Research Article

Common evolution of waprin and kunitz-like toxin families in Australian venomous snakes

L. St Pierre^{a,b,*}, S. T. Earl^{a,b}, I. Filippovichª, N. Sorokinaª, P. P. Masci^c, J. De Jersey^c and M. F. Lavin^{a,b}

^a The Queensland Cancer Fund Research Unit, The Queensland Institute of Medical Research, P.O. Box Royal Brisbane Hospital, Herston, Brisbane 4029 (Australia), Fax: 617 3362 0106, e-mail: Martin.Lavin@qimr.edu.au **b** Center for Clinical Research, The University of Queensland, Brisbane (Australia)

^c School of Medicine, Southern Clinical Division, The University of Queensland, Brisbane (Australia)

Received 14 September 2008; received after revision 08 October 2008; accepted 09 October 2008 Online First 3 November 2008

Abstract. The venoms of Australian snakes contain a myriad of pharmacologically active toxin components. This study describes the identification and comparative analysis of two distinct toxin families, the kunitztype serine protease inhibitors and waprins, and demonstrates a previously unknown evolutionary link between the two. Multiple cDNA and full-length gene isoforms were cloned and shown to be composed of three exons separated by two introns. A high degree of identity was observed solely within the first exon which coded for the propeptide sequence and its cleavage site, and indicates that each toxin family has arisen from a gene duplication event followed by diversification only within the portion of the gene coding for the functional toxin. It is proposed that while the mechanism of toxin secretion is highly conserved, diversification of mature toxin sequences allows for the existence of multiple protein isoforms in the venom to adapt to variations within the prey environment.

Keywords. Elapid, waprin, kunitz-like inhibitors, venom, toxin, gene.

Introduction

Australian elapid snake venoms contain a complex mixture of potent peptidic molecules that specifically target, and adversely effect, the homeostatic mechanisms of their prey. Numbered among these are a class of small molecules that demonstrate inhibitory activity towards proteases, most notably those involved in haemostatic processes [1]. These toxins demonstrate significant structural and functional identity to the kunitz class of serine protease inhibitors, the most well characterised member of which is bovine pancreatic trypsin inhibitor (BPTI) or aprotinin [2]. Kunitz-type inhibitors are approximately 7kDa in size and are typically characterised by a tight tertiary fold made up of three conserved disulfide bonds. They act as competitive antagonists binding to, and inhibiting, the active site of serine proteases such as trypsin and chymotrypsin [3].

Kunitz-type inhibitors have been isolated from the venom of a number of Elapidae and Viperidae snakes [4–7]. In addition to their ability to act as protease inhibitors, some snake venom kunitz-like homologs have evolved as neurotoxins by inhibiting calcium and potassium channels [8]. To date, two distinct types of kunitz inhibitors have been characterised from the venom of Australian elapid snakes. Taicatoxin serine protease inhibitor (TSPI) was purified as part of a * Corresponding author. multimeric complex from the venom of the coastal

taipan $(Oxyuranus\, such that is)$ and has been shown to have inhibitory activity towards both trypsin and chymotrypsin [9]. Subsequently, a second family of kunitz-type inhibitors were purified, cloned and characterised from the venom of the common brown snake (Pseudonaja textilis) [10, 11]. Two of these inhibitors, Textilinin-1 and -2, demonstrated potent inhibition of plasmin and are thought to be responsible for the anti-fibrinolytic properties of this venom [12].

More recently, a new class of small molecule peptides called waprins have been described from snake venom. First isolated from the venom of the green mamba (Naja nigricollis), waprins demonstrate structural homology to the whey acidic protein (WAP) family [13]. Members of the WAP family are typically 50 amino acids in approximate length with a conserved four disulfide bond arrangement with variable intervening residues. Identified in a number of tissue types from numerous organisms, they have been shown to have inhibitory activity towards a wide variety of proteases, including roles in the innate immune system and anti-bacterial effects [14–16]. The functional activity of venom waprins, however, still remains poorly characterised, with the only known activity to be the selective antimicrobial effects of Omwaprin recently reported from the Australian elapid Oxy uranus microlepidotus (inland taipan) [17]. Furthermore, there is little information on the evolutionary and genomic relationship of this potentially significant toxin family. Indeed, the recent description of the presence of waprins within the venom of five Colubrid snakes represents the first description of a cDNA sequence for this toxin family [18].

Advances in transcriptomic and proteomic technologies have contributed to the identification of some of the less abundant toxins within the venom of Australian snakes [19, 20]. However, to date only a handful of full-length genomic sequences have been identified for Australian elapids toxins, most notably for the phospholipase A_2 (PLA₂) and neurotoxin families from P. textilis [21–23]. This study describes the cloning and comparative analysis of two types of kunitz inhibitor along with waprin toxin sequences from a total of eleven Australian elapid snakes, demonstrating a common evolutionary relationship between the three types of toxin. This data is supported by identification of the genomic sequences of all three toxins, in conjunction with venom immunoblotting, and represents one of the most comprehensive comparative studies of the gene sequences from Australian elapids to date.

Method and Materials

RNA and DNA isolation. Venom glands were excised from Australian elapid snakes including the coastal taipan (Oxyuranus scutellatus), inland taipan (Oxyuranus microlepidotus), common brown snake (Pseudonaja textilis), red-bellied black snake (Pseudechis porphyriacus), mulga (Pseudechis australis), tiger snake (Notechis scutatus), rough-scaled snake (Tropidechis carinatus), Stephen's banded snake (Hoplocephalus stephensii), small-eyed snake (Rhinoplocephalus nigrescens), black whip snake (Demansia vestigiata) and copperhead snake (Austrelaps superbus), RNA isolated and cDNA synthesised as previously described [24]. In addition, genomic DNA was isolated from O. scutellatus, O. microlepidotus and P. textilis via the Tri Reagent method (Sigma, St. Louis, MO) with final samples resuspended in sterile water and stored at -20 \degree C.

