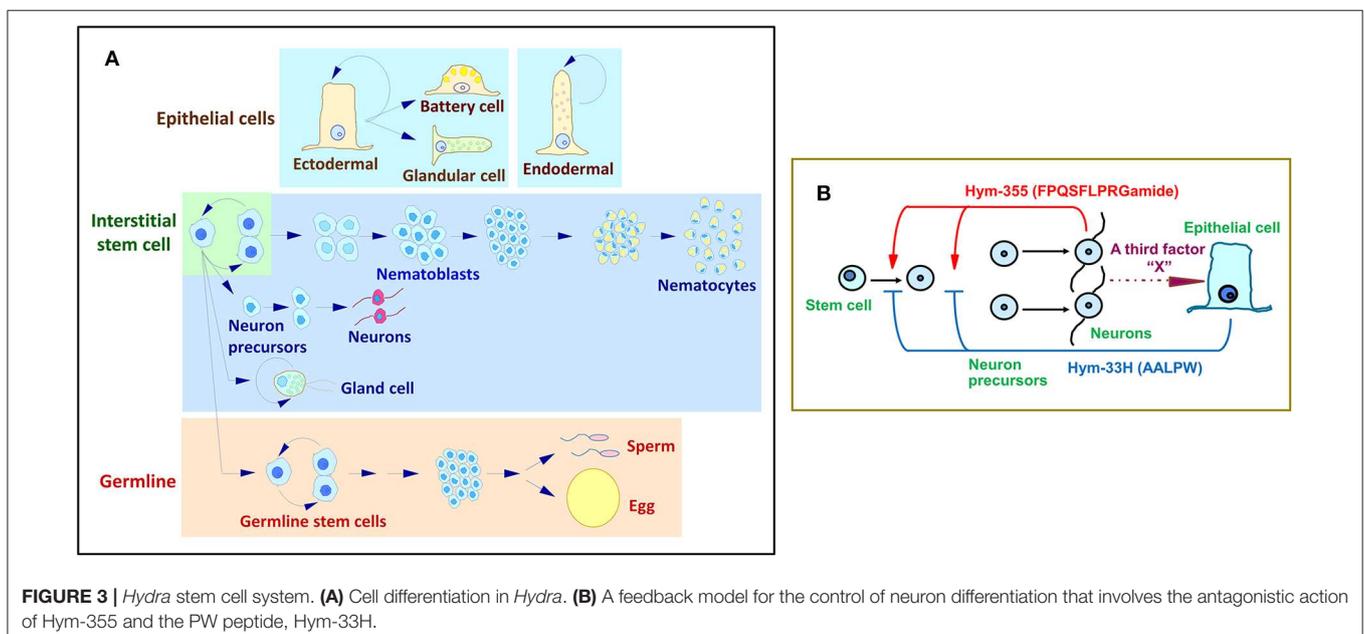
Name	Peptide sequence	Species	Reference
Hym-355	FPQSFLPRGamide	<i>Hydra magnipapillata</i>	(68)
PRGamide (ID:11)	GPRGGRATEFGPRGamide	<i>Nematostella vectensis</i>	(23)
PRGamide (ID:12)	GPRGGREVNLEGPRGamide		
Antho-RPamide	LPPGPLPRPamide	<i>Anthopleura elegantissima</i>	(52)
MIHs	WPRPamide	<i>Clytia hemisphaerica</i>	(69)
	WPRAamide	<i>Cladonema pacificum</i>	
	RPRPamide		
	RPRAamide		
	RPRGamide		
	RPRYamide		

MIHs, maturation-inducing hormones.

of nervous systems. Immunohistochemical staining suggests that oxytocin/vasopressin superfamily peptides exist in the *Hydra* nervous system (76, 77). Morishita and coworkers (78) purified two peptides, Hym-355 and SFLPRGamide, from *Hydra magnipapillata* using HPLC fractionation and immunologic assays. They demonstrated that the antigen for vasopressin-like immunoreactivity is Hym-355 in the *Hydra* nervous system. The C-terminal region of Hym-355 (PRGamide) is identical to that of vasopressin. Neither antibody against the two peptides discriminates one peptide from the other. Thus, Koizumi et al. (79) performed immunohistochemistry with an anti-Hym-355 antibody and demonstrated immunoreactivity in the nerve rings of *Cladonema radiatum* and *Turritopsis nutricula*. However, whether Hym-355 functions as a neurohypophysial hormone is not well-understood.

The tissue of *Hydra* undergoes continuous renewal (Figure 3A). The number of neurons remains constant. Two groups of peptides, Hym-355 and PW family peptides, regulate this state (7, 68, 80). PW family peptides share the same sequence of Pro-Trp and are identified as epitheliopeptides (81).

Hym-355 increases early neuron differentiation, and Hym-33H (AALPW) blocks neuron differentiation (68, 80). Simultaneous treatment with Hym-355 and Hym-33H results in a normal level of neuron differentiation. Taken together, the observations are consistent with a feedback model that modulates the homeostasis of neuronal differentiation in *Hydra* (Figure 3B) (68). This model suggests that Hym-355, which is synthesized by neurons, enhances early neuronal differentiation. To balance differentiation, epithelial cells produce PW peptides. A third factor termed as X in Figure 3B may control synthesis and secretion of PW family peptides. Hym-355, PW peptides, and the putative third factor may work together to maintain a constant neuronal density in *Hydra*. Hym-355 induces interstitial



stem cells to undergo neuron differentiation and also induces retractor muscle contraction in the sea anemone *Anthopleura fuscoviridis* (68).

A member of the GLWamide family, Hym-53 (NPYPGLWamide) (Table 2), and Hym-355 induce oocyte maturation and spawning, but the effect of Hym-53 is stronger than that of Hym-355. Hym-355-like immunoreactivity is observed in neurons in *Cytaeis* (63). Possibly, neurons expressing Hym-53- and Hym-355-like peptides contribute downstream of light receptors in oocyte maturation and spawning in *Cytaeis*. Takeda and coworkers demonstrated that endogenous peptides including W/RPRamide peptides are involved in oocyte maturation (Table 4) (69). RPRYamide, RPRGamide, WPRAamide, and RPRAamide may act as maturation-inducing hormones (MIHs) (Table 4) (69). Takeda et al. (69) also demonstrated that MIH peptides are synthesized by neurons in the gonad, and probably act on the oocyte surface. They propose that hydrozoan MIHs and neuropeptides are evolutionally

linked to regulate reproduction upstream of MIHs in bilaterian species (69).

FRamide Family

During research aimed at systematic identification of peptide signaling molecules in *Hydra* (7), two novel neuropeptides, FRamide-1 (IPTGTLIFRamide) and FRamide-2 (APGSLIFRamide), were identified (Table 5) (82). Among *Hydra* EST and genome databases (8), we can rapidly identify peptide transcripts and their genes. The two peptides and the single gene encoding both peptides were identified using this exact approach (Figure 2E).

FRamide-1 (IPTGTLIFRamide) and FRamide-2 (APGSLIFRamide) exhibit opposing effects even though they are encoded by the same gene. The former peptide evokes body column elongation due to endodermal muscle contraction, whereas the latter peptide evokes body column contraction due to ectodermal muscle contraction (82). Two explanations for these seemingly contradictory observations are possible. One possibility is that the release of each peptide is differentially regulated (83, 84), and the other possibility is that each peptide is processed in a different type of neuron (85). Additionally, the opposing effects of FRamide family peptides may be ligand binding affinity for one receptor (62). In higher animals, most neuropeptides bind to GPCRs that are localized at the

TABLE 5 | FRamide family peptides in Hydra.

Name	Peptide sequence	Species	Reference
FRamide-1	IPTGTLIFRamide	<i>Hydra magnipapillata</i>	(82)
FRamide-2	APGSLIFRamide		

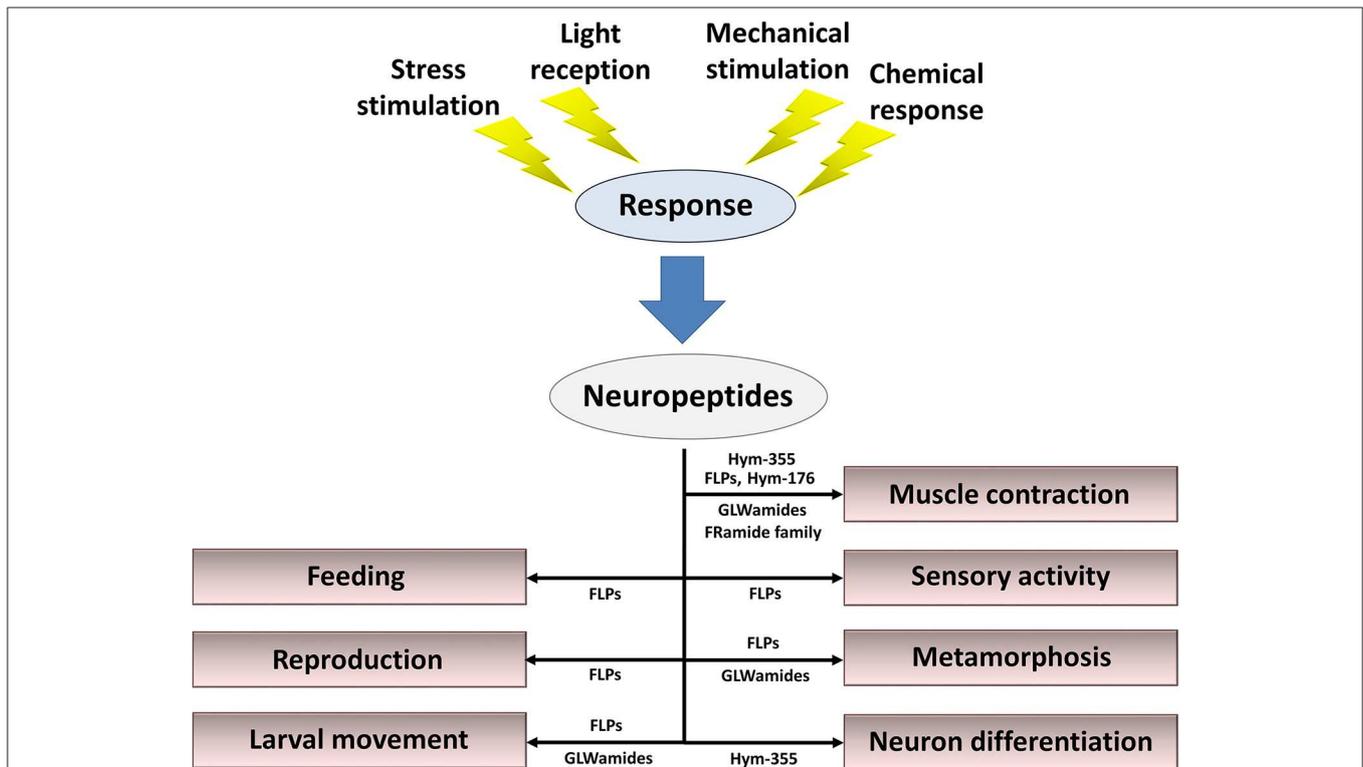


FIGURE 4 | Summary of roles for neuropeptides in the control of behavior, reproduction, metamorphosis, and tissue maintenance. Cnidarian peptide signaling molecules function together and/or separately to maintain the organism's lifestyle in response to stress stimulation, light reception, mechanical stimulation, and chemical stimulation.

target cell. To understand the opposite effects, identification of FRamide-specific receptors on the target cells is important.

