

Research article

Open Access

Role of accelerated segment switch in exons to alter targeting (ASSET) in the molecular evolution of snake venom proteins

Robin Doley¹, Stephen P Mackessy² and R Manjunatha Kini^{*1}

Address: ¹Protein Science Laboratory, Department of Biological Sciences, National University of Singapore, 117543, Singapore and ²School of Biological Sciences, University of Northern Colorado, Greeley, CO, 80639-0017, USA

Email: Robin Doley - doley@tezu.ernet.in; Stephen P Mackessy - Stephen.Mackessy@unco.edu; R Manjunatha Kini^{*} - dbskinim@nus.edu.sg

^{*} Corresponding author

Published: 30 June 2009

Received: 8 January 2009

BMC Evolutionary Biology 2009, 9:146 doi:10.1186/1471-2148-9-146

Accepted: 30 June 2009

This article is available from: <http://www.biomedcentral.com/1471-2148/9/146>

© 2009 Doley et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Snake venom toxins evolve more rapidly than other proteins through accelerated changes in the protein coding regions. Previously we have shown that accelerated segment switch in exons to alter targeting (ASSET) might play an important role in its functional evolution of viperid three-finger toxins. In this phenomenon, short sequences in exons are radically changed to unrelated sequences and hence affect the folding and functional properties of the toxins.

Results: Here we analyzed other snake venom protein families to elucidate the role of ASSET in their functional evolution. ASSET appears to be involved in the functional evolution of three-finger toxins to a greater extent than in several other venom protein families. ASSET leads to replacement of some of the critical amino acid residues that affect the biological function in three-finger toxins as well as change the conformation of the loop that is involved in binding to specific target sites.

Conclusion: ASSET could lead to novel functions in snake venom proteins. Among snake venom serine proteases, ASSET contributes to changes in three surface segments. One of these segments near the substrate binding region is known to affect substrate specificity, and its exchange may have significant implications for differences in isoform catalytic activity on specific target protein substrates. ASSET therefore plays an important role in functional diversification of snake venom proteins, in addition to accelerated point mutations in the protein coding regions. Accelerated point mutations lead to fine-tuning of target specificity, whereas ASSET leads to large-scale replacement of multiple functionally important residues, resulting in change or gain of functions.

Background

Snake venoms contain a mixture of proteins and polypeptides which exhibit various biochemical and pharmacological functions. These proteins and polypeptides are classified into non-enzymatic and enzymatic proteins which belong to a small number of superfamilies, such as three-finger toxins (3FTx), Kunitz-type serine protease inhibitors, phospholipase A₂(PLA₂) enzymes, serine proteases and metalloproteases [1-12]. Members of these

superfamilies have similar protein scaffolds but, at times, differ markedly in their biological effects. For example, members of 3FTx family exhibit a wide variety of specific pharmacologic effects by targeting various receptors and ion channels with high affinity and specificity. Short chain and long chain α -neurotoxins antagonize muscle nicotinic acetylcholine receptors [5,12], κ -bungarotoxins recognize neuronal nicotinic receptors [13], muscarinic toxins are selective agonists/antagonists of distinct sub-

types of muscarinic acetylcholine receptors [14], fasciculins inhibit acetylcholinesterase [15], calciseptine and related toxins block the L-type Ca^{2+} channels [16,16,17], cardiotoxins/cytotoxins exert their toxicity by forming pores in cell membranes [18], and dendroaspins are antagonists of various cell-adhesion processes [19].

Similarly, other venom proteins, such as the Kunitz-type serine protease inhibitors, have a conserved fold and are structurally similar to bovine pancreatic trypsin inhibitor (BPTI) [20]. They have been reported to inhibit proteolytic activity of trypsin or chymotrypsin specifically [3,21,21-23]. In addition, some inhibitor-like proteins specifically block potassium and calcium channels [24-27]. Snake venom PLA_2 isoenzymes, also characterized by a highly conserved fold, are known to induce various pharmacological activities such as neurotoxic, myotoxic, cardiotoxic, anticoagulant, and antiplatelet effects through specific interaction with their target proteins (for a review see [28]). Thus many subfamilies and isoforms of snake venom serine proteases and metalloproteases act on various components of the coagulation cascade and induce procoagulant or anticoagulant effects, as well as affect platelet aggregation, fibrinolytic and kallikrein-kinin systems [29-35]. The isoforms of the different superfamilies are known to evolve through a process of gene duplication followed by accelerated point mutations in the protein coding regions.

In venom proteins, comparisons of cDNA and gene sequences have shown that nonsynonymous nucleotide substitutions (leading to change in amino acid residues) are commonly greater than synonymous nucleotide substitutions (not producing change in amino acid residues) in the protein coding region compared to the non-coding (UTRs) and intron regions [36-38]. Thus, protein coding regions of genes encoding 3FTxs [39-41], Kunitz-type serine protease inhibitors [42], PLA_2 enzymes [7,43,44] and serine proteases [8] appear to be undergoing accelerated point mutations, resulting in numerous isoforms. Individual point mutations affect one residue at a time, leading to small change in the surface characteristics of a protein. Therefore, point mutations may contribute to fine tuning of toxin specificities by (a) improving the specificity towards a particular receptor or ion channel; (b) altering the specificity towards a closely related receptor or ion channel; and (c) modifying the species specificity. However, accelerated point mutations may not be sufficient to explain drastic changes in the molecular surface needed for the observed targeting of toxins with conserved scaffolds to diverse receptors or ion channels.

In a recent study, we identified five transcripts encoding 3FTxs from the cDNA library of venom gland tissues of *Sistrurus catenatus edwardsii* [45]. These transcripts showed

very low sequence similarity with elapid 3FTxs except for the conserved signal peptide and the number and position of cysteines. A systematic comparison of their sequences revealed that some of the segments in the mature proteins were 80–100% identical, whereas other segments were only 12.5–50% similar [46]. Some segments in the protein coding region seem to be exchanged with distinctly different segments, keeping the structural fold intact during their evolution. Interestingly, the segments in the introns of genes encoding these same proteins show high similarity (>85%) when present; and profound differences in segments appear to be restricted to exons only. Such switching of segments in the exon alters the surface topology and charge of the mature protein, which might alter the molecular targets of 3FTxs and contribute to the evolution of novel function. Therefore, we proposed that the phenomenon of accelerated segment switch in exons to alter targeting (ASSET) might play an important role in the evolution of 3FTxs in viperid snake venoms [46].

Here we have analyzed isoforms of 3FTxs from elapid snake venoms, as well as toxins from other protein superfamilies, to evaluate whether ASSET plays a role in their evolution. Elapid 3FTxs have been found to undergo changes due to ASSET as observed for viperid 3FTxs. Due to such exchange of segments, functionally important residues are changed, which might significantly affect their function. In some of these toxins, such change has led to the evolution of demonstrated novel functions [19,47], and thus the 3FTx toxin family seem to be functionally evolving through ASSET. In the Kunitz-type serine protease inhibitor family, ASSET does not seem to play an important role in evolution, even though there are multiple isoforms. The enzymatic families, such as PLA_2 and metalloproteases, appear to be evolving more through accelerated point mutations rather than ASSET. In these families, some of the segments seem to be exchanged during their evolution, but functional implication of such changes is not clearly understood. However, in the serine protease family, three segments near the substrate binding region have been found to be undergoing accelerated exchange of segments, and at least one of them may have significant implications for their substrate specificity.

Results and Discussion

Three-finger toxin (3FTx) family

3FTxs form a well-characterized superfamily of non-enzymatic proteins. They have a canonical three-finger fold of extending β -sheeted loops that is stabilized by four conserved disulphide bridges in the core region. Until recently, this family of proteins was thought to be present only in elapid venom [48]. However, 3FTxs have now been reported in colubrid venoms and in viperid venom gland transcriptomes as well [45,49-52]. Different iso-

forms of 3FTxs bind to various receptors/acceptors and exhibit diverse pharmacological functions despite their similar folding (for a review see [53]). Functionally important residues that are involved in interacting with the target receptors/ion channels generally reside in the tip of the loops [54]. Therefore the amino acid sequences, length and conformation of the loops play important role in their functional specificity (for reviews see [53,55]).