Identification of waprin toxin cDNA sequences. The mature protein sequence of native Omwaprin has previously been described from the venom of O. microlepidotus (P83952). Using this native sequence, forward $5'$ -AA(A/G) GA(C/T) (A/C)G(A/C/G/T) CC(A/C/G/T) AA(A/G) AA(A/G) CC(A/C/G/T) GG-3' and reverse $5'-TC(A/G)TC(C/T)TT(A/G)$ $CA(A/C/G/T) CC(A/G) TC(A/G) TT(A/G) CA-3'$ degenerate primers were designed to amplify a corresponding region of the cDNA sequence. PCR amplification was performed from O. microlepidotus cDNA with 1 unit of AmpliTaq Gold DNA polymerase buffered in 1x buffer, 2 mM $MgCl₂$ and 200 μ M d NTPs with 50 ρ mol of each primer (Applied Biosystems, Foster City, CA). The reaction was thermocycled 95 \degree C for 8 min followed by 30 cycles of 95 \degree C for 25 s, 53 \degree C for 25 s and 72 \degree C for 30 s with a final extension of 72 \degree C for 5 min. A single, 124bp product was visualised on a 1% TAE agarose gel, excised using a QIAex II gel extraction kit (Qiagen, Hilden, Germany) and cloned via the pGEM-T vector system (Promega, Madison, WI). The ligation mix was transformed into competent dH5a Escherichia coli, with clones selected on LB-Ampicillin $(50\mu g/mL)$ plates supplemented with IPTG and X-Gal. Multiple clones were isolated and sequenced with an ABI Big Dye Terminator cycle sequence ready reaction kit (Perkin-Elmer, Norwalk, CT) and alignments performed with BioEdit Software (Isis Pharmaceuticals Inc., Carlsbad, CA).

From the degenerate Omwaprin PCR product, internal forward (5-CAC TTC TGC TGC CCA GGA CAA CTC CAG-3) and reverse (5-CTG TGT CCC CCA CGC CCT CAA AAA CC-3) primers were designed for 5' and 3' RACE. Reactions were

performed with a SMART RACE cDNA Amplification kit according to manufacturer's instructions (Clontech, Palo Alto, CA), prior to cloning, sequencing and alignment as described above. After identification of the $5'$ and $3'$ untranslated regions of the Omwaprin cDNA sequence, forward (5-GAG CTT CAT CAT GTC TTC TGG-3) and reverse (5-GGT GGG TTT CCT TGT TGG-3) primers were designed to amplify the full-length cDNA sequence of waprins from a total of 10 Australian elapids. Amplification and cloning was performed as described above, with reactions thermocycled at $95 °C$ for 8 min followed by 32 cycles of 95 \degree C for 20 s, 58 \degree C for 20 s and 72 \degree C for 30 s with a final extension of 72 [°]C for 3 min.

Identification of kunitz-type inhibitor toxin cDNA sequences. Identification of Kunitz-type inhibitor cDNA sequences was performed by two different methods. Initially, full-length PCR was performed using forward (5'ATG TCT TCT GGA GGT-3') and reverse (5TCA GGC AGC ACA GGT-3) primers designed from the previously identified Textilinin-1 cDNA sequence from P. textilis (AF402324). The reaction was set up as previously described, and thermocycled at 95 \degree C for 8 min followed by 30 cycles of 95 °C for 25 s, 50 °C for 25 s and 72 °C for 1 min with a final extension of $72 \degree C$ for 5 min. Additional kunitz sequences were also identified via 3'RACE as described above, using a propeptide specific forward primer (5GAG CTT CAT CAT GTC TTC TGG AGG TCT TCT TC-3'). Full-length PCR and 3'RACE products from 11 Australian snake species were subsequently cloned and sequenced, and alignments performed.

Recombinant protein production. Recombinant protein samples used in antibody production, immunoblot detection and functional assays were produced by a number of methods. Recombinant Textilinin-1 was produced as an E. coli expression product by Bresa-Gen Ltd. (Australia) as previously described [12]. Recombinant omwaprin-b was initially produced as a bacterial GST-fusion product using the pGEX-6P-1 E. coli expression system (GE Healthcare, Rydalmere, Australia). Subsequent to cloning into the pGEX-6P-1 vector, recombinant omwaprin-b was expressed and purified according to the protocols of Frangioni and Neel [25]. Additionally, both omwaprin-b and TSPI were produced recombinantly as a yeast expression product using the EasySelect Pichia pastoris Expression Kit (Invitrogen) according to manufacturer's protocol. Expression constructs were generated using the cDNA sequence corresponding to the mature omwaprin-b and TSPI, and cloned into the pPICZ α A vector via XhoI and XbaI restriction sites. Constructs were then transformed into KM71H P. pastoris, positive clones selected, protein expression induced and recombinant protein purified by immobilised metal affinity chromotography with Ni-NTA nickel agarose beads according to manufacturer's instructions (Qiagen, Hilden, Germany).

Immunoblot detection of waprins and kunitz-type inhibitors in snake venom. Polyclonal antibodies were raised in rabbits which were injected with one of either three toxin antigens: native Textilinin purified from the venom of P. textilis, native TSPI purified from the venom of O. scutellatus and the recombinant omwaprin-GST fusion product generated from O. microlepidotus cDNA. The immunisation protocol was performed by the Institute of Medical Veterinary Sciences (Adelaide, Australia) and antibodies subsequently purified by affinity chromatography as previously described [26]. The presence of kunitz-type inhibitors, TSPI and waprins in the venom of Australian elapids was then determined via immunoblot analysis. Briefly, a total of 30 µg of lyophilized crude venom resuspended in 50% glycerol, 50% saline from O. microlepidotus, O. scutellatus and P. textilis were separated under reducing conditions on a 15% SDSpolyacrcylamide gel in addition to 5μ g of recombinant omwaprin-b, Textilinin-1 and TSPI as controls. Venom proteins were subsequently transferred to a nitrocellulose membrane at 100 V for 1 h at 4 \degree C and probed with one of the three primary antibodies described above. Immunoblots were then probed with anti-rabbit secondary antibody conjugated to horseradish peroxidase and the protein signal detected with enhanced chemiluminescence reagents (Perkin-Elmer).

