FEBS Letters 588 (2014) 740–745

FEBS Letters

journal homepage: www.FEBSLetters.org

Genes and evolution of two-domain toxins from lynx spider venom ☆

Maria Y. Sachkova [a,*], Anna A. Slavokhotova [a,b], Eugene V. Grishin [a], Alexander A. Vassilevski [a,*]

[a] M.M. Shemyakin and Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, 16/10 Miklukho-Maklaya, 117997 Moscow, Russian Federation
[b] N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences, 3 Gubkina, 119991 Moscow, Russian Federation

ARTICLE INFO

Article history:
Received 30 October 2013
Revised 2 January 2014
Accepted 10 January 2014
Available online 23 January 2014

Edited by Takashi Gojobori

Keywords:
Molecular evolution
Spider venom gland
Inhibitor cystine knot (ICK)
cDNA library
Toxin gene

ABSTRACT

Spiderines are comparatively long polypeptide toxins (~110 residues) from lynx spiders (genus *Oxyopes*). They are built of an N-terminal linear cationic domain (~40 residues) and a C-terminal knottin domain (~60 residues). The linear domain empowers spiderines with strong cytolytic activity. In the present work we report 16 novel spiderine sequences from *Oxyopes takobius* and *Oxyopes lineatus* classified into two subfamilies. Strikingly, negative selection acts on both linear and knottin domains. Genes encoding *Oxyopes* two-domain toxins were sequenced and found to be intronless. We further discuss a possible scenario of lynx spider modular toxin evolution.

1. Introduction

Spider venom is a source of biologically active polypeptides composing combinatorial libraries [1,2]. Toxin variability is not limited to multiplicity of amino acid sequences within the confines of the same structural motif. In addition to the more common one-domain peptide toxins, elaborate two-domain (or "modular") toxins have been recognized recently.

One-domain toxins usually possess the inhibitory cystine knot (ICK, or knottin) fold, or they are just short linear molecules adopting amphipathic α-helical conformation, with neurotoxic or cytolytic activities, correspondingly. Such "common" toxins may be used as building blocks to construct more complex two-domain molecules. Toxins corresponding to all possible combinations of these structural elements have been described in the venoms of spiders from different families: ICK-ICK (DkTx [3] and CpTx [4]), linear-linear (cyto-insectotoxins [5]), ICK-linear (LtTx, or latartoxins [6], and CsTx [7]), and linear-ICK (spiderines [8]). Intriguingly, different scorpions were also found to produce two-domain toxins referred to as scorpines and β-KTxs [9,10].

Spiderines (OtTx) from the lynx spider *Oxyopes takobius* are long polypeptides (~110 residues) built of an N-terminal linear domain (~40 residues) and a C-terminal ICK domain (~60 residues) linked by a short sequence resembling a mutated processing site [8]. OtTx possess strong cytolytic activity due to the linear domain. The functional role of the knottin domain is yet to be established. The "chimeric" structure of spiderines raises questions about their evolutionary origin.

Oxyopes lineatus is a species closely related to *O. takobius*. Interestingly, both spiders produce the same major single-domain ICK neurotoxin oxytoxin-1 (OxyTx 1) [11,12]. However, a variety of linear cytotoxins (oxyopinins) are present in *O. takobius* venom [11,13] but not in *O. lineatus* that conversely produces oxytoxin-2 (OxyTx 2) absent in *O. takobius* [12]. We therefore decided to mine *O. lineatus* venom glands for spiderines.

Genes encoding spider toxins are poorly studied in part due to the lack of data on spider genome sequences. In the present work we report sequences of genes encoding lynx spider two-domain toxins. Based on these data, we attempt to elucidate a possible scenario of *Oxyopes* two-domain toxin evolution.

2. Materials and methods

2.1. DNA and RNA purification from O. lineatus venom glands

Specimens of *O. lineatus* spider were collected in Kazakhstan by Andrey Feodorov (Fauna Laboratories, Ltd., Republic of Kazakhstan). Venom glands were excised and stored at −70 °C. ~100 mg

Abbreviations: gDNA, genomic DNA; ICK, inhibitor cystine knot; LRT, likelihood ratio test; ML, maximum likelihood; NG method, Nei–Gojobori method; PQM, processing quadruplet; RACE, rapid amplification of cDNA ends

☆ Nucleotide sequences reported in this paper have been deposited in GenBank with accession numbers KF766543–KF766563.