Amino acid sequences of 3FTxs isolated from venoms of the same genus of elapids were grouped together and analyzed for ASSET and point mutations (Figure 1). The segments with high identity (60–100%) are shown in similar colors, whereas those with low identity (13%–50%) are shown in dissimilar colors. For example, among *Naja* 3FTxs, the S2 segments in P29179, Q9YGI2 and Q9W713 show 60–100% identity (shown in turquoise), and the same segment shows 91.7% identity between O73856 and P01443 (shown in green). However, they share low identity with corresponding segments in other species (shown in various colors). Thus, *Naja* 3FTxs have seven different types of S2 segments. Overall, there are nine (S1–S9) distinct segments (Figure 1) with varied identities, and within the same genus some segments are represented by more than one type. S1 (signal peptide) and S8 segments are the most conserved among all elapid 3FTxs, whereas S5 and S9 are found only in long chain 3FTxs. The replacement of segments has been found to occur only in exons II and III (exon-intron boundaries are marked by red dashed lines), while exon I codes for the highly conserved signal peptide in all 3FTxs. Thus, as in the case of *Sistrurus* 3FTxs, elapid toxins also show discrete replacement of segments in protein coding regions. Such accelerated exchange of segments (ASSET) results in drastic changes in function (discussed below).

The three-finger folds of 3FTxs are held together by four conserved disulfide bridges. However, some 3FTxs have a fifth disulfide bridge in either the second loop (long chain neurotoxins and κ -neurotoxins [56]) or the first loop (non-conventional toxins [57]). The insertion of the fifth disulfide bridge in the second loop of long chain 3FTx is due to a change in the intron-exon boundary. This alteration in the intron-exon boundary is due to an insertion of a single nucleotide "A" in intron 2 which causes a shift in the splicing site [40], leading to the insertion of a short segment (S5) containing a cysteine residue. In the S4 segment, there is also a frame shift due to the deletion of a nucleotide, leading to a completely different sequence which also contains a new cysteine residue. Both cysteine residues form the fifth disulphide bridge and a cyclic structure in the second loop that is important to their binding to $\alpha 7$ receptors with high affinity [58]. The insertion of this short segment in long chain 3FTxs leads to a new function – binding to $\alpha 7$ receptors. In contrast, the

fifth disulphide bridge in the first loop of short chain 3FTxs is due to exchange of segments (ASSET). This additional fifth disulphide bridge does not change the overall fold but it causes subtle changes in the first loop which are known to have functional implications [56]. Further, the number of amino acid residues in this segment differs among the toxins and hence would lead to change in the length of the loop.

As mentioned above, the loops play a crucial role in the recognition of target receptors/ion channels. For example, the third loop in dendroaspin (or mambin; P01413) contains the "RGD" tripeptide sequence (underlined in Figure 1) which is known to bind to platelet glycoprotein (GP IIb-IIIa; $\alpha_{IIb}\beta_{III}$) and cause inhibition of platelet aggregation [19]. The S7 segment containing this sequence is replaced in other *Dendroaspis* toxins (Figure 1). This exchange of segment seems to be responsible for the loss of antiplatelet function in them. In one of the toxins (P25684), this segment is replaced with a segment containing the TAMW residues (underlined in Figure 1). In calciseptine, FS2 and other related toxins, this sequence is known to be involved in binding to L-type Ca^{2+} channels [47]. This segment in dendroaspin and FS2 shows a different conformation (Figure 2) which might also influence their function. As shown in Figure 1, the S7 segment is the most variable among the *Dendroaspis* toxins and may result in functional diversification due to accelerated exchange.

Short chain and long chain α -neurotoxins are known to antagonize muscle nicotinic acetylcholine receptors, resulting in flaccid paralysis [5,12]. The structure-function relationships of α -neurotoxins have been thoroughly studied using both chemical modification and genetic engineering approaches [54,59-61]. Unlike dendroaspin and calciseptine (Figure 1), the functional site in the neurotoxins is discontinuous and is distributed on all three loops [54]. In erabutoxin a (BAC78199), the important functional residues involved in binding to *Torpedo* electroplax or to muscle nAChR ($\alpha_2\beta\gamma\delta$) are Lys²⁷, Trp²⁹ (S3), Asp³¹, Phe³², Arg³³ (S5) and Lys⁴⁷ (S7) (underlined in Figure 1) [56]. Although Lys²⁷, Trp²⁹ and Arg³³ are conserved in all *Laticauda* toxins, the other critical residues (Asp³¹, Phe³² and Lys⁴⁷) are replaced via exchange of segments. We hypothesize that these segment exchanges may have a direct impact on their ability to bind to *Torpedo* or muscle ($\alpha_2\beta\gamma\delta$) receptors.

It is also important to note that there are minor changes in amino acid residues within the identical segments (highlighted in white in Additional file 1) and these changes are due to an accelerated rate of point mutations. Both ASSET and accelerated point mutations have contributed to the functional diversity of elapid 3FTxs; ASSET

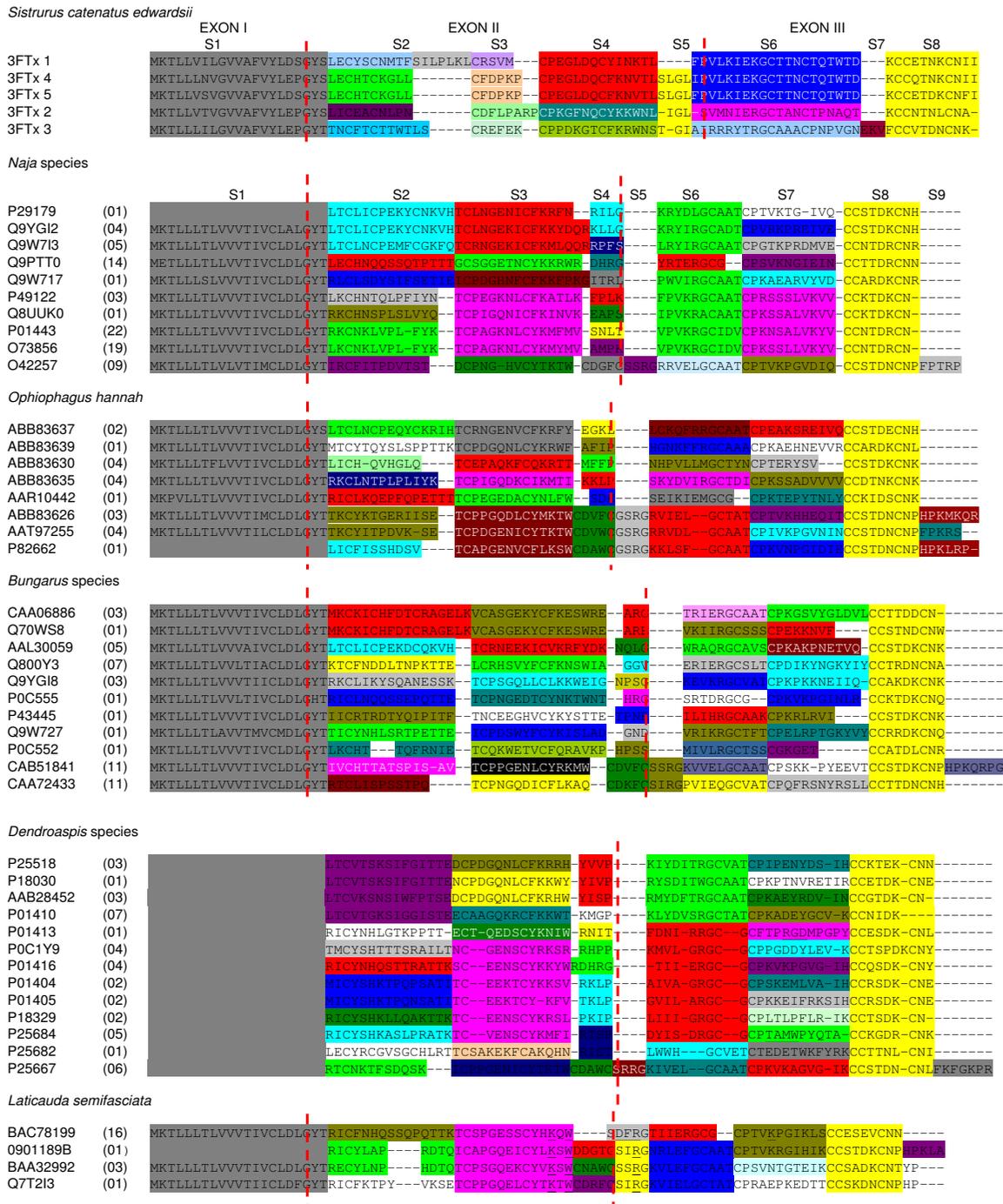


Figure 1
Alignment of three-finger toxins from snake venoms. Protein sequences were obtained from the NCBI database and are presented with their accession numbers. The signal peptide is similar in all the proteins and is shown in grey color; for those protein sequences whose signal peptide is not available, it is kept blank. The number of sequences of closely related homologues is shown in brackets. Based on structural identity, the mature protein is divided into several segments (S1–S9). To emphasize sequence similarities, segments with >50% identity are given similar colors, whereas <50% are given dissimilar colors. The gaps are inserted to optimize the alignments and shown as '-', and some of the functionally important amino acid residues are underlined (see text for details). The intron-exon boundary is marked with vertical red dashed lines.

