Identification of a Novel Family of Proteins in Snake Venoms

PURIFICATION AND STRUCTURAL CHARACTERIZATION OF NAWAPRIN FROM NAJA NIGRICOLLIS SNAKE VENOM*

Received for publication, May 21, 2003, and in revised form, July 22, 2003
Published, JBC Papers in Press, July 23, 2003, DOI 10.1074/jbc.M305322200

Allan M. Torres‡, Hui Y. Wong§, Malan Desai¶, Shabbir Moochhala¶, Philip W. Kuchel‡, and R. Manjunatha Kini¶**

From the ¶School of Molecular and Microbial Biosciences, University of Sydney, Sydney, New South Wales 2006, Australia, the §§Department of Pharmacy, Faculty of Science, National University of Singapore, Singapore, 119260 Singapore, the ¶¶Defence Medical Research Institute, Defence Science and Technology Agency, Singapore 117597, Singapore, and the ¶¶¶Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore 117543, Singapore

The three-dimensional structure of nawaprin has been determined by nuclear magnetic resonance spectroscopy. This 51-amino acid residue peptide was isolated from the venom of the spitting cobra, Naja nigricollis, and is the first member of a new family of snake venom proteins referred to as waprins. Nawaprin is relatively flat and disc-like in shape, characterized by a spiral backbone configuration that forms outer and inner circular segments. The two circular segments are held together by four disulfide bonds, three of which are clustered at the base of the molecule. The inner segment contains a short antiparallel β-sheet, whereas the outer segment is devoid of secondary structures except for a small turn or β10 helix. The structure of nawaprin is very similar to elafin, a human leukocyte elastase-specific inhibitor. Although substantial parts of the nawaprin molecule are well defined, the tips of the outer and inner circular segments, which are hypothesized to be critical for binding interactions, are apparently disordered, similar to that found in elafin. The amino acid residues in these important regions in nawaprin are different from those in elafin, suggesting that nawaprin is not an elastase-specific inhibitor and therefore has a different function in the snake venom.

Snake venoms are rich sources of pharmacologically active polypeptides and proteins. Some of these proteins exhibit enzymatic activities. These enzymes include phospholipase A₂, proteinase, nucleotidase, phosphodiesterase, and 1-amino acid oxidase. In addition to their catalytic properties that may contribute to the digestive action of the venom, these enzymes also induce various pharmacological effects including neurotoxic, myotoxic, cardiotoxic, hemorrhagic, hemolytic, procoagulant, and anticoagulant effects (1, 2). Several other snake venom proteins and polypeptides do not exhibit these and other enzymatic activities and thus are described as “nonenzymatic proteins.” These proteins include neurotoxins, cardiotoxins, myotoxins, ion channel inhibitors, and anticoagulant proteins (3, 4). Thus, snake venom proteins, whether they are enzymatic or nonenzymatic, have evolved as a complex mixture of proteins that target several tissues, organs, and physiological systems and interfere in their normal functions. Therefore snake venoms, when injected into a prey or victim, result in the simultaneous assault on various tissues, leading to multiple organ or system failure and often death.

A large number of protein toxins have been purified and characterized from snake venoms. These studies have shown that each venom contains over a hundred protein toxins. These toxins, however, belong to a very small number of superfamilies of proteins. For example, a single snake venom can contain as many as 15 isoforms of phospholipase A₂ (5–7). As one would expect, they share remarkable similarities in their primary, secondary, and tertiary structures. However, at times they differ from each other in their biological targeting and hence their pharmacological effects. Similarly, other enzymes as well as nonenzymatic proteins in snake venoms also exist in many isoforms (8) and can be classified in protein families. So far more than 1000 nonenzymatic proteins have been characterized, and these protein toxins are grouped into well recognized families as follows: 1) three-finger toxins (including neurotoxins and cardiotoxins), 2) serine proteinase inhibitors (including proteinase inhibitors and dengue toxins), 3) lectins, 4) sarafotoxins, 5) nerve growth factors, 6) atrial natriuretic peptides, 7) bradykinin-potentiating peptides, 8) disintegrins, and 9) helvepsins/CRISP (8–11). The members in each family of protein toxins have a similar molecular scaffold, but they exhibit multiple functions. Thus, it appears that during the evolution of venoms some of the molecular scaffolds have been “selected,” and various “functional sites” were generated by accelerated evolution to a common molecular scaffold. We are interested in the structure-function relationships of various families of toxins from snake and other venoms.

