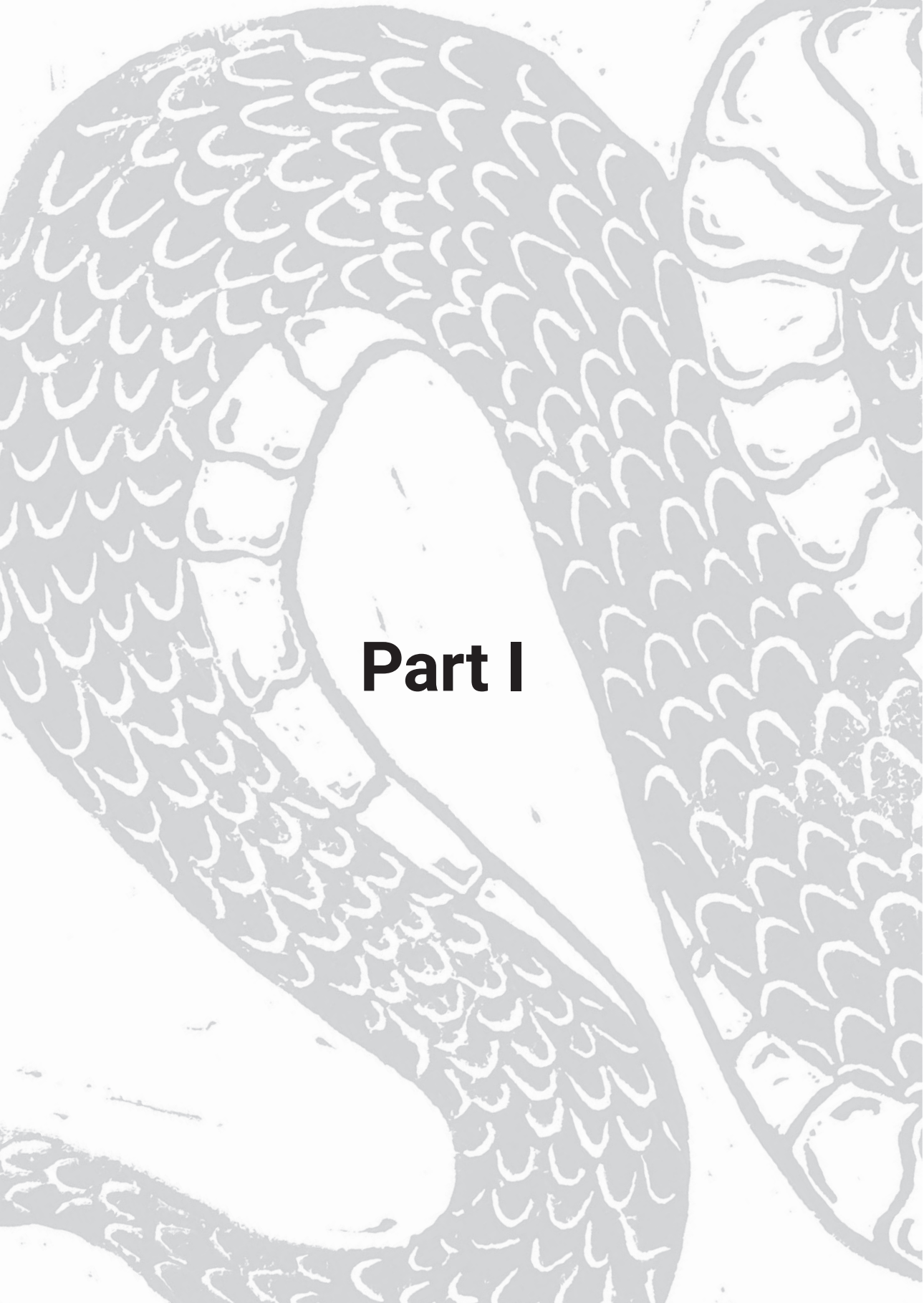
Lessons from snake venom: new insights into the structural and functional aspects of factor V and factor X

Verhoef, D.

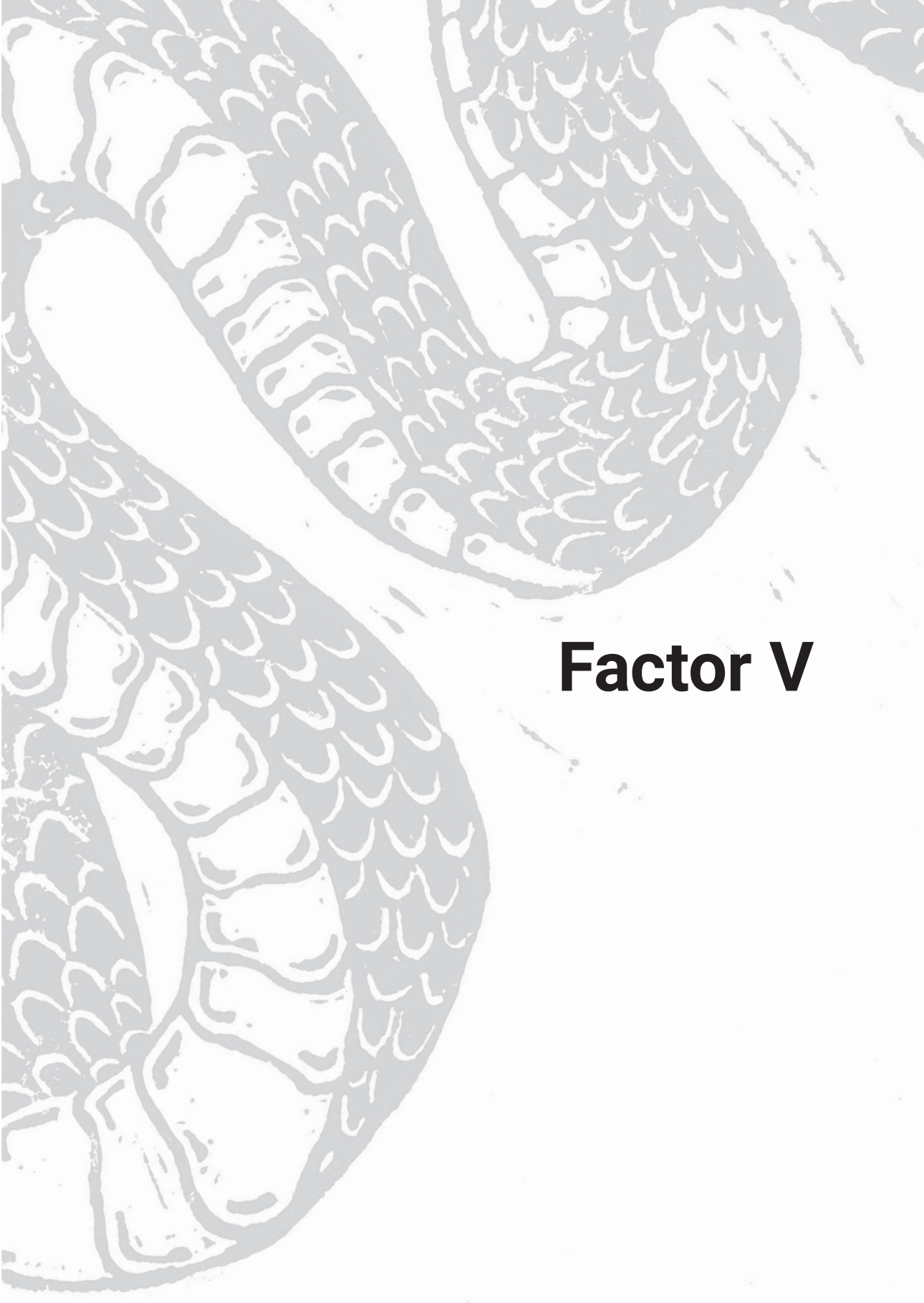
Citation
Verhoef, D. (2021, September 22). *Lessons from snake venom: new insights into the structural and functional aspects of factor V and factor X*. Retrieved from https://hdl.handle.net/1887/3213580

