Insect Biochemistry and Molecular Biology 86 (2017) 9–19

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

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

DINeR: Database for Insect Neuropeptide Research

Joseph G.C. Yeoh [a, 1], Aniruddha A. Pandit [a, 1], Meet Zandawala [b], Dick R. Nässel [b], Shireen-Anne Davies [a], Julian A.T. Dow [a, *]

[a] Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, G12 8QQ Glasgow, Scotland, UK
[b] Department of Zoology, Stockholm University, S-10691 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 27 January 2017
Received in revised form
3 May 2017
Accepted 8 May 2017
Available online 11 May 2017

Keywords:
Insect neuropeptide
Database
DINeR
Neuroendocrinology
Neuropeptide hormones
Neuropeptide signalling

ABSTRACT

Neuropeptides are responsible for regulating a variety of functions, including development, metabolism, water and ion homeostasis, and as neuromodulators in circuits of the central nervous system. Numerous neuropeptides have been identified and characterized. However, both discovery and functional characterization of neuropeptides across the massive Class Insecta has been sporadic. To leverage advances in post-genomic technologies for this rapidly growing field, insect neuroendocrinology requires a consolidated, comprehensive and standardised resource for managing neuropeptide information.

The Database for Insect Neuropeptide Research (DINeR) is a web-based database-application used for search and retrieval of neuropeptide information of various insect species detailing their isoform sequences, physiological functionality and images of their receptor-binding sites, in an intuitive, accessible and user-friendly format. The curated data includes representatives of 50 well described neuropeptide families from over 400 different insect species. Approximately 4700 FASTA formatted, neuropeptide isoform amino acid sequences and over 200 records of physiological functionality have been recorded based on published literature. Also available are images of neuropeptide receptor locations. In addition, the data include comprehensive summaries for each neuropeptide family, including their function, location, known functionality, as well as cladograms, sequence alignments and logos covering most insect orders. Moreover, we have adopted a standardised nomenclature to address inconsistent classification of neuropeptides.

As part of the H2020 nEUROSTRESSPEP project, the data will be actively maintained and curated, ensuring a comprehensive and standardised resource for the scientific community. DINeR is publicly available at the project website: http://www.neurostresspep.eu/diner/.

1. Introduction

Neuropeptides and neuropeptide hormones are synthesised by and released from neurons or neuroendocrine cells to trigger a physiological response. In insects, neuropeptides play an important role in coordinating complex homeostatic processes, such as development, metabolism, mating, water and ion homeostasis, reproduction, aggression and are also known to act as neuromodulators in circuits of the central nervous system (Caers et al., 2012; Nassel and Winther, 2010; Schoofs et al., 2017; Terhzaz et al., 2015). Since the discovery of the first insect neuropeptide, proctolin, in the American cockroach (Starratt and Brown, 1975), insect neuroendocrinology has progressed rapidly.

Neuropeptides are produced from larger precursor proteins which are known as prepropeptides (Fig. 1). Prepropeptides comprise of a signal peptide (which directs the protein to the secretory pathway), progenitors of mature peptides (the biologically-active peptides), spacer peptides (peptide fragments with no known biological function and non-conserved sequences) and cleavage sites (monobasic and dibasic) (Fig. 1). A useful website for predicting prepropeptide cleavage sites is Neuropred (Southey et al., 2006) (http://stagbeetle.animal.uiuc.edu/cgi-bin/neuropred.py). About 50 neuropeptide precursor-encoding genes are known in each insect species, with some species having a larger complement of precursors than others (Hauser et al., 2010). Each precursor can give rise to one or more mature neuropeptides or peptide

Abbreviations: DH31, diuretic hormone 31; DINeR, Database for Insect Neuropeptide Research.

* Corresponding author.
E-mail address: Julian.Dow@glasgow.ac.uk (J.A.T. Dow).

[1] Joseph G.C. Yeoh and Aniruddha A. Pandit contributed equally.