* Corresponding authors. Fax: +7 495 330 73 01.
E-mail addresses: sachkovamasha@mail.ru (M.Y. Sachkova), avas@ibch.ru (A.A. Vassilevski).

http://dx.doi.org/10.1016/j.febslet.2014.01.018
0014-5793/

of venom glands from several individuals were used to extract genomic DNA (gDNA) and total RNA. The gland tissue was dissolved at +55 °C in 500 μl of 50 mM Tris–HCl (pH 8.0), 200 mM NaCl, 100 mM EDTA, and 1% SDS solution containing 1 mg/ml proteinase K. 250 μl of 5 M NaCl were added and the pellet was removed. DNA/RNA mixture precipitation by ethanol was followed by phenol–chloroform extraction. After repeated ethanol precipitation the DNA/RNA mixture was dissolved in DNAase/RNAase-free water. The obtained gDNA was consequently used in PCR reactions. To obtain cDNA, mRNA was reverse transcribed by the MINT kit (Evrogen, Russia) following the manufacturer's recommendations. All cDNA sequences were flanked by adaptor sequences that were then used for rapid amplification of cDNA ends (RACE) (see below).

2.2. PCR amplification of cDNA and gDNA fragments and sequencing

Based on cDNA sequences from the *O. takobius* cDNA library, specific primers were designed to amplify two-domain toxin genes from *O. lineatus* (Table 1). 5′ and 3′ RACE were carried out with the universal primer T7cap and a corresponding specific primer. PCR products were cloned into the pAL-TA plasmid vector (Evrogen) and sequenced (see below). Using the newly established cDNA sequences we designed primers to amplify gDNA fragments encoding *O. lineatus* two-domain toxins (Table 1).

PCR reactions were performed with the Evrogen PCR kit. Purified PCR products were ligated into the pAL-TA plasmid vector that was used for transformation of competent *Escherichia coli* XL1-Blue cells. Positive clones were sequenced with the M13 forward primer. The Lasergene package (DNASTAR, USA) was used for analysis of DNA sequencing results and other manipulations with nucleotide sequences.

2.3. Analysis of cDNA sequences

O. takobius venom gland cDNA library was constructed in collaboration with DuPont Agriculture and Nutrition [8]. *O. lineatus* cDNA was synthesized as described above. To translate cDNA sequences *in silico*, the DNASTAR software was used. Signal peptides were identified by the SignalP 4.1 online tool (http://www.cbs.dtu.dk/services/SignalP/). Propeptide sequences were assigned as preceding the processing quadruplet motif (PQM) cleavage sites [14]. Multiple and pairwise alignments were constructed by the ClustalW program [15] using Vector NTI Suite 8 (Life Technologies, USA) and MEGA 5 software [16]. Search for putative splice sites was performed by the Fruitfly Splice Predictor (http://www.fruitfly.org/seq_tools/splice.html) [17].

The following analyses were performed using MEGA 5. For phylogenetic trees, the evolutionary history was inferred using the neighbor-joining method [18]. The bootstrap test (500 replicates) was performed to calculate the percentage of replicate trees in which the associated sequences clustered together. The evolutionary distances were computed using the maximum composite likelihood method [19]. Codon-based *Z*-test of Purifying/Positive Selection was performed using the Nei–Gojobori (NG) method [20].

Further phylogenetic analysis was executed with the maximum likelihood (ML) method by the CODEML program from the package PAML (version 4.7a) [21] using the graphical interface PAML X [22]. All alignment gaps were deleted for the analyses. For pairwise comparisons, the program was run with runmode = −2 and model M0. For likelihood ratio tests (LRTs), the program was run with runmode = 0 and five models: M0, M1a, M2a, M7, and M8. Branch lengths were estimated by the model M0 and used in all subsequent analyses. The natural logarithm of the maximum likelihood (ln L) was calculated for each model. LRTs were constructed for the pairs M1a/M2a and M7/M8. For each LRT, $2\Delta\ell = 2 \times (\ln L_1 - \ln L_0)$ was computed, where L_0 is the model that does not allow for positive selection (M1a and M7). $2\Delta\ell$ values were compared against χ^2 table for 2 degrees of freedom (critical χ^2 for 2 degrees of freedom and $P < 0.05$ or $P < 0.1$ is 5.9915 or 4.6052, correspondingly).

Similarity search was performed by the BLAST program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) in the Nucleotide and Non-redundant protein sequences NCBI databases. Nucleotide and amino acid sequences of separate domains were used as queries.