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3213580

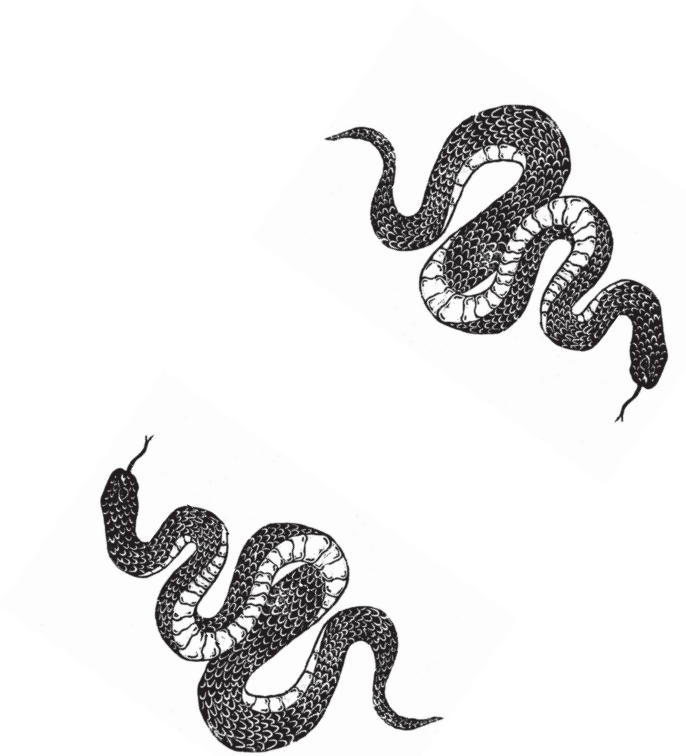
Note: To cite this publication please use the final published version (if applicable).



Part I



Factor V



Chapter 2

Evolutionary conservation of regulatory B-domain elements in blood coagulation factor V in snakes

Daniël Verhoef, Ka Lei Cheung, Hans L. Vos, Ryan J.R. McCleary, R. Manjunatha Kini, Pieter H. Reitsma, Mettine H.A. Bos.

Manuscript in preparation

Abstract

The B-domain of blood coagulation factor V (FV) is an essential element that negatively regulates FV activity through conserved auto-inhibitory regions. Interestingly, a FV sequence lacking the B-domain has previously been identified in an elapid snake, suggesting a divergent molecular mechanism of FV regulation. To investigate whether this unique feature hallmarks the entire Serpentes suborder, we uncovered the identity of the FV B-domain from twelve distinct snake species. To do so, cDNAs were generated from fresh-frozen liver tissue, and B-domain fragments were amplified employing oligonucleotides complementary to conserved FV-specific sequences. Sequence analysis revealed that auto-inhibitory B-domain regions were preserved with high sequence identity in all snake species. In addition, amplification reactions also generated a truncated B-domain fragment that corresponded to a shorter B-domain that lacked the auto-inhibitory regions. Alignment of the compiled snake FV transcripts uncovered two conserved pre-RNA splice sites within the B-domain-encoding exon that may be at the basis of the truncated B-domain transcript. Additional quantitative analysis by qPCR and capillary electrophoresis showed that while in most snakes the regular B-domain is more abundant (20-1000-fold) than the truncated version, the latter are almost equally expressed in snake venom FV-comprising elapids and highly abundant in viperid species. Taken together this indicates that, while the B-domain has been preserved as a potential regulator of FV activity in all members of the suborder Serpentes, alternative splicing of the B-domain may play a role as a disparate mechanism of FV regulation in some snake species.

Introduction

Blood coagulation factor V (FV) is an essential and highly conserved protein component of the vertebrate blood coagulation system [1, 2]. FV normally circulates in blood as an inactive procofactor (domain organization A1-A2-B-A3-C1-C2) and is only converted to its active cofactor form (FVa) in the event of vascular damage. Conversion of FV to FVa proceeds upon limited thrombin- or factor Xa (FXa)-mediated proteolysis of three peptide bonds (Arg^{709}, Arg^{1018}, and Arg^{1545}) that either flank or are within the B-domain sequence, resulting in removal of the central auto-inhibitory B-domain and expression of cofactor activity [3-6]. The auto-inhibitory B-domain normally maintains the procofactor state through a dedicated procofactor regulatory region (PRR), which is generally well conserved among vertebrate B-domain sequences unlike the remainder of the B-domain sequence. The PRR typically constitutes a basic region (BR) and an acidic region (AR) that cooperatively stabilize the procofactor state [7, 8]. Removal of the PRR is required for the expression of FVa cofactor activity. Once removed, FVa is able to assemble into the prothrombinase complex by binding to FXa on a negatively charged phospholipid surface in the presence of calcium ions [7, 8]. This is an essential step during hemostasis, as FVa effectively enhances the conversion of prothrombin into thrombin to physiologically relevant catalytic rates [9, 10].

The venom of the Australian common brown snake *Pseudonaja textilis* contains a powerful prothrombin activator (pseutarin C) that is structurally and functionally similar to the human FVa-FXa complex [11, 12]. The non-enzymatic subunit of pseutarin C (venom FV) is a unique FVa cofactor analogue that is constitutively active upon translation/secretion due to the absence of a regulatory B-domain sequence within its mRNA [13, 14]. Interestingly, a highly similar FV transcript comprising a truncated B-domain lacking the PRR has also been identified in liver tissue of *P. textilis* [15], suggesting a divergent molecular mechanism of FV regulation in hemostasis. On the other hand, whole genome sequencing of the venomous king cobra, *Ophiophagus hannah*, revealed the presence of a single full-length FV gene that included all putative auto-inhibitory B-domain elements [7, 8, 16]. It is therefore uncertain whether the absence of a regulatory B-domain sequence in liver FV is unique to *P. textilis* physiology or whether this trait is shared among other snakes. We have therefore aimed to uncover the identity of the FV B-domain in liver-derived transcripts from snakes belonging to various lineages and geographic regions, as well as venomous vs non-venomous species. By doing so, we have revealed that all investigated snakes include both full-length and truncated FV B-domain transcripts in their liver transcriptome. These findings suggest functional diversification of the FV gene in the suborder Serpentes.

Material and Methods

Materials and Reagents

Fordonia leucobalia tissue was collected locally through permits in Singapore by Ryan McCleary and Vic Toh under Singapore National Parks permit NP/RP10-095. Liver tissues from *Bungarus candidus*, *Oligodon octolineatus*, and *Sistrurus catenatus* were kindly gifted by drs. F.J. Vonk (Naturalis Biodiversity Center/ Amsterdam Institute of Molecular and Life Sciences, The Netherlands), M.A. Reza Tuhin (University of Rajshahi, Bangladesh), and S.J. Mackessy (University of Northern Colorado, USA). All other snake liver tissues were purchased from Venom Supplies Pty Ltd (Tanunda, Australia). RNA-Bee was obtained from Tel-Test Inc (Friendswood, TX, USA). Phusion high-fidelity DNA polymerase was obtained from New England Biolabs (Ipswich, MA, USA). HOT FIREPol DNA polymerase was obtained from Solis BioDyne (Tartu, Estonia). The Wizard SV gel and PCR clean up system was obtained from Promega (Madison, WI, USA). RNAse-free DNAse and RNeasy DNA removal columns were obtained from Qiagen (Hilden, Germany). Oligo-dT primers and Superscript II reverse transcriptase was obtained from Invitrogen (Thermo Scientific, Waltham, MA, USA). SYBR select qPCR master mix was from Applied Biosystems (Foster City, CA, USA). The Agilent DNA 1000 kit was from Agilent Technologies (Santa Clara, CA, USA).

