Toxicon 55 (2010) 1491–1509

Contents lists available at ScienceDirect

Toxicon

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

Conopeptide characterization and classifications: An analysis using ConoServer

Quentin Kaas, Jan-Christoph Westermann, David J. Craik*

The University of Queensland, Institute for Molecular Bioscience, Division of Chemistry and Structural Biology, Brisbane, 4072 QLD, Australia

ARTICLE INFO

Article history:
Received 29 October 2009
Received in revised form 25 February 2010
Accepted 1 March 2010
Available online 6 March 2010

Keywords:
Conus venom
Conopeptide
Conotoxin
ConoServer
Cone snail venom

ABSTRACT

Cone snails are carnivorous marine gastropods that have evolved potent venoms to capture their prey. These venoms comprise a rich and diverse cocktail of peptide toxins, or conopeptides, whose high diversity has arisen from an efficient hypermutation mechanism, combined with a high frequency of post-translational modifications. Conopeptides bind with high specificity to distinct membrane receptors, ion channels, and transporters of the central and muscular nervous system. As well as serving their natural function in prey capture, conopeptides have been utilized as versatile tools in neuroscience and have proven valuable as drug leads that target the nervous system in humans. This paper examines current knowledge on conopeptide sequences based on an analysis of gene and peptide sequences in ConoServer (http://www.conoserver.org), a specialized database of conopeptide sequences and three-dimensional structures. We describe updates to the content and organization of ConoServer and discuss correlations between gene superfamilies, cysteine frameworks, pharmacological families targeted by conopeptides, and the phylogeny, habitat, and diet of cone snails. The study identifies gaps in current knowledge of conopeptides and points to potential directions for future research.

1. Introduction

Marine snails of the *Conus* genus are a large family of carnivorous gastropods that possess a highly complex arsenal of toxins for prey capture and self defense. For over 30 years (Endean et al., 1974; Gray et al., 1981; Olivera and Cruz, 2001), cone snail toxins, or conopeptides, have stimulated interest in their remarkable molecular diversity and capacity to target neuroreceptors, ion channels and transporters, with both potency and specificity (Terlau and Olivera, 2004; Janes, 2005; Olivera et al., 2008). Conopeptides serve as valuable probes for neurophysiological studies (Olivera, 1997; Olivera and Cruz, 2001; Dutton and Craik, 2001; Lewis, 2009), and they provide lead compounds for drug discovery (Adams et al., 1999; Livett et al., 2004, 2006; Terlau and Olivera, 2004; Olivera, 2006; Craik and Adams, 2007; Vincler and McIntosh, 2007; Twede et al., 2009). For instance, the conopeptide MVIIA (Olivera et al., 1985) is used clinically, under the name "Prialt", for the treatment of neuropathic pain (Miljanich, 2004). Xen2174, an analog of the conopeptide MrIA from *Conus marmoreus* (McIntosh et al., 2000; Sharpe et al., 2001), entered Phase II clinical trials for the treatment of acute pain in September 2008 (Xenome Ltd, http://www.xenome.com).

Most conopeptides have a sequence length of 12–35 amino acids. They have a high frequency of post-translational modifications, which generates a rich chemical diversity that may in part explain their outstanding specificity for macromolecular targets. They are broadly divided into disulfide-rich conopeptides, also termed conotoxins, which have two or more disulfides, and disulfide-poor conopeptides, with none or one disulfide bond. Three other classification schemes are also used to describe different aspects of conopeptides, as illustrated in

* Corresponding author. Tel.: +61 7 3346 2019; fax: +61 7 3346 2101.
E-mail address: d.craik@imb.uq.edu.au (D.J. Craik).

0041-0101/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.toxicon.2010.03.002

Fig. 1. The "gene superfamily" classification scheme focuses on evolutionary relationships between conopeptides; the "cysteine framework" scheme sorts them according to the arrangement of cysteines (further explained in the Appendix); and the "pharmacological family" scheme reflects the target specificity of each conopeptide.

