General and Comparative Endocrinology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

General and Comparative Endocrinology

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

Short communication

Gonadotropin-releasing hormone by any other name would smell as sweet

Pei-San Tsai

Department of Integrative Physiology and Center for Neuroscience, University of Colorado, Boulder, CO 80309-0354, United States

ARTICLE INFO

Article history:
Received 28 July 2017
Revised 28 August 2017
Accepted 13 September 2017
Available online xxxx

Keywords:
GnRH superfamily
Nomenclature
Invertebrate GnRH
AKH
Corazonin
ACP

ABSTRACT

The goal of this article is to discuss the nomenclature of members of the gonadotropin-releasing hormone (GnRH) superfamily. This superfamily currently consists of 5 families: (1) vertebrate GnRH, (2) adipokinetic hormone, (3) corazonin, (4) adipokinetic hormone/corazonin-related peptide and (5) invertebrate GnRH (or GnRH/corazonin). The naming of some of these peptides, especially members of the invertebrate GnRH family, may not have reflected their true evolutionary origin, leading to some confusion and controversy. Using a few examples from the invertebrate GnRH family, this article proposes several peptide-naming criteria and discusses naming challenges and problem cases. It is recommended that the invertebrate GnRH family be renamed based on the naming criteria of (1) mature peptide structure, (2) prepropeptide phylogeny, and (3) receptor phylogeny. Following this approach, the names of the peptides should reflect their phylogeny, and if possible, delineate a monophyletic group.

1. Introduction

The evolution of peptides is a complicated and non-linear process. Functional and structural changes in peptides over time could obscure their evolutionary origins, especially in short peptides lacking sufficient amino acids for rigorous analysis. Consequently, peptides are often analyzed in conjunction with their receptors to provide a more complete evolutionary history, since ligands and receptors frequently co-evolve (Hauser et al., 2006; Hauser and Grimmelikhuijzen, 2014; Markov et al., 2008b; Moyle et al., 1994; Park et al., 2002; van Kesteren et al., 1996).

This review is written as a tribute to Dr. Stacia Sower's exemplary scientific career and her long-standing contributions to the study of agnathan reproduction. In one of her recent publications on the analysis of hormone evolution (Plachetzki et al., 2016), she and the co-authors discussed an important problem plaguing the field of comparative endocrinology in the past few years. The problem was how to properly name a large number of novel invertebrate peptides recently discovered by data mining and cloning. The peptides in question are those with some features of vertebrate gonadotropin-releasing hormone (GnRH) but have diverged sufficiently to become unrecognizable as authentic vertebrate GnRHs (Bigot et al., 2012; Iwakoshi et al., 2002; Roch et al., 2014; Zhang et al., 2008). Many of these invertebrate peptides have not undergone functional characterization and, although many of their putative receptors have been identified *in silico*, most ligands and receptors have not been functionally paired through receptor deorphanization. Adding to the confusion, these invertebrate peptides with GnRH-like features appear to segregate into four families with names that will be discussed later: (1) adipokinetic hormone (AKH), (2) corazonin (CRZ), (3) AKH/CRZ-related peptide (ACP), and (4) invertebrate GnRH. The operative word in the last sentence is "appear". Some of these peptides are clearly AKH, CRZ, or ACP, but some have traits that do not conform to the structural requirement of any single family, although they may be more related to a particular one (Hauser and Grimmelikhuijzen, 2014). Further, the use of the name "invertebrate GnRH" carries the connotation that this peptide family is somewhat more related to vertebrate GnRH than to other peptide families, a notion that is currently under debate (see Section 3).