Liver RNA isolation and cDNA synthesis

Liver samples (± 20 – 40 mg) from single individuals were stored at -80°C in RNAlater solution. Upon thawing, the samples were homogenized in RNA-Bee, and total RNA was isolated from homogenates using guanidium-phenol-chloroform extraction and isopropanol precipitation, after which RNA was dissolved in Tris-EDTA-RNAse inhibitor buffer. Spectrophotometric quantification and quality analysis of the sample was performed, and the integrity of the RNA was confirmed by agarose gel electrophoresis. Following genomic DNA removal, first-strand cDNA synthesis was performed using Superscript II reverse transcriptase according the manufacturer's guidelines.

cDNA amplification and Sanger sequencing

The cDNA templates were amplified using Phusion high-fidelity DNA polymerase. Primer annealing temperatures were separately optimized by Touchdown PCR for each cDNA template from individual snake species. Individual cDNA amplicons were isolated by agarose gel electrophoresis and purified using the Wizard SV gel and PCR clean up kit. Sanger sequencing was performed at BaseClear B.V. (Leiden, the Netherlands). Primer design and DNA/Protein sequence analysis was performed employing the Vector NTI 11.5 Advance software suite (Thermo Scientific, Waltham, MA, USA). An overview of all PCR and Sanger sequencing primers is provided (**Table 1**).

Table 1. Overview of primer sequences

cDNA amplification		
A2-F1	forward	5′ CTACTGTAACAATGGACAATC 3′
A3-R1	reverse	5′ GTAGTCCCAGAAAACTTCTTC 3′
Sanger sequencing		
Sanger-F1	forward	5′ CAGGAAACAGCTATGACCCATCTTGCTGGTCACAC 3′
Sanger-F2	forward	5′ CAGGAAACAGCTATGACCCTACTGTAACAATGGACAATC 3′
Sanger-F3	forward	5′ CAGGAAACAGCTATGACCCTAATGATAGCTTCCATGCTTGGGC 3′
Sanger-F4	forward	5′ CTAATGATAGCTTCCATGCTTGGGC 3′
Sanger-F5	forward	5′ CTACTGTAACAATGGACAATC 3′
Sanger-F6	forward	5′ GCTTCAATGCTTGGGCTTCG 3′
Sanger-F7	forward	5′ TTTTAATTTCAAATATCACTGC 3′
Sanger-F8	forward	5′ CAGAACAGCAGCAGAAATGCCTC 3′
Sanger-F9	forward	5′ CCAACCACACAAAATTG 3′
Sanger-F10	forward	5′ CCAGAAAGATGAGAACTGCACTAT 3′
Sanger-F11	forward	5′ GATAATCCATATACCACAAG 3′
Sanger-R1	reverse	5′ TGTAAAACGACGGCCAGTGTAGTCCCAGAAAACTTCTTC 3′
Sanger-R2	reverse	5′ TGTAAAACGACGGCCAGTCTCCTCATCAAAGACC 3′
Sanger-R3	reverse	5′ TGTAAAACGACGGCCAGTCCTGATAGTCACCATCTTC 3′
Sanger-R4	reverse	5′ CCTGATAGTCACCATCTTC 3′
Sanger-R5	reverse	5′ GTAGTCCCAGAAAACTTCTTC 3′
Sanger-R6	reverse	5′ GCTATGTCATCAGGATTACGTGCAC 3′
Sanger-R7	reverse	5′ TGTAAAACGACGGCCAGTGCAGTGATATTTGAAATTAAAA 3′
Sanger-R8	reverse	5′ GAGGCATTTCTGCTGCTGTTCTG 3′
Sanger-R9	reverse	5′ ATAGTGCAGTTCTCATCTTTCTGG 3′
RT-qPCR		
qPCR-F1	forward	5′ GGGTCAGTTGCTGAAGAAGAA′3
qPCR-F2	forward	5′ GGATCAGTTGCTGAAGAAGAA′3
qPCR-R1	reverse	5′ ACTGTCAATTCGAGGATCAGAAG 3′
qPCR-R2	reverse	5′ ACTGTCAAGTCGAGAATCAGAAG 3′
Multiplex-PCR		
A2-F2	forward	5′ GGGTCAGTTGCTGA 3′
AR*-R1	reverse	5′ GTCAAGTCGAGAATCAGA 3′
AR*-R2	reverse	5′ GTCAATTCGAGGATCAGA 3′
BR-R1	reverse	5′ TCCCAATTCTTCATTACC 3′

Real-time quantitative PCR (RT-qPCR)

Snake species-specific RT-qPCR primers were designed with the PRIMER EXPRESS software (Applied Biosystems, Foster City, CA, USA) (**Table 1**). RT-qPCR was performed using the SYBR select qPCR master mix and primer pair qPCR-F1/-R1 for *P. textilis*, *Oxyuranus scutellatus*, *Oligodon octolineatus*, *Psammophis schokari*, and *Bungarus candidus*; primer pair qPCR-F2/-R1 for *Tropidechis carinatus*, *Drysdalia coronoides*, *Fordonia leucobalia*, *Furina diadema*, and *Vermicella annulata;* and primer pair qPCR-F1/R2 for *Sistrurus catenatus* and *Echis carinatus*. RT-qPCR was performed on the Viia 7 real-time PCR system, and analysis of relative gene expression was carried out using the 2-ΔΔCT method according to the manufacturer's guidelines [17].

Multiplex PCR and automated capillary electrophoresis

Both the 2000 bp and 600 bp snake FV B-domain templates were amplified in a single reaction mixture using HOT FIREPol DNA polymerase and primers A2-F2/BR-R1/AR*-R1 for *P. textilis*, *O. scutellatus*, *D. coronoides* and *F. diadema* cDNA, and primers A2-F2/BR-R1/AR*-R2 for *P. schokari* and *S. catenatus* cDNA (**Table 1**). The reaction was performed with 200 nM of forward primer and 100 nM of both reverse primers at 57.1°C annealing temperature. Automated capillary electrophoresis was performed on the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) using the Agilent DNA 1000 Kit, and relative ratios were calculated based on a 25-1000 bp DNA standard (Agilent Technologies, Santa Clara, CA, USA) using the Bioanalyzer software package.