Functional assays. The activity of recombinant omwaprin-b was probed using a number of functional assays. The potential for omwaprin-b to act as an inhibitor of enzymatic activity was examined against a number of key enzymes within the haemostatic system including plasmin, trypsin, tissue plasminogen activator, kallikrein and urokinase. Tests of enzyme inhibition were performed as a chromogenic assay with appropriate controls as previously described [12]. The effects of omwaprin-b upon blood coagulation and fibrinolysis was examined using a thromboelastograph (Haemoscope Corporation, Niles, Illinois) in addition to tests of prothrombin time and activated partial thromboplastin time [27, 28]. Anti-bacterial effects of omwaprin-b were further examined using the gram positive bacterial strains Staphylococcus aureus, Staphylococcus epidermis and Streptococcus epidermis in addition to the gram negative strains Escherichia coli and Pseudomonas aeruginosa according to

4042 L. St Pierre et al. Elapid waprin and kunitz-like toxins

Figure 1. Full-length cDNA sequence of Omwaprin-b of Omwaprin-b (DQ917552) identified by $5'$ and 3 RACE and PCR from the venom gland of the inland taipan (Oxyuranus microlepidotus), aligned with the predicted translated protein sequence. mRNA features including 5' and 3' UTR as well as polyadenylation signal are indicated, as well as the propeptide and mature peptide sequences of the corresponding protein.

the protocols described by Nair et al. [17]. Potential neurotoxic effects of omwaprin-b were also examined via an ex vivo amphibian electrophysiological assay, testing inhibitory effects on end-plate potentials (EPPs) and minature EPPs at the toad neuromuscular-junction as previously described [29].

Identification of toxin genomic sequences. The full length gene sequences of the three toxin families from the eleven Australian snakes under investigation were identified by PCR and cloning from genomic DNA preparations. Waprin genomic sequences were amplified by PCR from all snakes with forward (5-GAG CTT CAT CAT GTC TTC TGG-3) and reverse (5- CTT CCA ATT CTG GTG GGT TTC CTT G-3) primers. Kunitz-type genes were amplified with forward (5-GAG CTT CAT CAT GTC TTC TGG AGG TCT TCT TC-3) and reverse (5-GAT CCA ATC CAG GAG GGT CTC CTC A-3) primer sequences, whilst the TPSI gene sequence was isolated with the same forward primer sequence, but specific reverse (5-TGC CAT GCA GAT TCT GCT AGG GTG-3) sequence. PCR products were subsequently amplified, cloned and sequenced and alignments performed as described above.

Results

Identification of waprin toxin cDNA sequences. The full-length cDNA clone of a member of the waprin family from the venom gland of *O. microlepidotus* was

determined via a combination of PCR, 5' and 3' RACE (Fig. 1). Differing by three amino acids in comparison to the mature protein sequence of the reported native peptide Omwaprin (P85932), this toxin was designated the name Omwaprin-b. Cloning of Omwaprin-b demonstrated the presence of a 24 amino acid propeptide leader sequence within the corresponding translated protein, along with a 10 nucleotide 5'UTR and 251 nucleotide 3'UTR. Subsequently, toxin gene homologues were cloned and sequenced from the venom glands of a total of nine other Australian elapids (Fig. 2). Alignment of the deduced waprin protein sequences demonstrates a significantly high degree of homology between the Australian snakes, with 76% overall identity and complete conservation of the cysteine residues involved in putative disulfide bond formation. Phylogenetic analysis of the waprin cDNA sequences with those recently reported from a number of colubrid snakes also indicate this high degree of identity, with clustering of the Australian elapid sequences into a single clade upon phylogenetic analysis (Supplementry Fig. 1A). Figure 3A demonstrates an alignment of the translated protein sequence of Omwaprin-b with the mature protein sequences of native Omwaprin and the first identified venom waprin sequence, Nawaprin, from N. nigricollis [13]. A significant degree of sequence variability between the two snake species was observed apart from conservation of a number of proline and cysteine residues, suggesting a highly divergent function for these molecules within the venom.

Figure 2. Alignment of the deduced amino acid sequence of waprins identified from a total of 10 Australian elapid snakes. Sequence identity is shown in black, gaps inserted for optimal alignment and the cleavage site between propeptide leader sequence and mature protein indicated. Conserved cysteine residues involved in putative disulfide bond formation are shaded gray and the Genbank accession numbers for each cDNA clone are provided at the end of each sequence.

Figure 3. (A) Alignment of the deduced protein sequence of Omwaprin-b from multiple cDNA clones with the native protein sequences of Omwaprin (P85932) from the venom of the inland taipan (Oxyuranus microlepidotus) and Nawaprin (P60589) from the Spitting cobra (Naja nigricollis). Note the presence of a 24 amino acid propeptide leader sequence in the Omwaprin precursor protein as well as conservation of a number of cysteine and proline residues between the Naja and Oxyuranus species. (B) Alignment of the deduced amino acid sequence from multiple cDNA clones of Omwaprin-b from O. microlepidotus (DQ917552), Taicatoxin Serine Protease Inhibitor-1 from O. scutellatus (EF112393) and Textilinin-1 from P. textilis (AF402324). Note complete conservation of the propeptide leader sequence, as well as the first four amino acids of the mature protein across the three distinct toxin families from the three different snake species despite significant differences in the mature toxin.

