Toxicon 51 (2008) 1479–1489

Contents lists available at ScienceDirect

Toxicon

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

Molecular diversification based on analysis of expressed sequence tags from the venom glands of the Chinese bird spider *Ornithoctonus huwena*☆, ☆☆

Liping Jiang [1], Li Peng [1], Jinjun Chen, Yongqun Zhang, Xia Xiong, Songping Liang *

Key Laboratory of Protein Chemistry and Developmental Biology of the Ministry of Education, College of Life Sciences, Hunan Normal University, Changsha 410081, PR China

ARTICLE INFO

Article history:
Received 20 January 2008
Received in revised form
22 March 2008
Accepted 25 March 2008
Available online 27 March 2008

Keywords:
Ornithoctonus huwena
Spider
Venom
Peptide toxin
Expressed sequence tag
Diversity
Evolution
Gene family

ABSTRACT

The bird spider *Ornithoctonus huwena* is one of the most venomous spiders in China. Its venom has been investigated but usually only the most abundant components have been analyzed. To characterize the primary structure of *O. huwena* toxins, a list of transcripts within the venom gland were made using the expressed sequence tag (EST) strategy. We generated 468 ESTs from a directional cDNA library of *O. huwena* venom glands. All ESTs were grouped into 24 clusters and 65 singletons, of which 68.00% of total ESTs belong to toxin-like sequences, 13.00% are similar to body peptide transcripts and 19.00% have no significant similarity to any known sequences. Precursors of all toxin-like sequences can be classified into eight different superfamilies (HWTX-I superfamily, HWTX-II superfamily, HWTX-X superfamily, HWTX-XIV superfamily, HWTX-XV superfamily, HWTX-XVI superfamily, HWTX-XVII superfamily, HWTX-XVIII superfamily) except HWTX-XI and HWTX-XIII, according to the identity of their precursor sequences. The results have predictive value for the discovery of various groups of pharmacologically distinct toxins in complex venoms, and for understanding the relationship of spider toxin evolution based on the diversification of cDNA sequences, primary structure of precursor peptides, three-dimensional structure motifs and biological functions.

1. Introduction

Spider venoms are a rich source of novel pharmacologically and agrochemically interesting compounds that have received increased attention from pharmacologists and biochemists in recent years (Escoubas, 2006; Escoubas et al., 2000). A conservative estimation of 500 peptides per venom would lead to a total of 19,000,000 toxins for the 38,000 known spider species (Escoubas, 2006). Spider venoms can be regarded as a complex natural library of polypeptide components.

The spider *Ornithoctonus huwena* is mainly distributed in the hilly areas of Guangxi and Yunnan in the south of China. The venom from the spider *O. huwena* contains a mixture of compounds. About 400 proteins and peptides from the venom can be separated and detected by 2D

☆ Sequence data from this article have been deposited with the GenBank Data Library under genbank:EU195228-EU195294.

☆☆ *Ethical statement*: On behalf of, and having obtained permission from all the authors, I declare that: (a) the material has not been published in whole or in part elsewhere; (b) the paper is not currently being considered for publication elsewhere; (c) all authors have been personally and actively involved in substantive work leading to the report, and will hold themselves jointly and individually responsible for its content; (d) all relevant ethical safeguards have been met in relation to patient or subject protection, or animal experimentation. I testify to the accuracy of the above on behalf of all the authors.

* Corresponding author. Tel.: +86 731 8872556; fax: +86 731 8861304.
E-mail address: liangsp@hunnu.edu.cn (S. Liang).

[1] These authors contributed equally to this paper.

0041-0101/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.toxicon.2008.03.024

electrophoresis. Using a combination of ion-exchange and reverse phase high performance liquid chromatography (HPLC), 14 toxins have been isolated and sequenced from this spider venom (Liang, 2004; Peng et al., 2001, 2002; Liu et al., 2006a, b; Shu and Liang, 1999; Shu et al., 2002; Lu et al., 1999). They have had different types of biological activities, including N-type Ca^{2+} channel inhibitor (HWTX-I and HWTX-X) (Peng et al., 2001; Liu et al., 2006a, b), insecticidal neurotoxins (HWTX-II, HWTX-VII and HWTX-VIII) (Shu and Liang, 1999; Shu et al., 2002), TTX-sensitive Na^{+} channel blocker (HWTX-IV) (Peng et al., 2002), the smallest lectin-like peptide (SHL-I) (Lu et al., 1999) and others. HWTX-I is the most abundant toxic component in the crude venom of *O. huwena*. The structure of it contains an 'inhibitor cystine-knot' (ICK) motif with a disulfide bonding pattern pairing of I–IV, II–V, III–VI (Qu et al., 1995). HWTX-II forms a I–III, II–V, IV–VI disulfide connectivity and disulfide-directed beta-hairpin (DDH) structural motifs (Shu et al., 2002).