Bioinformatics

For the pre-RNA spliceosome analysis of full-length snake B-domain DNA sequences, the Human Splicing Finder (HSF) software version 3.1 (http://www.umd.be/HSF3/index.html) was employed making use of a transcription database based on Ensembl release number 70 - January 2013 [18, 19]. The theoretical isoelectric point (pI) of snake specific sequences was calculated with the ProtParam tool on the ExPASy server (https://web.expasy.org/compute_pi/) [20]

Results

Evolutionary and geographic snake origins

To investigate the nature of the B-domain of liver-expressed snake FV, fresh-frozen liver tissues from twelve distinct snake species were obtained (**Table 2**). The collection included two elapid species known to produce venom analogues of FVa and FXa (*Pseu. textilis*—pseutarin C and *O. scutellanus*—scutarin C) [21, 22], and another elapid known to produce a venom analogue of FXa only (*T. carinatus*—trocarin D) [23]. Four additional elapid species that have not been reported to produce venom analogues of FVa or FXa were also included: three species that inhabit the same geographic region as *P. textilis* (*D. coronoides*, *F. diadema*, and *V. annulata*), and one species that is native to Southeast Asia and Indonesia (*B. candidus*). Evolutionarily distinct species of snakes (last shared ancestor >50 million years ago [24]) from different geographic regions were also included: a non-venomous colubrid species (*O. octolineatus*); a mildly venomous species of the family Psammophiidae (*P. schokari*); a potentially venomous species of the family Homalopsidae (*F. leucobalia*); and two venomous viperid species (*S. catenatus* and *E. carinatus*). Of the latter, *E. carinatus* is well known for its venom-derived metalloprotease ecarin, which catalyzes the conversion of prothrombin into thrombin independent from FVa and FXa [25].

Table 2. Names, family, and geographic distribution of the studied snakes

	Common name	Latin name	Family	Range
1	Common brown snake	Pseudonaja textilis	Elapidae	Eastern Australia
2	Coastal taipan	Oxyuranus scutellatus	Elapidae	Northern Australia and New Guinea
3	Rough-scaled snake	Tropidechis carinatus	Elapidae	Eastern Australia
4	White-lipped snake	Drysdalia coronoides	Elapidae	Wider Australia
5	Red-naped snake	Furina diadema	Elapidae	Eastern Australia
6	Bandy-bandy	Vermicella annulata	Elapidae	Eastern Australia
7	Malayan krait	Bungarus candidus	Elapidae	Southeast Asia to Indonesia.
8	Striped kukri snake	Oligodon octolineatus	Colubridae	Indonesia
9	Schokari sand racer	Psammophis schokari	Colubridae	Africa, Southwest Asia
10	Crab-eating water snake	Fordonia leucobalia	Colubridae	Southeast Asia to Northern Australia
11	Massasauga	Sistrurus catenatus	Viperidae	North America
12	Saw-scaled viper	Echis carinatus	Viperidae	Middle East and Central Asia

Multiple factor V B-domain amplicons with varying lengths are generated from snake cDNA

Fresh frozen liver tissue samples were individually processed for RNA extraction and cDNA generation. PCR amplification of the FV B-domain was enabled by designing primers based on sequence conservation between the known FV sequences from the anole lizard (*Anolis carolinensis*, uniprot: G1KK71) [26] and the elapid king cobra (*O. hannah*, uniprot: V8P243) [16]. For cDNA amplification, the forward primer was designed to bind a conserved stretch within the *A. carolinensis* and snake FV A2 domains, and similarly designed reverse primers targeted the FV A3-domain. Interestingly, PCR amplification of *P. textilis* and *O. scutellanus* cDNA using one primer set consistently produced two 'long' (~2000 bp) B-domain amplicons and one 'short' (600 bp) B-domain amplicon, irrespective of annealing temperature (**Figure 1**). Furthermore, a less abundant semi-long amplicon of approximately 1600 bp was generated. From all cDNA samples of the other snakes similarly sized B-domain amplicons were generated using the same primer pairs

Next, the relative ratio between the 600 bp and 2000 bp B-domain amplicons was examined in more detail employing multiplex PCR and automated capillary electrophoresis (Bioanalyzer). By doing so, we were able to confirm that the long amplicon was more abundant in most of the elapid species as well as the colubrid, psammophiid, and homalopsid species assessed (**Figure 2A,B**). However, considerable differences were noted in the ratio between the short and long amplicons among species belonging to the family Elapidae. For instance, the long B-domain transcript was only 2-fold more abundant in *P. textilis* and *O. scutellatus*, whereas it was 26-fold and 76-fold more expressed in *D. coronoides* and *F. diadema*, respectively. In contrast, in viperid snakes the short B-domain transcript was 12 to 30-fold more abundant relative to the long B-domain. Similar trends among species and between snake families were also observed employing RT-qPCR using specifically designed primer pairs (**Figure 2B**).

The full-length B-domain of snake FV contains conserved auto-inhibitory regions

Sanger sequencing of the 2000 bp and 600 bp B-domain amplicons demonstrated that each long B-domain transcript contained an open reading frame that encoded an amino acid sequence that was highly similar (>90% identity) between snakes. Interestingly, each of these B-domain sequences contained three putative thrombin/FXa cleavage motifs that were similar to the human FV Arg^{709}, Arg^{1018}, and Arg^{1545} motifs (**Figure 3**) [4]. In addition, each individual B-domain sequence contained two discrete regions that were rich in either basic residues (KRK…PKV; theoretical pI: 9.9; **Figure 3** Basic Region (BR)) or acidic residues (DYQ…PDD; theoretical pI: 3.7; **Figure 3** Acidic Region (AR)). The order of these elements was also similar to that of the human FV B-domain, suggesting functional

relevance as auto-inhibitory B-domain elements (**Figure 3**). Collectively, these data demonstrate that the *P. textilis* genome comprises a gene encoding for full-length FV. However, considerable differences were noted in the relative spacing of these elements along the snake FV polypeptide chain when compared to human FV. For example, the snake BR is located 100 residues further downstream from the Arg^{709} cleavage motif when aligned with the human BR. On the other hand, the spacer sequence that separates the Arg^{1018} cleavage motif from the snake AR is nearly 410 residues shorter in the snake B-domain (**Figure 3**). The full-length snake B-domain sequence (504 ± 4 residues) is thus considerably shorter than that of the human B-domain (834 residues) and resembles the B-domain sequence identified in the king cobra [16].

Similar to the large B-domain transcript, we uncovered that the short B-domain transcript contained an open reading frame. Following amino acid sequence alignment, all short B-domain transcripts were shown to encode B-domain sequences that were highly similar (>90% identity) to the truncated B-domain sequence that was previously identified in the *P. textilis* venom gland and liver tissues [13, 15]. The short B-domain sequence features only the Arg^{709} and Arg^{1545} cleavage motifs and a short 45-residue linker separating these two motifs (**Figure 4**). Notably, the first part of this linker is formed by a distinct stretch of residues that lies adjacent to the putative Arg^{709} cleavage motif (*P. textilis;* EEE…AHAS; **Figure 4**). The second part is formed by the last 14 residues from the putative snake AR sequence (*P. textilis;* DPR…PDD; **Figure 4**), the latter being over 500 amino acids downstream from the Arg^{709} motif in the full-length B-domain sequence (**Figure 3**). The complete linker sequence is negatively charged (theoretical pI: 3.9) and is therefore labeled 'alternative' Acidic Region (AR*).