Many of the early efforts in venom research were directed toward the isolation and characterization of either proteins that are found in abundance or the most toxic components. With the advent of more sophisticated purification techniques, there have been studies of new and interesting protein components that are found in smaller quantities. In this paper, we describe a novel toxin that is a member of a new family of snake venom toxins. Thus, we have isolated and purified nawaprin, the first member of this family, from Naja nigricollis venom.

* This work was supported by research grants from the Defense Science and Technology Agency of Singapore, Academic Research Grants, and the Australian Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1UDK) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

To whom correspondence should be addressed: Dept. of Biological Sciences, Faculty of Science, National University of Singapore, 14 Science Dr. 4, Singapore 117543, Singapore. Fax: 65-6779-2486; E-mail: dbkinim@nus.edu.sg.
The complete amino acid sequence and the solution structure of this toxin have been determined. Nawaprin is structurally similar to secretory leukocyte protease inhibitor (SLPI) and elafin, the tertiary structures of which have been studied by NMR (12) and X-ray crystallography (13). Both nawaprin and elafin contain four disulfide bonds and several proline residues. Elafin is a specific inhibitor of human leukocyte elastastase and porcine pancreatic elastase, the former of which is similar to secretory leukocyte proteinase inhibitor (14, 15). This new protein fold has also been used as a scaffold in the evolution of snake venom toxins and may be useful in the engineering of proteins with novel pharmacological actions.

**EXPERIMENTAL PROCEDURES**

**Materials**—Lyophilized crude *N. nigricollis* venom was obtained from Miami Serpentarium Laboratories (Miami, FL). Trypsin endopeptidase was purchased from Wako Pure Chemicals (Osaka, Japan). 4-Vinylpyridine was obtained from Sigma. Superdex 30 and Sephasil C18 columns were obtained from Amersham Biosciences.

**Isolation and Purification of Nawaprin from *N. nigricollis* Snake Venom**—Nawaprin was purified by a three-step purification process protocol; gel filtration of venom on a Superdex 30 column was followed by ion exchange chromatography on a UNO S6 column and HPLC on a Jupiter C18 column. Crude venom (200 mg) was loaded onto a Superdex 30 column (HiLoad™ 16/60 equilibrated with 50 mM Tris-HCl buffer, pH 7.5. The proteins were eluted with the same buffer at a flow rate of 1 ml/min on a fast performance liquid chromatography system (Amersham Biosciences). The protein elution was monitored at 280 nm. The fraction with the peak of interest (2–5 mg) was applied separately onto a UNO S6 cation exchange column (Bio-Rad) pre-equilibrated with 50 mM Tris-HCl buffer, pH 7.5 (Buffer A). The protein boundaries were eluted by a linear gradient of 1 M NaCl in Buffer A. Protein elution was carried out at a flow rate of 2 ml/min and monitored at 280 nm. The unbound fraction from the UNO S6 column was loaded onto a Jupiter C18, 10 µm (10 mm × 250 mm) column equilibrated with 0.1% (v/v) trifluoroacetic acid on a Vision Work station (PerkinElmer Life Sciences). The bound proteins were eluted using a linear gradient of 80% acetonitrile (ACN) in 0.1% (v/v) trifluoroacetic acid at a flow rate of 2 ml/min. The elution of proteins was monitored at 215 nm.

**Reduction and Pyridylethylation**—Purified protein was reduced and pyridylethyalted using procedures described earlier (16). Protein (0.5 mg) was dissolved in 500 µl of denaturant buffer 6 M guanidium hydrochloride, 0.25 M Tris-HCl, 1 mM EDTA, pH 8.5. After the addition of 10 µl of β-mercaptoethanol, the mixture was incubated under vacuum for 2 h at 37 °C. 4-Vinylpyridine (50 µl) was added to the mixture and kept at room temperature for 2 h. Pyridylethyalted protein was purified on a μ-RFC C/2/C18 (21.1 mm × 10 mm) column using ACN in 0.1% (v/v) trifluoroacetic acid at a flow rate of 200 µl/min.

