BMC Molecular Biology

Research article

Open Access

The venom gland transcriptome of the Desert Massasauga Rattlesnake (*Sistrurus catenatus edwardsii*): towards an understanding of venom composition among advanced snakes (Superfamily Colubroidea)

Susanta Pahari[1], Stephen P Mackessy[2] and R Manjunatha Kini*[3]

Address: [1]Center for Post Graduate Studies, Sri Bhagawan Mahaveer Jain College, 18/3, 9th Main, Jayanagar 3rd Block, Bangalore, India, [2]School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639-0017, USA and [3]Protein Science Laboratory, Department of Biological Sciences, National University of Singapore, Singapore 117543 and Deparment of Biochemistry, Virginia Commonwealth University, Medical college of Virginia, Richmond, VA 23298-0614, USA

Email: Susanta Pahari - susanta2001@yahoo.com; Stephen P Mackessy - Stephen.Mackessy@unco.edu; R Manjunatha Kini* - dbskinim@nus.edu.sg

* Corresponding author

Abstract

Background:

Results: ( ), *Sistrurus catenatus edwardsii* ( ), 576

Background

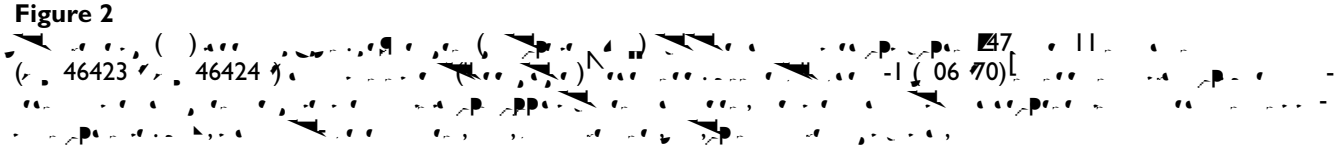
The advanced snakes (superfamily Colubroidea) consist

42 hypothetical ESTs (8.1%). Nine ESTs (1.7%) matched

Identification of toxin families

Serine proteinase

Figure 2



specificity towards specific receptors, e.g., replacement of R with a K in RGD motif of barbourin and ussuristatin 2 significantly increases the selectivity for $_{IIb}{}_{3}$ (fibrinogen receptor) without affecting its binding to $_{5}{}_{1}$ (fibronectin receptor) or $_{v}{}_{3}$ (vitronectin receptor) [62,63]. Additionally, the residues immediately adjacent to the RGD loop also influence both selectivity and affinity for integrin receptors [64,65]. For example, disintegrins with RGDW and recjr4/TT2 1orty for

activation [66-73]. All members of the PIII class of SVMPs have six conserved Cys residues at positions 126, 166, 168, 173, 190 and 206 in their metalloproteinase domain, and some isoforms have a seventh Cys residue at three variable positions (195, 181 or 100) [60,74]. The presence of the seventh Cys residue at position 195 (subgroup PIIIa) results in proteolysis/autolysis, producing a product comprised of the disintegrin-like and cysteine-rich domains (DC domain), whereas when it is present at position 181 (subgroup PIIIb), the formation of a homodimeric structure results [60]). We have not found any isoform having a Cys residue at position 100 (103 in

which matches with Catrin (AAO62995, 87% identity) from *C. atrox* venom. CRISPs are widely distributed in mammals, reptiles, amphibians, arthropods, nematodes, cone snails and plants, and they exhibit diverse biological functions [92]. They are single chain (MW of ~20–30 kDa), highly conserved proteins organized in three domains: a PR-1 (Pathogenesis Related proteins of group 1) domain, a hinge domain and a cysteine-rich domain (CRD). They contain 16 Cys residues forming eight conserved disulfide bonds. A few snake venom CRISPs have been shown to act upon various ion channels through the CRD domain [93-96]. However, the function of the majority of CRISPs from snake venom is unknown [97]. Therefore, it may be interesting to examine the biological properties of the CRISP found in *S. c. edwardsii* venom.