Identification of different kunitz-type toxin cDNA sequences. Given the known presence of kunitz-type inhibitor cDNA sequences within the venom glands of P. textilis and D. vestigiata, and given the conservation of their propeptide leader sequences with the waprin protein family, the existence of these toxins within other Australian elapids was investigated via a combination of 3'RACE and PCR for comparative analysis. Figure 4 demonstrates an alignment of kunitz-type inhibitor cDNA clones identified from a total of nine other Australian snakes with those reported from P. textilis and D. vestigiata. Note that multiple isoforms were identified in many of the snakes, with a high degree of identity within the propeptide leader sequence (a feature common to all Australian elapid toxin families) as well as complete conservation of cysteine residues involved in the formation of the

three disulfide bonds. Interestingly, the arginine residue present at position 19 of the mature protein sequences of Textilinin-1 and -2, which plays a key role in the plasmin inhibitory effects of these toxins (as well as in the bovine homolog, Aprotinin), is absent from every other kunitz-type inhibitor cloned from the other Australian elapids [10, 30]. Indeed, there is a significant degree of sequence variability within this region of the molecule, suggesting Australian snake venom kunitz-type inhibitors may have evolved to adopt a range of different activities. The Australian elapid kunitz-type inhibitors are, however, evolutionarily distinct from other kunitz-type inhibitors previously characterised from the venom glands of a number of vipers, colubrids and non-Australian elapids, clustering together in a phylogenetic analysis (Supplementary Fig. 1B).

A cDNA clone representing TSPI was also identified using the $3'RACE$ approach in $O.$ scutellatus. This sequence was consistent with the amino acid sequence for TSPI previously determined by Possani et al. [9]. Comparison between the TSPI precursor protein sequence and Textilinin-1 precursor protein sequence demonstrates significant variation between the two types of kunitz inhibitor (Fig. 3B). In particular, the TSPI precursor protein is 5 amino acids longer than Textilinin-1 and the other kunitz inhibitors from the Australian elapid snakes with only 68% amino acid sequence identity between TSPI and Textilinin-1. Interestingly, comparisons of Omwaprin-b with the precursor proteins of TSPI-1 from O. scutellatus and Textilinin-1 from *P. textilis* demonstrates 100% conservation of the propeptide sequences between these three toxins, as well as the first four amino acids of the mature protein, which represent the cleavage site for the leader sequence (Fig. 3B). This is despite a relatively low degree of sequence identity of 10% through the remainder of the mature protein. Conservation of the propeptide sequences suggests that all three toxin families are subject to a highly conserved mechanism of processing and secretion within the venom gland despite occurring in three different species, and also suggests a common point of evolution despite the significant degree of variability within the mature protein sequence.

Detection of waprins and kunitz-type inhibitors in venom. Since multiple cDNA isoforms for each of the waprin and kunitz-type serine protease inhibitor sequences were cloned from the venom glands of Australian elapids, their actual presence within the venom was confirmed by immunoblotting with antibodies raised against either native or recombinant forms of each toxin protein. Figure 5A demonstrates loading of an SDS-polyacrylamide gel with reduced venom from the two taipan species O. microlepidotus and O. scutellatus, as well as the brown snake P. textilis, along with recombinant forms of Omwaprin-b from O. microlepidotus, Textilinin-1 from P. textilis and TSPI from O. scutellatus. Affinity-purified antibody raised against recombinant Omwaprin-b detected the presence of this toxin, but failed to detect either recombinant Textilinin-1 or TSPI (Fig. 5B). A band of similar molecular size was also detected in the venom of O. microlepidotus, confirming the presence of Omwaprin-b in this snake. Interestingly, although a less abundant band was observed in the venom of the closely related O. scutellatus, no protein was detected from the venom of P. textilis supporting the absence of cDNA for this toxin within the venom gland of this snake. When *O. microlepidotus* venom was run under non-reducing conditions and an immunoblot analysis performed with the Omwaprin-b specific antibody, a band of similar molecular weight to the single peptide was detected, suggesting that Omwaprin-b does not form a covalent complex with any other toxin within the venom (result not shown).

Immunoblotting with an antibody raised against the native form of Textilinin from the venom of P. textilis detected the recombinant and native forms of this toxin but failed to detect either recombinant Omwaprin-b or TSPI from the taipans (Fig. 5C). A less prominent band was also observed in O. microlepidotus, probably representing some cross reactivity with the Microlepidins which were observed in the cDNA of this elapid. Similarly, an antibody raised against native TSPI from the venom of O. scutellatus was specific for the recombinant form of this toxin and only detected protein from the venom of this snake (Fig. 5D). This is in agreement with observations at both the cDNA and genomic DNA levels, which indicated that a TSPI-like gene was only present in O. scutellatus. Finally, immunoblotting with commercial CSL taipan antivenom demonstrated significant cross-reactivity against the denatured venoms of all three Australian snakes examined (Fig. 5E). This immuno-reactivity may be a result of the detection of either the waprins, the kunitz-type serine protease inhibitors or another toxin family of similar molecular size such as the α -neurotoxins. However, whilst the taipan anti-venom was strongly immuno-reactive with recombinant TSPI and somewhat reactive with the recombinant kunitz-type inhibitor Textilinin-1, it failed to detect recombinant Omwaprin-b, which may be a reflection of the low abundance of this toxin within the venom of the coastal taipan from which the anti-venom is derived.

Common genomic structures of toxin gene families. The multitude of toxin sequences observed in snake venoms is often the result of a process of gene duplication and diversification under positive selection [3]. To further characterise the evolutionary relationship between waprin and kunitz-type toxin families in Australian elapid venoms, full-length genomic sequences were amplified by PCR, cloned and sequenced from the same 11 snake species from which cDNA clones were isolated. Table 1 contains a list of the waprin, kunitz-type and TSPI-like genomic sequences identified from the Australian snakes, along with nomenclature, Genbank references, gene size and whether an exact corresponding cDNA sequence was also observed. Note that for many of the snake species, multiple genomic sequences were observed for each toxin family, suggesting that they have undergone gene duplication and divergence. A number of single nucleotide polymorphism differ-

Figure 4. Alignment of the deduced amino acid sequence of kunitz-type serine protease inhibitor cDNA clones from eleven Australian elapids. Note the presence of multiple toxin isoforms in most species. The propeptide and mature protein sequences have been indicated, with putative cysteine residues involved in disulfide bond formation shaded gray. Genbank accession numbers are provided at the end of each sequence and the P1 active site denoted by *.

ences between genomic and cDNA sequences were also evident.