**Chemical and Enzymatic Cleavage**—Peptides of pyridylethyalted protein were obtained by chemical cleavage using formic acid (Asp-specific) as described by Inglis (17). Briefly, the desalted protein sample (500 µg) was dissolved in 2% formic acid in a glass vial and then frozen. Subsequently, under vacuum, the vial was thawed at room temperature and then sealed off. The vial was then heated at 108 °C for 2 h and allowed to cool to room temperature. Peptide digestion of the pyridylethyalted protein was also obtained by enzymatic cleavage with trypsin. Pyridylethyalted protein (300 µg) was dissolved in 300 µl of 100 mM ammonium bicarbonate buffer and digested overnight by trypsin at 37 °C. The peptides generated by both formic acid and trypsinic digestion were separated by reverse phase HPLC on a Sephasil C18 (15 µg, 21 mm × 10 mm) column, equilibrated with 0.1% (v/v) trifluoroacetic acid. A linear gradient of 80% (v/v) ACN in 0.1% trifluoroacetic acid (v/v) was used to elute crude peptides.

**Mass Spectrometry**—The protein fractions eluted from the columns were screened for novel molecular weight peptides using matrix-assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI-TOF/TOF). A Voyager DE-STR Biospectrometry Work station (Applied Biosystems, Foster City, CA). Typically, 1–5 pmol/µl of the sample was co-crystallized with an equal volume of the matrix (10 mg/ml of α-cyano-4-

---

1 The abbreviations used are: SLPI, secretory leukocyte protease inhibitor; ACN, acetonitrile; ESI-MS, electrospray ionization mass spectrometry; HPLC, high performance liquid chromatography; WAP, wey acid protein; NOESY, nuclear overhauser enhancement spectroscopy.
spectrometry (data not shown). Peak 4 had a mass of ~5290 Da. The molecular size was less than three-finger toxins and serine proteinase inhibitors but larger than atrial natriuretic peptides (8). Thus, based on its mass, we had identified a polypeptide belonging to no other known family of snake venom proteins. Proteins in peak 4 were further separated on a cation exchange column, UNO S6 (Fig. 1B). The protein of interest did not bind to the column; it was eluted in the unbound fraction. The protein was further purified on a reverse phase column, Jupiter C18 (Fig. 1C). The major peak in the HPLC chromatogram had a molecular weight of 5288.50 ± 0.08 by ESI-MS (Fig. 1D). The overall yield of the protein varied from batch to batch of venom samples between 0.09 and 0.51% (n = 8).

Determination of the Amino Acid Sequence—Amino-terminal sequencing of the native protein was achieved by the Edman degradation, and it resulted in the identification of first 34 residues (Fig. 2A). To complete the sequence, the pyridylethylated proteins were 5288.12 (with the assumption that all eight cysteine residues are involved in disulfide bond formation) and 6137.37, respectively; these values matched the estimated masses determined by ESI-MS (Table I). BLAST search for sequence homology indicated that this protein belongs to the family of whey acidic proteins (WAP) (Fig. 2B). Since then, we have isolated and purified two other peptides from snake venoms that show similar mass and amino acid sequence. Although all of the cysteine residues are conserved in these proteins, the intercysteine segments are distinctly different.2 Because of their homology with WAPs, we have named this new family of snake venom proteins Waprins (WAP-related proteins) and named the protein from *N. nigricollis* venom *Nawaprin* (*Naja waprin*).

WAPs were the first members of this family to be isolated. They are small secretory proteins widely distributed in the whey of many species (27–29). They contain two four-disulfide core domains. Waprins are structurally closer to the epididymal secretory protein members of the WAP family (Fig. 2B). These proteins are specifically expressed in the vas deferens.

---

2 B. G. Fry, S. T. Hock, and R. M. Kini, unpublished observations.
A New Family of Snake Venom Protein

A

B

C

Determination of Solution Structure—The NOESY spectrum of nawaprin obtained at 25 °C and pH 3.1 showed wide dispersion of amide proton signals indicating β-sheet secondary structures (Fig. 3). The analysis of the spectra was, however, not straightforward because it was made difficult by a number complicating factors: These included the presence of seven proline residues, excessive line broadening for a number of peaks, and an extra set of small peaks suggesting the presence of minor conformations of the peptide in solution.