Bradykinin-potentiating peptide and C-type natriuretic peptide

We found a singleton (transcript abundance 0.28%; Additional file 2, Figure 1a) encoding a BPP-CNP (DQ464265) which showed 80% identity with a BPP-CNP precursor from *Lachesis muta* [98]. The BPP-CNP family of proteins lowers the blood pressure of prey during envenomation. Its low abundance in our library indicates that BPP-CNP may not have a significant role in envenomation by *Sistrurus*, unlike bites by other pitvipers (*Bothrops* and *Lachesis*) in Southern America [15,98].

Three-finger toxin like transcripts

We obtained three individual singletons (Additional file 2, Figure 1a) in the library (transcript abundance 0.83%) which belong to the 3FTx family of proteins. As 3FTxs are very uncommon in viperid venoms, using targeted approach we performed RT-PCR using a separate pool of RNA as template and sequenced 96 clones. We found a total of five isoforms of 3FTx-like trancripts (DQ464281,

Bayesian analyses and hence support the above conclusions (Additional file 6).

This family of proteins was not observed in a detailed proteomic characterization of *S. catenatus* and *S. miliarius barbouri* venoms [14]. cDNA libraries of other viper venom glands, including *B. jararacussu*, *B. insularis*, *A. acutus* and *Deinagkistrodon acutus*, do not show their presence [15,16,22,26]. This could be due to either low abundance transcripts and proteins or non-uniform recruitment of 3FTx into the venom proteome within Viperidae. In *S. c. edwardsii*, the low transcript abundance (0.83%) suggests that 3FTx are minor components of the mature venom.

In snake venoms, 3FTXs exhibit diverse pharmacological effects due to their ability to target various receptors and ion channels [103]. It is important to note that the -sheeted loops play crucial roles in binding to various targets, and these regions are the most variable among *S. c. edwardsii* 3FTXs. Further, the d_N/d_S ratio of 0.98.1()6.2(ta[(D)-6.3(e)-)

Novel toxin-like transcript

In our library, we obtained one singleton (Additional file 2, Figure 1a) (DQ464286, transcript abundance 0.28%) with an ORF encoding a signal peptide (24 residues) and a mature protein (128 residues). The putative mature protein is rich in Cys residues, similar to many other snake venom toxins. Its N-terminal domain matches with Kunitz/BPTI toxins (53–68% identity) and the middle domain matches with waprins (45–58% identity), and the novel transcript has an extended C-terminus (Figure 6). Both Kunitz/BPTI [104] and waprins [105,106] are found separately as single domain proteins in snake venoms. Two of the Cys residues, which form one of the four disulfide bonds in waprins, are missing in the new transcript (Figure 6). RT-PCR using a fresh RNA (other than used to make cDNA library) as template and sequencing experiments show the presence of this fused transcript in the venom gland and hence it is not an artifact due to template switching by the Reverse Transcriptase used for making the cDNA library [107-109]. Although a number of cDNA sequences of Kunitz/BPTI from snake venoms have been completed, none of them have the waprin domain and the C-terminal extension. Currently, cDNA sequences of waprins are not known. However, this is the first experimental evidence for the presence of a waprin domain (though fused with another toxin) in viperid venom.

The longer ORF having Kunitz/BPTI and waprin domains together could be due to the fusion of two individual genes encoding Kunitz/BPTI and waprin. Gene fusion mediated by exon shuffling (intron mediated recombination or retrotransposition) has been established as an essential genetic mechanism for the origin of new genes in invertebrates, vertebrates and plants [110,111]. Recently, a new genetic process, transcription-induced chimerism (TIC), in cases of tandemly located gene pairs has been shown to be responsible for gene fusion in the human genome, producing chimeric proteins [112,113]. It is not clear at this stage how this novel fused gene has originated in the snake venom gland. This fused transcript may code either for a precursor which is processed to form two individual classes of venom proteins (Kunitz/BPTI and waprin) or a novel toxin with two distinct domains and having a new biological function. It has been observed that new genes often give rise to new biological functions driven by adaptive Darwinian selection [114-116]. The mechanism of fusion of these apparently independent genes, the evolutionary trajectory of this fused gene and the potential new toxic function of the chimeric protein are all areas for future investigation.

Iron-binding protein

Four ESTs (Additional file 2, Figure 1a) (dbEST: SCEHYPO1, transcript abundance 1.11%) showed homology with an iron-binding protein with a potential signal peptide. Although most iron-binding proteins are generally categorized as storage protein, some of them, such as ovotransferrin and lactoferrin, have antimicrobial activities [117-119]. It is not clear whether or not this protein is found in the venom. However, omwaprin, a member of the waprin protein family, and the C-terminal region of a myotoxic PLA_2 were both shown to have antimicrobial activity [105,120].