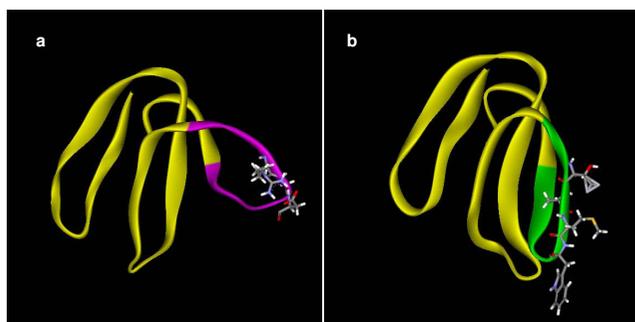


Figure 2
Ribbon model of (a) dendroaspin (PDB ID: 1DRS) and (b) FS2 (PDB ID: 1TFS). In dendroaspin, the segment CFT-PRGDMPGPY is shown in magenta, whereas in FS2, the segment CPTAMWVPYQTA is shown in green. Side chains of RGD and TAMW, the key residues in their functional motifs are shown. This segment exchange has profound effects on activity: dendroaspin is a potent antiplatelet protein interacting with $\alpha_{IIb}\beta_3$, whereas FS2 is a potent blocker of L-type Ca^{2+} channels.

leads to major changes in the surface properties, resulting in targeting of new receptors, while accelerated point mutations lead to fine-tuning of binding to the same receptors through minor alterations of the surface charge and topology.

Kunitz-type serine protease inhibitor family

Snake venom Kunitz-type serine protease inhibitors have been reported from both elapid and viperid venoms. Structurally, they are similar to Kunitz/BPTI inhibitors with a conserved fold stabilized by three disulphide bridges [6]. As with other toxin families, the isoforms are encoded by a multigene family and have evolved through gene duplication and positive selection [8]. The isoforms from the same genus are grouped together as above and analyzed for ASSET and accelerated point mutations (Additional file 2).

Though the snake venom Kunitz-type serine protease inhibitor family contains multiple isoforms, functionally they are not as diverse as other venom protein superfamilies, and they can be divided into either non-neurotoxic or neurotoxic homologs. Non-neurotoxic homologs inhibit either trypsin or chymotrypsin, while neurotoxic homologs act as calcium and potassium channel blockers which do not have protease inhibitory activity [26,27,42]. Structurally, both groups have a conserved fold similar to BPTI, but the inhibitor binding loops and the β turn regions have undergone adaptive evolution, resulting in new biological activities [6]. Analysis of the amino acid sequences of the isoforms shows that there is no radical change in the amino acid residues in the mature proteins,

as observed in 3FTxs. However, they have undergone adaptive evolution through accelerated point mutation (Additional file 2). In calcicludine and dendrotoxin-I, the N-terminal part and overall conformation play a significant role in calcium and potassium channel-blocking activity (Additional file 2). This has been demonstrated by synthesizing chimeras containing the N-terminal (1–30) of calcicludine and C-terminal (31–60) of dendrotoxin-I, and vice versa [24]. However, there are not enough Kunitz-type serine protease inhibitors and dendrotoxin sequences from *Dendroaspis* species in the database in order to determine if they have evolved through ASSET. Similarly, the B chain of β -bungarotoxin (from *Bungarus*) is also a Kunitz-type serine protease inhibitor but does not have protease inhibitory activity; however, it contributes to neurotoxicity [62]. The interaction of the B chain with the potassium channel was predicted to be localized opposite of the anti-protease loop, between residues 27–30 [63]. In addition to this, the mature protein shows accelerated point mutations which resulted in the introduction of a cysteine residue at the C-terminal end (underlined in Additional file 2). This extra cysteine residue forms the disulphide bridge with chain A [63]. Further, the C-terminal region of chain B shows a conformational change due to its interaction with the chain A and accounts for the lack of protease inhibitor activity [63]. Unlike 3FTxs, where ASSET has played an important role in the evolution of new functions, deviation of some members of Kunitz-type serine protease inhibitors from protease inhibitory activity is mainly due to accelerated point mutations. This might explain the low functional diversity in this group of toxins, even though they have multiple isoforms.

Phospholipase A₂ (PLA₂) family

PLA₂ enzymes are one of the best-studied hydrolytic enzymes and are found abundantly in nature. Snake venoms are a good source of these enzymes and often contain multiple isoenzymes. In addition to a role in the digestion of prey, they induce a wide variety of pharmacological effects in prey/victims (for a review see [64]). It has been well documented that accelerated point mutations have occurred in the protein coding regions, and this adaptive mode of evolution might also be responsible for acquisition of new functions [65]. We analyzed the elapid and viperids PLA₂ isoenzymes to determine if ASSET has played any role in the functional evolution of these toxins.

Comparison of the amino acid sequences of PLA₂ isoenzymes in all genera revealed that the N-terminal region seems to be undergoing exchange of segment except in *Naja* species (Figure 3; segments are shown in different colors). This segment is 13–14 amino acids long (forms helix B in several PLA₂ enzymes) and lies between the first helix and the calcium-binding loop. Differences in

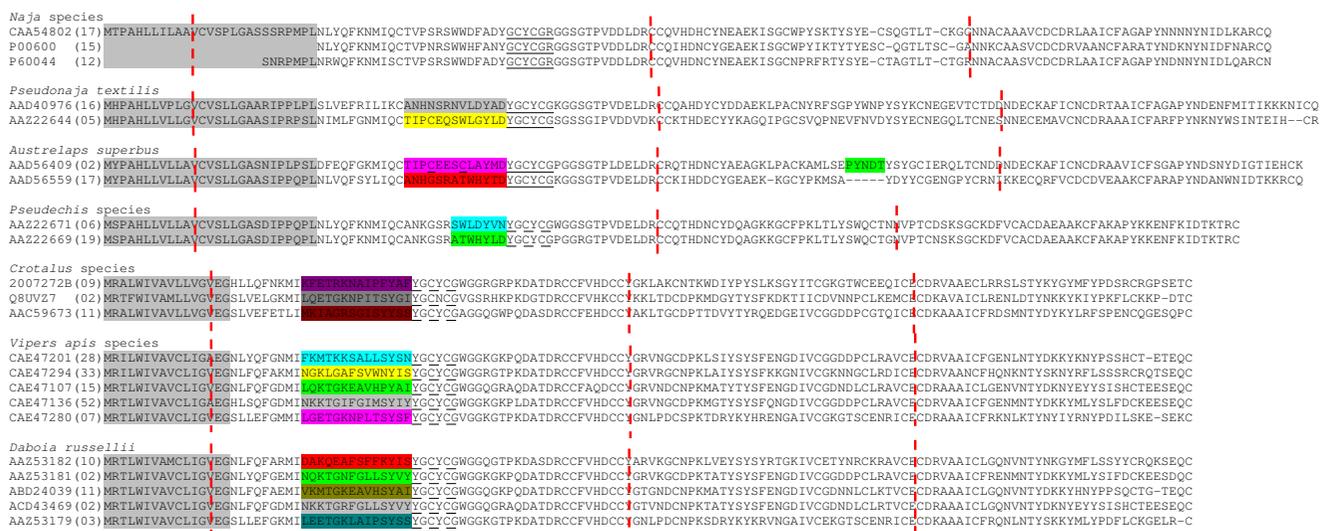


Figure 3
Alignment of phospholipase A₂ isoforms from snake venoms. Protein sequences were obtained from the NCBI database and are presented with their accession numbers. The signal peptide is highlighted in grey color. The number of sequences of closely related homologues is shown in brackets. The calcium binding loop is underlined and gaps are shown with "-". The region undergoing exchange of segment is given different color coding.