The seven proline residues in nawaprin presented a major difficulty in resonance assignment because it led to peak overlap in appropriate regions of the spectra. All proline residues displayed strong $d/H_2O$ $(i-1,i)$ connectivities, suggesting that they were mainly in the trans-conformation (40), but still the existence of minor cis-conformations could not be completely discounted. For example, Pro 31 showed an additional $d/H_2O$ $(i-1,i)$ cross-peak, although its intensity is very weak. This could explain the presence of minor sets of peaks in the spectra that could not be readily assigned to a specific residue in the sequence. The presence of minor conformations(s) in solution was confirmed by reverse phase HPLC of the repurified sample wherein two minor peaks, whose intensities were 5% of the large (major) peak, were detected.

In addition to these unwanted factors, broad peaks were also observed for many backbone amide protons that may suggest intermediate chemical exchange of protons with the aqueous solvent, slow conformational averaging, or flexibility in the molecule. Moreover, the backbone amide peaks of Lys21, Leu23, Cys41, Met 44, and Thr 45 were split, suggesting a slow-to-medium state of conformational exchange. In the final analysis, the difficulties encountered in the resonance assignments were

and the distal epididymis (30). Trappins (transglutaminase substrate and WAP domain-containing proteins) of the WAP family are “trapped” in the tissues through covalent cross-linking (31–33). They contain an amino-terminal transglutaminase substrate domain (also called the cementoin domain (34)) with a variable number of hexapeptide repeats with the consensus sequence GQDPVK. Trappins anchor the biologically active WAP motifs at appropriate sites in the extracellular matrix through this domain. Elafin/SKALP (skin-derived antileukoproteinase), SPAI-2 (Na+/K+-ATPase inhibitor), and porcine WAP-3 are some of the members of the trappin family, although elafin and SPAI were first isolated as soluble proteins (14, 36). In contrast, some of the WAP proteins such as SLPI and human seminal plasma inhibitor do not have the cementoin domain and are produced as secreted proteins (37–39). WAPs and other family members have one to three similar domains, whereas waprins contain a single four-disulfide core domain and are found in snake venoms in soluble form.

| Table I |

| Theoretical and experimentally determined masses of nawaprin and its peptides |
|-----------------|
| Molecular mass | Calculated$^a$ | Observed$^b$ |
| Native nawaprin | 5288.21 | 5288.5 ± 0.08 |
| Pyridylethylated nawaprin | 6137.37 | 6737.5 ± 0.2 |
| Formic acid peptide F1 (28–51) | 3001.66 | 3002.5 ± 0.8 |
| Formic acid peptide F2 (10–51) | 5114.28 | 5113.5 ± 1.1 |
| Tryptic peptide T3 (39–51) | 1537.84 | 1535.5 ± 0.5 |

$^a$ Masses were calculated from the amino acid sequences.

$^b$ Masses were determined by ESI-MS or matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

and the distal epididymis (30). Trappins (transglutaminase substrate and WAP domain-containing proteins) of the WAP family are “trapped” in the tissues through covalent cross-linking (31–33). They contain an amino-terminal transglutaminase substrate domain (also called the cementoin domain (34)) with a variable number of hexapeptide repeats with the consensus sequence GQDPVK. Trappins anchor the biologically active WAP motifs at appropriate sites in the extracellular matrix through this domain. Elafin/SKALP (skin-derived antileukoproteinase), SPAI-2 (Na+/K+-ATPase inhibitor), and porcine WAP-3 are some of the members of the trappin family, although elafin and SPAI were first isolated as soluble proteins (14, 36). In contrast, some of the WAP proteins such as SLPI and human seminal plasma inhibitor do not have the cementoin domain and are produced as secreted proteins (37–39). WAPs and other family members have one to three similar domains, whereas waprins contain a single four-disulfide core domain and are found in snake venoms in soluble form.
resolved by performing several homonuclear two-dimensional experiments at different temperatures and pH.