CONCLUSION

Neuropeptides released from nerve cells in response to a variety of stimuli are mandatory for fine-tuned regulation of behavior, reproduction, metamorphosis, and tissue maintenance (Figure 4). Here, I described 57 types of neuropeptides so far identified in cnidarians. However, the study of neuropeptides is still in its infancy. Additional novel peptides will likely be found (86), including neuropeptides, thus enabling elucidation of the mechanisms that regulate the physiology and development of cnidarians and increasing our understanding of peptide function in other species.

It is important to elucidate functional interaction between neuropeptides and receptors for the verification of their biological roles and evolutionary processes. However, no receptors for the neuropeptides remain to be identified in *Hydra* and cnidarians. Recently, Shiraishi and coworkers developed the machine-learning-assisted strategy for the identification of novel peptide–receptor pairs (87). As they indicate the multiplicity of use of the strategy, it is worth to use the strategy for increasing the receptor (especially GPCR) repertoire as many as possible on *Hydra* and cnidarians. When neuropeptide-GPCR pairs are efficiently and systematically elucidated in a

phylogenetically critical Hydrozoa *Hydra magnipapillata*, *Hydra* provides cnidarian perspectives into evolution of GPCRs.

The cells of *Hydra* are well-characterized and belong to the epithelial cell lineage and the interstitial stem cell lineage (Figure 3A). However, knowledge of the molecules and biochemical mechanisms of the cells remains limited. The single-cell RNA sequencing technique sheds light on the complete molecular diversity of the cells in *Hydra*. Siebert and coworkers (88) applied this approach to the homeostatic adult *Hydra*. They drew a molecular map of the *Hydra* nervous system and unlocked the door toward understanding the molecular basis of morphogenesis and regeneration in *Hydra*.

AUTHOR CONTRIBUTIONS

TT wrote the original review manuscript draft.

FUNDING

This work was supported by a Grant-in-Aid for Scientific Research (C) to TT (Grant Number 17K07495).

ACKNOWLEDGMENTS

The author acknowledges a grant from the JSPS KAKENHI to TT (Grant Number 17K07495).

REFERENCES

- Galliot B, Quiquand M, Ghila L, de Rosa R, Miljkovic-Licina M, Chera S. Origins of neurogenesis, a cnidarian view. *Dev. Biol.* (2009) 332:2–24. doi: 10.1016/j.ydbio.2009.05.563
- David CN, Gierer A. A cell cycle kinetics and development of *Hydra attenuate*. III. Nerve and nematocyte differentiation. *J Cell Sci.* (1974) 16:3375.
- Schmidt T, David CN. Gland cells in *Hydra*: Cell cycle kinetics and development. *J Cell Sci.* (1986) 85:197–215.
- Bosch TCG, David CN. Stem cells of *Hydra magnipapillata* can differentiate into somatic cells and germ line cells. *Dev Biol.* (1987) 121:182–91. doi: 10.1016/0012-1606(87)90151-5
- Grimmelikhuijzen CJP, Leviev I, Carstensen K. Peptides in the nervous system of cnidarians: structure, function and biosynthesis. *Int Rev Cytol.* (1996) 167:37–89. doi: 10.1016/S0074-7696(08)61345-5
- Kass-Simon G, Pierobon P. Cnidarian chemical neurotransmission, an updated overview. *Comp Biochem Physiol A Mol Integr Physiol.* (2007) 146:9–25. doi: 10.1016/j.cbpa.2006.09.008
- Takahashi T, Muneoka Y, Lohmann Y, deHaro MSL, Solleder G, Bosch TCG, et al. Systematic isolation of peptide signaling molecules regulating development in *Hydra*: LWamide and PW families. *Proc Natl Acad Sci USA.* (1997) 94:1241–6. doi: 10.1073/pnas.94.4.1241
- Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, et al. The dynamic genome of *Hydra*. *Nature.* (2010) 464:592–6. doi: 10.1038/nature08830
- Price DA, Greenberg MJ. Structure of a molluscan cardioexcitatory neuropeptide. *Science.* (1977) 97:670–1. doi: 10.1126/science.877582
- Price DA, Greenberg MJ. Purification and characterization of a cardioexcitatory neuropeptide from the central ganglia of a bivalve mollusk. *Prep Biochem.* (1977) 7:261–81. doi: 10.1080/00327487708061643
- Price DA, Greenberg MJ. The hunting of the FaRPs: The distribution of FMRFamide-related peptides. *Biol Bull.* (1989) 177:198–205. doi: 10.2307/1541933
- Espinoza E, Carrigan M, Thomas SG, Shaw G, Edison AS. A statistical view of FMRFamide neuropeptide diversity. *Mol Neurobiol.* (2000) 21:35–56. doi: 10.1385/MN:21:1-2:035
- Krajniak KG. Invertebrate FMRFamide related peptides *Protein Pept Lett.* (2013) 20:647–70. doi: 10.2174/0929866511320060005
- Grimmelikhuijzen CJP, Graff D. Isolation of pyroGlu-Gly-Arg-Phe-NH₂ (Antho-RFamide), a neuropeptide from sea anemones. *Proc Natl Acad Sci USA.* (1986) 83:9817–21. doi: 10.1073/pnas.83.24.9817
- Moosler A, Rhinehart KL, Grimmelikhuijzen CJP. Isolation of three novel peptides, the Cyanea-RFamides I-III, from scyphomedusae. *Biochem Biophys Res Commun.* (1997) 236:743–9. doi: 10.1006/bbrc.1997.7022
- Grimmelikhuijzen CJP, Hahn M, Rhinehart KL, Spencer AN. Isolation of pyroGlu-Leu-Leu-Gly-Gly-Arg-Phe-NH₂ (Pol-RFamide), a novel neuropeptide from hydromedusae. *Brain Res.* (1988) 475:198–203. doi: 10.1016/0006-8993(88)90219-3
- Grimmelikhuijzen CJP, Rhinehart KL, Spencer AN. Isolation of the neuropeptide less than Glu-Trp-Leu-Lys-Gly-Arg-Phe-NH₂ (Pol-RFamide II) from the hydromedusa *Polyorchis penicillatus*. *Biochem Biophys Res Commun.* (1992) 183:375–82. doi: 10.1016/0006-291X(92)90491-3
- Moosler A, Rhinehart KL, Grimmelikhuijzen CJP. Isolation of four novel neuropeptides, the Hydra-RFamides I-IV, from *Hydra magnipapillata*. *Biochem Biophys Res Commun.* (1996) 229:596–602. doi: 10.1006/bbrc.1996.1849
- Darmer D, Hauser F, Nothacker HP, Bosch TCG, Williamson M, Grimmelikhuijzen CJP. Three different prohormones yield a variety of Hydra-RFamide (Arg-Phe-NH₂) neuropeptide in *Hydra magnipapillata*. *Biochem J.* (1998) 332:403–12. doi: 10.1042/bj3320403
- Fujisawa T. *Hydra* peptide project 1993–2007. *Dev Growth Differ.* (2008) 50:5257–68. doi: 10.1111/j.1440-169X.2008.00997.x
- Gajewski M, Schmutzler C, Plickert G. Structure of neuropeptide precursors in cnidarian. *Ann N Y Acad Sci.* (1998) 839:311–5. doi: 10.1111/j.1749-6632.1998.tb10782.x