exchange within the same species seem to arise due to ASSET, as there is more than one amino acid replacement. As this segment lies near the calcium-binding region, it might affect catalysis due to changes in the conformation in this region. For example, in viperid PLA₂ enzymes, the sequence LEETGKLAIPSYSS (in AAZ53179) is replaced with an unrelated sequence VKMTGKEAVHSYAI (in ABD24039), which clearly imparts significant conformational changes (Figure 4). However, the impact of such replacements on catalysis is not clear. In elapids within the same genera, this segment is represented by two differ-

ent types, but in viperids there are several types. Interestingly, in AAD56409 and the γ subunit of taipoxin, as a result of this replacement, two cysteine residues are inserted. These additional cysteines form the extra disulphide bridge, in addition to the seven conserved disulphide bridges of PLA₂ enzymes [66]. Further, the residues in this short segment have been proposed to play crucial role in some of the pharmacological effects of PLA₂ enzymes [67]. Ammodytoxin A from the venom of *Vipera ammodytes ammodytes* interacts with human FXa through surface residues distributed in helices and the calcium binding loop. One of the helices (helix B), which is undergoing accelerated segment exchange, is involved in this interaction [67]. Replacement of critical amino acid residues through segment exchange can influence the biological activity of PLA₂ enzymes. Other than this segment, the mature protein in all the groups appears to be evolving via accelerated point mutations. Though elapid and viperid PLA₂ enzymes have evolved from different lineages, it is interesting to note that the exchange of the segment in both groups occurs at the same position. Such an exchange may impart conformational and/or functional changes.

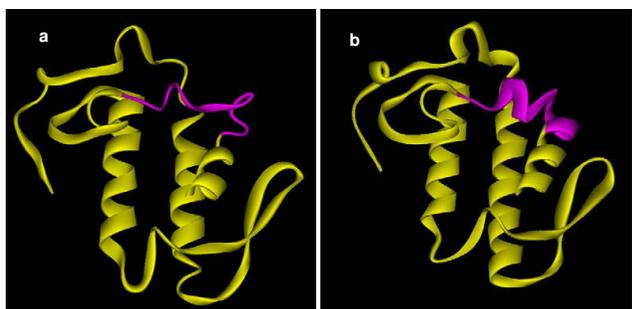


Figure 4
Ribbon model of (a) daboiatoxin (PDB ID: 2H4C) and (b) Vipera PLA₂ (PDB ID: 1Q6V) showing conformational change due to segment exchange. These changes may impart differences in biological function (not yet evaluated).

In addition to this exchange of a segment near the calcium binding region, the presence or absence of another segment has been observed in exon III of *Austrelaps superbus* PLA₂ enzymes (Figure 3, shown in green color). This segment represents the pancreatic loop, an ancestral feature found in the pancreatic PLA₂ enzymes. Pancreatic PLA₂

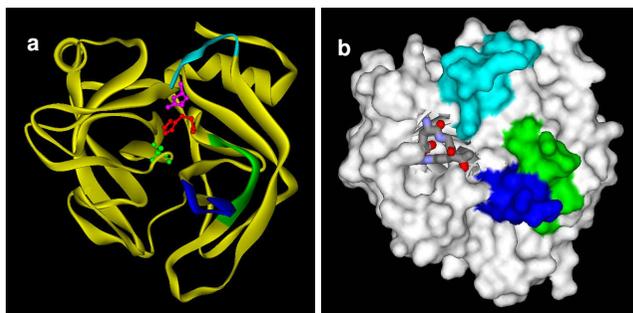


Figure 7
Ribbon (a) and surface (b) models of plasminogen activator (TSV-PA) (PDB ID: 1BQY). The segments that undergo exchange are shown in green, dark blue and turquoise color. The side chains of the active site residues are shown in the ribbon model. The substrate binding area is shown in turquoise color.

natus edwardsii and *Bothrops jararaca* obtained from the database to analyze for ASSET.

Comparison of the isoenzymes among the same group reveals that the signal peptide is highly conserved in all the groups, and the mature protein shows accelerated evolution. Particularly, three segments in the mature proteins seem to be undergoing accelerated exchange. These segments are identified in all the different genera and named as i, ii and iii (Figure 5; shown in different colors). Among them, segment iii is the most conserved segment and is represented by only four different types (shown in green, light blue, red and magenta, in Figure 5) across the genera whereas segments i and ii exist in several different forms. These segments were further analyzed to see if a similar pattern exists across different genera (Figure 6). Although some similar patterns emerge, the data strongly supports accelerated exchange of these segments. Further, the segment exchanges appear to be random and we could not link them through simple molecular phylogeny.

It is interesting to note that all three segment exchanges occur on the surface of the protein (Figure 7) which is important for substrate binding and interactions with ligands. SVSPs are all characterized by the presence of a highly reactive serine residue (Ser) in the active site. In addition, His57 and Asp102 present in the active cleft are important for catalysis (marked with an asterisk in Figure 6); this active cleft is located at the junction of two six-stranded β -barrels [77]. The entrance to this active site cleft is influenced by several surface loops of the protein. Asp97, present in segment iii, is the most important residue in substrate recognition, and a D97N mutant shows markedly decreased substrate binding capacity [78]. This residue (D97 in Q91516 (PSV-PA) is underlined in Figure

5. However, in other isoforms from *Trimeresurus*, this segment PNRKKDDE (shown in turquoise color) is replaced with either NCSNNFK (shown in red color) or LSSNNDIE (shown in green color) with the loss of critical Asp residue. Thus ASSET may have a direct impact on the substrate binding of these isoforms. On the other hand, changes of segment i and ii might not influence the catalytic activity, as they are away from the active cleft and are not involved in substrate binding (Figure 7). However, they might still contribute to targeting to various tissues or proteins as they are fully exposed on the surface of these enzymes. This family of proteins seems to be evolving via both ASSET and accelerated point mutations.

Metalloprotease family

Snake venom metalloproteases (SVMPs) are the main toxic components present in the venom of many Viperidae [2,11]. They are synthesized as zymogens containing multidomain precursors and undergo proteolytic processing. Structurally, they are categorized into four types, PI, PII, PIII and PIV, based on the presence of different domains and their quaternary structure [9,79]; the metalloprotease domain is present in all the subtypes. Catalytically it is the most important domain and contains the zinc binding site with a consensus sequence of HEXX-HXXGXXH [80]. This catalytic domain is known to play an important role in inducing hemorrhage during envenomation and is highly conserved [81]. SVMPs induce hemorrhage at the site of the bite by cleaving proteins in the basement membrane and subsequently weakening capillaries [82]. At present, not many isoforms of metalloproteases have been reported from the same species. Sequences from *Sistrurus catenatus edwardsii*, *Macrovipera lebetina* and *Echis ocellatus* were obtained from the database and analyzed for point mutations and ASSET (Figure 8). Most of the residues in the mature protein seem to be conserved, except for a few residues which have been replaced through accelerated point mutations. However, in the isoforms of *Sistrurus catenatus edwardsii* venom, the cysteine-rich domain seems to be undergoing exchange of segments (shown in red/green colors). Similarly, in the spacer region between the metalloprotease and the disintegrin-like domains of *Macrovipera lebetina* isoforms, 11 amino acid residues have been deleted in one of the isoforms (Figure 8). These two regions were analyzed in the crystal structure of catrocollastatin/vascular apoptosis-inducing protein (VAP) 2B from *Crotalus atrox* venom, which possesses metalloproteinase/disintegrin/cysteine-rich (MDC) domains (Figure 9). The segment represented by VGEECDCGTPE is a part of the shoulder domain containing one of the calcium binding regions, and this calcium binding region is absent in one of the isoforms of *Macrovipera lebetina* metalloprotease. The other segment, highlighted with red color, is the most variable and divergent among the ADAM/adamalysin/reprolysin protein

Sistrurus catenatus edwardsii

DQ464251 MIEVLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKYE
DQ464250 MIEVLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKYE
DQ464249 MIEVLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKYE
DQ464252 NVEKEDEAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
DQ464250 NVEKEDEAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
DQ464249 NVEKEDEAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
DQ464251 NQKLSGTIVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
DQ464250 NQKLSGTIVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
DQ464249 NQKLSGTIVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
DQ464251 ECDIAESCTGQSDAFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
DQ464250 ECDIAESCTGQSDAFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
DQ464249 ECDIAESCTGQSDAFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS

Macrovipera lebetina

AA38182 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
AA38181 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA66471 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
AA38182 YENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
AA38181 YENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA66471 YENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
AA38182 MCDPQYAGIVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
AA38181 MCDPQYAGIVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA66471 MCDPQYAGIVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
AA38182 NGECDVSLCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
AA38181 NGECDVSLCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA66471 NGECDVSLCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS

Echis ocellatus

CA301683 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA301682 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA301680 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA301688 NVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA301684 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA301679 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
BAA06910 MIVQLVLTICLAVFFYQSSII ILESGNVNDYEVVYPRKVTALPKGAVQYEDAMQYEFKNGEVSUHLLEKNKELFSENSETHYSPDGREITITYPADHCHYHGRINQDADSTASISACNGLKGFHLQ-GEMLYIEPLKLPDSEAHAVKY
CA301683 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA301682 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA301680 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA301688 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA301684 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA301679 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
BAA06910 KYENIEKEDAPKMGCVTQNMKSEYIKKFKSHLVVTAEQRLT--KRVLELVIVADHGMFKHNETL---RTWFFQMVNTINEIYIPLIRFALVGLIEIWSNGDKITVQSSAHNTLHFGNWRKRVLLKRRKHDAQLLTIDFDGRTVGLAHVSSMC
CA301683 AGMCSQERSVGVIEDYFVSDVHVAVI IAHMESHNLGIFHDVHCNGCAGNSIMYAKISNPPPMYSDCSQEQYQFLDNYKPDCTLIRPRTDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301682 SSMCQATRSVGVIEDYFVSDVHVAVI IAHMESHNLGIFHDVHCNGCAGNSIMYAKISNPPPMYSDCSQEQYQFLDNYKPDCTLIRPRTDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301680 ATMCHESELVGLVQDMPSEHVAAM IAHMESHNLGIFHDVHCNGCAGNSIMYAKISNPPPMYSDCSQEQYQFLDNYKPDCTLIRPRTDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301688 NSMCKPKYSVGVIEDYFVSDVHVAVI IAHMESHNLGIFHDVHCNGCAGNSIMYAKISNPPPMYSDCSQEQYQFLDNYKPDCTLIRPRTDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301684 GIMCQFQSAVQDHSIINLVMVTAHEHGHNLGMDHDIKYCTCGAPACVMTGNLSEQSYGFSNCSQNYQYFIINPNQCILNEPLDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301679 DMCKSDRSVGLIHDHSTTETFTV IAHMESHNLGIFHDVHCNGCAGNSIMYAKISNPPPMYSDCSQEQYQFLDNYKPDCTLIRPRTDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
BAA06910 YGCKSDRSVGLIHDHSTTETFTV IAHMESHNLGIFHDVHCNGCAGNSIMYAKISNPPPMYSDCSQEQYQFLDNYKPDCTLIRPRTDITVSPVPCGNELLEVGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301683 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301682 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301680 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301688 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301684 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
CA301679 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS
BAA06910 RGRINECDLPEYCTGQSAEFCPTDHLRNGQVCLNNGHYCYNGMCPIMDHQCIYALGADVADPACDFNLKQGDNFYCRREKARIFPCAFQDVCKGRFLCFEYVSGEEDCCGSPNTRCDPCCDAATCKQHSWVECEGCECQKFKGATCEAARS

Figure 8

Alignment of metalloprotease isoforms from snake venoms. Protein sequences were obtained from the NCBI database and are presented with their accession numbers. The signal peptide is highlighted in grey color and the zinc binding site is underlined. The number of sequences of closely related homologues is shown in brackets. Segments undergoing exchange are shown in different colors.

family and is known as hyper-variable region. This hyper-variable region has been predicted to be a potential exosite for target recognition [83]. This segment in one of the isoforms of *Sistrurus catenatus edwardsii* has been replaced with another segment (highlighted in grey color) and hence might recognize a different substrate or lose its substrate binding property.

Molecular mechanism of ASSET

Previously, we discussed the possible molecular mechanisms of ASSET, including splicing variation, recombination, accumulation of point mutations and independent recruitment events [46]. Nevertheless, we believe that none of these explanations are satisfactory. Splicing variations, such as alternative splicing and changes in the splicing site, can lead to insertion/deletion of alternate segments in the mature protein. However, all but one seg-

ment change occur within the exons and not at the intron-exon boundaries. In the long chain 3FTxs only, the insertion of a segment occurs at the intron-exon boundary due to a shift in the splicing site (discussed above). Genetic recombination might also give rise to replacement of segments in the mature protein. However, the segment exchanges observed in the venom proteins are too small, and canonical recombination processes cannot explain exchange of short segments. The possibility of accumulation of point mutations producing the observed change in segments cannot be ruled out unequivocally, as venom proteins have been well-documented to evolve through accelerated point mutations [7,8,37,39-43,84]. In such circumstances, this would have to occur over many generations to attain the observed change in segments, and intermediates would have to be selected via positive selection. Further, the same point mutations would have to

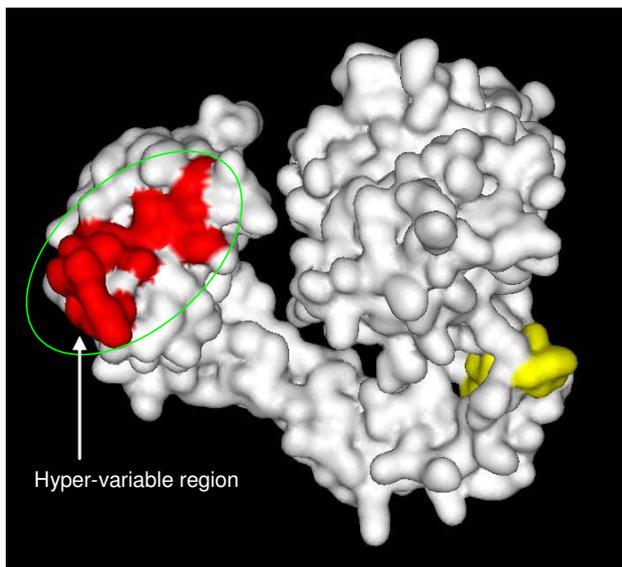


Figure 9
Surface model of catrocollastatin/vascular apoptosis-inducing protein (VAP) 2B (PDB ID: 2DW1) showing exchange of segments. Surface area shown in red is the hyper-variable segment found to be undergoing exchange in other metalloproteases whereas the yellow colored region is the shoulder domain.

occur independently in several unrelated lineages to produce the same segment composition. Thus, odds are against the accumulation of point mutations as an explanation. The 3FTxs showing these exchanges could have occurred through independent recruitment events, but significant similarities in protein sequence and gene structure (particularly high similarity among intron sequences) show that they have evolved from a common ancestor. The observed changes in protein segments could also be due to insertion/deletion of 1–2 nucleotides, resulting in a frame shift and hence altered protein sequence. In such a case, a similar number of nucleotide(s) must be removed/added in a downstream region, respectively, to get back into the original open reading frame. Therefore, we carefully analyzed each of these segment exchanges at the nucleotide level. Only one segment, RKCHNS-PLSLVYQ (S2 in Q8UUK0), is changed to RKC�KLVP-LFYK in P01443 (Figure 1A) due to deletion and insertion of nucleotides. In this case, there is an insertion of A at the 78th position (from the start codon) and a deletion of four nucleotides downstream (Additional file 3). Therefore, none of the above possibilities explain the observed exchange of segments in the exons. It is important to note that in spite of the segment exchanges, the cysteine residues, which maintain the three dimensional fold, are conserved in 3FTxs. Although the molecular mechanism of the exchange of segments is not yet understood, these

events clearly play a significant role in the functional evolution of some snake venom proteins.