Alignment of representative genomic sequences from a member of each toxin family revealed similar genomic structures composed of three exons separated by two introns (Fig. 6A). Of significant interest is the almost complete sequence conservation of the first exon, not only between snake species, but also between the waprin, kunitz-type and TSPI toxin families. This first exon represents the entire propeptide leader sequence as well as the first four amino acids of the mature protein (KDRP) corresponding to the cleavage site of the active toxin (Figs. 6B–D). Note in Figure 3B that it is this sequence that suggests an evolutionary link between the waprins and two types of kunitz serine protease inhibitor. Therefore, the significant degree of variation observed between the mature proteins of these toxin families can be solely attributed to divergence within the second exon only, with positive selection on the first exon for its almost complete conservation. Alignment of the genomic DNA sequence with the transcribed cDNA sequences allows determination of the intron/exon gene organisation, with all genomic sequences identified from the Australian elapids in this study observing the GT/AG rule of intron/exon boundaries.

The genomic sequence of Omwparin-c from O. microlepidotus contains three exons 95bp, 141bp and 34bp in size with two intervening introns 422bp and 1049bp in size (Fig. 6B). An alignment of the genomic sequence of Omwaprin-c with the cDNA sequence of Omwaprin-b indicates only a few, single nucleotide changes within exon-1. Note that the stop codon occurs just a single nucleotide from the end of the second exon, with the third exon represented by just 3' UTR sequence. The full-length genomic sequences from the other Australian snakes can be found in Supplementary Figure 2A–J. These include a sequence cloned from *P. textilis*, Textwaprin-a, despite no detection of a corresponding cDNA sequence in the venom gland via either PCR or 3' RACE methods,

Figure 5. Immunoblot detection of Australian snake venom proteins. (A) Coomassie stained gel demonstrating loading of 30 mg O. microlepidotus, O. scutellatus and P. textilis venom samples and 5 µg of recombinant omwaprin, textilinin and TSPI. The presence of specific toxin families was determined via western blot analysis using a range of antibodies, including those raised against (B) recombinant omwaprin, (C) native textilinin, (D) native TSPI and (E) CSL taipan anti-venom.

nor immunoblot detection of a native protein within the venom (Table 1 and Fig. 5B). This suggests that expression of Textwaprin-a is either negligible or nonexistent, which may be the result of potential alterations to the as yet unidentified promoter region.

The genomic sequences of the kunitz-type protease inhibitors, as exemplified by Textilinin-2 from P. textilis in Figure 6C, demonstrated overall similar structural organisation to the waprins, but were larger in size due to longer intronic sequences. The exception to this was the single genomic sequence identified from P. australis, whose second intron contained approximately a 1kb deletion (Table 1 and Supplementary Fig. 2B). The coding sequence for the mature protein of the kunitz type inhibitors also covered the entire second exon, with the penultimate codon and stop codon occurring at the beginning of the third exon, as opposed to the waprin sequences, whose stop codon was located within the second exon. This is reflected in the larger size of the mature kunitz-type inhibitors Table 1. List of the genomic sequences identified from Australian elapids for the waprin, kunitz-like and TSPI-like toxin families.

Table 1 (Continued)

Species	Toxin name	Genbank	Gene size	cDNA observed
Waprin				
	Vestiginin-3	EU401860	2138bp	Yes
Taicatoxin Serine Protease-like Inhibitors				
O. scutellatus	TSPI	EU401816	1210bp	Yes
A. superbus	BPTI-like toxin	EU401815	828bp	No

Table 2. Summary of known functional activities of waprins, the kunitz-type inhibitor textilinin, and TSPI identified from the venom of Australian elapid snakes.

compared to the waprin proteins by approximately nine amino acids. The overall domain arrangement, (including intron/exon boundaries sites) of the kunitztype inhibitor genes from Australian elapids closely reflects that of the only other known kunitz genes from two other snakes: a putative chymotrypsin inhibitor from the Taiwan cobra, Naja atra, and the B chain isoforms of the potassium channel blocker β bungarotoxin from the Taiwan banded krait, Bungarus multicinctus [5, 31, 32]. Indeed, the Textilinin-2 gene sequence demonstrates 87% identity to the N. atra sequence, whilst it is less similar to the β bungarotoxin B1 chain gene, primarily due to the presence of a greater than 1kb insertion within the second intron of this sequence. It is also notable that a genomic sequence was identified from each of the taipan species, O. microlepidotus (EU401845) and O. scutellatus (EU401842), which shared significant identity to the other kunitz-type inhibitor genomic structures, with the exception that they were missing the entire sequence corresponding to exon 2 (Supplementary Fig. 2K and L). These sequences probably reflect divergence and a loss of gene function; however, they may be involved in alternate splicing and processing of the other functional kunitz-type or even waprin gene products.

Although sharing overall domain arrangement with the waprins and other kunitz-type serine protease inhibitors, the TSPI gene sequence from O. scutellatus was significantly smaller (due to a shorter second intron) and contained more coding sequence within its third exon (Fig. 6D). Comparisons of intronic sequences between the three different toxin families indicated very little sequence homology. The exception to this was the first intron of Scuwaprin-a compared to the TSPI-1 genomic sequence from O. scutellatus, where identity over the first three quarters of the intron was as high as 94%, before dropping to 21% over the remainder of the intron, suggesting possible events of homologous recombination during gene evolution. It is interesting to note that it is the first intronic sequences of TSPI and waprins that are more closely related, despite the mature protein sequence of TSPI (and hence exon 2 sequences) being more similar to the kunitz-type inhibitors. A sole TSPI-like genomic sequence was also cloned from the DNA of the copperhead snake, A. superbus, despite no evidence for a cDNA sequence in the venom gland of this elapid (Supplementary Fig. 2V). This sequence demonstrated significant identity with TSPI-1 from O. scutellatus up until the start of intron 2, where homology drops off. The presence of a TSPI-like sequence within the copperhead snake suggests the evolution of this toxin gene occurred prior to the divergence of these species, with subsequent loss of function in A. superbus.