Identification of cellular transcripts

We obtained 106 clusters (transcript abundance 21%, 107 sequences) which are involved in various cellular functions, including transcription and translation, secretion, post-translational modification, general metabolism and other functions (Additional file 3, Figure 1b). Similar house-keeping protein products have been observed in other snake venom glands [13,15,22]. One of the ESTs (SCE438) matches (74%) a calcium- and integrin-binding protein which assists platelet spreading [121]. Although modulation of platelet and integrin functions is a key activity of several snake venom components, we do not

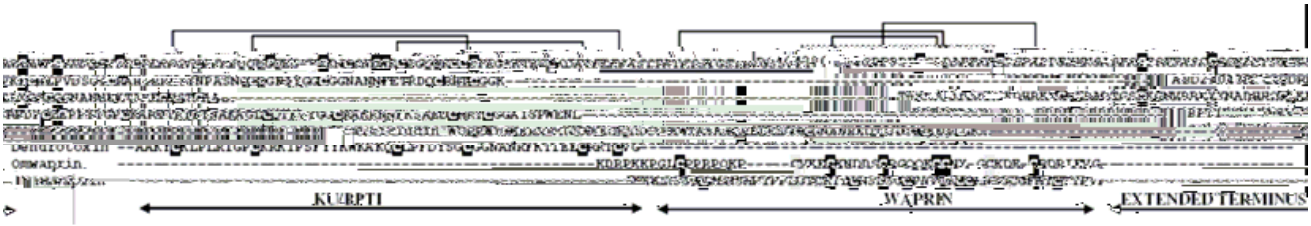


Figure 6

24043 (*Daboia russellii russellii*), 0 (, *Pseudonaja textilis textilis*), 165 (, *Dendroaspis angusticeps*), 00 1 (, *Dendroaspis polylepis polylepis*) 3 52 (, *Oxyuranus microlepidotus*), 605 (, *Naja nigricollis*).

believe that this protein is present in venom, as it lacks the signal peptide.

Overall, results from our cDNA library demonstrate extensive molecular diversity in the venom composition of *S. c. edwardsii*. Serine proteinase and metalloproteinase isoforms are the most abundant components and in the venom, they exert diverse pharmacological activities, particularly disrupting hemostasis. The numerous minor components likely play an ancillary role in envenomation. These diverse toxin isoforms, together with minor components, may be characteristic of venoms from species utilizing different prey types, such as lizards (ectotherms) and birds and mammals (endotherms) [122].

Venom composition and genetics of their origin

Snake venoms consist of a diverse range of pharmacologically active protein and peptide toxins which are primarily used in prey capture and secondarily as defense weapons. To date, the majority of the work on toxin identification and characterization has been concentrated on snakes of the families Elapidae and Viperidae because they are often abundant, produce high yields of venom and represent a significant risk to human health worldwide. Recent studies of venom transcriptomes and proteomes indicate that our knowledge of venom composition is partly limited by experimental detectability. For example, 3FTxs, which were thought to be found exclusively in elapid venoms, were detected in viperid venom gland transcriptomes only recently [[98] and in this study]. Similarly, CLP, thought to be limited to viperid venoms, have been detected recently in the venom gland of *Philodryas olfersii* (Colubridae) and *Bungarus* species (Elapidae) [27,106,123,124]. Further, a new family of low abundance toxin (vespryns) was identified in both elapid and viperid venoms [98,123], Therefore, with the application of advanced techniques like EST sequencing, "compositional specificities" between families of venomous snakes may become less distinct (Additional file 1). Multiple recruitment events may lead to an increase in the spectrum of known and unknown toxin families, decreasing the compositional specificities among venomous snakes. However, differential contribution of specific toxins to the overall expressed proteome of venomous snakes does lead to significant differences in venom composition between species.