ASSET occurs at the molecular surface

Surface residues of a protein molecule are important for their physicochemical properties as well as for their interactions with biomolecules, including other proteins. Accordingly, the alteration of the conformation and surface properties indeed affects the pharmacological properties of protein toxins. In an earlier paper, we showed that in PLA₂ enzymes the surface residues have undergone natural substitution 2.6–3.5 times faster than the buried residues and proposed that accelerated point mutations preferentially target the surface residues in PLA₂ enzymes, leading to the evolution of new isoforms with distinct functions [71]. As shown here, ASSET also targets surface residues in 3FTxs, PLA₂ enzymes, serine proteases and metalloproteases (Figure 2, 4, 7 and 9). Accelerated point mutations result in finer modifications to the surface topology and/or electrostatic potential, whereas ASSET drastically alters the surface, essentially instantaneously producing large-scale changes in the ligand interaction site(s). The molecular mechanisms of both accelerated point mutation and ASSET are not clearly understood, but both phenomena play a crucial role in the evolution of snake venom proteins.

Conclusion

Elapid 3FTxs, similar to viperid 3FTxs (Doley et al., 2008), evolve by both ASSET and accelerated point mutations. ASSET affects the entire mature protein of 3FTxs except for segment S8, which is highly conserved. In serine proteases, three of the surface segments are changed rapidly by ASSET, but the rest of the mature protein evolves only by accelerated point mutations. In PLA₂ enzymes and metalloproteases, only one and three surface segments (respectively) are changed via ASSET. In all these superfamilies of toxins, ASSET most likely affects their functional properties. However, serine protease inhibitors have evolved by only accelerated point mutations. We propose that ASSET occurs first, resulting in drastic changes in functionally important surface regions, followed by accelerated point mutations in those regions which fine-tune the target specificity. Although the molecular mechanisms of ASSET and accelerated point mutations are unknown, both contribute to the evolution of snake venom toxins and both help to explain the observed functional diversity of toxins and the evolution of new functions in snake venom protein superfamilies.

Methods

Sequence analysis and identification of segments

The protein and cDNA sequences were obtained from the NCBI database. Sequence alignments were done using the DNAMAN program and by manual examination. The

intron-exon boundary (marked by a red dashed line in the figures) was identified by comparing the gene and their respective cDNA sequences. In those toxins whose gene structure is not available, the intron-exon boundary was identified by comparing with other toxins whose boundary is known. These segments are identified by comparing with the corresponding sequences, and the point of deletion in amino acid sequence was identified as the boundary of most of the segments. We have analyzed the sequences in different species, but there is no absolute trend in defining the segment, other than high sequence identity. Color coding was used to distinguish segments with distinct % identity; segments with >50% identity are shown in the same color, whereas segments with <50% identity are shown in different colors. Ribbon and surface models were generated from PDB files using DS Viewer-Pro software.

Abbreviations used

ASSET: Accelerated Segment Switch in Exon to alter Targeting.

Authors' contributions

RD carried out the sequence analysis and drafted the manuscript. SPM participated in sequences analyses and writing of the manuscript. RMK designed the study, participated in sequence analyses and writing of the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Three-finger toxins from Naja species showing point mutations. The segments undergoing exchanges are given different color coding. The residues showing point mutations are shown in white and red type whereas those that are conserved are in black. The gaps are represented by "-". This demonstrates accelerated point mutations occur in various segments and the segment exchange is not due to these point mutations. Similarly three-finger toxins from other snake venoms as well as toxins from other families show point mutations through out various segments (data not shown).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-146-S1.pdf>]

Additional file 2

Alignment of Kunitz-type serine protease inhibitor isoforms from snake venoms. Protein sequences were obtained from the NCBI database and are presented with their accession numbers. Numbers of similar sequences are shown in brackets. They appear to evolve by accelerated point mutations in the mature protein (changes in residues are shown in red) but not by segment exchange.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-146-S2.pdf>]

Additional file 3

Segment S2 (Figure 1) of three-finger toxins Q8UUKO and P01443. Figure shows deletion and addition of nucleotides (red color) in the segment.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2148-9-146-S3.pdf>]

Acknowledgements

This work was supported by a grant to R.M. Kini from the Biomedical Research Council, Agency for Science, Technology and Research, Singapore.

References

- Angulo Y, Chaves E, Alape A, Rucavado A, Gutierrez JM, Lomonte B: **Isolation and characterization of a myotoxic phospholipase A2 from the venom of the arboreal snake Bothriechis (Bothrops) schlegelii from Costa Rica.** *Arch Biochem Biophys* 1997, **339**:260-266.
- Bjarnason JB, Fox JW: **Hemorrhagic metalloproteinases from snake venoms.** *Pharmacol Ther* 1994, **62**:325-372.
- Chang L, Chung C, Huang HB, Lin S: **Purification and characterization of a chymotrypsin inhibitor from the venom of Ophiophagus hannah (King Cobra).** *Biochem Biophys Res Commun* 2001, **283**:862-867.
- Chang LS, Wang JJ, Cheng YC, Chou WM: **Genetic organization of Bungarus multicinctus protease inhibitor-like proteins.** *Toxicon* 2008, **51**:1490-1495.
- Changeux JP: **The TiPS lecture. The nicotinic acetylcholine receptor: an allosteric protein prototype of ligand-gated ion channels.** *Trends Pharmacol Sci* 1990, **11**:485-492.
- Chen C, Hsu CH, Su NY, Lin YC, Chiou SH, Wu SH: **Solution structure of a Kunitz-type chymotrypsin inhibitor isolated from the elapid snake Bungarus fasciatus.** *J Biol Chem* 2001, **276**:45079-45087.
- Chijiwa T, Deshimaru M, Nobuhisa I, Nakai M, Ogawa T, Oda N, et al.: **Regional evolution of venom-gland phospholipase A2 isoenzymes of Trimeresurus flavoviridis snakes in the south-western islands of Japan.** *Biochem J* 2000, **347**:491-499.
- Deshimaru M, Ogawa T, Nakashima K, Nobuhisa I, Chijiwa T, Shimohigashi Y, et al.: **Accelerated evolution of crotalinae snake venom gland serine proteases.** *FEBS Lett* 1996, **397**:83-88.
- Fox JW, Serrano SM: **Structural considerations of the snake venom metalloproteinases, key members of the M12 reprolysins family of metalloproteinases.** *Toxicon* 2005, **45**:969-985.
- Gao R, Kini RM, Gopalakrishnakone P: **Purification, properties, and amino acid sequence of a hemoglobinuria-inducing phospholipase A(2), MiPLA-I, from Micropechis ikaheka venom.** *Arch Biochem Biophys* 1999, **369**:181-192.
- Mackessy SP: **Venom composition in rattlesnakes: trends and biological significance.** In *The Biology of Rattlesnakes* Edited by: Hayes WK, Beaman KR, Cardwell MD, Bush SP. California: Loma Linda University Press, Loma Linda; 2008:495-510.
- Tsetlin V: **Snake venom alpha-neurotoxins and other 'three-finger' proteins.** *Eur J Biochem* 1999, **264**:281-286.
- Grant GA, Chiappinelli VA: **kappa-Bungarotoxin: complete amino acid sequence of a neuronal nicotinic receptor probe.** *Biochemistry* 1985, **24**:1532-1537.
- Jerusalinsky D, Harvey AL: **Toxins from mamba venoms: small proteins with selectivities for different subtypes of muscarinic acetylcholine receptors.** *Trends Pharmacol Sci* 1994, **15**:424-430.
- Eastman J, Wilson EJ, Cervenansky C, Rosenberry TL: **Fasciculin 2 binds to the peripheral site on acetylcholinesterase and inhibits substrate hydrolysis by slowing a step involving proton transfer during enzyme acylation.** *J Biol Chem* 1995, **270**:19694-19701.