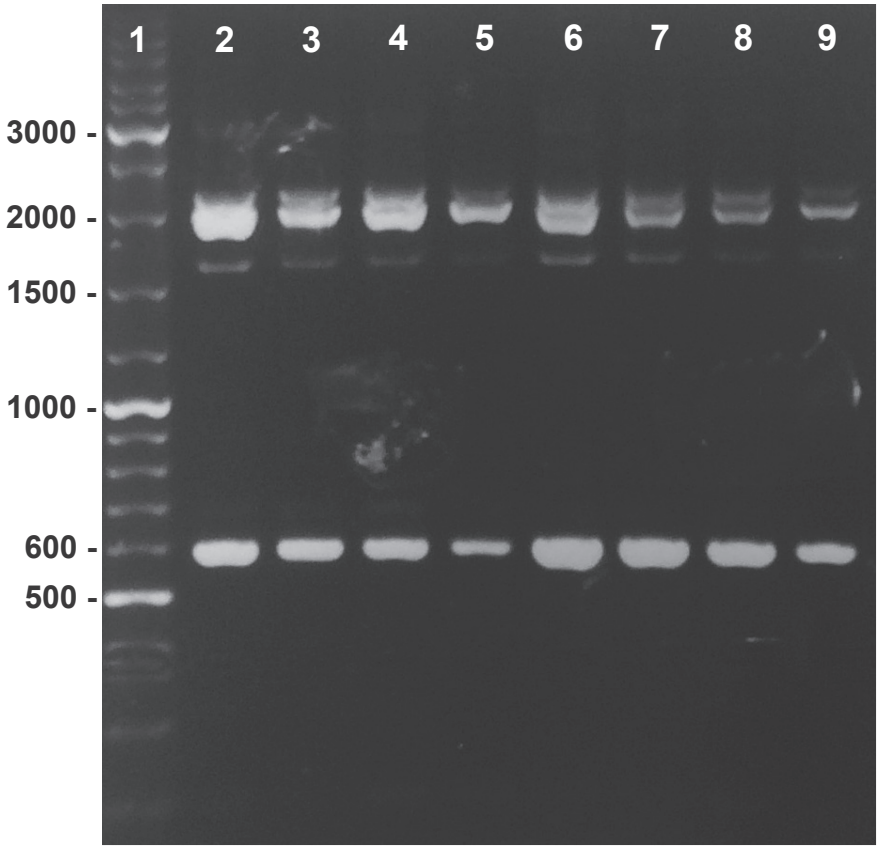
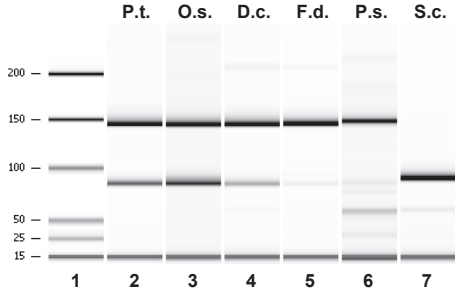


Figure 1. FV B-domain amplicons from *P. textilis* and *O. scutellanus*. Touchdown PCR amplicons from *P. textilis* and *O. scutellatus* liver-derived cDNA were amplified with forward primer A2-F1 and reverse primer A3-R1 at annealing temperatures (Ta) ranging between 57.9°C and 69.6°C and were subsequently run on a 1.0% Agarose gel. Lane 1: 1 kb GeneRuler Plus DNA Ladder, lanes 2-5: *P. textilis* amplicons, lanes 6-9: *O. scutellatus* amplicons, lanes 2,6: Ta 57.9°C, lanes 3,7: Ta 58.2°C °C, lane 4,8: Ta 63.0°C, and lane 5,9: Ta 64.5°C. The apparent bp lengths of the DNA standards are indicated.

A



B

	Snake	Bioanalyzer	RT-qPCR
1	*P. textilis*	1:2	1:3
2	*O. scutellatus*	1:2	1:4
3	*T. carinatus* *	-	1:197
4	*D. coronoides*	1:26	1:38
5	*F. diadema*	1:76	1:207
6	*V. annulata* *	1:4	1:17
7	*B. candidus* *	-	1:615
8	*O. octolineatus* *	1:7	1:636
9	*P. schokari*	1:8	1:453
10	*F. leucobalia* *	-	1:1032
11	*S. catenatus*	30:1	14:1
12	*E. carinatus* *	12:1	2:1

Figure 2. Ratios of 600 and 2000 bp cDNA amplicons. (***A***) Electropherogram of multiplex-PCR derived B-domain amplicons from six representative snake species as analyzed by automated capillary electrophoresis. The ratio of the short over the long snake FV B-domain transcript was quantified using a predefined standard (lane 1) and PCR samples were analyzed in parallel (lanes 2 – 7); The ~150 bp band corresponds to the full-length B-domain and the ~90 bp band corresponds to the truncated B-domain. The long and short B-domain templates were amplified using a forward primer that bound the snake A2-domain (A2-F2) and a pair of reverse primers that targeted either the full-length snake B-domain BR (BR-R1) or the truncated B-domain AR* (AR*-R1/AR*-R2), respectively (**Table 1**). Lane 1: 25-1000 bp DNA standard, lane 2: *P. textilis* (P.t.), lane 3: *O. scutellatus* (O.s.), lane 4: *D. coronoides* (D.s.), lane 5: *F. diadema* (F.d.), lane 6: *P. schokari* (P.s.), lane 7: *S. catenatus* (S.c.). (***B***) Summary of automated capillary electrophoresis ('Bioanalyzer') and RT-qPCR experiments showing the ratio between truncated and full-length B-domain transcripts. The data are representative for two experiments; the asterisk indicates cDNA species that have been tested once.

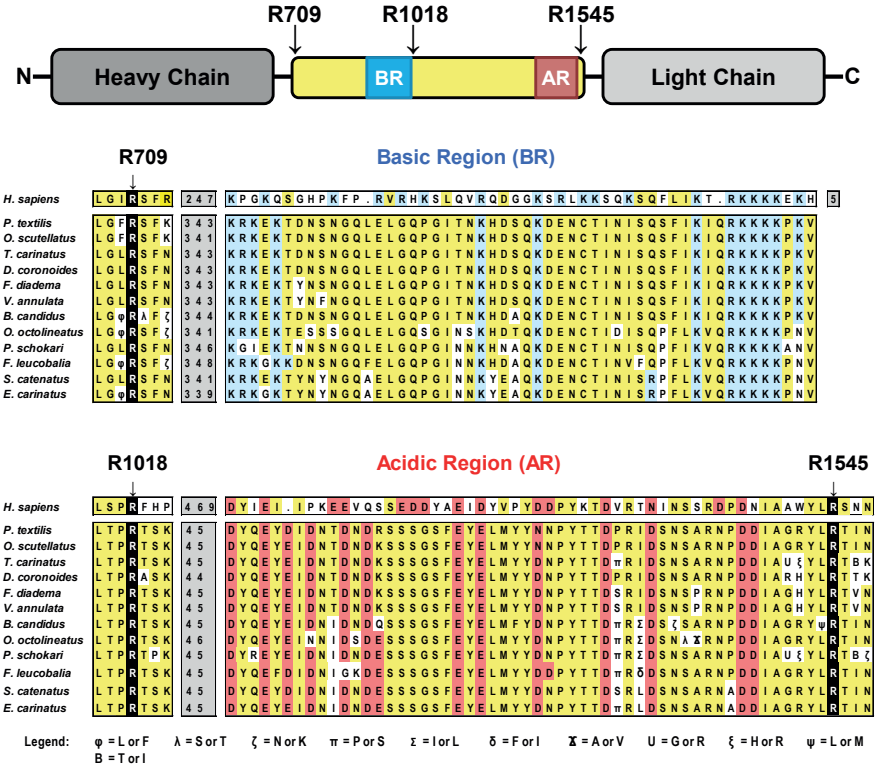


Figure 3. Conservation of regulatory B-domain elements within the full-length snake FV B-domain. ***Top:*** The human FV polypeptide chain is defined by an N-terminal heavy chain (A1-A2 domains; dark grey), a central regulatory B-domain (yellow) and a C-terminal light chain (A3-C1-C2 domains; light grey). The B-domain is removed after thrombin/FXa-mediated proteolysis; human cleavage sites are indicated above the B-domain. The blue box represents the Basic Region (BR; a.a. 963–1008) and the red box represents the Acidic Region (AR; a.a. 1493–1537). ***Middle:*** Amino acid alignment of the first cleavage motif (R709) and putative BR in full-length snake FV as identified through sequencing of 2000 bp cDNA amplicons from twelve snake species. The human sequence is shown for reference purposes, and the numbers within the grey boxes denote the number of amino acids between the regulatory elements. ***Bottom:*** Amino acid alignment of the second (R1018) and third (R1545) cleavage motifs and (putative) AR in human and snake FV, for which the mature human FV residue numbering is used. Greek letters represent amino acid discrepancies between a regular B-domain sequence (this figure) and a truncated B-domain sequence (**Figure 4)** of the same species, which are defined in the legend below the figure. Highlighted residues (yellow, red, or blue) represent identical or highly homologous amino acids at a particular location. Red highlighted residues are acidic and blue are basic. *P. textilis* refers to *P. textilis*.