A central theme in the evolution of venom systems is complete duplication of toxin genes, followed by accelerated evolution which favors nonsynonymous amino acid substitution towards neofunctionalization. Modification of selected surface areas of toxins [82] is responsible for producing the functional diversity in animal (invertebrates: snails and scorpions; vertebrates: snakes) toxin multigene families [125]. However, one important observation in the present report is the occurrence of a novel toxin-like transcript generated by fusion of two individual toxin genes, Kunitz/BPTI and waprin, in a snake venom gland. Though the mechanism for creation of this fused gene needs to be studied further, it clearly indicates that other genetic processes (gene shuffling or TIC) are also operating in the venom gland to create novel toxin genes. Genes originating by other genetic processes such as exon shuffling are recent [111], and therefore the addition of this fused toxin-like transcript to the venom proteome is perhaps new. At this stage, it is tempting to speculate that the origin of modular organization of different classes of SVMPs, which appears to be the result of gene fusion events, may be due to a genetic process other than gene duplication. SVMPs are very abundant toxins and carry out a principal role in envenomation by viperid snakes, and therefore studies of their genetic origin and organization will be of great interest. Circumstantial evidence of trans-splicing for the generation of serine proteinase isoforms in the venom gland of *V. lebetina* has been presented [126]. Kopelman *et al.*[127] have shown that alternative splicing and gene duplication are inversely correlated evolutionary mechanisms. According to Parra et al. [113], only 4–5% of the tandem gene pairs in the human genome can produce chimeric proteins. It is obvious that these alternative genetic processes responsible for expanding functional proteomes are uncommon among biological systems, and it is therefore not surprising in our case to have just a singleton of the fused transcript out of 576 ESTs (transcript abundance 0.28%). This also demonstrates that to detect rare genetic processes operating in the venom gland, the library generated must be of high quality and that subsequent analyses must be performed very carefully. In turn, these analyses help elucidate in detail the principles of evolution of snake venom transcriptome which have led to the evolutionary success of the advanced snakes [128].

Conclusion

The composition of snake venoms has been shown to be dependent on numerous factors, including phylogeny, diet, age, geography and even sex [129-132]. In general, greater similarity of venoms will be observed along broad phylogenetic lines (e.g., within-family than between-family). However, as this study has demonstrated, some toxins classically considered to occur in only one family, such as the 3FTxs, are actually broadly distributed among the advanced snakes (Colubroidea). The present capacity to detect low abundance toxins indicates a greater compositional similarity of venoms among advanced snakes than has been previously recognized. Further, we have demonstrated that in addition to gene duplication, exon shuffling or transcriptional splicing may also contribute to generating the diversity of toxins and toxin isoforms observed among snake venoms. Overall, the elucidation

of the venom gland transcriptome of *S. c. edwardsii* contributes to a broader picture of toxin expression which complements and extends proteomic analysis of this venom [85]. These approaches can lead to the identification of new toxins and provides mechanistic explanations for their evolution and diversification. An unresolved question involves the relationship between the venom gland transcriptome and how this is ultimately translated to the final proteome. This variable proteomic composition in turn determines the complex and often difficult to resolve sequelae which frequently develop following envenomation by the different species of venomous snakes.

Methods

Venom extraction and collection of venom glands

Specimens of *Sistrurus c. edwardsii* (Desert Massasauga) were collected in Lincoln County, Colorado, USA under permits granted by the Colorado Division of Wildlife to SPM (permits #0456, 06HP456). Venom was extracted from adult snakes using standard manual methods [133]; venoms were then centrifuged to remove particulates, frozen and lyophilized. Prior to gland removal, snakes were extracted of venom. Four days later, when mRNA levels are presumed maximal [134], two snakes were anesthetized with isofluorane and then sacrificed by decapitation. Glands were then rapidly dissected from the snakes, cut into small pieces and placed in approximately 0.5 mL RNAlater (Qiagen) and frozen at -80°C until used.

cDNA library construction and

Acknowledgements

36. [illegible] **Synopsis of recent developments in venomous snake systematics, No. 3.** *Toxicon* [illegible], **37:**1123-112[illegible].
37. [illegible] **Mitochondrial DNA sequences from dried snake venom: a DNA barcoding approach to the identification of venom samples.** *Toxicon* 2005,

7 [illegible] **A novel metalloprotease from Vipera lebetina venom induces human endothelial cell apoptosis.** *Toxicon* 2005, **46:**46-61.

tion and hyperalgesia in mice. *J Biol Chem* 2005,