The 34% sequence similarity between elafin and nawaprin, and the eight conserved cysteine residues, suggest equivalent cysteine pairing patterns (Fig. 2C). This was confirmed by NMR based on the characteristic NOEs between α and β protons of the bonded cysteine pairs. NOE connectivities that were observed included Cys30-H\(\alpha\)–Cys46-H\(\beta\), Cys30-H\(\beta\)–Cys46-H\(\alpha\), Cys80-H\(\alpha\)–Cys46-H\(\beta\), Cys80-H\(\beta\)–Cys46-H\(\alpha\), Cys7-H\(\alpha\)–Cys37-H\(\beta\), Cys7-H\(\beta\)–Cys37-H\(\alpha\), and Cys24-H\(\alpha\)–Cys36-H\(\beta\). The NOE cross-peaks linking Cys20–Cys41 were not observed, probably because of rather broad lines in the corresponding part of the spectrum. However, this pairing was easily established by elimination (because the three disulfide pairings were then known) and later during preliminary structure calculations, because considerable long range NOE connectivities were also observed among the protons of their neighboring residues (Cys41, Phe43, Thr22, and Lys21).

Structure Description—The structure of nawaprin in solution was characterized by the presence of both well and poorly defined regions, the extents of which were comparable in magnitude. Fig. 4A shows the ensemble of the best 20 structures superimposed over the backbone atoms N, C, and Ca of residues 2–8, 22–38, and 44–51 of the mean structure. B and C, ribbon diagram of the structure closest to the mean showing secondary structures and disulfide connectivities (in yellow). The two views are related by −90° rotation about the vertical axis. The figures were generated using MOLMOL (26).

![FIG. 4. Nawaprin structure. A, ensemble of the best 20 nawaprin structures superimposed to show the best fit over the backbone atoms N, C, and Ca of residues 2–8, 22–38, and 44–51 of the mean structure. B and C, ribbon diagram of the structure closest to the mean showing secondary structures and disulfide connectivities (in yellow). The two views are related by −90° rotation about the vertical axis. The figures were generated using MOLMOL (26).](image)

![FIG. 3. Fingerprint region of the NOESY spectrum of nawaprin at 25 °C and pH 3.1. Shown are the sequential NH-CαH connectivities for residues 26–30, 31–39, and 45–48. Intraresidue connectivities are labeled with their corresponding residue number.](image)

| TABLE II | Structural statistics for the ensemble of 20 nawaprin structures |
|----------|---------------------------------------------------------------|
| Quantity | Value             |
| Distance restraints |                     |
| Intraresidue \((i - j = 0)\) | 206 |
| Sequential \((i - j = 1)\) | 143 |
| Medium range \((i - j ≤ 5)\) | 24 |
| Long range \((i - j > 5)\) | 130 |
| Hydrogen bonds | 18 |
| Total | 521 |
| Dihedral angle restraints | 9 |
| Atomic root mean square deviation with the mean structure (Å) | 1.81 ± 0.58 |
| Backbone atoms (1–51) | 2.23 ± 0.48 |
| Heavy atoms (1–51) | 0.32 ± 0.07 |
| Backbone atoms (2–8, 22–38, and 44–51) | 0.68 ± 0.11 |
| Heavy atoms (2–8, 22–38, and 44–51) | |

residues were superimposed; this was reduced to 0.32 Å when only the well defined residues of 2–8, 22–38, and 44–51 were considered (Table II). Although the amino and carboxyl termini of nawaprin were relatively well defined, there was substantial disorder in the upper regions defined by residues 9–21, called the “outer loop,” and 39–43, called the “inner loop”; this suggests that these two regions of the molecule have higher flexibility than the rest. This apparent structural disorder reflected in the NMR spectra was mainly caused by the dearth of long...
range NOEs that would provide information on “connections” between the two loops. The few NOE connectivities that were observed were those between the protons of Ile$^{12}$-Cys$^{41}$ and Lys$^{21}$-Phe$^{43}$.