16. de Weille JR, Schweitz H, Maes P, Tartar A, Lazdunski M: **Calciseptine, a peptide isolated from black mamba venom, is a specific blocker of the L-type calcium channel.** *Proc Natl Acad Sci USA* 1991, **88**:2437-2440.
17. Albrand JP, Blackledge MJ, Pascaud F, Hollecker M, Marion D: **NMR and restrained molecular dynamics study of the three-dimensional solution structure of toxin FS2, a specific blocker of the L-type calcium channel, isolated from black mamba venom.** *Biochemistry* 1995, **34**:5923-5937.
18. Bilwes A, Rees B, Moras D, Menez R, Menez A: **X-ray structure at 1.55 Å of toxin gamma, a cardiotoxin from *Naja nigricollis* venom. Crystal packing reveals a model for insertion into membranes.** *J Mol Biol* 1994, **239**:122-136.
19. McDowell RS, Dennis MS, Louie A, Shuster M, Mulkerrin MG, Lazarus RA: **Mambin, a potent glycoprotein IIb-IIIa antagonist and platelet aggregation inhibitor structurally related to the short neurotoxins.** *Biochemistry* 1992, **31**:4766-4772.
20. Laskowski M Jr: **Protein inhibitors of serine proteinases—mechanism and classification.** *Adv Exp Med Biol* 1986, **199**:1-17.
21. Shafiqat J, Zaidi ZH, Jorvall H: **Purification and characterization of a chymotrypsin Kunitz inhibitor type of polypeptide from the venom of cobra (*Naja naja naja*).** *FEBS Lett* 1990, **275**:6-8.
22. Cheng YC, Yan FJ, Chang LS: **Taiwan cobra chymotrypsin inhibitor: cloning, functional expression and gene organization.** *Biochim Biophys Acta* 2005, **1747**:213-220.
23. Ritonja A, Meloun B, Gubensek F: **The primary structure of *Vipera ammodytes* venom trypsin inhibitor I.** *Biochim Biophys Acta* 1983, **748**:429-435.
24. Nishio H, Katoh E, Yamazaki T, Inui T, Nishiuchi Y, Kimura T: **Structure-activity relationships of calcicludine and dendrotoxin-I, homologous peptides acting on different targets, calcium and potassium channels.** *Biochem Biophys Res Commun* 1999, **262**:319-321.
25. Harvey AL, Karlsson E: **Protease inhibitor homologues from mamba venoms: facilitation of acetylcholine release and interactions with prejunctional blocking toxins.** *Br J Pharmacol* 1982, **77**:153-161.
26. Schweitz H, Heurteaux C, Bois P, Moinier D, Romey G, Lazdunski M: **Calcicludine, a venom peptide of the Kunitz-type protease inhibitor family, is a potent blocker of high-threshold Ca²⁺ channels with a high affinity for L-type channels in cerebellar granule neurons.** *Proc Natl Acad Sci USA* 1994, **91**:878-882.
27. Smith LA, Lafaye PJ, LaPenotiere HF, Spain T, Dolly JO: **Cloning and functional expression of dendrotoxin K from black mamba, a K⁺ channel blocker.** *Biochemistry* 1993, **32**:5692-5697.
28. Kini RM: **Excitement ahead: structure, function and mechanism of snake venom phospholipase A2 enzymes.** *Toxicon* 2003, **42**:827-840.
29. Serrano SM, Maroun RC: **Snake venom serine proteinases: sequence homology vs. substrate specificity, a paradox to be solved.** *Toxicon* 2005, **45**:1115-1132.
30. Shieh TC, Tanaka S, Kihara H, Ohno M, Makisumi S: **Purification and characterization of a coagulant enzyme from *Trimeresurus flavoviridis* venom.** *J Biochem* 1985, **98**:713-721.
31. Serrano SM, Mentele R, Sampaio CA, Fink E: **Purification, characterization, and amino acid sequence of a serine proteinase, PA-BJ, with platelet-aggregating activity from the venom of *Bothrops jararaca*.** *Biochemistry* 1995, **34**:7186-7193.
32. Bjarnason JB, Barish A, Dizenzo GS, Campbell R, Fox JW: **Kallikrein-like enzymes from *Crotalus atrox* venom.** *J Biol Chem* 1983, **258**:12566-12573.
33. Kisiel W, Kondo S, Smith KJ, McMullen BA, Smith LF: **Characterization of a protein C activator from *Agkistrodon contortrix contortrix* venom.** *J Biol Chem* 1987, **262**:12607-12613.
34. Kamiguti AS, Hay CR, Theakston RD, Zuzel M: **Insights into the mechanism of haemorrhage caused by snake venom metalloproteinases.** *Toxicon* 1996, **34**:627-642.
35. Kamiguti AS: **Platelets as targets of snake venom metalloproteinases.** *Toxicon* 2005, **45**:1041-1049.
36. Chuman Y, Nobuhisa I, Ogawa T, Deshimaru M, Chijiwa T, Tan NH, et al.: **Regional and accelerated molecular evolution in group I snake venom gland phospholipase A2 isozymes.** *Toxicon* 2000, **38**:449-462.
37. Nakashima K, Ogawa T, Oda N, Hattori M, Sakaki Y, Kihara H, et al.: **Accelerated evolution of *Trimeresurus flavoviridis* venom gland phospholipase A2 isozymes.** *Proc Natl Acad Sci USA* 1993, **90**:5964-5968.
38. Ogawa T, Oda N, Nakashima K, Sakaki H, Hattori M, Sakaki Y, et al.: **Unusually high conservation of untranslated sequences in cDNAs for *Trimeresurus flavoviridis* phospholipase A2 isozymes.** *Proc Natl Acad Sci USA* 1992, **89**:8557-8561.
39. Chang LS, Huang HB, Lin SR: **The multiplicity of cardiotoxins from *Naja naja atra* (Taiwan cobra) venom.** *Toxicon* 2000, **38**:1065-1076.
40. Fujimi TJ, Nakajyo T, Nishimura E, Ogura E, Tsuchiya T, Tamiya T: **Molecular evolution and diversification of snake toxin genes, revealed by analysis of intron sequences.** *Gene* 2003, **313**:111-118.
41. Gong N, Armugam A, Jeyaseelan K: **Molecular cloning, characterization and evolution of the gene encoding a new group of short-chain alpha-neurotoxins in an Australian elapid, *Pseudonaja textilis*.** *FEBS Lett* 2000, **473**:303-310.
42. Zupunski V, Kordis D, Gubensek F: **Adaptive evolution in the snake venom Kunitz/BPTI protein family.** *FEBS Lett* 2003, **547**:131-136.
43. Nakashima K, Nobuhisa I, Deshimaru M, Nakai M, Ogawa T, Shimohigashi Y, et al.: **Accelerated evolution in the protein-coding regions is universal in crotalinae snake venom gland phospholipase A2 isozyme genes.** *Proc Natl Acad Sci USA* 1995, **92**:5605-5609.
44. Ogawa T, Nakashima K, Oda N, Shimohigashi Y, Ohno M, Hattori S, et al.: ***Trimeresurus flavoviridis* venom gland phospholipase A2 isozymes genes have evolved via accelerated substitutions.** *J Mol Recognit* 1995, **8**:40-46.
45. Pahari S, Mackessy SP, Kini RM: **The venom gland transcriptome of the Desert Massasauga Rattlesnake (*Sistrurus catenatus edwardsii*): towards an understanding of venom composition among advanced snakes (Superfamily Colubroidea).** *BMC Mol Biol* 2007, **8**:115.
46. Doley R, Pahari S, Mackessy SP, Kini RM: **Accelerated exchange of exon segments in Viperid three-finger toxin genes (*Sistrurus catenatus edwardsii*; Desert Massasauga).** *BMC Evol Biol* 2008, **8**:196.
47. Kini RM, Caldwell RA, Wu QY, Baumgarten CM, Feher JJ, Evans HJ: **Flanking proline residues identify the L-type Ca²⁺ channel binding site of calciseptine and FS2.** *Biochemistry* 1998, **37**:9058-9063.
48. Fry BG, Wuster W, Kini RM, Brusic V, Khan A, Venkataraman D, et al.: **Molecular evolution and phylogeny of elapid snake venom three-finger toxins.** *J Mol Evol* 2003, **57**:110-129.
49. Fry BG, Lumsden NG, Wuster W, Wickramaratna JC, Hodgson WC, Kini RM: **Isolation of a neurotoxin (alpha-colubritoxin) from a nonvenomous colubrid: evidence for early origin of venom in snakes.** *J Mol Evol* 2003, **57**:446-452.
50. Junqueira-de-Azevedo IL, Ching AT, Carvalho E, Faria F, Nishiyama MY Jr, Ho PL, et al.: **Lachesis muta (Viperidae) cDNAs reveal diverging pit viper molecules and scaffolds typical of cobra (Elapidae) venoms: implications for snake toxin repertoire evolution.** *Genetics* 2006, **173**:877-889.
51. Pawlak J, Mackessy SP, Fry BG, Bhatia M, Mourier G, Fruchart-Gaillard C, et al.: **Denmotoxin, a three-finger toxin from the colubrid snake *Boiga dendrophila* (Mangrove Catsnake) with bird-specific activity.** *J Biol Chem* 2006, **281**:29030-29041.
52. Pawlak J, Mackessy SP, Sixberry NM, Stura EA, Le Du MH, Menez R, et al.: **Irditoxin, a novel covalently linked heterodimeric three-finger toxin with high taxon-specific neurotoxicity.** *FASEB J* 2009, **23**:534-545.
53. Kini RM: **Molecular moulds with multiple missions: functional sites in three-finger toxins.** *Clin Exp Pharmacol Physiol* 2002, **29**:815-822.
54. Tremeau O, Lemaire C, Drevet P, Pinkasfeld S, Ducancel F, Boulain JC, et al.: **Genetic engineering of snake toxins. The functional site of Erabutoxin a, as delineated by site-directed mutagenesis, includes variant residues.** *J Biol Chem* 1995, **270**:9362-9369.
55. Ohno M, Menez R, Ogawa T, Danse JM, Shimohigashi Y, Fromen C, et al.: **Molecular evolution of snake toxins: is the functional diversity of snake toxins associated with a mechanism of accelerated evolution?** *Prog Nucleic Acid Res Mol Biol* 1998, **59**:307-364.
56. Servent D, Menez A: **Snake neurotoxins that interact with nicotinic acetylcholine receptors.** In *Handbook of neurotoxicology*