22. Anctil M. Chemical transmission in the sea anemone *Nematostella vectensis*: a genomic perspective. *Comp Biochem Physiol D*. (2009) 4:268–89. doi: 10.1016/j.cbd.2009.07.001
23. Hayakawa E, Watanabe H, Menschaert G, Holstein TW, Baggerman G, Schoofs L. A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*. *PLoS ONE*. (2019) 14:e0215185. doi: 10.1371/journal.pone.0215185
24. Darmer D, Schmutzler C, Diekhoff D, Grimmelikhuijzen CJP. Primary structure of the precursor for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH₂). *Proc Natl Acad Sci USA*. (1991) 88:2555–9. doi: 10.1073/pnas.88.6.2555
25. Schmutzler C, Darmer D, Diekhoff D, Grimmelikhuijzen CJP. Identification of a novel type of processing sites in the precursor for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH₂) from *Anthopleura elegantissima*. *J Biol Chem*. (1992) 267:22534–41.
26. Reinscheid RK, Grimmelikhuijzen CJP. Primary structure of the precursor for the anthozoan neuropeptide Antho-RFamide from *Renilla kollikeri*: evidence for unusual processing enzymes. *J Neurochem*. (1994) 62:1214–22. doi: 10.1046/j.1471-4159.1994.62031214.x
27. Schmutzler C, Diekhoff D, Grimmelikhuijzen CJP. The primary structure of the Pol-RFamide neuropeptide precursor protein from the hydromedusa *Polyorchis penicillatus* indicates a novel processing proteinase activity. *Biochem J*. (1994) 299:431–6. doi: 10.1042/bj2990431
28. McFarlane ID, Graff D, Grimmelikhuijzen CJP. Excitatory actions of Antho-RFamide, an anthozoan neuropeptide, on muscles and conducting systems in the sea anemone. *Calliactis parasitica*. *J Exp Biol*. (1987) 133:157–68.
29. Anctil M, Grimmelikhuijzen CJP. Excitatory action of the native neuropeptide Antho-RFamide in muscles in the pennatulid *Renilla kollikeri*. *Gen Pharmacol*. (1989) 20:381–4. doi: 10.1016/0306-3623(89)90277-2
30. Shimizu H, Fujisawa T. Peduncle of *Hydra* and the heart of higher organisms share a common ancestral origin. *Genesis*. (2003) 36:182–6. doi: 10.1002/gene.10213
31. Cottrell GA, Green KA, Davis NW. The neuropeptide Phe-Met-Arg-Phe-NH₂(FMRFamide) can activate a ligand-gated ion channel in *Helix* neurons. *Pflugers Arch*. (1990) 416:612–4. doi: 10.1007/BF00382698
32. Lingueglia E, Champigny G, Lazdunski M, Barbry P. Cloning of the amiloride-sensitive FMRFamide peptide-gated sodium channel. *Nature*. (1995) 378:730–3. doi: 10.1038/378730a0
33. Golubovic A, Kuhn A, Williamson M, Kalbacher H, Holstein TW, Grimmelikhuijzen CJP, et al. A peptide-gated ion channel from the freshwater polyp *Hydra*. *J Biol Chem*. (2007) 282:35098–103. doi: 10.1074/jbc.M706849200
34. Dürrnagel S, Kuhn A, Tsiairis CD, Williamson M, Kalbacher H, Grimmelikhuijzen CJP, et al. Three homologous subunits form a high affinity peptide-gated ion channel in *Hydra*. *J Biol Chem*. (2010) 285:11958–65. doi: 10.1074/jbc.M109.059998
35. Assmann M, Kuhn A, Dürrnagel S, Holstein TW, Gründer S. The comprehensive analysis of DEG/ENaC subunits in *Hydra* reveals a large variety of peptide-gated channels, potentially involved in neuromuscular transmission. *BMC Biol*. (2014) 12:84. doi: 10.1186/s12915-014-0084-2
36. Tardent P. The cnidarian cnidocyte, a high tech cellular weaponry. *BioEssays*. (1995) 17:351–62. doi: 10.1002/bies.950170411
37. Holtmann M, Thurm U. Mono- and oligo-vesicular synapses and their connectivity in a Cnidarian sensory epithelium (*Coryne tubulosa*). *J Comp Neurol*. (2001) 432:537–49. doi: 10.1002/cne.1118
38. Westfall JA, Elliott CF, Carlin RW. Ultrastructural evidence for two-cell and three-cell neural pathways in the tentacle epidermis of the sea anemone *Aiptasia pallida*. *J Morphol*. (2002) 251:83–92. doi: 10.1002/jmor.1075
39. Anderson PA, Thompson LE, Money Penny CG. Evidence for a common pattern of peptidergic innervations of cnidocytes. *Biol Bull*. (2004) 207:141–6. doi: 10.2307/1543588
40. Nakanishi N, Hartenstein V, Jacobs DK. Development of the rhopalial nervous system in *Aurelia* sp. 1 (Cnidaria, Scyphozoa). *Dev Genes Evol*. (2009) 219:301–17. doi: 10.1007/s00427-009-0291-y
41. Tremblay ME, Henry J, Anctil M. Spawning and gamete follicle rupture in the cnidarian *Renilla Koellikeri*: effects of putative neurohormones. *Gen Comp Endocrinol*. (2004) 137:9–18. doi: 10.1016/j.ygcen.2004.02.009
42. Plickert G, Krohner M, Munck A. Cell proliferation and early differentiation during embryonic development and metamorphosis of *Hydractinia echinata*. *Development*. (1988) 103:795–803.
43. Katsukura Y, Ando H, David CN, Grimmelikhuijzen CJP, Sugiyama T. Control of planula migration by LWamide and RFamide neuropeptides in *Hydractinia echinata*. *J Exp Biol*. (2004) 207:1803–10. doi: 10.1242/jeb.00974
44. Leitz T. Metamorphosin A and related compounds: a novel family of neuropeptides with morphogenetic activity. *Ann N Y Acad Sci*. (1998) 839:105–10. doi: 10.1111/j.1749-6632.1998.tb10740.x
45. Leitz T. Induction of metamorphosis of the marine hydrozoan *Hydractinia echinata* Fleming, 1828. *Biofouling*. (1998) 12:173–87. doi: 10.1080/08927019809378353
46. Katsukura Y, David CN, Grimmelikhuijzen CJP, Sugiyama T. Inhibition of metamorphosis by RFamide neuropeptides in planula larvae of *Hydractinia echinata*. *Dev Genes Evol*. (2003) 213:579–86. doi: 10.1007/s00427-003-0361-5
47. Seipp S, Schmich J, Will B, Schetter, Plickert G, Leitz T. Neuronal cell death during metamorphosis of *Hydractinia echinata*. (Cnidaria, Hydrozoa). *Invertebr Neurosci*. (2010) 10:77–91. doi: 10.1007/s10158-010-0109-7
48. Leitz T, Morand K, Mann M. Metamorphosin A: a novel peptide controlling development of the lower metazoan *Hydractinia echinata* (Coelenterata, Hydrozoa). *Dev Biol*. (1994) 163:440–6. doi: 10.1006/dbio.1994.1160
49. Takahashi T, Kobayakawa Y, Muneoka Y, Fujisawa Y, Mohri S, Hatta M, et al. Identification of a new member of the GLWamide peptide family: physiological activity and cellular localization in cnidarian polyps. *Comp Biochem Physiol Part B*. (2003) 135:309–24. doi: 10.1016/S1096-4959(03)00088-5
50. Leviev I, Williamson M, Grimmelikhuijzen CJP. Molecular cloning of a preprohormone from *Hydra magnipapillata* containing multiple copies of Hydra-LWamide (Leu-Trp-NH₂) neuropeptides: evidence for processing at Ser and Asn residues. *J Neurochem*. (1997) 68:1319–25. doi: 10.1046/j.1471-4159.1997.68031319.x
51. Gajewski M, Leitz T, Schloherr J, Plickert G. LWamides from cnidaria constitute a novel family of neuropeptides with morphogenetic activity. *Roux's Arch Dev Biol*. (1996) 205:232–42. doi: 10.1007/BF00365801
52. Carstensen K, Rinehart KL, McFarlane ID, Graff D, Grimmelikhuijzen CJP. Isolation of Leu-Pro-Pro-Gly-Pro-Leu-Pro-Arg-Pro-NH₂ (Antho-RPamide), and N-terminally protected, biologically active neuropeptide from sea anemones. *Peptides*. (1992) 13:851–7. doi: 10.1016/0196-9781(92)90040-A
53. Leitz T, Lay M. Metamorphosin A is a neuropeptide. *Roux's Arch Dev Biol*. (1995) 204:276–9. doi: 10.1007/BF00208495
54. Takahashi T, Ohtani M, Muneoka Y, Aimoto S, Hatta M, Shimizu H, et al. Structure-activity relation of LWamide peptides synthesized with a multipetide synthesizer. In: Kitada C, editor. *Peptide Chemistry*. Osaka: Protein Research Foundation (1997). p. 193–6.
55. Takahashi T, Hatta M. The importance of GLWamide neuropeptides in Cnidarian development and physiology. *J Amino Acids*. (2011) 2011:1–8. doi: 10.4061/2011/424501
56. Schmich J, Trepel S, Leitz T. The role of GLWamides in metamorphosis of *Hydractinia echinata*. *Dev Genes Evol*. (1998) 208:267–73. doi: 10.1007/s004270050181
57. Iwao K, Fujisawa T, Hatta M. A cnidarian neuropeptide of the GLWamide family induces metamorphosis of reef-building corals in the genus *Acropora*. *Coral Reefs*. (2002) 21:127–9. doi: 10.1007/s00338-002-0219-8
58. Erwin PM, Szmant AM. Settlement induction of *Acropora* palmate planulae by a GLW-amide neuropeptide. *Coral Reefs*. (2010) 29:929–39. doi: 10.1007/s00338-010-0634-1
59. Marcum BA, Campbell RD. Development of *hydra*. lacking nerve and intestinal cells. *J Cell Sci*. (1978) 29:17–33.
60. Campbell RD. Elimination of *Hydra*. intestinal and nerve cells by means of colchicines. *J Cell Sci*. (1976) 21:1–13.
61. Ziegglänsberger W. Substance P and pain chronicity. *Cell Tissue Res*. (2019) 375:227–41. doi: 10.1007/s00441-018-2922-y
62. Ganjiwale A, Cowsik SM. Molecular recognition of tachykinin receptor selective agonists: insights from structural studies. *Mini Rev Med Chem*. (2013) 13:2036–46. doi: 10.2174/13895575113139990079
63. Takeda N, Kyoizuka K, Deguchi R. Increase in intracellular cAMP is a prerequisite signal for initiation of physiological oocyte meiosis

- maturation in the hydrozoan *Cytaeis uchidae*. *Dev Biol.* (2006) 298:248–58. doi: 10.1016/j.ydbio.2006.06.034
64. Takeda N, Nakajima Y, Koizumi O, Fujisawa T, Takahashi T, Matsumoto M, et al. Neuropeptides trigger oocyte maturation and subsequent spawning in the hydrozoan jellyfish *Cytaeis uchidae*. *Mol Reprod Dev.* (2013) 80:223–232. doi: 10.1002/mrd.22154
 65. Yum S, Takahashi T, Koizumi O, Ariura Y, Kobayakawa Y, Mohri S, et al. A novel neuropeptide, Hym-176, induces contraction of the ectodermal muscle in *Hydra magnipapillata*. *Biochem Biophys Res Commun.* (1998) 248:584–90. doi: 10.1006/bbrc.1998.8831
 66. Noro Y, Yum S, Fujisawa-Nishimiya C, Busse, C, Shimizu H, Mineta K, et al. Regionalized nervous system in *Hydra* and the mechanism of its development. *Gene Expr Patterns.* (2019) 31:42–59. doi: 10.1016/j.gep.2019.01.003
 67. Yum S, Takahashi T, Hatta M, Fujisawa T. The structure and expression of a prehormone of a neuropeptide, Hym-176 in *Hydra magnipapillata*. *FEBS Lett.* (1998) 439:31–4. doi: 10.1016/S0014-5793(98)01314-3
 68. Takahashi T, Koizumi O, Ariura Y, Romanovitch A, Bosch TCG, Kobayakawa Y, et al. A novel neuropeptide, Hym-355, positively regulates neuron differentiation in *Hydra*. *Development.* (2000) 127:997–1005. doi: 10.1016/S1095-6433(99)90367-7
 69. Takeda N, Kon Y, Artigas GQ, Lapebie P, Barreau C, Koizumi O, et al. Identification of jellyfish neuropeptides that act directly as oocyte maturation-inducing hormones. *Development.* (2018) 145:dev156786. doi: 10.1242/dev.156786
 70. Raina AK, Jaffe H, Kempe TG, Keim P, Blacher RW, Fales HM, et al. Identification of a neuropeptide hormone that regulates sex pheromone production in female moths. *Science.* (1989) 244:796–8. doi: 10.1126/science.244.4906.796
 71. Muneoka Y, Takahashi T, Kobayashi M, Ikeda T, Minakata H, Nomoto K. Phylogenetic aspects of structure and action of molluscan neuropeptides. In: Davey KG, Peter RE, Tobe SS, editors. *Perspectives in Comparative Endocrinology*. Toronto, ON: National Research Council of Canada (1994). p. 109–18.
 72. Morris HR, Panico M, Karplus A, Lloyd PE, Piniker B. Identification by FAB-MS of the structure of a new cardioactive peptide from *Aplysia*. *Nature.* (1982) 300:643–5. doi: 10.1038/300643a0
 73. Lloyd PE, Kupfermann I, Weiss KR. Sequence of small cardioactive peptide A: a second member of a class of neuropeptides in *Aplysia*. *Peptides.* (1987) 8:179–83. doi: 10.1016/0196-9781(87)90184-7
 74. Ikeda T, Kubota I, Miki W, Nose T, Takao T, Shimonishi Y, et al. Structures and actions of 20 novel neuropeptides isolated from the ventral nerve cords of an echiuroid worm, *Urechis unicinctus*. In: Yanaihara N, editor. *Peptide Chemistry 1992*. Leiden: Protein Research Foundation (1993). p. 583–5. doi: 10.1007/978-94-011-1474-5_168
 75. Minakata H, Ikeda T, Fujita T, Kiss T, Hiripi L, Muneoka Y, et al. Neuropeptides isolated from *Helix pomatia*. Part 2. FMRFamide-related peptides, S-Iamide peptides, FR peptides and others. In: Yanaihara N, editor. *Peptide Chemistry 1992*. Leiden: Protein Research Foundation (1993) 579–82. doi: 10.1007/978-94-011-1474-5_167
 76. Grimmelikhuijzen CJP, Dierickx K, Boer GJ. Oxytocin/vasopressin-like immunoreactivity is present in the nervous system of hydra. *Neuroscience.* (1982) 7:3191–9. doi: 10.1016/0306-4522(82)90241-X
 77. Koizumi O, Bode HR. Plasticity in the nervous system of adult hydra. III. Conversion of neurons to expression of a vasopressin-like immunoreactivity depends on axial location. *J Neurosci.* (1991) 11:2011–20. doi: 10.1523/JNEUROSCI.11-07-02011.1991
 78. Morishita F, Nitagai Y, Furukawa Y, Matsushima O, Takahashi T, Hatta M, et al. Identification of a vasopressin-like immunoreactive substance in hydra. *Peptides.* (2003) 24:17–26. doi: 10.1016/S0196-9781(02)0272-3
 79. Koizumi O, Hamada S, Minobe S, Hamaguchi-Hamada K, Kurumata-Shigeto M, Nakamura M. The nerve ring in cnidarians: its presence and structure in hydrozoan medusa. *Zoology.* (2015) 118:79–88. doi: 10.1016/j.zool.2014.10.001
 80. Takahashi T, Koizumi O, Hayakawa E, Minobe S, Suetsugu R, Kobayakawa Y, et al. Further characterization of the PW peptide family that inhibits neuron differentiation in *Hydra*. *Dev Genes Evol.* (2009) 219:119–29. doi: 10.1007/s00427-009-0272-1
 81. Takahashi T. Neuropeptides and epithelipeptides: structural and functional diversity in an ancestral metazoan *Hydra*. *Protein Pept Lett.* (2013) 20:671–80. doi: 10.2174/0929866511320060006
 82. Hayakawa E, Takahashi T, Nishimiya-Fujisawa C, Fujisawa T. A novel neuropeptide (FRamide) family identified by a peptidomic approach in *Hydra magnipapillata*. *FEBS J.* (2007) 274:5438–48. doi: 10.1111/j.1742-4658.2007.06071.x
 83. Eipper B, Mains RE. Structure and biosynthesis of proadrenocorticotropin/endorphin and released peptides. *Endocr Rev.* (1980) 1:1–27. doi: 10.1210/edrv-1-1-1
 84. Rosa PA, Policastro P, Herbert E. A cellular basis for the differences in regulation of synthesis and secretion of ACTH/endorphin peptides in anterior and intermediate lobes of the pituitary. *J Exp Biol.* (1980) 89:215–37.
 85. Klumperman J, Spijker S, van Minnen J, Sharp-Baker H, Smit AB, Geraerts WPM. Cell type-specific sorting of neuropeptides: a mechanism to modulate peptide composition of large-dense-core vesicles. *J Neurosci.* (1996) 16:7930–40. doi: 10.1523/JNEUROSCI.16-24-07930.1996
 86. Koch TL, Grimmelikhuijzen CJP. Global neuropeptide annotations from the genomes and transcriptomes of Cubozoa, Scyphozoa, Staurozoa (Cnidaria: Medusozoa), and Octocorallia (Chidaria: Anthozoa). *Front. Endocrinol.* (2019) 10:831. doi: 10.3389/fendo.2019.00831
 87. Shiraishi A, Okuda T, Miyasaka N, Osugi T, Okuno Y, Inoue J, et al. Repertoires of G protein-coupled receptors for *Ciona*-specific neuropeptides. *Proc. Natl. Acad. Sci. USA.* (2019) 116:7847–56. doi: 10.1073/pnas.1816640116
 88. Siebert S, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, et al. Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science.* (2019) 365:eaav9314. doi: 10.1126/science.aav9314