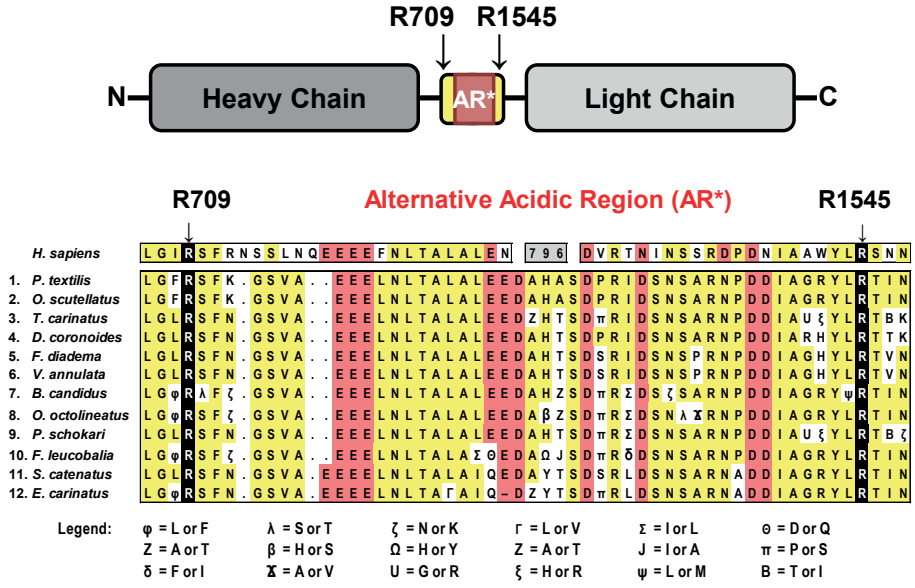


Figure 4. Sequence conservation within the truncated snake FV B-domains. ***Top:*** The truncated snake FV polypeptide chain is defined by an N-terminal heavy chain (A1-A2 domains; dark grey), a short central B-domain (± 46 residues; yellow) and a C-terminal light chain (A3-C1-C2 domains; light grey). The short B-domain retains the first (R709) and third (R1545) thrombin/FXa cleavage motifs and an alternative Acidic Region (AR*) that is represented by a red box. ***Bottom:*** Amino acid alignment of the R709 and R1545 thrombin cleavage motifs, for which the mature human FV residue numbering is used, and the putative AR* in truncated snake FV sequences as identified through sequencing of the ±600 bp amplicons from twelve snake species. The human FV B-domain sequence is shown for reference purposes, and the number within the grey box denotes the number of amino acids of the full-length human B-domain sequence that are not shown here. Greek letters represent amino acid discrepancies between the truncated B-domain sequence (this figure) and the regular B-domain sequence (**Figure 3)** of the same species, which are defined in the legend below the figure. Highlighted residues (yellow, red, or blue) represent identical or highly homologous amino acids at a particular location. Red highlighted residues are acidic and blue are basic. *P. textilis* refers to *P. textilis*.

Putative splice sites enable alternative splicing of the snake factor V B-domain

In order to investigate whether alternative splicing might account for the generation of a truncated B-domain transcript, we performed spliceosome analysis using the online splice site finder tool HSF [18]. Splice site analysis annotated several pre-mRNA exonic splicing enhancer motifs within the full-length snake B-domain transcripts. One such motif included a sequence that is typical for splicing in which the SRp40 protein assists, which was conserved in most of the species. Importantly, we observed that the precise location of the donor-acceptor sites was similar in each snake full-length B-domain transcript (**Figure 5**) and corresponded exactly with the sequence of the 600 bp fragments. Specifically, the snake FV sequence that is retained after putative splicing at the donor splice site precisely overlapped with the first part of the AR* sequence (AHAS, **Figure 4**). Similarly, the sequence that was retained after putative splicing at the acceptor splice site exactly matched the second part of the AR* sequence (DPRI, **Figure 4**). These findings suggest evolutionary conservation of the donor-acceptor motif in the B-domain of snake FV. Of note, this specifc donor-acceptor motif was not conserved in *O. octolineatus* and *E. carinatus* (**Figure 5**). Given that the truncated B-domain transcript was clearly identified in both of these snakes (**Figure 2A,B**), additional verification of these two B-domain sequences is required. Finally, as SR proteins involved in splicing are generally conserved among vertebrates [27, 28], spliceosome analysis was also performed on human, mouse, and anole lizard FV B-domain sequences. However, analogous exonic splicing enhancer motifs were not found within these sequences (data not shown).