3. Results

3.1. Two-domain toxins in lynx spiders

We have recently reported four cDNA sequences of *O. takobius* spiderines [8]. However, the variability of two-domain toxins is wider: in this work we found 9 new sequences belonging to the spiderine family in the *O. takobius* venom glands cDNA library and 6 novel cDNA sequences encoding *O. lineatus* two-domain toxins (OlTx) (GenBank Nos. KF766543–KF766559, KF766561) (see Fig. 1 for the phylogenetic tree).

All the cDNA sequences were translated into protein precursor sequences and processed *in silico* (an example for OlTx 2a is shown in Fig. 2). In total, 15 novel *Oxyopes* two-domain toxin sequences were established belonging to the spiderine family. As many toxins described earlier [1,23], spiderines possess different variation levels in signal, pro- and mature peptides. Signal peptides and propeptides are identical to each other or show a high similarity level (>74% and >62% of identical residues, correspondingly), while mature toxins are more variable (>42% identity). Mature polypeptides can be divided into two subfamilies differing in length: Ox-I (109–114 residues) and Ox-II (135–142 residues). Apart from OtTx 1a, 1b, 2a and 2b reported earlier, 10 newly established sequences were placed to the Ox-I subfamily: OtTx 1c–e, OtTx 2c–f, and OlTx 1a–c. OtTx 3a and 3b and OlTx 2a–c were placed to the Ox-II subfamily. Inside each subfamily, identity between toxins exceeds 68%, while between the subfamilies it is in the range of 42–50%. All Ox-I polypeptides (including OtTx reported earlier [8]) resemble complex precursors with mutated processing sites, but none of Ox-II polypeptides contains sequences of high similarity to the PQM processing site.

Length difference between the Ox-I and Ox-II subfamilies is due to the N-terminal domains of the mature polypeptides: they are composed of 47–55 and 69–73 residues, correspondingly (Fig. 3).

Table 1
Oligonocleotide primers used for amplification of cDNA and gDNA fragments encoding OlTx toxins.

Gene	Template	Direction	Sequence
OlTx1	gDNA	Forward	CGATTTCAAAATGAAGTTCTCTTTGG
		Reverse	TTGAGCAAGAATTCTATACAACTG
	cDNA	Forward	AAGGGTCTAGAGAAAGCAACACC
		Reverse	ATACAACTGGGAGGTCACAGGTG
OlTx2	gDNA, cDNA	Forward	AACATCGTTTTCAAAATGAAGATCG
		Reverse	AGTCTACTCAGCCGCAGGTTCCT
Both	cDNA	Universal primer T7cap	GTAATACGACTCACTATAGGGCAAGCAGTGGTAACAACGCAGAGT

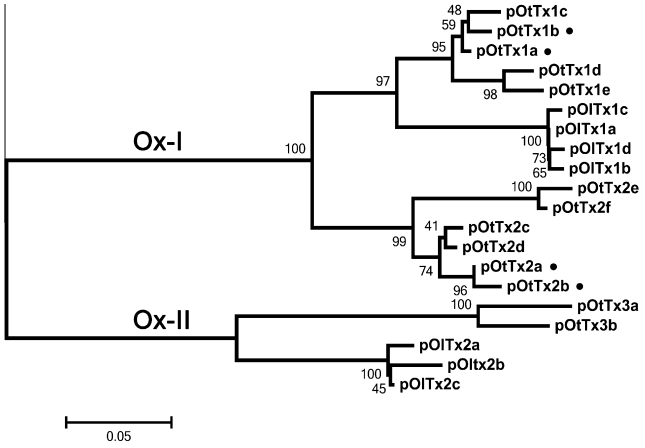

Fig. 1. Neighbour-joining phylogenetic tree generated from alignment of all spiderine precursors known to date. Sequences reported earlier [8] are marked by bold dots. The scale bar corresponds to the number of base substitutions per site along each branch. Bootstrap values are shown next to the branches.

In the linear domain of both Ox-I and Ox-II subfamilies, it is possible to localize "core" sequences that show moderate but significant similarity to each other (at least 30% identity). The N-terminal domains of Ox-II mature polypeptides additionally possess "extra" sequences that augment their length and contain several Gly-Lys-Ser repeats. Knottin domains are more conserved (at least 59% identity), and in some toxins they are even identical (on both protein and cDNA levels).