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Takahashi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Function and Distribution of the Wamide Neuropeptide Superfamily in Metazoans

Elizabeth A. Williams*

Living Systems Institute, University of Exeter, Exeter, United Kingdom

The Wamide neuropeptide superfamily is of interest due to its distinctive functions in regulating life cycle transitions, metamorphic hormone signaling, and several aspects of digestive system function, from gut muscle contraction to satiety and fat storage. Due to variation among researchers in naming conventions, a global view of Wamide signaling in animals in terms of conservation or diversification of function is currently lacking. Here, I summarize the phylogenetic distribution of Wamide neuropeptides based on current data and describe recent findings in the areas of Wamide receptors and biological functions. Common trends that emerge across Cnidarians and protostomes are the presence of multiple Wamide receptors within a single organism, and the fact that Wamide signaling likely functions across an extensive variety of biological systems, including visual, circadian, and reproductive systems. Important areas of focus for future research are the further identification of Wamide-receptor pairs, confirmation of the phylogenetic distribution of Wamides through largescale sequencing and mass spectrometry, and assignment of different functions to specific subsets of Wamide-expressing neurons. More extensive study of Wamide signaling throughout larval development in a greater number of phyla is also important in order to understand the role of Wamides in hormonal regulation. Defining the evolution and function of neuropeptide signaling in animal nervous systems will benefit from an increased understanding of Wamide function and signaling mechanisms in a wider variety of organisms, beyond the traditional model systems.

OPEN ACCESS

Edited by:

Lee E. Eiden,
National Institutes of Health (NIH),
United States

Reviewed by:

Meet Zandawala,
Brown University, United States
Taka-aki Koshimizu,
Jichi Medical University, Japan

*Correspondence:

Elizabeth A. Williams
e.williams2@exeter.ac.uk

Specialty section:

This article was submitted to
Neuroendocrine Science,
a section of the journal
Frontiers in Endocrinology

Received: 16 February 2020

Accepted: 01 May 2020

Published: 28 May 2020

Citation:

Williams EA (2020) Function and
Distribution of the Wamide
Neuropeptide Superfamily in
Metazoans. *Front. Endocrinol.* 11:344.
doi: 10.3389/fendo.2020.00344

Keywords: neuropeptide, wamide, myoinhibitory peptide, GLWamide, allatostatin B

INTRODUCTION

Neuropeptides are short peptidergic molecules released by animal neurons that act as modulators or hormones to regulate biological processes. These signaling molecules are notable for being present in the nervous system of early metazoans, and for their important functions in regulating animal behavior and physiology. Historically, neuropeptides have been named according to their function, or where function is unknown, according to repetitive conserved sequence motifs found in the precursor peptide. These naming strategies have the unfortunate consequence of often obscuring neuropeptide relationships across species or phyla. The Wamide neuropeptide superfamily is a striking example of this. Wamides are repetitive proneuropeptides that contain multiple cleavage sites flanking short, amidated active peptides with a conserved C-terminal tryptophan (W). This neuropeptide superfamily is of ancient origin, already present in the last common ancestor of cnidarians and protostomes (1). Depending on the species in which they were studied, Wamides have been referred to as myoinhibitory peptide (MIP), allatostatin B

(ASTB), prothoracicostatic peptide (PTSP), WWamide, GLWamide, or metamorphosin A (MMA). Even the name “Wamide” is not ideal for defining this neuropeptide family, since neuropeptides from other families may also contain a C-terminal amidated tryptophan residue. For example, adipokinetic hormone (AKH) in some insect species, molluscan APGWamides and echinoderm luqins, and some insect short neuropeptide F's (sNPF) also have C-terminal Wamide motifs, but phylogenetic analyses indicate that these neuropeptides belong to the AKH/corazonin/ACP/GnRH superfamily, the RYA/luqin, and the sNPF/prolactin superfamilies, respectively (2–5), therefore they are not considered here. In this mini review, I aim to unite recent knowledge of Wamide expression and function in diverse phyla to improve understanding of their evolutionary history and identify remaining knowledge gaps where future research could enlighten the evolution of Wamide function and mechanism of action.

A BRIEF HISTORY OF WAMIDE DISCOVERY

The first known Wamide discovered was locust myoinhibitory peptide (LOM-MIP). LOM-MIP was identified as a suppressor of visceral muscle contraction in hindgut and oviduct in 1991 (6). Subsequently, WWamide neuropeptide was identified in the mollusc *Achatina fulica* (7) and the first cnidarian GLWamide, known as metamorphosin A (MMA), was identified in *Anthopleura elegantissima* (8). Initially, both WWamide and MMA were considered novel peptides unrelated to insect MIPs. The term allatostatin B was used to describe MIPs which inhibited juvenile hormone III synthesis in the cricket *Gryllus bimaculatus* (9). In the silkworm *Bombyx mori*, MIPs were anointed “prothoracicostatic hormone” (PTSP), due to their inhibition of ecdysone synthesis in this species (10), although cricket ASTBs were also found to inhibit ovarian ecdysteroid synthesis prior to this (11). The first crustacean Wamide was identified in crab in 2005 and was noted to be related to insect ASTBs (12).

Fifteen years after the discovery of LOM-MIP, the term “Wamides” was first used to describe this neuropeptide superfamily in the construction of a metazoan neuropeptide database by Liu et al. (13), who grouped Arthropod MIP/ASTB/PTSP, cnidarian GLWamides, molluscan WWamides and nematode MIPs within the Wamide family. Annelid Wamides were identified in 2011 by Veenstra (14), who noted their link to insect and mollusc Wamides. Genome analysis of the gastropod *Lottia gigantea* reinforced the link between molluscan WWamides and insect ASTB (15). Largescale similarity-based clustering of metazoan neuropeptides revealed that Wamides form part of the ancient central cluster of repetitive proneuropeptides that give rise to short, amidated peptides, with representative sequences from annelids, molluscs, platyhelminths, nematodes, arthropods, and cnidarians (1).

PHYLOGENETIC DISTRIBUTION OF WAMIDES

Following initial discoveries of Wamides through peptide purification and sequencing, recent largescale genomic and transcriptomic analyses have enabled a more complete view of Wamide distribution throughout the animal kingdom (Figure 1). Wamide neuropeptides also occur in brachiopods (18), tardigrades and priapulids (19–21). Transcriptome analysis of xenacoelomorph neuropeptides did not find myoinhibitory peptide orthologs, although a putative MIP receptor was found in an acoele (22). A transcript fragment of a MIP can be found however in a *Xenoturbella bockii* transcriptome dataset used to identify homeobox genes (21), thus the presence of MIP signaling in xenacoelomorphs remains uncertain for now. Extensive transcriptome analyses recently confirmed the presence of Wamides in all molluscan groups, but also supported the absence of Wamides (ASTB) in phoronids, platyhelminths and rotifers, as well as ectoprocts and entoprocts (23).

Despite extensive sequencing data, Wamides have not been identified in deuterostomes, although phylogenetic analysis of GPCRs identified orphan human GPCR139 and GPCR142 as potential MIP receptor orthologs (24). Ligands for these receptors have not been confirmed *in vivo* but are suggested to be small peptides or amino acids, such as L-Trp (W) and L-Phe (F) (25, 26). Therefore, while a Wamide signaling system arose early in metazoan evolution, in a cnidarian/bilaterian ancestor, it appears to have been lost multiple times during evolution, including in ambulacrarians and tunicates. This trend also occurs within phyla, for example, although widespread among insects, Wamides have not been found in the honey bee, wasp, or leaf-cutting ants (27, 28). A precise picture of Wamide gain and loss throughout metazoan evolution requires further sequence data, particularly from lesser-studied spiralian and ecdysozoan phyla.