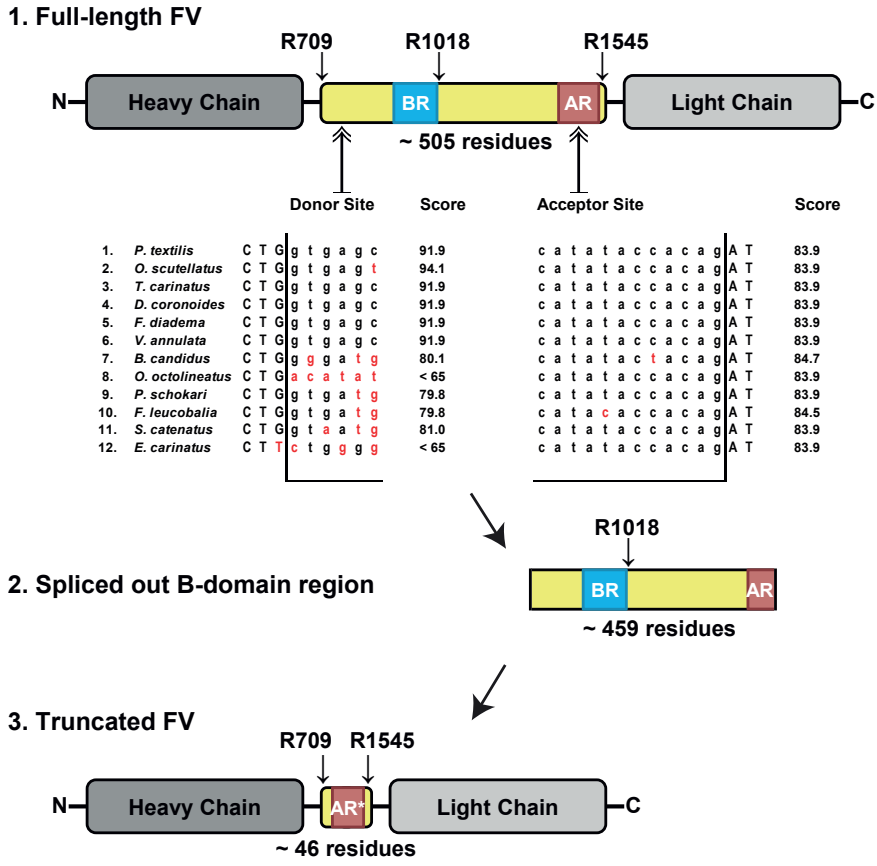


Figure 5. Putative SRp40-assisted donor and acceptor splices in the B-domain of snake FV. ***1:*** Splice site analysis of the full-length snake FV B-domain cDNA sequence revealed a putative donor site splice located C-terminal to the R709 cleavage site and a putative acceptor splice site located within the AR. The donor and acceptor splice sites are shown in the sequence alignment, dissimilar nucleotides are highlighted in red, and capital letters indicate sequences retained after splicing. Splice site probability scores assigned by HSF are specified, with the threshold for a putative donor or acceptor splice site being defined at 65; each HSF score above the threshold is considered a splice site. ***2:*** The regulatory B-domain elements that are lost after splicing of the snake FV transcript include the complete BR, the R1018 cleavage site, and part of the AR. ***3:*** The truncated snake FV transcript contains the R709 and R1545 cleavage consensus sites and an alternative acidic region (AR*) formed by a stretch of ±22 amino acids C-terminal to the R709 site and the last ±13 amino acid of the consensus AR (see **Figure 4)**. The mature human FV residue numbering is used to denote the thrombin/FXa cleavage motifs. *P. textilis* refers to *P. textilis*.

Discussion

In this study we have uncovered that the liver transcriptome of the Australian snake *P. textilis* contains two unique transcripts of FV. One transcript comprises a newly identified (full-length) FV sequence that includes a B-domain which contains all of the essential auto-inhibitory elements that comprise the FV PRR [7, 8]. Similar full-length FV B-domain transcripts were also identified in snakes from separate families, geographic regions, and venomous vs. non-venomous states, which suggest that the B-domain has been preserved as potential regulator of FV activity in all members of the suborder Serpentes. Similarly structured FV cDNAs have been found in all vertebrates, from fish to mammals. The other transcript incorporated a small ~46 amino acid B-domain sequence that is similar to the truncated B-domain FV sequence previously described for liver-expressed *P. textilis* FV [13, 15]. Remarkably, these small B-domain species share high sequence homology with the truncated B-domain from the FV-like component that is found in the venom of *P. textilis*, *O. scutellatus*, and *Oxyuranus microlepidotus* [13, 29, 30].

Through the use of bioinformatics we were able to identify the SRp40 pre-RNA splicing protein as a potential candidate for mediating the alternative splicing of the full-length snake FV B-domain sequence. Importantly, the location of the associated splice junction was similar in each B-domain sequence and corresponded to the exact sequence of the truncated B-domain. From these results we infer that the truncated FV B-domain transcript was generated through an alternative splicing event. Interestingly, apart from the full-length and truncated FV sequence, we also detected a B-domain amplicon of approximately 1600 bp in PCR reactions on *P. textilis* and *O. scutellanus* cDNA (**Figure 1**). While these amplicons might be the result of primer promiscuity or heteroduplex formation between the 2000 and 600 bp amplicon, they could alternatively represent less abundant FV splicing variants. Interestingly, a *P. textilis* venom gland transcriptome analysis performed by Viala et al. identified another splice variant of venom FV [31]. This splice variant was also detected at the protein level and comprised a unique A3-domain insertion of 34 residues upstream from the annealing site of our reverse primer (A3-R1; Table 1) used in cDNA amplification (**Figure 1**). The fact that this FV variant was not identified in the current study is not unexpected as we studied liver transcripts. In addition, it may be related to the general notion that venom genes undergo alternative splicing and accelerated evolution to give rise to a diverse toxin repertoire [32].

Quantitative analyses have shown that the full-length FV transcripts were more abundant than the truncated versions in most snake liver tissues. However, both transcripts were expressed at almost equal ratios in *P. textilis* and *O. scutellatus* (**Table 2**), which are those elapids that comprise truncated FV in their venom.

These results may be indicative of a link between the expression levels of truncated FV variants in the liver and venom gland [21]. On the other hand, our findings indicate that the truncated B-domain FV transcript was most abundant in liver tissue from the two viperid snakes, which have not been reported to comprise FV as toxin in their venom. At present, we have no explanation for this finding, other than that the Elapidae and Colubridae (including the families Psammophiidae and Homalopsidae, which were until recently considered colubrids) are evolutionarily closer (median distance of 50 million years) relative to the Viperidae (median distance of 57.5 million years to both Elapidae and Colubridae) [24]. We speculate that the Viperidae may have undergone a divergent molecular evolution, which could have resulted in different transcript profiles and abundances. However, clear evidence for this assumption is lacking thus far. Furthermore, recent studies into the venom toxins of the genus *Pseudonaja* have revealed an intricate link between venom FV expression and genus [33], geographic location [34], ontogenetic dietary shifts [35], and dietary niches [36]. A number of these factors might also affect the expression and alternative splicing of the liver FV transcript in *P. textilis* or possibly in other snake species. However, additional information on expression levels of blood coagulation factors in snakes is presently not available.

Recent studies have highlighted the significance of alternatively spliced human FV variants in the regulation of tissue factor pathway inhibitor (TFPI), a key regulatory plasma protein of the prothrombinase complex. It was shown that a splice variant of human FV, known as FV-short in which part of the B-domain comprising the BR is spliced out, interacts with TFPI through the AR region in the FV B-domain [37-39]. As we have previously shown that the truncated B-domain FV variant from *P. textilis* displays full cofactor activity [40] and makes up ~25-30% of total FV in *P. textilis* (**Figure 2B**), this would suggest that approximately one third of circulating FV species exist in an active cofactor state in *P. textilis*. As such, our findings could allude to an alternative mode of FV regulation in snakes. In this respect, it is attractive to consider TFPI as a regulator of truncated B-domain FV, potentially by interacting with the highly acidic stretch of amino acids that forms the AR* [39]. However, it remains unclear whether the truncated B-domain snake FV transcript is translated into a plasma protein. Additional research is therefore required to confirm the physiological presence of FV species in snake plasma and to assess plasma levels of key regulatory proteins of FV, such as TFPI.

In summary, we have revealed the presence of two distinct FV transcripts in the liver transcriptome of three major families of the suborder Serpentes: a full-length B-domain transcript that includes all putative regulatory B-domain elements and a truncated B-domain transcript that lacks these essential B-domain sequences. Online database search tools show that alternative splicing of the full-length FV transcript likely accounts for the presence of the truncated FV transcript in liver

tissues. Overall, these findings suggest that the snake FV gene has undergone unique functional adaptations, potentially in response to the extreme physiological and phenotypic adaptations that are typical to snakes [16, 41].

