Toxicon 107 (2015) 317–326

Contents lists available at ScienceDirect

Toxicon

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

Overview of the Knottin scorpion toxin-like peptides in scorpion venoms: Insights on their classification and evolution

Carlos E. Santibáñez-López*, Lourival D. Possani

Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, Apartado Postal 510-3, Cuernavaca Morelos 62210, Mexico

ARTICLE INFO

Article history:
Received 7 May 2015
Received in revised form
21 June 2015
Accepted 23 June 2015
Available online 14 July 2015

Keywords:
Bayesian inference
Evolution
Phylogenetics
Scorpion toxin

ABSTRACT

Scorpion venoms include several compounds with different pharmacological activities. Within these compounds, toxins affecting ion channels are among the most studied. They are all peptides that have been classified based on their 3D structure, chain size and function. Usually, they show a spatial arrangement characterized by the presence of a cysteine-stabilized alpha beta motif; most of them affect Na^+ and K^+ ion-channels. These features have been revised in several occasions before, but a complete phylogenetic analysis of the disulfide containing peptides is not been done. In the present contribution, two databases (Pfam and InterPro) including more than 800 toxins from different scorpions were analyzed. Pfam database included toxins from several organisms other than scorpions such as insects and plants, while InterPro included only scorpion toxins. Our results suggest that Na^+ toxins have evolved independently from those of K^+ toxins no matter the length of the peptidic chains. These preliminary results suggest that current classification needs a more detailed revision, in order to have better characterized toxin families, so the new peptides obtained from transcriptomic analyses would be properly classified.

1. Introduction

Scorpions are unique arachnids in many ways. They appeared in the Silurian period around 430 millions of years ago (Sissom, 1990; Dunlop, 2010; Waddington et al., 2015), and since then, they have evolved into very complex animals. For example, the only so far published scorpion genome (*Mesobuthus martensii*) contains 32,016 protein-coding genes, representing the biggest genome among sequenced arthropods. The genomic and transcriptomic analyses of this scorpion allowed studying important genetic features related with prey, behavior and detoxification (Cao et al., 2013). Also, the finding of a duplication of Hox genes paralogues, and their subsequent neofunctionalization suggests the existence of an extreme heteronomy in scorpions (Sharma et al., 2014), which makes them different from the rest of the arachnids.

Interesting aspects of scorpions include the fact that their venoms are among the most studied from venomous arthropods. Scorpion venoms are libraries of several different compounds, including organic molecules with enzymatic activity (hyaluronidase, phsopholipase, proteases), toxic and cytolytic peptides, biogenic amines, free amino acids, carbohydrates, lipids and others of unknown function. Their molecular weights are variable, from several thousand Daltons to short linear peptides (Smith et al., 2011; Ma et al., 2012; Rodríguez de la Vega et al., 2013). The presence of toxic peptides that affect many organisms, including humans, are one of the major reasons why there is such interest in studying scorpion venoms (Possani et al., 1999; Rodríguez de la Vega and Possani, 2004, 2005). In addition, recent discoveries show that among these components are substances with pharmacological activities that might be leading components for potential drug development (e.g. Rodríguez de la Vega et al., 2010; Ortiz et al., 2015 and references therein). In this contribution emphasis is placed on the bioactive peptides for which a phylogenetic analysis and evolutionary considerations are reviewed and reported.

1.1. Scorpion venom DBPs and NDBPs

Most of the studied scorpion venom peptides are ion channel blockers or modulators. They (mostly but not restrictedly) affect K^+

* Corresponding author. Instituto de Biotecnología, UNAM, Avenida Universidad, 2001 Apartado Postal 510-3, Cuernavaca, MOR 62210, Mexico.
E-mail address: cae@ibt.unam.mx (C.E. Santibáñez-López).

http://dx.doi.org/10.1016/j.toxicon.2015.06.029
0041-0101/

or Na^+ channels and fall into the 3001–4500 Da and 6001–7500 Da molecular weight ranges (Rodríguez de la Vega et al., 2013).

There is an important number of peptides described so far from scorpion venoms, which were proposed to be generically classified as disulfide bridge-containing peptides (DBPs or Knottin scorpion toxin-like peptides) and non-disulfide bridge-containing peptides (NDBPs; Zeng et al., 2005). This classification takes into consideration the three-dimensional folding of the peptides. However, other classifications suggested are based on the size of the peptide chain: short-chain peptides (e.g. K^+ blockers) and long-chain peptides (e.g. Na^+ channel modulators (Possani et al., 1999; Tytgat et al., 1999; Rodriguez de la Vega and Possani, 2004; Quintero-Hernández et al., 2013).

NDBPs represent a completely different arsenal with multifunctional activities such as antimicrobial, anticancer, hemolytic, anti-inflammatory and others (Almaaytah and Albalas, 2014). Mass fingerprint studies have suggested that these peptides represent nearly a third of all peptides in scorpion venoms (Valdez-Velazquez et al., 2013). NDBPs peptides are composed of 13–56 amino acid residues, exhibiting a marked diversity in their sequences. The majority of these peptides display an α-helical structure, with three subclasses regarding the organization of this structure within the body of their mature peptide (recently reviewed by Almaaytah and Albalas, 2014).

1.2. DBPs toxins

DBPs contain three to four disulfide bridges and are usually folded with a Cysteine-stabilized α/β (Csαβ) motif, which consists of an α-helix joined by disulfide bridges to a double or triple β-sheet structure (Rodriguez de la Vega et al., 2013; Mouhat et al., 2004). Peptides within this classification, in order of medical relevance, are the ones affecting Na^+, K^+, Ca^{++} and Cl^- channels, respectively (Possani et al., 1999; Quintero-Hernández et al., 2013).