3.2. Spiderine gene structure

Due to amplification and sequencing of OlTx gDNA fragments, genes encoding OlTx 1a, OlTx 1b, and OlTx 2a were identified. Sequence encoding OlTx 1d (GenBank No. KF766560) was found at the gDNA but not cDNA level. Similarly to most studied single-domain spider toxins [2], *O. lineatus* two-domain toxin genes were found intronless.

3.3. Evolutionary considerations on spiderines

We analyzed evolutionary rates of gene fragments encoding signal peptides, propeptides, and linear and knottin domains of spiderines (20 sequences total) by the Nei–Gojobori method (NG method). Nucleotide sequences encoding protein precursor fragments were aligned and all positions containing gaps were eliminated. Since linear domain sequences significantly differ in length, only their "core" parts were used in the analysis. Unexpectedly, strong purifying selection was found to act not only on signal peptides and propeptides, but also on the mature linear and knottin domains (Table 2). Pairwise sequence comparison detected positive selection ($P < 0.05$) acting on the N-terminal domains only in the pairs OtTx 2a–OtTx 2e and OtTx 2b–OtTx 2e. The latter conclusion does not pass the Bonferroni correction for multiple testing, and therefore the possibility that the result is false positive cannot be excluded.

Further pairwise comparison analysis by the ML method confirmed our observations. The vast majority of N-terminal sequence pairs were found to undergo purifying selection: $\omega \leqslant 0.3$ for 45.2% of the pairs and $0.3 < \omega < 0.9$ for 48.4% of the pairs. However, for the sequence pairs OtTx 2a–OtTx 2e and OtTx 2b–OtTx 2e, ω reached the value of 15.3, and it was >2 for four other pairs, which may be interpreted as positive selection.

Analogous analyses were performed for the C-terminal sequences as well as preprosequences, indicating tendency to evolve under negative selection and showing no sign of positive selection (see Supplementary Table). Thus, our data suggest that genes encoding *Oxyopes* two-domain toxins are well conserved, but positive selection may act on the N-terminal domain. This suggestion was further probed and confirmed by LRTs (Table 3). Both LRTs

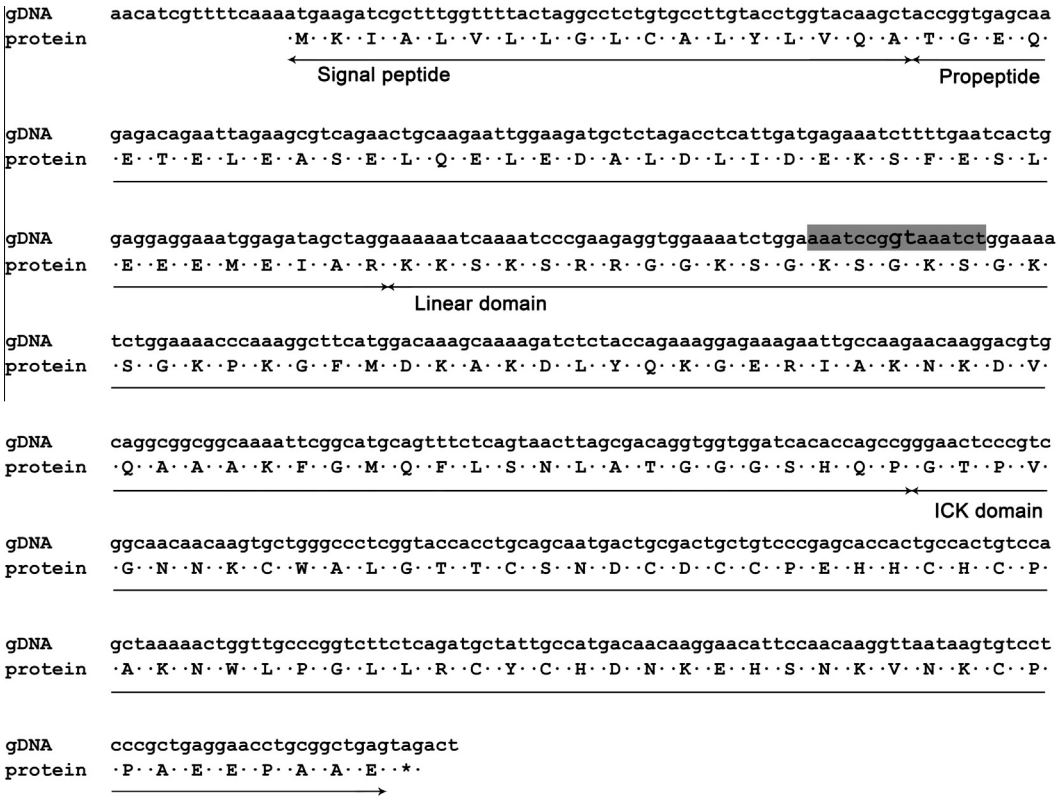

Fig. 2. Gene encoding OlTx 2a. Signal peptide, propeptide, and linear and ICK domains are indicated by arrows. A putative donor splice site is enclosed in a dark gray box.