http://dx.doi.org/10.1016/j.ibmb.2017.05.001

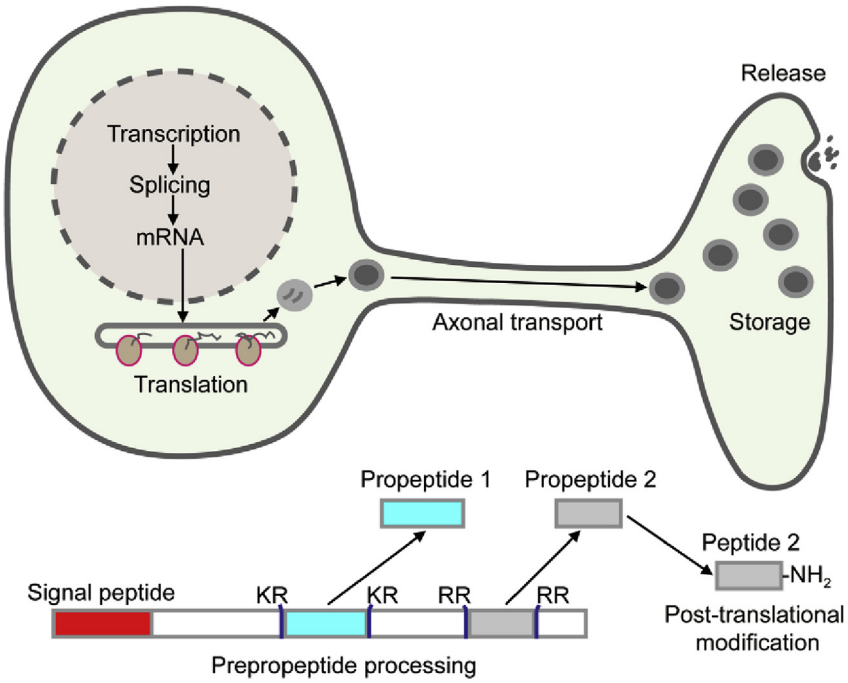

Fig. 1. Neuropeptide production starts in the nucleus and ends in the dense core vesicle. Neuropeptides are produced as part of larger precursor proteins, known as prepropeptides, which are encoded in the genome. These can give rise to one or several bioactive peptides. Neuropeptide encoding genes are transcribed in the nucleus. After splicing, mRNA is translated on ribosomes and with the aid of a signal peptide the immature prepropeptide is incorporated in the secretory pathway and ends up in vesicles. As the vesicles are transported to the axon termination the precursor is processed. The bioactive peptides are each surrounded by mono- or dibasic cleavage sites such as KR or RR shown here, that direct peptidases to enzymatically liberate the peptides. The white boxes represent non-conserved sequences (spacing regions) between peptide progenitors. Finally, posttranslational modifications may occur, such as C-terminal amidation (-NH_2) shown here. The mature neuropeptides are stored in dense core vesicles in the axon termination. A depolarization of the axon termination followed by Ca^{2+} influx triggers release of the peptide. This figure was redrawn and modified from Fig. 11.1 in Nässel and Larhammar (2013).

hormones. The number of mature peptides produced from a given precursor can vary from one insect species to another. These can either be (1) a set of very similar peptides with partly conserved sequences and thus similar receptor activation properties, or (2) in some cases peptides with distinct sequences and functions. Examples of the former are tachykinin-related peptides, AstBs and FMRFamides that exist in multiple closely related forms in the precursors. Examples of precursors containing peptides with distinct sequences and functions (bind distinct receptors) are those of CAPA/Pyrokinin, NPLP1 and vasopressin (Nassel and Winther, 2010; Stafflinger et al., 2008). There are also examples of peptides with similar sequences being produced by paralogs and splice variants. Insulin-like peptides are encoded by up to 38 paralogous genes in the moth *Bombyx mori* and 8 genes in *Drosophila* (Mizoguchi and Okamoto, 2013). The orcokinin gene in insects produces two different neuropeptide precursors by alternative splicing: orcokinin A and orcokinin B (Jiang et al., 2015; Sterkel et al., 2012). A typical prepropeptide and its biosynthesis and processing is shown in Fig. 1. At present, there are at least 50 well-described neuropeptide families identified from numerous species across different insect Orders. However, the wealth of neuropeptide information has generated a problem.

The naming of insect neuropeptide families has created confusion in the literature. Traditionally, neuropeptides were, in many cases, named after their first described function. However, neuropeptides can have multiple functions and thus, the same neuropeptide family might attract several names. For example, the first allatostatins identified (from *Diploptera punctata*) were named because they inhibited juvenile hormone production (Woodhead et al., 1989). Lorenz et al. (1995) then found similar neuropeptides with similar inhibitory properties in crickets, but with slightly different functional groups. Thus, this new group was named cricket type allatostatin or allatostatin B (AstB), and the original allatostatins were designated allatostatin A (AstA) or FGLa type allatostatin (FGLa/AST). However, a few years earlier, AstB orthologs were independently identified in the migratory locust, *Locusta migratoria,* found to have myoinhibiting properties, and named Myoinhibitory Peptide (or MIP) (Schoofs et al., 1991). Additionally, the allatostatin B/MIP peptides have also been shown to be the ancestral ligands for the sex peptide receptor (Kim et al., 2010; Poels et al., 2010). It is thus possible for researchers to encounter the same peptide family in entirely different contexts, and be unaware of the functional pleiotropy. Moreover, additional families of peptides with allatostatic activity, 'AstC' (also known as *Manduca* type allatostatin or PISCF/AST) and 'AstCC' (Kramer et al., 1991; Veenstra, 2009), have been found, further complicating the nomenclature. It is important to note here that although all the three allatostatins (AstA-AstC) may be found in a species, so far only one has been shown to display allatostatic properties in that species (Coast and Schooley, 2011; Nassel and Winther, 2010).

There is also a need to ensure that the different isoforms from the same species can be correctly identified and curated, as well as different isoforms in other species (interspecific isoform). The diuretic hormone, DH31, shows remarkable conservation throughout the insects. All DH31 sequences are 31 amino acids in length and the full sequence is important for DH31 to function (Zandawala, 2012). However, other neuropeptides show a higher degree of variability, both within and between species. For example, eight different kinin amino acid sequences have been found in the Madeiran cockroach, *Leucophaea maderae* (Holman