Wamide distribution among early branching metazoan phyla is of ongoing interest. Recent bioinformatic analyses indicate that among cnidarians, GLWamides are absent in class Scyphozoa, Staurozoa, and Octocorallia (29). The same analysis uncovers new G{A/V/T}Wamides in Cubozoa, Scyphozoa, and Staurozoa. How these new cnidarian Wamides relate to GLWamides and the rest of the Wamide neuropeptide superfamily awaits further investigation, however it is interesting to note that an {A/V/T}Wamide C-terminal motif also occurs in arthropod, mollusc and annelid Wamides (Figure 2). In the placozoan *Trichoplax adhaerens*, an RWamide precursor was found through bioinformatic prediction (30). Placozoan RWamide has strong resemblance to cnidarian GLWamides (Figure 2), however since some luqin/Ryamide and sNPF/prolactin family neuropeptides also show a conserved RWamide C-terminal motif (4, 5), it is currently unclear to which superfamily the placozoan RWamide belongs. Similarly, although putative neuropeptide precursors were found in ctenophores these did not have homology to other metazoan neuropeptides. However, since the cross-phylum conservation of neuropeptides can be limited to just a few residues (1, 31), the existence of Ctenophore Wamides remains a

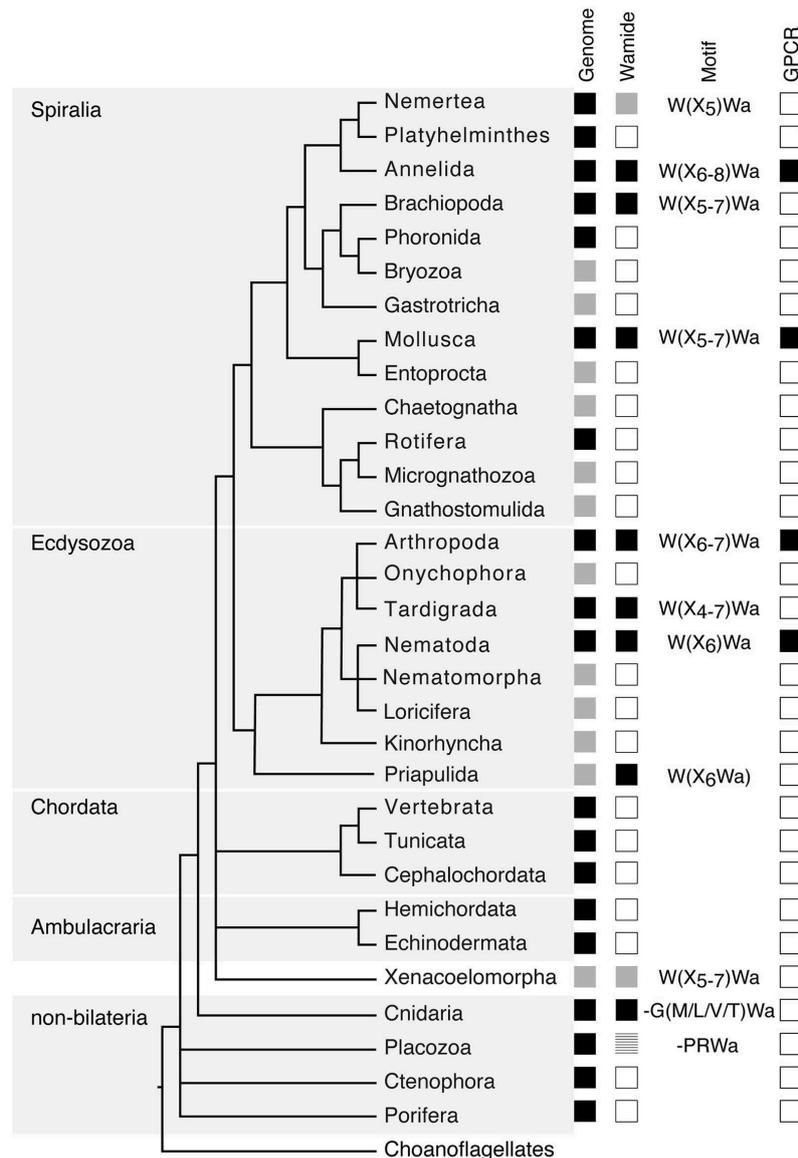
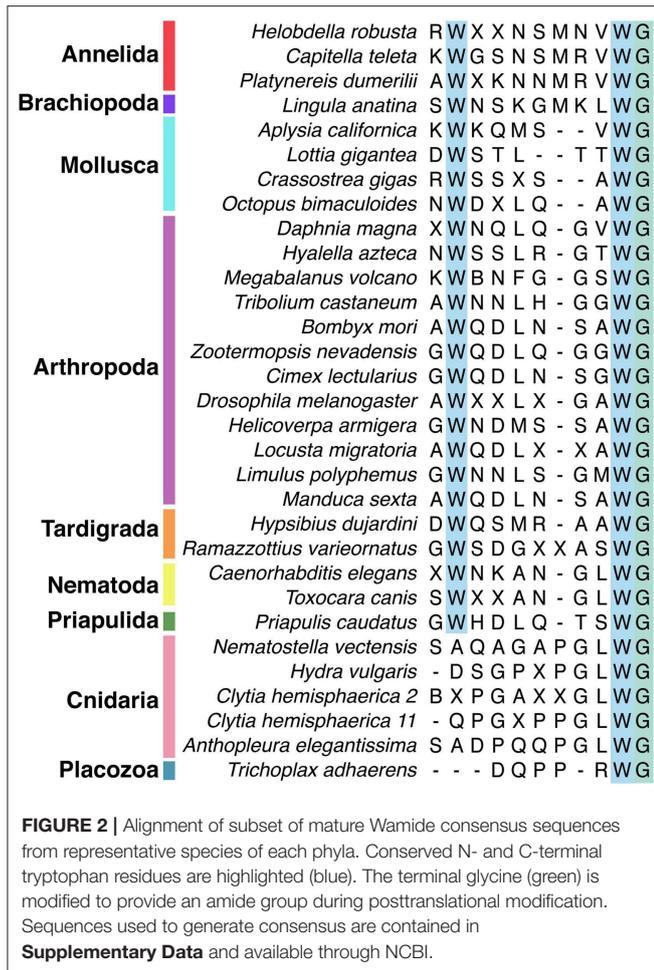


FIGURE 1 | Phylogenetic distribution of Wamide neuropeptides. In “Genome” column, black boxes indicate the presence of both genome and transcriptome data, gray boxes indicate the presence of transcriptome data only, without a published genome. In “Wamide” column, black boxes indicate confirmed presence of Wamide neuropeptide, gray boxes indicate putative Wamide sequence (partial sequence, sequence evidence only from transcriptome shotgun assembly), striped gray box indicates uncertainty of Wamide superfamily orthology, white boxes indicate lack of Wamide according to currently available sequence data. “Motif” column indicates conserved structure of Wamides within each phyla. In “GPCR” column, black boxes indicate presence of MIP/sex peptide receptor GPCR ortholog confirmed by receptor deorphanization assay, white boxes indicate currently no orthologous biochemically confirmed GPCR. Phylogeny and presence of genome based on Bezares et al. (16), Figure 6, with authors’ permission. Tree structure based on a phylogenomic study with Bayesian inference under the CAT + GTR + γ 4 model to suppress long-branch attraction artifacts (17).

possibility. No neuropeptide-like precursors have been identified in poriferans to date, despite neuropeptide-processing enzymes being present in the *Amphimedon queenslandica* genome (32). Further peptide and receptor characterization through mass spectrometry, largescale receptor deorphanization and functional studies are needed to clarify the occurrence and evolution of Wamide signaling in poriferans, ctenophores, placozoans, and cnidarians.

WAMIDE RECEPTORS

The most well-described Wamide receptor is a rhodopsin family G-protein-coupled receptor (GPCR). This was initially called the sex peptide receptor (SPR) after its first characterized ligand, the *Drosophila* sex peptide (33). Whereas sex peptides only occur in the family Drosophilidae and specifically activate the *Drosophila* receptor, *Drosophila* MIPs can activate orthologous receptors



from other insect groups and from *Aplysia californica*, a mollusc (34). Following this discovery, it was deduced that the sex peptide receptor's ancestral ligands were in fact MIPs (34, 35), causing a reassignment of name to myoinhibitory peptide receptor. Sex peptides differ significantly from MIPs in their primary amino acid sequence; they are larger (36aa cf. 9-12aa), lack C-terminal amidation, and contain a disulfide bridge. Like MIPs, however, sex peptides contain a pair of tryptophan (W) residues which are predicted to stabilize a beta-turn secondary structure adopted by both peptides (34). The conserved tryptophans are required for receptor binding in both sex peptide and MIPs, indicating that they may interact with a common binding site on the receptor (34).

Orthologs of the MIP receptor have been biochemically characterized in insects [fruitfly, mosquito (34), tick (36), kissing bug (37), silkworm (35)], a mollusc (34), nematode (38), and annelids (39). Receptor deorphanization assays with modified peptides show that the two conserved tryptophan residues are important for receptor activation. The C-terminal tryptophan residue is especially critical for receptor activation; when this residue is replaced with an alanine, all receptor activity is lost in both annelids and insects (34, 37, 39, 40).

Unlike protostomes, a cnidarian Wamide receptor has not yet been identified. Receptor identification based on phylogenetic analyses alone may prove difficult, as most cnidarian GPCRs are more closely related to each other than to specific nephrozoan receptors (22). A largescale combinatorial receptor de-orphanization approach, such as that used to identify several novel peptide-receptor pairs in *Platynereis dumerilii* (41), would be useful in this endeavor. Without a known receptor, cnidarian Wamides can not be definitively confirmed as orthologs of protostome Wamides. However, sequence similarity clustering (1), the conservation of the C-terminal GLWamide motif in nematode and some insect MIPs (**Figure 2** and **Supplementary Data**), the greater importance of the C-terminal tryptophan for receptor binding (see above), and the occasional occurrence of an N-terminal tryptophan in cnidarian GLWamides (e.g., *Hydra vulgaris*, **Supplementary Data**) are in support of orthology.

The presence of multiple Wamide receptors in the same organism is emerging as a common theme. *Caenorhabditis elegans* has three receptors related to arthropod MIP receptors (38). In *Drosophila melanogaster*, loss of the MIP/SPR GPCR does not affect the function of MIP in appetite control or female mating behavior, indicating that MIP may act through one or more additional receptors (42, 43). In the hydrozoan *Hydra magnipapillata*, the GLWamide Hym-248 activates both endodermal muscle contraction and sphincter muscle contraction, whereas other GLWamides activate only sphincter muscle contraction (44). This additional function of Hym-248 suggests that a receptor specific only to this peptide is expressed in endodermal muscle, while a more generalist GLWamide receptor is expressed in sphincter muscle (45). These findings indicate that a common evolutionary strategy for the diversification of Wamide peptide function is through the addition of receptors with varying binding specificities.

A new family of Wamide receptors has recently been identified in the marine worm *Platynereis dumerilii* (46). These receptors are peptide-gated ion channels from the degenerin/epithelial sodium channel family. The MIP-gated ion channel (MGIC) receptor is paralogous to the FMRamide-gated sodium channels found in snails and cnidarians (47, 48). The mechanism of MIPs binding to the MGIC receptor is similar to that of binding to the MIP GPCR in that the conserved tryptophans are essential for receptor activation. As with the MIP GPCR, replacement of the N-terminal tryptophan with an alanine reduced MGIC receptor activity, while replacement of the C-terminal tryptophan abolished MGIC receptor activity (46). Although all eleven mature MIP peptides from the same precursor activate both the MIP GPCR and the MGIC receptor, each mature peptide preferentially activates one of the receptor types, suggesting a mechanism for diversification of peptide function. Comparison of *in vivo* concentrations of the different mature MIPs to receptor activation concentrations determined *in vitro* would indicate whether different mature MIPs really activate both MGIC and GPCR receptor types *in vivo*. Phylogenetic analysis of peptide-gated ion channels identified nematode, amphioxus, and cnidarian sequences related to MGIC/FaNaC peptide-gated ion channels (46), however these putative peptide-gated ion

channels are yet to be assigned a peptidergic ligand through physiological assays.