Functional relationships of waprin and kunitz-type toxins. In an attempt to further characterise the functional role of the waprins within the venom of Australian snakes, a recombinant version of Omwaprin-b from O. microlepidotus was produced in both bacterial and yeast protein expression systems. These recombinant proteins were examined in a number of

B) O. microlepidotus Omwaprin-c Genomic Sequence

Figure 6. Toxin sequence genomic organisation. (A) Schematic structural organisation of the genomic sequences of waprins, kunitz-like inhibitors and TSPI from Australian snake venoms. Representative genomic sequences of (B) the waprin, Omwaprin-c (EU401817) from the inland taipan (O. microlepidotus), (C) the kunitz-type inhibitor, textilinin-2 (EU401838) from the common brown snake (P. textilis) and (D) taicatoxin serine protease inhibitor (EU401816) from the coastal taipan (O . *scutellatus*). For genomic sequence alignments of each toxin family for the remaining Australian elapid snakes (indicated in Table 1), refer to Supplementary Figure 2.

C) P. textilis Textilinin-2 Genomic Sequence

Figure 6. (continued)

functional assays, including tests for inhibitory activity against enzymes including plasmin, kallikrein, trypsin, tissue plasminogen activator and urokinase, an examination of their coagulopathic and fibrinolytic effects, tests of their antibacterial properties as well as an assay examining the potential neurotoxic effects of the molecule. None of these assays detected a positive activity with recombinant Omwaprin-b (results not shown). A summary of the observed and known activities of Omwaprin-b is compared to those previously reported from the kunitz-type inhibitor Textilinin, as well as TSPI, in Table 2. Despite the evolutionary relationship between these toxin families, the waprins do not appear to act as inhibitors of the same class of enzymes that the kunitz-type toxins are specific for.

D) O. scutellatus TSPI-1 Genomic Sequence

Figure 6. (continued)

Discussion

Snake venoms are a source of diverse, yet functionally specific bioactive molecules often encoded by multigene families. Amongst the most potent venoms in the world are those belonging to the Australian elapid snake family, which are known to contain multiple peptidic toxins with neurotoxic, coagulopathic, antifibrinolytic and myotoxic effects to name a few [1]. In this study, we described the cloning and comparative analysis of members of three types of small-molecule toxins from the venom glands of a total of eleven Australian elapids, the waprins along with two distinct types of kunitz-type inhibitor, and demonstrated a previously unrealised evolutionary relationship between the three gene families.

The waprins represent a recently characterised toxin family whose functional role within the venom remains to be fully elucidated. Previously, only a handful of waprin sequences have been identified from snake venoms, namely the native protein sequences of Nawaprin and Omwaprin, as well as a handful of cDNA sequences from a number of colubrid snakes [13, 17, 18]. This study represents the first description of a cDNA sequence of a waprin from an elapid snake. A high degree of sequence identity, and complete conservation of cysteine residues between the Australian snake waprin sequences suggest they adopt a similar conformational fold, and hence would be expected to have similar functional activities. Indeed, the three dimensional conformation and the presence of a number of N-terminal positively charged amino acid residues have previously been shown to be important for the antimicrobial activities of Omwaprin [17]. A recent report of a fused toxin protein cloned from the venom gland of the viper Sistrurus catenatus edwardsii (Desert Massasauga rattlesnake, DQ464286) demonstrates some sequence homology to the Australian elapid waprins [33]. This clone contains a leader sequence which shares 83% identity to the propeptide sequence reported here, followed by a fusion of two sequences that have homology to a kunitz-type inhibitor and a waprin, respectively. The authors propose that the fused toxin

is a result of exon shuffling or transcriptional splicing, and is further evidence for an evolutionary relationship between these two toxin families.

This study also reports the presence of multiple kunitz-type inhibitor cDNA isoforms from the venom glands of an additional nine Australian elapids. Sequence comparisons demonstrate a high degree of identity between the toxins, particularly within the cysteine residues, propeptide and C-terminal regions of the molecule. This identity, however, does not extend to the P1 antiproteinase site at amino acid position 19 of the mature protein, which dictates specificity towards serine proteases. None of the variants in the Australian snakes contain the arginine residue present within Textilinins-1 and -2 and aprotinin which have been demonstrated to be necessary for plasmin and trypsin inhibition [10, 30]. This is in keeping with the absence of plasmin inhibitory activity observed in the crude venom of all Australian snakes except for Oxyuranus and the Pseudonaja genera [11]. However, a number of clones contain leucine, phenylalanine and tyrosine residues within the P1 active site, which are known to be important for chymotrypsin inhibition [2]. The conservation of the number and spacing of cysteine residues across the Australian snake elapid kunitz-like inhibitor clones indicates that they adopt a similar tertiary fold; however, variations within key active sites of the molecule suggest they may have evolved to have altered functions or enzyme specificities within the venom. Indeed, a class of kunitz-type toxins from other snake venoms have lost their enzymatic inhibitory activity, and are instead neurotoxic, acting as K^+ and Ca^{2+} channel blockers [34]. The neurotoxic effects of the kunitz-type inhibitors from Australian elapid venoms remain to be determined, with the numerous cDNA isoforms identified in this study serving as an excellent platform for such an investigation. In addition to these kunitz-type inhibitors cloned from Australian snake venom glands was a second type of kunitz molecule, the serine protease inhibitor component of Taicatoxin, a multimeric Ca^{2+} channel blocker isolated from the venom gland of O. scutellatus [9, 35]. TSPI appears to be unique to this species venom, with immunoblot analysis, cDNA cloning and genomic DNA sequencing all failing to detect a homolog within the other elapid species. Hence, O. scutellatus appears to be unique in that its venom contains two types of kunitz-like inhibitors: TSPI and Scutellin.