Unlike many disulfide cross-linked polypeptides, nawaprin is not compact but is rather flat and disc-like (Fig. 4, B and C). The backbone configuration is essentially spiral in shape, characterized by outer and inner circular segments that are connected by disulfide bonds. The inner segment incorporates a small twisted antiparallel $\beta$-sheet (a $\beta$-hairpin) at residues 35–37 and 45–47, whereas the outer segment is devoid of any defined secondary structures except for some $\beta$-turns that are situated at residues 26–29, 27–30, and 31–34. Note that in seven of the 20 “best” structures in the ensemble, a continuous $3_10$ helix spanning residues 26–30 was found instead of the usual $\beta$-turn(s). The three disulfide bridges are clustered together at the “base” of the molecule, anchoring the lower inner loop to the two ends of the outer loop; the fourth disulfide bridge defined by Cys$^{20}$–Cys$^{41}$ holds the tips of the two loops together.

**DISCUSSION**

We have described the purification and three-dimensional structure determination of the first snake-toxin member, nawaprin from *N. nigricollis*, of a new family of proteins. Nawaprin and other waprins from snake venoms$^{2}$ are small proteins with ~50 amino acid residues that have a four-disulfide core domain structure, making them members of the WAP family of proteins (Fig. 2).

Comparison with Elafin and Other Proteins—A DALI algorithm (41) search for similar structures in the Protein Data Bank revealed that the overall fold of nawaprin has significant similarity to that of elafin. This is expected given a high degree of similarity to that of elafin. This is expected given a similarity to that of elafin. This is expected given a similarity to that of elafin. This is expected given a similarity to that of elafin.

![Fig. 5. Comparison of nawaprin and elafin structures.](image)

A, nawaprin and elafin structures superimposed over the backbone atoms of residues 2–9, 11–41, and 43–50 of nawaprin and residues 11–18, 19–49, and 50–57 of elafin. Nawaprin and elafin (1fle) are shown in black and light gray, respectively. The first 10 residues in elafin are not shown. B, same as in A but drawn as ribbon diagram. C and D, molecular surface of nawaprin (C) and elafin (D) highlighted to show electrostatic potential. Surfaces with positive, negative, and neutral electrostatic potentials are drawn in blue, red, and white, respectively. The two views are related by 180° rotation about the “virtual” vertical axis. The brackets in D indicate the primary binding sites in elafin.

Continuous region has higher mobility as evidenced by the lack of long range NOE connectivities in this part of the nawaprin molecule. NMR relaxation experiments could be used to probe the mobility of this region in the molecule.

**Functional Implications**—The tip of the outer circular segment in elafin, defined by residues 20–26, is important for its activity, because it is the primary binding segment that interacts with the active-site pocket in porcine pancreatic elastase (13). This binding segment in elafin is composed of at least seven residues, LIRCAM (boxed parts of loops 1 and 2 in Fig. 2C), six of which are hydrophobic; the presence of several hydrophobic residues in this region is known to be crucial for the activities of elafin and other protease inhibitors such as SLPI (12). This region also incorporates a disulfide bond that connects the outer segment to the inner core of the inhibitor. In the porcine pancreatic elastase-elafin complex (13), the primary binding loop (outer loop) in elafin is actually in an extended $\beta$-strand conformation, forming an antiparallel $\beta$-sheet with porcine pancreatic elastase through a series of hydrogen bonds. In free solution, however, this outer loop segment is disordered (12). Nawaprin in solution also has an apparently disordered outer loop segment similar to that in elafin. However, based on the sequence alignment in Fig. 2C alone, nawaprin does not have a fragment analogous to the primary binding segment.
defined by residues 20–26. The DALI algorithm finds that this primary binding segment in elafin, which is composed of LIR-CAML, is topologically similar to the segment defined by residues 12–18 in nawaprin, which is composed of MPIPPLG. Although these two segments are both hydrophobic, they are not sequentially similar to each other. Furthermore, the relative positions of the cysteine pairs that connect the tips of the outer and inner loops are also different in the two molecules.

Fig. 5 (C and D) show the electrostatic potential surfaces of nawaprin and elafin. One can clearly see that the charge distributions in the two molecules are different. Although the upper halves of the two molecules, which incorporate the inner and the outer loops, contain a number hydrophobic residues, one side of the nawaprin has a more hydrophobic upper part than the corresponding region in elafin. In addition to this, there is large continuous negative patch in nawaprin defined by Glu¹², Asp¹⁰, Asp²⁷, and the carboxyl-terminal Pro⁵¹, which is absent in elafin. In fact, part of this region in elafin is positively charged because it includes two lysine residues, Lys¹² and Lys⁴⁸. The difference in the nature of the side chains of the two molecules therefore suggests that nawaprin may not be a protease inhibitor, although its overall fold is very similar to that of elafin.