KNOWN FUNCTIONS OF WAMIDES

Throughout metazoans, Wamides have a wide variety of functions and often carry out multiple functions within the same organism. One known function of Wamides shared between cnidarians and protostomes is the regulation of life cycle transitions. Cnidarian GLWamides induce larval settlement and metamorphosis in the hydroid *Hydractinia echinata*, as well as the larvae of several coral species, and the hydrozoan *Clytia hemisphaerica* (8, 49–52). Knockdown of the GLWamide precursor gene in *Nematostella vectensis* showed that GLWamide is not necessary for metamorphosis, at least in this species, but plays a modulatory role in determining metamorphic timing (53). Similar to cnidarians, treatment of larvae of *Platynereis dumerilii* with synthetic MIP peptide induces settlement (39). Both cnidarian and *P. dumerilii* Wamide-expressing cells are sensory-neurosecretory, suggesting that this neuropeptide plays a role in activating a settlement and metamorphosis program in response to specific environmental cues, however the nature of the environmental cues that trigger Wamide peptide release, and the downstream signaling pathways activated or repressed by Wamides require further characterization. Microarray and RNA-Seq studies of GLWamide-treated branching coral *Acropora millepora* identified significant changes in transcription after peptide exposure (54, 55). These studies may be useful for identifying candidate genes and pathways for functional investigations of Wamide signaling networks acting in coral metamorphosis, now that tools for genome editing are available for coral (56). Curiously, the *Hydra* GLWamide, Hym-248, also promotes settlement and metamorphosis in two species of sponge (57). This suggests that although Wamides haven't been identified in sponges, there may be some overlap between the signal transduction pathways involved in sponge and cnidarian metamorphosis, particularly at the level of neuropeptide receptors.

Similar to regulating marine invertebrate larval settlement and metamorphosis, in some insect species, Wamides regulate levels of metamorphic hormones. In the silkworm, PTSPs suppress ecdysteroidogenesis in the prothoracic glands (10, 35, 58), while in crickets, allatostatin B inhibits juvenile hormone production (59, 60). MIPs also regulate aspects of the behavioral sequence underlying ecdysis in *D. melanogaster* and *Manduca sexta*, as part of a peptidergic signaling cascade initiated by ecdysis triggering hormone (61–64). However, functions of Wamides during insect larval development and how Wamides regulate hormone production or release still require further investigation. Do Wamides also regulate hormones similar to juvenile hormone or ecdysone in the induction of marine invertebrate metamorphosis? Juvenile hormone and its precursor, methyl farnesoate, can regulate larval metamorphosis in polychaetes and barnacles (65, 66) and ecdysone signaling may play a role in crustacean and mollusc metamorphosis (67), however Wamide function in mollusc and crustacean larval stages has not yet been reported, nor has the effect of Wamides on specific marine invertebrate hormones.

Perhaps the most widely conserved function of Wamides is the regulation of muscle contraction. In arthropods, MIPs inhibit muscle contraction in the hindgut, oviduct and heart/pyloric system (6, 68–72). Contrary to their name, MIPs promote muscle contractions in annelid gut muscles (73). Cnidarian GLWamides also promote muscle contractions in the longitudinal (ectodermal) muscles and circumferential (endoderm) muscles (44). A specialization of the muscle contraction function of some *Hydra* GLWamides is in inducing the contraction of sphincter muscles, which causes detachment of buds from a parental polyp, thereby linking Wamides to the regulation of asexual reproduction in *Hydra* (74). Mollusc WWamides inhibit the phasic contractions of the anterior byssus retractor muscle in mussel, but potentiate contractions of the penis and radula retractor muscles in land snail, as well as inducing contractions of the radula protractor muscle in a gastropod (7). Studies of the effects of Wamides on muscle contraction have revealed that Wamide effects are dependent on the current physiological state of the system (71), the concentration of peptide applied (7, 75), the type of receptor activated (45), and whether the peptide acts pre- or post-synaptically (7).

Additional to the regulation of muscle contraction in the digestive system, Wamides have been implicated in a variety of feeding-related functions, primarily in insects. In the kissing bug and ticks, MIP expression in the salivary glands suggests a role in salivation (36, 37, 76, 77). Activating MIP neurons in *D. melanogaster* adults decreased food intake and body weight and reduced the sensitivity of starved flies toward food (42). Also in *D. melanogaster*, MIPs modulate attraction to polyamine food odors in mated females (78). This sex-specific modulation is through an autocrine signaling mechanism, with MIP and the MIP/sex peptide receptor both expressed in taste and olfactory neurons. Recent findings from *C. elegans* show a role for MIP in aversive gustatory short-term learning, and long-term learning of salt avoidance behaviors (38). MIP's function in sensing specific food cues in adult fly and nematode shows interesting parallels with the role of MIP in regulating marine invertebrate metamorphosis. In both cases, MIP is expressed in neurons with chemosensory morphology and detection of specific cues and release of MIPs activates a switch in states. Cues for marine invertebrate metamorphosis are also often associated with preferred food sources of their future adult stages (79).

Beyond the regulation of life cycle transitions, muscle contraction and feeding/digestive, MIPs may also play a role in diverse biological systems, including reproductive, visual, and circadian systems. In female *D. melanogaster*, MIP-expressing abdominal interneurons enhance mating receptivity in mated females (43). Additionally, MIP-expressing interneurons in the central brain form part of a mechanosensory circuit that informs female mating decisions (80). Also in *D. melanogaster*, some optic lobe neurons express MIP, although it is not yet clear if these regulate signaling in the visual or circadian system, or both (81). MIP expression in *D. melanogaster* cycles with circadian rhythm, in line with the role of MIP in maintaining a sleep-like state in adults (82). Spatial expression of MIP in cockroach brains also suggests a role in the circadian system (83, 84). MIP expression in insects has some similarities to GLWamide expression in jellyfish.

For example, GLWamide expression is seen in the gonadal ectoderm cells of *Clytia hemisphaerica* and *Cytaeis uchidae* (85, 86), as well as the photoreceptive organs of the jellyfish *Cladonema radiatum*, *Aurelia aurita*, and *Tripedalia cystophora* (87). In both insects and cnidarians, further functional studies are needed to reveal the function of Wamides in reproduction, vision or circadian rhythm. These previous studies of spatial expression highlight the usefulness of detailed expression atlases that encompass different life cycle stages, sexes and whole-body analyses for additional species and phyla, to provide initial indications of Wamide function in distinct biological systems.

CONCLUSIONS AND FUTURE DIRECTIONS

Wamides are clearly neuropeptides of significant importance to nervous system signaling, with a role in diverse biological systems throughout an organism's life cycle. The fact that Wamide signaling is lost in some species or phyla indicates that Wamides function within more complex networks of neuropeptide signaling, sometimes playing a modulatory but non-essential role, and their function may be taken on or replaced by other neuropeptides. Several similarities are seen in both the function and spatial expression of Wamides in cnidarians and protostomes, supporting the definition of this neuropeptide superfamily, however the identification of cnidarian Wamide receptors will further enlighten the evolution of Wamide signaling in metazoans.

The most detailed studies of Wamide signaling to date have been carried out in model organisms *D. melanogaster* and *C. elegans* and these studies show that subsets of MIP-expressing neurons are likely different cell types (e.g., sensory neurons vs. interneurons) responsible for different aspects of Wamide function. It is therefore important for future studies aiming to uncover mechanisms of Wamide signaling to develop methods for manipulating specific subsets of Wamide-expressing neurons. One approach to this is through the development of libraries of reporter constructs driven by different promoters with a range of cell-specificities. Further understanding of Wamide signaling can also be achieved by analysis of genes co-expressed in Wamide- and Wamide receptor-expressing cells and more widespread analyses of Wamide receptor expression. These analyses can

indicate mechanisms of signaling, such as autoregulation in cells expressing both Wamide and receptor, or association with specific neurotransmitters, such as GABA or acetylcholine. They can also be used to generate maps of potential signaling cascades, as in insect ecdysis (61), through the comparison of expression of other neuropeptides and receptors. Receptor-ligand expression mapping based on single cell transcriptome data, as in (88), which should then be functionally tested, will facilitate these analyses.

Another aspect of Wamide signaling of importance for future studies is the identification of signaling cascades activated following Wamide release and receptor binding, in terms of which class of G protein is recruited, if indeed the Wamide signal is GPCR-mediated, and which downstream signaling pathways are activated or repressed. Again, single cell RNA-Seq analyses or precise spatial expression mapping can be used to identify which elements of the signaling pathway are present in the target cells of Wamide signaling. With both distinctive and conserved functions, the Wamide superfamily is an excellent model for studying the evolution of neuropeptide signaling and patterns of peptide-receptor coevolution in animal nervous systems.

AUTHOR CONTRIBUTIONS

EW conceived the study and wrote the paper.

FUNDING

EW is supported by a BBSRC David Phillips Fellowship (BB/T00990X/1).

ACKNOWLEDGMENTS

The author is grateful to members of the Jékely lab for informative discussions and valuable advice, and the two referees whose constructive comments improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2020.00344/full#supplementary-material>

REFERENCES

- Jekely G. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci USA*. (2013) 110:8702–7. doi: 10.1073/pnas.1221833110
- Tian S, Zandawala M, Beets I, Baytemur E, Slade SE, Scrivens JH, et al. Urbilaterian origin of paralogous GnRH and corazonin neuropeptide signalling pathways. *Sci Rep*. (2016) 6:28788. doi: 10.1038/srep28788
- Zandawala M, Tian S, Elphick MR. The evolution and nomenclature of GnRH-type and corazonin-type neuropeptide signaling systems. *Gen Comp Endocrinol*. (2018) 264:64–77. doi: 10.1016/j.ygcen.2017.06.007
- Yañez-Guerra LA, Delroisse J, Barreiro-Iglesias A, Slade SE, Scrivens JH, Elphick MR. Discovery and functional characterisation of a luqin-type neuropeptide signalling system in a deuterostome. *Sci Rep*. (2018) 8:7220. doi: 10.1038/s41598-018-25606-2
- Yañez-Guerra LA, Zhong X, Moghul I, Butts T, Zampronio CG, Jones AM, et al. Urbilaterian origin and evolution of sNPF-type neuropeptide signaling. *BioRxiv*. (2019). doi: 10.1101/712687
- Schoofs L, Holman GM, Hayes TK, Nachman RJ, De Loof A. Isolation, identification and synthesis of locusta myoinhibiting peptide (LOM-MIP), a novel biologically active neuropeptide from locusta migratoria. *Regul Pept*. (1991) 36:111–9. doi: 10.1016/0167-0115(91)90199-Q
- Minakata H, Ikeda T, Muneoka Y, Kobayashi M, Nomoto K. WWamide-1,–2 and–3: novel neuromodulatory peptides isolated from ganglia of the African giant snail, *Achatina fulica*. *FEBS Lett*. (1993) 323:104–8. doi: 10.1016/0014-5793(93)81458-C