1.2.1. Na^+ channel toxins

These "long-chain" peptides are responsible for the neurotoxic effects during scorpion envenomation. They have molecular weights in the range of 6500–8500 Da, containing from 58 to 76 amino acid residues (reviewed in Rodríguez de la Vega and Possani, 2005). The three dimensional arrangement of these peptides consists of a common structural core with six cysteines forming three conserved disulfide bridges, and a fourth disulfide bridge can be formed in three different arrangements (reviewed in Mouhat et al., 2004; Rodríguez de la Vega and Possani, 2004 and Quintero-Hernández et al., 2013). These toxins are also classified into two categories, based on their physiological effects and binding properties: α- and β-toxins. The α-toxins bind to the Na^+ channels at receptor site 3 on the extracellular surface of the channel, inhibiting the fast inactivation process. The β-toxins bind to receptor site 4 shifting the threshold of the channel activation (reviewed in Rodríguez de la Vega and Possani, 2007; Weinberger et al., 2010; Quintero-Hernández et al., 2013).

These toxins are also sub-classified according to their main targets. While α-toxins are divided into three distinct groups: 1) classical α-toxins (highly active only in mammals), 2) anti-insect α-NaScTXs (scorpion sodium channel-specific toxins), highly active only on insect voltage-gated sodium channels (VGSCs) and 3) α-like toxins (active on both insect and mammalian VGSCs; recently reviewed in Quintero-Hernández et al., 2013). β-toxins are subdivided into four classes (reviewed in Pedraza-Escalona and Possani, 2013; Quintero-Hernández et al., 2013): 1) anti-mammalian β-toxins, found only in venom of *Centruroides* species, 2) β-toxins active on insect and mammalian VGSCs, 3) anti-insect selective excitatory β-toxins and 4) anti-insect selective depressant toxins.

1.2.2. K^+ channel toxins

Toxic peptides acting on K^+ channels vary in length from 23 to more than 64 amino acids with estimated molecular weights usually around 4000 Da (Bergeron and Bingham, 2012). They are classified based on primary amino acid sequences and the cysteine pairing (Tytgat et al., 1999; Rodríguez de la Vega and Possani, 2004; Quintero-Hernández et al., 2013). Four families are recognized: α-, β-, γ- and κ-families, with the Csαβ motif (except for the κ-family).

The α-family is considered the largest within K^+ channel toxins, and it is subdivided into 30 subfamilies based on differences in the amino acid sequences (Diego-García et al., 2013; Quintero-Hernández et al., 2013). They adopt two different ways to interact with the K^+ channels: 1) the "pore plugin mode", in which the ion conduction in selective Kv1, KCa2 and KCa3 channels is blocked by a Lysine assisted by the aromatic side chain of a residue located some 7 amino acids apart; and 2) the "intermediate" mode, where a negatively charged extracellular loop of the KCa2 channel makes contact with a patch of basic residues of the toxin, stabilizing the binding (Rodríguez de la Vega et al., 2013; Quintero-Hernández et al., 2013).

Peptides of the β-family are long-chain toxins of 50–75 amino acid residues, which are further subdivided into three groups. The first two groups are characterized by similarity between the toxins and their apparent three dimensional structure (i.e. TsTX-Kβ related peptides, or BmTXKβ peptides), and the third group called Scorpine-like peptides or "orphan" peptides, which possess two structural and functional domains (see Diego-García et al., 2007; Quintero-Hernández et al., 2013). The Scorpine homologs are nowadays more studied due to their antimicrobial effects (i.e. Feng et al., 2013).

Peptides of the γ-family are also known as ergtoxins, because they block the ERG-K^+-channels of nerve, heart and endocrine cells (Bottiglieri et al., 2000). These short-chain toxins are 42–47 amino acid residues (but see also Corona et al., 2002) with four disulfide bridges, and the conserved Csαβ motif.

Finally, peptides of the κ-family, or hefutoxins, are short-chain toxins containing around 22 amino acid residues (Srinivasan et al., 2002). These toxins do not exhibit the conserved Csαβ but a Csαα motif (Quintero-Hernández et al., 2013).

1.2.3. Cl^- channel

Although some authors have shown some structural similarities between chlorotoxins and some α-KTXs toxins, (Huys et al., 2004); peptides of this family are short scorpion toxins of about 30–40 amino acid residues forming four disulfide bridges with a Csαβ motif, that affect chloride permeability (DeBin and Strichartz, 1991; Lippens et al., 1995; Fu et al., 2005; Rjeibi et al., 2011; Diego et al., 2014).

There are some studies (e.g. Peigneur et al., 2015) that found that some toxins (e.g. Ts1) exert a plethora of pharmacological effects on different Na_v channel isoforms (Peigneur et al., 2015). However, as we mention below, the current classification of toxins needs a major revision.

1.3. Knottin, scorpion toxin-like peptides domain

DBPs toxins are classified differently in two of the most important protein online resources: Pfam and InterPro (Table 1). Pfam (version 27.0) is a large collection of curated protein families based on UniProt (http://www.uniprot.org/), which are classified into conserved domains (Finn et al., 2014). In the other hand, InterPro is a resource that classifies proteins into families, and predicts the presence of domains and important sites. To do so,