A consensus that can be reached is that these invertebrate peptides share a common origin with vertebrate GnRHs and have evolved extensively in the bilaterian lineage (Hauser and Grimmelikhuijzen, 2014; Plachetzki et al., 2016). They, along with vertebrate GnRHs, can therefore all be considered as members of the GnRH superfamily. Through 700 million years of evolution, they underwent tremendous functional and structural diversification, with only vertebrate GnRH becoming highly specialized in the stimulation of gonadotropin secretion from the pituitary, an endocrine structure unique to the vertebrate lineage (Holland and Sower, 2010). The AKH and CRZ peptides were named based on the functions of the prototype peptides discovered (Stone

E-mail address: Pei-san.tsai@colorado.edu

https://doi.org/10.1016/j.ygcen.2017.09.010
0016-6480/

et al., 1976; Veenstra, 1989), but these names do not consistently reflect their functions in all taxa. The ACP peptides received their name because they have structural features of both AKH and CRZ, but they form a separate family because they have their own unique receptors (Hansen et al., 2010); no function has yet been assigned to this group. The ACP family will not be discussed further since its restricted presence in the arthropods (Zandawala et al., 2017) simplifies its nomenclature. Lastly, the invertebrate GnRH family was originally named because of structural similarities to vertebrate GnRH (Iwakoshi et al., 2002). However, they may be phylogenetically more related to CRZ and are sometimes referred to as CRZ/GnRH (Hauser and Grimmelikhuijzen, 2014).

The quick summary above reveals that there have been no consistent naming criteria for the invertebrate members of the GnRH superfamily, and understandably so. When these peptides were first discovered decades ago, our ability to understand their phylogenetic relationships was limited by the small number of sequences from a few taxa available at the time. As such, most were named using the initial naming criteria of the prototype peptides. Some might argue that naming is a task that gives a mere identity no different than Mary or Paul. However, in the broad context of comparative endocrinology, the name given to a peptide carries a profound implication in its lineage and function that will last for decades to influence our decision on how it should be studied (Markov et al., 2008a). The nomenclature of these peptides, especially within the vertebrate GnRH family initially discovered 40 some years ago (Amoss et al., 1971; Schally et al., 1971) has gone through several rounds of revision in an attempt to clarify their phylogenetic relationship (Fernald and White, 1999; Kah et al., 2007; Kim et al., 2011; Okubo and Nagahama, 2008). The revision during the transition has been substantial, but ultimately an established nomenclature of GnRH1, 2, and 3 for vertebrate GnRH, which now more accurately reflects their evolutionary history, is in place (Decatur et al., 2013; Roch et al., 2014).

The goals of the present article are to suggest several peptide-naming criteria and discuss naming challenges and problem cases using a few examples from the invertebrate GnRH family. This article does not intend to provide a comprehensive review of the GnRH superfamily members and their phylogenies; several recent reviews have already done so (Hauser and Grimmelikhuijzen, 2014; Roch et al., 2014; Zandawala et al., 2017). It is hoped that a neutral and objective nomenclature more reflective of the evolution of the GnRH superfamily will aid in the training of budding comparative endocrinologists and diffuse potential biases in future experimental design and data interpretation.

2. Naming criteria

Invertebrate peptides in the GnRH superfamily have been largely named using 4 criteria: (1) mature peptide structure, (2) prepropeptide (or propeptide) phylogeny, (3) receptor phylogeny, and (4) function. The first three criteria are, of course, interrelated, since a mature peptide's structure also reflects its phylogeny and receptor phylogeny. The last criterion is non-neutral and subjective (Markov et al., 2008a), and is often used to name only a prototype peptide. Following the naming of the original peptide based on function, related peptides that surfaced later were given the same name based on one or more of the first three criteria. For example, arthropod AKH was originally identified in the locust *Schistocerca gregaria* and named for its ability to mobilize lipid stores during flight (Stone et al., 1976). Additional arthropod peptides with structural features meeting the criteria of AKH were given the name AKH. These structural features include: (1) 8 or 10 amino acids in length, (2) a pyroglutamyl N terminus and an amidated C terminus, (3) an aliphatic or aromatic amino acid in Position 2, (4) a W reside in position 8 (W^8), (5) a F^4S^5, F^4T^5, or Y^4S^5 motif, and (6) W^8amide or $W^8G^9X^{10}$amide at the C terminus (Gade et al., 1997) (see Fig. 1). Phylogenetic analysis is difficult at the level of the peptide because of the short length of AKH, but phylogenetic analysis of the AKH prepropeptides or propeptides (Hauser and Grimmelikhuijzen, 2014; Roch et al., 2014) as well as a phylogenomic analysis (Plachetzki et al., 2016) largely support the naming of arthropod AKHs, since they all cluster into a distinct clade. Further, authenticated and putative AKH receptors (AKHR) also segregate into a distinct clade, demonstrating excellent hormone/receptor co-evolution (Hauser and Grimmelikhuijzen, 2014; Kavanaugh and Tsai, 2016; Li et al., 2016; Plachetzki et al., 2016; Roch et al., 2014; Tian et al., 2016; Zandawala et al., 2017). As such, the naming of arthropod AKH is supported by all of the first three naming criteria. The fact that the function "adipokinetic" does not apply to all forms of arthropod AKH (Hauser and Grimmelikhuijzen, 2014) has little consequence, since naming by function provides little insight except a key biological effect in a specific organism. In contrast, the grouping by mature peptide structure, prepropeptide phylogeny and receptor phylogeny clearly illuminates the evolutionary origin and history of the peptides as well as how they relate to each other.