8. Leitz T, Morand K, Mann M. Metamorphosin A: a novel peptide controlling development of the lower metazoan hydractinia echinata (coelenterata, hydrozoa). *Dev Biol.* (1994) 163:440–6. doi: 10.1006/dbio.1994.1160
9. Lorenz MW, Kellner R, Hoffmann KH. A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket, *gryllus bimaculatus*. *J Biol Chem.* (1995) 270:21103–8. doi: 10.1074/jbc.270.36.21103
10. Hua YJ, Tanaka Y, Nakamura K, Sakakibara M, Nagata S, Kataoka H. Identification of a prothoracicostatic peptide in the larval brain of the silkworm, *bombyx mori*. *J Biol Chem.* (1999) 274:31169–73. doi: 10.1074/jbc.274.44.31169
11. Lorenz MW, Lorenz JI, Treiblmaier K, Hoffmann K-H. *In vivo* effects of allatostatins in crickets, *gryllus bimaculatus* (sensifera: gryllidae). *Arch Insect Biochem Physiol.* (1998) 38:32–43. doi: 10.1002/(SICI)1520-6327(1998)38:1<32::AID-ARCH4>3.0.CO;2-X
12. Fu Q, Li L. *De novo* sequencing of neuropeptides using reductive isotopic methylation and investigation of ESI QTOF MS/MS fragmentation pattern of neuropeptides with N-terminal dimethylation. *Anal Chem.* (2005) 77:7783–95. doi: 10.1021/ac051324e
13. Liu F, Baggerman G, Schoofs L, Wets G. The construction of a bioactive peptide database in metazoa. *J Proteome Res.* (2008) 7:4119–31. doi: 10.1021/pr800037n
14. Veenstra JA. Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *capitella teleta* and *helobdella robusta*. *Gen Comp Endocrinol.* (2011) 171:160–75. doi: 10.1016/j.ygcen.2011.01.005
15. Veenstra JA. Neurohormones and neuropeptides encoded by the genome of *lottia gigantea*, with reference to other mollusks and insects. *Gen Comp Endocrinol.* (2010) 167:86–103. doi: 10.1016/j.ygcen.2010.02.010
16. Bezares-Calderón LA, Berger J, Jékely G. Diversity of cilia-based mechanosensory systems and their functions in marine animal behaviour. *Phil Trans R Soc B Biol Sci.* (2020) 375:20190376. doi: 10.1098/rstb.2019.0376
17. Laumer CE, Fernández R, Lemer S, Combsbosch D, Kocot KM, Riesgo A, et al. Revisiting metazoan phylogeny with genomic sampling of all phyla. *Phil Trans R Soc B Biol Sci.* (2019) 286:20190831. doi: 10.1098/rspb.2019.0831
18. Luo Y-J, Takeuchi T, Koyanagi R, Yamada L, Kanda M, Khalturina M, et al. The lingula genome provides insights into brachiopod evolution and the origin of phosphate biomineralization. *Nat Commun.* (2015) 6:8301. doi: 10.1038/ncomms9301
19. Koziol U. Precursors of neuropeptides and peptide hormones in the genomes of tardigrades. *Gen Comp Endocrinol.* (2018) 267:116–27. doi: 10.1016/j.ygcen.2018.06.012
20. Borner J, Rehm P, Schill RO, Ebersberger I, Burmester T. A transcriptome approach to ecdysozoan phylogeny. *Mol Phylogenet Evol.* (2014) 80:79–87. doi: 10.1016/j.ympev.2014.08.001
21. Brauchle M, Bilican A, Eyer C, Bailly X, Martínez P, Ladurner P, et al. Xenacoelomorpha survey reveals that all 11 animal homeobox gene classes were present in the first bilaterians. *Genome Biol Evol.* (2018) 10:2205–17. doi: 10.1093/gbe/evy170
22. Thiel D, Franz-Wachtel M, Aguilera F, Hejnol A. Xenacoelomorph neuropeptidomes reveal a major expansion of neuropeptide systems during early bilaterian evolution. *Mol Biol Evol.* (2018) 35:2528–43. doi: 10.1093/molbev/msy160
23. De Oliveira AL, Calcino A, Wanninger A. Extensive conservation of the proneuropeptide and peptide prohormone complement in mollusks. *Sci Rep.* (2019) 9:4846. doi: 10.1038/s41598-019-40949-0
24. Mirabeau O, Joly J-S. Molecular evolution of peptidergic signaling systems in bilaterians. *Proc Natl Acad Sci USA.* (2013) 110:E2028–37. doi: 10.1073/pnas.1219956110
25. Süsens U, Hermans-Borgmeyer I, Urny J, Schaller HC. Characterisation and differential expression of two very closely related G-protein-coupled receptors, GPR139 and GPR142, in mouse tissue and during mouse development. *Neuropharmacology.* (2006) 50:512–20. doi: 10.1016/j.neuropharm.2005.11.003
26. Isberg V, Andersen KB, Bisig C, Dietz GPH, Bräuner-Osborne H, Gloriam DE. Computer-Aided discovery of aromatic l- α -amino acids as agonists of the orphan g protein-coupled receptor GPR139. *J Chem Inform Model.* (2014) 54:1553–7. doi: 10.1021/ci500197a
27. Hauser F, Neupert S, Williamson M, Predel R, Tanaka Y, Grimmlikhuijzen CJP. Genomics and peptidomics of neuropeptides and protein hormones present in the parasitic wasp *nasomia vitripennis*. *J Proteome Res.* (2010) 9:5296–310. doi: 10.1021/pr100570j
28. Nygaard S, Zhang G, Schiött M, Li C, Wurm Y, Hu H, et al. The genome of the leaf-cutting ant *acromyrmex echinatio* suggests key adaptations to advanced social life and fungus farming. *Genome Res.* (2011) 21:1339–48. doi: 10.1101/gr.121392.111
29. Koch TL, Grimmlikhuijzen CJP. Global neuropeptide annotations from the genomes and transcriptomes of cubozoa, scyphozoa, staurozoa (cnidaria: medusozoa), and octocorallia (cnidaria: anthozoa). *Front Endocrinol.* (2019) 10:831. doi: 10.3389/fendo.2019.00831
30. Nikitin M. Bioinformatic prediction of *Trichoplax adhaerens* regulatory peptides. *Gen Comp Endocrinol.* (2015) 212:145–55. doi: 10.1016/j.ygcen.2014.03.049
31. Jékely G, Paps J, Nielsen C. The phylogenetic position of ctenophores and the origin(s) of nervous systems. *Evodevo.* (2015) 6:1. doi: 10.1186/2041-9139-6-1
32. Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier MEA, Mitros T, et al. The amphimedon queenslandica genome and the evolution of animal complexity. *Nature.* (2010) 466:720–6. doi: 10.1038/nature09201
33. Yapici N, Kim Y-J, Ribeiro C, Dickson BJ. A receptor that mediates the post-mating switch in *drosophila* reproductive behaviour. *Nature.* (2008) 451:33–7. doi: 10.1038/nature06483
34. Kim Y-J, Bartalska K, Audsley N, Yamanaka N, Yapici N, Lee J-Y, et al. MIPs are ancestral ligands for the sex peptide receptor. *Proc Natl Acad Sci USA.* (2010) 107:6520–5. doi: 10.1073/pnas.0914764107
35. Yamanaka N, Hua Y-J, Roller L, Spalovská-Valachová I, Mizoguchi A, Kataoka H, et al. Bombyx prothoracicostatic peptides activate the sex peptide receptor to regulate ecdysteroid biosynthesis. *Proc Natl Acad Sci USA.* (2010) 107:2060–5. doi: 10.1073/pnas.0907471107
36. Simo L, Koči J, Park Y. Receptors for the neuropeptides, myoinhibitory peptide and SIFamide, in control of the salivary glands of the blacklegged tick *Ixodes scapularis*. *Insect Biochem Mol Biol.* (2013) 43:376–87. doi: 10.1016/j.ibmb.2013.01.002
37. Paluzzi J-PV, Haddad AS, Sedra L, Orchard I, Lange AB. Functional characterization and expression analysis of the myoinhibiting peptide receptor in the chagas disease vector, *rhodnius prolixus*. *Mol Cell Endocrinol.* (2015) 399:143–53. doi: 10.1016/j.mce.2014.09.004
38. Peymen K, Wattereyne J, Borghgraef C, Van Sinay E, Beets I, Schoofs L. Myoinhibitory peptide signaling modulates aversive gustatory learning in *Caenorhabditis elegans*. *PLoS Genet.* (2019) 15:e1007945. doi: 10.1371/journal.pgen.1007945
39. Conzelmann M, Williams EA, Tunaru S, Randel N, Shahidi R, Asadulina A, et al. Conserved MIP receptor-ligand pair regulates platynereis larval settlement. *Proc Natl Acad Sci USA.* (2013) 110:8224–9. doi: 10.1073/pnas.1220285110
40. Vandersmissen HP, Nachman RJ, Vanden Broeck J. Sex peptides and MIPs can activate the same G protein-coupled receptor. *Gen Comparative Endocrinol.* (2013) 188:137–43. doi: 10.1016/j.ygcen.2013.02.014
41. Bauknecht P, Jékely G. Large-scale combinatorial deorphanization of platynereis neuropeptide GPCRs. *Cell Rep.* (2015) 12:684–93. doi: 10.1016/j.celrep.2015.06.052
42. Min S, Chae H-S, Jang Y-H, Choi S, Lee S, Jeong YT, et al. Identification of a peptidergic pathway critical to satiety responses in *drosophila*. *Curr Biol.* (2016) 26:814–20. doi: 10.1016/j.cub.2016.01.029
43. Jang Y-H, Chae H-S, Kim Y-J. Female-specific myoinhibitory peptide neurons regulate mating receptivity in *drosophila melanogaster*. *Nat Commun.* (2017) 8:1630. doi: 10.1038/s41467-017-01794-9
44. Takahashi T, Kobayakawa Y, Muneoka Y, Fujisawa Y, Mohri S, Hatta M, et al. Identification of a new member of the GLWamide peptide family: physiological activity and cellular localization in cnidarian polyps. *Comp Biochem Physiol B Biochem Mol Biol.* (2003) 135:309–24. doi: 10.1016/S1096-4959(03)00088-5
45. Takahashi T, Hatta M. The importance of glwamide neuropeptides in cnidarian development and physiology. *J Amino Acids.* (2011) 2011:1–8. doi: 10.4061/2011/424501