In this study we compared the genomic sequences of multiple isoforms of both waprin and kunitz-type inhibitors from a total of eleven Australian elapid snakes, and demonstrated a common evolutionary link between the three gene families. Molecular sequence data is a powerful phylogenetic tool, in that it can provide a measure of the order and timing of divergence between species [36]. Cloning of toxin cDNA sequences identified a highly conserved propeptide between the waprins and kunitz-type inhibitors not only within each snake species, but also between species. Identification of the corresponding gene sequences revealed that this propeptide region, along with the first four conserved amino acids of the mature protein, were encoded entirely by a single exon which was present and conserved across all species. The remainder of the gene sequences, however, were highly divergent between the waprins and kunitz-type inhibitors, although they did retain their overall genomic structural organisation of three exons separated by two introns. This suggests that the waprin and kunitz-type inhibitor gene sequences arose from a common ancestral gene with subsequent duplication and diversification. The significant degree of variation across the gene compared to the highly conserved first exon also suggests that this multigene family is subject to variable rates of evolution under selective pressure, as is observed for other snake venom toxin gene families such as the phospholipase A_2 (PLA₂) family [37, 38]. It is interesting to note that the sequence coding for the propeptide is not subject to the same degree of divergence as that coding for the functional mature toxin, and is hence, not subject to the same selective pressure of host-prey interaction.

Despite the significant clinical effects of envenomation from an Australian elapid, very little genomic sequence data is available for venom components for most species, being confined predominantly to a number of PLA_2 and neurotoxin sequences from P. textilis and the Factor X-like prothrombin activator Trocarin D gene from T. carinatus [21, 22, 39]. The genomic data from T. carinatus does, however, indicate that expression of Trocarin D was recruited to the venom gland via the duplication and divergence of the snake's own haemostatic blood coagulation Factor X gene typically expressed in the liver [39]. The recruitment and tissue specific expression of other venom gland toxins in Australian snakes, however, including the waprins and kunitz-type inhibitors, remains to be determined. It is interesting to note that the waprin toxin family demonstrates structural homology to other whey-acidic proteins present in milk, suggesting a bias towards this stable structural conformation in secreted fluids. Hence, this study describes for the first time the complete gene cloning and comparative analysis of a number of toxin isoforms from three different toxin types, demonstrating a hitherto unrealised evolutionary relationship which may have important implications for determining their function within the venom.

Electronic supplementary material. Supplementary material is available in the online version of this article at springerlink.com (DOI 10.1007/s00018-008-8573-5) and is accessible for authorized users.

Acknowledgements. We wish to thank Dr Katia Fisher for advice on the recombinant expression of protein in yeast as well as Mr Joe Sambono for the provision of venom glands. We also wish to acknowledge Dr Michael Batzloff and Mr Mitja Schenning for assistance in the anti-microbial and neurotoxic functional assays respectively. Financial assistance was provided by a joint collaboration between QRxPharma and the Australian Research Council.