The modest sequence similarity of 34% and the fact that both nawaprin and elafin incorporate several proline residues (including two consecutive proline residues in the external segments), share similar three-dimensional folds strongly suggest that two polypeptides have evolved from a common ancestral molecule.

Physiological Role of WAPs and Related Proteins—WAP is a major protein constituent in whey, and it is suggested to be the major food source for the young (44). Its secretion varies with the major protein constituent in whey, and it is suggested to be the major food source for the young (44). Its secretion varies with

Proteinase inhibitors of the WAP family play an important physiological role in regulating the activity of various proteinases. Generally, these inhibitors prevent the invasion of bacteria and other microbes. In addition, some of them play a specific role in host defense. For example, SLPI and elastase maintain the balance between proepithelin and epithelins and thus regulate innate immunity and wound healing (49). SLPI also acts as a potent anti-microbial agent that is a function that appears to be independent of its anti-proteinase activity (50). Mouse SWAM1 and SWAM2 have potent antibacterial activity, but they fail to inhibit elastase or cathepsin G (51). A caltrin-like protein secreted by guinea pig seminal vesicles inhibits Ca²⁺ uptake by spermatozoa (52). SPAI-1 inhibits Na⁺, K⁺-ATPase (36) but not proteinase (53). Therefore, based on their occurrence in snake venoms, wapprins may play a part in the "bait region" or "functional site" (9). In nature, this diversification could be achieved through gene duplication and accelerated evolution of WAP genes. Recent studies have shown that a locus on human chromosome 20 contains 14 genes encoding WAPs and related proteins, suggesting the evolution of WAP gene(s) by repeated duplications (56). Further, the region in exon 2 encoding the reactive site shows only 60–77% nucleotide identity compared with 97–98% identity in other regions (54). This suggests accelerated evolution of WAP genes.

In a similar fashion to the evolution of WAP proteins, several molecular scaffolds have been used during the evolution of “cocktails” of the toxins in snake venoms. The selected genes are duplicated several times, and the core of each protein scaffold is conserved, whereas the loops and surfaces are altered through mutations. As in the case of WAPs, some exons of the toxin genes of snake venom mutate more rapidly than their introns, thus speeding up the generation of new toxins (35). Therefore, members of a protein family share structurally important core residues including cysteine residues. However, the intercysteine loops show considerable differences in the sequence. This results in toxins with distinctly different molecular surfaces and hence different abilities to interact with target receptor/acceptor proteins. Hence they display differences in their biological properties (9).

So far, we have been able to isolate only a single waprin from each snake venom. However, three wapprins showed a conserved molecular framework with significantly different intercysteine loops. It will be interesting to search for other snake venom proteins with the WAP motif.

In summary, we have isolated and characterized a new structural family of snake venom proteins, wapprins. They contain a four-disulfide core structure and resemble the WAP structural fold. Furthermore, analysis of the structures indicates that wapprins could have a range of different biological properties.

REFERENCES

1. Bailey, G. S. (1998) Enzymes from Snake Venom, Alakan, Fort Collins, CO
2. Kini, R. M. (1997) Venom Phospholipase A2 Enzymes: Structure, Function and Mechanism, John Wiley & Sons, Chichester, UK
3. Harvey, A. L. (1991) Snake Toxins, Pergamon Press, New York
4. Tsunemi, M., Matsuura, Y., Sakakibara, S., and Katsube, Y. (1996) J. Mol. Biol. 261, 113–141
5. Wiedow, O., Schroder, J. M., Gregory, H., Young, J. A., and Christophers, E. (1999) Biochem. Biophys. Res. Commun. 265, 1479–1485
6. Joseph, J. S., Chung, M. C. M., Jeyaseelan, K., and Kini, R. M. (1999) Blood, 94, 621–631
7. Inglis, A. S. (1983) Methods Enzymol. 91, 324–332
8. Marion, D., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974
9. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 117, 479–485
A New Family of Snake Venom Protein