References

1. Jiang, Y. and R.F. Doolittle, *The evolution of vertebrate blood coagulation as viewed from a comparison of puffer fish and sea squirt genomes.* Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7527-32.
2. Davidson, C.J., et al., *Molecular evolution of the vertebrate blood coagulation network.* Thromb Haemost, 2003. **89**(3): p. 420-8.
3. Suzuki, K., B. Dahlback, and J. Stenflo, *Thrombin-catalyzed activation of human coagulation factor V.* J Biol Chem, 1982. **257**(11): p. 6556-64.
4. Jenny, R.J., et al., *Complete cDNA and derived amino acid sequence of human factor V.* Proc Natl Acad Sci U S A, 1987. **84**(14): p. 4846-50.
5. Corral-Rodriguez, M.A., et al., *Structural basis of thrombin-mediated factor V activation: the Glu666-Glu672 sequence is critical for processing at the heavy chain-B domain junction.* Blood, 2011. **117**(26): p. 7164-73.
6. Schuijt, T.J., et al., *Factor Xa activation of factor V is of paramount importance in initiating the coagulation system: lessons from a tick salivary protein.* Circulation, 2013. **128**(3): p. 254-66.
7. Bos, M.H. and R.M. Camire, *A bipartite autoinhibitory region within the B-domain suppresses function in factor V.* J Biol Chem, 2012. **287**(31): p. 26342-51.
8. Bunce, M.W., et al., *Restoring the procofactor state of factor Va-like variants by complementation with B-domain peptides.* J Biol Chem, 2013. **288**(42): p. 30151-60.
9. Mann, K.G., et al., *Surface-dependent reactions of the vitamin K-dependent enzyme complexes.* Blood, 1990. **76**(1): p. 1-16.
10. Rosing, J., et al., *The role of phospholipids and factor Va in the prothrombinase complex.* J Biol Chem, 1980. **255**(1): p. 274-83.
11. Rao, V.S. and R.M. Kini, *Pseutarin C, a prothrombin activator from Pseudonaja textilis venom: its structural and functional similarity to mammalian coagulation factor Xa-Va complex.* Thromb Haemost, 2002. **88**(4): p. 611-9.
12. Lechtenberg, B.C., et al., *Crystal structure of the prothrombinase complex from the venom of Pseudonaja textilis.* Blood, 2013. **122**(16): p. 2777-83.
13. Rao, V.S., S. Swarup, and R.M. Kini, *The nonenzymatic subunit of pseutarin C, a prothrombin activator from eastern brown snake (Pseudonaja textilis) venom, shows structural similarity to mammalian coagulation factor V.* Blood, 2003. **102**(4): p. 1347-54.
14. Bos, M.H., et al., *Venom factor V from the common brown snake escapes hemostatic regulation through procoagulant adaptations.* Blood, 2009. **114**(3): p. 686-92.
15. Minh Le, T.N., et al., *Gene duplication of coagulation factor V and origin of venom prothrombin activator in Pseudonaja textilis snake.* Thromb Haemost, 2005. **93**(3): p. 420-9.
16. Vonk, F.J., et al., *The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system.* Proc Natl Acad Sci U S A, 2013. **110**(51): p. 20651-6.

17. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
18. Desmet, F.O., et al., *Human Splicing Finder: an online bioinformatics tool to predict splicing signals.* Nucleic Acids Res, 2009. **37**(9): p. e67.
19. Flicek, P., et al., *Ensembl 2013.* Nucleic Acids Res, 2013. **41**(Database issue): p. D48-55.
20. Wilkins, M.R., et al., *Protein identification and analysis tools in the ExPASy server.* Methods Mol Biol, 1999. **112**: p. 531-52.
21. Masci, P.P., A.N. Whitaker, and J. de Jersey, *Purification and characterization of a prothrombin activator from the venom of the Australian brown snake, Pseudonaja textilis textilis.* Biochem Int, 1988. **17**(5): p. 825-35.
22. Speijer, H., et al., *Prothrombin activation by an activator from the venom of Oxyuranus scutellatus (Taipan snake).* J Biol Chem, 1986. **261**(28): p. 13258-67.
23. Joseph, J.S., et al., *Amino acid sequence of trocarin, a prothrombin activator from Tropidechis carinatus venom: its structural similarity to coagulation factor Xa.* Blood, 1999. **94**(2): p. 621-31.
24. Kumar, S., et al., *TimeTree: A Resource for Timelines, Timetrees, and Divergence Times.* Mol Biol Evol, 2017. **34**(7): p. 1812-1819.
25. Kornalik, F. and B. Blomback, *Prothrombin activation induced by Ecarin - a prothrombin converting enzyme from Echis carinatus venom.* Thromb Res, 1975. **6**(1): p. 57-63.
26. Alfoldi, J., et al., *The genome of the green anole lizard and a comparative analysis with birds and mammals.* Nature, 2011. **477**(7366): p. 587-91.
27. Goren, A., et al., *Comparative analysis identifies exonic splicing regulatory sequences--The complex definition of enhancers and silencers.* Mol Cell, 2006. **22**(6): p. 769-81.
28. Lev-Maor, G., et al., *The "alternative" choice of constitutive exons throughout evolution.* PLoS Genet, 2007. **3**(11): p. e203.
29. Welton, R.E. and J.N. Burnell, *Full length nucleotide sequence of a factor V-like subunit of oscutarin from Oxyuranus scutellatus scutellatus (coastal Taipan).* Toxicon, 2005. **46**(3): p. 328-36.
30. St Pierre, L., et al., *Comparative analysis of prothrombin activators from the venom of Australian elapids.* Mol Biol Evol, 2005. **22**(9): p. 1853-64.
31. Viala, V.L., et al., *Venomics of the Australian eastern brown snake (Pseudonaja textilis): Detection of new venom proteins and splicing variants.* Toxicon, 2015. **107**(Pt B): p. 252-65.
32. Ogawa, T.a.S., H., *Venomics Study of Protobothrops flavoviridis Snake: How Venom Proteins Have Evolved and Diversified?* Medical Toxicology, 2020. DOI: 10.5772/intechopen.91960
33. Reeks, T., et al., *Deep venomics of the Pseudonaja genus reveals inter- and intra-specific variation.* J Proteomics, 2016. **133**: p. 20-32.
34. McCleary, R.J., et al., *Proteomic comparisons of venoms of long-term captive and recently wild-caught Eastern brown snakes (Pseudonaja textilis) indicate venom does not change due to captivity.* J Proteomics, 2016. **144**: p. 51-62.

35. Cipriani, V., et al., *Correlation between ontogenetic dietary shifts and venom variation in Australian brown snakes (Pseudonaja).* Comp Biochem Physiol C Toxicol Pharmacol, 2017. **197**: p. 53-60.
36. Skejic, J., *Venoms of related mammal-eating species of taipans (Oxyuranus) and brown snakes (Pseudonaja) differ in composition of toxins involved in mammal poisoning.* BioRxiv, 2018. DOI: 10.1101/378141
37. Vincent, L.M., et al., *Coagulation factor V(A2440G) causes east Texas bleeding disorder via TFPIalpha.* J Clin Invest, 2013. **123**(9): p. 3777-87.
38. Cunha, M.L., et al., *A novel mutation in the F5 gene (factor V Amsterdam) associated with bleeding independent of factor V procoagulant function.* Blood, 2015. **125**(11): p. 1822-5.
39. Wood, J.P., et al., *Tissue factor pathway inhibitor-alpha inhibits prothrombinase during the initiation of blood coagulation.* Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17838-43.
40. Verhoef D., Y.X., Parthasarathy S., Reitsma P.H., Camire R.M., Bos M.H.A, *Functional implications of the unique disulfide bond in venom factor V from the Australian common brown snake Pseudonaja textilis.* Toxin Reviews, 2013. **33**(1-2): p. 37-41.
41. Castoe, T.A., et al., *The Burmese python genome reveals the molecular basis for extreme adaptation in snakes.* Proc Natl Acad Sci U S A, 2013. **110**(51): p. 20645-50.