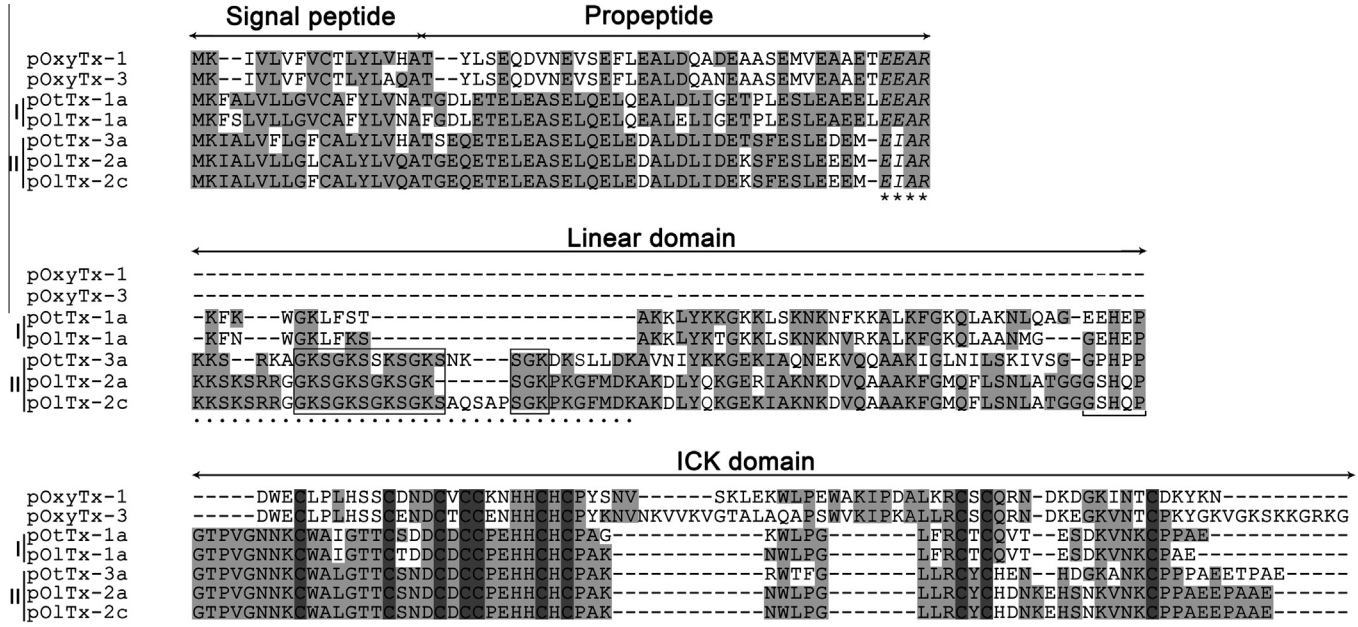

Fig. 3. Alignment of protein precursors of one-domain ICK toxins (OxyTx 1 and OxyTx 3) and two-domain spiderines belonging to the Ox-I (OtTx 1a and OlTx 1a) and Ox-II (OtTx 3a, OlTx 2a and OlTx 2c) subfamilies. Identical residues are shaded in light gray, conserved cysteine residues of the ICK domains are in dark gray. The PQM processing site is in italics and labeled by asterisks, linker sequence between domains is marked by a horizontal bracket. The "extra" sequence of linear domains is indicated by dots below, while the "core" sequence is the fragment between the "extra" and linker sequences. Gly-Lys-Ser repeats are placed in boxes.

Table 2
Results of codon-based *Z*-test of purifying selection for spiderine signal peptides, propeptides, and linear and knottin domains.