By overlapping the two partial cDNA sequences obtained by 3′ and 5′ rapid amplification of cDNA ends (RACE), eight cDNAs encoding seven toxins, HWTX-I, -II, -III, -IIIa, -IV, -V, VII and one lectin, SHL-I, from the spider *O. huwena*, were cloned and sequenced. This indicates that the seven peptides from *O. huwena* could be classified into two different superfamilies according to the prepro region of cDNA sequences (Liang, 2004; Diao et al., 2003).

Sequence determination and precise pharmacological analysis of peptide toxins from spiders or other venomous animals of interest are frequently restricted or unsuccessful because of the difficulties of obtaining sufficient venom material (Liang, 2004; Satake et al., 2004). To overcome these problems, the screening of venom gland cDNAs is a promising method with regard to efficiency of characterization of potential peptide toxins from venomous animals such as spider (Satake et al., 2004; Kozlov et al., 2005), cone snail (Pi et al., 2006), snake (Zhang et al., 2006; Wagstaff and Harrison, 2006; Junqueira-de-Azevedo Ide and Ho, 2002) and venomous fish (Magalhães et al., 2006). The EST approach provides a rapid and reliable method for gene discovery as well as a resource for the large-scale analysis of gene expression of known and unknown genes. To accomplish this task, a cDNA library was generated from venom glands and partial sequencing of the cDNAs was performed, generating hundreds of ESTs. In the current report we show the complete analysis of 468 ESTs from *O. huwena* venom glands, that can be classified into eight different superfamilies according to the identity of their precursor sequences.

2. Materials and methods

2.1. cDNA library construction

A directional full-length cDNA library was generated from the venom gland of *O. huwena*. Spiders *O. huwena* were collected in Guangxi Province, China. Spiders were milked to stimulate the production of messenger RNAs (mRNAs) in the venom glands. After 4 days, venom glands (about 1 g) of 20 individual spiders were isolated and immediately frozen in liquid nitrogen by grinding. The total RNA was extracted with Trizol Reagent (Invitrogen). The integrity of total RNA was checked by discerning the 28S and 18S bands of ribosomal RNA in a formaldehyde denaturing 1% agarose gel. The mRNA was isolated with an oligo-dT cellulose column (Promega). The PCR-based cDNA library was made following the instructions for the SMART cDNA library construction kit (Clontech, Palo Alto, CA, USA). Competent *Escherichia coli* XL1-Blue cells were transformed with the cDNA library plasmids to amplify the cDNA.

2.2. DNA sequencing and bioinformatics analysis

The resulting colonies were randomly picked, and the inserted cDNAs in the individual colonies were directly amplified by colony PCR using universal M13 forward and reverse primer sets. The PCR products were resolved by agarose gel electrophoresis to determine the size of each product. Selected clones with cDNAs >400 bp were conducted with standard M13 forward primers on an ABI 3730 automatic DNA sequencer according to the manufacturer's instructions (completed by Shanghai Sangon Biological Engineering Technology and Service Co., Ltd.). The clustering and assembling of ESTs was processed by a universal tool of the EGassembler (http://egassembler.hgc.jp/) (Masoudi-Nejad et al., 2006). *O. huwena* venom gland sequences (clusters and singletons) were searched against public databases (nr/NCBI, Swissprot+TREMBL/EMBL) using the BlastX program with the *e*-value cutoff set to $<10^{-5}$ (Benson et al., 2005) to identify putative functions of the new ESTs. The signal peptide was predicted with the SignalP 3.0 program (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al., 2004). The propeptide cleavage site was ascertained from the known start site of previously characterized mature toxins. The phylogenetic analysis of toxin protein families was conducted by MEGA 3.1 (Kumar et al., 2004).

3. Results

3.1. EST sequencing and clustering

The original cDNA library of the *O. huwena* venom gland contained 1.0×10^{6} independent clones according to the supplier's instructions, which could meet almost all requirements of obtaining a cDNA clone from low abundance mRNA. The average length of cloned cDNA in the library was about 550 bp, ranging from 0.25 to 1.2 kb. To extract the high quality sequence region, vector, primer sequences, poly(A) tails and low-quality sequences were removed. The average readable sequence length was approximately 467 bp. The ESTs assembling resulted in 24 clusters showing more than one EST and 65 singletons. Sixty-seven different toxin precursors deduced from the remaining 468 ESTs were deposited in the GenBank under EU195228 to EU195294. While 31 novel ESTs code different mature peptides.