46. Schmidt A, Bauknecht P, Williams EA, Augustynowski K, Gründer S, Jékely G. Dual signaling of wamide myoinhibitory peptides through a peptide-gated channel and a GPCR in *Platynereis*. *FASEB J.* (2018) 32:5338–49. doi: 10.1096/fj.201800274R
47. Lingueglia E, Champigny G, Lazdunski M, Barbry P. Cloning of the amiloride-sensitive FMRFamide peptide-gated sodium channel. *Nature.* (1995) 378:730–3. doi: 10.1038/378730a0
48. Golubovic A, Kuhn A, Williamson M, Kalbacher H, Holstein TW, Grimmelikhuijzen CJ, et al. A peptide-gated ion channel from the freshwater polyp hydra. *J Biol Chem.* (2007) 282:35098–103. doi: 10.1074/jbc.M706849200
49. Iwao K, Fujisawa T, Hatta M. A cnidarian neuropeptide of the GLWamide family induces metamorphosis of reef-building corals in the genus *Acropora*. *Coral Reefs.* (2002) 21:127–9. doi: 10.1007/s00338-002-0219-8
50. Erwin PM, Szmant AM. Settlement induction of *Acropora palmata* planulae by a GLW-amide neuropeptide. *Coral Reefs.* (2010) 29:929–39. doi: 10.1007/s00338-010-0634-1
51. Momose T, De Cian A, Shiba K, Inaba K, Giovannangeli C, Concordet J-P. High doses of CRISPR/Cas9 ribonucleoprotein efficiently induce gene knockout with low mosaicism in the hydrozoan clytia *hemisphaerica* through microhomology-mediated deletion. *Sci Rep.* (2018) 8:11734. doi: 10.1038/s41598-018-30188-0
52. Lechable M, Jan A, Weissbourd B, Uveira J, Gissat L, Collet S, et al. An improved whole life cycle culture protocol for the hydrozoan genetic model *Clytia hemisphaerica*. (2019). doi: 10.1101/852632
53. Nakanishi N, Martindale MQ. CRISPR knockouts reveal an endogenous role for ancient neuropeptides in regulating developmental timing in a sea anemone. *Elife.* (2018) 7:e39742. doi: 10.7554/eLife.39742
54. Grasso LC, Negri AP, Fôret S, Saint R, Hayward DC, Miller DJ, et al. The biology of coral metamorphosis: molecular responses of larvae to inducers of settlement and metamorphosis. *Dev Biol.* (2011) 353:411–9. doi: 10.1016/j.ydbio.2011.02.010
55. Meyer E, Aglyamova GV, Matz MV. Profiling gene expression responses of coral larvae (*Acropora millepora*) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. *Mol Ecol.* (2011) 20:3599–616. doi: 10.1111/j.1365-294X.2011.05205.x
56. Cleves PA, Strader ME, Bay LK, Pringle JR, Matz MV. CRISPR/Cas9-mediated genome editing in a reef-building coral. *Proc Natl Acad Sci USA.* (2018) 115:5235–40. doi: 10.1073/pnas.1722151115
57. Whalan S, Webster NS, Negri AP. Crustose coralline algae and a cnidarian neuropeptide trigger larval settlement in two coral reef sponges. *PLoS ONE.* (2012) 7:e30386. doi: 10.1371/journal.pone.0030386
58. Liu X, Tanaka Y, Song Q, Xu B, Hua Y. Bombyx mori prothoracicostatic peptide inhibits ecdysteroidogenesis *in vivo*. *Arch Insect Biochem Physiol.* (2004) 56:155–61. doi: 10.1002/arch.20005
59. Bendena WG, Donly BC, Tobe SS. Allatostatins: a growing family of neuropeptides with structural and functional diversity. *Ann NY Acad Sci.* (1999) 897:311–29. doi: 10.1111/j.1749-6632.1999.tb07902.x
60. Stay B, Tobe SS, Bendena WG. Allatostatins: identification, primary structures, functions and distribution. *Adv. Insect Physiol.* (1995) 267–337. doi: 10.1016/S0065-2806(08)60066-1
61. Kim Y-J, Zitnan D, Galizia CG, Cho K-H, Adams ME. A command chemical triggers an innate behavior by sequential activation of multiple peptidergic ensembles. *Curr Biol.* (2006) 16:1395–407. doi: 10.1016/j.cub.2006.06.027
62. Kim Y-J, Kim Y, Zitnan D, Cho K, Schooley DA, Mizoguchi A, et al. Central peptidergic ensembles associated with organization of an innate behavior. *Proc Natl Acad Sci USA.* (2006) 103:14211–6. doi: 10.1073/pnas.0603459103
63. Santos JG, Vömel M, Struck R, Homberg U, Nässel DR, Wegener C. Neuroarchitecture of peptidergic systems in the larval ventral ganglion of *Drosophila melanogaster*. *PLoS ONE.* (2007) 2:e695. doi: 10.1371/journal.pone.0000695
64. Davis NT. Localization of myoinhibitory peptide immunoreactivity in *manduca sexta* and *bombyx mori*, with indications that the peptide has a role in molting and ecdysis. *J Exp Biol.* (2003) 206:1449–60. doi: 10.1242/jeb.00234
65. Laufer H, Biggers WJ. Unifying concepts learned from methyl farnesoate for invertebrate reproduction and post-embryonic development. *Am Zool.* (2001) 41:442–57. doi: 10.1093/icb/41.3.442
66. Yamamoto H, Okino T, Yoshimura E, Tachibana A, Shimizu K, Fusetani N. Methyl farnesoate induces larval metamorphosis of the barnacle, *balanus amphitrite* via protein kinase C activation. *J Exp Zool.* (1997) 278:349–55. doi: 10.1002/(SICI)1097-010X(19970815)278:6<349::AID-JEZ2>3.0.CO;2-O
67. Joyce A, Vogeler S. Molluscan bivalve settlement and metamorphosis: neuroendocrine inducers and morphogenetic responses. *Aquaculture.* (2018) 487:64–82. doi: 10.1016/j.aquaculture.2018.01.002
68. Blackburn MB, Wagner RM, Kochansky JP, Harrison DJ, Thomas-Laemont P, Raina AK. The identification of two myoinhibitory peptides, with sequence similarities to the galanins, isolated from the ventral nerve cord of *manduca sexta*. *Regul Pept.* (1995) 57:213–9. doi: 10.1016/0167-0115(95)00034-9
69. Predel R, Rapus J, Eckert M. Myoinhibitory neuropeptides in the american cockroach. *Peptides.* (2001) 22:199–208. doi: 10.1016/S0196-9781(00)00383-1
70. Aguilar R, Maestro JL, Belles X. Effects of myoinhibitory peptides on food intake in the German cockroach. *Physiol Entomol.* (2006) 31:257–61. doi: 10.1111/j.1365-3032.2006.00513.x
71. Szabo TM, Chen R, Goeritz ML, Maloney RT, Tang LS, Li L, et al. Distribution and physiological effects of B-type allatostatins (myoinhibitory peptides, MIPs) in the stomatogastric nervous system of the crab *Cancer borealis*. *J Comp Neurol.* (2011) 519:2658–76. doi: 10.1002/cne.22654
72. 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) 101:1099–107. doi: 10.1111/j.1471-4159.2007.04482.x
73. Williams EA, Conzelmann M, Jékely G. Myoinhibitory peptide regulates feeding in the marine annelid *Platynereis*. *Front Zool.* (2015) 12:1. doi: 10.1186/s12983-014-0093-6
74. Takahashi T, Muneoka Y, Lohmann J, Lopez de Haro MS, Solleder G, Bosch TC, et al. Systematic isolation of peptide signal molecules regulating development in hydra: LWamide and PW families. *Proc Natl Acad Sci USA.* (1997) 94:1241–6. doi: 10.1073/pnas.94.4.1241
75. Katsukura Y, Ando H, David CN, Grimmelikhuijzen CJ, Sugiyama T. Control of planula migration by LWamide and RFamide neuropeptides in *Hydractinia echinata*. *J Exp Biol.* (2004) 207:1803–10. doi: 10.1242/jeb.00974
76. Simo L, Zitnan D, Park Y. Two novel neuropeptides in innervation of the salivary glands of the black-legged tick, *Ixodes scapularis*: myoinhibitory peptide and SIFamide. *J Comp Neurol.* (2009) 517:551–63. doi: 10.1002/cne.22182
77. Lange AB, Alim U, Vandersmissen HP, Mizoguchi A, Vanden Broeck J, Orchard I. The distribution and physiological effects of the myoinhibiting peptides in the kissing bug, *Rhodnius prolixus*. *Front Neurosci.* (2012) 6:98. doi: 10.3389/fnins.2012.00098
78. Hussain A, Üçpunar HK, Zhang M, Loschek LF, Grunwald Kadow IC. Neuropeptides modulate female chemosensory processing upon mating in *Drosophila*. *PLoS Biol.* (2016) 14:e1002455. doi: 10.1371/journal.pbio.1002455
79. Hadfield M, Paul V. Natural chemical cues for settlement and metamorphosis of marine-invertebrate larvae. *Marine Sci.* (2001) 431–61. doi: 10.1201/9781420036602.ch13
80. Shao L, Chung P, Wong A, Siwanowicz I, Kent CF, Long X, et al. A neural circuit encoding the experience of copulation in female *Drosophila*. *Neuron.* (2019) 102:1025–36. doi: 10.1016/j.neuron.2019.04.009
81. Kolodziejczyk A, Nässel DR. A novel wide-field neuron with branches in the lamina of the *Drosophila* visual system expresses myoinhibitory peptide and may be associated with the clock. *Cell Tissue Res.* (2011) 343:357–69. doi: 10.1007/s00441-010-1100-7
82. Oh Y, Yoon S-E, Zhang Q, Chae H-S, Daubnerová I, Shafer OT, et al. A homeostatic sleep-stabilizing pathway in *Drosophila* composed of the sex peptide receptor and its ligand, the myoinhibitory peptide. *PLoS Biol.* (2014) 12:e1001974. doi: 10.1371/journal.pbio.1001974
83. Schulze J, Neupert S, Schmidt L, Predel R, Lamkemeyer T, Homberg U, et al. Myoinhibitory peptides in the brain of the cockroach *Leucophaea maderae* and colocalization with pigment-dispersing factor in circadian pacemaker cells. *J Comp Neurol.* (2012) 520:1078–97. doi: 10.1002/cne.22785
84. Stengl M, Arendt A. Peptidergic circadian clock circuits in the *madeira* cockroach. *Curr Opin Neurobiol.* (2016) 41:44–52. doi: 10.1016/j.conb.2016.07.010
85. Takeda N, Nakajima Y, Koizumi O, Fujisawa T, Takahashi T, Matsumoto M, et al. Neuropeptides trigger oocyte maturation and subsequent spawning

- in the hydrozoan jellyfish *Cytaeis uchidae*. *Mol Reprod Dev.* (2013) 80:223–32. doi: 10.1002/mrd.22154
86. Takeda N, Kon Y, Quiroga Artigas G, Lapébie P, Barreau C, Koizumi O, et al. Identification of jellyfish neuropeptides that act directly as oocyte maturation-inducing hormones. *Development.* (2018) 145:140160. doi: 10.1101/140160
87. Plickert G, Schneider B. Neuropeptides and photic behavior in cnidaria. *Hydrobiologia.* (2004) 530–531:49–57. doi: 10.1007/s10750-004-2689-x
88. Williams EA, Verasztó C, Jasek S, Conzelmann M, Shahidi R, Bauknecht P, et al. Synaptic and peptidergic connectome of a neurosecretory center in the annelid brain. *Elife.* (2017) 6:e26349. doi: 10.7554/eLife.26349

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read for greatest visibility and readership



FAST PUBLICATION

Around 90 days from submission to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative, and constructive peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers acknowledged by name on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data and methods to enhance research reproducibility



DIGITAL PUBLISHING

Articles designed for optimal readership across devices



FOLLOW US

[@frontiersin](https://twitter.com/frontiersin)



IMPACT METRICS

Advanced article metrics track visibility across digital media



EXTENSIVE PROMOTION

Marketing and promotion of impactful research



LOOP RESEARCH NETWORK

Our network increases your article's readership