	P-value	$d_S - d_N$	Number of sites
Signal peptides	0.039	1.783	18
Propeptides	0.017	2.149	39
Linear domains	0.0001	3.951	46
Knottin domains	0.001	3.216	58

showed that sequences encoding mature polypeptides undergo positive selection, even though the M7/M8 LRT was less significant than the M1a/M2a LRT ($P < 0.1$ and $P < 0.05$, correspondingly). Models M2a and M8 allowing for positive selection indicated only one site Leu-14 (numbering according to OlTx 1a) to be positively selected ($P > 95\%$), and this site belongs to the N-terminal domain.

A search for genes related to ancestors of the mosaic spiderine toxin genes was performed. No significant similarity was detected between the N-terminal linear domain of spiderines and any other known protein. The C-terminal ICK domain, however, is significantly similar to oxytoxins OxyTx 1 and OxyTx 2 [11,12] (~45% identity). *O. takobius* cDNA library was further screened for sequences homologous to spiderines. We extracted sequences encoding protein precursors of linear toxins (including oxyopinins [11,13]) and knottin peptides with 10 cysteine residues (including OxyTx 1 and OxyTx 3, GenBank Nos. KF766562 and KF766563, correspondingly). Alignment of the prepropeptide and N-terminal domain sequences of spiderines to precursors of linear toxins by the BLASTP program showed no significant similarity. On the contrary, prepropeptide and C-terminal domain sequences of two-domain toxin precursors exhibit significant similarity to corresponding sequences of one-domain ICK toxin precursors (~40% identity) (Fig. 3).

Table 3
Results of likelihood ratio tests.

Model	lnL	$2\Delta\ell$	Sites under positive selection
M1a	−1627.942613	6.361392	14 L
M2a	−1624.761917		
M7	−1624.685734	4.749764	14 L
M8	−1622.310852		

4. Discussion

4.1. Diversity of lynx spider two-domain toxins

In the present work we found 16 novel sequences encoding spiderines (Fig. 1). The linear domains make a greater contribution to the variability than knottin domains (Fig. 3). A similar situation is noted among *Lachesana tarabaevi* ICK-linear latartoxins "inverted" with respect to spiderines. For example, for the LtTx 1a and LtTx 1b sequences, the identity is 64% for knottin domains and 33% for linear domains; for LtTx 2a and LtTx 2c, the values are 73% and 46%, correspondingly. It is likely that the higher variability of linear domains is due to their membrane-active function provided merely by characteristic distribution of charge and hydrophobicity that permits formation of amphiphilic α-helices and does not require fixation of specific residues. On the contrary, the 3D structure of knottin domains is maintained by a conserved cysteine motif, and moreover, interaction with a specific protein target necessitates a higher level of sequence conservation. Positive selection analyses revealed that only linear domains of spiderines show a weak tendency to fast adaptive evolution. Among two LRTs, only the M1a/M2a LRT was significant ($P < 0.05$), and only one site (Leu-14) situated in the N-terminal domain was shown to be positively selected. Knottin domains undergo strong purifying selection. Alterations of the "extra" sequence length contribute to the higher variability of linear domains.

To date there have been no reports on analysis of positive/purifying selection acting on linear spider toxins. Among ICK peptides, fast adaptive evolution was reported to act among paralogs of ω-atracotoxin Hv1a from Australian funnel-web spiders [1]. Short cysteine-rich toxins from scorpion and mollusk venoms are also thought to evolve rapidly [23,24]. Only the large protein toxins

from *Latrodectus* and *Steatoda* genera (α-latrotoxin and its homologs) are known to undergo purifying selection [25]. Tendency to negative selection acting on both spiderine domains may indicate their functional importance. OtTx 1a linear domain was shown to be a powerful cytolytic molecule [8], and the knottin domain is presumed to target calcium channels similarly to the homologous oxytoxins.

4.2. Evolution of spiderines

Among translated cDNA sequences of *O. takobius* knottin toxins we found those which putatively have common ancestors with the preprosequence and C-terminal domain of spiderines (Fig. 3). Emergence of *Oxyopes* two-domain toxins is therefore likely a result of an insertion of a DNA fragment encoding the linear domain into a one-domain ICK toxin gene between the parts encoding the preprosequence and mature chain. However, we cannot assume that the N-terminal linear domain was recruited from linear toxins, since its homologs were not detected in either the cDNA library or open databases. To date no spider genome has been sequenced, and we cannot conclude whether the N-terminal domain sequence originated *de novo*, although this scenario seems feasible and quite widespread in evolution [26].