- 1 Fry B. G. (1999) Structure-function properties of venom components from Australian elapids. Toxicon. 37, 11 – 32.
- 2 Laskowski M., Jr. and Kato I. (1980) Protein inhibitors of proteinases. Annu. Rev. Biochem. 49, 593 – 626.
- 3 Zupunski V., Kordis D., and Gubensek F. (2003) Adaptive evolution in the snake venom Kunitz/BPTI protein family. FEBS Lett. 547, 131 – 136.
- 4 Chang L., Chung C., Huang H. B., and Lin S. (2001) Purification and characterization of a chymotrypsin inhibitor from the venom of Ophiophagus hannah (King Cobra). Biochem. Biophys. Res. Commun. 283, 862 – 867.
- 5 Cheng Y. C., Yan F. J., and Chang L. S. (2005) Taiwan cobra chymotrypsin inhibitor: cloning, functional expression and gene organization. Biochim. Biophys. Acta. 1747, 213 – 220.
- 6 Shafqat J., Beg O. U., Yin S. J., Zaidi Z. H., and Jornvall H. (1990) Primary structure and functional properties of cobra (Naja naja naja) venom Kunitz-type trypsin inhibitor. Eur. J. Biochem. 194, 337 – 341.
- 7 Takahashi H., Iwanaga S., and Suzuki T. (1974) Distribution of proteinase inhibitors in snake venoms. Toxicon. 12, 193 – 197.
- 8 Cardle L. and Dufton M. J. (1997) Foci of amino acid residue conservation in the 3D structures of the Kunitz BPTI proteinase inhibitors: how do variants from snake venom differ? Protein Eng. 10, 131-136.
- 9 Possani L. D., Martin B. M., Yatani A., Mochca-Morales J., Zamudio F. Z., Gurrola G. B., and Brown A. M. (1992) Isolation and physiological characterization of taicatoxin, a complex toxin with specific effects on calcium channels. Toxicon. 30, 1343 – 1364.
- 10 Filippovich I., Sorokina N., Masci P. P., de Jersey J., Whitaker A. N., Winzor D. J., Gaffney P. J., and Lavin M. F. (2002) A family of textilinin genes, two of which encode proteins with antihaemorrhagic properties. Br. J. Haematol. 119, 376 – 384.
- 11 Masci P. P., Whitaker A. N., Sparrow L. G., de Jersey J., Winzor D. J., Watters D. J., Lavin M. F., and Gaffney P. J. (2000) Textilinins from Pseudonaja textilis textilis. Characterization of two plasmin inhibitors that reduce bleeding in an animal model. Blood Coagul. Fibrinolysis. 11, 385 – 393.
- 12 Flight S., Johnson L., Trabi M., Gaffney P., Lavin M., de Jersey J., and Masci P. (2005) Comparison of textilinin-1 with aprotinin as serine protease inhibitors and as antifibrinolytic agents. Pathophysiol. Haemost. Thromb. 34, 188 – 193.
- 13 Torres A. M., Wong H. Y., Desai M., Moochhala S., Kuchel P. W., and Kini R. M. (2003) Identification of a novel family of proteins in snake venoms. Purification and structural characterization of nawaprin from Naja nigricollis snake venom. J. Biol. Chem. 278, 40097 – 40104.
- 14 Araki K., Kuroki J., Ito O., Kuwada M., and Tachibana S. (1989) Novel peptide inhibitor (SPAI) of Na+, K+-ATPase from porcine intestine. Biochem. Biophys. Res. Commun. 164, $496 - 502$.
- 15 Hagiwara K., Kikuchi T., Endo Y., Huqun, Usui K., Takahashi M., Shibata N., Kusakabe T., Xin H., Hoshi S., MikiM., Inooka Tokue Y., and Nukiwa T. (2003) Mouse SWAM1 and SWAM2 are antibacterial proteins composed of a single whey acidic protein motif. J. Immunol. 170, 1973 – 1979
- 16 Wiedow O., Schroder J. M., Gregory H., Young J. A., and Christophers E. (1990) Elafin: an elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence. J. Biol. Chem. 265, 14791 – 14795.
- 17 Nair D. G., Fry B. G., Alewood P., Kumar P. P., and Kini R. M. (2007) Antimicrobial activity of omwaprin, a new member of the waprin family of snake venom proteins. Biochem. J. 402, $93 - 104.$
- 18 Fry B. G., Scheib H., van der Weerd L., Young B., McNaughtan J., Ramjan S. F., Vidal N., Poelmann R. E., and Norman J. A. (2008) Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). Mol. Cell. Proteomics. 7, 215 – 246.
- 19 Birrell G. W., Earl S. T., Wallis T. P., Masci P. P., de Jersey J., Gorman J. J., and Lavin M. F. (2007) The diversity of bioactive proteins in Australian snake venoms. Mol Cell Proteomics. 6, 973 – 986.
- 20 St Pierre L., Woods R., Earl S., Masci P. P., and Lavin M. F. (2005) Identification and analysis of venom gland-specific genes from the coastal taipan (Oxyuranus scutellatus) and related species. Cell. Mol. Life. Sci. 62, 2679 – 2693.
- 21 Armugam A., Gong N., Li X., Siew P. Y., Chai S. C., Nair R., and Jeyaseelan K. (2004) Group IB phospholipase A2 from Pseudonaja textilis. Arch. Biochem. Biophys. 421, 10-20.
- 22 Gong N., Armugam A., and Jeyaseelan K. (2000) Molecular cloning, characterization and evolution of the gene encoding a new group of short-chain alpha-neurotoxins in an Australian elapid, Pseudonaja textilis. FEBS Letters. 473, 303 – 310.
- 23 Gong N., Armugam A., Mirtschin P., and Jeyaseelan K. (2001) Cloning and characterization of the pseudonajatoxin b precursor. Biochem. J. 358, 647 – 656.
- 24 St Pierre L., Masci P. P., Filippovich I., Sorokina N., Marsh N., Miller D. J., and Lavin M.F. (2005) Comparative Analysis of Prothrombin Activators from the Venom of Australian Elapids. Mol. Biol. Evol. 22, 1853 – 1864.
- 25 Frangioni J. V. and Neel B. G. (1993) Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal. Biochem. 210, 179 – 187.
- 26 Bar-Peled M. and Raikhel N. V. (1996) A method for isolation and purification of specific antibodies to a protein fused to the GST. Anal. Biochem. 241, 140 – 142.
- 27 Filippovich I., Sorokina N., St Pierre L., Flight S., de Jersey J., Perry N., Masci P. P., and Lavin M.F. (2005) Cloning and functional expression of venom prothrombin activator protease from Pseudonaja textilis with whole blood procoagulant activity. Br. J. Haematol. 131, 237 – 246.
- 28 Langdell R. D., Wagner R. H., and Brinkhous K. M. (1953) Effect of antihemophilic factor on one-stage clotting tests; a presumptive test for hemophilia and a simple one-stage antihemophilic factor assy procedure. J. Lab. Clin. Med. 41, 637 – 647.
- 29 St Pierre L., Fischer H., Adams D. J., SchenningM., Lavidis N., de Jersey J., Masci P. P., and Lavin M. F. (2007) Distinct activities of novel neurotoxins from Australian venomous snakes for nicotinic acetylcholine receptors. Cell. Mol. Life. Sci. 64, 2829 – 2840.
- 30 Markwardt F. (1978) Pharmacological content of hyperproteolytic states in blood by synthetic inhibitors of serine proteinases. Haemostasis. 7, 177 – 182.
- 31 Cheng Y. C., Chen K. C., Lin S. K., and Chang L. S. (2006) Divergence of genes encoding B chains of beta-bungarotoxins. Toxicon. 47, 322 – 329.
- 32 Wu P. F. and Chang L. S. (2000) Genetic organization of A chain and B chain of beta-bungarotoxin from Taiwan banded krait (Bungarus multicinctus). A chain genes and B chain genes do not share a common origin. Eur. J. Biochem. 267, 4668 – 4675.
- 33 Pahari S., Mackessy S. P., and Kini R. M. (2007) 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. 8, 115.

- 34 Harvey A. L. (2001) Twenty years of dendrotoxins. Toxicon. 39, $15 - 26.$
- 35 Brown A. M., Yatani A., Lacerda A. E., Gurrola G. B., and Possani L. D. (1987) Neurotoxins that act selectively on voltage-dependent cardiac calcium channels. Circ Res. 61, $I6 - 19.$
- 36 Wuster W., Dumbrell A. J., Hay C., Pook C. E., Williams D. J., and Fry B. G. (2005) Snakes across the Strait: trans-Torresian phylogeographic relationships in three genera of Australasian snakes (Serpentes: Elapidae: Acanthophis, Oxyuranus, and Pseudechis). Mol. Phylogenet. Evol. 34, 1-14.
- 37 Kordis D. and Gubensek F. (2000) Adaptive evolution of animal toxin multigene families. Gene. 261, 43 – 52.
- 38 Moura-da-Silva A. M., Paine M. J., Diniz M. R., Theakston R. D., and Crampton J. M. (1995) The molecular cloning of a phospholipase A2 from Bothrops jararacussu snake venom: evolution of venom group II phospholipase A2's may imply gene duplications. J. Mol. Evol. 41, 174 – 179.
- 39 Reza M. A., Swarup S., and Kini R. M. (2007) Structure of two genes encoding parallel prothrombin activators in Tropidechis carinatus snake: gene duplication and recruitment of factor X gene to the venom gland. J. Thromb. Haemost. 5, 117 – 126.

To access this journal online: <http://www.birkhauser.ch/CMLS>