Another example is arthropod CRZ, which was first discovered and named for its ability to stimulate the heart of the American cockroach (Veenstra, 1989). Again, the naming of a prototype CRZ was based on a key function. Subsequently discovered CRZ-like molecules were named CRZ based on the largely invariant structural features of this peptide family in the arthropods. These features are (1) 11 amino acids in length, (2) the conservation of pyroglutamyl (p)Q^1, T^2, Y^5, S^6, G^8, and W^9 residues, and (3) C-terminal amidation. Again phylogenetic and phylogenomic analyses cluster these arthropod peptides and their receptors into distinct clades (Hauser and Grimmelikhuijzen, 2014; Kavanaugh and Tsai, 2016; Li et al., 2016; Plachetzki et al., 2016; Roch et al., 2014; Tian et al., 2016; Zandawala et al., 2017), supporting the their evolutionary origin from a common ancestor.

Unfortunately, the cases are not always as clear-cut as the examples above. The tight structural relationship among members of same peptide family, originally used to name AKH, CRZ, or ACP, begins to break down outside the phylum. Related peptides in other phyla may have structural features of one family or another, but may not fully satisfy all structural requisites of a single family. Under this condition, further clarification of the new peptides' evolutionary relationship with existing members by phylogenetic analyses of prepropeptides and cognate receptors is required before naming the new members. Sufficient naming criteria should

O. vulgaris GnRH	pQNYHFSNGWHPG amide
A. californica GnRH	pQNYHFSNGWYA amide
A. californica AKH	pQ-IHFSPDWGT amide
C. elegans AKH-GnRH	pQ-MTFTDQWT
D. melanogaster AKH	pQ-LTFSPDW-- amide
D. melanogaster CRZ	pQTFQYSRGWTN amide
Vertebrate GnRH2	pQ--HWSHGWYPG amide

Fig. 1. Deduced or verified amino acid sequences of representative members of the GnRH superfamily. *O. vulgaris* GnRH = *Octopus vulgaris* GnRH (oct-GnRH); *A. californica* GnRH = *Aplysia californica* GnRH (ap-GnRH); *A. californica* AKH = *Aplysia californica* AKH (ap-AKH); *C. elegans* AKH-GnRH = *Caenorhabditis elegans* AKH-GnRH (Ce-AKH-GnRH); *D. melanogaster* AKH and CRZ = *Drosophila melanogaster* AKH and CRZ. The names of invertebrate peptides are from the original publications (Iwakoshi et al., 2002; Johnson et al., 2014; Lindemans et al., 2009; Schaffer et al., 1990; Veenstra, 1994; Zhang et al., 2008). Amino acids generally conserved within the GnRH superfamily or among only the invertebrate members are indicated in red or green, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Tsai, P.-S. Gonadotropin-releasing hormone by any other name would smell as sweet. Gen. Comp. Endocrinol. (2017), https://doi.org/10.1016/j.ygcen.2017.09.010