Another possible origin of the N-terminal domain is an intronic sequence. Unlike spiders, many scorpion toxin genes contain introns, the junction site of which is situated at the signal and mature peptide boundary [27,28]. The only spider toxin gene known to contain introns possesses splice site junctions at the signal peptide–propeptide and propeptide–mature peptide boundaries [29]. *Oxyopes* two-domain toxin genes were found intronless. We may therefore hypothesize that the N-terminal domain evolved from an intronic sequence localized between exons encoding the propeptide and mature peptide. Favoring this hypothesis, a donor splice site was predicted at the 5′ terminus of some nucleotide sequences encoding Ox-II linear domains (Fig. 2). Supposedly, a mutation that destroyed the acceptor site led to splicing failure followed by intron evolution into a coding sequence. Many introns of scorpion toxin genes are A/T-rich [27,28]. If converted to a coding sequence, they would have a high content of Lys (AA[ag]), Phe (TT[tc]), Leu (CT[agtc] and TT[ag]) and Asp (AA[tc]) codons. In fact, the gene fragment encoding the N-terminal domain of spiderines is A/T-rich (A + T ∼ 60%). The N-terminal domain sequence of OtTx 1a, for example, contains 39% Lys, 15% Leu, 10% Phe, and 7% Asp (71% total).

To establish a more realistic hypothesis on origination of two-domain toxins, our knowledge of spider genomics must be significantly broadened. Here we analyzed spiderine variability and studied the structure of their genes. Genome sequencing of *Oxyopes* and related genera would permit an in-depth phylogenetic study and clarify routes that led to evolution of diverse two-domain toxins.

Acknowledgments

The authors thank the DuPont Agriculture & Nutrition staff and especially Maureen Dolan, Will Krespan, Bill McCutchen, and Rafi Herrmann for cDNA library construction and sequencing. We are grateful to Irena I. Artamonova (Vavilov Institute of General Genetics, Russian Academy of Sciences), Alexander Y. Panchin (Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences) and three anonymous reviewers for careful consideration of our manuscript and valuable comments that helped us to considerably improve the paper, especially its bioinformatics part. This work was supported by the Russian Foundation for Basic Research (Grant Nos. 11-04-00706 and 12-04-33151), the Program of Molecular and Cell Biology of the Russian Academy of Sciences, and by the Ministry of Education and Science of the Russian Federation (agreement number 8794). AAS and AAV are recipients of the stipend of the President of Russian Federation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2014.01.018.