As demonstrated by the existence of more than 3700 peer reviewed articles referring to conopeptides[1], the research community studying or using conopeptides is large and active. Based on early estimates of 100–200 distinct toxins (Olivera, 2002) in each of the >500 cone snail species (Röckel et al., 1995; Duda et al., 2009a), the number of conopeptides certainly exceeds 50,000 (Olivera, 2002). Currently, less than 2% of this proposed number has been discovered. Recent data even suggest that the number of species might have been dramatically underestimated, as the traditional identification method, based on the shell color patterns, has failed to distinguish between species clearly identified at the genetic level in a number of cases (Duda et al., 2008, 2009a). This observation and recent reports indicating that the number of conopeptides per species might exceed 1000 (Davis et al., 2009) suggest that conopeptide diversity may be much larger than expected. The large number of studies and the massive amount of available and forthcoming data on conopeptides prompted us to develop a database, ConoServer (http://www.conoserver.org), to help catalog the growing literature on conopeptide sequences and three-dimensional structures (Kaas et al., 2008). This database has now grown substantially and provides a valuable resource for analyzing the sequence and structural features of conopeptides.

This article presents an in-depth analysis of Cono Server's content. Fig. 1 gives an overview of conopeptide biosynthesis and physiological actions, and summarizes the layout of this article. After a brief introduction to conopeptides (Section 1), Section 2 describes the content and features of the database. Subsequent sections discuss conopeptide nucleic acid and protein precursor sequences (Section 3), mature toxin sequences (Section 4), three-dimensional structures (Section 5), and molecular targets (Section 6). The three major conopeptide classification schemes are evaluated in Sections 3, 4 and 6, and their relationships are discussed in Section 7. Numeric facts are stated widely throughout this article to provide a snapshot of the current state-of-the-art of the rapidly growing field of conopeptide research. Furthermore, we have attempted to identify potential directions for future research in areas that have been either overlooked or hardly studied. We hope that the discussions on conopeptide nomenclature and classification schemes will be of general interest to researchers in the broader toxinology field because a similar rapid growth of sequence data has also occurred for toxins extracted from other animals, including spiders, snakes and scorpions.

[1] Based on a search in pubmed performed for articles containing the terms "conotoxin(s)" or "conopeptide(s)".

2. ConoServer

As far as we are aware, ConoServer (http://www.conoserver.org) is the only public database that specializes in conopeptide sequences and three-dimensional structures (Kaas et al., 2008). It complements several other websites that provide valuable general information on cone snail species and their toxins. For instance, "the *Conus* Biodiversity website" (http://biology.burke.washington.edu/conus/) is a valuable resource for cone snail taxonomy and nomenclature. Bruce Livett's "cone shell conotoxins" website (http://grimwade.biochem.unimelb.edu.au/cone/) is another popular resource providing links, news, and major literature references on conopeptides. ConoServer was designed to address a lack of convenient access to the growing volume of sequence data on conopeptides, which is required to compute up-to-date statistics on conopeptide discovery, to name conopeptides and thus avoid duplication of names, and to assist in the coherent development of conopeptide classifications.

The technical description of the database structure has been published elsewhere in communication format (Kaas et al., 2008). Here, we provide the first complete description of the database content, the nomenclature scheme used in ConoServer, and the curation process. Two new features have been recently added to the database. The first is a description of cone snail species that was introduced to help correlate venom content and species characteristics. The second feature is a flag that allows users to discriminate conopeptides isolated at the peptide level or predicted from nucleic sequences. The species descriptors and experimental characterization flags are used extensively in the following analyses.

2.1. Database content

As of November 2009, ConoServer contains data for 3364 conopeptide sequences (1001 mature peptides, 943 protein precursors, 138 synthetic peptides, and 1282 patented sequences), 1653 nucleic acid sequences (including 738 patented sequences), and 126 three-dimensional structures from 85 species of cone snails. These data form the basis for the following analysis. It should be noted that some of the sequences and structures found in ConoServer have been directly submitted by their authors to public databases with no associated peer reviewed reference, but the majority of the data are associated with peer reviewed articles.

2.2. Curation process

ConoServer uses data retrieved from publicly accessible nucleic acid sequence databases, protein sequence databases, protein three-dimensional structure databases (NCBI databases, UniProtKB/Swiss-Prot, and wwPDB), and the refereed scientific literature. Data quality in ConoServer is ensured mainly through manual cross-checking of the original literature, and in several instances this has led to the correction of errors or ambiguities in other data collections. The inversion of toxin names (e.g. Lp1.8 and Lp1.7 (Yuan et al., 2007)) and sequence errors (e.g. PIIIE