- Edited by: Massaro EJ. Totowa, New Jersey: Humana Press; 2001:385-425.
57. Nirthanan S, Gopalakrishnakone P, Gwee MC, Khoo HE, Kini RM: **Non-conventional toxins from Elapid venoms.** *Toxicon* 2003, **41**:397-407.
 58. Servent D, Winckler-Dietrich V, Hu HY, Kessler P, Drevet P, Bertrand D, et al.: **Only snake curaremimetic toxins with a fifth disulfide bond have high affinity for the neuronal alpha7 nicotinic receptor.** *J Biol Chem* 1997, **272**:24279-24286.
 59. Antil-Delbeke S, Gaillard C, Tamiya T, Corringer PJ, Changeux JP, Servent D, et al.: **Molecular determinants by which a long chain toxin from snake venom interacts with the neuronal alpha 7-nicotinic acetylcholine receptor.** *J Biol Chem* 2000, **275**:29594-29601.
 60. Antil S, Servent D, Menez A: **Variability among the sites by which curaremimetic toxins bind to torpedo acetylcholine receptor, as revealed by identification of the functional residues of alpha-cobratoxin.** *J Biol Chem* 1999, **274**:34851-34858.
 61. Pillet L, Tremeau O, Ducancel F, Drevet P, Zinn-Justin S, Pinkasfeld S, et al.: **Genetic engineering of snake toxins. Role of invariant residues in the structural and functional properties of a curaremimetic toxin, as probed by site-directed mutagenesis.** *J Biol Chem* 1993, **268**:909-916.
 62. Kondo K, Narita K, Lee CY: **Amino acid sequences of the two polypeptide chains in beta I-bungarotoxin from the venom of Bungarus multicinctus.** *J Biochem (Tokyo)* 1978, **83**:101-115.
 63. Kwong PD, McDonald NQ, Sigler PB, Hendrickson WA: **Structure of beta 2-bungarotoxin: potassium channel binding by Kunitz modules and targeted phospholipase action.** *Structure* 1995, **3**:1109-1119.
 64. Kini RM: **Phospholipase A2 A Complex Multifunctional Protein Puzzle.** In *Venom Phospholipase A2 Enzymes: Structure, Function and Mechanism* Edited by: Kini RM. Chichester, England: John Wiley & Sons; 1997:1-28.
 65. Ogawa T, Chijiwa T, Oda-Ueda N, Ohno M: **Molecular diversity and accelerated evolution of C-type lectin-like proteins from snake venom.** *Toxicon* 2005, **45**:1-14.
 66. Fohlman J, Eaker D, Karlsson E, Thesleff S: **Taipoxin, an extremely potent presynaptic neurotoxin from the venom of the Australian snake taipan (Oxyuranus s. scutellatus). Isolation, characterization, quaternary structure and pharmacological properties.** *Eur J Biochem* 1976, **68**:457-469.
 67. Faure G, Gowda VT, Maroun RC: **Characterization of a human coagulation factor Xa-binding site on Viperidae snake venom phospholipases A2 by affinity binding studies and molecular bioinformatics.** *BMC Struct Biol* 2007, **7**:82.
 68. Kuipers OP, Thunnissen MM, de Geus P, Dijkstra BW, Drenth J, Verheij HM, et al.: **Enhanced activity and altered specificity of phospholipase A2 by deletion of a surface loop.** *Science* 1989, **244**:82-85.
 69. Nobuhisa I, Nakashima K, Deshimaru M, Ogawa T, Shimohigashi Y, Fukumaki Y, et al.: **Accelerated evolution of Trimeresurus okinavensis venom gland phospholipase A2 isozyme-encoding genes.** *Gene* 1996, **172**:267-272.
 70. Ogawa T, Nakashima K, Nobuhisa I, Deshimaru M, Shimohigashi Y, Fukumaki Y, et al.: **Accelerated evolution of snake venom phospholipase A2 isozymes for acquisition of diverse physiological functions.** *Toxicon* 1996, **34**:1229-1236.
 71. Kini RM, Chan YM: **Accelerated evolution and molecular surface of venom phospholipase A2 enzymes.** *J Mol Evol* 1999, **48**:125-132.
 72. Markland FS, Kettner C, Schiffman S, Shaw E, Bajwa SS, Reddy KN, et al.: **Kallikrein-like activity of crotalase, a snake venom enzyme that clots fibrinogen.** *Proc Natl Acad Sci USA* 1982, **79**:1688-1692.
 73. Perez AV, Rucavado A, Sanz L, Calvete JJ, Gutierrez JM: **Isolation and characterization of a serine proteinase with thrombin-like activity from the venom of the snake Bothrops asper.** *Braz J Med Biol Res* 2008, **41**:12-17.
 74. Samel M, Subbi J, Siigur J, Siigur E: **Biochemical characterization of fibrinogenolytic serine proteinases from Vipera lebetina snake venom.** *Toxicon* 2002, **40**:51-54.
 75. Wang YM, Wang SR, Tsai IH: **Serine protease isoforms of Deinagkistrodon acutus venom: cloning, sequencing and phylogenetic analysis.** *Biochem J* 2001, **354**:161-168.
 76. Itoh N, Tanaka N, Funakoshi I, Kawasaki T, Mihashi S, Yamashina I: **Organization of the gene for batroxobin, a thrombin-like snake venom enzyme. Homology with the trypsin/kallikrein gene family.** *J Biol Chem* 1988, **263**:7628-7631.
 77. Parry MA, Jacob U, Huber R, Wisner A, Bon C, Bode W: **The crystal structure of the novel snake venom plasminogen activator TSV-PA: a prototype structure for snake venom serine proteinases.** *Structure* 1998, **6**:1195-1206.
 78. Zhang Y, Wisner A, Maroun RC, Choumet V, Xiong Y, Bon C: **Trimeresurus stejnegeri snake venom plasminogen activator. Site-directed mutagenesis and molecular modeling.** *J Biol Chem* 1997, **272**:20531-20537.
 79. Fox JW, Serrano SM: **Insights into and speculations about snake venom metalloproteinase (SVMP) synthesis, folding and disulfide bond formation and their contribution to venom complexity.** *FEBS J* 2008, **275**:3016-3030.
 80. Bode W, Gomis-Ruth FX, Stockler W: **Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'.** *FEBS Lett* 1993, **331**:134-140.
 81. Tanjoni I, Butera D, Bento L, Della-Casa MS, Marques-Porto R, Takehara HA, et al.: **Snake venom metalloproteinases: structure/function relationships studies using monoclonal antibodies.** *Toxicon* 2003, **42**:801-808.
 82. Gutierrez JM, Rucavado A: **Snake venom metalloproteinases: their role in the pathogenesis of local tissue damage.** *Biochimie* 2000, **82**:841-850.
 83. Takeda S, Igarashi T, Mori H, Araki S: **Crystal structures of VAPI reveal ADAMs' MDC domain architecture and its unique C-shaped scaffold.** *EMBO J* 2006, **25**:2388-2396.
 84. Ohno M, Chijiwa T, Oda-Ueda N, Ogawa T, Hattori S: **Molecular evolution of myotoxic phospholipases A2 from snake venom.** *Toxicon* 2003, **42**:841-854.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