References

[1] Sollod, B.L., Wilson, D., Zhaxybayeva, O., Gogarten, J.P., Drinkwater, R. and King, G.F. (2005) Were arachnids the first to use combinatorial peptide libraries? Peptides 26, 131–139.
[2] Vassilevski, A.A., Kozlov, S.A. and Grishin, E.V. (2009) Molecular diversity of spider venom. Biochem. (Mosc.) 74, 1505–1534.
[3] Bohlen, C.J., Priel, A., Zhou, S., King, D., Siemens, J. and Julius, D. (2010) A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain. Cell 141, 834–845.
[4] Vassilevski, A.A., Fedorova, I.M., Maleeva, E.E., Korolkova, Y.V., Efimova, S.S., Samsonova, O.V., Schagina, L.V., Feofanov, A.V., Magazanik, L.G. and Grishin, E.V. (2010) Novel class of spider toxin: active principle from the yellow sac spider *Cheiracanthium punctorium* venom is a unique two-domain polypeptide. J. Biol. Chem. 285, 32293–32302.
[5] Vassilevski, A.A., Kozlov, S.A., Samsonova, O.V., Egorova, N.S., Karpunin, D.V., Pluzhnikov, K.A., Feofanov, A.V. and Grishin, E.V. (2008) Cyto-insectotoxins, a novel class of cytolytic and insecticidal peptides from spider venom. Biochem. J. 411, 687–696.
[6] Kuzmenkov, A.I., Fedorova, I.M., Vassilevski, A.A. and Grishin, E.V. (2013) Cysteine-rich toxins from *Lachesana tarabaevi* spider venom with amphiphilic C-terminal segments. Biochim. Biophys. Acta 1828, 724–731.
[7] Kuhn-Nentwig, L., Fedorova, I.M., Lüscher, B.P., Kopp, L.S., Trachsel, C., Schaller, J., Vu, X.L., Seebeck, T., Streitberger, K., Nentwig, W., Sigel, E. and Magazanik, L.G. (2012) A venom-derived neurotoxin, CsTx-1, from the spider *Cupiennius salei* exhibits cytolytic activities. J. Biol. Chem. 287, 25640–25649.
[8] Vassilevski, A.A., Sachkova, M.Y., Ignatova, A.A., Kozlov, S.A., Feofanov, A.V. and Grishin, E.V. (2013) Spider toxins comprising disulfide-rich and linear amphipathic domains: a new class of molecules identified in the lynx spider *Oxyopes takobius*. FEBS J. 280, 6247–6261.
[9] Zhu, S. and Tytgat, J. (2004) The scorpine family of defensins: gene structure, alternative polyadenylation and fold recognition. Cell. Mol. Life Sci. 61, 1751–1763.
[10] Diego-García, E., Schwartz, E.F., D'Suze, G., González, S.A.R., Batista, C.V.F., García, B.I., de la Vega, R.C.R. and Possani, L.D. (2007) Wide phylogenetic distribution of Scorpine and long-chain beta-KTx-like peptides in scorpion venoms: identification of "orphan" components. Peptides 28, 31–37.
[11] Corzo, G., Villegas, E., Gómez-Lagunas, F., Possani, L.D., Belokoneva, O.S. and Nakajima, T. (2002) Oxyopinins, large amphipathic peptides isolated from the venom of the wolf spider *Oxyopes kitabensis* with cytolytic properties and positive insecticidal cooperativity with spider neurotoxins. J. Biol. Chem. 277, 23627–23637.
[12] Villegas, E., Adachi-Akahane, S., Bosmans, F., Tytgat, J., Nakajima, T. and Corzo, G. (2008) Biochemical characterization of cysteine-rich peptides from *Oxyopes* sp. venom that block calcium ion channels. Toxicon 52, 228–236.
[13] Dubovskii, P.V., Vassilevski, A.A., Samsonova, O.V., Egorova, N.S., Kozlov, S.A., Feofanov, A.V., Arseniev, A.S. and Grishin, E.V. (2011) Novel lynx spider toxin shares common molecular architecture with defense peptides from frog skin. FEBS J. 278, 4382–4393.
[14] Kozlov, S., Malyavka, A., McCutchen, B., Lu, A., Schepers, E., Herrmann, R. and Grishin, E. (2005) A novel strategy for the identification of toxinlike structures in spider venom. Proteins 59, 131–140.
[15] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
[16] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
[17] Reese, M.G., Eeckman, F.H., Kulp, D. and Haussler, D. (1997) Improved splice site detection in Genie. J. Comput. Biol. 4, 311–323.
[18] Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
[19] Tamura, K., Nei, M. and Kumar, S. (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Natl. Acad. Sci. U.S.A. 101, 11030–11035.
[20] Nei, M. and Gojobori, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.
[21] Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.
[22] Xu, B. and Yang, Z. (2013) PAMLX: a graphical user interface for PAML. Mol. Biol. Evol. 30, 2723–2724.

[23] Conticello, S.G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z. and Fainzilber, M. (2001) Mechanisms for evolving hypervariability: the case of conopeptides. Mol. Biol. Evol. 18, 120–131.
[24] Kozminsky-Atias, A., Bar-Shalom, A., Mishmar, D. and Zilberberg, N. (2008) Assembling an arsenal, the scorpion way. BMC Evol. Biol. 8, 333.
[25] Garb, J.E. and Hayashi, C.Y. (2013) Molecular evolution of α-latrotoxin, the exceptionally potent vertebrate neurotoxin in black widow spider venom. Mol. Biol. Evol. 30, 999–1014.
[26] Wu, D.-D. and Zhang, Y.-P. (2013) Evolution and function of de novo originated genes. Mol. Phylogenet. Evol. 67, 541–545.
[27] Froy, O., Sagiv, T., Poreh, M., Urbach, D., Zilberberg, N. and Gurevitz, M. (1999) Dynamic diversification from a putative common ancestor of scorpion toxins affecting sodium, potassium, and chloride channels. J. Mol. Evol. 48, 187–196.
[28] Alami, M., Céard, B., Legros, C., Bougis, P.E. and Martin-Eauclaire, M.-F. (2006) Genomic characterisation of the toxin Amm VIII from the scorpion *Androctonus mauretanicus mauretanicus*. Toxicon 47, 531–536.
[29] Krapcho, K.J., Kral, R.M., Vanwagenen, B.C., Eppler, K.G. and Morgan, T.K. (1995) Characterization and cloning of insecticidal peptides from the primitive weaving spider *Diguetia canities*. Insect Biochem. Mol. Biol. 25, 991–1000.